Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Е. Д. Стройникова

ОСНОВЫ ПРИКЛАДНОЙ АЛГЕБРЫ

Рекомендовано Учебно-методическим объединением высших учебных заведений Республики Беларусь по естественнонаучному образованию в качестве учебно-методического пособия для студентов, обучающихся по специальности 1-31 03 04 «Информатика»

Минск БГУИР 2010

УДК 512(076)
ББК 22.14я73
С86

Рецензенты:
кафедра высшей математики Белорусского государственного экономического университета;
доцент кафедры высшей математики Белорусского государственного университета, кандидат физико-математических наук, доцент Н. Я. Радыно

Стройникова, Е. Д.

С86 Основы прикладной алгебры : учеб.-метод. пособие / Е. Д. Стройникова. – Минск : БГУИР, 2010. – 120 с. : ил.

ISBN 978-985-488-548-3.

Учебно-методическое пособие знакомит со специальными главами современной высшей алгебры и их приложениями к информатике, в особенности вопросами защиты информации от несанкционированного доступа. Изложены основы теории чисел и наиболее популярных алгебраических систем: групп, колец и полей. В качестве вспомогательного материала с использованием аппарата дискретной математики рассмотрены основные виды соответствий и более подробно – функциональных отображений, понятие мощности множества. В качестве приложений представлены решение диофантовых линейных уравнений, основные типы классических шифров и криптосистема RSA.

Предназначено для студентов высших учебных заведений, обучающихся по специальности «Информатика», также может быть рекомендовано студентам математических специальностей при изучении родственных курсов.

УДК 512(076)
ББК 22.14я73

ISBN 978-985-488-548-3

Оглавление

Предисловие

Учебно-методическое пособие «Основы прикладной алгебры» составлено в соответствии с государственным образовательным стандартом, типовым учебным планом и типовой учебной программой по дисциплине «Геометрия и алгебра» для высших учебных заведений Республики Беларусь, обеспечивающих подготовку по специальности 1-31 03 04 «Информатика». В основу данной книги положены односеместровый курс лекций для студентов специальности «Информатика» факультета компьютерных систем и сетей БГУИР, читавшийся впервые профессором В. А. Липницким и в течение ряда последних лет – автором, и соответствующая теоретическая часть электронного учебно-методического комплекса по геометрии и алгебре, составленного автором для студентов дистанционной формы обучения. Целью пособия является знакомство читателя со специальными главами высшей алгебры и их важными приложениями к информатике, в особенности вопросами защиты информации от несанкционированного доступа. Издание может быть полезно студентам вузов математических специальностей при изучении родственных курсов и преподавателям.

Учебно-методическое пособие состоит из четырех глав. В начале приводится список основных обозначений и сокращений, встречающихся в тексте, в порядке их появления. В пособии изложены основы теории чисел и наиболее популярных алгебраических систем: групп, колец и полей. Для изучения материала третьей и четвертой глав во второй главе с использованием аппарата дискретной математики рассмотрены основные виды соответствий и более подробно – функциональных отображений, понятие мощности множества. В качестве приложений представлены решение диофантовых линейных уравнений, основные типы классических шифров и широко распространенная криптосистема с открытым ключом RSA. Изложение материала проводится с полными доказательствами утверждений и свойств. Исключение составляют сложные доказательства нескольких теорем, выходящие за рамки программы курса, которые в книге не приводятся. Некоторые из них можно найти в других учебных пособиях и монографиях, список которых помещен в конце данного издания. Разобрано достаточно много примеров, иллюстрирующих изложенный теоретический материал и дающих образцы решения задач, в том числе и прикладной направленности. Также включены упражнения для самостоятельного решения, причем к упражнениям вычислительного характера даны ответы. Подобранные примеры и упражнения способствуют усвоению излагаемого материала и делают книгу удобной для самостоятельного изучения курса, в частности, для студентов заочной и дистанционной форм обучения.

Автор выражает глубокую благодарность доктору технических наук, профессору В. А. Липницкому за ряд ценных советов, использованных при написании книги. Автор искренне благодарит кандидата физико-математических наук, доцента Н. Я. Радыно и коллектив кафедры высшей математики БГЭУ (заведующий кафедрой доктор физико-математических наук, профессор М. П. Дымков) за обстоятельное рецензирование, способствовавшее улучшению содержания.

Основные обозначения

N – множество натуральных чисел

Z – множество целых чисел

Q – множество рациональных чисел

R – множество вещественных чисел

C – множество комплексных чисел

$A \subset B$ – A является собственным подмножеством множества B

$A \subseteq B$ – A является подмножеством множества B

$a \in A$ – элемент a принадлежит множеству A

$A_{>a}$ – подмножество числового множества $A \subseteq \mathbf{R}$, состоящее из всех чисел, больших a

$A_{<a}$ – подмножество числового множества $A \subseteq \mathbf{R}$, состоящее из всех чисел, меньших a

$A_{\geq a}$ – подмножество числового множества $A \subseteq \mathbf{R}$, состоящее из всех чисел, больших или равных a

$A_{\leq a}$ – подмножество числового множества $A \subseteq \mathbf{R}$, состоящее из всех чисел, меньших или равных a

$a_1, a_2, \ldots, a_n \in A$ – элементы $a_1, a_2, \ldots, a_n$ принадлежат множеству A, $n \in \mathbf{N}$

$|a|$ – модуль числа a

$\triangleright$ – начало доказательства утверждения (леммы, свойства, теоремы)

$\triangleleft$ – окончание доказательства утверждения (леммы, свойства, теоремы)

$\Rightarrow$ – символ следствия (импликации)

$\bullet$ – окончание примера

$a \vdots b$ – целое число a делится на целое число $b \neq 0$

$b \mid a$ – целое число $b \neq 0$ делит целое число a

$a \not\vdots b$ – целое число a не делится на целое число $b \neq 0$

$b \nmid a$ – целое число $b \neq 0$ не делит целое число a

$\forall$ – квантор общности, обозначающий слова «любой», «каждый»

& – символ конъюнкции, одновременного выполнения условий, обозначающий логическое «и»

$\Leftrightarrow$ – символ эквивалентности (равносильности)

$\vee$ – символ дизъюнкции, выполнения хотя бы одного из условий, обозначающий логическое «или»

$\operatorname{sgn}(a)$ – сигнум вещественного числа a

ОД – общий делитель

НОД – наибольший общий делитель

$(a_1, a_2, \ldots, a_n)$ – НОД целых чисел $a_1, a_2, \ldots, a_n$, $n \geq 2$

$i = \overline{m,n}$ – i принимает все целые значения от m до n

ОК – общее кратное

НОК – наименьшее общее кратное

$[a_1, a_2, \ldots, a_n]$ – НОК целых чисел $a_1, a_2, \ldots, a_n$, $n \geq 2$

$A \cup B$ – объединение множеств A и B

$\exists$ – квантор существования, обозначающий слова «существует», «найдется»

$[a; b]$ – отрезок (подмножество **R**, состоящее из всех чисел r, таких, что $a \leq r \leq b$)

$[a]$ – целая часть вещественного числа a

$k!$ – факториал целого неотрицательного числа k

$\pi(x)$ – количество всех простых чисел, меньших x для $x \in \mathbf{N}_{>2}$

$|$ – обозначение слов «такой, что», «такие, что» при описательном задании множеств

$\bar{\exists}$ – отрицание существования

$a \equiv b \pmod{m}$ – целые числа a и b сравнимы по модулю $m \in \mathbf{N}$

$\bar{i}$ – класс вычетов по натуральному модулю с представителем $i \in \mathbf{Z}$

$A = B$ – множества A и B равны

$\mathbf{Z}/m\mathbf{Z}$ – множество классов вычетов по натуральному модулю m

$A \neq B$ – множества A и B не равны

φ – функция Эйлера (тотиент-функция)

$q\colon X \to Y$ – существование соответствия q между множествами X и Y

$A \times B$ – прямое (декартово) произведение множеств A и B

$q = (X, Y, Q)$ – соответствие q с областью отправления X, областью прибытия Y и графиком Q

$D(q)$ – область определения соответствия q

$E(q)$ – область значений соответствия q

$q(x)$ – образ элемента x при соответствии q

$q^{-1}(y)$ – прообраз элемента y при соответствии q

$q(C)$ – образ множества C при соответствии q

$q^{-1}(D)$ – прообраз множества D при соответствии q

$p = q$ – отображения p и q равны

$q\colon x \mapsto y$ или $q(x) = y$ – функциональное соответствие q сопоставляет элементу x единственный элемент y

e_X – тождественная функция на множестве X

$\mathbf{R}^2$ – декартова вещественная плоскость

q^{-1} – соответствие, обратное соответствию q

$\varnothing$ – пустое множество

$A \cap B$ – пересечение множеств A и B

$g \circ f$ или gf – композиция функций f и g

$p \neq q$ – отображения p и q не равны

f_A – сужение функции f на множество A

$A \leftrightarrow B$ – множества A и B равномощны

$|A|$ – мощность множества A

A^n – n-я декартова степень множества A, $n \in \mathbf{N}$

$P(A)$ – множество-степень (булеан) множества A

$a \notin A$ – элемент a не принадлежит множеству A

C_n^k – число сочетаний из n элементов по k, биномиальный коэффициент

$A \backslash B$ – разность множеств A и B

$(a; b)$ – интервал (подмножество $\mathbf{R}$, состоящее из всех чисел r, таких, что $a < r < b$)

$[a; b)$ – полуинтервал (подмножество $\mathbf{R}$, состоящее из всех чисел r, таких, что $a \le r < b$)

$(a; b]$ – полуинтервал (подмножество $\mathbf{R}$, состоящее из всех чисел r, таких, что $a < r \le b$)

A^* – множество всех обратимых элементов относительно некоторой бинарной алгебраической операции, заданной на множестве A

$M_{m \times n}(K)$ – множество прямоугольных матриц порядка $m \times n$ с коэффициентами из множества K

$V_n(K)$ – множество n-мерных векторов с компонентами из множества K

$GL_n(K)$ – полная линейная группа квадратных матриц порядка n с коэффициентами из множества K

C_n – мультипликативная группа всех комплексных корней n-й степени из 1

$H \le G$ или $H < G$, если $H \subset G$, – H является подгруппой группы G

$SL_n(K)$ – специальная линейная группа квадратных матриц порядка n с коэффициентами из множества K

$\det(A)$ – определитель матрицы A

$< a >$ – циклическая группа (подгруппа), порожденная элементом a

$\mathrm{ord}(a)$ – порядок элемента a в группе

E_n – единичная квадратная матрица порядка n

aH – левый смежный класс с представителем a группы по подгруппе H

Ha – правый смежный класс с представителем a группы по подгруппе H

$G : H$ – индекс подгруппы H в группе G

$H \lhd G$ – H является нормальной подгруппой группы G

S_n – симметрическая группа степени n

$S(\Omega)$ – симметрическая группа множества Ω

$(i\ f(i)\ f^2(i) \dots f^{k-1}(i))$ – цикл длиной k в группе S_n

A_n – знакопеременная группа степени n

G/H – фактормножество (факторгруппа) группы G по нормальной подгруппе H

$\mathrm{Im} f$ – образ группового или кольцевого гомоморфизма f

$\mathrm{Ker} f$ – ядро группового или кольцевого гомоморфизма f

$(G_1, *) \cong (G_2, \bullet)$ или $G_1 \cong G_2$ – группы $(G_1, *)$ и $(G_2, \bullet)$ изоморфны

$\mathrm{Aut}(G)$ – множество всех автоморфизмов группы G

ЭЦП – электронная цифровая подпись

$M_n(K)$ – множество квадратных матриц порядка n с коэффициентами из множества K

$K[x]$ – множество полиномов от одной переменной x с коэффициентами из кольца K

K^* – мультипликативная группа ассоциативного кольца K с единицей

$L \le K$ или $L < K$, если $L \subset K$, – L является подкольцом кольца K

$J \lhd K$ – J является двусторонним идеалом кольца K

(a) – главный идеал коммутативного кольца, порожденный элементом a

$\deg f$ – степень многочлена $f(x)$

$g(x) \mid f(x)$ – многочлен $g(x) \neq 0$ делит многочлен $f(x)$

$f(x) \vdots g(x)$ – многочлен $f(x)$ делится на многочлен $g(x) \neq 0$

$g(x) \nmid f(x)$ – многочлен $g(x) \neq 0$ не делит многочлен $f(x)$

$f(x) \not\vdots g(x)$ – многочлен $f(x)$ не делится на многочлен $g(x) \neq 0$

$(f_1(x), f_2(x), \dots, f_s(x))$ – НОД многочленов $f_1(x), f_2(x), \dots, f_s(x)$, $s \ge 2$

$[f_1(x), f_2(x), \dots, f_s(x)]$ – НОК многочленов $f_1(x), f_2(x), \dots, f_s(x)$, $s \ge 2$

$\boldsymbol{F_q}$ – конечное поле из q элементов

$\bar{a}$ – класс вычетов с представителем a кольца по модулю двустороннего идеала

K/I – фактормножество (факторкольцо) кольца K по двустороннему идеалу I

$(K_1, +, \cdot) \cong (K_2, \dot{+}, \bullet)$ или $K_1 \cong K_2$ – кольца $(K_1, +, \cdot)$ и $(K_2, \dot{+}, \bullet)$ изоморфны

$\mathrm{Aut}(K)$ – множество всех автоморфизмов кольца K

$\mathrm{char}\, K$ – характеристика кольца K

$P(x)$ – множество рациональных функций кольца $P[x]$, где P – поле

Глава 1. Основы теории чисел

§1.1. Алгебраические операции на множестве целых чисел. Делимость целых чисел

Наиболее популярными числовыми множествами являются следующие:

N – множество натуральных чисел,

Z – множество целых чисел,

Q – множество рациональных чисел,

R – множество вещественных чисел,

C – множество комплексных чисел.

Они связаны следующим отношением включения: $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$.

Отношения >, <, ≥ и ≤ полностью упорядочивают множество **R**. Если $A \subseteq \mathbf{R}$ и $a \in A$, то в дальнейшем будем обозначать $A_{>a}$, $A_{<a}$, $A_{\geq a}$ и $A_{\leq a}$ подмножества A, состоящие из всех чисел, больших a, меньших a, больших или равных a и меньших или равных a соответственно.

Множество целых чисел **Z** – счетное, состоит из элементов 0, ±1, ±2,… …, ±n,… На нем определены две алгебраические операции – сложение и умножение. Эти операции обладают следующими общими свойствами (для любых a, b, $c \in \mathbf{Z}$):

1) ассоциативностью: $a+(b+c)=(a+b)+c$, $a\cdot(b\cdot c)=(a\cdot b)\cdot c$;

2) коммутативностью: $a+b=b+a$, $a\cdot b=b\cdot a$;

3) существованием единственных нейтральных элементов – 0 относительно сложения и 1 относительно умножения соответственно: $a+0=a$, $a\cdot 1=a$.

Кроме того, операция сложения обладает свойством:

4) для каждого $a \in \mathbf{Z}$ существует единственное $b \in \mathbf{Z}$, такое, что: $a+b=0$.

Ясно, что здесь $b=-a$ – противоположное целое. Это свойство позволяет ввести на **Z** вспомогательную операцию – вычитание. Так, $a-b=a+(-b)=c$ – целое число – разность чисел a и b, получаемое вычитанием b из a.

Умножение и сложение связаны свойством:

5) $(a+b)\cdot c=a\cdot c+b\cdot c$ – законом дистрибутивности умножения относительно сложения.

Аналог свойства 4 для умножения выполняется лишь для двух целых чисел: 1 и –1.

Вообще говоря, для каждого целого $a \neq 0$ существует обратное относительно умножения (т. е. такое число b, что $a\cdot b=1$), но оно является рациональным и не является целым числом при $a \neq \pm 1$. Следовательно, результат операции деления целого числа a на целое число $b \neq 0$ есть число рациональное и в редких случаях является целым. В общем же случае имеет место следующая теорема.

Теорема 1.1.1 (о делении с остатком). Для любых целых чисел a и b, $b \neq 0$, существуют единственные целые числа q и r, $0 \le r < |b|$, такие, что

$$a = b \cdot q + r. \qquad (1.1.1)$$

▷ Доказательство существования.

1. $a = 0,\ \ a = 0 \cdot b + 0$. В этом случае $q = r = 0$.

2. $a > 0,\ \ b > 0$.

$a - \underbrace{b - \ldots - b}_{q} = r,\ \ 0 \le r < b$ (вычитаем b из a столько раз, чтобы получилось число 0 или число $r < b$), значит, $a = bq + r,\ \ q \ge 0$.

3. $a < 0,\ \ b < 0$.

$a - \underbrace{b - \ldots - b}_{q} = r,\ \ 0 \le r < |b|$, следовательно, $a = bq + r,\ \ q \ge 0$.

4. $a > 0,\ \ b < 0$.

$a + \underbrace{b + \ldots + b}_{-q} = a - (\underbrace{-b - \ldots - b}_{-q}) = r,\ 0 \le r < |b|$, значит, $a = bq + r,\ q \le 0$.

5. $a < 0,\ \ b > 0$.

$a + \underbrace{b + \ldots + b}_{-q} = a - (\underbrace{-b - \ldots - b}_{-q}) = r,\ 0 \le r < |b|$, значит, $a = bq + r,\ q \le 0$.

Доказательство единственности.

Пусть $a = bq_1 + r_1$ и $a = bq_2 + r_2$. Вычтем из 1-го равенства 2-е:

$$0 = bq_1 - bq_2 + r_1 - r_2,$$
$$b(q_2 - q_1) = r_1 - r_2.$$

Пусть $q_1 \neq q_2$, тогда $|r_1 - r_2| \ge |b|$, но $0 \le r_1, r_2 < |b| \Rightarrow |r_1 - r_2| < |b|$. Значит, $q_1 = q_2$, поэтому и $r_1 = r_2$. ◁

Определение 1.1.1. В равенстве (1.1.1) r называют *остатком*, а q – *частным* (*неполным частным* при $r \neq 0$) от деления a на b.

Для нахождения частного и остатка можно использовать метод деления «уголком».

Пример 1.1.1.

1. $a = 437,\ b = 38$.

$$\begin{array}{r|l} 437 & 38 \\ \cline{2-2} -\ 38 & 11 \\ \cline{1-1} 57 & \\ -\ 38 & \\ \cline{1-1} 19 & \end{array}$$

$q = 11,\ r = 19$.

2. $a = -203,\ b = 16$.

$203 = 16 \cdot 12 + 11;$

$$\begin{array}{r|l} -\,203 & 16 \\ \cline{2-2} -\,16 & 12 \\ \cline{1-1} 43 & \\ -\,32 & \\ \cline{1-1} 11 & \end{array}$$

$-203 = 16\cdot(-12) - 11 = 16\cdot(-12) - 16 + 5 = 16\cdot(-13) + 5;$
$q = -13,\ r = 5.$

3. $a = -178,\ b = -11.$

$$\begin{array}{r|l} -\,178 & 11 \\ \cline{2-2} -\,11 & 16 \\ \cline{1-1} 68 & \\ -\,66 & \\ \cline{1-1} 2 & \end{array}$$

$178 = 11\cdot 16 + 2;$
$-178 = (-11)\cdot 16 - 2 = (-11)\cdot 16 - 11 + 9 = (-11)\cdot 17 + 9;$
$q = 17,\ r = 9.$

4. $a = 725,\ b = -53.$

$$\begin{array}{r|l} -\,725 & 53 \\ \cline{2-2} -\,53 & 13 \\ \cline{1-1} 195 & \\ -\,159 & \\ \cline{1-1} 36 & \end{array}$$

$725 = (-53)\cdot(-13) + 36;$
$q = -13,\ r = 36.\bullet$

Определение 1.1.2. Если в равенстве (1.1.1) $r = 0$, т. е. $a = b\cdot q$, то говорят, что *a делится* на b (и пишут $a \vdots b$), что *a* является *кратным* числа b, что b *делит a* (и пишут $b \mid a$), а также называют b *делителем* или *множителем* числа *a*. Все вышесказанное для b справедливо и для q при $q \neq 0$. Будем обозначать $a \not\vdots b$, если *a* не делится на *b*, и $b \nmid a$, если b не делит *a*.

Свойства делимости целых чисел

1. $b \mid 0$ для $\forall b \neq 0$: действительно, $0 = b\cdot 0$.

2. $\pm 1 \mid a$ для $\forall a \in \mathbf{Z}$, поскольку $a = 1\cdot a$ и $a = (-1)\cdot(-a)$.

3. $b \mid a \ \&\ a \mid c \Rightarrow b \mid c$ (свойство транзитивности).

$\triangleright\ a = bq,\ c = at = bqt = bs$, где $q, t, s \in \mathbf{Z}.\triangleleft$

4. $b \mid a \ \&\ a \mid b \Leftrightarrow |a| = |b|$.

$\triangleright\ (a = bq,\ b = at = bqt,\ q, t \in \mathbf{Z}) \Leftrightarrow (q = t = 1 \vee q = t = -1) \Leftrightarrow (a = b \vee a = -b) \Leftrightarrow |a| = |b|.\triangleleft$

5. $(b \mid a_1, \ldots, b \mid a_s) \Rightarrow b \mid (a_1k_1 + \ldots + a_sk_s)$ для $\forall s \in \mathbf{N}$ и $\forall k_1, \ldots, k_s \in \mathbf{Z}$.

$\triangleright\ a_1k_1 + ... + a_sk_s = bq_1k_1 + ... + bq_sk_s = b(q_1k_1 + ... + q_sk_s),\ q_1,..., q_s \in \mathbf{Z}. \triangleleft$

6. Если в равенстве $a_1 + ... + a_n = b_1 + ... + b_m$ при $\forall\, m,\ n \in \mathbf{N}$ все слагаемые, принадлежащие **Z**, делятся на d (кроме, быть может, одного a_i, $1 \le i \le n$), то и это слагаемое также делится на d.

$\triangleright\ a_i = b_1 + ... + b_m - a_1 - ... - a_{i-1} - a_{i+1} - ... - a_n = d(s_1 + ... + s_m - q_1 - ... - q_{i-1} - q_{i+1} - ... - q_n),\ s_1,..., s_m,\ q_1,..., q_{i-1}, q_{i+1},..., q_n \in \mathbf{Z}$. По свойству 5 имеем $d \mid a_i. \triangleleft$

7. $b \mid a \Leftrightarrow |b| \mid |a|$.

$\triangleright\ (a = bq,\ q \in \mathbf{Z}) \Leftrightarrow |a| = \operatorname{sgn}(a) = |b| \operatorname{sgn}(b) q \Leftrightarrow ((|a| = |b| \operatorname{sgn}(b) q / \operatorname{sgn}(a) = |b| qt,\ t = \pm 1,\ a \ne 0) \vee (|a| = |b| \operatorname{sgn}(b) q,\ a = 0)) \Leftrightarrow |b| \mid |a|. \triangleleft$

§1.2. Наибольший общий делитель целых чисел. Алгоритм Евклида. Соотношение Безу. Наименьшее общее кратное

Определение 1.2.1. Если целые числа a_1, a_2,…, a_n, $n \ge 2$, делятся на целое $d \ne 0$, то d называют *общим делителем* (*ОД*) этих чисел.

В дальнейшем речь пойдет только о натуральных делителях.

Определение 1.2.2. Максимальный из общих делителей целых чисел a_1, a_2,…, a_n, из которых хотя бы одно отлично от нуля, называется их *наибольшим общим делителем* (*НОД*) и обозначается НОД(a_1, a_2,…, a_n) или (если это не вызывает разночтений) (a_1, a_2,…, a_n).

Свойства НОД

1. $d \in \mathbf{N}$.

2. $d \mid a_i,\ i = \overline{1,n}$.

3. $\left(\forall\, \delta \in \mathbf{N}\colon \delta \mid a_i,\ i = \overline{1,n}\right) \Rightarrow \delta \mid d$.

Определение 1.2.2 НОД целых чисел a_1, a_2,…, a_n равносильно выполнению свойств 1–3.

4. НОД целых чисел единственный.

$\triangleright$ Пусть d и d_1 – два НОД целых чисел a_1, a_2,…, a_n, $n \ge 2$. Тогда по свойству 3 НОД $d_1 \mid d$ & $d \mid d_1 \Leftrightarrow |d| = |d_1|$ по свойству 4 делимости целых чисел. Так как d, $d_1 \in \mathbf{N}$, то $d = d_1. \triangleleft$

Очевидно, что если $b \mid a$, то $(a, b) = |b|$.

Теорема 1.2.1. Если $a = b \cdot q + r$, то $(a, b) = (b, r)$.

$\triangleright$ По свойству 6 делимости целых чисел $(a, b) \mid r \Rightarrow (a, b)$ – ОД b и $r \Rightarrow (a, b) \mid (b, r)$. По свойствам 5, 6 делимости целых чисел $(b, r) \mid a \Rightarrow (b, r)$ – ОД a и $b \Rightarrow (b, r) \mid (a, b)$. Согласно свойству 4 делимости целых чисел, поскольку $(a, b) > 0,\ (b, r) > 0,\ (a, b) = (b, r). \triangleleft$

Это наблюдение (теорема 1.2.1) позволило выдающемуся древнегреческому математику Евклиду (III в. до н. э.) обосновать следующий факт (теорема 1.2.2), являющийся по сути кратным применением теоремы 1.2.1.

Теорема 1.2.2 (Евклид). Наибольший общий делитель целых чисел a и b при условии, что $|a|>|b|$, $b \nmid a$, равен последнему отличному от нуля остатку от деления в цепочке равенств:

$$a = b \cdot q_1 + r_1,$$
$$b = r_1 \cdot q_2 + r_2, \text{ если } r_1 \neq 0,$$
$$\dots$$
$$r_{n-2} = r_{n-1} \cdot q_n + r_n, \text{ если } r_{n-1} \neq 0,$$
$$r_{n-1} = r_n \cdot q_{n+1}, \text{ если } r_n \neq 0.$$

То есть $r_n = (a, b)$.

▷ Согласно теореме 1.2.1 и поскольку $r_n > 0$, как остаток от деления, имеем

$$(a, b) = (b, r_1) = (r_1, r_2) = \dots = (r_{n-1}, r_n) = r_n.$$

Процесс получения (a, b) конечен, поскольку мы оперируем только с целыми числами и, начиная с деления r_1 на r_2, – с целыми положительными числами. Идет постоянное уменьшение остатков r_i: $0 \le r_i < r_{i-1}$, и за конечное число шагов будет достигнут остаток $r_{n+1} = 0$. ◁

Пример 1.2.1. Найти (72, 26).

$$72 = 26 \cdot 2 + 20,$$
$$26 = 20 \cdot 1 + 6,$$
$$20 = 6 \cdot 3 + 2,$$
$$6 = 2 \cdot 3.$$

Следовательно, (72, 26) = 2.●

Теорема 1.2.2 дает алгоритм Евклида нахождения НОД целых чисел. С помощью рекурсии он легко преобразуется в алгоритм нахождения НОД не только двух, но и большего количества целых чисел: $(a_1, \dots, a_{n-1}, a_n) = ((a_1, \dots, a_{n-1}), a_n)$, $n = 3, 4, \dots$ Алгоритм Евклида остается в классе самых быстрых алгоритмов нахождения НОД целых чисел.

Обратное применение цепочки равенств теоремы 1.2.2, равно как и выражение на каждом шаге остатка от деления в виде целочисленной линейной комбинации a и b – так называемый *расширенный алгоритм Евклида* – доказывает следующий факт (теорема 1.2.3).

Теорема 1.2.3. Если $d = (a, b)$, то существуют такие целые u и v, что выполняется соотношение

$$d = u \cdot a + v \cdot b.$$

▷ Если $a \mid b$, то $d = |a|$, $|a| = \begin{cases} a = 1 \cdot a + 0 \cdot b, & a > 0; \\ -a = (-1) \cdot a + 0 \cdot b, & a < 0. \end{cases}$

Если $b \mid a$, то $d = |b|$, $|b| = \begin{cases} b = 0 \cdot a + 1 \cdot b, & b > 0; \\ -b = 0 \cdot a + (-1) \cdot b, & b < 0. \end{cases}$

Если $a \nmid b$, $b \nmid a$, то, не ограничивая общности, можем считать $|a| > |b|$ и согласно теореме 1.2.2 получаем

$$d = r_n = r_{n-2} - q_n r_{n-1} = (r_{n-4} - q_{n-2} r_{n-3}) - q_n (r_{n-3} - q_{n-1} r_{n-2}) = \dots = ua + vb,$$

где $u, v \in \mathbf{Z}$. Используя расширенный алгоритм Евклида, получим те же значения u и v:

$$r_1 = a - q_1 b,$$
$$r_2 = b - q_2 r_1 = b - q_2(a - q_1 b) = -q_2 a + (1 + q_2 q_1) b,$$
$$\dots$$
$$d = r_n = r_{n-2} - q_n r_{n-1} = \dots = ua + vb. \triangleleft$$

Следствие. Пусть $d = (a_1, a_2, \dots, a_n)$. Тогда существуют такие $u_1, u_2, \dots, u_n \in \mathbf{Z}$, что

$$d = \sum_{i=1}^{n} u_i a_i. \qquad (1.2.1)$$

$\triangleright$ Для $n = 2$ равенство (1.2.1) справедливо согласно теореме 1.2.3. При $n \geq 2$ сделаем индуктивное предположение о справедливости равенства (1.2.1). Поскольку $(a_1, \dots, a_n, a_{n+1}) = \big((a_1, \dots, a_n), a_{n+1}\big) = (d_n, a_{n+1}) = \left(\sum_{i=1}^{n} v_i a_i, a_{n+1}\right)$, то согласно теореме 1.2.3 получаем соотношение: $(d_n, a_{n+1}) = ud_n + u_{n+1} a_{n+1}$. Тогда $(a_1, \dots, a_n, a_{n+1}) = u\sum_{i=1}^{n} v_i a_i + u_{n+1} a_{n+1} = \sum_{i=1}^{n+1} u_i a_i$. Таким образом, равенство (1.2.1) справедливо для любого натурального $n \geq 2. \triangleleft$

Определение 1.2.3. Равенство (1.2.1) называют *соотношением Безу* (Этьен Безу (1730–1783) – французский математик) для наибольшего общего делителя целых чисел.

Пример 1.2.2. Из примера 1.2.1 следует, что

$$2 = 20 + 6 \cdot (-3) = 20 + \big(26 + 20 \cdot (-1)\big) \cdot (-3) = 20 \cdot 4 + 26 \cdot (-3) =$$
$$= \big(72 + 26 \cdot (-2)\big) \cdot 4 + 26 \cdot (-3) = 72 \cdot 4 + 26 \cdot (-11).$$
$$2 = 72 \cdot u + 26 \cdot v,\ u = 4,\ v = -11. \bullet$$

Замечание. Коэффициенты $u_1, \dots, u_n$ в (1.2.1) не являются единственными с таким условием. Это следует из теории диофантовых линейных уравнений, которая будет рассмотрена ниже. Так, в примере 1.2.2 числа $u = -9$ и $v = 25$ также удовлетворяют соотношению Безу.

Определение 1.2.4. Если целые, отличные от нуля числа $a_1, a_2, \dots, a_n$, $n \geq 2$, делят целое m, то m называют *общим кратным* (*ОК*) этих чисел.

В дальнейшем речь пойдет только о натуральных кратных.

Определение 1.2.5. Минимальное из общих кратных целых чисел $a_1, a_2, \dots, a_n$ называется их *наименьшим общим кратным* (*НОК*) и обозначается НОК$(a_1, a_2, \dots, a_n)$ или (если это не вызывает разночтений) $[a_1, a_2, \dots, a_n]$.

Свойства НОК

1. $m \in \mathbf{N}$.

2. $m \vdots a_i,\ i = \overline{1,n}$.

3. $\big(\forall \gamma \in \mathbf{N} : \gamma \vdots a_i,\ i = \overline{1,n}\big) \Rightarrow \gamma \vdots m$.

Определение 1.2.5 НОК целых чисел $a_1, a_2,\ldots, a_n$ равносильно выполнению свойств 1–3.

4. Если НОК целых чисел существует, то оно единственное.

▷ Пусть $m,\ m_1$ – два НОК целых чисел $a_1, a_2,\ldots, a_n,\ n \geq 2$. По свойству 3 НОК $m_1 \vdots m\ \&\ m \vdots m_1 \Leftrightarrow |m| = |m_1|$ по свойству 4 делимости целых чисел. Поскольку $m,\ m_1 \in \mathbf{N}$, то $m = m_1$. ◁

Очевидно, что если $b \mid a$, то $[a, b] = |a|$ при $a \neq 0$.

С помощью рекурсии НОК вычисляется не только для двух, но и большего количества целых чисел: $[a_1,\ldots, a_{n-1}, a_n] = \big[[a_1,\ldots, a_{n-1}], a_n\big]$, $n = 3, 4,\ldots$

§1.3. Простые числа

Определение 1.3.1. Натуральное число $p > 1$ называется *простым*, если оно делится только на 1 и на само себя, в противном случае p называется *составным*. 1 не является ни простым, ни составным числом. Таким образом,

$$\mathbf{N} = \{1\} \cup \{\text{простые числа}\} \cup \{\text{составные числа}\}.$$

Очевидно, справедлива следующая теорема.

Теорема 1.3.1. Наименьший делитель $p > 1$ натурального числа $n > 1$ есть число простое, причем если n – составное число, то $p \leq \sqrt{n}$.

▷ 1. n – простое число $\Rightarrow p = n$.

2. n – составное число, пусть p – составное $\Rightarrow (\exists\, t \in \mathbf{N}\colon\ 1 < t < p,\ t \mid p) \Rightarrow$ $\Rightarrow t \mid n$, что противоречит минимальности делителя p.

Пусть теперь $(n = pn_1,\ 1 < p \leq n_1 < n) \Rightarrow p^2 \leq n \Rightarrow p \leq \sqrt{n}$. ◁

Свойства простых чисел

1. Для $\forall\, n \in \mathbf{N}$ и любого простого p $(n,\ p) = \begin{cases} 1, & p \nmid n; \\ p, & p \mid n. \end{cases}$

2. $(p_1, p_2) = 1$, если $p_1 \neq p_2$ – различные простые числа.

Заметим, что из соотношения $n = p \cdot q$ натуральных чисел при $n > 1$, $1 < p, q < n$ следует, что либо p, либо q принадлежит отрезку $\big[2;\ \sqrt{n}\,\big]$. Исторически первый метод проверки числа n на простоту заключался в делении его на простые числа, не превосходящие $\sqrt{n}$. Если среди таких чисел делителей не найдено, то n – простое число. Это один из вариантов так называемого «*решета*» Эратосфена, ставшего его наиболее знаменитым достижением. Эратосфен (ок. 276–194 до н. э.) – один из самых разносторонних ученых Античности.

$$\underline{2},\ \underline{3},\ 4,\ \underline{5},\ 6,\ \underline{7},\ 8,\ 9,\ 10,\ \underline{11},\ 12,\ldots,\ \big[\sqrt{n}\,\big],$$

где $\big[\sqrt{n}\,\big]$ – целая часть $\sqrt{n}$:

1) 2 – простое число, исключаем все числа, кратные 2 ($2 \mid n$ – да или нет?);

2) 3 – простое число, исключаем все числа, кратные 3 ($3 \mid n$ – да или нет?);

и т. д. до ближайшего к $\sqrt{n}$ простого числа.

В целом «решето» Эратосфена решает более общую задачу нахождения всех простых чисел на отрезке натурального ряда $[m; n]$, где $m < n$.

Теорема 1.3.2 (Евклид). Простых чисел бесконечно много.

▷ Пусть $p_1, p_2, \ldots, p_k$, $k \in \mathbf{N}$, – все простые числа. Рассмотрим число $N = p_1 \cdot p_2 \cdot \ldots \cdot p_k + 1$. N не делится ни на одно из p_i, $i = \overline{1,k}$, т. к. в противном случае для некоторого i $p_i \mid 1$ (свойство делимости целых чисел 6), что невозможно. Поэтому N является либо новым простым числом, либо составным, в последнем случае существует отличное от всех p_i простое число p, где $p \mid N$. ◁

Значение простых чисел в том, что они, согласно теореме 1.3.1, являются составными «кирпичиками» всех натуральных чисел. Распределение простых чисел среди натуральных весьма непредсказуемо, о чем свидетельствуют следующие две теоремы. Теорема 1.3.3, доказанная в 1852 г. выдающимся русским математиком и механиком Пафнутием Львовичем Чебышевым (1821–1894), приводится здесь без доказательства.

Теорема 1.3.3 (П. Л. Чебышев). Между натуральными числами k и $2k$, $k > 1$, обязательно найдутся простые.

Теорема 1.3.4. Для всякого натурального n существует отрезок $[k; k+n]$, $k \in \mathbf{N}$, натурального ряда, все числа которого составные.

▷ В самом деле, все следующие числа составные:

$$k = (n+2)! + 2, \ldots, k + n = (n+2)! + 2 + n.$$

Действительно, $k = 2 \cdot (3 \cdot \ldots \cdot (n+2) + 1), \ldots, k + n = (n+2) \cdot (2 \cdot \ldots \cdot (n+1) + 1)$ (здесь $k! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot k$ для $k \in \mathbf{N}$, $0! = 1$). ◁

Несмотря на теорему 1.3.4, для $x \in \mathbf{N}_{>2}$ количество $\pi(x)$ всех простых чисел, меньших x, подчинено достаточно равномерному закону.

Теорема 1.3.5 (Ж. Адамар, Ш. Ж. Валле-Пуссен).

$$\pi(x) \approx x / \ln x.$$

Любопытным фактом является утверждение следующей теоремы, доказанной в 1737 г. выдающимся швейцарским математиком Леонардом Эйлером (1707–1783).

Теорема 1.3.6 (Л. Эйлер). Ряд $1/2 + 1/3 + 1/5 + \ldots$ из чисел, обратных простым, – расходящийся.

Это означает, что сумма ряда равна $+\infty$. Тем не менее, его частичная сумма по всем известным простым числам (их примерно 50 млн) меньше 4.

Теоремы 1.3.5 и 1.3.6 приведены здесь без доказательства.

Также широко известно применение простых чисел в *криптографии* – науке о методах шифрования данных. Для криптографических целей нужны большие простые числа длиной более 80–90 десятичных знаков. Такие числа замечательно подходят для этих целей, т. к., во-первых, они встречаются довольно часто: аппроксимация была найдена французским математиком Жаком Адамаром (1865–1963) и бельгийским математиком Шарлем Жаном де ла Валле-Пуссеном (1866–1962), которые независимо друг от друга доказали в 1896 г., что простых чисел, меньших натурального числа x, $\pi(x) \approx x / \ln x$ (теорема 1.3.5).

Во-вторых, различные вычисления с простыми числами (модульная арифметика) дают хорошие *односторонние функции* (эффективно вычислимые, для задачи инвертирования которых не существует эффективных алгоритмов), позволяющие зашифровать преобразованное в числовой вид сообщение. Например, произведение двух простых чисел – качественная односторонняя функция.

Большой интерес представляют простые числа вида $2^p - 1$, которые названы в честь открывшего их французского математика, физика, философа и теолога Марена Мерсенна (1588–1648). Мерсенн открыл числа вида $2^p - 1$ в результате поиска *совершенных чисел* (так называются натуральные числа, которые равны сумме всех своих делителей, меньших данного числа, например, $496 = 1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248$). Одним из преимуществ чисел вида $2^p - 1$, благодаря которому при проверке их на простоту происходит отсев показателей p, является то, что числа такого вида могут быть простыми только тогда, когда p – простое число. Но не для любого простого p число $2^p - 1$ является простым, например, $2^{11} - 1 = 2047 = 23 \cdot 89$. При больших показателях p получаются огромные числа $2^p - 1$, поэтому все рекордные простые числа принадлежат классу *чисел Мерсенна*.

Французский математик Франсуа Эдуард Анатоль Люка (1842–1891) разработал методы для проверки чисел на простоту. В 1876 г. он доказал, что $2^{127} - 1$ – простое число. Оно оставалось самым большим известным простым числом Мерсенна в течение 75 лет. Позже, в 1930 г., американский математик, профессор университета в Беркли Деррик Генри Лемер (род. в 1905 г.) продолжил работу Люка и получил тест проверки чисел Мерсенна на простоту, называемый теперь *тестом Люка – Лемера*.

Поиском простых чисел Мерсенна занимается проект GIMPS (Great Internet Mersenne Primes Search), действующий с 1996 г. Именно участники этого проекта последнее время находят новые большие простые числа Мерсенна. Инструментом поиска служит программа Prime95 основателя GIMPS Джорджа Вольтмана (род. в 1957 г.), последняя версия которой – v25.7, где и используется тест Люка – Лемера. Исходный текст этой программы открыт для изучения и представляет собой высокооптимизированный ассемблерный код.

Открытое в 1999 г. 38-е простое число Мерсенна являлось самым большим простым числом на конец XX в. – это число $2^{6972593} - 1$, которое содержит 2098960 десятичных знаков. После этого было открыто еще девять больших простых чисел Мерсенна. По состоянию на конец 2009 г. самым большим из них является открытое в августе 2008 г. 45-е число $2^{43112609} - 1$, содержащее 12978189 десятичных знаков, последним из найденных является открытое в апреле 2009 г. 47-е число $2^{42643801} - 1$, содержащее 12837064 десятичных знака.

Обобщением теоремы 1.3.2 является следующая теорема, доказанная в 1837 г. известным немецким математиком, внесшим существенный вклад в математический анализ, теорию функций и теорию чисел, Петером Густавом Леженом-Дирихле (1805–1859), приводимая здесь без доказательства.

Теорема 1.3.7 (П. Дирихле). Всякая арифметическая прогрессия $\{a + b \cdot n \mid n \in \mathbf{Z}_{\geq 0}\}$, где a, b – фиксированные натуральные числа и $(a, b) = 1$, содержит бесконечно много простых чисел.

К сожалению, больше в теории чисел результатов, аналогичных теореме 1.3.7, нет. До сих пор существует много открытых вопросов относительно простых чисел, наиболее известные из которых были перечислены немецким математиком Эдмундом Георгом Германом Ландау (1877–1938) на V Международном математическом конгрессе в 1912 г.

Первая проблема Ландау, известная еще как *проблема Гольдбаха* (Кристиан Гольдбах (1690–1764), немецкий математик): доказать или опровергнуть, что каждое четное число, большее 2, может быть представлено в виде суммы двух простых чисел, а каждое нечетное число, большее 5, может быть представлено в виде суммы трех простых чисел. *Вторая проблема Ландау*: бесконечно ли множество *простых чисел-близнецов* (чисел типа p и $p+2$, например, 3 и 5, 5 и 7, 11 и 13, 17 и 19, 29 и 31,…, 1997 и 1999 и т. д.)? *Третья проблема Ландау*, известная также как *гипотеза Лежандра* (Адриен Мари Лежандр (1752–1833), французский математик): верно ли, что между n^2 и $(n+1)^2$ всегда найдется простое число? *Четвертая проблема Ландау*: бесконечно ли множество простых чисел вида $n^2 + 1$?

Открытыми проблемами являются бесконечность множества простых чисел Мерсенна, а также *чисел Ферма* $F_n = 2^{2^n} + 1$ и *чисел-значений многочлена Эйлера* $g(n) = n^2 + n + 41$, $n = 0, 1, 2,...$ Еще известный французский математик Пьер де Ферма (1601–1665) показал, что F_0, F_1, F_2, F_3, F_4 – простые числа, и выдвинул гипотезу, что все числа такого вида простые. Однако эта гипотеза была опровергнута в 1732 г. Леонардом Эйлером, нашедшим разложение числа F_5 на простые множители: $F_5 = 4294967297 = 641 \cdot 6700417$, и больше простых чисел вида F_n найдено не было. При $n = 0, 1, 2, 3,\ldots, 39$ многочлен $g(n)$ дает число простое, однако больше простых чисел такого вида не найдено, например, $g(40) = 41^2$.

§1.4. Взаимно простые числа. Основная теорема арифметики

Определение 1.4.1. Целые числа $a_1,..., a_n$, $n \geq 2$, называются *взаимно простыми*, если $(a_1,..., a_n) = 1$.

Такие числа не имеют общих простых делителей. Развитием теоремы 1.2.3 и следствия из нее является следующая теорема.

Теорема 1.4.1 (критерий взаимной простоты). Целые числа $a_1,..., a_n$, $n \geq 2$, взаимно просты тогда и только тогда, когда существуют такие целые числа $u_1,..., u_n$, что

$$\sum_{i=1}^{n} u_i a_i = 1.$$

▷ Необходимость. $(a_1,..., a_n)=1$, из соотношения Безу для НОД следует, что $\exists\ u_1,..., u_n \in \mathbf{Z}$, такие, что $\sum_{i=1}^{n} u_i a_i = 1$.

Достаточность. Пусть $\exists\ u_1,..., u_n \in \mathbf{Z}$, такие, что $\sum_{i=1}^{n} u_i a_i = 1$. Если $(a_1,..., a_n)=d$, то $d \mid 1$ (в силу свойства 5 делимости целых чисел), следовательно, $d=1$. ◁

Свойства взаимно простых чисел

1. $(a_1,..., a_n)=d \Leftrightarrow (a_1/d,..., a_n/d)=1$.

▷ Необходимость. Согласно следствию из теоремы 1.2.3 $\exists\ u_1,..., u_n \in \mathbf{Z}$, такие, что $d=u_1a_1+...+u_na_n$, следовательно, $1=u_1(a_1/d)+...+u_n(a_n/d)$, значит, $(a_1/d,..., a_n/d)=1$ по теореме 1.4.1.

Достаточность. $(a_1/d,..., a_n/d)=1$, следовательно, $\exists\ u_1,..., u_n \in \mathbf{Z}$, такие, что $1=u_1(a_1/d)+...+u_n(a_n/d)$ (согласно теореме 1.4.1), откуда $d=u_1a_1+\ldots+u_na_n$. Значит, $(a_1,..., a_n) \mid d$ по свойству 5 делимости целых чисел, а поскольку $d \mid a_i$, $i=\overline{1,n}$, то $d \mid (a_1,..., a_n)$ по определению НОД, следовательно, $(a_1,..., a_n)=d$. ◁

2. $c \mid ab$ & $(c, a) = 1 \Rightarrow c \mid b$.

▷ При $b \neq 0$ $\exists\ t, u, v \in \mathbf{Z}$, такие, что $ab=ct$ & $cu+av=1 \Rightarrow cub+ctv=b \Rightarrow$ $\Rightarrow c \mid b$ по свойству 5 делимости целых чисел. При $b=0$ доказательство очевидно согласно свойству 1 делимости целых чисел. ◁

3. $(a, b_1)=1$ & $(a, b_2)=1 \Leftrightarrow (a, b_1b_2)=1$.

▷ Необходимость.

$$\begin{cases} au_1+b_1v_1=1 \\ au_2+b_2v_2=1 \end{cases} \Rightarrow a^2u_1u_2+a(u_1b_2v_2+u_2b_1v_1)+b_1b_2v_1v_2=1 \Rightarrow au+b_1b_2v=1,$$

где u_1, v_1, u_2, v_2, u, $v \in \mathbf{Z}$. Значит, $(a, b_1b_2)=1$ согласно теореме 1.4.1.

Достаточность. Пусть $(a, b_1b_2)=1$. Если предположить, что $(a, b_1)=d_1>1$ или $(a, b_2)=d_2>1$, то согласно свойству 3 делимости целых чисел и определению НОД получим, что $d_1 \mid (a, b_1b_2)$ либо $d_2 \mid (a, b_1b_2)$. Это противоречит тому, что $(a, b_1b_2)=1$. Значит, $(a, b_1)=1$ & $(a, b_2)=1$. ◁

4. Следствие из свойства 3. Простое число p делит произведение $n_1 \cdot \ldots \cdot n_k$, $k \in \mathbf{N}$, тогда и только тогда, когда $\exists\ n_i$, $1 \le i \le k$, со свойством $p \mid n_i$.

▷ Необходимость. Если $\overline{\exists}\ n_i$ с указанным свойством, то по свойству 1 простых чисел $(p, n_i)=1$, $i=\overline{1,k}$, значит, $(p, n_1 \cdot \ldots \cdot n_k)=1$ по свойству 3 взаимно простых чисел, следовательно, $p \nmid n_1 \cdot \ldots \cdot n_k$ снова по свойству 1 простых чисел, что приводит к противоречию. Итак, $\exists\ n_i$, $1 \le i \le k$, со свойством $p \mid n_i$.

Достаточность. Если $\exists\, n_i$, $1 \le i \le k$, со свойством $p | n_i$, то $p \mid n_1 \cdot \ldots \cdot n_k$ по свойству 3 делимости целых чисел. ◁

С помощью свойств взаимно простых чисел и метода математической индукции доказывается следующая важная теорема.

Теорема 1.4.2 (основная теорема арифметики). Всякое натуральное число $n > 1$ однозначно раскладывается в произведение простых чисел с точностью до порядка следования множителей:

$$n = p_1 \cdot p_2 \cdot \ldots \cdot p_s,\ s \in \mathbf{N}. \qquad (1.4.1)$$

▷ $n = 2$ – база индукции. Предположим, что утверждение верно для $\forall k < n$.

Для простого числа n доказательство очевидно, $s = 1$. Пусть $n = n_1 \cdot n_2$ – составное число, $1 < n_1, n_2 < n$. Воспользуемся предположением индукции и разложим на простые множители n_1 и n_2, тем самым в произведение простых чисел будет разложено и n.

Пусть $n = p_1 \cdot p_2 \cdot \ldots \cdot p_s$ и $n = q_1 \cdot q_2 \cdot \ldots \cdot q_t$, $s, t \in \mathbf{N}$, – два разложения числа n на простые множители. Тогда, воспользовавшись свойством 4 взаимно простых чисел и свойством 1 простых чисел и последовательно сокращая обе части равенства $p_1 \cdot p_2 \cdot \ldots \cdot p_s = q_1 \cdot q_2 \cdot \ldots \cdot q_t$ на p_1, p_2, ..., p_s, приходим к выводу, что $s = t$ и с точностью до порядка следования множителей $p_i = q_i$, $i = \overline{1,s}$. Таким образом, разложение (1.4.1) числа n однозначно с точностью до порядка следования множителей. ◁

Определение 1.4.2. Если в равенстве (1.4.1) собрать одинаковые множители в степени, то получим *каноническое разложение натурального числа* $n > 1$: $n = p_1^{r_1} \cdot p_2^{r_2} \cdot \ldots \cdot p_t^{r_t}$, где $r_i \in \mathbf{N}$, $i = \overline{1,t}$, $p_i \ne p_j$ при $i \ne j$. Если z – целое, не равное 0, ±1 число, то $z = \operatorname{sgn}(z) \cdot p_1^{r_1} \cdot p_2^{r_2} \cdot \ldots \cdot p_t^{r_t}$ – *каноническое разложение* z.

Пример 1.4.1. Найдем разложения вида (1.4.1) и канонические разложения целых чисел.

1. $-484 = -2 \cdot 242 = -2 \cdot 2 \cdot 121 = -2 \cdot 2 \cdot 11 \cdot 11 = -2^2 \cdot 11^2$.

2. $2^{12} - 1 = 4095 = 3 \cdot 1365 = 3 \cdot 3 \cdot 455 = 3 \cdot 3 \cdot 5 \cdot 91 = 3 \cdot 3 \cdot 5 \cdot 7 \cdot 13 =$
$= 3^2 \cdot 5 \cdot 7 \cdot 13$. ●

По каноническому разложению целых чисел легко находятся их НОД и НОК, решаются иные задачи.

Пусть $a_1, \ldots, a_n$, $n \ge 2$, – ненулевые целые числа, хотя бы одно из которых отлично от ±1. Запишем их канонические разложения:

$$a_i = \operatorname{sgn}(a_i) \cdot p_1^{\alpha_{i1}} \cdot \ldots \cdot p_t^{\alpha_{it}},\ t \in \mathbf{N},\ \alpha_{ij} \in \mathbf{Z}_{\ge 0},\ j = \overline{1,t},\ i = \overline{1,n},$$

где $p_1, \ldots, p_t$ – простые числа, входящие во все разложения a_i, $i = \overline{1,n}$, причем в случае отсутствия множителя p_j в разложении a_i полагается $\alpha_{ij} = 0$. Тогда, исходя из определений, получаем следующие формулы для канонических разложений НОД и НОК чисел $a_1, \ldots, a_n$:

$$(a_1, \ldots, a_n) = p_1^{\gamma_1} \cdot \ldots \cdot p_t^{\gamma_t}, \text{ где } \gamma_j = \min_{1 \le i \le n} \alpha_{ij},\ j = \overline{1,t},$$

$$[a_1,..., a_n] = p_1^{\delta_1} \cdot ... \cdot p_t^{\delta_t}, \text{ где } \delta_j = \max_{1 \le i \le n} \alpha_{ij},\ j = \overline{1,t}.$$

§1.5. Диофантовы линейные уравнения

Диофант (ок. 325–409) – знаменитый александрийский математик. Ему бесспорно принадлежит заслуга введения в математику неопределенного анализа, он также ввел в алгебру буквенную символику. Именем Диофанта названы два больших раздела теории чисел – теория диофантовых уравнений и теория диофантовых приближений. Даты рождения и смерти этого ученого не вполне достоверны. По некоторым сведениям, Диофант жил в период ок. 250 г. Известно, что прожил он 84 года, как это видно из эпитафии, составленной в виде следующей задачи: «Диофант провел 6-ю часть жизни в младенчестве и 12-ю часть в юношеском возрасте; затем он женился и прожил в бездетном супружестве 7-ю часть жизни и еще 5 лет, после чего у него родился сын, достигший только половины возраста отца; отец же пережил сына на 4 года». Если обозначить за x количество лет жизни Диофанта, то можно составить следующее уравнение:

$$\frac{1}{6}x + \frac{1}{12}x + \frac{1}{7}x + 5 + \frac{1}{2}x + 4 = x.$$

Решая это уравнение, получим, что $x = 84$.

