Н.Коблиц

КУРС ТЕОРИИ ЧИСЕЛ И КРИПТОГРАФИИ

Москва: Научное изд-во ТВП, 2001, x+254 с.

Цель данной книги — ввести читателя в те области арифметики, как классические, так и самые современные, которые находятся в центре внимания приложений теории чисел, особенно криптографии. Предполагается, что знание высшей алгебры и теории чисел ограничено самым скромным знакомством с их основами; по этой причине излагаются также необходимые сведения из этих областей математики. Авторами избран алгоритмический подход, причем особое внимание уделяется оценкам эффективности методов, предлагаемых теорией. Особенностью книги является изложение совсем недавно разработанных приложений теории эллиптических кривых. Перевод на русский язык осуществлен с оригинала второго издания, существенно пересмотренного по сравнению с первым изданием и снабженного обновленным списком литературы. Каждая глава включает в себя тщательно составленную подборку задач, как правило, снабженных подробными указаниями и решениями.

Все это позволяет рекомендовать книгу не только в качестве ценного пособия для общетеоретической подготовки специалистов по защите информации, но и как полезный источник примеров практической применимости целого ряда абстрактных разделов математики и кибернетики. Книга прекрасно подходит и для самообразования.

СОДЕРЖАНИЕ

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ПРЕДИСЛОВИЕ

«...Гаусс, а также другие математики, по-видимому, справедливо утверждали, что существует, по крайней мере, одна наука (теория чисел), которая, ввиду своей отдаленности от обыденной человеческой деятельности, остается чистой и благородной.»

— G. H. Hardy, *A Mathematician's Apology*, 1940

Харди был бы удивлен и, возможно, огорчен ростом интереса к приложениям теории чисел в таких областях «обычной человеческой деятельности», как передача информации (коды, исправляющие ошибки) и криптография (секретные коды). Не прошло и полувека с того момента, как Харди написал приведенные выше строки, а уже не кажется невероятным (хотя пока этого не произошло), что АНБ (Агенство Национальной Безопасности, работа которого в области криптографии обеспечивает потребности правительства США) будет требовать предварительной экспертизы для разрешения на публикацию теоретических работ по определенным направлениям теории чисел.

Одной из причин, благодаря которым отдельные вопросы, интересовавшие специалистов по теории чисел, превратились в новое направление, названное «вычислительной теорией чисел», явился быстрый рост мощности и сложности компьютеров.

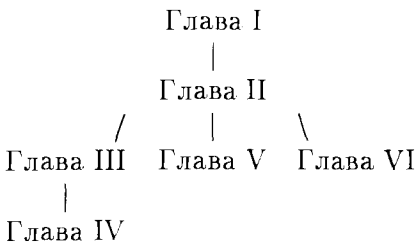
Книга почти не предполагает знания основ алгебры или теории чисел. Ее цель — ввести читателя в те области арифметики, как классические, так и самые современные, которые находятся в центре внимания приложений, особенно криптографии. По этой причине мы избрали алгоритмический подход, выделяя особенно вопрос оценки эффективности методов, предлагаемых теорией. Особенностью нашей книги является изложение совсем недавно разработанных приложений теории эллиптических кривых (глава VI). Эллиптические кривые долгое время были центральной темой некоторых направлений математики. Теперь оказалось, что арифметика эллиптических кривых может иметь и практические приложения.

Во все главы включены задачи, которые должны помочь читателям, интересующимся более широким кругом вопросов, полнее усвоить изучаемый материал.

Первые две главы посвящены общим вопросам. Студентам, не имевшим дела с алгеброй (расширениями полей, конечными полями) или элементарной теорией чисел (сравнениями), изложение покажется весьма сжатым, но они могут обратиться к более подробным курсам. С другой стороны, читатели с большей математической подготовкой, возможно, лишь бегло просмотрят эти главы, остановившись только на некоторых упражнениях.

При соответствующей подготовке слушателей первые пять глав можно взять за основу семестрового курса. Главы III–VI можно использовать в семестровом курсе, продолжающем семестровый курс элементарной теории чисел.

Зависимость между материалом глав следующая (исходя из ссылок в главах V и VI на предыдущий материал):



За основу этой книги взяты курсы, прочитанные в университете штата Вашингтон (Сиэтл) в 1985–86 годах и в Институте математики (Мадрас, Индия) в 1987 году. Мне приятно поблагодарить Гэри Нельсон и Дугласа Линда за апробацию рукописи и полезные замечания.

Рисунок на фронтисписе, принадлежащий профессору А. Т. Фоменко из Московского государственного университета, иллюстрирует тему книги. Обратите внимание, что коды десятичных цифр на стенах здания вовсе не случайны.

Эта книга посвящена памяти студентов Вьетнама, Никарагуа и Сальвадора, отдавших свои жизни в борьбе против американской агрессии. Авторский гонорар от продажи книги будет использован на покупку научной литературы для университетов и институтов этих трех стран.

Сиэтл, май 1987 года

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Вместе с возникновением в криптографии новых понятий и методов расширился и круг криптографических приложений теории чисел. В дополнение к элементарной и аналитической теории чисел все более широко используется алгебраическая теория чисел (тесты на простоту с применением сумм Гаусса и Якоби, криптосистемы, основанные на квадратичных полях, решето числового поля) и арифметическая алгебраическая геометрия (факторизация при помощи эллиптических кривых, криптосистемы, основанные на эллиптических и гиперэллиптических кривых и абелевых многообразиях). Некоторые из современных применений теории чисел в криптографии (наиболее значительное из них — способ разложения больших целых чисел на множители применением метода решета числового поля, разработанный после выхода первого издания) остались за пределами книги. Однако за счет небольшого увеличения объема книги удалось включить в нее некоторые новые темы, что позволило более полно отразить разнообразие применений теории чисел в этой увлекательной междисциплинарной области.

Следующий список суммирует основные изменения во втором издании.

— Внесены исправления и разъяснения, добавлены многочисленные ссылки.

— В главу VI добавлен раздел о доказательствах с нулевым разглашением и скрытой передаче.

— В главу V добавлен раздел о разложении на множители методом квадратичного решета.

— В главу VI включен раздел об использовании эллиптических кривых в тестах на простоту.

— Добавлено беглое обсуждение следующих вопросов: (n, k)-пороговые схемы, вероятностное шифрование, хеш-функции, рюкзачная криптосистема Чора–Райвеста и американский стандарт цифровой подписи DSS.

Сиэтл, май 1994 года

ГЛАВА I

НЕКОТОРЫЕ ВОПРОСЫ ЭЛЕМЕНТАРНОЙ ТЕОРИИ ЧИСЕЛ

Большинство понятий, рассматриваемых в настоящей главе, вероятно, хорошо известно читателям. Цель этой главы — напомнить те обозначения и факты из элементарной теории чисел, которые будут необходимы при работе с этой книгой. Большинство доказательств опускается, поскольку их можно найти почти в каждом учебнике по теории чисел. Одна тема, которая будет играть центральную роль далее, — оценивание числа двоичных операций, необходимых для решения различных теоретико-числовых задач на компьютере, — еще не вошла в стандартные курсы по элементарной теории чисел. Поэтому мы будем более детально рассматривать вопросы, связанные с оценками сложности, особенно в § 1.

§ 1. Временные оценки сложности арифметических операций

Числа в разных базах. Разложение неотрицательного целого числа n по основанию (базе) b представляет собой обозначение для n вида $(d_{k-1}d_{k-2}\dots d_1d_0)_b$, где d_i — цифры, т. е. символы целых чисел от 0 до $b-1$. Эта запись означает, что $n = d_{k-1}b^{k-1} + d_{k-2}b^{k-2} + \dots + d_1b + d_0$. Если цифра первого разряда отлична от нуля, то число n называют k-разрядным b-ичным числом. Любое целое число от b^{k-1} до $b^k - 1$ является k-разрядным по основанию b. Мы будем опускать скобки и индекс $(\dots)_b$ в случае десятичной системы счисления ($b = 10$) и в тех случаях, когда основание системы счисления ясно из контекста, особенно в случае двоичной системы ($b = 2$). Поскольку иногда удобно пользоваться системами, отличными от десятичной, приходится производить арифметические операции по произвольному основанию и переходить от одного основания к другому. Продемонстрируем это

на нескольких примерах.

З а м е ч а н и я. 1) Дробные числа также можно разлагать по любому основанию, т. е. представлять в виде

$$(d_{k-1}d_{k-2}\ldots d_1d_0, d_{-1}d_{-2}\ldots)_b.$$

2) При $b > 10$ принято использовать буквы для выражения цифр, больших девяти. Можно вообще использовать буквы вместо цифр.

П р и м е р 1. а) $(11001001)_2 = 201$.

b) При $b = 26$ будем использовать буквы латинского алфавита A–Z для записи цифр 0–25 соответственно. Тогда $(\mathrm{BAD})_{26} = 679$, тогда как $(\mathrm{B,AD})_{26} = 1\frac{3}{676}$.

П р и м е р 2. Умножить 160 на 199 в системе счисления по основанию 7.

Р е ш е н и е:

$$\begin{array}{r} 316 \\ \underline{403} \\ 1254 \\ \underline{16030} \\ 161554 \end{array}$$

П р и м е р 3. Разделить $(11001001)_2$ на $(100111)_2$ и $(\mathrm{HAPPY})_{26}$ на $(\mathrm{SAD})_{26}$.

Р е ш е н и е:

$$\begin{array}{llll} 11001001 & |\underline{100111} & \mathit{HAPPY} & |\mathit{SAD} \\ \underline{100111} & 101 & \mathit{GYBE} & \mathit{KD} \\ 101101 & & \mathit{COLY} & \\ \underline{100111} & & \mathit{CCAJ} & \\ 110 & \text{(ост)} & \mathit{MLP} & \text{(ост)} \end{array}$$

П р и м е р 4. Выразить 10^6 в системах счисления по основаниям 2, 7 и 26 (в последнем случае использовать буквенную запись).

Р е ш е н и е. Чтобы разложить число n по основанию b, надо сначала получить младший разряд, разделив n на b и взяв остаток от деления. Потом n заменяется частным от деления, описанный процесс повторяется и дает второй от конца знак разложения и т. д. В данном случае получаем

$$10^6 = (11110100001001000000)_2 = (11333311)_7 = (\mathrm{CEXHO})_{26}.$$

П р и м е р 5. Выразить $\pi = 3,1415926\ldots$ в двоичной системе счисления (выписывая 15 цифр после запятой) и по основанию 26 (выписывая 3 цифры после запятой).

Р е ш е н и е. Сначала представляется целая часть. Потом дробная часть умножается на основание b. Целая часть полученного числа дает d_{-1}, а к его дробной части применяется та же процедура, что последовательно дает $d_{-2}, d_{-3}, \ldots$. Таким способом получаем:

$$3,1415926\ldots = (11,001001000011111\ldots)_2 = (\mathrm{D},\mathrm{DRS}\ldots)_{26}.$$

Число разрядов. Как было упомянуто выше, целое число, удовлетворяющее неравенствам $b^{k-1} \leqslant n < b^k$, имеет k разрядов по основанию b. С помощью логарифмов это можно переписать в следующем виде («[ ]» обозначает целую часть числа):

$$\text{число разрядов } = [\log_b n] + 1 = \left[\frac{\log n}{\log b}\right] + 1$$

(здесь и всюду далее log понимается как натуральный логарифм $\log_e$).

Двоичные операции. Начнем с очень простой арифметической задачи сложения двух двоичных целых чисел, например,

$$\begin{array}{r} {\scriptstyle 1111} \\ 1111000 \\ +\,0011110 \\ \hline 10010110 \end{array}$$

Предположим, что оба числа имеют длину в k бит (слово «бит» является сокращением выражения «*bi*nary digi*t*»). Если запись одного из чисел короче, ее можно дополнить нужным числом нулей слева. Хотя в примере рассматриваются маленькие целые (складываются 120 и 30), следует иметь в виду, что число k может быть очень большим, скажем, 500 или 1000.

Детально проанализируем всю процедуру сложения. При сложении необходимо k раз повторить следующие шаги.

1. Посмотреть на верхний и нижний биты, а также проверить, имеется ли перенос единицы от сложения младших разрядов.

2. Если оба бита нулевые, а переноса нет, то в данном разряде суммы записываем нуль и двигаемся дальше.

3. Если либо a) оба бита нулевые и есть перенос, либо b) один бит — нуль, другой — единица и переноса нет, то записываем единицу и двигаемся дальше.

4. Если либо a) один бит — нуль, другой — единица и есть перенос, либо b) оба бита — единицы и переноса нет, то записываем нуль в данный разряд, записываем единицу переносов в следующий столбец и двигаемся дальше.

5. Если оба бита — единицы и есть перенос, то в данном разряде суммы записываем единицу, записываем единицу переносов в следующий столбец и двигаемся дальше.

Однократное выполнение этих шагов называется двоичной (битовой) операцией. Сложение двух k-разрядных двоичных чисел требует k двоичных операций. Мы увидим ниже, что и более сложные задачи тоже могут быть разбиты на двоичные операции. Время, которое расходует компьютер на решение задачи, по сути дела пропорционально числу двоичных операций. Конечно, константа пропорциональности — число наносекунд, расходуемых на одну двоичную операцию, — зависит от вида компьютера. (Сказанное является упрощением, так как это время может зависеть также от «технических» факторов, например, времени доступа к памяти.) Когда мы говорим об оценке времени работы, подразумевается оценка числа двоичных операций. В этих оценках мы будем пренебрегать временем, расходуемым на запись информации или на логические шаги, отличные от двоичных операций. На практике основное время занимает выполнение именно двоичных операций.

Теперь рассмотрим процесс умножения k-разрядного двоичного числа на l-разрядное двоичное число. Например,

$$\begin{array}{r} 11101 \\ \underline{\times\ \ 1101} \\ 11101 \\ 111010 \\ \underline{11101\ \ \ } \\ 101111001 \end{array}$$

Предположим, что мы пользуемся обычным способом умножения k-разрядного двоичного числа n и l-разрядного двоичного числа m. Мы получаем самое большее l строк (каждый нулевой бит числа m уменьшает это количество на единицу), где каждая строка содержит копию числа n, сдвинутую влево на некоторое расстояние, т. е. копию, дополненную нулями справа. Пусть имеется $l' \leqslant l$ строк. Поскольку мы ограничиваемся двоичными операциями, мы не можем сложить все строки сразу. Правильнее будет двигаться от второй строки к l'-й, складывая каждую строку с накопившейся суммой верхних строк. На каждом этапе сначала отмечаем, как далеко сдвинуто влево число n в рассматриваемой строке. Сносим вниз крайние правые разряды накопленной суммы верхних строк, а остальную часть записи накопленной суммы складываем (описанным выше способом) с числом n, что требует k двоичных операций. В примере выше 11101×1101 после сложения первых двух строк получаем 10010001, сносим вниз послед-

ние три разряда 001 и складываем остальное (т. е. 10010) с $n = 11101$. И наконец, к сумме $10010 + 11101 = 101111$ приписываем 001 и получаем 101111001 — сумму $l' = 3$ строк.

Это описание показывает, что задача умножения может быть разложена на $l' - 1$ сложений, по k двоичных операций каждое. Так как $l' - 1 < l' \leqslant l$, то получаем простую оценку

$$\text{Time (умножение } k\text{-разрядного и } l\text{-разрядного двоичных чисел)} < kl.$$

Здесь и далее $\text{Time}(A)$ обозначает число двоичных операций, необходимых для выполнения процедуры A.

Сделаем несколько замечаний об этом выводе оценки для числа двоичных операций, необходимых для двоичного умножения. Во-первых, как упоминалось выше, мы подсчитали только число двоичных операций. Мы пренебрегли временем на сдвиг числа n влево и временем на снос крайних правых разрядов накопленной суммы. На практике операции сдвига и копирования являются быстрыми по сравнению с большим числом производимых двоичных операций, так что можно без опаски проигнорировать их. Другими словами, мы *определим* «временную оценку (сложности)» арифметической задачи как верхнюю границу для числа двоичных операций без учета операций сдвига, копирования, обращения в память и т. п. Заметим, что мы могли бы применить эту же временную оценку к умножению k-разрядной и l-разрядной двоичных дробей. Единственное отличие состоит в необходимости правильного определения места для запятой, разделяющей целую и дробную части.

Во-вторых, если мы хотим получить простую и удобную в работе оценку, мы всегда должны предполагать, что имеем дело с «самым плохим случаем». Например, если двоичное разложение числа m имеет много нулей, то l' будет значительно меньше l. Поэтому можно использовать оценку Time (умножение k-разрядного и l-разрядного двоичных чисел)$< k\cdot$(число единиц в двоичном разложении m). Однако обычно вместо такого уточнения (понижения) нашей временной оценки удобнее пользоваться простой равномерной оценкой, зависящей лишь от длины записи n и m, а не от конкретных значений битов.

Как частный случай, имеем оценку:

$$\text{Time (умножение двух } k\text{-разрядных двоичных чисел)} < k^2.$$

Наконец, наша оценка kl может быть выражена в терминах n и m, если вспомнить приведенную выше формулу для числа разрядов, из которой следует, что $k = [\log_2 n] + 1 \leqslant \frac{\log n}{\log 2} + 1$ и $l = [\log_2 m] + 1 \leqslant \frac{\log m}{\log 2} + 1$.

Пример 6. Найти верхнюю границу для числа двоичных операций, необходимых для вычисления $n!$.

Решение. Используем следующую процедуру. Сначала умножим 2 на 3, затем результат умножим на 4, новый результат умножим на 5 и т. д., пока не получим $n!$. На $(j-1)$-м шаге $(j = 2, 3, \ldots, n-1)$ производится умножение $j!$ на $j+1$. Поэтому есть $n-2$ шагов, каждый из которых состоит в умножении частичного произведения $j!$ на очередное целое число. Частичные произведения быстро станут очень большими. В качестве оценки для числа разрядов частичного произведения в наихудшем случае возьмем число разрядов последнего произведения, т. е. числа $n!$.

При определении числа двоичных разрядов в произведении мы используем тот факт, что это число не превосходит суммы числа разрядов у сомножителей (см. выше рассуждение об умножении). Следовательно, произведение n целых k-разрядных чисел имеет не более nk разрядов. Таким образом, если n — двоичное k-разрядное число (тогда любое меньшее число имеет не больше k разрядов), то $n!$ имеет самое большее nk разрядов.

Итак, в каждом из $n-2$ умножений, необходимых при вычислении $n!$, мы умножаем не более чем k-разрядное целое число $j+1$ на не более чем nk-разрядное целое число $j!$. Это требует самое большее nk^2 двоичных операций. Всего таких умножений $n-2$. Поэтому общее число двоичных операций ограничено величиной $(n-2)nk^2 = n(n-2)([\log_2 n]+1)^2$, что приблизительно равно $n^2(\log_2 n)^2$.

Пример 7. Определить верхнюю границу для числа двоичных операций, необходимых для умножения многочлена $\sum a_i x^i$ степени не выше n_1 на многочлен $\sum b_j x^j$ степени не выше n_2; коэффициенты многочленов — положительные целые числа, не превосходящие m. Считается, что $n_2 \leqslant n_1$.

Решение. Для вычисления выражения $\sum_{i+j=\nu} a_i b_j$, являющегося коэффициентом при x^ν в произведении многочленов (здесь $0 \leqslant \nu \leqslant n_1 + n_2$), требуется не более $n_2 + 1$ умножений и n_2 сложений. Перемножаемые числа ограничены величиной m, а складываемые числа ограничены величиной m^2; но, так как мы должны сложить n_2 таких чисел, надо взять в качестве границы для слагаемых величину $n_2 m^2$. Таким образом, необходимое для вычисления коэффициента при x^ν число двоичных операций не превосходит

$$(n_2+1)(\log_2 m+1)^2 + n_2(\log_2(n_2 m^2)+1).$$

Поскольку в произведении многочленов имеется n_1+n_2+1 степеней x^ν,

наша временная оценка умножения многочленов выражается как

$$(n_1 + n_2 + 1)\left((n_2 + 1)(\log_2 m + 1)^2 + n_2(\log_2(n_2 m^2) + 1)\right).$$

Пренебрегая единицами, получаем менее строгую, но более компактную оценку

$$\frac{n_2(n_1 + n_2)}{\log 2}\left(\frac{(\log m)^2}{\log 2} + (\log n_2 + 2\log m)\right).$$

Замечание. Положим $n = n_1 \geqslant n_2$, и пусть $m \geqslant 16$ и $m \geqslant \sqrt{n_2}$ (что обычно выполняется на практике). Тогда последнее выражение можно упростить до $4n^2(\log_2 m)^2$. Этот пример показывает, что в общем случае нет единого «правильного ответа» на вопрос о нахождении границы для времени решения задачи. Можно искать границу как достаточно простую функцию от исходных данных (в рассмотренном случае это n_1, n_2 и m), которая для большинства исходных данных дает оценку, по порядку более или менее близкую к числу реально выполняемых двоичных операций. Поэтому, скажем, в примере 7 нет смысла заменять нашу оценку оценкой $4n^2 m$, так как при больших m она на много порядков больше реального числа операций.

До сих пор мы имели дело со сложением и умножением целых чисел. Две другие арифметические операции — вычитание и деление — имеют те же самые оценки временной сложности, что сложение и умножение соответственно: Time(вычитание k-разрядного двоичного числа из l-разрядного) $\leqslant \max(k, l)$, Time(деление k-разрядного двоичного числа на l-разрядное) $\leqslant kl$. Более точно, для описания вычитания надо расширить круг двоичных операций, включив в него операцию вычитания нуля или единицы из других нуля или единицы, возможно, с займом единицы из старшего разряда (см. пример 8).

Анализируя деление в двоичной системе, будем ориентироваться на иллюстрацию, подобную примеру 3. Пусть $k \geqslant l$ (если $k < l$, деление тривиально, т.е. частное равно нулю, а все делимое образует остаток). Нахождение частного и остатка требует самое большее $k - l + 1$ вычитаний. Каждое вычитание требует l или $l + 1$ двоичных операций, но в последнем случае в самом левом разряде разности будет стоять нуль, поэтому можно опустить одну двоичную операцию (считая, что это скорее операция «по учету данных», а не вычисление). Подобным образом мы игнорируем и другие технические детали, например, сравнение двоичных целых чисел (при определении минимального числа разрядов делимого, которые образуют число, большее

делителя), снос разрядов и т. п. Таким образом, наша оценка есть просто $(k-l+1)l$, что не больше kl.

Пример 8. Найти верхнюю границу для числа двоичных операций, необходимых для вычисления биномиального коэффициента $\binom{n}{m}$.

Решение. Так как $\binom{n}{m} = \binom{n}{n-m}$, то без потери общности можно предположить, что $m \leqslant n/2$. Будем использовать следующую процедуру вычисления $\binom{n}{m} = n(n-1)(n-2)\cdots(n-m+1)/(2\cdot 3\cdots m)$. Мы имеем $m-1$ умножений и последующие $m-1$ делений. Каждый раз максимально возможная величина первого числа при умножении и делении есть $n(n-1)(n-2)\cdots(n-m+1)$, а граница для второго числа есть n. Рассуждая аналогично примеру 6, убеждаемся в том, что граница для общего числа двоичных операций есть $2(m-1)m([\log_2 n]+1)^2$, что при больших m и n приблизительно равно $2m^2(\log_2 n)^2$.

Теперь обсудим весьма удобное обозначение для краткой записи временных оценок сложности.

O-большое. Пусть $f(n)$ и $g(n)$ — функции положительного целочисленного аргумента, принимающие положительные (не обязательно целые) значения при всех n. Скажем, что $f(n) = O(g(n))$ (или просто $f = O(g)$), если существует такая константа C, что $f(n)$ всегда меньше $C\,g(n)$. Например, $2n^2 + 3n - 3 = O(n^2)$ (действительно, нетрудно доказать, что левая часть меньше $3n^2$).

Поскольку мы хотим использовать обозначение O-большое в более общей ситуации, дадим более широкое определение. Рассмотрим f и g как функции нескольких переменных и не будем обращать внимание на соотношение между ними при небольших значениях n. Так же, как это делается в теории пределов в анализе, будем рассматривать лишь большие значения n.

Определение. Пусть $f(n_1, n_2, \ldots, n_r)$ и $g(n_1, n_2, \ldots, n_r)$ — две функции, определенные на наборах из r положительных целых чисел. Предположим, что существуют такие константы B и C, что, когда все n_j больше B, обе функции положительны и $f(n_1, n_2, \ldots, n_r) < Cg(n_1, n_2, \ldots, n_r)$. В этом случае говорим, что функция f ***ограничена*** функцией g, и пишем $f = O(g)$.

Заметим, что равенство в обозначении $f = O(g)$ следует, скорее, понимать как неравенство $<$, а O-большое — как некоторую мультипликативную константу.

Пример 9. а) Пусть $f(n)$ — ***произвольный*** многочлен степени d с положительным старшим коэффициентом. Тогда, как легко показать, $f(n) = O(n^d)$. В более общем случае можно показать, что

$f = O(g)$, если $f(n)/g(n)$ имеет конечный предел при $n \to \infty$.

b) Если ε — сколь угодно малое положительное число, можно показать, что $\log n = O(n^{\varepsilon})$ (т. е. при больших n функция log меньше любой степенной, сколь малой ни была бы степень). Это следует из равенства $\lim_{n\to\infty} \frac{\log n}{n^{\varepsilon}} = 0$, которое доказывается с помощью правила Лопиталя.

c) Если $f(n)$ обозначает число k двоичных разрядов числа n, то, как следует из приведенных выше формул для k, $f(n) = O(\log n)$. Заметим, что такое же соотношение выполнено, если $f(n)$ — число разрядов в разложении n по произвольному фиксированному основанию b. С другой стороны, если основание b не фиксированно, а может расти, и $f(n, b)$ — число разрядов в записи по основанию b, то $f(n, b) = O\left(\frac{\log n}{\log b}\right)$.

d) Имеем $\mathrm{Time}(n \cdot m) = O(\log n \cdot \log m)$, где в левой части стоит число двоичных операций, требующихся для умножения n на m.

e) В примере 6 можно записать $\mathrm{Time}(n!) = O((n \log n)^2)$.

f) В примере 7

$$\mathrm{Time}\left(\sum a_i x^i \cdot \sum b_j x^j\right) = O\left(n_1 n_2((\log m)^2 + \log(\min(n_1, n_2)))\right).$$

Функции $f(n)$ и $f(n_1, n_2, \ldots, n_r)$ у нас будут часто обозначать время, которое требуется для решения некоторой арифметической задачи с целым числом n или с набором целых чисел $n_1, n_2, \ldots, n_r$ в качестве исходных данных. Мы хотим получить наши оценки в виде достаточно простых функций $g(n)$. При этом желательно, чтобы функции $g(n)$ не давали чрезмерно завышенное представление о времени решения задач (хотя с чисто математической точки зрения замена функции $g(n)$ в соотношении $f = O(g)$ любой большей функцией корректна).

Образно говоря, соотношение $f(n) = O(n^d)$ показывает, что функция f растет приблизительно как d-я степень аргумента. Например, если $d = 3$, то оно говорит нам, что удвоение аргумента приведет к увеличению функции приблизительно в 8 раз. Соотношение $f(n) = O(\log^d n)$ (где $\log^d n$ означает $(\log n)^d$) показывает, что функция возрастает приблизительно как d-я степень числа двоичных разрядов в n. Это так, потому что с точностью до мультипликативной константы число бит равно приблизительно $\log n$ (а именно, $\log n / \log 2 = 1{,}4427 \log n$). Так, например, если $f(n) = O(\log^3 n)$, то удвоение числа бит в n (что гораздо сильнее увеличивает аргумент, нежели его удвоение) приводит к увеличению $f(n)$ приблизительно в 8 раз.

Заметим, что запись $f(n) = O(1)$ означает, что функция f ограничена некоторой константой.

З а м е ч а н и е. Как мы видели, при умножении двух чисел приблизительно одинакового размера можно использовать формулу Time (k-разрядное двоичное число $\cdot$ k-разрядное двоичное число) $= O(k^2)$. Следует заметить, что было предпринято много усилий по повышению скорости умножения двух k-разрядных двоичных чисел при больших k. Используя специальные приемы, много более сложные, чем обычный школьный способ умножения, математики смогли придумать процедуру умножения двух целых чисел из k бит, требующую всего $O(k \log k \log\log k)$ двоичных операций. Это лучше, чем $O(k^2)$, и даже лучше, чем $O(k^{1+\varepsilon})$ при любом сколь угодно малом $\varepsilon > 0$. Однако дальше мы будем пользоваться только приведенными выше грубыми оценками для времени, необходимого для умножения.

В общем случае, когда оценивается число двоичных операций, требующихся для решения какой-либо задачи, сначала надо определить и выписать подробную процедуру решения задачи. Конкретная пошаговая процедура выполнения вычислений называется *алгоритмом*. Конечно, может существовать много разных алгоритмов, выполняющих одну и ту же работу. Можно воспользоваться тем из них, который попроще в записи, или тем, который быстрее работает, или выбрать какой-то компромисс между простотой и быстродействием. Использованный выше алгоритм умножения n на m далек от самого быстрого из известных. Но вместе с этим он много быстрее метода повторного сложения (m-кратного сложения числа n с собой).

П р и м е р 10. Оценить время, требующееся для перевода числа из k бит в десятичную систему счисления.

Р е ш е н и е. Пусть n — целое из k бит, записанное в двоичной системе счисления. Алгоритм перевода следующий. Разделим n на $10 = (1010)_2$. Остаток от деления — который является одним из чисел 0, 1, 10, 11, 100, 101, 110, 111, 1000 или 1001 — даст содержимое d_0 разряда единиц. Частное от деления возьмем вместо n, поделим на $(1010)_2$ и возьмем остаток от этого деления в качестве d_1, а частное — в качестве делимого при следующем делении на $(1010)_2$. Это процесс должен повторяться столько раз, сколько десятичных разрядов содержится в числе n, т. е. $\left[\frac{\log n}{\log 10}\right] + 1 = O(k)$ раз. Тогда процесс будет завершен. (Наши записи мы могли вести и в десятичной системе, используя более привычные обозначения для остатков $0, 1, 2, 3, \ldots, 9$ вместо $0, 1, 10, 11, \ldots, 1001$.) Как много двоичных операций будет сделано? Мы сделали $O(k)$ делений, каждое из них требовало $O(4k)$ операций (делимое содержит не более k бит, а де-

литель $(1010)_2$ — 4 бита). Но $O(4k)$ эквивалентно $O(k)$ (постоянный множитель несуществен в обозначении O-большое). Поэтому общее число двоичных операций равно $O(k)\cdot O(k) = O(k^2)$. Если желательно выразить оценку в терминах n, а не k, то, используя равенство $k = O(\log n)$, можно записать

$$\text{Time}(\text{перевод } n \text{ в десятичную систему счисления}) = O(\log^2 n).$$

Пример 11. Оценить время, требующееся для перевода числа n из k бит в систему счисления по основанию b, которое может быть очень большим.

Решение. Используя алгоритм примера 10 (только делим теперь на число b из l бит), получаем, что каждое деление выполняется дольше (если l велико) и требует $O(kl)$ двоичных операций. А сколько раз придется делить? Следует заметить, что число разрядов при записи n в b-ичной системе равно $O(k/l)$ (см. пример 9(c)). Поэтому общее число двоичных операций во всех делениях равно $O(k/l)\cdot O(kl) = O(k^2)$. Это тот же ответ, что и в примере 10. Значит, наша оценка для времени перевода числа в новую систему счисления не зависит от основания системы (сколь бы большим оно ни было). Это так, поскольку увеличение времени для определения содержимого каждого разряда компенсируется уменьшением числа этих разрядов.

Пример 12. Выразить при помощи O-большого время, требующееся для вычисления a) $n!$, b) $\binom{n}{m}$ (см. примеры 6 и 8).

Решение. a) $O(n^2\log^2 n)$, b) $O(m^2\log^2 n)$.

В завершение дадим одно определение, являющееся фундаментальным в вычислительных науках и в теории алгоритмов.

Определение. Алгоритм для проведения вычислений, определяющихся целыми числами $n_1, n_2, \ldots, n_r$ из $k_1, k_2, \ldots, k_r$ бит, соответственно, называется *полиномиальным по времени* алгоритмом, если существуют такие целые числа $d_1, d_2, \ldots, d_r$, что число двоичных операций при работе этого алгоритма равно $O\left(k_1^{d_1}k_2^{d_2}\cdots k_r^{d_r}\right)$.

Таким образом, обычные арифметические операции $+$, $-$, $\times$, $\div$ дают примеры полиномиальных по времени алгоритмов. Еще один пример — перевод чисел из одной системы счисления в другую. С другой стороны, вычисление $n!$ не является такой операцией. (Однако, если требуется найти лишь заданное число значащих цифр, например, первые 1000 двоичных разрядов, то можно найти $n!$ за полиномиальное время при помощи формулы Стирлинга.)

УПРАЖНЕНИЯ

1. Умножить $(212)_3$ на $(122)_3$.

2. Разделить $(40122)_7$ на $(126)_7$.

3. Перемножить двоичные числа 101101 и 11001 и разделить 10011001 на 1011.

4. В системе счисления по основанию 26, представляя 0–25 буквами A–Z: а) умножить YES на NO, б) разделить JQVXHJ на WE.

5. Записать число $e = 2,7182818\ldots$ а) в двоичной системе счисления (с 15 знаками после запятой), б) по основанию 26 (с 3 знаками после запятой).

6. Под чисто периодической дробью с периодом f в b-ичной системе понимается число между 0 и 1, в b-ичной записи которого после запятой стоит периодическая последовательность периода f. Например, в десятичной системе счисления $1/3$ — чисто периодическая дробь с периодом 1, а $1/7$ — чисто периодическая дробь с периодом 6. Доказать, что несократимая дробь c/d между 0 и 1 является чисто периодической с периодом f в b-ичной системе тогда и только тогда, когда $b^f - 1$ есть кратное числа d.

7. а) В шестнадцатиричной системе с $b = 16$ буквы A–F изображают цифры от 10 до 15, соответственно. Разделить $(131B6C3)_{16}$ на $(1A2F)_{16}$.

б) Объяснить, как переводятся числа из двоичной системы в шестнадцатиричную и обратно и почему эти операции требуют много меньше времени, чем в полученной в примере 11 общей оценке для перевода из двоичной системы в b-ичную.

8. Описать двоичные операции при вычитании битов аналогично тому, как это было сделано в тексте для сложения (список из 5 возможностей).

9. а) Используя обозначение O-большое, оценить в терминах простой функции от n число двоичных операций при вычислении 3^n в двоичной системе.

б) То же самое для n^n.

10. Оценить в терминах простой функции от n и N число двоичных операций при вычислении N^n.

11. Для суммы квадратов первых n натуральных чисел справедлива формула

$$\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}.$$

а) Используя обозначение O-большое, оценить в терминах n число двоичных операций, требующихся для вычисления левой части этого равенства.

б) Оценить число двоичных операций, требующихся для вычисления правой части этого равенства.

12. Используя обозначение O-большое, оценить число двоичных операций, требующихся для умножения $r \times n$-матрицы на $n \times s$-матрицу, если все элементы матриц не превосходят m.

13. Цель этого упражнения — оценить в терминах n число двоичных операций, требующихся для вычисления произведения всех простых чисел, меньших n. При этом предполагается, что список всех простых чисел, не превосходящих n, уже составлен.

а) Обозначим $\pi(n)$ число простых чисел, не превосходящих n. Согласно теореме о распределении простых чисел $\pi(n)$ имеет асимптотику $n/\log n$. Это означает, что $\frac{\pi(n)}{n/\log n}$ стремится к 1 при $n \to \infty$. Используя этот факт, оценить число двоичных разрядов в произведении всех простых чисел, меньших n.

б) Найти границу для числа двоичных операций при любом из умножений, которые производятся при вычислении этого произведения.

в) Оценить число двоичных операций, требующихся для вычисления произведения всех простых чисел, меньших n.

14. а) Пусть надо проверить простоту числа n при помощи последовательного деления на все нечетные числа, не превосходящие $\sqrt{n}$. Оценить число двоичных операций.

б) В постановке п. а) дополнительно предположим, что имеется список простых чисел, не превосходящих $\sqrt{n}$. Тогда можно ограничиться делением на числа из этого списка (а не делить на все нечетные). Дать оценку числа двоичных операций в этом случае. Использовать теорему о распределении числа простых чисел.

15. Оценить время, необходимое для проверки того, имеет ли число n простой делитель, не превосходящий m. Предположим, что имеется список всех простых чисел, не превосходящих m. Опять воспользоваться теоремой о распределении числа простых чисел.

16. Пусть n — очень большое целое число, записанное в двоичной системе счисления. Найти простой алгоритм вычисления $[\sqrt{n}]$ за $O(\log^3 n)$ двоичных операций. Здесь [ ] — целая часть числа.

§ 2. Делимость и алгоритм Евклида

Делители и делимость. Для данных целых чисел a и b говорят, что a *делит* b (или b *делится на* a), и используют обозначение $a|b$, если существует такое целое число d, что $b = ad$. В этом случае a называют *делителем* b. Любое целое число $b > 1$ имеет, по крайней мере, два положительных делителя: 1 и b. Под *собственным делителем* b подразумевают любой положительный делитель, не равный b, а под *нетривиальным делителем* b — любой положительный делитель, не равный 1 или b. *Простым числом*, по определению, является целое число, большее 1, которое не имеет положительных делителей, отличных от 1 и самого себя; число называется *составным*, если оно имеет, по крайней мере, один нетривиальный делитель. Следующие свойства делимости легко выводятся непосредственно из определений:

1. Если $a|b$ и c — любое целое число, то $a|bc$.
2. Если $a|b$ и $b|c$, то $a|c$.
3. Если $a|b$ и $a|c$, то $a|b \pm c$.

Для простого p и целого неотрицательного числа α мы используем запись $p^\alpha \| b$, подразумевая, что p^α — наивысшая степень p, делящая b, т.е. $p^\alpha | b$, а $p^{\alpha+1} \not| b$. В этом случае говорят, что p^α *точно делит* b.

Основная теорема арифметики утверждает, что любое натуральное число n может быть записано в виде произведения простых чисел единственным образом (с точностью до перестановки сомножителей). Принято записывать это разложение в виде произведения различных простых сомножителей в соответствующих степенях, располагая простые числа в порядке возрастания. Например,

$4200 = 2^3 \cdot 3 \cdot 5^2 \cdot 7$.

Следующие два свойства делимости вытекают из основной теоремы.

4. Если простое число p делит ab, то $p|a$ или $p|b$.

5. Если $m|a$ и $n|a$ и если m и n не имеют общих делителей, больших 1, то $mn|a$.

Из единственности разложения целых чисел на простые множители следует также простой способ отыскания всех делителей n по его разложению. А именно, любой делитель d числа n должен быть произведением тех же простых сомножителей в степенях, не превышающих степени, точно делящие n. Таким образом, если $p^\alpha \| n$, то $p^\beta \| d$ для некоторого β, удовлетворяющего условию $0 \leqslant \beta \leqslant \alpha$. Например, для нахождения делителей числа 4200 нужно взять 2 в степени 0, 1, 2 или 3, умножить на 3 в степени 0 или 1, на 5 в степени 0, 1 или 2 и на 7 в степени 0 или 1. Число всех сомножителей, таким образом, есть произведение числа способов выбора степени для каждого простого сомножителя, которое, в свою очередь, равно $\alpha + 1$. Иначе говоря, число $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ имеет $(\alpha_1 + 1)(\alpha_2 + 1) \cdots (\alpha_r + 1)$ различных делителей. В частности, у числа 4200 их 48.

Наибольшим общим делителем двух данных целых чисел a и b (обозначаемым НОД (a, b) или просто (a, b)), не равных одновременно нулю, называется наибольшее целое число d, делящее и a, и b. Нетрудно доказать, что приведенное выше определение эквивалентно следующему: НОД (a, b) — это единственное положительное число, которое делит a и b и делится на любой другой их общий делитель.

Если имеются разложения на простые множители двух чисел a и b, — легко записать и НОД (a, b). Следует взять все простые числа, входящие в оба разложения, и каждое возвести в степень, равную минимуму из двух соответствующих показателей. Например, сравнивая разложение $10780 = 2^2 \cdot 5 \cdot 7^2$ с приведенным выше разложением числа 4200, получаем, что НОД $(4200, 10780) = 2^2 \cdot 5 \cdot 7 = 140$.

Наименьшее общее кратное чисел a и b обозначается НОК (a, b) и по определению является наименьшим целым положительным числом, делящимся как на a, так и на b. Имея разложение на простые множители чисел a и b, можно получить НОК (a, b), взяв все простые числа, входящие *хотя бы в одно* из разложений, и каждое возвести в степень, равную максимуму из двух показателей. Легко показать, что НОК $(a, b) = |ab|/$НОД (a, b).

Алгоритм Евклида. При работе с большими числами часто случается, что их разложение на простые множители неизвестно. В частности, важным направлением исследований в теории чисел явля-

ется отыскание быстрых методов разложения больших чисел на простые множители. К счастью, существует сравнительно быстрый способ нахождения НОД (a, b) даже в том случае, когда неизвестны простые делители a или b. Он называется *алгоритмом Евклида.*

Алгоритм Евклида работает следующим образом. Чтобы найти НОД (a, b), где $a > b$, сначала делят a на b и записывают частное q_1 и остаток r_1: $a = q_1 b + r_1$. Затем производят второе деление, в котором b играет роль a, а r_1 — роль b: $b = q_2 r_1 + r_2$. Затем делят r_1 на r_2. Деление продолжают, каждый раз деля предпоследний остаток на последний, получая новые частное и остаток. Когда в конце концов получается остаток, который является делителем предыдущего остатка, то этот последний ненулевой остаток и есть наибольший общий делитель чисел a и b.

Пример 1. Найти НОД $(1547, 560)$.

Решение.

$$\begin{aligned} 1547 &= 2 \cdot 560 + 427 \\ 560 &= 1 \cdot 427 + 133 \\ 427 &= 3 \cdot 133 + 28 \\ 133 &= 4 \cdot 28 + 21 \\ 28 &= 1 \cdot 21 + 7. \end{aligned}$$

Так как $7|21$, то НОД $(1547, 560) = 7$.

Предложение I. 2. 1. *Алгоритм Евклида всегда дает наибольший общий делитель за конечное число шагов. Кроме того, для $a > b$*

$$\text{Time}(\text{нахождение НОД } (a, b) \text{ по алгоритму Евклида}) = O(\log^3(a)).$$

Доказательство. Доказательство первого утверждения детально изложено во многих учебниках по элементарной теории чисел, поэтому изложим его кратко. Во-первых, легко показать, что остатки строго уменьшаются с каждым шагом и, следовательно, в конце концов достигнут нуля. Чтобы убедиться, что последний остаток есть НОД, воспользуемся вторым определением НОД. Получаем, что если некоторое число делит и a, и b, то оно должно делить r_1, а так как оно делит b и r_1, то оно должно делить r_2, и т. д. В конце концов, получаем, что оно должно делить последний ненулевой остаток. С другой стороны, переходя от последнего шага к предыдущему, легко видеть, что последний остаток должен делить все предыдущие остатки, а также a и b. Итак, это НОД, так как НОД — это единственное число, которое делит a и b и в то же время делится на любой их общий делитель.

Теперь докажем второе утверждение. Главный вопрос состоит в том, сколько раз производится деление. Покажем, что остатки не просто убывают, а убывают довольно быстро.

Лемма. $r_{j+2} < r_j/2$.

Доказательство леммы. Если $r_{j+1} \leqslant r_j/2$, то сразу получаем $r_{j+2} < r_{j+1} \leqslant r_j/2$. Пусть теперь $r_{j+1} > r_j/2$. В этом случае следующее деление дает $r_j = 1 \cdot r_{j+1} + r_{j+2}$ и $r_{j+2} = r_j - r_{j+1} < r_j/2$. Лемма доказана.

Вернемся к доказательству теоремы. Так как за каждые два шага остаток уменьшается, по крайней мере, вдвое и так как остаток не может стать меньше 1, то производится самое большее $2[\log_2 a]$ делений. Эта величина есть $O(\log a)$. Каждое деление производится над числами, не превышающими a, и, следовательно, проводится за $O(\log^2 a)$ двоичных операций. Таким образом, общее время работы составляет $O(\log a) \cdot O(\log^2 a) = O(\log^3 a)$, что и требовалось доказать.

Замечание. Если провести более тщательный подсчет числа двоичных операций, учитывая уменьшение чисел, участвующих в делении, можно улучшить оценку времени работы алгоритма Евклида до $O(\log^2 a)$.

Предложение I.2.2. *Пусть* $d = \text{НОД}\,(a, b)$, *где* $a > b$. *Тогда существуют такие целые числа* u *и* v, *что* $d = ua + bv$. *Другими словами, НОД двух чисел можно представить в виде линейной комбинации этих чисел с целыми коэффициентами. Кроме того, для нахождения чисел* u *и* v *достаточно* $O(\log^3 a)$ *двоичных операций.*

Схема доказательства. Проходя последовательность равенств в алгоритме Евклида снизу вверх и выражая каждый раз d через все более ранние остатки, в конце концов получаем выражение d через a и b. На каждом шаге необходимо произвести одно умножение и одно сложение или вычитание. Таким образом, как легко убедиться, число двоичных операций снова есть $O(\log^3 a)$.

Пример 1 (продолжение). Чтобы представить 7 в виде линейной комбинации 1547 и 560, запишем

$$\begin{aligned} 7 &= 28 - 1 \cdot 21 = 28 - 1 \cdot (133 - 4 \cdot 28) \\ &= 5 \cdot 28 - 1 \cdot 133 = 5 \cdot (427 - 3 \cdot 133) - 1 \cdot 133 \\ &= 5 \cdot 427 - 16 \cdot 133 = 5 \cdot 427 - 16 \cdot (560 - 1 \cdot 427) \\ &= 21 \cdot 427 - 16 \cdot 560 = 21 \cdot (1547 - 2 \cdot 560) - 16 \cdot 560 \\ &= 21 \cdot 1547 - 58 \cdot 560. \end{aligned}$$

Определение. Говорят, что два целых числа a и b *взаимно просты*, если $\text{НОД}\,(a, b) = 1$, т.е. если a и b не имеют общих делителей, больших 1.

Следствие. *Если a и b — взаимно простые числа и $a > b$, то можно найти представление 1 в виде целочисленной линейной комбинации этих чисел за полиномиальное время, точнее, за $O(\log^3 a)$ двоичных операций.*

О п р е д е л е н и е. Для любого целого положительного n *функция Эйлера* $\varphi(n)$ определяется как число неотрицательных целых b, меньших n и взаимно простых с n:

$$\varphi(n) \underset{\text{def}}{=} |\{0 \leqslant b < n \mid \text{НОД}\ (b, n) = 1\}|.$$

Легко проверить, что $\varphi(1) = 1$ и что $\varphi(p) = p - 1$ для любого простого p. Можно убедиться также, что для любого простого p

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha \left(1 - \frac{1}{p}\right).$$

Для этого достаточно заметить, что числа от 0 до $p^\alpha - 1$, которые *не* взаимно просты с p^α, — это в точности те числа, которые делятся на p, а их количество равно $p^{\alpha-1}$.

В следующем параграфе главе будет показано, что функция Эйлера $\varphi(n)$ обладает «мультипликативными свойствами», что дает возможность быстро вычислять значение $\varphi(n)$, используя разложение числа n на простые множители. А именно, если n записано в виде произведения степеней различных простых чисел p^α, то $\varphi(n)$ является произведением чисел $\varphi(p^\alpha)$.

УПРАЖНЕНИЯ

1. а) Доказать следующие свойства отношения $p^\alpha \| b$: (i) если $p^\alpha \| a$ и $p^\beta \| b$, то $p^{\alpha+\beta} \| ab$; (ii) если $p^\alpha \| a$, $p^\beta \| b$ и $\alpha < \beta$, то $p^\alpha \| a \pm b$.

б) Найти пример, опровергающий утверждение: если $p^\alpha \| a$ и $p^\alpha \| b$, то $p^\alpha \| a+b$.

2. Сколько делителей имеет число 945? Перечислить их все.

3. Пусть n — положительное нечетное число.

а) Доказать, что существует взаимно однозначное соответствие между такими делителями числа n, которые меньше $\sqrt{n}$, и такими, которые больше $\sqrt{n}$. (В этой части n может быть и четным.)

б) Доказать, что существует взаимно однозначное соответствие между множеством тех делителей числа n, которые не меньше $\sqrt{n}$, и множеством представлений n в виде разности $s^2 - t^2$ квадратов двух неотрицательных целых чисел. (Например, 15 имеет два делителя 5, 15, которые не меньше $\sqrt{15}$, и $15 = 4^2 - 1^2 = 8^2 - 7^2$.)

в) Укажите все возможные способы записи числа 945 в виде разности квадратов двух неотрицательных целых чисел.

4. а) Показать, что степень простого p, точно делящая $n!$, равна $[n/p]+[n/p^2]+[n/p^3]+\cdots$. (Заметим, что эта сумма конечна.)

б) Найти степени чисел 2, 3, 5, 7, точно делящие 100!, а затем выписать разложение 100! на простые множители.

в) Пусть $S_b(n)$ означает сумму цифр числа n в системе счисления по основанию b. Доказать, что степень числа 2, точно делящая $n!$, равна $n - S_2(n)$. Вывести аналогичную формулу для произвольного простого p.

5. Найти $d = \text{НОД}\,(360, 294)$ двумя способами: а) найдя разложение обоих чисел на простые множители, найти затем разложение d; и б) при помощи алгоритма Евклида.

6. При помощи алгоритма Евклида найти наибольший общий делитель для следующих четырех пар чисел и выразить его как линейную комбинацию этих чисел с целыми коэффициентами. а) 26, 19. б) 187, 34. в) 841, 160. г) 2613, 2171.

7. Часто можно ускорить поиск НОД по алгоритму Евклида, если использовать деление с отрицательным остатком, т. е. выбирать из равенств $r_j = q_{j+2}r_{j+1} - r_{j+2}$ и $r_j = q_{j+2}r_{j+1} + r_{j+2}$ то, в котором значение r_{j+2} наименьшее. При этом всегда будет $r_{j+2} \leqslant r_{j+1}/2$. Решить все четыре примера из упражнения 6, используя этот метод.

8. а) Доказать, что следующий алгоритм находит $d = \text{НОД}\,(a, b)$ за конечное число шагов. Сначала заметим, что $\text{НОД}\,(a, b) = \text{НОД}\,(|a|, |b|)$, т. е. без ограничения общности можно предполагать, что a и b — положительные числа. Если a и b оба четные, положим $d = 2d'$, где $d' = \text{НОД}\,(a/2, b/2)$. Если одно из этих двух чисел нечетное, а другое (скажем, b) четное, то положим $d = d'$, где $d' = \text{НОД}\,(a, b/2)$. Если оба числа нечетные и различные, скажем, $a > b$, то положим $d = d'$, где $d' = \text{НОД}\,(a - b, b)$. Наконец, если $a = b$, положим $d = a$. Повторяем этот процесс, пока не получаем последний случай (когда числа равны между собой).

б) Воспользоваться алгоритмом из пункта а) для нахождения НОД (2613, 2171), производя все действия в двоичной системе счисления, т. е. найти

$$\text{НОД}\,((101000110101)_2, (100001111011)_2).$$

в) Доказать, что алгоритм в пункте а) требует всего $O(\log^2 a)$ двоичных операций (если $a > b$).

г) Почему этот алгоритм в приведенном выше виде не всегда лучше алгоритма Евклида?

9. Пусть a много больше b. Найти временную оценку сложности вычисления НОД (a, b) в терминах O-большого, которая была бы лучше, чем $O(\log^3 a)$.

10. Цель этой задачи — найти «наилучшую» оценку для числа делений в алгоритме Евклида. *Числа Фибоначчи* можно определить правилом $f_1 = 1$, $f_2 = 1$, $f_{n+1} = f_n + f_{n-1}$ для $n \geqslant 2$, или, что то же самое, матричным равенством

$$\begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n.$$

а) Допустим, что $a > b > 0$ и для нахождения НОД (a, b) по алгоритму Евклида (в стандартном варианте с неотрицательными остатками) требуется k делений. Показать, что $a \geqslant f_{k+2}$.

б) Используя матричное определение для f_n, доказать, что

$$f_n = \frac{\alpha^n - \alpha'^n}{\sqrt{5}}, \quad \text{где} \quad \alpha = \frac{1 + \sqrt{5}}{2}, \quad \alpha' = \frac{1 - \sqrt{5}}{2}.$$

в) Используя пункты а) и б), найти верхнюю оценку для k в терминах a. Сравнить с оценкой, получающейся из доказательства предложения I. 2. 1.

11. Цель этой задачи — найти оценку времени, необходимого для вычисления НОД (a, b) (где $a > b$), которая лучше оценки, полученной в предложении I. 2. 1.

а) Показать, что число двоичных операций, требующихся для проведения деления $a = qb + r$, есть $O((\log b)(1 + \log q))$.

б) Применяя результат а) к каждому из $O(\log a)$ делений вида $r_{i-1} = q_{i+1}r_i + r_{i+1}$, получить временную оценку в виде $O((\log b)(\log a))$.

12. Рассмотрим многочлены с вещественными коэффициентами (можно рассматривать ту же задачу для многочленов с коэффициентами из произвольного *поля*). Пусть есть два многочлена f и g. Будем говорить, что $f|g$, если существует такой многочлен h, что $g = fh$. Определим НОД (f, g) аналогично тому, как это делалось для целых чисел, а именно, как многочлен наибольшей степени, делящий f и g. Многочлен НОД (f, g) определяется не однозначно, так как, умножая его на любую ненулевую константу, можно получить другой многочлен с теми же свойствами. Однако единственности можно добиться, если потребовать, чтобы многочлен НОД был *нормированным*, т. е. имел старший коэффициент 1. Многочлены f и g называются взаимно простыми, если их НОД есть многочлен-константа 1. Постройте процедуру нахождения НОД для многочленов, полностью аналогичную алгоритму Евклида для целых чисел, и примените ее для нахождения а) НОД $(x^4 + x^2 + 1, x^2 + 1)$ и б) НОД $(x^4 - 4x^3 + 6x^2 - 4x + 1, x^3 - x^2 + x - 1)$. В каждом случае найдите многочлены $u(x)$ и $v(x)$, представляющие НОД в виде $u(x)f(x) + v(x)g(x)$.

13. Из алгебры известно, что многочлен имеет кратный корень тогда и только тогда, когда он имеет общий делитель со своей производной, в этом случае кратные корни многочлена $f(x)$ являются корнями многочлена НОД (f, f'). Найти кратные корни многочлена $x^4 - 2x^3 - x^2 + 2x + 1$.

14. (Прежде, чем приступить к этому упражнению, вспомните правила действий с комплексными числами. Заметьте, что так как $(a + bi)(a - bi)$ есть вещественное число $a^2 + b^2$, то можно производить деление по формуле $(c+di)/(a+bi) = (c + di)(a - bi)/(a^2 + b^2)$.) *Гауссовыми числами* называются комплексные числа, у которых действительная и мнимая части суть целые числа. На комплексной плоскости они изображаются вершинами квадратов, которые образуют целочисленную решетку. Если α и β — гауссовы числа, то говорят, что $\alpha|\beta$, если существует такое гауссово число γ, что $\beta = \alpha\gamma$. Определим НОД (α, β) как максимальное по модулю гауссово число δ, делящее как α, так и β. (Напомним, что модуль $|\delta|$ — это расстояние от точки δ на комплексной плоскости до точки 0, т. е. квадратный корень из суммы квадратов действительной и мнимой частей). НОД определен не единственным образом, потому что его можно умножить на ± 1 или $\pm i$ и получить другое δ с тем же модулем и тоже делящее и α, и β. Это дает четыре возможности. Мы будем рассматривать любое из этих четырех чисел как НОД.

Заметим, что любое комплексное число может быть записано в виде суммы гауссова числа и комплексного числа, действительная и мнимая части которого заключены в промежутке от $-1/2$ до $1/2$. Показать, что поэтому мы можем разделить одно гауссово число α на другое гауссово число β и получить частное и остаток в виде гауссовых чисел, причем остаток по модулю будет меньше β. Используйте этот факт, чтобы сформулировать алгоритм Евклида для получения НОД двух гауссовых чисел. Воспользуйтесь этим алгоритмом для нахождения а) НОД $(5 + 6i, 3 - 2i)$ и б) НОД $(7 - 11i, 8 - 9i)$. В каждом случае представьте НОД в виде линейной комбинации $u\alpha + v\beta$, где u и v являются гауссовыми числами.

15. Последняя задача дает эффективный способ представления некоторых больших простых чисел в виде суммы двух квадратов. Например, предположим,

что p — простое число, делящее число вида b^6+1. Мы хотим записать p в виде $p=c^2+d^2$, где c и d — целые. Это эквивалентно нахождению нетривиального гауссова делителя числа p, так как $c^2+d^2=(c+di)(c-di)$. Мы можем действовать следующим образом. Заметим, что

$$b^6+1=(b^2+1)(b^4-b^2+1) \quad \text{и} \quad b^4-b^2+1=(b^2-1)^2+b^2.$$

По четвертому свойству делимости простое число p должно делить один из множителей в правой части первого равенства. Если $p|(b^2+1)=(b+i)(b-i)$, то НОД $(p, b+i)$ дает нам нужное $c+di$. Если $p|(b^4-b^2+1)=((b^2-1)+bi)((b^2-1)-bi)$, то НОД $(p,(b^2-1)+bi)$ дает нам $c+di$.

Пример. Простое число 12277 делит второй сомножитель в произведении $20^6+1=(20^2+1)(20^4-20^2+1)$. Поэтому найдем НОД $(12277, 399+20i)$:

$$12277=(31-2i)(399+20i)+(-132+178i),$$

$$399+20i=(-1-i)(-132+178i)+(89+66i),$$

$$-132+178i=(2i)(89+66i).$$

Таким образом, НОД есть $89+66i$, т. е. $12277=89^2+66^2$.

а) Воспользовавшись равенством $19^6+1=2\cdot 13^2\cdot 181\cdot 769$ и алгоритмом Евклида для гауссовых чисел, представить 769 в виде суммы двух квадратов.

б) Аналогично представить простое число 3877 в виде суммы двух квадратов, зная, что 3877 делит 15^6+1.

в) Зная, что простое число 38737 делит $2^{36}+1$, представить 38737 в виде суммы двух квадратов.

§ 3. Сравнения

Основные свойства. Пусть даны три целых числа a, b и m; говорят, что *a сравнимо с b по модулю m*, если разность $a-b$ делится на m. Записывают это так: $a\equiv b \pmod m$. Число m называется *модулем* сравнения. Из определения легко выводятся следующие свойства:

1. (i) $a\equiv a \pmod m$; (ii) $a\equiv b \pmod m$ тогда и только тогда, когда $b\equiv a \pmod m$; (iii) если $a\equiv b \pmod m$ и $b\equiv c \pmod m$, то $a\equiv c \pmod m$. Для фиксированного m свойства (i)–(iii) означают, что сравнимость по модулю m есть *отношение эквивалентности*.

2. Для фиксированного m каждый *класс эквивалентности* по этому отношению обладает в точности одним представителем в множестве чисел от 0 до $m-1$. (Другими словами, любое целое число сравнимо по модулю m ровно с одним целым числом в промежутке от 0 до $m-1$). Множество классов эквивалентности (они называются *классами вычетов*) будет обозначаться $\mathbf{Z}/m\mathbf{Z}$. Любое множество представителей (по одному от каждого класса — *прим. перев.*) называется *полной системой вычетов по модулю m*.

3. Если $a \equiv b \pmod m$ и $c \equiv d \pmod m$, то $a \pm c \equiv b \pm d \pmod m$ и $ac \equiv bd \pmod m$. Другими словами, сравнения (по одному и тому же модулю) можно складывать, вычитать и перемножать. Таким образом, множество $\mathbf{Z}/m\mathbf{Z}$ является *коммутативным кольцом*, т. е. классы вычетов можно складывать, вычитать и перемножать (причем результат не зависит от того, какие представители классов эквивалентности используются), и эти операции удовлетворяют обычным аксиомам ассоциативности, коммутативности, существования противоположного элемента и т. д.

4. Если $a \equiv b \pmod m$, то $a \equiv b \pmod d$ для любого делителя $d|m$.

5. Если $a \equiv b \pmod m$, $a \equiv b \pmod n$ и если m и n взаимно просты, то $a \equiv b \pmod{mn}$. (См. пятое свойство делимости в § I.2.)

Предложение I.3.1. *Обратимыми по умножению являются те элементы из* $\mathbf{Z}/m\mathbf{Z}$, *представители которых взаимно просты с* m, *т. е. числа* a, *для которых существует такое* b, *что* $ab \equiv 1 \pmod m$, — *это те и только те числа* a, *для которых* НОД $(a, m) = 1$. *Кроме того, если* НОД $(a, m) = 1$, *то обратный элемент* b *может быть найден за* $O(\log^3 m)$ *двоичных операций.*

Доказательство. Если бы $d =$ НОД (a, m) был больше 1, то ни для какого b сравнение $ab \equiv 1 \pmod m$ не могло бы выполняться, так как в противном случае d делил бы $ab - 1$ и, следовательно, $d|1$. Обратно, если НОД $(a, m) = 1$, то согласно свойству 2 можно считать, что $a < m$. Тогда в соответствии с предложением I.2.2 существуют целые числа u и v, которые можно найти за $O(\log^3 m)$ двоичных операций и для которых $au + vm = 1$. Полагая $b = u$, получаем, что $m|1 - au = 1 - ab$, что и требовалось.

Замечание. Если НОД $(a, m) = 1$, то под отрицательной степенью $a^{-n} \pmod n$ подразумевают n-ю степень обратного класса вычетов, т. е. класс вычетов, содержащий n-ю степень любого целого числа b такого, что $ab \equiv 1 \pmod m$.

Пример 1. Найти $160^{-1} \pmod{841}$, т. е. элемент, обратный к 160 по модулю 841.

Решение. Воспользовавшись результатом упражнения 6 с) из предыдущего раздела, получим ответ 205.

Следствие 1. *Если* p — *простое число, то каждый ненулевой класс вычетов по модулю* p *имеет обратный, который может быть найден за* $O(\log^3 p)$ *двоичных операций. Кольцо* $\mathbf{Z}/p\mathbf{Z}$ *является полем. Мы часто будем обозначать это поле* $\mathbf{F}_p$ *и называть полем из* p *элементов.*

Следствие 2. *Пусть требуется решить линейное сравнение* $ax \equiv b \pmod m$; *не теряя общности, можно считать, что* $0 \leqslant$

$a, b < m$. *Тогда если* НОД $(a, m) = 1$, *то сравнение имеет решение* x_0, *которое можно найти за* $O(\log^3 m)$ *двоичных операций, и любое решение имеет вид* $x = x_0 + mn$, *где* n — *некоторое целое число. Если же* НОД $(a, m) = d$, *то сравнение имеет решение тогда и только тогда, когда* $d|b$, *и в этом случае данное сравнение эквивалентно* (*в смысле совпадения множества решений*) *сравнению* $a'x \equiv b' \pmod{m'}$, *где* $a' = a/d$, $b' = b/d$, $m' = m/d$.

Первое следствие есть просто частный случай предложения I. 3. 1. Второе следствие легко вывести из предложения I. 3. 1 и определений. Так же, как при решении аналогичных линейных уравнений с вещественными коэффициентами, для решения линейных уравнений в $\mathbf{Z}/m\mathbf{Z}$ умножают обе части на элемент, мультипликативно обратный коэффициенту при неизвестном.

Вообще, при вычислениях по модулю m аналогом понятия «ненулевой» является «взаимно простой с m». Выше было показано, что так же, как уравнения, сравнения можно складывать, вычитать и умножать (см. свойство 3 для сравнений). Их можно также делить, если «знаменатель» взаимно прост с m.

Следствие 3. *Если* $a \equiv b \pmod{m}$, $c \equiv d \pmod{m}$ *и* НОД $(c, m) = 1$ (*в этом случае, конечно,* НОД $(d, m) = 1$), *то* $ac^{-1} \equiv bd^{-1} \pmod{m}$ (*где* c^{-1} *и* d^{-1} *обозначают целые числа, являющиеся обратными к* c *и* d *по модулю* m).

Для доказательства следствия 3 рассмотрим сравнения $c(ac^{-1} - bd^{-1}) \equiv (acc^{-1} - bdd^{-1}) \equiv a - b \equiv 0 \pmod{m}$, и так как m не имеет общих делителей с c, то m должно делить $ac^{-1} - bd^{-1}$.

Предложение I. 3. 2 (Малая теорема Ферма). *Пусть* p — *простое число. Любое целое число* a *удовлетворяет сравнению* $a^p \equiv a \pmod{p}$, *и любое целое число* a, *не делящееся на* p, *удовлетворяет сравнению* $a^{p-1} \equiv 1 \pmod{p}$.

Доказательство. Предположим сначала, что $p \nmid a$. Покажем, что $0a, 1a, 2a, 3a, \ldots, (p-1)a$ есть полная система вычетов по модулю p. В самом деле, если бы какие-либо два из этих чисел, скажем ia и ja, принадлежали одному классу вычетов, т. е. $ia \equiv ja \pmod{p}$, это означало бы, что $p|(i-j)a$, а так как a не делится на p, мы получили бы $p|i-j$. Так как i и j меньше p, это возможно лишь в случае $i = j$. Таким образом, числа $a, 2a, \ldots, (p-1)a$, рассматриваемые по модулю p, являются перестановкой чисел $1, 2, \ldots, p-1$. Отсюда следует, что произведение всех чисел первого набора сравнимо по модулю p с произведением чисел второго набора, т. е. $a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$. Получаем, что $p|(p-1)!(a^{p-1}-1)$. Так как $(p-1)!$ не делится на p, получим $p|a^{p-1}-1$, что и требовалось доказать. Наконец, если умно-

жить обе части сравнения $a^{p-1} \equiv 1 \pmod{p}$ на a, получится первое сравнение из утверждения теоремы для случая, когда a не делится на p. Если же a делится на p, то сравнение $a^p \equiv a \pmod{p}$ является тривиальным, так как обе части его сравнимы с нулем по модулю p. Это завершает доказательство предложения.

Следствие. *Если a не делится на p и если $n \equiv m \pmod{(p-1)}$, то $a^n \equiv a^m \pmod{p}$.*

Доказательство следствия. Пусть $n > m$. Так как $p-1 | n-m$, то $n = m + c(p-1)$ для некоторого положительного целого c. Умножив почленно c сравнений $a^{p-1} \equiv 1 \pmod{p}$ и сравнение $a^m \equiv a^m \pmod{p}$, получим искомый результат: $a^n \equiv a^m \pmod{p}$.

Пример 2. Найти последнюю цифру в семиричной записи числа $2^{1000000}$.

Решение. Пусть $p = 7$. Так как 1000000 при делении на $p - 1 = 6$ дает остаток 4, получаем $2^{1000000} \equiv 2^4 = 16 \equiv 2 \pmod{7}$, и, значит, последняя цифра равна 2.

Предложение I.3.3 (Китайская теорема об остатках). *Пусть требуется решить систему сравнений по различным модулям:*

$$\begin{aligned} x &\equiv a_1 \pmod{m_1}, \\ x &\equiv a_2 \pmod{m_2}, \\ &\cdots\cdots \\ x &\equiv a_r \pmod{m_r}, \end{aligned}$$

причем любые два модуля взаимно просты: НОД $(m_i, m_j) = 1$ *для $i \neq j$. Тогда эта система разрешима и любые два решения сравнимы по модулю $M = m_1 m_2 \cdots m_r$.*

Доказательство. Сначала докажем единственность по модулю M (последнее утверждение теоремы). Пусть x' и x'' — два решения системы. Положим $x = x' - x''$. Тогда x сравним с нулем по любому модулю m_i, а значит, и по модулю M (по пятому свойству сравнений). Теперь покажем, как найти решение x.

Обозначим через $M_i = M/m_i$ произведение всех модулей, кроме i-го. Очевидно, что НОД $(m_i, M_i) = 1$ и, следовательно, существует такое целое N_i, что $M_i N_i \equiv 1 \pmod{m_i}$ (число N_i может быть найдено, например, по алгоритму Евклида). Положим теперь $x = \sum_i a_i M_i N_i$. Тогда для каждого i все слагаемые в этой сумме, за исключением i-го, делятся на m_i, так как $m_i | M_j$ для всех $i \neq j$. Таким образом, для каждого i имеем $x \equiv a_i M_i N_i \equiv a_i \pmod{m_i}$, что и требовалось доказать.

Следствие. *Функция Эйлера обладает свойством «мультипликативности», т. е. $\varphi(mn) = \varphi(m)\varphi(n)$, если* НОД $(m, n) = 1$.

Доказательство следствия. Для доказательства необходимо подсчитать количество целых чисел между нулем и $mn-1$, не имеющих общих делителей с mn. Для каждого j из этого множества обозначим через j_1 наименьший неотрицательный вычет по модулю m (т.е. $0 \leqslant j_1 < m$ и $j \equiv j_1 \pmod m$) и через j_2 наименьший неотрицательный вычет по модулю n (т.е. $0 \leqslant j_2 < n$ и $j \equiv j_2 \pmod n$). Из китайской теоремы об остатках следует, что каждой паре j_1, j_2 соответствует одно и только одно число j в промежутке от 0 до $mn-1$, для которого $j \equiv j_1 \pmod m$ и $j \equiv j_2 \pmod n$. Заметим, что число j не имеет общих делителей с mn тогда и только тогда, когда оно не имеет общих делителей с m (это эквивалентно взаимной простоте j_1 и m) и не имеет общих делителей с n (что эквивалентно взаимной простоте j_2 и n). Таким образом, числа j, взаимно простые с mn, находятся во взаимно однозначном соответствии с парами j_1, j_2, для которых $0 \leqslant j_1 < m$, НОД $(j_1, m) = 1$, $0 \leqslant j_2 < n$, НОД $(j_2, n) = 1$. Число возможных значений для j_1 равно $\varphi(m)$, а для j_2 равно $\varphi(n)$. Итак, число пар равно $\varphi(m)\varphi(n)$. Следствие доказано.

Поскольку каждое число n может быть представлено в виде произведения степеней различных простых чисел и уже установлена формула $\varphi(p^\alpha) = p^\alpha(1-p^{-1})$, следствие означает, что при $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$

$$\begin{aligned}\varphi(n) &= p_1^{\alpha_1}\left(1-\frac{1}{p_1}\right) p_2^{\alpha_2}\left(1-\frac{1}{p_2}\right)\cdots p_r^{\alpha_r}\left(1-\frac{1}{p_r}\right)\\ &= n\prod_{p\mid n}\left(1-\frac{1}{p}\right).\end{aligned}$$

Из формулы для $\varphi(n)$ выводится следующее утверждение, которое будет использовано при обсуждении криптосистемы RSA с открытым ключом.

Предложение I.3.4. *Пусть известно, что n есть произведение двух простых чисел. Тогда, зная эти числа p и q, можно найти $\varphi(n)$ и обратно, зная n и $\varphi(n)$, можно найти p и q. Точнее, $\varphi(n)$ можно вычислить по p и q за $O(\log n)$ двоичных операций, а числа p и q можно вычислить по n и $\varphi(n)$ за $O(\log^3 n)$ двоичных операций.*

Доказательство. Утверждение очевидно, если n четно, так как в этом случае $p = 2$, $q = n/2$ и $\varphi(n) = n/2 - 1$; поэтому предположим, что n нечетно. В силу мультипликативности функции φ для $n = pq$ получаем $\varphi(n) = (p-1)(q-1) = n + 1 - (p+q)$. Таким образом, значение $\varphi(n)$ может быть получено из чисел p и q при помощи одного сложения и одного вычитания. Обратно, предположим, что известны n и $\varphi(n)$, и надо найти p и q. Для неизвестных величин p и q известны их произведение n и сумма $p + q = n + 1 - \varphi(n)$.

Обозначим последнее выражение через $2b$ ($p+q$ — число четное). Но два числа, произведение которых равно n, а сумма $2b$, должны быть корнями квадратного уравнения $z^2 - 2bz + n = 0$. Итак, p и q равны $b \pm \sqrt{b^2 - n}$. Наибольшее время при вычислении занимает процедура извлечения квадратного корня: в соответствии с упражнением 16 из § I.1 на нее потребуется $O(\log^3 n)$ двоичных операций. Предложение доказано.

Перейдем к обобщению малой теоремы Ферма, принадлежащему Эйлеру.

Предложение I.3.5. *Если* НОД $(a, m) = 1$, *то* $a^{\varphi(m)} \equiv 1 \pmod m$.

Доказательство. Сначала докажем это утверждение в случае, когда m есть степень простого числа: $m = p^\alpha$. Проведем индукцию по α. При $\alpha = 1$ получается малая теорема Ферма (предложение I.3.2). Пусть $\alpha \geqslant 2$ и формула верна для $(\alpha-1)$-й степени p. Тогда $a^{p^{\alpha-1}-p^{\alpha-2}} = 1 + p^{\alpha-1}b$ для некоторого целого b. Возводя обе части этого равенства в степень p и используя тот факт, что все биномиальные коэффициенты в выражении $(1+x)^p$, кроме первого и последнего, делятся на p, получаем, что $a^{p^\alpha - p^{\alpha-1}}$ на 1 больше суммы, каждое слагаемое которой делится на p^α. Другими словами, $a^{\varphi(p^\alpha)} - 1$ делится на p^α. Итак, теорема доказана для степеней простых чисел.

Наконец, из мультипликативности функции φ следует, что если $p^\alpha \| m$, то $a^{\varphi(m)} \equiv 1 \pmod{p^\alpha}$ (достаточно возвести обе части сравнения $a^{\varphi(p^\alpha)} \equiv 1 \pmod{p^\alpha}$ в соответствующую степень). Поскольку это верно для любого $p^\alpha \| m$ и поскольку степени различных простых чисел взаимно просты, из свойства 5 сравнений следует, что $a^{\varphi(m)} \equiv 1 \pmod m$.

Следствие. *Если* НОД $(a, m) = 1$ *и если* n' — *наименьший неотрицательный вычет* n *по модулю* $\varphi(m)$, *то* $a^n \equiv a^{n'} \pmod m$.

Это следствие доказывается так же, как следствие из предложения I.3.2.

Замечание. Как видно из доказательства предложения I.3.5, существует и меньшая степень a, сравнимая с 1 по модулю m: наименьшее общее кратное показателей степеней, сравнимых с 1 по модулю p^α для каждого $p^\alpha \| m$. Например, $a^{12} \equiv 1 \pmod{105}$ для a, взаимно простого с 105, потому что 12 кратно $3-1$, $5-1$ и $7-1$. Заметим, что $\varphi(105) = 48$. А вот еще один пример.

Пример 3. Вычислить $2^{1000000} \pmod{77}$.

Решение. Так как 30 есть наименьшее общее кратное чисел $\varphi(7) = 6$ и $\varphi(11) = 10$, на основании приведенного выше замечания получим $2^{30} \equiv 1 \pmod{77}$. Так как $1000000 = 30 \cdot 33333 + 10$, получим

$2^{1000000} \equiv 2^{10} \equiv 23 \pmod{77}$. Другой способ решения заключается в том, чтобы вычислить $2^{1000000} \pmod 7$ (так как $1000000 = 6 \cdot 166666 + 4$, то это $2^4 \equiv 2$), то же проделать для $2^{1000000} \pmod{11}$ (так как 1000000 делится на $11 - 1$, это 1) а затем, используя китайскую теорему об остатках, в промежутке от 0 до 76 найти такое x, что $x \equiv 2 \pmod 7$ и $x \equiv 1 \pmod{11}$.

Возведение вычетов в степень методом повторного возведения в квадрат. Основным действием, которое часто встречается в арифметике вычетов, является отыскание значения $b^n \pmod m$ (т.е. нахождение наименьшего неотрицательного вычета) в случае, когда m и n очень велики. Существует способ, гораздо более эффективный, чем вычисление b^n последовательным умножением на b. В дальнейшем будем предполагать, что $b < m$ и что, вычислив произведение, мы тут же приводим его по модулю m (т.е. заменяем произведение его наименьшим неотрицательным вычетом). Тогда в процессе вычислений никогда не возникнут числа, превосходящие m^2. Теперь опишем алгоритм.

Используем символ a для обозначения промежуточного результата умножения. В конце a примет значение наименьшего неотрицательного вычета величины $b^n \pmod m$. Вначале положим $a = 1$. Пусть $n_0, n_1, \ldots, n_{k-1}$ — цифры двоичной записи числа n, т.е. $n = n_0 + 2n_1 + 4n_2 + \cdots + 2^{k-1}n_{k-1}$. Каждое n_j равно 0 или 1. Если $n_0 = 1$, заменим a на b (в противном случае оставим $a = 1$). Затем возведем b в квадрат и положим $b_1 \equiv b^2 \pmod m$ (т.е. b_1 есть наименьший неотрицательный вычет b^2 по модулю m). Если $n_1 = 1$, умножим a на b_1 (и приведем по модулю m); в противном случае сохраним прежнее значение a. Затем возведем в квадрат b_1 и положим $b_2 \equiv b_1^2 \pmod m$. Если $n_2 = 1$, умножаем a на b_2, в противном случае сохраняем прежнее значение a. Продолжаем далее по тому же правилу. Таким образом, на j-м шаге вычисляем $b_j \equiv b_1^{2^j} \pmod m$. Если $n_j = 1$, т.е. если 2^j входит в двоичное представление числа n, используем b_j как множитель для вычисления нового значения a (и не делаем этого при $n_j = 0$). Легко убедиться, что после $(k-1)$-го шага получим $a \equiv b^n \pmod m$.

Сколько двоичных операций потребует алгоритм? На каждом шаге необходимо выполнить 1 или 2 умножения чисел, которые меньше m^2. И потребуется выполнить $k - 1$ шагов. Так как на каждом шаге требуется $O(\log^2(m^2)) = O(\log^2 m)$ двоичных операций, то получаем следующую оценку.

Предложение I. 3.6. $\mathrm{Time}(b^n \bmod m) = O((\log n)(\log^2 m))$.

Замечание. Если n очень велико, можно попробовать вос-

пользоваться следствием из предложения I. 3. 5, заменив n его наименьшим неотрицательным вычетом по модулю $\varphi(m)$. Но для этого нужно знать $\varphi(m)$. Если $\varphi(m)$ известно и НОД $(b, m) = 1$, то можно заменить n его наименьшим неотрицательным вычетом по модулю $\varphi(m)$, и оценка в правой части предложения I. 3. 6 примет вид $O(\log^3 m + (\log n)\log m)$ (Слагаемое $(\log n)\log m$ соответствует вычислению остатка от деления n на $\varphi(m)$. — *Прим. ред.*)

Применяя еще раз мультипликативность функции Эйлера $\varphi(n)$, выведем формулу, которая будет использована в начале главы II.

***Предложение I. 3. 7.** $\sum_{d|n} \varphi(d) = n$.

Доказательство. Обозначим через $f(n)$ левую часть этого равенства, т. е. пусть $f(n)$ — сумма значений $\varphi(d)$ по всем делителям d числа n (включая 1 и n). Требуется доказать, что $f(n) = n$. Покажем сначала, что функция $f(n)$ мультипликативна, т. е. $f(mn) = f(m)f(n)$ при НОД $(m, n) = 1$. Заметим сначала, что любой делитель d числа mn может быть единственным образом представлен в виде произведения $d_1 d_2$, где $d_1 | m$, $d_2 | n$. Так как НОД $(d_1, d_2) = 1$, то $\varphi(d) = \varphi(d_1)\varphi(d_2)$ в силу мультипликативности функции φ. Мы получим все возможные делители d числа mn, взяв все возможные пары d_1, d_2, где $d_1 | m$, $d_2 | n$. Таким образом, $f(mn) = \sum_{d_1|m} \sum_{d_2|n} \varphi(d_1)\varphi(d_2) = (\sum_{d_1|m} \varphi(d_1))(\sum_{d_2|n} \varphi(d_2)) = f(m)f(n)$. Теперь для доказательства теоремы предположим, что $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ есть разложение n на простые множители. Так как f мультипликативна, то $f(n)$ есть произведение чисел вида $f(p^\alpha)$. Поэтому достаточно доказать предложение для p^α, т. е. доказать, что $f(p^\alpha) = p^\alpha$. Но все делители числа p^α имеют вид p^j, где $0 \leqslant j \leqslant \alpha$, и поэтому $f(p^\alpha) = \sum_{j=0}^{\alpha} \varphi(p^j) = 1 + \sum_{j=1}^{\alpha} (p^j - p^{j-1}) = p^\alpha$. Это доказывает теорему для чисел вида p^α, а значит, и для всех n.

УПРАЖНЕНИЯ

1. Опишите все решения следующих сравнений:

а) $3x \equiv 4 \pmod 7$; г) $27x \equiv 25 \pmod{256}$;

б) $3x \equiv 4 \pmod{12}$; д) $27x \equiv 72 \pmod{900}$;

в) $9x \equiv 12 \pmod{21}$; е) $103x \equiv 612 \pmod{676}$.

2. Какой цифрой может заканчиваться полный квадрат в шестнадцатиричной системе счисления (см. упражнение 7 к § I. 1)?

3. Какой цифрой может заканчиваться произведение двух последовательных нечетных положительных чисел в двенадцатиричной системе счисления?

4. Доказать, что в десятичной системе счисления целое число тогда и только тогда делится на 3, когда сумма его цифр делится на 3, и что число делится на 9 тогда и только тогда, когда сумма его цифр делится на 9.

5. Доказать, что $n^5 - n$ всегда делится на 30.

6. Чтобы выложить плиткой пол размером 8 футов × 9 футов, вы купили 72 плитки по цене, которую не можете вспомнить. Общая сумма покупки, проставленная в вашем чеке, меньше \$100, однако первая и последняя цифры неразборчивы. Запись выглядит так: \$?0.6?. Сколько стоила одна плитка?

7. а) Пусть m есть либо степень p^α простого числа $p > 2$, либо удвоенная степень простого нечетного числа. Доказать, что если $x^2 \equiv 1 \pmod m$, то либо $x \equiv 1 \pmod m$, либо $x \equiv -1 \pmod m$.

б) Доказать, что утверждение п. а) неверно, если m не представимо в виде p^α или $2p^\alpha$ и $m \neq 4$.

в) Доказать, что если m — нечетное число, которое делится на r различных простых чисел, то сравнение $x^2 \equiv 1 \pmod m$ имеет 2^r различных решений между 0 и m.

8. Доказать теорему Вильсона, которая утверждает, что $(p-1)! \equiv -1 \pmod p$ для любого простого p. Доказать, что $(m-1)!$ не сравнимо с -1 по модулю m, если m *не является* простым.

9. Найти трехзначное (в десятичной записи) целое число, которое дает остаток 4 при делении на 7, 9 и 11.

10. Найти наименьшее целое положительное число, которое при делении на 11 дает остаток 1, при делении на 12 — остаток 2, а при делении на 13 — остаток 3.

11. Найти наименьшее неотрицательное решение каждой из следующих систем сравнений:

а)
$$\begin{aligned} x &\equiv 2 \pmod 3 \\ x &\equiv 3 \pmod 5 \\ x &\equiv 4 \pmod{11} \\ x &\equiv 5 \pmod{16} \end{aligned}$$

б)
$$\begin{aligned} x &\equiv 12 \pmod{31} \\ x &\equiv 87 \pmod{127} \\ x &\equiv 91 \pmod{255} \end{aligned}$$

в)
$$\begin{aligned} 19x &\equiv 103 \pmod{900} \\ 10x &\equiv 511 \pmod{841} \end{aligned}$$

12. Предположим, что трехзначное (в десятичной записи) положительное целое число, дающее остаток 7 при делении на 9 и 10 и остаток 3 при делении на 11, является делителем некоторого шестизначного натурального числа, которое при делении на 9, 10 и 11 дает, соответственно, остатки 8, 7 и 1. Найти частное.

13. В условиях предложения I.3.3 предположим, что $0 \leqslant \alpha_j < m_j < B$ для всех j, где B — некоторое число, превосходящее все модули. Предположим, что r также велико. Найти оценку для числа двоичных операций, необходимых для решения системы. Оценка должна быть функцией B и r и должна охватывать случаи, когда r либо очень мало, либо очень велико по сравнению с числом двоичных знаков в записи B.

14. Применить метод повторного возведения в квадрат для нахождения числа $38^{75} \pmod{103}$.

15. Сберегает ли время метод повторного возведения в квадрат при обычных арифметических операциях с целыми числами (в отличие от операций по модулю)? Объясните ответ, используя оценки вида O-большое.

16. Заметим, что для a, взаимно простого с p, a^{p-2} является обратным элементом к a в кольце вычетов по модулю p. Предположим, что p очень велико. Сравнить метод повторного возведения в квадрат для вычисления a^{p-2} с алгоритмом Евклида (как эффективные способы отыскания $a^{-1} \pmod p$) в случаях, когда: а) a имеет примерно столько же цифр, что и p, и б) когда a много меньше p.

17. Найти $\varphi(m)$ для всех m от 90 до 100.

18. Перечислите все $\varphi(n)$, для которых $\varphi(n) \leqslant 12$, и докажите, что ваш список полон.

19. Предположим, что n не является полным квадратом и что $n-1 > \varphi(n) > n - n^{2/3}$. Доказать, что n есть произведение двух различных простых чисел.

20. Если $m \geqslant 8$ есть степень 2, то показатель степени в предложении I.3.5 можно заменить на $\varphi(m)/2$. Доказать.

21. Пусть $m = 7785562197230017200 = 2^4 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 19 \cdot 31 \cdot 37 \cdot 41 \cdot 61 \cdot 73 \cdot 181$.

а) Найти наименьший неотрицательный вычет числа 6647^{362} по модулю m.

б) Пусть a — целое положительное число, меньшее m и взаимно простое с m. Сначала найдите такое положительное число k, меньшее 500, что a^k сравнимо с a^{-1} по модулю m. Затем опишите алгоритм возведения a в эту степень k, использующий операции по модулю m. Сколько умножений и делений потребует этот алгоритм (приведение по модулю m рассматривать как одно деление). Какое максимальное количество двоичных разрядов может быть в записи чисел, с которыми вы будете работать?

Наконец, укажите хорошую оценку числа двоичных операций, необходимых для отыскания $a^{-1} \pmod m$ этим способом (ваш ответ должен быть конкретным числом — без использования обозначений O-большое).

22. Проведите другое доказательство предложения I.3.7 следующим способом. Для каждого делителя d числа n обозначим через S_d подмножество в $\mathbf{Z}/n\mathbf{Z}$, состоящее из всех кратных n/d. Таким образом, S_d содержит d элементов.

а) Докажите, что S_d имеет $\varphi(d)$ различных элементов x, которые *порождают* S_d в том смысле, что кратные x (рассматриваемые по модулю n) дают все элементы из S_d.

б) Докажите, что каждый элемент x порождает одно из подмножеств S_d и, следовательно, число элементов в $\mathbf{Z}/n\mathbf{Z}$ равно сумме (взятой по всем делителям d) чисел элементов, порождающих S_d. Отсюда и из п. а) следует предложение I.3.7.

23. а) Используя основную теорему арифметики, докажите, что $\prod 1/(1-p^{-1})$, где произведение берется по всем простым p, стремится к бесконечности.

б) Используя результат а), докажите, что ряд, состоящий из дробей, обратных к простым числам, расходится.

в) Найти такую последовательность n_j, стремящуюся к ∞, для которой $\lim_{j\to\infty} \varphi(n_j)/n_j = 1$, и последовательность n_j, для которой $\lim_{j\to\infty} \varphi(n_j)/n_j = 0$.

24. Пусть N — очень большое секретное число, используемое для разблокирования ракетной системы, т. е. знание N позволяет запустить ракету. Предположим, что есть командующий генерал и n генерал-лейтенантов. Чтобы застраховаться от ситуации, когда командующий генерал (который знает N) будет выведен из строя, желательно, чтобы каждый из генерал-лейтенантов обладал частичной информацией относительно N, достаточной для того, чтобы любые три генерал-лейтенанта могли бы, действуя согласованно, запустить ракеты (но никакие два из них не могли бы сделать этого).

а) Пусть $p_1, \ldots, p_n$ — различные простые числа, каждое из которых больше, $\sqrt[3]{N}$, но значительно меньше $\sqrt{N}$. Используя эти числа p_i, опишите частичную информацию о числе N, которую нужно предоставить генерал-лейтенантам.

б) Обобщите этот способ на случай, когда требуется, чтобы доступ к системе открывался при согласованном действии любых k генерал-лейтенантов ($k \geqslant 2$), а никакие $k-1$ из них не могли бы разблокировать систему. Структура такого типа при n участниках называется (n, k)-пороговой системой разделения секрета.

§ 4. Некоторые применения к разложению на множители

Предложение I.4.1. *Для любого целого b и любого целого положительного n число $b^n - 1$ делится на $b-1$, и частное имеет вид $b^{n-1} + b^{n-2} + \cdots + b^2 + b + 1$.*

Доказательство. Существует полиномиальное тождество, вытекающее из того, что 1 есть корень многочлена $x^n - 1$ и, следовательно, разность $x - 1$ делит $x^n - 1$. При делении многочленов получается $x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x^2 + x + 1)$. (Это тождество можно получить также, умножая x на $x^{n-1} + x^{n-2} + \cdots + x^2 + x + 1$ и затем вычитая $x^{n-1} + x^{n-2} + \cdots + x^2 + x + 1$; после приведения подобных членов получится $x^n - 1$.) Подстановкой b вместо x доказательство завершается.

Другое доказательство состоит в переходе к системе счисления по основанию b. Запись числа $b^n - 1$ в такой системе состоит из n цифр $b-1$ (например, $10^6 - 1 = 999999$). С другой стороны, число $b^{n-1} + b^{n-2} + \cdots + b^2 + b + 1$ состоит из n единиц. Умножая $111\ldots111$ на однозначное число $b-1$, получим $((b-1)(b-1)(b-1)\cdots(b-1)(b-1)(b-1))_b = b^n - 1$.

Следствие. *Для любого целого b и любых целых положительных n и m выполняется равенство*

$$b^{mn} - 1 = (b^m - 1)(b^{m(n-1)} + b^{m(n-2)} + \cdots + b^{2m} + b^m + 1).$$

Доказательство. Просто подставим b^m вместо b в последнем предложении.

В качестве примера применения этого следствия находим, что $2^{35} - 1$ делится на $2^5 - 1 = 31$ и на $2^7 - 1 = 127$. Действительно, положим $b = 2$ и либо $m = 5$, $n = 7$, либо $m = 7$, $n = 5$.

Предложение I.4.2. *Пусть b и m взаимно просты, а числа a и c — целые положительные. Если $b^a \equiv 1 \pmod m$, $b^c \equiv 1 \pmod m$ и $d = \text{НОД}\,(a, c)$, то $b^d \equiv 1 \pmod m$.*

Доказательство. Используя алгоритм Евклида, можно записать d в виде $d = ua + vc$ с целыми коэффициентами u и v. Легко проверить, что одно из двух чисел u и v положительно, а другое отрицательно или нуль. Положим для определенности $u > 0$, $v \leqslant 0$. Возведем теперь обе части сравнения $b^a \equiv 1 \pmod m$ в степень u, а обе части сравнения $b^c \equiv 1 \pmod m$ возведем в степень $-v$. Поделив полученные сравнения, получим $b^{au-c(-v)} \equiv 1 \pmod m$. Но $au + cv = d$; таким образом, утверждение доказано.

Предложение I.4.3. *Если p — простой делитель числа $b^n - 1$, то либо* (i) *$p \mid b^d - 1$ для некоторого собственного делителя d числа*

n, либо (ii) $p \equiv 1 \pmod{n}$. *Если* $p > 2$, *а* n *нечетно, то в случае* (ii) $p \equiv 1 \pmod{2n}$.

Доказательство. По условию $b^n \equiv 1 \pmod{n}$; кроме того, по малой теореме Ферма $b^{p-1} \equiv 1 \pmod{p}$. По предыдущему предложению это означает, что $b^d \equiv 1 \pmod{p}$, где $d = \text{НОД}\,(n, p-1)$. Если $d < n$, это означает, что $p | b^d - 1$ для некоторого собственного делителя d числа n, т.е. имеет место случай (i). С другой стороны, если $d = n$, то поскольку $d | p-1$, имеет место сравнение $p \equiv 1 \pmod{n}$. Наконец, если p и n нечетны и $n | p-1$ (т.е. имеет место случай (ii)), то, очевидно, $2n | p-1$.

Покажем теперь, как можно применять это предложение при разложении на множители некоторых больших целых чисел.

Примеры.

1. Разложить на множители $2^{11} - 1 = 2047$. Если $p | 2^{11} - 1$, то согласно нашему предложению $p \equiv 1 \pmod{22}$. Поэтому проверяем $p = 23, 67, 89, \ldots$ (на самом деле достаточно проверить лишь числа до $\sqrt{2047} = 45, \ldots$). Сразу получаем разложение $2047 = 23 \cdot 89$. Совершенно аналогичным способом легко показать, что $2^{13} - 1 = 8191$ есть простое число. Простые числа вида $2^n - 1$ называются «простыми числами Мерсенна».

2. Разложить на множители $3^{12} - 1 = 531440$. В соответствии с предложением I.4.3 сначала проверяем делители чисел $3^1 - 1, 3^2 - 1, 3^3 - 1, 3^4 - 1$ и те делители числа $3^6 - 1 = (3^3 - 1)(3^3 + 1)$, которых не было в разложении $3^3 - 1$. Так получаем $2^4 \cdot 5 \cdot 7 \cdot 13$. Так как $531440/(2^4 \cdot 5 \cdot 7 \cdot 13) = 73$ есть простое число, задача решена. Заметим, что, как и следовало ожидать, простое число, отсутствующее среди делителей чисел вида $3^d - 1$, где d — собственный делитель числа 12, а именно 73, — сравнимо с единицей по модулю 12.

3. Разложить на множители $2^{35} - 1 = 34359738367$. Сначала рассмотрим делители чисел $2^d - 1$ для $d = 1, 5, 7$. Это даст нам простые множители 31 и 127. Имеем $(2^{35} - 1)/(31 \cdot 127) = 8727391$. По нашему предложению все остальные простые делители должны быть сравнимы с 1 по модулю 70. Поэтому проверим, не являются ли $71, 211, 281, \ldots$ делителями числа 8727391. Может показаться, что необходимо проверить все такие числа до $\sqrt{8727391} \approx 2954$. Однако сразу получается $8727391 = 71 \cdot 122921$, и остается проверить лишь числа до $\sqrt{122921} \approx 350$. Получаем, что 122921 — простое число. Итак, $2^{35} - 1 = 31 \cdot 71 \cdot 127 \cdot 122921$ — разложение на простые множители.

Замечание. Каким образом в примере 3 можно производить вычисления на калькуляторе, который показывает, скажем, всего 8 десятичных знаков? Нужно просто разбивать числа на части. На-

пример, при вычислении 2^{35} мы достигаем предела возможностей нашего калькулятора уже при $2^{26} = 67108864$. Чтобы умножить это число на $2^9 = 512$, запишем $2^{35} = 512 \cdot (67108 \cdot 1000 + 864) = 34359296 \cdot 1000 + 442368 = 34359738368$. Далее, когда приходится делить $2^{35} - 1$ на $31 \cdot 127 = 3937$, сначала надо разделить 34359738 на 3937, взять целую часть дроби $\left[\frac{34359738}{3937}\right] = 8727$, а затем записать $34359738 = 3937 \cdot 8727 + 1539$. Далее,

$$\frac{34359738367}{3937} = \frac{(3937 \cdot 8727 + 1539) \cdot 1000 + 367}{3937} = 8727000 + \frac{1539367}{3937} = 8727391.$$

УПРАЖНЕНИЯ

1. Докажите двумя различными способами, что если n нечетно, то $b^n + 1 = (b+1)(b^{n-1} - b^{n-2} + \cdots + b^2 - b + 1)$. В одном доказательстве используйте полиномиальное тождество, в другом — вычисления в системе счисления по основанию b.

2. Доказать, что если $2^n - 1$ — простое, то n — простое, а если $2^n + 1$ — простое, то n есть степень двойки. В первом случае простые числа называются простыми числами Мерсенна, как уже упоминалось выше, а во втором случае — простыми числами Ферма. Первые несколько чисел Мерсенна — это 3, 7, 31, 127; первые несколько чисел Ферма — это 3, 5, 17, 257.

3. Предположим, что b взаимно просто с m, где $m > 2$, и что a и c — положительные целые числа. Доказать, что если $b^a \equiv -1 \pmod m$, $b^c \equiv \pm 1 \pmod m$ и $d = \text{НОД}\,(a, c)$, то $b^d \equiv -1 \pmod m$ и a/d нечетно.

4. Доказать, что если $p | b^n + 1$, то либо (i) $p | b^d + 1$ для некоторого собственного делителя d числа n, для которого n/d нечетно, либо (ii) $p \equiv 1 \pmod{2n}$.

