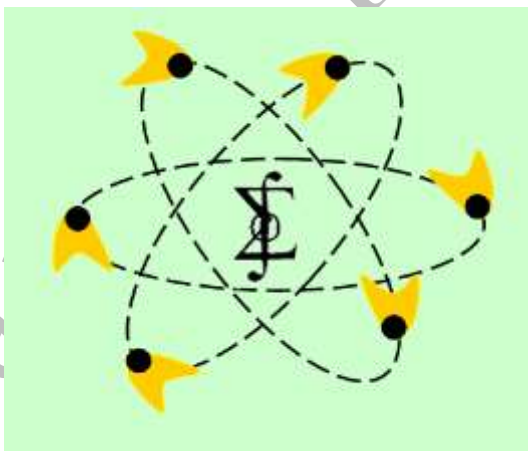
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

им. В.Ф. УТКИНА

М.Е. ИЛЬИН, К.А. ЦИПОРКОВА

ТЕОРЕТИКО-ЧИСЛОВЫЕ МЕТОДЫ В КРИПТОГРАФИИ ЧАСТЬ 1



Рязань 2020

Министерство науки и высшего образования Российской Федерации
Рязанский государственный радиотехнический университет
им. В.Ф. Уткина

М.Е. ИЛЬИН, К.А. ЦИПОРКОВА

ТЕОРЕТИКО-ЧИСЛОВЫЕ МЕТОДЫ В КРИПТОГРАФИИ ЧАСТЬ 1

Учебное пособие

Рязань 2020

УДК 511.21+681.3.06

Теоретико-числовые методы в криптографии. Ч. 1: учеб. пособие / М.Е. Ильин, К.А. Ципоркова; Рязан. гос. радиотехн. ун-т. – Рязань, 2020. – 112 с.

Рассматривается основой математический аппарат теории чисел, используемый в криптографии для построения криптографических систем защиты информации с открытым ключом. Приведены теоремы, тесты и примеры, поясняющие свойства простых и составных чисел. Алгоритмы, включенные в пособие, позволяют студентам получить навыки практического применения математического аппарата вплоть до написания программ, реализующих учебные варианты алгоритмов шифрования.

Предназначено для студентов специальности 10.05.01 «Компьютерная безопасность», а также для студентов всех направлений и специальностей, рабочие программы которых включают дисциплины, имеющие отношение к криптографии, защите информации и теории чисел.

Табл. 14. Ил. 1. Библиогр.: 6 назв.

Теория чисел, простое, составное число, алгоритм деления, алгоритм Эвклида, алгоритм и теоремы Ферма, псевдопростые числа, китайская теорема об остатках, арифметические группы, тотиент, теорема Лагранжа

Печатается по решению редакционно-издательского совета Рязанского государственного радиотехнического университета.

Рецензент: кафедра информационной безопасности Рязанского государственного радиотехнического университета (зав. кафедрой канд. техн. наук В.Н. Пржегорлинский)

Оглавление

Введение

Рассматривается математический аппарат, используемый для построения систем шифрования с открытым ключом. Наиболее известная из таких систем - RSA. В самой простой формулировке построение системы шифрования сводится к следующим действиям:

1) для реализации системы RSA необходимо иметь два больших простых числа p и q;

2) для кодирования сообщений с помощью RSA следует воспользоваться числом $n = p \cdot q$;

3) для декодирования шифровки RSA уже необходимы исходные числа p и q;

4) безопасность системы RSA (сложность декодирования) определяется тем, что разложить число n на множители p и q трудно, поскольку они очень велики и процесс разложения достаточно продолжителен.

Для понимания модели системы шифрования RSA и её реализации необходимо знать свойства целых чисел. Особенно тесно связаны с ней два вопроса:

1) как эффективно раскладывать данное целое число на множители?

2) как доказать, что данное целое число простое?

Область математики, занимающаяся изучением свойств целых чисел, называется теорией чисел, математический аппарат которой был завершен к концу девятнадцатого века усилиями древнегреческих математиков Ферма, Эйлера и Гаусса [1].

1. Основные алгоритмы

В теории чисел известны два самых фундаментальных алгоритма — алгоритм мягкого деления или деления с остатком и алгоритм Эвклида. Алгоритм деления предназначен для вычисления неполного частного и остатка при делении двух целых чисел. Алгоритм Эвклида вычисляет наибольший общий делитель двух целых чисел. Эти алгоритмы в теории чисел применяются самостоятельно и как составные части ряда теорем и тестов [1, 2].

При построении алгоритма необходимо определить его «ввод» и «вывод». Ввод соответствует совокупности исходных данных, над которыми действует алгоритм. Вывод — результату, который следует получить. Собственно алгоритм — это набор действий, которые нужно совершить над вводом, чтобы получить вывод.

Для теоремы, связанной с алгоритмом, ввод алгоритма соответствует предположению теоремы, а вывод — заключению. Если алгоритм сформулирован и необходимо проверить его корректность, то необходимо ответить на следующие вопросы.

1. Всегда ли при исполнении инструкций будет получен результат за конечное время?
2. Совпадает ли результат с ожидаемым?

Под доказательством мы понимаем логическое рассуждение, в основе которого лежат базисные факты, или аксиомы, о которых мы договорились заранее. Для большинства алгоритмов в качестве аксиом выступают элементарные свойства целых чисел.

1.1. Алгоритм деления

Проанализируем алгоритм. Нас интересует деление целых чисел, поэтому задача состоит в том, чтобы найти неполное частное и остаток от деления двух положительных целых чисел. При этих словах приходит на ум схема деления «уголком», подобная:

1	2	3	4	5	4
1	0	8		2	2
	1	5	4		
	1	0	8		
		4	6		

В этом примере мы делим 1234 на 54; неполное частное оказалось равным 22, а остаток равен 46. Вводом алгоритма служат

делимое и делитель; в приведённом примере они равны соответственно 1234 и 54. Вывод состоит из частного и остатка, значения которых в примере 22 и 46.

В общем случае ввод алгоритма деления состоит из двух положительных целых чисел a и b. Поделив a на b, получаем при выводе числа q и r, которые связаны с a и b следующим соотношением:

$$a = \mathrm{q} \cdot \mathrm{b} + \mathrm{r}, 0 \leq \mathrm{r} < \mathrm{b}. \quad (1.1)$$

Здесь q - неполное частное, а r — остаток от деления. У этого определения есть простая интерпретация, которую стоит иметь в виду. Последовательно вычитаем из делимого делитель. Процесс вычитания продолжается до тех пор, пока делимое станет меньше делителя. Эта интерпретация приводит к простейшей реализации алгоритма деления.

Алгоритм 1.1 (алгоритм деления)

Ввод: натуральные числа a и b.

Вывод: неотрицательные целые числа q и r, для которых выполнено равенство: $a = q \cdot b + r, 0 \leq r < b$.

Шаг 1. Положить $Q = 0$ и $R = a$.

Шаг 2. Если $R < b$, то сообщить: «частное равно Q, а остаток равен R» и остановиться; в противном случае перейти к шагу 3.

Шаг 3. Если $R > b$, то вычесть b из R, увеличить Q на 1 и возвратиться к шагу 2.

Предположим, например, что $a > b$. Тогда после первого прохода через шаг 3 мы получим $Q = 1$ и $R = a - b$. Если $a — b \geq b$, то, согласно алгоритму, мы должны выполнить шаг 3 ещё раз. Проделав это, мы получим $Q = 2$ и $R = a — 2b$, и т.д. Почему такой процесс не может повторяться бесконечно? Другими словами, почему алгоритм прекращает свою работу? Заметим, что в результате последовательного применения шага 3 мы получаем следующую последовательность значений переменной R:

Начальное значение	1-й цикл	2-й цикл	3-й цикл	…
a	$a - b$	$a - 2b$	$a - 3b$	…

Это убывающая последовательность положительных целых чисел. Поскольку количество целых чисел между a и 0 конечно, то последовательность остатков $a, a - 2b, a - 3b, \cdots$ с необходимостью приводит к числу, меньшему b. Тогда на шаге 2 алгоритма работа

останавливается и алгоритм выводит значения переменных R и Q. Следовательно алгоритм завершает работу.

1.2. Теорема деления

Сформулируем теорему, соответствующую алгоритму деления.

Теорема 1.1 (деление). *Пусть a и b — натуральные числа. Тогда существует единственная пара неотрицательных целых чисел q и r таких, что*

$$a = \mathrm{q} \cdot \mathrm{b} + \mathrm{r}, 0 \leq \mathrm{r} < \mathrm{b}. \quad (1.2)$$

Доказательство. Доказательство существования обосновано приведенным выше алгоритмом деления 1.1. Более того, он же определяет и правило вычисления чисел q и r.

Доказательство единственности. Пусть даны два числа a и $b, b \neq 0$. Допустим противное, что существуют, по крайней мере, две различные пары чисел q и r, а также q' и r' таких, что одновременно выполняются различные тождества

$$a \equiv qb + r, 0 \leq r < b;$$
$$a \equiv q'b + r', 0 \leq r' < b.$$

Выразим из этих тождеств остатки. Имеем $r \equiv a - qb$ и $r' \equiv a - q'b$. Вычитая эти равенства друг из друга, получаем:

$$r - r' \equiv (a - qb) - (a - q'b) = (q' - q)b.$$

Рассмотрим по отдельности левые и правые части этого тождества. Левая часть тождества – целое число, и в силу определения остатка (1.2) справедливы неравенства $0 \leq r < b$ и $0 \leq r' < b$. Это означает, что разность $r - r'$ заведомо больше $-b$ и одновременно меньше b. То есть подчинена двойному неравенству

$$-b < r - r' < b.$$

Правая часть тождества $(q' - q)b$ – целое число, кратное b. Это означает, что и левая часть тождества также кратна b. Единственное целое число, кратное b и принадлежащее множеству $(-b, b)$, — это число ноль. Другими словами, выполняется тождество $r - r' \equiv 0$. Это, в свою очередь, означает, что справедливо и другое тождество

$$0 \equiv r - r' \equiv (a - qb) - (a - q'b) = (q' - q)b$$

или

$$(q' - q)b \equiv 0.$$

Поскольку $b \neq 0$, то $q = q'$. Теорема доказана.

1.3. Алгоритм Эвклида

Алгоритм Эвклида предназначен для вычисления наибольшего общего делителя двух натуральных чисел. Введем основные термины.

Определение 1.1. *Целое число b делит целое число a, если существует ещё одно целое число c такое, что $a = b \cdot c$.*

В этом случае говорят также, что b является делителем или множителем числа a, а a, в свою очередь, — кратным числу b. Определение дает и метод проверки делителя. Для этого достаточно подсчитать остаток от деления a на b и проверить, равен ли он нулю.

Определение 1.2. *Пусть a и b — целые числа. Наибольший общий делитель чисел a и b — это наибольшее целое число d, на которое и a, и b делятся; тогда пишут:*

$$d = \text{НОД}(a, b). \qquad (1.3)$$

Определение 1.3. *Целые числа называются взаимно простыми, если их наибольший общий делитель равен единице НОД$(a, b) = 1$.*

Определение наибольшего общего делителя подсказывает следующий алгоритм его вычисления. Если числа a и b заданы, то найдём все положительные делители числа a и все положительные делители числа b. Выберем все числа, входящие в оба множества, и возьмём наибольшее из них. Оно и будет наибольшим общим делителем. Эта процедура проста, при больших a и b не эффективна, и неизвестно ни одного простого алгоритма разложения целых чисел на множители. Существует эффективный способ поиска наибольшего общего делителя - алгоритм Эвклида, который работает следующим образом.

Разделим a на b с остатком; назовём этот остаток r_1. Если $r_1 \neq \neq 0$, то разделим b на r_1 с остатком; пусть r_2 — остаток второго деления. Аналогично, если $r_2 \neq 0$, то разделим r_1 на r_2 и получим новый остаток r_3.

Таким образом, i-й цикл алгоритма состоит из одного деления с остатком, причём делимое равно остатку, полученному в $(i — 2)$-м цикле, а делитель — остатку, полученному в $(i — 1)$-м цикле. Цикл повторяется до тех пор, пока мы не получим нулевого остатка; наименьший ненулевой остаток - наибольшим общим делителем чисел a и b.

Пример 1.1. *Вычислить наибольший общий делитель чисел 1234 и 54 с помощью алгоритма Эвклида.*

Решение. Построим последовательность делений с остатком.

Шаг 1. Выполняем деление числа 1234 на число 54. Получаем $1234 = 22 \cdot 54 + 46$.

Шаг 2. Выполняем деление числа 54 на число 46. Получаем $54 = 1 \times 46 + 8$.

Шаг 3. Выполняем деление числа 46 на число 8. Получаем $46 = 5 \times 8 + 6$.

Шаг 4. Выполняем деление числа 8 на число 6. Получаем $8 = 1 \cdot 6 + 2$.

Шаг 5. Выполняем деление числа 6 на число 2. Получаем $6 = 3 \cdot 2 + 0$.

Последний ненулевой остаток на шаге 4 равен 2, поэтому верно

$$\text{НОД}(1234,54) = 2.$$

Отметим, что неполные частные не принимают непосредственного участия в подсчёте наибольшего общего делителя.

Алгоритм 1.2 (алгоритм Эвклида)

Ввод: натуральные числа a и b, $a > b$.

Вывод: наибольший общий делитель чисел a и b.

Шаг 1. Положить $A = a$ и $R = B = b$.

Шаг 2. Заменить значение R остатком от деления A на B и перейти к шагу 3.

Шаг 3. Если $R = 0$, то сообщить: «наибольший общий делитель чисел a и b равен B» и остановиться; в противном случае перейти к шагу 4.

Шаг 4. Заменить значение A на значение B, значение B на значение R и возвратиться к шагу 2.

Таким образом, для вычисления наибольшего общего делителя нам необходимо лишь выполнить несколько делений с остатком. Обоснуем истинность алгоритма Эвклида. Для этого докажем вспомогательное утверждение.

Лемма 1.1. *Пусть a и b — натуральные числа. Предположим, что существуют такие целые числа g и s, при которых $a = b \cdot g + s$. Тогда выполняется равенство:*

$$\text{НОД}(a, b) = \text{НОД}(b, s). \qquad (1.4)$$

Доказательство. Введем обозначения

$$d_1 = \text{НОД}(a, b), d_2 = \text{НОД}(b, s).$$

Необходимо доказать, что $d_1 \equiv d_2$. Покажем, что выполняется неравенство $d_1 \leq d_2$. Из равенства $d_1 = НОД(a, b)$ следует, что d_1 делит как a, так и b. Это означает, что существуют такие натуральные числа u и v, что выполняются равенства

$$a = d_1 \cdot u \text{ и } b = d_1 \cdot v.$$

Подставим эти значения в равенство $a = b \cdot g + s$ и получим:

$$d_1 \cdot u = d_1 \cdot v \cdot g + s.$$

Выразим из равенства величину s:

$$s = d_1 u — d_1 v \cdot g = d_1(u — vg).$$

Последнее равенство означает, что s делится на d_1. Поскольку $d_1 = НОД(a, b)$, то d_1 делит b. То есть число d_1 одновременно делит и s, и b. Поэтому d_1 является общим делителем чисел b и s. Однако d_2 - наибольший из таких общих делителей, поэтому $d_1 \leq d_2$.

Аналогично доказывается и неравенство $d_2 \leq d_1$. Два неравенства $d_1 \leq d_2$ и $d_2 \leq d_1$ справедливы тогда и только тогда, когда выполняется тождество

$$d_1 \equiv d_2.$$

Лемма доказана.

Отметим, что при доказательстве леммы существенно используется соотношение $a = b \cdot g + s$.

Теорема 1.2 (алгоритм Эвклида). *Последний отличный от нуля остаток алгоритма Эвклида – наибольший общий делитель натуральных чисел a и b.*

Доказательство. Необходимо убедиться, что в последовательности остатков заведомо появится ноль и наибольший общий делитель совпадет с последним остатком, отличным от нуля. Обоснуем появление нуля в качестве остатка и тем самым покажем, что алгоритм завершит работу. Допустим, что для того чтобы найти наибольший общий делитель чисел a и b, выполнены деления:

$$a = b \cdot q_1 + r_1, 0 \leq r_1 < b;$$
$$b = r_1 \cdot q_2 + r_2, 0 \leq r_2 < r_1;$$
$$r_1 = r_2 \cdot q_3 + r_3, 0 \leq r_3 < r_2;$$
$$r_2 = r_3 \cdot q_4 + r_4, 0 \leq r_4 < r_3.$$

В правых частях неравенств расположена последовательность остатков. Всякий остаток заведомо меньше предыдущего, а также все остатки неотрицательны. Переписав неравенства последовательно, получим цепочку:

$$0 \le \cdots < r_3 < r_2 < r_1 < \mathrm{b}. \tag{1.5}$$

Поскольку между нулем и b расположено конечное число целых чисел, то последовательность остатков не может продолжаться бесконечно, в конце её может стоять только нуль, и, следовательно, алгоритм остановится. Конечность работы алгоритма обоснована.

Докажем, что последний ненулевой остаток – искомый наибольший общий делитель. Применим алгоритм к целым числам $0 < b < a$, предположим, что остаток после n-го деления равен нулю, получим следующую последовательность делений

$$\begin{cases} a = b \cdot q_1 + r_1, 0 \le r_1 < \mathrm{b}; \\ b = r_1 \cdot q_2 + r_2, 0 \le r_2 < r_1; \\ r_1 = r_2 \cdot q_3 + r_3, 0 \le r_3 < r_2; \\ r_2 = r_3 \cdot q_4 + r_4, 0 \le r_4 < r_3; \\ \cdots \\ r_{n-4} = r_{n-3} \cdot q_{n-2} + r_{n-2}, 0 \le r_{n-2} < r_{n-3}; \\ r_{n-3} = r_{n-2} \cdot q_{n-1} + r_{n-1}, 0 \le r_{n-1} < r_{n-2}; \\ r_{n-2} = r_{n-1} \cdot q_n, r_n = 0. \end{cases} \tag{1.6}$$

Обратим внимание на левые части равенств. Последнее равенство означает, что r_{n-2} делится на r_{n-1}. Поэтому наибольший общий делитель этих двух чисел равен r_{n-1}. Другими словами,

$$\text{НОД}(r_{n-2}, r_{n-1}) = r_{n-1}.$$

Теперь воспользуемся леммой 1.1. Применяя её к предпоследнему равенству, мы заключаем, что

$$\text{НОД}(r_{n-3}, r_{n-2}) = \text{НОД}(r_{n-2}, r_{n-1}) = r_{n-1},$$

причём последняя величина по-прежнему равна r_{n-1}. Повторное применение леммы 1.1 к полученному равенству даёт:

$$\text{НОД}(r_{n-4}, r_{n-3}) = \text{НОД}(r_{n-3}, r_{n-2}) = r_{n-1},$$

что опять равно r_{n-1}. Действуя таким же образом до первой строки (1.6), получаем искомый результат $НОД(a, b) = r_{n-1}$. Теорема доказана полностью.

Получим верхнюю оценку числа делений, необходимых для вычисления наибольшего общего делителя. Вернёмся к неравенствам (1.6). Каждое число в последовательности остатков строго меньше

предыдущего. Поэтому наибольшее возможное значение остатка в каждом делении на единицу меньше значения остатка на предыдущем делении. Если бы в каждом цикле это наибольшее возможное значение достигалось, то для получения нулевого остатка нам потребовалось бы b делений. Это и есть наихудший возможный случай. Поэтому при применении алгоритма Эвклида к паре чисел a и b число делений заведомо не превосходит b.

Нетрудно заметить, что при $b > 3$ число делений всегда меньше b. Зафиксируем число n. Найдем наименьшие взаимно простые a и b, для которых вычисление $НОД(a, b)$ требует n делений.

Для того, чтобы числа a и b были минимально возможными, частные на каждом шаге тоже должны быть минимально возможными. Поскольку делитель меньше делимого, то наименьшее возможное частное двух целых чисел равно 1. Допустим, что выполнено n делений до получения нулевого остатка. Тогда последовательность остатков имеет вид (1.6). В наихудшем возможном случае все частные равны 1. Запишем теперь все деления (1.6), начиная с последнего. В силу взаимной простоты чисел мы получим:

$$r_{n-2} = r_{n-1} \cdot 1 + 0;$$
$$r_{n-3} = r_{n-2} \cdot 1 + r_{n-1};$$
$$r_{n-4} = r_{n-3} \cdot 1 + r_{n-2};$$
$$r_{n-5} = r_{n-4} \cdot 1 + r_{n-3};$$
$$\ldots$$
$$a = b \cdot 1 + r_1.$$

Выпишем элементы этой последовательности остатков для $n = 11$:

$$0,1,1,2,3,5,8,13,21,34,55.$$

Значит, наименьшая пара взаимно простых чисел a и b, для подсчёта наибольшего общего делителя которых необходимо 11 делений с остатком, это $a = 55$ и $b = 34$. Заметьте, что, хотя число $b = 34$ и наименьшее, все равно оно больше, чем $n = 11$. Приведённая последовательность — начало последовательности Фибоначчи.

1.4. Расширенный алгоритм Эвклида

Алгоритм Эвклида допускает модификацию, в которой наибольший общий делитель — лишь часть выходных данных.

Теорема 1.3 (расширенный алгоритм Эвклида). *Пусть a и b — натуральные числа, а d — их наибольший общий делитель. Тогда существуют два целых числа α и β таких, что*

$$\alpha \cdot a + \beta \cdot b = d. \tag{1.7}$$

Отметим, что (за исключением нескольких тривиальных случаев) a и β имеют противоположные знаки. Для вычисления этих чисел добавим дополнительные инструкции в обычный алгоритм Эвклида 1.2 так, чтобы d, α и β подсчитывались одновременно. Основная идея алгоритма: записать каждый ненулевой остаток в виде (1.7).

Допустим, что для вычисления наибольшего общего делителя чисел a и b выполнена последовательность делений (1.6). Дополним её записью предполагаемого представления остатка:

$$\begin{cases} a = b \cdot q_1 + r_1, r_1 = \mathrm{a} \cdot x_1 + b \cdot y_1; \\ \mathrm{b} = r_1 \cdot q_2 + r_2, r_2 = \mathrm{a} \cdot x_2 + b \cdot y_2; \\ \mathrm{r}_1 = r_2 \cdot q_3 + r_3, r_3 = \mathrm{a} \cdot x_3 + b \cdot y_3; \\ \mathrm{r}_2 = r_3 \cdot q_4 + r_4, r_4 = \mathrm{a} \cdot x_4 + b \cdot y_4; \\ \dots \\ \mathrm{r}_{n-4} = r_{n-3} \cdot q_{n-2} + r_{n-2}, r_{n-2} = \mathrm{a} \cdot x_{n-2} + b \cdot y_{n-2}; \\ \mathrm{r}_{n-3} = r_{n-2} \cdot q_{n-1} + r_{n-1}, r_{n-1} = \mathrm{a} \cdot x_{n-1} + b \cdot y_{n-1}; \\ \mathrm{r}_{n-2} = r_{n-1} \cdot q_n, r_n = 0. \end{cases} \tag{1.8}$$

Числа $x_1, x_2, \dots, x_{n-1}$ и $y_1, y_2, \dots, y_{n-1}$ необходимо найти. Сведем всю информацию в таблицу.

п/п	Остатки	Частные	x	y
1	a		x_{-1}	y_{-1}
2	b		x_0	y_0
3	r_1	q_1	x_1	y_1
4	r_2	q_2	x_2	y_2
5	r_3	q_3	x_3	y_3
⋮	⋮	⋮	⋮	⋮
n-2	r_{n-2}	q_{n-2}	x_{n-2}	y_{n-2}
n-1	r_{n-1}	q_{n-1}	x_{n-1}	y_{n-1}

Таблица начинается с двух вспомогательных строк, которые используются для запуска алгоритма.

Для доказательства теоремы следует указать правила заполнения граф x и y. Пусть таблица заполнена до некоторой, скажем до $(j-1)$-й, строки. Для заполнения j-й строки необходимо, прежде

всего, разделить r_{j-2} на r_{j-1}. В результате получатся r_j и q_j — первые два элемента j-й строки. Не будем забывать, что

$$r_{j-2} = r_{j-1} \cdot q_j + r_j, 0 \le r_j < r_{j-1}.$$

Таким образом,

$$r_{\text{j}} = r_{j-2} - r_{j-1} \cdot q_j. \tag{1.9}$$

В строчках $j-1$ и $j-2$ известны величины $x_{j-2}, x_{j-1}, y_{j-2}, y_{j-1}$, а следовательно, определяются и остатки

$$r_{j-2} = a \cdot x_{j-2} + b \cdot y_{j-2}, r_{j-1} = a \cdot x_{j-1} + b \cdot y_{j-1}.$$

Подставляя эти значения в (1.9), получаем:

$$r_j = \left(a \cdot x_{j-2} + b \cdot y_{j-2}\right) - \left(a \cdot x_{j-1} + b \cdot y_{j-1}\right) \cdot q_j$$
$$= a \cdot \left(x_{j-2} - q_j \cdot x_{j-1}\right) + b \cdot \left(y_{j-2} - q_j \cdot y_{j-1}\right)$$

Поэтому

$$x_j = x_{j-2} - q_j \cdot x_{j-1}, y_j = y_{j-2} - q_j \cdot y_{j-1}.$$

Заметим, что для вычисления x_j и y_j необходимы только частное q_j и данные из двух строк таблицы, непосредственно предшествующих j-й строке.

Получена рекуррентная процедура, для запуска которой необходимы начальные данные. Именно для них добавлены две строки в таблицу. Определим в них значения x и y. Придавая им тот же смысл, что и в остальных строках, получаем:

$$a = a \cdot x_{-1} + b \cdot y_{-1}, b = a \cdot x_0 + b \cdot y_0.$$

Поэтому следует положить

$$x_{-1} = 1, y_{-1} = 0, x_0 = 0, y_0 = 1,$$

и процедура может быть запущена. В результате цепочки делений с остатком мы получим равенство $НОД(a, b) = r_{n-1}$ и вычислим такие целые числа x_{n-1} и y_{n-1}, что

$$d = r_{n-1} = a \cdot x_{n-1} + b \cdot y_{n-1}.$$

Значит, $\alpha = x_{n-1}$ и $\beta = y_{n-1}$. Заметим, что, зная α и $d = r_{n-1}$, β определяется выражением

$$\beta = \frac{d - a \cdot \alpha}{b}.$$

Поэтому достаточно вычислять только первые три графы таблицы.

Пример 1.2. *С помощью расширенного алгоритма Эвклида вычислить НОД*$(1234, 54)$.

Решение. Для поиска наибольшего общего делителя составим таблицу:

п/п	Остатки	Частные	x	y
1	1234		1	0
2	54		0	1
3	46	22	$1-22\cdot 0=1$	$0-22\cdot 1=-22$
4	8	1	$0-1\cdot 1=-1$	$1-1\cdot(-22)=23$
5	6	5	$1-5\cdot(-1)=$ $=6$	$-22-5\cdot 23=$ $=-137$
6	2	1	$-1-1\cdot 6=$ $=-7$	$23-1\cdot(-137)=$ $=160$
7	0	3		

Поэтому $\alpha=-7,\beta=160$ и справедливо равенство

$$(-7)\cdot 1234 + 160\cdot 54 = 2.$$

Расширенный алгоритм Эвклида представляет собой просто алгоритм Эвклида из раздела 1.3, дополненный инструкциями для вычисления значений x и y. Поэтому он останавливается и наибольший общий делитель составляет часть выходных данных. Кроме того, в каждой строке числа из столбцов x и y удовлетворяют равенству вида (1.7), в котором число d заменено остатком из соответствующей строки. В частности, равенство (1.7) выполняется, если в качестве α и β взять числа из столбцов x и y строки, отвечающей последнему ненулевому остатку.

Замечание. Пара чисел α,β, упомянутая в теореме 1.3, не единственная. На самом деле таких пар бесконечно много. Например, возьмём α и β, для которых $\alpha\cdot a+\beta\cdot b\equiv d$, и рассмотрим какое-нибудь целое k. Тогда получаем тождество

$$(\alpha+k\cdot b)\cdot a+(\beta-k\cdot a)\cdot b\equiv d.$$

2. Факторизация

Методы анализа и синтеза – одни из основных методологических подходов в научных исследованиях, в том числе и в теории чисел: произвольное целое число раскладывается в произведение простых. Разложение служит главным инструментом в доказательстве многих свойств простых чисел. Найти разложение данного числа, однако, не всегда легко. Если число очень велико, то процедура его

разложения предъявляет большие требования к мощности компьютера [3].

2.1. Теорема разложения

Начнём с определения основных терминов и понятий.

Определение 2.1. *Целое число p называется простым, если* $p \neq \neq \pm 1$ *и единственными его делителями являются числа* ± 1 *и* $\pm p$.

Так, числа $2,3,5$ и -7 простые, а число $45 = 5 \cdot 9$ — нет. Почти всюду будем использовать определение в несколько более узком смысле, называя простым положительное простое число.

Определение 2.2. *Целое число, отличное от* ± 1 *и непростое, называется составным, или разложимым.*

Для составного числа n существуют такие целые числа a и b, что $1 < a, b < n$ и $n = a \cdot b$. Значит, число 45 составное.

Заметим, что числа ± 1 не являются ни составными, ни простыми. Они относятся к третьей группе — это единственные целые числа, у которых есть целые обратные.

Теорема 2.1 (разложение на множители). *Всякое целое число* $n \geq 2$ *единственным образом записывается в виде произведения:*

$$n = p_1^{e_1} \cdot p_2^{e_2} \cdot \cdots \cdot p_k^{e_k}, \qquad (2.1)$$

где $p_1, p_2, \ldots, p_k$ *— простые числа,* $1 < p_1 < p_2 < \cdots < p_k$ *и* $e_1, e_2, \ldots, e_k$ *— натуральные числа.*

Эта теорема настолько важна, что её иногда называют основной теоремой арифметики.

Показатели $e_1, e_2, \ldots, e_k$ называются кратностями простых множителей в разложении числа n. Другими словами, кратностью множителя p_i в разложении числа n называется наибольшее число e_i такое, что n делится на $p_i^{e_i}$. Отметим также, что у числа n есть k различных простых делителей, однако общее число его простых делителей равно $e_1 + e_2 + \cdots + e_k$.

Теорема содержит два различных утверждения. Во-первых, всякое целое число можно представить в виде произведения степеней простых чисел. Во-вторых, простые сомножители и их степени в разложении определены однозначно. Значит, нам нужно доказать два положения: разложение существует, и оно единственно. Мы докажем их по отдельности.

После того, как теорема о разложении сформулирована, понятно, почему числа ± 1 не следует считать простыми. Если включить их в число простых, то разложение на простые сомножители потеряет свойство единственности. Действительно, если число 1 простое, то 2 и $1^2 \cdot 2$ — два различных разложения числа 2 на простые сомножители. Это даёт бесконечно много разложений для любого целого числа. Поэтому числа ± 1 исключены из определения простого числа.

2.2. Существование разложения

В этом разделе показывается, что всякое целое число $n \geq 2$ может быть записано в виде произведения простых. Для доказательства этого утверждения мы приводим алгоритм, получающий на входе целое число $n \geq 2$ и выдающий простые сомножители числа n и их кратности [2]. В качестве предварительного шага построим алгоритм, выходом которого является какой-нибудь простой делитель числа n.

Для поиска простых делителей проще всего воспользоваться следующим алгоритмом. Попробуем разделить n на все целые числа от 2 до $n-1$ подряд. Если одно из них делит n, то число n составное, и мы нашли наименьший из его делителей. В противном случае число n простое. Кроме того, если n составное, то найденный нами делитель обязан быть простым.

Убедимся в справедливости этого положения. Пусть f - целое число, такое что $1 < f < n$. Предположим, что f - наименьший делитель числа n и пусть $f' > 1$ — делитель числа f. По определению делимости, существуют такие натуральные числа a и b, что

$$n = f \cdot a, f = f' \cdot b.$$

Значит, $n = f' \cdot (a \cdot b)$ и f' также является делителем числа n. Поскольку f - наименьший делитель n, должно выполняться неравенство $f \leq f'$, но f' делит f, поэтому $f' \leq f$. Эти неравенства означают, что $f \equiv f'$. Тем самым мы доказали, что если число $f' \neq 1$ делит f, то оно совпадает с f. Значит, f простое.