Определение 1.5.1. *Диофантовым линейным уравнением* называется уравнение вида

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b, \qquad (1.5.1)$$

где $n \in \mathbf{N}$, коэффициенты и правая часть $a_1, a_2,..., a_n, b \in \mathbf{Z}$, $\exists\ a_i \neq 0,\ 1 \le i \le n$, а решения $(x_1,\ x_2,...,\ x_n)$ – множества упорядоченных наборов из n элементов – также ищутся в целых числах.

Теорема 1.5.1. Уравнение (1.5.1) разрешимо в целых числах тогда и только тогда, когда $\text{НОД}(a_1,..., a_n) \mid b$.

▷ Необходимость. Пусть $(k_1,..., k_n)$ – целочисленное решение уравнения (1.5.1). Тогда справедливо равенство $\sum_{i=1}^{n} a_i k_i = b$, следовательно, по свойству 5 делимости целых чисел $\text{НОД}(a_1,..., a_n) \mid b$.

Достаточность. Пусть $d = \text{НОД}(a_1,..., a_n),\ d \mid b \Rightarrow (b = dt,\ t \in \mathbf{Z})$. Тогда по следствию из теоремы 1.2.3 (соотношение Безу для НОД) $\exists\ u_1,..., u_n \in \mathbf{Z}$, такие, что $\sum_{i=1}^{n} a_i u_i = d$, следовательно, умножая последнее равенство на t при $t \neq 0$, получаем $\sum_{i=1}^{n} a_i t u_i = b$, т. к. $tu_i \in \mathbf{Z},\ i = \overline{1,n}$, то $(tu_1,..., tu_n)$ – решение уравнения (1.5.1). При $t = 0$ получаем, что (0,…, 0) – решение уравнения (1.5.1). ◁

Рассмотрим частный случай уравнения (1.5.1) – уравнение с двумя неизвестными:

$$ax + by = c, \tag{1.5.2}$$

где $a, b, c \in \mathbf{Z}$, $a \neq 0$ либо $b \neq 0$, а пары-решения (x, y) ищутся в целых числах.

Теорема 1.5.2. Если НОД$(a, b) \nmid c$, то уравнение (1.5.2) не имеет решений в целых числах. При НОД$(a, b) \mid c$, $a \neq 0$, $b \neq 0$ множество всех целочисленных пар-решений уравнения (1.5.2) имеет вид

$$\left\{\left(x_0 + \frac{b}{d}t,\, y_0 - \frac{a}{d}t\right) \mid t \in \mathbf{Z}\right\}, \tag{1.5.3}$$

где $d = \text{НОД}(a, b)$;

$(x_0,\, y_0)$ – некоторое частное решение уравнения (1.5.2).

При $a \mid c$, $a \neq 0$, $b = 0$ множество всех целочисленных пар-решений уравнения (1.5.2) имеет вид

$$\left\{\left(\frac{c}{a}, t\right) \mid t \in \mathbf{Z}\right\}. \tag{1.5.3′}$$

При $b \mid c$, $a = 0$, $b \neq 0$ множество всех целочисленных пар-решений уравнения (1.5.2) имеет вид

$$\left\{\left(t, \frac{c}{b}\right) \mid t \in \mathbf{Z}\right\}. \tag{1.5.3″}$$

▷ Если НОД$(a, b) \nmid c$, то уравнение (1.5.2) как частный случай уравнения (1.5.1) не имеет решений в целых числах по теореме 1.5.1.

Пусть теперь НОД$(a, b) = d$ и $d \mid c$. Тогда уравнение (1.5.2) имеет решения в целых числах согласно теореме 1.5.1.

Пусть $a \neq 0$, $b \neq 0$. Так как $d = \text{НОД}(a, b, c)$, то уравнение (1.5.2) эквивалентно уравнению

$$\frac{a}{d}x + \frac{b}{d}y = \frac{c}{d}. \tag{1.5.4}$$

Пусть $(x_0,\, y_0)$ – некоторое частное целочисленное решение уравнения (1.5.2), тогда справедливо соотношение

$$\frac{a}{d}x_0 + \frac{b}{d}y_0 = \frac{c}{d}. \tag{1.5.5}$$

Если (x, y) – целочисленное решение уравнения (1.5.2), то для данной пары выполняется соотношение (1.5.4). Вычтем почленно из равенства (1.5.4) равенство (1.5.5), получим равенство

$$\frac{a}{d}(x - x_0) + \frac{b}{d}(y - y_0) = 0,$$

которое эквивалентно соотношению

$$\frac{a}{d}(x - x_0) = \frac{b}{d}(y_0 - y).$$

Так как по свойству 1 взаимно простых чисел $\text{НОД}\left(\frac{a}{d}, \frac{b}{d}\right) = 1$, то по свойству 2 взаимно простых чисел $\frac{a}{d} | (y_0 - y)$. Следовательно, существует $t \in \mathbf{Z}$, такое, что $\frac{a}{d}t = y_0 - y$, откуда $y = y_0 - \frac{a}{d}t$. Тогда имеем равенство

$$\frac{a}{d}(x - x_0) = \frac{b}{d}\frac{a}{d}t \Leftrightarrow x - x_0 = \frac{b}{d}t,$$

откуда $x = x_0 + \frac{b}{d}t$. Итак, получаем, что пара $(x, y) = \left(x_0 + \frac{b}{d}t,\ y_0 - \frac{a}{d}t\right)$, являющаяся целочисленным решением уравнения (1.5.2), имеет вид (1.5.3).

Рассмотрим теперь целочисленные пары $\left(x_0 + \frac{b}{d}t,\ y_0 - \frac{a}{d}t\right)$ при произвольных $t \in \mathbf{Z}$, где (x_0, y_0) – некоторое частное целочисленное решение уравнения (1.5.2), а значит, и уравнения (1.5.4). Подставим $x = x_0 + \frac{b}{d}t,\ y = y_0 - \frac{a}{d}t$ в левую часть уравнения (1.5.2). Получим

$$a\left(x_0 + \frac{b}{d}t\right) + b\left(y_0 - \frac{a}{d}t\right) = ax_0 + by_0 = c,$$

т. к. (x_0, y_0) удовлетворяет соотношению (1.5.2). Итак, множество (1.5.3) является множеством всех целочисленных решений уравнения (1.5.2).

При $a \neq 0$, $b = 0$ и $a = 0$, $b \neq 0$ получаем частный случай уравнения (1.5.2) – линейное уравнение с одним неизвестным в целых числах.

Пусть $a | c$, $a \neq 0$, $b = 0$. Тогда уравнение (1.5.2) имеет вид $ax = c \Leftrightarrow x = \frac{c}{a}$, при этом y может быть произвольным целым числом. И множество всех целочисленных решений (1.5.2) имеет в этом случае вид (1.5.3′).

Пусть $b | c$, $a = 0$, $b \neq 0$. Тогда уравнение (1.5.2) имеет вид $by = c \Leftrightarrow y = \frac{c}{b}$, при этом x может быть произвольным целым числом. И множество всех целочисленных решений (1.5.2) имеет в этом случае вид (1.5.3″). ◁

Алгоритм решения диофантовых линейных уравнений с двумя неизвестными

1. Если $\text{НОД}(a, b) \nmid c$, то уравнение (1.5.2) не имеет решений в целых числах.

2. При $a | c$, $a \neq 0$, $b = 0$ множество всех целочисленных решений уравнения (1.5.2) имеет вид (1.5.3′).

3. При $b | c$, $a = 0$, $b \neq 0$ множество всех целочисленных решений уравнения (1.5.2) имеет вид (1.5.3″).

4. При $\text{НОД}(a, b) = d$, $d | c$, $a \neq 0$, $b \neq 0$ приходим к уравнению (1.5.4) и переходим к шагу 5.

5. Поскольку $\text{НОД}\left(\frac{a}{d}, \frac{b}{d}\right) = 1$, то по теореме 1.4.1 (критерий взаимной простоты) существуют и находятся, например, по расширенному алгоритму Евклида или с помощью обратного применения цепочки равенств алгоритма Евклида $u_0, v_0 \in \mathbf{Z}$, такие, что выполняется соотношение Безу:

$$\frac{a}{d}u_0 + \frac{b}{d}v_0 = 1. \qquad (1.5.6)$$

Далее переходим к шагу 6.

6. Умножим обе части равенства (1.5.6) на $\frac{c}{d}$, получим

$$\frac{a}{d}\frac{c}{d}u_0 + \frac{b}{d}\frac{c}{d}v_0 = \frac{c}{d},$$

откуда $x_0 = \frac{c}{d}u_0$, $y_0 = \frac{c}{d}v_0$ – частное решение уравнения (1.5.4), а значит, и уравнения (1.5.2). Переходим к шагу 7.

7. По формуле (1.5.3) находим все множество целочисленных решений уравнения (1.5.2).

Многие старинные способы «угадывания» числа и месяца рождения основывались на умении решать диофантовы линейные уравнения. Например, чтобы определить число и месяц рождения вашего собеседника, вам достаточно узнать у него сумму, получаемую от сложения двух произведений: даты дня рождения (x) на 12 и номера месяца (y) на 31. Причем из всех решений выбирается то единственное, для которого $1 \le x \le 31$, $1 \le y \le 12$. Здесь НОД(12, 31) = 1.

Упражнение 1.5.1. Определите число и месяц рождения человека, у которого указанная выше сумма равна 67.

Ответ. 3 января.

С помощью диофантовых линейных уравнений решаются и другие задачи.

Упражнение 1.5.2. Сколько существует способов составления отрезка длиной 1 м из отрезков длинами 7 см и 12 см?

Ответ. Единственный способ: 4 отрезка по 7 см и 6 отрезков по 12 см.

§1.6. Сравнения целых чисел

Теорема 1.6.1. Пусть m – натуральное число. Для любых целых чисел a и b следующие условия равносильны:

1) a и b имеют одинаковые остатки от деления на m;

2) $a - b$ делится на m, т. е. $a - b = m \cdot q$, для подходящего целого q;

3) $a = b + m \cdot q$ для некоторого целого q.

▷ **1) ⇒ 2).** Пусть $a = m \cdot q_1 + r$, $b = m \cdot q_2 + r$, $0 \le r < m$, $q_1, q_2 \in \mathbf{Z}$. Тогда $a - b = m \cdot q_1 - m \cdot q_2 = m \cdot (q_1 - q_2) \Rightarrow (a - b) \vdots m$.

2) ⇒ 3). $(a - b) \vdots m \Rightarrow (a - b = m \cdot q, q \in \mathbf{Z}) \Rightarrow a = b + m \cdot q$.

3) $\Rightarrow$ 1). Пусть $a = b + m \cdot q$, $a = m \cdot q_1 + r_1$, $b = m \cdot q_2 + r_2$, $0 \le r_1, r_2 < m$, $q, q_1, q_2 \in \mathbf{Z}$. Тогда $m \cdot q_1 + r_1 = m \cdot q_2 + r_2 + m \cdot q \Rightarrow r_1 - r_2 = m \cdot (q_2 + q - q_1) \Rightarrow$ $\Rightarrow m | (r_1 - r_2) \Rightarrow r_1 = r_2$. $\triangleleft$

Определение 1.6.1. Целые числа a и b называются *сравнимыми по модулю m*, если они удовлетворяют условиям теоремы 1.6.1. Этот факт обозначают формулой $a \equiv b \pmod m$ или $a \equiv b\ (m)$. Данное соотношение между целыми числами называют *сравнением по модулю m*.

Пример 1.6.1. $-6 \equiv 9 \equiv 14 \equiv 24 \equiv 39 \equiv 4 \pmod 5$.$\bullet$

Свойства сравнений

1. Для всякого целого c $a \equiv b \pmod m \Leftrightarrow a \pm c \equiv b \pm c \pmod m$, т. е. к обеим частям сравнения можно прибавить или вычесть из обеих частей одно и то же целое число.

$\triangleright (a = b + mq, q \in \mathbf{Z}) \Leftrightarrow a \pm c = b \pm c + mq \Leftrightarrow a \pm c \equiv b \pm c \pmod m$. $\triangleleft$

2. Сравнения можно почленно складывать и вычитать: $a \equiv b \pmod m$ & $c \equiv$ $\equiv d \pmod m \Rightarrow a \pm c \equiv b \pm d \pmod m$.

$\triangleright (a = b + mq_1$ & $c = d + mq_2,\ q_1, q_2 \in \mathbf{Z}) \Rightarrow (a \pm c = b \pm d + m(q_1 \pm q_2),$ $q_1 \pm q_2 \in \mathbf{Z}) \Leftrightarrow a \pm c \equiv b \pm d \pmod m$. $\triangleleft$

3. Сравнения можно почленно перемножать: $a \equiv b \pmod m$ & $c \equiv$ $\equiv d \pmod m \Rightarrow a \cdot c \equiv b \cdot d \pmod m$.

$\triangleright (a = b + mq_1$ & $c = d + mq_2,\ q_1, q_2 \in \mathbf{Z}) \Rightarrow (a \cdot c = (b + mq_1) \cdot (d + mq_2) =$ $= b \cdot d + m(q_1 d + bq_2 + mq_1 q_2) = b \cdot d + m \cdot t,\ t \in \mathbf{Z}) \Leftrightarrow a \cdot c \equiv b \cdot d \pmod m$. $\triangleleft$

4. Следствие из свойства 3 сравнений. Сравнения можно почленно возводить в любую натуральную степень: $a \equiv b \pmod m \Rightarrow a^n \equiv b^n \pmod m$ для $\forall\, n \in \mathbf{N}$.

5. Если в сравнении числа a, b, m имеют общий делитель d, то на него обе части сравнения и модуль можно сократить: $a \equiv b \pmod m \Leftrightarrow a/d \equiv b/d \pmod{m/d}$.

$\triangleright (a = b + mq, q \in \mathbf{Z}) \Leftrightarrow a/d = b/d + (m/d)q \Leftrightarrow a/d \equiv b/d \pmod{m/d}$. $\triangleleft$

6. Обе части сравнения можно сократить на их общий множитель, взаимно простой с модулем: если $a = a_1 \cdot d$, $b = b_1 \cdot d$, $(d, m) = 1$, то $a \equiv b \pmod m \Leftrightarrow$ $\Leftrightarrow a_1 \equiv b_1 \pmod m$.

$\triangleright (a_1 \cdot d = b_1 \cdot d + mq,\ q \in \mathbf{Z}) \Rightarrow d | mq$ по свойству 6 делимости целых чисел, но т. к. $(d, m) = 1$, то $d | q$ по свойству 2 взаимно простых чисел, значит, $q = q_1 \cdot d$, $q_1 \in \mathbf{Z}$. Таким образом, $a \equiv b \pmod m \Leftrightarrow a_1 \cdot d = b_1 \cdot d + m \cdot q_1 \cdot d \Leftrightarrow$ $\Leftrightarrow a_1 = b_1 + m \cdot q_1 \Leftrightarrow a_1 \equiv b_1 \pmod m$. $\triangleleft$

Легко проверяются следующие три свойства.

7. Рефлексивность: для любого целого a и всякого натурального m $a \equiv a \pmod m$.

$\triangleright$ Действительно, $a - a = 0$ и $m | 0$ по свойству 1 делимости целых чисел. $\triangleleft$

8. Симметричность: $a \equiv b \pmod m \Leftrightarrow b \equiv a \pmod m$.

$\triangleright m | (a - b) \Leftrightarrow m | (b - a)$ по свойству 7 делимости целых чисел. $\triangleleft$

9. Транзитивность: $a \equiv b \pmod m$ & $b \equiv c \pmod m \Rightarrow a \equiv c \pmod m$.

$\triangleright\,(a-c=(a-b)+(b-c),\ m\,|\,(a-b)\ \&\ m\,|\,(b-c))\Rightarrow m\,|\,(a-c)$ по свойствам 5 и 6 делимости целых чисел. $\triangleleft$

Свойства 7–9 означают, что отношение сравнимости на множестве целых чисел **Z** есть отношение эквивалентности. Это означает, что **Z** разбивается на непересекающиеся классы попарно сравнимых друг с другом по модулю $m\in\mathbf{N}$ целых чисел.

Каждый класс сравнимых друг с другом целых чисел характеризуется общими свойствами представителей этого класса. Например, все они имеют один и тот же остаток от деления на модуль; все они в силу теорем 1.6.1 и 1.2.1 имеют одинаковый наибольший общий делитель с этим модулем:

$$a=mq_1+r,\ b=mq_2+r,\ \text{где}\ q_1,q_2\in\mathbf{Z},\ 0\le r<m;$$

$$(a=mq+b,\ \text{где}\ q\in\mathbf{Z})\Rightarrow(a,\,m)=(m,\,b).$$

§1.7. Множество классов вычетов. Функция Эйлера

При делении целых чисел на натуральное число m существует m различных остатков: 0, 1, 2,…, $m-1$. Соответственно этим остаткам **Z** разбивается на m непересекающихся классов сравнимых друг с другом чисел, имеющих, как отмечено в §1.6, один и тот же остаток от деления на m. В соответствии с остатками от деления на m будем обозначать эти классы $\bar{0},\bar{1},\ldots,\overline{m-1}$. Таким образом, класс $\bar{i}=\{m\cdot q+i\mid q\in\mathbf{Z}\}$ для каждого $i=0,1,2,\ldots,m-1$. Любой представитель класса однозначно определяет свой класс, т. е. для каждого $m\cdot q+i$ класс $\overline{m\cdot q+i}=\bar{i}$.

Поскольку английское residue – «остаток» – переводится на русский язык еще и как «вычет», то множество всех классов сравнимых друг с другом чисел по модулю m называют *множеством классов вычетов по модулю m* и обозначают $\mathbf{Z}/m\mathbf{Z}$. Таким образом, $\mathbf{Z}/m\mathbf{Z}=\{\bar{0},\bar{1},\ldots,\overline{m-1}\}$ – множество из m элементов. Заметим, что для любых классов $\bar{k},\bar{l}\in\mathbf{Z}/m\mathbf{Z}$ и для произвольных $k_1,k_2\in\bar{k}$, $l_1,l_2\in\bar{l}$ суммы k_1+l_1 и k_2+l_2 принадлежат одному классу из $\mathbf{Z}/m\mathbf{Z}$, т. к. эти суммы сравнимы друг с другом по модулю m согласно свойству 2 сравнений из §1.6. Аналогично, согласно свойству 3 сравнений из §1.6 произведения $k_1\cdot l_1$ и $k_2\cdot l_2$ находятся в одном классе из $\mathbf{Z}/m\mathbf{Z}$.

$$k_1\equiv k_2\ (\mathrm{mod}\ m)\ \&\ l_1\equiv l_2\ (\mathrm{mod}\ m)\Rightarrow k_1+l_1\equiv k_2+l_2\ (\mathrm{mod}\ m)\ \&\ k_1\cdot l_1\equiv k_2\cdot l_2\ (\mathrm{mod}\ m).$$

Определим операции сложения и умножения на $\mathbf{Z}/m\mathbf{Z}$. Полагаем, что сумма $\bar{k}+\bar{l}=\bar{z}$, где $\bar{z}$ – такой единственный класс из $\mathbf{Z}/m\mathbf{Z}$, в который попадают все суммы $k+l$ для $k\in\bar{k}$, $l\in\bar{l}$, а произведение $\bar{k}\cdot\bar{l}=\bar{s}$ – тот единственный класс из $\mathbf{Z}/m\mathbf{Z}$, в который попадают все произведения $k\cdot l$ для $k\in\bar{k}$, $l\in\bar{l}$:

$$\bar{k}+\bar{l}=\{k+l\mid k\in\bar{k},\,l\in\bar{l}\},\ \bar{k}\cdot\bar{l}=\{k\cdot l\mid k\in\bar{k},\,l\in\bar{l}\}.$$

Поскольку сложение и умножение в $\mathbf{Z}/m\mathbf{Z}$ однозначно определяются сложением и умножением представителей классов вычетов, то свойства 1–5 операций сложения и умножения в $\mathbf{Z}$ из §1.1 справедливы и в $\mathbf{Z}/m\mathbf{Z}$:

1) $\bar{k} + (\bar{l} + \bar{r}) = (\bar{k} + \bar{l}) + \bar{r}$, $\bar{k} \cdot (\bar{l} \cdot \bar{r}) = (\bar{k} \cdot \bar{l}) \cdot \bar{r}$ – ассоциативность;

2) $\bar{k} + \bar{l} = \bar{l} + \bar{k}$, $\bar{k} \cdot \bar{l} = \bar{l} \cdot \bar{k}$ – коммутативность;

3) существуют единственные нейтральные элементы: $\bar{k} + \bar{0} = \bar{k}$, $\bar{k} \cdot \bar{1} = \bar{k}$ для $\forall\ \bar{k} \in \mathbf{Z}/m$.

Пусть $\bar{n}$ и $\bar{t}$ – нейтральные элементы относительно сложения и умножения соответственно. Тогда

$$\bar{n} + \bar{0} = \bar{n}\ \&\ \bar{n} + \bar{0} = \bar{0} \Rightarrow \bar{n} = \bar{0};\ \bar{t} \cdot \bar{1} = \bar{t}\ \&\ \bar{t} \cdot \bar{1} = \bar{1} \Rightarrow \bar{t} = \bar{1}.$$

При $m = 1$ класс вычетов $\bar{0}$ – нейтральный элемент относительно сложения и умножения, поскольку $\bar{0} = \bar{1}$;

4) для всякого $\bar{k} \in \mathbf{Z}/m\mathbf{Z}$ существует единственный противоположный класс $\bar{l} = -\bar{k}$, такой, что $\bar{k} + \bar{l} = \bar{0}$, при этом $\bar{l} = \overline{m - k}$.

Пусть $\bar{k} + \bar{l}_1 = \bar{0}$, $\bar{k} + \bar{l}_2 = \bar{0}$. Тогда

$$\bar{l}_1 + \bar{k} + \bar{l}_2 = \bar{l}_1 + (\bar{k} + \bar{l}_2) = \bar{l}_1 = (\bar{l}_1 + \bar{k}) + \bar{l}_2 = \bar{l}_2 \Rightarrow \bar{l}_1 = \bar{l}_2;$$

5) $(\bar{k} + \bar{l}) \cdot \bar{r} = \bar{k} \cdot \bar{r} + \bar{l} \cdot \bar{r}$ – дистрибутивность умножения относительно сложения.

Свойства 1, 2, 5 выполняются для всех $\bar{k}, \bar{l}, \bar{r} \in \mathbf{Z}/m\mathbf{Z}$.

Определение 1.7.1. Элемент $\bar{k} \in \mathbf{Z}/m\mathbf{Z}$ называется *обратимым*, если найдется такой класс $\bar{l} \in \mathbf{Z}/m\mathbf{Z}$, что $\bar{k} \cdot \bar{l} = \bar{1}$. Тогда $\bar{l} = \bar{k}^{-1}$ называют *классом, обратным* $\bar{k}$.

Так как $\bar{1} = \bar{0}$ при $m = 1$, то $\mathbf{Z}/\mathbf{Z} = \{\bar{0}\}$ и состоит из одного обратимого класса вычетов.

Из ассоциативности умножения вытекает, что если $\bar{k}$ – обратимый класс, то для него обратный класс определен однозначно.

Пусть $\bar{k} \cdot \bar{l}_1 = \bar{1}$, $\bar{k} \cdot \bar{l}_2 = \bar{1}$. Тогда

$$\bar{l}_1 \cdot \bar{k} \cdot \bar{l}_2 = \bar{l}_1 \cdot (\bar{k} \cdot \bar{l}_2) = \bar{l}_1 = (\bar{l}_1 \cdot \bar{k}) \cdot \bar{l}_2 = \bar{l}_2 \Rightarrow \bar{l}_1 = \bar{l}_2.$$

Лемма 1.7.1. Пусть $\bar{k} \in \mathbf{Z}/m\mathbf{Z}$ – такой класс, что $(k, m) = 1$. Тогда:

1) для каждого $\bar{l} \neq \bar{0}$ $\bar{k} \cdot \bar{l} \neq \bar{0}$;

2) $\bar{k} \cdot \bar{l}_1 \neq \bar{k} \cdot \bar{l}_2$, если $\bar{l}_1 \neq \bar{l}_2$;

3) $\bar{k}$ – обратимый класс в $\mathbf{Z}/m\mathbf{Z}$.

▷ **1)** $(k, m) = 1$. Если бы $\exists\, l$, $0 < l < m$, такое, что $\bar{k} \cdot \bar{l} = \bar{0}$, то $m \mid k \cdot l$, но по свойству 2 взаимной простоты $m \mid l$, что невозможно;

2) если бы $\bar{k} \cdot \bar{l}_1 = \bar{k} \cdot \bar{l}_2$ при $\bar{l}_1 \neq \bar{l}_2$, то $\bar{k} \cdot (\bar{l}_1 - \bar{l}_2) = \bar{0}$, т. е. $m \mid k \cdot (l_1 - l_2)$, в силу свойства 2 взаимной простоты $m \mid (l_1 - l_2)$, что невозможно, поскольку $0 < |l_1 - l_2| < m$;

3) по критерию взаимной простоты (теорема 1.4.1) $\exists\, u, v \in \mathbf{Z}$, такие, что $k \cdot u + m \cdot v = 1 \Leftrightarrow k \cdot u = 1 - m \cdot v \Leftrightarrow \overline{k} \cdot \overline{u} = \overline{1}$ согласно условию 3 теоремы 1.6.1, т. е. $\overline{k}$ – обратимый класс, $\overline{u}$ – обратный ему класс. ◁

Лемма 1.7.2. Пусть $\overline{k} \in \mathbf{Z}/m\mathbf{Z}$ – такой класс, что $(k, m) = d > 1$. Тогда:

1) существует $\overline{l} \neq \overline{0}$, такой, что $\overline{k} \cdot \overline{l} = \overline{0}$;

2) существуют $\overline{l_1} \neq \overline{l_2}$, такие, что $\overline{k} \cdot \overline{l_1} = \overline{k} \cdot \overline{l_2}$;

3) для всех $\overline{l} \neq \overline{0}$ $\overline{k} \cdot \overline{l} \neq \overline{1}$, т. е. класс $\overline{k}$ необратим в $\mathbf{Z}/m\mathbf{Z}$.

▷ **1)** так как $(k, m) = d > 1$, то $m = d \cdot l$, где $1 \le l < m$, $k = d \cdot k_1$, $(k_1, l) = 1$ по свойству 1 взаимно простых чисел. Тогда $k \cdot l = (d \cdot k_1) \cdot l = k_1 \cdot (d \cdot l) = k_1 \cdot m$, что означает $\overline{k} \cdot \overline{l} = \overline{0}$ при $\overline{l} \neq \overline{0}$;

2) существует $\overline{l} \neq \overline{0}$ с условием $\overline{k} \cdot \overline{l} = \overline{0}$ согласно утверждению 1 данной леммы. Рассмотрим $\overline{l_1} \in \mathbf{Z}/m\mathbf{Z}$, построим класс $\overline{l_2} = \overline{l_1 + l} = \overline{l_1} + \overline{l} \in \mathbf{Z}/m\mathbf{Z}$, $\overline{l_1} \neq \overline{l_2}$. Тогда $\overline{k} \cdot \overline{l_2} = \overline{k} \cdot \left(\overline{l_1} + \overline{l}\right) = \overline{k} \cdot \overline{l_1} + \overline{0} = \overline{k} \cdot \overline{l_1}$;

3) если бы $\overline{k}$ был обратим в $\mathbf{Z}/m\mathbf{Z}$, то существовал бы $\overline{u} \in \mathbf{Z}/m\mathbf{Z}$ со свойством $\overline{k} \cdot \overline{u} = \overline{1}$, т. е. согласно условию 3 теоремы 1.6.1 $(k \cdot u = 1 + m \cdot q,\ q \in \mathbf{Z}) \Leftrightarrow$ $\Leftrightarrow k \cdot u - m \cdot q = 1 \Leftrightarrow (k, m) = 1$ (по критерию взаимной простоты), что не так. ◁

Из лемм 1.7.1 и 1.7.2 вытекает следующая теорема.

Теорема 1.7.1. Класс $\overline{k} \in \mathbf{Z}/m\mathbf{Z}$ обратим тогда и только тогда, когда $(k, m) = 1$. Произведение обратимых классов есть обратимый класс.

▷ Первое утверждение следует из утверждений 3 лемм 1.7.1 и 1.7.2.

Пусть $n \ge 2$ и $\overline{k_1}, \ldots, \overline{k_n}$ – обратимые классы, $\overline{k_1}^{-1}, \ldots, \overline{k_n}^{-1}$ – соответственно обратные им классы. Тогда, поскольку операция умножения в $\mathbf{Z}/m\mathbf{Z}$ коммутативна,

$$\left(\overline{k_1} \cdot \ldots \cdot \overline{k_n}\right) \cdot \left(\overline{k_1}^{-1} \cdot \ldots \cdot \overline{k_n}^{-1}\right) = \left(\overline{k_1} \cdot \overline{k_1}^{-1}\right) \cdot \ldots \cdot \left(\overline{k_n} \cdot \overline{k_n}^{-1}\right) = \overline{1} \cdot \ldots \cdot \overline{1} = \overline{1}.$$

Следовательно, $\overline{k_1 \cdot \ldots \cdot k_n}^{-1} = \left(\overline{k_1} \cdot \ldots \cdot \overline{k_n}\right)^{-1} = \overline{k_1}^{-1} \cdot \ldots \cdot \overline{k_n}^{-1}$. ◁

Следствие. Если p – простое число, то в $\mathbf{Z}/p\mathbf{Z}$ каждый ненулевой класс обратим.

Поскольку $\mathbf{Z}/m\mathbf{Z}$ – конечное множество, то сложение и умножение на нем можно задавать поэлементно в виде таблиц. На пересечении i-й строки и j-го столбца таблицы записывается $x_i \circ x_j$ (где $\circ$ – знак операции). Так задавать алгебраические операции на конечном множестве впервые предложил известный английский математик XIX в. Артур Кэли (1821–1895), поэтому такие таблицы называются *таблицами Кэли*.

Пример 1.7.1. Построим таблицы Кэли сложения и умножения в $\mathbf{Z}/3\mathbf{Z}$ и в $\mathbf{Z}/4\mathbf{Z}$ – рис. 1.7.1 и 1.7.2 соответственно – и найдем обратимые элементы.

$$\mathbf{Z}/3\mathbf{Z} = \left\{\overline{0}, \overline{1}, \overline{2}\right\} = \{3\mathbf{Z},\ 3\mathbf{Z} + 1,\ 3\mathbf{Z} + 2\};$$

$$\mathbf{Z}/4\mathbf{Z} = \left\{\overline{0}, \overline{1}, \overline{2}, \overline{3}\right\} = \{4\mathbf{Z},\ 4\mathbf{Z} + 1,\ 4\mathbf{Z} + 2,\ 4\mathbf{Z} + 3\}.$$

+	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

а

×	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

б

Рис. 1.7.1

Из таблицы умножения на рис. 1.7.1, б видно, что $\bar{1}$ и $\bar{2}$ обратны сами себе.

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

а

×	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

б

Рис. 1.7.2

Из таблицы умножения на рис. 1.7.2, б видно, что $\bar{1}$ и $\bar{3}$ обратны сами себе, а $\bar{2}$ необратим.●

Определение 1.7.2. *Функция Эйлера* (или *тотиент-функция*) φ ставит в соответствие каждому натуральному $m > 1$ количество натуральных чисел, меньших m и взаимно простых с m.

Эта функция обладает следующими свойствами.

Теорема 1.7.2 (о вычислении значений функции Эйлера).

1. $\varphi(p) = p - 1$ для каждого простого числа p.

2. $\varphi(p^s) = p^s - p^{s-1}$, $\forall s \in \mathbf{N}$.

3. Если $(n, m) = 1$, то $\varphi(n \cdot m) = \varphi(n) \cdot \varphi(m)$.

4. Если $n = p_1^{s_1} \cdot \ldots \cdot p_t^{s_t}$ – каноническое разложение числа n, то

$$\varphi(n) = \left(p_1^{s_1} - p_1^{s_1-1}\right) \cdot \ldots \cdot \left(p_t^{s_t} - p_t^{s_t-1}\right) = n \cdot (1 - 1/p_1) \cdot \ldots \cdot (1 - 1/p_t).$$

▹ **1.** Количество чисел множества $\{1, 2, 3, \ldots, p-1\}$, взаимно простых с p, равно $p - 1$, т. к. любое число этого множества взаимно просто с p.

2. Рассмотрим множество $\{1, 2, 3, \ldots, p-1, p, p+1, \ldots, 2p, \ldots, 3p, \ldots, p^{s-1}, \ldots, p^s\}$. Разобьем это множество на подмножества $\{1, 2, 3, \ldots, p-1, p\}$, $\{p+1, \ldots, 2p\}, \ldots$, $\{p^s - p + 1,\ p^s - p + 2, \ldots,\ p^s - p + p = p^s\}$. В каждом подмножестве число элементов, взаимно простых с p, равно $p - 1$. Всего таких подмножеств p^{s-1}. Поэтому $\varphi(p^s) = (p-1) \cdot p^{s-1} = p^s - p^{s-1}$.

3. Рассмотрим множество $\{1, 2, \ldots, n, n+1, \ldots, 2n, \ldots, (m-1)n, \ldots, mn\}$. Разобьем его на подмножества $\{1, 2, 3, \ldots, n\}$, $\{n+1, n+2, \ldots, 2n\}$, $\{2n+1, \ldots, 3n\}, \ldots$ $\ldots$, $\{(m-2)n+1, \ldots, (m-1)n\}$, $\{(m-1)n+1, \ldots, (m-1)n+n = mn\}$. Числа, взаимно простые с $m \cdot n$, взаимно просты с m и n одновременно. В первом подмножестве

количество чисел, взаимно простых с n, равно $\varphi(n)$ по определению. Во всех подмножествах числа имеют вид $b = kn + r$, $1 \le r \le n$, $k = \overline{0, m-1}$, $(b, n) = (n, r)$ (теорема 1.2.1), т. е. количество чисел, взаимно простых с n, тоже равно $\varphi(n)$. Всего элементов, взаимно простых с n, в исходном множестве – $\varphi(n) \cdot m$. Пусть $d_1, ..., d_{\varphi(n)}$ – числа первого подмножества, взаимно простые с n. Тогда $kn + d_i \not\equiv ln + d_i \pmod m$, $0 \le k, l \le m-1$, $k \ne l$, т. к. $(k-l)n \not\equiv 0 \pmod m$, ведь $(n, m) = 1$, $0 < |k-l| < m$. Таким образом, $\left\{\overline{kn + d_i},\ k = \overline{0, m-1}\right\} = \mathbf{Z}/m\mathbf{Z}$ для фиксированного d_i, $i = \overline{1, \varphi(n)}$. Все множество чисел, взаимно простых с n, разобьем на $\varphi(n)$ подмножеств по m элементов в каждом, вычеты которых по модулю m все различны. В каждом таком подмножестве $\varphi(m)$ элементов, взаимно простых с m, поэтому всего $\varphi(n) \cdot \varphi(m)$ элементов, взаимно простых с n и m одновременно. Значит, $\varphi(n \cdot m) = \varphi(n) \cdot \varphi(m)$.

4. Пусть $n = p_1^{s_1} \cdot ... \cdot p_t^{s_t}$ – каноническое разложение числа n. Тогда

$$\varphi(n) = \varphi\left(p_1^{s_1} \cdot ... \cdot p_t^{s_t}\right) = \varphi\left(p_1^{s_1}\right) \cdot ... \cdot \varphi\left(p_t^{s_t}\right) = \left(p_1^{s_1} - p_1^{s_1-1}\right) \cdot ... \cdot \left(p_t^{s_t} - p_t^{s_t-1}\right) =$$
$$= p_1^{s_1} \cdot ... \cdot p_t^{s_t} \cdot (1 - 1/p_1) \cdot ... \cdot (1 - 1/p_t) = n \cdot (1 - 1/p_1) \cdot ... \cdot (1 - 1/p_t). \triangleleft$$

Пример 1.7.2. $n = 72 = 2^3 \cdot 3^2$.

Следовательно, $\varphi(72) = (2^3 - 2^2) \cdot (3^2 - 3) = 72 \cdot (1 - 1/2) \cdot (1 - 1/3) = 24$.●

Из теоремы 1.7.1 следует, что при $m \in \mathbf{N}_{>1}$ в $\mathbf{Z}/m\mathbf{Z}$ имеется в точности $\varphi(m)$ обратимых классов.

Пример 1.7.3. $\varphi(18) = \varphi(2 \cdot 3^2) = 6$, значит, в $\mathbf{Z}/18\mathbf{Z}$ имеется в точности 6 обратимых элементов. Ими являются классы $\bar{1}, \bar{5}, \bar{7}, \overline{11}, \overline{13}, \overline{17}$.●

Теорема 1.7.3 (Л. Эйлер). Для $\forall\, m \in \mathbf{N}_{>1}$ и $\forall\, a \in \mathbf{Z}$ $(a, m) = 1$ тогда и только тогда, когда $a^{\varphi(m)} \equiv 1 \pmod m$.

▷ Необходимость. Если $(a, m) = 1$, то $\bar{a}$ – обратимый элемент $\mathbf{Z}/m\mathbf{Z}$ по теореме 1.7.1. Пусть $\overline{k_1}, ..., \overline{k_{\varphi(m)}}$ – все обратимые элементы в $\mathbf{Z}/m\mathbf{Z}$. Рассмотрим

$$\left(\bar{a} \cdot \overline{k_1}\right) \cdot ... \cdot \left(\bar{a} \cdot \overline{k_{\varphi(m)}}\right) = \bar{a}^{\varphi(m)} \cdot \overline{k_1} \cdot ... \cdot \overline{k_{\varphi(m)}} = \overline{k_1} \cdot ... \cdot \overline{k_{\varphi(m)}}.$$

Действительно, $\bar{a} \cdot \overline{k_i} \ne \bar{a} \cdot \overline{k_j}$ при $i \ne j$ по лемме 1.7.1, тогда $\bar{a} \cdot \overline{k_i}$, $i = \overline{1, \varphi(m)}$, – все различные обратимые классы по теореме 1.7.1. Следовательно,

$$a^{\varphi(m)} \cdot k_1 \cdot ... \cdot k_{\varphi(m)} \equiv k_1 \cdot ... \cdot k_{\varphi(m)} \pmod m.$$

Так как $(k_1, m) = ... = (k_{\varphi(m)}, m) = 1$, то по свойству 6 сравнений $a^{\varphi(m)} \equiv 1 \pmod m$.

Достаточность. $\varphi(m) \ge 1$, значит, $a^{\varphi(m)} \equiv 1 \pmod m \Rightarrow \bar{a} \cdot \bar{a}^{\varphi(m)-1} = \bar{1} \Rightarrow \bar{a}$ – обратимый элемент в $\mathbf{Z}/m\mathbf{Z} \Rightarrow (a, m) = 1$ (по теореме 1.7.1). ◁

Следствие. В $\mathbf{Z}/m\mathbf{Z}$ при $m \in \mathbf{N}_{>1}$ всякий обратимый элемент $\bar{k}$ обладает свойствами:

1) $\bar{k}^{\varphi(m)} = \bar{1}$;

2) обратным классу $\bar{k}$ является класс вычетов $\overline{k^{\varphi(m)-1}}$.

Теорема 1.7.4 (малая теорема Ферма). Пусть p – простое число. Число $a \in \mathbf{Z}$ не делится на p тогда и только тогда, когда $a^{p-1} \equiv 1 \pmod{p}$.

▷ Доказательство следует из теоремы Эйлера, поскольку $\varphi(p) = p - 1$. ◁

Следствие 1. В $\mathbf{Z}/p\mathbf{Z}$ при простом p обратным классу $\overline{k} \neq \overline{0}$ является класс вычетов $\overline{k^{p-2}}$.

Следствие 2. Если $a^{m-1} \not\equiv 1 \pmod{m}$, $m \in \mathbf{N}_{>1}$, для некоторого целого a при $(a, m) = 1$, то m – составное число.

Этот факт часто используется для проверки числа на простоту. Он позволяет установить наличие нетривиальных множителей данного числа m, не находя ни одного из таких множителей.

Решение линейных сравнений в целых числах

Рассмотрим сравнение вида

$$ax \equiv b \pmod{m}, \qquad (1.7.1)$$

где $a, b \in \mathbf{Z}$, $m \in \mathbf{N}$, x – искомое значение в $\mathbf{Z}$.

1. Пусть $(a, m) = 1$. Тогда сравнение (1.7.1) имеет в качестве множества решений $\overline{x}$ – единственный класс вычетов в $\mathbf{Z}/m\mathbf{Z}$. По теореме 1.7.1 $\overline{a}$ – обратимый класс в $\mathbf{Z}/m\mathbf{Z}$ и $\overline{a}^{-1}$ – единственный обратный $\overline{a}$ класс. В $\mathbf{Z}/m\mathbf{Z}$ сравнение (1.7.1) соответствует уравнению $\overline{a} \cdot \overline{x} = \overline{b}$. Умножив обе его части на $\overline{a}^{-1}$, получим $\overline{x} = \overline{a}^{-1} \cdot \overline{b}$. Тогда $\overline{x} = \{x + mq \mid q \in \mathbf{Z}\}$ – множество решений (1.7.1). Представитель класса $\overline{a}^{-1}$ можно найти, определив коэффициент при a в соотношении Безу для 1 и чисел a и m, или согласно следствию из теоремы 1.7.3 $\overline{a}^{-1} = \overline{a^{\varphi(m)-1}}$.

2. Пусть $(a, m) = d > 1$ и $d \nmid b$, тогда сравнение (1.7.1) не имеет решений в $\mathbf{Z}$, поскольку не выполняется свойство 6 делимости целых чисел: $ax = b + mq$, $q \in \mathbf{Z}$, $d \mid a$, $d \mid m$, но $d \nmid b$.

3. Пусть $(a, m) = d > 1$ и $d \mid b$. Тогда разделим обе части (1.7.1) и m на d согласно свойству 5 сравнений. Получим сравнение

$$a_1 x \equiv b_1 \pmod{m_1}, \qquad (1.7.2)$$

где $a_1 = a/d$, $b_1 = b/d$, $m_1 = m/d$, и $(a_1, m_1) = 1$ по свойству 1 взаимно простых чисел. Сравнение (1.7.2) имеет в качестве множества решений единственный класс вычетов $\overline{x_0}$ в $\mathbf{Z}/m_1\mathbf{Z}$ согласно случаю 1. Числа из $\overline{x_0} = \{x_0 + m_1 q \mid q \in \mathbf{Z}\}$ являются решениями (1.7.1): $a(x_0 + m_1 q) = da_1 x_0 + ma_1 q \equiv db_1 + dm_1 t_0 \equiv b \pmod{m}$ для $\forall\, q \in \mathbf{Z}$ и фиксированного $t_0 \in \mathbf{Z}$. Пусть $x = x_0 + y$ – решение (1.7.1), тогда $ax \equiv ax_0 \equiv b \pmod{m}$. Это возможно, когда $ay \equiv 0 \pmod{m} \Leftrightarrow m \mid ay \Leftrightarrow m_1 \mid a_1 y$ и $m_1 \mid y$ по свойству 2 взаимно простых чисел, значит, $x \in \overline{x_0}$. Итак, числа из $\overline{x_0}$, и только они, являются решениями сравнения (1.7.1). Поскольку $y \equiv 0 \pmod{m_1}$, то $y \equiv 0,\ m_1,\ 2m_1, \ldots,\ (d-1)m_1 \pmod{m}$ в соответствии с остатками от деления $q \in \mathbf{Z}$ на d. Все множество решений сравнения (1.7.1) – это объединение d множеств классов вычетов по модулю m: $\overline{x_0}, \overline{x_0 + m_1}, \overline{x_0 + 2m_1}, \ldots, \overline{x_0 + (d-1)m_1}$.

Глава 2. Отображения и их свойства

§2.1. Соответствия, отображения, функции

Определение 2.1.1. Пусть X и Y – два непустых множества. Если определен способ сопоставления элементов Y элементам X, то говорят, что между множествами X и Y установлено *соответствие*. Если обозначить соответствие буквой q, то запись $q: X \to Y$ обозначает существование данного соответствия между множествами X и Y. При этом совершенно не обязательно, чтобы в сопоставлении участвовали все элементы множеств X и Y. Для того чтобы задать соответствие между множествами X и Y, нужно задать множество $Q \subseteq X \times Y = \{(x, y) \mid x \in X, y \in Y\}$, определяющее закон, по которому осуществляется соответствие, т. е. перечисляющий все пары (x, y), участвующие в сопоставлении.

Таким образом, соответствие, обозначаемое q, представляет собой тройку множеств:

$$q = (X, Y, Q),$$

в которой $Q \subseteq X \times Y$. В этом выражении первую компоненту X называют *областью отправления соответствия*, вторую компоненту Y – *областью прибытия соответствия*, третью компоненту Q – *графиком соответствия*. Термин «график» будет более подробно разъяснен при рассмотрении частного вида соответствия, называемого функцией.

Кроме рассмотренных множеств X, Y и Q с каждым соответствием q неразрывно связаны еще два множества: множество $D(q) = \{x \in X \mid (x, y) \in Q\}$, называемое *областью определения соответствия*, которое состоит из всех элементов множества X, участвующих в сопоставлении, и множество $E(q) = \{y \in Y \mid (x, y) \in Q\}$, называемое *областью значений соответствия*, которое состоит из всех элементов множества Y, участвующих в сопоставлении. Если $(x, y) \in Q$, то говорят, что *элемент y соответствует элементу x*. Геометрически этот факт удобно изображать стрелкой, направленной от x к y.

Множество всех $y \in E(q)$, соответствующих фиксированному элементу $x \in D(q)$, называется *образом x* в Y при соответствии q и обозначается $q(x)$. Множество всех $x \in D(q)$, которым соответствует фиксированный элемент $y \in E(q)$, называется *прообразом y* в X при соответствии q и обозначается $q^{-1}(y)$. Если $C \subseteq D(q)$, то *образом множества C* $q(C)$ называется объединение образов всех элементов из C. Аналогично определяется *прообраз множества D* $q^{-1}(D)$ для любого $D \subseteq E(q)$ как объединение прообразов всех элементов из D.

Определение 2.1.2. Если $D(q) = X$, то соответствие q называется *всюду определенным* или *отображением X в Y* (в противном случае соответствие называется *частичным*). Если $E(q) = Y$, то соответствие q называется *сюръективным* (*сюръекцией*) на Y. Соответствие q называется *инъективным* (*инъекцией*), если

любые различные x_1 и x_2 из $D(q)$ имеют различные образы и любые различные y_1 и y_2 из $E(q)$ имеют различные прообразы при соответствии q.

Два отображения p и q называются *равными* (обозначение $p = q$), если их области определения – одно и то же множество X, и для любого $x \in X$ $p(x) = q(x)$.

Отображение, для которого область определения и область прибытия являются одним и тем же множеством X, часто называют *преобразованием множества X*.

Определение 2.1.3. Соответствие q называется *функциональным* (или *однозначным*), если образом любого элемента $x \in D(q)$ является единственный элемент $y \in E(q)$, что обычно записывается как $q : x \mapsto y$ или $q(x) = y$. Соответствие q между множествами X и Y называется *взаимно однозначным* или *биективным* (*биекцией*, иногда *1-1 соответствием*), если оно всюду определено, сюръективно и инъективно. Однозначное отображение называется *функцией*. Функция называется *инъективной*, если различным x_1 и x_2 из X соответствуют различные y_1 и y_2 из Y, и *сюръективной*, если она сюръективна как соответствие. Функция называется *биективной*, если она одновременно инъективна и сюръективна.

Определение 2.1.4. Пусть $f: X \to Y$ – функция. Каждому элементу $x \in X$ f ставит в соответствие единственный элемент $y \in Y$: $f(x) = y$. Элемент x называется *аргументом функции*, y – *значением функции* на x. Если $E(f)$ состоит из единственного элемента, то f называется *функцией-константой*. *Тождественной функцией* на множестве X называется функция $e_X: X \to X$, такая, что $e_X(x) = x$ для любого $x \in X$. Если X, $Y \subseteq \mathbf{R}$, то функцию f называют *вещественной*.

Определение 2.1.5. Если f – вещественная функция, то упорядоченные пары $(x, f(x))$ можно изобразить в виде точек на плоскости $\mathbf{R}^2$. Полная совокупность таких точек будет представлять собой *график функции f*.

Например, график вещественной функции $f: X \to Y$, где $X = [-1; 1]$, $Y = [0; 1]$, $Q_f = \{(x, y) \in X \times Y \mid x^2 + y^2 = 1\} = \Gamma$ представлен на рис. 2.1.1.

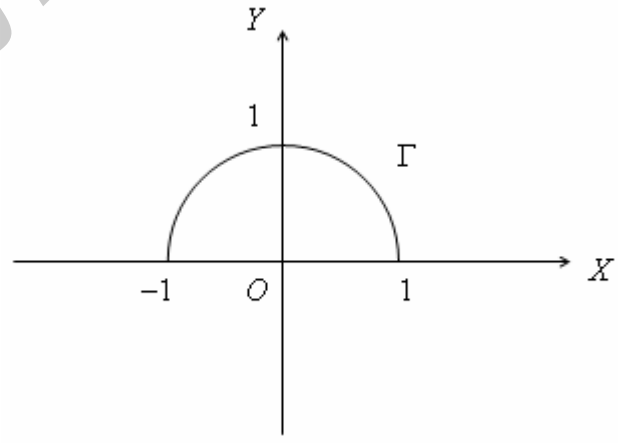


Рис. 2.1.1

Определение 2.1.6. Для каждого соответствия $q = (X, Y, Q)$ с $Q \subseteq X \times Y$ существует *обратное соответствие*, которое получится, если данное соответствие q рассматривать в обратном направлении, т. е. определять элементы $x \in X$, с которыми сопоставляются элементы $y \in Y$. Соответствие, обратное соответствию q, будем обозначать

$$q^{-1} = (Y, X, Q^{-1}),$$

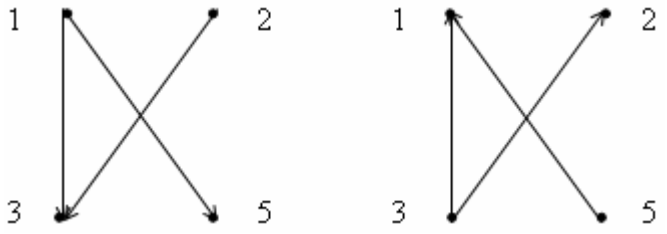
где $Q^{-1} \subseteq Y \times X$. Геометрическое представление обратного соответствия получается путем изменения направления стрелок в геометрическом представлении прямого соответствия. Отсюда следует, что обратным соответствием для обратного соответствия будет прямое соответствие:

$$(q^{-1})^{-1} = q.$$

Пример 2.1.1. Пусть $X = \{1, 2\}$, $Y = \{3, 5\}$, так что $X \times Y = \{(1, 3), (1, 5), (2, 3), (2, 5)\}$. Это множество дает возможность получить $2^4 = 16$ различных соответствий. Графики соответствий: $Q_0 = \{(\,)\} = \varnothing$, $Q_1 = \{(1, 3)\}$, $Q_2 = \{(1, 5)\}$, $Q_3 = \{(2, 3)\}$, $Q_4 = \{(2, 5)\}$, $Q_5 = \{(1, 3), (1, 5)\}$, $Q_6 = \{(1, 3), (2, 3)\}$, $Q_7 = \{(1, 3), (2, 5)\}$, $Q_8 = \{(1, 5), (2, 3)\}$, $Q_9 = \{(1, 5), (2, 5)\}$, $Q_{10} = \{(2, 3), (2, 5)\}$, $Q_{11} = \{(1, 3), (1, 5), (2, 3)\}$, $Q_{12} = \{(1, 3), (1, 5), (2, 5)\}$, $Q_{13} = \{(1, 3), (2, 3), (2, 5)\}$, $Q_{14} = \{(1, 5), (2, 3), (2, 5)\}$, $Q_{15} = \{(1, 3), (1, 5), (2, 3), (2, 5)\} = X \times Y$. Обозначим q_i соответствие с графиком Q_i, $i = \overline{0,15}$.

Областью определения соответствия q_9 является $\{1, 2\} = X$, поэтому q_9 – отображение. Областью значений q_9 является $\{5\} \neq Y$, поэтому q_9 несюръективно. Элемент 5 является образом 1 и 2 при соответствии q_9, поэтому q_9 неинъективно. Образом множества X при соответствии q_9 является множество $\{5\}$. Прообразом множества $\{5\}$ при соответствии q_9 является множество X. Соответствие q_9 функционально, т. к. каждый из элементов 1 и 2 имеет единственный образ – элемент 5.

На рис. 2.1.2, а изображено геометрическое представление соответствия q_{11}. Графиком обратного соответствия q_{11}^{-1} является множество $Q_{11}^{-1} = \{(3, 1), (5, 1), (3, 2)\}$. Геометрическое представление q_{11}^{-1} приведено на рис. 2.1.2, б.



а **б**

Рис. 2.1.2

Отображениями являются соответствия q_6–q_9, q_{11}–q_{15}. Сюръективными соответствиями являются q_5, q_7, q_8, q_{10}–q_{15}. Функциональными соответствиями яв-

ляются q_1–q_4, q_6–q_9. Инъективные соответствия: q_1–q_4, q_7, q_8. Функциями являются q_6–q_9. Биективные функции: q_7, q_8.•

В дальнейшем будем рассматривать только функциональные соответствия, которые также иногда для краткости будем называть функциями.