5. Положим $m = 2^{24} + 1 = 16777217$. а) Найти простое число Ферма, делящее m. б) Доказать, что любой другой простой делитель сравним с 1 по модулю 48. в) Разложить m на простые множители.

6. Разложить на множители $3^{15} - 1$ и $3^{24} - 1$.

7. Разложить на множители $5^{12} - 1$.

8. Разложить на множители $10^5 - 1$, $10^6 - 1$ и $10^8 - 1$.

9. Разложить на множители $2^{33} - 1$ и $2^{21} - 1$.

10. Разложить на множители $2^{15} - 1$, $2^{30} - 1$ и $2^{60} - 1$.

11. а) Доказать, что если $d = \text{НОД}\,(m, n)$ и $a > 1$ есть целое число, то $\text{НОД}\,(a^m - 1, a^n - 1) = a^d - 1$.

б) Пусть нужно перемножить числа a и b, записывающиеся в двоичной системе счисления k битами, причем k очень велико. Пусть l — фиксированное целое число, много меньшее k, и $r = [4k/l] + 1$. Выберем множество чисел m_i, $1 \leqslant i \leqslant r$, так, что $l/2 < m_i < l$ для всех i и $\text{НОД}\,(m_i, m_j) = 1$ при $i \neq j$. Предположим, что большое число a представлено r-набором $(a_1, \ldots, a_r)$, где a_i — наименьший положительный вычет числа a по модулю $2^{m_i} - 1$. Доказать, что каждое из чисел a, b и ab однозначно определяется своим r-набором, и оценить число двоичных операций, необходимых для построения r-набора числа ab по r-наборам чисел a и b.

ЛИТЕРАТУРА к ГЛАВЕ I

1. *Brillhart J., Lehmer D. H., Selfridge J. L., Tuckerman B., Wagstaff S. S., Jr.* Factorizations of $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$, up to High Powers. Providence: Amer. Math. Soc., 1983.
2. *Dickson L. E.* History of the Theory of Numbers. 3 volumes. N.Y.: Chelsea, 1952.
3. *Guy R. K.* Unsolved Problems in Number Theory. Heidelberg etc.: Springer, 1982.
4. *Hardy G. H., Wright E. M.* An Introduction to the Theory of Numbers. 5th ed. Oxford: Oxford Univ. Press, 1979.
5. *LeVeque W. J.* Fundamentals of Number Theory. Reading: Addison–Wesley, 1977.
6. *Rademacher H.* Lectures on Elementary Number Theory. Krieger, 1977.
7. *Rosen K. H.* Elementary Number Theory and Its Applications. 3rd ed. Reading: Addison–Wesley.
8. *Schroeder M. R.* Number Theory in Science and Communications. 2nd ed. Heidelberg etc.: Springer, 1986.
9. *Shanks D.* Solved and Unsolved Problems in Number Theory. 3rd ed. N.Y.: Chelsea, 1985.
10. *Sierpiński W.* A Selection of Problems in the Theory of Numbers. Oxford: Pergamon, 1964.
11. *Spencer D. D.* Computers in Number Theory. Oxford–N.Y.: Computer Science Press, 1982.

ГЛАВА II

КОНЕЧНЫЕ ПОЛЯ И КВАДРАТИЧНЫЕ ВЫЧЕТЫ

В этой главе мы предполагаем знакомство читателя с основными определениями и свойствами полей. Кратко напомним то, что нам потребуется в дальнейшем.

1. *Поле* — это множество $\mathbf{F}$ с операциями *умножения* и *сложения*, которые удовлетворяют обычным правилам: как умножение, так и сложение ассоциативны и коммутативны; справедлив дистрибутивный закон; существует аддитивная единица 0 и мультипликативная единица 1; существуют аддитивные обратный элемент и для всех отличных от нуля элементов — мультипликативные обратные элементы. Следующие примеры полей являются основными во многих областях математики: (1) поле $\mathbf{Q}$, состоящее из всех рациональных чисел; (2) поле $\mathbf{R}$ вещественных чисел; (3) поле $\mathbf{C}$ комплексных чисел; (4) поле $\mathbf{Z}/p\mathbf{Z}$ классов вычетов целых чисел по модулю простого числа p.

2. *Векторное пространство* можно определить над любым полем $\mathbf{F}$ при помощи тех же свойств, которые используются при определении векторного пространства над вещественными числами. Любое векторное пространство имеет *базис*, и число элементов в базисе называется его *размерностью*. *Расширение*, т. е. большее поле, содержащее $\mathbf{F}$, является векторным пространством над $\mathbf{F}$. Мы называем его *конечным расширением*, если оно представляет собой конечномерное векторное пространство над $\mathbf{F}$. *Степень* конечного расширения — это его размерность как векторного пространства. Один из способов получить расширение поля $\mathbf{F}$ — это присоединить к нему новый элемент: мы обозначаем через $\mathbf{K} = \mathbf{F}(\alpha)$ поле, состоящее из всех рациональных выражений, образованных элементом α и элементами $\mathbf{F}$.

3. Аналогично, *кольцо многочленов* можно определить над любым полем $\mathbf{F}$. Оно обозначается $\mathbf{F}[X]$ и состоит из всех конечных сумм степеней X с коэффициентами из $\mathbf{F}$. Многочлены в $\mathbf{F}[X]$ складываются и перемножаются так же, как многочлены с вещественными коэффици-

ентами. *Степень* d многочлена — это наибольшая степень X, которая встречается с ненулевым коэффициентом; в *нормированном* многочлене коэффициент при X^d есть 1. Будем говорить, что g *делит* f, где $f, g \in \mathbf{F}[X]$, если существует такой многочлен $h \in \mathbf{F}[X]$, что $f = gh$. *Неприводимые* многочлены $f \in \mathbf{F}[X]$ — это те многочлены, которые не делятся ни на какие многочлены меньшей степени, кроме констант; они играют ту же роль среди многочленов, что простые числа среди целых. Кольцо многочленов обладает свойством *единственности разложения*, т. е. любой нормированный многочлен разлагается, притом единственным (с точностью до порядка сомножителей) способом, в произведение нормированных неприводимых многочленов. (Ненормированный многочлен может быть однозначно представлен в виде такого произведения, умноженного на константу.)

4. Элемент α в некотором расширении $\mathbf{K}$, содержащем $\mathbf{F}$, называется *алгебраическим* над $\mathbf{F}$, если существует многочлен с коэффициентами из $\mathbf{F}$, обращающийся в 0 при подстановке в него α. В этом случае существует *единственный* нормированный неприводимый многочлен в $\mathbf{F}[X]$, корнем которого является α (и всякий другой многочлен из $\mathbf{F}[X]$, корнем которого является α, должен делиться на этот приведенный неприводимый многочлен). Если этот нормированный неприводимый многочлен имеет степень d, то любой элемент $\mathbf{F}(\alpha)$ (т. е. любое рациональное выражение, включающее в себя степени α и элементы из $\mathbf{F}$) можно представить как линейную комбинацию степеней $1, \alpha, \alpha^2, \ldots, \alpha^{d-1}$. Таким образом, эти степени α образуют базис поля $\mathbf{F}(\alpha)$ над $\mathbf{F}$, и степень расширения, полученного присоединением α, равна степени нормированного неприводимого многочлена с корнем α. Любой другой корень α' того же неприводимого многочлена называется *сопряженным* к α над $\mathbf{F}$. Поля $\mathbf{F}(\alpha)$ и $\mathbf{F}(\alpha')$ *изоморфны*: можно отобразить любое выражение, включающее в себя α, в то же самое выражение, отличающееся лишь заменой α на α'. Слово «изоморфно» означает, что мы имеем взаимно однозначное соответствие, которое сохраняет сложение и умножение. В некоторых случаях поля $\mathbf{F}(\alpha)$ и $\mathbf{F}(\alpha')$ совпадают, и тогда мы получаем автоморфизм поля. Например, $\sqrt{2}$ имеет сопряженный элемент, а именно, $-\sqrt{2}$, над $\mathbf{Q}$, и отображение $a + b\sqrt{2} \to a - b\sqrt{2}$ есть автоморфизм поля $\mathbf{Q}(\sqrt{2})$ (которое состоит из всех вещественных чисел вида $a + b\sqrt{2}$, где a и b — рациональные числа). Если все сопряженные к α числа принадлежат полю $\mathbf{F}(\alpha)$, то поле $\mathbf{F}(\alpha)$ называется *расширением Галуа*.

5. *Производная* многочлена определяется посредством правила nX^{n-1} (а не как предел, так как предельные переходы не имеют смысла, если в $\mathbf{F}$ не определены расстояние или топология). Многочлен f

степени d может не иметь корня $r \in \mathbf{F}$, т. е. значения, которое дает 0 при подстановке в многочлен вместо X. Если корень r существует, то многочлен первой степени $X - r$ делит f; если $(X-r)^m$ — наибольшая степень $X-r$, которая делит f, то мы говорим, что r есть корень *кратности* m. Ввиду однозначности разложения на множители общее число корней f в $\mathbf{F}$, с учетом кратностей, не может превосходить d. Если многочлен $f \in \mathbf{F}[X]$ имеет кратный корень r, то r должен быть корнем *наибольшего общего делителя* f и его производной f' (см. упражнение 13 к § I. 2).

6. Для любого многочлена $f(X) \in \mathbf{F}[X]$ существует такое расширение $\mathbf{K}$ поля $\mathbf{F}$, что $f(X)$ разлагается на произведение линейных сомножителей (или, что равносильно, многочлен f степени d имеет d корней в $\mathbf{K}$ с учетом кратностей). Более того, существует наименьшее расширение $\mathbf{F}$, содержащее эти корни; оно называется *полем разложения* f. Поле разложения $\mathbf{K}$ определяется однозначно с точностью до изоморфизма, т. е. если $\mathbf{K}'$ — другое поле с этими свойствами, то существует взаимно однозначное соответствие $\mathbf{K} \xrightarrow{\sim} \mathbf{K}'$, сохраняющее сложение и умножение. Например, $\mathbf{Q}(\sqrt{2})$ есть поле разложения $f(X) = X^2 - 2$, а чтобы получить поле разложения $f(X) = X^3 - 2$, нужно к $\mathbf{Q}$ присоединить как $\sqrt[3]{2}$, так и $\sqrt{-3}$.

7. Если сложение с собой мультипликативной единицы 1 в поле $\mathbf{F}$ никогда не дает аддитивную единицу 0, то говорят, что $\mathbf{F}$ имеет *характеристику* 0. В этом случае $\mathbf{F}$ содержит в себе копию поля рациональных чисел. В противном случае существует такое простое число p, что сумма p слагаемых $1 + 1 + \cdots + 1$ равна 0; тогда p называется *характеристикой* поля $\mathbf{F}$. В этом случае $\mathbf{F}$ содержит в себе копию поля $\mathbf{Z}/p\mathbf{Z}$ (см. следствие 1 предложения I. 3. 1), которое называется *простым полем*.

§ 1. Конечные поля

Пусть $\mathbf{F}_q$ обозначает поле, состоящее из конечного числа q элементов. Ясно, что конечное поле не может иметь характеристику 0, так что характеристика $\mathbf{F}_q$ — простое число p. Тогда $\mathbf{F}_q$ содержит в себе простое поле $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ и, следовательно, является векторным пространством (разумеется, конечномерным) над $\mathbf{F}_p$. Пусть f обозначает его размерность как $\mathbf{F}_p$-векторного пространства. Выбор базиса позволяет нам установить взаимно однозначное соответствие между элементами этого f-мерного векторного пространства и множеством всех f-выборок элементов из $\mathbf{F}_p$. Поэтому в $\mathbf{F}_q$ должно быть p^f элементов. Таким образом, q — *степень характеристики* p.

Мы вскоре увидим, что для каждой степени $q = p^f$ простого числа p существует единственное (с точностью до изоморфизма) поле из q элементов. Но сначала мы изучим мультипликативный *порядок* элементов в множестве $\mathbf{F}_q^*$ ненулевых элементов нашего конечного поля. Под «порядком» ненулевого элемента мы подразумеваем наименьшую положительную его степень, равную 1.

Существование мультипликативных образующих конечных полей. В $\mathbf{F}_q$ имеется $q-1$ ненулевых элементов. Согласно определению поля они образуют *абелеву группу* по умножению. Это означает, что произведение двух ненулевых элементов — не нуль, выполняются ассоциативный и коммутативный законы, существует единичный элемент 1 и любой ненулевой элемент имеет обратный. Число элементов в группе делится на порядок любого элемента — это факт, общий для всех конечных групп. В целях полноты изложения мы даем его доказательство для нашей группы $\mathbf{F}_q^*$.

Предложение II.1.1. *Порядок любого $a \in \mathbf{F}_q^*$ делит $q-1$.*

Первое доказательство. Пусть d — наименьшая степень a, для которой $a^d = 1$. (Такая степень, действительно, существует, так как ввиду конечности $\mathbf{F}_q^*$ степени элемента a не могут быть все различны между собой, и если $a^i = a^j$, $j > i$, то $a^{j-i} = 1$.) Пусть S обозначает множество $\{1, a, a^2, \ldots, a^{d-1}\}$ всех различных степеней a. Для любого $b \in \mathbf{F}_q^*$ пусть bS обозначает смежный класс, состоящий из всех элементов вида ba^j (например, $1S = S$). Нетрудно заметить, что любые два смежных класса либо совпадают, либо не содержат общих элементов. (Действительно, пусть элемент b_1a^i из b_1S принадлежит также b_2S, т.е. $b_1a^i = b_2a^j$. Тогда для любого $b_1a^{i'}$ из b_1S имеем $b_1a^{i'} = b_1a^{i+i'-i} = b_2a^{j+i'-i}$.) Каждый смежный класс состоит из d элементов. Объединение всех смежных классов исчерпывает $\mathbf{F}_q^*$ — это означает, что $\mathbf{F}_q^*$ есть объединение попарно непересекающихся d-элементных подмножеств. Следовательно, $d|(q-1)$.

Второе доказательство. Покажем, во-первых, что $a^{q-1} = 1$. Чтобы убедиться в этом, напишем произведение всех ненулевых элементов из $\mathbf{F}_q$. Таковых имеется $q-1$. Так как при умножении на a любые два различных элемента остаются различными, то, умножив каждый элемент произведения на a, мы вновь получаем то же произведение (при другом порядке сомножителей). Однако при этом мы умножили произведение на a^{q-1}. Следовательно, $a^{q-1} = 1$ (ср. с доказательством предложения I.3.2). Пусть теперь d — порядок a, т.е. наименьшая положительная степень a, которая дает 1. Если бы $q-1$ не делилось на d, т.е. $q-1 = bd + r$, $0 < r < d$, то

меньшее, чем d, число r обладало бы свойством $a^r = a^{q-1-bd} = 1$. Но это противоречит минимальности d. Доказательство завершено.

О п р е д е л е н и е. *Образующим элементом* g конечного поля $\mathbf{F}_q$ называется элемент порядка $q-1$; это равносильно тому, что степени g пробегают все элементы $\mathbf{F}_q^*$.

Следующее предложение описывает одно из наиболее фундаментальных свойств конечных полей. Оно утверждает, что отличные от нуля элементы любого конечного поля образуют *циклическую группу*, т. е. все они суть степени одного и того же элемента.

Предложение II. 1. 2. *Каждое конечное поле имеет образующий элемент. Если g — образующий элемент $\mathbf{F}_q^*$, то g^j — также образующий тогда и только тогда, когда* НОД $(j, q-1) = 1$. *В частности, всего в $\mathbf{F}_q$ имеется $\varphi(q-1)$ различных образующих элементов.*

Д о к а з а т е л ь с т в о. Предположим, что $a \in \mathbf{F}_q^*$ имеет порядок d, т. е. $a^d = 1$ и никакая меньшая степень a не равна 1. Согласно предложению II. 1. 1 число d делит $q-1$. Так как a^d есть наименьшая степень, равная 1, все элементы $a, a^2, \ldots, a^d = 1$ различны между собой. Мы утверждаем, что элементы порядка d — это в точности те $\varphi(d)$ значений a^j, для которых НОД $(j, d) = 1$. Во-первых, так как все d различных степеней различны, они представляют собой множество всех корней уравнения $X^d = 1$ (см. п. 5 списка свойств полей в начале главы). Любой элемент порядка d должен тем самым быть среди степеней a. Однако не каждая степень a имеет порядок d, так как при НОД $(j, d) = d' > 1$ элемент a^j имеет меньший порядок: числа d/d' и j/d' — целые, и поэтому $(a^j)^{d/d'} = (a^d)^{j/d'} = 1$. Обратно, теперь покажем, что a^j при НОД $(j, d) = 1$ имеет порядок d. Действительно, допустим, что НОД $(j, d) = 1$ и a^j имеет порядок $d'' < d$. Тогда $(a^{d''})^j = (a^{d''})^d = 1$, следовательно, $(a^{d''})^{\text{НОД}(j,d)} = a^{d''} = 1$ (доказывается аналогично предложению I. 4. 2). Но $a^{d''} \neq 1$, так как a имеет порядок d. Таким образом, a^j имеет порядок d тогда и только тогда, когда НОД $(j, d) = 1$.

Это значит, что если имеется элемент порядка d, то существует в точности $\varphi(d)$ элементов порядка d. Итак, для всякого $d|(q-1)$ существуют лишь две возможности: либо элементов порядка d нет, либо таких элементов в точности $\varphi(d)$.

Но каждый элемент имеет некоторый порядок $d|(q-1)$, и имеется либо 0, либо $\varphi(d)$ элементов порядка d. Согласно предложению I. 3. 7 справедливо равенство $\sum_{d|(q-1)} \varphi(d) = q-1$, правая часть которого совпадает с числом элементов в $\mathbf{F}_q^*$. Отсюда и из того, что каждый элемент имеет некоторый порядок, следует, что для каждого $d|(q-1)$

всегда имеется в точности $\varphi(d)$ (и никогда — нуль) элементов порядка d. В частности, существует в точности $\varphi(q-1)$ элементов порядка $q-1$. Как мы показали выше, если элемент g имеет порядок $q-1$, то все остальные элементы порядка $q-1$ — это такие степени g^j, что НОД $(j, q-1)=1$. Предложение доказано.

Следствие. *Для каждого простого p существует такое целое число g, что его степени пробегают все ненулевые классы вычетов по модулю p.*

Пример 1. Мы можем получить все вычеты по модулю 19 от 1 до 18, взяв степени 2. А именно, последовательные степени 2, приведенные по модулю 19 —это 2, 4, 8, 16, 13, 7, 14, 9, 18, 17, 15, 11, 3, 6, 12, 5, 10, 1.

Во многих ситуациях при работе с такими конечными полями, как $\mathbf{F}_p$, p — простое число, бывает полезно найти образующий элемент. Что получится, если выбрать элемент $g \in \mathbf{F}_p^*$ случайно? Какова вероятность того, что он будет образующим? Другими словами, какую часть от ненулевых элементов составляют образующие? В силу предложения II. 1. 2, эта часть равна $\varphi(p-1)/(p-1)$. Согласно формуле для $\varphi(p-1)$ из следствия предложения I. 3. 3 эта дробь равна $\prod(1-l^{-1})$, где произведение берется по всем простым делителям l числа $p-1$. Таким образом, шансы найти образующий элемент путем случайного угадывания сильно зависят от разложения числа $p-1$. Например, справедливо следующее утверждение.

Предложение II. 1. 3. *Существует такая последовательность простых чисел p, что случайно выбранный в $\mathbf{F}_p^*$ элемент g оказывается образующим с вероятностью, стремящейся к нулю с ростом p.*

Доказательство. Пусть $\{n_j\}_{j=2,3,\ldots}$ — любая последовательность натуральных чисел, которые с ростом j к ∞ делятся на все большие количества последовательных простых чисел $2, 3, 5, 7, \ldots$ Можно взять, например, $n_j = j!$. Выберем в качестве p_j любое такое простое число, что $p_j \equiv 1 \pmod{n_j}$. Почему такой выбор возможен? Это следует из *теоремы Дирихле о простых числах в арифметической прогрессии*, которая утверждает, что *если целые числа k и n взаимно просты, то существует бесконечно много простых чисел, сравнимых с k по модулю n.* (Более того, простые числа «равномерно распределены» среди возможных вычетов $k \bmod n$, т. е. доля тех простых, которые сравнимы с данным k по модулю n, составляет $1/\varphi(n)$. Но этот факт нам здесь не понадобится.) Тогда множество простых, делящих $p_j - 1$, содержит в себе простые, делящие n_j, и, значит,

$$\frac{\varphi(p_j-1)}{p_j-1} \leqslant \prod_{\text{простые } l \mid n_j} \left(1-\frac{1}{l}\right).$$

При $j \to \infty$ это произведение стремится к $\prod_{\text{все простые } l}(1-l^{-1})$, которое есть 0 (см. упражнение 23 § I.3). Предложение доказано.

Существование и единственность конечных полей с числом элементов, равным степени простого числа.

Как существование, так и единственность мы установим, доказав, что конечное поле с $q = p^f$ элементами есть поле разложения многочлена $X^q - X$. Следующее предложение показывает, что для каждого числа q, равного степени простого числа, существует одно и (с точностью до изоморфизма) только одно конечное поле с q элементами.

Предложение II.1.4. *Если* $\mathbf{F}_q$ *есть поле с* $q = p^f$ *элементами, то каждый его элемент удовлетворяет уравнению* $X^q - X = 0$ *и* $\mathbf{F}_q$ *— это в точности множество корней этого уравнения. Обратно, для любой степени простого* $q = p^f$ *поле разложения над* $\mathbf{F}_p$ *многочлена* $X^q - X$ *есть поле из* q *элементов.*

Доказательство. Предположим сначала, что $\mathbf{F}_q$ — конечное поле. Так как порядок каждого ненулевого элемента делит $q-1$, такой элемент должен удовлетворять уравнению $X^{q-1} - 1 = 0$, а следовательно, и уравнению $X^q - X = 0$. Конечно, элемент 0 также удовлетворяет последнему уравнению. Таким образом, все элементы поля — корни многочлена $X^q - X$ степени q. Так как этот многочлен не может иметь более q корней, его корни — в точности элементы $\mathbf{F}_q$. Заметим: это означает, что $\mathbf{F}_q$ есть поле разложения для многочлена $X^q - X$, т.е. наименьшее расширение $\mathbf{F}_p$, которое содержит все его корни.

Обратно, пусть $q = p^f$ есть степень простого числа p и $\mathbf{F}$ — поле разложения над $\mathbf{F}_p$ многочлена $X^q - X$. Заметим, что $X^q - X$ имеет производную $qX^{q-1} - 1 = -1$ (так как q кратно p и, следовательно, равно нулю в $\mathbf{F}_p$). Поэтому $X^q - X$ не имеет общих корней с производной (у которой вообще корней нет) и, следовательно, не имеет кратных корней. Поэтому $\mathbf{F}$ должно содержать, по крайней мере, q различных корней $X^q - X$. Однако мы утверждаем, что множество из q корней уже составляет поле. Ключевым является тот факт, что сумма и произведение корней многочлена $X^q - X$ — снова его корни. Это нетрудно проверить для произведения: если a и b — корни многочлена, то $a^q = a$, $b^q = b$ и, значит, $(ab)^q - ab = a^q b^q - ab = ab - ab = 0$, т.е. ab — корень $X^q - X$. Чтобы убедиться в том, что сумма $a + b$ также есть корень $X^q - X$, мы установим важное свойство полей ха-

рактеристики p.

Лемма. *Для любых элементов a, b поля характеристики p выполняется соотношение $(a+b)^p = a^p + b^p$.*

Утверждение леммы следует из того, что все промежуточные члены в биномиальном разложении $\sum_{j=0}^{p} \binom{p}{j} a^j b^{p-j}$ обращаются в нуль, так как $p!/((p-j)!j!)$ делится на p, если $0 < j < p$.

Последовательное применение леммы дает: $a^p + b^p = (a+b)^p$, $a^{p^2} + b^{p^2} = (a^p + b^p)^p = (a+b)^{p^2}, \ldots, a^q + b^q = (a+b)^q$. Таким образом, если $a^q = a$ и $b^q = b$, то $(a+b)^q = a^q + b^q = a + b$, и $a+b$ есть корень $X^q - X$. Мы заключаем теперь, что множество из q корней есть наименьшее поле, содержащее все корни многочлена $X^q - X$, т. е. поле разложения этого многочлена есть поле из q элементов. Предложение доказано.

По ходу доказательства мы показали, что возведение элементов в степень p сохраняет сложение и умножение. В следующем предложении мы выведем другие важные следствия этого факта.

Предложение II. 1. 5. *Пусть $\mathbf{F}_q$ — конечное поле из $q = p^f$ элементов и пусть σ — отображение $\mathbf{F}_q$, при котором элементу a сопоставляется a^p: $\sigma(a) = a^p$. Тогда σ есть* ***автоморфизм*** *поля $\mathbf{F}_q$ (т. е. взаимно однозначное отображение поля в себя, сохраняющее сложение и умножение). Элементы $\mathbf{F}_q$, не изменяющиеся при действии σ, — это в точности элементы простого поля $\mathbf{F}_p$; f-я степень σ есть тождественное отображение, и меньшие степени σ этим свойством не обладают.*

Доказательство. Отображение возведения в степень всегда сохраняет умножение, а сохранение сложения вытекает из леммы в доказательстве предложения II. 1. 4. Заметим, что j-я степень σ (результат j-кратного применения σ) есть отображение σ^j: $a \mapsto a^{p^j}$. Таким образом, неподвижные элементы отображения σ^j — это корни $X^{p^j} - X$. При $j = 1$ они совпадают с p элементами простого поля (это — частный случай $q = p$ предложения II. 1. 4, а именно, малая теорема Ферма). Неподвижные элементы σ^f — это корни $X^q - X$, т. е. все элементы $\mathbf{F}_q$. Так как f-я степень σ — тождественное отображение, σ должно быть взаимно однозначным. Обратным для σ является отображение σ^{f-1}: $a \mapsto a^{p^{f-1}}$. Никакая меньшая f степень σ не является тождественным отображением, так как при $j < f$ не все элементы $\mathbf{F}_q$ являются корнями многочлена $X^{p^j} - X$. Это завершает доказательство.

Предложение II. 1. 6. *Если (в обозначениях предложения* II. 1. 5*) α — любой элемент $\mathbf{F}_q$, то все сопряженные с α над $\mathbf{F}_p$ элементы (т. е. элементы F_q, являющиеся вместе с α корнями нормированно-*

го неприводимого многочлена с коэффициентами из $\mathbf{F}_p$*) имеют вид* $\sigma^j(\alpha) = \alpha^{p^j}$.

Доказательство. Пусть поле $\mathbf{F}_p(\alpha)$ как расширение $\mathbf{F}_p$ имеет степень d, т.е. представляет собой копию $\mathbf{F}_{p^d}$. Тогда α удовлетворяет уравнению $X^{p^d} - X = 0$ и $\alpha^{p^j} - \alpha \neq 0$, каково бы ни было $j < d$. Таким образом, повторным применением σ к α получаем d различных элементов. Теперь достаточно показать, что все они являются корнями того же самого нормированного неприводимого многочлена, что и α; в этом случае они составляют d его корней. Чтобы сделать это, достаточно показать, что если α — корень многочлена $f(X) \in \mathbf{F}_p[X]$, то и $f(\alpha^p) = 0$. Пусть $f(X) = \sum a_j X^j$, $a_j \in \mathbf{F}_p$. Тогда $0 = f(\alpha) = \sum a_j \alpha^j$. Возводя обе части в степень p, согласно лемме получаем $0 = \sum (a_j \alpha^j)^p$. Но по малой теореме Ферма $a_j^p = a_j$, и мы имеем $0 = \sum a_j (\alpha^p)^j = f(\alpha^p)$, что и требовалось доказать.

Явные построения. До сих пор наши рассмотрения конечных полей имели теоретический характер. Единственным реальным примером были поля вида $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$. Теперь мы обсудим, как следует работать с конечными расширениями $\mathbf{F}_p$. Здесь нам следует повторить, как мы обращались с такими расширениями поля рациональных чисел $\mathbf{Q}$, как поле $\mathbf{Q}(\sqrt{2})$. Это поле получалось, когда мы брали корень α уравнения $X^2 - 2 = 0$. Затем мы рассматривали выражения вида $a + b\alpha$, $a, b \in \mathbf{Q}$, которые обычным образом складывали и перемножали, лишь заменяя каждый раз α^2 на 2 (в случае $\mathbf{Q}(\sqrt[3]{2})$ работаем с выражениями вида $a + b\alpha + c\alpha^2$, при умножении заменяя всякий раз α^3 на 2). Тот же самый общий подход можно применить к конечным полям.

Пример 2. Для построения $\mathbf{F}_9$ берем любой приведенный квадратичный многочлен в $\mathbf{F}_3[X]$, не имеющий корней в $\mathbf{F}_3$. Перебирая все наборы коэффициентов и проверяя, не являются ли $0, \pm 1 \in \mathbf{F}_3$ корнями получающихся многочленов, мы обнаруживаем лишь три приведенных неприводимых квадратичных многочлена: $X^2 + 1$, $X^2 \pm X - 1$. Если мы, например, берем корень многочлена $X^2 + 1$ (его лучше обозначить i, а не α, — ведь мы фактически присоединяем $\sqrt{-1}$), то элементами $\mathbf{F}_9$ будут все возможные комбинации $a + bi$, где $a, b = 0, \pm 1$. Арифметические действия в $\mathbf{F}_9$ поэтому аналогичны арифметике гауссовых целых чисел (см. упражнение 14 к § I.2) с той лишь разницей, что коэффициенты a, b принадлежат крошечному полю $\mathbf{F}_3$.

Заметим, что элемент i, который мы присоединим, не является образующим, так как его порядок 4, а не $q - 1 = 8$. Если, однако, мы присоединяем корень α многочлена $X^2 - X - 1$, то можем получить все

ненулевые элементы $\mathbf{F}_9$, взяв последовательные степени α (напомним, что α^2 всегда следует заменять на $\alpha+1$, так как α — корень уравнения $X^2=X+1$): $\alpha^1=\alpha$, $\alpha^2=\alpha+1$, $\alpha^3=-\alpha+1$, $\alpha^4=-1$, $\alpha^5=-\alpha$, $\alpha^6=-\alpha-1$, $\alpha^7=\alpha-1$, $\alpha^8=1$. Мы иногда будем называть такие многочлены, как X^2-X-1, *примитивными*, имея в виду, что каждый корень такого неприводимого многочлена является образующим элементом мультипликативной группы ненулевых элементов поля. В $\mathbf{F}_9^*$ имеется $4=\varphi(8)$ образующих (по предложению II.1.2): два корня X^2-X-1 и два корня X^2+X-1 (второй из корней X^2-X-1 сопряжен с α: $\sigma(\alpha)=\alpha^3=-\alpha+1$). Остальные 4 ненулевых элемента — это $\pm i=\pm(\alpha+1)$ (корни X^2+1) и 2 ненулевых элемента ± 1 поля $\mathbf{F}_3$ (корни нормированных многочленов $X+1$ и $X-1$ степени 1).

Вообще, в любом конечном поле $\mathbf{F}_q$, $q=p^f$, всякий ненулевой элемент α удовлетворяет единственному неприводимому нормированному многочлену над $\mathbf{F}_p$ некоторой степени d. Тогда поле $\mathbf{F}_p(\alpha)$, полученное присоединением этого элемента к простому полю, есть расширение степени d, содержащееся в $\mathbf{F}_q$, т.е. это копия $\mathbf{F}_{p^d}$. Так как большое поле $\mathbf{F}_{p^f}$ содержит $\mathbf{F}_{p^d}$ и поэтому является векторным пространством над $\mathbf{F}_{p^d}$ некоторой размерности f', то число элементов в $\mathbf{F}_{p^f}$ равно $(p^d)^{f'}=p^f$, т.е. $f=df'$. Следовательно, $d|f$. Обратно, для любого $d|f$ конечное поле $\mathbf{F}_{p^d}$ содержится в $\mathbf{F}_{p^f}$, так как любое решение уравнения $X^{p^d}=X$ есть решение $X^{p^f}=X$ (чтобы убедиться в этом, заметим, что если в левой части уравнения $X^{p^d}=X$ заменить $f'=f/d$ раз X на X^{p^d}, то получится соотношение $X^{p^f}=X$). Таким образом, мы доказали следующий результат.

Предложение II.1.7. *Подполя $\mathbf{F}_{p^f}$ — это поля $\mathbf{F}_{p^d}$ для $d|f$. Если элемент $\alpha\in\mathbf{F}_{p^f}$ присоединяется к $\mathbf{F}_p$, то получается одно из этих полей.*

Теперь нетрудно доказать формулу, которая используется при нахождении числа неприводимых многочленов данной степени.

Предложение II.1.8. *Для любого $q=p^f$ многочлен X^q-X разлагается в $\mathbf{F}_p[X]$ в произведение всех нормированных неприводимых многочленов степеней d, делящих f.*

Доказательство. Если присоединить к $\mathbf{F}_p$ корень α любого нормированного неприводимого многочлена $h(x)$ степени $d|f$, мы получаем копию $\mathbf{F}_{p^d}$, которая содержится в $\mathbf{F}_{p^f}$. Так как α — корень X^q-X, этот многочлен делится на $h(x)$. Обратно, пусть $h(x)$ — нормированный неприводимый многочлен, делящий X^q-X. Тогда корни $h(x)$ лежат в $\mathbf{F}_q$ (поскольку там лежат все корни X^q-X). Поэтому степень $h(x)$ делит f, согласно предложению II.1.7, так как присо-

единение корня дает подполе в $\mathbf{F}_q$. Таким образом, нормированные неприводимые многочлены, делящие $X^q - X$, — это в точности те, степени которых делят f. Как мы видели, $X^q - X$ не имеет кратных корней, и это означает, что $X^q - X$ есть произведение всех таких неприводимых многочленов, что и требовалось доказать.

Следствие. *Если f есть простое число, то в $\mathbf{F}_p[X]$ существует всего $(p^f - p)/f$ неприводимых нормированных многочленов степени f.*

Заметим, что $(p^f - p)/f$ при простом f — целое число, так как малая теорема Ферма гарантирует нам, что $p^f \equiv p \pmod f$. Пусть n — число неприводимых нормированных многочленов степени f. Согласно предложению, многочлен $X^{p^f} - X$ степени p^f есть произведение n многочленов степени f и p неприводимых многочленов $X - a$, $a \in \mathbf{F}$, степени 1. Приравнивая показатели степеней, получаем равенство $p^f = nf + p$, из которого следует искомая формула для n.

Предположим теперь, что f — не обязательно простое число. Через n_d обозначим число неприводимых нормированных многочленов степени d над $\mathbf{F}_p$. Тогда имеем $n_f = (p^f - \sum d\, n_d)/f$, где суммирование распространяется на все делители d числа f, $d < f$.

Мы распространим теперь временные оценки из главы I, касающиеся арифметики по модулю p, на конечные поля общего вида.

Предложение II.1.9. *Пусть $\mathbf{F}_q$, где $q = p^f$, есть конечное поле и $F(X)$ — неприводимый многочлен степени f над $\mathbf{F}_p$. Тогда два элемента из $\mathbf{F}_q$ можно перемножить или поделить за $O(\log^3 q)$ двоичных операций. Если k — целое положительное число, то элемент поля $\mathbf{F}_q$ можно возвести в степень k за $O((\log k)(\log^3 q))$ двоичных операций.*

Доказательство. Элемент $\mathbf{F}_q$ — это многочлен с коэффициентами из $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, рассматриваемый по модулю $F(X)$. Чтобы перемножить два таких элемента, мы сначала перемножаем многочлены. Это требует $O(f^2)$ умножений целых чисел по модулю p (и последующего сложения целых чисел по модулю p, что занимает значительно меньше времени). Затем мы делим полученный многочлен на $F(X)$: остаток от деления и дает искомое произведение. Деление многочленов требует $O(f)$ делений целых чисел по модулю p и $O(f^2)$ умножений целых чисел по модулю p. Так как умножение по модулю p требует $O(\log^2 p)$ двоичных операций, а деление по модулю p (при использовании, например, алгоритма Евклида) — $O(\log^3 p)$ двоичных операций (см. следствие из предложения I.2.2), общее число двоичных операций равно $O(f^2 \log^2 p + f \log^3 p) = O((f \log p)^3) = O(\log^3 q)$. Что-

бы установить такую же оценку для деления, достаточно показать, что обратный элемент можно найти за время $O(\log^3 q)$. Используя алгоритм Евклида для многочленов над полем $\mathbf{F}_p$ (см. упражнение 12 § 1.2), мы должны записать 1 как линейную комбинацию данного нам элемента из $\mathbf{F}_q$ (т.е. многочлена степени меньше f) и фиксированного многочлена $F(X)$ степени f. Это включает в себя $O(f)$ делений многочленов степени меньше f, а каждое деление многочленов требует $O(f^2 \log^2 p + f \log^3 p) = O(f^2 \log^3 p)$ двоичных операций. Таким образом, общее время составляет $O(f^3 \log^3 p) = O(\log^3 q)$. Наконец, возведение в k-ю степень можно производить методом повторного возведения в квадрат тем же способом, что и возведение в степень по модулю (см. конец § I.3). Для этого нужно $O(\log k)$ умножений (или возведений в квадрат) элементов из $\mathbf{F}_q$ и, следовательно, $O((\log k)(\log^3 q))$ двоичных операций. Доказательство завершено.

В заключение этого параграфа мы дадим пример вычислений с многочленами над конечными полями. Для иллюстрации мы выбрали пример над самым маленьким (но, быть может, наиболее важным) конечным полем — 2-элементным полем $\mathbf{F}_2 = \{0, 1\}$. Многочлен в $\mathbf{F}_2[X]$ — это просто сумма нескольких степеней X. В некотором смысле многочлены над $\mathbf{F}_p$ похожи на разложения целых чисел в p-ичной системе счисления: коэффициенты многочлена аналогичны значениям разрядов. Например, при бинарном разложении целое число записывается как сумма степеней 2 (с коэффициентами 0 или 1) — и точно так же многочлен над $\mathbf{F}_2$ — это сумма степеней X. Однако это сравнение может вводить в заблуждение. Например, сумма любого числа многочленов степени d есть многочлен степени не больше d. В то же время сумма нескольких d-битовых чисел будет числом, имеющим более d двоичных разрядов.

Пример 3. Пусть $f(X) = X^4 + X^3 + X^2 + 1$, $g = X^3 + 1 \in \mathbf{F}_2[X]$. Найти НОД (f, g), используя алгоритм Евклида для многочленов, и представить этот НОД в виде $u(X)f(X) + v(X)g(X)$.

Решение. Деление многочленов дает нам приведенную ниже последовательность равенств, которая приводит к заключению: НОД $(f, g) = X + 1$. Другая последовательность равенств позволяет нам, действуя в обратном направлении, выразить $X + 1$ в виде линейной комбинации f и g. (Заметим, кстати, что в поле характеристики 2 сложение эквивалентно вычитанию: $a - b = a + b - 2b = a + b$.) Имеем

$$\begin{aligned} f &= (X+1)g + X^2 + X, \\ g &= (X+1)(X^2+X) + (X+1), \\ X^2 + X &= X(X+1), \end{aligned}$$

и затем

$$\begin{aligned} X+1 &= g+(X+1)(X^2+X) = \\ &= g+(X+1)(f+(X+1)g) = \\ &= (X+1)f+(X^2)g. \end{aligned}$$

УПРАЖНЕНИЯ

1. Для $p = 2, 3, 5, 7, 11, 13, 17$ найти наименьшее положительное целое число, которое порождает $\mathbf{F}_p^*$, и определить, сколько среди чисел $1, 2, 3, \ldots, p-1$ образующих.

2. Пусть $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ обозначает множество всех обратимых (т. е. не делящихся на p) вычетов по модулю p^α. **Внимание:** следует различать множество вычетов $\mathbf{Z}/p^\alpha\mathbf{Z}$ (в котором $p^\alpha - p^{\alpha-1}$ обратимых элементов) и поле $\mathbf{F}_{p^\alpha}$ (в котором каждый ненулевой элемент обратим). Они совпадают лишь при $\alpha = 1$.

а) Пусть $p > 2$ и g — целое число, порождающее $\mathbf{F}_p^*$. Пусть α — любое целое число, большее 1. Показать, что либо g, либо $(p+1)g$ порождают $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$. Таким образом, $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ — ***циклическая группа.***

б) Показать, что при $\alpha > 2$ группа $(\mathbf{Z}/2^\alpha\mathbf{Z})^*$ ***нециклическая***, однако число 5 порождает ***подгруппу***, состоящую из половины ее элементов, а именно, из элементов, сравнимых с 1 по модулю 4.

3. Сколько элементов в наименьшем расширении $\mathbf{F}_5$, содержащем все корни многочленов X^2+X+1 и X^3+X+1?

4. Для каждой степени $d \leqslant 6$ найти число всех неприводимых многочленов над $\mathbf{F}_2$ степени d и построить их список.

5. Для каждого $d \leqslant 6$ найти число нормированных неприводимых многочленов над $\mathbf{F}_3$ степени d и для $d \leqslant 3$ построить их список.

6. Предположим, что f есть степень простого числа l. Найти простую формулу для числа нормированных неприводимых многочленов над $\mathbf{F}_p$.

7. При помощи обобщения алгоритма Евклида на многочлены (см. упражнение 12 к § I. 2) найти НОД (f, g) для $f, g \in \mathbf{F}_p[X]$ в каждом из следующих примеров. В каждом случае выразить НОД (f, g) как комбинацию f и g, т. е. в виде $d(X) = u(X)f(X) + v(X)g(X)$.

а) $f = X^3+X+1$, $g = X^2+X+1$, $p = 2$;

б) $f = X^6+X^5+X^4+X^3+X^2+X+1$, $g = X^4+X^2+X+1$, $p = 2$;

в) $f = X^3-X+1$, $g = X^2+1$, $p = 3$;

г) $f = X^5+X^4+X^3-X^2-X+1$, $g = X^3+X^2+X+1$, $p = 3$;

д) $f = X^5+88X^4+73X^3+83X^2+51X+67$, $g = X^3+97X^2+40X+38$, $p = 101$.

8. Вычислив НОД (f, f') (см. упражнение 13 к § I. 2), найти все кратные корни $f(X) = X^7+X^5+X^4-X^3-X^2-X+1 \in \mathbf{F}_3[X]$ в его поле разложения.

9. Предположим, что $\alpha \in \mathbf{F}_{p^2}$ удовлетворяет уравнению $X^2+aX+b = 0$, где $a, b \in \mathbf{F}_p$.

а) Доказать, что α^p также удовлетворяет этому уравнению.

б) Доказать, что если $\alpha \notin \mathbf{F}_p$, то $a = -\alpha-\alpha^p$ и $b = \alpha^{p+1}$. в) Доказать, что если $\alpha \notin \mathbf{F}_p$, а $c, d \in \mathbf{F}_p$, то $(\alpha c+d)^{p+1} = d^2-acd+bc^2 \in \mathbf{F}_p$.

г) Пусть i — квадратный корень из -1 в $\mathbf{F}_{19^2}$. Использовать пункт в), чтобы найти $(2+3i)^{101}$ (т. е. представить его в виде $a+bi$, $a, b \in \mathbf{F}_{19}$).

10. Пусть d — максимум степеней двух многочленов $f, g \in \mathbf{F}_p[X]$. В терминах d и p оценить число двоичных операций, необходимых для вычисления НОД (f, g) по алгоритму Евклида.

11. Для каждого из следующих полей $\mathbf{F}_q$, где $q = p^f$, найти неприводимый многочлен с коэффициентами из простого поля, корень которого α был бы примитивен (т. е. порождал бы $\mathbf{F}_q^*$), и выписать все степени α в виде многочленов от α степени меньше f. а) $\mathbf{F}_4$. б) $\mathbf{F}_8$. в) $\mathbf{F}_{27}$. г) $\mathbf{F}_{25}$.

12. Пусть $F(X) \in \mathbf{F}_2[X]$ есть примитивный неприводимый многочлен степени f. Это означает, что если α — его корень, то степенями α исчерпываются все элементы $\mathbf{F}_{2^f}^*$. Используя обозначение O-большое, оценить (в терминах f) число двоичных операций, необходимое для записи каждой степени α в виде многочлена от α степени меньше f.

13. а) При каких условиях на p и f *каждый* элемент $\mathbf{F}_{p^f}$, кроме 0 и 1, является образующим $\mathbf{F}_{p^f}^*$?

б) При каких условиях каждый элемент, отличный от 0 и 1, является либо образующим, либо квадратом образующего?

14. Доказать, что для любого фиксированного p существует такая последовательность $q_j = p^{f_j}$ степеней p, что случайный элемент $\mathbf{F}_{q_j}$ является образующим $\mathbf{F}_{q_j}^*$ с вероятностью, стремящейся к 0 при $j \to \infty$.

15. Какие многочлены в $\mathbf{F}_p[X]$ имеют тождественно нулевую производную?

16. Пусть σ — автоморфизм поля $\mathbf{F}_q$ из предложения II.1.5. Доказать, что множество элементов, остающихся неподвижными при действии σ^j, — это поле $\mathbf{F}_{p^d}$, где $d = \text{НОД}\ (j, f)$.

17. Доказать, что если b — образующий $\mathbf{F}_{p^n}^*$ и $d|n$, то $b^{(p^n-1)/(p^d-1)}$ — образующий $\mathbf{F}_{p^d}^*$.

§ 2. Квадратичные вычеты и закон взаимности

Корни из единицы. Во многих ситуациях полезно знать решения уравнения $X^n = 1$. Предположим, что мы работаем в конечном поле $\mathbf{F}_q$. Мы сейчас найдем ответ на вопрос: сколько корней n-й степени из единицы содержится в $\mathbf{F}_q$?

Предложение II.2.1. *Пусть g — образующий элемент поля $\mathbf{F}_q$. Элемент g^j является корнем n-й степени из единицы тогда и только тогда, когда $nj \equiv 0 \pmod{q-1}$. Число корней n-й степени из единицы есть* НОД $(n, q-1)$. *В частности, $\mathbf{F}_q$ содержит* ***примитивный*** *корень n-й степени из единицы (т. е. такой элемент ξ, что степени ξ пробегают все корни n-й степени из единицы) тогда и только тогда, когда $n|(q-1)$. Если ξ — примитивный корень n-й степени из единицы в $\mathbf{F}_q$, то ξ^j — также примитивный корень n-й степени из единицы тогда и только тогда, когда* НОД $(j, n) = 1$.

Доказательство. Любой элемент $\mathbf{F}_q^*$ можно записать как степень g^j образующего элемента g. Степень элемента g равна 1 в том и только том случае, когда ее показатель делится на $q-1$. Таким образом, g^j — корень n-й степени из единицы, если и только если $j_n \equiv$

$0 \pmod{q-1}$. Далее, пусть $d = \text{НОД}\,(n, q-1)$. Согласно следствию 2 из предложения I.3.1 сравнение $nj \equiv 0 \pmod{q-1}$ относительно j эквивалентно сравнению $\frac{n}{d}j \equiv 0 \pmod{\frac{q-1}{d}}$. Так как $\frac{n}{d}$ и $\frac{q-1}{d}$ взаимно просты, последнее сравнение эквивалентно тому, что j кратно $(q-1)/d$. Другими словами, d различных степеней элемента $g^{(q-1)/d}$ — это в точности корни n-й степени из 1. Число таких корней равно n тогда и только тогда, когда $d = n$, т.е. $n|(q-1)$. Наконец, если $n|(q-1)$, то полагаем $\xi = g^{(q-1)/n}$. Тогда $\xi^j = 1$ тогда и только тогда, когда $n|j$. Далее, k-я степень ξ^j равна 1 тогда и только тогда, когда $kj \equiv 0 \pmod{n}$. Нетрудно заметить, что ξ^j имеет порядок n (т.е. последнее сравнение не выполняется ни при каком положительном $k < n$) тогда и только тогда, когда j и n взаимно просты. Таким образом, при $n|(q-1)$ существует $\varphi(n)$ различных примитивных корней n-й степени из единицы. Доказательство завершено.

Следствие 1. *Если* НОД $(n, q-1) = 1$, *то* 1 *есть единственный корень n-й степени из единицы.*

Следствие 2. *Элемент* $-1 \in \mathbf{F}_q$ *имеет квадратный корень в* $\mathbf{F}_q$ *в том и только том случае, когда* $q \equiv 1 \pmod{4}$.

Следствие 1 — частный случай предложения. Чтобы доказать следствие 2, заметим, что квадратный корень из -1 — это то же самое, что примитивный корень четвертой степени из 1, а необходимое и достаточное условие его существования — это $4|(q-1)$.

Следствие 2 утверждает, что при $q \equiv 3 \pmod{4}$ мы всегда можем получить квадратичное расширение $\mathbf{F}_{q^2}$, присоединяя корень многочлена $X^2 + 1$, т.е. путем рассмотрения выражений типа «гауссовых целых чисел» $a + bi$. Мы проделали это при $q = 3$ в предыдущем параграфе.

Предположим, например, что p — простое число и $p \equiv 3 \pmod{4}$. Имеется красивый способ рассмотрения поля $\mathbf{F}_{p^2}$, который обобщается также на другие ситуации. Пусть R обозначает кольцо гауссовых целых чисел (см. упражнение 14 к §I.2). Иногда используется запись $R = \mathbf{Z} + \mathbf{Z}i$, подразумевающая, что R состоит из всех линейных комбинаций элементов 1 и i с целыми коэффициентами. Если m — произвольное гауссово число, $\alpha = a + bi$, $\beta = c + di$ — два гауссовых целых, то пишем $\alpha \equiv \beta \pmod{m}$, если $\alpha - \beta$ делится на m, т.е. если частное $(\alpha - \beta)/m$ есть целое гауссово число. Мы можем рассматривать множество R/mR совершенно так же, как кольцо классов вычетов в случае обычных целых чисел: при сложении и умножении получающийся класс вычетов не зависит от того, какие представители были выбраны в слагаемых (или сомножителях). Если теперь $m = p + 0i$ есть простое число и $p \equiv 3 \pmod{4}$, то, как нетрудно

заметить, R/pR — не что иное, как поле $\mathbf{F}_{p^2}$.

Квадратичные вычеты. Пусть p — нечетное простое число, т.е. $p > 2$. Нам интересно знать, какие из ненулевых элементов $\{1, 2, \ldots, p-1\}$ в $\mathbf{F}_p$ являются квадратами. Если $a \in \mathbf{F}_p^*$ — квадрат, скажем, $a = b^2$, то a имеет в точности 2 квадратных корня $\pm b$ (так как уравнение $X^2 - a = 0$ имеет в поле не более двух решений). Таким образом, все квадраты в $\mathbf{F}_p^*$ можно найти, вычисляя b^2 по модулю p для $b = 1, 2, \ldots, (p-1)/2$ (каждое из остальных чисел сравнимо с $-b$ для некоторого из этих b), т.е. в точности половина элементов $\mathbf{F}_p^*$ — квадраты. Например, квадраты в $\mathbf{F}_{11}$ — это $1^2 = 1$, $2^2 = 4$, $3^2 = 9$, $4^2 = 5$, $5^2 = 3$. Квадраты в $\mathbf{F}_p$ называются *квадратичными вычетами* по модулю p. Остальные ненулевые элементы называются *невычетами*. Для $p = 11$ невычеты — это 2, 6, 7, 8, 10. Имеется $(p-1)/2$ вычетов и $(p-1)/2$ невычетов.

Если g — образующий в $\mathbf{F}_p$, то любой элемент можно записать как g^j. Тогда квадрат любого элемента имеет вид g^j с четным j. Обратно, любой элемент g^j, где j четно, есть квадрат: $g^j = (\pm g^{j/2})^2$.

Символ Лежандра. Пусть a — целое число и $p > 2$ — простое число. *Символ Лежандра* $\left(\frac{a}{p}\right)$ определим равенством

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & \text{если } p|a, \\ 1, & \text{если } a \text{ — квадратичный вычет по модулю } p, \\ -1, & \text{если } a \text{ — невычет по модулю } p. \end{cases}$$

Таким образом, символ Лежандра — это просто способ обозначать, является данное целое число квадратичным вычетом или невычетом по модулю p.

Предложение II. 2. 2. *Имеет место соотношение*

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

Доказательство. Если a делится на p, то обе части соотношения сравнимы с нулем по модулю p. Пусть $p \nmid a$. По малой теореме Ферма квадрат $a^{(p-1)/2}$ в $\mathbf{F}_p$ есть 1. Поэтому $a^{(p-1)/2} = \pm 1$. Пусть g — образующий в $\mathbf{F}_p^*$ и $a = g^j$. Как мы видели, a — вычет тогда и только тогда, когда j четно. Далее, $a^{(p-1)/2} = g^{j(p-1)/2}$ равно 1 в том и только том случае, когда $j\frac{p-1}{2}$ делится на $p-1$, т.е. когда j четно. Таким образом, обе части в сравнении равны ± 1 в $\mathbf{F}_p$, и каждая из них равна $+1$ тогда и только тогда, когда j четно. Предложение доказано.

Предложение II. 2. 3. *Символ Лежандра обладает следующими свойствами*:

a) $\left(\frac{a}{p}\right)$ *зависит только от класса вычетов по модулю* p, *к которому принадлежит* a;

b) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$;

c) *если* b *и* p *взаимно просты, то* $\left(\frac{ab^2}{p}\right) = \left(\frac{a}{p}\right)$;

d) $\left(\frac{1}{p}\right) = 1$ *и* $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$.

Доказательство. Свойство a) непосредственно следует из определения. Свойство b) следует из предложения II. 2. 2, так как правая часть сравнима с $a^{(p-1)/2}b^{(p-1)/2} = (ab)^{(p-1)/2}$, как и левая. Свойство c) следует непосредственно из свойства b). Первое равенство свойства d) очевидно, так как $1^2 = 1$. Второе равенство d) вытекает из следствия 2 предложения II. 2. 1 (или из предложения II. 2. 2 с $a = -1$). Предложение доказано.

Часть b) предложения II. 2. 3 показывает, что можно определить, является ли a квадратичным вычетом по модулю p, т. е. вычислить $\left(\frac{a}{p}\right)$, если разложить на множители число a и найти символы Лежандра сомножителей. Первый шаг к этому — записать a как произведение степени 2 и нечетного числа. Рассмотрим теперь, как вычисляется $\left(\frac{2}{p}\right)$.

Предложение II. 2. 4. *Имеет место соотношение*

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = \begin{cases} 1, & \text{если } p \equiv \pm 1 \pmod 8, \\ -1, & \text{если } p \equiv \pm 3 \pmod 8. \end{cases}$$

Доказательство. Пусть $f(n) = (-1)^{(n^2-1)/8}$ при n нечетном, $f(n) = 0$ при n четном. Мы хотим показать, что $\left(\frac{2}{p}\right) = f(p)$. Из различных доказательств мы выберем эффективный метод, основанный на том, что нам уже известно о конечных полях. Так как $p^2 \equiv 1 \pmod 8$ для любого нечетного простого p, то поле $\mathbf{F}_{p^2}$ содержит примитивный корень 8-й степени из единицы. Обозначим его ξ. Заметим, что $\xi^4 = -1$. Положим $G = \sum_{j=0}^{7} f(j)\,\xi^j$ (G — пример того, что называется *гауссовой суммой*). Тогда $G = \xi - \xi^3 - \xi^5 + \xi^7 = 2(\xi - \xi^3)$ (так как $\xi^5 = \xi^4\xi = -\xi$, $\xi^7 = -\xi^3$) и $G^2 = 4(\xi^2 - 2\xi^4 + \xi^6) = 8$. Поэтому в поле $\mathbf{F}_{p^2}$ имеем:

$$G^p = (G^2)^{(p-1)/2}G = 8^{(p-1)/2}G = \left(\frac{8}{p}\right)G = \left(\frac{2}{p}\right)G$$

согласно предложению II. 2. 2 и предложению II. 2. 3 c). С другой стороны, используя определение G, соотношение $(a+b)^p = a^p + b^p$, справедливое в $\mathbf{F}_{p^2}$, а также очевидное равенство $f(j)^p = f(j)$, находим, что $G^p = \sum_{j=0}^{7} f(j)\xi^{pj}$. Легко проверить, что $f(j) = f(p)f(pj)$. Произведем теперь замену переменных $j' = pj$ (ясно, что по модулю 8 индекс j' так же, как и j, пробегает $0, 1, \ldots, 7$). Тогда

$$G^p = \sum_{j=0}^{7} f(p)f(pj)\xi^{pj} = f(p)\sum_{j'=0}^{7} f(j')\xi^{j'} = f(p)\,G.$$

Сравнение двух равенств для G^p дает желаемый результат (деление на G возможно, так как $G^2 = 8$, т. е. $G \neq 0$ в $\mathbf{F}_{p^2}$).

Далее мы будем иметь дело с простыми нечетными делителями числа a. Обозначим такой нечетный делитель через q. **Внимание**: в оставшейся части параграфа q всегда будет обозначать простое нечетное число, отличное от p, а не степень p, как это было в предыдущем параграфе.

Число a в силу предложения II. 2. 3 a) мы всегда можем считать меньшим p, поэтому и его простые делители q будут меньше p. Следующее предложение — фундаментальный квадратичный закон взаимности — связывает $\left(\frac{q}{p}\right)$ и $\left(\frac{p}{q}\right)$. Последний символ Лежандра считать легче, так как можно заменить p его наименьшим положительным вычетом по модулю q, следовательно, свести к символу Лежандра для меньших чисел. Квадратичный закон взаимности утверждает, что $\left(\frac{q}{p}\right)$ и $\left(\frac{p}{q}\right)$ не совпадают между собой, только если и p, и q сравнимы с 3 по модулю 4: в этом случае они противоположны по знаку. Это условие можно выразить формулой, используя тот факт, что $(p-1)(q-1)/4$ нечетно лишь тогда, когда $p, q \equiv 3 \pmod 4$, а в остальных случаях четно.

Предложение II. 2. 5 (Квадратичный закон взаимности). *Пусть p, q — два нечетных простых числа. Тогда*

$$\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{p}{q}\right) = \begin{cases} -\left(\frac{p}{q}\right), & \text{если } p \equiv q \equiv 3 \pmod 4; \\ \left(\frac{p}{q}\right) & \text{в остальных случаях.} \end{cases}$$

Доказательство. Опубликовано несколько дюжин доказательств квадратичного закона взаимности. Мы дадим весьма короткое доказательство в духе предыдущего предложения, используя конечные поля. Пусть f — такая степень p, что $p^f \equiv 1 \pmod q$. Мы можем, например, всегда полагать $f = q - 1$. Как мы установили в

начале параграфа (предложение II. 2. 1), поле $\mathbf{F}_{p^f}$ содержит примитивный корень q-й степени из единицы, который мы обозначим ξ (напомним, что здесь q обозначает отличное от p простое число, а не p^f). Определим «гауссову сумму» G формулой $G = \sum_{j=0}^{q-1} \left(\frac{j}{q}\right) \xi^j$. Ниже мы докажем лемму о том, что $G^2 = (-1)^{(q-1)/2} q$. Покажем сначала, как с ее помощью получается нужный нам результат. Это доказательство очень похоже на доказательство предложения II. 2. 4. Во-первых (применяя лемму, которая будет доказана ниже), мы находим, что

$$G^p = (G^2)^{(p-1)/2} G = \left((-1)^{(q-1)/2} q\right)^{(p-1)/2} G$$
$$= (-1)^{(p-1)(q-1)/4} q^{(p-1)/2} G = (-1)^{(p-1)(q-1)/4} \left(\frac{q}{p}\right) G$$

согласно предложению II. 2. 2 с q вместо a (напомним, что мы действуем в поле $\mathbf{F}_{p^f}$ характеристики p и поэтому сравнения по модулю p становятся равенствами). С другой стороны, из определения G, равенства $(a+b)^p = a^p + b^p$ в поле $\mathbf{F}_{p^f}$ и очевидного соотношения $\left(\frac{j}{q}\right)^p = \left(\frac{j}{q}\right)$ имеем

$$G^p = \sum_{j=0}^{q-1} \left(\frac{j}{q}\right) \xi^{pj} = \sum_{j=0}^{q-1} \left(\frac{p}{q}\right) \left(\frac{pj}{q}\right) \xi^{pj},$$

в силу частей b) и c) предложения II. 2. 3. Вынося $\left(\frac{p}{q}\right)$ за знак суммы и меняя порядок суммирования ($j \to j' = pj$), приходим к соотношению $G^p = \left(\frac{p}{q}\right) G$. Приравнивая два выражения для G^p и производя деление на G (что возможно, так как $G^2 = \pm q$ и потому не равно нулю в $\mathbf{F}_{p^f}$), получаем квадратичный закон взаимности.

Итак, остается лишь доказать следующую лемму.

Лемма. *Справедливо равенство* $G^2 = (-1)^{(q-1)/2} q$.

Доказательство. Воспользуемся определением G, в одном из сомножителей G заменим индекс суммирования j на $-k$ и заметим, что суммирование можно производить, начиная с 1, так как $\left(\frac{0}{p}\right) = 0$. Мы можем записать

$$G^2 = \sum_{j,k=1}^{q-1} \left(\frac{j}{q}\right) \xi^j \left(\frac{-k}{q}\right) \xi^{-k} = \left(\frac{-1}{q}\right) \sum_{j=1}^{q-1} \sum_{k=1}^{q-1} \left(\frac{jk}{q}\right) \xi^{j-k}$$
$$= (-1)^{(q-1)/2} \sum_{j=1}^{q-1} \sum_{k=1}^{q-1} \left(\frac{j^2 k}{q}\right) \xi^{j(1-k)};$$

здесь мы воспользовались частью d) предложения II.2.3, чтобы заменить $\left(\frac{-1}{q}\right)$ на $(-1)^{(q-1)/2}$ и произвели замену переменных $k \to jk$ во внутреннем суммировании (так как для каждого фиксированного j элементы jk пробегают вместе с k все ненулевые вычеты по модулю q, а слагаемые зависят только от соответствующих вычетов по модулю q). Мы далее используем часть c) предложения II.2.3, изменяем порядок суммирования и выносим $\left(\frac{k}{q}\right)$ за пределы внутренней суммы по j, что дает $G^2 = \sum_k \left(\frac{k}{q}\right) \sum_j \xi^{j(1-k)}$. В обеих двойных суммах суммирование распространяется от 1 до $q-1$. Но мы можем, не изменяя величины общей суммы, добавить слагаемые с $j = 0$: это равносильно прибавлению величины $\sum_{k=1}^{q-1} \left(\frac{k}{q}\right)$, которая равна нулю, так как число вычетов совпадает с числом невычетов. Поэтому двойная сумма принимает вид

$$\sum_{k=1}^{q-1} \left(\frac{k}{q}\right) \sum_{j=0}^{q-1} \xi^{j(1-k)}.$$

Но при каждом $k \neq 1$ внутренняя сумма равна 0 как сумма всех различных степеней примитивного корня ξ' из 1. В этом легко убедиться, заметив, что умножение такой суммы на ξ' сводится к перестановке слагаемых, т. е. произведение такой суммы и $\xi' - 1 \neq 0$ равно 0. Следовательно, вклад дает только слагаемое с $k = 1$, и мы в итоге получаем

$$G^2 = (-1)^{(q-1)/2} \left(\frac{1}{q}\right) \sum_{j=0}^{q-1} \xi^0 = (-1)^{(q-1)/2} q.$$

Лемма доказана и, таким образом, доказан квадратичный закон взаимности.

Пример 1. Определить, является ли 7411 квадратичным вычетом по модулю простого числа 9283.

Решение. Так как 7411 и 9283 — простые числа, сравнимые с 3 по модулю 4, имеем $\left(\frac{7411}{9283}\right) = -\left(\frac{9283}{7411}\right) = -\left(\frac{1872}{7411}\right)$ согласно предложению II.2.3, часть a). Так как $1872 = 2^4 \cdot 3^2 \cdot 13$, согласно предложению II.2.3, часть c), имеем $-\left(\frac{1872}{7411}\right) = -\left(\frac{13}{7411}\right)$. Применяем теперь квадратичный закон взаимности, учитывая, что $13 \equiv 1 \pmod 4$: $-\left(\frac{13}{7411}\right) = -\left(\frac{7411}{13}\right) = -\left(\frac{1}{13}\right) = -1$. Тем самым мы видим, что 7411 — невычет по модулю 9283.

Неудобство при подобном вычислении символа Лежандра заключается в том, что каждый раз верхнее число надо разлагать на множители, чтобы иметь возможность применить предложение II.2.5. Если наши числа астрономически велики, то это потребует чрезмерных затрат времени. К счастью, можно обойтись без какого бы то ни было

разложения (исключая выделение степеней 2, что сделать очень легко). Мы докажем сейчас обобщение квадратичного закона взаимности, применимое ко всем нечетным положительным целым числам, не обязательно простым. Но сначала введем определение, обобщающее определение символа Лежандра.

Символ Якоби. Пусть a — целое число, n — любое нечетное натуральное число. Пусть $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ есть разложение n на простые множители. *Символ Якоби* $\left(\frac{a}{n}\right)$ определяется как произведение символов Лежандра по всем простым делителям числа n:

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{\alpha_1} \cdots \left(\frac{a}{p_r}\right)^{\alpha_r}.$$

Важно иметь в виду: равенство $\left(\frac{a}{n}\right) = 1$ для составного n *не* означает, что a есть квадрат по модулю n. Например, $\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)(-1) = 1$, однако нет такого целого числа, квадрат которого был бы сравним с 2 по модулю 15.

Обобщим теперь предложения II.2.4 и II.2.5 на символ Якоби.

Предложение II.2.6. *Для любого нечетного натурального n*

$$\left(\frac{2}{n}\right) = (-1)^{(n^2-1)/8}.$$

Доказательство. Пусть $f(n)$, как и в доказательстве предложения II.2.4, обозначает функцию в правой части равенства. Легко заметить, что $f(n_1 n_2) = f(n_1) f(n_2)$ для любых двух нечетных чисел n_1 и n_2 (достаточно рассмотреть все возможности для n_1 и n_2 как вычетов по модулю 8). Но отсюда следует, что правая часть равенства — это $f(p_1)^{\alpha_1} \cdots f(p_r)^{\alpha_r} = \left(\frac{2}{p_1}\right)^{\alpha_1} \cdots \left(\frac{2}{p_r}\right)^{\alpha_r}$ (по предложению II.2.4), что, по определению, равно $\left(\frac{2}{n}\right)$.

Предложение II.2.7. *Для любых двух натуральных нечетных чисел m и n имеем*

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

Доказательство. Заметим, во-первых, что если m и n имеют общий множитель, то по определению символов Лежандра и Якоби обе части равны нулю. Поэтому можно считать, что НОД $(m, n) = 1$. Затем мы записываем m и n в виде произведений простых чисел: $m = p_1 p_2 \cdots p_r$, $n = q_1 q_2 \cdots q_s$ (в записях могут встречаться повторения, если m и n делятся на квадраты). Чтобы перейти от $\left(\frac{m}{n}\right) = \prod_{i,j} \left(\frac{p_i}{q_j}\right)$ к $\left(\frac{n}{m}\right) = \prod_{i,j} \left(\frac{q_j}{p_i}\right)$, мы должны rs раз применить

квадратичный закон взаимности для символа Лежандра. Число -1 получится столько же раз, сколько встретится случаев, когда как p_i в разложении m, так и q_j в разложении n сравнимы с 3 по модулю 4; а число таких случаев равно произведению числа простых сомножителей в разложении m, сравнимых с 3 по модулю 4, и числа таких же простых сомножителей в разложении n. Значит, $\left(\frac{m}{n}\right) = \left(\frac{n}{m}\right)$ всегда, кроме случая, когда число простых сомножителей, сравнимых с 3 по модулю 4, в каждом разложении нечетно; в последнем случае $\left(\frac{m}{n}\right) = -\left(\frac{n}{m}\right)$. Но произведение нечетных простых (в частности, m или n) сравнимо с 3 по модулю 4 тогда и только тогда, когда оно содержит нечетное число простых, сравнимых с 3 по модулю 4. Итак, в результате перехода от $\left(\frac{m}{n}\right)$ к $\left(\frac{n}{m}\right)$ коэффициент -1 может возникнуть только при $m \equiv n \equiv 3 \pmod 4$. Тем самым мы получили закон взаимности для символа Якоби.

Пример 2. Вернемся к примеру 1: покажем, как можно вычислить символ Лежандра, не разлагая 1872 на множители полностью, но лишь выделяя в нем максимальную степень 2. По квадратичному закону взаимности для символа Якоби имеем:

$$-\left(\frac{1872}{7411}\right) = -\left(\frac{16}{7411}\right)\left(\frac{117}{7411}\right) = -\left(\frac{7411}{117}\right) = -\left(\frac{40}{117}\right);$$

это равняется $-\left(\frac{2}{117}\right)\left(\frac{5}{117}\right) = \left(\frac{5}{117}\right) = \left(\frac{117}{5}\right) = \left(\frac{2}{5}\right) = -1$.

Квадратные корни в кольце вычетов по модулю p. Используя квадратичный закон взаимности, можно быстро определить, является ли данное целое число a квадратичным вычетом или невычетом по модулю p. Однако если a — вычет, то этот факт не помогает решить сравнение $x^2 \equiv a \pmod p$, — он лишь означает, что решение существует. В заключение этого параграфа мы приведем алгоритм, позволяющий для данного квадратичного вычета a найти его квадратный корень, если известен какой-либо *не*вычет n.

Предположим, что p — нечетное простое число и предположим, что нам известен квадратичный невычет n по модулю p. Пусть a — такое целое число, что $\left(\frac{a}{p}\right) = 1$. Мы хотим найти такое целое число x, что $x^2 \equiv a \pmod p$. Для этого выполняем следующие действия. Во-первых, записываем $p-1$ в виде $2^\alpha s$ (s нечетно). Вычисляем n^s по модулю p и обозначаем его через b. Далее вычисляем $a^{(s+1)/2}$ по модулю p и обозначаем его через r. Покажем, что r в определенном смысле близко к квадратному корню из a. Точнее, покажем, что отношение r^2 к a есть корень $2^{\alpha-1}$-й степени из единицы по модулю p. Для краткости мы вместо сравнений по модулю p будем писать равенства

и через a^{-1} обозначать обратный элемент к a по модулю p. Тогда

$$(a^{-1}r^2)^{2^{\alpha-1}} = a^{s2^{\alpha-1}} = a^{(p-1)/2} = \left(\frac{a}{p}\right) = 1.$$

Теперь нужно умножением r на некоторый корень 2^α-й степени из 1 получить такой x, чтобы выполнялось равенство $x^2/a = 1$. Убедимся, прежде всего, что b есть *примитивный* корень 2^α-й степени из единицы, т. е. что каждый корень 2^α-й степени из единицы можно представить как степень b. Сначала заметим, что b есть корень 2^α-й степени из единицы: $b^{2^\alpha} = n^{s2^\alpha} = n^{p-1} = 1$. Если бы корень b не был примитивным, то существовала бы некоторая меньшая его степень (делитель 2^α), возведение в которую дало бы 1. Но тогда b был бы четной степенью примитивного корня 2^α-й степени из 1, т. е. квадратом в $\mathbf{F}_p^*$, что невозможно, так как $b = n^s$, n — невычет, а s нечетно, и потому $\left(\frac{b}{p}\right) = \left(\frac{n}{p}\right)^s = -1$. Таким образом, b — примитивный корень 2^α-й степени из единицы. Остается найти подходящую степень b^j, где $0 \leqslant j < 2^\alpha$, так, чтобы $x = b^j r$ оказалось искомым квадратным корнем из a. Для этого представим j в двоичной записи $j = j_0 + 2j_1 + \cdots + 2^{\alpha-2}j_{\alpha-2}$ и покажем, что можно последовательно определять, какие значения (0 или 1) принимают $j_0, j_1, \ldots$ (Заметим, что всегда можно предполагать, что $j < 2^{\alpha-1}$, так как $b^{2^{\alpha-1}} = -1$, и, следовательно, изменение j на $2^{\alpha-1}$ дает значение j', для которого $b^{j'}r$ является другим квадратным корнем из a.) Вот индуктивная процедура для нахождения двоичных цифр в разложении j.

1. Возводим r^2/a в степень $2^{\alpha-2}$. Как мы доказали, квадрат этого элемента есть 1. Следовательно, мы получаем ± 1. Если мы получили 1, то полагаем $j_0 = 0$, если мы получили -1, то полагаем $j_0 = 1$. Заметим, что j_0 тем самым выбрано так, что $(b^{j_0}r)^2/a$ является корнем $2^{\alpha-2}$-й степени из единицы.

2. Пусть такие $j_0, j_1, \ldots, j_{k-1}$, что $(b^{j_0+2j_1+\cdots+2^{k-1}j_{k-1}}r)^2/a$ является корнем $2^{\alpha-k-1}$-й степени из 1, уже найдены, и мы хотим определить j_k. Возводим полученный корень в степень $2^{\alpha-k-2}$ и выбираем j_k в зависимости от того, получается при этом $+1$ или -1:

$$\text{если } \left(\frac{(b^{j_0+2j_1+\cdots+2^{k-1}j_{k-1}}r)^2}{a}\right)^{2^{\alpha-k-2}} = \begin{cases} 1 \\ -1 \end{cases},$$

то берем $j_k = \begin{cases} 0 \\ 1 \end{cases}$ соответственно. Легко проверить, что при таком выборе j_k «исправленное» значение оказывается ближе к квадратному корню из a, т. е. $(b^{j_0+2j_1+\cdots+2^k j_k}r)^2/a$ — корень $2^{\alpha-k-2}$-й степени из 1.

Наконец, когда мы дойдем до $k = \alpha - 2$ и найдем $j_{\alpha-2}$, мы получим, что

$$(b^{j_0+2j_1+\cdots+2^{\alpha-2}j_{\alpha-2}}r)^2/a = 1,$$

т.е. $b^j r$ — искомый квадратный корень из a, что и требовалось.

Пример 3. Используем предложенный алгоритм, чтобы найти квадратный корень из $a = 186$ по модулю $p = 401$.

Решение. Число $n = 3$ — невычет. Имеем $p - 1 = 2^4 \cdot 25$, поэтому $b = 3^{25} = 268$ и $r = a^{13} = 103$ (вместо сравнений по модулю p пишем равенства). Вычислив $a^{-1} = 235$, получим, что число $r^2/a = 98$ должно быть корнем степени 8 из единицы. Мы подсчитываем далее, что $98^4 = -1$. Следовательно, $j_0 = 1$. Затем находим, что $(br)^2/a = -1$. Так как квадрат этого числа равен 1, то $j_1 = 0$ и, наконец, $j_2 = 1$. Итак, $j = 5$ и искомый квадратный корень есть $b^5 r = 304$.

Замечания. 1. Наиболее простым этот алгоритм оказывается в случае, когда $p \equiv 3 \pmod 4$, p — простое. Тогда $\alpha = 1$, $s = (p-1)/2$ и $(s+1)/2 = (p+1)/4$. Легко проверить, что $x = r = a^{(p+1)/4}$ — искомый квадратный корень.

2. Оценим время работы этого алгоритма. Предположим, что невычет n известен с самого начала. Действия по нахождению s, b и $r = a^{(s+1)/2}$ (проводимые в кольце вычетов по модулю p) требуют не более $O(\log^3 p)$ двоичных операций (см. предложение I.3.6). Далее, при нахождении j наиболее трудоемкая часть k-го шага индукции — возведение в степень $2^{\alpha-k-2}$ — требует $\alpha - k - 2$ возведений в квадрат по модулю p чисел, меньших p. Так как $\alpha - k - 2 < \alpha$, для каждого шага справедлива оценка $O(\alpha \log^2 p)$. Так как всего шагов $\alpha - 1$, окончательная оценка имеет вид $O(\log^3 p + \alpha^2 \log^2 p) = O((\log p + \alpha^2)\log^2 p)$. В худшем случае (если $p-1$ лишь небольшим множителем отличается от степени 2) это составляет $O(\log^4 p)$, так как $\alpha < \log_2 p = O(\log p)$. Таким образом, если дан невычет n по модулю p, то можно извлечь квадратный корень по модулю p за полиномиальное время (ограниченное 4-й степенью числа бит в записи p).

3. В настоящее время не известны строгие (не использующие так называемую «гипотезу Римана») доказательства существования алгоритма, который находит за полиномиальное время невычет по модулю p. Однако, для любого $\varepsilon > 0$ существует алгоритм, который за полиномиальное время находит невычет с вероятностью, большей $1 - \varepsilon$. А именно, случайно выбранное число n, $0 < n < p$, с вероятностью $1/2$ является невычетом, и это можно проверить за полиномиальное время (см. упражнение 17 ниже). Если мы проделаем это для более

чем $\log_2(1/\varepsilon)$ различных случайно выбранных n, то с вероятностью, большей, чем $1-\varepsilon$, по крайней мере, одно из них будет невычетом.

УПРАЖНЕНИЯ

1. Составить таблицу всех квадратичных вычетов и невычетов по модулю p для $p = 3, 5, 7, 13, 17, 19$.

2. Предположим, что $p|2^{2^k}+1$, где $k > 1$.

а) Используя упражнение 4 § I.4 показать, что $p \equiv 1 \pmod{2^{k+1}}$.

б) Используя предложение II.2.4, доказать, что $p \equiv 1 \pmod{2^{k+2}}$.

в) Используя часть б), показать, что $2^{16}+1$ — простое число.

3. Сколько корней степени 84 из единицы содержится в поле из 11^3 элементов?

4. Доказать, что $(\frac{-2}{p}) = 1$, если $p \equiv 1$ или $3 \pmod 8$; $(\frac{-2}{p}) = -1$, если $p \equiv 5$ или $7 \pmod 8$.

5. Найти $(\frac{91}{167})$, используя закон квадратичной взаимности.

6. Найти гауссову сумму $G = \sum_{j=1}^{q-1}(\frac{j}{q})\xi^j$ (где ξ — корень q-й степени из 1 в поле $\mathbf{F}_{p^f}$, где $p^f \equiv 1 \pmod q$), если

а) $q = 7$, $p = 29$, $f = 1$, $\xi = 7$;

б) $q = 5$, $p = 19$, $f = 2$, $\xi = 2 - 4i$, где i — корень X^2+1;

в) $q = 7$, $p = 13$, $f = 2$, $\xi = 4 + \alpha$, где α — корень $X^2 - 2$.

7. Пусть $m = a^4 + 1$, $a \geqslant 2$. Найти такое натуральное x между 0 и $m/2$, что $x^2 \equiv 2 \pmod m$. Воспользовавшись этим, найти $\sqrt{2}$ в $\mathbf{F}_p$ в каждом из случаев: $p = 17$, 257, 65537 (простые числа Ферма); $p = 41 = (3^4+1)/2$; $p = 1297$; $p = 1201$. (Указание: см. доказательство предложения II.2.4.)

8. Пусть p и q — простые числа, $q \equiv 1 \pmod p$. Пусть ξ — примитивный корень p-й степени из 1 в $\mathbf{F}_q$. В терминах ξ найти формулу для квадратного корня из $(\frac{-1}{p})p$ в $\mathbf{F}_q$.

9. а) Пусть $m = a^p - 1$, где p — нечетное простое число и $a \geqslant 2$. Найти такое натуральное число x между 0 и $m/2$, что $x^2 \equiv (\frac{-1}{p})p \pmod m$. Использовать это для нахождения $\sqrt{5}$ в $\mathbf{F}_{31}$, $\sqrt{-7}$ в $\mathbf{F}_{127}$, $\sqrt{13}$ в $\mathbf{F}_{8191}$, $\sqrt{-7}$ в $\mathbf{F}_{1093}$.

б) Пусть $q = 2^p - 1$ есть простое число Мерсенна. Найти выражение для такого наименьшего натурального числа b, что $b^2 \equiv (\frac{-1}{p})p \pmod q$.

10. Вычислить символ Лежандра $(\frac{1801}{8191})$: а) используя квадратичный закон взаимности только для символа Лежандра (т.е. разлагая на множители все появляющиеся числа), б) не разлагая на множители нечетные числа, т.е. используя квадратичный закон взаимности для символа Якоби.

11. Вычислить следующие символы Лежандра:

а) $(\frac{11}{37})$; б) $(\frac{19}{31})$; в) $(\frac{97}{101})$; г) $(\frac{31}{167})$; д) $(\frac{5}{160465489})$; е) $(\frac{3083}{3911})$; ж) $(\frac{43691}{65537})$.

12. а) Пусть p — нечетное простое. Доказать, что -3 является вычетом в $\mathbf{F}_p$ тогда и только тогда, когда $p \equiv 1 \pmod 3$.

б) Показать, что 3 является невычетом по модулю любого простого числа Мерсенна, большего 3.

13. Найти условие на последнюю десятичную цифру простого числа p, эквивалентное тому, что 5 — квадрат в $\mathbf{F}_p$.

14. Доказать, что квадратичный вычет никогда не может быть образующим в $\mathbf{F}_p^*$.

15. Пусть p — простое число Ферма.

а) Показать, что любой квадратичный невычет есть образующий в $\mathbf{F}_p^*$.

б) Показать, что 5 — образующий в $\mathbf{F}_p^*$, если $p \neq 5$.

в) Показать, что 7 — образующий в $\mathbf{F}_p^*$, если $p \neq 3$.

16. Пусть p — простое число Мерсенна. Положим $q = p^2$. Пусть i — корень $X^2 + 1 = 0$, так что $\mathbf{F}_q = \mathbf{F}_p(i)$.

а) Предположим, что целое число $a^2 + b^2$ — образующий элемент в $\mathbf{F}_p^*$. Показать, что тогда $a + bi$ — образующий для $\mathbf{F}_q^*$.

б) Показать, что как $4 + i$, так и $3 + 2i$ являются образующими для $\mathbf{F}_{31^2}^*$.

17. Пусть p — нечетное простое число, a — целое число между 1 и $p - 1$. Оценить в терминах p число двоичных операций, требующихся для вычисления $(\frac{a}{p})$: а) при использовании квадратичного закона взаимности для символа Якоби и б) при использовании предложений II.2.2 и I.3.6.

18. а) Пусть p — нечетное простое число, a, b, c — целые числа и $p \nmid a$. Доказать, что число решений $x \in \{0, 1, 2, \ldots, p-1\}$ сравнения $ax^2 + bx + c \equiv 0 \pmod{p}$ дается формулой $1 + (\frac{D}{p})$, где $D = b^2 - 4ac$ есть дискриминант.

б) Сколько решений в $\mathbf{F}_{83}$ имеется для каждого из следующих уравнений: 1) $X^2 + 1 = 0$; 2) $X^2 + X + 1 = 0$; 3) $X^2 + 21X - 11 = 0$; 4) $X^2 + X + 21 = 0$; 5) $X^2 - 4X - 13 = 0$?

в) Сколько решений в $\mathbf{F}_{97}$ имеется для каждого из уравнений в части б)?

19. Пусть $p = 2081$ и n — наименьший положительный невычет по модулю p. Найти n и, используя описанный в § II.2 метод, вычислить квадратный корень из 302 по модулю p.

20. Пусть $m = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ — разложение на простые множители целого нечетного числа m. Предположим, что a — целое число, взаимно простое с m и являющееся квадратом по модулю m. Ваша цель — найти такое целое x, что $x^2 \equiv a \pmod{m}$. Предположим, что для каждого j мы знаем некоторый невычет по модулю p_j, т. е. такое целое n_j, что $(\frac{n_j}{p_j}) = -1$.

а) Для каждого $j = 1, \ldots, r$ фиксируем $p = p_j$ и $\alpha = \alpha_j$. Предположим, что, пользуясь методом § II.2, мы нашли такое x_0, что $x_0^2 \equiv a \pmod{p}$. Показать, как можно найти такое $x = x_0 + x_1 p + x_2 p^2 + \cdots + x_{\alpha-1} p^{\alpha-1}$, что $x^2 \equiv a \pmod{p^\alpha}$.

б) Описать, как найти затем такое x, что $x^2 \equiv a \pmod{m}$. Метод, описанный в этом упражнении, называется «подъемом» квадратного корня от $\mathbf{F}_{p_j}$ $(1 \leqslant j \leqslant r)$ до $\mathbf{Z}/m\mathbf{Z}$.

21. Выше мы видели, что если n — нечетное простое число, и НОД $(b, n) = 1$, то

$$b^{(n-1)/2} \equiv \left(\frac{b}{n}\right) \pmod{n}. \qquad (*)$$

Цель этого упражнения — показать, что если n — нечетное составное число, то соотношение $(*)$ не выполняется не менее чем для 50% таких b, что НОД $(b, n) = 1$.

а) Доказать, что если соотношение $(*)$ выполняется для b_1 и не выполняется для b_2, то оно не выполняется для $b_1 b_2$. Используя это, показать, что если $(*)$ не выполняется хотя бы для одного b, то число элементов b, для которых $(*)$ не выполняется, не меньше числа элементов, для которых оно справедливо.

б) Пусть n делится на квадрат простого числа p. Показать, как найти такое целое число b, взаимно простое с n, что $b^{(n-1)/2} \not\equiv \pm 1 \pmod{n}$.

в) Пусть n — произведение различных простых чисел и p — одно из них, а b удовлетворяет условиям: $(\frac{b}{p}) = -1$ и $b \equiv 1 \pmod{n/p}$. Доказать, что для b соотношение $(*)$ не выполняется. Показать, что такое b всегда существует.

22. Объяснить, почему следующий вероятностный алгоритм дает квадратный корень из a по модулю p. Выбираем $t \in \mathbf{F}_p$ случайно до тех пор, пока не окажется,

что $t^2 - a$ является невычетом по модулю p. Пусть $\alpha = \sqrt{t^2 - a}$ обозначает элемент в квадратичном расширении $\mathbf{F}_{p^2}$. Затем вычисляем $b = (t + \alpha)^{(p+1)/2}$. Показать, что $b \in \mathbf{F}_p$ и что $b^2 = a$.

23. Предположим, что p — простое число, $p \equiv 1 \pmod 4$ и известен квадратичный невычет n по модулю p. Описать алгоритм, дающий представление $p = c^2 + d^2$ и имеющий временную оценку $O(\log^3 p)$.

ЛИТЕРАТУРА к ГЛАВЕ II

1. *Adleman L., Manders K., Miller G.* On taking roots in finite fields. — In: Proceedings of the 20th Annual Symposium on the Foundations of Computer Science, 1979, p. 175–178.
2. *Berlekamp E. R.* Factoring polynomials over large finite fields. — Math. Comp., 1970, v. 24, p. 713–735.
3. *Blake I., Gao X., Menezes A., Mullen R., Vanstone S., Yaghoobian T.* Applications of Finite Fields. Dordrecht etc.: Kluwer Acad. Publ., 1992.
4. *Gauss C. F.* Disquisitiones Arithmeticae. London–New Haven: Yale Univ. Press, 1966.
5. *Grosswald E.* Topics from the Theory of Numbers. 2nd ed. Basel: Birkhäuser, 1984.
6. *Herstein I. N.* Topics in Algebra. 2nd ed. N.Y.–Chichester: Wiley, 1975.
7. *Ireland K., Rosen M. I.* A Classical Introduction to Modern Number Theory. 2nd ed. Heidelberg etc.: Springer, 1990.
8. *Ленг С.* Алгебра. М.: Мир, 1968.
9. *Лидл Р., Нидеррайтер Г.* Конечные поля. М.: Мир, 1988.
10. *Pless V.* Introduction to the Theory of Error-Correcting Codes. N.Y.–Chichester: Wiley, 1982.
11. *Shanks D.* Solved and Unsolved Problems in Number Theory. 3rd ed. N.Y.: Chelsea Publ. Co., 1985.
12*. *Виноградов И. М.* Основы теории чисел. М.: Наука, 1972.
13*. *Боревич З. И., Шафаревич И. Р.* Теория чисел. М.: Наука, 1985.

* Добавлено при переводе. — *Прим. ред.*

ГЛАВА III

КРИПТОГРАФИЯ

§ 1. Некоторые простые криптосистемы

Основные обозначения. Криптография изучает методы пересылки сообщений в замаскированном виде, при которых только намеченные отправителем получатели могут удалить маскировку и прочитать сообщение. Предназначенное для пересылки сообщение называется *открытым текстом*, а замаскированное сообщение — *шифрованным текстом* или кратко *шифртекстом*. Открытый текст и шифртекст записываются в некоторых *алфавитах*; обычно, но не всегда, эти алфавиты совпадают и состоят из некоторого числа N *букв*. Термин «буква» (или «символ») может относится не только к обычным буквам, но также к цифрам, к пробелам, к знакам пунктуации и ко всяким другим символам, используемым при записи сообщения. (Если мы не включим, например, пробелы, то все слова слипнутся и сообщение будет трудно читать.) Процесс преобразования открытого текста в шифртекст называется *шифрованием* (или зашифрованием), а обратная процедура называется *дешифрованием* (или расшифрованием).

Открытый и шифрованный тексты разбиваются на *элементарные сообщения* («элементы»). Элементом может быть отдельная буква, пара букв (*биграмма*), тройка букв (*триграмма*) или даже блок из 50 букв. *Шифрующее преобразование* является функцией, которая преобразует элемент открытого текста в элемент шифртекста. Другими словами, это — отображение f из множества $\mathcal{P}$ всех возможных элементов открытого текста в множество C всех возможных элементов шифртекста. Будем всегда предполагать, что это отображение взаимно однозначное, т. е. для одного элемента шифртекста существует один и только один элемент открытого текста, из которого элемент шифртекста получается при шифровании. *Дешифрующее преобразование* действует в обратном направлении, это — функция f^{-1}, вос-

станавливающая открытый текст по шифртексту. Всю эту ситуацию можно схематически изобразить диаграммой

$$\mathcal{P} \xrightarrow{f} \mathcal{C} \xrightarrow{f^{-1}} \mathcal{P}.$$

Любая такая конструкция называется *криптосистемой*.

Первый шаг в создании криптосистемы состоит в обозначении всех возможных элементов открытого и шифрованного текстов математическими объектами, на которых легко строить функции. Часто такими объектами являются числа из некоторого диапазона. Например, если элементами открытого и шифрованного текстов являются отдельные буквы латинского 26-буквенного алфавита от A до Z, то можно обозначать буквы числами $0, 1, \ldots, 25$, которые будем называть их «числовыми эквивалентами». Так, вместо A мы пишем 0, вместо S — 18, а вместо X — 23. Если же, например, элементарными сообщениями являются биграммы букв 27-буквенного алфавита, состоящего из букв A–Z и пробела, то можно сопоставить пробелу числовой эквивалент 26 (следующий за эквивалентом для Z), а в качестве обозначения для биграммы, состоящей из двух букв с эквивалентами $x, y \in \{0, 1, \ldots, 26\}$, взять целое число

$$27x + y \in \{0, 1, \ldots, 728\}.$$

Таким образом, мы рассматриваем отдельные буквы как разряды 27-ричного числа, а саму биграмму — как двузначное число по основанию 27. Например, биграмме «NO» соответствует целое число $27 \cdot 13 + 14 = 365$. Аналогично, если в качестве элементов текста используются триграммы, их можно обозначать числами $729x + 27y + z \in \{0, 1, \ldots, 19682\}$. В общем случае k-буквенные блоки N-буквенного алфавита можно обозначать целыми числами между 0 и $N^k - 1$, рассматривая каждый такой блок как k-разрядное целое число в системе счисления с основанием N.

Иногда удобнее обозначать элементарные сообщения иными математическими объектами, нежели целые числа, например, векторами или точками некоторой кривой. Но в данной главе в качестве обозначений используются только целые числа.

Примеры. Начнем со случая, когда элементами сообщений (как открытого, так и шифрованного) являются буквы N-буквенного алфавита, обозначенные числами $0, 1, 2, \ldots, N - 1$. Тогда по определению преобразование шифрования является перестановкой этих N чисел.

Для ускорения шифрования и дешифрования удобно иметь относительно простое правило реализации такой перестановки. Один из

способов — рассматривать множество целых $\{0, 1, 2, \ldots, N-1\}$ как $\mathbf{Z}/N\mathbf{Z}$ и использовать на нем операции сложения и умножения по модулю N.

Пример 1. Предположим, что используется 26-буквенный алфавит A–Z с числовыми эквивалентами 0–25. Пусть буква $P \in \{0, 1, \ldots, 25\}$ представляет собой элемент открытого текста. Определим функцию f из множества $\{0, 1, \ldots, 25\}$ на себя правилом

$$f(P) = \begin{cases} P+3, & \text{если} \quad P < 23, \\ P-23, & \text{если} \quad x \geqslant 23. \end{cases}$$

Другими словами, f просто прибавляет 3 по модулю 26: $f(P) \equiv P + 3 \pmod{26}$. Последнее определение, использующее модульную арифметику, удобнее для записи и для работы с ним. Так, в данной системе для зашифрования слова «YES» сначала мы преобразуем буквы в цифры: 24 4 18, затем прибавляем 3 по модулю 26: 1 7 21 и переходим обратно к буквам: «BHV». Для дешифровки сообщения надо вычесть 3 по модулю 26. Например, шифртекст «ZKB» переходит в открытый текст «WHY». Эта криптосистема, по-видимому, использовалась в Древнем Риме Юлием Цезарем, который, как предполагается, изобрел ее сам.

Пример 1 можно обобщить следующим образом. Предположим, что используется N-буквенный алфавит с числовыми эквивалентами $0, 1, \ldots, N-1$. Пусть b — фиксированное целое. Под преобразованием *сдвига* мы понимаем функцию f, определенную по правилу $C = f(P) \equiv P + b \pmod{N}$. Шифрсистема Юлия Цезаря отвечает случаю $N = 26$, $b = 3$. Для дешифрования элемента шифртекста $C \in \{0, 1, \ldots, N-1\}$ надо просто вычислить $P = f^{-1}(C) \equiv C - b \pmod{N}$.

Теперь предположим, что вы не обладаете информацией о правилах зашифрования и дешифрования, но, тем не менее, хотите прочитать кодированное сообщение. Такое действие называется *вскрытием* кода, а наука о вскрытии кодов называется *криптоанализом*.

Для вскрытия криптосистемы нужна информация двух видов. Первый вид — это информация о природе (*структуре*) криптосистемы. Например, может быть известно, что криптосистема использует преобразование сдвига отдельных букв 26-буквенного алфавита A–Z с числовыми эквивалентами 0–25 соответственно. Второй вид информации — это информация о конкретном выборе некоторых параметров рассматриваемой криптосистемы. В нашем примере второй тип информации — это значение параметра сдвига b. Имея эту информацию, можно по формулам $C \equiv P + b \pmod{N}$ и $P \equiv C - b \pmod{N}$ производить шифрование и дешифрование.

Будем всегда предполагать, что общая структура системы известна. На практике пользователи часто имеют оборудование для шифрования и дешифрования, предназначенное для реализации только одной криптосистемы. Со временем возможна утечка информации о типе используемой системы. Поэтому для повышения надежности производят частую смену выбираемых параметров системы. Например, два пользователя криптосистемы с преобразованием сдвига могут встречаться раз в год. При каждой встрече они согласовывают список из 52 вариантов выбора параметра b, по одному варианту на каждую неделю следующего года.

Параметр b (более сложные криптосистемы обычно используют несколько параметров) называется *ключом* или, более точно, *ключом шифрования.*

Пример 2. Так, предположим, что перехвачено сообщение «FQOCUDEM», которое, как нам известно, зашифровано с использованием преобразования сдвига на буквах 26-буквенного алфавита, как в примере выше. Нам остается найти b. Сделать это можно при помощи *частотного анализа.* Он работает следующим образом. Предположим, что нами перехвачен длинный отрезок шифртекста, скажем, несколько сотен букв. Мы знаем, что «E» — наиболее часто встречающаяся буква английского языка. Поэтому разумно предположить, что наиболее часто встречающаяся в шифртексте буква является результатом шифрования буквы «E». Пусть чаще других в шифртексте встречается буква «U». Это значит, что сдвиг преобразует «E»$=4$ в «U»$=20$, т. е. $20 \equiv 4 + b \pmod{26}$, так что $b = 16$. Чтобы дешифровать сообщение, остается вычесть 16 (по модулю 26) из числовых эквивалентов «FQOCUDEM»:

$$\text{«FQOCUDEM»} = 5\,16\,14\,2\,20\,3\,4\,12 \mapsto 15\,0\,24\,12\,4\,13\,14\,22 = \text{«PAYMENOW»}.$$

Если шифрование сдвигом применяется к буквам 26-буквенного алфавита, то нет необходимости иметь длинный отрезок шифртекста, позволяющий найти наиболее часто встречаемую букву. В конце концов, для b имеется всего 26 возможностей, и можно просто попробовать их все. Скорее всего, лишь одному из значений b будет соответствовать осмысленное сообщение, такое b и есть ключ шифрования.

