О вычислении классических сумм Якобсталя

Н. Н. Осипов

В статье рассказывается о быстрых алгоритмах вычисления классических сумм Якобсталя. Эти алгоритмы основаны на нетривиальной связи сумм Якобсталя с представлением простых чисел $p \equiv 1 \pmod 4$ и $p \equiv 1 \pmod 3$ бинарными квадратичными формами $A^2 + B^2$ и $A^2 + 3B^2$ соответственно.

Введение

Пусть $p > 2$ — нечётное простое число. Сумма вида

$$\phi(n) = \sum_{x=0}^{p-1} \left(\frac{x^3 + nx}{p} \right) \tag{0.1}$$

называется *суммой Якобсталя* (она впервые возникла и изучалась в работе Э. Якобсталя 1907 года [21]). Здесь n — произвольное целое число, а $\left(\frac{a}{p}\right)$ обозначает *символ Лежандра* a по p: если $a \equiv 0 \pmod p$, то $\left(\frac{a}{p}\right) = 0$, иначе

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{если сравнение } x^2 \equiv a \pmod p \text{ разрешимо,} \\ -1 & \text{в противном случае.} \end{cases}$$

Основные свойства символа Лежандра излагаются практически в любом учебнике по элементарной теории чисел (см., например, [10]). В частности, имеет место сравнение

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

(так называемый *критерий Эйлера*), с помощью которого можно быстро вычислить символ Лежандра $\left(\frac{a}{p}\right)$ даже при больших p[1)].

[1)] Речь идёт о применении хорошо известного *бинарного алгоритма* возведения в степень [34]. Более эффективный алгоритм основан на замене символа Лежандра его обобщением — *символом Якоби*, который быстро вычисляется с помощью *квадратичного закона взаимности* (см. [10, гл. 6]).

Многие авторы изучали и более общие суммы вида

$$\phi_l(n)=\sum_{x=0}^{p-1}\Big(\frac{x^{l+1}+nx}{p}\Big),\quad l\in\mathbb{N},\tag{0.2}$$

которые также называются суммами Якобсталя (подробности см. в статье [30] или, начиная с п. 5.49, в хорошо известной монографии [9], где основательно освещена и история вопроса).

В данной статье будет рассказано, как вычисляются суммы Якобсталя в случаях $l=2$ и $l=3$[2)]. При $n\not\equiv 0 \pmod p$ имеем

$$\phi_3(n)=\sum_{x=1}^{p-1}\Big(\frac{x^4+nx}{p}\Big)=\sum_{x=1}^{p-1}\Big(\frac{x^{-4}+nx^{-1}}{p}\Big)=\sum_{x=1}^{p-1}\Big(\frac{1+nx^3}{p}\Big)=\Big(\frac{n}{p}\Big)\sum_{x=0}^{p-1}\Big(\frac{x^3+n^{-1}}{p}\Big),$$

где $^{-1}$ означает взятие обратного по модулю p. Вместо суммы $\phi_3(n)$ (которая впервые, по-видимому, изучалась в работе [27]) нам будет удобнее рассматривать сумму

$$\psi(n)=\sum_{x=0}^{p-1}\Big(\frac{x^3+n}{p}\Big).\tag{0.3}$$

Суммы Якобсталя (0.1) и (0.3) имеют непосредственное отношение к следующей важной задаче (в том числе для приложений в криптографии [7, гл. VI]). Рассмотрим *эллиптическую кривую* E, заданную уравнением

$$y^2=x^3+ax+b,\tag{0.4}$$

над полем $\mathbb{Z}_p$ классов вычетов по модулю p (предполагается, что p не является делителем дискриминанта $\Delta=-4a^3-27b^2$). Как известно, на множестве $\mathbb{Z}_p$-точек (x,y) этой кривой вместе с формальной бесконечно удалённой точкой ∞ можно ввести операцию сложения, относительно которой это множество превращается в *абелеву группу* (при этом ∞ играет роль нулевого элемента). Порядок $N_p(E)$ этой группы выражается формулой

$$N_p(E)=\sum_{x=0}^{p-1}\Big(1+\Big(\frac{x^3+ax+b}{p}\Big)\Big)+1=p+1+\sum_{x=0}^{p-1}\Big(\frac{x^3+ax+b}{p}\Big).$$

Согласно *теореме Хассе* о числе точек эллиптической кривой над конечным полем, имеет место неравенство $|N_p(E)-p-1|<2p^{1/2}$ (теорема 10.5 в книге [6]). Нахождение точного значения $N_p(E)$ при больших p является содержательной задачей, для решения которой существует довольно

[2)] Случай $l=1$ совсем прост (см. далее лемму 1.5).

нетривиальный *алгоритм Шуфа* (см. оригинальную работу [26]; описание алгоритма можно быстро найти по ссылкам [33, 35]). В частном случае, когда $b = 0$ или $a = 0$, имеется более простой способ найти значение $N_p(E)$, поскольку для сумм $\phi(n)$ и $\psi(n)$, как выясняется, есть быстрый практический алгоритм вычисления[3)].

Цель настоящей статьи — рассказать о практически пригодных (быстро работающих даже для больших p) алгоритмах вычисления сумм Якобсталя $\phi(n)$ и $\psi(n)$ при любом n. С теоретической основой этих алгоритмов можно познакомиться по книгам [1] (см. §§ 18.3, 18.4) и [13] (см. теоремы 6.2.9 и 6.2.10). Описание самих алгоритмов с иллюстрирующими примерами можно найти, например, в статье [22]. Для доказательства утверждений, на которых основаны алгоритмы, обычно привлекают мощный аппарат *сумм Гаусса* и *сумм Якоби*, но мы воспользуемся более элементарными средствами в духе идей пионерской работы [21].

§ 1. Предварительные сведения

В этом разделе мы сообщим предварительные сведения, необходимые для дальнейшего. Те утверждения, которые можно быстро (и не сильно отвлекаясь) доказать, мы приведём с доказательством, а в остальном ограничимся комментариями и ссылками на соответствующую литературу.

1.1. Алгоритм Корнаккиа

Ключевым моментом в задаче быстрого вычисления классических сумм Якобсталя $\phi(n)$ и $\psi(n)$ является неожиданная и нетривиальная связь с задачей представления простого числа p в виде $p = A^2 + B^2$ (для суммы $\phi(n)$) или $p = A^2 + 3B^2$ (для суммы $\psi(n)$). Последняя задача интересна и сама по себе, но для нас принципиально то, что она имеет быстрый алгоритм решения, причём даже в следующей более общей постановке.

Пусть d — фиксированное натуральное число. Требуется представить простое число $p > d$ в виде

$$p = A^2 + dB^2, \tag{1.1}$$

где A и B — некоторые натуральные числа. Конечно, для данного d представление (1.1) возможно не для любого простого числа p, так как есть необходимое условие — разрешимость сравнения

$$x^2 + d \equiv 0 \pmod{p}. \tag{1.2}$$

[3)] В случае суммы (0.1) см. иллюстрирующий пример 1.4 в книге [23, гл. 6].

Последнее легко выяснить с помощью квадратичного закона взаимности. Например, при $d = 1$ сравнение (1.2) разрешимо тогда и только тогда, когда $p \equiv 1 \pmod 4$. Вместе с тем, дать несложное описание тех простых p, которые можно представить в виде (1.1), удаётся далеко не для всех d (см. по этому поводу книгу [18]).

