Ш.Т. ИШМУХАМЕТОВ

МЕТОДЫ ФАКТОРИЗАЦИИ НАТУРАЛЬНЫХ ЧИСЕЛ

Казанский университет

2011

УДК 511, 519.6

Печатается по решению редакционно-издательского совета ФГАОУВПО
«КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»,
методической комиссии факультета вычислительной
математики и кибернетики, протокол № 6 от 21 января 2011 г.

Автор-составитель —
докт. физ. мат. наук, проф. каф. САИТ КФУ
Ш.Т. Ишмухаметов

Рецензент —
докт. техн. наук, проф. В.М. Захаров
(Казанский государственный технический университет им. А.Н.Туполева)

Ишмухаметов Ш.Т.
Методы факторизации натуральных чисел: учебное пособие /
Ш.Т. Ишмухаметов.– Казань: Казан. ун. 2011.– 190 с.

Факторизацией натурального числа называется разложение этого числа в произведение простых сомножителей. Эта задача имеет большую вычислительную сложность. Один из самых популярных методов криптографии с открытым ключом, метод RSA, основан на трудоемкости задачи факторизации длинных целых чисел. Другими важными проблемами теории чисел, имеющими важные приложения на практике, являются проблемы проверки простоты целого числа и построения больших простых чисел.

В этой книге мы даем описание наиболее известных методов проверки простоты натуральных чисел и факторизации, включая самые быстрые на сегодняшний день метод эллиптических кривых Х. Ленстры, метод квадратичного решета К. Померанца и метод решета числового поля Д. Полларда.

Предназначено для студентов старших курсов факультета вычислительной математики и кибернетики.

Содержание

Введение

> The problem of distinguishing prime numbers from composite numbers and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. Further, the dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated.
>
> Karl Friedrich Gauss «Disquisitiones Arithmeticae» (1801)

> Проблема различения простых и составных чисел и разложения последних в их главные факторы, как известно, является одной из самых важных и полезных в арифметике. Далее, достоинство самой науки требует, чтобы все возможные средства были исследованы для решения этой проблемы, столь изящной и настолько знаменитой.
>
> Карл Фридрих Гаусс «Арифметические исследования» (1801)

В 1977 г. трое ученых Рональд Райвест (Ronald Linn Rivest), Ади Шамир (Adi Shamir) и Леонард Адлеман (Leonard Adleman) из Массачусетского Технологического Института (MIT) опубликовали в журнале Scientific American новый алгоритм шифрования, основанный на идее двухключевого шифрования, названный по первым буквам фамилий авторов методом RSA.

В этом методе известным параметром служит некоторое целое число n большой длины (обычно 1024 или 2048 бита), являющееся произведением двух простых чисел p и q. Эти числа p и q являлись секретными параметрами метода, и для взлома системы RSA было достаточно найти множители p и q, т.е выполнить разложение числа n на простые сомножители.

На момент опубликования алгоритма RSA было известны лишь небольшое количество алгоритмов факторизации, самым известным из которых являлся метод Ферма. Эти методы позволяли на тот день факторизовать числа, состоящие не более чем из 25 – 30 цифр. Поэтому использование в качестве n натурального числа, имеющего более 100 десятичных знаков, гарантированно обеспечивало безопасность шифрования

этим методом. Сами создатели метода предложили всей математической общественности для тестового взлома 129-значное десятичное число, пообещав за его разложение условное вознаграждение в $100. Масла в огонь подлила также опубликованная в 1977 г. в журнале Sci.Amer. статья известного математика и популяризатора Мартина Гарднера «A new kind of cipher that would take millions of years to break» («Новый алгоритм шифрования, для взлома которого потребуется миллионы лет») [24].

Вызов, брошенный всему миру, не остался незамеченным. В соревнование по поиску быстрых алгоритмов факторизации включилось огромное количество людей, среди которых были известные математики, специалисты по теории чисел и криптографы. В результате этой гонки были созданы несколько алгоритмов факторизации, обогативших алгоритмическую теорию чисел рядом замечательных идей. В конце 80-х г. XX столетия были разработаны самые быстрыми на сегодняшний день алгоритмы факторизации - метод эллиптических кривых (the Elliptic Curves method)(Х.Ленстра [31]), метод квадратичного решета (the Quadratic Sieve QS) (C.Pomerance [46]) и метод решета числового поля (the Number Field Sieve NFS) (J. Pollard [44]). Также Джоном Поллардом были разработаны более медленные, но намного более простые методы, названные ρ – метод и $(p-1)$ – метод Полларда (см. [43]). Фактически, появилось новое направление в современной алгоритмической теории чисел — исследование методов проверки простоты целых чисел и методов нахождения делителей непростых (составных) целых чисел.

Сама же история с разложением 129-значного числа создателей метода RSA закончилась в 1994 г., когда с помощью алгоритма квадратичного решета, реализованного в сети коллективом авторов, возглавляемым А.Ленстрой, было выполнено разложение этого числа на сомножители. Эта процедура потребовала колоссальных усилий. Была задействована сеть, состоящая из 1600 компьютеров, которые проработав 220 дней, подготовили систему линейных уравнений, содержащую более 0,5 млн неизвестных. Потом эта система была решена с помощью суперкомпьютера за 2 дня вычислений.

В настоящее время исследования в области построения быстрых алгоритмов факторизации интенсивно ведутся во всем мире. Ежегодно проводятся десятки конференций по этой тематике, достигаются новые рекорды факторизации длинных чисел, исследуются известные проблемы алгоритмической теории чисел и ставятся новые проблемы. Недавно (в конце 2009 г.) коллективом европейских ученых, возглавляемым Торштеном Кляйнъюнгом [30], был установлен новый рекорд по разложению 768-битового натурального числа с помощью метода решета числового поля. Предыдущий рекорд в 512-бит был установлен в 2000 г., т.е. переход от 512-битовых к 768-битовым числам потребовал почти 10 лет. Поэтому следующий рекорд в 1024 бита при сохранении прежних темпов роста исследований планируется выполнить не ранее, чем в 2020 г.

Наша страна практически полностью устранилась от участия в этом соревновании, что объясняется отсутствием источников финансирования подобных проектов. Другой причиной является отсутствие литературы на русском языке по наиболее современным методам факторизации, какими являются метод квадратичного решета и метод решета числового поля. Некоторые сведения можно получить из монографий А.В.Черемушкина «Лекции по арифметическим функциям в криптографии», МЦНМО, 2002, [79] и О.Н.Василенко «Теоретико-числовые алгоритмы в криптографии», МЦНМО, 2003, [64]. Однако, значительный прогресс, достигнутый в этой области за последние годы, не был отражен в перечисленных изданиях.

Данная книга, сокращая дефицит в этой области, дает первоначальные сведения об основных методах факторизации, используемых в современной теории числовых алгоритмов.

Автор просит читателей присылать замечания и предложения по улучшению содержимого книги, а также сведения об возможных опечатках, встречающихся в тексте, по адресу: ishm2010@yandex.ru

1. Простые числа

Натуральное число называется *простым*, если оно делится только на себя и на 1. Число, не являющееся простым, называется *составным*.

В этой главе мы дадим описание простейших алгоритмов проверки простоты. Однако, прежде чем приступить к описанию основных алгоритмов, дадим некоторые сведения из элементарной теории чисел.

1.1. Модулярная арифметика

Пусть $\mathbf{Z}$ обозначает множество целых чисел, а $p \geq 2$–целое число. Условие, что целое число a делится нацело на b будем записывать через $b|a$.

Определение 1.1. *Говорят, что два целых числа a и b* сравнимы *по модулю p, записывается,*

$$a \equiv b \,(\bmod p),$$

если $p|(a-b)$.

Отношение сравнения по модулю натурального числа обладает следующими свойствами:

1. Симметричность: $a \equiv a \,(\bmod p)$.
2. Рефлексивность: $a \equiv b \,(\bmod p) \to b \equiv a \,(\bmod p)$.
3. Транзитивность: $a \equiv b \,(\bmod p) \,\&\, b \equiv c \,(\bmod p) \to a \equiv c \,(\bmod p)$.

Значит отношение сравнения по модулю является отношением эквивалентности на множестве целых чисел. Классы эквивалентности, образованные целыми числами по этому отношению, называются *вычетами*. Вычет, содержащий число k, обозначается $\overline{k}$. Множество классов вычетов по модулю числа натурального $n > 0$ содержит ровно n элементов, записываемых как $\mathbf{Z}_n = \{\overline{0}, \overline{1}, \ldots \overline{n-1}\}$.

Над вычетами можно выполнять арифметические операции сложения, вычитания, умножения и возведения в степень, а если число n простое или

является некоторой степенью простого числа, то и деление. Будем обозначать множество вычетов по модулю n через $\mathbf{Z}_n$.

Отметим, что для любых a и b выполняется формула $\overline{a} + \overline{b} = \overline{a+b}$ (то же для других перечисленных выше операции), поэтому операции над вычетами выполняются как над обычными числами, приводя результат к значению, принадлежащему интервалу $[0,\, n-1]$ путем выполнения операции вычисления остатка от деления результата на число n (т.е. операции, обозначаемой $\bmod n$). Например, в множестве $\mathbf{Z}_7$ $\overline{2} \cdot \overline{5} = \overline{10} = \overline{3}\,(\bmod 7)$. Чаще пишут просто: $2 \cdot 5 = 3\,(\bmod 7)$.

Множество классов вычетов по модулю n образует структуру, являющуюся *кольцом*. Кольцом K называется непустое множество элементов, на котором определены две арифметические операции *сложения* $+$ и *умножения* $\cdot$, относительно которых выполняются следующие формулы:

1. Ассоциативность по сложению: $(\forall a,\, b,\, c \in K)\ a + (b + c) = (a + b) + c$,
2. Существование нулевого элемента: $(\exists \mathbf{0} \in K)(\forall a \in K)\ a + \mathbf{0} = \mathbf{0} + a = a$,
3. Существование обратного элемента: $(\forall a \in K)(\exists b \in K)\ a + b = b + a = \mathbf{0}$,
4. Ассоциативность по умножению: $(\forall a,\, b,\, c \in K)\, a \cdot (b \cdot c) = (a \cdot b) \cdot c$,
5. Дистрибутивность: $(\forall a,\, b,\, c \in K)\ a \cdot (b + c) = a \cdot b + a \cdot c$,
$(b + c) \cdot a = b \cdot a + c \cdot a$.

Обратный по сложению к a элемент обозначается через $(-a)$. Множество элементов, удовлетворяющих только первым трем свойствам, называется *группой*. Если в группе $< G, + >$ выполняется свойство коммутативности $a + b = b + a$, то группа называется *коммутативной* или *абелевой*. Очевидно, что группа по сложению кольца $\mathbf{Z}_n$ является абелевой группой.

Если модуль n является простым числом или степенью простого числа $n = p^k$, где $k \geq 1$–натуральное число, то множество ненулевых элементов кольца $\mathbf{Z}_n$ (обозначаемое через $\mathbf{Z}_n^*$) образует коммутативную группу по умножению, т.е. существует нейтральный элемент $\mathbf{1}$ $a \cdot \mathbf{1} = \mathbf{1} \cdot a$, и для каждого элемента a имеется обратный по умножению a^{-1} со свойством $a \cdot a^{-1} = \mathbf{1}$.

Алгебраические структуры, содержащие абелеву группы по сложению и группу по умножению, связанные законами дистрибутивности, называются *полями*. Конечные поля называют также полями Галуа по имени гениального французского математика Эвариста Галуа (1811 – 1832), исследовавшего эти поля, и обозначают $GF(q)$. Более подробные сведения о конечных полях читатель может получить из монографии Р.Лидла и Г.Нидеррайтера «Конечные поля» [73].

Пусть G–произвольная группа по умножению.

Определение 1.2. *Порядком элемента a группы G (обозначается через $ord_G(a)$) называется* наименьшее *число k такое, что $a^k = 1$. Порядком группы называется число ее элементов.*

Следующее свойство, связывающее порядки элементов с порядком группы, широко используется в различных алгоритмах, описанных ниже. Эта теорема была доказана знаменитым французским математиком Жозефом Луи Ланранжем (1736–1813).

Теорема 1.1. (Лагранж). *Порядок любого элемента конечной группы является делителем порядка группы.*

Доказательство. Пусть элемент a конечной группы $< G, \cdot >$ имеет порядок $k > 1$. Тогда элементы $a,\ a^2,\ \ldots,\ a^{k-1},\ a^k = 1$ различны и сами образуют группу A, содержащую k элементов и являющуюся подгруппой G. Различные смежные классы $b \cdot A$ для $b \in G$ имеют также мощность k, а объединение их дает в совокупности группу G. Значит, число элементов G равно $k \cdot m$, где m–число смежных классов, откуда вытекает утверждение теоремы.

Пример. Рассмотрим кольцо $\mathbf{Z}_p$ при $p = 29$. Ненулевые элементы этого кольца образуют группу по умножению, порядок которой равен $p - 1 = 28$. По теореме Лагранжа порядок любого элемента a этой группы является делителем 28, т.е. может принимать одно из следующих значений: 1, 2, 4, 7, 14 и 28.

Элемент $a \in G$ называется *примитивным* элементом или *генератором* группы, если его порядок $ord_G(a)$ равен порядку группы.

Не любая группа имеет генератор. Группа, в которой есть генератор, порождается одним элементом и называется *циклической*.

Большинство операций в кольце вычетов $\mathbf{Z}_n$ можно выполнять, выполнив сначала действия с числами, а затем находя остаток от деления результата на модуль n. Однако с операциями возведения в степень и вычисления дискретного (модулярного) логарифма, такой порядок является очень неэффективным. Например, если мы захотим вычислять $2^{199} \pmod{1003}$, используя калькулятор, входящий в состав операционной системы Windows, то результат окажется неверным. В то же время эту же операции несложно выполнить с помощью алгоритма *быстрого возведения в степень по модулю* заданного натурального числа, который мы сейчас опишем.

1.2. Алгоритм быстрого возведения в степень по модулю

Предположим, что требуется вычислить $z = a^b \bmod n$. Рассмотрим следующий алгоритм:

1. Представим b в двоичный системе исчисления: $b = (b_0\, b_1 \ldots b_k)_2$, $b_i \in \{0, 1\}$. Например, $199 = 11000111_2$,

2. Заполним следующую таблицу

b	b_0	b_1	...	b_k
a	a_0	a_1	...	a_k

где $a_0 = a$, $a_{i+1} = \begin{cases} a_i^2 \bmod n, \text{ если } b_{i+1} = 0, \\ a_i^2 \cdot a \bmod n, \text{ если } b_{i+1} = 1 \end{cases}$ для $i \geq 0$.

Результат появится в последней ячейке второй строки.

Пример. Вычислить $2^{199} \bmod 1003$:

b	1	1	0	0	0	1	1	1
c	2	8	64	84	35	444	93	247

Ответ: $2^{199} \bmod 1003 = 247$.

1.3. Решето Эратосфена и критерии простоты

Очевидно, что любое простое число, не равное 2, является нечетным. Существуют признаки делимости целых чисел на различные простые числа, например, чтобы число в десятичном виде делилось на 3 и 9 достаточно, чтобы сумма его цифр делилась на 3 и 9 соответственно. Чтобы число делилось на 5, достаточно, что его последняя цифра была 0 или 5.

Такие частные признаки делимости можно использовать, если нужно уменьшить множество кандидатов проверки на простоту или отсечь заведомо составные числа. Альтернативным способом получения простых чисел является *решето Эратосфена*, приписываемое древнегреческому ученому Эратосфену Киренскому, жившему примерно в 276 - 194 г. до н.э.

Для нахождения множества простых до заранее выбранной верхней границы B мы сначала выписываем последовательность всех нечетных чисел от 3 до B. Затем выбираем первое число в списке, т.е. тройку, и оставляя его в списке, вычеркиваем все кратные 3, начиная с 6. Потом переходим ко второму числу списка (пятерке) и вычеркиваем его кратные, оставив саму пятерку и т.д., пока не дойдем до конца списка. В оставшемся списке будут только простые числа.

Малая теорема Ферма

Знаменитый французский математик Пьер Ферма (1601–1665) доказал теорему, которая известна как *малая теорема Ферма.*

Теорема 1.2. (Малая теорема Ферма) *Если число p–простое, то для любого натурального числа, не сравнимого с p выполняется сравнение*

$$a^{p-1} \equiv 1 \,(\bmod\, p) \tag{1.1}$$

Эта теорема является частным случаем теоремы Лагранжа (теор.1.1). Действительно, при простом p множество ненулевых элементов кольца Z_p образует группу по умножению, имеющую $p-1$ элемент. Будем обозначать это множество через Z_p^*. По теореме Лагранжа порядок любого элемента $a \in Z_p^*$ является делителем порядка $p-1$, откуда $a^{p-1} \equiv 1 \,(\bmod\, p)$.

Из теоремы Ферма сразу следует, что если для некоторого $a < p$ выполнено условие $a^{p-1} \not\equiv 1 \,(\bmod p)$, тогда число p является составным. Однако обращение малой теоремы Ферма не верно – существуют составные числа p, для которых выполняется условие Ферма для каждого a, не сравнимого с p. Такие числа называются числами Кармайкла и будут рассмотрены в разделе 1.19.

Следующая теорема дает критерий проверки простоты числа p и одновременно примитивности корня a:

Теорема 1.3. *(Критерий примитивности и простоты). Если для некоторых a и p выполнены условия:*

1. $a^{p-1} \equiv 1(\bmod p)$,
2. $a^{(p-1)/q} \not\equiv 1(\bmod p)$ *для* $\forall q|(p-1)$,

тогда число p –простое, и a является примитивным корнем поля GF_p (т.е. генератором группы по умножению поля GF_p).

Пример. $n = 1\,022\,333\,835\,329\,657, \quad n-1 = 2 \cdot 2957 \cdot 146\,063 \cdot 292\,877.$

$$3^{n-1} \equiv 1(\bmod n),$$
$$3^{(n-1)/2} \equiv -1 \,(\bmod n),$$
$$3^{(n-1)/2597} \equiv 324224767363906 \,(\bmod n),$$
$$3^{(n-1)/146\,063} \equiv 697302646321792 \,(\bmod n),$$
$$3^{(n-1)/292\,877} \equiv 736785752408036 \,(\bmod n).$$

Поэтому число n в нашем примере является простым, а 3 является примитивным корнем поля Галуа GF_n.

Отметим, что разбиение $n-1$ в произведение простых сомножителей само является очень сложной задачей, поэтому для длинных чисел этот критерий простоты неприменим.

1.4. Метод пробных делений

Метод пробных делений (the trial division) является наиболее простым методом проверки простоты входного составного числа n или нахождения

его делителей. Будем использовать обозначение $\lfloor x \rfloor$ для функции floor(x), равной наибольшему целому числу, не превышающему x (округление вниз). Аналогично, $\lceil x \rceil$ используется для обозначения функции ceil(x), равной наименьшему целому числу, большему или равному x (округление вверх).

Для этого в цикле выполняется пробное деление n на все целые числа от 2 до $\sqrt{n}$:

```
int Tr_div(int n)
{
for(int i = 2; i < ⌊√n⌋; i++)
if (n%i == 0) return i;
return 0}
```

Каждое деление имеет асимптотическую сложность $O(log^2 n)$, поэтому общая сложность метода может быть оценена как $O(n^{1/2} log^2 n)$. Обозначим через L длину двоичного представления числа n, $L = \lceil \log_2 n \rceil$. Тогда можно записать последнюю оценку в более стандартном для теории вычислимости виде:

$$T(n) = O(L^2 \cdot e^{L/2}). \tag{1.2}$$

Значит, алгоритм пробных делений имеет экспоненциальную оценку относительно длины входного числа, поэтому этот метод не может быть использован для тестирования больших чисел.

1.5. Решето Аткина

Решето Аткина — быстрый современный алгоритм нахождения всех простых чисел до заданного целого числа. Это оптимизированная версия старинного решета Эратосфена: решето Аткина проделывает некоторую предварительную работу, а затем вычеркивает числа, кратные квадрату простых. Алгоритм был создан А. Аткиным (A. Atkin) и Д. Бернштейном (D. Bernstein).

Ниже представлена упрощенная версия кода, иллюстрирующая основную идею алгоритма – использование квадратичных форм.

```
int limit = 1000;
int sqr_lim;  bool  is_prime[1001];  int x2,  y2;  int i, j; int n;
// Инициализация решета
sqr_lim =  (int) sqrt((long double) limit);
for (i = 0; i <= limit; i++) is_prime[i] = false;
is_prime[2] = true; is_prime[3] = true;
// Предположительно простые - это целые с нечетным числом
// представлений в данных квадратных формах.
// x2 и y2 - это квадраты i и j (оптимизация).
x2 = 0;
for (i = 1;  i <= sqr_lim; i++) {
    x2 + = 2 * i - 1;
    y2 = 0;
    for (j = 1;  j <= sqr_lim;  j++) {
        y2 += 2 * j - 1;
        n = 4 * x2 + y2;
        if ((n <= limit) && (n % 12 == 1 || n % 12 == 5))
            is_prime[n] = ! is_prime[n];
        // n = 3 * x2 + y2;
        n -= x2; // Оптимизация
        if ((n <= limit) && (n % 12 == 7))
            is_prime[n] = ! is_prime[n];
        // n = 3 * x2 - y2;
        n -= 2 * y2; // Оптимизация
        if ((i > j) && (n <= limit) && (n % 12 == 11))
            is_prime[n] = ! is_prime[n];
        }
    }
// Отсеиваем квадраты простых чисел в интервале [5,  √limit].
// (основной этап не может их отсеять)
for (i = 5;  i <= sqr_lim;  i++) {
        if (is_prime[i]) {
            n = i * i;
        for (j = n;  j <= limit;  j += n) {
            is_prime[j] = false;
        }
    }
}
// Вывод списка простых чисел в консоль.
```

```
printf("2, 3, 5");
for (i = 6; i <= limit; i++) {
// добавлена проверка делимости на 3 и 5. В оригинальной
// версии алгоритма потребности в ней нет.
        if (is_prime[i]) && (i % 3 <> 0) && (i % 5 <> 0){
            printf(" % d " , i); }
}
```

Обоснование алгоритма. Алгоритм полностью игнорирует любые числа, которые делятся на три, пять и семь. Все числа, равные (по модулю 60) 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56 или 58, делятся на два и заведомо не простые. Все числа, равные (по модулю 60) 3, 9, 15, 21, 27, 33, 39, 45, 51 или 57, делятся на три и тоже не являются простыми. Все числа, равные (по модулю 60) 5, 25, 35 или 55, делятся на пять и также не простые. Все эти остатки (по модулю 60) игнорируются.

Все числа, равные (по модулю 60) 1, 13, 17, 29, 37, 41, 49 или 53, имеют остаток от деления на 4 равный 1. Эти числа являются простыми тогда и только тогда, когда количество решений уравнения $4x^2 + y^2 = n$ нечётно и само число не является квадратом (squarefree).

Числа, равные (по модулю 60) 7, 19, 31 или 43, имеют остаток от деления на 6 равный 1. Эти числа являются простыми тогда и только тогда, когда количество решений уравнения $3x^2 + y^2 = n$ нечётно и само число не является квадратом.

Числа, равные (по модулю 60) 11, 23, 47 или 59, имеют остаток от деления на 12 равный 11. Эти числа являются простыми тогда и только тогда, когда количество решений уравнения $3x^2 - y^2 = n$ нечётно и само число не является квадратом.

Ни одно из рассматриваемых чисел не делится на 2, 3 или 5, значит они не могут делиться и на их квадраты. Поэтому проверка того, что число не является квадратом, не включает чисел 22, 32 и 52.

Оценка сложности решета Аткина

По оценке авторов алгоритм имеет асимптотическую сложность

$$O\left(\frac{n}{\ln\ln n}\right)$$

и требует $O(n^{1/2+o(1)})$ бит памяти. Ранее были известны столь же асимптотически быстрые алгоритмы, но они требовали существенно больше памяти.

1.6. Тест Поклингтона

Если у числа $n-1$ найдено один или несколько простых делителей, то это позволяет ограничить область значений простых делителей числа n или даже показать, что n является простым. Следующая теорема подтверждает это наблюдение:

Теорема 1.4. *(H.C. Pocklington). Пусть $n-1 = F \cdot R$, и полное разложение множителя F на простые множители известно. Тогда, если для некоторого $a < n$ выполняются условия:*

1. $a^{n-1} \equiv 1 \pmod{n}$,
2. *Н.О.Д.*$(a^{(n-1)/q}, n) \neq 1$ *для любого* $q|F$,

тогда любой делитель числа n сравним с 1 по модулю F.

Доказательство. Пусть p–простой делитель числа n. Из п.1 предположений теоремы следует, что порядок k элемента a^R в мультипликативной группе поля GF_p является делителем $(n-1)/F = R$. Из второго пункта предположений следует, что k не может быть собственным делителем, т.е. $k = F$. Отсюда $F|(p-1)$, т.е. $p = 1 + m \cdot F$ для некоторого целого m.

Следствие. Если $F > \sqrt{n}$, тогда число n–простое.

Действительно, в этом случае, любой нетривиальный делитель p числа n должен быть больше $\sqrt{n}$, что невозможно.

Пример. Пусть $n = 618\,970\,019\,642\,690\,137\,449\,462\,111$. Число $n-1$ имеет полное разложение вида

$$n - 1 = 2 \cdot 3 \cdot 5 \cdot 17 \cdot 23 \cdot 89 \cdot 353 \cdot 397 \cdot 683 \cdot 2113 \cdot 2\,931\,542\,417.$$

Отметим, что наибольший делитель $n-1$, равный $2\,931\,542\,417$, меньше $\lfloor\sqrt{n}\rfloor = 24\,879\,108\,095\,803$.

Базис $a = 2$ не подходит по условию теоремы, т.к. $2^{(n-1)/q} \equiv 1 \,(\bmod n)$ для всех делителей q. Выполним тест с базой $a = 3$:

$$m = 3^{(n-1)/p} - 1 \equiv 180\ 591\ 065\ 836\ 317\ 083\ 554\ 066\ 745 \neq \pm 1 \,(\bmod n),$$

и, Н.О.Д.$(n,\, m) = 1$. Значит, возможные делители числа n имеют вид $1 + k \cdot p < \sqrt{n}$, откуда, $0 < k < 8486$. Простым перебором всех k можно убедиться, что n не имеет простых делителей, и, значит, является простым.

Можно было также вместо выполнения делений применить теорему для F, равного произведению трех наибольших делителей $n-1$ (для них годится та же база $a = 3$), тогда $F > \lfloor\sqrt{n}\rfloor$, откуда сразу следует, что n–простое.

1.7. Генерация простых чисел

Во многих задачах теории чисел и криптографии возникает необходимость генерации простых чисел заданного размера. Одним из хороших способов порождения простых чисел является способ получения простых чисел с помощью какой-нибудь формулы. Например, еще Л.Эйлер предложил многочлен

$$p(x) = x^2 + x + 41,$$

значения которого в первых 40 членах натурального ряда дают простые числа.

Впрочем, есть гораздо более сильный результат Юрия Матиясевича [1970], который доказал, что существует многочлен с целыми коэффициентами от нескольких переменных, значениями которого будут в точности все простые числа. Приведем этот многочлен (материал взят из лекций С.В. Сизого по теории чисел [77]):

$$F(a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s,t,u,v,w,x,y,z) =$$
$$(k+2)(1-(wz+h+j-q)^2-(2n+p+q+z-e)^2-(a^2y^2-y^2+1-x^2)^2-$$
$$-((e^4+2e^3)(a+1)^2-o^2)^2-(16(k+1)^3(k+2)(n+1)^2+1-f^2)^2-$$
$$-(((a+u^4-u^2a)^2-1)(n+4dy)^2+1-(x+cu)^2)^2-$$
$$-(ai+k+1-l-i^2-((gk+2g+k+1)(h+j)+h-z)^2-$$
$$-(16r^2y^4(a2-1)+1-u^2)^2-(p-m+l(a-n-1)+$$
$$+b(2an+2a-n^2-2n-2))^2-(z-pm+pla-p^2l+t(2ap-p^2-1))^2-$$
$$-(q-x+y(a-p-1)+s(2ap+2a-p^2-2p-2))^2-$$
$$-(a^2l^2-l^2+1-mr)^2-(n+l+v-y)^2)$$

Следует также упомянуть, что с помощью этого многочлена Ю. Матиясевич сумел решить знаменитую 10-ю проблему Гильберта, доказав, что не существует общего алгоритма для определения, имеет ли произвольное диофантово уравнение решение в целых числах. Этот результат послужил развитию целой серии результатов о неразрешимых проблемах в различных областях математики.

Генерация простых чисел, основанная на тесте Поклингтона

Рассмотрим один способ генерации больших простых чисел, основанный на тесте Поклингтона (стр.19) Пусть задано простое число p:

1. Выберем случайным образом чётное число R на промежутке $p \leq R \leq 4p+2$ и определим $n = pR+1$.

2. Проверим число n на отсутствие малых простых делителей, разделив его на малые простые числа.

3. Выполним для числа n тест Миллера-Рабина (см.ниже на с.27) с использованием нескольких различных баз $a < p$. Если при одном из тестов

выяснится, что n–составное число, то выберем новое значение R и повторим вычисления.

Оценка эффективности этого метода зависит от плотности распределения простых чисел и расстояния между соседними простыми числами. Вопрос этот является вовсе не простым (см.разд.1.13) и зависит от справедливости обобщенной гипотезы Римана (ОГР). Если допустить справедливость ОГР, то этот алгоритм является полиномиальным.

Генерация простых чисел с использованием решета Эратосфена

Способ, описанный в предыдущем параграфе, дает большой разброс в поиске простых чисел. Иногда, требуется найти простое число $p(x)$, следующее за данным целым числом x (в пакете Wolfram Mathematica есть специальная функция, вычисляющая $p(x)$ для очень больших чисел x). Для нахождения простых чисел, следующих за заданным числом x, можно использовать следующий алгоритм:

1. Выписываем множество чисел $\{n,\ n+2,\ n+4,\ ...,\ n+2m\}$, где $n \geq x$– наименьшее нечетное число, $n+2m$ – верхняя граница интервала.

2. Выполним просеивание этого интервала с использованием решета Эратосфена или Аткина (см. с.16) с помощью множества небольших простых чисел $\{3, 5, ..., p_k\}$, ограниченного сверху границей B. При $B = 10$, отсеется примерно половина кандидатов. При $p_k < 1000$, будет отсеяно $5/6$ всех кандидатов.

3. Далее, проверим оставшиеся кандидаты с помощью теста Миллера–Рабина.

Пример. При $x \approx 10^{260}$ была выбрана база для просеивания, ограниченная сверху числом $B = 1000$. После 67 попыток было найдено простое число $x + 782$. При $B = 8000$ число попыток уменьшилось до 50, а при $B = 50\,000$ до 36. Значит, большое увеличение границы B не обеспечивает выигрыша алгоритма в целом, т.к. время, сэкономленное за счет уменьшения числа проб, не компенсирует время, затраченное на выполнение

просеивания.

1.8. Расширенный алгоритм Евклида

Расширенный алгоритм Евклида (РАЕ) используется во многих криптографических и теоретико–числовых алгоритмах. Он состоит из двух частей. В первой части алгоритма для заданных целых чисел A и B вычисляется их наибольший общий делитель (greatest common divisor d) d. Вычисление Н.О.Д. натуральных чисел A и B выполняется по рекуррентной формуле:

$$\text{Н.О.Д.}(A,\, B) = \text{Н.О.Д.}(B,\, A \bmod B), \tag{1.3}$$

где $A \bmod B$ означает операцию вычисления остатка при целочисленном делении A на B. Производится последовательное использование этой формулы, пока остаток от деления первого операнда на второй не станет равным 0. Последнее ненулевое значение второго операнда и есть искомый общий делитель:

```
int Euclid(int A, B)
{
while (A mod B !=0) {
   int C=A mod B;
    A=B; B=C ; }
return B;
}
```

Для решения уравнений вида $Ax+By = d$, где A, B – заданные числа, а d – их наибольший общий делитель, используется *расширенный* алгоритм Евклида. Первая часть РАЕ в результате которой мы находим Н.О.Д. d, выполняется также, как описано выше. Значения A, B, а также целую часть и остаток от деления A на B сохраняются в таблице, содержащей 4 столбца.

Во второй части работы алгоритма к таблице добавляются два новых столбца, озаглавленных x и y. Поместим в последнюю строчку столбцов x и y значения 0 и 1. Затем, считая значения x_{i+1} y_{i+1} известными,

последовательно вычисляем значения x_i и y_i, $i \geq 0$, по формулам:

$$x_i = y_{i+1}, \qquad y_i = x_{i+1} - y_{i+1} \cdot (A\,div\,B)_i$$

Пример. Решить уравнение $72x + 25y = 1$. Помещаем в первую строчку значения $A = 72$, $B = 25$. Вычисляем $A\,mod\,B$ – остаток от деления A на B, и $A\,div\,B$ – целую часть от деления A на B. Потом переносим значения B и $A \bmod B$ на строчку вниз и на одну клетку влево. Повторяем вычисления во второй строке. Продолжаем вычисления, пока значение в столбце $A\,mod\,B$ не станет равным 0. Тогда заносим в последнюю строчку столбцов x и y значения 0 и 1, и ведем вычисление снизу вверх по формулам, описанным выше.

A	B	A mod B	A div B	x	y
72	25	22	2	8	-25
25	22	3	1	-7	8
22	3	1	7	1	-7
3	1	0	3	0	1

Ответ: Н.О.Д.$(72,\ 25) = 1$ – последнее значение в столбце B. Пара $(x, y) = (8,\ -25)$, дающая решение уравнению $72x + 25y = 1$, берется из первой строки таблицы.

Оценка производительности алгоритма Евклида

Расширенный алгоритм Евклида используется во многих криптографических методах, поэтому оценка его производительности играет важную роль в расчетах эффективности криптографических алгоритмов.

Основополагающим фактором в оценке РАЕ является число итераций в главном цикле вычисления новой пары $(A;\ B)$, или, другими словами, число строчек в таблице вычисления вычислений. Чтобы оценить это число, заметим, что на шаге k произвольной итерации возможны два случая:

Случай 1. $B < A/2$. На шаге $k + 1$ новое значение A, равное предыдущему B, будет меньше, чем $A/2$.

Случай 2. $B \geq A/2$. Остаток $r = A \bmod B = A - B$ будет меньше $A/2$, и на шаге $k+2$ новое значение A станет равным остатку $r < A/2$.

В любом случае, после каждой пары итераций первый аргумент A уменьшается более, чем в 2 раза, значит, общее число итераций не может быть больше, чем $2\log_2 A$. Число операций на каждой итерации постоянно (как при прямом ходе, так и при подъеме при вычислении коэффициентов уравнения $Ax+By=d$), поэтому, оценка РАЕ равна $O(L)$, где $L = \lceil \log_2 A \rceil$ – длина двоичного представления меньшего из чисел.

Эта оценка не является завышенной, т.к. существует последовательность *чисел Фибоначчи*, $\{F_n\}$, на парах соседних элементов которых и достигается эта верхняя оценка. Последовательность или ряд Фибоначчи определяется следующими формулами:

$$F_0 = 1, \quad F_1 = 1, \quad F_{n+2} = F_n + F_{n+1}, \ n \geq 2.$$

Выпишем начальный интервал этого ряда:

$S = \{1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ 233,\ 377,\ 610,\ 987,\ 1597,\ 2584,\ 4181\}$.

1.9. Символ Лежандра

Определение 1.3. *Пусть $n > 1$ –целое число. Число a, принадлежащее интервалу $[0,\ n-1]$ называется* квадратичным вычетом *по модулю n, если найдется целое число x такое, что $x^2 \equiv a \pmod n$.*

Если такого x не существует, то a называется *квадратичным невычетом.* Отметим, что ровно половина элементов из интервала $[0,\ n-1]$ является квадратичными вычетами.

Условия того, является ли a квадратичным вычетом по простому модулю p, проверяется с помощью, так называемого, символа Лежандра:

$$\left(\frac{a}{p}\right) = \begin{cases} 1, \text{ если } (\exists\, x)\, x^2 \equiv a \bmod p, \\ -1, \text{ если не}(\exists\, x)\, x^2 \equiv a \bmod p, \\ 0, \text{ если } p \mid a. \end{cases} \tag{1.4}$$

Вычисление символа Лежандра может быть выполнено по следующей формуле, полученной Леонардом Эйлером:

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \mod n. \tag{1.5}$$

Однако использование этой формулы на практике сопряжено с вычислениями больших степеней, поэтому предпочтительнее пользоваться законом квадратичной взаимности, доказанный Карлом Гауссом в возрасте 17 лет.

Закон квадратичной взаимности: для любых нечетных простых чисел p и q выполняется формула

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)(-1)^{(p-1)(q-1)/4}.$$

Иначе говоря,

$$\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right), \text{ если } p \equiv q \equiv 3 \bmod 4, \quad \text{и} \quad \left(\frac{q}{p}\right) = \left(\frac{p}{q}\right), \text{ иначе.}$$

Гаусс в течение своей жизни неоднократно возвращался к этому закону и получил несколько его доказательств, основанных на совершенно различных идеях.

Для быстрого вычисления символа Лежандра полезными являются также следующие формулы:

$$\left(\frac{q}{p}\right) = \left(\frac{q \bmod p}{p}\right), \qquad \left(\frac{q \cdot r}{p}\right) = \left(\frac{q}{p}\right) \cdot \left(\frac{r}{p}\right), \qquad \left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} \mod n.$$

Пример. Вычислить $(15/17)$:

$$\left(\frac{15}{17}\right) = \left(\frac{3}{17}\right) \cdot \left(\frac{5}{17}\right) = \left(\frac{2}{3}\right) \cdot \left(\frac{2}{5}\right) = (-1) \cdot (-1)^3 = 1$$

Для составных чисел n используется символ Якоби, который является обобщением символа Лежандра на произвольные целые числа и обладает следующим свойством:

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{r_1} \cdot \left(\frac{a}{p_1}\right)^{r_2} \cdot \ldots \left(\frac{a}{p_k}\right)^{r_k}, \tag{1.6}$$

где $n = p_1^{r_1} \cdot p_2^{r_2} \cdot \quad \ldots \, p_k^{r_k}$–разложение n в произведение степеней простых чисел.

1.10. Тест простоты Миллера–Рабина

В качестве критерия проверки, является ли заданное число n простым или составным, может служить следующая теорема:

Теорема 1.5. *(Критерий непростоты) Нечетное число $n \geq 3$ является* составным *тогда и только тогда, когда n является либо полным квадратом, либо найдутся два натуральных числа x и y такие, что*

$$x \not\equiv \pm y \,(\bmod n),\ \text{и}\ x^2 \equiv y^2 \,(\bmod n). \tag{1.7}$$

Доказательство. Если выполняются условия (1.7) , то Н.О.Д.$(n, x^2 - y^2) \neq 1,\ \neq n$. Обратно, если $n = p \cdot q$, где $p > q$, то определим $x = (p+q)/2$, $y = (p-q)/2$. Очевидно, x и y удовлетворяют (1.7).

На критерии непростоты основан известный вероятностный тест Миллера–Рабина, который старается найти пару x и $y = 1$, удовлетворяющие критерию непростоты.

Пусть n–число, которое необходимо проверить на простоту. Представим $n-1$ в виде $n-1 = 2^s \cdot d$, где d–нечетно. Назовем произвольное число $a \in \mathbf{Z}_n^*$ *свидетелем простоты* n, если выполняет одно из следующих условий:

$$1.\ x = a^d \equiv \pm 1 \,(\bmod n),\ \text{или}$$

$$2.\ (\exists k,\ 0 < k < s)\ x^{2^k} \equiv -1 \,(\bmod n). \tag{1.8}$$

В противном случае, назовем a *свидетелем непростоты* n.

Докажем сначала, что если для числа n найдется хотя–бы один свидетель его непростоты, то n–составное.

Действительно, пусть для некоторого $a \in \mathbf{Z}_n^*$ не выполнен ни один из п.1–2 условий (1.8), тогда последовательность

$$x_0 = a^d (\bmod n),\ x_1 = x_0^2 (\bmod n),\ \ldots,\ x_{s-1} = x_{s-2}^2 (\bmod n)$$

не содержит -1. Вычислим x_s, равное $x_{s-1}^2 (\bmod\ n)$. Оно не равно 1. Но если n–простое, то $x_s = a^{d \cdot 2^s} = a^{n-1}$ должно равняться по малой теореме Ферма единице. Значит, число n – составное.

Для оценки эффективности теста Миллера–Рабина определим понятие функции Эйлера.

Функция Эйлера $\varphi(n)$

Леонард Эйлер ввел в обиход функцию φ, определенную на целых положительных числах, значением которой на аргументе n является количество положительных чисел, меньших n и взаимно-простых с n.

Очевидно, что для всех $n > 1$ $\varphi(n) < n$. Для простого числа p значение $\varphi(n)$ равно $p-1$, а для числа n, являющегося произведением двух сомножителей $n = n_1 \cdot n_2$, $\varphi(n) = \varphi(n_1) \cdot \varphi(n_2)$.

Рабин доказал теорему о том, что если нечетное число $n > 2$–составное, то множество свидетелей его простоты имеет мощность не более $\varphi(n)/4 < n/4$. Отсюда следует, что если при проверке k произвольно выбранных чисел $a < n$ все они окажутся свидетелями простоты n, то n–простое с вероятностью ошибки, не превышающей 4^{-k}. На этом наблюдении строится следующий тест Миллера–Рабина.

Тест Миллера–Рабина

Пусть число $n > 2$–нечетно и $n-1 = 2^s \cdot d$, где d–нечетно. Для каждого числа a от 2 до $r+1$, где r – число проверок в тесте, выполним следующие действия:

1. Вычислим $x_0 = a^d (\mathrm{mod}\, n)$.

2. Проверим условие $x_0 \in \{1, n-1\}$. Если оно выполнится, тогда a–свидетель простоты. Перейдем к следующему a.

3. Иначе проверим, содержится ли число $n-1$ в последовательность $\{x_1,\ x_2,\ ...,\ x_{s-1}\}$, где каждый последующий x вычисляется по формуле $x_{i+1} = x_i^2 \,(\mathrm{mod}\, n)$.

Если ответ положительный, то a–свидетель простоты. Перейдем к следующему $a \le r+1$.

Иначе, найден свидетель непростоты n. Завершаем тест с сообщением «число n–составное».

Если после r проверок окажется r свидетелей простоты, то заканчиваем тест с сообщением «n–вероятно простое».

Пример. Пусть $n = 1729$–число Кармайкла (см. стр. 50). Разложим $n - 1 = 2^6 \cdot 3^3$. Выполним тест Миллера–Рабина для $a = 2$:

$x_0 = 2^{27} \bmod 1729 = 645 \neq 1, \neq n - 1,$

$x_1 = x_0^2 \bmod 1729 = 645^2 \bmod 1729 = 1065.$

$x_2 = x_1^2 \bmod 1729 = 1065^2 \bmod 1729 = 1.$

Последующие элементы $\{x_i\}$ для $i = 3, 4, 5$ равны 1, и последовательность $\{x_1,\ x_2,\ ...,\ x_{s-1}\}$, не содержит $n - 1$. Значит, 2 является свидетелем непростоты n, и $n = 1729$ – составное число.

Оценка эффективности теста Миллера–Рабина

Следующая лемма была доказана Рабином в предположении справедливости обобщенной гипотезы Римана о распределении простых чисел (с. 33):

Лемма 1.1. *Пусть число n–нечетное число, и $n-1 = 2^s \cdot d$, где d–нечетно. Если для всех x, $0 < x < 2 \cdot (log_2 n)^2$ выполняется $x^d \equiv 1 (\mathrm{mod}\, n)$, или $x^{2^k \cdot d} \equiv -1 (\mathrm{mod}\, n)$ для некоторого $0 \leq k < s$. Тогда число n является простым.*

Оценка, приведенная в этой лемме, является полиномиальной, однако, с теоретической точки зрения она не может быть использована, пока не доказана обобщенная гипотеза Римана, которая на сегодняшний день является самой известной из нерешенных «проблем тысячелетия». Кроме того, для практических расчетов эта оценка является сильно завышенной. Вместо нее обычно используется граница порядка $O(log_2 n)$ (см. [50]).

1.11. Вероятностный тест простоты Соловея–Штрассена

Тест Соловея — Штрассена опирается на малую теорему Ферма (разд. 1.2) и свойства символа Якоби (разд. 1.9):

Теорема 1.6. *Если n — нечетное составное число, то количество целых чисел a, взаимно простых с n и меньших n, удовлетворяющих сравнению*

$$a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \pmod{n}, \tag{1.9}$$

не превосходит $n/2$.

Алгоритм Соловея — Штрассена

Сначала для алгоритм Соловея — Штрассена выбирается целое число $k \geq 1$. Тест проверки простоты числа n состоит из k отдельных раундов. В каждом раунде выполняются следующие действия:

1. Случайным образом выбирается число $a < n$, и вычисляется $d = \text{Н.О.Д.}(a, n)$.

2. Если $d > 1$, то выносится решение о том, что n составное. Иначе проверяется сравнение (1.9). Если оно не выполнено, то n - составное. Иначе, a является свидетелем простоты числа n.

Если после завершения k раундов найдено k свидетелей простоты, то делаем заключение «n–вероятно простое число».

Вычислительная сложность и эффективность теста

В каждом раунде вероятность отсеять составное число больше $1/2$, поэтому через k раундов тест Соловея–Штрассена определяет простое число с вероятностью ошибки, меньшей 2^{-k}. Поэтому этот тест сравним по эффективности с тестом Ферма, но имеет преимущество перед тестом Ферма в том, что он отсеивает все числа Кармайкла (см.разд. 1.19).

С другой стороны, он проигрывает тесту Миллера–Рабина, который за k раундов имеет ошибку, меньшую 4^{-k}.

Общая вычислительная сложность алгоритма оценивается как $O(k \log_2 n)$.

1.12. Полиномиальный критерий простоты AKS

Одной из важных проблем, долгое время стоявших перед исследователями, была проблема построения детерминированного алгоритма проверки простоты натуральных чисел, имеющего полиномиальную оценку времени работы. Алгоритм Миллера–Рабина, упомянутый в предыдущем разделе, имеет полиномиальную оценку, но не является детерминированным. Другие тесты, например тест Поклингтона (разд.1.6), являются детерминированными, но не имеют полиномиальной оценки.