Значит, этот тип криптосистем слишком прост, чтобы быть хорошим. Вскрыть его очень легко. Его можно усовершенствовать, используя более широкий класс преобразований $\mathbf{Z}/N\mathbf{Z}$, называемых *аффинными* отображениями: $C \equiv aP+b \pmod{N}$, где a и b — фиксированные целые числа (вместе они образуют ключ шифрования). Если,

например, опять имея дело с 26-буквенным алфавитом, мы хотим зашифровать сообщение «PAYMENOW», используя аффинное отображение с ключом шифрования $a = 7$ и $b = 12$, то получим

$$15\,0\,24\,12\,4\,13\,14\,22 \mapsto 13\,12\,24\,18\,14\,25\,6\,10 = \text{«NMYSOZGK»}.$$

Чтобы расшифровать сообщение, зашифрованное применением аффинного отображения $C \equiv aP + b \pmod N$, нужно просто выразить P через C: $P \equiv a'C + b' \pmod N$, где $a' = a^{-1}$ есть обратное к a по модулю N число, а b' равно $-a^{-1}b$. Заметим, что это возможно лишь при НОД $(a, N) = 1$. В противном случае нельзя выразить P через C: если НОД $(a, N) > 1$, то, как нетрудно убедиться, одной букве шифртекста отвечает несколько букв открытого текста, и поэтому нельзя однозначно восстановить открытый текст по шифрованному. По определению такое преобразование не является шифрующим, так как последнее обязано быть взаимно однозначным, т. е. открытый текст должен определяться шифртекстом однозначно. Итак, аффинная криптосистема над N-буквенным алфавитом с параметрами $a \in (\mathbf{Z}/N\mathbf{Z})^*$ и $b \in \mathbf{Z}/N\mathbf{Z}$ задается правилами

$$C \equiv aP + b \pmod N, \quad P \equiv a'C + b' \pmod N,$$

где $a' = a^{-1}$ в $(\mathbf{Z}/N\mathbf{Z})^*$, $b' = -a^{-1}b$.

Как частный случай аффинной криптосистемы при $a = 1$ получаем преобразования сдвига. В другом частном случае при $b = 0$ получаем $P \equiv aC \pmod N$, $C \equiv a^{-1}P \pmod N$. Такое преобразование называется *линейным*. Оно отображает сумму в сумму, т. е. если P_1 при шифровании переходит в C_1, а P_2 — в C_2, то $P_1 + P_2$ переходит в $C_1 + C_2$ (сложение, конечно, производится по модулю N).

Пусть теперь нам известно, что перехваченное сообщение зашифровано применением аффинного отображения букв N-буквенного алфавита. Мы хотим определить ключ a, b, чтобы прочитать это сообщение. Для этого нужно знать, как зашифровываются какие-нибудь две буквы.

Пример 3. По-прежнему используем 26-буквенный алфавит. Допустим, что в шифртексте чаще всего встречается буква «K», а вторая по встречаемости буква — «D». Разумно предположить, что ими зашифрованы две наиболее часто встречающиеся буквы английского языка «E» и «T». Заменяя буквы их числовыми эквивалентами и подставляя последние в формулу дешифрования, получаем:

$$10a' + b' \equiv 4 \pmod{26},$$
$$3a' + b' \equiv 19 \pmod{26}.$$

Мы имеем два сравнения с двумя неизвестными a' и b'. Кратчайший способ решения — вычесть одно сравнение из другого, чтобы исключить b'. Получаем $7a' \equiv 11 \pmod{26}$ и $a' \equiv 7^{-1}11 \equiv 9 \pmod{26}$. Наконец, находим b', подставив это значение a' в одно из сравнений: $b' \equiv 4 - 10a' \equiv 18 \pmod{26}$. Итак, сообщение может быть дешифровано применением формулы $P \equiv 9C + 18 \pmod{26}$.

Из линейной алгебры известно, что n уравнений позволяют определить n неизвестных только тогда, когда эти уравнения независимы (т.е. когда определитель системы отличен от нуля). Например, в случае двух уравнений с двумя неизвестными это означает, что прямолинейные графики уравнений пересекаются в одной точке (не параллельны). В нашем случае при попытке провести криптоанализ аффинной системы на базе информации о двух наиболее часто встречаемых буквах шифртекста может оказаться, что два сравнения не могут быть однозначно разрешены относительно a' и b'.

Пример 4. Пусть имеется отрезок шифртекста, полученного применением аффинного преобразования 28-буквенного алфавита, содержащего буквы A–Z, пробел и знак ?, причем буквы имеют численные эквиваленты 0–25, пробел=26, ?=27. Пусть частотный анализ показал, что чаще всего в шифртексте встречаются «B» и «?» (в указанном порядке). Поскольку самыми частыми буквами в английском письменном тексте в таком 28-буквенном алфавите являются пробел и «E» (в указанном порядке), мы предполагаем, что пробел при шифровании переходит в «B», а «E» — в «?». Это приводит к двум сравнениям $a' + b' \equiv 26 \pmod{28}$, $27a' + b' \equiv 4 \pmod{28}$. Вычитая одно из другого, получаем $2a' \equiv 22 \pmod{28}$, что эквивалентно сравнению $a' \equiv 11 \pmod{14}$. Это означает, что $a' \equiv 11$ или $25 \pmod{28}$, а тогда $b' \equiv 15$ или $1 \pmod{28}$ соответственно. Оба допустимых аффинных преобразования $11C + 15$ и $25C + 1$ дают пробел и «E» как буквы открытого текста, отвечающие «B»и «?» соответственно. Теперь можно испытать обе возможности и определить, какой вариант дает осмысленное сообщение. А можно продолжить частотный анализ. Пусть «I» — третья по частоте встречаемости буква шифртекста. Используя тот факт, что «T» — третья по частоте встречаемости буква английского языка (из наших 28 букв), получаем третье сравнение $8a' + b' \equiv 19 \pmod{28}$. Эта дополнительная информация достаточна для того, чтобы выбрать из двух аффинных отображений правильное: $11C + 15$.

Преобразования биграмм. Предположим теперь, что элементами открытого и шифрованного текстов являются двухбуквенные блоки, называемые *биграммами*. Это значит, что открытый текст

разбивается на двухбуквенные сегменты. Если открытый текст состоит из нечетного числа букв, то, чтобы получить целое число биграмм, добавим к концу текста еще одну букву, выбрав ее так, чтобы не исказить смысл, например, добавим пробел, если он содержится в нашем алфавите, и «X» или «Q», если имеем дело с 26-буквенным алфавитом.

Каждой биграмме приписывается далее ее числовой эквивалент. Простейший способ — взять его в виде $xN + y$ где x — числовой эквивалент первой буквы биграммы, y — числовой эквивалент второй буквы биграммы, а N — число букв в алфавите, т.е. рассматривать биграмму как запись двузначного целого числа в системе счисления с основанием N. Это дает взаимно однозначное соответствие между множеством всех биграмм в N-буквенном алфавите и множеством всех неотрицательных целых, меньших N^2. Выше мы рассмотрели частный случай этой процедуры при $N = 27$.

Следующий шаг — выбор шифрующего преобразования, т.е. перестановки целых чисел $\{0, 1, 2, \ldots, N^2 - 1\}$. Примером простейших шифрующих преобразований служат *аффинные* преобразования: указанное множество целых чисел отождествляется с $\mathbf{Z}/N^2\mathbf{Z}$ и при шифровании P переходит в неотрицательное целое число, меньшее N^2 и удовлетворяющее сравнению $C \equiv aP + b \pmod{N^2}$. Здесь, как и раньше, число a должно не иметь общих множителей с N (что означает отсутствие общих множителей и с N^2) для того, чтобы существовало обратное преобразование, указывающее способ дешифрования: $P \equiv a'C + b' \pmod{N^2}$, где $a' \equiv a^{-1} \pmod{N^2}$, $b' \equiv -a^{-1}b \pmod{N^2}$. Мы преобразуем C в двухбуквенный блок шифртекста, записывая его в виде $C = x'N + y'$ и затем выписывая буквы с числовыми эквивалентами x' и y'.

Пример 5. Пусть мы имеем дело с 26-буквенным алфавитом и используем биграммное шифрующее преобразование $C \equiv 159P + 580 \pmod{676}$. Тогда биграмма «NO» имеет числовой эквивалент $13 \cdot 26 + 14 = 352$, и ей соответствует биграмма шифртекста $159 \cdot 352 + 580 \equiv 440 \pmod{676}$, что есть «QY». Биграмма «ON» имеет числовой эквивалент 377 и переходит в 359=«NV». Заметим, что биграмма преобразуется как единое целое, и между зашифрованными биграммами, имеющими общую букву или даже составленными из одних и тех же букв, но в разном порядке, нет явной связи.

Для того чтобы вскрыть биграммную систему, использующую аффинное преобразование $C \equiv aP+b \pmod{N^2}$, надо знать шифртекст, соответствующий двум разным элементам открытого текста. Так как элементами текста являются биграммы, то частотный анализ означа-

ет выделение в длинном отрезке шифртекста двухбуквенных блоков, встречающихся чаще других (разумеется, надо считать только элементы текста и игнорировать пары, образованные буквами соседних элементов), и сравнение с известными частотами биграмм в английских текстах (записанных в том же алфавите). Например, если используется 26-буквенный алфавит, то статистический анализ, скорее всего, покажет, что наиболее частыми биграммами являются «TH» и «HE» (в указанном порядке). Информации о двух парах отвечающих друг другу биграмм открытого и шифрованного текстов зачастую (но не всегда) бывает достаточно для определения a и b.

Пример 6. Вы знаете, что ваш противник использует криптосистему с 27-буквенным алфавитом, в которой буквы A–Z имеют числовые эквиваленты 0–25 и пробел=26. Каждой биграмме отвечает числовой эквивалент — целое число между 0 и $728 = 27^2 - 1$, определяемое по правилу $27x + y$, где x и y — числовые эквиваленты букв биграммы. Пусть анализ длинного шифртекста показал, что чаще всего в нем встречаются биграммы «ZA», «IA» и «IW» (в указанном порядке). Предположим, что самыми частыми биграммами в английском языке (в текстах в нашем 27-буквенном алфавите) являются биграммы «E » (т. е. E и пробел), «S » и « T». Мы знаем, что криптосистема использует аффинное шифрующее преобразование по модулю 729. Найти ключ дешифрования и прочитать сообщение «NDXBHO». Найти также ключ шифрования.

Решение. Мы знаем, что открытый текст шифруется по правилу $C \equiv aP + b \pmod{729}$ и что шифртекст может быть расшифрован по правилу $P \equiv a'C + b' \pmod{729}$, где a, b образуют ключ шифрования, а a', b' — ключ дешифрования. Сначала мы хотим найти a' и b'. Нам известно, как расшифровываются три биграммы. Заменив эти биграммы их числовыми эквивалентами, получим три сравнения

$$675a' + b' \equiv 134 \pmod{729},$$
$$216a' + b' \equiv 512 \pmod{729},$$
$$238a' + b' \equiv 721 \pmod{729}.$$

Если исключить b', взяв разность первых двух сравнений, то получим $459a' \equiv 351 \pmod{729}$, что не дает единственного решения $a' \pmod{729}$ (имеется 27 решений). Будет лучше, если вычесть из первого сравнения третье, что даст $437a' \equiv 142 \pmod{729}$. Для решения последнего сравнения надо найти число, обратное к 437 по модулю 729. Воспользуемся алгоритмом Евклида:

$$
\begin{aligned}
729 &= 437 + 292\\
437 &= 292 + 145\\
292 &= 2 \cdot 145 + 2\\
145 &= 72 \cdot 2 + 1
\end{aligned}
$$

и затем

$$
\begin{aligned}
1 &= 145 - 72 \cdot 2\\
&= 145 - 72\,(292 - 2 \cdot 145)\\
&= 145 \cdot 145 - 72 \cdot 292\\
&= 145\,(437 - 292) - 72 \cdot 292\\
&= 145 \cdot 437 - 217 \cdot 292\\
&= 145 \cdot 437 - 217\,(729 - 437)\\
&\equiv 362 \cdot 437 \pmod{729}.
\end{aligned}
$$

Итак, $a' \equiv 362 \cdot 142 \equiv 374 \pmod{729}$, и, следовательно, $b' \equiv 134 - 675 \cdot 374 \equiv 647 \pmod{729}$. Теперь применим преобразование дешифрования к биграммам «ND», «XB» и «HO» нашего сообщения (им соответствуют целые числа 354, 622 и 203 соответственно) и получим числа 365, 724 и 24. Записав $365 = 13 \cdot 27 + 14$, $724 = 26 \cdot 27 + 22$, $24 = 0 \cdot 27 + 24$, мы соединим биграммы открытого текста в сообщение «NO WAY». Наконец, для нахождения ключа шифрования мы вычисляем $a \equiv a'^{-1} \equiv 374^{-1} \equiv 614 \pmod{729}$ (снова используя алгоритм Евклида) и $b \equiv -a'^{-1}b' \equiv -614 \cdot 647 \equiv 47 \pmod{729}$.

Замечание. Хотя аффинные криптосистемы с биграммами (т. е. по модулю N^2) лучше аналогичных однобуквенных систем (т. е. по модулю N), они также имеют недостатки. Заметим, что вторая буква каждой биграммы шифртекста зависит только от второй буквы биграммы открытого текста. Так получается из-за того, что эта вторая буква определяется значением $C = aP + b \pmod{N^2}$ по модулю N, а оно зависит только от P по модулю N, т. е. только от второй буквы биграммы открытого текста. Поэтому можно получить важную информацию (а именно, значения параметров a и b по модулю N) из частотного анализа четных букв шифртекста. Подобное замечание справедливо и для аффинных преобразований k-буквенных блоков по модулю N^k.

УПРАЖНЕНИЯ

1. В некоторых компьютерных сетях принято в том случае, если надо послать сообщение, которое может шокировать кого-либо (например, скабрезную шутку), шифровать буквы (но не пробелы или знаки препинания) применением преобразования $C \equiv P + b \pmod{26}$. Тогда любой при желании сможет легко дешифровать текст, но никто не будет вынужден смотреть на сообщение, которое бьет по нервам. Используя частотный анализ для нахождения b, дешифруйте заключительную фразу (она приводится на английском языке) в следующем анекдоте.

На международном конгрессе хирургов представители различных стран хвастались последними достижениями в области пересадки органов. Особенно гордились французы, американцы и русские. Французский хирург сказал: «Мы пришили оторванную ногу бегуну, а год спустя он принял участие в национальных соревнованиях по бегу на 1000 метров». «Используя новейшие достижения хирургии, — заявил русский хирург, — мы смогли пришить на место потерянную руку атлету, а через год он этой самой рукой установил мировой рекорд в толкании ядра». Но все они смущенно умолкли, когда американец, не желая отставать от других, объявил, что «Jr frjrq n fzvyr ba n ubefr'f nff, naq n lrne yngre vg jnf fyrpgrq Cerfvqrag!».

Примечание. В записи использован 26-буквенный алфавит, а пробелы и знаки препинания использованы для облегчения чтения.

2. Используя частотный метод, провести криптоанализ и дешифровать сообщение, зашифрованное применением преобразования сдвига букв 26-буквенного алфавита:

$$PXPXKXENVDRUXVTNLXHYMXGMAXYKXJN$$

$$XGVRFXMAHWGXXWLEHGZXKVBIAXKMXQM$$

3. В 27-буквенном алфавите (пробел=26), используя аффинное шифрующее преобразование с ключом $a = 13$, $b = 9$, зашифровать сообщение «HELP ME».

4. В длинном отрезке шифртекста, полученного применением аффинного отображения букв 26-буквенного алфавита, чаще всех встречаются буквы «Y» и «V» (в указанном порядке). Предполагая, что эти элементы шифртекста получены шифрованием букв «E» и «T» соответственно, прочитать сообщение «QAOOYQQEVHEQV».

5. Вы анализируете аффинное шифрующее преобразование букв 37-буквенного алфавита, включающего в себя цифры 0–9, помеченные сами собой (т. е. целыми числами 0–9), буквы A–Z, имеющие числовые эквиваленты 10–35 соответственно, и пробел=36. Перехвачено сообщение «OH7F86BB46R3627O266BB9» (здесь O обозначает букву, а не цифру). Вам известно, что открытый текст заканчивается подписью «007» (ноль-ноль-семь). Что это за сообщение?

6. Вы перехватили шифртекст «QFJDFOHFXOL», полученный применением аффинного преобразования букв 27-буквенного алфавита (пробел = 26). Вам известно, что первым словом является слово «I» (за I следует пробел!). Определить ключ шифрования и прочитать сообщение.

7. а) Сколько существует различных преобразований сдвига для N-буквенного алфавита?

б) Найти формулу для числа различных аффинных шифрующих преобразований при N-буквенном алфавите.

в) Сколько существует аффинных преобразований при N=26, 27, 29, 30?

8. Говорят, что элемент P открытого текста *неподвижен* при данном шифрующем преобразовании f, если $f(P) = P$. Предположим, что используется аффинное шифрующее преобразование букв N-буквенного алфавита. Будем считать также, что это отображение *не* является сдвигом, т. е. $a \neq 1$.

а) Доказать, что если N — простое число, то всегда существует ровно одна неподвижная буква.

б) Доказать (для любого N), что если наше аффинное преобразование линейно, т. е. $b = 0$, то имеется, по крайней мере, одна неподвижная буква, и что если N четно, то линейное шифрующее преобразование имеет, по крайней мере, две неподвижных буквы.

в) Привести пример (при некотором N) аффинного преобразования, не имеющего неподвижных букв.

9. Пусть элементами сообщения являются биграммы N-буквенного алфавита. Найти формулу для числа различных аффинных шифрующих преобразований. Сколько их при N=26, 27, 29, 30?

10. Перехвачено шифрованное сообщение «PWULPZTQAWHF», полученное применением аффинного отображения биграмм 26-буквенного алфавита, где, как и ранее, биграмме из букв с числовыми эквивалентами x и y соответствует целое число $26x + y$. Статистический анализ предшествующих шифртекстов, полученных применением того же отображения, показал, что в шифртексте чаще всего встречаются биграммы «IX» и «TQ» (в указанном порядке). Известно, что наиболее часто встречающимися биграммами английского текста являются «TH» и «HE» (в указанном порядке).

а) Найти ключ дешифрования и прочитать сообщение.

б) Вы решили захватить получателя сообщения, но так, чтобы отправитель не заподозрил неладное. Поэтому вы хотите выдать себя за сообщника отправителя и отвечаете «GOODWORK» («отлично»). Найти ключ шифрования и соответствующий шифртекст.

11. Вы перехватили сообщение «DXM SCE DCCUVGX », зашифрованное применением аффинного отображения биграмм 30-буквенного алфавита, в котором буквы A–Z имеют числовые эквиваленты 0–25, пробел= 26, ? = 27, ! = 28, ' = 29. Частотный анализ показывает, что в предыдущих шифртекстах чаще всех встречались биграммы «M », «U » и «IH» (в указанном порядке). Допустим, что в английском языке (для данного алфавита) самыми частыми являются биграммы «E », «S » и « T» (в указанном порядке).

а) Определить ключ дешифрования и прочитать сообщение.

б) Найти ключ шифрования и зашифровать сообщение «YES I'M JOKING!».

12. Описанные приемы, разумеется, можно использовать при любых алфавитах. Рассмотрим пример для русского алфавита. Используем следующие числовые эквиваленты для букв кириллицы

А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й
0	1	2	3	4	5	6	7	8	9	10
К	Л	М	Н	О	П	Р	С	Т	У	Ф
11	12	13	14	15	16	17	18	19	20	21
Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
22	23	24	25	26	27	28	29	30	31	32

Предположим, что перехвачено кодированное сообщение «ЦНТИ», зашифрованное применением аффинного отображения биграмм 33-буквенного русского алфавита. Частотный анализ предыдущих шифртекстов показал, что чаще всего в нем встречаются биграммы «ЦА» и «ЫТ» (в указанном порядке). Допустим, что двумя самыми часто встречающимися биграммами в русском языке являются «НО» и «ЕТ» (в указанном порядке). Найти ключ дешифрования и выписать открытый текст сообщения.

13. Пусть, как и в упражнении 8, *неподвижным* элементом открытого текста называется элемент, не изменяющийся при данном шифрующем преобразовании. Найти все неподвижные биграммы для шифрующего преобразования примера 11.

14. Под *композицией* (или *произведением*) двух криптосистем понимается криптосистема, в которой открытый текст шифруется первой криптосистемой, а полученный шифртекст рассматривается как открытый текст для второй криптосистемы, т. е. шифруется еще раз с помощью второй криптосистемы. При более строгом изложении следует потребовать, чтобы множество $\mathcal{C}_1$ элементов шифртекста первой криптосистемы содержалось в множестве элементов открытого текста второй криптосистемы. Пусть f_1 и f_2 — шифрующие функции, тогда композиция криптосистем задается шифрующей функцией $f = f_2 \circ f_1$. Если обозначить через I («intermediate text» — промежуточный текст) элемент шифртекста первой системы, а через $\mathcal{I} = \mathcal{C}_1$ обозначить множество элементов промежуточного текста, то композицию криптосистем можно схематически изобразить диаграммой

$$\mathcal{P} \xrightarrow{f_1} \mathcal{I} \xrightarrow{f_2} \mathcal{C}.$$

Доказать, что:

а) композиция двух шифрующих преобразований сдвига также является шифрующим преобразованием сдвига;

б) композиция двух шифрующих линейных преобразований является шифрующим линейным преобразованием;

в) композиция двух аффинных шифрующих преобразований является аффинным шифрующим преобразованием.

15. Рассмотрим немного более сложную криптосистему, в которой открытый и шифрованный тексты записываются в разных алфавитах. Мы возьмем для открытого текста N-буквенный алфавит, а для шифртекста — M-буквенный, $M > N$. Как обычно, мы рассматриваем биграммы в N-буквенном алфавите как двузначные числа в системе счисления с основанием N, т. е. как целые числа между 0 и $N^2 - 1$. Аналогично, биграммы в M-буквенном алфавите рассматриваются как целые числа между 0 и $M^2 - 1$. Теперь выберем любое целое L между N^2 и M^2: $N^2 \leqslant L \leqslant M^2$. Выберем также целые числа a и b с НОД $(a, L) = 1$. Биграмма P открытого текста шифруется по правилу $C \equiv aP + b \pmod{L}$ (где C — наименьший неотрицательный вычет по модулю L, удовлетворяющий этому сравнению). Здесь $\mathcal{P}$ — множество всех возможных биграмм P — состоит из всех целых чисел из диапазона от 0 до $N^2 - 1$. Множество $\mathcal{C}$ всех возможных биграмм C шифртекста (т. е. биграмм в большем алфавите) составляет лишь часть множества всех целых из диапазона от 0 до $M^2 - 1$ и на самом деле представляет собой подмножество множества тех целых чисел, меньших L, которые получаются применением шифрующего преобразования ко всем возможным биграммам открытого текста. Пусть открытый текст записан в 27-буквенном алфавите упражнения 3, а шифртекст — в 30-буквенном алфавите упражнения 11. Предположим, что $L = 853$. Предположим также, что самые часто встречающиеся биграммы открытого текста

«E » и «S » при шифровании переходят в биграммы «FQ» и «LE» соответственно. Найти ключ дешифрования и прочитать сообщение «YAVAOCH'D!».

16. Продолжая упражнение 15, покажем, как можно без особого труда построить криптоситему, вскрыть которую существенно труднее. Пусть f_1 — криптосистема описанного в упражнении 15 типа, т.е. заданная правилом $f_1(P) \equiv a_1P + b_1 \pmod{L_1}$, и пусть f_2 — другая криптосистема того же типа. Параметры N и M у этих систем одни и те же, но параметры a, b и L у каждой системы свои. Пусть $L_2 > L_1$. Построим *композицию* (произведение) этих двух криптоситем (см. упражнение 14), т.е. будем зашифровывать элемент P открытого текста по правилам

$$I \equiv a_1P + b_1 \pmod{L_1},$$

$$C \equiv a_2I + b_2 \pmod{L_2}.$$

(В первом случае I — неотрицательное целое число, меньшее L_1 и удовлетворяющее первому сравнению, во втором правиле C меньше L_2.) Поскольку модули L_1 и L_2 различны, утверждения из упражнения 14 в) здесь неприменимы, и эта композиция криптосистем, вообще говоря, не является аффинной системой. Здесь мы предполагаем, что алфавиты из N и M букв заданы, но можно произвольным образом выбирать параметры a_1, b_1, L_1, a_2, b_2, L_2, подчиненные, разумеется, условиям $N^2 \leqslant L_1 < L_2 \leqslant M^2$, НОД $(a_1, L_1) = 1$, НОД $(a_2, L_2) = 1$. Таким образом, ключ состоит из шестерки параметров $\{a_1, b_1, L_1, a_2, b_2, L_2\}$. Пусть, как и в упражнении 15, алфавиты открытого и шифрованного текстов содержат по 27 и 30 букв, соответственно. Пусть ключ шифрования имеет вид $\{247, 109, 757, 675, 402, 881\}$. Объяснить, как проводить *де*шифрование, и дешифровать сообщение «D!RAJ'KCTN».

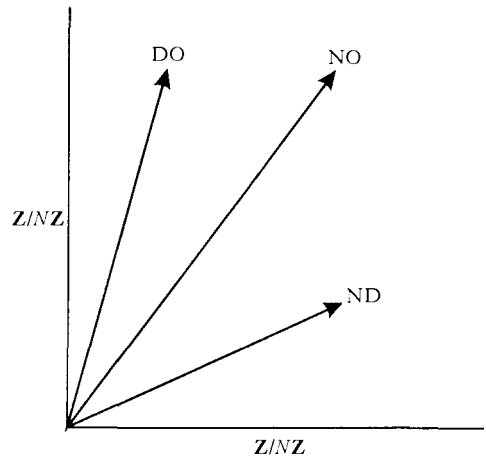
§ 2. Шифрующие матрицы

Пусть имеется N-буквенный алфавит и нам надо передавать биграммы (двухбуквенные блоки) как элементы сообщения. В §1 было показано, что каждой биграмме можно сопоставить целое число по модулю N^2, т.е. элемент из $\mathbf{Z}/N^2\mathbf{Z}$. Другая возможность состоит в сопоставлении биграмме некоторого *вектора*, т.е. пары целых чисел $\binom{x}{y}$ с x и y, рассматриваемыми по модулю N. Например, если имеется 26-буквенный алфавит A–Z с числовыми эквивалентами 0–25 соответственно, то биграмме NO отвечает вектор $\binom{13}{14}$, как показано на диаграмме.

Мы изображаем каждую биграмму как узел в квадратной решетке размера $N \times N$. Наше изображение аналогично обычной координатной плоскости с той разницей, что каждая из осей соответствует не множеству вещественных чисел, а множеству $\mathbf{Z}/N\mathbf{Z}$. Аналогично обозначению $\mathbf{R}^2$ для обычной плоскости будем использовать обозначение $(\mathbf{Z}/N\mathbf{Z})^2$ для этой решетки размера $N \times N$.

Изобразив биграммы векторами (точками на плоскости), мы можем интерпретировать «шифрующее преобразование» как перестановку узлов нашей $N \times N$-решетки. Точнее, шифрующее отображе-

ние — это взаимно однозначное отображение $(\mathbf{Z}/N\mathbf{Z})^2$ в себя.



Замечание. В течение нескольких веков одним из самых распространенных способов шифрования был так называемый «шифр Виженера». Его можно описать следующим образом. При некотором заданном k рассмотрим блоки из k букв как векторы в $(\mathbf{Z}/N\mathbf{Z})^k$. Зафиксируем некоторый вектор $b \in (\mathbf{Z}/N\mathbf{Z})^k$ (обычно выбирается вектор, соответствующий какому-нибудь легко запоминаемому «ключевому слову»). Шифрование проводится посредством векторного сдвига $C = P + b$, где элемент шифртекста C и элемент открытого текста P представляют собой векторы с целочисленными координатами по модулю N. К сожалению, эта криптосистема вскрывается почти так же просто, как сдвиг отдельных букв (см. пример 1 в предыдущем разделе). А именно, при известных или угаданных N и k нужно просто разбить текст на блоки по k букв и провести частотный анализ первых букв в каждом блоке, чтобы определить первую компоненту вектора b. Затем исследуются вторые буквы блоков, и т. д.

Сведения из линейной алгебры. Напомним, как проводятся операции с векторами в вещественной двумерной плоскости и с 2×2-матрицами из вещественных чисел. Напомним, что, имея 2×2-матрицу $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ и вектор $\binom{x}{y}$ (мы будем записывать векторы как столбцы), можно получить новый вектор, *применив матрицу к вектору* по правилу

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

При фиксированной матрице эта функция, переводящая вектор в вектор, называется *линейным преобразованием*, так как она перестановочна с операциями сложения векторов и умножения векторов на кон-

станты. В этих обозначениях любая система уравнений вида $ax+by=e$, $cx+dy=f$ эквивалентна матричному уравнению $AX=B$, где A обозначает матрицу

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

X обозначает вектор неизвестных $\binom{x}{y}$, а B — вектор констант $\binom{e}{f}$. Словесно эту систему уравнений можно интерпретировать как задачу поиска вектора, при «умножении» которого на заданную матрицу получается заданный вектор. Таким образом, эта система аналогична уравнению $ax=b$, которое решается умножением обеих частей на a^{-1} (если $a \neq 0$). Аналогичный вид имеет один из способов решения матричного уравнения $AX=B$: находится обратная к матрице A матрица A^{-1}, умножение на которую обеих частей уравнения дает единственное векторное решение $X=A^{-1}B$.

Под обратной к матрице A понимается такая матрица, умножение которой на A дает единичную матрицу

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(применение этой матрицы к любому вектору оставляет этот вектор неизменным). Но не все матрицы имеют обратные. Нетрудно показать, что матрица

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

имеет обратную тогда и только тогда, когда отличен от нуля ее *определитель* $D \stackrel{\text{def}}{=} ad-bc$. В этом случае обратная матрица равна

$$\frac{1}{D}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\begin{pmatrix} D^{-1}d & -D^{-1}b \\ -D^{-1}c & D^{-1}a \end{pmatrix}.$$

Сущетвует три возможных вида множества решений системы уравнений $AX=B$. Во-первых, если определитель D отличен от нуля, то система имеет единственное решение $X=\binom{x}{y}$. Если же $D=0$, то либо решений нет, либо их бесконечно много. Эти три варианта имеют простую геометрическую интерпретацию. Два уравнения соответствуют двум прямым на координатной плоскости. Если $D \neq 0$, то прямые пересекаются в единственной точке (x,y). В противном случае они либо параллельны и не пересекаются вовсе (уравнения системы не имеют общих решений), либо совпадают (уравнения системы имеют бесконечно много общих решений).

Пусть имеется набор векторов $X_1 = \binom{x_1}{y_1}, \ldots, X_k = \binom{x_k}{y_k}$, расположенных в столбцах $2 \times k$-матрицы. Определим произведение матриц

$$AX = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \ldots x_k \\ y_1 \ldots y_k \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} ax_1 + by_1 & \ldots & ax_k + by_k \\ cx_1 + dy_1 & \ldots & cx_k + dy_k \end{pmatrix},$$

т. е. матрица A просто применяется поочередно к каждому столбцу, давая при этом очередной столбец произведения. Например, произведение 2×2-матриц вычисляется по формуле

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}.$$

Аналогичными свойствами обладают 3×3-матрицы, которые применяются к трехмерным векторам, и т. д. Однако формулы для определителя и обратной матрицы усложняются. На этом мы завершаем краткий обзор линейной алгебры над множеством вещественных чисел.

Линейная алгебра по модулю N. В §1, имея дело с отдельными символами и шифрующими отображениями $\mathbf{Z}/N\mathbf{Z}$, мы обнаружили два удобных для использования класса отображений:

(a) «линейные» отображения $C = aP$, где a обратимо в $\mathbf{Z}/N\mathbf{Z}$,

(b) «аффинные» отображения $C = aP + b$, где a обратимо в $\mathbf{Z}/N\mathbf{Z}$.

Аналогичная ситуация возникает, когда элементами сообщения являются биграммы-векторы. Сначала рассмотрим линейные отображения. При работе с $(\mathbf{Z}/N\mathbf{Z})^2$, в отличие от случая $\mathbf{Z}/N\mathbf{Z}$, вместо целого a мы используем 2×2-матрицу, которую будем обозначать через A. Сначала выясним, матрицы какого типа нам потребуются.

Пусть R — любое коммутативное кольцо, т. е. множество с умножением и сложением, удовлетворяющими тем же правилам, что и в поле, за исключением того, что *допускается* существование ненулевых элементов, не имеющих обратных по умножению. Например, $\mathbf{Z}/N\mathbf{Z}$ — всегда кольцо, но оно не поле, если число N не является простым. Через R^* обозначим подмножество обратимых элементов R. Например, $(\mathbf{Z}/N\mathbf{Z})^* = \{0 < j < N \mid \text{НОД}\ (j, N) = 1\}$.

Для коммутативного кольца R обозначим через $M_2(R)$ множество всех 2×2-матриц с элементами из R с обычными для матриц операциями сложения и умножения. Мы называем $M_2(R)$ ***кольцом матриц над*** R; $M_2(R)$ — кольцо, но *не*коммутативное, так как произведение матриц зависит от порядка сомножителей.

Ранее в этом разделе рассматривался случай, когда $R = \mathbf{R}$ есть кольцо (более того, поле) вещественных чисел. Напомним, что

матрица

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

с вещественными элементами a, b, c, d имеет обратную тогда и только тогда, когда определитель $D = ad - bc$ отличен от нуля. В этом случае обратная матрица имеет вид

$$\begin{pmatrix} D^{-1}d & -D^{-1}b \\ -D^{-1}c & D^{-1}a \end{pmatrix}.$$

Такая же ситуация имеет место для произвольного кольца R.

Именно, пусть

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(R)$$

и $D = \det A \stackrel{\text{def}}{=} ad - bc$ принадлежит R^*. Пусть D^{-1} обозначает обратный по умножению элемент к D в R. Тогда

$$\begin{pmatrix} D^{-1}d & -D^{-1}b \\ -D^{-1}c & D^{-1}a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} D^{-1}(da - bc) & 0 \\ 0 & D^{-1}(-cb + ad) \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Тот же результат

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

мы получим, умножив матрицы в обратном порядке. Таким образом, обратная к A матрица задается той же самой формулой, что и для матриц из вещественных чисел:

$$A^{-1} = \begin{pmatrix} D^{-1}d & -D^{-1}b \\ -D^{-1}c & D^{-1}a \end{pmatrix}.$$

Пример 1. Найти обратную к матрице

$$A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix} \in M_2(\mathbf{Z}/26\mathbf{Z}).$$

Решение. Здесь $D = 2 \cdot 8 - 3 \cdot 7 = -5 = 21$ в $\mathbf{Z}/26\mathbf{Z}$. Так как НОД $(21, 26) = 1$, то определитель D обратим, а именно, $21^{-1} = 5$. Таким образом,

$$A^{-1} = \begin{pmatrix} 5 \cdot 8 & -5 \cdot 3 \\ -5 \cdot 7 & 5 \cdot 2 \end{pmatrix} = \begin{pmatrix} 40 & -15 \\ -35 & 10 \end{pmatrix} = \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}.$$

Проверим:

$$\begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 105 & 130 \\ 104 & 131 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Здесь, поскольку операции проводятся в $\mathbf{Z}/26\mathbf{Z}$, знак «=» означает, что элементы матриц сравнимы по модулю 26.

Как и в случае вещественных чисел, 2×2-матрицу

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

с элементами из кольца R можно умножить на вектор-столбец $\binom{x}{y}$ с $x, y \in R$ и получить новый вектор $\binom{x'}{y'}$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

Это отображение является линейным: оно переводит линейную комбинацию $\binom{k_1 x_1 + k_2 x_2}{k_1 y_1 + k_2 y_2}$, где k_1 и k_2 принадлежат R, в $\binom{k_1 x_1' + k_2 x_2'}{k_1 y_1' + k_2 y_2'}$. Единственное отличие рассматриваемого случая от предыдущего состоит в том, что вместо вещественных чисел здесь мы имеем дело с элементами кольца R.

Нас особенно будет интересовать случай $R = \mathbf{Z}/N\mathbf{Z}$. Следующее предложение формулируется именно для этого случая, хотя аналогичное утверждение верно для любого R.

Предложение III. 2. 1. *Пусть*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbf{Z}/N\mathbf{Z}) \quad \text{и} \quad D = ad - bc.$$

Следующие утверждения эквивалентны:

a) НОД $(D, N) = 1$;

b) *матрица A имеет обратную*;

c) *если хотя бы один из элементов $x, y \in \mathbf{Z}/N\mathbf{Z}$ отличен от нуля, то $A\binom{x}{y} \neq \binom{0}{0}$.*

d) *матрица A задает взаимно однозначное отображение множества $(\mathbf{Z}/N\mathbf{Z})^2$ на себя.*

Доказательство. Мы уже видели, что a) $\Longrightarrow$ b). Поэтому достаточно показать, что b) $\Longrightarrow$ d) $\Longrightarrow$ c) $\Longrightarrow$ a).

Предположим, что выполнено b). Тогда выполнено и утверждение d), так как A^{-1} задает обратное отображение $\binom{x'}{y'}$ в $\binom{x}{y}$. Далее, если выполнено d), то $\binom{x}{y} \neq \binom{0}{0}$ влечет за собой, что $A\binom{x}{y} \neq A\binom{0}{0} = \binom{0}{0}$, и, следовательно, c) выполнено. Наконец, докажем, что c) $\Longrightarrow$ a),

рассуждением от противного. Итак, пусть a) неверно. Положим $m = \text{НОД}\,(D, N) > 1$ и $m' = N/m$. Возможны три случая.

Случай 1. Если все четыре элемента матрицы A делятся на m, то возьмем $\binom{x}{y} = \binom{m'}{m'}$, что приведет к противоречию со свойством c).

Случай 2. Если хотя бы один из элементов a и b не делится на m, то возьмем $\binom{x}{y} = \binom{-bm'}{am'}$. Тогда

$$A\binom{x}{y} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\binom{-bm'}{am'} = \binom{-abm' + bam'}{-cbm' + dam'} = \binom{0}{Dm'} = \binom{0}{0},$$

так как $m|D$ и, следовательно, $N = mm'|Dm'$.

Случай 3. Если хотя бы один из элементов c и d не делится на m, то возьмем $\binom{x}{y} = \binom{dm'}{-cm'}$ и продолжим рассуждения, как в предыдущем случае.

Рассмотренные случаи исчерпывают все возможности. Таким образом, нарушение a) влечет за собой нарушение c). Это завершает доказательство предложения.

Пример 2. Решить следующие системы сравнений:

a)
$$2x + 3y \equiv 1 \pmod{26},$$
$$7x + 8y \equiv 2 \pmod{26};$$

b)
$$x + 3y \equiv 1 \pmod{26},$$
$$7x + 9y \equiv 2 \pmod{26};$$

c)
$$x + 3y \equiv 1 \pmod{26},$$
$$7x + 9y \equiv 1 \pmod{26}.$$

Решение. В матричной форме система a) имеет вид $AX \equiv B \pmod{26}$, где A — матрица из примера 1, $X = \binom{x}{y}$, $B = \binom{1}{2}$. Получаем единственное решение

$$X \equiv A^{-1}B \equiv \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}\binom{1}{2} \equiv \binom{10}{11} \pmod{26}.$$

Матрица систем b) и c) не имеет обратной по модулю 26, так как ее определитель равен 14 и имеет с модулем 26 общий множитель 2. Однако можно перейти к модулю 13, найти решение системы по этому модулю и проверить, не дает ли оно решений по модулю 26. По модулю 13 получаем

$$\binom{x}{y} \equiv \begin{pmatrix} 9 & 10 \\ 6 & 1 \end{pmatrix}\binom{e}{f}$$

(где $\binom{e}{f} = \binom{1}{2}$ для системы b) и $\binom{1}{1}$ для системы c)). Это дает $\binom{x}{y} \equiv \binom{3}{8}$ и $\binom{6}{7}$ (mod 13) соответственно. Проверяя варианты, возникающие при переходе к модулю 26, убеждаемся, что система b) решений *не* имеет, а система c) имеет два решения $x = 6$, $y = 7$ и $x = 19$, $y = 20$.

Другой способ решения систем уравнений (иногда более предпочтительный, особенно при необратимой матрице) состоит в исключении одного из неизвестных (так, в системах b) и c) можно было вычесть семь раз первое сравнение из второго).

Возвращаясь к криптографии, видим, что согласно предложению III. 2. 1 можно построить преобразование шифрования наших биграмм-векторов, используя матрицы $A \in M_2(\mathbf{Z}/N\mathbf{Z})$, у которых определитель не имеет общих множителей с N:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad D = ad - bc, \quad \text{НОД}\ (D, N) = 1.$$

А именно, элемент открытого текста $P = \binom{x}{y}$ преобразуется в элемент шифртекста $C = \binom{x'}{y'}$ по правилу

$$C = AP, \quad \text{т. е.} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Чтобы дешифровать сообщение, просто применяем обратную матрицу: $P = A^{-1}AP = A^{-1}C$, т. е.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} D^{-1}d & -D^{-1}b \\ -D^{-1}c & D^{-1}a \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Пример 3. В 26-буквенном алфавите, используя матрицу A примера 1, зашифровать элемент «NO».

Решение. Имеем

$$AP = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 13 \\ 14 \end{pmatrix} = \begin{pmatrix} 68 \\ 203 \end{pmatrix} = \begin{pmatrix} 16 \\ 21 \end{pmatrix},$$

и $C = AP$ есть «QV».

Замечание. Чтобы зашифровать последовательность $P = P_1P_2P_3 \ldots P_k$ из k биграмм открытого текста, можно построить $2 \times k$-матрицу из k векторов-столбцов (обозначим ее также через P) и, умножив 2×2-матрицу A на $2 \times k$-матрицу P, получить $2 \times k$-матрицу $C = AP$ из шифрованных биграмм-векторов.

Пример 4. Продолжая пример 3, зашифровать открытый текст «NOANSWER».

Р е ш е н и е. Числовым эквивалентом для «NOANSWER» является последовательность векторов $\binom{13}{14}\binom{0}{13}\binom{18}{22}\binom{4}{17}$. Имеем

$$C = AP = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 13 & 0 & 18 & 4 \\ 14 & 13 & 22 & 17 \end{pmatrix} = \begin{pmatrix} 68 & 39 & 102 & 59 \\ 203 & 104 & 302 & 164 \end{pmatrix}$$
$$= \begin{pmatrix} 16 & 13 & 24 & 7 \\ 21 & 0 & 16 & 8 \end{pmatrix},$$

т. е. шифрованное сообщение имеет вид «QVNAYQHI».

П р и м е р 5. В ситуации примеров 3 и 4 дешифровать сообщение «FWMDIQ».

Р е ш е н и е. Имеем

$$P = A^{-1}C = \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}\begin{pmatrix} 5 & 12 & 8 \\ 22 & 3 & 16 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 19 & 2 \\ 19 & 0 & 10 \end{pmatrix} = \text{«ATTACK»}.$$

Как и в § 1, будем считать, что у нас имеется некоторая ограниченная информация, которую мы хотим использовать для дешифрования отрезка шифртекста. Нам известно, что «противник» использует биграммы-векторы и линейное шифрующее преобразование $C = AP$ в N-буквенном алфавите. Однако мы не знаем ни ключа шифрования (матрицы A), ни ключа дешифрования (матрицы A^{-1}). Но предположим, что у нас есть возможность найти две пары биграмм открытого и шифрованного текстов $C_1 = AP_1$ и $C_2 = AP_2$. Возможно, эти сведения нам удалось получить частотным анализом биграмм в длинном отрезке шифртекста или каким-либо образом стало известно, во что шифруется некоторый четырехбуквенный отрезок текста. В этом случае для определения матриц A и A^{-1} мы можем сделать следующее. Объединим два столбца P_1 и P_2 в 2×2-матрицу P и таким же образом поступим со столбцами шифртекста. Мы получим уравнение для 2×2-матриц: $C = AP$, где C и P нам известны, а A является неизвестной. Можно решить это уравнение, домножив обе части на P^{-1}:

$$A = APP^{-1} = CP^{-1}.$$

Аналогично, из уравнения $P = A^{-1}C$ находится A^{-1}: $A^{-1} = PC^{-1}$.

П р и м е р 6. Пусть известно, что противник использует при шифровании 2×2-матрицу с элементами из 29-буквенного алфавита, где A–Z имеют обычные числовые эквиваленты, пробел=26, ?=27, !=28. Принято сообщение

«GFPYJP X?UYXSTLADPLW»,

причем предполагается, что последние пять букв открытого текста — подпись «KARLA». Этих букв хватает для построения лишь двух биграмм. Итак, биграммам «DP» и «LW» шифртекста в открытом тексте отвечают биграммы «AR» и «LA» соответственно. Значит, матрица P, составленная из «AR» и «LA», получается как результат умножения неизвестной матрицы дешифрования A^{-1} на матрицу C, составленную из «DP» и «LW»:

$$\begin{pmatrix} 0 & 11 \\ 17 & 0 \end{pmatrix} = A^{-1} \begin{pmatrix} 3 & 11 \\ 15 & 22 \end{pmatrix}.$$

Таким образом,

$$A^{-1} = \begin{pmatrix} 0 & 11 \\ 17 & 0 \end{pmatrix} \begin{pmatrix} 3 & 11 \\ 15 & 22 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 11 \\ 17 & 0 \end{pmatrix} \begin{pmatrix} 3 & 13 \\ 23 & 7 \end{pmatrix} = \begin{pmatrix} 21 & 19 \\ 22 & 18 \end{pmatrix},$$

и все открытое сообщение имеет вид

$$\begin{pmatrix} 21 & 19 \\ 22 & 18 \end{pmatrix} \begin{pmatrix} 6 & 15 & 9 & 26 & 27 & 24 & 18 & 11 & 3 & 11 \\ 5 & 24 & 15 & 23 & 20 & 23 & 19 & 0 & 15 & 22 \end{pmatrix}$$
$$= \begin{pmatrix} 18 & 17 & 10 & 26 & 19 & 13 & 14 & 28 & 0 & 11 \\ 19 & 8 & 4 & 0 & 26 & 14 & 13 & 10 & 17 & 0 \end{pmatrix}$$
$$= \text{«STRIKE AT NOON!KARLA»}.$$

З а м е ч а н и е. Чтобы описанный метод работал, нужна обратимость матрицы P, образованной из двух известных биграмм открытого текста, т. е. определитель D не должен иметь общих множителей с числом букв N. Что же делать, если нам не повезет? Если известна еще одна пара биграмм открытого и шифрованного текстов, можно попробовать использовать ее в матрицах P и C вместо первых или вторых столбцов в надежде получить обратимую матрицу. Пусть у нас нет такой возможности или ни одна известная пара не дает обратимой матрицы P. Тогда мы не можем определить матрицу A^{-1} точно. Однако имеющейся у нас информации достаточно для того, чтобы значительно сократить число возможных вариантов для дешифрующей матрицы. Проиллюстрируем это на примере. (См. также упражнения в конце раздела.)

П р и м е р 7. Известно, что противник использует шифрующую матрицу A с элементами из 26-буквенного алфавита. Перехвачено сообщение «WKNCCHSSJH». Известно также, что первое слово — это «GIVE». Мы хотим найти матрицу дешифрования A^{-1} и прочитать сообщение.

Р е ш е н и е. Воспользовавшись процедурой примера 6 и записав

$$P = \text{«GIVE»} = \begin{pmatrix} 6 & 21 \\ 8 & 4 \end{pmatrix},$$

$$C = \text{«WKNC»} = \begin{pmatrix} 22 & 13 \\ 10 & 2 \end{pmatrix}, \qquad A^{-1} = PC^{-1},$$

мы сразу столкнемся с трудностями, так как $\det C = 18$ и НОД $(18, 26) = 2$. Мы можем продолжить следующим образом. Пусть $\overline{A}$ обозначает приведенную по модулю 13 матрицу A. Аналогичный смысл имеют обозначения $\overline{P}$ и $\overline{C}$. Рассматривая эти матрицы в $M_2(\mathbf{Z}/13\mathbf{Z})$, можно найти C^{-1} (точнее, $\overline{C}^{-1}$), так как НОД $(\det C, 13) = 1$. Поэтому из равенства $\overline{P} = \overline{A}^{-1}\overline{C}$ получим

$$\overline{A}^{-1} = \overline{P}\,\overline{C}^{-1} = \begin{pmatrix} 6 & 8 \\ 8 & 4 \end{pmatrix}\begin{pmatrix} 9 & 0 \\ 10 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 4 \\ 3 & 2 \end{pmatrix}.$$

Таким образом, элементы матрицы A^{-1}, являющиеся целыми по модулю 26, должны приводится по модулю 13 к элементам матрицы

$$\begin{pmatrix} 2 & 4 \\ 3 & 2 \end{pmatrix}.$$

Следовательно, для каждого элемента этой матрицы есть по две возможности. Точнее,

$$A^{-1} = \begin{pmatrix} 2 & 4 \\ 3 & 2 \end{pmatrix} + 13A_1,$$

где $A_1 \in M_2(\mathbf{Z}/2\mathbf{Z})$ есть 2×2-матрица из нулей и единиц. Поэтому остается $2^4 = 16$ возможных вариантов. Однако, что самое главное, так как матрица A^{-1} обратима, ее определитель должен быть взаимно прост с 26 и, в частности, быть нечетным. Это соображение оставляет возможными всего шесть вариантов для A_1. Далее, когда мы подставим

$$\begin{pmatrix} 2 & 4 \\ 3 & 2 \end{pmatrix} + 13A_1$$

вместо A^{-1} в уравнение

$$A^{-1}\begin{pmatrix} 22 & 13 \\ 10 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 21 \\ 8 & 4 \end{pmatrix}$$

(последнее равенство имеет смысл поэлементного сравнения по модулю 26), мы исключим все возможности, кроме двух следующих:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

т. е.

$$A^{-1} = \begin{pmatrix} 15 & 4 \\ 16 & 15 \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} 15 & 17 \\ 16 & 15 \end{pmatrix}.$$

Попытка дешифрования с первой матрицей приводит к заведомо сомнительному варианту «GIVEGHEMHP». Дешифрование с матрицей

$$A^{-1} = \begin{pmatrix} 15 & 17 \\ 16 & 15 \end{pmatrix}$$

дает ответ «GIVETHEMUP». Значит, последний вариант — правильный. Этот метод предполагает некоторое количество проб и ошибок, но в любом случае это лучше проверки всех 157 248 вариантов матрицы дешифрования $A^{-1} \in M_2(\mathbf{Z}/26\mathbf{Z})^*$.

Замечание. В примере 7, возможно, удобнее было бы подправить элементы матрицы A^{-1}, прибавив к ним кратные 13 так, чтобы все они делились на два, т. е. определить матрицу A_1 равенством

$$A^{-1} = \begin{pmatrix} 2 & 4 \\ 16 & 2 \end{pmatrix} + 13A_1.$$

Тогда можно искать матрицу A_1, оперируя по модулю 2, так как $A_1 C \equiv P \pmod 2$.

Шифрующие аффинные преобразования. Более общий способ шифрования биграммы-вектора $P = \binom{x}{y}$ состоит в умножении его на 2×2-матрицу $A = \binom{a\ b}{c\ d} \in M_2(\mathbf{Z}/N\mathbf{Z})$ и прибавлении к результату вектора констант $B = \binom{e}{f}$:

$$C = AP + B,$$

т. е.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} ax + by + e \\ cx + dy + f \end{pmatrix}.$$

Это преобразование называется *аффинным* отображением. Оно аналогично шифрующей функции $C = aP + b$ из §1, где рассматривались однобуквенные элементы сообщений. Разумеется, как и раньше, знак равенства понимается как сравнимость элементов по модулю N.

Обратное преобразование, выражающее P через C, может быть получено вычитанием вектора B из обеих частей и последующим применением к ним A^{-1}:

$$P = A^{-1}C - A^{-1}B.$$

Оно также является аффинным преобразованием $P = A'C + B'$ с $A' = A^{-1}$ и $B' = -A^{-1}B$. Заметим, что для однозначности дешифрования требуется обратимость матрицы A.

Пусть известно, что противник использует аффинное шифрующее преобразование биграмм-векторов из N-буквенного алфавита. Для нахождения A и B (или $A' = A^{-1}$ и $B' = -A^{-1}B$) теперь нужны, по

крайней мере, *три* пары биграмм. Пусть, например, известно, что биграммы C_1, C_2, C_3 шифртекста отвечают биграммам открытого текста P_1, P_2, P_3:

$$P_1 = A'C_1 + B',$$
$$P_2 = A'C_2 + B',$$
$$P_3 = A'C_3 + B'.$$

Для нахождения A' и B' мы можем сделать следующее. Вычтем последнее уравнение из первых двух и образуем 2×2-матрицу P из столбцов $P_1 - P_3$ и $P_2 - P_3$, а 2×2-матрицу C — из столбцов $C_1 - C_3$ и $C_2 - C_3$. Получим матричное уравнение $P = A'C$, которое, как и в случае линейных преобразований, может быть разрешено относительно A' (при условии, что матрица C обратима). Наконец, определив $A' = A^{-1}$, находим из любого из приведенных выше уравнений B', например, $B' = P_1 - A'C_1$.

УПРАЖНЕНИЯ

1. Используя частотный метод, дешифровать сообщение в 26-буквенном алфавите, закодированное шифром Виженера с трехбуквенным ключевым словом. Сделать это следующим образом. Для определения первой буквы ключевого слова использовать последовательность из каждой третьей буквы сообщения, начиная с первой. Не надо предполагать, что самой частой буквой сообщения обязательно является буква «E». Следует просмотреть в качестве варианта шифра для буквы «E» четыре наиболее часто встречающихся буквы. Если при выборе варианта для «E» оказывается, что хотя бы одна из трех других часто встречающихся букв отвечает сравнительно редкой букве, скажем, «Q» или «Z», то это свидетельствует о неправильном выборе шифра для «E». Исключая такие варианты, найти шифр для «E», затем букву ключевого слова, использованного при шифровании. Таким способом определить ключевое слово и дешифровать сообщение

AWYVPQCTBLWYLPASQJWUPGBUSHFACELDLLDLWLBWA

FAHSEBYJXXACELWCJTQMARKDDLWCSXBUDLKDPLXSEQCJT

NWPRWSRGBCLWPGJEZIFWIMJDLLDAGCQMAYLTGLPPJXTWS

GFRMVTLGUYUXJAIGWHCPXQLTBXDPVTAGSGFVRZTWTGM

MVFLXRLDKWPRLWCSXPHDPLPKSHQGULMBZWGQAPQCTBA

URZTWSHQMBCVXAGJJVGCSSGLIFWNQSXBFDGSHIWSFGLRZ

TWEPLSVCVIFWNQSXBOWCFHMETRZXLYPPJXTWSGFRMVTR

ZTWHWMFTBOPQZXLYIMFPLVWYVIFWDPAVGFPJETQKPEWG

CSSRGIFWB

2. Найти обратные к следующим матрицам по модулю N. Записать элементы обратных матриц как неотрицательные целые, меньшие N.

а) $\begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix} \pmod 5$, б) $\begin{pmatrix} 1 & 3 \\ 4 & 3 \end{pmatrix} \pmod{29}$, в) $\begin{pmatrix} 15 & 17 \\ 4 & 9 \end{pmatrix} \pmod{26}$,

г) $\begin{pmatrix} 40 & 0 \\ 0 & 21 \end{pmatrix} \pmod{841}$, д) $\begin{pmatrix} 197 & 62 \\ 603 & 271 \end{pmatrix} \pmod{841}$.

В упражнениях 3–5 найдите все решения $\binom{x}{y}$ по модулю N, рассматривая x и y как неотрицательные целые числа, меньшие N.

3.

а) $x + 4y \equiv 1 \pmod 9$
$5x + 7y \equiv 1 \pmod 9$

б) $x + 4y \equiv 1 \pmod 9$
$5x + 8y \equiv 1 \pmod 9$

в) $x + 4y \equiv 1 \pmod 9$
$5x + 8y \equiv 2 \pmod 9$

г) $x + 4y \equiv 0 \pmod 9$
$5x + 8y \equiv 0 \pmod 9$

4.

а) $17x + 11y \equiv 7 \pmod{29}$
$13x + 10y \equiv 8 \pmod{29}$

б) $17x + 11y \equiv 0 \pmod{29}$
$13x + 10y \equiv 0 \pmod{29}$

в) $9x + 20y \equiv 0 \pmod{29}$
$16x + 13y \equiv 0 \pmod{29}$

г) $9x + 20y \equiv 10 \pmod{29}$
$16x + 13y \equiv 21 \pmod{29}$

д) $9x + 20y \equiv 1 \pmod{29}$
$16x + 13y \equiv 2 \pmod{29}$

5.

а) $480x + 971y \equiv 416 \pmod{1111}$
$297x + 398y \equiv 319 \pmod{1111}$

б) $480x + 971y \equiv 109 \pmod{1111}$
$297x + 398y \equiv 906 \pmod{1111}$

в) $480x + 971y \equiv 0 \pmod{1111}$
$297x + 398y \equiv 0 \pmod{1111}$

г) $480x + 971y \equiv 0 \pmod{1111}$
$298x + 398y \equiv 0 \pmod{1111}$

д) $480x + 971y \equiv 648 \pmod{1111}$
$298x + 398y \equiv 1004 \pmod{1111}$

6. Числа *Фибоначчи* определяются правилами $f_1 = 1$, $f_2 = 1$, $f_3 = 2$, $f_{n+1} = f_n + f_{n-1}$ при $n > 1$, или, в эквивалентной форме, матричным уравнением

$$\begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n$$

(см. упражнение 10 к § I.2). Используя определение в матричной форме, доказать, что f_n четно тогда и только тогда, когда n делится на 3. Обобщить этот результат, доказав, что f_n делится на a тогда и только тогда, когда n делится на b, при

следующих a и b: а) $a = 2$, $b = 3$; б) $a = 3$, $b = 4$; в) $a = 5$, $b = 5$; г) $a = 7$, $b = 8$; д) $a = 8$, $b = 6$; е) $a = 11$, $b = 10$.

7. Перехвачено сообщение «SONAFQCHMWPTVEVY», зашифрованное *линейным* отображением биграмм-векторов, причем известно, что отправитель использовал обычный 26-буквенный алфавит A–Z с числовыми эквивалентами 0–25 соответственно. Предварительный статистический анализ длинного отрезка перехваченного шифртекста показал, что самыми частыми биграммами в шифртексте являются «KH» и «XW» (в указанном порядке). Предположим, что эти биграммы отвечают биграммам «TH» и «HE» соответственно, как самым частым биграммам в большинстве длинных открытых текстов, относящихся к предполагаемому предмету переписки. Найти дешифрующую матрицу и прочитать сообщение.

8. Перехвачено сообщение «ZRIXXYVBMNPO», зашифрованное линейным отображением биграмм-векторов при использовании 27-буквенного алфавита, причем буквы A–Z имеют числовые эквиваленты 0–25 соответственно, пробел = 26. Обнаружено, что чаще всего в шифртексте встречаются биграммы «PK» и «RZ». Предположим, что они отвечают самым частым биграммам 27-буквенного алфавита «E » (E и пробел) и «S » соответственно. Найти дешифрующую матрицу и прочитать сообщение.

9. Перехвачено сообщение «!IWGV1EX!ZRADRYD», зашифрованное линейным шифрующим отображением биграмм-векторов при использовании 29-буквенного алфавита, причем буквы A–Z имеют числовые эквиваленты 0–25 соответственно, пробел = 26, ? = 27, ! = 28. Известно, что последние пять букв открытого текста — подпись отправителя «MARIA».

а) Найти дешифрующую матрицу и прочитать сообщение.

б) Найти шифрующую матрицу и, представившись другом Марии по имени Джо, зашифровать сообщение «DAMN FOG! JO».

10. В этом упражнении снова имеем дело с кириллицей (см. упражнение 12 предыдущего раздела). Мы используем 34-буквенный алфавит с традиционными числовыми эквивалентами для букв, пробел = 33. Предположим, что наиболее частыми биграммами в русском языке являются «НО» и «ЕТ». Пусть известно, что в длинных отрезках шифртекста чаще всех встречаются биграммы «ЮТ» и «ЧМ». Известно также, что шифрование производится линейным преобразованием биграмм-векторов для 34-буквенного алфавита. Прочитать перехваченное сообщение «СХНСЪШОНЩЗ».

11. Доказать, что *композиция* (см. упражнение 14 предыдущего раздела) криптосистемы с шифрующей матрицей $A_1 \in M_2(\mathbf{Z}/N\mathbf{Z})^*$ и криптосистемы с шифрующей матрицей $A_2 \in M_2(\mathbf{Z}/N\mathbf{Z})^*$ также является криптосистемой с линейным шифрующим преобразованием.

12. Для того чтобы затруднить вскрытие криптосистемы, решено сначала зашифровать биграмму-вектор в 26-буквенном алфавите, используя матрицу

$$\begin{pmatrix} 3 & 11 \\ 4 & 15 \end{pmatrix}$$

по модулю 26, а потом к результату применить матрицу

$$\begin{pmatrix} 10 & 15 \\ 5 & 9 \end{pmatrix}$$

по модулю 29. (Заметим, что последовательное применение двух матриц по одному модулю, как показано в упражнении 11, эквивалентно применению одной матрицы,

применение двух матриц по разным модулям дает более сложное преобразование.) Таким образом, открытый текст будет записан в 26-буквенном алфавите, а шифртекст — в 29-буквенном алфавите упражнения 9.

а) Зашифровать сообщение «SEND».

б) Описать, как производится дешифрование, и дешифровать сообщение «ZMOY».

13. Доказать, что если биграммы-векторы зашифровываются по формуле $C = AP$ с необратимой матрицей $A \in M_2(\mathbf{Z}/N\mathbf{Z})$, то любой посланный шифртекст может быть дешифрован, по крайней мере, в два различных открытых текста.

14. Перехвачено сообщение «S GNLIKD?KOZQLLIOMKUL.VY» (пробел после S является частью сообщения). Предположим, что использовано линейное шифрующее преобразование $C = AP$ в 30-буквенном алфавите, в котором A–Z имеют числовые эквиваленты 0–25, пробел = 26, . = 27, , = 28, ? = 29. Известно также, что последние шесть букв открытого текста — подпись KARLA и точка. Найти дешифрующую матрицу A^{-1} и весь открытый текст.

15. Перехвачено сообщение «KVW? TA!KJB?FVR». (Пробелы после ? и R входят в сообщение, а завершающая точка — нет). Известно, что использовано линейное шифрующее преобразование в 30-буквенном алфавите, в котором A–Z имеют числовые обозначения 0–25, пробел = 26, ? = 27, ! = 28, . = 29. Известно также, что первые шесть букв сообщения — это «C.I.A.». Найти дешифрующую матрицу A^{-1} и весь открытый текст.

16. Предположим, что $N = nm$, где НОД $(m, n) = 1$. Любую матрицу $A \in M_2(\mathbf{Z}/N\mathbf{Z})$ можно вложить в $M_2(\mathbf{Z}/m\mathbf{Z})$ или $M_2(\mathbf{Z}/n\mathbf{Z})$ простым приведением элементов по модулю m или n. Пусть $\overline{A}$ и $\widetilde{A}$ обозначают соответствующие матрицы в $M_2(\mathbf{Z}/m\mathbf{Z})$ или $M_2(\mathbf{Z}/n\mathbf{Z})$.

а) Доказать, что отображение, сопоставляющее матрице A пару $(\overline{A}, \widetilde{A})$, является взаимно однозначным соответствием между $M_2(\mathbf{Z}/N\mathbf{Z})$ и множеством $M_2(\mathbf{Z}/m\mathbf{Z}) \times M_2(\mathbf{Z}/n\mathbf{Z})$ всех пар матриц, одна из которых является матрицей по модулю m, а другая — по модулю n.

б) Доказать, что это отображение задает взаимно однозначное соответствие между множеством $M_2(\mathbf{Z}/N\mathbf{Z})^*$ обратимых матриц по модулю N и множеством $M_2(\mathbf{Z}/m\mathbf{Z})^* \times M_2(\mathbf{Z}/n\mathbf{Z})^*$.

17. При простом p найти двумя разными способами число элементов множества $M_2(\mathbf{Z}/p\mathbf{Z})^*$ и убедиться в совпадении результатов:

а) подсчитать число решений в $\mathbf{F}_p$ уравнения $ad - bc = 0$ и вычесть эту величину из числа элементов множества $M_2(\mathbf{Z}/p\mathbf{Z})$;

б) любая $A \in M_2(\mathbf{Z}/p\mathbf{Z})^*$ переводит векторы $\binom{1}{0}$ и $\binom{0}{1}$ в пару линейно независимых векторов, из которых первый — любой ненулевой вектор, а второй не пропорционален первому; подсчитать число вариантов.

18. Доказать, что матрица в $M_2(\mathbf{Z}/p^\alpha\mathbf{Z})$ обратима тогда и только тогда, когда результат ее приведения по модулю p обратим в $M_2(\mathbf{Z}/p\mathbf{Z})$. Затем найти число элементов $M_2(\mathbf{Z}/p^\alpha\mathbf{Z})^*$.

19. Используя упражнения 16–18, найти формулу для числа элементов множества $M_2(\mathbf{Z}/N\mathbf{Z})^*$. Обозначим это число через $\varphi_2(N)$. Использовать формулу $\varphi(N) = N\prod_{p|n}(1 - p^{-1})$ для числа $\varphi(N)$ элементов множества $(\mathbf{Z}/N\mathbf{Z})^*$. Записать формулу для $\varphi_2(N)$ в аналогичном виде. Сколько существует шифрующих 2×2-матриц при $N = 26$, 29, 30?

20. Пусть $\varphi_k(N)$ обозначает число обратимых $k \times k$-матриц с элементами из $(\mathbf{Z}/N\mathbf{Z})$. *Угадайте* формулу для $\varphi_k(N)$. Эту формулу нетрудно доказать методом упражнения 16 б).

З а м е ч а н и е. Прием, использованный в упражнениях 16–20, типичен для доказательств и вычислений в кольцах вычетов по модулю N. Сначала, используя свойство мультипликативности, можно свести задачу к случаю, когда модуль является степенью простого числа. Затем при помощи процедуры «подъема» (другой пример этой процедуры см. в упражнении 20 к §II.2) задача приводится к случаю простого модуля. Теперь мы имеем дело с *полем* $\mathbf{F}_p$. Работая в полях, можно в большей степени использовать геометрическую интуицию (как, например, в упражнении 17б) выше). Все положения линейной алгебры, изученные нами в случае вещественных чисел, дословно переносятся на любое поле. Например, сравнение вида $ax + by \equiv c \pmod p$ может быть изображено как «прямая» в «плоскости» над полем $\mathbf{F}_p$. Прямая, соответствующая второму сравнению, либо пересечется с первой прямой в единственной точке, либо будет параллельной первой прямой, либо совпадет с ней. В случае сравнений по составному модулю имеются и другие возможности, возникающие, когда определитель матрицы коэффициентов имеет нетривиальный общий множитель с N.

21. Сколько существует аффинных шифрующих преобразований для биграмм в N-буквенном алфавите? Сколько их при N=26, 29, 30?

22. Предположим, что надо найти дешифрующую матрицу $A^{-1} \in M_2(\mathbf{Z}/N\mathbf{Z})^*$ по уравнению $P = A^{-1}C$, где P и C получены из двух известных пар биграмм открытого и шифрованного текстов. Пусть НОД $(\det C, N) = p$, где p — простое число, только первая степень которого делит N. Положим $n = N/p$.

а) Найти число вариантов для A^{-1}, остающихся при решении сравнения $P \equiv A^{-1}C \pmod n$ с учетом того, что $p \not| \det A^{-1}$.

б) Предположим, что *не* все элементы матрицы C делятся на p. Как можно воспользоваться сравнением $P \equiv A^{-1}C \pmod p$ для уменьшения числа вариантов значений A^{-1}? Сколько вариантов остается после этого? В упражнениях 8 и 15 это применялось в случае $p = 2$.

23. Требуется определить шифрующую 2×2-матрицу A по модулю 30. Имеется две пары биграмм открытого и шифрованного текстов в 30-буквенном алфавите, которые позволяют записать соотношение $AP \equiv C \pmod{30}$, где

$$P = \begin{pmatrix} 2 & 3 \\ 2 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 17 & 8 \\ 8 & 29 \end{pmatrix}.$$

а) Используя приведение по модулю 10, записать матрицу A в виде $A \equiv A_0 + 10A_1 \pmod{30}$, где A_1 — неизвестная матрица по модулю 3 (с 0, 1 и 2 в качестве элементов) и A_0 — матрица, получаемая в результате вычислений по модулю 10. Выбрать A_0 так, чтобы все ее элементы лежали между 0 и 29 и делились на 3.

б) Используя приведение по модулю 3, найти второй столбец матрицы A_1.

в) Сколько есть вариантов для исходной матрицы A? Перечислить их.

24. Пусть

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbf{Z}/N\mathbf{Z})^*$$

— матрица линейного шифрующего преобразования биграмм N-буквенного алфавита. Под *неподвижной биграммой* матрицы A мы понимаем биграмму-вектор P, которому соответствует вектор шифртекста $C = P$, т. е. $AP = P$. В этом упражнении предполагается, что матрица A не является единичной. (Нет ведь никакого смысла рассматривать шифрующее преобразование, которое ничего не скрывает.)

а) Показать, что биграмма «AA»= $\binom{0}{0}$ всегда неподвижна. Какой должна быть матрица

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

для того, чтобы «AA» была ее ***единственной*** неподвижной биграммой?

б) Пусть N — простое число и «AA» — не единственная неподвижная биграмма. Доказать, что имеется ровно N неподвижных биграмм.

25. Перехвачено сообщение

«WUXHURWZNQR XVUEXU!JHALGQGJ?»,

которое зашифровано *аффинным* преобразованием векторов $\binom{x}{y}$ в 841-буквенном алфавите. Здесь числовыми эквивалентами биграмм являются числа $x = 29x_1 + x_2$, где x_1 — эквивалент первой и x_2 — эквивалент второй букв биграммы (29 букв алфавита занумерованы, как в упражнении 9). Таким образом, каждый блок из четырех букв дает столбец $\binom{x}{y}$: две первые буквы дают целое x, а две другие дают y. Известно также, что последние 12 букв приведенного выше шифртекста отвечают подписи «HEADQUARTERS».

а) Найти дешифрующее преобразование и прочитать сообщение.

б) Найти шифрующее преобразование и построить шифрованное сообщение от имени штаба следующего содержания: слова «CANCEL LAST ORDER!» после двух пробелов сопровождаются подписью «HEADQUARTERS».

26. Сколько имеется аффинных шифрующих преобразований в ситуации упражнения 25 (с 841-элементным алфавитом биграмм)?

27. Сколько имеется аффинных шифрующих преобразований для *три*грамм (векторов из 3 элементов) 26-буквенного алфавита?

28. Перехвачено сообщение

«FBRTLWUGAJQINZTHHXTEPHBNXSW»,

зашифрованное линейным шифрующим преобразованием триграмм 26-буквенного алфавита A–Z с числовыми эквивалентами 0–25. Известно, что последние три триграммы — это подпись отправителя «JAMESBOND». Найти дешифрующую матрицу и прочитать сообщение.

ЛИТЕРАТУРА к ГЛАВЕ III

1. *Hill L. S.* Concerning certain linear transformation apparatus of cryptography. — Amer. Math. Monthly, 1931, v. 38, p. 135–154.
2. *Kahn D.* The Codebreakers, the Story of Secret Writing. N.Y. etc.: Macmillan, 1967.
3. *Rosen K. H.* Elementary Number Theory and Its Applications. 3rd ed. Reading: Addison-Wesley, 1993.

ОТКРЫТЫЙ КЛЮЧ

§ 1. Суть криптографии с открытым ключом

Напомним, что криптосистема состоит из взаимно однозначного шифрующего преобразования f из множества $\mathcal{P}$ всех возможных элементов открытого текста в множество $\mathcal{C}$ всех возможных элементов шифртекста. На самом деле термин «криптосистема» чаще применяется к целому семейству таких преобразований, зависящих от выбора некоторых *параметров* (от них могут зависеть как отображение f, так и множества $\mathcal{P}$ и $\mathcal{C}$). Например, при фиксированном N-буквенном алфавите (с раз и навсегда выбранными числовыми эквивалентами) можно рассмотреть аффинную криптосистему (или «семейство криптосистем»), которая при каждом $a \in (\mathbf{Z}/N\mathbf{Z})^*$ и $b \in \mathbf{Z}/N\mathbf{Z}$ является отображением из $\mathcal{P} = \mathbf{Z}/N\mathbf{Z}$ в $\mathcal{C} = \mathbf{Z}/N\mathbf{Z}$, заданным формулой $C \equiv aP + b \pmod N$. В этом примере множества $\mathcal{P}$ и $\mathcal{C}$ фиксированны (так как N — фиксированное число), но шифрующее преобразование f зависит от выбора параметров a и b. Поэтому шифрующее преобразование можно задать: (i) алгоритмом, единым для всего семейства, и (ii) значениями параметров. Значения параметров называются *ключом шифрования* K_E. В нашем примере K_E — это пара (a, b). На практике считается, что алгоритм известен (т.е. общий вид процедуры шифрования сохранить в тайне нельзя). Однако ключи легко меняются и, если это необходимо, держатся в секрете.

Для дешифрования (т.е. вычисления f^{-1}) также необходимы алгоритм и ключ. Этот ключ называется *ключом дешифрования* K_D. В нашем примере аффинной криптосистемы дешифрование производится также аффинным преобразованием $P \equiv a^{-1}C - a^{-1}b \pmod N$, т.е. алгоритм дешифрования совпадает с алгоритмом шифрования, но с другим ключом, а именно, $(a^{-1}, -a^{-1}b)$. (В некоторых криптосистемах алгоритм дешифрования, как и ключ, отличается от алгоритма шифрования.) Мы всегда будем предполагать, что алгоритмы ши-

фрования и дешифрования общеизвестны, а скрыты могут быть лишь ключи K_E и K_D.

Пусть кто-либо пытается использовать для засекречивания связи описанную выше аффинную криптосистему $C \equiv aP + b$. Как мы видели в § III.1, эту систему нетрудно вскрыть, если используются однобуквенные элементы текста в N-буквенном алфавите. Немного сложнее вскрыть эту систему при использовании биграмм, что можно рассматривать как использование N^2-буквенного алфавита. Еще лучше использовать блоки из k букв с числовыми эквивалентами из $\mathbf{Z}/N^k\mathbf{Z}$. При $k > 3$ использовать частотный анализ уже затруднительно, так как число возможных k-грамм очень велико и многие из них можно с равным основанием считать наиболее часто встречающимися. Если идти по пути увеличения k, то надо учитывать затраты времени на решение различных арифметических задач (самая важная из них — нахождение a^{-1} с помощью алгоритма Евклида) как при выборе ключей, так и при шифровании и дешифровании каждого сообщения. В частности, полезно иметь оценки типа O-большое (при увеличении количественных параметров, т.е. когда криптосистема становится «большой») времени, необходимого для: а) шифрования при известном K_E, б) дешифрования при известном K_D, в) вскрытия кода шифрования без знания K_E, г) вскрытия кода дешифрования без знания K_D.

Во всех приведенных выше примерах, как и во всех криптосистемах, использовавшихся более пятнадцати лет назад, не было необходимости задавать ключ дешифрования, если ключ (и алгоритм) шифрования были известны. Даже имея дело с такими большими числами, как N^k при весьма большом k, можно определить ключ дешифрования по ключу шифрования за время, близкое по порядку ко времени работы стандартных алгоритмов. Например, в случае аффинного преобразования в $\mathbf{Z}/N^k\mathbf{Z}$, зная ключ шифрования $K_E = (a, b)$, можно вычислить ключ дешифрования $K_D = (a^{-1} \pmod{N^k}, -a^{-1}b \pmod{N^k})$ с помощью алгоритма Евклида за $O(\log^3 N^k)$ двоичных операций.

Таким образом, в традиционных криптосистемах любой, кто имел возможность дешифровать сообщение, мог с небольшими усилиями или совсем без них определить ключ шифрования. Казалось наивностью или глупостью думать, что человек, вскрывший шифр, может, тем не менее, не знать ключа шифрования. Об этом свидетельствует следующий отрывок из автобиографии хорошо известного исторического персонажа.

«... Пять или шесть недель спустя она (мадам Дюрфе) спросила меня, расшифровал ли я послание, записанное с помощью транс-

мутации. Я ответил утвердительно.

— Извините меня, сэр, но я уверена, что такое без ключа невозможно.

— Мадам! Вы хотите, чтобы я назвал вам ключ?

— Будьте добры.

Тогда я сказал ей ключевое слово, которое не принадлежало ни одному языку, и увидел ее изумление. Она ответила мне, что это невозможно, что она полагала себя единственным обладателем этого слова, которое она держала в своей памяти и никогда не записывала.

Я мог бы сказать ей правду — что те же самые вычисления, при помощи которых я расшифровал послание, позволили мне найти это слово, — но что-то заставило меня сказать, что его открыл мне дух. Это лживое признание приворожило мадам Дюрфе ко мне. В тот день я стал властителем ее души и злоупотребил своей властью. До сих пор мне неприятно и стыдно вспоминать об этом эпизоде, и я рассказал о нем лишь потому, что обязался писать в своих мемуарах одну лишь правду...»

— Казанова, 1757, перевод отрывка, цитированного по книге D. Kahn «Codebreakers».

Ситуация оставалась без изменений почти 220 лет после этого разговора между Казановой и мадам Дюрфе: для любой криптосхемы знание способа шифрования и знание способа дешифрования рассматривались как по сути дела эквивалентные. Однако в 1976 году Диффи и Хеллман открыли принципиально новый тип криптосистем и изобрели «криптографию с открытым ключом».

По определению, криптосистема с открытым ключом обладает тем свойством, что знание шифрующего преобразования не позволяет по ключу шифрования найти ключ дешифрования, избежав чрезвычайно длинных вычислений. Другими словами, шифрующая функция $f\colon \mathcal{P} \to \mathcal{C}$ легко вычисляется, если ключ шифрования K_E известен, но вычислять значения обратной функции $f^{-1}\colon \mathcal{C} \to \mathcal{P}$ очень сложно. С точки зрения практической вычислимости это значит, что функция f необратима (без дополнительной информации — ключа дешифрования K_D). Такую функцию будем называть *функцией с замком*[*]. Таким образом, функция f с замком — это легко вычислимая функция, для которой обратную функцию f^{-1} вычислить трудно, если не иметь некоторой дополнительной информации сверх той, что используется

[*] Для английского термина «trapdoor function» в русскоязычной литературе используется также перевод «функция с лазейкой». — *Прим. ред.*

при вычислении f. Обратная функция f^{-1}, однако, легко вычислима, если известна дополнительная информация K_D — ключ дешифрования.

Очень близким является понятие *однонаправленной функции*. Это такая функция, которую легко вычислить, но для которой обратную функцию f^{-1} вычислить трудно даже при наличии некоторой дополнительной информации. В то время как понятие функции с замком появилось впервые в 1978 году в связи с разработкой криптосистемы RSA с открытым ключом, понятие однонаправленной функции немного старше. По-видимому, впервые использование однонаправленных функций в криптографии было описано в книге Уилкеса о системах разделения времени, опубликованной в 1968 году. Автор описывает новый *однонаправленный шифр*, использованный Нидхамом для того, чтобы компьютер мог проверять пароли, не храня информацию, которая позволила бы злоумышленнику выдавать себя за легального пользователя.

В системе Нидхама при установлении или смене пароля он немедленно шифруется и хранится в компьютере в зашифрованном виде. Когда пароль вводится в ответ на запрос при идентификации пользователя, пароль снова шифруется и результат сравнивается с хранящейся версией. Для злоумышленника нет прямого смысла добывать список хранящихся зашифрованных паролей, так как прежде, чем им пользоваться, эти пароли надо расшифровать. Для этого, даже имея доступ к компьютеру и зная во всех деталях алгоритм шифрования, придется затратить массу времени для процедуры дешифрования.

В 1974 году Дж. Парди впервые подробно описал такую однонаправленную функцию. Исходный пароль и зашифрованный пароль рассматриваются как целые числа по большому простому модулю p, а однонаправленное отображение $\mathbf{F}_p \to \mathbf{F}_p$ задается многочленом $f(x)$, который нетрудно вычислить на компьютере, но обратить его за разумное время невозможно. Парди использовал $p = 2^{64} - 59$, $f(x) = x^{2^{24}+17} + a_1 x^{2^{24}+3} + a_2 x^3 + a_3 x^2 + a_4 x + a_5$, где коэффициентами являются произвольные 19-разрядные целые числа.

Данные выше определения криптосистемы с открытым ключом, однонаправленной функции и функции с замком строгими с математической точки зрения не являются. Они опираются на понятие «практической вычислимости». Но это понятие — эмпирическое, на его содержание влияют как достижения вычислительной техники (например, появление компьютеров с параллельными процессорами), так и открытие новых алгоритмов, ускоряющих решение арифметических

задач. Поэтому шифрующее преобразование, с полным основанием рассматривавшееся в 1994 году как однонаправленная функция или функция с замком, может потерять этот статус в 2004 или 2994 году.

Не исключено, что для некоторых преобразований можно *доказать*, что они являются функциями с замком. Например, может существовать теорема, дающая нетривиальную нижнюю границу для числа двоичных операций, необходимых («в среднем», при случайном выборе ключевых параметров) для того, чтобы описать и реализовать дешифрующий алгоритм, не зная ключ дешифрования. При этом следовало бы допускать возможность использования большого числа соответствующих друг другу элементов открытого и шифрованного текстов (как это делалось при частотном анализе простых систем в главе III), так как, по определению системы с открытым ключом кто угодно может выработать произвольно большое число таких пар. Необходимо было бы также учесть возможность использования «вероятностных» методов, которые, не гарантируя вскрытия кода при однократном применении, рано или поздно дают нужный результат, если применять их много раз. (Примеры вероятностных алгоритмов будут приведены в следующей главе.) К сожалению, ни одной такой теоремы не доказано ни для одной из функций, используемых как шифрующие отображения. Таким образом, хотя сейчас есть много криптосистем, представляющихся заслуживающими названия «систем с открытым ключом», нет ни одной криптосистемы, для которой это свойство было бы *доказано*.

Смысл названия «открытый ключ» состоит в том, что информация, используемая при отправке секретных сообщений — ключ шифрования K_E — может быть раскрыта без риска, что кто-либо получит возможность прочесть секретные сообщения. Так, пусть имеется некоторая группа пользователей криптосистемы, каждый из которых хочет иметь возможность принимать и дешифрировать конфиденциальные сообщения от любого другого без участия третьих лиц (как пользователей, так и посторонних). Какой-нибудь центральный узел может собрать ключи шифрования $K_{E,A}$ от каждого пользователя A и опубликовать их в «телефонной книге» в следующем виде

AAA Banking Company	(9974398087453939, 2975290017591012)
Aardvark, Aaron	(8870004228331, 7234752637937)
...	...

Желающий послать сообщение просто должен найти ключ шифрования получателя в такой «телефонной книге» и использовать его в общем шифрующем алгоритме как ключевой параметр. Только этот получатель имеет в своем распоряжении ключ дешифрования, необхо-

димый для прочтения сообщения.

Раньше никто бы не подумал, что системы этого типа обладают какими-либо особенно поразительными преимуществами. Традиционно криптография использовалась, главным образом, военными и дипломатами. Обычно требовалось распределить систему ключей внутри небольшой строго определенной группы пользователей, и эта система ключей периодически (с помощью курьеров) обновлялась, чтобы лишить противника возможности вскрыть ее.

Однако в последние годы сфера реальных и потенциальных приложений криптографии расширилась и включила в себя ряд других областей, в которых системы связи играют важнейшую роль: создание и хранение баз конфиденциальной информации, электронные финансовые расчеты и т. д. Зачастую имеется большая сеть пользователей, любые два из которых должны иметь возможность сделать переписку между ними секретной как для других пользователей сети, так и для посторонних злоумышленников. Два участника могут в какой-то момент установить секретную связь, позднее один из них может захотеть послать конфиденциальное сообщение третьему. Таким образом, множество участников, между которыми поддерживается секретная связь, может постоянно изменяться. Было бы неудобно каждый раз обмениваться ключами со всеми потенциальными корреспондентами.

Заметим, что система с открытым ключом позволяет двум участникам начать секретный обмен данными без предварительного контакта, без взаимной проверки и без предварительного обмена какой-либо информацией. Вся необходимая для посылки шифрованного сообщения информация является общедоступной.

Классическая версия открытого ключа. Под *классической* криптосистемой (иначе, системой с *секретным ключом* или *симметричной* криптосистемой) понимается криптосистема, в которой, имея информацию о преобразовании шифрования, можно реализовать преобразование дешифрования примерно за такое же время, что и преобразование шифрования. Все криптосистемы главы III были классическими. Бывает, что дешифрование в них требует несколько большего времени, поскольку приходится использовать алгоритм Евклида, чтобы найти обратное число по модулю N, или обращать матрицу (а обращение $k \times k$-матриц при больших k также занимает значительное время); тем не менее, это дополнительное время не слишком велико. (Более того, это дополнительное время используется лишь один раз — при определении K_D — после чего дешифрование становится не сложнее шифрования.) Например, может понадобиться лишь $O(\log^2 B)$ двоичных операций для шифрования одного элемента сооб-

щения и $O(\log^3 B)$ двоичных операций для определения K_D по K_E. Здесь B — граница для размера ключевого параметра. Заметьте, какую роль играют здесь оценки вида O-большое.

С другой стороны, если бы время шифрования было полиномиально по $\log B$, а время дешифрования (при известном K_E, но не K_D) полиномиально по B, а не $\log B$, то это, скорее, была бы система с открытым ключом, а не классическая криптосистема.

Аутентикация.[*)] Часто одной из наиболее важных частей сообщения является *подпись*. Человеческая подпись, написанная машинальным росчерком пера, которую трудно повторить, позволяет быть уверенным, что сообщение получено действительно от лица, чье имя написано ниже. Если сообщение имеет особенную важность, могут потребоваться дополнительные способы *заверения подлинности* (аутентикации) сообщения. А в электронных средствах связи, где нельзя передать физическую подпись, приходится полагаться совершенно на другие методы. Например, когда сотрудник компании хочет снять деньги со счета компании по телефону, его часто просят сообщить некоторую личную информацию (например, девичью фамилию матери), которую знают как сотрудник, так и банк (эти данные подаются при открытии счета), но навряд ли знает посторонний.

Криптография с открытым ключом предоставляет очень простой способ доказать, что вы — это именно вы, которым не может воспользоваться никто другой, желающий выдать себя за вас. Пусть A (Алиса) и B (Боб) — два пользователя системы. Пусть f_A — шифрующее преобразование, которым должен воспользоваться любой, кто отправляет сообщение Алисе, а f_B — соответствующее преобразование для Боба. Для простоты считаем, что множества $\mathcal{P}$ и $\mathcal{C}$ элементов открытого и шифрованного текстов совпадают и одинаковы у всех пользователей. Пусть P — подпись Алисы (включающая в себя, возможно, идентификационный номер Алисы, время отправления сообщения и т. п.). Если Алиса просто пошлет Бобу некоторое сообщение $f_B(P)$, то (поскольку способ вычисления $f_B(P)$ *общеизвестен*) нет способа проверки, гарантирующего, что подпись принадлежит Алисе. Однако, если в начале или конце сообщения будет помещено $f_B f_A^{-1}(P)$, то Боб, применив f_B^{-1}, дешифрует сообщение, включая этот добавок, и все превратится в открытый текст, за исключением добавка, который примет вид $f_A^{-1}(P)$. Так как Боб знает, что сообщение должно

---

*) Английский термин authentication означает удостоверение (или доказательство) подлинности. В русскоязычной литературе часто используется его искаженная калька «аутентификация». — *Прим. ред.*

было быть послано от Алисы, то он применит к добавку f_A (ключ Алисы открыт и ему известен) и получит P. Так как никто, кроме Алисы, не может воспользоваться функцией f_A^{-1}, обратной к f_A, то он удостоверяется в том, что сообщение послано Алисой.

Хеш-функции. Обычно документы подписывают с помощью *хеш-функции*. Образно говоря, хеш-функция — это легко вычислимое отображение $f\colon x \mapsto h$ очень длинного входа x во много более короткий выход h (например, отрезок в 10^6 бит отображается в отрезок длиной 150 или 200 бит), обладающее следующим свойством: *вычислительная сложность нахождения двух различных входов x и x', для которых $f(x) = f(x')$*, чрезвычайно велика. Если «подпись» Алисы содержит значение хеш-функции $h = f(x)$, где x — полный текст ее сообщения, то Боб может не только проверить, что сообщение действительно послано Алисой, но и убедиться, что оно не было изменено при передаче. А именно, Боб применяет хеш-функцию к дешифрованному открытому тексту и проверяет, совпадает ли результат с величиной h из подписи Алисы. По предположению, никто не в состоянии изменить x, не меняя $h = f(x)$.

Обмен ключами. Существующие криптосистемы с открытым ключом работают медленнее, чем современные системы классического типа. Число элементов открытого текста, которые могут быть переданы за одну секунду, у них меньше. Однако, если группа пользователей предпочитает традиционный тип криптосистем, она может воспользоваться криптосистемой с открытым ключом как вспомогательным средством для рассылки друг другу своих ключей $K = (K_E, K_D)$ для классической системы. Таким образом, можно соблюдать основные правила классической криптосистемы и периодически проводить обмен ключами при помощи сравнительно медленной системы с открытым ключом, в то время как основной поток сообщений пересылается с использованием более быстрых прежних методов.

Вероятностное шифрование. Большинство криптосистем для передачи сообщений, основанных на теоретико-числовых идеях, является *детерминированными* в том смысле, что один и тот же открытый текст всегда шифруется в один и тот же шифртекст. Такое шифрование имеет два недостатка: (1) если злоумышленник знает, что имеется мало возможных вариантов для открытого текста (например, «да» или «нет»), то он может зашифровать все эти варианты и определить, который из них составляет сообщение, и (2) очень трудно *доказать* содержательное утверждение о надежности системы детерминированного шифрования. По этим причинам было введено

понятие *вероятностного шифрования*. В этой книге мы не будем обсуждать этот вопрос подробнее или приводить примеры. Ознакомиться с ним можно по основополагающим работам Голдвассера и Микали (Proceedings of the 14th ACM Symposium on the Theory of Computing, 1982, с. 365–377, и J. Comput. System Sci., 1984, т. 28, с. 270–299).

УПРАЖНЕНИЯ

1. Пусть m пользователей хотят обеспечить возможность связи каждого с каждым с помощью классической криптосистемы. При этом каждый настаивает на том, чтобы его переписка с любым другим пользователем была недоступна для остальных $m-2$ пользователей. Сколько для этого потребуется ключей $K=(K_E, K_D)$? Сколько потребуется ключей, если в этой ситуации используется система с открытым ключом? Сколько требуется ключей для криптосистем каждого типа при $m=1000$?

2. Пусть в сети, обслуживающей инвесторов и биржевых брокеров, используется криптосистема с открытым ключом. Инвесторы опасаются, что их брокеры будут покупать акции без согласования с ними (чтобы присвоить комиссионные), а если деньги инвестора будут потеряны, заявят, что они лишь выполняли их инструкции (предъявив в доказательство шифрованное распоряжение о покупке акций, полученное якобы от инвестора). Брокеры, со своей стороны, опасаются, что в случае, если они купят акции согласно инструкции клиента, а акции упадут в цене, инвестор может заявить, что никакой инструкции не посылал или что она послана посторонним или самим брокером. Объяснить, как можно решить эту проблему применением криптографии с открытым ключом, так что когда все они потянут друг друга в суд, можно будет доказать, кто именно виноват в неосторожном вложении денег и последующей их потере. (Предполагается, что в случае тяжбы между инвестором A и брокером B суду будут предъявлены вся относящаяся к делу информация о шифровании и дешифровании — ключи $K_A=(K_{E,A}, K_{D,A})$ и $K_B=(K_{E,B}, K_{D,B})$, а также программное обеспечение, необходимое для шифрования и дешифрования.)

3. Предположим, что две страны A и B пытаются достичь соглашения о запрещении подземных ядерных испытаний. Ни одна из сторон не доверяет другой, имея на это веские причины. Тем не менее, они вынуждены согласиться разместить в различных точках территории обеих стран системы контрольного оборудования. Каждый набор оборудования состоит из сложного сейсмографа, небольшого компьютера для обработки наблюдений сейсмографа и составления сообщений, а также радиопередатчика. Объяснить, как криптография с открытым ключом позволяет удовлетворить всем следующим условиям (которые, на первый взгляд, кажутся несовместимыми).

а) Страна A настаивает на ознакомлении с содержанием (открытым текстом) всех сообщений, отправляемых с ее территории, чтобы быть уверенной в том, что оборудование не используется страной B в целях шпионажа.

б) Страна B настаивает на том, чтобы страна A не могла сфабриковать сообщение, передаваемое оборудованием с его территории (т. е. сообщение типа «все нормально», когда сейсмограф зафиксировал нарушение договора).

в) Страна A настаивает на том, чтобы в случае, если страна B сделает ложное заявление о получении от оборудования сигнала о нарушении договора, любая

заинтересованная третья страна была в состоянии определить, что в реальности сообщения о нарушении не посылалось.

г) Условия а)–в) с переменой сторон.

д) Контрольное оборудование в обеих странах должно быть идентичным и созданным совместно учеными обеих стран.

4. В этой задаче требуется с помощью какой-нибудь функции с замком типа два-в-один организовать игру в монетку на большом расстоянии между игроками. Например, двое, находясь в различных частях света, собираются сыграть партию в шахматы по переписке или по телефону и им надо определить, кто играет белыми. Или, при подготовке к международному хоккейному матчу, представители команд решили бросить монетку для определения того, кто будет считаться хозяином, но не хотят ради этого встречаться или привлекать третью сторону.

Под системой функций с замком типа два-в-один мы понимаем пару алгоритмов: алгоритм, который при заданном ключе K_E подходящего типа строит такую функцию $f\colon \mathcal{P} \to \mathcal{C}$, что каждому элементу c образа f отвечают ровно два таких прообраза $p_1, p_2 \in \mathcal{P}$, что $f(p_1) = f(p_2) = c$, и алгоритм, который при заданном ключе K_D, «обратном к K_E», может найти оба прообраза для любого c из образа f. Здесь предполагается, что вычислительная сложность нахождения K_D по K_E чрезвычайно *велика*. *Заметим*, что *при заданном элементе* $p_1 \in \mathcal{P}$ *можно найти* другой элемент p_2, имеющий тот же образ (т. е. найти оба обратных к $f(p_1)$), если знать как K_E, так и K_D. Однако мы предполагаем, что, зная лишь K_E, практически невозможно найти p_2 ни для какого p_1.

Предположим, что игроки А (Анюта) и В (Бьерн) хотят использовать этот подход для игры в монетку. Игрок А вырабатывает пару ключей K_E и K_D и посылает K_E (но *не* K_D) игроку В. Объяснить процедуру, предоставляющую каждому из игроков равные шансы на «выигрыш» (дать подходящее определение «выигрыша») и надежную защиту от обмана.

ЛИТЕРАТУРА к § 1 ГЛАВЫ IV

1. *Blum M.* Coin-flipping by telephone — a protocol for solving imposible problems. — IEEE Proc., Spring Compcon., p. 133–137.
2. *Diffie W., Hellman M. E.* New directions in cryptography. — IEEE Trans. Inform. Theory IT-22, 1976, p. 644–654.
3. *Chaum D.* Achieving electronic privacy. — Sci. American, 1992, v. 267, p. 96–101.
4. *Goldwasser S.* The search for provably secure cryptosystems. — Crypt. Comput. Number Theory, Proc. Symp. Appl. Math., 1990, v. 42, p. 89–113.
5. *Hellman M. E.* The mathematics of public-key cryptography. — Sci. American, 1979, v. 241, p. 146–157.
6. *Kranakis E.* Primality and Cryptography. N.-Y. etc.: Wiley, 1986.
7. *Rivest R.* Cryptography. — In: Handbook of Theoretical Computer Science. Vol. A. Amsterdam: Elsevier, 1990, p. 717–755.
8. *Ruggiu G.* Cryptology and complexity theories. — In: Advances in Cryptology. Proceedings of Eurocrypt 84. Heidelberg etc.: Springer, 1985, p. 3–9.

§ 2. Криптосистема RSA

При подборе функции с замком f для системы с открытым ключом желательно, чтобы идея шифрования была проста концептуально и несложна в реализации. С другой стороны, хотелось бы иметь убедительное эмпирическое обоснование — основанное на многолетних попытках построения алгоритма для f^{-1} — свидетельствующее о том, что дешифрование реально не осуществимо без знания секретного ключа дешифрования. По этой причине естественно обратить внимание на старинную проблему теории чисел — задачу полной факторизации большого составного целого числа при неизвестных заранее простых множителях. Успех так называемой криптосистемы «RSA» (названной по именам ее создателей: Rivest, Shamir и Adleman), являющейся одной из самых старых (16 лет) и самых популярных криптосистем с открытым ключом, определен чрезвычайной трудностью этой задачи.

Теперь опишем, как работает криптосистема RSA. Сначала каждый пользователь выбирает два очень больших простых числа p и q (имеющих, скажем, по 100 десятичных разрядов каждое) и вычисляет $n = pq$. Зная факторизацию числа n, несложно вычислить $\varphi(n) = (p-1)(q-1) = n + 1 - p - q$. Затем пользователь выбирает случайно целое число e между 1 и $\varphi(n)$, которое взаимно просто с $\varphi(n)$.