Прежде, чем перейти к описанию алгоритма, следует определить, как долго следует производить поиск делителя? Очевидно, что $n-1$ – максимально возможный кандидат: делитель числа не может превышать его самого. Это верхняя граница поиска делителей.

Однако можно утверждать, что не следует искать делители, превышающие $\sqrt{n}$.

Последнее утверждение вновь вытекает из того, что алгоритм ищет наименьший делитель числа n, больший чем 1. Поэтому нам необходимо показать только, что наименьший делитель $f > 1$ числа n удовлетворяет неравенству $f \leq \sqrt{n}$. Это утверждение легко проверить. Пусть $n = f \cdot a$. Поскольку $f > 1$ — наименьший делитель числа n, имеем $f \leq a$. Теперь $a = \frac{n}{f}$ и, следовательно, $f \leq \frac{n}{f}$, откуда вытекает, что $f^2 \leq \sqrt{n}$. Другими словами, $f \leq \sqrt{n}$, что и требовалось доказать. Обозначим через $\lfloor a \rfloor$ целую часть вещественного числа a. Другими словами, $\lfloor a \rfloor$ — наибольшее целое число, меньшее или равное a. Так, $\lfloor \pi \rfloor = 3$ и $\lfloor \sqrt{2} \rfloor = 1$. Заметим, что для целого числа $r \leq a$ выполняется неравенство $r \leq \lfloor a \rfloor$. Значит, для реализации описанного выше алгоритма разложения нам достаточно знать только значение $\lfloor \sqrt{n} \rfloor$.

Алгоритм выполняет проверку (начиная с 2 и пробегая по натуральным числам, не превышающим $\sqrt{n}$) того, делится ли n на очередное число. Для составного числа n, таким образом, будет найден его наименьший делитель, больший или равный 2. Как мы уже доказали, этот делитель будет обязательно простым. Если в процессе проверки ни один делитель не будет обнаружен, то само n — простое.

Алгоритм 2.1 (алгоритм разложения путём деления методом проб)

Ввод: натуральное число n.

Вывод: натуральное число $f > 1$ — наименьший простой делитель числа n — или сообщение о том, что n простое.

Шаг 1. Положить $F = 2$.

Шаг 2. Если $\frac{n}{F}$ целое, то сообщить: «F является делителем числа n» и завершить работу; в противном случае перейти к шагу 3.

Шаг 3. Увеличить F на единицу и перейти к шагу 4.

Шаг 4. Если $F > \lfloor \sqrt{n} \rfloor$, то сообщить: «$n$ простое» и завершить работу; в противном случае перейти к шагу 2.

Итак, приведен способ определения, является ли число $n > 2$ простым, и подсчитать его делитель, если n составное. Разумеется, если n простое, то мы находим и его разложение.

Однако, если n составное, то необходимо найти все его делители и указать их кратности. Для этого достаточно применить описанный алгоритм несколько раз.

Предположим, что в результате применения предыдущего алгоритма к n мы нашли делитель q_1 этого числа. Тогда q_1 — его наименьший простой множитель. Применим тот же алгоритм к числу $\frac{n}{q_1}$. Предположим, что число $\frac{n}{q_1}$ составное и q_2 — его наименьший простой делитель. Ясно, что $q_2 \geq q_1$.

Заметим, однако, что эти делители могут оказаться равными. Такая возможность реализуется, если n делится на q_1^2. Продолжая в том же духе, мы применяем алгоритм к $\frac{n}{q_1 q_2}$ и т.д.

В результате мы получаем последовательность простых чисел:

$$q_1 \leq q_2 \leq q_3 \leq \cdots \leq q_s,$$

каждое из которых является делителем числа n. Этой последовательности отвечает последовательность частных

$$\frac{n}{q_1} > \frac{n}{q_1 q_2} > \frac{n}{q_1 q_2 q_3} > \cdots > \frac{n}{q_1 q_2 \cdots q_s}.$$

Заметим, что это строго убывающая последовательность натуральных чисел, каждое из которых соответствует применению алгоритма деления методом проб к n. Поскольку множество натуральных чисел, меньших n, конечно, после нескольких шагов мы получим полное разложение n на простые множители. Несложно проверить, что последнее число в последовательности частных равно 1, что и служит критерием завершения работы.

Предположим, теперь, что мы хотим представить разложение в том виде, в котором оно записано в теореме о разложении. Все простые делители нам известны, осталось лишь подсчитать их кратности. Для этого нужно вычислить, сколько раз каждый простой сомножитель встречается в указанной последовательности простых чисел.

Пример 2.1. *Найти разложение числа* $n = 450$ *в произведение простых множителей.*

Решение. Воспользуемся алгоритмом деления методом проб. Алгоритм методом проб даёт первый наименьший простой множитель 2. Повторное применение алгоритма, на этот раз к частному $\frac{450}{2} = 225$, дает множитель 3. Значит, делитель 3 числа 225 является и делителем числа 450. Применим алгоритм к числу $\frac{225}{3} = 75$. Вновь наименьшим делителем оказывается 3. Значит, 450 делится на 3^2. Еще два выполнения алгоритма деления методом проб показывают, что 25 делится на 5^2, причём частное равно 1. Поэтому мы нашли полное разложение: $450 = 2 \cdot 3^2 \cdot 5^2$.

2.3. Эффективность алгоритма деления методом проб

Описанный в предыдущем параграфе алгоритм легко понять и запрограммировать, однако он оказывается чрезвычайно неэффективным. Поскольку для поиска разложения мы многократно используем алгоритм деления методом проб, следует оценить эффективность последнего. Проиллюстрируем это простым примером.

При применении алгоритма деления методом проб к натуральному числу $n > 2$ худшим случаем оказывается тот, когда n простое. В этом случае алгоритм выполняет до остановки $\lfloor \sqrt{n} \rfloor$ циклов. Для упрощения вычислений предположим, что n простое и в нем не меньше ста цифр. Сколько времени потребуется для подтверждения простоты числа n с помощью алгоритма деления методом проб?

Мы предполагаем, что $n > 10^{100}$, т.е. $\sqrt{n} > 10^{50}$. Значит, мы должны повторить цикл, по меньшей мере, 10^{50} раз. Чтобы прикинуть, сколько это займёт времени, допустим, что наш компьютер выполняет 10^{10} делений в секунду. Здесь мы предполагаем, что никаких других операций, кроме операции деления, в цикле нет. Безусловно, это далеко не так, однако сделаем такое допущение. Разделив первое число на второе, мы заключаем, что компьютеру потребуется 10^{40} секунд на проверку простоты числа n. Простой подсчёт показывает, что на это уйдёт приблизительно 10^{31} лет. По современным стандартам это чересчур большой срок. Чтобы представить его себе, вспомним, что согласно наиболее распространённой точке зрения, Большой Взрыв произошёл около $2 \cdot 10^{11}$ лет назад.

Означает ли это, что алгоритм деления методом проб бесполезен? Разумеется, нет. Допустим, что у раскладываемого числа есть простой множитель, скажем, меньший чем 10^6. Тогда алгоритм деления методом проб быстро его отыщет.

С другой стороны, если у нас есть основания полагать, что тестируемое число простое, то алгоритм деления методом проб не выглядит наилучшим решением. Есть много других алгоритмов разложения целых чисел, эффективность которых зависит от типа введённого числа. Так, алгоритм раздела 2.2 очень хорош для чисел с маленькими простыми делителями. В следующем разделе рассмотрим алгоритм Ферма, который наиболее эффективен для таких чисел n, у которых есть делитель (не обязательно простой), не сильно превышающий $\sqrt{n}$.

Следует помнить, что эффективный алгоритм разложения на множители случайно выбранного натурального числа неизвестен. Неясно только, действительно ли он не существует или пока он не найден.

2.4. Алгоритм Ферма разложения на множители

Алгоритм раздела 2.2 эффективен в случае, если у числа n, разложение которого мы ищем, есть малый простой делитель. В настоящем параграфе мы рассмотрим алгоритм, который эффективен, когда у числа n есть делитель (не обязательно простой), незначительно превосходящий $\sqrt{n}$. Идея алгоритма восходит к Ферма, и она требует гораздо большей изобретательности, чем алгоритм деления методом проб.

Предположим, что n нечетное. В противном случае число 2 было бы его делителем. Идея алгоритма состоит в том, чтобы попробовать представить n в виде $n = x^2 - y^2$, где x, y — неотрицательные целые числа. Если такие числа найдены, то

$$n = x^2 - y^2 = (x - y)(x + y).$$

Значит, $x - y$ и $x + y$ являются делителями числа n.

Допустим, что компьютер умеет вычислять целую часть числа $\sqrt{n}$. Проще всего применять алгоритм Ферма к полным квадратам. Если $n = r^2$ для некоторого целого числа r, то r является делителем n. Тогда $x = r$ и $y = 0$. С другой стороны, если $y > 0$, то

$$x = \sqrt{n + y^2} > \sqrt{n}.$$

Алгоритм 2.2 (алгоритм Ферма разложения на множители)

Ввод: нечётное натуральное число n.

Вывод: множитель числа n или сообщение о том, что n простое.

Шаг 1. Положить $x = \lfloor\sqrt{n}\rfloor$. Если $n = x^2$, то x является делителем числа n и работа алгоритма останавливается; в противном случае увеличить x на 1 и перейти к шагу 2.

Шаг 2. Если $x = \frac{n+1}{2}$, то число n простое и работа алгоритма останавливается; в противном случае вычислить $y = \sqrt{x^2 - n}$.

Шаг 3. Если число y целое (т.е., если $\lfloor\sqrt{n}\rfloor^2 = x^2 - n$), то n раскладывается в произведение $(x + y)(x — y)$ и работа алгоритма останавливается; в противном случае увеличить x на 1 и перейти к шагу 2.

Как показывает следующий пример, этот алгоритм достаточно прост в реализации.

Пример 2.2. *Разложить на множители с помощью алгоритма Ферма число* $n = 1342127$.

Решение. Присвоим переменной x целую часть числа $\sqrt{n}$. Она равна $x = 1158$. Выполняется соотношение

$$x^2 = 1158^2 = 1340964 < 1342127.$$

Поэтому мы должны увеличить x на 1. Мы продолжим этот процесс до тех пор, пока число $y = \sqrt{x^2 - n}$ не станет целым или не будет выполняться равенство $x = \frac{n+1}{2}$. Заметим, что в нашем примере $\frac{n+1}{2} = 671064$. Значения переменных x и y после завершения каждого цикла приведены в таблице.

п/п	x	$y = \sqrt{x^2 - n}$
1	1159	33.97…
2	1160	58.93…
3	1161	76.11…
4	1162	90.09…
5	1163	102.18…
6	1164	113

Таким образом, на шестом цикле мы получили целое число. Значит, искомые числа $x = 1164$ и $y = 113$. Соответствующие множители равны $x + y = 1277, x - y = 1051$.

2.5. Доказательство корректности алгоритма Ферма

Теперь мы должны доказать, что алгоритм Ферма выполняет свою задачу и что он всегда завершает работу. При доказательстве удобно разделить случай составного и случай простого числа n на входе. В первом случае мы должны показать, что существует натуральное число x, такое что $\lfloor\sqrt{n}\rfloor \le x < \frac{n+1}{2}$ и $\sqrt{x^2 - n}$ — целое число. Это означает, что если n составное, то алгоритм всегда находит делитель, меньший x, прежде, чем x становится равным $\frac{n+1}{2}$. А если n простое, то мы должны проверить, что число $\sqrt{x^2 - n}$ не может быть целым при $x < \frac{n+1}{2}$.

Предположим, что n можно представить в виде произведения $n = a \cdot b$, где $a \le b$. Мы хотим найти такие целые числа x и y, чтобы выполнялось $n = x^2 - y^2$. Другими словами,

$$n = a \cdot b = (x - y) \cdot (x + y) = x^2 - y^2.$$

Поскольку $x - y \le x + y$, разумно положить $a = x - y$ и $b = x + y$. Решая эту систему уравнений относительно двух неизвестных, получаем:

$$x = \frac{b + a}{2}, y = \frac{b - a}{2}.$$

Действительно, непосредственные вычисления показывают, что

$$\left(\frac{b + a}{2}\right)^2 - \left(\frac{b - a}{2}\right)^2 = a \cdot b = n. \qquad (2.2)$$

Отметим, что x и y должны быть целыми, а это значит, что оба числа $b + a$ и $b - a$ должны быть четными. Именно поэтому мы и требуем, чтобы n было нечетным; тогда каждое из чисел a и b, будучи делителем n, также нечетно, а значит, и $b + a$, и $b - a$ - чётные. Если число n четное, то алгоритм может работать неправильно. Если, например, $n = 2 \cdot k$ для нечётного k, то алгоритм никогда не завершит свою работу.

Если n простое, то единственные возможные значения a и b — это $a = 1, b = n$. Таким образом, $x = \frac{n+1}{2}$, и это наименьшее значение x, для которого число $\sqrt{x^2 - n}$ целое.

Рассмотрим теперь, что происходит, если n составное. Если $a = b$, то алгоритм находит делитель на шаге 1. Значит, мы можем предполагать, что n составное и не является полным квадратом. Другими словами, $1 < a < b < n$. Мы утверждаем, что тогда алгоритм останавливается, так как

$$\lfloor\sqrt{n}\rfloor < \frac{a+b}{2} < \frac{n+1}{2}. \qquad (2.3)$$

Начнём с доказательства неравенств. Правая часть неравенства означает, что $a + b < n + 1$. Заменяя n на $a \cdot b$ и вычитая из обеих частей $b + 1$, мы приходим к неравенству $a - 1 < a \cdot b - b$. Однако $a > 1$, поэтому обе части неравенства можно разделить на $a - 1$. В результате мы получаем $1 < b$. Это рассуждение показывает, что неравенство $1 < b$ эквивалентно неравенству $a + b < n + 1$. Поскольку $1 < a < b$ по предположению, мы показали, что $\frac{a+b}{2} < \frac{n+1}{2}$.

Рассмотрим теперь левую часть неравенства (2.3). Отметим для начала, что, поскольку $\lfloor\sqrt{n}\rfloor \leq \sqrt{n}$, нам достаточно доказать, что $\sqrt{n} \leq \frac{a+b}{2}$. Ясно, что последнее неравенство выполняется тогда и только тогда, когда $n \leq \frac{(a+b)^2}{4}$. Однако формула (2.2) даёт

$$\frac{(b+a)^2}{4} - n = \frac{(b-a)^2}{4}$$

и правая часть неотрицательна. Тем самым мы доказали неравенство $\frac{(b+a)^2}{4} - n \geq 0$, равносильное исходному.

Вернёмся к алгоритму. Напомним, что значение переменной x вначале равно $\lfloor\sqrt{n}\rfloor$, а затем оно увеличивается на 1 при каждом выполнении цикла. Поэтому из неравенства (2.3) следует, что, если число n составное, то алгоритм дойдёт до $\frac{a+b}{2}$ раньше, чем до $\frac{n+1}{2}$. Однако при $x = \frac{a+b}{2}$ получаем:

$$y^2 = \frac{(b+a)^2}{4} - n = \frac{(b-a)^2}{4}.$$

Таким образом, достигнув этого значения, алгоритм завершит работу, а его вывод будет состоять из множителей a и b. Поэтому, если n составное, то алгоритм всегда останавливается при некотором $x < \frac{n+1}{2}$, вычислив два делителя этого числа.

Заметим, что данное составное число n можно представить в виде $n = a \cdot b$, где $1 < a < b < n$, возможно, несколькими различными способами. Какое из разложений находит алгоритм Ферма? Поиск начинается при $x = \lfloor \sqrt{n} \rfloor$, и x увеличивается на каждом шаге. Значит, найденные алгоритмом делители a и b таковы, что разность

$$\frac{a+b}{2} - \lfloor \sqrt{n} \rfloor -$$

наименьшая из возможных. Теорема доказана.

Замечание. Алгоритм Ферма подсказывает правила выбора двух простых множителей числа n. Если нам удастся разложить n, то шифр будет взломан. Алгоритм деления методом проб мог бы создать иллюзию, что, выбрав большие простые множители, мы могли бы добиться того, что n сложно разложить. Однако это не так. Если множители большие, но их разность мала, то n очень легко разложить на множители алгоритмом Ферма.

2.6. Фундаментальное свойство простых чисел

Для доказательства единственности разложения целого числа в произведение простых нам понадобится следующее фундаментальное свойство простых чисел. Это свойство доказывается в данном разделе, а последующие разделы посвящены некоторым его приложениям. Начнём с леммы, которая послужит первым применением расширенного алгоритма Эвклида.

Лемма 2.1. *Пусть a, b и c натуральные числа, причем a и b взаимно просты. Тогда:*

1) если произведение $a \cdot c$ делится на b, то c делится на b;

2) если c делится на a и на b, то c делится на $a \cdot b$.

Доказательство. Докажем сначала утверждение (1). По предположению, числа a и b взаимно просты, т.е. $НОД(a, b) = 1$. Тогда результатом работы расширенного алгоритма Эвклида служат такие числа α и β, что

$$\alpha \cdot a + \beta \cdot b = 1.$$

Умножив обе части последнего равенства на c, получим:

$$\alpha \cdot a \cdot c + \beta \cdot b \cdot c = c.$$

Второе слагаемое в левой части заведомо делится на b, но то же справедливо и для первого слагаемого. Действительно, оно делится на $a \cdot c$, а поэтому, по предположению, и на b. Таким образом, вся сумма в левой части делится на b, а поскольку она равна c, то первое утверждение доказано.

Выведем теперь (2) из утверждения (1). Раз c делится на a, то существует такое натуральное t, что $c = a \cdot t$. Однако c делится и на b, и поэтому из (1) следует, что t делится на b, так как a и b взаимно просты. Значит, $t = b \cdot k$ для некоторого целого k. Поэтому число

$$c = a \cdot t = a \cdot (b \cdot k) = (a \cdot b) \cdot k$$

делится на $a \cdot b$, что и утверждалось в пункте. Лемма доказана.

Эта лемма будет использоваться очень часто, начиная с доказательства следующего свойства простых чисел, которое сформулировано в виде предложения 30 книги VII Эвклидовых «Начал». Это свойство настолько важное, что его называют фундаментальным свойством простых чисел.

Теорема 2.2 (фундаментальное свойство простых чисел). *Если произведение натуральных чисел $a \cdot b$ делится на простое число p, то либо a делится на p, либо b делится на p.*

Доказательство. Фундаментальное свойство доказывается с помощью леммы 2.1. По предположению, $a \cdot b$ делится на p. Если a делится на p, то доказательство завершено. Предположим, что a не делится на p. Поскольку число p простое, это означает, что $\text{НОД}(a, p) = 1$. Из первого утверждения леммы 2.1 следует ($a \cdot b$ делится на p, но a и p взаимно просты), что b делится на p. Теорема доказана.

2.7. Иррациональные числа

В этом разделе мы рассматриваем одно из приложений фундаментального свойства простых чисел, доказанного в разделе 2.6.

Теорема 2.3. *Если число p простое, то число $\sqrt{p}$ иррациональное, то есть не представимо в виде рациональной дроби.*

Доказательство. Воспользуемся доказательством от противного. Допустим, что $\sqrt{p}$ является рациональной дробью. Это значит, что существуют натуральные числа a и b такие, что

$$\sqrt{p} = \frac{a}{b}. \quad (2.4)$$

Более того, можно считать, что дробь записана в приведённом виде, т.е. $НОД(a, b) = 1$. В таком виде можно записать каждую дробь: для этого достаточно произвольную дробь сократить на наибольший общий делитель числителя и знаменателя. Возведём обе части равенства (2.4) в квадрат. Получим:

$$p = \frac{a^2}{b^2} \Leftrightarrow a^2 = p \cdot b^2. \quad (2.5)$$

Значит, a^2 делится на p. Согласно фундаментальному свойству простых чисел, это означает, что a делится на p. Поэтому существует такое число c, что $a = p \cdot c$. Подставляя последнее выражение в (2.5), получаем:

$$b^2 \cdot p = p^2 \cdot \mathrm{c}^2.$$

Сокращая на p, мы заключаем, что b^2 должно делиться на p. Повторное использование фундаментального свойства простых чисел даёт, что b делится на p. Значит, и a, и b делятся на p. Однако это невозможно, поскольку $НОД(a, b) = 1$. Тем самым мы пришли к противоречию, и $\sqrt{p}$ не может быть рациональной дробью. Теорема доказана.

2.8. Единственность разложения

Обоснуем единственность представления натурального числа в виде, указанном в разделе 2.1.

Воспользуемся доказательством от противного, допустив существование натуральных чисел, больших 2, допускающих более одного разложения: n — наименьшее натуральное число, у которого есть, по крайней мере, два различных разложения

$$n = p_1^{e_1} \cdot p_2^{e_2} \cdot \dots \cdot p_k^{e_k} = q_1^{r_1} \cdot q_2^{r_2} \cdot \dots \cdot q_s^{r_s}, \quad (2.6)$$

где $p_1, p_2, \dots, p_k, q_1, q_2, \dots, q_s$ — простые, а $e_1, e_2, \dots, e_k, r_1, r_2, \dots, r_s$ — натуральные числа. Кроме того, мы предполагаем, что эти два разложения различны. Заметим, что такое может произойти по двум причинам. Во-первых, в одном из разложений могут присутствовать простые множители, которых нет в другом. Во-вторых, даже если набор простых множителей в обоих случаях одинаков, их кратности могут быть различными (2.6).

Из левой части разложения (2.6), следует, что n делится на p_1. Но $n = q_1^{r_1} \cdot q_2^{r_2} \cdot \cdots \cdot q_s^{r_s}$. Многократное применение фундаментального свойства простых чисел показывает, что один из сомножителей в $q_1^{r_1} \cdot q_2^{r_2} \cdot \cdots \cdot q_s^{r_s}$ должен делиться на p_1, а значит, одно из чисел q_t должно делиться на p_1. Но простое число делится на другое простое, только, если они равны. Значит, $p_1 = q_j$ для некоторого $j, 1 \le j \le s$. Поэтому в правой части разложения числа n (2.6) заменим q_j на p_1:

$$n = p_1^{e_1} \cdot p_2^{e_2} \cdot \cdots \cdot p_k^{e_k} = q_1^{r_1} \cdot q_2^{r_2} \cdot \ldots \cdot q_j^{r_j} \cdot \ldots \cdot q_s^{r_s} = \\ = q_1^{r_1} \cdot q_2^{r_2} \cdot \ldots \cdot p_1^{r_j} \cdot \ldots \cdot q_s^{r_s}.$$

Теперь на p_1 можно сократить, поскольку оно входит в оба разложения в положительной степени. В результате получим:

$$p_1^{e_1-1} \cdot p_2^{e_2} \cdot \cdots \cdot p_k^{e_k} = q_1^{r_1} \cdot q_2^{r_2} \cdot \ldots \cdot p_1^{r_j-1} \cdot \ldots \cdot q_s^{r_s},$$

т.е. два разложения на множители некоторого нового натурального числа, которое мы обозначим через m. Однако эти разложения не могут быть различными. Действительно, мы выбрали в качестве n наименьшее положительное число с двумя различными разложениями, а $m = \frac{n}{p_1} < n$. Раз полученные разложения совпадают, то $j = 1$, т.е. $p_1 = q_1$, а кроме того, $k = s$. Далее,

$$p_2 = q_2, p_3 = q_3, \ldots, p_k = q_k$$

и кратности каждого простого числа в обоих разложениях равны

$$e_1 - 1 = r_1 - 1, e_2 = r_2, \ldots, e_k = e_k.$$

Однако из этих равенств вытекает, что разложения в (2.6) тождественны и мы пришли к противоречию. Значит, представление натурального числа в виде, указанном в теореме 2.1, единственно. Теорема доказана полностью.

3. Простые числа

В первых двух разделах были рассмотрены некоторые свойства простых чисел. Для безопасной реализации RSA необходимо подбирать большие простые числа. Приступаем к решению этой задачи. Решето Эратосфена — старейший из известных методов нахождения простых чисел.

3.1. Полиномиальная формула

Представляет интерес поиск функции (правила) получения простых чисел («формула простых чисел»).

Определение 3.1. *Функция $f: N \to N$ называется формулой простых чисел, если $f(m)$ — простое число для каждого $m \in \mathbb{Z}$.*

Это определение слишком общее. Вместо «формулы простых чисел» будем искать функции, среди значений которых часто встречаются простые числа. Поскольку многочлен — простейшая из возможных функций, стоит начать с вопроса: существует ли полиномиальная формула простых чисел? Как следует из определения, данного выше, многочлен

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

с целыми коэффициентами $a_n, a_{n-1}, \cdots, a_1, a_0$ является формулой простых чисел, если число $f(m)$ простое для каждого положительного целого m. Рассмотрим с многочлен $f(x) = x^2 + 1$. Результат представлен в следующей таблице.

x	F(x)	Простое	x	F(x)	Простое
1	2	да	6	37	да
2	5	да	7	50	нет
3	10	нет	8	65	нет
4	17	да	9	82	нет
5	26	нет	10	101	да

Заметим, что, если x нечётно, то $f(x)$ чётно. Таким образом, $f(x)$ — всегда чётное, а значит, и составное число для нечётных значений x (за исключением $x = 1$, так как $f(1) = 2$).

Значит, если $x > 1$ и $f(x)$ — простое, то x обязательно чётное. Следовательно, если бы числа $f(x)$ были простыми при каждом чётном x, то многочлен $f(2x)$ был бы формулой простых чисел. К сожалению, это не так; например, $f(8) = 65$ — составное число. Итак, многочлен $f(x) = x^2 + 1$ не даёт формулу простых чисел в смысле данного определения. Такая ситуация не исключение.

Теорема 3.1. *Для данного многочлена f(x) существует бесконечно много положительных целых чисел m, при которых число $f(m)$ составное.*

Доказательство. Проведем доказательство только для многочленов степени 2. Общий случай доказывается аналогично.

Пусть $f(x) = ax^2 + bx + c$ — многочлен с целыми коэффициентами a, b, c. Мы можем предполагать, что $a > 0$. Это означает, что $f(x)$ положителен для достаточно больших значений переменной x. Допустим, что существует натуральное число m, что $p = f(m)$ — простое.

Пусть h — произвольное натуральное число. Вычислим значение $f(m + hp) = f(m + hp) = a(m + hp)^2 + b(m + hp) + c$.
Раскрывая квадрат и собирая вместе члены, содержащие p, получаем:

$$f(m + hp) = (am^2 + bm + c) + p(amh + aph^2 + bh).$$

Заметим, что выражение в первой скобке равно $f(m) = p$, так что

$$f(m + hp) = p(1 + 2amh + aph^2 + bh). \qquad (3.1)$$

Формула (3.1) показывает, что число $f(m + hp)$ разложимо. Действительно, оно равно произведению p и натурального числа большего единицы. То есть необходимо найти значения h, при которых

$$1 + 2amh + aph^2 + bh > 1.$$

Последнее неравенство эквивалентно

$$2amh + aph^2 + bh > 0.$$

Поскольку число h положительно (по предположению), требуемое неравенство выполнено только, если справедливо

$$2amh + aph^2 + bh > 0 \Leftrightarrow h > \frac{-b - 2am}{ap}.$$

Заметим, что $-b - 2am$ является положительным числом в том случае, если b отрицательно и меньше чем $-2am$. Теорема доказана.

Смысл теоремы состоит в том, что, если $f(x) = ax^2 + bx + c$ — многочлен с целыми коэффициентами и положительным a, и число - $f(m) = p$ — простое, то $f(m + hp)$ является составным для всех $h > \frac{-b-2am}{ap}$. В частности, найдётся бесконечно много натуральных значений x, при которых $f(x)$ — составное число. Доказанная теорема означает, что не существует многочлена одного переменного, все значения которого простые числа. Однако есть многочлены нескольких переменных, все положительные значения которых — простые числа.

3.2. Экспоненциальные формулы. Числа Мерсенна

Существуют две экспоненциальные формулы огромной исторической важности. Обе изучались математиками XVII и XVIII веков, в особенности Ферма и Эйлером. Вот эти формулы:

$$M(n) = 2^n - 1, F(n) = 2^{2^n} + 1, \qquad (3.2)$$

где n — натуральное. Числа из первой формулы называются числами Мерсенна, а из второй — числами Ферма.

Вопрос о том, при каких значениях n числа Мерсенна просты, восходит к математикам античной Греции. В пифагорейском мистицизме число называлось совершенным, если оно равняется полусумме своих положительных делителей. Например, делители числа 6 — это 1, 2, 3 и 6. Складывая их, получаем:

$$1 + 2 + 3 + 6 = 12 = 2 \cdot 6.$$

Следовательно, 6 — совершенное число. Конечно, никакое простое число не будет совершенным. Действительно, делители простого числа p — это 1 и р, и $1 + p < 2p$, поскольку $p > 1$.

Эвклид знал, что число $2^{n-1}(2^n - 1)$ совершенно, если $2^n - 1$ простое. Нетрудно показать, что все чётные совершенные числа имеют такой вид, но доказан этот факт был только Эйлером в восемнадцатом веке. Формула Эвклида сводит задачу о поиске чётных совершенных чисел к нахождению простых чисел Мерсенна: Марэн Мерсенн - священник и математик-любитель семнадцатого века. Числа вида $2^n - 1$ обязаны своим именем утверждению Мерсенна о том, что они просты, в случае

$$n = 2;\ 3;\ 5;\ 7;\ 13;\ 17;\ 19;\ 31;\ 67;\ 127;\ 257,$$

и являются составными для всех остальных 44 положительных простых n, меньших 257.

Первое важное замечание: Мерсенн рассматривал значения функции $2^n - 1$ только при простых n. Действительно, если n — составное, то такое же и $M(n)$. Предположим, что $n = rs$, $1 < r$ и $s < n$, тогда

$$\mathrm{M}(n) = 2^n - 1 = 2^{rs} - 1 = \\ = (2^r - 1)\left(2^{r(s-1)} + 2^{r(s-2)} + \cdots + 2^r + 1\right).$$

Следовательно, если r делит n, то $M(r)$ делит $M(n)$. Второй важный момент заключается в том, что обратное неверно. Иначе говоря, если n — простое, то $M(n)$ не обязано быть простым. Мы

видим из списка Мерсенна, что $M(11)$ должно быть составным. Это легко проверить:

$$M(11) = 2047 = 23 \cdot 89.$$

Как часто бывало, Мерсенн не привёл доказательства своего утверждения, что дало повод для сомнений в его истинности и оставило широкое поле деятельности для математиков. В поисках простых чисел Мерсенна принял участие и Эйлер. В 1732 году он нашёл два «новых простых» числа: $M(67)$ и $M(257)$, отсутствовавших в списке Мерсенна. Позже выяснилось, что в этом случае Эйлер был не прав. Сейчас известно, что, кроме $M(61)$, в списке пропущены простые числа $M(89)$ и $M(107)$ и присутствуют составные числа $M(67)$ и (257).

При доказательстве простоты чисел Мерсенна Ферма использовал метод разложения на множители, который будет описан далее. В наше время используется гораздо более эффективный тест Люка - Лемера. С помощью этого теста в 1998 году было показано, что число Мерсенна $M(3021377)$ простое. Оно состоит из 1819 050 знаков и является наибольшим из простых чисел, известных на конец двадцатого века.

3.3. Экспоненциальная формула. Числа Ферма

История чисел Ферма очень схожа с историей чисел Мерсенна. Ферма знал, что, если $2^m + 1$ простое, то m должно быть степенью двойки. Поэтому, интересуясь простыми, необходимо смотреть только на числа вида $2^{2^n} + 1$. В письме, адресованном шевалье Френикл (Prenicle), другому математику-любителю, Ферма выписал эти числа для $n = 0, 1, \ldots, 6$:

$$3; 5; 17; 257; 65537; 4294967297; \ 18446744073709551617.$$

Затем он предположил, что все числа вида $2^{2^n} + 1$ простые. Как ни странно, Ферма, кажется, не пытался разлагать на множители эти числа методом, аналогичным использованному им для разложения чисел Мерсенна. Если бы он применил этот метод, то увидел бы, что число $F(5)$ составное. Необходимую проверку произвел Эйлер.

В отличие от чисел Мерсенна, о которых известно, что они являются богатым источником больших простых чисел, среди чисел Ферма простых известно очень мало. Фактически все известные простые числа Ферма — это $F(0), \ldots, F(4)$ — список, неизменный со времён Ферма. Конечно, вычислять числа Ферма при «больших»

значениях n очень трудно. В конце концов, формула, описывающая эти числа, — двойная экспонента, т.е. экспонента от экспоненты.

