Бессалов А.В.

ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ В ФОРМЕ ЭДВАРДСА И КРИПТОГРАФИЯ

МОНОГРАФИЯ

КИЕВ

«Політехніка»

2017

УДК 621.391.15 : 519.7

Рецензенты:

Задирака В.К., д.ф.м.н., профессор, академик НАН Украины

Горбенко И.Д., д.т.н., профессор, академик АНПРЭ

Бессалов А.В.

Эллиптические кривые в форме Эдвардса и криптография: монография. – Киев: ІВЦ «Видавництво «Політехника»», 2017. –272с.

ISBN

Исследованы свойства нового класса эллиптических кривых в форме Эдвардса, которые полезны для решения задач асимметричной криптографии. Обоснована новая классификация кривых в обобщенной форме Эдвардса над конечным полем нечетной характеристики с разбиением их на три непересекающихся класса в зависимости от свойств квадратичности их параметров a и d. Дан анализ свойств циклических полных кривых и нециклических скрученных кривых Эдвардса над простым полем, доказаны 12 теорем о новых сойствах этих кривых. На основе свойства делимости точек кривой на 2 предложен оригинальный метод нахождения порядка точек кривой, в сотни раз более производительный в сравнении со стандартным. Проведен сравнительный анализ скорости экспоненцирования точки для кривых в форме Эдвардса и Вейерштрасса с выигрышем первых в 1.5 – 1.6 раза. Рассчитаны и табулированы общесистемные параметры криптостойких полных кривых Эдвардса над простым полем и расширениями малых простых полей, а также скрученных кривых Эдвардса над простым полем. Они предлагаются для внедрения в новый национальный стандарт цифровой подписи. Приведен обзор асимметричных криптопротоколов и известных стандартов цифровой подписи на эллиптических кривых.

Для студентов, аспирантов, программистов и ученых, специализирующихся в области асимметричной криптогафии и безопасности информации.

УДК 621.391.15 : 519.7

ПРЕДИСЛОВИЕ

Название данной книги содержит ключевые слова широко известного сегодня понятия «эллиптическая криптография». После появления идеи этой технологии в знаменитых работах В.Миллера [13] и Н.Коблица [14] минуло 3 десятипетия, в течение которых криптосистемы на эллиптических кривых уверенно вытесняют традиционные и устаревшие асимметричные криптосистемы RSA и Эль-Гамаля с арифметикой в конечном кольце и конечном поле. Конкурентоспособность последних упала по понятным причинам: их субэкспоненциальная сложность требует использования модулей в тысячи бит, что быстро переполняет как временные ресурсы, так и ресурсы памяти. Криптосистемы на эллиптических кривых с их экспоненциальной сложностью выигрывают в сравнении с ними практически на порядок в быстродействии и длине модуля поля. Например, кривая P-384 стандарта США FIPS-186-2-2000 с модулем в 384 бита [78] и RSA-7068 с длиной модуля 7068 бита имеют одинаковый уровень безопасности в 192 бита симметричного шифрования [15]. При заданном уровне безопасности это фактически в десятки раз ускоряет и удешевляет функционирование криптопротоколов. В этой связи после значительного прогресса в исследованиях уязвимых типов кривых и эффективности ее арифметики с началом 21-го столетия начался очень активный процесс стандартизации алгоритмов и протоколов асимметричных криптосистем на эллиптических кривых (ECC – Elliptic Curve Cryptosystems) [73 – 86]. За прошедшие полтора десятилетия не выявлено известных случаев взлома ECC по причине появления нового метода криптоатаки на стандартные кривые и алгоритмы.

В настоящее время перспективы ECC оцениваются во взаимосвязи с прогнозируемым появлением постквантовой криптографии (PQC – Postquantum Cryptography). По мнению ведущих ученых в области эллиптической криптографии Коблица и Менезеса [15] появления доступных квантовых компьютеров по самым оптимистичным прогнозам следует ожидать не ранее чем через 15-20 лет. В то же время последняя версия стандарта США FIPS-186-4 [79] 2013-го года до сих пор рекомендует кривые, рассчитанные математиками АНБ еще в 1997 году и вошедшие в

национальный стандарт FIPS-186-2 [78] 2000-го года. В августе 2015 года АНБ США анонсировало форсирование работ криптографов страны в области стандартизации криптопримитивов PQC. Обсуждая различные аспекты политики АНБ, авторы обзора [15] считают, что давно пора обновить эллиптические кривые 1997 года более эффективными и быстрыми, в частности, кривыми Эдвардса с рекордной производительностью.

В нашей стране с 2002 года действует национальный стандарт цифровой подписи ДСТУ 4145-2002 [82] с арифметикой несуперсингулярных эллиптических кривых над расширениями полей характеристики 2. Кроме того, в 2014 году в качестве национального утвержден международный стандарт ISO/IEC 14888-1,2,3:2008 [83 – 85]. По оценкам большого числа экспертов кривые над полями характеристики 2 имеют немало уязвимых мест и их следует избегать в криптосистемах [15]. Второй стандарт ISO/IEC также не содержит рекомендаций по новым кривым. Можно констатировать, что как американские, так и наши стандарты сильно устарели и требуют абгрейда.

Автор настоящей книги вместе со своими сотрудниками, учениками и аспирантами с 2010 года занимается исследованиями новых свойств эллиптических кривых в форме Эдвардса [30 – 66]. С самого начала мы были заинтригованы уникальностью этих свойств и красотой ее математики с двухсотлетней историей. У нас нет сомнений в перспективности применения этой наиболее быстрой и эффективной технологии в задачах асимметричной эллиптической криптографии. Как и во всем мире, стандарты в нашей стране следует периодически обновлять. Есть надежда, что данная монография будет способствовать этой цели.

Выражаю большую благодарность рецензентам книги академику НАНУ профессору В.К.Задираке и академику АНПРЭ профессору И.Д. Горбенко за труд по ее рецензированию и полезные замечания, А.А.Дихтенко за работу по многочисленным расчетам общесистемных параметров криптостойких кривых Эдвардса, Л.В.Ковальчук за плодотворное научное сотрудничество. Очень признателен аспирантке О.В.Цыганковой, с которой нами опубликовано наибольшее число статей по теме работы, за обсуждение и анимацию некоторых результатов , а также за помощь в подготовке рукописи.

ВВЕДЕНИЕ

Среди изобилия различных форм представления эллиптических кривых [5] в последние годы появилась новая форма (точнее, давно забытая старая), представленная в работе профессора математики Университета Нью-Йорка Гарольда Эдвардса [1]. Изучая труды Эйлера, Гаусса и Абеля двухсотлетней давности, он обнаружил уравнение, которое с помощью рациональных преобразований приводится к уравнению канонической эллиптической кривой в форме Вейерштрасса. Эдвардсу удалось найти закон сложения точек для этой кривой и доказать ее изоморфизм с кривой Вейерштрасса. Эти два результата дали ученым веские основания называть предложенную форму уравнения кривыми в форме Эдвардса. Вскоре оказалось, что новый класс кривых обладает рядом замечательных свойств. Эти свойства сразу были замечены и исследованы криптографами. Первой очень конструктивной работой в развитие этого направления следует отметить статью Даниэля Бернстейна и Тани Ланге [2] (за ней последовала серия работ этих и других авторов). Авторы [2] проанализировали свойства кривых Эдвардса над конечным полем характеристики, не равной 2. Они модифицировали оригинальную кривую Эдвардса, ввели новый параметр кривой d как квадратичный невычет поля и получили закон сложения точек для модифицированной кривой. Эта модификация позволила перейти от нециклической оригинальной кривой Эдвардса с особыми точками к циклической кривой без особых точек. Немаловажное достоинство кривых Эдвардса – наличие одного параметра d вместо двух для кривой Вейерштрасса. Далее авторы [2] доказали, что наряду со свойствами полноты и универсальности закона сложения, заменой точки на бесконечности аффинной точкой (нейтральный элеиент группы точек), кривые в форме Эдвардса среди известных являются рекордно производительными: в проективных координатах групповые операции сложения и удвоения точек выполняются минимальным числом полевых операций. По нашим оценкам, переход от канонических эллиптических кривых в форме Вейерштрасса на новую технологию кривых в форме Эдвардса дает выигрыш в скорости

экспоненцирования точки кривой не менее чем в 1.5 – 1.6 раза. Иными словами, банк, например, за один промежуток времени может заверить своей подписью не 100, а 150 платежных документов. Пропорционально удешевляется услуга. Последнее свойство делает их особенно привлекательными для криптографической защиты информации, где, как и везде, время – деньги.

Уникальность кривых в форме Эдвардса состоит прежде всего в том, что все ее точки, включая нейтральный элемент группы, являются не особыми. Это сразу снимает проблему программирования операций, включающих особую точку с бесконечными координатами, что характерно для всех традиционных кривых в форме Веерштрасса. Безусловно, это еще в большей степени ускоряет программную реализацию и выполнение криптоалгоритмов.

Не будет преувеличением отметить, что наибольший вклад в развитие теории кривых Эдвардса в интересах криптографии внес американский ученый Даниэль Бернстейн со своими соавторами. Наряду с вышесказанным о статье [2] в его последующих работах впервые определены и исследованы свойства скрученных кривых Эдвардса [3], бинарных кривых Эдвардса [7,8], предложена арифметика проективных инвертированных координат для полных кривых Эдвардса [5] с рекордно минимальной сложностью групповой операции сложения точек. Без этих замечательных результатов оригинальные кривые Эдвардса скорее всего остались бы красивым математическим открытием, не интересным для криптоприложений.

Монография содержит 6 глав. В первых 4 главах анализируются известные и новые свойства кривых в форме Эдвардса над простым полем характеристики, не равной 2. Доказывается изоморфизм между четвертью всех кривых в форме Вейерштрасса и кривыми Эдвардса (глава 1), обосновывается и предлагается новая классификация кривых в обобщенной форме Эдвардса (глава 2), доказываются новые свойства полных кривых Эдвардса (глава 3) и скрученных кривых Эдвардса (глава 4). Основным акцентом исследований является поиск кривых порядка $N_E = 4n$ (n – простое чисдо), пригодных для криптографических приложений и стандартов. Предложен метод определения порядка точки кривой Эдвардса на основе свойства делимости точки на 2. Метод обеспечивает выигрыш в экономии

вычислений в сотни раз по сравнению со стандартным. Приводятся также наши оценки производительности экспоненцирования точки кривой Эдвардса из разных классов с выигрышем в 1.5 – 1.6 раза в сравнении с канонической формой кривой. Формулируются и доказываются 12 новых теорем и 9 утверждений о свойствах точек малых и больших порядков в классах полных и скрученных кривых Эдвардса. Рассчитаны и сведены в таблицы общесистемные параметры криптостойких полных и скрученных кривых Эдвардса над простым полем, пригодных для новых стандартов цифровой подписи и распределения ключей. Найдены формулы расчета точного числа несуперсингулярных кривых Эдвардса, в том числе кривых с порядком $N_E = 4n$. В главе 5 приводятся метод и наши результаты поика криптостойких кривых Эдвардса над расширениями малых полей характеристик 5, 7 и 2. Для полей нечетной характеристики такие кривые были найдены с помощью кривых минимального порядка над базовым полем с последующим расширением поля. Над полями характеристики 2 параметры криптостойких кривых Эдвардса были определены на основе их изоморфизма со стандартными кривыми в форме Вейерштрасса. В заключительной 6-й главе рассмотрены основные криптопротоколы и стандарты эллиптической криптографии. Лейтмотивом ее является то, что стандарты асимметричной криптографии практически не обновлялись с начала 21-го века. Если в США в 2013 году появилась новая редакция стандарта FIPS-186-4 с содержанием в основном старого стандарта FIPS-186-2-2000, то в Украине смело пользуются стандартом выпуска 2002 года с самыми небезопасными эллиптическими кривыми над полями характеристики 2.

ГЛАВА 1

ИЗОМОРФИЗМ ЭЛЛИПТИЧЕСКИХ КРИВЫХ В ФОРМЕ ВЕЙЕРШТРАССА И ЭДВАРДСА

Определение эллиптической кривой над полем К в обобщенной форме Вейерштрасса базируется на уравнении 3-й степени

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \text{ , } a_i \in \text{К.}$$

Кривая Эдвардса вида

$$x^2 + y^2 = 1 + dx^2y^2$$

описывается уравнением 4-й степени, что может вызвать сомнения, что между этими уравнениями есть что-то общее, а кривая Эдвардса – вовсе не эллиптическая. Между тем Эдвардсу впервые удалось доказать, что последнее уравнение рациональным преобразованием сводится к частному виду уравнения в форме Вейерштрасса с ограниченными свойствами. С этим мы и хотим разобраться в этой главе.

В начале настоящей главы расмотрены несколько примеров из теории чисел, приводящих к эллиптическим кривым разных классов, среди которых могут возникать и кривые, изоморфные кривым Эдвардса (раздел 1.1). В разделе 1.2 рассмотрены оригинальные кривые Эдвардса, а в следующем разделе 1.3 – их модификация Бернстейном и Ланге. Наиболее важные свойства этих кривых – полнота закона сложения и его универсальность – доказываются в теоремах 1.1. и 1.2 [2]. Далее в разделах 1.5, 1.6 подробно излагается преобразование канонической кривой в форме Вейерштрасса в форму Эдвардса (в работе [2] изоморфизм сразу строится из формы Монтгомери), доказана теорема о необходимых и достаточных условиях существования точек 8-го порядка (раздел 1.7), даны простые оценки для расчета числа изоморфизмов и пар кривых кручения (раздел 1.8), рассмотрены примеры. Изоморфные кривые могут найти различные приложения, например, при формировании псевдослучайных последовательностей, обновлении

параметров в протоколах и пр. Для нас в первую очередь важно, что изоморфизм между четвертой частью всех кривых в форме Вейерштрасса и кривыми Эдвардса позволяет в итоге получить совершенно новое качество и уникальные свойства последних. В последнем разделе 1.9 на основе условий существования 2-х точек 4-го порядка предлагается алгоритм поиска приемлемой для криптографии канонической кривой и построения изоморфной ей кривой Эдвардса.

1.1. Некоторые задачи теории чисел и эллиптические кривые

История математики полна примерами классических задач теории чисел, приводящих к той или иной форме эллиптической кривой. Ниже мы приводим несколько таких примеров, которые иогут оказаться полезными при исследовании свойств кривых в форме Эдвардса.

Гипотеза Ферма

Еще в 17-м веке Ферма утверждал (правда, без доказательства), что уравнение

$$a^n + b^n = c^n,\ abc \neq 0, n \geq 3,$$

не имеет целочисленных решений. Это утверждение еще называют последней теоремой Ферма [24,26].

Пусть $n = 3,\ a + b \neq 0$ и примем

$$x = \frac{12c}{a+b}, \qquad y = \frac{36(a-b)}{a+b}.$$

Тогда с учетом равенства $a^3 + b^3 = (a + b)(a^2 - ab + b^2)$ получим

$$x^3 - 2^4 3^3 = 2^4 3^3 \left(\frac{4(a^3 + b^3)}{(a+b)^3} - 1\right) = 2^4 3^3 \left(\frac{4(a^2 - ab + b^3)}{(a+b)^2} - 1\right) =$$

$$= 2^4 3^4 \frac{(a-b)^2}{(a+b)^2} = y^2.$$

Итак, кубическое уравнение Ферма сводится к частноиу виду эллиптической кривой

$$y^2 = x^3 - 2^4 3^3.$$

Известно [16], что существуют лишь 2 рациональных решения этого уравнения $(x, y) = (12, \pm 36)$. Из $y = 36$ следует, что $b = 0$, тогда как значение $y = -36$ справедливо лишь при $a = 0$. Таким образом, при $n = 3$ не существует целых решений уравнения Ферма, таких что $abc \neq 0$.

Уравнение Ферма 4-й степени $a^4 + b^4 = c^4$ также сводится к уравнению эллиптической кривой заменой

$$x = \frac{2(b^2+c^2)}{a^2}, \qquad y = \frac{4b(b^2+c^2)}{a^3}.$$

Действительно, преобразование кубического полинома $x^3 - 4x$ с учетом уравнения Ферма $a^4 + b^4 = c^4$ дает

$$x^3 - 4x = x(x-2)(x+2) = \frac{2(b^2+c^2)}{a^2}\left(\frac{2(b^2+c^2)}{a^2} - 2\right)\left(\frac{2(b^2+c^2)}{a^2} + 2\right) =$$

$$= \frac{2^3}{a^6}(b^2+c^2)((b^2+c^2)^2 - a^4) = \frac{2^3}{a^6}(b^2+c^2)(b^4+c^4+2b^2c^2 - a^4) =$$

$$= \frac{2^3}{a^6}(b^2+c^2)(2b^4+2b^2c^2) = \frac{2^4}{a^6}b^2(b^2+c^2)^3 = y^2.$$

Таким образом, получили эллиптическую кривую вида

$$y^2 = x^3 - 4x,$$

рациональными точками которой являются лишь 3 точки 2-го порядка $(0,0), (-2,0)$ и $(2,0)$ (это доказывается в главе 8 [16]). Для этих точек из $y = 0$ следует $b = 0$, что вновь исключает нетривиальные целые решения уравнения Ферма $a^4 + b^4 = c^4$, $abc \neq 0$. Этот частный случай последней теоремы Ферма был доказан еще Эйлером [26] без привлечения понятия эллиптической кривой. Уравнение $y^2 = x^3 - 4x$ изоморфными преобразованиями приводит к вырожденной паре кривых кручения в классе скрученных кривых Эдвардса (глава 4).

Следующий пример будет также полезен для темы нашего исследования.

Пирамида из шаров.

Построим пирамиду из шаров (например, бильярдных), в вершине которой лежит один шар, под которым в квадрат уложены 2^2 шара, последние опираются на квадрат из 3^2 шаров и т.д. Если $x-$ число этажей такой пирамиды, то суммарное число шаров равно

$$S(x) = 1 + 2^2 + 3^2 + \cdots + x^2 = \frac{1}{3}x(x+1)\left(x+\frac{1}{2}\right).$$

Действительно, $S(1) = 1, S(2) = 5, S(3) = 14$ и т. д. Предположим, что приведенное выше равенство справедливо при натуральном x и докажем, что тогда оно справедливо и при $x+1$, т.е.

$$S(x+1) = 1 + 2^2 + 3^2 + \cdots + (x+1)^2 = \frac{1}{3}(x+1)(x+2)\left(x+\frac{3}{2}\right).$$

Рассмотрим разность

$$S(x+1) - S(x) = \frac{1}{3}(x+1)\left((x+2)\left(x+\frac{3}{2}\right) - x\left(x+\frac{1}{2}\right)\right) =$$

$$= \frac{1}{3}(x+1)\left(x^2 + \frac{7}{2}x + 3 - x^2 - \frac{1}{2}x\right) = (x+1)^2.$$

Таким образом, представление $S(x)$ вышеприведеным полиноиом 3-й степени справедливо при всех натуральных значениях x. В задаче пирамиды из шаров требуется найти число этажей x, при котором все шары пирамиды укладываются в квадрат со стороной y, т.е. $S(x) = y^2$ или

$$y^2 = \frac{1}{3}x(x+1)\left(x+\frac{1}{2}\right). \qquad (1.1)$$

Это уравнение является частным примером неприведенной формы эллиптической кривой [22], если допустить, что x является элементом некоторого поля. Над полем рациональных чисел **Q** кривая (1.1) имеет 3 точки 2-го порядка (0,0), $(-1,0)$ и $\left(-\frac{1}{2},0\right)$. Тривиальными решениями задачи о целых решениях уравнения (1.1) являются точки $(0,0)$и $(1,1)$. Над полем вещественных чисел **R** график этой кривой приведен на рис.1.1.

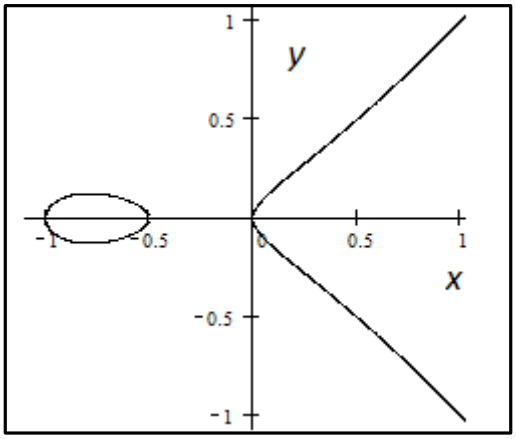


Рис.1.1. График кривой (1.1) над полем **R**

Кроме точек 2-го порядка, рациональными точками кривой (1.1) являются также точки $\pm P = \left(\frac{1}{2}, \pm\frac{1}{2}\right)$, $\pm Q = (1, \pm 1)$. Если провести прямую через любую пару этих рациональных точек, она пересечет кривую (1.1) в рациональной точке [22,24] . Например, прямая через точки $\left(\frac{1}{2}, \frac{1}{2}\right)$ и $(1,1)$ проходит через тривиальную точку $(0,0)$. Нетривиальную рациональную точку можно получить как третью точку пересечения кривой прямой линией, проходящей через точки $-P = \left(\frac{1}{2}, -\frac{1}{2}\right)$ и $Q = (1,1)$. Эта прямая описывается уравнением $y = 3x - 2$. Подставляя эту функцию в (1.1), получим

$$(3x-2)^2 = \frac{1}{3}x^3 + \frac{1}{2}x^2 + \frac{1}{6}x \Rightarrow \frac{1}{3}x^3 - \frac{17}{2}x^2 + \frac{37}{6}x - 4 = 0.$$

С другой стороны, корнями этого кубического уравнения являются x-координаты точек $-P, Q$ и 3-й точки $R = (x_3, y_3)$ пересечения кривой (1.1) и прямой $y = 3x - 2$, так что

$$\frac{1}{3}\left(x - \frac{1}{2}\right)(x-1)(x-x_3) = 0 \Rightarrow \frac{1}{3}x^3 - \frac{1}{3}\left(x_3 + \frac{3}{2}\right)x^2 + \cdots = 0$$

Из равенства коэффициентов при x^2 в двух последних кубическоих уравнениях получаем $\left(x_3 + \frac{3}{2}\right) = \frac{51}{2} \Rightarrow x_3 = 24$. Тогда $y_3 = 3x_3 - 2 = 70$. Единственным нетривиальным целочисленным решением уравнения (1.1) является точка $R = (24,70)$, при этом

$$1 + 2^2 + 3^2 + \cdots + 24^2 = 70^2 = 4900.$$

Итак, в один квадрат со стороной 70 шаров можно уложить все 4900 шаров пирамиды, имеющей 24 этажа и квадрат основания со стороной 24 шара.

Уравнение (1.1) легко приводится к форме Монтгомери и форме Вейерштрасса [16,17]. Заменой координат $x \to \frac{U}{2},\ y \to \frac{V}{2\sqrt{6}}$ получаем

$$V^2 = U(U+1)(U+2) = U^3 + 3U^2 + 2U. \qquad (1.2)$$

Пусть поле $F(\sqrt{2})$ содержит элемент $\mathrm{C} = \sqrt{2}$ и $C^2 = 2$. Тогда заменой $U \to \mathrm{C}u,\ V^2 \to C^3 v^2$ приходим к частному виду уравнения в форме Монтгомери

$$v^2 = u^3 + 3C^{-1}u^2 + u. \qquad (1.3)$$

Эта форма особенно удобна при изоморфном преобразовании ее в форму Эдвардса.

Уравнение (1.2) простым смещением всех корней $U \to X - 1, V \to Y$ приводится к частному виду эллиптической кривой в форме Вейерштрасса и Монтгомери

$$Y^2 = X(X-1)(X+1) = X^3 - X. \qquad (1.4)$$

Характерным для этой формы кривой является отсутствие 2-й степени в кубическом полиноме уравнения кривой (т.е. след этого полинома равен 0), при этом сумма корней кубики равна 0 [22,29]. График кривой (1.4) аналогичен рис.1.1 со смещением начала координат в точку (0,0) и преобразованием масштаба. Точки кривых (1.1) и (1.4) связаны линейным преобразованием координат $X = 2x + 1,\ Y = 2\sqrt{6}y.$ В частности, в координатах уравнения (1.4) точки $-P' = \left(2, -\sqrt{6}\right), Q' = \left(3, 2\sqrt{6}\right).$

Задачу о пирамиде из шаров можно также решить, складывая иррациональные точки $-P'$ и Q' по формуле суммы двух разных точек кривой (1.3) [29]

$$(X_1, Y_1) + (X_2, Y_2) = (X_3, Y_3) = \left(\lambda^2 - X_1 - X_2, -Y_1 - \lambda(X_3 - X_1)\right),$$

где $\lambda = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{3\sqrt{6}}{3-2} = 3\sqrt{6}.$

Отсюда $X_3 = 49, Y_3 = \sqrt{6} - 3\sqrt{6}(49-2) = -140\sqrt{6}.$ Возвращаясь к координатам кривой (1.1), получаем $x_3 = \frac{X_3 - 1}{2} = 24,\ y_3 = \frac{Y_3}{2\sqrt{6}} = -70.$ Точка пересечения кривой и прямой линии, проходящей через точки $-P$ и Q, симметрична относительно оси x и имеет координаты $(24,70)$. Решение совпадает с предыдущим.

Пример «пирамида из шаров» полезен в нашем исследовании: он, как и предыдущий пример, приводит к вырожденной паре кривых кручения Эдвардса (глава 4).

Задача о конгруэнтных числах

Исторически к некоторому виду эллиптических кривых над полем рациональных чисел привели исследования Пифагора, Евклида, Диофанта, Ферма при решении задачи о прямоугольном треугольнике с целочисленными длинами сторон («пифагоровы тройки»). Ее развитием стала задача о *конгруэнтных числах*.

Определение 1.1. *Натуральное число* n *называется конгруэнтным, если оно является площадью прямоугольного треугольника с рациональными длинами сторон.*

Способ построения пифагоровых троек весьма прост. Пусть *X, Y, Z* – стороны прямоугольного треугольника, рис.1.2, с площадью $n = \frac{1}{2}XY$. Нормируем все стороны к гипотенузе, тогда $v = Y/Z$, $u = X/Z$ и $v^2 + u^2 = 1$. В единичной окружности на рис.1.2 обозначим углы 2α (между гипотенузой и катетом *X*) и $\alpha < \pi/4$. Зададим целые числа *a* и *b* такие, что $a > b$ и $tg\alpha = b/a$. Тогда

$$v = \sin 2\alpha = \frac{2tg\,\alpha}{1+tg^2\alpha} = \frac{2ab}{a^2+b^2},$$

$$u = \cos 2\alpha = \frac{1-tg^2\alpha}{1+tg^2\alpha} = \frac{a^2-b^2}{a^2+b^2}.$$

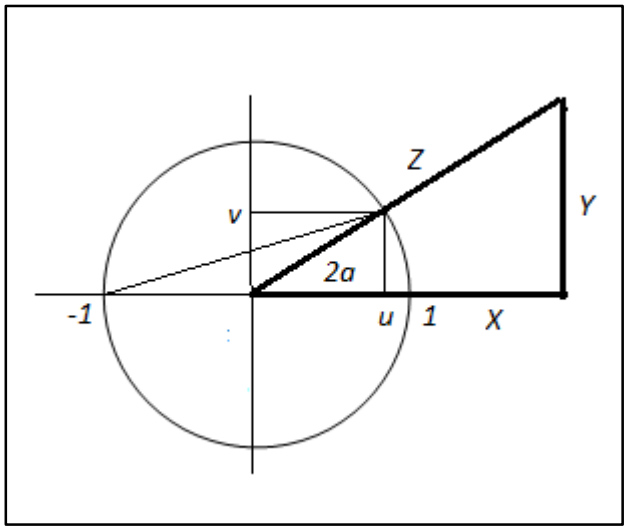


Рис.1..2. Геометрические построения для получения пифагоровых троек

Теперь из $X = a^2 - b^2$, $Y = 2ab$, $Z = a^2 + b^2$ легко получим целочисленные стороны прямоугольных треугольников. Например, при $b = 1$ первые 6 значений сторон треугольников и их площадей представлены в таблице 1.1. Значения n здесь являются по определению конгруэнтными числами. Вместе с тем кроме целочисленных сторон такие числа образуются рациональными сторонами. Скажем, площадь прямоугольного треугольника со сторонами 3/2, 20/3, 41/6 равна 5. Оказывается, это минимальное из конгруэнтных чисел.

Таблица 1.1. Некоторые примеры пифагоровых троек при $b = 1$ и значения параметра n

a	2	3	4	5	6	7
X	3	8	15	24	35	48
Y	4	6	8	10	12	14
Z	5	10	17	26	37	50
n	6	24	60	120	210	336

Пусть $x = (Z/2)^2$ – некоторое рациональное число и катет $X < Y$. Из системы уравнений

$$\begin{cases} \frac{1}{4}(X^2 + Y^2) = x, \\ \frac{1}{2}XY = n, \end{cases}$$

получим

$$\begin{cases} X = \sqrt{x+n} - \sqrt{x-n}, \\ Y = \sqrt{x+n} + \sqrt{x-n}, \\ Z = 2\sqrt{x}. \end{cases}$$

Отсюда доказывается утверждение, что n конгруэнтно тогда и только тогда, когда все три числа $x, x + n$ и $x – n$ являются квадратами рациональных чисел (см. [23,67]). Например, при $n = 5$ и $x = 1681/144$ имеем $x – n = 961/144$, $x + n = 2401/144$, т.е. все три числа являются квадратами.

Последнюю систему уравнений можно также переписать в виде

$$t = \sqrt{x^2 - n^2} = \frac{1}{4}(Y^2 - X^2), \qquad \sqrt{x} = \frac{Z}{2}.$$

Введя новую переменную $y = t\sqrt{x}$ и возводя ее в квадрат, получаем кубическое уравнение

$$y^2 = x(x^2 – n^2) = x^3 – n^2x, \qquad (1.5)$$

которое является частным случаем уравнения эллиптической кривой с целыми коэффициентами. Одна из рациональных точек этой кривой с координатой $x = (X^2 + Y^2)/4$ задает прямоугольный треугольник с рациональными сторонами и площадью n (если n – конгруэнтное число). Нахождение других рациональных решений уравнения (1.5) является, разумеется, более общей

задачей, чем задача о конгруэнтных числах. Кривая (1.5), как и кривые 2-х предыдущих примеров, изоморфна скрученной кривой Эдвардса.

1.2. Эллиптические кривые в оригинальной форме Эдвардса

В пионерской работе профессора университета Нью-Йорка Гарольда Эдвардса [1] рассматривались свойства эллиптической кривой в форме

$$x^2 + y^2 = e^2(1 + x^2y^2), \tag{1.6}$$

близкая к которой еще около 2-х веков назад встречалась в работах Эйлера и Гаусса (при $e = 1$ и замене знака «+» на «–» в правой части). Эти математики еще не знали, что уравнение (1.6) можно назвать эллиптической кривой, так как понятие это сформировалось почти век спустя после введения закона сложения точек кривой с образованием структуры абелевой группы точек. Эдвардсу впервые удалось доказать, что уравнение (1.6) описывает кривую, изоморфную кривой в форме Вейерштрасса, и получить закон сложения ее точек

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1y_2 + x_2y_1}{e(1+x_1x_2y_1y_2)}, \frac{y_1y_2 - x_1x_2}{e(1-x_1x_2y_1y_2)}\right). \tag{1.7}$$

Кривые (1.6) в этой связи называют *оригинальной формой Эдвардса*. Заметим, что нейтральным элементом здесь является точка $O = (0, e)$, а обратная точка определена как $-(x_1, y_1) = (-x_1, y_1)$. Из (1.7) следует, что $(x_1, y_1) + (0, e) = (x_1, y_1)$ и $(x_1, y_1) + (-x_1, y_1) = (0, e)$.

Кривые (1.6) существуют над всеми полями с нулевой характеристикой и над конечными полями F_p^m характеристики $p \neq 2$. Для них всегда существует точка 2-го порядка такая, что $2D_0 = O$. При совпадении слагаемых точек универсального закона (1.7) получаем в частном случае закон удвоения

$$2(x_1, y_1) = \left(\frac{2x_1y_1}{e(1+x_1^2y_1^2)}, \frac{y_1^2 - x_1^2}{e(1-x_1^2y_1^2)}\right). \tag{1.8}$$

Подставляя в правую часть точку $O = (0, e)$, получим решение для точки 2-го порядка $D_0 = (0, -e)$.

Для задач криптографии могут оказаться интересны лишь кривые вида (1.6) над полем F_q конечного порядка $q = p^m$. Очевидно, заменой $x \to \frac{x}{e}$ кривая (1.6) записывается в изоморфной форме

$$x^2 + y^2 = 1 + e^4 x^2 y^2 \quad \Rightarrow \quad y^2 = \frac{1-x^2}{1-e^4 x^2}, e^4 \neq 1. \qquad (1.9)$$

При $e = 1$ при всех значениях x имеем 2 решения $y = \pm 1$, и порядок такой кривой $N_E = 2q$ выходит за границы Хассе [29], кривая не является эллиптической. Кроме того, возникают особые случаи в законе удвоения. Например, удвоение точки $P = (1,1)$, являющейся решенем уравнения (1.6), порождает неопределенность 0/0 для y -координаты в (1.8). Следует поэтому принять $e^4 \neq 1$,, тогда число решений уравнения (1.9) ограничивается числом элементов e^4 поля, порождающих квадраты в правой части уравнения.

Для точки F_0 4-го порядка кривой (1.6), принимая $2F_0 = D_0$, получим согласно (1.8)

$$\frac{2x_1 y_1}{e(1+x_1^2 y_1^2)} = 0, \qquad \frac{y_1^2 - x_1^2}{e(1-x_1^2 y_1^2)} = -e.$$

Отсюда $y_1^2 = 0 \Rightarrow x_1^2 = e^2 \Rightarrow x_1 = \pm$e. Итак, для кривой (1.6) при $e^4 \neq 1$ над конечным полем характеристики $p \neq 2,3$, всегда существуют 2 точки 4-го порядка $\pm F_0 = (\pm e, 0)$. Сразу заметим, что найденными выше не ограничиваются все точки 2-го и 4-го порядков. В частности, всегда имеются еще две особые точки 2-го порядка (на бесконечности), и кривая (1.6) является нециклической (с тремя точками 2-го порядка).

Действительно, из уравнения кривой (1.6) справедливо

$$y^2 = \frac{e^2 - x^2}{1 - e^2 x^2}, \qquad x^2 = \frac{e^2 - y^2}{1 - e^2 y^2}.$$

При нулевых значениях знаменателей этих равенств получаем 4 особые точки кривой: $F_{1,2} = (\pm e^{-1}, \infty)$, $D_{1,2} = (\infty, \pm \mathrm{e}^{-1})$. Здесь знаком "$\infty$" обозначено деление на 0. Хотя в конечном поле эти элементы не определены, но в групповых операциях (1.7) и (1.8), имеющих вид рациональных функций, обе координаты точек входят в числители и знаменатели. Это позволяет пользоваться формулами (1.7) и (1.8) в особых точках, принимая правила

обычного предельного перехода (к этому вопросу мы вернемся в главе 2). Тогда с помощью (1.8) получим

$$2D_{1,2} = 2(\infty, \pm e^{-1}) = \left(0, \frac{\infty^2}{e e^{-2} \infty^2}\right) = (0, e) = O.$$

$$2F_{1,2} = 2(\pm e^{-1}, \infty) = \left(0, \frac{\infty^2}{-e e^{-2} \infty^2}, \right) = (0, -e) = D_0.$$

Отсюда следует, что особые точки $D_{1,2}$ имеют порядок 2, а особые точки $F_{1,2}$ – порядок 4. Оригинальные кривые Эдвардса, таким образом, имеют свойства нециклических кривых.

Аналогично, для точки S 8-го порядка с учетом равенства $2S = F$ имеем

$$\frac{2x_1y_1}{e(1 + x_1^2y_1^2)} = e, \quad \frac{y_1^2 - x_1^2}{e(1 - x_1^2y_1^2)} = 0.$$

Тогда с учетом (1.6)

$$y_1^2 = x_1^2 \Rightarrow x_1^4 - 2e^{-2}x_1^2 + 1 = 0 \Rightarrow x_1^2 = e^{-2}\left(1 \pm \sqrt{1 - e^4}\right).$$

Здесь мы видим, что точки 8-го порядка существуют лишь в случае, когда выражение в скобках существует и является квадратом. Далее в главе 2 мы покажем, что уравнение (1.6) приводит к так называемым *квадратичным кривым Эдвардса*, с минимальным кофактором 8 порядка кривой $N_E = 8n$ (n – нечетное). Это определяется нециклической структурой данной кривой с возникновением 3-х точек 2-го порядка, в результате чего с учетом точек 4-го порядка кривая всегда содержит нециклическую подгруппу 8-го порядка. Наличие точек 8-го порядка лишь увеличивает значение кофактора порядка кривой до 16, 32 и т.д. Кроме того, как мы убедились, среди точек 2-го и 4-го порядков кривой (1.6) имеются особые точки (на бесконечности), для которых следует вводить арифметику сложения точек. Отмеченные недостатки кривых в оригинальной форме Эдвардса делают их практически неинтересными для криптографических приложений.

1.3. Эллиптические кривые в форме Эдвардса с модификацией Бернстейна-Ланге

Вскоре после работы [1] появилась замечательная работа специалистов по криптографии Даниэля Бернстейна и Тани Ланге [2], в которой предложена модификация кривой (1.6) с введением неквадратичного параметра d над конечным полем F_p^m характеристики $p \neq 2$ вида

$$E: \quad x^2 + y^2 = e^2(1 + dx^2y^2), \quad d(1 - de^4) \neq 0, \left(\frac{d}{p}\right) = -1, \qquad (1.10)$$

где $\left(\frac{d}{p}\right) -$ символ Лежандра, и параметр $d -$ квадратичный невычет [29,67]. Универсальный закон сложения для точек этой кривой имеет вид

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1y_2 + x_2y_1}{e(1+dx_1x_2y_1y_2)}, \frac{y_1y_2 - x_1x_2}{e(1-dx_1x_2y_1y_2)}\right). \qquad (1.11)$$

Закон удвоения для совпадающих точек, соответственно, записывается как

$$2(x_1, y_1) = \left(\frac{2x_1y_1}{e(1+dx_1^2y_1^2)}, \frac{y_1^2 - x_1^2}{e(1-dx_1^2y_1^2)}\right). \qquad (1.12)$$

Принципиальными отличиями кривой (1.10) от (1.6) являются циклическая структура группы точек (в отношении точек 2-го порядка) и отсутствие особых точек (с делением на 0 в законе сложения). Последнее свойство определено в [2] как *полнота закона сложения*. Как и для кривой (1.6), обратная точка определена как $-(x_1, y_1) = (-x_1, y_1)$, нулем группы точек (нейтральным элементом аддитивной группы точек) здесь является точка $O = (0, e)$, но существуют лишь единственная точка 2-го порядка $D = (0, -e)$ и ровно 2 точки 4-го порядка $\pm F = (\pm e, 0)$.

Важной является доказанная в [2] теорема о полноте закона сложения.

Теорема 1.1. *Для любых пар точек кривой* (1.10) *знаменатели закона сложения* (1.11) *не обраются в нуль*: $dx_1x_2y_1y_2 \neq \pm 1$.

Доказательство.

Допустим обратное, например, $dx_1x_2y_1y_2 = 1$. Тогда $x_1y_1 = \frac{1}{dx_2y_2}$ и

$$x_1^2 + y_1^2 = e^2\left(1 + \frac{1}{dx_2^2\, y_2^2}\right) = \frac{x_2^2+y_2^2}{dx_2^2\, y_2^2},$$

$$(x_1 + y_1)^2 = x_1^2 + y_1^2 + 2x_1y_1 = \frac{x_2^2+y_2^2}{dx_2^2\, y_2^2} + \frac{2}{dx_2y_2} = \frac{(x_2^2+y_2^2+2x_2y_2)}{dx_2^2\, y_2^2} = \frac{(x_2+y_2)^2}{dx_2^2\, y_2^2}.$$

Аналогично получим

$$(x_1 - y_1)^2 = x_1^2 + y_1^2 - 2x_1y_1 = = \frac{(x_2^2 + y_2^2 - 2x_2y_2)}{dx_2^2\, y_2^2} = \frac{(x_2 - y_2)^2}{dx_2^2\, y_2^2}.$$

Оба равенства в силу неквадратичности параметра d выполняются лишь тогда, когда одновременно $x_1 + y_1 = 0$ и $x_1 - y_1 = 0$, и, следовательно, $x_1 = y_1 = 0$. Но кривая (1.10) не содержит такой точки. Итак, допущение $dx_1x_2y_1y_2 = 1$ приводит к противоречию.

Подобные же результаты получим при втором допущении $dx_1x_2y_1y_2 = -1$:

$$(x_1 + y_1)^2 = x_1^2 + y_1^2 + 2x_1y_1 = \frac{x_2^2 + y_2^2}{dx_2^2\, y_2^2} - \frac{2}{dx_2y_2} = \frac{(x_2^2 + y_2^2 - 2x_2y_2)}{dx_2^2\, y_2^2} = \frac{(x_2 - y_2)^2}{dx_2^2\, y_2^2},$$

$$(x_1 - y_1)^2 = x_1^2 + y_1^2 - 2x_1y_1 = \frac{x_2^2 + y_2^2}{dx_2^2\, y_2^2} + \frac{2}{dx_2y_2} = \frac{(x_2^2 + y_2^2 + 2x_2y_2)}{dx_2^2\, y_2^2} = \frac{(x_2 + y_2)^2}{dx_2^2\, y_2^2},$$

приводящие к тому же выводу. Теорема доказана. ▲

Сформулированное в теореме 1.1 свойство авторы [2] назвали *полнотой* закона сложения. В этой связи в работе [3] они отнесли их к классу *полных кривых Эдвардса* (*complete Edwards curve*). Далее в этой главе мы рассматриваем некоторые из свойств лишь этих кривых.

Справедливость закона сложения (1.11) доказывается следующей теоремой 1.2 [2]. Мы ниже дадим ее более детальное по сравнению с [2] и понятное доказательство. Для упрощения записи в громоздких выкладках мы здесь избегаем индексации точек путем замены $(x_1, y_1) \to (x, y)$, $(x_2, y_2) \to (u, v)$, $(x_3, y_3) \to (X, Y)$.

Теорема 1.2. *Пусть* $P = (x, y)$ и $Q = (u, v) -$ *точки кривой* (1.10) *и выполняются равенства* $x^2 + y^2 = e^2(1 + dx^2y^2)$, $u^2 + v^2 = e^2(1 + du^2v^2)$. *Определим* $X = \frac{(xy+ uv)}{e(1+dxyuv)}$, $Y = \frac{(yv- xu)}{e(1-dxyuv)}$. *Тогда* $X^2 + Y^2 = e^2(1 + dX^2\, Y^2)$.

Доказательство. Примем сначала $e = 1$ и обозначим

$$a = xv,\ b = yu,\quad c = yv,\quad h = xu,\quad \Rightarrow\quad ab = ch\,. \tag{1.13}$$

Из уравнения кривой при *e* =1 определим единицу как равенство

$$V = X^2 + Y^2 - dX^2Y^2 =$$

$$\frac{(a+\ b)^2(1-dab)^2 + (c-h)^2(a+\ b)^2(1+dab)^2 - d(a+\ b)^2(c-h)^2}{(1-d^2a^2b^2)^2}.$$

Его можно привести к виду

$$V = \frac{(1+d^2a^2b^2)(a^2+b^2+c^2+h^2) - d(a^2+b^2)(c^2+h^2) - 4da^2b^2}{(1-d^2a^2b^2)^2}.$$

С учетом (1.13)

$$(a^2+b^2+c^2+h^2) = (x^2+y^2)(u^2+v^2),$$

$$(a^2+b^2)(c^2+h^2) - 4da^2b^2 = x^2y^2(u^2+v^2)^2 + u^2v^2(x^2+y^2)^2.$$

Тогда

$$V = \frac{(1+d^2a^2b^2)(x^2+y^2)(u^2+v^2) - d[x^2y^2(u^2+v^2)^2 + u^2v^2(x^2+y^2)^2]}{(1-d^2a^2b^2)^2}.$$

Числитель этого выражения можно преобразовать как

$$d^2a^2b^2(x^2+y^2)(u^2+v^2) -\ dx^2y^2(u^2+v^2)^2 +$$

$$+(x^2+y^2)(u^2+v^2) - du^2v^2(x^2+y^2)^2 = d^2a^2b^2(x^2+y^2)(u^2+v^2) +$$

$$+(x^2+y^2)\big(u^2+v^2-du^2v^2(x^2+y^2)\big) - dx^2y^2(u^2+v^2)^2 =$$

$$= \big(x^2+y^2-dx^2y^2(u^2+v^2)\big)\big(\,u^2+v^2-du^2v^2(x^2+y^2)\big).$$

Итак,

$$V = \frac{\big(x^2+y^2-dx^2y^2\big(u^2+v^2\big)\big)\big(\,u^2+v^2-du^2v^2\big(x^2+y^2\big)\big)}{(1-d^2a^2b^2)^2}. \tag{1.14}$$

Из определения точек кривой в общем случае имеем:

$x^2 + y^2 = e^2(1 + dx^2y^2), \qquad u^2 + v^2 = e^2(1 + du^2v^2).$

Умножим второе равенство на $-dx^2y^2$ и вычтем результат из первого, тогда получим

$$x^2 + y^2 - dx^2y^2(u^2 + v^2) = e^2(1 - d^2x^2y^2u^2v^2).$$

Аналогичный результат получим умножением первого равенства на $-du^2v^2$ с вычитанием результата из второго

$$u^2 + v^2 - du^2v^2(x^2 + y^2) = e^2(1 - d^2x^2y^2u^2v^2).$$

Подставляя эти соотношения в (1.14), получим $V = e^4$. Возвращаясь к началу доказательства, заметим, что величина V была определена как

$$V = X^2 + Y^2 - dX^2Y^2 = 1$$

из уравнения кривой при $e = 1$. Для этого случая теорема доказана. Переходя к общему случаю линейным преобразованием координат $X \to eX,\ Y \to eY$, получим

$$e^2X^2 + e^2Y^2 - e^4dX^2Y^2 = e^4 \quad \Rightarrow \quad X^2 + Y^2 = e^2(1 + dX^2Y^2).$$

Теорема доказана. ▲

Надо полчеркнуть, что уравнение (1.10) содержит избыточный параметр e. В самом деле, подстановкой $\frac{x}{e} \to x,\ \ \frac{y}{e} \to y$ получим изоморфную кривую в форме Эдвардса

$$x^2 + y^2 = (1 + d'x^2y^2), \quad d' = de^4.$$

Такая форма кривой Эдвардса, определяемая единственным параметром d', проще и предпочтительней (1.10), так как сокращает 2 умножения в поле. Поэтому с точностью до изоморфизмов в формуле (1.10) чаще всего принимают $e = 1$.

1.4. Трансформация эллиптической кривой в форме Эдвардса в форму Вейерштрасса

Каноническая форма эллиптической кривой (или форма Вейерштрасса (Weierstrass)) над конечным полем F_q ($q = p^m,\ p \neq 2$) имеет вид

$$W: \quad v^2 = u^3 + au + b, \quad a, b \in F_q\ . \tag{1.15}$$

Изоморфизм между кривыми в форме Вейерштрасса и Эдвардса в форме (1.10) может обеспечить не всякая пара параметров a, b в (1.15). Необходимым и достаточным условием изоморфизма является существование на кривой (1.15) ровно 2-х точек $\pm F$ 4-го порядка (и, соответственно, единственной точки D 2-го порядка).

Один из видов преобразования кривой ($W \rightarrow E$) рассмотрен в работе [16]. Кривая (1.15) в ней представлена в форме

$$v^2 = (u - 1 - e^4 d)(u^2 - 4e^4 d). \tag{1.16}$$

Сразу заметим, что парабола $(u^2 - 4e^4 d)$ в правой части уравнения не имеет корней в поле F_q в силу неквадратичности параметра d, поэтому кривая (1.16) имеет единственную точку 2-го порядка $(1 + e^4 d, 0)$.

Рациональная замена переменных

$$u = \frac{-2e(w-e)}{x^2},\ v = \frac{4e^2(w-e)+2ex^2(e^4 d+1)}{x^3},\ w = (e^2 dx^2 - 1)y, \tag{1.17}$$

трансформирует кривую (1.10) в частную форму (1.16) Вейерштрасса. Уравнение (1.10) можно записать как

$$x_1^2 - e^2 = y_1^2(e^2 dx_1^2 - 1) = \frac{w^2}{e^2 dx_1^2 - 1}.$$

Тогда

$$w^2 = (x_1^2 - e^2)(e^2 dx_1^2 - 1).$$

В работе [15] формулируется (без доказательства) утверждение 2.18: *при выполненнии замены* (1.17) *кривая* (1.10) *над конечным полем* F_q *изоморфна кривой* (1.16) *в частной форме Вейерштрасса*. Доказательство его требует очень громоздких преобразований и мы его опускаем. Справедливость утверждения, однако, без труда проверяется частными примерами.

Более прозрачным является подход, основанный на использовании преобразования эллиптической кривой вида (1.15) в форму Монтгомери [2]. Мы обсуждаем его в следующем параграфе. В отличие от [2], мы начинаем с канонической кривой в форме Вейерштрасса (1.15).

1.5. Трансформация эллиптической кривой в форме Вейерштрасса в форму Монтгомери

В работах [30,41–42] дан детальный анализ параметров a, b кривой (1.15), порождающих изоморфные кривые в форме Монтгомери и Эдвардса. Уравнение эллиптической кривой в форме Монтгомери имеет вид

M: $\quad v^2 = u^3 + Au^2 + Gu, \quad A,G \in F_q.$ (1.18)

Пусть c – единственный корень кубического полинома (кубики) в правой части уравнения (1.15). Тогда это уравнение можно переписать в виде

$Y^2 = (X - c)(X^2 + cX + a + c^2),\ b = -c(a + c^2).$ (1.19)

Из равенства $c^3 + ac + b = 0$ в этом уравнении следует, что c – корень кубики. Заменой $X - c \to u,\ Y \to v$ получим уравнение в форме Монтгомери (1.18), в котором

$v^2 = u^3 + 3cu^2 + (a + 3c^2)u \ \to A = 3c,\ G = (a + 3c^2),$ (1.20)

Далее вместо пары параметров a, b нам будет удобно использовать параметры a, c, при этом в соответствии с (1.19) $b = -c(a + c^2)$.

Определим условия, накладываемые на параметры a, c, при которых имеется единственная точка 2-го порядка и ровно 2 точки 4-го порядка. Второй задачей в этом разделе будет нахождение зависимости между параметрами a и c канонической формы эллиптической кривой и параметром d кривой $x^2 + y^2 = 1 + dx^2y^2$ в форме Эдвардса.

Теорема 1.3. *Необходимыми и достаточными условиями существования единственной точки* 2*-го и двух точек* 4*-го порядков кривой* (1.20) *являются:*

(i) $\left(\frac{-(3c^2+4a)}{p}\right) = -1,$ (ii) $\left(\frac{(3c^2+a)}{p}\right) = 1.$

Доказательство этой теоремы мы дадим в главе 3 (теорема 3.6). Здесь лишь отметим, что условие (i) определяет единственность точки 2-го порядка, а условие (ii) – наличие ровно 2-х точек 4-го порядка.

В процессе доказательства получены квадраты для координат точек 4-го порядка

$$u_1^2 = 3c^2 + a = \delta, \qquad v_1^2 = u_1^2(2u_1 + 3c). \tag{1.21}$$

Отсюда следует, что параметр G в (1.18) и (1.20) должен быть квадратом, или

$$\left(\frac{\delta}{p}\right) = \left(\frac{(3c^2+a)}{p}\right) = 1.$$

Из последнего выражения в (1.21) можно теперь получить

$$3\mathrm{c} = \frac{v_1^2}{u_1^3}\left(1 - 2\frac{u_1^3}{v_1^2}\right)u_1 = 2\frac{1+d}{1-d}u_1, \qquad d = 1 - 4\frac{u_1^3}{v_1^2}. \tag{1.22}$$

Первая формула в (1.22) позволяют выразить параметр d через параметры a и c канонической формы кривой

$$d = \frac{3c-2u_1}{3c+2u_1}, \quad u_1 = (-1)^s\sqrt{3c^2 + a},\ s \in \{0,1\}, \tag{1.23}$$

Итак, с учетом (1.21) и (1.22) коэффициенты в уравнении (1.20) равны

$$A = 3c = 2\frac{1+d}{1-d}u_1, \qquad G = (a + 3c^2) = u_1^2.$$

Тогда это уравнение принимает вид

$$v^2 = u^3 + 2\frac{1+d}{1-d}u_1u^2 + u_1^2u.$$

Заменой $v^2 \to \frac{1}{1-d}v^2$, делением правой части на u_1^3 и заменой $\frac{u}{u_1} \to u$ уравнение (1.20) в форме Монтгомери теперь может быть приведено к виду, зависящему лишь от одного параметра d

$$M\text{:} \qquad \frac{1}{1-d}v^2 = u^3 + 2\frac{1+d}{1-d}u^2 + u. \tag{1.24}$$

Если левую часть этого уравнения умножить на квадратичный невычет d, то соответствущие решения уравнения (точки кривой) обращаются в «не решения» (дырки) и наоборот. Это справедливо для всех точек, кроме точек 2-го порядка, сохраняющих свои u-координаты. Тогда получаем уравнение для кривой *квадратичного кручения* (*quadratic twist*) [2,3,29]

$$M^t\text{:} \qquad \frac{d}{1-d}v^2 = u^3 + 2\frac{1+d}{1-d}u^2 + u. \tag{1.25}$$

Пару кривых (1.24), (1.25) еще называют *парой кривых кручения.* Переход от одной из кривых к другой осуществляется простой заменой $d \to d^{-1}$. Действительно, уравнение (1.24) после такой замены имеет вид

$$-\frac{d}{1-d}v^2 = u^3 - 2\frac{1+d}{1-d}u^2 + u.$$

Тогда, подставляя $(-u) \to u$, получим уравнение (1.25). Если порядок кривой (1.24) равен $N_E = q + 1 - t$, то порядок кривой кручения $N_E^t = q + 1 + t$, (где t – след уравнения Фробениуса [22]) симметричен относительно среднего значения $q + 1$.

Заметим, что для кривых Эдвардса порядок кривой $N_E = 0 mod 4$, поэтому след уравнения Фробениуса *t* может быть равен 0 лишь для значений модуля $p = 3\ mod 4$. В этом случае элемент поля (– 1) является квадратичным невычетом, и при значении $d = d^{-1} = -1$ пара кривых кручения вырождается в одну суперсингулярную кривую с порядком $N_E = q + 1$. Это следует также из уравнения (1.24), которое при $d = -1$ принимает вид $v^2 = u^3 + u$. В форме Вейерштрасса (1.15) это кривая с коэффициентом $b = 0$.

Ограничения для параметра d. Так как $u_1 \neq 0, d \neq 1$. Если допустить, что $d = 0$, то в уравнении (1.24) появляются кратные корни кубики, т.е. нарушается несингулярность кривой [22]. Требование единственности точки второго порядка эквивалентно тому, что дискриминант правой части уравнения (1.24) или (1.25) после выделения корня $u = 0$ должен быть невычетом

$$\Delta = 4\left(\left(\frac{1+d}{1-d}\right)^2 - 1\right) = \frac{16d}{(1-d)^2} \neq C^2 \Rightarrow \left(\frac{d}{p}\right) = -1.$$

Отсюда следует, что d – квадратичный невычет в поле F_q.

Форма кривой (1.24) с помощью сравнительно несложной замены переменных $(u.v) \to (x, y)$ [2,30] приводится к изоморфной кривой в форме Эдвардса ($M \to E$).

1.6. Трансформация кривой Монтгомери в форму Эдвардса

Прямое и обратное преобразование координат кривых в форме Монтгомери и форме Эдвардса $(u, v) \Leftrightarrow (x, y)$ задается рациональными функциями [2]

$$x = 2\frac{u}{v}, \qquad y = \frac{u-1}{u+1}, \tag{1.26}$$

$$u = \frac{1+y}{1-y}, \qquad v = 2\frac{1+y}{1-y}\,x. \tag{1.27}$$

Умножение уравнения (1.24) на $(1 - d)/u^2$ дает

$$\left(\frac{v}{u}\right)^2 = (1-d)(u + u^{-1}) + 2(1+d).$$

С учетом (1.26) и (1.27) получим

$$\frac{2}{x^2} = (1+d) + (1-d)\frac{1+y^2}{1-y^2} \Rightarrow$$

$$\Rightarrow 2(1-y^2) = x^2(1-y^2)(1+d) + x^2(1+y^2)(1-d) = 2x^2 - 2dx^2y^2.$$

Отсюда получаем уравнение кривой Эдвардса (1.10) при $e = 1$

$$E: \qquad x^2 + y^2 = 1 + dx^2y^2.$$

Как отмечалось в разделе 1.3, оно легко может быть преобразовано в уравнение изоморфной кривой (1.10) с произвольным значением e. Эта кривая изоморфна (или *бирационально эквивалентна* [2]) кривой Монтгомери (1.24). Как следует из этого раздела, на основе простых рациональных замен координат (1.26) и (1.27) трансформация ($M \rightarrow E$) гораздо проще, чем трансформация из формы Вейерштрасса ($W \rightarrow E$), которая обсуждалась в разделе 1.4.

1.7. Точки малых порядков кривой Эдвардса

Любая кривая Эдвардса в форме (1.10) при $e = 1$ (далее без оговорок принимаем это значение) содержит одну точку 2-го порядка $D = (0, -1)$ и 2

точки 4-го порядка $\pm F = (\pm 1, 0)$. Эти точки вместе с точкой $O = (0, 1)$ являются исключительными точками, лежащими на осях x, y (т.е. других точек на осях не существует).

Интересными являются следующие свойства точек 2-го и 4-го порядков, вытекающие из закона сложения (1.11):

$$\begin{aligned} P + D &= (x_1, y_1) + (0, -1) = (-x_1, -y_1) = P^*, \\ P + F &= (x_1, y_1) + (1,0) = (y_1, -x_1), \\ P - F &= (x_1, y_1) + (-1,0) = (-y_1, x_1). \end{aligned} \qquad (1.28)$$

Мы видим, что сложение точки P с точкой D инвертирует знаки обеих координат исходной точки, тогда как ее сложение с точками $\pm F$ меняет координаты точки P местами с инверсией одного из знаков.

В работе [30] впервые дан анализ условий существования точек 8-го порядка кривой (1.5). Здесь мы дадим более строгое доказательство.

Теорема 1.4. *Необходимым и достаточным условием существования* 4-х точек *8-го порядка кривой* (1.10) *является*

$$\left(\frac{1-d}{p}\right) = 1.$$

Доказательство. *Необходимость.* Пусть $Ord(S) = 8$, тогда $2S = F$. В соответствии с (1.12) для координат $S = (x_1, y_1)$ имеем

$$\frac{2x_1y_1}{(1+dx_1^2y_1^2)} = 1, \quad \frac{y_1^2 - x_1^2}{(1-dx_1^2y_1^2)} = 0.$$

Отсюда $y_1^2 = x_1^2 \Rightarrow dx_1^4 - 2x_1^2 + 1 = 0 \Rightarrow x_1^2 = 1 \pm \sqrt{1-d}$.

Следовательно, условие $\left(\frac{1-d}{p}\right) = 1$ теоремы является необходимым условием существования координат $x_1 = \pm y_1$.

Достаточность. Докажем, что условие теоремы всегда порожлает ровно 4 точки 8-го порядка. Так как произведение $\left(1 + \sqrt{1-d}\right)\left(1 - \sqrt{1-d}\right) = d$, то одно из значений в равенстве $x_1^2 = 1 \pm \sqrt{1-d}$ является квадратичным вычетом, другое – квадратичным невычетом. Выбирая двоичным δ квадрат из этой альтернативы $\left(1 + (-1)^{\delta}\sqrt{1-d}\right)$, $\delta \in \{0,1\}$, получим 4 точки $(\pm x_1, \pm x_1)$ 8-го порядка. Теорема доказана. ▲

В отношении точек 8-го порядка сформулируем и докажем следующие утверждения.

Утверждение 1.1. *Пусть* $p \equiv 1 mod 4$. *Тогда одна из кривых пары кручения Эдвардса, для которой* $\left(\frac{1-d}{p}\right) = 1$, *имеет точки* 8-го *порядка* ($N_E \equiv 0 mod 8$), *а другая кривая – не имеет* ($N_E \equiv 4 mod 8$).