Замечание. Когда мы говорим «случайно», то всегда подразумеваем, что число выбрано с помощью датчика случайных (или псевдослучайных) чисел, т. е. компьютерной программы, генерирующей последовательность цифр, которую никто не может повторить или предугадать, и которая, по-видимому, имеет те же статистические свойства, что и истинно случайная последовательность. Много написано об эффективных и надежных способах генерации случайных чисел, но здесь мы не будем касаться этого вопроса. В системе RSA генератор случайных чисел используется не только для выбора e, но и для выбора больших простых чисел p и q (чтобы никто не смог предугадать наш выбор, взяв таблицы простых чисел специального вида, например, простых Мерсенна или делителей чисел вида $b^k \pm 1$ при малых b и сравнительно малых k). Что понимается под «случайно выработанным» простым числом? Это следующее. Сначала вырабатывается большое случайное целое число m. Если m четное, то оно заменяется на $m+1$. Потом, чтобы проверить, является ли это нечетное число простым, используется *тест на простоту числа* (тесты на простоту подробно изучаются в следующей главе). Если число m не является простым, то проверяются $m+2$, потом $m+4$ и т. д., пока не будет получено первое простое число, превосходящее m, которое и берется в качестве «случайного» простого. Поскольку по теореме

о простых числах (формулировку см. в упражнении 13 раздела I.1) доля простых среди целых вблизи m составляет примерно $1/\log m$, можно ожидать, что для нахождения первого простого, большего или равного m, потребуется проверить $O(\log m)$ чисел.

Подобным образом вырабатывается «случайное» число e, взаимно простое с $\varphi(n)$. Сначала вырабатывается случайное (нечетное) целое число подходящего размера, которое последовательно увеличивается, пока не будет найдено e с НОД $(e, \varphi(n)) = 1$. (Можно поступать по-другому: проводить проверку на простоту, пока не найдется простое e, скажем, между $\max(p, q)$ и $\varphi(n)$; такое простое обязательно удовлетворяет условию НОД $(e, \varphi(n)) = 1$.)

Итак, каждый пользователь A выбирает два простых числа p_A, q_A, а вслед за этим — случайное число e_A, которое не имеет общих множителей с $(p_A - 1)(q_A - 1)$. Далее, A вычисляет $n_A = p_A q_A$, $\varphi(n_A) = n_A + 1 - p_A - q_A$ и число, обратное относительно умножения к e_A по модулю $\varphi(n_A)$: $d_A \stackrel{\text{def}}{=} e_A^{-1} \pmod{\varphi(n_A)}$. Ключ шифрования $K_{E,A} = (n_A, e_A)$ делается открытым, а ключ дешифрования $K_{D,A} = (n_A, d_A)$ — секретным. Шифрующее преобразование — это отображение $\mathbf{Z}/n_A\mathbf{Z}$ в себя по формуле $f(P) \equiv P^{e_A} \pmod{n_A}$. Дешифрующее преобразование — это отображение $\mathbf{Z}/n_A\mathbf{Z}$ в себя по формуле $f^{-1}(C) \equiv C^{d_A} \pmod{n_A}$. Нетрудно заметить, что согласно нашему выбору d_A эти два отображения взаимно обратны. А именно, последовательное применение в любом порядке f и f^{-1} приводит к возведению в степень $d_A e_A$. Поскольку $d_A e_A$ дает при делении на $\varphi(n_A)$ остаток 1, это эквивалентно возведению в первую степень (см. следствие из предложения I.3.5, которое гарантирует это, когда P не имеет общих множителей с n_A; случай НОД $(P, n_A) > 1$ рассматривается ниже в упражнении 6).

В приведенном описании множества $\mathcal{P}$ и $\mathcal{C}$ элементов открытого и шифрованного текстов удовлетворяли условию $\mathcal{P} = \mathcal{C}$ и были различными для разных пользователей. На практике может возникнуть необходимость выбирать $\mathcal{P}$ и $\mathcal{C}$ единообразно для всей системы. Пусть, например, используется N-буквенный алфавит и $k < l$ — такие натуральные числа, что N^k и N^l имеют приблизительно по 200 десятичных разрядов каждое. В качестве элементов открытого текста рассмотрим все блоки по k букв, которые интерпретируются как k-разрядные N-ичные числа, т. е. им присваиваются числовые эквиваленты из диапазона между 0 и $N^k - 1$. Аналогично, элементами шифртекста будут блоки из l букв того же алфавита. Каждый пользователь A должен выбрать большие простые числа p_A и q_A так, чтобы $n_A = p_A q_A$ удовлетворяло неравенствам $N^k < n_A < N^l$. Тогда любой

элемент открытого текста, т.е. целое число, меньшее N^k, соответствует вычету из $\mathbf{Z}/n_A\mathbf{Z}$ (для пользователя A) и, так как $n_A < N^l$, то образ $f(P) \in \mathbf{Z}/n_A\mathbf{Z}$ может быть записан в виде l-буквенного блока единственным образом. (При этом возникают не все l-буквенные блоки, а только те, которые отвечают целым, меньшим значения n_A, выбранного данным пользователем.)

Пример 1. Ради читателей, не имеющих под рукой компьютера и программного обеспечения, позволяющего производить вычисления с большим числом знаков, мы отойдем от реальности и будем использовать в большинстве наших примеров сравнительно небольшие целые числа. Выберем $N = 26$, $k = 3$, $l = 4$. Таким образом, открытый текст составлен из триграмм, а шифртекст — из четырехграмм 26-буквенного алфавита. Чтобы послать сообщение «YES» пользователю A с ключом шифрования $(n_A, e_A) = (46927, 39423)$, мы определяем числовой эквивалент для «YES»: $24 \cdot 26^2 + 4 \cdot 26 + 18 = 16346$, и затем вычисляем $16346^{39423} \pmod{46927}$, что есть $21166 = 1 \cdot 26^3 + 5 \cdot 26^2 + 8 \cdot 26 + 2$=«BFIC». Адресат A знает ключ дешифрования $(n_A, d_A) = (46927, 26767)$ и вычисляет $21166^{26767} \pmod{46927} = 16346$ =«YES». Как же пользователь A построил свои ключи? Сначала он перемножил простые числа $p_A = 281$ и $q_A = 167$ и получил n_A, затем случайно (но при условии, что НОД $(e_A, 280)$ = НОД $(e_A, 166) = 1$) было выбрано e_A. В итоге было найдено $d_A = e_A^{-1} \pmod{280 \cdot 166}$. Числа p_A, q_A, d_A остались секретными.

Сколь громоздки вычисления в примере 1? Самым трудоемким по времени является возведение в степень по модулю, т.е. вычисление $16346^{39423} \pmod{46927}$. Но оно может быть выполнено методом повторного возведения в квадрат (см. §I.3) за $O(k^3)$ двоичных операций, где k — число бит в наших целых числах. Если бы использовались очень большие целые числа, то для пользователя A самым трудоемким этапом мог бы оказаться поиск пары больших простых чисел p_A и q_A. Чтобы ускорить поиск очень больших простых чисел, нужно использовать эффективные тесты на простоту. Такие тесты описываются в следующей главе.

Замечания. 1. При выборе p и q пользователь A должен позаботиться о выполнении ряда условий. Самые важные из них следующие: эти простые числа не должны быть слишком близки друг к другу (например, одно должно быть на несколько десятичных разрядов длиннее другого), числа $p-1$ и $q-1$ должны иметь очень маленький наибольший общий делитель и каждое из них должно иметь хотя бы один большой простой делитель. Некоторые причины появления таких условий указаны в последующих упражнениях. Разумеется, если

будет изобретен метод факторизации, который быстро находит ответ при некоторых ограничениях на p и q, то будущим пользователям RSA придется накладывать на выбор p и q дополнительные условия.

2. В § I.3 мы видели, что, когда n является произведением двух простых чисел p и q, знание $\varphi(n)$ эквивалентно знанию разложения. Предположим теперь, что мы решили вскрыть систему RSA посредством нахождения такого натурального d, что $a^{ed} \equiv a \pmod{n}$ для всех a, взаимно простых с n. Это эквивалентно тому, что $ed-1$ кратно наименьшему общему кратному чисел $p-1$ и $q-1$. Знание $m = ed-1$ несет в себе меньше информации, чем точное знание $\varphi(n)$. Однако сейчас мы приведем процедуру, позволяющую по m с большой вероятностью разложить n.

Итак, предположим, что даны целое число n, которое является произведением двух неизвестных простых чисел, и такое целое число m, что $a^m \equiv 1 \pmod{n}$ для всех a, взаимно простых с n. Заметим, что такое m обязательно четно (в чем можно убедиться, взяв $a = -1$). Сначала проверим, обладает ли тем же самым свойством число $m/2$, в случае чего можно заменить m на $m/2$. Если $a^{m/2}$ не сравнимо с 1 по модулю n для некоторых a, взаимно простых с n, то $a^{m/2} \not\equiv 1 \pmod{n}$, по крайней мере, для половины элементов множества $(\mathbf{Z}/n\mathbf{Z})^*$ (это доказывается точно так же, как пункт а) упражнения 21 к § II.2). Таким образом, если мы выберем случайно несколько дюжин значений a и обнаружим, что во всех этих случаях $a^{m/2} \equiv 1 \pmod{n}$, то можно с большой надежностью считать, что это сравнение выполнено для всех a, взаимно простых с n, и можно снова заменить m на $m/2$. Эта процедура повторяется до тех пор, пока сравнение не перестанет выполняться. Тогда имеется две возможности.

1) $m/2$ кратно одному из чисел $p-1$ или $q-1$ (скажем, $p-1$), но не обоим. В этом случае $a^{m/2}$ всегда сравнимо с 1 по модулю p и точно в половине случаев сравнимо по модулю q с -1, а не с 1.

2) $m/2$ не кратно ни одному из чисел $p-1$ и $q-1$. В этом случае $a^{m/2}$ сравнимо с 1 как по модулю p, так и по модулю q (а значит, и по модулю n) ровно в 25% случаев, сравнимо с -1 как по модулю p, так и по модулю q ровно в 25% случаев, а в остальных 50% случаев выбора a оно сравнимо с 1 по одному модулю и с -1 по другому.

Таким образом, случайно выбирая значения a, мы с большой вероятностью быстро найдем такое a, при котором $a^{m/2} - 1$ делится только на одно из наших двух простых чисел (скажем, p). (Случайно выбранное a с вероятностью 50% удовлетворяет этому условию.) Как только такое a найдено, мы сразу можем разложить n, так как НОД $(n, a^{m/2} - 1) = p$.

Приведенная процедура дает пример *вероятностного алгоритма*. Мы еще встретимся с вероятностными алгоритмами в следующей главе.

3. Как послать подпись в системе RSA? При обсуждении проблемы аутентикации в предыдущем разделе мы предполагали для простоты, что $\mathcal{P} = \mathcal{C}$. Для системы RSA ситуация немного сложнее. Существует способ, который позволяет избежать сложностей с неодинаковостью чисел n_A и размеров блоков (число букв k в элементе открытого текста меньше числа букв l в элементе шифртекста). Предположим, как и в предыдущем разделе, что Алиса посылает свою подпись (некоторый открытый текст P) Бобу. Ей известны ключ шифрования Боба $K_{E,B} = (n_B, e_B)$ и собственный ключ дешифрования $K_{D,A} = (n_A, d_A)$. Она должна послать $f_B f_A^{-1}(P)$, если $n_A < n_B$, и $f_A^{-1} f_B(P)$, если $n_A > n_B$. В первом случае она берет наименьший положительный вычет P^{d_A} по модулю n_A, затем, приведя это число по модулю n_B, она вычисляет $(P^{d_A} \pmod{n_A})^{e_B} \pmod{n_B}$, что и посылает в качестве элемента шифртекста. Если $n_A > n_B$, то она сначала вычисляет $P^{e_B} \pmod{n_B}$ и потом возводит это выражение в степень d_A по модулю n_A. Ясно, что Боб, чтобы проверить подлинность сообщения, в первом случае использует возведение в степень d_B по модулю n_B, а потом в степень e_A по модулю n_A. Во втором случае эти операции применяются в обратном порядке.

УПРАЖНЕНИЯ

1. Пусть для открытых и шифрованных текстов во всех случаях используется 40-буквенный алфавит с числовыми эквивалентами 0–25 для A–Z, пробел = 26, . = 27, ? = 28, \$ = 29 и с числовыми эквивалентами 30–39 для цифр 0–9. Пусть элементами открытого текста являются биграммы, а шифртекста — триграммы (т. е. $k = 2$, $l = 3$, $40^2 < n_A < 40^3$ для всех n_A).

а) Послать сообщение «SEND \$7500» пользователю с ключом шифрования $(n_A, e_A) = (2047, 179)$.

б) Вскрыть код, разложив n_A на множители и вычислив затем ключ дешифрования (n_A, d_A).

в) Объяснить, почему аналитик может сравнительно быстро определить ключ дешифрования, не прибегая к факторизации n_A. Другими словами, почему число 2047, помимо своего малого размера, является очень плохим вариантом для n_A?

2. Попробуйте вскрыть код, в котором используется ключ шифрования $(n_A, e_A) = (536813567, 3602561)$. Используйте для разложения на множители числа n_A компьютер и простейший известный алгоритм, состоящий в делении на все нечетные числа по очереди. Если нельзя воспользоваться компьютером, попробуйте угадать простой делитель n_A, пробуя специальные классы простых чисел. После факторизации n_A найдите ключ дешифрования. После этого дешифруйте сообщение «BNBPPKZAVQZLBJ» в предположении, что открытый текст состоит из шестибуквенных блоков в обычном 26-буквенном алфавите (преобразуемых в целые от 0

до 26^6-1 обычным способом), а шифртекст состоит из семибуквенных блоков в том же самом алфавите. Из этого упражнения должно стать ясно, что даже 29-битовое число n_A слишком мало.

3. Пусть элементы открытого и шифрованного текстов — это триграммы, но открытый текст написан в 27-буквенном алфавите (буквы A–Z и пробел $=26$), а шифртекст записан в 28-буквенном алфавите, полученном из 27-буквенного добавлением знака «/» с числовым эквивалентом 27. Предполагается, что каждый пользователь A выбирает n_A в диапазоне между $27^3 = 19683$ и $28^3 = 21952$, так что триграмма открытого текста в 27-буквенном алфавите соответствует вычету P по модулю n_A. Тогда $C \equiv P^{e_A} \pmod{n_A}$ соответствует триграмме шифртекста в 28-буквенном алфавите.

а) Пусть ваш ключ дешифрования — это $K_D = (n, d) = (21583, 20787)$. Дешифровать сообщение «YSNAUOZHXXH » (в конце сообщения стоит один пробел).

б) В условиях а) известно, что $\varphi(n) = 21280$. Найти: 1) $e \equiv d^{-1} \pmod{\varphi(n)}$; 2) разложение числа n.

4. Показать, что 35-битовое целое 23360947609 является очень плохим вариантом для $n = pq$ из-за того, что два его простых делителя слишком близки друг к другу. Иначе говоря, показать, что n легко разлагается «методом Ферма», состоящим в следующем. Заметим, что если $n = pq$ (скажем, $p > q$), то $n = (\frac{p+q}{2})^2 - (\frac{p-q}{2})^2$. Если p и q близки, то $s = (p-q)/2$ мало и $t = (p+q)/2$ есть целое число, лишь немного большее $\sqrt{n}$, причем $t^2 - n$ является полным квадратом. Проверяя подряд целые числа $t > \sqrt{n}$, вы довольно быстро найдете такое, для которого $n = t^2 - s^2$. Тогда $p = t+s$ и $q = t-s$ (см. упражнение 3 к § I.2 и § 3 главы V.)

5. Предположим, что имеется быстрый (вероятностный) алгоритм решения уравнения $x^2 \equiv a \pmod{p}$ при любом простом p и любом квадратичном вычете a. Например, выбирая случайно целые числа и вычисляя символ Лежандра, с большой вероятностью можно найти невычет. Затем можно применить алгоритм, описанный в §II.2. Предположим также, что не существует хорошего алгоритма для решения уравнения $x^2 \equiv a \pmod{n}$, где a является квадратом по модулю n, а $n = pq$ есть произведение двух больших простых чисел, причем разложение числа n на множители неизвестно (при известных множителях можно найти квадратный корень по каждому из модулей и потом воспользоваться китайской теоремой об остатках для нахождения квадратного корня по модулю n). Пусть оба числа p и q не сравнимы с 1 по модулю 4. Положим $K_E = n$ и $K_D = \{p, q\}$. Рассмотрим множество $\mathcal{P} = \mathcal{C} = (\mathbf{Z}/n\mathbf{Z})^* / \pm 1$, которое получается из множества вычетов, взаимно простых с n, «склеиванием» противоположных элементов x и $-x$. Пусть $f\colon \mathcal{P} \to \mathcal{C}$ — это отображение $x \mapsto x^2 \pmod{n}$. Показать, что это отображение является примером отображения упражнения 4 из предыдущего раздела. Оно дает способ реализации игры в монетку на большом расстоянии.

6. Пусть n — целое число, свободное от квадратов (т. е. произведение различных простых). Пусть d и e — такие положительные целые числа, что $de - 1$ делится на $p-1$ для каждого простого делителя p числа n. (Например, так будет, если $de \equiv 1 \pmod{\varphi(n)}$.) Доказать, что $a^{de} \equiv a \pmod{n}$ для *любого* целого a (вне зависимости от того, имеет ли оно общие делители с n).

7. Доказать утверждения замечания 2 о вероятностях выполнения различных сравнений в пп. 1) и 2).

ЛИТЕРАТУРА к § 2 ГЛАВЫ IV

1. *Adleman L. M., Rivest R. L., Shamir A.* A method for obtaining digital signatures and public-key cryptosystems. — Comm. ACM, 1978, v. 21, p. 120–126.
2. *Rivest R. L.* RSA chips (past/present/future). — In: Advances in Cryptology. Proceedings of Eurocrypt 84. Heidelberg etc.: Springer, 1985, p. 159–165.
3. *Gordon J. A.* Strong primes are easy to find. — In: Advances in Cryptology. Proceedings of Eurocrypt 84. Heidelberg etc.: Springer, 1985, p. 216–223.

§ 3. Дискретное логарифмирование

Обсуждавшаяся в предыдущем разделе система RSA основана на том, что найти и перемножить два больших простых числа значительно проще, чем выполнить обратную операцию (имея n, найти эти два простых числа). Имеются и другие фундаментальные теоретико-числовые процедуры, обладающие, по всей видимости, свойствами однонаправленности или функции с замком. Одной из самых важных таких процедур является возведение в степень в большом конечном поле.

В поле вещественных чисел возведение в степень (нахождение b^x с заданной точностью) не намного проще обратной операции (нахождения $\log_b x$ с заданной точностью). Рассмотрим теперь конечную группу, например, $(\mathbf{Z}/n\mathbf{Z})^*$ или $\mathbf{F}_q^*$ (с умножением в качестве групповой операции). Метод повторного возведения в квадрат (см. § I. 3), позволяет вычислить b^x при больших x сравнительно быстро (за полиномиальное по $\log x$ время). Пусть теперь задан элемент y, допускающий представление b^x (считается, что «основание» b задано); как найти степень элемента b, в которой он равен y, т. е. вычислить $x = \log_b y$ (здесь «log» имеет отличный от обычного, хотя и близкий, смысл)? Этот вопрос называется «задачей дискретного логарифмирования». Слово «дискретное» подчеркивает отличие случая конечной группы постановки от классического непрерывного случая.

О п р е д е л е н и е. Пусть G — конечная группа, b — элемент группы G и y — элемент группы G, являющийся степенью b. Любое целое число x, для которого $b^x = y$, называется *дискретным логарифмом* y по основанию b.

П р и м е р 1. Возьмем $G = \mathbf{F}_{19}^* = (\mathbf{Z}/19\mathbf{Z})^*$ и порождающий элемент 2 в качестве b (см. пример 1 в § II.1). Тогда дискретный логарифм числа 7 по основанию 2 равен 6.

П р и м е р 2. Если $\alpha \in \mathbf{F}_9^*$ есть корень уравнения $X^2 - X - 1$ (см. пример 2 в § II. 1), то дискретный логарифм числа -1 по основанию α равен 4.

В конце раздела будет приведен краткий обзор известных алгоритмов решения задачи дискретного логарифмирования в конечных полях. Сначала же опишем несколько криптосистем с открытым ключом и систем специального назначения с открытым ключом, основанных на вычислительной сложности задачи дискретного логарифмирования в конечных полях.

Система Диффи–Хеллмана обмена ключами. Так как криптосистемы с открытым ключом значительно медленнее классических криптосистем (по крайней мере, при нынешнем состоянии науки и техники), то разумнее использовать их в качестве дополнения к классическим криптосистемам, с помощью которых и передаются сообщения. Зато процедуру согласования ключей классической криптосистемы можно очень эффективно реализовать с помощью системы с открытым ключом. Первая такая детально проработанная схема, предложенная Диффи и Хеллманом, основана на задаче дискретного логарифмирования.

Пусть ключом классической криптосистемы является большое случайно выбранное натуральное число (или набор таких чисел). Пусть, например, используется матричное аффинное преобразование пар биграмм (см. § III. 2)

$$C \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} P + \begin{pmatrix} e \\ f \end{pmatrix} \pmod{N^2}$$

где $0 \leqslant a, b, c, d, e, f < N^2$ и P — вектор-столбец, составленный из числовых эквивалентов двух соседних биграмм (составляющих вместе четырехбуквенный блок) открытого текста в N-буквенном алфавите. Выбрав случайно целое число k между 0 и N^{12}, мы сразу определим a, b, c, d, e, f как значения шести разрядов числа k в N^2-ичной системе счисления. (При этом надо проверить, что $ad - bc$ обратимо по модулю N^2, т. е. что это число не имеет общих множителей с N; в противном случае выбирается новое k.)

Заметим, что выбор случайного целого в некотором интервале эквивалентен выбору случайного элемента большого конечного поля приблизительно такого же размера. Пусть, например, мы хотим выбрать случайное положительное $k < N^{12}$ и наше конечное поле простое и состоит из p элементов. Элементам поля $\mathbf{F}_p$ сопоставим, как обычно, целые от 0 до $p - 1$. Если полученное таким образом целое не меньше N^{12}, то приводим его по модулю N^{12}.

Если же наше поле имеет вид $\mathbf{F}_{p^f}$, то сначала выбирается $\mathbf{F}_p$-базис этого поля. Таким образом, каждому элементу поля сопоставляется набор из f элементов поля $\mathbf{F}_p$. Затем, рассматривая составляющие

набора как разряды целого числа, записанного в p-ичной системе счисления, получаем целое число, меньшее p^f. **Предупреждение.** Эта процедура устанавливает взаимно однозначное соответствие между $\mathbf{F}_{p^f}$ и $\mathbf{Z}/p^f\mathbf{Z} = \{0, 1, \ldots, p^f - 1\}$. Эти два множества имеют совершенно различную структуру относительно операций сложения и умножения. Первое является *полем*, т.е. все $p^f - 1$ ненулевых элементов имеют обратные, а второе — *кольцом*, в котором это свойство нарушается для p^{f-1} из p^f элементов (для элементов, кратных p).

Теперь опишем метод построения случайного элемента большого конечного поля $\mathbf{F}_q$, разработанный Диффи и Хеллманом. Пусть значение q общеизвестно, т.е. все знают, из какого конечного поля выбирается ключ. Пусть, далее, g — некоторый фиксированный элемент $\mathbf{F}_q$, который также не является секретным. В идеале, g должен быть порождающим элементом группы $\mathbf{F}_q^*$. Однако в нашем случае в этом нет необходимости, так как описываемый ниже метод построения ключа дает только те элементы $\mathbf{F}_q$, которые являются степенями элемента g. Если же мы хотим, чтобы случайный элемент мог принимать любые значения из $\mathbf{F}_q^*$, то g должен быть порождающим элементом.

Пусть два пользователя (Аида и Бернардо) хотят согласовать ключ — случайный элемент из $\mathbf{F}_q^*$, — посредством которого они будут зашифровывать переписку между собой. Аида выбирает случайное число a между 1 и $q - 1$, которое она держит в секрете, и вычисляет $g^a \in \mathbf{F}_q$, которое объявляет открыто. Бернардо делает то же самое: он случайно выбирает b и объявляет g^b. В качестве секретного ключа используется g^{ab}. Оба пользователя могут вычислить этот ключ. Например, Аида знает g^b (это открытая информация) и свой собственный секретный ключ a. Однако посторонние знают только g^a и g^b. Если для мультипликативной группы $\mathbf{F}_q^*$ выполнено следующее предположение, то посторонние не смогут определить ключ.

Предположение Диффи–Хеллмана. *Сложность вычисления g^{ab} по g^a и g^b чрезвычайно велика.*

Предположение Диффи–Хеллмана априори не слабее предположения о чрезвычайной сложности дискретного логарифмирования в конечной группе. Если бы можно было вычислять дискретные логарифмы, то, очевидно, предположение Диффи–Хеллмана было бы неверным. Некоторые считают, что справедливо и обратное, однако пока этот вопрос остается открытым. Другими словами, никто еще не предложил способ получения g^{ab} из g^a и g^b без использования a и b. Однако вполне возможно, что такой способ существует.

Пример 3. Пусть при шифровании используется преобразование сдвига однобуквенных элементов в 26-буквенном алфавите (см.

пример 1 в § III. 1): $C \equiv P + B \pmod{26}$. (Мы используем для параметра сдвига обозначение B, чтобы отличить его от символа b в предыдущем абзаце.) В качестве B возьмем наименьший неотрицательный вычет по модулю 26 случайного элемента из $\mathbf{F}_{53}$. Пусть $g = 2$ (это порождающий элемент в $\mathbf{F}_{53}$). Предположим, что Аида выбрала случайно $a = 29$ и ей известен открытый ключ Бернардо 2^b, скажем, $12 \in \mathbf{F}_{53}$. Тогда она находит, что ключом шифрования является $12^{29} = 21 \in \mathbf{F}_{53}$, т.е. $B = 21$. В то же время она публикует свой открытый ключ $2^{29} = 45$. Поэтому Бернардо тоже может найти ключ $B = 21$ возведением 45 в степень b (его секретный показатель $b = 19$). Конечно, система с таким маленьким полем не дает надежной защиты: посторонний может легко найти дискретный логарифм по основанию 2 для 12 или 45 по модулю 53. И уж, во всяком случае, не является надежным шифрование сдвигом для однобуквенных элементов сообщения. Этот пример лишь иллюстрирует механизм данной системы обмена ключами.

Криптосистема Мэсси–Омуры для передачи сообщений. Предположим, что все согласились использовать конечное поле $\mathbf{F}_q$; оно зафиксировано и общеизвестно. Каждый пользователь системы втайне выбирает такое случайное целое e между 0 и $q-1$, что НОД $(e, q-1) = 1$, и с помощью алгоритма Евклида вычисляет обратное к нему число $d \equiv e^{-1} \pmod{q-1}$, т.е. $de \equiv 1 \pmod{q-1}$. Если Алиса (пользователь А) намерена передать сообщение P Бобу, она сначала посылает ему элемент P^{e_A}. Это послание для Боба бессодержательно, так как он не знает d_A (или e_A, что то же самое) и не может восстановить P. Не пытаясь понять смысл сообщения, Боб возводит его в *свою степень* e_B и отсылает $P^{e_A e_B}$ обратно Алисе. Третий шаг состоит в том, что Алиса несколько «распутывает» сообщение, возводя его в степень d_A; так как $P^{d_A e_A} = P$ (по предложению II.1.1), то, по сути дела, она отсылает Бобу P^{e_B}, который теперь может прочитать сообщение, возведя его в степень d_B.

Весьма простая идея этой системы может быть реализована и в схемах, где используются процедуры, отличные от возведения в степень в конечных полях. Однако не все так просто. Прежде всего, отметим, что при использовании системы Мэсси–Омуры абсолютно необходима хорошая схема подписи сообщений. Без нее любое постороннее лицо C, которому сообщение P не предназначено, может представиться Бобом и вернуть Алисе сообщение $P^{e_A e_C}$. Алиса, не зная, что оно прислано самозванцем, пользовавшимся своим ключом e_C, продолжит процедуру, возведя сообщение в степень d_A, дав тем самым C возможность прочитать сообщение. Поэтому промежуточное сообще-

ние $P^{e_A e_B}$ от Боба Алисе должно сопровождаться аутентикацией, т. е. некоторым сообщением в схеме подписи, которое может послать только Боб.

Кроме того, важно, чтобы такие пользователи, как B или C, которые после дешифрования различных сообщений P будут знать пары (P, P^{e_A}), не могли воспользоваться этим для определения e_A. Так, если бы Боб мог решать задачу дискретного логарифмирования в $\mathbf{F}_q^*$ и по P и P^{e_A} определять, каким должно быть e_A, то он мог бы легко вычислить $d_A \equiv e_A^{-1} \pmod{q-1}$ и затем перехватывать и читать все дальнейшие сообщения Алисы, кому бы они ни были адресованы.

Криптосистема Эль-Гамаля. Сначала зафиксируем очень большое конечное поле $\mathbf{F}_q$ и элемент $g \in \mathbf{F}_q^*$ (желательно, хотя и не обязательно, чтобы он был порождающим). Предположим, что используются элементы открытого текста с численными эквивалентами P в $\mathbf{F}_q$. Каждый пользователь A выбирает случайно целое число $a = a_A$, скажем, из диапазона $0 < a < q-1$. Это секретный ключ дешифрования. Открытым ключом шифрования является элемент $g^a \in \mathbf{F}_q$.

Чтобы передать сообщение P пользователю A, мы выбираем случайно целое число k и посылаем A следующую пару элементов из $\mathbf{F}_q$:

$$(g^k, Pg^{ak}).$$

Заметим, что вычислить g^{ak} можно, не зная a, просто возведя g^a в степень k. Теперь A, зная a, может по этой паре раскрыть P, возведя первый элемент g^k в a-ю степень и разделив на результат второй элемент (или, что эквивалентно, возведя g^k в степень $q-1-a$ и умножив на второй элемент). Другими словами, наше послание состоит из замаскированного сообщения (P «несет маску» g^{ak}) и «ключа», а именно, g^k, которым можно снять маску (но воспользоваться ключом может лишь тот, кто знает a).

Тот, кто умеет решать задачу дискретного логарифмирования в $\mathbf{F}_q$, вскроет эту криптосистему, определив ключ дешифрования a по ключу шифрования g^a. Вообще говоря, может существовать способ определения g^{ak} по g^k и g^a, а значит, и вскрытия шифра, не связанный с дискретным логарифмированием. Однако, как уже упоминалось при обсуждении системы Диффи–Хеллмана, считается, что нет способа получить g^{ak} из g^k и g^a, не решив, по существу, задачи дискретного логарифмирования.

Стандарт цифровой подписи. В 1991 году Национальный институт стандартов и технологий (НИСТ) при правительстве США

предложил стандарт цифровой подписи (Digital Signature Standard, сокращенно, DSS). Как ожидается, роль DSS будет аналогична роли принятого много раньше стандарта шифрования данных (DES), т.е. он будет стандартом цифровой подписи для использования как государственными, так и коммерческими организациями. В отличие от DES, который является классической криптосистемой с секретным ключом, при построении цифровой подписи необходимо использовать криптосистему с открытым ключом. НИСТ заложил в основу своей схемы подписи задачу дискретного логарифмирования в простом конечном поле. DSS очень похож на схему подписи, предложенную первоначально Шнорром (см. ссылки ниже). Он похож также на схему подписи Эль-Гамаля (см. упражнение 9 ниже). Теперь расскажем, как работает DSS.

Чтобы получить возможность подписывать сообщения, каждым пользователем (Алисой) осуществляется следующая последовательность действий: 1) выбирается простое число q приблизительно в 160 бит (для этого используются генератор случайных чисел и тест на простоту); 2) выбирается другое простое число p, сравнимое с 1 по модулю q размером приблизительно в 512 бит; 3) выбирается порождающий элемент g единственной циклической подгруппы в $\mathbf{F}_p^*$ порядка q (т.е. вычисляется $g \equiv g_0^{(p-1)/q} \pmod{p}$ для случайного целого g_0; если $g \neq 1$, то он и будет порождающим элементом); 4) выбирается секретный ключ как случайное целое число x из диапазона $0 < x < q$. В качестве открытого ключа берется $y = g^x \pmod{p}$.

Пусть Алиса хочет подписать сообщение. Сначала она применяет к своему открытому тексту хеш-функцию (см. § 1) и получает целое число h из диапазона $0 < h < q$. Затем она выбирает случайное целое k из того же диапазона и вычисляет $g^k \pmod{p}$. Пусть r — наименьший неотрицательный вычет по модулю q последнего выражения (значит, g^k сначала вычисляется по модулю p, а результат приводится по меньшему простому модулю q). Наконец, Алиса находит такое целое s, что $sk \equiv h + xr \pmod{q}$. Пара (r, s) вычетов по модулю q является ее подписью.

Чтобы проверить подпись, получатель Боб вычисляет $u_1 \equiv s^{-1}h \pmod{q}$ и $u_2 \equiv s^{-1}r \pmod{q}$. Потом он вычисляет $g^{u_1}y^{u_2} \pmod{p}$. Если результат сравним с r по модулю q, то подпись считается подлинной. (Заметим, что $g^{u_1}y^{u_2} \equiv g^{s^{-1}(h+xr)} \equiv g^k \pmod{p}$.)

Достоинством этой схемы является сравнительно малый размер подписи, состоящей из двух чисел по 160 бит (размер q). С другой стороны, надежность системы зависит, очевидно, от сложности решения задачи дискретного логарифмирования в большом конечном поле $\mathbf{F}_p$.

Хотя для вскрытия рассмотренной системы достаточно решения этой задачи в меньшей подгруппе, порожденной g, представляется маловероятным, что эта задача намного проще задачи дискретного логарифмирования для всей группы $\mathbf{F}_p^*$. Таким образом, DSS, по-видимому, имеет очень высокий уровень надежности при небольшом размере подписи и малом времени на вычисления.

Алгоритмы дискретного логарифмирования в конечных полях. Сначала мы рассмотрим случай, когда все простые делители числа $q-1$ малы. Такие числа $q-1$ будем называть «гладкими». Для них существует быстрый алгоритм вычисления дискретного логарифма по основанию b элемента из $\mathbf{F}_q^*$. Для простоты будем считать, что b является порождающим элементом группы $\mathbf{F}_q^*$. Теперь опишем этот алгоритм, предложенный Силвером, Полигом и Хеллманом.

Сначала для каждого простого p, делящего $q-1$, вычисляются корни p-й степени из единицы: $r_{p,j} = b^{j(q-1)/p}$ при $j = 0, 1, \ldots, p-1$. (Как обычно, при возведении b в большую степень используется метод последовательного возведения в квадрат.) Эта таблица значений $\{r_{p,j}\}$ является основой для вычисления дискретного логарифма любого $y \in \mathbf{F}_q^*$. (Заметим, что при фиксированном b эти вычисления производятся лишь один раз, и одна и та же таблица используется при любом y.)

Наша цель — найти такой элемент x, $0 \leqslant x < q-1$, что $b^x = y$. Пусть $q-1 = \prod_p p^\alpha$ есть разложение числа $q-1$ на простые множители. Достаточно найти $x \pmod{p^\alpha}$ для каждого p, делящего $q-1$; после этого x однозначно определяется применением алгоритма из доказательства китайской теоремы об остатках (предложение I.3.3). Итак, мы зафиксируем простое p, делящее $q-1$, и покажем, как найти $x \pmod{p^\alpha}$.

Предположим, что $x \equiv x_0 + x_1 p + \cdots + x_{\alpha-1} p^{\alpha-1} \pmod{p^\alpha}$ с $0 \leqslant x_i < p$. Для того чтобы найти x_0, мы вычисляем $y^{(q-1)/p}$. Это корень p-й степени из единицы, так как $y^{q-1} = 1$. Из равенства $y = b^x$ следует, что $y^{(q-1)/p} = b^{x(q-1)/p} = b^{x_0(q-1)/p} = r_{p,x_0}$. Сравниваем $y^{(q-1)/p}$ с $\{r_{p,j}\}_{0 \leqslant j < p}$ и полагаем x_0 равным тому значению j, при котором $y^{(q-1)/p} = r_{p,j}$.

Далее, чтобы найти x_1, мы заменяем y на $y_1 = y/b^{x_0}$. Тогда y_1 имеет дискретный логарифм $x - x_0 \equiv x_1 p + \cdots + x_{\alpha-1} p^{\alpha-1} \pmod{p^\alpha}$. Так как y_1 — это p-я степень, то получаем $y_1^{(q-1)/p} = 1$ и $y_1^{(q-1)/p^2} = b^{(x-x_0)(q-1)/p^2} = b^{(x_1 + x_2 p + \cdots)(q-1)/p} = b^{x_1(q-1)/p} = r_{p,x_1}$. Значит, мы можем сравнить $y_1^{(q-1)/p^2}$ с $\{r_{p,j}\}$ и положить x_1 равным тому j, при котором $y_1^{(q-1)/p^2} = r_{p,j}$.

Теперь должно быть понятно, как можно продолжить этот процесс, чтобы получить все $x_0, x_1, \ldots, x_{\alpha-1}$. А именно, для каждого $i = 1, 2, \ldots, \alpha - 1$ полагаем

$$y_i = y/b^{x_0 + x_1 p + \cdots + x_{i-1} p^{i-1}};$$

дискретный логарифм этого выражения сравним по модулю p^α с $x_i p^i + \cdots + x_{\alpha-1} p^{\alpha-1}$. Так как y_i является p^i-й степенью, то $y_i^{(q-1)/p^i} = 1$ и $y_i^{(q-1)/p^{i+1}} = b^{(x_i + x_{i+1} p + \cdots)(q-1)/p} = b^{x_i (q-1)/p} = r_{p,x_i}$. Поэтому мы полагаем x_i равным тому j, при котором $y_i^{(q-1)/p^{i+1}} = r_{p,j}$.

Завершив этот процесс, мы получим $x \pmod{p^\alpha}$. Повторив такие вычисления для каждого $p | (q-1)$, мы воспользуемся китайской теоремой об остатках и найдем x.

Этот алгоритм хорошо работает, когда все простые делители числа $q-1$ малы. Очевидно, однако, что для составления таблицы $\{r_{p,j}\}$ и сравнения величин $y_i^{(q-1)/p^{i+1}}$ с ее элементами потребуется много времени, если $q-1$ имеет большой простой делитель. (Под «большим» понимается число хотя бы в 20 разрядов. Если $p | (q-1)$ по порядку меньше 10^{20}, то можно использовать комбинацию алгоритма Силвера–Полига–Хеллмана и алгоритма, использующего метод «больших и малых шагов» Шэнкса; см. с. 9 и с. 575–576 второго тома книги Кнута.)

Пример 4. Найти дискретный логарифм 28 по основанию 2 в $\mathbf{F}_{37}^*$, используя алгоритм Силвера–Полига–Хеллмана (2 является порождающим элементом в $\mathbf{F}_{37}^*$).

Решение. Имеем $37 - 1 = 2^2 \cdot 3^2$. Вычисляем $2^{18} \equiv 1 \pmod{37}$ и $r_{2,0} = 1$, $r_{2,1} = -1$. (При $p = 2$ всегда $\{r_{2,j}\} = \{\pm 1\}$.) Далее, $2^{36/3} \equiv 26 \pmod{37}$, $2^{2 \cdot 36/3} \equiv 10 \pmod{37}$ и $\{r_{3,j}\} = \{1, 26, 10\}$. Пусть теперь $28 \equiv 2^x \pmod{37}$. Сначала возьмем $p = 2$ и найдем вычет $x \pmod 4$, который запишем как $x_0 + 2x_1$. Имеем $28^{36/2} \equiv 1 \pmod{37}$, и, следовательно, $x_0 = 0$. Далее, $28^{36/4} \equiv -1 \pmod{37}$, и поэтому $x_1 = 1$, т. е. $x \equiv 2 \pmod 4$. Теперь берем $p = 3$ и находим вычет $x \pmod 9$, который записываем как $x_0 + 3x_1$. (Разумеется, для каждого p значения x_i свои.) Чтобы найти x_0, вычисляем $28^{36/3} \equiv 26 \pmod{37}$. Таким образом, $x_0 = 1$. Затем вычисляем $(28/2)^{36/9} = 14^4 \equiv 10 \pmod{37}$ и получаем $x_1 = 2$. Итак, $x \equiv 1 + 2 \cdot 3 = 7 \pmod 9$. Остается найти единственный элемент $x \pmod{36}$, при котором $x \equiv 2 \pmod 4$ и $x \equiv 7 \pmod 9$. Это $x = 34$. Таким образом, $28 = 2^{34}$ в $\mathbf{F}_{37}^*$.

Индексный алгоритм дискретного логарифмирования. Читатель может пропустить этот раздел или прочитать его бегло и вер-

нуться к нему потом, при работе с § V.3, поскольку индексный алгоритм вычисления дискретных логарифмов в конечных полях имеет много общего с методом факторных баз для разложения больших целых чисел.

Здесь мы будем предполагать, что $q = p^n$ есть весьма большая степень малого простого p и b — порождающий элемент группы $\mathbf{F}_q^*$. Индексный алгоритм для любого $y \in \mathbf{F}_q^*$ находит такой вычет $x \pmod{q-1}$, что $y = b^x$.

Пусть $f(X) \in \mathbf{F}_p[X]$ — неприводимый многочлен степени n. Тогда $\mathbf{F}_q$ изоморфно кольцу вычетов $\mathbf{F}_p[X]/f(X)$. Любой элемент $\mathbf{F}_p[X]/f(X)$ может быть записан (единственным образом) как многочлен $a(X) \in \mathbf{F}_p[X]$ степени не выше $n-1$. В частности, основание $b = b(X)$ — тоже многочлен: «константы» являются элементами в $\mathbf{F}_p \subset \mathbf{F}_q$.

Заметим сначала, что $b' = b^{(q-1)/(p-1)}$ является порождающим элементом группы $\mathbf{F}_p^*$ (см. упражнение 17 к § II.1). Значит, решение задачи дискретного логарифмирования в $\mathbf{F}_p^*$ (по основанию b') дает значения дискретных логарифмов этих констант по основанию b. Так как по условию p мало, то таблицу таких дискретных логарифмов легко составить. В важном частном случае $p = 2$ единственной ненулевой константой является 1, и ее дискретный логарифм по любому основанию равен 0. Далее будем предполагать, что можно легко найти дискретный логарифм любой константы.

До конца этого раздела будем через $\operatorname{ind} a(X)$ (от слова «index») обозначать дискретный логарифм $a(X) \in \mathbf{F}_q^*$ по основанию $b(X)$. Основание $b(X)$ зафиксировано, поэтому оно не вошло в обозначение.

Индексный алгоритм состоит из двух основных этапов. Первый этап мы называем «предварительным», поскольку он не зависит от элемента $y(X) \in \mathbf{F}_q^*$, логарифм которого мы хотим вычислить. Этот этап выполняется один раз и его результаты можно многократно использовать при вычислении дискретных логарифмов по фиксированному основанию $b(X)$. (Напомним, что аналогичный предварительный этап имеется и в алгоритме Силвера–Полига–Хеллмана — это составление таблицы $\{r_{p,j}\}$.)

Сначала выберем подмножество $B \subset \mathbf{F}_q$, которое будет служить нам «базисом». Обычно B состоит из всех нормированных неприводимых многочленов над $\mathbf{F}_p$ степени не выше m, где $m < n$ определяется некоторым оптимальным образом так, чтобы множество B имело подходящий размер $h = \#(B)$ промежуточного порядка между $p = \#(\mathbf{F}_p)$ и $q = p^n = \#(\mathbf{F}_q)$. Предварительный этап состоит в нахождении дискретных логарифмов для всех $a(X) \in B$ и заключается в следующем.

Выберем случайное целое t между 1 и $q-2$ и вычислим $b^t \in \mathbf{F}_q$, т. е. вычислим такой многочлен $c(X) \in \mathbf{F}_p[X]$ степени ниже n, что

$$c(X) \equiv b(X)^t \pmod{f(X)}.$$

(При этом используется метод повторного возведения в квадрат с приведением после каждого шага по модулю $f(X)$.) Вынеся за скобки старший коэффициент c_0 в $c(x)$, проверяем, можно ли полученный нормированный многочлен представить в виде произведения многочленов $a(X) \in B$, т. е. может ли $c(X)$ быть записан в виде

$$c(X) = c_0 \prod_{a \in B} a(X)^{\alpha_{c,a}}.$$

Один из способов такой проверки — просмотреть все $a(X) \in B$, деля $c(X)$ последовательно на $a(X)^{\alpha_{c,a}}$ (где $\alpha_{c,a}$ — наибольшая степень $a(X)$, на которую делится $c(X)$ в $\mathbf{F}_p[X]$). Если после всех этих делений от многочлена остается лишь константа c_0, то $c(X)$ имеет указанную выше форму. В противном случае все повторяется для нового случайно выбранного целого t. (Другой — иногда более быстрый — способ проверки того, разлагается ли $c(X)$ в произведение $a(X) \in B$, состоит в нахождении всех делителей $c(X)$ с помощью какого-нибудь алгоритма разложения элементов $\mathbf{F}_p[X]$. Например, удобный алгоритм Берлекампа описан в § 4.6.2 второго тома книги Кнута.)

Теперь допустим, что найден элемент $c(X) \equiv b(X)^t \pmod{f(X)}$, имеющий разложение искомого типа. Взяв дискретный логарифм от обеих частей этого равенства, получаем

$$\operatorname{ind} c(X) - \operatorname{ind} c_0 = \sum_{a \in B} \alpha_{c,a} \operatorname{ind} a(X),$$

где равенство должно пониматься как сравнение по модулю $q-1$ (так как дискретный логарифм определен только по модулю $q-1$). Левая часть этого равенства известна: $\operatorname{ind} c(X) = t$, а дискретные логарифмы констант мы считаем известными. Коэффициенты $\alpha_{c,a}$ в правой части также известны. Неизвестными являются h значений $\operatorname{ind} a(X)$, $a(X) \in B$, в правой части.

Таким образом, мы получили линейное уравнение с h неизвестными в $\mathbf{Z}/(q-1)\mathbf{Z}$. Представим себе, что мы продолжаем выбирать случайные целые t, пока не получим много разных $c(X)$, разложимых в произведение многочленов $a(X)$. Как только мы получим h независимых сравнений вида

$$t - \operatorname{ind} c_0 \equiv \sum_{a \in B} \alpha_{c,a} \operatorname{ind} a(X) \pmod{q-1}$$

(здесь «независимость» означает, что определитель матрицы $\{\alpha_{c,a}\}$ взаимно прост с $q-1$), мы сможем решить эту систему относительно неизвестных по модулю $q-1$. (Элементы линейной алгебры по $\bmod N = q-1$ изложены в § III.2.) На этом предварительная стадия индексного алгоритма завершается. Она создала большую «базу данных», а именно, значений дискретных логарифмов для всех $a(X) \in B$, с помощью которых можно вычислить дискретный логарифм любого элемента из $\mathbf{F}_q^*$.

Прежде чем приступить к описанию второй фазы индексного алгоритма, обсудим способ выбора m, который не был конкретизирован при определении $B \subset \mathbf{F}_p[X]$ как множества всех нормированных неприводимых многочленов степени не выше m. Размер h множества B быстро растет с ростом m. Например, при простом m (см. следствие из предложения II.1.8) существует $(p^m - p)/m$ нормированных неприводимых многочленов только степени m. Так как необходимо найти не менее h различных $c(X)$, которые дадут систему из h линейных сравнений относительно h неизвестных $\operatorname{ind} a(X)$, и эту систему надо будет решать, то желательно, чтобы h (и, значит, m) были не слишком большими. С другой стороны, если m мало, то «типичный» нормированный многочлен $c_0^{-1}c(X)$ степени не выше $n-1$, скорее всего, не разложится в произведение многочленов $a(X)$ степени не выше m и будет иметь, по крайней мере, один неприводимый множитель степени выше m. Таким образом, при малом m придется слишком долго ждать даже первого появления t, при котором $c(X) \equiv b(X)^t \pmod{f(X)}$ имеет разложение желаемого вида. Итак, m должно быть, с одной стороны, не слишком мало, а с другой — заметно меньше n. Оптимальный выбор m, зависящий, разумеется, от p и n, требует глубокого исследования вероятностей и оценок трудоемкости, что выходит за рамки данной книги. Ограничимся примером: при $p=2$ и $n=127$ наилучшим значением является $m=17$ (тогда $h=16510$). Величина $q=2^{127}$ берется часто, так как $\#(\mathbf{F}_{2^{127}}^*) = 2^{127}-1$ есть простое число Мерсенна.

Теперь вернемся к индексному алгоритму и опишем его заключительный этап. Пусть требуется вычислить дискретный логарифм элемента $y(X) \in \mathbf{F}_q^*$ и на первом этапе найдены величины $\operatorname{ind} a(X)$ для всех $a(X) \in B$. Опять выбираем случайно t между 1 и $q-2$ и вычисляем $y_1 = yb^t$, т. е. единственный многочлен $y_1(X) \in \mathbf{F}_p[X]$ степени ниже n, удовлетворяющий условию $y_1(X) \equiv y(X)b(X)^t \pmod{f(X)}$. Как и на первом этапе алгоритма, проверяем, разлагается ли $y_1(X)$ в произведение константы c_0 и степеней многочленов $a(X) \in B$. Если это не так, то случайно выбираем новое t и т. д., пока не будет найдено такое целое t, что $y_1(X) \equiv y_0 \prod_{a \in B} a(X)^{\alpha_a}$. Как только это произойдет,

цель достигнута, так как по определению y_1 выполняются равенства $\operatorname{ind} y = \operatorname{ind} y_1 - t$ и $\operatorname{ind} y_1 = \operatorname{ind} y_0 + \sum \alpha_a \operatorname{ind} a(X)$. Все слагаемые члены в правой части последнего равенства известны. Это завершает описание индексного алгоритма.

Следует заметить, что Д. Копперсмит предложил метод, который в наиболее интересном случае $p = 2$ значительно ускоряет процесс вычисления дискретного логарифма. Поэтому криптосистемы, использующие дискретное логарифмирование в $\mathbf{F}_{2^n}^*$, считаются ненадежными, если n меньше 1000. Несмотря на это, поля $\mathbf{F}_{2^n}$ часто используются ввиду удобства для программирования. Хороший обзор этих вопросов (охватывающий то, что было известно к 1985 году) имеется в статье А. Одлыжко (см. приведенный ниже список литературы).

Если $q = p^n$ есть степень нечетного простого числа, имеющая длину k бит, то оказывается, что время решения задачи дискретного логарифмирования в $\mathbf{F}_q^*$ сравнимо по порядку со временем работы алгоритмов разложения на множители k-битового целого числа. Поэтому с эмпирической точки зрения задача дискретного логарифмирования, по-видимому, столь же трудна, как и задача разложения на множители (хотя еще никому не удалось доказать теорему такого рода). Действительно, когда в следующей главе будут рассматриваться алгоритмы разложения на множители, мы увидим, что один из основных методов разложения больших целых чисел поразительно похож на индексный алгоритм дискретного логарифмирования.

Поэтому еще рано судить о том, какие из криптосистем: типа RSA (основанные на сложности разложения целых чисел) или системы дискретного логарифмирования, окажутся более надежными.

УПРАЖНЕНИЯ

Замечание. Упражнения 4, 6, 7в) и 8 предназначены для читателей, располагающих компьютерами с программами умножения многократной точности. (Необходимы лишь программы для вычисления $a^b \pmod m$ для очень больших целых чисел a, b, m. Напомним, что $a^{-1} \pmod p$ может быть вычислено как a^{p-2}.)

1. Если приходится проводить много вычислений в некотором не очень большом конечном поле $\mathbf{F}_q$, то можно сэкономить время, составив полную «таблицу логарифмов». Для этого выбирается порождающий элемент g в $\mathbf{F}_q$ и строится таблица пар n, g^n для всех n от 1 до $q-1$. К двум столбцам этой таблицы присоединяются третий и четвертый столбцы, содержащие все пары $a, \log_g a$. Для этого располагаем элементы a из $\mathbf{F}_q^*$ в третьем столбце в некотором удобном для нас порядке, а затем просматриваем первые два столбца, и ставим в четвертый столбец рядом с a то значение n, при котором $g^n = a$. Например, для $\mathbf{F}_9$ (см. пример 2 в

§II.1) выбираем g как корень α многочлена $X^2 - X - 1$ и строим таблицу

n	g^n	a	$\log_g a$
1	α	1	8
2	$\alpha+1$	-1	4
3	$-\alpha+1$	α	1
4	-1	$\alpha+1$	2
5	$-\alpha$	$\alpha-1$	7
6	$-\alpha-1$	$-\alpha$	5
7	$\alpha-1$	$-\alpha+1$	3
8	1	$-\alpha-1$	6

После этого умножение и деление сводятся к сложению или вычитанию по модулю $q-1$ и использованию таблицы. Например, для умножения $\alpha-1$ на $-\alpha-1$ надо найти эти два числа в третьем столбце, сложить отвечающие им логарифмы: $7+6 \equiv 5 \pmod 8$, и, наконец, найти ответ $-\alpha$ во втором столбце рядом с числом 5 из первого столбца.

а) Построить таблицу логарифмов для $\mathbf{F}_{31}^*$ и воспользоваться ею для вычисления $16 \cdot 17$, $19 \cdot 13$, $1/17$, $20/23$.

б) Построить таблицу логарифмов для $\mathbf{F}_8^*$ и воспользоваться ею для вычисления следующих элементов (здесь α — корень X^3+X+1): $(\alpha+1)(\alpha^2+\alpha)$, $(\alpha^2+\alpha+1)(\alpha^2+1)$, $1/(\alpha^2+1)$, $\alpha/(\alpha^2+\alpha+1)$. Ответы должны быть записаны в виде многочленов от α не выше второй степени.

2. На первый взгляд кажется, что при постановке задачи дискретного логарифмирования вместо $\mathbf{F}_q^*$ можно использовать циклическую группу $(\mathbf{Z}/p^\alpha\mathbf{Z})^*$ (см. упражнение 2 а) к § II.1). Однако задача дискретного логарифмирования для $\mathbf{Z}/p^\alpha\mathbf{Z}$ при $\alpha > 1$ даже при больших α лишь немногим сложнее задачи логарифмирования при $\alpha = 1$ (т.е. для $\mathbf{F}_p$). Точнее, используя приемы, описанные в этом упражнении, можно доказать, что если мы можем решать задачу дискретного логарифмирования «по модулю p», то для логарифмирования «по модулю p^α» остается произвести лишь небольшую дополнительную работу за полиномиальное по $\log p^\alpha = \alpha \log p$ время. (Напомним, что пока нет алгоритма дискретного логарифмирования «по модулю p» при больших p за полиномиальное относительно $\log p$ время, и специалисты сомневаются в том, что такой алгоритм существует.) В этом упражнении мы увидим, что при $p = 3$ существует несложный алгоритм дискретного логарифмирования «по модулю 3^α» за полиномиальное по α время.

Итак, возьмем $g = 2$ (легко показать, что 2 — порождающий элемент $(\mathbf{Z}/3^\alpha\mathbf{Z})^*$ при любом α). Пусть a — некоторое целое не делящееся на 3 число, и нужно решить сравнение $2^x \equiv a \pmod{3^\alpha}$. Показать, что приводимый ниже алгоритм всегда позволяет найти x за полиномиальное относительно α время, и оценить (в терминах O-большое) необходимое для этого число двоичных операций.

1) Показать, что эта задача дискретного логарифмирования эквивалентна решению сравнения $2^x a \equiv 1$. Далее показать, что без потери общности можно считать, что $a \equiv 1 \pmod 3$, а x — четное число. Поэтому исходное сравнение можно заменить сравнением $4^x a \equiv 1 \pmod{3^\alpha}$.

2) Запишем $x = x_0 + 3x_1 + \cdots + 3^{\alpha-2}x_{\alpha-2}$, где x_j — разряды числа x в троичной системе счисления. Возьмем $x_{-1} = 0$. Тогда сравнение

$$4^{x_0+3x_1+\cdots+3^{j-2}x_{j-2}} a \equiv 1 \pmod{3^j} \qquad (*)_j$$

выполнено при $j = 1$. Положим $g_1 = 4$. В процессе работы в качестве промежуточных результатов наш алгоритм будет вычислять $g_j \stackrel{\text{def}}{=} 4^{3^{j-1}} \pmod{3^\alpha}$.

Положим $a_1 = a$. При $j > 1$ определим a_j как наименьший положительный вычет по модулю 3^α числа $4^{x_0+3x_1+\cdots+3^{j-2}x_{j-2}}a$. По ходу алгоритма мы будем последовательно вычислять a_j.

3) Пусть $j > 1$ и уже найдены такие $x_0, \ldots, x_{j-3}$, что выполнено сравнение $(*)_{j-1}$. Предположим также, что мы уже вычислили $g_{j-1} = 4^{3^{j-2}} \pmod{3^\alpha}$ и a_{j-1}. Сначала полагаем x_{j-2} равным $(1-a_{j-1})/3^{j-1}$ по модулю 3. (Заметим, что в силу $(*)_{j-1}$ имеем $a_{j-1} \equiv 1 \pmod{3^{j-1}}$.) Далее вычисляем $a_j = g_{j-1}^{x_{j-2}} a_{j-1} \pmod{3^\alpha}$. Наконец, если $j < \alpha$, вычисляем g_j, возводя g_{j-1} в третью степень в кольце вычетов по модулю 3^α.

4) Если $j = \alpha$, то алгоритм свою работу заканчивает.

3. Вы с вашим другом договорились поддерживать связь, используя аффинное преобразование шифрования $C \equiv AP + B \pmod{N}$ (см. примеры 3 и 4 в § III.1; в этих примерах коэффициенты преобразования обозначались строчными буквами a и b). Элементами сообщений являются отдельные буквы 31-буквенного алфавита, в котором A–Z имеют числовые эквиваленты 0–25, пробел $= 26$, $. = 27$, $? = 28$, $! = 29$, $' = 30$. Вы рассматриваете ключ $K_E = (A, B)$ как элемент $A + Bi$ поля из 31^2 элементов (i обозначает корень из -1 в этом поле). Вы также договорились обменяться ключами с помощью системы Диффи–Хеллмана и взять $g = 4 + i$. Затем вы выбираете случайно секретное целое число $a = 209$. Ваш друг прислал вам свое значение $g^b = 1 + 19i$.

а) Определить ключ шифрования.

б) Какой элемент из $\mathbf{F}_{961}$ вы должны послать вашему другу для того, чтобы он мог определить этот ключ?

в) Найти преобразование дешифрования.

г) Прочитать сообщение «BUVCFIWOUJTZ!H».

4. Вы получили шифртекст «VHNHDOAM», зашифрованный с помощью матрицы размера (2×2)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

применяемой к биграммам обычного 26-буквенного алфавита. Эта матрица шифрования была построена с помощью обмена ключами по методу Диффи–Хеллмана следующим образом. Имеется простое поле из 3602561 элементов. Ваш корреспондент послал вам $g^b = 983776$. Вы выбрали случайно показатель $a = 1082389$. Наконец, вы договорились получать матрицу из ключевого числа $K_E \in \mathbf{F}_{3602561}$, выражая наименьший неотрицательный вычет K_E по модулю 26^4 в форме $a \cdot 26^3 + b \cdot 26^2 + c \cdot 26 + d$, где a, b, c, d — содержимое разрядов в записи вычета в системе счисления по основанию 26. Если полученная таким образом матрица необратима по модулю 26, то заменяем K_E на $K_E + 1$ и повторяем процесс построения снова. В качестве матрицы шифрования возьмем первую обратимую матрицу, встретившуюся среди построенных.

а) Воспользоваться имеющейся информацией для определения матрицы шифрования.

б) Найти матрицу дешифрования и прочитать сообщение.

5. Пусть каждый пользователь A имеет секретную пару преобразований f_A и f_A^{-1} из $\mathcal{P}$ в $\mathcal{P}$, где $\mathcal{P}$ — заданное множество элементов открытого текста. Для защиты информации используется метод Мэсси–Омуры, а именно, пользователь

А посылает пользователю В сообщение $f_A(P)$, получает от В назад $f_B(f_A(P))$ и т. д. Указать условия на совокупность f_A, необходимые для работоспособности этой системы.

6. Пусть p — простое число Ферма 65537 и $g = 5$. Вы приняли сообщение (29095, 23846), которое построено с помощью криптосистемы Эль-Гамаля в $\mathbf{F}_p^*$ и вашего открытого ключа g^a. Ваш секретный ключ для дешифрования $a = 13908$. Вы договорились преобразовывать целые из $\mathbf{F}_p$ в триграммы 31-буквенного алфавита из упражнения 3, записывая их в системе счисления по основанию 31 и принимая содержимое разрядов в этой записи (от старшего разряда к младшему) в качестве числовых эквивалентов букв триграммы. Дешифровать принятое сообщение.

7. а) Показать, что выбор $\mathbf{F}_p$ с простым числом Ферма $p = 2^{2^k} + 1$ является исключительно неудачным, построив алгоритм, производящий дискретное логарифмирование в $\mathbf{F}_p^*$ за полиномиальное по $\log p$ время. Пусть g — порождающий элемент (например, 5 или 7, как показано в упражнении 15 к § II. 2) и для заданного a требуется найти такое x, что $0 \leqslant x < p - 1 = 2^{2^k}$ и $g^x \equiv a \pmod p$). Записать x в двоичной системе и построить алгоритм по образцу алгоритма извлечения квадратного корня по модулю p, описанному в конце § II. 2.

б) Найти оценку вида O-большое (в терминах p) для числа двоичных операций, необходимых для нахождения x с помощью алгоритма пункта а).

в) Использовать алгоритм пункта а) для нахождения величины k в упражнении 6.

8. Пусть элементами открытого текста являются 18-буквенные блоки обычного 26-буквенного алфавита, числовыми эквивалентами которых являются 18-разрядные — в системе счисления по основанию 26 — целые числа (разряды записываются в порядке убывания степеней основания). Получено сообщение

$$(82746592004375034872957717, 16406376843791542595481935 1),$$

зашифрованное с помощью криптосистемы Эль-Гамаля в простом поле из 29726270500913900677161 1927 элементов и вашего открытого ключа g^a. Ваш секретный ключ $a = 10384756843984756438549809$. Дешифровать полученное сообщение.

9. Ниже описана схема (также предложенная Эль-Гамалем) посылки подписи, использующая большое простое конечное поле $\mathbf{F}_p$. Объяснить, почему Алиса может выполнить все необходимые для пересылки подписи шаги за полиномиальное по $\log p$ время, почему Боб может проверить, что это именно Алиса послала свою подпись, и почему эта система не дает защиты от злоумышленника, если он может решать задачу дискретного логарифмирования в $\mathbf{F}_p^*$.

Пусть зафиксированы и общеизвестны p и элемент $g \in \mathbf{F}_p^*$. Каждый пользователь А выбирает случайно и сохраняет в секрете целое $a_A, 0 < a_A < p - 1$, а публикует $y_A = g^{a_A}$.

Для посылки своей подписи, состоящей из элементов с числовыми эквивалентами S из диапазона $0 \leqslant S < p - 1$, Алиса сначала выбирает случайно целое число k, взаимно простое с $p - 1$. Она вычисляет $r = g^k \pmod p$ и после этого решает сравнение $g^S \equiv y_A^r r^x \pmod p$ относительно x. Затем она посылает пару (r, x) Бобу вместе со своей подписью S. Борис проверяет, что в самом деле $g^S \equiv y_A^r r^x \pmod p$. Убедившись в этом, он может быть вполне уверен, что именно Алиса прислала сообщение S.

10. Используя алгоритм Силвера–Полига–Хеллмана, найти дискретный логарифм числа 153 по основанию 2 в $\mathbf{F}_{181}^*$ (2 — порождающий элемент в $\mathbf{F}_{181}^*$).

11. а) Какова вероятность (в процентах) того, что случайный многочлен над $\mathbf{F}_2$ степени 10 разлагается в произведение многочленов степени не выше второй? Какова вероятность (в процентах) того, что случайный ненулевой многочлен над $\mathbf{F}_2$ степени не выше 10 разлагается в такое произведение?

б) Какова вероятность того, что случайный нормированный многочлен над $\mathbf{F}_3$ десятой степени разлагается в произведение многочленов степени не выше второй? Какова вероятность того, что случайный нормированный многочлен над $\mathbf{F}_3$ степени не выше 10 разлагается в такое произведение?

12. При $n > m \geqslant 1$ обозначим $P_p(n, m)$ вероятность того, что случайный нормированный многочлен над $\mathbf{F}_p$ степени не выше n является произведением неприводимых сомножителей степени не выше m.

а) Доказать, что при любых фиксированных n и m существует $P(n, m) = \lim_{p\to\infty} P_p(n, m)$ и $0 < P(n, m) < 1$.

б) Найти точное выражение для $P(n, 2)$.

в) Вычислить $P(n, 2)$ при всех $n \leqslant 7$.

ЛИТЕРАТУРА к § 3 ГЛАВЫ IV

1. *Adleman L. M.* A subexponential algorithm for the discrete logarithm problem with applications to cryptography. — In: Proceedings of the 20th Annual Symposium on the Foundations of Computer Science, 1979, p. 55–60.
2. *Adleman L. M., DeMarrais J.* A subexponential algorithm for discrete logarithms over all finite fields. — Math. Comput., 1993, v. 61, p. 1–15.
3. *Coppersmith D.* Fast evaluation of logarithms in fields of characteristic two. — IEEE Trans. Inform. Theory IT-30, 1984, p. 587–594.
4. *Coppersmith D., Odlyzko A., Schroeppel R.* Discrete logarithms in $GF(p)$. — Algorithmica, 1986, v. 1, p. 1–15.
5. *Diffie W., Hellman M. E.* New directions in cryptography. — IEEE Trans. Inform. Theory IT-22, 1976, p. 644–654.
6. *ElGamal T.* A public key cryptosystem and a signature scheme based on discrete logarithms. — IEEE Trans. Inform. Theory IT-31, 1985, p. 469–472.
7. *ElGamal T.* A subexponential-time algorithm for computing discrete logarithms over $GF(p^2)$. — IEEE Trans. Inform. Theory IT-31, 1985, p. 473–481.
8. *Fellows M., Koblitz N.* Fixed-parameter complexity and cryptography. — In: Proceedings of the Tenth International Symposium on Applied Algebra, Algebraic Algorithms and Error Correcting Codes (San Juan, Puerto Rico), 1993.
9. *Gordon D.* Discrete logarithms in $GF(p)$ using the number field sieve. — SIAM J. Discrete Math., 1993, v. 6, p. 124–138.
10. *Gordon D., McCurley K.* Massively parallel computation of discrete logarithms. — Advances in Cryptology — Crypto '92. Berlin etc.: Springer, 1992.
11. *Knuth D. E.* The Art of Computer Programming. V. II. Reading etc.: Addison-Wesley, 1973.
12. *LaMacchia B., Odlyzko A.* Computation of discrete logarithms in prime fields. — Designs, Codes and Cryptography, 1991, v. 1, p. 47–62.

13. *Massey J. L.* Logarithms in finite cyclic groups — cryptographic issues. — In: Proceedings of the 4th Benelux Symposium on Information Theory, 1983, p. 17–25.
14. *McCurley K.* The discrete logarithm problem. — Crypt. Comput. Number Theory, Proc. Symp. Appl. Math., 1990, v. 42, p. 49–74.
15. *Odlyzko A. M.* Discrete logarithms in finite fields and their cryptographic significance. — In: Advances in Cryptology. Proceedings of Eurocrypt 84. Heidelberg etc.: Springer, 1985, p. 224–314.
16. *Wah P. K. S., Wang M. Z.* Realization and application of the Massey–Omura lock. — In: Proceedings of the International Zürich Seminar, 1984, p. 175–182.

§ 4. Задача о рюкзаке

В этом разделе мы опишем иной тип криптосистем с открытым ключом, использующих так называемую «задачу о рюкзаке». Представьте себе, что у вас есть большой рюкзак объема V, с которым вы собираетесь идти в долгий поход по диким местам. Имеется большое число предметов (скажем, их k штук) объемов v_i, $i = 0, \ldots, k-1$, соответственно, которые можно положить в рюкзак. Предположим, что вы опытный упаковщик рюкзаков и можете заполнить его, не оставляя свободного места. Чтобы взять максимально возможный груз, вам надо определить совокупность предметов, которая полностью заполнит рюкзак. Другими словами, требуется найти такое подмножество $I \subset \{1, \ldots, k\}$, что $\sum_{i \in I} v_i = V$, если оно существует. Это *общая задача о рюкзаке*. Далее мы предполагаем, что V и все v_i являются натуральными числами. Тогда задача может быть поставлена в следующей форме.

Задача о рюкзаке. Пусть заданы множество $\{v_i\}$ из k натуральных чисел и целое число V. Требуется найти такое k-разрядное двоичное число $n = (\varepsilon_{k-1}\varepsilon_{k-2}\ldots\varepsilon_1\varepsilon_0)_2$ (где $\varepsilon_i \in \{0,1\}$ суть значения разрядов в двоичной записи числа n), что $\sum_{i=0}^{k-1} \varepsilon_i v_i = V$, если такое n существует.

Заметим, что, в зависимости от значений набора $\{v_i\}$ и числа V, такого решения может не быть вообще, решений может быть несколько или решение будет единственным.

Частным случаем задачи о рюкзаке является *задача о рюкзаке с быстрорастущим набором*. Это случай, когда величины v_i, будучи упорядоченными в порядке возрастания, обладают тем свойством, что каждое число больше *суммы* всех предыдущих.

Пример 1. Набор $\{2, 3, 7, 15, 31\}$ является быстрорастущим набором.

Известно, что общая задача о рюкзаке относится к классу очень

трудных задач, называемых «NP-полными задачами». Это означает, что она эквивалентна по сложности известной «задаче о коммивояжере». В частности, если верна основная гипотеза теории сложности, в чем уверено большинство специалистов, то не существует алгоритма, который бы решал произвольную задачу о рюкзаке за время, полиномиальное по k и по $\log B$, где B — граница для значений V и v_i.

Однако задача о рюкзаке с быстрорастущим набором несравненно проще. А именно, просмотрим (упорядоченный по убыванию) список v_i, начиная с наибольшего, до тех пор, пока не найдем первый элемент, который не превосходит V. Соответствующее значение i мы включим в наше множество I (т. е. положим $\varepsilon_i = 1$), заменим V на $V - v_i$ и продолжим простмотр списка v_i по убыванию, пока опять не найдем элемент, не превосходящий эту разность. Продолжая действовать таким способом, мы, очевидно, построим подмножество элементов v_i, в сумме равных V, или же исчерпаем все элементы v_i, не доведя разность $V - \sum_{i\in I} v_i$ до нуля, что будет в случае, когда задача не имеет решения. Теперь мы запишем алгоритм решения в более формальном виде, удобном для преобразования в компьютерную программу.

Следующий алгоритм (полиномиальной сложности) решает задачу о рюкзаке при заданном быстрорастущем наборе $\{v_i\}$ и целом V.

1. Положим W равным V и j равным k.

2. Начиная с ε_{j-1} и последовательно уменьшая j, полагаем все $\varepsilon_j = 0$ до тех пор, пока не придем к первому такому значению i (обозначаем его через i_0), что $v_{i_0} \leqslant W$. Положим $\varepsilon_{i_0} = 1$.

3. Заменим W на $W - v_{i_0}$, положим $j = i_0$ и, если $W > 0$, то переходим к шагу 2.

4. Если $W = 0$, то цель достигнута. Если $W > 0$ и все оставшиеся v_i больше W, то ясно, что решения $n = (\varepsilon_{k-1} \ldots \varepsilon_0)_2$ нашей задачи не существует. Заметим, что решение, если оно есть, единственно.

Пример 2. Пусть v_i те же, что в примере 1, а $V = 24$. Тогда, проходя пятерку $\{2, 3, 7, 15, 31\}$ справа налево, мы видим, что $\varepsilon_4 = 0$, $\varepsilon_3 = 1$ (в этот момент мы заменяем 24 на $24 - 15 = 9$), $\varepsilon_2 = 1$ (в этот момент мы заменяем 9 на $9 - 7 = 2$), $\varepsilon_1 = 0$, $\varepsilon_0 = 1$. Таким образом, $n = (01101)_2 = 13$.

Теперь мы опишем, как построить рюкзачную криптосистему (называемую также системой Меркля–Хеллмана). Сначала предположим, что элементы открытого текста имеют в качестве своих числовых эквивалентов k-разрядные двоичные числа P. Например, если мы имеем дело с отдельными буквами 26-буквенного алфавита, то каждой букве естественным образом ставится в соответствие свое пятиразрядное двоичное число от $0 = (00000)_2$ до $25 = (11001)_2$.

Далее каждый пользователь выбирает быстрорастущий набор

$\{v_0, \ldots, v_{k-1}\}$, целое число m, большее $v_0 + \cdots + v_{k-1}$, и взаимно простое с m целое число a, $0 < a < m$. Это делается с помощью случайной процедуры. Например, мы можем выбрать произвольно последовательность из $k+1$ положительных целых чисел $z_0, \ldots, z_k$, меньших некоторого подходящего предела, положить $v_0 = z_0$, $v_i = z_i + v_{i-1} + v_{i-2} + \cdots + v_0$ для $i = 1, \ldots, k-1$ и $m = z_k + \sum_{i=0}^{k-1} v_i$. Затем мы можем выбрать случайно положительное $a_0 < m$ и выбрать a как наименьшее целое, большее или равное a_0, которое взаимно просто с m. После этого вычисляются $b = a^{-1} \pmod m$ (т.е. b является наименьшим положительным целым, удовлетворяющим условию $ab \equiv 1 \pmod m$)) и k-элементный набор $\{w_i\}$, определяемый равенствами $w_i \equiv av_i \pmod m$ (т.е. w_i является наименьшим положительным вычетом av_i по модулю m). Пользователь держит числа v_i, m, a и b в секрете, а набор $\{w_i\}$ делает общеизвестным. Таким образом, ключом шифрования K_E является набор $\{w_0, \ldots, w_{k-1}\}$, а ключом дешифрования K_D — пара (b, m) (которая вместе с ключом шифрования позволяет определить набор $\{v_0, \ldots, v_{k-1}\}$).

Желающий передать сообщение $P = (\varepsilon_{k-1}, \ldots, \varepsilon_1, \varepsilon_0)$ пользователю с ключом шифрования $\{w_i\}$ вычисляет $C = f(P) = \sum_{i=0}^{k-1} \varepsilon_i w_i$ и передает это целое число.

Чтобы прочитать это сообщение, пользователь сначала находит наименьший положительный вычет V числа bC по модулю m. Так как $bC \equiv \sum \varepsilon_i b w_i \equiv \sum \varepsilon_i v_i \pmod m$ (поскольку $bw_i \equiv bav_i \equiv v_i \pmod m$)), то $V = \sum \varepsilon_i v_i$. (Здесь мы воспользовались тем фактом, что каждое из условий $V < m$ и $\sum \varepsilon_i v_i \leqslant \sum v_i < m$ превращает сравнение по модулю m в равенство.) Теперь можно воспользоваться приведенным выше алгоритмом для задачи о рюкзаке с быстрорастущим набором и найти единственное решение $(\varepsilon_{k-1} \ldots \varepsilon_0)_2 = P$ задачи о подмножестве $\{v_i\}$ с суммой, равной V. Так мы восстанавливаем сообщение P.

Заметим, что злоумышленнику, знающему только набор $\{w_i\}$, придется решать задачу о рюкзаке $C = \sum \varepsilon_i w_i$ уже *не* с быстрорастущим набором, поскольку свойство быстрого роста $\{v_i\}$ разрушается умножением на a и приведением по модулю m. Поэтому приведенный выше алгоритм уже нельзя использовать, и возникающая перед злоумышленником задача кажется значительно более сложной. Мы вернемся к этому позднее.

Пример 3. Предположим, что элементами открытого текста являются буквы 26-буквенного алфавита, которым, как и выше, отвечают двоичные числа от $0 = (00000)_2$ до $25 = (11001)_2$, а наш секретный ключ дешифрования — это быстрорастущий набор из примера 1. Выберем $m = 61$, $a = 17$. Тогда $b = 18$ и ключ шифрования — это

набор (34, 51, 58, 11, 39). Чтобы послать сообщение «WHY», наш корреспондент должен вычислить «W»$= (10110)_2 \mapsto 51 + 58 + 39 = 148$, «$H$»$= (00111)_2 \mapsto 34 + 51 + 58 = 143$, «$Y$»$= (11000)_2 \mapsto 11 + 39 = 50$. Чтобы прочитать сообщение 148, 143, 50, мы сначала умножаем эти числа на 18 и приводим результаты по модулю 61; получаем 41, 12, 46. Далее, взяв поочередно $V = 41$, $V = 12$ и $V = 46$ и проведя такие же рассуждения, как в примере 2, мы восстановим сообщение $(10110)_2$, $(00111)_2$, $(11000)_2$.

Конечно, использование однобуквенных элементов открытого текста со столь малым $k = 5$ не обеспечивает надежной защиты. Пример 3 призван лишь проиллюстрировать механизм системы.

Одно время многие специалисты очень оптимистично оценивали возможности рюкзачных криптосистем. Раз задача раскрытия системы принадлежит классу очень трудных задач (NP-полные задачи), то, считали они, система должна быть надежной.

Однако в этих рассуждениях крылась ошибка. Порождаемые такими системами задачи о рюкзаке $C = \sum \varepsilon_i w_i$, хотя и не соответствуют быстрорастущим наборам, но имеют весьма специфический тип, а именно, получаются из задач с быстрорастущим набором простым преобразованием: умножением всех величин на a и приведением по модулю m. В 1982 году Шамир нашел полиномиальный по k алгоритм решения задач о рюкзаке такого типа. Поэтому исходная криптосистема Меркля–Хеллмана не может считаться надежной криптосистемой с открытым ключом.

Одним из способов защиты от алгоритма Шамира может быть некоторое усложнение рюкзачной системы последовательностью преобразований вида $x \mapsto ax \pmod m$. Например, можно просто применить два преобразования с параметрами (a_1, m_1) и (a_2, m_2) соответственно. Мы сначала заменяем нашу быстрорастущую последовательность $\{v_i\}$ на $\{w_i\}$, где w_i — наименьший положительный вычет $a_1 v_i \pmod{m_1}$, а затем получаем третью последовательность $\{u_i\}$, взяв наименьшие положительные вычеты $u_i = a_2 w_i \pmod{m_2}$. Здесь мы выбираем m_1, m_2 и a_1, a_2 случайно так, чтобы они удовлетворяли условиям $m_1 > \sum v_i$, $m_2 > km_1$ и НОД $(a_1, m_1) =$ НОД $(a_2, m_2) = 1$. Тогда открытым ключом будет набор величин u_i, а шифрующей функцией — функция $C = f(P) = \sum_{i=0}^{k-1} \varepsilon_i u_i$, где $P = (\varepsilon_{k-1} \cdots \varepsilon_0)_2$. При дешифровании шифртекста используется ключ $K_D = (b_1, m_1, b_2, m_2)$ (где $b_1 = a_1^{-1} \pmod{m_1}$ и $b_2 = a_2^{-1} \pmod{m_2}$). Сначала вычисляется наименьший положительный вычет $b_2 C$ по модулю m_2, затем результат домножается на b_1 и приводится по модулю m_1. Так как $b_2 C \equiv \sum \varepsilon_i w_i \pmod{m_2}$ и $m_2 > km_1 > \sum w_i$, то результат приведе-

ния b_2C по модулю m_2 *совпадает* с $\sum \varepsilon_i w_i$. Потом, когда мы берем $b_1 \sum \varepsilon_i w_i \pmod{m_1}$, получаем $\sum \varepsilon_i v_i$, откуда можно определить ε_i, используя приведенный выше алгоритм для задачи о рюкзаке с быстрорастущим набором.

В настоящее время еще не найден полиномиальный алгоритм решения итерированной задачи о рюкзаке (т.е. криптосистемы с открытым ключом, описанной в предыдущем абзаце). Однако Брикелю с соавторами удалось обобщить алгоритм Шамира и показать, что к итерированной рюкзачной криптосистеме применимы эффективные методы криптоанализа. Во всяком случае, после достижения Шамира большинство экспертов потеряло доверие к стойкости криптосистем с открытым ключом данного типа.

Одна до сих пор не вскрытая рюкзачная система. Теперь опишем метод передачи сообщений, основанный на однонаправленной функции рюкзачного типа, который использует многочлены над конечным полем. Эту криптосистему предложили Чор и Райвест; мы изложим слегка упрощенный (и менее эффективный) вариант их конструкции.

Опять предположим, что Алиса хочет принимать сообщения в форме наборов из k битов $\varepsilon_0, \ldots, \varepsilon_{k-1}$. (Число k Алиса выбирает так, чтобы выполнялись указанные ниже условия.) Как и раньше, ее открытым ключом является последовательность натуральных чисел $v_0, \ldots, v_{k-1}$, построенная описываемым ниже способом. Но теперь Боб должен посылать ей не только целое число $c = \sum \varepsilon_j v_j$, но и сумму разрядов $c' = \sum \varepsilon_j$.

Алиса строит последовательность v_j следующим образом. Выбор всех параметров, которые упоминаются в этом абзаце, можно делать секретным образом, так как Бобу для посылки сообщения необходим лишь итоговый набор $v_0, \ldots, v_{k-1}$. Сначала Алиса выбирает такую степень простого $q = p^f$, что $q - 1$ не имеет больших простых делителей (в этом случае возможно практическое вычисление логарифмов в группе $\mathbf{F}_q^*$, см. § 3), причем как p, так и f имеют умеренные значения (например, записываются двумя или тремя десятичными разрядами). Так, в работе Чора и Райвеста 1988 г. рассматривалась величина $q = 197^{24}$. Далее Алиса выбирает нормированный неразложимый многочлен $F(X) \in \mathbf{F}_p[X]$ степени f, так что $\mathbf{F}_q$ может рассматриваться как $\mathbf{F}_p[X]/F[X]$. Она выбирает также образующий элемент g в $\mathbf{F}_q^*$ и целое число z. Алиса производит выбор F, g и z каким-нибудь случайным образом.

Пусть $t \in \mathbf{F}_q = \mathbf{F}_p[X]/F(X)$ обозначает класс вычетов, содержащий X. Алиса выбирает в качестве k любое целое число, меньшее p и

f одновременно. Для $j = 0, \ldots, k-1$ она вычисляет такие неотрицательные целые $b_j < q-1$, что $g^{b_j} = t + j$ (согласно предположению, Алиса может легко вычислять дискретные логарифмы в $\mathbf{F}_q^*$). Наконец, Алиса выбирает случайно перестановку π элементов множества $\{0, \ldots, k-1\}$ и полагает v_j равным наименьшему неотрицательному вычету $b_{\pi(j)} + z$ по модулю $q-1$. Она объявляет $(v_0, \ldots, v_{k-1})$ своим открытым ключом.

Расшифрование происходит следующим образом. Получив от Боба c и c', Алиса сначала вычисляет элемент $g^{c-zc'}$, который единственным образом представляется в виде многочлена $G(X) \in \mathbf{F}_p[X]$ степени ниже f. Но она знает, что этот элемент равен также $\prod g^{\varepsilon_j b_{\pi(j)}} = \prod (t+\pi(j))^{\varepsilon_j}$, что соответствует многочлену $\prod (X+\pi(j))^{\varepsilon_j}$. Поскольку многочлены $G(X)$ и $\prod (X+\pi(j))^{\varepsilon_j}$ имеют степень ниже f и представляют собой один и тот же элемент по модулю $F(X)$, должно выполняться равенство

$$G(X) = \prod (X + \pi(j))^{\varepsilon_j}.$$

Теперь Алиса может найти величины ε_j, разлагая $G(X)$ на множители (для чего имеется эффективный алгоритм; см. второй том книги Кнута).

УПРАЖНЕНИЯ

1. Для каждого из приводимых вариантов наборов и «объемов» определить, является ли набор быстрорастущим, имеет ли задача о рюкзаке решение и, если имеет, то сколько:

а) $\{2, 3, 7, 20, 35, 69\}$, $V = 45$; б) $\{1, 2, 5, 9, 20, 49\}$, $V = 73$;

в) $\{1, 3, 7, 12, 22, 45\}$, $V = 67$; г) $\{2, 3, 6, 11, 21, 40\}$, $V = 39$;

д) $\{4, 5, 10, 30, 50, 101\}$, $V = 186$; е) $\{3, 5, 8, 15, 28, 60\}$, $V = 43$.

2. а) Показать, что быстрорастущая последовательность с наименьшими членами — это последовательность $1, 2, 4, 8, \ldots, 2^i, \ldots$

б) Показать, что задача о рюкзаке с этим быстрорастущим набором всегда имеет решение n, а именно, $n = V$, и что не существует других быстрорастущих наборов, при которых задача всегда имеет решение.

3. Показать, что всякая последовательность натуральных чисел $\{v_i\}$ со свойством $v_{i+1} \geqslant 2v_i$ при всех i является быстрорастущей.

4. Предположим, что элементами открытого текста являются буквы 26-буквенного алфавита A–Z, которым отвечают числа от 0 до 25. Вы принимаете последовательность элементов шифртекста 14, 25, 89, 3, 65, 24, 3, 49, 89, 24, 41, 25, 68, 41, 71. Открытым ключом является последовательность $\{57, 14, 3, 24, 8\}$, а секретным ключом — пара $b = 23$, $m = 61$.

а) Попробуйте дешифровать сообщение, не пользуясь ключом дешифрования; проверьте результат, используя ключ дешифрирования и алгоритм решения задач о рюкзаке с быстрорастущим набором.

б) Использовать открытый ключ для посылки сообщения «TENFOUR».

5. Предположим, что элементами открытого текста являются тройки букв 32-буквенного алфавита, где буквам A–Z отвечают числа 0–25, пробел = 26, ? = 27, ! = 28, . = 29, '= 30, $ = 31. Вы получили последовательность элементов шифртекста 152472, 116116, 68546, 165420, 168261. Открытым ключом является набор {24038, 29756, 34172, 34286, 38334, 1824, 18255, 19723, 143, 17146, 35366, 11204, 32395, 12958, 6479}, а секретным ключом — пара $b = 30966$, $m = 47107$. Расшифруйте сообщение.

6. Предположим, что элементами открытого текста являются биграммы букв 32-буквенного алфавита из примера 5. Вы получили последовательность элементов шифртекста 33219, 7067, 18127, 43099, 37953, которые зашифрованы с помощью двухступенчатого рюкзачного ключа {23161, 6726, 4326, 16848, 21805, 11073, 120, 15708, 2608, 341}. Секретный ключ: $b_1 = 533$, $m_1 = 2617$, $b_2 = 10175$, $m_2 = 27103$. Расшифруйте сообщение.

ЛИТЕРАТУРА к § 4 ГЛАВЫ IV

1. *Brickell E.* Breaking iterated knapsacks. — In: Advances in Cryptology. Crypto'84. Heidelberg etc.: Springer, 1985, p. 342–358.
2. *Brickell E., Odlyzko A.* Cryptanalysis: A survey of recent results. — Proc. IEEE, 1988, v. 76, p. 578–593.
3. *Chor B., Rivest R.* A knapsack-type public key cryptosystem based on arithmetic in finite fields. — In: Advances in Cryptology. Crypto'84. Heidelberg etc.: Springer, 1985, p. 54–65; revised version in IEEE Trans. Inform. Theory IT-34, 1988, p. 901–909.
4. *Garey M. R., Johnson D. S.* Computers and Intractability: A Guide to the Theory of NP-Completeness. San-Francisco: W. H. Freeman, 1979. (Русск. перев. — *Гэри М., Джонсон Д. С.* Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982.)
5. *Goodman R. M. F., McAuley A. J.* A new trapdoor knapsack public key cryptosystem. — In: Advances in Cryptology. Proceedings of Eurocrypt 84. Heidelberg etc.: Springer, 1985, p. 150–158.
6. *Hellman M. E.* The mathematics of public-key cryptology. — Sci. American, 1979, v. 241, p. 146–157.
7. *Hellman M. E., Merkle R. C.* Hiding information and signatures in trapdoor knapsack. — IEEE Trans. Inform. Theory IT-24, 1978, p. 525–530.
8. *Odlyzko A.* The rise and fall of knapsack cryptosystems. — Crypt. Comput. Number Theory, Proc. Symp. Appl. Math., 1990, v. 42, p. 75–88.
9. *Schnorr C.* Efficient identification and signatures for smart cards. — In: Advances in Cryptology. Crypto '89. Heidelberg etc.: Springer, 1990, p. 239–251.
10. *Shamir A.* A polynomial time algorithm for breaking the basic Merkle–Hellman cryptosystem. — In: Proceedings of the 23rd Annual Symposium on the Foundations of Computer Science, 1982, p. 145–152.
11. *van Oorschot P.* A comparison of practical public-key cryptosystems based on integer factorisation and discrete logarithms. — In: Contemporary Cryptology: The Science of Information Integrity./Ed. by G. Simmons. Piscataway: IEEE Press, 1992, p. 289–322.

§ 5. Протоколы с нулевым разглашением и скрытая передача

«Нулевое разглашение» — это название криптографического понятия, введенного в начале восьмидесятых годов в связи со следующей задачей. Пусть кто-либо желает убедить кого-то в том, что он умеет делать нечто, например, решать некоторое уравнение, доказывать теорему или разгадывать головоломку, не передавая при этом никакой информации о способе решения. Возможно ли это? Как можно убедить кого-либо в том, что вы имеете некоторое решение, не демонстрируя его? Удивительно, но во многих ситуациях это вполне возможно.

«Доказывающая», которую мы назовем Пикара, знает решение. «Проверяющий», которого мы назовем Вивалес, — это тот, кого нужно убедить в том, что Пикара знает решение, не дав ему при этом даже смутного представления о том, в чем состоит это решение.

В этом разделе мы сначала приведем простой и наглядный пример интерактивного (т.е. в режиме двусторонней связи между Пикарой и Вивалесом) доказательства с нулевым разглашением. Этот пример связан с раскрашиванием карты и не имеет отношения к теории чисел. Затем мы рассмотрим другой пример: как доказать, что вы нашли дискретный логарифм, никак не помогая проверяющему найти его значение. Далее мы обсудим понятие «скрытой передачи», с помощью которого можно строить неинтерактивные доказательства с нулевым разглашением. Наконец, мы используем скрытую передачу для доказательства знания разложения на простые множители с нулевым разглашением.

Раскраска карты. Рассмотрим первый пример. В настоящее время уже известно, что любую плоскую карту можно раскрасить четырьмя цветами. Некоторые карты можно раскрасить тремя цветами, а некоторые нет. Предположим, что Пикара имеет сложную карту, которую после долгих усилий она научилась раскрашивать только тремя цветами (r — красный, b — голубой, g — зеленый). Как ей убедить Вивалеса в том, что она умеет делать это, не давая ему намека, облегчающего ему поиск раскраски?

Сначала мы переформулируем задачу на языке теории графов.

О п р е д е л е н и е. *Графом* называется пара, состоящая из множества V, элементы которого называются «вершинами», и некоторого подмножества E множества всех (неупорядоченных) пар элементов множества V. Элементы E называются «ребрами». «Ребро» $e = \{u, v\}$, где $u, v \in V$, можно изображать линией, соединяющей вершины u и v.

О п р е д е л е н и е. Говорят, что граф *можно раскрасить* в цвета r, b, g, если существует такая функция $f\colon V \to \{r, b, g\}$, что $\{u, v\} \in$

$E \Longrightarrow f(u) \neq f(v)$, т. е. никакое ребро не соединяет вершины одного и того же цвета.

Задача о трех красках состоит в определении того, можно ли данный граф раскрасить тремя цветами r, b, g.

Чтобы переформулировать задачу о раскрашивании карты в задачу о раскрашивании графа, следует просто взять в качестве V множество стран (изображая их теперь точками) и «соединять» две страны ребром в том и только том случае, если они имеют общую границу.

Задача о трех красках имеет два приятных свойства, делающих ее удобной для обсуждения многих вопросов: 1) она наглядна; 2) это NP-полная задача (по поводу этого понятия см. § 4). Из свойства NP-полноты следует, что верификацию с нулевым разглашением любой NP-задачи можно свести к верификации с нулевым разглашением задачи о раскрашиваемости в три цвета.

Однако это не означает, что после построения верификации с нулевым разглашением некоторой NP-полной задачи P_1 (скажем, раскрашиваемости в три цвета), ничего не стоит построить верификацию с нулевым разглашением для другой NP-полной задачи P_2. Напротив, в процессе сведения задачи P_2 к задаче P_1 обычно значительно увеличивается размер входных данных. Поэтому более эффективной процедурой, скорее всего, будет прямая верификация с нулевым разглашением задачи P_2, нежели сведение P_2 к P_1 с последующей верификацией с нулевым разглашением P_1. Так, ниже мы приведем непосредственное доказательство с нулевым разглашением для возможности дискретного логарифмирования. Было бы в высшей степени неэффективно строить такое доказательство сведением задачи дискретного логарифмирования к задаче о трех красках для некоторого графа.

Доказательство с нулевым разглашением для задачи о трех красках. Предположим, что Пикара имеет некоторый граф. Мы будем представлять его вершины в виде маленьких цветных лампочек, помещенных внутри шариков, соединенных между собой трубками, изображающими ребра. Лампочки могут испускать красный, голубой или зеленый свет. У Пикары есть: 1) устройство A, позволяющее переключить любую вершину на любой из трех цветов, 2) устройство B, позволяющее нажатием одной кнопки совершить случайную перестановку на множестве цветов и переключение всех вершин в соответствии с этой перестановкой. Например, если устройство B выбрало транспозицию красного и голубого цветов, то все вершины, светившие голубым светом, переключаются на красный, все вершины, светившие красным, переключаются на голубой свет, а вершины, светившие зеленым светом, сохраняют его. Вивалес не управляет устрой-

ством B и даже не знает, какая перестановка произошла.

Предположим далее, что поверхность шариков не позволяет увидеть свет лампочек снаружи. Однако, если проникнуть внутрь трубки, соединяющей две вершины, то будут видны огни находящихся в этих вершинах лампочек (и только они).

Пусть Пикара знает раскрашивание в три цвета и использует устройство A для установки соответствующих цветов в вершинах. Вот процедура, призванная убедить Вивалеса в том, что Пикара действительно знает такое раскрашивание.

1. Вивалес может заглянуть в любую трубку и увидеть цвета вершин на ее концах. Убедившись, что эти цвета различны, он получит определенное подтверждение того, что Пикара действительно знает правильную раскраску (напомним, что «правильной» является раскраска, при которой смежные вершины имеют разные цвета).

2. Далее Пикара нажимает кнопку устройства B и переставляет цвета.

3. Вивалес снова заглядывает в одну из трубок.

4. Пикара и Вивалес поочередно повторяют шаги 2 и 3 до тех пор, пока Вивалес не проверит все трубки (или, по его желанию, пока он не проверит все трубки несколько раз — он может подозревать, что Пикара мошенничает при переключении лампочек в вершинах ранее проверенных ребер).

При недолгом размышлении становятся понятными следующие две вещи. 1) Если бы Пикара не знала правильного раскрашивания в три цвета, то она не смогла дурачить Вивалеса — рано или поздно с помощью шага 3 он обнаружил бы пару смежных вершин одного цвета. 2) Благодаря случайному переключению цветов Вивалес ничего не узнает о реальной раскраске, он лишь убедится в том, что она Пикаре удалась. Если он тоже хочет решить задачу о трех красках для данного графа, то после описанной процедуры эта задача останется для него столь же трудной, как и до того.

Чтобы доказать, что Вивалес ничего не узнал о раскрашивании, приведем следующее рассуждение. Пусть некоторое третье лицо, скажем, Клайд, не знает, как раскрасить граф тремя красками, но зато *знает* заранее, в какую трубку заглянет Вивалес. Тогда Клайд, зажигая на концах этого ребра лампочки разного цвета, может добиться в точности того же эффекта, что и Пикара. Таким образом, информация, которую Вивалес получает от Клайда, неотличима от той, которую он мог бы получить от Пикары. Но Клайд едва ли может сообщить ему что-нибудь полезное о раскрашивании, так как сам ничего об этом не знает. Говорят, что Клайд «имитирует» действия Пикары. Этот прием имитации является стандартным способом до-

казательства того, что данный протокол является доказательством с нулевым разглашением.

Доказательство знания одного дискретного логарифма с нулевым разглашением. Как и в § 3, предположим, что G — конечная группа (относительно умножения) из N элементов, b — фиксированный элемент G и y — элемент G, для которого Пикара вычислила дискретный логарифм по основанию b, т. е. решила уравнение $b^x = y$ относительно положительного целого x. Ей надо доказать Вивалесу, что она знает это решение, не давая никакой зацепки для определения этого решения. Сначала предположим, что Вивалесу известен порядок N рассматриваемой группы. Вот последовательность шагов, которые они оба должны сделать.