В 2004 г. тремя молодыми индийскими математиками Агравелой, Каялом и Саксеной ([1]) был разработан детерминированный полиномиальный безусловный тест AKS проверки простоты заданного натурального числа. Тест AKS основывается на следующей теореме:

Теорема 1.7. *(Agrawel, Kayal, Saxena [2004].) Пусть n–нечетное натуральное число, r–простое число и выполнены условия:*

1. *Число n не делится ни на одно из чисел, меньших или равных r,*
2. *Порядок n в мультипликативной группе $\mathbf{Z}_p^*$ поля GF_p не меньше $(\log_2(n))^2$,*
3. *Для всех a, $0 \le a \le r$, выполнена формула*

$$(X+a)^n \equiv X^n + a \text{ в кольце многочленов } \mathbf{Z}_n[X]/\frac{X^r-1}{X-1}\ . \qquad (1.10)$$

Тогда число n является простым.

В этой теореме используются вычисления в кольце многочленов $\mathbf{Z}_n[X]$ с коэффициентами, ограниченными сверху числом n, факторизованных по модулю многочлена деления круга

$$\Phi_r(X) = \frac{X^r-1}{X-1} = X^{r-1} + X^{r-2} + \ ... \ + X \ + \ 1.$$

Конечно, если n–просто, то эквивалентность $(X+a)^n \equiv X^n + a \bmod n$ в силу малой теоремы Ферма выполняется и в кольце $\mathbf{Z}_n[X]$, однако эти

вычисления слишком громоздки, чтобы их можно было реально выполнить. Суть замечательной идеи Агравелы, Каялы и Саксены состояла в том, чтобы заменить кольцо $\mathbf{Z}_n[X]$ на гораздо меньшее кольцо $\mathbf{Z}_n[X]/\Phi_r(X)$.

На этой теореме основан следующий тест проверки простоты числа n:

1. Проверим, что n не является полным квадратом,

2. Используя числа $r = 2, 3, 5, \ldots$, найдем наименьшее простое число r такое, что r не является делителем n, и не является делителем $n^i - 1$ для всех $i \in \{0, 1, 2, \ldots, (\log_2 n)^2)\}$.

3. Проверим, что выполнены условия пункта 3 теоремы.

Если эти условия выполнены, то n–простое, иначе n–составное.

Замечание. Несмотря на то, что тест AKS явился решением крупной и долгостоявшей научной проблемы, он является не слишком удобным с практической точки зрения. Проверка условий пункта 3 является настолько громоздкой, что общая оценка времени работы алгоритма достигает $\mathrm{O}(\log^{18} n)$ (см. Д. Вентури. Лекции по алгоритмической теории чисел [53]). Поэтому этот тест следует применять лишь в тех случаях, когда надо получить *гарантированное* доказательство того, что число n является простым.

1.13. Распределение простых чисел

Известно, что простых чисел бесконечно много. Действительно, согласно известной теореме Евклида, если $\{2, 3, 5, \ldots, p_m \leq B\}$ – множество всех простых чисел, меньших некоторой границы B, то число M

$$M = \prod_{i=1}^{m} p_i + 1$$

снова является простым. Этим методом можно порождать сколь угодно большие числа. Недостатком такого метода будет только то, что члены последовательности будут расположены слишком редко.

Между тем, простых чисел достаточно много. Обозначим через $\pi(x)$ количество простых чисел на интервале от 1 до x.

Еще в 1796 г. французский математик Адриен Мари Лежандр (1752–1833) предположил, что функция $\pi(x)$ может быть приближена выражением

$$\pi(x) \approx \frac{x}{\ln x} - B, \text{ где } B \approx 1,08.$$

В середине XIX столетия П.Л. Чебышев (1821–1894) указал формулу для количества простых чисел $\pi(x)$ на интервале от 1 до x:

$$A \cdot \frac{x}{\ln x} < \pi(x) < B \cdot \frac{x}{\ln x},$$

где $A = 0,921$, $B = 1,06$, и доказал, что предел отношения $\pi(x)$ к $x/\ln x$, если он существует, равен 1.

Полвека спустя в 1896 г., развивая идеи работ Чебышева, Адамар и Валле-Пуссен одновременно и независимо доказывают теорему о том, что предел отношения $\pi(x)$ к $x/\ln x$ существует и равен 1.

Приведем здесь таблицу распределения функции $\pi(x)$ на интервалах $[1; 10^k]$ для начальных значений k:

x	10^2	10^3	10^4	10^6	10^8	10^{12}
$\pi(x)$	25	168	1 229	78 498	5 761 455	37 607 912 018

Распределение простых чисел связано также со знаменитой *гипотезой Римана*. Эта гипотеза вошла в список 7 задач тысячелетия, за решение каждой из которых Математический институт Клэя (Clay Mathematics Institute, Кембридж, Массачусетс) выдвинул премию в 1 млн. долл. США, и касается распределения нулей дзета-функции Римана, являющейся решением функционального уравнения

$$\zeta(s) = 2^s \pi^{s-1} \cdot \sin\frac{\pi s}{2} \cdot \Gamma(1-s) \cdot \zeta(1-s),$$

или представленной явным выражением

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s},$$

где $\mu(n)$–функция Мебиуса.

В 1901 г. Хельге фон Кох показал, что гипотеза Римана эквивалентна следующему утверждению о распределении простых чисел:

$$\pi(x) = \int_2^x \frac{dt}{\ln t} + O(\sqrt{x}\ln x) \text{ при } x \to \infty$$

1.14. Извлечение квадратного корня в конечных полях

В реализациях методов квадратичного решета и решета числового поля, описываемых в 4-й и 5-й главах, будет использован алгоритм извлечения квадратного корня в конечных полях, разработанный Шенксом и Тоннелли. Опишем данный алгоритм в этом параграфе.

Рассмотрим конечное поле GF_p, $p > 2$, и элемент a, являющийся квадратичным вычетом по модулю p. Требуется найти x такое, что $a \equiv x^2 \,(\bmod p)$.

Представим число $p-1$ в виде $p-1 = 2^r \cdot s$, где s–нечетно. Заметим, что поскольку $p-1$–четно, $r \geq 1$. Пусть z–квадратичный невычет по модулю p. Вычислим $y = z^s \mod p$. Поскольку порядок любого элемента является делителем числа $2^r \cdot s$, то порядок y является делителем 2^r, откуда $y^{2^r} \equiv 1 \,(\bmod p)$. Можно также показать, что $y^{2^{r-1}} \equiv -1 \,(\bmod p)$, т.е. порядок элемента y равен в точности 2^r. Вычислим далее элементы

$$\lambda_0 = a^s \,(\bmod\ p),\ \ w_0 = a^{(s+1)/2} \,(\bmod\ p). \tag{1.11}$$

Заметим, что

$$w_0^2 \equiv a \cdot \lambda_0 \,(\bmod p) \ \text{ и } \ x^2 \equiv a (\bmod p) \to x^{2s} \equiv a^s = \lambda_0 (\bmod p). \tag{1.12}$$

Поскольку порядок элемента x^s является делителем 2^r, то порядок λ_0 является делителем 2^{r-1}. Идея метода Шенкса–Тоннелли состоит в построении последовательности пар чисел (λ_i, w_i), удовлетворяющих условию

$$w_i^2 \equiv a \cdot \lambda_i \,(\bmod p),\ \ i = 0, 1, 2, \ldots, \tag{1.13}$$

причем порядок λ_{i+1} является собственным делителем порядка λ_i, до тех пор, пока порядок очередного λ_i не окажется равным 0. Тогда для найденного i выполнятся условия $\lambda_i = 1$ и

$$w_i^2 \equiv a \ (\bmod p),$$

откуда $x = w_i$ является искомым корнем.

Поскольку исходные значения (λ_0, w_0), удовлетворяющие (1.13), согласно (1.12), уже определены, то осталось просто описать формулы для вычисления значений (λ_{i+1}, w_{i+1}):

$$\lambda_{i+1} = \lambda_i \cdot y^{2^{r-m}}, \quad w_{i+1} = w_i \cdot y^{2^{r-m-1}}, \tag{1.14}$$

где 2^m–порядок элемента λ_i.

Пример. Рассмотрим пример вычисления квадратного корня из $a = 2$ в простом поле GF_p при $p = 41$:

1. Имеем, $p - 1 = 40 = 2^3 \cdot 5$, откуда, $s = 5$, $r = 3$.

2. Вычислим исходные значения (λ_0, w_0) по формулам (1.11):

$$\lambda_0 = a^s \ (\bmod \ p) = 2^5 \ (\bmod \ 41) = 32,$$

$$w_0 = a^{(s+1)/2} \ (\bmod \ p) = 2^3 \ (\bmod \ 41) = 8.$$

3. Найдем порядок элемента λ_0:

$$\lambda_0^2 \bmod \ p = 32^2 \ \bmod \ 41 = 40 \equiv -1 \ (\bmod \ 41), \ \lambda_0^4 \equiv 1 \ \bmod \ p.$$

Отсюда, $ord(\lambda_0) = 2^m = 4$, $m = 2$.

4. Будем искать квадратичный невычет. Вычислим символ Лежандра для $z = 3$:

$$\left(\frac{z}{p}\right) = \left(\frac{3}{41}\right) = \left(\frac{41 \bmod 3}{3}\right)(-1)^{(41-1)(3-1)/2} = \left(\frac{2}{3}\right) = -1,$$

значит, $z = 3$ является квадратичным невычетом и может быть использован для вычисления пар (λ_{i+1}, w_{i+1}).

5. Найдем $y = z^s (\bmod\ p) = 3^5 (\bmod\ 41) = 38$.

6. Вычислим степень, в которую надо возводить y:

$$d = 2^{r-m} = 2^{3-2} = 2, \quad y^d = 3^2 = 9.$$

7. Вычислим $\lambda_1 = \lambda_0 \cdot y^d (\bmod\ p) = 32 \cdot 9 (\bmod\ 41) = 1$, $w_1 = w_0 \cdot y^{d-1} (\bmod\ p) = 8 \cdot 3 (\bmod\ 41) = 24$. Поскольку очередное λ_i оказалось равным 1, то процедура закончена. Корень $x = w_1 = 24$. Выполним проверку:

$$x^2 \bmod\ p = 24^2 \bmod\ 41 = 2 = a.$$

1.15. Китайская теорема об остатках

Китайская теорема об остатках позволяет вычислить целое число, если известны его остатки по нескольким простым модулям. Впервые эта теорема была упомянута в трактате китайского математика Сунь Цзы, примерно в третьим веком до н.э.

Теорема 1.8. *Если натуральные числа $m_1, m_2, ..., m_n$ попарно взаимно просты, то для любых целых $r_1, r_2, ..., r_n$ таких, что $0 \le r_1 < m_i$ при всех i, найдётся число x, которое при делении на m_i даёт остаток r_i при всех $1 \le i \le n$. Более того, любые два таких числа x_1 и x_2 удовлетворяют уравнению*

$$x_1 \equiv x_2 \ (\bmod m), \ \textit{где}\ m = m_1 \cdot m_2 \cdot ... \cdot m_n.$$

В трактате другого китайского математика Джу Шао Квина (Jiushao Qin) (1247 г. н.э.) дается формула для вычисления числа x, удовлетворяющего теореме:

$$x = \sum_{i=0}^{n} r_i \cdot e_i, \ \text{где}\ e_i = \frac{m}{m_i} \cdot \left(\left(\frac{m}{m_i} \right)^{-1} \bmod\ m_i \right), \ 1 \le i \le n. \qquad (1.15)$$

Заметим, что поскольку число m_i взаимно просто с m/m_i , то обратное число в формуле для e_i всегда существует для $1 \le i \le n$. Кроме того, имеют место равенства

$$\begin{cases} e_i \cdot e_i \equiv e_i \ (\bmod m), \\ e_i \cdot e_j \equiv 0 \ (\bmod m) \text{ при } i \ne j, \end{cases}$$

т.е. компоненты e_i взаимно ортогональны по модулю m.

Алгоритм Гарнера

Для вычисления x может быть использован алгоритм Гарнера (Garner' algorithm), согласно которому x можно вычислить как n-й член последовательности $\{x_i\}$. Последовательности $\{x_i\}$, $\{y_i\}$ строятся по следующим формулам:

$$\begin{cases} y_1 = x_1 = r_1, \\ y_{i+1} = \frac{r_{i+1} - x_i}{m_1 \cdot m_2 \cdot \ldots \cdot m_i} \mod m_{i+1}, \\ x_{i+1} = x_i + y_{i+1} \cdot m_1 \cdot m_2 \cdot \ldots \cdot m_i. \end{cases} \tag{1.16}$$

Достоинство этого алгоритма заключается в том, что вычисление каждого последующей пары $(x_{i+1},\, y_{i+1})$ использует только одно предыдущее значение $(x_i,\, y_i)$, что позволяет последовательно уточнять значения корня x.

Пример. Найти наименьшее положительное x, удовлетворяющее системе уравнение:

$$\begin{cases} x \equiv 2\,(\bmod 3\,) \\ x \equiv 5\,(\bmod 7\,) \\ x \equiv 4\,(\bmod 11\,). \end{cases}$$

Решение. В нашем примере $m_1 = 2$, $m_2 = 7$, $m_3 = 11$, $r_1 = 2$, $r_2 = 5$, $r_3 = 4$. Будем вычислять последовательно y_i и x_i, $i = 1, 2, 3$:

$y_1 = x_1 = 2,$
$y_2 = (r_2 - x_1) \cdot (m_1)^{-1} \mod m_2 = (5-2) \cdot (3)^{-1} \mod 7 = 1$
$x_2 = x_1 + (y_2 \cdot m_1 \mod m_2) = 2 + (1 \cdot 3 \mod 7) = 5,$
$y_3 = (r_3 - x_2) \cdot (m_1 \cdot m_2)^{-1} \mod m_3 = (4-5) \cdot 21^{-1} \mod 11 = 1,$
$x_3 = x_2 + y_3 \cdot m_1 \cdot m_2 = 5 + 1 \cdot 3 \cdot 7 = 26.$

Ответ: $x = 26$.

1.16. π, e и другие известные константы

В этом разделе приведем значения наиболее известных числовых констант математики. Большинство из них были введены Леонардом Эйлером.

Число π

$\pi = 3,14159265358979323846264338327950288419716939937510\ldots$

– математическая константа, выражающая отношение длины окружности к длине её диаметра. Обозначается буквой греческого алфавита «пи». Существует несколько различных представлений π в виде ряда. Приведем одно из них, открытую в 1997 г. Саймоном Плаффом (Simon Plouffe):

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16k^2} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right)$$

Основание натурального логарифма e

$e = 2,71828182845904523536028747135266249775 7\ldots$

– математическая константа, основание натурального логарифма, трансцендентное число. Иногда число e называют числом Эйлера или числом Непера. Обозначается строчной латинской буквой «e». Число e может быть определено несколькими способами:

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \quad \text{или} \quad e = \sum_{n=1}^{\infty} \frac{1}{n!}$$

Знаменитая формула Эйлера связывает три замечательных числа e, π и мнимую единицу $i = \sqrt{-1}$:

$$e^{ix} = \cos x + i \sin x, \quad \text{откуда,} \quad e^{i\pi} = -1$$

Другие константы

$\sqrt{2} = 1,41421356237309504880168872421 0\ldots$

$\ln 2 = 0,69314718055994530941723212145 8\ldots$

$\boxed{\log_2 e = 1,44269504088896340735992468100...}$

$\boxed{\ln 10 = 2,30258509299404568401799145468 4...}$

1.17. Открытые проблемы теории чисел

До сих пор в математике существует много открытых проблем, связанных с простыми числами. Наиболее известные из которых были перечислены немецким математиком Эдмундом Ландау (1877–1938) на Пятом Международном математическом конгрессе в 1912 г. (Эдмунд Ландау являлся отдаленным родственником известного советского физика Льва Давыдовича Ландау (1908–1968)). Ни одна из этих проблем, сформулированных Э.Ландау, не решена до сих пор.

- Проблема Гольдбаха (первая проблема Ландау): доказать или опровергнуть, что каждое чётное число, большее двух, может быть представлено в виде суммы двух простых чисел, а каждое нечётное число, большее 5, может быть представлено в виде суммы трёх простых чисел.

- Вторая проблема Ландау: бесконечно ли множество «простых близнецов» (т.е. пар простых чисел, разность между которыми равна 2) ?

- Гипотеза Лежандра (третья проблема Ландау): верно ли, что между n^2 и $(n+1)^2$ всегда найдётся простое число?

- Четвёртая проблема Ландау: бесконечно ли множество простых чисел вида n^2+1?

Открытой проблемой является также существование бесконечного количества простых чисел во многих целочисленных последовательностях, включая числа Фибоначчи, числа Ферма и т. д. Рассмотрим в следующих разделах две знаменитые гипотезы теории чисел – гипотезу Гольдбаха–Эйлера и гипотезу о простых числах – близнецах.

Гипотеза Гольдбаха

В 1742 г. прусский математик Кристиан Гольдбах послал письмо Леонарду Эйлеру, в котором он высказал следующее предположение:

Каждое нечётное число большее 5 можно представить в виде суммы трёх простых чисел.

Эйлер заинтересовался проблемой и выдвинул более сильную гипотезу:

Каждое чётное число большее двух можно представить в виде суммы двух простых чисел.

Первое утверждение называется тернарной проблемой Гольдбаха, второе — бинарной проблемой Гольдбаха (или проблемой Эйлера). Из справедливости утверждения бинарной проблемы Гольдбаха автоматически следует справедливость тернарной проблемы Гольдбаха: если каждое чётное число > 2 есть сумма двух простых чисел, то добавляя 3 к каждому чётному числу, можно получить все нечётные числа > 5.

В 1923 г. математики Харди и Литлвуд показали, что в случае справедливости обобщенной гипотезы Римана тернарная проблема Гольдбаха верна для всех достаточно больших нечётных чисел.

В 1937 г. советский математик, академик Иван Матвеевич Виноградов (1891–2007) представил доказательство, не зависящее от справедливости гипотезы Римана того, что любое достаточно большое нечётное число может быть представлено в виде суммы трёх простых. Сам Виноградов не дал явной оценки для этого «достаточно большого числа», но его студент К. Бороздин доказал, что оно не превышает $3^{3^{15}}$. Это число содержит 6 миллионов цифр, что делает практически невозможным прямую проверку всех меньших простых чисел. Тем не менее, за решение этой проблемы Виноградов получил Сталинскую премию и звание Героя Социалистического Труда.

В дальнейшем результат Виноградова многократно улучшали, пока в 1989 Ванг и Чен не опустили нижнюю грань до $3,3 \cdot 10^{43000}$, что, по-прежнему находится вне пределов досягаемости для явной проверки даже при современном развитии вычислительной техники.

В 1997 г. Дезуйе, Эффингер, Тэ Риле и Зиновьев показали, что обобщённая гипотеза Римана влечёт справедливость слабой проблемы Гольдбаха. Они доказали её справедливость для чисел превышающих 10^{20}, в то время как справедливость утверждения для меньших чисел легко устанавливается на компьютере. Таким образом, тернарная проблема Гольдбаха решена полностью.

Бинарная проблема Гольдбаха остается, по-прежнему, далекой от решения. Виноградов в 1937 г. и Теодор Эстерманн в 1938 г. показали, что почти все чётные числа представимы в виде суммы двух простых чисел (доля не представимых, если они есть, стремится к нулю). Этот результат немного усилен в 1975 г. Х. Монтгомери (H. Montgomery) и Р.Ч. Воном (R.C. Vaughan). Они показали, что существуют положительные константы c и C, такие что количество чётных чисел, не больших N, не представимых в виде суммы двух простых чисел, не превышает $C \cdot N^{1-c}$.

В 1939 г. Шнирельман доказал, что любое чётное число представимо в виде суммы не более 300 000 простых чисел. Этот результат многократно улучшался. В 1995 г. Ремер (Ramare) доказал, что любое чётное число — сумма не более 6 простых чисел.

В 1966 г. Чэнь Цзинжунь (Chen Jingrun) доказал, что любое достаточно большое чётное число представимо или в виде суммы двух простых чисел, или же в виде суммы простого числа и полупростого (произведения двух простых чисел). Например, $100 = 23 + 7 \cdot 11$.

На июль 2008 г. бинарная гипотеза Гольдбаха была проверена для всех чётных чисел, не превышающих $1,2 \cdot 10^{18}$.

Бинарная гипотеза Гольдбаха может быть переформулирована как утверждение о неразрешимости диофантова уравнения 4-й степени некоторого специального вида.

Гипотеза о бесконечности пар простых чисел–близнецов

Близнецами (twins) называются пары, состоящие из простых чисел, разность которых равна 2. Неизвестно, является ли число таких пар

конечным или бесконечным. Также неизвестна формула числа таких пар на начальном отрезке натурального ряда длины X.

Примерами пар близнецов являются пары $(3, 5)$, $(101, 103)$ и пара $(65\ 516\ 468\ 355 \cdot 2^{333\,333} \pm 1)$. Последняя пара является самой большой из известных на сегодняшний день пар близнецов. Отметим, что «троек» простых чисел с расстоянием 2 известна только одна – $(3, 5\,7)$.

Значительный шаг к решению проблемы близнецов сделали в 2005 г. Ден Голдстоун, Янос Пинц и Ким Ялдрим (Dan Goldston, Janos Pintz and Cem Yildirim), которые доказали формулу

$$\liminf \frac{p_{n+1} - p_n}{\log p_n} = 0, \tag{1.17}$$

где p_n–обозначает n–е простое число. Этот результат означает, что для любого сколь угодно малого ε и сколь угодно большого x найдется номер $n > x$ такой, что $p_{n+1} - p_n < \varepsilon \cdot \log p_n$, откуда следует, что функция $f(x) = min\{p_{n+1} - p_n \,|\, p_n > x\}$ возрастает медленнее, чем функция $\log x$.

Это доказательство опубликовано в электронном архиве Front For ArXiv и в статье «Are there infinitely many twin primes?» на сайтах http://front.math.ucdavis.edu/0710.2728 и http://www.math.sjsu.edu/ goldston/twinprimes.pdf

Впрочем пока этот интересный результат не будет опубликован в солидном математическом издании, его нельзя считать полностью достоверным.

1.18. Великая теорема Ферма

Великой или последней теоремой Ферма называется утверждение о том, что уравнение

$$x^n + y^n = z^n \tag{1.18}$$

не имеет решения в целых числа при $n \geq 3$.

В общем виде теорема была сформулирована Пьером Ферма в 1637 г. на полях «Арифметики» Диофанта:

Наоборот, невозможно разложить куб на два куба, биквадрат на два биквадрата и вообще никакую степень, большую квадрата, на две степени с тем же показателем. Я нашел этому поистине чудесное доказательство, но поля книги слишком узки для него.

Несколько позже сам Ферма опубликовал доказательство частного случая для $n = 4$, что добавляет сомнений в том, что у него было доказательство общего случая, иначе он непременно упомянул бы о нём в этой статье.

Эйлер в 1770 г. доказал теорему для случая $n = 3$, Дирихле и Лежандр в 1825 г. – для $n = 5$, Ламе – для $n = 7$. Куммер показал, что теорема верна для всех простых n, меньших 100, за возможным исключением иррегулярных простых чисел.

Простое число называется *иррегулярным*, если простое число классов идеалов кругового поля $R(\mathrm{e}^{2\pi/p})$ делится на p. Все остальные простые нечетные числа называются регулярными. В первой сотне иррегулярными числами являются 37, 59 и 67.

Над полным доказательством Великой теоремы работало немало выдающихся математиков и множество любителей; считается, что теорема стоит на первом месте по количеству некорректных «доказательств». В конце XIX века французская Академия Наук отказалась принимать новые доказательства теоремы Ферма. Среди приводимых доказательств было много курьезных. Например, один из корреспондентов написал: «Я нашел доказательство Великой теоремы Ферма. Главная идея – перенести слагаемое z^n в левую часть».

В 1972 г. советский математический журнал «Квант», публикуя статью о теореме Ферма, сопроводил ее следующей припиской: «Редакция «Кванта» со своей стороны считает необходимым известить читателей, что письма с проектами доказательств теоремы Ферма рассматриваться (и возвращаться) не будут».

Немецкий математик Эдмунд Ландау, которому очень докучали «ферматисты», чтобы не отвлекаться от основной работы, заказал несколько сотен бланков со следующим текстом: «Уважаемый . . . ! Благодарю Вас за

присланную Вами рукопись с доказательством Великой теоремы Ферма. Первая ошибка находится на стр. ... в строке ... ». Находить ошибку и заполнять пробелы в бланке он поручал своим аспирантам.

В 1908 г. немецкий любитель математики Вольфскель завещал 100 000 немецких марок тому, кто докажет теорему Ферма. Однако после Первой мировой войны премия обесценилась.

Тем не менее многочисленные усилия специалистов и дилетантов привели к получению многих важных результатов в этом направлении.

В 1980-х г. появился новый подход к решению проблемы. Из гипотезы Морделла, доказанной Фальтингсом в 1983 г., следует, что уравнение $a^n+b^n = c^n$ при $n > 3$ может иметь лишь конечное число взаимно простых решений. Последний, но самый важный шаг в доказательстве теоремы Ферма сделал английский математик сэр Эндрю Уайлс (Andrew John Wiles).

В январе 1993 г. Уайлсу показалось, что он доказал теорему Ферма. Он поделился полученным доказательством с коллегой по Принстонскому университету Ником Катцем. Спустя три месяца Уайлс изложил основные идеи доказательства в трех лекциях, которые состоялись в Кембриджском университете 21, 22 и 23 июня 1993 г. Математический мир был восхищен успехом Уайлса, и все с нетерпением ждали публикации текста доказательства. Комиссия в Геттингене была оповещена о возможном лауреате ее премии. Однако публикация затягивалась. Как оказалось Уайлс направил 200-страничное доказательство в журнал «Inventiones Mathematicae», где работу отдали на проверку сразу шести рецензентам. Одним из рецензентов был Ник Катц – он и обнаружил пробел в доказательстве. Это скрывалось до тех пор, пока 4 декабря 1993 г. Уайлс не признал о возникших пробелах в доказательстве. Это еще не было очередным поражением в решении проблемы Ферма, но в тот момент никто не был, включая Уайлса, уверен в обратном. В январе 1994 г. Уайлс пригласил для совместной работы своего ученика Ричарда Тэйлора. В результате совместной работы, как пишет сам Уайлс, 19 сентября 1994 г. он понял, что наконец теорема Ферма окончательно доказана. В мае 1995 г. в журнале «Annals of Mathematics» были опубликованы две статьи общим объемом

в 130 страниц. Первой шла статья Уайлса «Modular elliptic curves and Fermat's Last Theorem», поступившая в редакцию 14 октября 1994 г., второй – совместная статья Уайлса и Тэйлора «Ring-theoretic properties of certain Hecke algebras», поступившая в редакцию 7 октября 1994 г. В совокупности они давали доказательство гипотезы Таниямы–Симуры–Вейля, из которого следует доказательство Великой теоремы Ферма. Так счастливо завершилась 357-летняя история Великой теоремы Ферма. 27 октября 1995 г. Уайлс был награжден призом Ферма в Тулузе, посетил городок Бомон-де-Ломань, где родился Ферма, и его могилу, на надгробии которой высечена в виде формулы великая теорема Ферма. Так как условия конкурса на премию Вольфскеля были выполнены Уайлсом полностью, то он через два года после публикации – 27 июня 1997 г. получил награду. К этому времени Королевское научное общество было переименовано в Геттингенскую академию наук, а премия составила 75000 немецких марок.

1.19. Числа Ферма, Мерсенна и Кармайкла

В истории развития методов факторизации важную роль сыграли числа специальной формы, на которых испытывались те или алгоритмы проверки простоты и факторизации. Среди этих чисел в первую очередь надо выделить числа вида $a^n \pm 1$. Прежде, чем рассматривать эти числа, напомним формулу геометрической прогрессии:

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \ ... \ + x + 1)$$

Если n–нечетно, то, заменяя x на $-x$ и умножая все равенство на -1, получим следующую формулу:

$$x^n + 1 = (x+1)(x^{n-1} - x^{n-2} + \ ... \ - x + 1) \tag{1.19}$$

Справедлива следующая теорема:

Теорема 1.9. *Если a, n–целые числа, причем a–четно, $n \geq 2$, тогда*

a^n+1 –простое число $\rightarrow n = 2^k$ для целого числа k.

Доказательство. Пусть $n = t \cdot u$, где u–нечетно. Используя формулу (1.2) при $x = a^t$, получим

$$a^n + 1 = x^u + 1 = (x+1)(x^{t-1} - x^{t-2} + \ ... \ - x + 1)$$

где первый множитель $x + 1 = a^t + 1$ является нетривиальным делителем $a^n + 1$. Теорема доказана.

Числа Ферма

Числа вида $F_n = 2^{2^n} + 1$ впервые начал изучать Пьер Ферма, поэтому эти числа называются *числами Ферма.* Он высказал гипотезу, что все эти числа являются простыми, но не смог ни доказать, ни опровергнуть это утверждение. Первые 5 чисел Ферма действительно являются простыми: $F_0 = 3,\ F_1 = 5,\ F_2 = 17,\ F_3 = 257,\ F_4 = 65537.$

Гипотеза Ферма о простоте чисел F_n была отвергнута в 1732 г. другим выдающимся математиком Леонардом Эйлером (1707–1783) , который нашел разложение шестого числа Ферма:

$$2^{2^5} + 1 = 4\,294\,967\,297 = 641 \cdot 6\,700\,417$$

Леонард Эйлер, швейцарец по происхождению, был блестящим ученым, написавшим за свою жизнь более 800 сочинений В возрасте 20 лет он переехал из Швейцарии в Россию, куда был приглашен Петром I в период организации Российской академии наук, где и прожил почти всю свою жизнь (за исключением периода с 1741 г. по 1766 г., когда он жил и работал в Пруссии, в Берлине). С 1766 г. до своей смерти в 1783 г. он жил в Санкт-Петербурге, имел 13 детей (выжило 5), младший сын Кристоф был военным артиллеристом, подполковником прусской армии, после возвращения Эйлера в Россию, некоторое время служил в прусской армии, но потом по просьбе Екатерины II был принят в русскую армию, в которой дослужился до генерал-лейтенанта, был командиром Сестрорецкого оружейного завода. Много потомков Эйлера живет и работает в России в наше время.

Теорема 1.10. *(о делителях чисел Ферма (Эйлер-Лука)). Пусть* $F_n = 2^{2^n} + 1$, $n \geq 1$, *и* p*–простой делитель* F_n, *тогда* $p \equiv 1 \pmod{2^{n+2}}$.

Доказательство. Пусть r–простой делитель F_n, и пусть h–наименьшее натуральное число такое, что $2^h \equiv 1 (\mathrm{mod}\ r)$. Т.к. $2^{2^n} \equiv 1 (\mathrm{mod}\, r)$, то $h = 2^{n+1}$. Далее, поскольку по малой теореме Ферма $2^{r-1} \equiv 1 (\mathrm{mod}\, r)$, то h является делителем $r - 1$. Т.к. $n \geq 2$, то $r \equiv 1 (\mathrm{mod}\ 8)$. Из последнего условия следует (см.стр.25), что 2 является квадратичным вычетом по модулю r, откуда, $2^{(r-1)/2} \equiv -1 (\mathrm{mod}\, r)$, что влечет утверждение теоремы.

Пример. Рассмотрим $F_5 = 2^{32} + 1$. Любой его делитель по теореме должен иметь вид $1 + 128k$. И действительно, $F_5 = 4\,294\,967\,297 = 641 \cdot 6\,700\,417 = (1 + 5 \cdot 128) \cdot (1 + 52347 \cdot 128)$.

Сведения о последних разложениях чисел Ферма можно найти на сайте ([28]): **http://www.prothsearch.net/fermat.html**

Ниже мы выпишем таблицу разложений первых чисел Ферма:

$F_0 = 3$

$F_1 = 5$

$F_2 = 17$

$F_3 = 257$

$F_4 = 65537$

$F_5 = 641 \cdot 6700417$

$F_6 = 274177 \cdot 67280421310721$

$F_7 = 59649589127497217 \cdot 5704689200685129054721$

$F_8 = 1238926361552897 \cdot P62$

$F_9 = 2424833 \cdot 7455602825647884208337395736200454918783366342657 \cdot P99$

$F_{10} = 45592577 \cdot 6487031809 \cdot 4659775785220018543264560743076778192897 \cdot$
$\cdot P252$

$F_{11} = 319489 \cdot 974849 \cdot 167988556341760475137 \cdot 3560841906445833920513 \cdot$
$\cdot P564$

Здесь запись Pk с индексом k, например $P564$, обозначает известное простое число длины k в десятичной записи.

Первым в последовательности чисел Ферма, которые еще не полностью факторизованы на сегодняшний день, является F_{12}. Совсем недавно, в марте 2010 г. Майклом Вангом (Michael Vang) был найден шестой фактор F_{12}, равный

$$P_6 = 1735323021042959457913309969912316298948244452089 9 \cdot 2^{15}+1.$$

Интересно, что предыдущие делители F_{12} были найдены в 1877 г., 1903 г. (два делителя), 1974 г. и 1986 г. соответственно. Однако, остался еще один составной кофактор этого числа, состоящий из 1033 десятичных цифр, который еще не разложен до конца.

Есть частичные сведения также о числах Ферма с большими номерами, некоторые из которых просто огромные! Например, Янг (J.Young) показал, что число F_{213319} имеет кофактор $3 \cdot 2^{213321} + 1$.

Числа Мерсенна

Числами Мерсенна называются числа вида $M_p = 2^p - 1$ с простыми индексами p. Начальная последовательность чисел Мерсенна имеет вид:

1, 3, 7, 31, 127, 2047, 8191, 131071, 524287, 8388607, 536870911, 2147483647.

Теорема 1.11. *(о делителях чисел Мерсенна). Пусть $p \geq 3$ –простое число, q –делитель $M_p = 2^p - 1$, тогда $q \equiv 1 \pmod{2p}$ и $q \equiv \pm 1 \pmod 8$.*

Доказательство. По предположению, $2^p \equiv 1 \pmod q$ и по малой теореме Ферма, $2^{q-1} \equiv 1 \pmod q$. Поэтому, порядок 2 (т.е. наименьший показатель t такой, что $2^t \equiv 1 \pmod q$ обозначается $ord_q(2)$) является одновременно делителем p и $q-1$. Он не может быть равен 1, поэтому он равен p, откуда $p|(q-1)$, т.е. $q = 1 + mp$. Т.к. оба числа p и q–нечетные, то $m = 2k$–четно, откуда $q = 1 + 2kp$. Значит, p–делитель $(q-1)/2$. Т.к $2^p \equiv 1 \pmod q$, то $2^{(q-1)/2} \equiv 1 \pmod q$, откуда символ Лежандра $(2/q) = 1$ и $q \equiv \pm 1 \pmod 8$.

Числа Мерсенна получили известность в связи с эффективным *критерием простоты Люка — Лемера*, благодаря которому простые числа

Мерсенна давно удерживают лидерство как самые большие известные простые числа. Этот тест был придуман Люка (Lucas) в 1878 г. и в 1930 г. усовершенствован Лемером (Lehmer).

Тест Люка — Лемера базируется на том, что простота числа Мерсенна M_p влечёт простоту числа p, и следующем утверждении:

Тест Люка. Для простого числа $p \geq 3$ число M_p является простым тогда и только тогда, когда оно делит число $L_p - 1$, где числа L_k определяются рекуррентным соотношением:

$$L_1 = 4, \;\; L_{k+1} = L_k^2 - 2$$

Для установления простоты M_p достаточно вычислять последовательность чисел по модулю числа M_p, длина которого ограничена p битами. Последнее число в этой последовательности $L_p - 1 (\mathrm{mod} M_p)$ называется вычетом Люка — Лемера. Таким образом, число Мерсенна M_p является простым тогда и только тогда, когда число p простое и вычет Люка — Лемера равен нулю.

Еще в 1876 г. Люка с помощью этого критерия установил, что число

$$2^{127} - 1 = 170141183460469231731687303715884105727$$

является простым. Это число оставалось самым большим известным простым числом до 1951 г., т.е. на протяжении 75 лет.

К сентябрю 2004 г. было найдено 44 простых числа Мерсенна, 4 сентября 2006 г. было аннонсировано 45-е число $2^{32\,582\,657} - 1$. На сегодняшний день известно 47 простых чисел Мерсенна. Самым большим известным простым числом является число Мерсенна $M_p = 2^{43112609} - 1$, найденное в августе 2008 г. в рамках проекта распределённых вычислений GIMPS. Длина его составляет 12978189 десятичных цифр, что позволило GIMPS в 2009 г. получить премию в 100000 долларов США, назначенную сообществом Electronic Frontier Foundation за нахождение простого числа, длина которого превышает 10 миллионов десятичных цифр.

В теории чисел изучаются также *обобщенные* числа Мерсенна. Такими являются числа вида

$$M(k,n) = k \cdot 2^n \pm 1,$$

где n–простое число, а k–небольшое простое число.

Числа Кармайкла

Малая теорема Ферма утверждает, что $a^{p-1} \equiv 1 \,(\mathrm{mod}\ p\,)$ для простых p и произвольных натуральных a, взаимно–простых с p. Обратное утверждение неверно, есть небольшое количество составных чисел n таких, что для всех не сравнимых с n чисел a выполняется $a^{n-1} \equiv 1(\ \mathrm{mod}\ n)$. Такие числа называются числами Кармайкла (Carmichael' Numbers).

Эквивалентное определение чисел Кармайкла дает критерий Корсельта.

Теорема 1.12. *(Корсельт, 1899) Составное число n является числом Кармайкла тогда и только тогда, когда n свободно от квадратов, и для каждого простого делителя p числа n число $p-1$ делит число $n-1$.*

В частности, из теоремы Корсельта следует, что все числа Кармайкла нечётны, так как любое чётное составное число, свободное от квадратов, имеет, по крайней мере, один нечётный простой делитель, и поэтому из делимости $(p-1) \,|\, (n-1)$ следует, что чётное делит нечётное, что невозможно.

Корсельт был первым, кто заметил это свойство, но он так и не смог найти какие-либо примеры. В 1910 г. Кармайкл нашел первое и наименьшее такое число 561. Последовательность чисел Кармайкла начинается так:

561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341, 41041, 46657, 52633, 62745, 63973, 75361, 101101, 115921, 126217, 162401, 172081, 188461, 252601, 278545, 294409, 314821, 334153, 340561, 399001, 410041, 449065, 488881, 512461.

Числа Каннингама

Числа Каннингама (Cunningham' Numbers) – числа вида $b^n \pm a^n$. Проект Каннингама – это проект создания таблиц разложения чисел

Каннингама на простые множители (при $a = 1$). Алан Каннингам (Allan Cunningham), английский математик (1848–1928), составил первые таблицы разложения в 1925 г. (совместно с Г. Вуделом H. Woodall). Проект был продолжен Брильхартом, Лемером, Селфриджем, Такерманом и Вагстаффом (John Brillhart, D.H. Lehmer, J.L. Selfridge, Bryant Tuckerman, and S.S. Wagstaff, Jr.) в 1988 г. Последний (третий) вариант таблиц Каннингама опубликован AMS в 2002 г. в книге «Factorizations of $b^n \pm 1, b = 2, 3, 5, 6, 7, 10, 11, 12$ Up to High Powers», Third Edition. Доступна по адресу:

http://www.ams.org/publications/online-books/conm22-index

2. Простые алгоритмы факторизации

Факторизацией целого числа называется его разложение в произведение простых сомножителей. Такое разложение, согласно основной теореме арифметики, всегда существует и является единственным (с точностью до порядка следования множителей). Все методы факторизации в зависимости от их производительности можно разбить на две группы: экспоненциальные методы и субэкспоненциальные методы. Все эти методы достаточно трудоемки, поэтому требуют значительных вычислительных ресурсов для чисел большой длины. Однако теоретическое обоснование необходимой сложности таких вычислений или, другими словами, существование высоких нижних оценок не доказано, поэтому вопрос о существовании алгоритма факторизации с полиномиальной сложностью на классическом компьютере для выполнения факторизации является одной из важных открытых проблем современной теории чисел. В то же время факторизация с полиномиальной сложностью возможна на квантовом компьютере с помощью алгоритма Шора.

В этой главе мы дадим описание наиболее известных алгоритмов факторизации, имеющих экспоненциальную оценку сходимости.

2.1. Метод Ферма

Пусть $n = p \cdot q$ – известное целое число, являющееся произведением двух неизвестных простых чисел p и q, которые требуется найти. Большинство современных методов факторизации основано на идее, предложенной еще Пьером Ферма, заключающейся в поиске пар натуральных чисел A и B таких, что выполняется соотношение:

$$n = A^2 - B^2. \tag{2.20}$$

Алгоритм Ферма может быть описан следующим образом:

1. Вычислим целую часть от квадратного корня из n:

$$m = \lceil \sqrt{n} \rceil.$$

2. Для $x = 1, 2, \ldots$ будем вычислять значения

$$q(x) = (m + x)^2 - n, \tag{2.21}$$

до тех пор, пока очередное значение $q(x)$ не окажется равным полному квадрату.

3. Пусть $q(x)$ является полным квадратом, например, числа B: $q(x) = B^2$. Определим $A = m + x$, откуда из равенства $A^2 - n = B^2$ найдем $n = A^2 - B^2 = (A + B) \cdot (A - B)$, и искомые делители p и q вычисляются, как $p = A + B$, $q = A - B$.

Пример. Пусть факторизуемое число $n = 19\,691$. Вычислим $m = \lfloor \sqrt{n} \rfloor = 140$. Представим процедуру вычисления делителей n в виде таблицы:

x	y	$\sqrt{y}$
1	190	13,78
2	473	21,75
3	758	27,53
4	1045	32,33
5	1334	36,52
6	1625	40,31
7	1918	43,79
8	2213	47,04
9	2510	50,10
10	2809	53

Из последнего столбца получим: $(140 + 10)^2 - n = 53^2$, откуда $n = 150^2 - 53^2 = 203 \cdot 97$. Итак, $19\,691 = 203 \cdot 97$, и вычисление потребовало 10 итераций, в каждой из которых было выполнено 1 возведение в степень, 1 вычитание и одно вычисление квадратного корня, т.е. константное число операций.

Оценка производительности метода Ферма

В наихудшем случае, когда q близко к 1, а p близко к n, алгоритм будет работать даже хуже, чем метод пробных делений. Действительно, $A = (p + q)/2$, откуда число итераций в методе Ферма равно

$$Iter(n) = \frac{p+q}{2} - \lfloor n^{1/2} \rfloor \approx \frac{n}{2} - \lfloor n^{1/2} \rfloor,$$

т.е. имеет порядок $0(n)$. Для того, чтобы метод Ферма работал не хуже, чем метод пробных делений необходимо, чтобы $Iter(n)$ было меньше $n^{1/2}$, откуда больший делитель $p < 4n^{1/2}$.

Таким образом, как и метод пробного деления, алгоритм Ферма имеет экспоненциальную оценку и не эффективен для разложения длинных чисел.

Можно улучшить метод Ферма, выполнив сначала пробное деление числа n на числа от 2 до некоторой константы B, исключив тем самым малые делители n до B включительно, и только потом выполнить поиск методом Ферма.

2.2. $(p-1)$–метод Полларда

Этот метод был разработан английским математиком Джоном Поллардом в 1974 г. и опубликован в [43].

Пусть n—факторизуемое число, а $1 < p < n$– его простой делитель. Согласно малой теореме Ферма, для любого a, $1 \leq a < p$, выполняется условие $a^{p-1} \equiv 1(\bmod p)$.

Это же сравнение выполнится, если вместо степени $p-1$ взять произвольное натуральное число M кратное $p-1$, т.к. если $M = (p-1)\cdot k$, то $a^M = (a^{p-1})^k \equiv 1^k \equiv 1(\bmod p)$. Последнее условие эквивалентно $a^M - 1 = pr$ для некоторого целого r. Отсюда, если p является делителем числа n, тогда p является делителем наибольшего общего делителя Н.О.Д.$(n, a^M - 1)$ и совпадет с Н.О.Д.$(n, a^M - 1)$, если $a^M - 1 < n$. Пусть

$$p - 1 = p_i^{r_1} \cdot p_2^{r_2} \cdot \ ... \ \cdot p_t^{r^t}. \tag{2.22}$$

Идея $(p-1)$–метод Полларда состоит в чтобы выбрать M в виде произведения как можно большего числа простых сомножителей или их

степеней так, чтобы M делилось на каждый сомножитель $p_i^{r_i}$, входящий в разложение (2.22). Тогда, Н.О.Д.$(n, a^M - 1)$ даст искомый делитель. Алгоритм состоит из двух стадий:

Первая стадия (p-1)–алгоритма Полларда

1. Сначала выберем границу B_1.
2. Определим множество P, состоящее из простых чисел и их степеней, меньших границы B_1:

$$P = \{p_1^{r_1},\ p_2^{r_2},\ \ldots p_k^{r_k}\},\ \ p_i^{r_i} < B_1.$$

3. Вычислим произведение

$$M = M(B_1) = \prod_{p_i^{r_i} \in P} p_i^{r_i}$$

4. Выберем произвольное число a, например 2, и вычислим $a^M \mod n$.
5. Вычислим Н.О.Д.$(n, a^M - 1)$, который, если повезет даст искомый делитель числа n.

Пример. Факторизовать $n = 10\,001$. Выберем $B = 10$, тогда, $M(B_1) = 2^3 \cdot 3^2 \cdot 5 \cdot 7 = 2520$. Вычислим, $2^{2520} \mod 10\,001 = 3579$. Найдем, Н.О.Д.$(n, a^M - 1) =$ Н.О.Д.$(10\,001, 3578) = 73$.

Заметим, что при большом значении B_1 число $M(B_1)$ может оказаться чрезвычайно большим (оно сравнимо с $B_1!$). В таких случаях лучше разбить полное произведение $M(B_1)$ на l блоков, состоящих из примерно одинакового числа сомножителей, и вычислив числа M_i как произведение элементов блока i, представить $M(B)$ в виде $M_1 \cdot M_2 \cdot \ldots \cdot M_l$. Затем, можно вычислить $a^{M(B)}$ как предел последовательности $\{a_i\}$, где $a_1 = a^{M_1} \pmod n$, а последующие a_i вычисляются по формуле:

$$a_{i+1} = a_i^{M_{i+1}} \mod n,\ \ i < l.$$

В этом случае, все вычисления будут выполняться с числами, сравнимыми по модулю с числом n.

Вторая стадия (p-1)–алгоритма Полларда

Если в результате первого этапа алгоритм не выдает требуемого делителя, то можно либо увеличить границу B_1, либо начать вторую стадию работы алгоритма.

Вторая стадия алгоритма предполагает, что существует только один простой множитель q числа $p-1$, значение которого больше границы B_1. Выберем новую границу $B_2 \gg B_1$, например, $B_2 = B^2$. Обозначим через b число $a^{M(B)} \mod n$, вычисленное на первой стадии работы алгоритма.

Выпишем последовательность $q_0 < q_1 < ... < q_s$ всех простых чисел на интервале $[B; B_2]$. Для построения этого множества можно воспользоваться решетом Эратосфена, либо решетом Аткина (см.гл.1).