Перечислим основные способы задания функций.

1. Наиболее простой способ задания функций – это *табличный*. Таблицы при этом представляют собой конечные списки пар $(x, f(x))$. Однако таким способом могут быть заданы только функции, определенные на конечных множествах.

2. Другим не менее известным способом задания функций является *аналитический*, или *формула*, описывающая функцию с помощью суперпозиции других (исходных) функций. Если способ вычисления исходных функций известен, то формула задает процедуру вычисления данной функции как некоторую последовательность вычислений исходных функций.

Иногда для разных подмножеств множества X при задании функции приходится пользоваться различными формулами. Пусть $A_i \subset X$, $i = \overline{1,n}$, $X = A_1 \cup \ldots \ldots \cup A_n$, $A_i \cap A_j = \varnothing$ при $i \neq j$. Обозначим через $f_i(x)$ формулу, определяющую y при $x \in A_i$, $i = \overline{1,n}$. Тогда функция f, определенная на всем множестве X, задается так:

$$f(x) = \begin{cases} f_1(x) \text{ при } x \in A_1; \\ f_2(x) \text{ при } x \in A_2; \\ \ldots \\ f_n(x) \text{ при } x \in A_n. \end{cases}$$

Так, функцию $f(x) = |x|$ на $\mathbf{R} = \mathbf{R}_{\geq 0} \cup \mathbf{R}_{<0}$ можно задать в виде

$$|x| = \begin{cases} x \text{ при } x \geq 0; \\ -x \text{ при } x < 0. \end{cases}$$

3. Если f – вещественная функция, то она может быть задана графически на плоскости $\mathbf{R}^2$, как сказано ранее в определении 2.1.5.

4. Вычисления функций по таблицам, формулам, а также с помощью графиков являются частными видами вычислительных процедур. Существуют вычислительные процедуры, не относящиеся к указанным трем видам. Среди них особенно следует выделить рекурсивные процедуры. *Рекурсивная процедура* задает функцию f, определенную на множестве $\mathbf{N}$ ($\mathbf{Z}_{\geq 0}$), следующим образом: 1) задается значение $f(1)$ ($f(0)$); 2) значение $f(n+1)$ определяется через суперпозицию $f(n)$ и других функций, считающихся известными. Простейшим примером рекурсивной процедуры является вычисление функции $n!$: 1) $0! = 1$; 2) $(n+1)! = n!(n+1)$. Для вычисления $(n+1)!$ при $n \in \mathbf{N}$ требуется $n-1$ умножение, т. е. число вычислительных шагов увеличивается с ростом аргумента.

Пример 2.1.2. Пусть $X = Y = \mathbf{R}$, $f: X \to Y$, где $f: x \mapsto x^3$.

Здесь f является числовой функцией, поскольку $D(f) = \mathbf{R}$ и данное отображение однозначно.

Пусть $x_1 \neq x_2$, тогда ${x_1}^3 \neq {x_2}^3$, потому что иначе получаем ${x_1}^3 - {x_2}^3 = (x_1 - x_2)\left({x_1}^2 + x_1 \cdot x_2 + {x_2}^2\right) = 0$. Поскольку $x_1 \neq x_2$, получаем, что ${x_1}^2 + x_1 \cdot x_2 + {x_2}^2 = 0$. Но при $x_2 \neq 0$ дискриминант $D = {x_2}^2 - 4{x_2}^2 = -3{x_2}^2 < 0$ и равенство ${x_1}^3 = {x_2}^3$ невозможно ни для каких $x_1 \in \mathbf{R}$. Если $x_2 = 0$, то $x_1 = 0$ и имеем противоречие тому, что $x_1 \neq x_2$. Итак, f – инъективная функция.

Для $\forall\, y \in \mathbf{R}$ $y = x^3$ при $x = \sqrt[3]{y} \in \mathbf{R}$. Поэтому f – сюръективная функция.

Значит, данная функция $f\colon X \to Y$ биективна.●

Определение 2.1.7. Пусть даны две функции $f\colon X \to Y_1$ и $g\colon Y_2 \to Z$, $Y_1 \subseteq Y_2$. Функция $h\colon X \to Z$ называется *композицией функций* f и g (обозначение $h = g \circ f$ или $h = gf$), если h – последовательное применение функций f и g: для любого $x \in X$ $h(x) = g(f(x))$. Часто говорят, что функция h получена *подстановкой* f в g.

$$gf : x \mapsto z,\ z = gf(x) = g(f(x)),$$

$$gf : x \xrightarrow{\ f\ } y \xrightarrow{\ g\ } z.$$

Аналогично, по индукции определяется композиция n отображений для любого натурального числа $n \geq 2$.

Пример 2.1.3. $f\colon \mathbf{R} \to \mathbf{R}$, $g\colon \mathbf{R} \to \mathbf{R}$, где $f(x) = \sin x$, $g(x) = \sqrt{x^2 - 5x + 9}$.

$D(f) = \mathbf{R}$. Поскольку дискриминант $D = 25 - 36 = -11 < 0$, то $x^2 - 5x + 9 > 0$ при $\forall\, x \in \mathbf{R}$ и функция g всюду на $\mathbf{R}$ определена. $D(g) = \mathbf{R}$ также. Построим композиции функций gf и fg. Очевидно, что $D(gf) = D(fg) = \mathbf{R}$.

$$gf(x) = \sqrt{\sin^2 x - 5\sin x + 9}\text{, при } x = 0\ gf(0) = 3;$$

$$fg(x) = \sin\sqrt{x^2 - 5x + 9}\text{, при } x = 0\ gf(0) = \sin 3 \neq 3\text{, поскольку } E(f) = [-1;\ 1].$$

Итак, $gf \neq fg$.●

Лемма 2.1.1. Пусть даны функции $f\colon X \to Y$ и $g\colon Y \to X$, причем их композиция $gf = e_X$. Тогда f – инъекция, а g – сюръекция.

▷ Пусть f не является инъекцией. Тогда найдутся два элемента $x_1, x_2 \in X$, такие, что $x_1 \neq x_2$, но $f(x_1) = f(x_2) = y \in Y$. Тогда $gf(x_1) = gf(x_2) = g(y)$, но т. к. $gf = e_X$, то $gf(x_1) = x_1$, а $gf(x_2) = x_2$, и мы получили противоречие тому, что $x_1 \neq x_2$. Итак, f – инъекция.

Теперь докажем, что g – сюръекция. Действительно, для любого $x \in X$ $x = gf(x) = g(y)$ для некоторого $y \in Y$. ◁

Для обратной функции удобно использовать еще одно определение.

Определение 2.1.8. Пусть функция $f\colon X \to Y$. Функция $f^{-1}\colon Y \to X$ называется *обратной* для функции f, если $f^{-1}f = e_X$, а $ff^{-1} = e_Y$, где e_X и e_Y – тождественные функции на множествах X и Y соответственно.

Рассмотрим пример *кодирования*, под которым будем понимать представление сообщений в форме, удобной для передачи по определенному каналу связи. Тогда для кодирующей функции обратной будет декодирующая функция. Если кодирующая функция несюръективна, то декодирующая функция не всюду определена.

При аналитическом задании функции f принято аргумент как прямой, так и обратной функции обозначать одной и той же буквой, например, x. Поэтому для нахождения обратной функции следует уравнение $y = f(x)$ разрешить (если это возможно) относительно x и поменять обозначения, заменив x на y и y на x. При этом формула для обратной функции запишется в виде $y = f^{-1}(x)$.

Теорема 2.1.1 (критерий существования обратной функции). Функция $f: X \to Y$ имеет обратную тогда и только тогда, когда f – биекция.

▷ Необходимость. Пусть функция f имеет обратную $f^{-1}: Y \to X$. Тогда по определению обратной функции $ff^{-1} = e_Y$. Следовательно, по лемме 2.1.1 f – сюръекция. Также по определению обратной функции $f^{-1}f = e_X$. Тогда по лемме 2.1.1 f – инъекция. Одновременно будучи сюръекцией и инъекцией, f – биекция.

Достаточность. Пусть функция f биективна. Поскольку f сюръективна, для любого $y \in Y$ существует $x \in X$, такой, что $f(x) = y$. Положим $f^{-1}(y) = x$. Значит, существует отображение $f^{-1}: Y \to X$. Так как f инъективна, любой элемент $y \in Y$ имеет единственный прообраз $x \in X$. Поэтому f^{-1} – функция. Очевидно, что f^{-1} является обратной для f функцией. ◁

Свяжем с функцией еще одно понятие.

Определение 2.1.9. Пусть $f: X \to Y$ – произвольная функция, $A \subset X$ – произвольное непустое собственное подмножество X. *Сужением функции f на множество A* называют функцию f_A, график которой Q_{f_A} состоит из тех и только тех пар (x, y) графика Q_f функции f, в которых $x \in A$, а значит, $(x, y) \in A \times Y$. Таким образом, $Q_{f_A} = Q_f \cap A \times Y$.

Операцию сужения функции часто используют для табличного задания функций с бесконечной областью определения X. В качестве множества A рассматривают обычно выборку равноотстоящих значений x множества X. Получаемое при этом сужение f_A функции f уже легко представить в виде таблицы. По этому принципу построены таблицы логарифмов, тригонометрических функций и некоторые другие.

Может так случиться, что сама функция, заданная на множестве X, не имеет обратной, но сужение этой функции на некоторое подмножество множества X, на котором она инъективна, уже имеет обратную функцию, определенную на области значений исходной функции. Например, $f(x) = x^2$, $f: \mathbf{R} \to \mathbf{R}_{\geq 0}$, обратной функции на $\mathbf{R}_{\geq 0}$ не имеет, т. к. обратное отображение каждому $y \in \mathbf{R}_{>0}$ ставит в соответствие два значения: $\sqrt{y}$ и $-\sqrt{y}$, где $\sqrt{\ }$ – арифметический квадратный корень. Такое отображение нефункционально. Но отображения $f_+^{-1}: \mathbf{R}_{\geq 0} \to \mathbf{R}_{\geq 0}$, где $f_+^{-1}(x) = \sqrt{x}$, и $f_-^{-1}: \mathbf{R}_{\geq 0} \to \mathbf{R}_{\leq 0}$, где $f_-^{-1}(x) = -\sqrt{x}$, являются функциями, обратными соответственно сужениям функции f на множества $\mathbf{R}_{\geq 0}$ и $\mathbf{R}_{\leq 0}$.

Пример 2.1.4. Рассмотрим примеры функциональных преобразований множеств X с различными свойствами.

1. $X = \mathbf{R}, f: X \to X,\ f(x) = \sqrt[3]{x^2 - 5}$. Функциональное отображение неинъективно, т. к., например, $f(1) = f(-1) = \sqrt[3]{-4}$, и несюръективно, т. к. нет, например, такого $x \in \mathbf{R}$, чтобы $f(x) = \sqrt[3]{-6}$.

2. $X = \mathbf{N},\ f: X \to X,\ f(n) = 5 \cdot n$. Функциональное отображение несюръективно, т. к., например, 1, 2, 3, 4 не имеют прообразов в **N**, и инъективно, поскольку при $n_1 \neq n_2\ \ f(n_1) = 5 \cdot n_1 \neq 5 \cdot n_2 = f(n_2)$, потому что иначе получим $5 \cdot (n_1 - n_2) = 0$, чего быть не может.

3. $X = \mathbf{Z}/6\mathbf{Z}, f: X \to X,\ f(\overline{n}) = \overline{5 \cdot n}$. (5, 6) = 1, тогда по лемме 1.7.1 f – биекция:

$$f(\overline{0}) = \overline{0},\ f(\overline{1}) = \overline{5},\ f(\overline{2}) = \overline{4},\ f(\overline{3}) = \overline{3},\ f(\overline{4}) = \overline{2},\ f(\overline{5}) = \overline{1}.$$

4. $X = \mathbf{Z}/6\mathbf{Z}, f: X \to X,\ f(\overline{n}) = \overline{3 \cdot n}$. (3, 6) = 3 > 1, тогда по лемме 1.7.2 f – не инъекция и не сюръекция:

$$f(\overline{0}) = \overline{0},\ f(\overline{1}) = \overline{3},\ f(\overline{2}) = \overline{0},\ f(\overline{3}) = \overline{3},\ f(\overline{4}) = \overline{0},\ f(\overline{5}) = \overline{3}.$$

5. $X = \mathbf{C},\ f: X \to X,\ f(x) = x^3$. Функция f сюръективна, поскольку для $\forall z \in \mathbf{C}\ \ \exists \sqrt[3]{z} \in \mathbf{C},$ такой, что $f(\sqrt[3]{z}) = (\sqrt[3]{z})^3 = z$. Но отображение неинъективно, потому что при $z \neq 0\ \exists$ щ$_k = \sqrt[3]{|z|}\, e^{2k\pi i/3}$, $k = \overline{0,2},$ – три различных значения аргумента со свойством $f(\text{щ}_k) = z$.

6. $X = \mathbf{R}, f: X \to X,\ f(x) = x^3$. f – биекция, как это было показано в примере 2.1.2. Сюръективность можно доказать и так: f – непрерывная на **R** функция, и $\lim\limits_{x \to \pm\infty} f(x) = \lim\limits_{x \to \pm\infty} x^3 = \pm\infty$. Отображение инъективно, потому что f – монотонно возрастающая функция на X: производная $f'(x) = 3x^2 > 0$ при $x \neq 0$, причем $f(0) = 0, f(x) \neq 0$ при $x \neq 0$.

7. $X = \mathbf{Z}, f: X \to X,\ f(x) = x^3$. Функция инъективна, потому что она инъективна на **R**, как было показано в п. 6 данного примера, а $\mathbf{Z} \subset \mathbf{R}$. Но отображение не является сюръективным на **Z**, поскольку, например, $\overline{\exists}\, x \in \mathbf{Z}$, такого, что $f(x) = x^3 = 2$.●

Свойства функций и их композиций

1. Композиция сюръективных функций сюръективна (следует из определений 2.1.7 и 2.1.3).

2. Композиция инъективных функций инъективна (следует из определений 2.1.7 и 2.1.3).

3. Композиция биективных функций биективна (следует из свойств 1 и 2).

4. Композиция функций в общем случае некоммутативна (пример 2.1.3).

5. Композиция функций ассоциативна.

▷ Пусть $Y_1 \subseteq Y_2,\ Z_1 \subseteq Z_2,\ f: X \to Y_1,\ g: Y_2 \to Z_1,\ h: Z_2 \to W$ – произвольные функции. Рассмотрим произвольный элемент $x \in X$ и $f(x) = y, g(y) = z, h(z) = w$.

Тогда

$$h(gf)(x) = h(gf(x)) = h(g(y)) = h(z) = w,$$
$$(hg)f(x) = hg(f(x)) = hg(y) = h(g(y)) = h(z) = w.$$

То есть для любого $x \in X$ $h(gf)(x) = (hg)f(x)$. Значит, $h(gf) = (hg)f$. ◁

6. Относительно операции композиции функций, являющихся преобразованиями одного множества X, имеется нейтральный элемент – e_X.

▷ Пусть $f: X \to X$ – произвольная функция. Тогда $e_X f(x) = e_X(f(x)) = f(x)$, $fe_X(x) = f(e_X(x)) = f(x)$ для любого $x \in X$. Итак, $e_X f = fe_X = f$ для любой функции $f: X \to X$. ◁

7. Функция, обратная биекции, сама является биекцией.

▷ Пусть $f: X \to Y$ – биекция. Тогда по теореме 2.1.1 для f существует обратная функция $f^{-1}: Y \to X$. Поскольку для функции f^{-1} обратная функция существует (ею является f), то снова же по теореме 2.1.1 f^{-1} – биекция. ◁

Приведем теорему о связи свойств инъективности и сюръективности функции, являющейся преобразованием конечного множества, что уже можно было заметить в частных случаях, в пп. 3, 4 примера 2.1.4.

Теорема 2.1.2. Пусть A – конечное множество и функция $f: A \to A$. f – сюръекция тогда и только тогда, когда f – инъекция.

▷ Необходимость. Пусть $A = \{a_1, \ldots, a_n\}$, $n \in \mathbf{N}$, и $f: A \to A$ – сюръекция. Тогда $E(f) = \{f(a_1), \ldots, f(a_n)\} = \{a_1, \ldots, a_n\}$, т. е. $E(f) = A$. Если бы f не была инъекцией, то в A нашлось бы по крайней мере два элемента $a_i \neq a_j$, таких, что $f(a_i) = f(a_j)$. Но тогда бы мы пришли к противоречию, получив, что $E(f) \subset A$ и, значит, $E(f) \neq A$.

Достаточность. Пусть $A = \{a_1, \ldots, a_n\}$, $n \in \mathbf{N}$, и $f: A \to A$ – инъекция. Тогда для различных a_i и a_j $f(a_i)$ и $f(a_j)$ различны. $E(f) \subseteq A$, и количества элементов в $E(f)$ и A совпадают и равны n. Поэтому $E(f) = A$. Значит, f – сюръекция. ◁

§2.2. Взаимно однозначное соответствие. Мощность множества. Конечные, счетные, несчетные, континуальные множества и их свойства

Иногда бывает необходимо сопоставлять друг с другом элементы некоторых множеств. Рассмотрим, например, два множества: стадо из четырехсот овец и рощу из четырехсот деревьев. Эти множества можно попарно сопоставить друг с другом, привязав овец к деревьям, так что каждая овца и каждое дерево будут в точности принадлежать одной определенной паре. Ни одно из данных множеств не может быть по такому принципу попарно сопоставлено, например, с рощей из семисот деревьев или с множеством из трехсот камней.

Для взаимно однозначного соответствия можно дать следующее определение.

Определение 2.2.1. *Взаимно однозначным соответствием* между двумя непустыми множествами A и B называется такое правило (или закон) f, по кото-

рому каждому элементу $a \in A$ сопоставляется единственный элемент $f(a) \in B$ и для любого элемента $b \in B$ существует единственный элемент $a \in A$, такой, что $f(a) = b$, другими словами, функция f задает биекцию между A и B.

Таким образом, кодирование можно рассматривать как взаимно однозначное отображение множества сообщений во множество сигналов или *кодовых последовательностей* (*кодовых слов*).

Определение 2.2.2. Множества A и B называются *равномощными* (обозначение $A \leftrightarrow B$), если между ними можно установить взаимно однозначное соответствие.

Очевидно, что если множество A равномощно B, а B равномощно C, то A равномощно C, т. е. $A \leftrightarrow B \;\&\; B \leftrightarrow C \Rightarrow A \leftrightarrow C$, – свойство транзитивности.

Определение 2.2.3. Число элементов в конечном множестве A называется *мощностью* A и часто обозначается $|A|$. Пустое множество, т. е. не содержащее элементов, относят к конечным, оно является множеством мощности 0: $|\varnothing| = 0$.

Мощность бесконечного множества – более сложное понятие. Оно будет рассмотрено ниже.

Теорема 2.2.1. Между непустыми конечными множествами A и B существует взаимно однозначное соответствие тогда и только тогда, когда $|A| = |B|$.

▷ Необходимость. Пусть $f: A \to B$ задает взаимно однозначное соответствие между множествами. Пусть $|A| \neq |B|$, допустим, $|A| < |B|$. Тогда, поскольку для любого $a \in A$ существует единственный элемент $b \in B$, такой, что $f(a) = b$, для некоторого $c \in B$ не найдется элементов из A с таким свойством, а это противоречит определению взаимно однозначного соответствия (сюръективности). Предположим теперь, что $|B| < |A|$. Тогда, поскольку f определено всюду на A, $\exists\, a_1, a_2 \in A$, такие, что $f(a_1) = f(a_2) = b \in B$. Снова получаем противоречие с тем, что f задает взаимно однозначное соответствие (противоречие инъективности). Итак, $|A| = |B|$.

Достаточность. Пусть $|A| = |B| = n \in \mathbf{N}$. Тогда $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_n\}$. Рассмотрим правило f, по которому каждому элементу $a_i \in A$ ставится в соответствие единственный элемент $b_i \in B$. Тогда для $\forall\, b_i \in B$ существует единственный элемент $a_i \in A$, такой, что $f(a_i) = b_i$. Значит, f задает взаимно однозначное соответствие между множествами A и B. ◁

В связи с этой теоремой становится понятно, почему два первых множества из примера в начале параграфа находятся во взаимно однозначном соответствии друг с другом, но не являются равномощными третьему и четвертому множествам. Сколько же существует таких взаимно однозначных соответствий для двух n-элементных множеств A и B?

Теорема 2.2.2. Общее число взаимно однозначных соответствий для двух n-элементных множеств равно $n!$.

▷ Следуя доказательству достаточности теоремы 2.2.1, первый элемент множества A может быть сопоставлен с любым из n элементов множества B. Для каждого из n таких сопоставлений второй элемент множества A может быть сопоставлен с любым из оставшихся $n-1$ элементов множества B и т. д.

После того как $n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot 2$ таких сопоставлений проведено для $n-1$ элементов множества A, в каждом случае последний элемент этого множества будет сопоставляться с единственным оставшимся элементом множества B. Таким образом, общее число взаимно однозначных соответствий для n-элементных множеств будет равно

$$n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot 2\cdot 1 = n!.\ \triangleleft$$

Теорема 2.2.3. Пусть $A_1,\ldots, A_n$ – конечные множества и $|A_i| = m_i,\ i=\overline{1,n}$. Тогда мощность множества $A_1\times A_2\times\ldots\times A_n = \{(a_1, a_2,\ldots, a_n) \mid a_i\in A_i,\ i=\overline{1,n}\}$ равна произведению мощностей множеств $A_1, A_2,\ldots, A_n$:

$$|A_1\times A_2\times\ldots\times A_n| = m_1\cdot m_2\cdot\ldots\cdot m_n.$$

$\triangleright$ Эту теорему докажем методом математической индукции. Для $n = 1$ теорема тривиально верна. Предположим, что она верна для $n = k$, и докажем ее справедливость для $n = k+1$. По индуктивному предположению $|A_1\times A_2\times\ldots\times A_k| = m_1\cdot m_2\cdot\ldots\cdot m_k$. Возьмем любой вектор $(a_1, a_2,\ldots, a_k)$ из $A_1\times A_2\times\ldots\times A_k$ и припишем справа элемент $a_{k+1}\in A_{k+1}$. Таким образом, из всех $m_1\cdot m_2\cdot\ldots\cdot m_k$ векторов из $A_1\times A_2\times\ldots\times A_k$ приписыванием справа элемента из A_{k+1} можно получить $m_1\cdot\ldots\cdot m_k\cdot m_{k+1}$ векторов из $A_1\times A_2\times\ldots\times A_k\times A_{k+1}$, причем все они различны и никаких других векторов в $A_1\times A_2\times\ldots\times A_{k+1}$ не содержится. Поэтому и для $n = k+1$ теорема верна и, следовательно, верна для любых $n\in\mathbf{N}$. $\triangleleft$

Если $A_i = A,\ i=\overline{1,n}$, то $\underbrace{A\times A\times\ldots\times A}_{n} = A^n = \{(a_1, a_2,\ldots, a_n) \mid a_i\in A,\ i=\overline{1,n}\}$.

Следствие. $|A^n| = |A|^n$ для любого конечного множества A и любого $n\in\mathbf{N}$.

Теорема 2.2.3 и следствие из нее лежат в основе очень многих комбинаторных фактов.

Определение 2.2.4. Пусть A – некоторое множество. *Множеством-степенью* или *булеаном множества A* называется множество $P(A) = \{X \mid X\subseteq A\}$ – множество всех подмножеств множества A.

Вернемся к теореме 2.2.1. Во-первых, она позволяет установить равенство мощностей двух конечных множеств, не вычисляя этих мощностей, а во-вторых, часто дает возможность определить мощность множества, установив его взаимно однозначное соответствие с множеством, мощность которого известна или легко вычисляется. В качестве иллюстрации этого приема докажем важную теорему о числе подмножеств конечного множества.

Теорема 2.2.4. Для любого конечного множества A, $|A| = n\in\mathbf{Z}_{\geq 0}$, число всех подмножеств A равно 2^n, т. е. $|P(A)| = 2^n$.

$\triangleright$ $|\varnothing| = 0$. Очевидно, что $P(\varnothing) = \{\varnothing\}$ и $|P(\varnothing)| = 1 = 2^0$.

Пусть $n\in\mathbf{N}$. Рассмотрим множество B_n всех двоичных слов (векторов) из нулей и единиц длиной n. Каждому подмножеству $B\subseteq A$ сопоставим двоичное слово S_B длиной n, построенное по следующему правилу:

$$S_B = (\alpha_1,\ldots, \alpha_n),\ \text{где}\ \alpha_i = \begin{cases} 1, & \text{если } a_i\in B,\\ 0, & \text{если } a_i\notin B, \end{cases}\ i=\overline{1,n}.$$

Таким образом, каждому подмножеству множества A ставится в соответствие единственное двоичное слово длиной n и каждое двоичное слово длиной n поставлено в соответствие единственному подмножеству множества A. Данное сопоставление задает взаимно однозначное соответствие между булеаном множества A $P(A)$ и множеством B_n. Согласно теореме 2.2.1 $|P(A)| = |B_n|$. А т. к. B_n является n-й декартовой степенью двухэлементного множества $\{0, 1\}$, то в силу следствия из теоремы 2.2.3 $|B_n| = 2^n$. Итак, $|P(A)| = 2^n$.◁

Пусть требуется определить число всех k-элементных подмножеств n-элементного множества A, $n \in \mathbf{Z}_{\geq 0}$. Очевидно, что k должно удовлетворять неравенству $0 \leq k \leq n$. Если $A = \varnothing$, то $n = k = 0$ и число 0-элементных подмножеств равно $|P(\varnothing)| = 1$. Пусть теперь $n \in \mathbf{N}$. Поскольку, как уже было сказано в доказательстве теоремы 2.2.4, множества $P(A)$ и B_n равномощны, то число всех k-элементных подмножеств в A равно числу всех двоичных слов длиной n, в которых k символов равны 1, а остальные равны 0. Это число равно числу сочетаний из n элементов по k: $C_n^k = \dfrac{n!}{k!(n-k)!}$. Когда $n = k = 0$, получаем $C_0^0 = 1$, т. е. частный случай этой формулы. Этот факт также является иллюстрацией применения теоремы 2.2.1.

Если множества являются бесконечными, то при установлении между ними взаимно однозначного соответствия мы наталкиваемся на трудности, связанные с необходимостью оперировать с бесконечно большим числом элементов множества. За основу для сопоставления бесконечных множеств принято брать множество натуральных чисел $\mathbf{N} = \{1, 2, 3,\ldots, n,\ldots\}$.

Определение 2.2.5. Множество, равномощное множеству натуральных чисел $\mathbf{N}$, называется *счетным*.

Вообще любое бесконечное подмножество множества $\mathbf{N}$ счетно. Действительно, пусть $N' \subset \mathbf{N}$. Выберем в N' наименьший элемент и обозначим его n_1; в $N' \setminus \{n_1\}$ выберем наименьший элемент и обозначим его n_2; наименьший элемент в $N' \setminus \{n_1, n_2\}$ обозначим n_3 и т. д. Для всякого натурального числа n имеется лишь конечное множество меньших натуральных чисел: $\varnothing$, если $n = 1$, и $\{1, 2,\ldots, n-1\}$, если $1 < n$. Таким образом, любой элемент N' рано или поздно получит свой номер. Эта нумерация, т. е. функция f: $N' \to \mathbf{N}$, $f(n_i) = i$, и есть взаимно однозначное соответствие между N' и $\mathbf{N}$.

Счетным является множество всех целых чисел $\mathbf{Z}$, хотя $\mathbf{N} \subset \mathbf{Z}$. Это можно установить, рассмотрев взаимно однозначное соответствие f: $\mathbf{Z} \to \mathbf{N}$, представленное на рис. 2.2.1.

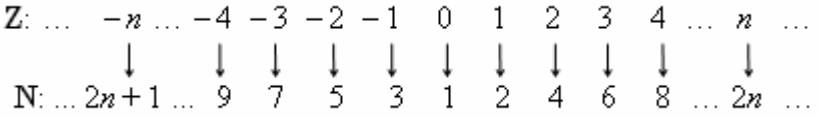


Рис. 2.2.1

$$f(z) = \begin{cases} 2z, & z \in \mathbf{N}; \\ 2|z|+1, & z \in \mathbf{Z}_{\le 0}. \end{cases}$$

Еще более удивителен тот факт, что счетным является множество всех рациональных чисел **Q**. Докажем сначала счетность множества $\mathbf{Q}_{>0}$. $\mathbf{Q}_{>0} = \{m/n \mid m, n \in \mathbf{N},\ \text{НОД}(m, n) = 1\}$. Представим множество $\mathbf{Q}_{>0}$ в виде следующей таблицы на рис. 2.2.2:

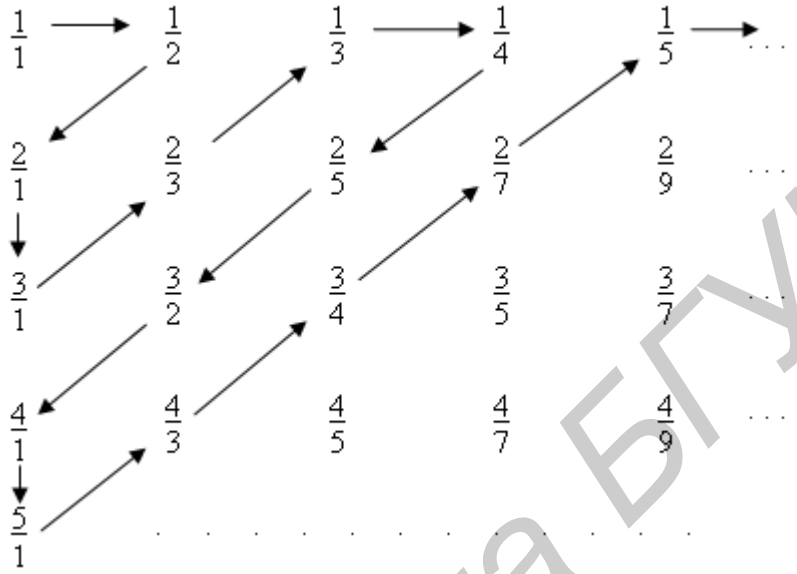


Рис. 2.2.2

Обходя данную таблицу по направлению стрелок, начиная от 1/1 = 1, приходим к последовательности 1, 1/2, 2, 3, 2/3, 1/3, 1/4, 2/5, 3/2, 4, 5, 4/3, 3/4, 2/7, 1/5,…, позволяющей занумеровать все эти числа. Обход таблицы можно осуществить и другими способами, например, по «квадратам». Итак, задано правило, по которому каждому числу $q \in \mathbf{Q}_{>0}$ ставится в соответствие единственное число $n \in \mathbf{N}$, являющееся номером q в данной последовательности. И для любого числа $n \in \mathbf{N}$ существует единственное число $q \in \mathbf{Q}_{>0}$, такое, что q имеет номер n в данной последовательности. Итак, $\mathbf{Q}_{>0}$ равномощно **N**.

Теперь рассмотрим множество всех рациональных чисел **Q**. Было показано выше, что $\mathbf{Q}_{>0} = \{q_1, q_2, q_3, \ldots, q_m, \ldots\} = \{q_i \mid i \in \mathbf{N}\}$. Обозначим теперь q_0 число 0. Построим взаимно однозначное соответствие f: $\mathbf{Q} \to \mathbf{N}$ таким образом, как показано на рис. 2.2.3.

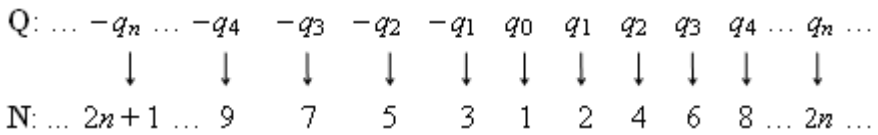


Рис. 2.2.3

$$f(q)=\begin{cases}2n,\ q=q_n\in\mathbf{Q}_{>0};\\ 2n+1,\ q=-q_n\in\mathbf{Q}_{\le 0}.\end{cases}$$

Из приведенных примеров видно одно из замечательных свойств бесконечных множеств, которое не имеет места в случае конечных множеств, – возможность приведения во взаимно однозначное соответствие бесконечного множества с его же бесконечным подмножеством.

Объединение конечного числа счетных множеств $A_1, A_2,\ldots, A_k$, $k\in\mathbf{N}$, счетно. Действительно, перенумеруем сначала все первые элементы множеств, затем – вторые и т. д. Объединение счетного множества конечных множеств также счетно (сначала нумеруем все элементы первого множества, затем все элементы второго множества и т. д.). Из последнего утверждения следует, что множество всех слов в любом конечном алфавите счетно. Менее очевидно, что счетно и объединение счетного множества счетных множеств. Примером такого объединения является множество $\bigcup\limits_{i\in\mathbf{N}}\mathbf{N}^i$ всех векторов с натуральными компонентами, где $\mathbf{N}^i$ – множество таких векторов длиной i.

Определение 2.2.6. Если бесконечное множество не равномощно множеству **N**, то такое множество называется *несчетным*.

Существование несчетных множеств следует из теоремы, доказанной в 1874 г. известным немецким математиком Георгом Кантором (1845–1918).

Теорема 2.2.5 (Г. Кантор). Множество всех действительных чисел интервала (0; 1) несчетно.

▷ Заметим, что любое число рассматриваемого интервала представляет собой конечную или бесконечную десятичную дробь вида $0,\alpha_1\alpha_2\alpha_3\ldots$ и может быть представлено точкой интервала вещественной прямой. Предположим, что множество всех вещественных чисел интервала (0; 1) счетно (это множество является бесконечным) и существует нумерация чисел данного множества. Расположим все числа, представленные бесконечными десятичными дробями (если дробь конечна, то, начиная с некоторого номера, в ее десятичных разрядах записываются только нули), в порядке этой нумерации:

$$0,\alpha_{11}\alpha_{12}\alpha_{13}\ldots$$
$$0,\alpha_{21}\alpha_{22}\alpha_{23}\ldots$$
$$0,\alpha_{31}\alpha_{32}\alpha_{33}\ldots$$
$$\ldots\ldots\ldots\ldots..$$

Рассмотрим любую бесконечную десятичную дробь $0,\beta_1\beta_2\beta_3\ldots$, такую, что $\beta_1\neq\alpha_{11}$, $\beta_2\neq\alpha_{22}$, $\beta_3\neq\alpha_{33}$ и т. д. Эта дробь не может войти в указанную последовательность, т. к. от первого числа она отличается первой цифрой, от второго числа – второй цифрой и т. д. Следовательно, все вещественные числа интервала (0; 1) не могут быть пронумерованы и данное множество несчетно. ◁

Определение 2.2.7. Мощность множества всех действительных чисел интервала (0; 1) называется *континуумом*, а множества такой мощности – *континуальными*. Метод, использованный при доказательстве теоремы 2.2.5, называется *диагональным методом Кантора*.

Необычные свойства несчетных множеств проявляются в том, что рассмотренный интервал (0; 1) может быть приведен во взаимно однозначное соответствие с полуинтервалами [0; 1) (рис. 2.2.4), (0; 1], отрезком [0; 1], а также множествами $(a; b)$, $(a; b]$, $[a; b)$, $[a; b]$, где $a, b \in \mathbf{R}$, $a < b$, и всем множеством $\mathbf{R}$.

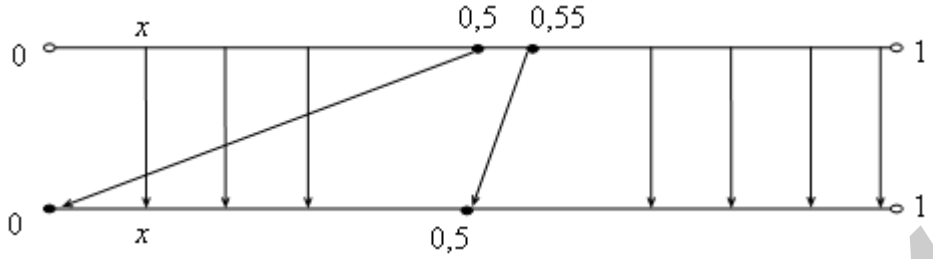


Рис. 2.2.4

$$f(x) = \begin{cases} x,\ x \neq 0,\underbrace{5\ldots5}_{n},\ n \in \mathbf{N}; \\ 0,\ x = 0{,}5; \\ 0,\underbrace{5\ldots5}_{n},\ x = 0,\underbrace{5\ldots5}_{n+1},\ n \in \mathbf{N}. \end{cases} \qquad f : (0;1) \to [0;1).$$

Аналогично доказывается, что $(0; 1) \leftrightarrow (0; 1]$, $(0; 1] \leftrightarrow [0; 1]$, $[0; 1) \leftrightarrow [0; 1]$. Поэтому $(0; 1) \leftrightarrow [0; 1]$ согласно свойству транзитивности взаимно однозначного соответствия.

На рис. 2.2.5 приведено непосредственное доказательство того, что $(0; 1) \leftrightarrow [0; 1]$.

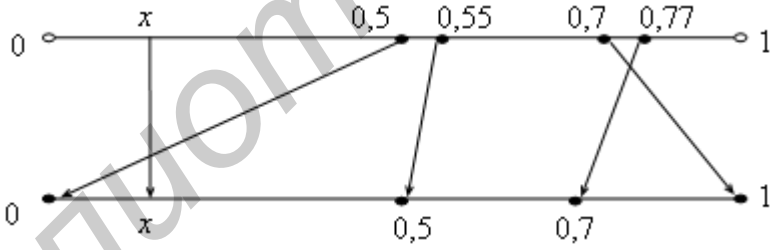


Рис. 2.2.5

$$f(x) = \begin{cases} x,\ x \neq 0,\underbrace{5\ldots5}_{n},\ 0,\underbrace{7\ldots7}_{n},\ n \in \mathbf{N}; \\ 0,\ x = 0{,}5; \\ 1,\ x = 0{,}7; \\ 0,\underbrace{5\ldots5}_{n},\ x = 0,\underbrace{5\ldots5}_{n+1},\ n \in \mathbf{N}; \\ 0,\underbrace{7\ldots7}_{n},\ x = 0,\underbrace{7\ldots7}_{n+1},\ n \in \mathbf{N}. \end{cases} \qquad f : (0;1) \to [0;1].$$

Аналогично можно показать, что $(a; b) \leftrightarrow (a; b]$, $(a; b) \leftrightarrow [a; b)$, $(a; b] \leftrightarrow$ $\leftrightarrow [a; b]$, $[a; b) \leftrightarrow [a; b]$, $(a; b) \leftrightarrow [a; b]$ для произвольных $a, b \in \mathbf{R}$, таких, что $a < b$.

Взаимно однозначное соответствие $(0; 1) \to (a; b)$ можно осуществить с помощью центральной проекции (рис. 2.2.6). Это же соответствие можно задать аналитически: $f(x) = (b - a)x + a$, $\forall\, x \in (0; 1)$.

В общем случае взаимно однозначное соответствие $f: (c; d) \to (a; b)$ для $\forall\, a, b, c, d \in \mathbf{R}$, $a < b$, $c < d$, задается аналитически следующим образом:

$$f(x) = \frac{b-a}{d-c}(x - c) + a\,,\ \forall\, x \in (c; d).$$

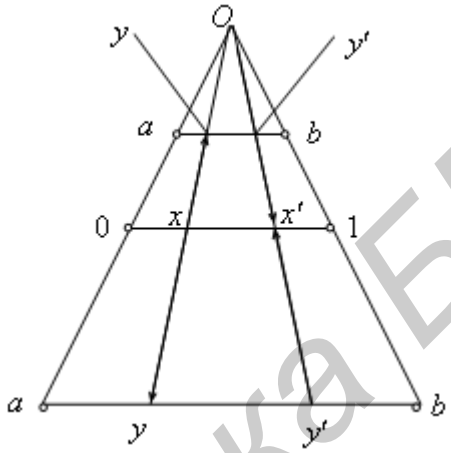


Рис. 2.2.6

Таким образом, $(0; 1) \leftrightarrow (-\pi/2; \pi/2)$. Рассмотрим биекцию $f: (-\pi/2; \pi/2) \to \mathbf{R}$, где $f(x) = \operatorname{tg} x$, $\forall\, x \in (-\pi/2; \pi/2)$. Получаем, что $(-\pi/2; \pi/2) \leftrightarrow \mathbf{R}$. Итак, $(0; 1) \leftrightarrow \mathbf{R}$ и $\mathbf{R}$ является континуальным множеством.

Интервал $(-\pi/2; \pi/2)$ можно представить полуокружностью длиной π. Тогда графически биекцию $\operatorname{tg}: (-\pi/2; \pi/2) \to \mathbf{R}$ можно представить в виде центральной проекции, как показано на рис. 2.2.7.

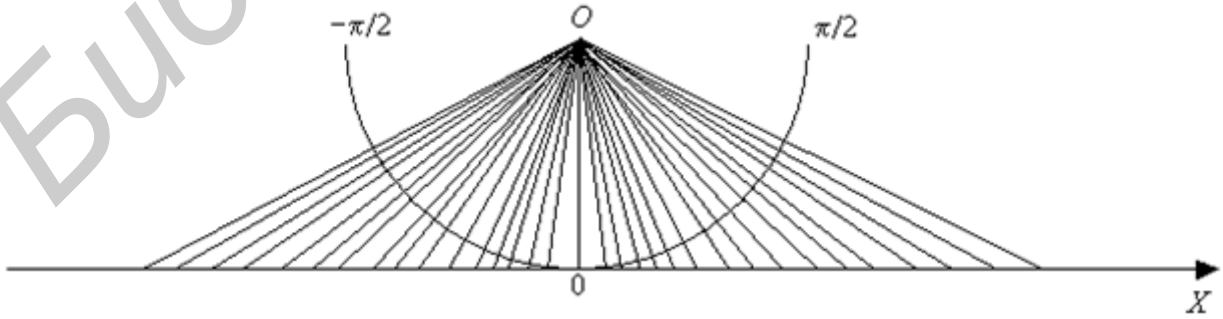


Рис. 2.2.7

Г. Кантор доказал, что $\mathbf{R}^2 \leftrightarrow \mathbf{R}$ и вообще $\mathbf{R}^n \leftrightarrow \mathbf{R}$ для любого $n \in \mathbf{N}$.

Множество всех подмножеств счетного множества континуально. Это становится ясным, если воспользоваться, как и в теореме 2.2.4, сопоставлением подмножеству последовательности (но теперь уже бесконечной!) из нулей и единиц: на i-м месте стоит 1, если i-й элемент множества принадлежит данному подмножеству, и 0 в противном случае. Получаем взаимно однозначное соответствие между собственными подмножествами счетного множества и правильными двоичными дробями, все множество которых, в свою очередь, взаимно однозначно соответствует множеству всех вещественных чисел интервала (0; 1). Как показано в теории множеств (с помощью метода, аналогичного диагональному методу Кантора), для множества любой мощности его булеан имеет более высокую мощность. Поэтому не существует множества максимальной мощности. Хорошо известным в истории математики является парадокс Кантора «множество всех множеств». По смыслу своего описания оно должно содержать все мыслимые множества и, следовательно, иметь максимальную мощность, что противоречит результатам теории множеств. Однако само это «множество всех множеств» содержится в своем булеане в качестве элемента и поэтому не может иметь максимальную мощность. И, таким образом, кажущееся противоречие разрешается.

§2.3. Классические шифры

Шифрование – это частный случай кодирования, преобразование сообщений в формулы, которые обеспечивают защиту информации от несанкционированного доступа.

Определение 2.3.1. *Криптографические преобразования шифрования* (*зашифровывания*) и *расшифровывания* определяются как взаимно обратные функции:

$$f_L(A) = B,\ g_L(B) = A,$$

где L – *ключ* – параметр шифра, определяющий выбор конкретного преобразования, известный отправителю и адресату;

A и B – соответственно исходное и закодированное сообщение, или *открытый текст* и *шифртекст* (*криптограмма*).

Совокупность преобразований f_L и набор ключей, которым они соответствуют, будем называть *шифром*. *Классическими шифрами* принято называть симметричные блочные шифры, т. е. те, которые для шифрования и расшифровывания используют один и тот же ключ и шифруют информацию блоками.

Определение 2.3.2. *Криптология* – наука, исследующая криптографические преобразования. В криптологии различают следующие направления: *криптографию* и *криптоанализ*. *Криптография* изучает принципы, средства и методы преобразования данных с целью сокрытия их информационного содержания, предотвращения их необнаружимой модификации и (или) несанкционированного использования. *Криптоанализ* – наука о методах получения исходного значения зашифрованной информации без возможности доступа к секрет-

ной информации (ключу), необходимой для этого. В большинстве случаев здесь речь идет о методах нахождения ключа. Но иногда под термином «криптоанализ» также понимают оценку сильных и слабых сторон методов шифрования и криптографических алгоритмов.

Для определения исходного текста по шифрованному при неизвестном ключе возможны два подхода: первый – определить ключ и затем найти исходное сообщение расшифровыванием, второй – найти исходное сообщение без определения ключа.

Определение 2.3.3. Получение открытого сообщения без заранее известного ключа по шифрованному называется *дешифрованием*, или *вскрытием шифра*, в отличие от процесса *расшифровывания*, когда ключ известен. Под *стойкостью шифра*, как правило, понимается способность противостоять попыткам произвести его вскрытие.

Криптография – одна из старейших наук, ее история насчитывает несколько тысяч лет. В XV–XVIII вв. в математике были заложены основы аппарата, применяемого для анализа шифров и дешифрования.

Определение 2.3.4. *Шифрами перестановки* называют такие шифры, преобразования из которых приводят к изменению только порядка следования символов исходного сообщения. *Шифрами замены* называют такие шифры, преобразования из которых приводят к замене каждого символа открытого сообщения на другие символы – *шифробозначения*, причем порядок следования шифробозначений совпадает с порядком следования соответствующих им символов открытого сообщения.

Шифры замены

Шифр простой однобуквенной замены. Рассмотрим на примере русского алфавита следующую таблицу (см. табл. 2.3.1).

Таблица 2.3.1

А	Б	В	Г	Д	…	Э	Ю	Я
f(А)	*f*(Б)	*f*(В)	*f*(Г)	*f*(Д)	…	*f*(Э)	*f*(Ю)	*f*(Я)

Вторая строка в табл. 2.3.1 представляет собой перестановку букв алфавита первой строки. При расшифровывании и шифровании надо помнить вторую строку, т. е. ключ. Обычно ее запомнить сложно, поэтому всегда пытались придумать какое-либо правило.

Одним из древнейших шифров, известных истории, был *шифр Цезаря*, для которого вторая строка в табл. 2.3.1 является циклически сдвинутой на определенное число позиций первой строкой, т. е. последовательностью, записанной в алфавитном порядке, но начинающейся не с буквы «А». Шифр назван так в честь римского императора Гая Юлия Цезаря (102 или 100 – 44 гг. до н. э.), использовавшего его для секретной переписки. Такой шифр, описанный в книге «Записки о галльской войне», реализует следующее преобразование открытого текста: каждая буква открытого текста заменяется третьей после нее буквой

в алфавите, который считается написанным по кругу, т. е. на примере русского алфавита за буквой «Я» следует снова буква «А». Отметим, что Цезарь заменял букву третьей после нее буквой, но в целом можно заменять и какой-нибудь другой буквой алфавита. Главное, чтобы тот, кому посылается шифрованное сообщение, знал эту величину сдвига. Итак, чтобы запомнить ключ, надо знать первую букву второй строки табл. 2.3.1. Однако такой шифр обладает большим недостатком: число различных ключей равно числу букв в алфавите. Перебрав эти варианты, можно однозначно восстановить отправленное сообщение.

Пример 2.3.1. Зашифровывание фразы на латинском языке осуществляется в два этапа. На первом этапе каждая буква текста заменяется на следующую в алфавитном порядке («Z» заменяется на «A»). На втором этапе применяется шифр простой замены с неизвестным ключом. Его применение заключается в замене каждой буквы шифрованного текста буквой того же алфавита, при этом разные буквы заменяются разными буквами. Ключом такого шифра является таблица, в которой указано, какой буквой надо заменить каждую букву алфавита. По данному шифртексту

OSZJX FXRF YOQJSZ RAYFJ

требуется восстановить отправленное сообщение, если известно, что для использованного (неизвестного) ключа результат шифрования не зависит от порядка выполнения указанных этапов при любом отправленном сообщении. Пробелы разделяют слова, при зашифровывании пробел остается пробелом. Известно также, что в результате зашифровывания «A» $\mapsto$ «F».

Занумеруем буквы латинского алфавита от 0 до 23, как указано ниже, в табл. 2.3.2.

Таблица 2.3.2

Латинский алфавит																							
A	B	C	D	E	F	G	H	I	J	L	M	N	O	P	Q	R	S	T	U	V	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Порядковые номера букв																							

Пусть x – некоторое число от 0 до 23, $f(x)$ – число, в которое переходит x на втором этапе. Тогда перестановочность этапов можно записать в виде $f(x+1)=f(x)+1$, т. е. $f(x+1)-f(x)=1$, значит, соседние числа x и $x+1$ на втором этапе переходят в соседние числа $f(x)$ и $f(x)+1$, отсюда следует, что второй этап – тоже циклический сдвиг. Последовательное применение $f(x)+1$ двух сдвигов – сдвиг. Итак, мы имеем классический шифр Цезаря. Остается рассмотреть 24 варианта различных сдвигов. Но поскольку в условии указано, что в результате зашифровывания «A» $\mapsto$ «F», то получаем, что зашифровывание представляет собой циклический сдвиг на 5 позиций вправо. Осложнения, связанные с переходом «Z» в «A», устраняются либо переходом к остаткам при делении на 24, либо выписыванием после буквы «Z» второй раз алфавита A B … Z, т. е. операции выполняются в $\mathbf{Z}/24\mathbf{Z}$. Итак, для расшифровывания

фразы нужно каждую букву полученного сообщения сдвинуть циклически на 5 позиций влево, а пробелы оставить на месте.

Итак, получаем следующее исходное сообщение:

INTER ARMA SILENT MUSAE

(и́нтэр а́рма си́лент му́зэ – «когда гремит оружие, музы молчат»).•

Все шифры замены имеют серьезный недостаток – с использованием одного ключа нельзя зашифровать достаточно длинные сообщения. Поэтому, как правило, шифры замены применяются в комбинации с другими шифрами, чаще всего – с шифрами перестановки.

Шифры перестановки

Шифры перестановки изменяют только порядок следования символов текста, но не изменяют их самих. Поэтому для расшифровывания нужно знать подстановку – биекцию множества символов сообщения, задающую преобразование. Всего существует $n!$ подстановок на n-элементном множестве. С увеличением числа n значение $n!$ растет очень быстро. При больших n для приближенного вычисления $n!$ можно пользоваться следующей известной формулой, названной в честь шотландского математика Джеймса Стирлинга (1692–1770):

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \text{ где } e = 2{,}718281828\ldots$$

Широкое распространение получили шифры перестановки, использующие некоторую геометрическую фигуру. Рассмотрим некоторые из них подробнее.

1. Одним из самых первых шифровальных приспособлений был жезл («сцитала»), применявшийся еще во времена войны Спарты против Афин в V в. до н. э. и связанный с именем спартанского полководца Лисандра. Жезл имел форму цилиндра, на который виток к витку наматывалась узкая папирусная лента (без просветов и нахлестов), а затем на этой ленте вдоль его оси записывался необходимый для передачи текст. Лента разматывалась, и получалось (для непосвященных), что поперек нее в беспорядке написаны какие-то буквы. Затем лента отправлялась адресату, который, имея цилиндр точно такого же диаметра, наматывал ленту на него и прочитывал сообщение вдоль оси. Ясно, что такой способ шифрования осуществляет перестановку местами букв сообщения.

Шифр «сцитала» реализует не более n перестановок, где n – длина сообщения. Действительно, этот шифр, как нетрудно видеть, эквивалентен следующему шифру маршрутной перестановки: в таблицу, состоящую из m столбцов, построчно записывают сообщение, после чего выписывают буквы по столбцам. Число задействованных столбцов таблицы не может превосходить длины сообщения. Имеются еще и чисто физические ограничения, накладываемые реализацией шифра «сцитала». Естественно предположить, что диаметр жезла не должен превосходить 10 см. Тогда при высоте строки в 1 см на одном витке такого жезла уместится не более 32 букв ($10\pi < 32$). Таким образом, число перестановок, реализуемых «сциталой», вряд ли превосходит 32.

2. Шифр «поворотная решетка». Идея использования поворотной решетки как средства шифрования принадлежит итальянскому математику и философу Джероламо Кардано (1501 или 1506 – 1576). В его честь данный способ шифрования также носит название «сетки (решетки) Кардано». Для использования шифра под названием «поворотная решетка» изготавливается трафарет из прямоугольного листа клетчатой бумаги размером $2m \times 2k$ клеток. В трафарете вырезано mk клеток так, что при наложении его на чистый лист бумаги того же размера четырьмя возможными способами его вырезы полностью покрывают всю площадь листа. Буквы сообщения последовательно вписываются в вырезы трафарета (по строкам, в каждой строке слева направо) при каждом из четырех его возможных положений в заранее установленном порядке. Получатель сообщения, имеющий точно такую же решетку, без труда прочтет исходный текст, наложив решетку на шифртекст по порядку четырьмя способами.