В интересующих нас случаях $d = 1$ и $d = 3$ такое описание есть. Для $d = 1$ его доставляет знаменитая *теорема Ферма — Эйлера*, которая гласит: всякое простое число $p \equiv 1 \pmod 4$ допускает представление в виде суммы двух квадратов. Для $d = 3$ имеет место аналогичное утверждение: каждое простое число $p \equiv 1 \pmod 3$ допускает представление в виде суммы квадрата и утроенного квадрата. Оба утверждения могут быть легко получены на основе *факториальности* колец целых чисел соответствующих мнимых квадратичных полей (см., например, доказательство теоремы Ферма — Эйлера в учебнике [8, гл. 4, § 2, с. 153]; вообще, эту классическую теорему можно доказать многими способами, среди которых есть и весьма экзотические [31]).

Для сравнения рассмотрим $d = 11$. Легко проверить, что для простого $p > 11$ сравнение (1.2) разрешимо тогда и только тогда, когда $p \equiv 1, 3, 4, 5, 9 \pmod{11}$. Но последнее условие ещё не гарантирует возможность представления (1.1): например, для $p = 23$ оно выполнено, однако равенство $23 = A^2 + 11B^2$ невозможно. Какие же простые числа $p > 11$ можно представить в виде $p = A^2 + 11B^2$? Известен только такой ответ: те и только те, для которых сравнение

$$x^3 - x^2 - x - 1 \equiv 0 \pmod p$$

имеет три решения. Этот критерий весьма нетривиален, но он не упрощает проверку[4)].

Прежде чем переходить к изложению *алгоритма Корнаккиа* (см. п. 1.5.2 в книге [17]), решающего для данного простого p вопрос о представлении в виде (1.1), докажем следующий важный факт.

Лемма 1.1. *Если представление* (1.1) *возможно, то оно однозначно*[5)].

Доказательство. Будем рассуждать от противного. Пусть

$$p = A^2 + dB^2 = X^2 + dY^2,$$

где A, B, X, Y — натуральные числа, причём $A < X$, $B > Y$. Поскольку

$$B^2X^2 - A^2Y^2 \equiv B^2X^2 + dB^2Y^2 = B^2(X^2 + dY^2) = B^2p \equiv 0 \pmod p,$$

имеем одно из двух сравнений

$$BX \pm AY \equiv 0 \pmod p.$$

[4)] Подробности и переформулировку критерия в терминах *чисел Трибоначчи* см. в [19].

[5)] В случае $d = 1$ мы не обращаем внимание на порядок слагаемых.

Так как $BX \pm AY > 0$, получаем $BX \pm AY \geqslant p$. Следовательно,

$$p^2 \leqslant (BX \pm AY)^2 \leqslant (A^2 + B^2)(X^2 + Y^2) \leqslant p^2,$$

причём равенство возможно только при $d=1$ и в этом случае вектор (X, Y) пропорционален одному из векторов $(B, \pm A)$, откуда $X = B$ и $Y = A$. □

Для успешной работы алгоритма Корнаккиа требуется предварительно найти некоторое решение $x = m_0$ сравнения (1.2). В случае больших простых p это тоже содержательная задача, для решения которой можно применить общий *вероятностный алгоритм* решения уравнений над полем $\mathbb{Z}_p$ или специфические методы извлечения квадратных корней по модулю p (см., например, §§ 6.1, 6.2 в книге [2]). Вероятностные алгоритмы такого рода на практике работают довольно быстро.

Пусть $\left(-\frac{d}{p}\right) = 1$, т. е. сравнение (1.2) разрешимо. Тогда для $p = 4k + 3$ его решения можно найти по явной формуле $x \equiv \pm(-d)^{k+1} \pmod p$, поскольку

$$x^2 \equiv (-d)^{2k+2} = (-d)^{(p+1)/2} \equiv -d \pmod p.$$

В случае $p \equiv 1 \pmod 4$ можно применить *алгоритм Тонелли — Шенкса* [36], для чего предварительно потребуется найти какой-нибудь *квадратичный невычет* b по модулю p. На практике можно просто «подбросить монетку»: выбрать случайный вычет b и вычислить символ Лежандра $\left(\frac{b}{p}\right)$: с вероятностью $1/2$ он будет равен -1. В случаях $p \equiv 5 \pmod 8$ и $p \equiv 5 \pmod{12}$ можно взять $b = 2$ и $b = 3$ соответственно.

Сравнение (1.2) можно также решить с помощью *алгоритма Чиполлы* [32]. В этом алгоритме вычисления производятся в некотором *квадратичном расширении* поля $\mathbb{Z}_p$. Предварительно необходимо найти вычет b, для которого $b^2 + 4d$ является квадратичным невычетом по модулю p (понятно, что и здесь применимы вероятностные соображения).

Будем считать, что $0 < m_0 < p/2$. Алгоритм Корнаккиа состоит в следующем.

(а) Положим $r_0 = p$, $r_1 = m_0$.

(б) Применяя *алгоритм Евклида* к числам r_0 и r_1, будем вычислять остатки $r_2, \ldots, r_s$ (r_{i+1} — остаток от деления r_{i-1} на r_i) до тех пор, пока не получим неравенство $r_s^2 < p$.

(в) Если $(p - r_s^2)/d = t^2$ для некоторого натурального t, то (1.1) имеет место для $(A, B) = (r_s, t)$. Иначе представление (1.1) невозможно.

В случае $d = 1$ шаг (в) можно упростить: вычислим ещё один остаток r_{s+1}, и тогда $(A, B) = (r_s, r_{s+1})$. На практике алгоритм Корнаккиа довольно быстро решает вопрос о представлении (1.1), поскольку последовательность остатков $\{r_i\}$ экспоненциально убывает.

Пример 1.1. Пусть $d = 3$, $p = 2017$. Тогда $m_0 = 589$. Имеем

$$r_1 = p \bmod m_0 = 250, \quad r_2 = m_0 \bmod r_1 = 89, \quad r_3 = r_1 \bmod r_2 = 72,$$
$$r_4 = r_2 \bmod r_3 = 17, \quad 17^2 < 2017, \quad \frac{2017 - 17^2}{3} = 24^2.$$

Таким образом, $2017 = 17^2 + 3 \cdot 24^2$.

Обоснование корректности алгоритма Корнаккиа представляет интересную тему для отдельной статьи, и мы не будем здесь этим заниматься (читателя отсылаем к статье [12] как содержащей наиболее компактное изложение). Отметим только, что данный алгоритм работает[6)] и для составных p (при условии, что удалось найти все решения сравнения (1.2), а это сложная проблема в случае составного модуля).

Случай $d = 1$ издавна привлекал внимание. Известно несколько алгоритмов для представления простых чисел суммой двух квадратов, появившихся до алгоритма Корнаккиа (1908 год). В первую очередь следует упомянуть *алгоритм Эрмита — Серре* 1848 года (см. [20, 28]), использующий разложение рационального числа m_0/p в цепную дробь. По существу, алгоритм Корнаккиа в случае $d = 1$ является укороченным вариантом алгоритма Эрмита — Серре. На языке цепных дробей формулируется и схожий *алгоритм Смита* (1855 год), детальное изложение которого можно найти в статье [16]. Исторически первым был *алгоритм Лежандра* (1808 год), который использует разложение в (периодическую) цепную дробь квадратичной иррациональности $p^{1/2}$. Для больших p этот алгоритм, вообще говоря, неэффективен, так как период цепной дроби может оказаться настолько длинным, что его невозможно будет выписать.

Краткое изложение упомянутых алгоритмов (с поясняющими примерами) есть в главе V книги [5]. В статье [14] приводится обоснование усовершенствованной версии алгоритма Эрмита — Серре. Другое обоснование можно найти в статье [4].

1.2. Евклидовы кольца и алгоритм Евклида

Для некоторых d вопрос о получении представления (1.1) можно решать другим (тоже вполне эффективным на практике) способом. Нам понадобится понятие *евклидова кольца* (см., например, учебник [3]). Предполагая d свободным от квадратов, рассмотрим *кольцо целых чисел*

$$\mathbb{Z}[\omega] = \{x + y\omega : (x, y) \in \mathbb{Z}^2\}$$

[6)] В случае составного p алгоритм Корнаккиа находит только *примитивные представления* (1.1), т. е. с дополнительным условием НОД $(A, B) = 1$ (оно автоматически выполнено для простых p).