Доказательство. Пусть выполняется условие теоремы 1.4 и существуют 4 точки 8-го порядка. Отсюда следует, что ($N_E \equiv 0 mod 8$). Перейдем к кривой кручения обращением $d \to d^{-1}$. Тогда элемент поля

$$1 - d^{-1} = (-1)\frac{1-d}{d} \qquad (1.29)$$

при $p \equiv 1 mod 4$ является невычетом, так как (-1) – вычет в поле [29], а d – невычет. Для кривой кручения, таким образом, условие теоремы 1.4 не выполняется, точки 8-го порядка отсутствуют и порядок кривой кратен 4: ($N_E \equiv 4 mod 8$).

Утверждение 1.2. Пусть $p \equiv 3 \mathrm{mod} 4$. *Тогда при* $\left(\frac{1-d}{p}\right) = 1$ *обе кривые пары кручения Эдвардса имеют точки* 8-*го порядка* ($N_E \equiv 0 mod 8$), *а при* $\left(\frac{1-d}{p}\right) = -1$ *– не имеют* ($N_E \equiv 4 mod 8$).

Доказательство. Пусть выполняется условие теоремы 1.4 и существуют 4 точки 8-го порядка. Отсюда следует, что ($N_E \equiv 0 mod 8$). Перейдем к кривой кручения обращением $d \to d^{-1}$. Тогда элемент поля (1.29) при $p \equiv 3 mod 4$ является квадратом, так как (-1) и d – квадратичные невычеты в поле [29]. Для кривой кручения, таким образом, также выполняется условие теоремы 1.4 и ($N_E \equiv 0 mod 8$). В противном случае при $\left(\frac{1-d}{p}\right) = -1$ приходим к обратным выводам и ($N_E \equiv 4 mod 8$).

Для последнего случая ($p \equiv 3 \mathrm{mod} 4$) особым является значение невычета $d = -1$. При этом $d^{-1} = d$, и пара кручения вырождается в одну суперсингулярную кривую с порядком $N_E = q + 1$ и следом Фробениуса $t = 0$.

На основе утверждений 1.1 и 1.2 можно прийти к выводу, что приблизительно половина кривых имеет точки 8-го порядка (и, возможно, более высоких степеней 2), а половина – нет. Для криптоприложений, как известно, следует выбирать кривые с минимальным кофактором простого числа, т.е. избегать кривых с точками 8-го порядка. Для этого, как следует из анализа выше, целесообразно ввести дополнительное ограничение на

параметр кривой d: произведение $d(1-d)$ должно быть квадратичным вычетом в мультипликативной группе поля F_q.

Кроме 4-х точек на осях x и y и, возможно, точек 8-го порядка на прямых $y = \pm x$, остальные точки кривой Эдвардса собираются в семейства по 8 точек $(\pm x_i, \pm y_i), (\pm y_i, \pm x_i)$, расположенных на концентрических окружностях с центром в начале координат. Любая из этих точек в соответствии с (1.28) определяет все семейство. Сумма любой пары точек этого семейства дает одну из точек O, D или $\pm F$. Вместе с тем порядки точек всегда совпадают только для обратных точек. Подробнее этот вопрос будет рассмотрен в главе 3.

Пример 1.1. Каноническая эллиптическая кривая в форме Вейерштрасса (1.15) $y^2 = (x^3 + 2x + 5) mod 13$ имеет порядок $N_E = 12$ и единственный корень с $= 8$ кубики. Ее параметры a и c отвечают условиям теоремы 1.3. Изоморфная ей кривая Монтгомери согласно (1.20) имеет вид $v^2 = u^3 - 2u^2 - u$. С помощью (1.21), (1.23) находим ее точки 4-го порядка $\pm F = (8, \pm 5)$ и параметр $d = 8$. Итак, изоморфная канонической кривой W кривая Эдвардса E имеет вид $x^2 + y^2 = (1 + 8x^2y^2) mod 13$ и порядок $N_E = 12$ (след уравнения Фробениуса $t = 2$). Ее точки представлены на рис.1.1a. На рис.1.1b даны точки кривой кручения с параметром $d' = d^{-1} = 5$ и порядком $N_E^t = p + 1 + t = 16$. Эти изображения с двойной симметрией напоминают калейдоскоп.

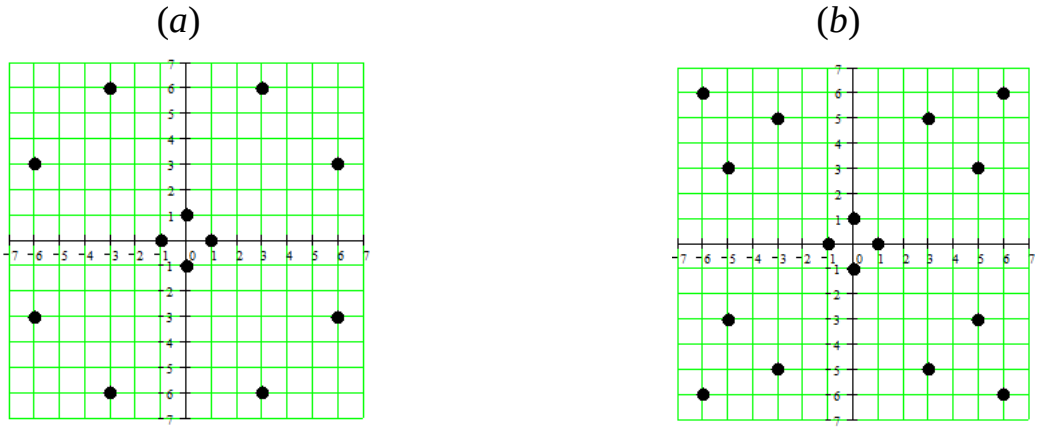


Рис.1.1.Точечные графики кривых пары кручения над полем $\mathbf{F}_{13}$ при $d = 8\,(a)$ и $d^{-1} = 5\,(b)$

Первая кривая содержит 2 точки 3-го порядка. Из равенства $2Q = -Q$ формулы сложения (1.11) и уравнения кривой (1.10) получим равенство, связывающее координаты x_Q, y_Q точки 3-го порядка

$$\frac{2x_Q y_Q}{1+dx_Q^2 y_Q^2} = -x_Q \Rightarrow \; y_Q^2 + x_Q^2 + 2y_Q = 0, \Rightarrow \; y_Q = -1 \pm \sqrt{1-x_Q^2}.$$

Для нашего примера точку 3-го порядка легко найти: при $x_Q = 3$, $y_Q = -1 \pm 2$, таких точек не существует; при $x_Q = 6$, $y_Q = -1 \pm 4$, тогда точка $Q = (6,3)$ – одна из точек 3-го порядка (второй является точка $-Q = (-6,3)$). Точки $\pm R = \pm Q + D = (\pm 6, -3)$ имеют порядок 6, и, следовательно, 4 точки с координатами $x_Q = \pm 3$ – порядок 12. Любая из последних генерирует всю группу. В главе 3 мы рассмотрим оригинальный метод нахождения точек максимального порядка, не требующий рутинной процедуры выполнения скалярного произведения *kP* точки *P*.

Кривая кручения (рис.1.1*b*) с порядком $N_E^t = 16$ имеет точки 8-го порядка ($\pm 6, \pm 6$). В соответствии с утверждением 1.1 для ее параметра $d' = d^{-1} = 5$ выполняется свойство $(1 - d') = 9 mod 13$ – квадратичный вычет в поле F_{13}.

1.8. Число изоморфизмов и пар кручения кривых Эдвардса над простым полем

В некоторых криптографических задачах, использующих изоморфизмы, требуется знать мощность множества изоморфных преобразований или ее оценки. Для кривых в форме Эдвардса над простым полем эта задача решается просто.

При $e = 1$ число различных кривых равно числу квадратичных невычетов d поля $\mu = \frac{p-1}{2}$, тогда общее число различных кривых при всех значениях e^2 равно M = μ^2 . Соответственно, для каждой кривой с фиксированным d при всех e^2 имеется ровно μ изоморфных кривых. Число пар кривых кручения при $p \equiv 1 mod 4$ равно $\mu/2$, а при $p \equiv 3 \mathrm{mod} 4$ $-(\mu - 2)/2$.

В последнем случае число пар уменьшается на 1, так как при $d = -1$ пара вырождается в одну кривую.

Пример 1.2. Расширим пример 1.1 и получим все кривые Эдвардса над полем F_{13}. В таблицу 1.1 сведены параметры a и b для $(p - 1)/2 = 6$ канонических кривых W с двумя точками 4-го порядка, параметры d изоморфных им кривых Эдвардса, параметры t и N_E этих кривых. Середина таблицы является осью симметрии для пар кручения с параметрами $\pm t$ и взаимообратными значениями d.

Таблица 1.1. Параметры и порядки всех кривых в формах Вейерштрасса и Эдвардса при $p = 13$.

a	2	1	2	7	4	7
b	1	2	5	1	3	5
d	2	6	8	5	11	7
t	6	2	2	– 2	– 2	– 6
N_E	8	12	12	16	16	20

Каждая из приведенных в таблице кривых Эдвардса может быть преобразована в $\mu = 6$ изоморфных кривых умножением правой части уравнения (1.10) на различные значения e^2. В итоге имеем $\mu^2 = 36$ кривых, что составляет ¼ всех эллиптических кривых с параметрами $a, b \neq 0$ (их число для поля F_{13} равно 12^2). Такой же результат $\left(\frac{p-1}{2}\right)^2 \frac{1}{(p-1)^2} = \frac{1}{4}$ имеет место при произвольном значении p. Строго говоря, 25% – это нижняя граница оценки доли кривых Эдвардса, изоморфных всем кривым в форме Вейерштрасса, так как из множества последних следует исключить незначительное число несингулярных кривых с дискриминантом $4a^3 + 27b^2 = 0$.

Несмотря на уменьшение вчетверо пространства всех кривых и наличие минимального кофактора 4 у порядка кривой кривые Эдвардса безусловно

являются перспективным направлением эллиптической криптографии. В первую очередь их отличает наивысшая среди известных производительность выполнения групповой операции в проективных координатах [2], универсальность и полнота закона сложения. В существующих стандартах имеются канонические кривые с кофактором 4 (в частности, над полями характеристики 2) [29], поэтому нет оснований сомневаться в целесообразности внедрения технологии кривых Эдвардса в новые стандарты асимметричной криптографии.

1.9. Алгоритм построения кривых Эдвардса над простым полем, изоморфных кривым в форме Вейерштрасса

Результаты, полученные в разделе 1.5, позволяют построить алгоритм вычисления параметров кривых Эдвардса, изоморфных каноническим эллиптическим кривым в форме Вейерштрасса. Решение этой задачи опубликовано в работах [39,46].

Задача состоит в том, чтобы на первом этапе найти параметры приемлемой для криптографических приложений кривой (1.15)) в форме Вейерштрасса, а на втором этапе – рассчитать параметр d изоморфной ей кривой Эдвардса.

Согласно теореме 1.3, условия существования точек 2-го и 4-го порядков и можно выразить через символы Лежандра как

$$\text{(i)}\ \left(\frac{-\left(3c^2+4a\right)}{p}\right)=-1\ ,\ \text{(ii)}\ \left(\frac{\delta}{p}\right)=\left(\frac{3c^2+a}{p}\right)=1\ ,\quad b=-c(a+c^2). \qquad (1.30)$$

Мы полагаем, что все параметры ненулевые: $a\neq 0, c\neq 0 \Rightarrow b\neq 0$. Тем самым мы сразу исключаем некоторые слабые суперсингулярные кривые. Рассмотрим простой пример.

Пример 1.3. Требуется найти кривую с двумя точками 4-го порядка над полем $\boldsymbol{F_7}$. Примем $c=1$ и вычислим аргументы функций (1.30) для всех ненулевых значений a (таблица 1.2). Так как $p\equiv 3 \bmod 4$,

(-1) – квадратичный невычет в поле [29], поэтому $\left(3c^2+4a\right)$ должен быть вычетом, как и второй параметр δ.

Таблица 1.2. Значения параметров *a,c* для проверки выполнения условий (1.30)

a	1	2	3	4	5	6
$\left(3c^2+4a\right)$	0	4	1	5	2	6
$\delta=\left(3c^2+a\right)$	4	5	6	0	1	2

Из таблицы видим, что условия (1.30) для совместных вычетов совпадают лишь при одном значении $a=5$, при этом $b=-c^3-ac=1$. Тогда имеем кривую в форме Вейерштрасса $y^2=x^3+5x+1$ порядка $N_E=12$ (след Фробениуса $t=-4$). Ее единственная точка второго порядка $D=(1,0)$, а координаты точки 4-го порядка в соответствии с (1.20), (1.21) равны:

$$x_1=c\pm\sqrt{\delta}=1\pm1 \quad \Rightarrow x_1=0,$$

$$y_1{}^2=\delta\left(\pm2\sqrt{\delta}+3c\right)=1, \Rightarrow y_1=\pm1.$$

Здесь решения, не лежащие на кривой, отбрасываются.

Для найденной кривой легко построить кривую кручения $y^2=x^3+3x+6$ порядка $N_E=4$ и параметром $t=4$ (см.[29]). Здесь корень кубики смещается $(c=3)$, но свойства (1.30) выполняются и имеются лишь 2 точки 4-го порядка.

Вообще над полем F_7 существует 6 кривых с ненулевыми параметрами a и b и двумя точками 4-го порядка. Их параметры c, a и b вместе с порядками N_E кривых приведены в таблице 1.2. Здесь слева даны параметры трех изоморфных кривых порядка 12 с корнями $c=1,2,4$, а справа – их кривые кручения порядка 4 с корнями $c=3,6,5$ (они, разумеется, также изоморфны).

Таблица 1.3. Параметры кривых W_p и $W_p^{\,t}$ при $p = 7$.

Параметры кривой W_p					Параметры кривой кручения $W_p^{\,t}$			
c	a	b	N_E		c	a	b	N_E
1	5	1	12		3	3	6	4
2	6	1	12		6	5	6	4
4	3	1	12		5	6	6	4

Можно заметить, все $(p-1)$ ненулевых значений корня c могут дать, по крайней мере, $\frac{(p-1)}{2}$ значений параметра a, так как в (1.30) решение определяется квадратом c^2. Поэтому число решений можно вдвое сократить, переходя (при необходимости) к кривой кручения [29].

Будем использовать условия (1.30), в которых обозначим:

1. При $p = 3 mod 4$

$$\begin{cases} -(3c^2 + 4a) = A^2, \\ 3c^2 + a = B^2, \end{cases} \Rightarrow \quad \begin{cases} a = 3^{-1}\left(A^2 - B^2\right), \\ c^2 = 9^{-1}\left(4B^2 - A^2\right) \end{cases} \tag{1.31}$$

2. При $p = 1 mod 4$

$$\begin{cases} \left(3c^2 + 4a\right)h = A^2, \\ 3c^2 + a = B^2, \end{cases} \Rightarrow \quad \begin{cases} a = (3h)^{-1}\left(A^2 - hB^2\right), \quad \left(\frac{h}{p}\right) = -1, \\ c^2 = 9^{-1}\left(4hB^2 - A^2\right). \end{cases} \tag{1.32}$$

Эти равенства записаны на основе (1.30) и свойств элемента поля (– 1), который при $p \equiv 3 mod 4$ является квадратичным невычетом, а при $p \equiv 1 mod 4$ – квадратичным вычетом. В этой связи в равенства (1.32) вписывается произвольный сомножитель h со свойствами квадратичного невычета.

Формулы (1.30) – (1.32) конструктивны, так как позволяют рассчитывать параметры a и $\pm c$ кривой (и, соответственно, $\pm b$) при заданных значениях пар квадратичных вычетов $\left(A^2, B^2\right)$. Объектом поиска является кривая порядка $N_E = 4n$ $n -$ большое простое число. На основе (1.30) – (1.32) можно

предложить следующий алгоритм построения канонических кривых в форме Вейерштрасса с двумя точками 4-го порядка, и далее, изоморфной кривой Эдвардса:

1. В поле F_p задаем произвольное значение пары квадратичных вычетов $\left(A^2, B^2\right)$ и согласно (1.31) или (1.32) рассчитываем параметры a и c^2. Если вычисленное значение c^2 – невычет, меняем параметр B^2 и повторяем расчеты.

2. Если вычисленное c^2 – квадратичный вычет, находим 2 кривые с параметрами $(a, \pm c)$ и $(a, \pm b)$. Значение параметра b рассчитываем в соответствии с (1.30).

3. Находим координаты точки 4-го порядка (для построения изоморфной кривой Эдвардса).

4. Вычисляем порядок одной из кривых и, в случае неприемлемого порядка, рассчитываем порядок кривой кручения. Если приемлемое решение не найдено, переходим к другой паре значений $\left(A^2, B^2\right)$ (возвращаемся в п.1).

5. По формуле (1.23) находим параметр изоморфной кривой Эдвардса.

Разумеется, можно модифицировать данный алгоритм, фиксируя, например, параметр c^2, после чего требовать выполнения условий (1.30). Однако в предложенном виде алгоритм быстрее приводит к кривой с двумя точками 4-го порядка. Далее, как описано в [39], строится изоморфная кривая в форме Эдвардса. В работе [44] использован метод с некоторыми модификациями рассмотренного выше.

ГЛАВА 2

КЛАССИФИКАЦИЯ КРИВЫХ В ОБОБЩЕННОЙ ФОРМЕ ЭДВАРДСА

В настоящей главе мы исследуем новые свойства кривых в форме Эдвардса над простыми конечными полями характеристики $p \neq 2$. Полезными для криптоприложений могут быть как различные классы кривых Эдвардса над простым полем (главы 1 – 4), так и кривые Эдвардса над расширениями малых простых полей (глава 5). Для изучения свойств кривых Эдвардса прежде всего следует их классифицировать. Определения классов кривых Эдвардса в известной работе [3] являются некорректными, так как порождают пересекающиеся классы. Для разделения кривых Эдвардса на непересекающиеся классы, имеющие существенно различные свойства циклических и нециклических групп, мы в данной главе строим новую классификацию кривых в обобщенной форме Эдвардса [54,56]. На ее основе в следующих главах 3 – 5 будет дан анализ свойств циклических и нециклических кривых в форме Эдвардса и предложены оригинальные методы нахождения порядков точек.

В работе [3] авторы расширили класс кривых Эдвардса с модификацией Бернстейна-Ланге [2] введением нового параметра a в уравнение (1.10) кривой и снятием ограничения на неквадратичность ее параметра d. Они назвали этот класс *скрученными кривыми Эдвардса* (*twisted Edwards curves*), а модифицированные Бернстейном и Ланге кривые, определенные в [2] – *полными кривыми Эдвардса* ($a = 1, \left(\frac{d}{p}\right) = -1$). Третий термин – *кривые Эдвардса* – определен авторами для кривых (1.10) без ограничения на свойство неквадратичности ее параметра d. Единственным корректно определенным классом при таких определениях является лишь класс полных кривых Эдвардса. Действительно, класс скрученных кривых в частных случаях включает в себя два других класса, а в класс кривых Эдвардса входит подкласс полных кривых Эдвардса. Это не может не внести путаницу при изучении свойств кривых разных классов. В работе [3] был дан анализ некоторых свойств этих кривых, сделана попытка (на наш взгляд, неудачная) дать

классификацию кривых в форме Эдвардса и привести статистику распределения порядков кривых, относящихся к разным классам этих кривых, при небольших значениях модуля $p = 1009$ и $p = 1019$. Мы обнаружили ряд некорректных утверждений авторов [3] и некорректные результаты в статистике распределения порядков кривых [3, раздел 4] .

Кривые, объединяющие все три класса, мы далее называем кривыми *в обобщенной форме Эдвардса*. По терминологии работы [3] они определены как скрученные кривые Эдвардса. Поскольку, как отмечалось, это порождает три пересекающихся класса кривых, в статистических таблицах раздела 4 [3] одни и те же кривые попадают в разные классы. Это дает недостоверную статистику. Наш анализ показал, что введение нового параметра a в уравнение (1.10) кривой оправдан лишь в случае $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) = -1$, что дает основания оставить термин «скрученные кривые Эдвардса» только для этого случая. Два других класса кривых Эдвардса по новой классификации изоморфны кривым с параметром $a = 1$.

В данной главе мы даем анализ свойств точек порядков 2, 4 и 8 кривых в обобщенной форме Эдвардса, строим новую классификацию этих кривых и получаем точные формулы для числа таких кривых с порядком $4n$. В разделе 2.1 предложена модификация закона сложения точек кривых с заменой обозначения координат $(x \leftrightarrow y)$. В следующем разделе 2.2 вводится арифметика для групповых операций с особыми точками этих кривых, дан анализ точек малых порядков и формулы, связывающие их с другими точками кривой. В разделе 2.3 обсуждается некорректность ряла утверждений, классификации кривых и статистики их порядков в [3], предложена классификация кривых в обобщенной форме Эдвардса с разбиением на три непересекающиеся класса в зависимости от квадратичности параметров a и d кривой. Дан анализ некоторых свойств кривых всех 3-х классов и возможных значений порядков этих кривых. Наконец, в разделе 2.4 мы получили формулы для расчета точного числа кривых различных классов с минимальным кофактором 4 порядка кривой при $p \equiv 1 mod 4$ и $p \equiv 3 mod 4$.

Начиная с этой главы, мы будем пользоваться модифицированным законом сложения точек, предложенным нами в [50].

2.1. Модификация закона сложения точек кривой в обобщенной форме Эдвардса

В работе [3] *скрученные кривые Эдвардса* (*twisted Edwards curves*) определены как обобщение кривых Эдвардса $x^2 + y^2 = (1 + d'x^2y^2)$ [2] введением нового параметра a в уравнение

$$ax^2 + y^2 = (1 + dx^2y^2), \quad a, d \in F_p^*, d \neq 1, \ a \neq d, \ p \neq 2.$$

Наряду с этим авторы [3] сняли ограничения на пару параметров a и d, допуская любые значения $\left(\frac{ad}{p}\right) = \pm 1$. При $a = 1$ такая кривая получила в [3] название *кривой Эдвардса*, а если у нее d – квадратичный невычет (т.е. $\left(\frac{d}{p}\right) = -1$), то – *полной кривой Эдвардса*. Этот термин связан с полнотой закона сложения точек кривой [2].

Если в соответствии с вышеприведенными определениями обозначить класс скрученных кривых Эдвардса как *TEC*, класс кривых Эдвардса как *EC*, а класс полных кривых Эдвардса как *CEC*, то имеет место включение $CEC \subset EC \subset TEC$. Другими словами, полные кривые Эдвардса можно назвать и кривыми Эдвардса, и скрученными кривыми Эдвардса. Полагаем, что для приведенного выше уравнения следует ввести термин «кривые в обобщенной форме Эдвардса».

В работе [50] мы предложили поменять местами *x* и *y* координаты в форме кривой Эдвардса с целью сохранения горизонтальной симметрии обратных точек, принятой в теории эллиптических кривых. Опираясь на это свойство, определим *кривую в обобщенной форме Эдвардса* уравнением

$$E_{a,d}: \ x^2 + ay^2 = (1 + dx^2y^2), \quad a, d \in F_p^*, d \neq 1, \ a \neq d, \ p \neq 2. \qquad (2.1)$$

Этим термином мы далее заменяем термин «скрученная кривая Эдвардса», определенный в [3], а последний термин будет отнесен к частному классу кривых (2.1). Тогда модифицированный универсальный закон сложения точек кривой (2.1) имеет вид

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1x_2 - ay_1y_2}{1 - dx_1x_2y_1y_2}, \frac{x_1y_2 + x_2y_1}{1 + dx_1x_2y_1y_2}\right). \qquad (2.2)$$

При совпадении двух точек получим из (2) закон удвоения точек

$$2(x_1, y_1) = \left(\frac{x_1^2 - ay_1^2}{1 - dx_1^2y_1^2}, \frac{2x_1y_1}{1 + dx_1^2y_1^2}\right). \qquad (2.3)$$

Использование модифицированных законов (2.2), (2.3) позволяет сохранить общепринятую горизонтальную симметрию (относительно оси x) обратных точек. Нейтральный элемент группы здесь равен $O = (1, 0)$. Определяя теперь обратную точку как $-P = -(x_1, y_1) = (x_1, -y_1)$, получим согласно (2.1) $(x_1, y_1) + (x_1, -y_1) = (1, 0) = O$. Кроме нейтрального элемента O на оси x также всегда лежит точка $D_0 = (-1, 0)$ второго порядка, для которой в соответствии с (2.3) $2D_0 = (1, 0) = O$. В зависимости от свойств параметров a и d можно получить еще 2 особые точки второго порядка и 2 или 4 или 6 точек 4-го порядка. Как следует из (2.1), на оси y могут лежать точки $\pm F_0 = (0, \pm 1/\sqrt{a})$ 4-го порядка, для которых $\pm 2F_0 = D_0 = (-1, 0)$. Эти точки существуют над полем $\mathbf{F}_p$, если параметр, a является квадратичным вычетом.

2.2. Свойства точек порядков 2, 4, 8 кривых в обобщенной форме Эдвардса

Кривые Эдвардса, которые мы рассматривали в главе 1, имеют простую циклическую структуру с одной точкой 2-го порядка, двумя точками 4-го порядка, четырьмя точками 8-го порядка и т.д. (при условии их существования). Уравнение (2.1) может породить кривые с более сложной нециклической структурой точек 2-го порядка, содержащие три точки 2-го порядка, 0, 2, 4, 6 или 8 точек 4-го порядка, до 12 точек 8-го порядка. Их анализ усложняется тем, что 2 точки 2-го порядка и возможные 2 точки 4-го порядка являются особыми, т.е. одна из их координат не определена в конечном поле (что возникает при делении на 0).

Из уравнения (2.1) определим квадраты

$$x^2 = \frac{1-ay^2}{1-dy^2}, \qquad y^2 = \frac{1-x^2}{a-dx^2},$$

порождающие при нулевых знаменателях в этих выражениях особые точки на бесконечности (знак "∞" мы далее ставим при делении на 0):

$$D_{1,2} = \left(\pm\sqrt{\frac{a}{d}}, \infty\right), \quad \pm F_1 = \left(\infty, \pm\frac{1}{\sqrt{d}}\right). \tag{2.4}$$

Они возникают в случаях $\left(\frac{ad}{p}\right) = 1$ и $\left(\frac{d}{p}\right) = 1$ соответственно.

Введем формальную арифметику с особыми точками (2.4) кривой (забудем, что на 0 делить нельзя). Так как в наших обозначениях $\infty = \frac{1}{0}$ и $0 = \frac{1}{\infty}$, появление бесконечной координаты в (2.2) или в (2.3) равнозначно умножению числителей и знаменателей на 0 или 0^2. При этом остаются лишь слагаемые, являющиеся сомножителями при знаке ∞. Это отвечает правилам обычного предельного перехода. В частности, с помощью закона удвоения (2.3) легко проверить, что $2D_{1,2} = 0,\ \pm 2F_1 = D_0 = (-1, 0)$. Например, в первом случае

$$2\left(\pm\sqrt{\frac{a}{d}}, \infty\right) = \left(\frac{\frac{a}{d} - a\cdot\infty^2}{1 - \frac{da}{d}\infty^2},\ \frac{\pm 2\sqrt{\frac{a}{d}}\infty}{1 + \frac{da}{d}\infty^2}\right) = \left(\frac{0^2\cdot\frac{a}{d} - a}{0^2 1 - a},\ \frac{0\cdot\left(\pm 2\sqrt{\frac{a}{d}}\right)}{0^2 1 + a}\right) = (1,0).$$

Иными словами, при выполнении условий их существования особые точки $D_{1,2}$ есть точки 2-го порядка, а особые точки $\pm 2F_1$ – точки 4-го порядка.

Кроме перечисленных, точки 4-го порядка могут существовать как не особые при ненулевых координатах *x* и *y*.

Дадим анализ некоторых новых свойств точек 4-го и 8-го порядков [54,56].

Теорема 2.1. *Неособые точки 4-го порядка*

$$\pm F_2 = \left(\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right),\ \pm F_3 = \left(-\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right)$$

кривой в форме (2.1) *при* $x \neq 0$ *существуют тогда и только тогда, когда выполняются условия*:

(i) при $p \equiv 3 mod 4$: $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) - 1$;

(ii) при $p \equiv 1 mod 4$: $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) = 1,\ \frac{a}{d} = c^4$.

Доказательство. *Необходимость.* Особые точки $\pm F_1 = \left(\infty, \pm\frac{1}{\sqrt{d}}\right)$ из формул (2.4), возникающие при $\left(\frac{d}{p}\right) = 1$, исключаются из рассмотрения в соответствии с формулировкой теоремы. Не рассматриваются также точки $\pm F_0 = (0, \pm 1/\sqrt{a})$ при $x = 0$. Положим $2F_2 = 2(x_1, y_1) = D_1$. Тогда согласно (2.3) и (2.4) запишем два уравнения:

$$\frac{x_1^2 - a y_1^2}{(1 - d x_1^2 y_1^2)} = \sqrt{\frac{a}{d}}, \qquad \frac{2x_1 y_1}{(1 + d x_1^2 y_1^2)} = \infty.$$

Отсюда $(1 + dx_1^2 y_1^2) = 0, \Rightarrow\ x_1^2 + ay_1^2 = 0,\ \Rightarrow\ x_1^2 = -ay_1^2$. Из $x_1 \neq 0$ следует $y_1 \neq 0$. Здесь второе равенство записано на основании уравнения (2.1) кривой. Согласно первому из уравнений и равенства $x_1^2 = -ay_1^2$ имеем

$$\frac{2x_1^2}{1+\frac{d}{a}x_1^4} = \sqrt{\frac{a}{d}} \quad \Rightarrow\ dx_1^4 - 2\sqrt{ad}\,x_1^2 + a = 0 \quad \Rightarrow\ x_1^2 = \sqrt{\frac{a}{d}},\ y_1^2 = -\frac{1}{\sqrt{ad}}.$$

Итак, получаем 4 точки с координатами:

$$\pm F_2 = \left(\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right),\ \pm F_3 = \left(-\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right). \qquad (2.5)$$

При $p \equiv 3 mod 4$ элемент (-1) есть квадратичный невычет [29], тогда $(-a)$ – квадратичный вычет в условиях (i) и равенство $x_1^2 = -ay_1^2$ связывает квадраты координат точки F_2. Пусть β – примитивный элемент мультипликативной группы $\mathbf{F}_p^*$, и β^2 – квадрат этой группы, тогда при условии (i) имеем $\beta^2 = \beta^2\beta^{p-1} = \beta^{2+4k+2} = \beta^{4(k+1)}$. Значит, любой квадрат имеет квадратные корни и корни 4-й степени при $p \equiv 3 mod 4$. Необходимость существования первых координат в (2.5) с учетом условий (i) доказана. Учитывая условия (i) и принимая значение $\left(\frac{-\sqrt{ad}}{p}\right) = 1$ (т.е. как квадратичного вычета, при этом $\sqrt{ad}$ – квадратичный невычет), получаем по 2 решения для вторых координат в точках (2.5). Так как квадраты *ad* и *a/d* имеют корни 4-й степени, такие точки в условиях теоремы существуют. Необходимость условий (i) теоремы доказана.

При $p \equiv 1 mod 4$ элемент (-1) есть квадратичный вычет [29], тогда равенство $x_1^2 = -ay_1^2$ выполняется при $\left(\frac{a}{p}\right) = 1$. Для квадрата мультипликативной группы имеем $\beta^2 = \beta^2\beta^{p-1} = \beta^{2+4k} = \beta^{2(2k+1)}$ Для этого случая при $\beta = c^2$ число элементов c^4 при всех ненулевых значениях *c* равно $(p-1)/4$. Обе координаты точек (2.5) существуют, если $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) = 1$, и $\frac{a}{d} = c^4$. Тогда и для второй координаты справедливо $\frac{1}{ad} = \frac{c^4}{a^2} = e^4$. Итак, необходимость условий (ii) теоремы доказана.

Достаточность. Пусть выполняются условия (i) или (ii). Тогда существуют 4 точки $\pm F_{2,3} = \left(\pm\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right)$, для которых согласно (2.3)

получим $\pm 2F_{2,3} = D_{1,2}$. Так как удвоение точек $F_{2,3}$ 4-го порядка дает точки 2-го порядка, то определенные координатами (2.5) точки есть точки 4-го порядка. Это доказывает достаточность условий теоремы. ▲

Точки $\pm F_{2,3}$ можно рассматривать как точки деления на два точек 2-го порядка $D_{1,2}/2$ [57, 58].

Пример 2.1. Для кривой $x^2 - y^2 = (1 + 3x^2y^2) mod 7$ (здесь $a = -1, d = = 3$ – квадратичные невычеты при $p = 7 \equiv 3 mod 4$ и выполняются условия (i) теоремы 2.1) точки 4-го порядка имеют координаты $F_{2,3} = (\pm 2, \pm 2)$. При удвоении их согласно (2.3) получим $2F_2 = \left(\sqrt{\frac{a}{d}},\ \infty\right) = (\pm 3, \infty) = D_{1,2}$. Порядок N_E этой кривой, включающей точки $O, D_{0,1,2}, \pm F_{2,3}$, равен 8, группа точек нециклическая с типом $T = (2, 2^2)$.

Пример 2.2. В условиях (ii) теоремы 2.1 рассмотрим кривую $x^2 + y^2 = (1 + 3x^2y^2) mod 13$ (здесь $a = 1,\ d = 3$ – квадратичные вычеты при $p = 13 \equiv 1 mod 4$). Согласно (2.5) находим точки 4-го порядка $F_{2,3} = (\pm 6, \pm 4)$. Кроме того, согласно (2.4) кривая имеет две особые точки 4-го порядка $\pm F_1 = \left(\infty, \pm \frac{1}{\sqrt{d}}\right) = (\infty, \pm 3)$. Подстановка значений координат точек $F_{2,3}$ в уравнение кривой дает $6^2+4^2=1+3\cdot 6^2\cdot 4^2 = 0$. Удвоение точек $F_{2,3}$ согласно (2.3) дает точки $2F_2 = \left(\pm\sqrt{\frac{a}{d}},\ \infty\right) = (\pm 3, \infty) = D_{1,2}$. Эта кривая имеет порядок $N_E = 16$, является нециклической с типом $T = (2^2, 2^2)$.

Утверждение 2.1. *Все кривые Едвардса* (2.1) с условиями $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) - 1$ *при* $p \equiv 1 mod 4$ *имеют порядок* $N_E = 4n$ ($n -$ *нечетное*).

Доказательство. В условиях $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) - 1$ теоремы 2.1 при $p \equiv 1 mod 4$ кривая не содержит точек 4-го порядка, но включает нециклическую подгруппу 4-го порядка точек 2-го порядка $G_4 = \{O, D_0, D_1, D_2\}$. Следовательно, порядки всех других точек могут быть равными n и $2n$ (вместе с возможными нечетными сомножителями n). Итак, подгруппа G_4 есть подгруппа

минимального четного порядка 4 кривой, и порядок кривой $N_E = 4n$. Утверждение доказано. ▲

В следующем разделе 2.3 кривые с условиями (i) теоремы 2.1 мы отнесем к классу скрученных кривых Едвардса. Из утверждения 2.1 следует важный результат: все скрученные кривые Едвардса при $p \equiv 1 mod 4$ имеют порядок $N_E = 4n$.

Найдем условия существования точек 8-го порядка, порожденных делением на 2 точки F_0.

Теорема 2.2 *Необходимыми и достаточными условиями существования точек 8-го порядка кривой* (2.1), *порожденных делением на* 2 *точки* F_0, *являются:*

(i) при $\left(\frac{ad}{p}\right) = -1$: $\left(\frac{a}{p}\right) = 1$, $\left(\frac{1-\frac{d}{a}}{p}\right) = 1$;

(ii) при $\left(\frac{ad}{p}\right) = 1$: $\left(\frac{a}{p}\right) = 1$, $\left(\frac{1-\frac{d}{a}}{p}\right) = 1$ и $\left(\frac{1+\sqrt{1-\frac{d}{a}}}{p}\right) = 1$.

Доказательство. *Необходимость.* Пусть $S = (x_1, y_1)$ – точка 8-го порядка, тогда $2S = F_0 = (0, 1/\sqrt{a})$ – одна из точек 4-го порядка на оси y. Согласно (2.3) и координат $\boldsymbol{F_0}$ имеем

$$\frac{x_1^2 - a y_1^2}{(1 - d x_1^2 y_1^2)} = 0\,, \qquad \frac{2 x_1 y_1}{(1 + d x_1^2 y_1^2)} = \frac{1}{\sqrt{a}}\,. \tag{2.6}$$

Тогда $x_1^2 = a y_1^2$, $\Rightarrow$ $\frac{d}{a} x_1^4 - 2 x_1^2 + 1 = 0$ $\Rightarrow$ $x_{1,2}^2 = \frac{a}{d}\left(1 \pm \sqrt{1 - \frac{d}{a}}\right)$.

Координаты точек S_k, k = 1..4, или k = 1..8 определяются из

$$\boldsymbol{S}_k = \left(\pm \left(\frac{a}{d} \left(1 \pm \sqrt{1 - \frac{d}{a}} \right) \right)^{1/2}, \pm \left(\frac{1}{d} \left(1 \pm \sqrt{1 - \frac{d}{a}} \right) \right)^{1/2} \right). \tag{2.7}$$

Так как справедливо

$$\left(1 + \sqrt{1 - \frac{d}{a}}\right)\left(1 - \sqrt{1 - \frac{d}{a}}\right) = \frac{d}{a}, \tag{2.8}$$

то при $\left(\frac{ad}{p}\right) = -1$ и $\left(\frac{1-\frac{d}{a}}{p}\right) = 1$ либо $\left(1 + \sqrt{1 - \frac{d}{a}}\right)$ является квадратом, либо $\left(1 - \sqrt{1 - \frac{d}{a}}\right)$. Умножая квадратичный невычет из этой альтернативы

на невычет $\frac{a}{d}$, получим значение x_1^2 координаты одной из точек S_k. Извлекая из квадрата x_1^2 два корня, определяем значения координат $\pm x_1$ в (2.7). Необходимыми условиями существования $\pm x_1$ являются условия (i) теоремы. Учитывая условие $\left(\frac{a}{p}\right) = 1$ и разделив эти значения на $\sqrt{a}$, получим координаты $\pm y_1$ точек 8-го порядка. Число точек 8-го порядка для данного случая равно 4. Первое из необходимых условий теоремы (i) доказано.

При $\left(\frac{ad}{p}\right) = 1$ (условия (ii) теоремы) оба значения в скобках (2.8) есть квадратичные вычеты или невычеты. Так как сомножитель $\frac{a}{d}$ квадрата x_1^2 является квадратом, то вместе с условием $\left(\frac{1-\frac{d}{a}}{p}\right) = 1$ должно выполняться $\left(\frac{1+\sqrt{1-\frac{d}{a}}}{p}\right) = 1$, (и, соответственно, $\left(\frac{1-\sqrt{1-\frac{d}{a}}}{p}\right) = 1$). Необходимость условий (i) и (ii) доказана.

Достаточность. Пусть выполняются условия (i). Тогда, выбирая квадратичный невычет из 2-х значений $\left(1 \pm \sqrt{1-\frac{d}{a}}\right)$, определяем координаты 4-х точек из (2.7). Для них $2S = F_0 = (0,1/\sqrt{a})$, т.е. в этом случае 4 точки 8-го порядка существуют.

Пусть выполняются условия (ii). Так как оба значения $\left(1 \pm \sqrt{1-\frac{d}{a}}\right)$ в этом случае являются либо квадратичными вычетами, либо невычетами, то с учетом $\left(\frac{a}{p}\right) = 1$ получаем обе координаты 8-ми точек 8-го порядка (2.7). Увеличение вдвое числа точек связано с нециклической структурой точек четного порядка для этого случая. Итак, 8 точек 8-го порядка в условиях (ii) теоремы существуют. Теорема доказана. ▲

Теорема 2.2 не исчерпывает всех возможных точек 8-го порядка, так как при $\left(\frac{ad}{p}\right) = 1$ могут возникать особые точки 4-го порядка (2.4) и неособые точки 4-го порядка (2.5), для которых деление на 2 может также породить точки 8-го порядка.

В приведенном выше примере 2.1 кривой с $a = -1$ и $d = 3$ при $p = 7$ оба параметра – квадратичные невычеты и нарушается одно из условий (ii) теоремы 2.2 $\left(\frac{a}{p}\right) = 1$. Хотя порядок кривой равен 8, точек 8-го порядка она не содержит. Для кривой $x^2 + y^2 = (1 + 3x^2y^2) mod 13$ примера 2.2 значение $1 - \frac{d}{a} = 11 -$ квадратичный невычет, условия (ii) теоремы 2.2 не выполняются и точки 8-го порядка (2.7) не существуют.

При условии существования особых точек (2.4) вместе с точками $D_0 = (-1, 0)$, $\pm F_0 = (0, \pm 1/\sqrt{a})$, принимая правила предельного перехода в (2.2), можно найти координаты сумм:

$$(x_1, y_1) + (-1,\ 0) = (-x_1, -y_1),$$

$$(x_1, y_1) + \left(\sqrt{\frac{a}{d}}, \infty\right) = \left(\sqrt{\frac{a}{d}} \cdot x_1^{-1},\ \frac{1}{\sqrt{ad}} \cdot y_1^{-1}\right),$$

$$(x_1, y_1) + \left(-\sqrt{\frac{a}{d}}, \infty\right) = \left(-\sqrt{\frac{a}{d}} \cdot x_1^{-1}, -\frac{1}{\sqrt{ad}} \cdot y_1^{-1}\right),$$

$$(x_1, y_1) + \left(\infty, \frac{1}{\sqrt{d}}\right) = \left(-\frac{1}{\sqrt{d}} \cdot y_1^{-1},\ \frac{1}{\sqrt{d}} \cdot x_1^{-1}\right),$$

$$(x_1, y_1) + \left(\infty, -\frac{1}{\sqrt{d}}\right) = \left(\frac{1}{\sqrt{d}} \cdot y_1^{-1}, -\frac{1}{\sqrt{d}} \cdot x_1^{-1}\right).$$

Все найденные суммы удовлетворяют уравнению (2.1) при подстановке, т.е. являются точками кривой.

Подчеркнем, что использование правил предельного перехода сохраняет операцию сложения любых пар точек, включая особые. Это позволяет говорить об изоморфизме кривых в форме Монтгомери и Эдвардса, который рассматривается в следующем разделе.

2.3. Новая классификация кривых в обобщенной форме Эдвардса

Как уже отмечалось, в работе [3] впервые введено понятие скрученной кривой Эдвардса. Нам представляется, в ней имеются некорректные определения, утверждения и результаты, которые мы ниже обсуждаем [54,56].

Основные теоремы в работе [3] опираются на бирациональную эквивалентность между кривыми (2.1) и кривыми в форме Монтгомери

$$M_{A,B}\colon Bv^2 = u^3 + Au^2 + u,\ A = 2\frac{a+d}{a-d},\ B = \frac{4}{a-d},\ a = \frac{A+2}{B},\ d = \frac{A-2}{B},$$
$$A^2 \neq 4. \qquad (2.9)$$

Она основана на замене координат с помощью рациональных функций

$$y = \frac{u}{v},\ x = \frac{u-1}{u+1} \Rightarrow u = \frac{1+x}{1-x},\quad v = \frac{u}{y}. \qquad (2.10)$$

В работе [3] доказана теорема 3.2: *любая скрученная кривая Эдвардса* (2.1) *бирационально эквивалентна кривой* (2.9) *в форме Монтгомери*.

Так как нам придется обращаться к паре квадратичного кручения (*quadratic twist* [3]), мы также проведем отображение точек (2.9) в точки кривой (2.1).

Разделим уравнение (2.9) на v^2 и с учетом (2.10) получим

$$\frac{4}{(a-d)}\cdot\frac{1}{y^2} = u + u^{-1} + 2\,\frac{a+d}{a-d},\quad \Rightarrow \qquad \frac{2}{(a-d)}\cdot\frac{1}{y^2} = \frac{1+x^2}{1-x^2} + \frac{a+d}{a-d}.$$

Отсюда

$$\frac{2(1-x^2)}{y^2} = (1+x^2)(a-d) + (1-x^2)(a+d),$$

и, наконец, получим изоморфную кривой (2.9) кривую в форме (2.1)

$$M_{A,B} \sim E_{a,d} \quad : \qquad (1-x^2) = y^2(a-dx^2).$$

Нетрудно с помощью (2.10) осуществить и обратное преобразование. Имеет место взаимно однозначное отображение точек $(u_1, v_1) \leftrightarrow (x_1, y_1)$. Если для любой пары точек принять операцию сложения (2.2) с включением особых точек (см 2.2), то можно утверждать, что кривые (2.1) и (2.9) изоморфны: $M_{A,B} \sim E_{a,d}$.

Перейдем к парам квадратичного кручения. Пусть $\left(\frac{c}{p}\right) = -1$, тогда кривая кручения для кривой (2.9) в форме Монтгомери имеет вид

$$M^t_{cB,A}:\quad cBv^2 = u^3 + Au^2 + u,\ A = 2\frac{a+d}{a-d},\ B = \frac{4}{a-d}.$$

Изоморфная ей кривая в обобщенной форме Эдвардса (2.1), как можно видеть из выполненных выше преобразований, записывается как

$$E^t_{ca,cd} \sim M^t_{cB,A}: (1-x^2) = cy^2(a-dx^2) = y^2(ca-cdx^2),\ \left(\frac{c}{p}\right) = -1.$$

Иначе говоря, для построения пары квадратичного кручения к кривой в форме (2.1) следует перейти к новым параметрам кривой (2.1) в форме Эдвардса $a' = ca, d' = cd$, при этом квадратичные вычеты обращаются в невычеты и обратно, а квадратичное кручение кривой (2.1) определяется как

$$E_{a,d}^t \sim E_{ca,cd}, \ \left(\frac{c}{p}\right) = -1. \qquad (2.11)$$

Во 2-м разделе работы [3] утверждается, что кривая $E_{1,d/a}$ есть пара квадратичного кручения (*quadratic twist*) кривой $E_{a,d}$, т.е. $E_{a,d}^t \sim E_{1,d/a}$. Видимо, следует признать это утверждение в общем случае некорректным. Как следует из нашего анализа, оно справедливо лишь при $\left(\frac{a}{p}\right) = -1$, если принять $c = a^{-1}$. При $\left(\frac{a}{p}\right) = 1$ кривые $E_{a,d}$ и $E_{1,d/a}$ изоморфны: $E_{a,d} \sim E_{1,d/a}$ Здесь же авторы [3] заключают, что кривая $E_{1,d}$ есть квадратичное кручение кривой $E_{1,1/d}$, ссылаясь на известный факт из работы [2]. Но в [2] это справедливо в условиях $\left(\frac{d}{p}\right) = -1$, тогда как в [3] допускается $\left(\frac{d}{p}\right) = 1$, и тогда эта пара кривых изоморфна: $E_{1,d} \sim E_{1,1/d}$ Действительно, заменив $d \to d^{-1}$ в уравнении (2.1) или (2,9) при $a = 1$, получим изоморфную кривую при условии $\left(\frac{d}{p}\right) = 1$.

Чтобы классифицировать кривые в обобщенной форме Эдвардса с разбиением на непересекающиеся классы, рассмотрим все 4 сочетания для пар параметров a и d кривой (2.1). Мы задаем их условиями $C1$ и $C2$ и разбиваем на пары.

$C1$: $\left(\frac{ad}{p}\right) = -1$.

$C1.1$: $\left(\frac{a}{p}\right) = 1, \ \left(\frac{d}{p}\right) = -1$. Согласно (2.1) и (2.2) в этом случае на кривой (2.1) имеется единственная точка $D_0 = (-1, 0)$ 2-го порядка и 2 точки 4-го порядка $\pm F_0 = (0, \pm 1/\sqrt{a})$. В соответствии с (2.10) им отвечают точки кривой Монтгомери (2.9) $D_{M0} = (0,0), \ \pm F_{M0} = (1, \pm\sqrt{a})$. Этот случай определен в работе [2]. Здесь заменой $(x, y) \to (X, Y/\sqrt{a})$ получаем изоморфную кривой (2.1) полную кривую Эдвардса $X^2 + Y^2 = 1 + d'X^2Y^2$, $d' = \frac{d}{a}, \ \left(\frac{d'}{p}\right) = -1$. Итак, для этого случая имеет место изоморфизм $E_{a,d} \sim E_{1,d/a}$.

*C*1.2: $\left(\frac{a}{p}\right) = -1$, $\left(\frac{d}{p}\right) = 1$. Здесь параметры a и d просто меняются местами. С помощью замены $(x, y) \to (1/X, Y)$ получим изоморфную (2.1) кривую $X^2 + dY^2 = 1 + aX^2Y^2$. Ее квадратичное кручение образуется заменой $d' = cd$, $a' = ca$, $\left(\frac{c}{p}\right) = -1$, при этом $E^t_{d,a} \sim E_{c\,d,\,c\,a}$ и мы попадаем в условия *C*1.1. Согласно *C*1.1 справедливо $E_{d,a} \sim E_{1,a/d}$. Таким образом, пара кривых $E_{d,a} \sim E^t_{a,d}$, отвечающих условиям *C*1.1 и *C*1.2, образуют пару квадратичного кручения. Этот вывод обобщает известный из [2] результат: $E^t_{1,d} \sim E_{1,1/d}$.

Итак, рассмотренные в *C*1 условия для a и d порождают класс изоморфизмов *полных кривых Эдвардса*, и каждая кривая в условиях *C*1.1 заменой $d \to d^{-1}$ отображается в кривую квадратичного кручения *C*1.2 и обратно.

*C*2: $\left(\frac{ad}{p}\right) = 1$.

*C*2.1: $\left(\frac{a}{p}\right) = -1$, $\left(\frac{d}{p}\right) = -1$. Согласно (2.9) имеем $(Bad)^2 = (A + 2)(A - 2)$ и, следовательно, дискриминант квадратного уравнения в правой части (2.9) ($A^2 - 4$) является квадратом. Тогда уравнение $u^3 + Au^2 + u = 0$ имеет 3 корня в поле $\mathbf{F}_p$: $\{0, -(A \pm \sqrt{A^2 - 4})/2\}$, а кривая Монтгомери содержит 3 точки 2-го порядка: $D_{M0} = (0,0)$, $D_{M1,2} = ((\mathrm{A} \pm \sqrt{A^2 - 4})/2, 0)$. Преобразованием координат (2.10) точка D_{M0} кривой (2.9) переходит в точку $D_0 = (-1, 0)$ кривой (2.1), а две другие точки $D_{\mathrm{M1,2}}$ отображаются в 2 точки 2-го порядка $D_{1,2} = \left(\pm\sqrt{\frac{a}{d}}, \infty\right)$ с делением на 0 y-координаты $y = u/v$. Точки 4-го порядка $\pm F_0 = \left(0, \pm\frac{1}{\sqrt{a}}\right)$ на оси y и особые точки $\pm F_1 = \left(\infty, \pm\frac{1}{\sqrt{d}}\right)$ для этого случая не существуют. Согласно теореме 2.1, кривая (2.1) имеет точки 4-го порядка (2.5) лишь при $p \equiv 3 \bmod 4$. На основе замены $(x, y) \to (1/X, Y)$ и умножения на X^2 получаем изоморфизм $E_{a,d} \sim E_{d,a}$.

*C*2.2: $\left(\frac{a}{p}\right) = 1$, $\left(\frac{d}{p}\right) = 1$. Как и в предыдущем случае, имеются 3 точки 2-го порядка с теми же координатами, что и в *C*2.1. Кроме того, имеются точки 4-го порядка $\pm F_0 = \left(0, \pm\frac{1}{\sqrt{a}}\right)$ на оси y кривой (2.1). Вместе с тем возникают особые точки 4-го порядка (2.4) $\pm F_1 = \left(\infty, \pm\frac{1}{\sqrt{d}}\right)$. При $p \equiv 1 \bmod 4$ и $ad = c^4$

согласно теореме 2.1 есть также 4 точки $\pm F_{2,3}$ (2.5). Всего, как видим, всегда имеется 4 или 8 точек 4-го порядка. Для данного случая преобразование координат $(x, y) \to (X/\sqrt{a}, Y)$ дает изоморфную кривой (2.1) кривую $X^2 + Y^2 = 1 + d'X^2Y^2$, где $d'= d/a$ и имеет место изоморфизм $E_{a,d} \sim E_{1,d/a}$.

Из (2.11) очевидно, что кривые с условиями *C*2.1. и *C*2.2. образуют пару квадратичного кручения, т.е. $E^t_{a,d} \sim E_{ca,cd}$, $\left(\frac{c}{p}\right) = -1$.

Проведенный выше анализ свойств кривых в обобщенной форме Эдвардса (2.1) в условиях *C*1 и *C*2 приводит к выводам:

1. Кривые с условиями *C*1 являются циклическими полными кривыми Эдвардса. Замена условий *C*1.1 на *C*1.2 (или замена $a \leftrightarrow d$) порождает пару квадратичного кручения. Эти кривые изоморфны кривым с параметром $a = 1$.
2. Кривые с условиями C2 являются нециклическими кривыми Эдвардса, включающими 3 точки 2-го порядка, из которых две точки являются особыми. Замена условий *C*2.1 на *C*2.2 (или замена $a \to ca,\ d \to cd$) также порождает пару квадратичного кручения. Замена $a \leftrightarrow d$ внутри этих условий порождает изоморфную кривую : $E_{a,d} \sim E_{d,a}$.
3. Кривые с условиями *C*2.1 не содержат точек 4-го порядка при $p \equiv 1 mod 4$.
4. Кривые с условиями *C*2.2 всегда содержат 4 или 8 точек 4-го порядка, среди которых 2 точки – особые.
5. Введение нового параметра a в обобщенную форму кривых Эдвардса оправдывается лишь исключительно в условиях *C*2.1. Все другие условия приводят к изоморфным кривым с параметром $a = 1$.

Последний вывод дает основания для того, чтобы термин «скрученные кривые Эдвардса» использовать лишь для кривых с условиями *C*2.1. Вместе с тем остальные кривые, изоморфные кривым с параметром $a = 1$, следует разбить на 2 класса: полные кривые Эдвардса (с параметром $\left(\frac{d}{p}\right) = -1$) и квадратичные кривые Эдвардса (с параметром $\left(\frac{d}{p}\right) = 1$). Лишь последний термин вместе с термином «кривые в обобщенной форме Эдвардса» являются новыми по сравнению с работой [3], но наши определения классов принципиально отличаются от приведенных в этой работе.

Итак, мы разбиваем все кривые в обобщенной форме Эдвардса (2.1) на 3 непересекающиеся класса:

- *полные кривые Эдвардса* (с условиями *C*1: $\left(\frac{ad}{p}\right) = -1$);
- *скрученные кривые Эдвардса* (с условиями *C*2.1: $\left(\frac{a}{p}\right) = -1, \left(\frac{d}{p}\right) = -1$);
- *квадратичные кривые Эдвардса* (с условиями *C*2.2: $\left(\frac{a}{p}\right) = 1, \left(\frac{d}{p}\right) = 1.$).

Основные свойства этих трех классов кривых приведены в таблице 2.1.

Таблица 2.1. Классификация и свойства кривых в обобщенной форме Эдвардса

Усло-вия	Свой-ства	Класс кривых Эдвардса	Структу-ра	Точки 2-го порядка	Точки 4-го порядка	Изоморфизм	Класс изомор-физмов
C1	$\left(\frac{ad}{p}\right) = -1$						
C1.1	$\left(\frac{a}{p}\right) = 1,$ $\left(\frac{d}{p}\right) = -1$	*полные*	*циклич.*	$\boldsymbol{D_0}=(-1,0)$	$\pm\boldsymbol{F_0}=(0, \pm 1/\sqrt{a})$	$X^2+Y^2=1+d'X^2Y^2,$ $d'=d/a \Rightarrow \left(\frac{d'}{p}\right)= -1$ $(x,y)\rightarrow(X,Y/\sqrt{a})$	$E_{a,d}$~$E_{1,d/a}$.
C1.2	$\left(\frac{a}{p}\right) = -1,$ $\left(\frac{d}{p}\right) = 1$	*полные*	*циклич.*	$\boldsymbol{D_0}=(-1,0)$	$\pm\boldsymbol{F_0}=(0, \pm 1/\sqrt{d})$	$X^2+dY^2=1+aX^2Y^2$ $d'=cd, a'=ca, \left(\frac{c}{p}\right) =$ $= -1,$ $(x,y)\rightarrow(1/X, Y/\sqrt{d})$ $X^2+Y^2=1+\frac{a'}{d}X^2Y^2$	$E^t_{1,d/a}$~$E_{1,a/d}$
C2	$\left(\frac{ad}{p}\right) = 1$	новые классы					

C2.1	$\left(\frac{a}{p}\right) = -1,$ $\left(\frac{d}{p}\right) = -1$	*Скручен-ные*	*нециклич.*	$\boldsymbol{D_0}$=(−1,0) $\boldsymbol{D_{1,2}}$ = $\left(\pm\sqrt{\frac{a}{d}}, \infty\right)$	$\pm\boldsymbol{F_{2,3}}$ $= \left(\pm\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right)$ (p ≡3mod4)	$(x,y) \rightarrow (1/X, Y)$	$E_{a,d}$~$E_{d,a}$
C2.2	$\left(\frac{a}{p}\right) = 1,$ $\left(\frac{d}{p}\right) = 1$	*Квадратич-ные*	*нециклич.*	$\boldsymbol{D_0}$=(−1,0) $\boldsymbol{D_{1,2}}$ = $\left(\pm\sqrt{\frac{a}{d}}, \infty\right)$	$\pm\boldsymbol{F_0}$=(0, $\pm 1/\sqrt{a}$), $\pm\boldsymbol{F_1}$= $\left(\infty, \pm\frac{1}{\sqrt{d}}\right)$, $\pm\boldsymbol{F_{2,3}}$	$X^2+Y^2=1+d'X^2Y^2$, $d'=d/a \Rightarrow \left(\frac{d'}{p}\right)= 1$ $(x,y) \rightarrow (X, Y/\sqrt{a})$	$E_{a,d}$~$E_{d,a}$ $E_{a,d}$~$E_{1,d/a}$ $E_{a,d}$~$E_{1,a/d}$

В разделе 1.1 мы показали, что кривые в оригинальной форме Эдвардса изоморфны кривым $x^2 + y^2 = 1 + e^4x^2y^2$, и, таким образом, при $e^4 = d \neq 1$ относятся к классу квадратичных кривых Эдвардса.

Обратимся к классификации кривых в форме Эдвардса, данной в работе [3]. В статистические таблицы порядков кривых в [3, раздел 4] вошли классы кривых Эдвардса (Edwards curve), полных кривых Эдвардса (complete Edwards curve), скрученных кривых Эдвардса (twisted Edwards curve). Как отмечено в разделе 2.1, кривые Эдвардса в работе [3] определены как $E_{1,d}$ без ограничений на квадратичность параметра d. По нашей классификации это объединяет два непересекающихся класса: полных и квадратичных кривых Эдвардса. Скрученные кривые Эдвардса в [3] определены как кривые в обобщенной форме Эдвардса (2.1) (по нашей терминологии) и, следовательно, включают все три класса, определенные нами. Непонятно, как с такой классификацией в [3] можно строить таблицы распределения числа кривых, входящих в разные классы. Одни и те же кривые при этом будут учитываться дважды или трижды.

В этом свете нельзя признать корректной статистику порядков кривых, приведенной в [3, раздел 4]. Это, разумеется, связано с пересечением классов кривых, определенных в этой работе. Неясно, откуда при p≡3mod4 в таблице порядков кривых (p=1019) возникает, например, 236 скрученных кривых Эдвардса с кофактором 4. Этот результат противоречит теореме 3.4 работы [3] и нашей теореме 2.1. Значит, они заимствованы из кривых, изоморфных полным кривым Эдвардса, т.е. одни и те же кривые регистрируются в разных

классах. Определения классов кривых Эдвардса, принятые в [3], в принципе не могут дать достоверной статистики.

Поэтому мы предлагаем новую классификацию кривых в обобщенной форме Эдвардса (2.1) с разбиением их на 3 непересекающихся класса. Далее мы опираемся на нашу классификацию [53].

В работе [3] доказаны теоремы 3.3 – 3.5. о бирациональной эквивалентности кривых в форме Эдвардса и Монтгомери. В теореме 3.3 [3] доказано, что кривые Эдвардса $E_{1,d}$ и Монтгомери $M_{A,B}$ бирационально эквивалентны лишь при наличии в них точек 4-го порядка. Далее в теореме 3.4 доказана бирациональная эквивалентность этих кривых и наличие в них точек 4-го порядка при $p \equiv 3mod4$. В частности, для скрученных кривых Эдвардса (с условиями *C*2.1) порядок кривой $N_E \equiv 0\mathrm{mod}8$. Действительно, для нее парой квадратичного кручения является кривая с условиями *C*2.2, имеющая подгруппы 8-го порядка. Следовательно, ее порядок $N_E \equiv 0\mathrm{mod}8$. Тогда сумма числа точек пары кривых кручения при $p \equiv 3mod4$ равна $N_E + N_E^t = 2(p+1) = 2(4k+3+1)$. Отсюда следует $N_E^t = 0mod8$.