Можно доказать, что число возможных трафаретов, т. е. количество ключей шифра «поворотная решетка», составляет $T = 4^{mk}$. Этот шифр предназначен для сообщений длиной $n = 4mk$. Число всех перестановок в тексте такой длины составит $4mk!$, что во много раз больше числа T. Однако уже при размере трафарета 8×8 число возможных решеток превосходит 4 млрд.

3. Широко распространена разновидность шифра маршрутной перестановки, называемая «вертикальной перестановкой». Здесь снова используется прямоугольник, в который сообщение вписывается обычным способом (по строкам слева направо). Для получения горизонтальной криптограммы по строкам слева направо выписываются буквы, записанные по вертикали в столбцах, а столбцы при этом берутся в порядке, определяемом ключом.

Число ключей шифра «вертикальная перестановка» не более $m!$, где m – число столбцов таблицы. Как правило, m гораздо меньше, чем длина текста n (сообщение укладывается в несколько строк по m букв), а значит, и $m!$ много меньше $n!$. В случае когда ключ «вертикальной перестановки» не рекомендуется записывать, его можно извлекать из какого-то легко запоминающегося слова или предложения, содержащего m букв. Наиболее распространенный способ состоит в том, чтобы приписывать буквам числа в соответствии с обычным алфавитным порядком букв. Например, пусть ключевым словом будет «перестановка». Присутствующая в нем буква «А» получает номер 1. Если какая-то буква входит в ключевое слово несколько раз, то ее появления нумеруются последовательно слева направо. Поэтому второе вхождение буквы «А» получает номер 2. Поскольку буквы «Б» в этом слове нет, то буква «В» получает номер 3 и т. д. Процесс продолжается до тех пор, пока все буквы не получат номера. Таким образом, мы получаем ключ в соответствии с табл. 2.3.3.

Таблица 2.3.3

Исходный номер столбца											
1	2	3	4	5	6	7	8	9	10	11	12
П	Е	Р	Е	С	Т	А	Н	О	В	К	А
9	4	10	5	11	12	1	7	8	3	6	2
Номер столбца после перестановки											

Данный ключ определяет перестановку столбцов: $1 \mapsto 9,\ 2 \mapsto 4,\ 3 \mapsto 10,\ 4 \mapsto 5,\ 5 \mapsto 11,\ 6 \mapsto 12,\ 7 \mapsto 1,\ 8 \mapsto 7,\ 9 \mapsto 8,\ 10 \mapsto 3,\ 11 \mapsto 6,\ 12 \mapsto 2$. Получатель сообщения должен выполнить обратную перестановку строк: $1 \mapsto 7,\ 2 \mapsto 12,\ 3 \mapsto 10,\ 4 \mapsto 2,\ 5 \mapsto 4,\ 6 \mapsto 11,\ 7 \mapsto 8,\ 8 \mapsto 9,\ 9 \mapsto 1,\ 10 \mapsto 3,\ 11 \mapsto 5,\ 12 \mapsto 6$, затем прочитать текст по столбцам сверху вниз в порядке их следования или транспонировать матрицу и прочитать текст по строкам слева направо в порядке их следования.

Многоалфавитные шифры замены

В европейских странах многоалфавитные шифры были изобретены в эпоху Возрождения, когда развитие торговли потребовало надежных способов защиты информации. Шифр многоалфавитной замены впервые был предложен Леоном Батиста Альберти (1404–1472), итальянским ученым, архитектором, писателем и музыкантом, который в 1466 г. представил трактат о шифрах, где рассматривались различные способы шифрования, в том числе маскировка открытого текста в некотором вспомогательном тексте. Альберти изобрел, по его собственным словам, «шифр, достойный королей», – многоалфавитный шифр, реализованный в виде шифровального диска. Многоалфавитная замена может быть описана таблицей шифрования. Суть алгоритма шифрования заключается в том, что здесь используется несколько замен в соответствии с ключом. Позднее Альберти изобрел код с перешифровкой. Его изобретение значительно опередило свое время, поскольку такой тип шифра стал широко применяться в странах Европы только спустя 400 лет.

В 1518 г. в Германии появляется первая печатная книга по криптографии «Полиграфия». Ее автор, аббат Иоганнес Тритемий (1462–1516), развивает идею Альберти о многоалфавитной замене. Следующим этапом развития криптографии можно считать 1563 г., когда в своей книге «О тайной переписке» итальянский естествоиспытатель, основатель одного из первых в мире научных обществ «Секретная академия» Джованни Батиста Порта (1535–1615) описал *биграммный шифр*, в котором по определенной схеме осуществляется замена не одной буквы, а пары букв одним знаком. В своем исследовании по криптологии Порта впервые произвел систематизацию всех известных в то время шифров.

Еще один известный криптолог XVI в., французский дипломат Блез де Виженер (1523–1596), работавший в Риме, познакомился с трудами Альберти, Тритемия и Порта, детально проверил их идеи и создал на их основе новый шифр полиалфавитной замены с так называемым подвижным ключом. Хотя и Альберти, и Тритемий, и Порта внесли значительный вклад в изобретение дан-

ного шифра, он известен как *шифр Виженера*, в честь ученого, который придал ему окончательный вид.

Дополним естественный порядок букв в алфавите. Будем считать, что за последней буквой алфавита следует его первая буква. Расположим буквы на окружности в естественном порядке по часовой стрелке. Тогда значения относительных порядковых номеров (относительно фиксированной буквы) букв алфавита из n элементов совпадают со значениями всевозможных остатков от деления целых чисел на натуральное число n.

Определение 2.3.5. Число $D(N_1, N_2)$, равное порядковому номеру буквы с естественным номером N_1 относительно буквы с порядковым номером N_2 в алфавите, называется *знаком гаммы*.

Обозначим $r_m(N)$ остаток от деления целого N на $m \in \mathbf{N}$. Таким образом, $D(N_1, N_2) = r_n(N_1 - N_2)$, где n – число букв в алфавите.

Введем следующие обозначения:

$$N_1 [-] N_2 = r_n(N_1 - N_2),$$
$$N_1 [+] N_2 = r_n(N_1 + N_2).$$

Тогда

$$D(N_1, N_2) = N_1 [-] N_2,$$
$$N_1 = N_2 [+] D(N_1, N_2).$$

Для рассматриваемого шифра характерно то, что буквы открытого текста, зашифрованные одним и тем же знаком гаммы, по сути, зашифрованы одним и тем же шифром простой замены. Например, *ключевая таблица* этого шифра простой замены при знаке гаммы, равном 1, для русского алфавита имеет вид табл. 2.3.4.

Таблица 2.3.4

А	Б	В	Г	Д	…	Ь	Э	Ю	Я
Б	В	Г	Д	Е	…	Э	Ю	Я	А

Вторую строку ключевой таблицы называют *алфавитом шифрования, соответствующим данному знаку гаммы*. Поскольку в рассматриваемом шифре возможны все значения гаммы от 0 до 29 (русский алфавит без «Ё», «Й», «Ъ»), то данный шифр можно рассматривать как 30-алфавитный шифр замены. Если каждому из этих алфавитов поставить в соответствие его первую букву, то каждый знак гаммы можно заменить порядковым номером этой буквы в исходном алфавите. В этом случае ключ рассматриваемого шифра можно взаимно однозначно заменить соответствующим словом в этом же алфавите.

Такой многоалфавитный шифр замены был описан в 1585 г. Блезом де Виженером в его книге «Трактат о шифрах». Взяв за основу тайный алфавит Юлия Цезаря, Виженер построил по изобретенной Тритемием квадратной схеме четырехугольник, который впоследствии получил название *таблицы Виженера*. Все алфавиты шифрования относительно латинского алфавита были сведены Виженером в таблицу. Над таблицей располагался открытый текст, а сле-

ва от нее – ключ. Ниже приведена таблица Виженера для современного латинского алфавита – табл. 2.3.5, она составлена из списка 26 алфавитов шифрования, расположенных горизонтально. Каждый алфавит циклически сдвинут относительно находящегося над ним на одну букву влево.

Таблица 2.3.5

A	B	C	D	E	…	W	X	Y	Z
B	C	D	E	F	…	X	Y	Z	A
C	D	E	F	G	…	Y	Z	A	B
…	…	…	…	…	…	…	…	…	…
Z	A	B	C	D	…	V	W	X	Y

Способ зашифровывания с помощью таблиц Виженера заключается в том, что первый из алфавитов соответствует исходному алфавиту открытого текста, а букве ключевого слова соответствует алфавит шифрования из данного списка, начинающийся с этой буквы. Буква зашифрованного текста находится в алфавите шифрования на месте, соответствующем данной букве открытого текста. Рассмотренный ранее шифр Цезаря – частный случай, соответствующий ключевой последовательности, состоящей из одной буквы. Например,

УНИВЕРСИТЕТ – открытый текст,

ВЛТВЛТВЛТВЛ – периодическая ключевая последовательность,

ХЧЫДРВУУГЗЭ – зашифрованный текст.

Простота построения таблиц Виженера делает эту систему привлекательной для практического использования. Виженер использовал идею алгоритма многоалфавитной замены для создания шифрующего устройства, состоящего из вращающихся колес. Изобретенная им система кодов применяется в современных шифровальных машинах. Виженером было издано около 20 книг по криптологии. Среди них уже упоминавшийся «Трактат о шифрах», который содержит разделы, посвященные полиалфавитным системам, методам шифрования, а также созданию ключей. Еще одно важное усовершенствование многоалфавитных шифров замены, состоящее в идее использования в качестве ключа текста самого сообщения или же шифрованного текста, принадлежит Кардано и Виженеру. Такой шифр был назван *самоключом*.

Итак, первая идея вскрытия зашифрованного текста при таком методе шифрования состоит в использовании вероятного слова, т. е. слова, которое с большой вероятностью может содержаться в данном открытом тексте или в ключевой последовательности. Речь идет в том числе и о словах, часто встречающихся в любых открытых текстах, к ним, например, относятся слова «который», «тогда», «что», «если», приставки и предлоги «при», «пре», «под» и т. п. В вышеупомянутой книге «О тайной переписке» Порта привел примеры списков вероятных слов из различных областей знания, существенно предвосхитив то, что впоследствии криптологами было названо *методом вероятного слова*.

Вторая идея основана на том, что буквы открытого сообщения находятся в тексте на вполне определенных позициях. Если разность номеров их позиций

окажется кратной периоду ключевой последовательности, то стоящие на этих позициях буквы будут зашифрованы одним и тем же знаком гаммы. Это означает, что определенные части открытого текста окажутся зашифрованными шифром Цезаря. Эту идею можно использовать для определения периода ключа многоалфавитного шифра замены.

На протяжении почти трехсот лет шифр Виженера считался практически невскрываемым. Впервые метод разгадки шифра Виженера предложил в 1863 г. в своей книге «Тайнопись и искусство дешифровки» немецкий криптоаналитик и археолог, отставной офицер прусской армии Фридрих Вильгельм Казиски (1805–1881). С тех пор этот алгоритм вскрытия многоалфавитных шифров с ключом неизвестного периода известен как *тест Казиски*. Данный тест, по сути, заключается в том, что в зашифрованном тексте надо найти пары одинаковых отрезков из не менее чем трех символов, вычислить разности номеров позиций их начал и определить общие делители найденных разностей. Как правило, один из этих общих делителей равен периоду ключевой последовательности. Если нам известна длина ключевой фразы, то можно разбить текст на несколько фрагментов, для каждого из которых применяется шифр Цезаря. К каждому из фрагментов затем можно применить метод частотного анализа, как в случае обычных шифров однобуквенной замены. Таким образом, разгадывание шифра Виженера сводится на 90 % к нахождению длины ключевой последовательности. Конечно, этот метод применим, только если текст достаточно большой, а ключевая фраза значительно короче его.

В 1920 г. известный американский криптолог и криптоаналитик, один из основоположников современной криптографии, организатор и первый директор американской службы сигнальной разведки Уильям Фредерик Фридман (1891–1969) предложил другой, более простой и в то же время более продуктивный метод определения длины ключевой фразы. Этот метод является одним из самых ярких достижений классической криптографии. Метод Фридмана основан на подсчете так называемого *индекса совпадения*. Зашифрованное сообщение переписывается с циклическим сдвигом текста на несколько позиций. Далее полученные таким образом сообщения переписываются одно над другим и подсчитывается число совпавших букв в верхней и нижней строках. Вычисляется индекс совпадения, равный отношению числа совпадений к длине сообщения. Для текстов на русском языке индекс совпадения составляет в среднем примерно 6 %. Но для абсолютно случайной последовательности русских букв индекс совпадения значительно ниже! Поскольку всего в русском алфавите 32 буквы (русский алфавит без «Ё»), то вероятность совпадения составляет 1/32, или примерно 0,031, – чуть больше 3 %. На этом соображении и основан метод Фридмана. Зашифрованный текст записывается со сдвигами 2, 3, 4 и т. д., и для каждого сдвига вычисляется индекс совпадения. Если индекс совпадения колеблется между 0 и 5 %, то, скорее всего, размер сдвига не соответствует длине ключевой последовательности. Длина ключевой последовательности найдена, как только индекс совпадения скачком возрастет до 6 % и более.

Понять, почему индекс совпадения резко возрастает, когда величина сдвига становится кратной длине ключевой фразы, совсем не трудно. В случае любого шифра простой замены, в частности, шифра Цезаря, индекс совпадения для зашифрованного текста такой же, как и для исходного текста. При сдвиге текста, зашифрованного методом Виженера, на длину ключевой фразы мы сравниваем между собой буквы, которые шифруются одним и тем же шифром Цезаря. Следовательно, индекс совпадения получится в точности таким же, как и у незашифрованного текста при сдвиге на длину ключевой последовательности.

В заключение отметим, что теоретическое открытие, оказавшее серьезное влияние на развитие криптографии, было сделано в работах известного американского математика и инженера, одного из создателей математической теории информации Клода Элвуда Шеннона (1916–2001) «Теория связи в секретных системах» (1949 г.) и выдающегося советского и российского ученого в области радиотехники, радиосвязи и радиоастрономии Владимира Александровича Котельникова (1908–2005) «Основные положения автоматической шифровки» (1941 г.). В данных работах были сформулированы и доказаны необходимые и достаточные условия недешифруемости системы шифра. Они заключаются в том, что получение противником шифртекста не изменяет вероятностей используемых ключей. При этом было установлено, что единственным недешифруемым шифром является так называемая *лента одноразового использования*, когда открытый текст шифруется с помощью случайного ключа такой же длины. Однако это делает абсолютно стойкий шифр очень дорогим в эксплуатации.

Глава 3. Элементы теории групп

§3.1. Понятие алгебраической системы

Определение 3.1.1. Пусть X – непустое множество. *Бинарной алгебраической операцией* на множестве X называется всякое правило f, по которому каждой упорядоченной паре (x, y) элементов $x, y \in X$ ставится в соответствие один вполне определенный элемент z из X. Таким образом, функция $f: X^2 \to X$, $f : (x, y) \mapsto z$.

Аналогично можно определить n-арную алгебраическую операцию для $\forall\, n \in \mathbf{N}$. В дальнейшем будем рассматривать только бинарные алгебраические операции. Обычно для обозначения операций используются знаки $*$, $\times$, $\bullet$, $\circ$, $+$ и т. п. Воспользуемся первым из обозначений, тогда в определении 3.1.1 $z = x * y$.

Определение 3.1.2. Если на множестве X задана одна или несколько алгебраических операций, то говорят, что X есть *алгебраическая система с данными операциями*.

Пример 3.1.1. $(\mathbf{N}, +)$, $(\mathbf{Z}, +, \cdot)$, $(\mathbf{Q}, +, \cdot)$, $(\mathbf{R}, +, \cdot)$, $(\mathbf{Z}/n\mathbf{Z}, +, \cdot)$, $\forall\, n \in \mathbf{N}$, – алгебраические системы с бинарными операциями сложения и умножения.●

Алгебраические системы различают по количеству и свойствам операций.

Определение 3.1.3. Алгебраическая система $(X, *)$ на множестве X с одной алгебраической операцией $*$ называется *группоидом*. Если у группоида $(X, *)$ операция $*$ ассоциативна: $a * (b * c) = (a * b) * c$ для любых $a, b, c \in X$, то такую алгебраическую систему называют *полугруппой*. *Моноидом* $(X, *)$ называют полугруппу с *единицей*, или *нейтральным элементом*, т. е. таким элементом $e \in X$, что $e * x = x * e = x$ для каждого $x \in X$.

Из определения нейтрального элемента следует его единственность в моноиде: если e_1 и e_2 – нейтральные элементы, то $e_1 = e_1 * e_2 = e_2$.

Пример 3.1.2.

1. $(\mathbf{Z}, -)$ – группоид, но не полугруппа, поскольку операция вычитания неассоциативна: $(5 - 2) - 1 = 3 - 1 = 2$, а $5 - (2 - 1) = 5 - 1 = 4$.

2. $(\mathbf{N}, +)$ – полугруппа, но не моноид, поскольку в $\mathbf{N}$ нет нейтрального элемента относительно сложения.

3. $(\mathbf{Z}, \cdot)$, $(\mathbf{Z}/n\mathbf{Z}, \cdot)$, $\forall\, n \in \mathbf{N}$, – моноиды.●

Определение 3.1.4. Моноид $(X, *)$, у которого каждый элемент обратим, т. е. для всякого $x \in X$ существует *обратный элемент* $y \in X$, такой, что $x * y = y * x = e$ (тогда пишут $y = x^{-1}$), называется *группой*.

Из ассоциативности операции $*$ и определения 3.1.4 следует, что обратный элемент в группе $(X, *)$ единственный для каждого $x \in X$: пусть y_1 и y_2 – обратные элементы для x, тогда

$$y_1 = y_1 * e = y_1 * (x * y_2) = (y_1 * x) * y_2 = e * y_2 = y_2 .$$

Пример 3.1.3. Примерами групп являются моноиды: $(\mathbf{Z}, +)$, $(\mathbf{Q}^*, \cdot)$, $(\mathbf{R}^*, \cdot)$, $(\mathbf{C}^*, \cdot)$, $(\{-1, 1\}, \cdot)$, $(\mathbf{Z}/n\mathbf{Z}, +)$, $(\mathbf{Z}/n\mathbf{Z}^*, \cdot)$ для $\forall\, n \in \mathbf{N}$. Здесь и в дальнейшем A^* обозначает множество всех обратимых элементов относительно некоторой бинарной алгебраической операции, заданной на множестве A. В частности, $\mathbf{Q}^* = \mathbf{Q}\backslash\{0\}$, $\mathbf{R}^* = \mathbf{R}\backslash\{0\}$, $\mathbf{C}^* = \mathbf{C}\backslash\{0\}$, $\mathbf{Z}/n\mathbf{Z}^* = \{\overline{k} \in \mathbf{Z}/n\mathbf{Z} \mid \text{НОД}(k, n) = 1\}$, $\mathbf{Z}/p\mathbf{Z}^* = (\mathbf{Z}/p\mathbf{Z}) \setminus \{\overline{0}\}$ для любого простого числа p.•

В следующей главе будут рассмотрены кольца, поля, тела – алгебраические системы с двумя бинарными алгебраическими операциями. В данной главе остановимся на рассмотрении наиболее популярных алгебраических систем с одной операцией – группах.

§3.2. Группы, их основные типы и свойства

Дадим независимое определение группы.

Определение 3.2.1. *Группой* называется непустое множество G с определенной на нем бинарной алгебраической операцией $\bullet$, которая обладает свойствами:

1) ассоциативностью – $(a\bullet b)\bullet c = a\bullet(b\bullet c)$ для любых $a, b, c \in G$;

2) существованием нейтрального элемента, т. е. такого элемента $e \in G$, что $g\bullet e = e\bullet g = g$ для каждого $g \in G$;

3) наличием для каждого элемента $g \in G$ обратного, т. е. такого элемента $h \in G$, что $g\bullet h = h\bullet g = e$.

В любой группе нейтральный элемент и обратный для каждого элемента единственны в силу их определений и ассоциативности операции. Знак групповой операции $\bullet$, как и знак умножения, в записи можно опускать.

Определение 3.2.2. *Абелевой*, по фамилии норвежского математика Нильса Абеля (1802–1829), или *коммутативной* называется группа $(G, \bullet)$ со свойством $a\bullet b = b\bullet a$ для произвольных $a, b \in G$. В противном случае группа называется *неабелевой* или *некоммутативной*.

Определение 3.2.3. Группа относительно операции сложения называется *аддитивной группой*. Все аддитивные группы являются абелевыми. Нейтральный элемент аддитивной группы называют *нулем* и обозначают символом 0, а обратный элемент для элемента a – *противоположным* и обозначают $-a$.

Пример 3.2.1. Аддитивными являются группы $(\mathbf{Z}, +)$, $(\mathbf{Q}, +)$, $(\mathbf{R}, +)$, $(\mathbf{C}, +)$, $(\mathbf{Z}/n\mathbf{Z}, +)$ для $\forall\, n \in \mathbf{N}$.•

Пример 3.2.2. Пусть K – одно из множеств: $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ или $\mathbf{Z}/k\mathbf{Z}$ для $\forall\, k \in \mathbf{N}$. К аддитивным относятся группы $(M_{m\times n}(K), +)$ – множество прямоугольных матриц порядка $m \times n$ ($m, n \in \mathbf{N}$ – произвольные фиксированные числа) с коэффициентами из множества K относительно операции сложения матриц и $(V_n(K), +)$ – множество n-мерных векторов с компонентами из множества

K относительно операции векторного сложения ($n \in \mathbf{N}$ – произвольное фиксированное число).•

Определение 3.2.4. Группа относительно операции умножения называется *мультипликативной группой*. Нейтральный элемент мультипликативной группы называют *единицей* и часто обозначают символом 1, а обратный элемент для элемента a обозначают a^{-1}.

Пример 3.2.3. Абелевыми мультипликативными группами являются $(\mathbf{Q}^*, \cdot)$, $(\mathbf{R}^*, \cdot)$, $(\mathbf{C}^*, \cdot)$, $(\{-1, 1\}, \cdot)$, $(\mathbf{Z}/n\mathbf{Z}^*, \cdot)$ для $\forall\, n \in \mathbf{N}$.•

Пример 3.2.4. Пусть K – одно из множеств: **Q**, **R** или **C**. *Полной линейной группой* $GL_n(K)$ (от англ. general linear group – «полная линейная группа») называется множество всех квадратных матриц порядка $n \in \mathbf{N}$ с коэффициентами из K и ненулевым определителем с операцией матричного умножения. $GL_n(K)$ является неабелевой мультипликативной группой при $n \geq 2$, т. к. произведение матриц некоммутативно в общем случае.•

Определение 3.2.5. Группа $(G, \bullet)$ называется *конечной*, если G – конечное множество, в противном случае – *бесконечной*. *Порядком* конечной группы $|G|$ называется мощность множества G.

Алгебраическая операция в конечной группе может быть задана таблицей Кэли.

Пример 3.2.5.

1. Конечными являются группы: $(\mathbf{Z}/n\mathbf{Z}, +)$ порядка n, $(\mathbf{Z}/n\mathbf{Z}^*, \cdot)$ порядка 1 при $n = 1$ и порядка $\varphi(n)$ при $n \geq 2$, $(M_{m\times n}(\mathbf{Z}/k\mathbf{Z}), +)$ порядка k^{mn}, $(V_n(\mathbf{Z}/k\mathbf{Z}), +)$ порядка k^n.

2. Бесконечные группы: $(K, +)$, $(M_{m\times n}(K), +)$, $(V_n(K), +)$, где K – одно из множеств: **Z**, **Q**, **R** или **C**, а также $(K^*, \cdot)$ и $GL_n(K)$, где K – одно из множеств: **Q**, **R** или **C**.•

Для каждого $n \in \mathbf{N}$ существуют абелевы аддитивная и мультипликативная группы порядка n. Примерами таких групп являются $(\mathbf{Z}/n\mathbf{Z}, +)$ и C_n – мультипликативная группа всех комплексных корней n-й степени из 1, т. е. чисел $\omega_k = e^{2k\pi i/n}$, $k = \overline{0, n-1}$, в частности, $C_2 = (\{-1, 1\}, \cdot)$.

§3.3. Подгруппы

Определение 3.3.1. *Подгруппа* группы $(G, \bullet)$ – это непустое подмножество H множества G, которое в свою очередь является группой относительно той же бинарной алгебраической операции. Этот факт обозначают так: $H \leq G$ или $H < G$, если $H \subset G$.

Очевидно, что $G \leq G$ для произвольной группы $(G, \bullet)$.

Пример 3.3.1. $(\mathbf{Z}, +) < (\mathbf{Q}, +) < (\mathbf{R}, +) < (\mathbf{C}, +)$, $(\mathbf{Q}^*, \cdot) < (\mathbf{R}^*, \cdot) < (\mathbf{C}^*, \cdot)$.•

Пример 3.3.2. $C_m < C_n < (\mathbf{C}^*, \cdot)$ для всех $m, n \in \mathbf{N}$, где $m < n$ и $m \mid n$.•

Пример 3.3.3. Обозначим $m\mathbf{Z} = \{mq \mid q \in \mathbf{Z}\}$ для фиксированного $m \in \mathbf{Z}$. Тогда $m\mathbf{Z} \le n\mathbf{Z} \Leftrightarrow n \mid m$. В частности, $\ldots < (2^k\mathbf{Z}, +) < \ldots < (4\mathbf{Z}, +) < (2\mathbf{Z}, +) < (\mathbf{Z}, +)$, $k \in \mathbf{Z}_{\ge 0}$.•

Теорема 3.3.1 (критерий подгруппы). Непустое подмножество H группы $(G, \bullet)$ является подгруппой тогда и только тогда, когда для произвольных $a, b \in H$ выполняется условие: $a \bullet b^{-1} \in H$.

▷ Необходимость. Пусть $H \le G$, тогда для $\forall\, a, c \in H$ $a \bullet c \in H$. Поскольку для $\forall\, b \in H$ $\exists\, b^{-1} \in H$, то для $\forall\, a, b \in H$ $a \bullet b^{-1} \in H$.

Достаточность. Для $\forall\, a, b \in H$ $a \bullet b^{-1} \in H$. Отсюда следует выполнение следующих свойств:

1) для $\forall\, a \in H$ $a \bullet a^{-1} = e \in H$ – нейтральный элемент группы $(G, \bullet)$ находится в H;

2) для $\forall\, a \in H$ $e \bullet a^{-1} = a^{-1} \in H$, т. е. для каждого элемента из H обратный ему элемент также находится во множестве H;

3) для $\forall\, a \in H$, $\forall\, b \in H$ $\exists\, b^{-1} \in H \Rightarrow a \bullet (b^{-1})^{-1} = a \bullet b \in H$ – определена бинарная алгебраическая операция на H, которая совпадает с операцией на G;

4) для $\forall\, a, b, c \in H$ $a \bullet (b \bullet c) = (a \bullet b) \bullet c$ – операция на H ассоциативна, т. к. она ассоциативна на G.

Выполнение свойств 1–4 равносильно тому, что $(H, \bullet)$ является группой согласно определению 3.2.1. Таким образом, $H \le G$. ◁

В силу теоремы 3.3.1 в любой группе $(G, \bullet)$ подмножество $\{e\}$, состоящее из одного нейтрального элемента e этой группы, является подгруппой.

Определение 3.3.2. Подгруппа H группы G называется *собственной* (или *нетривиальной*), если $H \ne G$ и $H \ne \{e\}$.

Упражнение 3.3.1. Проверить, что пересечение подгрупп произвольной группы – подгруппа этой группы.

Пример 3.3.4. Пусть K – одно из множеств: **Z**, **Q**, **R** или **C**. *Специальной линейной группой* $SL_n(K)$ (от англ. special linear group – «специальная линейная группа») называется множество всех квадратных матриц порядка $n \in \mathbf{N}$ с коэффициентами из K и определителем, равным 1, с операцией матричного умножения. С помощью критерия подгруппы легко убедиться в том, что при $K \ne \mathbf{Z}$ $SL_n(K) < GL_n(K)$, но $SL_n(\mathbf{Z}) < GL_n(\mathbf{Q})$. Действительно, для $\forall\, A, B \in SL_n(K)$ в силу свойств определителей

$$\det(A \cdot B^{-1}) = \det(A) \cdot (\det(B))^{-1} = 1 \cdot 1 = 1,$$

откуда следует, что $A \cdot B^{-1} \in SL_n(K)$.•

Теорема 3.3.2. Пусть a – любой фиксированный элемент произвольной группы $(G, \bullet)$, и пусть $H = \{a^0 = e, a, a^2, \ldots, a^{-1}, a^{-2}, \ldots\} = \{a^z \mid z \in \mathbf{Z}\}$ – множество всевозможных целых степеней элемента a. Тогда H – подгруппа группы G, причем абелева.

$\triangleright H \neq \varnothing$, т. к. $e \in H$, и $H \subseteq G$. Так как $a^k \bullet a^{-n} = a^{k-n} \in H$, поскольку для $\forall k, n \in \mathbf{Z}\ k - n \in \mathbf{Z}$, то по критерию подгруппы $H \leq G$. В силу того что $a^m \bullet a^n = a^n \bullet a^m = a^{m+n}$ для $\forall m, n \in \mathbf{Z}$, H – абелева подгруппа группы G.$\triangleleft$

Определение 3.3.3. Подгруппа H из теоремы 3.3.2 называется *циклической подгруппой*, *порожденной элементом* a, и обозначается $< a >$. Если найдется $b \in G$, такой, что $G = < b >$, то такую группу называют *циклической*, а b – ее *образующим элементом*.

Пример 3.3.5.

1. $(\mathbf{Z}, +) = < 1 > = < -1 >$.

2. $(\mathbf{Z}/n\mathbf{Z}, +) = < \overline{k} >$ и $C_n = < e^{2k\pi/n} >$ для всех $k = \overline{0, n-1}$ с условием $(k, n) = 1$, таким образом, количество образующих элементов равно 1 при $n = 1$ и $\varphi(n)$ при $n \geq 2$.$\bullet$

Определение 3.3.4. Пусть e – нейтральный элемент группы. Элемент группы a называется *элементом бесконечного порядка*, если для $\forall n \in \mathbf{N}\ a^n \neq e$. Элемент a называется *элементом конечного порядка* $n \in \mathbf{N}$, если $a^n = e$, но $a^k \neq e$ для $\forall k \in \mathbf{N}_{<n}$. Будем обозначать порядок элемента a $\mathrm{ord}(a)$ (от англ. order – «порядок»).

Очевидно, что в любой группе $\mathrm{ord}(e) = 1$.

Пример 3.3.6.

1. Для $\forall z \in \mathbf{Z}\backslash\{0\}$ $\mathrm{ord}(z) = \infty$ в $(\mathbf{Z}, +)$, поскольку при $\forall n \in \mathbf{N}\ n \cdot z \neq 0$.

2. Для всех $k = \overline{0, n-1}$ $\mathrm{ord}(\overline{k}) = \mathrm{ord}(e^{2k\pi i/n}) = n/(k, n)$ в $(\mathbf{Z}/n\mathbf{Z}, +)$ и в C_n.$\bullet$

Теорема 3.3.3. Пусть элемент $a \in G$ обладает свойством $\mathrm{ord}(a) = n$ для некоторого $n \in \mathbf{N}$. Тогда $< a > = \{a, a^2, ..., a^n = e\}$ – циклическая подгруппа группы G, имеющая порядок n.

$\triangleright$ Пусть $H = \{a, a^2, ..., a^n = e\}$, тогда $|H| = n$, т. к. $\mathrm{ord}(a) = n$.

$< a > = \{a^0 = e, a, a^2, ..., a^{-1}, a^{-2}, ...\} = \{a^z \mid z \in \mathbf{Z}\}$. Очевидно, $H \subseteq < a >$.

Рассмотрим $\forall z \in \mathbf{Z}$. Если $0 \leq z < n$, то $a^z \in H$. Если $z \geq n$ или $z < 0$, то по теореме 1.1.1 существуют такие $q, r \in \mathbf{Z}$, где $0 \leq r < n$, что $a^z = a^{nq+r} = a^r \in H$, поскольку $a^{nq} = e$. Значит, $< a > \subseteq H$.

Итак, $H = < a >$ – циклическая подгруппа, порожденная элементом a.$\triangleleft$

Теорема 3.3.4. Всякая подгруппа H циклической группы G является циклической.

$\triangleright$ Пусть $G = < a >$.

1. $H = \{e\} \Rightarrow H = < e >$.

2. $H \neq \{e\}$. Для $\forall b \in H\ \exists m \in \mathbf{Z}$, такой, что $b = a^m$, т. к. $H \leq G$ и $G = < a >$. Поскольку $H \leq G$, $a^{-m} \in H$. Таким образом, в H содержатся некоторые степени элемента a с натуральными показателями. Из всех таких показателей выберем наименьший. Пусть это будет $d \in \mathbf{N}$. Докажем, что $H = < a^d >$.

$< a^d > \subseteq H$, т. к. $a^d \in H$, и поскольку $H \leq G$, то $a^{dz} \in H$ для $\forall z \in \mathbf{Z}$.

Рассмотрим теперь $\forall\, h \in H.\ h = a^s = a^{dq+r}$, где $q, r \in \mathbf{Z},\ 0 \le r < d$, согласно теореме 1.1.1, поскольку $s \in \mathbf{Z}$. Тогда $a^r = a^{-dq} \bullet h \in H$, т. к. $a^{-dq}, h \in H$. Если $r \neq 0$, то $r < d$ и степень элемента a с натуральным показателем, меньшим, чем d, принадлежит H, что приводит к противоречию. Значит, $r = 0$ и для некоторого $q \in \mathbf{Z}\ \ h = a^{dq} \in < a^d >$. Поэтому $H \subseteq < a^d >$.

Итак, $H = < a^d >$ – циклическая подгруппа группы G. ◁

Обозначим E_n единичную квадратную матрицу порядка $n \in \mathbf{N}$.

Пример 3.3.7. Рассмотрим матрицу $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL_2(\mathbf{Z})$. Здесь $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$,…, таким образом, $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \neq E_2$ при $\forall\, n \in \mathbf{N}$. Степени матрицы A попарно различны и образуют бесконечную последовательность. Следовательно, $\mathrm{ord}(A) = \infty$ и циклическая подгруппа $< A > = \{A^z \mid z \in \mathbf{Z}\}$, порожденная матрицей A, в группе $SL_2(\mathbf{Z})$ является бесконечной.

$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $A^{-2} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$,…, $A^{-n} = \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix}$ для $\forall\, n \in \mathbf{N}$. Таким образом, $A^z = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}$ при $\forall\, z \in \mathbf{Z}$. ●

Пример 3.3.8. Матрица $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in SL_2(\mathbf{Z})$ имеет степени $B^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $B^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2$. Поэтому $\mathrm{ord}(B) = 4$ в группе $SL_2(\mathbf{Z})$. Согласно теореме 3.3.3 подгруппа $< B > = \{B, B^2, B^3, B^4 = E_2\}$ является конечной подгруппой порядка 4 в группе $SL_2(\mathbf{Z})$. ●

Отметим, что чаще группы не являются циклическими. Например, все некоммутативные группы не могут быть циклическими. Циклическими не являются группы, заданные на несчетных множествах, например, $(\mathbf{R}, +)$, $(\mathbf{C}, +)$, $(\mathbf{R}^*, \cdot)$ и $(\mathbf{C}^*, \cdot)$.

Пример 3.3.9. Несмотря на то, что $\mathbf{Q}^* \leftrightarrow \mathbf{Q} \leftrightarrow \mathbf{Z}$, абелевы группы $(\mathbf{Q}, +)$ и $(\mathbf{Q}^*, \cdot)$ также нециклические. Ведь для каждого фиксированного $q = m/n \in \mathbf{Q}^*$, где $m \in \mathbf{Z},\ n \in \mathbf{N},\ (m, n) = 1$, подгруппа $< q > = \{0, \pm m/n, \pm 2m/n, \pm 3m/n, \ldots\}$ не содержит рациональных несократимых дробей r/s при $r \in \mathbf{Z},\ s \in \mathbf{N}_{>n}$, следовательно, $< q > \neq \mathbf{Q}$. А подгруппа $< q > = \{1, (m/n)^{\pm 1}, (m/n)^{\pm 2}, (m/n)^{\pm 3}, \ldots\}$ не содержит $r/s \in \mathbf{Q}^*$, где $r \in \mathbf{Z},\ s \in \mathbf{N}_{>1},\ (r, s) = 1$ и $(s, n) = (s, m) = 1$, следовательно, $< q > \neq \mathbf{Q}^*$. ●

§3.4. Смежные классы. Теорема Лагранжа. Нормальные подгруппы

Определение 3.4.1. Пусть H – подгруппа группы G и $a \in G$. Обозначим aH множество элементов $\{a \bullet h \mid h \in H\}$ и назовем его *левым смежным классом* группы G по подгруппе H. Элемент a называется *представителем левого смежного класса* aH.

Если существует $b \in G$, $b \notin H \cup aH$, то можно построить левый смежный класс bH и т. д.

Аналогично строятся *правые смежные классы* $Ha = \{h \bullet a \mid h \in H\}$ с представителями $a \in G$.

Теорема 3.4.1. Пусть H – подгруппа группы G. Тогда справедливы утверждения:

1) каждый элемент $g \in G$ принадлежит какому-нибудь левому (правому) смежному классу по подгруппе H;

2) два элемента $a, b \in G$ принадлежат одному левому (правому) смежному классу тогда и только тогда, когда $a^{-1} \bullet b \in H$ $(b \bullet a^{-1} \in H)$;

3) любые два левых (правых) смежных класса либо не пересекаются, либо совпадают;

4) G есть объединение попарно непересекающихся левых (правых) смежных классов по подгруппе H.

$\triangleright$ **1)** для $\forall\, g \in G$ $g = g \bullet e$, т. к. $e \in H$, то $g \in \{g \bullet h \mid h \in H\} = gH$. Аналогично, $g = e \bullet g \Rightarrow g \in \{h \bullet g \mid h \in H\} = Hg$;

2) $a \in aH$, $b \in bH$, a и b принадлежат одному и тому же левому смежному классу $\Leftrightarrow$ ($\exists\, h \in H$, такой, что $b = a \bullet h \Leftrightarrow a^{-1} \bullet b = h$) $\Leftrightarrow a^{-1} \bullet b \in H$.

$a \in Ha$, $b \in Hb$, a и b принадлежат одному и тому же правому смежному классу $\Leftrightarrow$ ($\exists\, h \in H$, такой, что $b = h \bullet a \Leftrightarrow b \bullet a^{-1} = h$) $\Leftrightarrow b \bullet a^{-1} \in H$;

3) пусть $c \in aH \cap bH \Leftrightarrow$ для некоторых $h_1, h_2 \in H$ $c = a \bullet h_1 = b \bullet h_2 \Leftrightarrow b^{-1} \bullet a = h_2 \bullet h_1^{-1} \in H \Leftrightarrow a, b$ принадлежат одному и тому же левому смежному классу согласно утверждению 2, т. е. $aH = bH$.

Пусть $c \in Ha \cap Hb \Leftrightarrow c = h_1 \bullet a = h_2 \bullet b \Leftrightarrow a \bullet b^{-1} = h_1^{-1} \bullet h_2 \in H \Leftrightarrow Ha = Hb$.

Либо $aH \cap bH = \varnothing$ ($Ha \cap Hb = \varnothing$);

4) следует из утверждений 1 и 3 данной теоремы. $\triangleleft$

Определение 3.4.2. Мощность множества всех различных левых (правых) смежных классов группы G по подгруппе H называется *индексом* подгруппы H в группе G и обозначается $G : H$.

К важнейшим в теории групп относится следующая теорема, доказанная известным французским математиком и механиком Жозефом Лагранжем (1736–1813).

Теорема 3.4.2 (Ж. Лагранж). Порядок конечной группы равен произведению порядка и индекса любой ее подгруппы.

$\triangleright$ Пусть $H \leq G$, $|G| = n$, $|H| = k$, $n, k \in \mathbf{N}$, и $H = \{e = h_1, h_2, \ldots, h_k\}$.

Для $\forall a \in G$ $aH = \{a, a \bullet h_2, \ldots, a \bullet h_k\}$, $Ha = \{a, h_2 \bullet a, \ldots, h_k \bullet a\}$. Покажем, что $|H| = |aH| = |Ha|$. Действительно, $a \bullet h_i = a \bullet h_j \Leftrightarrow a^{-1} \bullet (a \bullet h_i) = a^{-1} \bullet (a \bullet h_j) \Leftrightarrow h_i = h_j$, аналогично, $h_i \bullet a = h_j \bullet a \Leftrightarrow h_i = h_j$ для всех $i, j = \overline{1,k}$. В общем случае, даже если подгруппа H бесконечна, $H \leftrightarrow aH \leftrightarrow Ha$.

Согласно утверждению 4 теоремы 3.4.1 G есть объединение попарно непересекающихся левых (правых) смежных классов по подгруппе H. Таким образом, $G = \bigcup\limits_{a \in G} aH = \bigcup\limits_{a \in G} Ha$ (объединение всех различных смежных классов) $\Rightarrow$ $\Rightarrow |G| = |H| \cdot (G : H)$ и $G : H = n/k$. $\triangleleft$

Следствие 1. Порядок конечной группы делится на порядок любой ее подгруппы.

Следствие 2. Если G – конечная группа порядка n, то порядок любого элемента группы делит порядок группы и для каждого $a \in G$ $a^n = e$.

$\triangleright$ Пусть $|G| = n$. Рассмотрим $< a >$ для $\forall a \in G$, очевидно, что $\text{ord}(a) \le n$. $|< a >| = \text{ord}(a)$ согласно теореме 3.3.3. $< a > \le G$, значит, по следствию 1 из теоремы 3.4.2 $\text{ord}(a) \,|\, n$. Отсюда следует, что $a^n = a^{\text{ord}(a) \cdot m} = \left(a^{\text{ord}(a)}\right)^m = e^m = e$, где $m \in \mathbf{N}$. $\triangleleft$

Следствие 3. Любая группа простого порядка является циклической и не содержит собственных подгрупп.

$\triangleright$ Пусть порядок группы G $|G| = p$ – простое число. Рассмотрим $\forall a \in$ $\in G \backslash \{e\}$. Согласно следствию 1 из теоремы 3.4.2 $|< a >| \,|\, p$, откуда $|< a >| = 1$ или $|< a >| = p$. Но в первом случае $< a > = < e >$ и $a = e$, что не так. Поэтому $|< a >| = p$, значит, $G = < a >$.

По следствию 1 из теоремы 3.4.2 порядок любой подгруппы H в G равен 1 или p, т. е. $H = \{e\}$ или $H = G$. $\triangleleft$

Так как для произвольной группы $(G, \bullet)$ и $\forall a \in G$, следуя доказательству теоремы 3.4.2, $G \leftrightarrow aG \leftrightarrow Ga$, то $aG = G = Ga$ и $G : G = 1$.

Пример 3.4.1. Рассмотрим $(n\mathbf{Z}, +) \le (\mathbf{Z}, +)$ для $\forall n \in \mathbf{N}$ и найдем все смежные классы $\mathbf{Z}$ по $n\mathbf{Z}$. Поскольку $(\mathbf{Z}, +)$ – абелева группа, то для одинаковых представителей левые смежные классы совпадают с соответствующими правыми. Таким образом, для $\forall z \in \mathbf{Z}$ $z = nq + r$, где $q, r \in \mathbf{Z}$, $0 \le r < n$, согласно теореме 1.1.1, и $z + n\mathbf{Z} = n\mathbf{Z} + z = \{nm + r \mid m \in \mathbf{Z}\} = \overline{r}$ в $\mathbf{Z}/n\mathbf{Z}$. В частности, $n\mathbf{Z} = \overline{0}$ в $\mathbf{Z}/n\mathbf{Z}$. Итак, $\{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\} = \mathbf{Z}/n\mathbf{Z}$ – множество всех различных левых (правых) смежных классов группы $(\mathbf{Z}, +)$ по подгруппе $(n\mathbf{Z}, +)$. Значит, $\mathbf{Z} : n\mathbf{Z} = n$.

Здесь $(\mathbf{Z}, +)$ – бесконечная группа, $(n\mathbf{Z}, +)$ – ее бесконечная подгруппа, а индекс $\mathbf{Z} : n\mathbf{Z}$ конечен. $\bullet$

Пример 3.4.2. $(V_4(\mathbf{Z}/2\mathbf{Z}), +)$ – аддитивная группа всех строк-векторов с четырьмя координатами из $\mathbf{Z}/2\mathbf{Z}$. Рассмотрим следующее подмножество в $V_4(\mathbf{Z}/2\mathbf{Z})$: $H = \{0 = (\overline{0}, \overline{0}, \overline{0}, \overline{0}), e_1 = (\overline{1}, \overline{0}, \overline{1}, \overline{1}), e_2 = (\overline{0}, \overline{1}, \overline{0}, \overline{1}), e_3 = (\overline{1}, \overline{1}, \overline{1}, \overline{0})\}$. Легко проверить, например, по критерию подгруппы (теорема 3.3.1), что $H < V_4(\mathbf{Z}/2\mathbf{Z})$. $|V_4(\mathbf{Z}/2\mathbf{Z})| = 2^4 = 16$, $|H| = 4$, значит, $V_4(\mathbf{Z}/2\mathbf{Z}) : H = 16 : 4 = 4$. Для

одинаковых представителей левые смежные классы совпадают с правыми, поскольку $(V_4(\mathbf{Z}/2\mathbf{Z}), +)$ – абелева группа.

Построим табл. 3.4.1 смежных классов группы $(V_4(\mathbf{Z}/2\mathbf{Z}), +)$ по подгруппе $(H, +)$.

Таблица 3.4.1

Смежные классы	Элементы смежных классов			
класс $a + H$	$a + 0$	$a + e_1$	$a + e_2$	$a + e_3$
$0 + H = H$	$(\bar{0}, \bar{0}, \bar{0}, \bar{0})$	$(\bar{1}, \bar{0}, \bar{1}, \bar{1})$	$(\bar{0}, \bar{1}, \bar{0}, \bar{1})$	$(\bar{1}, \bar{1}, \bar{1}, \bar{0})$
$(\bar{1}, \bar{0}, \bar{0}, \bar{0}) + H$	$(\bar{1}, \bar{0}, \bar{0}, \bar{0})$	$(\bar{0}, \bar{0}, \bar{1}, \bar{1})$	$(\bar{1}, \bar{1}, \bar{0}, \bar{1})$	$(\bar{0}, \bar{1}, \bar{1}, \bar{0})$
$(\bar{0}, \bar{1}, \bar{0}, \bar{0}) + H$	$(\bar{0}, \bar{1}, \bar{0}, \bar{0})$	$(\bar{1}, \bar{1}, \bar{1}, \bar{1})$	$(\bar{0}, \bar{0}, \bar{0}, \bar{1})$	$(\bar{1}, \bar{0}, \bar{1}, \bar{0})$
$(\bar{0}, \bar{0}, \bar{1}, \bar{0}) + H$	$(\bar{0}, \bar{0}, \bar{1}, \bar{0})$	$(\bar{1}, \bar{0}, \bar{0}, \bar{1})$	$(\bar{0}, \bar{1}, \bar{1}, \bar{1})$	$(\bar{1}, \bar{1}, \bar{0}, \bar{0})$

●

Таблицы смежных классов играют важную роль в теории и практике помехоустойчивого кодирования при коррекции и обнаружении ошибок (например, в телекоммуникационных, радиотехнических и вычислительных системах для синхронизации, передачи информации, локации и диагностики). Так, табл. 3.4.1 можно рассматривать в качестве *таблицы декодирования*, или *стандартного расположения*, для кода, слова которого образуют подгруппу H группы $(V_n(P), +)$, где P – конечное поле, $|P| = q \in \mathbf{N}$. В данном случае $n = 4$, P – $(\mathbf{Z}/2\mathbf{Z}, +, \cdot)$. Из соображений простоты и надежности большинство систем связи конструируется для передачи двоичных последовательностей: $q = 2$. При передаче кодовых слов по каналам связи на них в результате помех могут налагаться векторы ошибок из $(V_n(P), +)$, так что принятое сообщение представляется в виде суммы в $(V_n(P), +)$ кодового слова и вектора ошибок. При математическом анализе систем связи обычно рассматривается модель q-ичного симметричного канала, когда каждый символ принимается правильно с вероятностью p и ошибочно с вероятностью $(1-p)/(q-1)$ и ошибки в передаче последовательных символов происходят независимо. Тогда при построении таблиц декодирования в качестве представителей смежных классов группы по подгруппе берутся *лидеры* – представители смежных классов с минимальным *весом*, т. е. числом ненулевых компонент (при наличии нескольких таких векторов выбирается любой из них случайным образом). Метод декодирования принятого слова *по максимуму правдоподобия* заключается в отыскании данного слова в таблице и выборе в качестве переданного кодового слова в том же столбце и в первой строке, при этом вектором ошибок считается лидер смежного класса, находящийся в той же строке и в первом столбце. Данный метод сводится к выбору ближайшего к полученному вектору по *расстоянию*, т. е. весу разности, кодового слова в качестве результата декодирования. Такой выбор оптимален, поскольку он минимизирует вероятность ошибки при передаче по q-ичному симметричному каналу. При этом ошибка декодирования происходит тогда и только тогда, когда вектор ошибок не является лидером смежного класса, и незамеченными остаются в точности те векторы ошибок, которые сами являются ненулевыми кодовыми словами.

Определение 3.4.3. Подгруппа H группы G называется *нормальной*, если для всякого $a \in G$ $aH = Ha$, т. е. каждый левый смежный класс по подгруппе H совпадает с правым смежным классом с тем же представителем. В этом случае используется обозначение $H \triangleleft G$.

Ясно, что у абелевых групп все подгруппы нормальны.

В любой группе $(G, \bullet)$ тривиальные подгруппы $\{e\}$ и G являются нормальными, т. к. для $\forall\, a \in G\ a\{e\} = \{a\} = \{e\}a$ и $aG = G = Ga$. Итак, $\{e\} \lhd G$ и $G \lhd G$.

Пример 3.4.3. Если $G : H = 2$, то $H \lhd G$ для произвольной группы $(G, \bullet)$. Поскольку для $\forall\, g \in G \backslash H\ \ H \cap gH = H \cap Hg = \varnothing$ и $G = H \cup gH = H \cup Hg$ согласно утверждениям 3 и 4 теоремы 3.4.1, следовательно, $gH = Hg$ и, таким образом, $aH = Ha$ для $\forall\, a \in G$.•

Теорема 3.4.3 (критерий нормальной подгруппы). $H \lhd G$ тогда и только тогда, когда для каждого $a \in G\ \ aHa^{-1} = H$ (для $\forall\, a \in G, \forall\, h \in H\ a \bullet h \bullet a^{-1} \in H$).

$\triangleright$ Необходимость. Пусть $H \lhd G$, тогда для $\forall a \in G\ \ aH = Ha$. Значит, для $\forall\, h \in H\ \exists\, h_1 \in H$, такой, что $a \bullet h = h_1 \bullet a \Leftrightarrow a \bullet h \bullet a^{-1} = h_1 \in H$. Таким образом, для $\forall\, a \in G\ aHa^{-1} = \{a \bullet h \bullet a^{-1} \mid h \in H\} \subseteq H$. Поскольку для всех $h_i,\ h_j \in H$ справедливо $a \bullet h_i \bullet a^{-1} = a \bullet h_j \bullet a^{-1} \Leftrightarrow a^{-1} \bullet (a \bullet h_i \bullet a^{-1}) \bullet a = a^{-1} \bullet (a \bullet h_j \bullet a^{-1}) \bullet a \Leftrightarrow h_i = h_j$, то $|aHa^{-1}| = |H|$ и $aHa^{-1} = H$.

Достаточность. Пусть для $\forall\, a \in G,\ \forall\, h \in H\ a \bullet h \bullet a^{-1} \in H$ $(aHa^{-1} = H)$. Докажем, что $aH \subseteq Ha$ и $Ha \subseteq aH$. Для $\forall\, b \in aH\ \exists\, h_1 \in H$, такой, что $b = a \bullet h_1 \Leftrightarrow b \bullet a^{-1} = a \bullet h_1 \bullet a^{-1} = h_2 \in H \Rightarrow \exists\, h_2 \in H$, такой, что $b = h_2 \bullet a \in Ha$. Значит, $aH \subseteq Ha$. Для $\forall\, c \in Ha\ \exists\, h_3 \in H$, такой, что $c = h_3 \bullet a \Leftrightarrow a^{-1} \bullet c = a^{-1} \bullet h_3 \bullet a = h_4 \in H \Rightarrow \exists\, h_4 \in H$, такой, что $c = a \bullet h_4 \in aH$. Значит, $Ha \subseteq aH$. Итак, $aH \subseteq Ha$ и $Ha \subseteq aH$, следовательно, $aH = Ha$ для $\forall\, a \in G$ и $H \lhd G$. $\triangleleft$

Пример 3.4.4.

1. Если K – одно из множеств: $\mathbf{Q}$, $\mathbf{R}$ или $\mathbf{C}$, то $SL_n(K) \lhd GL_n(K)$ для $\forall\, n \in \mathbf{N}$, поскольку для произвольных матриц $A \in GL_n(K)$ и $B \in SL_n(K)\ A \cdot B \cdot A^{-1} \in GL_n(K)$ и в силу свойств определителей

$$\det(A \cdot B \cdot A^{-1}) = \det(A) \cdot \det(B) \cdot (\det(A))^{-1} = \det(B) = 1,$$

откуда следует, что $A \cdot B \cdot A^{-1} \in SL_n(K)$.