мнимого квадратичного поля $\mathbb{Q}(\sqrt{-d})$. Здесь

$$\omega = \begin{cases} \dfrac{1+\sqrt{-d}}{2} & \text{при } -d \equiv 1 \pmod 4, \\ \sqrt{-d} & \text{при } -d \not\equiv 1 \pmod 4. \end{cases}$$

Пусть $N(\gamma)$ — *норма* числа $\gamma = x + y\omega \in \mathbb{Q}(\sqrt{-d})$, т. е.

$$N(\gamma) = |\gamma|^2 = \gamma\overline{\gamma} = \begin{cases} \dfrac{x^2+xy+(d+1)y^2}{4} & \text{при } -d \equiv 1 \pmod 4, \\ x^2+dy^2 & \text{при } -d \not\equiv 1 \pmod 4 \end{cases}$$

(здесь и далее черта сверху означает сопряжение в поле комплексных чисел $\mathbb{C}$).

Следующий критерий очевидным образом вытекает из определений и свойства мультипликативности нормы: $N(\gamma_1\gamma_2) = N(\gamma_1)N(\gamma_2)$.

Лемма 1.2. *В кольце* $\mathbb{Z}[\omega]$ *можно делить с остатком относительно нормы* $N(\cdot)$ *тогда и только тогда, когда для любого* $\gamma \in \mathbb{Q}(\sqrt{-d})$ *найдётся такое* $\gamma^* \in \mathbb{Z}[\omega]$, *что*

$$N(\gamma - \gamma^*) < 1. \tag{1.3}$$

Опираясь на лемму 1.2, нетрудно установить евклидовость некоторых колец $\mathbb{Z}[\omega]$:

Лемма 1.3. *Кольцо* $\mathbb{Z}[\omega]$ *евклидово при* $d \in \{1, 2, 3, 7, 11\}$.

Доказательство. При естественном геометрическом изображении чисел кольца $\mathbb{Z}[\omega]$ на комплексной плоскости $\mathbb{C}$ получится *решётка* с базисом 1 и ω. Нужно убедиться, что базисный параллелограмм можно покрыть открытыми кругами единичного радиуса с центрами в вершинах 0, 1, ω и $1+\omega$ этого параллелограмма. Для указанных значений d это более или менее очевидно из элементарно-геометрических соображений. □

На самом деле в лемме 1.3 перечислены все евклидовы кольца: как бы мы ни вводили *евклидову норму* в кольце $\mathbb{Z}[\omega]$, других случаев евклидовости не появится (это несложное упражнение мы оставляем заинтересованному читателю). Случаи $d = 1$ (*целые гауссовы числа*) и $d = 3$ (*целые числа Эйзенштейна*) хорошо известны. Оценку (1.3) в этих случаях можно уточнить:

$$N(\gamma - \gamma^*) \leqslant \begin{cases} 1/2 & \text{при } d = 1, \\ 1/3 & \text{при } d = 3. \end{cases}$$

Для $d \in \{1, 2, 3, 7, 11\}$ задачу о представлении (1.1) можно решать следующим образом. Пусть простое число p таково, что сравнение (1.2)

разрешимо и $x = m_0$ — одно из решений. Идея состоит в том, чтобы рассмотреть $\delta = \text{НОД}\,(p, m_0 + \sqrt{-d})$ в кольце $\mathbb{Z}[\omega]$.

Лемма 1.4. $N(\delta) = p$.

Доказательство. Имеем

$$(m_0 + \sqrt{-d})(m_0 - \sqrt{-d}) = pl$$

для некоторого целого l. Отсюда видно, что $p > 2$ не является простым элементом евклидова (и потому факториального) кольца $\mathbb{Z}[\omega]$, ибо ни одно из чисел $m_0 \pm \sqrt{-d}$ не делится на p. Пусть $p = \pi\xi$, где $\pi,\ \xi \in \mathbb{Z}[\omega]$ и π — простой элемент. Тогда $p^2 = N(p) = N(\pi)N(\xi)$, откуда $N(\pi) = N(\xi) = p$ и $\xi = \overline{\pi}$. Таким образом, $p = \pi\overline{\pi}$ есть произведение двух простых (возможно, ассоциированных) элементов. Теперь уже легко установить, что δ ассоциировано либо с π, либо с $\overline{\pi}$. В обоих случаях $N(\delta) = p$. □

Теперь мы можем вычислить $\delta = x_0 + y_0\omega$ с помощью алгоритма Евклида для кольца $\mathbb{Z}[\omega]$ (это делается быстро даже для больших p). Тогда при $d \in \{1, 2\}$ имеем $p = x_0^2 + dy_0^2$, и задача о представлении (1.1) решена. Если же $d \in \{3, 7, 11\}$, то

$$p = x_0^2 + x_0y_0 + \frac{d+1}{4}y_0^2 = \left(x_0 + \frac{y_0}{2}\right)^2 + \frac{dy_0^2}{4},$$

и требуемое представление возможно только при чётном y_0.

При $d = 3$ заменой δ на $\omega\delta$ или на $\omega^2\delta$ можно сделать y_0 чётным. При $d = 7$ число y_0 обязано быть чётным. А вот при $d = 11$ число y_0 уже может оказаться нечётным, и тогда искомое представление (1.1) будет невозможным. Например, для $p = 23$ имеем

$$23 = 4^2 + 4 \cdot 1 + 3 \cdot 1^2 = 5^2 + 5 \cdot (-1) + 3 \cdot (-1)^2,$$

так что $y_0 = \pm 1$ — нечётное число.

Пример 1.2. Пусть $d = 1$, $p = 2017$. Тогда можно взять $m_0 = 229$. Положим $\rho_0 = 2017$, $\rho_1 = 229 + \sqrt{-1}$ и вычислим $\delta = \text{НОД}\,(\rho_0, \rho_1)$. Имеем

$$\frac{\rho_0}{\rho_1} = \frac{229 - \sqrt{-1}}{26} \approx 8{,}80 - 0{,}03\sqrt{-1}, \quad \gamma^* = 9,$$

так что $\rho_2 = \rho_0 - \rho_1\gamma^* = -44 - 9\sqrt{-1}$. Далее находим

$$\frac{\rho_1}{\rho_2} = -5 + \sqrt{-1} = \gamma^*$$

и $\rho_3 = \rho_1 - \rho_2\gamma^* = 0$. Значит, $\delta = \rho_2 = -44 - 9\sqrt{-1}$ и $2017 = 9^2 + 44^2$.

1.3. Сумма символов Лежандра

Пусть b и c — целые числа. Следующее утверждение о значении суммы

$$S(b,c)=\sum_{x=0}^{p-1}\Big(\frac{x^2+bx+c}{p}\Big)$$

составляет основу всех дальнейших вычислений.

Лемма 1.5. *Справедливо равенство*

$$S(b,c)=\begin{cases} p-1, & \text{если } D\equiv 0 \pmod p,\\ -1, & \text{если } D\not\equiv 0 \pmod p,\end{cases} \tag{1.4}$$

где $D=b^2-4c$.