При $p \equiv 1mod4$ аналогично получим $N_E + N_E^t = 2(p+1) = 2(4k+1+$ $+1)$, тогда $N_E + N_E^t = 0mod4$, и при $N_E = 0mod8$ для квадратичной кривой Эдвардса порядок скрученной кривой Эдвардса $N_E^t = 0mod4$. Ясно, что в этом случае скрученная кривая имеет 3 точки 2-го порядка и не имеет точек 4-го порядка. Это же утверждает условие (i) теоремы 2.1. При этом нет изоморфизма скрученной кривой Эдвардса с кривой $E_{1,d/a}$, имеющей точки 4-го порядка (теорема 3.5 [3]). Итак, скрученные кривые Эдвардса с минимальным кофактором 4 порядка $N_E = 4n$ существуют лишь для половины возможных значений модуля $p \equiv 1mod4$.

Пример 2.3. Рассмотрим кривую Монтгомери (2.9) $v^2 = u^3 + 9u^2 + u$, p = 17 [3], для которой $A = 9, B = 1, a = 11, d = 7$. Здесь параметры a и d являются квадратичными невычетами по модулю 17, что отвечает условиям *C*2.1. Согласно (2.1) получаем уравнение изоморфной скрученной кривой Эдвардса $x^2 + ay^2 = (1 + dx^2y^2)$. Кривая Монтгомери $v^2 = u(u+3)(u+$ $+6)$ имеет порядок $N_E = 20$, содержит 3 точки 2-го порядка $(0,0), (0,14)$ и $(0,11)$ и не имеет точек 4-го порядка. Она является нециклической с типом

T = (2,2,5) и представляется прямой суммой циклических подгрупп 2-го и 10-го порядков. Ясно, что она содержит 3 различные подгруппы точек 10 порядка (всего имеется 4 точки 5-го и 12 точек 10-го порядков). Если уравнение изоморфной кривой Эдвардса $E_{11,7}$ записать как $x^2 = \frac{(y^2-1)}{(7y^2-11)}$, то и числитель, и знаменатель здесь обращаются в 0 при соответственно $y^2 = 1$ и $y^2 = 4$. В первом случае вместе с точкой O получаем обязательную точку 2-го порядка $D_0 = (-1, 0)$. При $y^2 = 4$ возникают особые точки 2-го порядка $D_{1,2} = (\pm 2, \infty)$. Согласно (2.10) $x = (u-1)/(u+1)$ и эти значения отвечают корням $\{-3, -6\}$ кубического уравнения в $M_{9,1}$, т.е точкам 2-го порядка $(0,14), (0,11)$ кривой Монтгомери. Например, одной из точек кривой $E_{11,7}$ является точка P = (8,1). Тогда $2P$ = (3, –5,), $4P$ = (–4,6), $8P$ = (3,5). Так как $8P = -2P$, то $10P$ = O и порядок этой точки Ord$P$ = 10. Но в подгруппу <P> входит особая точка 2-го порядка $5P = 4P + P = (2, \infty)$. Приняв генератором подгруппы 5-го порядка точку $G = 2P$ простого порядка 5, можно в подгруппе точек <G>, не включающей особых точек, использовать арифметику кривых с групповой операцией (2.2) без особенностей. Это справедливо для любых точек нечетного порядка.

Для криптографических приложений следует искать кривые Эдвардса порядка $N_E = 4n$ с минимальным кофактором 4 при нечетном n, из которых отбираются кривые с простым n. Среди полных кривых Эдвардса (условия $C1$) практически половина имеют порядок $4n$ (n – нечетное). Они являются циклическими, и их порядки пробегают все кратные 4-м числа в границах Хассе. Квадратичные кривые Эдвардса являются нециклическими с тремя точками 2-го порядка и четырьмя или 8-ю точками 4-го порядка. Отсюда следует, что они содержат нециклическую подгруппу, изоморфную Z/2 × Z/4 порядка 8, а порядок этих кривых имеет минимальный кофактор 8. Поэтому кривые порядка $N_E = 4n$ наряду с полными кривыми Эдвардса можно искать лишь среди скрученных кривых в условиях $C2.1$. Согласно утверждению 2.1 и приведенного выше анализа, при $p \equiv 1 mod 4$ все скрученные кривые имеют порядок $N_E = 4n$.

2.4. Число кривых в обобщенной форме Эдвардса порядка 4*n*

Изучив свойства кривых в обобщенной форме Эдвардса, нет смысла изучать статистику порядков этих кривых, как было сделано в [3]. На основе нашей классификации в разделе 2.3 и свойств кривых мы можем найти точное число кривых (2.1) (с точностью до изоморфизма при e =1) порядка $4n$ (n – нечетное) [56]. Для этого рассмотрим 2 случая.

A. $p \equiv 1 \bmod 4$.

Для полных кривых Эдвардса с условием *C*.1 число всех кривых равно числу квадратичных невычетов $(p-1)/2$. Так как для пары квадратичного кручения справедливо $N_E + N_E^t = 2(p+1) = 0 mod 4$, то из $N_E = p + 1 - t = = 0 mod 4 N_E$ и $p + 1 \equiv 2 mod 4$ следует $\pm t \equiv 2 mod 4$. При этом $N_E^t = 0 mod 8$. Итак, если порядок одной из кривых имеет минимальный кофактор 4, то порядок кривой кручения имеет минимальный кофактор 8 и наоборот. Поскольку каждой кривой отвечает одна кривая кручения с инверсией $d \to d^{-1}$, то число полных кривых Эдвардса с минимальным кофактором $M_{A1} = M_{1.1} + M_{1.2} = (p-1)/4$. Злесь обозначены $M_{i,k}$ – число кривых в классах с условиями *C*i.k раздела 2.3, $i, k = 1,2$.

Кроме этого, при $p \equiv 1 \bmod 4$ кривыми с минимальным кофактором 4 являются все скрученные кривые (утверждение 2.1). Их число найдем, пользуясь их кривыми кручения.

Для кривых с условиями *C*.2.2 классификации квадратичные кривые Эдвардса строятся с помощью квадратов $\delta = \left(\frac{d}{a}\right)$, из которых выбрасывается 1, так что остается $(p-3)/2$ квадратичных вычетов. Так как инверсия $\delta \to \delta^{-1}$ дает изоморфную кривую, следует найти число изоморфизмов. Так как элемент (-1) является квадратом при $p \equiv 1 \bmod 4$ и он совпадает со своей инверсией, он вместо пары изоморфных кривых порождает одну кривую. Тогда число пар изоморфных кривых равно $(p-5)/4$. Добавляя к этому числу кривую с $\delta = -1$, получаем число изоморфизмов квадратичных кривых Эдвардса с минимальным кофактором 8 $M_{2.2} = (p-1)/4$. Переход к скрученным кривым Эдвардса с условиями *C*2.1 с минимальным кофактором 4 как квадратичному кручению кривых с условиями *C*2.2 дает то же число кривых $M_{2.1} = (p-1)/4$. Все скрученные кривые Эдвардса при $p \equiv 1 \bmod 4$

имеют порядок $N_E = 4n$. Таким образом, в условиях $C.2$ число кривых с порядком $4n$ равно $M_{A2} = (p-1)/4$. Общее число кривых в форме (2.1) порядка $4n$ при $p \equiv 1 \bmod 4$ равно $M_A = M_{A1} + M_{A2} = (p-1)/2$.

В. $p \equiv 3 \bmod 4$.

Для этого случая кривые (2.1) порядка $4n$ существуют лишь в классе полных кривых Эдвардса (условия $C1$). Любая кривая при этом содержит ровно 2 точки 4-го порядка и половина кривых – 4 точки 8-го порядка. Из $(p-1)/2$ квадратичных невычетов d мультипликативной группы $\mathbf{F}_p^*$ в соответствии с условием теоремы 1.3 следует оставить значения, для которых $\left(\frac{1-d}{p}\right) = -1$. Иными словами, следует найти число пар произведений $d(1-d)$, в которых оба сомножителя – квадратичные невычеты. Подобная задача рассматривалась в работе [44] автором данной работы. Введем обозначение N для квадратичного невычета, S – для квадратичного вычета, при этом (NN), (SS), (NS), (SN) – число пар в схеме Гаусса для всех произведений $m(m+1)$, $m = 1, 2, 3, \ldots, p-1$ [67]. Перепишем $d(1-d) = -d'(d'+1)$, $d' = d-1$, что отвечает схеме Гаусса. В этой схеме следует найти число пар (SN), так как $(-d')$ – квадратичный вычет, а $(d'+1)$ – невычет. Согласно формулы (15) в [44] получим искомое число $(SN) = \frac{p-\varepsilon}{4}, \varepsilon = (-1)^{\frac{p-1}{2}}$. В нашем случае при $p \equiv 3 \bmod 4$, $\varepsilon = -1$ и $(SN) = \frac{p+1}{4}$. Таким образом, число кривых (2.1) порядка $4n$ для данного случая $M_B = (p+1)/4$.

Итак, практически половина всех кривых (2.1) в обобщенной форме Эдвардса при $p \equiv 1 \bmod 4$ и четверть их при $p \equiv 3 \bmod 4$ имеет минимальный четный кофактор 4 порядка кривой. Их число в классе полных кривых Эдвардса вдвое больше, чем в классе скрученных кривых Эдвардса.

Заметим, что введение нового параметра a в обобщенную форму (2.1) кривой Эдвардса лишь в 1.5 раза расширяет множество кривых в форме Эдвардса с минимальным кофактором 4. Множество скрученных кривых с этим свойством существует лишь при $p \equiv 1 \bmod 4$. Позитивным аргументом в пользу скрученных кривых Эдвардса является то, что при $p \equiv 1 \bmod 4$ все они имеют порядок $4n$, что упрощает поиск полезных для криптосистем кривых. Максимальный порядок точки такой кривой равен $2n$, что позволяет найти генератор G криптосистемы одним удвоением случайной точки кривой.

ГЛАВА 3

ПОЛНЫЕ КРИВЫЕ ЭДВАРДСА НАД ПРОСТЫМ ПОЛЕМ

Рассматриваемый в данной главе класс полных кривых Эдвардса обладает рядом замечательных свойств. В сравнении с кривыми в форме Вейерштрасса он прежде всего выигрывает в быстродействия арифметики экспоненцирования точки кривой на 50 – 60% [31,53]. К этому бесспорному преимуществу следует добавить наличие аффинных координат нейтральной точки группы и универсальность закона сложения, что еще в большей степени ускоряет программную реализацию криптоалгоритмов. Хотя в сегодняшних явно устаревших стандартах криптосистем (в том числе стандарте США FIPS-186-2) еще используются кривые над простым полем, рассчитанные в 1997 году [15], внедрение новой технологии полных кривых Эдвардса обещает в итоге дать большой экономический выигрыш.

Симметрия точек кривых Эдвардса относительно обеих координатных осей влечет за собой интересные и удобные свойства этих кривых. Исключая бесполезные изоморфные кривые, в кривых Эдвардса достаточно использовать один параметр d вместо обычных двух параметров a и b кривой в канонической форме Вейерштрасса. В этой главе наряду с известными мы рассматриваем ряд новых свойств полных кривых Эдвардса над простым полем. В частности, интересные свойства возникают при введении операции, обратной удвоению точки кривой – деления точки на 2 (или извлечения квадратного корня в терминах мультипликативной группы). Мы обнаружили простое условие делимости на 2 для точек полной кривой Эдвардса большого порядка (более 4-го) [34,51]. Оно формулируется и доказывается в теореме 3.1 (раздел 3.2). В теореме 3.2 доказано важное свойство, связывающее обе координаты таких точек. Этот подход позже был обобщен в работах Л.В.Ковальчук и А.Ю.Беспалова, которые решили задачу деления точки на любое натуральное число. Их результаты [55] мы приводим в разделе 3.3. В разделе 3.4 обсуждаются некоторые слабые суперсингулярные кривые Эдвардса, порождающие вырожденные пары кривых кручения, и доказывается теорема 3.5 об условиях их существования [51]. Далее предлагается простой метод нахождения порядка точки на полной кривой

Эдвардса (раздел 3.5), и на основе взаимосвязи точек семейства – метод реконструкции точек kP скалярного произведения (раздел 3.6). Доказаны также 2 утверждения о порядках точек кривой. Предложен метод, который при знании всего 1/8 части точек кривой Эдвардса позволяет реконструировать все остальные точки этой кривой, заданные скалярным произведением kP. В разделе 3.7 доказывается теорема о точном числе полных кривых Эдвардса, изоморфных кривым в форме Вейерштрасса с ненулевыми параметрами a и b. В разделах 3.8, 3.9 проведен сравнительный анализ производительности экспоненцирования точки полной кривой Эдвардса и кривой в форме Вейерштрасса. Наконец, в последнем разделе 3.10 мы приводим результаты вычислений общесистемных параметров криптостойких полных кривых Эдвардса, пригодных для стандартизации.

3.1. Общие свойства полных кривых Эдвардса

Полные кривые Эдвардса были определены в работах [2,3] и раздепе 1.3 настоящей работы. Уравнение (1.10) при $e = 1$ с точностью до изоморфизма задает полную кривую Эдвардса вида

$$\boldsymbol{E}: \quad x^2 + y^2 = 1 + dx^2y^2, \quad d(1-d) \neq 0, \left(\frac{d}{p}\right) = -1, \tag{3.1}$$

где параметр $d-$ квадратичный невычет. В интересах криптографических приложений мы в этой главе будем использовать простые поля F_p, хотя многие результаты обобщаются на расширенные поля F_p^m характеристики $p \neq 2$.

В разделе 2.1 мы модифицировали закон сложения точек, поменяв местами координаты $x \leftrightarrow y$. Универсальный модифицированный закон сложения (2.2) для точек кривой (3.1) имеет вид

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1x_2 - y_1y_2}{1 - dx_1x_2y_1y_2}, \frac{x_1y_2 + x_2y_1}{1 + dx_1x_2y_1y_2}\right). \tag{3.2}$$

Закон удвоения точки (x_1, y_1), соответственно, записывается как

$$2(x_1, y_1) = \left(\frac{x_1^2 - y_1^2}{1 - dx_1^2y_1^2}, \frac{2x_1y_1}{1 + dx_1^2y_1^2},\right). \tag{3.3}$$

Напомним, что удобство законов (3.2), (3.3) – в горизонтальной симметрии обратных точек, при этом $-P = -(x_1, y_1) = (x_1, -y_1)$. Нейтральный элемент

группы теперь равен $O = (1, 0)$, точка $D = (-1, 0)$ – точка второго порядка, и $\pm \boldsymbol{F} = (0, \pm 1)$ – точки 4-го порядка.

Среди общесистемных параметров криптосистемы на эллиптических кривых важнейшим элементом является ее генератор как точка достаточно большого простого порядка n. При использовании кривых Эдвардса над простым полем порядок кривой $N_E = 4n$, где n – большое простое число [37]. После нахождения случайной точки $Q = (x_Q, y_Q)$ кривой генератор криптосистемы порядка n нетрудно найти как точку $G = (x_G, y_G) = 4Q$, для чего потребуется два удвоения (т.е. две групповые операции). В данной главе мы показываем, что задача нахождения генератора решается проще – двумя операциями в поле и одним удвоением в группе точек.

Семейством точек большого порядка мы называем 8 точек кривой, лежащих на одной окружности с радиусом, большим 1. В разделе 3.4 дан анализ свойств точек семейства, на основе которых удается без групповых операций находить точки различных порядков и реконструировать точки скалярного произведения.

Идея и метод определения порядков точек кривых Эдвардса рассматривались в работе [34]. Для этого мы привлекали решение задачи, обратной удвоению точки: деление точки на 2. В настоящей главе мы приводим новый подход к решению этой задачи и доказываем необходимое и достаточное условие делимости точки на 2 [51]. Это условие позволило сформировать простой алгоритм вычисления точек требуемого порядка для использования в криптосистемах [48 – 51].

Заметим, что каждая не базовая точка (x_1, y_1), $x_1^2 \neq y_1^2$ порождает семейство из 8 точек $(\pm x_1, \pm y_1)$, $(\pm y_1, \pm x_1,)$, лежащих симметрично на одной окружности радиуса $\sqrt{x_1^2 + y_1^2}$ (по 2 в каждом квадранте). Все они связаны между собой через 3 базовых точки: D и $\pm F$. Согласно формулы (3.2) имеем:

$$\begin{aligned} P + D &= (x_1, y_1) + (-1, 0) = (-x_1, -y_1) = P^*, \\ P + F &= (x_1, y_1) + (0, 1) = (-y_1, x_1\ ), \\ P - F &= (x_1, y_1) + (0, -1) = (y_1, -x_1\ ). \end{aligned} \qquad (3.4)$$

Остальные 4 точки семейства формируются аналогично обратной точкой $-P$.

Некоторые известные свойства кривых (3.1) уже обсуждались в главе 1. Рассмотрим далее ряд новых свойств полных кривых Эдвардса.

3.2. Необходимое и достаточное условие делимости точки полной кривой Эдвардса на два

Впервые эта задача для кривых (3.1) рассматривалась в работах автора [34,35]. Далее в работах [51,55] она получила более строгое решение с доказательсвами теорем и методом нахождения координат двух точек деления на 2.

Пусть $P = (x_1, y_1)$ и $2P = (a, b)$. В этом случае можно записать обратную удвоению (3.2) точки операцию деления точки на 2 как $\frac{(a,b)}{2} = P$. Вторым решением операции деления на 2 будет точка $\frac{(a,b)}{2} = P + D$, где D – точка 2-го порядка. Согласно (3.4) $P + D = (-x_1, -y_1) = P^*$. Ясно, что удвоение этих двух точек дает один результат $2P = 2P^*$. Деление на 2 точки аддитивной группы имеет аналогию с извлечением корня квадратного из элемента мультипликативной группы поля характеристики $p \neq 2$. С этими операциями связаны родственные проблемы дискретного логарифмирования [59].

Воспользуемся формулой удвоения (3.2). Исключим из рассмотрения 4 точки кривой (3.1), лежащие на окружности радиуса 1: нуль группы равен $O = (1, 0)$, точку $D = (-1, 0)$ второго порядка, и $\pm\boldsymbol{F} = (0, \pm 1)$ – точки 4-го порядка. Обозначим $Z = y_1/x_1$, $V = y_1 x_1 \neq 0$. Согласно (3.1) вторую координату b в (3.3) можно выразить двумя формулами

$$\frac{2x_1y_1}{x_1^2+y_1^2} = \frac{2Z}{1+Z^2} = b, \qquad \frac{2x_1y_1}{1+dx_1^2y_1^2} = \frac{2V}{1+dV^2} = b\,. \tag{3.5}$$

Тогда с учетом (3.5) и введенных обозначений для одной точки P кривой, не лежащей на окружности радиуса 1, одновременно справедливы два квадратных уравнения

$$Z^2 - 2b^{-1}Z + 1 = 0, \qquad dV^2 - 2b^{-1}V + 1 = 0, \qquad b \neq 0,1 \tag{3.6}$$

с дискриминантами

$$\Delta_1 = 4b^{-2}(1 - b^2), \quad \Delta_2 = 4b^{-2}(1 - db^2), \tag{3.7}$$

и решениями

$$Z_{1,2} = b^{-1}(1 \pm \sqrt{1 - b^2}) \qquad V_{1,2} = (db)^{-1}\left(1 \pm \sqrt{1 - db^2}\right). \tag{3.8}$$

На основе формул (3.5) – (3.8) сформулируем и докажем следующую теорему.

Теорема 3.1. *Для любой точки* (a,b) *кривой Эдвардса* (3.1)*, не лежащей на окружности радиуса* 1*, существуют* 2 *точки деления* $(a,b)/2 \in \{P, P+D\}$ *тогда и только тогда, когда* $\left(\frac{1-b^2}{p}\right) = 1$. *При* $\left(\frac{1-b^2}{p}\right) = -1$ *точка* (a,b) *на* 2 *не делится*.

Доказательство.

Необходимость. Удвоение любой точки $P = (x_1, y_1)$ с ненулевыми координатами согласно закону (3.3) порождает единственную точку $2P = (a,b)$, причем координаты точек P и $2P$ являются решениями двух квадратных уравнений (3.6) в поле F_p. Необходимым условием существования решения первого из уравнений (3.6), как следует из (3.7), является то, что элемент поля $(1 - b^2)$ есть ненулевой квадрат в этом поле, т.е. $\left(\frac{1-b^2}{p}\right) = 1$. При выполнении этого условия кроме точки P, для которой $2P = (a,b)$, существует точка $P^* = P + D = (-x_1, -y_1)$, для которой $2P^* = 2P + +2D = (a,b)$ с учетом $2D = O$. При $\left(\frac{1-b^2}{p}\right) = -1$ уравнение (3.6) решений в поле F_p не имеет и точек деления на 2 не существует. Необходимость доказана.

Достаточность. Для любой не базовой точки P кривой (3.1), для которой имеет место равенства (3.5), справедливы оба тождества (3.6). Достаточно

потребовать, чтобы один из дискриминантов (3.7) был квадратом, из чего сразу следует, что и второй дискриминант есть квадрат. Пусть (a,b) – точка кривой (3.1). Тогда равенство $a^2 + b^2 = 1 + da^2b^2$ можно записать как $(1 - b^2) = a^2(1 - db^2)$. Отсюда видим, что для любой точки (a, b) кривой величины $(1 - b^2)$ и $(1 - db^2)$ являются обе квадратичными вычетами, либо – невычетами. В первом случае существуют две точки деления $(a, b)/2$, в противном случае – нет. Теорема доказана. ▲

Найдем теперь координаты точек деления. Эта задача решалась в работах [34,55]. Пусть получены все 4 решения (3.8). Запишем равенства

$$\left(1 + \sqrt{1 - b^2}\right)\left(1 - \sqrt{1 - b^2}\right) = b^2, \left(1 + \sqrt{1 - db^2}\right)\left(1 - \sqrt{1 - db^2}\right) = db^2.$$

Из решений (3.8) и последних равенств можно получить

$$Z_1 Z_{,2} = 1,\ V_1 V_2 = d^{-1} \quad \Rightarrow \quad Z_{,2} = {Z_1}^{-1},\ V_2 = d^{-1}{V_1}^{-1} \qquad (3.9)$$

Как видим, оба взаимно-обратных решения Z_i либо квадраты, либо – квадратичные невычеты, для второго же равенства одно из решений – квадратичный вычет, другое –невычет.

Примем, например, что Z_1 и V_1 – оба квадраты или оба – квадратичные невычеты. Тогда из определений $Z = \frac{y_1}{x_1},\ V = y_1 x_1$ получим равенства

$$x_1^2 = \frac{V_1}{Z_1},\ y_1^2 = Z_1 V_1, \quad x_2^2 = \frac{V_1}{Z_2}, \quad y_2^2 = Z_2 V_1. \qquad (3.10)$$

в которые включены лишь один из корней V_1 и оба корня Z_1 и Z_2. Если учесть (3.9), то возникает лишь 2 значения в (3.10), но с неоднозначностью координат, так как

$$x_2^2 = \frac{V_1}{Z_2} = Z_1 V_1 = y_1^2,\ y_2^2 = Z_2 V_1 = \frac{V_1}{Z_1} = x_1^2\ . \qquad (3.11)$$

Для устранения неоднозначности следует использовать вторую координату a точки $2P$ в законе (3.3)

$$\frac{x_1^2 - y_1^2}{1 - dx_1^2 y_1^2} = a \ \Rightarrow \ x_1^2 - y_1^2 = a(1 - dx_1^2 y_1^2). \qquad (3.12)$$

Отсюда видно, что замена координат $(x \leftrightarrow y)$ меняет знак в правой части. Это позволяет отобрать нужную пару значений из (3.10). Если справедливо $x_1^2 = \frac{Z_1}{V_1},\ y_1^2 = Z_1 V_1$, то решениями будут 4 точки: (x_1, y_1), $(-x_1, -y_1)$, $(x_1, -y_1)$ и $(-x_1, y_1)$. Из них только для двух точек выполняется $V_1 = x_1, y_1$. Последний тест оставляет два решения для точек деления на 2: $(a, b)/2 \in \{(x_1, y_1) = P, (-x_1, -y_1) = P + D\}$.

Обозначим

$$\delta^2 = (1 - b^2) \Rightarrow \delta = \sqrt{1 - b^2},$$

тогда с учетом уравнения (3.1) имеем $(1 - b^2) = a^2(1 - db^2)$, и решения (3.8) можно записать

$$Z_{1,2} = b^{-1}(1 \pm \delta), \qquad V_{1,2} = (dab)^{-1}(a \pm \delta). \tag{3.13}$$

Квадраты координат точек деления на 2 (3.11) теперь можно выразить

$$x_2^2 = Z_1 V_1 = (dab^2)^{-1}(1 + \delta)(a \pm \delta) = y_1^2, \tag{3.14}$$

$$y_2^2 = Z_2 V_1 = (dab^2)^{-1}(1 - \delta)(a \pm \delta) = x_1^2. \tag{3.15}$$

Пример 3.1. Пусть точка $P = (x_1, y_1) = (2,4)$ есть точка кривой $x^2 + y^2 = (1 + 7x^2y^2)$ над полем F_{13}, с порядком $N_E = 20$. Согласно (3.2) вычисляем $2P = (a, b) = (5,6)$, и параметр $\delta = \pm 2$. Из (3.13) получим $Z_{1,2} = 6^{-1}(1 \pm 2) = \{2^{-1}, 2\}$, $V_{1,2} = (7 * 5 * 6)^{-1}(5 \pm 2) = \{8,10\}$, $V_1 = 8$. Здесь $\{2^{-1}, 2, 8\}$ – квадратичные невычеты. По формулам (3.14) находим $x_2^2 = Z_1 V_1 = 2^{-1} * 8 = 4 = y_1^2$, $y_2^2 = Z_2 V_1 = 2 * 8 = 3 = x_1^2$. Примем $x_1^2 = 3, y_1^2 = 4$. Тестируем согласно (3.12): $3 - 4 = 5(1 - 7 * 3 * 4y_1^2) = 1$. Здесь знаки левой и правой части обратны, выбор неверен. Тогда правильный выбор квадратов координат: $x_2^2 = 4$, $y_2^2 = 3$. Решениями являются 4 точки $(2,4), (-2,-4), (2,-4), (-2,4)$. Из них только для 2-х первых точек выполняется равенство $V_1 = 8$. Итак, получили $\frac{(5,6)}{2} = \{(2,4), (-2,-4)\}$.

Данный метод нахождения координат точек деления на два более подробно изложен в следующем разделе. Мы видим, что он довольно трудоемок. В разделе 3.3 мы обсудим этот метод вычисления точек $\frac{(a,b)}{2}$.

Стоит подчеркнуть, что для наших задач важен не столько результат выполнения операции деления $\frac{(a,b)}{2}$, сколько констатация факта, делится или не делится точка на 2 (или существует ли корень квадратный из точки в терминах мультипликативной группы). Это свойство точек позволяет без использоавния групповых операций находить точки максимального порядка

$4n$ полной кривой Эдвардса. Для 4-х точек полной кривой Эдвардса на осях x и y на 2 всегда делится точка D, так что $D/2 = \pm F$ (или $\pm 2F = D$), и точки $\pm F$, если существуют точки 8-го порядка $(\pm x_1, \pm x_1)$.

В следующей теореме доказываются новые свойства координат точки полной кривой Эдвардса.

Теорема 3.2. *Для любой не базовой точки* (x_1, y_1) *кривой* (3.1) *справедливо равенство* $\left(\frac{1-x_1^2}{p}\right)\left(\frac{1-y_1^2}{p}\right) = \left(\frac{1-d}{p}\right)$.

Доказательство.

Для точки (x_1, y_1) с учетом определения (3.1) запишем произведение

$(1 - y_1^2)(1 - x_1^2) = 1 + x_1^2 y_1^2 - x_1^2 - y_1^2 = y_1^2 - d y_1^2 = (1 - d)\, x_1^2 y_1^2$.

Тогда из последнего соотношения сразу следует, что произведение $(1 - y_1^2)(1 - x_1^2)$ является квадратичным невычетом при $\left(\frac{1-d}{p}\right) = -1$, и наоборот, что и доказывает утверждение теоремы. ▲

Теорема 3.2 легко обобщается и на изоморфные кривые (1.10) с параметром $e \neq 1$. Действительно, с помощью замены $u = x/e$, $v = y/e$, $d' = de^4$ получаем уравнение изоморфной (1.10) кривой $u^2 + v^2 = 1 + d'u^2v^2$. Для него условие теоремы справедливо после замены $(x,y) \to (u,v)$ и $d \to d'$.

Для кривых Эдвардса, не имеющих точек 8-го порядка, элемент $(1 - d)$ является квадратичным невычетом [30]. Тогда из теоремы 3.2 следует, что любая точка такой кривой, не лежащая на осях x и y, имеет пару значений $(1 - y_1^2)$ и $(1 - x_1^2)$, одно из которых есть квадратичный вычет, а другое – квадратичный невычет. В частности, для точки максимального порядка $4n$ элемент $(1 - y_1^2)$ – квадратичный невычет, а $(1 - x_1^2)$ – – квадратичный вычет.

Определение координат точек деления на 2 рассмотрено также в работе [34]. В следующем разделе рассмотрим обобщение понятия делимости точки на произвольное натуральное число.

3.3. Извлечение корней произвольной степени из точки полной кривой Эдвардса

В этом разделе приведем интересные оригинальные результаты, полученные Л.В.Ковальчук и А.Ю.Беспаловым в работе [55], с сохранением введенных в ней понятий и обозначений.

Как правило, множество точек кривой Эдвардса (3.1) образует циклическую группу, порядок которой кратен 4. В криптографических приложениях используются исключительно такие кривые Эдвардса, порядки которых $N_E = 4n$, где n – большое (от 190 бит) простое число. Поэтому далее в этом разделе мы будем рассматривать только кривые такого порядка.

Согласно свойствам циклической группы [20], в группе E_p кроме одной точки второго порядка и двух точек четвертого порядка имеется $\varphi(n) = n - -1$ точек порядка n, столько же точек порядка $2n$, и $\varphi(4n) = 2(n-1)$ точек порядка $4n$. Для криптографических приложений используется подгруппа кривой E_p, имеющая порядок n. Она состоит из нейтрального элемента $O = (1,0)$ и всех точек порядка n. Образующий элемент G этой подгруппы мы называем генератором криптосистемы (базовой точкой кривой E_p).

Для дальнейшего изложения нам понадобятся такие определения.

Определение 3.1: Пусть $P \in E_p$, $k \in N$. Будем говорить, что точка P делится на k, если $\exists R \in E_p : P = kR$, где под выражением kR мы понимаем k-кратное сложение R (как принято говорить, умножение точки R на скаляр k, или – скалярное умножение).

Множество точек E_p, которые делятся на k, будем обозначать $T_k(E_p)$ или $T_k(p)$.

Определение 3.2: Пусть $P \in T_k(E_p)$. Будем говорить, что точка R является корнем k-ой степени из P, если $kR = P$.

Цель данного раздела – сформулировать критерии делимости точки на k, $1 \le k \le 4n$, с условиями, удобными для вычислений, и показать, как эти критерии могут быть использованы в криптографических приложениях. Заметим, что для 3-х значений k, таких, что $k \mid 4n$, справедлив следующий критерий делимости: $P \in T_k\left(E_p\right) \Leftrightarrow \frac{4n}{k} P = O$.

3.3.1. Критерий делимости точки на 2

В работе [51] было сформулировано и доказано необходимое и достаточное условие делимости точки кривой на 2. Здесь мы еще раз приведем его доказательство, дополненное алгоритмом вычисления координат точек деления [55].

Теорема 3.3: критерий делимости точки на 2. *Пусть* $P = (a,b) \in E_p$. Тогда *следующие условия равносильны*:

1) $P \in T_2\left(E_p\right)$;

2) $$\left(\frac{1-b^2}{p}\right) = 1 \qquad (3.16)$$

Доказательство.

1. Докажем, что из условия 1) следует условие 2). Пусть $P = (a,b) \in T_2\left(E_p\right)$, то есть $\exists R = (x,y) : P = 2R$. Тогда, согласно (3.1)

$$a^2 + b^2 = 1 + da^2b^2 \qquad (3.17)$$

и, согласно (3.1) и (3.3), для координат точки R справедлива система уравнений:

$$\begin{cases} x^2+y^2=1+dx^2y^2; \\ \dfrac{2xy}{x^2+y^2}=\dfrac{2xy}{1+dx^2y^2}=b; \\ \dfrac{x^2-y^2}{1-dx^2y^2}=a, \end{cases} \tag{3.18}$$

Из 1-го и 2-го уравнений системы (3.18) получаем $\dfrac{2xy}{x^2+y^2}=b$, откуда $2\dfrac{y}{x}=b\left(1+\left(\dfrac{y}{x}\right)^2\right)$. Обозначим в последнем уравнении $V=\dfrac{y}{x}$, получим уравнение $2V=b\left(1+V^2\right)$, или

$$V^2-2b^{-1}V+1=0. \tag{3.19}$$

Вследствие условия 1) теоремы, уравнение (3.19) имеет решение. Как следствие, его дискриминант является квадратичным вычетом по $mod\ p$, то есть $D=4b^{-2}-4=4\left(b^{-2}-1\right)=\dfrac{4\left(1-b^2\right)}{b^2}$ – квадратичный вычет по модулю p. Так как $4b^{-2}\in Q_p$, то последнее условие эквивалентно условию $\left(\dfrac{D}{p}\right)=\left(\dfrac{1-b^2}{p}\right)=1$, то есть выполняется условие 2) теоремы 3.3.

2. Докажем, что из условия 2) следует условие 1).

Пусть $\left(\dfrac{1-b^2}{p}\right)=1$. Покажем, что $\exists x,y\in F_p$, которые являются решениями системы (3.18) при заданных $a,b\in F_p$, $P=(a,b)\in T_2\left(E_p\right)$.

Из условия 2) получаем, что уравнение имеет два решения:

$$V_{1,2}=\frac{2b^{-1}\pm 2b^{-1}\sqrt{1-b^2}}{2}=b^{-1}\left(1\pm\sqrt{1-b^2}\right). \tag{3.20}$$

Поскольку $P=(a,b)\in E_p$, то есть выполняется условие 1), получаем $\frac{1-b^2}{1-db^2}=a^2$, откуда $(1-db^2)(1-b^2)\in Q_p$, и, как следствие, $\left(\frac{1-db^2}{p}\right)=\left(\frac{1-b^2}{p}\right)=1$.

Поэтому уравнение

$$Z^2-\frac{2}{bd}Z+\frac{1}{d}=0, \tag{3.21}$$

дискриминант которого равен $D=\frac{4}{b^2d^2}-\frac{4}{d}=\frac{4-4b^2d}{b^2d^2}=\frac{4}{b^2d^2}(1-b^2d)$, также имеет два решения:

$$Z_{1,2}=\frac{\frac{2}{bd}\pm\frac{2}{bd}\sqrt{1-b^2d}}{2}=\frac{1}{bd}\left(1\pm\sqrt{1-b^2d}\right). \tag{3.22}$$

При этом

$$V_1\cdot V_2=1\in Q_p,\ Z_1Z_2=\frac{1}{d}\notin Q_p, \tag{3.23}$$

то есть корни V_1 и V_2 одновременно либо квадратичные вычеты, либо квадратичные невычеты, а из корней Z_1 и Z_2 один всегда квадратичный вычет, а второй – квадратичный невычет.

Далее, уравнения (3.19) и (3.21) эквивалентны, соответственно, уравнениям

$$\frac{2V}{1+V^2}=b\ , \tag{3.24}$$

и

$$\frac{2Z}{1+dZ^2}=b, \tag{3.25}$$

которые выполняются для V_1,V_2 и Z_1,Z_2 соответственно.

Если $V_1 \in Q_p$ (при этом также $V_2 \in Q_p$, как показано выше), то обозначим Z_1 – тот из корней уравнения (3.21), который является квадратичным вычетом. В ином случае (если $V_1 \notin Q_p$), обозначим Z_1 – тот из корней уравнения, который является квадратичным невычетом. Тогда $V_1 Z_1 \in Q_p$, $V_2 Z_1 \in Q_p$.

Обозначим

$$y_1 = \sqrt{V_1 Z_1}, \quad y_2 = -y_1, \quad y_3 = \sqrt{V_2 Z_1}, \quad y_4 = -y_3,$$

$$x_1 = \sqrt{\frac{Z_1}{V_1}}, \; x_2 = -x_1, \quad x_3 = \sqrt{\frac{Z_1}{V_2}}, \quad x_4 = -x_3.$$

Тогда

$$x_i y_i = Z_1, i = \overline{1,4}; \tag{3.26}$$

$$y_i / x_i = V_1, i = \overline{1,2}; \; y_i / x_i = V_2, i = 3,4. \tag{3.27}$$

Подставив формулы в уравнение (3.21), получаем

$$\frac{2 x_i y_i}{1 + d x_i^2 y_i^2} = b, \tag{3.28}$$

то есть $(x_i, y_i), \quad i = \overline{1,4}$ являются решениями 2-го уравнения в системе (3.18).

Аналогично, подставив в (3.19), получаем $2\frac{y_i}{x_i} = b\left(1 + \left(\frac{y_i}{x_i}\right)^2\right)$, откуда

$$\frac{2 x_i y_i}{x_i^2 + y_i^2} = b. \tag{3.29}$$

Приравняв левые части в (3.28) и (3.29), получаем

$$x_i^2 + y_i^2 = 1 + d x_i^2 y_i^2, \tag{3.30}$$

то есть пары (x_i, y_i) являются решениями 1-го уравнения системы (3.18).

Заметим, что если (x, y) являются решениями 1-го и 2-го уравнений системы (3.18), то пары $(y, x), (-x, -y), (-y, -x)$ также являются их решениями. Именно эти пары мы получили в формулах (3.26) и (3.27). Однако, у точки $P = (a, b) \in T_2(p)$ существует всего два корня 2-й степени. Чтобы избавиться от двух лишних точек, используем 3-е уравнение системы. Сначала покажем, что $(x_i, y_i), \quad i = \overline{1,4}$ также являются решениями системы

$$a^2 = \left(\frac{x^2 - y^2}{1 - dx^2 y^2} \right)^2 . \tag{3.31}$$

Действительно, из (3.1) следует, что

$$a^2 = \frac{1 - b^2}{1 - db^2}, \tag{3.32}$$

а из 2-го уравнения системы (3.18), получаем

$$b^2 = \frac{4x^2 y^2}{(1 + dx^2 y^2)^2} . \tag{3.33}$$

Подставив (3.33) в правую часть (3.32) , получаем

$$a^2 = \frac{1 - \dfrac{4x^2 y^2}{(1 + dx^2 y^2)^2}}{1 - \dfrac{4dx^2 y^2}{(1 + dx^2 y^2)^2}} = \frac{(1 + dx^2 y^2)^2 - 4x^2 y^2}{(1 + dx^2 y^2)^2 - 4dx^2 y^2} = \frac{(x^2 + y^2)^2 - 4x^2 y^2}{(1 + dx^2 y^2)^2 - 4dx^2 y^2} =$$

$$= \frac{(x^2 - y^2)^2}{(1 - dx^2 y^2)^2} = \left(\frac{x^2 - y^2}{1 - dx^2 y^2} \right)^2 ,$$

откуда следует (3.31).

Поэтому выполняется ровно одно из равенств: либо $a = \dfrac{x^2 - y^2}{1 - dx^2 y^2}$, либо

$$a = \frac{y^2 - x^2}{1 - dx^2 y^2} .$$

Тогдау из всех пар вида

$$(x,y),(-x,-y),(y,x),(-y,-x), \tag{3.34}$$

которые являются решениями 1-го и 2-го уравнений системы (3.18), только две будут решениями 3-го уравнения этой системы. Исходя из вида левой части 3-го уравнения (3.18), делаем вывод, что это пары вида $R_1=(x,y)$ и $R_2=(-x,-y)$.

Именно эти точки R_1 и R_2 являются корнями второй степени из точки $P=(a,b)$.

Теорема 3.3 доказана. ▲

Обозначим Z – тот корень уравнения, для которого

$$VZ \in Q_p, \tag{3.35}$$

где V – любой из корней уравнения.

Следствие 1: Пусть $P=(a,b)\in T_2(E_p)$, $R=(x,y)\in E_p$, $P=2R$. Тогда в наших обозначениях $y^2=\dfrac{(a+1)d\cdot Z^2-a+1}{2}$.

Доказательство**:** Так как $P=(a,b)\in T_2(E_p)$, то, согласно теореме 3.3, существуют решения V_1,V_2 уравнения (3.19) и решения Z_1,Z_2 уравнения (3.21). Вследствие (3.20) и (3.21), либо $V_1Z_1\in Q_p$, либо $V_1Z_2\in Q_p$, поэтому мы всегда можем выбрать Z в соответствии с (3.35). Если $P=2R$, где $R=(x,y)$, то, согласно теореме 3.3, (x,y) является решением системы (3.18). Тогда, из 3-го уравнения системы (3.18), учитывая (3.26), получаем $a=\dfrac{x^2-y^2}{1-dZ^2}$, откуда

$$x^2-y^2=a(1-dZ^2), \tag{3.36}$$

а из 1-го уравнения системы (3.18) получаем

$$x^2 + y^2 = 1 + dZ^2. \qquad (3.37)$$

Вычитая из (3.37) (3.36), получаем $2y^2 = 1 + dZ^2 - a + adZ^2$, откуда

$$y^2 = \frac{(a+1)dZ^2 - a + 1}{2}. \qquad (3.38)$$

Следствие доказано. ▲

Следствие 2: Пусть $P = (a,b) \in T_2(E_p)$, V и Z выбраны согласно (3.20) (3.22) и (3.35), $R = (x, y) \in E_p$, $P = 2R$. Тогда

$$1 - y^2 = \frac{(a+1)(1 - dZ^2)}{2}. \qquad (3.39)$$

Доказательство выполняется соответствующим применением (3.38)

Следствием теоремы 3.3 также можно считать следующий алгоритм вычисления корня 2-ой степени из точки $P = (a,b) \in T_2(E_p)$.

Алгоритм 1: Вычисление корня 2-ой степени из точки $P = (a,b) \in T_2(E_p)$.

Вход: a, b (такие, что $1 - b^2 \in Q_p$).

1. Вычислить $V_1 = b^{-1}\left(1 - \sqrt{1-b^2}\right)$,

$$V_1 Z_1 = \frac{1}{b^2 d}\left(1 - \sqrt{1-b^2}\right)\left(1 - \sqrt{1-bd^2}\right),\ Z_1 = (bd)^{-1}\left(1 - \sqrt{1-bd^2}\right)$$

(или $Z_1 = bd^{-1}\left(1 + \sqrt{1-b^2 d}\right)$).

2. Если $V_1 Z_1 \notin Q_p$, то $Z_1 = (bd)^{-1}\left(1 + \sqrt{1-bd^2}\right)$

(или $Z_1 = bd^{-1}\left(1 + a^{-2}\sqrt{1-b^2}\right)$).

3. Вычислить $y = \sqrt{\frac{(a+1)dZ_1^2 - a + 1}{2}}$.

4. Вычислить $x = y^{-1} Z_1$.

Выход: $R_1 = (x, y), \quad R_2 = (-x, -y)$.

Вычислительная сложность: Алгоритм использует 16 операций умножения (каждая порядка $(\log p)^2$ битовых операций), 6 операций вычисления обратного по модулю (каждая порядка $(\log p)^3$ битовых операций), 3 операции вычисления квадратного корня по $mod\ p$ (каждая порядка $(\log p)^3$ битовых операций), а также 10 операций сложения и вычитания, время работы которых значительно меньше. Итак, вычислительную сложность алгоритма можно оценить как $O(\log^3 p)$.

3.3.2. Критерий делимости точки на 4

Теперь мы готовы сформулировать критерий делимости точки $P = (a, b)$ на 4.

Теорема 3.4: Пусть $P = (a, b) \in T_2(E_p)$. Обозначим s_1 произвольный корень из $1 - b^2$, а s_2 – корень из $1 - b^2 d$, такой, что $(1 - s_1)(1 - s_2) \notin Q_p$.

Тогда, следующие условия равносильны:

1) $P \in T_4(E_p)$

2) $(a+1) s_2 (1 - s_2) \notin Q_p$ (3.40)

Доказательство:

Пусть $P = 2Q$, где $Q = (x, y)$.

Заметим, что вследствие (3.20), (3.22) и (3.23)

$$(1 - s_1)(1 + s_1) = b^2 \in Q_p, \tag{3.41}$$

$$(1 + s)(1 - s) = db^2 \notin Q_p, \tag{3.42}$$

где s – произвольный (любой из двух возможных) корень из $1 - b^2 d$.

Поэтому для произвольного корня s_1 из величины $1-b^2$ всегда будет существовать единственный корень s_2 из величины $1-b^2 d$ такой, что

$$(1-s_1)(1-s_2) \notin Q_p,$$

и условие теоремы корректно. В принятых обозначениях и согласно (3.20), (3.22), (3.41) и (3.42) последнее выражение эквивалентно тому, что $Z = \frac{1-s_2}{bd}$. Тогда, согласно (3.28) и с учётом (3.39), имеем $1-y^2 = \frac{a+1}{2}\left(1-dZ^2\right) = \frac{a+1}{2} s_2(1-s_2) \cdot \frac{2}{b^2 d} = \frac{(a+1)s_2(1-s_2)}{b^2 d}$.

Поскольку по предположению $d \notin Q_p$, то условие $\left(\frac{1-y^2}{p}\right) = 1$ равносильно условию $(a+1)s_2(1-s_2) \notin Q_p$. Теорема доказана. ▲

Замечание 1:

Алгоритмы получения корня 4-й степени из точки $P \in T_4(E_p)$ могут быть реализованы либо непосредственно, с использованием теоремы , либо последовательным двукратным применением алгоритма 1.

Замечание 2:

Если $|E_p| = 4n$, где n – (большое) простое число, то для любой точки $P = (a,b)$ выполнено одно из двух условий: либо $1-b^2 \in Q_p$, либо $1-a^2 \in Q_p$. Это эквивалентно тому, что $P = (a,b) \in T_2(E_p)$, либо $P' = (b,a) \in T_2(E_p)$.

3.3.3. Сравнительный анализ алгоритмов генерации базовой точки кривой Эдвардса

Рассмотрим три алгоритма генерации базовой точки – «классический» (применяется, например, в алгоритме ДСТУ 4145-2002, [82]), алгоритм, который основывается на теореме 1, и алгоритм, который основывается на

теореме 2. Проведём их сравнительный анализ по быстродействию и некоторым другим факторам.

Алгоритм 2 (ДСТУ 4145-2002):

Вход: эллиптическая кривая $E(F_p)$.

1. Случайно выбрать точку $P=(x,y)\in E(F_p)$.
2. Вычислить nP.
3. Если $nP\neq O$, возвращаемся к шагу 3.2.

Выход: $G=P=(x,y)$ – базовая точка.

Вычислительная сложность. Алгоритм использует примерно $\log p$ сложений точек. При каждом сложении точек выполняется 6 умножений ($6\log^2 p$ битовых операций) [22], 6 делений с остатком ($6\log^2 p$ битовых операций) и два алгоритма Евклида ($2\log^3 p$ битовых операций).

Поэтому общее время работы алгоритма составляет $96\log^3 p+4\log^4 p$ (с учетом того, что среднее количество шагов до успеха равно четырём).

Следующий алгоритм 3 основан на теореме 3.3. В п.2 алгоритма 3 проверяется выполнение условия 2 теоремы 3.3; если оно выполняется, то, согласно теореме, полученная точка делится на 2, и для построения базовой точки её достаточно удвоить. Если же условие теоремы 3.3 не выполняется, то точка, полученная перестановкой координат, будет делиться на 2, а точка, полученная в результате её удвоения, будет делиться на 4, т.е. будет базовой точкой.

Алгоритм 3:

Вход: эллиптическая кривая $E(F_p)$.

1. Случайно выбрать точку $P=(a,b)\in E(F_p)$.

2. Если $a\in\{0,1,-1\}$, то перейти к п.1.

3. Если $1-b^2\notin Q_p$, то $c\leftarrow a, a\leftarrow b, b\leftarrow c$.

4. Вычислить $P\leftarrow 2P$.

Выход: $G=P$ – базовая точка.

Вычислительная сложность. Алгоритм использует одно умножение и одно деление с остатком ($2\log^2 p$ битовых операций), одну проверку квадратичности ($2\log^3 p$ битовых операций) и одно удвоение точки ($12\log^2 p+2\log^3 p$ битовых операций).

Всего $4\log^3 p+14\log^2 p$ битовых операций.

Следующий алгоритм построен по аналогии с алгоритмом 3, но использует теорему 3.4. Он является, фактически, алгоритмом проверки делимости точки на 4. Действительно, любая точка $G\in T_4(E_p)$, такая, что $G\neq O$, где $O=(1,0)$, является базовой точкой кривой.

Алгоритм 4:

Вход: эллиптическая кривая $E(F_p)$.

1. Случайно выбрать точку $P=(a,b)\in E(F_p)$.

2. Если $a\in\{0,1,-1\}$, то перейти к п.1.

3. Если $1-b^2\notin Q_p$, то $c\leftarrow a,\ a\leftarrow b,\ b\leftarrow c$.

4. Вычислить $s_1=\sqrt{1-b^2}$, $s_2=\sqrt{1-db^2}$ (любые из двух возможных корней).

5. Если $(1-s_1)(1-s_2)\in Q_p$, то $s_2\leftarrow p-s_2$.

6. Если $(a+1)s_2(1-s_2)\in Q_p$, то перейти к п.1.

Выход: G – базовая точка.

Вычислительная сложность. Алгоритм использует 2 умножения и 2 деления с остатком ($4\log^2 p$ битовых операций), одно вычисление корня ($2\log^3 p$ битовых операций) и две проверки квадратичности ($4\log^3 p$ битовых операций). Следует заметить, что алгоритм 4 является вероятностным; при этом среднее количество шагов до успеха равно двум. Поэтому время его работы составляет $12\log^3 p + 8\log^2 p$ битовых операций.

3.3.4. Алгоритмы вычисления корней других степеней

В этом разделе мы приведем критерии и алгоритмы вычисления корней (точек деления) степени $n,\quad 2n,\quad 4n$ и степени k, где $1 < k < 4n, (k,4n) = 1$. Хотя эти критерии и не имеют таких очевидных приложений, как критерии делимости точки на 2 и на 4, но без них вопросы делимости точек кривой и вычисления точек деления не будут решены полностью.

Заметим, что корни степени $2n$ та $4n$ можно вычислить, используя последовательно алгоритмы вычисления корня степени 2 и степени n. Также будут приведены соответствующие критерии делимости точки.

Поскольку все точки кривой E_p образуют циклическую группу порядка $4n$, то справедливыми будут следующие утверждения:

- ровно две точки кривой E_p делятся на $2n$ – это точка второго порядка $D = (-1,0)$ и точка первого порядка $O = (1,0)$;
- ровно одна точка кривой E_p делится на $4n$ – это точка $O = (1,0)$;
- ровно четыре точки кривой делятся на n – это точки $D = (1,0),\quad F = (0,1),\quad -F = (0,-1)$ и $O = (1,0)$;
- каждая точка кривой делится на k, где $1 < k < 4n$, $(k,4n) = 1$.

Кроме того, как было доказано раньше, ровно n точек делятся на 4 (это все базовые точки и точка $O = (1,0)$), и ровно $2n$ точек делятся на 2 (это все

базовые точки G, точка $O=(1,0)$ и все точки вида $2P$, где P –точка порядка $4n$).

Из всего приведенного выше следуют критерии и алгоритмы делимости точек кривой на $n,\quad 2n,\quad 4n$ и на k , где $1<k<4n,\quad (k,4n)=1.$

Приведем теперь алгоритмы вычисления корней соответствующих степеней из точек кривой. Для них нам понадобится точка $P=(x,y)$, которая является образующим элементом группы E_p . Она может быть получена стандартным алгоритмом для нахождения образующего элемента группы, или как корень 4-й степени из базовой точки G .

Алгоритм 5. Вычисление корня степени $2n$ из точки $Z\in T_{2n}(E_p)$.

Вход: точка $Z=D=(-1,0)$ (или $Z=O=(0,-1)$).

1. Если $Z=D$, то $S=G$, иначе $S=2G$.
2. Выход S .

Вместо алгоритма вычисления корня степени $4n$, заметим, что корнем степени $4n$ из точки $O=(0,-1)$ является любая точка кривой.

Алгоритм 6. Вычисления корня степени k из точки кривой, где $1<k<n$, $(k,4n)=1.$

Вход: произвольная точка $Z\in E_p$.

1. Используя алгоритм Евклида, вычислить $u,v\in Z$ такие, что $uk+vn=1.$
2. Вычислить $u=u \bmod n$.
3. $S=uZ$.
4. Выход S .

Наиболее существенными математическими результатами этого раздела можно считать:

- критерии делимости точки кривой Эдвардса на 2 и на 4, а также соответствующие алгоритмы деления точек;
- критерий делимости точки на n и на произвольное число k, взаимно простое с $4n$, где n – простое число, равное порядку циклической подгруппы группы точек кривой Едвардса.

Практическими результатами, которые основываются на перечисленных выше математических результатах, являются, в первую очередь, новые алгоритмы генерации базовой точки кривой Эдвардса. Также приведен сравнительный анализ новых и классических алгоритмов генерации базовой точки. Кроме этого, получены алгоритмы вычисления корня произвольной степени из точки кривой.

- Алгоритмы 3 и 4 имеют одинаковый порядок временной сложности (то есть если учитывать лишь степень полинома, описывающего время работы, без учета мультипликативной константы) и оба они значительно быстрее алгоритма 2.
- Если использовать более точные оценки, то в порядке снижения быстродействия алгоритмы располагаются следующим образом:
  - Алгоритм 3
  - Алгоритм 4
  - Алгоритм 2
- Хотя алгоритм 4 немного уступает алгоритму 3 по быстродействию, у него есть одно бесспорное преимущество – этот алгоритм использует только операции в конечном поле, над которым задана кривая, и не использует арифметику на самой кривой.

Заметим, что приведенные в этом разделе оценки сложности алгоритмов построены для групповых операций в аффинных координатах. На практике

для того, чтобы избавиться от трудоемкой инверсии элементов поля, используются проективные координаты, что всегда дает выигрыш в скорости. В разделах 3.8, 3.9 мы будем давать оценки производительности групповых операций на кривых Эдвардса в проективных координатах.

3.4. Вырожденные пары кривых кручения

Переход к кривой кручения для формы (3.1) Эдвардса осуществляется простой заменой $d \to d^{-1}$ [2], тогда порядки пары этих кривых $N_E = p + 1 \pm t$. Пара кручения называется вырожденной, если след Фробениуса $t = 0$, порядок обеих кривых $N_E = p + 1$. Такая кривая относится к классу суперсингулярных кривых. Этот случай возможен лишь при $p \equiv 3 \bmod 4$ и тогда $4|(p + 1)$. Например, при $d = d^{-1} = -1$ имеем тривиальный случай вырожденной пары кручения. Автор обнаружил еще один нетривиальный пример вырожденной пары кручения для кривой Эдвардса [49]. Докажем следующую теорему.

Теорема 3.5. *При* $p \equiv 3 \bmod 4$ *и* $p \equiv \pm 3 \bmod 8$ *пара кривых кручения в форме Эдвардса над* $\mathbf{F}_p$ *с параметрами* $d \in \{2, 2^{-1}\}$ *является вырожденной с порядком* $N_E = p + 1$.

Доказательство.

Первое условие теоремы обсуждалось выше и связано с делимостью порядка кривой на 4, или $4|(p + 1)$. При выполнении второго условия элемент 2 поля F_p не является квадратом, т.е. $\left(\frac{2}{p}\right) = -1$ [29]. Требуется доказать, что при $d = 2$ оба уравнения пары кривых кручения имеют одинаковый порядок $p + 1$.

Для всех точек кривой (3.1), исключая две общие точки O и D с координатами $x = \pm 1$, $y = 0$, можно записать равенство

$$E: \qquad y^{-2}(x^2 - 1) + 1 = dx^2 .$$

Для кривой кручения после замены $d \to d^{-1}$ имеем

$$E^t: \qquad y^{-2}(x^2 - 1) + 1 = d^{-1}x^2$$

Умножив последнее равенство на (–d), получим

$$-dy^{-2} = -1 + (d-1)V^{-1},\ V = x^2 - 1,$$

причем в левой части имеем квадрат, так как (–d) – квадратичный вычет при $p \equiv 3 \bmod 4$. При $d = 2$ эти уравнения имеют вид:

$$E: \qquad y^{-2} = 2 + V^{-1}, \tag{3.43}$$

$$E^t: \qquad -2y^{-2} = -1 + V^{-1}, \qquad V = x^2 - 1. \tag{3.44}$$

Покажем, что оба уравнения дают одинаковое число решений. Левые части этих равенств пробегают все значения ненулевых квадратов. Число решений (3.43) и (3.44) определяется числом квадратов правой части равенств. При всех $x^2 \neq 1$ переменная V^{-1} пробегает возможные ненулевые значения из множества $\{1, 2, 3, \ldots, p-1\}$, среди элементов которого $(p-1)/2$ квадратичных вычетов. Область возможных значений величины $(2 + V^{-1})$ в уравнении (3.43) смещается к величинам $\{3, 4, 5, \ldots, p-1, 0, 1\}$, среди которых элемент 0 заменил квадратичный невычет 2 исходного множества V^{-1}. Соответственно, в уравнении (3.44) область возможных значений величины $(-1 + V^{-1})$ включает элементы $\{0, 1, 2, 3, \ldots, p-2\}$ с вытеснением элементом 0 квадратичного невычета (–1). Отсюда следует, что число ненулевых квадратичных вычетов в обоих смещенных множествах одинаково.

Определим число квадратов в множестве $V = x^2 - 1,\ \forall x \neq 1$. Такая задача рассмотрена в [44] и в настоящей работе (раздел 3.7.2). Исключим из рассмотрения еще 2 точки кривой $\pm F = (0, \pm 1)$,с координатой $x = 0$ дающие тривиальные решения в (3.43) и (3.44). Согласно леммы 3.1 и формулы (3.59) раздела 3.7.2 число ненулевых квадратичных вычетов в множестве V равно $(p-3)/4$. Такое же число квадратичных вычетов содержит множество $V^{-1} = z^2 - 1,\ \forall z \neq 0,1$. Тогда правая часть (3.43) после подстановки $V^{-1} = z^2 - 1$ примет вид $z^2 + 1$, а правая часть (3.44) – вид $z^2 - 2$. Оба равенства согласно леммы 3.2 дают в наших условиях одинаковое число квадратов, равное $(p-3)/4$. Так как каждое решение для квадратов в (3.43) и (3.44) с исключением базовых точек O, D, $\pm F$ дает по 4 точки $(\pm x, \pm y)$, получаем $(p-3)$ точек, удовлетворяющих равенствам (3.43) и (3.44). Добавляя 4 отброшенные при анализе точки O = (1, 0), D = (– 1, 0) и $\pm F = (0, \pm 1)$, получаем порядок обеих кривых $N_E = p + 1$. Теорема доказана. ▲

Значениями $d = -1, 2$ и 2^{-1} не исчерпывается перечень суперсингулярных кривых Эдвардса. В работе [47] доказано, что если элемент 3 поля F_p является квадратичным вычетом при $p \equiv 3 \bmod 4$, то параметр $d = -(\sqrt{3}\pm 2)/(-\sqrt{3}\pm 2)$ также порождает суперсингулярную кривую.

3.5. Методы нахождения точек заданного порядка

Кривые Эдвардса подходят для использования в криптосистемах, если их порядок $N = 4n$, где n – большое простое число ($n > 2^{163}$). Если порядок генератора кривой $\mathrm{Ord}P = 4n$, то генератор криптосистемы $G = 4P$ имеет порядок n. Точки 8-го порядка отсутствуют согласно теореме 1.4, если $(1 - d)$ – квадратичный невычет [29].

Утверждение 3.3. *На полной кривой Эдвардса* (3.1) *порядка* $4n$ *существуют точки деления на 2 для всех точек, кроме точек* $< P >$ *максимального порядка и точек* $\pm F$ *четвертого порядка.*

Доказательство. Каждой точке kP кривой отвечает скалярный множитель k как элемент кольца целых чисел Z_N с операциями по модулю $N = 4n$. Все нечетные элементы $k \in \{1, 3, 5, \ldots, 4n - 1\}$ кольца $\mathbf{Z}_N$, которым отвечают точки кривой максимального порядка $4n$ и порядка 4 ($\pm \boldsymbol{F} = \pm nP$), не делятся на 2 в кольце $\mathbf{Z}_N$. С другой стороны, все четные элементы кольца $2s$ при делении на два по модулю N (или умножении на 2^{-1}) дают два значения s и $s + N/2$, удвоение которых по модулю N дает вновь $2s = k$. Возвращаясь к точкам kP кривой, заключаем, что утверждение 3.3 доказано.