2. $SL_n(\mathbf{Z})$ не является нормальной подгруппой $GL_n(\mathbf{Q})$ при $n \geq 2$. Рассмотрим, например, матрицы $A = \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix} \in GL_2(\mathbf{Q})$ и $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \in SL_2(\mathbf{Z})$. Тогда $A^{-1} = \begin{pmatrix} -1/5 & 4/5 \\ 2/5 & -3/5 \end{pmatrix}$ и $A \cdot B \cdot A^{-1} = \begin{pmatrix} 17/5 & -18/5 \\ 8/5 & -7/5 \end{pmatrix} \notin SL_2(\mathbf{Z})$, но $A \cdot B \cdot A^{-1} \in SL_2(\mathbf{Q})$.•

§3.5. Симметрические группы

Определение 3.5.1. Пусть Ω – конечное множество из n элементов для произвольного $n \in \mathbf{N}$. Поскольку природа его элементов для нас несущественна, удобно считать, что $\Omega = \{1, 2, \ldots, n\}$. Всякая биекция, т. е. взаимно однозначное преобразование Ω, называется *подстановкой* на множестве Ω. В развернутой и наглядной форме подстановку $f : i \mapsto f(i),\ i = 1, 2, \ldots, n$, удобно изображать в виде двухстрочной таблицы: $f = \begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}$. В этой

таблице каждый i-й столбец четко указывает, в какой элемент $f(i)$ преобразуется элемент i, $1 \le i \le n$.

Композиция подстановок, как композиция биективных преобразований, является подстановкой. Подстановки перемножаются в соответствии с общим правилом композиции функций: $gf(i) = g(f(i))$, $i = 1, 2, \ldots, n$.

Пример 3.5.1. Пусть $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$. Тогда

$$fg = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{matrix} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 4 & 1 & 2 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 4 & 2 & 3 & 1 \end{matrix} \begin{matrix} g: \\ \\ \\ f: \\ \\ \end{matrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix},$$

аналогично, $gf = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$. Как видим, $fg \ne gf$, т. е. композиция подстановок не обладает свойством коммутативности.●

Очевидно, тождественная подстановка $e = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}$ играет роль нейтрального элемента относительно операции композиции подстановок. Как известно, композиция функций является ассоциативной операцией, поэтому и композиция подстановок ассоциативна. Каждая подстановка – обратимая функция. Чтобы найти для подстановки f обратную подстановку f^{-1}, достаточно в таблице $\begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}$ переставить строки местами, а затем столбцы упорядочить по возрастанию элементов первой строки.

Пример 3.5.2. Для подстановки $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ найдем обратную подстановку f^{-1}. Получаем, что $f^{-1} = \begin{pmatrix} 3 & 1 & 4 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$. Чтобы убедиться в правильности вычислений, найдем композиции этих двух подстановок в том и другом порядке:

$$f^{-1}f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = e;$$

$$ff^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = e.\bullet$$

Таким образом, все подстановки на множестве Ω образуют в общем случае некоммутативную группу с операцией композиции подстановок.

Определение 3.5.2. Группу всех подстановок на n-элементном множестве относительно операции композиции называют *симметрической группой степени n* и обозначают S_n.

Несложно видеть, что все вышесказанное справедливо не только для конечного, но и для произвольного непустого множества Ω. Таким образом, множество всех биективных преобразований Ω образует группу относительно операции композиции, обозначаемую $S(\Omega)$ и называемую *симметрической группой множества Ω.*

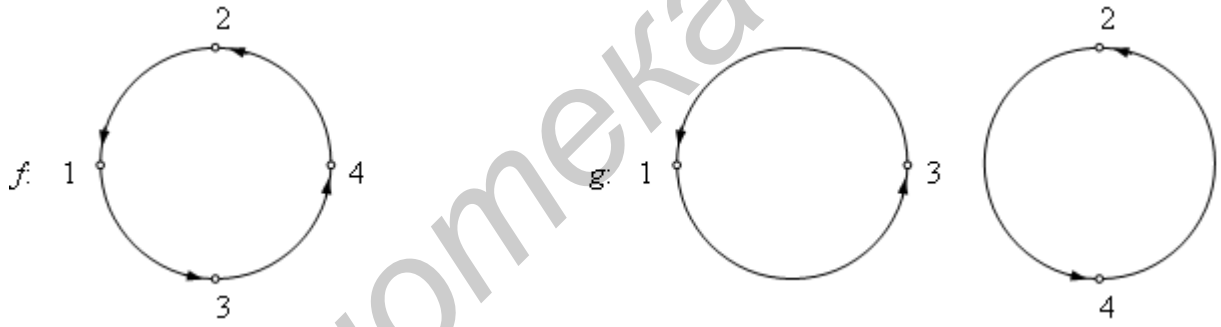
Теорема 3.5.1. Порядок группы S_n равен $n!$.

▷ Исходя из определения подстановки как взаимно однозначного преобразования n-элементного множества, утверждение справедливо согласно теореме 2.2.2. ◁

Таким образом, $|S_2| = 2$, $|S_3| = 6$, $|S_4| = 24$, $|S_5| = 120$ и т. д.

Разложим подстановки из S_n в произведение более простых подстановок. Сначала поясним суть разложения, рассмотрев подстановки f и g из примера 3.5.1.

Пример 3.5.3. Подстановка f кратко записывается в виде $f = (1\ 3\ 4\ 2)$ или, что то же самое, в виде $f = (3\ 4\ 2\ 1) = (4\ 2\ 1\ 3) = (2\ 1\ 3\ 4)$ и носит название *цикла длиной 4*. Подстановка $g = (1\ 3)(2\ 4)$ – произведение двух независимых (непересекающихся) циклов (1 3) и (2 4) длиной 2. Разложение в произведение независимых циклов подстановок f и g изображено на рис. 3.5.1, а и 3.5.1, б соответственно.●



а **б**

Рис. 3.5.1

Перейдем теперь к общему случаю. Пусть S_n – произвольная симметрическая группа степени n и f – произвольная подстановка из S_n. Пусть $\Gamma_f = \langle f \rangle$ – циклическая подгруппа в S_n, порожденная подстановкой f. Элементы i и j множества $\Omega = \{1, 2, \ldots, n\}$ назовем *Γ_f-эквивалентными*, если найдется такая подстановка $g = f^k$ для подходящего целого k, что $g(i) = j$. Тогда $g^{-1}(j) = i$. Очевидно, отношение Γ_f-эквивалентности на множестве Ω рефлексивно $(f^0(i) = i)$, симметрично $(f^k(i) = j \Rightarrow f^{-k}(j) = i)$ и транзитивно $(f^k(i) = j \ \& \ f^l(j) = m \Rightarrow f^l f^k(i) = f^l(f^k(i)) = f^{k+l}(i) = m)$. Тогда Ω разбивается на попарно непересекающиеся классы Γ_f-эквивалентных друг другу элементов, называемых *f-орбитами* (или *Γ_f-орбитами*): $\Omega = \Omega_1 \cup \ldots \cup \Omega_p$. Каждый элемент $i \in \Omega$ принадлежит в точности одной f-орбите, и если $i \in \Omega_k$, то Ω_k состоит из образов элемента i при действии степеней f: $i, f(i), f^2(i), \ldots, f^{l_k - 1}(i)$, где $l_k = |\Omega_k|$ – длина f-орбиты Ω_k. Очевидно,

$l_k \le \mathrm{ord}(f) = |<f>|$, в дальнейшем будет показано, что $l_k \,|\, \mathrm{ord}(f)$. Ясно, что $f^{l_k}(i) = i$, причем l_k – наименьшее натуральное число с таким свойством.

Определение 3.5.3. Положим

$$f_k = (i\ f(i)\ f^2(i) \dots f^{l_k-1}(i)) = \begin{pmatrix} i & f(i) & \dots & f^{l_k-1}(i) \\ f(i) & f^2(i) & \dots & i \end{pmatrix}.$$

Тем самым получим подстановку, называемую *циклом длиной* l_k, действующую тождественно на остальных элементах из Ω.

Итак, f_k действует как f на Ω_k и тождественно на $\Omega \backslash \Omega_k$. Это дает основание считать циклы f_k и f_m при $k \ne m$ независимыми, или непересекающимися (т. к. они действуют на непересекающихся множествах Ω_k и Ω_m). Таким образом, разбиению $\Omega = \Omega_1 \cup \dots \cup \Omega_p$ соответствует разложение подстановки f в произведение: $f = f_1 f_2 \dots f_p$, при этом циклы-сомножители независимы и перестановочны. Естественно в произведении $f = f_1 f_2 \dots f_p$ опускать сомножители, соответствующие орбитам Ω_i из одного элемента, т. к. $f_i = e$ – тождественная подстановка на Ω.

Пример 3.5.4. Подстановку $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 4 & 5 & 6 & 1 & 7 & 2 \end{pmatrix} \in S_8$ можно разложить в произведение независимых циклов следующим образом: $f = (1\ 3\ 4\ 5\ 6)(2\ 8)(7) = (1\ 3\ 4\ 5\ 6)(2\ 8)$.•

Итак, выше была доказана следующая теорема.

Теорема 3.5.2. Каждая подстановка $f \in S_n$, $f \ne e$, является произведением независимых циклов длинами $l_k \ge 2$. Это разложение в произведение определено однозначно с точностью до порядка следования циклов.

Следствие. Порядок подстановки $f \in S_n$ (порядок циклической группы Γ_f) равен наименьшему общему кратному длин независимых циклов, входящих в разложение f.

▷ Пусть $f_k = \begin{pmatrix} i & f(i) & \dots & f^{l_k-1}(i) \\ f(i) & f^2(i) & \dots & i \end{pmatrix} = (i\ f(i)\ f^2(i) \dots f^{l_k-1}(i))$ – цикл длиной l_k. Найдем $\mathrm{ord}(f_k)$ в группе S_n.

$$f_k^{\,2} = \begin{pmatrix} i & f(i) & \dots & f^{l_k-1}(i) \\ f^2(i) & f^3(i) & \dots & f(i) \end{pmatrix} = (i\ f^2(i)\ f^4(i) \dots f^{l_k-2}(i)),$$

$$f_k^{\,3} = \begin{pmatrix} i & f(i) & \dots & f^{l_k-1}(i) \\ f^3(i) & f^4(i) & \dots & f^2(i) \end{pmatrix} = (i\ f^3(i)\ f^6(i) \dots f^{l_k-3}(i))$$

и т. д.,

$$f_k^{\,l_k} = \begin{pmatrix} i & f(i) & \dots & f^{l_k-1}(i) \\ i & f(i) & \dots & f^{l_k-1}(i) \end{pmatrix} = e.$$

Отсюда видно, что l_k – наименьшее натуральное число, такое, что $f_k^{\,l_k} = e$. Значит, $\mathrm{ord}(f_k) = l_k$. Итак, порядок любого цикла равен длине этого цикла.

Имеем $\mathrm{ord}(e)=1$. Согласно теореме 3.5.2 каждая подстановка $f \in S_n$, $f \neq e$, является произведением независимых циклов длинами $l_k \geq 2$, и это разложение определено однозначно с точностью до порядка следования циклов. Порядок подстановки делится на порядок каждого цикла, который входит в ее разложение, причем это наименьшее число с таким свойством по определению порядка элемента группы. Значит, порядок такой подстановки равен НОК порядков этих циклов, т. е. длин этих циклов. ◁

Определение 3.5.4. Цикл длиной 2 называется *транспозицией*.

Теорема 3.5.3. Каждая подстановка $f \in S_n$, $f \neq e$, разлагается в произведение транспозиций.

▷ Согласно теореме 3.5.2 каждая подстановка $f \in S_n$, $f \neq e$, является произведением независимых циклов длинами $l_k \geq 2$. Для доказательства данной теоремы достаточно убедиться в том, что каждый цикл длиной $l_k \geq 2$ можно разложить в произведение транспозиций следующим образом:

$$f_k = (i\ f(i)\ f^2(i) \ldots f^{l_k-1}(i)) = (i\ f^{l_k-1}(i))(i\ f^{l_k-2}(i))(i\ f^{l_k-3}(i))\ldots(i\ f(i)).$$

Это проверяется непосредственным умножением транспозиций. Но данный способ разложения цикла длиной $l_k \geq 2$ в произведение транспозиций не является единственным. Итак, каждая подстановка $f \in S_n$, $f \neq e$, разлагается в произведение транспозиций. ◁

Пример 3.5.5. Разложим подстановку $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 5 & 1 & 6 & 7 & 8 & 9 & 3 & 2 \end{pmatrix} \in S_9$ в произведение независимых циклов и транспозиций:

$$g = (1\ 4\ 6\ 8\ 3)(2\ 5\ 7\ 9) = (1\ 3)(1\ 8)(1\ 6)(1\ 4)(2\ 9)(2\ 7)(2\ 5) =$$
$$= (3\ 1\ 4\ 6\ 8)(5\ 7\ 9\ 2) = (3\ 8)(3\ 6)(3\ 4)(3\ 1)(5\ 2)(4\ 6)(5\ 9)(4\ 6)(5\ 7).$$

На данном примере убеждаемся в том, что разложение подстановки в произведение транспозиций неоднозначно.●

Важно отметить, что любые два разложения одной и той же подстановки в произведение транспозиций содержат либо четное, либо нечетное число сомножителей.

Определение 3.5.5. Подстановка называется *четной*, если она разлагается в произведение четного числа транспозиций, в противном случае она называется *нечетной*. Тождественную подстановку также относят к четным, полагая, что в ее разложении нуль транспозиций.

Таким образом, характер четности фиксированной подстановки не изменяется в зависимости от различных разложений в произведение транспозиций.

Теорема 3.5.4. Все четные подстановки группы S_n образуют подгруппу A_n; при $n = 1$ $A_n = S_n$, при $n \geq 2$ $|A_n| = n!/2$.

▷ Когда $n = 1$, $A_n = S_n = \{e\}$, поэтому, очевидно, $A_n \leq S_n$.

При $n \geq 2$ покажем, что $A_n \leq S_n$ согласно критерию подгруппы (теорема 3.3.1). Рассмотрим $\forall f, g \in A_n$ и разложим их в произведения транспозиций:

$$f = t_1 \ldots t_{2k},\ g = \tau_1 \ldots \tau_{2l}, \text{ где } k, l \in \mathbf{Z}_{\geq 0}.$$
$$g^{-1} = (\tau_1 \ldots \tau_{2l})^{-1} = \tau_{2l}^{-1} \ldots \tau_1^{-1} = \tau_{2l} \ldots \tau_1.$$

Тогда $fg^{-1} = t_1 \dots t_{2k}\tau_{2l} \dots \tau_1 \in A_n$, поскольку в разложении количество всех транспозиций-сомножителей равно $2k + 2l = 2(k + l)$ – четное число. Итак, $A_n \le S_n$.

Пусть $n \ge 2$. Рассмотрим смежный класс τA_n, где τ – некоторая транспозиция. Получим класс нечетных подстановок: $\tau A_n \subseteq S_n \backslash A_n$. Допустим, $|A_n| = |\tau A_n| = a$, а $|S_n \backslash A_n| = n! - a = b$, тогда $a \le b$. Теперь рассмотрим множество $S_n \backslash A_n$ всех нечетных подстановок и применим к ним одну и ту же транспозицию τ – получим столько же четных подстановок, т. е. $|S_n \backslash A_n| = |\tau(S_n \backslash A_n)| = b$ и $\tau(S_n \backslash A_n) \subseteq A_n$, поэтому $b \le a$. Итак, $a = b$. Множества всех четных и нечетных подстановок не пересекаются и в объединении дают все множество S_n. Значит, $|A_n| = |S_n \backslash A_n| = n!/2$. $\triangleleft$

Следствие. A_n – нормальная подгруппа группы S_n; при $n \ge 2$ $S_n : A_n = 2$.

$\triangleright$ При $n = 1$ $A_n = S_n = \{e\}$, поэтому, очевидно, $A_n \triangleleft S_n$.

При $n \ge 2$ согласно теоремам 3.5.4 и 3.4.2 $S_n : A_n = |S_n| / |A_n| = n!/(n!/2) = 2$. И, как было показано в примере 3.4.3, любая подгруппа индекса 2 нормальна. Следовательно, $A_n \triangleleft S_n$ и в этом случае. $\triangleleft$

Определение 3.5.6. Подгруппа A_n всех четных подстановок группы S_n называется *знакопеременной группой степени n*.

§3.6. Факторгруппы

Определение 3.6.1. Пусть $(G, \bullet)$ – группа и H – ее нормальная подгруппа. Множество $\{H, aH, bH, \dots\}$ всех левых или, что то же самое, правых смежных классов (в этом параграфе будем говорить просто «смежных классов») называется *фактормножеством* группы G по подгруппе H и обозначается G/H.

Теорема 3.6.1. Пусть $H \triangleleft G$. Тогда фактормножество G/H является группой относительно индуцированной операции, определенной формулой

$$aH \odot bH = (a \bullet b)H. \qquad (3.6.1)$$

$\triangleright$ Пусть $aH, bH \in G/H$ – произвольные смежные классы, где $a, b \in G$ – фиксированные элементы группы $(G, \bullet)$. Тогда для любых $h_1, h_2 \in H$ и подходящего $h_3 \in H$

$$(a \bullet h_1) \bullet (b \bullet h_2) = a \bullet (h_1 \bullet b) \bullet h_2 = a \bullet (b \bullet h_3) \bullet h_2 = (a \bullet b) \bullet (h_3 \bullet h_2) \in (a \bullet b)H.$$

То есть результат операции $\bullet$ для элементов $a \bullet h_1$ и $b \bullet h_2$ из смежных классов не зависит от выбора элементов $h_1, h_2 \in H$ и принадлежит смежному классу $(a \bullet b)H$, определяемому представителями a и b. Значит, формулой (3.6.1) определена бинарная алгебраическая операция на множестве G/H.

Далее проверим для $(G/H, \odot)$ определение 3.2.1 группы:

1) для любых смежных классов $aH, bH, cH \in G/H$

$$(aH \odot bH) \odot cH = (a \bullet b)H \odot cH = \big((a \bullet b) \bullet c\big)H,$$
$$aH \odot (bH \odot cH) = aH \odot (b \bullet c)H = \big(a \bullet (b \bullet c)\big)H,$$
$$\text{значит, } (aH \odot bH) \odot cH = aH \odot (bH \odot cH).$$

То есть индуцированная операция $\odot$ на факторможестве G/H ассоциативна, поскольку групповая операция $\bullet$ на G ассоциативна;

2) $H \in G/H$ – нейтральный элемент относительно данной индуцированной операции. Действительно, для $\forall\, aH \in G/H$ имеем

$$aH \odot H = aH \odot eH = (a \bullet e)H = aH,\ H \odot aH = eH \odot aH = (e \bullet a)H = aH;$$

3) для $\forall\, aH \in G/H\ \exists\, (aH)^{-1} = a^{-1}H \in G/H$ – обратный элемент относительно данной индуцированной операции. Действительно,

$$aH \odot a^{-1}H = (a \bullet a^{-1})H = eH = H,\ a^{-1}H \odot aH = (a^{-1} \bullet a)H = eH = H.$$

Итак, согласно определению 3.2.1 $(G/H, \odot)$ – группа. $\triangleleft$

Определение 3.6.2. Группа $(G/H, \odot)$ из теоремы 3.6.1 называется *факторгруппой* группы G по нормальной подгруппе H.

Свойства факторгрупп

1. Факторгруппа абелевой группы является абелевой.

$\triangleright$ Пусть $(G, \bullet)$ – абелева группа, $H \triangleleft G$ (любая подгруппа абелевой группы нормальна согласно определению 3.4.3). Для $\forall\, aH, bH \in G/H$ имеем

$$aH \odot bH = (a \bullet b)H = (b \bullet a)H = bH \odot aH,$$

поскольку $(G, \bullet)$ – абелева группа. Итак, $(G/H, \odot)$ – абелева группа. $\triangleleft$

2. Факторгруппа циклической группы является циклической.

$\triangleright$ Пусть $G = < a >$ – циклическая группа, $H \triangleleft G$ (любая подгруппа циклической группы нормальна, поскольку циклическая группа является абелевой). Докажем, что $G/H = < aH >$.

$< aH > \subseteq G/H$, т. к. для $\forall\, k \in \mathbf{Z}$ $(aH)^k = a^kH \in G/H$. Рассмотрим теперь $\forall\, bH \in G/H$, $b \in G$. Так как $G = < a >$ – циклическая группа, то $b = a^m$ для некоторого $m \in \mathbf{Z}$. Тогда $bH = a^mH = (aH)^m \in < aH > \Rightarrow G/H \subseteq < aH >$. Итак, $G/H = < aH >$. $\triangleleft$

Пример 3.6.1. Абелева группа $(\mathbf{Z}, +)$ содержит для всякого $n \in \mathbf{N}$ нормальную подгруппу $(n\mathbf{Z}, +)$, как было показано в примере 3.4.1. Следовательно, факторгруппа $(\mathbf{Z}/n\mathbf{Z}, \oplus)$ – группа классов вычетов с индуцированной операцией сложения. Для всех представителей $i, j = \overline{0, n-1}$ имеем

$\bar{i} \oplus \bar{j} = (i + n\mathbf{Z}) \oplus (j + n\mathbf{Z}) = (i + j) + n\mathbf{Z} = \overline{i + j}$ – сложение классов вычетов.

$(\mathbf{Z}/n\mathbf{Z}, \oplus)$ – абелева и циклическая группа порядка n по свойствам 1 и 2 факторгрупп, поскольку $(\mathbf{Z}, +)$ – абелева и циклическая группа. Поскольку $(\mathbf{Z}, +) = < 1 > = < -1 >$ (п. 1 примера 3.3.5), то, следуя доказательству свойства 2 факторгрупп, $\bar{1}$ и $-\bar{1} = \overline{n-1}$ – образующие элементы $(\mathbf{Z}/n\mathbf{Z}, \oplus)$ $\left(\bar{1} = -\bar{1}\right.$ при $\left. n \le 2\right)$. Вообще же, как было указано в п. 2 примера 3.3.5, $(\mathbf{Z}/n\mathbf{Z}, \oplus) = < \bar{k} >$ для всех $k = \overline{0, n-1}$ с условием $(k, n) = 1$, и количество всех образующих элементов равно 1 при $n = 1$ и $\varphi(n)$ при $n \ge 2$, где φ – функция Эйлера.

Индуцированная операция сложения в $(\mathbf{Z}/n\mathbf{Z}, \oplus)$ может быть задана таблицей Кэли, как отмечалось в §1.7 и в §3.2.$\bullet$

§3.7. Гомоморфизмы групп

Определение 3.7.1. Пусть $(G_1, *)$ и $(G_2, \bullet)$ – две группы. Всякая функция $f: G_1 \to G_2$ и сохраняющая операцию, т. е. обладающая свойством $f(g_1 * g_2) = f(g_1) \bullet f(g_2)$ для $\forall\, g_1, g_2 \in G_1$, называется *гомоморфизмом* из группы G_1 в группу G_2.

Определение 3.7.2. *Образом гомоморфизма* $f: G_1 \to G_2$ называется множество $\mathrm{Im} f = \{f(g) \in G_2 \mid g \in G_1\} = E(f)$, т. е. образ группы G_1 при гомоморфизме f, иногда используется обозначение $\mathrm{Im}\, G_1$ (от англ. image – «образ»).

Определение 3.7.3. *Ядром гомоморфизма* $f: G_1 \to G_2$ называется множество $\mathrm{Ker} f = \{g \in G_1 \mid f(g) = e_2\}$, где e_2 – нейтральный элемент группы $(G_2, \bullet)$ (от англ. kernel – «ядро»).

Свойства гомоморфизмов групп

1. Пусть $(G_1, *)$, $(G_2, \bullet)$, $(G_3, \circ)$ – группы. Если $f: G_1 \to G_2$ и $\psi: G_2 \to G_3$ – гомоморфизмы групп, то $\psi f: G_1 \to G_3$ – гомоморфизм групп.

$\triangleright$ $\psi f: G_1 \to G_3$ – функция по определению композиции функций. Для $\forall\, g_1, g_2 \in G_1$, поскольку $f: G_1 \to G_2$ и $\psi: G_2 \to G_3$ – гомоморфизмы групп, имеем

$$\psi f(g_1 * g_2) = \psi\big(f(g_1 * g_2)\big) = \psi\big(f(g_1) \bullet f(g_2)\big) =$$
$$= \psi\big(f(g_1)\big) \circ \psi\big(f(g_2)\big) = \psi f(g_1) \circ \psi f(g_2).$$

Итак, $\psi f: G_1 \to G_3$ – гомоморфизм групп. $\triangleleft$

2. Если $f: G_1 \to G_2$ – гомоморфизм групп, то:

1) $f(e_1) = e_2$, т. е. f нейтральный элемент первой группы переводит в нейтральный элемент второй группы;

2) для $\forall\, g \in G_1$ $f(g^{-1}) = f(g)^{-1}$, т. е. образ обратного элемента при гомоморфизме есть обратный элемент для образа;

3) $\mathrm{Im} f \le G_2$.

$\triangleright$ $\mathrm{Im} f \subseteq G_2$, $\mathrm{Im} f \neq \varnothing$ по определению, поскольку $G_1 \neq \varnothing$. Так как $f: G_1 \to G_2$ – гомоморфизм групп, то для $\forall f(g_1), f(g_2) \in \mathrm{Im} f$ $\exists\, g_3 = g_1 * g_2 \in G_1$, такой, что $f(g_1) \bullet f(g_2) = f(g_1 * g_2) = f(g_3) \in \mathrm{Im} f$, следовательно, определена бинарная алгебраическая операция $\bullet$ на $\mathrm{Im} f$. Данная операция на $\mathrm{Im} f$ совпадает с групповой операцией на G_2, поэтому она ассоциативна.

Для $\forall\, f(g) \in \mathrm{Im} f$ справедливы равенства $f(e_1) \bullet f(g) = f(e_1 * g) = f(g)$ и $f(g) \bullet f(e_1) = f(g * e_1) = f(g)$, т. е. $f(e_1)$ – нейтральный элемент в $\mathrm{Im} f$. Из первого равенства получаем, что $e_2 \bullet f(g) = f(e_1) \bullet f(g) \Leftrightarrow e_2 = f(e_1)$, поскольку $\exists\, f(g)^{-1} \in G_2$. Тем самым доказано утверждение 1.

Рассмотрим произвольный элемент $g \in G_1$. Тогда получаем, что

$$f(g) \bullet f(g^{-1}) = f(g * g^{-1}) = f(e_1) = e_2,$$

$$f(g^{-1}) \bullet f(g) = f(g^{-1} * g) = f(e_1) = e_2.$$

Это означает, что $f(g^{-1}) = f(g)^{-1}$, т. е. доказано утверждение 2 и для каждого элемента из Imf существует обратный элемент в Imf.

Итак, принимая во внимание все вышесказанное, согласно определениям 3.3.1 и 3.2.1 Im$f \le G_2$, и утверждение 3 также доказано. ◁

3. Если $f : G_1 \to G_2$ – гомоморфизм групп, то Ker$f \lhd G_1$ – ядро гомоморфизма является нормальной подгруппой группы G_1.

▷ Если $f : G_1 \to G_2$ – гомоморфизм групп, то Ker$f \subseteq G_1$ и Ker$f \neq \varnothing$, т. к. $e_1 \in$ Ker$f$ согласно свойству 2. Рассмотрим $g_1 * g_2^{-1}$ для $\forall g_1, g_2 \in$ Ker$f$.

$$f(g_1 * g_2^{-1}) = f(g_1) \bullet f(g_2^{-1}) = f(g_1) \bullet f(g_2)^{-1} = e_2 \bullet e_2^{-1} = e_2.$$

Следовательно, $g_1 * g_2^{-1} \in$ Ker$f$. Значит, по теореме 3.3.1 Ker$f \le G_1$.

Пусть g – произвольный элемент из G_1 и h – произвольный элемент из Ker$f$. Тогда получаем, что $f(g * h * g^{-1}) = f(g) \bullet e_2 \bullet f(g)^{-1} = e_2$. Значит, $g * h * g^{-1} \in$ Ker$f$. Тогда по теореме 3.4.3 Ker$f \lhd G_1$. ◁

Определение 3.7.4. Гомоморфизм $f : G_1 \to G_2$ называется *мономорфизмом* (или *вложением*), если f – инъективная функция; другими словами, если $g_1, g_2 \in G_1$ и $g_1 \neq g_2$, то $f(g_1) \neq f(g_2)$.

Приведем здесь еще одно важное свойство гомоморфизмов групп.

4. Гомоморфизм $f : G_1 \to G_2$ является мономорфизмом тогда и только тогда, когда Ker$f = \{e_1\}$.

▷ Необходимость. Пусть $f : G_1 \to G_2$ является мономорфизмом. Очевидно, что $\{e_1\} \subseteq$ Ker$f$. Допустим, $\exists g \in$ Ker$f$, но $g \neq e_1$. Тогда получаем $f(e_1) = f(g) = e_2$, что противоречит инъективности f. Значит, Ker$f = \{e_1\}$.

Достаточность. Пусть $f : G_1 \to G_2$ – гомоморфизм групп и Ker$f = \{e_1\}$. Допустим, f не является мономорфизмом. Тогда найдутся $g_1, g_2 \in G_1$ с условием $g_1 \neq g_2$, но $f(g_1) = f(g_2)$. Отсюда, поскольку существует $f(g_2)^{-1} \in G_2$, получаем $f(g_1) \bullet f(g_2)^{-1} = e_2$. Согласно свойству 2 и определению гомоморфизмов групп $f(g_1) \bullet f(g_2)^{-1} = f(g_1 * g_2^{-1}) = e_2$. Значит, $g_1 * g_2^{-1} \in$ Ker$f$, откуда имеем $g_1 * g_2^{-1} = e_1 \Leftrightarrow g_1 = g_2$, что противоречит неравенству этих элементов. Таким образом, $f : G_1 \to G_2$ – мономорфизм. ◁

Определение 3.7.5. Гомоморфизм групп $f : G_1 \to G_2$ называется *эпиморфизмом*, если f – сюръективная функция, другими словами, если Im$f = G_2$.

Теорема 3.7.1 (первая теорема о гомоморфизмах групп). Пусть $(G, \bullet)$ – группа, $H \lhd G$, $(G/H, \odot)$ – факторгруппа с индуцированной операцией. Тогда существует эпиморфизм $f : G \to G/H$, такой, что Ker$f = H$.

▷ Построим функцию $f : G \to G/H$, такую, что для $\forall g \in G$ $f(g) = gH$. f – сюръекция, поскольку для любого смежного класса $lH \in G/H$ $\exists f^{-1}(lH) = l \in G$.

Для $\forall g_1, g_2 \in G \quad f(g_1 \bullet g_2) = (g_1 \bullet g_2)H = g_1H \odot g_2H = f(g_1) \odot f(g_2)$. Значит, f – гомоморфизм из G в G/H.

$\mathrm{Ker} f = \{g \in G \mid f(g) = gH = H\}$. Если $h \in H$, то $f(h) = hH = H$ по свойству смежных классов (утверждение 3 теоремы 3.4.1), т. е. $H \subseteq \mathrm{Ker} f$. Если $g \notin H$, то $f(g) = gH \neq H$ (опять же согласно утверждению 3 теоремы 3.4.1) и $g \notin \mathrm{Ker} f$. Значит, $\mathrm{Ker} f \subseteq H$. Итак, $\mathrm{Ker} f = H$. $\triangleleft$

Определение 3.7.6. Гомоморфизм из теоремы 3.7.1 $f : G \to G/H$, где $H \triangleleft G$, $f(g) = gH$ для $\forall g \in G$, называется *естественным* (или *каноническим*) *гомоморфизмом*.

Определение 3.7.7. Биективный гомоморфизм групп называется *изоморфизмом*. Изоморфные группы будем обозначать $(G_1, *) \cong (G_2, \bullet)$ или $G_1 \cong G_2$.

Несложно видеть, что отношение изоморфизма на множестве всех групп является отношением эквивалентности: оно рефлексивно, симметрично и транзитивно. Таким образом, все множество групп разбивается на непересекающиеся классы попарно изоморфных объектов. В математике изоморфные объекты отождествляются. Основная цель теории групп – классифицировать все группы с точностью до изоморфизма.

Теорема 3.7.2 (вторая теорема о гомоморфизмах групп). Пусть $(G_1, \bullet)$ и $(G_2, *)$ – две группы, $\psi: G_1 \to G_2$ – гомоморфизм групп. Тогда $G_1/\mathrm{Ker}\ \psi \cong \mathrm{Im}\ \psi$.

$\triangleright \mathrm{Im}\ \psi \leq G_2$ (свойство 2 гомоморфизмов групп), и $\mathrm{Ker} f \triangleleft G_1$ (свойство 3 гомоморфизмов групп). Построим функцию $f: G_1/\mathrm{Ker}\ \psi \to \mathrm{Im}\ \psi$, полагая $f(g\mathrm{Ker}\ \psi) = \psi(g)$. Если $g_1 = g \bullet h_1$, где $h_1 \in \mathrm{Ker}\ \psi$, – другой представитель смежного класса $g\mathrm{Ker}\ \psi$, то $g\mathrm{Ker}\ \psi = g_1\mathrm{Ker}\ \psi$ по свойству смежных классов (утверждение 3 теоремы 3.4.1). И, поскольку ψ – гомоморфизм, получаем, что

$$f(g_1\mathrm{Ker}\ \psi) = \psi(g_1) = \psi(g \bullet h_1) = \psi(g) * \psi(h_1) = \psi(g) = f(g\mathrm{Ker}\ \psi),$$

т. к. $h_1 \in \mathrm{Ker}\ \psi$. То есть значение функции f не зависит от выбора представителя смежного класса $g\mathrm{Ker}\ \psi$.

f – сюръекция, т. к. для $\forall\ \psi(g) \in \mathrm{Im}\ \psi\ \ \exists f^{-1}(\psi(g)) = g\mathrm{Ker}\ \psi \in G_1/\mathrm{Ker}\ \psi$.

Пусть $g_1\mathrm{Ker}\ \psi \neq g_2\mathrm{Ker}\ \psi$ – различные смежные классы. Если предположить, что $f(g_1\mathrm{Ker}\ \psi) = f(g_2\mathrm{Ker}\ \psi)$, то получаем

$$\psi(g_1) = \psi(g_2) \Leftrightarrow \psi(g_1) * \psi(g_2)^{-1} = e_2 \Leftrightarrow \psi(g_1 \bullet g_2^{-1}) = e_2 \Leftrightarrow$$
$$\Leftrightarrow g_1 \bullet g_2^{-1} \in \mathrm{Ker}\ \psi \Leftrightarrow g_1\mathrm{Ker}\ \psi = g_2\mathrm{Ker}\ \psi,$$

согласно утверждению 2 теоремы 3.4.1, и приходим к противоречию. Итак, $f(g_1\mathrm{Ker}\ \psi) \neq f(g_2\mathrm{Ker}\ \psi)$. Значит, f – инъекция.

Пусть $(G_1/\mathrm{Ker}\ \psi, \odot)$ – факторгруппа. Тогда

$$f(g_1\mathrm{Ker}\ \psi \odot g_2\mathrm{Ker}\ \psi) = f((g_1 \bullet g_2)\mathrm{Ker}\ \psi) = \psi(g_1 \bullet g_2) =$$
$$= \psi(g_1) * \psi(g_2) = f(g_1\mathrm{Ker}\ \psi) * f(g_2\mathrm{Ker}\ \psi)$$

для любых смежных классов $g_1\mathrm{Ker}\ \psi$, $g_2\mathrm{Ker}\ \psi \in G_1/\mathrm{Ker}\ \psi$. Следовательно, $f: G_1/\mathrm{Ker}\ \psi \to \mathrm{Im}\ \psi$ – гомоморфизм групп.

Итак, поскольку f – биекция, приходим к заключению, что f – изоморфизм групп. Значит, $G_1/\mathrm{Ker}\ \psi \cong \mathrm{Im}\ \psi$. ◁

Теорема 3.7.3. Все циклические группы одного и того же порядка изоморфны.

▷ Пусть $|G_1| = |G_2|$, $G_1 = < a >$, $G_2 = < b >$. Причем ord(a) и ord(b) могут быть одновременно конечны и равны либо одновременно бесконечны. Построим функцию $f : G_1 \to G_2$, где $f(a^k) = b^k$, $\forall k \in \mathbf{Z}$. Пусть $a^n \neq a^m$, тогда $a^{n-m} \neq e_1$ и ord(a) $\nmid (n-m)$, значит, $f(a^n) = b^n \neq b^m = f(a^m)$, поскольку ord($a$) = ord($b$). Для $\forall\, b^k \in < b > \ \exists\, f^{-1}(b^k) = a^k \in < a >$. Таким образом, f – биекция. f – гомоморфизм групп, поскольку

$$f(a^n * a^m) = f(a^{n+m}) = b^{n+m} = b^n \bullet b^m = f(a^n) \bullet f(a^m),\ \forall n, m \in \mathbf{Z}.$$

Итак, $G_1 \cong G_2$. ◁

Из данной теоремы и примеров 3.2.5 и 3.3.5 следует, что все бесконечные циклические группы изоморфны группе ($\mathbf{Z}$, +), а все конечные циклические группы порядка $n \in \mathbf{N}$ изоморфны ($\mathbf{Z}/n\mathbf{Z}$, +).

Теорема 3.7.4 (А. Кэли). Всякая группа (G, •) изоморфна некоторой подгруппе симметрической группы $S(G)$. В частности, всякая конечная группа порядка n изоморфна некоторой подгруппе группы S_n.

▷ Любому элементу $a \in G$ поставим в соответствие функциональное преобразование $\psi_a : G \to G$, т. е. $f(a) = \psi_a$, где $\psi_a(g) = a \bullet g$ для $\forall\, g \in G$.

Для $\forall\, h \in G$ если $a \bullet g = h$, то $g = a^{-1} \bullet h \in G$ и $h = \psi_a(g)$, следовательно, ψ_a – сюръекция. Если $g_1, g_2 \in G$ и $g_1 \neq g_2$, то $\psi_a(g_1) \neq \psi_a(g_2)$, т. к. иначе $a \bullet g_1 = a \bullet g_2 \Leftrightarrow a^{-1} \bullet (a \bullet g_1) = a^{-1} \bullet (a \bullet g_2) \Leftrightarrow g_1 = g_2$ – получаем противоречие, следовательно, ψ_a – инъекция. Значит, $\psi_a : G \to G$ – биекция, $\psi_a \in S(G)$.

Итак, $f : G \to S(G)$. Для $\forall\, a, b \in G$ $f(a \bullet b) = \psi_{a \bullet b}$, $f(a) \circ f(b) = \psi_a \circ \psi_b$. Для $\forall\, g \in G$ $\psi_{a \bullet b}(g) = (a \bullet b) \bullet g$, $\psi_a(\psi_b(g)) = a \bullet \psi_b(g) = a \bullet (b \bullet g)$. В силу ассоциативности групповой операции на G получаем, что $\psi_{a \bullet b} = \psi_a \circ \psi_b$, т. е. $f(a \bullet b) = f(a) \circ f(b)$. Следовательно, f – гомоморфизм групп.

$\mathrm{Ker} f = \{a \in G \mid f(a) = \psi_a = e_G\}$, где e_G – тождественная функция на G.

$$\psi_a = e_G \Leftrightarrow \text{для } \forall\, g \in G \ \ \psi_a(g) = g \Leftrightarrow a \bullet g = g \Leftrightarrow (a \bullet g) \bullet g^{-1} = e \Leftrightarrow a = e.$$

Итак, получаем, что $\mathrm{Ker} f = \{e\}$, где e – нейтральный элемент группы (G, •). По теореме 3.7.2 о гомоморфизмах групп $G/\mathrm{Ker} f \cong \mathrm{Im} f$. Поскольку $G/\{e\} = G$, то $G \cong \mathrm{Im} f$, а $\mathrm{Im} f \leq S(G)$ по свойству 2 гомоморфизмов групп. Следовательно, группа (G, •) изоморфна некоторой подгруппе симметрической группы $S(G)$.

Если $|G| = n$, то $S(G) \cong S_n$ и в данном случае по свойству транзитивности изоморфизма получаем, что группа (G, •) изоморфна некоторой подгруппе группы S_n. ◁

Теорема 3.7.5. Существует мономорфизм $f: S_n \to GL_n(\mathbf{Q})$ для всех $n \in \mathbf{N}$, причем такой, что $\det(f(\sigma)) = 1$, если σ – четная подстановка, и $\det(f(\sigma)) = -1$, если σ – нечетная подстановка.

▷ Пусть функция f ставит подстановке σ в соответствие *мономиальную*, или *перестановочную*, *матрицу* $f(\sigma)$, полученную из единичной матрицы E_n соответствующей σ перестановкой строк.

Очевидно, что если $\sigma_1, \sigma_2 \in S_n$ и $\sigma_1 \neq \sigma_2$, то $f(\sigma_1) \neq f(\sigma_2)$. Поэтому функция f инъективна.

При данном отображении f транспозиции соответствует матрица, полученная из E_n одной перестановкой строк. Определитель матрицы меняет знак на противоположный при одной перестановке ее строк, а $\det(E_n) = 1$. Следовательно, если $\sigma \in A_n$, то $\det(f(\sigma)) = 1$, и если $\sigma \in S_n \backslash A_n$, то $\det(f(\sigma)) = -1$.

Определим порядок вычисления композиции подстановок в S_n в соответствии с порядком вычисления произведения матриц в $GL_n(\mathbf{Q})$. Внешней в композиции подстановок будем считать подстановку, стоящую слева. При вычислении произведения матриц строки матрицы, стоящей слева, скалярно умножаются на столбцы матрицы, стоящей справа. Для случая вычисления композиции подстановок и произведения матриц в обратном порядке все рассуждения полностью аналогичны. Пусть $\sigma_1, \sigma_2 \in S_n$. Тогда $f(\sigma_2\sigma_1)$ – мономиальная матрица, полученная из E_n перестановкой строк, соответствующей подстановке $\sigma_2\sigma_1$. Произведение матриц $f(\sigma_2) \cdot f(\sigma_1)$ – мономиальная матрица, полученная из $f(\sigma_1)$ перестановкой строк, соответствующей подстановке σ_2. $f(\sigma_1)$ – мономиальная матрица, полученная из E_n перестановкой строк, соответствующей подстановке σ_1. Поэтому $f(\sigma_2\sigma_1) = f(\sigma_2) \cdot f(\sigma_1)$. Значит, $f: S_n \to GL_n(\mathbf{Q})$ – мономорфизм групп. ◁

Определение 3.7.8. Если гомоморфизм $f: G \to G$ является преобразованием группы G, то его называют *эндоморфизмом*. *Автоморфизм* – это биективный эндоморфизм группы.

Теорема 3.7.6. Множество $\mathrm{Aut}(G)$ всех автоморфизмов произвольной группы $(G, \bullet)$ является группой относительно операции композиции функций.

▷ Композиция гомоморфизмов – гомоморфизм (свойство 1 гомоморфизмов), значит, композиция эндоморфизмов – эндоморфизм. Композиция биекций – биекция, поэтому композиция автоморфизмов – автоморфизм. Композиция автоморфизмов ассоциативна, как композиция функций.

Тождественная функция e_G на множестве G, очевидно, – тождественный автоморфизм. Поэтому $\mathrm{Aut}(G) \neq \varnothing$ для произвольной группы $(G, \bullet)$. e_G – нейтральный элемент относительно операции композиции в $\mathrm{Aut}(G)$. Так как автоморфизм – биекция, для него существует обратная функция – тоже биекция. Рассмотрим $\forall f \in \mathrm{Aut}(G)$. Для $\forall a, b \in G$ $\exists a_1, b_1 \in G$, такие, что $a = f(a_1)$, $b = f(b_1)$, или $a_1 = f^{-1}(a)$, $b_1 = f^{-1}(b)$. Тогда

$$f^{-1}(a \bullet b) = f^{-1}(f(a_1) \bullet f(b_1)) = f^{-1}(f(a_1 \bullet b_1)) = a_1 \bullet b_1 = f^{-1}(a) \bullet f^{-1}(b),$$

поэтому $f^{-1} \in \mathrm{Aut}(G)$.

Итак, доказано, что Aut(G) – группа относительно операции композиции автоморфизмов. ◁

Теорема 3.7.7. Пусть (G, •) – конечная абелева группа порядка $n \in \mathbf{N}$. Тогда для всякого целого k, такого, что (k, n) = 1, функция $f: G \to G$, где $f(g) = g^k$ для $\forall\, g \in G$, есть автоморфизм данной группы.

▷ Для $\forall\, g \in G$ и $\forall\, k \in \mathbf{Z}$ $g^k \in G$, следовательно, $f: G \to G$, где $f(g) = g^k$, – функция. Очевидно, $(g_1 \bullet g_2)^0 = g_1{}^0 \bullet g_2{}^0 = e$ для $\forall\, g_1, g_2 \in G$.

$$f(g_1 \bullet g_2) = (g_1 \bullet g_2)^k = \underbrace{(g_1 \bullet g_2)^{\operatorname{sgn}(k)} \bullet \ldots \bullet (g_1 \bullet g_2)^{\operatorname{sgn}(k)}}_{|k|} = \Big(\underbrace{g_1{}^{\operatorname{sgn}(k)} \bullet \ldots \bullet g_1{}^{\operatorname{sgn}(k)}}_{|k|}\Big) \bullet$$

$$\bullet \Big(\underbrace{g_2{}^{\operatorname{sgn}(k)} \bullet \ldots \bullet g_2{}^{\operatorname{sgn}(k)}}_{|k|}\Big) = g_1{}^k \bullet g_2{}^k = f(g_1) \bullet f(g_2)$$

для $\forall\, g_1, g_2 \in G$ и $\forall\, k \in \mathbf{Z}\backslash\{0\}$, т. к. ($G$, •) – абелева группа. Значит, f – эндоморфизм G при $\forall\, k \in \mathbf{Z}$.

Согласно теореме 1.1.1 $k = n \cdot q + r$ для подходящих $q, r \in \mathbf{Z}$, причем $r = 0$ при $n = 1$, при $n > 1$ $0 < r < n$ и (r, n) = 1, поскольку (k, n) = 1. Тогда для $\forall\, g \in G$ $f(g) = g^{n \cdot q + r} = g^r$, т. к. $g^n = e$ согласно следствию 2 из теоремы 3.4.2 Лагранжа.

Если $n = 1$, то $r = 0$, и если $n = 2$, то $r = 1$, поэтому в данных случаях $f(g) = g$ для $\forall\, g \in G$, т. е. $f = e_G$ – тождественный автоморфизм группы (G, •).

Пусть $n \neq 1, 2$. Для $\forall\, g_1, g_2 \in G$, $g_1 \neq g_2$, рассмотрим $f(g_1)$ и $f(g_2)$. Если предположить, что $f(g_1) = f(g_2)$, то $g_1{}^r = g_2{}^r \Leftrightarrow g_1{}^r \bullet (g_2{}^r)^{-1} = (g_1 \bullet g_2{}^{-1})^r = e$, но $g_1 \bullet g_2{}^{-1} \neq e$, т. к. $g_1 \neq g_2$. Пусть $\operatorname{ord}(g_1 \bullet g_2{}^{-1}) = m > 1$, $m \mid r$ в силу определения порядка элемента группы. Также $m \mid n$ согласно следствию 2 из теоремы 3.4.2 Лагранжа. Итак, $m > 1$ – общий делитель r и n. Получаем противоречие тому, что (r, n) = 1. Таким образом, $f(g_1) \neq f(g_2)$ и f – инъективное функциональное преобразование конечного множества, а значит, f сюръективно по теореме 2.1.2 и биективно по определению. Итак, f – автоморфизм группы (G, •). ◁

§3.8. Криптосистема RSA и система электронной цифровой подписи на ее основе

Определение 3.8.1. *Секретность данных* – свойство данных быть известными и доступными только той группе субъектов, для кого они предназначены. *Криптографическая защита* – это защита информационных процессов от целенаправленных попыток отклонить их от нормальных условий протекания, базирующаяся на криптографических преобразованиях данных. *Имитозащитой* называется защита от навязывания ложной информации. Имитозащита достигается обычно за счет включения в пакет передаваемых данных *имитовставки* – блока информации, зависящего от ключа и данных. *Криптографическим алгоритмом* называется алгоритм преобразования данных, являющийся полностью или частично секретным либо использующий при работе набор секретных па-

раметров. Дополнительно к криптографическим алгоритмам относят алгоритмы, не использующие секретные параметры, но применяющиеся в единой технологической цепочке с криптографическими алгоритмами. *Криптографический протокол* – набор правил, регламентирующих использование криптографических преобразований и алгоритмов в информационных процессах.

Определение 3.8.2. *Криптографическая система* представляет собой набор обратимых криптографических преобразований, или алгоритмов, предназначенных для работы в единой технологической цепочке с целью решения определенной задачи защиты информационного процесса. *Криптопакет* – конкретная программная реализация криптосистемы. *Секретность* – свойство криптосистемы обеспечивать секретность защищаемых данных.

Криптосистема выполняет три основные функции: усиление защищенности данных, облегчение работы с криптоалгоритмами со стороны человека и обеспечение совместимости потока данных с другим программным обеспечением. К основным причинам ненадежности криптосистем следует отнести невозможность применения абсолютно стойких криптоалгоритмов (алгоритм шифрования считается стойким до тех пор, пока не будет доказано обратное), ошибки в реализации и неправильное применение криптоалгоритмов, человеческий фактор.

Новейшие методы защиты информации от несанкционированного доступа называют *групповыми*, т. к. они базируются на теории групп. Ярким примером тому является *криптосистема RSA*, предложенная в 1977 г. исследователями Массачусетского технологического института (США) и получившая название в честь ее создателей (по первым буквам их фамилий). Ими являются Рональд Ривест (род. в 1947 г.) – американский специалист по криптографии, Ади Шамир (род. в 1952 г.) – израильский ученый в области теории вычислительных систем, и Леонард Адлеман (род. в 1945 г.) – американский ученый-теоретик в области компьютерных наук.

Суть криптосистемы RSA заключается в следующем. Находятся два больших простых числа (60–70 десятичных знаков) p и q. Вычисляется $n = p \cdot q$. Тогда $\varphi(n) = (p-1)\cdot(q-1)$, где φ – функция Эйлера. Фиксируется натуральное число e с условиями $e < \varphi(n)$ и НОД$\left(e, \varphi(n)\right) = 1$. Пара (e, n) называется *открытым ключом*. Передаваемая информация шифруется в виде натурального числа c-сообщения, где $c < n$ и НОД$(c, n) = 1$. Тогда $\bar{c}$ – обратимый класс в $\mathbf{Z}/n\mathbf{Z}$, т. е. элемент абелевой мультипликативной группы $\mathbf{Z}/n\mathbf{Z}^*$ порядка $\varphi(n)$. Сообщение c шифруется и передается числом $m \equiv c^e \pmod{n}$, т. е. $\bar{m}$ – e-я степень $\bar{c}$ в $\mathbf{Z}/n\mathbf{Z}$. Согласно теореме 3.7.7 возведение в e-ю степень является автоморфизмом группы $(\mathbf{Z}/n\mathbf{Z}^*, \cdot)$.

Адресат получает сообщение m, знает n и e. Он должен также знать *секретный ключ* – такое натуральное число d, $d < \varphi(n)$, что $e \cdot d \equiv 1 \left(\mathrm{mod}\ \varphi(n)\right)$. Значит, $e \cdot d = \varphi(n) \cdot k + 1$ для некоторого натурального числа k. Тогда в $\mathbf{Z}/n\mathbf{Z}$ согласно следствию 2 из теоремы 3.4.2 Лагранжа имеем следующее равенство:

$\overline{m}^{d} = \overline{c}^{e \cdot d} = \overline{c}^{\varphi(n) \cdot k + 1} = \left(\overline{c}^{\varphi(n)}\right)^{k} \cdot \overline{c} = \overline{1} \cdot \overline{c} = \overline{c}$. Чтобы расшифровать m, адресат должен возвести m в d-ю степень по модулю n – это простая задача. Для возведения в степень удобно использовать двоичное представление числа d.

Криптоаналитик для расшифровки сообщения m должен разложить n на множители p и q. Тогда вычисляется $\varphi(n)$ и d легко находится по значению e из открытого ключа: $\overline{d} = \overline{e}^{-1}$ в $\mathbf{Z}/\varphi(n)\mathbf{Z}$. Именно *факторизация n* (разложение на простые множители) и составляет основную сложность. На настоящий момент неизвестны *полиномиальные алгоритмы* факторизации натуральных чисел, хотя и не доказано, что таких алгоритмов не существует. Здесь речь идет о полиномиальной зависимости времени работы алгоритма от логарифма проверяемого числа, т. е. от количества его цифр. Только в 2002 г. было конструктивно доказано, что задача проверки натурального числа на простоту в общем виде полиномиально разрешима. Таким образом, поиск эффективного метода факторизации чисел оказался более сложной алгоритмической проблемой. На предполагаемой вычислительной сложности задачи факторизации и базируется стойкость криптосистемы RSA.

Чтобы продемонстрировать стойкость своей криптосистемы, изобретатели зашифровали фразу на английском языке, содержащую оксюморон – сочетание слов с противоположным значением: «The magic words are squeamish ossifrage». Что можно перевести так: «Волшебные слова – привередливый бородач» (английское слово «ossifrage» – устаревшее название бородача, или ягнятника, – птицы семейства ястребиных, известной своей неприхотливостью и образом жизни в суровых условиях). В качестве n выступало 129-значное число и в качестве e – 4-значное число, сообщение m было 128-значным числом. Американский математик, писатель, популяризатор науки и всемирно известный специалист по математическим головоломкам Мартин Гарднер (род. в 1914 г.) опубликовал этот криптотекст в журнале «Scientific American» в августе 1977 г., компанией «RSA Data Security» была назначена награда в 100 долларов тому, кто расшифрует текст. В русском переводе заглавие статьи Гарднера звучит так: «Новый вид шифра, для взлома которого потребуются миллионы лет». Именно эта статья сыграла важнейшую роль в распространении информации об RSA, привлекла к криптографии внимание широких кругов неспециалистов и фактически способствовала бурному прогрессу этой области, произошедшему в последовавшие 20 лет.