1. Пикара вырабатывает случайное натуральное число $e < N$ и посылает Вивалесу $b' = b^e$.

2. Вивалес бросает монету. Если выподает «решка», то Пикара должна раскрыть значение e и Вивалес может проверить, что действительно $b' = b^e$.

3. Если выпадает «орел», то Пикара должна сообщить наименьший положительный вычет числа $x + e$ по модулю N, а Вивалес проверяет равенство $yb' = b^{x+e}$.

4. Шаги 1 – 3 повторяются до тех пор, пока Вивалес не поверит, что Пикара знает значение x дискретного логарифма.

Заметим, что если Пикара в действительности не знает величины x, то она может давать правильный ответ не более чем при одном варианте выпадения монеты. Если при выполнении шага 1 она ожидает, что выпадет решка, то она может, не зная x, послать Вивалесу именно $b' = b^e$. Если она ожидает, что выпадет орел, то она должна послать Вивалесу $b' = b^e/y$ (тогда на шаге 3 вместо $x + e$ она сможет послать просто e), но если в этом случае выпадет решка, то она не сможет дать правильный ответ, так как не знает, в какую степень нужно возвести b, чтобы получить b'.

Свойство отсутствия разглашения у этого протокола может быть обосновано с помощью приема имитации. А именно, пусть Клайд не знает дискретного логарифма y по основанию b, но зато *знает* заранее, как выпадет монета. Тогда он может имитировать шаги Пикары (посылая $b' = b^e$ перед выпадением решки и $b' = b^e/y$ перед выпадением орла). Получаемая в этом случае Вивалесом информация будет неотличима от информации, которую он мог бы получить от Пикары. При этом Клайд не может сообщить Вивалесу ничего полезного о значении дискретного логарифма, так как сам не знает его.

В упражнениях мы изучим ситуацию, когда Вивалес не знает

величину N. Например, он может знать, что $G = (\mathbf{Z}/M\mathbf{Z})^*$, но не знать разложения числа M. (Напомним, что если M — произведение двух простых чисел, то знание разложения эквивалентно знанию $N = \varphi(M)$, см. § I.3.) Тогда Пикара (или изображающий ее Клайд), которая пользуется величиной N на шаге 1, не должна давать Вивалесу никакой информации о N (иначе это не будет доказательством с нулевым разглашением). Это условие может показаться излишним, но любой вправе потребовать, чтобы передавалась лишь самая минимальная информация.

Скрытая передача. «Каналом скрытой передачи» информации от Пикары к Вивалесу называется система пересылки от Пикары к Вивалесу двух пакетов зашифрованной информации, удовлетворяющая следующим условиям.

1. Вивалес может дешифровать и прочитать только один пакет из двух.

2. Пикара не знает, который именно пакет он может дешифровать.

3. И Пикара, и Вивалес должны быть уверены, что условия 1 и 2 выполнены.

На первый взгляд, эти требования производят странное впечатление. Однако такой канал оказывается одним из фундаментальных понятий криптографии. Скоро мы увидим, как с его помощью строится неинтерактивное (т. е. без взаимодействия сторон) доказательство с нулевым разглашением. Но сначала мы опишем способ построения такого канала, использующий сложность решения задачи дискретного логарифмирования.

Точнее, пусть имеются такие большое конечное поле $\mathbf{F}_q$ и фиксированный элемент b мультипликативной группы $\mathbf{F}_q^*$, что не существует практически реализуемого способа вычисления b^{xy} по значениям b^x и b^y. Это — условие Диффи–Хеллмана, и считают, что оно выполняется, если задача дискретного логарифмирования в $\mathbf{F}_q^*$ трудноразрешима (см. § 3).

Пусть, далее, существует легко вычислимое (и легко обратимое) отображение ψ нашего конечного поля в n-мерное векторное пространство $\mathbf{F}_2^n$ над полем $\mathbf{F}_2$. Пусть образ $\mathbf{F}_q$ при этом отображении содержит все элементы $\mathbf{F}_2^{n-1}$ (т. е. все векторы, заканчивающиеся нулем). Например, если q взаимно просто с p, то можно определить n условием $2^{n-1} < p < 2^n$ и отображать элемент $\mathbf{F}_q$ (например, неотрицательное целое, меньшее p) в отвечающую ему последовательность двоичных знаков.

Пусть элементами сообщения являются наборы по n бит, т. е. элементы $m \in \mathbf{F}_2^n$. Наконец, зафиксируем раз и навсегда некоторый эле-

мент $C \in \mathbf{F}_q^*$, выбранный так, чтобы никто не знал его дискретного логарифма. (Напомним, что по предположению задача дискретного логарифмирования трудноразрешима в $\mathbf{F}_q^*$.) Элемент C может быть предложен «Центром доверия» (т. е. специальным посредником, которому доверяют обе стороны) или выбран с помощью случайной или интерактивной процедуры с участием Пикары и Вивалеса.

Скрытая передача происходит следующим образом. Вивалес выбирает случайно целое число $x, 0 < x < q-1$, и элемент $i \in \{1,2\}$. Далее символы x и i будут обозначать фиксированные целые числа из множеств $\{1,\ldots,q-2\}$ и $\{1,2\}$ соответственно. Вивалес полагает $\beta_i = b^x$ и $\beta_{3-i} = C/b^x$. Он объявляет своим открытым ключом пару (β_1, β_2) и сохраняет секретными величины x и i. Заметим, что Вивалес не может знать дискретный логарифм β_{3-i}, который мы обозначим через x': если бы он его знал, то он мог бы вычислить дискретный логарифм $C = \beta_i\beta_{3-i}$, что противоречит нашим предположениям.

Пусть у Пикары имеется два элемента текста: $m_1 \in \mathbf{F}_2^n$ из первого пакета и $m_2 \in \mathbf{F}_2^n$ из второго пакета. Она выбирает случайно два целых числа $0 < y_1, y_2 < q-1$ и посылает Вивалесу две пары элементов: b^{y_1}, $b^{y_2} \in \mathbf{F}_q^*$ и $\alpha_1 = m_1 + \psi(\beta_1^{y_1})$, $\alpha_2 = m_2 + \psi(\beta_2^{y_2}) \in \mathbf{F}_2^n$ (сложение производится в n-мерном векторном пространстве $\mathbf{F}_2^n$; такая операция известна как «исключающее или»). Пикара оставляет y_1 и y_2 секретными.

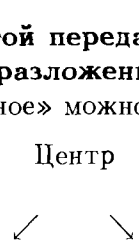
Поскольку $\beta_i^{y_i} = \left(b^{y_i}\right)^x$ и Вивалес знает числа b^{y_i} и x, он может легко найти $\psi(\beta_i^{y_i})$, а следовательно, и $m_i = \alpha_i + \psi(\beta_i^{y_i})$. Однако, если бы он захотел найти m_{3-i}, то ему пришлось бы найти величину $\beta_{3-i}^{y_{3-i}} = b^{x'y_{3-i}}$, зная лишь $b^{y_{3-i}}$ и $b^{x'}$ и не зная ни y_{3-i}, ни x'. Поэтому (в силу условия Диффи–Хеллмана) найти m_{3-i} он не может.

Заметим, что Пикара может легко проверить, что $\beta_1\beta_2 = C$ и, следовательно, убедиться, что Вивалес не знает дискретных логарифмов сразу обоих элементов своего открытого ключа (β_1, β_2). Так как в интересах Вивалеса иметь всю возможную информацию, Пикара может быть уверена в том, что дискретный логарифм одного из элементов этой пары он знает. Но у Пикары нет возможности различить β_1 и β_2, чтобы узнать, какое из них Вивалес получил как b^x, а какое — как C/b^x. Значит, как для Вивалеса, так и для Пикары условия 1 и 2 выполнены.

Если один ключ β_1 и β_2 (т. е. одни и те же x и i) используется для последовательности пар (m_1, m_2), то Пикара знает, что дешифрованные Вивалесом элементы m_i занимают в парах m_1, m_2 одинаковые места. Если требуется, чтобы новая последовательность элементов была послана независимым (от предыдущей) образом, то Вивалес должен

выбрать случайно новые величины x и i и послать новый открытый ключ β_1 и β_2.

Использование скрытой передачи для неинтерактивного доказательства того, что разложение на множители известно. Смысл слова «неинтерактивное» можно представить диаграммой



Здесь «Центр доверия» можно понимать как источник случайных битов, посылаемых одновременно Пикаре и Вивалесу (Центр может перед рассылкой производить над битами какие-нибудь арифметические операции). Комбинация этих битов и реакция на них Пикары (сообщение, посланное ею Вивалесу) должны быть достаточны для того, чтобы Вивалес поверил, что она на самом деле сделала то, о чем сообщает (т. е. чтобы вероятность того, что Вивалеса обманули, была экспоненциально мала).

Под «неинтерактивностью» понимается то, что в процессе проверки Вивалес не переписывается с Пикарой. Однако допускается, чтобы в самом начале процедуры Пикара получила длинную последовательность открытых ключей скрытой передачи (β_1, β_2) для Вивалеса, описанных выше. Это не рассматривается как переписка Вивалеса с Пикарой. Фактически, теми же самыми открытыми ключами (в соответствии со смыслом термина «открытый») может пользоваться каждый, кто захочет играть роль Пикары. И Пикара может использовать одну и ту же последовательность открытых ключей во многих различных доказательствах с нулевым разглашением, которые она направляет Вивалесу.

Теперь опишем процедуру, с помощью которой Пикара может убедить Вивалеса в том, что она знает разложение на множители целого числа $n = pq$, не сообщая ему никакой информации о значениях сомножителей. Мы будем использовать эквивалентность знания p и q возможности извлечения из произвольного числа квадратного корня в арифметике по модулю $n = pq$ (см. упражнение 5 ниже). Процедура имеет следующий вид.

1. Центр вырабатывает случайно целое число x и посылает Пикаре и Вивалесу наименьший неотрицательный вычет из x^2 по модулю n; полагаем $y = x^2 \pmod n$.

2. Пикара находит четыре квадратных корня из y по модулю n, а именно, $\pm x, \pm x'$. Она произвольным образом выбирает из них один;

обозначаем его через x_0.

3. Пикара выбирает случайно целое число r и посылает Вивалесу число $s = r^2 \pmod n$. Она также вычисляет $m_1 = r \pmod n$ и $m_2 = x_0 r \pmod n$ и посылает эти два сообщения Вивалесу с помощью скрытой передачи.

4. Вивалес может прочитать только одно из этих сообщений. Он проверяет, что квадрат этого числа по модулю n равен s (если его случайное i равно 1) или равен ys (если $i = 2$).

5. Шаги 1–4 повторяются (с различными открытыми ключами (β_1, β_2)). Если Пикара выдержит T проверок, то Вивалес убедится (с вероятностью ошибки $1 - 2^{-T}$) в том, что Пикара действительно знает это разложение.

УПРАЖНЕНИЯ

1. Рассмотрим процедуру доказательства с нулевым разглашением знания одного дискретного логарифма. Пусть Пикара в действительности не знает значение дискретного логарифма. Какова вероятность того, что она сможет успешно обманывать Вивалеса при T-кратном повторении шагов 1–3 этой процедуры?

2. Опять рассмотрим процедуру доказательства с нулевым разглашением знания одного дискретного логарифма. Пусть Вивалес не знает величину N.

а) Объяснить, почему описанный в тексте протокол не обладает свойством нулевого разглашения.

б) Как Пикара может уменьшить количество информации о N, поступающей к Вивалесу?

3. Пусть Пикара не знает N и поэтому на шаге 1 выбирает случайное число $e < B$, где B — граница для возможного значения N. На шаге 3 она посылает просто число $x + e$ вместо наименьшего положительного вычета из $x + e$ по модулю N. Объяснить, почему это не будет доказательством с нулевым разглашением? Почему эта процедура, проделанная Клайдом, не будет правильно имитировать поведение Пикары?

4. Объяснить, как можно использовать приведенную в тексте процедуру доказательства с нулевым разглашением знания одного дискретного логарифма в системе электронной идентификации с открытым ключом. (Это означает, что Пикара убеждает Вивалеса в том, что она действительно Пикара.)

5. Объяснить, почему умение извлекать квадратные корни по модулю $n = pq$ по существу эквивалентно знанию разложения числа n на простые множители.

6. Может ли один и тот же открытый ключ (β_1, β_2) для скрытой передачи использоваться несколькими лицами для представления Вивалесу доказательств с нулевым разглашением того, что они все, независимо друг от друга, знают разложение на простые множители одного и того же числа? Считается, что все они имеют доступ ко всем передаваемым сообщениям.

7. Используя скрытую передачу, построить неинтерактивное доказательство с нулевым разглашением знания дискретного логарифма. (Предполагается, что порядок N группы известен всем.)

8. Недавно была предложена следующая схема протокола с нулевым разглашением, позволяющего Пикаре показать Вивалесу, что она знает делители p и q

целого числа n, о котором известно, что оно является произведением двух простых чисел, сравнимых с 3 по модулю 4. Найти основной изъян этой схемы.

Шаг 1. Вивалес, который знает n, но не знает p и q, выбирает случайно целое число x. Он вычисляет наименьший неотрицательный вычет числа x^4 по модулю n и посылает результат (который мы обозначим через y) Пикаре.

Шаг 2. Получив y, Пикара вычисляет квадратный корень в кольце вычетов по модулю n (что ей нетрудно сделать, так как она знает разложение n, см. упражнение 5 выше). Из четырех возможных квадратных корней она (единственным образом) выбирает тот, который является квадратичным вычетом как по модулю p, так и по модулю q. Это число должно быть наименьшим положительным вычетом из x^2 по модулю n. Она посылает его Вивалесу.

Шаг 3. Вивалес проверяет, что полученное число на самом деле является вычетом числа x^2 по модулю n. Тем самым он убеждается, что Пикара действительно извлекла квадратный корень по модулю n, что возможно лишь при знании разложения числа n.

9. Найти недостаток следующей процедуры доказательства с нулевым разглашением знания разложения. Пусть n является произведением двух простых p и q. Предположим, что «Центр доверия» вырабатывает бесконечную последовательность $y_1, y_2, \ldots$ квадратов случайных чисел (приведенных по модулю n), как это описано в тексте. Поочередно для каждого y_i Пикара находит один из его квадратных корней x_i и посылает его Вивалесу, который проверяет, что $x_i^2 \equiv y_i \pmod{n}$.

ЛИТЕРАТУРА к § 5 ГЛАВЫ IV

1. *Bellare M., Micali S.* Non-interactive oblivious transfer and applications. — In: Advances in Cryptology: Crypto'89. Heidelberg etc.: Springer, p. 547–557.
2. *Ben-Or M., Goldreich O., Goldwasser S., Håstad J., Kilian J., Micali S., Rogaway P.* Everything provable is provable in zero-knowledge. — In: Advances in Cryptology: Crypto'88. Heidelberg etc.: Springer, 1990, p. 37–56.
3. *Blum M., Feldman P., Micali S.* Non-interactive zero-knowledge proofs and their applications. — In: Proceedings of the 20th ACM Symposium on the Theory of Computing, 1988.
4. *Chaum D., Evertse J.-H., van de Graaf J., Peralta R.* Demonstrating possession of a discrete logarithm without revealing it. — In: Advances in Cryptology: Crypto'86. Heidelberg etc.: Springer, 1987, p. 200–212.
5. *Garey M. R., Johnson D. S.* Computers and intractability: A Guide to the Theory of NP-completeness. San-Francisco: W. H. Freeman, 1979. (Русск. перев. — *Гэри М., Джонсон Д. С.* Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982.)
6. *Goldwasser S., Micali S., Rackoff C.* The knowledge complexity of interactive proof systems. — SIAM J. Comput., 1989, v. 18, p. 186–208.
7. *Kilian J.* Founding cryptography on oblivious transfer. — In: Proceedings of the 20th ACM Symposium on the Theory of Computing, 1988, p. 20–31.
8. *Rabin M.* How to exchange secrets by obvilious transfer. — Techn. Report TR-81. Aiken Computation Laboratory, Harvard University, 1981.
9. *Shamir A.* The search for provably secure identification schemes. — In: Proceedings of the International Congress of Mathematicians, 1986, p. 1488–1495.

ГЛАВА V

ПРОСТОТА И ФАКТОРИЗАЦИЯ

Во многих случаях требуется выяснить, является ли большое число n простым. Например, в системе открытого ключа RSA и различных системах, основанных на задаче дискретного логарифмирования в конечных полях, нам нужно найти большое «случайное» простое число. Один из подходов к этому заключается в том, чтобы выбрать большое нечетное целое число n_0, используя генератор случайных чисел, и затем проверять $n_0, n_0 + 2, \ldots$ на простоту до тех пор, пока мы не найдем первое простое число, большее или равное n_0. Другой тип использования тестов на простоту — выяснение того, является ли некоторое число весьма специального типа простым. Например, нас может интересовать, является ли число $2^f - 1$ для данного большого числа f простым числом Мерсенна: в поле из 2^f элементов, как мы видели (см. упражнение 13 а) к § II. 1) все элементы, отличные от 0 и 1, являются образующими $\mathbf{F}_{2^f}^*$ в том и только том случае, когда $2^f - 1$ есть простое число.

Тест на простоту представляет собой критерий того, что число n *не* является простым. Если n «проходит» этот тест, то оно, *возможно*, простое число. Если оно «проходит» целый набор тестов на простоту, то весьма вероятно, что оно действительно является простым. С другой стороны, если n не проходит хотя бы одного теста на простоту, то оно совершенно определенно является составным. Однако при этом остается нерешенной трудная задача нахождения простых делителей числа n (задача факторизации). В общем случае для разложения на множители большого числа, о котором известно, что оно составное (поскольку оно не прошло теста на простоту), требуется гораздо больше времени, чем для нахождения простого числа того же порядка величины. (Это — эмпирическое утверждение, а не теорема; ни одного утверждения такого рода не доказано). Надежность криптосистемы RSA основывается на том предположении, что значительно

легче найти два чрезвычайно больших простых числа p и q, чем, зная $n = pq$, но не p или q, найти делители числа n. После рассмотрения тестов на простоту в § 1 мы опишем три различных метода факторизации в §§ 2–5.

§ 1. Псевдопростые числа

Замечали ли вы, что не делается попыток найти большие числа, которые *не являются* простыми? Было бы вам интересно услышать в выпуске новостей сообщение о том, что «Сегодня Отдел вычислительных наук в Вашингтонском университете объявил, что $2^{58,111,625,031} + 8$ — четное число. Это — самое большое не простое число, известное ныне»
— надпись в душевой комнате Вашингтонского университета.

«Явление, вероятность которого 10^{-50}, так никогда бы и не проявилось, или, во всяком случае, никогда бы не было замечено»
— Э. Борель *Вероятность и жизнь*.

Пусть n — большое нечетное число, и мы хотим определить, является ли n простым. Простейший тест на простоту — «пробы делением». Это значит, что вы берете нечетное целое число m и проверяете, делится ли n на m. Если $m \neq 1, m \neq n$ и $m|n$, то n — составное; в противном случае n проходит тест на простоту «деление на m». Если по мере того, как m пробегает одно за другим нечетные числа, начиная с 3, число n проходит все пробы делением, то становится все более и более вероятным, что n — простое число. Мы полностью убедимся в простоте n, если m достигнет значения $\sqrt{n}$. Конечно, это чрезвычайно неэффективный способ проверки простоты числа n. Другие тесты, описанные в этом параграфе, значительно быстрее.

Почти все известные эффективные тесты на простоту аналогичны в общих чертах следующему.

Если n — простое число, то согласно малой теореме Ферма для любого такого b, что НОД $(b, n) = 1$,

$$b^{n-1} \equiv 1 \pmod{n}. \tag{1}$$

Если n — *не* простое число, то (1) тоже может выполняться (хотя это маловероятно).

О п р е д е л е н и е. Если n — нечетное составное число, b — целое число, НОД $(n, b) = 1$, и (1) выполняется, то n называется *псевдопростым числом по основанию b*.

Другими словами, «псевдопростое» число — это число n, которое «претендует» быть простым, проходя тест (1).

Пример 1. Число $n = 91$ — псевдопростое по основанию $b = 3$, так как $3^{90} \equiv 1 \pmod{91}$. Однако, 91 *не* есть псевдопростое число по основанию 2, так как $2^{90} \equiv 64 \pmod{91}$. Если бы мы еще не знали, что 91 составное число, то соотношение $2^{90} \not\equiv 1 \pmod{91}$ доказало бы нам это.

Предложение V.1.1. *Пусть n — нечетное составное число.*

a) *Число n тогда и только тогда является псевдопростым по основанию b, для которого* НОД $(b, n) = 1$, *когда порядок b в группе $(\mathbf{Z}/n\mathbf{Z})^*$ (т. е. наименьшее натуральное d, при котором $b^d \equiv 1 \pmod{n}$) делит $n - 1$.*

b) *Если n является псевдопростым по основаниям b_1 и b_2, для которых* НОД $(b_1, n) =$ НОД $(b_2, n) = 1$, *то n — псевдопростое по основаниям $b_1 b_2$ и $b_1 b_2^{-1}$ (b_2^{-1} обозначает элемент, обратный к b_2 по модулю n).*

c) *Если n не является псевдопростым хотя бы по одному основанию $b \in (\mathbf{Z}/n\mathbf{Z})^*$, то n не проходит тест* (1), *по крайней мере, для половины оснований $b \in (\mathbf{Z}/n\mathbf{Z})^*$.*

Доказательство. Части a) и b) очевидны, и мы оставляем их проверку читателю.

c) Пусть $\{b_1, b_2, \ldots, b_s\}$ — множество всех оснований, по которым n — псевдопростое число, т. е. множество всех таких натуральных чисел, меньших n, для которых выполняется соотношение (1). Пусть b — фиксированное основание, по которому n не является псевдопростым. Если n было бы псевдопростым хотя бы по одному основанию bb_i, то, ввиду части b), оно было бы псевдопростым по основанию $b \equiv (bb_i)b_i^{-1} \pmod{n}$, что противоречит выбору b. Таким образом, для s различных оснований $\{bb_1, bb_2, \ldots, bb_s\}$ число n не проходит тест (1). Поэтому в группе $(\mathbf{Z}/n\mathbf{Z})^*$ число оснований, по которым n не есть псевдопростое, не меньше числа оснований, для которых выполняется (1). Доказательство завершено.

Таким образом, если n удовлетворяет условию (1) не для *всех* возможных b с НОД $(b, n) = 1$, то при случайно выбранном b с вероятностью, не меньшей 50%, условие (1) для n не выполняется. Допустим теперь, что мы хотим выяснить, является ли данное большое нечетное число n простым. Мы можем случайно выбрать b в интервале $0 < b < n$. Вначале находим $d =$ НОД (b, n), применяя алгоритм Евклида. Если $d > 1$, то n — не простое число, и мы фактически уже нашли его нетривиальный делитель $d|n$. Если $d = 1$, возводим b в степень $n - 1$ (используя метод повторного возведения в квадрат, см. § I.3). Если (1) не выполняется, то n —составное число. Если (1) выполняется, то мы получаем некоторое подтверждение того, что n —

простое число. Мы выбираем тогда другое b и повторяем весь процесс. Если (1) нарушается для какого-либо b, то можно остановиться, убедившись в том, что n — составное число. Предположим, что мы сделали k попыток с различными b и нашли, что n — псевдопростое число по всем k основаниям. По предложению V. 1. 1 вероятность того, что составное n пройдет k тестов, не превышает $1/2^k$ (*если только n не таково*, что (1) выполняется для каждого $b \in (\mathbf{Z}/n\mathbf{Z})^*$). Если k велико, мы можем быть уверены «с большой вероятностью», что n — простое (если только n не является псевдопростым по всем основаниям). Этот метод нахождения простых чисел называется *вероятностным* методом. Он отличается от *детерминистического* метода: слово «детерминистический» значит, что метод со 100%-й вероятностью определяет, является число n составным или простым.

Существуют ли такие составные n, что (1) выполняется для каждого b, взаимно простого с n? В таком случае вероятностный метод не может обнаружить, что n — составное (если только нам не повезет выбрать такое b, что НОД $(b, n) > 1$). Ответ на вопрос известен: такие числа существуют и называются числами Кармайкла.

О п р е д е л е н и е. *Числом Кармайкла* называется составное число n, для которого (1) выполняется при любом $b \in (\mathbf{Z}/n\mathbf{Z})^*$.

Предложение V. 1. 2. *Пусть n — нечетное составное число.*

а) *Если n делится на квадрат целого числа, большего* 1, *то n не есть число Кармайкла.*

b) *Если n свободно от квадратов, то n — число Кармайкла в том и только том случае, когда $(p-1)|(n-1)$ для всякого простого делителя p числа n.*

Д о к а з а т е л ь с т в о. а) Предположим, что $p^2|n$. Пусть g — образующий по модулю p^2, т. е. такое целое число, что $g^{p(p-1)}$ есть наименьшая степень g, сравнимая с 1 по модулю p^2. Согласно упражнению 2 к § II. 1 такое g всегда существует. Пусть n' обозначает произведение всех простых делителей числа n, отличных от p. По китайской теореме об остатках существует решение системы из двух сравнений: $b \equiv g \pmod{p^2}$ и $b \equiv 1 \pmod{n'}$. Тогда b, как и g, есть образующий по модулю p^2, и, кроме того, НОД $(b, n) = 1$, так как b не делится ни на p, ни на простые делители числа n'.

Покажем, что n не есть псевдопростое число по основанию b. Действительно, заметим, что если (1) выполняется, то автоматически $b^{n-1} \equiv 1 \pmod{p^2}$, так как $p^2|n$. Но в этом случае $p(p-1)|(n-1)$, так как $p(p-1)$ — порядок b по модулю p^2. Однако $n - 1 \equiv -1 \pmod{p}$, так как $p|n$, а это означает, что $n - 1$ не делится на $p(p-1)$. Это противоречие доказывает, что b — такое основание, по которому n не

является псевдопростым.

b) Предположим сначала, что $(p-1)|(n-1)$ для каждого простого делителя p числа n. Пусть b — любое такое основание, что НОД $(b,n)=1$. Тогда для всякого простого $p|n$ мы имеем: b^{n-1} есть степень b^{p-1} и, значит, $b^{n-1}\equiv 1 \pmod p$. Таким образом, $b^{n-1}-1$ делится на каждый из простых делителей числа n, следовательно, делится и на их произведение, которое равно n. Поэтому (1) выполняется для всех оснований b. Обратно, предположим, что существует такое простое число $p|n$, что $n-1$ не делится на $p-1$. Пусть g — целое число, порождающее $(\mathbf{Z}/p\mathbf{Z})^*$. Как в доказательстве части a), найдем такое целое число b, что $b\equiv g \pmod p$, $b\equiv 1 \pmod{n/p}$. Тогда НОД $(b,n)=1$ и $b^{n-1}\equiv g^{n-1} \pmod p$. Однако $g^{n-1}\not\equiv 1 \pmod p$, так как $n-1$ не делится на порядок $p-1$ числа g по модулю p. Стало быть, $b^{n-1}\not\equiv 1 \pmod p$ и (1) не может выполняться. Доказательство предложения завершено.

Пример 2. Число $n=561=3\cdot 11\cdot 17$ есть число Кармайкла, так как $n-1=560$ делится на $3-1=2$, $11-1=10$, $17-1=16$. В упражнениях мы увидим, что 561 — наименьшее число Кармайкла.

Предложение V.1.3. *Число Кармайкла должно быть произведением не менее* 3 *различных простых чисел.*

Доказательство. Согласно предложению V.1.2 число Кармайкла должно быть произведением различных простых чисел. Следовательно, остается лишь исключить возможность того, что $n=pq$ есть произведение двух различных простых.

Пусть $p<q$. Тогда, если бы n было числом Кармайкла, мы имели бы $n-1\equiv 0 \pmod{q-1}$ по части b) предложения V.1.2. Однако $n-1=p(q-1+1)-1\equiv p-1 \pmod{q-1}$, и так как $0<p-1<q-1$, то правая часть не сравнима с нулем по модулю $q-1$. Полученное противоречие с частью b) предложения V.1.2 завершает доказательство.

Замечание. Лишь недавно Альфорд, Грэнвиль и Померанц доказали, что существует бесконечно много чисел Кармайкла. См. отчет Грэнвиля в Notices of the Amer. Math. Soc., 1992, т. 39, с. 696–700.

Эйлеровы псевдопростые. Пусть n — нечетное целое число и $\left(\frac{b}{n}\right)$ обозначает символ Якоби (см. § II.2). Если n — простое число, то по предложению II.2.2

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

для любого целого b. С другой стороны, если n — составное, то согласно упражнению 21 к § II.2 не менее 50% всех $b\in(\mathbf{Z}/n\mathbf{Z})^*$ не удо-

влетворяют (2). Эти два факта позволяют построить эффективный вероятностный тест для проверки простоты большого числа n.

Начнем со следующего определения.

О п р е д е л е н и е. Если n — нечетное составное число, а b — такое целое число, что НОД $(b, n) = 1$ и выполняется (2), то n называется *эйлеровым псевдопростым числом по основанию* b.

Предложение V.1.4. *Если n — эйлерово псевдопростое число по основанию b, то оно — псевдопростое по основанию b.*

Д о к а з а т е л ь с т в о. Мы должны доказать, что если выполнено (2), то выполнено и (1). Но это очевидно: достаточно возвести обе части (2) в квадрат.

П р и м е р 3. Предложение, обратное к V.1.4, неверно. Так, в примере 1 мы видели, что 91 — псевдопростое по основанию 3. Однако $3^{45} \equiv 27 \pmod{91}$, так что (2) не выполняется для $n = 91$, $b = 3$. (Заметим, что знание порядка b в $(\mathbf{Z}/91\mathbf{Z})^*$ упрощает возведение b в высокую степень по модулю 91: так как $3^6 \equiv 1 \pmod{91}$, то $3^{45} \equiv 3^3 \pmod{91}$.) Число 10 дает пример основания, по которому 91 — эйлерово псевдопростое число: $10^{45} \equiv 10^3 \equiv -1 \pmod{91}$ и $\left(\frac{10}{91}\right) = -1$.

П р и м е р 4. Легко убедиться, что любое нечетное составное n является псевдопростым по основаниям ± 1; в дальнейшем мы будем исключать эти «тривиальные» основания b.

Опишем теперь **тест Соловея–Штрассена**. Пусть дано нечетное натуральное число и мы хотим знать, простое оно или составное. Выбираем случайно k натуральных чисел b, меньших n. Для каждого из них вычисляем значения обеих частей (2). Вычисление $b^{(n-1)/2}$ проводится за $O(\log^3 n)$ двоичных операций при использовании метода повторного возведения в квадрат (предложение I.3.6); вычисление символа Якоби в правой части также проводится за $O(\log^3 n)$ двоичных операций (см. упражнение 17 к § II.2). Если эти два значения не сравнимы по модулю n, то мы убедились, что n — составное число, и прекращаем тест. В противном случае переходим к следующему b. Вероятность того, что при составном n сравнение (2) выполняется для всех k случайно выбранных b, не превышает $1/2^k$. Таким образом, тест Соловея–Штрассена — вероятностный алгоритм, который приводит либо к выводу о том, что n — составное число, либо к выводу о том, что n — «вероятно» простое.

Отметим, что не существует эйлеровых псевдопростых чисел, аналогичных псевдопростым числам Кармайкла: для *любого* составного n критерий (2) не выполняется, по меньшей мере, для половины всех возможных b.

Сильно псевдопростые. Рассмотрим теперь еще один вид критериев простоты, который в определенном смысле даже лучше теста Соловея–Штрассена, основанного на определении псевдопростоты по Эйлеру. Это — тест Миллера–Рабина, основанный на вводимом ниже понятии «сильной псевдопростоты». Предположим, что n — большое нечетное натуральное число и $b \in (\mathbf{Z}/n\mathbf{Z})^*$. Пусть, далее, n — псевдопростое по основанию b, т. е. $b^{n-1} \equiv 1 \pmod{n}$. Идея критерия сильной псевдопростоты такова. Пусть $n-1 = 2^s t$, t — нечетно. Если последовательно вычислять $b^{(n-1)/2}$, $b^{(n-1)/4}, \ldots, b^{(n-1)/2^s}$, то при простом n первым элементом, отличным от 1, должен быть элемент -1, так как при простом n единственными решениями сравнения $X^2 - 1 \equiv 0 \pmod{n}$ являются $+1$ и -1. Практически действия выполняются «в обратном направлении». Полагаем $n-1 = 2^s t$, t — нечетно. Вычисляем b^t по модулю n. Если $b^t \not\equiv 1 \pmod{n}$, то возводим b^t в квадрат по модулю n, получая $b^{2t} \pmod{n}$, затем вновь возводим в квадрат и т. д. до тех пор, пока не получим 1: $b^{2^r t} \equiv 1 \pmod{n}$. Тогда, если n — простое, предыдущим числом должно быть -1, в противном случае мы получаем доказательство того, что n составное.

Определение. Пусть n — нечетное составное число и $n - 1 = 2^s t$, t — нечетно. Пусть $b \in (\mathbf{Z}/n\mathbf{Z})^*$. Если n и b удовлетворяют одному из условий:

1) $b^t \equiv 1 \pmod{n}$;

2) существует такое r, $0 \leqslant r < s$, что

$$b^{2^r t} \equiv -1 \pmod{n}, \tag{3}$$

то n называется *сильно псевдопростым* по основанию b.

Предложение V.1.5. *Если $n \equiv 3 \pmod{4}$, то n — сильно псевдопростое число по основанию b тогда и только тогда, когда оно — эйлерово псевдопростое по основанию b.*

Доказательство. Так как в этом случае $s = 1$ и $t = (n-1)/2$, мы видим, что n — сильно псевдопростое по основанию b тогда и только тогда, когда $b^{(n-1)/2} \equiv \pm 1 \pmod{n}$. Если n —эйлерово псевдопростое, то это условие выполнено по определению. Обратно, пусть $b^{(n-1)/2} \equiv \pm 1 \pmod{n}$. Мы должны показать, что ± 1 справа — это $\left(\frac{b}{n}\right)$. Но если $n \equiv 3 \pmod{4}$, то $\pm 1 = \left(\frac{\pm 1}{n}\right)$. Таким образом,

$$\left(\frac{b}{n}\right) = \left(\frac{b(b^2)^{(n-3)/4}}{n}\right) = \left(\frac{b^{(n-1)/2}}{n}\right) \equiv b^{(n-1)/2} \pmod{n},$$

что и требовалось.

Следующие два важных предложения несколько труднее для доказательства.

Предложение V.1.6. *Если n — сильно псевдопростое по основанию b, то оно также эйлерово псевдопростое по основанию b.*

Предложение V.1.7. *Если n — нечетное составное целое число, то n — сильно псевдопростое по основанию b не более чем для 25% чисел b, $0 < b < n$.*

Замечание. Как мы увидим в упражнениях, в общем случае обратное утверждение к предложению V.1.6 неверно.

Перед доказательством этих двух предложений мы опишем **тест Миллера–Рабина** на простоту. Предположим, что мы хотим определить, является большое натуральное число n простым или составным. Представим $n-1$ в виде $2^s t$, t нечетно, и выберем случайное целое число b, $0 < b < n$. Сначала вычисляем b^t по модулю n. Если получается ± 1, то заключаем, что n прошло тест (3) при данном b, и производим новый случайный выбор b. В противном случае возводим b^t в квадрат по модулю n, результат вновь возводим в квадрат по модулю n и продолжаем так до тех пор, пока не получим -1. Если это происходит, то мы считаем, что n прошло тест. Если же этого не происходит, т.е. если мы получаем $b^{2^{r+1}} \equiv 1 \pmod{n}$, в то время как $b^{2^r} \not\equiv -1 \pmod{n}$, то n не проходит тест, и это доказывает, что n — составное число. Если мы k раз случайно выбирали разные основания b и n каждый раз проходило соответствующий тест, то, согласно предложению V.1.7, число n имеет не более одного шанса из 4^k быть составным. Это следует из того факта, что если n — составное, то не более $1/4$ всех возможных оснований b, $0 < b < n$, удовлетворяет (3). Заметим, что это несколько лучше, чем в тесте Соловея–Штрассена, где аналогичная оценка шансов — 1 из 2^k (мы увидим в одном из упражнений, что существуют составные числа n, являющиеся псевдопростыми эйлеровыми относительно половины всех возможных оснований).

Переходим к доказательствам предложений V.1.6 и V.1.7.

Доказательство предложения V.1.6. Пусть n и b удовлетворяют (3). Докажем, что они удовлетворяют (2). Полагаем $n-1 = 2^s t$, t нечетное.

Случай 1. Пусть $b^t \equiv 1 \pmod{n}$. Тогда левая часть (2) есть, очевидно, 1. Нам нужно показать, что $\left(\frac{b}{n}\right) = 1$. Но $1 = \left(\frac{1}{n}\right) = \left(\frac{b^t}{n}\right) = \left(\frac{b}{n}\right)^t$. Так как t нечетно, то $\left(\frac{b}{n}\right) = 1$.

Случай 2. Теперь предположим, что $b^{(n-1)/2} \equiv -1 \pmod{n}$. Мы должны показать, что $\left(\frac{b}{n}\right) = -1$. Пусть p — любой простой делитель n. Представим $p-1$ в виде $2^{s'} t'$, t' нечетно.

Утверждение. *Справедливы соотношения* $s' \geqslant s$ *и*

$$\left(\frac{b}{p}\right) = \begin{cases} -1, & \text{если } s' = s, \\ 1, & \text{если } s' > s. \end{cases}$$

Доказательство утверждения. Так как $b^{(n-1)/2} = b^{2^{s-1}t} \equiv -1 \pmod{n}$, то, возведя обе части в степень t', получаем $(b^{2^{s-1}t'})^t \equiv -1 \pmod{n}$. Так как $p|n$, это сравнение справедливо и по модулю p. В случае $s' < s$ мы имели бы $b^{p-1} = b^{2^{s'}t'} \not\equiv 1 \pmod{p}$, что противоречит малой теореме Ферма. Поэтому $s' \geqslant s$. Если $s' = s$, то $(b^{2^{s'-1}t'})^t \equiv -1 \pmod{p}$. Поэтому $\left(\frac{b}{p}\right) \equiv b^{(p-1)/2} = b^{2^{s'-1}t'} \equiv -1 \pmod{p}$. Пусть теперь $s' > s$. Возведя обе части сравнения $b^{2^{s-1}t'} \equiv -1 \pmod{p}$ в степень $2^{s'-s}$, получаем $b^{2^{s'-1}t'} \equiv b^{(p-1)/2} \equiv \left(\frac{b}{p}\right) \equiv 1 \pmod{p}$. Утверждение доказано.

Вернемся к доказательству случая 2 предложения V.1.6. Запишем n в виде произведения простых чисел (*не* обязательно различных): $n = \prod p$. Представим также каждое $p-1$ в виде $p-1 = 2^{s'}t'$, t' нечетно. Как было показано, $s' \geqslant s$ и $\left(\frac{b}{p}\right) = -1$ лишь при $s' = s$. Число таких случаев обозначим через k (учитывая кратность, т.е. k со свойством $s' = s$ дает вклад α, если n делится на p^α и не делится на $p^{\alpha+1}$). Тогда $\left(\frac{b}{n}\right) = \prod\left(\frac{b}{p}\right) = (-1)^k$. Далее, если p — такой делитель n, что $s' > s$, то $p \equiv 1 \pmod{2^{s+1}}$, а для остальных k делителей p числа n имеем $s' = s$ и $p \equiv 1 + t'2^s \equiv 1 + 2^s \pmod{2^{s+1}}$. Так как $n = 1 + 2^s t \equiv 1 + 2^s \pmod{2^{s+1}}$, то $n \equiv 1 + 2^s \equiv \prod p \equiv (1+2^s)^k \equiv 1 + k2^s \pmod{2^{s+1}}$ (последнее — по формуле бинома Ньютона). Это показывает, что k должно быть нечетным, т.е. $\left(\frac{b}{n}\right) = (-1)^k = -1$, что и требовалось доказать.

Случай 3. Предположим, наконец, что $b^{2^{r-1}t} \equiv -1 \pmod{n}$ для некоторого r, $0 < r < s$ (используем $r-1$ вместо r в (3)). Имеем тогда $b^{(n-1)/2} \equiv 1 \pmod{n}$, и нам нужно показать, что $\left(\frac{b}{n}\right) = 1$ в случае 3.

Пусть опять p — простой делитель n, $p - 1 = 2^{s'}t'$, t' нечетно.

Утверждение. *Справедливы соотношения* $s' \geqslant r$ *и*

$$\left(\frac{b}{p}\right) = \begin{cases} -1, & \text{если } s' = r, \\ 1, & \text{если } s' > r. \end{cases}$$

Доказательство этого утверждения идентично доказательству соответствующего утверждения в случае 2.

Для доказательства предложения в случае 3 обозначим через k число простых p (не обязательно различных) в произведении $n = \prod p$,

для которых выполнена первая альтернатива, т. е. $s' = r$. Тогда, как и в случае 2, мы, очевидно, имеем $\left(\frac{b}{n}\right) = (-1)^k$. С другой стороны, так как $n = 1 + 2^s t \equiv 1 \pmod{2^{r+1}}$ и $n = \prod p \equiv (1+2^r)^k \pmod{2^{r+1}}$, то k должно быть четным, т. е. $\left(\frac{b}{n}\right) = 1$. Это завершает доказательство предложения V. 1. 6.

Перед доказательством предложения V. 1. 7 мы докажем одну лемму общего характера о числе решений уравнения $x^k = 1$ в «циклической группе», содержащей m элементов. Мы уже однажды встречались с этой леммой в начале § II. 2; данное ниже доказательство следует сравнить с доказательством предложения II. 2. 1.

Лемма 1. *Пусть* $d = \text{НОД}\,(k, m)$. *В группе* $\{g, g^2, g^3, \ldots, g^m = 1\}$ *порядка* m *имеется ровно* d *элементов, удовлетворяющих уравнению* $x^k = 1$.

Доказательство. Элемент g^j удовлетворяет этому уравнению тогда и только тогда, когда $g^{jk} = 1$, т. е. в том и только том случае, когда $jk \equiv 0 \pmod{m}$, т. е. $m | jk$. Это эквивалентно условию $\frac{m}{d} | j\frac{k}{d}$. Но m/d и k/d взаимно просты, поэтому j кратно m/d. Среди чисел $\{1, 2, \ldots, m\}$ имеется в точности d таких значений j.

Нам нужна еще одна лемма, которая доказывается аналогично предыдущей.

Лемма 2. *Пусть* p — *нечетное простое число,* $p - 1 = 2^{s'} t'$, *где* t' *нечетно. Тогда число* $x \in (\mathbf{Z}/p\mathbf{Z})^*$, *удовлетворяющих уравнению* $x^{2^r t} \equiv -1 \pmod{p}$ (t *нечетно*), *равно нулю, если* $r \geqslant s'$, *и равно* $2^r \text{НОД}\,(t, t')$, *если* $r < s'$.

Доказательство. Пусть g — образующий элемент группы $(\mathbf{Z}/p\mathbf{Z})^*$ и $x = g^j$, $0 \leqslant j < p - 1$. Так как $g^{(p-1)/2} \equiv -1 \pmod{p}$ и $p - 1 = 2^{s'} t'$, то сравнение из условия леммы эквивалентно сравнению $2^r t j \equiv 2^{s'-1} t' \pmod{2^{s'} t'}$ (с неизвестным j). Ясно, что при $r > s' - 1$ оно не имеет решений. Если же $r \leqslant s' - 1$, то делим обе части сравнения и модуль на НОД модуля и коэффициента при j, т. е. на $2^r d$, где $d = \text{НОД}\,(t, t')$. Получившееся сравнение по модулю $2^{s'-r}\left(\frac{t'}{d}\right)$ имеет единственное решение. Поэтому исходное сравнение имеет $2^r d$ решений, как и утверждалось. Лемма доказана.

Доказательство предложения V. 1. 7.

Случай 1. Предположим сначала, что n делится на квадрат простого числа p, т. е. $p^\alpha || n$, $\alpha \geqslant 2$. Покажем, что в этом случае n не может быть даже псевдопростым (не говоря уже о сильной псевдопростоте) для более, чем $(n-1)/4$ оснований b, $0 < b < n$. Для этого предположим, от противного, что $b^{n-1} \equiv 1 \pmod{n}$, и, значит, $b^{n-1} \equiv 1 \pmod{p^2}$, и найдем условия, которым должен удовлетворять

вычет b по модулю p^2. Напомним, что $(\mathbf{Z}/p^2\mathbf{Z})^*$ — циклическая группа порядка $p(p-1)$ (см. упражнение 2 к § II. 1), т. е. существует такое целое число g, что $(\mathbf{Z}/p^2\mathbf{Z})^* = \{g, g^2, g^3, \ldots, g^{p(p-1)}\}$. Согласно лемме 1, число решений сравнения $b^{n-1} \equiv 1 \pmod{p^2}$ в $(\mathbf{Z}/p^2\mathbf{Z})^*$ равно $d = \text{НОД}\,(p(p-1), n-1)$. Но p — делитель n, поэтому p не делит $n-1$ и, следовательно, p не делит d. Стало быть, $d \leqslant p-1$. Поэтому доля не делящихся на p^2 чисел в интервале от 0 до n, обладающих свойством $b^{n-1} \equiv 1 \pmod{p^2}$, не превышает $\frac{p-1}{p^2-1} = \frac{1}{p+1} \leqslant \frac{1}{4}$. Так как эта величина оценивает сверху долю таких оснований b, что $0 < b < n$ и $b^{n-1} \equiv 1 \pmod{n}$, то n может быть псевдопростым не больше, чем по четверти оснований b, $0 < b < n$. В случае 1 предложение доказано.

Замечание. Эта верхняя 25%-я граница достигается в случае 1 при $n = 9$ (т. е. 9 есть сильное псевдопростое для двух из 8 возможных значений b, именно, для $b = \pm 1$).

Случай 2. Теперь предположим, что n — произведение двух различных простых чисел p и q: $n = pq$*. Пусть $p - 1 = 2^{s'}t'$, t' нечетно, $q - 1 = 2^{s''}t''$, t'' нечетно. Без ущерба для общности можем предположить, что $s' \leqslant s''$. Чтобы элемент $b \in (\mathbf{Z}/n\mathbf{Z})^*$ мог служить основанием, по которому n является сильно псевдопростым, должно случиться одно из двух: 1) $b^t \equiv 1 \pmod{p}$ и $b^t \equiv 1 \pmod{q}$ или 2) $b^{2^r t} \equiv -1 \pmod{p}$ и $b^{2^r t} \equiv -1 \pmod{q}$ для некоторого r, $0 \leqslant r < s$. По лемме 1 число оснований b, для которых реализуется первая возможность, равно $\text{НОД}\,(t, t') \cdot \text{НОД}\,(t, t'')$ (произведению чисел классов вычетов по модулям p и q); ясно, что это число не превосходит $t't''$. По лемме 2 для каждого $r < \min(s', s'') = s'$ число таких b, что $b^{2^r t} \equiv -1 \pmod{n}$, равно $2^r\,\text{НОД}\,(t, t') \cdot 2^r\,\text{НОД}\,(t, t'') < 4^r t't''$. Так как $n - 1 > \varphi(n) = 2^{s'+s''}t't''$, то доля оснований $b \in (\mathbf{Z}/n\mathbf{Z})^*$, по которым n — сильно псевдопростое, не превышает

$$\frac{t't'' + t't'' + 4t't'' + 4^2t't'' + \cdots + 4^{s'-1}t't''}{2^{s'+s''}t't''} = 2^{-s'-s''}\left(1 + \frac{4^{s'}-1}{4-1}\right).$$

Если $s'' > s'$, то это не превосходит $2^{-2s'-1}(\frac{2}{3} + \frac{4^{s'}}{3}) < 2^{-3}\frac{2}{3} + \frac{1}{6} = \frac{1}{4}$, что и требовалось доказать. С другой стороны, если $s' = s''$, то одно из неравенств $\text{НОД}\,(t, t') \leqslant t'$, $\text{НОД}\,(t, t'') \leqslant t''$ должно быть строгим: если бы мы имели одновременно $t' | t$ и $t'' | t$, то из сравнения $n - 1 = 2^s t = pq - 1 \equiv q - 1 \pmod{t'}$ следовало бы, что $t' | (q-1) = 2^{s''}t''$, т. е. $t' | t''$; аналогично $t'' | t'$ и, значит, $t' = t''$. Но тогда $p = q$, вопреки

---

* В дальнейшем предполагается, что $n - 1 = 2^s t$, t нечетно. — *Прим. ред.*

предположению. В строгом неравенстве вида НОД $(t, t^*) < t^*$ левая часть составляет не более $1/3$ от правой (так как все числа в этом неравенстве — нечетные). Таким образом, в нашем случае можно заменить $t't''$ на $\frac{1}{3}t't''$ в числителе оценки для числа оснований b, по которым n является сильно псевдопростым. Это приводит к следующей оценке сверху доли тех оснований b, $0 < b < n$, по которым n — сильно псевдопростое:

$$\frac{1}{3}2^{-2s'}\left(\frac{2}{3}+\frac{4^{s'}}{3}\right) \leqslant \frac{1}{18}+\frac{1}{9}=\frac{1}{6}<\frac{1}{4},$$

что и требовалось. Это завершает доказательство предложения в случае 2.

С л у ч а й 3. Предположим, наконец, что n — произведение более, чем двух различных простых чисел: $n = p_1 p_2 \cdots p_k$, $k \geqslant 3$. Положим $p_j - 1 = 2^{s_j} t_j$, где t_j нечетно, и будем действовать в точности так же, как в случае 2. Без ущерба для общности можно предполагать, что $s_1 \leqslant s_j$, $j = 2, \ldots, k$. Мы получаем следующую верхнюю оценку доли возможных оснований b, по которым n — сильно псевдопростое число:

$$2^{-s_1-s_2-\cdots-s_k}\left(1+\frac{2^{ks_1}-1}{2^k-1}\right) \leqslant 2^{-ks_1}\left(\frac{2^k-2}{2^k-1}+\frac{2^{ks_1}}{2^k-1}\right)$$

$$= 2^{-ks_1}\frac{2^k-2}{2^k-1}+\frac{1}{2^k-1} \leqslant 2^{-k}\frac{2^k-2}{2^k-1}+\frac{1}{2^k-1}=2^{1-k}\leqslant\frac{1}{4},$$

так как $k \geqslant 3$ в случае 3. Тем самым доказательство предложения V.1.7 закончено.

З а м е ч а н и я. 1. Практически не возникает необходимости брать большое количество оснований b, чтобы быть «почти» уверенным, что число n простое, если оно сильно псевдопростое по каждому из оснований b. Например, было показано, что существует лишь одно составное число, меньшее $2{,}5 \cdot 10^{10}$, а именно, 3215031751, которое является сильно псевдопростым по каждому из оснований $b = 2, 3, 5, 7$.

2. Полностью полагаться на вероятностный тест не вполне надежно. Несмотря на заверение Эмиля Бореля, цитированное в начале параграфа, было бы неплохо иметь быстрые методы *доказательства* того, что данное число n действительно простое (особенно, если важно с практической или теоретической точки зрения знать, что данное n действительно простое). Например, допустим, что существует такое небольшое число B (зависящее от n), что любое составное n не является сильно псевдопростым хотя бы по одному основанию $b < B$.

Если бы мы знали такое B, то для абсолютной уверенности в простоте n было бы достаточно провести тест (3) для первых B оснований.

Подобное утверждение имеет место, однако оно является следствием недоказанного предположения, называемого «обобщенной гипотезой Римана». Обычная гипотеза Римана утверждает, что все комплексные нули так называемой дзета-функции Римана $\zeta(s)$ (определяемой при $s > 1$ как сумма дробей $1/n^s$), которые лежат в «критической полосе» (где действительная часть s находится между 0 и 1), должны лежать на «критической прямой» (действительная часть s равна $\frac{1}{2}$). «Обобщенная гипотеза Римана» — это то же самое утверждение для некоторых обобщений $\zeta(s)$, называемых «L-рядами Дирихле». Следующее предложение, доказательство которого выходит за рамки этой книги, показывает, что условие (3) в критерии Миллера–Рабина дает *детерминистический* тест на простоту, временная сложность которого является полиномиальной функцией от $\log n$, если справедлива обобщенная гипотеза Римана (ОГР).

Если ОГР *справедлива и* n *— составное нечетное число, то* n *не удовлетворяет условию* (3) *хотя бы для одного основания* b, *меньшего* $2\log^2 n$.

3. В 1980-е гг. был разработан эффективный детерминистический тест на простоту, который хотя и не является, строго говоря, полиномиальным по $\log n$, на практике может доказать простоту чисел, имеющих свыше ста цифр в десятичной записи, в течение каких-то секунд (на современных больших компьютерах). Этот метод Адлемана–Померанца–Рамли (Adleman–Pomerance–Rumely) и Коэна–Ленстры (Cohen–Lenstra) основывается на тех же идеях, что и рассмотренные выше тесты на простоту, с той разницей, что он использует аналоги малой теоремы Ферма в расширениях полей рациональных чисел. Основную роль играют гауссовы суммы (некоторые их типы были введены в § II.2 для доказательства закона квадратичной взаимности) и тесно с ними связанные «суммы Якоби». Подробное рассмотрение их метода завело бы нас слишком далеко. Обстоятельное и не слишком сложное изложение имеется в статье Коэна и Ленстры в «Mathematics of Computation».

УПРАЖНЕНИЯ

1. а) **Найти все основания b, для которых 15 — псевдопростое число (не включать тривиальные основания ± 1).**

б) **Найти все основания, для которых 21 — псевдопростое число.**

в) **Доказать, что существует 36 оснований $b \in (\mathbf{Z}/91\mathbf{Z})^*$ (т. е. половина всех возможных оснований), для которых 91 — псевдопростое число.**

г) **Обобщая часть в), показать, что если p и $2p-1$ — простые числа и $n =$**

$p(2p-1)$, то n — псевдопростое для 50% возможных оснований b, а именно, для тех b, которые являются квадратичными вычетами по модулю $2p-1$.

2. Пусть n — нечетное составное натуральное число и пусть НОД $(b, n) = 1$.

а) Показать, что если p — простой делитель n и $n' = n/p$, то n — псевдопростое по основанию b только при $b^{n'-1} \equiv 1 \pmod{p}$.

б) Доказать, что никакое целое число $n = 3p$ ($p > 3$ есть простое число) не может быть псевдопростым по основаниям 2, 5 или 7.

в) Доказать, что никакое целое $n = 5p$ ($p > 5$ есть простое число) не может быть псевдопростым по основаниям 2, 3 или 7.

г) Доказать, что 91 — наименьшее псевдопростое число по основанию 3.

3. Пусть p — простое число. Показать, что p^2 — псевдопростое по основанию b тогда и только тогда, когда $b^{p-1} \equiv 1 \pmod{p^2}$.

4. а) Найти наименьшее псевдопростое по основанию 5.

б) Найти наименьшее псевдопростое по основанию 2.

5. Пусть $n = pq$ есть произведение двух различных простых чисел.

а) Пусть $d =$ НОД $(p-1, q-1)$. Доказать, что n — псевдопростое по основанию b в том и только том случае, когда $b^d \equiv 1 \pmod{n}$. Выразить через d число различных оснований, по которым n — псевдопростое.

б) Сколько имеется оснований, по которым n — псевдопростое, если $q = 2p+1$? Привести их полный список (в терминах p).

в) Пусть $n = 341$. Какова вероятность того, что случайно выбранное b, взаимно простое с n, будет основанием, по которому n — псевдопростое число?

6. Показать, что если n — псевдопростое по основанию $b \in (\mathbf{Z}/n\mathbf{Z})^*$, то оно будет также псевдопростым по основаниям $-b$ и b^{-1}.

7. а) Доказать, что если n — псевдопростое по основанию 2, то $N = 2^n - 1$ — также псевдопростое по тому же основанию.

б) Доказать, что если n — псевдопростое по основанию b и НОД $(b-1, n) = 1$, то целое число $N = (b^n - 1)/(b-1)$ — псевдопростое по основанию b.

в) Доказать, что существует бесконечно много псевдопростых чисел по основаниям 2, 3, 5.

г) Привести пример, показывающий, что без условия НОД $(b-1, n) = 1$ часть б) неверна.

8. Пусть $b > 1$ есть любое натуральное число, p — нечетное простое, не делящее $b-1$, b, $b+1$. Положим $n = (b^{2p} - 1)/(b^2 - 1)$.

а) Показать, что n — составное.

б) Показать, что $2p|(n-1)$.

в) Показать, что n — псевдопростое по основанию b; вывести отсюда, что для любого основания b существует бесконечно много псевдопростых по этому основанию.

9. а) Использовать тест (1), чтобы показать, что число $2047 = 2^{11} - 1$ — составное.

б) Объяснить, почему никогда не следует пытаться использовать тест (1) с $b = 2$ для проверки на простоту чисел Ферма $2^{2^k} + 1$ или чисел Мерсенна $2^p - 1$. Как обстоит дело с тестом (2) по основанию $b = 2$? Как обстоит дело с тестом (3) по основанию $b = 2$?

10. Пусть m — такое натуральное число, что $6m+1$, $12m+1$, $18m+1$ — простые числа. Положим $n = (6m+1)(12m+1)(18m+1)$. Доказать, что n — число Кармайкла.

Замечание. Неизвестно, существует ли бесконечно много чисел Кармайкла вида $(6m+1)(12m+1)(18m+1)$. Эвристические соображения, однако, показывают, что это так.

11. Показать, что следующие натуральные числа — это числа Кармайкла: $1105 = 5 \cdot 13 \cdot 17$; $1729 = 7 \cdot 13 \cdot 19$; $2465 = 5 \cdot 17 \cdot 29$; $2821 = 7 \cdot 13 \cdot 31$; $6601 = 7 \cdot 23 \cdot 41$; $29341 = 13 \cdot 37 \cdot 61$; $172081 = 7 \cdot 13 \cdot 31 \cdot 61$; $278545 = 5 \cdot 17 \cdot 29 \cdot 113$.

12. а) Найти все числа Кармайкла вида $3pq$ (p, q — простые).

б) Найти все числа Кармайкла вида $5pq$ (p, q — простые).

в) Доказать, что для любого фиксированного простого r имеется лишь конечное множество чисел Кармайкла вида rpq (p, q — простые).

13. Доказать, что 561 — наименьшее число Кармайкла.

14. Привести пример составного числа n и основания b таких, что $b^{(n-1)/2} \equiv \pm 1 \pmod{n}$, однако n не есть эйлерово псевдопростое число по основанию b.

15. а) Доказать, что если n — эйлерово псевдопростое по основанию $b \in (\mathbf{Z}/n\mathbf{Z})^*$, то оно также эйлерово псевдопростое по основаниям $-b$ и b^{-1}.

б) Доказать, что если n — эйлерово псевдопростое по основаниям b_1 и b_2, то оно также эйлерово псевдопростое по основанию $b_1 b_2$.

16. Пусть n имеет вид $p(2p-1)$, как в упражнении 1 г).

а) Доказать, что n — эйлерово псевдопростое по 25% возможных оснований $b \in (\mathbf{Z}/n\mathbf{Z})^*$.

б) Найти класс чисел n такого типа, являющихся сильно псевдопростыми по 25% возможных оснований.

17. Пусть n имеет вид $(6m+1)(12m+1)(18m+1)$, как в упражнении 10.

а) Доказать, что если m нечетно, то n — эйлерово псевдопростое по 50% всех возможных оснований $b \in (\mathbf{Z}/n\mathbf{Z})^*$.

б) Если m четно, то n — эйлерово псевдопростое по 25% возможных оснований $b \in (\mathbf{Z}/n\mathbf{Z})^*$.

18. а) Используя обозначение O-большое, оценить число двоичных операций в ситуации, когда тест Миллера–Рабина применяется столько раз, что вероятность прохождения составного числа n через эти тесты меньше $1/m$ (n и m очень велики).

б) Предполагая справедливой обобщенную гипотезу Римана, оценить число двоичных операций в ситуации, когда тест Миллера–Рабина применяется столько раз, сколько необходимо для достаточной уверенности в том, что прошедшее все тесты число n является простым.

19. а) Доказать, что если n — псевдопростое по основанию 2, то $N = 2^n - 1$ является сильно псевдопростым и эйлеровым псевдопростым по основанию 2.

б) Доказать, что существует бесконечно много сильно псевдопростых и эйлеровых псевдопростых по основанию 2.

20. Доказать, что если n — сильно псевдопростое по основанию b, то оно сильно псевдопростое также по основанию b^k для любого целого k.

21. Пусть n — число Кармайкла 561.

а) Найти число оснований $b \in (\mathbf{Z}/561\mathbf{Z})^*$, для которых 561 — эйлерово псевдопростое.

б) Найти число оснований, для которых 561 — сильно псевдопростое. Привести их список.

22. Доказать, что если $n = p^\alpha$ есть степень простого числа, где $\alpha > 1$, то n — сильно псевдопростое по основанию b тогда и только тогда, когда оно — псевдопростое по основанию b.

23. а) Показать, что 65 — сильно псевдопростое число по основанию 8 и по основанию 18, однако не является таковым по основанию $14 \equiv 8 \cdot 18 \pmod{65}$.

б) Для любого нечетного составного числа n пусть (*) обозначает утверждение: если n — сильно псевдопростое по основанию b_1 и по основанию b_2, то оно — сильно псевдопростое по основанию $b_1 b_2$ (другими словами, свойство сильной простоты сохраняется при перемножении оснований). Доказать, что (*) справедливо тогда и только тогда, когда n — либо степень простого числа, либо делится на простое число, сравнимое с 3 по модулю 4.

24. а) Доказать, что если найдется такое b, что n — псевдопростое число по основанию b, но *не* является сильно псевдопростым по b, то можно быстро найти нетривиальный делитель числа n.

б) Объяснить, как можно защититься от этого при выборе $n = pq$ в криптосистеме RSA.

З а м е ч а н и е. Во многих тестах на простоту оказывается, что если составное n проходит начальный тест, но не проходит какого-либо из последующих, полученная информация не только доказывает, что n — составное число, но и позволяет легко найти его нетривиальный множитель. Упражнение 24 — тому пример: если n проходит тест на псевдопростоту по основанию b, а затем не проходит тест на сильную псевдопростоту по основанию b, то можно найти делитель n. Однако это не означает, что тесты на простоту можно использовать в задаче факторизации. Если n — большое составное число (например, произведение больших случайно выбранных простых чисел p и q), то очень маловероятно, что мы найдем основание b, по которому n — псевдопростое (упражнение 5 а) дает представление о вероятности выбора такого b). Таким образом, различные изощренные тесты на псевдопростоту позволяют лишь подтвердить простоту действительно простого числа. На практике составное число, которое мы хотим разложить на множители, не будет проходить любой отдельный тест на простоту, но сами тесты не помогут нам найти его множители.

ЛИТЕРАТУРА к § 1 ГЛАВЫ V

1. *Adleman L. M., Pomerance C., Rumely R. S.* On distinguishing prime numbers from composite numbers. — Ann. Math., 1983, v. 117, p. 173–206.
2. *Cohen H., Lenstra H. W., Jr.* Primality testing and Jacobi sums. — Math. Comp., 1984, v. 42, p. 297–330.
3. *Dixon J. D.* Factorization and primality tests. — Amer. Math. Monthly, 1984, v. 91, p. 333–352.
4. *Kranakis E.* Primality and Cryptography. N.Y. etc.: Wiley, 1986.
5. *Lenstra A.* Primality testing. — Crypt. Comput. Number Theory, Proc. Symp. Appl. Math., 1990, v. 42, p. 13–25.
6. *Miller G. L.* Riemann's hypothesis and test for primality. — In: Proceedings of the 7th Annual ACM Symposium on the Theory of Computing, p. 234–239.
7. *Pomerance C.* Recent developments in primality testing. — Math. Intelligencer, 1981, v. 3, p. 97–105.
8. *Pomerance C.* The search for prime numbers. — Sci. American, 1982, v. 247, p. 136–147.
9. *Rabin M. O.* Probabilistic algorithms for testing primality. — J. Number Theory, 1980, v. 12, p. 128–138.

10. *Solovay R., Strassen V.* A fast Monte Carlo test for primality. — SIAM J. Comput., 1977, v. 6, p. 84–85; erratum: 1978, v. 7, p. 118.
11. *Wagon S.* Primality testing. — Math. Intelligencer, 1986, v. 8, № 3, p. 58–61.

§ 2. Ро-метод

Предположим, что нам известно, что некоторое большое нечетное число n составное. Например, оно не прошло один из тестов на простоту из § 1. Как уже упоминалось, это не значит, что мы имеем какое-либо представление о его делителях. Из рассмотренных нами методов проверки простоты лишь самый медленный — последовательное деление n на простые числа, меньшие $\sqrt{n}$, — одновременно с установлением непростоты конкретно указывает делитель числа n. Все более быстрые алгоритмы проверки простоты лишь устанавливают факт существования делителей числа n, но не дают о них информации.

Метод пробных делений на простые числа, меньшие $\sqrt{n}$, может потребовать более $O(\sqrt{n})$ двоичных операций. Простейший алгоритм, работающий существенно быстрее, — это «ро-метод» факторизации Полларда (называемый также «методом Монте-Карло»).

Первый шаг ро-метода заключается в выборе легко вычислимого отображения кольца $\mathbf{Z}/n\mathbf{Z}$ в себя, а именно, простого многочлена с целыми коэффициентами, например, $f(x) = x^2 + 1$. Затем выбираем некоторое конкретное значение x_0 (это может быть 1 или 2 или случайно порожденное целое число) и последовательно вычисляем итерации $x_1 = f(x_0)$, $x_2 = f(f(x_0))$, $x_3 = f(f(f(x_0)))$ и т. д. Таким образом, мы полагаем

$$x_{j+1} = f(x_j), \qquad j = 0, 1, 2, \ldots$$

После этого производится сравнение различных x_j с целью найти два, принадлежащих различным классам вычетов по модулю n, но одному классу вычетов по модулю некоторого делителя числа n. Как только такие x_j, x_k найдены, сразу находим собственный делитель числа n как НОД $(x_j - x_k, n)$.

Пример 1. Разложим на множители число 91, взяв $f(x) = x^2 + 1$, $x_0 = 1$. Имеем $x_1 = 2$, $x_2 = 5$, $x_3 = 26$ и т. д. Находим НОД $(x_3 - x_2, n) =$ НОД $(21, 91) = 7$, т. е. 7 — делитель. Конечно, этот пример тривиален: делитель 7 быстрее было найти пробным делением.

В ро-методе важно выбрать многочлен $f(x)$, отображающий $\mathbf{Z}/n\mathbf{Z}$ в себя нерегулярным, «случайным» образом. Так, вскоре мы убедимся, что $f(x)$ должен не быть линейным многочленом и должен не порождать взаимно однозначное отображение.

Предположим теперь, что $f(x)$ — «случайное» отображение $\mathbf{Z}/n\mathbf{Z}$ в себя, и оценим, как долго нам придется ждать появления двух таких итераций x_j и x_k, что $x_j - x_k$ и n имеют нетривиальный общий делитель. Для этого для фиксированного (но не известного нам) делителя r числа n найдем типичную (*в среднем* по всем отображениям $\mathbf{Z}/n\mathbf{Z}$ в себя и по всем значениям x_0) величину первого такого значения k, что $x_k \equiv x_j \pmod r$ при некотором $j < k$. Иными словами, мы рассматриваем $f(x)$ как отображение $\mathbf{Z}/r\mathbf{Z}$ в себя и интересуемся, сколько потребуется итераций до возникновения первого повторения $x_j = x_k$ в $\mathbf{Z}/r\mathbf{Z}$.

Предложение V.2.1. *Пусть S — множество из r элементов. Для отображения f множества S в себя и элемента $x_0 \in S$ полагаем $x_{j+1} = f(x_j)$, $j = 0, 1, 2, \ldots$ Пусть λ — положительное вещественное число и пусть $l = 1 + [\sqrt{2\lambda r}]$. Тогда доля пар (f, x_0), для которых все $x_0, x_1, \ldots, x_l$ различны (где f пробегает все отображения из S в S, а x_0 пробегает все элементы S), меньше $e^{-\lambda}$.*

Доказательство. Общее число рассматриваемых пар равно r^{r+1}, так как существует всего r вариантов выбора для x_0 и для каждого из r различных элементов $x \in S$ имеется r вариантов выбора $f(x)$. Сколько существует таких пар (f, x_0), для которых $x_0, x_1, \ldots, x_l$ различны? Значение x_0 можно выбрать r способами, значение $f(x_0)$ можно выбрать $r-1$ способами (так как $f(x_0) \neq x_0$), значение $f(x_1)$ можно выбрать $r-2$ способами и т. д., пока функция f не будет определена для всех $x_0, x_1, \ldots, x_{l-1}$. Значения функции f на остальных значениях аргумента произвольны и имеется r^{r-l} вариантов их выбора. Поэтому общее число таких способов выбора x_0 и задания функции $f(x)$, что все $x_0, x_1, \ldots, x_l$ различны, равно

$$r^{r-l} \prod_{j=0}^{l} (r-j),$$

а доля тех пар, которые обладают указанным свойством (отношение приведенного выражения к r^{r+1}), равна

$$r^{-l-1} \prod_{j=0}^{l} (r-j) = \prod_{j=1}^{l} \left(1 - \frac{j}{r}\right).$$

Предложение утверждает, что логарифм этого выражения меньше $-\lambda$ (где $l = 1 + [\sqrt{2\lambda r}]$). Чтобы доказать это, возьмем логарифм произведения в правой части и используем неравенство $\log(1-x) < -x$ при $0 < x < 1$ (на геометрическом языке это означает, что логарифмическая кривая лежит ниже касательной к ней в точке $(1, 0)$). Используя

формулу для суммы первых l натуральных чисел, получаем

$$\log\left(\prod_{j=1}^{l}\left(1-\frac{j}{r}\right)\right) < \sum_{j=1}^{l} -\frac{j}{r} = \frac{-l(l+1)}{2r} < \frac{-l^2}{2r} < \frac{-(\sqrt{2\lambda r})^2}{2r} = -\lambda,$$

что и требовалось показать. Предложение доказано.

Предложение V. 2. 1 дает оценку для вероятного времени работы ро-метода *при условии*, что выбранный многочлен ведет себя как «случайное» отображение $\mathbf{Z}/r\mathbf{Z}$ в себя. Прежде чем обсуждать эту оценку, мы изложим более эффективную версию ро-метода.

Напомним, что ро-метод состоит в последовательном вычислении $x_k = f(x_{k-1})$ и сравнении x_k с ранее полученными x_j до тех пор, пока не найдется пара с НОД $(x_k - x_j, n) = r > 1$. Однако с ростом k вычисление НОД $(x_k - x_j, n)$ для всех $j < k$ становится очень трудоемким. Покажем теперь, как надо действовать, чтобы при каждом k вычислять лишь один НОД. Сначала заметим, что если при некоторых k_0, j_0 для делителя r числа n выполняется сравнение $x_{k_0} \equiv x_{j_0} \pmod r$, то $x_k \equiv x_j \pmod r$ для любой последующей пары индексов j, k, имеющих ту же разность: $k - j = k_0 - j_0 = m$. Чтобы убедиться в этом, достаточно к сравнению $x_{k_0} \equiv x_{j_0} \pmod r$ применить m раз многочлен f.

Опишем теперь работу алгоритма. Последовательно вычисляя x_k, на каждом шаге делаем следующее. Пусть k является $(h+1)$-разрядным целым числом в двоичной системе счисления, т. е. $2^h \leqslant k \leqslant 2^{h+1}$. Пусть j — наибольшее число из h двоичных разрядов: $j = 2^h - 1$. Сравниваем x_k с x_j, т. е. вычисляем НОД $(x_k - x_j, n)$. Если в результате получаем нетривиальный делитель числа n, то останавливаемся, если нет, то переходим к $k+1$.

Преимущество этой модификации состоит в том, что при каждом k вычисляется лишь один НОД, а недостаток — в том, что алгоритм может пропустить первый такой момент k_0, что НОД $(x_{k_0} - x_{j_0}, n) = r > 1$ при некотором $j_0 < k_0$. Однако довольно быстро пара элементов x_k, x_j, разность которых имеет нетривиальный общий множитель с n, все же будет обнаружена. А именно, пусть $j = 2^{h+1} - 1$ и $k = j + (k_0 - j_0)$. Тогда j — максимальное $(h+1)$-разрядное число, k состоит из $h+2$ разрядов, при этом НОД $(x_k - x_j, n) > 1$. Заметим, что $k < 2^{h+2} = 4 \cdot 2^h \leqslant 4k_0$.

П р и м е р 2. Вернемся к примеру 1, но теперь будем сравнивать каждое x_k лишь с тем x_j, для которого j является наибольшим из меньших k чисел вида $2^h - 1$. Для $n = 91$, $f(x) = x^2 + 1$, $x_0 = 1$ имеем $x_1 = 2$, $x_2 = 5$, $x_3 = 26$, как и выше, и $x_4 = 40$ (так как $26^2 + 1 \equiv 40 \pmod{91}$). Следуя описанному выше алгоритму, находим

делитель числа n, вычисляя НОД $(x_4 - x_3, n) =$ НОД $(14, 91) = 7$.

Пример 3. Разложить на множители число 4087, используя $f(x) = x^2 + x + 1$ и $x_0 = 2$.

Решение. Проводим вычисления в следующем порядке:

$$x_1 = f(2) = 7; \quad \text{НОД}\,(x_1 - x_0, n) = \text{НОД}\,(7 - 2, 4087) = 1;$$

$$x_2 = f(7) = 57; \quad \text{НОД}\,(x_2 - x_1, n) = \text{НОД}\,(57 - 7, 4087) = 1;$$

$$x_3 = f(57) = 3307; \quad \text{НОД}\,(x_3 - x_1, n) = \text{НОД}\,(3307 - 7, 4087) = 1;$$

$$x_4 \equiv f(3307) \equiv 2745 \pmod{4087};$$

$$\text{НОД}\,(x_4 - x_3, n) = \text{НОД}\,(2745 - 3307, 4087) = 1;$$

$$x_5 \equiv f(2745) \equiv 1343 \pmod{4087};$$

$$\text{НОД}\,(x_5 - x_3, n) = \text{НОД}\,(1343 - 3307, 4087) = 1;$$

$$x_6 \equiv f(1343) \equiv 2626 \pmod{4087};$$

$$\text{НОД}\,(x_6 - x_3, n) = \text{НОД}\,(2626 - 3307, 4087) = 1;$$

$$x_7 \equiv f(2626) \equiv 3734 \pmod{4087};$$

$$\text{НОД}\,(x_7 - x_3, n) = \text{НОД}\,(3734 - 3307, 4087) = 61.$$

Итак, мы получили $4087 = 61 \cdot 67$. Разложение найдено.

Предложение V.2.2. *Пусть n — нечетное составное целое число и r — его нетривиальный делитель, не превосходящий $\sqrt{n}$ (т. е. $r|n$, $1 < r < \sqrt{n}$; мы хотим найти такое r). Пусть пара (f, x_0), состоящая из многочлена f с целыми коэффициентами и начального значения x_0, выбрана так, что ведет себя подобно «случайной» паре (f, x_0) в смысле предложения* V.2.1 *(f отображает $\mathbf{Z}/r\mathbf{Z}$ в себя, x_0 — целое число). Тогда ро-метод выявляет делитель r за $O(\sqrt[4]{n}\log^3 n)$ двоичных операций с высокой степенью вероятности. Точнее, существует такая константа C, что для любого вещественного положительного числа λ вероятность того, что ро-метод не найдет нетривиального делителя числа n за $C\sqrt{\lambda}\sqrt[4]{n}\log^3 n$ двоичных операций, меньше $e^{-\lambda}$.*

Доказательство. Пусть C_1 — такая константа, что НОД $(y - z, n)$ можно вычислить за $C_1 \log^3 n$ двоичных операций, если $y, z \leqslant n$ (см. § I.3). Пусть C_2 — такая константа, что наименьший неотрицательный вычет из $f(x)$ по модулю n можно вычислить за $C_2 \log^2 n$ двоичных операций, если $x < n$ (см. § I.1). Если k_0 — первый из индексов, для которых существует $j_0 < k_0$ со свойством $x_{k_0} \equiv x_{j_0} \pmod{r}$, то ро-метод в том виде, как он был описан выше, позволяет найти r на k-м шаге, где $k < 4k_0$. (Строго говоря, может

случиться, что НОД $(x_k - x_j, n) > r$, т. е. НОД $((x_k - x_j)/r, n/r) > 1$. Однако вероятность случайно выбрать целое число, не взаимно простое с n/r, мала, особенно если n — произведение небольшого количества больших простых чисел. Поэтому мы пренебрежем этой возможностью, которая, в худшем случае, потребовала бы увеличения константы в предположениях предложения V. 2. 2.)

Таким образом, число двоичных операций, необходимых для нахождения r, ограничено величиной $4k_0(C_1 \log^3 n + C_2 \log^2 n)$. Согласно предложению V.2.1 вероятность того, что $k_0 > 1 + \sqrt{2\lambda r}$, меньше $e^{-\lambda}$. Если $k_0 \leqslant 1 + \sqrt{2\lambda r}$, то число двоичных операций, необходимых для нахождения r, оценивается (с учетом того, что $r < \sqrt{n}$) величиной

$$4(1+\sqrt{2\lambda r})(C_1 \log^3 n + C_2 \log^2 n) < 4(1+\sqrt{2}\sqrt{\lambda}\sqrt[4]{n})(C_1 \log^3 n + C_2 \log^2 n).$$

Если мы выберем C несколько больше, чем $4\sqrt{2}(C_1 + C_2)$ (позаботившись о прибавлении 1), получим, что, как и утверждалось, делитель r будет найден за $C\sqrt{\lambda}\sqrt[4]{n} \log^3 n$ двоичных операций, если только мы не сделаем неудачного выбора пары (f, x_0), вероятность чего меньше $e^{-\lambda}$.

Замечания. 1. Основное предположение, лежащее в основе ро-метода, состоит в том, что можно найти многочлены, которые ведут себя подобно случайным отображениям в смысле предложения V. 2. 1. Этого доказано не было. Однако практический опыт разложения чисел ро-методом позволяет предположить, что «случайный» многочлен ведет себя подобно случайному отображению и что есть очень простые многочлены (чаще всего используется $f(x) = x^2 + 1$), обладающие этим свойством «случайности».

2. Согласно предложению V. 2. 2, если мы выберем λ достаточно большим, чтобы быть уверенными в успехе (например, выберем $\lambda = 9$, тогда $e^{-\lambda} \approx 0,0001$), то мы почти наверное разложим число n за $3C\sqrt[4]{n} \log^3 n$ двоичных операций.

УПРАЖНЕНИЯ

В упражнениях 1–4 применить ро-метод с указанными $f(x)$ и x_0 для разложения заданного числа n. В каждом случае сравнивать x_k лишь с x_j, для которого $j = 2^h - 1$, если k является $(h+1)$-разрядным двоичным числом.

1. $x^2 - 1$, $x_0 = 2$, $n = 91$.
2. $x^2 + 1$, $x_0 = 1$, $n = 8051$.
3. $x^2 - 1$, $x_0 = 5$, $n = 7031$.
4. $x^3 + x + 1$, $x_0 = 1$, $n = 2701$.
5. Пусть множество S состоит из r элементов, а отображение f в паре (f, x_0) пробегает все *биекции* S на себя (т. е. $f : S \longrightarrow S$ есть взаимно однозначное соответствие S с самим собой, т. е. никакие два разных элемента не имеют одинакового

образа при отображении f). Как и прежде, пусть $x_{j+1} = f(x_j)$, $j = 0, 1, \ldots$ Для каждой пары (f, x_0) пусть k обозначает первый из таких индексов, для которого найдется $j < k$ со свойством $x_j = x_k$. Доказать, что:

а) k не превосходит r и принимает каждое значение от 1 до r с вероятностью $1/r$;

б) среднее значение для k равно $(r + 1)/2$ (среднее берется по всем парам (f, x_0), где f — биекция).

6. Используя упражнение 5, объяснить, почему в ро-методе в качестве $f(x)$ никогда не следует брать линейный многочлен $ax + b$.

7. Предположим, что ро-метод используется для разложения числа, имеющего простой делитель r. В качестве функции для итерации решено взять $f(x) = x^2$. (Это плохой выбор, что станет ясно ниже.) Нас интересует первое значение k, при котором $x_k \equiv x_j \pmod r$ для некоторого $j < k$, т. е. первое значение k, при котором *не* все $x_0, x_1, \ldots, x_k$ различны (*по модулю* r). Предположим, что в качестве x_0 выбран порождающий элемент группы $(\mathbf{Z}/r\mathbf{Z})^*$. Положим $r - 1 = 2^s t$, где t нечетно.

а) Написать сравнение по модулю $r - 1$, эквивалентное соотношению $x_k = x_l$ (равенство обозначает сравнение по модулю r).

б) Найти первые значения k и l, для которых выполнено условие из а), выразив ответ в терминах s и двоичного разложения дроби $1/t$.

в) Сколь велико должно быть k по сравнению с r? Почему сделанный выбор функции $f(x)$ является плохим для ро-метода?

ЛИТЕРАТУРА к § 2 ГЛАВЫ V

1. *Blair W. D.*, *Lacampagne C. B.*, *Selfridge J. L.* Factoring large numbers on a pocket calculator. — Amer. Math. Monthly, 1986, v. 93, p. 802–808.
2. *Brent R. P.* An improved Monte Carlo factorization algorithm. — BIT, 1980, v. 20, p. 176–184.
3. *Brent R. P.*, *Pollard J. M.* Factorization of the eighth Fermat number. — Math. Comp., 1981, v. 36, p. 627–630.
4. *Guy R. K.* How to factor a number. — In: Proceedings of the 5th Manitoba Conference on Numerical Mathematics, 1975, p. 49–89.
5. *Pollard J. M.* A Monte Carlo method for factorization. — BIT, 1975, v. 15, p. 331–334.

§ 3. Факторизация Ферма и факторные базы

Факторизация Ферма. Как мы видели раньше (см. упражнение 3 к § I. 2 и упражнение 4 к § IV. 2), существует способ факторизации составного числа n, который эффективен, если n — произведение двух целых чисел, близких одно к другому. Этот метод, называемый «факторизацией Ферма», основывается на том факте, что n в этом случае равно разности двух квадратов, значение одного из которых очень невелико.

Предложение V.3.1. *Пусть n — натуральное число. Существует взаимно однозначное соответствие между факторизациями n в виде $n = ab$, где $a \geqslant b > 0$, и представлениями числа n в виде $t^2 - s^2$, где s и t — неотрицательные целые числа. Это соответствие дается уравнениями:*

$$t = \frac{a+b}{2}, \quad s = \frac{a-b}{2}, \quad a = t+s, \quad b = t-s.$$

Доказательство. Если дана такая факторизация, мы пишем $n = ab = ((a+b)/2)^2 - ((a-b)/2)^2$, таким образом, мы получаем представление в виде разности двух квадратов. Обратно, если дано, что $n = t^2 - s^2$, мы можем разложить правую часть на множители: $(t+s)(t-s)$. Приведенные в предложении уравнения дают в явном виде взаимно однозначное соответствие между двумя способами представления числа n.

Если $n = ab$, где a и b близки друг к другу, то $s = (a-b)/2$ мало, и таким образом, t лишь немного больше $\sqrt{n}$. В этом случае мы можем найти a и b, пробуя все значения для t, начиная с $[\sqrt{n}] + 1$, пока не найдем такое значение, при котором $t^2 - n = s^2$ есть полный квадрат.

Везде в дальнейшем мы будем предполагать, что n никогда не является полным квадратом; тем самым нам не нужно беспокоиться о тривиальных исключениях в тех или иных процедурах и утверждениях.

Пример 1. Разложить на множители 200819.

Решение. Имеем $[\sqrt{200819}] + 1 = 449$. Находим, что $449^2 - 200819 = 782$, и это не есть полный квадрат. Далее, мы пробуем $t = 450$: $450^2 - 200819 = 1681 = 41^2$. Итак, $200819 = 450^2 - 41^2 = (450 + 41)(450 - 41) = 491 \cdot 409$.

Заметим, что если даже a и b в разложении $n = ab$ не близки, то метод факторизации Ферма все равно найдет a и b, но только после большого числа проб для $t = [\sqrt{n}] + 1, [\sqrt{n}] + 2, \ldots$ Существует обобщение метода факторизации Ферма, которое зачастую работает лучше в такой ситуации. Мы выбираем малое k, последовательно полагаем $t = [\sqrt{kn}] + 1, [\sqrt{kn}] + 2$ и т. д. до тех пор, пока не получим такое t, для которого $t^2 - kn = s^2$ является полным квадратом. Тогда $(t+s)(t-s) = kn$, и, таким образом, $t+s$ имеет с n общий нетривиальный делитель. Он может быть найден путем вычисления НОД $(t+s, n)$.

Пример 2. Разложить на множители 141467.

Решение. Если мы попытаемся использовать факторизацию Ферма, полагая $t = 377, 378, \ldots$, то через некоторое время мы утомимся пробовать различные t. Однако, если мы будем пробовать $t = [\sqrt{3n}]+1 = 652, \ldots$, мы скоро обнаружим, что $655^2 - 3 \cdot 141467 = 68^2$;

тогда мы вычислим НОД $(655+68, 141467) = 241$. Мы получаем тем самым, что $141467 = 241 \cdot 587$. Обобщенная факторизация Ферма при $k = 3$ дала результат потому, что в разложении $n = ab$ множитель b близок к $3a$. При $k = 3$ нам пришлось пробовать только четыре значения t, тогда как при простой факторизации Ферма (т.е. с $k = 1$) пришлось бы испробовать 38 значений t.

Факторные базы. Еще одно обобщение идеи, лежащей в основе факторизации Ферма, приводит к значительно более эффективному методу факторизации. А именно, если нам удается получить сравнение вида $t^2 \equiv s^2 \pmod n$ с $t \not\equiv \pm s \pmod n$, то можно сразу найти множитель числа n, вычисляя НОД $(t+s, n)$ (или НОД $(t-s, n)$). Действительно, n делит $t^2 - s^2 = (t+s)(t-s)$, но не делит ни $t+s$, ни $t-s$; таким образом, НОД $(t+s, n)$ должен быть собственным делителем a числа n, и тогда $b = n/a$ делит НОД $(t-s, n)$.

Пример 3. Предположим, что мы хотим разложить на множители 4633 и замечаем, что $118^2 \equiv 25 \equiv 5^2 \pmod{4633}$. Тогда мы находим, что НОД $(118+5, 4633) = 41$, а НОД $(118-5, 4633) = 113$, и $4633 = 41 \cdot 113$.

Скептик, скорее всего, усомнится в том, что в примере 3 можно было угадать число 118, наименьший положительный вычет квадрата которого является полным квадратом. Возможно ли при случайном выборе разных b быстро получить такое значение, что наименьший положительный вычет b^2 по модулю n есть полный квадрат? Это крайне маловероятно, если n велико. Поэтому возникает необходимость обобщить этот подход таким образом, чтобы он допускал значительно большую свободу при выборе тех b, для которых мы рассматриваем $b^2 \pmod n$. Идея заключается в выборе нескольких b_i, обладающих тем свойством, что $b_i^2 \pmod n$ есть произведение степеней малых простых чисел, а произведение некоторого подмножества b_i дает число, сравнимое по модулю n с полным квадратом. Перейдем теперь к деталям.

Под «наименьшим абсолютным вычетом» числа a по модулю n мы понимаем целое число в интервале от $-n/2$ до $n/2$, сравнимое с a. Мы будем обозначать его $a \pmod n$.

Определение. *Факторной базой* называется множество $B = \{p_1, p_2, \ldots, p_h\}$, состоящее из различных простых чисел, кроме числа p_1, которое может равняться -1. Мы скажем, что квадрат числа b есть *B-число* (при заданном n), если наименьший абсолютный вычет $b^2 \pmod n$ можно записать как произведение чисел из B.

Пример 4. Пусть $n = 4633$ и $B = \{-1, 2, 3\}$. Тогда квадраты трех целых чисел — 67, 68, 69 являются B-числами. Действительно, $67^2 \equiv -144 \pmod{4633}$, $68^2 \equiv -9 \pmod{4633}$, $69^2 \equiv 128 \pmod{4633}$.

Пусть $\mathbf{F}_2^h$ обозначает векторное пространство над полем из двух элементов, которое состоит из наборов нулей и единиц, по h в каждом наборе. Если дано n и факторная база B, состоящая из h чисел, то каждому B-числу можно сопоставить некоторый вектор $\vec{\varepsilon} \in \mathbf{F}_2^h$. А именно, запишем $b^2 \pmod n$ в виде $\prod_{j=1}^h p_j^{\alpha_j}$ и положим j-ю компоненту ε_j равной $\alpha_j \pmod 2$, т. е. $\varepsilon_j = 0$, если α_j четно, и $\varepsilon_j = 1$, если α_j нечетно.

Пример 5. В ситуации примера 4 вектор, соответствующий 67 — это $(1,0,0)$, числу 68 соответствует также $(1,0,0)$, а числу 69 — вектор $(0,1,0)$.

Предположим, что у нас есть такое множество B-чисел $b_i^2 \pmod n$, что сумма соответствующих векторов $\vec{\varepsilon} = (\varepsilon_{i1}, \ldots, \varepsilon_{ih})$ есть нулевой вектор в $\mathbf{F}_2^h$. Тогда произведение наименьших абсолютных вычетов b_i^2 равно произведению *четных* степеней всех p_j, входящих в B, т. е. если для каждого i обозначим через a_i наименьший абсолютный вычет $b_i^2 \pmod n$ и положим $a_i = \prod_{j=1}^h p_j^{\alpha_{ij}}$, то получим

$$\prod a_i = \prod_{j=1}^h p_j^{\sum_i \alpha_{ij}};$$

показатель каждого p_j в правой части — четное число. Тогда правая часть равенства есть квадрат числа $\prod_j p_j^{\gamma_j}$, где $\gamma_j = \frac{1}{2}\sum_i \alpha_{ij}$. Таким образом, наименьшие положительные вычеты $b = \prod_i b_i \pmod n$ и $c = \prod_j p_j^{\gamma_j} \pmod n$, полученные совершенно различными путями (одно — как произведение чисел b_i, другое — как произведение чисел p_j), имеют квадраты, сравнимые по модулю n.

Если оказывается, что $b \equiv \pm c \pmod m$, то нам не повезло, и мы должны начать снова, с другой совокупностью B-чисел, у которых соответствующие им векторы дают в сумме нуль. Так произойдет, например, если мы по недосмотру выберем все $b_i < \sqrt{n/2}$; все векторы будут тогда нулевыми, и мы придем в итоге к тривиальному сравнению.

Однако при составном n и более случайном выборе b_i следует ожидать, что b и c будут сравнимыми (с точностью до ± 1) по модулю n не более чем в 50% случаев. Дело в том, что квадрат по модулю n имеет $2^r \geqslant 4$ квадратных корней, если n имеет $r \geqslant 2$ различных простых множителей (см. упражнение 7 к § I.3); значит, случайный квадратный корень из b^2 лишь с вероятностью $2/2^r \leqslant \frac{1}{2}$ совпадает с b или $-b$. А если у нас есть такие b и c, что $b^2 \equiv c^2 \pmod n$, но $b \not\equiv \pm c \pmod n$, мы сразу можем найти нетривиальный множитель НОД $(b + c, n)$. Таким образом, если мы k раз повторяем описанную

процедуру нахождения b и c для получения пары, дающей нетривиальный делитель n, то вероятность неудачи в результате всех этих k попыток не превосходит $1/2^k$.

Как же практически выбирать факторную базу B и b_i? Один из методов: в качестве B взять первые h простых чисел (или первые $h-1$ простых и $p_1 = -1$); числа b_i выбирать случайно, пока не найдется несколько таких, что их квадраты — B-числа. Другой метод: выбрать сначала несколько чисел b_i, для которых $b_i^2 \pmod n$ (наименьшие абсолютные вычеты) малы по абсолютной величине (например, выбрать b_i близкими к $\sqrt{kn}$ для малых k или так, как будет описано в § 4). Затем выбираем в качестве B небольшую совокупность малых простых чисел (включая в него также $p_1 = -1$), чтобы несколько $b_i^2 \pmod n$ можно было выразить через числа из B.

Пример 6. В ситуации примеров 4–5 мы выбирали 67 и 68, так как они близки к $\sqrt{4633}$. Обнаружив, что $67^2 \equiv -144 \pmod{4633}$, $68^2 \equiv -9 \pmod{4633}$, мы положили $B = \{-1, 2, 3\}$. Как мы видели, числам 67 и 68 соответствуют векторы $(1,0,0)$, и их сумма — нулевой вектор. Поэтому $b = 67 \cdot 68 \equiv -77 \pmod{4633}$ и $c = 2^{\gamma_2} 3^{\gamma_3}$ (степенью -1 в c мы можем пренебречь), т. е. $c = 36$. Так как $-77 \not\equiv \pm 36 \pmod{4633}$, то мы находим делитель, вычисляя НОД $(-77 + 36, 4633) = 41$.

Сколько нужно иметь b_i, чтобы можно было найти сумму $\vec{\varepsilon_i}$, равную нулевому вектору? Другими словами, какой объем должна иметь совокупность векторов из $\mathbf{F}_2^h$, чтобы она содержала в себе подмножество векторов с нулевой суммой, т. е. чтобы она была *линейно зависимой над полем* $\mathbf{F}_2$. Согласно основам линейной алгебры (справедливым над полем $\mathbf{F}_2$ так же, как и над полем вещественных чисел) любые $h+1$ векторов линейно зависимы. Поэтому в худшем случае нам нужно образовать $h+1$ различных B-чисел, чтобы найти первый пример сравнения вида $\prod_i b_i^2 \equiv (\prod_j p_j^{\gamma_j})^2 \pmod n$. (Пример 6 показывает, что линейно зависимые векторы могут получиться и раньше; в этом примере $h = 3$, а мы остановились, найдя два B-числа). Если h велико, то, возможно, поверхностный просмотр не позволит выделить подмножество векторов с нулевой суммой. В этом случае нам следует записать векторы в виде строк матрицы и с помощью методов линейной алгебры (обнуляя строки) найти множество линейно зависимых строк.

Пример 7. Пусть $n = 4633$. Найти такую наименьшую факторную базу B, чтобы квадраты 68, 69 и 96 были B-числами, и затем разложить 4633.

Решение. Как мы уже убедились, $68^2 \pmod n$ и $69^2 \pmod n$ являются произведениями чисел $-1, 2$ и 3; так как $96^2 \equiv -50 \pmod n$,

наименьшие абсолютные вычеты всех трех квадратов можно выразить через элементы факторной базы $B=\{-1,2,3,5\}$. Мы уже нашли векторы $\varepsilon_1=(1,0,0,0)$ и $\varepsilon_2=(0,1,0,0)$, отвечающие, соответственно, 68 и 69. Так как $96^2\equiv -50 \pmod{4633}$, то $\varepsilon_3=(1,1,0,0)$. Эти три вектора имеют нулевую сумму, поэтому возьмем $b=68\cdot 69\cdot 96\equiv 1031 \pmod{4633}$ и $c=2^4\cdot 3\cdot 5$. Находим: НОД $(1031+240,4633)=41$.

Примеры 6 и 7 демонстрируют систематический способ поиска нескольких b_i таких, что наименьшие абсолютные вычеты $b_i^2 \pmod n$ оказываются произведениями малых простых чисел. Вероятность того, что $b_i^2 \pmod n$ есть произведение малых простых, увеличивается, если этот вычет мал по абсолютному значению. Таким образом, мы можем последовательно пробовать целые числа b_i, близкие к $\sqrt{kn}$ для малых значений k. Например, мы можем выбрать $[\sqrt{kn}], [\sqrt{kn}]+1$ для $k=1,2,\ldots$

Пример 8. Разложим на множители $n=1829$, выбирая в качестве b_i все такие числа $[\sqrt{1829k}]$, $[\sqrt{1829k}]+1$, $k=1,2,\ldots$, что $b_i^2 \pmod n$ есть произведение простых чисел, меньших 20. Представляем такие числа в виде $b_i^2 \pmod n=\prod_j p_j^{\alpha_{ij}}$ и составляем таблицу из α_{ij}. Рассмотрев $k=1,2,3,4$, получаем таблицу, в которой первым элементом j-го столбца является p_j, а элемент этого столбца в строке, соответствующей значению b_i, — это показатель, с которым p_j входит в разложение $b_i^2 \pmod n$.

b_i	-1	2	3	5	7	11	13
42	1	–	–	1	–	–	1
43	–	2	–	1	–	–	–
61	–	–	2	–	1	–	–
74	1	–	–	–	–	1	–
85	1	–	–	–	1	–	1
86	–	4	–	1	–	–	–

Ищем теперь множества строк, у которых суммы элементов в каждом столбце четны. Сразу видно, что сумма 2-й и 6-й строк четная: $-6-2---$. Это дает сравнение $(b_2b_6)^2\equiv(2^{6/2}5^{2/2})^2 \pmod n$, т.е. $(43\cdot 86)^2\equiv 40^2 \pmod{1829}$. Но так как $43\cdot 86\equiv 40 \pmod{1829}$, мы нашли лишь тривиальное соотношение. Поэтому приходится искать другие подмножества строк, суммы которых содержат лишь четные элементы. Мы замечаем, что таковы строки 1, 2, 3 и 5; их сумма $22222-2$; это дает сравнение $(42\cdot 43\cdot 61\cdot 85)^2\equiv(2\cdot 3\cdot 5\cdot 7\cdot 13)^2 \pmod n$, т.е. $1459^2\equiv 901^2 \pmod{1829}$. Мы получаем отсюда, что делителем 1829 является НОД $(1459+901,1829)=59$.