Если случайная точка кривой Q имеет порядок $2n$, то обе точки деления на 2 $\{Q/2, Q/2+D\}$ имеют максимальный порядок $4n$. Если точка Q имеет порядок n, то порядки точек деления на 2 $\{Q/2, Q/2+D\}$ равны n и $2n$, соответственно.

Прикладное значение доказанной в разделе 3.2 теоремы 3.1 очевидно. Для нахождения порядка точек кривой Эдвардса не требуется вычислять скалярное произведение nQ. Если у случайной точки кривой (x_Q, y_Q) величина $(1 - y_Q^2)$ – квадратичный невычет, то $\mathrm{Ord}(Q) = 4n$. В противном случае порядок точки равен n или $2n$. Получить такую точку можно непосредственно, меняя

местами координаты x_Q и y_Q (теорема 3.2). После этого генератор G порядка n находится одним удвоением $G = 2Q$.

3.6. Взаимосвязь семейств точек больших порядков. Реконструкция точек kP кривой Эдвардса

На основе симметрии восьми точек полной кривой Эдвардса $(\pm x_1, \pm y_1), (\pm y_1, \pm x_1)$ семейства точек, лежащих на одной окружности, и формул (3.4), связывающих эти точки, можно построить алгоритм нахождения всех точек скалярного произведения kP точки P, если ивестен сегмент 1/8 части всех точек. Мы предложили этот метод в работе [51]. Проиллюстрируем этот метод примером.

Пример 3.3.. Рассмотрим кривые Эдвардса с модулем $p = 19$, для которого выполняются оба условия теоремы 3.5. При $d \in \{-1, 2, 2^{-1}\}$ имеем суперсингулярные кривые с порядком $N_E = p + 1 = 20$. Исключим также кривые с порядком, кратным 8, для которых $(1 - d)$ – квадратичный вычет. Получаем две кривые с параметрами $d = 8$ и $d^{-1} = 12$, которые дают пару кривых кручения с порядками 28 и 12 (для них $t = \pm 8$). Точки первой из них представлены на рис.3.1.

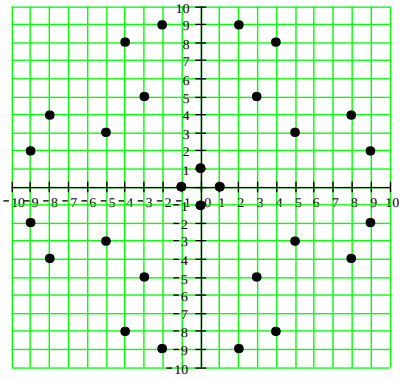


Рис.3.1. График полной Эдвардса при $p = 19$, $d = 8$, $N_E = 28$

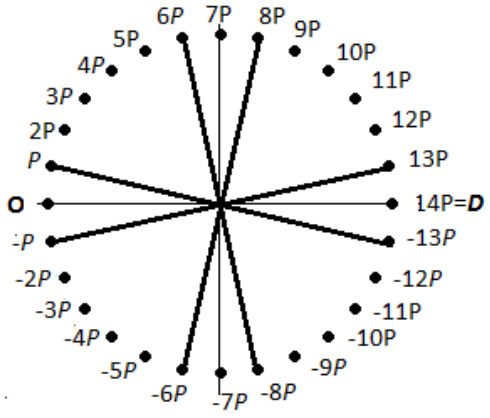


Рис.3.2. Колесо точек циклической кривой порядка $N_E = 28$

Обозначим P = (2,9), Q = (3,5), R = (4,8), S = (5,3). T = (8,4), U = (9,2) – точки первого квадранта. Здесь точками максимального порядка 28 являются точки P, Q, R, для которых значения $(1 - y^2)$ являются квадратичными невычетами. Всех таких точек $\varphi(28) = 12$, по 3 точки в каждом квадранте. Кроме них, имеется 6 точек 14-го и 6 точек 7-го порядков. Удвоение точек P, Q, R согласно (3.3) дает точки 14-го порядка $2P$ = (–8,4) = – T*, $2Q$ = (–9,2) = –U*, $2R$ = (5, –3) = – S. Итак, в первом квадранте имеем одну точку S 14-го порядка, и 2 точки T и U 7-го порядка.

Циклическую группу точек кривой kP можно представить в виде последовательности точек на окружности в порядке роста скалярного числа $k = 0, 1, 2, \ldots, N_E - 1$ по часовой стрелке. Для нашего примера такая точечная окружность представлена на рис.3.2. Назовем этот график колесом точек. Точки колеса, соединенные диаметральными линиями, связаны как P и $P^* = P + D$. Для любой не базовой точки семейство из 8 связанных линиями точек на рис.3.2 лежат на одной окружности на графике кривой рис.3.1.

Знание около 1/8 части всех точек позволяет реконструировать все другие точки кривой. Пусть точка P порождает все точки кривой и известны 4 точки: P = (2,9), $2P$ = (–8,4), $4P$ = (–5,3), $7P$ = $–F$ = (0, –1). В силу свойства $(x_1, y_1) + (–y_1, –x_1) = (0, –1) = –F$, легко находятся точки $6P$ = (–9, –2), $5P$ = (–4,8), $3P$ = (–3,5), меняя местами координаты $x \leftrightarrow y$ и их знаки соответственно точек P, $2P$ и $4P$. Координаты точек kP при k = 0...14 представлены в таблице 3.1. .

Таблица 3.1. Координаты точек kP полной кривой Эдвардса при $d = 8$, $N_E = 28$

kP	O	P	$2P$	$3P$	$4P$	$5P$	$6P$	$7P$	$8P$	$9P$	$10P$	$11P$	$12P$	$13P$	$14P$
x_k	1	2	–8	–3	–5	–4	–9	0	9	4	5	3	8	–2	–1
y_k	0	9	4	5	3	8	–2	–1	–2	8	3	5	4	9	0

Для определения координат точек правее точки 4-го порядка мы используем свойство $P + D = P^* = (–x_1, –y_1)$ или $P – P^* = D = 14P$. Например, точка $13P$, вертикально симметричная точке P и равная – P*, имеет координаты $(–x_1, y_1)$. В таблице 3.1 хорошо видна симметрия (антисимметрия)

координат точек верхней половины рис.3.2: все y-координаты симметричны относительно точки $7P$, тогда как x-координаты обратны по знаку. Точки нижней половины колеса рис.3.2 обратны точкам верхней половины с инверсией знака y-координаты. Например, точка $17P = 28P - 11P = -11P =$ $=(3, -5)$.

Итак, при известных 4-х точках (причем одна из них базовая $-F$) мы без вычислений получили координаты всех 28 точек kP кривой Эдвардса. Этот метод годится очевидным образом для кривой любого порядка, при этом предвычисления состоят в расчете координат точек kP для $k = 2, 3, \ldots, (n-1)/2$, что составляет практически 1/8-ю часть порядка кривой.

Возвращаясь к графику кривой на рис.3.1, мы находим в таблице 3.1 все ее точки как скалярное произведение kP. Точки первого квадранта $Q = (3,5) = 11P$, $R = (4,8) = 9P$ имеют порядок 28, точка $S = (5,3) = 10P$ имеет порядок 14, а две точки $U = (9,2) = -8P =$ и $T = (8,4) = 12P$– порядок 7. Это отвечает выводам предыдущего анализа.

График точек кривой на рис.3.1 является точечным и, естественно, не отвечает термину «кривая Эдвардса». Соавтор данного анализа О.В.Цыганкова предложила интересную анимацию этого графика кусочно-ломаной кривой, в которой последовательно соединяются прямыми линиями точки скалярного произведения $kP, k = 1,2,3, \ldots, N = 28$. Всего имеется $\varphi(28)=12$ точек 28-го порядка, которые могут «нарисовать» 6 различных графиков (обратные точки дают один рисунок). Это точки $P^{\langle 1 \rangle} = (2,9)$, $P^{\langle 2 \rangle} =$ $= (3,5), P^{\langle 3 \rangle} = (4,8), P^{\langle 4 \rangle} = (-4,8), P^{\langle 5 \rangle} = (-3,5), P^{\langle 6 \rangle} = (-2,9)$.
Нарисованные этими 6-ю генераторами графики представлены на рис. 3.3. Так как анимационная идея пришла в воображение ее автора в канун Рождества, мы называем эти удивительные рисунки «рождественскими снежинками».

Утверждение 3.2. *Для кривой Эдвардса порядка* $4n$ *любое семейство из* 8 *точек* $(\pm x_1, \pm y_1), (\pm y_1, \pm x_1)$, *лежащих на одной окружности, содержит* 4 *точки порядка* $4n$, 2 *точки порядка* $2n$ и 2 *точки порядка* n.

Доказательство. Пусть $\mathrm{Ord}(kP) = 4n$, тогда пары точек $\pm kP$ в левой и $\pm kP^*$ в правой части колеса точек рис.2 имеют одинаковый порядок $4n$.

В верхней части колеса точек имеем точки $nP \pm kP$, причем $(n \pm k)$ – четные числа, одно из которых сравнимо с 0mod4, а второе – с 2mod4. Отсюда следует, что порядки этих точек равны n и $2n$.

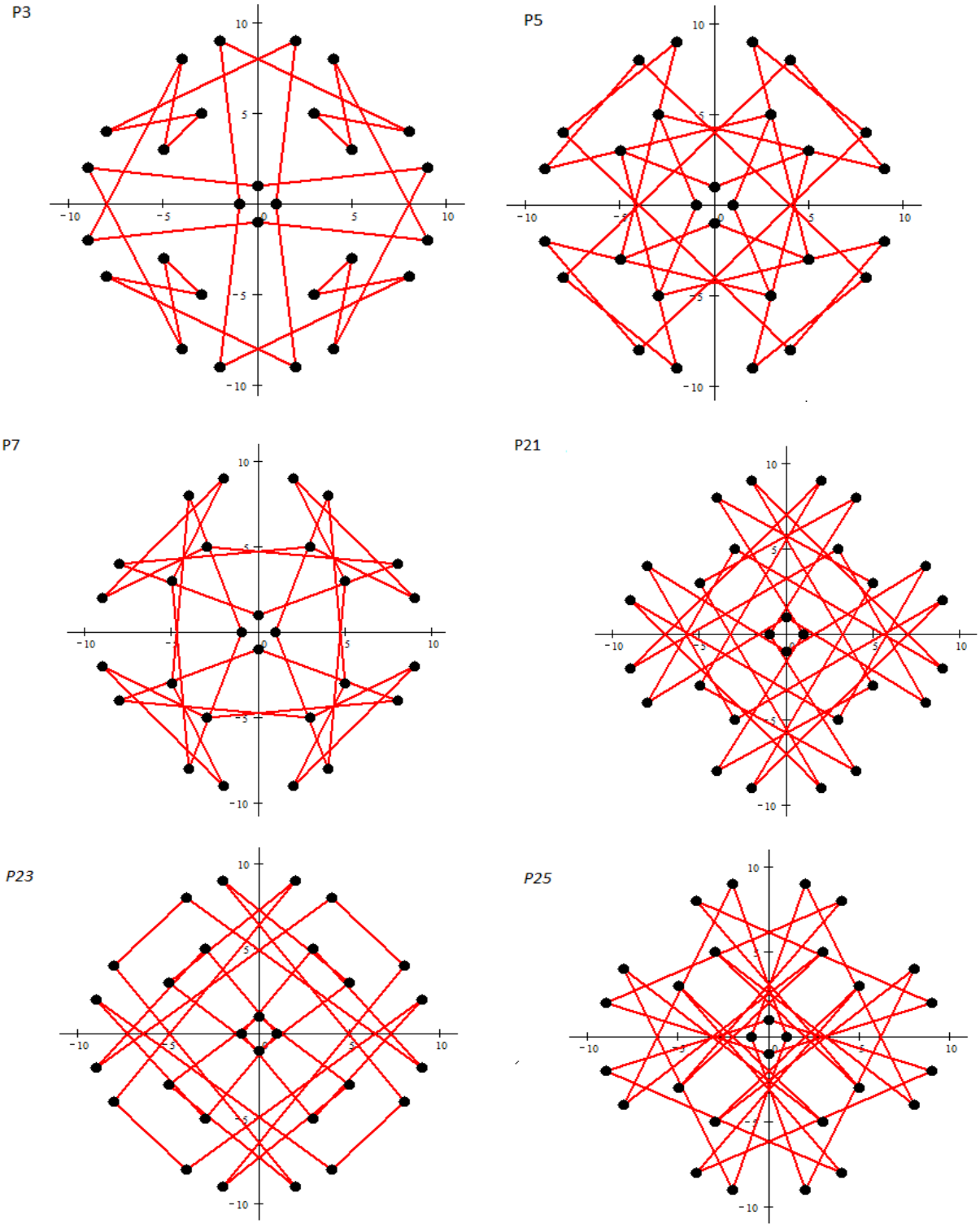


Рис.3,3. Последовательно соединенные точки скалярного произведения $kP^{\langle i \rangle}$ для 6 генераторов $P^{\langle i \rangle}$ кривой Эдвардса порядка 28 над полем F_{19}

Пусть теперь Ord($\pm kP$) = 2n, тогда точки $\pm kP^* = \pm kP + \boldsymbol{D}$ имеют порядок n, так как $n(\pm kP + D) = \pm nkP + nD = \pm D + D = O$. Точки $nP \pm kP$ в верхней части рис.2 имеют сомножителями ($n \pm k$) – нечетные числа, поэтому их порядки (и, соответственно, обратных им точек) максимальны и равны 4n.

Наконец, пусть Ord($\pm kP$) = n, тогда точки $\pm kP^* = \pm kP + \boldsymbol{D}$ имеют порядок 2n, так как $2n(\pm kP + D) = O$. По аналогии с предыдущим абзацем остальные 4 точки имеют порядок 4n. Утверждение 2 доказано.

Замечание. Приведенные выше свойства кривой Эдвардса не должны снижать сложности вычисления дискретного логарифма в группе точек <G> простого порядка n. Действительно, согласно утверждению 3.2 из 8-ми точек

каждого семейства на колесе точек рис.3.2 лишь 2 обратных точки имеют порядок n подгруппы <G>. Поэтому, как и для кривых в канонической форме, сложность DLP [29] здесь снижается лишь вдвое за счет обратных точек. Тем не менее, эти свойства могут послужить основой для поиска новых методов решения проблемы дискретного логарифма.

Заметим, что существует лишь 2 точки максимального порядка, порождающие известный генератор G подгруппы точек простого порядка n – это точки P и P^*, для которых $2P^* = 2P, G = 4P$. Все четные точки колеса рис.3.2 при переходе к порождающей точке P^* сохраняют свои координаты, а нечетные $P^*, 3P^*, 5P^*, \ldots$ меняют знаки обеих координат.

3.7. Точное число полных кривых Эдвардса, изоморфных кривым в канонической форме с ненулевыми параметрами

Различные этапы решения этой задачи отражены в работах автора [41,42,47]. В итоговой работе [44] дано ее наиболее строгое математическое решение, доказаны важные для нее 2 леммы из теории чисел и впервые получены формулы для числа канонических кривых со свойствами (1.23) и, соответственно, изоморфных им кривых Эдвардса. Напомним, что в разделе 1.4 мы ввели зависимый от традиционных параметров (a, b) кривой в канонической форме параметр c как единственный в поле F_p корень

кубического уравнения. Там же даны необходимые и достаточные условия существования одной точки 2-го порядка и 2-х точек 4-го порядка этой кривой. При анализе их использовалась кривая в форме Монтгомери. В разделе 1.8 на их основе получена система линейных уравнений для нахождения неизвестных параметров a и c^2, в уравнения которой входят квадратичные вычеты и невычеты. Для нахождения точного числа канонических кривых, изоморфных кривым Эдвардса, нам потребовалось сформулировать и доказать 2 леммы о числе решений уравнений, связывающих суммы квадратичных вычетов и невычетов. Доказательства опираются на схему Гаусса распределения квадратичных вычетов [66]. В итоге авторам [44] удалось найти формулы расчета точного числа кривых с ненулевыми параметрами a и b и заданными свойствами над любым простым конечным полем F_p характеристики $p > 3$. Известно, что при $ab = 0$ кривая в форме Вейерштрасса становится суперсингулярной со слабыми криптографическими свойствами [29]. Поэтому при нахождении числа кривых мы исключили такие кривые. Кроме того, предложен модифицированный алгоритм поиска кривых с хорошими криптографическими свойствами, изоморфных кривым Эдвардса.

3.7.1. Условия существования ровно двух точек четвертого порядка для эллиптической кривой в форме Вейерштрасса

Логика решения задачи нахождения числа кривых с заданными свойствами требует вновь обратиться к условиям существования ровно 2-х точек 4-го порядка. В разделе 1.5 мы сформулировали (без доказательства) теорему 1.3. Более логичным мы посчитали доказать ее здесь. Мы не используем изначально кривых в форме Монтгомери, поэтому приведенный ниже анализ и доказательство оправданы иными путями получения условий существования 2-х точек 4-го порядка.

Каноническая форма кривой над полем характеристики $p > 3$ описывается известным уравнением в форме Вейерштрасса [16]

$$W_p: \quad y^2 = x^3 + ax + b, \quad \Delta = 4a^3 + 27b^2 \neq 0, \quad a,b \in F_p. \tag{3.45}$$

Согласно определению, операция удвоения точки $P=(x_1,y_1)$, которая дает координаты точки $2P=(x_3,y_3)$, задается следующим образом [29]:

$$\begin{cases} x_3=\nu^2-2x_1, \\ y_3=-y_1-\nu(x_3-x_1), \qquad \nu=\dfrac{3x_1^{\,2}+a}{2y_1}. \end{cases} \tag{3.46}$$

Далее нам понадобятся следующие стандартные обозначения. Множество квадратичных вычетов по модулю простого числа p Q_p определяется как:

$$Q_p=\left\{x\in F_p \middle| \left(\frac{x}{p}\right)=1\right\},$$

где $\left(\dfrac{x}{p}\right)$ – символ Лежандра:

В данной работе мы будем рассматривать только такие кривые (3.45), порядок которых делится на 4. Ясно, что в этом случае кривая обязательно имеет точку второго порядка. Согласно (3.46), точка $P=(x_1,y_1)$ будет точкой второго порядка тогда и только тогда, если $y_1=0$ (в этом случае при вычислении точки $2P$ в (3.46) возникает деление на 0), т.е. точка второго порядка будет иметь координаты $(c,0)$, для некоторого $c\in F_p$. Подставляя в уравнение кривой (3.45) значение $y=0$, получаем, что c – корень уравнения $x^3+ax+b=0$ в поле F_p (который обязательно существует, вследствие существования точки второго порядка). Тогда в наших обозначениях уравнение (3.45) можно переписать в виде

$$y^2=(x-c)(x^2+cx+a+c^2), \text{ где } b=-c^3-ac, \quad c\in F_p. \tag{3.47}$$

Как упоминалось ранее, кривая в канонической форме изоморфна кривой Эдвардса в том и только в том случае, если она содержит ровно две точки четвертого порядка. Следующая теорема дает необходимые и достаточные условия (в терминах параметров кривой (3.45)) существования на кривой E_p ровно двух таких точек.

Теорема 3.6: *необходимыми и достаточными условиями существования ровно двух точек четвертого порядка на кривой* E_p *является:*

$$\text{(i)}\ \left(\frac{-\left(3c^2+4a\right)}{p}\right)=-1, \qquad \text{(ii)}\ \left(\frac{\delta}{p}\right)=1, \text{ где } \delta=3c^2+a. \tag{3.48}$$

Доказательство. *Необходимость.* Предположим, что кривая имеет две точки четвертого порядка, и покажем, что при этом выполняются условия (3.48). Пусть на кривой (3.45) существует ровно две точки четвертого порядка. Тогда, очевидно, она не может содержать более одной точки второго порядка (т.к., согласно определению порядка точки кривой, сумма точек четвертого и второго порядка будет точкой четвертого порядка). Следовательно, парабола в правой части (3.47) не имеет корней в поле F_p, т.е. дискриминант соответствующего квадратного уравнения является квадратичным невычетом. Данный дискриминант равен

$$\Delta = c^2-4\left(a+c^2\right)=-\left(3c^2+4a\right);$$

и, поскольку он квадратичный невычет, то

$$\left(\frac{-\left(3c^2+4a\right)}{p}\right)=-1.$$

Необходимость первого условия (i) в (3.48) доказана. Заметим, что условие $\left(3c^2+4a\right)\neq 0$, которое следует из пункта (i) формулы (3.48), исключает кратные корни кубического уравнения и, тем самым, сингулярные кривые с j-инвариантом $j(W)=0$ [16,29].

Пусть $P=(x_1,y_1)$ – точка 4-го порядка. Тогда при ее удвоении, согласно (3.46), получаем точку второго порядка $D=(c,0)$:

$$\begin{cases} \left(\dfrac{3x_1^{\,2}+a}{2y_1}\right)^2-2x_1=c; \\ -y_1-\left(\dfrac{3x_1^{\,2}+a}{2y_1}\right)(c-x_1)=0. \end{cases} \tag{3.49}$$

Из первого уравнения этой системы получаем

$$y_1^{\,2} = \frac{\left(3x_1^{\,2} + a\right)^2}{4\left(c + 2x_1\right)},$$

а из второго – выражение

$$y_1^{\,2} = -\frac{\left(3x_1^{\,2} + a\right)\left(c - x_1\right)}{2}.$$

Приравнивая правые части этих выражений и сокращая на множитель $3x_1^{\,2} + a$, получим квадратное уравнение для координаты x_1 этой точки:

$$x_1^{\,2} - 2cx_1 - (2c^2 + a) = 0. \qquad (3.50)$$

Корни этого уравнения существуют (вследствие существования точек 4-го порядка), следовательно, дискриминант δ данного уравнения либо равен нулю, либо является квадратичным вычетом. Если дискриминант равен нулю, то, при $y = 0$, уравнение (3.47) принимает вид:

$$\left(x - c\right)^2\left(x + 2c\right) = 0,$$

вследствие чего на кривой существуют две кратные точки второго порядка и кривая является сингулярной. Мы их не рассматриваем. Следовательно, дискриминант уравнения (3.50) $\delta = 3c^2 + a$ является квадратичным вычетом, т.е. выполняется условие (ii) формулы (3.48). Необходимость условий (3.48) доказана.

Достаточность. Пусть выполняются условия (3.48). Покажем, что при этом существует ровно два решения системы (3.49). Путем преобразований уравнений данной системы получим уравнение (3.50) с одной переменной x_1. Поскольку выполняется условие (ii) формулы (3.48), то существует два корня данного уравнения:

$$x_1^{(1),(2)} = c \pm \sqrt{\delta} = c \pm \sqrt{3c^2 + a}. \qquad (3.51)$$

Подставляя эти выражения во второе уравнение системы (3.49), получаем:

$$\begin{aligned} y_1^{\,2} &= -2^{-1}\left(3\left(c \pm \sqrt{\delta}\right)^2 + a\right)\cdot\left(\mp\sqrt{\delta}\right) = \pm 2^{-1}\sqrt{\delta}\left(3\left(c^2 \pm 2c\sqrt{\delta} + \delta\right) + a\right) = \\ &= \pm\, 2^{-1}\sqrt{\delta}\left(3c^2 \pm 6c\sqrt{\delta} + 3\delta + a\right) = \pm\, 2^{-1}\sqrt{\delta}\left(\pm\, 6c\sqrt{\delta} + 4\delta\right) = \\ &= \pm\sqrt{\delta}\left(\pm\, 3c\sqrt{\delta} + 2\delta\right) = \pm\delta\left(\pm\, 3c + 2\sqrt{\delta}\right) = \delta\left(3c \pm 2\sqrt{\delta}\right). \end{aligned} \qquad (3.52)$$

Из (3.52) следует, что решение системы (3.49) существует тогда и только тогда, когда хотя бы одно из выражений

$$3c - 2\sqrt{\delta}, \quad 3c + 2\sqrt{\delta} \tag{3.53}$$

является квадратичным вычетом (т.к., по условию, $\left(\frac{\delta}{p}\right) = 1$).

Покажем, что в нашем случае в точности одно из выражений (3.51) будет квадратичным вычетом. Действительно, перемножив эти выражения, получим:

$$\left(3c - 2\sqrt{\delta}\right) \cdot \left(3c + 2\sqrt{\delta}\right) = 9c^2 - 4\delta = -\left(3c^2 + 4a\right),$$

что, согласно пункту (i) условия (3.48), является квадратичным невычетом. Отсюда, вследствие свойства мультипликативности символа Лежандра, следует, что в точности одно из выражений (3.53) является квадратичным вычетом.

Если $3c + 2\sqrt{\delta}$ является квадратичным вычетом, то существуют координаты

$$y_1^{(1),(2)} = \pm\sqrt{\delta\left(3c + 2\sqrt{\delta}\right)};$$

в противном случае существуют

$$y_1^{(1),(2)} = \pm\sqrt{\delta\left(3c - 2\sqrt{\delta}\right)},$$

и тогда либо пары $\left(3c - 2\sqrt{\delta}, \ \pm\sqrt{\delta\left(3c + 2\sqrt{\delta}\right)}\right)$, либо пары $\left(3c + 2\sqrt{\delta}, \ \pm\sqrt{\delta\left(3c - 2\sqrt{\delta}\right)}\right)$, соответственно, являются двумя решениями системы (3.49). Непосредственной проверкой (подстановкой в уравнение (3.45) или в уравнение (3.47) убеждаемся, что данные пары являются решениями уравнений (3.45) и (3.47), следовательно, каждая пара задает координаты некоторой точки данной кривой. Вследствие выполнения (3.49), каждая из двух полученных точек является точкой 4-го порядка. Других решений система (3.49) не имеет, следовательно, на кривой (3.49) существуют ровно две точки четвертого порядка, что и доказывает достаточность условий (3.48).

Теорема 3.6 доказана. ▲

Выше мы отмечали, что кривые с нулевыми значениями параметров a или b обладают плохими криптографическими свойствами и, соответственно,

не используются в криптографических приложениях [29]. Поэтому их следует исключить из наших оценок. Возникает вопрос: как много имеется кривых (3.45) с ненулевыми параметрами a и b, для которых существует изоморфизм с кривыми Эдвардса, при различных значениях порядка поля p? Другими словами, сколько кривых Эдвардса с указанными свойствами существует над простым конечным полем?

3.7.2. Точное число кривых в канонической форме с ненулевыми параметрами, изоморфных кривым Эдвардса

Для определения точного числа эллиптических кривых в форме Вейерштрасса (3.45), имеющих ровно 2 точки 4-го порядка, необходимо обратиться к некоторым результатам теории чисел. Г. Дэвенпорт в своей работе [67] приводит блестящее доказательство распределения квадратичных вычетов, полученное Гауссом. Рассмотрим схему Гаусса и итоги его анализа.

Произведение $n(n+1) \bmod p, \quad n = 1,2,\ldots,p-1$, включает составляющие SS (оба сомножителя – квадратичные вычеты S) с общим числом $(SS), NN$ (оба сомножителя – квадратичные невычеты N) с числом (NN), и смешанные пары SN и NS с числом (SN) и (NS). Гаусс доказал, что имеет место система уравнений:

$$(SS)+(NN)=(p-2-\varepsilon)/2, \text{ где } \varepsilon=(-1)^{(p-1)/2}; \tag{3.54}$$

$$(NS)+(NN)=(p-2+\varepsilon)/2; \tag{3.55}$$

$$(SS)+(NS)=(p-3)/2; \tag{3.56}$$

$$(SN)+(NN)=(p-1)/2; \tag{3.57}$$

$$(SS)+(NN)-(SN)-(NS)=-1. \tag{3.58}$$

Здесь первые 4 уравнения не являются линейно независимыми (суммы первой и второй пар уравнений совпадают), поэтому добавлено 5-е уравнение. Из этой системы легко найти любую из 4-х неизвестных. Комбинируя (3.54) – (3.58), можно получить:

$$(SS)=(p-4-\varepsilon)/4, \quad (SN)=(p-\varepsilon)/4; \tag{3.59}$$

$$(NS)=(NN)=(p-2+\varepsilon)/4,\ \text{где}\ \varepsilon=(-1)^{(p-1)/2}. \qquad (3.60)$$

Очевидно, что сумма $(SS)+(NN)+(SN)+(NS)$ при $n=1,2,\ldots,p-2$, равна $p-2$. Заметим, что при этом $n(n+1)\bmod p\neq 0$.

Для ответа на поставленный вопрос потребуются два результата, которые мы докажем в двух приведенных ниже леммах.

Для произвольного $C\in F_p^{\ *}$ обозначим $F_C(X)=(X^2-C^2)\bmod p$.

Лемма 3.1: *для всех* $X\in F_p^*$ *число разных значений функции* $F_C(X)$, *принадлежащих множеству квадратичных вычетов* Q_p, *равно* (SS), *т.е.*

$$\left|\{F_C(X): X\in F_p^{\ *}\}\cap Q_p\right|=(SS).$$

Доказательство Воспользуемся схемой Гаусса для произведения $n(n-1)\bmod p$. Функцию $F_C(X)$ при $C\neq 0$ представим как $F_C(X)=C^2\left((XC^{-1})^2-1\right)$. Мы видим, что если X пробегает все ненулевые значения $1,2,\ldots,p-1\in F_p^{\ *}$, то $(XC^{-1})^2$ пробегает все значения квадратов из Q_p. В то же время значение функции $F_C(X)$ является квадратичным вычетом в F_p^* тогда и только тогда, когда $\left((XC^{-1})^2-1\right)\bmod p\in Q_p$. Отсюда следует (при $z=(XC^{-1})^2$):

$$\left|\{F_C(X): F_C(X)\in Q_p\}\right|=\left|\{z\in F_p^{\ *}: z\in Q_p\wedge z-1\in Q_p\}\right|=(SS).$$

Лемма 3.1 доказана. ▲

На основе этой леммы и формулы (3.59) мы можем утверждать, что число значений функции $F_C(X)$, являющихся квадратичными вычетами, при любом фиксированном ненулевом $C\in F_p^{\ *}$ равно

$$\frac{p-4-(-1)^{\frac{p-1}{2}}}{4}.$$

Далее, для произвольного $C\in F_p\setminus Q_p$ обозначим

$$G_C(X)=(X^2+C)\bmod p.$$

Лемма 3.2: *для всех* $X\in F_p^*$ *число разных значений функции* $G_C(X)$, *принадлежащих множеству квадратичных вычетов* Q_p, *равно* (NS), *т.е.*

$$\left|\left\{G_C(X): X \in F_p^*\right\} \cap Q_p\right| = (NS).$$

Доказательство. Воспользовавшись тем, что $X \neq 0$ и разделив функцию $G_C(X)$ на X^2, получим равенство

$$CX^{-2} + 1 = G_C(X)X^{-2}.$$

При переборе всех ненулевых значений $X \in F_p^*$ элемент CX^{-2} пробегает все значения квадратичных невычетов из $\overline{Q_p}$. При этом из нашего уравнения следует, что если $G_C(X) \in Q_p$, то и $CX^{-2} + 1 \in Q_P$. Поэтому при $z = CX^{-2}$:

$$\left|\left\{G_C(X): G_C(X) \in Q_p\right\}\right| = \left|\left\{z \in F_p^* : z \in Q_p \wedge z+1 \in Q_p\right\}\right| = (NS) = (NN),$$

согласно (3.60).

Лемма 3.2 доказана. ▲

На основе формулы (3.60) и леммы 3.2 можно утверждать, что число значений функции $G_C(X)$, являющихся квадратичными вычетами, при любом фиксированном ненулевом квадратичном невычете $C \in F_p^*$ равно

$$|G_C(X)| = (NS) = \left(p - 2 + (-1)^{(p-1)/2}\right)/4.$$

Перейдем теперь к вычислению числа кривых с ненулевыми параметрами a и b, изоморфных кривым Эдвардса. Мы исключаем кривые с параметрами $a = 0$ или $b = 0$, так как эти значения параметров порождают криптографически слабые кривые с j-инвариантом, равным 0 или 12^3, соответственно [15]. Случай $a = b = 0$ дает сингулярную кривую и, разумеется, неприемлем.

Теорема 3.7: *число кривых* (3.45) в *форме Вейерштрасса с параметрами* $a \neq 0$ *и* $b \neq 0$ *над полем* F_p *с двумя точками* 4-го *порядка определяется следующими формулами*:

1) *при* $p \equiv 3 \pmod 4$:

$$(\alpha)\quad M_\alpha = \frac{(p-1)(p-7)}{4},\ \textit{если}\ \left(\frac{3}{p}\right) = 1,$$

$$(\beta)\quad M_\beta = \frac{(p-1)(p-3)}{4}\ ,\ \textit{если}\ \left(\frac{3}{p}\right) = -1;$$

2) *при* $p \equiv 1 (\mathrm{mod}\, 4)$:

$$(\gamma)\quad M_\gamma = \frac{(p-1)^2}{4}.$$

Доказательство.

1. Пусть $p \equiv 3 (\mathrm{mod}\, 4)$, тогда (–1) – квадратичный невычет по модулю p [29], т.е. $\left(\frac{-1}{p}\right) = -1$, и тогда утверждение пункта *а*) в условиях (3.48) эквивалентно утверждению

$$\left(\frac{3c^2 + 4a}{p}\right) = 1.$$

Т.е. оба элемента $3c^2 + 4a$ и $3c^2 + a$ поля F_p являются квадратичными вычетами, следовательно, по определению квадратичного вычета, условие (3.48) будет эквивалентно следующему условию:

$$\exists\, A, B \in F_p^* : \begin{cases} 3c^2 + 4a = A^2; \\ 3c^2 + a = B^2. \end{cases}$$

Решая полученную невырожденную систему уравнений над полем F_p (линейных относительно переменных a и c^2), получаем:

$$a = 3^{-1}\left(A^2 - B^2\right), \qquad c^2 = 9^{-1}\left(4B^2 - A^2\right). \tag{3.61}$$

Для кривых с параметрами $a \neq 0$ и $b \neq 0$ квадратичные вычеты $A^2 \neq B^2$ и, кроме того, $4B^2 \neq A^2$ (т.е. нулевые значения a и c^2 отбрасываются), так как из равенств $c = 0$ и $b = -c^3 - ac$ следует равенство $b = 0$. Из (3.48) следует, что $A^2 \neq 0$ и $B^2 \neq 0$. Как видим из (3.61), решение для c существует тогда и только тогда, когда $4B^2 - A^2$ является квадратичным вычетом по модулю p.

Построим квадратную таблицу 3.2 из упорядоченных $\frac{p-1}{2}$ значений всех B^2 (по столбцам) и A^2 (по строкам). В клетки таблицы запишем значения $4B^2 - A^2$ из (3.61), так что на главной диагонали оказываются элементы $3A^2$, которые отбрасываются вследствие условия $A^2 \neq B^2$. Кроме того, в каждой

строке имеем ровно один нулевой элемент, который также отбрасывается. Требуется найти число ν ненулевых недиагональных квадратичных вычетов в строке, при которых существует значение c, согласно (3.61). Общее число таких элементов по всем строкам, очевидно, равно $\mu = \frac{\nu(p-1)}{2}$.

Из доказанной нами леммы 3.2 следует, что число ненулевых квадратичных вычетов в каждой строке таблицы при $p \equiv 3 (\mathrm{mod}\, 4)$ равно $\frac{p-3}{4}$. В таблице 3.2 таких вычетов по два в каждой строке. На главной диагонали со значениями $3A^2$ имеем квадратичные вычеты, если 3 – квадратный вычет в поле, и невычеты в противном случае. Поскольку диагональные элементы отбрасываются, в каждой строке остается $\nu = \frac{p-3}{4} - 1 = \frac{p-7}{4}$ элементов при $\left(\frac{3}{p}\right) = 1$ и $\nu = \frac{p-3}{4}$ при $\left(\frac{3}{p}\right) = -1$. Общее число пар (A, B) по всем строкам таблицы, при которых существует значение c в (3.61), таким образом, равно:

$$\mu_\alpha = \frac{(p-1)(p-7)}{8} \text{ при } \left(\frac{3}{p}\right) = 1 \text{ и } \mu_\beta = \frac{(p-1)(p-3)}{8} \text{ при } \left(\frac{3}{p}\right) = -1.$$

Число эллиптических кривых M_α, M_β с заданными свойствами вдвое больше количества этих пар, так как каждому значению для c^2 отвечают два корня $\pm c$ и, соответственно, два коэффициента кривой $\pm b$. Мы доказали два первых утверждения теоремы 3.7 при $p \equiv 3 (\mathrm{mod}\, 4)$.

Заметим, что условие $\left(\frac{3}{p}\right) = 1$ всегда выполняется при $p \equiv \pm 1 (\mathrm{mod}\, 12)$ [67]. В частности, 3 является квадратичным вычетом при $p = 11, 13, 23, 47$ т.д.

2. Пусть теперь $p \equiv 1 (\mathrm{mod}\, 4)$, тогда число -1 является квадратичным вычетом по модулю p, т.е. $\left(\frac{-1}{p}\right) = 1$ [29].

Тогда утверждение пункта *a*) в условии (4) эквивалентно утверждению

$$\left(\frac{3c^2 + 4a}{p}\right) = -1,$$

а само условие (4) эквивалентно следующему условию:

$$\exists\, A \in \overline{Q_p},\ \exists\, B \in F_p^{\,*} : \begin{cases} 3c^2 + 4a = A, & \left(\dfrac{A}{p}\right) = -1, \\ 3c^2 + a = B^2. \end{cases}$$

Единственное решение данной невырожденной системы (относительно переменных a и c^2) имеет вид

$$a = 3^{-1}\left(A - B^2\right),\ c^2 = 9^{-1}\left(4B^2 - A\right), \tag{3.62}$$

а решение относительно переменных a и c существует тогда и только тогда, когда $4B^2 - A$ является квадратичным вычетом.

Здесь, как видим, переменные a и c^2 не могут принимать нулевые значения. Нам остается лишь найти число квадратичных вычетов в таблице ненулевых значений выражения $\left(4B^2 - A\right)$ при различных A и B^2. Если принять $B^2 = 0$, то в формуле для c^2 мы вновь получим квадратичный невычет в правой части, поэтому и в данном случае учитываем лишь ненулевые элементы A и B^2.

Подобно п.1, построим квадратную таблицу 3.3 из $(p-1)/2$ значений всех квадратичных вычетов B^2 (по столбцам) и невычетов A (по строкам). В клетки таблицы запишем значения $\left(4B^2 - A\right)$ из (3.62), все не равные нулю. Необходимо найти число ν квадратичных вычетов в строке, при которых существует значение c, согласно (3.62), и умножить это значение на число строк.

Из леммы 3.2 следует, что выражение $\left(4B^2 - A\right)$ с ненулевыми квадратичными вычетами B^2 и фиксированным невычетом A принимает $\nu_\gamma = \left(p - 2 + (-1)^{(p-1)/2}\right)/4$ значений на множестве квадратичных вычетов. Это значение равно числу квадратов в каждой строке таблицы, тогда, с учетом того, что $(-1)^{(p-1)/2} = 1$ при $p \equiv 1 (\mathrm{mod}\, 4)$, получаем общее число квадратичных вычетов в таблице:

$$\mu_\gamma = \nu_\gamma (p-1)/2 = (p-1)^2/8.$$

Как отмечалось выше, число кривых M_γ с заданными свойствами вдвое превосходит μ_γ. Итак, теорема 3.7 доказана. ▲

Пример 3.4: приведем примеры построения таблиц, используемых в доказательстве теоремы 3.7, для значений выражений $4B^2 - A^2$ и $4B^2 - A$.

Построим таблицу значений выражения $4B^2 - A^2$ для $p = 11$ (таблица 3.2).

Таблица 3.2. Значения выражения $4B^2 - A^2$, $p = 11$

A^2 \ B^2	1	4	9	5	3
1	3	4	2	8	0
4	0	1	10	5	8
9	6	7	5	0	3
5	10	0	9	4	7
3	1	2	0	6	9

Построим таблицу значений выражения $4B^2 - A$ для $p = 13$ (таблица 3.3).

Таблица 3.3. Значения выражения $4B^2 - A$, $p = 13$

A \ B^2	1	4	9	3	12	10
2	2	1	8	10	7	12
5	12	11	5	7	4	9
6	11	10	4	6	3	8
7	10	9	3	5	2	7
8	9	8	2	4	1	6
11	6	5	12	1	11	3

Следующая таблица (таблица 3.4) иллюстрирует количество кривых, изоморфным кривым Эдвардса (и, соответственно, кривых Эдвардса),

рассчитанное по формулам, доказанным в теореме 3.7, при значениях $p = 7,\ \ 11,\ \ 13, \ldots, 47$.

Таблица 3.4. Количество кривых Эдвардса над полями характеристики p

p	7	11	13	17	19	23	29	31	37	41	43	47
M	6	10	36	64	72	88	196	210	324	400	420	529

По найденным параметрам a и c^2 можно с помощью (1.16) рассчитать параметр d для пары кручения полной кривой Эдвардса

$$d = \frac{3c-2\delta}{3c+2\delta}, \qquad \delta = \pm\sqrt{3c^2 + a}, \tag{3.63}$$

Пример 3.5: требуется найти кривую с двумя точками 4-го порядка над полем F_{11}. Примем, с учетом данных таблицы 3.2, $A^2 = 1, B^2 = 4$. Тогда, согласно (3.61), $c^2 = 9$ – квадратичный вычет в заданном поле, $a = 10$ и $b = \pm c\left(c^2 + a\right) = \pm 2$. Получили пару кривых кручения $y^2 = x^3 + 10x \pm 2$ с порядками $N_E = 8$ и $N_E^t = 16$. Их точки второго порядка $D = (-3,0)$ и $D^t = (3,0)$, а координаты двух точек 4-го порядка первой кривой в соответствии с (3.51), (3.52) равны: $x_1 = 6, y_1 \pm 5$. С помощью (3.63) находим пару кручения изоморфных кривых Эдвардса с уравнением кривой $X^2 + Y^2 = 1 + dX^2Y^2$, которая при $d = 8$ имеет порядок $N_E = 8$, а при $d = 8^{-1} = 7$ – порядок $N_E = 16$.

Вообще нал полем F_{11} существует, как следует из таблицы 3.4, 10 кривых с ненулевыми параметрами a и b и двумя точками 4-го порядка.

Так как общее число всех кривых с ненулевыми a и b, исключая кривые с нулевым дискриминантом, близко к $(p-1)^2$, количество кривых, изоморфных кривым Эдвардса, для больших полей практически равна четверти всех эллиптических кривых, т.е. доля кривых Эдвардса среди всех эллиптических кривых примерно одна четверть.

Формулы (3.61), (3.62) позволяют рассчитывать параметры a и $\pm c$ кривой (и, соответственно, $\pm b$) при заданных значениях пар квадратичных вычетов (A^2, B^2).

На основании условий (3.48) и формул (3.61) – (3.63) можно построить алгоритм вычисления параметров полных кривых Эдвардса [43] .

В следующих двух разделах рассмотрим важнейшее преимущество кривых в форме Эдвардса – их рекордное быстродействие.

3.8. Сложность групповых операций для точек полной кривой Эдвардса в проективных координатах

В работе [2] впервые был дан анализ сложности выполнения групповых операций на кривой в форме Эдвардса в проективных координатах и доказан существенный выигрыш по сравнению с аналогичными операциями на кривой в форме Вейерштрасса. Следуя этой работе, обозначим сложности выполнения полевых операций в поле $\mathbf{F}_q$ как сложности: M – умножения, S – возведения в квадрат, I – инверсии, U – умножения на параметр кривой). Как известно [15], самой трудоемкой операцией арифметики эллиптических кривых является инверсия элементов, оцениваемая порядком $I \cong (10 - 50M)$. Чтобы избавиться от инверсии, при выполнении крилтопротоколов переходят от двумерных аффинных координат к 3-х и даже 4-хмерным координатам, среди которых наиболее распространенными являются проективные координаты. Такой переход практически всегда обеспечивает значительный выигрыш в производительности вычислений [29].

3.8.1. Сложность групповых операций на полной кривой Эдвардса в проективных координатах

1. Сложение точек

Наличие 2-х инверсий в законе сложения (3.2) заставляет обращаться к проективным координатам [2]. Введение третьей координаты используется с целью замены инверсии другими операциями, главным образом, несколькими умножениями. Введем третью координату Z как общий знаменатель в (3.2). Обозначим $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$, тогда гомогениосное уравнение кривой (3.1) в проективных координатах имеет вид

$$(X^2 + Y^2)Z^2 = Z^4 + dX^2Y^2, \quad X = xZ, \quad Y = yZ.$$

Сумма двух точек в проективных координатах $(X:Y:Z)$ теперь записывается как $(X_1 : Y_1 : Z_1) + (X_2 : Y_2 : Z_2) = (X_3 : Y_3 : Z_3)$. С учетом подстановок выразим координаты суммарной точки согласно (3.2):

$$y_3 = \frac{Y_3}{Z_3} = \frac{\left(\frac{X_1Y_2}{Z_1Z_2} + \frac{X_2Y_1}{Z_1Z_2}\right)\left(1 - d\frac{X_1X_2Y_1Y_2}{{Z_1}^2{Z_2}^2}\right)}{\left(1 + d\frac{X_1X_2Y_1Y_2}{{Z_1}^2{Z_2}^2}\right)\left(1 - d\frac{X_1X_2Y_1Y_2}{{Z_1}^2{Z_2}^2}\right)} =$$

$$= \frac{Z_1Z_2\left(Z_1^2Z_2^2 - dX_1X_2Y_1Y_2\right)(X_1Y_2 + X_2Y_1)}{\left(Z_1^2Z_2^2 + dX_1X_2Y_1Y_2\right)\left(Z_1^2Z_2^2 - dX_1X_2Y_1Y_2\right)}.$$

$$x_3 = \frac{X_3}{Z_3} = \frac{Z_1Z_2(Z_1^2Z_2^2 + dX_1X_2Y_1Y_2)(X_1X_2 - Y_1Y_2)}{(Z_1^2Z_2^2 + dX_1X_2Y_1Y_2)(Z_1^2Z_2^2 - dX_1X_2Y_1Y_2)}$$

Обозначим:

$A = Z_1Z_2;\ B = A^2;\ \ C = X_1X_2;\ D = Y_1Y_2; E = dCD;\ \ F = B - E; G = B + E$

Тогда

$Y_3 = A{\cdot}F{\cdot}\ ((X_1 + Y_1){\cdot}(X_2 + Y_2) - C - D)$,

$X_3 = A{\cdot}G{\cdot}(D - C)$,

$Z_3 = F{\cdot}G$.

Подсчет числа элементарных операций здесь дает 10 умножений M, одно возведение в квадрат S и одно умножение на параметр d кривой. Игнорируя простую операцию сложения (вычитания) в поле, находим сложность вычисления суммы различных точек через сложности полевых операций $V_E = 10M + 1S + 1U$. Заметим, что сложность возведения в квадрат оценивается приблизительно как $1\mathrm{S} \cong \frac{2}{3}\mathrm{M}$ [2].

2. Удвоение точек

Используя уравнение кривой (3.1), закон удвоения (3.3) запишем в форме, не зависящей от параметра d

$$2(x_1, y_1) = \left(\frac{x_1^2 - y_1^2}{2 - x_1^2 - y_1^2}, \frac{2x_1y_1}{x_1^2 + y_1^2},\right).$$

Тогда координаты точки удвоения огласно (3.3):

$$x_3 = \frac{X_3}{Z_3} = \frac{\left(\left(\frac{X_1}{Z_1}\right)^2 - \left(\frac{Y_1}{Z_1}\right)^2\right)\left(\left(\frac{X_1}{Z_1}\right)^2 + \left(\frac{Y_1}{Z_1}\right)^2\right)}{\left(2 - \left(\frac{X_1}{Z_1}\right)^2 - \left(\frac{Y_1}{Z_1}\right)^2\right)\left(\left(\frac{X_1}{Z_1}\right)^2 + \left(\frac{Y_1}{Z_1}\right)^2\right)} = \frac{(X_1^2 - Y_1^2)(X_1^2 + Y_1^2)}{(2Z_1^2 - X_1^2 - Y_1^2)(X_1^2 + Y_1^2)},$$

$$y_3 = \frac{Y_3}{Z_3} = \frac{2\frac{X_1}{Z_1}\frac{Y_1}{Z_1}\left(2 - \left(\frac{X_1}{Z_1}\right)^2 - \left(\frac{Y_1}{Z_1}\right)^2\right)}{\left(2 - \left(\frac{X_1}{Z_1}\right)^2 - \left(\frac{Y_1}{Z_1}\right)^2\right)\left(\left(\frac{X_1}{Z_1}\right)^2 + \left(\frac{Y_1}{Z_1}\right)^2\right)} = \frac{2X_1Y_1(X_1^2 + Y_1^2)}{(2Z_1^2 - X_1^2 - Y_1^2)(X_1^2 + Y_1^2)}$$

Обозначим

$$A = X_1^2, B = Y_1^2, C = Z_1^2, D = (A + B), E = (A - B), F = 2C - A - B.$$

Тогда

$$X_3 = D \cdot E,$$

$$Y_3 = 2X \cdot Y \cdot F,$$

$$Z_3 = D \cdot F.$$

Подсчет числа возведений в квадрат и умножений в поле дает суммарную сложность удвоения $T_{E1} = 4M + 3S$. Так как возведение в квадрат проще умножения (приблизительно на треть), можно модифицировать вычисления, представив $2X_1Y_1 = (X_1 + Y_1)^2 - X_1^2 - Y_1^2$. Обозначим далее $G = (X_1 + Y_1)^2$, $H = G - D$. Теперь координата $Y_3 = H \cdot F$ вычисляется не двумя, а одним умножением, но появляется еще одно возведение в квадрат в величине G. Итак, обменяв одно умножение на возведение в квадрат, получаем сложность удвоения точки на кривой Эдвардса $T_E = 3M + 4S$. Она почти вдвое меньше соответствующей сложности удвоения на кривой Вейерштрасса.

3.8.2. Сложность групповых операций на полной кривой Эдвардса в инвертированных проективных координатах

Авторы [5] предложили взамен классических проективных координатах использовать оригинальные инвертированные проективные координаты, что позволяет сократить одну операцию $1M$ при сложении точек кривой (3.1).

Введем формальную замену координат $x = \frac{Z}{X},\ y = \frac{Z}{Y}$, тогда уравнение (3.1)

$$\frac{Z^2}{X^2} + \frac{Z^2}{Y^2} = 1 + \frac{dZ^4}{X^2Y^2}$$

можно переписать в форме проективной кривой с инверсией координат

$$Z^2(X^2 + Y^2) = X^2Y^2 + dZ^4, \quad XYZ \neq 0. \tag{3.64}$$

Исключительными точками этого уравнения, как видим, являются 4 точки кривой (3.1) с нулевыми координатами (0,±1) и (±1,0) четного порядка и O. Их появление в законе сложения (3.2) точек для кривой (3.64) ведет к сбою программы вычислений, поэтому они должны быть исключены. В криптосистеме используются точки G нечетного простого порядка n, в подгуппе которых имеется лишь одна исключительная точка O = (1,0). Поэтому для вычисления скалярного произведения $kG, k \neq n$, использование инвертированных проективных координат допустимо. Напротив, для проверки порядка точки их использование проблематично (но не невозможно).

Рассмотрим оценки сложности групповых операций для кривой (3.64).

1. Сложение точек

Подстановка $x = \frac{Z}{X},\ y = \frac{Z}{Y}$ в формулу сложения точек (3.2) при выполнении условия (3.64) для пары точек $X_{1,2}Y_{1,2}Z_{1,2} \neq 0$ дает равенство

$$\left(\frac{Z_1}{X_1}, \frac{Z_1}{Y_1}\right) + \left(\frac{Z_2}{X_2}, \frac{Z_2}{Y_2}\right) = \left(\frac{(Y_1Y_2 - X_1X_2)Z_1Z_2}{X_1X_2Y_1Y_2 - dZ_1^2Z_2^2}, \frac{(X_1Y_2 + X_2Y_1)Z_1Z_2}{X_1X_2Y_1Y_2 + dZ_1^2Z_2^2}\right) = \left(\frac{Z_3}{X_3}, \frac{Z_3}{Y_3}\right).$$

Отсюда

$$X_3 = (Y_1Y_2 - X_1X_2)(X_1X_2Y_1Y_2 - dZ_1^2Z_2^2),$$

$$X_3 = (X_1Y_2 + X_2Y_1)(X_1X_2Y_1Y_2 + dZ_1^2Z_2^2),$$

$$Z_3 = (Y_1Y_2 - X_1X_2)(X_1Y_2 + X_2Y_1)Z_1Z_2.$$

Обозначим:

$$A = Z_1Z_2, B = dA^2,\ C = X_1X_2,\ D = Y_1Y_2,\ E = CD,\ F = C - D,$$

$$G = (X_1 + Y_1)(X_2 + Y_2) - C - D.$$

Тогда

$$X_3 = (E + B)F,$$

$$Y_3 = (E - B)G,$$

$$Z_3 = A \cdot F \cdot G.$$

Подсчет числа полевых операций M, S и U в групповой операции сложения точек в инвертированных проективных координатах дает суммарную сложность $V_{EI} = 9M + 1S + 1U$. По сравнению с проективными координатами получаем экономию на одну операцию $1M$. Это рекордно низкая сложность для сложения точек [5].

1. Удвоение точек

Воспользуемся теперь формулой удвоения (3.3), тогда подстановка $x = \frac{Z}{X},\ y = \frac{Z}{Y}$ приводит к равенству

$$2\left(\frac{Z_1}{X_1}, \frac{Z_1}{Y_1}\right) = \left(\frac{(Y_1^2 - X_1^2)Z_1^2}{X_1^2Y_1^2 - dZ_1^4}, \frac{2X_1Y_1Z_1^2}{X_1^2Y_1^2 + dZ_1^4}\right) = \left(\frac{Z_3}{X_3}, \frac{Z_3}{Y_3}\right).$$

С учетом уравнения кривой (3.64), можно записать знаменатели координат в более простой форме

$$2\left(\frac{Z_1}{X_1}, \frac{Z_1}{Y_1}\right) = \left(\frac{(Y_1^2 - X_1^2)}{X_1^2 + Y_1^2 - 2dZ_1^2}, \frac{2X_1Y_1}{X_1^2 + Y_1^2}\right) = \left(\frac{Z_3}{X_3}, \frac{Z_3}{Y_3}\right)$$

Из этого равенства имеем

$$X_3 = 2X_1Y_1(X_1^2 + Y_1^2 - 2dZ_1^2),$$

$$Y_3 = (X_1^2 + Y_1^2)(Y_1^2 - X_1^2),$$

$$Z_3 = 2X_1Y_1(Y_1^2 - X_1^2).$$

Пусть

$$A = X_1^2,\ B = Y_1^2, C = A + B,\ \ D = A - B,\ \ E = (X_1 + Y_1)^2 - C$$

Тогда $X_3 = E\cdot(C - 2dZ_1^2),\ Y_3 = C\cdot D,\ Z_3 = D\cdot E.$

Итак, суммарная сложность удвоения точки в инвертированных проективных координатах оценивается величиной $T_{EI} = 3M + 4S + 1U$. Это на одну операцию $1U$ больше, чем в классических проективных координатах. Так как оперрация удвоения при экспоненцировании точки выполняется в среднем вдвое чаще, чем сложение, выигрыш при сложении будет скомпенсирован проигрышем при удвоении.

3.8.3. Сложность групповых операций на кривой Вейерштрасса

1. Сложение точек

Обратимся теперь к эллиптической кривой в форме Вейерштрасса над полем F_p

$$W: \qquad y^2 = x^3 + ax + b,$$

с законом сложения различных точек [28]

$$(x_1, y_1) + (x_2, y_2) = \left(\left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2\ , -y_1 - \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x_3 - x_1)\ \right)$$

В проективных координатах с заменой $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$ имеем

$$\frac{\frac{Y_2}{Z_2} - \frac{Y_1}{Z_1}}{\frac{X_2}{Z_2} - \frac{X_1}{Z_1}} = \frac{u}{v}.\quad u = Y_2Z_1 - Y_1Z_2\ ,\qquad v = X_2Z_1 - X_1Z_2.$$

Тогда

$$\frac{X_3}{Z_3} = \left(\frac{u}{v}\right)^2 - \frac{X_1}{Z_1} - \frac{X_2}{Z_2} = \frac{Z_1Z_2u^2 - v^2(X_1Z_2 + X_2Z_1)}{Z_1Z_2v^2} = \frac{vg}{Z_3},$$

где

$$Z_3 = Z_1 Z_2 v^3, \qquad g = Z_1 Z_2 u^2 - v^3 - 2v^2 X_1 Z_2.$$

Далее

$$\frac{Y_3}{Z_3} = -\frac{Y_1}{Z_1} + \left(\frac{u}{v}\right)\left(\frac{X_1}{Z_1} - \frac{vg}{Z_1 Z_2 v^3}\right) = \frac{-Y_1 Z_2 v^3 + u(X_1 Z_2 v^2 - g)}{Z_1 Z_2 v^3}$$

Итак, окончательно получаем:

$$X_3 = vg,$$
$$Y_3 = -Y_1 Z_2 v^3 + u(X_1 Z_2 v^2 - g),$$
$$Z_3 = Z_1 Z_2 v^3.$$

Подсчет числа операций в этмх выражениях дает сложность вычисления суммы точек канонической кривой W: $V_W = 12M + 2S$. Аналогичный расчет для удвоения точек приводит к результату [29] $T_W = 7M + 5S$.

3.9. Сравнительный анализ быстродействия экспоненцирования точки для кривых в форме Эдвардса и Вейерштрасса

Данный анализ был нами впервые проведен в работе [31]. Наиболее значимой и трудоемкой операцией во всех известных криптопротоколах является скалярное произведение kP точки P (или экспоненцирование точки в терминах мультипликативной группы). Другие операции, выполняемые в поле F_p, занимают незначительную долю вычислительных ресурсов при выполнении протокола. Поэтому оценка выигрыша в производительности экспоненцирования точки кривой может быть принята как грубая оценка соответствующего выигрыша для протокола в целом.

Приведем сложности всех полевых операций к сложности умножения M. Принимая вычислительную сложность возведения в квадрат 1S = 0.67M [2], а умножения на параметр кривой $1U$ = 0.5M, получим оценки сложности сложения и удвоения на кривой Эдвардса $V_E = 10M + 1S + 1U = 11{,}17M$, $T_E = 3M + 4S = 5.67M$. Удвоение в проективных координатах, как видим, практически вдвое быстрее сложения. При использовании инвертированных

проективных координат получаем $V_{EI} = 9M + 1S + 1U = 10{,}17M$, $T_{EI} = 3M + 4S + 1U = 6.17M$. Для эллиптической кривой в форме Вейерштрасса имеем, соответственно, $V_W = 12M + 2S = 13.33M$, $T_W = 7M + 5S = 10.33M$.

При вычислении скалярного произведения kP точки P число k представляется в двоичной форме, тогда работает алгоритм последовательного сложения-удвоения (схема Горнера) [29]. Пусть ν_1 – относительная частота знаков «1» в двоичной последовательности числа k. В схеме экспоненцирования точки на каждом шаге выполняется удвоение, и лишь при знаке «1» этой последовательности выполняется сложение точек. Тогда в общей форме выигрыш в производительности вычисления скалярного произведения на кривой Эдвардса в сравнении с тем же вычислением на кривой в форме Вейерштрасса равен

$$\gamma(v_1) = \frac{T_W + v_1 V_W}{T_E + v_1 V_E} = \frac{10.33 + 13.33\, v_1}{5.67 + 11.17\, v_1}. \qquad (3.65)$$

Для вычислений в инвертированных проективных координатах получаем

$$\gamma(v_1) = \frac{T_W + v_1 V_W}{T_{EI} + v_1 V_{EI}} = \frac{10.33 + 13.33\, v_1}{6.17 + 10.17\, v_1}. \qquad (3.66)$$

Функция $\gamma(v_1)$ является монотонно-убывающей от максимального значения $\gamma(0) = 1.83$ до минимального $\gamma(1) = 1.41$ (для формулы (3.65)). Наиболее вероятным является среднее значение плотности знаков «1» $v_1 = 0.5$, при этом средний выигрыш вычислений в проективных координатах равен $\gamma_{\text{ср}} = 1.51$, в инвертированных проективных координатах он дает то же значение $\gamma_{\text{ср}} = 1.51$. Как уже отмечалось, выигрыш в сложности при сложении точек компенсируется проигрышем при их удвоении.