Доказательство. Выделяя в выражении x^2+bx+c полный квадрат, нетрудно обнаружить, что

$$S(b,c)=S(0,-D)=\sum_{x=0}^{p-1}\Big(\frac{x^2-D}{p}\Big).$$

Сначала вычислим $S(0,-D)$ по модулю p, воспользовавшись критерием Эйлера. Имеем

$$S(0,-D)\equiv\sum_{x=0}^{p-1}(x^2-D)^{(p-1)/2}=\sum_{s=0}^{p^*}C_{p^*}^s(-D)^{p^*-s}\sum_{x=0}^{p-1}x^{2s}\pmod p,$$

где $p^*=(p-1)/2$. Далее нам понадобится следующая формула для степенной суммы по модулю p (здесь t — целое число):

$$\sum_{x=1}^{p-1}x^t\equiv_p\begin{cases}-1, & \text{если } t \text{ делится на } p-1,\\ 0 & \text{иначе}\end{cases}$$

(несложное доказательство, основанное на цикличности *мультипликативной группы* поля $\mathbb{Z}_p$, предоставляется читателю). Тогда

$$\sum_{x=0}^{p-1}x^{2s}\equiv_p\begin{cases}0 & \text{при } 0\leqslant s<p^*,\\ -1 & \text{при } s=p^*.\end{cases}$$

Следовательно, $S(0,-D)\equiv -1 \pmod p$. Теперь, поскольку $|S(0,-D)|\leqslant p$, получим $S(0,-D)\in\{-1,p-1\}$ для любого D. Ясно, что $S(0,0)=p-1$. Но

$$\sum_{D=0}^{p-1}S(0,-D)=\sum_{x=0}^{p-1}\sum_{D=0}^{p-1}\Big(\frac{x^2-D}{p}\Big)=0,$$

поэтому $S(0,-1)=S(0,-2)=\ldots=S(0,-p+1)=-1$. □

Существуют и другие способы доказательства формулы (1.4) (в нашем доказательстве мы следовали Якобсталю [21]). Вычисление суммы $S(0,-D)$ эквивалентно подсчёту числа точек гиперболы $y^2 = x^2 - D$ над полем $\mathbb{Z}_p$. С точки зрения элементарной алгебраической геометрии, это кривая второго порядка, которая допускает *рациональную параметризацию*. В данном случае удобно перейти к новым переменным $u = x - y$, $v = x + y$ и записать уравнение в виде

$$uv = D.$$

Линейная замена $(x, y) \to (u, v)$ биективна, а число решений (u, v) последнего уравнения над полем $\mathbb{Z}_p$ легко находится в зависимости от D. Таким образом, имеем

$$\sum_{x=0}^{p-1}\Big(1 + \Big(\frac{x^2-D}{p}\Big)\Big) = \begin{cases} 2p-1, & \text{если } D \equiv 0 \pmod p, \\ p-1, & \text{если } D \not\equiv 0 \pmod p. \end{cases}$$

Отсюда и следует равенство (1.4).

§ 2. Вычисление суммы $\phi(n)$

При замене x на $-x$ выражение $x^3 + nx$ меняет знак, что позволяет записать сумму $\phi(n)$ в виде

$$\phi(n) = \Big(1 + \Big(\frac{-1}{p}\Big)\Big) \sum_{x=1}^{(p-1)/2} \Big(\frac{x^3+nx}{p}\Big).$$

В случае $p \equiv 3 \pmod 4$ имеем $\Big(-\frac{1}{p}\Big) = -1$, поэтому $\phi(n) = 0$ для любого n. В частности, для эллиптической кривой E, заданной уравнением

$$y^2 = x^3 + nx, \tag{2.1}$$

имеем $N_p(E) = p + 1$.

Далее в этом разделе будем рассматривать только случай $p \equiv 1 \pmod 4$. В этом случае $\Big(-\frac{1}{p}\Big) = 1$ и $\phi(n)$ чётно при любом n. Ясно также, что $\phi(n) = 0$ при $n \equiv 0 \pmod p$.

Введём обозначение: $k = (p-1)/4$, так что $p = 4k + 1$ и $(p-1)/2 = 2k$.

2.1. Абсолютное значение $\phi(n)$

Пусть $m \not\equiv 0 \pmod p$. Имеем

$$\phi(nm^2) = \sum_{x=0}^{p-1}\Big(\frac{x^3+nm^2x}{p}\Big) = \Big(\frac{m}{p}\Big)\sum_{x=0}^{p-1}\Big(\frac{x^3+nx}{p}\Big) = \Big(\frac{m}{p}\Big)\phi(n).$$

Как следствие, получим

$$|\phi(nm^2)| = |\phi(n)|.$$

Отсюда видно, что $|\phi(n)|$ при $n \not\equiv 0 \pmod p$ может принимать только два значения: одно для квадратичных вычетов n по модулю p, другое — для квадратичных невычетов. Более конкретно можно выразиться так: либо $|\phi(n)| = |\phi(1)|$, либо $|\phi(n)| = |\phi(g)|$, где g — какой-нибудь *первообразный корень* по модулю p.

Рассмотрим частный случай $m = m_0$, где $m_0^2 \equiv -1 \pmod p$. Имеем

$$\phi(-n) = \phi(nm_0^2) = \left(\frac{m_0}{p}\right)\phi(n) = (-1)^k\phi(n),$$

поскольку по критерию Эйлера

$$\left(\frac{m_0}{p}\right) \equiv m_0^{2k} \equiv (-1)^k \pmod p.$$

В частности, $\phi(1) = (-1)^k\phi(-1)$ (это равенство нам понадобится в подразделе 2.3).

2.2. Значение $\phi(n)$ по модулю p

Важным шагом на пути вычисления $\phi(n)$ является вычисление $\phi(n) \bmod p$. Применим ту же технику, что и при доказательстве леммы (1.5):

$$\phi(n) \equiv \sum_{x=0}^{p-1}(x^3+nx)^{(p-1)/2} = \sum_{x=0}^{p-1}\sum_{s=0}^{2k} C_{2k}^s x^{2k+2s} n^{2k-s} = \\ = \sum_{s=0}^{2k} C_{2k}^s n^{2k-s} \sum_{x=0}^{p-1} x^{2k+2s} \equiv -n^k C_{2k}^k \pmod p.$$

Как следствие, получим

$$\frac{\phi(n)}{2} \equiv n^k a \pmod p, \tag{2.2}$$

где введено обозначение $a = \phi(1)/2$. Формула (2.2) в виде формулы (10) есть в статье [24] (см. также статью [30], где по модулю p найдены суммы Якобсталя (0.2) для любого l).

Главное наблюдение: если удастся вычислить a, то с помощью сравнения (2.2) можно однозначно и быстро найти $\phi(n)/2$ для любого n, поскольку a priori имеет место оценка

$$\left|\frac{\phi(n)}{2}\right| \leqslant \frac{p-1}{2}$$

и $n^k \bmod p$ быстро вычисляется с помощью бинарного алгоритма.

Таким образом, всё сводится к нахождению числа a. Как это можно сделать, довольно понятно написано ещё самим Якобсталем в статье [21]. Следующие подразделы 2.3 и 2.4 представляют собой пересказ соответствующей части этой статьи.

2.3. Значение $\phi(1)/2$ по модулю 4

Якобсталь фактически вычислил значение $a = \phi(1)/2$ по модулю 8 (см. ниже (2.5)). Удобно сначала рассмотреть $a' = \phi(-1)/2$. Имеем

$$2a' = \sum_{x=0}^{p-1}\Big(\frac{x^3-x}{p}\Big) = \sum_{x=1}^{p-3}\Big(\frac{x}{p}\Big)\Big(\frac{x+1}{p}\Big)\Big(\frac{x+2}{p}\Big). \tag{2.3}$$

Пусть $N_p^{(3)}$ — число тех $x \in \{1, 2, \ldots, p-3\}$, для которых

$$\Big(\frac{x}{p}\Big) = \Big(\frac{x+1}{p}\Big) = \Big(\frac{x+2}{p}\Big) = -1.$$

Тогда

$$8N_p^{(3)} = \sum_{x=1}^{p-3}\Big(1-\Big(\frac{x}{p}\Big)\Big)\Big(1-\Big(\frac{x+1}{p}\Big)\Big)\Big(1-\Big(\frac{x+2}{p}\Big)\Big). \tag{2.4}$$

Кроме того, число $N_p^{(3)}$ чётно (поскольку из $x \in N_p^{(3)}$ следует $p-x-2 \in N_p^{(3)}$, так что все элементы $N_p^{(3)}$ разбиваются на пары). Теперь из равенств (2.3) и (2.4) с помощью леммы 1.5 можно вывести равенство $8N_p^{(3)} = p-3-2a'$ (это тривиальное, но несколько громоздкое и скучное упражнение решается исключительно усилием воли; вслед за Якобсталем мы предоставим это читателю).