Текст был расшифрован лишь в апреле 1994 г. В течение шести месяцев были задействованы ресурсы 1600 компьютеров более 20 стран (через Интернет). Организаторы вычислений использовали суперкомпьютер – MasPar. Данные были занесены в 0–1 матрицу из 188346 строк и 188146 столбцов. Файл с этой матрицей превосходил 4 Гбайта, причем каждый бит был существенным. 129-значное число n было разложено на 64- и 65-значные множители p и q. Полученное за расшифровку символическое денежное вознаграждение было пожаловано координаторами вычислительного проекта, одним из которых был известный современный голландский математик и криптоаналитик Арьен Ленстра (род. в 1956 г.), корпорации «The Free Software Foundation».

Определение 3.8.3. *Симметричным* называется шифр, использующий для зашифровывания и расшифровывания либо один и тот же ключ, либо формально различные ключи, допускающие простое преобразование одного в другой. Преобразование шифрования может быть *симметричным* или *асимметричным* (*односторонним*) относительно преобразования расшифровывания. Это важное свойство функции преобразования определяет два класса криптосистем: 1) *сим-*

метричные (или *одноключевые*), 2) *асимметричные* (или *двухключевые*, иначе *криптосистемы с открытым ключом*). *Симметричной* называется криптосистема, реализующая алгоритмы шифрования (симметричные шифры), выполняемые с использованием одного набора параметров – одного ключа. *Асимметричной* называется криптосистема, содержащая преобразования (алгоритмы), наборы параметров которых различны и таковы, что по одному из них вычислительно невозможно определить другие параметры. Таким образом, в симметричной криптосистеме секретный ключ передают отправителю и получателю по защищенному каналу распространения ключей, например, такому, как курьерская служба. В асимметричной же криптосистеме по незащищенному каналу передают только открытый ключ, а секретный сохраняют на месте его генерации. *Криптография с секретным ключом* (или *одноключевая*, *традиционная криптография*) – раздел криптографии, исследующий и разрабатывающий симметричные криптографические системы. *Криптография с открытым ключом* (или *двухключевая*, *современная криптография*) – раздел криптографии, изучающий и разрабатывающий асимметричные криптографические системы.

Примерами известных симметричных алгоритмов шифрования являются DES (англ. Data Encryption Standard), ГОСТ 28147-89, IDEA (англ. International Data Encryption Algorithm), Blowfish, Twofish, CAST, AES (англ. Advanced Encryption Standard), или Rijndael, а асимметричных алгоритмов – RSA и El-Gamal (алгоритм Тахера Эль-Гамаля (род. в 1955 г.), американского ученого арабского происхождения). Во многих странах приняты национальные стандарты шифрования. Так, ГОСТ 28147-89 – стандарт Российской Федерации на шифрование и имитозащиту данных – стал в 1990 г. официальным стандартом СССР, а позже, после распада СССР, федеральным стандартом Российской Федерации и также стандартом СНГ. DES – федеральный стандарт шифрования США в 1977–2001 гг. AES в настоящее время (с 2002 г.) является федеральным стандартом шифрования США.

На начало 2001 г. криптосистема RSA являлась наиболее широко распространенной в мире асимметричной криптосистемой. Криптосистема RSA используется в самых различных продуктах, на различных платформах и во многих отраслях. В настоящее время криптосистема RSA встраивается во многие коммерческие продукты, число которых постоянно увеличивается. Также ее используют операционные системы Microsoft, Apple, Sun и Novell. В аппаратном исполнении RSA-алгоритм применяется в защищенных телефонах, на сетевых платах Ethernet, на смарт-картах, широко используется в криптографическом оборудовании. Кроме того, алгоритм входит в состав всех основных протоколов для защищенных коммуникаций Интернета, в том числе IPSEC (англ. Internet Protocol Security), S/MIME, SSL, S/WAN и TLS (которым предполагается заменить SSL), а также в стандарт PKCS, применяемый в важных приложениях. Также RSA используется во многих учреждениях, например, в правительственных службах, в большинстве корпораций, в государственных лабораториях и университетах. На осень 2000 г. технологии с применением алгоритма RSA были лицензированы более чем 700 компаниями. Технологию шифрования RSA BSAFE используют около 500 млн пользователей всего мира, и это количество имеет явную тенденцию к увеличению по мере роста Интернета.

Определение 3.8.4. *Электронная цифровая подпись* (*ЭЦП*) – последовательность символов, полученная в результате криптографического преобразования электронных данных, которая добавляется к блоку данных и позволяет получателю блока проверить источник и целостность данных и защититься от подделки. ЭЦП используется в качестве аналога обычной подписи, которая устанавливает подлинность какого-либо документа или договора с использованием электронных и компьютерных средств.

Системы ЭЦП являются разделом криптографии. Первые варианты цифровой подписи были реализованы с помощью симметричных криптосистем. Современные процедуры создания и проверки ЭЦП основаны на шифровании с открытым ключом. Технология применения системы ЭЦП предполагает наличие сети абонентов, посылающих друг другу подписанные электронные документы. Для каждого абонента генерируется пара ключей: секретный и открытый. Секретный ключ хранится абонентом в тайне и используется им для формирования ЭЦП. Открытый ключ известен всем другим пользователям и предназначен для проверки ЭЦП получателем подписанного электронного документа. Иначе говоря, открытый ключ является необходимым инструментом, позволяющим проверить подлинность электронного документа и автора подписи. Открытый ключ не позволяет вычислить секретный ключ.

Для генерации пары ключей (секретного и открытого) в алгоритмах ЭЦП используются разные математические схемы, основанные на применении односторонних функций. Эти схемы разделяются на две группы. В основе такого разделения лежат обеспечивающие стойкость алгоритмов известные сложные вычислительные задачи: задача факторизации больших чисел и задача *дискретного логарифмирования* (восстановления показателя по известному основанию и его степени с использованием модульной арифметики). Первой и наиболее известной во всем мире конкретной системой ЭЦП стала система RSA. Более надежный и удобный для реализации на персональных компьютерах ЭЦП алгоритм был разработан в 1984 г. Тахером Эль-Гамалем и получил название EGSA (англ. El-Gamal Signature Algorithm). Идея EGSA основана на том, что для обоснования практической невозможности фальсификации ЭЦП может быть использована более сложная вычислительная задача, чем факторизация большого числа, – задача дискретного логарифмирования в простом поле $(\mathbf{Z}/p\mathbf{Z}, +, \cdot)$, где p – большое простое число. Кроме того, Эль-Гамалю удалось избежать явной слабости алгоритма ЭЦП RSA, связанной с возможностью подделки ЭЦП под некоторыми сообщениями без определения секретного ключа.

Развитием алгоритма EGSA стали DSA (англ. Digital Signature Algorithm), предложенный в 1991 г. в США для использования в федеральном стандарте цифровой подписи DSS (англ. Digital Signature Standard) с последующими изменениями в 1993 г. и в 1996 г., и алгоритмы цифровой подписи, определенные российским стандартом ГОСТ Р 34.10-94 и стандартом Республики Беларусь СТБ 1176.2-99. Алгоритмы цифровых подписей ECDSA (англ. Elliptic Curve Digital Signature Algorithm) и ГОСТ Р 34.10-2001 (новый стандарт цифровой подписи в Российской Федерации с 2002 г.) являются усовершенствованиями алгоритмов цифровых подписей DSA и ГОСТ Р 34.10-94 соответственно и по сравнению

с последними обладают большей стойкостью, основанной на сложности решения задачи дискретного логарифмирования в группе точек эллиптической кривой над простым полем.

Недостатки защищенной *аутентификации* (установления подлинности) были главным препятствием для замены бумажного документооборота электронным; почти везде контракты, чеки, официальные письма, юридические документы все еще выполняются на бумаге. Именно это – необходимость элементов бумажного документооборота – не позволяло полностью перейти к электронным транзакциям. Предлагаемая RSA цифровая подпись – инструмент, который позволит перевести наиболее существенные бумажные документопотоки в электронно-цифровой вид. Благодаря цифровым подписям многие документы – паспорта, избирательные бюллетени, завещания, договора аренды – теперь смогут существовать в электронной форме, а любая бумажная версия будет в этом случае только копией электронного оригинала. Все это стало возможным благодаря стандарту цифровых подписей RSA.

Единая система открытого ключа (англ. public key) допускает обмен документами с ЭЦП между пользователями из разных государств, использующими различное программное обеспечение на различных платформах; такая возможность насущно необходима для развития электронной коммерции. Распространение системы RSA настолько велико, что ее учитывают при создании новых стандартов и зачастую называют стандартом де-факто. Вне зависимости от официальных стандартов существование такого стандарта чрезвычайно важно для развития электронной коммерции и вообще экономики. Множество разрабатываемых в настоящее время стандартов включают в себя либо сам алгоритм RSA или его поддержку, либо рекомендуют криптосистему RSA для обеспечения секретности и (или) аутентификации. Например, включают в себя систему RSA рекомендации IEEE P1363 и WAP WTLS. Стандарт ISO 9796 описывает RSA как совместимый криптографический алгоритм, соответствующий стандарту безопасности ITU-T X.509. Кроме этого, криптосистема RSA является частью стандартов SWIFT, ANSI X9.31 rDSA и проекта стандарта X9.44 для американских банков. Австралийский стандарт управления ключами AS2805.6.5.3 также включает систему RSA. При разработке стандартов цифровых подписей в первую очередь в 1997 г. был разработан стандарт ANSI X9.30, поддерживающий DSS. Годом позже был введен ANSI X9.31, в котором сделан акцент на цифровых подписях RSA, что отвечает фактически сложившейся ситуации, в частности, для финансовых учреждений.

Глава 4. Введение в теорию колец и полей

§4.1. Кольца, их основные типы и свойства

Определение 4.1.1. *Кольцо* $(K, +, \cdot)$ – это алгебраическая система с непустым множеством K и двумя бинарными алгебраическими операциями на нем, которые будем называть *сложением* и *умножением*. Кольцо является абелевой аддитивной группой, а умножение и сложение связаны законами дистрибутивности: $(a+b)\cdot c = a\cdot c + b\cdot c$ и $c\cdot(a+b) = c\cdot a + c\cdot b$ для произвольных $a, b, c \in K$. Нейтральный элемент аддитивной группы кольца $(K, +)$ называют *нулем* и часто обозначают символом 0.

Определение 4.1.2. Кольцо $(K, +, \cdot)$ называется *конечным*, если K – конечное множество, в противном случае – *бесконечным*. *Порядком* конечного кольца $|K|$ называется мощность множества K. Основная классификация колец ведется по свойствам операции умножения. Различают *ассоциативные* кольца, когда операция умножения ассоциативна, и *неассоциативные* кольца соответственно. Ассоциативные кольца делятся на *кольца с единицей* (обладающие нейтральным элементом относительно умножения, который часто будем обозначать символом 1) и *без единицы*; *коммутативные* (операция умножения коммутативна) и *некоммутативные* соответственно.

Пример 4.1.1. Приведем примеры колец с различными свойствами.

1. $(\mathbf{Z}, +, \cdot)$, $(\mathbf{Q}, +, \cdot)$, $(\mathbf{R}, +, \cdot)$, $(\mathbf{C}, +, \cdot)$ – соответственно кольца целых, рациональных, вещественных и комплексных чисел с обычными операциями сложения и умножения. Данные кольца называются *числовыми* и являются бесконечными ассоциативными и коммутативными кольцами с единицей 1.

2. $(\mathbf{Z}/n\mathbf{Z}, +, \cdot)$ – кольцо классов вычетов по модулю $\forall\, n \in \mathbf{N}$ с операциями сложения и умножения. Данное кольцо является конечным порядка n, ассоциативным и коммутативным кольцом с единицей $\bar{1}$.

3. Множество $M_n(K)$ всех квадратных матриц фиксированного порядка $n \in \mathbf{N}$ с коэффициентами из кольца $(K, +, \cdot)$ с операциями матричного сложения и умножения. В частности, K может быть равно $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ (в этом случае $M_n(K)$ бесконечно) или $\mathbf{Z}/k\mathbf{Z}$ при $\forall\, k \in \mathbf{N}$ $\left(\text{тогда } |M_n(K)| = k^{n^2}\right)$. Кольцо $(M_n(K), +, \cdot)$ ассоциативно, коммутативно при $n = 1$ и некоммутативно при $n \geq 2$ и $k > 1$, обладает единицей E_n (E_n – нулевая матрица, если $k = 1$).

4. Множество всех вещественных функций, определенных на фиксированном интервале $(a; b)$ вещественной числовой прямой, с обычными операциями сложения и умножения функций является бесконечным ассоциативным и коммутативным кольцом с единицей – функцией-константой 1.

5. Множество всех полиномов (многочленов) $K[x]$ с коэффициентами из ассоциативного и коммутативного кольца $(K, +, \cdot)$ от одной переменной x с естественными операциями сложения и умножения полиномов является беско-

нечным при $K \neq \{0\}$, ассоциативным и коммутативным кольцом. $(K[x], +, \cdot)$ обладает единицей $1 \in K$ тогда и только тогда, когда $(K, +, \cdot)$ обладает единицей. В частности, примерами таких колец полиномов с единицами являются $\mathbf{Z}[x]$, $\mathbf{Q}[x]$, $\mathbf{R}[x]$, $\mathbf{C}[x]$, $\mathbf{Z}/n\mathbf{Z}[x]$ при $\forall\, n \in \mathbf{N}$.

6. Кольцо векторов $(V_3(\mathbf{R}), +, \times)$ с операциями сложения и векторного умножения является бесконечным неассоциативным кольцом, поскольку здесь $(i \times i) \times k = 0 \times k = 0$, $i \times (i \times k) = i \times (-j) = -k \neq 0$, где i, j, k – базисные орты $V_3(\mathbf{R})$.

7. Кольцо $(\{0\}, +, \cdot)$ с операциями сложения и умножения: $0 + 0 = 0$, $0 \cdot 0 = 0$. Данное кольцо имеет порядок 1, ассоциативно, коммутативно и обладает единицей 0.●

Теорема 4.1.1. Пусть $(K, +, \cdot)$ – ассоциативное кольцо с единицей. Тогда множество K^* обратимых относительно умножения элементов кольца K – мультипликативная группа.

▷Проверим выполнение определения 3.2.1 группы. Пусть $a, b \in K^*$. Покажем, что $a \cdot b \in K^*$. $\exists\, (a \cdot b)^{-1} = b^{-1} \cdot a^{-1} \in K$. Действительно,

$$(a \cdot b) \cdot (b^{-1} \cdot a^{-1}) = a \cdot (b \cdot b^{-1}) \cdot a^{-1} = a \cdot 1 \cdot a^{-1} = 1,$$
$$(b^{-1} \cdot a^{-1}) \cdot (a \cdot b) = b^{-1} \cdot (a^{-1} \cdot a) \cdot b = b^{-1} \cdot 1 \cdot b = 1,$$

где $a^{-1}, b^{-1} \in K$ – обратные элементы для a и b соответственно. Далее:

1) умножение в K^* ассоциативно, т. к. K – ассоциативное кольцо;

2) $(1^{-1} = 1\colon\ 1 \cdot 1 = 1) \Rightarrow 1 \in K^*$, 1 – нейтральный элемент относительно умножения в K^*;

3) для $\forall\, a \in K^*$, $a^{-1} \in K^*$, т. к. $(a^{-1}) \cdot a = a \cdot (a^{-1}) = 1 \Leftrightarrow (a^{-1})^{-1} = a$.◁

Определение 4.1.3. Множество K^* обратимых относительно умножения элементов ассоциативного кольца с единицей $(K, +, \cdot)$ называют *мультипликативной группой кольца*.

Пример 4.1.2. Приведем примеры мультипликативных групп различных ассоциативных колец с единицами.

1. $\mathbf{Z}^* = \{1, -1\}$.

2. $M_n(\mathbf{Q})^* = GL_n(\mathbf{Q})$, $M_n(\mathbf{R})^* = GL_n(\mathbf{R})$, $M_n(\mathbf{C})^* = GL_n(\mathbf{C})$ для $\forall\, n \in \mathbf{N}$.

3. $\mathbf{Z}/n\mathbf{Z}^*$ – множество обратимых классов вычетов, $\mathbf{Z}/n\mathbf{Z}^* = \left\{ \bar{k} \mid (k, n) = 1,\ 0 \le k < n \right\}$. $\mathbf{Z}/\mathbf{Z}^* = \mathbf{Z}/\mathbf{Z} = \left\{ \bar{0} \right\}$, при $n > 1$ $|\, \mathbf{Z}/n\mathbf{Z}^* \,| = \varphi(n)$, где φ – функция Эйлера.

4. $\{0\}^* = \{0\}$, т. к. в данном случае $1 = 0$.●

Определение 4.1.4. Если в ассоциативном кольце $(K, +, \cdot)$ с единицей группа $K^* = K \backslash \{0\}$, где 0 – нейтральный элемент относительно сложения, то такое кольцо называют *телом* или *алгеброй с делением*. Коммутативное тело называется *полем*.

Из данного определения очевидно, что в теле $K^* \neq \varnothing$ и $1 \in K^*$, значит, $1 \neq 0$, поэтому минимальное тело, являющееся полем, состоит из двух элементов: 0 и 1.

Пример 4.1.3.

1. (**Q**, +, ·), (**R**, +, ·), (**C**, +, ·) – соответственно *числовые поля* рациональных, вещественных и комплексных чисел.

2. (**Z**/p**Z**, +, ·) – конечное поле из p элементов, если p – простое число. Например, (**Z**/2**Z**, +, ·) – минимальное поле из двух элементов.

3. Некоммутативным телом является *тело кватернионов* – совокупность всех *кватернионов*, т. е. выражений вида $h = a + bi + cj + dk$, где $a, b, c, d \in \mathbf{R}$, $i^2 = j^2 = k^2 = -1$, $i \cdot j = k = -j \cdot i$, $j \cdot k = i = -k \cdot j$, $i \cdot k = -j = -k \cdot i$, с операциями сложения и умножения. Это 4-мерное линейное векторное пространство над **R** с базисом $\{1, i, j, k\}$. Кватернионы складываются и перемножаются почленно с учетом указанных выше формул. Для всякого $h \neq 0$ обратный кватернион имеет вид $h^{-1} = \frac{a - bi - cj - dk}{a^2 + b^2 + c^2 + d^2}$.●

Различают кольца с делителями нуля и кольца без делителей нуля.

Определение 4.1.5. Если в кольце найдутся ненулевые элементы a и b, такие, что $a \cdot b = 0$, то их называют *делителями нуля*, а само кольцо – *кольцом с делителями нуля*. В противном случае кольцо называется *кольцом без делителей нуля*.

Пример 4.1.4.

1. Кольца (**Z**, +, ·), (**Q**, +, ·), (**R**, +, ·), (**C**, +, ·) – кольца без делителей нуля.

2. В кольце $(V_3(\mathbf{R}), +, \times)$ каждый ненулевой вектор является делителем нуля, поскольку $v \times v = 0$ для всех $v \in V_3(\mathbf{R})$.

3. В кольце матриц $(M_3(\mathbf{Z}), +, \cdot)$ примерами делителей нуля являются матрицы $A = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ и $B = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{pmatrix}$, т. к. $A \cdot B = O$ (нулевая матрица).

В общем случае при $n \geq 2$ кольцо $(M_n(K), +, \cdot)$, когда K равно **Z**, **Q**, **R**, **C** или **Z**/k**Z** при $\forall\, k \in \mathbf{N}_{\geq 2}$, имеет делители нуля.

4. В кольце (**Z**/n**Z**, +, ·) с составным $n = k \cdot m$, где $1 < k, m < n$, классы вычетов $\bar{k}$ и $\bar{m}$ являются делителями нуля, т. к. $\bar{k} \cdot \bar{m} = \overline{k \cdot m} = \bar{n} = \bar{0}$.●

Ниже приведем основные свойства колец и полей.

Свойства колец

1. В любом кольце $(K, +, \cdot)$ уравнение $a + x = b$ имеет единственное решение для $\forall\, a, b \in K$:

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

▷ Поскольку $(K, +)$ – группа, для каждого $a \in K$ существует противоположный элемент $-a \in K$, такой, что $a + (-a) = 0$. Прибавив к обеим частям уравнения $-a$, получим решение $x = b + (-a) \in K$.

Если $y \in K$ – другое решение данного уравнения, то $a + x = b$ и $a + y = b \Rightarrow$ $\Rightarrow a + x = a + y$, тогда, прибавляя $-a$ к обеим частям последнего равенства, получаем $x = y$. ◁

Определение 4.1.6. Сумма $a+(-b)$ называется *разностью* элементов a и b и обозначается $a-b$. Для $\forall\,(a, b)\in K\ a-b\in K$. Таким образом, $\forall\,(a, b)\mapsto a-b$ – определена бинарная алгебраическая операция *вычитания* в любом кольце.

2. Умножение в кольце дистрибутивно относительно вычитания:

$$a\cdot(b-c)=a\cdot b-a\cdot c \text{ и } (b-c)\cdot a=b\cdot a-c\cdot a \text{ для } \forall\, a, b, c\in K.$$

▷ Умножение в кольце дистрибутивно относительно сложения. Поэтому

$$a\cdot(b-c)+a\cdot c=a\cdot(b-c+c)=a\cdot b\Rightarrow a\cdot(b-c)=a\cdot b-a\cdot c;$$

$$(b-c)\cdot a+c\cdot a=(b-c+c)\cdot a=b\cdot a\Rightarrow (b-c)\cdot a=b\cdot a-c\cdot a.\ \triangleleft$$

3. Свойство нуля: $a\cdot 0=0\cdot a=0$ для $\forall\, a\in K$.

▷ $a\cdot 0=a\cdot(a-a)=a\cdot a-a\cdot a=0$; $0\cdot a=(a-a)\cdot a=a\cdot a-a\cdot a=0.\ \triangleleft$

4. Правила знаков:

$$a\cdot(-b)=(-a)\cdot b=-(a\cdot b),\ (-a)\cdot(-b)=a\cdot b \text{ для } \forall\, a, b\in K.$$

▷ Согласно законам дистрибутивности и свойству 3 колец имеем:

$$a\cdot(-b)+a\cdot b=a\cdot(-b+b)=a\cdot 0=0\Rightarrow a\cdot(-b)=-(a\cdot b);$$

$$(-a)\cdot b+a\cdot b=(-a+a)\cdot b=0\cdot b=0\Rightarrow (-a)\cdot b=-(a\cdot b);$$

$$(-a)\cdot(-b)+\big(-(a\cdot b)\big)=(-a)\cdot(-b)+(-a)\cdot b=(-a)\cdot(-b+b)=(-a)\cdot 0=0\Rightarrow$$

$$\Rightarrow (-a)\cdot(-b)=-\big(-(a\cdot b)\big)=a\cdot b.\ \triangleleft$$

5. Если $|K|>1$, то в кольце $(K, +, \cdot)$ с единицей $1\neq 0$.

▷ $0\in K$, т. к. $(K, +)$ – группа. Поскольку $|K|>1$, то в $K\ \exists\, a\neq 0$. Пусть $1=0$, тогда $a\cdot 1=a\cdot 0\ \Rightarrow\ a=0$, т. к. $a\cdot 1=a$, а $a\cdot 0=0$ по свойству 3 колец. Полученное противоречие доказывает, что $1\neq 0.\ \triangleleft$

6. В ассоциативном кольце $(K, +, \cdot)$ с единицей делители нуля необратимы.

▷ Если бы для ненулевых элементов $a, b\in K$ произведение $a\cdot b=0$ и существовал бы $b^{-1}\in K$, то $a\cdot b\cdot b^{-1}=0$. С другой стороны, $a\cdot b\cdot b^{-1}=a\cdot\big(b\cdot b^{-1}\big)=$ $=a\cdot 1=a$, откуда $a=0$ – противоречие тому, что $a\neq 0$. Аналогично, если сделать предположение, что существует $a^{-1}\in K$, то получим противоречие тому, что $b\neq 0.\ \triangleleft$

Свойства полей

Для полей справедливы все свойства колец. Будем обозначать в дальнейшем поле $(P, +, \cdot)$. Также отметим специфические свойства полей.

1. В поле нет делителей нуля.

▷ Поскольку по определению поле является ассоциативным кольцом с единицей, в нем согласно свойству 6 колец делители нуля необратимы. Но т. к., с другой стороны, по определению поля $P^*=P\backslash\{0\}$, то в поле $(P, +, \cdot)$ отсутствуют делители нуля. $\triangleleft$

2. В поле $(P, +, \cdot)$ уравнение $a\cdot x=b$ для $\forall\, a\in P^*$, $\forall\, b\in P$ имеет единственное решение:

$$x=a^{-1}\cdot b.$$

▷ Так как $a\in P^*$, то $\exists\, a^{-1}\in P^*$. Умножая a на $a^{-1}\cdot b$, получаем b, поэтому $x=a^{-1}\cdot b\in P$ – решение уравнения. Если $y\in P$ – другое решение, то $a\cdot x=a\cdot y$, умножая слева обе части равенства на a^{-1}, получаем, что $x=y.\ \triangleleft$

§4.2. Подкольца. Идеалы колец

Определение 4.2.1. *Подкольцо* кольца $(K, +, \cdot)$ – это подгруппа L аддитивной группы $(K, +)$, в свою очередь являющаяся кольцом, т. е. для $\forall\, l_1, l_2 \in L$ выполняются свойства:

1) $l_1 - l_2 \in L$;

2) $l_1 \cdot l_2 \in L$.

Обозначения подкольца: $L \le K$ или $L < K$, если $L \subset K$.

Очевидно, что для любого кольца $(K, +, \cdot)$ подмножества $\{0\}$ и K являются подкольцами.

Определение 4.2.2. Подкольцо L кольца K называется *собственным* (или *нетривиальным*), если $L \ne K$ и $L \ne \{0\}$.

Пример 4.2.1.

1. $(n\mathbf{Z}, +, \cdot)$ при $\forall\, n \in \mathbf{Z}$ – подкольцо кольца $(\mathbf{Z}, +, \cdot)$. Это подкольцо – кольцо без единицы при $|n| > 1$, хотя само кольцо $(\mathbf{Z}, +, \cdot)$ обладает единицей 1.

2. $(\mathbf{Z}, +, \cdot) < (\mathbf{Q}, +, \cdot) < (\mathbf{R}, +, \cdot) < (\mathbf{C}, +, \cdot)$.

3. $(M_n(\mathbf{Z}), +, \cdot) < (M_n(\mathbf{Q}), +, \cdot) < (M_n(\mathbf{R}), +, \cdot) < (M_n(\mathbf{C}), +, \cdot)$ для $\forall\, n \in \mathbf{N}$.●

Упражнение 4.2.1. Проверить, что матричное ассоциативное, некоммутативное кольцо $M_2(\mathbf{C})$ с единицей E_2 и делителями нуля содержит подкольцо матриц $C = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbf{C} \right\}$, а C, в свою очередь, содержит подкольцо скалярных матриц $S = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \mid a \in \mathbf{C} \right\}$. Показать, что C – коммутативное кольцо с единицей и без делителей нуля, а S – поле. Проверить также, что подкольцами $M_2(\mathbf{C})$ являются следующие подмножества матриц: $J_1 = \left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \mid a, b \in \mathbf{C} \right\}$, $J_2 = \left\{ \begin{pmatrix} 0 & c \\ 0 & d \end{pmatrix} \mid c, d \in \mathbf{C} \right\}$, $J_3 = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbf{C} \right\}$, $J_4 = \left\{ \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix} \mid c, d \in \mathbf{C} \right\}$. Доказать, что J_1–J_4 – некоммутативные кольца без единицы и с делителями нуля.

Определение 4.2.3. Непустое подмножество Δ, $\Delta \subseteq P$, поля $(P, +, \cdot)$ называют *подполем* поля P, если Δ само является полем:

1) Δ – подкольцо кольца P;

2) для $\forall\, \delta \in \Delta \backslash \{0\} \Rightarrow \delta^{-1} \in \Delta$.

В этом случае поле $(P, +, \cdot)$ называется *расширением* поля $(\Delta, +, \cdot)$. Последовательность расширений полей $P_1 \le P_2 \le \ldots \le P_n$, где $n \in \mathbf{N}_{\ge 3}$, называется *башней расширений* полей.

Пример 4.2.2. $(\mathbf{Q}, +, \cdot)$ – подполе $(\mathbf{R}, +, \cdot)$ и $(\mathbf{C}, +, \cdot)$, а $(\mathbf{R}, +, \cdot)$ – подполе $(\mathbf{C}, +, \cdot)$ согласно п. 2 примера 4.2.1 и определению 4.2.3. Поэтому $\mathbf{Q} < \mathbf{R} < \mathbf{C}$ – башня расширений числовых полей.●

В теории колец наибольшее значение имеют подкольца специального вида, которые носят названия *идеалов*.

Определение 4.2.4. Подкольцо J кольца $(K, +, \cdot)$ называется *левым идеалом* кольца K, если для любых $k \in K$ и $j \in J$ выполняется условие $k \cdot j \in J$, т. е. $KJ = \{k \cdot j \mid k \in K, j \in J\} \subseteq J$. Если же $JK = \{j \cdot k \mid j \in J, k \in K\} \subseteq J$, то подкольцо J называют *правым идеалом* кольца K. *Двусторонний идеал* кольца – идеал, являющийся одновременно и левым, и правым. Для двустороннего идеала J кольца K будет использоваться обозначение $J \lhd K$.

Ясно, что в коммутативном кольце все идеалы двусторонние, в этом случае будем называть их просто *идеалами* кольца.

Легко видеть, что $\{0\}$ и K – тривиальные двусторонние идеалы любого кольца $(K, +, \cdot)$.

Пример 4.2.3.

1. $m\mathbf{Z} = \{m \cdot q \mid q \in \mathbf{Z}\}$ – идеал коммутативного кольца целых чисел $(\mathbf{Z}, +, \cdot)$ для всякого целого числа m. Отметим, что для $\forall\, m \in \mathbf{Z}$ справедливо равенство множеств $-m\mathbf{Z} = m\mathbf{Z}$. Очевидно, что $m\mathbf{Z} \neq \mathbf{Z}$, если $|m| \neq 1$. Можно, например, построить следующие бесконечные цепочки идеалов кольца $\mathbf{Z}$:

$\mathbf{Z} \rhd 2\mathbf{Z} \rhd 4\mathbf{Z} \rhd 8\mathbf{Z} \rhd 16\mathbf{Z} \rhd \ldots \rhd 2^n\mathbf{Z} \rhd \ldots$, $\mathbf{Z} \rhd 2\mathbf{Z} \rhd 6\mathbf{Z} \rhd 18\mathbf{Z} \rhd \ldots \rhd 2 \cdot 3^n\mathbf{Z} \rhd \ldots$, $n \in \mathbf{Z}_{\geq 0}$.

2. В коммутативном кольце $(\mathbf{Z}/n\mathbf{Z}, +, \cdot)$ с составным $n = pq$, где $1 < p, q < n$, подмножество классов вычетов $J\overline{p} = \{\overline{p}, \overline{2p}, \ldots, \overline{(q-1)p}, \overline{0}\}$ замкнуто относительно операций вычитания и умножения классов вычетов и, следовательно, образует подкольцо. Легко видеть, что $J\overline{p}$ – идеал кольца $\mathbf{Z}/n\mathbf{Z}$. Аналогично, идеалом того же кольца является подмножество $J\overline{q} = \{\overline{q}, \overline{2q}, \ldots, \overline{(p-1)q}, \overline{0}\}$. •

Упражнение 4.2.2. Проверить, что в матричном кольце $M_2(\mathbf{C})$ (упражнение 4.2.1) подкольца J_1 и J_2 – левые, но не правые идеалы, а J_3 и J_4 – правые, но не левые идеалы.

Обозначим ma для $m \in \mathbf{Z}$ целочисленное кратное элемента a кольца $(K, +, \cdot)$. То есть $0a = 0 \cdot a = 0$ и $ma = \underbrace{\operatorname{sgn}(m)a + \ldots + \operatorname{sgn}(m)a}_{|m|}$ при $m \in \mathbf{Z}\backslash\{0\}$. Здесь $\operatorname{sgn}(m)a = \begin{cases} a, & m > 0; \\ -a, & m < 0. \end{cases}$ Очевидно, что $ma \in K$ для всех $a \in K$.

Определение 4.2.5. Для каждого элемента a ассоциативного кольца $(K, +, \cdot)$ подмножество $Ka + \mathbf{Z}a = \{k \cdot a + ma \mid k \in K, m \in \mathbf{Z}\}$ есть левый идеал кольца K, называемый *главным левым идеалом*, *порожденным элементом* a. Аналогично определяется *главный правый идеал*, *порожденный элементом* a: $aK + \mathbf{Z}a = \{a \cdot k + ma \mid k \in K, m \in \mathbf{Z}\}$. *Двусторонний главный идеал*, *порожденный элементом* a, является одновременно левым и правым, т. е. имеет вид $KaK + Ka + aK + \mathbf{Z}a = \{k_1 \cdot a \cdot k_1' + \ldots + k_n \cdot a \cdot k_n' + k' \cdot a + a \cdot k'' + ma \mid n \in \mathbf{N}, \quad m \in \mathbf{Z}, k_i, k_i', k', k'' \in K, i = \overline{1,n}\}$. Элемент a в этих случаях называется *образующим главного идеала*.

Когда $(K, +, \cdot)$ – ассоциативное кольцо с единицей, для каждого $a \in K$ главные левый, правый и двусторонний идеалы, порожденные a, выглядят соответственно следующим образом:

$$Ka + \mathbf{Z}a = Ka = \{k \cdot a \mid k \in K\},\ aK + \mathbf{Z}a = aK = \{a \cdot k \mid k \in K\},$$

$$KaK + Ka + aK + \mathbf{Z}a = KaK = \left\{ k_1 \cdot a \cdot k_1' + \ldots + k_n \cdot a \cdot k_n' \mid n \in \mathbf{N},\ k_i, k_i' \in K,\ i = \overline{1,n} \right\}.$$

В коммутативном кольце понятия главных левого, правого и двустороннего идеалов с образующим элементом a эквивалентны. В таком случае для каждого фиксированного элемента a данные идеалы представляют собой одно и то же множество, которое будем называть просто *главным идеалом*, *порожденным элементом a*, и обозначать (a).

В собственном идеале J не может быть обратимых элементов кольца K, т. к. иначе $1 \in J$, тогда $J = K$, что приводит к противоречию. Поскольку в поле все элементы, кроме 0, обратимы, то в поле нет собственных идеалов. С другой стороны, необратимые элементы ассоциативного кольца, в том числе и делители нуля, порождают главные собственные идеалы. Отсюда, в частности, также следует, что в поле нет делителей нуля (свойство 1 полей).

Определение 4.2.6. Ассоциативное кольцо, в котором каждый идеал является главным, называется *кольцом главных идеалов*. В некоммутативном случае различают *кольцо главных левых идеалов* и *кольцо главных правых идеалов*.

Пример 4.2.4. Рассмотрим $(P[x, y], +, \cdot)$ – кольцо полиномов от двух переменных x и y над произвольным полем P. Здесь $P[x, y]$ представляет собой множество всех конечных сумм одночленов вида $bx^k y^l$, где $b \in P$ и $k, l \in \mathbf{Z}_{\geq 0}$, причем в записи многочлена подобные члены должны быть приведены. Складываются и умножаются полиномы из $P[x, y]$ почленно. Таким образом, несложно видеть, что $(P[x, y], +, \cdot)$ является бесконечным ассоциативным и коммутативным кольцом с единицей ($1 \in P$). В данном кольце всякий главный идеал, порожденный фиксированным многочленом $f(x, y)$ имеет следующий вид: $(f(x, y)) = \{g(x, y) \cdot f(x, y) \mid g(x, y) \in P[x, y]\}$.

Подмножество $J = \{g(x, y) \cdot x + h(x, y) \cdot y \mid g(x, y), h(x, y) \in P[x, y]\}$ является идеалом кольца $(P[x, y], +, \cdot)$ согласно определению 4.2.4. Если бы $J = (f(x, y))$ для некоторого $f(x, y) \in P[x, y]$, то x и y, как элементы J, были бы соответственно представимы в виде $x = p(x, y) \cdot f(x, y)$ и $y = q(x, y) \cdot f(x, y)$ для подходящих $p(x, y), q(x, y) \in P[x, y]$. По правилу умножения многочленов из $P[x, y]$ это возможно только при $f(x, y) = \alpha \in P^*$, но $\alpha \notin J$ ни для каких $\alpha \in P^*$. Данные рассуждения показывают, что идеал J не является главным. Поэтому $(P[x, y], +, \cdot)$ не является кольцом главных идеалов.●

Теорема 4.2.1. $(\mathbf{Z}, +, \cdot)$ – кольцо главных идеалов.

▷ В $(\mathbf{Z}, +, \cdot)$, как в коммутативном кольце, все идеалы являются двусторонними и понятия главных левого, правого и двустороннего идеалов эквивалентны. Поскольку $(\mathbf{Z}, +, \cdot)$ – кольцо с единицей, все главные идеалы в нем имеют вид $(a) = \{k \cdot a \mid k \in \mathbf{Z}\}$, где $a \in \mathbf{Z}$.

1. $J = \{0\}$ – главный идеал. Действительно, $\{0\} = (0) = \{k \cdot 0 \mid k \in \mathbf{Z}\}$.

2. Пусть теперь идеал $J \neq \{0\}$. Докажем, что если t – наименьшее натуральное число из J, то $J = (t)$.

Так как $t \in J$, то $(t) = \{k \cdot t \mid k \in \mathbf{Z}\} \subseteq J$. Пусть $m \in J$, согласно теореме 1.1.1 справедливо представление $m = q \cdot t + r$, где $q, r \in \mathbf{Z}$, $0 \le r < t$. Тогда $m - q \cdot t = r$, и если $r \neq 0$, то в J содержится натуральное число r, меньшее t, что невозможно. Итак, $m = q \cdot t$, таким образом, $m \in (t)$, следовательно, $J \subseteq (t)$. Итак, $J = (t)$. ◁

Замечание. Согласно теореме 4.2.1 все идеалы кольца $(\mathbf{Z}, +, \cdot)$ имеют вид (t), где $t \in \mathbf{Z}$. Поскольку для $\forall\, t \in \mathbf{Z}$ $(t) = t\mathbf{Z} = -t\mathbf{Z} = (-t)$, как уже отмечалось в п. 1 примера 4.2.3, в дальнейшем можно ограничиться обозначением идеалов кольца $(\mathbf{Z}, +, \cdot)$ вида (t), где $t \in \mathbf{Z}_{\ge 0}$.

Пример 4.2.5. При $\forall\, n \in \mathbf{N}_{>1}$ $(n\mathbf{Z}, +, \cdot)$ – коммутативное кольцо без единицы, как уже отмечалось в п. 1 примера 4.2.1. Поэтому все главные идеалы в нем имеют вид $(a) = n\mathbf{Z}a + \mathbf{Z}a = \mathbf{Z}a = \{ka \mid k \in \mathbf{Z}\}$, где $a \in n\mathbf{Z}$, при этом можно брать $a \in n\mathbf{Z}_{\ge 0}$, т. к. $(a) = (-a)$. Подчеркнем, что $(a) \neq n\mathbf{Z}a$, поскольку $1 \notin n\mathbf{Z}$.

Аналогично доказательству теоремы 4.2.1 можно показать, что $\{0\} = (0)$ и для любого идеала $J \neq \{0\}$ данного кольца $J = (nt)$, где t – минимальное натуральное число со свойством $nt \in J$. Таким образом, $(n\mathbf{Z}, +, \cdot)$ – кольцо главных идеалов.●

Определение 4.2.7. Собственный идеал кольца называется *максимальным*, если он не содержится ни в каком другом собственном идеале данного кольца.

Рассмотрим вопрос о максимальности идеалов в кольце $(\mathbf{Z}, +, \cdot)$, которое согласно теореме 4.2.1 является кольцом главных идеалов. Очевидно, что идеал (m) является собственным в $(\mathbf{Z}, +, \cdot)$ тогда и только тогда, когда $m > 1$.

Теорема 4.2.2. Идеал (m) максимален в $(\mathbf{Z}, +, \cdot)$ тогда и только тогда, когда m – простое число.

▷ Необходимость. Пусть (m) – максимальный идеал в $(\mathbf{Z}, +, \cdot)$. Если m – составное число, то $m = n \cdot q$, где $1 < n$, $q < m$, и $m \in (n)$, но $n \notin (m)$, поскольку $m \nmid n$. Следовательно, $(m) \subset (n) \neq \mathbf{Z}$, что противоречит максимальности идеала (m). Поэтому m – простое число.

Достаточность. Пусть m – простое число. Рассмотрим идеал (m) в $(\mathbf{Z}, +, \cdot)$. Если (m) – немаксимальный идеал, то $\exists\, n \in \mathbf{N}_{>1}$, такое, что $(m) \subset (n) \neq \mathbf{Z}$. Тогда $m = n \cdot q$ для некоторого $q \in \mathbf{N}_{>1}$, что противоречит тому, что число m простое. Значит, (m) – максимальный идеал в $(\mathbf{Z}, +, \cdot)$. ◁

§4.3. Кольцо полиномов от одной переменной над полем

Рассмотрим множество $P[x] = \left\{ f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \mid n \in \mathbf{Z}_{\ge 0},\ a_i \in P,\ i = \overline{0,n} \right\}$ – множество полиномов (многочленов) от одной переменной x с коэффициентами из поля P. $(P[x], +, \cdot)$ является кольцом с обычными операциями почленного сложения и умножения полиномов. Это ассоциативное и

коммутативное кольцо с единицей ($1 \in P$) и без делителей нуля, что следует из свойств умножения, правила вычисления старшего коэффициента в $P[x]$ и свойств поля P. Условимся степень многочлена $f(x)$ обозначать $\deg f$ (от англ. degree – «степень»). Тогда $\deg(f \cdot g) = \deg f + \deg g$ для всех $f(x), g(x) \in P[x] \setminus \{0\}$. Степень нулевого многочлена часто считают неопределенной, но иногда будем считать ее нулевой, как степень многочлена, являющегося элементом поля P ($0 \in P$). По своим свойствам кольцо $P[x]$ близко к кольцу целых чисел.

Теорема 4.3.1. Обратимыми в кольце $(P[x], +, \cdot)$ являются многочлены нулевой степени, отличные от нуля, и только они, т. е. $P[x]^* = P^*$.

▷ Очевидно, что $P^* \subseteq P[x]^*$. Если $f(x) \in P[x]^*$ и $\deg f \geq 1$, то при умножении на $f(x)$ любого многочлена, отличного от нуля, согласно правилу вычисления старшего коэффициента получим многочлен не ниже первой степени и, таким образом, не сможем получить $1 \in P$ – многочлен нулевой степени. Следовательно, $\deg f = 0$, значит, $f(x) \in P^*$. Поэтому $P[x]^* \subseteq P^*$. Итак, $P[x]^* = P^*$. ◁

Как и для **Z**, для кольца полиномов от одной переменной над полем имеют место следующие две теоремы.

Теорема 4.3.2 (о делении с остатком). Для любых двух многочленов $f(x)$ и $g(x) \neq 0$ из $P[x]$ существуют единственные многочлены $q(x)$ и $r(x)$ из $P[x]$, такие, что $r(x) = 0$ либо $0 \leq \deg r < \deg g$, и выполняется равенство

$$f(x) = g(x) \cdot q(x) + r(x). \qquad (4.3.1)$$

▷ Докажем существование пары многочленов $q(x)$ и $r(x)$, рассмотрев следующие случаи.

1. Если $f(x) = 0$, то, очевидно, $q(x) = 0$, $r(x) = 0$.

2. Если $\deg f < \deg g$, то $f(x) = 0 \cdot g(x) + f(x)$. Тогда $q(x) = 0$, $r(x) = f(x)$.

3. Если $\deg f \geq \deg g$, то для нахождения $q(x)$ и $r(x)$ используется, например, известный метод деления «уголком» $f(x)$ на $g(x)$, аналогичный методу деления целых чисел.

Теперь докажем единственность пары многочленов $q(x)$ и $r(x)$. Пусть $f(x) = g(x) \cdot q_1(x) + r_1(x) = g(x) \cdot q_2(x) + r_2(x)$, тогда $g(x) \cdot (q_1(x) - q_2(x)) = r_2(x) - r_1(x)$. $r_1(x) = r_2(x)$, т. к. иначе $r_1(x) \neq r_2(x)$ влечет $q_1(x) \neq q_2(x)$ и получаем противоречие с тем, что $0 \leq \deg(r_2 - r_1) < \deg g$. Отсюда следует, что $q_1(x) = q_2(x)$, поскольку $g(x) \neq 0$. ◁

Теорема 4.3.3. $(P[x], +, \cdot)$ – кольцо главных идеалов.

▷ В кольце $(P[x], +, \cdot)$, как в коммутативном кольце, все идеалы являются двусторонними и понятия главных левого, правого и двустороннего идеалов эквивалентны. Поскольку $(P[x], +, \cdot)$ – кольцо с единицей, все главные идеалы в нем имеют вид $(f(x)) = \{g(x) \cdot f(x) \mid g(x) \in P[x]\}$, где $f(x) \in P[x]$.

1. Если идеал $J = \{0\}$, то $\{0\} = (0) = \{g(x) \cdot 0 \mid g(x) \in P[x]\}$ – главный идеал.

2. Если идеал $J \neq \{0\}$, то в J найдем многочлен наименьшей степени, пусть это будет $d(x)$. Покажем, что $J = (d(x))$.

Очевидно, что $(d(x)) \subseteq J$ по определению идеалов J и $(d(x))$.

Пусть $f(x) \in J$, тогда согласно теореме 4.3.2 существует единственная пара $q(x)$ и $r(x)$ из $P[x]$, таких, что выполняется равенство $f(x) = q(x) \cdot d(x) + r(x)$, причем $r(x) = 0$ либо $0 \le \deg r < \deg d$. Значит, $r(x) = f(x) - q(x) \cdot d(x) \in J$. Если $r(x) \ne 0$, то $\deg r < \deg d$, а это противоречит тому, что $d(x)$ – многочлен минимальной степени в J. Поэтому $r(x) = 0$ и $f(x) \in (d(x))$, значит, $J \subseteq (d(x))$.

Итак, $J = (d(x))$. ◁

Определение 4.3.1. В равенстве (4.3.1) $q(x)$ называют *частным* (*неполным частным*, если $r(x) \ne 0$), а $r(x)$ – *остатком* от деления $f(x)$ на $g(x)$. Если $r(x) = 0$, то говорят, что $g(x)$, $q(x) \ne 0$ – *делители* (или *множители*) полинома $f(x)$. Также в этом случае говорят, что $g(x)$ и $q(x) \ne 0$ *делят* $f(x)$, и обозначают $g(x) \mid f(x)$, или, что $f(x)$ *кратно* (*делится* на) $g(x)$ и $q(x) \ne 0$, и обозначают $f(x) \vdots g(x)$.

Если $g(x) \mid f(x)$ и $0 < \deg g < \deg f$, то многочлен $g(x)$ называют *нетривиальным делителем* многочлена $f(x)$. Очевидно, произвольный элемент $\alpha \in P^*$ является делителем любого многочлена $f(x)$ из $P[x]$. При $f(x) \ne 0$ $\alpha^{-1} f(x)$ – также делитель $f(x)$. Поэтому такие делители называют *тривиальными*.

Будем обозначать $g(x) \nmid f(x)$, если $g(x) \ne 0$ не делит $f(x)$, и $f(x) \not\vdots g(x)$, если $f(x)$ не делится на $g(x) \ne 0$.

Для многочленов из $P[x]$ справедливы свойства делимости, аналогичные свойствам делимости в **Z**.

Свойства делимости многочленов

1. Для $\forall\, g(x) \ne 0$ $g(x) \mid 0$.

▷ $0 = g(x) \cdot 0$. ◁

2. Для $\forall\, \alpha \in P^*$ и $\forall\, f(x) \in P[x]$ $\alpha \mid f(x)$.

▷ $f(x) = \alpha \cdot (\alpha^{-1} \cdot f(x))$. ◁

3. $g(x) \mid f(x)$ & $f(x) \mid h(x) \Rightarrow g(x) \mid h(x)$ – свойство транзитивности.

▷ $h(x) = f(x)q(x) = g(x)u(x)q(x)$ для соответствующих $q(x)$, $u(x) \in P[x]$. ◁

4. $g(x) \mid f(x)$ & $f(x) \mid g(x) \Leftrightarrow f(x) = \alpha g(x)$ для некоторого $\alpha \in P^*$.

▷ $f(x) = g(x)q(x)$ & $g(x) = f(x)u(x) \Leftrightarrow f(x) = f(x)u(x)q(x) \Leftrightarrow q(x)u(x) = 1 \Leftrightarrow (q(x), u(x) \in P^*) \Leftrightarrow q(x) = \alpha$ & $u(x) = \alpha^{-1}$, где $\alpha \in P^*$. ◁

5. Если $g(x) \mid f_i(x)$, $i = \overline{1,k}$, то $g(x) \mid \sum_{i=1}^{k} u_i(x) f_i(x)$ для $\forall\, k \in \mathbf{N}$, $\forall\, u_i(x) \in P[x]$.

▷ $\sum_{i=1}^{k} u_i(x) f_i(x) = \sum_{i=1}^{k} u_i(x) q_i(x) g(x) = \left(\sum_{i=1}^{k} u_i(x) q_i(x) \right) g(x) = u(x) g(x)$, причем $u(x) \in P[x]$. ◁

6. Если в равенстве $f_1(x) + \ldots + f_k(x) = g_1(x) + \ldots + g_n(x)$, $\forall k$, $n \in \mathbf{N}$, все многочлены, кроме, быть может, одного делятся на $d(x)$, то и этот многочлен также делится на $d(x)$.

▷ Не ограничивая общности, можем считать, что этим многочленом является $f_k(x)$. Поскольку $g_i(x) = q_i(x) d(x)$, $q_i(x) \in P[x]$, $i = \overline{1,n}$, а также $f_j(x) = p_j(x) d(x)$, $p_j(x) \in P[x]$, $j = \overline{1,k-1}$, то справедливо равенство

$$f_k(x) = (q_1(x) + ... + q_n(x) - p_1(x) - ... - p_{k-1}(x))d(x) = u(x)d(x),$$

где $u(x) \in P[x]$, следовательно, $d(x) \mid f_k(x)$. ◁

7. $f(x) \mid g(x) \Leftrightarrow \alpha f(x) \mid \beta g(x)$ для $\forall\, \alpha, \beta \in P^*$.

▷ $g(x) = q(x)f(x)$ для соответствующего $q(x) \in P[x]$. Тогда для $\forall\, \alpha, \beta \in P^*$ последнее равенство равносильно $\beta g(x) = \beta\alpha^{-1}q(x)\alpha f(x) \Leftrightarrow \alpha f(x) \mid \beta g(x)$. ◁

Определение 4.3.2. *Общим делителем* (*ОД*) многочленов $f_1(x)$, $f_2(x)$,… …, $f_s(x) \in P[x]$, $\forall\, s \in \mathbf{N}_{\geq 2}$, называется многочлен $d(x) \in P[x]$, такой, что $d(x) \mid f_i(x)$, $i = \overline{1,s}$. *Наибольшим общим делителем* (*НОД*) многочленов $f_1(x)$, $f_2(x)$,…, $f_s(x)$, хотя бы один из которых отличен от нулевого, называется такой их нормированный общий делитель (со старшим коэффициентом 1), который делится на любой другой их общий делитель. Таким образом, нормированность гарантирует однозначность НОД. Его обозначают НОД$(f_1(x), f_2(x),\ldots, f_s(x))$ или, если это не вызывает разночтений, $(f_1(x), f_2(x),\ldots, f_s(x))$.

Очевидно, что если $g(x) \mid f(x)$, то $(g(x), f(x)) = b_m^{-1}g(x)$, где b_m – старший коэффициент $g(x)$, $\deg g = m$.

Аналогом теоремы 1.2.1 для $P[x]$ является следующая теорема.

Теорема 4.3.4. Если $f(x) = g(x)q(x) + r(x)$, то $(f(x), g(x)) = (g(x), r(x))$.

▷ Доказательство следует из определения НОД и свойств 5 и 6 делимости многочленов. ◁

Теорема 4.3.5. НОД многочленов $f(x)$ и $g(x)$ из кольца $P[x]$ при $g(x) \nmid f(x)$ и $f(x) \nmid g(x)$, $\deg g \leq \deg f$, совпадает с нормированным последним отличным от нуля остатком от деления $r_n(x)$ из следующей цепочки равенств:

$$\begin{aligned}
f(x) &= g(x)q_1(x) + r_1(x),\\
g(x) &= r_1(x)q_2(x) + r_2(x), \text{ если } r_1(x) \neq 0,\\
r_1(x) &= r_2(x)q_3(x) + r_3(x), \text{ если } r_2(x) \neq 0,\\
&\ldots\\
r_{n-2}(x) &= r_{n-1}(x)q_n(x) + r_n(x), \text{ если } r_{n-1}(x) \neq 0,\\
r_{n-1}(x) &= r_n(x)q_{n+1}(x), \text{ если } r_n(x) \neq 0.
\end{aligned} \qquad (4.3.2)$$

▷ Согласно теореме 4.3.4 имеем

$$(f(x), g(x)) = (g(x), r_1(x)) = (r_1(x), r_2(x)) = ... = (r_{n-1}(x), r_n(x)) = \alpha r_n(x),\ \alpha \in P^*,$$

где α является обратным элементом в P для старшего коэффициента многочлена $r_n(x)$, чтобы НОД был нормированным. Согласно теореме 4.3.2 $0 \leq \deg r_i < \deg r_{i-1}$, $i = \overline{1,n}$, где $r_0(x) = g(x)$. Процесс конечен, т. к. степень остатка постоянно уменьшается до 0. ◁

Теорема 4.3.5 представляет собой аналог для $P[x]$ алгоритма Евклида нахождения НОД целых чисел (теоремы 1.2.2).