3.4. Прайморіальная формула

Напомним, что факториалом натурального числа n называется произведение всех положительных натуральных чисел, не превосходящих n.

Определение 3.2. *Праймориал $p^{\#}$ простого $p > 0$ - произведение всех простых чисел, меньших или равных p.*

Например, $2^{\#} = 2$ и $5^{\#} = 2 \cdot 3 \cdot 5 = 30$. Заметим, что, если q — следующее после p простое число, то

$$q^{\#} = p^{\#} \cdot q.$$

Рассмотрим числа вида $p^{\#} + 1$, которые сведем в таблицу:

p	$p^{\#}$	$p^{\#}+1$	p	$p^{\#}$	$p^{\#}+1$
1	2	3	4	5	6
2	2	3	7	210	211
3	6	7	11	2310	2311
5	30	31			

Все числа в третьей и шестой графах таблицы — простые. Может ли это быть просто совпадением? Если этот вопрос вселяет в вас надежду на то, что все числа вида $p^{\#} + 1$ простые, вам бы следовало попытаться заполнить следующую строку таблицы. На самом деле

$$13^{\#} + 1 = 30031 = 59 \cdot 509$$

является составным числом. Справедливо следующее утверждение.

Теорема 3.2. *Натуральное число $p^{\#} + 1$ не имеет делителей, меньших или равных p.*

Доказательство. Допустим противное. Пусть $q \leq p$ — простой делитель числа $p^{\#} + 1$. Поскольку $p^{\#}$ — произведение всех простых чисел вплоть до p, q должно также делить $p^{\#}$. Следовательно, q делит разность

$$\left(\mathrm{p}^{\#} + 1\right) - \mathrm{p}^{\#} = 1.$$

Значит, $q = 1$, что противоречит его простоте. В итоге мы получаем: наименьший делитель числа $p^{\#} + 1$ должен быть больше p. Теорема доказана.

Теорема 3.2 позволяет реализовать следующий алгоритм построения больших простых чисел. Пусть известны все простые

числа вплоть до числа p. Вычисляем $p^{\#}+1$. Если оно простое, то найдено следующее простое число. Если нет, то найдём его наименьший простой делитель; он должен быть больше, чем p. В любом случае, мы нашли простое число, большее p.

Описанный подход плох по нескольким причинам. Наиболее очевидная из них — необходимость разложения на простые множители числа $p^{\#}+1$. Даже для сравнительно небольших значений p прайморіал $p^{\#}$ огромен. С другой стороны, число $p^{\#}+1$ может оказаться простым. Далее покажем, что существуют способы проверки простоты больших чисел, которые не используют разложения на множители. Простое число, представимое в форме $p^{\#}+1$, называется прайморіально простым.

Конечно, остаётся подход к проверке простоты, заключающийся в систематических попытках подобрать подходящий делитель числа. Как мы видели в разделе 3.3, этот алгоритм весьма неэффективен. Специальный алгоритм тестирования простоты чисел вида $p^{\#}+1$ будет изучаться далее. Несмотря на то, что он очень удобен, пока найдено только 16 прайморіально простых чисел, наибольшее из которых соответствует $p=24027$ и насчитывает 10387 знаков.

3.5. Бесконечность множества простых чисел

Истинная причина, по которой мы так долго и детально разбирались с прайморіальной формулой, состоит в том, что она даёт доказательство следующего фундаментального результата.

Теорема 3.3. *Простых чисел бесконечно много.*

Доказательство (Эвклид). Предположим, множество простых чисел конечно. Это означает, что существует наибольшее простое число, скажем, p. Другими словами, мы предполагаем, что все числа, большие p, — составные. Однако из теоремы 3.2 следует, что число $p^{\#}+1$ не может иметь простых делителей, меньших или равных p. Отсюда получаем, что у числа $p^{\#}+1$ нет простых делителей, т.е. оно само является простым числом. А поскольку $p^{\#}+1>p$, мы получаем противоречие с предположением: p — наибольшее простое число. Итак, простых чисел должно существовать бесконечно много. Теорема доказана.

Было найдено много других доказательств бесконечности ряда простых чисел. Доказательство Эйлера (1737 года) носит особый характер. Оно оказалось тем подходом, из которого были получены многие результаты в теории чисел. Подобно доказательству Эвклида здесь также применен метод от противного.

Доказательство (Эйлер). Допустим, что множество простых чисел конечно, и пусть p — наибольшее из них. Рассмотрим произведение множителей

$$P = \frac{1}{1-\frac{1}{2}} \cdot \frac{1}{1-\frac{1}{3}} \cdot \frac{1}{1-\frac{1}{5}} \cdot \ldots \cdot \frac{1}{1-\frac{1}{p}} \quad (3.3)$$

по одному для каждого простого числа. Это произведение равно некоторому положительному вещественному числу. Более того, аккуратно перемножая члены произведения, можно показать, что

$$P = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots, \quad (3.4)$$

где теперь у нас есть своё слагаемое для каждого натурального числа. Такое представление числа (3.3) в виде (3.4) будет объяснено далее.

Нетрудно убедиться, что сумма (3.4), соответствующая P, не может быть равной ни одному вещественному числу [5]. Заметим, что справедливы неравенства

$$\frac{1}{3} + \frac{1}{4} \geq 2 \cdot \frac{1}{4} = \frac{1}{2};$$

$$\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \geq 4 \cdot \frac{1}{8} = \frac{1}{2};$$

$$\frac{1}{2^{n-1}+1} + \frac{1}{2^{n-1}+2} + \cdots + \frac{1}{2^n} \geq 2^{n-1} \cdot \frac{1}{2^n} = \frac{1}{2}.$$

Следовательно,

$$P = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots > 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{\mathrm{n}}} + \cdots \geq \mathrm{n} \cdot \frac{1}{2} = \frac{n}{2}$$

для любого данного натурального n. Таким образом, P больше любого наперед заданного числа, поэтому оно не может быть каким-то вещественным числом.

Окончательно имеем, что число P (3.3) - конечно, а, с другой стороны, P (3.4) - не является конечным числом. Полученное противоречие доказывает, что простых чисел бесконечно много. Теперь

вернёмся к доказательству равенства (3.4). Фактически нам нужно показать, что произведение (3.3) совпадает с бесконечной суммой (3.4). Дробь $\frac{1}{1-q}$ – [при $q \in (0,1)$] можно интерпретировать как сумму бесконечной геометрической прогрессии со знаменателем q:

$$\frac{1}{1-q} = 1 + q + q^2 + q^3 + q^4 + \cdots + q^n + \cdots.$$

Таким образом, равенство (3.3) можно переписать в виде:

$$\mathrm{P} = \left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right)\left(1 + \frac{1}{3} + \frac{1}{3^2} + \cdots\right) \ldots \left(1 + \frac{1}{\mathrm{p}} + \frac{1}{\mathrm{p}^2} + \cdots\right).$$

Раскроем скобки в этом произведении. Такую операцию можно сделать двумя способами. Один из них заключается в последовательном перемножении скобок: сначала раскрываем произведение первых двух скобок, затем полученный результат умножаем на третью и т.д. Это привычный способ, но длинный и малоэффективный. Действительно, мы не знаем точное количество перемножаемых скобок, да и каждая скобка представляет собой бесконечную сумму.

Другой способ, менее привычный, но более грамотный, основан на простом наблюдении: если аккуратно раскрыть скобки и учесть структуру сомножителей, то получили бы бесконечную сумму произведений вида

$$\frac{1}{2^{r_1}3^{r_2}5^{r_3}\cdots p^{r_s}} (r_i \geq 0).$$

По одному сомножителю из каждой скобки (значение показателя $r_i = 0$ говорит о том, что из i-й скобки в качестве сомножителя мы берем 1). Итак, число P из равенства (3.3) представляет собой бесконечную сумму слагаемых вида

$$\frac{1}{2^{r_1}3^{r_2}5^{r_3}\cdots p^{r_s}}, r_1 \geq 0, r_2 \geq 0, \ldots, r_s \geq 0. \qquad (3.5)$$

Упорядочим эти слагаемые по убыванию. Самым большим будет член $\frac{1}{2^0 3^0 5^0 \cdots p^0} = 1$, затем идёт $\frac{1}{2^1 3^0 5^0 \cdots p^0} = \frac{1}{2}$, потом $\frac{1}{2^0 3^1 5^0 \cdots p^0} = \frac{1}{3}$, после него — $\frac{1}{2^2 3^0 5^0 \cdots p^0} = \frac{1}{4}$ и т.д. Как видите, мы получили начало бесконечной суммы (3.4). Покажем, что число P совпадает с этой суммой. Для этого необходимо проверить две вещи:

1) для любого натурального числа n среди дробей (3.5) найдётся такая, что

$$\frac{1}{n} = \frac{1}{2^{r_1}3^{r_2}5^{r\_3}\cdots p^{r\_s}};$$

2) каждая дробь вида $\frac{1}{n}$ встречается в бесконечной сумме, представляющей P, только один раз.

Начнём с доказательства второго утверждения. Предположим, что нашлись две разные дроби, равные $\frac{1}{n}$.

$$\frac{1}{n} = \frac{1}{2^{r_1}3^{r_2}5^{r\_3}\cdots p^{r\_s}} = \frac{1}{2^{k_1}3^{k_2}5^{k\_3}\cdots p^{k_s}}.$$

Поскольку числители всех этих дробей равны 1, мы получаем равенство знаменателей:

$$n = 2^{r_1}3^{r_2}5^{r\_3}\cdots p^{r_s} = 2^{k_1}3^{k_2}5^{k_3}\cdots p^{k_s}.$$

Как первое произведение в этом равенстве, так и второе целиком состоит из простых чисел, т.е. мы двумя разными способами разложили натуральное число n на простые сомножители, что противоречит основной теореме арифметики 2.1. Значит, $r_1 = k_1, r_2 = = k_2, \dots, r_s = k_s$, т.е. дроби совпадали изначально. Таким образом, утверждение (2) доказано.

Перейдём к доказательству утверждения (1). По основной теореме арифметики любое натуральное число n представляется в виде произведения простых сомножителей: $n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_m^{\alpha_m}$. С другой стороны, знаменатель каждой из дробей (3.5) — это произведение степеней всех простых чисел: $2^{r_1} \cdot 3^{r_2} \cdots p^{r_s}$, причем их показателями могут быть любые неотрицательные целые числа. Таким образом, чтобы приравнять дробь вида (3.5) к числу n, достаточно выбрать подходящие показатели r_i. Докажем, что следующие дроби могут быть равными

$$\frac{1}{p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_m^{\alpha_m}} = \frac{1}{2^{r_1} \cdot 3^{r_2} \cdots p^{r_s}},$$

если в первой из них присутствуют только некоторые простые числа, в то время как во второй — все. Напомним, что показатели степеней r_i могут обращаться в нуль. Поэтому возьмём нулевые показатели для всех тех простых чисел, которых нет в знаменателе левой дроби. Итак, первое утверждение доказано. То есть мы доказали, что произведение (3.3) и бесконечная сумма (3.4) равны одному и тому же числу P. Теорема доказана.

Замечание. При доказательстве теоремы предполагалось, что результат суммирования бесконечных сумм положительных чисел не зависит от порядка суммирования (теорема Римана [5]).

Тот факт, что простых чисел бесконечно много, ставит много интересных проблем. Например, что можно сказать об их распределении? Растёт или убывает «плотность» простых чисел, когда мы переходим ко все большим и большим числам? Существует ли возможность измерить эту «плотность»? Наилучший способ точно сформулировать проблему о распределении простых чисел состоит в использовании $\pi(x)$-функции, значение которой равно количеству простых чисел, не превосходящих x.

Оценка функции $\pi(x)$ — важная задача теории чисел. Справедливо предельное равенство

$$\lim_{x\to\infty} \frac{\pi(x)\cdot \ln x}{x} = 1,$$

из которого следует, что выполняется $\pi(x) \approx \frac{x}{\ln x}$ при достаточно большой величине x. Например, если $x = 10^{16}$, то разность

$$\pi(x) - \left[\frac{x}{\ln x}\right] = 7804289\,844393$$

будет величиной порядка 10^{13}. Поскольку в этом случае $\frac{x}{\ln x}$ имеет порядок 10^{14}, то ошибка действительно значительна. Есть много других простых функций, дающих неплохую аппроксимацию функции $\pi(x)$ при больших x [1, 2].

3.6. Решето Эратосфена

Решето Эратосфена — это старейший из известных способов получения простых чисел. Решето или правило получило своё имя потому, что, когда оно применяется к списку натуральных чисел, составные числа просеиваются, а простые задерживаются.

Прежде всего, цель решета — определить все положительные простые числа, меньшие некоторой верхней границы $n > 0$, которую мы предполагаем целой. Чтобы использовать метод решета, как это делал Эратосфен (имея только карандаш и бумагу), мы поступаем следующим образом. Сначала выписываем все нечётные целые между 3 и n. Причина, по которой в списке нет четных, заключается в том, что, кроме 2, среди них нет простых чисел.

Теперь мы начинаем просеивать список: исключать составные числа. Первое число в нем 3. Начиная со следующего числа в списке, а это число 5, мы вычёркиваем из него каждое третье число. Проделав это до конца, мы вычеркнем все числа из списка, кратные 3 и большие самой тройки.

Теперь выберем наименьшее число из списка, превосходящее 3, которое ещё не было вычеркнуто. Таким будет 5, а следующее за ним число — 7. Вычёркиваем каждое пятое число из нашего списка, начиная с 7. Таким образом, все числа, кратные 5, будут вычеркнуты. Продолжаем эту процедуру, пока не дойдём до n. Заметим, что, если мы собираемся вычеркивать каждое p-е число, то всегда надо начинать отсчёт с числа $p + 2$, даже когда это число было вычеркнуто на предыдущих шагах. Например, для $n = 41$ список нечетных чисел выглядит так:

3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37,39,41.

Вычеркнув каждое третье число начиная с 5, мы получим:

3,5,7,~~9~~,11,13,~~15~~,17,19,~~21~~,23,25,~~27~~,29,31,~~33~~,35,37,~~39~~,41.

Теперь мы вычёркиваем каждое пятое число, начиная с 7, что даёт:

3, 5, 7,~~9~~,11,13,~~15~~,17,19,~~21~~,23,~~25~~,~~27~~,29,31,~~33~~,~~35~~,37,~~39~~,41.

Мы должны бы теперь вычеркнуть каждое седьмое число, начиная с 9. Но если мы это сделаем, то никакие новые числа не отсеются. Далее нам нужно бы вычеркнуть каждое одиннадцатое число, начиная с 13, но это опять не даст никакого эффекта. На самом деле ни одно из чисел, оставшихся в списке после вычёркивания каждого пятого, не будет позже вычеркнуто ни на каком этапе просеивания. Положительные нечётные простые числа, не превосходящие 41, это

3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41.

В приведённом примере есть пара важных обстоятельств, которые необходимо учесть. Во-первых, хотя нам следовало повторять процесс вычёркивания вплоть до граничного числа n (41 в нашем примере), мы исключили из всех составных чисел к тому моменту, когда отсеяли кратные 5, и все остальные просеивания оказались излишними. Во-вторых, некоторые числа вычёркивались больше чем один раз. Такое произошло, например, с 15. Первый раз оно было вычеркнуто, когда мы отсеивали кратные 3. Но 15 также делится на 5, поэтому оно было вычеркнуто снова при отсеивании кратных 5.

Посмотрим, как можно увеличить эффективность решета с учетом этих двух замечаний. Начнём со второго из них, т.е. посмотрим, можем ли мы так все организовать, чтобы каждое число вычёркивалось только один раз. К сожалению, ответ на этот вопрос отрицателен: нет хорошего способа выполнить это. Но часть улучшений могут быть реализованы.

Предположим, что мы отсеиваем числа, кратные какому-то простому p. Учитывая наше описание решета, нам следовало бы вычеркивать каждое p-е число, начиная с $(p + 2)$ — следующего за p числа в списке. Простейшее усовершенствование — это начинать процесс вычеркивания не с $p + 2$, а с наименьшего числа, кратного p, которое не делится на простое число, меньшее p.

Найдём его. Положительные числа, кратные p, записываются в виде $k \cdot p$, где n — натуральное. Если $k < p$, то $k \cdot p$ делится на число, меньшее p, а именно на k. Значит, первое кратное p число, которое не делится на простое число, меньшее p, есть p^2. Так что, достаточно вычёркивать каждое p-е число, начиная с p^2. Однако необходимо обратить внимание на то, что даже после этого нововведения останутся числа, которые будут вычёркиваться не единожды.

Что касается другого замечания, можем ли мы закончить отсев раньше, чем дойдём до n-го шага? На этот раз ответ положителен, он следует из того, что мы только что сделали. Допустим, например, что мы вычёркиваем каждое p-е число. Как мы только что видели, первое число, подлежащее вычёркиванию, равно p^2. Но если $p^2 > n$, такого числа нет в списке. В итоге нам нужно вычёркивать каждое p-е до тех пор, пока $p \leq \sqrt{n}$. Поскольку p — целое, это равносильно тому, что $p \leq \lfloor \sqrt{n} \rfloor$. В примере, разобранном выше, $\lfloor \sqrt{41} \rfloor = 6$. Вот почему исключение кратных 3 и 5 было достаточно для вылавливания всех составных чисел из списка.

Обсудим компьютерную реализацию решета. Список нечётных чисел представляется одномерным массивом, который также принято называть вектором. Напомним, что с каждой ячейкой вектора ассоциируется два числа. Одно из них — значение ячейки, а другое идентифицирует её местоположение в векторе. Например, в векторе

(a *b* c *d e f g*)
↑

величина ячейки, отмеченной стрелкой, равна b, а ее индекс — двум, поскольку она стоит на втором месте в векторе.

Вернёмся к решету Эратосфена. Предположим, что мы хотим найти все простые числа, не превосходящие нечётного целого n. Сначала мы должны построить вектор с $\frac{n-1}{2}$ ячейками, по одной для каждого нечётного целого $2j + 1$. Ячейки будут принимать одно из двух возможных значений: 1 или 0. Если значение ячейки равно 0, то нечетное число, ею представленное, было вычеркнуто на каком-то предыдущем шаге процесса просеивания. Итак, начальное значение каждой ячейки равно 1, поскольку ещё нет никаких вычеркнутых чисел. Чтобы «вычеркнуть» число $2j + 1$, нужно заменить 1, стоящую в j-й ячейке вектора, на 0. Конечно, эта ячейка могла быть «вычеркнута» на предыдущем шаге решета. В этом случае ее значение уже равно 0 и не будет меняться в процессе выполнения следующих шагов алгоритма.

Теперь мы приведём алгоритм решета Эратосфена, описанный выше. Он содержит оба усовершенствования, которые мы обсуждали, т.е. каждое p-е число вычёркивается начиная с p^2, а алгоритм заканчивается, как только p превысит $\sqrt{n}$.

Алгоритм 3.1 (решето Эратосфена)

Ввод: нечётное натуральное n.

Вывод: список всех нечётных положительных простых чисел, меньших или равных n.

Шаг 1. Начинаем с создания вектора $\bar{v}$ с $\frac{n-1}{2}$ ячейками, каждой из которых присвоено значение 1, и полагаем $P = 3$.

Шаг 2. Если $P^2 > n$, выписываем все числа $2j + 1$, для которых значение j-й ячейки вектора равно 1, и останавливаемся; в противном случае переходим к шагу 3.

Шаг 3. Если значение ячейки вектора $\bar{v}$ с номером $\frac{P-1}{2}$ равно 0, увеличиваем P на 2 и возвращаемся к шагу 2; в противном случае переходим к шагу 4.

Шаг 4. Присваиваем новой переменной T значение P^2; замещаем нулём значение ячейки вектора $\bar{v}$ под номером $\frac{T-1}{2}$ и увеличиваем T на $2P$; повторяем эти два шага до тех пор, пока $T \leq n$, затем увеличиваем P на 2 и возвращаемся к шагу 2.

Заметим, что на последнем шаге мы увеличиваем T на 2P, а не на P, как можно было бы ожидать. Мы делаем это потому, что вектор $\bar{v}$ представляет набор нечетных чисел, так что как T, так и P нечётны. Поэтому, если мы вычёркиваем каждое p-е число, то число, которое будет вычеркнуто после T, есть $T + 2P$.

Может показаться, что существует простое изменение описанной выше процедуры, которое ускорит алгоритм. Способ, с помощью которого мы избавляемся от просмотренных составных чисел в векторе, состоит в замене 1, стоящей в соответствующей ячейке, на 0. Но почему, если мы не заботимся о составных числах, просто не исключить их из вектора?

К сожалению, мы не можем этого сделать. Неприятность в том, что метод, с помощью которого мы узнаем, кратно ли число, соответствующее данной ячейке, некоторому p, зависит от номера ячейки. Другими словами, числа, кратные p, встречаются в каждой p-й позиции вектора. Если мы удалим некоторые числа из списка, то алгоритм, который мы описали, перестанет работать.

Подобно всем алгоритмам, решето Эратосфена имеет ограничения. Например, оно неэффективно при поиске очень больших простых чисел. Напомним, однако, что цель этого алгоритма — поиск всех простых меньших, чем определённая верхняя граница. Ясно, что такой поиск невыполним, если граница слишком велика. Суммируя ограничения, накладываемые назначением алгоритма, отметим два его слабых места: решето требует существенного объема памяти компьютера и огромного числа повторений отдельных циклов программы. К его достоинствам можно отнести простоту реализации и, кроме того, нам не нужно вычислять отдельные делители.

4. Арифметика остатков

Большинство алгоритмов, описанных в предыдущих разделах, проверяют делимость чисел непосредственным делением, убеждаясь, что в остатке действительно получается 0. Однако сейчас построены более эффективные методы, одним из которых доказано, что $5 \cdot 2^{23473} + 1$ — множитель числа $F(23471)$. Так как эти числа слишком большие, мы полагаем, что даже проверка справедливости этого утверждения непосредственным делением займёт очень много времени. Можно ли оценить, насколько велико число $F(23471)$. Пользуясь логарифмами, легко показать, что в нем более чем 10^{7063}

знаков. То есть количество знаков в $F(23471)$ больше числа элементарных частиц в видимой части вселенной. Не приходится и говорить о проверке непосредственным делением того факта, что $5 \times \times 2^{23473} + 1$ — множитель числа $F(23471)$. Как же тогда это сделать?

Выход из этой дилеммы заключается в использовании арифметики остатков, являющейся темой данного раздела. При решении вопросов делимости техника арифметики остатков играет основную роль. Она полезна также при вычислениях, имеющих отношение к эффекту периодичности.

4.1. Отношение эквивалентности

Арифметику остатков лучше всего вводить с помощью отношения эквивалентности. Поскольку такие отношения будут играть важную роль как в этом разделе, так и далее, стоит подробно разобрать это основное понятие.

Пусть $\mathbb{X}$ — конечное или бесконечное множество. Отношением на $\mathbb{X}$ называется правило, по которому «сравниваются» его элементы. Заметим, что для определения отношения мы должны чётко задать само множество; другими словами, нам должно быть ясно, какие элементы нужно сравнивать.

Рассмотрим несколько примеров. На множестве целых чисел есть много простых отношений - вроде «равно», «не равно», «меньше, чем», «меньше или равно». На множестве цветных мячей у нас есть отношение «тот же цвет». Последний пример, ввиду своей конкретности, хорош для запоминания в качестве модельного случая. Кстати, мы предполагаем, что каждый мяч из множества окрашен только в один цвет, пёстрые мячи мы не рассматриваем.

Отношение эквивалентности — это отношение весьма специфичного вида. Возвращаясь к общим определениям, предположим, что $\mathbb{X}$ — множество, в котором было определено отношение. Для обозначения эквивалентности обычно употребляют значок «~».

Определение 4.1. *Отношение «~» назовем отношением эквивалентности, если для всех $x, y, z \in \mathbb{X}$ выполнены свойства:*

1) $x \sim x$;

2) если $x \sim$ y, то y $\sim x$;

3) если $x \sim$ y и y $\sim z$, *то $x \sim z$.*

Первое свойство называется рефлексивностью. Оно говорит, что, когда мы имеем отношение эквивалентности, любой элемент эквивалентен сам себе. Это свойство верно для равенства целых чисел: любое целое число равно самому себе. Но оно не выполнено для отношения «<». Поэтому «<» на множестве $\mathbb{Z}$ не является отношением эквивалентности.

Второе свойство называется симметричностью. Отношение «<» на множестве целых чисел не симметрично. Действительно, $2 < 3$, в то время как неравенство $3 < 2$ ложно. С другой стороны, отношение «≤» на $\mathbb{Z}$ рефлексивно, но не симметрично.

Третье — свойство транзитивности. На множестве целых чисел отношения «равно», «меньше, чем», «меньше или равно» — транзитивны. А вот «не равно» этим свойством не обладает. Действительно, $2 \neq 3$ и $3 \neq 2$, но из этих неравенств не следует $2 \neq 2$. Добавим, что «≠» симметрично, но не рефлексивно.

Мы преднамеренно привели примеры отношений, которые не удовлетворяют этим свойствам, потому что это единственный путь к пониманию их действительного смысла. Именно владение примерами и контрпримерами обеспечивает успех в усвоении новых понятий. Равенство целых чисел, очевидно, удовлетворяет всем свойствам, выписанным выше. Отношение «тот же цвет» на множестве цветных мячей — ещё один простой и, пожалуй, самый яркий пример. Среди примеров отношения эквивалентности на множестве многоугольников находятся такие отношения, как «одинаковое число сторон» и «одна и та же площадь».

Отношение эквивалентности используют для классификации элементов данного множества, группируя их в подмножества по принципу схожести свойств. Естественное разбиение множества, индуцированное отношением эквивалентности, называется разбиением на классы эквивалентности. Пусть на множестве $\mathbb{X}$ задано отношение эквивалентности «$\sim$» и x - элемент этого множества.

Определение 4.2. *Классом эквивалентности элемента x называется подмножество в $\mathbb{X}$ всех элементов, эквивалентных x относительно «$\sim$».*

Обозначив класс эквивалентности элемента x символом $\bar{x}$, можно записать:

$$\bar{x} = \{y \in \mathbb{X} \mid y \sim x\}. \qquad (4.1)$$

Приведём простой пример. Обозначим символом $\mathbb{M}$ множество цветных мячей с отношением эквивалентности «тот же цвет». Класс эквивалентности красного мяча в $\mathbb{M}$ состоит из всех красных мячей, содержащихся в $\mathbb{M}$.

Одно из свойств классов эквивалентности настолько важно, что мы назовём его основным принципом классов эквивалентности. Принцип гласит, что любой элемент класса эквивалентности — хороший представитель всего класса. Иначе говоря, зная один элемент из класса эквивалентности, можно немедленно восстановить этот класс полностью. Этот факт бросается в глаза, когда мы имеем дело с множеством $\mathbb{M}$ цветных мячей и отношением «тот же цвет». Вернёмся к абстрактному множеству $\mathbb{X}$ с отношением эквивалентности «$\sim$».

Теорема 4.1 (равенство классов эквивалентности). *Если y — элемент из класса эквивалентности x, то классы эквивалентности x и y совпадают.*

То же самое можно выразить короче: если

$$x \in \mathbb{X}, y \in \bar{x} \Rightarrow \bar{x} = \bar{y}.$$

Доказательство. Воспользуемся определяющими свойствами отношения эквивалентности. Если $y \in \bar{x}$, то, по определению класса эквивалентности, $y \sim x$. Ввиду симметричности, $x \sim y$. Но если $z \in \bar{x}$, то и $z \sim x$. Тогда свойство транзитивности влечет $z \sim y$, т.е. $z \in \bar{y}$. Мы доказали включение: $\bar{x} \subseteq \bar{y}$. Похожее рассуждение доказывает обратное включение: $\bar{y} \subseteq \bar{x}$. Теорема доказана.

Вероятно, это все может показаться несколько педантичным. Но основной принцип — такой источник неразберихи и ошибок, что нам не стоит жалеть усилий на прояснение его точного смысла. Кроме того, полезно осознать, что он непосредственно следует из определения отношения эквивалентности. Кстати, о педантичности: вы поняли, что свойство $x \in \bar{x}$ вытекает из рефлексивности?

Основной принцип приводит к важнейшему свойству отношения эквивалентности. Как и раньше, пусть $\mathbb{X}$ — множество с отношением эквивалентности «$\sim$»; тогда:

1) $\mathbb{X}$ — объединение своих классов эквивалентности относительно «$\sim$»;

2) два разных класса эквивалентности не могут иметь общего элемента.

Первое утверждение следует из часто упоминаемого факта: класс эквивалентности элемента x содержит сам этот элемент. Для доказательства второго предположим, что элементы $x, y, z \in \mathbb{X}$ и $z \in \in \bar{x} \cap \bar{y}$. Так как $z \in \bar{x}$, то по основному принципу $\bar{z} = \bar{x}$. Аналогично $\bar{z} = \bar{y}$. Так что, $\bar{x} = \bar{y}$. Заметим, что свойства (1) и (2) означают, что множество $\mathbb{X}$ разбито на непересекающиеся подмножества, классы эквивалентности. Другими словами, мы имеем дело с разбиением множества $\mathbb{X}$.

Множество, составленное из классов эквивалентности множества $\mathbb{X}$ относительно отношения эквивалентности «$\sim$», имеет специальное название: фактормножество $\mathbb{X}$ по отношению «$\sim$». Отметим, что элементы фактормножества — это подмножества в $\mathbb{X}$. Поэтому фактормножество не является подмножеством в $\mathbb{X}$.

Воспользуемся введенными понятиями для пояснения ситуации с рациональными дробями. Традиционно считается, что дробь состоит из двух целых чисел, одно из которых (знаменатель) должно быть натуральным числом. Конечно, такая конструкция воспринимается как частное. Но, с другой стороны, рациональная дробь в действительности — упорядоченная пара чисел (координатный столбец вектора), одно из которых заведомо не равно нулю. Однако такое определение некорректно.

В математике две упорядоченные пары равны, если они имеют одинаковые первый и второй элементы. Так, пары $(2,4)$ и $(1,2)$ неравны. Но дроби $\frac{2}{4}$ и $\frac{1}{2}$ равны; так что дроби — не пары чисел. Это элементы фактормножества.

Рассмотрим множество $\mathbb{Q}$ пар целых (a, b) с $b \neq 0$, которое может быть представлено в виде декартова произведения

$$\mathbb{Q} = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}).$$

Две пары (a, b) и (a', b') целых чисел можно теперь назвать эквивалентными, если $a \cdot b' = a' \cdot b$. Легко проверить, что это отношение эквивалентности, а дробь — класс эквивалентности множества $\mathbb{Q}$ относительно этого отношения. Следовательно, $a \,/\, b$ означает не пару (a, b), а бесконечное множество всех пар из $\mathbb{Q}$, эквивалентных (a, b). Итак, множество $\mathbb{Q}$ рациональных чисел — это фактормножество множества $\mathbb{Q}$ по только что определённому отношению эквивалентности.

Приведенные рассуждения показывают, что арифметические операции над рациональными числами сводятся к операциям не над множествами упорядоченных пар целых чисел, а над элементами соответствующих фактормножеств. Необходимо только подобрать элементы в фактормножествах – заменить один элемент на другой. Этот элемент полностью описывает в целом класс эквивалентности. Более того, нас устроит любой элемент класса.

Итак, вы можете оперировать с $1/2$, как обычно, так же, как если бы это была пара чисел. Мы вспоминаем, что рациональная дробь — это класс эквивалентности, только когда (в процессе вычислений) оказывается, что дробь можно сократить или расширить. В этот момент вы заменяете одного представителя класса эквивалентности на другой для упрощения вычислений.

В следующем разделе определятся отношение эквивалентности на множестве $\mathbb{Z}$, а фактормножество этого отношения играет абсолютно фундаментальную роль в этом курсе. Как и в случае дробей, классы эквивалентности будут бесконечны, а нам предстоит делать вычисления с ними.

4.2. Сравнения

Проанализируем знакомые 24-часовые часы с точки зрения определенного выше отношения эквивалентности. Когда некто говорит «один час», мы не можем понять, подразумевает ли говорящий это время сегодня, вчера или завтра. Поэтому «один час» — это не момент времени, а класс эквивалентности таких моментов. Объясним более подробно. Сначала разделим временной континуум на равные интервалы, скажем, часы. После этого определим отношение эквивалентности: два момента, отличающиеся на эти 24 равных интервала, эквивалентны. Теперь один час — это класс эквивалентности моментов по специальному отношению эквивалентности. С одной стороны, такой подход усложняет очевидное, но с другой — крайне полезно попрактиковаться с феноменом цикличности.