Как следствие, приходим к сравнению

$$a' \equiv \frac{p-3}{2} = 2k-1 \pmod 8.$$

Далее имеем

$$a = (-1)^k a' \equiv (-1)^k(2k-1) \pmod 8, \tag{2.5}$$

откуда $a \equiv -1 \pmod 4$, поскольку $(-1)^k(2k-1) \equiv -1 \pmod 4$ при любом k.

Таким образом, если нам удастся вычислить $|a|$, то с помощью полученного сравнения можно однозначно определить и само число a.

2.4. Связь с представлением $p = A^2 + B^2$

Последнее вычисление, которые мы предпримем, является, пожалуй, самым важным: оно приводит к формуле (2.6), лежащей в основе любого эффективного алгоритма вычисления $\phi(n)$.

Читателю предлагается убедиться в справедливости следующей цепочки равенств:

$$\begin{aligned}\sum_{n=0}^{p-1}\phi^2(n) &= \sum_{n=0}^{p-1}\sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\Big(\frac{x^3+nx}{p}\Big)\Big(\frac{y^3+ny}{p}\Big) = \\ &= \sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\Big(\frac{xy}{p}\Big)\sum_{n=0}^{p-1}\Big(\frac{n^2+(x^2+y^2)n+x^2y^2}{p}\Big) = \\ &= \sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\Big(\frac{xy}{p}\Big)\sum_{n=0}^{p-1}\Big(\frac{n^2-(x^2-y^2)^2}{p}\Big) = \\ &= p\sum_{x=1}^{p-1}\Big(\frac{x^2}{p}\Big)+p\sum_{x=1}^{p-1}\Big(\frac{-x^2}{p}\Big)-\sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\Big(\frac{xy}{p}\Big) = 2p(p-1).\end{aligned}$$

Теперь, учитывая результат подраздела 2.1, получим

$$\sum_{n=0}^{p-1}\phi^2(n) = \frac{p-1}{2}\big(\phi^2(1)+\phi^2(g)\big) = 2p(p-1),$$

где g — первообразный корень по модулю p. Отсюда

$$p = \Big(\frac{\phi(1)}{2}\Big)^2+\Big(\frac{\phi(g)}{2}\Big)^2. \tag{2.6}$$

В частности, $|\phi(1)| < 2p^{1/2}$ и $|\phi(g)| < 2p^{1/2}$, так что $|\phi(n)| < 2p^{1/2}$ вообще для любого $n \not\equiv 0 \pmod p$. Тем самым полностью доказано утверждение теоремы Хассе для эллиптической кривой E, заданной уравнением (2.1).

Равенство (2.6) есть не что иное, как представление числа p в виде суммы двух квадратов. Таким образом, нам осталось найти альтернативный (и эффективный на практике) способ получения такого представления. Как мы уже видели (см. подразделы 1.1 и 1.2), такой способ есть, и даже не один.

2.5. Быстрый алгоритм вычисления $\phi(n)$

Представленные выше факты приводят к следующему алгоритму вычисления $\phi(n)$ при $n \not\equiv 0 \pmod p$, работающему в случае больших p.

(а) С помощью какого-нибудь быстрого алгоритма (см. подразделы 1.1 и 1.2) находим представление простого $p = 4k+1$ в виде суммы двух квадратов:

$$p = A^2+B^2, \tag{2.7}$$

где A, B — натуральные числа, при этом A нечётно. Как видно из леммы 1.1, такие числа A и B определяются числом p однозначно.

(б) Затем находим $a = \phi(1)/2$, исходя из условий

$$|a| = A, \quad a \equiv -1 \pmod 4.$$

(в) Наконец, для заданного n находим $\phi(n)$, исходя из условий

$$\frac{\phi(n)}{2} \equiv n^k a \pmod p, \quad \left|\frac{\phi(n)}{2}\right| \leqslant \frac{p-1}{2}.$$

Проиллюстрируем данный алгоритм одним примером.

Пример 2.1. Пусть $p{=}2017{=}4{\cdot}504{+}1$ и $n{=}-37$. Имеем $2017{=}9^2{+}44^2$ (см. пример 1.2), так что $A = 9$ и, таким образом, $a = -9$. Тогда

$$\frac{\phi(-37)}{2} \equiv (-37)^{504} \cdot (-9) \equiv 1973 \pmod{2017}.$$

Поскольку $1973 > 1008 = (2017-1)/2$, получим

$$\frac{\phi(-37)}{2} = 1973 - 2017 = -44,$$

откуда $\phi(-37) = -88$.

Упражнение 2.1. Для простого числа $p = 4k+1$ опишите алгоритм быстрого вычисления биномиального коэффициента C_{2k}^k по модулю p.

Указание. Воспользуйтесь сравнением $C_{2k}^k \equiv -\phi(1) \pmod p$.

Упражнение 2.2. Докажите, что для простого числа $p = 4k+1$ представление (2.7) может быть найдено с помощью следующих *формул Гаусса*:

$$A \equiv \frac{1}{2} C_{2k}^k \pmod p, \quad B \equiv (2k)!A \pmod p,$$

где (не обязательно положительные) числа A, B удовлетворяют дополнительным условиям $|A| \leqslant (p-1)/2$, $|B| \leqslant (p-1)/2$.

Комментарий. Формулы Гаусса непригодны для непосредственного разложения больших чисел p в сумму двух квадратов, поскольку непонятно, как быстро вычислить C_{2k}^k по модулю p, не прибегая к алгоритму из упражнения 2.1.

Упражнение 2.3. Пусть $N_p(G)$ — число точек эллиптической кривой G, заданной уравнением $u^2v^2 + u^2 + v^2 = 1$, над полем $\mathbb{Z}_p$. Докажите, что

$$N_p(G) = N_p(E) - 2\Big(\frac{-1}{p}\Big) = p + 1 + \phi(4) - 2\Big(\frac{-1}{p}\Big), \tag{2.8}$$

где E — кривая (2.1) с $n = 4$.

Указание. Воспользуйтесь *бирациональной заменой*

$$u=\frac{2x}{y},\quad v=\frac{x-2}{x+2}\quad\leftrightarrows\quad x=\frac{2(1+v)}{1-v},\quad y=\frac{4(1+v)}{u(1-v)}.$$

Комментарий. В формуле (2.8) учитываются и две точки кривой G на бесконечности (это точки пересечения G с бесконечно удалённой прямой). Формула для $N_p(G)$ впервые встречается в дневниках Гаусса как предположение, которое позднее было доказано (см. по этому поводу [15], а также комментарий на стр. 97 в книге [11]).

Для сравнения укажем ещё один алгоритм для вычисления $\phi(n)$, который можно предложить на основе теоремы 6.2.1 из книги [13]. Он также опирается на представление (2.7), однако теперь в вычислениях будут участвовать как A, так и B.

(а) Выберем какой-нибудь первообразный корень g по модулю p.