НОД многочленов вычисляется рекурсивно для $\forall\, s \geq 3$:

$$(f_1(x), f_2(x),\ldots, f_s(x)) = ((f_1(x), f_2(x),\ldots, f_{s-1}(x)), f_s(x)).$$

Пример 4.3.1. Найдем наибольший общий делитель многочленов $f(x) = x^4 + 3x^3 - x^2 - 4x - 3$ и $g(x) = 3x^3 + 10x^2 + 2x - 3$ в кольце $\mathbf{Q}[x]$.

Используя, например, метод деления «уголком» для многочленов, получаем цепочку равенств согласно алгоритму Евклида:

$$f(x) = g(x)(1/3x - 1/9) - 5/9(x^2 + 5x + 6) = g(x)q_1(x) - 5/9\tilde{r}_1(x),$$
$$g(x) = \tilde{r}_1(x)(3x - 5) + 9(x + 3) = \tilde{r}_1(x)q_2(x) + 9\tilde{r}_2(x),$$
$$\tilde{r}_1(x) = \tilde{r}_2(x)(x + 2) = \tilde{r}_2(x)q_3(x).$$

Таким образом, $(f(x), g(x)) = \tilde{r}_2(x) = 1/9r_2(x) = x + 3$.●

Теорема 4.3.6. Если $d(x) = (f(x), g(x))$, то в $P[x]$ существуют многочлены $u(x)$ и $v(x)$, такие, что $d(x) = u(x)f(x) + v(x)g(x)$.

▷ Если $g(x) \mid f(x)$, то, как указывалось выше, $(f(x), g(x)) = b_m^{-1}g(x) = 0 \cdot f(x) + b_m^{-1}g(x)$, где b_m – старший коэффициент $g(x)$, $\deg g = m$. Если $f(x) \mid g(x)$, то по аналогии $(f(x), g(x)) = a_n^{-1}f(x) = a_n^{-1}f(x) + 0 \cdot g(x)$, где a_n – старший коэффициент $f(x)$, $\deg f = n$.

Пусть теперь $g(x) \nmid f(x)$, $f(x) \nmid g(x)$ и, не ограничивая общности, можно считать, что $\deg g \le \deg f$. Случай $\deg f \le \deg g$ рассматривается аналогично. Согласно теореме 4.3.5 и обратному применению цепочки равенств (4.3.2) имеем

$$d(x) = \alpha r_n(x) = \alpha r_{n-2}(x) - \alpha r_{n-1}(x)q_n(x) = \alpha r_{n-4}(x) - \alpha r_{n-3}(x)q_{n-2}(x) -$$
$$- \alpha q_n(x)(r_{n-3}(x) - r_{n-2}(x)q_{n-1}(x)) = \ldots = u(x)f(x) + v(x)g(x),$$

где $u(x)$, $v(x) \in P[x]$. Используя расширенный алгоритм Евклида, из цепочки равенств (4.3.2) получим в соотношении для $d(x)$, $f(x)$ и $g(x)$ те же многочлены $u(x)$ и $v(x)$:

$$r_1(x) = f(x) - g(x)q_1(x),$$
$$r_2(x) = g(x) - r_1(x)q_2(x) = g(x) - q_2(x)(f(x) - g(x)q_1(x)) =$$
$$= -q_2(x)f(x) + (1 + q_2(x)q_1(x))g(x),$$
$$\ldots$$
$$d(x) = \alpha r_n(x) = \alpha r_{n-2}(x) - \alpha r_{n-1}(x)q_n(x) = \ldots = u(x)f(x) + v(x)g(x). \triangleleft$$

Следствие. Если $d(x) = (f_1(x), f_2(x), \ldots, f_s(x))$, $s \ge 2$, то в $P[x]$ существуют многочлены $u_1(x)$, $u_2(x), \ldots, u_s(x)$, такие, что

$$d(x) = \sum_{i=1}^{s} u_i(x)f_i(x). \qquad (4.3.3)$$

▷ Доказательство проводится методом математической индукции с использованием теоремы 4.3.6 и рекурсивного вычисления НОД многочленов. ◁

Определение 4.3.3. Равенство (4.3.3) называют *соотношением Безу* для наибольшего общего делителя многочленов.

Замечание. Многочлены $u_1(x), u_2(x), \ldots, u_s(x) \in P[x]$ в (4.3.3) не являются единственными с таким условием. Так, например, легко видеть, что при $s = 2$, если $d(x) = u_1(x)f_1(x) + u_2(x)f_2(x)$, то для любого многочлена $q(x) \in P[x]$ многочлены из $P[x]$ $u_1(x) + q(x)f_2(x)/d(x)$ и $u_2(x) - q(x)f_1(x)/d(x)$ также являются коэффициентами данного соотношения Безу при $f_1(x)$ и $f_2(x)$ соответственно.

Пример 4.3.2. Найдем $u(x), v(x) \in \mathbf{Q}[x]$ в соотношении Безу для НОД $d(x) = x + 3$ многочленов $f(x)$ и $g(x)$ из примера 4.3.1.

$$d(x) = 1/9\,g(x) - 1/9\,\tilde{r}_1(x)(3x-5) = 1/9\,g(x) - 1/5(3x-5)\big(-f(x) +$$
$$+ g(x)(1/3x - 1/9)\big) = (3/5x - 1)f(x) + \big(-1/5x^2 + 2/5x\big)g(x).$$

Таким образом, $u(x) = 3/5x - 1$, $v(x) = -1/5x^2 + 2/5x$.

Согласно приведенному выше замечанию для всех $q(x) \in \mathbf{Q}[x]$ $u_q(x) = u(x) + q(x)g(x)/d(x)$ и $v_q(x) = v(x) - q(x)f(x)/d(x)$ также являются коэффициентами данного соотношения Безу при $f(x)$ и $g(x)$ соответственно. Например, при $q(x) = 1$ получаем $u_1(x) = u(x) + g(x)/d(x) = 3x^2 + 8/5x - 2$ и $v_1(x) = v(x) - f(x)/d(x) = -x^3 - 1/5x^2 + 7/5x + 1$.●

Определение 4.3.4. *Общим кратным* (*ОК*) ненулевых многочленов $f_1(x)$, $f_2(x),\ldots, f_s(x) \in P[x]$, $\forall\, s \in \mathbf{N}_{\geq 2}$, называется многочлен $m(x) \in P[x]$, такой, что $m(x) \vdots f_i(x)$, $i = \overline{1,s}$. *Наименьшим общим кратным* (*НОК*) ненулевых многочленов $f_1(x)$, $f_2(x),\ldots, f_s(x)$ называется такое их нормированное общее кратное (со старшим коэффициентом 1), на которое делится любое другое их общее кратное. Таким образом, нормированность гарантирует однозначность НОК. Его обозначают НОК$\big(f_1(x), f_2(x),\ldots, f_s(x)\big)$ или, если это не вызывает разночтений, $\big[f_1(x), f_2(x),\ldots, f_s(x)\big]$.

Очевидно, что если $g(x) \,|\, f(x)$, то при $f(x) \neq 0$ $\big[g(x), f(x)\big] = a_n^{-1} f(x)$, где a_n – старший коэффициент $f(x)$, $\deg f = n$.

НОК многочленов вычисляется рекурсивно для $\forall\, s \geq 3$:

$$\big[f_1(x), f_2(x),\ldots, f_s(x)\big] = \big[\big[f_1(x), f_2(x),\ldots, f_{s-1}(x)\big], f_s(x)\big].$$

Определение 4.3.5. Если $\big(f_1(x), f_2(x),\ldots, f_s(x)\big) = 1$, $s \geq 2$, то многочлены $f_1(x)$, $f_2(x),\ldots, f_s(x)$ называются *взаимно простыми*.

Теорема 4.3.7 (критерий взаимной простоты многочленов). Многочлены $f_1(x)$, $f_2(x),\ldots, f_s(x) \in P[x]$, $s \geq 2$, взаимно просты тогда и только тогда, когда найдутся такие многочлены $u_1(x)$, $u_2(x),\ldots, u_s(x) \in P[x]$, что $\sum\limits_{i=1}^{s} u_i(x) f_i(x) = 1$.

▷ Необходимость. Если $\big(f_1(x), f_2(x),\ldots, f_s(x)\big) = 1$, то по следствию из теоремы 4.3.6 существуют многочлены $u_1(x)$, $u_2(x),\ldots, u_s(x) \in P[x]$, такие, что $d(x) = 1 = \sum\limits_{i=1}^{s} u_i(x) f_i(x)$.

Достаточность. Если данное соотношение справедливо и $d(x) = \big(f_1(x), f_2(x),\ldots, f_s(x)\big)$, то по свойству 5 делимости многочленов $d(x) \,|\, 1$, следовательно, $\deg d = 0$. Так как старший коэффициент $d(x)$ равен 1, то $d(x) = 1$, что означает взаимную простоту многочленов $f_1(x), f_2(x),\ldots, f_s(x)$ по определению. ◁

Из этого критерия вытекает ряд важных следствий как свойств взаимно простых многочленов.

Следствие 1. Если произведение многочленов $f(x)$ и $g(x)$ делится на многочлен $h(x)$ и $\big(f(x), h(x)\big) = 1$, то $g(x)$ делится на $h(x)$.

$\triangleright\ \exists\ u(x), v(x) \in P[x]$, где $u(x)f(x) + v(x)h(x) = 1$. Умножим на $g(x)$ обе части данного равенства. Получим $u(x)f(x)g(x) + v(x)h(x)g(x) = g(x)$, откуда имеем

$$\big(u(x)q(x) + v(x)g(x)\big)h(x) = g(x),$$

поскольку $\exists\ q(x) \in P[x]$ с условием $f(x)g(x) = q(x)h(x)$. Значит, $h(x) \mid g(x). \triangleleft$

Следствие 2. Многочлен $f(x)$ взаимно прост с каждым из многочленов $g(x)$ и $h(x)$ тогда и только тогда, когда он взаимно прост и с их произведением.

$\triangleright$ Необходимость. Так как $\big(f(x), g(x)\big) = 1$ и $\big(f(x), h(x)\big) = 1$, то согласно теореме 4.3.7 $\exists\, u_1(x), v_1(x), u_2(x), v_2(x) \in P[x]$, такие, что выполняются равенства $u_1(x)f(x) + v_1(x)g(x) = 1$, $u_2(x)f(x) + v_2(x)h(x) = 1$. Тогда, перемножая почленно данные равенства, получим

$$\big(u_1(x)u_2(x)f(x) + u_1(x)v_2(x)h(x) + v_1(x)u_2(x)g(x)\big)f(x) + v_1(x)v_2(x)g(x)h(x) = 1,$$

т. е. выполняется критерий взаимной простоты для многочленов $f(x)$ и $g(x)h(x)$.

Достаточность. Пусть $\big(f(x), g(x)h(x)\big) = 1$. Если предположить, что $\big(f(x), g(x)\big) = d_1(x) \neq 1$ или $\big(f(x), h(x)\big) = d_2(x) \neq 1$, то согласно свойству 3 делимости многочленов и определению НОД получим, что $d_1(x) \mid \big(f(x), g(x)h(x)\big)$ либо $d_2(x) \mid \big(f(x), g(x)h(x)\big)$. Это противоречит тому, что $\big(f(x), g(x)h(x)\big) = 1$. Значит, $\big(f(x), g(x)\big) = 1$ и $\big(f(x), h(x)\big) = 1. \triangleleft$

§4.4. Неприводимость над полем и корни полиномов

Определение 4.4.1. Многочлен $f(x)$ степени $n \in \mathbf{N}$ называется *неприводимым в* $P[x]$ (*неприводимым над полем* P), если в любом его представлении вида $f(x) = g(x)q(x)$ для $g(x), q(x) \in P[x]$ его сомножители тривиальны, т. е. один из них является элементом из P^*. В противном случае многочлен $f(x)$ называется *приводимым в кольце* $P[x]$ (*приводимым над полем* P).

Очевидно, что все многочлены первой степени являются неприводимыми в $P[x]$ согласно данному определению.

Неприводимые многочлены в кольце $P[x]$ служат аналогом простых чисел в кольце $\mathbf{Z}$, что показывают следующие две важные теоремы, соответствующие основной теореме арифметики (теорема 1.4.2) и теореме Евклида о бесконечности множества всех простых чисел (теорема 1.3.2).

Теорема 4.4.1 (о разложении на множители). Всякий многочлен $f(x) \in P[x]$ степени $n \geq 1$ представим в виде

$$f(x) = a_n p_1(x) p_2(x) \ldots p_s(x), \qquad (4.4.1)$$

где $p_i(x),\ i = \overline{1, s}$, – неприводимые над P полиномы со старшими коэффициентами, равными 1;

a_n – старший коэффициент $f(x)$.

Такое представление единственно с точностью до порядка следования множителей.

▷ Пусть $f(x)$ – многочлен первой степени. Он является неприводимым согласно определению. Тогда $f(x) = a_1(a_1^{-1}f(x))$ и $a_1^{-1}f(x)$ – неприводимый многочлен первой степени со старшим коэффициентом, равным 1. Итак, $n = 1$ – база индукции.

Пусть утверждение о существовании такого разложения верно для любых натуральных $k < n$. Рассмотрим многочлен $f(x)$ n-й степени. В случае неприводимости $f(x)$ имеем $f(x) = a_n(a_n^{-1}f(x))$ и $a_n^{-1}f(x)$ – неприводимый многочлен n-й степени со старшим коэффициентом, равным 1. Если $f(x)$ – приводимый над полем P многочлен, то справедливо разложение $f(x) = g(x)h(x)$ для некоторых многочленов $g(x)$ и $h(x)$ из $P[x]$ с условием $1 \le \deg g, \deg h < n$. Тогда воспользуемся предположением индукции для многочленов $g(x)$ и $h(x)$ и получим требуемое разложение для $f(x)$, поскольку старший коэффициент $f(x)$ равен произведению старших коэффициентов $g(x)$ и $h(x)$. Итак, существование разложения (4.4.1) верно для произвольного многочлена $f(x) \in P[x]$ натуральной степени n.

Докажем теперь единственность разложения (4.4.1) с точностью до порядка сомножителей для произвольного многочлена $f(x) \in P[x]$ степени $n \in \mathbf{N}$. Пусть для $f(x)$ справедливы два разложения:

$$f(x) = a_n p_1(x)...p_s(x) = a_n q_1(x)...q_t(x).$$

Тогда согласно следствию 1 из теоремы 4.3.7 и неприводимости полиномов-сомножителей со старшими коэффициентами, равными 1, в обеих частях равенства получаем, что

$$p_i(x) \mid q_1(x)...q_t(x) \Rightarrow p_i(x) \mid q_j(x) \Rightarrow (p_i(x) = q_j(x),\ i = 1,...,s,\ j = 1,...,t) \Rightarrow s = t.$$

Таким образом, сомножители в обеих частях равенства совпадают. ◁

Определение 4.4.2. Если в разложении (4.4.1) объединить одинаковые сомножители-многочлены в степени, то получится разложение, называемое *каноническим разложением многочлена $f(x)$ над полем P*:

$$f(x) = a_n p_1^{\gamma_1}(x) p_2^{\gamma_2}(x)...p_s^{\gamma_s}(x),\ p_i(x) \ne p_j(x),\ i \ne j.$$

По каноническому разложению многочленов легко находится их наибольший общий делитель, наименьшее общее кратное, решаются иные задачи.

Упражнение 4.4.1. Получить формулы, аналогичные формулам из §1.4, для канонических разложений НОД и НОК по каноническим разложениям многочленов.

Теорема 4.4.2. В $P[x]$ существует бесконечно много неприводимых многочленов.

▷ Пусть $p_1(x),..., p_k(x),\ k \in \mathbf{N},$ – все неприводимые над P многочлены. Рассмотрим многочлен $p_1(x)...p_k(x) + 1 = f(x)$. Тогда $p_i(x) \nmid f(x),\ i = \overline{1,k}$, т. к. иначе $p_i(x) \mid 1$ согласно свойству 6 делимости многочленов, что невозможно. Значит, $f(x)$ либо сам неприводим над P, либо имеет делители, отличные от $p_i(x)$, которые неприводимы согласно теореме 4.4.1. ◁

Рассмотрим вопрос о максимальности идеалов в кольце $(P[x], +, \cdot)$, которое согласно теореме 4.3.3 является кольцом главных идеалов. Очевидно, что идеал $(f(x))$ является собственным в $(P[x], +, \cdot)$ тогда и только тогда, когда $\deg f \geq 1$.

Теорема 4.4.3. Идеал $(f(x))$ максимален в $(P[x], +, \cdot)$ тогда и только тогда, когда $f(x)$ неприводим над полем P.

▷ Необходимость. Пусть $(f(x))$ – максимальный идеал в $(P[x], +, \cdot)$. Если $f(x)$ приводим над P, то $f(x) = g(x) \cdot h(x)$, $g(x)$, $h(x) \in P[x]$, $1 \leq \deg g$, $\deg h < \deg f$. Тогда $f(x) \in (g(x))$, но $g(x) \notin (f(x))$, т. к. $f(x) \nmid g(x)$. Следовательно, $(f(x)) \subset \subset (g(x)) \neq P[x]$, что противоречит максимальности идеала $(f(x))$. Поэтому $f(x)$ – неприводимый над полем P многочлен.

Достаточность. Пусть $f(x)$ неприводим над полем P. Рассмотрим идеал $(f(x))$ в $(P[x], +, \cdot)$. Если $(f(x))$ – немаксимальный идеал, то $\exists\, g(x) \in P[x]$, такой, что $1 \leq \deg g < \deg f$ и $(f(x)) \subset (g(x)) \neq P[x]$. Тогда $f(x) = g(x) \cdot h(x)$ для некоторого $h(x) \in P[x]$, где $1 \leq \deg h < \deg f$, что противоречит тому, что многочлен $f(x)$ неприводим над P. Значит, $(f(x))$ – максимальный идеал $(P[x], +, \cdot)$. ◁

Данная теорема является аналогом теоремы 4.2.2 для кольца $(P[x], +, \cdot)$.

Определение 4.4.3. Элемент $c \in P$ или некоторого расширения поля P называется *корнем многочлена* $f(x) \in P[x]$, если $f(c) = 0$. Поле P называется *алгебраически замкнутым*, если всякий многочлен $f(x) \in P[x]$ при $\deg f \geq 1$ имеет в P корень.

Теорема 4.4.4 (Э. Безу). Для $\forall f(x) \in P[x]$, $\forall c \in P$ $f(c)$ равно остатку от деления $f(x)$ на $x - c$.

▷ По теореме 4.3.2 справедливо равенство $f(x) = q(x)(x - c) + r$, где $q(x) \in \in P[x]$, $r \in P$. Тогда $f(c) = q(c)(c - c) + r = r$.

Итак, $f(x) = q(x)(x - c) + f(c)$ для $\forall c \in P$. ◁

Следствие 1 (критерий корня). Элемент $c \in P$ является корнем многочлена $f(x) \in P[x]$ тогда и только тогда, когда $(x - c) \mid f(x)$.

▷ Необходимость. $f(c) = 0 \Rightarrow f(x) = q(x)(x - c) \Rightarrow (x - c) \mid f(x)$.

Достаточность. $(x - c) \mid f(x) \Rightarrow f(x) = q(x)(x - c) \Rightarrow f(c) = 0 \Rightarrow c$ – корень $f(x)$. ◁

Следствие 2. Неприводимый полином кольца $P[x]$ степени $n \geq 2$ не имеет корней в поле P.

▷ Если бы неприводимый полином кольца $P[x]$ степени $n \geq 2$ имел корень $c \in P$, то данный полином делился бы на полином первой степени $x - c$, что противоречит неприводимости в $P[x]$. ◁

Следствие 3. Определение алгебраически замкнутого поля P эквивалентно тому, что любой многочлен $f(x) \in P[x]$ при $\deg f \geq 1$ имеет все корни в P. Неприводимыми над алгебраически замкнутым полем являются многочлены первой степени, и только они.

▷ Необходимость. Пусть $\deg f = n \geq 1$, значит, $f(c_1) = 0$ для некоторого $c_1 \in P$. Тогда по следствию 1 из теоремы 4.4.4 $(x - c_1) \mid f(x)$. Разделим $f(x)$ на $x - c_1$, получим в частном $f_1(x) \in P[x]$, где $\deg f_1 = n - 1$. Отсюда следует, что $f(x)$ при-

водим над P при $n > 1$. Если $\deg f_1 \geq 1$, то $\exists\, c_2 \in P$, такой, что $f_1(c_2) = 0$, значит, и $f(c_2) = 0$, поскольку $f(x) = (x - c_1)f_1(x)$. Согласно следствию 1 из теоремы 4.4.4 $f(x) = (x - c_1)(x - c_2)f_2(x)$, где $f_2(x) \in P[x]$ и $\deg f_2 = n - 2$, и т. д. Продолжая процесс деления на линейные множители, на n-м шаге получим следующее разложение $f(x)$ над полем P:

$$f(x) = a_n(x - c_1)(x - c_2)\ldots(x - c_n), \qquad (4.4.2)$$

где a_n – старший коэффициент $f(x)$;

$c_1, c_2, \ldots, c_n$ – все корни $f(x)$, среди которых могут быть и равные.

Других корней ни в P, ни в каком его расширении многочлен $f(x)$ не имеет согласно следствию 1 из теоремы 4.4.4 и единственности разложения (4.4.2) с точностью до порядка сомножителей по теореме 4.4.1.

Достаточность очевидна. ◁

Определение 4.4.4. Пусть $c \in P$ – корень полинома $f(x)$. *Кратностью корня* c называется такое число $k \in \mathbf{N}$, что $(x - c)^k \mid f(x)$, а $(x - c)^{k+1} \nmid f(x)$. При $k = 1$ корень называется *простым*, при $k > 1$ – *кратным*.

Следствие 4. Если P – алгебраически замкнутое поле, то для любого многочлена $f(x) \in P[x]$ при $\deg f = n \in \mathbf{N}$ справедливо каноническое разложение:

$$f(x) = a_n(x - \alpha_1)^{k_1}\ldots(x - \alpha_s)^{k_s}, \qquad (4.4.3)$$

где a_n – старший коэффициент $f(x)$;

$\alpha_1, \ldots, \alpha_s$ – все различные корни $f(x)$ в P соответственно кратностей $k_1, \ldots, k_s$.

В этом случае

$$k_1 + \ldots + k_s = n,$$

т. е. число всех корней ненулевого многочлена в алгебраически замкнутом поле (с учетом их кратностей) равно степени данного многочлена.

▷ Разложение (4.4.3) получается, если объединить в правой части (4.4.2) одинаковые множители в степени. Приравнивая степени многочленов в обеих частях (4.4.3), приходим к равенству $k_1 + \ldots + k_s = n$. ◁

Теорема 4.4.5 (о числе корней многочлена). Пусть $f(x) \in P[x]$ и $\deg f = n \in \mathbf{N}$. Если $\alpha_1, \ldots, \alpha_s$ – различные корни $f(x)$ в поле P соответственно кратностей $k_1, \ldots, k_s$, где $s \in \mathbf{N}$, то $f(x)$ делится на произведение $(x - \alpha_1)^{k_1}\ldots(x - \alpha_s)^{k_s}$. При этом справедливо следующее неравенство:

$$k_1 + \ldots + k_s \leq n,$$

которое означает, что число корней ненулевого многочлена в поле (с учетом их кратностей) не превосходит степени данного многочлена.

▷ Все многочлены $(x - \alpha_i)^{k_i}$, $i = \overline{1, s}$, попарно взаимно просты для различных i и по следствию 1 из теоремы 4.4.4 каждый из них делит $f(x)$. Согласно следствиям 1 и 2 из теоремы 4.3.7 $(x - \alpha_1)^{k_1}\ldots(x - \alpha_s)^{k_s} \mid f(x)$ в $P[x]$. Поэтому $f(x) = g(x)u(x)$, где $g(x) = (x - \alpha_1)^{k_1}\ldots(x - \alpha_s)^{k_s}$ и $u(x) \in P[x]$. Так как $\deg u \geq 0$ и $\deg f = \deg g + \deg u$, то $k_1 + \ldots + k_s = \deg g \leq \deg f = n$. ◁

Удобной для деления многочлена $f(x) \in P[x]$, $\deg f = n \in \mathbf{N}$, на двучлен (бином) $x - c$, где $c \in P$, является *схема Горнера*, названная в честь английского ма-

тематика Вильямса Джорджа Горнера (1786–1837). Пусть $f(x) = g(x)(x-c) + b_0$, где $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0,\ a_n \neq 0,\ g(x) = b_n x^{n-1} + b_{n-1}x^{n-2} + \ldots + b_2 x + b_1$, $b_0 = f(c)$. Вычисление коэффициентов частного от деления $g(x)$ производится рекуррентно по следующим формулам:

$$b_n = a_n,$$
$$b_{n-1} = a_{n-1} + c \cdot b_n,$$
$$b_{n-2} = a_{n-2} + c \cdot b_{n-1},$$
$$\ldots$$
$$b_0 = a_0 + c \cdot b_1,$$

таким образом, $b_i = a_i + c \cdot b_{i+1},\ i = n-1, \ldots, 0$. Результаты вычислений по схеме Горнера обычно представляются в виде таблицы, изображенной на рис. 4.4.1, которая имеет более компактную форму по сравнению с традиционной записью метода деления на двучлен «уголком».

	a_n	a_{n-1}	a_{n-2}	…	a_1	a_0
c	a_n	b_{n-1}	b_{n-2}	…	b_1	b_0

Рис. 4.4.1

Схема Горнера представляет собой алгоритм вычисления значения многочлена $f(x)$ при заданном значении переменной $x = c$, а также позволяет осуществить проверку, является ли c корнем $f(x)$. С помощью схемы Горнера можно определять кратность корня многочлена $f(x)$, находить коэффициенты при разложении $f(x)$ по степеням $x - c$. В этих случаях частное $g(x)$ при $\deg g \geq 1$ делится на $x - c$ и аналогично заполняется следующая строка таблицы на рис. 4.4.1, но число столбцов становится на один крайний правый меньше, и т. д.

Важными частными случаями поля P являются числовые поля: **C**, **R** и **Q**. Отметим, что $(\mathbf{Q}[x], +, \cdot) < (\mathbf{R}[x], +, \cdot) < (\mathbf{C}[x], +, \cdot)$.

Теорема 4.4.6 (основная теорема алгебры). Поле комплексных чисел **C** алгебраически замкнуто.

Данная теорема приводится здесь без доказательства.

Следствие 1. Любой многочлен $f(x) \in \mathbf{C}[x]$ при $\deg f \geq 1$ имеет все корни в **C** и справедливо каноническое разложение (4.4.3). Неприводимыми над **C** являются многочлены первой степени, и только они.

▷ Доказательство вытекает из алгебраической замкнутости **C** и следствий 3 и 4 из теоремы 4.4.4. ◁

Следствие 2. Любой многочлен $f(x) \in \mathbf{R}[x]$ при $\deg f > 2$ приводи́м над **R**. Неприводимыми над **R** являются многочлены первой степени и второй степени с отрицательными дискриминантами, и только они.

▷ Если $\alpha \in \mathbf{C} \setminus \mathbf{R}$ и $f(\alpha) = 0$, то $f(\overline{\alpha}) = 0$ (здесь черта сверху обозначает комплексное сопряжение) для $f(x) \in \mathbf{R}[x]$ при $\deg f = n \geq 2$. Действительно, пусть $f(\alpha) = a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 = 0$. Тогда $\overline{f(\alpha)} = \overline{a_n\alpha^n + \ldots + a_1\alpha + a_0} =$

$= a_n\overline{\alpha}^n + ... + a_1\overline{\alpha} + a_0 = 0,$ т. к. $\overline{0} = 0,$ следовательно, $\overline{\alpha}$ – корень $f(x)$. Поэтому комплексных, но не действительных корней для $f(x) \in \mathbf{R}[x]$ всегда имеется пара. Отсюда по следствию 1 из теоремы 4.4.4 получаем, что такие корни α и $\overline{\alpha}$ имеют одинаковые кратности в разложении (4.4.3) $f(x)$ над $\mathbf{C}$. Пусть $\alpha = a + bi$, где $a, b \in \mathbf{R}$, $i^2 = -1$, тогда

$$(x-\alpha)(x-\overline{\alpha}) = (x-a-bi)(x-a+bi) = (x-a)^2 - (bi)^2 = (x-a)^2 + b^2 \in \mathbf{R}[x].$$

Значит, если $f(\alpha) = 0$ при $\alpha \in \mathbf{C} \setminus \mathbf{R}$, то $\left((x-a)^2 + b^2\right) | f(x)$. В случае же, когда все корни $f(x) \in \mathbf{R}[x] \setminus \{0\}$ действительные, справедливо разложение (4.4.3) над $\mathbf{R}$. Таким образом, если $\deg f > 2$, то $f(x)$ приводим над $\mathbf{R}$.

Если $\deg f = 1$, то $f(x)$, очевидно, неприводим над $\mathbf{R}$. Если $\deg f = 2$, то $f(x)$ неприводим над $\mathbf{R}$ в том и только том случае, если он не имеет действительных корней. А это возможно тогда и только тогда, когда дискриминант $f(x)$ отрицателен. $\triangleleft$

Любой многочлен из $\mathbf{Q}[x]$ домножением на НОК знаменателей его коэффициентов приводится к многочлену из $\mathbf{Z}[x]$. Следовательно, вопросы о существовании корня в $\mathbf{Q}$ и неприводимости многочлена в $\mathbf{Q}[x]$ сводятся к вопросам о существовании корня в $\mathbf{Q}$ и неприводимости полученного таким образом многочлена из $\mathbf{Z}[x]$ над $\mathbf{Q}$.

Остановимся сначала на вопросе о корнях в $\mathbf{Q}$ полиномов из $\mathbf{Z}[x]$. Если полином имеет вид $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_l x^l$, где $\deg f = n \geq 1$, $0 < l \leq n$, $a_l \neq 0$, то, очевидно, 0 является его корнем кратности l. В таком случае можно разделить $f(x)$ на x^l и рассматривать вопрос о корнях полученного частного, если его степень ненулевая.

Теорема 4.4.7. Пусть $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \mathbf{Z}[x]$, где $n \geq 1$, $a_n \neq 0$, $a_0 \neq 0$. Тогда если $m/k \in \mathbf{Q}$, где $m \in \mathbf{Z}$, $k \in \mathbf{N}$, $(m,k) = 1$, является корнем $f(x)$, то $m \mid a_0$ и $k \mid a_n$.

$\triangleright$ Так как $f(m/k) = 0 \Leftrightarrow a_n(m/k)^n + ... + a_1(m/k) + a_0 = 0$, то, умножая обе части последнего равенства на k^n, получим $a_n m^n + a_{n-1}m^{n-1}k + ... + a_1 m k^{n-1} + a_0 k^n = 0$. Тогда по свойству 6 делимости целых чисел $m \mid a_0 k^n$ и $k \mid a_n m^n$. Поскольку $(m,k) = 1$, по свойству 3 взаимно простых чисел $(m, k^n) = (m^n, k) = 1$, тогда по свойству 2 взаимно простых чисел $m \mid a_0$ и $k \mid a_n$. $\triangleleft$

Следствие. Если $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \mathbf{Z}[x]$, где $n \geq 1$, $a_0 \neq 0$, то все рациональные корни $f(x)$ являются целочисленными и находятся среди делителей a_0.

$\triangleright$ В этом случае если $m/k \in \mathbf{Q}$, где $m \in \mathbf{Z}$, $k \in \mathbf{N}$, $(m,k) = 1$, является корнем $f(x)$, то согласно теореме 4.4.7 $m \mid a_0$ и $k \mid 1$, т. е. $k = 1$ и $m/k = m \in \mathbf{Z}$. $\triangleleft$

Замечание. Теорема 4.4.7 дает только необходимое, но не достаточное условие существования корня в $\mathbf{Q}$ у многочлена из $\mathbf{Z}[x]$ с $a_0 \neq 0$. Но она позволяет непосредственной проверкой найти все корни многочлена из $\mathbf{Z}[x]$ в $\mathbf{Q}$ или доказать, что их нет.

Пример 4.4.1. Многочлен $f(x) = x^2 - 3$ не имеет корней в **Q** согласно следствию из теоремы 4.4.7, в чем можно легко убедиться непосредственной подстановкой вместо x ± 1 и ± 3, являющихся делителями $a_0 = -3$. Поэтому $f(x)$ неприводим над **Q**. Но $f(x)$ приводи́м над **R** согласно следствию 2 из теоремы 4.4.6, как имеющий дискриминант $D = 12 > 0$. $\pm\sqrt{3}$ – корни $f(x)$ в **R**, поэтому $f(x) = \left(x - \sqrt{3}\right)\left(x + \sqrt{3}\right)$ – разложение вида (4.4.3) над **R**.•

Упражнение 4.4.2. Убедиться в том, что многочлен $f(x) = 8x^5 - 14x^4 - 77x^3 + 128x^2 + 45x - 18$ имеет все пять простых корней в **Q**, используя теорему 4.4.7.

Ответ. Корнями $f(x)$ являются $2, \pm 3, -1/2, 1/4$.

Рассмотрим теперь один достаточный признак неприводимости многочлена с целочисленными коэффициентами над полем **Q**. Данный признак является частным и наиболее распространенным случаем признака Эйзенштейна, названного в честь Фердинанда Эйзенштейна (1823–1852) – немецкого математика и философа.

Отметим предварительно следующее важное свойство, которое вытекает из свойства 7 делимости многочленов и свойств взаимно простых чисел в **Z**: если $f(x) \in \mathbf{Z}[x]$ допускает разложение на множители из $\mathbf{Q}[x]$, то $f(x)$ допускает разложение на множители из $\mathbf{Z}[x]$. Отсюда следует, что $f(x) \in \mathbf{Z}[x]$ неприводим над **Q** тогда и только тогда, когда $f(x)$ неприводим над **Z**, т. е. не допускает разложения на нетривиальные множители с коэффициентами из **Z**.

Теорема 4.4.8 (признак Эйзенштейна). Пусть $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathbf{Z}[x]$, где $n > 1$, $a_n \neq 0$. Если можно подобрать простое число p, удовлетворяющее условиям:

1) $p \nmid a_n$,

2) $p \mid a_i$, $i = \overline{0, n-1}$,

3) $p^2 \nmid a_0$,

то $f(x)$ неприводим над **Q**.

▷ Пусть простое число p, удовлетворяющее данным условиям, существует. Допустим, $f(x)$ приводи́м над **Q**: $f(x) = (b_m x^m + \ldots + b_1 x + b_0)(c_k x^k + \ldots + c_1 x + c_0)$ – разложение $f(x)$ на нетривиальные множители с коэффициентами из **Z**. Тогда $a_0 = b_0 c_0$ & $p \mid a_0 \Rightarrow$ пусть $p \mid b_0$ согласно свойству 4 взаимно простых чисел, но $p^2 \nmid a_0 \Rightarrow (p, c_0) = 1$ согласно свойству 1 простых чисел. Из $a_1 = b_1 c_0 + b_0 c_1$ & & $p \mid a_1 \Rightarrow p \mid b_1 c_0$ по свойству 6 делимости целых чисел. Так как $(p, c_0) = 1$, то $p \mid b_1$ по свойству 2 взаимно простых чисел. Далее, $a_2 = b_2 c_0 + b_1 c_1 + b_0 c_2 \Rightarrow$ $\Rightarrow p \mid b_2, \ldots, p \mid b_m$ по аналогичным рассуждениям. Но $a_n = b_m c_k \Rightarrow p \mid a_n$ по свойству 3 делимости целых чисел, чего не может быть по условию. Полученное противоречие доказывает, что $f(x)$ неприводим над **Q**. ◁

Замечание. Выполнение условий признака Эйзенштейна достаточно, но не необходимо для неприводимости многочлена $f(x) \in \mathbf{Z}[x]$ при $\deg f \geq 2$ над **Q**.

Пример 4.4.2.

1. Для $f(x) = x^3 + 1$ условия теоремы 4.4.8 не выполняются и $f(x)$ приводи́м над **Q**. $f(x) = (x+1)(x^2 - x + 1)$ – разложение вида (4.4.1) на неприводимые множители над **Q**.

2. Для $f(x) = x^2 + 1$ условия теоремы 4.4.8 также не выполняются, но $f(x)$ неприводим над **Q**. $f(x)$ неприводим над **R** согласно следствию 2 из теоремы 4.4.6, как имеющий дискриминант $D = -4 < 0$. $f(x)$ приводи́м только над **C** согласно следствию 1 из теоремы 4.4.6, поскольку $\deg f = 2$. Так как $\pm i$, где $i^2 = -1$, – корни $f(x)$ в **C**, то $f(x) = (x - i)(x + i)$ – разложение вида (4.4.1) над **C**, которое единственно в силу теоремы 4.4.1.●

Структура неприводимых многочленов существенно зависит от поля P. Так, неприводимыми в кольце $\mathbf{C}[x]$ являются только многочлены первой степени согласно следствию 1 из основной теоремы алгебры (теорема 4.4.6). Отсюда согласно следствию 2 из теоремы 4.4.6 вытекает, что в $\mathbf{R}[x]$ неприводимыми являются лишь многочлены первой степени, а также второй степени с отрицательными дискриминантами. Что касается кольца $\mathbf{Q}[x]$, то здесь для каждого натурального n существует бесконечно много неприводимых многочленов степени n. К примеру, по признаку Эйзенштейна таковыми являются многочлены $x^n \pm p$, где $n > 1$, а p – простое число. Как известно, простых чисел бесконечно много (теорема 1.3.2 Евклида).

Важными частными случаями поля P являются также конечные поля. Конечные поля получили широкое применение в современной теории кодирования и занимают заметное место в современной математике. Это основной инструмент при построении помехоустойчивых кодов и дискретных сигналов, алгоритмов их обработки, многих современных криптосистем. Конечные поля часто называют *полями Галуа* в честь выдающегося французского математика, основоположника современной алгебры Эвариста Галуа (1811–1832). Как уже отмечалось в п. 2 примера 4.1.3, для любого простого числа p $(\mathbf{Z}/p\mathbf{Z}, +, \cdot)$ – конечное поле из p элементов. В дальнейшем для конечного поля, состоящего из q элементов, будет использоваться обозначение $\boldsymbol{F}_q$. Отметим без доказательства два следующих факта.

1. Поле $\boldsymbol{F}_q$ существует тогда и только тогда, когда $q = p^n$, где p – простое число, $n \in \mathbf{N}$.

2. В кольце $\boldsymbol{F}_q[x]$ существуют неприводимые многочлены любой натуральной степени n.

В $\boldsymbol{F}_q[x]$ множество всех многочленов фиксированной степени n имеет лишь конечную мощность, равную $(q-1)q^n$, как будет показано в §4.5, поэтому и число неприводимых многочленов данной степени всегда конечно, в отличие, например, от $\mathbf{Q}[x]$.

Согласно определению 4.4.1, теореме 4.4.1 и следствиям 1 и 2 из теоремы 4.4.4, неприводимыми в $P[x]$ являются все многочлены первой степени и все многочлены второй и третьей степеней, не имеющие корней в поле P, т. е. не имеющие делителей первой степени в $P[x]$. Неприводимыми многочленами

четвертой степени в $P[x]$ являются все, не имеющие корней в P, а также не кратные нормированным неприводимым многочленам второй степени из $P[x]$, и т. д. В общем случае неприводимыми над P многочленами n-й степени являются все, не имеющие в $P[x]$ нормированных неприводимых делителей степеней k, где $1 \le k \le [n/2]$. Данная процедура «просеивания» для определения неприводимых многочленов напоминает «решето» Эратосфена для проверки натуральных чисел на простоту. Она удобна для поиска неприводимых многочленов и разложения многочленов на неприводимые множители над конечными полями.

Пример 4.4.3. Используя аналог «решета» Эратосфена, приведем список всех неприводимых многочленов в кольце $\mathbf{Z}/2\mathbf{Z}[x]$ степеней, меньших шести:

1) x;
2) $x+\bar{1}$;
3) $x^2+x+\bar{1}$;
4) $x^3+x+\bar{1}$;
5) $x^3+x^2+\bar{1}$;
6) $x^4+x+\bar{1}$;
7) $x^4+x^3+\bar{1}$;
8) $x^4+x^3+x^2+x+\bar{1}$;
9) $x^5+x^2+\bar{1}$;
10) $x^5+x^3+\bar{1}$;
11) $x^5+x^3+x^2+x+\bar{1}$;
12) $x^5+x^4+x^2+x+\bar{1}$;
13) $x^5+x^4+x^3+x+\bar{1}$;
14) $x^5+x^4+x^3+x^2+\bar{1}$.●

§4.5. Факторкольца

Пусть $(K, +, \cdot)$ – кольцо с двусторонним идеалом I, при этом I является нормальной подгруппой аддитивной абелевой группы $(K, +)$. На K может быть определено отношение сравнения по модулю I.

Определение 4.5.1. Элементы $a, b \in K$ называются *сравнимыми по модулю идеала* I, если $a - b \in I$, т. е. $a = b + i$ для некоторого $i \in I$.

Это отношение является отношением эквивалентности (рефлексивно, симметрично, транзитивно) и, следовательно, разбивает K на непересекающиеся классы эквивалентности сравнимых по модулю I элементов – *классы вычетов* по модулю идеала I. Для $\forall\, a \in K$ *класс вычетов с представителем* a имеет вид $\bar{a} = a + I = \{a + i \mid i \in I\}$. Множество всех различных классов вычетов $K/I = \{\bar{0} = I, \bar{a} = a + I, \bar{b} = b + I, \ldots\}$ называется *факторможеством* кольца K по идеалу I.

Теорема 4.5.1. Факторможество K/I кольца $(K, +, \cdot)$ по двустороннему идеалу I является кольцом относительно индуцированных операций $\oplus$ и $\otimes$ классов вычетов по модулю идеала I:

$$\bar{a} \oplus \bar{b} = \overline{a+b},$$
$$\bar{a} \otimes \bar{b} = \overline{a \cdot b}$$

для $\forall\, \bar{a}, \bar{b} \in K / I$.

▷ Введем бинарную индуцированную операцию $\oplus$ на K/I: $\bar{a} \oplus \bar{b} = \overline{a+b}$. Тогда результат данной операции не зависит от выбора представителей классов вычетов по модулю идеала I:

$$\overline{a+i_1} \oplus \overline{b+i_2} = \overline{a+b+(i_1+i_2)} = \overline{a+b},\ \forall i_1,\ i_2 \in I,$$

поскольку сложение в $(K, +, \cdot)$ ассоциативно, коммутативно и $i_1 + i_2 \in I$.

Введем также бинарную индуцированную операцию $\otimes$ на K/I: $\bar{a} \otimes \bar{b} = \overline{a \cdot b}$. Тогда результат данной операции не зависит от выбора представителей классов вычетов по модулю идеала I:

$$\overline{a+i_1} \otimes \overline{b+i_2} = \overline{a \cdot b + a \cdot i_2 + i_1 \cdot b + i_1 \cdot i_2} = \overline{a \cdot b},\ \forall i_1,\ i_2 \in I,$$

поскольку умножение дистрибутивно относительно сложения в $(K, +, \cdot)$ и $a \cdot i_2 + i_1 \cdot b + i_1 \cdot i_2 \in I$.

$(I, +) \lhd (K, +) \Rightarrow (K/I, \oplus)$ – абелева группа, как факторгруппа абелевой группы. Можно отметить, что $\bar{0} = I$.

$(\bar{a} \oplus \bar{b}) \otimes \bar{c} = \overline{a+b} \otimes \bar{c} = \overline{(a+b) \cdot c} = \overline{a \cdot c + b \cdot c}$, т. к. умножение в $(K, +, \cdot)$ дистрибутивно относительно сложения. $\bar{a} \otimes \bar{c} \oplus \bar{b} \otimes \bar{c} = \overline{a \cdot c} \oplus \overline{b \cdot c} = \overline{a \cdot c + b \cdot c}$. То есть для $\forall \bar{a},\ \bar{b},\ \bar{c} \in K/I$ выполняется равенство

$$(\bar{a} \oplus \bar{b}) \otimes \bar{c} = \bar{a} \otimes \bar{c} \oplus \bar{b} \otimes \bar{c}.$$

Аналогично можно показать, что для $\forall \bar{a},\ \bar{b},\ \bar{c} \in K/I$ выполняется второй закон дистрибутивности умножения относительно сложения:

$$\bar{c} \otimes (\bar{a} \oplus \bar{b}) = \bar{c} \otimes \bar{a} \oplus \bar{c} \otimes \bar{b}.$$

Таким образом, $(K/I, \oplus, \otimes)$ – кольцо. ◁

Определение 4.5.2. Кольцо $(K/I, \oplus, \otimes)$ классов вычетов по модулю двустороннего идеала I кольца $(K, +, \cdot)$ с индуцированными операциями сложения и умножения называется *факторкольцом* кольца K по идеалу I.

Поскольку операции $\oplus$ и $\otimes$ полностью определяются сложением и умножением в кольце $(K, +, \cdot)$, то свойства этих операций переносятся на $(K/I, \oplus, \otimes)$. Если K – ассоциативное кольцо, кольцо с единицей, коммутативное кольцо, то и $(K/I, \oplus, \otimes)$ также будет соответственно ассоциативным кольцом, кольцом с единицей $\bar{1}$, коммутативным кольцом.

Пример 4.5.1. Кольцо классов вычетов $(\mathbf{Z}/n\mathbf{Z}, +, \cdot)$ можно рассматривать как факторкольцо $(\mathbf{Z}/n\mathbf{Z}, \oplus, \otimes)$ кольца $(\mathbf{Z}, +, \cdot)$ по идеалу $I = n\mathbf{Z} = (n)$, где $n \in \mathbf{Z}_{\geq 0}$. Сложение и умножение в данном факторкольце при $n \geq 1$ могут быть заданы таблицами Кэли, об этом уже говорилось в §1.7 и в §3.6.●

Пример 4.5.2. Возьмем в кольце $P[x]$ произвольный собственный идеал $I = (f(x))$ для некоторого полинома $f(x)$ степени $n \geq 1$. Согласно теореме 4.3.2 в один класс вычетов по модулю I попадают те и только те полиномы, которые имеют один и тот же остаток $r(x)$ от деления на $f(x)$. Следовательно, фактормножество $P[x]/I$ состоит из классов $\overline{r(x)} = r(x) + I = \{r(x) + f(x)q(x) \mid q(x) \in P[x]\}$, где $r(x) = 0$ либо $0 \leq \deg r < \deg f$. По теореме 4.5.1 результатом сложения или умно-

жения таких классов вычетов является соответственно класс, представитель которого есть остаток от деления на $f(x)$ суммы или произведения представителей данных классов в $P[x]$.

Из сказанного выше следует, в частности, что факторможество $\boldsymbol{F}_q[x]/(f(x))$ конечно. Поэтому сложение и умножение в факторкольце $(\boldsymbol{F}_q[x]/(f(x)), \oplus, \otimes)$ можно задать в виде таблиц Кэли.●

Пример 4.5.3. Следуя рассуждениям из примера 4.5.2, получаем, что ассоциативное и коммутативное кольцо с единицей $\boldsymbol{F}_2[x]/(x^2+x+1)$ состоит из классов вычетов $\bar{0}$, $\bar{1}$, $\bar{x}$, $\overline{x+1}$ по модулю главного идеала (x^2+x+1) кольца $\boldsymbol{F}_2[x]$. Здесь поле $\boldsymbol{F}_2$ состоит из двух элементов: 0 и 1, причем $1+1=0$. Построим таблицы Кэли сложения и умножения в этом факторкольце (см. рис. 4.5.1, а и 4.5.1, б соответственно).

$\oplus$	$\bar{0}$	$\bar{1}$	$\bar{x}$	$\overline{x+1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{x}$	$\overline{x+1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$	$\overline{x+1}$	$\bar{x}$
$\bar{x}$	$\bar{x}$	$\overline{x+1}$	$\bar{0}$	$\bar{1}$
$\overline{x+1}$	$\overline{x+1}$	$\bar{x}$	$\bar{1}$	$\bar{0}$

а

$\otimes$	$\bar{0}$	$\bar{1}$	$\bar{x}$	$\overline{x+1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{x}$	$\overline{x+1}$
$\bar{x}$	$\bar{0}$	$\bar{x}$	$\overline{x+1}$	$\bar{1}$
$\overline{x+1}$	$\bar{0}$	$\overline{x+1}$	$\bar{1}$	$\bar{x}$

б

Рис. 4.5.1

Из таблицы умножения на рис. 4.5.1, б следует, что в факторкольце $\boldsymbol{F}_2[x]/(x^2+x+1)$ все ненулевые классы вычетов по модулю идеала (x^2+x+1) обратимы. Следовательно, $(\boldsymbol{F}_2[x]/(x^2+x+1), \oplus, \otimes)$ – поле из четырех элементов, являющееся расширением поля $(\overline{\boldsymbol{F}_2}, \oplus, \otimes)$, где $\overline{\boldsymbol{F}_2} = \{\bar{0}, \bar{1}\}$. Здесь $\bar{0}$ и $\bar{1}$ обозначают соответственно классы вычетов с представителями 0 и 1 по модулю идеала (x^2+x+1) кольца $\boldsymbol{F}_2[x]$.●

Лемма 4.5.1. Мощность фактормножества $\boldsymbol{F}_q[x]/(f(x))$ при $\deg f = n \geq 1$ равна q^n.

▷ О структуре $\boldsymbol{F}_q[x]/(f(x))$ уже говорилось в примере 4.5.2. Поэтому мощность $\boldsymbol{F}_q[x]/(f(x))$ равна количеству всех многочленов из $\boldsymbol{F}_q[x]$, степени которых меньше n. Это количество равно

$$q + q(q-1) + q^2(q-1) + \ldots + q^{n-1}(q-1) = (q-1)(1+q+\ldots+q^{n-1}) + 1 = q^n,$$

поскольку количество многочленов 0-й степени равно $|\boldsymbol{F}_q| = q$, а количество многочленов i-й степени равно $q^i(q-1)$, где $i = \overline{1, n-1}$. Старшие коэффициенты многочлена степени $i \geq 1$ принимают $q-1$ значений (отличны от 0), остальные коэффициенты принимают q различных значений из поля $\boldsymbol{F}_q$. ◁

В частности, число элементов в $\boldsymbol{F}_2[x]/(f(x))$ при $\deg f = n \geq 1$ равно 2^n. В примере 4.5.3 $n = 2$, поэтому поле $(\boldsymbol{F}_2[x]/(x^2+x+1), \oplus, \otimes)$ состоит из четырех элементов.

Важным критерием поля для факторкольца является следующая теорема.

Теорема 4.5.2. Факторкольцо ассоциативного и коммутативного кольца с единицей $(K, +, \cdot)$ по собственному идеалу I является полем тогда и только тогда, когда I – максимальный идеал.

▷ Необходимость. Пусть $(K/I, \oplus, \otimes)$ – поле и допустим, что I – немаксимальный собственный идеал кольца $(K, +, \cdot)$. Тогда $I \subset J \subset K$ для некоторого собственного идеала J. Рассмотрим множество $\overline{J} = \{\overline{j} \in K/I \mid j \in J\} \subset K/I$. Несложно проверить, что $\overline{J} \lhd K/I$, поскольку $J \lhd K$. Но $\overline{J} \neq \overline{0}$, т. к. $I \subset J$, также $\overline{J} \neq K/I$, т. к. $1 \notin J$ и, следовательно, $\overline{1} \notin \overline{J}$. Поэтому $\overline{J}$ – собственный идеал поля $(K/I, \oplus, \otimes)$. С другой стороны, поле не может иметь собственных идеалов. Полученное противоречие доказывает, что I – максимальный идеал кольца $(K, +, \cdot)$.

Достаточность. Факторкольцо ассоциативного и коммутативного кольца с единицей $(K, +, \cdot)$ по идеалу I является ассоциативным и коммутативным кольцом с единицей $\overline{1}$ согласно свойствам факторколец. Пусть I – максимальный идеал кольца K. Значит, $\overline{1} \neq \overline{0}$, поскольку $\overline{0} = I$, $\overline{1} = 1 + I$, $I \neq K$, $1 \notin I$, и $|K/I| > 1$. Рассмотрим произвольный ненулевой класс вычетов $\overline{x} \in K/I$, т. е. его представитель $x \notin I$. Подмножество $I + Kx = \{i + k \cdot x \mid i \in I, k \in K\}$ является идеалом в K, что несложно проверить, и $I \subset I + Kx$, а поскольку I – максимальный идеал, то $I + Kx = K$. Значит, справедливо представление $i + k \cdot x = 1$ для некоторых $i \in I$, $k \in K \setminus I$. Данное представление влечет равенство $\overline{k} \otimes \overline{x} = \overline{1}$ в факторкольце K/I, что означает обратимость класса $\overline{x}$ в K/I. Итак, $(K/I, \oplus, \otimes)$ – поле. ◁

Приведем важные следствия из теоремы 4.5.2 для колец $\mathbf{Z}$ и $P[x]$.