Поскольку наличие в последовательности $\{q_i\}$ нескольких составных чисел не испортит работы алгоритма, можно выполнить только частичное просеивание, отсеяв числа, кратные небольшим простым числам. Это ускорит общую работу алгоритма.

Если искомый множитель $p-1$ равен q_i, то для нахождения делителя n, необходимо вычислить $c_i = b^{q_i} \mod n$, и найти Н.О.Д.$(n, c_i - 1)$. Поскольку, значение q неизвестно, мы должны выполнить последние две операции с каждым числом q_i из интервала $[B_1; B_2]$. Полларда предложил следующий вариант организации этой процедуры. Обозначим через δ_i разность между соседними простыми числами $\delta_i = q_{i+1} - q_i$. Возможные значения, принимаемые d_i, лежат в небольшом множестве $D = \{2, 4, ..., 2t\}$. Можно заранее вычислить все значения $b^\delta \mod n$ для $\delta \in D$ и сохранить полученные числа в массиве. Вторая стадия алгоритма выполняется следующим образом:

1. Вычислим сначала $c_0 = b^{q_0} \mod n$, и найдем $d =$Н.О.Д.$(n, c_0 - 1)$.

2. Если $d = 1$, то вычислим следующее $c_1 = b^{q_1} \mod n$ и $d =$Н.О.Д.$(n, c_1 - 1)$ и т.д.

3. Каждое последующее значение c_{i+1} вычисляется по формуле

$$b^{q_{i+1}} \mod n = b^{q_i+\delta_i} \mod n = b^{q_i} \cdot b^{\delta_i} \mod n = c_i \cdot b^{\delta_i} \mod n. \qquad (2.23)$$

Поскольку все значения $b^{\delta_i} \bmod n$ заранее вычислены, то для вычисления очередного значения c_{i+1} достаточно одной операции умножения и вычисления остатка по модулю n. Поэтому вторая стадия алгоритма Полларда выполняется очень быстро.

Оценка эффективности $(p-1)$–метода Полларда

Сделаем расчет времени работы алгоритма при условии, что параметры B_1 и B_2 выбраны. Время выполнения первой стадии зависит от числа простых чисел и их степеней на интервале $[2; B]$. Число простых чисел оценивается величиной $\pi(B_1)$, приближенно равной по теореме Чебышева числу $B_1/\ln B_1$. Для каждой степени p^r, меньшей B, производится r возведений в степень по модулю по алгоритму, описанному на стр. 13, и требует $\log_2 p \le \log_2 B_1$ операций возведения в квадрат и умножений по модулю числа n. Поэтому общее число операций можно оценить величиной $O(B_1 \log B_1 \log^2 N)$. Метод очень быстро находит простые факторы малой и средней величины (до 20-25 десятичных цифр). Текущим рекордом для $(p-1)$–метода является простой делитель числа $960^{119}-1$, состоящий из 66 десятичных цифр, установленный Т. Нохара (T. Nohara) в 2006 г.

Использование второй стадии позволяет увеличить эффективность метода. По оценке Монтгомери, вторая стадия алгоритма требует

$$O(\log^2 B_2) \,+\, O(\log q_{\pi(B_1)}) \,+\, 2(\pi(B_2) - \pi(B_1))$$

умножений по модулю n и вычислений Н.О.Д. с n. Отбрасывая слагаемые меньшего порядка, получим оценку $O(\pi(B_2))$.

Условие сходимости $(p-1)$–метода Полларда

Пусть p–наименьший из делителей n и q^t–наибольшая степень простого числа, входящего в разложение $p-1$. Иначе говоря, q^t максимально среди всех степеней $q_i^{t_i} \mid p-1$. Отметим, что оценка сложности $(p-1)$–алгоритма определяется не размером факторизуемого числа n, а размером сомножителя q^t чисел $p-1$ для $p \mid n$.

Если $q^t \leq B_1$, тогда вычисление закончится на первом этапе алгоритма. Иначе, для успеха алгоритма необходимо, чтобы выполнилось $q^t \leq B_2$, а все степени простых делителей $(p-1)$ вида q^r кроме последнего были меньше B_1. Кроме того необходимо, чтобы среди делителей $p-1$ не нашлось множителей вида r^k при $k \geq 2$, находящегося между B_1 и B_2. Для таких r^k имеем два неравенства:

$$r^{k-1} \leq B, \quad B_1 < r^k < B_2,$$

откуда

$$B_1^{1/k} < p < min\{B^{1/(k-1)},\ B_2^{1/k}\}. \tag{2.24}$$

При $B_2 = cB_1$ и $k = 2$ уравнения (2.24) приобретут вид:

$$\sqrt{B_1} < r < min\{B_1,\ \sqrt{cB_1}\}.$$

Поскольку значение границы B_2 обычно выбирается так, чтобы выполнялось $B_2 \leq B_1^2$, последнее уравнение эквивалентно

$$\sqrt{B_1} < r\ <\ c_1\sqrt{B_1},\ \text{где } c_1 = \sqrt{c}. \tag{2.25}$$

Общая доля чисел, удовлетворяющих (2.25), невелика и значительно меньше числа простых чисел из интервала $(B_1,\ B_2)$, поэтому ими можно пренебречь. Однако, если не пренебрегать этими числами и добавить в алгоритм дополнительный цикл по элементам r, удовлетворяющим условию (2.25), общее время алгоритма увеличится незначительно. Поскольку размер наибольшей степени q^t сильно зависит от степени гладкости числа $p-1$, поэтому эффективность $(p-1)$–метода Полларда сильно зависит от исходного числа n, изменяясь в широких пределах при различных n одинаковой длины. Поэтому одной из рекомендаций метода RSA является выбор n так, чтобы $p-1$ имел хотя-бы один большой делитель, превышающий размер границы B_2, до которой возможно выполнение реальных вычислений по $(p-1)$–методу Полларда.

Скажем несколько слов о выборе исходного значения параметра a. Если $p-1$ имеет большое число различных делителей, то найдется много различных чисел $a < n$, для которых $a^k \equiv 1 (\mathrm{mod} p)$

выполнится для $k < p-1$. Найдутся даже такие $a < n$, что уже $a^2 \equiv 1 (\mathrm{mod} p)$. Для таких a скорость схождения метода будет значительно выше. Поэтому, можно ускорить сходимости $(p-1)$–метода Полларда, запуская алгоритм с несколькими значениями a. В статье «Pollard ρ on the Play Station 3», размещенной на сайте http://www.hyperelliptic.org/tanja/SHARCS/slides09/03-bos.pdf, приводятся примеры программирования алгоритмов факторизации, включая ρ и $(p-1)$ методы Полларда с возможностью распараллеливания на игровой консоли Sony Play Station 3, имеющей 8 сопроцессоров.

Пример

Пусть $p = 29$–делитель n, тогда, $p - 1 = 28 = 2^2 \cdot 7$. Для каждого $a \leq 28$ найдем наименьшее k такое, что $a^k \equiv 1 (\mathrm{mod}\ p)$. Приведем фрагмент полученной таблицы:

a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
k	28	28	14	14	14	7	28	14	28	28	4	14	28	28	7	4

Среди 28 значений $a < 29$ окажется 12 значение с показателем $k = 28$, по 6 значений с показателем $k = 14$ и $k = 7$, 2 значения с $k = 4$, по одному с $k = 2$ и $k = 1$. Математическое ожидание наименьшего показателя k равно:

$$M[k] = (12 \cdot 28 + 6 \cdot 14 + 6 \cdot 7 + 2 \cdot 4 + 2 \cdot 1)/28 \approx 16,85$$

Таким образом, выбирая удачное $a < n$, можно получить значительный выигрыш по сравнению с произвольным.

Дальнейшие улучшения алгоритма

Для ускорения всех вычислений Полларда предложил использовать быстрое преобразование Фурье (a Fast Fourier Transform) для всех основных операций.

Дальнейшие улучшения алгоритма были предложены П.Монтгомери [36]. Он заметил, что на второй стадии алгоритма большую часть времени отнимает вычисление для каждого простого числа q_i из интервала $[B_1; B_2]$

Н.О.Д. $(n, c_i - 1)$, где $c_i = b^{q_i}$. Монтгомери предложил объединять несколько соседних элементов c_i в блоки G_j и вычислять сначала произведение $h = \prod(c_i - 1) \mod n$ для всех $c_i \in G_j$, а потом Н.О.Д. (n, h). Если Н.О.Д. $(n, c_i - 1) > 1$, то таковым будет и Н.О.Д. (n, h). Эти и другие улучшения, найденные Монтгомери, позволяют в несколько раз сократить общее время работы алгоритма.

На сайте www.loria.fr/∼zimmerma/records/Pminus1.html можно найти таблицу рекордов разложения натуральных чисел, установленных с помощью $(p-1)$–метода Полларда.

Упражнение.

Сформируйте множество E_n всех составных чисел n одинаковой длины, являющихся произведением двух простых чисел. Найдите для каждого числа n математическое ожидание $M[k]$ наименьшего показателя k для множества элементов $a < p$, где p– наименьший делитель n. Распределите все числа из множества E_n в классы в зависимости от значения $M[k]$. Сделайте вывод о доле составных чисел, которые могут быть быстро разложены в произведение простых сомножителей с помощью $(p-1)$–метода Полларда.

В главе 2 мы увидим, что идея $(p-1)$-метода Полларда была использована Х.Ленстрой для построения нового более быстрого метода факторизации с использованием эллиптических кривых.

2.3. $(p+1)$–метод Вильямса

Определение. Последовательностью Люка (Lucas) назовем реккурентную последовательность u_n, определяемую соотношениями:

$$u_0 = 0, \quad u_1 = u, \quad u_{n+1} = P \cdot u_n - Q \cdot u_{n-1}, \tag{2.26}$$

где P, Q – фиксированные целые числа.

$(p+1)$–метод Вильямса (Williams) похож на $(p-1)$–метод Полларда и основан на предположении гладкости числа $p+1$. Пусть p–простой делитель

факторизуемого числа n, и выполнено разложение $p+1$

$$p+1=\prod_{i=1}^{k} q_i^{a_i}.$$

Обозначим через $B = \max\{q_i^{a_i} | 1 \leq i \leq k\}$. По-прежнему будем называть натуральное число r B–степенно-гладким, если наибольшая степень сомножителя $p_i^{a_i}$ в разложении r на простые множители, не превышает B. Таким образом, определенное выше число B является наименьшим числом, для которого $p+1$ является B–степенно-гладким. Отметим, что поскольку p не известно, то и B так же не известно.

Алгоритм Вильямса заключается в следующем:

1. Выбираем некоторое число B, являющее верхней границей для рассматриваемых простых чисел и их степеней.

2. Строим последовательность простых чисел $2 < 3 < 5 < ... < p_m$, меньших B и последовательность степеней a_i такую, что $p_i^{a_i} < B$.

3. Полагаем число $R = \prod_{i=1}^{m} q_i^{a_i}$. Если p является B–степенно-гладким, то R делится на p.

4. Выбираем случайным образом числа P и Q и строим последовательность чисел Люка, пока не вычислим u_R.

5. Далее вычислим Н.О.Д.$(n, u_R) = d$. Если $1 < d < n$, то задача решена.

Доказано, что если Q взаимно просто с p и

$$\left(\frac{P^2-4Q}{p}\right) = -1,$$

то свойства последовательности Люка обеспечивают нахождение нетривиального делителя числа n.

2.4. ρ-метод Полларда

Этот метод был разработан Джоном Поллардом в 1975 г. Пусть n – число, которое следует разложить. ρ-метод Полларда работает следующим образом:

1. Выбираем небольшое число x_0 и строим последовательность чисел $\{x_n\}$, $n = 0, , 1, 2, \ldots$, определяя каждое следующее x_{n+1} по формуле $x_{n+1} = (x_n^2 - 1) \,(\mathrm{mod}\, n)$.

2. Одновременно на каждом шаге i вычисляем Н.О.Д d числа n и всевозможных разностей $|x_i - x_j|$, где $j < i$.

3. Когда будет найдем $d =$Н.О.Д.$(n,\, |x_i - x_j|)$, отличный от 1, вычисление заканчивается. Найденное d является делителем n. Если n/d не является простым числом, то процедуру можно продолжить, взяв вместо n число n/d.

Вместо функции $F(x) = (x^2 - 1) \bmod n$ в вычислении x_{n+1} можно взять другой многочлен, например, $x^2 + 1$ или произвольный многочлен 2-й степени $F(x) = ax^2 + bx + c$.

Недостатком данного варианта метода является необходимость хранить большое число предыдущих значений x_j. Заметим, что если

$$(x_j - x_i) \equiv 0 \,(\,\mathrm{mod}\, p), \text{ то } (f(x_j) - f(x_i)) \equiv 0 \,(\,\mathrm{mod}\, p),$$

поэтому, если пара $(x_i,\, x_j)$ дает нам решение, то решение даст любая пара $(x_{i+k},\, x_{j+k})$.

Поэтому, нет необходимости проверять все пары $(x_i,\, x_j)$, а можно ограничиться парами виды $(x_i,\, x_j)$, где $j = 2^k$, и k пробегает набор последовательных значений $1, 2, 3, \ldots$, а i принимает значения из интервала $[2^k + 1; 2^{k+1}]$. Например, при $k = 3$ $j = 2^3 = 8$, а $i \in [9; 16]$.

```
int ρ-Pollard (int n)
{ int x = random (1, n-2);
int y = 1; int i = 0; int stage = 2;
while(Н.О.Д.(n, abs(x − y)) = 1)
{
   if (i = = stage ){
      y = x;
       stage = stage*2; }
   x=x ∗ x + 1(mod n);
   i=i + 1;
}
return Н.О.Д.(n, abs(x − y)); }
```

В этом варианте вычисление требует хранить в памяти всего три переменные n, x и y, что выгодно отличает этот метод от других методов факторизации.

Еще одна вариация ρ-метода Полларда была разработана Флойдом (Floyd). Согласно Флойду, значение y обновляется на каждом шаге по формуле $y = F^2(y) = F(F(y))$, поэтому на шаге i будут получены значения $x_i = F^i(x_0)$, $y_i = x_{2i} = F^{2i}(x_0)$, и Н.О.Д. на этом шаге вычисляется между n и $y - x$.

Обоснование ρ-метода Полларда

Приведем обоснование этого метода и оценим его трудоемкость. Оценка основывается на известном «парадоксе дня рождения».

Теорема 2.1. *(Парадокс дня рождения) Пусть* $\lambda > 0$. *Для случайной выборки из* $l + 1$ *элементов, каждый из которых меньше* q, *где* $l = \sqrt{2\lambda q}$, *вероятность того, что два элемента окажутся равными*

$$p > 1 - e^{-\lambda}.$$

Отметим, что вероятность $p = 0,5$ в парадоксе дня рождения достигается при $\lambda \approx 0,69$.

Пусть последовательность $\{u_n\}$ состоит из разностей $|x_i - x_j|$, проверяемых в ходе работы алгоритма. Определим новую последовательность $\{z_n\}$, где $z_n = u_n \bmod q$, q – меньший из делителей n. Все члены последовательности $\{z_n\}$ меньше $\sqrt{n}$. Если рассматривать $\{z_n\}$ как случайную последовательность чисел, меньших q, то, согласно парадоксу близнецов, вероятность того, что среди первых $l+1$ ее членов попадутся два одинаковых, превысит $1/2$ при $\lambda \approx 0,69$, тогда l должно быть не меньше $\sqrt{2\lambda q} \approx \sqrt{1.4q} \approx 1,18\sqrt{q}$.

Если $z_i = z_j$, тогда $x_i - x_j \equiv 0 \bmod q \to x_i - x_j = kq$ для некоторого $k \in \mathbf{Z}$. Если $x_i \neq x_j$, что выполняется с большой вероятностью, то искомый делитель q числа n будет найден как Н.О.Д.$(n, x_i - x_j)$. Поскольку $\sqrt{q} \leq n^{1/4}$, то с вероятностью, превышающей 0,5, делитель n может быть найден за $1,18 \cdot n^{1/4}$ итераций.

Таким образом, ρ-метод Полларда является вероятностным методом, позволяющим найти нетривиальный делитель q числа n за $O(q^{1/2}) \leq O(n^{1/4})$ итераций. Сложность вычисления нетривиального делителя в этом методе зависит только от размера этого делителя, а не от размера числа n. Поэтому, ρ–метод Полларда применим в тех случаях, когда другие методы факторизации, зависящие от размера n, становятся не эффективными.

Отметим, что в некоторых случаях, последовательность $\{y_n\}$ может зацикливаться (т.е. на некотором шаге t появляется $x_t = x_0$, после чего последовательность повторяется), тогда надо поменять исходный элемент x_0 или полином $F(x)$ на какой–нибудь другой.

Упражнения.

1. Подберите несколько составных чисел n одинаковой длины, и выполните пробное разложение этих чисел методом Полларда. Вычислите среднее время (количество итераций) до нахождения нетривиального делителя n. Как сильно отличается время вычисления для различных n?

2. Выполните упражнение 1 с алгоритмом Флойда. Сравните среднее время вычисления делителя в первом и втором случаях.

2.5. ρ-метод Полларда для вычисления дискретного логарифма

Проблема дискретного логарифма (Discrete Logarithm Problem DLP) состоит в вычислении в конечном поле F_q с образующей g для произвольного элемента t наименьшего числа k такого, что $g^k = t$. Хотя эта проблема не связана непосредственно с проблемой факторизации целых чисел, она играет важную роль в криптографии. При длине ключа L проблема DLP имеет такую же сложность решения, как и проблема факторизации числа длины L, поэтому на проблеме вычисления DLP построено много криптографических протоколов, в том числе, известные протоколы Диффи-Хелмана вычисления общего секретного ключа и Эль-Гамаля электронной цифровой подписи.

Существует большое число различных методов для решения этой задачи. В главе 5 книги Л.Вашингтона [54] дано описание основных алгоритмов для ДЛЭК. В этом разделе мы рассмотрим ρ-метод Полларда для DLP, который играет здесь ту же роль, что и ρ-метод Полларда для проблемы факторизации. Группу по умножению поля F_p, p–простое число, обозначим через $F_p^* = \{1, 2 \dots p-1\}$. Напомним, что элемент $g \in F_p^*$ называется генератором поля, если любой элемент $t \in F_p^*$ равен некоторой степени элемента g: $t = g^k$. Пусть g – (какой-нибудь) генератор этой группы, и пусть t - произвольный элемент F_p^*.

Для нахождения неизвестного показателя k такого, что $g^k = t$, будем строить последовательность пар (a_i, b_i) чисел по модулю $p-1$ и последовательность x_i чисел по модулю p такую что $x_i = t^{a_i} g^{b_i}$. Определим начальные значения $a_0 = b_0 = 0$, $x_0 = 1$. Вычисление последующих членов последовательностей будем выполнять по формулам:

$$(a_{i+1}, b_{i+1}) = \begin{cases} (a_i + 1,\ b_i) \bmod (p-1), & \text{если } 0 < x_i < p/3, \\ (2a_i,\ 2b_i) \bmod (p-1), & \text{если } p/3 < x_i < 2p/3, \\ (a_i,\ b_i + 1) \bmod (p-1) & \text{если } 2p/3 < x_i < p, \end{cases} \tag{2.27}$$

и, соответственно,

$$x_{i+1} = \begin{cases} tx_i \bmod p, & \text{если } 0 < x_i < p/3, \\ x_i^2 \bmod p, & \text{если } p/3 < x_i < 2p/3, \\ gx_i \bmod p & \text{если } 2p/3 < x_i < p, \end{cases} \tag{2.28}$$

Эта последовательность вычисляется до тех пор, пока не появятся номера i, j такие, что $x_i = x_j$. Тогда, $t^{a_i} g^{b_i} = t^{a_j} g^{b_j}$, откуда,

$$(a_j - a_i)k \equiv b_i - b_j \pmod{(p-1)} \tag{2.29}$$

Если Н.О.Д.$(a_j - a_i,\ p-1) = 1$, тогда множитель k в уравнении (2.29) может быть найден с использованием обобщенного алгоритма Евклида, решив в целых числах уравнение

$$x(a_j - a_i) + y(p-1) = b_i - b_j \tag{2.30}$$

относительно x, y и определяя $k = x \mod (p-1)$.

Если же Н.О.Д.$(a_j - a_i,\ p-1) = d > 1$, тогда, уравнение (2.30) по-прежнему, разрешимо, но дает решение нашего уравнения с точностью до слагаемого кратного $(p-1)/d$, т.е. решение имеет вид

$$x = x_0 + m(p-1)/d \tag{2.31}$$

где $m \in [0, d-1]$–целое число. Если множитель d - мал, то решение будет найдено подстановкой чисел (2.31) в уравнение $g^X \equiv t \mod t$.

Так же, как в ρ–методе факторизации, в этом алгоритме можно использовать модификацию Флойда, вычисляя на i-м шаге одновременно тройку (a_i, b_i, x_i) и тройку (a_{2i}, b_{2i}, x_{2i}), пока не дойдем до шага i, на котором $x_i = x_{2i}$. В этом варианте опять не надо хранить в памяти на шаге i все тройки (a_j, b_j, x_j) для $j \le i$, а достаточно сохранять две тройки (a_i, b_i, x_i) и (a_{2i}, b_{2i}, x_{2i}).

Пример. Рассмотрим поле F_p при $p = 43$. Элемент $g = 2$ не является генератором по критерию Поклингтона (см.стр.19), т.к. $2^{14} \bmod 43 = 1$. Возьмем в качестве генератора элемент $g = 3$ и решим уравнение

$$3^X \bmod 43 = 15. \tag{2.32}$$

Итак, $p = 43$, $g = 3$, $t = 15$. Определим $(_0, b_0, x_0) = (0, 0, 1)$, и будем строить две последовательности (a_i, b_i, x_i) и (a_{2i}, b_{2i}, x_{2i}) по формулам (2.27) и (2.28):

i	a_i	b_i	x_i	a_{2i}	b_{2i}	x_{2i}
1	2	0	10	6	0	11
2	3	0	21	7	1	22
3	6	0	11	15	2	36
4	7	0	36	30	6	11
5	7	1	22	31	7	22

На 5-м шаге значения x_i и x_{2i} совпали. Составим уравнение (2.30):

$$x(31-7)+y(43-1) \equiv 1-7 \pmod{42}, \quad \text{или,} \quad 24x+42y \equiv 36 \pmod{42}.$$

Вычислим Н.О.Д.$(a_j - a_i,\, p-1)$ =Н.О.Д.$(24,\, 42) = 6 \neq 1$.

Поделим все коэффициенты уравнения на 6:

$$4x + 7y \equiv 6 \pmod{42}.$$

С помощью расширенного алгоритма Евклида (см. разд 1.8) решим уравнение $4x + 7y = 1$, подставляя вместо A и B коэффициенты этого уравнения (поменяв их местами, чтобы A было больше B):

A	B	A mod B	A div B	y	x
7	4	3	1	-1	2
4	3	1	1	1	-1
3	1	0	3	0	1

Таким образом, $7 \cdot (-1) + 4 \cdot 2 = 1$, или, $7 \cdot (-6) + 4 \cdot 12 = 6$. Положим, $x_0 = x = 12$. Поскольку, $d > 1$, то корень X уравнения (2.32) определяется с точностью до слагаемое $(p-1)/d = 7$, т.е. имеет вид $X = x_0 + 7k$, где $k \in \mathbf{Z}$. Будем подставлять в (2.32) последовательно числа 12, 19, 25, ..., пока не получим тождество $3^{25} \mod 43 = 15$. Задача решена.

2.6. Факторизация с использованием непрерывных дробей

Один из методов факторизации основан на использовании *непрерывных дробей.* Рассмотрим это важное понятие.

Пусть α — вещественное положительное число. Обозначим буквой q_0 наибольшее целое число, не превосходящее α. При нецелом α имеем $\alpha = q_0 + \frac{1}{\alpha_1}$, $\alpha_1 > 1$. Точно так же при нецелых $\alpha_1, \ldots, \alpha_{s-1}$ имеем

$$\begin{aligned} &\alpha_1 = q_1 + \tfrac{1}{\alpha_2},\ \alpha_2 > 1, \\ &\ldots \\ &\alpha_{s-1} = q_{s-1} + \tfrac{1}{\alpha_s},\ \alpha_s > 1. \end{aligned}$$

Эта процедура даем нам разложение α в непрерывную дробь:

$$\alpha = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \ldots + \cfrac{1}{q_{s-1} + \cfrac{1}{\alpha_s}}}}. \tag{2.33}$$

Будем использовать запись $\alpha = [q_0, q_1, q_2, \ldots q_{s-1}]$ для сокращенного обозначения формулы (2.33).

Если α — иррациональное, то и всякое α_s — иррациональное, и указанный процесс может быть неограниченно продолжен. Если же число α — рационально, то процесс будет конечен и может быть выполнен с помощью алгоритма Евклида.

Пример 1. Разложить дробь $\alpha = \frac{72}{25}$

.

A	B	$A \bmod B$	$q_i = \lfloor A/B \rfloor$
72	25	22	2
25	22	3	1
22	3	1	7
3	1	0	3

$$\frac{72}{25} = 2 + \frac{1}{1 + \frac{1}{7 + \frac{1}{3}}} = [2, 1, 7, 3]$$

Величины $\delta_0 = q_0$, $\delta_1 = [q_0, q_1]$, $\ldots$, $\delta_s = [q_0, q_1, ..., q_s]$, $\ldots$ называются подходящими дробями. Для любых действительных α существует последовательность подходящих дробей такая что $\alpha = \lim_{n\to\infty} \delta_n$.

Пример 2. Приблизить $\alpha = \sqrt{14}$ последовательностью подходящих дробей

Решение. Обозначим через $q_0 = [\sqrt{14}] = 3$. Тогда

$$\alpha = 3+(\sqrt{14}-3) = q_0+\frac{1}{\alpha_1}.$$

$$\alpha_1 = \frac{1}{\sqrt{14}-3} = \frac{\sqrt{14}+3}{5} = 1+\frac{\sqrt{14}-2}{5},\ \ q_1 = 1.$$

$$\alpha_2 = \frac{5}{\sqrt{14}-2} = \frac{\sqrt{14}+2}{2} = 2+\frac{\sqrt{14}-2}{2},\ \ q_2 = 2.$$

$$\alpha_3 = \frac{2}{\sqrt{14}-2} = \frac{\sqrt{14}+2}{5} = 1+\frac{\sqrt{14}-3}{5},\ \ q_3 = 1.$$

$$\alpha_4 = \frac{5}{\sqrt{14}-3} = \sqrt{14}+3 = 6+(\sqrt{14}-3),\ \ q_4 = 6.$$

Последовательность подходящих дробей имеет вид:

$$\delta_0 = 3,\ \ \delta_1 = 3+1/1 = 4,\ \ \delta_2 = 3+\frac{1}{1+1/2} \approx 3.667,\ \ \delta_3 = 3+\frac{1}{1+\frac{1}{2+1/2}} \approx 3.75,$$

$$\delta_4 = 3\frac{20}{27} \approx 3.741\ldots.$$

Можно видеть, что четные члены этой последовательности дают значение $\sqrt{14}$ с недостатком, а нечетные–с избытком, а разности $|\delta_{s+1} - \delta_s|$ убывают и стремятся к нулю.

Вычисление подходящих дробей для квадратных иррациональностей

Квадратными иррациональностями будем называться уравнения вида

$$Ax^2 + Bx + C = 0, \text{ где } A,\ B,\ C \in \mathbf{Z}, \text{ и } D = B^2 - 4AC > 0. \tag{2.34}$$

Корни уравнения (2.34) вычисляются по формуле

$$\alpha = \frac{-B \pm \sqrt{D}}{2A}. \tag{2.35}$$

Вводя обозначения $P = -B$, $Q = 2A$ и рассматривая только больший корень, придем к выражению

$$\alpha = \frac{P + \sqrt{D}}{Q} \tag{2.36}$$

Обозначим через $m = [\sqrt{D}]$. Тогда первая подходящая дробь для корня α имеет вид:

$$\delta_0 = \frac{P + m}{Q} \tag{2.37}$$

Найдем вторую подходящую дробь для корня α. Для этого выделим целую часть $r = [P + \sqrt{D}/Q]$ из (2.35). Дробь, обратная к остатку, будет иметь вид:

$$\alpha_1 = \frac{1}{\alpha - q_0} = \frac{Q}{P + \sqrt{D} - r \cdot Q)} = \frac{Q}{\sqrt{D} - (r \cdot Q - P)} \tag{2.38}$$

Домножим числитель и знаменатель последней дроби на выражение, сопряженное к знаменателю. Получим:

$$\alpha_1 = \frac{Q \cdot (\sqrt{D} + (r \cdot Q - P))}{D - (r \cdot Q - P)^2} \tag{2.39}$$

Поскольку $Q|(D - P^2)$, то знаменатель делится на Q, и после сокращения получим:

$$\alpha_1 = \frac{\sqrt{D} + (r \cdot Q - P)}{Q'} = \frac{P' + \sqrt{D}}{Q'} \tag{2.40}$$

где $P' = rQ - P$, $Q' = (D - (r \cdot Q - P)^2)/Q = (D - (P')^2)/Q$.

Это дает нам реккурентные формулы для вычисления последовательностей $\{P_n\}$, $\{Q_n\}$ и $\{q_n\}$:

$$P_0 = -B,\ \ Q_0 = 2B,\ \ r_0 = \left[\frac{P_0 + \sqrt{D}}{Q_0}\right],$$

$$P_{j+1} = r_j \cdot Q_j - P_j,\ \ Q_{j+1} = (D - P_{j+1}^2)/Q_j,\ \ r_{j+1} = \left[\frac{P_{j+1} + \sqrt{D}}{Q_{j+1}}\right]. \tag{2.41}$$

Теперь подходящие дроби для корня α можно вычислить по формулам:

$$\delta_0 = r_0, \qquad \delta_1 = r_0 + \frac{1}{r_1} = \frac{1 + r_0 \cdot r_1}{r_1},$$

$$\delta_{j+1} = \frac{p_{j+1}}{q_{j+1}} = \frac{r_{j+1} \cdot p_j + p_{j-1}}{r_{j+1} \cdot q_j + q_{j-1}} \tag{2.42}$$

Отметим, что для частного случая $\alpha = \sqrt{n}$, следует взять $D = n$, $P_0 = 0$, $Q_0 = 1$. Можно также для сокращения выкладок на один шаг, взять сразу $P_0 = [\sqrt{D}]$, $Q_0 = D$. Для вычисления целой части r_{j+1} в формулах (2.42) необходимо брать приближение $\sqrt{D} \approx m = [\sqrt{D}]$ при $Q_j > 0$, и $\sqrt{D} \approx m+1$ при $Q_j < 0$.

Замечание. Поскольку, все члены последовательностей $\{P_j\}$ и $\{P_j\}$ ограничены сверху $[\sqrt{D}]$, то число различных пар (P_j, Q_j)–конечно, и найдется номер k такой, что $(P_0, Q_0) = (P_k, Q_k)$. Наименьшее из таких k называется *периодом* последовательности непрерывных дробей.

2.7. Уравнение Пелла

Уравнением *Пелла* называется диофантово уравнение

$$x^2 - ny^2 = \pm 1 \tag{2.43}$$

Решение этого уравнения из-за ошибки Эйлера приписывается англичанину Джону Пеллу (J. Pell 1611-1685), хотя общее решение этого уравнения было впервые найдено в Европе другим англичанином, лордом Вильямсом Браункером (Williams Brouncker 1620–1684), который использовал это

уравнение для получения приближения числа π непрерывными дробями. Однако, гораздо ранее это уравнение уже решалось индийским математиком и астрономом Брахмагуптой (Brahmagupta 598–668).

Решения уравнения Пелла при 1 в правой части определяются следующей теоремой:

Теорема 2.2. *Пусть $n > 0$ –натуральное число, не являющееся полным квадратом. Уравнение Пелла $x^2 - ny^2 = 1$ всегда имеет множество решений (x, y), наименьшее из которое является подходящей дробью $\{x_k/y_k\}$, сходящихся к действительному числу $\sqrt{n}$, где номер k –период последовательности непрерывных дробей. Любое другое решение уравнения является степенью наименьшего решения: $x_t + y_t\sqrt{n} = (x_0 + y_0 \cdot \sqrt{n})^t$, $t = 0, 1, 2, \ldots$.*

Пример 3. Решить уравнение $x^2 - ny^2 = 1$ при $n = 14$.

(1) $\sqrt{14} = 3+(\sqrt{14}-3)$, (2) $\dfrac{1}{\sqrt{14}-3} = \dfrac{3+\sqrt{14}}{5} = 1+\dfrac{\sqrt{14}-2}{5}$,

(3) $\dfrac{5}{\sqrt{14}-2} = \dfrac{2+\sqrt{14}}{2} = 2+\dfrac{\sqrt{14}-2}{2}$, (4) $\dfrac{2}{\sqrt{14}-2} = \dfrac{2+\sqrt{14}}{5} = 1+\dfrac{\sqrt{14}-3}{5}$,

(5) $\dfrac{5}{\sqrt{14}-3} = 3+\sqrt{14} = 6+(\sqrt{14}-3)$.

Период последовательности непрерывных дробей равен 4. Найдем теперь подходящую дробь для $k = 4$ (последнее вычисление не учитывается):

$$3 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1}}} = 3 + \frac{1}{1 + \frac{1}{3}} = 3 + \frac{3}{4} = \frac{15}{4}.$$

Отсюда $x_1 = 15$, $y_1 = 4$ – наименьшее решение уравнения. Проверим решение
$15^2 - 14 \cdot 4^2 = 225 - 224 = 1$. Все решения можно найти, возводя в степень t двучлен $(x_0 + y_0\sqrt{n})$, $t = 0, 1\,2, 3 \ldots$. Например, при $t = 2$ получим $(15 + 4\sqrt{14})^2 = 449 + 120\sqrt{14}$, т.е. $(x_2, y_2) = (449, 120)$.

Отметим, что для вычисления подходящей дроби мы использовали последовательность целых частей дробей α_i $\{r_i\} = \{3, 1, 2, 1, 6 \ldots\}$. Эта

процедура не слишком удобна, поэтому для вычисления прямого вычисления подходящих дробей удобнее реккурентная формула (2.41), приведенная в предыдущем параграфе.

Уравнение Пелла при –1 в правой части не всегда имеет решение, что устанавливается следующей теоремой:

Теорема 2.3. *Пусть* $n > 0$ *–натуральное число, не являющееся полным квадратом. Уравнение Пелла* $x^2 - ny^2 = -1$ *имеет непустое множество решений в том и только в том случае, если период последовательности непрерывных дробей является нечетным числом* k*.*

Если (x_1, y_1) *–наименьшее (нетривиальное) решение* $x^2 - ny^2 = -1$*, то любое другое решение* (x_t, y_t) *можно получить возведением в нечетную степень двучлена* $(x_1 + y_1 \cdot \sqrt{n})$*. При возведении двучлена* $(x_1 + y_1 \cdot \sqrt{n})$ *в четную степень будут получены решения уравнения* $x^2 - ny^2 = 1$*.*

Пример 4. Решить уравнение $x^2 - ny^2 = -1$ при $n = 29$.

Раскладывая действительное число $\sqrt{29}$ в последовательность непрерывных дробей, найдем наименьший период $k = 5$. Решение, соответствующее 5-й подходящей дроби, имеет вид $(x_5, y_5) = (70, 13)$. Проверим уравнение $70^2 - 13^2 \cdot 29 = -1$. Квадрат двучлена $70+13\sqrt{29}$ дает пару $(x_2, y_2) = (9801, 1820)$, являющуюся решением уравнения $x^2 - 29y^2 = 1$.

Описание алгоритма факторизации

Метод факторизации с использованием непрерывных дробей (the continued fraction factorization method CFRAC) был разработан в 1975 г. Мориссоном и Брильхартом (см.[40]).

Пусть n– натуральное число, которое требуется разложить. Рассматривается уравнение Пелла

$$x^2 - y^2 \cdot n = 1$$

и строится последовательность подходящих дробей $\{p_i/q_i\}$, $i = 1, 2, 3, \ldots$

для квадратного корня $\sqrt{n}$ по формулам:

$$P_k = \begin{cases} 0, & \text{если } k = 0 \\ \lfloor \sqrt{n} \rfloor, & \text{если } k = 1 \\ r_{k-1} \cdot Q_{k-1} - P_{k-1}, & \text{если } k \geq 2. \end{cases} \tag{2.44}$$

$$Q_k = \begin{cases} 0, & \text{если } k = 0 \\ n - r_0^2, & \text{если } k = 1 \\ Q_{k-2} + r_{k-1} \cdot (P_{k-1} - P_k), & \text{если } k \geq 2. \end{cases} \tag{2.45}$$

$$r_k = \begin{cases} \lfloor \sqrt{n} \rfloor, & \text{если } k = 0 \\ \lfloor \frac{r_0 + P_k}{Q_k} \rfloor, & \text{если } k \geq 2. \end{cases} \tag{2.46}$$

$$p_k = \begin{cases} r_0, & \text{если } k = 0 \\ 1 + r_0 \cdot r_1, & \text{если } k = 1 \\ r_k \cdot p_{k-1} + p_{k-2}, & \text{если } k \geq 2. \end{cases} \tag{2.47}$$

$$q_k = \begin{cases} 1, & \text{если } k = 0 \\ r_1, & \text{если } k = 1 \\ r_k \cdot q_{k-1} + q_{k-2}, & \text{если } k \geq 2. \end{cases} \tag{2.48}$$

Процесс продолжается до тех пор, пока выражение

$$S_i = p_i^2 - q_i^2 \cdot n$$

не окажется равным полному квадрату $S_i = B^2$. Тогда, $q_i^2 \cdot n = p_i^2 - B^2 = (p_i + B)(p_i - B)$, и с большой долей вероятности n будем иметь нетривиальный общий делитель либо с $(p_i + B)$, либо с $(p_i - B)$. Этот делитель можно найти, вычисляя Н.О.Д.$(n,\, p_i \pm B)$. Если же нам не повезет, то следует продолжить поиск следующего квадрата Q_i, По теореме 2.2 такое решение когда-нибудь появится.

Приведем теперь пример разложения числа $n = 11111$ с использованием метода непрерывных дробей:

k	P	Q	r	p	q	$p^2 - nq^2$
0	0	1	105	105	1	-86
1	105	86	2	211	2	77
2	67	77	2	1527	5	-46
3	87	46	4	2319	22	37
4	97	37	5	12122	115	-91
5	88	91	2	26563	252	25

Значение выражения $p^2 - nq^2$ в последней строке стало равно полному квадрату 25, откуда $252n = 26563^2 - 5^2$. Вычисляя с помощью алгоритма Евклида d=Н.О.Д$(n,\, 26563 + 5)$ =Н.О.Д$(26568,\, 11111)$, найдем нетривиальный делитель $d = 41$ числа n.

Заметим, что наша таблица позволяет также вычислить значение квадратного корня из $n = 11\,111$. Действительно, точное значение $\sqrt{11\,111} = 105,408\,728\,3...$, а последняя подходящая дробь дает отношение $p_5/q_5 = 105,408\,730\,1$, которое отличается от точного меньше, чем на $2 \cdot 10^{-6}$.

Замечание. Отметим здесь, что сами авторы метода CFRAC строили последовательность дробей, сходящуюся к корню $\sqrt{n}$, итерационным методом Ньютона. Лишь позже было замечено, что использовать непрерывные дроби удобнее. В 4-ой главе мы рассмотрим использование непрерывных дробей в методе квадратичного решета.

2.8. Факторизация с использованием квадратичных форм

Метод факторизации, основанный на квадратичных бинарных формах, получил название SQUFOF от английского SQUare FOrm Factorization и был разработан в 1975 г. Даниелем Шенксом (D. Shanks), хотя сам Шенкс не опубликовал ни одной статьи, посвященной этому методу. Подробнее об истории этого метода и о его связи с методом непрерывных дробей можно узнать из статьи Говера и Вагстаффа (J. Gover, S.S. Wagstaff [25]).

Этот метод занимает свою нишу в классификации различных методов

факторизации, являясь самым чемпионом для чисел от 10^{10} до 10^{18}. Кроме того, этот метод используется как вспомогательный при разложении делителей больших чисел типа чисел Ферма.

Определение. Бинарной квадратичной формой называется полином от двух переменных x и y:

$$f(x,y) = ax^2 + bxy + cy^2 = (x \;\; y) \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Будем записывать квадратичную форму $f = ax^2 + bxy + cy^2$ в более кратком виде $f = (a,b,c)$. *Дискриминантом* формы (a,b,c) называется число $D = b^2 - 4ac$. Формы с отрицательным дискриминантом называются *определенными* – *definite*, а формы с положительным дискриминантом *неопределенными* – *indefinite*. В методе Шенкса используются только неопределенные формы.

Определение. Две квадратичные формы $f = (a,b,c)$ и $g = (p,q,r)$ называются *эквивалентными*, если найдется целочисленная матрица

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

с определителем равным 1, переводящая матрицу f в матрицу g:

$$\begin{pmatrix} p & q \\ 0 & r \end{pmatrix} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \cdot \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

Поскольку эквивалентное преобразование не меняет детерминант формы, необходимым условием эквивалентности двух форм является равенства их детерминантов. Однако, обратное неверно. Среди форм с одинаковым дискриминантом может найтись конечное число неэквивалентных.

Среди класса всех квадратичных форм, эквивалентных данной, имеют особое значение так называемые *редуцированные* формы. Форма $f = (a,b,c)$ называется *редуцированной*, если выполняется неравенство

$$|\sqrt{D} - 2|a|| < b < \sqrt{D}. \tag{2.49}$$

Метод получения редуцированной формы, эквивалентной данной, был найден еще Карлом Гауссом и состоит в последовательном применении оператора *редукции* $f = \rho(f)$, пока f не станет редуцированной:

Определение. Оператор редукции $\rho(f)$– это оператор на классе квадратичных бинарных форм с целыми коэффициентами, который определяется следующим образом:

$$\rho(f) = \rho(a, b, c) = \left(c,\ r,\ \frac{r^2 - D}{4c}\right), \tag{2.50}$$

где $r = r(-b, c)$–целое число, однозначно определяемое следующими условиями:

1. $r \equiv -b \,(\bmod 2c)$,
2. $-|c| < r \leq |c|$, если $\sqrt{D} < |c|$,
3. $\sqrt{D} - 2|c| < r < \sqrt{D}$, если $|c| < \sqrt{D}$.

Результат применения оператора ρ к форме f n раз записывается $\rho^n(f)$. Обратный к ρ оператор ρ^{-1} определяется следующим образом:

$$\rho^{-1}(a, b, c) = \left(\frac{r^2 - D}{4a},\ r,\ a\right), \tag{2.51}$$

где $r = r(-b, c)$ определено как и раньше. Справедлива следующую теорему:

Теорема 2.4. *Каждая форма f эквивалентна некоторой редуцированной форме, и любая редуцированная форма для f равна $\rho^k(f)$ для некоторого положительного k. Если f–редуцирована, то $\rho(f)$ также редуцирована.*

Определение. Две формы $(a,\, b,\, c)$ и $(c,\, b',\, c')$ называются *смежными (adjacent)*, если выполняется условие $b + b' \equiv 0 \,(\bmod 2c)$. Каждая форма $(a,\, b,\, c)$ эквивалентна некоторой смежной форме, а именно,

$$(a,\, b,\, c) \sim (c,\, -b,\ a) \tag{2.52}$$

Также на множестве эквивалентных форм можно определить операцию умножения (композицию), тогда, если коэффициенты a и b

взаимно–просты,

$$(a,\, b,\, ac)^2 \sim (a^2,\, b,\, c) \tag{2.53}$$

Если в квадратичной форме $f = (a,\, b,\, c^2)$, третий коэффициент является полным квадратом, то такую форму будем называть *квадратной (square)*. Из квадратной формы f можно извлечь квадратный корень. Для вычисления корня заменим форму f на эквивалентную ей смежную форму $f \sim (c^2,\, -b,\, a)$, потом извлечем квадратный корень по формулам (2.53). Получим:

$$g = f^{1/2} = \sqrt{(a,\, b,\, c^2)} = \sqrt{(c^2,\, -b,\, a)} = (c,\, -b,\, ac). \tag{2.54}$$

Определение. Форма вида (k, kn, c) называется *неоднозначной (ambiguous)*. Если форма – неоднозначна, то ее определитель делится на k: $D = (kn)^2 - 4kc = k(kn^2 - 4c)$.

Идея метода Шенкса состоит в сопоставлении числу n, которое надо разложить, квадратичной бинарной формы f с дискриминантом $D = 4n$, с которой потом выполняется серия эквивалентных преобразований и переход от формы f к неоднозначной форме (a', b', c'). Тогда, Н.О.Д.$(a',\, b')$ будет являться делителем n. Более подробно алгоритм может быть записан в следующем виде:

Вход: Нечетное составное число n, которое требуется факторизовать. Если $n \bmod 4 = 1$, заменим n на $2n$. Теперь $n \bmod 4 = 2$ или 3. Последнее свойство нужно, чтобы определитель квадратичной формы был фундаментальным, что обеспечивает сходимость метода.

Выход: Нетривиальный делитель n.

1. Определим исходную квадратичную форму $f = (1,\, 2b,\, b^2 - D)$, с дискриминантом $D = 4n$, где $b = \lfloor \sqrt{n} \rfloor$.

2. Выполним цикл редуцирований $f = \rho(f)$, пока форма f не станет квадратной:

   while not (f square) **do** $f = \rho(f)$;

3. Вычислим квадратный корень из f по формулам (2.54):

   $g = (a',\, b',\, c') = f^{1/2}$

4. Выполним цикл редуцирований $g = \rho(g)$, пока значение второго коэффициента не стабилизируется $b'_{i+1} = b'_i$. Число итераций m этого цикла должно быть примерно равно половине от числа итераций первого цикла. Последнее значение a' даст делитель числа n (возможно тривиальный).

Дадим теперь полное описание этого алгоритма с соответствующими формулами. Отметим, что хотя теоретическая часть алгоритма связана с эквивалентными преобразования квадратичных форм, практическая часть алгоритма выполняется по формулам (2.44)–(2.46) вычисления коэффициентов P, Q и r метода непрерывных дробей без обращения к формам. Каждая итерация цикла соответствует одной операции применения оператора редукции к соответствующей форме. При необходимости можно восстановить соответствующие формы $f_k = (a_k,\, b_k,\, c_k)$ по формулам:

$$(a_k,\, b_k,\, c_k) = \big((-1)^{k-1}Q_{k-1},\ 2P_k,\ (-1)^k Q_k\big) \tag{2.55}$$

Ниже мы дадим описание алгоритма SQUFOF для практической реализации:

Алгоритм SQUFOF нахождения нетривиального делителя n

Вход: Составное число n.
Выход: Нетривиальный делитель n.