Теперь мы будем рассматривать похожее отношение эквивалентности, определённое на множестве целых чисел. Выберем натуральное число n, которое с этого момента будет фиксировано. Оно называется периодом, или модулем отношения эквивалентности, которое мы собираемся определить.

Построим отношение эквивалентности на множестве $\mathbb{Z}$, постулируя, что каждое n-е число (начиная с 0) входит в один и тот же класс эквивалентности.

Определение 4.3. *Любые два целых числа, отличающиеся друг от друга на число, кратное n, эквивалентны.*

Более формально два целых числа a и b сравнимы по модулю n, если $a - b$ делится на n; в таком случае мы пишем:

$$a \equiv b \ (mod\ n). \qquad (4.2)$$

Пример 4.1. *Проверить выполнение отношения эквивалентности в следующих случаях: 1) $a = 10, b = 0, n = 5$; 2) $a = 14, b = 27, n = 5$; 3) $a = 32, b = 13, n = 7$.*

Решение. Выполним сравнения. 1) $10 \equiv 0\ (mod\ 5)$. Числа 10 и 0 сравнимы по модулю 5. 2) $14 \equiv 4\ (mod\ 5)$, $27 \equiv 2\ (mod\ 5)$. Числа 14 и 27 не сравнимы по модулю 5. 3) $32 \equiv 4\ (mod\ 7)$, $13 \equiv \equiv 6\ (mod\ 7)$. Числа 32 и 13 не сравнимы по модулю 7.

Заметим, что числа, сравнимые по некоторому модулю, не обязательно сравнимы по другому модулю. Так, 21 сравнимо с 1 по модулю 5; но эти числа не сравнимы по модулю 7, потому что разность $21 — 1 = 20$ не кратна 7.

Проверим, что сравнимость по модулю n является отношением эквивалентности. Рефлексивность. Необходимо доказать, что любое целое число a удовлетворяет сравнению $a \equiv a\ (mod\ n)$. Это будет так, если $a - a$ кратно n. Но $a - a = 0$ кратно любому целому числу. Значит, сравнение по модулю n рефлексивно. Симметричность. Пусть для некоторых целых a и b справедливо сравнение $a \equiv \equiv b\ (mod\ n)$. Это означает, что $a - b = kn$ для какого-то целого k. Умножая это равенство на -1, получаем:

$$b - a = -(a - b) = (-k)n,$$

что также кратно n. Поэтому $b \equiv a\ (mod\ n)$, и мы доказали, что сравнение по модулю n симметрично.

Наконец, транзитивность. Предположим, что $a \equiv b\ (mod\ n)$ и $b \equiv c\ (mod\ n)$, где a, b и c — целые числа. По определению, эти сравнения говорят, что разности $a - b$ и $b - c$ кратны n. Но при сложении кратных n получается число, делящееся на n. Поэтому $(a - -b) + (b - c) = (a - c)$ кратно n. Другими словами, $a \equiv c\ (mod\ n)$, как мы и хотели показать.

Все три свойства для сравнения по модулю n выполняются. Следовательно, оно — отношение эквивалентности.

Определение 4.4. *Факторможество $\mathbb{Z}$ по отношению сравнения по модулю n называется множеством вычетов по модулю n и обозначается символом $\mathbb{Z}_n$.*

Из определения следует, что элементы $\mathbb{Z}_n$ — подмножества в $\mathbb{Z}$, т.е. классы эквивалентности в $\mathbb{Z}$ сравнимых чисел по модулю n. Отождествим между собой элементы этих классов.

Пусть $a \in \mathbb{Z}$. Класс $\bar{a}$ образован всеми целыми b, для которых $b - a$ кратно n, т.е. $b - a = k \cdot n$ для некоторого $k \in \mathbb{Z}$. Таким образом, класс эквивалентности числа a описывается формулой:

$$\bar{a} = \{a + k \cdot n | k \in \mathbb{Z}\}. \qquad (4.3)$$

Заметим, что $\bar{0}$ состоит из всех чисел, кратных n и каждый класс эквивалентности — бесконечное множество.

Мы видели, что в $\bar{a}$ бесконечно много элементов, каждый из которых полностью представляет класс целиком. Отсюда резонный вопрос: можно ли простым способом найти наименьшее натуральное число, представляющее число a? Ответ — да. Нам нужно только поделить число a на n. Обозначим через r остаток, а через q неполное частное этого деления; тогда

$$a = n \cdot q + r, 0 \leq r \leq n - 1. \qquad (4.4)$$

Следовательно, разность $a - r = nq$ кратна n. Поэтому $a \equiv \equiv r \ (mod\ n)$.

Определение 4.5. *Число r (4.4) называется вычетом числа a по модулю n.*

На самом деле мы доказали больше. Действительно, мы показали, что любое целое число сравнимо с одним из целых, лежащим между 0 и $n - 1$. В частности, факторможество $\mathbb{Z}_n$ имеет не более n классов $0, \dots n - 1$. Для уверенности в том, что оно имеет ровно n различных классов сопряжённости по модулю n, нам нужно проверить, что никакие два из выписанных классов не могут быть равными. Каждый класс сопряжённости представляется неотрицательным целым числом, меньшим n. Если бы классы совпадали, их представители были бы сравнимы по модулю n, т.е. разность двух разных неотрицательных целых чисел, меньших n, делилась бы на n. Этого не может быть. Так что, классы $0, \dots, n - 1$ на самом деле различны. Итак,

$$\mathbb{Z}_n = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}.$$

Говорят, что класс эквивалентности $\bar{a}$ записан в приведённом виде, если $0 \le a < n$. Как и в случае дробей, класс сопряжённости удобно представлять именно в таком виде. Этому есть две причины. Первая: меньший представитель класса выбрать легче, чем больший. Вторая: если два класса записаны в приведённом виде, то очень легко определить, равны они или нет (они будут равны тогда и только тогда, когда их представители совпадают). Разумеется, утверждение неверно, если классы записаны не в приведённом виде.

Имеется наглядная геометрическая интерпретация, изображающая множество $\mathbb{Z}_n$. Вспомним геометрическое изображение множества $\mathbb{Z}$. Целые числа — это точки, равномерно расположенные вдоль прямой линии – числовой оси. Где-то на этой прямой стоит точка, изображающая 0, так что отрицательные целые числа находятся слева от нее, а положительные — справа. Сравнение по модулю n отождествляет между собой все числа, отличающиеся на kn ($k \in \mathbb{Z}$), и поэтому, подходя к n, мы должны вернуться опять к 0.

Если прямую целых чисел представить себе как гибкую проволоку, то можно взять точку, помеченную n, и склеить ее с 0, что даст окружность. Продолжая наматывать прямую целых чисел на получившуюся окружность, вы увидите, что числа, сравнимые по модулю n, попадут в одни и те же точки на окружности. В связи с этим $\mathbb{Z}_n$ можно представлять себе как окружность, на которой через равные промежутки отмечены n классов сопряженности.

4.3. Арифметика остатков

Геометрическая картинка, приведённая в конце предыдущего параграфа, помогает нам дать простое описание сложения в $\mathbb{Z}_n$. Представьте себе n классов эквивалентности $\mathbb{Z}_n$ цифрами, отмеченными на циферблате часов. Предположим, что 0 находится в верхней точке круга (что соответствует 12-ти часам), в то время как остальные классы распределены вдоль границы циферблата через равные промежутки. Пусть эти часы имеют только одну стрелку, закреплённую в центре циферблата, и мы можем указывать ею на наши классы. Хотелось бы превратить такие часы в приспособление для вычисления сумм в $\mathbb{Z}_n$. Такой прием – аналог счёта на палочках для традиционной арифметики.

Итак, предположим, что нужно сложить $\bar{a}$ и $\bar{b}$, два класса из $\mathbb{Z}_n$. Будем предполагать, что оба класса записаны в приведённой форме, т.е. a и b неотрицательны и меньше n. Процесс вычисления $a + b$ следующий. Поместим стрелку часов в точку, помеченную a, затем передвинем ее по часовой стрелке на b позиций. Теперь стрелка будет указывать сумму $a + b$.

Пример 4.2. *Найти сумму чисел* $\bar{4}$ *и* $\bar{5}$ *в* $\mathbb{Z}_8$.

Решение. В нашем примере часы имеют 8-часовой формат. Поместим стрелку часов на 5 и передвинем ее на 4 позиции вперед. Стрелка минует 0 и остановится на 1. Значит $\bar{4} + \bar{5} = \bar{1}$ в $\mathbb{Z}_8$.

Требуется определить математический способ вычисления сумм в $\mathbb{Z}_n$. Пусть $\bar{a}$ и $\bar{b}$ — классы в $\mathbb{Z}_n$, которые нам предстоит сложить. Операция построения суммы определяется формулой:

$$\bar{a} + \bar{b} = \overline{a + b}. \qquad (4.5)$$

Интерпретация этой формулы требует некоторой осторожности. Слева в ней стоит сумма двух классов $\mathbb{Z}_n$, а справа мы имеем класс, соответствующий сумме двух целых чисел. Итак, сложение классов определяется в терминах операции, которую мы хорошо знаем: сложение целых чисел.

Вернёмся к примеру сложения в $\mathbb{Z}_8$, которое мы выполнили с помощью прибора. Мы хотим к $\bar{4}$ прибавить $\bar{5}$. Учитывая формулу (4.5), сначала сложим 4 и 5; их сумма, очевидно, равна 9. Отсюда следует, что $\bar{4} + \bar{5} = \bar{9}$. На первый взгляд кажется, что результат суммирования отличается от предыдущего. Но следует учесть, что $9 - 1 = 8$, т.е. $\bar{9} = \bar{1}$.

Последний пример указывает на одну важную проблему. Мы видели, что класс эквивалентности может быть представлен любым своим элементом; это основной принцип раздела 4.1. Но, складывая два класса, мы сначала суммируем их представителей, а затем выбираем соответствующий класс.

Рассмотрим ещё раз сумму $\bar{5}$ и $\bar{4}$ в $\mathbb{Z}_8$. Следуя сформулированному выше правилу, мы нашли, что их сумма равна $\bar{9}$. Однако $\overline{13} = \bar{5}$ и $\overline{12} = \bar{4}$, и формула говорит нам, что, если сложить $\overline{13}$ и $\overline{12}$, то получится $\overline{25}$. Так, поначалу может показаться, что, выбирая различные представители классов, мы получаем разные суммы. Но это

только кажется. На основании того, что $25 - 9 = 16$ делится на 8, мы заключаем: $\overline{25} = \bar{9} = \bar{1}$ в $\mathbb{Z}_8$.

Один из путей решения проблемы мог бы состоять в записи классов в приведённом виде перед их сложением. Это неудобно и совсем необязательно. Как подсказывает разобранный выше пример, результат суммирования не зависит от выбора представителей классов. Это очень важно и должно быть проверено в деталях.

Теорема 4.2. *Пусть $\bar{a}$ и $\bar{b}$ - два класса в $\mathbb{Z}_n$. Если $\bar{a} = \bar{a}'$ и $\bar{b} = \bar{b}'$, то справедливо равенство*

$$\overline{a + b} = \overline{a' + b'}.$$

Доказательство. Равенство $\bar{a} = \bar{a}'$ означает, что $a - a'$ кратно n, и то же самое верно для $b - b'$. Сумма кратных n снова кратна n, откуда число

$$(a - a') + (b - b') = (a + b) - (a' + b')$$

должно быть кратно n. Следовательно, $\overline{a + b} = \overline{a' + b'}$, что и требовалось доказать.

Вычитание классов определяется аналогичным образом и совсем нетрудно. Посмотрим, как должно определяться умножение.

Теорема 4.3. *Пусть $\bar{a}$ и $\bar{b}$ - два класса в $\mathbb{Z}_n$. Тогда справедливо равенство*

$$\bar{a} \cdot \bar{b} = \overline{a \cdot b}. \qquad (4.6)$$

Доказательство. Как и в случае суммы, мы должны быть уверены, что определение умножения не зависит от выбора представителей классов. Итак, предположим, что $\bar{a} = \bar{a}'$ и $\bar{b} = \bar{b}'$. Нужно проверить, что $\overline{a \cdot b} = \overline{a' \cdot b'}$. Равенство $\bar{a} = \bar{a}'$ влечёт: $a - a'$ кратно n; скажем, $a = a' + r \cdot n$ для некоторого целого r. Аналогично $b = b' + s \cdot n$ для некоторого целого s. Умножая a на b, получаем:

$$ab = (a' + r \cdot n)(b' + s \cdot n) = a'b' + (a's + rb' + srn)n.$$

Значит, $ab - a'b'$ кратно n и $\overline{ab} = \overline{a'b'}$. Теорема доказана.

Зная теперь, как складываются и перемножаются классы, любопытно выяснить, так же ли ведут себя эти операции, как их тёзки в $\mathbb{Z}$. Пусть $\bar{a}, \bar{b}$ и $\bar{c}$ — классы в $\mathbb{Z}_n$. Сложение классов обладает следующими свойствами:

$$(\bar{a} + \bar{b}) + \bar{c} = \bar{a} + (\bar{b} + \bar{c});$$
$$\bar{a} + \bar{b} = \bar{b} + \bar{a};$$

$$\bar{a} + \bar{0} = \bar{a};$$
$$\bar{a} + \overline{-a} = \bar{0}.$$

Элемент $\overline{-a}$ называется противоположным a. Заметим, что, если a записан в приведенном виде, то приведенным видом класса $\overline{-a}$ будет $\overline{n-a}$. Умножение классов обладает следующими свойствами:

$$(\bar{a} \cdot \bar{b}) \cdot \bar{c} = \bar{a} \cdot (\bar{b} \cdot \bar{c});$$
$$\bar{a} \cdot \bar{b} = \bar{b} \cdot \bar{a};$$
$$\bar{a} \cdot \bar{1} = \bar{a}.$$

Итак, каждое свойство умножения соответствует свойству сложения, за одним исключением: существование противоположного элемента. Мы вернёмся к этому вопросу в разделе 4.7, где обсудим деление классов.

Есть ещё свойство дистрибутивности

$$\bar{a} \cdot (\bar{b} + \bar{c}) = \bar{a} \cdot \bar{b} + \bar{a} \cdot \bar{c}.$$

Эти свойства довольно легко следуют из их аналогов для сложения и умножения целых, поэтому мы опустим доказательства. Читателю не составит труда найти их самостоятельно.

Операции в $\mathbb{Z}_n$, очевидно, ведут себя так же, как их двойники в $\mathbb{Z}$. Однако «слепо» полагаться на эту аналогию довольно опасно, поскольку может сформироваться ложное чувство уверенности, приводящее к ошибкам. Действительно, одно ключевое свойство целых чисел не выполняется для сравнений по модулю n. Следующий пример проясняет этот факт. Рассмотрим классы $\overline{2}$ и $\overline{3}$ в $\mathbb{Z}_6$. Они оба отличны от нуля, однако

$$\overline{2} \cdot \overline{3} = \overline{6} = \overline{0},$$

т.е. произведение двух ненулевых элементов из $\mathbb{Z}_6$ может равняться нулю. Такого, конечно, не бывает в $\mathbb{Z}$.

В качестве важного следствия этого примера получаем, что в $\mathbb{Z}_n$ не всегда можно сокращать на ненулевой элемент. Другими словами, если $\bar{a} \neq \bar{0}$, то не всегда справедливо рассуждение:

$$\bar{a} \cdot \bar{b} = \bar{a} \cdot \bar{c} \Rightarrow \bar{b} = \bar{c}.$$

Так, например, $\overline{2} \cdot \overline{3} = \overline{2} \cdot \overline{0}$, но $\overline{3} \neq \overline{0}$. Мы вернемся к этому вопросу в разделе 4.6, рассмотрев сначала некоторые приложения арифметики остатков.

4.4. Критерий делимости

Большинство людей помнят из начальной школы, что число делится на 3, если сумма цифр в его десятичной записи делится на 3. Но почему это верно? Мы легко можем это доказать, используя сравнения по модулю 3. Напомним, что число делится на 3 тогда и только тогда, когда оно сравнимо с 0 по модулю 3. Значит, на языке арифметики остатков критерий делимости на 3 утверждает, что число сравнимо с 0 по модулю 3, если и только если то же самое справедливо для суммы его цифр. Последнее утверждение мы сейчас и докажем.

Пусть a — целое число и $a_0, a_1, \dots, a_n$ — его цифры в десятичной записи. Иначе говоря,

$$a = a_n 10^n + a_{n-1} 10^{n-1} + \cdots + a_1 10 + a_0,$$

где $0 \le a_i < 9$ для $i = 0,1, \dots, n$. В предыдущем разделе мы показали, что умножение не зависит от выбора представителей классов по модулю n. Поскольку $10 \equiv 1 \ (mod\ 3)$, независимость от выбора говорит нам, что 10^k сравнимо с 1^k по модулю 3 для любого положительного целого к. Другими словами, любая степень 10 имеет вычет 1 по модулю 3. Значит,

$$a \equiv a_n + a_{n-1} + \cdots + a_1 + a_0 \ (mod\ 3).$$

Отсюда немедленно следует, что $a \equiv 0 \ (mod\ 3)$ тогда и только тогда, когда $a_n + a_{n-1} + \cdots + a_1 + a_0 \equiv 0 (mod\ 3)$. А это как раз то, что нужно было доказать.

Заметим, что все проделанные вычисления останутся верными после замены 3 на 9, потому что $10 \equiv 1 \ (mod\ 9)$. Значит, целое число делится на 9, если и только если на 9 делится сумма его цифр в десятичной записи.

Применим похожие аргументы к другому числу, например к 11. Снова будем предполагать, что

$$a = a_n 10^n + a_{n-1} 10^{n-1} + \cdots + a_1 10 + a_0,$$

где $a_n, a_{n-1}, \dots, a_1, a_0$ — цифры записи числа a. Поскольку $10 = -1 \ (mod\ 11)$, то

$$10^k \equiv (-1)^k (mod\ 11)$$

будет либо 1 (если k чётно), либо -1 (если k нечётно). Поэтому

$$a \equiv a_n (-1)^n + a_{n-1} (-1)^{n-1} + \cdots - a_1 + a_0 \ (mod\ 11).$$

Говоря человеческим языком, число делится на 11 тогда и только тогда, когда на 11 делится альтернированная сумма его цифр.

Например, 3443 делится на 11, потому что $3 + 4 + 4 - 3 = 0$ делится на 11.

Критерии делимости на 2 и 5 слишком очевидны, чтобы приводить их доказательства. Таким образом мы нашли простые критерии делимости на все простые числа от 2 до 11, исключая 7. Разберёмся, что произойдёт в случае применения того же подхода к 7.

Мы уже знаем из предыдущего примера, что та часть рассуждений, которая зависит от модуля, состоит в вычислении степеней 10. На этот раз $10 \equiv 3\ (mod\ 7)$, а степени 3 не так легко вычисляются, как степени 1 или —1. Попытаемся найти эти степени при малых показателях. Все дальнейшие сравнения сделаны по модулю 7.

$$10^2 \equiv 3^2 \equiv 2;$$
$$10^3 \equiv 10 \cdot 10^2 \equiv 3 \cdot 2 \equiv 6 \equiv -1;$$
$$10^4 \equiv 10 \cdot 10^3 \equiv (-1) \cdot 3 \equiv 4;$$
$$10^5 \equiv 10 \cdot 10^4 \equiv 3 \cdot 4 \equiv 5;$$
$$10^6 \equiv 10 \cdot 10^5 \equiv 3 \cdot 5 \equiv 1.$$

Заметим, что последний вычет $10^0 \equiv 1$. Это означает, что вычеты будут циклически, с периодом 6, повторяться. Эти вычисления показывают, что критерий делимости на 7 несколько более сложен для запоминания, чем на 3 и 11. Поскольку мы уже довольно далеко продвинулись в работе над этим критерием, точно сформулируем его в простом случае. Предположим, что

$$a = a_2 10^2 + a_1 10 + a_0, 0 \le a_0, a_1, a_2 \le 9.$$

Используя вычеты степеней 10, вычисленные только что, мы имеем

$$a \equiv a_2 10^2 + a_1 10 + a_0 \equiv 2a_2 + 3a_1 + a_0\ (mod\ 7),$$

на 7 делится выражение $2a_2 + 3a_1 + a_0$. Например, 231 делится на 7 ввиду делимости $2 \cdot 2 + 3 \cdot 3 + 1 = 14$ на 7.

4.5. Степени

Во многих приложениях мы будем встречаться со следующей задачей: пусть a, k и n — натуральные числа; найти остаток от деления a^k на n. Если k очень большое, то может оказаться невозможным даже пересчитать знаки в a^k, как в примере из начала этого раздела. Однако мы можем упростить эту проблему, используя арифметику остатков.

Начнём с простого примера. Предположим, что мы хотим найти остаток от деления 10^{135} на 7. Мы видели в предыдущем параграфе,

что $10^6 \equiv 1\ (mod\ 7)$. Деля 135 на 6, мы найдём $135 = 6 \cdot 22 + 3$. Полученное равенство даёт следующие сравнения по модулю 7:

$$10^{135} \equiv (10^6)^{22} \cdot 10^3 \equiv (1)^{22} \cdot 10^3 \equiv\ 6.$$

Значит, остаток от деления 10^{135} на 7 равен 6.

Вычисления не всегда оказываются такими лёгкими. Например, какой остаток получится при делении 3^{64} на 31? Вычислив несколько степеней 3 по модулю 31, мы быстро найдём, что $3^3 \equiv \equiv -4\ (mod\ 31)$. Вместо вычисления более высоких степеней в надежде, что одна из них станет снова 1, воспользуемся уже имеющейся информацией. Так как $64 = 3 \cdot 21 + 1$, мы получаем сравнение по модулю 31:

$$3^{64} \equiv (3^3)^{21} \cdot 3 \equiv (-4)^{21} \cdot 3 \equiv -(2)^{42} \cdot 3.$$

Мы ещё не получили искомый остаток, а только степень 2, которая лежит между нами и нашей целью. Удачно, что $2^5 \equiv \equiv 1\ (mod\ 31)$. Так как $45 = 8 \cdot 5 + 2$, мы находим:

$$3^{64} \equiv -2^{42} \cdot 3 \equiv -(2^5)^8 \cdot 2^2 \cdot 3 \equiv -12\ (mod\ 31).$$

Но $-12 \equiv\ 19\ (mod\ 31)$, так что остаток от деления 3^{64} на 31 равен 19.

Были бы вычисления легче, продолжай мы находить степени 3 до тех пор, пока не встретится 1? Ответ — нет, в чем вы сами можете убедиться, попытавшись найти наименьшее натуральное r, для которого $3^r \equiv 1\ (mod\ 31)$.

Допустим, что мы хотим найти остаток от деления 6^{35} на 16. В этом случае бесполезно пытаться отыскать наименьшую степень 6, сравнимую с 1 по модулю 16: ее просто нет. Действительно,

$$6^4 \equiv 2^4 \cdot 3^4 \equiv 0\ \cdot 3^4 \equiv\ 0\ (mod\ 16),$$

откуда

$$6^{35} \equiv 6^4 \cdot 6^{31} \equiv 0\ (mod\ 16).$$

Эти примеры иллюстрируют некоторые приемы, используемые для облегчения подсчёта вычетов степеней по модулю n. Другие приёмы появятся в последующих разделах. Конечно, компьютеру не нужны никакие трюки. Это не говорит о том, что компьютер не использует арифметику остатков для таких вычислений; на самом деле он ее использует. Быстрый алгоритм, вычисляющий степени по модулю n, приведен в упражнениях. Его можно использовать для доказательства делимости числа $F(23471)$ на $5 \cdot 2^{23473} + 1$ — достижение, которое раньше казалось невыполнимой задачей.

4.6. Диофантовы уравнения

Далее мы используем сравнения для доказательства отсутствия решений у некоторых диофантовых уравнений. Диофантово уравнение — это полиномиальное уравнение с несколькими неизвестными и целыми коэффициентами. Примеры: $3x - 2y = 1$, $x^3 + y^3 = z^2$, $x^3 - 117y^3 = 5$. Когда говорят о решениях диофантова уравнения, обычно имеют в виду целочисленные решения. Эти уравнения названы по имени греческого математика Диофанта из Александрии, жившего около 250 г. н.э. В своей «Арифметике» Диофант подробно обсуждает проблему поиска решений неопределённых уравнений. Однако он искал рациональные решения, а не целые, что мы обычно делаем сегодня.

Поскольку эти уравнения зависят от нескольких переменных, они могут иметь бесконечно много решений. Например, для любого целого k числа $x = 1 + 2k$ и $y = 1 + 3k$ удовлетворяют уравнению $3x - 2y = 1$. Уравнение $x^3 + y^3 = z^3$ — частный случай великой Теоремы Ферма, о которой упомянули в конце второго раздела. Как мы видели, если x, y и z — целые числа, удовлетворяющие этому уравнению, то одно из них должно быть равно нулю. Это специальный случай теоремы, впервые доказанный Эйлером в 1770 году. Вспомните, что Ферма сформулировал свою великую Теорему на полях принадлежащей ему копии Диофантовой «Арифметики».

Пример 4.3. *Решить уравнение* $x^3 - 117y^3 = 5$ *в целых числах.*

Решение. Воспользуемся сравнением по модулю 9 и покажем, что решения нет. Доказательство «от противного». Предположим, что уравнение $x^3 - 117y^3 = 5$ имеет целочисленное решение, т.е. существуют такие целые x_0 и y_0, что $x_0^3 - 117y_0^3 = 5$. Так как все числа в этом выражении целые, мы можем рассмотреть его по модулю 9. Число 117 делится на 9, поэтому

$$x_0^3 \equiv x_0^3 - 117y_0^3 \equiv 5 \ (mod\ 9).$$

Следовательно, если это уравнение имеет целое решение (x_0, y_0), то $x_0^3 \equiv 5 \ (mod\ 9)$. Возможна ли такая ситуация? Чтобы разобраться, напомним: каждое целое число по модулю 9 имеет вычет, лежащий между 0 и 8. Значит, нам будет достаточно вычислить кубы по модулю 9 каждого из этих вычетов.

$\bar{x}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{8}$
$\overline{x^3}$	$\bar{0}$	$\bar{1}$	$\bar{8}$	$\bar{0}$	$\bar{1}$	$\bar{8}$	$\bar{0}$	$\bar{1}$	$\bar{8}$

Простой взгляд на таблицу показывает, что вычет куба целого числа по модулю 9 может быть равен только 0, 1 или 8. В частности, нет такого целого x_0, куб которого сравним с 5 по модулю 9: $x^3 \equiv$ $\equiv 5\ (mod\ 9)$. Следовательно, $x^3 - 117y^3 = 5$ не может иметь целочисленных решений, что и требовалось доказать.

4.7. Деление по модулю n

Настало время вернуться к проблеме деления классов в $\mathbb{Z}_n$. Но сначала рассмотрим тот же самый вопрос в более знакомой ситуации. Пусть a и b — вещественные числа. Один из способов деления a на b состоит в умножении a на $\frac{1}{b}$. Число $\frac{1}{b}$ называется обратным к b и однозначно определяется как решение уравнения $b \cdot x = 1$. С практической точки зрения этот способ не облегчает дела, потому что для вычисления $\frac{1}{b}$ мы все равно должны разделить 1 на b. Однако с теоретической точки зрения иногда бывает лучше рассуждать об обратном элементе, нежели о делении. Наконец, $\frac{1}{b}$ существует только, если $b \neq 0$, поскольку уравнение $0 \cdot x = 1$ не имеет решений. Вернёмся к $\mathbb{Z}_n$, где n — фиксированное натуральное число. Допустим, что $\bar{a} \in \mathbb{Z}_n$.

Определение 4.6. *Назовём $\bar{\alpha} \in \mathbb{Z}_n$ обратным к $\bar{a}$, а сам элемент $\bar{a}$ обратимым, если в Z_n справедливо равенство: $\bar{a} \cdot \bar{\alpha} = \bar{1}$.*

Ясно, что $\bar{0}$ не имеет обратного в $\mathbb{Z}_n$. К сожалению, $\bar{0}$ может быть не единственным элементом в $\mathbb{Z}_n$, не имеющим обратного. Необходимо рассмотреть этот момент очень подробно.

Предположим, что $\bar{a} \in \mathbb{Z}_n$ имеет обратный элемент $\bar{\alpha}$, и посмотрим, что мы с этого будем иметь. Из уравнения

$$\bar{a} \cdot \bar{\alpha} = \bar{1}$$

следует, что $a \cdot \alpha$ делится на n. Иначе говоря,

$$a \cdot \alpha + k \cdot n = 1 \qquad (4.7)$$

для некоторого целого k. Равенство (4.7) влечет: $\text{НОД}(a; n) = 1$. Таким образом, мы заключаем, что, если $\bar{a}$ имеет обратный элемент в Z_n, то $\text{НОД}(a; n) = 1$.

Верно ли обратное? Для ответа предположим, что a — такое целое число, для которого $\text{НОД}(a; n) = 1$. Равенство (4.7) подсказывает применить расширенный алгоритм Евклида 1.4 к числам a и n. Этот алгоритм даст нам такие целые числа α и β, что

$$a \cdot \alpha + \beta \cdot n = 1.$$

Полученное уравнение эквивалентно тождеству

$$\bar{a} \cdot \bar{\alpha} = \bar{1}$$

в $\mathbb{Z}_n$. Значит, класс $\bar{a}$, найденный с помощью расширенного алгоритма Евклида, является обратным к $\bar{a}$ в $\mathbb{Z}_n$. Итак, если $\text{НОД}(a, n) = 1$, то $\bar{a}$ обратим в $\mathbb{Z}_n$. Подведем итог в следующей теореме.

Теорема 4.4 (обратимость). *Класс $\bar{a}$ обратим в $\mathbb{Z}_n$ тогда и только тогда, когда целые a и n взаимно просты.*

Рассуждения, приведённые выше, являются конструктивным доказательством теоремы обратимости в том смысле, что они представляют собой процедуру для проверки существования обратного элемента и его вычисления, если он есть. Конечно, эта процедура — непосредственное применение расширенного алгоритма Евклида. Например, обладает ли $\bar{3}$ обратным в $\mathbb{Z}_{32}$? И, если «да», то чему он равен? Применяя расширенный алгоритм Евклида к числам 32 и 3, мы найдём, что $\text{НОД}(3,32) = 1$ и

$$3 \cdot 11 + (-1) \cdot 32 = 1.$$

Так что, обратный элемент существует. Переписывание этого уравнения по модулю 32 приводит к равенству: $\bar{3} \cdot \overline{11} = \bar{1}$. Итак, $\overline{11}$ — обратный элемент к $\bar{3}$ в $\mathbb{Z}_{32}$.

Множество обратимых элементов в $\mathbb{Z}_n$, обозначающееся символом $\mathbb{U}(n)$, играет ключевую роль в последующих разделах курса. По теореме обратимости мы имеем:

$$\mathbb{U}(n) = \{\bar{a} \in Z_n | \text{НОД}(a, n) = 1\}. \qquad (4.8)$$

Для простого числа p множество $\mathbb{U}(p)$ вычислить очень легко, потому что в этом случае условие $\text{НОД}(a, p)$ равносильно такому: p не делит a. Поскольку это справедливо для всех натуральных чисел, меньших p, имеем $\mathbb{U}(p) = \mathbb{Z}_p \setminus \{0\}$.

К сожалению, такое простое описание множества $\mathbb{U}(p)$ верно только для простого p. Если n — составное и $1 < k < n$ — его множитель, то $\text{НОД}(k, n) = k \neq 1$, так что k не обратим в $\mathbb{Z}_n$. Два простых примера.

$$\mathbb{U}(4) = \{\bar{1}, \bar{3}\}, \mathbb{U}(8) = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}\}.$$

Ключевое свойство множества $\mathbb{U}(n)$ состоит в том, что оно содержит произведение любых своих элементов.