Алгоритм факторных баз. Опишем теперь систематический метод факторизации очень большого n при помощи *случайного* выбора чисел b_i. Выбираем целое число y «промежуточной» величины, например, если n — 50-значное десятичное число, то выбираем в качестве y число с 5 или 6 десятичными знаками. Пусть B состоит из -1 и всех простых чисел, не превосходящих y. Выбираем случайным образом много b_i и пытаемся разложить $b_i^2 \pmod n$ (наименьшие абсолютные вычеты) в произведение чисел из B. Получив большое количество B-чисел $b_i^2 \pmod n$ (достаточно найти $\pi(y)+2$ таких чисел, где $\pi(y)$ — число простых, не превосходящих y), вычисляем соответствующие векторы в $\mathbf{F}_2^h$ (где $h = \pi(y)+1$) и с помощью обнуления строк находим такое подмножество b_i, что сумма соответствующих векторов $\vec{\varepsilon_i}$ дает нулевой вектор. Затем образуем $b = \prod_i b_i \pmod n$ и $c = \prod_j p_j^{\gamma_j} \pmod n$ так, как это было описано выше. Тогда $b^2 \equiv c^2 \pmod n$. Если $b \equiv \pm c \pmod n$, повторяем действия с начала, с другой совокупностью B-чисел (или, что более эффективно, выбираем в матрице из $\vec{\varepsilon_i}$ другое подмножество строк с нулевой суммой, находя, если необходимо, несколько большее количество B-чисел и соответствующих им строк). Когда мы, наконец, получим $b^2 \equiv c^2 \pmod n$ и $b \not\equiv \pm c \pmod n$, вычисляем НОД $(b+c, n)$, который и будет нетривиальным делителем n.

Эвристическая временная оценка. Мы дадим теперь очень грубый вывод оценки числа двоичных операций, требующихся для факторизации *очень* большого n при применении описанного выше алгоритма. Мы воспользуемся несколькими упрощающими предположениями и приближенными соотношениями; в любом случае наш результат — лишь вероятностная оценка. Если нам очень не повезло при выборе b_i, алгоритм будет работать дольше.

Нам потребуются следующие факты.

Ф а к т 1 (формула Стирлинга). Приближенное значение $\log n!$ есть $n \log n - n$.

Слово «приближенное» означает, что разность величин растет при $n \to \infty$ много медленнее, чем n. Это соотношение можно доказать, заметив, что $\log n!$ — сумма Римана для определенного интеграла $\int_1^n (\log x)\,dx = n \log n - n + 1$, построенная по крайним правым точкам разбиения $[1, n]$ на отрезки единичной длины.

Ф а к т 2. Если N — натуральное число и $u > 0$, то общее число таких N-выборок $(\alpha_1, \alpha_2, \ldots, \alpha_N)$, состоящих из неотрицательных целых чисел, что $\sum_{j=1}^N \alpha_j \leqslant u$, равно биномиальному коэффициенту $\binom{[u]+N}{N}$.

Здесь $[u]$ обозначает функцию целой части (наибольшее целое число, не превосходящее u). Факт 2 можно доказать, если сопоставить каждой N-выборке α некоторую N-выборку β из чисел $1, 2, \ldots, N+[u]$ по правилам: $\beta_1 = \alpha_1 + 1$ и $\beta_{j+1} = \beta_j + \alpha_{j+1} + 1$ для $j \geqslant 1$, т.е. мы выбираем числа β_j так, что между β_{j-1} и β_j имеется α_j чисел. Это дает взаимно однозначное соответствие между множеством N-выборок α с $\alpha_1 + \cdots + \alpha_N \leqslant u$ и множеством способов выбора N чисел из совокупности $N + [u]$ чисел.

Основным этапом при оценке времени работы нашего алгоритма является оценка вероятности того, что случайное число, меньшее x, будет произведением простых чисел, меньших y (где y — число, значительно меньшее x). Прежде всего, обозначим через u дробь $\frac{\log x}{\log y}$. Таким образом, если x есть r-разрядное целое число в двоичной записи, а y — s-разрядное целое число в двоичной записи, то u близко к отношению r/s.

По ходу оценок удобно делать некоторые упрощения, пренебрегая малыми членами. Для этого будем предполагать, что u *значительно* меньше y. Как обычно, мы обозначаем через $\pi(y)$ количество простых чисел, не превосходящих y. Так как по теореме о простых числах $\pi(y)$ приближенно равно $y/\log y$, мы предполагаем также, что u значительно меньше $\pi(y)$. В типичном практическом применении алгоритма мы можем выбирать для x, y, u приблизительно такие по величине значения:

$x \approx 10^{48}$;

$y \approx 10^6$ (так что $\pi(y) \approx 7 \cdot 10^4$ и $\log y \approx 14$);

$u \approx 8$.

Как обычно, пусть $\Psi(x, y)$ обозначает число всех целых, не превосходящих x, которые не делятся ни на какие простые числа, большие y, т.е. число целых, которые можно записать в виде $\prod p_j^{\alpha_j} \leqslant x$, где произведение берется по всем простым, не превосходящим y, и α_j — неотрицательные целые числа. Существует очевидное взаимно однозначное соответствие между такими $\pi(y)$-выборками неотрицательных целых чисел α_j, что $\prod_j p_j^{\alpha_j} \leqslant x$, и целыми числами, не превосходящими x, которые не делятся на простые числа, большие y. Логарифмируя, находим, что $\Psi(x, y)$ равно числу целочисленных решений α_j неравенства $\sum_{j=1}^{\pi(y)} \alpha_j \log p_j \leqslant \log x$. Заметим теперь, что большинство простых p_j имеет логарифмы, не намного меньшие, чем $\log y$. Действительно, большинство простых, меньших y, имеют почти то же число цифр, что и y; лишь относительно немногие имеют существенно меньшее число цифр и, следовательно, много меньший логарифм. Поэтому мы позволим себе в предыдущем неравенстве заменить $\log p_j$ на $\log y$. Поделив обе части получившегося неравенства

на $\log y$ и заменив $\log x/\log y$ на u, мы можем сказать, что $\Psi(x,y)$ примерно равно числу решений неравенства $\sum_{j=1}^{\pi(y)} \alpha_j \leqslant u$.

Сделаем теперь еще одно важное упрощение: заменим число переменных $\pi(y)$ на y. Это может, на первый взгляд, показаться довольнотаки опрометчивой модификацией нашей задачи. В самом деле, при замене $\pi(y)$ на y появляются дополнительные ненулевые слагаемые; однако оказывается, что они сокращаются, и итоговый результат получается тем же самым, что при значительно более точной аппроксимации $\Psi(x,y)$. Таким образом, мы будем предполагать, что $\Psi(x,y)$ приближенно равно числу неотрицательных целочисленных решений неравенства $\sum_{i_j=1}^{y} \alpha_j \leqslant u$.

Используя факт 2 с $N = y$, мы получаем, что $\Psi(x,y)$ приближенно равно $\binom{[u]+y}{y}$. Оценим теперь величину $\log(\Psi(x,y)/x)$ — логарифм вероятности того, что случайно взятое целое число между 1 и x представляется в виде произведения простых чисел, не превосходящих y. Заметим теперь, что $\log x = u \log y$ по определению u, и применим полученное приближение для $\Psi(x,y)$ и факт 1:

$$\begin{aligned}\log\left(\frac{\Psi(x,y)}{x}\right) &\approx \log\left(\frac{([u]+y)!}{[u]!y!}\right) - u\log y \\ &\approx ([u]+y)\log([u]+y) - ([u]+y) - \\ &\quad - ([u]\log[u] - [u]) - (y\log y - y) - u\log y.\end{aligned}$$

Сделаем еще несколько приближений. Во-первых, заменим $[u]$ на u. Затем, учитывая, что u значительно меньше y (в силу нашего предположения), заменим $\log(y+u)$ на $\log y$. После сокращения получим:

$$\log\left(\frac{\Psi(x,y)}{x}\right) \approx -u\log u,$$

т. е.

$$\frac{\Psi(x,y)}{x} \approx u^{-u}.$$

Отсюда, например, следует, что если $x \approx 10^{48}$ и $y \approx 10^6$ (как выше), то вероятность того, что случайно выбранное целое число между 1 и x есть произведение простых чисел, не превосходящих y, равна примерно $1/8^8$.

Теперь мы можем оценить число двоичных операций, используемых описанным выше алгоритмом факторных баз. Для простоты будем предполагать, что наша факторная база B состоит из первых $h = \pi(y)$ простых чисел, т. е. из всех простых, не превосходящих y. Для упрощения анализа мы будем также предполагать, что -1 не

включается в B и что мы рассматриваем наименьшие положительные вычеты (а не наименьшие абсолютные вычеты) $b_i^2 \pmod n$.

Итак, мы оцениваем число двоичных операций, требующихся для выполнения следующих шагов.

Ш а г 1. Выбираем случайно числа b_i в интервале от 1 до n; находим наименьший неотрицательный вычет $b_i^2 \pmod n$ и выражаем его, если это возможно, в виде произведения простых чисел, не превосходящих y; останавливаемся тогда, когда удается набрать $\pi(y)+1$ различных b_i, для которых $b_i^2 \pmod n$ записываются как такие произведения.

Ш а г 2. Находим множество из линейно зависимых строк в соответствующей $(0,1)$-матрице размера $(\pi(y)+1)\times\pi(y)$, чтобы получить сравнение вида $b^2 \equiv c^2 \pmod n$.

Ш а г 3. Если $b \equiv \pm c \pmod n$, то повторяем шаги 1 и 2 с новыми b_i до тех пор, пока не получим сравнение $b^2 \equiv c^2 \pmod n$ с $b \not\equiv \pm c \pmod n$. Тогда находим нетривиальный множитель n, вычисляя НОД $(b+c, n)$.

Если предполагать, что $b_i^2 \pmod n$ (понимаемые как наименьшие неотрицательные вычеты по модулю n) распределены равномерно между 1 и n, то в соответствии с приведенными выше соображениями нам придется выбрать порядка u^u (где $u = \log n/\log y$) чисел до появления такого b_i, что $b_i^2 \pmod n$ есть произведение простых чисел, не превосходящих y. Мы решим позднее, как выбрать y, чтобы минимизировать время ожидания. Суть дела заключается в том, что, выбирая y большим, мы делаем u^u малым и поэтому будем часто встречать b_i, для которых $b_i^2 \pmod n$ — произведение простых, не превосходящих y. Однако в этом случае разложение $b_i^2 \pmod n$ в произведение таких простых (для чего нам придется проводить $\pi(y)+1$ делений) и последующее обнуление строк матрицы становятся слишком трудоемкими операциями. С другой стороны, если выбрать y очень малым, то последние задачи станут легче, однако придется слишком долго ждать появления такого b_i, что $b_i^2 \pmod n$ разлагается в произведение простых чисел, не превосходящих y, так как в этом случае u^u станет очень большим. Следовательно, y нужно выбирать в некоторой промежуточной области, чтобы избежать этих крайностей.

Чтобы решить вопрос о выборе y, мы сначала получим грубую оценку числа двоичных операций в терминах y (и, конечно, n). Затем мы минимизируем ее по y (используя элементарные вычисления и некоторые упрощающие приближения) и найдем временную оценку для того y, который дает минимальное время работы.

Предположим, что n — r-разрядное, а y — s-разрядное числа;

тогда u весьма близко к r/s. Прежде всего, сколько нужно двоичных операций для каждой проверки случайно выбранного b_i? Мы утверждаем, что это число операций полиномиально по r и y, т. е. что оно есть $O(r^l l^{ks})$ для некоторых (довольно малых) целых чисел k и l. Для порождения случайного бита требуется некоторое фиксированное время и, таким образом, нужно $O(r)$ двоичных операций, чтобы породить случайное целое число между 1 и n. Далее, вычисление $b_i^2 \pmod n$ требует $O(r^2)$ двоичных операций. Затем нам придется последовательно делить $b_i^2 \pmod n$ на все простые числа, не превосходящие y, на которые оно делится без остатка (или на степени таких простых), надеясь после всех делений получить 1. Простой способ проделать это (хоть и не самый эффективный) — делить на два и затем на все нечетные простые p от 3 до y, записывая по ходу дела, на какую степень p происходит деление без остатка. Заметим, что если p — не простое, то оно не делит без остатка, так как его простые сомножители были уже удалены из $b_i^2 \pmod n$. Так как деление r-разрядного целого числа на s-разрядное целое число в двоичной системе счисления проводится за время $O(rs)$, то каждая проверка случайно выбранного b_i требует $O(rsy)$ двоичных операций.

При выполнении шага 1 придется проверить приблизительно $u^u(\pi(y)+1)$ значений b_i, чтобы найти $\pi(y)+1$ таких значений b_i, что $b_i^2 \pmod n$ представляется в виде произведения простых, не превосходящих y. Так как $\pi(y) \approx \frac{y}{\log y} = O(y/s)$, то на шаг 1 затрачивается $O(u^u r y^2)$ двоичных операций.

Шаг 2 включает в себя операции со сложностью, полиномиальной по y и r (такие, как преобразования матриц и нахождение b и c по модулю n). Таким образом, шаг 2 выполняется за $O(y^j r^h)$ двоичных операций для некоторых целых j и h. При каждом выполнении шагов 1–2 вероятность успеха, т. е. нахождения $b \not\equiv \pm c \pmod n$, не меньше 50%. Точнее, вероятность успеха равна 50%, если n делится в точности на два разных простых числа, и больше, если n имеет более двух простых делителей. Поэтому, если нас устраивает, скажем, вероятность $1 - 2^{-50}$ нахождения нетривиального делителя n, то шаги 1–2 нам достаточно пройти 50 раз. Принимая эту величину достаточной для практических целей, при подходящих целых k и h получаем окончательную оценку

$$O(50(u^u r^2 y^2 + y^j r^h)) = O(r^h u^u y^j) = O(r^h u^u e^{ks}) = O(r^h (r/s)^{r/s} e^{ks}).$$

Мы найдем теперь y — или, что равносильно, s, — для которого эта оценка минимальна. Так как число r бит в записи n фиксированно, это означает минимизацию $(r/s)^{r/s} e^{ks}$ по s, или, что эквивалентно,

минимизацию его логарифма, который равен $\frac{r}{s}\log\frac{r}{s}+ks$. Для этого мы полагаем

$$0=\frac{d}{ds}\left(\frac{r}{s}\log\frac{r}{s}+ks\right)=-\frac{r}{s^2}\left(\log\frac{r}{s}+1\right)+k\approx-\frac{r}{s^2}\log\frac{r}{s}+k,$$

т.е. мы выбираем s так, чтобы ks было примерно равно $\frac{r}{s}\log\frac{r}{s}$, другими словами, чтобы два сомножителя в $(r/s)^{r/s}e^{ks}$ были примерно равны между собой. Так как k — константа, из последнего приближенного равенства следует, что порядки s^2 и $r\log(r/s)=r(\log r-\log s)$ одинаковы; поэтому s имеет порядок величины между $\sqrt{r}$ и $\sqrt{r\log r}$. Значит, $\log s$ близок к $\frac{1}{2}\log r$. Делая подстановку $\log s\approx\frac{1}{2}\log r$, мы приводим полученное выше соотношение к виду

$$0\approx-\frac{r}{2s^2}\log r+k\quad\text{или}\quad s\approx\sqrt{\frac{r}{2k}\log r}.$$

Оценим теперь время работы при этом значении s. Так как при оптимальном выборе s множители $(r/s)^{r/s}$ и e^{ks} приблизительно равны по величине, то временная оценка упрощается: $O(e^{2ks})=O(e^{\sqrt{2k}\sqrt{r\log r}})$. Заменяя константу $\sqrt{2k}$ на C, мы, наконец, приходим к следующей оценке числа двоичных операций, затраченных на факторизацию r-разрядного двоичного числа n: $O(e^{C\sqrt{r\log r}})$.

Приведенные выше выкладки не являются строгими. Мы не пытались обосновывать наши упрощения или оценивать погрешности в приближенных равенствах. К тому же как сам алгоритм, так и наша оценка времени его работы — вероятностные.

До последнего времени, пока не был изобретен метод решета в числовом поле (см. замечание в конце § 5), все оценки времени работы самых лучших универсальных алгоритмов факторизации имели вид $O(e^{C\sqrt{r\log r}})$. В некоторых случаях оценки доказывались строго, а в иных — основывались на правдоподобных, но не доказанных предположениях. Оценки для различных алгоритмов различались, в основном, константой C в показателе. В этом отношении история задачи факторизации совершенно не похожа на историю рассмотренных в § 1 критериев простоты, где улучшение временных оценок (в особенности в детерминистических тестах на простоту) было прямо-таки поразительным. Подробный обзор и сравнение методов факторизации, которые были известны к началу 80-х гг., имеется в статье Померанца (Pomerance) 1982 г., указанной в приводимом ниже списке литературы.

Замечание. Так как $r = O(\log n)$, приведенную выше временную оценку можно выразить в виде

$$\text{Time}(\text{Разложить } n) = O(e^{C\sqrt{\log n \log\log n}}).$$

За исключением метода решета в числовых полях, эвристические оценки времени работы всех асимптотически быстрых универсальных факторизационных алгоритмов имеет указанный выше вид при $C = 1 + \varepsilon$, где ε произвольно мало.

Применения к RSA. Напомним, что надежность криптосистемы с открытым ключом RSA (см. §IV.2) основана на том, что разложение на множители очень большого числа n вида $n = pq$ требует значительно больших затрат времени, чем действия, которые должны выполнять законные пользователи системы и сложность которых (как при проверке простоты) является полиномиальной или почти полиномиальной функцией от числа r бит в n. Мы только что видели, почему при анализе факторизационных алгоритмов часто возникают временные оценки вида $O(e^{C\sqrt{r\log r}})$. Так как полиномиальную функцию от r можно записать в виде $O(e^{C\log r})$, то для больших r время, затрачиваемое на факторизацию, много больше времени работы полиномиальных или почти полиномиальных алгоритмов. (Однако факторизационные алгоритмы с временной оценкой $O(e^{C\sqrt{r\log r}})$ для больших r лучше, чем ро-метод, временная оценка которого — приблизительно $O(\sqrt[4]{n}) = O(e^{Cr})$, где $C = \frac{1}{4}\log 2$.)

Наконец, отметим, что вопрос о замене $\sqrt{r\log r}$ в показателе меньшей функцией от r — это не единственный вопрос, имеющий практическое значение при оценке надежности системы RSA. В конце концов, полиномиальная функция от числа r бит становится значительно меньшей, чем $C_1 e^{C_2\sqrt{r\log r}}$, лишь когда r велико, а насколько велики такие значения r, в сильной степени зависит от констант C_1 и C_2. Поэтому даже построение алгоритма факторизации с той же временной оценкой, но с меньшими константами, повлияло бы на оценку надежности системы RSA с открытым ключом.

УПРАЖНЕНИЯ

1. Использовать факторизацию Ферма, чтобы разложить на множители:
а) 8633; б) 809009; в) 92296873; г) 88169891; д) 4061.

2. Показать, что если делитель числа n отстоит от $\sqrt{n}$ не более, чем на $\sqrt[4]{n}$, то факторизация Ферма дает результат при первой попытке (т. е. для $t = [\sqrt{n}] + 1$).

3. а) Доказать, что обобщенная факторизация Ферма не позволяет найти делитель большого нечетного числа n, если выбрано $k = 2$ или если k — любое целое число, делящееся на 2, но не на 4.

б) Доказать, что если $k = 4$ и если обобщенная факторизация Ферма работает для некоторого t, то и «простая» (с $k = 1$) факторизация Ферма работает столь же хорошо.

4. Использовать обобщенную факторизацию Ферма, чтобы разложить на множители:

а) 68987; б) 29895581; в) 19578079; г) 17018759.

5. Пусть $n = 2701$. Использовать B-числа $52^2, 53^2 \pmod n$ (при подходящей факторной базе B), чтобы разложить 2701. Какие векторы $\vec{\varepsilon}$ соответствуют 52 и 53?

6. Пусть $n = 4633$. Использовать 68, 152 и 153 с соответствующей факторной базой B, чтобы разложить 4633. Каковы соответствующие векторы?

7. а) Доказать, что $\log n! - (n \log n - n) = O(\log n)$.

б) Вывести более точную оценку $\log n! - ((n + \frac{1}{2}) \log n - n) = O(1)$.

в) Каково математическое ожидание $\log j$ для случайно выбранного j в интервале от 1 до y?

8. а) Какова вероятность того, что k случайно выбранных векторов в $\mathbf{F}_2^n$ линейно независимы (при $k \leqslant n$)?

б) Какова вероятность того, что 5 случайно выбранных векторов в $\mathbf{F}_2^5$ образуют базис?

9. Пусть n — r-разрядное двоичное целое число. На какой множитель умножается каждое из выражений $\sqrt[4]{n}$ (которое входит во временную оценку для ро-метода) и $e^{\sqrt{r \log r}}$ (которое входит во временную оценку для метода факторных баз), когда n возрастает от 50-значного десятичного числа до 100-значного?

10. а) Пусть $f(s)$ — положительная монотонно убывающая функция, а $g(s)$ — положительная монотонно возрастающая функция на некотором интервале. Предположим, что $f(s_0) = g(s_0)$. Доказать, что функция $h(s) = f(s) + g(s)$ «по существу» достигает минимума в точке s_0 в том смысле, что минимальное значение $h(s)$ лежит между $h(s_0)$ и $\frac{1}{2}h(s_0)$.

б) Предположим, что $f(s) > 1$ — монотонно убывающая функция, $g(s) > 1$ — монотонно возрастающая функция на интервале и $f(s_0) = g(s_0)$. Доказать, что $h(s) = f(s)g(s)$ «по существу» достигает минимума в точке s_0 в том смысле, что минимальное значение $h(s)$ лежит между $h(s_0)$ и $\sqrt{h(s_0)}$.

в) Используя часть б), показать, что функция $h(s) = (r/s)^{r/s} e^{ks}$ на интервале $(0, r)$ (здесь k и r — положительные константы) «по существу» достигает своего минимального значения при $(r/s)^{r/s} = e^{ks}$.

ЛИТЕРАТУРА к § 3 ГЛАВЫ V

1. *Dickson L. E.* History of the Theory of Numbers. Vol. 1. N.Y.: Chelsea, 1952.
2. *Kraitchik M.* Théorie des nombers. Vol. 2. Paris–Montrouge: Gauthier–Villars, 1926.
3. *Lehman R. S.* Factoring large integers. — Math. Comp., 1974, v. 28, p. 637–646.
4. *Pomerance C.* Analysis and comparison of some integer factoring algorithms. — Computational Methods in Number Theory. Part I. Amsterdam: Mathematic Centrum, 1982.

§ 4. Метод цепных дробей

В последнем параграфе мы видели, что эффективность метода факторных баз для нахождения нетривиального делителя большого составного натурального числа n зависит от наличия хорошего метода получения целых чисел b между 1 и n, для которых наименьший абсолютный вычет $b^2 \pmod n$ разлагается в произведение малых простых чисел. Вероятность такого разложения тем больше, чем меньше значение наименьшего абсолютного вычета $b^2 \pmod n$. В этом параграфе мы опишем метод (восходящий к Лежандру) поиска многих b с тем свойством, что $|b^2 \pmod n| < 2\sqrt{n}$. Этот метод использует «цепные дроби». Поэтому мы начнем с краткого введения в теорию представления вещественных чисел в виде цепных дробей. Наше описание коснется лишь свойств, которые потребуются в дальнейшем. Более подробное изложение теории цепных дробей можно найти, например, в классической книге Давенпорта (Davenport), см. ссылки в конце параграфа*.

Цепные дроби. Для вещественного числа x построим его разложение в виде цепной дроби следующим образом. Пусть $a_0 = [x]$ ($[x]$ — наибольшее целое число, не превосходящее x). Положим $x_0 = x - a_0$; пусть $a_1 = [1/x]$ и $x_1 = 1/x_0 - a_1$; для $i > 1$ положим $a_i = [1/x_{i-1}]$ и $x_i = 1/x_{i-1} - a_i$. Когда оказывается, что $1/x_{i-1}$ — целое число, то $x_i = 0$ и процесс прекращается. Нетрудно понять, что процесс заканчивается в том и только том случае, когда x — рациональное число (потому что в этом случае дроби x_i — рациональные числа с уменьшающимися знаменателями). В силу построения чисел $a_0, a_1, \ldots, a_i$ для каждого i мы можем записать:

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots \cfrac{1}{a_i + x_i}}},$$

что обычно записывается в более компактной форме:

$$x = a_0 + \frac{1}{a_1+}\,\frac{1}{a_2+}\,\frac{1}{a_3+}\cdots\frac{1}{a_i + x_i}.$$

Предположим, что x — *иррациональное* вещественное число. Если мы выполняем описанное разложение до i-го члена и затем удаляем x_i,

---

* На русском языке теория цепных дробей излагается в книгах А. Я. Хинчина «Цепные дроби» (М.: ГИФМЛ, 1961), Дж. В. С. Касселса «Введение в теорию диофантовых приближений» (М.: ИЛ, 1961) и др. — *Прим. ред.*

то получаем рациональное число b_i/c_i, называемое i-м *приближением* цепной дроби x:

$$\frac{b_i}{c_i} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\frac{1}{a_3+}\cdots\frac{1}{a_{i-1}+}\frac{1}{a_i}.$$

Предложение V.4.1. *В обозначениях, введенных выше, имеем:*
a) $\frac{b_0}{c_0} = \frac{a_0}{1}$; $\frac{b_1}{c_1} = \frac{a_0a_1+1}{a_1}$; $\frac{b_i}{c_i} = \frac{a_ib_{i-1}+b_{i-2}}{a_ic_{i-1}+c_{i-2}}$ *для* $i \geqslant 2$;
b) *в п.* a) *дроби в правых частях — несократимые, т. е. если* $b_i = a_ib_{i-1} + b_{i-2}$ *и* $c_i = a_ic_{i-1} + c_{i-2}$, *то* НОД $(b_i, c_i) = 1$;
c) $b_ic_{i-1} - b_{i-1}c_i = (-1)^{i-1}$ *для* $i \geqslant 1$.

Доказательство. Определим последовательности $\{b_i\}$ и $\{c_i\}$ с помощью соотношений из части a) (приравнивая соответствующие числители и знаменатели — *прим. ред.*); докажем по индукции, что тогда b_i/c_i — i-е приближение. Докажем это, не предполагая, что a_i — целые числа, т. е. докажем, что для любых вещественных чисел a_i отношение b_i/c_i с b_i и c_i, определенными формулами в части a), равно $a_0 + \frac{1}{a_1+}\cdots\frac{1}{a_i}$. База индукции ($i = 0, 1, 2$) тривиальна. Предположим теперь, что утверждение справедливо вплоть до i-го приближения, и докажем его для $(i+1)$-го приближения. Заметим, что мы получим $(i+1)$-е приближение, заменяя a_i на $a_i + 1/a_{i+1}$ в формуле, которая выражает числитель и знаменатель i-го приближения в терминах $(i-1)$-го и $(i-2)$-го приближений. Таким образом, $(i+1)$-е приближение в силу предположения индукции равно

$$\frac{(a_i + \frac{1}{a_{i+1}})b_{i-1} + b_{i-2}}{(a_i + \frac{1}{a_{i+1}})c_{i-1} + c_{i-2}} = \frac{a_{i+1}(a_ib_{i-1} + b_{i-2}) + b_{i-1}}{a_{i+1}(a_ic_{i-1} + c_{i-2}) + c_{i-1}} = \frac{a_{i+1}b_i + b_{i-1}}{a_{i+1}c_i + c_{i-1}}.$$

Тем самым часть a) доказана.

Часть c) также легко доказать по индукции. Шаг индукции проводится следующим образом:

$$\begin{aligned} b_{i+1}c_i - b_ic_{i+1} &= (a_{i+1}b_i + b_{i-1})c_i - b_i(a_{i+1}c_i + c_{i-1}) = b_{i-1}c_i - b_ic_{i-1} \\ &= -(-1)^{i-1} = (-1)^i, \end{aligned}$$

Таким образом, из того, что часть c) справедлива для i, следует соответствующее утверждение для $i+1$.

Наконец, часть b) следует из части c), так как любой общий делитель b_i и c_i должен быть делителем числа $(-1)^{i-1}$, равного ± 1. Предложение доказано.

Если разделить равенство в предложении V.4.1 c) на c_ic_{i-1}, то получим:

$$\frac{b_i}{c_i} - \frac{b_{i-1}}{c_{i-1}} = \frac{(-1)^{i-1}}{c_ic_{i-1}}.$$

Так как c_i, очевидно, образуют строго возрастающую последовательность натуральных чисел, это равенство показывает, что последовательность приближений ведет себя подобно знакопеременному ряду, т. е. она колеблется вперед и назад с уменьшающейся амплитудой, и, значит, последовательность приближений сходится к пределу.

Наконец, нетрудно заметить, что предел последовательности есть само число x, для которого строится разложение. Чтобы показать это, заметим, что x получается, если в $(i+1)$-м приближении заменить a_{i+1} на $1/x_i$. Таким образом, по предложению V.4.1 а) (с $i+1$ вместо i и заменой a_{i+1} на $1/x_i$) имеем:

$$x = \frac{b_i/x_i + b_{i-1}}{c_i/x_i + c_{i-1}} = \frac{b_i + x_i b_{i-1}}{c_i + x_i c_{i-1}},$$

при этом x оказывается в промежутке между b_{i-1}/c_{i-1} и b_i/c_i. (Чтобы убедиться в этом, рассмотрим два вектора $\mathbf{u} = (b_i, c_i)$ и $\mathbf{v} = (b_{i-1}, c_{i-1})$ в плоскости, оба расположенные в одном и том же квадранте; заметим, что направление вектора $\mathbf{u} + x_i\mathbf{v}$ является промежуточным между направлениями векторов $\mathbf{u}$ и $\mathbf{v}$.) Стало быть, последовательность b_i/c_i колеблется вокруг x и сходится к x.

Цепные дроби имеют много интересных свойств, что обусловило их применения в различных областях математики. Например, они доставляют «наилучший возможный» способ рациональной аппроксимации вещественных чисел (в том смысле, что любое рациональное число, более близкое к x, чем b_i/c_i, должно иметь знаменатель, больший, чем c_i). Другое свойство аналогично тому факту, что в представлении вещественного числа x в десятичной (или b-ичной) системе знаки периодически повторяются тогда и только тогда, когда x рационально. В разложении x в цепную дробь, как мы видели, последовательность целых чисел a_i обрывается в том и только том случае, когда x рационально. Можно показать, что последовательность целых чисел a_i становится периодической последовательностью в том и только том случае, когда x — квадратичная иррациональность, т. е. имеет вид $x_1 + x_2\sqrt{n}$, где x_1 и x_2 — рациональны, а n не есть полный квадрат. Это — известная теорема Лагранжа.

Пример 1. Выписывая по порядку первые члены разложения $\sqrt{3}$ в цепную дробь, получаем

$$\sqrt{3} = 1 + \frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{2+}\,\frac{1}{1+}\,\frac{1}{2+}\cdots.$$

Естественно предположить, что a_i попеременно принимают значения 1 и 2. Докажем это. Пусть x равно бесконечной цепной дроби в правой части с чередующимися 1 и 2. Согласно определению цепной дроби

для x мы, очевидно, имеем $x = 1 + \frac{1}{1+\frac{1}{1+x}}$. Упрощая это рациональное выражение и умножая затем обе части равенства на $x+2$, получаем $2x + x^2 = 3 + 2x$, т.е. $x = \sqrt{3}$.

Предложение V.4.2. *Пусть $x > 1$ — вещественное число, при разложении в цепную дробь имеющее приближения b_i/c_i.*

Тогда $|b_i^2 - x^2 c_i^2| < 2x$ при любом i.

Доказательство. Как мы видели, x лежит в промежутке между b_i/c_i и b_{i+1}/c_{i+1}. Разность между этими величинами по абсолютной величине равна $\frac{1}{c_i c_{i+1}}$ (предложение V.4.1 c)). Поэтому

$$|b_i^2 - x^2 c_i^2| = c_i^2 \left|x - \frac{b_i}{c_i}\right| \left|x + \frac{b_i}{c_i}\right| < c_i^2 \frac{1}{c_i c_{i+1}} \left(x + \left(x + \frac{1}{c_i c_{i+1}}\right)\right).$$

Следовательно,

$$|b_i^2 - x^2 c_i^2| - 2x < 2x\left(-1 + \frac{c_i}{c_{i+1}} + \frac{1}{2x c_{i+1}^2}\right) < 2x\left(-1 + \frac{c_i}{c_{i+1}} + \frac{1}{c_{i+1}}\right)$$
$$< 2x\left(-1 + \frac{c_{i+1}}{c_{i+1}}\right) = 0.$$

Предложение доказано.

Предложение V.4.3. *Пусть n — натуральное число, не являющееся полным квадратом. Пусть b_i/c_i — приближения при разложении $\sqrt{n}$ в цепную дробь. Тогда наименьший по абсолютному значению вычет $b_i^2 \pmod n$ (расположенный между $-n/2$ и $n/2$) меньше $2\sqrt{n}$.*

Доказательство. Применим предложение V.4.2 при $x = \sqrt{n}$. Тогда $b_i^2 \equiv b_i^2 - n c_i^2 \pmod n$, а это последнее число меньше $2\sqrt{n}$ по абсолютной величине.

Предложение V.4.3 — ключевое для алгоритма цепных дробей. Оно показывает, что можно найти последовательность чисел b_i, квадраты которых имеют малые вычеты по модулю n. Для этого нужно брать числители приближений при разложении $\sqrt{n}$ в цепную дробь. Заметим, что нам нужно находить не сами приближения, а только их числители b_i, и то лишь по модулю n. Поэтому быстрый рост числителей и знаменателей приближений не является препятствием. Нам никогда не придется работать с числами, большими n^2 (при умножении чисел по модулю n).

Опишем теперь последовательно, как работает алгоритм цепных дробей. Он отличается от метода факторных баз из § 3 лишь использованием предложения V.4.3 вместо случайного выбора чисел b_i.

Алгоритм разложения на множители при помощи цепных дробей. Пусть n — целое число, которое нужно разложить на

множители. Все вычисления будем проводить по модулю n, т. е. произведения и суммы целых чисел будем приводить по модулю n к их наименьшему неотрицательному вычету (или к наименьшему абсолютному вычету в шаге 3). Вначале полагаем $b_{-1} = 1$, $b_0 = a_0 = [\sqrt{n}]$, $x_0 = \sqrt{n} - a_0$. Вычисляем $b_0^2 \pmod n$ (фактически $b_0^2 - n$). Далее для $i = 1, 2, \ldots$ последовательно выполняем следующие действия.

1) Полагаем $a_i = [1/x_{i-1}]$ и затем $x_i = 1/x_{i-1} - a_i$.

2) Полагаем $b_i = a_i b_{i-1} + b_{i-2}$ (приводим по модулю n).

3) Вычисляем $b_i^2 \pmod n$. Проделав это для нескольких i, обращаем внимание на те из чисел на шаге 3, абсолютная величина которых представляется в виде произведения малых простых чисел. Пусть факторная база B состоит из -1 и простых чисел, которые встречаются более чем в одном из $b_i^2 \pmod n$ (или входят только в одно из $b_i^2 \pmod n$ в четной степени). Составляем список всех тех чисел $b_i^2 \pmod n$, которые являются B-числами, и соответствующих векторов $\vec{\varepsilon}_i$ из нулей и единиц. Если возможно, находим подмножество векторов с нулевой суммой. Полагаем $b = \prod b_i$ (проводя вычисления по модулю n; произведение берется по подмножеству с $\sum \vec{\varepsilon}_i = 0$). Полагаем $c = \prod p_j^{\gamma_j}$, где p_j — элементы из B (исключая -1), и $\gamma_j = \frac{1}{2} \sum \alpha_{ij}$ (где сумма берется по тому же подмножеству индексов i, см. § 3). Если $b \not\equiv \pm c \pmod n$, то НОД $(b + c, n)$ — нетривиальный делитель n. Если $b \equiv \pm c \pmod n$, то ищем другое подмножество индексов i, для которого $\sum \vec{\varepsilon}_i = 0$. Если невозможно найти такое подмножество индексов i, что $\sum \vec{\varepsilon}_i = 0$, то нужно продолжать, вычисляя новые a_i, b_i, $b_i^2 \pmod n$ и при необходимости расширяя факторную базу B.

Замечание. Чтобы упростить вычисление $c = \prod p_j^{\gamma_j}$, удобно для каждого B-числа $b_i^2 \pmod n$ записывать вектор $\vec{\alpha}_i = (\ldots, \alpha_{ij}, \ldots)$, а не $\vec{\varepsilon}_i$, который представляет собой вектор $\vec{\alpha}_i$, приведенный по модулю 2.

Пример 2. Используем изложенный алгоритм для разложения 9073 на множители.

Решение. Вначале выпишем по порядку a_i и b_i (где b_i — наименьший неотрицательный вычет $a_i b_{i-1} + b_{i-2}$ по модулю n), а также наименьшие абсолютные вычеты $b_i^2 \pmod n$:

i	0	1	2	3	4
a_i	95	3	1	26	2
b_i	95	286	381	1119	2619
$b_i^2 \pmod n$	-48	139	-7	87	-27

Рассматривая последнюю строку таблицы, мы видим, что в качестве B разумно взять множество $\{-1, 2, 3, 7\}$. Тогда $b_i^2 \pmod n$ — B-

числа при $i = 0, 2, 4$. Соответствующие векторы $\vec{\alpha}_i$ равны $(1,4,1,0)$, $(1,0,0,1)$ и $(1,0,3,0)$. Сумма первого и третьего — нулевой вектор по модулю 2. Поэтому полагаем $b = 95 \cdot 2619 \equiv 3834 \pmod{9073}$ и $c = 2^2 \cdot 3^2 = 36$. Таким образом, $3834^2 \equiv 36^2 \pmod{9073}$. Так как $3834 \not\equiv 36 \pmod{9073}$, мы получаем нетривиальный делитель: НОД $(3834 + 36, 9073) = 43$. Итак, $9073 = 43 \cdot 211$.

Пример 3. Разложить на множители 17873.

Решение. Как и в примере 2, мы начинаем с таблицы:

i	0	1	2	3	4	5
a_i	133	1	2	4	2	3
b_i	133	134	401	1738	3877	13369
$b_i^2 \pmod n$	−184	83	−56	107	−64	161

Если положить $B = \{-1, 2, 7, 23\}$, то B-числа получаются при $i = 0, 2, 4, 5$; соответствующие векторы $\vec{\alpha}_i$ равны $(1,3,0,1)$, $(1,3,1,0)$, $(1,6,0,0)$ и $(0,0,1,1)$. Сумма 1-го, 2-го и 4-го из этих векторов — нулевой вектор по модулю 2. Однако если вычислить $b = 133 \cdot 401 \cdot 13369 \equiv 1288 \pmod{17873}$ и $c = 2^3 \cdot 7 \cdot 23 = 1288$, то окажется, что $b \equiv c \pmod{17873}$. Поэтому приходится продолжить поиск B-чисел, которым соответствуют векторы с суммой, равной нулю по модулю 2. Продолжая таблицу, получаем:

i	6	7	8
a_i	1	2	1
b_i	17246	12115	11488
$b_i^2 \pmod n$	−77	149	−88

Если мы теперь расширим B, включив туда еще простое число 11, т.е. $B = \{-1, 2, 7, 11, 23\}$, то для $i = 0, 2, 4, 5, 6, 8$ получаем B-числа со следующими векторами $\vec{\alpha}_i$: $(1,3,0,0,1)$, $(1,3,1,0,0)$, $(1,6,0,0,0)$, $(0,0,1,0,1)$, $(1,0,1,1,0)$, $(1,3,0,1,0)$. Замечаем, что сумма 2-го, 3-го, 5-го и 6-го векторов равна нулю по модулю 2. Это дает $b = 7272$, $c = 4928$, и мы находим, наконец, нетривиальный делитель: НОД $(7272 + 4928, 17873) = 61$. Таким образом, $17873 = 61 \cdot 293$.

УПРАЖНЕНИЯ

1. Найти представление в виде цепной дроби следующих рациональных чисел: а) 45/89; б) 55/89; в) 1,13.

2. а) Предположим, что x — вещественное число, в представлении которого в виде бесконечной цепной дроби

$$x = a + \frac{1}{a+}\,\frac{1}{a+}\,\frac{1}{a+}\,\frac{1}{a+}\cdots$$

бесконечно много раз повторяется натуральное число a. Как записать это вещественное число x в простом виде?

б) Доказать, что при $a = 1$ в части а) x есть золотое сечение, а числители и знаменатели его приближений суть числа Фибоначчи.

3. Разложить в цепную дробь число e и попытаться угадать схему образования последовательности a_i.

4. Объяснить, почему в алгоритме цепных дробей нецелесообразно включать в факторную базу такие простые числа p, что $(\frac{n}{p}) = -1$.

5. Следуя примерам 2 и 3, применить алгоритм цепных дробей для разложения на множители следующих чисел: а) 9509; б) 13561; в) 8777; г) 14429; д) 12403; е) 14527; ж) 10123; з) 12449; и) 9353; к) 25511; л) 17873.

ЛИТЕРАТУРА к § 4 ГЛАВЫ V

1. *Davenport H.* The Higher Arithmetic. 5th ed. N.Y.: Cambridge Univ. Press, 1982.
2. *Knuth D.* The Art of Computer Programming. Vol. 2. Reading: Addison-Wesley, 1973.
3. *Lehmer D. H., Powers R. E.* On factoring large numbers. — Bull. Amer. Math. Soc., 1931, v. 37, p. 770–776.
4. *Morrison M. A., Brillhart J.* A method of factoring and the factorization of F_7. — Math. Comp., 1975, v. 29, p. 183–205.
5. *Pomerance C., Wagstaff S. S., Jr.* Implementation of the continued fraction integer factoring algorithm. — In: Proceedings of the 12th Winnipeg Conference on Numerical Methods and Computing, 1983.
6. *Wunderlich M. C.* A running time analysis of Brillhart's continued fraction factoring method. — Lect. Notes Math., 1979, B. 751, S. 328–342.
7. *Wunderlich M. C.* Implementing the continued fraction factoring algorithm on parallel machines. — Math. Comp., 1985, v. 44, p. 251–260.

§ 5. Метод квадратичного решета

Метод квадратичного решета для факторизации больших целых чисел, разработанный Померанцем в начале 80-х гг., долгое время превосходил другие методы факторизации целых чисел n общего вида, не имеющих простых делителей, порядок величины которых значительно меньше $\sqrt{n}$. (Для целых n специального вида могут существовать более быстрые методы с узкой областью применимости, а для чисел n, имеющих простые делители, много меньшие $\sqrt{n}$, более быстрым является метод факторизации на эллиптических кривых, описанный в § VI. 4. См. также обсуждение метода числового решета в конце данного параграфа.)

Квадратичное решето представляет собой вариант метода факторных баз § 3. В качестве факторной базы B мы берем множество простых чисел, состоящее из $p = 2$ и всех таких нечетных простых

чисел p, не превосходящих заданной границы P (выбираемой из соображений оптимальности), что n— квадратичный вычет по модулю p. Множество целых чисел S, в котором мы ищем B-числа (напомним, что B-число — это целое, делящееся только на простые числа из B), — то же самое, что используется при факторизации Ферма (см. § 3), а именно:

$$S = \{t^2 - n \mid [\sqrt{n}] + 1 \leqslant t \leqslant [\sqrt{n}] + A\}$$

при надлежаще выбранной границе A.

Основная идея метода состоит в следующем: вместо того, чтобы брать одно за другим каждое $s \in S$ и делить его на простые числа из B, чтобы определить, является ли B-числом число s, мы берем одно за другим каждое $p \in B$ и проверяем делимость на p (и его степени) одновременно для всех $s \in S$. Слово «решето» соответствует этой идее. Известное «решето Эратосфена» можно использовать для составления списка всех простых p, не превосходящих A. Например, чтобы составить список простых чисел, не превосходящих 1000, берется список всех целых, не превосходящих 1000, и затем для каждого $p = 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31$ из него удаляются все числа, кратные p и большие p; они «проваливаются в отверстия решета, удаленные одно от другого на расстояние p». Оставшиеся числа — простые.

Изложим в общих чертах способ реализации рассматриваемого метода и затем приведем пример. Описанная ниже версия — это лишь один из возможных вариантов метода, и не обязательно самый эффективный. Кроме того, в нашем примере число n, которое нужно разложить (а также соответствующие числа в упражнениях в конце параграфа) взяты из области $\approx 10^6$, чтобы избежать работы с большими матрицами. Однако такие n слишком малы для того, чтобы продемонстрировать, как решето позволяет экономить время при нахождении большого множества B-чисел.

Итак, предположим, что дано нечетное составное число n.

1. Выбираем границы P и A порядка величины приблизительно

$$e^{\sqrt{\log n \log\log n}}.$$

Число A должно быть больше P, но не превышать сравнительно малой его степени, например, $P < A < P^2$.

С функцией $\exp\{\sqrt{\log n \log\log n}\}$, традиционно обозначаемой $L(n)$, мы уже сталкивались в этой главе. Порядок ее роста находится в промежутке между многочленом от $\log n$ и многочленом от n. Если $n \approx 10^6$, то $L(n) \approx 400$. В примерах, приведенных ниже, выберем $P = 50$, $A = 500$.

2. Для $t = [\sqrt{n}]+1, [\sqrt{n}]+2, \ldots, [\sqrt{n}]+A$ выписываем в столбец по порядку целые числа $t^2 - n$.

3. Для каждого нечетного простого числа $p \leqslant P$ проверяем условие $(\frac{n}{p}) = 1$ (см. § II. 2); если оно не выполняется, то удаляем p из факторной базы.

4. Предполагая, что p — такое нечетное простое число, что n — квадратичный вычет по модулю p (случай $p = 2$ разберем далее отдельно), решаем уравнение $t^2 \equiv n \pmod{p^\beta}$ для $\beta = 1, 2, \ldots$ методом упражнения 20 к § II. 2. Берем значения β в порядке возрастания, пока не окажется, что уравнение не имеет решений t, сравнимых по модулю p^β с каким-либо из чисел в области $[\sqrt{n}]+1 \leqslant t \leqslant [\sqrt{n}]+A$. Обозначим через β наибольшее из таких чисел, для которых в указанной области найдется число t со свойством $t^2 \equiv n \pmod{p^\beta}$. Пусть t_1 и t_2 — два решения $t^2 \equiv n \pmod{p^\beta}$ и $t_2 \equiv -t_1 \pmod{p^\beta}$ (мы не требуем, чтобы t_1 и t_2 принадлежали отрезку $[[\sqrt{n}]+1, [\sqrt{n}]+A]$).

5. При том же самом значении p просматриваем список значений $t^2 - n$, полученный в п. 2. В столбце, соответствующем p, ставим 1 против всех значений $t^2 - n$, для которых t отличается от t_1 на некоторое кратное p, после этого заменяем 1 на 2 для всех таких значений $t^2 - n$, что t отличается от t_1 на кратное p^2, затем заменяем 2 на 3 у всех значений $t^2 - n$, для которых t отличается от t_1 на кратное p^3, и так далее до p^β. Затем делаем то же самое с t_2 вместо t_1. Наибольшим числом, которое появляется в этом столбце, будет β.

6. Когда производим действия в п. 5, каждый раз, когда ставим 1 или заменяем 1 на 2, 2 на 3 и т. д., делим соответствующее число $t^2 - n$ на p и сохраняем полученный результат.

7. В столбце под $p = 2$ при $n \not\equiv 1 \pmod 8$ просто ставим 1 против $t^2 - n$ с нечетным t и делим соответствующее $t^2 - n$ на 2. При $n \equiv 1 \pmod 8$ решаем уравнение $t^2 \equiv n \pmod{2^\beta}$ и продолжаем в точности так же, как в случае нечетного p (за исключением того, что при $\beta \geqslant 3$ уравнение будет иметь 4 различных решения t_1, t_2, t_3, t_4).

8. Когда указанные действия будут проведены для всех простых чисел, не превосходящих P, отбросим все $t^2 - n$, кроме тех, которые обратились в 1 после деления на все степени p, не превосходящих P. Тогда получится таблица того же вида, что в примере 9 из § 3, в которой столбец, помеченный b_i, будет содержать все такие значения t из интервала $[[\sqrt{n}]+1, [\sqrt{n}]+A]$, что $t^2 - n$ есть B-число, а остальные столбцы будут соответствовать тем значениям $p \leqslant P$, для которых n — квадратичный вычет.

9. Оставшаяся часть процедуры в точности совпадает с процедурой из § 3.

Пример. Давайте попытаемся разложить на множители $n = 1042387$, выбрав границы $P = 50$ и $A = 500$. Здесь $[\sqrt{n}] = 1020$. Наша факторная база состоит из 8 простых чисел $\{2, 3, 11, 17, 19, 23, 43, 47\}$, для которых 1042387 — квадратичный вычет. Так как $n \not\equiv 1 \pmod 8$, то в столбце, соответствующем $p = 2$, стоят лишь 0 и 1, причем 1 проставляется во всех случаях, когда t нечетно, $1021 \leqslant t \leqslant 1520$.

Опишем подробно, как образуется столбец для $p = 3$. Мы хотим найти решение $t_1 = t_{1,0} + t_{1,1}3 + t_{1,2}3^2 + \cdots + t_{1,\beta-1}3^{\beta-1}$ сравнения $t^2 \equiv 1042387 \pmod{3^\beta}$, где $t_{1,j} \in \{0, 1, 2\}$ (за другое решение t_2 можно принять $t_2 = 3^\beta - t_1$). Мы можем, очевидно, взять $t_{1,0} = 1$. (Для каждого из наших 8 простых чисел первый шаг — решение сравнения $t^2 \equiv 1042387 \pmod p$ — можно сделать быстро методом проб и ошибок; если бы мы работали с бóльшими простыми числами, мы могли бы использовать процедуру, описанную в конце § II.2.) Затем мы работаем по модулю 9: $(1 + 3t_{1,1})^2 \equiv 1042387 \equiv 7 \pmod 9$, откуда $6t_{1,1} \equiv 6 \pmod 9$, т.е. $2t_{1,1} \equiv 2 \pmod 3$ и $t_{1,1} = 1$. Далее, по модулю 27: $(1 + 3 + 9t_{1,2})^2 \equiv 1042387 \equiv 25 \pmod{27}$, т.е. $16 + 18t_{1,2} \equiv 25 \pmod{27}$ и $2t_{1,2} \equiv 1 \pmod 3$, так что $t_{1,2} = 2$. Затем по модулю 81: $(1 + 3 + 18 + 27t_{1,3})^2 \equiv 1042387 \equiv 79 \pmod{81}$, откуда следует, что $t_{1,3} = 0$. Продолжая действовать так же до 3^7, мы находим решение (в обозначениях § I.1 для чисел, записываемых в троичной системе): $t_1 \equiv (210211)_3 \pmod{3^7}$ и $t_2 \equiv (2012012)_3 \pmod{3^7}$. Однако в интервале между 1021 и 1520 нет числа t, сравнимого с t_1 или t_2 по модулю 3^7. Таким образом, мы имеем $\beta = 6$, и можно взять $t_1 = (210211)_3 = 589 \equiv 1318 \pmod{3^6}$ и $t_2 = 3^6 - t_1 = 140 \equiv 1112 \pmod{3^5}$ (заметим, что в отрезке $[1021, 1520]$ нет числа t, сравнимого с t_2 по модулю 3^6).

Построим теперь наше «решето» для $p = 3$. Начиная с 1318, делаем «прыжки» на 3 вниз, пока не достигнем 1021, и такие же «прыжки» вверх, пока не достигнем 1519; каждый раз ставим 1 в столбец, если $t^2 - n$ делится на 3, и записываем результат деления (на самом деле для нечетного t число, которое мы делим на 3, равно $(t^2 - n)/2$, так как мы уже делили $t^2 - n$ пополам, когда формировали столбец из 0 и 1 для $p = 2$). Затем мы делаем то же самое, «прыгая» на 9, изменяя каждый раз 1 на 2 в столбце под 3, деля еще раз на 3 то, что осталось от соответствующего $t^2 - n$, и записывая результат. Проделываем аналогичную процедуру со скачками на 27, 81, 243, 729 (для 729 скачка не существует — мы только изменяем 5 на 6 в графе, соответствующей 1318, и делим еще раз на 3 то, что осталось от $1318^2 - 1042387$). Наконец, мы проделываем те же шаги с $t_2 = 1112$ вместо $t_1 = 1318$, на

сей раз останавливаясь на скачках величины 243.

После прохождения этой процедуры для остальных 6 чисел нашей факторной базы мы получаем таблицу размера 500×8 из показателей степеней, в которой каждая строка соответствует значению t от 1021 до 1520. Теперь мы отбрасываем все строки, для которых $t^2 - n$ не превратилось в 1 при повторном делении на степени p в процессе формирования таблицы, т. е. мы берем только строки, для которых $t^2 - n$ есть B-число. В рассматриваемом примере с $n = 1042387$ останется следующая таблица (с прочерками на месте нулевых показателей степени).

t	$t^2 - n$	2	3	11	17	19	23	43	47
1021	54	1	3	–	–	–	–	–	–
1027	12342	1	1	2	1	–	–	–	–
1030	18513	–	2	2	1	–	–	–	–
1061	83334	1	1	–	1	1	–	1	–
1112	194157	–	5	–	1	–	–	–	1
1129	232254	1	3	1	1	–	1	–	–
1148	275517	–	2	3	–	–	1	–	–
1175	338238	1	2	–	–	1	1	1	–
1217	438702	1	1	1	2	–	1	–	–
1390	889713	–	2	2	–	1	–	1	–
1520	1268013	–	1	–	1	–	2	–	1

Продолжая так же, как мы делали в примере 9 из § 3, ищем теперь соотношения по модулю 2 между строками этой матрицы, т. е., двигаясь сверху вниз, ищем такие наборы строк, суммы элементов которых в каждом столбце четны. Первый обнаруживаемый набор такого типа — это первые три строки, сумма которых — удвоенная строка 1 3 2 1 – – – – . Таким образом, мы получаем сравнение

$$(1021 \cdot 1027 \cdot 1030)^2 \equiv (2 \cdot 3^3 \cdot 11^2 \cdot 17)^2 \pmod{1042387}.$$

Однако, хотя нам и повезло быстро найти подмножество строк, линейно зависимых по модулю 2, эта удача неполная: оба возводимые в квадрат числа в последнем сравнении сравнимы с 111078 по модулю 1042387, и мы получаем лишь тривиальную факторизацию. Продвигаясь вниз по таблице, мы находим еще несколько наборов линейно зависимых строк, которые также не дают нетривиальной факторизации. Наконец, когда мы уже почти готовы сдаться и начать все сначала с увеличенным A, мы замечаем, что последняя строка, соответствующая нашему самому последнему значению t, зависит от предшествующих строк. А именно, по модулю 2 она равна 5-й строке. Это дает нам $(1112 \cdot 1520)^2 \equiv (3^3 \cdot 17 \cdot 23 \cdot 47)^2 \pmod{1042387}$,

т. е. $647853^2 \equiv 496179^2 \pmod{1042387}$, и мы получаем нетривиальный делитель: НОД $(647853 - 496179, 1042387) = 1487$.

Используя некоторые правдоподобные предположения, можно показать, что математическое ожидание времени работы метода разложения при помощи квадратичного решета имеет порядок

$$O(e^{(1+\varepsilon)\sqrt{\log n \log\log n}}) \quad \text{для любого} \quad \varepsilon > 0.$$

Метод использует очень большой объем памяти (также порядка $\exp\{C\sqrt{\log n \log\log n}\}$). Подробное обсуждение характеристик алгоритма факторизации при помощи квадратичного решета (и некоторых других) можно найти в статье Померанца из списка литературы к этому параграфу.

Решето в числовом поле. До недавнего времени все оценки времени работы лучших универсальных алгоритмов разложения на множители имели вид

$$\exp\{O(\sqrt{\log n \log\log n})\}.$$

Некоторые даже предполагали, что эта функция n является естественной нижней границей для времени работы. Однако в течение последних нескольких лет был разработан новый метод, названный *решетом в числовом поле*. Эвристическая оценка времени его работы значительно лучше (асимптотически):

$$\exp\{O((\log n)^{1/3}(\log\log n)^{2/3})\}.$$

Практически он оказался самым быстрым методом для разложения на множители чисел, которые имеют величину, максимально допустимую для современных (1994 г.) методов факторизации, или немного превышают ее, т. е. для чисел, имеющих в десятичной системе счисления немного больше 150 знаков.

Факторизационный алгоритм решета в числовом поле во многом подобен ранее рассмотренным алгоритмам, в которых проводится поиск наборов сравнений, дающих соотношение вида $x^2 \equiv y^2 \pmod{n}$. Однако в нем используется «факторная база» в кольце целых чисел некоторого надлежаще выбранного поля алгебраических чисел. Таким образом, наряду с основной техникой квадратичного решета, в этом методе факторизации применяется также теория алгебраических чисел. Видимо, это наиболее сложный из известных факторизационных алгоритмов. Мы дадим только его набросок.

Основные требования алгоритма можно кратко описать следующим образом. Если нужно разложить на множители число n, то нужно

выбрать некоторую степень d и выразить n как значение при некотором целом m неприводимого нормированного многочлена степени d:

$$n = f(m) = m^d + a_{d-1}m^{d-1} + a_{d-2}m^{d-2} + \cdots + a_1 m + a_0,$$

где m и a_k — целые числа, порядок величины которых — $O(n^{1/d})$. Один из способов найти такой многочлен — это взять в качестве m целую часть корня степени d из n и затем разложить n по основанию m. Для 125-значных чисел анализ алгоритма показывает, что d должно равняться 5, и потому m и коэффициенты должны быть примерно 25-значными.

Решето числового поля тогда отыскивает (с помощью того же процесса, что и в случае квадратичного решета) максимально возможное число таких пар (a, b), что как число $a + bm$, так и число

$$b^d f(-a/b) = (-a)^d + a_{d-1}(-a)^{d-1}b + a_{d-2}(-a)^{d-2}b^2 + \cdots - a_1 a b^{d-1} + a_0 b^d$$

— гладкие относительно данной факторной базы (т. е. делятся лишь на простые числа, принадлежащие факторной базе). Подробности того, как это делается и как это приводит к факторизации n, можно найти в книге [3] из списка литературы к данному параграфу. Чтобы процедура была успешной, нужно, чтобы доля гладких чисел среди значений многочлена f была приблизительно такой же, как доля гладких чисел среди всех чисел той же длины. Хотя это выглядит правдоподобно и выполняется во всех примерах, которые были просчитаны, тем не менее, утверждение выглядит чрезвычайно трудным для доказательства. Поскольку оценка времени работы алгоритма основана на этом недоказанном предположении, она является эвристической. Хотя это обстоятельство не очень существенно для практической факторизации чисел, оно указывает на некоторые важные нерешенные задачи теоретического анализа асимптотической сложности факторизации.

Автор хотел бы поблагодарить Дж. Булера (Joe Buhler), предоставившего для настоящей книги это краткое описание решета над числовым полем.

УПРАЖНЕНИЯ

1. В приведенном примере найти все отношения линейной зависимости по модулю 2 между строками матрицы и показать, что если $P = 50$ и $A \leqslant 499$, то получить нетривиальную факторизацию 1042387 этим методом невозможно.

2. Пусть $n \to \infty$. Предположим, что P и A всегда выбираются величинами одинакового порядка (например, что $c_1 \leqslant (\log A)/(\log P) \leqslant c_2$ для некоторых положительных констант c_1 и c_2). Какой из шагов 1–7 изложенной выше версии метода квадратичного решета является асимптотически наиболее трудоемким? Дать в терминах O-большого оценку числа двоичных операций на этом шаге.

3. Использовать метод этого параграфа с $P = 50$ и $A = 500$, чтобы разложить на множители: а) 1046603; б) 1059691; в) 998771.

ЛИТЕРАТУРА к § 5 ГЛАВЫ V

1. *Caron T., Silverman R.* Parallel implementation of the quadratic sieve. — J. Supercomputing, 1988, v. 1, p. 273–290.
2. *Gerver J. L.* Factoring large numbers with a quadratic sieve. — Math. Comp., 1983, v. 41, p. 287–294.
3. The Development of the Number Field Sieve./ Ed. by A. Lenstra, H. W. Lenstra, Jr.. Berlin–Heidelberg etc.: Springer, 1993.
4. *Lenstra H. W., Jr., Pomerance C.* A rigorous time bound for factoring integers. — J. Amer. Math. Soc., 1992, v. 5, p. 483–516.
5. *Pomerance C.* Analysis and comparison of some integer factoring algorithms. — In: Computational Methods in Number Theory./ Ed. by H. W. Lenstra, Jr., R. Tijdeman. Amsterdam: Mathematisch Centrum, 1982, p. 89–139.
6. *Pomerance C.* Factoring. — Crypt. Comput. Number Theory, Proc. Symp. Appl. Math., 1990, v. 42, p. 27–47.

ГЛАВА VI

ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ

В последнее время одна из областей теории чисел и алгебраической геометрии — эллиптические кривые (точнее, теория эллиптических кривых над конечными полями) — нашла применение в криптографии. Основная причина этого состоит в том, что эллиптические кривые над конечными полями доставляют неисчерпаемый источник конечных абелевых групп, которые (даже если они велики) удобны для вычислений и обладают богатой структурой. Ранее (в § IV.3) мы работали с мультипликативными группами полей. Во многих отношениях эллиптические кривые — естественный аналог этих групп, но более удобный, так как существует бóльшая свобода в выборе эллиптической кривой, чем в выборе конечного поля.

Начнем с изложения основных определений и свойств эллиптических кривых. Мы ограничимся минимальным числом основных фактов, необходимых для понимания приложений к криптографии (см. §§ 2–4), уделяя больше внимания примерам и конкретным описаниям и меньше заботясь о доказательствах и общности. Более систематическое изложение этого предмета можно найти по ссылкам в конце § 1.

§ 1. Основные факты

В этом параграфе мы предполагаем, что K — поле: либо поле $\mathbf{R}$ вещественных чисел, либо поле $\mathbf{Q}$ рациональных чисел, либо поле $\mathbf{C}$ комплексных чисел, либо поле $\mathbf{F}_q$ из $q = p^r$ элементов.

О п р е д е л е н и е. Пусть K — поле характеристики, отличной от 2, 3, и x^3+ax+b (где $a, b \in K$) — кубический многочлен без кратных корней. *Эллиптическая кривая над* K — это множество точек (x, y), $x, y \in K$, удовлетворяющих уравнению

$$y^2 = x^3 + ax + b, \tag{1}$$

вместе с единственным элементом, обозначаемым O и называемым «точка в бесконечности» (о ней подробнее будет сказано ниже).

Если K — поле характеристики 2, то *эллиптическая кривая над* K — это множество точек, удовлетворяющих уравнению либо типа

$$y^2 + cy = x^3 + ax + b, \tag{2а}$$

либо типа

$$y^2 + xy = x^3 + ax^2 + b \tag{2б}$$

(здесь кубические многочлены в правых частях могут иметь кратные корни), вместе с «точкой в бесконечности» O.

Если K — поле характеристики 3, то *эллиптическая кривая над* K — это множество точек, удовлетворяющих уравнению

$$y^2 = x^3 + ax^2 + bx + c \tag{3}$$

(где кубический многочлен справа не имеет кратных корней), вместе с «точкой в бесконечности» O.

З а м е ч а н и я. 1. Имеется общая форма уравнения эллиптической кривой, которая применима при любом поле: $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$; в случае, когда $\operatorname{char} K \neq 2$, ее можно привести к виду $y^2 = x^3 + ax^2 + bx + c$ (или к виду $y^2 = x^3 + bx + c$, если $K > 3$). В случае, когда поле K имеет характеристику 2, это уравнение преобразуется либо к виду (2а), либо к виду (2б).

2. Если $F(x, y) = 0$ — неявное уравнение, выражающее y как функцию x в (1) (или в (2), (3)), т. е. $F(x, y) = y^2 - x^3 - ax - b$ (или $F(x, y) = y^2 + cy + x^3 + ax + b, y^2 + xy + x^3 + ax + b, y^2 - x^3 - ax^2 - bx - c$), то точка (x, y) этой кривой называется *неособенной* (или *гладкой*) точкой, если, по крайней мере, одна из частных производных $\partial F/\partial x$, $\partial F/\partial y$ в этой точке не равна нулю. (Производные многочленов можно определить обычными формулами над любым полем; см. 5-й абзац в начале главы II.)

Нетрудно показать, что условие отсутствия кратных корней у кубических многочленов в правой части в (1) и (3) эквивалентно требованию, чтобы все точки кривой были неособенными.

Эллиптические кривые над вещественными числами. Перед обсуждением конкретных примеров эллиптических кривых над разными полями мы отметим чрезвычайно важное свойство множества точек эллиптической кривой: они образуют абелеву группу. Чтобы объяснить наглядно, как это получается, мы временно будем полагать, что $K = \mathbf{R}$, т. е. что эллиптическая кривая — обычная плоская кривая (с добавлением еще одной точки O «в бесконечности»).

О п р е д е л е н и е. Пусть E — эллиптическая кривая над вещественными числами, и пусть P и Q — две точки на E. Определим точки $-P$ и $P+Q$ по следующим правилам.

1. Если P — точка в бесконечности O, то $-P=O$ и $P+Q=Q$, т. е. O — тождественный элемент по сложению («нулевой элемент») группы точек. В следующих пунктах предполагается, что ни P, ни Q не являются точками в бесконечности.

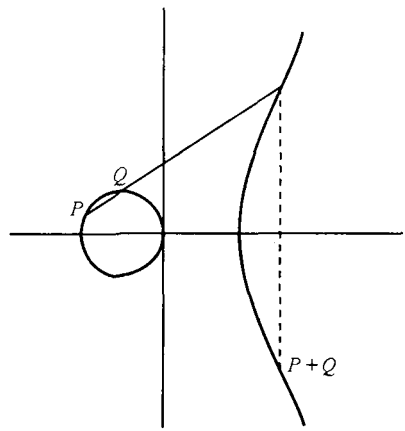
2. Точки $P=(x,y)$ и $-P$ имеют одинаковые x-координаты, а их y-координаты различаются только знаком, т. е. $-(x,y)=(x,-y)$. Из (1) сразу следует, что $(x,-y)$ — также точка на E.

3. Если P и Q имеют различные x-координаты, то прямая $l=\overline{PQ}$ имеет с E еще в точности одну точку пересечения R (за исключением двух случаев: когда она оказывается касательной в P, и мы тогда полагаем $R=P$, или касательной в Q, и мы тогда полагаем $R=Q$). Определяем теперь $P+Q$ как точку $-R$, т. е. как отражение от оси x третьей точки пересечения. Геометрическое построение, дающее $P+Q$, приводится ниже в примере 1.

4. Если $Q=-P$ (т. е. x-координата Q та же, что и у P, а y-координата отличается лишь знаком), то полагаем $P+Q=O$ (точке в бесконечности; это является следствием правила 1).

5. Остается возможность $P=Q$. Тогда считаем, что l — касательная к кривой в точке P. Пусть R — единственная другая точка пересечения l с E. Полагаем $P+Q=-R$ (в качестве R берем P, если касательная прямая в P имеет «двойное касание», т. е. если P есть точка перегиба кривой).

П р и м е р 1. На рисунке справа изображены эллиптическая кривая $y^2=x^3-x$ в плоскости xy и типичный случай сложения точек P и Q. Чтобы найти $P+Q$, проводим прямую $\overline{PQ}$ и в качестве $P+Q$ берем точку, симметричную относительно оси x третьей точке, определяемой пересечением прямой PQ и кривой. Если бы P совпадала с Q, т. е. если бы нам нужно было найти $2P$, мы использовали бы касательную к кривой в P; тогда точка $2P$ симметрична третьей точке, в которой эта касательная пересекает кривую.



Теперь мы покажем, почему существует в точности еще одна точка, где прямая l, проходящая через P и Q, пересекает кривую; заодно мы выведем формулу для координат этой третьей точки и тем самым — для координат $P+Q$.

Пусть (x_1, y_1), (x_2, y_2) и (x_3, y_3) обозначают координаты соответственно P, Q и $P+Q$. Мы хотим выразить x_3, y_3 через x_1, y_1, x_2, y_2.

Предположим, что мы находимся в ситуации п. 3 определения $P+Q$, и пусть $y = \alpha x + \beta$ есть уравнение прямой, проходящей через P и Q (в этой ситуации она не вертикальна). Тогда $\alpha = (y_2 - y_1)/(x_2 - x_1)$ и $\beta = y_1 - \alpha x_1$. Точка на l, т. е. точка $(x, \alpha x + \beta)$, лежит на эллиптической кривой тогда и только тогда, когда $(\alpha x + \beta)^2 = x^3 + ax + b$. Таким образом, каждому корню кубического многочлена $x^3 - (\alpha x + \beta)^2 + ax + b$ соответствует точка пересечения. Мы уже знаем, что имеется два корня x_1 и x_2, так как $(x_1, \alpha x_1 + \beta)$, $(x_2, \alpha x_2 + \beta)$ — точки P, Q на кривой. Так как сумма корней нормированного многочлена равна взятому с обратным знаком коэффициенту при второй по старшинству степени многочлена, то в нашем случае третий корень — это $x_3 = \alpha^2 - x_1 - x_2$. Тем самым получаем выражение для x_3, и, следовательно, $P + Q = (x_3, -(\alpha x_3 + \beta))$ или, в терминах x_1, y_1, x_2, y_2,

$$\begin{aligned} x_3 &= \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2, \\ y_3 &= -y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_1 - x_3). \end{aligned} \tag{4}$$

Ситуация в п. 5 аналогична, только теперь α — производная dy/dx в P. Дифференцирование неявной функции, заданной уравнением (1), приводит к формуле $\alpha = (3x_1^2 + a)/(2y_1)$, и мы получаем следующие формулы для координат удвоенной точки P:

$$\begin{aligned} x_3 &= \left(\frac{3x_1^2 + a}{2y_1}\right)^2 - 2x_1, \\ y_3 &= -y_1 + \frac{3x_1^2 + a}{2y_1}(x_1 - x_3). \end{aligned} \tag{5}$$

Пример 2. На эллиптической кривой $y^2 = x^3 - 36x$ пусть $P = (-3, 9)$ и $Q = (-2, 8)$. Найти $P + Q$ и $2P$.

Решение. Подстановка $x_1 = -3, y_1 = 9, x_2 = -2, y_2 = 8$ в первое из уравнений (4) дает $x_3 = 6$; тогда второе из уравнений (4) дает $y_3 = 0$. Далее, подставляя $x_1 = -3, y_1 = 9, a = -36$ в первое из уравнений (5), получаем для x-координаты точки $2P$ значение $25/4$, а второе из уравнений (5) дает для y-координаты значение $-35/8$.

Существует несколько способов доказать, что принятое выше определение $P+Q$ превращает множество точек на эллиптической кривой в абелеву группу. Можно использовать результаты из проективной геометрии, из комплексного анализа двоякопериодических функций или алгебраическое доказательство, использующее теорию дивизоров на кривых. Доказательства каждого из этих типов можно найти по ссылкам в конце параграфа.

Если n — целое число, то, как и в любой абелевой группе, nP обозначает сумму n точек P при $n>0$ и сумму $|n|$ точек $-P$, если $n \leqslant 0$.

Мы пока что мало сказали о «точке в бесконечности» O. По определению, это — тождественный элемент группового закона. На рисунке (см. выше) следует себе представлять ее расположенной на оси y в предельном направлении, определяемом все более «крутыми» касательными к кривой. Она является «третьей точкой пересечения» с кривой для любой вертикальной прямой: такая прямая пересекается с кривой в точках вида $(x_1, y_1), (x_1, -y_1)$ и O. Мы изложим сейчас более естественный способ введения точки O.

Под *проективной плоскостью* мы понимаем множество классов эквивалентности троек (X, Y, Z) (не все компоненты равны нулю), при этом две тройки называются эквивалентными, если одна из них — скалярное кратное другой, т. е. $(\lambda X, \lambda Y, \lambda Z) \sim (X, Y, Z)$. Такой класс эквивалентности называется *проективной точкой*. Если проективная точка имеет ненулевую компоненту Z, то существует, причем только одна, тройка в ее классе эквивалентности, имеющая вид $(x, y, 1)$: просто полагаем $x = X/Z, y = Y/Z$. Тем самым проективную плоскость можно представить как объединение всех точек (x, y) обычной («аффинной») плоскости с точками, для которых $Z = 0$. Эти последние точки составляют то, что называется *бесконечно удаленной прямой*; наглядно ее можно себе представить на плоскости как «горизонт». Любому алгебраическому уравнению $F(x, y) = 0$ кривой в аффинной плоскости отвечает уравнение $\widetilde{F}(X, Y, Z) = 0$, которому удовлетворяют соответствующие проективные точки: нужно заменить x на X/Z, y — на Y/Z и умножить на подходящую степень Z, чтобы освободиться от знаменателей. Например, если применить эту процедуру к аффинному уравнению (1) эллиптической кривой, то получится «проективное уравнение» $Y^2Z = X^3 + aXZ^2 + bZ^3$. Этому уравнению удовлетворяют все проективные точки (X, Y, Z) с $Z \neq 0$, для которых соответствующие аффинные точки (x, y), где $x = X/Z, y = Y/Z$, удовлетворяют (1). Помимо них, какие еще точки бесконечно удаленной прямой удовлетворяют последнему уравнению? Если положить

в уравнении $Z = 0$, то уравнение примет вид $X^3 = 0$, т. е. $X = 0$. Но единственный класс эквивалентности троек с $X = 0, Z = 0$ — это класс тройки $(0, 1, 0)$. Это и есть точка, которую мы обозначили O; она лежит на пересечении оси y с бесконечно удаленной прямой.

Эллиптические кривые над комплексными числами. Алгебраические формулы (4)–(5) для сложения точек на эллиптической кривой над вещественными числами на самом деле имеют смысл над любым полем. (В полях характеристики 2 или 3 можно вывести аналогичные равенства, исходя из уравнений (2) или (3).) Можно показать, что эти формулы выражают закон абелевой группы на эллиптической кривой над любым полем.

В частности, пусть E — эллиптическая кривая, определенная над полем $\mathbf{C}$ комплексных чисел, т. е. E — множество пар (x, y) комплексных чисел, удовлетворяющих уравнению (1), вместе с точкой в бесконечности O. Мы называем E «кривой», однако с точки зрения обычных геометрических представлений она двумерна, т. е. представляет собой поверхность в 4-мерном вещественном пространстве, координатами в котором являются действительные и мнимые части x и y. Покажем теперь, как можно наглядно представить себе E в качестве поверхности.

Пусть L — *решетка* в комплексной плоскости. Это означает, что L — абелева группа, состоящая из всех целочисленных линейных комбинаций двух данных комплексных чисел ω_1 и ω_2 (где ω_1 и ω_2 «заметают» плоскость, т. е. не лежат на одной прямой, проходящей через начало координат): $L = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$. Например, если $\omega_1 = 1, \omega_2 = i$, то L — множество всех гауссовых целых чисел, т. е. квадратная сетка, состоящая из всех комплексных чисел с целыми действительными и мнимыми частями.

Если задана эллиптическая кривая (1) над комплексными числами, то, как оказывается, существуют решетка L и функция комплексного переменного, называемая «$\wp$-функцией Вейерштрасса» и обозначаемая $\wp_L(z)$, со следующими свойствами:

1. $\wp(z)$ аналитична всюду, кроме точек L, в каждой из которых имеет полюс второго порядка;

2. $\wp(z)$ удовлетворяет дифференциальному уравнению $\wp'^2 = \wp^3 + a\wp + b$ и, следовательно, при любом $z \notin L$ точка $(\wp(z), \wp'(z))$ лежит на эллиптической кривой E;

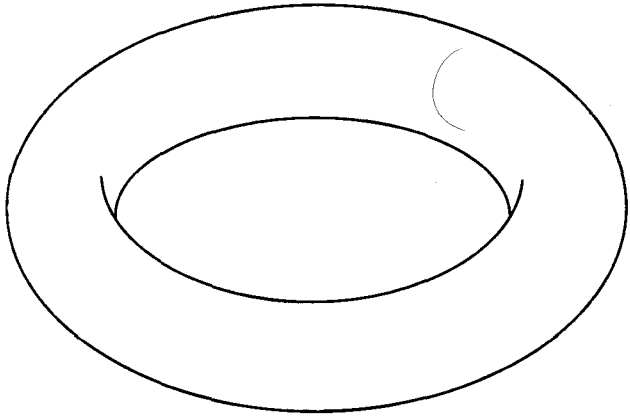
3. два комплексных числа z_1 и z_2 дают одну и ту же точку $(\wp(z), \wp'(z))$ на E тогда и только тогда, когда $z_1 - z_2 \in L$;

4. отображение, которое любой точке $z \notin L$ ставит в соответствие точку $(\wp(z), \wp'(z))$ на E, а любой точке $z \in L$ — точку в бесконечности

O, дает взаимно однозначное соответствие между E и факторгруппой $\mathbf{C}/L$ комплексной плоскости по подгруппе L;

5. это взаимно однозначное соответствие есть изоморфизм абелевых групп, иными словами, если z_1 соответствует точке $P \in E$, а z_2 — точке $Q \in E$, то $z_1 + z_2$ соответствует точке $P + Q$.

Таким образом, можно представлять себе абелеву группу E как комплексную плоскость «по модулю» некоторой решетки. Чтобы эту последнюю группу изобразить наглядно, заметим, что у каждого класса эквивалентности $z + L$ существует один и только один представитель в «фундаментальном параллелограмме», состоящем из комплексных чисел вида $a\omega_1 + b\omega_2$, $0 \leqslant a, b < 1$ (если, например, L — гауссовы числа, то фундаментальный параллелограмм — это единичный квадрат). Так как разность между противоположными точками на параллельных сторонах границы параллелограмма есть точка решетки, они равны в $\mathbf{C}/L$, и их можно считать «склеенными». Наглядно это означает, что мы сгибаем параллелограмм так, чтобы одна из сторон соприкоснулась с противоположной (получая при этом часть цилиндра), и затем, вновь сгибая полученную цилиндрическую трубку, склеиваем противоположные окружности — и получаем «тор» (бублик), изображенный ниже.



Как группа, тор есть произведение двух экземпляров группы окружности, т. е. его точки можно параметризовать парой углов (α, β). (Точнее, если тор получен из решетки $L = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$, то следует представить элемент из $\mathbf{C}/L$ в виде $a\omega_1 + b\omega_2$, полагая $a = 2\pi\alpha$, $b = 2\pi\beta$.) Таким образом, можно рассматривать эллиптическую кривую над комплексными числами как двумерное обобщение окружности в вещественной плоскости. Фактически эта аналогия идет значительно дальше, чем может показаться. «Эллиптические функции» (которые показывают, как по точке $(x, y) \in E$ найти то комплексное число

z, для которого $(x, y) = (\wp(z), \wp'(z)))$, как оказывается, имеют свойства, аналогичные свойствам известной функции Arcsin (которая показывает, как найти вещественное число, которое соответствует точке единичной окружности при «наматывании» вещественной прямой на окружность). При рассмотрении эллиптических кривых с точки зрения теории алгебраических чисел обнаруживается глубокая аналогия между координатами «точек, делящих эллиптические кривые на n частей» (т. е. таких точек P, что $nP = O$) и точками, делящими на n частей единичную окружность (которые соответствуют корням степени n из единицы в комплексной плоскости). Более подробные сведения об этом, а также об определении $\wp$-функции Вейерштрасса и ее свойствах можно найти по ссылкам в конце параграфа.

Эллиптические кривые над рациональными числами. Если в уравнении (1) a и b — рациональные числа, то естественно рассматривать рациональные решения (x, y), т. е. эллиптическую кривую над полем $\mathbf{Q}$ рациональных чисел. Теория эллиптических кривых над рациональными числами очень обширна. Было доказано, что соответствующие абелевы группы являются конечно порожденными (теорема Морделла). Это означает, что каждая из таких групп есть сумма конечной «подгруппы кручения» (точек конечного порядка) и подгруппы, порожденной конечным числом точек бесконечного порядка. Число (минимальное) образующих бесконечной части называется *рангом* r; оно равно нулю тогда и только тогда, когда вся группа конечна. Изучение ранга r и других свойств группы точек эллиптической кривой над $\mathbf{Q}$ связано со многими интересными вопросами теории чисел и алгебраической геометрии. Например, известный с древних времен вопрос «Существует ли прямоугольный треугольник с рациональными сторонами, площадь которого равна данному целому n?» эквивалентен следующему: «Верно ли, что ранг эллиптической кривой $y^2 = x^3 - n^2x$ больше нуля?» Случай $n = 6$ и прямоугольного треугольника со сторонами 3, 4 и 5 соответствует точке $P = (-3, 9)$ из примера 2, которая является точкой бесконечного порядка на эллиптической кривой $y^2 = x^3 - 36x$. За более подробной информацией об этом предмете мы вновь отсылаем читателя к литературе в конце параграфа.

Точки конечного порядка. *Порядком* N точки P на эллиптической кривой называется такое наименьшее натуральное число, что $NP = O$; конечно, такого конечного N может и не существовать. Часто требуется найти точки конечного порядка на эллиптической кривой, в особенности, на эллиптических кривых, определенных над полем $\mathbf{Q}$.

Пример 3. Найти порядок точки $P = (2,3)$ на $y^2 = x^3 + 1$.

Решение. Применяя (5), находим, что $2P = (0,1)$, и вновь с помощью (5), что $4P = 2(2P) = (0,-1)$. Поэтому $4P = -2P$ и, следовательно, $6P = O$. Тем самым порядок P может быть равен 2, 3 или 6. Но $2P = (0,1) \neq O$, а если бы P имела порядок 3, то было бы $4P = P$, что неверно. Итак, P имеет порядок 6.

Эллиптические кривые над конечным полем. До конца параграфа предполагаем, что K — конечное поле $\mathbf{F}_q$ с $q = p^r$ элементами. Пусть E — эллиптическая кривая, определенная над $\mathbf{F}_q$. Если $p = 2$ или 3, то E задается уравнением вида (2) или (3) соответственно.

Легко усмотреть, что эллиптическая кривая может иметь не более $2q+1$ различных $\mathbf{F}_q$-точек, т.е. точку в бесконечности и не более чем $2q$ пар (x,y), $x, y \in \mathbf{F}_q$, удовлетворяющих (1) (или (2) или (3), если $p = 2$ или 3). А именно, для каждого из q возможных значений x имеется не более двух значений y, удовлетворяющих (1).

Но так как лишь у половины элементов $\mathbf{F}_q^*$ имеются квадратные корни, естественно ожидать (если бы $x^3 + ax + b$ были случайными элементами поля), что количество $\mathbf{F}_q$-точек примерно вдвое меньше этого числа. Точнее, пусть χ — квадратичный характер $\mathbf{F}_q$. Это — отображение, переводящее $x \in \mathbf{F}_q^*$ в 1 или -1 в зависимости от того, является или нет элемент x квадратом элемента из $\mathbf{F}_q$ (полагаем также $\chi(0) = 0$). Например, если q — это простое число p, то $\chi(x) = \left(\frac{x}{p}\right)$ есть символ Лежандра (см. § II.2). В любом случае число решений $y \in \mathbf{F}_q$ уравнения $y^2 = u$ равно $1 + \chi(u)$ и, значит, число решений (1) (с учетом точки в бесконечности) равно

$$1 + \sum_{x \in \mathbf{F}_q} (1 + \chi(x^3 + ax + b)) = q + 1 + \sum_{x \in \mathbf{F}_q} \chi(x^3 + ax + b). \qquad (6)$$

Следует ожидать, что $\chi(x^3 + ax + b)$ одинаково часто принимает значения $+1$ и -1. Вычисление суммы очень похоже на «случайное блуждание», когда мы подбрасываем монету q раз, продвигаясь на шаг вперед, если выпал герб, и назад — если решетка. В теории вероятностей подсчитывается, что после q бросаний случайное блуждание оказывается на расстоянии порядка $\sqrt{q}$ от исходного положения. Сумма $\sum \chi(x^3 + ax + b)$ ведет себя в чем-то аналогично случайному блужданию. Точнее, удалось доказать, что она ограничена величиной $2\sqrt{q}$. Этот результат — теорема Хассе; ее доказательство см. в § V.1 указанной в списке литературы книги Силвермена (J. Silverman) об эллиптических кривых.

Теорема Хассе. *Пусть N — число $\mathbf{F}_q$-точек на эллиптической кривой, определенной над $\mathbf{F}_q$. Тогда*

$$|N-(q+1)| \leqslant 2\sqrt{q}.$$

В дополнение к числу N элементов на эллиптической кривой над $\mathbf{F}_q$ нам бывает нужно знать строение этой абелевой группы. Она — не обязательно циклическая, но можно показать, что она — всегда произведение двух циклических групп. Это означает, что группа изоморфна произведению p-примарных групп вида $\mathbf{Z}/p^{\alpha}\mathbf{Z} \times \mathbf{Z}/p^{\beta}\mathbf{Z}$, где произведение берется по всем простым делителям N (здесь $\alpha \geqslant 1, \beta \geqslant 0$). Под *типом* абелевой группы $\mathbf{F}_q$-точек на E мы понимаем список $(\ldots, p^{\alpha}, p^{\beta}, \ldots)_{p|N}$ порядков циклических p-примарных сомножителей в упомянутом представлении в виде произведения (если $\beta = 0$, p^{β} опускаем). Найти тип не всегда легко.

Пример 4. Найти тип для кривой $y^2 = x^3 - x$ над $\mathbf{F}_{71}$.

Решение. Находим сначала число точек N. Заметим, что в сумме (6) члены для x и для $-x$ сокращаются друг с другом: $\chi((-x)^3-(-x)) = \chi(-1)\chi(x^3-x) = -\chi(x^3-x)$, так как $71 \equiv 3 \pmod 4$, и потому $\chi(-1) = -1$. Следовательно, $N = q+1 = 72$. Далее, имеется в точности четыре точки порядка 2 (включая «бесконечную» точку O): они соответствуют корням многочлена $x^3 - x = x(x+1)(x-1)$ (см. упражнение 4а ниже). Это означает, что 2-примарная часть группы имеет тип (4, 2) и, таким образом, тип группы — это (4, 2, 3, 3) или (4, 2, 9), в зависимости от того, равно 9 или 3 число точек порядка 3. Следовательно, остается выяснить, существует ли 9 точек порядка 3. Заметим, что условие $3P = O$ при $P \neq O$ эквивалентно условию $2P = \pm P$, т. е. тому, что x-координаты точек $2P$ и P одинаковы. Согласно (5) это означает, что $\left((3x^2-1)^2/(2y)\right)^2 - 2x = x$, т. е. $(3x^2-1)^2 = 12xy^2 = 12x^4 - 12x^2$. Упрощая, получаем $3x^4 - 6x^2 - 1 = 0$. Это уравнение имеет, самое большее, 4 корня в $\mathbf{F}_{71}$. Каждый корень может давать не более двух точек (при $y = \pm\sqrt{x^3-x}$, если x^3-x есть квадрат по модулю 71), и если было бы 4 корня, то получилось бы 9 точек порядка 3 (включая точку O). В противном случае должно было бы быть меньше 9 точек порядка 3 (и, стало быть, в точности 3 таких точки). Но если корень x биквадратного уравнения таков, что $x^3 - x$ есть квадрат по модулю 71, то для другого его корня $-x$ получаем, что $(-x)^3 - (-x) = -(-x^3 - x)$ не есть квадрат. Значит, число точек порядка 3 не может быть равно 9 и потому тип группы — (4, 2, 9).

Расширения конечных полей, гипотеза Вейля. Если эллиптическая кривая определена над $\mathbf{F}_q$, она определена также над $\mathbf{F}_{q^r}$, $r = 1, 2, \ldots$, и имеет смысл рассматривать $\mathbf{F}_{q^r}$-точки, т.е. решения (1) над расширениями поля. Отправляясь от поля $\mathbf{F}_q$ как поля, над которым задана E, определяем число N_r как число $\mathbf{F}_{q^r}$-точек на E. (Таким образом, $N_1 = N$ есть число точек с координатами в нашем «основном поле» $\mathbf{F}_q$.)

Для чисел N_r можно рассмотреть «производящий ряд» $Z(T; E/\mathbf{F}_q)$ формальный степенной ряд в $\mathbf{Q}[[T]]$:

$$Z(T; E/\mathbf{F}_q) = e^{\sum N_r T^r/r}, \tag{7}$$

где T — неизвестная; обозначение $E/\mathbf{F}_q$ указывает эллиптическую кривую и поле, которое мы берем в качестве основного, а сумма в правой части берется по всем $r = 1, 2, \ldots$ Можно показать, что ряд справа (рассматриваемый как бесконечное произведение экспоненциальных степенных рядов $e^{N_r T^r/r}$) имеет положительные целые коэффициенты. Этот степенной ряд называется *дзета-функцией* эллиптической кривой (над $\mathbf{F}_q$) и представляет собой весьма важный объект, связанный с E.

«Гипотеза Вейля» (ныне теорема Делиня (P. Deligne)) в значительно более общем контексте (алгебраические многообразия произвольной размерности) утверждает, что дзета-функция имеет весьма специальный вид. В случае эллиптической кривой $E/\mathbf{F}_q$ Вейль (A. Weil) доказал следующее утверждение.

Гипотеза (теорема) Вейля для эллиптической кривой. *Дзета-функция есть рациональная функция от T вида*

$$Z(T; E/\mathbf{F}_q) = \frac{1 - aT + qT^2}{(1 - T)(1 - qT)}, \tag{8}$$

где от эллиптической кривой E зависит лишь целое число a. Значение a связано с числом $N = N_1$ соотношением $N = q + 1 - a$. Кроме того, дискриминант квадратного трехчлена в числителе отрицателен (т. е. $a^2 < 4q$ — теорема Хассе), таким образом, этот трехчлен имеет два комплексно сопряженных корня α и β, оба по модулю равные $\sqrt{q}$ (точнее, корнями являются $1/\alpha$ и $1/\beta$, а α, β — корни «возвратного» уравнения).

Доказательство см. в § V.2 упомянутой выше книги Силвермена.

Замечание. Если записать числитель (8) в виде $(1-\alpha T)(1-\beta T)$ и затем взять производную от логарифмов обеих частей (заменяя

левую часть по формуле (7), определяющей дзета-функцию), то нетрудно убедиться, что формула (8) эквивалентна последовательности соотношений

$$N_r = q^r + 1 - \alpha^r - \beta^r, \qquad r = 1, 2, \ldots$$

Так как α и β, наряду с a, определяются значением $N = N_1$, то число точек над $\mathbf{F}_q$ однозначно определяет число точек над любым его расширением. Таким образом, теорему Вейля для эллиптических кривых можно использовать, в частности, для нахождения числа точек над расширениями высокой степени.

Пример 5. Легко вычисляется дзета-функция эллиптической кривой $y^2 + y = x^3$ над $\mathbf{F}_2$, так как имеется всего три $\mathbf{F}_2$-точки. Она равна $(1 + 2T^2)/((1 - T)(1 - 2T))$. Таким образом, обратные корни числителя — это $\pm i\sqrt{2}$. Отсюда следует формула

$$N_r = \begin{cases} 2^r + 1, & \text{если } r \text{ нечетно,} \\ 2^r + 1 - 2(-2)^{r/2}, & \text{если } r \text{ четно.} \end{cases} \tag{9}$$

В заключение этого параграфа заметим, что существует много аналогий между группой $\mathbf{F}_q$-точек на эллиптической кривой и мультипликативной группой $\mathbf{F}_q^*$. Например, по теореме Хассе они имеют примерно одинаковое число элементов. Однако абелевы группы, которые строятся по эллиптическим кривым, имеют одно значительное преимущество, которое объясняет их ценность для криптографии: для одного и того же (большого) q существует богатый выбор различных эллиптических кривых с разными значениями N. Эллиптические кривые составляют богатый источник «естественно возникающих» конечных абелевых групп. Мы воспользуемся этой возможностью в последующих трех параграфах.

УПРАЖНЕНИЯ

1. Пусть E — эллиптическая кривая, определенная над $\mathbf{C}$, уравнение (1) которой имеет коэффициенты $a, b \in \mathbf{R}$; тогда точки E с вещественными координатами образуют подгруппу. Описать все возможные виды структуры такой подгруппы комплексной кривой E (которая как группа изоморфна произведению окружности на себя). Приведите пример для каждой из них.

2. Сколько точек P порядка n (т.е. таких, что $nP = O$) имеется на эллиптической кривой над $\mathbf{C}$? Сколько таких точек на эллиптической кривой над $\mathbf{R}$?

3. Привести пример эллиптической кривой над $\mathbf{R}$, имеющей в точности 2 точки порядка 2, и пример кривой, имеющей в точности 4 точки порядка 2.

4. Пусть P — точка на эллиптической кривой над $\mathbf{R}$. Предположим, что P не есть точка в бесконечности. Найти геометрическое условие, эквивалентное тому, что P — точка порядка а) 2; б) 3; в) 4.

5. Каждая из следующих точек имеет конечный порядок на данной эллиптической кривой над Q. Найти в каждом случае порядок P.

а) $P = (0, 16)$ на $y^2 = x^3 + 256$.

б) $P = (\frac{1}{2}, \frac{1}{2})$ на $y^2 = x^3 + \frac{1}{4}x$.

в) $P = (3, 8)$ на $y^2 = x^3 - 43x + 166$.

г) $P = (0, 0)$ на $y^2 + y = x^3 - x^2$ (уравнение можно привести к виду (1) заменой переменных $y \to y - \frac{1}{2}$, $x \to x + \frac{1}{3}$).

6. Вывести формулы сложения, аналогичные (4)–(5), для эллиптических кривых над полем характеристики 2, 3 (см. уравнения (2)–(3)).

7. Доказать, что число $\mathbf{F}_q$-точек на каждой из следующих эллиптических кривых равно $q + 1$:

а) $y^2 = x^3 - x$, когда $q \equiv 3 \pmod 4$;

б) $y^2 = x^3 - 1$, когда $q \equiv 2 \pmod 3$ (q нечетно);

с) $y^2 + y = x^3$, когда $q \equiv 2 \pmod 3$ (здесь q может быть четным).

8. Для всех степеней нечетных простых чисел $q = p^r$ до 27 включительно найти порядок и тип группы $\mathbf{F}_q$-точек на эллиптических кривых $y^2 = x^3 - x$ и $y^2 = x^3 - 1$ (в последнем случае — при $p \neq 3$). В некоторых случаях вам нужно будет проверить, сколько точек имеют порядок 3 или 4.

9. Пусть $q = 2^r$ и пусть эллиптическая кривая E над $\mathbf{F}_q$ имеет уравнение $y^2 + y = x^3$.

а) Выразить координаты $-P$ и $2P$ в терминах координат P.

б) Показать, что при $q = 16$ каждая $P \in E$ есть точка порядка 3.

в) Показать, что любая точка E с координатами в $\mathbf{F}_{16}$ фактически есть точка с координатами в $\mathbf{F}_4$. Далее с помощью теоремы Хассе при $q = 4$ и 16 определить число точек на кривой.

10. Вычислить дзета-функцию двух кривых из упражнения 8 над $\mathbf{F}_p$ для $p = 5, 7, 11, 13$.

11. Вычислить дзета-функцию кривой $y^2 + y = x^3 - x + 1$ над $\mathbf{F}_p$ для $p = 2$ и 3. (Сначала покажите, что в обоих случаях $N_1 = 1$.) Пусть $N(x) = x \cdot \overline{x}$ обозначает норму комплексного числа. В терминах нормы найти простую формулу для N_r.

ЛИТЕРАТУРА к § 1 ГЛАВЫ VI

1. *Fulton W.* Algebraic Curves. Menlo Park: Benjamin, 1969.
2. *Husemöller D.* Elliptic Curves. Heidelberg etc.: Springer, 1987.
3. *Koblitz N.* Introduction to Elliptic Curves and Modular Forms. 2nd ed. Heidelberg etc.: Springer, 1993.
4. *Koblitz N.* Why study equations over finite fields? — Math. Mag., 1982, v. 55, p. 144–149.
5. *Lang S.* Elliptic Curves: Diophantine Analysis. Heidelberg etc.: Springer, 1978.
6. *Silverman J.* The Arithmetic of Elliptic Curves. Heidelberg etc.: Springer, 1986.