I. Инициализация алгоритма.

1. Проверим, является ли n полным квадратом. Если да, то вычислим $d = \sqrt{n}$, и завершим вычисление. Иначе, перейдем к следующему пункту.

2. Если $n \equiv 1 \pmod 4$, тогда заменим n на $2n$. Определим $D = 4n$, $q_0 = \lfloor \sqrt{D} \rfloor$.

3. Определим исходные значения параметров P, Q, r:

$$P_0 = 0, \quad Q_0 = 1, \quad r_0 = P_1 = \lfloor \sqrt{n} \rfloor, \quad Q_1 = n - r_0^2, \quad r_1 = \lfloor 2r_0/Q_1 \rfloor.$$

II. Первый цикл. Последующие значения параметров P и Q будем вычислять по формулам формулами (2.44)-(2.46) вычисления частичных дробей в методе факторизации непрерывных дробей CFRAC (раздел 2.6):

$$P_k = r_{k-1} \cdot Q_{k-1} - P_{k-1}, \; Q_k = Q_{k-2} + (P_{k-1} - P_k) \cdot r_{k-1}, \; r_k = \left\lfloor \frac{P_k + \lfloor \sqrt{n} \rfloor}{Q_k} \right\rfloor, \; k \geq 2. \tag{2.56}$$

Продолжим вычисления коэффициентов P_k, Q_k и r_k, $k = 2, 3, \ldots$ до тех пор, пока не найдем Q_k, являющееся полным квадратом. Это должно произойти при некотором четном k. Пусть $Q_k = d^2$ для целого $d > 0$. Перейдем к следующему циклу.

III. Второй цикл.

Начнем цикл вычислений новых параметров P'_j, Q'_j, r'_j, $j = 0, 1, 2, \ldots$. Формулы для реализации второго цикла останутся такими же, как раньше. Изменятся только начальные значения параметров P', Q' и r':

$$P'_0 = -P_k, \;\; Q'_0 = d, \;\; r'_0 = \left\lfloor \frac{P'_0 + \lfloor \sqrt{n} \rfloor}{Q'_0} \right\rfloor, \;\; P'_1 = r'_0 \cdot Q'_0 - P'_0, \;\; Q'_1 = (N - P_1'^2)/Q'_0.$$

$$P'_j = r'_{j-1} \cdot Q'_{j-1} - P'_{j-1}, \; Q'_j = Q'_{j-2} + (P'_{j-1} - P'_j) \cdot r'_{j-1}, \; r'_j = \left\lfloor \frac{P'_j + \lfloor \sqrt{n} \rfloor}{Q'_j} \right\rfloor, \; j \geq 2.$$

Вычисление следует продолжать, пока два подряд идущих значения P'_j и P'_{j+1} не окажутся равными. Тогда, значение Q_j даст искомый делитель числа n.

Описание алгоритма Шенкса закончено.

Пример факторизации по методу Шенкса

Выполним факторизацию нашего примера из предыдущего раздела $n = 11\,111$.

Инициализация. Найдем $n \bmod 4 = 3$, $r_0 = \lfloor \sqrt{n} \rfloor = 105$.

Цикл 1. Составим таблицу вычисления значений коэффициентов P, Q и r по формулам (2.56). Вычисление заканчивается после того, как в столбце Q появится полный квадрат:

k	P	Q	r
0	0	1	105
1	105	86	2
2	67	77	2
3	87	46	4
4	97	37	5
5	88	91	2
6	94	**25**	7

В столбце Q найден полный квадрат $d^2 = 25$, откуда находим значение $d = 5$, используемое во втором цикле.

Цикл 2. Выполняем второй цикл вычислений. Цикл продолжается, пока в столбце P' не появятся подряд два одинаковых значения. Тогда значение Q_{j-1}, находящееся в предпоследней строке столбца Q, является искомым делителем числа n.

j	P'	Q'	r'
0	-94	5	2
1	104	59	3
2	73	98	1
3	25	107	1
4	82	**41**	4
5	82	107	1

В столбце P были найдены два подряд идущих значения 82. Завершаем вычисление. Предпоследнее значение в столбце Q содержит искомый делитель 41 числа $n = 11\,111$.

Отметим, что теоретическое число итераций 2-ого цикла должно быть примерно равно половине от числа итераций первого цикла (в нашем примере $j = 4$ оказалось на единицу больше половины $k = 6$).

Оценка сходимости метода Шенкса

Согласно расчетам, выполненным самим Шенксом, число итераций первого и второго циклов определяется числом w сомножителей числа n и равно примерно

$$\frac{C}{2^w - 2} \cdot n^{1/4},$$

где C–константа, равная примерно $2,4$ для первого цикла итераций. Таким образом, метод квадратичных форм Шенкса имеет асимптотическую сложность $\mathrm{O}(n^{1/4})$ и является, как уже упоминалось раньше, наиболее быстрым методом в классе алгоритмов для разложения чисел длиной до 18 десятичных знаков.

3. Эллиптические кривые и их приложения в криптографии

Хотя эллиптические кривые (Elliptic Curves) исследовались на протяжении более сотни лет, интерес к ним проявляли исключительно узкие специалисты в области теории чисел. Так было примерно до 1985 г., пока одновременно и независимо Н. Коблиц (N. Coblitz) и В. Миллер (V. Miller) не предложили использовать эллиптические кривые для построения криптосистем с открытым ключом.

После этого интерес к эллиптических кривых стал расти в геометрической прогрессии. Были найдены приложения инструмента ЭК в разных областях криптографии таких, как теория кодирования, генерация псевдослучайных последовательностей, алгоритмическая теория чисел для построения тестов на простоту и, наконец, для создания одного из самых красивых методов факторизации целых чисел (Х. Ленстра [31]).

Метод факторизации Ленстры можно рассматривать как модификацию $(p-1)$–метода Полларда. Он является самым быстрым среди всех методов, упомянутых ранее. Как и $(p-1)$–метод Полларда, сложность этого метода определяется величиной не самого числа n, а величиной его наименьшего делителя, поэтому, даже если число n очень велико и недоступно другим алгоритмам, оно может быть проверено с помощью метода факторизации эллиптических МФЭК. Подобно $(p-1)$–методу Полларда (см.разд. 2.2), МФЭК состоит из двух стадий. Первая стадия алгоритма была разработана самим Ленстрой и имеет единственный вариант. Вторая стадия имеет несколько вариаций. Одна из них, основанная на парадоксе близнецов, была предложена Брентом. [8].

В этой главе мы рассмотрим основные свойства эллиптических кривых и их приложения в теории чисел и криптографии. Среди литературы на русском языке, относящейся к теме эллиптический кривых отметим, в первую очередь, книгу Н. Коблица «Курс теории чисел и криптографии» [69] и две книги А. Болотова, С. Гашкова, А. Фролова и А. Часовских под названием «Элементарное введение в эллиптическую криптографию»

[60] и «Алгоритмические основы эллиптической криптографии» [61]. На английском языке следует отметить, в первую очередь, 2-е издание книги Л. Вашингтона «Elliptic Curves Number Theory and Cryptography». [54] Хороший обзор по алгоритмам использования эллиптических кривых в криптографии можно найти в [16].

Начнем наше изложение с определения эллиптической кривой над конечным полем.

3.1. Определение эллиптической кривой

Определение. Пусть F_q, $q = p^k$, конечное поле характеристики $p \geq 2$. Эллиптической кривой над полем F_q называется множество точек $(x, y) \in F_q \oplus F_q$, удовлетворяющих уравнению Вейерштрассе

$$y^2 + ay + b = x^3 + cx^2 + dx + e \pmod{q}. \tag{3.57}$$

Кроме того, к множеству точек ЭК добавляется специальная точка, обозначаемая через ∞ и называемая точкой в бесконечности.

Если характеристика поля $p \geq 3$ (а именно этот случай для нас наиболее интересен), уравнение (3.57) может быть преобразовано путем замены переменных в уравнение

$$y^2 = x^3 + ax + b \pmod{q}, \tag{3.58}$$

где $a, b \in F_q$. Дополнительным требованием на параметры a и b является условие $4a^3 + 27b^2 \neq 0$, в силу которого дискриминант полинома $x^3 + ax + b$ не равен 0, и полином не имеет кратных корней.

Поскольку все операции в уравнении ЭК выполняются по модулю числа q, знак равенства в уравнении (3.58) следовало бы заменить знаком эквивалентности $\equiv$, однако, следуя традициям записи уравнения ЭК, мы используем знак $=$.

На множестве точек E эллиптической кривой можно определить групповую операцию сложения $+$, с помощью которой это множество становится аддитивной абелевой группой с точкой ∞ в качестве нуля.

Пусть $P=(x,y)\in E$, тогда обратной к точке P является точка $-P=(x,-y)$. Сумма $P+(-P)=\infty$. Сумма точек $P=(x_1,y_1)$ и $Q=(x_2,y_2)$, где $P\neq -Q$, вычисляется по формулам:

$$\begin{aligned} x_3 &= \lambda^2 - x_1 - x_2 \\ y_3 &= \lambda(x_1 - x_3) - y_1 \\ \lambda &= \begin{cases} \frac{y_2-y_1}{x_2-x_1}, \text{ если } P\neq Q, \\ \frac{3x_1^2+a}{2y_1} \text{ если } P=Q \end{cases} \end{aligned} \tag{3.59}$$

Группа точек эллиптической кривой над полем F_q обозначается символом $E(F_q)$, а ее мощность (количество элементов) символом $\#E(F_q)$. Известно, что $E(F_q)\cong C_{n_1}\oplus C_{n_1}$, где C_n - циклическая группа порядка n, n_2 делит n_1, и n_2 делит $q-1$.

Пример. Пусть $E(F_q)$ - группа точек кривой $y^2=x^3+x+1$ над полем F_{23}. Эта группа является циклической с генератором $P(0,1)$. Рассмотрим все кратные kP точки P:

$P(0,1)$	$2P=(6,-4)$	$3P=(3,-10)$	$4P=(-10,-7)$
$5P=(-5,3)$	$6P=(7,11)$	$7P=(11,3)$	$8P=(5,-4)$
$9P=(-4,-5)$	$10P=(12,14)$	$11P=(1,-7)$	$12P=(-6,-3)$
$13P=(9,-7)$	$14P=(4,10)$	$15P=(9,7)$	$16P=(-6,3)$
$17P=(1,7)$	$18P=(12,-4)$	$19P=(-4,5)$	$20P=(5,4)$
$21P=(11,-3)$	$22P=(7,-11)$	$23P=(-5,-3)$	$24P=(10,7)$
$25P=(3,10)$	$26P=(6,4)$	$27P=(0,-1)$	$28P=(\infty)$

Таким образом, данная кривая содержит 28 точек. *Порядок точки* A — это наименьшее натуральное число k такое, что $kA=\infty$. Порядок любой точки является делителем порядка числа точек, поэтому, порядок любой точки на данной кривой принадлежит множеству $\{1,\,2,\,4,\,7,\,14,\,28\}$.

Пример. Найти сумму точек $3P=(3,-10)$ и $7P=(11,3)$.

Решение. Вычислим $\lambda=(y_2-y_1)/(x_2-x_1)$: $y_2-y_1=3-(-10)=13$, $x_2-x_1=11-3=8$. Поскольку $8^{-1}\equiv 3\,(\mathrm{mod}\,23)$, то $\lambda=13/8=$

$13 \cdot 3 \bmod 23 = 16$. Теперь, $x_3 = \lambda^2 - x_1 - x_2 = (16^2 - 3 - 11) \bmod 23 = 12$, а $y_3 = \lambda(x_1 - x_3) - y_1 = (16 - 12 - (-10)) \bmod 23 = 14$.

Ответ: $(3, -10) + (11, 3) = (12, 14)$.

Замечание. При выполнении удвоения точки или вычислении суммы необходимо вычисление обратного элемента в поле F_q. Эта операция выполняется с помощью обобщенного алгоритма Евклида (см.разд.1.8).

Вычисление кратного kQ заданной точки Q

Поскольку, арифметика эллиптических кривых не содержит прямых формул для вычисления кратного kQ для заданной точки $Q(x_1, y_1)$, то эту операцию выполняют с использованием операций сложения, вычитания и удвоения точки. Для этого надо представить число k в двоичной системе исчисления $k = b_t\, b_{t-1} \ldots b_0$, $b_i \in \{0, 1\}$, потом вычислить все точки $2Q$, $4Q$, ... , $2^t \cdot Q$ и подсчитать сумму тех точек $2^i \cdot Q$, для которых $b_i = 1$.

Пример. Пусть $k = 13$. В двоичной системе $k = 1\,1\,0\,1_2$, тогда, $13Q = 8Q + 4Q + Q$. Эту же точку можно вычислить как $16Q - 2Q - Q$.

Преобразование уравнения кривой в проективные координаты

Нахождение суммы и кратного точек ЭК требует вычисления обратного элемента в конечном поле. Это трудоемкая операция, требующая использование обобщенного алгоритма Евклида. Можно однако избавиться от частого использования этой операции, если рассматривать уравнение эллиптической кривой в трехмерных проективных координатах. Рассмотрим новое уравнение кривой

$$Y^2 Z = X^3 + aXZ^2 + Z^3, \tag{3.60}$$

которое получено из уравнения (3.58) при переходе к проективным координатам. В этом случае кривая $E(F_q)$ рассматривается как множество классов эквивалентности (X, Y, X) с отношением эквивалентности $(X, Y, X) \sim (aX, aY, aX)$. Будем называть такие классы точками и обозначать их любым представителем (X, Y, X) класса. Бесконечно

удаленной точки ∞ будет соответствовать класс $(0,1,0)$ проективной плоскости. Если $P = (X,Y,Z) \neq \infty$, тогда сопоставляя точке P точку $P' = (x,y)$, $x = X/Z$, $y = Y/Z$, получим взаимно-однозначное соответствие между точками между точками исходной ЭК и классами проективной ЭК.

Для получения формул сложения и удвоения точек ЭК в проективных координатах заметим, что коэффициент λ входит в формулы (3.58) во второй и третьей степени и содержит знаменатель $2y_1$, поэтому, чтобы в результате знаменатель сократился полностью, надо домножить полученные координаты новой точки (X', Y', Z') на коэффициент $8y_1^3$. После соответствующих преобразований, получим формулы для вычисления операции удвоения точек в проективных координатах, не использующую вычисление обратного элемента:

$$\begin{cases} X_3 = 2Y_1Z_1((3X_1^2 + aZ_1^2)^2 - 8X_1Y_1^2Z_1)(\mathrm{mod}p) \\ Y_3 = 4Y_1^2Z_1(3X_1(3X_1^2 + aZ_1^2) - 2Y_1^2Z_1) - (3X_1^2 + aZ_1^2)^3(\mathrm{mod}p) \\ Z_3 = 8Y_1^3Z_1^3(\mathrm{mod}p) \end{cases} \tag{3.61}$$

При сложении двух точек примем $Z_3 = Z_1Z_2(X_1X_2 - X_1Z_2)^3(\mathrm{mod}p)$, откуда две другие координаты будут находиться по формулам:

$$\begin{cases} X_3 = (X_2Z_1 - X_1Z_2)[Z_1Z_2(Y_2Z_1 - Y_1Z_2)^2 - (X_2Z_1 + X_1Z_2)\cdot \\ (X_2Z_1 - X_1Z_2)^2](\mathrm{mod}p) \\ Y_3 = (X_2Z_1 - X_1Z_2)^2[Y_2Z_1(Y_2Z_1 + 2Y_1Z_2) - Y_1Z_2(X_1Z_2 + 2X_2Z_1)] - \\ -Z_1Z_2(Y_2Z_1 - Y_1Z_2)^3(\mathrm{mod}p) \end{cases} \tag{3.62}$$

3.2. Число точек эллиптической кривой

Одной из трудных проблем, имеющих важное значение для приложений, является проблема вычисления количества точек на эллиптической кривой. Известное неравенство Хассе (Hasse) утверждает, что

$$\#E(F_q) = q + 1 - t, \tag{3.63}$$

где $|t| \leq 2\sqrt{q}$.

Если выполнено условие $p \mid t$, то кривая называется *суперсингулярной (supersingular)*, иначе кривая называется *обыкновенной (ordinary)*. Отметим, что условие $p \mid t$ при $p \geq 5$ эквивалентно условию $t = 0$.

Неравенство Хассе вытекает из уравнения

$$\#E(F_q) = p^k + 1 - \sum_{x \in F_{p^k}} \chi(x^3 + ax + b), \tag{3.64}$$

где $\chi(z)$ – квадратичный характер в поле F_q (иными словами, $\chi(z) = 1, -1, 0$ в зависимости от того, является z квадратичным вычетом, квадратичным невычетом или равен 0). Напомним, что квадратичные вычеты можно вычислять с помощью символа Лежандра (см. раздел 1.9). Однако практически формула (3.64) не применима, поскольку выполнение расчетов с ее использованием занимает слишком много времени.

Уравнение для числа точек эллиптических кривых описывается теоремой В. Вотерхауза, см. монографию Л. Вашингтона [54], теор. 4.3, с. 98:

Теорема 3.1. *(W. Waterhous). Пусть дано поле F_q, $q = p^n$ и число $N = q + 1 - t$. Над полем F_q существует эллиптическая кривая с числом точек*

$\#E(F_{p^k}) = N \Leftrightarrow t \leq 2\sqrt{q}$ *и выполнено одно из следующих условий:*

1. *Н.О.Д.(a, p)=1,*
2. *n –четно, и* $t = \pm 2\sqrt{q}$,
3. *n –четно,* $p \not\equiv 1 \pmod 3$, *и* $t = \pm\sqrt{q}$,
4. *n –четно,* $p \not\equiv 1 \pmod 4$, *и* $t = 0$,
5. *n –нечетно,* $p = 2$ *или 3, и* $t = \pm p^{(n+1)/2}$,
6. *n –нечетно, и* $t = 0$.

Из неравенства Хассе вытекает, что число точек на эллиптической кривой отличается от мощности поля $q = p^n$ самое большее на величину t меньшего порядка $O(q^{1/2})$. Однако вычисления в абелевой группе точек

эллиптической кривой более громоздкие, чем в конечных полях. А это значит, что для произвольной точки G вычисление множителя k такого, что $G = kP$, где P генератор точек кривой, т.е. решение проблемы, аналогичной вычислению дискретного логарифма в конечных полях, решается более трудоемко. Поэтому на группах точек эллиптических кривых можно строить криптографические протоколы типа протокола Диффи-Хелмана выработки общего секретного ключа, электронной цифровой подписи или шифрования информации, выбирая размерность ключа (определяемую здесь размерностью поля F_{p^k}) меньшей длины. Было подсчитано, что длина ключа в 160 бит на эллиптических кривых соответствует ключу длины 1024 бита в методе RSA (см.[74], с.132).

3.3. Алгоритм факторизации Ленстры

Перейдем далее к методу Ленстры факторизации целых чисел. Пусть n – составное число, для которого требуется найти наименьший делитель p. В отличии от других методов факторизации, рассматриваемых здесь, производительность метода Ленстры зависит только от размера этого наименьшего множителя, а не от размерности числа n.

Рассмотрим множество $Z_n = \{0, 1, 2, ..., n-1\}$ как основное множество для координат точек эллиптической кривой $EC(Z_n)$: $y^2 = x^3 + ax + b$. В строгом математическом смысле эта кривая не будет эллиптической кривой (Ленстра назвал такую кривую *псевдокривой*), т.к F не является полем, и, значит, в нем не всегда выполнимы операции нахождения обратного элемента, необходимые для нахождения суммы точек кривой. Однако Ленстра заметил, что невозможность вычисления суммы двух точек $P(x_1, y_1)$ и $Q(x_2, y_2)$ означает, что разность первых координат $x_2 - x_1$ должны равняться 0 по модулю одного из делителей n, тогда, вычисляя наибольший общий делитель Н.О.Д. $(n, x_2 - x_1)$, мы легко найдем искомый делитель.

Суть алгоритма Ленстры заключается в выборе на псевдокривой $EC(Z_n)$ произвольной базовой точки P_0 и домножении ее на всевозможные

простые числа и их степени пока не получим

$$kP_0 = \infty \pmod p, \tag{3.65}$$

где p–один из делителей n.

Замечание 1. Поскольку, ни один из делителей n нам заранее не известен, то условие выполнения (3.65) невозможно проверить, поэтому признаком успешного завершения работы алгоритма является выполнение условия Н.О.Д. $(n, C) = d > 1$ при очередном вычислении коэффициента λ в операции удвоения или сложения точек при вычислении очередного кратного C точки P_0.

Замечание 2. Работа алгоритма состоит из двух стадий, называемых этапом 1 и этапом 2 (stage-one and stage-two). На первом этапе существенную роль играет настраиваемый параметр B_1, называемый ограничителем 1 этапа (stage-one limit). По сути алгоритм Ленстры является полным аналогом $(p-1)$–алгоритма Полларда (см.разд 2.2), где операция возведения в степень простого числа p заменена операцией домножения точки ЭК на множитель p. В остальном, организация работы первого и второго этапов может быть выполнена полностью аналогично работе $(p-1)$–метода.

Описание первой стадии алгоритма

I. *Инициализация:*

1. Выберем некоторое значение B_1, например, $B_1 = 10000$.

2. Выберем случайным образом числа $x, y, a \in [0, n-1]$.

3. Вычислим $b = y^2 - x^3 - ax \mod n$ и $g =$ Н.О.Д. $(n, 4a^3 + 27b^2)$. Если $g = n$, возвращаемся к п.2. Если $1 < g < n$, тогда прекратим вычисление – делитель найден. Иначе, определим кривую $E : y^2 = x^3 + ax + b$ и базовую точку-генератор $P_0(x, y)$.

4. Присвоим изменяемому параметру $P(x, y)$ начальное значение, равное P_0.

II. *Вычисление:*

1. Для каждого простого числа $p < B_1$ найдем наибольшую степень r такую, что $p^r < B_1$. Выполним цикл for $(j = 0; j < r; j++)$ $P = p \cdot P$, в результате которого точка P домножится на p^r. Каждое умножение на p выполняется с помощью алгоритма нахождения кратного точки, описанного на с.86.

2. Продолжим вычисление до тех пор, пока не будут пройдены все простые числа, меньшие B_1, или не найдется шаг, на котором выполнится условие Н.О.Д $(n, P) = d > 1$.

Если выполнится последнее условие, то искомый делитель n найден.

Иначе, либо увеличиваем B_1 и повторяем все заново, либо переходим ко второй стадии алгоритма.

Вторая стадия алгоритма

На второй стадии алгоритма предполагается, что остался один большой простой множитель $q > B_1$ такой, что домножение точки P, полученной в конце первого этапа, приведет к выполнению условия (3.65).

1. Выберем новую границу B_2, и выпишем все простые числа из интервала $[B_1;\ B_2] : \{q_1, q_2, ..., q_m\}$.

2. Будем последовательно вычислять точки $q_1 \cdot P,\ q_2 \cdot P,\ q_3 \cdot P,\ ...$ пока не дойдем до границы B_2, либо не выполнится условие (3.65).

Как и в $(p-1)$–методе Полларда, чтобы вычислить очередную точку $q_{i+1} P$ достаточно прибавить к ранее вычисленной точке $q_i P$ точку $\delta_i P$, где $\delta_i = q_{i+1} - q_i$. Поскольку простые числа расположены достаточно близко друг к другу, различных точек вида $\delta_i P$ будем немного. Их можно вычислить заранее и расположить в некотором массиве. Тогда каждое новое вычисление $q_i P$ можно выполнить с помощью только одной операции сложения. Поэтому вторая часть алгоритма выполняется очень быстро.

В наиболее простом варианте реализации второй стадии можно вычислить только одну точку $2P$ и прибавлять ее к точке $q_1 P$ пока не получим требуемое условие (3.65).

Пример 1. Пусть требуется разложить число $n = 455\,839$. Выберем эллиптическую кривую

$$y^2 = x^3 + 5x - 5,$$

точку $P = (1, 1)$ на ней и постараемся вычислить $10!\,P$.

1. Найдем сначала $2P$. Тангенс угла наклона касательной λ в т.P равен $\lambda = (3 \cdot 2 + 5)/(2y) = 4$ и координаты $P_2 = 2P = (x_2, y_2) = (14, -53) \pmod{n}$.

2. Вычислим далее, $P_3 = 3(2P) = 3P_2$. Прямой формулы для вычисления точки $3P_2$ нет, поэтому придется вычислить сначала $2P_2$, затем получить $3P_2$, суммируя точки $2P_2$ и P_2. Получим $2P_2 = (259\,851, 116\,255)$, $3P_2 = (195\,045,\ 123\,227)$.

3. Продолжая эту процедуру вычислим $4!P$, потом $5!P$ и т.д. При вычислении $8!P$ знаменатель λ станет равным 599 и вычисление Н.О.Д.$(n, 599)$ даст значение $d = 599$. Отсюда 599 является делителем n, и деля n на 599 найдем второй делитель n: $455839 = 599 \cdot 761$.

Причина, по которой процесс сошелся при вычислении $8!P$, состоит в том, что кривая $y^2 = x^3 + 5x - 5 \pmod{599}$ содержит $640 = 27 \cdot 5$ точек. Вторя кривая $y^2 = x^3 + 5x - 5 \pmod{761}$ содержит $640 = 27 \cdot 5$ точек. Число 8! делится на 640, но не делится на 777. Поэтому первым появился делитель $p = 599$.

Анализ метода Ленстры

Проведем теперь анализ метода Ленстры и оценим условия, при которых он будет успешно завершен. До сих пор все вычисления проводились по модулю числа n, однако, если координаты полученных точек вычислять по модулю p, являющегося делителем n, тогда условием успешного завершения алгоритма будет, очевидно, условие

$$kP = \infty, \text{ где } k = \prod_{p_i^{a_i} \le B_1} p_i^{a_i}, \tag{3.66}$$

и эллиптическая кривая $y^2 = x^3 + ax + b$ рассматривается в конечном поле F_p.

Пусть $l = \#E(F_p)$ число точек этой кривой. По неравенству Хассе $l \in [p + 1 - 2\sqrt(p), p + 1 + 2\sqrt(p)]$. Для любой точки $Q(x, y)$ выполняется условие $lQ = \infty$, поэтому, для того, что алгоритм Ленстры успешно завершился, необходимо, чтобы множитель k в уравнении (3.66) делился на порядок кривой l. Последнее условие будет выполнено, если все делители l не превышают границы B_1.

Дадим здесь определение *гладкости* целого числа (smoothness), которое будет широко использоваться в последующих разделах. Пусть B – некоторое положительное целое число. Произвольное целое число x называется B – гладким, если все делители x по модулю не превышают B. Например, $x = 2^5 \cdot 5 \cdot 13^2$ является B-гладким для любого $B \geq 13$.

Это условие гладкости является более слабым, чем требуется в алгоритме Ленстры. Для успешного завершения этого алгоритма *необходимо*, чтобы все делители числа l вида p^r, кроме последнего, были меньше границы B_1, а наибольший делитель p^r имел степень $r = 1$ и был меньше границы B_2. Например, для $\#EC(F_p) = 2^5 \cdot 5 \cdot 13^2 \cdot 233$ границы B_1, B_2 должны удовлетворять условиям $B_1 \geq 13^2 = 169$, $B_2 \geq 233$.

Число l, любой делитель которого вида p^r, где p–простое число, меньше границы B, называется B*–гладкостепенным.* .

Отметим, что необходимая граница для степеней делителей l существенно зависит от значения $\#EC(F_p)$, которое, в свою очередь, определяется коэффициентами a и b кривой. К сожалению, нет никакого регулярного способа выбрать кривую с наименьшим значением максимальной степени делителя $\#EC(F_p)$.

Пример 2. Рассмотрим простое число $p = 1007$ и вычислим наименьшие значения границ B_1, B_2 для каждого целого числа k из интервала $[1001, 1013]$, окружающего p. Получим:

k	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013
$\mathbf{B}_1$	7	3	1	4	5	2	1	16	1	5	1	11	1
$\mathbf{B}_2$	143	167	1003	251	67	503	1007	16	1009	101	1011	23	1013

Этот пример показывает, что факторизацию числа n, имеющего в разложении множитель $p = 1007$, может выполнить при $B_1 = B_2 = 16$, а может потребовать границы B_2, сравнимой с числом 1007, в зависимости от выбранной кривой. Отметим также, что границу B_1 в большинстве случаев можно взять очень небольшой (наибольшее значение = 16), а граница B_2 почти всегда очень большая.

Поэтому процедуру факторизации на ЭК следует *всегда* выполнять одновременно с несколькими различными кривыми.

Классический алгоритм Ленстры 1987 г. завершается в своей первой стадии. Последующие улучшения этого алгоритма, выполненные Монтгомери, Брентом и др., позволили решать задачу факторизации n даже в случае, если порядок l содержит более одного множителя, превышающего B_1. Описание алгоритма Монтгомери для вычисления кратного точки ЭК можно найти в книге Болотова и др. [61], а улучшения Монгомери к методу Ленстры в статьях [36] и [37].

Оценка эффективности метода эллиптических кривых Ленстры

Пусть наименьший множитель числа n равен p. Тогда, время работы алгоритма Ленстры можно оценить величиной

$$exp\left(\sqrt{2} + o(1)\sqrt{\ln p \ln\ln p)}\right), \tag{3.67}$$

которая выполняется в случае, если граница B_1 выбрана близко к величине

$$exp\left(\sqrt{2}/2 + o(1)\sqrt{\ln p \ln\ln p)}\right).$$

Поскольку значение множителя p неизвестно, то выбор значения B_1 выполняется эмпирически, что несколько ухудшает практическую оценку сходимости метода Ленстры. Отметим, что добавление в алгоритм Ленстры второй стадии вычислений сохраняет общую асимптотическую оценку,

хотя обеспечивает большой практический прирост скорости сходимости алгоритма.

Если сравнивать метод эллиптических кривых с другими методами факторизации, то метод Ленстры относится к классу субэкспоненциальных методов факторизации, а, значит, работает быстрее любого метода, упомянутого во второй главе.

Если сравнивать его с методом квадратичного решета QS и методом решета числового поля NFS, то все зависит от размера наименьшего делителя числа n. Если число n выбрано по методу RSA как произведение двух простых чисел примерно одинаковой длины, то метод ЕК имеет ту же оценку, что и метод квадратичного решета, но уступает методу решета числового поля.

Однако если n имеет размерность, превышающую рекордные показателя для методов QS и NFS, (напомним, что последнее рекордное разложение чисел RSA с использованием NFS относится к числу длины 768 бит), то единственная надежда найти делитель n может выполнена только с помощью метода эллиптических кривых.

3.4. Криптографические протоколы на эллиптических кривых

Рассмотрим наиболее известные варианты использования эллиптических кривых в криптографии.

Протокол Диффи-Хелмана

Протокол Диффи-Хельмана используется для генерации двумя удаленными абонентами общего секретного ключа в условиях незащищенного канала связи.

Сначала выбирается простое число p $\approx 2^{160}$ и параметры a и b эллиптической кривой. Этим задается множество точек ЭК $E_p(a, b)$. Затем на $E_p(a, b)$ выбирается генерирующая точка $G = (x_1, y_1)$. При выборе G важно, чтобы наименьшее значение n, при котором $nG = 0$, оказалось

очень большим простым числом. Точка G называется *базовой* точкой. Параметры $E_p(a,b)$ и координаты базовой точки криптосистемы являются открытыми параметрами, известными всем участникам. Обмен ключами между пользователями A и B производится по следующей схеме:

1. Участник A выбирает целое число $n_A < n$. Это число является закрытым ключом участника A. Затем участник A вычисляет открытый ключ $P_A = n_A G$, который представляет собой некоторую точку на $E_p(a,b)$.

2. Точно так же участник B выбирает закрытый ключ n_B и вычисляет открытый ключ – точку на кривой $P_B = n_B G$.

3. Участники обмениваются открытыми ключами, после чего вычисляют общий секретный ключ $K_{A,B}$ по следующей схеме:

   Участник A вычисляет точку эллиптической кривой ЭК $K_{A,B} = n_A \cdot P_B$, являющуюся требуемым общим ключом. Участник B находит ключ по формуле $K_{A,B} = n_B \cdot P_A$. Равенство ключей обеспечивается соотношением $K_{A,B} = n_A \cdot P_B = n_A \cdot (n_B \cdot G) = n_B \cdot (n_A \cdot G)$. Отметим, что поскольку точка на ЭК имеет две координаты, то можно в качестве ключа брать либо только координату x, либо только координату y, либо их сумму $x + y$.

Возможный противник, зная известные параметры n_A, n_B и G, не сможет вычислить значение общего ключа, т.к. для этого ему надо решить задачу дискретного логарифмирования на эллиптической кривой (т.е. найти кратное k по координатам т. kG и G).

Замечание. Для суперсингулярных кривых в 1993 году был найден алгоритм Менезеса, Окатамо и Ванстоуна (так называемая MOV–атака) ([34]), основанный на преобразовании Вейля-Тейта, позволяющий свести задачу дискретного логарифмирования на ЭК (ДЛЭК) к задаче дискретного логарифмирования в конечном поле (ДЛКР), где эта задача может быть решена намного эффективнее.

Поэтому суперсингулярные кривые перестали использоваться в протоколах построения электронной цифровой подписи ЭЦП и шифрования. Однако в 2000 году А. Джоукс ([27]) нашел замечательные применения преобразованию Вейля-Тейта в криптографии, и на сегодняшний день эта тематика является одной из самых популярных а криптографии. Мы рассмотрим алгоритм Менезеса, Окатамо и Ванстоуна в разделе 3.5.

Шифрование сообщений с использованием эллиптических кривых

Рассмотрим самый простой подход к шифрованию/дешифрованию секретных сообщений с использованием эллиптических кривых. Задача состоит в том, чтобы зашифровать сообщение M, которое может быть представлено в виде точки на эллиптической кривой $P_m(x, y)$. Как и в случае обмена ключом, в системе шифрования/дешифрования в качестве параметров рассматривается эллиптическая кривая $E_p(a, b)$ и базовая точка G на ней. Участник B выбирает закрытый ключ n_B, представляющий собой целое число от 2 до n, где n–порядок точки G и вычисляет открытый ключ $P_B = n_B \cdot G$, являющийся точкой на кривой. Участник A выбирает случайное целое положительное число k и вычисляет зашифрованное сообщение C_m, являющееся парой точек на эллиптической кривой:

$$Cm = \{k \cdot G, P_m + k \cdot P_B\}.$$

Чтобы дешифровать сообщение, участник B вычитает из второй точки произведение первой точки на свой закрытый ключ:

$$P_m + k \cdot P_B - nB \cdot (k \cdot G) = P_m + k \cdot (nB \cdot G) - nB \cdot (k \cdot G) = P_m.$$

Участник A зашифровал сообщение P_m добавлением к нему kP_B. Никто не знает значения k, поэтому, хотя P_B и является открытым ключом, никто не знает $k \cdot P_B$. Противнику для восстановления сообщения придется вычислить k, зная G и $k \cdot G$. Сделать это будет нелегко, т.к. надо вычислить дискретный логарифм. Получатель также не знает k, но ему в качестве подсказки посылается $k \cdot G$. Умножив $k \cdot G$ на свой закрытый ключ, получатель получит значение, которое было добавлено отправителем

к незашифрованному сообщению. Тем самым получатель, не зная k, но имея свой закрытый ключ, может восстановить незашифрованное сообщение.

Замечание. Здесь надо еще сказать, как кодировать текстовое сообщение точками кривой. Для этого сообщение разбивается на отдельные символы, и каждый символ заменяется его числовым кодом. Можно использовать какую-нибудь стандартную кодировку типа ASCII, DOS-866 или WIN1251 либо просто перенумеровать все буквы последовательными цифрами.

Далее надо каждому коду (т.е. числу от 0 до 255) сопоставить какую-нибудь точку на кривой. Самый простой вариант – это составить таблицу, сопоставив символу с кодом k координату x точки kG. Другой вариант – воспользоваться тем свойством, для любого числа $x \leq [n/2]$ на эллиптической кривой с высокой вероятностью найдется точка с координатами $U_1(2x, y_1)$ или $U_2(2x+1, y_2)$, которая будет служить образом для кода x. Действительно, зная координаты (x', y') любой из точек U_1 или U_2 можно восстановить x, вычисляя целую часть от $x'/2$.

Построение электронной цифровой подписи с использованием ЭК

Алгоритм ECDSA (Elliptic Curve Digest Signature Algorithm) принят в качестве стандартов ANSI X9F1 и IEEE P1363. Создание ключей:

1. Выбирается эллиптическая кривая $E_p(a, b)$. Число точек на ней должно делиться на большое простое число n.
2. Выбирается базовая точка $G \in E_p(a, b)$ порядка n, $n \cdot G = \infty$.
3. Выбирается случайное число $d \in (1, n)$.
4. Вычисляется $Q = d \cdot G$.
5. Закрытым ключом является d, открытым ключом - кортеж $< a, b, G, n, Q >$.

Создание подписи:

1. Выбирается случайное число $k \in (1, n)$.
2. Вычисляется $k \cdot G = (x_1,\, y_1)$ и $r = x_1 \,(\mathrm{mod}\, n)$.
3. Проверяется условие $r \neq 0$, так как иначе подпись не будет зависеть от закрытого ключа. Если $r = 0$, то выбирается другое случайное число k.
4. Вычисляется $k^{-1} \,(\mathrm{mod}\, n)$.
5. Вычисляется $s = k^{-1} \cdot (\mathrm{H}(M) + dr) \;(\mathrm{mod}\, n)$.
6. Проверяется условие $s \neq 0$, так как в этом случае необходимого для проверки подписи числа $s^{-1} \,(\mathrm{mod}\, n)$ не существует. Если $s = 0$, то выбирается другое случайное число k.

Подписью для сообщения M является пара чисел (r, s).

Проверка подписи:

1. Проверим, что числа r и s принадлежат диапазону чисел $(1, n)$. В противном случае результат проверки отрицательный, и подпись отвергается.
2. Вычислить $w = s^{-1} \,(\mathrm{mod}\, n)$, и $H(M)$,
3. Вычислить $u_1 = H(M)w \,(\mathrm{mod}\, n)$, и $u_2 = rw \,(\mathrm{mod}\, n)$
4. Вычислить $u_1 P + u_2 Q = (x_0, y_0)$, $v = x_0 \,(\mathrm{mod}\, n)$
5. Подпись верна в том и только том случае, когда $v = r$.

Дополнительную информацию об использовании эллиптических кривых в криптографии можно найти в книгах [21] и [37].

3.5. Спаривание Вейля-Тейта

Широкое использование эллиптических кривых в криптографии основано на том свойстве, что задача дискретного логарифмирования на эллиптических кривых (задача ДЛЭК) является более трудоемкой, чем задача дискретного логарифмирования в конечных полях ДЛКП. Это позволяет использовать ключи меньшей длины по сравнению с ключами методов RSA и Эль–Гамаля (160 бит против 1024 бит), что уменьшает требования на вычислительные системы, выполняющие шифрование.

Однако, в 1993 году А. Менезес, Т. Окамото и С. Вэнстоун [34] показали, что задача ДЛЭК сводится к задаче ДЛКП в некотором конечном расширении исходного поля $GF(q)$ эллиптической кривой. Идея сведения основана на спаривании Вейля (Weil's Pairing) по имени выдающегося французского математика Андре Вейля (1906–1998), известного своими трудами в области алгебраической геометрии.

Пусть задано уравнение ЭК $E:\ y^2 = x^3+ax+b$ над полем F_q, $q = p^m$. Напомним, что алгебраическим замыканием поля K называется множество корней уравнений с коэффициентами из этого поля (обозначается $\overline{K}$).

Пусть n–целое положительное число, взаимно-простое с p. Определим множество $E[n]$ как множество точек кривой E порядка n над алгебраическим замыканием $\overline{F_q}$ исходного поля F_q , т.е. множество точек $P(x,y) \in EC(\overline{F_q})$, удовлетворяющих $nP = \infty$.

Хотя исходное поле F_q конечно, его замыкание бесконечно. Однако, множество $E[n]$ содержит конечное число элементов (можно доказать, что оно изоморфно группе $\mathbf{Z}_n \oplus \mathbf{Z}_n$) и, значит, содержится в некотором конечном расширении F_{q^k} исходного поля. Степень k называется *степенью вложения*. Эта степень может быть определена как наименьшее положительное число со свойством $n \mid (q^k - 1)$. Определим μ_n как множество корней n-й степени из 1, содержащихся в F_{q^k}.

Отображение Вейля представляет собой билинейное отображение

$$e:\ E[n] \oplus E[n] \to \mu_n, \tag{3.68}$$

обладающее следующими свойствами:

- (билинейность) $e(A+B,C)=e(A,C)+e(B,C)$,
  $e(A,B+C)=e(A,B)+e(A,C)$,
- $e(P,P)=1$ для любого $P\in E[n]$,
- (невырожденность) $(\exists P,Q\in E[n])\ e(P,Q)\neq 1$,
- (вычислимость) $e(X,Y)$ может быть эффективно вычислено.

Если степень вложения k принимает небольшие значения (до $k=6$), то для поиска ключа шифрования вместо решения задачи ДЛЭК можно решать более легкую задачу вычисления дискретного логарифма в конечном поле размерности q^k. Таким образом, эллиптические кривые, допускающие вложение в конечные поля с небольшой степенью k, не могут быть использованы в криптографии. Таковыми, например, являются все суперсингулярные кривые, имеющие степень вложения $k\in\{1,2,3,4,5,6\}$.

Кривая $E=EC(GF_{p^r})$ называется *суперсингулярной*, если ее мощность $\#E=p^r+1-t$, и $p\,|\,t$.

Примером суперсингулярной кривой может служить кривая $E:\ y^2=x^3+1\ (\mathrm{mod}\ p)$, если характеристика поля $p\equiv 2\ (\mathrm{mod}3)$, тогда E содержит $p+1$ элемент, $t=0$ и E имеет степень вложения, равную 2.

Способ вычисления дискретного логарифма на ЭК, использующий сведение Вейля, получил называние MOV-атаки (MOV-attack) по заглавным буквам фамилий изобретателей

Многие протоколы, использующие шифрование и электронные цифровые подписи на эллиптических кривых, специально запрещают использование суперсингулярных кривых. Таким образом, суперсингулярные кривые были изъяты из криптографии.

Однако в 2002 году А.Джоукс [27] нашел неожиданное применение спариванию Вейля и суперсингулярным кривым для построения однораундового протокола выработки общего секретного ключа на основе метода Диффи-Хелмана. Далее были найдены и другие, не менее интересные приложения такие, как, например, построение открытого ключа пользователя на основе его общеизвестных идентификационных данных

таких, как, например, имя или адрес электронной почты (identity based open keys) (см. Advances in Elliptic Curve [4]).

Решение проблемы дискретного логарифмирования с помощью MOV–алгоритма

Описание этого алгоритма можно найти в главе 5 книги Л. Вашингтона [54].

Пусть заданы эллиптическая кривая $EC:\ y^2 = x^3 + ax + b\ (\mathrm{mod} p^r)$, и точки $P,\ Q \in EC$ порядка n, где n–простое число, причем существует m такое, что $Q = mP$. Требуется найти множитель m. Отображение Вейля будем обозначать через $e(X, Y)$. Алгоритм вычисления m заключается в следующем:

1. Находим случайную точку $T \in EC(F_{q^k})$.

2. Находим порядок M точки T.

3. Находим $d =$Н.О.Д.(n, M). Если $d = 1$, то возвращаемся к п.1. Иначе, перейдем к следующему пункту. Определим, что в этом случае т.T имеет порядок n.

4. Вычислим $a = e(P, T)$ и с $= e(Q, T)$.

5. Вычисляя дискретный логарифм в поле F_{q^k}, найдем искомый множитель m.

Отметим, что можно выполнять этот алгоритм с составным n, тогда число d может оказаться собственным делителем n и найденный множитель окажется равным $m \bmod d$. В этом случае можно повторять вычисление с различными точками T_i, вычисляя $m_i = m \bmod d_i$ до тех пор, пока произведение различных d_i не станет больше или равно n. После этого можно найти m с помощью китайской теоремы об остатках.

Замечание. Если речь идет о произвольной точке Q, то прежде, чем вычислять дискретный логарифм, полезно знать, найдется ли такое m, что $Q = mP$. Эту проверку можно выполнить, используя следующее утверждение:

Теорема 3.2. *Для произвольной т.*$Q \in EC(F_{q^k})$ *найдется число* m *такое,*

что $Q = mP$ *в том и только в том случае, если выполняются два условия:*

1. $nQ = \infty$,
2. $e(P, Q) = 1$.

3.6. Дивизоры

Построение отображения Вейля и родственного ему отображения Тейта основано на теории дивизоров (делителей) алгебраических кривых, разработанной Андре Вейлем. Приведем здесь основные сведения из этой теории. Более подробный материал можно найти в книге Л. Вашингтона [54].

Идея понятия дивизора основано на том наблюдении, что коэффициенты любого полинома можно вычислить с точностью до ненулевого множителя, зная корни этого многочлена и их кратность. Действительно, если многочлен $P(x)$ имеет своими корнями кратности r_i элементы x_i, то

$$P(x) = a \cdot \prod (x - x_i)^{r_i}.$$

В нашем случае класс изучаемых функций состоит из дробно-рациональных функций над эллиптическими кривыми, т.е. отношений двух многочленов от двух переменных x и y, определенных на точках некоторой эллиптической кривой.

Пусть теперь $E:\ y^2 = x^3 + ax + b$–эллиптическая кривая над полем K, а $f(x, y):\ E \to K$–дробно-рациональная функция. Если f – не константа, то существует не более конечного числа точек $P \in E$, в которых $f(P) = 0$ или $f(P) = \infty$. Точки первого вида называются *нулями функции* f, а второго – *полюсами* f.

С точностью до ненулевого множителя функцию f можно задать, перечисляя все ее нули и полюсы и задавая их кратность. Если f имеет нуль (полюс) кратности k в точке P, то f можно представить в виде произведения $f = u_P^k \cdot g$, где u_P имеет в точке P нуль (полюс) первого порядка, а $g(P) \neq 0, \neq \infty$. Функция u_P называется *униформизатором* функции f в точке P

Пример. Рассмотрим кривую $y^2 = x^3 - x$ и функцию $f(x, y) = x/y$. Перепишем f в виде

$$f(x,y) = \frac{x}{y} = \frac{xy}{y^2} = \frac{xy}{x^3 - x} = \frac{y}{x^2 - 1} = y \cdot \frac{1}{x^2 - 1}.$$

Из последнего представления видим, что точка $P(0,0)$ является нулем 1-о порядка функции $f(x,y) = x/y$, а функция $u(x,y) = y$ ее униформизатором в точке $P(0,0)$.