Если использовать вместо двоичного представления числа k произведения kP троичное NAF(k) $k_i \in \{0,1,-1\}$[29], то можно снизить среднее число ненулевых компонент в числе k до 1/3. По формуле (3.65) при $v_1 = 1/3$ это даст максимальное значение среднего выигрыша в проективных координатах $\gamma_{\text{ср}} = 1.574$, а в инвертированных проективных координатах

$\gamma_{\text{ср}} = 1.545$. Этот результат понятен, так как преобладание невыгодных удвоений снижает эффективность инвертированных проективных координат.

Заметим, что приведенные результаты относительно нижней границы $\gamma(v_1)$ в некоторой степени условны, так как мы приняли сложность умножения на параметр кривой $1U$ = 0.5M. В частных случаях параметр d, может принимать малые значения, тогда величиной $1U$ вообще можно пренебречь. В таком случае знаменатель в (3.65) равен $(5.67 + 10.67\, v_1)$, при этом средний выигрыш $\gamma_{\text{ср}} = 1.54$. Соответственно, при троичном NAF(k) представлении числа k средний выигрыш достигает значения $\gamma_{\text{ср}} = 1.6$. Для инвертированных координат знаменатель в (3.66) становится равным $(5.67 + 9.67\, v_1)$, при этом двоичное кодирование дает средний выигрыш $\gamma_{\text{ср}} = 1.62$, а троичное – $\gamma_{\text{ср}} = 1.66$.

Одним из выводов данного раздела является неэффективность в общем случае инвертированных проективных координат при экспоненцировании точки. Вычисления в них, кроме того, требует обходить точки (±1,0), (0,±1). Однако при малых значениях параметра d они обеспечивают рекордное значение среднего выигрыша $\gamma_{\text{ср}} = 1.66$.

Таким образом, наряду с более эффективным и лаконичным программированием, технология полных кривых Эдвардса позволяет в 1.5 – 1.6 раза ускорить выполнение криптопротоколов (а значит, сэкономить пропорциональное количество денег) в сравнении с теми кривыми, которые сегодня рекомендуются устаревшими стандартами [см.15,29] . Как известно [15], в них рекомендуются кривые, параметры которых рассчитывались еще в 1997 году (в нашей стране – в 2001году). Хочется надеяться, что столь очевидная экономическая выгода наряду с требованиями безопасности ускорят стандартизацию в нашей стране технологии кривых Эдвардса для решения задач асимметричной криптографии.

3.10. Вычисление общесистемных параметров криптостойких полных кривых Эдвардса

Целевой задачей исследования свойств кривых в форме Эдвардса является внедрение этой несомненно эффективной технологии в проекты новых

стандартов эллиптической криптографии. Существующие стандарты [29] рекомендуют, как правило, набор кривых над конечными полями в широком диапазоне требований по их криптостойкости. Например, первый из таких стандартов – национальный стандарт США FIPS-186-2-2000 [78] – рекомендует 5 кривых в форме Вейерштрасса над простыми полями с длинами модулей от 192 до 521 бита . Мы в данном разделе взяли за основу те же поля и решили задачу вычисления общесистемных параметров криптостойких полных кривых Эдвардса порядка $4n$ с простым порядком n точки G – генератора криптосистемы. Эти результаты опубликованы в работе [37] и обсуждались на конференциях [61,62].

Поиск кривых Эдвардса, приемлемых для криптографии, представляет собой нетривиальную задачу. Ключевым моментом в ней является расчет порядка кривой, заданной над конечным полем. В данной работе поставлена задача поиска полных кривых Эдвардса с почти простым значением порядка $4n$ над большими простыми полями. В разделе 3.10.1 мы обсуждаем содержание проблемы определения порядка кривой в форме Эдвардса и кратко описываем возможные пути и алгоритм ее решения. В разделе 3.10.2 мы приводим методы расчетов и наши результаты вычислений общесистемных параметров 40 кривых Эдвардса над простыми полями $\mathbf{F}_p$ с модулями p длиной 192, 224, 256 и 384 бит [37]. Порядок $4n$ предложенных кривых содержит простой сомножитель n, близкий по величине к величине соответствующего поля. Таким образом, найденные кривые удовлетворяют современным требованиям к порядку генератора криптосистемы и с успехом могут применяться на практике и в проектируемых криптосистемах.

В работе [2] доказано, что для любой кривой, записанной в форме (3.1), найдется изоморфная эллиптическая кривая в канонической форме над полем $\mathbf{F}_p$. Однако, в известных стандартах шифрования на эллиптических кривых [28] не содержится кривых над простыми полями с кофактором порядка, равным 4. Это не позволяет преобразовать рекомендуемые современными стандартами кривые непосредственно в форму (3.1). В этой связи для криптографических приложений следует провести поиск кривых Эдвардса над простыми полями с приемлемым значением порядка $4n$.

3.10.1. Алгоритм поиска криптостойкой полной кривой Эдвардса на базе изоморфной кривой в форме Вейерштрасса

В разделах 1.8 и 3.7 мы уже приводили алгоритмы поиска криптостойкой полной кривой Эдвардса с вводом нового параметра c в уравнение канонической кривой. Здесь мы интегрируем эти алгоритмы в более общий алгоритм, использующий прямые формулы изоморфного преобразования кривой в форме Вейерштрасса в кривую Эдвардса. Как нам уже известно из главы 1, для каждой кривой (3.1) в форме Эдвардса E найдется изоморфная ей кривая в форме Вейерштрасса W вида

$$W: \qquad v^2 = u^3 + au + b. \tag{3.65}$$

Соответствующий изоморфизм между точками кривых E и W задается рациональными функциями [10]:

$$u = \frac{(5-d)+(1-5d)y}{12(1-y)}, \quad v = \frac{(1-d)+(1+y)}{4x(1-y)}, \quad (y-1) \neq 0. \tag{3.66}$$

Хотя они сложней формул (1.20) для кривых в форме Монтгомери, они задают прямой изоморфизм ($W \to E$) без использования промежуточной формы Монтгомери ($W \to M \to E$).

Четыре точки пересечения с осями координат преобразуются следующим образом:

$(x, y) = (0, 1) \to (u, v) = O,$

$(x, y) = (0, -1) \to (u, v) = \left(\frac{1+d}{6}, 0\right)$ при $x = 0.$

$(x, y) = (\pm 1, 0) \to (u, v) = \left(\frac{5-d}{12}, \pm\frac{1-d}{4}\right)$ при y$= \pm 1$

Коэффициенты кривой W выражаются через параметр кривой E следующим образом [10]:

$$a = -\frac{(1+14d+d^2)}{48}, \qquad b = -\frac{(1-33d-33d^2+d^3)}{864}. \tag{3.67}$$

Для обратного преобразования справедливо:

$$x = \frac{6u-(1+d)}{6v},$$

$$y = \frac{12u+d-5}{12u+1-5d},$$

при $6v(12u + 1 - 5d) \neq 0$,

$(u, v) = \left(\frac{1+d}{6}, 0\right) \rightarrow (x, y) = (0, -1)$, при $v = 0$,

$(u, v) = O \rightarrow (x, y) = (0, 1)$. (3.68)

Одним из способов расчета порядка кривой Эдвардса является адаптация соответствующих методов нахождения порядка канонических эллиптических кривых (таких как алгоритмы Скуфа, SEA, Satoh). Используя соотношения (3.66) – (3.68), для кривых в форме Эдвардса определяется последовательность полиномов деления [10], с помощью которых методом SEA может быть вычислен порядок рассматриваемой кривой Эдвардса. С другой стороны, подсчитать порядок кривой Эдвардса можно посредством изоморфного перехода к канонической форме с последующим нахождением порядка кривой по известным алгоритмам.

Второй сценарий был использован для поиска кривых над простыми полями, приведенных в разделе 3.10.2. Выбрав произвольно параметр $d \neq A^2$ в поле F_p и, используя формулы (3.67), получим изоморфную эллиптическую кривую в форме Вейерштрасса. Заметим, что при заданном ограничении на параметр d кривой, кубика правой части уравнения (3.65) будет иметь единственный корень c, и будет выполняться также условие (ii) существования двух точек 4-го порядка (3.48) теоремы 3.6. Порядок N_E рассматриваемой кривой считаем приемлемым, если число $n = N_E/4$ простое, лежащее приблизительно в пределах 180 – 600бит. Такая кривая может быть рекомендована к применению в криптопротоколах.

Для построения криптографической системы на полученной кривой Эдвардса необходимо определить генерирующую точку G порядка n. Задавая произвольно координату x и вычисляя из уравнения кривой (3.1) значение y получим произвольную точку Q кривой Эдвардса. Если $x \neq 0$ и $x \neq \pm 1$ (вероятность этого события ничтожно мала), порядок точки Q может быть равен $n = N_E/4$, $2n$ или $4n = N_E$. Тогда генератором криптосистемы будет

точка Q, $2Q$ или $4Q$ соответственно. В разделе 3.5 мы предложили метод нахождения генератора с помощью одного удвоения точки Q.

3.10.2. Общесистемные параметры криптостойких полных кривых Эдвардса

В данном разделе мы рассматриваем простые поля с модулями

$$p_{192} = 2^{192} - 2^{64} - 1,$$
$$p_{224} = 2^{224} - 2^{96} + 1,$$
$$p_{256} = 2^{256} - 2^{224} + 2^{192} + 2^{96} - 1,$$
$$p_{384} = 2^{384} - 2^{128} - 2^{96} + 2^{32} - 1,$$

рекомендуемые стандартом FIPS – 186 – 2 – 2000 [78], и приводим перечень кривых в форме Эдвардса почти простого порядка $N_E = 4n$(n – простое) над каждым из полей. Данные приведены в таблицах 3.5 – 3.8. Наряду с этим, в таблицах также содержатся общесистемные параметры для реализации шифрования с помощью кривой Эдвардса, а именно, порядок $n = N_E/4$ и координаты (x_G, y_G) генератора G криптосистемы для каждой кривой.

В каждой из приведенных ниже таблиц содержится по 10 кривых Эдвардса над соответствующим полем с параметрами d различной битовой длины. Порядок кривых сравним по длине с длиной рассматриваемого поля. Расчеты производились посредством прикладных программ, основанных на использовании функций библиотеки MIRACL.

Таблица 3.5. Кривые Эдвардса почти простого порядка над полем с модулем p_{192}

$\boldsymbol{p} =$	**FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFEFFFFFFFFFFFFFFFF**
$d =$	28453E
$n =$	4000000000000000000000001AEAD1229D137F564D7FF6D5
$x_G =$	AE709D07B2D112CECD4A7AE103757F2C101D054ACB1A0F17
$y_G =$	8BDF8D5A994AAB1E1455889D28E29CA3EC68548D3F7CA32F
$d =$	6DBA6A

$n =$	3FFFFFFFFFFFFFFFFFFFFFFFEA75D4027230DD4DFFDB0455
$x_G =$	44F083BB00E51AD91A2743284D31F57EE5C84826FCC91F4B
$y_G =$	15FC16E5870524E0DBBE9EC8BB9F066C02A02B1978D4E029
$d =$	CE37DC
$n =$	400000000000000000000000002618428A483A133570834389
$x_G =$	508AEB91AB9A230C377BD82BDDBE5B75CD38CDF1DA407A3C
$y_G =$	6E12D78C595D0ECC1614BFD04616CE45D6BB7D8607BB2ED7
$d =$	111DB4A
$n =$	40000000000000000000000004F2EA0DD45656E7E6066E8C9
$x_G =$	1603A0943674DA7D8476608E764BA7A63DBEBE4D8E549165
$y_G =$	148D5DA4654152C9196699074374EE091FE789A0CC7EC60
$d =$	12CCB98
$n =$	4000000000000000000000006AF187D9A93890273705F7E9
$x_G =$	6AC4FE0FEF645CA5754BFB0BD05967A418FCCAAF20E473A9
$y_G =$	B2F122E8816B142765B0FABCD7834CDC2D1AAC7EA779750B
$d =$	BFFAB4397C01049E12A46027E7E14D7A666240E6831D926
$n =$	400000000000000000000000022F2FB1AD838D6680571D7BB
$x_G =$	B46D3402D2A7F4CE30C5AA7839B0CE8D957240684824C726
$y_G =$	F0B24C20880DE704123ADF2A44189480DDFFDEEAF002B238
$d =$	CAE5B27FA36F62C6EC5EEC6ABE086C2BFF6A9F029CE42C88
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFBD410EA5F77DD3ECCBD0FC45
$x_G =$	72F6A4BADEB19F876258926630ECA22F1D9FEF34CE53440C
$y_G =$	852D88019C7A20987BEE784A9EC09FF213639E59A3C3A51E
$d =$	2FDE6BF890C01A5FEB8B8D976676DDAC8C3349FE0C89959A
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFC9B03E26E57C1A34245E23ED
$x_G =$	F2FA1F9E30FDF9C5CD1E84D287CE5F2DE9283B077C2C83FA
$y_G =$	419A50AF660CBF77699FBECBF7181F7F15C0DCF31523171F
$d =$	83E48CF4C49FDDEF6E5D812D3FBD06054835C85D6DD283BC

$n =$	4000000000000000000000002640829409C3B60282409D8B
$x_G =$	5BCFE7F1248E5A43CDF179D95334FE45061F0ECB599020B2
$y_G =$	FDEB479483DB5D1743AE7496B9A32B5CF774B8D8CD13272F
$d =$	CFC522DAB7BE9E92FA78DDA10CE941CC7108A299FB4A79C9
$n =$	4000000000000000000000006B362B3138DB79B9866A5BD7
$x_G =$	BE321DFF94A4E7DA940394E9C9EC1F2BC4B29F93302BDA3B
$y_G =$	B5E1CE7F685CEFE82284E5027404FE6E8D3C11E7D37092F3

Таблица 3.6. Кривые Эдвардса почти простого порядка над полем с модулем p_{224}

$\boldsymbol{p} =$	**FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF000000000000000000000001**
$d =$	3608425
$n =$	4000000000000000000000000020BBEC47CEDB34DD05BCB6B7E619
$x_G =$	C448CA02660F57204FF1BDE2B5CC3E25606A7460399FEA3DA9A06383
$y_G =$	319117770D6FC7FE35F6A02905FE1F363156BD2E5B75BB89A64CAFAB
$d =$	42E5CF5
$n =$	400000000000000000000000070D1AD037FD1F2585B37C3CD8E75
$x_G =$	74E8DA9676A0AA64C73460EFBA56F04A5D69FB39DC03A0D53B9E99A9
$y_G =$	A70D758DE2B7CADCEED9D6315F44987AF89B9D1D3BB62B54574971D5
$d =$	148CD57
$n =$	40000000000000000000000001BE5276AAC300270278233024CF9
$x_G =$	EEA0CBD1BFB937E1C83C12E90200429DD3F74854256DA5E249B04A2B
$y_G =$	512F075B98B7281E07BDABF68CE2BF4A30200DB1A9029ABDAAAEDEEA
$d =$	2C4C61C
$n =$	400000000000000000000000054547A5F0105F649731ECC684CF
$x_G =$	142954F6B9702D136634AD2F6574ED5CF24E2D7D8AEB0B855105451A
$y_G =$	6EB430B9BE6B0A78D41BCBCFC3A207D0A813C8AB052BFD23529D23E9
$d =$	2DD96A9
$n =$	400000000000000000000000024E3FE6372F8DB77945FFF06024D

$x_G =$	C62E8C8EB908FFC6DAE6D79D335618D62855B9A14F49DF1DEE0A414E
$y_G =$	6955A4D991CFE25E6EE21B64FC61763C16634CA234222A0E752F4E85
$d =$	1968A1D7A81AF44243D0F60F186A519FB95BF7C2FE2C037798D70FE9
$n =$	400000000000000000000000000003FCA1753F6F54C2A869DC8C5DBD9
$x_G =$	6F39E721F18128534C6C0339A982011683D9676565CAFD942F8C7761
$y_G =$	2E5ECC135081FE6FF9E0F7020BBC7EFD9DFB18A47B81DB825FC26817
$d =$	60B4BE2A9399C390AA051FC8B4706856E1CE2A03791C1B5A2C88DD40
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFFFFF5722B78ED1CA9B003572246F325
$x_G =$	60D6B139288C1B171E8FE28EC068130CABDA0A1A48C1DF7280A14211
$y_G =$	198E5F9E549D6C080CCCA7886E6BC880623B5AB1070CB62A81377BC0
$d =$	E3850F8A3D93759F172493E9339B29DB7D345BE57CFCDCE545EEB1DC
$n =$	400000000000000000000000000005B3B3E37B9EDE0A1B45BEC627F5B
$x_G =$	DB921A45CBF9979A3425F0556165DEF73ED6792034A9D9A5DC180FC9
$y_G =$	A3A87DA7A0307AD196B2CEC840C3E17147C570EBE9CC432DBF941CD7
$d =$	6E6607C2DCE3392468455ED0DBFBA5442AD87C7B8CE484C1B85D592
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFFFFFE623254CA68DF10B09B50D57327B
$x_G =$	96C7C00D85C8BB5B53F74E6EF38EFEEB458FE9434290046114E0E906
$y_G =$	A2410947B7C669193EC0B39001FC786B0558C3DBB6B64BC9677A7016
$d =$	2BFE721565BE78DE5DABD588BCE5E62C7127A0BEC33A2FF844E0B79D
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFFFFCCA2B4C227A60815F6A6359225E9
$x_G =$	DA2B86A81DBF8B887D3A3552A26CF3679DC73C300998BD82EEC1F49C
$y_G =$	D817BCDA64B37A31B2A15975F472A18F8D226446D1473CB3FB38D7C

Таблица 3.7. Кривые Эдвардса почти простого порядка над полем с модулем p_{256}

$\boldsymbol{p} =$	**FFFFFFFF00000001000000000000000000000000FFFFFFFFFFFFFFFFFFFFFFFF**
$d =$	72A38
$n =$	3FFFFFFFC00000003FFFFFFFFFFFFFFFBA76FA29C30CC3AA4954B53EDBE54D75
$x_G =$	894F8283626AEE6848515DDDC3B8DBB3D5302DEE0EE75080D6753E4D39BA5AB2

$y_G =$	EA612346223F6480CBBAFA39DB95D54D21469DD3074A957EFDA4FD79FEB630B5
$d =$	2EBFA9
$n =$	3FFFFFFFC0000000400000000000000005EFF905BA96F95CE79513CBE0CC53D1F
$x_G =$	363D655BF3F221256F032FC791B06149C14ACFFD92B59C84D1D3B817A9E622D2
$y_G =$	A52438C53DDCA661685B1F235EF0F1D280A493C33153AD691097AEC67A62C564
$d =$	805294
$n =$	3FFFFFFFC000000040000000000000051814C8E8360B7C96A8419F38B8039F1
$x_G =$	EEBABE42482F67FECF7F9D2D49A4430372D7678CF7EDAB4B9184D42BD93F390C
$y_G =$	E2A059E8C776F46028BA9265E20C09A785ECEEFB162562DA2AF78BC412D2D3CC
$d =$	9855C9
$n =$	3FFFFFFFC000000040000000000000002C1564946E895DAE0EB7EB501C62C62F
$x_G =$	D139CE35F84CFA17E8ED28E083F07C708303AE788477118C5AFE86D313D443BD
$y_G =$	F5F4A5309EEA75820AEE714A8D99CD22CE2539E73D2C6688B8480E18F1D384D
$d =$	BE957B
$n =$	3FFFFFFFC000000040000000000000005DACD9D621DE8444CB5625A8193E1D81
$x_G =$	B1550150B88C76AEC1CD0B0EC2008D1D73D086A0FD63103A0875EF574F0FF113
$y_G =$	306021DA97347B1DA05C98CE858B5C69ED901187DB03F68C399B694003FACA96
$d =$	4A5084ABF5DFFAC393E29A8BF045F4AA94C80F55414AFDAB8517CE769130DC82
$n =$	3FFFFFFFC00000003FFFFFFFFFFFFFFF9B91248CDD0CAF813BD0F8B9F32C892B
$x_G =$	96540A02F26ED385B606A7956E5BA33B8D5A9E47FCE718C752562E2F2A6891E2
$y_G =$	ACDE96C64793B77B72DC71EB12508AAC629C4CCAA281BF164B73030EF382D9DE
$d =$	340B9C82867CD106FA7E1B11E9415A1099BB48C9A28F3B21A77E70C20E528839
$n =$	3FFFFFFFC00000003FFFFFFFFFFFFFFFFBE08C2E4E3DDB90FBAB306B69633071
$x_G =$	490659C09655C82EECDC109AC5F29E85DC94882E30916E9F21273AFE9D1B01AC
$y_G =$	9BB0A4609F48A8029AD9014BE7184E36A120B5789799DE52F4E7D8D90386F786
$d =$	32BBA52848792541A79779B4CFFF035903D2112814334C1BE3556A670A9D73D1
$n =$	3FFFFFFFC00000003FFFFFFFFFFFFFFFC0285F94E407693E8624559A0D329309
$x_G =$	9D39DB8088196DE49C37D787D64EEB9ECA6A762BE77D0FCBD0430DCCBB057C2E
$y_G =$	62ABD075DA7C26034514079E7B5000D4493BC4318C5A3E3856FAAB8724C5C0F5
$d =$	8884FA14412BC9B62B0C467262DA6BD8F529C01AAD09545B2A294D479B254B2F
$n =$	3FFFFFFFC00000003FFFFFFFFFFFFFFFC7D37AFE15BFD6994B8E6427BA368D27
$x_G =$	D5D245310DD89B88DFF59C2CAEC8DBE62000E53D0BB4051247B97DDF420158C

$y_G =$	444B435A604729B76A17E8F4B93A47DC62D7777C1B056DB727675D627CF749F1
$d =$	E49478138FFE7946F81C3AC55BBBE4D41B4DF660A3B118B0075676506425D9FA
$n =$	3FFFFFFFC0000000400000000000000300FC8D622E987518514DFF16654AAFB
$x_G =$	92967B57C8CAB56FBDE3C3B0568C336BE2FBBB875F3983281D38B173C80971B9
$y_G =$	F3270D7C8A827B913A0B10079555D7D2E3A1C64578A49A77A9C1E6B45C186E90

Таблица 3.8. Кривые Эдвардса почти простого порядка над полем с модулем p_{384}

$\boldsymbol{p} =$	**FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF EFFFFFFFF0000000000000000FFFFFFFF**
$d =$	12593A
$n =$	40000000000000000000000000000000000000000000000005063576B5A9A0C3A23E9510EA680650B4884E63A2968DD71
$x_G =$	1FC0E8E61F599813E376D11F7510D77F177C2F1CDE19FD14D63A2EC5EAD4D0DED1BD06676CCF365243BF3C0675A31B62
$y_G =$	F52B4FA352B257D7A102FA45C56A50CCBDB3DEC053D5610EDBD0188C11F321F28A43E2FC50395E4A8BD0029AE71D51AA
$d =$	46959A
$n =$	4000000000000000000000000000000000000000000000000C06E6B5CE557E3DDEC6844B3470B819A9E1C0ED5F348517
$x_G =$	4567EA938DECB2BC17F9ABDEDBD77AD188FEFD6E244FD8A12C95A365F9C241E13888506D9D1FB4F8F9B8AA49098C3DF0
$y_G =$	121D587D9CEDEB34DD331E150D9FB2A23B314A354411C4F4A5AFF18685F6EDA1D557EC99F830785420A4DAD15897814E
$d =$	5FD247
$n =$	4000000000000000000000000000000000000000000000000A6A021C752BB157244080A30D578824A05A3BFFF4E67575
$x_G =$	8947F1251392317F3C0E127F7D337B1EE67A120BC69B5E1D91B6C7DC004904C1D42013DE9F12176B6EB34AD5CAC2D485
$y_G =$	B934A355EC5F900123E5EFFDB74502E084A94E59BFDF80E2B0979D02F971F0FE93A4ECE5FFC82C34724A2063836DA847
$d =$	7B35A8
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF96861B6F4CC029EE5B29F5BE0649FA4B3A605D3EC2CFD8F

$x_G =$	712160E70ED57AEC69A66E55A3A285DBFA712F9ACF9B334AC881E5066E55A415BE5935881C7DF7AF90EF6558C380037
$y_G =$	F1FF700AB04A2CA0F40FD94F1E32F61A3D4638570A28EB07AD1D281DAAC0293B946205098F71AC9BEBC8616F1B1DD3E8
$d =$	9302CE
$n =$	4000000000000000000000000000000000000000000000000422338834D697FC3149FC70C79239975452EA01F75E2B699
$x_G =$	E9FDD336AC467447D6DB0627563C181663797C6D87F8E66A3E61BE69E53F66CCC4847DF28A19235D9377F4A33FEE0F69
$y_G =$	CF9AC7F9EEB6528DF53C80BF02E3A83D07614F2708148B4F7F1907398DE79019256D32E535219D71F173DEAE9621716B
$d =$	4CCEDF781B3680122FCFC696374F2ACA57E8F7514293BDD81F5040F0858023ECD295E6835FC14021B1AFFAF8065139C9
$n =$	400000000000000000000000000000000000000000000000046824B5EE5BADD70EB246617456AD087BADECDA7CD2A6E81
$x_G =$	36AFEDFF2308CF4C2F2052B0BF3068F0B4EE9759729A40491EF5D5E75F23E74A6F8528D276E6C528F487DE35E71454CB
$y_G =$	94D6B53F630A6D41E4DD65CD195839CC24DE7B6A71C3DC0A814E379445C67C6604B8034F7C5580AC0FBE771FBA1E9E36

$d =$	3462B5F8DD2744C13ED8618BAFC853F60F83134B936F3DD2867562D1CEEA866E590D702CE8C11262193F077A798918DE
$n =$	400000000000000000000000000000000000000000000000339419162A90FA9A9CA9C0C12B231CF5D95CC5E241D18C0F
$x_G =$	387CCD92DDCF7203D579EC6D158AD3A4DBAB83BBA4E153FB130F380425FD9F73AABEE0F2BE3BCA53D9E1D798DB016F3C
$y_G =$	50E827613BC65DEC52FD123387279D4E9E27102CE6B9119DC656632ED043D3961C72EB7C67D690D970CC65DD2F7C5702
$d =$	941F6CCE649868BD59F335D3293FBB201E3C811ADE19779EEFD8FB21AC1A466377365A8763793A3B3BB2B6575D9893D1
$n =$	4000000000000000000000000000000000000000000000002E0E36DDDDC5D4EA80D339D03926CA39AE3A4739EF81A1C9

$x_G =$	647D270CE4CB1F9038EEE75ED9570A3360AC68DE443BE52D23A734966F813D690615BE4284BDD78CF89B4ED2D02BE08A
$y_G =$	32247232CFA0A712AE3B029B1DDF8C5FE86A78E5728CB4D4386026C21A92297C271C81294C7B9DC18E245F6D87706865
$d =$	584D94FEA391957947760106F180D57CFE8E85109656E4B0BF6FCBF69D62F9A649AEF53CDA9B9F3B2BB247DEB6AB0C70
$n =$	3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF8C7F4A394F5C92E8C6FB2FD97B1657DF603DB9677EEE6DD9
$x_G =$	4CAFBA0D62C538B7CC74F8274978AF35C43E2360BA6606FD379086367898EFD57040629FD9AC9721D65BCB511603BA00
$y_G =$	E2C024B79576CC47776562F9203B166E1C8300AD5C7F4D0B04FE0DBAF92FEC3DD0E5CE1F93288308D2C5CFF5B24DEC9A

На сегодняшний день открытой остается задача адаптации алгоритмов вычисления порядка кривой для кривых в форме Эдвардса. Широко используемые для этой задачи методы SEA и Satoh работают с кривыми в форме Вейерштрасса над большими простыми полями и расширенными полями характеристики 2. Однако приемлемые полные кривые Эдвардса над простыми полями можно получить посредством трансформации кривой Эдвардса в изоморфную кривую в форме Вейерштрасса с последующим определением порядка кривой в форме Вейерштрасса.

Таким способом в данной работе получено 40 кривых Эдвардса над простыми полями с модулями p_{192}, p_{224}, p_{256}, p_{384}. Порядок кривых, приведенных в разделе 3.10.2, имеет минимально возможный кофактор, равный 4, и простой сомножитель n, сравнимый по длине с длиной соответствующего поля. Это делает возможным применение полученных кривых в проектируемых криптосистемах, а общесистемные параметры криптостойких кривых Эдвардса могут быть рекомендованы для стандартизации. Благодаря выигрышу в быстродействии (в среднем в 1.5 – 1.6 раза, см. раздел 3.9) и удобству программирования криптосистем технология полных кривых Эдвардса должна стать эффективной альтернативой канонической форме эллиптических кривых.

ГЛАВА 4

СКРУЧЕННЫЕ КРИВЫЕ ЭДВАРДСА НАД ПРОСТЫМ ПОЛЕМ

В главе 2 мы определили класс скрученных кривых Эдвардса как один из 3-х классов кривых в обобщенной форме Эдвардса (2.1) с параметрами a и d, обдадающими свойствами квадратичных невычетов конечного поля характеристики $p \neq 2$. Как мы показали, лишь для этого класса оправдано введение нового параметра a в уравнение (2.1) кривой, предложенное в работе [3]. Остальные кривые изоморфны кривым в обобщенной форме Эдвардса с коэффициентом $a = 1$. Напомним, что по нащей классификации наряду с непересекающимися классами полных и скрученных кривых Эдвардса в главе 2 мы определили класс квадратичных кривых Эдвардса (у них параметры a и $d -$ квадратичные вычеты). Последний класс с характерными для него свойствами нецикличности, наличием 4-х особых точек, множеством точек 4-го порядка может быть полезен лишь теоретически, но не представляет практического интереса для криптографии. В разделе 2.4 нам он понадобился лишь как класс кривых, являющихся квадратичным кручением класса скрученных кривых Эдвардса.

Анализ некоторых свойств класса скрученных кривых Эдвардса уже был дан в главе 2. Главные отличия скрученных кривых Эдвардса в сравнении с полными – это нециклическая структура группы точек 2-го порядка и наличие среди них двух особых точек с делением на 0 y-координаты. На первый взгляд представляется, что эти особенности вместе с нарушением свойства полноты закона сложения точек делают проблематичным их применение в задачах эллиптической криптографии. Здесь мы постараемся развеять это предубеждение и даже найти некоторые бесспорные преимущества этих кривых. Главной идеей этой главы является возможность эффективного использования скрученных кривых Эдвардса в задачах практической криптографии.

В данной главе проводится более детальный анализ двух свойств этого класса эллиптических кривых: производительности вычислений и условий

делимости точек на 2 с возникновением при этом 4-х решений (квадратных корней) для одной точки, из которой извлекаются квадратные корни.

После работы [3] дальнейший прогресс в изучении новых свойств кривых (2.1) в обобщенной форме Эдвардса получен в работе [4], в которой найдены альтернативные формулы для закона сложения точек кривой, определены особые точки для этого закона и предложен метод расчета координат суммы точек в расширенных (четырехмерных) проективных координатах. Авторам работы [4] удалось снизить число операций при сложении разных точек с $10\mathrm{M} + 2S + 2U$ до $9M + 1U$ (M – одно умножение в поле, S –возведение в квадрат, U – умножение на параметр кривой). Вместе с тем удвоение точки по альтернативным формулам [4] порождает особенности (деление на 0), что заставляет переходить обратно к оригинальным формулам удвоения. Возникает вопрос: имеется ли вообще выигрыш в производительности вычислений при использовании альтернативных формул [4]. Мы исследовали этот вопрос в работе [53] и приводим здесь наши выводы.

В данной главе мы приводим критический анализ некоторых свойств скрученных кривых Эдвардса (разделы 4.1, 4.2), производительности вычисления скалярного произведения точек на такой кривой (разделы 4.3– 4.5) . Далее мы даем сравнительную оценку производительности вычисления скалярного произведения точки кривой в форме Эдвардса и в канонической форме при выполнении операций в проективных координатах [3] и в расширенных проективных координатах [4]. Мы показываем, что полученный авторами [4] выигрыш в вычислении суммы точек практически компенсируется проигрышем при удвоении точки, в итоге общий выигрыш либо незначителен, либо отсутствует. Предложен метод минимизации сложности групповых операций на скрученных кривых Эдвардса путем выбора минимального численного значения параметра a, при этом достигается такое же быстродействие экспоненцирования точки, как у полных кривых Эдвардса (раздел 4.5). Средний показатель выигрыша в экспоненцировании точки на скрученных кривых Эдвардса в сравнении с кривыми в канонической форме приблизительно равен 1.5. В разделах 4.6 – 4.7 впервые получены необходимые и достаточные условия делимости на два точек скрученной кривой Эдвардса и рассмотрены методы нахождения порядка точек кривой.

В заключительном разделе 4.8 приводятся табулированные результаты расчетов общесистемных параметров криптостойких скрученных кривых Эдвардса над простым полем во всем диапазоне стандартных значений модуля поля.

4.1. Определение и свойства скрученных кривых Эдвардса

В работе [3] скрученные кривые Эдвардса (twisted Edwards curves) были определены как обобщение кривых Эдвардса $x^2 + y^2 = 1 + dx^2y^2$ путем ввода нового параметра a в уравнение

$$ax^2 + y^2 = 1 + dx^2y^2,\ a \neq d,\ a,d \in \mathrm{F_p}^*,\ d \neq 1,\ p \neq 2.$$

без ограничения на квадратичность (неквадратичность) параметров a и d. Как уже обсуждалось в главе 2, такое определение нельзя признать корректным, так как оно порождает 3 различных непересекающихся класса кривых с разными свойствами. При этом два из этих классов изоморфны кривым с параметром $a = 1$. Ввод нового параметра $a \neq 1$ не дает подобного изоморфизма в единственном случае, когда оба параметра не являются квадратами: $\left(\frac{a}{p}\right) = -1,\ \left(\frac{d}{p}\right) = -1$. Именно этот случай и определяет отдельный класс скрученных кривых Эдвардса (по нашей классификации в главе 2). Начиная с главы 2, мы пользуемся модифицированными законами сложения (2.2) и удвоения точек (2.3). Они обеспечивают сохранение общепринятой горизонтальной симметрии обратных точек. Все кривые Эдвардса, объединяющие различные непересекающиеся классы кривых, мы определяем кривыми в обобщенной форме Эдвардса с уравнением (2.1)

$$E_{a,d}:\ x^2 + ay^2 = 1 + dx^2y^2,\ \ a,d \in F_p^*, d \neq 1,\ a\ \neq d,\ p \neq 2.$$

Частным классом таких кривых является *скрученная кривая Эдвардса* с ограничениями на параметры a и d

$$E_{a,d}: x^2 + ay^2 = (1 + dx^2y^2),\ a,d \in F_p^*, d \neq 1, a\ \neq d,\ \left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) = -1. \quad (4.1)$$

В настоящей главе используется определение (4.1) скрученных кривых Эдвардса.

Модифицированные законы сложения и удвоения точек кривой (4.1) имеют вид:

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1x_2 - ay_1y_2}{(1-dx_1x_2y_1y_2)}, \frac{x_1y_2 + x_2y_1}{(1+dx_1x_2y_1y_2)}\right). \quad (4.2)$$

$$2(x_1, y_1) = \left(\frac{x_1^2 - ay_1^2}{(1-dx_1^2y_1^2)}, \frac{2x_1y_1}{(1+dx_1^2y_1^2)}\right). \quad (4.3)$$

При $y = 0$ в (4.1) получим $x_{1,2} = \pm 1$. На оси x лежат 2 точки: нейтральный элемент группы $O = (1, 0)$ и точка $D_0 = (-1, 0)$ 2-го порядка. При $x = 0$ в (4.1) имеем $y_{1,2} = \pm 1/\sqrt{a}$. Таким образом, в силу неквадратичности параметра a точки 4-го порядка $\pm F_0 = (0, \pm 1/\sqrt{a})$ на кривой (4.1) не существуют.

Интересным примером, приводящим к скрученной кривой Эдвардса, является кривая (1.1) и изоморфная ей кривая (1.4). Эти уравнения возникли из задачи «пирамида из шаров» в разделе 1.1. Уравнение (1.4) $Y^2 = X^3 - X$ над простым полем F_p является частным случаем кривой в форме Монтгомери (2.9) при $A = 0$, $a = -d$, и $a - d = 2a$. Эта кривая содержит 3 точки 2-го порядка (0,0), $(-1,0)$ и (1,0) и при $p = 3 mod 4$ имеет порядок $N_E = p + 1$ [29]. В этом случае пара кривых кручения вырождается в одну суперсингулярную кривую.

4.2. Свойства точек порядков 2, 4, 8 на скрученных кривых Эдвардса

Из уравнения (4.1) определяем:

$$x^2 = \frac{1-ay^2}{1-dy^2}, \qquad y^2 = \frac{1-x^2}{a-dx^2}.$$

Особые точки второго порядка (точки на бесконечности, при этом знаку "∞" отвечает деление на 0) в условиях кривой (4.1) возникают лишь для y -координаиы:

$$D_{1,2} = \left(\pm\sqrt{\frac{a}{d}}, \infty\right). \quad (4.4)$$

В главе 2 мы ввели арифметику с особыми точками (4.4) кривой, которая отвечает правилам предельного перехода для неопределенностей типа $\left(\frac{\infty}{\infty}\right)$. С ее помощью можно проверить, что $2D_{1,2} = 0$. Итак, кривая (4.1), как нам известно из главы 2, является нециклической с тремя точками 2-го порядка: $D_0 = (-1, 0)$, и $D_{1,2} = \left(\pm\sqrt{\frac{a}{d}}, \infty\right)$. Отсюда следует, что минимальный кофактор порядка кривой равен 4, т.е. $N_E = 0mod4$.

Если кривая (4.1) имеет точки 4-го порядка, минимальный кофактор порядка кривой N_E станет равным 8 (или $N_E = 0mod8$), при наличии точек 8-го порядка – 16, и т.д. Требуется найти условия, порождающие такие точки на кривой (4.1).

Дадим анализ некоторых новых свойств точек 4-го и 8-го порядков [56,58] на кривой (4.1).

Точки 4-го порядка могут рассматриваться как точки деления на два точек 2-го порядка.

Утверждение 4.1. Для точки 2-го порядка $D_0 = (-1, 0)$ кривой (4.1) не существует точек деления на 2.

Доказательство. Допустим обратное, что существует точка 4-го порядка $F = (x_1, y_1)$, такая, что $2F = D_0 = (-1, 0)$. Согласно закону удвоения (4.3) из равенства

$$\frac{2x_1y_1}{(1 + dx_1^2y_1^2)} = 0 \Rightarrow x_1y_1 = 0,$$

т.е. точка 4-го порядка может лежать лишь на оси x ($y_1 = 0$) или на оси y ($x_1 = 0$). Но из уравнения кривой (4.1) и законов (4.2) и (4.3) следует, что на оси x исключительными являются 2 точки: $O = (1, 0)$ и точка $D_0 = (-1, 0)$ 2-го порядка. На оси y уравнение (4.1) решения не имеет. Следовательно, не существует точки $F = (x_1, y_1)$, удвоение которой дает точку $D_0 = (-1, 0)$, и наше допущение неверно. Утверждение доказано. ▲

Дальнейший анализ показал, что точки 4-го порядка кривой (4.1) могут порождаться лишь делением на 2 особых точек 2-го порядка (4.4). В разделе 2.1 доказана теорема 2.1 о необходимых и достаточных условиях существования точек 4-го порядка кривой (2.1), которые приводят к скрученной кривой (4.1). Доказательство ее построено на делении на 2 особых точек (4.4) (см. раздел 2.1).

Теорема 4.1. *Точки 4-го порядка скрученной кривой в форме* (4.1) *при* $x \neq 0$ *существуют тогда и только тогда, когда выполняется условие*:

$$p \equiv 3 mod 4.$$

При выполнении условий теоремы координаты 4-х точек 4-го порядка согласно (2.5) определяются как:

$$\pm F_2 = \left(\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right), \ \pm F_3 = \left(-\sqrt[4]{\frac{a}{d}}, \pm\sqrt{\frac{-1}{\sqrt{ad}}}\right). \qquad (4.5)$$

Эта теорема является частью теоремы 2.1, в которой учтены лишь условия (i), справедливые для скрученной кривой Едвардса. Ее доказательство дано в разделе 2.2.

Точки $\pm F_{2,3}$ можно рассматривать как точки деления на два точек 2-го порядка $D_{1,2}/2$ [57].

Утверждение 4.2. *Все скрученные кривые Едвардса* (4.1) *при* $p \equiv 1 mod 4$ *имеют порядок* $N_E = 4n$, а при $p \equiv 3 mod 4$ порядок кривой $N_E = 0 mod 8$.

Доказательство. Условия (i) теоремы 2.1 определяет класс скрученных кривых Едвардса (4.1). При $p \equiv 1 mod 4$ кривая не содержит точек 4-го порядка, так как не выполняется условие (ii) теоремы. Однако кривая (4.1) включает нециклическую подгруппу 4-го порядка точек 2-го порядка $G_4 = \{O, D_0, D_1, D_2\}$. Следовательно, порядки всех других точек могут быть равными n и $2n$ (вместе с возможными нечетными сомножителями n). Итак, подгруппа G_4 есть подгруппа минимального четного порядка 4 кривой, и порядок кривой $N_E = 4n$.

При $p \equiv 3 mod 4$ согласно теореме 2.1 кривая (4.1) содержит точки (4.5) 4-го порядка, суммирование которых с точкой D_0 2-го порядка из подгруппы

G_4 образует нециклическую группу 8-го порядка точек 2-го и 4-го порядков. Отсюда следует, что $N_E = 0 mod 8$. Утверждение доказано.

Итак, все скрученные кривые Едвардса при $p \equiv 1 mod 4$ имеют порядок $N_E = 4n$. Их можно рекомендовать для криптографии.

Найдем условия существования точек 8-го порядка. Согласно теореме 2.2 раздела 2.1, для скрученных кривых Едвардса ни одно из необходимых условий (i) и (ii) теоремы не выполняется. Это следует из того, что на скрученных кривых точки $\pm F_0$ 4-го порядка не существуют. Поэтому точки 8-го порядка на кривых (4.1) следует искать как точки деления на 2 точек 4-го порядка (4.5) лишь при $p \equiv 3 mod 4$.

Теорема 4.2. *Необходимыми условиями существования точек 8-го порядка скрученной кривой* (4.1) *являются:*

$$\text{(i)} \quad \left(\frac{1+\sqrt{\frac{a}{d}}}{p}\right) = 1, \qquad \text{(ii)} \quad p \equiv 3 mod 4.$$

Доказательство. Пусть $S = (x_1, y_1)$ – точка 8-го порядка, тогда $2S = F_2 = (X, \mathrm{Y})$ – точка 4-го порядка. Согласно (4.1), (4.3) и второй координаты (4.5) точки $\boldsymbol{F_2}$ имеем

$$\frac{2x_1y_1}{(1+dx_1^2y_1^2)} = \frac{2x_1y_1}{(x_1^2+ay_1^2)} = \mathrm{Y} = \sqrt{\frac{-1}{\sqrt{ad}}}. \tag{4.6}$$

Обозначим $Z = \frac{y_1}{x_1}$, $V = x_1y_1$. Тогда второе уравнение в (4.6) с учетом (4.1) можно записать в виде двух квадратных уравнений:

$$aZ^2 - 2Y^{-1}Z + 1 = 0, \quad dV^2 - 2Y^{-1}V + 1 = 0.$$

Их дискриминанты после подстановки Y^2 из (4.6) равны

$$\Delta_1 = 4Y^{-2}(1 - a\,Y^2), \qquad \Delta_2 = 4Y^{-2}(1 - d\,Y^2). \tag{4.7}$$

Уравнение (4.1) в точке (X, Y) можно записать как

$$(1 - a\,Y^2) = X^2(1 - d\,Y^2).$$

Отсюда следует, что если один из дискриминанатов $\Delta_{1,2}$ является квадратом, то и другой – также квадрат (и наоборот). Потребуем, например, чтобы

элемент поля $(1 - a\,Y^2)$ был квадратом, тогда с учетом (4.6) и (4.7) квадратом должен быть сомножитель $\left(1 + \sqrt{\frac{a}{d}}\right)$ дискриминанта Δ_1. Необходимое условие (i) теоремы доказано. Условие (ii) является необходимым условием существования точек $F_{2,3}$ 4-го порядка, деление которых на 2 порождает точки 8-го порядка. Теорема доказана. ▲

Пример 4.1. Дана скрученная кривая $x^2 - y^2 = (1 + 2x^2y^2) mod 11$. Для нее $a = -1, d = 2 -$ квадратичные невычеты при p = 11, причем выполняются оба условия теоремы 4.1. Она имеет 3 точки 2-го порядка $D_0 = (-1,0), D_2 = (-4, \infty), D_1 = (4, \infty)$, а 4 точки 4-го порядка имеют координаты $\pm F_{2,3} = (\pm 2, \pm 2)$. При удвоении согласно (4.3) получим $2\boldsymbol{F_2}$ = $(4, \infty)$ = D_1 . Порядок этой кривой $N_E = 16$, группа точек нециклическая с типом T = (2, 2^3). Она содержит кроме точек O, $D_{0,1,2}$, $\pm F_{2,3}$, 8 точек 8-го порядка $(\pm 3, \pm 1), (\pm 5, \pm 4)$. Заметим, что эта кривая содержит 2 циклические подгруппы 4-го порядка и 2 циклические подгруппы 8-го порядка. Генераторами циклических подгрупп 4-го порядка здесь являются, например, точки F_2 = (2, 2) и F_3 = (−2, −2). Эти циклические подгруппы с последовательным экспоненцированием элементов записываются как:

$$G_4' = \{O,\ F_2 = (2,2),\ 2F_2 = (4,\infty) = D_2,\ 3F_2 = (2,-2) = -F_2\},$$
$$G_4'' = \{O,\ F_3 = (-2,-2), 2F_3 = (-4,\infty) = D_1,\ 3F_3 = (-2,2) = -F_3\}.$$

Отсюда видно, что все элементы группы G_4'', кроме нейтрального, имеют обратные по сравнению с элементами группы G_4' знаки обеих координат. Это отвечает известному свойству сложения произвольной точки с точкой 2-го порядка $D_0 = (-1,0)$: $(x_1, y_1) + (-1, 0) = (-x_1, -y_1)$. Другими словами, элементы подгрупп G_4' и G_4'' связаны точкой D_0, а прямая сумма двух циклических подгрупп 2-го и 4-го порядков $\{O, D_0\} \oplus G_4'$ образует нециклическую группу 8-го порядка точек 2-го и 4-го порядков. Ясно, что увеличение вдвое порядка этой группы возникло в связи с тем, что точка D_0 лежит за пределами подгрупп G_4' и G_4''. Подобный же анализ можно провести с двумя циклическими подгруппами G_8' и G_4'' 8-го порядка, также не включающими точки D_0. При этом кривая E 16-го порядка представляется прямой суммой циклических подгрупп $\{O, D_0\} \oplus G_8'$.

Часто точки кривой (4.1) приходится складывать с точками 2-го порядка $D_0 = (-1, 0)$, $D_1 = (4, \infty), D_2 = (-4, \infty)$. Принимая правила предельного перехода в (4.2), можно найти координаты сумм:

$$(x_1, y_1) + (-1,\ 0) = (-x_1, -y_1),$$

$$(x_1, y_1) + \left(\sqrt{\frac{a}{d}}, \infty\right) = \left(\sqrt{\frac{a}{d}} \cdot x_1^{-1},\ \frac{1}{\sqrt{ad}} \cdot y_1^{-1}\right), \tag{4.8}$$

$$(x_1, y_1) + \left(-\sqrt{\frac{a}{d}}, \infty\right) = \left(-\sqrt{\frac{a}{d}} \cdot x_1^{-1}, -\frac{1}{\sqrt{ad}} \cdot y_1^{-1}\right).$$

Все найденные суммы удовлетворяют уравнению (4.1).

4.3. Производительность экспоненцирования точек на скрученной кривой Эдвардса

В этом разделе мы проведем анализ сложности групповых операций на скрученной кривой Эдвардса в классических проективнх координатах и в так называемых расширенных проективных координатах (для альтернативного закона сложения [4]), после чего дадим сравнительные оценки производительности экспоненцирования точек для кривых в форме Эдвардса и Вейерштрасса.

4.3.1. Альтернативный закон сложения точек

Авторы статьи [4], изыскивая возможности ускорения групповых операций на скрученных кривых Эдвардса, нашли интересный резерв для решения этой задачи. Выразив параметры a и d через координаты складываемых точек, они получили альтернативные формулы для закона сложения точек, в частности

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1 y_1 + x_2 y_2}{y_1 y_2 + a x_1 x_2},\ \frac{x_1 y_1 - x_2 y_2}{x_1 y_2 - x_2 y_1}\right) = (x_3, y_3). \tag{4.9}$$

Его достоинством является независимость координат от обоих параметров a и d кривой (4.1). Вместе с тем этот закон не работает при удвоении точек, так

как во второй координате возникает неопределенность 0/0. Поэтому при удвоении точек приходится возвращаться к обычной формуле (4.3). Здесь основная проблема связана с частым переходом к различным проективным координатам (3-х и 4-х мерным, что будет понятным после ввода расширенных проективных координат ниже).

Хотя альтернативный закон (4.9) уже в общем случае не является полным (существуют особые точки, обращающие знаменатели в 0), для точек нечетного порядка особых точек нет и формулы (4.9) вполне конструктивны. Мы обсудим этот вопрос в разделе 4.4.

4.3.2. Особые точки альтернативного закона сложения

Из формулы (4.5) для *y*-координаты сразу видно, что при удвоении точки знаменатель обращается в 0. Поэтому при удвоении точки следует пользоваться законом (4.2). Но и при сложении разных точек также возникают особые пары точек, обращающие знаменатели *x*-координаты и *y*-координаты в ноль. В [4] доказана теорема:

Теорема 4.3. *Пусть имеем скрученную кривую Эдвардса* $E_{a,d}$ (4.1). *Для фиксированной точки кривой* $P = (x_1, y_1)$ *найдется такая точка* $Q = (x_2, y_2)$, для которой:

1). $y_1y_2 + ax_1x_2 = 0$ тогда и только тогда, когда $Q \in S_x$, где

$$S_x = \{\left(\frac{y_1}{\sqrt{a}}, -x_1\sqrt{a}\right), \left(-\frac{y_1}{\sqrt{a}}, x_1\sqrt{a}\right), \left(\frac{1}{x_1\sqrt{ad}}, -\frac{\sqrt{a}}{y_1\sqrt{d}}\right), \left(\frac{-1}{x_1\sqrt{ad}}, \frac{\sqrt{a}}{y_1\sqrt{d}}\right)\};$$

2). $x_1y_2 - y_1x_2 = 0$ тогда и только тогда, когда $Q \in S_y$, где

$$S_y = \{(x_1, y_1), (-x_1, -y_1), \left(\frac{1}{y_1\sqrt{d}}, \frac{1}{x_1\sqrt{d}}\right), \left(\frac{-1}{y_1\sqrt{d}}, \frac{-1}{x_1\sqrt{d}}\right)\}.$$

Рассмотрим случай $\left(\frac{a}{p}\right) = 1$, $\left(\frac{d}{p}\right) = -1$. Тогда в каждом из множеств S_x, S_y остается по две первых точки. Их координаты для множества S_x определяются как $Q = P \pm F$, где $\pm F = (\pm\frac{1}{\sqrt{a}}, 0)$ – точки 4-го порядка. Тогда здесь возникает особенность при $Q + P = 2P \pm F$. Координаты точек для множества S_x определяются как $Q = P$ и $Q = P + D$, где $D = (0, 1)$ – точка 2-го порядка. Здесь особенность возникает при $Q + P = 2P$ и при $Q + P = 2P + D$. Мы видим, что все особые случаи порождаются удвоением точки P с возможным суммированием с ним точек 4-го или 2-го порядков.

Если P – точка нечетного порядка n, то Ord($2P$) = n , так как $n2P$ = O. Отсюда следует, что Ord($2P \pm F$) = $4n$ и Ord($2P + D$) = $2n$. Другими словами, особенности в рассматриваемом случае могут возникать лишь при сложении разных точек четных порядков (мы исключаем удвоение для закона (4.5)). Вообще говоря, при вычислении скалярного произведения kP при больших значениях k для каждой точки P большого порядка (возможно четного) может существовать всего 3 точки Q таких, что сумма $Q + P$ не определена, а вероятность такого события ничтожна. В криптосистеме с генератором G простого порядка n суммирование любых разных точек из группы $< G >$ с помощью формулы (4.5) особенностей не порождает. Это всегда справедливо для всех точек нечетного порядка.

4.3.3. Сложность групповых операций на скрученной кривой Эдвардса

1. Сложение точек (закон сложения (4.2))

Для полных кривых Эдвардса этот анализ приведен в разделе 3.8.1. Так как в уравнении кривой (4.1) появился новый параметр a, требуется оценить, насколько он увеличивает вычислительные затраты. Введем третью координату Z как общий знаменатель в (4.2). Пусть $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$, тогда однородное уравнение кривой (4.1) в проективных координатах имеет вид

$$(X^2 + aY^2)Z^2 = Z^4 + dX^2Y^2, \quad X = xZ, \quad Y = yZ.$$

Сумма двух точек теперь записывается как $(X_1 : Y_1 : Z_1) + (X_2 : Y_2 : Z_2) = (X_3 : Y_3 : Z_3)$. С учетом подстановок выразим координаты суммарной точки согласно (4.2):

$$x_3 = \frac{X_3}{Z_3} = \frac{Z_1Z_2\left(Z_1^2Z_2^2 + dX_1X_2Y_1Y_2\right)(X_1X_2 - aY_1Y_2)}{\left(Z_1^2Z_2^2 + dX_1X_2Y_1Y_2\right)\left(Z_1^2Z_2^2 - dX_1X_2Y_1Y_2\right)},$$

$$y_3 = \frac{Y_3}{Z_3} = \frac{Z_1Z_2\left(Z_1^2Z_2^2 - dX_1X_2Y_1Y_2\right)(X_1Y_2 + X_2Y_1)}{\left(Z_1^2Z_2^2 + dX_1X_2Y_1Y_2\right)\left(Z_1^2Z_2^2 - dX_1X_2Y_1Y_2\right)}.$$

Обозначим:

$A = Z_1Z_2;\ B = A^2;\ \ C = X_1X_2;\ D = aY_1Y_2; E = dCD;$

$F = B - E;\ G = B + E$

Тогда

$Y_3 = A{\cdot}F{\cdot}\ ((X_1 + Y_1){\cdot}(X_2 + Y_2) - C - D),$

$X_3 = A{\cdot}G{\cdot}(D - C),$

$Z_3 = F{\cdot}G.$

Подсчет числа элементарных операций здесь дает 10 умножений M, одно возведение в квадрат S и 2 умножения на параметры a и d кривой. Итак, находим сложность вычисления суммы различных точек, выраженную через число умножений и возведений в квадрат в поле $V_E = 10M + 1S + 2U$ [3].

2. Удвоение точек (закон удвоения (4.3))

Используя уравнение кривой (4.1), закон удвоения (4.3) запишем в форме, не зависящей от параметра d

$$2(x_1, y_1) = \left(\frac{x_1^2 - ay_1^2}{2 - x_1^2 - ay_1^2}, \frac{2x_1y_1}{x_1^2 + ay_1^2},\right).$$

Тогда координаты точки удвоения согласно (3.3):

$$x_3 = \frac{X_3}{Z_3} = \frac{\left(\left(\frac{X_1}{Z_1}\right)^2 - a\left(\frac{Y_1}{Z_1}\right)^2\right)\left(\left(\frac{X_1}{Z_1}\right)^2 + a\left(\frac{Y_1}{Z_1}\right)^2\right)}{\left(2 - \left(\frac{X_1}{Z_1}\right)^2 - a\left(\frac{Y_1}{Z_1}\right)^2\right)\left(\left(\frac{X_1}{Z_1}\right)^2 + a\left(\frac{Y_1}{Z_1}\right)^2\right)}$$

$$= \frac{(X_1^2 - aY_1^2)(X_1^2 + aY_1^2)}{(2Z_1^2 - X_1^2 - aY_1^2)(X_1^2 + aY_1^2)},$$

$$y_3 = \frac{Y_3}{Z_3} = \frac{2\frac{X_1Y_1}{Z_1Z_1}\left(2 - \left(\frac{X_1}{Z_1}\right)^2 - a\left(\frac{Y_1}{Z_1}\right)^2\right)}{\left(2 - \left(\frac{X_1}{Z_1}\right)^2 - a\left(\frac{Y_1}{Z_1}\right)^2\right)\left(\left(\frac{X_1}{Z_1}\right)^2 + a\left(\frac{Y_1}{Z_1}\right)^2\right)} = \frac{2X_1Y_1(2Z_1^2 - X_1^2 - aY_1^2)}{(2Z_1^2 - X_1^2 - aY_1^2)(X_1^2 + aY_1^2)}.$$

Обозначим

$A = X_1^2, B = Y_1^2, C = aY_1^2,\ D = Z_1^2, E = (A + C), F = (A - C), G = 2D - A - C,$

$H = (X_1 + Y_1)^2 \Rightarrow 2X_1 \cdot Y_1 = H - A - B.$

Тогда

$$X_3 = E \cdot F,$$

$$Y_3 = 2X_1Y_1 \cdot G,$$

$$Z_3 = E \cdot G.$$

Подсчет числа возведений в квадрат и умножений в поле дает суммарную сложность группового удвоения $T_E = 3M + 4S + 1U$ [3].

3. Сложность сложения (альтернативный закон (4.9))

Вводя расширенные проективные координаты $(X:Y:T:Z)$, авторам [4] удалось сократить число полевых операций при сложении 2-х разных точек до $9\mathrm{M} + 1U$ в сравнении со сложностью $10\mathrm{M} + 1S + 2U$ при реализации сложения по формуле (4.2) [3]. Рассмотрим этот метод.

При Z ≠ 0 зададим четырехмерные проективные координаты $(X:Y:T:Z)$, подстановкой в (4.9) *x = X/Z*, *y = Y/Z*, *t = xy/Z*, *T = XY/Z*. Тогда

$$\frac{X_3}{Z_3} = \frac{(T_1Z_2 + Z_1T_2)}{(Y_1Y_2 + aX_1X_2)}, \qquad \frac{Y_3}{Z_3} = \frac{(T_1Z_2 - Z_1T_2)}{(X_1Y_2 - Y_1X_2)}.$$

Отсюда

$$\begin{aligned} X_3 &= (X_1Y_2 - Y_1X_2)((T_1Z_2 + Z_1T_2), \\ Y_3 &= (Y_1Y_2 + aX_1X_2)((T_1Z_2 - Z_1T_2), \\ T_3 &= (T_1Z_2 + Z_1T_2)((T_1Z_2 - Z_1T_2), \\ Z_3 &= (Y_1Y_2 + aX_1X_2)((X_1Y_2 - Y_1X_2). \end{aligned} \qquad (4.10)$$

Пусть $A = X_1X_2$, $B = Y_1Y_2$, $C = T_1Z_2$, $D = Z_1T_2$, $E = C + D$, $F = C - D$,

$G = B + aA$, $H = (X_1 - Y_1)(X_2 + Y_2) - A + B$.

Тогда

$X_3 = EH, \quad Y_3 = GF, \quad T_3 = EF, \quad Z_3 = GH.$

Мы видим, что сложность групповой операции сложения разных точек составляет $V_E^* = 9\text{M} + 1U$. Если параметр a = ± 1 или мал, сложность оценивается как 9M. При удвоении точки кривой Эдвардса в трехмерных проективных координатах сложность минимальна и составляет $T_E = 3\text{M} + 4\text{S} + + 1U$[3]. В работе [4] показано, что в расширенных проективных координатах сложность удвоения возрастает на одну операцию умножения $T_E^* = 4\text{M} + 4S + 1U$.

Значения сложности групповых операций для полных кривых Эдвардса (результаты раздела 3.8.1) и скрученных кривых Эдвардса (для обычных проективных и расширенных проективных координат, раздел 4.2.2) приведены в таблице 4.1.

Таблица 4.1

Класс кривых, координаты	Сложность групповой операции	
	Сложение точек	Удвоение точек
Полные кривые Эдвардса, проективные координаты	$10\text{M} + 1S + 1U$	$3\text{M} + 4S$
Скрученные кривые Эдвардса, проективные координаты	$10\text{M} + 1S + 2U$	$3\text{M} + 4S + 1U$
Скрученные кривые Эдвардса, расширенные проективные координаты	$9\text{M} + 1U$	$4\text{M} + 4S + 1U$

Наименьших вычислительных затрат, как следует из таблицы, требуют операции на полных кривых Эдвардса. Особенно они выигрывают при удвоении, которое обходится без операции умножения на параметр кривой $1U$.

4.4. Сравнительный анализ производительности экспоненцирования точки на скрученной кривой Эдвардса и кривой в форме Вейерштрасса

Ниже приводятся сравнительные оценки быстродействия арифметики на скрученной кривой Эдвардса и канонической кривой для двух записей закона сложения точек.