§ 2. Криптосистемы на эллиптических кривых

В § IV.3 мы показали, как можно конечную абелеву группу $\mathbf{F}_q^*$ — мультипликативную группу конечного поля — использовать для создания криптосистемы с открытым ключом. Точнее, именно слож-

ность решения задачи дискретного логарифмирования приводит к криптосистемам, обсуждавшимся в § IV.3. Цель этого параграфа — построение аналогичных систем с открытым ключом, основанных на конечной абелевой группе эллиптической кривой, определенной над $\mathbf{F}_q$.

Прежде чем описывать криптосистемы, нужно обсудить некоторые вспомогательные понятия.

Кратные точки. *Для эллиптических кривых аналогом умножения двух элементов группы* $\mathbf{F}_q^*$ *служит сложение двух точек эллиптической кривой* E, *определенной над* $\mathbf{F}_q$. Таким образом, аналог возведения в степень k элемента из $\mathbf{F}_q^*$ — это умножение точки $P \in E$ на целое число k. Возведение в k-ю степень в $\mathbf{F}_q^*$ методом повторного возведения в квадрат можно осуществить за $O(\log k \log^3 q)$ двоичных операций (см. предложение II.1.9). Аналогично, мы покажем, что кратное $kP \in E$ можно найти за $O(\log k \log^3 q)$ двоичных операций методом повторного удвоения.

Пример 1. Чтобы найти $100P$, записываем $100P = 2(2(P + 2(2(2(P + 2P)))))$ и приходим к цели, производя 6 удвоений и 2 сложения точек на кривой.

Предложение VI.2.1. *Пусть эллиптическая кривая* E *определена уравнением Вейерштрасса (уравнением* (1),(2) *или* (3) *из предыдущего параграфа) над конечным полем* $\mathbf{F}_q$. *Если задана точка* P *на* E, *то координаты* kP *можно вычислить за* $O(\log k \log^3 q)$ *двоичных операций.*

Доказательство. Заметим, что вычисление координат суммы двух точек по уравнениям (4)–(5) (или по аналогичным уравнениям из упражнения 6 к § 1) требует не более 20 действий в $\mathbf{F}_q$ (умножений, делений, сложений или вычитаний). Поэтому, согласно предложению II.1.9, сложение двух точек (или удвоение точки) занимает время $O(\log^3 q)$. Так как при методе повторного удвоения выполняется $O(\log k)$ одинаковых шагов (см. доказательство предложения I.3.6), мы получаем, что координаты точки kP вычисляются за $O(\log k \log^3 q)$ двоичных операций.

Замечание 1. Оценки времени работы в предложении VI.2.1 не являются наилучшими, особенно для конечных полей характеристики $p = 2$. Нам, однако, достаточно этих оценок, которые следуют из наиболее очевидных алгоритмов арифметики в конечных полях.

Замечание 2. Если известно число N точек на нашей эллиптической кривой E и если $k > N$, то в силу равенства $NP = O$ мы можем заменить k его наименьшим неотрицательным вычетом по модулю N; в этом случае временная оценка заменяется на $O(\log^4 q)$

(напомним, что $N \leqslant q + 1 + 2\sqrt{q} = O(q)$). Рене Шуф (R. Schoof) предложил алгоритм, вычисляющий N за $O(\log^8 q)$ двоичных операций.

Представление открытого текста. Мы намереваемся кодировать наши открытые тексты точками некоторой заданной эллиптической кривой E, определенной над конечным полем $\mathbf{F}_q$. Мы хотим это осуществить простым и систематическим способом так, чтобы открытый текст m (который можно рассматривать как целое число из некоторого интервала) можно было легко прочитать, зная координаты соответствующей точки P_m. Заметим, что это «кодирование» — не то же самое, что засекречивание. Позднее мы будем рассматривать способы шифрования точек P_m открытого текста. Однако законный пользователь системы должен быть в состоянии восстановить m после дешифрования точки шифртекста.

Следует сделать два замечания. Во-первых, не известно *детерминистического* полиномиального (по $\log q$) алгоритма для выписывания большого числа точек произвольной эллиптической кривой E над $\mathbf{F}_q$. Однако, как мы увидим далее, существуют вероятностные алгоритмы с малой вероятностью неудачи. Во-вторых, порождать случайные точки на E недостаточно: чтобы закодировать большое число возможных сообщений m, необходим какой-то систематический способ порождения точек, которые были бы связаны с m определенным образом, например, чтобы x-координата имела с m простую связь.

Приведем один возможный вероятностный метод представления открытых текстов как точек на эллиптической кривой E, определенной над $\mathbf{F}_q$, где $q = p^r$ предполагается большим (и нечетным (см. ниже упражнение 8 при $q = 2^r$)). Пусть $\varkappa$ — достаточно большое целое число, настолько, что можно считать допустимым, если попытка представить в нужном нам виде элемент («слово») открытого текста m оказывается неудачной в одном случае из $2^{\varkappa}$; практически достаточно $\varkappa = 30$ или, на худой конец, $\varkappa = 50$. Пусть элементы нашего сообщения m — целые числа, $0 \leqslant m \leqslant M$. Предположим также, что выбранное нами конечное поле имеет q элементов, $q > M\varkappa$. Записываем целые числа от 1 до $M\varkappa$ в виде $m\varkappa + j$, где $1 \leqslant j \leqslant \varkappa$, и устанавливаем взаимно однозначное соответствие между такими числами и некоторым множеством элементов из $\mathbf{F}_q$. Например, можно записать такое число как r-значное число в p-ичной системе счисления и, отождествляя цифры в этой записи с элементами $\mathbf{F}_p \cong \mathbf{Z}/p\mathbf{Z}$, рассматривать их как коэффициенты многочлена степени не выше $r-1$ над $\mathbf{F}_p$, соответствующего элементу поля $\mathbf{F}_q$. Таким образом, числу $(a_{r-1}a_{r-2}\ldots a_1a_0)_p$ ставится в соответствие многочлен $\sum_{i=0}^{r-1} a_i X^i$, который, будучи рассмотрен по модулю некоторого фиксированного неприводимого многочлена степе-

ни r над $\mathbf{F}_p$, дает элемент $\mathbf{F}_q$.

Итак, при данном m при каждом $j = 1, 2, \ldots, \varkappa$ мы получаем элемент x из $\mathbf{F}_q$, соответствующий $m\varkappa + j$. Для такого x вычисляем правую часть уравнения

$$y^2 = f(x) = x^3 + ax + b$$

и пытаемся найти квадратный корень из $f(x)$, используя метод, изложенный в конце § II. 2. (Хотя этот алгоритм был приведен для простого поля $\mathbf{F}_p$, он дословно переносится на любое конечное поле $\mathbf{F}_q$. Чтобы воспользоваться им, нужно иметь элемент g в этом поле, не являющийся квадратом; его можно легко найти с помощью вероятностного алгоритма.) Если мы находим такой y, что $y^2 = f(x)$, то берем $P_m = (x, y)$. Если $f(x)$ не оказывается квадратом, то увеличиваем j на 1 и повторяем попытку с соответствующим значением x. Если мы находим x, для которого $f(x)$ — квадрат, прежде, чем j превысит $\varkappa$, то мы можем восстановить m по известной точке (x, y) с помощью формулы $m = [(\widetilde{x} - j)/\varkappa]$, где $\widetilde{x}$ — целое число, соответствующее x при установленном взаимно однозначном соответствии между целыми числами и элементами $\mathbf{F}_q$. Так как $f(x)$ — квадрат приблизительно в 50% случаев, то вероятность того, что метод не приведет к результату и мы не найдем точки P_m с x-координатой, соответствующей целому числу $\widetilde{x}$ между $m\varkappa + 1$ и $m\varkappa + \varkappa$, равна примерно $2^{-\varkappa}$. (Точнее, вероятность того, что $f(x)$ есть квадрат, фактически равна $N/(2q)$, однако $N/(2q)$ очень близко к $1/2$.)

Дискретный логарифм на E. В § IV. 3 мы рассматривали криптосистемы с открытым ключом, основанные на задаче дискретного логарифмирования в мультипликативной группе конечного поля. Теперь мы сделаем то же самое в группе (относительно сложения точек) эллиптической кривой, определенной над конечным полем $\mathbf{F}_q$.

О п р е д е л е н и е. Пусть E — эллиптическая кривая над $\mathbf{F}_q$ и B — точка на E. *Задачей дискретного логарифмирования* на E (с основанием B) называется задача нахождения для данной точки $P \in E$ такого целого числа $x \in \mathbf{Z}$ (если оно существует), что $xB = P$.

Вполне возможно, что задача дискретного логарифмирования на эллиптической кривой окажется более трудной для решения, чем задача дискретного логарифмирования в конечных полях. Наиболее сильные методы, разработанные для конечных полей, по-видимому, не работают в случае эллиптических кривых. Это обстоятельство по-особенному отчетливо проявляется в случае полей характеристики 2. Как объясняется в обзорной статье Одлыжко, упомянутой в списке литературы, специальные методы решения задачи дискретного лога-

рифмирования в $\mathbf{F}_{2^r}^*$ позволяют сравнительно легко вычислять дискретные логарифмы и, следовательно, вскрывать криптосистемы, рассмотренные в § IV. 3, если r не выбрано очень большим. Аналогичные системы, использующие эллиптические кривые над $\mathbf{F}_{2^r}$ (см. ниже), судя по всему, являются надежными при значительно меньших значениях r. Так как имеются практические причины (связанные с устройством ЭВМ и программированием) предпочтительности арифметических операций над полями $\mathbf{F}_{2^r}$, криптосистемы с открытым ключом, рассматриваемые ниже, могут оказаться более удобными для практического применения, чем системы, основанные на задаче дискретного логарифмирования в $\mathbf{F}_q^*$.

До 1990 г. единственными известными алгоритмами дискретного логарифмирования на эллиптических кривых были те, которые работают в любой группе и не используют особенности ее строения. Эти алгоритмы с экспоненциальным временем работы применимы к случаям, когда порядок группы делится на большое простое число. Однако впоследствии Менезес (Menezes), Окамото (Okamoto) и Вэнстон (Vanstone) предложили новый подход к задаче дискретного логарифмирования на эллиптической кривой E, определенной над $\mathbf{F}_q$. А именно, они использовали спаривание Вейля (см. § III. 8 учебника Силвермена (Silverman) из списка литературы к § 1) для вложения группы E в мультипликативную группу некоторого расширения $\mathbf{F}_{q^k}$ поля $\mathbf{F}_q$. Это вложение сводит задачу дискретного логарифмирования на E к соответствующей задаче для $\mathbf{F}_{q^k}^*$.

Однако такое сведение полезно, лишь если степень k расширения мала. Фактически единственный вид эллиптических кривых, для которых k мало, — это так называемые «суперсингулярные» эллиптические кривые, наиболее известными примерами которых являются кривые вида $y^2 = x^3 + ax$ над полем $\mathbf{F}_q$ характеристики $p \equiv -1 \pmod 4$ и кривые вида $y^2 = x^3 + b$ над полем $\mathbf{F}_q$, когда $p \equiv -1 \pmod 3$. Как правило, эллиптические кривые не являются суперсингулярными. Для них такое сведение почти никогда не приводит к субэкспоненциальным алгоритмам (см. мою работу в J. of Cryptology, приведенную в списке литературы).

Таким образом, основные преимущества криптосистем на эллиптических кривых заключаются в том, что не известны субэкспоненциальные алгоритмы вскрытия этих систем, если в них не используются суперсингулярные кривые, а также кривые, порядки которых делятся на большое простое число.

Теперь мы опишем аналоги систем с открытым ключом из § IV. 3, основанные на задаче дискретного логарифмирования на эллиптической кривой, определенной над конечным полем $\mathbf{F}_q$.

Аналог ключевого обмена Диффи–Хеллмана. Предположим, что Аида и Бернардо хотят договориться о ключе, которым будут впоследствии пользоваться в некоторой классической криптосистеме. Прежде всего они открыто выбирают какое-либо конечное поле $\mathbf{F}_q$ и какую-либо эллиптическую кривую E над ним. Их ключ строится по случайной точке P на этой эллиптической кривой. Если у них есть случайная точка P, то, например, ее x-координата дает случайный элемент $\mathbf{F}_q$, который можно затем преобразовать в r-разрядное целое число в p-ичной системе счисления (где $q = p^r$), и это число может служить ключом в их классической криптосистеме. (Здесь мы пользуемся словом «случайный» в неточном смысле; мы лишь хотим сказать, что выбор из некоторого большого множества допустимых ключей произволен и непредсказуем). Они должны выбрать точку P так, чтобы все их сообщения друг другу были открытыми и все же никто, кроме них двоих, ничего бы не знал о P.

Аида и Бернардо первым делом открыто выбирают точку $B \in E$ в качестве «основания»; B играет ту же роль, что образующий g в системе Диффи–Хеллмана для конечных полей. Мы, однако, не требуем, чтобы B была образующим элементом в группе точек кривой E. Эта группа может и не быть циклической. Даже если она циклическая, мы не хотим тратить время на проверку того, что B — образующий элемент (или даже на нахождение общего числа N точек, которое нам не понадобится в последующем). Нам хотелось бы, чтобы порожденная B подгруппа была большой, предпочтительно того же порядка величины, что и сама E. Позже мы обсудим этот вопрос. Пока что предположим, что B — взятая открыто точка на E весьма большого порядка (равного либо N, либо большому делителю N).

Чтобы образовать ключ, Аида вначале случайным образом выбирает целое число a, сравнимое по порядку величины с q (которое близко к N); это число она держит в секрете. Она вычисляет $aB \in E$ и передает эту точку открыто. Бернардо делает то же самое: он выбирает случайно b и открыто передает $bB \in E$. Тогда используемый ими секретный ключ — это $P = abB \in E$. Оба пользователя могут вычислить этот ключ. Например, Аида знает bB (точка была передана открыто) и свое собственное секретное a. Однако любая третья сторона знает лишь aB и bB. Кроме решения задачи дискретного логарифмирования — нахождения a по B и aB (или нахождения b по B и bB), — по-видимому, нет способа найти abB, зная лишь aB и bB.

Аналог системы Мэсси–Омуры. Как и в случае конечного поля, это — криптосистема с открытым ключом для передачи элементов сообщения m, которые мы теперь предположим представлен-

ными точками P_m фиксированной (и не скрываемой) эллиптической кривой E над $\mathbf{F}_q$ (q берется большим). Предполагается также, что общее число N точек на E вычислено и не составляет секрета. Каждый пользователь системы секретно выбирает такое целое случайное число e между 1 и N, что НОД $(e, N) = 1$. Используя алгоритм Евклида, он находит затем обратное e^{-1} к числу e по модулю N, т.е. такое целое число d, что $de \equiv 1 \pmod N$. Если Алиса хочет послать Бобу сообщение P_m, то она сначала посылает ему точку $e_A P_m$ (индекс A указывает на пользователя Алису). Это ничего не говорит Бобу, который, не зная ни e_A, ни d_A, не может восстановить P_m. Однако, не придавая этому значения, он умножает ее на *свое* e_B и посылает обратно Алисе $e_B e_A P_m$. На третьем шаге Алиса должна частично раскрыть свое сообщение, умножив $e_B e_A P_m$ на d_A. Так как $NP_m = O$ и $d_A e_A \equiv 1 \pmod N$, при этом получается точка $e_B P_m$, которую Алиса возвращает Бобу. Тот может теперь прочитать сообщение, умножив точку $e_B P_m$ на d_B.

Заметим, что злоумышленник может знать $e_A P_m$, $e_B e_A P_m$ и $e_B P_m$. Если бы он умел решать задачу дискретного логарифмирования на E, то он мог бы определить e_B по первым двум точкам, вычислить $d_B \equiv e_B^{-1} \pmod N$ и $P_m = d_B e_B P_m$.

Аналог системы Эль-Гамаля. Это — другая криптосистема с открытым ключом для передачи сообщений P_m. Как и в описанной выше системе ключевого обмена, мы исходим из данных несекретных: а) конечного поля $\mathbf{F}_q$, б) определенной над ним эллиптической кривой E и в) точки-«основания» B на ней. (Знать общее число N точек на E нам не нужно.) Каждый из пользователей выбирает случайное целое число a, которое держит в секрете, затем находит и делает общедоступной точку aB.

Чтобы послать Бьорну сообщение P_m, Анюта выбирает случайно целое число k и посылает пару точек $(kB, P_m + k(a_B B))$ (где $a_B B$ — открытый ключ Бьорна). Чтобы прочитать сообщение, Бьорн умножает первую точку из полученной пары на свое секретное число a_B и вычитает результат умножения из второй точки:

$$P_m + k(a_B B) - a_B(kB) = P_m.$$

Таким образом, Анюта посылает замаскированное P_m вместе с «подсказкой» kB, при помощи которой можно снять «маску» $ka_B B$, если знать секретное число a_B. Злоумышленник, который умеет решать задачу дискретного логарифмирования на E, может, конечно, найти a_B, зная $a_B B$ и B.

Выбор кривой и точки. Существуют различные способы выбора эллиптической кривой и (в системах Диффи–Хеллмана и Эль-Гамаля) точки B на ней.

Случайный выбор (E, B). Взяв какое-либо большое конечное поле $\mathbf{F}_q$, можно следующим образом осуществить одновременный выбор E и $B = (x, y) \in E$. (Будем предполагать, что характеристика p поля $\mathbf{F}_q$ больше 3, так что эллиптическая кривая задана уравнением (1) из § 1; при $q = 2^r$ или 3^r нетрудно сделать очевидные изменения в дальнейшем изложении.) Выбираем сначала случайным образом три элемента из $\mathbf{F}_q^*$ в качестве x, y, a. Далее полагаем $b = y^2 - (x^3 + ax)$. Убеждаемся в том, что кубический многочлен $x^3 + ax + b$ не имеет кратных корней, что равносильно проверке условия $4a^3 + 27b^2 \neq 0$. (Если это условие не выполняется, берем другую случайную тройку x, y, a.) Полагаем $B = (x, y)$. Тогда B — точка на эллиптической кривой $y^2 = x^3 + ax + b$.

Число N точек кривой можно найти несколькими способами. Первый полиномиальный алгоритм для вычисления $\# E$, построенный Рене Шуфом (René Schoof), является даже детерминистическим. Он основывается на нахождении значения $\# E$ по модулю l для всех простых чисел l, меньших некоторой границы. Для этого анализируется действие автоморфизма Фробениуса (отображения в p-ю степень) на точках порядка l.

В статье Шуфа оценка времени работы была фактически $O(\log^8 q)$, т. е. хотя и полиномиальной, однако быстро растущей. Вначале казалось, что алгоритм не имеет практического значения. Однако с тех пор многие пытались повысить скорость алгоритма Шуфа (Миллер (V. Miller), Элкис (N. Elkis), Бухман (J. Buchman), Мюллер (V. Müller), Менезес (A. Menezes), Чарлап (L. Charlap), Коули (R. Coley) и Роббинс (D. Robbins)). Кроме того, Эткин (A. O. L. Atkin) разработал другой метод, который, хотя и не гарантирует полиномиального времени работы, на практике дает весьма удовлетворительные результаты. В результате всех этих усилий стало возможным вычислять порядок произвольной эллиптической кривой над $\mathbf{F}_q$, если q — степень простого числа, записываемая 50 или даже 100 знаками. Некоторые методы нахождения числа точек на эллиптической кривой рассматриваются в работах из списка литературы в конце этого параграфа.

Следует также отметить, что хотя для реализации систем Диффи–Хеллмана или Эль-Гамаля знать N не нужно, на практике желательно быть уверенным в их надежности, которая зависит от того, имеет ли N большой простой делитель. Если N есть произведение малых простых чисел, то для решения задачи дискретного логарифми-

рования можно использовать метод Полига–Силвера–Хеллмана (см. § IV.3). Заметим, что метод Полига–Силвера–Хеллмана решает задачу дискретного логарифмирования в любой конечной абелевой группе (в отличие от также рассмотренного в § IV.3 алгоритма вычисления индекса, который зависит от особенностей $\mathbf{F}_q^*$). Таким образом, нужно знать, что N не есть произведение малых простых чисел и не похоже, что это можно узнать, если не найти фактически значение N.

Редукция глобальной пары (E, B) по модулю p. Упомянем теперь второй способ нахождения пары, состоящей из эллиптической кривой и точки на ней. Выберем сначала раз и навсегда «глобальную» эллиптическую кривую и точку бесконечного порядка на ней. Итак, пусть E — эллиптическая кривая, определенная над полем рациональных чисел (или, для большей общности, можно было бы использовать эллиптическую кривую, определенную над некоторым числовым полем), и B — точка бесконечного порядка на E.

Пример 2. Точка $B = (0, 0)$ является точкой бесконечного порядка на эллиптической кривой E: $y^2 + y = x^3 - x$ и фактически порождает всю группу рациональных точек на E.

Пример 3. Точка $B = (0, 0)$ является точкой бесконечного порядка на E: $y^2 + y = x^3 + x^2$ и порождает всю группу рациональных точек.

Далее, мы выбираем большое простое число p (или, если наша эллиптическая кривая определена над расширением K поля $\mathbf{Q}$, выбираем некоторый простой идеал в K) и рассматриваем *редукцию* E и B по модулю p. Точнее, для всех p, за исключением нескольких малых простых чисел, коэффициенты в уравнении для E имеют взаимно простые с p знаменатели и, следовательно, могут рассматриваться как коэффициенты в уравнении по модулю p. Если сделать замену переменных, приведя полученное уравнение над $\mathbf{F}_p$ к виду $y^2 = x^3 + ax + b$, то кубический многочлен в правой части не будет иметь кратных корней (за исключением нескольких малых простых p) и дает поэтому эллиптическую кривую над $\mathbf{F}_p$ (которую мы будем обозначать $E \pmod p$). Координаты точки B, будучи также приведенными по модулю p, дают точку на эллиптической кривой $E \pmod p$, которую мы будем обозначать $B \pmod p$.

При использовании этого второго способа мы раз и навсегда фиксируем E и B и за счет этого получаем много различных возможностей посредством изменения простого p.

Порядок точки B. С какой вероятностью «случайная» точка B на «случайной» эллиптической кривой оказывается порождающим

элементом? Или, в случае нашего второго метода выбора (E, B), какова вероятность того, что (для случайного p) точка B при редукции по модулю p дает образующий элемент кривой $E \pmod p$? Этот вопрос близок к следующему вопросу о мультипликативных группах конечных полей: пусть целое b фиксированно, а простое p случайно; какова вероятность того, что b — образующий в $\mathbf{F}_p^*$? Вопрос изучался как для конечных полей, так и для эллиптических кривых. Более подробное изложение можно найти в работе Гупты (Gupta) и Мурти (Murty), приведенной в списке литературы.

Как упоминалось выше, описанные криптосистемы могут быть надежными, даже если точка B не является порождающим элементом. Фактически нужно, чтобы в циклической группе, порождаемой B, задача дискретного логарифмирования не была эффективно разрешима. Это будет так (т. е. все известные методы решения задачи дискретного логарифмирования в произвольной абелевой группе оказываются слишком медленными), если порядок B делится на очень большое простое число, скажем, имеющее порядок величины, близкий к N.

Один из способов гарантировать, что наш выбор B является надлежащим (а фактически, что B порождает эллиптическую кривую) — это взять такую эллиптическую кривую и такое конечное поле, чтобы число точек N было простым числом. Тогда всякая точка $B \neq O$ будет порождающим элементом. Если использовать первый из описанных выше методов, то при фиксированном $\mathbf{F}_p$ можно продолжать выбор пар (E, B), пока не найдется такая, для которой число точек на E есть простое число (что можно определить одним из тестов на простоту, обсуждавшихся в § V. 1). Если применять второй метод, то для фиксированной глобальной эллиптической кривой E над $\mathbf{Q}$ можно продолжать выбирать простые p, пока не найдем кривую $E \pmod p$, число точек на которой — простое. Как долго нам придется ждать? Этот вопрос аналогичен следующему вопросу о группах $\mathbf{F}_p^*$: является ли $(p-1)/2$ простым числом, т. е. верно ли, что любой элемент, отличный от ± 1, — либо порождающий, либо квадрат порождающего элемента (см. упражнение 13 к § II. 1)? Ни для эллиптических кривых, ни для конечных полей вопрос пока не получил явного ответа, однако в обоих случаях предполагается, что вероятность выбора p с требующимся свойством есть $O(1/\log p)$.

Замечание. Для того чтобы $E \pmod p$ имела простой порядок N при большом p, надо выбирать E так, чтобы она имела тривиальное кручение, т. е. чтобы на ней не было точек конечного порядка, кроме O. В противном случае N будет делиться на порядок периодической подгруппы.

УПРАЖНЕНИЯ

1. Построить вероятностный алгоритм для нахождения элемента, не являющегося квадратом в $\mathbf{F}_q$.

2. Описать полиномиальный *детерминистический* алгоритм для представления открытых текстов в виде точек на эллиптических кривых в следующих случаях:

(а) E имеет уравнение $y^2 = x^3 - x$ и $q \equiv 3 \pmod 4$;

(б) E имеет уравнение $y^2 + y = x^3$ и $q \equiv 2 \pmod 3$.

3. Пусть E — эллиптическая кривая $y^2 + y = x^3 - x$, определенная над полем из $p = 751$ элементов (заменой переменных вида $y' = y + 376$ это уравнение приводится к виду (1) из § 1). Эта кривая содержит $N = 727$ точек. Предположим, что передаваемые элементы открытого текста — это десятичные цифры 0–9 и буквы A–Z с числовыми эквивалентами 10–35 соответственно. Берем $\varkappa = 20$.

а) Применяя изложенный метод, написать сообщение «STOP007» в виде последовательности семи точек на кривой.

б) Перевести последовательность точек (361, 383), (241, 605), (201, 380), (461, 467), (581, 395) в ответное сообщение.

4. Пусть E — эллиптическая кривая, определенная над $\mathbf{Q}$, и p — большое простое число, в частности, настолько большое, что, приводя уравнение $y^2 = x^3 + ax + b$ по модулю p, мы получаем эллиптическую кривую над $\mathbf{F}_p$. Показать, что

а) если многочлен $x^3 + ax + b$ разлагается по модулю p на три линейных множителя, то $E \pmod p$ — не циклическая;

б) если этот многочлен имеет корень по модулю p, то число N элементов на $E \pmod p$ четно.

5. Пусть E — эллиптическая кривая из примера 5 в § 1. Полагаем $q = 2^r$, и пусть N_r — число $\mathbf{F}_{2^r}$-точек на E.

а) Показать, что при $r > 1$ число N_r не может быть простым.

б) В случае $4|r$ найти условия, эквивалентные тому, что N_r делится на $(r/4)$-разрядное или $(r/4+1)$-разрядное простое число в двоичной записи.

6. Пусть E — эллиптическая кривая, определенная над $\mathbf{F}_p$, и N_r обозначает число $\mathbf{F}_{p^r}$-точек на E.

а) Доказать, что при $r > 1$ и $p > 3$ число N_r не может быть простым.

б) Привести контрпримеры к части а) в случаях $p = 2$ и $p = 3$.

7. а) Найти эллиптическую кривую E, определенную над $\mathbf{F}_4$, которая имеет лишь одну $\mathbf{F}_4$-точку (точку в бесконечности O).

б) Показать, что число $\mathbf{F}_{4^r}$-точек на кривой из части а) есть квадрат числа Мерсенна $2^r - 1$.

в) Найти простую формулу удвоения $\mathbf{F}_{4^r}$-точки на этой кривой.

г) Доказать, что если $2^r - 1$ — простое число Мерсенна, то всякая $\mathbf{F}_{4^r}$-точка (за исключением O) имеет порядок в точности $2^r - 1$.

8. Пусть r нечетно и пусть K обозначает поле $\mathbf{F}_{2^r}$. Для $z \in K$ пусть $g(z)$ обозначает $\sum_{j=0}^{(r-1)/2} z^{2^{2j}}$, а $\operatorname{tr} z$ (называемый «следом») обозначает $\sum_{j=0}^{r-1} z^{2^j}$.

а) Доказать, что $\operatorname{tr} z \in \mathbf{F}_2$; $\operatorname{tr}(z_1 + z_2) = \operatorname{tr} z_1 + \operatorname{tr} z_2$; $\operatorname{tr} 1 = 1$; $g(z) + g(z)^2 = z + \operatorname{tr} z$.

б) Показать, что $\operatorname{tr} z = 0$ для одной половины элементов из K и $\operatorname{tr} z = 1$ для другой половины.

в) Описать вероятностный алгоритм порождения $\mathbf{F}_{2^r}$-точек на эллиптической кривой $y^2 + y = x^3 + ax + b$.

9. Пусть E — эллиптическая кривая $y^2 = x^3 + ax + b$ с $a, b \in \mathbf{Z}$. Пусть $P \in E$. Пусть $p > 3$ обозначает такое простое число, что ни $4a^3 + 27b^2$, ни знаменатели x- и y-координат точки P не делятся на p. Показать, что порядок точки $P \pmod p$ на эллиптической кривой $E \pmod p$ — это такое наименьшее натуральное число k, что либо 1) $kP = O$ на E, либо 2) p делит знаменатель координат kP.

10. Пусть E — эллиптическая кривая $y^2 + y = x^3 - x$, определенная над $\mathbf{Q}$, и пусть $P = (0, 0)$. Вычисляя $2^j P$ для $j = 1, 2, \ldots$, найти пример такого простого p, что $E \pmod p$ *не* порождается точкой $P \pmod p$. (Замечание. Можно показать, что P *порождает* группу рациональных точек на E.)

11. Использовать построенный по эллиптической кривой аналог системы Эль-Гамаля, чтобы послать сообщение упражнения 3 а, выбирая E и p так же, как в упражнении 3, и полагая $B = (0, 0)$. Предположим, что открытый ключ вашего корреспондента — это точка (201, 380) и что ваша последовательность случайных значений k — это 386, 209, 118, 589, 312, 483, 335 (каждое из них используется для посылки одного элемента сообщения). Какую последовательность из 7 пар точек вы посылаете?

Заметим, что в этом упражнении мы использовали довольно малое значение p; более реальный пример такого рода, с которым можно встретиться на практике, потребовал бы работы с числами, записанными с помощью нескольких дюжин десятичных цифр.

ЛИТЕРАТУРА к § 2 ГЛАВЫ VI

1. *Agnew G., Mullin R., Vanstone S. A.* An implementation of elliptic curve cryptosystems over $\mathbf{F}_{2^{155}}$. — IEEE J. Select. Areas Comm., 1993, v. 11, p. 804–813.
2. *Gupta R., Murty M. R.* Primitive points on elliptic curves. — Compositio Math., 1986, v. 58, p. 13–44.
3. *Koblitz N.* Elliptic curve cryptosystems. — Math. Comp., 1987, v. 48.
4. *Koblitz N.* Primality of the number of points on an elliptic curve over a finite field. — Pacific J. Math., 1988, v. 131, p. 157–165.
5. *Koblitz N.* Constructing elliptic curve cryptosystems in characteristic 2. — In: Advances in Cryptology —Crypto'90. Heidelberg etc.: Springer, 1991, v. 156–167.
6. *Koblitz N.* Elliptic curve implementation of zero-knowledge blobs. — J. Cryptology, 1991, v. 4, p. 207–213.
7. *Koblitz N.* CM-curves with good cryptographic properties. — In: Advances in Cryptology —Crypto'91. Heidelberg etc.: Springer, 1992, p. 279–287.
8. *Lenstra H. W., Jr.* Elliptic curves and number-theoretic algorithms. — Report 86–19. Amsterdam: Mathematisch Instituut, Universiteit van Amsterdam, 1986.
9. *Menezes A.* Elliptic Curve Public Key Cryptosystem. Dordrecht: Kluwer Acad. Publ., 1993.
10. *Menezes A., Okamoto T., Vanstone S. A.* Reducing elliptic curve logarithms to logarithms in a finite field. — IEEE Trans. Inform. Theory IT-39, 1993, p. 1639–1646.
11. *Menezes A., Vanstone S., Zuccherato R.* Counting points on elliptic curves over $\mathbf{F}_{2^m}$. — Math. Comp., 1993, v. 60, p. 407–420.

12. *Miller V.* Use of elliptic curves in cryptography. — In: Abstracts for Crypto 85, 1985.
13. *Odlyzko A. M.* Discrete logarithms in finite fields and their cryptographic significance. — In: Advances in Cryptology, Proceedings of Eurocrypt 84. Heidelberg etc.: Springer, 1985, p. 224–314.
14. *Schoof R.* Elliptic curves over finite fields and the computation of square roots mod p. — Math. Comp., 1985, v. 44, p. 483–494.

§ 3. Критерий простоты, использующий эллиптические кривые

Критерий простоты, использующий эллиптические кривые, который предложили Гольдвассер (S. Goldwasser), Килиан (J. Kilian) и (в другом варианте) Эткин (A. O. L. Atkin), представляет собой аналог следующего связанного с группой $(\mathbf{Z}/n\mathbf{Z})^*$ критерия простоты Поклингтона (H. Pocklington).

Предложение VI. 3. 1. *Пусть n — натуральное число. Предположим, что имеется простой делитель q числа $n-1$, больший $\sqrt{n}-1$. Если существует такое целое число a, что* 1) $a^{n-1} \equiv 1 \pmod{n}$ *и* 2) НОД $(a^{(n-1)/q} - 1, n) = 1$, *то n — простое число.*

Доказательство. Если n — не простое, то существует простой делитель p числа n, не превосходящий $\sqrt{n}$. Так как $q > p - 1$, имеем НОД $(q, p-1) = 1$, и, следовательно, существует такое целое u, что $uq \equiv 1 \pmod{p-1}$. Тогда $a^{(n-1)/q} \equiv a^{uq(n-1)/q} = a^{u(n-1)} \equiv 1 \pmod{p}$ (по условию 1)). Однако это противоречит условию 2).

Замечания. Это — отличный тест при условии, что $n-1$ делится на простое число $q > \sqrt{n} - 1$ и что мы в самом деле можем найти это q (и доказать его простоту). В противном случае им пользоваться нельзя. (Впрочем, это не совсем так — существует более общий вариант, применимый тогда, когда известны большой делитель $n-1$ и известно его разложение на простые множители, см. упражнение 2 ниже.)

Отметим, что этот критерий простоты является вероятностным лишь в том смысле, что случайно выбранное a может либо удовлетворять, либо не удовлетворять условию 2) (но, конечно, a обязано удовлетворять условию 1), иначе n — не простое). Однако, как только такое a найдено (обычно достаточно взять $a = 2$), критерий доказывает, что n — простое число. В отличие от тестов § V. 1 (тестов Соловея–Штрассена и Миллера–Рабина), заключение теста Поклингтона — совершенно определенное: n — простое число (а не «вероятно простое»).

Критерий простоты, использующий эллиптические кривые, основан на аналогичном предложении, в котором предполагается заданным уравнение $y^2 = x^3 + ax + b$, рассматриваемое по модулю n. Иначе говоря, целые числа a, b рассматриваются как вычеты по модулю n и E обозначает множество всех пар (x, y) целых чисел из $\mathbf{Z}/n\mathbf{Z}$, удовлетворяющих этому уравнению, вместе с символом O, который мы называем «точкой в бесконечности». Если n — простое число (а это почти всегда так, ибо практически мы рассматриваем лишь числа, прошедшие тесты на вероятную простоту из § V. 1), то E — эллиптическая кривая с нейтральным элементом O.

Прежде чем формулировать аналог предложения VI. 3. 1 для E, заметим, что даже не зная, что n — простое число, можно для сложения элементов из E применять формулы из § 1. При сложении двух точек (или удвоении точки) возможны три случая: 1) мы получаем корректно определенную точку; 2) если точки имеют вид (x, y) и $(x, -y)$ по модулю n, то мы получаем точку в бесконечности; 3) формулы оказываются неопределенными, поскольку встретился знаменатель, не обратимый по модулю n. Однако случай 3) означает, что n — составное число, и мы можем найти его нетривиальный делитель, взяв НОД от n и этого знаменателя. Таким образом, без ущерба для общности мы можем в дальнейшем предполагать, что случай 3) исключен.

Можно показать, что если P — элемент E по модулю n, то даже при составном n ответ, который наш алгоритм дает для mP, не зависит от последовательности сложений или удвоений точек (а priori это не очевидно). Однако этот факт ниже не используется. Достаточно условиться обозначать mP *любую* точку, полученную из P путем вычислений по модулю n по формулам из § 1.

Вполне аналогично тому, как сложение точек по модулю n можно проводить без предположения о простоте n, можно применить к нашему множеству E по модулю n алгоритм нахождения числа точек на эллиптической кривой (например, метод Шуфа). Мы либо получим некоторое целое m, которое при простом n равно числу точек *эллиптической кривой* E, либо столкнемся с неопределенным выражением, знаменатель которого имеет с n нетривиальный общий делитель. Как и в случае сложения точек, без ущерба для общности можно предположить, что последний случай никогда не встречается.

Такое m будет играть роль $n - 1$ в предложении VI. 3. 1; отметим, что $n - 1$ — это порядок группы $(\mathbf{Z}/n\mathbf{Z})^*$, если n — простое число.

Теперь сформулируем аналог критерия Поклингтона для эллиптических кривых.

Предложение VI. 3. 2. *Пусть n — натуральное число. Пусть E — множество, определяемое уравнением $y^2 = x^3 + ax + b$ по моду-*

лю n, как описано выше (и НОД $(4a^3+27b^2, n) = 1$ *— прим. перев.). Пусть m — целое число. Предположим, что у m есть простой делитель q, больший $(n^{1/4}+1)^2$. Если существует такая точка P на E, что* 1) $mP = O$ *и* 2) $\frac{m}{q}P$ *определена и не равна O, то n — простое число.*

Доказательство (ср. с доказательством предложения VI. 3. 1). Если n — не простое, то у него есть простой делитель $p \leqslant \sqrt{n}$. Пусть E' — эллиптическая кривая, заданная тем же уравнением, что и E, но рассматриваемая по модулю p, и пусть m' — порядок группы E'. По теореме Хассе $m' \leqslant p+1+2\sqrt{p} = (\sqrt{p}+1)^2 \leqslant (n^{1/4}+1)^2 < q$ и поэтому НОД$(q, m') = 1$. Следовательно, существует такое целое u, что $uq \equiv 1 \pmod{m'}$. Пусть $P' \in E'$ — это точка P, рассматриваемая по модулю p. Тогда в E' имеем $\frac{m}{q}P' = uq\frac{m}{q}P' = umP' = O$ ввиду 1), так как mP' получается в точности тем же путем, что и mP, только все действия производятся не по модулю n, а по модулю его делителя p. Но это противоречит 2), так как если $\frac{m}{q}P$ определена и не равна O по модулю n, то, производя при вычислении $\frac{m}{q}P'$ те же действия, только по модулю p, мы должны получить $\frac{m}{q}P' \neq O$. Предложение доказано.

Это предложение приводит к алгоритму для доказательства простоты данного натурального числа n (о котором уже известно, что оно «вероятно простое»). Действуем следующим образом. Выбираем случайно три целых числа a, x, y по модулю n и полагаем $b \equiv y^2 - x^3 - ax \pmod{n}$. (Если при этом НОД $(4a^3+27b^2, n) > 1$, то выбираем другую тройку (a, x, y). — *Прим. перев.*) Тогда $P = (x, y)$ есть элемент множества E, которое определяется уравнением $y^2 = x^3 + ax + b$. Используем метод Шуфа (или другой метод для подсчета числа точек на эллиптической кривой), чтобы найти число m, которое при простом n равно числу точек эллиптической кривой E над $\mathbf{F}_n$. Если мы не можем представить m в виде $m = kq$, где $k \geqslant 2$ — малое целое число, а q — «вероятно простое» (т. е. прошедшее один из тестов § V. 1), то выбираем другую случайную тройку (a, x, y) и повторяем действия с начала. Предположим, что мы получили, наконец, эллиптическую кривую, для которой m имеет искомый вид. Тогда с помощью формул из § VI. 1 (работая по модулю n) находим mP и kP. Если мы в какой-то момент получаем неопределенное выражение — либо при нахождении кратной для точки P, либо при использовании алгоритма Шуфа, то мы сразу находим нетривиальный делитель числа n. Мы можем предположить, что этого не случается. Если $mP \neq O$, то n — составное (так как, если бы n было простым, группа E имела бы порядок m и

любой элемент E при умножении на m обращался бы в нуль). Если $kP = O$ (что весьма маловероятно), то нам не повезло и нужно начинать заново с другой тройкой. Но если $mP = O$ и $kP \neq O$, то согласно предложению VI. 3. 2 число n — простое, если большой делитель q числа m — действительно простое число (а мы знаем только, что q — «вероятно простое»). Это сводит задачу к доказательству простоты числа q, которое не превосходит $n/2$. Мы затем начинаем снова с q вместо n. Таким образом, мы получаем рекуррентную процедуру с t применениями критерия простоты, где t не больше $\log_2 n$. В конце работы мы получаем число q_t, о котором знаем, что оно простое; отсюда следует, что предыдущее q_{t-1} было на самом деле простым (а не только «вероятно простым»), то же самое справедливо для q_{t-2} и далее, до $q_1 = q$, и, наконец, получаем, что само n — действительно простое. Этим описание критерия простоты, использующего эллиптические кривые, завершено.

С данным критерием связаны две трудности, одна — практическая, а другая — теоретическая. Во-первых, хотя алгоритм Шуфа и работает полиномиальное время от $\log n$, для практических применений он очень сложен. В последнее время были достигнуты некоторые успехи в его разработке и упрощении, но все равно приходится подсчитывать число точек на большом количестве кривых E, пока мы не найдем, наконец, кривую, у которой m имеет нужный нам вид — $m = kq$. Чтобы справиться с этой проблемой, Эткин (A. O. L. Atkin) разработал вариант критерия простоты, использующего эллиптические кривые, с тщательным подбором эллиптических кривых с комплексным умножением, для которых значительно легче находить число точек в их редукциях по модулю n. Более подробно о методе Эткина см. статью Ленстры и Ленстры из приведенного ниже списка литературы.

Вторая трудность — теоретическая. Чтобы *найти эллиптическую* кривую E над $\mathbf{F}_n$ (предполагая, что n — простое), у которой число точек — «почти простое» (т. е. имеет вид $m = qk$ для малого k и простого q), мы должны знать кое-что о распределении простых чисел (вернее, «почти простых») в интервале $p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p}$, где по теореме Хассе находится число m. Пока что не доказаны теоремы, гарантирующие, что в результате полиномиально зависящего от $\log n$ числа попыток с высокой вероятностью будет обнаружена кривая E, число точек которой имеет указанный вид и лежит в интервале столь (относительно) малой длины. Однако существует весьма правдоподобная гипотеза, которая могла бы это гарантировать, чего вполне достаточно для практических целей. Построение доказуемо полиномиального вероятностного алгоритма значительно сложнее: такой критерий простоты, разработанный Адлеманом (L. Adleman) и Хуа-

ном (M. Huang), использует двумерные абелевы многообразия, представляющие собой обобщение эллиптических кривых на размерность 2. Однако этот алгоритм очень громоздок и совершенно неприменим на практике.

УПРАЖНЕНИЯ

1. а) Пусть в критерии простоты Поклингтона n — простое число, $n-1$ делится на простое число q, как в предложении VI.3.1, и a выбирается случайно в $(\mathbf{Z}/n\mathbf{Z})^*$. Какова вероятность того, что a будет удовлетворять условиям этого предложения?

б) Пусть в критерии простоты, использующем эллиптические кривые, n — простое число, имеется эллиптическая кривая, порядок которой делится на простое число q, как в предложении VI.3.2, и P — случайно выбранная точка на ней. Какова вероятность того, что P будет удовлетворять условиям этого предложения?

2. Обобщить критерий простоты Поклингтона на случай, когда известен делитель s числа $n-1$, больший $\sqrt{n}-1$, и известны все его простые делители q. Условие 2) должно выполняться для всех $q|s$.

3. а) (Критерий простоты Пепина (Pépin) для чисел Ферма). Доказать, что число Ферма $n=2^{2^k}+1$ простое тогда и только тогда, когда существует такое целое число a, что $a^{2^{2^k-1}} \equiv -1 \pmod{n}$. Доказать, что если n — простое, то 50% из всех $a \in (\mathbf{Z}/n\mathbf{Z})^*$ обладают этим свойством. Доказать также, что если $k>1$, то в качестве a всегда можно выбрать одно из чисел 3, 5, 7.

б) Доказать, что число Мерсенна $n=2^p-1$ является простым в том и только том случае, когда на кривой $E\colon y^2=x^3+x \pmod{n}$ найдется такая точка $P=(x,y)$, что 1) при вычислении $2^{p-1}P$ не возникают необратимые по модулю n знаменатели и 2) y-координата точки $2^{p-1}P$ равна нулю. Для этого покажите, во-первых, что если $n=2^p-1$ является простым, то группа точек на $E \pmod{n}$ — циклическая порядка 2^p и 50% всех $P \in E \pmod{n}$ имеют отмеченные выше свойства 1)–2). Объясните, как можно порождать случайные точки $E \pmod{n}$. Вы можете пользоваться любым алгоритмом, который предполагает, что $b^{n-1} \equiv 1 \pmod{n}$ (т. е. что n — псевдопростое по различным основаниям b), так как, если вы когда-либо столкнетесь с b, для которого это сравнение не выполняется, то ваш критерий прекратит работу, выдав заключение, что n — составное число.

Заметим, что этот критерий простоты — вероятностный в том смысле, что если n — простое число, то нет никаких гарантий относительно времени, за которое нужная нам точка P обнаружится. Но, как только такая точка P найдена, критерий заключает, что n — простое число. В этом отношении он отличается от критериев псевдопростоты из § V.1. Обобщенную процедуру проверки на простоту любого нечетного n можно найти в работе Босма (W. Bosma) из приведенного ниже списка литературы.

ЛИТЕРАТУРА к § 3 ГЛАВЫ VI

1. *Adleman L., Huang M.* Recognizing primes in random polynomial time. — In: Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, p. 462–469.

2. *Bosma W.* Primality testing using elliptic curves. — Report 85–12. Amsterdam: Mathematisch Instituut, Universiteit van Amsterdam, 1985.
3. *Goldwasser S., Kilian J.* Almost all primes can be quickly certified. — In: Proceedings of the 18th Annual ACM Symposium on Theory of Computing, 1986, p. 316–329.
4. *Lenstra A. K., Lenstra H. W., Jr.* Algorithms in number theory. — Technical Report 87–008. Chicago: University of Chicago, 1987.
5. *Morain F.* Implementation of the Goldwasser–Kilian–Atkin primality testing algorithm. — INRIA report 911, 1988.
6. *Pocklington H.* The determination of the prime and composite nature of large numbers by Fermat's theorem. — Proc. Cambridge Phil. Soc., 1914–16, v. 18, p. 29–30.
7. *Schoof R.* Elliptic curves over finite fields and the computation of square roots mod p. — Math. Comp., 1985, v. 44, p. 483–494.

§ 4. Разложение на множители при помощи эллиптических кривых

Основной причиной повышенного интереса части криптографов к эллиптическим кривым послужило недавно найденное Ленстрой (H. W. Lenstra) остроумное применение эллиптических кривых в новом методе факторизации, которой во многих отношениях лучше существовавших ранее. С точки зрения практики, это продвижение не настолько значительно, чтобы угрожать надежности криптосистем, основанных на трудности задачи разложения на множители (оценки времени имеют тот же вид, с которым мы уже встречались в § V.3). Тем не менее, открытие более совершенного метода, использующего неожиданный новый инструментарий, предостерегает от благодушия по поводу невозможности существенных продвижений в задаче разложения на множители. Цель этого последнего параграфа — описание метода Ленстры.

Предваряя изложение алгоритма Ленстры разложения на множители с помощью эллиптических кривых, мы напомним один классический способ факторизации, аналогичный методу Ленстры.

$(p-1)$-метод Полларда. Пусть мы хотим разложить на множители составное число n и предполагаем, что p — его (еще не известный) простой делитель. Если p таков, что $p-1$ не имеет больших простых делителей, то p можно найти следующим способом.

1. Выбираем целое число k, кратное всем или большинству целых чисел, меньших некоторой границы B. Например, в качестве k можно взять $B!$ или общее наименьшее кратное всех целых чисел, не превосходящих B.

2. Выбираем целое число a между 2 и $n-2$. Например, a может быть равно 2, 3 или случайно выбранному целому числу.

3. Вычисляем a^k по модулю n повторным возведением в квадрат.

4. Вычисляем $d = \text{НОД}\,(a^k - 1, n)$, используя алгоритм Евклида и вычет $a^k \pmod n$ из шага 3.

5. Если d не есть нетривиальный делитель n, то повторяем все шаги с новым a и/или новым k.

Чтобы понять, при каких условиях этот алгоритм будет работать, предположим, что k делится на все натуральные числа, не превосходящие B. Далее, пусть p — простой делитель n, для которого $p-1$ представляется в виде произведения степеней небольших простых чисел, причем каждый из сомножителей меньше B. Отсюда следует, что k кратно $p-1$ (так как оно кратно каждому из сомножителей в разложении $p-1$). Поэтому, по малой теореме Ферма, имеем $a^k \equiv 1 \pmod p$. Тогда НОД $(a^k - 1, n)$ делится на p. Таким образом, единственное препятствие, которое может помешать нам получить на шаге 4 нетривиальный делитель n, — это случай, когда $a^k \equiv 1 \pmod n$.

Пример 1. Разложим указанным методом $n = 540143$. Выбираем $B = 8$ (и, следовательно, k равно 840 — общему наименьшему кратному $1, 2, \ldots, 8$) и $a = 2$. Находим, что $2^{840} \pmod n$ — это 53047 и НОД $(53046, n) = 421$. Тем самым приходим к разложению $540143 = 421 \cdot 1283$.

Основная слабость метода Полларда, очевидно, проявляется при попытках его применения в тех случаях, когда все простые делители p числа n таковы, что $p-1$ делится на относительно большое простое число (или на большую степень простого числа).

Пример 2. Пусть $n = 491389$. Вряд ли мы найдем нетривиальный делитель, если выберем $B < 191$. Причина этому — разложение $n = 383 \cdot 1283$: мы имеем $383 - 1 = 2 \cdot 191$, $1283 - 1 = 2 \cdot 641$ (как 191, так и 641 — простые числа). За исключением $a \equiv 0, \pm 1 \pmod{383}$, все другие a имеют порядки по модулю 383 либо 191, либо 382; за исключением $a \equiv 0, \pm 1 \pmod{1283}$, все другие значения a имеют порядки по модулю 1283 либо 641, либо 1282. Таким образом, если k не делится на 191 (или 641) то, скорее всего, на шаге 4 мы вновь и вновь будем получать НОД $(a^k - 1, n) = 1$.

Основной недостаток $(p-1)$-метода Полларда состоит в том, что в нем используется только группа $(\mathbf{Z}/p\mathbf{Z})^*$ (точнее, различные такие группы, когда p пробегает простые делители n). Для данного n эти группы фиксированны. Если порядок каждой из них делится на большое простое число, метод не работает.

Принципиальное отличие метода Ленстры, как мы увидим, со-

стоит в том, что при использовании эллиптических кривых над полем $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ мы неожиданно получаем целый сонм групп и поэтому можем рассчитывать найти среди них такую, у которой порядок не делится на большое простое число или большую степень простого числа.

Мы начнем описание алгоритма Ленстры с некоторых комментариев по поводу редукции точек на эллиптических кривых по модулю n, где n — составное число (в отличие от § 2, где мы работали по модулю простых чисел и в конечных полях).

Эллиптические кривые — редукция по модулю n. До конца этого параграфа мы будем обозначать через n нечетное составное число и через p (пока неизвестный) простой делитель n. Будем считать, что $p > 3$. Пусть m — целое число и x_1, x_2 — два рациональных числа со знаменателями, взаимно простыми с m; будем писать $x_1 \equiv x_2 \pmod m$, если числитель разности $x_1 - x_2$, записанной в виде несократимой дроби, делится на m. Для любого рационального числа x_1 со знаменателем, взаимно простым с m, существует такое однозначно определенное целое x_2 между 0 и $m-1$ («наименьший неотрицательный вычет»), что $x_1 \equiv x_2 \pmod m$. Иногда мы будем обозначать наименьший неотрицательный вычет как $x_1 \pmod m$.

Пусть даны уравнение вида $y^2 = x^3 + ax + b$, $a, b \in \mathbf{Z}$, и удовлетворяющая ему точка $P = (x, y)$. На практике эллиптическая кривая E вместе с точкой P будут порождаться некоторым «случайным» образом. Например, можно выбрать три случайных целых числа a, x, y из некоторой области и положить затем $b = y^2 - x^3 - ax$. Будем предполагать, что кубический многочлен $x^3 + ax + b$ имеет различные корни, т. е. что $4a^3 + 27b^2 \neq 0$; это условие выполняется почти всегда, если коэффициенты выбираются описанным выше случайным образом. Для простоты мы предполагаем также в дальнейшем, что $4a^3 + 27b^2$ не имеет с n общих множителей; другими словами, что $x^3 + ax + b$ не имеет кратных корней по модулю p для любого простого делителя p числа n. Практически, выбрав a, b, мы можем проверить это, найдя НОД $(4a^3 + 27b^2, n)$. Если это число больше 1, то либо $n | (4a^3 + 27b^2)$ (и тогда нам следует выбрать другие a и b), либо мы уже обнаружили нетривиальный делитель n (и тогда задача решена). Итак, мы будем предполагать, что НОД $(4a^3 + 27b^2, n) = 1$.

Пусть нам нужно найти кратную kP для точки P методом повторного удвоения, изложенным в § VI. 2. Есть много различных способов сделать это за $O(\log k)$ шагов, каждый из которых — либо удвоение точки, либо сложение двух различных точек. Можно, например, записать k в двоичной системе счисления: $k = a_0 + a_1 2 + \cdots + a_{m-1} 2^{m-1}$

и затем последовательно удваивать P, прибавляя 2^jP к уже накопленной сумме, если соответствующий бит a_j равен 1. Можно также разложить k в произведение простых чисел $l_1, l_2, \ldots$ (расположенных, скажем, в неубывающем порядке) и затем последовательно вычислять l_1P, $l_2(l_1P)$ и т. д. Здесь каждая точка l_jP_j, где $P_j = l_{j-1}l_{j-2}\ldots l_1P$, вычисляется повторным удвоением с использованием двоичной записи чисел l_j.

Пусть выбран какой-либо из таких приемов вычисления кратных kP. Будем рассматривать точку P и все ее кратные по модулю n. Это означает, что мы полагаем $P \pmod n = (x \pmod n, y \pmod n)$ и всякий раз при вычислении кратной точки kP мы фактически вычисляем лишь вычеты ее координат по модулю n. Все вычисления при удвоении и сложении точек можно проводить по модулю n, если выполнено нетривиальное условие: все знаменатели должны быть взаимно просты с n.

Предложение VI. 4. 1. *Пусть E — эллиптическая кривая с уравнением $y^2 = x^3 + ax + b$, где $a, b \in \mathbf{Z}$ и* НОД $(4a^3 + 27b^2, n) = 1$. *Пусть P_1 и P_2 — две точки на E, у которых знаменатели координат взаимно просты с n, и $P_1 \neq -P_2$. Тогда $P_1 + P_2 \in E$ имеет координаты, у которых знаменатели взаимно просты с n, в том и только том случае, когда у n нет простого делителя p, для которого сумма точек $P_1 \pmod p$ и $P_2 \pmod p$ на эллиптической кривой $E \pmod p$ была бы равна точке в бесконечности $O \pmod p \in E \pmod p$. Здесь $E \pmod p$ обозначает эллиптическую кривую над $\mathbf{F}_p$, полученную приведением по модулю p коэффициентов уравнения $y^2 = x^3 + ax + b$.*

Доказательство. Пусть все точки $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ и $P_1 + P_2 \in E$ поначалу имеют координаты со знаменателями, взаимно простыми с n. Требуется доказать, что для любого простого делителя p числа n сумма $P_1 \pmod p + P_2 \pmod p \neq O \pmod p$. Если $x_1 \not\equiv x_2 \pmod p$, то в соответствии с описанием закона сложения на $E \pmod p$ мы сразу же заключаем, что $P_1 \pmod p + P_2 \pmod p \neq O \pmod p$. Теперь предположим, что $x_1 \equiv x_2 \pmod p$. При $P_1 = P_2$ координаты точки $P_1 + P_2 = 2P_1$ определяются формулой (5) из § 1, и $2P_1 \pmod p$ находится по той же формуле с заменой каждого члена его вычетом по модулю p. Нам нужно показать, что знаменатель $2y_1$ дроби в правой части (5) не делится на p. Но если бы он делился на p, то $3x_1^2 + a$ делилось бы на p (так как знаменатель x-координаты точки $2P_1$ на p не делится). Тогда x_1 был бы корнем по модулю p как многочлена $x^3 + ax + b$, так и его производной $3x^2 + a$, что противоречит нашему предположению об отсутствии кратных корней по модулю p у этого многочлена. Теперь пусть $P_1 \neq P_2$. Так как $x_2 \equiv x_1 \pmod p$

и $x_2 \not\equiv x_1$, мы можем записать $x_2 = x_1 + p^r x$, где $r \geqslant 1$ выбрано так, что ни числитель, ни знаменатель x не делятся на p. По предположению знаменатели координат точки $P_1 + P_2$ не делятся на p, поэтому из формулы (4) в § 1 следует, что y_2 имеет вид $y_1 + p^r y$. С другой стороны,

$$\begin{aligned} y_2^2 &= (x_1 + p^r x)^3 + a(x_1 + p^r x) + b \\ &\equiv x_1^3 + ax_1 + b + p^r x(3x_1^2 + a) = y_1^2 + p^r x(3x_1^2 + a) \pmod{p^{r+1}}. \end{aligned} \tag{1}$$

Так как $x_2 \equiv x_1 \pmod p$ и $y_2 \equiv y_1 \pmod p$, то $P_1 \pmod p = P_2 \pmod p$, т. е. $P_1 \pmod p + P_2 \pmod p = 2P_1 \pmod p$. Очевидно, что $2P_1 \pmod p = O \pmod p$ тогда и только тогда, когда $y_1 \equiv y_2 \equiv 0 \pmod p$. Если это сравнение выполнено, то $y_2^2 - y_1^2 = (y_2 - y_1)(y_2 + y_1)$ должно делиться на p^{r+1} (т. е. должен делиться на p^{r+1} числитель этого рационального числа). Поэтому из сравнения (1) следовало бы, что $3x_1^2 + a \equiv 0 \pmod p$. Но это невозможно, поскольку $x^3 + ax + b$ не имеет кратных корней по модулю p, т. е. x_1 не может быть общим корнем этого многочлена и его производной. Значит, $P_1 \pmod p + P_2 \pmod p \not\equiv O \pmod p$, как и утверждалось.

Обратно, пусть $P_1 \pmod p + P_2 \pmod p \not\equiv O \pmod p$ для каждого простого делителя p числа n. Покажем, что знаменатели координат $P_1 + P_2$ взаимно просты с n, т. е. что знаменатели не делятся ни на какой простой делитель p числа n. Фиксируем некоторое $p|n$. Формулы (4), (5) из § 1 показывают, что если $x_2 \not\equiv x_1 \pmod p$, то делящихся на p знаменателей нет. Поэтому предположим, что $x_2 \equiv x_1 \pmod p$. Тогда $y_2 \equiv \pm y_1 \pmod p$; но так как $P_1 \pmod p + P_2 \pmod p \not\equiv O \pmod p$, то $y_2 \equiv y_1 \not\equiv 0 \pmod p$. При $P_2 = P_1$ формула (5) из § 1 вместе с условием $y_1 \not\equiv 0 \pmod p$ показывает, что знаменатели координат точки $P_1 + P_2 = 2P_1$ взаимно просты с p. Наконец, если $P_2 \neq P_1$, мы вновь пишем $x_2 = x_1 + p^r x$ с x, не делящимся на p, и, используя сравнение (1), получаем $(y_2^2 - y_1^2)/(x_2 - x_1) \equiv 3x_1^2 + a \pmod p$. Так как p не делит $y_1 + y_2 \equiv 2y_1 \pmod p$, отсюда следует, что знаменатель числа $\frac{y_2^2 - y_1^2}{(y_1+y_2)(x_2-x_1)} = \frac{y_2 - y_1}{x_2 - x_1}$ не делится на p, и в силу формулы (4) из § 1 знаменатели координат точки $P_1 + P_2$ не делятся на p. Теорема доказана.

Метод Ленстры. Пусть дано составное целое нечетное число n и нам нужно найти его нетривиальный делитель $d|n$, $1 < d < n$. Берем сначала какую-либо эллиптическую кривую E: $y^2 = x^3 + ax + b$ с целыми a, b и точку $P = (x, y)$ на ней. Пару (E, P) можно выбирать случайным образом или использовать любой детерминистический метод, который порождает много таких пар (как в примере 4 ниже). Мы

попытаемся с помощью E и P разложить n способом, описанным ниже; если попытка окажется неудачной, то возьмем другую пару (E, P) и будем продолжать до тех пор, пока не найдем делитель $d|n$. Если вероятность неудачи равна $\rho < 1$, то вероятность того, что все h последовательно выбранных пар (E, P) окажутся неудачными, равна ρ^h, т. е. очень мала для больших h. Таким образом, с высокой вероятностью мы разложим n за приемлемое число шагов.

Имея пару (E, P), мы выбираем целое число k, которое делится на степени малых простых чисел (т. е. не больших некоторого B), не превосходящие C. Таким образом, мы полагаем

$$k = \prod_{l \leqslant B} l^{\alpha_l}, \tag{2}$$

где $\alpha_l = [(\log C)/(\log l)]$ есть такой наибольший возможный показатель, что $l^{\alpha_l} \leqslant C$. Далее мы пытаемся вычислить kP, выполняя все действия по модулю n. Результаты этих вычислений не представляют для нас интереса до тех пор, пока при попытке найти число, обратное к $x_2 - x_1$ в формуле (4) из § 1 или к $2y_1$ в (5), мы не столкнемся с числом, которое *не* взаимно просто с n. Согласно предложению VI. 4. 1 это произойдет, когда появится такая точка $k_1 P$ (частичная сумма в процессе вычисления kP), что $k_1(P \pmod p) = O \pmod p$ для некоторого простого делителя $p|n$, иными словами, порядок $P \pmod p$ в группе $E \pmod p$ является делителем k_1. Когда мы, применяя алгоритм Евклида, пытаемся обратить знаменатель, делящийся на p, мы вместо этого находим НОД числа n и этого знаменателя. Этот НОД и будет искомым собственным делителем n, если он отличен от n, т. е. если знаменатель не делится на само n. Однако такая делимость означала бы, по предложению VI. 4. 1, что $k_1 P \pmod p = O \pmod p$ для *всех* простых делителей p числа n, а это очень маловероятно, если n имеет не менее двух очень больших простых делителей. Таким образом, почти с полной гарантией при попытке вычисления $k_1 P$ по модулю n для k_1, которое делится на порядок точки $P \pmod p$ при некотором $p|n$, мы получим собственный делитель числа n.

Отметим сходство с $(p-1)$-методом Полларда. Вместо группы $(\mathbf{Z}/p\mathbf{Z})^*$ мы используем группу $E \pmod p$. Однако на сей раз, если наш выбор E неудачен, т. е. для каждого $p|n$ порядок группы $E \pmod p$ делится на большое простое число (и, следовательно, $kP \pmod p$, скорее всего, не совпадает с $O \pmod p$ для k, определенного формулой (2)), то нам нужно просто отбросить ее и выбрать другую эллиптическую кривую E вместе с точкой $P \in E$. В методе Полларда такой возможности не было.

Алгоритм. Пусть n — нечетное составное натуральное число. Опишем теперь вероятностный метод Ленстры разложения n на множители.

Пусть у нас есть метод порождения пар (E, P), состоящих из эллиптической кривой $y^2 = x^3 + ax + b$, $a, b \in \mathbf{Z}$, и точки $P = (x, y) \in E$. Если дана такая пара, то мы проводим описанную ниже последовательность действий. Если посредством этой процедуры не удается получить нетривиальный делитель n, то мы образуем новую пару (E, P) и повторяем процесс.

До начала работы с кривой E по модулю n нужно проверить, что это — действительно эллиптическая кривая по модулю любого $p|n$, т. е. что кубический многочлен в правой части уравнения не имеет кратных корней по модулю p. Это справедливо тогда и только тогда, когда дискриминант $4a^3 + 27b^2$ взаимно прост с n. Таким образом, если НОД $(4a^3 + 27b^2, n) = 1$, мы можем действовать. Конечно, если НОД не равен 1 или n, мы получаем искомый делитель n. Если этот НОД равен n, то нам следует выбрать другую эллиптическую кривую.

Далее, предположим, что мы выбрали два натуральных числа B и C. Здесь B — максимальная величина простого делителя того целого числа k, на которое мы будем умножать точку P. Чем больше B, тем больше вероятность того, что $kP \pmod p = O \pmod p$ для нашей пары (E, P) и некоторого $p|n$; с другой стороны, чем больше B, тем дольше придется вычислять $kP \pmod p$. Поэтому B нужно выбирать таким образом, чтобы (по нашей оценке) минимизировать время работы. Число C должно служить как бы верхней границей для того простого делителя $p|n$, с которым мы надеемся в конце концов получить соотношение $kP \pmod p = O \pmod p$. Затем мы выбираем k по формуле (2), т. е. представленным в виде произведения степеней простых чисел, не превосходящих B, каждая из которых не превосходит C. Тогда, согласно теореме Хассе, если $p + 1 + 2\sqrt{p} < C$ и порядок кривой $E \pmod p$ не делится ни на какое простое число, большее B, то k кратно этому порядку и потому $kP \pmod p = O \pmod p$.

Пример 3. Предположим, что мы выбрали $B = 20$ и хотим разложить на множители 10-значное (в десятичной системе) число n, которое может быть произведением двух 5-значных простых чисел (т. е. n не делится ни на какое простое число, записываемое менее, чем 5 знаками). Тогда выбираем $C = 100700$ и $k = 2^{16} \cdot 3^{10} \cdot 5^7 \cdot 7^5 \cdot 11^4 \cdot 13^4 \cdot 17^4 \cdot 19^3$.

Вернемся к описанию алгоритма. Работая по модулю n, пытаемся вычислять kP следующим образом. Используя метод повторного удвоения, находим $2P, 2(2P), 2(4P), \ldots, 2^{\alpha_2}P$, вслед за этим $3(2^{\alpha_2})P$,

$3(3\cdot 2^{\alpha_2}P),\ldots,3^{\alpha_3}2^{\alpha_2}P$ и т. д., пока, наконец, не получим $\prod_{l\leqslant B} l^{\alpha_l}P$ (умножаем, последовательно переходя от самых малых простых делителей l числа k к самым большим). В этих вычислениях при каждом делении по модулю n применяем алгоритм Евклида для нахождения обратного элемента. Если на какой-либо стадии алгоритм Евклида не дает обратного элемента, то либо найден нетривиальный делитель n, либо в качестве НОД n и этого знаменателя получено само n. В последнем случае следует возвратиться к началу и взять другую пару (E,P). Если алгоритм Евклида всегда дает обратный элемент — и, таким образом, $kP \pmod n$ вычисляется, — следует возвратиться к началу и взять другую пару (E,P). Описание алгоритма завершено.

Пример 4. Используем семейство эллиптических кривых $y^2 = x^3+ax-a$, $a=1,2,\ldots$, каждая из которых содержит точку $P=(1,1)$. Перед использованием a мы должны проверить, что дискриминант $4a^3+27a^2$ взаимно прост с n. Попытаемся разложить $n=5429$, полагая $B=3$ и $C=92$. (В этом примере и в приведенных ниже упражнениях мы иллюстрируем метод, используя небольшие значения n. Конечно, на практике достоинства метода проявляются лишь при несравненно больших n). Здесь наш выбор C объясняется желанием найти простой делитель, который может принимать значения до $\sqrt{n}\approx 73$; для $p=73$ верхняя оценка общего числа $\mathbf{F}_p$-точек — это $74+2\sqrt{73}<92$. Согласно (2) выбираем $k=2^6\cdot 3^4$. При каждом значении a мы последовательно умножаем шесть раз точку P на 2 и затем четыре раза на 3, работая по модулю n на эллиптической кривой $y^2=x^3+ax-a$. Когда $a=1$, все вычисления проходят гладко и тем самым выясняется, что $3^42^6P \pmod p$ — конечная точка на $E \pmod p$ для каждого $p|n$. Следующая попытка — с $a=2$. В этом случае при вычислении $3^22^6P \pmod p$ мы получаем знаменатель, у которого НОД с n равен 61, т. е. находим собственный делитель n. Таким образом, порядок точки $(1,1)$ на кривой $y^2=x^3+2x-2$ по модулю 61 является делителем 3^22^6 (см. упражнение 5 ниже). Итак, вторая попытка оказалась успешной. Кстати, если мы возьмем $a=3$, то при вычислении 3^42^6P найдем другой простой делитель числа, равный 89. (Обычно, но не всегда, этот метод дает наименьший простой делитель.)

Время работы. Центральным вопросом при оценке времени работы является вычисление вероятности того, что при заданном p и заданной границе B (выбранной некоторым оптимальным образом) порядок N случайной эллиптической кривой по модулю p не делится ни на какое простое число, большее B. Известно, что порядки N эллиптических кривых, находящиеся (по теореме Хассе) в интервале $[p+1-2\sqrt{p}, p+1+2\sqrt{p}]$, распределены в нем довольно равномерно (с

тем исключением, что вблизи концов интервала плотность значений N падает). Значит, эта вероятность приближенно равна вероятности того, что случайно выбранное целое число величины примерно p не делится ни на какое простое число, большее B. Мы уже видели при выводе эвристической временной оценки в § V. 3, что эта вероятность есть примерно u^{-u}, где $u = (\log p)/(\log B)$. Это приводит к оценке вида $O(e^{C\sqrt{r \log r}})$, где r — число бит в n. Подробный вывод оценки времени работы можно найти в статье Ленстры.

Точнее, предположим, что n — натуральное число, которое не есть степень простого и не делится на 2 и на 3. Используя правдоподобное предположение о распределении целых чисел, не делящихся на простые числа, большие B, в малом интервале вокруг p, Ленстра доказывает следующую вероятностную временную оценку для числа двоичных операций, требующихся для получения нетривиального делителя n:

$$e^{\sqrt{(2+\varepsilon) \log p \log \log p}}, \tag{3}$$

где p — наименьший простой делитель n и ε сколь угодно близко к нулю для достаточно больших p. Так как всегда $p < \sqrt{n}$, из (3) получается также оценка

$$e^{\sqrt{(1+\varepsilon) \log n \log \log n}}. \tag{4}$$

Оценка (4) имеет тот же самый вид, что и (эвристические) временные оценки для лучших из известных общих методов факторизации. Однако метод Ленстры имеет следующий ряд преимуществ перед своими конкурентами.

1. Это — единственный метод, время работы которого существенно уменьшается, если n делится на простое число, значительно меньшее $\sqrt{n}$.

2. По этой причине его можно использовать в сочетании с другими методами факторизации, когда требуется разлагать на множители какие-либо вспомогательные числа. (Например, в методе цепных дробей из § V. 4 нам нужно полное разложение числа $b_i^2 \pmod n$, если оно есть произведение относительно малых простых чисел.)

3. В отличие от своих конкурентов он использует небольшой объем памяти.

Однако самой замечательной особенностью факторизационного алгоритма Ленстры является исторически первое использование эллиптических кривых — интенсивно изучаемого объекта современной теории чисел и алгебраической геометрии, обладающего богатой структурой. Это показывает возможность появления новых методов раз-

ложения на множители, использующих неожиданные построения из доселе весьма далеких областей математики.

УПРАЖНЕНИЯ

1. Используя метод Полларда с $k = 840$ и $a = 2$, попытаться разложить на множители $n = 53467$. Затем сделать попытку с $a = 3$.

2. Предположим, что лишь один из простых делителей p числа n таков, что $p - 1$ не делится на большие простые числа. Предположим, что в алгоритме Полларда берется значение k, которое не кратно $p-1$, и пробуются различные значения a. Оцените в терминах k и $p - 1$ вероятность получить $d = p$ на шаге 4.

3. Для следующих значений p и B найти (используя, если необходимо, компьютер) долю целых чисел между $p + 1 - 2\sqrt{p}$ и $p + 1 + 2\sqrt{p}$, которые не имеют простых делителей, больших B: а) $p = 109$, $B = 3$; б) $p = 109$, $B = 19$; в) $p = 1009$, $B = 19$; г) $p = 1009$, $B = 97$; д) $p = 9973$, $B = 97$.

4. Каждое из чисел n в упражнении 5 из § V.4 имеет делитель $p < 100$. В каждом из случаев а)–л) найти этот делитель методом Ленстры эллиптических кривых, выбирая $B = 5$, $C = 120$, $P = (1, 1)$ и E: $y^2 = x^3 + ax - a$ при $a = 1, 2, \ldots$ (взяв те a, для которых дискриминант взаимно прост с n). Каково в каждом из случаев первое значение a, для которого вы находите делитель, и каково при этом первое значение k_1, для которого найденный делитель появляется как НОД (знаменатель, n) при вычислении $k_1 P$?

5. Пусть k задано равенством (2). Предположим, что вы находите делитель числа n в процессе вычисления $k_1 P$ по модулю n, где k_1 — частичное произведение в (2). (Напомним, что мы вычисляем kP, умножая последовательно на числа l в порядке возрастания.) Докажите, что $k_1 P \pmod p = O \pmod p$ для некоторого $p|n$, исключив тем самым возможность получения знаменателя, не взаимно простого с n, при вычислении $l((k_1/l)P)$ на одном из этапов метода удвоения, предшествующем последнему шагу.

6. а) Предположим, что для любого $a \in \mathbf{Z}$ имеется эффективный способ нахождения такой точки $P = (x, y)$, что $y^2 \equiv x^3 + ax \pmod n$. Объясните, чем плоха идея использовать для разложения n эллиптическую кривую $y^2 = x^3 + ax$ с различными a.

б) Тот же вопрос для семейства эллиптических кривых $y^2 = x^3 + b$ с различными b.

7. Предположим, что вы хотите немного увеличить вероятность того, что порядок $E \pmod p$ для некоторого $p|N$ есть произведение малых простых множителей, обеспечивая делимость этого порядка на 4. Опишите, как это сделать.

ЛИТЕРАТУРА к § 4 ГЛАВЫ VI

1. *Lenstra H. W., Jr.* Factoring integers with elliptic curves. — Ann. Math., 1987, v. 126, p. 649–673.
2. *Montgomery P.* Speeding the Pollard and elliptic curves methods of factorization. — Math. Comp., 1987, v. 48, p. 243–264.
3. *Pollard J. M.* Theorems on factorization and primality testing. — Proc. Cambridge Phil. Soc., 1974, v. 76, p. 521–528.

ОТВЕТЫ К УПРАЖНЕНИЯМ

§ I. 1

1. $(112111)_3$.

2. $(260\frac{12}{126})_7$.

3. 10001100101; $1101\frac{1010}{1011}$.

4. MPJNS; LIKE$\frac{\text{IT}}{\text{WE}}$ (другими словами, JQVXHJ=WE·LIKE+IT).

5. а) 10.101101111110000. б) C.SRO.

6. Если $b^f - 1$ кратно d, то дробь можно записать в виде $a/(b^f - 1)$, где a — целое число не более чем из f разрядов. Теперь воспользуйтесь формулой для суммы геометрической прогрессии с первым членом ab^{-f} и знаменателем b^{-f}. Обратно, при заданном чисто периодическом разложении x с периодом f покажите, что $b^f x$ отличается от x на f-значное целое a. Это означает, что $x = a/(b^f - 1)$.

7. а) $(\text{BAD})_{16}$. б) Никакого деления не требуется. Например, при переходе от двоичной к шестнадцатиричной системе просто разбиваем число (начав справа) на блоки по четыре разряда и рассматриваем каждую четверку как шестнадцатиричное число (или заменяем на один из символов 0–9, A–F).

8. 1) Посмотреть на верхний и нижний биты и проверить, не занималась ли единица для вычитания в младшем разряде. 2) Если биты одинаковы и единица не занималась или если верхний бит — это 1, а нижний — это 0 и единица занималась, то пишем 0 и идем дальше. 3) Если верхний бит — это 1, а нижний — это 0 и единица не занималась, то пишем 1 и идем дальше. 4) Если верхний бит — это 0, а нижний — это 1 и единица занималась, то занимаем единицу в старшем разряде, в текуший разряд разности помещаем 0 и идем дальше. 5) Если оба бита одинаковы и единица занималась или сверху стоит 0, а снизу 1 и единица не занималась, то занимаем единицу в старшем разряде, в текуший разряд разности помещаем 1 и идем дальше.

9. а) Требуется $n - 1$ умножений; в каждом из них промежуточное произведение 3^j имеет не более $O(n)$ разрядов и 3 имеет 2 разряда, поэтому требуется $O(n)$ двоичных операций. Таким образом, в итоге получаем $O(n^2)$. б) Здесь промежуточное произведение имеет $O(n \log n)$ разрядов. Поэтому каждое произведение требует $O(n \log^2 n)$ двоичных операций. Общая трудоемкость составляет $O(n^2 \log^2 n)$ двоичных операций.

10. $O(n^2 \log^2 N)$.

11. а) $O(n \log^2 n)$. б) $O(\log^2 n)$.

12. $O(rsn(\log^2 m + \log n))$.

13. а) Произведение $O(n/\log n)$ чисел из $O(\log n)$ разрядов каждое состоит $O(n/\log n) \cdot O(\log n) = O(n)$ разрядов. б) $O(n \log n)$. в) $O(n^2)$.

14. а) $O(\sqrt{n} \log^2 n)$. б) $O(\sqrt{n} \log n)$.

15. $O(m \log n)$.

16. Предположим, что n состоит из $k+1$ бит. В качестве первого приближения для $m = [\sqrt{n}]$ возьмем единицу с $[k/2]$ нулями за ней. Находим знаки разрядов числа m, сдвигаясь от единицы слева направо и поочередно пробуя заменить 0 на 1. Если при этой замене квадрат полученного числа m становится больше n, то в этом разряде оставляем 0.

§ I.2

1. б) Простой контрпример: пусть $b = -a$.

2. 16 делителей: 1, 3, 5, 7, 9, 15, 21, 27, 35, 45, 63, 105, 135, 189, 315, 945.

3. а) Если $a|n$, то пишем $n = ab$ и полагаем $a \longleftrightarrow b$. б) Имея $n = ab$, где $a \geqslant b$, полагаем $s = (a+b)/2$ и $t = (a-b)/2$. Обратно, имея $n = s^2 - t^2$, положим $a = s + t$, $b = s - t$. в) $473^2 - 472^2$, $159^2 - 156^2$, $97^2 - 92^2$, $71^2 - 64^2$, $57^2 - 48^2$, $39^2 - 24^2$, $33^2 - 12^2$, $31^2 - 4^2$.

4. б)$=100! = 2^{97} \cdot 3^{48} \cdot 5^{24} \cdot 7^{16} \cdot 11^9 \cdot 13^7 \cdot 17^5 \cdot 19^5 \cdot 23^4 \cdot 29^3 \cdot 31^3 \cdot 37^2 \cdot 41^2 \cdot 43^2 \cdot 47^2 \cdot 53 \cdot 59 \cdot 61 \cdot 67 \cdot 71 \cdot 73 \cdot 79 \cdot 83 \cdot 89 \cdot 97$. в) Формула имеет вид $(n - S_p(n))/(p-1)$. Чтобы доказать это, запишем $n = d_{k-1}p^{k-1} + \cdots + d_1 p + d_0$ и заметим, что $[n/p^j] = d_{k-1}p^{k-1-j} + \cdots + d_{j+1}p + d_j$ для каждого j. Далее воспользуемся формулой из пункта а).

6. а) $1 = 11 \cdot 19 - 8 \cdot 26$. б) $17 = 1 \cdot 187 - 5 \cdot 34$. в) $1 = 205 \cdot 160 - 39 \cdot 841$. г) $13 = 65 \cdot 2171 - 54 \cdot 2613$.

7. Для примера сравниваем два способа в случае г):

$$\begin{aligned} 2613 &= 2171 + 442 \\ 2171 &= 4 \cdot 442 + 403 \\ 442 &= 403 + 39 \\ 403 &= 10 \cdot 39 + 13 \\ 39 &= 3 \cdot 13 \end{aligned} \qquad \begin{aligned} 2613 &= 2172 + 442 \\ 2171 &= 5 \cdot 442 - 39 \\ 442 &= 11 \cdot 39 + 13 \\ 39 &= 3 \cdot 13 \end{aligned}$$

8. б)

$$\begin{aligned} &\text{НОД}\,(1010000110101, 100001111011) \\ &= \text{НОД}\,(110111010, 100001111011) \\ &= \text{НОД}\,(11011101, 100001111011) = \text{НОД}\,(11011101, 11110011110) \\ &= \text{НОД}\,(11011101, 1111001111) = \text{НОД}\,(11011101, 1011110010) \\ &= \text{НОД}\,(11011101, 101111001) = \text{НОД}\,(11011101, 10011100) \\ &= \text{НОД}\,(11011101, 100111) = \text{НОД}\,(10110110, 100111) \\ &= \text{НОД}\,(1011011, 100111) = \text{НОД}\,(110100, 100111) \\ &= \text{НОД}\,(1101, 100111) = \text{НОД}\,(1101, 11010) \\ &= \text{НОД}\,(1101, 1101) = 1101. \end{aligned}$$

в) Рассмотреть произведение ab и показать, что через каждые два шага произведение двух чисел, для которых вычисляется наибольший общий делитель, уменьшается, по крайней мере, в два раза. Таким образом, потребуется $O(\log a)$ шагов. Каждый шаг — это не более чем вычитание, требующее $O(\log a)$ двоичных операций (заметим, что ни делений, ни умножений алгоритм не содержит). г) Этот алгоритм не позволяет выразить НОД в виде целочисленной линейной комбинации двух исходных чисел. Однако его можно модифицировать так, что это будет возможным, см. статью G. H. Norton «Extending the binary GCD algorithm» в книге Algebraic Algorithms and Error Correcting Codes. Heidelberg etc.: Springer, 1986, с. 363–372.

9. $O(\log a \log b + \log^3 b)$.

10. а) Остаток уменьшается медленнее всего, когда все частные равны единице. б) Пишем $\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right) = BAB^{-1}$, где $A = \left(\begin{smallmatrix}\alpha&0\\0&\alpha'\end{smallmatrix}\right)$ — диагональная матрица из собственных значений, а столбцы матрицы B являются собственными векторами, т.е. $B = \left(\begin{smallmatrix}\alpha&\alpha\\1&1\end{smallmatrix}\right)$. в) Так как $\sqrt{5}a \geqslant \sqrt{5}f_{k+2} = \alpha^{k+2} - \alpha'^{k+2} > \alpha^{k+2} - 1$, то $k < (\log(1+\sqrt{5}a))/(\log\alpha) - 2$. Можно также получить более простую оценку $k < (\log a)/(\log\alpha)$. Последняя равна $1,44042\ldots \cdot \log_2 a$, в то время как оценка из доказательства предложения I.2.1 равна $2\log_2 a$.

11. б) К сумме величин $(\log r_i)(1+\log q_{i+1})$ применить неравенства $r_i \leqslant b$ и $\prod q_{i+1} \leqslant a$. Вывести, что эта сумма ограничена величиной $O((\log b)(\log a + \log a))$.

12. а) $x^4 + x^2 + 1 = (x^2)(x^2+1) + 1$, $1 = 1(x^4 + x^2 + 1) - x^2(x^2+1)$. б) $x^4 - 4x^3 + 6x^2 - 4x + 1 = (x-3)(x^3 - x^2 + x - 1) + (2x^2 - 2)$, $x^3 - x^2 + x - 1 = (\frac{1}{2}x - \frac{1}{2})(2x^2 - 2) + (2x - 2)$, $2x^2 - 2 = (x+1)(2x-2)$, так что НОД равен $x-1$; $x - 1 = (-\frac{1}{4}x + \frac{1}{4})f + (\frac{1}{4}x^2 - x + \frac{5}{4})g$.

13. НОД $(f, f') = x^2 - x - 1$, кратные корни — это $(1 \pm \sqrt{5})/2$ («золотое сечение» и сопряженное с ним число).

14. а) $5 + 6i = 2i(3-2i) + 1$, $1 = 1(5+6i) - 2i(3-2i)$. б) $8 - 19i = 2(7-11i) + (-6+3i)$, $7 - 11i = (-2+i)(-6+3i) + (-2+i)$, $-6 + 3i = 3(-2+i)$, так что НОД равен $-2+i$; $-2 + i = (-3+2i)(7-11i) + (2-i)(8-19i)$.

15. а) $12^2 + 25^2$. б) $54^2 + 31^2$. в) $116^2 + 159^2$.

§ I.3

1. а) $x = 6 + 7n$ при произвольном целом n. б) Нет решений. в) То же самое, что в а). г) $219 + 256n$. д) $36 + 100n$. е) $636 + 676n$.

2. 0, 1, 4, 9.

3. 3, B.

4. Разность между $n = 10^{k-1}d_{k-1} + \cdots + 10d_1 + d_0$ и суммой цифр $d_{k-1} + \cdots + d_1 + d_0$ есть сумма слагаемых, кратных числам вида $10^j - 1$, которые делятся на 9.

5. Доказать, что это число делится на 2, на 3 и на 5.

6. Пусть эти две цифры есть x и y. Тогда 72, а следовательно, и 8, и 9 делят сумму $1000x + 60 + y$ центов. Таким образом, $8|(60+y)$, а это означает, что $y = 4$. Далее, $9|(1000x + 64)$, что сравнимо с $x+1$ по модулю 9. Поэтому $x = 8$. Итак, каждая плитка стоит \$ 1,12.

7. а) Например, предположим, что $m = 2p^\alpha$. Так как $m|(x^2-1) = (x+1)(x-1)$, то произведение чисел $x+1$ и $x-1$ должно делиться на p^α. Но так как $p \geqslant 3$, то p не может делить и $x+1$, и $x-1$ (которые отличаются всего на два), и значит, p^α должно делить одно из этих чисел. Если $p^\alpha|(x+1)$, то это означает, что $x \equiv -1 \pmod{p^\alpha}$; если $p^\alpha|(x-1)$, то $x \equiv 1 \pmod{p^\alpha}$. Наконец, так как $2|(x^2-1)$, то x должно быть нечетным, т.е. $x \equiv 1 \equiv -1 \pmod 2$. Итак, по свойству 5 для сравнений либо $x \equiv 1 \pmod{2p^\alpha}$, либо $x \equiv -1 \pmod{2p^\alpha}$. б) Во-первых, если $m \geqslant 8$ есть степень двойки, то легко показать, что $x = m/2 + 1$ приводит к противоречию с частью а). Затем предположим, что m не является степенью простого числа (или удвоенной степенью простого числа) и что $p^\alpha \parallel m$. Положим $m' = m/p^\alpha$. Воспользуемся китайской теоремой об остатках, чтобы найти такое x, которое было бы сравнимо с 1 по модулю p^α и сравнимо с -1 по модулю m'. Показать, что это противоречит пункту а).

8. Разобьем все числа от 1 до $p-1$ на пары взаимно обратных по умножению. На основании упражнения 7а) лишь 1 и -1 совпадают со своими обратными.

Таким образом, когда эти $p-1$ чисел перемножаются, каждая пара дает в произведении единицу и остаются еще 1 и -1.

9. Разумеется, 4 обладает нужными свойствами, но не является трехзначным числом. В силу последнего утверждения китайской теоремы об остатках любое другое число, имеющее нужные остатки, должно отличаться от 4 на число, кратное $7 \cdot 9 \cdot 11 = 693$. Единственное такое трехзначное число — это $4 + 693 = 697$.

10. Можно применить китайскую теорему об остатках к сравнениям $x \equiv 1 \pmod{11}$, $x \equiv 2 \pmod{12}$, $x \equiv 3 \pmod{13}$. Другой способ: можно заметить, что -10 дает нужные остатки, и затем, рассуждая, как в упражнении 9, получить $-10 + 11 \cdot 12 \cdot 13 = 1706$.

11. а) 1973. б) 63841. в) 58837.

12. Частное дает остатки 5, 1, 4 при делении на 9, 10, 11, и, следовательно (по китайской теореме об остатках), имеет вид $851 + 990m$. Соответственно, делитель имеет вид $817 + 990n$. Так как делитель трехзначный, то $n = 0$. Так как произведение шестизначное, то и $m = 0$. Ответ: 851.

13. Больше всего времени при применении китайской теоремы об остатках занимают: 1) вычисление M; 2) вычисление $M_i = M/m_i$ для каждого из r различных модулей; 3) отыскание обратного для M_i по модулю m_i при каждом i; 4) умножение $a_i M_i N_i$ в формуле для x при каждом i; 5) деление полученного x на M для получения наименьшего неотрицательного значения. Для числа двоичных цифр в записи чисел m_i, a_i, N_i имеем оценку $O(\log B)$, а для числа двоичных цифр в записи чисел M и M_i имеем оценку $O(r \log B)$. Это дает оценку $O(r^2 \log^2 B)$ для числа двоичных операций при выполнении 1)–2), 4)–5). В 3) необходимо выполнить $O(r^2 \log^2 B)$ двоичных операций для приведения каждого M_i по модулю соответствующего m_i перед вычислением обратных и затем $O(r \log^3 B)$ двоичных операций для нахождения всех r обратных по алгоритму Евклида. Это дает в итоге оценку $O(r(r + \log B)\log^2 B)$. Который из двух членов этой оценки ($r^2 \log^2 B$ или $r \log^3 B$) является главным, зависит от соотношения величин r и $\log B$ (т. е. от числа уравнений и числа двоичных разрядов в модулях).

14. $38^{1+2+2^3+2^6} \equiv 38 \cdot 2 \cdot 16 \cdot 63 \equiv 79 \pmod{103}$.

15. Если использовать оценку $O(k^2)$ для числа операций при одном умножении k-значных двоичных чисел (как уже делалось), то никакого выигрыша в оценке мы не получим. Действительно, самое последнее умножение требует времени $O((n \log b)^2)$, что совпадает с оценкой для умножения b на себя n раз. Трудность состоит в том, что, в отличие от вычислений в кольце вычетов, в методе повторного возведения в квадрат мы подходим к концу, оперируя с парами очень больших чисел, что сводит на нет выигрыш от того, что количество умножений значительно меньше. Однако если бы мы воспользовались более рациональным способом умножения двух k-значных двоичных чисел, например, алгоритмом, требующим лишь $O(k \log k \log\log k)$ двоичных операций, это дало бы выигрыш во времени.

16. а) Повторное возведение в квадрат требует $O(\log^3 p)$ двоичных операций, в то время как для алгоритма Евклида может быть доказана оценка $O(\log^2 p)$. б) Повторное возведение в квадрат, по-прежнему, требует времени $O(\log^3 p)$, а когда мы выполним первый шаг алгоритма Евклида — деление p на a (что потребует $O(\log p \log a)$ двоичных операций) — оставшаяся часть алгоритма Евклида потребует $O(\log^2 a)$ двоичных операций. Таким образом, алгоритм Евклида действует быстрее, особенно при очень малых по сравнению с p числах a.

17.

n	90	91	92	93	94	95	96	97	98	99	100
$\varphi(n)$	24	72	44	60	46	72	32	96	42	60	40

18. Не существует такого n, для которого $\varphi(n)$ является нечетным числом,

большим 1; $\varphi(n) = 1$ для $n = 1$ или 2; $\varphi(n) = 2$ для $n = 3, 4, 6$; $\varphi(n) = 4$ для $n = 5, 8, 10, 12$; $\varphi(n) = 6$ для $n = 7, 9, 14, 18$; $\varphi(n) = 8$ для $n = 15, 16, 20, 24, 30$; $\varphi(n) = 10$ для $n = 11, 22$; $\varphi(n) = 12$ для $n = 13, 21, 26, 28, 36, 42$. Чтобы доказать, например, что здесь перечислены все n, для которых $\varphi(n) = 12$, сравним все допустимые разложения 12 на множители (в которых разрешается использовать множитель 1, но не 3) с формулой $\varphi(\prod p^\alpha) = \prod(p^\alpha - p^{\alpha-1})$. Получим $1 \cdot 2 \cdot 6$, $1 \cdot 12$, $2 \cdot 6$, 12. Первое произведение дает $2 \cdot 3 \cdot 7$, второе дает $2 \cdot 13$, третье дает (3 или 4)·7 и $4 \cdot 9$ и, наконец, четвертое дает 13.

19. Число n не может быть простым, так как в этом случае $\varphi(n) = n - 1$. По предположению n не является квадратом простого числа. Если n не является произведением двух различных простых чисел, то оно должно быть произведением не менее трех простых (не обязательно различных) чисел. Пусть p — наименьший из сомножителей. Тогда $p \leqslant n^{1/3}$, и $\varphi(n) \leqslant n(1 - p^{-1}) \leqslant n(1 - n^{-1/3}) = n - n^{2/3}$, что противоречит условию.

20. Показать, что квадрат любого нечетного числа сравним с 1 по модулю 8, а затем воспользоваться индукцией, как это сделано в первой части доказательства предложения I. 3. 5.

21. а) Заметим, что 360 кратно функции $\varphi(p^\alpha)$ для каждого $p^\alpha \| m$. В соответствии с замечанием перед примером 3 в тексте это означает, что $6647^{362} \equiv 6647^2 \equiv 44182609 \pmod m$. (Здесь мы используем тот факт, что НОД $(6647, m) = 1$, так как $6647 = 17^2 \cdot 23$.) б) Возведем a в степень 359 по модулю m методом повторного возведения в квадрат. Так как $m = (101100111)_2$, то нам понадобится 8 возведений в квадрат и 5 умножений (целых чисел, имеющих не более 63 двоичных разрядов). Каждая из этих операций сопровождается делением (в худшем случае 126-разрядного целого на 63-разрядное). Таким образом, число двоичных операций не превосходит $13 \cdot 63 \cdot 63 + 13 \cdot 64 \cdot 63 = 104013$.

22. а) Показать, что если $x = j \cdot \frac{n}{d}$, то x порождает S_d тогда и только тогда, когда НОД $(x, d) = 1$. Заметим, что j принимает значения $0, 1, \ldots, d - 1$. б) Разобьем множество элементов $\mathbf{Z}/n\mathbf{Z}$ на подмножества в соответствии с тем, какие множества S_d эти элементы порождают. Согласно п. а) подмножество, порождающее данное S_d, содержит $\varphi(d)$ элементов.

23. а) Представим каждый член произведения в виде суммы геометрической прогрессии $(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots)$. При перемножении таких скобок получаются все возможные дроби со знаменателями вида $p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}$. В силу основной теоремы арифметики каждое натуральное число n имеет единственное разложение такого вида. Следовательно, рассматриваемое произведение является суммой гармонического ряда $\sum_{n=1}^{\infty} \frac{1}{n}$, который, как известно, расходится. б) Сначала докажите, что $x > -\frac{1}{2}\log(1 - x)$ при $x \leqslant 1/2$ (см. график логарифма). Примените это неравенство к $x = 1/p$ и сравните $\sum \frac{1}{p}$ с логарифмом произведения из пункта а). в) Для каждой последовательности простых чисел n, стремящейся к бесконечности, $\frac{\varphi(n)}{n} = 1 - \frac{1}{n} \to 1$, а для каждой последовательности чисел n, делящихся на все возрастающее число последовательных простых (например, для $n_j = j!$), имеем $\frac{\varphi(n)}{n} = \prod_{p|n}(1 - \frac{1}{p}) \to \prod_{\text{все } p}(1 - \frac{1}{p}) = 0$ на основании пункта а).

24. а) Сообщить i-му генерал-лейтенанту p_i и вычет числа N по модулю p_i и воспользоваться китайской теоремой об остатках. б) Выбрать каждое p_i большим, чем $\sqrt[k]{N}$, но значительно меньшим, чем $\sqrt[k-1]{N}$.

§ I. 4

3. Рассуждая так же, как при доказательстве последнего предложения, показать, что $b^d \equiv \pm 1 \pmod m$. Но так как $(b^d)^{a/d} \equiv -1 \pmod m$, то $b^d \equiv -1 \pmod m$ и a/d нечетно.

4. Использовать упражнение 3 с $a = n$ и $c = (p-1)/2$.

5. а) $2^8 + 1 = 257$. б) Воспользоваться упражнением 4. с) $m = 97 \cdot 257 \cdot 673$.

6. $2 \cdot 11^2 \cdot 13 \cdot 4561$, $2^5 \cdot 5 \cdot 7 \cdot 13 \cdot 41 \cdot 73 \cdot 6481$.

7. $2^4 \cdot 3^2 \cdot 7 \cdot 13 \cdot 31 \cdot 601$.

8. $3^2 \cdot 41 \cdot 271$, $3^3 \cdot 7 \cdot 11 \cdot 13 \cdot 37$, $3^2 \cdot 11 \cdot 73 \cdot 101 \cdot 137$.