Пусть M_1–множество нулей, а M_2– множество полюсов функции $f(x,y)$. Сопоставим функции f формальное выражение

$$f(x,y) \sim \sum_{P \in M_1} r_P [P] - \sum_{P \in M_2} r_P [P], \tag{3.69}$$

где r_P– кратность нуля (полюса) P.

Определение 3.1. *Пусть $E: y^2 = x^3 + ax + b$ – эллиптическая кривая над полем k. Дивизором D над кривой E называется формальная сумма вида*

$$D = \sum_{P \in E} r_P [P],$$

в которой коэффициенты r_P – целые числа (положительные или отрицательные) и число слагаемых с ненулевым коэффициентом r_P – конечно.

Множество точек P, для которых $r_P \neq 0$, называется носителем (support) *дивизора D и обозначается $supp(D)$. Целое число $k = \sum r_P$, $P \in supp(D)$, называется степенью D и обозначатся $deg(D)$. Точка эллиптической кривой, равная $\sum_{P \in E} r_P \cdot P$, называется* суммой *дивизора D и обозначается $sum(D)$.*

Сумма дивизоров определяется естественным образом. Множество дивизоров эллиптической кривой образует аддитивную группу относительно операции сложения, а нулем является дивизор, у которого все коэффициенты равны 0. В группе дивизоров наиболее важную роль играют дивизоры функций, которые называются главными дивизорами (principal divisors).

Вычислим дивизор прямой $l : ax + by + c$, проходящей через две заданные точки $P_1(x_1, y_1)$ и $P_2(x_2, y_2)$ эллиптической кривой E. Если l не является касательной в т.P_1 и P_2, то она пересекает E и в третьей т.$P_3(x_3, y_3)$, а также в бесконечно удаленной точке ∞. В точках P_1, P_2 и P_3 прямая l имеет нули 1 порядка, а в т. ∞ – полюс 3 порядка. Чтобы увидеть это, перепишем уравнение ЭК $y^2 = x^3 + Ax + B$ в следующем виде:

$$\left(\frac{x}{y}\right)^2 = x^{-1}\left(1 + \frac{A}{x^2} + \frac{B}{x^3}\right)^{-1}, \tag{3.70}$$

откуда

$$x^{-1} = \left(\frac{x}{y}\right)^2 \cdot \left(1 + \frac{A}{x^2} + \frac{B}{x^3}\right). \tag{3.71}$$

Из уравнения (3.70) следует, что x/y обращается в 0 в т.∞, а уравнение (3.71) показывает, что функция x/y является униформизатором x^{-1} в т.∞ и т.∞ является нулем второго порядка для x^{-1}. Значит т.∞ является полюсом 2 порядка для x. Так как $y = x \cdot (y/x)$, то т.∞ является полюсом 3 порядка для y и для функции $l = Ax + By + C$. Отсюда дивизор прямой l имеет вид

$$div(l_{P_1,P_2}) = 1[P_1] + 1[P_2] + 1[P_3] - 3[\infty]. \tag{3.72}$$

Проведем через т.P_3 вертикальную прямую $v = x - x_3$. Она проходит через т.$P_3(x_3, y_3)$, $-P_3(x_3, -y_3)$ и т.∞, а ее дивизор имеет вид

$$div(v_{P_3}) = 1[P_3] + 1[-P_3] - 2[\infty]. \tag{3.73}$$

Из формул (3.72) и (3.73) получим

$$div\left(\frac{Ax + By + C}{x - x_3}\right) = div(Ax+By+C) - div(x-x_3) = [P_1]+[P_2]-[-P_3]-[\infty].$$

Так как $P_1 + P_2 = -P_3$ на кривой E, то последнюю формулу можно переписать в виде

$$[P_1] + [P_2] = [P_1 + P_2] + [\infty] + div\left(\frac{Ax + By + C}{x - x_3}\right). \tag{3.74}$$

Из формул (3.72) и (3.73) можно видеть, что согласно определению 3.1 степени прямых l_{P_1,P_2} и v_{P_3} равны 0, а их сумма равна ∞, что является примером общего факта, выражаемого следующей теоремой:

Теорема 3.3. *Дивизор D эллиптической кривой E, имеющий степень 0, является дивизором некоторой функции тогда и только тогда, когда $sum(D) = \infty$.*

Пример нахождения функции по заданному дивизору

Формула (3.74) дает способ нахождения функции f для заданного дивизора D, удовлетворяющего теореме 3.3. Вычислим функцию f на ЭК $E:\ y^2 = x^3 + 4x (\mathrm{mod}\ 11)$, дивизор которой имеет вид

$$D = [(0,0)] + [(2,4)] + [(4,5)] + [(6,3)] - 4[\infty].$$

Прямая l, проходящая через т.$(0,0)$ и $(2,4)$ имеет вид $l = y - 2x$, причем т.$(2,4)$ является нулем 2 порядка, откуда

$$div(y - 2x) = [(0,0)] + 2[(2,4)] - 3[\infty].$$

Вертикальная прямая через т.$(2,4)$ имеет вид $v = x - 2$ и

$$div(x - 2) = [(2,4] + [(2,-4)] - 2[\infty].$$

Значит,

$$[(0,0)] + [(2,4)] = [(2,-4)] + [\infty] + div\left(\frac{y - 2x}{x - 2}\right).$$

Аналогично,

$$[(4,5)] + [(6,3)] = [(2,4)] + [\infty] + div\left(\frac{y + x + 2}{x - 2}\right),$$

откуда

$$D = [(2,-4)] + div\left(\frac{y - 2x}{x - 2}\right) + [(2,4)] + div\left(\frac{y + x + 2}{x - 2}\right) - 2[\infty].$$

Поскольку $(2,-4)] + (2,4)] = div(x-2) + 2[\infty]$, то получим

$$D = div(x-2) + div\left(\frac{y-2x}{x-2}\right) + div\left(\frac{y+x+2}{x-2}\right) =$$

$$= div\left(\frac{(y-2x)(y+x+2)}{x-2}\right).$$

Если раскрыть скобки в числителе и заменить слагаемое y^2 на x^3+4x, то вынося $x-2$ за скобки, получим $(y-2x)(y+x+2) = (x-2)(x^2-y)$, откуда

$$D = div(x^2-y).$$

Функции от дивизоров

Отображение, задаваемое формулой (3.76), является групповым гомоморфизмом из аддитивной группы дивизоров в мультипликативную группу поля K, т.к.

$$f(D_1+D_2) = f(D_1)\cdot f(D_2), \quad f(D_1-D_2) = \frac{f(D_1)}{f(D_2)}. \tag{3.75}$$

Распространяя формулы (3.75) на произвольныедивизоры, получим формулу

$$f(\sum kP) = \prod f(P)^k. \tag{3.76}$$

Следующая теорема носит название закона взаимности Вейля (Weil reciprocity).

Теорема 3.4. *Если f и g – функции на эллиптической кривой такие, что $div(f)$ и $div(g)$ не имеют общих точек, тогда выполняется следующая формула:*

$$f(div(g)) = g(div(f)).$$

Определение отображений Вейля и Тейта

Дадим в этом разделе точные определения отображений Вейля и Тейта (Weil' and Tate' Pairings). Пусть $E:\ y^2 = x^3+ax+b$ – эллиптическая кривая

над алгебраически замкнутым полем K, n–положительное целое число и $E[n]$ – подгруппа точек кривой E порядка n:

$$E[n] = \{P \in E \mid n \cdot P = \infty\}.$$

Эта подгруппа изоморфна аддитивной группе $\mathbf{Z}_n \times \mathbf{Z}_n$.

Пусть т.$T \in E[n]$. Рассмотрим дивизор $D = n[T] - n[\infty]$. Его степень равна 0, а сумма ∞. По теореме 3.3 найдется функция f, дивизор которой равен D:

$$div(f_T) = n[T] - n[\infty]. \tag{3.77}$$

Будем называть функцию f_T, удовлетворяющую (3.77), функцией Вейля. Пусть т.$P \in E[n]$ не принадлежит орбите т.T, т.е. не совпадает ни с каким кратным kT, $k \leq n$, точки T. Рассмотрим дивизоры

$$D_S = [S] - [\infty], \quad D_T = [T + R] - [R], \tag{3.78}$$

где R–произвольно выбранная точка $E[n]$.

Определение 3.2. *Отображение (спаривание) Вейля – это билинейное отображение*

$$e_n : \ E[n] \times E[n] \to \mu_n, \tag{3.79}$$

где μ_n–подгруппа по умножению корней n–й степени из 1 поля K, задаваемое следующей формулой:

$$e_n(T, S) = \frac{f_T(D_S)}{f_S(D_T)} = \frac{f_T([S] - [\infty])}{f_S([T + R] - [R])}. \tag{3.80}$$

Используя формулы (3.75), можно переписать (3.80) в виде

$$e_n(T, S) = \frac{f_T(R) f_T(S)}{f_S(T + R) f_T(\infty)}. \tag{3.81}$$

Можно доказать, что преобразование Вейля не зависит от выбора т.R, поэтому в определении (3.78) в качестве R можно взять любую точку из $E[n]$. В книге Л.Вашингтона ([54]) отображение Вейля задается обратным

отображением по отношению к формуле (3.78), полученным при перестановке местами аргументов T и S. Это не влияет на свойства этого преобразования. Рассмотрим пример вычисления отображения Вейля.

Пример. Пусть E–эллиптическая кривая над полем F_7, заданная уравнением

$$y^2 = x^3 + 2.$$

Имеем, $E(F_7)[3] \simeq \mathbf{Z}_3 \oplus \mathbf{Z}_3$. Вычислим $e_3((5,1),\,(0,3))$.

Определим $S = (0,3)$, $T = (5,1)$ и $R = (6,1)$. Тогда $D_S = [(0,3)] - [\infty]$, $D_T = [(3,6)] - [(6,1)] = [(5,1) + (6,1)] - [(6,1)]$. Также как и в предыдущем разделе, найдем функцию Вейля (3.77) для точек S и T:

$$f_{(0,3)} = y - 3, \quad f_{(5,1)} = \frac{4x - y + 1}{5x - y - 1}.$$

Далее,

$$f_{(0,3)}(D_T) = \frac{f_{(0,3)}(3,6)}{f_{(0,3)}(6,1)} = \frac{6-3}{1-3} \equiv 2 \pmod 7.$$

Аналогично,

$$f_{(5,1)}(D_S) = 4.$$

Отсюда

$$e_3((5,1),\,(0,3)) = \frac{2}{4} \equiv 4 \pmod 7.$$

Отметим, что 4 является кубическим корнем из 1, т.к. $4^3 = 64 \equiv 1 \pmod 7$.

Определим далее отображение Тейта. Первым аргументом преобразования Тейта по-прежнему является произвольная т.$T \in E[n]$. Обозначим через nE множество точек $\{nQ \mid Q \in E\}$, а через E/nE множество классов эквивалентности кривой E по множеству nE.

Определение 3.3. *Отображение (спаривание) Тейта – это билинейное отображение*

$$\tau_n : \; E[n] \times E/E[n] \to \mathbf{F}^*_{q^k} \times \mathbf{F}^*_{q^k} \setminus \mu_n, \tag{3.82}$$

где μ_n *–подгруппа по умножению корней* n *–й степени из 1 поля* F_{q^k}, *задаваемое следующей формулой:*

$$\tau_n(T,S) = \frac{f_T(S+R)}{f_T(R)}, \tag{3.83}$$

где $R \notin \{T,\ -S,\ T-S,\ \infty\}$.

Одним из важных отличий отображение Тейта является то, что оно не вырождено (не равно 1) при $P = Q$. Это позволяет вычислить множитель m такой, что $Q = mP$ за одно вычисление. Действительно,

$$\tau(P,Q) = \tau(P,mP) = \tau(P,P)^m = b \ \ (\bmod\ q).$$

Чтобы найти теперь m, достаточно вычислить дискретный логарифм $\log_a b\,(\mathrm{mod} q)$, где $a = \tau(P,P)$, в поле $K = F_q$.

Отметим, что значение преобразования Тейта $\tau(P,\,Q)$ определяется точками P и Q не однозначно, а с точностью до множителя из группы μ_n. Чтобы получить уникальное значение, элемент $\tau(P,\,Q)$ возводят в степень $(q^k-1)/n)$. Обозначим эту функцию через τ_{un}:

$$\tau_{un}(P,Q) = \tau(P,Q)^{(q^k-1)/n}. \tag{3.84}$$

Алгоритм Миллера

Главной проблемой в вычислении преобразований Вейля и Тейта является нахождение функции f, дивизор которой совпадает с заданным дивизором D. Пусть т.$T \in E[n]$. В этом разделе будем обозначать функцию Вейля (3.77) с дивизором $n[T]-n[\infty]$ через $f_{n,T}$, подчеркивая ее зависимость от порядка n т.T. Определим вспомогательные дивизоры

$$D_j = j[S+R] - j[R] - [jS] + [\infty],$$

которые удовлетворяют условиям теоремы (3.3). Обозначим через $f_{j,T}$ функцию, дивизор которой равен D_j. Эти функции называются функциями Миллера.

Функцию Вейля $f_{n,P}(Q)$ можно вычислить с помощью рекурсивного алгоритма Миллера, основанного на вычислении промежуточных функций Миллера $f_{j,P}(Q)$ для $j < n$ по следующей формуле:

$$f_{1,T}(Q) = 1 \text{ для любой т.} Q \in E(K),$$

$$f_{i+j,T}(Q) = f_{i,T}(Q) \cdot f_{j,T}(Q) \cdot \left. \frac{l_{i,j}}{v_{i+j}} \right|_Q, \tag{3.85}$$

где $l_{i,j} = Ax + By + C$– уравнение прямой, проходящей через т.iT и jT, $v_{i+j} = x - x_0$– уравнение вертикальной прямой, проходящей через т. $R = (i + j)T$.

Приведем формулы для вычисления коэффициентов A, B и C прямой $l_{P,Q}$, проходящей через т.$P(x_1, y_1)$ и $Q(x_2, y_2)$:

1. $P = Q$. Угловой коэффициент λ наклона касательной равен

$$\lambda = (3x_1^2 + a)/(2y_1) \,(\bmod p). \tag{3.86}$$

2. $P \neq Q$. Угловой коэффициент λ равен в этом случае

$$\lambda = (y_2 - y_1)/(x_2 - x_1) \,(\bmod p). \tag{3.87}$$

В обоих случаях уравнение прямой, проходящей через точку $P(x_1, y_1)$ и имеющей коэффициент наклона λ, имеет вид $y - y_1 = \lambda \cdot (x - x_1)$, откуда получим уравнение l:

$$l = y - \lambda x + (\lambda x_1 - y_1). \tag{3.88}$$

Последними выпишем формулы для вычисления координат суммы точек $P + Q = (x_3, y_3)$ (формулы для удвоенной точки можно получить, приравнивая $x_2 = x_1$):

$$\begin{cases} x_3 = \lambda^2 - x_1 - x_2 \\ y_3 = -y_1 + \lambda(x_1 - x_3). \end{cases} \tag{3.89}$$

Алгоритм Миллера вычисления функции Вейля $f_{P,n}$

1. Найдем бинарное представление числа $n = (n_t \dots n_0)_2$.

2. Определим исходные значения переменной точки Z и функции f равными P и 1 соответственно.

3. Выполняем цикл по i от $i = t - 1$ до $i = 0$:

- Установим

$$f = f^2 \cdot l_{Z,Z}/v_{2Z},$$
$$Z = 2Z.$$

- Если $n_i = 1$, тогда выполним операцию сложения $P + Z$:

$$f = f^2 \cdot l_{P,Z}/v_{P+Z},$$
$$Z = P + Z.$$

4. Определим выходное значение функции Вейля $f_{P,n} = f$.

Пример 1. Дана кривая $y^2 = x^3 + 11$ над полем F_{31}. Она содержит 25 точек и изоморфна группе $\mathbf{Z}_5 \times \mathbf{Z}_5$. Эта группа порождается точками $P = (2; 9)$ и $Q = (3; 10)$, имеющими порядок $n = 5$. Степень вложения $k = 1$, т.к. $p^1 - 1 = 30$ делится на $n = 5$. Вычислим функцию Вейля $f_{5,P}$, используя алгоритм Миллера:

1. Найдем двоичное представление $n = 5 = (101)_2$, $t = 2$.

2. Положим $Z = (2; 9)$. Выполним вычисления шага 3 алгоритма Миллера при $i = t - 1 = 1$.

$$\lambda = 3 \cdot 2^2/(2 \cdot 9) \bmod 31 = 2/3 \bmod 31 = 2 \cdot 21 \bmod 31 = 11.$$
$$l = y - \lambda x + (\lambda x_1 - y_1) = y - 11x + 11 \cdot 2 - 9 = y - 11x + 13.$$
$$Z = 2Z = (\lambda^2 - 2x_1;\ -y_1 - \lambda(x_2 - x_1)) = (24;\ 28)$$
$$v = x - 24 \equiv x + 7.$$
$$f_{2,P} = (-11x + y + 13)/(x + 7).$$

Проверим условие $n_i = 1$. Т.к. $n_i = 0$, то операция сложения на i–м шаге не выполняется. Переходим к следующей итерации при $i = 0$.

3. $Z = (24, 28)$. Выполним операцию удвоения т. Z: $\lambda = 22$, $l_{2,2} = 9x + y + 4$, $2Z = (2;\ 22)$, $v_4 = x - 2$. Отсюда

$$f_{4,P} = f_{2,P}^2 \cdot \frac{9x + y + 4}{x - 2} = \frac{(-11x + y + 13)^2(9x + y + 4)}{(x + 7)^2(x - 2)}$$

Т.к. $n_i = 1$, то выполняем вторую часть шага 3 алгоритма Миллера. Это вычисление приводит к т. $5P = \infty$ и $l = x - 2$, $v = 1$, откуда

$$f_{5,P} = f_{4,P} \cdot (x - 2) = \frac{(-11x + y + 13)^2(9x + y + 4)}{(x + 7)^2}$$

Вычислим значение преобразования Тейта $\tau(P, Q)$, взяв $Q = (3;\ 10)$. Для этого потребуется вспомогательная точка R. Возьмем, например, $R = Q$. Вычислим сумму $S = 2Q = (-1;\ 14)$. Вычислим функцию Вейля в точках S и R:

$$f(S) = f(-1;\ 14) = 20, \quad f(Q) = f(3;\ 10) = 10. \quad \tau(P, Q) = 20/10(\ \mathrm{mod}\ 31) = 2.$$

Снова вычислим значение преобразования Тейта $\tau(P, Q)$, взяв $R = 2Q$. Возьмем, например, $R = 2Q = (-1;\ 14)$. Вычислим сумму $S = 2Q + Q = (-1, 17)$. Тогда

$$f(S) = f(-1;\ 17) = 23, \quad f(2Q) = 20. \quad \tau(P, Q) = 23/20 = 12(\ \mathrm{mod}\ 31) = 2.$$

Мы видим, что значение преобразования Тейта зависит от выбора точки R. Для получения уникального значения необходимо полученное значение возвести в степень $(q^k - 1)/n$. В нашем примере оно равно $(31 - 1)/5 = 6$. Имеем,

$$2^6\ (\mathrm{mod}\ 31) = 2, \quad 12^6\ (\mathrm{mod}\ 31) = 2.$$

Пример 2. Даны две точки $P = (2;\ 9)$ и $xP = (24;\ 3)$. Найти кратное x на эллиптической кривой $y^2 = x^3 + 11(\ \mathrm{mod}\ 31)$ из предыдущего примера.

Решение. Определим функцию $f_{P,5}$ так же, как в предыдущем примере. Вычислим $\tau(P, P)$, взяв $R = (15,\ 10)$:

$$S = P + R = (2;\ 9) + (15,\ 10) = (3,\ 10), \quad f_{P,5}(S) = 30, \quad f_{P,5}(R) = 7,$$

$\tau(P,P) = 30/7 \equiv 22(\mod 31),\ a = \tau_{un}(P,P) = 22^6 \equiv 8(\mod 31).$

Вычислим $\tau(P, xP)$, взяв по-прежнему $R = (15,\ 10)$:

$S' = xP + R = (24;\ 3) + (15,\ 10) = (6,\ 14),\quad f_{P,5}(S') = 29,$

$\tau(P, xP) = 29/7 \equiv 13(\mod 31),\ b = \tau_{un}(P, xP) = 13^6 \equiv 16\,(\mod 31).$

Теперь для нахождения x надо вычислить $x = \log_a b(\mathrm{mod}p) = \log_8 16\,(\,\mathrm{mod}\,31)$. Найдем x простым перебором:

$8^2 \mod 31 = 2,\quad 8^3 \mod 31 = 16,$ значит $x = 3.$

4. Метод квадратичного решета

Метод квадратичного решета занимает вторую строчку в списке самых быстрых алгоритмов факторизации, уступая только методу решета числового поля. Рассмотрим основные идеи этого метода.

4.1. Идея Мориса Крейтчика и алгоритм Диксона

В 20-х г. XX столетия Морис Крейтчик (Maurice Kraitchik), обобщая идею Ферма, предложил вместо пар чисел, удовлетворяющих уравнение $A^2 - B^2 = n$ (2.20), искать пары чисел, удовлетворяющих более общему уравнению

$$A^2 \equiv B^2 (\mathrm{mod}\ n) \tag{4.90}$$

Крейтчик заметил, что многие из чисел $q(x)$ в (2.21) раскладываются в произведение небольших простых чисел (т.е. являются гладкими по современной терминологии). Рассмотрим пример из статьи Померанца [46]:

Пример. Пусть число $n = 2041$, которое требуется разложить. Ближайшим к n числом, являющимся полным квадратом, является $m = 45$, $m^2 = 2025$. Рассмотрим последовательность пар чисел $\{x;\, q(x)\}$, где

$$q(x) = (m + x)^2 - n, \tag{4.91}$$

а x принимает близкие к 0 значения:
$\{(-2;\, -192),\, (-1;\, -105), (0;\, -16), (1;\, 75),\, (2;\, 168),\, (3;\, 263),\, (4;\, 360),\, (5;\, 459),\, (6;\, 560)\}$.

Можно заметить, что вторые координаты некоторых из этих пар разложимы в произведение небольших простых чисел:
$-192 = -2^6 \times 3,\ -105 = -3 \times 5 \times 7,\ -16 = -2^4, 75 = 3 \times 5^2,\ 168 = 2^3 \times 3 \times 7, 360 = 2^3 \times 3^2 \times 5,\ 560 = 2^5 \times 5 \times 7$.

Выберем множество небольших простых чисел и назовем его факторной базой $FB = \{2, 3, 5, 7\}$. Любое целое число, разложимое в произведение множителей из факторной базы, назовем *гладким* относительно этой факторной базы. Таким образом, перечисленные выше числа являются гладкими. Когда факторная база выбрана, любое гладкое число можно

представить вектором показателей степеней $v = (r_1, r_2, \ldots r_k)$, k – размер факторной базы. Например, числу 560 соответствует вектор $v = (5, 0, 1, 1)$.

Далее Крейтчик заметил, что можно перемножить некоторые гладкие числа так, чтобы получить полные квадраты:

$$75 \times 168 \times 360 \times 560 = 50400^2, \quad (-192) \times (-16) \times 75 = 480^2.$$

Определение 4.1. *Пара целых чисел (A, B) называется гладкой (относительно факторной базы F), если:*

1. Выполняется соотношение $A^2 \equiv B \,(\mathrm{mod} n)$,

2. B – раскладывается в произведение элементов из F.

Согласно этому определению, множество $M = \{(-2, -192), (-1, -105), (0, -16), (2, 168), (4, 360), (5, 560)\}$, построенное в примере, состоит из гладких пар.

Как использовать гладкие пары? Заметим, что при умножении чисел, их вектора показателей степеней складываются. Чтобы произведение гладких чисел было полным квадратом, необходимо подобрать такую комбинацию сомножителей, чтобы сумма векторов показателей имела только четные координаты. Например, произведению $75 \times 168 \times 360 \times 560$ соответствует сумма векторов $(0, 1, 2, 0) + (3, 1, 0, 1) + (3, 2, 1, 0) + (5, 0, 1, 1) = (8, 4, 4, 2)$.

Вместо целочисленных координат векторов степеней и их сумм достаточно рассматривать их остатки по модулю 2, т.е. выполнять все вычисления в поле $F_2 = \{0, 1\}$, тогда произведение элементов M является полным квадратом тогда и только тогда, когда вектор суммы по модулю 2 всех векторов–показателей является нулевым вектором. Множество всевозможных векторов размерности k над полем $F_2 = \{0, 1\}$ образует линейное пространство L_k размерности k, поэтому любое множество векторов, содержащее больше k элементов, является линейно зависимым, и существует нетривиальная линейная комбинация таких векторов, равная нулевому вектору. Коэффициенты этой линейной комбинации можно найти, решая систему линейных уравнений, составленную из коэффициентов степеней взятых по модулю 2.

Поскольку коэффициенты этой линейной комбинации равны либо 0, либо 1, то она представляет собой просто сумму некоторого подмножества векторов. Отсюда следует тот факт, что для нахождения нетривиального подмножества гладких чисел, произведение которых есть полный квадрат, достаточно иметь $k+1$ гладкое число, где k – размер факторной базы.

Когда набор гладких пар, содержащий не менее $k+1$ элемента, найден, то решение уравнения 4.90 можно найти, решая соответствующую систему линейных уравнений, например, с помощью метода исключения неизвестных Гаусса.

В 1981 г. Д. Диксон [22] предложил метод факторизации, использующий идеи Крейтчика. Однако, хотя этот метод обеспечивал значительное преимущество по отношению к предшествующим методам, он работал медленно, т.к. пробное деление на элементы факторной базы занимало слишком много времени.

4.2. Метод Померанца

В 1982 г. Карл Померанц предложил новую идею поиска гладких чисел, используя метод, похожий на решето Эратосфена. Идея Померанца заключалась в том, что если для простого числа $p \in FB$ найден аргумент x такой, что $q(x) \equiv 0 \,(\bmod\ p)$, то на p будут делиться все элементы $q(y)$, где y отличающиеся от x на аргумент, кратный p, т.е. $y = x + k \cdot p$, $k \in \mathbf{Z}$. Поэтому, если для данного p найден корень x уравнения $q(x) \equiv 0 \,(\bmod\ p)$, то для всех y, $y \equiv x \,(\mathrm{mod}\, p)$ будем также выполнено условие $q(y) \equiv 0 \,(\mathrm{mod}\, p)$.

Алгоритм Померанца работает следующим образом:

1. Выбирается некоторый числовой интервал $[-L;\, L]$, называемый интервалом просеивания,

2. В массив целых чисел $W[-L\,..\,L]$ заносятся значения полинома $q(x)$ для $x \in [-L; L]$,

3. Для каждого числа p из факторной базы FB ищутся числа $0 \le x < p$

такие, что выполняется сравнение

$$q(x) \equiv 0 \bmod p \tag{4.92}$$

Отметим, что это уравнение может иметь либо 2 решения, либо ни одного.

4. Для каждого решения x уравнения (4.92) в цикле просматриваются числа вида $x_k = x + kp$ из интервала $[-L;\, L]$, $k \in \mathbf{Z}$, и выполняется деление элементов $W[x_k]$ на p.

Процедура повторяется для всех чисел p из факторной базы и их степеней $p^k < B$, ограниченных сверху границей B. В результате выполнения процедуры просеивания некоторые элементы $W[x]$ массива W станут равными ± 1. Соответствующие пары $(x, q(x))$ являются гладкими. Для надежного нахождения нетривиального решения уравнения (4.90) необходимо, что число найденных гладких пар превышало, по-крайней мере, на 1 размер факторной базы. Тогда из коэффициентов разложения гладких чисел на элементы факторной базы можно составить систему недоопределенную систему линейных уравнений, которая будет иметь нетривиальное решение.

Следует отметить, что наличие нетривиального решения гарантирует нам только нахождение пары натуральных чисел $(A;\, B)$, удовлетворяющих (4.90). Однако не все пары такого рода дают искомое решение, поэтому для надежности получения нетривиальных делителей n число гладких чисел должно быть больше размерности факторной базы на число $k > 1$, например, на $k = 3$ или $k = 4$.

Дополнительным достоинством метода Померанца явилась возможность распараллеливания вычисления и запуска алгоритма для одновременного просеивания на нескольких компьютерах. Метод Померанца получил название метода квадратичного решета (the Quadratic Sieve). Используя этот метод, в 1994 г. Аткинс, Граф, Лейланд и Ленстра сумели разложить 129-значное число, предложенное создателями RSA.

Процедура просеивания является самой затратной по времени частью алгоритма квадратичного решета. В ходе ее выполнения ищутся пары чисел (A, B), удовлетворяющих

$$A^2 \equiv B \bmod n. \tag{4.93}$$

Размер факторной базы является одним из ключевых параметров алгоритма просеивания, определяющих его эффективность. Если ее размер является слишком маленьким, то гладкие числа будут попадаться очень редко или не будут попадаться вообще. При недостаточном размере факторной базы приходится выбирать большой радиус L интервала просеивания, что влечет увеличение общего времени вычисления.

Если же размер факторной базы слишком велик, то требуется поиск большого числа гладких чисел, что также увеличивает общее время вычисления. Отметим, что для разложения 129-значного числа создателей RSA потребовалась факторная база, содержащая 524338 простых чисел.

Обозначим размер факторной базы через k. Для разложения числа n, необходимо отыскать k' гладких пар $(A; B)$, $k' \geq k + 2$. После этого формируется система линейных уравнений размерности $k \times k'$ с коэффициентами из поля $F_2 = \{0, 1\}$ и нулевым столбцом свободных членов. Система – не доопределена и потому имеет нетривиальное решение, представляющее множество M первых аргументов гладких пар $(x, q(x))$. Тогда пара (A, B), удовлетворяющая уравнению (4.90), находится как

$$A = \prod_{x \in M} (x + m) \bmod n, \quad B = \prod_{x \in M} q(x) \bmod n. \tag{4.94}$$

Теперь делители p исходного n можно найти, вычисляя Н.О.Д.$(n, A \pm B)$. В некоторых случаях оба найденных делителя оказываются тривиальными (равными 1 или n), тогда нужно искать другое решение системы и новую пару (A, B). Опишем процедуру построения факторной базы более подробно.

4.3. Построение факторной базы

Общая последовательность действий в процедуре подготовки факторной базы такова:

I. Инициализация:

1. Составляем множество из всех простых чисел, меньших некоторой верхней границы B: $FB = \{2, 3, 5, \ldots, p_k\}$. Верхняя граница при $n \approx 10^{100}$ выбирается равной $10^6 - 10^7$.

2. Фильтруем факторную базу, оставив в FB только такие элементы p, для которых n является квадратичным вычетом по модулю p. Иначе говоря, уравнение

$$n = k^2 \bmod p$$

должно иметь целое решение k. Такие простые p отбираются с помощью вычисления символа Лежандра. Сложность вычисления функции Лежандра имеет оценку $O(\log n \log p)$ (см. [17], с. 29-31), т.е. полиномиальна по отношению к длине числа n.

3. Следующий шаг заключается в формировании генерирующего полинома

$$q(x) = (x + m)^2 - n = x^2 + 2mx - a, \tag{4.95}$$

где $m = [\sqrt{n}]$ и $a = n - m^2$.

4. Потом для каждого $p \in FB$, нам нужно найти корни $r_1^{(p)}$, $r_2^{(p)}$ выражения

$$q(x) = 0(\mathrm{mod} p) \tag{4.96}$$

Для небольших p это можно сделать простой подстановкой чисел $0, 1, \ldots, (p-1)/2$ в уравнение (4.96), пока решение не будет найдено. Для больших p алгоритм Шенкса–Тонелли (D. Shanks, Tonelli), описанный в разд. 1.14, позволяет найти корни за время $O(\log^2 p)$. Если

один корень найден, то второй легко находится с помощью теоремы Виета: $r_1^{(p)} + r_2^{(p)} = 2m(\mathrm{mod}p)$. Предположим теперь, что для всех $p \in FB$ тройки $< p, r_1^{(p)}, r_2^{(p)} >$ уже найдены.

Последним шагом в формировании факторной базы является нахождение решения уравнений

$$q(x) = 0(\mathrm{mod}\, p^k) \tag{4.97}$$

для степеней $k > 1, \quad p^k < B$. Рассмотрим эту процедуру отдельно для случаев $p = 2$ и $p > 2$:

p=2. Рассмотрим два подслучая:

a) свободный член a уравнения (4.95) нечетен. Тогда сравнение (4.97) возможно только для нечетных x, причем если $x = 2y + 1$, тогда

$$(2y+1)^2 + 2m(2y+1) - a \equiv 0 \bmod 4 \leftrightarrow 1 + 2m - a \equiv 0 \bmod 4. \tag{4.98}$$

Если последнее уравнение не выполнено, тогда уравнение (4.95) не имеет решения при $k > 1$. Если сравнение (4.98) выполнено, то надо подставить $y = 2y + 1$ в (4.97), поделить получившееся уравнение на 4 и повторить те же действия с новым уравнением. Ниже мы рассмотрим пример решения уравнения (4.97).

b) свободный член a уравнения (4.95) четен. Тогда, сравнение (4.97) возможно только для четных x, причем если $x = 2y$, тогда

$$4y^2 + 4my - a \equiv 0 \bmod 4 \leftrightarrow a \equiv 0 \bmod 4. \tag{4.99}$$

Рассмотрим следующий пример:

Входные данные: n = 31593021658093170959102286152343 77.

Выберем $m \approx \sqrt{n} = 56207669990930215$, тогда $a = n - m^2 = 603298676152881520$

Случай $p = 2$

Для наглядности предположим, что ограничивающая константа $B = 1500$, тогда максимальная степень двойки, меньшая B, равна 10: $2^{10} = 1024 < B$.

Вычислим $m \bmod 1024 = 897$, и $a \bmod 1024 = 856$. Перепишем уравнение (4.99) в виде:

$$x^2 + 1614x - 856 \equiv 0 (\bmod 2^k).$$

В примере $a \equiv 0 (\bmod 4)$, поэтому решение существует только при $x = 2y$. Подставим $x = 2y$ в последнее уравнение:

$$4y^2 + 4 \cdot 807x - 856 \equiv 0 \,(\bmod 2^k).$$

Поделим последнее уравнение на 4:

$$y^2 + 807y - 214 \equiv 0 \,(\bmod 2^{k-2}). \tag{4.100}$$

Рассмотрим каждую последующую степень:

1. k=3. Заменим коэффициенты в (4.100) на их остатки по модулю 8. Получим уравнение

$$y^2 + y \equiv 0 \,(\bmod 2),$$

которое выполняется для всех y.

2. k=4. Заменим коэффициенты в (4.98) на их остатки по модулю 16. Получим уравнение

$$y^2 + 3y - 2 \equiv 0 \,(\bmod 2),$$

которое выполняется при $y \equiv 2 \,(\bmod 4)$ или $y \equiv 3 \,(\bmod 4)$.

3. Для последующих $k > 4$ корни уравнения (4.98) будем искать по индукции.

Любое решение должно иметь вид $z = y + 2^k \cdot r$, где y – решение уравнения (4.99) для предыдущей степени, r - целое число, которое будем искать. Подставим $z = y + 2^k \cdot r$ в уравнение $s(z) = z^2 + 807z - 214 \equiv 0 \,(\bmod\ 2^{k+1})$:

$$(y + 2^k \cdot r)^2 + 807(y + 2^k \cdot r) - 214 \equiv 0 \,(\bmod\ 2^{k+1})$$

или

$$s(y) + 807 \cdot 2^k \cdot r - 214 \equiv 0 \,(\bmod\ 2^{k+1}).$$

Учитывая, что $s(y) \equiv 0 \,(\mathrm{mod}\, 2^k)$, поделим последнее уравнение на 2^k:

$$f + r \equiv 0 \,(\mathrm{mod}\ 2), \text{ где } f = [s(y)/2^k] \,(\mathrm{mod}\, 2). \tag{4.101}$$

Итак, получим искомое решение в виде $z = y + 2^k \cdot f$, где f вычисляется по формулам (4.101).

Выполним соответствующие расчеты для нашего примера для степени $k + 1 = 5$ при $y \in \{2, 3\}$:

$f = ((y^2 + 807y - 214) \bmod 8)/4 = ((y^2 + 7y - 6) \bmod 8)/4$

1. $y = 2$, $f = ((4 + 14 - 6) \bmod 8)/4 = 1$. Получим первый корень $z_1 = 2 + 4f = 6$ и серию решений $z = 6 + 8t$, $t \in Z$.

1. $y = 3$, $f = ((9 + 21 - 6) \bmod 8)/4 = 0$. Получим второй корень $z_2 = 3 + 4f = 3$ и серию решений $z = 3 + 8t$, $t \in Z$.

Случай p > 2

Рассмотрим уравнение

$$q(x) = (x + m)^2 - n = x^2 + 2mx - a \equiv 0 \,(\mathrm{mod}\, p^k), \tag{4.102}$$

при $p > 2$.

Обозначим через z квадратный корень из n по модулю p. Тогда, уравнение (4.102) при $k = 1$ имеет корни

$$x = (-m \pm z)(\mathrm{mod}\, p).$$

Предположим теперь, что найдены корни x уравнения (4.102) для произвольного $k \geq 1$. Найдем корни для следующего значения $k + 1$. Эти корни имеют вид $y = x + p^{k+1} \cdot r$, где r – неизвестное значение. Сделаем подстановку y в уравнение (4.102):

$$(x + p^{k+1} \cdot r)^2 + 2m(x + p^{k+1} \cdot r) - a \equiv 0 \,(\mathrm{mod}\, p^{k+1}).$$

Перепишем последнее уравнение в виде:

$$q(x) + 2(x + m) \cdot p^k \cdot r \equiv 0 \,(\mathrm{mod}\, p^{k+1}).$$

Учитывая равенство $q(x) \equiv 0 \,(\mathrm{mod}\, p^k)$, поделим последнее уравнение на p^k. Получим:

$$f + 2(x+m)r \equiv 0 \,(\mathrm{mod}\, p),$$

где f – остаток от деления $q(x)/p^k$ на p.

Если $f = 0$, тогда $r = 0$ и $y = x$. В противном случае, с помощью расширенного алгоритма Евклида следует вычислить элемент u, обратный к элементу $2(x+m)$ в поле F_p, тогда решение y находится по формуле:

$$y = x - ufp^{k+1}. \tag{4.103}$$

II. *Процедура просеивания*

В ходе *инициализации* процедуры просеивания выбирается радиус просеивания L. Далее формируется массив W действительных чисел размерности $2L$. Вместо значений многочлена $q(x)$ в массив W будем помещать логарифмы значений $q(x)$: $W[x] = logq(x) \approx log(2m) + logx$ при $x \in [-L;\, L]$, взятые с небольшой точностью. Позднее мы оценим допустимую погрешность вычисления логарифма. Для нашего примера, когда n имеет 34 десятичных разряда, в значении логарифма достаточно брать два знака после десятичной точки.

Сама организация просеивания не имеет каких-то особых трудностей, а представляет собой двойной цикл в ходе которого перебираются все простые числа из факторной базы, для каждого из которых потом выполняется цикл по элементам массива W. Например, если тройка $< 19, 3, 10 >$ принадлежит факторной базе FB, это означает, что на 19 делятся $q(3)$, $q(10)$, а также все $q(y)$, $y \in [-L,\, L]$, $y \equiv 3 \bmod 19$ или $y \equiv 10 \bmod 19$. Можно организовать вычисление, найдя наименьшее $x \equiv 3 \bmod 19$, большее $-L$, вычесть из $W[x]$ $log19$, затем, присвоив x новое значение $x+19$, опять вычесть из $W[x]$ $log19$ и т.д. пока не дойдем до правого конца интервала $[-L,\, L]$.

После окончания просеивания по первым степеням элементов факторной базы, следует перейти к их степеням. Пусть, например, кортеж $< 19, 22, 105, 2 >$ принадлежит таблице степеней факторной базы. Тогда все числа $q(x)$ такие, что $x \equiv 22 \bmod 361$ или $x \equiv 105 \bmod 361$ делятся на

$361 = 19^2$. Но поскольку в ходе просеивания по первым степеням простых чисел, эти числа em уже были поделены на 19, то надо пробежать по интервалу $[-L;\ L]$ с шагом 361 и делить не на 361, а на 19.

Если $q(x)$ является гладким числом, тогда после выполнения полного просеивания, соответствующее значение $W[x]$ станет числом, близким по модулю к 0 (погрешность возникает из-за ошибок округления). Оценим сейчас общую погрешность, предположим что значения логарифмов вычисляются с точностью до ϵ, а факторная база ограничена сверху числом B.

На интервале $[-L;\ L]$ полином $q(x) = x^2 + 2mx - a$ принимает максимальное по модулю значение на концах интервала. Коэффициенты m и a имеют размерность $0(n^{1/2})$. Считается, что L на несколько порядков меньше m, поэтому максимальное значение $q(x)$ имеет порядок $0(L \cdot\ n^{1/2})$. Если считать среднее значение размера простого числа, входящего в разложение гладкого числа $q(x)$, равным B/c, где c некоторая константа, принимающая небольшие значения, например, $c = 10$, тогда можно подсчитать, сколько сомножителей входит в разложение максимального гладкого числа $q(x)$ на интервале $[-L;\ L]$:

$$k \approx \log_{B/c}(L^2 \cdot n^{1/2}).$$

В нашем примере с 34-значным числом n факторная база была ограничена сверху числом $B = 10^4,\ c = 10$, а $L = 2 \cdot 10^6$, тогда,

$$k \approx \log_{10^3}(4 \cdot 10^{12} \cdot 10^{17}) \approx 10.$$

Таким образом, в нашем примере гладкие числа представляют собой произведения, состоящие, в среднем, из 10 сомножителей.

Построение множества векторов показателей

В результате самой трудоемкой и затратной по времени процедуры просеивания в алгоритме квадратичного решета мы получим множество чисел $Smooth = \{x_1,\ x_2,\ ...,\ x_k\}$, принадлежащих интервалу просеивания $[-L,\ L]$, для которых значение полинома просеивания $q(x)$ является

гладким. Мощность этого множества значительно меньше радиуса L интервала просеивания.

Поскольку при первичном просеивании делители элементов $q(x_i)$ не накапливаются, т.к. это требует огромного количества памяти, то необходимо вторичное просеивание по элементам множества $Smooth$. Вторичное просеивание выполняется по тому же алгоритму, что и основное, но вместо всего интервала просеивания $[-L, L]$ пробегаются только элементы множества $Smooth$. Поэтому время выполнения этой стадии намного меньше времени первого просеивания.

Для хранения векторов степеней разложения элементов $Smooth$ необходимо определить массив $Vec[1..k, 0..s_z]$, где константа s_z равна размеру факторной базы + 1. Элемент $Vec[i, 0]$ могут принимать два значения:

$$Vec[i, 0] = \begin{cases} 0, \text{ если } q(x_i) > 0, \\ 1, \text{ иначе.} \end{cases}$$

Другие значения $Vec[i, j]$ равны степени, в которой элемент факторной базы FB p_j входит в разложение значения $q(x_i)$. Максимальные значения $Vec[i, j]$ достигаются на начальных значениях $j = 1, 2 \ldots$ и не превышает $\log_2 q(L)$.

4.4. Решение системы линейных уравнений

В результате предыдущего шага была сформирована недоопределенная система линейных уравнений с нулевой правой частью и коэффициентами из поля $F_2 = \{0, 1\}$, число уравнений m которой строго меньше числа неизвестных k.

$$A_{m \times k} \cdot X = 0. \tag{4.104}$$

Матрица системы получена соединением столбцов, каждый из который представляет вектор разложения j-о гладкого числа по факторной базе по модулю 2. Система обладает как минимум $k - m$ нетривиальными векторами решения. Особенностью системы является сильная ее разреженность, особенно в нижней ее части, соответствующей степеням старших простых чисел из факторной базы (см. пример в следующем разделе). Такая система

может быть легко решена с использованием обычного метода Гаусса. Сложность заключается в том, что размерность системы чрезвычайно велика и достигает величины $10^6 \times 10^6$ и более. Поэтому для решения соответствующей системы ЛУ используется специальные методы решения разреженных систем ЛУ, например, блочный метод Лантцоша (Lanszos' Block method). Питер Монтгомери ([38]) применил его для решения СЛАУ в методе решета числового поля. Описание этого метода можно также найти в обзоре ([23]).

Дадим краткое описание решения системы методом Гаусса. Алгоритм Гаусса состоит из двух этапов. На первом этапе матрица приводится к левотреугольному виду с нулями ниже главной диагонали, а на втором этапе – добиваемся, чтобы нули стояли выше главной диагонали. Первый этап метода Гаусса выполняется в цикле по строкам i матрицы системы $SystMatr[1..m, 1..k]$. На i-м шаге 1 этапа сначала выбирается ведущий элемент, отличный от 0. Если элемент $SystMatr[i, i] = 1$, он и будет ведущим элементом. Если же $SystMatr[i, i] = 0$, то организуем цикл по элементам i-й строки, находящимся правее $SystMatr[i, i]$, проверяя условие $SystMatr[i, j] = 1$ для $i < j \leq k$. Если элемент $SystMatr[i, j] = 1$ существует, то переставим i и j столбцы местами, делая элемент $SystMatr[i, i]$ ненулевым.

Возможна ситуация, когда такого ненулевого элемента $SystMatr[i, j]$ не существует. По конструкции левее $SystMatr[i, i]$ находятся также нули, поэтому этот случай означает, что строка i полностью нулевая. Такую строчку можно выбросить из матрицы без ущерба для дальнейшего вычисления.

4.5. Оценка сложности метода квадратичного решета

Мы используем здесь оценки Померанса [47]. Пусть число ε, принадлежащее интервалу $(0, 1)$ таково, что радиус интервала просеивания L оценивается величиной $L + N^{\varepsilon}$. Наибольшее значение полинома просеивания $q(x)$ наблюдается на границах интервала $[-L, L]$ и равно

$max\, q(x) \approx 2n^{1/2+\varepsilon}$. Оценим вероятность того, что случайно выбранное значение $q(x)$, $x \in [-L, L]$, является B–гладким, где B– верхняя граница факторной базы. Обозначим через $\psi(X, B)$ число B-гладких чисел в интервале $[1, X]$.