4.4.1. Производителность экспоненцирования точки в расширенных проективных координатах (альтернативный закон сложения)

Оценим выигрыш в производительности при вычислении скалярного произведения на скрученной кривой Эдвардса в расширенных проективных координатах по сравнению с аналогичной процедурой на канонической кривой в проективных координатах. Эти оценки рассматривались в работе [53].

Расчет числа операций при вычислении суммы точек кривой в форме Вейерштрасса W дает сложность $V_W = 12M + 2S$. Аналогичный расчет для удвоения точек приводит к результату $T_W = 7\text{M} + 5\text{S}$ [29] .

На скрученной кривой Эдвардса E в расширенных проективных координатах имеем согласно раздела 4.2.2 соответствующие показатели сложности $V_E^* = 9\text{M} + 1U$, $T_E^* = 4\text{M} + 4S + 1U$.

Принимая вычислительную сложность возведения в квадрат $1S = 0.67M$, а умножения на параметр кривой $1U = 0.5M$, получим оценки сложности сложения и удвоения на кривой в форме Вейерштрасса $V_W = 13.33M$, $T_W = 10.33M$. Аналогичные показатели для скрученной кривой Эдвардса $V_E^* = 9.5\text{M}, T_E^* = 7.17\text{M}$.

Как и в разделе 3.8, выигрыш в производительности экспоненцирования точки на скрученной кривой Эдвардса в сравнении с кривой в форме Вейерштрасса определяется коэффициентом

$$\gamma(v) = \frac{T_W + vV_W}{T_E^* + vV_E^*}, \qquad (4.11)$$

где ν – относительная частота знаков 1 в двоичной последовательности k скалярного произведения kP точки P. Подставляя сюда найденные оценки, получим

$$\gamma(v) = \frac{10.33 + v13.33}{7.17 + v9.5}.$$

В среднем при равновероятных 0 и 1 в двоичной записи числа k ($v = 0.5$) получаем среднее значение выигрыша $\gamma(0.5) = 1.426$.

Заменяя двоичное представление числа k произведения kP троичным NAF(k) $k_i \in \{0,1,-1\}$[29], можно снизить среднее число ненулевых компонент в числе k до 1/3. Для этого случая получим максимальное значение среднего выигрыша в расширенных проективных координатах $\gamma(0.33) = 1.429$.

4.4.2. Производительность экспоненцирования точки в проективных координатах (классический закон сложения)

В отличие от предыдущего подраздела здесь мы анализируем выигрыш в производительности экспоненцирования точки на скрученной кривой Эдвардса с использованием закона сложения (4.2) и результатов раздела 4.2.2.

На скрученной кривой Эдвардса E в проективных координатах имеем согласно раздела 4.2.2 следующие показатели сложности: $V_E = 10M + 1S + 2U$, $T_E = 3M + 4S + 1U$. Пересчет этих результатов в эквивалентное число умножений в поле дает значения: $V_E = 11.67M$, $T_E = 6.17M$. С учетом этих оценок коэффициент (4.11) равен

$$\gamma(v) = \frac{10.33 + v13.33}{6.17 + v11.67}.$$

Тогда при двоичном и троичном представлениях числа k получим по аналогии с предыдущим подразделом средние значения выигрышей: $\gamma(0.5) = 1.416$, $\gamma(0.33) = 1.47$.

Для сравнения выигрыша производительности экспоненцирования точки кривых Эдвардса по сравнению с кривыми в канонической форме Вейерштрасса мы свели результаты расчетов коэффициентов $\gamma(v)$ в таблицу 4.2.

Таблица 4.2. Сравнительная оценка выигрыша в скорости экспоненцирования точки для скрученных кривых Эдвардса

Класс кривых, тип координат	Выигрыш $\gamma(\nu)$	
	($\nu = 0.5$)-двоичное k	($\nu = 0.33$)-троичное k
Полные кривые Эдвардса, проективные координаты	1.51	1.574
Скрученные кривые Эдвардса, проективные координаты	1.416	1.47
Скрученные кривые Эдвардса, расширенные проективные координаты	1.426	1.429

Итак, использование закона сложения (4.5) и расширенных проективных координат дает весьма незначительный прирост производительности вычислений на кривой Эдвардса по сравнению с полным универсальным законом сложения (4.2). При троичном экспоненцировании точки скрученной кривой Эдвардса классический закон сложения (4.2) в проективных координатах дает небольшой выигрыш по сравнению с альтернативным законом сложения (4.5) в расширенных координатах. Наилучшие результаты в быстродействии экспоненцирования, как и ожидалось, обеспечивают полные кривые Эдвардса, не имеющие избыточного параметра a. Следует подчеркнуть, что по сравнению с каноническими кривыми обе арифметики дают прирост скорости экспоненцирования приблизительно в 1.5 раза.

4.5. Метод достижения минимальной сложности групповых операций на скрученной кривой Эдвардса

Как следует из таблицы 4.1, ввод дополнительного парметра a в уравнение скрученной кривой (4.1) увеличивает вычислительные затраты сложения точек на одну операцию $1U$ и удвоения точек на $1U$ в сравнении с полной кривой Эдвардса. В этом подразделе мы предлагаем простой способ, как можно избавиться от этих дополнительных затрат и достичь максимальной производительности экспоненцирования точки на скрученной кривой Эдвардса.

В главе 2 отмечалось, что квадратичное кручение любой скрученной кривой Эдвардса дает квадратичную кривую Эдвардса и обратно: $E_{a,d}^{t} \sim E_{\,c\,a,\,c\,d}$ (здесь $\left(\frac{c}{p}\right) = -1, \left(\frac{ad}{p}\right) = 1$). Кроме того, внутри классов с условиями *C*2.1 и *C*2.2 существует изоморфизм кривых $E_{a,d} \sim E_{d,a}$. Он легко доказывается, например, на основе замены $(x, y) \to (1/X, Y)$ и умножения на X^2.

Свойства изоморфизма и квадратичного кручения можно обосновать также, используя *j*-инвариант кривой в обобщенной форме Эдвардса [3,12]

$$j(a,d) = \frac{16(a^2+d^2+14ad)^3}{ad(a-d)^4}, \quad ad(a-d) \neq 0. \qquad (4.12)$$

Как известно [16,22,68], изоморфные кривые (с порядком $N_E = p + 1 - t$) и кривые квадратичного кручения (с порядком $N_E^t = p + 1 + t$) имеют один и тот же *j*-инвариант. Из (4.12) сразу следуют свойства симметрии *j*-инварианта относительно переменных ca, cd и их инверсий:

$$j(a,d) = j(d,a), \qquad (4.13)$$
$$j(a,d) = j(ca,cd), \qquad (4.14)$$
$$j(a,d) = j(a^{-1},d^{-1}), \qquad (4.15)$$
$$j(a,d) = j(1,d/a) = j(1,a/d). \qquad (4.16)$$

Внутри класса скрученных кривых Эдвардса нет пар квадратичного кручения, но для каждой кривой имеется изоморфная кривая со свойством (4.13) или $E_{a,d} \sim E_{d,a}$, причем $\left(\frac{a}{p}\right) = -1, \left(\frac{d}{p}\right) = -1$.

Идея состоит в том, что при поиске подходящей для криптографии скрученной кривой Эдвардса нет смысла в переборе различных значений параметров a и d. Можно зафиксировать один из этих параметров (например, параметр *a*) и варьировать другой в области его допустимых значений. Если задать этот фиксированный параметр на минимальном числовом уровне $a = \alpha$, таком что $\left(\frac{\alpha}{p}\right) = -1$, то можно сэкономить полевую операцию $1U$ (умножение на параметр *a* кривой (4.1)) при сложении точек и удвоении точки кривой. Например, $\alpha = 2$, тогда двукратное сложение (тождественное

умножению на 2) можно считать «бесплатной» операцией. Практически то же справедливо при $\alpha = 3,5$ и других малых по сравнению с n чисел. При этом достигается минимальная сложность групповой операции, равная сложности операции для полной кривой Эдвардса.

Нам требуется доказать, что при фиксации параметра $a = \alpha$, перебор всех допустимых параметров d дает все возможные значения j-инварианта и, соответственно, порядков скрученной кривой.

Утверждение 4.3. *При фиксированном значении параметра* $a = \alpha$ *кривой* (4.1) ее j*-инвариант* $j(\alpha, d)$ *принимает* $(p-1)/4$ *возможных значений при* $p = 1mod4$ *и* $(p-3)/4$ *возможных значений при* $p = 3mod4$ *при всех* $\left(\frac{d}{p}\right) = -1, d \neq \alpha$.

Доказательство. Рассмотрим квадратичные кривые Эдвардса с параметрами $\left(\frac{a}{p}\right) = 1, \left(\frac{d}{p}\right) = 1$. Для любой такой кривой существует изоморфизм $E_{a,d}$~$E_{1,d/a}$~$E_{1,a/d}$. Обозначим $\delta = \frac{d}{a}$. Из всех $\frac{(p-1)}{2}$ квадратов мультипликативной группы параметр $\delta \neq 1$ принимает ровно $\frac{(p-3)}{2}$ допустимых значения. При $p = 1mod4$ для каждого δ, кроме квадрата $\delta = -1$, существует пара изоморфных кривых $E_{1,\delta}$~$E_{1,\delta^{-1}}$. Случай $\delta = -1 = \delta^{-1}$ вырождает пару изоморфных кривых в одну кривую. Тогда j-инвариант (4.16) кривой $E_{1,\delta}$ принимает ровно $\frac{(p-1)}{4}$ значения. При $p = 3mod4$ имеется ровно $\frac{(p-1)}{2}$ квадратов и $\frac{(p-3)}{2}$ допустимых значений δ. В этом случае элемент (-1) является квадратичным невычетом и существует ровно $\frac{(p-3)}{4}$ пар изоморфных кривых и такое же число j-инвариантов.

Парой кручения каждой квадратичной кривой Эдвардса является скрученная кривая (4.1), т.е. $E^t_{a,d}$~$E_{ca,cd}$, $\left(\frac{c}{p}\right) = -1$. Следовательно, число изоморфных пар скрученных кривых , равное числу j-инвариантов с теми же значениями, что и для квадратичных кривых Эдвардса, также равно $\frac{(p-1)}{4}$ при $p = 1mod4$ и $\frac{(p-3)}{4}$ при $p = 3mod4$. Осталось доказать, что все изоморфные пары скрученных кривых Эдвардса могут быть получены при одном фиксированном значении параметра $a = \alpha$.

Воспользуемся свойсвами (4.13), (4.14), и умножим параметры второго j-инварианта на $\mathrm{c} = \frac{a}{d}$, тогда

$$j(a,d) = j(d,a) = j\left(a, \frac{a^2}{d}\right). \qquad (4.17)$$

Отсюда следует, что любая пара изоморфных скрученных кривых Эдвардса определяется единственным параметром $a = \alpha$ и множеством всех пар квадратичных невычетов d и $d^{-1}, d \neq \alpha$. Объемы множеств таких пар, как и число изоморфизмов скрученных кривых Эдвардса, остается таким же, как и для квадратичных кривых Эдвардса. Утверждение доказано. ▲

Как ранее отмечалось, все скрученные кривые Эдвардса имеют порядок $4n$ при $p = 1 mod 4$, поэтому нам интересен для криптографии лишь этот случай. Минимальное числовое значение квадратичного невычета $\alpha = 2$ существует лишь при $p = \pm 3 mod 8$ [29]. Следующее желаемое значение квадратичного невычета $\alpha = 3$ требует выполнения $p = \pm 5 mod 12$ [67]. В таблице 4.3 приведены для примера простые числа $p = 1 mod 4$, для которых $\alpha = 2$ и $\alpha = 3$ (соответствующие столбцы помечены знаком +).

Таблица 4.3

p	13	17	29	37	41	53	61	73	89	97
α=2	+		+	+		+	+			
α=3		+			+				+	

Хотя эта выборка из первой сотни простых чисел с заданными свойствами не репрезентативна, можно сделать предположение, что около 80% простых чисел $p = 1 mod 4$ являются модулями полей, содержащих квадратичные невычеты 2 или 3. В других случаях всегда можно найти минимальное значение параметра $a = \alpha$, что позволяет пренебречь сложностью операции $1U$ в оценках сложности групповых операций сложения и удвоения точек.

Пример 4.2. Пусть $p = 29$ и $a = \alpha = 2$. Согласно формулы (4.12) можно сначала найти j-инвариант единственной квадратичной кривой Эдвардса

$j(1,-1) = 17$ и соответствующей скрученной кривой Эдвардса $j(2,-2) = 17$. Далее в соответствии с утверждением 4.3 и формул (4.17) находим 6 j-инвариантов для изоморфных пар скрученных кривых Эдвардса

$$j(2,3) = j(2,11) = 18, \quad j(2,8) = j(2,15) = 12, \quad j(2,10) = j(2,12) = 16,$$

$$j(2,14) = j(2,21) = 18, \quad j(2,18) = j(2,26) = 23, \quad j(2,19) = j(2,17) = 18.$$

Все скрученные кривые Эдвардса с одинаковым j-инвариантом имеют одинаковый порядок. Но число допустимых порядков в нашем примере почти вдвое меньше числа различных вычисленных j-инвариантов. Действительно, все скрученные кривые при $p = 1 mod 4$ имеют минимальный кофактор 4 порядка кривой. В границах Хассе $[p \pm 2\sqrt{p}]$ имеются лишь 3 значения таких порядков $N_E \in \{20,\ 28,\ 36\}$. В данном примере получены такие результаты:

$N_E = 20$ при $j(2,-2) = 17$,
$N_E = 28$ при $j(2,d) \in \{16, 18\}$,
$N_E = 36$ при $j(2,d) \in \{12, 23\}$.

Подчеркнем, что кривые с одинаковым j-инвариантом не обязательно изоморфны, но всегда имеют одинаковый порядок. В то же время одинаковый порядок могут иметь кривые с разными значениями j-инвариантов и, разумеется, неизоморфные кривые.

Итак, задавая в скрученной кривой Эдвардса минимальное неквадратичное значение параметра $a = \alpha$, можно достичь максимальной производительности вычисления групповых операций и экспоненцирования точек, равной производительности вычислений на полной кривой Эдвардса с параметром $a = 1$.

4.6. Необходимые и достаточные условия делимости точки скрученной кривой Эдвардса на 2

Задача деления точки на 2 на полной кривой Эдвардса, обратная удвоению, рассматривалась нами в работах [34,35]. Для скрученной кривой Эдвардса эта задача сложнее, вместо двух корней для каждой делимой на 2

точки образуется по 4 корня. Усложняются и условия делимости на 2. Результаты нашего анализа опубликованы в работе [58]. Ниже мы приводим ее решение для скрученной кривой Эдвардса с параметрами $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) = -1$.

Пусть $P = (x_1, y_1)$, и $2P = R = (X,Y)$. В этом случае можно записать обратную удвоению (4.3) точки операцию деления точки на 2 как $(X, Y)/2 = P$. Для нециклической скрученной кривой с тремя точками 2-го порядка существуют всего 4 решения операции деления на 2: $(X, Y)/2 \in \{P, P + D_0, P + D_1, P + D_2\}$. Ясно, что удвоение любой из этих точек дает один результат $2P$. Деление на 2 точки аддитивной группы имеет аналогию с извлечением корня квадратного из элемента мультипликативной группы поля характеристики $p \neq 2$, однако вместо двух квадратных корней здесь появляется 4 корня.

Исключим из рассмотрения нейтральный элемент O, точку 2-го порядка $D_0 = (-1, 0)$ кривой (4.1) с нулевой Y-координатой. На оси y (при $x = 0$) точек скрученной кривой Эдвардса не существует. Кроме того, не рассматриваем особых точек 2-го порядка D_1 и D_2 с бесконечной Y-координатой. Для остальных точек воспользуемся формулой удвоения (4.3). Согласно (4.1) вторую координату Y в (4.3) можно выразить двумя формулами:

$$\frac{2x_1y_1}{x_1^2+ay_1^2} = Y, \qquad \frac{2x_1y_1}{1+dx_1^2y_1^2} = Y\,.$$

Обозначим $Z = y_1/x_1$, $V = y_1x_1$, причем $Z, V \neq 0$. Деление на x_1 корректно, так как не существует точек при $x = 0$. Тогда с учетом введенных обозначений для любой не особой точки P кривой, не лежащей на окружности радиуса 1, одновременно справедливы два квадратных уравнения:

$$aZ^2 - 2Y^{-1}Z + 1 = 0, \quad dV^2 - 2Y^{-1}V + 1 = 0, \quad (X,Y) \neq D_{1,2,3} \qquad (4.18)$$

с дискриминантами:

$$\Delta_1 = 4Y^{-2}\,(1 - aY^2), \qquad \Delta_2 = 4Y^{-2}\,(1 - dY^2), \quad Y \neq 0, \qquad (4.19)$$

и решениями:

$$Z_{1,2} = (aY)^{-1}(1 \pm \sqrt{1 - aY^2}), \quad V_{1,2} = (dY)^{-1}\,(1 \pm \sqrt{1 - dY^2}). \quad (4.20)$$

Эти решения имеют свойства:

$$Z_1 Z_2 = a^{-1}, \quad V_1 V_2 = d^{-1}. \tag{4.21}$$

В частности, для скрученной кривой с параметрами $\left(\frac{a}{p}\right) = \left(\frac{d}{p}\right) = -1$ одно из решений в (4.20) всегда есть квадратичный вычет, а другое – квадратичный невычет. Далее в соответствии с (4.21) полагаем, что Z_1 и V_1 – квадратичные вычеты (соответственно, Z_2 и V_2 – квадратичные невычеты).

Изложенное выше позволяет сформулировать и доказать следующую теорему.

Теорема 4.4. *Для любой точки* $(X, Y) \neq 0$, $D_{0,1,2}$ *скрученной кривой Эдвардса* (4.1) *существуют* 4 *точки деления на два* $(X,Y)/2 \in \{P, P+D_{0,1,2}\}$ *тогда и только тогда, когда выполняются условия*:

$$\text{(i)} \quad \left(\frac{1-aY^2}{p}\right) = 1, \quad \text{(ii)} \quad V_1(Z_1 - Y) = XZ_1(V_1 - Y).$$

При невыполнении любого из них точка (X,Y) *на* 2 *не делится*.

Доказательство.

Необходимость. Удвоение любой точки $(x_1, y_1) = P$ согласно закону (4.3) порождает единственную точку $2P = (X, Y)$, причем координаты точек P и $2P$ являются решениями двух квадратных уравнений (4.18) в поле F_p. Как следует из (4.19), необходимым условием существования решения первого из уравнений (4.18) является то, что элемент поля $(1 - aY^2)$ есть ненулевой квадрат в этом поле, т.е. $\left(\frac{1-aY^2}{p}\right) = 1$.. Это доказывает необходимость условия (i) теоремы. Это условие определяется лишь одной координатой точки (X, Y). Из (4.1) очевидно, что кривая имеет 2 точки $(\pm X, Y)$, из которых равенство $2P = (X, Y)$ справедливо лишь для одной. Тогда согласно (4.3) на основе вычисленных по формулам (4.20) значений Z_1 и V_1 получим

$$\frac{x_1^2 - ay_1^2}{(1-dx_1^2y_1^2)} = X \;\Rightarrow\; \frac{x_1^2(1-aZ_1^2)}{1-dV_1^2} = X \;\Rightarrow\; \left(\frac{V_1}{Z_1}\right)\frac{1-aZ_1^2}{1-dV_1^2} = X. \tag{4.22}$$

Из уравнений (4.18) справедливы равенства

$$aZ_1^2 - 1 = 2(Z_1 Y^{-1} - 1), \qquad dV_1^2 - 1 = 2(V_1 Y^{-1} - 1).$$

Тогда вместо (4.22) можно записать

$$V_1(Z_1 - Y) = XZ_1(V_1 - Y).$$

Последнее равенство определяет второе необходимое условие (ii) теоремы. По сути оно позволяет при выполнении условия (i) определить уникальное значение координаты X точки, которая делится на 2.

Достаточность. Пусть для координат точки $R = (X,Y)$ выполняются условия теоремы. Так как уравнение (4.1) можно записать в форме $(1 - aY^2) = = X^2(1 - -dY^2)$, то для любой точки из условия $\left(\frac{1-aY^2}{p}\right) = 1$ следует $\left(\frac{1-dY^2}{p}\right) = 1$. Итак, при выполнении условия (i) теоремы существуют все 4 решения (4.20) уравнений (4.18), из которых Z_1 и V_1 – квадратичные вычеты, а Z_2 и V_2 – квадратичные невычеты. Тогда существуют решения для квадратов $y_{1,2}^2 = Z_{1,2}V_{1,2}$, $x_{1,2}^2 = \frac{V_{1,2}}{Z_{1,2}}$ координат точек $R/2$.. Их корни порождают 8 точек $(\pm x_1, \pm y_1)$, $(\pm x_2, \pm y_2)$. Из них только для 4-х точек справедливы равенства $x_1 y_1 = V_1$, $x_2, y_2 = V_2$, при выполнении которых отбираются точки $P_1 = (x_1, y_1)$, $P_1^* = (-x_1, -y_1)$, $P_2 = (x_2, y_2)$, $P_2^* = (-x_2, -y_2)$. Такие же результаты справедливы при использовании равенств $Z_i = y_i/x_i$ $i = 1,2$. Как следует из (4.8), точки P_i и P_i^* связаны точкой D_0, а точки P_1 и P_2 – точкой D_1 или D_2. Итак, при выполнении условия (i) теоремы существуют 4 точки деления на два:

$$\{P_1 = (x_1, y_1), P_1^* = (-x_1, -y_1), P_2 = (x_2, y_2), P_2^* = (-x_2, -y_2)\}.$$

Удвоение любой из них дает уникальную точку $2P_i = 2P_i^* = (X, Y), i = 1,2$. Тот же результат согласно (4.8) получим для точек $\{P_1, P_1 + D_{0,1,2}\}$. Условие (ii) теоремы, записанное для точки $P_1 = (x_1, y_1)$, является достаточным, так как его выполнение тождественно выполнению такого условия для всех 4-х точек. Действительно, если выполняется условие (ii), то справедливо также равенство $V_2(Z_2 - Y) = XZ_2(V_2 - Y)$ для второй пары решений (4.20). При $\left(\frac{1-aY^2}{p}\right) = -1$ оба уравнения (4.18) решений в поле $\mathbf{F}_p$ не имеют и точек

деления на 2 не существует. При проверке второго условия теоремы точка (–*X*,*Y*), удовлетворяющая условию (i), но не отвечающая условию (ii), также на 2 не делится. Теорема доказана. ▲

Исключение точек $O, D_{0,1,2}$ из условий теоремы не означает, что для них не существует условий делимости на 2. В частности, для точек O и D_0 анализ прост. В этом случае согласно формулы удвоения (4.3) $x_1 y_1 = 0$, т.е. $x_1 = 0$ или $y_1 = 0$. Так как на оси *y* (при *x* = 0) точек согласно (4.1) нет, то $y_1 = 0$, и, следовательно, $x_1^2 = 1$. Это же равенство получаем с помощью формулы удвоения (4.3) для *X*-координаты точки O = (1,0), тогда как для *X*-координаты точки $D_0 = (-1,0)$ получаем $x_1^2 = -1$. Итак, точка O делится на 2, а точка D_0 – не делится. Второе справедливо для кривой (4.1) в связи с отсутствием точек на оси *y*.

Интересно, что тот же вывод следует для этих точек при проверке условий теоремы 4.4. Действительно, простейшей точкой, имеющей 4 корня, является нейтральный элемент $O = (1,0)$, для которой выполняются $\left(\frac{1}{p}\right) = 1$ (условие (i)) и равенство $V_1 Z_1 = Z_1 V_1$ (условие (ii)). Здесь точки деления на 2 равны $(1,0)/2 \in \{O, D_{0,1,2}\}$. Для точки D_0 условие (i) $\left(\frac{1}{p}\right) = 1$ выполняется, а условие (ii) приводит к $V_1 Z_1 = -Z_1 V_1$. Согласно (4.21), все сомножители здесь ненулевые и равенство не выполняется.

Точки D_1, D_2 лежат за пределами теоремы 4.4, но условия делимости их на 2 найдены нами в теореме 2.1.

Утверждение 4.4. *Для скрученной кривой Эдвардса порядка 4*n *с простым* n *на* 2 *делятся лишь точки подгруппы <G> порядка* n*, причем одна из точек* $\frac{G}{2}$ *имеет порядок* n*, и три точки – порядок* 2n*.*

Доказательство. Нециклическая скрученная кривая Эдвардса с тремя точками второго порядка содержит ровно (*n* – 1) точек простого порядка *n* и 3(*n* – 1) точек порядка 2*n*. Следовательно, существуют ровно 3 точки порядка 2*n*, удвоение которых даст уникальную точку подгруппы <G> порядка *n*. В подгруппе < G > простого порядка *n* деление на 2 точки тождественно умножению точки на единственный элемент $2^{-1} = \frac{n+1}{2}$ поля F_n, что

порождает при делении на 2 одну точку порядка n. Остальные три точки $\frac{G}{2}$ имеют порядок $2n$. Утверждение доказано. ▲

В следующей теореме определяется новое свойство обеих координат точки скрученной кривой Эдвардса.

Теорема 4.5. *Для любой точки* $(X, Y) \neq 0, D_{0,1,2}$ *кривой* (4.1) *справедливо равенство*

$$\left(\frac{1-X^2}{p}\right)\left(\frac{1-aY^2}{p}\right) = \left(\frac{a-d}{p}\right).$$

Доказательство.

Исключим из рассмотрения точки 2-го порядка $D_{0,1,2}$ и точку О. Для всех других точек (X, Y) с учетом определения (4.1) запишем произведение

$$(1 - X^2)\,(1 - aY^2) = 1 - X^2 - aY^2 + aX^2Y^2 = (a - d)\,X^{2\,2}Y^2\,.$$

Отсюда сразу следует, что произведение $(1 - X^2)(1 - a\,Y^2)$ является квадратичным невычетом при $\left(\frac{a-d}{p}\right) = -1$ и наоборот, что и доказывает утверждение теоремы. Для точек О и D_0 имеет место равенства $X^2 = 1,\ Y = 0,$ при этом равенство в утверждении теоремы не выполняется. Для особых точек $D_{1,2}$ координата $Y = \infty$, и равенство в утверждении теоремы не определено. Для остальных точек теорема доказана. ▲

Следует заметить, что теорема 4.5 справедлива не только для класса скрученных кривых, но и для всех кривых (2.1) в обобщенной форме Эдвардса. Эта теорема позволяет заменить тестирование величины $(1 - aY^2)$ тестированием значения $(1 - X^2)$ точки (X,Y). Докажем два следствия трех теорем: теоремы 2.1, 4.4 и 4.5.

Следствие 4.1. *Достаточным условием того, что точка* $(X, Y) \neq 0, D_{0,1,2}$ *скрученной кривой Эдвардса почти простого порядка* 4n *является точкой максимального порядка* 2n*, является условие*

$$\left(\frac{1-X^2}{p}\right) = -\left(\frac{a-d}{p}\right). \qquad (4.23)$$

Доказательство.

Из теоремы 2.1 следует, что порядок скрученной кривой (4.1) при $p = 1 mod 4$ равен $4n$ (n – простое по условию), а максимальный порядок точки равен $2n$. Такая точка не делится на 2 (утверждение 4.4) и не выполняется условие (i) теоремы 4.4, т.е выполняется $\left(\frac{1-aY^2}{p}\right) = -1$ или $\left(\frac{1-aY^2}{p}\right) = 0$. Второе равенство исключается, так как не существует элемент поля $\sqrt{a}$. Исключаются также точки O и точки второго порядка D_i, для которых $X^2 \in \{1, \frac{a}{d}\}$. Тогда согласно теореме 4.5 и с учетом $\left(\frac{1-aY^2}{p}\right) = -1$ для точек максимального порядка $2n$ достаточно выполнить условие $\left(\frac{1-X^2}{p}\right) = -\left(\frac{a-d}{p}\right)$. Следствие 4.1 доказано. ▲

Условие следствия 4.1 всегда порождает точки порядка $2n$ и потому является достаточным, но не необходимым. При невыполнении этого условия половина точек также имеет порядок $2n$.

Следствие 4.2. *Семейство точек* (±*X*, ±*Y*) *скрученной кривой Эдвардса почти простого порядка* 4n *содержит* 2 *точки порядка* n *и* 2 *точки порядка* 2n *тогда и только тогда, когда выполняется условие*

$$\left(\frac{1-X^2}{p}\right) = \left(\frac{a-d}{p}\right). \qquad (4.24)$$

Доказательство. *Необходимость.* Согласно условию (i) теоремы 4.4 и теореме 4.5 условие (4.24) является необходимым условием делимости точки на 2. Рассмотрим 2 точки семейства $R = (X, Y)$ и $-R^* = (-X, Y) = -(R + D_0)$. Необходимое условие (ii) теоремы 4.4 выполняется лишь для одной из этих точек, поэтому одна из них имеет порядок n, а другая – порядок $2n$.

Достаточность. Пусть выполняется условие теоремы. Тогда согласно теореме 4.4 существуют решения для точек деления на 2 и одна из точек (X, Y) или $(-X, Y)$ делится на 2 и, следовательно, имеет порядок n, другая – порядок $2n$. Обратные точки этой пары имеют те же порядки n и $2n$, что отвечает утверждению 4.4. Следствие 4.2 доказано. ▲

Ясно, что для половины всех точек $(X, Y) \neq 0, D_{0,1,2}$ выполняется условие (4.23), для другой половины – условие (4.24). Только тогда точки порядка n составляют четверть всех точек больших порядков (утверждение 4.4).

4.7. **Методы определения порядков точек скрученной кривой Эдвардса**

Как и везде в этой работе, мы ищем для криптосистем кривые Эдвардса с минимальным кофактором 4 порядка кривой $N_E = 4n$, где n – достаточно большое простое число. Такое значение N_E называют почти простым числом. Для скрученных кривых Эдвардса все кривые имеют минимальный кофактор 4 при $p \equiv 1 \bmod 4$, остается лишь подобрать кривую с простым значением n. Далее полагаем, что n – простое число. Особенностью этих кривых является то, что максимальный порядок точки равен $2n$ вместо $4n$ для циклических кривых. Вместе с тем число точек порядка $2n$ втрое больше, чем порядка n. Для нахождения генератора криптосистемы как точки $G = 2P$ достаточно удвоить любую (кроме точек D_i) случайную точку кривой, что требует выполнения одной групповой операции. Альтернативным может быть метод, основанный на свойстве делимости точки на 2 [58].

Согласно утверждению 4.4, для точек максимального порядка $2n$ кривой Эдвардса не существует точек деления на 2, но они существуют для точек порядка n. Другими словами, если для случайной точки $R = (X, Y)$ выполняется условие (4.23) следствия 4.1, то порядок такой точки равен $2n$. В противном случае (с вероятностью 1/2) выполняется условие (4.24), тогда порядок точки равен n или $2n$. При выполнении обоих условий теоремы 4.4 порядок точки равен n. Ее можно принять генератором G криптосистемы.

Таким образом, для определения порядка точек скрученной кривой Эдвардса вовсе не требуется выполнять сложную операцию скалярного произведения nR, что рекомендуется существующими стандартами. Точка (X,Y) максимального порядка $2n$ определяется в соответствии со следствием 4.1 лишь одним достаточным условием (4.23) следствия 4.1.

Итак, вместо в среднем $1.5\log(n)$ групповых операций удвоения и сложения при вычислении скалярного произведения nR [29], со сложностью порядка 10 операций в поле каждая, требуется выполнить всего 2 полевые операции: возведение в квадрат S и нахождение символа Лежандра. Большой выигрыш в экономии вычислений очевиден. Этот же метод для полных кривых Эдвардса описан в нашей работе [51].

Для определения генератора G криптосистемы порядка n как точки скрученной кривой Эдвардса порядка $4n$ можно использовать следующие методы:

M1. Вычислить скалярное произведение nR случайной точки R, тогда при $nR = O$ искомая точка найдена: $R = G$.

M2. Удвоить случайную точку $R \neq D_i$ тогда $2R = G$.

M3. Проверить условия (i) и (ii) теоремы 4.4 делимости точки R на 2, при их выполнении $R = G$.

M4. Проверить условие (4.23) следствия 4.1. При его выполнении вычислить $G = R + D_1$.

Метод M1 используется во всех стандартах эллиптической криптографии и является очень трудоемким. Он требует $\log(n)$ удвоений точки R со сложностью $T_E = 6.17M$ и в среднем $0.5\log(n)$ сложений точек со сложностью $V_E = 11.67M$ (раздел 4.4.2) с суммарной сложностью экспоненцирования точки $S1 = \log(n)(T_E + 0.5V_E) = 12\text{M}\log(n)$. Предлагаемый здесь метод удвоения M2 имеет сложность $T_E = 6.17M$, что дает выигрыш в производительности вычисления генератора криптосистемы в $\frac{S1}{T_E} = 1.95\log(n)$ раз по сравнению со стандартным. Например, для модуля длиной в 300 бит этот выигрыш близок к 600. Важно, что выигрыш в производительности экспоненцирования точки пропорционален длине модуля $\log(n)$.

Приведенные выше оценки справедливы при выполнении групповых операций в проективных координатах, не привлекающих инверсии элементов поля. Сложность метода М1 в аффинных координатах с учетом 2-х инверсий в каждой групповой операции приблизительно оценивается временем работы $4\log^4(p)$, где $p \approx n$ [55].

Метод М3 основан на свойстве делимости точки R на 2 (теорема 4.4). Вычисления по формулам (4.20) и условию (ii) теоремы оцениваются сложностью $S3 = 1L + 1S + 8M + 2I + 2RT$, где составляющие обозначают оценки сложности вычисления символа Лежандра (L), возведения в квадрат (S), умножения (M), инверсии (I) и извлечения корня квадратного (RT) соответственно. Три из этих полевых операций (L, I, RT) имеют суммарную сложность порядка $5\log^3(p)$ [23], а умножение M и возведение в квадрат S – сложность порядка $\log^2(p)$. Удвоение же точки в проективных координатах имеет сложность всего $T_E = 6.17M$. Очевидно, что по сравнению со вторым методом метод М3 существенно проигрывает в вычислительной сложности. Вместе с тем по сравнению с М1 с вычислениями в аффинных координатах время работы уменьшается в $4\log^4(p)/ 5\log^3(p) = 0.8\log(p)$ раз.

Если ограничиться условием (4.24) (условие (i) теоремы 4.4), то точка $R = (X, Y)$ имеет порядок n с вероятностью ½. Остается проблема выбора генератора из двух точек $(\pm X, Y)$ с порядками n и $2n$.

Метод М4 является альтернативой удвоению точки. Он позволяет найти точку порядка n с помощью точки R порядка $2n$ (что всегда имеет место при выполнении (4.23)), для которой $nR = D_1$, тогда $R + D_1 = (n + 1)R$ является точкой порядка n. Как следует из (4.8), для вычисления координат этой точки потребуется выполнить $(1L + 2I + 2M)$ полевых операций, что вместе с проверкой квадратичности делает этот метод менее эффективным, чем метод удвоения М2. Если, однако, удвоение точки выполнять в аффинных координатах согласно (4.3) со сложностью $(2I + 3M + 3S)$, также требующего двух инверсий, то методы М2 и М4 соизмеримы по сложности.

Пример 4.3. При p = 17 скрученная кривая Эдвардса $x^2 + 3y^2 = 1 + 6x^2y^2$ имеет порядок N_E = 20 (здесь a = 3 и d = 6 – квадратичные невычеты). Кривая отвечает условиям теоремы 2.1, имеет 3

точки 2-го порядка $D_0 = (-1,0)$, $D_1 = (-3,\infty)$ и $D_2 = (3,\infty)$ и не имеет точек 4-го порядка. Она содержит 4 точки 5-го и 12 точек 10-го порядков. Точечный график этой кривой изображен на рис.4.1(a). Особые точки $D_{1,2}$здесь имеют бесконечные y-координаты, а их x-координаты обозначены кружками. Прямые суммы подгруппы точек 5-го порядка с тремя подгруппами 2-го порядка образуют 3 подгруппы точек 10-го порядка. Они представлены как точки скалярных произведений kP_0, kP_1 и kP_2 в таблице 4.4. Здесь индексы генераторов подгрупп P_i совпадают с индексами точек второго порядка D_i, так что $D_i = 5P_i$, i = 0,1,2. Каждый из генераторов образует одну циклическую подгруппу 10-го порядка на колесе точек, рис.4.1(b). Всего для нашей нециклической кривой можно построить 3 таких колеса с разными генераторами P_i, координаты которых занимают первый столбец таблицы 4.4 (при k = 1). Кривая имеет тип (2,2,5) [29] и может быть представлена прямой суммой циклических подгрупп G_2 и G_{10} 2-го и 10-го порядков соответственно, включающих разные точки 2-го порядка.

(a) (b)

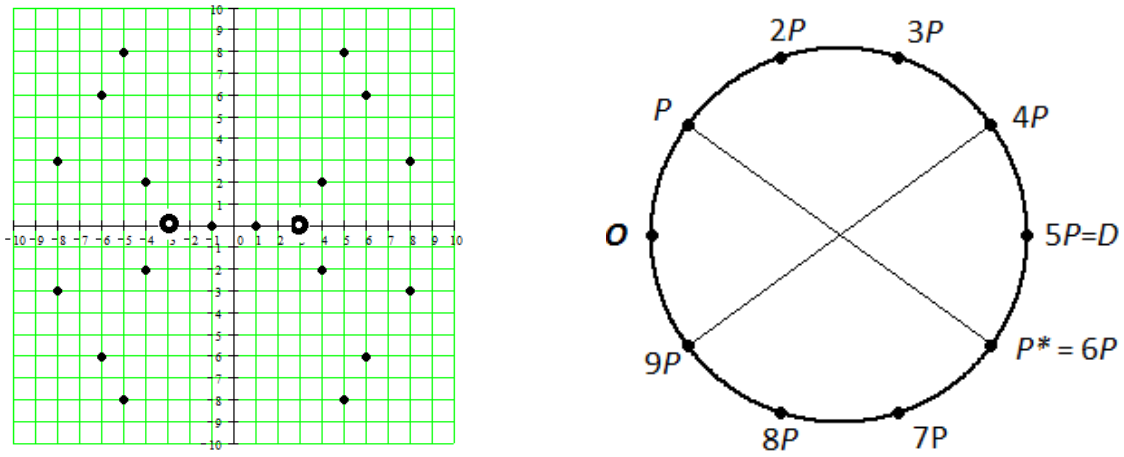


Рис.4.1. График точек кривой $x^2 + 3y^2 = (1 + 6x^2y^2) mod 17$ (a) и колесо точек скалярного произведения этой кривой kP, $k = 0..9$ (b)

Таблица 4.4. Координаты точек $kP_{1,2}$ циклических подгрупп 10-го порядка кривой Эдвардса $x^2 + 3y^2 = (1 + 6x^2y^2) mod 17$.

	k	1	2	3	4	5	6	7	8	9	10

kP_0	x_k	5	–8	8	–5	–1	–5	8	–8	5	1
	y_k	–8	3	3	–8	0	8	–3	–3	8	0
kP_1	x_k	4	–8	–6	–5	–3	–5	–6	–8	4	1
	y_k	2	3	6	–8	∞	8	–6	–3	–2	0
kP_2	x_k	–4	–8	6	–5	3	–5	6	–8	–4	1
	y_k	–2	3	–6	–8	∞	8	6	–3	2	0

Пусть $P_1 = (4,2),\ G = 2P_1 = (-8,3)$. Точка G имеет 5-й порядок и, как и все 4 точки 5-го порядка, делится на 2, образуя 4 корня 2-й степени $\frac{G}{2} \in \{P_1, P_1 + D_{0,1,2}\}$. Действительно, согласно (4.19) и условию (i) теоремы 4.4 $\Delta_1 = (1 - 3\cdot 3^2) = 8$ – квадратичный вычет, тогда и $\Delta_2 = (1 - 6\cdot 3^2) = 15$ – - также квадратичный вычет, и существуют решения (4.20) $Z_{1,2} = (3)^{-2}(1 \pm 7) \in \{-8, -5\}$, $V_{1,2} = (1 \pm 7) \in \{8, -6\}$. Здесь Z_1 и V_1 – квадратичные вычеты, а Z_2, V_2 – квадратичные невычеты в поле F_{17}. Легко проверить выполнение условия (ii) теоремы 4.4 : $8(-8-3) = (8 — 3) \Rightarrow 14 = 14$. Определяем далее квадраты $Z_1 V_1 = y_1^2 = 4, Z_2 V_2 = y_2^2 = 13$, $\frac{V_1}{Z_1} = x_1^2 = 16$, $\frac{V_2}{Z_2} = x_2^2 = 8$. Отсюда $(\pm x_1, \pm y_1) = (\pm 4,\ \pm 2)$, $(\pm x_2, \pm y_2) = (\pm 5,\ \pm 8)$. Из этих 8 точек отбираем 4 точки, для которых $x_1\, y_1 = V_1 = 8,\ x_2\, y_2 = V_2 = -6,$ и получаем точки $P_1 = (4,2)$, $P_1^* = (-4,-2),\ P_2 = (5,-8), P_2^* = (-5,8)$. Из них 3 точки 10-го порядка, и одна точка $(-5,8)$ – 5-го порядка. В поле F_5 деление на 2 есть умножение на $2^{-1} = 3$, поэтому данная точка $(-5,8) = 3G$. Это видно и из таблицы 4.4, из которой также легко определяются все 4 точки деления на 2 или корни квадратные $\frac{G}{2} = \{(4,2), (-4,-2), (5,-8), (-5,8)\}$.

Для нахождения порядка случайной точки, например, $R = (4,2)$, согласно условий (4.23), (4.24) вычисляем $(1 - 4^2) = 2$ – квадратичный вычет. Так как (a – d) = 3 – 6 = 14 – квадратичный невычет, условие (i) теоремы 4.4 не выполняется и точки $(4,2),\ (-4,-2)$ имеют максимальный порядок $2n$ =10. Это отвечает следствию 4.1. Все точки с X-координатами, для которых $(1 - X^2)$ – квадратичный вычет, имеют порядок 10. Это имеет место для

$X = \pm 4$ и $X = \pm 6$. Эти точки, как следует из таблицы 4.4, включены в 2 подгруппы, генерируемые точками P_1 *и* P_2.

Альтернативный случай, если $(1 - X^2) -$ квадратичный невычет, отвечает условию (4.24) следствия 4.2. В нашем примере это выполняется для значений координат $X = \pm 5$ и $X = \pm 8$. Все точки, удовлетворяющие условию (4.24), входят в подгруппу таблицы 4.4, генерируемую точкой P_0 и содержащей точку D_0. Для этого случая порядок случайной точки (X,Y) с равной вероятностью имеет порядок 5 или 10. В таблице 4.4 точка $(5,8)$ имеет порядок 10, а точка $(-5,8) -$ порядок 5.

Точки колеса, соединенные диаметральными линиями, связаны как P_i и $P_i + D_i$. Для любой пары точек $\pm P_i$ имеется пара связанных с ними линиями на рис.4.1(b) точек. Порядки точек этих пар могут совпадать и быть равными $2n$, либо отличаться в 2 раза (быть равными n и $2n$). Первое имеет место для двух подгрупп $<P_{1,2}>$, а второе – для подгруппы $<P_0>$.

Чтобы заполнить таблицу 4.4, содержащую 30 точек и 60 значений координат, не требуется вычислять все точки, надо лишь знать их свойства. Сначала можно определить, например, точки деления на 2: $\frac{(-8,3)}{2} = \{(4,2), (-4,-2), (5,-8), (-5,8)\}$. Первые 3 из них есть точки 10-го порядка, они записываются в столбец таблицы при k = 1. Так как удвоение каждой из них дает одну точку $(-8,3)$ 5-го порядка, во 2-й столбец таблицы при k = 2 записываем координаты этой точки. Точки 9-го и 8-го столбцов таблицы обратны точкам 1-го и 2-го столбцов, поэтому координаты последних переписываются сюда с инверсией знаков y-координат. Далее мы пользуемся свойствами сумм точки (x_1, y_1) с точками 2-го порядка (4.8):

$$(x_1, y_1) + (-1, 0) = (-x_1, -y_1),$$

$$(x_1, y_1) + \left(\pm\sqrt{\frac{a}{d}}, \infty\right) = \left(\pm\sqrt{\frac{a}{d}} \cdot x_1^{-1},\ \pm\frac{1}{\sqrt{ad}} \cdot y_1^{-1}\right).$$

Точку $D_0 = (-1, 0)$ содердит лишь верхняя циклическая подгруппа точек в таблице 4.4, так как две нижние порождены точками с инверсией знаков обеих координат и, следовательно, содержат особые точки $(\pm 3, \infty)$ 2-го порядка. В 5-й столбец таблицы переписываем координаты соответствующих точек 2-го

порядка. Так как $P_0 + D_0 = 6P_0$ и $2P_2 + D_0 = 7P_2$, то в 6й и 7-й столбцы верхней подгруппы таблицы переписываем координаты первых двух с инверсией обоих знаков. Обратные им точки определяются, соответственно, при $k = 4$ и $k = 3$, где меняются знаки y-координат. Итак, верхнее колесо точек kP_0 полностью определено. При четных k в два нижних ряда точек kP_1 и kP_2 переписывается те же координаты точек 5-го порядка. Для этих двух подгрупп остаются не заполненными лишь 3-й и 7-й столбцы. Так как $3P_1 = 8P_1 + D_1$, то мы вычисляем координаты точки $3P_1$ 10-го порядка как

$$(-8,-3) + (-3,\infty) = (-3 \cdot x_1^{-1}, -y_1^{-1}) = (-6,6).$$

В этом же столбце при $k = 3$ для второй подгруппы точек вписываем точку $3P_2 = (6,-6)$ с инверсией обоих знаков, а при $k = 7$ для обеих подгрупп записываем координаты обратных точек $7P_1 = -3P_1$ и $7P_2 = -3P_2$.

Для больших полей рассмотренный пример можно обобщить на метод, позволяющий при известном сегменте из ¼ точек kP_i кривой найти практически без вычислений координаты всех других точек. Ясно, что такая возможность объясняется наличием на кривой семейств из 4-х точек ($\pm x_1, \pm y_1$). Наибольший практический интерес из циклических подгрупп порядка $2n$, представляет подгруппа kP_0, содержащая точку $nP_0 = D_0$. Для всех ее точек выполняется условие (4.24) следствия (4.2). В этой подгруппе выполняется свойство полноты закона сложения, так как она не содержит особых точек. По своим свойствам она близка к полной кривой Эдвардса с тем отличием, что в ней нет точек порядка 4 и $4n$.

Рассмотрим некоторые свойства семейств из 4-х точек ($\pm x_1, \pm y_1$) кривой (4.1) порядка $4n$ (n – простое), лежащих на одной окружности радиуса $\sqrt{x_1^2 + y_1^2} \neq 1$.

Из колеса циклической подгруппы точек kP_0 на рис.4.1(b) становится ясно, что для скрученных кривых порядка $4n$ с максимальным порядком точки $2n$ любое семейство из 4-х точек ($\pm x_1, \pm y_1$) содержит по 2 точки порядков n и $2n$ (они связаны диаметральными линиями). Такие точки имеют обратные знаки x-координат, так как инверсия знака у-координаты дает обратную точку того же порядка.

Колесо циклической подгруппы точек kP_1 на рис.4.1(b), включающее особую точку $nP_1 = D_1$, связывает линиями 2 точки порядков n и $2n$, не входящих в одно семейство ($\pm x_1$, $\pm y_1$). Для подгрупп kP_1 и kP_2 все точки семейства имеют одинаковый порядок $2n$. В соответствии с формулами (4.8) суммирование с точкой D_1 инвертирует с весом координаты точки (x_1, y_1). Например, согласно таблице 4.4 $(4,2) + (-3, \infty) = (-5,8)$, т.е. точка 10 порядка $(4,2)$ преобразуется в точку 5-го порядка $(-5,8)$. Это является идеей метода М4.

Заметим, что точки порядка n не образуют семейств ($\pm x_1$, $\pm y_1$). Таким образом, в любом семействе точек либо все точки имеют порядок $2n$ (при условии (4.23)), либо семейство точек содержит по 2 точки порядка n и $2n$ с разными знаками x-координат (при условии (4.24)). При известной точке максимального порядка P_i можно найти точку порядка n двумя способами: удвоением этой точки $2P_i$ или сложением ее с точкой 2-го порядка $P_i + D_i = (n+1)P_i$.

Докажем следующее утверждение.

Утверждение 4.5. *Число семейств точек* ($\pm x_1$, $\pm y_1$) *скрученной кривой Эдвардса почти простого порядка* $4n$, *содержащих точки простого порядка* n, *равно числу семейств, не содержащих таких точек и равно* $(n-1)/2$.

Доказательство. В семейства точек кривой (4.1) не входят точки O и три точки D_i 2-го порядка, так что остается $4(n-1)$ точек, входящих в $(n-1)$ семейство. Так как нециклическая кривая почти простого порядка $4n$ содержит ровно $(n-1)$ точку порядка n и $3(n-1)$ точек порядка $2n$ (теорема 4.4), то число семейств, содержащих точки порядка n, равно $(n-1)/2$. Утверждение доказано. ▲

Возвращаясь к примеру 4.3 и таблице 4.4, мы видим, что условие (4.24) следствия 4.2 определяет семейства точек, входящих в подгруппу, генерируемую точкой $P_{0.}$ Если известно, что точка $P_{0.} = (5,-8)$ имеет порядок $2n$=10, то генератор порядка n=5 легко находится как точка $P_{0.} + D_{0.} = 6P_{0.} = (-5,8)$, входящая в то же семейство с инверсией координаты X.

Несмотря на взаимосвязь 4-х точек ($\pm x_1, \pm y_1$), включающих точки порядка n, сложность вычисления дискретного логарифма в подгруппе точек $<G>$ простого порядка n не снижается. Как и для кривых в форме Вейерштрасса, она снижается лишь вдвое за счет обратных точек.

Можно заключить, что циклические подгруппы простого порядка n скрученных кривых Эдвардса порядка $4n$ при $p \equiv 1 mod 4$. не только приемлемы, но и могут быть предпочтительными для использования в криптосистемах. Эти подгруппы обходят особые точки кривой, в них имеет место универсальность и полнота закона сложения точек. Они удобны для программирования и, как и полные кривые Эдвардса, имеют минимальную вычислительную сложность (при минимальном значении параметра $a = \alpha$). Свойства делимости точки на два на таких кривых позволяет находить точки заданного порядка в сотни раз быстрее, чем предлагают стандартные алгоритмы.

В сравнении с полными кривыми Эдвардса, скрученные кривые не содержат точек порядка $4n$, что позволяет найти генератор простого порядка n всего одной операцией удвоения случайной точки без всяких предварительных вычислений. Это дает выигрыш приблизительно в $2\log(n)$ раз (т.е. в сотни раз) по сравнению со стандартным экспоненцированием точки. При $p \equiv 1 mod 4$ все скрученные кривые Эдвардса имеют порпядок $4n$, что упрощает поиск криптостойких кривых. Альтернативный закон сложения точек (4.9) полезен при сложении разных точек нечетного порядка и практически не дает выигрыша в быстродействии при вычислении скалярного произведения. По сравнению с эллиптической кривой в форме Вейерштрасса быстродействие операции экспоненцирования точки на скрученной кривой Эдвардса с минимальным параметром a возрастает приблизительно в 1.5 – 1.6 раза.

4.8. Результаты расчета общесистемных параметров криптостойких скрученных кривых Эдвардса с минимальной сложностью

Авторским коллективом работы [66] была поставлена и успешно решена задача поиска и расчета общесистемных параметров приемлемых для стандартизации скрученных кривых Эдвардса над простым полем.

В данном разделе мы рассматриваем простые поля с модулями длиной 192, 224, 256, 384 и 521 бит, которые рекомендуются стандартом FIPS–186–2-2000 [78] (и его последней версии FIPS–186–4-2013 [79]), и приводим перечень параметров скрученных кривых Эдвардса почти простого порядка $N_E = 4n$ (n – простое) над каждым из полей. Результаты расчетов общесистемных параметров кривых в шестнадцатеричной системе чисел сведены в таблицы 3 – 7. Здесь модули длины L обозначены как p_L. Модули полей $p \equiv 1 mod 4$ выбирались как простые числа с малым двоичным весом Хэмминга 3..5. Для каждой кривой приведены значения p, порядки $n = N_E/4$ генератора G криптосистемы и его координаты (x_G, y_G), а также значения параметров a и d.

Надо заметить, что параметр $a = 2$ является квадратичным невычетом лишь при $p = \pm 3 mod 8$. Это значит, что двоичное представление числа p заканчивается тремя младшими разрядами $101 = 5_{10}$ или $011 = 3_{10}$, а все более старшие разряды дают $0 mod 8$. В нашем алгоритме случайного поиска простых чисел с малым весом лишь одно значение $p = 2^{255} + 2^{38} + 2^2 + 1$ в таблице 5 отвечает этому условию (здесь $a = 2$), поэтому практически все кривые имеют минимальный параметр $a = 3$.

Проверка чисел p и n на простоту производилась с помощью тестов Миллера-Рабина и Лукаса-Лемера, реализованных в языках программирования C#, Java и в системе Wolfram Mathematica.

Вычисление символов Лежандра для нахождения подходящих параметров a и d производилось с помощью библиотечных функций языка Java и системы Wolfram Mathematica.

Порядки эллиптических кривых рассчитывались по алгоритму SEA (Schoof- Elkies -Atkin), реализованном в библиотеке PARI/GP.

Генераторы G порядка n были найдены удвоением случайной точки, удовлетворяющей уравнению (4.1), с использованием системы Wolfram Mathematica и языка Java. Одного удвоения достаточно, так как на

нециклической скрученной кривой порядка $4n$ максимальный порядок точки равен $2n$.

В каждой из приведенных ниже таблиц 4.3 – 4.7 содержатся параметры 5 скрученных кривых Эдвардса с минимальным значением параметра $a = 2, 3$ или 5. Далее параметр d подбирался как наименьшее из значений, при котором порядок кривой $4n$ становился почти простым. Порядок кривых по длине сравним с длиной поля.