Теорема 4.5. *Если $\bar{a}$ и $\bar{b}$ — обратимые классы в $\mathbb{Z}_n$, то произведение $\bar{a} \cdot \bar{b}$ тоже обратимо в $\mathbb{Z}_n$.*

Доказательство. Пусть $\bar{\alpha}$ — обратный элемент для $\bar{a}$, а $\bar{\beta}$ — для $\bar{b}$ в $\mathbb{Z}_n$. Тогда обратным к произведению $\bar{a} \cdot \bar{b}$ будет $\bar{\alpha} \cdot \bar{\beta}$. Для доказательства этого допущения воспользуемся определением:

$$(\bar{a} \cdot \bar{b})(\bar{\alpha} \cdot \bar{\beta}) = (\bar{a} \cdot \bar{\alpha})(\bar{b} \cdot \bar{\beta}) = \bar{1} \cdot \bar{1} = \bar{1}.$$

Теорема доказана. Мы еще будем периодически возвращаться к множеству $\mathbb{U}(n)$ в последующих разделах.

Вернёмся к проблеме делимости $\bar{a}$ на $\bar{b}$ в $\mathbb{Z}_n$, с которой мы начинали этот раздел. Прежде всего, нам нужно узнать, имеет ли $\bar{b}$ обратный элемент в $\mathbb{Z}_n$. Если имеет, то мы найдем его с помощью расширенного алгоритма Евклида. Пусть, например, это будет $\bar{\beta}$. Чтобы разделить $\bar{a}$ на $\bar{b}$, мы вычислим произведение $\bar{a} \cdot \bar{\beta}$.

Пример 4.4. *Найти результат деления $\bar{2}$ на $\bar{3}$ в $\mathbb{Z}_8$.*

Решение. После применения алгоритма Евклида к 3 и 8 мы имеем: $\text{НОД}(3,8) = 1$. Тогда справедливо $\bar{3} \cdot \bar{3} = \overline{3 \cdot 3} = \bar{9} = \bar{1}$. Число 3 само себе обратимо. Поэтому результат деления $\bar{2}$ на $\bar{3}$ в $\mathbb{Z}_8$ равен $\frac{\bar{2}}{\bar{3}} = \bar{2} \cdot \bar{3} = \bar{6}$.

Применим результаты этого параграфа к решению линейных сравнений в $\mathbb{Z}_n$. Линейное сравнение — это уравнение вида

$$ax \equiv b \ (mod\ n), \qquad (4.9)$$

где $a, b \in \mathbb{Z}$. Если бы это было линейное уравнение над вещественными числами, нам следовало бы поделить его на a. Попытаемся использовать ту же идею здесь. Предполагая, что $\text{НОД}(a, n) = 1$, мы заключаем (по теореме обратимости), что существует такой $\alpha \in \mathbb{Z}$, что $a \cdot \alpha \equiv 1 \ (mod\ n)$. Умножая обе стороны уравнения (4.9) на α, получаем:

$$x \equiv \alpha \cdot a \cdot x \equiv \ \alpha \cdot b \ (mod\ n),$$

и уравнение решено. Например, для решения сравнения $7x \equiv$ $\equiv 3 \ (mod\ 15)$ мы сначала ищем обратный элемент к 7 по модулю 15. Так как $15 - 2 \cdot 7 = 1$, обратным к 7 по модулю 15 будет $-2 \equiv$ $\equiv 13 \ (mod\ 15)$. Умножая сравнение $7x \equiv 3 \ (mod\ 15)$ на 13, получаем

$$x \equiv 13 \cdot 3 \equiv 39 \equiv \ 9 \ (mod\ 15),$$

что является искомым решением.

Заметим, метод, применяемый к решению линейных сравнений, показывает, что в случае НОД$(a, n) = 1$ сравнение $a \cdot x \equiv b\ (mod\ n)$ имеет одно и только одно решение по модулю n. Иначе говоря, несмотря на бесконечность числа целых, удовлетворяющих данному уравнению, все они сравнимы друг с другом по модулю n. Это замечание существенное. Уравнение $2 \cdot x \equiv 1\ (mod\ 8)$, например, не имеет решений вовсе.

5. Теорема Ферма

Теперь, зная базисные факты арифметики остатков, мы готовы вернуться к изучению простых чисел. Основной результат этой главы — очень полезная теорема, впервые доказанная Ферма. Она непосредственно следует из более общей теоремы теории групп, которую мы изучим в разделе 8. Здесь же мы, следуя рассуждениям Ферма, даём её прямое доказательство методом математической индукции. С описания такого метода доказательства мы и начнём.

5.1. Метод математической индукции

Теорема 5.1 (принцип математической индукции). *Пусть для каждого натурального числа n существует утверждение (предикат) $S(n)$, обладающее следующими двумя свойствами:*

1) $S(1)$ верно;

2) если $S(k)$ верно для натурального k, то $S(k+1)$ также верно. Тогда $S(n)$ справедливо для всех натуральных n.

Доказательство. Допустим, у нас есть утверждение, обладающее свойствами (1) и (2) принципа. Второе из них говорит нам, что если для некоторого натурального k мы можем показать справедливость $S(k)$, то $S(k+1)$ тоже будет верным. По первому свойству $S(1)$ — истинное утверждение. Поэтому, применяя (2) при $k = 1$, мы заключаем, что $S(1+1)$ верно. Теперь, зная об истинности $S(2)$, мы опять можем применить (2), но на этот раз при $k = 2$. Получим истинность $S(2+1)$. Таким образом, выбрав произвольное натуральное n, мы можем продолжать действовать подобным образом, пока не достигнем $S(n)$. Поэтому утверждение должно быть справедливым для каждого положительного целого n. Принцип доказан.

5.2. Теорема Ферма

Теорему, которую мы хотим доказать, иногда называют «малой теоремой Ферма». Она говорит, что если p простое число и a — любое целое, то p делит $a^p - a$. Частный случай этой теоремы известен многие сотни лет, но Ферма, кажется, был первым, кто доказал теорему в полной общности. Сформулируем эту теорему в терминах сравнений. Первоначально докажем вспомогательное утверждение.

Лемма 5.1. *Пусть p — положительное простое число, a и b — целые; тогда*

$$(a+b)^p \equiv a^p + b^p \pmod{p}. \qquad (5.1)$$

Доказательство. Как следует из формулы бинома Ньютона, для произвольного натурального p выполняется тождество

$$(a+b)^p = \sum_{k=1}^{p-1} C_p^k a^{p-k} b^k = a^p + b^p + \sum_{k=1}^{p-1} C_p^k a^{p-k} b^k.$$

Таким образом, для доказательства леммы достаточно показать, что каждое слагаемое, стоящее под знаком суммы, заведомо делится на p. По определению биномиального коэффициента получаем:

$$C_p^k = \frac{p!}{k!\,(p-\mathrm{k})!} = \frac{p(p-1)(p-2)\cdots(p-\mathrm{k}+1)}{k!}.$$

Поскольку биномиальные коэффициенты — целые числа, знаменатель дроби должен делить числитель. Кроме того, если $1 \le k \le p-1$, то p не является делителем числа $k!$. Поэтому сомножитель p из числителя дроби не может сократиться ни с каким сомножителем знаменателя. Значит, знаменатель обязан делить произведение $(p-1)(p-2)\cdots(p-k+1)$. Отсюда C_p^k — число, кратное p, то есть

$$C_p^k \equiv 0\ (mod\ p), 1 < k < p$$

или

$$C_p^k a^{p-k} b^k \equiv 0\ (mod\ p), 1 < k < p.$$

Таким образом, каждое слагаемое суммы кратно p, а значит, сама сумма сравнима с нулем по модулю p:

$$\sum_{k=1}^{p-1} C_p^k a^{p-k} b^k \equiv 0\ (mod\ p),$$

что и требовалось доказать.

Теорема 5.2 (Ферма). *Пусть p — положительное простое число и a — целое; тогда*

$$a^p \equiv a \pmod{p}. \qquad (5.2)$$

Доказательство. Воспользуемся методом математической индукции, для чего необходимо найти утверждение $S(n)$ (предикат), к которому можно применить метод. Вот этот предикат:

$$n^p \equiv n(mod\ p)$$

для натурального n.

Отметим, что утверждение $S(n)$ говорит о верности сравнения только для натуральных чисел, т.е. положительных целых n. Таким образом, мы не доказываем теоремы в полной общности. Однако любое целое число сравнимо по модулю p с каким-то неотрицательным целым, меньшим p. Поэтому достаточно проверить теорему для $0 \le a \le p-1$. В частности, теорема будет доказана, если показать справедливость утверждений $S(n)$ для всех $n \ge 1$.

Разумеется, $S(1) = 0$ - истинное высказывание, поскольку $1^1 = = 1$. Таким образом, база индукции выполнена. Для индуктивного перехода от $S(n)$ к $S(n+1)$ нам нужно выявить связь между этими утверждениями. Теперь можно вернуться к доказательству теоремы Ферма. Предположение индукции: $n^p \equiv n\ (mod\ p)$ для некоторого натурального n. В качестве индуктивного перехода нам нужно показать, что $(n+1)^p \equiv n+1\ (mod\ p)$. По лемме

$$(n+1)^p \equiv n^p + 1^p \equiv n^p + 1\ (mod\ p).$$

По предположению индукции, n^p можно заменить на n. Проделав это, мы получим: $(n+1)^p \equiv n^p + 1 \equiv n + 1\ (mod\ p)$, что и требовалось доказать.

Применим теорему для упрощения вычислений степеней по модулю p - проблемы, с которой мы уже сталкивались. Сначала переформулируем теорему 5.2 в более удобной форме.

Теорема 5.3 (Ферма). *Пусть p — положительное простое число и a — целое, не делящееся на p; тогда*

$$a^{p-1} \equiv 1 \pmod{p}. \qquad (5.3)$$

Доказательство. Предположим, теперь, что a не делится на p. Тогда, ввиду простоты p, числа a и p взаимно просты, и из теоремы обратимости следует, что a обратимо по модулю p. Пусть a' —

обратный к a элемент. Умножая на a' сравнение $a^p \equiv a\ (mod\ p)$, имеем:

$$a' \cdot a \cdot a^{p-1} \equiv a' \cdot a\ (mod\ p).$$

Но $a' \cdot a \equiv 1\ (mod$ р), и окончательно $a^{p-1} \equiv 1\ (mod\ p)$. Теорема доказана.

Задача, к которой мы хотели бы применить теорему Ферма, формулируется следующим образом. Для трех данных натуральных чисел a, k и p таких, что $k > p - 1$, найти вычет a^k по модулю p.

Если p делит a, то вычет равен 0. Поэтому без ограничения общности можно предполагать, что p не делит a. Разделим k на p— 1 с остатком: $k = (p-1)q + r$, где q и r — неотрицательные целые числа, причем $0 \le r \le p - 1$. Следовательно,

$$a^k \equiv a^{(p-1)q+r} \equiv (a^{p-1})^q a^r\ (mod\ p).$$

По теореме Ферма, $a^{p-1} \equiv 1\ (mod\ p)$. Значит, $a^k \equiv a^r\ (mod\ p)$, и достаточно делать вычисления только для экспонент, показатель которых меньше $p - 1$.

Пример 5.1. *Найти вычет числа* $2^{5432675}$ *по модулю 13.*

Решение. Метод предыдущего раздела предписывал вычислить несколько степеней двойки по модулю 13, прежде чем что-нибудь получить. Посмотрим, что будет, если применить теорему Ферма. Сначала найдём остаток от деления 5432675 на 13— 1 = 12. Он равен 11. Затем, рассуждая, как в предыдущих абзацах, мы имеем:

$$2^{5432675} \equiv 2^{11}\ (mod\ 13).$$

Непосредственный подсчёт даёт окончательный ответ: $2^{11} \equiv$ $\equiv 7\ (mod\ 13)$.

5.3. Вычисление корней

Остаётся ещё один вопрос, который нужно рассмотреть. Существует ли натуральное k, меньшее чем $p - 1$, при котором $a^k \equiv$ $\equiv 1\ (mod\ p)$ для всех целых чисел a, не делящихся на p?

Можно попытаться ответить на него следующим образом. Хорошо известная теорема алгебры говорит, что полиномиальное уравнение не может иметь корней больше, чем его степень. А поскольку, по предположению, сравнение $a^k \equiv 1\ (mod\ p)$ справедливо для всех целых чисел, взаимно простых с p, корни полиномиального уравнения $x^k = \bar{1}$ в $\mathbb{Z}_p$ - это $\bar{1}, \bar{2}, \ldots, \overline{p-1}$.

Таким образом, уравнение имеет $p-1$ разных корня. По упомянутой теореме получаем оценку: $k \geq p-1$. Так что, ответ на поставленный вопрос отрицателен.

Хотя эти рассуждения корректны, они прячут истинную природу проблемы в теореме, с помощью который мы решили этот вопрос. Когда мы говорим о полиномиальном уравнении, мы обычно имеем в виду уравнение с вещественными или комплексными коэффициентами. А под «корнем» мы подразумеваем вещественный или комплексный корень. Но коэффициенты уравнения из нашего рассуждения, как и его корни, суть элементы $\mathbb{Z}_p$. Нам нужно показать, что теорема о числе корней полиномиального уравнения справедлива и в этом случае. Хотя эта теорема верна в случае простого модуля, для составного модуля она ложна!

Лемма 5.2. *Пусть $h(x)$ — многочлен степени m с целыми коэффициентами. Для любого целого числа α найдётся многочлен $q(x)$ степени $m-1$, удовлетворяющий условию:*

$$h(x) = q(x) \cdot (x-\alpha) + h(\alpha).$$

Доказательство. Воспользуемся математической индукцией. Пусть $S(k)$ – утверждение (предикат) истинный на многочленах степени k. Непосредственно проверяем для $k=1$. $S(1)$ истинное высказывание. Многочлен первой степени с целыми коэффициентами $h(x) = ax + b$. Поэтому выполняется

$$h(x) = ax + b = a \cdot (x-\alpha) + \alpha \cdot a + b = a \cdot (x-\alpha) + h(\alpha).$$

Пусть верны $S(k)$ утверждения для $k = 1,2,\dots,m-2,m-1$. Докажем, что справедливо утверждение леммы и для многочлена степени m, то есть $S(m)$. Рассмотрим многочлен с целыми коэффициентами степени m:

$$h(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0,$$

где $a_m \neq 0$. Обозначим через $g(x)$ разность

$$h(x) - a_m x^{m-1} \cdot (x-\alpha),$$

то есть

$$g(x) = (a_{m-1} + a_m\alpha)x^{m-1} + a_{m-2}x^{m-2} + \cdots + a_1 x + a_0.$$

Из записи видно, что степень многочлена $g(x)$ меньше или равна $m-1$: степень будет в точности равна $m-1$ только, если выполняется неравенство $a_{m-1} + a_m\alpha \neq 0$.

Так как степень многочлена $g(x)$ меньше или равна $m-1$, предположение индукции влечёт

$$g(x) = p(x)(x - \alpha) + g(\alpha),$$

где $p(x)$ — многочлен с целыми коэффициентами, степень которого на единицу меньше, чем степень $g(x)$. Ввиду равенства $g(\alpha) = h(\alpha)$, получаем:

$$g(x) = p(x)(x - \alpha) + h(\alpha).$$

Но $h(x) = g(x) + a_m x^{m-1}(x - \alpha)$, так что

$$h(x) = (p(x) + a_m x^{m-1})(x - \alpha) + h(\alpha).$$

Наконец, $p(x) + a_m x^{m-1}$ имеет степень в точности $m - 1$, потому что степень $p(x)$ меньше $m - 1$. Лемма доказана.

Замечание. Есть альтернативное доказательство этой леммы, использующее деление многочленов. Как и прежде, пусть $h(x)$ — многочлен степени m с целыми коэффициентами. Разделив $h(x)$ на $x - \alpha$, мы найдем такие многочлены $q(x)$ и $r(x)$, что

$$h(x) = q(x)(x - \alpha) + r(x),$$

причём либо $r(x) = 0$, либо его степень меньше степени $x - \alpha$. Значит, его степень должна быть равной 0, т.е. $r(x) = c$ — целое число. Заменим в полученном выражении x на α:

$$h(\alpha) = q(\alpha)(\alpha - \alpha) + c.$$

Таким образом, мы можем его переписать в виде:

$$h(x) = q(x)(x - \alpha) + h(\alpha),$$

что завершает доказательство.

Теорема 5.4. *Пусть $f(x)$ — многочлен степени k с целыми коэффициентами и старшим коэффициентом 1. Если p — простое число, то $f(x)$ имеет не более k корней в $\mathbb{Z}_p$.*

Доказательство. Обратим внимание на два момента. Во-первых, модуль должен быть простым, поскольку нам нужно следующее свойство: если $a \cdot b \equiv 0 \ (mod\ p)$, то или $a \equiv 0 \ (mod\ p)$ или $b \equiv 0 \ (mod\ p)$. Это фактически фундаментальное свойство простых чисел, сформулированное на языке сравнений.

Применим метод математической индукции по n (степени многочлена) с помощью леммы 5.2. Если $n = 1$, то $f(x) = x + b$. И соответствующее уравнение имеет только одно решение $\overline{-b}$ в $\mathbb{Z}_p$. Итак, многочлен степени 1 имеет единственный корень в $\mathbb{Z}_p$, и теорема в этом случае верна.

Допустим, что любой многочлен степени $k - 1$ со старшим коэффициентом 1 имеет не более $k - 1$ корня в $\mathbb{Z}_p$. Это предположение

индукции. Мы хотим показать, что данное предположение влечёт аналогичное утверждение для многочлена степени k со старшим коэффициентом 1.

Итак, пусть $f(x)$ — многочлен степени k с целыми коэффициентами, старший из которых равен 1. Если $f(x)$ не имеет корней в $\mathbb{Z}_p$, то доказывать нечего, ибо $k \geq 0$. Такие многочлены в действительности существуют; соответствующий пример приведён после доказательства теоремы. Следовательно, мы можем предполагать, что $f(x)$ имеет корень $\bar{\alpha} \in \mathbb{Z}_p$; иначе говоря, $f(\alpha) \equiv 0\ (mod\ p)$. По лемме 5.2

$$f(x) = (x - \alpha)q(x) + f(\alpha), \tag{5.4}$$

где степень $q(x)$ равна $k - 1$. Так как старшие коэффициенты многочленов $f(x)$ и $x - a$ равны 1, то же самое справедливо и для многочлена $q(x)$. Значит, мы можем применить индуктивное предположение к $q(x)$.

Рассматривая равенство (5.4) по модулю p, мы получаем:

$$f(x) \equiv (x - \alpha)q(x)\ (mod\ p). \tag{5.5}$$

Пусть $\bar{\beta} \neq \bar{\alpha}$ — ещё один корень $f(x)$ в $\mathbb{Z}_p$. Это означает, что

$$f(\beta) \equiv (mod\ p), \alpha - \beta \equiv (mod\ p).$$

Заменяя x на β в (5.5) и используя эти сравнения, мы имеем:

$$0 \equiv f(\beta) \equiv (\beta - \alpha)\mathrm{q}(\beta)\ (mod\ p).$$

Так как p простое, то это влечёт сравнение $q(\beta) \equiv\ 0\ (mod\ p)$. Мы заключаем, что, если $\bar{\beta}$ — корень многочлена $f(x)$, отличный от $\bar{\alpha}$, то $\bar{\beta}$ является и корнем $q(x)$ в $\mathbb{Z}_p$. Иначе говоря, у $f(x)$ только на один корень больше (в Z_p), чем у $q(x)$. А по предположению индукции у последнего многочлена не более чем $k - 1$ различных корней в $\mathbb{Z}_p$. Следовательно, $f(x)$ не может иметь больше k разных корней, что завершает доказательство теоремы.

Рассмотрим несколько примеров. Первый из них — многочлен $f(x) = x^2 + 3$. Он удовлетворяет всем условиям теоремы, но не имеет корней по модулю 5. Действительно, единственно возможные вычеты квадратов целых чисел по модулю 5 — это 1 и 4, откуда немедленно следует наше утверждение. Это как раз тот пример, о котором упоминалось в доказательстве теоремы.

Второй пример иллюстрирует, что происходит, когда мы пытаемся искать корни многочлена по составному модулю. Корни

многочлена $x^2 - 170$ в $\mathbb{Z}_{385}$ – это числа $\overline{95}, \overline{150}, \overline{235}, \overline{290}$, что легко проверить. Итак, мы представили полиномиальное уравнение степени 2 с четырьмя корнями. Это, конечно, не противоречит нашей теореме, поскольку 385 составное число.

6. Псевдопростые числа

В разделе покажем, как теорема Ферма помогает выяснить, является ли данное число составным, не раскладывая его на множители. Раздел заканчивается обсуждением стратегий различных систем символьных вычислений, используемых для проверки чисел на простоту или разложимость.

6.1. Свидетель псевдопростого числа

Как утверждает теорема Ферма, для целого числа a, не делящегося на простое p, имеет место сравнение: $a^{p-1} \equiv 1\ (mod\ p)$. Предположим, что нам нужно узнать, является ли данное нечетное число n простым. Допустим также, что нам как-нибудь удалось найти целое b, не делящееся на n и удовлетворяющее условию: $b^{n-1} \not\equiv \not\equiv 1\ (mod\ n)$. В этом случае теорема Ферма говорит нам, что n не может быть простым. Такое число b будем называть свидетелем разложимости n. Итак, у нас есть метод тестирования числа на наличие нетривиальных делителей, при котором не нужно разлагать число на множители. Трудность применения этого теста заключается в том, что он не будет работать до тех пор, пока мы не найдем свидетеля; а это требует времени и везения. Но как мы дальше увидим, более вероятно, что такой свидетель будет найден, чем не найден.

Заметим, что для поиска свидетеля b нам нет нужды просматривать все целые числа. Действительно, поскольку мы работаем с сравнениями по модулю n, можно ограничить поиск числами b, лежащими в интервале $0 \leq b \leq n - 1$. Имеет смысл исключить ещё 0 (поскольку b не должно делиться на n) и 1 [ибо $1^{n-1} \equiv 1\ (mod\ n)$ для всех n]. Более того, из нечетности n вытекает сравнение: $(n-1)^{n-1} \equiv 1\ (mod\ n)$, т.е. сравнение выполнено и для $n - 1$. Итак, можно предполагать, что искомый свидетель b удовлетворяет неравенству: $1 < b < n - 1$. Прежде, чем мы будем применять этот тест, сформулируем его в виде теоремы.

Теорема 6.1 (тест на разложимость Лейбница). *Пусть* n *— нечётное натуральное число. Если найдётся такое целое число* b*, что:*

1) $1 < b < n - 1$*;*

2) $b^{n-1} \not\equiv 1\ (mod\ n)$*;*

то n — разложимое, т.е. составное число.

Напомним, что повторяющиеся единицы $R(n)$ определяются формулой

$$R(n) = \frac{10^n - 1}{9}.$$

Другими словами, это целые числа, в десятичной записи которых встречаются только 1. Мы уже видели, что $R(n)$ составное число для разложимых n [4]. Однако для простого числа 229 у нас до сих пор не было возможности определить, простое $R(229)$ или составное. Кроме того, это число состоит из более чем 200 знаков, так что раскладывать его на множители слишком утомительно. Вместо этого применим тест на разложимость с $b = 2$. С помощью системы символьных вычислений легко найти вычет $2^{R(229)-1}$ по модулю $R(229)$. Он равен

104516500584333397781753768885982835488612737233884898570848288405666898406290825536552313452374268256539145527606121567512885287283062854774198632697829520351103663852079821692412346101479040743884170069248576365931104545032921 7

и не сравним с 1 по модулю $R(229)$. Так что, $R(229)$ — составное число. Поскольку нас больше интересуют простые числа, нежели составные, резонно задать вопрос: можно ли использовать теорему Ферма для доказательства простоты чисел? Более точно допустим, что n — положительное нечетное целое число, для которого $b^{n-1} \equiv \equiv 1(mod\ n)$ при некотором целом числе b, удовлетворяющем неравенству $1 < b < n - 1$; обязательно ли n будет простым числом?

Лейбниц полагал, что ответ утвердительный. Он использовал это соображение как тест на простоту, всегда выбирая 2 в качестве b для упрощения вычислений. К сожалению, Лейбниц был не прав. Например, имеет место сравнение $2^{340} \equiv 1\ (mod\ 341)$ и, по Лейбницу, число 341 должно быть простым. Но $341 = 11 \cdot 31$ — составное. Числа, дающие ложный «положительный» результат в этом тесте, известны как псевдопростые.

Определение 6.1. *Нечетное составное натуральное число n, удовлетворяющее сравнению* $b^{n-1} \equiv 1(mod\ n)$ *для некоторого целого* b *из интервала* $(1; n-1)$*, называется псевдопростым по основанию* b*.*

Следовательно, 341 — псевдопростое по основанию 2.

Несомненно, тест Лейбница все-таки полезен, хотя и не абсолютно точен. Для малых целых, выбранных наугад, тест чаще дает правильный ответ, чем ошибается. Чтобы убедиться в этом, подсчитаем простые и псевдопростые числа по основанию 2, не превышающие какой-нибудь подходящей грани. Например, между 1 и 10^9 лежит 50847544 простых и только 5597 псевдопростых по основанию 2. Таким образом, число из этого промежутка, выдержавшее тест Лейбница, будет скорее простым, нежели псевдопростым по основанию 2.

Кроме того, мы применяли этот тест только по одному основанию, а если использовать несколько разных оснований, то количество неопределяемых составных чисел значительно уменьшится. Например, $3^{340} = 56\ (mod\ 341)$, так что число 3 (свидетель) свидетельствует о разложимости числа 341. На самом деле между 1 и 10^9 находится 1272 псевдопростых по основаниям 2 и 3 и только 685 псевдопростых по основаниям 2, 3 и 5.

Так как нам нужно примерять тест только по конечному числу оснований: $2,3, \ldots, n-2$, возникает вопрос: может ли n быть составным, оставаясь псевдопростым по всем этим основаниям? Пусть $n > 2$ предположим, что $b^{n-1} \equiv 1\ (mod\ n)$ для некоторого $1 < b < n-1$. Поскольку $b^{n-1} = b \cdot b^{n-2}$, то из предположения следует обратимость b по модулю n. А теорема обратимости утверждает, что такое возможно лишь в случае $НОД(b, n) = 1$. Поэтому, если b — составное, а один из делителей b делит n, то $b^{n-1} \not\equiv 1\ (mod\ n)$. В частности, любой делитель n свидетельствует о разложимости n. Итак, ответ на вышесформулированный вопрос отрицателен.

Какой можно сделать вывод из результатов этого раздела? Напомним, что наша цель — найти эффективный способ определения, является ли данное конкретное число простым. Если число имеет небольшой делитель, его можно найти с помощью алгоритма деления методом проб из раздела 2.3. Так что, на практике тест на разложимость следует применять только в том случае, когда

предварительные рассмотрения показали, что данное число не имеет малых делителей. Таким образом, вопрос, заданный выше, имеет небольшой практический интерес. Было бы быстрее найти делитель, чем пытаться проверить, является ли большое число n псевдопростым по всем основаниям между 2 и $n-2$.

6.2. Числа Кармайкла

Как мы показали в конце последнего раздела, составное число n не является псевдопростым по основанию b, если числа n и b имеют общий нетривиальный делитель. К сожалению, эта информация не очень полезна. Практически, чтобы ограничить объем вычислений разумными рамками, мы выделяем несколько оснований среди небольших простых чисел. И, если наименьший делитель числа b очень большой, то все выбранные основания будут взаимно просты с n. Поэтому вопрос, который нам стоило бы задать, нужно нацелить на усовершенствование метода, приведенного в конце предыдущего раздела. А именно, может ли составное нечетное число n быть псевдопростым по всем взаимно простым с ним основаниям b? Забегая вперед, скажем, что ответ: «да».

Заметим, что, если число b взаимно просто с n, то сравнение $b^n \equiv b\ (mod\ n)$ равносильно $b^{n-1} \equiv 1\ (mod\ n)$. Это позволяет поставить вопрос чуть более строго: существует ли такое нечетное натуральное n, которое, будучи составным, удовлетворяет сравнениям $b^n \equiv b\ (mod\ n)$ для всех целых b? Одно из преимуществ такой формулировки вопроса заключается в отсутствии необходимости каких-либо предварительных предположений относительно b. Первым привел пример таких чисел n математик Р. Д. Кармайкл в своей работе, опубликованной в 1912 году, поэтому они и называются числами Кармайкла [6].

Поскольку эти числа играют важную роль во многом из того, о чем нам еще предстоит говорить, хорошо бы дать их формальное определение.

Определение 6.2. *Нечетное натуральное n называется числом Кармайкла, если оно составное и* $b^n \equiv b\ (mod\ n)$ *для всех целых* b.

Конечно, достаточно проверить это сравнение только для чисел, удовлетворяющих неравенству $1 < b < n-1$, поскольку мы работаем по модулю n.

Как показал сам Кармайкл, наименьшее из чисел, открытых им, равно 561. В принципе мы можем проверить этот факт, исходя из определения. Однако даже для относительно небольших чисел такая процедура очень длинна и скучна. Действительно, чтобы доказать прямо из определения, что число 561 — число Кармайкла, нам потребуется проверять истинность сравнения $b^{561} \equiv b \ (mod\ 561)$ для $b = 2,3,4, \ldots ,559$, т.е. всего — 557 раз. Это может показаться не такой уж тяжелой работой, если у вас есть компьютер, а что делать, когда у вас такой кандидат на число Кармайкла:

349 407 515 342 287 435 050 603 204 719 587 201?

Самое подходящее время воспользоваться методами, обсуждёнными ранее. Попытаемся найти обходной путь для доказательства того, что 561 — число Кармайкла. Прежде всего, заметим, что оно довольно легко раскладывается на простые множители:

$$561 = 3 \cdot 11 \cdot 17.$$

Теперь мы хотим проверить сравнение

$$b^{561} \equiv b \ (mod\ 561) \qquad (6.1)$$

для некоторого целого b. Наша стратегия состоит в демонстрации делимости разности $b^{561} - b$ на 3, 11 и 17. Так как это различные простые числа, то лемма из главы 3 говорит нам, что тогда их произведение тоже должно делить $b^{561} - b$. Но упомянутое произведение равно 561, так что (6.1) этим будет доказано.

Чтобы сделать нашу стратегию рабочей, нужно доказать, что $b^{561} - b$ делится на каждый простой множитель числа 561. В этом нам поможет теорема Ферма. Мы приведем подробное доказательство делимости разности на 17, оставив случаи делимости на 3 и на 11 в качестве полезных упражнений.

Итак, мы хотим доказать, что 17 делит разность $b^{561} - b$ или на языке сравнений:

$$b^{561} \equiv b \ (mod\ 17). \qquad (6.2)$$

Следует рассмотреть два случая.

1. Число 17 делит b. В этой ситуации обе части уравнения (6.2) сравнимы с 0 по модулю 17, т.е. сравнение справедливо.

2. Число 17 не делит b. Тогда теорема Ферма влечет равенство: $b^{16} \equiv 1 \ (mod\ 17)$. Прежде чем применить этот факт к сравнению (6.2), нам нужно найти остаток от деления 561 на 16. Но $561 = 35 \times$ $\times 16 + 1$, поэтому

$$b^{561} \equiv (b^{16})^{35} \cdot b \equiv b \ (mod\ 17).$$

Заметим, теорема Ферма существенно сократила наши вычисления благодаря тому, что остаток от деления 561 на 16 оказался равным 1. Удачно, что остатки от деления 561 на $2(2 = 3— 1)$ и на $10(10 = 11— 1)$ тоже равны 1. Так что, вычисления, которые предстоит сделать для 3 и 11, — дословное повторение проведённых.

Успех стратегии обязан двум свойствам числа 561. Первое: деление 561 на разность каждого из его делителей и единицы даёт в остатке 1. Второе: каждый простой делитель числа 561 входит в его разложение на простые множители с кратностью 1. Что же это: нам очень повезло с выбором примера или числа Кармайкла встречаются крайне редко? Реальность оказывается ещё более удивительной. Существует бесконечно много чисел Кармайкла, и все они обладают свойствами, столь облегчившими наши вычисления с 561.

Теорема 6.2 (Корселт). *Нечётное натуральное число n является числом Кармайкла тогда и только тогда, когда для каждого его простого делителя p выполнены следующие два условия:*

1) p^2 не делит n;

2) $p - 1$ делит $n - 1$.