Вычислим сначала значение $\psi(X, B)$ при $B = x^{1/2}$:

$$\psi(X, X^{1/2}) = X - \sum_{X^{1/2}<p\le X} [X/p], \tag{4.105}$$

где p пробегает все простые числа из интервала $[1, X]$. Действительно, все числа, меньшие $X^{1/2}$ по определению $X^{1/2}$ – гладкие. Число из второй половины интервала $[1, X]$ не будет гладким, если оно кратно какому-нибудь простому числу p, $X^{1/2} < p < X$. Число таких кратных для каждого p равно $\lfloor X/p \rfloor$ (т.е. целой части от частного X/p). По теореме Чебышева, число простых чисел в интервале $[1, X]$, обозначаемое через $\pi(x)$, примерно равно $X/\ln X$. Отсюда

$$\psi(X, X^{1/2}) = X\left(1 \quad - \sum_{X^{1/2}<p\le X} 1/p\right) + 0(X/\ln X). \tag{4.106}$$

По теореме Мертенса,

$$\sum_{p\le t} 1/p = \ln\ln t + C + 0(1/\ln t), \tag{4.107}$$

для некоторой константы C. Используя эту теорему, получим

$$\sum_{X^{1/2}<p\le X} 1/p = \sum_{p\le X} 1/p - \sum_{p\le X^{1/2}} 1/p = \ln\ln X - \ln\ln(X^{1/2}) + 0(1/\ln X^{1/2}) =$$

$$= \ln 2 + 0(1/\ln X).$$

Подставляя последнее равенство в (4.106), получим

$$\psi(X, X^{1/2}) = (1 - \ln 2)X + 0(1/\ln X^{1/2}), \tag{4.108}$$

откуда

$$\frac{\psi(X, X^{1/2})}{X} \sim 1 - \ln 2 \text{ при } x \to \infty \tag{4.109}$$

Отсюда, около 30 % чисел, меньших X, является $\sqrt{X}$–гладкими.

Для произвольной границы $B = X^{1/u}$ выполняется формула К.Дикмана (K. Dickman [1930]):

$$\frac{\psi(X, X^{1/u})}{X} \sim \rho(u), \tag{4.110}$$

где $\rho(u)$ для $u > 0$ функция Дикмана–де Брюина. Эта функция является непрерывным решением дифференциального уравнения

$u\rho'(u)+\rho(u-1) = 0$ при $u > 1$ с начальным условием $\rho(u) \equiv 1$ на интервале $[0, 1]$.

Ее значение приблизительно равно $\rho(u) \approx u^{-u}$. Дадим здесь таблицу начальных значений этой функции

u	2	3	4	5	6	7	8
$\rho(u)$	0.25	$3,7 \cdot 10^{-2}$	$3,9 \cdot 10^{-4}$	$3,2 \cdot 10^{-4}$	$2,1 \cdot 10^{-5}$	$1,2 \cdot 10^{-6}$	$5,96 \cdot 10^{-8}$

Заметим, что величина $X/\psi(X, X^{1/u})$ равна среднему числу попыток для нахождения одного гладкого числа. Каждая попытка может быть оценена в $\ln \ln B$ элементарных операций. Нам требуется найти столько гладких чисел, сколько элементов в факторной базе, т.е. около $\pi(B)$ (в действительности, факторная база содержит примерно половину от этого числа, т.к. только половина p является квадратичными вычетами по модулю n). Отсюда можно подсчитать, что общее число операций $T(u)$ примерно равно

$$T(u) = \pi(B) \ln \ln B X/\psi(X, X^{1/u}) \approx X^{1/X} u^{1/u}. \tag{4.111}$$

Мы должны минимизировать эту величину. Вычислим ее логарифм:

$$\ln T(u) = \frac{1}{u} \ln X + u \ln u.$$

Производная этой функции равна 0, если $u^2(\ln u + 1) = \ln X$, откуда

$$u \sim (2 \ln X / \ln \ln X)^{1/2} \text{ и } B \approx e^{\sqrt{2 \ln X \ln \ln X)}}. \tag{4.112}$$

При выборе B равным выражению (4.112) при $X = n$, число элементарных операций метода квадратичного решета разложения числа n

оценивается величиной

$$T(n) = e^{c\sqrt{\ln n \ln\ln n}} \text{ при } c \in (1,\, 2). \tag{4.113}$$

Обозначим через $L(k, n)$ функцию

$$L_n(\alpha; c) = exp\left((c + o(1))(\ln n)^{\alpha}(\ln\ln n)^{1-\alpha}\right), \tag{4.114}$$

тогда оценка производительности метода квадратичного решета перепишется в виде

$$T(n) = L_n(1/2; c) = exp\left((c + o(1))(\ln n)^{1/2}(\ln\ln n)^{1/2}\right), \tag{4.115}$$

4.6. Пример факторизации методом квадратичного решета

Рассмотрим следующий пример факторизации натурального числа. Пусть $n = 750513679$, $m \approx \sqrt{n} = 27396$, $a = n - m^2 = -27137$. Полином просеивания (4.95) получает вид:

$$q(x) = x^2 + 54792x - 27137 \tag{4.116}$$

Замечание. Для таких маленьких чисел n оценка (4.111) является завышенной. Действительно, значение B равное $exp\left(\sqrt{2\ln n \ln\ln n)}\right) \approx 10^4$ слишком велико для нашего примера. Установим значение B равным 100.

I. ***Инициализация***

1. Определим $B = 100$. Число простых чисел, не превышающих B, равно 25. Для каждого такого такого p найдем остаток $n_p = n \bmod p$ и вычислим символ Лежандра $g = Leg(n_p,\, p)$. Если $g <> 1$, выбросим p из факторной базы. После выполнения такой фильтрации останется 14 простых чисел

$$FB = \{2,\, 3,\, 5,\, 11,\, 17,\, 23,\, 29,\, 43,\, 47,\, 53,\, 59,\, 61,\, 67,\, 83\}.$$

2. Для каждого элемента p факторной базы найдем, пользуясь алгоритм Шенкса, корень уравнения $x^2 \equiv n_p (\mathrm{mod}\, p)$, где $n_p = n \bmod p$,

и занесем найденные пары (x, p) в массив Roots:

$$Roots = \{(1,2),(3,1),(5,2),(11,5),(17,5),(23,9),(29,1),(43,35),(47,17),$$

$$(53,36),(59,7),(61,22),(67,39),(83,17)\}.$$

3. Заполним массив $Roots2$, размещая в нем тройки $< x, p, r >$, где x – корень уравнения $q(x) \equiv 0(\mathrm{mod}\, p^r)$ для $2 \le r \le k$, $p^k \le B$.

4. Выберем радиус интервала просеивания L так, чтобы число гладких чисел на интервале $[-L;\, L]$ немного превышало размер факторной базы. Поскольку размер факторной базы равен 14, то число гладких чисел должно быть не меньше 16. В нашем примере, экспериментальным путем выбрано $L = 300$.

5. Определим массив $W[-L\,..\,L]$ из действительных чисел и занесем в него логарифмы значения $q(x)$, $x \in [-L,\, L]$. Основание логарифма можно взять произвольным, например, $a = 2$ или $e = 2,71828...$. Требуемая точность вычисления логарифма зависит от длины n и не является высокой. В большинстве случаев достаточно стандартной точности реализации с 5–6 знаками после запятой. В нашем примере достаточно хранить 2 знака после запятой.

6. Подготовим действительный массив $LogFB[1..14]$, состоящий из логарифмов соответствующих элементов факторной базы: $LogFB[i] = \log p_i$.

II. ***Просеивание***

1. Выполним первое просеивание, в результате которого будут найдены гладкие числа из интервала $[-L\,..\,L]$. Для этого выполняется двойной цикл, в котором внешний цикл берется по элементам факторной базы, а внутренний – по числам $x \in [-L\,..\,L]$, удовлетворяющим уравнению $x \equiv r(\mathrm{mod}p)$, где r–корень уравнения $q(x) \equiv 0\,(\mathrm{mod}\, p)$. Напомним, что все такие корни занесены в массив $Roots[1..14]$.

2. Цикл просеивания для произвольного p может быть выполнен следующим образом. Ищем корни $\{r_1,\, r_2\}$ уравнения $q(x) \equiv 0\,(\mathrm{mod}p)$.

Для каждого r_i найдем наименьшее $x \in [-L, L]$, удовлетворяющее $x \equiv r(\mathrm{mod}\, p)$, и выполним следующий цикл:

```
while (x ≤ L) {
W[x] = W[x] − log p;
x = x + p; }
```

Закончив просеивание по элементам $p \in FB$, выполним просеиванию по степеням элементов факторной базы. При этом надо помнить, что если даже x–корень кратности $k > 1$, вычитание $W[x] = W[x] - \log p$ выполняется по-прежнему как для простого корня, т.к. к моменту рассмотрения корня x кратности k операция вычитания для $W[x]$ уже выполнялась $k - 1$ раз.

3. После завершения просеивания по степеням p получим в массиве $W[-L, L]$ несколько элементов, примерно равных 0 (с точностью, например, до 1 цифры после запятой). Занесем все x со значением $W[x] \approx 0$ в новый массив Smooth, который назовем массивом гладких чисел. В нашем примере множество Smooth будет состоять из 16 чисел:

$$\{-224, -166, -155, -99, -77, -40, -23, -21, -13, -12, 11, 22, 32, 41, 46, 268\}.$$

4. Поскольку при первом просеивании мы не сохраняли данные об элементах p, входящих в разложения $q(x)$, то необходимо повторное просеивание уже не по всему интервалу $[-L, L]$, а только по элементам множества Smooth. Это просеивание выполняется намного быстрее. В результате будет сформирован массив $Vec[1..16, 0..14]$, соответствующий матрице системы. Первая координата массива $Vec[i, j]$ равна номеру гладкого числа из множества Smooth, а вторая координата используется для хранения степени, в которой p_j входит в разложение числа x_i. $Vec[i, 0]$ обозначает знак $q(x_i)$, равный 0, если $q(x_i) > 0$, и равный 1, в противном случае. В нашем примере второе просеивание даст следующие результаты:

x	q(x)	Вектор разложения
-224	-12196095	$-3 \times 5 \times 23^2 \times 29 \times 53$
-166	-9040779	$-3^2 \times 11 \times 29 \times 47 \times 67$
-155	-5769579	$-3 \times 17 \times 29 \times 47 \times 83$
-99	-5387470	$-2 \times 5 \times 11 \times 17 \times 43 \times 67$
-77	-4185918	$-2 \times 3^8 \times 11 \times 29$
-40	-2162943	$-3^7 \times 23 \times 43$
-23	-1232550	$-2 \times 3^3 \times 5^2 \times 11 \times 83$
-21	-1123054	$-2 \times 17^2 \times 29 \times 67$
-13	-684990	$-2 \times 3^3 \times 5 \times 43 \times 59$
-12	-630223	$-11 \times 23 \times 47 \times 53$
11	629970	$2 \times 3 \times 5 \times 11 \times 23 \times 83$
22	1233045	$3^2 \times 5 \times 11 \times 47 \times 53$
32	1781505	$3^2 \times 5 \times 11 \times 59 \times 61$
41	2275290	$2 \times 3^4 \times 5 \times 53^2$
46	2549685	$3 \times 5 \times 43 \times 59 \times 67$
268	14783217	$3 \times 17^4 \times 59$

5. Для каждого $p \in$ FB подсчитаем количество гладких чисел, в разложение которых этот p входит в нечетной степени:

p	2	3	5	11	17	23	29	43	47	53	59	61	67	83
#p	6	9	8	7	2	4	4	4	4	3	4	1	4	3

Из таблицы 1 видно, что множитель $p = 61$ вошло только в одно гладкое число $q(32)$. Удалим 61 из факторной базы, а число $q(32)$ из множества Smooth. Остальные числа представим в виде векторов, состоящих из показателей степеней (первая координата представляет знак числа равный 1, если число - отрицательное, и 0, если число - положительное).

6. Сформируем матрицу системы A размерности 14×15, полагая

$$A(i+1,\, j) = \begin{cases} 1, \text{ если } p_i \text{ входит в разложение} q(x_j) \text{ в нечетной степени,} \\ 0, \text{ в противном случае.} \end{cases}$$

Первая строка матрицы определяет знаки соответствующих $q(x_j)$.

III. ***Решение системы***

Система уравнений с основной матрицей A и нулевым столбцом свободных членов содержит 15 неизвестных и 14 уравнений, т.е. является недоопределенной. Заменим все коэффициенты матрицы A их остатками по модулю 2. Тогда, все коэффициенты системы примут значения либо 0, либо 1. Можно рассматривать столбцы системы, как векторы длины 14. Множество столбцов состоит из 15 вектором, которые, очевидно, будут линейно зависимы, поэтому, можно выбрать непустое подмножество столбцов, нетривиальная линейная комбинация которых равна нулевому вектору. Поскольку, коэффициенты линейной комбинации равны либо 0, либо 1, то линейная комбинация является просто суммой векторов (по модулю 2).

Коэффициенты этой линейной комбинации могут быть найдены либо методом Гаусса, либо методом Лантцоша (см. [38]). В следующем разделе рассмотрим пример решения системы методом Гаусса.

4.7. Пример решения системы методом Гаусса

Рассмотрим пример нахождения решения системы методом Гаусса для $n = 2041$. Константу m положим равной 45, тогда $n = m^2 + 16$ и полином $q(x)$ равен $(x + m)^2 - n = x^2 + 90x - 16$. Определим $L = 5$ и вычислим значения $q(x)$ на интервале [-5; 5]:

x	-5	-4	-3	-2	-1	0	1	2	3	4	5
q(x)	-441	-360	-277	-192	-105	-16	75	168	263	360	459

Выберем факторную базу $FB = \{2, 3, 5, 7\}$ и выполним просеивание $q(x)$ на интервала [-5; 5] по этой факторной базе. Получим следующее представление:

x	-5	-4	-2	-1	0	1	2	4
q(x)	$-3^2 \cdot 7^2$	$-2^3 \cdot 3^2 \cdot 5$	$-2^6 \cdot 3$	$-3 \cdot 5 \cdot 7$	-2^4	$3 \cdot 5^2$	$2^3 \cdot 3 \cdot 7$	$2^3 \cdot 3^2 \cdot 5$

Таблица 1. Гладкие числа.

Перепишем эти разложения в виде матрицы, в которой столбец $\pm$

обозначает знак числа $q(x)$ и равен 1, если $q(x) < 0$, и 0, в противном случае. Заменим коэффициенты матрицы их остатками по модулю 2. Получим:

x	$\pm$	2	3	5	7
-441	1	7	0	0	2
-360	1	3	2	1	0
-192	1	6	1	0	0
-105	1	0	1	1	1
-16	1	4	0	0	0
75	0	0	1	2	0
168	0	3	0	0	1
360	0	3	2	1	0

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Припишем к полученной матрице единичную матрицу размерности 8×8 и будем выполнять эквивалентные преобразования со строками расширенной матрицы (прибавление ведущей строки к нижележащим строкам, и перестановку строк) с целью приведения матрицы к треугольному виду:

$$\left(\begin{array}{cccccccccccccc} 1 & 2 & 3 & 4 & 5 & -5 & -4 & -2 & -1 & 0 & 1 & 2 & 4 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccccccccccccc} 1 & 2 & 3 & 4 & 5 & -5 & -4 & -2 & -1 & 0 & 1 & 2 & 4 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

Последние три строчки содержат различные решение системы. Для нахождения какого-нибудь решения возьмем одну из этих строчек, например, строчку 6, и выделим столбцы, в которых в этой строке содержатся единицы: $X = \{-2,\, 0,\, 1\}$. Вычислим

$$A = \prod_{x \in X} (x+m) = (45-2)\cdot 45\cdot(45+1) = 89010, \qquad B^2 = \prod_{x \in X} q(x) = 480^2$$

Вычисляя теперь с помощью алгоритма Евклида Н.О.Д. $(n,\, A - B)$=Н.О.Д. $(2401,\, 88530) = 13$, найдем искомый делитель числа n. Второй делитель $q = 157$ можно вычислить как Н.О.Д. $(n,\, A+B)$, либо просто деля n на 13. То же решение можно получить, используя строку 7, а строка 8 дает

только тривиальные решения:

$$\text{Н.О.Д. } (n,\, A+B) = n,\ \text{Н.О.Д. } (n,\, A-B) = 1$$

Программа решения системы методом Гаусса

Приведем текст программы на языке Pascal для описанного выше алгоритма (исходные параметры mt, nt – число строк и столбцов соответственно). Матрица A – матрица системы размерности $mt \times nt$, а E – единичная матрица размерности $mt \times nt$. Массив $OrdRow[1..mt]$ нумерует строки системы. Исходные значения его элементов равны $OrdRow[i] = i$ для $1 \le i \le mt$. При необходимости перестановки строк с целью вывода на позицию $A[k,k]$ ненулевого элемента, мы просто меняем ссылки на строки в массиве $OrdRow$:

```
Procedure Gauss(mt,nt:int64;var A,E:SysMatr);
var
    i,j,k,c:integer;
Begin
  k:=1; { k enumerates rows of matrix }
  While k<=mt do
  begin
    j:=k;
       {Search for a non-zero element in the column k:}
    while (A[OrdRow[j],k]=0) And (j<=mt) do inc(j);
    if j>mt then { Case when all elements below A[k,k] are 0 }
    begin
       inc(k); continue;
    end;
    If j>k then { Case A[k,j]>0.Exchange k and j rows }
    begin
       c := OrdRow[j]; OrdRow[j]:=OrdRow[k];OrdRow[k]:=c;
    end;
   i:=k;
```

```
        {Making column k equal to 0 below  A[k,k]}
      while i < mt do
      begin
        inc(i);
        If A[OrdRow[i], k] = 1 Then
        begin
        For j := 1 To nt do
            A[OrdRow[i], j] := (A[OrdRow[i], j] + A[OrdRow[k], j])mod2;
        For j := 1 To mt do
            E[OrdRow[i], j] := (E[OrdRow[i], j] + E[OrdRow[k], j])mod2;
        end;
       end;
       inc(k);
   End; { of cycle by k}
End;
```

4.8. Вариация множителя в методе квадратичного решета

Одним из важных улучшений в методе квадратичного решета является идея *вариация большого множителя* – Large Prime Variations (LPV). Кратко эта идея состоит в следующем.

В процессе просеивания накапливается много значений полиномов $q(x)$, $x \in [-L; L]$, которые представимы в виде $q(x) = P_x \cdot C_x$, где множитель C_x является гладким числом, а P_x представляет собой множитель, превышающий границу гладкости B, но меньший B^2 такой, что P_x не содержит делителей из факторной базы. Множитель P_x является простым числом, т.к. иначе его можно было бы разложить в произведение двух сомножителей, меньших B. Назовем все такие значения *полугладкими.*

Предположим, что мы находимся на этапе, когда первое просеивание, определяющее номера x гладких чисел, завершено. Выполним цикл сортировки по всем $x \in [-L; L]$, для которых значение остаток значения $y = q(x)$ после всех делений на числа из факторной базы стал меньше B^2, упорядочив все полугладкие числа по возрастанию множителя P_x. Предположим, что найдется несколько значений y_1, y_2, ..., y_k, где $k \geq 2$,

$y_i = g(x_i)$, $-L \le x_i \le L$, с одинаковым значением множителя P_x. Тогда, можно добавить $k-1$ новых чисел вида

$g_1 = y_1 \cdot y_2$, $g_2 = y_1 \cdot y_3$, ..., $g_{k-1} = y_1 \cdot y_k$,

к множеству гладких чисел, поскольку каждое $g_j = P_x^2 \cdot C_1 \cdot C_{j+1}$, и, значит, может использоваться на следующем этапе для построения системы линейных уравнений.

Отметим, что дополнительное время, необходимое для сортировки полугладких элементов, является ничтожно малым на фоне полного времени просеивания, а выигрыш является значительным, поскольку не меняя размерности факторной базы мы можем получать большее число гладких чисел. Таким образом, используя вариацию больших множителей, можно увеличить выход гладких чисел на единицу простого числа факторной базы, что позволяет уменьшить границу гладкости B для факторной базы и уменьшить все основные затраты по поиску решения.

Граница B^2, предложенная для поиска полугладких чисел, является слишком большой, и вероятность того, что найдутся два полугладких числа с одинаковым множителем P_x, значительно уменьшается с ростом P_x, поэтому многие практические реализации квадратичного решета ограничивали поиск полугладких чисел на значительно меньшем интервале, например, на интервале $[B;\, kB]$, где множитель k находится в пределах от 10 до 100.

При разложении тестового 129-разрядного числа, предложенного создателями RSA, Х. Ленстра и М. Манассе в 1994 г. использовали дальнейшее развитие идеи LPV с двумя большими множителями вместо одного. Хотя такая реализация требует намного больше дополнительной работы, но обеспечивает больше возможностей для выбора полугладких чисел. В этом варианте граница полугладкости может быть доведена (хотя бы теоретически) до B^3, и рассматриваются числа, имеющие либо простой множитель в интервале $[B;\, B^2]$, либо составной множитель в интервале $[B^2;\, B^3]$. Для проверки того, что число с множителем в интервале $[B^2;\, B^3]$ является составным, используется самая простая проверка вида $2^{y-1} \not\equiv y$ метода Ферма. Если число y не проходит эту проверку, то

оно, вероятно, простое, и отбрасывается (хотя, при этом одновременно отбрасываются и многие составные числа). Если же число проходит проверку, то оно является детерминированно составным числом и подвергается разложению на множители, меньшие B^2 каким-нибудь несложным методом факторизации типа ρ–метода Полларда.

Дальнейшее развитие идеи больших множителей до трех или более множителей вызывает дополнительные сложности и считается неэффективным.

4.9. Варианты метода квадратичного решета с возможностью распараллеливания

Одним из благоприятствующих факторов, выделивших метод квадратичного решета среди других методов, явилась возможность распараллеливания процесса вычисления на несколько компьютеров (процессоров). Для этого П.Монтгомери была предложена идея метода квадратичного решета с использованием *множества полиномов (multiple polynomial quadratic sieve MPQS).*

Рассмотрим полином:

$$z_{a,b}(x) = (ax+b)^2 - n = a^2x^2 + 2abx + b^2 - n, \quad \text{где } a, b \in \mathbf{Z}. \tag{4.117}$$

Как и раньше, просеивание будет выполняться при различных значениях x в интервале $[-L, L]$. Наибольшее значение полинома $z_{a,b}(x)$ при $x \in [-L, L]$ примерно равно $a^2K^2 - n$, наименьшее примерно равно $-n$. Если взять $a \approx \sqrt{2n/M}$, то наибольшее и наименьшее значение станут примерно равны по абсолютному значению. Число b нужно выбрать таким образом, чтобы $b^2 - n$ делилось на a, то есть выполнялось $b^2 - n = ac$, $c \in \mathbf{Z}$. Тогда полином (4.117) можно записать в виде

$$z_{a,b}(x) = a(ax^2 + 2bx + c).$$

Это гарантирует, что каждое просеиваемое значение будет делиться на a. Пусть $a = t^2$, $t \in \mathbf{Z}$. Если некоторое произведение значений полинома

$q(x) = ax^2 + 2bx + c$ является полным квадратом

$$\prod_{x \in M} (ax^2 + 2bx + c) = A^2,$$

тогда полным квадратом будет и произведение

$$\prod_{x \in M} z_{a,b}(x) = \prod_{x \in M} t^2(ax^2 + 2bx + c) = (tA)^2$$

Просеивание будет проводиться среди $2L + 1$ значений полинома $q(x) = z_{a,b}(x)/a = ax^2 + 2bx + c$, каждый из которых ограничен сверху значением $L\sqrt{n/2}$. Делимость $b^2 - n$ на a означает, что $b^2 \equiv n (\mathrm{mod}\ t^2)$. Если выбрать число t простым, так чтобы n было квадратичным вычетом по модулю t, то b можно легко вычислить по алгоритму Шенкса.

Следующий вопрос состоит в том, насколько большим нужно брать L. При большом значении L на каждом полиноме будет просеиваться много чисел большого размера, среди которых гладких чисел будет сравнительно мало, что нежелательно. Если взять L малым, то числа для просеивания также будут небольшими, но нужно генерировать большое количество полиномов, подбирая коэффициенты a,b. Но для каждого нового полинома нужно заново формировать факторную базу, т.е. заново выполнять этап инициализации, что также является затратной процедурой. Таким образом, если L слишком мало, то большая часть времени будет тратиться не на просеивание, а на инициализацию полиномов. Если удастся сократить время инициализации, то можно будет использовать меньшие значения L.

Самоинициализирующееся квадратичное решето

Алгоритм самоинициализирующегося квадратичного решета (self initializing quadratic sieve) позволяет менять полиномы за малое время, что дает возможность использовать меньшие значения L, нежели в алгоритме квадратичного решета с множеством полиномов. Основная идея метода состоит в том, чтобы использовать несколько полиномов с одним и тем же значение параметра a и различными b, $0 < b < a/2$.

Если значения a, b, c выбраны, то этап инициализации заключается в решении уравнения

$$ax^2 + 2bx + c \equiv 0 \,(\bmod p) \tag{4.118}$$

для каждого простого $p < B$, где B–граница факторной базы, для которого уравнение

$$t^2 \equiv n (\bmod p) \tag{4.119}$$

имеет решение. Если уравнение (4.119) разрешимо, то обозначим через t_p его корень (любой из двух возможных). Тогда решения уравнения (4.118) можно записать в виде

$$r_1 = (-b + t_p)a^{-1} \,(\bmod p), \;\; r_2 = (-b - t_p)a^{-1} \,(\bmod p). \tag{4.120}$$

Таким образом, основная работа на этапе инициализации состоит в вычислении $a^{-1} \,(\bmod p)$ для каждого простого числа $p \in FB$.

Зададим полином $z_{a,b}(x) = a(ax^2 + 2bx + c)$, в котором a является произведением нескольких простых чисел из факторной базы. Полином $z_{a,b}(x)$ будет гладким тогда и только тогда, когда гладким будет значение выражения $ax^2 + 2bx + c$. Кроме того для каждого значения a можно найти несколько соответствующих ему значений коэффициента b. Пусть $a = q_1 \cdot q_2 \cdot, \ldots q_s$, $q_i \in FB$. Для того чтобы сгенерировать полином нужно найти b, $b^2 \equiv n \,(\bmod a)$. Существует 2^s значений $b \,(\bmod a)$, удовлетворяющих этому условию. Но только половина из них может быть использована, так как $z_{a,b}(x)$ дает те же значения вычетов, что и $z_{a,-b}(x)$. Таким образом можно найти 2^{s-1} полиномов при неизменном значении коэффициента a.

Нужно найти целые числа B_i, $1 \le i \le s$, удовлетворяющие условиям:

$$B_i^2 \equiv n \;(\bmod q_i), \quad B_i \equiv 0 \;(\bmod a/q_i) \tag{4.121}$$

тогда по китайской теореме об остатках любое значение b можно записать в виде

$$b = \pm B_1 \pm B_2 \ldots \pm B_s \,(\bmod a), \tag{4.122}$$

что позволяет быстро вычислить все требуемые значения b по найденным значениям B_i.

Идея самоинициализирующегося метода решета была предложена Померансом, Смитом и Тулером в 1988 ([48]). Вильям Харт (William Hart) разработал пакет программ SIMPQS для реализации этого метода (доступен код на сайте http://www.friedspace.com/QS), использующий библиотеку работы с длинными числами GMP).

4.10. Метод Занга (Zhang' Special QS)

Метод Занга (Zhang' Special Quadratic Sieve) был разработан в 1998 г. ([55]). Идея этого метода состояла в использовании многочленов более высокой степени, чем квадратные, также являющихся полными квадратными по модулю n. Пусть n–факторизуемое число, а m–натуральное число, равное примерно $n^{1/3}$. Разложим число n в m-й системе координат:

$$n = m^3 + a_2 m^2 + a_1 + a_0. \tag{4.123}$$

Разложение считается более хорошим, если коэффициенты a_i принимают меньшие значения по абсолютной величине. Выбор m позволяет изменять значения коэффициентов. Пусть b_2, b_1, b_0– изменяемые параметры по множеству $\mathbf{Z}$, и рассмотрим

$$x = b_2 m^2 + b_1 m + b_0. \tag{4.124}$$

Имеем следующие соотношения:

$$\begin{aligned} m^3 &\equiv -a_2 m^2 - a_1 m - a_0 \ (\bmod n) \\ m^4 &\equiv (a_2^2 - a_1) m^2 + (a_1 a_2 - a_0) m + a_0 a_2 \ (\bmod n), \end{aligned} \tag{4.125}$$

откуда получим,

$$x^2 \equiv c_2 m^2 + c_1 m + c_0, \ (\bmod n), \tag{4.126}$$

где

$$c_2 = (a_2^2 - a_1)b_2^2 - 2a_2b_1b_2 + b_1^2 + 2b_0b_2$$
$$c_1 = (a_1a_2 - a_0)b_2^2 - 2a_1b_1b_2 + b_1^2 + 2b_0b_1$$
$$c_0 = a_0a_2b_2^2 - 2a_0b_1b_2 + b_0^2.$$

Подберем параметры b_0, b_1, b_2 так, чтобы коэффициент c_2 был равен нулю. Для этого положим

$$b_2 = 2, \quad b_1 = 2t, \quad b_0 = a_1 - a_2^2 + 2a_2t - t^2, \tag{4.127}$$

где t принимает произвольное целые значения. Подставляя значения (4.127) в формулы (4.125), получим эквивалентность

$$x(t)^2 \equiv y(t) \ (\bmod n)$$

где

$$x(t) = 2m^2 + 2mt + a_1 - a_2^2 + 2a_2t - t^2$$
$$y(t) = (4a_1a_2 - 4a_0 - (4a_1 + 4a_2^2)\,t + 8a_2t^2 - 4t^3)\,m +$$
$$+4a_0a_2 - 8a_0t + (a_1 - a_2^2 + 2a_2t - t^2)^2).$$

В этом случае можно, меняя значения t на некотором интервале просеивания, искать пары чисел (x, y) так же, как в обычном методе квадратичного решета.

Метод Занга существенно зависит от значений коэффициентов в формуле (4.123), поэтому имеет ограниченную область применения. Его можно использовать для разложения чисел специального вида.

Рассмотрим, например, число $n = 2^{601} - 1$. Используя метод разложения, зависящий от младшего сомножителя, например, метод эллиптических кривых, найдем два младших делителя n:

$$n = 3607 \cdot 64\,863\,527 \cdot n_0,$$

где n_0–составное 170-разрядное десятичное число. Рассмотрим число

$$4n = 2^{603} - 4 = (2^{201})^3 - 4 = m^3 - 4,$$

где $m = 2^{201}$. Тогда формулы для $x(t)$, $y(t)$ приобретают вид:

$$x(t) = 2m^2 + 2mt - t^2, \quad y(t) = (16 - 4t^3)m + 32t + t^4).$$

Рост значений полинома $y(t)$ зависит от значений t в 4-й степени, поэтому t не может быть слишком большим. Так же, как и в методе квадратичного решета, в методе Занга можно использовать различные полиномы, что несколько улучшает общую сходимость метода. Однако, в целом, этот метод метод работает не лучше, чем метод квадратичного решета с многими полиномами, а если значения коэффициентов a_i велики, то уступает ему. В случае чисел специального вида метод Занга уступает методу решета числового поля, о котором мы будем говорить в следующей главе. Поэтому метод Занга не получил широкого распространения и не использовался в известных проектах разложений больших чисел.

5. Метод решета числового поля

В этой главе мы дадим описание самого быстрого на сегодняшний день алгоритма факторизации натуральных чисел – метода решета числового поля. Идея этого метода принадлежит Джону Полларду, который в 1988 г. распространил среди коллег письмо (см. [44]), в котором предложил выполнять просеивание не в кольце целых чисел Z, как в методе квадратичного решета, а в алгебраическом числовом поле. Первоначально этот метод можно было использовать только для разложения чисел специального вида $2^n \pm c$, поэтому метод получит название «специального решета числового поля» (the Special Number Field Sieve SFNS).

Практическая реализация идеи Полларда была выполнена А. Ленстрой, Х. Ленстрой, М. Манассе и Д. Поллардом (см.[32]) в 1990 г., когда с помощью этого метода было факторизовано 9–е число Ферма 2^{2^9}. Также были факторизованы некоторые числа вида $b^c \pm 1$ из проекта Каннингама (см.[11]).

Вскоре было замечено, идею Полларда можно использовать и для разложения произвольных чисел. Этот метод получил название обобщенного решета числового поля (the General Number Field Sieve GFNS). Была найдена эвристическая оценка нового метода, которая равнялась $L_n(1/3, c)$ в терминах функции $L_n(\alpha, c)$ (4.114), т.е. множитель α был уменьшен со значения $\alpha = 1/2$ в методе квадратичного решета до $\alpha = 1/3$ в решете числового поля. Это обеспечило новому методу значительный выигрыш по отношению ко всем другим ранее известным методам.

Сборник статей под редакцией А. Ленстры и Х. Ленстры [33], опубликованный издательством Springer в 1993 г., подвел итоги раннего развития этого метода. Перечислим основные источники литературы, в которых можно найти описание метода решета числового поля: [12], [13], [21], [23], [32], [46], [47], [64].

Опишем в следующем разделе основные идеи GNFS. Мы будем придерживаться, в основном, книги Бригса [12].

5.1. Базовый алгоритм решета числового поля

Пусть n–нечетное составное число, которое необходимо факторизовать. Главным недостатком метода квадратичного решета заключается в том, что значения полинома просеивания $q(x) = (m+x)^2 - n$ растут очень быстро при увеличении аргумента x на интервале просеивания $[-L;\, L]$. Например, при $L = O(10^{10})$ и $n = O(10^{100})$ значение $q(x)$ достигает значений $O(10^{60})$. Ясно, что факторная база для таких чисел должна была быть также огромной.

Революционная идея Джона Полларда состояла в предложении заменить полином 2–й степени $q(x) = (x + m)^2 - n$, используемый в квадратичном решете, произвольным полиномом с целыми коэффициентами $P_d(x)$ степени $d \geq 3$, удовлетворяющим условию $P_d(m) = n$ для некоторого целого m, а просеивание по множеству целых чисел $\mathbf{Z}$ – просеиванием в кольце $\mathbf{Z}[\theta]$, полученном присоединением к кольцу $\mathbf{Z}$ *целого алгебраического* числа θ, являющегося корнем полинома $P_d(x)$. Вместо простых натуральных чисел соответствующая факторная база будет состоять теперь из неразложимых простых элементов кольца целых алгебраических чисел, а проверка делимости будет выполняться с использованием нормы алгебраического числа. Математическая теория метода решета числового поля строится на основе теории делимости в алгебраических числовых полях, основные положения которой мы дадим в приложении А.

Выигрыш в использовании такой идеи состоит, в первую очередь, в том, что условие $P_d(m) = n$ для некоторого целого m, накладываемое на полином $P_d(x)$ может быть выполнено при значительно меньших значениях коэффициентов полинома $P_d(x)$, по сравнению с коэффициентами полинома $q(x)$, используемого в методе квадратичного решета. Это влечет уменьшение абсолютных значений полинома в области просеивания и общее увеличение эффективности метода.

Описание этапов алгоритма

Ниже мы дадим описание основных этапов алгоритма решета числового числового поля, опуская детали, которые будут описаны в последующих

разделах:

1. Выберем степень неприводимого многочлена $d \geq 3$ (можно взять $d = 2$, но в этом случае никакого выигрыша по отношению к методу квадратичного решета не будет).

2. Выберем целое число m, $\lfloor n^{1/(d+1)} \rfloor < m < \lfloor n^{1/d} \rfloor$, и разложим n по основанию m:

$$n = m^d + a_{d-1}m^{d-1} + \ ... \ + a_0. \tag{5.128}$$

   Отметим, что в базовой версии алгоритма старший коэффициент многочлена $f_1(x)$ всегда равен 1, однако, на с. 169 мы обсудим, как использовать многочлены с произвольным старшим коэффициентом.

3. С разложением (5.128) свяжем неприводимый в кольце $Z[x]$ (кольцо полиномов от переменной x с целыми коэффициентами) полином

$$f_1(x) = x^d + a_{d-1}x^{d-1} + \ ... \ + a_0. \tag{5.129}$$

4. Определим полином просеивания $F_1(a, b)$ как однородный полином от двух переменных a и b:

$$F_1(a, b) = b^d \cdot f_1(a/b) = a^d + a_{d-1}a^{d-1}b + a_{d-2}a^{d-1}b^2 + \ ... \ + a_0 b^d. \tag{5.130}$$

   Забегая вперед, скажем, что значение $F_1(a, b)$ равно норме полинома $a - bx$ в алгебраическом числовом поле $\mathbf{Q}[\theta]$, полученном добавление к полю рациональных чисел $\mathbf{Q}$ в общем случае комплексного корня θ многочлена $f_1(x)$ (см.стр. 184). Свойство коммутативности нормы $Nr(h_1(x) \cdot h_2(x)) = (Nr(h_1(x)) \cdot Nr(h_2(x))$ позволяет вместо разложения многочленов из кольца $\mathbf{Z}[\theta]$ выполнять разложение их норм.

5. Также определим второй многочлен $f_2(x) = x - m$ и соответствующий однородный многочлен $F_2(a, b) = a - bm$. Главное требование к выбору пары $(f_1,\ f_2)$ состоит в том, что должно выполняться условие:

$$f_1(m) \equiv f_2(m) \ (\bmod n), \tag{5.131}$$

которое, очевидно выполняется в нашем случае, т.к. первый многочлен в т. m равен n, а второй – нулю.

6. Выберем два положительных числа L_1 и L_2, которые определяют некоторую прямоугольную область $SR = \{1 \le b \le L_1,\ -L_2 \le a \le L_2\}$, называемую областью просеивания (sieve region).

7. Пусть θ–корень многочлена $f_1(x)$. Рассмотрим кольцо многочленов $\mathbf{Z}[\theta]$ (практически корень θ не вычисляется, а используется только для формального описания алгоритма). Определим *алгебраическую факторную базу* $FB1$, состоящую из многочленов первого порядка вида $a - b\theta$ с нормой (5.130), являющейся простым числом. Такие многочлены являются простыми неразложимыми элементами в кольце алгебраических целых поля $K = \mathbf{Q}[\theta]$. Абсолютные величины норм многочленов из факторной базы FB_1 ограничим сверху некоторой константой B_1.

8. Одновременно определим *рациональную факторную базу* FB_2, состоящую из всех простых чисел, ограниченных сверху второй константой B_2.

9. Напомним, что произвольный элемент a кольца K называется квадратичным вычетом, если найдется элемент $x \in K$ такой, что $x^2 = a$. Чтобы иметь возможность проверять на заключительной стадии алгоритма, является ли найденный в ходе просеивания многочлен полным квадратом, определим сравнительно небольшое множество многочленов 1 порядка $c - d\theta$, норма которых также является простым числом. Обозначим это множество как FB_3. Оно должно удовлетворять условию $FB_1 \cap FB_3 = \emptyset$ и называется *факторной базой квадратичных характеров (the Quadratic Character Base).*

10. Далее выполняется одновременное просеивание многочленов $\{a - b\theta \mid (a,b) \in SR\}$ по факторной базе FB_1 и целых чисел $\{a - bm \mid (a,b) \in SR\}$ по факторной базе FB_2 с целью получения

множества M, состоящего из *гладких пар* (a, b). Пара (a, b) называется гладкой, если Н.О.Д.$(a, b) = 1$, и полином $a - b\theta$ и число $a - bm$ раскладываются полностью по соответствующим факторным базам FB_1 и FB_2. Число гладких пар в множестве M должно превышать суммарную мощность трех факторных баз, по-крайней мере, на две единицы.

11. На следующем шаге ищется подмножество $S \subseteq M$ такое, что произведение всех пар

$$\prod_{(a,b)\in S} Nr(a - b\theta) = H^2, \text{ для } H \in \mathbf{Z}, \text{ и } \prod_{(a,b)\in S} (a - bm) = B^2, \ B \in \mathbf{Z}.$$

Для нахождения множества S составляется, как и в методе квадратичного решета, система линейных алгебраических уравнений с коэффициентами из множества $F_2 = \{0, 1\}$, решением которой и будут номера множества S.

12. Далее формируем многочлен

$$g(\theta) = (f_1'(\theta))^2 \cdot \prod_{(a,b)\in S} (a - b\theta), \tag{5.132}$$

где $f_1'(x)$–производная многочлена $f_1(x)$.

13. Если вся процедура была выполнена корректно, то многочлен $g(\theta)$ является полным квадратом в кольце полиномов $\mathbf{Z}[\theta]$. Извлекаем квадратные корни из многочлена $g(\theta)$ и целого числа B^2, находя некоторый многочлен $\alpha(\theta)$ и число B.

14. Заменяем многочлен $\alpha(\theta)$ числом $\alpha(m)$. Отображение $\phi : \theta \to m$ является кольцевым гомоморфизмом кольца алгебраических целых чисел $\mathbf{Z}_K$ в кольцо $\mathbf{Z}$, откуда получим соотношение:

$$A^2 = g(m)^2 \equiv (\phi(g(\alpha)^2) \equiv \phi\left((f_1'(\theta))^2 \cdot \prod_{(a,b)\in S} (a - b\theta)\right) \equiv$$

$$\equiv (f_1'(m))^2 \cdot \prod_{(a,b)\in S} (a - bm) \equiv (f_1'(m))^2 \cdot C^2 \ (\mod n) \tag{5.133}$$

Определив $B = f'(m) \cdot C$, найдем пару целых чисел (A, B), удовлетворяющих условию

$$A^2 \equiv B^2 \ (\mathrm{mod}\, n).$$

Теперь можно найти делитель числа n так же, как в методе квадратичного решета, вычисляя Н.О.Д. $(n,\ A \pm B)$.

В следующей части мы дадим более подробное описание основных шагов метода решета числового поля, выполнив все необходимые расчеты для числа $n = 45113$. Пример взят из книги М.Бриггса [12].

Пример

Найти разложения числа $n = 45\,113$. Вычислим $n^{1/3} = 35,4...$, и определим $m = 35$. Раскладывая n в m–й системе исчисления, получим

$$n = m^3 + m^2 + 28m + 33.$$

Получившийся кубический многочлен имеет большие коэффициенты, близкие к максимальному значению $m - 1 = 34$. Вычтем m из последнего коэффициента, одновременно добавляя 1 к коэффициенту при m в 1-й степени. Получим:

$$n = m^3 + m^2 + 29m - 2.$$

Вычтем m из предпоследнего коэффициента, одновременно добавляя 1 к коэффициенту при m в квадрате. Получим:

$$n = m^3 + 2m^2 - 6m - 2.$$

Такой многочлен выглядит намного лучше с точки зрения меньшего роста значений $F_1(a, b)$ в области просеивания, что позволит найти больше гладких чисел при неизменном размере области просеивания. Меняя также основание m от значений $m^{1/(d+1)}$ до $m^{1/d}$, можно также менять значения коэффициентов.

Операция уменьшения коэффициента a_i на величину m с последующим добавлением 1 к a_{i+1}, эквивалентна добавления к $f_1(x)$

многочлена $x^i(x-m)$. Обобщая эту идею, можно заменить, что сложение $f_1(x)$ с полиномом вида $g(x)\cdot f_2(x)$, где $g(x)$–произвольный многочлен с целыми коэффициентами, сохраняет свойство (5.131).

Таким образом, изменение основания m и добавление (вычитание) из $f_1(x)$ полиномов вида $g(x)\cdot f_2(x)$ позволяет порождать многочисленные допустимые полиномиальные пары $(f_1(x),\, f_2(x))$, среди которых можно выбрать пару с наименьшим размахом коэффициентов. Проблема выбора наилучшего полинома подробно исследуется в диссертации Б.Мерфи ([41]). Мы обратимся к этой проблеме в разделе 5.7. В нашем примере остановим свой выбор на $m=31$ и полиноме разложения

$$n = m^3 + 15m^2 + 29m + 8.$$

Объясним на этом примере все последующие алгоритмы метода числового решета.

5.2. Выбор факторных баз

Рассмотрим сначала процедуру построения рациональной факторной базы FB_2. Эта база будет использоваться для разложения чисел вида $a-bm$ в множестве $\mathbf{Z}$, поэтому множество FB_2 определяется равным множеству всех простых чисел, ограниченных сверху константой B_1. Граница B_1 для рекордных разложений доходит до гигантских значений $10^6 - 10^7$. В нашем примере для $n = 45113$ определим $B_1 = 30$ и $FB_2 = \{2,\, 3,\, 5,\, 7,\, 11,\, 13,\, 17,\, 19,\, 23,\, 29\}$, состоящим из 10 простых чисел.

Построение алгебраической факторной базы

Алгебраическая факторная база состоит из линейных многочленов $c-d\theta$, порождающих простые идеалы в кольце целых алгебраических чисел $\mathbf{Z}_K$. Построение такой факторной базы является очень сложной задачей, однако, следующая теорема (М. Бригс[12], теор.3.1.7) позволяет перейти от неприводимых многочленов и порождаемых ими простых идеалов к парам натуральных чисел $(p,\, r)$:

Теорема 5.1. *Множество простых идеалов кольца* Z_K *находится во взаимно–однозначном соответствии с множеством пар положительных целых чисел* $(p,\, r)$ *таких, что* p*–простое число,* $0 \le r < p$*, и* $f_1(r) \equiv 0\,(\bmod p)$.

Благодаря этому результату, можно не выписывать явно простые идеалы кольца $\mathbf{Z}_K$, а просто установить границу B_1 и искать все пары $(p,\, r)$, где $p \le B_1$–простое число, а $r \in [0,\, p-1]$, и $f_1(r) \bmod p = 1$.

В нашем примере установим границу $B_1 = 103$. Элементы FB_1 можно найти простым перебором (табл.1).

Таблица 1

Алгебраическая факторная база

$(p,\, r)$	$(p,\, r)$	$(p,\, r)$	$(p,\, r)$
$(2, 0)$	(41,19)	(67,44)	(89,62)
$(7, 6)$	(43,13)	(73,50)	(89,73)
(17,13)	(53,1)	(79,23)	(97,28)
(23,11)	(61,46)	(79,47)	(101,87)
(29,26)	(67,2)	(79,73)	(103,47)
(31,18)	(67,6)	(89,28)	

Простой перебор всех допустимых пар работает неэффективно для больших чисел p. Опишем здесь один алгоритм, позволяющий улучшить поиск корней многочлена $f_1(x) \bmod p$. Этот алгоритм основывается на следующей теореме:

Теорема 5.2. *В конечном поле* $GF_q = GF_{p^k}$ *многочлен* $x^q - x$ *полностью раскладывается на линейные множители*

$$x^q - x = \prod_{i=0}^{q-1}(x - i). \tag{5.134}$$

Алгоритм вычисления корней по модулю простого числа p работает следующим образом:

1. Ищем $g(x) = \text{Н.О.Д.}(f_1(x), x^p - x)$. Целью этого шага является отсечение части полинома $f_1(x)$, имеющей корни по модулю p. Например, если окажется, что $g(x) = 1$, тогда $f_1(x)$ не имеет корней по модулю p.

2. По теореме 5.2 $g(x-b) \mid x^p - x$ и

$$x^p - x = x(x^{(p-1)/2} + 1)(x^{(p-1)/2} - 1) \text{ для любого } b,\ 0 \le b < p \qquad (5.135)$$

3. Используя свойство 5.135 отделим корни $g(x) \bmod p$.