Таблица 4.3. Скрученные кривые Эдвардса почти простого порядка над полем с модулем p_{192}

$p = 2^{191} + 2^{16} + 2^{12} + 1$ p = 8000000000000000000000000000000000000000011001 n = 2000000000000000000000004447D62B952200D604D2BEB9 a = 3 d = DD x_G = 79FD21DC7BF961D56CDD092798C3016F11C2034F154A94B5 y_G = 17B25C0DADFD379FF3D7CD1EAEFD83328B945410BE11D5DB
$p = 2^{191} + 2^{45} + 2^{41} + 1$ p = 8000000000000000000000000000000000220000000001 n = 1FFFFFFFFFFFFFFFFFFFFFFEAA490B04DCC3760DB69D80B a = 5 d = B7 x_G = 41D4C37EE995324B600442F0DCE31A0EA61C08C92D80110B y_G = 3EC06FDFB33AFA858E30D898F9C18606CA62E79F6175FD80
$p = 2^{191} + 2^{46} + 2^{36} + 1$ p= 8000000000000000000000000000000000401000000001 n = 1FFFFFFFFFFFFFFFFFFFFFFC219C1BBD9EF98692F7D3077 a = 3 d = 4D x_G = 50707DB9B16BF11C894E2A2F25A2887A6AB3A5921E9AACDE y_G = 1BEF37EF93DBA0D7203F96BD14FCDA0E6B1CF118866982B2
$p = 2^{191} + 2^{49} + 2^{27} + 1$ p = 8000000000000000000000000000000002000008000001 n = 1FFFFFFFFFFFFFFFFFFFFFFD4622687DA3DDCBF47806983 a = D d = AC x_G = 3BFDC9D07301A8A3A9BEC18540B0EEDC0F7C3CD21F652EFE y_G = 3C48E70B7F04C446A8CC208696DA2592A56FB29C79888D52
$p = 2^{191} + 2^{158} + 1$

p = 8000000040000000000000000000000000000000000001 n = 2000000010000000000000004DE2B37DC449E40E33C414B9 a = 5 d = 800000003FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFC0F x_G = 209295DA12CEDAE57617AF4911C57DDCE7043EB18687E13C y_G = 3847775699DF8A7F431D8DA8FE993A28A6D4F108B6502917

Таблица 4.4 Скрученные кривые Эдвардса почти простого порядка над полем с модулем p_{224}

p= $2^{223} + 2^{24} + 2^{20} + 1$ p = 8000000000000000000000000000000000000000000001100001 n = 2000000000000000000000000003F330C2860F36EC9E831F641A2B9 a = 3 d = 93 x_G = 6C7CC9C10F4259CBEB0D1973AF0E4FC64AE442A301A90DFEEB5BC081 y_G = 2D4F15BA1A686CAEDB9D43F9525BF78683DAA39B82301FCA8A7874BE
p = $2^{223} + 2^{38} + 2^{36} + 1$ p = 8000000000000000000000000000000000000000005000000001 n = 1FFFFFFFFFFFFFFFFFFFFFFFFFFFBCE784913191A6362578E1CD28C9 a = 3 d = 215 x_G = 60536D73F2A4EF1F54C1048734301E01306FF7F331719201335D5A55 y_G = DDBFD9D1AFD09CD322F639F524CC60F9A7A727139F56BBF8D127940
p = $2^{223} + 2^{61} + 2^{41} + 1$ p = 8000000000000000000000000000000000000002000020000000001 n = 2000000000000000000000000002165B89B7402F30BF010CB396EA9 a = 5 d = 3D x_G = 59507C9FEF622459507C9FEF622459507C9FEF623AAD7109DDBA6AF7 y_G = 1366A5B9781878270245AAA111E53CDC079B0322CB4A8C805309452A
p = $2^{223} + 2^{72} + 2^{20} + 1$ p = 8000000000000000000000000000000000001000000000000100001 n = 1FFFFFFFFFFFFFFFFFFFFFFFFFFE0DAC7710567C631D4CA120783E9 a = 3 d = 15C x_G = 5C611714E8FD05D966F07BD978DF524642C21CF3BDBB6BA1FE037DD8 y_G = 3499A41BF2C767BD41A045CDB7285F9E4907598421C0B78D29D5A7E0
p = $2^{223} + 2^{66} + 2^{14} + 1$ p = 8000000000000000000000000000000000000040000000000004001 n = 2000000000000000000000000014E4DE701954CB43404E060EAC01 a = 3 d = 26

x_G = 2A795ABA13A5D9DC2BEB32468049F7E8E287393711EA0A66DCBFF040 y_G = 532C172BF052220CDE0B63A2F421DB65E8A676B96D7FA0307DDC0E53

Таблица 4.5. Скрученные кривые Эдвардса почти простого порядка над полем с модулем p_{256}

p = $2^{255} + 2^{46} + 2^{42} + 1$ p = 8000000000000000000000000000000000000000000000000000440000000001 n = 2000000000000000000000000000000330A60EA43DF5957AC18C44AB8122EB7 a = 3 d = 1AF x_G = 25091FBF427205B62204FD7FE48236752C1FE497EF2DB3197938BA36D9A27554 y_G = 52A7749533218F16BB1F32137CE731DC0F26149CE2E1CF1378B74E066B952532
p = $2^{255} + 2^{41} + 1$ p = 8000000000000000000000000000000000000000000000000020000000001 n = 200000000000000000000000000002498CDC14EB2676199F9EFB8C86EA9D1 a = 3 d = BC x_G = 1306A0056F9B6F44758D8146286E140B8D2A4C7179CCB2B515E9EAE4A679F81 y_G = 33CD1A3853C6059B39DE2485F320CC00D97E1BB77D5C79025BF01D1F3C0D8868
p = $2^{255} + 2^{66} + 2^{60} + 1$ p = 8000000000000000000000000000000000000000000000041000000000000001 n = 200000000000000000000000000002BC056BB954BD3CF60CD27D370FF3265 a = 3 d = 6C x_G = 7E5D3187BA7FF18EF3066E57C722DCE95279A01945D6B4C2E918B56C32FF35D1 y_G = 3AC8B7A5A64DA05FF1F28870506E451F103DA6EE32FAB89D3D903E073660572E
p = $2^{255} + 2^{38} + 2^{2} + 1$ p = 8000000000000000000000000000000000000000000000000004000000005 n = 2000000000000000000000000000031F23720CCCC83EF9F44858B6E952E4F a = 2 d = 1CC

x_G = 25091FBF427205B62204FD7FE48236752C1FE497EF2DB3197938BA36D9A27554 y_G = 52A7749533218F16BB1F32137CE731DC0F26149CE2E1CF1378B74E066B95253
p = $2^{255} + 2^{70} + 2^{66} + 1$ p = 8000000000000000000000000000000000000000000044000000000000001 n = 2000000000000000000000000000002B81184994F3F0880BB2B78314AFD2E9 a = 3 d = 13A x_G = 1C32148FB9F80F78D2E84879553296DA0858C29EFEAD3EDCCBFC55E8FEE15C85 y_G = 2D4ED6C59052B9113E76AEDACBC668F3BB0EE5819C5B5221A311118DECAAE65E

Таблица 4.6. Скрученные кривые Эдвардса почти простого порядка над полем с модулем p_{384}

p = $2^{383} + 2^{155} + 1$ p = 800000000000000000000000000000000000000000000000000000800000000000000000000000000001 n = 1FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFE935B9B743CFBDC890CDB6FC507DB7744384DF7B5FAF81D1 a = 3 d = EC x_G = 5B83FE601E531F45B83FE601E531F45B83FE601E531F45B83FE601E537AC9B6A004550477AC9B6A004550477AC9B69FF y_G = 5FF608B37764C7FE1719D75DA24646D8869E60637F0C477A548C46BDECB0270BF2C840E083654F63367D737325053472
n = 1FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFD2E846F4DD24CB4DC3014E1B9B39C56950B3F5FF1870523 a = 3 d = 10A x_G = 2301C9793214FC6152B4CEED1379498ACA6F87089186C2301C97932151F631C2E01020FDA9C3F995C9A81DB118C9FA39 y_G = 79AFEE927FA08107917851B6F7DD952DC852C052D6A2DDE0E0E1AF9EA781CDC8AF1ED4E9B024BFA5FAB992443F80C745

p = $2^{383} + 2^{233} + 1$ p= 80000000000000000000000000000000000002000000000000000000000000000000 0000000000000000000000000001 n = 200000000000000000000000000000000000007FFFFFFFFFCEA783C2F07E63118A 47FDB07FBAD3E742437E35EA11EA81 a = 3 d = 15A x_G = 45FB6177745030C78D712C115E5354E6F6AB496E668C3DB6A55EB1EF2F941805 2C6FF89BF253B3813F48B5134EFD5220 y_G = 35525C097B1F9D1DB5ABB1695B62EF88E79DDCE007ACD8F167E7022A2EC379 AF0A7C184F8C09934F230A30522FA27EE3
n = 200000000000000000000000000000000000007FFFFFFFFFC4E3F96E2B485B72D B4DC623D7ECECFE59A562991D60A279 a = 3 d = 7B0 x_G = 5EE2C5033D54291ACD632030F57D0AD23E537766CAEF2EC8E7117FA4D247D71 E488F0FAA45B5A1E8FEA7E6617BFA140E y_G = 39949C9783884C7491E5D3F0D33704A278A6D4DF52C56C9B83FDC59BCF9282A 42BFB06B377CF4E00933842AD9B65511
n = 200000000000000000000000000000000000007FFFFFFFFFFB43327D7C94B4BEA 93CCCFCA103CF64D6A40B8F00E9016B a = 3 d = 33B x_G = 56499011F2275F6A3E9421923C69AA0F8C47A25BD16EF68635243D6B1867A732 DC11CA8C8CEFF30132EFDF33C9B2B969 y_G =46D521F9D3A98FF01C4D5676A6CBB88A78AB84F2502A0A30E9A51A64A26D60 FD47F8B02E0B5F1713783CFC878DAF6108

Таблица 4.7. Скрученные кривые Эдвардса почти простого порядка над полем с модулем p_{521}

p = $2^{520} + 2^{66} + 2^{65} + 1$ p = 10000000000000000000000000000000000000000000000000000000000000000000 00000000000000000000000000000000000000000000006000 0000000000001

n =
4000000000000000000000000000000000000000000000000000000001E5
EE0522024E5E359E578FDD4AAE5B3388CCE2795EC03189792778B80EDB1251
a = 3
d = 2A8
x_G =
BA5B99F7375F86F6D700E0DCD8E750E64FE9EA4F44B5C4CF022B45EDD2B819
62129B6AED970FCAF0754EA418054E5AAB55FD9911C
3BBAA99DB75F86F6D700E0DCD
y_G =
91E0646D69810E9FC606F87EA7BFA7140B1446EEA937536CDDC166C301BB0EC
E279DA15838913154E7BEDE69C7AC43D2756A279005
625F52F9D145A41E77E31D4B

$p = 2^{520} + 2^{80} + 2^{39} + 1$

p =

1000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000010000000008000000001

n =

3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
FFFFD4F3EBC074E5DCC909AAE45C02D90E60764C007

6526D0C9B9F66517F45515A05

a = 3

d = 87
x_G =
6287A808BC52FD146441B870582B459B154149EC3E444F88F3950A07DF4AC946
781D7517D86287A808BC52FD146441B870582B459B154
1AC73E64D0C0D347D72A7
y_G =
6A20EC64258785F16B2A400BF0C6D2E90C8080AE25A96F465029EE2A148A0EF2
BF0DF9E087F1E02487738C927BAF4B0C50A9349BA68
FAF1D3B24BB6E0D8E991B1

$p = 2^{520} + 2^{110} + 2^{4} + 1$

p =

1000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000040000000000000

0000000000011

n =

3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
FFFBE822ED9ECC69628BD3C3257DB51B5FC7477094B

E7A8266F720BFB22D9AA02123

a = 5

d = D2
x_G =
DE935EF49AF7A4D7BD26BDE935EF49AF7A4D7BD26BDE935EF49AF7A4D7BD2
6BDE935EF49AF7A4D7BD26BDE935EF49AF7A4D7B
D5E62C0F3160798B03CC581E62C1C
y_G =
7CD02BB98782E14860C91A53E907B7C4FDAA1C6898375790CB3FDED238FD9A
069B8C53B67AC5DC0C8B782D74624A69D91450D20C
C9DCA8B76DB97994008CB492F1

p = $2^{520} + 2^{348} + 2^{278} + 1$
p =
10000000000000000000000000000000000000000000100000000000000000040000000
00000000000000000000000000000000000000000000000000
0000000000001
n =
40000000000000000000000000000000000000000040000000000000000FFFFDC
B98F10F9CF70F01865640CBDD9D354DDFB78BBD8EE3F20DD416BE23E0E1D08
D
a = 5
d = 53
x_G =
D8F7F6D0EEC7BFB687763DFDB43BB1EFEDA1DD8F7F7A9E6BE90A54F35F4888
E598AE7E472CC573F239662B9F91CB315CFC8E598AE7E472CC573F239662B9F
91CB2
y_G =
FDB9481BBAFC22AD99D0C546F5B6C6809013308A794235AEE551A74142CBE46
95FD23708A6C27631A6AF3DF61BB4F83C481614132F713D0B860EAB37B7F7474
8D2

p = $2^{520} + 2^{73} + 1$

p =

10000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000200000

0000000000001

n =

```
3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
FFFD34B63B11FF960C064B6AEA3B6851C23CCDC50B6

E6921941026DF9E4883D1C157

a = 7

d =
10000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000001FFFF
FFFFFFFFFE786B
xG =
4CFD4C477A7FABB556313753CD284B6B2315E866A9C3A38F399E19FCEC6BA07
96367AFBD7B6AAE33967252EA110AF880CB1038A2E
699072C8FACACCF0F919B3C67
yG =
74E83FF02F90FB98F028FC1D0BD94A0093E94E6250BA9DDEE68F1DE65242E49
6D90B9CEAE817717DFB7EC4E9E2D7A0004FDBB99B2
79AA75BF2E785DB9A5D415EBA
```

В заключение отметим, что предлагаемые для стандартизации и имплементации скрученные кривые Эдвардса имеют рекордную скорость экспоненцирования точки. Большинство из рассчитанных кривых наряду с минимальным значением параметра $a = 3$ имеют всего 2-х или 3-х разрядное шестнадцатеричное значение второго параметра d, что практически позволяет пренебречь сложностью операций $1U$ и $2U$ для скрученных кривых в таблице 4.1. Оценки сложности сложения точек $V_E = 10M + 1S + 2U$ и удвоения точки $T_E = 3M + 4S + 1U$ достигают в нашем случае нижних границ $V_E = 10M + 1S = \frac{32}{3}M = 10.67M,$, $T_E = 3M + 4S = \frac{17}{3}M = 5.67M$ если принять $S = \frac{2}{3}M$ [2]. Кроме того, в отличие от предыдущей работы [37] с параметрами полной кривой Эдвардса, здесь найдены кривые для модулей поля p_{521} с наивысшим стандартным уровнем стойкости.

ГЛАВА 5

ПОЛНЫЕ КРИВЫЕ ЭДВАРДСА НАД РАСШИРЕНИЯМИ МАЛЫХ ПРОСТЫХ ПОЛЕЙ

В настоящей главе рассматриваются полные кривые Эдвардса над расширениями простого поля F_p, причем мы ограничиваемся малыми характеристиками p, равными 2, 5 и 7. В разделе 5.1 мы рассматриваем кривые над расширенными полями нечетных характеристик 5 и 7, а в разделе 5.2 – кривые над расширенными полями характеристики 2. Первые описываются модифицированными уравнениями (1.10) кривой Эдвардса, введенными Бернстейном и Ланге в работе [2]. Здесь нами была поставлена задача: для единственной кривой минимального порядка 4 над основным полем найти простое расширение степени m, при котором будет получен приемлемый для стандартизации псевдопростой порядок $4n$ кривой. После этого вычисляются координаты генератора криптосистемы и другие общесистемные параметры. Результаты решения этой задачи опубликованы в работах [32,33,36,60]. Кривые Эдвардса над полями характеристики 2 (раздел 5.2) имеют совершенно особые формы уравнений и законов сложения точек, впервые исследованные в работах [7,8]. Для этих кривых с двумя и одним параметром А.Дихтенко решена задача вычисления общесистемных параметров криптостойких кривых Эдвардса [40,43], изоморфных кривым стандартов FIPS-186-2-2000 (и последней версии этого стандарта FIPS-186-4-2013) и ДСТУ 4145-2002 [78,79,82].

5.1. Полные кривые Эдвардса над полями характеристик 5 и 7

Нами предлагается наиболее простой путь нахождения кривой Эдвардса почти простого порядка $4n$ [32]. Наподобие с кривыми Коблица над полями характеристики 2, мы предлагаем найти две полные кривые Эдвардса минимального порядка $N_{E1} = 4$ над малыми простыми полями F_5 и F_7, после чего найти порядки этих кривых над расширениями степени m полей F_5 и F_7 с последующим отбором при простых m подходящего почти простого порядка кривой $4n$. Под подходящим имеется в виду простое значение n длины $\log n$ в пределах стандартных величин (порядка 200 – 600 бит). В результате нами были определены несколько кривых в области известных криптографических стандартов.

5.1.1. Поиск кривых Эдвардса почти простого порядка и результаты

Форма полных кривых Эдвардса над конечными полями характеристики $p > 3$ имеет вид, определяеиый лишь одним параметром d в уравнении (3.1)

$$x^2 + y^2 = (1 + d\,x^2y^2), \qquad d(1-d) \neq 0, \left(\frac{d}{p}\right) = -1.$$

Пусть кривая определена над полем F_5, здесь допустимыми значениями параметра d являются квадратичные невычеты 2 и 3. Они являются мультипликативно обратными, поэтому образуют пару кривых кручения [2]. Границы Хассе $p + 1 \pm 2\sqrt{p}$ при $p = 5$ лежат в интервале [2,10], в пределах которого для кривых Эдвардса допустимы лишь 2 значения порядка N_E кривой, равные 4 и 8. Согласно теореме 1.4 точки 8-го порядка существуют тогда и только тогда, когда $(1-d)$ – квадратичный вычет, и не существуют в противном случае. При $d = 2$ значение $(1 - d) = 4 \bmod 5$ – квадратичный вычет, и соответствующая кривая имеет порядок 8. При $d = 3$ значение $(1 - d) = 3 \bmod 5$ – квадратичный невычет, при этом кривая Эдвардса (3.1) имеет минимальный порядок $N_{E1} = 4$. Она содержит лишь 4 точки любой полной кривой Эдвардса $(0, \pm 1)$, $(\pm 1, 0)$ на осях x и y.

Итак, мы принимаем $d = 3$, тогда из $N_{E1} = p + 1 - t_1 = 4$ след уравнения Фробениуса $t_1 = 2$. Рассчитаем порядки кривых над расширениями $F_p{}^m$ по известной формуле [29]

$$N_{Em} = p^m + 1 - t_m, \qquad (5.1)$$

где для определения параметра t_m воспользуемся рекуррентной зависимостью

$$t_m = t_1\, t_{m-1} - p t_{m-2}, \; m = 2, 3, \ldots, \; t_0 = 2. \qquad (5.2)$$

Результаты расчетов по формулам (5.1), (5.2) с отбором простых значений $n = N_{Em}/4$ приведены в таблице 5.1. Во второй колонке таблицы даны целые части значений для длины модуля поля $m_b = \lfloor m \log p / \log 2 \rfloor$ в битах.

Тестирование числа n на простоту с помощью теста Миллера-Рабина осуществлялось специальной прикладной программой.

В границах Хассе имеется еще одна кривая с минимальным порядком $N_{E1} = p + 1 - t_1 = 4$ при $p = 7$ и $t_1 = 4$. Она также имеет параметр $d = 3$, который является квадратичным невычетом в поле F_7, причем $1 - d = 5$ – также невычет. Простые сомножители n порядков этой кривой над расширениями F_7^m, рассчитанные с помощью (5.1), (5.2), даны в таблице 5.2.

Таблица 5.1 ($p = 5$)

m	m_b	$n = N_{Em}/4$
3	7	37
5	11	761
17	39	190734426721
47	109	177635683940025049111870902558317
53	123	27755575615628913519432138978855094 01
181	420	81566305849981556583878676365706844446264553225862081846982955622470058935583394181280598166864036391710622583401627348551324 1
227	527	11591269220898191830411672692336373479273639933618096882665747059117441687798840670250687806029382008026655960498496355087266800506918498606995903214468432291 7
353	819	13625471488026082303712171891991388314389109549794181122960160293908508251985766836112118027927542086233890704552817681219819158519647915156383473783742883700653042365583720331179910890621621002009304697009015594466023580409118149203179025776784 01

Таблица 5.2 ($p = 7$)

m	m_b	$n = N_{Em}/4$
5	14	4261
7	19	205759
17	47	58157621574673

43	120	5459535939979491492246532674488972 83
47	132	131083457918907590863454504379855878 2183
127	356	531362731142004177110825957764740560832984541840999627025991600224016573324879563993413337967881303987543591
223	626	715818522269416229332997374116591979329811044151737634516615270731955989279240178038393960757114885677425855487376571651860602081282044454562195975459126950384575131473354470967163835260 39

Приходится констатировать, что наши априорные ожидания достаточно большого числа приемлемых для криптографии кривых Эдвардса над расширениями малых простых полей характеристики p > 3 не подтвердились. Как следует из таблиц 5.1 и 5.2, в границах стандартных требований к порядку генератора криптосистемы и близким к нему расширением 2^{mb} ($m_b \cong 180 \ldots 600$) мы нашли всего три кривые Эдвардса: две кривые над полем $F_5{}^m$ со степенями m = 181 и m = 227, и одну кривую над полем $F_7{}^m$ со степенью m = 127. К ним, правда, мы добавляем еще 2 кривые с m = 353 (при p = 5) и с m = 223 (при p = 7), т.е. с завышенным в сравнении со стандартным уровнем стойкости и значением m_b > 600 (очень скоро он перестанет быть завышенным). По сравнению с кривыми Коблица над полем $F_2{}^m$, число пригодных для стандартизации кривых снижается в 2 – 3 раза. Известные кривые Коблица американского стандарта FIPS-186-2-2000 [78] имеют степени расширения m = 163, 233, 283, 409, 571, т.е. практически в каждой сотне значений m.

5.1.2 Вычисление параметров генератора криптосистемы на кривой Эдвардса над расширениями простых полей характеристик 5 и 7

В этом разделе рассмотрим методику и результаты вычислений координат точек G как генераторов простого порядка n. Эта работа была выполнена А.Дихтенко и А.Яценко и опубликована в работах [36,60].

В качестве примера рассмотрим определение координат точек кривой Эдвардса простого порядка n (таблица 5.1) над полем F_5^{181}. Для данного расширения был найден примитивный полином минимального веса 5 $P(z) = z^{181} + z^3 + z^2 + 3z + 3$, который используется в арифметике поля при сложении точек кривой. Выбираем случайную координату

$x = x(z) =$ [3 0 2 4 0 2 3 1 1 4 0 0 1 0 3 1 4 3 2 3 0 4 3 1 4 4 3 0 1 4 1 4 2 3 4 4 2 1 4 4 1 3 1 0 1 1 0 0 3 4 4 1 3 3 0 1 2 3 0 0 3 3 2 1 0 3 4 2 3 1 4 4 3 1 1 2 2 1 3 1 1 2 2 3 2 0 4 3 0 2 2 0 4 0 2 0 3 3 1 0 2 2 4 0 1 3 3 2 4 2 2 1 1 0 1 2 2 2 3 1 4 3 4 4 3 3 0 3 1 1 3 2 4 0 2 2 3 1 1 1 1 2 1 3 2 3 3 3 4 0 4 2 4 2 3 2 0 2 1 1 4 3 0 4 2 4 4 1 2 0 3 0 1 3 1 2 2 4 2 1],

вычисляем согласно (3.1) значение

$y^2 = (1 - x^2)\cdot(1 - 3x^2)^{-1} =$ [0 3 2 2 4 1 0 0 4 2 2 4 3 0 0 4 1 1 1 0 2 4 4 4 2 1 3 1 3 4 0 4 2 3 1 4 0 0 0 4 1 4 2 1 3 4 0 3 3 2 1 4 3 4 4 0 3 0 0 0 1 3 0 1 0 1 3 1 3 4 4 4 0 1 2 3 3 4 3 2 1 4 3 2 4 2 1 0 3 1 0 4 3 4 1 3 1 1 0 1 1 2 3 2 3 3 2 4 3 0 1 4 3 1 0 0 4 3 1 0 4 2 3 1 4 4 4 1 0 4 0 0 2 3 1 4 2 1 0 4 4 2 4 0 3 1 3 1 2 3 2 3 2 0 1 3 1 1 4 1 1 4 3 2 0 1 3 0 4 3 2 2 0 2 0 1 0 4 0 2 4]

(младшая степень – слева). Определение квадратного корня из элемента y^2 выполняем с помощью экспоненцирования [29]. В нашем случае $p = 5 \equiv 1 \bmod 4$, $q = 5^{181} \equiv 5 \bmod 8$. В мультипликативной группе поля F_q , если $a = y^2$ – квадратичний вычет, имеем элементи подгруппы F_5^*

$$a^{\frac{q-1}{2}} = 1, \quad a^{\frac{q-1}{4}} = \pm 1 = \delta, \quad \delta^{\frac{1}{2}} = \pm 2.$$

Тогда

$$a = \delta a \cdot a^{\frac{q-1}{4}} = \delta \cdot a^{\frac{q+3}{4}} \qquad \Rightarrow \; y = \delta^{\frac{1}{2}} \cdot a^{\frac{q+3}{8}}.$$

С помощью этой формулы получаем

$y =$ [2 0 1 1 3 0 3 2 3 2 3 0 4 3 4 1 2 0 2 1 4 3 2 2 0 1 2 3 2 1 0 0 2 2 3 4 1 4 1 4 2 3 3 2 1 0 0 2 3 0 4 2 2 0 0 1 2 2 3 1 2 1 2 4 3 4 2 3 2 4 0 2 1 2 1 3 4 4 2 2 4 2 1 1 3 0 4 4 3 3 3 0 4 4 0 0 0 0 1 4 4 1 0 0 0 4 1 4 1 2 4 2 2 2 2 3 2 1 1 1 3 2 2 2 1 2 4 0 0 4 0 1 4 1 0 4 0 4 3 1 1 0 1 3 1 2 0 3 2 2 1 4 4 3 0 1 2 4 3 0 0 4 4 3 0 1 0 2 3 2 2 0 1 1 0 2 2 3 1 3 3].

Подстановка найденных координат $(x, y) = P$ в уравнение (3.1) дает тождество, поэтому эта точка существует. Далее, скалярное умножение точки P на порядок n из таблицы 5.1 дает точку $F = (0, 1)$ 4-го порядка, поэтому генератором криптосистемы порядка n является точка $G = 4P$. Ее координаты:

X = [0 3 2 0 1 1 2 2 1 4 2 3 0 4 2 1 1 3 0 3 1 3 2 4 2 2 3 1 0 3 4 0 1 4 4 0 4 0 3 3 4 0 1 4 1 0 1 2 4 0 2 3 3 2
4 3 1 4 0 2 3 3 3 4 3 4 3 1 2 3 0 1 2 4 3 2 0 3 3 0 0 2 0 4 2 3 1 1 1 1 2 3 0 4 3 4 0 4 4 3 0 1 2 1 1 1 4 4 1 0 2
3 1 1 0 0 4 1 4 4 2 4 1 0 0 4 3 0 1 3 3 1 4 1 4 4 1 1 1 0 3 3 0 4 2 3 1 0 1 1 3 1 3 0 4 1 2 3 1 2 1 3 4 1 0 3 1 4
0 4 1 1 2 4 4 4 2 0 0 2] ,

Y = [3 4 2 2 4 2 4 2 3 3 2 4 1 0 2 0 3 3 0 3 4 3 3 1 1 0 2 2 4 2 4 2 1 2 1 2 2 0 1 0 2 4 3 4 4 3 4 2 1 1 4 1 3 0
0 4 1 3 1 3 2 1 1 3 3 1 1 4 0 2 2 3 2 0 1 0 4 0 2 4 4 1 4 0 3 0 3 4 2 3 3 3 0 1 3 3 0 3 3 4 2 3 3 0 1 4 4 3 2 3 3
0 1 4 0 3 4 2 3 3 1 1 3 1 1 0 3 1 1 4 1 4 2 1 0 3 3 3 1 1 1 1 4 0 2 0 0 3 3 4 2 0 4 1 4 0 2 3 1 1 4 0 2 3 0 0 3 4
2 0 0 3 0 2 0 4 1 1 4 0 4] .

В случае, если при тестировании умножением на n получим точку 2-го порядка, генератор определяется как $G = 2P$.

В работе [36] детально рассмотрен один из возможных способов нахождения кривых Эдвардса вида (3.1), в границах приемлемых криптографичных значений параметров (таблицы 5.1 и 5.2). Кривая $x^2 + y^2 = 1 + 3x^2y^2$ имеет почти простое значение порядка $4n$, которое удовлетворяет стандартным требованиям к порядку генератора криптосистемы при расширениях m полей F_5 и F_7, приведенных в таблице 5.3. Здесь же даны значения битовой длины m_b этих полей и значения простых чисел $n = N_{Em}/4$.

Таблица 5.3. Расширения полей F_5 и F_7 и простые порядки n генераторов G криптосистемы

F_p^m	m_b	$n = N_{Em}/4$
F_5^{181}	420	4D1E1043D31FB1CC9B562A717B3C43259476330974981C14F25E03 EACA14C7378C72BEB6F54DB72B8180B352DF12BA34CC023C219
F_5^{227}	527	21C529DD78FA571E196B3EBB0D20429C476A1848CAB5E0E8A121 378DE187888F99D299F404EE4F9BC974D5035A62AC9F5E1E0DA2 9A510B4012E23ECD15909A4B1065
F_5^{353}	819	13625471488026082303712171891991388314389109547941811229601602939085082519857668361121180279275420862338907045528 17

		6812198191585196479151563834737837428837006530423655837203 3117991089062162100200930469700901559446602358040911814920 317902577678401
F_7^{127}	356	5CAC4104D859A6DF582D5731211D9947A4AE9CFD1F4E3648997D 050DCE03624B891381F19AA1824CF98DE5637
F_7^{223}	626	7158185222694162293329973741165919793298110441517376345166 1527073195598927924017803839396075711488567742585548737657 1651860602081282044454562195975459126950384575131473354470 96716383526039

Дальнейшая реализация рекомендованных в [32] кривых Эдвардса над расширениями полей F_5 и F_7 представляет собой два последовательных этапа:

- поиск примитивных полиномов P(z) для полей F_5^{181}, F_5^{277},F_7^{127} и построение соответствующей арифметики этих полей;

- вычисление координат генератора абелевой группы точек кривой в соответствии с арифметикой полей F_5^{181}, F_5^{277} и F_7^{127}.

С помощью прикладной программы был получен ряд примитивних полиномов указанных полей, среди которых ми выбрали полиномы минимального веса (отметим, что для случая поля F_7^{127} существуют примитивные полиномы наименьшего веса – триномы). В общем случае точками кривой будут пары (*x, y*) элементов поля F_p^m, для которых виполняется равенство (3.1). Для поиска генератора подгруппы точек исследуемой кривой Едвардса $x^2 + y^2 = 1 + 3x^2y^2$, выбираєм случайную координату x из элементов соответствующего поля и вычисляем значение $a = \frac{1-x^2}{1-3x^2}$. Если $\left(\frac{a}{p}\right) = 1$, то определение квадратного корня из элемента *a* в расширенном поле осуществляется с помощью экспоненцирования [28].

В случае поля характеристики 5: $q = 5^{181} \equiv 5 \bmod 8$ или $q = 5^{227} \equiv 5 \bmod 8$. В мультипликативной группе поля F_q, если $a = y^2$ – квадратичний вычет, имеем элементи подгруппы F_5^*:

$$a^{\frac{q-1}{2}} = 1, \qquad a^{\frac{q-1}{4}} = \pm 1 = \delta, \qquad \delta^{\frac{1}{2}} = \pm 2.$$

Тогда $a = \delta a \cdot a^{\frac{q-1}{4}} = \delta \cdot a^{\frac{q+3}{4}} \Rightarrow \quad y = \delta^{\frac{1}{2}} \cdot a^{\frac{q+3}{8}}$.

Для поля характеристики 7: $q = 7^{127} \equiv 3 \bmod 4$.

Аналогично, так как $a = y^2$ – квадратичний вычет, имеем:

$$a^{\frac{q-1}{2}} = 1, \qquad a^{\frac{q-1}{2}} a = a^{\frac{q+1}{2}} = a, \qquad \Rightarrow \qquad y = a^{\frac{q+1}{4}}$$

Таким образом, получаем пару координат (*x, y*), удовлетворяющую равенству $x^2 + y^2 = 1 + 3x^2y^2$, и значит, точка $Q = (x, y)$ лежит на кривой Эдвардса. Умножив Q на величину n из таблицы 5.3, можем получить точку $O = (1, 0)$, точку $D = (-1, 0)$ 2-го порядка или точки $\pm F = (0, \pm 1)$ 4-го порядка. В первом случае генератором G подгруппы точек кривой Эдвардса простого порядка будет точка $G = Q = (x, y)$, в других – генератор G определяеся как $G = 2Q$ или $G = 4Q$ соответственно. Результаты вычислений, а именно, примитивные полиномы и генераторы подгруппы точек кривой $x^2 + y^2 = 1 + 3x^2y^2$ простого порядка для соответствующих полей, приведены в таблицях 5.4 – 5.7 (младшие степени векторов – слева).

Таблица 5.4. Параметры кривой Эдвардса $x^2 + y^2 = 1 + 3x^2y^2$ над полем F_5^{181}.

$P(z) = z^{181} + z^3 + z^2 + 3z + 3$

x=[020142034310022224101012221304140301432203243032241124434204012141101011240333421403412430442412314131110013401212233320143114004323232130032424012224444043224043044333224012421344 4]

y=[332430012113102123101022332221434201344433301244110443241322234411431010032114420334344112412432431021014432304241344110320103214110041311411143304243313330304410134112442244300230 4]

$P(z) = z^{181} + z^3 + 4z^2 + 3z + 2$

x=[00322134112010012143103323241140323012224131132321024424240130
32143111433023110100343142203133440433332220322322444424114233140
30420411224034442134440131343004334020340401303330243]

y=[03043020420114103344012120100442012323311044140132023413041313
24303323141123240023041121001203142020432030102410043113312224344
23031243233323200023131134221110113111233041334110303]

$P(z) = z^{181} + 2z^4 + z^3 + 2z^2 + 2$

x=[22004023004104340442013341242213300221442130010210021300001434
24042034331303014110433014433341334034302432140034332144023404131
03210401232142442030124341024334330424232440120113002]

y=[32430013442442200430444314301132321021422011020423000330330433
22004242123134203212442331122041111003213041131012213042202403102
04210400144112114132110343442043222324113020213312223 2]

$P(z) = z^{181} + 3z^4 + z^3 + 3z^2 + 3$

x=[01143320410224102443132333314340403023030212110414003223413021
22032133124143101103200223340012212404414411342003224204233430203
43343324131140104122114122431314220110124242443242042]

y=[31141331400440010243444221440141143122241040233232242110412014
43032413340203330424022312133120114324233002130000332113130200022
31440302241203004141444211110400022431042343111443331]

$P(z) = z^{181} + z^5 + z^3 + 2z + 2$

x=[23013124041144401400433102343011223012242120003242244333124304
32103412342314223402334440402311200211443033043003241213132010242
43440011311440201424314041042203010202041413441220344]

y=[03343412304143114142243340110340412334204423313344234344244322
33244031334231414340110030124414333340232211410342123441444003341
21032240203210003304242103013330244110134334240110411 2]

$P(z) = z^{181} + z^5 + z^3 + 2z + 3$ x=[033331143444211132440330113010331322120203042204021130003110224122324111232304143213011432430321040003402334313404120314141110203301342312133204014433102201121414213103210020233233] y=[224433021012423102133424420211322011420114010032223220143234001020013040323402413104011423002000431000143241344412311141324132431001304331413224301101321240311333112331101113212222]
$P(z) = z^{181} + z^5 + 2z^4 + 3z^3 + 2$ x=[401313203242141100414140341402132031340204211010304301304433000241303202204322333210310143001114430042433302110323403441142111010403133420102301120222303041212423022004240014014144 2] y=[233004302440242400114114043112223221431440010340210110330414113414342233432443313200244201210031432134431420414013303432040211000102400144423003012304204124411022242214442113402104 2]
$P(z) = z^{181} + z^5 + 3z^4 + 3z^3 + 3$ x=[430231143030042104122442004211031442233231212331100100323003341201344433331120310002101312011223223204122134310412131024210020010344000212342212231041402210200331004344113222024213] y=[220214013224431013231414442230343012344044044401130023141411412230431411443324442143420222112422320020233343333111224420014110214141313101211024203211233332014432024110101244422032]
$P(z) = z^{181} + 2z^5 + 2z^3 + 4z + 2$ x=[213303120043101410430214113000223042401423430440104441233234204013241021021410012212131113130013010300034431014044244234032001424420410222124313100214302032044141310412102201001122 4] y=[333342223403012114404422142010023210021141130430442302023204431414313444410321233000203120022141222334233342304003212003112104210341331403421343332031320420414242343223111134313142 4]

P(z) = $z^{181} + 2z^5 + 2z^3 + 4z + 3$ *x*=[041004310220124400010323000323444403432134210112041043030341412422420242102402133113211112200032214301020314430023124230400401241320121003143044221313232340414001111113200242424002 4] *y*=[103300132311434420343104341031424340043124401420403234203202244214313030440040144030112014201104040324134020031112010040201101100101322211004030140442442444443041223413101210230300 2]

Таблица 5.5. Параметры кривой Эдвардса $x^2 + y^2 = 1 + 3x^2y^2$ над полем F_5^{277}.

P(z) = $z^{227} + z^3 + 2z^2 + z + 2$ *x*=[200434204200342242421132330113311033122431404333242244010231330340440143341310142343240422122012144022221230022024030021341441000322330142321441233132042423021432414321221131340244221020101130123333102342434124103011001024411 2] *y*=[143314021300210304111310442344114031131013143133441204112044300113411410202331023320443333224142334442403242311333313133304320431140021013242301132431231300041404210042241140132201331324024442100212031201421440342020413001121 24]
P(z) = $z^{227} + z^3 + 3z^2 + z + 3$ *x*=[320221201114411222324310314021323031404042223141002332301313302110244332031130143141303444041130332024410412211212401323224002210342140411344214404423041332423202410344324144233101420230340104323034103143333441231103241330144 44] *y*=[001120210000330303040420144233441044014422043244201001232342113432204102340213022403420404340301130140402401133341212021323300414322143310300223222304313024443401212311042303131101220244224230313201403111311123434242202004444 1]

P(z) = $z^{227} + z^4 + 2z^2 + z + 2$ *x*=[00130341131312012421212040140240313014202340431420010044200313404102301334141242324123210333142033414113314031314413000113222221440212004022243013143332123403114333141444340241300030204403341112121434014142324330221001112410233] *y*=[00314323301034220420433200330304332423230043240411103020101034122314222141341104433434122243404422030404334223334303034040424401041124242344400104324232112301300142220320443040141332230222013431233200332300420214144122220040011]
P(z) = $z^{227} + 2z^4 + 4z^3 + 2z + 3$ *x*=[04143040131323211331302033023204443224124423021343210341003414444033213324434332221104012340233034200120441212232122404211210033411123013102023002023120343430143204313011133221134001430444231340213300421434310222012014001304312] *y*=[42230343300243410212030123020242100011302124303210440141002443241020441144112034004242310442233112340411042343130421314112244031221111113214212420111422144013404042341302424143014134400303140332312122304404124120142141042421 12]
P(z) = $z^{227} + 3z^4 + 4z^3 + 2z + 2$ *x*=[42431120100132410101402123140402300013433034141124410421142232234430434120401121023214001134341202002213410434423032400123012013430200241103103223122442420240324142044104100240004233131420424303114010414324324130040420431221312] *y*=[23232102342014112231203433413112230330443100403233204023040124003030212110124431223141434223110141130003014112014223414410344240443220114242140311242030003020022043303103223344013230202242002334432013101442113041341131031430241]

$P(z) = z^{227} + 4z^4 + 3z^2 + z + 3$

x=[243033421310442312042403420110322200130221242234112430021030413342324212123023424304443244001144241113413322320440241231131304432000202323210224144301214224122412033423144302111444400110310141404400033033330022313312132242401]

y=[401022101330113440340210024210200341414203031014213313241442214103200103211430141340300330320234102343303040432013201203312232103030322434122132300314311023431121244414324304420332312342311423001432412431304222440220303131 01]

$P(z) = z^{227} + z^6 + 2z^3 + 3z + 2$

x=[413434123431201213333420003131303322304231314021411432000102431113042211332410113244243034333023221221220221333300134243440100214244422311231332420421242412424432200444101312142321302243332041010141441413231121131433404403203 04]

y=[423413402412444223004404402422213302120013124120411142213312041344430004401110020001413123442441040342243112224040204311103334410120140334010311000202121302010211221240124412101431243430143004014113103204341100044424 1002004334]

$P(z) = z^{227} + z^6 + 3z^3 + z + 3$

x=[134003443320223441424103313002311340123430324400402142103143341443424123022430331020320004130121120330004413112413210034323320231024041434233314012240000000403410341301212112102204101143024103140214210312044343242341234304144 22]

y=[021033404141101144123033113412100310032433031043221113032433240302440031030231042324202212343322231110203431341212204320221304330001212034320103201103303013312331320101203244031423413102202003001413141442124221212340420412320 21]

P(z) = $z^{227} + 2z^6 + z^4 + 2z^2 + 2$ *x*=[32430432204441144130401211141034021102003130444001114011304301313433441123000142403411000400241433323230400134043303013343004030211413131231430001030022322410141340242404322113122423244222040423221042213312140403221123420220143] *y*=[10420433333004020211302010113111131024121232430033304242243030003102014401040010430212034430124413432331321310232423322241243123140220131422102112422100014203040212131300001014434224221141411234112422113200341203313003133044 24]
P(z) = $z^{227} + 2z^6 + 3z^4 + 4z^2 + 3$ *x*=[02124203121121324411200224140411102320211132444243110341144403332421234003140123120411042423013102433344041411420133023011103223400200431141020403033114412130332424342132443301332100234223114013114320024111430010043412323030441] *y*=[43424210212234423234101402131204120131221240003311022232424130102024343134034323320333441003124123322112414420123021202100333444011423323000043442110340122330042040310412322120443044300230112104222243003134330440302014342021402]

Таблица 5.6. Параметры кривой Эдвардса $x^2 + y^2 = 1 + 3x^2y^2$ над полем F_7^{127}.

P(z) = $z^{127} + 3z^2 + 2$ *x*=[46044006603145305201405123204222530031016222516201005661204244422522523363324113266525222005454623240563146654416124642102126225] *y*=[13151463044055366051631435240114440055064050055036454015644623565455240161623205625064212024436215201514620643652053153153046045]

$P(z) = z^{127} + 5z^2 + 4$ x=[036554333652105643411546203513264545351511610136523305665530511625524564144054211114445363065516550405625360361336651022653213 6] y=[132053546202660426010154442621356120411534111560222032320333343324562615312121326606566114006006315112441333550621423132645246 5]
$P(z) = z^{127} + 2z^2 + 2z + 4$ x=[646631565224643604246133226210612643251551466125466503401655264606562606053306010012415243655321122425463616534232436635331053 3] y=[054230045346246346451443465641064641421513031154655553405522145145425424155136320501546005445220511541534315522544013432301661 4]
$P(z) = z^{127} + 4z^2 + 2z + 2$ x=[156036134334532014564403432461341616623261162625645626110654136116401500643545203224414354302031524052223361061603162304503464 6] y=[225651365154041032143546141001026204244326102044342215055265206316101201033041310341355324424250211600206356043453305605321324 5]
$P(z) = z^{127} + 2z^3 + 6z^2 + 2$ x=[510232400251512003015131465556251123335143134253126154046246333635444641124021432311045466445624116331516646433452526221032242 2] y=[314546003640022304353165222000356536306533263361055343302601643622336405351324050122611535522560114361501505141233060455355530 5]

$P(z) = z^{127} + 4z^3 + 3z^2 + 4$ x=[032104460221551530616360444453465435545665322340211542662646430065534316213363304523543423304163211302330065104135163465126042 1] y=[643004566234311166623635400353406535235400160150044555651513462330355264665540264160105564053203333363313136064113103645665611 3]
$P(z) = z^{127} + 3z^5 + 2z^2 + 4$ x=[360554063253063002002062140020353366635121004451151221431264632043325236211316313312131064046010503011264510662402135401436311 6] y=[502334232404301404025500036021355020350553350330566212632040005330062002331604551532526031661063104051361050150236622362112661 6]
$P(z) = z^{127} + 6z^5 + 4z^2 + 2$ x=[263536326365544262432531352065003352434164663011216645330110565464143210235525622154142303024501161450602100416105443561563056 2] y=[406432032231634653246024514250335146034624533454121205465461245335003004365653411544013653656556131646645012205146262300316223 3]

С практической точки зрения вопрос о реальных оценках производительности криптосистем на кривой Эдвардса над расширениями малых полей по сравнению с кривыми над большими простыми полями требует экспериментальных исследований. Однако теоретический анализ позволяют утверждать, что параметры, приводимые в таблицах 5.4–5.6, обеспечат максимальную производительность криптосистем при заданном уровне стойкости, определяемым порядком n генератора G криптосистемы.

Полученные параметры можно рассматривать как эквивалентные относительно производительности при фиксированной стойкости системы.

Исключением является случай поля F_7^{127}: для этого поля найдены два примитивных полинома минимального веса 3, поэтому соответствующая арифметика поля является более эффективной по сравнению с решениями, построенными с полиномами большего веса. Незначительные потери стойкости криптосистем [33] на кривой Эдвардса $x^2 + y^2 = 1 + 3x^2y^2$ над полями F_5^{181}, F_5^{227} и F_7^{127} составляют 2 бита в каждом из этих трех случаев, что является несущественным по сравнению с значениями соответствующих полей в битах.

Следует заключить, что найденные кривые с минимальным и простым значением параметра $d = 3$ обеспечивают при заданной стойкости наивысшую производительность вычислений групповых операций. В операции сложения разных точек мы экономим на одной полевой операции умножения $1U$ на параметр кривой (раздел 3.8), так как умножение на $d = 3$ заменяется двукратным сложением в поле, т.е. практически бесплатной операцией. В разделе 3.9 мы привели оценки сложности сложения и удвоения на полной кривой Эдвардса $V_E = 10M + 1S + 1U = 11{,}17M$, $T_E = 3M + 4S = 5.67M$. С учетом экономии в операции $1U = 0.5M$ получим $V_E = 10{,}67M$, $T_E = 5.67M$, тогда в соответствии с формулой (3.65) максимальный выигрыш производительности экспоненцирования точки в сравнении с кривой в форме Вейерштрасса над тем же полем достигает значения $\gamma(0.33) = 1.6$. Это второй результат для значения выигрыша после рекордного для инвертированных проективных координат ($\gamma(0.33) = 1.66$, см. раздел 3.8). Арифметика вычислений в расширениях малых полей часто эффективней арифметики в простых полях большой характеристики [15,71,72]. Полагаем, что их можно рекомендовать как для проектов будущих стандартов, так, возможно, и для использования в высокоскоростных криптопротоколах уже сегодня.

5.2. Кривые Эдвардса над расширенными полями характеристики 2

При построении криптосистем на эллиптических кривых наиболее широко используются кривые над большими простыми полями F_p и кривые над расширенными полями F_2^m характеристики 2. Сравнивая их

производительность, нельзя гарантировать преимущества какой-либо из них, так как быстродействие криптосистемы зависит от множества факторов, и выбор между кривыми исходит из требований к параметрам и возможностей реализации. Кривые над полями F_{2^m} в форме Вейерштрасса рекомендуются действующими стандартами и успешно применяются в протоколах шифрования [29]. Известно большое число усовершенствованных алгоритмов вычислений в этих полях.

Перспективным классом эллиптических кривых сегодня является форма Эдвардса [2,3]. Для полей F_{2^m} соответствующие формы уравнений и законов сложения точек получены Бернстейном, Ланге и Фарашахи в работе [7]. В продолжение ряда работ [7,8] здесь мы кратко описываем циклические кривые Эдвардса, заданные над расширенными полями характеристики 2 . По форме уравнения кривой Эдвардса над полями четных и нечетных характеристик существенно отличаются, но есть ряд свойств, сходных для кривых этих двух типов. Главные преимущества кривых Эдвардса – высокая производительность и удобство реализации благодаря полноте закона сложения и наличию аффинных координат нуля группы точек кривой [7]. В данном разделе мы приводим формулы изоморфного преобразования канонической кривой и кривой Эдвардса для случая полей характеристики 2. Взяв за основу действующие стандарты ДСТУ 4145 – 2002 [82] и FIPS 186-2 – 2000 [78], мы вычислили параметры изоморфных кривых Эдвардса и координаты генераторов криптосистем над различными расширениями полей характеристики 2. Их перечень дан в разделе 5.2.2. Простой порядок генераторов криптосистем в предложенных кривых сравним с величиной соответствующего поля и удовлетворяет стандартным требованиям.

Рассмотрим F_{2^m} – конечное поле характеристики 2 и d_1, d_2 - пару элементов этого поля, для которых справедливо $d_1 \neq 0$ и $d_2 \neq d_1^{\,2} + d_1$. Тогда кривая Эдвардса над полем F_{2^m} в аффинных координатах будет задаваться уравнением [7, 8]

$$E_{d_1,d_2} : d_1(x+y) + d_2(x^2+y^2) = xy + xy(x+y) + x^2y^2 \qquad (5.3)$$

Подобно кривым Эдвардса над полями нечетных характеристик [2,3,4], у каждой кривой Эдвардса вида (5.3) существуют обязательные точки. В данном случае таких точек две: $O = (0,0)$ – нейтральный элемент относительно сложения и $D = (1,1)$ – точка второго порядка кривой (5.3). Отсюда следует, что при отсутствии точек 4-го порядка минимальный кофактор порядка циклической кривой Эдвардса над полем F_{2^m} равен 2. Кроме того, для таких кривых справедливо свойство покоординатной симметрии и, при заданных ограничениях на параметры d_1, d_2 , кривая вида (5.3) не имеет точек сингулярности [7,8]. Авторы [7] также рассматривают альтернативную форму записи кривой Эдвардса над расширенным полем характеристики 2. Переход к этой форме может быть осуществлен посредством изоморфного преобразования кривой (5.3) по правилу: $(x, y) \to (x, y+1)$.

Правила сложения точек (x_1, y_1) и (x_2, y_2) кривой вида (5.3) задается формулами [7]:

$$(x_1, y_1) + (x_2, y_2) = (x_3, y_3),$$

$$x_3 = \frac{d_1(x_1 + x_2) + d_2(x_1 + y_1)(x_2 + y_2) + (x_1 + x_1^2)(x_2(y_1 + y_2 + 1) + y_1 y_2)}{d_1 + (x_1 + x_1^2)(x_2 + y_2)},$$

$$y_3 = \frac{d_1(y_1 + y_2) + d_2(x_1 + y_1)(x_2 + y_2) + (y_1 + y_1^2)(y_2(x_1 + x_2 + 1) + x_1 x_2)}{d_1 + (y_1 + y_1^2)(x_2 + y_2)}. \qquad (5.4)$$

При удвоении точки можно получить формулы, требующие лишь одной инверсии элемента поля:

$$2(x_1, y_1) = (x_3, y_3),$$

$$x_3 = 1 + \frac{d_1 + d_2(x_1^2 + y_1^2) + y_1^2 + y_1^4}{d_1 + x_1^2 + y_1^2 + (d_2/d_1)(x_1^4 + y_1^4)},$$

$$y_3 = 1 + \frac{d_1 + d_2(x_1^2 + y_1^2) + x_1^2 + x_1^4}{d_1 + x_1^2 + y_1^2 + (d_2/d_1)(x_1^4 + y_1^4)}. \qquad (5.5)$$

Из (5.3) и (5.4) следует, что точка (y_1, x_1) будет обратной к точке (x_1, y_1) [7]. С помощью (5.4) можно также убедиться в том, что $(x_1, y_1) + (0,0) = (x_1, y_1)$ для произвольной точки (x_1, y_1). Более сильное ограничение на параметр $d_2 \neq t^2 + t$, при $t \in F_{2^m}$ обеспечивает полноту закона сложения (5.4). Это означает, что приведенные формулы справедливы для всех пар точек (x_1, y_1), (x_2, y_2), включая равные, обратные точки и нуль группы *O*. Полнота закона

сложения – главное преимущество кривых вида (5.3) перед другими формами представления эллиптических кривых с точки зрения практических приложений. Далее будем рассматривать кривые, для которых выполняется условие $d_2 \neq t^2 + t$.

5.2.1. Изоморфизм между несуперсингулярной канонической кривой и кривой Эдвардса над полем F_2^m

Несуперсингулярная эллиптическая кривая в канонической форме Вейерштрасса над расширенным полем характеристики 2 имеет вид [6,29]

$$v^2 + uv = u^3 + a_2 u^2 + a_6, \quad a_6 \neq 0. \qquad (5.6)$$

Нулем группы точек эллиптической кривой (5.6) относительно сложения является точка O (точка на бесконечности).

В отличие от случая кривых над полями нечетной характеристики, где класс канонических эллиптических кривых шире соответствующего класса кривых Эдвардса, каждая каноническая кривая вида (5.6) изоморфна некоторой кривой Эдвардса вида (5.3). Изоморфное преобразование от формы кривой (5.3) к форме (5.6) осуществляется посредством следующих формул [7]:

$$u = d_1(d_1^2 + d_1 + d_2)\frac{(x+y)}{(xy + d_1(x+y))},$$

$$v = d_1(d_1^2 + d_1 + d_2)\left(\frac{x}{(xy + d_1(x+y))} + d_1 + 1\right),$$

$$a_2 = (d_1^2 + d_2), \qquad a_6 = d_1{}^4(d_1{}^4 + d_1{}^2 + d_2{}^2). \qquad (5.7)$$

Обратное преобразоване задается формулами

$$x = \frac{d_1(u + d_1^2 + d_1 + d_2)}{(u + v + (d_1{}^2 + d_1)(d_1^2 + d_1 + d_2))}, \; y = \frac{d_1(u + d_1^2 + d_1 + d_2)}{(v + (d_1{}^2 + d_1)(d_1^2 + d_1 + d_2))}. \qquad (5.8)$$

Используя приведенные выше формулы, можно осуществить изоморфный переход от известных канонических эллиптических кривых

(которые применяются на практике и рекомендованы действующими стандартами [78,82]) к кривым в форме Эдвардса. В следующем пункте мы приводим перечень таких кривых и кратко описываем механизм соответствующего перехода.

5.2.2 Вычисление параметров криптостойких кривых Эдвардса над расширенными полями характеристики 2

Эта задача была выполнена А.Дихтенко, ее результаты опубликованы в работе [40]. Для нахождения параметров кривых Эдвардса над расширенными полями характеристики 2 мы взяли за основу канонические эллиптические кривые действующего украинского стандарта ДСТУ 4145 – 2002 и кривые стандарта FIPS 186-2 – 2000. Зная значения коэффициентов a_2, a_6 канонической эллиптической кривой над заданным полем, вычисляем параметры d_1, d_2 следующим образом [7]:

1. Выбираем некоторый случайный параметр d_1 так, чтобы выполнялись условия $Tr(d_1) = Tr(a_2) + 1$ и $Tr\left(\frac{\sqrt{a_6}}{d_1^{\,2}}\right) = 1$.

2. Рассчитываем значение второго параметра по формуле $d_2 = d_1^{\,2} + d_1 + \frac{\sqrt{a_6}}{d_1^{\,2}}$.

В итоге получим параметры кривой в форме Эдвардса, изоморфной данной над исходным полем. Результаты расчетов приведены в таблицах 5.7 и 5.8 в шестнадцатиричной системе. Здесь мы также приводим координаты генератора (x_G, y_G) для каждой из полученных кривых Эдвардса. Поиск генератора осуществлялся аналогично работе [7]. В таблице 5.7 содержатся кривые для случая американского национального стандарта FIPS 186-2 – 2000 [78], в таблице 5.8 – кривые для случая украинского стандарта ДСТУ 4145 – 2002 [82].

Таблица 5.7. Кривые Эдвардса для случая стандарта FIPS 186-2 – 2000

B-163:	$P(x) = x^{163} + x^7 + x^6 + x^3 + 1$, n=400000000000000000292FE7E70C12A4234C33
d_1	4
d_2	6AC25B85BADF8927593D21C366DA89C03969F3494
x_G	687540D72D2908FD3841129BCA7958858B6C5C2BE
y_G	699BB09A27DAD0407498C71F6F64BC9209430B2F0
B-233:	$\boldsymbol{P(x) = x^{233} + x^{74} + 1}$, n=1000000000000000000000000000013E974E72F8A6922031D2603CFE0D7
d_1	2
d_2	61FE1589EE5E1D39D1FB8C781B5C72ABBA94BC8494F97E51
x_G	107D76A232756C62932C89C879CDD4A452072D06663DDE171
y_G	11A4B8C4C816DE81462BD446A6BE32B22750BB2FA6AF5183
B-283:	$P(x) = x^{283} + x^{12} + x^7 + x^5 + 1$, n=3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFEF90399660FC938A90165B042A7CEFADB307
d_1	8
d_2	6DCAF32715E4AC7AFA0660C227EDBE0D68BF48828C11093285F020AB85D10ABB7A6EB70
x_G	6AD770485BE05F6282F2C9702B8B6EEF71CDB66E3B47BCEB988045F1CB8DE8F82A5CC31
y_G	148E3745CBC60604095D85F56B6E41328A91E7341D954AC04AE63BF3B1AD12CCD68B76E
B-409:	$\boldsymbol{P(x) = x^{409} + x^{87} + 1}$, n=1000000000000000000000000000000000000000000000000001E2AAD6A612F33307BE5FA47C3C9E052F838164CD37D9A21173
d_1	2
d_2	1A64D7DF91DA3B8BBC8BE6D28A7D4F2D764EACBB42B5F395F06C9927F7A2565433D95D9DCCC9B14A2F2389EE1CAB33C2F0
x_G	C76FF6F8C191F824D47A1858A8ABD7307288305F929812EC9F9DC117FF39AAC2CA34F8C69BE6DA8741585D1015254D75B5149

y_G	1F6AF8ADC131E10CC369DDF66212EF8E3904F35F641EAFBF 9A3C907E9B08BB01793C25A7F7F6A8EEE144824D4DA16B146B24
B-571:	$P(x) = x^{571} + x^{10} + x^5 + x^2 + 1$, n=3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF FFFFFFFFFFFFFFFFFFFFFFE661CE18FF55987308059B186823851EC 7DD9CA1161DE93D5174D66E8382E9BB2FE84E47
d_1	4
d_2	7F32D556640C20B5DD739A058DFFD58268D41C59135429EB0 41D7AA1255902E6362C4800A874AB0B60536B58460CD20C06F034
x_G	8452FAF52887F029A9F928ED8D074ADEEBC27AED0C7F0824 52AA37A72C7E642281893E50AFD96AB2B68FAD5E4DC7DD0941
y_G	7DAFEB6704D7FE96DF3611B29A4FBBC8DE308DABD10C55 55901F01A813D5F68C135ECADF8721900FB4272EE988D88C301C8

Таблица 5.8. Кривые Эдвардса для случая стандарта ДСТУ 4145 – 2002

m = 163	$P(x) = x^{163} + x^7 + x^6 + x^3 + 1$, n=40000000000000000002BEC12BE2262D39BCF14D
d_1	8
d_2	768F690F32DC034F5DEDB24C8FB319ED8B486B3DA
x_G	559682C8E8BBC689464D0D7E621E98BBCD8DABD2
y_G	B7F0D7B1B708E25CAA1BE95F45837DCAECC5C0F0
m = 167	$P(x) = x^{167} + x^6 + 1$, n=3FFFFFFFFFFFFFFFFFFFFFB12EBCC7D7F29FF7701F
d_1	2
d_2	7545508021921A6B02567AA6A4417DC34053497001
x_G	4F162059495517E97160CD0F5E9A7479EA77878A34
y_G	22222912562574D05A1F8C5C99521C737C2114BCEE
m = 173	$P(x) = x^{173} + x^{10} + x^2 + x + 1$, n=8000000000000000000189B4E67606E3825BB2831
d_1	1
d_2	1D182BC81F177A6A25AA0137888AF3E6A3DCFBAF1DDD
x_G	B3F998920A5710E0ADD293A70643A69699E49293F11
y_G	1EABC4D59CB70A65D6C5E3540BF8B28DC5CB07AD585A
m = 179	$P(x) = x^{179} + x^4 + x^2 + x + 1$, n=3FFFFFFFFFFFFFFFFFFFFFFB981960435FE5AB64236EF

d_1	2
d_2	34437707C8E046F16D6C49D2815725284CE51C7F993D2
x_G	57A15986F8B562BE752B53CA3BF400CC21368BCA9F693
y_G	369D420DD0B4D1E8CB620D1226E7D8C9FDFE138F617C0
m = 191	$P(x) = x^{191} + x^9 + 1$, n=4000000000000000000000069A779CAC1DABC6788F7474F
d_1	8
d_2	3E9D061E228EDCA3D70943FBA58882CEED93F23D2F0888
x_G	4753A3DEE912C24C9298B9F06AA7E27583949BF09462C67
y_G	389B5804EE0CD12CB108D2034603B930855979CB5AD3077
m = 233	$P(x) = x^{233} + x^9 + x^4 + x + 1$, n=1000000000000000000000000000013E974E72F8A6922031D2603CFE0D7
d_1	4
d_2	168D50202F1C874FE7BCE08CF64ABF78D959DB56A53EE3
x_G	F2F608558FED69CC31BDAA9BB4842265F9E9AD8457596272C00AB6DBB76D5BE58ACED3FD4B0794E0ECE21E84A41
y_G	
m = 257	$P(x) = x^{257} + x^{12} + 1$, n=800000000000000000000000000000006759213AF182E987D3E17714907D470D
d_1	1
d_2	129412CB0FA992A6B6A6BEFEF740F83E1AE6C17BE4D4F36
x_G	9FD3727044903DF7449203878D034D87964052D64664D583A0ED69B37D4D6A61
y_G	1A898787363EBB480F20A1AC759F498378375F5066468A137135346F50E16373
m = 307	$P(x) = x^{307} + x^8 + x^4 + x^2 + 1$, n=3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFC079C2F3825DA70D390FBBA588D4604022B7B7
d_1	6
d_2	6A577FE886FF3BD047B314D9E094DCDC2E9F042EE3338869E776CB6075C4A0A169BE053417AD7

x_G	6AAD22CC933A1BE6349B802D17DA70E48C480A6FA4C6C099F39CFEB270178D49A64BF92A8D262
y_G	1055BE77CBB48AF00CF6A4C7E6D27EBA27D77442C4D0A8B99893D978369163FD8A908125CCD8
m = 367	$P(x) = x^{367} + x^{21} + 1$, n=400000000000000000000000000000000000000000009C300B75A3FA824F22428FD28CE8812245EF44049B2D49
d_1	2
d_2	47AE47377615E10041C7DF34C9FC6EFB9A589B6405C6BC15BA9AAC186B3CE4C09446C5AB4DFFC199CD94F1C37373
x_G	7934BB36E0E76EDC55C2BC68476B3E8157C6C731EF07B9D0B9D31623CEAC1A46C4A7B6B0EBAFC2C16D93E3794997
y_G	4471E442DC9FC47DFB60CE95E312B2209FC50000CBEABA9294DC2E7D612CFC5246894A55B9702DE055892B2B7DE2
m = 431	$P(x) = x^{431} + x^5 + x^3 + x + 1$, n=3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFBA3175458009A8C0A724F02F81AA8A1FCBAF80D90C7A95110504CF
d_1	8
d_2	26C08DD2E5BE99638D8123FAAA6FFEE4965F85CBDA46C906F1A42F66150C49928B00BE2B4A6A3E9D1F7AE09D9984B681
x_G	198E2AD22D178117E7FEBF738913EFAF8C655BC49E96E3F314B3D788671C872ECFE6D93BBE561A114E6B4A251503BC2F5
y_G	61A19032C14D0A9CE7A5A0289AF4299C9FEC94BEB2A5F86F8B095FEB7657403A00EF759C9BA4DEDE26D4C5CDF018DF47245781C23EE

Всего нами было найдено еще по пять изоморфных кривых Эдвардса для каждого случая – в общей сложности получено 90 кривых для различных степеней m расширения поля F_{2^m}. Для поля определенной битовой длины полученные кривые Эдвардса равнозначны стандартным каноническим с точки зрения стойкости и обладают высокой производительностью. Приведенные в данной работе кривые Эдвардса с малыми значениями параметра d_1 требуют выполнения меньшего числа операций в поле, что дает дополнительный выигрыш.

Итак, в данном разделе мы рассмотрели альтернативную форму эллиптической кривой над расширенным полем характеристики 2, а именно, кривую в форме Эдвардса. Она имеет ряд характерных свойств, интересных для практического применения рассматриваемых кривых. В разделе даны формулы изоморфного преобразования между формой Вейерштрасса и формой Эдвардса кривой над полями характеристики 2. На основе общесистемных параметров двух действующих стандартов (ДСТУ 4145 – 2002 и FIPS 186-2 – 2000) мы приводим кривые Эдвардса, изоморфные каноническим эллиптическим кривым в этих стандартах. Для каждой из найденных кривых Эдвардса рассчитаны и приведены в соответствующих таблицах координаты генератора криптосистемы. Полученные кривые Эдвардса могут быть рекомендованы к реализации в проектируемых криптосистемах.

Главными преимуществами кривых в форме Эдвардса перед каноническими эллиптическими кривыми являются высокая скорость вычислений, полнота закона сложения и наличие аффинных координат нейтрального элемента аддитивной группы точек кривой.

5.2.3. Кривые Эдвардса с одним параметром над полями характеристики 2

Задачей исследования этого раздела является поиск бинарных кривых Эдвардса, наиболее приемлемых для криптографии [43]. Анализ оценок

сложности операций сложения и удвоения точек кривой Эдвардса над полем $F_2{}^m$ приводит к выводу, что наибольшая производительность присуща кривым с одним параметром $d = d_1 = d_2$. Между несуперсингулярными кривыми и кривыми Эдвардса в общем виде над полями $F_2{}^m$ существует изоморфизм [7]. В этом разделе мы находим условия, при которых для данной эллиптической кривой найдется изоморфная кривая Эдвардса с одним параметром d. Для известных канонических кривых из национальных стандартов (ДСТУ 4145 – 2002 [82] и FIPS 186-2 – 2000 [78]), удовлетворяющих полученным условиям, мы нашли изоморфные кривые Эдвардса с одним параметром d . В случае ДСТУ 4145 – 2002 таких кривых две, в американском стандарте FIPS 186-2 – 2000 данные условия выполняются для четырех кривых Коблица с параметром $a = 0$.

Приведем оценки сложности выполнения групповых операций на кривой Эдвардса в проективных координатах [7].

Очевидно, что производительность криптосистемы в значительной мере зависит от ее параметров, и в случае кривых над полями $F_2{}^m$, актуален вопрос нахождения таких коэффициентов d_1, d_2 кривой Эдвардса, при которых будет достигаться максимальная скорость выполнения операций.