(б) Определим a_4, исходя из условий

$$a_4\equiv-\Big(\frac{2}{p}\Big)\pmod 4,\quad |a_4|=A.$$

(в) Затем вычислим b_4, исходя из условий

$$b_4\equiv g^k a_4\pmod p,\quad |b_4|=B.$$

(г) Далее найдём l, для которого $g^l\equiv n\pmod p$, и вычислим $r=l \bmod 4$.

(д) Тогда

$$\phi(n)=\begin{cases}2(-1)^k a_4, & \text{если } r=0,\\ 2(-1)^k b_4, & \text{если } r=1,\\ 2(-1)^{k+1} a_4, & \text{если } r=2,\\ 2(-1)^{k+1} b_4, & \text{если } r=3.\end{cases}\tag{2.9}$$

Для доказательства корректности этого алгоритма достаточно проверить, что для числа $\phi(n)$, найденного по формуле (2.9), выполняется сравнение (2.2). Проще всего рассмотреть все ситуации в зависимости от значений $(2/p)$ и r. Пусть, например, мы имеем

$$\Big(\frac{2}{p}\Big)=1,\quad r=3.$$

В этом случае $a_4=a$, а число k чётно. Тогда по формуле (2.9) получим

$$\frac{\phi(n)}{2}=-b_4\equiv-g^k a\pmod p.$$

Нам нужно убедиться, что $-g^k a\equiv n^k a\pmod p$. Действительно, имеем

$$n^k\equiv g^{kl}=g^{k(4m+3)}\equiv g^{3k}\equiv-g^k\pmod p,$$

поскольку $g^{2k}=g^{(p-1)/2}\equiv-1\pmod p$.

По очевидным причинам (сначала требуется найти g, а затем решать вычислительно сложную задачу *дискретного логарифмирования*, чтобы найти l и r) данный алгоритм не будет эффективным при больших p.

§ 3. Вычисление суммы $\psi(n)$

В этом разделе мы рассмотрим сумму (0.3), вычисление которой можно организовать по тому же сценарию, что и вычисление суммы (0.1). Поэтому мы подробно опишем только наиболее сложные этапы вычисления, предоставив читателю самому восстановить детали в остальных случаях.

I. Сначала заметим, что при $p \equiv 2 \pmod 3$ отображение

$$f : \mathbb{Z}_p \to \mathbb{Z}_p, \quad f(x) = x^3,$$

биективно (очередное упражнение для читателя), поэтому

$$\psi(n) = \sum_{x=0}^{p-1} \Big(\frac{x^3+n}{p}\Big) = \sum_{y=0}^{p-1} \Big(\frac{y+n}{p}\Big) = 0.$$

Далее пусть $p \equiv 1 \pmod 3$. Положим $k = (p-1)/6$, так что $p = 6k+1$ и $(p-1)/2 = 3k$. Пусть также $n \not\equiv 0 \pmod p$ (иначе, очевидно, $\psi(n) = 0$).

Так же как и в подразделе 2.1 легко доказать, что

$$\psi(nm^3) = \Big(\frac{m}{p}\Big)\psi(n) \tag{3.1}$$

при любом $m \not\equiv 0 \pmod p$. Отсюда следует, что $|\psi(n)|$ принимает одно из трёх значений $|\psi(g^u)|$, $|\psi(g^v)|$ и $|\psi(g^w)|$, где u, v и w — какая-нибудь полная система вычетов по модулю 3 (как и выше, g — фиксированный первообразный корень по модулю p). Далее мы возьмём $u = 0$, $v = 2$, $w = 4$.

II. Теперь заметим, что число $\psi(n)$ чётно тогда и только тогда, когда

$$n^k \equiv \pm 1 \pmod p. \tag{3.2}$$

Действительно, при выполнении условия (3.2) сравнение

$$x^3 + n \equiv 0 \pmod p$$

имеет ровно три решения, а иначе оно неразрешимо. Значит, среди слагаемых $\Big(\frac{x^3+n}{p}\Big)$ суммы $\psi(n)$ число плюс-минус единиц равно либо $p-3$, либо p, что и доказывает утверждение. В частности, число $\psi(1)$ чётно, а числа $\psi(g^2)$ и $\psi(g^4)$ нечётны.

III. Следующий шаг — *вычисление* $\psi(n)$ по модулю p. Имеем

$$\psi(n) \equiv \sum_{x=0}^{p-1}(x^3+n)^{(p-1)/2} = \sum_{x=0}^{p-1}\sum_{s=0}^{3k} C_{3k}^s x^{3s} n^{3k-s} = \\ = \sum_{s=0}^{3k} C_{3k}^s n^{3k-s} \sum_{x=0}^{p-1} x^{3s} \equiv -n^k C_{3k}^{2k} \pmod p.$$

В частности, видно, что всегда $\psi(n) \not\equiv 0 \pmod p$. Кроме того,

$$\psi(n) \equiv 2n^k a \pmod p, \tag{3.3}$$

где используется обозначение $a = \psi(1)/2$. Если удастся вычислить число a, то, учитывая чётность числа $\psi(n)$, с помощью сравнения (3.3) и неравенства $|\psi(n)| \leqslant p$ мы сможем однозначно определить $\psi(n)$.

IV. Как мы уже знаем, число a — целое. Более того, справедливо сравнение

$$a \equiv -1 \pmod 3. \tag{3.4}$$

В самом деле, имеем

$$\psi(n) = \Big(\frac{n}{p}\Big) + \sum_{x=1}^{p-1}\Big(\frac{x^3+n}{p}\Big) = \Big(\frac{n}{p}\Big) + \sum_{y=1}^{p-1} N_p(y)\Big(\frac{y+n}{p}\Big),$$

где $N_p(y)$ — число решений сравнения $x^3 \equiv y \pmod p$. При $y \not\equiv 0 \pmod p$ имеем $N_p(y) \in \{0, 3\}$, поэтому

$$\psi(n) \equiv \Big(\frac{n}{p}\Big) \pmod 3.$$

В частности, при $n = 1$ отсюда следует сравнение (3.4).

V. Осталось самое сложное — найти связь числа a с представлением

$$p = A^2 + 3B^2, \tag{3.5}$$

где натуральные A и B однозначно определены числом p (лемма 1.1). Мы докажем, что $|a| = A$ и, следовательно, $a = \pm A$, где знак выбирается в соответствии с (3.4).

С этой целью вычислим две суммы:

$$S_1 = \sum_{n=0}^{p-1} \psi(n^2), \quad S_2 = \sum_{n=0}^{p-1} \psi^2(n).$$

Основным инструментом при вычислении будет лемма 1.5.

Первая сумма вычисляется так:

$$S_1 = \sum_{n=0}^{p-1}\sum_{x=0}^{p-1}\Big(\frac{x^3+n^2}{p}\Big) = \sum_{x=0}^{p-1}\sum_{n=0}^{p-1}\Big(\frac{n^2+x^3}{p}\Big) = p-1+(-1)(p-1) = 0.$$

С другой стороны, имеем

$$S_1 = 2\sum_{j=0}^{3k-1}\psi(g^{2j}) = 2\sum_{l=0}^{k-1}\big(\psi(g^{6l})+\psi(g^{6l+2})+\psi(g^{6l+4})\big) = \\ = 2k\big(\psi(1)+\psi(g^2)+\psi(g^4)\big),$$

поскольку

$$\psi(g^{6l}) = \psi(1), \quad \psi(g^{6l+2}) = \psi(g^2), \quad \psi(g^{6l+4}) = \psi(g^4)$$