Рассмотрим этот алгоритм на примере многочлена $f_1(x) = x^3 + 15x^2 + 29x + 8$, $p = 67$:

1. Вычислим $g(x) = \text{Н.О.Д.}(f_1(x), x^p - x) = \text{Н.О.Д.}(, x^67 - x) = x^3 + 15x^2 + 29x + 8$. Получили $g(x) = f_1(x)$, значит, $f_1(x)$ содержит все три возможных корня $f_1(x)$ по модулю 67.

2. При $b = 0$ равенство (5.135) получит вид

$$g(x) = x^3+15x^2+29x+8 \ \mid \ x(x^{33}+1)(x^{33}-1)$$

Так как $g(0) = 8 \neq 0$, $x = 0$ не является корнем $g(x)$.

3. Вычислим $\text{Н.О.Д.}(g(x), x^{33}+x) = x^2+21x+21$, и $\text{Н.О.Д.}(g(x), x^{33} - x) = x + 61$, откуда найдем корень $x = -61 \equiv 6 \pmod{67}$. Два оставшихся корня являются корнями многочлена $g_1(x) = x^2 + 21x + 21 \equiv x^2 + 21x - 46 \pmod{67}$. Их можно, например, вычисляя с помощью алгоритма Шенкса–Тонелли квадратный корень из дискриминанта этого уравнения $D = (-21)^2 + 4 \cdot 46 = 2998 \equiv 50 \pmod{67}$.

Повторим это вычисление с другими значениями b. Подставляя в равенство (5.135) $b = 1$, получим $\text{Н.О.Д.}(g(x-1), x^{33} + x) = g(x-1)$ и $\text{Н.О.Д.}(g(x-1), x^{33} - x) = 1$. Значит, расщепления $g_1(x_1)$ не происходит и $b = 1$ не решает нашей задачи.

Подставим $b = 2$. Получим $\text{Н.О.Д.}(g(x-2), x^{33} + x) = x + 21$ и $\text{Н.О.Д.}(g(x-2), x^{33}-x) = x+63$, откуда $x = -21 \equiv 46 \pmod{67}$, $x = -63 \equiv 6 \pmod{67}$. Числа $x = 44$ и $x = 2$ являются корнями $f_1(x) \pmod{67}$. Все три корня найдены.

Построение алгебраической базы характеров

Построение третьей факторной базы, состоящей из квадратичных характеров, выполняется также, как и алгебраической. Выбирается новый ограничитель $B_3 > B_2$ и рассматриваются все простые числа от B_2 до B_3. Для рекордных разложений B_3 выбирается так, чтобы в интервал $[B_2, B_3]$ попало $10^4 - 10^5$ простых чисел. Выбор размерности B_3 определяет степень уверенности в том, что найденное в результате просеивания полином является полным квадратом (см. Buchler et alt. [13]).

В нашем примере возьмем $B_3 = 109$, и построим квадратичную факторную базу способом, описанным ранее (см.табл.2).

Таблица 2

Факторная база квадратичных характеров

(p, r)	(p, r)	(p, r)
(107,4)	(107,80)	(109,99)
(107,8)	(109,52)	

Расчет необходимого количества гладких пар

Вычислим теперь суммарный объем трех факторных баз $s = 10 + 23 + 5 = 38$. Значит, область просеивания должна содержать не менее 40 элементов (добавляется один столбец для хранения знака числа $a - bm$ плюс дополнительная единица для того, чтобы матрица системы была недоопределена и имела ненулевое решение (см.разд. 4.4). Полученное значение влияет на размер области просеивания, в которой ищутся гладкие пары, и определяет размер системы линейных уравнений, формируемой для нахождения решения.

5.3. Просеивание в решете числового поля

Процедура просеивания в методе решета числового поля, в целом, мало чем отличается от аналогичной процедуры в методе квадратичного решета.

Область просеивания представляет собой прямоугольник вида

$$\mathrm{S}P = \{(a,\, b) \mid 1 \le b \le L_2;\ \ -L_1 \le a \le L_1\}. \tag{5.136}$$

Линейное просеивание выполняется в виде двойного цикла, в котором внешней переменной является переменная b, принимающая последовательно значения $1,\, 2,\, ...,\, L_2$, и с каждым значением b_1 выполняется внутренний цикл по a, изменяющемся от $-L_1$ до L_1.

Джон Поллард в работе [45], опубликованной в 1993 г., предложил вместо линейного использовать *решеточное* просеивание, которое заключается в том, что выбирается несколько больших p из пар (p, r), содержащихся в алгебраической факторной базе. Такие p называются специальными числами. Просеивание выполняется теперь не вдоль прямых $b = i,\ |a| \le L_1$, а вдоль прямых

$$L_{p,r} = \{(a, b) \in SR \,|\, a - br \equiv 0 \ (\mathrm{mod}\ p)\}.$$

Смысл такого просеивания в том, что на указанной прямой значения $F_1(a, b) \equiv 0 \,(\mathrm{mod}\ p)$, т.е. делятся нацело на p, и можно выполнять просеивание по значениям $F_1(a, b)/p$, которые принимают заведомо меньшие по модулю значения, чем исходные $F_1(a, b)$. Решеточное просеивание позволяет в несколько раз увеличить эффективность процедуры просеивания, однако область просеивания должна быть не прямоугольником, а иметь более сложную форму, что усложняет процедуру.

Следует заметить, что в рекордных проектах разложениях присутствуют обе эти стратегии, поскольку линейное просеивание также дает много гладких пар на некоторых специально выбранных прямых.

В работе ([67]) автор этой монографии со своими соавторами предлагает другую организацию просеивания, рассматривая область прилегающую к прямой $a = x_0 b$, где x_0—действительный корень полинома $f_1(x)$. Идея такого просеивания состоит в том, что вдоль действительного корня значения полинома $F_1(a, b)$ растут на порядок медленнее, чем в произвольном другом направлении.

Первичное просеивание для $n = 45113$

Для нашего примера $n = 45113$ выберем линейное просеивание со значением $L_1 = 1000$ и b, принимающим значения 1, 2, ..., пока не будет найдено более 40 гладких чисел. Поместим найденные гладкие числа в табл.3.

Таблица 3

Гладкие числа, найденные на этапе предпросеивания

$(a,\ b)$	$(a,\ b)$	$(a,\ b)$	$(a,\ b)$	$(a,\ b)$	$(a,\ b)$	$(a,\ b)$
(73,1)	(2,1)	(1,1)	(2,1)	(-3,1)	(-4,1)	(-8,1)
(-13,1)	(-14,1)	(-15,1)	(-32,1)	(-56,1)	(-61,1)	(-104,1)
(-116,1)	(5,2)	(-3,2)	(-25,2)	(-33,2)	(8,3)	(-2,3)
(-17,3)	(-19,4)	(-48,5)	(-54,5)	(-313,5)	(43,6)	(8,7)
(-11,7)	(-38,7)	(-44,9)	(-4,11)	(-119,11)	(-856,11)	(-536, 15)
(-5,17)	(-5,31)	(-9,32)	(202,43)	(-24,55)		

Этот этап называется предварительным просеиванием, поскольку он позволяет только найти *номера* гладких пар, а не сами разложения (хранить вектора разложения для всех чисел из интервала просеивания является неэффективным). Для нахождения самих векторов разложений нужно выполнить повторное просеивания только уже не по области просеивания, а по множеству найденных номеров гладких чисел. В результате будут найдены представления каждого из чисел $F_1(a,b)$ и $F_2(a,b)$ уже в виде произведения элементов соответствующих факторных баз.

Вторичное просеивание для $n = 45113$

Выполнив вторичное просеивание найдем множество гладких пар M. Рассмотрим пример элемента такого множества.

$$F_2(8,3) = 8-3m = -85 = -2^0 \cdot 3^0 \cdot 5^1 \cdot 7^0 \cdot 11^0 \cdot 13^0 \cdot 17^1 \cdot 19^0 \cdot 23^0 \cdot 29^0,$$

что соответствует вектору разложения

$$v(8,3) = (1,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0)$$

Первая координата отражает знак числа и равно 0 для положительных $a - bm$, и 1, для отрицательных.

Аналогично для алгебраической факторной базы

$Nr(a,b) = F_1(8,3) = 5696 = 2^6 \cdot 89^1$.

Полная длина вектора разложения равна 24.

Формирование системы линейных уравнений

В предыдущем разделе была описана процедура просеивания, в результате которой было получено множество S гладких пар (a,b), а также разложения значений $F_1(a,b)$ и $F_2(a,b)$ в произведения элементов факторных баз FB_1 и FB_2.

Для формирования системы линейных уравнений нам также понадобится вычислить значение квадратичного характера для каждой пары $(a,b) \in S$ и каждой пары (p,r) из квадратичной факторной базы. Это означает, что для всех $(a,\, b) \in S$, $(p,r) \in FB_3$, мы должны вычислить символ Лежандра

$$\left(\frac{a-br}{p}\right). \tag{5.137}$$

Теперь можно сформировать матрицу системы. Она будет содержать $|S|$ строк по числу найденных гладких пар. Каждая строка содержит $k = s_1 + s_2 + s_3 + 2$ компоненты, принадлежащие множеству $F_2 = \{0,1\}$. Здесь s_i–положительное число, обозначающее число элементов в факторной базе FB_i.

Вектор, соответствующий паре (a,b), состоит из следующих компонент. Первая компонента соответствует знаку числа $a - bm$– положительному числу отвечает 0, а отрицательному –1. Далее идет блок из s_1 элементов, соответствующий разложению $a - bm$ по FB_2. Все элементы этого вектора заменятся их остатками по модулю 2. Дальше идут компоненты вектора разложения числа $F_1(a,b)$ по факторной базе FB_1, взятые по модулю 2. Наконец, последний блок из s_3 элементов соответствует

значениям символа Лежандра, вычисленным по формуле (5.137). Отметим, что компонента для знака $F_1(a,b)$ отсутствует, поскольку полученное произведение будет полным квадратом.

Решение системы линейных уравнений

Само решение системы не отличает от решения в случае метода квадратичного решета (см.разд.4.4). Опять можно выполнить либо использование стандартной процедуры исключения переменных методом Гаусса, либо использовать гораздо более сложный метод Лантцоша (см.[38]). Мы не будем здесь останавливаться на особенностях решения системы, а сделаем предположение о том, что решение уже найдено. Выпишем такое решение для нашего примера $n = 45113$, взятого из диссертации М.Бриггса [12].

Таблица 4

Элементы множества M, дающие решение системы

$(a,\ b)$	$(a,\ b)$	$(a,\ b)$	$(a,\ b)$	$(a,\ b)$
(1,1)	(-104,1)	(8,3)	(43,6)	(-856,11)
(-3,1)	(-3,2)	(-48,5)	(8,7)	
(-13,1)	(-25,2)	(-54,5)	(-11,7)	

Обозначим через $\gamma(x)$ многочлен, равный произведению всех многочленов $a - bx$, где пара (a,b) принадлежит найденному множеству решений S:

$$\gamma(x) = \prod_{x\in S}(a-bx) = 2051543129764485x^2+15388377355799440x+$$
$$+24765692886531904 \qquad (5.138)$$

Сделаем теперь несколько замечаний о найденном решении. Пока что мы можем только утверждать, что норма многочлена $Nr(\gamma(x))$ является полным квадратом. Очевидно, если многочлен $\gamma(x) = \alpha^2(x)$ является

полным квадратом, то его $Nr(g(x))$ будет полным квадратом в кольце целых чисел $\mathbf{Z}$. Однако обратное соотношение верно не всегда.

Пример. Рассмотрим числовое поле $Q[\sqrt{3}]$, получаемое при рассмотрении неприводимого полинома $f(x) = x^2 - 3$. Элемент $v = 2 + \sqrt{3}$ принадлежит $Q[\sqrt{3}]$ и имеет координаты $(2, 1)$ в базисе $B = (1; \sqrt{3})$. Его норма, определяемая по формуле (А.169), равна $Nr(v) = v_1^2 - 3v_2^2 = 1$. Таким образом, норма элемента является полным квадратом в Z, а сам полином v не является квадратом в $Q[\sqrt{3}]$, т.к. $v = w^2$ для $w = (\sqrt{6} + \sqrt{2})/2$, и $w \notin Q[\sqrt{3}]$. Рассмотрим причины, по которым это может произойти:

1. Первая причина состоит в том, что теорема об однозначном разложении (теор. А.7) выполняется в кольце целых алгебраических чисел $\mathbf{Z}_K$ поля $Q[\theta]$, а мы работает в меньшем кольце полиномов с целыми коэффициентами $\mathbf{Z}[\theta]$ от переменной θ, являющемся в общем случае собственным подмножеством кольца $\mathbf{Z}_K$.

Пример. В поле $K = Q[\sqrt{5}]$ элемент $g = (-1 + \sqrt{5})/2$ является целым, т.к. он является корнем минимального многочлена $f(x) = x^2 + x - 1$, однако не принадлежит кольцу $\mathbf{Z}[\sqrt{5}]$.

К счастью, эту причину можно устранить, воспользовавшись следующей теоремой:

Теорема 5.3. *Пусть $K = Q[\theta]$ – алгебраическое числовое поле, полученное присоединением к Q корня неприводимого в Q многочлена $f(x)$. Тогда для любого полинома $\alpha(x)$, принадлежащего кольцу $\mathbf{Z}_K$, многочлен $h(x) = \alpha(x) \cdot f'(x)$ принадлежит кольцу $\mathbf{Z}[\alpha]$, где $f'(x)$ –производная многочлена $f(x)$.*

Значит, чтобы устранить неполноту кольца $\mathbf{Z}[\theta]$ в GNFS, достаточно домножить многочлен $\gamma(x) = \prod_{x \in S}(a - b\alpha)$ на квадрат производной $f_1(x)$:

$$g(x) = (f_1'(x))^2 \cdot \gamma(x). \tag{5.139}$$

Вычислим значение многочлена $g(x)$ для нашего базового примера $n = 45113$:

$$g(x) = (f_1'(x))^2 \cdot \gamma(x) = (3x^2 + 30x + 29)^2 \cdot \prod_{x \in S}(a - bx) \equiv$$

$$\equiv 22455983949710645412x^2+54100105785512562427x+22939402657683071224 \pmod{(x^3+15x^2+29x+8)} \tag{5.140}$$

2. Вторая причина того, что многочлен $\gamma(x)$ может не быть полным квадратом, заключается в существовании нетривиальных элементов кольца $\mathbf{Z}(\theta)$, имеющих норму, равную 1 или -1.

Элемент v кольца $Z[\theta]$ числового поля $Q[\theta]$ называется *единицей*, если его норма $Nr(v)$ равна $\pm$. Например, элемент $v = 2+\sqrt{3}$ в поле $Q[\sqrt{3}]$ имеет норму 1 и является единицей. Множество единиц алгебраического числового поля образует группу по умножению (см.[62], гл.2, §4). Строение этой группы описывается теоремой Дирихле ([62], с.133), которая утверждает, что любой ее элемент ε является произведение вида

$$\varepsilon = \zeta \cdot \varepsilon_1^{r_1} \cdot ... \cdot \varepsilon_k^{r_k}, \tag{5.141}$$

где ζ – некоторый корень из 1, а набор порождающих ε_1, ..., $\varepsilon_1^{r_k}$ содержит $k = s+t-1$ элементов, s и t – число действительных и пар мнимых корней многочлена $f(x)$.

Таким образом, если норма многочлена $g(x)$ является полным квадратом в Z, то $g(x)$ является произведением квадрата $\alpha^2(x)$ на многочлен $h(x)$ с нормой, равной 1. Построить базис множества единиц в произвольном числовом поле является задачей нетривиальной, поэтому в статье ([13]) была предложена идея введения *квадратичных характеров*. Рассмотрим эту идею более подробно.

Напомним, что символ Лежандра позволяет определить, является ли произвольный элемент конечного поля квадратичным вычетом, т.е. квадратом другого элемента. В нашем случае если многочлен $\gamma(x) = \prod_{x\in S}(a - bx) = \alpha(x)^2$ является полным квадратом, то выполняется следующая теорема:

Теорема 5.4. *Если идеал первого порядка $c-dx$, определяемый парой целых чисел (q, s), не делит никакой идеал $a - bx$, входящий в произведение*

элементов M, *и* $f'(s) \not\equiv 0 \pmod q$, *тогда*

$$\prod_{x \in S} \left(\frac{a - bs}{q} \right) = 1. \tag{5.142}$$

Если произведение $\prod_{x \in S}(a - bx)$ не является полным квадратом, то найдется пара чисел (q, s), удовлетворяющая условию теоремы, для которой символ Лежандра $(a - bs/q)$ окажется равным 0. Квадратичная факторная база, состоящая из подобных пар (q, s), как раз и предназначена для того, чтобы предотвратить такую возможность. Поскольку мы не можем проверить условие (5.142) на бесконечном множестве пар, то существует вероятность, что найденное в ходе просеивания множество S, не дает полный квадрат, но с ростом размера квадратичной факторной базы, такая вероятность стремится к нулю.

Таким образом, при достаточном большом размере факторной базы квадратичных характеров, единицы поля $K[\theta]$ будут входить в многочлен $g(x)$ только в четной степени.

3. Третьей причиной, по которой $g(x)$ может не быть полным квадратом, заключается в том, что теорема об однозначном разложении целых алгебраических чисел выполняется в кольце идеалов кольца Z_K, а идеалы Z_K не всегда являются главными, т.е. не могут порождаться элементом вида $a - b\theta$, $a, b \in \mathbf{Z}$.

5.4. Вычисление квадратного корня

Как только множество M пар целых чисел (a, b), образующих решение, найдено, необходимо вычислить многочлен $\alpha(x)$, квадрат которого равен

$$g(x) = (f'(x))^2 \cdot \prod_{x \in S}(a - bx) \tag{5.143}$$

Заметим, что подобной задачи не было ни в одном из предыдущих методов факторизации, включая метод квадратичного решета, и она появляется только в методе решета числового поля. При небольших значений коэффициентов вычисление квадратного корня в кольце

полиномов не представляет особой трудности, и выполняется примерно такими же методами, как вычисление квадратного корня в кольце целых чисел или конечных полях, например, с использование алгоритма Шенкса–Тоннелли (см.разд.1.14).

В случае рекордных значений факторизуемых чисел, значения коэффициентов полиномов достигают нескольких миллионов цифр, и задача извлечения корня становится чрезвычайно сложной и трудоемкой задачей, сравнимой по затратам ресурсов с задачей просеивания. Если же степень полинома просеивания становится больше 7, то время, затрачиваемое на вычисление корня, превышает время просеивания. Поэтому нахождение эффективного алгоритма вычисления квадратного корня играет чрезвычайно важную роль в решете числового поля. Мы рассмотрим здесь два основных метода, используемых в практике GNFS.

Вычисление корня путем редукции к подполям

Это метод был предложен в работе 1993 г. Жаном Ковейном (Jean M. Couveignes) [20] (см.также П.Монгомери [39]). Суть метода состоит в том, что вычисление квадратного корня выполняется в конечных полях по нескольким различным простым модулям p. Если рассматриваемое p является инертным, т.е. многочлен $g(x) \bmod p$ (т.е. многочлен, полученный из $g(x)$ путем замены всех его коэффициентов их остатками по модулю p) неприводим в поле F_p, и $\alpha_p^2(x) = g(x) \bmod p$, тогда коэффициенты искомого корня из полинома $g(x)$ совпадут с коэффициентами $\alpha_p(x)$ по модулю p. Вычислив несколько таких частичных корней по различным модулям, можно потом восстановить полный полином $\alpha(x)$ с помощью китайской теоремы об остатках. Число различных p должно быть таким, чтобы произведение всех этих p превышало максимальный по абсолютной величине коэффициент полинома $\alpha(x)$. Ниже мы дадим пример вычисления корня из полинома по методу Ковейна для $n = 45113$.

Метод Ковейна работает только для нечетных степеней d полинома $f_1(x)$. Кроме того, в редких случаях возможна ситуация, когда для многочлена нет подходящих простых чисел p, для который $g(x) \bmod p$

неприводим. Для такого случая метод Ковейна также не будет работать. В статье Бухлера, Ленстры и Померанса [13] приведены возможные варианты вычисления корня в этом случае.

Вычисление корня методом подъема Гензеля

Идея подъема Гензеля (Hensel's Lifting) уже использовалась в нашей книге при построении факторной базы в методе квадратичного решета для вычисления корней полинома просеивания $q(x)$ по степени p^{k+1}, если корни $q(x) \equiv 0 \,(\bmod\, p^k)$ известны.

Первоначально выбирается простой модуль p, для которого многочлен $g(x) \bmod p$ является неприводимым. Так же, как и в методе Ковейна, сначала вычисляется корень $\alpha_1(x)$ из $q(x) \,(\mathrm{mod} p)$. Предположим, что найден корень $\alpha_k(x)$ из $q(x) \,(\bmod p^k)$ и требуется найти корень $\alpha_{k+1}(x)$ по модулю p^{k+1}. Коэффициенты $\alpha_{k+1}(x)$ сравнимы с коэффициентами $\alpha_k(x)$ по модулю p^k. Отсюда, каждый коэффициент $a_i, k+1$ имеет вид $a_i, k+1+t_i \cdot p$, где t_i–целое число из интервала $[0;\, p-1]$. Подставляя выражения для всех $a_i, k+1$ в уравнение $\alpha_{k+1}^2(x) = g(x) \bmod p^{k+1}$, получим уравнения, позволяющие найти все значения t_i, а зная их, вычислить коэффициенты многочлена $\alpha_{k+1}(x)$. Этот метод применим для любой степени многочлена $f_1(x)$, однако при увеличении степени k коэффициенты $\alpha_k(x)$ достигают огромных значений, что делает вычисления новых полиномов $\alpha_{k+1}(x)$ очень трудоемким.

5.5. Пример вычисления квадратного корня и оценка его сложности

В этом разделе рассмотрим пример вычисления корня из многочлена 5.140, определенного ранее для $n = 45113$, методом редукции к конечным подполям. Этот алгоритм состоит из трех основных шагов.

1. Выбор подходящих простых p

Первая часть алгоритма состоит в выборе подходящих p, для которых многочлен $g_p(x) = f_1(x) \bmod p = x^3 + 15x^2 + 29x + 8 \bmod p$ является неприводимым в поле F_p. Выберем значение p равным 9929. Для проверки неприводимости полинома $g_p(x)$ воспользуемся следующими двумя теоремами (напомним, что унитарным называется многочлен со старшим коэффициентом равным 1):

Теорема 5.5. *Пусть $q = p^d$–положительная степень простого числа. Многочлен*

$$w_q(x) = x^{p^d} - x \tag{5.144}$$

является произведением всех унитарных неприводимых многочленов над Z/pZ, степень которых делит d.

Теорема 5.6. *Пусть p–простое число. Многочлен $f(x)$ степени d над Z/pZ является неприводимым тогда и только тогда, когда выполнены следующие два условия:*

1. *$x^{p^d} - x$ делится на $f(x)$,*
2. *Н.О.Д.$(x^{p^{d/p_i}} - x,\ f(x)) = 1$ для всех простых p_i, делящих d.*

1. Первым шагом алгоритма является вычисление $h_p(x) = f_1(x) \bmod p$. В нашем примере, $h_p(x)$ совпадает с $f_1(x)$.

2. Проверим условие $h_p(x) \mid x^{9929^3} - x$. Это условие проверяется непосредственным делением $x^{9929^3} - x$ на $h_p(x)$. Условие выполнено.

3. Проверим условие 2 теоремы 5.6 для единственного нетривиального простого делителя $p_1 = 3$. Надо проверить, что Н.О.Д. $(x^{9929}, h_p(x)) = 1$. Для этого вычислим

$$(x^{9929} - x) \bmod h_p(x) = 7449x^2 + 4697x + 5984$$

и проверим, что Н.О.Д. $(7449x^2 + 4697x + 5984,\ x^3 + 15x^2 + 29x + 8)$ равен 1.

Таким образом, число $p = 9929$ подходит для вычисления корня. Таким же способом проверим еще два простых числа $p = 9851$ и $p = 9907$.

Примером того, что не каждое простое p удовлетворяет теореме 5.6, является число $p = 9923$. Н.О.Д. $(x^p - x,\ \ h_p(x)) = x - 847 \neq 1$.

2. Вычисление квадратного корня по алгоритму Шенкса-Тоннелли

Во второй части алгоритма мы должны вычислить квадратные корни из $g_p(x) = g(x) \bmod p$ для каждого простого p, отобранного в первой части работы алгоритма ($g(x)$ определен формулой 5.140). Приведем это вычисления для $p = 9929$:

1. $a(x) = g(x) \mod 9929 = 2027x^2 + 3891x + 6659$.

2. Вычислим $q = p^3 = 978\,850\,872\,089$. Разложим $q - 1$ в произведение четного и нечетного чисел $q - 1 = 2^3 \cdot 122\,356\,359\,011$, откуда $r = 3$, $s = 122\,356\,359\,011$. Используем такие же обозначения, как в (1.14).

3. Найдем квадратичный невычет в кольце многочленов по модулю $g_p(x)$ с коэффициентами из Z/pZ. Таким невычетом является $z(x) = x + 1$. Для проверки вычислим $z(x)^{(q-1)/2} \pmod{h_p(x)} = 9928 \equiv -1 \pmod p$.

4. Вычислим $y(x) = (x + 1)^s \pmod{h_p(x)} = 1273$.

5. Вычислим $\lambda_0(x) = (a(x))^s \pmod{h_p(x)} =$

   $= (2027x^2 + 3891x + 6659)^{122\,356\,359\,011} \pmod{x^3 + 15x^2 + 29x + 8} =$
   $= 9928 \equiv -1 \pmod{9929}$.

6. Вычислим $w_0 = (a(x))^{(s+1)/2} =$

   $= (2027x^2 + 3891x + 6659)^{61\,178\,179\,506} \pmod{(x^3 + 15x^2 + 29x + 8)} =$
   $= 2124x^2 + 5715x + 4075$.

7. Поскольку $\lambda_0^2 \equiv 1 \pmod{9929}$, то порядок λ_0 равен 2, откуда $m = 1$. Степень $k = 2^{d-m} = 4$.

8. Вычислим λ_1 и w_1:

$$\lambda_1 = \lambda_0 \cdot y^k (\mathrm{mod}\, h_p(x)) = 1$$
$$w_1 = w_0 \cdot y^{k-1} (\mathrm{mod}\, h_p(x)) = 6527x^2 + 8769x + 6852.$$

Поскольку $w_1 = 1$, то вычисление закончено. Искомый корень из $g_p(x) = 2027x^2 + 3891x + 6659$ равен $6527x^2 + 8769x + 6852$.

Наконец последним шагом является вычисление корня $x_p = g_p(x) \mod p = (2027 \cdot 31^2 + 3891 \cdot 31 + 6659) \mod 9929 = 5694$.

3. Нахождение полного корня с помощью китайской теоремы об остатков

В предыдущей секции был вычислен частичный квадратный корень из произведения линейных многочленов $a - bx$ для $x \in M$ по модулю $p = 9929$ и найдено его значение при $x = m$.

Выполним такую же операцию по оставшимся модулям $p = 9851$ и $p = 9907$. Поскольку это вычисление полностью повторяет вычисление предыдущей секции, мы его пропустим. Выпишем окончательные значения:

p	$g_p(x)$	$g_p(m)$
9851	$7462x^2 + 5679x + 4037$	5694
9907	$5126x^2 + 5072x + 3125$	4152
9929	$3402x^2 + 1160x + 3125$	3077

Теперь надо вычислить полное значение корня $g(m) \mod n$, используя значения $g(m)$ по частичным модулям 9851, 9907 и 9929. Иначе говоря, надо решить систему сравнений:

$$\begin{cases} x \equiv 5694 \,(\mathrm{mod}\, 9851), \\ x \equiv 4152 \,(\mathrm{mod}\, 9907), \\ x \equiv 3077 \,(\mathrm{mod}\, 9929). \end{cases}$$

Для решения этой задачи воспользуемся китайской теоремой об остатках (разд. 1.15).

Применяя эту теорему, найдем значение корня $x^* = g(m) \bmod n = 694683807559 \bmod 45113 = 43992$.

Недостатком этого метода вычисления квадратного корня из полинома является его ограничения, заключающиеся в том, что он применим только в случае полинома $f_1(x)$ нечетной степени и выбор соответствующей системы простых чисел, по которым $f_1(x) \bmod p$ является неприводимым, не всегда возможен.

В работе Бухлера, Ленстры и Померанса ([13]) дается оценка сложности алгоритма вычисления квадратного корня. Она зависит от выбранного алгоритма и техники реализации операций с длинными числами. Если операции умножения полиномов и вычисления остатков по модулю $f(x)$ выполнены с помощью дискретного преобразования Фурье, то время вычисления квадратного корня определяется с помощью следующей оценки:

$$T(n) = y^{1+o(1)}, \tag{5.145}$$

где y–верхняя граница для параметров a, b области просеивания, зависящая от числа n и степени d многочлена $f_1(x)$. Ее оптимальное значение равно

$$\log y = \left(\frac{1}{2} + o(1)\right)\left(d\log s + \sqrt{(d\log d)^2 + 4\log(n^{1/d})\log\log(n^{1/d}).}\right) \tag{5.146}$$

В указанной работе дается описание еще одного алгоритма вычисления квадратного корня, который позволяет вычислять корень путем последовательного вычисления системы вспомогательных полиномов $(\mu_i(s), \nu_i(s))$. Пусть $S = \{(a_i, b_i)\}_{i=1}^{k}$–решение системы, тогда определим систему многочленов (μ_i, ν_i) следующим образом:

$$(\mu_0(s), \nu_0(s) = (1,\ 1)).$$

$$\mu_i = \begin{cases} \mu_{i-1}/(a - b\theta), \text{ если } (a - b\theta) \mid \mu_{i-1}, \\ \mu_{i-1} \cdot (a - b\theta), \text{ иначе.} \end{cases} \tag{5.147}$$

$$\nu_i = \begin{cases} \nu_{i-1} \cdot (a - b\theta), \text{ если } (a - b\theta) \mid \mu_{i-1}, \\ \nu_{i-1}, \text{ иначе.} \end{cases} \tag{5.148}$$

Теперь выражений для многочлена $g(x)$ (5.143) получит вид

$$g(\alpha) = (f'(\alpha))^2 \cdot \prod_{x \in S} (a - b\alpha) = (f'(\alpha))^2 \mu_s \nu_s^2 \tag{5.149}$$

и теперь вместо извлечение квадратного корня из $g(x)$ достаточно вычислить корень из μ_s и перемножить три полинома $f'(x)$, $\sqrt{\mu_s}$ и ν_s.

Авторы этого метода заметили, что практическая польза от этой идеи вычисления корня полностью не ясна, и окончательный вердикт должен быть вынесен после ее реализации.

4. Завершение вычисления

Теперь мы можем завершить полную задачу факторизации числа $n = 45113$.

Вычислим сначала произведение

$$y^2 = f_1'(m)^2 \cdot \prod_{x \in S} (a - bm) = (3 \cdot 31^2 + 30 \cdot 31 + 29)^2 \cdot (-1+31)(3+31)(13+31)(104+31) \cdot$$

$$(3+2 \cdot 31)(-8+3 \cdot 31)(48+5 \cdot 31)(54+5 \cdot 31)(-43+6 \cdot 31)(-8+6 \cdot 31)(-8+7 \cdot 31)(11+7 \cdot 31) \cdot$$

$$(856+11 \cdot 31) = 3842^2 \cdot 31746503388600^2 \mod n,$$

откуда $y = 3824 \cdot 31746503388600 \mod 45113 = 15160$. Пара чисел $(x, y) = (43992, 15160)$ удовлетворяет уравнению $x^2 \equiv y^2 \mod n$, откуда $x^2 - y^2 = (x + y)(x - y) = (43992 + 15160)(43992 - 15160) = 59152 \cdot 28832$.

Вычисляя теперь с помощью алгоритма Евклида наибольший общий делитель Н.О.Д.$(n,\, x \pm y)$, найдем делители $n = 45113$:

Н.О.Д.$(n, x + y) =$ Н.О.Д.$(45113, 59152) = 229$,

Н.О.Д.$(n, x - y) =$Н.О.Д.$(45113, 28832) = 197$.

5.6. Оценка сложности решета числового поля

Общее время работы алгоритма решета числового поля зависит от времени работы составляющих его частей, из которых наибольший вес имеют время первичного просеивания, в ходе которого ищутся номера

гладких пар для составления системы линейных уравнений и время вычисления квадратного корня из полинома в пространстве $Z[x]/(f_1(x))$. Все остальные составляющие алгоритма влияют значительно меньше на производительность GNFS.

Приведем оценку метода решета числового поля, взятую из ([13]) на основе функции $L_n(\alpha; c)$, определенной на с.130. Эта оценка выполнена при условии, что степень d и граница области просеивания y выбраны, как указано ниже:

$$d = \left(3^{1/3} + o(1)\right) (\log n / \log\log n)^{1/3},\ n > d^{2d^2} > 1,$$
$$u = y = L_n\left(1/3, (8/9)^{1/3} + o(1)\right). \tag{5.150}$$

При выполнении условий (5.150) и времени вычисления корня (5.146) время работы алгоритма решета числового поля оценивается величиной

$$T(n) = L_n\left(1/3, (64/9)^{1/3} + o(1).\right) \tag{5.151}$$

Отметим, что приближенное значение константы $(64/9)^{1/3}$ равно $1,92$. Таким образом, уменьшение значения показателя степени в наиболее важном сомножителе $\log n$ функции $L_n(\alpha; c)$ от значения $1/2$ в методе квадратичного решета до $1/3$ в методе решета числового поля дает тот прогресс, который обеспечивает приоритет этого метода над методом квадратичного решета и всеми остальными методами факторизации, известными на сегодняшний день.

5.7. Улучшение алгоритма выбора полиномиальной пары

Основную часть расчета в методе GNFS занимает процедура просеивания. Эффективность этой процедуры зависит от величины коэффициентов многочленов $f_1(x)$ и $f_2(x)$, а также наличия у них корней по модулю небольших простых чисел. Множество подходящих пар определяется путем вариации значения параметра m и добавления к

полиному $f_1(x)$ произвольной комбинации вида $g(x) \cdot f_2(x)$, где степень $g(x) \leq d-1$.

Выбор подходящих m и $g(x)$ в базовом алгоритме GNFS можно выполнить за сравнительно небольшое время в итерационной процедуре (см.Murphy [41]). Этот выбор ограничен многочленами $f_1(x)$ и $f_2(x)$, имеющими старший коэффициент, равный 1 (такие многочлены называются *унитарными*).

Значительно больший набор возможностей появляется, если рассматривать произвольные не унитарные многочлены $f_1(x)$ и $f_2(x)$. Впервые возможность использования не унитарных полиномов описана в статье 1993 г. Бухлера, Ленстры и Померанса [13], где были предложены три возможных варианта расширения множества применяемых полиномов. Опишем эти варианты:

1. Старший коэффициент c_d полинома $f_1(x)$–произвольное целое число

Если взять равным, например, произведению небольших простых сомножителей, то однородный многочлен $F_1(a,b)$ будем делиться на p для каждого делителя p коэффициента c_d и b, кратного p. Это способствует увеличению количества гладких пар в области просеивания. Проанализируем, что потребуется теперь изменить в наших формулах.

Пусть $c_d \neq 1$, и $\theta \in \mathbf{C}$–корень $f_1(x)$. Тогда $\beta = c_d \cdot \theta$—алгебраическое целое. Действительно, β—корень многочлена с целыми коэффициентами

$$H(x) = c_d^{d-1} f_1(x) = x^d + c_{d-1}x^{d-1} + c_d\, c_{d-2}x^{d-2} + \ \ldots\, c_d^{d-1}\, c_0.$$

Отсюда, если S-множество пар (a,b) такое, что $\prod_{(a,b)\in S}(a - b\theta)$—квадрат в $\mathbf{Q}(\alpha)$ и S содержит *четное* число пар, тогда

$$(H'(c_d\,\theta))^2 \cdot \prod_{(a,b)\in S} (ac_d - bc_d\theta)$$

является квадратом в $\mathbf{Z}[c_d\theta]$, например, γ^2. Преобразуем множитель $H'(c_d x)$

в выражении для $\gamma^2(x)$:

$$H'(c_d x) = c_d^{d-1} \cdot f_1'(x) = c_d^{d-1} \cdot F_1'(x,1) = \frac{1}{c_d} \cdot F_1'(x, c_d) = \frac{1}{c_d} \cdot \sum_{i=1}^{d-1} i \cdot c_i \, x^i \, c_d^{d-1-i},$$

откуда выражение для многочлена $\gamma^2(\theta)$ получит окончательный вид:

$$\gamma^2(\theta) = \frac{1}{c_d^2} \cdot (F_1'(x, c_d))^2 \cdot \prod_{(a,b) \in S} (a c_d - b c_d \theta). \tag{5.152}$$

Найдем целые коэффициенты многочлена $\gamma(x) = \sum_{i-0}^{d-1}$ согласно описанному раньше алгоритму вычислению квадратного корня, причем базисом будет служить теперь множество $\{1, c_d\theta, \ldots (c_d\theta)^{d-1}\}$. Замена θ на $c_d\, m$ в многочлене $g(x)$ позволяет найти делитель числа n также, как и в случае унитарного многочлена $f_1(x)$. Ниже мы опишем необходимые изменения в базовом алгоритме GNFS для наиболее общего варианта 3.

2. Замена условия $f_1(m) = n$ условием $F_1(m_1, m_2) = n$

Поскольку $f_1(m_1) = F_1(m_1, 1)$, то при $m_2 = 1$ этот вариант совпадает со стандартным. Этот вариант дает возможность увеличить количество допустимых полиномиальных пар, среди которых можно выбрать полиномы с меньшими по модулю значениями коэффициентов. В статье [13] предложен следующий вариант выбора m_1, m_2:

Полагаем $c_d = 1$. Далее выберем $m_1 \approx n^{1/(d+1)}$ так, чтобы разность $n - m_1^d$ имела делитель $m_2 \approx n^{1/(d+1)}$, выполняя пробное деление или факторизацию $n - m_1^d$ с помощью эллиптических кривых. Далее раскладываем $(n - m_1^d)/m_2$ по степеням m_1, m_2:

$$\frac{n - m_1^d}{m_2} = c_{d-1} m_1^{d-1} + \ldots + c_1 m_1 m_2^{d-2} + c_0 m_2^{d-1}. \tag{5.153}$$

3. Алгоритм Кляйнъюнга

В упомянутой статье [13] была упомянута возможность использования общего варианта, при котором оба параметра c_d не m_2 равны 1, но было предложено никакого алгоритма для получения такого представления,

однако, очевидно, одновременная вариация двух параметров дает большую возможность в выборе хорошей полиномиальной пары. Этот вариант построения полиномиальной пары получил развитие в более поздних работах. Мы опишем вариант алгоритма построения полиномов f_1 и f_2, предложенный в 2006 г. Т. Кляйнъюнгом [29]:

1. Выбираем параметр m_2, $1 \le m_2 \ll n^{1/d}$.

В дальнейшем для совмещения обозначения со статьей Кляйнъюнга будем использовать вместо m_2 обозначение p. Значение p выбирается равным произведению нескольких небольших простых чисел p_i, удовлетворяющих условию $p_i \equiv 1 \pmod{d}$. Вместо m_1 будем писать просто m.

2. Выбираем старший коэффициент a_d и $m \approx (n/a_d)^{1/d}$ так, чтобы выполнялась конгруэнтность

$$a_d \cdot m^d \equiv n \pmod{p} \tag{5.154}$$

Кляйнъюнг предлагает выбирать c_d взаимно-простым с p, тогда уравнение $c_d\, x^d \equiv n \pmod{p}$ либо не имеет ни одного решения, либо имеет d решений.

3. Обозначим $r_d = n$. Далее будем вычислять параметры r_i, c_i для $d > i \ge 0$ по реккурентным формулам:

$$r_i = \frac{r_{i+1} - c_{i+1}\, m^{i+1}}{p}, \quad c_i = \frac{r_i}{m^i} + \delta_i, \tag{5.155}$$

где $0 \le \delta_i < p$ выбирается так, чтобы выполнить условие $r_i \equiv c_i\, m^i \pmod{p}$. Тогда, для всех i выполнится условие

$$r_i = \sum_{j=0}^{i} c_j\, m^j\, p^{i-j}, \tag{5.156}$$

а при $i = d$ выполнится $r_d = n = p^d \cdot f_1(m/p) = F_1(m, p)$.

Полином $f_1(x)$ оказывается определенным в этой конструкции, а значение второго полинома $f_2(x)$ определяем равным $px - m$. Полиномиальная пара определена. Этим заканчивается описание алгоритма Кляйнъюнга.

Практическая реализация алгоритма Кляйнъюнга

В практической реализации алгоритма выбирается сначала параметр p как произведение нескольких средних простых чисел p_i, сравнимых с 1 по модулю d. Потом следует выбрать пару (c_d, m) в ходе итерационной процедуры изменения параметра c_d с некоторым шагом.

В рекордном разложении 512-битового числа RSA [15] степень полинома была выбрана равной $d = 5$, параметр p равным произведению 7 простых чисел и вспомогательного множителя p_0, а шаг перебора c_d равным 60.

Для каждого значения c_d решается сравнение (5.154). Если оно разрешимо, то находится d решений x и рассматривается d соответствующих значений для m таких, что разность $n - a_d\, m^d$ имеет наименьшие значения по абсолютной величине. Вычисляется коэффициент $c_{d-1} = (n - a_d\, m^d)/p$, и оставляются только те решения, для которых значение $|c_{d-1}|$ меньше некоторой наперед выбранной границы.

Выбрав c_{d-1}, можно вычислить c_{d-2} с помощью реккурентных формул (5.155). Здесь также есть небольшой выбор при выборе δ_i. Если старшие коэффициенты оказались большими по абсолютной величине, то нет смысла их рассматривать, поэтому такие полиномы сразу отбрасываются, и выполняется приращение очередной коэффициента c_d.

Лучшим считается полином, у которого все коэффициенты принимают небольшие по модулю значения, и кроме того, полином имеет много корней по модулю небольших простых чисел. Последний фактор учесть достаточно сложно, поскольку он достаточно случаен. В статье Кляйнъюнга [29] приводится интегральная мера для числа гладких пар, попадающих в область просеивания A:

$$\frac{6}{\pi^2} \int_A \rho\left(\frac{\log F_1(x,y) + \alpha_1}{\log B_1}\right) \rho\left(\frac{\log F_2(x,y) + \alpha_2}{\log B_2}\right) dxdy \qquad (5.157)$$

Здесь ρ обозначает функцию Дикмана-де Брюина (см.стр.129), коэффициент $6/\pi^2$ оценивает долю пар (a, b) со взаимно-простыми аргументами, а параметры α_i введены для учета корней полиномов F_1, F_2 по небольшим

простым модулям:

$$\alpha_i = \sum_{small\ p} \left(1 - r(F_i, p)\frac{p}{p+1}\right)\frac{\log p}{p-1} \quad \text{для } i \in \{1, 2\}, \tag{5.158}$$

Второй сомножитель в интеграле (5.157) мало зависит от выбора параметров полинома f_1, поэтому им можно пренебречь. Дальнейшее упрощение состоит в рассмотрении выражения

$$\alpha_1 + \frac{1}{2}\log\left(\int_A F_1^2(x, y)dxdy\right), \tag{5.159}$$

которое надо минимизировать в отличие от предыдущих мер.

Таким образом, в алгоритме Кляйнъюнга перебираются всевозможные полиномы $f_1(x)$, удовлетворяющие условию $n = p^d \cdot f_1(m/p)$ и ищутся тот, который доставляет минимум выражению (5.159). Приведем выражение для наилучшего полинома для числа RSA-512, найденное по этому алгоритму:

$$\begin{aligned} f_1(x) = {} & 498520x^5 + 15578368316860x^4 - 5137488762804904 87x^3 - \\ & -10211574130795357032973 44x^2 - 39893111467231678678251 29900x + \\ & +14658919460374074323550710377995600, \\ f_2(x) = {} & 8794555574829559x - 29394756538965034296055627 0613. \end{aligned}$$

В качестве критики этого подхода отметим, что указанная мера зависит от области просеивания, вид которой существенно влияет на количество генерируемых гладких пар. Например, вдоль прямой $x_0 = b/a$, где x_0— действительный корень полинома число гладких пар значительно возрастает. В статье Ш.Т.Ишмухаметова, Д.Б.Зиятдинова и Р.Г.Рубцовой [68] предлагается другой подход к оценке полиномов.

Изменения в базовом алгоритме GNFS

Обсудим изменения, которые следует внести в базовый алгоритм GNFS, описанный на стр.146, при изменении параметров c_d и $m_2 = p$:

1. Эти изменения касаются, в первую очередь, однородных многочленов $F_1(a, b)$ и $F_2(a, b)$, которые в новом варианте получают вид :

$$F_1(a, b) = c_d a^d + c_{d-1}a^{d-1}b + ... + c_0 b^d \qquad F_2(a, b) = am_2 - bm_1. \tag{5.160}$$

2. Стадии просеивания, формирования системы линейных уравнений и построения множества S гладких пар (a, b) не изменятся.

3. Изменится формула (5.132) для определения $g^2(x)$, как полинома с целыми коэффициентами степени $d-1$:

$$g^2(\theta) = (f_1'(\theta)/c_d)^2 \cdot \prod_{(a,b)\in S} (c_d\, a - b\theta) \tag{5.161}$$

4. Извлечение квадратного корня $g(x)$ по модулю многочлена $f_1(x)$ остается прежним. Вычисляем коэффициенты $g(x)$ и число v:

$$g(x) = \prod_{i=0}^{d-1} b_i x^i, \quad v = \prod_{i=0}^{d-1} b_i m_1^i m_2^{d-1-i} \mod n. \tag{5.162}$$

5. Вычисляем число D^2 и C:

$$D^2 = \prod_{(a,b)\in S} (am_1 - bm_2\theta), \quad C = D \mod n. \tag{5.163}$$

6. Вычисляем числа $A = m_2^{\#S/2} \cdot v \mod n$ и $B = c_d^{d-2+\#S/2} \cdot C \mod n$.

7. Находим делители n, вычисляя Н.О.Д.$(n,\, A \pm B)$.

5.8. Заключение

Мы завершили описание основных алгоритмов вычислительной теории чисел, связанных с факторизацией. Нельзя считать эту тему закрытой, поскольку многие крупные открытия и новые достижения в этой области появляются снова и снова. Эти достижения обусловлены не только развитием вычислительной мощности компьютеров и сетей, но и развитием тех областей теоретической математики, которые до недавнего времени служили областью интересов лишь профессиональных математиков– специалистов в абстрактной алгебре и теории чисел.