9. $7 \cdot 23 \cdot 89 \cdot 599479$; $7^2 \cdot 127 \cdot 337$ (этот пример показывает, что для простого $p|(b^d - 1)$ в предложении I. 4. 3 число $b^n - 1$ может делиться на p в большей степени, чем $b^d - 1$).

10. $7 \cdot 31 \cdot 151$, $3^2 \cdot 7 \cdot 11 \cdot 31 \cdot 151 \cdot 331$, $3^2 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 31 \cdot 41 \cdot 61 \cdot 151 \cdot 331 \cdot 1321$.

11. а) Применить одновременно алгоритм Евклида для нахождения НОД $(a^m - 1, a^n - 1)$ и НОД (m, n). Заметить, что на каждом шаге остаток в первом применении алгоритма Евклида имеет вид $a^r - 1$, где r — остаток во втором применении алгоритма Евклида. В частности, на первом шаге при делении $a^m - 1$ на $a^n - 1$ получается остаток $a^r - 1$, где r — остаток от деления m на n. б) Согласно части а) и китайской теореме об остатках никакие два числа в промежутке от 0 до $\prod(2^{m_i} - 1)$ не могут иметь два одинаковых набора остатков. Это произведение больше, чем $2^{rl/2} > 2^{2k} > ab$. При оценке времени получаем r умножений чисел, имеющих не более l цифр в двоичной записи. Это дает $O(rl^2) = O(kl)$ двоичных операций, что в r раз лучше, чем дает обычное умножение a на b, требующее $O(k^2)$ операций.

§ II. 1

1.

простое p	2	3	5	7	11	13	17
наименьший образующий	1	2	2	3	2	2	3
число образующих	1	1	2	2	4	4	8

2. а) При $g^{p-1} \equiv 1 \pmod{p^2}$ заменить g на $(1+p)g$ и показать, что $g^{p-1} = 1 + g_1 p$, где g_1 и p взаимно просты. Далее, если $g^j \equiv 1 \pmod{p^\alpha}$, то сначала показать, что $(p-1)|j$, т. е. $j = (p-1)j_1$ и, таким образом, $(1+g_1 p)^{j_1} \equiv 1 \pmod{p^\alpha}$. Пользуясь равенством $(1+g_1 p)^{j_1} = 1 + j_1 g_1 p +$ кратные p^2, показать, что j_1 делится на $p^{\alpha-1}$. б) Первая часть — см. упражнение 20 к § I. 3; вторая часть (сводящаяся к доказательству того, что 5^j не сравнимо с 1 по модулю 2^α, если только $2^{\alpha-2}$ не делит j) аналогична части а).

3. 5^6.

4. 2 для $d = 1$: X, $X + 1$; 1 для $d = 2$: $X^2 + X + 1$; 2 для $d = 3$: $X^3 + X^2 + 1$, $X^3 + X + 1$; 3 для $d = 4$: $X^4 + X^3 + 1$, $X^4 + X + 1$, $X^4 + X^3 + X^2 + X + 1$; 6 для $d = 5$: $X^5 + X^3 + 1$, $X^5 + X^2 + 1$, $X^5 + X^4 + X^3 + X^2 + 1$, $X^5 + X^4 + X^3 + X + 1$, $X^5 + X^4 + X^2 + X + 1$, $X^5 + X^3 + X^2 + X + 1$; 9 для $d = 6$: $X^6 + X^5 + 1$, $X^6 + X^3 + 1$, $X^6 + X + 1$, $X^6 + X^5 + X^4 + X^2 + 1$, $X^6 + X^5 + X^4 + X + 1$, $X^6 + X^5 + X^3 + X^2 + 1$, $X^6 + X^5 + X^2 + X + 1$, $X^6 + X^4 + X^3 + X + 1$, $X^6 + X^4 + X^2 + X + 1$.

5. 3 для $d = 1$: X, $X \pm 1$; 3 для $d = 2$: $X^2 + 1$, $X^2 \pm X + 1$; 8 для $d = 3$: $X^3 + X^2 \pm (X - 1)$, $X^3 - X^2 \pm (X + 1)$, $X^3 \pm (X^2 - 1)$, $X^3 - X \pm 1$; 18 для $d = 4$; 48 для $d = 5$; 116 для $d = 6$.

6. $(p^f - p^{f/l})/f$.

7. а) НОД $= 1 = X_g^2 + (X+1)f$. б) НОД $= X^3 + X^2 + 1 = f + (X^2 + X)g$. в) НОД $= 1 = (X-1)f - (X^2 - X + 1)g$. г) НОД $= X + 1 = (X-1)f - (X^3 - X^2 + 1)g$. д) НОД $= X + 78 = (50X + 20)f + (51X^3 + 26X^2 + 27X + 4)g$.

8. Так как НОД $(f, f') = X^2 + 1$, кратные корни — это $\pm\alpha^2$, где α — образующий элемент $\mathbf{F}_9^*$.

9. а) Возвести обе части $0 = \alpha^2 + b\alpha + c$ в p-ю степень и, используя равенства $b^p = b$ и $c^p = c$, получить $(\alpha^p)^2 + b\alpha^p + c = 0$. б) Два различных корня многочлена равны α и α^p. Следовательно, a — это взятая с минусом сумма, а b — произведение этих корней. в) $(c\alpha + d)^{p+1} = (c\alpha^p + d)(c\alpha + d)$; далее результат получается перемножением элементов в скобках с использованием б). г) $(2 + 3i)^{5(19+1)+1} = (2^2 + 3^2)^5(2 + 3i) = 14(2 + 3i) = 9 + 4i$

10. При каждом делении многочленов (сначала f на g, далее r_j на r_{j+1}) сначала нужно находить обратный по модулю p к старшему коэффициенту r_{j+1} (на это требуется $O(\log^3 p)$ двоичных операций), а затем нужно осуществить $O(d^2)$ умножений целых чисел по модулю p, каждое из которых производится за $O(\log^2 p)$ двоичных операций. Таким образом, на каждое деление уходит $O(\log^3 p + d^2 \log^2 p)$ двоичных операций, а на весь алгоритм Евклида — $O(d)O(\log^2 p(\log p + d^2)) = O(d \log^2 p(\log p + d^2))$ двоичных операций. (Это выражение можно упростить, записав его как $O(d \log^3 p)$, если d растет не быстрее, чем $\sqrt{\log p}$, или $O(d^3 \log^2 p)$, если p растет не быстрее, чем e^{d^2}.)

11. а) Пусть α — корень $X^2 + X + 1 = 0$, т. е. три последовательные степени α — это α, $\alpha+1$ и 1. б) Пусть α — корень $X^3 + X + 1 = 0$; тогда 7 последовательных степеней α — это $\alpha, \alpha^2, \alpha+1, \alpha^2+\alpha, \alpha^2+\alpha+1, \alpha^2+1, 1$. в) Пусть α — корень $X^3 - X - 1 = 0$; тогда 26 последовательных степеней α — это $\alpha, \alpha^2, \alpha+1, \alpha^2+\alpha, \alpha^2+\alpha+1, \alpha^2-\alpha+1, -\alpha^2-\alpha+1, -\alpha^2-1, -\alpha+1, -\alpha^2+\alpha, \alpha^2-\alpha-1, -\alpha^2+1, -1$ и далее те же 13 элементов, в которых все «+» заменены «−» и наоборот. г) Пусть α — корень $X^2 - X + 2 = 0$; тогда 24 последовательные степени α — это $\alpha, \alpha-2, -\alpha-2, 2\alpha+2, -\alpha+1, 2$, затем те же 6 элементов, умноженные на 2, затем умноженные на -1, затем умноженные на -2, — итого все 24 степени α.

12. $O(f2^f)$. Действительно, для каждой из $O(2^f)$ степеней α нужно каждое предыдущее выражение в виде многочлена от α умножить на α и, если возникнет α^f, то заменить его в полученном выражении, добавив к оставшимся членам многочлен, выражающий α^f через более низкие степени α; все это составляет $O(f)$ двоичных операций.

13. а) $p = 2$ и $2^f - 1$ — простое число Мерсенна (см. пример 1 и упражнение 2 к § I.4). б) Ясно, что условия выполняются в случае а), а также когда $p = 3$ и $(3^f - 1)/2$ — простое число (требуется, как в п. а), чтобы f было простым числом, однако этого недостаточно, как показывает пример $f = 5$) или когда $p = 2p' + 1$ при $f = 1$ и p' простом. Впрочем, неизвестно, является ли бесконечным число простых полей, удовлетворяющих условиям пунктов а) или б) (хотя предполагают, что это именно так). Простые p', для которых $2p' + 1$ — такое простое, называются «простыми Жермен» по имени Софи Жермен, которая в 1823 г. доказала, что большая теорема Ферма справедлива, если в показателе — такое простое число.

14. Выбрать последовательность n_j так, чтобы $\varphi(n_j)/n_j \to 0$ при $j \to \infty$ (см. упражнение 23 к § I.3) и p не делило никакое n_j, и положить f_j равным *порядку* p по модулю n_j, т. е. наименьшему числу, для которого $p^{f_j} \equiv 1 \pmod{n_j}$.

15. Любой многочлен, в котором X^j встречается с ненулевым коэффициентом лишь при $p|j$.

16. Свести к случаю $j = d$, показав, что если $\sigma^j(a) = a$ и $\sigma^f(a) = a$, то и $\sigma^d(a) = a$ (см. доказательство предложения I.4.2). Заметить, что поле $\mathbf{F}_{p^d}$,

которое есть поле разложения $X^{p^d} - X$, содержится в $\mathbf{F}_q$, так как любой корень этого многочлена также удовлетворяет уравнению $X^q - X$ (чтобы заметить это, нужно f/d раз возвести обе части равенства $a^{p^d} = a$ в степень p^d).

17. Во-первых, доказать, что элемент $b' = b^{(p^n-1)/(p^d-1)}$ принадлежит $\mathbf{F}_{p^d}$, показав, что он неподвижен относительно σ^d (т. е. возведя его в степень p^d). Далее убедиться, что он — образующий, так как степени $(b')^j$, $j = 0, 1, \ldots, p^d - 2$, все различны. Это следует из того, что первые $p^n - 1$ степеней элемента b различны.

§ II. 2

1. Множества вычетов: $\{1\}$ для $p = 3$, $\{1, 4\}$ для $p = 5$, $\{1, 2, 4\}$ для $p = 7$, $\{1, 3, 4, 9, 10, 12\}$ для $p = 13$, $\{1, 2, 4, 8, 9, 13, 15, 16\}$ для $p = 17$, $\{1, 4, 5, 6, 7, 9, 11, 16, 17\}$ для $p = 19$.

2. б) Из части а) и предложений II. 2. 2 и II. 2. 4 следует, что $(\frac{2}{p}) = 1 \equiv 2^{(p-1)/2} \pmod p$. Это означает, что $((p-1)/2^l)$-я степень 2 сравнима с -1 по модулю p для некоторого $l \geqslant 2$. Так как $2^{2^k} \equiv -1 \pmod p$, можно показать, что НОД $((p-1)/2^l, 2^k) = 2^k$. Отсюда сразу следует, что $p \equiv 1 \pmod{2^{k+l}}$. в) Единственным простым числом, сравнимым с 1 по модулю 64 и меньшим, чем $\sqrt{65537}$, является 193, однако оно не делит 65537.

3. НОД $(84, 1330) = 14$.

4. Использовать равенство $(\frac{-2}{p}) = (\frac{-1}{p})(\frac{2}{p})$ и рассмотреть 4 возможных случая $p \equiv 1, 3, 5, 7 \pmod 8$.

5. $(\frac{91}{167}) = (\frac{7}{167})(\frac{13}{167}) = -(\frac{167}{7})(\frac{167}{13}) = -(\frac{-1}{7})(\frac{-2}{13}) = -(-1)(-1) = -1$.

6. а) 14. б) 9. в) 9α.

7. $a^3 - a$ (см. доказательство предложения II. 2. 4); 6, 60, 4080, 24, 210, 336.

8. Примитивный p-й корень ξ из единицы в $\mathbf{F}_q$ существует, так как $q \equiv 1 \pmod p$. Далее, квадрат числа $G = \sum_{j=1}^{p-1} (\frac{j}{p})\xi^j$ равен $(\frac{-1}{p})p$ (см. лемму в доказательстве предложения II. 2. 5).

9. а) $(\frac{-1}{p})\sum_{j=1}^{p-1}(\frac{j}{p})a^j$; 6, 45, 3126, 906 (в последнем случае использовать равенство $1093 = (3^7 - 1)/2$). б) Пусть $G = \sum_{j=1}^{p-1}(\frac{j}{p})2^j$. Тогда наименьший положительный квадратный корень из $(\frac{-1}{p})p$ по модулю $2^p - 1$ равен g, если $p \equiv 5 \pmod 8$, равен $-g$, если $p \equiv 3 \pmod 8$, равен $p + g$, если $p \equiv 7 \pmod 8$, и равен $p - g$, если $p \equiv 1 \pmod 8$.

10. а) $(\frac{1801}{8191}) = (\frac{8191}{1801}) = (\frac{987}{1801}) = (\frac{3}{1801})(\frac{7}{1801})(\frac{47}{1801}) = (\frac{1}{3})(\frac{2}{7})(\frac{15}{47}) = 1 \cdot 1 \cdot (\frac{3}{47})(\frac{5}{47}) = -(\frac{2}{3})(\frac{2}{5}) = -1$. б) $(\frac{987}{1801}) = (\frac{1801}{987}) = (\frac{2 \cdot 407}{987}) = -(-1)(\frac{987}{407}) = (\frac{173}{407}) = (\frac{407}{173}) = (\frac{61}{173}) = (\frac{173}{61}) = (\frac{51}{61}) = (\frac{61}{51}) = (\frac{2 \cdot 5}{51}) = -(\frac{5}{51}) = -(\frac{51}{5}) = -1$.

11. а) 1. б) 1. в) 1. г) 1. д) 1. е) 1. ж) -1.

12. а) $(\frac{-3}{p}) = (\frac{-1}{p})(\frac{3}{p}) = (-1)^{(p-1)/2}(-1)^{(3-1)(p-1)/4}(\frac{p}{3}) = (\frac{p}{3})$; это равно 1 тогда и только тогда, когда $p \equiv 1 \pmod 3$. б) $(\frac{3}{2^p-1}) = -(\frac{2^p-1}{3}) = -(\frac{1}{3}) = -1$.

13. Последняя десятичная цифра должна быть 1 или 9.

14. Любая степень вычета есть вычет, поэтому никакой невычет не может встретиться среди этих степеней, а это означает, что вычет не может быть образующим.

15. а) Так как $p-1$ есть степень 2, то порядок всякого элемента g есть степень 2. Если g — невычет, то $-1 = (\frac{g}{p}) \equiv g^{(p-1)/2} \pmod p$. Поэтому порядок g не может быть меньше $p - 1$. б) Если $k > 1$ и $p = 2^{2^k} + 1$, то $p \equiv 2 \pmod 5$ (так как степень, в которую возводится 2, делится на 4). Поэтому $(\frac{5}{p}) = (\frac{p}{5}) = (\frac{2}{5}) = -1$ и б) следует из а). в) Аналогично б): так как 2 возводится в степень, не делящуюся

на 3, то $p-1 \equiv 2$ или 4 (mod 7), следовательно, $p \equiv 3$ или 5 (mod 7). Поэтому $(\frac{7}{p}) = (\frac{p}{7}) = -1$, и в) следует из а).

16. а) Имеем: $(a+bi)^{p+1} = (a^p + b^p i^p)(a+bi) = (a-bi)(a+bi) = a^2 + b^2$. **Утверждение.** *Если* $(a+bi)^m \in \mathbf{F}_p$, *то* $(p+1)|m$. Чтобы доказать это утверждение, положим $d = \text{НОД}(m, p+1)$. Используя те же аргументы, что и в доказательстве предложения I. 4. 2, видим, что $(a+bi)^d \in \mathbf{F}_p$. Однако, так как $p+1$ есть степень 2, то в случае $d < p+1$ мы находим, что $(a+bi)^{(p+1)/2}$ — элемент $\mathbf{F}_p$, квадрат которого есть a^2+b^2. Но a^2+b^2 есть невычет (см. упражнение 14), поэтому $d = p+1$ и $(p+1)|m$. Пусть теперь n таково, что $(a+bi)^n = 1$. Тогда, по доказанному утверждению, $n = (p+1)n'$. Следовательно, $(a^2+b^2)^{n'} = 1$; так как a^2+b^2 — образующий $\mathbf{F}_p^*$, то $(p-1)|n'$. б) Достаточно показать, что 17 и 13 — образующие $\mathbf{F}_{31}$.

17. В обоих случаях получаем $O(\log^3 p)$. Заметим, однако, что предложение II. 2. 2 применимо к $(\frac{a}{n})$ лишь тогда, когда $n = p$ есть простое число. В то же время метод из части а) применим при любом нечетном натуральном n. Заметим также, что время работы при а) может быть сокращено до $O(\log^2 p)$ методом, использованным в упражнении 11 к § I. 2.

18. а) Показать, что число решений квадратичного сравнения равно числу решений сравнения $x^2 \equiv D \pmod p$. Имеется одно решение, если $D = 0$, ни одного, если D — невычет, и 2, если D — вычет. б) 0, 0, 2, 1, 2. в) 2, 2, 1, 0, 0.

19. $n = 3$, $p - 1 = 2^5 \cdot 65$, $r \equiv 302^{33} \equiv 203 \pmod p$ (мы вычисляем 302^{33} методом повторного возведения в квадрат, последовательно возводя в квадрат 5 раз и затем умножая результат на 302). Также методом повторного возведения в квадрат вычисляем $b \equiv n^{65} \equiv 888 \pmod p$. Наконец, берем $j = 2^2$, т. е. $\sqrt{302} \pmod p \equiv b^4 r \equiv 1292 \pmod p$.

20. а) Использовать индукцию по α. Чтобы перейти от $\alpha-1$ к α, предположить, что $\tilde{x}$ — такое $(\alpha-1)$-значное число в системе счисления с основанием p, что $\tilde{x}^2 \equiv a \pmod{p^{\alpha-1}}$. Чтобы определить последнюю цифру $x_{\alpha-1} \in \{0, 1, \ldots, p-1\}$ искомого решения $x = \tilde{x} + x_{\alpha-1}p^{\alpha-1}$, записываем $\tilde{x}^2 = a + bp^{\alpha-1}$ для некоторого целого числа b и затем действуем по модулю p^α: $x^2 = (\tilde{x} + x_{\alpha-1}p^{\alpha-1})^2 \equiv \tilde{x}^2 + 2x_0 x_{\alpha-1}p^{\alpha-1} = a + p^{\alpha-1}(b + 2x_0 x_{\alpha-1})$. Таким образом, достаточно выбрать $x_{\alpha-1} \equiv -(2x_0)^{-1}b \pmod p$ (заметим, что $2x_0$ обратимо по модулю p, так как p нечетно, а $x_0^2 \equiv a \pmod p$, где a взаимно просто с p). б) Использовать китайскую теорему об остатках, чтобы найти x, который сравним по каждому модулю p^α с квадратным корнем, найденным в части а).

21. а) Если бы $(*)$ было справедливо для b_1 и $b_1 b_2$, то, поделив одно сравнение на другое (это возможно, так как функции как справа, так и слева мультипликативны), мы получим, что $(*)$ выполняется также для b_2. С другой стороны, если бы $(*)$ не выполнялось для некоторого b, то множество элементов вида bc, состоящее из произведений b на все элементы c, для которых $(*)$ выполняется, состояло бы из элементов, не удовлетворяющих $(*)$. б) Взять, например, $b = 1 + n/p$, где $p^2|n$. Тогда $(\frac{b}{n}) = 1$, но $b^j \equiv 1$, только если $p|j$. Однако для $j = (n-1)/2$ это не так. в) Покажем, что $(\frac{b}{n}) = -1$, однако $b^{(n-1)/2} \equiv 1 \pmod{n/p}$ и, следовательно, сравнение $b^{(n-1)/2} \equiv -1 \pmod{n/p}$ (и, тем более, такое сравнение по модулю n) не может иметь места. Далее, пусть a_1 — любой невычет по модулю p и $a_2 = 1$. Теперь примените китайскую теорему об остатках, чтобы получить решение b для пары сравнений $x \equiv a_1 \pmod p$, $x \equiv a_2 \pmod{n/p}$.

22. $b^2 = (t+\alpha)^p(t+\alpha) = (t+\alpha^p)(t+\alpha) = (t-\alpha)(t+\alpha) = t^2 - \alpha^2 = a$, где третье из равенств следует из того, что $\alpha = \sqrt{t^2-a}$ сопряжено с $\alpha^p = -\sqrt{t^2-a}$;

заметим, что $b \in \mathbf{F}_p$, так как по условию a имеет два квадратных корня в $\mathbf{F}_p$, и потому его квадратные корни в $\mathbf{F}_{p^2}$ на самом деле лежат в $\mathbf{F}_p$.

23. Пусть b — наименьший положительный вычет $n^{(p-1)/4}$ по модулю p. Тогда b — квадратный корень из -1 по модулю p, т. е. $p|(b^2+1)$. Далее вычислить $c+di = \text{НОД}\,(p, b+i)$ (см. упражнение 14 к § I. 2).

§ III. 1

1. «We sewed a smile on a horse's ass, and a year later it was elected President.»

2. Вывести равенство $b = 19$ из того, что буква «X» — самая частая в шифртексте. Сообщение имеет вид: WEWERELUCKYBECAUSEOFTENTHEFREQ UENCYMETHODNEEDSLONGERCIPHERTEXT.

3. THRPXDH.

4. SUCCESSATLAST.

5. AGENT 006 IS DEAD 007.

6. Есть 9 возможных вариантов для a' и b': $a' = 1, 4, 7, 10, 13, 16, 19, 22, 25$ и $b' = 21, 6, 18, 3, 15, 0, 12, 24, 9$ соответственно. Ничего не остается, как просто перепробовать все эти 9 вариантов. Получаем следующие варианты открытого текста: «I DY IB RIF», «I PS IH RIX», «I AM IN RIO», «I MG IT RIF», «I YA IZ RIX», «I JV IE RIO», «I VP IK RIF», «I GJ IQ RIX», «I SD IW RIO». Осмысленный текст дает только третий вариант: $P \equiv 7C + 18 \pmod{27}$.

7. а) N. б) $N\varphi(N) = N^2 \prod_{p|N}(1 - \frac{1}{p})$. в) 312, 486, 812, 240.

8. а) Если $a \neq 1$, то сравнение $(a-1)P \equiv -b \pmod{N}$ имеет ровно одно решение в поле $\mathbf{F}_N = \mathbf{Z}/N\mathbf{Z}$. б) $P = 0$ всегда фиксированно. При четном N (тогда a нечетно) сравнение $(a-1)P \equiv 0 \pmod{N}$ имеет, по крайней мере, два решения: $P = 0$ и $P = N/2$. в) N четно, b нечетно. Более общий пример: b не делится на НОД $(a-1, N)$.

9. $N^2\varphi(N^2) = N^4 \prod_{p|N}(1 - \frac{1}{p})$; 210912; 354294; 682892; 216000.

10. а) $a' = 435$, $b' = 64$; «FOUNDTHEGOLD». б) $a = 115$, $b = 76$; «AWOFUWAE».

11. а) Из первых двух сравнений ключ найти невозможно, но, взяв разность первого и третьего сравнений, получим $139a' \equiv 247 \pmod{900}$, что дает $a' = 73$ и $b' = 768$. Сообщение: «ARE YOU JOKING?». б) $a = 37$, $b = 384$; «FWU ORI DCCUVGA».

12. «CCCP».

13. Сравнение $P \equiv 37P + 384 \pmod{900}$ влечет за собой $3P \equiv 43 \pmod{75}$; таких биграмм нет.

14. а) Композиция $I \equiv P + b_1 \pmod{N}$ и $C \equiv I + b_2 \pmod{N}$ — это $C \equiv P + b \pmod{N}$ с $b = b_1 + b_2$. б) Композиция будет линейным преобразованием с $a = a_1 \cdot a_2$. в) Композиция будет аффинным преобразованием с $a = a_1 \cdot a_2$ и $b = a_2 \cdot b_1 + b_2$.

15. $P \equiv 642C + 187 \pmod{853}$; «DUMB IDEA».

16. Сначала найти, что $I \equiv 201C + 250 \pmod{881}$, а потом — что $P \equiv 331I + 257 \pmod{757}$; «NO RETREAT».

§ III. 2

1. При шифровании использовано ключевое слово «SPY». Открытый текст (здесь для удобства вставлены пробелы и знаки пунктуации): «I had asked that a cable from Washington to New Delhi summarizing the results of the aid consortium be repeated to me through the Toronto Consulate. It arrived in code; no facilities

existed for decoding. They brought it to me at the airport — a mass of numbers. I asked if they assumed I could read it. They said no. I asked how they managed. They said when something arrived in code, they phoned Washington and had the original message read to them.» (Джон Кеннет Гэлбрайт «Дневник посла»; цитируется по работе *Mellen G. E.* Cryptology, computers and common sense. — In: Computers and Security. Vol. III).

2. а) $\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$. б) $\begin{pmatrix} 19 & 10 \\ 23 & 16 \end{pmatrix}$. в) $\begin{pmatrix} 11 & 11 \\ 24 & 1 \end{pmatrix}$. г) $\begin{pmatrix} 820 & 0 \\ 0 & 801 \end{pmatrix}$. д) $\begin{pmatrix} 127 & 303 \\ 546 & 353 \end{pmatrix}$.

3. а) $\binom{6}{1}$. б) Решений нет (умножение второго сравнения на 2 и вычитание результата из первого дает $6y \equiv 8 \pmod 9$, что невозможно, так как влечет за собой $3|8$). в) $\binom{6}{1}$, $\binom{3}{4}$, $\binom{0}{7}$. г) $\binom{0}{0}$, $\binom{6}{3}$, $\binom{3}{6}$.

4. а) $\binom{9}{21}$. б) $\binom{0}{0}$. в) Любой вектор с $y = x$, т.е. $\binom{0}{0}$, $\binom{1}{1}$, $\binom{2}{2}$ и т.д. г) Любой вектор вида $\binom{n}{15+n}$. д) Решений нет.

5. а) $\binom{787}{759}$. б) $\binom{626}{233}$. в) $\binom{0}{0}$. г) $\binom{0}{0}$, $\binom{101}{505}$, $\binom{202}{1010}$, $\binom{303}{404}$, $\binom{404}{909}$, $\binom{505}{303}$, $\binom{606}{808}$, $\binom{707}{202}$, $\binom{808}{707}$, $\binom{909}{101}$, $\binom{1010}{606}$. д) Сложить вектор $\binom{31}{800}$ с любым из 11 векторов п. г) и привести по модулю 1111.

6. Воспользоваться методом математической индукции: сначала проверить, что утверждение верно при $n = 1, 2, \ldots, b$, а затем показать, что если оно верно при n, то оно верно при $n + b$. А именно, убедиться в том, что

$$\begin{aligned}\begin{pmatrix} f_{n+b+1} & f_{n+b} \\ f_{n+b} & f_{n+b-1} \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n+b} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{b} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n} \\ &= \begin{pmatrix} f_{b+1} & f_b \\ f_b & f_{b-1} \end{pmatrix} \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} cf_{n+1} & cf_n \\ cf_n & cf_{n-1} \end{pmatrix} \pmod a,\end{aligned}$$

где $c \in (\mathbf{Z}/a\mathbf{Z})^*$, и воспользоваться предположением индукции. (Можно показать, что для *любого* целого a найдется такое b, что $a|f_n \Leftrightarrow b|n$, и что если $a = p^\alpha$ при простом $p \neq 5$, то b делит $p^{\alpha-1}(p^2-1)$; доказательство использует некоторые сведения из алгебраической теории чисел для квадратичного расширения поля вещественных чисел, порожденного отношением «золотого сечения» $(1+\sqrt{5})/2$; заметим, что это число и сопряженное с ним число $(1-\sqrt{5})/2$ являются собственными значениями матрицы из определения чисел Фибоначчи.)

7. $A^{-1} = \begin{pmatrix} 23 & 7 \\ 18 & 5 \end{pmatrix}$, «SENATORTOOK.»

8. $A^{-1} = \begin{pmatrix} 22 & 16 \\ 21 & 17 \end{pmatrix}$, «MEET AT NOON.»

9. $A^{-1} = \begin{pmatrix} 22 & 20 \\ 28 & 8 \end{pmatrix}$, «WHY NO GO? MARIA»; $A = \begin{pmatrix} 3 & 7 \\ 4 & 1 \end{pmatrix}$, «JMLD W EFWJV».

10. «СЛАВА КПСС».

11. Композиция криптосистем имеет матрицу шифрования A_2A_1.

12. а) «?CVK». б) Применить матрицу $\begin{pmatrix} 18 & 28 \\ 19 & 20 \end{pmatrix}$ к вектору шифртекста, действуя в кольце вычетов по модулю 29, а к результату применить матрицу $\begin{pmatrix} 15 & 15 \\ 22 & 3 \end{pmatrix}$,

действуя в кольце вычетов по модулю 26; «STOP».

13. В силу предложения III.2.1 (а именно, если нарушается b), то нарушается и c)) существует ненулевой вектор, переводимый матрицей A в $\binom{0}{0}$. Такую биграмму-вектор можно прибавить к любой биграмме-вектору открытого текста без изменения соответствующего шифртекста.

14. Шифртекст имеет вид $\begin{pmatrix} 18 & 6 & 11 & 10 & 29 & 14 & 16 & 11 & 14 & 10 & 11 & 21 \\ 26 & 13 & 8 & 3 & 10 & 25 & 11 & 8 & 12 & 20 & 27 & 24 \end{pmatrix}$, а последние три столбца (10-й, 11-й и 12-й) открытого текста имеют вид $\begin{pmatrix} 10 & 17 & 0 \\ 0 & 11 & 27 \end{pmatrix}$. Определитель матрицы, образованной 10-м и 11-м столбцами открытого текста, равен 20 (mod 30), что не имеет обратного по модулю 30, но имеет обратное по модулю 3. Определитель матрицы, образованной 11-м и 12-м столбцами открытого текста, равен 9 (mod 30) и не обратим по модулю 30, но обратим по модулю 10. Производя в первом случае операции по модулю 3, получаем $A^{-1} \pmod 3 = \begin{pmatrix} 10 & 17 \\ 0 & 11 \end{pmatrix} \cdot \begin{pmatrix} 10 & 11 \\ 20 & 27 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Аналогично, во втором случае, производя операции по модулю 10, получаем $A^{-1} \equiv \begin{pmatrix} 4 & 9 \\ 5 & 8 \end{pmatrix}$. По китайской теореме об остатках имеется единственная матрица $A^{-1} \pmod{30}$, удовлетворяющая этим двум сравнениям: $A^{-1} = \begin{pmatrix} 4 & 9 \\ 25 & 28 \end{pmatrix}$. Открытый текст «GIVE THE PLANS TO KARLA.».

15. Шифртекст имеет вид $\begin{pmatrix} 10 & 22 & 26 & 0 & 10 & 1 & 5 & 17 \\ 21 & 27 & 19 & 28 & 9 & 27 & 21 & 26 \end{pmatrix}$, а первые три столбца открытого текста имеют вид $\begin{pmatrix} 2 & 8 & 0 \\ 29 & 29 & 29 \end{pmatrix}$. При использовании равенства $A^{-1} = PC^{-1}$ учесть, что наибольший общий делитель определителя матрицы из первых двух столбцов шифртекста и 30 равен 6. Лучше использовать 1-й и 3-й столбцы: $\det \begin{pmatrix} 10 & 26 \\ 21 & 19 \end{pmatrix} = 4$ и НОД $(4, 30) = 2$. Взяв эту матрицу в качестве C и оперируя по модулю 15, получить равенство $A^{-1} = \begin{pmatrix} 2 & 2 \\ 8 & 4 \end{pmatrix} + 15A_1$, где $A_1 \in M_2(\mathbf{Z}/2\mathbf{Z})$. Воспользовавшись тем, что $A^{-1}\begin{pmatrix} 10 & 22 & 26 \\ 21 & 27 & 19 \end{pmatrix} = \begin{pmatrix} 2 & 8 & 0 \\ 29 & 29 & 29 \end{pmatrix}$, и тем, что $\det A^{-1}$ нечетен, показать, что либо $A^{-1} = \begin{pmatrix} 17 & 2 \\ 8 & 19 \end{pmatrix}$, либо $A^{-1} = \begin{pmatrix} 17 & 2 \\ 23 & 19 \end{pmatrix}$. В первом варианте получаем открытый текст «C.I.A. WILLLHTLA», а во втором — открытый текст «C.I.A. WILL HELP».

16. Использовать китайскую теорему об остатках.

17. $(p^2 - 1)(p^2 - p)$.

18. Определитель взаимно прост с p^α в том и только том случае, когда он взаимно прост с p; $p^{4\alpha-3}(p^2 - 1)(p - 1)$.

19. $N^4 \prod_{p|N}(1 - \frac{1}{p})(1 - \frac{1}{p^2})$; 157248, 682080, 138240.

20. $N^{(k^2)} \prod_{p|N}((1 - \frac{1}{p})(1 - \frac{1}{p^2}) \cdots (1 - \frac{1}{p^k}))$.

21. $N^6 \prod_{p|N}(1 - \frac{1}{p})(1 - \frac{1}{p^2})$; 106 299 648; 573 629 280; 124 416 000.

22. а) $(p^2 - 1)(p^2 - p)$. б) $p^2 - p$.

23. а) $A_0 = \begin{pmatrix} 21 & 27 \\ 18 & 27 \end{pmatrix}$. б) $\binom{1}{1}$. в) Шесть матриц (ср. с ответом к упражнению

22 б) при $p = 3$): $A = \begin{pmatrix} a & 7 \\ c & 7 \end{pmatrix}$, где $\binom{a}{c} = \binom{21}{28}, \binom{21}{8}, \binom{1}{18}, \binom{1}{8}, \binom{11}{18}, \binom{11}{28}$.

24. а) НОД $(\det(A-I), N) = 1$, где $\det(A-I) = (a-1)(d-1) - bc$ (применить утверждение об эквивалентности условий а) и с) в предложении III. 2. 1 с заменой A на $A - I = \begin{pmatrix} a-1 & b \\ c & d-1 \end{pmatrix}$). б) Пусть $\mathbf{F}_N$ — поле $\mathbf{Z}/N\mathbf{Z}$. Биграммы образуют двумерное векторное пространство, а фиксированные биграммы — подпространство в нем. Любое подпространство, содержащее более одного вектора, либо одномерно (и содержит N векторов), либо содержит все биграммы (тогда $A = I$).

25. а) $P = A'C + B'$, $A' = \begin{pmatrix} 14 & 781 \\ 821 & 206 \end{pmatrix}$, $B' = \binom{322}{202}$; «HIT ARMY BASE! HEADQUARTERS». б) $C = AP+B$, $A = \begin{pmatrix} 103 & 30 \\ 10 & 7 \end{pmatrix}$, $B = \binom{301}{412}$; «!NJUFYKTEGOUL IB!VFEXU!JHALGQGJ?».

26. $29^8(29^2 - 1)(29^2 - 29)$=341 208 073 352 438 880.

27. 91 617 661 629 000 000.

28. $A^{-1} = \begin{pmatrix} 18 & 21 & 19 \\ 13 & 18 & 3 \\ 3 & 19 & 11 \end{pmatrix}$; «SENDROSESANDCAVIARJAMESBOND».

§ IV.1

1. $\binom{m}{2} = m(m-1)/2$ для классической криптосистемы и m для системы с открытым ключом. При $m = 1000$ это дает 499500 против 1000.

2. Один из возможных способов таков. Инвесторы и брокеры используют систему с $\mathcal{P} = \mathcal{C}$. Пользователь A посылает сообщение пользователю B, предварительно преобразуя каждый элемент сообщения P в $f_B f_A^{-1}(P)$. Каждое сообщение включает в себя свой идентификационный номер. Пользователь B должен сразу же уведомить о получении сообщения с указанием идентификационного номера сообщения, полученного от A. При этом каждый элемент P уведомляющего сообщения перед отправлением преобразуется в $f_A f_B^{-1}(P)$ (аналогично двойному шифрованию, использованному A при посылке исходного сообщения). Если A не получил уведомление от B, то он посылает свое сообщение повторно, и так до тех пор, пока не получит ответ. Потом, когда из-за потери капиталов или по другим причинам возникнет спор о том, кто именно и какое послал сообщение, брокер сможет доказать, что сообщение было послано A, ибо никто, кроме A (и суда), не может создать сообщение, читаемое с помощью преобразования $f_A f_B^{-1}$. Аналогично, A сможет доказать, что сообщение с данным идентификационным номером было принято пользователем B (так как никто другой не мог бы послать уведомительное сообщение) и от B можно потребовать предоставить полученное сообщение суду.

3. В данном случае подойдет криптосистема с открытым ключом, использующая случайные целые числа (удовлетворяющие, возможно, некоторым условиям) для построения ключей шифрования и дешифрования при помощи какого-нибудь алгоритма. Компьютер программируется на генерацию случайных целых чисел, которые используются для получения пары ключей $K = (K_E, K_D)$. Компьютер передает K_D (*но не* K_E) во внешний мир и сохраняет K_E (*но не* K_D) у себя. Поэтому любой может читать его сообщения, но никто не сможет создать сообщение, дешифруемое с помощью ключа K_D. (Эта ситуация обратна по отношению к обычной криптографии с открытым ключом. Там любой может послать сообщение, но только пользователь может прочитать его.) Для совместно работающих ученых вполне возможно запрограммировать компьютер так, чтобы никто не мог

ни предсказать генерируемые случайные числа, ни продублировать его работу на своем собственном компьютере. (Отметим глубокий реализм этого примера, который предполагает, что обе страны абсолютно не доверяют друг другу, но в то же время полностью доверяют компьютерам.)

4. Бьерн выбирает случайно элемент $p \in \mathcal{P}$, вычисляет $c = f(p)$ и посылает число c Анюте. Анюта вычисляет два прообраза p_1 и p_2 и посылает один из них, скажем, p_1, Бьерну. Если $p_1 \neq p$, то Бьерн может назвать оба прообраза p_1 и $p_2 = p$. В этом случае считается, что он выигрывает. В противном случае выигрывает Анюта. Если Анюта выиграла, то она должна предоставить второй прообраз, чтобы Бьерн мог проверить равенство $f(p_2) = c$ и удостовериться, что Анюта не смошенничала, выбрав неправильный ключ, для которого c имеет только один прообраз. (Анюте невыгодно выбирать ключ, при котором у каждого c имеется более двух прообразов, так как это уменьшает ее шансы послать Бьерну тот прообраз, который он уже знает.)

§IV.2

1. а) BH A 2AUCAJEAR0. б) $2047 = 23 \cdot 89$ (см. пример 1 в §I.4), $d_A = 411$. в) так как $\varphi(23)$ и $\varphi(89)$ имеют небольшое наименьшее общее кратное 88, то любой обратный к 179 по модулю 88 (например, 59) будет действовать как d_A.

2. Число n_A является произведением простого числа Мерсенна 8191 и простого числа Ферма 65537. Это чрезвычайно плохой выбор. Имеем $d_A = 201934721$, «DUMPTHESTOCK».

3. а) STOP PAYMENT. б) (1) 6043; (2) $n = 113 \cdot 191$.

4. С третьей попытки ($t = 152843,\ 152844,\ 152845$) находим $t^2 - n = 804^2$, тогда $p = 152845 + 804 = 153649$, $q = 152845 - 804 = 152041$.

5. Чтобы показать невозможность вычисления ассоциированного элемента в $\mathcal{P}$ (т. е. элемента, имеющего тот же самый образ, что и данный элемент), предположим, что некто, знающий лишь K_E (т. е. число n, но не его разложение) получил вторую пару $\pm x_2$ с тем же квадратом по модулю n, что и $\pm x_1$. Показать, что тогда НОД $(x_1 + x_2, n)$ есть либо p, либо q. Другими словами, отыскание *одной* лишь пары ассоциированных элементов в $(\mathbf{Z}/n\mathbf{Z})^* / \pm 1$ равносильно разложению n на множители.

6. Достаточно показать, что $a^{de} \equiv a \pmod{p}$ для любого целого a и каждого простого делителя p числа n. Это очевидно, если $p|a$. В других случаях воспользоваться малой теоремой Ферма (предложение I.3.2).

7. Если $m/2 \equiv (p-1)/2 \pmod{p-1}$, то $a^{m/2} \equiv \left(\frac{a}{p}\right)$, что в половине случаев равно $+1$, а в другой половине случаев равно -1. В случае (2) при помощи китайской теоремы об остатках показать, что событие, состоящее в том, что элемент в $(\mathbf{Z}/n\mathbf{Z})^*$ является вычетом по модулю p, и событие, состоящее в том, что он является вычетом по модулю q, не зависят друг от друга. Таким образом, ситуация случая (2) подобна независимому бросанию двух монет.

§IV.3

1. а) 24, 30, 11, 13. б) $1, \alpha^2 + \alpha, \alpha, \alpha + 1$.

2. 1) Чтобы оправдать переход к сравнению $2^x a \equiv 1$, заметим, что если $x < \varphi(3^\alpha)$ является решением $2^x a \equiv 1 \pmod{3^\alpha}$, то $\varphi(3^\alpha) - x$ является решением исходного сравнения. Если $a \equiv 2 \pmod 3$, то решаем $2^x(2a) \equiv 1 \pmod{3^\alpha}$, где $2a \equiv 1 \pmod 3$. Тогда $x + 1$ является решением исходного сравнения. Если $a \equiv 1 \pmod 3$, то решение x должно быть четным, так как $2^{\text{нечетн}} \equiv 2 \pmod 3$.

3) Для доказательства того, что $(*)_j$ выполняется при $x_{j-2} = (1 - a_{j-1})/3^{j-1}$,

вычислить левую часть $(*)_j$ по модулю 3^j. Она равна $a_{j-1}g_{j-1}^{x_{j-2}} \equiv (1 - 3^{j-1}x_{j-2})g_{j-1}^{x_{j-2}}$. Далее посредством биномиального разложения показать, что $(1+3)^{3^{j-2}x_{j-2}} \equiv 1 + 3^{j-1}x_{j-2} \pmod{3^j}$. Таким образом, левая часть в $(*)_j$ сравнима с $(1 - x_{j-2}^2 3^{2(j-1)})$, т.е. с 1 по модулю 3^j. Наконец, чтобы оценить число двоичных операций, заметить, что при выполнении шага 3) каждый раз производится по паре умножений и делений целых чисел по $O(\alpha)$ бит, т.е. на каждом шаге производится $O(\alpha^2)$ двоичных операций. Таким образом, вся процедура требует $O(\alpha^3)$ двоичных операций.

3. а) Чтобы упростить вычисление $(g^b)^a$ в $\mathbf{F}_{31^2}$, воспользоваться тем, что $(c+di)^{32} = c^2 + d^2$. Ответ: $A + Bi = 26 + 28i$. б) $20 + 13i$. в) $P \equiv 6C + 18 \pmod{31}$. г) YOU'RE JOKING!

4. а) $K_E = 1951280$, наименьший неотрицательный вычет этого числа по модулю 26^4 равен $7 \cdot 26^3 + 0 \cdot 26^2 + 13 \cdot 26 + 6$; однако к нему надо прибавить единицу, чтобы получить обратимую матрицу шифрования $\left(\begin{smallmatrix} 7 & 0 \\ 13 & 7 \end{smallmatrix}\right)$; б) $\left(\begin{smallmatrix} 15 & 0 \\ 13 & 15 \end{smallmatrix}\right)$, DONOTPAY.

5. Преобразования f_A должны коммутировать, т.е. $f_A f_B = f_B f_A$ для всех пар пользователей A и B. Нужно использовать их вместе с хорошей схемой подписи (как указано в условии), и не должно существовать практической возможности определить ключ для f_A по паре $(P, f_A(P))$. Например, преобразования сдвига $f_A(P) = P + b$ или линейное $f_A(P) = aP$ удовлетворяют первому условию, но не удовлетворяют второму, так как по паре $(P, P+b)$ (или (P, aP)) легко находится b (или a). Пример в тексте удовлетворяет этим условиям, так как по нашим предположениям задача дискретного логарифмирования не может быть решена за разумное время.

6. $P = 6229 = \text{«GO!»}$.

7. а) Сначала заменить x на $p - 1 - x$ и свести задачу к эквивалентному сравнению $g^x a \equiv 1 \pmod{p}$. Положить $l = 2^k$ и $x = x_0 + 2x_1 + \cdots + 2^{l-1}x_{l-1}$. Пусть $g_j \equiv g^{2^j} \pmod{p}$ и $a_j \equiv g^{x_0 + 2x_1 + \cdots + 2^{j-1}x_{j-1}} a \pmod{p}$ (в качестве a_0 берется a). На j-м шаге вычислить $a_{j-1}^{2^{k-j}} = \pm 1$ и положить $x_{j-1} = 0$ при $+1$ и $x_{j-1} = 1$ при -1; вычислить также $g_j = g_{j-1}^2$ и $a_j = g_{j-1}^{x_{j-1}}$. При $j = l$ работа заканчивается. б) $O(\log^4 p)$. в) $k = 7912$.

8. THEYREFUSEOURTERMS.

9. Чтобы найти x, Алиса сводит сравнение $g^S \equiv y^r r^x \equiv g^{ar+kx}$ к сравнению $S \equiv ar + kx \pmod{p-1}$, которое имеет решение $x \equiv k^{-1}(S - ar) \pmod{p-1}$. Боб знает p, g и $y = y_A$ и может проверить, что $g^S \equiv y^r r^x \pmod{p}$, так как он получил пару (r, x) вместе с S. Наконец, любой, кто умеет решать задачу дискретного логарифмирования, может определить a по g и y и подделать подпись, найдя x.

10. 107.

11. а) $9/128 = 7,03\%$, $160/1023 = 15,64\%$. б) $70/2187 = 3,20\%$, $1805/29524 = 6,11\%$ (см. следствие из предложения II.1.8).

12. а) Число нормированных многочленов равно $(p^{n+1} - 1)/(p-1) \approx p^n$. Числом произведений степени ниже n можно пренебречь. Число n_f неприводимых нормированных многочленов степени f равно $\frac{1}{f}(p^f - \sum_{d<f, d|f} dn_d) \approx \frac{p^f}{f}$. Число произведений степени n равно тогда следующей сумме по всем разбиениям числа $n = \sum_{d=1}^m i_d d$ $(i_d \geqslant 0)$:

$$\sum \binom{n_1 + i_1 - 1}{i_1} \cdots \binom{n_m + i_m - 1}{i_m} \approx p^n \sum \frac{1}{2^{i_2} 3^{i_3} \cdots m^{i_m} i_1! i_2! \cdots i_m!}.$$

Таким образом,

$$P(n,m)=\sum\Bigl(\prod_{d=1}^{m} d^{i_d} i_d!\Bigr)^{-1}.$$

Это выражение, очевидно, положительно. Чтобы проверить оценку $P(n,m)<1$, заметим, что имеется приблизительно p^n/n нормированных неприводимых многочленов степени n, поэтому вероятность того, что нормированный многочлен не разлагается в произведение нужного вида, не меньше $1/n$. б) $\sum_{i+2j=n,\ 0\leqslant i,j}(2^j i!j!)^{-1}$. в) $P(3,2)=2/3$, $P(4,2)=5/12$, $P(5,2)=13/60$, $P(6,2)=19/180$, $P(7,2)=29/630$.

§IV.4

1. а) Да, 1. б) Да, 0. в) Нет, 2. г) Нет, 0. д) Да, 1. е) Нет, 1.

2. а) Использовать индукцию по k. б) Для доказательства второй части взять $v_i>1+v_{i-1}+\cdots+v_0$ и положить $V=v_i-1$.

3. Использовать индукцию.

4. а) INTERCEPTCONVOY. б) 89, 3, 25, 11, 41, 60, 65.

5. FORMULA STOLEN!

6. BRIBE HIM!

§IV.5

1. Вероятность успешного T-кратного обмана равна 2^{-T} и при $T\to\infty$ стремится к 0.

2. а) Числа e и $e+x$ по модулю N, которые Вивалес получает на шагах 2 и 3 процедуры, принадлежат диапазону от 0 до $N-1$. При большом числе испытаний Вивалес получит довольно хорошее представление о величине N. б) Положить N' очень большим целым, кратным N, и заменить N на N' в описании шагов 1 и 3.

3. Величины, получаемые Вивалесом на шаге 3, представляют собой верхние оценки для x. Величины, которые на шаге 3 посылает Клайд, не ограничены снизу, в отличие от величин $x+e$, посылаемых Пикарой.

4. Открытым ключом Пикары должно быть y; подписание документа состоит в том, что Пикара убеждает Вивалеса, что ей известно значение дискретного логарифма этого числа.

5. Знание разложения на простые множители позволяет извлекать квадратный корень, используя метод, приведенный в конце §II.2, вместе с китайской теоремой об остатках (см. также упражнение 5 к §IV.2). Обратно, допустим, что нам известен алгоритм извлечения квадратного корня. Тогда выберем случайное число x и применим этот алгоритм к наименьшему неотрицательному вычету x^2 по модулю n. В результате получим такое x', что $x'^2\equiv x^2 \pmod n$. С вероятностью 0,5 оказывается, что $x'\not\equiv\pm x \pmod n$. В этом случае сразу получаем нетривиальный делитель как НОД $(x'+x,n)$. Повторив процедуру T раз, с вероятностью $1-2^{-T}$ получим искомое разложение.

6. Да. Предположим, что некто Пикара$_2$, играющая роль Пикары, перехватила сообщение $b^{y_1}, b^{y_2}, \alpha_1, \alpha_2$, посланное Пикарой Вивалесу, и с помощью этого хочет обмануть Вивалеса, внушив ему, что она сама знает разложение числа n (или раскрашивание в три краски, или дискретный логарифм числа и т. п.). Представим себе теперь, что Вивалес не удовлетворяется получением от Пикары$_2$ точным повторением четверки, посланной Пикарой. Не зная ни секретных случайных чисел y_1, y_2 Пикары, ни ее сообщений m_1, m_2, ни дискретного логарифма хотя бы одного из чисел β_1 и β_2, Пикара$_2$ не сможет построить другую четверку, которая создала бы у Вивалеса впечатление, что она знает разложение.

7. Пикара выбирает случайно $0 \leqslant x' < N$ и посылает Вивалесу $y' = b^{x'}$. Два сообщения для скрытой передачи — это $m_1 = x'$ и $m_2 = x + x' \pmod N$. Вивалес проверяет выполнение условий либо $b^{x'} = y'$, либо $b^{x+x'} = yy'$. Если процедура повторяется T раз, то Пикаре может повезти лишь в одном из 2^T случаев (т. е. Пикаре удастся убедить Вивалеса в том, что она знает дискретный логарифм y, когда на самом деле не знает его).

8. Вивалес довольно легко может заставить Пикару выдать разложение числа n следующим образом. Он выбирает случайно целые числа z до тех пор, пока не найдет z, символ Якоби которого по модулю n равен -1. Тогда он посылает Пикаре $y = z^2 \pmod n$. Пикара отвечает величиной x^2 квадратного корня из $y \pmod n$, отличного от $\pm z$. Теперь Вивалес может найти нетривиальный делитель числа n, а именно, НОД $(x^2 + z, n)$.

9. Доказательство отсутствия разглашения при передаче, использующее введение «имитатора» Клайда, здесь не работает. Другая проблема состоит в том, что Пикара должна быть уверена, что каждое y_i получено из Центра, а не от Вивалеса, выдающего себя за Центр.

§ V.1

1. а) 4, 11. б) 8, 13. в) См. часть г). г) Показать, что $n-1 \equiv p-1 \pmod{2p-2}$; таким образом, $b^{n-1} \equiv 1 \pmod p$ и $b^{n-1} \equiv b^{(2p-1-1)/2} \equiv (\frac{b}{2p-1}) \pmod{2p-1}$. Тогда $b^{n-1} \equiv 1 \pmod{p(2p-1)}$ равносильно $(\frac{b}{2p-1}) = 1$.

2. а) Использовать сравнение $n = n'p = n'(p-1+1) \equiv n' \pmod{p-1}$. б) Использовать п. а) с $n' = 3$, чтобы заключить, что p должно быть делителем $2^2-1, 5^2-1, 7^2-1$. в) p должно быть делителем $2^4-1, 5^4-1, 7^4-1$. г) Согласно п. б) любое меньшее n должно быть произведением двух простых чисел, больших 5. Затем проверить 49 и 77.

3. Поделить сравнение (1) с $n = p^2$ на сравнение $b^{p^2-p} \equiv 1 \pmod{p^2}$, которое всегда выполняется согласно теореме Эйлера (предложение I.3.5).

4. а) 217. б) 341.

5. а) Предположим сначала, что n — псевдопростое по основанию b. Так как $n-1 = pq-1 \equiv q-1 \pmod{p-1}$, то $b^{q-1} \equiv 1 \pmod p$. Однако по малой теореме Ферма всегда $b^{p-1} \equiv 1 \pmod p$. Так как d представляется как целочисленная линейная комбинация $p-1$ и $q-1$, то $b^d \equiv 1 \pmod p$. Меняя ролями p и q, получаем, что $b^d \equiv 1 \pmod q$, откуда следует, что $b^d \equiv 1 \pmod n$. Обратное утверждение доказывается аналогично (фактически — легче). В $(\mathbf{Z}/n\mathbf{Z})^*$ имеется d^2 оснований. б) Четыре: $\pm 1, \pm(4p+1)$. в) $d^2/\varphi(341) = 100/300 = 1/3$.

7. а) См. п. б). б) Имеем $N-1 = b(b^{n-1}-1)/(b-1)$. Так как n есть псевдопростое по основанию b, числитель делится на n; знаменатель взаимно прост с n, поэтому $n|(N-1)$. Так как $b^n \equiv 1 \pmod N$ (а именно, $(b-1)N = b^n-1$), имеем $b^{N-1} \equiv 1 \pmod N$. Нужно еще убедиться, что N составное и нечетное. То, что N составное, следует из того, что составным по предположению является n (см. следствие предложения I.4.1). Разложение $N = b^{n-1}+b^{n-2}+\cdots+b+1$ показывает, что N нечетно (b при этом может быть как четным, так и нечетным). в) Начиная соответственно с 341, 91, 217, использовать п. б) для построения возрастающей последовательности псевдопростых. Заметить, что условие НОД $(b-1, n) = 1$ всегда выполнено при $b = 2, 3, 5$. г) Число 15 — псевдопростое по основанию 4, однако $N = (4^{15}-1)/3$ таковым не является. Чтобы убедиться в этом, заметьте, что 4 имеет в $(\mathbf{Z}/N\mathbf{Z})^*$ порядок 15, однако $N-1 = 4(4^{14}-1)/3$ не делится на 3, не говоря уже о 15.

8. а) $n = (\frac{b^p-1}{b-1})(\frac{b^p+1}{b+1})$. б) Заметьте, что n нечетно (см. указание к 7б)), следовательно, $2|(n-1)$. Далее, так как $(n-1)(b^2-1) = b^2(b^{2(p-1)}-1) \equiv 0 \pmod p$ и p не делит $(b+1)(b-1) = b^2-1$, то $p|(n-1)$. в) Так как n — нечетное составное число, $b^{2p} \equiv 1 \pmod n$ и $2p|(n-1)$, то n — псевдопростое по основанию b. Так как существует бесконечно много простых чисел, больших $b+1$, то можно получить бесконечно много псевдопростых чисел по основанию b.

9. а) $3^{2046} \equiv 1013 \pmod{2047}$, поэтому (1) не выполняется при $b = 3$. б) Каждое такое составное число будет псевдопростым по основанию 2. Действительно, пусть $n = 2^{2^k}+1$ есть число Ферма. Тогда $2^{2^k} \equiv -1 \pmod n$. Поэтому $2^{n-1} \equiv 1 \pmod n$ (в этом можно убедиться повторным возведением в квадрат). Для $n = 2^p - 1$ имеем $n-1 = 2(2^{p-1}-1) \equiv 0 \pmod p$, и поэтому из $2^p = n+1 \equiv 1 \pmod n$ следует, что $2^{n-1} \equiv 1 \pmod n$. Тест (2) при $b = 2$ также не будет работать, так как обе части будут равны 1, даже если n — составное. Тест (3) при $b = 2$ также не работает для числа Ферма, поскольку $2^{2^k} \equiv -1 \pmod n$; для числа Мерсенна это следует из предложения V.1.5.

10. Раскрывая круглые скобки, показать, что $n-1$ делится на $36m$ и, следовательно, на $6m, 12m, 18m$.

12. Метод для решения пп. а) и б) упражнения дается в п. в). а) $561 = 3\cdot 11\cdot 17$. б) $1105 = 5\cdot 13\cdot 17$, $2465 = 5\cdot 17\cdot 29$, $10585 = 5\cdot 29\cdot 73$. в) Пусть $p < q$. Так как $(q-1)|(rpq-1) \equiv rp-1 \pmod{q-1}$, должно быть $rp-1 = a(q-1)$, $1 < a < r$. Аналогично, $(p-1)|(rq-1)$, и потому $(p-1)|a(rq-1) = r(aq)-a = r(a+rp-1)-a \equiv (r-1)(a+r) \pmod{p-1}$. Таким образом, при фиксированном r и любом фиксированном a, $2 \leqslant a \leqslant r-1$, для простого p существует лишь конечное число возможностей, а именно, $p-1$ должно быть делителем $(r-1)(a+r)$. Если p выбрано, то q однозначно определяется условием $rp-1 = a(q-1)$. Разумеется, не каждая пара a и p дает число Кармайкла, так как, например, a может не быть делителем $rp-1$.

13. Любое число Кармайкла, не перечисленное в упражнениях 12а) и 12б), должно быть произведением не менее чем трех различных простых чисел, каждое из которых не меньше 7.

14. $n = 21$, $b = 8$.

16. а) В упражнении 1г) нужно проверять только те b, для которых $b^{p-1} \equiv (\frac{b}{2p-1}) \equiv 1 \pmod{2p-1}$. Так как $n-1 \equiv p-1 \pmod{2p-2}$, то $b^{(n-1)/2} \equiv b^{(p-1)/2} \pmod p$ и $\pmod{2p-1}$; следовательно, $b^{(n-1)/2} \equiv b^{(p-1)/2} \pmod n$. Далее, $(\frac{b}{n}) = (\frac{b}{2p-1})(\frac{b}{p}) = (\frac{b}{p}) \equiv b^{(p-1)/2} \pmod p$, следовательно, условие (2) эквивалентно $b^{(p-1)/2} \equiv (\frac{b}{p}) \pmod{2p-1}$. Оно выполняется в точности для половины всех b, для которых $b^{p-1} \equiv 1 \pmod{2p-1}$ (так как в $(\mathbf{Z}/(2p-1)\mathbf{Z})^*$ такой элемент b должен быть такой степенью g^j образующего элемента g, что $\frac{p-1}{2}j \equiv 0 \pmod 4$, если $(\frac{b}{p}) = 1$, $\frac{p-1}{2}j \equiv 2 \pmod 4$, если $(\frac{b}{p}) = -1$). б) $n = p(2p-1)$, где $p \equiv 3 \pmod 4$ (см. предложение V.1.5).

17. Найти n по модулю $72m$: $n \equiv 36m^2+36m+1$. Таким образом, $\frac{n-1}{2} \equiv 18m(m+1) \pmod{36m}$. Если m нечетно, отсюда следует, что всегда $b^{(n-1)/2} \equiv 1 \pmod n$ (так как $(p-1)|36m$ для каждого $p|n$). Поэтому (2) выполняется тогда и только тогда, когда $(\frac{b}{n}) = 1$, т.е. в 50% случаев. Если m четно, то, по-прежнему, $b^{(n-1)/2} \equiv 1 \pmod{6m+1}$ и $\pmod{18m+1}$, в то время как $b^{(n-1)/2} \equiv b^{6m} \equiv (\frac{b}{12m+1}) \pmod{12m+1}$. В этом случае (2) выполнено тогда и только тогда, когда $(\frac{b}{12m+1}) = 1$ (так что $b^{(n-1)/2} \equiv 1 \pmod n$), а также $(\frac{b}{n}) = 1$, т.е. в 25% случаев.

18. а) $O(\log^3 n \log m)$. б) $O(\log^5 n)$.

19. а) N — составное число, поскольку таковым является n (по следствию из предложения I.4.1); далее методом упражнения 9 убедиться, что $2^{(N-1)/2} = 2^{2^{n-1}-1} \equiv 1 \pmod N$. Но так как $N \equiv -1 \pmod 8$, то $(\frac{2}{N}) = 1$. Таким образом, N есть эйлерово псевдопростое; по предложению V.1.5 оно также и сильно псевдопростое. б) Использовать тот же подход, что и в упражнении 7 в).

20. Если в (3) реализуется первая из возможностей, то, очевидно, $(b^k)^t \equiv 1 \pmod n$. Теперь предположим, что $b^{2^r t} \equiv -1 \pmod n$. Положить $k = 2^i j$, где j нечетно. Если $i > r$, то $(b^k)^t \equiv 1 \pmod n$; если $i \leqslant r$, то $(b^k)^{2^{r-i}t} = (b^{2^r t})^j \equiv (-1)^j = -1 \pmod n$.

21. а) Показать, что необходимые и достаточные условия на b — это $(\frac{b}{17}) = 1$, $(\frac{b}{561}) = 1$. Оба эти условия выполняются в 25% случаев, т. е. для 80 оснований в $(\mathbf{Z}/561\mathbf{Z})^*$. б) Так как $b^{70} \equiv 1 \pmod 3$ и $\pmod{11}$, то 561 — сильно псевдопростое по основанию b в том и только том случае, когда $b^{35} \equiv \pm 1 \pmod{561}$, т. е. тогда и только тогда, когда либо 1) $b \equiv 1 \pmod 3$, $b \equiv 1 \pmod{17}$, $(\frac{b}{11}) = 1$, либо 2) $b \equiv -1 \pmod 3$, $b \equiv -1 \pmod{17}$, $(\frac{b}{11}) = -1$. Согласно китайской теореме об остатках существует 10 таких оснований (по 5 в случаях 1) и 2)). Помимо тривиальных ± 1, это $b = 50, 101, 103, 256, 305, 458, 460, 511$.

22. Использовать упражнение 7 а) к §I.3, согласно которому квадратными корнями из единицы являются лишь ± 1.

23. а) $8^2 \equiv 18^2 \equiv -1 \pmod{65}$; $14^2 \equiv 1 \pmod{65}$, но $14^1 \not\equiv \pm 1 \pmod{65}$. б) Случай, когда n — степень простого числа, рассматривается как следствие из предыдущего упражнения. Предположим поэтому, что n не есть степень простого числа. Во-первых, если $p|n$, $p \equiv 3 \pmod 4$, то никакое целое число в четной степени не сравнимо с -1 по модулю n, так как -1 не есть квадратичный вычет по модулю p. Следовательно, в данном случае условие сильной псевдопростоты представимо в виде $b^t \equiv \pm 1 \pmod n$. Это условие обладает, очевидно, свойством мультипликативности. Далее, предположим, что $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, где $p_j \equiv 1 \pmod 4$, $j = 1, 2, \ldots, r$. Пусть $\pm a_j$ — два квадратных корня из -1 по модулю $p_j^{\alpha_j}$ (квадратный корень по модулю p_j можно преобразовать в квадратный корень по модулю $p_j^{\alpha_j}$, см. упражнение 20 к §II.2). Тогда любое $b \equiv \pm a_j \pmod{p_j^{\alpha_j}}$ является основанием, по которому n — сильно псевдопростое, так как тогда $b^{2t} \equiv (-1)^t \equiv -1 \pmod n$. Выбрать b_1 и b_2 так, чтобы $b_1 = a_j \pmod{p_j^{\alpha_j}}$, $b_2 \equiv (-1)^{\varepsilon_j} a_j \pmod{p_j^{\alpha_j}}$, $j = 1, \ldots, r$, где $(\varepsilon_1, \ldots, \varepsilon_r)$ — любой набор из 0 и 1, отличный от $(0, \ldots, 0)$ и $(1, \ldots, 1)$. Далее показать, что если $b = b_1 b_2$, то $b^{2t} \equiv 1 \pmod n$, а $b^t \equiv b \not\equiv \pm 1 \pmod n$.

24. а) В этом случае получается число c, не равное ± 1 и квадрат которого равен 1; тогда НОД $(c+1, n)$ — нетривиальный делитель n. б) Выбрать p и q так, чтобы $p-1$ и $q-1$ не имели большого общего делителя (см. выше упражнение 5).

§V.2

1. НОД $(x_5 - x_3, n) =$ НОД $(21 - 63, 91) = 7$; $91 = 7 \cdot 13$.
2. НОД $(x_6 - x_3, n) =$ НОД $(2839 - 26, 8051) = 97$; $8051 = 83 \cdot 97$.
3. НОД $(x_9 - x_7, n) =$ НОД $(869 - 3397, 7031) = 79$; $7031 = 79 \cdot 89$.
4. НОД $(x_6 - x_3, n) =$ НОД $(630 - 112, 2701) = 37$; $2701 = 37 \cdot 73$.
5. Доказать индукцией по k, что для $1 \leqslant k \leqslant r$ вероятность того, что $x_0, x_1, \ldots, x_{k-1}$ все различны, а x_k равен одному из предыдущих x_j, равна $1/r$. При $k = 1$ вероятность того, что $f(x_0) = x_0$, равна $1/r$. Шаг индукции проводим следующим образом. По предположению индукции вероятность того, что k не меньше минимального значения t, удовлетворяющего условию $x_t = x_j$ при не-

некотором $j < k$, равна $1 - \frac{k-1}{r} = \frac{r-(k-1)}{r}$. Если это событие происходит, то для $f(x_{k-1})$ существует $r-k+1$ возможных значений, так как f — биекция и потому $f(x_{k-1})$ не может принимать ни одно из значений $f(x_j)$, $0 \leqslant j \leqslant k-2$. Из $r-(k-1)$ возможных значений одно — это x_0, а другие не равны $x_0, x_1, \ldots, x_{k-1}$. Таким образом, лишь в одном из $r-k+1$ случаев значение x_k совпадает с более ранним значением (а именно, когда $x_k = f(x_{k-1}) = x_0$). Вероятность того, что одновременно произойдут *два* события, — 1) ни один из $x_1, \ldots, x_{k-1}$ не совпадает со своими предшественниками и 2) $x_k = x_0$ — равна произведению вероятностей этих событий, т.е. $\frac{r-(k-1)}{r} \cdot \frac{1}{r-(k-1)} = \frac{1}{r}$. б) Так как все значения от 1 до r равновероятны, среднее равно $\frac{1}{r}\sum_{k=1}^{r} k = \frac{1}{r} r(r+1)/2 = \frac{r+1}{2}$.

6. Предположим, что a и n взаимно просты (если бы это было не так, мы сразу нашли бы делитель n, вычислив НОД (a, n), и могли бы не использовать ро-метод). Тогда отображение $f\colon x \to ax + b$ есть биекция $\mathbf{Z}/r\mathbf{Z}$ на себя для любого $r|n$. Поэтому математическое ожидание числа шагов до первого повторения какого-либо значения по модулю r — величина порядка $r/2$ (по упражнению 5 б)), а не $\sqrt{r}$, т.е. значительно больше.

7. а) $2^k \equiv 2^l \pmod{r-1}$. б) $l = s$ и $k = s + m$, где m — *порядок* 2 по модулю t, т.е. такое наименьшее натуральное число, что $2^m \equiv 1 \pmod{t}$. Число m есть также *период* двоичного разложения $1/t$; в этом можно убедиться, перейдя от равенства $2^m - 1 = ut$ к $1/t = u\sum_{i=1}^{\infty} 2^{-mi}$. в) Нередки случаи, когда k имеет порядок почти такой же величины, что и r; например, если $r-1$ есть удвоенное простое число, а 2 — образующий по модулю этого простого числа (тогда $s = 1$, $m = (r-3)/2$).

§ V.3

1. а) (Использовать $t = [\sqrt{n}] + 1 = 93$) $89 \cdot 97$. б) (Использовать $t = [\sqrt{n}] + 4 = 903$) $823 \cdot 983$. в) (Использовать $t = [\sqrt{n}] + 6 = 9613$) $9277 \cdot 9949$. г) (Использовать $t = [\sqrt{n}] + 1 = 9390$) $9343 \cdot 9437$. е) (Использовать $t = [\sqrt{n}] + 8 = 75$) $43 \cdot 107$.

2. Пусть $n = ab$ и $a > b$. Если $a < \sqrt{n} + \sqrt[4]{n}$, то $b = n/a > n/(\sqrt{n} + \sqrt[4]{n}) > \sqrt{n} - \sqrt[4]{n}$. С другой стороны, если мы начинаем с $b > \sqrt{n} - \sqrt[4]{n}$, то должно быть $a < \sqrt{n} + \sqrt[4]{n} + 2$, иначе было бы $n = ab > (\sqrt{n} + \sqrt[4]{n} + 2)(\sqrt{n} - \sqrt[4]{n}) = n + \sqrt{n} - 2\sqrt[4]{n} > n$ (если $n > 15$; для меньших значений n упражнение 2 проверяется отдельно). Таким образом, в обоих случаях $a - b < 2(\sqrt[4]{n} + 1)$. Но если факторизация Ферма «не срабатывает» для первого значения t, то s и t, соответствующие разложению $n = ab$, удовлетворяют условию $t > \sqrt{n} + 1$ и, следовательно, $s = \sqrt{t^2 - n} > \sqrt{(\sqrt{n}+1)^2 - n} = \sqrt{2\sqrt{n} + 1} > \sqrt{2}\sqrt[4]{n}$, что при $n > 33$ противоречит неравенству $s = \frac{a-b}{2} < \sqrt[4]{n} + 1$.

3. а) Мы имели бы $t^2 - s^2 = kn \equiv 2 \pmod 4$, но по модулю 4 разность квадратов не может быть сравнима с 2. б) Мы имели бы $t^2 - s^2 = 4n \equiv 4 \pmod 8$. Это может быть, лишь когда s и t четны. Но тогда $(t/2)^2 - n = (s/2)^2$ и простая факторизация Ферма «сработала» бы так же хорошо.

4. а) (Использовать $t = [\sqrt{3n}] + 1 = 455$) $149 \cdot 463$. б) (Использовать $t = [\sqrt{3n}] + 2 = 9472$) $3217 \cdot 9293$. в) (Использовать $t = [\sqrt{5n}] + 1 = 9894$) $1973 \cdot 9923$. г) (Использовать $t = [\sqrt{5n}] + 2 = 9226$) $1877 \cdot 9067$.

5. $B = \{2, 3\}$; векторы $(0, 1)$ и $(0, 1)$; $b = 52 \cdot 53 \pmod n = 55$, $c = 2 \cdot 3^2 = 18$; НОД $(55 + 18, 2701) = 73$; $2701 = 37 \cdot 73$.

6. $B = \{-1, 2, 3, 61\}$; векторы $(1, 0, 0, 0)$, $(1, 0, 0, 1)$ и $(0, 0, 0, 1)$; $b = 68 \cdot 152 \cdot 153 \pmod n = 1555$; $c = 2 \cdot 3 \cdot 61 = 366$; НОД $(1555 + 366, 4633) = 113$; $4633 = 41 \cdot 113$.

7. а) Оценить разность, взяв сумму площадей «треугольных» областей между графиком $\log x$ и прямоугольниками римановой суммы. б) Сравнить $\int_1^n \log x\, dx$ с

суммой площадей трапеций, верхние стороны которых соединяют точки $(j, \log j)$, и показать, что площадь между кривой и трапециями ограничена константой. в) $\lim_{y\to\infty}(\frac{1}{y}\log(y!) - (\log y - 1)) = 0$; итак, ответ: $\log y - 1$.

8. а) $(1-2^{-n})(1-2^{-n+1})\cdots(1-2^{-n+k-1})$. б) 0,298.

9. Выражение в оценке сложности ро-метода становится в $3{,}2\cdot 10^{12}$ раз больше, в оценке сложности метода факторных раз в $2{,}6\cdot 10^6$ раз больше.

10. а) Если $s < s_0$, то $h(s) \geqslant f(s) > f(s_0) = \frac{1}{2}h(s_0)$, а если $s > s_0$, то $h(s) \geqslant g(s) > g(s_0) = \frac{1}{2}h(s_0)$. б) Применить часть а) к $\log f(s)$ и $\log g(s)$.

§ V.4

1. а) $\frac{1}{1+}\frac{1}{1+}\frac{1}{44}$. б) $\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+}\frac{1}{1+1}$. в) $1+\frac{1}{7+}\frac{1}{1+}\frac{1}{2+}\frac{1}{4}$.

2. а) Так как $x = a + \frac{1}{x}$, то число x должно быть положительным корнем уравнения $x^2 - ax - 1 = 0$, т. е. $x = (a+\sqrt{a^2+4})/2$. б) Так как все числа a_i равны 1, то рекуррентное соотношение для числителей и знаменателей приближений то же самое, что и для чисел Фибоначчи.

3. $2+\frac{1}{1+}\frac{1}{2+}\frac{1}{1+}\frac{1}{1+}\frac{1}{4+}\frac{1}{1+}\frac{1}{1+}\frac{1}{6}\cdots$; можно показать, что a_i при $i \equiv 2 \pmod 3$ — последовательные четные числа и $a_i = 1$ для всех остальных i.

4. Для каждого b_i разность $b_i^2 - c_i^2 n$ есть наименьший абсолютный вычет b_i^2 по модулю n. Если p делит его, то $b_i^2 \equiv c_i^2 n \pmod p$. Но отсюда следует, что n — квадратичный вычет по модулю p.

5. Следующие ниже таблицы составлены для стольких первых значений i, что наименьшие абсолютные вычеты $b_0^2, \ldots, b_i^2$ обеспечивают факторизацию n. В четырех случаях (пп. ж), з), к), л)) существуют меньшие значения i, при которых можно найти такие подмножества вычетов, что сумма соответствующих им векторов $\vec{\varepsilon_i}$ равна нулю; однако в этих случаях мы приходим к сравнениям $b \equiv \pm c \pmod n$.

а)

i	0	1	2	3
a_i	97	1	1	17
b_i	97	98	195	3413
$b_i^2 \pmod n$	-100	95	-11	44

$B = \{-1, 2, 5, 11\}, b = 97\cdot 195\cdot 3413, c = 2^2\cdot 5\cdot 11$, НОД $(b+c, n) = 257$.

б)

i	0	1	2	3
a_i	116	2	4	1
b_i	116	233	1048	1281
$b_i^2 \pmod n$	-105	45	-137	80

$B = \{2, 3, 5\}, b = 233\cdot 1281, c = 2^2\cdot 3\cdot 5$, НОД $(b+c, n) = 191$.

в)

i	0	1	2
a_i	93	1	2
b_i	93	94	281
$b_i^2 \pmod n$	-128	59	-32

$B = \{-1, 2\}, b = 93\cdot 281, c = 2^6$, НОД $(b+c, n) = 67$.

г)

i	0	1	2
a_i	120	8	3
b_i	120	961	3003
$b_i^2 \pmod n$	-29	65	-116

$B = \{-1, 2, 29\}$, $b = 120 \cdot 3003$, $c = 2 \cdot 29$, НОД $(b + c, n) = 307$.

д)

i	0	1	2	3	4	5	6
a_i	111	2	1	2	2	7	1
b_i	111	223	334	891	2116	3300	5416
$b_i^2 \pmod n$	-82	117	-71	89	-27	116	-39

$B = \{-1, 3, 13\}$, $b = 233 \cdot 2116 \cdot 5416$, $c = 3^3 \cdot 13$, НОД $(b + c, n) = 157$.

е)

i	0	1	2	3	4	5
a_i	120	1	1	8	2	2
b_i	120	121	241	2049	4339	10727
$b_i^2 \pmod n$	-127	114	-27	98	-71	162

$B = \{-1, 2, 3, 7\}$, $b = 2049 \cdot 10727$, $c = 2 \cdot 3^2 \cdot 7$, НОД $(b + c, n) = 199$.

ж)

i	0	1	2	3	4	5
a_i	100	1	1	1	1	2
b_i	100	101	201	302	503	1308
$b_i^2 \pmod n$	-123	78	-91	97	-66	77

$B = \{-1, 2, 3, 7, 11, 13\}$, $b = 101 \cdot 201 \cdot 503 \cdot 1308$,

$c = 2 \cdot 3 \cdot 7 \cdot 11 \cdot 13$, НОД $(b + c, n) = 191$.

з)

i	0	1	2	3	4	5	6	7	8	9
a_i	111	1	1	2	1	4	1	6	2	1
b_i	111	112	223	558	781	3682	4463	5562	3138	8700
$b_i^2 \pmod n$	-128	95	-67	139	-40	163	-31	79	-115	80

$B = \{-1, 2, 5\}$, $b = 111 \cdot 781 \cdot 8700$, $c = 2^7 \cdot 5$, НОД $(b + c, n) = 59$.

и)

i	0	1	2	3	4	5	6	7	8
a_i	96	1	2	2	5	1	1	1	1
b_i	96	97	290	677	3675	4352	8027	3026	1700
$b_i^2 \pmod n$	-137	56	-77	32	-107	79	-88	89	-77

$B = \{-1, 2, 7, 11\}$, $b = 290 \cdot 1700$, $c = 7 \cdot 11$, НОД $(b + c, n) = 47$.

к)

i	0	1	2	3	4	5	6	7	8	9
a_i	159	1	2	1	1	2	4	1	5	1
b_i	159	160	479	639	1118	2875	12618	15493	13550	3532
$b_i^2 \pmod n$	-230	89	-158	145	-115	61	-227	50	-167	145

$$B = \{-1, 2, 5, 23, 29\}, b = 639 \cdot 3532, c = 5 \cdot 29, \text{НОД } (b + c, n) = 97.$$

л)

i	0	1	2	3	4	5	6	7	8
a_i	133	1	2	4	2	3	1	2	1
b_i	133	134	401	1738	3877	13369	17246	12115	11488
$b_i^2 \pmod n$	-184	83	-56	107	-64	161	-77	149	-88

$$B = \{-1, 2, 7, 11, 23\}, b = 401 \cdot 3877 \cdot 17246 \cdot 11488, c = 2^6 \cdot 7 \cdot 11, \text{НОД } (b + c, n) = 61.$$

§ V.5

2. Наиболее трудоемкий — шаг 6). Время работы ограничивается величиной

$$O\Big(\sum_{\text{простые } p \leqslant P} \frac{A}{p} \log p \log n\Big) = O(A \log n \log P \log\log P).$$

(Вопрос касался только шагов 1)–7); другим трудоемким этапом для очень больших n является нахождение линейно зависимых строк по модулю 2 в матрице показателей, соответствующих B-числам среди чисел $t^2 - n$.)

3. а)

t	$t^2 - n$	2	13	17	19	29	37	41	47
1030	14297	–	–	1	–	2	–	–	–
1319	693158	1	–	1	1	1	1	–	–
1370	830297	–	2	3	–	–	–	–	–
1493	1182446	1	–	–	1	2	1	–	–

Строки 1 и 3 зависимы и приводят к разложению $1879 \cdot 557$.

б)

t	t^2-n	2	3	5	13	17	19	23	31	37	41
1030	1209	–	1	–	1	–	–	–	1	–	–
1043	28158	1	1	–	1	–	2	–	–	–	–
1046	34425	–	4	2	–	1	–	–	–	–	–
1047	36518	1	–	–	–	–	1	–	2	–	–
1079	104550	1	1	2	–	1	–	–	–	–	1
1096	141525	–	2	2	–	1	–	–	–	1	–
1123	201438	1	2	–	–	–	2	–	1	–	–
1141	242190	1	4	1	1	–	–	1	–	–	–
1154	272025	–	3	2	1	–	–	–	1	–	–
1161	288230	1	–	1	–	–	1	–	–	1	1
1199	377910	1	2	1	1	1	1	–	–	–	–
1233	460598	1	–	–	–	1	1	1	1	–	–
1251	505310	1	–	1	3	–	–	1	–	–	–
1271	555750	1	2	3	1	–	1	–	–	–	–
1284	588965	–	–	1	2	1	–	–	–	–	1
1309	653790	1	1	1	–	–	1	–	1	1	–
1325	695934	1	2	–	–	–	–	1	–	–	2
1366	806265	–	2	1	–	–	1	1	–	–	1
1371	819950	1	–	2	–	–	–	2	1	–	–
1420	956709	–	2	–	2	1	–	–	–	1	–
1504	1202325	–	1	2	–	1	–	1	–	–	1

Строки 1, 2 и 7 зависимы по модулю 2, но не приводят к нетривиальному делителю. Строки 1 и 9 зависимы и приводят к разложению $1787 \cdot 593$.

в)

t	t^2-n	2	5	7	11	17	19	37	43	47
1001	3230	1	1	–	–	1	1	–	–	–
1003	7238	1	–	1	1	–	–	–	–	1
1004	9245	–	1	–	–	–	–	–	2	–
1018	37553	–	–	–	–	1	–	–	–	2
1039	80750	1	3	–	–	1	1	–	–	–
1056	116365	–	1	–	–	1	–	2	–	–
1069	143990	1	1	1	2	1	–	–	–	–
1086	180625	–	4	–	–	2	–	–	–	–
1090	189329	–	–	1	–	1	–	1	1	–
1146	314545	–	1	1	1	–	1	–	1	–
1164	356125	–	3	1	1	–	–	1	–	–
1191	419710	1	1	–	–	–	1	–	–	2
1241	541310	1	1	1	1	–	1	1	–	–
1311	719950	1	2	1	2	1	–	–	–	–
1426	1034705	–	1	1	–	1	–	1	–	1

Строки 1 и 5 зависимы и приводят к разложению $661 \cdot 1511$.

§ VI.1

1. Либо группа окружности (если вещественная кривая имеет одну связную компоненту), либо произведение группы окружности и двуэлементной группы (если она имеет две связные компоненты). Пример первого случая — кривая с уравнением $y^2 = x^3 + x$, пример второго — кривая с уравнением $y^2 = x^3 - x$ (для уравнения вида (1) это определяется тем, имеет ли кубический многочлен в правой части 1 или 3 вещественных корня).

2. n^2 комплексных точек порядка n; n вещественных точек порядка n, если n нечетно, и либо n, либо $2n$, если n четно, в зависимости от того, имеет вещественная кривая одну или две компоненты.

3. Те же самые примеры, что и в упражнении 1.

4. а) На оси x. б) Точка перегиба. в) Точка, где прямая линия, проходящая через точку пересечения кривой с осью x, касается кривой (в дополнение к точкам из п. а)).

5. а) 3. б) 4. в) 7. г) 5.

6. Характеристика 2: $x_3 = \frac{y_1^2+y_2^2}{x_1^2+x_2^2} + x_1 + x_2$, $y_3 = c + y_1 + \frac{y_1+y_2}{x_1+x_2}(x_1 + x_3)$, а если $P = Q$, то $x_3 = \frac{x_1^4+a^2}{c^2}$, $y_3 = c + y_1 + \frac{x_1^2+a}{c}(x_1 + x_3)$; для уравнения (2б) $x_3 = \frac{y_1^2+y_2^2}{x_1^2+x_2^2} + \frac{y_1+y_2}{x_1+x_2} + x_1 + x_2 + a$, $y_3 = \frac{y_1+y_2}{x_1+x_2}(x_1 + x_3) + x_3 + y_1$, а если $P = Q$, то $x_3 = x_1^2 + \frac{b}{x_1^2}$, $y_3 = x_1^2 + (x_1 + \frac{y_1}{x_1})x_3 + x_3$; характеристика 3: $x_3 = (\frac{y_2-y_1}{x_2-x_1})^2 - a - x_1 - x_2$, $y_3 = -y_1 + \frac{y_2-y_1}{x_2-x_1}(x_1 - x_3)$, а если $P = Q$, то $x_3 = (\frac{ax_1-b}{y_1})^2 - a + x_1$, $y_3 = -y_1 + \frac{ax_1-b}{y_1}(x_1 - x_3)$.

7. а) Показать, что для каждой пары $\{a, -a\}$ в точности одно из значений $x = \pm a$ приводит к двум решениям (x, y) уравнения (значение $x = 0$ и точку в бесконечности рассмотреть отдельно). б)–в) Использовать тот факт, что при $q \equiv 2 \pmod 3$ отображение $x \to x^3$ поля $\mathbf{F}_q$ на себя взаимно однозначно.

8. В следующей таблице приводятся типы абелевых групп для каждого значения q и каждой из двух эллиптических кривых:

q	3	5	7	9	11	13
$y^2 = x^3 - x$	$(2,2)$	$(4,2)$	$(4,2)$	$(4,4)$	$(2,2,3)$	$(4,2)$
$y^2 = x^3 - 1$	$--$	$(2,3)$	$(2,2)$	$--$	$(4,3)$	$(2,2,3)$

q	17	19	23	25	27
$y^2 = x^3 - x$	$(4,4)$	$(2,2,5)$	$(4,2,3)$	$(8,4)$	$(2,2,7)$
$y^2 = x^3 - 1$	$(2,9)$	$(2,2,7)$	$(8,3)$	$(2,2,3,3)$	$--$

9. а) Пусть $P = (x, y)$. Тогда $-P = (x, y + 1)$, $2P = (x^4, y^4 + 1)$. б) Имеем $2(2P) = (x^{16}, y^{16} + 1 + 1) = (x^{16}, y^{16}) = (x, y) = P$. в) По п. б) имеем: $2P = -P$, т. е. $(x^4, y^4 + 1) = (x, y + 1)$; это, однако, означает, что $x^4 = x$, $y^4 = y$, а потому x и y принадлежат полю из 4 элементов. По теореме Хассе число N точек кривой отличается от 4+1 не более, чем на $2\sqrt{4} = 4$, и от 16+1 не более, чем на $2\sqrt{16} = 8$. Поэтому $N = 9$.

10. Знаменатель дзета-функции равен $(1 - T)(1 - pT)$ во всех случаях; в следующей таблице указаны числители для $p = 5, 7, 11, 13$:

$y^2 = x^3 - x$	$1 + 2T + 5T^2$	$1 + 7T^2$	$1 + 11T^2$	$1 - 6T + 13T^2$
$y^2 = x^3 - 1$	$1 + 5T^2$	$1 - 4T + 7T^2$	$1 + 11T^2$	$1 - 2T + 13T^2$

11. В обоих случаях уравнение не имеет решений (x, y) в $\mathbf{F}_p$. Поэтому единственной $\mathbf{F}_p$-точкой является точка в бесконечности. Числители дзета-функций равны, соответственно, $1 - 2T + 2T^2$ и $1 - 3T + 3T^2$. Тогда $N_r = \mathbf{N}((1+i)^r - 1)$ и, соответственно, $\mathbf{N}((1+\omega)^r - 1)$, где $\omega = (-1 + i\sqrt{3})/2$.

§ VI.2

1. Выбирать случайно элементы из $\mathbf{F}_q$, пока не попадется такой g, что $g^{(q-1)/2} = -1$ (а не $+1$).

2. Пусть $x \in \mathbf{F}_q$ соответствует m. а) Положить $f(x) = x^3 - x$. Заметить, что в точности одно из значений $f(x), f(-x) = -f(x)$ — квадрат. Положим $y = f(x)^{(q+1)/4}$. Показать затем, что либо (x, y), либо $(-x, y)$ — точка на кривой. б) Выбрать любой y и положить $x = (y^2 + y)^{(2-q)/3}$ (если $y = 0$ или -1, то положить $x = 0$); показать, что (x, y) лежит на кривой.

3. а) Последовательность точек (x, y) — следующая: $(562, 576), (581, 395), (484, 214), (501, 220), (1, 0), (1, 0), (144, 565)$. б) ICANT (I can't).

4. а) Группа $E \pmod p$ содержит нециклическую группу порядка 4 с тремя точками порядка 2. б) $E \pmod p$ содержит точку порядка 2, следовательно, ее порядок четен.

5. Использовать формулы из примера 5 в § 1. а) Использовать сравнение по модулю 3, чтобы показать, что в обоих случаях (r четно и r нечетно) имеем $3|N_r$. б) Если $4|r$, то $N_r = (2^{r/2} - 1)^2 = (2^{r/4} - 1)^2(2^{r/4} + 1)^2$; N_r делится на $(r/4)$-битовое простое число тогда и только тогда, когда $r/4$ — простое число, для которого $2^{r/4} - 1$ есть простое число Мерсенна; N_r делится на $(r/4 + 1)$-битовое простое число тогда и только тогда, когда $r/4 = 2^k$ и $2^{2^k} + 1$ — простое число Ферма.

6. а) $\mathbf{F}_p$-точки образуют собственную подгруппу $\mathbf{F}_{p^r}$-точек (по теореме Хассе), и эта подгруппа имеет более одного элемента (также по теореме Хассе). Таким образом, N_r имеет собственный делитель. б) В обоих случаях пусть E задано уравнением $y^2 + y = x^3 - x + 1$; легко проверить, что ни над $\mathbf{F}_2$, ни над $\mathbf{F}_3$ у кривой нет точек, кроме точки в бесконечности. Таким образом, рассуждения, проведенные в п. а), не применимы. Можно показать, что если $p = 2$, то $N_2 = 5, N_3 = 13, N_5 = 41, N_7 = 113, N_{11} = 2113$ (заметим, что дзета-функция равна $(1 - 2T + 2T^2)/((1 - T)(1 - 2T))$; для r простого N_r — простое число тогда и только тогда, когда так называемое «комплексное число Мерсенна» $(1 + i)^r - 1$ является простым гауссовым целым числом, или, что эквивалентно, тогда и только тогда, когда $2^r + 1 - (\frac{2}{r})2^{(r+1)/2}$ — простое число, где $(\frac{2}{r})$ — символ Лежандра); если $p = 3$, то $N_2 = 7, N_5 = 271, N_7 = 2269$ (здесь дзета-функция равна $(1 - 3T + 3T^2)/((1 - T)(1 - 3T))$).

7. а) $y^2 + y = x^3 + \alpha$, где α — любой элемент $\mathbf{F}_4$, не принадлежащий $\mathbf{F}_2$. б) Дзета-функция равна $(1 - 4T + 4T^2)/((1 - T)(1 - 4T))$ и обратные элементы к обоим корням числителя равны 2; далее использовать замечание в конце § 1. в) Удвоенная точка (x, y) — это (x^4, y^4) (возведение в 4-ю степень есть «отображение Фробениуса», т. е. образующий группы Галуа $\mathbf{F}_{4^r}$ над $\mathbf{F}_4$). г) «r-кратное» удвоение любой точки — это $(x^{4^r}, y^{4^r}) = (x, y)$, т. е. любая $P \in E$ удовлетворяет соотношению $2^r P = P$.

8. а) Воспользоваться тем, что x принадлежит $\mathbf{F}_2$ в том и только том случае, когда $x^2 = x$, а также тем, что $(a + b)^2 = a^2 + b^2$ в поле характеристики 2. б) Отображение $z \to 1 + z$ задает взаимно однозначное соответствие элементов с $\operatorname{tr} z = 0$ и элементов с $\operatorname{tr} z = 1$. в) Выбрать случайным образом $x \in \mathbf{F}_{2^r}$, затем подставить $z = x^3 + ax + b$ в $g(z)$; если z оказывается среди тех 50% элементов,

для которых $\operatorname{tr} z = 0$, то получаем точку $(x, g(z))$ на кривой.

9. При работе с E по модулю p пользуются теми же формулами (4)–(5) из § 1. Точка в бесконечности получается тогда, когда складываются точки меньшей кратности $kP = k_1 P + k_2 P$, где $k_1 P$ и $k_2 P$ — точки, которые после приведения по модулю p имеют одинаковые x-координаты и противоположные по знаку y-координаты. Это эквивалентно условиям 1)–2) в тексте упражнения.

10. Знаменатель точки $8P$ делится на 23, поэтому $P \pmod{23}$ согласно упражнению 9 имеет порядок 8 на $E \pmod{23}$. Однако из теоремы Хассе вытекает, что $E \pmod{23}$ имеет более восьми точек.

11. (676,182), (385,703), (595,454), (212,625), (261,87), (77,369), (126,100), (66,589), (551,606), (501,530), (97,91), (733,110), (63,313), (380,530).

§ VI.3

1. а) $1 - 1/q$. б) $1 - 1/q$.

3. а) Если $n = 2^{2^k} + 1$ — простое число, то любое a с $\left(\frac{a}{n}\right) = -1$ обладает указанным свойством. Относительно $a = 3, 5, 7$ см. упражнение 15 к § II.2. С другой стороны, если p — собственный простой делитель n и если $a^{2^{2^k-1}} \equiv -1$, то 2^{2^k}, а не $2^{2^{k-1}}$, является кратным порядка элемента a по модулю p_j, т. е. его порядок равен $2^{2^k} = n - 1 > p - 1$, что невозможно. б) Предположим сначала, что $n = 2^p - 1$ есть простое число. Группа $E \pmod{n}$ состоит из 2^p точек, см. упражнение 7 а) к § VI.1. То, что эта группа — циклическая, следует из того, что имеется только две точки порядка 2: действительно, многочлен $x^3 + x$ имеет лишь один корень по модулю n. Отсюда следует, что любые из тех 50% точек, которые порождают $E \pmod{n}$ (т. е. не получаются удвоением какой-либо точки), обладают свойствами 1)–2). Обратно, предположим, что n имеет собственный простой делитель l. Если бы P обладала свойствами 1)–2), то на $E \pmod{l}$ порядок P делил бы 2^p, но не 2^{p-1}, т. е. он был бы равен 2^p. Но тогда $2^p = n + 1$ было бы делителем порядка группы $E \pmod{l}$, что противоречит теореме Хассе, утверждающей, что этот порядок не превосходит $l + 2\sqrt{l} + 1$. Для порождения случайных точек на $E \pmod{n}$ выберем случайным образом $x \in \mathbf{Z}/n\mathbf{Z}$. Если окажется, что $b = x^3 + x$ есть квадрат по модулю n, то полагаем $y = b^{(n+1)/4}$, и тогда $y^2 \equiv b \cdot b^{(n-1)/2} \equiv x^3 + x$ (см. замечание 1 в конце § II.2).

§ VI.4

1. НОД $(2^k - 1, n) = n$, но НОД $(3^k - 1, n) = 127$, $n = 127 \cdot 421$.

2. Вероятность того, что случайный вычет $a \in (\mathbf{Z}/p\mathbf{Z})^*$ удовлетворяет условию $p \mid (a^k - 1)$, равна НОД $(k, p-1)/(p-1)$. Так как имеется мало шансов, что $a^k - 1$ будет делиться на какой-либо другой делитель n, то это же число служит оценкой вероятности того, что НОД $(a^k - 1, n) = p$.

3. а) 3 из 41. б) 22 из 41. в) 25 из 127. г) 68 из 127. д) 105 из 399.

4. Выбрать $k = 2^6 \cdot 3^4 \cdot 5^2$. Далее указаны: первое значение a, для которого метод дает делитель, сам этот делитель, значение k_1, на котором алгоритм заканчивается. а) $1, 37, 2^3$. б) $2, 71, 2^6 \cdot 3^4 \cdot 5$. в) $1, 67, 2^6 \cdot 3^4 \cdot 5$. г) $1, 47, 2^6 \cdot 3$. д) $2, 79, 2^6 \cdot 3^4 \cdot 5^2$. е) $1, 73, 2^6 \cdot 3$. ж) $5, 53, 2^2$. з) $4, 59, 2^6 \cdot 3^2$. и) $1, 47, 2^6 \cdot 3$. к) $3, 97, 2^6 \cdot 3$. л) $1, 61, 2^6 \cdot 3^4 \cdot 5^2$.

5. Если бы реализовалась последняя возможность, то это означало бы, что $l'(k_1/l)P \pmod{p} = O \pmod{p}$ для некоторого $l' < l$, тогда как $(k_1/l)P \pmod{p} \neq O \pmod{p}$. Однако l' — это произведение простых чисел $l^* < l$, а наш выбор показателей в (2) гарантировал для каждого такого l^*, что наивысшая степень l^*,

которая могла бы делить порядок $P \pmod p$ в $E \pmod p$, заведомо содержится в $(l^*)^{\alpha_{l^*}}$, т.е. в k_1/l.

6. а) Если n делится лишь на простые числа p, сравнимые с 3 по модулю 4, то для каждого $p|n$ на кривой $E \pmod p$ всегда имеется в точности $p+1$ точек (см. упражнение 7 а) к § 1 для случая $a = -1$; однако те же аргументы применимы при любом a). Если при этом $p+1$ для каждого $p|n$ делится на большое простое число, то изменение a не даст результата. б) Если n делится лишь на простые числа p, сравнимые с 2 по модулю 3, то всегда имеется $p+1$ точек (см. упражнение 7 б) к § 1); и опять, если $p+1$ для каждого $p|n$ делится на большое простое число, то изменение b не приведет к результату.

7. Порождать пары (E, P), где E имеет уравнение $y^2 = x(x-a)(x-b)$. Тогда E имеет четыре точки порядка 2 (включая точку в бесконечности) (см. упражнение 4 а) к § VI. 1); для этого можно сначала случайным образом выбирать a, x, y_0 и затем полагать $y = x(x-a)y_0$ и $b = x - yy_0$.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