(см. формулу (3.1)). В качестве следствия получим равенство

$$\psi(1)+\psi(g^2)+\psi(g^4) = 0. \tag{3.6}$$

Вычисление второй суммы можно организовать следующим образом:

$$S_2 = \sum_{n=0}^{p-1}\sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\Big(\frac{x^3+n}{p}\Big)\Big(\frac{y^3+n}{p}\Big) = \\ = \sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\sum_{n=0}^{p-1}\Big(\frac{n^2+(x^3+y^3)n+x^3y^3}{p}\Big) = \\ = \sum_{x=0}^{p-1}\sum_{y=0}^{p-1}\sum_{n=0}^{p-1}\Big(\frac{n^2-(x^3-y^3)^2}{p}\Big) = (p-1)N_p+(-1)(p^2-N_p),$$

где N_p — число решений (x, y) сравнения $x^3 - y^3 \equiv 0 \pmod p$. Нетрудно видеть, что

$$N_p = 3(p-1)+1 = 3p-2,$$

поэтому окончательно получим

$$S_2 = 2p(p-1).$$

С другой стороны, имеем

$$S_2 = \sum_{n=0}^{p-1}\psi^2(n) = 2k\big(\psi^2(1)+\psi^2(g^2)+\psi^2(g^4)\big).$$

Следовательно,

$$\psi^2(1)+\psi^2(g^2)+\psi^2(g^4) = 6p. \tag{3.7}$$

Наконец, из полученных равенств (3.6) и (3.7) следует равенство

$$\psi^2(1) + \psi(1)\psi(g^2) + \psi^2(g^2) = 3p,$$

которое можно переписать в виде

$$p = \Big(\frac{\psi(1)}{2}\Big)^2 + 3\Big(\frac{\psi(1) + 2\psi(g^2)}{6}\Big)^2.$$

Таким образом, $|a| = A$ в силу единственности представления (3.5).

VI. На основе сказанного выше можно предложить следующий алгоритм вычисления $\psi(n)$ при любом $n \not\equiv 0 \pmod p$.

(а) С помощью одного из алгоритмов (см. подразделы 1.1 и 1.2) находим представление простого $p = 6k + 1$ в виде (3.5).

(б) Находим $a = \psi(1)/2$, исходя из равенства $|a| = A$ и сравнения (3.4).

(в) С помощью сравнения (3.2) определяем чётность $\psi(n)$, а затем находим $\psi(n)$, опираясь на сравнение (3.3) и неравенство $|\psi(n)| \leqslant p - 1$.

Пример 3.1. Пусть $p = 2017 = 6 \cdot 336 + 1$ и $n = -432$. Имеем

$$2017 = 17^2 + 3 \cdot 24^2$$

(см. пример 1.1), так что $A = 17$ и, следовательно, $a = 17$. Поскольку

$$(-432)^{336} \equiv 1 \pmod{2017},$$

число $\psi(-432)$ чётно, при этом

$$\psi(-432) \equiv 2 \cdot (-432)^{336} \cdot 17 \equiv 34 \pmod{2017}.$$

Значит, $\psi(-432) = 34$.

Упражнение 3.1. Докажите утверждение теоремы Хассе для эллиптической кривой E, заданной уравнением

$$y^2 = x^3 + n. \tag{3.8}$$

Упражнение 3.2. Пусть $p > 3$ и $N_p(F)$ — число точек (эллиптической) *кривой Ферма* F, заданной уравнением $u^3 + v^3 = 1$, над полем $\mathbb{Z}_p$. Докажите формулу

$$N_p(F) = N_p(E) = p + 1 + \psi(-432), \tag{3.9}$$

где E — кривая (3.8) с $n = -432$.

Указание. Воспользуйтесь бирациональной заменой

$$u = \frac{36 + y}{6x}, \quad v = \frac{36 - y}{6x} \quad \leftrightarrows \quad x = \frac{12}{u + v}, \quad y = \frac{36(u - v)}{u + v}.$$

Комментарий. Формула (3.9) учитывает точки (одну или три) кривой F на бесконечности. При $p \equiv 1 \pmod 3$ имеем

$$\psi(-432) = \psi((-3)^3 \cdot 16) = \left(\frac{-3}{p}\right)\psi(16) = \psi(16),$$

а если $p \equiv 2 \pmod 3$, то $\psi(-432) = 0$.

Упражнение 3.3. Докажите *теорему Гаусса* (см. § 8.3 в книге [1], а также § 2 главы IV в книге [29]):
если $p \equiv 1 \pmod 3$ и в представлении

$$4p = C^2 + 27D^2 \tag{3.10}$$

имеем $C \equiv 1 \pmod 3$, то

$$N_p(F) = p + 1 + C.$$

Попутно выясните, когда число 2 будет *кубическим вычетом* по модулю p.

Решение. Прежде всего заметим, что представление (3.10) возможно, при этом числа C и D определены однозначно с точностью до знака (это можно вывести из аналогичного утверждения о представлении (3.5)). Как следствие, сравнение $C \equiv 1 \pmod 3$ определяет число C однозначно.

I. Пусть $N_p = N_p(F) - 3 = p - 2 + \psi(16)$ — число решений сравнения

$$u^3 + v^3 \equiv 1 \pmod p.$$

Тогда, как нетрудно обнаружить, $N_p \equiv 6 \pmod 9$. Следовательно, верно сравнение

$$4p + 4\psi(16) \equiv 5 \pmod 9.$$

Поскольку

$$\psi(16) \equiv \psi(1) \equiv \psi(g^2) \equiv \psi(g^4) \equiv 1 \pmod 3,$$

возможны три случая: либо $\psi(16) = \psi(1)$, либо $\psi(16) = \psi(g^2)$, либо $\psi(16) = \psi(g^4)$.

1) Пусть $\psi(16) = \psi(1)$. Имеем

$$4p + 4\psi(16) = \psi^2(1) + 3\left(\frac{\psi(1) + 2\psi(g^2)}{3}\right)^2 + 4\psi(1) \equiv 5 \pmod 9.$$

Так как $\psi(1) \equiv 1 \pmod 3$, имеем $\psi^2(1) + 4\psi(1) \equiv 5 \pmod 9$. Значит,

$$\frac{\psi(1) + 2\psi(g^2)}{3} \equiv 0 \pmod 3$$

и, таким образом, $C = \psi(1)$.

2) Пусть теперь $\psi(16) = \psi(g^2)$. В этом случае имеем

$$4p + 4\psi(16) = \psi^2(g^2) + 3\left(\frac{2\psi(1) + \psi(g^2)}{3}\right)^2 + 4\psi(g^2) \equiv 5 \pmod 9.$$

Как и выше, отсюда следует, что

$$\frac{2\psi(1) + \psi(g^2)}{3} \equiv 0 \pmod 3$$

и, таким образом, $C = \psi(g^2)$.

3) В случае $\psi(16) = \psi(g^4)$ рассуждения аналогичны и мы получим $C = \psi(g^4)$.

II. Очевидно, число 2 является кубическим вычетом по модулю p тогда и только тогда, когда $\psi(16) = \psi(1)$. Поскольку число $\psi(1)$ чётно, а числа $\psi(g^2)$ и $\psi(g^4)$ нечётны, последнее условие выполнено тогда и только тогда, когда $C \equiv 0 \pmod 2$ в представлении (3.10) (или, в терминах представления (3.5), когда $B \equiv 0 \pmod 3$).

§ 4. Заключительные замечания

Важно подчеркнуть, что рассмотренные нами быстрые алгоритмы вычисления сумм Якобсталя $\phi(n)$ и $\psi(n)$ имеют вероятностную природу, поскольку их ключевой ингредиент — некоторое решение $x = m_0$ сравнения (1.2) — предполагается находить вероятностными методами. В связи с этим возникает вопрос: существуют ли быстрые *детерминированные* алгоритмы извлечения квадратных корней по модулю простого числа p? Ответ, как выясняется, положительный, но неожиданный: для решения этой задачи можно приспособить уже упомянутый нами алгоритм Шуфа, который находит число точек эллиптической кривой над конечным полем и имеет *полиномиальную* сложность (см. оригинальную работу [26]). При этом даже не приходится опираться на правдоподобные, но ещё не доказанные гипотезы в теории чисел, как это иногда бывает.