Доказательство. Покажем для начала, что если число n удовлетворяет условиям (1) и (2) теоремы, то оно является числом Кармайкла. Для этого мы применим стратегию, разработанную в предыдущих вычислениях с числом 561. Предположим, что p — простой делитель числа n. Покажем, что

$$b^n \equiv b \ (mod\ p).$$

Если b делится на p, то обе части в сравнении сравнимы с нулём по модулю p, и доказывать нечего. Предположим теперь, что p не делит b. Как следует из теоремы Ферма, $b^{p-1} \equiv 1 \ (mod\ p)$. Прежде чем применить это соображение к предыдущему выражению, мы должны найти остаток от деления n на $p - 1$. Но $p - 1$ делит $n - 1$ согласно условию (2) теоремы, т.е. $n - 1 = (p - 1)q$ для некоторого целого q и

$$n = (n - 1) + 1 = (p - 1)q + 1.$$

Значит,

$$b^n = (b^{n-1})^q \cdot b \equiv b \ (mod\ p),$$

где второе сравнение следует из теоремы Ферма. Суммируя все вышесказанное, получаем, что, если p простой делитель числа n, то $b^n \equiv b\ (mod\ p)$ для любого целого b. Ввиду условия (1) теоремы, $n = p_1 \cdots p_k$, где $p_1, \ldots, p_k$ — попарно различные простые числа. Мы уже видели, что $b^n - b$ делится на каждое из этих чисел, а, поскольку все они различны, то, учитывая лемму 2.1, можно сделать вывод: $b^n - b$ делится на их произведение $p_1 \cdots p_k = n$. Иначе говоря, $b^n \equiv b\ (mod\ n)$. Поскольку вычисления справедливы для любого целого b, то n является числом Кармайкла.

Теперь следует показать, что всякое число Кармайкла удовлетворяет условиям (1) и (2) теоремы. Сделаем это методом «от противного». Сначала, предположив, что n - число Кармайкла, получим, что делимость n на p^2 приводит к противоречию. Этим мы докажем, что числа Кармайкла удовлетворяют условию (1) теоремы. Так как n — число Кармайкла, противоречие получится, если мы найдём целое b, для которого $b^n \not\equiv b\ (mod\ n)$. Положим $b = p$. Тогда

$$p^n - p = p(p^{n-1} - 1).$$

Но p не делит $p^{n-1} - 1$, поэтому p^2 не может делить $p^n - p$. Иначе говоря, $p^n \not\equiv p\ (mod\ n)$. Это противоречит предположению о том, что n — число Кармайкла.

Для завершения доказательства нам осталось показать, что числа Кармайкла удовлетворяют условию (2) теоремы. Однако для этого необходима теорема о примитивных корнях, которая будет доказана далее. К сожалению, для проверки данного натурального n на принадлежность к числам Кармайкла с использованием теоремы Корселта нам нужно разложить число на простые множители, что довольно сложно в случае больших чисел. Тем не менее, довольно часто бывает так, что очень большие числа Кармайкла имеют много небольших делителей. Например, число с 36 знаками, приведённое в начале этого параграфа, является наименьшим числом Кармайкла с 20 простыми делителями. Его разложение на множители выглядит следующим образом:

$$11 \cdot 13 \cdot 17 \cdot 19 \cdot 29 \cdot 31 \cdot 37 \cdot 41 \cdot 43 \cdot 61 \cdot 71 \cdot 73 \cdot 97 \cdot 101 \cdot 109 \times$$
$$\times 113 \cdot 151 \cdot 181 \cdot 193 \cdot 641.$$

Используя программу символьных вычислений, теперь можно быстро проверить, что это действительно число Кармайкла.

6.3. Тест Миллера

В разделе 5.2 мы видели, что теорема Ферма предлагает способ проверки разложимости данного числа, не требуя поиска его делителей. Однако этот подход не всегда работает и, если очень не повезёт, может потерпеть неудачу. Далее (раздел 6.2) мы пытались объяснить, что означает «невезение» в данном контексте. Было обнаружено, что числа Кармайкла ведут себя как простые в том смысле, что мы не можем установить их разложимость методом, развитым в конце параграфа. Но этот тест можно так усовершенствовать, что его не смогут обмануть даже числа Кармайкла. Новый тест ввёл Миллер (G. L. Milller) в 1976 году.

Пусть $n > 0$ — нечётное целое. Выберем целое число b, удовлетворяющее неравенству $1 < b < n - 1$, и назовем его, как и раньше, основанием. Поскольку n — нечётное, $n - 1$ должно быть четным числом. Первый шаг теста Миллера состоит в определении такого показателя $k \geq 1$, для которого $n - 1 = 2^k q$, где q не делится на 2. Иначе говоря, мы должны найти наибольшую степень двойки, делящую $n - 1$, и соответствующее частное q.

Далее тест вычисляет вычеты по модулю n у следующей последовательности степеней:

$$b^q, b^{2q}, b^{4q}, \ldots, b^{2^{k-1}q}, b^{2^k q}.$$

Разберёмся, какими свойствами обладает эта последовательность в случае простого n. Итак, пока не будет оговорено особо, n — простое число. Теорема Ферма говорит нам, что

$$b^{2^k q} \equiv b^{n-1} \equiv 1 \,(mod\ n).$$

Значит, если n — простое, то последний вычет в последовательности всегда равен 1. Конечно, единица среди вычетов может встретиться и раньше. Пусть j — наименьший показатель, для которого $b^{2^j q} \equiv 1 \,(mod\ n)$. Если $j \geq 1$, то

$$b^{2^j q} - 1 = \left(b^{2^{j-1} q} - 1\right)\left(b^{2^{j-1} q} + 1\right).$$

По предположению, число n — простое и (так как оно делит разность квадратов $b^{2^j q} - 1$) делит либо $b^{2^{j-1} q} - 1$, либо $b^{2^{j-1} q} + 1$. С другой стороны, в силу выбора показателя j число n не может делить $b^{2^{j-1} q} - 1$. Остается только одна возможность: n — делитель числа $b^{2^{j-1} q} - 1$, т.е. $b^{2^{j-1} q} \equiv -1 \,(mod\ n)$. Эти рассуждения

показывают, что в случае простого n среди последовательности степеней:

$$b^q, b^{2q}, b^{4q}, \dots, b^{2^{k-1}q}$$

найдется, по крайней мере, одна, сравнимая с -1 по модулю n. Хорошо, но не слишком. Дело в том, что наше рассуждение основывалось на предположении: $j > 0$. Если $j = 0$, то $b^q \equiv 1\ (mod\ n)$. А у нас нет простого способа разложения числа $b^q - 1$ на множители, поскольку q нечетно. Значит, если n - простое, то с последовательностью вычетов по модулю n, о которой мы говорили, должно произойти одно из двух: либо первый же вычет равен 1, либо среди них появится $n - 1$. В противном случае (не будет ни того, ни другого) число n должно быть составным.

Последовательность вычетов, используемая в тесте Миллера, довольно легко вычисляется, потому что каждый вычет (за исключением первого) - квадрат предыдущего. В самом деле, $b^{2^j q} = \left(b^{2^{j-1}q}\right)^2$ при $j > 1$. Отсюда вытекает, что, как только в последовательности вычетов по модулю n встретится $n - 1$, все остальные вычеты будут равны 1.

У нас появился еще один тест, который позволяет нам показать разложимость данного числа, но работающий только при удачном стечении обстоятельств. Тем не менее, тест Миллера более эффективен, нежели предложенный в разделе 6.1. Чтобы понять, почему, заметим, что последовательность степеней, выписанная для псевдопростого n по основанию b, должна иметь член, сравнимый с 1 по модулю n. А так как n не простое, то существует хороший шанс, что этой степени не будет предшествовать другая, сравнимая с $n - 1$. В этом случае тест Миллера обнаружит разложимость числа. Приведем алгоритм, реализующий тест Миллера.

Алгоритм 6.1 (тест Миллера)

Ввод: нечетное натуральное n и основание b, где $1 < b < n - 1$.

Вывод: одно из двух сообщений: «n составное» или «ничего определенного сказать нельзя».

Шаг 1. Последовательно делим $n - 1$ на 2, пока не получим нечетного частного. В результате найдем положительное целое k и нечетное q, для которых $n - 1 = 2^k q$.

Шаг 2. Присвоим i нулевое значение, а r - значение вычета b^q по модулю n.

Шаг 3. Если $i = 0$ и $r = 1$ или $i > 0$, а $r = n - 1$, то вывести сообщение: «ничего определенного сказать нельзя»; в противном случае переходим к шагу 4.

Шаг 4. Увеличиваем i на 1 и заменяем r на r^2 по модулю n; переходим к шагу 5.

Шаг 5. Если $i < k$, то возвращаемся к шагу 3; в противном случае выдаем сообщение: «n составное».

В случае неопределенного сообщения возможны две ситуации: либо n простое, либо составное. К сожалению, второй случай реально встречается. Рассмотрим несколько примеров, причем начнем с хороших новостей.

Мы видели в разделе 6.1, что 341 — псевдопростое по основанию 2, так что это хороший объект для теста Миллера. Прежде всего, $340 = 2^2 \cdot 85$. Теперь нам нужно найти вычеты по модулю 341 у степеней 2 с показателями 85 и 170:

$$2^{85} \equiv 32\ (mod\ 341), 2^{170} \equiv 32^2 \equiv 1\ (mod\ 341).$$

Это позволяет тесту дать заключение: число составное.

Более наглядный пример дает нам число Кармайкла 561. Подвергнем его тесту Миллера по основанию 2. Простые вычисления показывают, что $560 = 2^4 \cdot 35$. Выпишем последовательность вычетов степеней 2 по модулю 561:

Степени	35	$2 \cdot 35$	$2^2 \cdot 35$	$2^3 \cdot 35$
Вычеты	263	166	67	1

И в этом случае тест сообщит нам о разложимости числа.

Хотя 561 — число Кармайкла, мы обнаружили, что оно составное, применив наименьшее из возможных оснований. Теперь плохие новости. Применим тест Миллера по основанию 7 к 25. Поскольку $24 = 2^3 \cdot 3$, последовательность степеней и их остатков имеет вид:

Степени	3	$2 \cdot 3$	$2^2 \cdot 3$
Вычеты	18	24	1

Так что тест здесь не может сказать ничего определенного, несмотря на то, что разложимость числа 25 очевидна. Конечно, основание 7 мы выбрали не случайно, если бы основанием было 2, то тест бы распознал в 25 составное число.

Пусть $n > 0$ — нечетное целое число и $1 < b < n - 1$. Если n составное, а тест Миллера его не распознал, то оно называется строго псевдопростым по основанию b. Вышеприведенный пример показывает, что 25 — строго псевдопростое по основанию 7. Легко увидеть, что строго псевдопростое число по основанию b является псевдопростым по этому же основанию.

Как мы уже отметили, число 25 не является строго псевдопростым по основанию 2. Наименьшее строго псевдопростое число по основанию 2 — это 2047. Более того, существует только 1282 строго псевдопростых чисел по основанию 2, лежащих между 1 и 10^9, что дает хорошее представление об эффективности данного теста. Конечно, можно применять тест Миллера по нескольким основаниям, что существенно увеличит его эффективность. Например, наименьшее строго псевдопростое одновременно по основаниям 2, 3 и 5 — это 25326001.

Более того, не существует «строго Кармайкловских чисел». Это следует из результата М. О. Рабина (Rabin).

Теорема 6.3 (Рабин). *Пусть $n > 0$ — нечетное натуральное число. Если тест Миллера, примененный к n для более чем $\frac{n}{4}$ оснований, лежащих между 1 и $n - 1$, не распознает в n составного числа, то n — простое.*

Замечание. В случае больших n применение теста Миллера по $\frac{n}{4}$ основаниям займет очень много времени. Вопреки этому замечанию, наиболее практичные тесты на простоту, как мы увидим в следующем параграфе, основываются на теореме Рабина.

6.4. Тестирование простоты. Системы символьных вычислений

Многие системы символьных вычислений имеют простую команду для проверки, является ли данное число простым. Самое удивительное при этом, что ответ выдается почти мгновенно, даже если проверяемое число было довольно большим. Так происходит из-за того, что большинство программ основаны на тесте Миллера, применяемом по большому числу оснований.

Рациональное зерно в таком использовании теста Миллера следует из теоремы Рабина. Пусть n — нечетное составное число. Основание b между 1 и $n - 1$ выберем случайным образом. Из теоремы

Рабина следует, что вероятность неопределенного результата при применении теста Миллера к n и b не превосходит

$$\frac{\frac{n}{4}}{n} = \frac{1}{4}.$$

Так что правдоподобно предположение: число n в этом случае будет составным с вероятностью $\frac{1}{4}$. Если теперь выбрать k различных оснований, то вероятность станет равной $\frac{1}{4^k}$. Поэтому, беря все больше и больше оснований, мы можем сделать вероятность сколь угодно малой.

Эти рассуждения приводят к вероятностному тесту Рабина на простоту. Поскольку он вероятностный, необходимо выбрать вероятность ошибки. Предположим, нам нужно, чтобы вероятность ошибки не превосходила ε, и пусть k — такое натуральное число, для которого $\frac{1}{4^k} < \varepsilon$. Тогда тест Рабина выбирает k оснований и применяет тест Миллера по каждому из них. Из вышеприведенных рассуждений вытекает, что в случае неопределенного ответа на тест Миллера по всем этим основаниям вероятность того, что проверяемое число все же окажется составным, будет меньше или равна $\frac{1}{4^k}$, т.е. меньше требуемого ε. Как же выбираются основания? Конечно, удобнее было бы брать основания поменьше, иначе вычисления, необходимые для теста Миллера, будут занимать слишком много времени. Как правило, выбирают первые k положительных простых чисел. Естественно, мы хотим быть абсолютно уверены в том, что числа, прошедшие тест, будут простыми. Для этого необходимо выбрать ε очень маленьким. Например, пусть $\varepsilon = 10^{-20}$. Так как $\frac{1}{4^{40}}$ — величина порядка 10^{-24}, нам нужно выбрать 40 оснований, чтобы вероятность ошибки была меньше 10^{-20}.

Теорема 6.4. *Для любого конечного набора оснований найдется бесконечно много чисел Кармайкла, строго псевдопростых по всем основаниям.*

7. Системы сравнений

В этом разделе рассмотрим метод решения систем линейных сравнений, называемый китайским алгоритмом остатков. В одном

из примеров будет показано, как этот алгоритм применяется для передачи ключа к шифру нескольким людям.

7.1. Линейные уравнения

Начнем со случая одного линейного уравнения:

$$a \cdot x \equiv b \ (mod\ n), \tag{7.1}$$

в котором n — натуральное число. В разделе 4.7 мы видели, что это уравнение легко решается, если $\text{НОД}(a, n) = 1$. По теореме обратимости, взаимная простота чисел a и n влечет обратимость $\bar{a}$ в $\mathbb{Z}_n$. Пусть $\bar{\alpha}$ — соответствующий обратный элемент. Умножая на него сравнение (7.1), получаем:

$$\alpha \cdot (a \cdot x) \equiv \alpha \cdot b \ (mod\ n).$$

Так как $\alpha \cdot a \equiv 1 \ (mod\ n)$, отсюда следует сравнение

$$x \equiv \alpha \cdot b \ (mod\ n),$$

что дает решение уравнения. В частности, если n — простое и $a \not\equiv \not\equiv 0 \ (mod\ n)$, уравнение (7.1) всегда имеет решение.

Предположим, что $\bar{a}$ не обратим в $\mathbb{Z}_n$. Напомним, что это равносильно условию $\text{НОД}(a, n) \neq 1$. Наличие решения у (7.1) означает, что найдутся элементы $x, y \in \mathbb{Z}_n$, для которых

$$a \cdot x - n \cdot y = b. \tag{7.2}$$

А это возможно только, если $\text{НОД}(a, n)$ делит b. Итак, если уравнение (7.1) имеет решение, то b делится на $\text{НОД}(a, n)$. Разумеется, в случае обратимости класса $\bar{a}$ в $\mathbb{Z}_n$ необходимое условие выполнено, поскольку тогда $\text{НОД}(a, n) = 1$.

Проверим, что обратное тоже верно. Пусть $d = \text{НОД}(a, n)$ делит b. Тогда $a = da', b = db'$ и $n = dn'$ для некоторых натуральных a', b' и n'. Сокращение равенства (7.2) на d дает:

$$a'x - n'y = b',$$

что равносильно сравнению $a'x \equiv b'(mod\ n')$. Отметим, что сравнение берется по модулю n' — делителю исходного модуля n. Более того, $\text{НОД}(a', n') = 1$, так что новое сравнение должно иметь решение. Итак, мы доказали, что, если $\text{НОД}(a, n)$ делит b, то множество решений уравнения (7.1) непусто. Суммируя вышесказанное, формулируем.

Теорема 7.1. *Уравнение (7.1) имеет решение тогда и только тогда, когда* $\text{НОД}(a, n)$ *делит* b.

Кроме того, приведенный метод решения линейных сравнений легко применим, поскольку использует только расширенный алгоритм Эвклида. Однако, когда решения будут получены, могут возникнуть вопросы к его корректности.

Пример 7.1. *Решить сравнение* $6x \equiv 4\ (mod\ 8)$.

Решение. Поскольку $\text{НОД}(6,8) = 2 \neq 1$, то $\bar{6}$ не имеет обратного в $\mathbb{Z}_8$. Если данное сравнение имеет решение, то найдутся такие целые x и y, для которых $6x - 8y = 4$. Разделим это равенство на 2: $3x - 4y = 2$, что равносильно сравнению $3x \equiv 2\ (mod\ 4)$. Но $\bar{3}$ сам себе обратен в $\mathbb{Z}_4$. Умножая последнее сравнение на 3, мы приходим к решению:

$$x \equiv 2\ (mod\ 4).$$

Это не совсем то, что нужно. Действительно, мы начинали со сравнения по модулю 8, и решение тоже необходимо найти по модулю 8, а не по модулю 4, как получено. Скорректируем решение. Как следует из полученного выражения, решение x сравнения $6x \equiv 4\ (mod\ 8)$ записывается в виде $x = 2 + 4k$ для некоторого $k \in \mathbb{Z}$. Если k четно, то $x \equiv 2\ (mod\ 8)$ — одно из решений. С другой стороны, если k нечетно, то $k = 2m + 1$ и $x = 6 + 8m$. Так что, $x \equiv 6\ (mod\ 8)$ — другое решение. Более того, поскольку k может быть либо четным, либо нечетным, существуют только эти возможности. Следовательно, уравнение $\bar{6} \cdot \bar{x} = \bar{4}$ имеет ровно два разных решения в $\mathbb{Z}_8$, а именно $\bar{2}$ и $\bar{6}$.

Это пример линейного сравнения с двумя решениями. Как мы видели в разделе 5.3, такое произошло потому, что модуль сравнения был составным.

7.2. Пример решения системы линейных уравнений

В этом параграфе мы описываем один из методов решения систем линейных сравнений. Это очень древний алгоритм. Он применялся еще в античности для решения проблем астрономии. Мы начнем с задачи, сформулированной на современном языке, которая могла бы рассматриваться древними астрономами.

Пример 7.2. *Три спутника пересекут меридиан города Лидса сегодня ночью: первый — в 1 ночи, второй — в 4 утра, а третий — в 8 утра. У каждого спутника свой период обращения. Первому на полный оборот вокруг Земли требуется 13 часов, второму — 15, а*

третьему — 19 часов. Сколько часов пройдет (от полуночи) до того момента, когда спутники одновременно пересекут меридиан Лидса?

Решение. Запишем условия задачи в терминах сравнений. Пусть x — количество часов, которые пройдут с 12 часов ночи до момента одновременного прохождения спутниками над меридианом Лидса. Первый спутник пересекает этот меридиан каждые 13 часов, начиная с часу ночи. Это можно записать как $x = 1 + 13t$ для некоторого целого t. Другими словами, $x \equiv 1(\text{mod } 13)$. Соответствующие уравнения для остальных спутников имеют вид:

$$x \equiv 4(mod\ 15), x \equiv 8\ (mod\ 19).$$

Таким образом, три спутника одновременно пересекут меридиан Лидса через x часов, если x удовлетворяет этим трем уравнениям. Следовательно, для ответа на поставленный вопрос достаточно решить систему сравнений:

$$\begin{cases} x \equiv 1(mod\ 13), \\ x \equiv 4(mod\ 15), \\ x \equiv 8(mod\ 19). \end{cases} \tag{7.3}$$

Заметим, что мы не можем складывать или вычитать отдельные уравнения системы, поскольку модули сравнений в них разные. Будем решать эту задачу, пользуясь определением, переходя от сравнений к уравнениям в целых числах. Так, сравнение $x \equiv 1(mod\ 13)$ соответствует диофантову уравнению: $x = 1 + 13t$. Заменяя x во втором сравнении системы на $1 + 13t$, получаем:

$$1 + 13t \equiv 4\ (mod\ 15) \Rightarrow 13t \equiv 3\ (mod\ 15).$$

Но 13 обратимо по модулю 15, обратный к нему элемент — это 7. Умножая последнее сравнение на 7 и переходя в нем к вычетам по модулю 15, имеем:

$$t \equiv 6(mod\ 15).$$

Значит, t может быть записан в виде: $t = 6 + 15u$ для какого-то целого u. Следовательно,

$$x = 1 + 13t = 1 + 13(6 + 15u) = 79 + 195u.$$

Заметим, что все числа вида $79 + 195u$ являются целыми решениями первых двух сравнений системы (7.3). Наконец, подставим в третье сравнение вместо x выражение $79 + 195u$:

$$79 + 195u \equiv 8\ (mod\ 19) \Rightarrow 5u \equiv 5\ (mod\ 19).$$

Ввиду обратимости остатка 5 по модулю 19, на него можно сократить и увидеть, что $u \equiv 1\ (mod\ 19)$. Переписывая это сравнение как диофантово уравнение, мы получаем $u = 1 + 19v$ для некоторого целого v. Итак, окончательно имеем:

$$x = 79 + 195u = 79 + 195(1 + 19v) = 274 + 3705v.$$

Какой отсюда можно сделать вывод относительно спутников? Напомним, что x — количество часов, которые пройдут от полуночи до момента одновременного прохождения спутников над меридианом Лидса. Поэтому нам нужно было найти наименьшее натуральное значение переменной x, удовлетворяющее системе (7.3). Мы это сделали. Поскольку решение системы: $x = 274 + 3705v$, то ответ: 274. Итак, спутники одновременно пройдут над меридианом Лидса через 274 часа после 0 часов сегодняшней ночи, что соответствует 11 дням и 10 часам. Но общее решение системы дает больше информации. Прибавляя к 274 любое кратное 3705, мы получаем другое решение системы. Иначе говоря, спутники одновременно пересекают означенный меридиан каждые 3705 часов после первого такого момента, что соответствует 154 дням и 9 часам.

В следующем разделе мы проведем детальный анализ примененного метода решения системы линейных сравнений. Заметим, что мы решали эту систему трех сравнений, рассматривая по два сравнения за раз. Действительно, сначала мы получили решение первых двух сравнений: $x = 79 + 195u$, что равносильно $x \equiv \equiv 79\ (mod\ 195)$. Для поиска решений третьего сравнения мы решаем другую систему двух уравнений, а именно:

$$\begin{cases} x = 79\ (mod\ 195), \\ x = 8\ (mod\ 19). \end{cases}$$

В общей ситуации нам предстоит решать несколько систем двух сравнений. Поэтому в следующем параграфе мы детально проанализируем алгоритм решения систем только двух сравнений.

7.3. Китайский алгоритм остатков. Простые корни

Китайский алгоритм остатков так назван потому, что впервые был найден в «Учебнике математики мастера Сана», написанном между 287 и 473 годами нашей эры.

Китайский алгоритм остатков — это простое обобщение метода, использованного при решении системы уравнений раздела 7.2.

Подробному изучению алгоритма и посвящен настоящий параграф. Рассмотрим систему:

$$\begin{cases} x \equiv a (mod\ m), \\ x \equiv b (mod\ n). \end{cases} \quad (7.4)$$

Как и в разделе 7.2, из первого сравнения следует, что $x = a + my$, где y — целое число. Подставляя вместо x во втором сравнении $a + my$, мы получаем: $a + my \equiv b\ (mod\ n)$. Другими словами,

$$my \equiv (b - a)(mod\ n). \quad (7.5)$$

Но благодаря разделу 7.1 мы знаем, что это сравнение имеет решение тогда и только тогда, когда наибольший общий делитель m и n делит $b - a$. Для уверенности в том, что это условие выполнено, достаточно предположить, что НОД$(n, m) = 1$. Предположение равносильно тому, что m имеет обратный элемент в $\mathbb{Z}_n$; скажем, $\bar{\alpha}$.

Теперь сравнение (7.5) легко решается. Умножая обе его части на α, получаем: $y \equiv \alpha(b - a)\ (mod\ n)$. Следовательно, $y = \alpha(b - -a) + nz$, где z — целое. Так как $x = a + my$, то

$$x = a + m\alpha(b - a) + mnz.$$

Но $\overline{\alpha} \cdot \overline{m} = \overline{1}$ в $\mathbb{Z}_n$ Значит, найдется такое целое β, при котором $1 — \alpha m = \beta n$. Итак,

$$x = a(1 - m\alpha) + m\alpha b + mnz = a\beta n + m\alpha b + mnz.$$

Преимущество такой записи решения в том, что α и β легко вычисляются. Действительно, $1 = \alpha m + \beta n$, так что α и β находятся расширенным алгоритмом Эвклида, примененным к m и n. В итоге, если НОД$(m, n) = 1$, то при любом целом k число $a\beta n + b\alpha m + +kmn$ является решением системы (7.4).

А сколько решений имеет такая система сравнений? Бесконечно много, если мы имеем в виду целочисленные решения. Ведь при каждом конкретном выборе z мы получаем новое решение согласно формуле, выписанной выше. Рассмотрим этот момент более подробно. Предположим, что целые x и y удовлетворяют системе (7.4). Тогда $x \equiv a\ (mod\ m)$ и $y \equiv a\ (mod\ m)$. Разность этих сравнений приводит к соотношению: $x - y \equiv 0\ (mod\ m)$, т.е. $x - y$ делится на m. Проделывая то же самое со вторым сравнением, получаем, что $x - y$ делится на n. А ввиду взаимной простоты m и n и леммы 2.1, это доказывает делимость $x - y$ на mn. Значит, если x и y — целые решения системы (7.4), то $x \equiv y\ (mod\ mn)$. Поэтому, хотя система и имеет бесконечно много решений, все они сравнимы друг с другом

по модулю mn. Другими словами, система имеет только одно решение в $\mathbb{Z}_{mn}$. Но нельзя забывать, что все наши выводы справедливы только благодаря предположению: НОД$(m, n) = 1$. Сведем все полученные рассуждения в следующую теорему.

Теорема 7.2 (Китайская теорема об остатках). *Пусть m и n — взаимно простые натуральные числа. Система линейных уравнений*

$$\begin{cases} x \equiv a(mod\ m), \\ x \equiv b(mod\ n) \end{cases}$$

имеет одно и только одно решение в $\mathbb{Z}_{mn}$.

Существует геометрическая интерпретация этой теоремы. Предположим, что у вас есть таблица с $m \cdot n$ клетками. Столбцы таблицы пронумерованы элементами $\mathbb{Z}_m$, а строки — элементами $\mathbb{Z}_n$. В клетку таблицы, стоящую на пересечении столбца, соответствующего $\bar{a} \in \mathbb{Z}_m$, и строки, соответствующей $\bar{b} \in \mathbb{Z}_n$, поместим целое число x, удовлетворяющее следующим условиям:

$$0 \le x \le m \cdot n - 1, x \equiv a(mod\ m), x \equiv b(mod\ n).$$

Скажем, что соответствующая клетка таблицы имеет координаты $(\bar{a}, \bar{b})$. Поскольку $0 \le x \le m \cdot n - 1$, мы можем считать число x представителем класса сопряженности по модулю $m \cdot n$, т.е. реально x представляет некоторый класс $\bar{x} \in \mathbb{Z}_{mn}$.

Что говорит об этой таблице китайская теорема об остатках? Предположим, что НОД$(m, n) = 1$. Из теоремы следует, что каждая клетка таблицы соответствует в точности одному целому, заключенному между 0 и $m \cdot n - 1$, т.е. одному классу из $\mathbb{Z}_{mn}$. Таким образом, в разных клетках стоят разные числа и наоборот. Еще раз напомним, что модули у нас взаимно просты. Приведем таблицу для $m = 4$ и $n = 5$.

	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{5}$	$\overline{10}$	$\overline{15}$
$\bar{1}$	$\overline{16}$	$\bar{1}$	$\bar{6}$	$\overline{11}$
$\bar{2}$	$\overline{12}$	$\overline{17}$	$\bar{2}$	$\bar{7}$
$\bar{3}$	$\bar{8}$	$\overline{13}$	$\overline{18}$	$\bar{3}$
$\bar{4}$	$\bar{4}$	$\bar{9}$	$\overline{14}$	$\overline{19}$

Заметим, что таблица соответствует прямому произведению $\mathbb{Z}_4 \times \mathbb{Z}_5$. С первого взгляда может показаться, что для заполнения

таблицы необходимо решить 20 систем линейных сравнений. Но китайская теорема об остатках подсказывает заполнять таблицу «в обратном порядке». Для целого числа x между 0 и $mn-1$ найдем место в таблице, вычисляя его вычеты по модулю m и n. Например, в нашем случае вычет 14 по модулю 4 равен 2, а по модулю 5 - 4. Поэтому ему соответствует клетка с координатами $(\overline{2},\overline{4})$.

Фактически существует возможность заполнить всю таблицу целиком, не производя вычислений для отдельных чисел! Чтобы понять, как, вспомним, что существует геометрическая интерпретация множества $\mathbb{Z}_4$: четыре точки, расположенные на равных расстояниях друг от друга вдоль окружности, каждая из которых представляет один класс в $\mathbb{Z}_4$. Похожая картинка есть и для $\mathbb{Z}_5$.

В действительности нашу таблицу можно интерпретировать как плоскую карту или план, представляющий некую расположенную в пространстве двумерную поверхность. Чтобы найти эту поверхность, поступаем следующим образом. Поскольку классы $\mathbb{Z}_4$ (горизонтальные координаты) можно считать расположенными вдоль окружности, мы склеим правую сторону таблицы с левой. Получится цилиндр. Но классы $\mathbb{Z}_5$ (вертикальные координаты) также удобно представлять точками на окружности. Поэтому верх таблицы нужно склеить с низом. В результате получится поверхность, которая называется тором, напоминающая по форме бублик или баранку.

Вернемся к задаче о заполнении таблицы. Поскольку числа 0, 1, 2 и 3 меньше и четырех, и пяти, они совпадают со своими вычетами по обоим этим модулям. Поэтому нам не нужно делать каких-либо вычислений для определения их координат, и мы сразу можем поместить их в таблицу:

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$			
$\overline{1}$		$\overline{1}$		
$\overline{2}$			$\overline{2}$	
$\overline{3}$				$\overline{3}$
$\overline{4}$				

Заметим, что, расставляя по очереди эти классы по клеткам таблицы, мы, начав с левого верхнего угла таблицы, переходили каждый раз на одну клетку вправо и одну вниз. А, проставив четыре

первых класса, уперлись в правую границу таблицы. Если бы в таблице был еще один столбец, то, придерживаясь сформулированного правила, мы поместили бы в нем 4, но на одну строчку ниже 3, т.е. в последней строке. Однако у нас нет этого лишнего столбца. Воспользуемся геометрической интерпретацией таблицы. Склеив ее вертикальные границы, мы видим, что первый левый столбец таблицы можно считать идущим сразу за ее последним правым столбцом. В терминах несклеенной таблицы это означает, что мы должны переместиться с последнего столбца на первый, одновременно спустившись на одну строчку вниз. Следовательно, 4 нужно поставить на пересечение первого столбца и последней строки таблицы:

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$			
$\overline{1}$		$\overline{1}$		
$\overline{2}$			$\overline{2}$	
$\overline{3}$				$\overline{3}$
$\overline{4}$	$\overline{4}$			

Кажется, у нас появилась новая проблема: мы дошли до нижней границы и опять не можем двигаться дальше. Но в силу геометрической картинки нижняя и верхняя границы таблицы склеены, и мы можем перейти к первой строке таблицы. Только теперь нужно сместиться на один столбец вправо относительно предыдущего положения. Сделав это, в нашем случае мы получим:

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{5}$		
$\overline{1}$		$\overline{1}$		
$\overline{2}$			$\overline{2}$	
$\overline{3}$				$\overline{3}$
$\overline{4}$	$\overline{4}$			

Мы можем повторять этот процесс, пока не заполним все клетки таблицы.