Следствие 1. Факторкольцо $(\mathbf{Z}/n\mathbf{Z}, \oplus, \otimes)$ при $n > 1$ является полем тогда и только тогда, когда n – простое число.

▷ Доказательство вытекает из теорем 4.5.2 и 4.2.2. ◁

Следствие 2. Факторкольцо $(P[x]/(f(x)), \oplus, \otimes)$ при $\deg f \geq 1$ является полем тогда и только тогда, когда $f(x)$ – неприводимый полином над полем P.

▷ Доказательство вытекает из теорем 4.5.2 и 4.4.3. ◁

В рассмотренном выше примере 4.5.3 полином $x^2 + x + 1$ неприводим над $\boldsymbol{F_2}$ (поскольку он не имеет корней в $\boldsymbol{F_2}$), поэтому $(\boldsymbol{F_2}[x]/(x^2 + x + 1), \oplus, \otimes)$ – поле. Данное следствие 2 позволяет с использованием неприводимого в $P[x]$ полинома $f(x)$ степени $n \geq 1$ построить поле $(P[x]/(f(x)), \oplus, \otimes)$, являющееся расширением поля $(\overline{P}, \oplus, \otimes)$, где $\overline{P} = \{\overline{\alpha} \mid \alpha \in P\}$. Очевидно, что $P \leftrightarrow \overline{P}$ и $P[x]/(f(x)) \leftrightarrow P[x]/(g(x))$, если $f(x)$ и $g(x)$ – два неприводимых многочлена одной степени n. В случае когда $P = \boldsymbol{F_q}$, поле $(\boldsymbol{F_q}[x]/(f(x)), \oplus, \otimes)$ является конечным, состоящим из q^n элементов согласно лемме 4.5.1. Так, с использованием простейших полей $\boldsymbol{F_p}$, например, $(\mathbf{Z}/p\mathbf{Z}, +, \cdot)$, где p – простые числа, и неприводимых над $\boldsymbol{F_p}$ многочленов можно построить конечные поля любых заданных порядков $q = p^m$, $m \in \mathbf{N}$ (как отмечалось в конце §4.4).

§4.6. Гомоморфизмы колец

Определение 4.6.1. Пусть $(K_1, +, \cdot)$ и $(K_2, \tilde{+}, \bullet)$ – два кольца. Всякая функция $f: K_1 \to K_2$, обладающая свойствами:

1) $f(a+b) = f(a) \tilde{+} f(b)$,

2) $f(a \cdot b) = f(a) \bullet f(b)$

для $\forall\, a, b \in K_1$, называется *гомоморфизмом колец*.

Определение 4.6.2. Инъективный гомоморфизм колец называется *мономорфизмом* (или *вложением*), сюръективный – *эпиморфизмом*, биективный – *изоморфизмом*. Всякий гомоморфизм $f: K \to K$ называют *эндоморфизмом* кольца K, а взаимно однозначный эндоморфизм – *автоморфизмом*.

Изоморфные кольца будем обозначать $(K_1, +, \cdot) \cong (K_2, \tilde{+}, \bullet)$ или $K_1 \cong K_2$.

Несложно видеть, что отношение изоморфизма на множестве всех колец является отношением эквивалентности: оно рефлексивно, симметрично и транзитивно. Таким образом, все множество колец разбивается на непересекающиеся классы попарно изоморфных объектов. Важной задачей теории колец является классификация всех колец с точностью до изоморфизма.

Пример 4.6.1.

1. Аннулятор кольца K: $f(a) = 0$ для всякого $a \in K$ – тривиальный пример эндоморфизма произвольного кольца K.

2. Другой тривиальный пример эндоморфизма, являющегося автоморфизмом, – тождественное преобразование e_K любого кольца K.•

Пример 4.6.2. Пусть $\mathbf{C}$ – поле комплексных чисел, $f: \mathbf{C} \to \mathbf{C}$. Для каждого $z = a + bi$ из $\mathbf{C}$, где $a, b \in \mathbf{R}$, $i^2 = -1$, положим $f(z) = a - bi = \bar{z}$. Очевидно, операция комплексного сопряжения взаимно однозначна и обладает свойствами 1 и 2 из определения гомоморфизма. Следовательно, f – автоморфизм поля комплексных чисел, причем $f^2 = e_{\mathbf{C}}$.•

Определение 4.6.3. Пусть $f: K_1 \to K_2$ – гомоморфизм колец. *Образом гомоморфизма* $f: K_1 \to K_2$ называется множество $\operatorname{Im} f = \{f(a) \in K_2 \mid a \in K_1\} = E(f)$ (иногда еще обозначается $\operatorname{Im} K_1$). *Ядром гомоморфизма* f называется множество $\operatorname{Ker} f = \{a \in K_1 \mid f(a) = 0_{K_2}\}$, где 0_{K_2} – нейтральный элемент группы $(K_2, \tilde{+})$.

Очевидно, что когда $f: K_1 \to K_2$ – мономорфизм колец, $K_1 \cong \operatorname{Im} f$.

Пример 4.6.3. Пусть $(P, +, \cdot)$ – произвольное поле и $g(x)$ – неприводимый над P многочлен степени n из $P[x]$. Тогда по следствию 2 из теоремы 4.5.2 факторкольцо $(P[x]/(g(x)), \oplus, \otimes)$ является полем. Несложно проверить, что функция $f: P \to P[x]/(g(x))$, где $f(\alpha) = (\overline{\alpha})$ для всех $\alpha \in P$, – мономорфизм данных полей. Следуя рассуждениям в конце §4.5, $(P, +, \cdot) \cong (\overline{P}, \oplus, \otimes)$, т. к. $\operatorname{Im} f = \overline{P}$. Значит, $(P, +, \cdot)$ с точностью до изоморфизма отождествляется с подполем в $(P[x]/(g(x)), \oplus, \otimes)$.•

Свойства гомоморфизмов колец

1. Композиция гомоморфизмов колец – гомоморфизм колец.

▷ Пусть $(K_1, +, \cdot)$, $(K_2, \tilde{+}, \bullet)$ и $(K_3, \hat{+}, \circ)$ – кольца, $f: K_1 \to K_2$, $\psi: K_2 \to K_3$ – гомоморфизмы колец, $\psi f: K_1 \to K_3$ – композиция функций. Тогда для $\forall\, a, b \in K_1$ выполняются равенства:

1) $\psi f(a+b) = \psi\big(f(a+b)\big) = \psi\big(f(a) \tilde{+} f(b)\big) = \psi\big(f(a)\big) \hat{+} \psi\big(f(b)\big) = \psi f(a) \hat{+} \psi f(b)$;

2) $\psi f(a \cdot b) = \psi\big(f(a \cdot b)\big) = \psi\big(f(a) \bullet f(b)\big) = \psi\big(f(a)\big) \circ \psi\big(f(b)\big) = \psi f(a) \circ \psi f(b)$.

Итак, $\psi f: K_1 \to K_3$ – гомоморфизм колец. ◁

2. Если $f: K_1 \to K_2$ – гомоморфизм колец, то $\mathrm{Im} f$ – подкольцо K_2.

▷ $\mathrm{Im} f \subseteq K_2$ и $\mathrm{Im} f \neq \varnothing$, т. к. $K_1 \neq \varnothing$: $\exists\, k \in K_1 \Rightarrow \exists f(k) \in \mathrm{Im} f$. $(K_1, +)$ и $(K_2, \tilde{+})$ – группы, f – гомоморфизм данных аддитивных групп. Поэтому по свойству 2 гомоморфизмов групп $(\mathrm{Im} f, \tilde{+}) \leq (K_2, \tilde{+})$. Для $\forall f(a), f(b) \in \mathrm{Im} f$ выполняется условие $f(a) \bullet f(b) = f(a \cdot b) \in \mathrm{Im} f$, поскольку f – гомоморфизм колец. Следовательно, $\mathrm{Im} f$ – подкольцо K_2. ◁

3. Если $f: K_1 \to K_2$ – гомоморфизм колец, то $f(0_{K_1}) = 0_{K_2}$ и $f(-a) = \tilde{-} f(a)$ для $\forall\, a \in K_1$.

▷ Так как $f: K_1 \to K_2$ – гомоморфизм соответствующих аддитивных групп колец $(K_1, +)$ и $(K_2, \tilde{+})$, то по свойству 2 гомоморфизмов групп имеем $f(0_{K_1}) = 0_{K_2}$ и $f(-a) = \tilde{-} f(a)$ для $\forall\, a \in K_1$. ◁

4. Если $f: K_1 \to K_2$ – гомоморфизм колец, то $f(1_{K_1}) = 1_{\mathrm{Im} f}$ и для $\forall\, a \in K_1^*$ $f(a^{-1}) = f(a)^{-1}$ в $\mathrm{Im} f$, где 1_{K_1} – единица K_1 и $1_{\mathrm{Im} f}$ – единица $\mathrm{Im} f$.

▷ Согласно свойству 2 гомоморфизмов колец $\mathrm{Im} f \leq K_2$. Если 1_{K_1} – единица K_1, то для $\forall\, a \in K_1$, поскольку f – гомоморфизм, выполняются равенства

$$f(a \cdot 1_{K_1}) = f(a) = f(a) \bullet f(1_{K_1}),\ f(1_{K_1} \cdot a) = f(a) = f(1_{K_1}) \bullet f(a).$$

То есть $f(1_{K_1})$ – единица $\mathrm{Im} f$. Для $\forall\, a \in K_1^*$ выполняются равенства $f(a \cdot a^{-1}) = f(1_{K_1}) = 1_{\mathrm{Im} f} = f(a) \bullet f(a^{-1})$ и $f(a^{-1} \cdot a) = f(1_{K_1}) = 1_{\mathrm{Im} f} = f(a^{-1}) \bullet f(a)$, следовательно, $f(a^{-1}) = f(a)^{-1}$ в $\mathrm{Im} f$. ◁

5. Если $f: K_1 \to K_2$ – гомоморфизм колец, то $\mathrm{Ker} f$ – двусторонний идеал K_1.

▷ $\mathrm{Ker} f \subseteq K_1$ и $\mathrm{Ker} f \neq \varnothing$, т. к. $0_{K_1} \in \mathrm{Ker} f$ по свойству 3 гомоморфизмов колец. Для $\forall\, a, b \in \mathrm{Ker} f$ $f(a-b) = f\big(a + (-b)\big) = f(a) \tilde{-} f(b) = 0_{K_2} \tilde{-} 0_{K_2} = 0_{K_2}$ по свойству 3 гомоморфизмов колец, следовательно, $a - b \in \mathrm{Ker} f$. Далее, для $\forall\, a \in \mathrm{Ker} f$, $\forall\, k \in K_1$ $f(a \cdot k) = f(a) \bullet f(k) = 0_{K_2} \bullet f(k) = 0_{K_2}$, значит, $a \cdot k \in \mathrm{Ker} f$, $f(k \cdot a) = f(k) \bullet f(a) = f(k) \bullet 0_{K_2} = 0_{K_2}$, значит, $k \cdot a \in \mathrm{Ker} f$. Итак, $\mathrm{Ker} f \lhd K_1$. ◁

Пример 4.6.4. Рассмотрим функцию $f: \mathbf{Z}/28\mathbf{Z} \to \mathbf{Z}/28\mathbf{Z}$, где $f(\bar{a}) = \bar{8} \cdot \bar{a}$. f – эндоморфизм кольца $(\mathbf{Z}/28\mathbf{Z}, +, \cdot)$, т. к. для любых $\bar{a}, \bar{b} \in \mathbf{Z}/28\mathbf{Z}$ имеем:

1) $f(\bar{a} + \bar{b}) = \bar{8} \cdot (\bar{a} + \bar{b}) = \bar{8} \cdot \bar{a} + \bar{8} \cdot \bar{b} = f(\bar{a}) + f(\bar{b})$;

2) $f(\bar{a} \cdot \bar{b}) = \bar{8} \cdot (\bar{a} \cdot \bar{b}) = (\bar{8} \cdot \bar{8}) \cdot (\bar{a} \cdot \bar{b}) = (\bar{8} \cdot \bar{a}) \cdot (\bar{8} \cdot \bar{b}) = f(\bar{a}) \cdot f(\bar{b})$.

$$\mathrm{Im}f = \{\bar{0}, \bar{4}, \bar{8}, \overline{12}, \overline{16}, \overline{20}, \overline{24}\}, \ \mathrm{Ker}f = \{\bar{0}, \bar{7}, \overline{14}, \overline{21}\}.$$

Можно заметить, что $f(\bar{c}) = \bar{8} \cdot \bar{c} = \overline{8 \cdot c} = \bar{c} = e_{\mathrm{Im}f}(\bar{c})$ для $\forall\, \bar{c} \in \mathrm{Im}f$, поскольку $f(\bar{1}) = \bar{8} = 1_{\mathrm{Im}f}$. $\mathrm{Im}f$ и $\mathrm{Ker}f$, как подгруппы циклической группы ($\mathbf{Z}/28\mathbf{Z}$, +), являются главными идеалами кольца $\mathbf{Z}/28\mathbf{Z}$, т. е. $\mathrm{Im}f = (\bar{k})$, где $\mathrm{ord}(\bar{k}) = 7 = 28/(k, 28)$, откуда $(k, 28) = 4$ и $\mathrm{Ker}f = (\bar{l})$, где $\mathrm{ord}(\bar{l}) = 4 = 28/(l, 28)$, значит, $(l, 28) = 7$. Таким образом,

$$\mathrm{Im}f = (\bar{4}) = (\bar{8}) = (\overline{12}) = (\overline{16}) = (\overline{20}) = (\overline{24}), \ \mathrm{Ker}f = (\bar{7}) = (\overline{21}).$$

Здесь $\mathrm{Im}f$ – подкольцо $\mathbf{Z}/28\mathbf{Z}$, но $1_{\mathrm{Im}f} \neq 1_{\mathbf{Z}/28\mathbf{Z}} = \bar{1}$, поскольку $\bar{1} \notin \mathrm{Im}f$.●

6. Гомоморфизм колец $f: K_1 \to K_2$ является мономорфизмом тогда и только тогда, когда $\mathrm{Ker}f = \{0_{K_1}\}$.

▷ Доказательство вытекает из свойства 4 гомоморфизмов групп, поскольку $f: K_1 \to K_2$ – гомоморфизм соответствующих групп $(K_1, +)$ и $(K_2, \breve{+})$. ◁

Из свойств 5 и 6 гомоморфизмов колец следует, что любой гомоморфизм $f: K_1 \to K_2$, где K_1 – поле, является либо нулевым, либо инъективным (т. к. поля не имеют нетривиальных идеалов). Гомоморфизмы позволяют произвести отождествления изоморфных полей, установить между полями отношения частичного порядка – по включению.

Аналогично теореме 3.7.6 можно доказать, что множество $\mathrm{Aut}(K)$ всех автоморфизмов произвольного кольца $(K, +, \cdot)$ является группой относительно операции композиции функций.

Теорема 4.6.1 (первая теорема о гомоморфизмах колец). Пусть $(K, +, \cdot)$ – кольцо, $I \lhd K$ – двусторонний идеал, $(K/I, \oplus, \otimes)$ – факторкольцо с индуцированными операциями. Тогда существует эпиморфизм колец $f: K \to K/I$, для которого $\mathrm{Ker}f = I$.

▷ Построим функцию $f: K \to K/I$, где $f(a) = \bar{a} = a + I$ для $\forall\, a \in K$. f – сюръекция: для $\forall\, \bar{a} \in K/I \ \exists f^{-1}(\bar{a}) = a \in K$. Согласно построению факторкольца $(K/I, \oplus, \otimes)$ для $\forall\, a, b \in K$ имеем:

1) $f(a + b) = \overline{a + b} = \bar{a} \oplus \bar{b} = f(a) \oplus f(b)$;

2) $f(a \cdot b) = \overline{a \cdot b} = \bar{a} \otimes \bar{b} = f(a) \otimes f(b)$.

Поэтому f – гомоморфизм колец.

Для $\forall\, i \in I \ f(i) = I = \bar{0} \Rightarrow I \subseteq \mathrm{Ker}f$. Пусть $a \in \mathrm{Ker}f$, значит, $f(a) = \bar{0} = I$, но если $a \notin I$, то $f(a) = a + I \neq I$, т. к. различные классы вычетов по модулю идеала не пересекаются. Получаем противоречие. Значит, $a \in I$ и $\mathrm{Ker}f \subseteq I$. Итак, $\mathrm{Ker}f = I$. ◁

Определение 4.6.4. Гомоморфизм колец $f: K \to K/I$, где $I \lhd K$, при котором $f(a) = \bar{a}$ для $\forall\, a \in K$, называется *естественным* (или *каноническим*) гомоморфизмом.

Теорема 4.6.2 (вторая теорема о гомоморфизмах колец). Пусть $\psi\colon K_1 \to K_2$ – гомоморфизм кольца $(K_1, +, \cdot)$ в кольцо $(K_2, \dot{+}, \bullet)$. Тогда $K_1/\operatorname{Ker}\psi \cong \operatorname{Im}\psi$.

$\triangleright \operatorname{Im}\psi \leq K_2$ (свойство 2 гомоморфизмов колец), $\operatorname{Ker}\psi \lhd K_1$ (свойство 5 гомоморфизмов колец). Построим функцию $f\colon K_1/\operatorname{Ker}\psi \to \operatorname{Im}\psi$, где $f(\bar{a}) = \psi(a)$ для $\forall\, \bar{a} \in K_1/\operatorname{Ker}\psi$. Функция f не зависит от выбора представителя класса вычетов $\bar{a} \in K_1/\operatorname{Ker}\psi$: пусть $a_1 = a + i \in \bar{a}$, $i \in \operatorname{Ker}\psi$, тогда получаем, что

$$f(\overline{a_1}) = f(\overline{a+i}) = \psi(a+i) = \psi(a) \dot{+} \psi(i) = \psi(a) \dot{+} 0_{K_2} = \psi(a) = f(\bar{a}).$$

f – сюръективная функция, т. к. для $\forall\, \psi(a) \in \operatorname{Im}\psi \;\; \exists f^{-1}(\psi(a)) = a + \operatorname{Ker}\psi = \bar{a} \in K_1/\operatorname{Ker}\psi$. Пусть $\bar{a} \neq \bar{b}$, т. е. $a - b \notin \operatorname{Ker}\psi$, тогда $f(\bar{a}) = \psi(a) \neq \psi(b) = f(\bar{b})$. Так как иначе $0_{K_2} = \psi(a) \dot{-} \psi(b) = \psi(a-b) \Rightarrow a - b \in \operatorname{Ker}\text{ш} \Rightarrow \bar{a} = \bar{b}$, что приводит к противоречию тому, что $\bar{a} \neq \bar{b}$. Значит, f – инъекция. Итак, $f\colon K_1/\operatorname{Ker}\psi \to \operatorname{Im}\psi$ – биективная функция. Для $\forall\, \bar{a}, \bar{b} \in K_1/\operatorname{Ker}\text{ш}$ имеем:

1) $f(\bar{a} \oplus \bar{b}) = f(\overline{a+b}) = \psi(a+b) = \psi(a) \dot{+} \psi(b) = f(\bar{a}) \dot{+} f(\bar{b})$;

2) $f(\bar{a} \otimes \bar{b}) = f(\overline{a \cdot b}) = \psi(a \cdot b) = \psi(a) \bullet \psi(b) = f(\bar{a}) \bullet f(\bar{b})$.

Значит, f – гомоморфизм колец. Таким образом, f – изоморфизм колец $(K_1/\operatorname{Ker}\psi, \oplus, \otimes)$ и $(\operatorname{Im}\psi, \dot{+}, \bullet)$. $\triangleleft$

Поскольку для произвольного кольца $(K, +, \cdot)$ $K/\{0\} = K$, из свойства 6 гомоморфизмов колец и теоремы 4.6.2 следует, что любой инъективный эндоморфизм кольца является автоморфизмом. В частности, любой ненулевой эндоморфизм поля является его автоморфизмом.

Пример 4.6.5. Пусть $(P, +, \cdot)$ – поле, $(P[x], +, \cdot)$ – кольцо полиномов над полем P. $\alpha \in P$ – фиксированный элемент поля. Рассмотрим функцию $\text{ш}\colon P[x] \to P$, где $\psi(g(x)) = g(\alpha), \forall\, g(x) \in P[x]$. Тогда для $\forall\, g(x), h(x) \in P[x]$ справедливы равенства:

1) $\psi(g(x) + h(x)) = g(\alpha) + h(\alpha) = \psi(g(x)) + \psi(h(x))$;

2) $\psi(g(x) \cdot h(x)) = g(\alpha) \cdot h(\alpha) = \psi(g(x)) \cdot \psi(h(x))$.

Следовательно, ψ – гомоморфизм колец.

$$\operatorname{Ker}\psi = \{g(x) \in P[x] \mid g(\alpha) = 0\} = \{g(x) \in P[x] \mid g(x) \vdots (x-\alpha)\} = (x-\alpha)$$

согласно следствию 1 из теоремы 4.4.4 Безу. Функция ψ сюръективна, т. к. для $\forall \beta \in P \Rightarrow \beta \in P[x]$, $\psi(\beta) = \beta$ и $\beta \in \psi^{-1}(\beta)$. Значит, $\operatorname{Im}\psi = P$. Таким образом, по теореме 4.6.2 $(P[x]/(x-\alpha), \oplus, \otimes) \cong (P, +, \cdot)$.$\bullet$

Развитием следствия 2 из теоремы 4.5.2 является следующая теорема.

Теорема 4.6.3 (теорема существования корня). Для всякого неприводимого полинома $g(x) \in P[x]$ степени $n \in \mathbf{N}$ существует расширение поля P, содержащее корень этого полинома и изоморфное полю $(P[x]/(g(x)), \oplus, \otimes)$.

▷ Факторкольцо $(P[x]/(g(x)), \oplus, \otimes)$ является полем согласно следствию 2 из теоремы 4.5.2. Подполе $\overline{P} = \{\overline{a} \mid a \in P\}$ в $P[x]/(g(x))$ изоморфно полю P, очевидно, изоморфизм задает функция $\psi: P \to P[x]/(g(x))$, где $\psi(a) = \overline{a}$, являющаяся вложением P в $P[x]/(g(x))$, как уже было показано в примере 4.6.3. Пусть $g(x) = a_n x^n + \ldots + a_1 x + a_0,\ a_i \in P,\ i = \overline{0,n}$, тогда в поле $P[x]/(g(x))$ $\overline{g(x)} = \overline{0}$. Но поскольку $\overline{g(x)} = \overline{a_n x^n + \ldots + a_1 x + a_0} = \overline{a_n}\,\overline{x}^n \oplus \ldots \oplus \overline{a_1}\,\overline{x} \oplus \overline{a_0}$, то $\overline{x}$ является корнем полинома $\overline{a_n} x^n \oplus \ldots \oplus \overline{a_1} x \oplus \overline{a_0} \in \overline{P}[x]$. Рассмотрим теперь множество S, удовлетворяющее условиям $S \cap P = \varnothing,\ |S| = \left|(P[x]/(g(x)))\backslash \overline{P}\right| \neq 0$ при $n > 1$. Пусть $F = S \cup P$, при $n = 1$ $F = P$. Зададим на F структуру поля, продолжив мономорфизм ш до изоморфизма F на $P[x]/(g(x))$. Если $b, c \in F$, то полагаем

$$b + c = \psi^{-1}(\psi(b) \oplus \psi(c)),\ b \cdot c = \psi^{-1}(\psi(b) \otimes \psi(c)).$$

При ограничениях на P эти операции совпадают соответственно с заданными операциями сложения и умножения в P, и ясно, что P – подполе F. Положим $\alpha = \psi^{-1}(\overline{x})$, тогда $\psi(g(\alpha)) = \psi(a_n\alpha^n + \ldots + a_1\alpha + a_0) = \overline{a_n}\,\psi(\alpha)^n \oplus \ldots \oplus \overline{a_1}\,\psi(\alpha) \oplus \overline{a_0} = \overline{a_n}\,\overline{x}^n \oplus \ldots \oplus \overline{a_1}\,\overline{x} \oplus \overline{a_0} = \overline{0}$ и, поскольку ψ – изоморфизм F на $P[x]/(g(x))$, то $g(\alpha) = 0$ в поле $(F, +, \cdot)$. Значит, построенное поле является расширением поля P, содержащим корень α полинома $g(x)$. ◁

Следствие 1. Для любого полинома $g(x) \in P[x]$ степени $n \in \mathbf{N}$ существует расширение поля P, содержащее корень этого полинома.

▷ По теореме 4.4.1 полином $g(x)$ однозначно разлагается на множители: $g(x) = a_n p_1(x) p_2(x) \ldots p_s(x)$, где $p_i(x),\ i = \overline{1,s}$, – неприводимые над P полиномы со старшими коэффициентами, равными 1; a_n – старший коэффициент $g(x)$. Согласно теореме 4.6.3 для каждого $p_i(x),\ i = \overline{1,s}$, существует расширение поля P, содержащее корень данного полинома, являющийся и корнем полинома $g(x)$ в соответствии со следствием 1 из теоремы 4.4.4 Безу и свойством 3 делимости полиномов. ◁

Замечание. Тот факт, что расширение P' поля P содержит корень α полинома $g(x) \in P[x]$, вовсе не означает, что P' содержит все корни этого полинома.

Пример 4.6.6. Полином $g(x) = x^4 - 2$ неприводим над $\mathbf{Q}$ по признаку Эйзенштейна: $p = 2$. В поле $\mathbf{C}$ данный полином имеет четыре простых корня: $\sqrt[4]{2}$, $\sqrt[4]{2}i$, $-\sqrt[4]{2}$ и $-\sqrt[4]{2}i$. Но поле $\mathbf{Q}(\sqrt[4]{2}) = \{a + b\sqrt[4]{2} + c\sqrt{2} + d\sqrt[4]{2}\sqrt{2} \mid a, b, c, d \in \mathbf{Q}\}$, являющееся расширением $\mathbf{Q}$ и подполем $\mathbf{C}$, содержит только первый и третий из этих корней, но не все четыре. ●

Следствие 2. Для любого полинома $g(x) \in P[x]$ степени $n \in \mathbf{N}$ существует расширение поля P, содержащее все корни $g(x)$.

▷ Доказательство проведем методом индукции по степени полинома $g(x)$. Если $\deg g = 1$, то $g(x) = ax + b,\ a \neq 0$, и $-a^{-1} \cdot b \in P$ – единственный корень $g(x)$, значит, P – искомое поле. Предположим, что утверждение верно для всех мно-

гочленов степеней, меньших фиксированного $n \in \mathbf{N}_{>1}$, с коэффициентами из произвольных полей.

Пусть теперь $\deg g = n > 1$. Тогда по следствию 1 из теоремы 4.6.3 существует расширение P_1 поля P, содержащее корень α полинома $g(x)$. Согласно следствию 1 из теоремы 4.4.4 в $P_1[x]$ $g(x) = (x - \alpha)h(x)$, где $h(x) \in P_1[x]$ и $\deg h = n - 1$. По предположению индукции существует расширение P_2 поля P_1, содержащее все корни $h(x)$. Так как P_2 является расширением поля P и содержит все корни полинома $g(x)$ (оно содержит α и все корни полинома $h(x)$), то P_2 и есть искомое расширение. ◁

Следствие 3. Если поле P содержит все корни многочлена $g(x) \in P[x]$ при $\deg g = n \in \mathbf{N}$, в частности, когда P – алгебраически замкнутое поле, то справедливо каноническое разложение:

$$g(x) = a_n(x - \alpha_1)^{k_1} \ldots (x - \alpha_s)^{k_s},$$

где a_n – старший коэффициент $g(x)$;

$\alpha_1, \ldots, \alpha_s$ – все различные корни $g(x)$ в P соответственно кратностей $k_1, \ldots, k_s$.

В этом случае выполняется равенство $k_1 + \ldots + k_s = n$, т. е. число всех корней ненулевого многочлена (с учетом их кратностей) равно степени данного многочлена.

▷ Доказательство аналогично доказательству следствий 3 и 4 из теоремы 4.4.4 с использованием следствия 1 из теоремы 4.4.4, свойства 3 делимости многочленов и теоремы 4.4.1. ◁

Теорема 4.6.4. Пусть $f : K \to K'$ – эпиморфизм колец, $\operatorname{Ker} f = I$. Тогда существует взаимно однозначное и сохраняющее включения соответствие между всеми идеалами U' кольца $(K', \dot{+}, \bullet)$ и идеалами U кольца $(K, +, \cdot)$, содержащими I, такое, что $f(U) = U'$, $f^{-1}(U') = U$.

▷ $\operatorname{Im} f = K'$, $\operatorname{Ker} f = I$. Пусть U' – произвольный идеал кольца $(K', \dot{+}, \bullet)$, соответственно левый, правый, двусторонний. Рассмотрим прообраз идеала U' при отображении f – $f^{-1}(U') \subseteq K$. Тогда для $\forall a, b \in f^{-1}(U')$, $\forall k \in K$ выполняются следующие свойства:

1) $a - b \in f^{-1}(U')$, т. к. $f(a - b) \in U'$. Поскольку $f(a), f(b) \in U'$ и U' – идеал кольца K', значит, $f(a) \dot{-} f(b) \in U'$, а $f(a) \dot{-} f(b) = f(a - b)$;

2) $k \cdot a \in f^{-1}(U')$, если U' – левый идеал, $a \cdot k \in f^{-1}(U')$, если U' – правый идеал, $k \cdot a$, $a \cdot k \in f^{-1}(U')$, если U' – двусторонний идеал. Так как соответственно $f(k \cdot a) \in U'$, $f(a \cdot k) \in U'$, $f(k \cdot a)$, $f(a \cdot k) \in U'$, поскольку $f(a) \in U'$, $f(k) \in K'$ и соответственно $f(k) \bullet f(a) \in U'$, $f(a) \bullet f(k) \in U'$, $f(k) \bullet f(a)$, $f(a) \bullet f(k) \in U'$, а $f(k) \bullet f(a) = f(k \cdot a)$, $f(a) \bullet f(k) = f(a \cdot k)$.

Таким образом, $f^{-1}(U')$ является соответственно левым, правым, двусторонним идеалом кольца $(K, +, \cdot)$ согласно определению. Поскольку $0_{K'} \in U'$, то $f^{-1}(0_{K'}) = I \subseteq f^{-1}(U')$. Если U_1', U_2' – идеалы кольца $(K', \dot{+}, \bullet)$, причем

$U_1' \subseteq U_2'$, то для $\forall a' \in U_1' \Rightarrow a' \in U_2'$. Значит, для $\forall a \in f^{-1}(U_1')$ $\exists a' \in U_1'$, такое, что $f(a) = a'$. А поскольку $a' \in U_2'$, то $a \in f^{-1}(a') \subseteq f^{-1}(U_2')$, следовательно, $f^{-1}(U_1') \subseteq f^{-1}(U_2')$. Итак, существует функция, заданная на множестве всех идеалов кольца $(K', \dot{+}, \bullet)$, которая каждому левому, правому, двустороннему идеалу ставит в соответствие его прообраз при отображении f, являющийся соответственно левым, правым, двусторонним идеалом кольца $(K, +, \cdot)$, содержащим I, и данная функция сохраняет включения идеалов.

Пусть теперь U – произвольный идеал кольца $(K, +, \cdot)$, соответственно левый, правый, двусторонний. Рассмотрим образ идеала U при отображении f – $f(U) \subseteq K'$. Тогда для $\forall f(a), f(b) \in f(U)$, $\forall k' = f(k) \in K'$ выполняются следующие свойства:

1) $f(a) \dot{-} f(b) \in f(U)$, т. к. $a - b \in U$. Поскольку $a, b \in U$ и U – идеал кольца K, значит, $f(a-b) \in f(U)$, а $f(a-b) = f(a) \dot{-} f(b)$;

2) $k' \bullet f(a) \in f(U)$, если U – левый идеал, $f(a) \bullet k' \in f(U)$, если U – правый идеал, $k' \bullet f(a), f(a) \bullet k' \in f(U)$, если U – двусторонний идеал. Так как соответственно $k \cdot a \in U$, $a \cdot k \in U$, $k \cdot a$, $a \cdot k \in U$, поскольку $a \in U, k \in K$ и соответственно $f(k \cdot a) \in f(U)$, $f(a \cdot k) \in f(U)$, $f(k \cdot a), f(a \cdot k) \in f(U)$, а $f(k \cdot a) = f(k) \bullet f(a) = k' \bullet f(a)$, $f(a \cdot k) = f(a) \bullet f(k) = f(a) \bullet k'$.

Таким образом, $f(U)$ является соответственно левым, правым, двусторонним идеалом кольца $(K', \dot{+}, \bullet)$ согласно определению. Если U_1, U_2 – идеалы кольца $(K, +, \cdot)$, причем $U_1 \subseteq U_2$, то для $\forall a \in U_1 \Rightarrow a \in U_2$. Значит, для $\forall a' \in f(U_1)$ $\exists a \in U_1$, такое, что $f(a) = a'$. А поскольку $a \in U_2$, то $f(a) = a' \in f(U_2)$, следовательно, $f(U_1) \subseteq f(U_2)$. Для выполнения условия $f(U) = U'$, $f^{-1}(U') = U$, где U' – идеал кольца $(K', \dot{+}, \bullet)$, необходимо и достаточно, чтобы идеал U кольца $(K, +, \cdot)$ содержал I.

Итак, показано, что существует взаимно однозначное и сохраняющее включения соответствие между всеми левыми, правыми, двусторонними идеалами кольца $(K', \dot{+}, \bullet)$ и их прообразами при отображении f, являющимися соответственно левыми, правыми, двусторонними идеалами кольца $(K, +, \cdot)$, содержащими I. $\triangleleft$

Замечание. В ходе доказательства теоремы 4.6.4 было показано, что при произвольном гомоморфизме колец $f: K_1 \to K_2$ образ любого левого, правого, двустороннего идеала кольца $(K_1, +, \cdot)$ является соответственно левым, правым, двусторонним идеалом кольца $(\mathrm{Im}f, \dot{+}, \bullet)$. Также было показано, что прообраз любого левого, правого, двустороннего идеала кольца $(\mathrm{Im}f, \dot{+}, \bullet)$ является соответственно левым, правым, двусторонним идеалом кольца $(K_1, +, \cdot)$, содержащим $\mathrm{Ker}f$.

Следствие. При произвольном гомоморфизме колец $\psi: K_1 \to K_2$ существует взаимно однозначное и сохраняющее включения соответствие между всеми

левыми, правыми, двусторонними идеалами кольца (K_1/Ker ψ, $\oplus$, $\otimes$) и соответственно левыми, правыми, двусторонними идеалами кольца (Im ψ, $\breve{+}$, $\bullet$).

▷ Согласно теореме 4.6.2 существует кольцевой изоморфизм f: K_1/Ker $\psi \to$ Im ψ, для которого Ker$f$ = Ker $\psi = \overline{0}$. Применяя теорему 4.6.4, получаем, что f задает требуемое взаимно однозначное соответствие между идеалами $\overline{U}$ кольца (K_1/Ker ψ, $\oplus$, $\otimes$) и U' кольца (Im ψ, $\breve{+}$, $\bullet$), такое, что $f(\overline{U}) = U'$, $f^{-1}(U') = \overline{U}$. ◁

§4.7. Характеристика. Минимальные поля

Пусть в этом параграфе (K, +, ·) – ассоциативное и коммутативное кольцо с единицей $e \neq 0$, без делителей нуля.

Функция f: $\mathbf{Z} \to K$, где $f(z) = ze$ для $\forall\, z \in \mathbf{Z}$, является, что легко проверить, кольцевым гомоморфизмом с Ker$f$ = (n) при некотором $n \in \mathbf{Z}_{\geq 0}$. Согласно теореме 4.6.2 $\mathbf{Z}/(n) = \mathbf{Z}/n\mathbf{Z} \cong$ Imf, где Imf – некоторое подкольцо в K. Поскольку в (K, +, ·) нет делителей нуля, в (Imf, +, ·) и в ($\mathbf{Z}/n\mathbf{Z}$, $\oplus$, $\otimes$) также нет делителей нуля, поэтому либо $n = 0$, либо $n = p$ – простое число. В первом случае K содержит в качестве подкольца кольцо, изоморфное $\mathbf{Z}$ и часто отождествляемое с $\mathbf{Z}$. Во втором случае K содержит в качестве подкольца изоморфный образ ($\mathbf{Z}/p\mathbf{Z}$, $\oplus$, $\otimes$), являющийся полем $\boldsymbol{F_p}$.

Определение 4.7.1. Говорят, что кольцо (K, +, ·) имеет *характеристику 0* в первом случае и p во втором случае.

Таким образом, и всякое поле имеет характеристику 0 или p, где p – простое число. Можно дать и другое, эквивалентное определение характеристики кольца и непосредственно показать, что в случае, когда характеристика ненулевая, она является простым числом.

Определение 4.7.2. Кольцо (K, +, ·) имеет *характеристику, равную 0*, если для каждого $n \in \mathbf{N}$ $ne \neq 0$. Если для кольца (K, +, ·) существует такое $n \in \mathbf{N}$, что $ne = 0$, причем n – наименьшее натуральное число с таким свойством, то говорят, что *характеристика кольца равна n*. Характеристику кольца K будем обозначать char K (от англ. characteristic – «характеристика»).

Теорема 4.7.1. Если характеристика кольца (K, +, ·) отлична от 0, то она является простым числом.

▷ Пусть характеристика кольца $n \in \mathbf{N}$ является составным числом: $n = n_1 n_2$, где $1 < n_1, n_2 < n$. Тогда

$$ne = (n_1 n_2)e = n_1(n_2 e) = (n_1 e) \cdot (n_2 e) = 0.$$

Так как n – минимальное натуральное число со свойством $ne = 0$, то $n_1 e \neq 0$, $n_2 e \neq 0$ и получаем противоречие тому, что в кольце нет делителей нуля. Поэтому n – простое число. ◁

Пример 4.7.1. Поле $\boldsymbol{F_p} \cong \mathbf{Z}/p\mathbf{Z}$, где p – простое число, $f(\overline{k}) = ke$ для всех $\overline{k} \in \mathbf{Z}/p\mathbf{Z}$, f – изоморфизм полей, и char $\boldsymbol{F_p} = p$. В самом деле, $\mathbf{Z}/p\mathbf{Z}$, а значит, и

$\boldsymbol{F_p}$ являются аддитивными группами порядка p и, следовательно, по следствию 3 из теоремы 3.4.2 Лагранжа, циклическими группами порядка p, порожденными, в частности, классом вычетов $\bar{1}$ и единицей $e = f(\bar{1})$ поля $\boldsymbol{F_p}$ соответственно. Следовательно, согласно структуре циклических групп имеем

$$\mathbf{Z}/p\mathbf{Z} = \{\bar{1},\ \bar{1} \oplus \bar{1} = \bar{2},\ \bar{1} \oplus \bar{1} \oplus \bar{1} = \bar{3},\ldots,\ p\bar{1} = \bar{p} = \bar{0}\},$$

$$\boldsymbol{F_p} = \{e,\ e + e = 2e,\ e + e + e = 3e,\ldots,\ pe = 0\}.$$

Это и означает, что $\operatorname{char} \boldsymbol{F_p} = \operatorname{char} \mathbf{Z}/p\mathbf{Z} = p$. Любое конечное поле $\boldsymbol{F_q}$, где $q = p^m$, $m \in \mathbf{N}$, также имеет характеристику p, поскольку оно содержит $\boldsymbol{F_p}$ в качестве подполя и e – единица $\boldsymbol{F_p}$ и $\boldsymbol{F_q}$.●

Пример 4.7.2. Кольцо $\mathbf{Z}$, а также поля $\mathbf{Q}$, $\mathbf{R}$ и $\mathbf{C}$ имеют, очевидно, характеристику 0.●

Пример 4.7.3. Пусть $(P, +, \cdot)$ – поле и ψ – вложение P в факторкольцо $(P[x]/(g(x)), \oplus, \otimes)$ для некоторого полинома $g(x) \in P[x]$, $\deg g \geq 1$, где $\psi(\alpha) = \bar{\alpha}$ для $\forall\, \alpha \in P$. Тогда $P[x]/(g(x))$ – кольцо той же характеристики, что и P, поскольку $\psi(e) = \bar{e}$ и $n\bar{e} = n\,\psi(e) = \psi(ne) = \bar{0} \Leftrightarrow ne = 0$ при $n \in \mathbf{N}$.●

Пример 4.7.4. Для произвольного поля P кольцо полиномов $P[x]$ и *поле рациональных функций* (или *поле частных*) кольца $P[x]$, т. е. поле $P(x) = \{g(x)/h(x) = g(x) \cdot h(x)^{-1} \mid g(x), h(x) \in P[x],\ h(x) \neq 0\}$ с естественными операциями сложения и умножения дробей, имеют ту же характеристику, что и P. Действительно, $P < P[x] < P(x)$ и e – единица P, $P[x]$ и $P(x)$. В частности, $\operatorname{char} \boldsymbol{F_q}[x] = \operatorname{char} \boldsymbol{F_q}(x) = p$ при $q = p^m$ и $\forall\, m \in \mathbf{N}$. Итак, существуют бесконечные кольца и поля конечной характеристики.●

Пусть $\operatorname{char} K = p > 0$. Тогда для любого $n \in \mathbf{Z}$ $n = pq + r$, где $q \in \mathbf{Z}$, $0 \leq r < p$, согласно теореме 1.1.1, и для каждого $\alpha \in K$ $n\alpha = r\alpha$, поскольку $(pq)\alpha = p(q\alpha) = (pe) \cdot (q\alpha) = 0$. Отсюда получаем следующую полезную формулу бинома Ньютона $(a + b)^n$ для частного случая натуральной степени $n = p^m$ в кольце K. Формула названа в честь великого английского физика, математика и астронома Исаака Ньютона (1643–1727), обобщившего ее для произвольных комплексных показателей степени n.

Теорема 4.7.2. В кольце $(K, +, \cdot)$ с $\operatorname{char} K = p > 0$ справедлива формула

$$(a + b)^{p^m} = a^{p^m} + b^{p^m} \tag{4.7.1}$$

для любых $a, b \in K$ и для всякого $m \in \mathbf{N}$.

▷ Докажем справедливость формулы (4.7.1) методом математической индукции. При $m = 1$ формула бинома Ньютона имеет вид $(a + b)^p = \sum_{k=0}^{p} C_p^k a^{p-k} b^k$, где $C_p^0 = C_p^p = 1$, а все биномиальные коэффициенты $C_p^k = \dfrac{p!}{k!(p-k)!}$ для $1 \leq k < p$ являются натуральными числами и делятся на p, поскольку $p \mid p!$, а $(p, k!) = (p, (p-k)!) = 1$ в силу простоты p. Поэтому для любых $a, b \in K$ формула имеет вид

$$(a+b)^p = a^p + b^p. \qquad (4.7.2)$$

Предположим, что формула (4.7.1) верна для $\forall\, m \in \mathbf{N}$. Докажем ее справедливость для $m+1$. Поскольку $(a+b)^{p^{m+1}} = \left((a+b)^{p^m}\right)^p$, то согласно предположению индукции и формуле (4.7.2) при $m = 1$ получаем

$$\left((a+b)^{p^m}\right)^p = \left(a^{p^m} + b^{p^m}\right)^p = \left(a^{p^m}\right)^p + \left(b^{p^m}\right)^p = a^{p^{m+1}} + b^{p^{m+1}}.$$

Таким образом, формула (4.7.1) верна для любых $a, b \in K$ и $m \in \mathbf{N}$. ◁

Определение 4.7.3. *Минимальным* (или *простым*) называется поле, не содержащее собственных подполей.

Теорема 4.7.3. Поле **Q** – минимальное поле характеристики 0, для любого простого p $\mathbf{Z}/p\mathbf{Z}$ – минимальное поле характеристики p. Поля **Q** и $\mathbf{Z}/p\mathbf{Z}$ не имеют других автоморфизмов, кроме тождественного. В любом поле P существует в точности одно минимальное подполе, не имеющее других автоморфизмов, кроме тождественного, изоморфное либо **Q**, либо $\mathbf{Z}/p\mathbf{Z}$ в зависимости от того, равна ли характеристика поля P 0 или p.

▷ Предположим, что P_0 – подполе поля **Q**. Тогда $1 \in P_0$ и для $\forall\, m \in \mathbf{Z} \Rightarrow$ $\Rightarrow\ m \in P_0$, для $\forall\, n \in \mathbf{Z}\backslash\{0\} \ \Rightarrow\ n^{-1} \in P_0$, поэтому $m \cdot n^{-1} \in P_0$. Таким образом, $\mathbf{Q} \subseteq P_0$, следовательно, $P_0 = \mathbf{Q}$. Итак, **Q** – минимальное поле характеристики 0.

Пусть p – произвольное простое число. Предположим, что P_0 – подполе поля $\mathbf{Z}/p\mathbf{Z}$. Тогда $\bar{0} \in P_0$, $\bar{1} \in P_0$, следовательно, $\bar{2} = \bar{1} \oplus \bar{1} \in P_0, \ldots,$ $\overline{p-1} = \underbrace{\bar{1} \oplus \bar{1} \oplus \ldots \oplus \bar{1}}_{p-1} \in P_0$. Таким образом, $\mathbf{Z}/p\mathbf{Z} \subseteq P_0$, следовательно, $P_0 = \mathbf{Z}/p\mathbf{Z}$. Итак, $\mathbf{Z}/p\mathbf{Z}$ – минимальное поле характеристики p.

Согласно определению автоморфизма и свойствам 3 и 4 гомоморфизмов колец получаем следующее. Если $f\colon \mathbf{Q} \to \mathbf{Q}$ – автоморфизм, то $f(1) = 1$, $f(m) = mf(1) = m$ для $\forall\, m \in \mathbf{Z}$, $f(n^{-1}) = f(n)^{-1} = n^{-1}$ для $\forall\, n \in \mathbf{N}$, следовательно, для $\forall\, m \cdot n^{-1} \in \mathbf{Q}$ $f(m \cdot n^{-1}) = f(m) \cdot f(n^{-1}) = m \cdot n^{-1}$ и f – тождественный автоморфизм **Q**. Если $f\colon \mathbf{Z}/p\mathbf{Z} \to \mathbf{Z}/p\mathbf{Z}$ – автоморфизм, то $f(\bar{0}) = \bar{0}$, $f(\bar{1}) = \bar{1}$, тогда $f(\bar{k}) = f(\underbrace{\bar{1} \oplus \ldots \oplus \bar{1}}_{k}) = kf(\bar{1}) = k\bar{1} = \bar{k}$, $k = \overline{2, p-1}$, поэтому f – тождественный автоморфизм поля $\mathbf{Z}/p\mathbf{Z}$.

Пусть теперь P – произвольное поле. Как уже было показано в начале данного параграфа, если $\operatorname{char} P = 0$, то P содержит в качестве подкольца изоморфный образ кольца **Z**, а если $\operatorname{char} P = p$, то P содержит в качестве подполя изоморфный образ поля $\mathbf{Z}/p\mathbf{Z}$, обозначаемый $\boldsymbol{F_p}$. Если $\operatorname{char} P = 0$, то мономорфизм $f\colon \mathbf{Z} \to P$ можно продолжить до мономорфизма $\tilde{f}\colon \mathbf{Q} \to P$, положив $\tilde{f}(n^{-1}) = f(n)^{-1} = (ne)^{-1}$ для $\forall\, n \in \mathbf{Z}\backslash\{0\}$. Тогда $\operatorname{Im}\tilde{f}$ – подполе поля P и $\operatorname{Im}\tilde{f} \cong \mathbf{Q}$. Аналогично приведенным ранее доказательствам для **Q** и $\mathbf{Z}/p\mathbf{Z}$ можно показать, что подполя $\operatorname{Im}\tilde{f} = \{(me) \cdot (ne)^{-1} \mid m \in \mathbf{Z}, n \in \mathbf{N}\}$ и $\boldsymbol{F_p} = \{0, e, 2e, \ldots, (p-1)e\}$ также не имеют других автоморфизмов, кроме тождественного.

Если бы в поле P существовало несколько различных минимальных подполей, то легко проверить по определению подполя, что их пересечение было бы также подполем P и собственным подполем минимальных подполей, что противоречит определению минимального поля. Значит, минимальное подполе произвольного поля P единственно. $\triangleleft$

Литература

Основная

1. Биркгоф, Г. Современная прикладная алгебра / Г. Биркгоф, Т. К. Барти ; пер. с англ. – 2-е изд., стереотип. – СПб. : Лань, 2005. – 400 с.

2. Введение в криптографию / В. В. Ященко [и др.] ; под общ. ред. В. В. Ященко. – 3-е изд., перераб. – М. : МЦНМО, 2003. – 400 с.

3. Винберг, Э. Б. Алгебра многочленов : учеб. пособие / Э. Б. Винберг. – М. : Просвещение, 1980. – 176 с.

4. Виноградов, И. М. Основы теории чисел / И. М. Виноградов. – 9-е изд., перераб. – М. : Наука, 1981. – 176 с.

5. Каргополов, М. И. Основы теории групп / М. И. Каргополов, Ю. И. Мерзляков. – М. : Наука, 1972. – 240 с.

6. Карпов, В. Г. Математическая логика и дискретная математика : учеб. пособие / В. Г. Карпов, В. А. Мощенский. – Минск : Выш. шк., 1977. – 254 с.

7. Кострикин, А. И. Введение в алгебру. В 3 ч. Ч. 1 : Основы алгебры / А. И. Кострикин. – 3-е изд. – М. : Физматлит, 2004. – 272 с.

8. Коутинхо, С. Введение в теорию чисел. Алгоритм RSA / С. Коутинхо ; пер. с англ. – М. : Постмаркет, 2001. – 328 с.

9. Кузнецов, О. П. Дискретная математика для инженера / О. П. Кузнецов, Г. М. Адельсон-Вельский. – М. : Энергия, 1980. – 344 с.

10. Курош, А. Г. Курс высшей алгебры : учебник для вузов / А. Г. Курош. – 17-е изд., стереотип. – СПб. : Лань, 2008. – 432 с.

11. Липницкий, В. А. Современная прикладная алгебра. Математические основы защиты информации от помех и несанкционированного доступа : учеб. пособие / В. А. Липницкий. – Минск : БГУИР, 2005. – 88 с.

12. Математические и компьютерные основы криптологии : учеб. пособие / Ю. С. Харин [и др.]. – Минск : Новое знание, 2003. – 382 с.

13. Проскуряков, И. В. Сборник задач по линейной алгебре / И. В. Проскуряков. – 12-е изд., стереотип. – СПб. : Лань, 2008. – 480 с.

Дополнительная

1. Айерлэнд, К. Классическое введение в современную теорию чисел / К. Айерлэнд, М. Роузен ; пер. с англ. – М. : Мир, 1987. – 416 с.

2. Аршинов, Н. Н. Коды и математика / Н. Н. Аршинов, Л. Е. Садовский. – М. : Наука, 1983. – 124 с.

3. Баричев, С. Г. Основы современной криптографии : учеб. пособие для вузов / С. Г. Баричев, В. В. Гончаров, Р. Е. Серов. – 2-е изд., испр. и доп. – М. : Горячая линия – Телеком, 2002. – 175 с.

4. Бухштаб, А. А. Теория чисел / А. А. Бухштаб. – 3-е изд., стереотип. – М. : URSS, 2008. – 384 с.

5. Василенко, О. Н. Теоретико-числовые алгоритмы в криптографии / О. Н. Василенко. – М. : МЦНМО, 2003. – 326 с.

6. Коршунов, Ю. М. Математические основы кибернетики : учеб. пособие для вузов / Ю. М. Коршунов. – М. : Энергоатомиздат, 1987. – 496 с.

7. Ленг, С. Алгебра / С. Ленг ; пер. с англ. – М. : Мир, 1968. – 564 с.

8. Лидл, Р. Конечные поля. В 2 т. / Р. Лидл, Г. Нидеррайтер ; пер. с англ. – М. : Мир, 1988. – 820 с.

9. Морелос-Сарагоса, Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение / Р. Морелос-Сарагоса ; пер. с англ. – М. : Техносфера, 2005. – 320 с.

10. Ноден, П. Алгебраическая алгоритмика / П. Ноден, К. Китте ; пер. с англ. – М. : Мир, 1999. – 720 с.

11. Прасолов, В. В. Многочлены / В. В. Прасолов. – 3-е изд., испр. – М. : МЦНМО, 2003. – 336 с.

12. Романец, Ю. В. Защита информации в компьютерных системах и сетях / Ю. В. Романец, П. А. Тимофеев, В. Ф. Шаньгин ; под ред. В. Ф. Шаньгина. – 2-е изд., перераб. и доп. – М. : Радио и связь, 2001. – 376 с.

13. Сигорский, В. П. Математический аппарат инженера / В. П. Сигорский. – 2-е изд., стереотип. – Киев : Техника, 1977. – 768 с.

Учебное издание

Стройникова Елена Дмитриевна

ОСНОВЫ ПРИКЛАДНОЙ АЛГЕБРЫ

Учебно-методическое пособие

Редактор *Г. С. Корбут*
Корректор *Е. Н. Батурчик*
Компьютерная верстка *Е. Д. Стройникова*

Подписано в печать 05.10.2010. Формат 60x84 1/16. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 7,21. Уч.-изд. л. 6,3. Тираж 150 экз. Заказ 690.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
ЛИ №02330/0494371 от 16.03.2009. ЛП №02330/0494175 от 03.04.2009.
220013, Минск, П. Бровки, 6