Так, в предположении верности *обобщённой гипотезы Римана* квадратичный невычет b для алгоритма Тонелли — Шенкса можно найти простым перебором за полиномиальное время. Как следствие, алгоритм[7)] Тонелли — Шенкса становится детерминированным и полиномиальным.

Таким образом, если нас интересует безусловный (не апеллирующий ни к каким гипотезам), детерминированный и полиномиальный алгоритм вычисления классических сумм Якобсталя, то на данный момент

[7)] К слову, этот алгоритм правильнее было бы называть алгоритмом Тонелли, ибо Шенкс лишь переоткрыл его спустя 80 лет, не наведя исторические справки по анекдотичной причине (см. [36]).

можно предложить только алгоритм Шуфа. Который, однако, довольно сложен теоретически и, как уже отмечалось, на практике проигрывает более наивным вероятностным алгоритмам.

Классические суммы Якобсталя отвечают за число точек на очень специальных эллиптических кривых (0.4) над полем $\mathbb{Z}_p$. Про такие кривые говорят, что они допускают *комплексное умножение* (что это такое, можно узнать только основательно погрузившись в теорию эллиптических кривых). Естественно, что в *системах компьютерной алгебры* (например PARI/GP) для подсчёта точек на эллиптических кривых с комплексным умножением над конечными полями используются быстрые практичные алгоритмы типа тех, что были рассмотрены в настоящей статье[8)]. И только в более сложных случаях применяется алгоритм Шуфа и его модификации (подробности можно узнать по ссылке [37]).

Фундаментальное изучение общих сумм Якобсталя $\phi_l(n)$ (включая и рассмотренные нами случаи первых значений l) возможно на основе более сложных конструкций — сумм Гаусса и Якоби. Систематическое изложение результатов в этой области читатель может найти в монографии [13]. Для первоначального знакомства с указанными суммами могут быть полезны соответствующие разделы в книгах [1] и [9].

Автор выражает благодарность А. В. Устинову за содержательные замечания и комментарии по теме статьи.

Список литературы

[1] *Айерлэнд К.*, *Роузен М.* Классическое введение в современную теорию чисел. М.: Мир, 1987.

[2] *Василенко О. Н.* Теоретико-числовые алгоритмы в криптографии. М.: МЦНМО, 2006.

[3] *Винберг Э. Б.* Курс алгебры. М.: МЦНМО, 2019.

[4] *Вялый М. Н.* О представлении чисел в виде суммы двух квадратов // Математическое просвещение. Сер. 3. Вып. 10. М.: МЦНМО. 2006. С. 190–194.

[5] *Дэвенпорт Г.* Высшая арифметика. М.: Наука, 1965.

[6] *Кнэпп Э.* Эллиптические кривые. М.: Факториал Пресс, 2004.

[7] *Коблиц Н.* Курс теории чисел и криптографии. М.: ТВП, 2001.

[8] *Кострикин А. И.* Введение в алгебру. Часть 3. Основные структуры. М.: МЦНМО, 2018.

[9] *Лидл Р.*, *Нидеррайтер Г.* Конечные поля. М.: Мир, 1988.

[8)] На самом деле всё не так просто и совсем не элементарно (читатель может заглянуть в статью [25]).

[10] *Нестеренко Ю. В.* Теория чисел. М.: Академия, 2008.

[11] *Степанов С. А.* Арифметика алгебраических кривых. М.: Наука, 1991.

[12] *Basilla J. M.* On the solution of $x^2 + dy^2 = m$ // Proc. Japan Acad. Ser. A Math. Sci 2004. Vol. 80, № 5. P. 40–41.

[13] *Berndt B. C.*, *Evans R. J.*, *Williams K. S.* Gauss and Jacobi sums. New York: John Wiley & Sons, Inc., 1998.

[14] *Brillhart J.* Note on representing a prime as a sum of two squares // Math. Comp. 1972. Vol. 26. P. 1011–1013.

[15] *Chowla S.* The last entry in Gauss's diary // Proc. Nat. Acad. Sci. U.S.A. 1949. Vol. 35, № 5. P. 244–246.

[16] *Clarke F. W.*, *Everitt W. N.*, *Littlejohn L. L.*, *Vorster S. J. R.* H. J. S. Smith and the Fermat two squares theorem // Amer. Math. Monthly. 1999. Vol. 106, № 7. P. 652–665.

[17] *Cohen H.* A course in computational algebraic number theory. Berlin: Springer-Verlag, 1993. (Grad. Texts Math.; Vol. 138).

[18] *Cox D. A.* Primes of the form $x^2 + ny^2$. New York: John Wiley & Sons, Inc., 1989.

[19] *Evink T.*, *Helminck P. A.* Tribonacci numbers and primes of the form $p = x^2 + 11y^2$ // https://arxiv.org/abs/1801.04605

[20] *Hermite C.* Note au sujet de l'article précédent // J. Math. Pures Appl. 1848. Vol. 13. P. 15.

[21] *Jacobsthal E.* Über die Darstellung der Primzahlen der Form $4n + 1$ als Summe zweier Quadrate // J. Reine Angew. Math. 1907. Bd. 132. S. 238–246.

[22] *Katre S. A.* Jacobsthal sums in terms of quadratic partitions of a prime // Number theory (Ootacamund, 1984). Berlin: Springer-Verlag, 1985. (Lect. Notes Math.; Vol. 1122). P. 153–162.

[23] *Koblitz N.* Algebraic aspects of cryptography. Berlin: Springer-Verlag, 1998. (Alg. Comp. Math.; Vol. 3).

[24] *Lehmer E.* On Euler's criterion // J. Austral. Math. Soc. 1959/1961. Vol. 1, part 1. P. 64–70.

[25] *Rubin K.*, *Silverberg A.* Point counting on reductions of CM elliptic curves // J. Number Theory. 2009. Vol. 129, № 12. P. 2903–2923.

[26] *Schoof R.* Elliptic curves over finite fields and the computation of square roots mod p // Math. Comp. 1985. Vol. 44, №. 170. P. 483–494.

[27] *von Schrutka L.* Ein Beweis für die Zerlegbarkeit der Primzahlen von der Form $6n + 1$ in ein einfaches und ein dreifaches Quadrat // J. Reine Angew. Math. 1911. Bd. 140. S. 252–265.

[28] *Serret J.-A.* Sur un théorème relatif aux nombres entieres // J. Math. Pures Appl. 1848. Vol. 13. P. 12–14.

[29] *Silverman J. H.*, *Tate J.* Rational points on elliptic curves. New York: Springer-Verlag, 1992. (Undergrad. Texts in Math.).

[30] *Whiteman A. L.* Cyclotomy and Jacobsthal sums // Amer. J. Math. 1952. Vol. 74, № 1. P. 89–99.

[31] *Zagier D.* A one-sentence proof that every prime $p \equiv 1 \pmod 4$ is a sum of two squares // Amer. Math. Monthly. 1990. Vol. 97, № 2. P. 144.

[32] https://ru.wikipedia.org/wiki/Алгоритм_Чиполлы

[33] https://en.wikipedia.org/wiki/Counting_points_on_elliptic_curves

[34] https://ru.wikipedia.org/wiki/Алгоритмы_быстрого_возведения_в_степень

[35] https://ru.wikipedia.org/wiki/Алгоритм_Шуфа

[36] https://ru.wikipedia.org/wiki/Алгоритм_Тонелли-Шенкса

[37] https://pari.math.u-bordeaux.fr/dochtml/html/Elliptic_curves.html

Николай Николаевич Осипов, Сибирский федеральный университет, Красноярск

nnosipov@rambler.ru