7.4. Китайский алгоритм остатков. Общий случай

Мы в деталях проанализировали решение системы линейных сравнений для взаимно простых модулей, потому что именно этот случай встретится нам в следующих разделах. Однако китайский

алгоритм остатков можно также использовать и при решении систем, у которых модули не взаимно простые. Но в этом случае при решении линейных сравнений требуется предельная внимательность, возрастающая с каждым шагом алгоритма. Достаточно продемонстрировать один пример. Рассмотрим систему:

$$\begin{cases} x \equiv 3 \pmod{12}, \\ x \equiv 19 \pmod{8}. \end{cases}$$

Из первого сравнения мы получаем: $x = 3 + 12y$ для некоторого целого y. Подставляя найденное выражение для x во второе сравнение системы, имеем: $12y \equiv 16\ (mod\ 8)$. Так как НОД$(12,8) = 4$ делит 16, последнее сравнение должно иметь решение. Действительно, его целочисленный эквивалент имеет вид: $12y - 8z = 16$. Разделим это равенство на 4: $3y - 2z = 4$, т.е. $3y \equiv 4\ (mod\ 2)$. Но $3 \equiv 1\ (mod\ 2)$, а $4 \equiv 0\ (mod\ 2)$, так что $y \equiv 0\ (mod\ 2)$. Следовательно, $y = 2k$ для некоторого целого k. Наконец, подставляя $2k$ вместо y в равенство $x = 3 + 12y$, находим $x = 3 + 24k$. Итак, данная система имеет единственное решение по модулю 24. Однако $8 \cdot 12 = 96$. Так какое же отношение число 24 имеет к модулям 8 и 12? Число 24 – наименьшее общее кратное этих чисел.

Для любой пары не взаимно простых модулей всегда можно написать систему сравнений, не имеющую ни одного решения. С точки зрения геометрической интерпретации (7.4) это означает, что если модули обладают общим нетривиальным делителем, то в соответствующей таблице всегда останутся незаполненные клетки.

Еще раз повторим, нет никакой необходимости делать какие-либо вычисления для заполнения таблицы. Просто мы должны проставлять числа 0,1, ..., начиная с левой верхней клетки, сдвигаясь каждый раз на один столбец вправо и одну строку вниз, не забывая «перепрыгивать» справа налево и снизу вверх при приближении к соответствующей границе таблицы.

Заполняя таким образом таблицу для не взаимно простых модулей, мы вернемся в клетку с координатами $(\bar{0}, \bar{0})$, не перебрав всех $mn - 1$ чисел. Это объясняет, почему некоторые клетки таблицы останутся пустыми. Для $m = 4$ и $n = 6$ таблица выглядит следующим образом:

	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$		$\bar{6}$	
$\bar{1}$		$\bar{1}$		$\bar{7}$
$\bar{2}$	$\bar{8}$		$\bar{2}$	
$\bar{3}$		$\bar{9}$		$\bar{3}$
$\bar{4}$	$\bar{4}$		$\overline{10}$	
$\bar{5}$		$\bar{5}$		$\overline{11}$

7.5. Применения китайской теоремы об остатках

Для систем сравнений, число уравнений в которых больше двух, есть свой вариант китайской теоремы об остатках. Мы ее сформулируем без доказательства, поскольку это просто еще одно приложение китайского алгоритма остатков. Сначала определение.

Определение 7.1. *Натуральные числа* $n_1, \dots, n_k$ *называются попарно взаимно простыми, если* $\text{НОД}(n_i, n_j) = 1$ *для любой пары чисел* $i \neq j$.

Например, три числа n_1, n_2, n_3 — попарно взаимно простые, если $\text{НОД}(n_1, n_2) = 1, \text{НОД}(n_1, n_3) = 1$ и $\text{НОД}(n_2, n_3) = 1$.

Теорема 7.3 (Китайская теорема об остатках). *Пусть* $n_1, \dots, n_k$ *- попарно взаимно простые натуральные числа. Тогда всякая система*

$$\begin{cases} x \equiv a_1 \ (\text{mod } \text{n}_1), \\ x \equiv a_2 \ (\text{mod } \text{n}_2), \\ \quad \dots \\ x \equiv a_\text{k} \ (\text{mod } \text{n}_\text{k}) \end{cases}$$

имеет одно и только одно решение в $\mathbb{Z}_{n_1 \dots n_k}$.

Применив эту версию теоремы, мы можем упростить вычисление вычетов степеней по модулю n, если известно его разложение на простые множители. Мы будем также предполагать, что каждый простой множитель входит в это разложение с кратностью 1, потому что именно в этом случае метод наиболее эффективен.

Допустим, что разложение имеет вид: $n = p_1 \cdots p_k$, где $p_1, \dots, p_k$— простые числа. Для целых a и m мы сначала находим вычет a^m по каждому модулю p_i. Если простые множители не слишком велики, то вычисления будут очень быстрыми даже для

больших m и a, поскольку нам помогает это делать теорема Ферма. Предположим, что мы уже сделали эти вычисления, причем

$$a^m \equiv r_1 (mod\ p_1), 0 \le r_1 < p_1,$$
$$a^m \equiv r_2 (mod\ p_2), 0 \le r_2 < p_2,$$
$$\dots$$
$$a^m \equiv r_k (mod\ p_k), 0 \le r_k < p_k.$$

Поэтому для определения вычета a^m по модулю n нам нужно только решить систему сравнений:

$$\begin{cases} x \equiv \mathrm{r}_1 \pmod{\mathrm{p}_1}, \\ x \equiv \mathrm{r}_2 \pmod{\mathrm{p}_2}, \\ \quad \dots \\ x \equiv \mathrm{r}_\mathrm{k} \pmod{\mathrm{p}_\mathrm{k}}. \end{cases}$$

Заметим, что модули системы — различные простые числа, поэтому они попарно взаимно просты. Значит, по китайской теореме об остатках, система всегда имеет решение, скажем, $r, 0 \le r < n$. Более того, любые два таких решения сравнимы по модулю $p_1 \cdots p_k = n$. Так как a^m — тоже решение системы, имеем $a^m \equiv \equiv r(mod\ n)$. Следовательно, r — вычет a^m по модулю n.

Пример 7.3. *Найти вычет числа* 2^{6754} *по модулю 1155.*

Решение. Обозначим $x = 2^{6754}$. Раскладывая 1155 на множители, найдем: $1155 = 3 \cdot 5 \cdot 7 \cdot 11$. Применив теорему Ферма к каждому из этих простых чисел, получим:

$$2^{6754} = 1\ (mod\ 3),$$
$$2^{6754} = 4\ (mod\ 5),$$
$$2^{6754} = 2\ (mod\ 7),$$
$$2^{6754} = 5\ (mod\ 11).$$

Таким образом, для нахождения числа x следует решить систему:

$$\begin{cases} x = 1 \pmod 3, \\ x = 4 \pmod 5, \\ x = 2 \pmod 7, \\ x = 5 \pmod{11} \end{cases}$$

с помощью китайского алгоритма остатков. Первое сравнение приводит к выражению: $x = 1 + 3y$, поэтому второе сравнение дает:

$$1 + 3y \equiv 4\ (mod\ 5) \Rightarrow y \equiv 1\ (mod\ 5).$$

Число 3 обратимо по модулю 5 и на него можно сократить обе части сравнения. Итак, $x = 4 + 15z$.

Подставляя вместо x его выражение через z в третье сравнение и решая его, получаем: $x = 79 + 105t$. Наконец, решая четвертое сравнение относительно i, мы имеем $t \equiv 6\ (mod\ 11)$. Значит, $x = = 709 + 1155u$ и 709 — искомый вычет числа 2^{6754} по модулю 1155.

Следующий пример моделирует процедуру допуска к секретным сведениям, основанную на китайской теореме об остатках. Представьте себе следующую ситуацию. Подвал банка должен открываться каждый день. В банке служат пять старших кассиров, имеющих доступ к подвалу. По причинам безопасности руководство банка предпочитает систему, требующую присутствия хотя бы двух из этой пятерки для возможности открыть подвал. Проблема в том, чтобы подвал могли открыть любые два старших кассира. Рассмотрим эту проблему в более общем виде.

Пример 7.4. *Для того, чтобы открыть подвал банка, необходимо знать код, который является натуральным числом s. Необходимо распределить этот код между n сотрудниками банка так, чтобы каждый из них знал что-то об s. Открыть подвал должно быть невозможно, если в банке присутствуют менее k сотрудников кассиров, где $k \geq 2$ — натуральнее число, меньшее n.*

Решение. Назовем такую частичную информацию фрагментом кода. Открыть подвал невозможно, если в банке присутствуют меньше k сотрудников банка. Это условие реализуется распределением информации о коде таким образом, что:

1) число s заведомо определяется, если известно k или более фрагментов;

2) число s трудно определимо, если известно менее k фрагментов.

Фрагменты кода, известные каждому из сотрудников банка, — это в действительности элементы множества $\mathbb{S}$, состоящего из n упорядоченных пар натуральных чисел. Чтобы построить $\mathbb{S}$, выберем сначала множество $\mathbb{L}$ из n попарно взаимно простых чисел. Пусть N — произведение наименьших k из них, а M — произведение $k - -1$ наибольших. Будем говорить, что k является порогом для $\mathbb{L}$, если $M < N$. Из этого условия следует, что произведение любых k (или более) элементов из $\mathbb{L}$ всегда больше, чем N, а произведение $k - 1$ (или менее) его элементов — всегда меньше M.

Предположим, код s выбран так, что $M < s < N$, а множество $\mathbb{S}$ состоит из пар (m, s_m), где $m \in \mathbb{L}$, s_m — вычет числа s по модулю m. Эти пары и являются теми фрагментами кода, которые сообщаются старшим кассирам. Тот факт, что множество $\mathbb{L}$ имеет порог $k \geq 2$, обеспечивает неравенство $s > m$ для каждого $m \in \mathbb{L}$. В частности, $s_m < s$ для любого $m \in \mathbb{L}$.

Что произойдет, если k или более сотрудников банка находятся в банке? В этом случае известны $t(\geq k)$ пар из множества $\mathbb{S}$. Обозначив эти пары через $(m_1, s_1), \ldots, (m_t, s_t)$, рассмотрим систему сравнений:

$$\begin{cases} x \equiv s_1 (mod\ m_1), \\ x \equiv s_2 (mod\ m_2), \\ \ldots \\ x \equiv s_t (mod\ m_t). \end{cases} \tag{7.6}$$

Элементы множества $\mathbb{L}$ попарно взаимно просты. Значит, по китайской теореме об остатках, эта система имеет решение $0 \leq x_0 < m_1 \cdots m_t$. Но совпадает ли x_0 с s? Это как раз та причина, по которой мы накладывали требование: $\mathbb{L}$ имеет порог k. Поскольку $t > k$, то наше требование влечет:

$$m_1 \cdots m_t \geq N > s.$$

Но s тоже удовлетворяет системе (7.6), и по китайской теореме об остатках

$$x_0 \equiv s\ (mod\ m_1 \cdots m_t).$$

А так как s и x_0 — натуральные числа, меньшие $m_1 \cdots m_t$, то $s = x_0$.

Предположим, теперь, что в банке находится менее k сотрудников банка. Несмотря на то, что t теперь меньше k, мы все равно сможем решить систему (7.6). Пусть x_0 — наименьшее неотрицательное решение, тогда $0 \leq x_0 < m_1 \cdots m_t$. Но произведение меньшего, чем k количества элементов из $\mathbb{L}$ всегда меньше M; так что, $x_0 < M < s$. Следовательно, решения системы недостаточно для восстановления кода s. Однако как x_0, так и s — решения системы (7.6), поэтому

$$s = x_0 + y \cdot (m_1 \cdots m_t),$$

где y — некоторое натуральное число. Неравенство

$$x_0 < M < s < N$$

влечет

$$\frac{M - x_0}{m_1 \cdots m_t} \leq y \leq \frac{s - x_0}{m_1 \cdots m_t} \leq \frac{N - x_0}{m_1 \cdots m_t}.$$

Приходим к выводу: если $< k$, то для восстановления кода s нам предстоит отыскивать недостающий множитель y среди более чем

$$d = \left\lfloor \frac{N - M}{M} \right\rfloor$$

целых чисел. Выбрав модули так, чтобы d оказалось очень большим, мы сделаем задачу поиска у практически нерешаемой.

Для завершения разбора задачи осталось осветить один вопрос: можно ли найти множество $\mathbb{L}$, удовлетворяющее всем необходимым требованиям? Ответ на него положителен, но нуждается в сведениях о распределении простых чисел, которые выходят за рамки данного курса.

Сформулируем результат. Для построения системы защиты требуются начальные данные: число n сотрудников, имеющих доступ в подвал банка, и наименьшее число k из них, присутствия которых в банке достаточно для открытия подвала. Первое число определяет размер множества $\mathbb{L}$, а второе — его порог k. Далее нам нужно подобрать множество $\mathbb{L}$ из n элементов с порогом k (эту часть конструкции мы подробно не обсуждали) и вычислить M и N, определенные выше. Напомним, что $\mathbb{L}$ нужно выбирать с таким расчетом, чтобы число d, о котором мы говорили, было как можно больше; в противном случае код может быть разгадан простым перебором. Код s — натуральное число, которое выбирается лежащим между M и N. Теперь можно вычислить элементы множества $\mathbb{S}$ и сообщить их сотрудникам. Конечно, безопасность этой схемы зависит от того, насколько велико k, уменьшающее вероятность, что одновременно k кассиров из одного банка окажутся нечестными. Если это все-таки произойдет, то нам придется утешать себя мыслью, что не существует систем с абсолютной безопасностью.

Пример 7.5. *Допустим, что в банке работают 5 сотрудников банка и из соображений безопасности, по крайней мере, двое из них должны присутствовать при открытии подвала. Найти фрагменты кода.*

Решение. По условиям задачи множество $\mathbb{L}$ должно состоять из пяти элементов, а его порог равен 2. Выбрав элементы $\mathbb{L}$ среди малых простых чисел, получим:

$$\mathbb{L} = \{11,13,17,19,23\}.$$

Произведение двух наименьших чисел этого множества равно $N = 11 \cdot 13 = 143$. С другой стороны, поскольку $k = 2$, произведение $k - 1$ наибольших простых из $\mathbb{L}$ в действительности равно его максимальному элементу. Таким образом, $M = 23$ и L имеет порог 2. Код s может быть любым целым числом, лежащим между 23 и 143. Пусть $s = 30$. Тогда

$$\mathbb{S} = \{(11,19), (13,17), (17,13), (19,11), (23,7)\}.$$

Наконец, что будет, если в банке присутствуют сотрудники банка с фрагментами (17,13) и (23,7)? Код из их фрагментов получается как наименьшее число, удовлетворяющее системе:

$$\begin{cases} x \equiv 13 \ (mod\ 17), \\ x \equiv 7 \ (mod\ 23). \end{cases}$$

Легко увидеть, что таким числом будет 30. Этот код корректен, он позволяет открыть подвал.

8. Группы

В этом разделе вводится понятия «группа» и «подгруппа» и доказывается теорема Лагранжа. Группы — один из «таксономических классов», которые используются при классификации математических структур с общими характеризующими свойствами. Группа — это абстрактный объект, наделенный определенными (групповыми) свойствами. Необходимость наделения этими свойствами объясняется кругом решаемых задач. Для лучшего понимания приведем примеры, которые включают в себя группы симметрии многоугольников и группы обратимых целых чисел по модулю n. Именно группы обратимых целых чисел позволяют построить набор тестов на простоту.

8.1. Термины и определения

Группа обладает двумя основными составляющими: непустым множеством и операцией, определенной на этом множестве. Обозначим множество буквой $\mathbb{G}$, а операцию — символом «$\star$». Тогда группа $\langle \mathbb{G}, \star \rangle$. Под операцией мы понимаем правило, согласно которому любым двум элементам a и b множества $\mathbb{G}$ ставится в соответствие третий элемент из $\mathbb{G}$, обозначаемый символом $a \star b$.

В математике довольно часто изучаются множества, снабженные подобной операцией. Наиболее известные примеры: натуральные числа с операцией сложения, целые числа с операцией

сложения, рациональные числа с операцией умножения, векторы в трехмерном пространстве с операцией векторного умножения. Однако не любое множество с операцией является группой.

Определение 8.1. *Множество $\mathbb{G}$ с операцией «$\star$» называется группой только в том случае, если операция удовлетворяет следующим условиям:*

1) ассоциативность: для любых элементов $a, b, c \in \mathbb{G}$

$$a \star (b \star c) = (a \star b) \star c;$$

2) единичный элемент: существует такой элемент $e \in \mathbb{G}$, что при любом $a \in \mathbb{G}$

$$a \star e = e \star a = a;$$

3) обратный элемент: для каждого $a \in \mathbb{G}$ найдется элемент $a' \in \in \mathbb{G}$, называемый обратным к элементу a, удовлетворяющий соотношению

$$a \star a' = a' \star a = e.$$

Заметим, что определение не предполагает выполнения коммутативности операции, т.е. равенства $a \star b = b \star a$ для любой пары $a, b \in \mathbb{G}$. Те группы, у которых групповая операция коммутативна, называются абелевыми.

Нередко множества, снабженные операцией, не являются группами. Например, сложение натуральных чисел $\mathbb{N}$ ассоциативно и имеет 0 в качестве единичного элемента. Однако единственное обратимое натуральное число (т.е. число, имеющее обратный элемент по сложению) — это 0, поскольку отрицательные числа не принадлежат множеству $\mathbb{N}$; значит, натуральные числа не образуют группу.

Множество векторов трехмерного пространства с векторным произведением — более показательный пример. Оно не является группой потому, что векторное произведение не ассоциативно. Более того, скалярное произведение векторов вообще не является групповой операцией: результатом скалярного умножения двух векторов является число, а не вектор.

С другой стороны, группы встречаются среди наиболее известных множеств с операциями. Например, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ и $\mathbb{C}$, множества целых, рациональных, действительных и комплексных чисел - группы относительно операции сложения. Ввиду того, что только два целых числа (± 1) обратимы относительно умножения, множество $\mathbb{Z}$ не является группой по умножению, так же, как и $\mathbb{Q}$, $\mathbb{R}$, и $\mathbb{C}$ из-за

невозможности деления на 0. Однако, если исключить 0 из этих множеств, то они станут группами. Таким образом, $\mathbb{Q} \setminus \{0\}$, $\mathbb{R} \setminus \{0\}$ и $\mathbb{C} \setminus \{0\}$ — группы относительно умножения.

Для любого положительного числа n множество $\mathbb{Z}_n$ образует группу по сложению. Множество квадратных $n \times n$ матриц с вещественными коэффициентами и операцией сложения матриц — тоже группа, в то время как обратимые матрицы (т.е. матрицы с ненулевым определителем) образуют группу относительно умножения матриц. Заметим, что последний пример определяет неабелеву группу, ибо произведение матриц некоммутативно.

Определение 8.2. *Число элементов в группе называется ее порядком.*

Все группы в приведенных примерах бесконечны, за исключением $\mathbb{Z}_n$, чей порядок равен n. Другая хорошо известная конечная группа — это множество $\{-1,1\}$ с операцией умножения целых чисел; ее порядок равен 2. Обратите внимание, что, говоря о «группе», фактически указывают на некотором множество и операцию, в ней введенную. Это общий подход, и в дальнейшем будем его придерживаться. Многие результаты этого раздела будут относиться к «общим» группам. Поэтому во избежание недоразумений удобно использовать универсальный символ «$\star$» для обозначения операции в общей группе. Однако при этом мы будем употреблять термины «умножить» и «умножение», даже если «$\star$» обозначает совсем другую операцию.

8.2. Симметрии

Одно из наиболее важных приложений групп — изучение симметрии. Можно было бы даже сказать, что группы — это интерпретация понятия «симметрии» на языке математики.

В современном математическом языке термин «симметрия» закреплен за двумя разными понятиями. С первым из них знакомятся в средней школе. Это «осевая симметрия» и «центральная симметрия». Второе понятие обсуждается в пособии. Во избежание недоразумений отметим, что, если употребляется слово «симметрия» в его первом значении, ему всегда будет предшествовать прилагательное «осевая» или «центральная».

Поэтому нет ничего удивительного в том, что группы играют ключевую роль во многих дисциплинах, для которых симметрии

носят основополагающий характер: в геометрии, кристаллографии геологии, физике. В геометрии под симметрией понимают такое преобразование, которое, будучи применено к геометрическому объекту (фигуре), не меняет его внешнего вида. Лучшим способом представить это может быть следующий. Допустим, что есть некоторая геометрическая фигура, над которой выполняются операции или действия. Если нет возможности определить, были применены к фигуре преобразования или нет, то такое преобразование называется симметрией.

Рассмотрим простейший пример. Попытаемся найти все симметрии равностороннего треугольника. Прежде всего, у нас есть три поворота вокруг центра треугольника против часовой стрелки: на углы в $120°, 240°$ и $360°$. Последний из них совпадает с поворотом на $0°$. Кроме того, есть еще три осевых симметрии относительно биссектрис (высот, медиан) треугольника. Ясно, что перечисленные шесть преобразований удовлетворяют критерию, сформулированному выше, и, как следствие, относятся к симметриям правильного треугольника. Более того, можно показать, что других симметрий у треугольника нет.

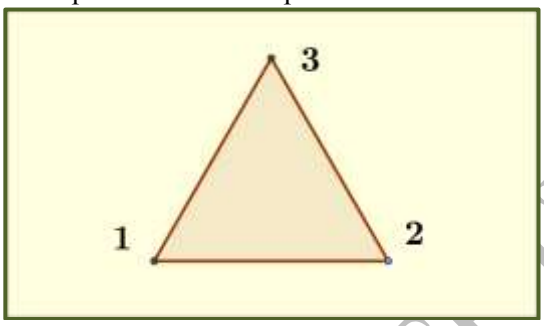
Множество действий (элементов) описано. Оно включает три поворота и три осевые симметрии. Если же представлять симметрию как преобразование совокупности точек, образующих треугольник, то естественный кандидат на операцию — композиция преобразований, т.е. симметрии. Поскольку композиция отображений ассоциативна, первое условие, накладываемое на операцию в группе, очевидно, выполнено. Роль единичного элемента играет поворот на $0°$, который на самом деле вообще ничего не делает с треугольником.

Нетрудно определить и обратный. Обратным к повороту на $120°$ будет поворот на $240°$ и наоборот, поскольку выполняется равенство $120 + 240 = 360$, а поворот на $360°$, по существу, то же самое, что поворот на $0°$.

Каждая из осевых симметрий очевидным образом обратна сама себе. Итак, все симметрии треугольника, описанные выше, имеют обратные элементы. Поэтому множество всех симметрий равностороннего треугольника с операцией композиции является группой порядка 6, которую обычно обозначают символом $\mathbb{D}_3$.

Пронумеруем вершины равностороннего треугольника, как показано на рисунке: 1 и 2 — вершины при основании, а 3 - вершина,

противоположная основанию. Нумерация вершин дает возможность описать симметрии треугольника как перестановки вершин. Например, при повороте на 120° каждая вершина переходит в соседнюю по направлению против часовой стрелки.



Правильный треугольник

Существует очень удобное обозначение для записи описанного правила - перестановка:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

В первой строке указано первоначальное состояние вершин треугольника. Вторая строка – конечное положение вершины после преобразования. Приведенная запись показывает, что преобразование треугольника таково, что вершина 1 переходит на место, занимавшееся ранее вершиной 2, вершина 2 переходит на место третьей, а вершина 3 становится на место первой. В верхней строчке нашего обозначения преобразований треугольника всегда будут стоять 1, 2 и 3 в стандартном порядке; а в нижней строчке мы будем записывать места, на которых окажутся соответствующие вершины после применения преобразования к треугольнику. Заметим, что места носят номера вершин, которые занимали их до преобразования.

Рассмотрим другой пример. Осевую симметрию относительно биссектрисы утла при вершине 3 можно записать в виде:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Если через ρ обозначить поворот на 120°, то поворот на 240° можно записать как

$$\rho^2 = \rho \cdot \rho,$$

а $\rho^3 = e$ — единичный элемент группы - поворот на 360°. Пусть σ обозначает любую из осевых симметрий. Тогда $\sigma^2 = e$. Мы хотели бы сопоставить «произведение» $\sigma \cdot \rho$ и соответствующее ему преобразование треугольника. Заметим, что $\sigma \cdot \rho$ не может равняться ρ^2. Действительно, «умножая» с правой стороны предполагаемое равенство $\sigma \cdot \rho = \rho^2$ на ρ^2 и учитывая соотношение $\rho^3 = e$, мы получаем: $\sigma = \rho$, т.е. противоречие. Аналогично можно показать, что $\sigma \times \rho \neq e, \sigma \cdot \rho \neq \rho$. Значит, $\sigma \cdot \rho$ не поворот, а какая-то из осевых симметрий. Однако $\sigma \cdot \rho \neq \rho$, поскольку предположение $\sigma \cdot \rho = \rho$ влечет ложное равенство $\rho = e$. Итак, $\sigma \cdot \rho$ — осевая симметрия, отличная от σ.

Пусть осевая симметрия σ, которую мы только что рассматривали, оставляет неподвижной вершину 3. Чтобы подчеркнуть этот факт, обозначим ее через σ_3. Вернемся к нашему вопросу. Поворот ρ переводит вершину 1 на место вершины 2, в то время как $\sigma_3 \cdot \rho$ перемещает вершину 2 на место 1. Таким образом, композиция этих преобразований оставляет вершину 1 на месте, т.е. $\sigma_3 \cdot \rho(1) = 1$. Обозначив символом σ_1 осевую симметрию, оставляющую неподвижной вершину 1, получим соотношение: $\sigma_3 \cdot \rho = \sigma_1$.

Мы могли бы вычислить $\sigma_3 \cdot \rho$ другим способом, используя обозначения перестановок вершин, введенные выше. Так,

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \rho = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Прежде чем выполнять вычисления, заметим: запись $\sigma_3 \times \times \rho(1)$ означает, что ρ применяется к 1 раньше σ, т.е.

$$1 \overset{\rho}{\to} 2 \overset{\sigma_3}{\to} 1.$$

Следовательно,

$$\sigma_3 \rho = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \sigma_1.$$

Применяя только основные свойства групп, мы можем, опираясь на доказанное равенство $\sigma_3 \cdot \rho = \sigma_1$, вычислить несколько соотношений между элементами $\mathbb{D}_3$, используя только определение группы. Например, пусть $\mathbb{G}$ — группа с операцией $\star$. Если $x, y \in \mathbb{G}$, то обратным элементом к произведению элементов $x \star y$ будет $y' \star \star x'$. Для проверки этого утверждения достаточно перемножить элементы:

$$(x \star y) \star (y' \star x') = x \star (y \star y') \star x' = x \star e \star x' = x \star x' = e.$$

Учитывая доказанный простой факт из теории групп и сохраняя штрих для обозначения обратных элементов, имеем:

$$(\sigma_3 \cdot \rho)' = \rho' \cdot \sigma_3' = \rho^2 \cdot \sigma_3.$$

Но мы уже установили тождество: $\sigma_3 \cdot \rho = \sigma_1$. А так как $\sigma_1^2 = e$, заключаем, что $\rho^2\sigma_3 = \sigma_1$. Поэтому

$$\sigma_3 \cdot \rho = \sigma_1 \neq \rho^2 \cdot \sigma_3 \neq \rho \cdot \sigma_3.$$

В частности, $\mathbb{D}_3$ не абелева группа.

Равенство $\sigma_3\rho = \sigma_1$ влечет много других соотношений. Умножая его слева на σ_3 и учитывая тождество $\sigma_3^2 = e$, получаем: $\rho = \sigma_3\sigma_1$ в то время как умножение его справа на ρ^2 дает: $\sigma_3 = \sigma_1\rho^2$. Заметим, что, поскольку операция в $\mathbb{D}_3$ не коммутативна, мы должны оговаривать, с какой стороны умножаем равенство на элемент.

Продолжая эти вычисления достаточно долго, мы можем заполнить таблицу умножения для $\mathbb{D}_3$. В общей ситуации таблица умножения конечной группы представляет собой таблицу, строки и столбцы которой «пронумерованы» элементами группы, и, если операция в группе обозначена «$\star$», то в клетке таблицы, находящейся на пересечении строки x и столбца y, стоит произведение $x \star y$. Таблица умножения группы $\mathbb{D}_3$ имеет вид:

	e	ρ	ρ^2	σ_1	σ_2	σ_3
e	e	ρ	ρ^2	σ_1	σ_2	σ_3
ρ	ρ	ρ^2	e	σ_3	σ_1	σ_2
ρ^2	ρ^2	e	ρ	σ_2	σ_3	σ_1
σ_1	σ_1	σ_2	σ_3	e	ρ	ρ^2
σ_2	σ_2	σ_3	σ_1	ρ^2	e	ρ
σ_3	σ_3	σ_1	σ_2	ρ	ρ^2	e

Заметим, что ни в строках, ни в столбцах выписанной таблицы нет повторяющихся элементов. Это общий факт, имеющий место для любой группы. Действительно, предположим, что у нас есть группа $\mathbb{G}$ с операцией $\star$. Значения клеток строки таблицы умножения, отвечающей элементу $a \in \mathbb{G}$, имеют вид: $a \star x$; для некоторых $x \in \mathbb{G}$. Допустим теперь, что существуют такие $x \in \mathbb{G}$, для которых $a \star x = a \star y$. Тогда

$$x = a' \star (a \star x) = a' \star (a \star y) = y,$$

где a' — обратный к a элемент группы. Итак, значения клеток $a \star x$ и $a \star y$ совпадают тогда и только тогда, когда они лежат в одном столбце. В частности, все элементы одной строки таблицы умножения группы $\mathbb{G}$ должны быть различны. Аналогичные рассуждения доказывают соответствующий результат для столбцов.

Группа симметрии правильного n-угольника, обозначаемая $\mathbb{G}_n$, имеет порядок $2n$ и порождается поворотом ρ на $\frac{360}{n}$ градусов и одной из его осевых симметрий. Если через σ обозначить осевую симметрию, то

$$\sigma \cdot \rho = \rho^{n-1} \cdot \sigma.$$

Такая группа называется группой диэдра порядка $2n$. Для проверки того факта, что группа диэдра действительно описывает симметрии правильного многоугольника, необходимо привлечь аппарат линейной алгебры.

8.3. Арифметические группы

Теория групп обязана своим рождением теории полиномиальных уравнений. Нас интересуют простые числа и разложения целых в произведение простых сомножителей. Эти арифметические свойства связаны с мультипликативной структурой множества $\mathbb{Z}$. В данном разделе мы вводим группы, призванные помочь в изучении этих свойств.

Пусть n — натуральное число. Напомним (см. раздел 4.7), что символом $\mathbb{U}(n)$ мы обозначаем множество обратимых элементов в $\mathbb{Z}_n$, т.е.

$$\mathbb{U}(n) = \{\bar{a} \in \mathbb{Z}_n | \text{НОД}(a, n) = 1\}.$$

Покажем, что это множество образует группу относительно операции умножения классов в $\mathbb{Z}_n$.

Мы знаем, что произведение двух элементов из $\mathbb{Z}_n$ тоже принадлежит $\mathbb{Z}_n$. Так ли это для элементов $\mathbb{U}(n)$? Точнее говоря, верно ли, что произведение любых двух элементов из $\mathbb{U}(n)$ остается в множестве $\mathbb{U}(n)$? Если нет, то умножение классов не является операцией на множестве $\mathbb{U}(n)$ в смысле определения из раздела 8.1.

Другими словами, нам нужно проверить, что произведение двух обратимых классов из $\mathbb{Z}_n$ тоже обратимо. Мы уже сделали это в разделе 4.7, но рассуждение настолько просто, а результат так важен, что мы повторим его здесь. Обозначим, как обычно, обратные

элементы к $\bar{a}, \bar{b} \in \mathbb{U}(n)$ через $\overline{a'}$ и $\overline{b'}$ соответственно. Тогда $\bar{a} \cdot \bar{b}$ обратим, причем обратный к нему элемент равен $\overline{a'} \cdot \overline{b'}$. Для проверки достаточно вычислить произведение:

$$\left(\overline{a \cdot} \, \overline{b}\right) \cdot \left(\overline{a' \cdot} \, \overline{b'}\right) = (\bar{a} \cdot \overline{a'}) \cdot \left(\bar{b} \cdot \overline{b'}\right) = \overline{1}.$$

Теперь мы уверены, что умножение чисел по модулю n задает групповую операцию на множестве $\mathbb{U}(n)$, и нам остается показать, что она удовлетворяет условиям из определения группы. Ассоциативность очевидным образом выполнена, поскольку умножение элементов $\mathbb{Z}_n$ ассоциативно. Единичным элементом служит $\overline{1}$: это обратимый элемент в $\mathbb{Z}_n$, и поэтому принадлежит он $\mathbb{U}(n)$. Тот факт, что любой элемент $\mathbb{U}(n)$ имеет обратный, непосредственно следует из определения $\mathbb{U}(n)$. Итак, множество $\mathbb{U}(n)$ с операцией умножения — на самом деле группа.