Задачи элементарной теории чисел, вычислительные задачи криптографии подстегивает интерес большого круга лиц к математике и теории чисел, превращая их из дилетантов в профессионалы. Я надеюсь, что эта книга послужит решению этой важной задачи.

A. Приложение. Алгебраические числовые поля

В этой части мы дадим введение в теорию алгебраических числовых полей, необходимое для понимания метода решета числового поля, а также определения алгебраических понятий, использовавшихся в других алгоритмах. С общей теорией алгебраических числовых полей можно познакомится по учебникам Ван дер Вардена «Алгебра» [63], «Теория чисел» [62] З.И. Боревича и И.Р. Шафаревича, а также «Лекциям об алгебраических числах» Ю.В. Нестеренко [75], выложенным в Интернете. Из книг, изданных за рубежом, можно рекомендовать 3-е издание известного учебника «Алгебраические числовые поля и последняя теорема Ферма» И. Стюарта и Д. Толла, 2002 г., в котором дает ясное и наглядное описание всех трудных алгебраических понятий с сопровождением изложения многочисленными примерами и объяснениями.

Основные определения

Определение A.1. *Кольцом называется непустое множество элементов C, на элементах которого определены алгебраические операции сложения $+$, относительно которой C является абелевой группой с нейтральным элементом 0, и умножения $*$, относительно которой C является моноидом, т.е. выполняется свойство ассоциативности $a*(b*c) = (a*b)*c$. Операции сложения и умножения связаны законами дистрибутивности* $a * (b + c) = a * b + a * c, \quad (b + c) * a = b * a + c * a$.

Кольцо называется *коммутативным*, если операция умножения является коммутативной. Все рассматриваемые здесь кольца являются коммутативными.

Примерами колец могут служить кольцо целых чисел $\mathbf{Z}$, а также кольцо полиномов $Z[x]$ от одной переменной x с коэффициентами из Z. Эти кольца содержат единичный элемент 1, который обладает свойством нейтральности по умножению: для любого элемента a кольца выполняется $a * 1 = 1 * a = a$. Такие кольца называются кольцами с *единицей*. Другие кольца, например, кольцо $a\mathbf{Z}$, состоящее из всех кратных $a * k$, $k \in \mathbf{Z}$,

произвольного $a \neq 1$, не содержит единицы и называется кольцом *без единицы.* Напомним также определение поля.

Определение А.2. *Полем называется непустое множество элементов K, на элементах которого определены алгебраические операции сложения $+$ и умножения $*$,относительно которых K является коммутативными группами, связанные законами дистрибутивности.*

В произвольном поле любой ненулевой элемент a имеет обратные по сложению $-a$ и умножению a^{-1}, поэтому, в полях возможны все 4 арифметические операции – сложения, вычитания, умножения и деления. Примерами числовых полей являются поле рациональных чисел $\mathbf{Q}$, действительных чисел $\mathbf{R}$, комплексных чисел $\mathbf{C}$. Эти поля содержат бесконечное число элементов (поле $\mathbf{Q}$–счетно, все остальные поля имеют мощность континуума).

Существует также конечные поля, например, поле $GF_p = \{\overline{0}, \overline{1}, \overline{2}, ..., \overline{p-1}\}$ классов вычетов кольца $\mathbf{Z}$ по простому модулю p, а также поля GF_{p^k}, состоящие из p^k элементов, где p–простое число, а k–положительное целое число.

Любое коммутативное кольцо с единицей C может быть погружено в некоторое минимальное поле $K \supset C$, состоящее из элементов вида a/b, где $a, b \in C$. Такое поле называется полем частных кольца C. Примером может служить поле рациональных чисел Q, являющееся полем частных для кольца целых чисел Z.

Множество многочленов $K[x]$ с коэффициентами из поля K образует кольцо с обычными операциями сложения и умножения многочленов. Роль простых чисел в этом кольце играют неприводимые многочлены.

Определение А.3. *Многочлен $f \in K[x]$ называется* неприводимым *над K, если он не может быть разложен в произведение двух многочленов $h \in K[x]$ и $g \in K[x]$, отличных от константных многочлена.*

Модули над кольцом

Определение А.4. *Модулем M над кольцом R называется абелева группа с операцией умножения на элементы кольца R: $R \times M \to M$, которая удовлетворяет следующим условиям:*

1. $\forall m \in M, \forall r_1, r_2 \in R, (r_1 r_2)m = r_1(r_2),$
2. $\exists 1 \in M, \ \forall m \in M \ 1 \cdot m = m \cdot 1 = m,$
3. $\forall m_1, m_2 \in M, \forall r \in R, r(m_1 + m_2) = rm_1 + rm_2,$
4. $\forall m \in M, \forall r_1, r_2 \in R, (r_1 + r_2)m = r_1 m + r_2 m.$

Любое кольцо можно рассматривать как модуль над самим собой. Другим важным примером модуля является модуль $Z[x]$ многочленов над кольцом Z.

Модуль M над кольцом R называется конечнопорожденным, если найдется конечная совокупность элементов $a_1, a_2, ..., a_k \in M$ такая, что любой элемент x модуля M можно представить в виде линейной комбинации $x = c_1 \cdot a_1 + c_2 \cdot a_2, ..., c_k \cdot a_k$ для некоторых $c_i \in R$. Если при этом элементы $a_1, a_2, ..., a_k$ линейно независимы (т.е. никакая их линейная комбинация с коэффициентами из R, среди которых есть хотя-бы один ненулевой элемент, не равна 0), то это множество называется *базисом* модуля M. Число элементов базиса называется размерностью модуля.

Примеры. 1. Кольцо $Z[\theta]$, где θ– алгебраический элемент над Q, является конечным модулем с размерностью, равной степени d алгебраического элемента и базисом $1, \theta, \theta^2, ... \theta^{d-1}$.

2. Гауссовым кольцом называется кольцо $\mathbf{Z}[i]$, состоящее из элементов вида $a + bi$, где $i = \sqrt{-1}$–мнимая единица, $a, b \in \mathbf{Z}$. Минимальный многочлен для i, равный $x^2 + 1$, имеет степень 2. Любой модуль в этом кольце конечнопорожден и имеет размерность, не превышающую степень минимального многочлена =2. Примером произвольного модуля в $\mathbf{Z}[i]$ является модуль $(3, 2i)$, состоящий из чисел вида $a + bi$, где a кратно 3, а b–четно.

A.1. Алгебраические расширения числовых полей

Определение A.5. *Пусть K–числовое поле. Элемент θ называется* алгебраическим *над полем K, если он является корнем какого-нибудь многочлена $f(x)$ с коэффициентами из поля K. Многочлен наименьшей степени со старшим коэффициентом 1, имеющий θ своим корнем, называется* минимальным *многочленом для θ.* Степенью *алгебраического элемента θ называется степень его минимального многочлена. Числами,* сопряженными *с θ, называются все остальные корни минимального многочлена.*

Минимальный многочлен любого алгебраического числа неприводим и не имеет кратных корней. Примерами алгебраических элементов над полем $\mathbf{Q}$ являются числа 7, $\sqrt{2}$, $3i$ с минимальными многочленами $f(x) = x - 7$, $x^2 - 2$ и $x^2 + 9$ соответственно.

Определение A.6. *Минимальное поле, содержащее исходное поле K и конечный набор алгебраических над K элементов $\{\theta_1, \theta_2, \dots \theta_r, \}$, называется конечным алгебраическим расширением поля K и обозначается $K[\theta_1, \theta_2, \dots \theta_r]$. Теорема о примитивном элементе утверждает, что любое конечное алгебраическое расширения порождается одним алгебраическим элементом, который называется* примитивным *элементом этого расширения. Иначе говоря, для любого расширения $K[\theta_1, \theta_2, \dots \theta_r]$ существует алгебраический над K элемент θ, такой, что $K[\theta_1, \theta_2, \dots \theta_r] = K[\theta]$.*

Определение A.7. Алгебраическим числовым полем *называется произвольное конечное расширение поля рациональных чисел* $\mathbf{Q}$.

Пример. Присоединим к Q число $i = \sqrt{-1}$. Число i является корнем неприводимого многочлена $x^2 + 1$, и поле $K = Q(i)$–алгебраическое числового поле. Это поле изоморфно полю многочленов 1-степени $ax + b$ с рациональными коэффициентами. Произведение $(2x-1)(x+3)$ вычисляется по модулю x^2+1, т.е. $(2x-1)(x+3) = 2x^2+5x-3 \mod (x^2+1) = 2x^2+5x-3-$

$2(x^2+1) = 5x-5$. Обратный к $g(x)$ элемент можно найти, решая с помощью расширенного алгоритма Евклида, уравнение $u(x) \cdot f(x) + w(x) \cdot g(x) = 1$, и полагая $g^{-1}(x) = w(x)$.

Исходя из сказанного, любое алгебраическое числовое поле $K = Q(\alpha)$ можно представлять, как конечномерное векторное пространство с базисом $B = \{1, x, x^2, ..., x^{d-1}\}$ и коэффициентами из поля Q, где d–степень минимального многочлена элемента α.

Однако в методе решета числового поля рассматриваемые только многочлены с целыми коэффициентами. Следовательно, главным объектом изучения являются элементы кольца многочленов $Z(\alpha)$ для некоторого целого алгебраического числа α. Алгебраическое число называется целым, если его минимальный многочлен имеет целые коэффициенты.

Содержанием многочлена $g(x)$ (*content* of polynomial $g(x)$) называется наибольший общий делитель всех его ненулевых коэффициентов, будем обозначать содержание $content(g)$. Если $content(g) = 1$, то такой многочлен называется *примитивным*.

Многочлен $g(x) \in \mathbf{Z}[x]$ называется *унитарным*, если его старший коэффициент равен 1. Кольцо $\mathbf{Z}[x]$ является бесконечномерным с базисом $B = \{1, x, x^2, ...\}$. Однако, в дальнейшем, речь будет идти только о кольце вычетов многочленов из $\mathbf{Z}[x]$, взятых по модулю неприводимого многочлена $f(x)$. Такая структура обозначается как $\mathbf{Z}[x]/(f(x))$, где $(f(x))$– это идеал $\mathbf{Z}[x]$, порождаемый многочленом $f(x)$. Это кольцо изоморфно кольцу многочленов $\mathbf{Z}[\alpha]$, где α–корень многочлена $f(x)$.

А.2. Идеалы коммутативных колец

Определение А.8. *Пусть M – непустое подмножество элементов кольца C.* Идеалом *коммутативного кольца $< C, +, * >$, порождаемым множеством M, называется наименьшее множество элементов, содержащее M и замкнутое относительно операций сложения и умножения на элементы из всего кольца C:*

1. $(\forall x, y \in I)\ x + y \in I$

2. $(\forall x \in C)(\forall y \in I)\ x * y \in I$.

Очевидно, что любой идеал сам является кольцом.

Любой идеал кольца R можно также рассматривать как подмодуль кольца R.

Определение А.9. *Идеал I кольца C называется главным, если он порождается одним элементом:* $I = (a)$.

Примерами идеалов в коммутативных кольцах могут служить:

1. Нулевой идеал, состоящий из одного нуля.

2. Единичный идеал, содержащий все элементы кольца.

3. Главные идеалы (a), порожденные элементом $a \in C$ и состоящие из всевозможных выражений вида $ra + ka$, где $r \in C$, $k \in \mathbf{Z}$.

Идеал I называется простым, если для любых a и $b \in C$ из $ab \in I$ и $a \notin I$ следует, что $b \in I$. Идеал I называется примарным, если для любых a и $b \in C$, если $ab \in I$, и $a \notin I$, тогда для некоторого натурального n $b^n \in I$. Каждый простой идеал является примарным, но обратное не верно.

Идеалы колец играют ту же роль, что и нормальные подгруппы в группах, т.е. служат ядрами гомоморфизмов. Простые идеалы соответствуют простым целым числам, а примарные идеалы – степеням простых чисел кольца $\mathbf{Z}$.

Суммой идеалов I_1 и I_2 кольца C (обозначается $I_1 + I_2$) называется идеал, порожденный элементами, принадлежащими как I_1, так и I_2. Произведением идеалов I_1 и I_2 кольца C (обозначается $I_1 \cdot I_2$) называется идеал, порожденный всевозможными произведения вида ab, где $a \in I_1$ и $b \in I_2$.

Например, если в кольце Z идеалы I_1, I_2 равны соответственно $I_1 = (6)$, $I_1 = (9)$, то $I_1 + I_2 = (3)$, $I_1 \cdot I_2 = (18)$.

Любой идеал I кольца C определяет разбиение кольца на смежные классы или классы *вычетов* по идеалу I. Элементы a и b кольца C

называются *сравнимыми по идеалу* I, если их разность $a - b$ принадлежит I, записывается

$$a \equiv b \pmod{I}.$$

Обозначим через $[a]$ класс вычетов, содержащий элемент a. Над классами вычетов можно выполнять те же операции, что и с элементами кольца, определяя $[a] + [b] = [a + b]$, $[a] * [b] = [a * b]$. Нетрудно доказать, что эти операции определены корректно на классах вычетов, т.е. результат не зависит от выбора представителей классов.

Множество классов вычетов также является кольцом, которое называется фактор-кольцом кольца C по идеалу I и обозначается C/I.

Определение A.10. *Произвольный непустой набор X поля K порождает ненулевой идеал I относительно операций сложения и умножения на элементы K. Минимальный набор X, порождающий I, называется базисом идеала I. Число элементов базиса называется* размерностью *базиса.*

Нетрудно видеть, что в кольце $\mathbf{Z}[\theta]$, размерность базиса любого идеала не может превышать степень алгебраического числа d. Значит, любой идеал в $\mathbf{Z}[\theta]$ обладает конечным базисом. Кольца, в которых любой идеал является конечнопорожденным, называются *нетеровыми (noetherian)* по имени Эмми Нетер-Amalie Emmy Noether (1882–1935)., считающейся самым крупным математиком среди женщин-математиков всех времен и народов.

A.3. Целые алгебраические числа

Напомним, что алгебраическое число α называется целым, если его минимальный многочлен имеет целые коэффициенты. Зафиксируем алгебраическое числовое поле $K = Q(\alpha)$, определяемое минимальным многочленом $f(x) = x^d + a_{d-1}x^{d-1} + \ldots a_0$ с целыми коэффициентами, и обозначим через Z_K множество целых алгебраических чисел, принадлежащих полю K.

Очевидно, элемент α принадлежит Z_K. Также кольцу Z_K принадлежат многочлены от α с целыми коэффициентами. Если два

элемента принадлежат Z_K, то их сумма и произведение также принадлежат Z_K. Справедлива следующая теорема:

Теорема А.1. *Множество Z_K является кольцом, содержащим кольцо $\mathbf{Z}[\alpha]$ многочленов с целыми коэффициентами в виде подкольца.*

При этом Z_K не обязано совпадать с кольцом $\mathbf{Z}[\alpha]$. Например, в поле $K = Q[\sqrt{5}]$ элемент $g = (-1+\sqrt{5})/2$ является целым, т.к. он является корнем (минимального) многочлена $f(x) = x^2+x-1$, однако не принадлежит кольцу $\mathbf{Z}[\sqrt{5}]$.

Понятие целого алгебраического числа является очень важным для теории алгебраических числовых полей, поскольку теорема об однозначном разложении целых элементов поля K может выполняться только в кольце целых алгебраических чисел:

Теорема А.2. *Если в произвольном кольце $C \subset K$ выполняется теорема об однозначном разложении на множители, то кольцо C –цельнозамкнуто, т.е. содержит корни всех неприводимых унитарных многочленов с коэффициентами из C.*

С другой стороны, разница между кольцами Z_K и $\mathbf{Z}[\alpha]$ не столь велика, и любое целое алгебраическое число умножением на некоторый фиксированный элемент кольца $\mathbf{Z}[\alpha]$ может быть преобразовано в элемент $\mathbf{Z}[\alpha]$:

Теорема А.3. *Для любого $g(\alpha) \in Z_K$, $g(\alpha) \cdot f'(\alpha)$ лежит в $\mathbf{Z}[\alpha]$. Здесь $f(x)$ –минимальный многочлен элемента α.*

Поговорим теперь о возможности однозначного разложения целых элементов поля K в виде произведения простых элементов.

Определение А.11. *Коммутативное кольцо называется областью целостности (integral domain), если в нем нет делителей нуля, т.е. ненулевых элементов a, b, произведение которых равно 0. Будем называть произвольное коммутативное кольцо* дедекиндовым, *если оно является областью целостности и любой элемент этого кольца однозначно представим в виде произведения простых элементов этого кольца.*

Классическим примером дедекиндового кольца является кольцо целых чисел **Z**. Более общим примером дедекиндового кольца является теорема об однозначном разложении в произвольном *евклидовом* кольце.

Определение A.12. *Кольцо C называется евклидовым, если оно является областью целостности, и определена евклидова функция нормы Nr, удовлетворяющая следующему свойству:*

$$(\forall a, b \in C)(\exists k, r \in C)a = k \cdot b + r,$$

причем, $Nr(r) < Nr(b)$.

Для колец многочленов понятием евклидовой нормы может служить норма многочлена, определяемая ниже.

A.4. Норма полинома

Рассмотрим кольцо $Z(\alpha)$, где, по-прежнему, α–примитивный корень неприводимого многочлена $f(x)$ степени d с целыми коэффициентами. Будем рассматривать $Z(\alpha)$ как d–мерное векторное пространство над кольцом Z. Стандартным базисом этого пространства является базис $B = (x^{d-1}, x^{d-2}, \ldots 1)$. Если стандартный базис B зафиксирован, то нормой вектора h в пространстве V будем называть определитель матрицы линейного преобразования H, определяемого умножением вектора h на элементы базиса (т.е. матрицы преобразования, переводящего систему векторов B в систему $h \cdot B$).

$$Nr(h) = det(H) \tag{A.164}$$

Очевидным следствием определения нормы является ее мультипликативность:

$$Nr(g \cdot h) = Nr(g) \cdot Nr(h) \tag{A.165}$$

Пример вычисления нормы

Пусть $f(x) = x^3 + a_2x^2 + a_1x + a_0$. Вычислим норму многочлена первой степени $h = x - b$. Умножив h на базис $B = (x^2,\, x,\, 1)$ получим вектор

$$B \cdot h = (x^3 - bx^2,\, x^2 - bx,\, x - b).$$

Только первая координаты вышла за пределы пространства полиномов 2-й степени, поэтому вычитая из первой координаты вектора $B \cdot h$ полином $f(x)$, получим вектор из пространства V:

$$B \cdot h = (-a_2x^2 - a_1x - a_0 - bx^2,\, x^2 - bx,\, x - b).$$

Перепишем последнее равенство в виде:

$$B \cdot h = (x^2,\, x,\, 1) \cdot H = (x^2,\, x,\, 1) \cdot \begin{pmatrix} -a_2 - b & 1 & 0 \\ -a_1 & -b & 1 \\ -a_0 & 0 & -b \end{pmatrix} \qquad \text{(A.166)}$$

Значит, матрица H этого преобразования равна:

$$\begin{pmatrix} -a_2 - b & 1 & 0 \\ -a_1 & -b & 1 \\ -a_0 & 0 & -b \end{pmatrix}$$

Определитель этой матрицы равен $b^3 + a_2b^2 + a_1b + a_0 = f(b)$. Значит, норма $Nr(x - b)$ многочлена $x - b$ равна значению многочлена $f(x)$ в т. $x = b$.

Норма полинома $a - bx$

Проведя аналогичные вычисления, нетрудно вычислить норму произвольного многочлена 1-й степени $a - bx$ в поле K, определяемом произвольным унитарным многочленом $f_1(x) = x^d + a_{d-1}x^{d-1} + \ ... \ + a_0$:

$$Nr(a - bx) = -Nr(b) \cdot Nr\left(x - \frac{a}{b}\right) = -b^d \cdot f_1\left(\frac{a}{b}\right) \qquad \text{(A.167)}$$

Отметим, что пространство $Z[x]/(f_2(x))$ представляет собой множество полиномов 0-й степени, т.е. просто кольцо целых чисел.

Значение полинома $g(x) = a - bx$ в этом кольце равно остатку от деления на $f_2(x)$ и равно

$$g(x) = a - bx \mod f_2(x) = a - bx \mod (x - m) = a - bm = g(m),$$

т.е. совпадает со значением полинома в т. m.

Норма полинома $g(x) = a - bx$ в $Z[x]/(f_2(x))$ равна соответственно

$$Nr(a - bx) = -b \cdot f_2(a/b) = -b \cdot (a/b - m) = -(a - bm) = -g(m) \quad \text{(A.168)}$$

т.е. совпадает с точностью до знака с самим полиномом.

Норма многочленов может принимать как положительные, так и отрицательные значения. Для простоты удалим знак '–' перед определением нормы в обоих выражениях и обозначим полученные функции через F_1 и F_2:

$$F_1(a, b) = a^d \cdot f_1\left(\frac{b}{a}\right) = a_d \cdot a^d + a_{d-1}a^{d-1}b + \ ... \ + a_0 b^d, \quad \text{(A.169)}$$

$$F_2(a, b) = a - bm. \quad \text{(A.170)}$$

A.5. Теория делимости в алгебраических числовых полях

Представление целых чисел в виде произведения простых чисел и их степеней, являясь предметом этой монографии, дает классический пример однозначного разложения. Такие представления также возможны в произвольных евклидовых колец, т.е. кольцах, обладающих евклидовой нормой (A.12). Однако, однозначное разложение возможно не только в евклидовых кольцах, но и в кольцах более общего вида – нетеревых кольцах (A.2), обладающих конечным базисом. Поскольку, кольцо целых элементов Z_K алгебраического числового поля $K = Q(\alpha)$ являются нетеровым, то теория делимости для такого кольца является наиболее важной в контексте этой монографии.

Напомним, что для метода решета числового поля требуется только однозначная разложимость в кольце $Z[\alpha]$, где α–корень неприводимого

многочлена. Однако, кольцо $Z[\alpha]$ не является цельнозамкнутым (см.refClosedRing), что является необходимым условием однозначной разложимости. Поэтому вместо кольца $Z[\alpha]$ рассматривается его алгебраическое целое замыкание $Z_K \supseteq Z[\alpha]$, для которого и строится теория делимости.

Простые и неразложимые элементы

В кольце целых чисел понятия простоты и неразложимости являлись эквивалентными. Действительно, натуральное число p является простым, если оно удовлетворяет любому из следующих двух свойств:

1. *Неразложимость:* Если $p = a \cdot b$, то либо $a = \pm 1$, либо $b = \pm 1$,
2. *Простота:* Если $a \cdot b$ делится на p, то либо a делится на p, либо b делится на p.

В кольцах более общего вида эти свойства – не эквивалентны. Свойство простоты является более общим, и из него можно вывести неразложимость. Обратное не всегда верно. Действительно, рассмотрим кольцо $\mathbf{Z}[\sqrt{-6}]$. В нем число 6 имеет два различных представления $6 = 2 \cdot 3$, и $6 = \sqrt{-6} \cdot \sqrt{-6}$. Нетрудно проверить, что все три элемента 2, 3, и $\sqrt{-6}$ неразложимы в кольце $\mathbf{Z}[\sqrt{-6}]$, однако, свойство простоты не выполнено: произведение $2 \cdot 3$ делится на $\sqrt{-6}$, но ни один из сомножителей 2, 3 не делится на $\sqrt{-6}$.

Пока эта неэквивалентность не была замечена и точно сформулирована в 1844 г. Эйзенштейном, многие ранние доказательства относительно делимости элементов в кольцах были неверными. Этой ошибки не избежали даже крупные математики такие как, например, Л.Эйлер. Некоторые из его доказательств о единственности разложений оказались неверными. С другой стороны, Гаусс сумел избежать этой ошибки, приведя строгое доказательство однозначности разложения в кольце целых гауссовых чисел $\mathbf{Z}[i]$.

Требование конечномерности кольца Z_K

Свойство конечномерности идеалов кольца Z_K является необходимым условием для возможности разложения элементов Z_K в произведение простых элементов. Действительно, например, в кольце всех целых алгебраических элементов, обладающем бесконечным базисом, разложение на простые элементы невозможно, потому, что простых элементов нет. Например, для любого целого алгебраического числа β число $\beta_1 = \sqrt{\beta}$ является снова целым алгебраическим числом, из которого опять можно извлечь квадратный корень $\beta_2 = \sqrt{\beta_1}$, являющийся целым алгебраическим числом, и т.д. Этот процесс можно продолжить до бесконечности.

Важным следствием свойства конечномерности является конечность любой последовательности расширяющихся идеалов:

Теорема А.4. *В кольце Z_K любая последовательность идеалов $I_0 \subseteq I_1 \subseteq I_2 \subseteq \ldots$, начиная с некоторого номера, является тождественной.*

Например, в кольце Z_K, где $K = Q[\sqrt{5}]$, можно построить цепочку расширяющихся идеалов $(2) \subset (2, \sqrt{5}) \subset (1, \sqrt{5}) \subset ((1+\sqrt{5})/2, (1-\sqrt{5})/2)$. Последним в этой цепочке указан полный базис (integral basis) кольца Z_K, после которого цепочка расширяться больше не может. Докажем последнее. Произвольный элемент $Q[\sqrt{5}]$ имеет вид $p + q\sqrt{5}$.Его минимальный многочлен имеет вид

$$(x - p + q\sqrt{5})(x - p - q\sqrt{5}) = x^2 - 2pt + (p^2 - 5q^2).$$

Если $p + q\sqrt{5} \in Z_K$, то коэффициенты $-2p$ и $p^2 - 5q^2$ должны быть целыми числами. Отсюда, $2p = a$, $p^2 - 5q^2 = m$, где $a, m \in Z$. Если a–четно, то $p \in Z$, $5q^2 \in Z$, откуда, $q \in Z$. Если a–нечетно, то $20q^2 \in Z$, откуда $2q \in Z$. Отсюда, произвольный элемент Z_K имеет вид $(a + b\sqrt{5})/2$, где $a, b \in Z$.

Однозначное разложение в кольцах

Элемент e кольца Z_K называется единицей, если он обратим, т.е. существует $e^{-1} \in Z_K$. Обратимыми являются в точности те элементы Z_K, норма которых равна ± 1.

Поскольку, если e– единица Z_K, то для любого $a \in Z_K$ выполняется $ae = a$, то однозначность разложения для элементов Z_K возможно только с точностью до умножения на единицы кольца Z_K. Элементы a и b называются *ассоциированными*, если $a = be$, где e–единица кольца Z_K.

Определение А.13. *Будем говорить, что в кольце C деление* однозначно, *если в тех случаях, когда деление элементов кольца возможно, оно однозначно. Иначе говоря, для любых $a, b, c, d \in C$, если выполняется $a = bc$, $a = bd$, и $b \neq 0$, тогда элементы c и d ассоциированы $c = ed$.*

Ранее на примере кольца $\mathbf{Z}[\sqrt{-6}]$ было показано, что отсутствие неразложимых элементов, не являющихся простыми, является необходимым условием однозначности разложения. Однако это условие является, в свою очередь, и достаточным:

Теорема А.5. *Область целостности R с конечным базисом и однозначным делением является дедекиндовым кольцом тогда и только тогда, когда каждый неразложимый элемент является простым.*

Доказательство. Предположим, что каждый неразложимый элемент R является простым. Наличие конечного базиса позволяет представить произвольный элемент $x \in R$ в виде

$$x = e_1 p_1 p_2 \ldots p_k,$$

где e–единица, а p_i– неразложимые (и, значит, простые) элементы R. Предположим, что такое представление не единственно:

$$x = e_1 p_1 p_2 \ldots p_k = e_2 q_1 q_2 \ldots q_m, \tag{А.171}$$

Докажем индукцией по k, что $k = m$ и каждый элемент p_i ассоциируется с каким-нибудь элементом q_j. При $k = 0$ утверждение очевидно. При $k > 0$ имеем $p_k \mid q_1 q_2 \ldots q_m$. В силу простоты p_k найдется j такое, что $p_k \mid q_j$, и $p_k = eq_j$. Сокращая (А.171) на p_k, получим равенство,

$$e_1 p_1 p_2 \ldots p_{k-1} = e_2 q_1 \ldots q_{j-1} e q_{j+1} \ldots q_m, \tag{А.172}$$

для которого уже выполняется индукционное предположение. Теорема доказана.

Отметим, что однозначность деления здесь была нужна, чтобы выполнить сокращение.

Однозначность разложения идеалов в кольце Z_k

Пример неоднозначного разложения в кольце $\mathbf{Z}[\sqrt{-6}]$ показывает, что не во всех алгебраических числовых полях выполняется однозначное разложение на простые элементы. Выходом из этой проблемы является предложение выполнять разложение не самих элементов, а идеалов, порождаемых этими элементами.

Куммер предложил погрузить кольцо алгебраических элементов в более широкое кольцо, элементы которого он назвал *идеальными числами*. Дедекинд ввел понятие *идеала*.

Идея Куммера состояла в том, что если, например, элемент a представим в виде $a = p_1 \cdot p_2$ и $a = q_1 \cdot q_2$, то найдется расширение L поля K, в кольце целых чисел которого, имеется разложение $p_1 = b_1 \cdot b_2$, $p_2 = b_3 \cdot b_4$, $q_1 = b_1 \cdot b_3$, $q_2 = b_2 \cdot b_2$, тогда,

$$a = p_1 \cdot p_2 = (p_1 p_2) \cdot (p_3 p_4) = (p_1 p_3) \cdot (p_2 p_4) = q_1 \cdot q_2.$$

Пример. В поле $Q(\sqrt{15})$ элемент 10 имеет два различных разложения:

$$10 = 2 \cdot 5 = (5 + \sqrt{15})(5 - \sqrt{15}).$$

Ясно, что элементы 2 и 5 не ассоциированы с элементами $5 \pm \sqrt{15}$. Рассмотрим более широкое поле $L = Q(\sqrt{3}, \sqrt{5})$. В кольце целых чисел поля L выполняется разложение:

$$10 = (\sqrt{5})(\sqrt{5})(\sqrt{5}+\sqrt{3})(\sqrt{5}-\sqrt{3}). \qquad \text{(A.173)}$$

Дедекинд заметим, что разложение (A.173) выполняется уже в кольце целых чисел Z_K поля K, просто соответствующие идеалы не являются главными (т.е. имеют более одного генератора), поэтому описать их в K

намного труднее. Действительно, Обозначим через $I_1 - I_3$, полученные в результате пересечения кольца Z_K с идеалами $(\sqrt{5})$, $(\sqrt{5}+\sqrt{3})$ и $(\sqrt{5}-\sqrt{3})$. Тогда, в Z_K имеет место однозначное разложение на простые элементы:

$$10 = (I_1)^2 \cdot I_2 \cdot I_3, \tag{A.174}$$

и все идеалы $I_1 - I_3$ являются нетривиальными и не главными.

Рассмотрим, например, идеал $I_2 = Z_K \cap (\sqrt{5}+\sqrt{3})$. Элементы $\sqrt{5}(\sqrt{5}+\sqrt{3}) = \sqrt{15}+5$ и $\sqrt{3}(\sqrt{5}+\sqrt{3}) = \sqrt{15}+3$ принадлежат I_2. Также I_2 содержит их разность

$$(\sqrt{15}+5) - (\sqrt{15}+3) = 2.$$

Если бы идеал I_2 был бы основным, то он имел бы вид $(a + b\sqrt{15})$, и $(2) \subset a + b\sqrt{15}$, тогда норма 2, равная 4, должна делиться на норму $ab\sqrt{15}$, равную $a^2 - 15b^2$. Поскольку идеал I_2 не тривиален, его норма не равна ± 1, значит,

$$a^2 - 15b^2 = \pm 2.$$

Находя остаток в этом равенстве по модулю 5, получим $a^2 \equiv \pm 2 \,(\mod 5)$, что невозможно.

Таким образом, можно раскладывать идеалы кольца Z_K в произведение простых идеалов, однако простые идеалы в этом разложении не обязаны быть главными, а могут иметь два или более генератора.

Сформулируем это факт в виде следующей теоремы:

Теорема А.6. *Кольцо Z_K не является кольцом главных генераторов.*

Наконец, сформулируем основной результат этого раздела - теорему об однозначном разложении в кольце алгебраических целых чисел (теор.5.5 в [52]):

Теорема А.7. *Каждый ненулевой идеал I кольца целых алгебраических чисел Z_k представим в виде произведения простых идеалов, и это представление однозначно с точностью до порядка сомножителей.*

Список литературы

[1] Agrawal M. *PRIMES is in P* / M.Agrawal, N.Kayal, N.Saxena.– Annals of Mathematics.– 2004, v.160, p. 781–793.

[2] Atkin A. *Prime sieves using binary quadratic forms*/ A. Atkin, D. Bernstein.– http://cr.yp.to/papers/primesieves–19990826.pdf

[3] Bach E. *Factoring with cyclotomic polynomials* / E. Bach, J. Shallit.– Math. Comp. 1989. v.52(185), p. 201–219.

[4] Blake A.(ed). *Advances in Elliptic Curve Cryptography.* / A. Blake(ed).– London Mathematical Society Lecture Note Series. 317, Cambridge Univ.Press, 2005, 281 p.

[5] Boender H. *The number of relations in the Quadratic Sieve Algorithm* / H. Boender NM-R9622, The Netherlands, 1996, p. 1–22.

[6] Brent R.P. *An improved Monte Carlo factorization algorithm*/ R.P. Brent.– BIT, 1980, v.20, p. 176—184.

[7] Brent R.P. *Factorization of the eighth Fermat number* / R.P. Brent, J.M. Pollard.– Math. Comp, 1981, v.36, p. 627-— 630.

[8] Brent R.P. *Some integer factorization algorithms using elliptic curves*/ R.P. Brent.– Austral.Comput.Sci.Comm, 1986, v.8, p. 149–163.

[9] Brent R.P. *Factorization of the tenth Fermat number* / R.P.Brent.– Math. Comp, 1999, v.68, p. 429–451.

[10] Brent R.P. *Some parallel algorithms for integer factorisation* / R.P. Brent.– Lect.Notes in Comp.Sci, 1999, v.1685, p. 1–22.

[11] Brillihart J. *Factorisations of* $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ *up to high powers* / J. Brillihart, D.H. Lehmer, S.Wafstaff.– Contemporary Mathematics, **22**, Th.Edit., AMS, Providence, 2005, 327 p.

[12] Briggs M. *An Introduction to the General Number Field Sieve* / M. Briggs.– Master's Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1998, p. 1–84.

[13] Buhler J.P. *Factoring integers with the number field sieve* / J. P. Buhler, H. W. Lenstra, C. Pomerance.– in The Developement of the Number Field Sieve, Springer–Verlag, Berlin, Germany, 1993, p. 50–94.

[14] Buhler J. *Algorithmic Number Theory: Proc. ANTS-III* / J.P. Buhler(ed.).– Portland, OR, v.1423, Lect.Not.Comp.Sci. Springer–Verlag, 1998, 640 p.

[15] Cavallar S. *Factorization of 512-bit RSA-modulus* / S. Cavallar, W.M. Lioen, H.J.te Riele, B. Dodson, A.K. Lenstra, P.L. Montgomery, B. Murphy et al.– CWI Report MAS-R0007, February 2000, 18 p.

[16] Coblitz N. *The state of elliptic cryptography* / N.Coblitz, A.Menezes, S.Vanstone.– Design, Codes and Cryptography, 19, Kluwer Publ. 2000, p. 103–123.

[17] Cohen H. *A course in computational algebraic number theory* / H. Cohen.– Springer–Verlag, Berlin, 1993, 545 p.

[18] Coppersmith D. *Fast evaluation of discrete logarithms in fields of characteristic two*/ D. Coppersmith.– IEEE Trans Inform. Theory, 1984, v.30(4), p. 587–594.

[19] Coppersmith D. *Solving homogeneous linear equations over GF (2) via block Wiedemann algorithm* / D. Coppersmith.– Math. Comp. 1994, v.62, p. 333–350

[20] Couveignes J.M. *Computing a square root for the number field sieve* / Jean Marc Couveignes.– in [33], p. 95–102

[21] Crandall R. *The prime numbers: a computational perspertive* / R. Crandall, C. Pomerance.– sec.ed. Springer–Verlag, Berlin, 2005, 604 p.

[22] Dixon J.D. *Asymptotically fast factorization of integers* / J.D. Dixon.– Math. Comp. 36, 1981, p. 255–260.

[23] Elkenbracht-Huising M. *An implementation of the Number Field Sieve* / M. Elkenbracht-Huising.– Experimental Mathematics, 1996, v.5, p. 231—-253.

[24] Gardner M. *A new kind of cipher that would take millions years to break* / M. Gardner.– Sci. Amer. 1977, p. 120–124.

[25] Gower J. *Square form factorization*/ J. Gower, S.S. Wagstaff Jr.– Mathematics of Computation, v.77 (2008), p. 551–588.

[26] Hackmann P. *Elementary Number Theory* / P. Hackmann.– HHH Publ, 2007, 411 p.

[27] Joux A. *A one round protocol for tripartie Diffie-Hellman.* / A. Joux.– Algorithmic Number Theory: 4-th International Symposium, ANT–IV, Lecture Notes in Computer Science, v.1838(2000), Springer–Verlag, p. 385–393.

[28] Keller W. *Prime factors* $k \cdot 2^n + 1$ *of Fermat numbers* F_m *and complete factoring status* / W. Keller.– http://www.prothsearch.net/fermat.html

[29] Kleinjung T. *On Polynomial Selection for the General Number Field Sieve*/ T. Kleinjung.– Math. Comp. 75 (2006), 2037–2047 p.

[30] Kleinjung T. *Factorization of a 768-bit RSA modulus* / T. Kleinjung et alt.– Scientific Report, 2010, 22 p.

[31] Lenstra H.W. *Factoring integers with elliptic curves* / H.W. Lenstra.– Ann.Math. v.126 (1987), p. 649–674.

[32] Lenstra A.K. *Factoring integers with the number field sieve* / A. K. Lenstra, H.W. Lenstra,Jr, M.S. Manasse, J.M. Pollard.– in [33] p. 11–42.

[33] Lenstra A. *The Development of the Number Field Sieve* / A. Lenstra and H. Lenstra (eds.).– Lect.Not.in Math.**1554**, Springer–Verlag, Berlin, 1993, 139 p.

[34] Menezes A. *Reducing Elliptic Curve Logarithms to a Finite Field* / A. Menezes, T. Okamoto, S. Vanstone.– IEEE Trans. Info. Theory, v.39, 1993, p. 1639–1646.

[35] Menezes A. *Elliptic Curve Public Key Cryptosystems* / A. Menezes.– 1993, 144 p.

[36] Montgomery P.L. *Speeding the Pollard and Elliptic Curve Methods of Factorization.*/P.L. Montgomery.– Mathematics of Computation, v.48, iss.177, 1987, p.234–264.

[37] Montgomery P.L. *An FFT-extension of the Elliptic Curve Method of Factirization* / P.L. Montgomery.– Doctoral Dissertation, 1992, Univ.Calif. USA, 118 p.

[38] Montgomery P.L. *A block Lanczos algorithm for finding dependences over GF(2)*/ P.L. Montgomery.– in Advances in Cryptology: Eurocrypt'95, Lect.Notes in Comp.Sci. **921**, Springer–Verlag, Berlin, p. 106–120.

[39] Montgomery P.L. *Square roots of products of algebraic numbers.*/P.L. Montgomery.– 1997, 24 p. http://ftp.cwi.nl/pub/pmontgom/sqrt.ps.gz.

[40] Morrison M.A. *A Method of Factoring and the Factorization of F7*/ M.A. Morrison, J. Brillhart.– Mathematics of Computation, AMS, 29 (129),January 1975, p.183–205.

[41] Murphy D.A. *Polynomial selection for the number field sieve.*/ B.A. Murphy. – Doctoral Thesis, Australia, 1999, 142 p.

[42] Niven I. *An introduction to the number theory*/ I. Niven, H. Zuckrman, H. Mongomery. – Willey Publ., 5-th edition, 1991, 541 p.

[43] Pollard J.M. *Theorems on factorization and primality testing* / J.M. Pollard. – Proc.Cambridge Phil.Society. 1974, v.76, p. 521-578.

[44] Pollard J.M. *Factoring with cubic numbers./* J.M. Pollard. – in Lenstra et alt[1993], p. 4-10.

[45] Pollard J.M. *The lattice sieve./* J.M. Pollard. – in Lenstra et alt[1993], p. 43-49.

[46] Pomerance C. *Tale of Two Sieves/* C. Pomerance. – Notices of AMS, 1996, P. 1473–1485.

[47] Pomerance C. *Smooth Numbers and the Quadratic Sieve* / C. Pomerance. – MSRI publications, **v.44** – 2008, p. 69–82.

[48] Pomerance C. *A pipeline architecture for factoring large integers with the quadratic sieve algorithm./* C. Pomerance, J. Smith, R. Tuler. – SIAM J. Comput., 17:387–403, 1988. Special issue on cryptography.

[49] Ribenboim P. *The New Book Of Prime Number Records,/* P. Ribenboim. – 3rd ed. Springer, 1996, 541 p.

[50] Schoof R. *Four primarity testing algorithms./* R. Schoof. – in *Surveys in Algorithmic Number Theory, ed.J.B.Buchler, P.Stevenhagen*, Math.Sci.Res.Inst.Publ. 44, Cambridge Univ.Press, New York, 2008, p.101-126.

[51] Shoup V. *A Computational Introduction to Number Theory and Algebra/* V. Shoup. – Cambridge University Press, Sec.Edition, 2005, 600 p. http://shoup.net/ntb/

[52] Stewart I. *Algebraic Number Theory and Fermat's Last Theorem* / I. Stewart, D. Tall. – Third Ed., Massachusetts:AK Peters, 2002, 314 p.

[53] Venturi D. *Lecture Notes on Algorithmic Number Theory./* D. Venturi. – Springer-Verlag, New-York, Berlin, 2009, 217 p.

[54] Washington L. *Elliptic Curves Number Theory and Cryptography* /L. Washington. – Series Discrete Mathematics and Its Applications, Chapman & Hall/CRC,second ed. 2008, 524 p.

[55] Zhang M. *Factorization of the Numbers of the Form* $\mathbf{m^3 + c_2 m^2 + c_1 m + c_0}$. / M. Zhang. – in [14],P.131-136.

[56] Аграновский А.В. *Практическая криптография: алгоритмы и их программирование* / А.В. Аграновский, Р.А. Хади.– М.: Солон-Пресс, 2009, 256 с.

[57] Айерленд К. *Классическое введение в современную теорию чисел.* / К. Айерленд, М. Роузен. – М.: Мир, 1987, 428 с.

[58] Акритас А. *Основы компьютерной алгебры и приложениями.* / А. Акритас. – М.: Мир, 1994, 544 с.

[59] Богопольский О.В. *Алгоритмическая теория чисел и элементы криптографии.* / О.В. Богопольский.– Спецкурс для студентов НГУ, Новосибирск, 2005, 35 с. / http://math.nsc.ru/ bogopolski/Articles/SpezkNumber.pdf

[60] Болотов А.А. *Элементарное введение в эллиптическую криптографию: протоколы криптографии на эллиптических кривых.* / А.А. Болотов, С.Б. Гашков, А.Б. Фролов. – М.:КомКнига, 2004, 280 с.

[61] Болотов А.А. *Алгоритмические основы эллиптической криптографии.* / А.А. Болотов, С.Б. Гашков, А.Б. Фролов, Часовских А.А.. – М.:РГСУ, 2004, 499 с.

[62] Боревич З.И. *Теория чисел.* / З.И. Боревич, И.Р. Шафаревич. – 3-е издание, М.: Наука, 1985, 504 с.

[63] Ван дер Варден Б.Л. *Алгебра.*/ Б.Л.ван дер Варден. – изд.2, М.: Наука, 1979, 623 с.

[64] Василенко О.Н. *Теоретико-числовые алгоритмы в криптографии*/ О.Н. Василенко. – МЦНМО, 2003, 326 с.

[65] Вельщенбах М. *Криптография на C и C++ в действии: учебное пособие* / М. Вельщенбах. – М.: Триумф, 2008, 464 с.

[66] Захаров В.М. *Вычисления в конечных полях: уч.–метод. пособие* / В.М. Захаров, Б.Ф. Эминов. – Казань: КГТУ им. А.Н.Туполева, 2010, 132 с.

[67] Ишмухаметов Ш.Т. *Об одном подходе к проблеме факторизации натуральных чисел* / Ш.Т. Ишмухаметов, А.А. Бойко, Д.Б. Зиятдинов. – Известия вузов. Математика, №4, 2011, 15-22 с.

[68] Ишмухаметов Ш.Т. *О проблеме выбора полинома в методе решета числового поля* / Ш.Т. Ишмухаметов, Д.Б. Зиятдинов, Р.Г. Рубцова. – Труды III Всероссийской конференции "Информационные технологии в системе социально–экономической безопасности России и ее регионов", Казань, 2010, 177-183 с.

[69] Коблиц Н. *Курс теории чисел и криптографии* / Н. Коблиц. – М.: ТВП, 2001, 260 с.

[70] Корешков Н.А. *Теория чисел.*/Н.А. Корешков. – Уч.-мет. пособие, Казань, КФУ, 2010, 35 с.

[71] Кормен Т. *Алгоритмы: построение и анализ* /Т. Кормен, Ч. Лейзерсон, Р. Ривест. – М.: МЦНМО, 1999.

[72] Лазарева С.В. *Математические основы криптологии: тесты простоты и факторизация* / С.В. Лазарева, А.А. Овчинников. Учебное пособие, Санкт-Петербург, СПбГУАП, 2006, 65 с.

[73] Лидл Р. *Конечные поля*/Р. Лидл,Г. Нидеррайтер.– Т. 1, 2. М.: Мир, 1988, 428 с.

[74] Молдовян Н.А. *Криптография. От примитивов к синтезу алгоритмов* / Н.А.Молдовян, А.А. Молдовян,М.А. Еремеев. – БХВ-Петербург, 2004, 446 с.

[75] Нестеренко Ю.В. *Лекции об алгебраических числах*/ Ю.В. Нестеренко. – Москва, МГУ, лекции, 80 с.

[76] Нестеренко Ю.В. *Теория чисел*/ Ю.В. Нестеренко. – Москва, Изд.Центр Академия, 2008, 273 с.

[77] Сизый С.В. *Лекции по теории чисел: учебное пособие для математических специальностей* / С.В. Сизый.– Екатеринбург, УрГУ, 1999, 136 с.

[78] Эстарбю О. *Прямые методы для разреженных матриц.* /О. Эстербю, З. Златев.– М.: Мир, 1987, 118 с.

[79] Черемушкин А.В. *Лекции по арифметическим функциям в криптографии* / А.В. Черемушкин.– М.: МЦНМО, 2002.

Предметный указатель