Очевидно, что группа $\mathbb{U}(n)$ имеет конечный порядок: она является подмножеством в множестве $\mathbb{Z}_n$, состоящем из n классов. Однако в следующих разделах нам потребуется точная информация о порядке $\mathbb{U}(n)$. На самом деле порядок группы $\mathbb{U}(n)$ настолько часто используется в приложениях, что получил специальное обозначение: $\varphi(n)$. Таким образом, мы имеем функцию, сопоставляющую каждое натуральное число n с количеством элементов множества $\mathbb{U}(n)$. Она называется функцией Эйлера или тотиентом.

Хотелось бы сразу найти общую формулу для $\varphi(n)$, но начнем мы с некоторых частных случаев. Предположим, что p — простое. Тогда каждое натуральное число, меньшее p, взаимно просто с ним. Значит,

$$\mathbb{U}(n) = \mathbb{Z}_n \setminus \{0\}$$

состоит из $p - 1$ элемента, т.е. $\varphi(p) = n - 1$.

Также легко вычисляется $\varphi\left(p^k\right)$ с простым p. Все, что мы должны для этого сделать, — подсчитать количество натуральных чисел, меньших p^k, чей наибольший общий делитель с p^k равен 1. Но ввиду простоты числа p равенство $\text{НОД}\left(a, p^k\right) = 1$ равносильно тому, что p не делит a. Следовательно, достаточно подсчитать количество натуральных чисел, меньших p^k, не делящихся на p. Однако проще найти числа, делящиеся на p. Действительно, если $0 < a < < p^k$ делится на p, то

$$a = p \cdot b, 0 < b < p^{k-1}.$$

Таким образом, среди $p^k - 1$ натуральных чисел, меньших p^k, имеется ровно $p^{k-1} - 1$, делящихся на p. Значит, $p^k - p^{k-1}$ - количество натуральных чисел из того же интервала, не делящихся на p. Итак,

$$\varphi(p^k) = p^k - p^{k-1}.$$

Для вывода общей формулы потребуется доказать следующую лемму.

Лемма 8.1. *Класс $\bar{x}$ принадлежит группе $\mathbb{U}(m \cdot n)$ тогда и только тогда, когда $\bar{a} \in \mathbb{U}(m)$ и $\bar{b} \in \mathbb{U}(n)$.*

Доказательство. Пусть сначала $\bar{x} \in \mathbb{U}(m \cdot n)$. Тогда класс $\bar{x}$ имеет обратный $\overline{x'} \in \mathbb{U}(m \cdot n)$, т.е. $x \cdot x' \equiv 1 (mod\, m \cdot n)$. Но это сравнение справедливо в том и только том случае, когда $x \cdot x' - 1$ делится на $m \cdot n$. В частности, $x \cdot x' - 1$ должно делиться на m, и поэтому $x \cdot x' \equiv 1\ (mod\ m)$. Ввиду выбора $\bar{x}$, справедливо сравнение x $\equiv a\ (mod\ m)$. Значит, $a \cdot x' \equiv 1 (mod\, m)$ и $\bar{a}$ обратим в $\mathbb{Z}_m$. Аналогичные рассуждения показывают, что $\bar{b}$ обратим в $\mathbb{Z}_n$.

Для доказательства обратного утверждения предположим, что $\bar{x}$ — элемент из $\mathbb{Z}_{m \cdot n}$, чьи координаты удовлетворяют условиям: $\bar{a} \in$ $\in \mathbb{U}(m), \bar{b} \in \mathbb{U}(n)$. Нам предстоит показать, что $\bar{x}$ — обратимый элемент $\mathbb{Z}_{m \cdot n}$. По предположению $\bar{a}$ имеет обратный элемент $\overline{a'}$ в $\mathbb{Z}_m$ и $\bar{b}$ имеет обратный $\overline{b'}$ в $\mathbb{Z}_n$. Если $\bar{x} \in \mathbb{Z}_{m \cdot n}$ обратим, то где-то в таблице должен находиться обратный к нему. Чему равны его координаты? Естественно ожидать, что ими будут классы $\overline{a'}$ и $\overline{b'}$. Итак, пусть $0 \le y \le m \cdot n - 1$ — такое целое число, что

$$\begin{cases} \text{y} \equiv \text{a}'\ (\text{mod m}), \\ \text{y} \equiv \text{b}'\ (\text{mod n}). \end{cases}$$

Докажем, что $\bar{y} \in \mathbb{Z}_{m \cdot n}$ — обратный класс для $\bar{x}$. Поскольку $x \equiv$ $\equiv a\ (mod\ m)$ и $\text{y} \equiv a' (mod\ m)$, имеем:

$$x \cdot y \equiv a \cdot a' \equiv 1\ (mod\ m).$$

Значит, $x \cdot y - 1$ делится на m. Дословно те же рассуждения показывают, что $x \cdot y - 1$ делится на n. Но НОД$(m, n) = 1$, что в силу леммы раздела 2.6 обеспечивает нам делимость $x \cdot y - 1$ на произведение $m \cdot n$. Другими словами, $\bar{x} \cdot \bar{y} = \bar{1}$ в $\mathbb{Z}_{m \cdot n}$. Лемма доказана полностью.

Теорема 8.1 (правило вычисление тотиена). *Если m и n — взаимно простые натуральные числа, то*

$$\varphi(m \cdot n) = \varphi(m) \cdot \varphi(n).$$

Доказательство. Будем использовать геометрическую интерпретацию китайской теоремы об остатках, т.е. таблицу раздела 7.3. Напомним, как она строится. Мы начинаем с двух натуральных чисел, удовлетворяющих условию НОД$(m, n) = 1$. Далее чертим таблицу с m столбцами и n строками. Каждый столбец помечаем неотрицательным целым числом (слева направо, в порядке возрастания чисел) и интерпретируем эти числа как классы из $\mathbb{Z}_m$. Аналогично «нумеруем» строки таблицы элементами $\mathbb{Z}_n$. Итак, получилась таблица с $m \times n$ клетками. В каждую из них, стоящую на пересечении столбца a и строки b, мы ставим натуральное число x, удовлетворяющее системе сравнений:

$$\begin{cases} x \equiv \text{a (mod m)}, \\ x \equiv \text{b (mod n)}, \end{cases}$$

и неравенству: $0 \le x \le m \cdot n - 1$. Целые числа a и b мы называем координатами вычета x. Поскольку m и n взаимно просты, китайская теорема об остатках гарантирует нам, что значение каждой клетки таблицы однозначно определяется сформулированными условиями. Мы будем интерпретировать x как класс в $\mathbb{Z}_{m \cdot n}$.

По определению значение функции $\varphi(mn)$ равно числу элементов в $\mathbb{U}(m \cdot n)$. Следовательно, необходимо подсчитать число клеток таблицы, координаты которых лежат соответственно в $\mathbb{U}(m)$ и $\mathbb{U}(n)$. Но число элементов $\mathbb{U}(m)$ равно $\varphi(m)$, а $\mathbb{U}(n)$ имеет $\varphi(n)$ элементов. Значит, искомое число клеток равно $\varphi(m) \cdot \varphi(n)$. Итак, $\varphi(m \cdot n) = \varphi(m) \cdot \varphi(n)$, и теорема доказана.

Замечание. Предположение теоремы о взаимной простоте m и n необходимо. Без него теорема просто не верна. Если, например, $m = m = p$, то $\varphi(m \cdot n) = \varphi(p^2) = p \cdot (p - 1)$, но $\varphi(m \cdot n) = \varphi(m) \cdot \varphi(n) = \varphi(p)^2 = (p - 1)^2$.

Следствие. *Если натуральное число* n *представимо в виде разложения в произведение простых множителей:*

$$n = p_1^{\varepsilon_1} \cdots p_k^{\varepsilon_k},$$

где $0 < p_1 < \cdots < p_k$*, то выполняется*

$$\varphi(n) = p_1^{\varepsilon_1 - 1} \cdots p_k^{\varepsilon_k - 1}(p_1 - 1) \cdots (p_k - 1).$$

Доказательство. В силу взаимной простоты имеем:

$$\varphi(n) = \varphi\left(p_1^{\varepsilon_1} \cdots p_k^{\varepsilon_k}\right) = \varphi\left(p_1^{\varepsilon_1}\right) \cdots \varphi\left(p_k^{\varepsilon_k}\right).$$

Учитывая формулу для φ в случае степени простого числа, выведенную выше, мы получаем искомое выражение. Следствие доказано.

Пример 8.1. *Вычислить* $\varphi(120)$.

Решение. Запишем разложение числа $n = 120$ в произведение простых сомножителей $n = 120 = 2^3 \cdot 3 \cdot 5$. Тогда выполняется

$$\varphi(120) = 2^2(2-1)(3-1)(5-1) = 32.$$

Заметим, что для применения формулы нам необходимо знать разложение числа n на простые множители. Трудность вычисления $\varphi(n)$ для большого целого n обеспечивает безопасность системы шифрования RSA.

8.4. Подгруппы

Если подмножество $\mathbb{H}$ группы $\mathbb{G}$ является группой относительно операции в $\mathbb{G}$, то его называют подгруппой группы $\mathbb{G}$. Ввиду важности понятия подгруппы для дальнейшего выпишем его более формальное определение.

Определение 8.3. *Пусть* $\mathbb{G}$ *— группа с операцией* $\star$. *Непустое подмножество* $\mathbb{H} \subset \mathbb{G}$ *называется подгруппой группы* $\mathbb{G}$, *если:*

1) $a \star b \in \mathbb{H}$ *для любой пары* $a, b \in \mathbb{H}$;

2) единичный элемент группы $\mathbb{G}$ *содержится в* $\mathbb{H}$;

3) $a \in \mathbb{H}$, *то обратный к нему элемент* a' *также входит в* $\mathbb{H}$.

По терминологии, введенной в разделе 8.1, условие (1) означает, что «$\star$» (операция на группе $\mathbb{G}$) является также операцией на множестве $\mathbb{H}$.

Знакомство с примерами начнем с подгрупп группы $\mathbb{Z}$ по сложению. Для натурального n обозначим через $n\mathbb{Z}$ множество всех чисел, кратных n, как положительных, так и отрицательных. Является ли $n\mathbb{Z}$ подгруппой в $\mathbb{Z}$? Для начала отметим, что сумма чисел, кратных n, тоже кратна n, так что условие (1) выполнено. Единичным элементом группы $\mathbb{Z}$ служит 0, кратный любому числу, поэтому $0 \in n\mathbb{Z}$. Наконец, $-a \cdot n = (-a) \cdot n \in n\mathbb{Z}$. Поэтому обратный к каждому элементу из $n\mathbb{Z}$ тоже принадлежит $n\mathbb{Z}$. Таким образом, $n\mathbb{Z}$ действительно подгруппа в $\mathbb{Z}$.

Отметим, что любая группа имеет, по крайней мере, две подгруппы. Одна из них совпадает со всей группой, а другая состоит только из единичного элемента. Среди примеров раздела 8.1

встречается несколько подгрупп. Так, относительно сложения, $\mathbb{Z}$ является подгруппой группы $\mathbb{Q}$, которая, в свою очередь, подгруппа в $\mathbb{R}$. Более того, множество вещественных чисел $\mathbb{R}$ с операцией сложения — подгруппа в группе комплексных чисел $\mathbb{C}$. Если же рассматривать умножение, то множество ненулевых рациональных чисел — подгруппа в $\mathbb{R} \setminus \{0\}$, которая является подгруппой в $\mathbb{C} \setminus \{0\}$.

С другой стороны, $\mathbb{Q} \setminus \{0\}$ (группа по умножению) лежит внутри $\mathbb{Q}$ (группы по сложению). Несмотря на это, мы не можем называть $\mathbb{Q} \setminus \{0\}$ подгруппой в $\mathbb{Q}$, поскольку на них определены разные операции.

Большой интерес, особенно с точки зрения криптографии, представляют конечные группы. Тот факт, что данное конечное подмножество группы является подгруппой, влечет неожиданные ограничения на количество его элементов, т.е. его порядок, которые облегчают процесс поиска подгрупп в конечной группе. Мы изучим простейшее из таких соотношений, которое называется теоремой Лагранжа: порядок любой подгруппы конечной группы делит порядок всей группы.

Проясним, в чем суть теоремы Лагранжа. Пусть $\mathbb{H}$ - подмножество конечной группы $\mathbb{G}$. Тогда, очевидно, количество элементов в $\mathbb{H}$ не больше, чем в $\mathbb{G}$ (равно, если $\mathbb{H} = \mathbb{G}$). Теорема Лагранжа утверждает, что, если к тому же $\mathbb{H}$ — подгруппа в $\mathbb{G}$, то порядок $\mathbb{H}$ обязан делить порядок $\mathbb{G}$. Обратное утверждение неверно.

Рассмотрим $\mathbb{D}_3$ группу симметрии правильного треугольника, порядок которой равен 6. По теореме Лагранжа список порядков ее подгрупп ограничивается числами: 1, 2, 3 и 6. Любая подгруппа должна содержать единичный элемент e, но это не исключает возможности, что какая-то подгруппа состоит только из него. Таким образом, в $\mathbb{D}_3$ есть подгруппа $\{e\}$ порядка 1. Далее, подгруппа может совпадать со всей группой $\mathbb{G}$, при этом, естественно, ее порядок будет равен 6. Нам осталось решить задачу о выделении подгрупп в $\mathbb{D}_3$ порядков 2 и 3.

8.5. Циклические подгруппы

Как обычно, обозначим символом «$\star$» операцию на конечной группе $\mathbb{G}$. Определим k-ю степень элемента a этой группы стандартным образом:

$$a^k = \underbrace{a \star a \star \cdots \star a}_{k \text{ раз}}$$

и рассмотрим множество его степеней:

$$\mathbb{H} = \{e, a, a^2, a^3, \ldots\}.$$

Очевидно, $\mathbb{H}$ — конечное множество в силу конечности группы $\mathbb{G}$. Но такое может быть только в том случае, если существуют равные степени a с разными показателями. Другими словами, найдутся такие натуральные $n > m$, что $a^m = a^n$.

Пусть a' — обратный элемент к a. Умножая обе части равенства $a^m = a^n$ на $(a')^m$, получаем: $a^{n-m} = e$ — единичный элемент. Следовательно, для данного элемента $a \in \mathbb{G}$ найдется натуральное число k, для которого $a^k = e$. Таким образом,

$$a \star a^{k-1} = a^k = e,$$

т.е. обратный элемент к a совпадает с a^{k-1}, степенью самого a. В частности, обратный к a элемент принадлежит множеству $\mathbb{H}$. Поскольку при умножении двух степеней элемента a снова получается его степень, то это означает, что множество $\mathbb{H}$ — подгруппа $\mathbb{G}$.

Чему равен порядок группы $\mathbb{H}$? Предположим, что k — наименьшее натуральное число, обладающее свойством $a^k = e$. Если $n > k$, то мы можем разделить n на k с остатком: $n = kq + r$, где $0 \le r \le k - 1$. Значит,

$$a^n = a^{kq+r} = \left(a^k\right)^q \star a^r.$$

Но $a^k = e$, так что $a^n = a^r$. Иначе говоря, любая степень a, показатель которой больше k, равна степени с некоторым меньшим показателем. Следовательно,

$$\mathbb{H} = \{e, a, a^2, \ldots, a^{k-1}\}.$$

Более того, все эти элементы различны. Действительно, если $r \le \le s < k$ и $a^r = a^s$, то $a^{s-r} = e$. С другой стороны, по предположению, $r - s$ меньше k — наименьшего натурального числа, обладающего таким свойством. Значит, $r - s = 0$, т.е. $r = s$. Таким образом, порядок группы $\mathbb{H}$ равен k.

Итак, у нас появился простой метод построения подгрупп в данной группе $\mathbb{G}$. Выберем произвольный элемент $a \in \mathbb{G}$, тогда:

1) множество $\mathbb{H}$ степеней a является подгруппой в $\mathbb{G}$;

2) порядок группы $\mathbb{H}$ равен наименьшему натуральному k, для которого $a^k = e$.

Введем понятие циклической подгруппы.

Определение 8.4. *Подгруппа $\mathbb{H}$, состоящая из всех степеней одного элемента a, называется циклической подгруппой группы $\mathbb{G}$, а элемент a — ее образующей. Наименьшее натуральное k, удовлетворяющее соотношению $a^k = e$, называется порядком элемента a.*

И, как нетрудно убедиться, порядок циклической подгруппы совпадает с порядком ее образующей.

В качестве простой иллюстрации применения сформулированного метода определим структуру группы $\mathbb{G}$, порядок которой — простое число p. Примером такой группы служит $\mathbb{Z}_p$ с операцией сложения. Предположим, что $\mathbb{H}$ — какая-то подгруппа в $\mathbb{G}$. По теореме Лагранжа порядок $\mathbb{H}$ должен делить порядок $\mathbb{G}$, равный p. Так как p простое, то порядком $\mathbb{H}$ может быть только либо 1, либо p. В первом случае $\mathbb{H} = \{e\}$, а во втором $\mathbb{H} = \mathbb{G}$. Таким образом, в группе $\mathbb{G}$ есть только две подгруппы. Теперь выберем произвольный элемент $a \neq e$ в $\mathbb{G}$ и обозначим через $\mathbb{H}$ циклическую подгруппу, им порожденную. Так как $e \neq a \in \mathbb{H}$, то $\mathbb{H}$ не может состоять только из единичного элемента и, значит, $\mathbb{H} = \mathbb{G}$. В частности, $\mathbb{G}$ — циклическая группа, причем любой ее элемент, отличный от e, является образующей. Суммируем эти результаты в следующей теореме.

Теорема 8.2 (группа простого порядка). *Если порядок группы $\mathbb{G}$ прост, то:*

1) $\mathbb{G}$ циклическая;

2) $\mathbb{G}$ имеет только две подгруппы: саму $\mathbb{G}$ и $\{e\}$;

3) любой элемент группы $\mathbb{G}$, за исключением e, является образующей всей группы.

Итак, любая группа простого порядка — циклическая. Обратное же, вообще говоря, неверно. Например, порядок группы $\mathbb{U}(5)$ равен $\varphi(5) = 4$, но она циклическая и $\overline{2}$ — ее образующая.

8.6. Примеры построения подгрупп

Применим результаты предыдущего параграфа к определению всех подгрупп группы $\mathbb{D}_3$. Мы уже отмечали в 8.5, что для этого достаточно найти в ней подгруппы порядков 2 и 3. Но 2 и 3 — простые числа, так что, согласно теореме 8.2, соответствующие подгруппы должны быть циклическими. Более того, они полностью

определяются своими образующими. Поэтому нам достаточно понять, какие из элементов $\mathbb{D}_3$ имеют порядок 2, а какие — 3.

Так как $\rho^2 \neq e$ и $\rho^3 = e$, то порядок ρ равен 3. Кроме того,

$$(\rho^2)^2 = \rho^3 \cdot \rho = \rho, (\rho^2)^3 = (\rho^3)^2 = e.$$

Значит, ρ^2 также имеет порядок 3. Поскольку каждая из осевых симметрий сама себе обратна, ее порядок равен 2. Циклическая подгруппа, порожденная ρ, имеет вид:

$$\mathbb{R} = \{e, \rho, \rho^2\}.$$

Откуда следует, что она совпадает с подгруппой, порожденной ρ^2. Следовательно, $\mathbb{R}$ — единственная подгруппа в $\mathbb{D}_3$, имеющая порядок 3. Каждая из осевых симметрий задает свою подгруппу в $\mathbb{D}_3$ порядка 2, а именно:

$$\{e, \sigma_1\}, \{e, \sigma_2\}, \{e, \sigma_3\}.$$

Итак, мы доказали, что кроме $\{e\}$ и $\mathbb{D}_3$, в группе $\mathbb{D}_3$ существуют только циклические подгруппы. Следует ли отсюда, что $\mathbb{D}_3$ имеет только циклические подгруппы? На самом деле нет, поскольку $\mathbb{D}_3$, будучи подгруппой в себе самой, не является циклической. Действительно, если бы она была циклической, то ее образующая имела бы порядок 6. Но, как мы видели, любой из ее элементов имеет порядок 1, 2 или 3. Однако все собственные подгруппы в $\mathbb{D}_3$ — циклические.

Определение 8.5. *Подгруппа $\mathbb{H}$ в $\mathbb{G}$ называется собственной, если она содержит хотя бы один элемент, отличный от единичного, и не совпадает при этом со всей группой $\mathbb{G}$: $\mathbb{H} \neq \{e\}$ и $\mathbb{H} \neq \mathbb{G}$.*

Далее мы хотим привести пример группы, которая не только сама не является циклической, но и содержит собственную нециклическую подгруппу. Рассмотрим

$$\mathbb{U}(16) = \{\overline{1}, \overline{3}, \overline{5}, \overline{7}, \overline{9}, \overline{11}, \overline{13}, \overline{15}\}$$

с операцией умножения по модулю 16. Эта группа имеет порядок $\varphi(16) = 8$. По теореме Лагранжа у нее могут быть подгруппы только порядков 1, 2, 4 или 8. Подгруппы порядков 1 и 8 определяются непосредственно, это $\{\overline{1}\}$ и $\mathbb{U}(16)$ соответственно.

Для выделения остальных подгрупп мы должны вычислить порядки каждого элемента из $\mathbb{U}(16)$. Это можно сделать довольно быстро. Элементы $\overline{7}, \overline{9}$ и $\overline{15}$ имеют порядок 2, а элементы $\overline{3}, \overline{5}, \overline{11}$ и $\overline{13}$ — порядок 4. Значит, в группе $\mathbb{U}(16)$ нет элемента порядка 8. В частности, она не является циклической.

Можно ли утверждать, что любая собственная подгруппа в $\mathbb{U}(16)$ циклическая? Напомним, что подгруппа простого порядка должна быть циклической. Поэтому порядок нециклической подгруппы обязан быть составным числом, меньшим восьми, и одновременно делить 8. Значит, если такая подгруппа существует, то ее порядок равен 4. Кроме того, порядок любого ее элемента должен быть меньше 4, поскольку в противном случае она будет циклической. В силу теоремы Лагранжа отсюда вытекает, что все элементы искомой подгруппы, кроме $\overline{1}$, должны иметь порядок 2. Но $\mathbb{U}(16)$ имеет в точности 3 таких элемента, которые вместе с единичным образуют множество из четырех элементов, а именно:

$$\{\overline{1}, \overline{7}, \overline{9}, \overline{15}\}.$$

Легко проверить, что оно является искомой подгруппой. Итак, $\mathbb{U}(16)$ имеет собственную нециклическую подгруппу порядка 4.

Мы воспользуемся результатами предыдущего параграфа для обобщения теоремы Ферма на случай составных модулей.

Теорема 8.3 (Эйлер). *Пусть a и n — натуральные числа. Если $n > 0$ и $\text{НОД}(a, n) = 1$, то*

$$a^{\varphi(n)} \equiv 1 (mod\ n).$$

Доказательство. Воспользуемся теоремой Лагранжа. Из взаимной простоты a и n следует, что $\bar{a} \in \mathbb{U}(n)$. По теореме Лагранжа порядок элемента a делит порядок группы $\mathbb{U}(n)$, который равен $\varphi(n)$. Обозначив через k порядок элемента $\bar{a}$, получим: $\varphi(n) = kr$ для некоторого натурального r. Значит,

$$(\bar{a})^{\varphi(n)} = \left(\bar{a}^k\right)^r = \overline{1},$$

что равносильно сравнению из теоремы Эйлера.

8.7. Теорема Лагранжа

Введем определение отношения эквивалентности, которое будем применять при доказательстве теоремы Лагранжа.

Определение 8.6. *Пусть $\mathbb{G}$ - группа с операцией $\star$ и $\mathbb{H}$ — ее подгруппа. Говорят, что элементы $x, y \in \mathbb{G}$ сравнимы по модулю $\mathbb{H}$, если*

$$x \star y' \in \mathbb{H},$$

где y' — обратный к y элемент группы. В этом случае записывают $x \equiv y\ (mod\ \mathbb{H})$.

Сравнение по модулю n, определенное в разделе 4, является частным случаем этого отношения. Действительно, пусть $\mathbb{G}$ - группа $\mathbb{Z}$ по сложению и $\mathbb{H}$ — ее подмножество, состоящее из всех чисел (как положительных, так и отрицательных), кратных n. Поскольку операция на $\mathbb{Z}$ — сложение, обратным элементом к y будет $y' = -y$. Таким образом, согласно определению, x = y $(mod\ \mathbb{H})$ тогда и только тогда, когда x— $y \in \mathbb{H}$, т.е. $x - y$ кратно n. Следовательно, в этом примере сравнение по модулю $\mathbb{H}$ эквивалентно сравнению по модулю n.

Вернемся теперь к общему случаю сравнения по модулю $\mathbb{H}$ в группе $\mathbb{G}$. Проверим выполнение свойств введенным отношением эквивалентности. Пусть $x, y, z \in \mathbb{G}$.

Рефлексивность. Нам нужно показать, что $x \equiv x\ (mod\ \mathbb{H})$. По определению сравнения это равносильно условию: $x \star x' \in \mathbb{H}$, которое, в свою очередь, следует из того, что $x \star x' = e$ и $\mathbb{H}$ — подгруппа.

Симметричность. Сравнение $x \equiv y\ (mod\ \mathbb{H})$ означает, что $x \star \star y' \in \mathbb{H}$. Но вместе с любым элементом подгруппы ей принадлежит и обратный. Поэтому элемент $y \star x'$, обратный к $x \star y'$, тоже содержится в $\mathbb{H}$ и, следовательно, $y \equiv x\ (mod\ \mathbb{H})$, что доказывает симметричность сравнения.

Транзитивность. Допустим, $x \equiv y\ (mod\ \mathbb{H})$ и $y \equiv z\ (mod\ \mathbb{H})$. Эти сравнения переписываются в виде: $x \star y' \in \mathbb{H}$ и $y \star z' \in \mathbb{H}$ соответственно. Так как $\mathbb{H}$ — подгруппа, то

$$x \star z' = (x \star y') \star (y \star z') \in \mathbb{H}.$$

Отсюда $x \equiv z\ (mod\ \mathbb{H})$, и последнее свойство доказано.

Итак, сравнение по модулю $\mathbb{H}$ рефлексивно, симметрично и транзитивно, поэтому оно является отношением эквивалентности. Заметим, что существует точное соответствие между условием «$\mathbb{H}$ — подгруппа в $\mathbb{H}$» и свойствами, которые интерпретируют сравнение по модулю $\mathbb{H}$ в отношение эквивалентности.

Теперь, зная, что сравнение по модулю $\mathbb{H}$ задает отношение эквивалентности, найдем класс эквивалентности элемента $x \in \mathbb{G}$ относительно этого отношения. По определению искомый класс равен

$$\{\, y \in \mathbb{G} \mid y \equiv x\ (mod\ \mathbb{H}) \,\}.$$

Но условие $y \equiv x\ (mod\ \mathbb{H})$ — это то же самое, что $y \star x' \in \mathbb{H}$. Поэтому $y = h \star x$ для некоторого элемента $h \in \mathbb{H}$. Следовательно, класс эквивалентности элемента x может быть переписан в виде:

$$\{ h \star x \mid h \in \mathbb{H} \}.$$

Общепринятое название $\mathbb{H} \star x$ для такого класса - правый смежный класс группы $\mathbb{G}$ по подгруппе $\mathbb{H}$. Заметим, что класс эквивалентности единичного элемента e совпадает с самой подгруппой $\mathbb{H}$.

Напомним формулировку теоремы.

Теорема 8.4 (Лагранж). *Порядок любой подгруппы конечной группы делит порядок всей группы.*

Доказательство. Пусть $\mathbb{G}$ — конечная группа с операцией $\star$ и $\mathbb{H}$ — ее подгруппа. Подсчитаем сначала количество элементов в классе эквивалентности по модулю $\mathbb{H}$. Если $x \in \mathbb{G}$, то его класс эквивалентности совпадает с

$$\mathbb{H} \star x = \{ h \star x \mid h \in \mathbb{H} \}.$$

Покажем, что множества $\mathbb{H} \star x$ и $\mathbb{H}$ имеют одинаковое количество элементов. Так как элементы множества $\mathbb{H} \star x$ получаются умножением элементов $\mathbb{H}$ на фиксированный элемент $x \in \mathbb{G}$, их не может быть больше, чем в $\mathbb{H}$. Предположим, теперь, что $h_1 \star x = h_2 \star x$ для каких-то двух $h_1, h_2 \in \mathbb{H}$. Тогда

$$h_1 = (h_1 \star x) \star x' = (h_2 \star x) \star x' = h_2,$$

где x' - обратный элемент к x. Это означает, что различные элементы подгруппы $\mathbb{H}$ при умножении их справа на x дают различные элементы в $\mathbb{H} \star x$. Итак, доказано, что число элементов любого класса $\mathbb{H} \star x$ совпадает с порядком подгруппы $\mathbb{H}$.

В итоге имеем:

1) сравнение по модулю $\mathbb{H}$ является отношением эквивалентности, оно разбивает всю группу $\mathbb{G}$ на различные непересекающиеся классы эквивалентности. Таким образом, порядок группы $\mathbb{G}$ равен сумме количеств элементов в каждом из этих классов;

2) установлено, что все эти классы содержат одно и то же число элементов, равное порядку подгруппы $\mathbb{H}$.

Значит, порядок группы $\mathbb{G}$ равен порядку $\mathbb{H}$, умноженному на количество различных классов эквивалентности. В частности, порядок подгруппы $\mathbb{H}$ делит порядок группы $\mathbb{G}$. Теорема доказана.

Замечание. Утверждение, обратное теореме Лагранжа, ложно. Если $\mathbb{G}$ - группа порядка n и k - один из делителей числа n, то группа $\mathbb{G}$ вовсе не обязана иметь подгруппу порядка k. Например, порядок группы симметрии правильного тетраэдра равен 12, но в ней нет подгруппы порядка 6. Тем не менее, если p - простой делитель порядка группы, то в $\mathbb{G}$ найдется подгруппа порядка k (теорема Коши).

Библиографический список

1. Теория чисел: учеб. пособие / А.А. Бухштаб. – Изд. 3-е стер.- СПб.: Лань, 2008.-383 с.

2. Нестеренко Ю.В. Теория чисел: учебник для студ. высш. учеб. заведений / Ю.В. Нестеренко. – М.: Издательский центр «Академия», 2008. – 272 с.

3. Ильин М.Е. Криптографическая защита информации в объектах информационной инфраструктуры: учебник для студ. учреждений сред. проф. образования / М.Е. Ильин, Т.И. Калинкина, В.Н. Пржегорлинский; под ред. В.Н. Пржегорлинского. – М.: Издательский центр «Академия», 2020.- 288 с.

4. Оре Ойстин. Приглашение в теорию чисел: пер. с англ. Изд. 2-е, стер.-М.: Едиториал УРСС, 2003.-128 с.

5. Пискунов Н.С. Дифференциальное и интегральное исчисление для втузов, т. 2: учеб. пособие для втузов. – 13-е изд.- М.: Наука, Главная редакция физико-математической литературы, 1985.- 560 с.

6. Carmichael R. D. On composite numbers P which satisfy the Fermat congruence $a^{p-1} \equiv 1\ (mod\ p)$. The Amer. Math. Monthly 19 (1912), 22-27.

И л ь и н Михаил Евгеньевич
Ц и п о р к о в а Ксения Андреевна

Теоретико-числовые методы в криптографии. Часть 1

Редактор М.Е. Цветкова
Корректор С.В. Макушина
Подписано в печать 10.07.20. Формат бумаги 60×84 1/16.
Бумага писчая. Печать трафаретная. Усл. печ. л. 7,0.
Тираж 30 экз. Заказ
Рязанский государственный радиотехнический университет.
390005, Рязань, ул. Гагарина, 59/1.
Редакционно-издательский центр РГРТУ.