Проблема инверсии в формулах сложения (5.4) и удвоения (5.5) для кривой, заданной в аффинных координатах, решается переходом к проективным координатам. Подставив в уравнение (5.3) $x = \frac{X}{Z}, y = \frac{Y}{Z}$ и умножив обе его части на Z^4 (при $Z \neq 0$) получим однородное уравнение кривой Эдвардса над полем $F_2{}^m$ с теми же параметрами d_1, d_2:

$$d_1(X+Y)Z^3 + d_2(X^2+Y^2)Z^2 = XYZ^2 + XY(X+Y)Z + X^2Y^2,$$
$$X, Y, Z \in F_{2^m} \qquad (5.9)$$

Помимо точек вида $(\alpha X : \alpha Y : \alpha Z)$ при $Z \neq 0$ и $a \in F_2{}^{m*}$, которые соответствуют точкам (x, y) аффинного представления, уравнению (5.9) удовлетворяют еще две точки с проективными координатами $(1:0:0)$ и $(0:1:0)$. Обе являются сингулярными.

Согласно [7], сложение в проективных координатах для кривых Эдвардса в общем случае реализуется за $V_{E d_1, d_2} = 21M + 1S + 4U$ операций в поле. Аналогичная величина для удвоения составляет и $W_{E d_1, d_2} = 2M + 6S + 3U$

операций. Здесь M, S, U – сложность умножения, возведения в квадрат и умножения на параметры d_1, d_2 в поле $F_2{}^m$. Главным преимуществом кривых Эдвардса над полями $F_2{}^m$, как отмечалось, является полнота и универсальность закона сложения (5.4) [7, 8]. Производительность же арифметики данных кривых в общем случае не является максимальной [7]. Однако приведенные оценки сложности можно улучшить, если принять значения параметров кривой Эдвардса над полем $F_2{}^m$ равными между собой. Другими словами, при $d_1 = d_2 = d$ имеем кривую в аффинных координатах вида

$$E_d: \quad d(x + y + x^2 + y^2) = xy + xy(x + y) + x^2 y^2 \tag{5.10}$$

с соответствующим представлением в проективных координатах:

$$d\left((X + Y)Z^3 + (X^2 + Y^2)Z^2\right) = XYZ^2 + XY(X + Y)Z + X^2Y^2,$$
$$d, X, Y, Z \in F_{2^m},\ d \neq 0 \text{ и } d \neq t^2 + t,\ \forall t \in F_{2^m}. \tag{5.11}$$

Для этого случая формулы сложения и удвоения будут иметь меньшую сложность: $V_{Ed} = 16M + 1S + 4U$ и $W_{Ed} = 2M + 5S + 2U$ операций в поле $F_2{}^m$ [6].

Логично поставить вопрос, при каких условиях для данной канонической кривой можно найти изоморфную кривую Эдвардса вида (5.10) над полем F_{2^m} и как связаны параметры таких кривых.

Каноническая эллиптическая кривая (или несуперсингулярная кривая) задана над полем $F_2{}^m$ аффинным уравнением

$$v^2 + uv = u^3 + a_2u^2 + a_6, \qquad a_6 \neq 0. \tag{5.12}$$

При построении кривых Эдвардса вида (5.3), изоморфных кривым вида (5.12) в работе [39] мы выбирали значение параметра d_1 так, чтобы выполнялись два условия: $Tr(d_1) = Tr(a_2) + 1$ и $Tr\left(\frac{\sqrt{a_6}}{d_1{}^2}\right) = 1$. Далее вычисляли

значение другого параметра по формуле $d_2 = d_1^{\,2} + d_1 + \dfrac{\sqrt{a_6}}{d_1^{\,2}}$ [7]. Пусть для кривой (5.12) существует изоморфная кривая Эдвардса вида (5.10). Тогда, принимая $d_1 = d_2 = d$, получим систему

$$\begin{cases} d = d^2 + d + \dfrac{\sqrt{a_6}}{d^2}, \\ Tr(d) = Tr(a_2) + 1, \\ Tr(\dfrac{\sqrt{a_6}}{d^2}) = 1. \end{cases} \qquad (5.13)$$

Возьмем функцию следа от обеих частей первого уравнения системы, тогда с учетом $Tr(d) = Tr(d^2)$ получим $Tr(d) = Tr(\dfrac{\sqrt{a_6}}{d^2})$. Теперь из 2-го и 3-го уравнений системы (5.13) сразу следует, что

$$Tr(a_2) = 0. \qquad (5.14)$$

Первая формула в системе (5.13) позволяет вычислить для изоморфной кривой Эдвардса единственное значение параметра *d*

$$d^2 + \frac{\sqrt{a_6}}{d^2} = 0, \qquad \Rightarrow \qquad d = \sqrt[8]{a_6}. \qquad (5.15)$$

Условие $d \neq 0$ в (5.10) и существование квадратного корня у каждого элемента поля $F_2{}^m$ обеспечивает разрешимость данного равенства.

Равенства (5.14), (5.15) задают изоморфизм между кривыми вида (5.12) и (5.10). Из всех несуперсингулярных кривых ровно половина кривых со следом $Tr(a_2) = 0$ отвечает этим условиям. Так как 8 не делит порядок мультипликативной группы поля $(2^m - 1)$, для каждого ненулевого параметра a_6 кривой (5.10) существует единственное значение параметра $d = \sqrt[8]{a_6}$ изоморфной кривой Эдвардса и обратно: $a_6 = d^8$ – единственное значение для каждого *d*.

Из (5.14) следует, что все несуперсингулярные кривые (5.12) и

изоморфные им кривые Эдвардса вида (5.10) имеют порядок, кратный 4 [29]. В качестве примера можем взять две кривые действующего стандарта ДСТУ 4145 – 2002 над полями со степенью расширения $m = 173$, $m = 257$ соответственно, которые удовлетворяют условиям (5.14), (5.15). В таблице 5.9 приведены рассчитанные параметры кривых в форме Эдвардса вида $d(x + y + x^2 + y^2) = xy + xy(x + y) + x^2 y^2$, изоморфных данным кривым в канонической форме над соответствующим полем, а также координаты генераторов криптосистемы.

Таблица 5.9. Кривые Эдвардса с одним параметром, изоморфные каноническим кривым, рекомендованным стандартом ДСТУ 4145 – 2002 над полями $F_2{}^m$ при $m = 173,\ m = 257$.

$m =$ 173	$P(x) = x^{173} + x^{10} + x^2 + x + 1,$
	n=800000000000000000000189B4E67606E3825BB2831
d	194E92F31C2F97583B5B079A66498651CFF964D4EDC1
x_G	7F533BD52D2DDEAA0A26A9B1547AD25257F26E7E1BD
y_G	1310000C9C9510CA7EEAE41922062276B7C0EAE85512
$m =$ 257	$P(x) = x^{257} + x^{12} + 1,$
	n=800000000000000000000000000000006759213AF182E987D3E17
d	1A561C01C65FDA18821A29F0299C88C26C8A85C4F5F326B
x_G	12F1EFDA68924370FA29955383A6FCAB88CAD1E67BF6141
y_G	18E4BCBC54AD766E6834C60BBC8630661D42C8B075E024A

Взяв за основу американский стандарт FIPS 186-2 – 2000 и принимая во внимание условия (5.14), (5.15), можем заметить, что каждой из 4-х кривых Коблица с параметром $a = 0$ изоморфна кривая Эдвардса с параметрами $d_1 = d_2 = 1$, над различными полями F_{2^m} простых степеней m расширения. В таблице 5.10 мы приводим координаты генераторов криптосистемы на кривых Эдвардса вида $x + y + x^2 + y^2 = xy + xy(x + y) + x^2 y^2$ изоморфных данным кривым Коблица над соответствующим полем.

Таблица 5.10. Кривые Эдвардса с одним параметром, изоморфные кривым Коблица стандарта FIPS 186-2 – 2000 над полями $F_2{}^m$ при $m = 233$, $m = 283$, $m = 409$, $m = 571$

K-233:	$P(x) = x^{233} + x^{74} + 1$, n=8000000000000000000000000000069D5BB915BCD46EFB1AD5F173ABDF
x_G	FACA9831F3C99277CF551679FE7245D52B11A048FC45C0B78
y_G	19BCD68FB073EBC8437C67422B12751BB1279087397553C881
K-283:	$P(x) = x^{283} + x^{12} + x^7 + x^5 + 1$, n=1FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFE9AE2ED07577265DFF7F94451E061E163C61
x_G	5AD205F99526557BC0731B3991187B66FCE4D30386D0864AE5318A21FFB4A96BD88CE9B
y_G	10142CE6C38325511BB593DEA87393A70D9CDB25D097DEF2
K-409:	$P(x) = x^{409} + x^{87} + 1$, n=7FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFE5F83B2D4EA20400EC4557D5ED3E3E7CA5B4B5C83B8E01E5FCF
x_G	1CE08B54A8E78B891EB7E8D1DF23A74B0A4976E3C2FD33E984FEA99D2CD03AA3633127ADA8858E3854E7EBDF442F300FAE4
y_G	CCFE11FE0FD07C5F196CCA8672276D10C9489D344C33E230185BA43CDF3F505BE98285BCD2547BB274707811736A323B7E6FE
K-571:	$P(x) = x^{571} + x^{10} + x^5 + x^2 + 1$, n=2000000000000000000000000000000000000000000000000000000000000000000000000000131850E1F19A63E4B391A8DB917F4138B630D84BE5D639381E91DEB45CFE778F637C1001
x_G	1653830605B1343AC3BB092259163566A374B13FF921A3EE4DD9C8B189443DE17B83D18B37B79870DFE06ABF1DCD13F22B4EC
y_G	7CC37E95C30C6C3DC8286F24323FAFF14BF2EC60F0574DEBAA9F2AD2F08622CBCFEBA00DE29CA7D6B31E8F4AE96D3B9857

Закон сложения для кривых Коблица не обладает свойством полноты и универсальности (в отличие от случая кривых Эдвардса), что можно трактовать как их недостаток. Однако, сложность групповой операции в случае кривых Коблица все же будет меньшей, чем в случае изоморфных кривых Эдвардса, поэтому нельзя сделать однозначный вывод о превосходстве одной формы рассматриваемых кривых над другой.

Исходя из имеющихся оценок сложности групповой операции [7], а также формул изоморфного преобразования [7, 8] между кривыми Эдвардса и кривыми в форме Вейерштрасса над полями F_{2^m}, мы получили условия существования кривой Эдвардса с одним параметром, изоморфной кривой в канонической форме. Далее, мы вычислили искомые значения параметров, соответствующие двум кривым из стандарта ДСТУ 4145 – 2002 (при $m = 173$ и $m = 257$) и четырем кривым стандарта FIPS 186-2 – 2000 (при $m = 233$, $m = 283$, $m = 409$, $m = 571$).

В итоге можно констатировать, что сравнительно немного кривых над полями F_{2^m} из рассматриваемых стандартов удовлетворяют условию (5.14). Для нахождения большего числа быстрых кривых Эдвардса необходимо разработать методы поиска новых таких кривых с почти простым значением порядка.

ГЛАВА 6

ПРОТОКОЛЫ КРИПТОСИСТЕМ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Специфика асимметричных криптосистем определила их три основные задачи: распределение ключей по открытым каналам, направленное шифрование и цифровая подпись. Уже в исторической работе Диффи и Хэллмана [69] предложена схема с разделением секрета для обмена ключом симметричной криптосистемы по открытому каналу, использующая одностороннюю функцию экспоненцирования $g^x \bmod p$ в простом поле. Ее стойкость, как известно, основана на проблеме дискретного

логарифмирования (*DLP*) в простом поле $\boldsymbol{F}_p$. Решение двух других задач с помощью этой функции было впервые дано Эль-Гамалем в [70], где для построения односторонней функции «с лазейкой» предложена идея использовать два ключа. Появившиеся почти одновременно криптосистемы на эллиптических кривых [13,14] наследуют основные протоколы Эль-Гамаля-подобных систем, поскольку их алгебраические структуры построены на конечных группах элементов. Особенность эллиптических кривых состоит в том, что все криптоалгоритмы с алгеброй поля адаптируются к аддитивной группе точек кривой с операцией скалярного произведения $kG \bmod n$ точки вместо экспоненцирования в мультипликативной группе поля $g^k \bmod p$. При этом параметры шифрования могут оказаться как точкой кривой, так и элементом поля, определяющим, например, параметр цифровой подписи или одну из координат точки. Мы рассмотрим основные протоколы в порядке нарастания сложности.

Для функционирования криптосистемы все ее пользователи снабжаются общесистемными параметрами (ОСП), определяемыми на этапе разработки. К ним относятся:

- конечное поле F_q, над которым определена кривая ($q = p^m$ или $q = 2^m$, p и m – простые числа);
- уравнение кривой в форме Вейерштрасса W (или Эдвардса E) с известными коэффициентами a и b (или a и d);
- порядок кривой $N_E = 2^i cn$ с простым множителем $n > 2^{180}$ и малым кофактором $2^i c$, c – нечетное число 1 или 3, $i = 0, 1, 2$;
- координаты точки G – генератора криптосистемы порядка n.

Кроме того, каждый пользователь генерирует с помощью датчика случайных чисел долговременный секретный ключ e как целое число в интервале $0 < e < n - 1$ и рассчитывает открытый ключ как точку кривой $Q = eG$. Первый хранится в секрете, тогда как второй после сертификации доступен другим пользователям криптосистемы.

После задания и вычисления ОСП при проектировании криптосистемы последние проходят ряд тестов на стойкость к наиболее известным атакам. Аномальные кривые над простым полем, суперсингулярные кривые (кривые с нулевым j-инвариантом), и кривые, не прошедшие теста на MOV-атаку,

являются криптографически слабыми и не рекомендуются к применению современными стандартами.

Материал этой главы в основном опирается на на обзор протоколов и стандартов, рассмотренных в предыдущей работе [29] автора.

6.1. Протоколы распределения ключей

Воспользовавшись аналогией между экспоненцированием элементов в мультипликативной группе поля и k-кратным сложением точки кривой, нетрудно построить протокол Диффи-Хэллмана [69] обмена секретным ключом по открытому каналу на основе технологии эллиптических кривых.

Пользователи A и B хотят выработать общий секретный ключ K_{AB} для симметричного шифрования своей секретной переписки. Им известны ОСП и открытые ключи $Q_A = (x_A,y_A)$ и $Q_B = (x_B,y_B)$ друг друга. Их секретные ключи e_A и e_B они могут использовать для формирования общего секретного ключа по схеме разделения секрета. Для этого используется следующий неинтерактивный протокол:

1. Пользователь A вычисляет точку $V = e_AQ_B = e_Ae_BG$.
2. Пользователь B вычисляет точку $V = e_BQ_A = e_Be_AG$. Результаты вычислений A и B совпадают в силу коммутативности абелевой группы точек.
3. В качестве общего секретного ключа оба пользователя могут принять, например, x-координату точки $V = (x_V, y_V)$, т.е. $k_{AB} = x_V$.

В интерактивном протоколе формирования разовых ключей симметричного шифрования для одного сеанса секретной связи пользователи A и B используют не долговременные ключи друг друга, а разовые секретные ключи k. При этом они выполняют следующие действия:

1. Пользователь A генерирует случайное секретное число $0 < k_A < n$, вычисляет точку $R_A = k_AG$ и отправляет ее пользователю B.

2. Пользователь B генерирует случайное секретное число $0 < k_B < n$, вычисляет точку $R_B = k_BG$ и отправляет ее пользователю A.

3. Пользователь A вычисляет точку $k_AR_B = k_Ak_BG = V_{AB}$, аналогично пользователь B вычисляет точку $k_BR_A = k_Bk_AG = V_{BA}$. В качестве общего

секретного ключа они могут принять x-координату точки $V_{AB} = (x_V, y_V)$, т.е. $k_{AB} = x_V$. Длина ключа, как видим, определяется размером поля в m бит.

Если третье лицо (злоумышленник С) попытается определить общий секрет пользователей, то ему потребуется вычислить $k_A k_B$ по известным $k_A G$, $k_B G$ и G. Эта проблема называется проблемой Диффи-Хэллмана для криптосистем на эллиптических кривых. Для ее решения приходится дважды решать проблему дискретного логарифмирования на эллиптической кривой.

Последняя из схем весьма распространена, однако она совершенно не защищена от противника С, который имеет доступ к каналу связи и может подменять пересылаемые точки R_A и (или) R_B своими точками $R_C = k_C G$. Он, таким образом, может либо выступать от имени одного из пользователей, установив секретную связь с другим, либо, контролируя канал, быть транслятором их переписки, свободно расшифровывая и читая все сообщения. Такого активного криптоаналитика С называют «man in between». Этот пример показывает, что для защиты от перехвата и подлога чрезвычайно важной становится задача аутентификации (установления подлинности) пользователей. Рассмотрим протокол, в котором предусмотрена такого рода защита. Он описан, например, в проекте стандарта IEEE P1363 [74,75].

Elliptic Curve Key Establishment Protocol (ECKEP), предложенный Менезисом, Кью и Ванстоуном (A. Menezes, M.Qu, S. Vanstone) и называемый MQV-протоколом, предполагает использование как долговременных ключей e_A и e_B, так и разовых ключей k_A, k_B, пользователей. Для формирования общего секретного ключа K_{AB} пользователи выполняют следующие действия:

1. Пользователь А:

(a) Генерирует случайное целое число $0 < k_A < n$.

(b) Вычисляет точку $R_A = k_A G$.

(c) Вычисляет точку $Y_A = k_A Q_B = (x_1, y_1)$.

(d) Вычисляет целое число $s_A = (k_A + e_A x_A x_1) \bmod n$.

(e) Отправляет точку R_A пользователю В.

2. Пользователь B:

(a) Генерирует случайное целое число $0 < k_B < n$.

(b) Вычисляет точку $R_B = k_B G$.

(c) Вычисляет точку $Y_B = k_B Q_A = (x_2, y_2)$.

(d) Вычисляет целое число $s_B = (k_B + e_B x_B x_2) \bmod n$.

(e) Отправляет точку R_B пользователю A.

3. Пользователь A:

(a) Вычисляет точку $Y_B = e_A R_B. = (x_2, y_2)$.

(b) Вычисляет точку $K = e_A (R_B + x_B x_2 Q_B)$.

4. Пользователь B:

(a) Вычисляет точку $Y_A = e_B R_A. = (x_1, y_1)$.

(b) Вычисляет точку $K = s_B (R_A + x_A x_1 Q_A)$.

Нам остается показать, что в результате вычислений в соответствии с п.3(b) и 4(b) пользователи получают одну и ту же точку. Действительно, с учетом равенства $Q_B = e_B G$ и согласно соотношений 2 (b), (d) и 3 (b) имеем

$$K = s_A (R_B + x_B x_2 Q_B) = s_A (k_B + e_B x_B x_2)\mathrm{G} = s_A s_B G. \qquad (6.1)$$

Аналогично вычисления пользователя B согласно 4 (b) дают

$$K = s_B (R_A + x_A x_1 Q_A) = s_B (k_A + e_A x_A x_1)\mathrm{G} = s_B s_A G. \qquad (6.2)$$

Ясно, что в связи с коммутативностью операции сложения в группе W или E результаты совпадают. Как описывалось ранее, координаты секретной точки K могут быть известным способом преобразованы в разовый ключ k_{AB} симметричного шифрования.

Отметим, что описанный протокол можно модифицировать без расчета параметров s_A, s_B согласно п.1(d) и 2(d), а вычисляя точку K тождественными операциями с точками кривой W. Для этого в п.3 пользователь А вычисляет точку $V = (R_B + x_B x_2 Q_B)$, а затем точки $k_A V$ и $e_A x_A x_1 V$. Сумма двух последних точек, очевидно, дает точку K. Так же действует и пользователь В. Этот алгоритм, однако, более трудоемок, так как операции с точками сложней операций в поле $\mathbf{F}_q$.

Обсудим теперь, какие проблемы при выполнении этого протокола возникли у активного злоумышленника С («man in between»). Если бы пользователи просто складывали свои секретные ключи, т. е. $s_{A,B} = (k_{A,B} + e_{A,B}) \bmod n.$, то, очевидно, ситуация для С оставалась бы такой же благоприятной, как и в предпоследнем протоколе. Идея защиты от навязывания противником С своего разового ключа k_C состоит в том, что соотношение

$$s_A = (k_A + e_A x_A x_1) \bmod n. \qquad (6.3)$$

согласно п.1(d) нелинейно связывает ключи k_A , e_A и e_B , так как $x_1 = f(k_A, e_B)$. Тем самым противник С лишается возможности свободной подтасовки ключа k_C.. Чтобы рассчитать свое значение (6.3), в котором ему известно лишь x_A, противнику С придется определить долговременные ключи e_A и e_B, т.е. дважды вычислить дискретный логарифм в группе W (или E).

6.2. Протокол направленного шифрования

Направленное шифрование является специфической задачей асимметричной криптографии. Пользуясь для этой цели открытым ключом Q_A пользователя А, любой другой пользователь может зашифровать и послать

ему секретное сообщение. Расшифровать это сообщение способен лишь пользователь А с помощью своего секретного ключа e_A. Схема шифрования может быть построена по аналогии со схемой Эль-Гамаля [70] с тем отличием, что вместо мультипликативной группы поля $\mathbf{F}_p^*$ используется подгруппа <G> точек простого порядка n кривой W. Кроме того, приходится решать задачи согласования арифметики верхнего и нижнего уровней при шифровании сообщения.

Протокол ECES (Elliptic Curve Encryption Scheme), описанный в [75,76], предполагает следующие действия при передаче секретного сообщения M от пользователя В пользователю А:

1. Пользователь В:

(a) Генерирует случайное целое число $0 < k_B < n$.

(b) Вычисляет точку $R_B = k_B G = (x_1, y_1)$..

(c) Вычисляет точку $Y_B = k_B Q_A = (x_2, y_2)$.

(d) Представляет блок сообщения M в виде элементов (m_1, m_2) поля $\mathbf{F}_q$.

(e) Вычисляет пару элементов поля $(c_1, c_2) = (x_2 + m_1, x_2 + m_2)$.

(f) Отправляет пользователю А зашифрованное сообщение (x_1, y_1, c_1, c_2).

2. Пользователь А:

(a) Вычисляет точку $Y_B = d_A R_B = e_A (x_1, y_1) = (x_2, y_2)$..

(b) Расшифровывает сообщение $m_1 = c_1 - x_2$, $m_2 = c_2 - x_2$.

Отметим, что сложение (вычитание) в поле $\mathbf{F}_q$ можно заменить умножением (делением), что, естественно, усложняет вычисления.

Из данного протокола видно, что для реализации направленного шифрования приходится использовать пару секретных ключей: разовый ключ k_B пользователя В и долговременный ключ e_A пользователя А. Первый из них

выполняет функцию лазейки в односторонней функции, а второй делает возможным расшифрование. Пара ключей k_B, e_A по сути реализуют схему разделения секрета при ключевом обмене. Нетрудно видеть, что в данной схеме отсутствует аутентификация пользователя B, т.е. от его имени может отправлять сообщения любой пользователь. Защита от этой угрозы предусмотрена в алгоритмах цифровой подписи.

6.3. Алгоритмы цифровой подписи

При передаче электронных документов (файлов) по открытым сетям возникают задачи сохранения целостности документа (защиты от модификаций и подмен) и аутентификации (установления подлинности) отправителя этого документа. В принципе обе задачи можно решить одним методом, обратным направленному шифрованию: зашифровать пересылаемый документ секретным ключом пользователя A (отправителя). Такой шифртекст будет обладать следующим свойством: сформировать его может лишь пользователь A, а прочитать – все другие пользователи криптосистемы. Для этого они расшифровывают шифртекст с помощью общеизвестного открытого ключа пользователя A.

Подобное решение может быть приемлемым лишь при малых размерах документа *M*, не превышающих размерности поля. При асимметричном шифровании больших текстов приходится затрачивать на два-три порядка больше времени, чем при симметричном. Поэтому более рациональным решением можно считать двухэтапную процедуру: сжатие текста *M* до размеров поля с помощью однонаправленной хэш-функции (хэш-кода) *h*(M) на первом этапе, и шифрование *h*(*M*) секретным ключом пользователя A на втором. Такой зашифрованный хзш-код получил название *цифровой подписи DS* (*DS* – Digital Signature). Он обычно пересылается вместе с сообщением *M* в виде заключающей приставки, выполняя таким образом функцию обычной подписи под документом.

Цифровую подпись (ЦП) определяют как добавляемый к сообщению *M* блок данных *DS* небольшого размера, полученный в результате криптопреобразования *M* с помощью секретного ключа *e* отправителя и

позволяющий получателю удостовериться в целостности сообщения и подлинности источника. Цифровая подпись защищает получателя от угрозы отказа отправителя от своего сообщения M, а также отправителя – от подлога или фальсификации со стороны получателя. Главное свойство ЦП: сформировать подпись может лишь один пользователь (обладатель секретного ключа), а прочитать подпись могут все другие пользователи с помощью открытого ключа отправителя. Основные требования к цифровой подписи [16 – 19,28]:

- функция выработки цифровой подписи должна быть односторонней (сложно обратимой);
- ключи выработки цифровой подписи должны быть конфиденциальными ключами отправителя, а ключи проверки – открытыми ключами отправителя;
- цифровая подпись должна быть чувствительной к малейшим изменениям в тексте (в том числе случайным ошибкам в канале);
- вероятность формирования двух одинаковых подписей под двумя различными сообщениями должна быть ничтожно мала (порядка 2^{-m} при размерности поля в m бит);
- цифровые подписи для одного сообщения, переданные в разное время или различными терминалами, должны отличаться;
- вычислительная сложность формирования и проверки подписи должна быть полиномиальной и приблизительно одинаковой;
- уровень стойкости цифровой подписи должен быть дифференцируемым в зависимости от задаваемых требований.

Большинство алгоритмов цифровой подписи, привлекающих аппарат эллиптических кривых, базируется на схеме Эль-Гамаля [70]. В 1994 году был принят первый национальный стандарт цифровой подписи США FIPS 186 [73], рекомендующий алгоритм DSA (Digital Signature Algorithm) как модификацию схемы Эль-Гамаля в конечном поле. С 1999 года появились национальные и международные стандарты криптосистем на эллиптических кривых. Вместе с тем существующие стандарты и их проекты [74–85]

рекомендуют различные алгоритмы с разной степенью безопасности и вычислительной сложности, что расширяет возможности выбора при разработке криптосистемы. Рассмотрим наиболее распространенные алгоритмы.

Полагаем, что всем пользователям криптосистемы известны общесистемные параметры (ОСП): конечное поле $\mathbf{F}_q$, $q = p$ или $q = 2^m$, над которым определена кривая W или E с генератором – точкой G порядка n. Кроме того, все пользователи знают и применяют для сжатия сообщения одну и ту же одностороннюю криптографическую функцию хэширования $h(M)$. Для формирования подписи отправитель А сообщения M использует свой долговременный секретный ключ e_A.

Основные требования к функции хэширования:

- хэш-функция должна быть чувствительной к любым изменениям в тексте M (вставкам, пробелам, перестановкам, заменам и т.д.);

- хэш-функция должна быть однонаправленной (односторонней), т.е. задача подбора сообщения M под известное значение хэш-функции должна быть вычислительно неразрешима;

- хэш-функция должна быть однозначным образом сообщения *M;*

- вероятность коллизий (совпадений) хэш-кодов различных сообщений должна быть ничтожно мала (порядка 2^{-logq});

- алгоритм хєширования должен быть открытым, однако либо M, либо $h(M)$ должно быть скрыто от злоумышленника (зашифровано).

Часто для этих целей шифруется сообщение M. Для противодействия ряду других угроз осуществляется сертификация открытого ключа и идентификатора отправителя центром распределения ключей (сертификат подписывается секретным ключом центра).

Процедура цифровой подписи включает:

- формирование и проверку параметров пользователей;

- формирование (generation) цифровой подписи;

• проверку (verification) цифровой подписи.

6.3.1. ECDSA (Elliptic Curve Digital Signature Algorithm)

Этот алгоритм является адаптированным к арифметике эллиптических кривых алгоритмом цифровой подписи (DSA), принятого в американском стандарте DSS (Digital Signature Standard) FIPS-186 [73] и утвержденного Национальным Институтом Стандартов и Технологий (NIST) США в 1994 году. Алгоритм ECDSA был впервые предложен С. Ванстоуном еще в 1992 году, после чего он прошел длительный этап исследований на стойкость и различного рода усовершенствований и доработок. Лишь в конце 20-го века он был утвержден в ряде стандартав цифровой подписи: международном стандарте ISO/IEC CD 15946 1999 году [76], американском стандарте для финансовых служб X9.62 ANSI (American National Standard Institute) [77] в январе 1999 года, американском национальном стандарте FIPS-186-2 NIST (National Institute of Standards and Technology) [78] в январе 2000 года, проекте стандарта P1363 IEEE (Institute of Electrical and Electronics Engineers) [74] и в стандарте P1363 IEEE [75] в 2000 году и других. Хотя в последующие годы эти стандарты модифицировались, рекомендуемые алгоритмы и ОСП эллиптических кривых в них в основном остались прежними. Примером этого является последняя версия американского стандарта FIPS-186-4 PUB- 2013 [79].

Параметры пользователя

Пользователь А:

• генерирует и хранит в секрете долговременный секретный ключ как целое число $0 < e_A < n$;

• вычисляет открытый ключ как точку кривой $Q_A = e_A G$. Открытый ключ доступен для всех пользователей системы.

Формирование ЦП

Пользователь А:

1. Вычисляет хэш-код сообщения М как целое число $h = h(M),\ h < n$.
2. Генерирует случайное целое число $0 < k_A < n$.
3. Вычисляет точку $R = k_A G = (x_1, y_1)$.
4. Вычисляет параметр $r = \pi(R) \bmod n$. При $r = 0$ возврат в п.2.
5. Вычисляет обратный элемент k_A^{-1} простого поля $\mathbf{F}_n$.
6. Вычисляет параметр $s = k_A^{-1}(h + e_A r) \bmod n$. При $s = 0$ возврат в п.2.
7. Направляет пользователю В подписанное сообщение (M, r, s), в котором $DS = (r, s)$ – цифровая подпись.

Заметим, что в п.4 преобразование точки R в целое число предполагает, что ее x-координата как элемент x_1 поля $\mathbf{F}_q$ каким-то способом переводится в целое число $\bar{x}_1$ с последующей редукцией по модулю n ($r = \pi(R) = \bar{x}_1 \bmod n$).

Для данной схемы можно записать ключевое уравнение

$$sk_A = (h + e_A r) \bmod n, \qquad (6.4)$$

связывающее параметры s и r с разовым и долговременным ключами k_A и e_A. Оно используется при формировании и проверке цифровой подписи. Параметр s определен в п.6 протокола на основе этого уравнения.

Проверка ЦП

Пользователь В проверяет цифровую подпись пользователя А, имея в распоряжении следующую информацию: открытый ключ пользователя А Q_A, общесистемные параметры (ОСП), алгоритм хэширования $h(M)$ и подписанное сообщение (M, r, s). Суть проверки состоит в вычислении на основе известных данных параметра r' и сравнении его с принятым значением r.

Умножив (6.4) на инверсию s^{-1} второго параметра подписи и учитывая, что $Q_A = e_A G$, для точек криптосистемы в результате экспоненцирования (скалярного произведения) получим равенство

$$k_A G = s^{-1} hG + s^{-1} r\, e_A G = uG + vQ_A, \qquad (6.5)$$

$$u = s^{-1} h \bmod n, \quad v = s^{-1} r \bmod n.$$

Согласно пп.3 и 4 протокола формирования левая часть этого равенства определяет точку $R = (x_1, y_1)$ и, соответственно, параметр $r = \bar{x}_1 \bmod n$. Правая часть равенства включает известные получателю данные, которые он использует для вычисления параметра r' (он может оказаться отличным от параметра r при модификациях сообщения M и ошибках в канале связи). Итак, протокол проверки ЦП на основе (6.5) включает следующие вычисления.

Пользователь В:

1. Вычисляет хэш-код полученного сообщения M: $h = h(M)$, $h < n$.
2. Вычисляет обратный элемент $s^{-1} \bmod n$ поля $\boldsymbol{F}_n$.
3. Вычисляет параметры $u = s^{-1} h \bmod n$, $v = s^{-1} r \bmod n$.
4. Вычисляет точку $R' = uG + vQ_A = (x_1', y_1')$.
5. Вычисляет параметр $r' = \pi(R') = \bar{x}_1' \bmod n$.
6. Сравнивает вычисленное r' и принятое значения r. При равенстве $r' = r$ цифровая подпись верна, в противном случае она отвергается.

В результате проверки пользователь В удостоверяется в подлинности отправителя А и целостности сообщения M.

В ряде проектов и стандартов определение параметра r подписи не регламентируется, а задается функцией $r = \pi(x_1, y_1) \bmod n$. Это, в частности, позволяет избежать неоднозначности определения $r = \bar{x}_1 \bmod n$ в связи с

наличием обратной точки $(n - k_A)G = -k_AG$, имеющей ту же x-координату, что и точка k_AG (и, следовательно, совпадающий параметр r).

Некоторые атаки на ECDSA

В качестве злоумышленника в асимметричной криптосистеме может выступать как третье лицо (криптоаналитик, хакер, квакер), так и любой законный пользователь системы. Существует большой перечень возможных угроз со стороны всех пользователей открытой компьютерной сети [15-18,27]. В частности, отправитель А может:

- отказаться от факта передачи подписанного сообщения (ренегатство А);
- не передать подписанное сообщение и утверждать, что он его отправил;
- передать подписанное сообщение в момент времени t_1 и утверждать, что он его отправил в момент t_2;
- передать подписанное сообщение M и утверждать, что в действительности передал подписанное сообщение M' (подмена А) и др.

С другой стороны, недобросовестный (или преступный) получатель В может совершить следующие действия:

- отказаться от факта получения подписанного сообщения (ренегатство В);
- сфальсифицировать получение подписанного сообщения, сформировав и подписав его (подлог В);
- получить подписанное сообщение в момент времени t_1 и утверждать, что получил его в момент t_2;
- получить подписанное сообщение M и утверждать, что в действительности получил модифицированное подписанное сообщение M' (подмена В) и др.

Кроме санкционированных пользователей в криптосистеме может совершать те или иные правонарушения несанкционированный злоумышленник (противник, аналитик С):

- перехватить сообщение, возможно, модифицировать его и с задержкой переслать адресату (пассивный или активный перехват);

- сформировать ложное сообщение, подписать его от имени законного пользователя и направить другому пользователю (маскарад С);

- попытаться на основе криптоанализа определить ключи законного пользователя, в результате получив возможность подписывать его документы и читать его секретную информацию;

- осуществить повторную передачу документа пользователя А пользователю В и др.

Алгоритмы и протоколы цифровой подписи должны обеспечивать безопасность любого пользователя в отношении всех этих и других угроз. Для разрешения споров между сторонами существует независимый арбитр-эксперт, который на основании всех несекретных данных должен доказать вину какого-то пользователя или констатировать другие события (скажем, взлом ключа неизвестным злоумышленником). Юридические отношения сторон и экспертиза в спорных ситуациях должны регламентироваться законом о цифровой подписи [86].

Одной из самых известных атак на цифровую подпись является атака при повторном использовании разового ключа. В этом случае пользователь А для двух разных сообщений M_1 и M_2 (с соответствующими хэш-кодами h_1 и h_2) при формировании подписи повторно использует одно значение ключа k_A. В соответствии с (6.4) имеем

$$s_1 k_A = (h_1 + e_A r) \bmod n, \quad s_2 k_A = (h_2 + e_A r) \bmod n,$$

откуда

$$k_A = (h_1 - h_2)/(s_1 - s_2) \bmod n,$$

$$e_A = r^{-1} (s_1 k_A - h_1) \bmod n.$$

Таким образом, повторное использование разового ключа позволяет легко определить пару секретных ключей пользователя А и реализует угрозу «полное раскрытие». Знание злоумышленником долговременного ключа пользователя А дает ему возможность читать его секретные письма и отправлять сообщения и распоряжения от его имени. Поэтому смена разового ключа с уничтожением предыдущего является необходимым условием безопасности для каждого пользователя криптосистемы. Вместе с тем рекомендуется периодически менять и долговременный ключ. Понятно, что с увеличением срока действия ключа e растет вероятность его компрометации в силу возможных ошибок или других действий персонала.

Другой возможной угрозой является так называемая «селективная подделка». Она реализуется, если злоумышленнику удается сформировать на основе ложного сообщения M' ту же цифровую подпись, что и для истинного сообщения M. Поскольку асимметричная криптография работает в схеме взаимного недоверия, злоумышленником может быть и законный пользователь А криптосистемы, подписывающий свои сообщения. Он, как отмечалось выше, может отказаться от того, что передал подписанное сообщение M, и утверждать, что на самом деле передал M' (скажем, изменив на знак переводимую сумму денег). Подпись при этом может оказаться неизменной.

Проиллюстрируем, какие возможности для этого возникают в схеме ECDSA в связи с обратной точкой $(-k_A G)$, порождающей то же значение параметра r , что и точка $k_A G$. Пусть $h_1 = h(M_1)$ и $h_2 = h(M_2)$, тогда одинаковые параметры (r, s) подписи формируются при выполнении равенств (6.4)

$$sk_A = (h_1 + e_A r) \bmod n,$$

$$-sk_A = (h_2 + e_A r) \bmod n.$$

Отсюда в результате суммирования получим $h_2 = -(h_1 + 2e_A r) \bmod n$. Пользователь А может вычислить это единственное значение хэш-кода ложного сообщения и попытаться подобрать для него сообщение M_2. Эта

задача, однако, относится к практически нерешаемым, так как хэш-функция должна отвечать требованиям к односторонним функциям.

Более реальной может оказаться угроза селективной подделки, если пользователь А на момент смены ключа e_A получит возможность вместо случайного ключа выбрать рассчитанное значение $e_A = (h_1 + h_2)/2r$, которое для двух различных сообщений M_1 и M_2 и заранее выбранного ключа k_A (определяющего параметр r) формирует одну и ту же подпись. Отправив сообщение M_1, он затем может утверждать, что в действительности отправил M_2, т.е. совершить подмену отправителя. Угроза может быть реализована в случае, если оба ключа генерируются не случайно, а заготавливаются и рассчитываются заранее. Это должно накладывать жесткие требования на протоколы генерации параметров подписи.

Если злоумышленником является субъект, не знающий ключа e_A, для него возникает уже две сложные задачи: наряду с подбором сообщения M под известную хэш-функцию $h(M)$ надо найти значение e_A. Так как этот ключ можно определить на основе открытого ключа $Q_A = e_A G$, достигается это решением проблемы дискретного логарифма на кривой (ECDLP). Для алгоритма ECDSA имеется строгое доказательство того, что стойкость его против атаки на секретный ключ e сводится к решению ECDLP [89, 91].

При атаке на секретный ключ e аналитик может иметь в распоряжении последовательность M_1, M_2,…, M_L подписанных сообщений, для которых в соответствии с (6.4)

$$e = (k_i A_i - B_i) \bmod n, \quad A_i = s_i r_i^{-1}, \quad B_i = e_i r_i^{-1}, \quad i = 1, 2, \ldots, L.$$

Здесь A_i, B_i – известные аналитику после вычислений параметры (при условии, что он знает алгоритм хэширования и ОСП), а k_i и e – неизвестные. Для полного раскрытия необходимо, таким образом, решить систему из L уравнений с $L + 1$ неизвестным. Стойкость подписи можно повысить благодаря шифрованию сообщений, после чего неизвестными аналитику становятся и значения h_i. В этом случае число неизвестных в нашей системе почти удваивается и становится равным $2L + 1$.

6.3.2. ECSS (Elliptic Curve Signature Scheme)

Параметры пользователя

Пользователь А:

- генерирует и хранит в секрете долговременный секретный ключ как целое число $1 < e_A < n$;
- вычисляет открытый ключ как точку кривой $Q_A = e_A G$. Открытый ключ доступен всем пользователям криптосистемы.

Формирование ЦП

Пользователь А:

1. Вычисляет хэш-код сообщения М: $h = h(\text{M})$, $h < n$.
2. Генерирует случайное целое число $0 < k_A < n$.
3. Вычисляет точку $R = k_A G = (x_1, y_1)$.
4. Преобразует двоичную запись элемента x_1 в целое число $\bar{x}_1$.
5. Вычисляет параметр $r = (h + \bar{x}_1) \bmod q$. При $r = 0$ возврат в п.2.
6. Вычисляет параметр $s = (k_A - e_A r) \bmod n$. При $s = 0$ возврат в п.2.
7. Направляет пользователю В подписанное сообщение (М, r, s), в котором $DS = (r, s)$ – электронная цифровая подпись.

Для данной схемы справедливо ключевое уравнение

$$k_A = (s + e_A r) \bmod n, \qquad (6.6)$$

в котором параметр r связывает хэш-функцию h и функцию от точки $k_A G$. Оно используется при формировании и проверке цифровой подписи.

Проверка ЦП

Пользователь B проверяет цифровую подпись пользователя A на основе тех же данных, что и в схеме ECDSA. По аналогии с этой схемой проверка заключается в вычислении на основе известных данных параметра r' и сравнении его с принятым значением r.

Произведя скалярное умножение в соответствии с левой и правой частью (6.6) на генератор G и учитывая, что $Q_A = e_A G$, для точек криптосистемы получим равенство

$$k_A G = sG + r\, e_A G = sG + rQ_A, \qquad (6.7)$$

Согласно п.3 и 4 алгоритма формирования левая часть этого равенства определяет точку $R = (x_1, y_1)$ и, следовательно, параметр $\bar{x}_1$. Правая часть равенства содержит известные получателю данные, которые он использует для вычисления параметра r'. Итак, алгоритм проверки ЦП на основе (6.7) включает следующие действия.

Пользователь B:

1. Вычисляет хэш-код полученного сообщения M как целое число $h = h(M)$, $h < n$.
2. Вычисляет точку $R' = sG + rQ_A = (x_1', y_1')$.
3. Определяет целое число $\bar{x}_1'$
4. Вычисляет параметр $r' = (\bar{x}_1' + h) \bmod n$.
5. Сравнивает вычисленное и принятое значения $r' = r$ (?). При равенстве цифровая подпись верна, в противном случае она отвергается.

В результате проверки пользователь B удостоверяется в подлинности отправителя A и целостности сообщения M.

Нетрудно видеть, что как при формировании, так и при верификации подписи в этой схеме не требуется вычислять обратный элемент поля $\boldsymbol{F}_n$. Это, естественно, повышает производительность вычислений и уменьшает время

этих процедур. При напряженном трафике подписываемых документов этот фактор является достаточно важным.

6.3.3. EC-GDSA (Elliptic Curve ElGamal Digital Signature Algorithm)

Параметры пользователя

Пользователь А:

• генерирует случайное целое число $t < n$ и вычисляет долговременный секретный ключ $e_A = t^{-1} \bmod n$;

• вычисляет открытый ключ как точку кривой $Q_A = tG = e_A^{-1}G$. Открытый ключ доступен всем пользователям криптосистемы.

Формирование ЦП

Пользователь А:

1. Вычисляет хэш-код сообщения М: $h = h(M)$, $h < n$.
2. Генерирует случайное целое число $0 < k_A < n$.
3. Вычисляет точку $R = k_A G = (x_1, y_1)$.
4. Вычисляет параметр $r = \pi(kG) \bmod n$. При $r = 0$ возврат в п.2.
5. Вычисляет параметр $s = e_A (k_A r - h) \bmod n$. При $s = 0$ возврат в п.2.
6. Направляет пользователю В подписанное сообщение (M, r, s), в котором $DS = (r, s)$ – электронная цифровая подпись.

Заметим, что п.4 предполагает, что координаты x_1, y_1 точки R каким-то способом переводится в целое число с редукцией по модулю n.

Для данной схемы ключевое уравнение имеет вид

$$k_A r = (h + st) \bmod n, \quad t = e_A^{-1} . \qquad (6.8)$$

Оно связывает параметры s и r с k_A и e_A. и используется при формировании и проверке цифровой подписи. Параметр s в п.5 определен на основе этого уравнения.

Проверка ЦП

Пользователь В проверяет цифровую подпись пользователя А на основе известных данных, как и в ECDSA. Процедура проверки заключается в вычислении параметра r' и сравнении его с принятым значением r.

Умножив (6.8) на инверсию r^{-1} первого параметра подписи и учитывая, что $Q_A = tG$, для точек криптосистемы имеем

$$k_AG = r^{-1}hG + r^{-1}s\,tG = uG + vQ_A, \qquad (6.9)$$

$$u = r^{-1}h \bmod n, \quad v = r^{-1}s \bmod n.$$

Как видим, в этом выражении по сравнении с (6.4) параметры s и r поменялись местами, а вместо e_A стоит его инверсия t. Левая часть этого равенства, как и ранее, определяет точку $R = (x_1, y_1)$ и, далее, параметр $r = \pi(\text{kG}) \bmod n$. Правая часть равенства (6.9) используется для вычисления параметра r'. Итак, проверка ЭЦП на основе (6.9) включает следующие вычисления.

Пользователь В:

1. Вычисляет хэш-код полученного сообщения М: $h = h(M)$, $h < n$.
2. Вычисляет обратный элемент $r^{-1} \bmod n$ поля $\mathbf{F}_n$.
3. Вычисляет параметры $u = r^{-1}h \bmod n$, $v = r^{-1}s \bmod n$.
4. Вычисляет точку $R' = uG + vQ_A = (x_1', y_1')$.
5. Вычисляет параметр $r' = \pi(R') \bmod n$.
6. Сравнивает вычисленное и принятое значения: $r' = r$ (?). При равенстве цифровая подпись верна, в противном случае она отвергается.

6.3.4. EC- KCDSA

Параметры пользователя

Пользователь А:

• генерирует случайный секретный ключ $e_A < n$ и вычисляет обратный элемент $t = e_A^{-1} \bmod n$;

• вычисляет открытый ключ как точку кривой $Q_A = e_A^{-1}G$. Открытый ключ доступен для всех пользователей системы;

• вычисляет хэш-код $z_A = h(C_A) < n$. данных сертификата пользователя C_A, содержащего идентификатор (имя и др.,), открытый ключ Q_A и ОСП.

Формирование ЦП

Пользователь А:

1. Генерирует случайное целое число $0 < k_A < n$.
2. Вычисляет точку $R = k_AG = (x_1, y_1)$.
3. Вычисляет хэш-код $r = h(k_AG) = h(x_1||y_1)$ (знак || означает конкатенацию или присоединение к строке x_1 строки y_1).
4. Вычисляет хэш-код $H = h(z_A \| M)$.
5. Вычисляет параметр $h = r \wedge H \bmod n$ (знак ^ означает «исключающее или» для побитовой записи r и H).
6. Вычисляет параметр $s = e_A(k_A - h) \bmod n$.
7. Направляет пользователю В подписанное сообщение (M, r, s), в котором $DS = (r, s)$ –цифровая подпись.

В п.3, таким образом, хэшируется конкатенация побитовой записи координат x_1 и y_1 точки R , причем $h < n$.

Для данной схемы ключевое уравнение имеет вид

$$k_A = (h + s\, e_A^{-1}) \bmod n, \qquad (6.10)$$

которое используется при формировании и проверке цифровой подписи. Параметр s в п.6 определен на основе этого уравнения.

Проверка ЦП

Пользователь В проверяет цифровую подпись пользователя А на основе известных данных подобно тому, как в ECDSA. Процедура проверки заключается в вычислении параметра r' и сравнении его с принятым значением r.

Согласно (6.10) с учетом того, что $Q_A = d_A^{-1} G$, для точек криптосистемы имеем

$$k_A G = hG + s\, e_A^{-1} G = hG + sQ_A, \qquad (6.11)$$

Как видим, это выражение вычисляется сравнительно просто и не требует при проверке вычисления обратного элемента (эта операция e_A^{-1} вообще вынесена за пределы протокола формирования и проверки подписи). Левая часть этого равенства, как и ранее, определяет точку $R = (x_1, y_1)$. Правая часть равенства используется для вычисления параметра r'. Итак, проверка ЦП на основе (6.11) включает следующие вычисления.

Пользователь В:

1. Вычисляет хэш-код $H = h(z_A \| M)$.
2. Вычисляет параметр $h = r \wedge H \bmod n$.
3. Вычисляет точку $R' = hG + sQ_A = (x_1', y_1')$.
4. Вычисляет хэш-код $r' = h(x_1' \| y_1')$ координат точки R'.
5. Сравнивает вычисленное r' и принятое значения r. При равенстве $r' = r$ цифровая подпись верна, в противном случае она отвергается.

6.4. Некоторые стандарты криптосистем на эллиптических кривых

6.4.1. Международный стандарт ISO/IEC 15945:2002.

Этот стандарт [95] прошел длительный этап согласования во многих странах мира (Канада, США, Великобритания, Германия, Франция, Бельгия, Италия, Польша, Корея, Япония) и после внесения модификаций и поправок утвержден Объединенным Техническим Комитетом (JTC - Joint Technical Committee) в 2002году. Он содержит три части:

1. Общие положения.
2. Цифровая подпись.
3. Распределение ключей.

Часть 1. Общие положения

В стандарте определены кривые над полями $\mathbf{F}(p)$, $\mathbf{F}(2^m)$ и $\mathbf{F}(p^m)$, $(p > 3)$, общесистемные параметры ОСП (Domain Parameters), а также процедуры их проверки и генерации конфиденциальных и общих ключей пользователей. Кроме наиболее распространенных на практике полей, как видим, допускается более общий случай – расширенное поле $\mathbf{F}(p^m)$ характеристики p.

Определение кривых и групповых законов сложения в стандарте даются традиционно (см. гл.1– 5). Поэтому мы здесь рассмотрим лишь некоторые ограничения на ОСП и ключи пользователей.

А. Проверка общесистемных параметров

При использовании кривой порядка $N_E = cn$ над полем $\mathbf{F}_p{}^m$ $(p > 3, m \geq 1)$ с генератором $G = (x_{G,}, y_G)$ порядка n пользователь должен:

1. Проверить, что p^m – степень нечетного простого числа.

2. Проверить, что параметры кривой $a, b, x_{G,}, y_G$ являются элементами поля $\mathbf{F}_p{}^m$.

3. Если эллиптическая кривая генерируется по случайному закону, проверить, что a и b были соответствующим образом получены из случайной двоичной последовательности SEED.

4. Проверить, что $(4a^3 + 27b^2) \neq 0$ в поле. $\mathbf{F}_p{}^m$.

5. Проверить, что $y_G{}^2 = x_G{}^3 + ax_G + b$ в поле $\mathbf{F}_p{}^m$.

6. Проверить, что n – простое число и $n > 4\sqrt{p^m}$.

7. Проверить, что $nG = O$.

8. Вычислить $c' = \left\lfloor (\sqrt{p^m} + 1)^2 / n \right\rfloor$ и проверить, что $c = c'$ (это верно лишь при выполнении п.6).

9. Проверить свойства кривой с целью исключения известных криптографически слабых кривых:

- проверить выполнение MOV – условия, исключающего суперсингулярные кривые: $p^{mk} \neq 1 \bmod n$, $k = 1, 2, 3, \ldots, B$, $B \geq 20$.
- проверить, что кривая не является аномальной, т.е. $N_E \neq p^m$.

Если хотя бы одна из проверок дала отрицательный результат, соответствующая кривая отбраковывается и не может применяться для криптографических задач. Отметим, что условие 6 является довольно слабым для порядка n генератора криптосистемы. В реальных криптосистемах n имеет приблизительно тот же порядок, что и поле (т.е. кофактор c имеет максимальную разрядность лишь 2 бита). Стойкость к MOV-атаке в этом стандарте задается невысокой по сравнению с другими стандартами нижней границей значений $B = 20$.

При использовании кривой порядка $N_E = cn$ над полем $\mathbf{F}_2{}^m$ ($m > 1$) с генератором $G = (x_G, y_G)$ порядка n пользователь должен:

1. Проверить, что $q = 2^m$ для заданного m.

2. Проверить, что a, b, x_G , y_G являются двоичными векторами длины m бит.

3. Если эллиптическая кривая генерируется по случайному закону, проверить, что a и b были соответствующим образом получены из случайной двоичной последовательности SEED.

4. Проверить, что $b \neq 0$.

5. Проверить, что $y_G^2 + x_G y_G = x_G^3 + a x_G^2 + b$ в $\mathbf{F}_2{}^m$..

6. Проверить, что n – простое число, и $n > 2^{160}$ и $n > 4\sqrt{2^m}$.

7. Проверить, что $nG = O$.

8. Вычислить $c' = \lfloor (\sqrt{2^m} + 1)^2 / n \rfloor$ и проверить, что $c = c'$.

9. Проверить свойства кривой с целью исключения известных криптографически слабых кривых:

- проверить выполнение MOV – условия, исключающего суперсингулярные кривые и слабые несуперсингулярные кривые: $2^{mk} \neq 1 \bmod n$, $k = 1, 2, 3, \ldots, B$, $B \geq 20$.

- проверить, что кривая не является аномальной в расширении $\mathbf{F}_2{}^m$ и число точек $N_E \neq 2^m$.

Если любая из проверок в этом перечне дала негативный результат, то ОСП должны рассматриваться как недействительные.

Отметим, что аномальные кривые над полем $\mathbf{F}_2{}^m$ определяются шире, чем кривые с порядком 2^m. К ним часто относят все кривые с коэффициентами $a, b \in \mathrm{F}_2$, $b = 1$ (кривые Коблица). Строго говоря, кривые Коблица аномальны лишь над полем $\mathbf{F}_2$, но впоследствии эти кривые над расширением $\mathbf{F}_2{}^m$ стали также называть аномальными [91,92]. Среди них имеются достаточно стойкие кривые, рекомендованные, например, в стандарте FIPS 186-2. Следует также заметить, что стандарт ISO не ограничивает выбор расширения m поля $\mathbf{F}_2{}^m$, т.е.

допускаются к применению составные значения m. Это может привести к угрозе взлома криптосистемы методом спуска Вейля [93, 94].

В. Генерация конфиденциальных ключей и проверка открытых ключей

Опишем процедуру генерации ключевой пары и проверки открытого ключа при заданных ОСП.

Конфиденциальный и открытый ключи при заданных ОСП формируются следующим образом:

1. Выбрать случайное или псевдослучайное целое число e в интервале $[1, n-1]$. Число e должно быть защищено от неавторизованных искажений и быть непредсказуемым.

2. Вычислить точку $Q = (x_Q, y_Q) = eG$.

3. Ключевой парой является (Q, e), где Q используется как открытый ключ, а e – конфиденциальный ключ.

Согласно п.1, возможность предварительного расчета ключа e и его подтасовки должна быть исключена протоколом.

Проверка открытого ключа при заданных ОСП включает следующие действия:

1. Проверить, что Q не является точкой на бесконечности О.

2. Проверить, что координаты x_Q и y_Q точки Q являются элементами поля $\mathbf{F}_q$.

3. Если $q = p^m$ $(p > 3)$, проверить, что $y_Q^2 = x_Q^3 + ax_Q + b$ в $\mathbf{F}_p{}^m$.

4. Если $q = 2^m$, проверить, что $y_Q^2 + x_Q y_Q = x_Q^3 + ax_Q^2 + b$ в $\mathbf{F}_2{}^m$.

5. Проверить, что $nQ = O$.

Если кандидат открытого ключа не прошел хотя бы одну проверку, он считается недействительным, как и все криптографические операции, в которые он вовлекался. Кроме того, использование недействительного открытого ключа может привести к компрометации (раскрытию) секретного ключа пользователя. Поэтому перед использованием открытого ключа пользователь должен убедиться в том, что он является действительным.

В части 1 стандарта имеются обширные приложения с основами теории конечных полей и эллиптических кривых, их описание и арифметика как в аффинных координатах, так и в проективных. Рассмотрены все три типа полей: $\mathbf{F}_p$, $\mathbf{F}_{2^m}$ и $\mathbf{F}_{p^m}$ ($p > 3$). Приведены некоторые сведения по методам вычисления скалярного умножения kG (в стандарте – integer multiplication on an EC), решения ECDLP, MOV- условие, методы компрессии точки кривой.

Часть 2. Цифровая подпись

В этой части рассмотрена общая модель цифровой подписи с генерацией ОСП и параметров пользователей, процедурами формирования и проверки ЦП. Далее даны три алгоритма ЦП: EC-GDSA (предложен Германией по схеме Эль-Гамаля с модификацией), ECDSA, EC-KCDSA (предложен Кореей). Эти алгоритмы были рассмотрены в пп.6.3.1, 6.3.3 и 6.3.4.

В Приложении А дана сравнительная характеристика алгоритмов ECDSA и EC-KCDSA. В таблице 6.1 приведены для сравнения основные параметры подписи согласно этим алгоритмам, а в таблице 6.2 – число операций при генерации и проверке подписи.

Таблица 6.1.Сравнительная характеристика алгоритмов ECDSA и EC-KCDSA

	EC-DSA	EC-KCDSA
Параметры безопасности	n, $h(\,)$	
Условие на n	$n \geq 2^{\lvert h(\,)\rvert}$	$n > 2^{\lvert h(\,)\rvert - 1}$
Секретный ключ	$e_A \in [1, \cdots, n-1]$	

Вычисление открытого ключа	$Q_A = e_A G$	$Q_A = (e_A^{-1} \bmod n)G$
Формирование ЦП	$k \in [1, \cdots, n-1]$ $r = \pi(kG) \bmod n$ $s = (e_A r + h(M))k^{-1} \bmod n$	$k \in [1, \cdots, n-1]$ $r = h(kG)$ $s = e_A(k - r \text{ XOR } h(z_A \| M)) \bmod n$
Размер подписи	$0 < r < n, 0 < s < n$	$0 < s < n, \quad 0 \le r < 2^{\|h()\|}$
Проверка ЦП	$u_1 = s'^{-1} h(M') \bmod n$ $u_2 = s'^{-1} r' \bmod n$ $\pi(u_1 G + u_2 P_A) \bmod n =$ $= r'$?	$h' = r' \text{ XOR } h(z_A \| M') \bmod n$ $h(s'G + h'G) = r'$?

Таблица 6.2. Число операций в алгоритмах ECDSA и EC-KCDSA при формировании и проверке ЦП

Процедура	Операция	ECDSA	EC-KCDSA
Формирование ЦП	$h(\,)$	1	2
	$\pi(\,)$	1	0
	$k^{-1} \bmod n$	1	0
	Умножение в $\mathbf{F}_n$	2	1
	Сложение в $\mathbf{F}_n$	1	1
	Скалярное	1	1
Проверка ЦП	$h(\,)$	1	2
	$\pi(\,)$	1	0
	$s^{-1} \bmod n$	1	0
	Умножение в $\mathbf{F}_n$	2	0
	Скалярное	2	2
	Сложение точек	1	1

Как следует из таблицы 6.2, ECDSA конвертирует точку кривой в целое число с помощью функции $\pi(\cdot)$, тогда как EC-KCDSA взамен $\pi(\cdot)$ использует хэш-функцию $h(\cdot)$ (см 6.3.4). Вычислительная сложность хэш-функции выше, чем конверсии точки. Однако, доля вычислений хэш-функции невелика в общем объеме вычислений. Более того, применение хэш-функций делает модель подписи EC-KCDSA доказуемо стойкой, в то время как функция

конверсии этого не обеспечивает. Наряду с этим, EC-KCDSA не требует вычисления инверсий элементов поля $\mathbf{F}_n$ ни при формировании, ни при проверке подписи. В общей схеме подписи наиболее трудоемкие вычисления занимают сравнительно немного времени, если используются обычные компьютеры, однако для таких средств, как смарт-карты, они могут оказаться весьма сложными. Таким образом, в первом случае EC-KCDSA может быть более эффективной с вычислительной точки зрения, чем ECDSA.

В Приложении В части 2 стандарта приводятся численные примеры с конкретными значениями параметров криптосистемы и процедурами генерации и верификации подписи.

Часть 3. Распределение ключей

После введения общих определений и принципов разделения секрета в стандарте приводится ряд известных схем и механизмов формирования общего секретного ключа пользователей А и В для использования в симметричной системе шифрования. Рассмотрены неинтерактивные и интерактивные методы ключевого обмена.

А. Неинтерактивное распределение ключей Диффи-Хэллмана

В начале п.6.1 мы привели простейший протокол ключевого обмена Диффи-Хэллмана на основе долговременных секретных и открытых ключей пользователей А и В ($S_{AB} = e_A Q_B = e_B Q_A$). Поскольку для вычисления общего ключа не требуется пересылать партнерам никаких сообщений, такой обмен называется неинтерактивным.

Если аутентификация открытых ключей не проведена, существует очевидная угроза подмены открытого ключа Q. Злоумышленник может попытаться, например, заменить открытый ключ какого-то пользователя точкой Q' малого порядка. Если порядок кривой $N_E = cn$ с малым кофактором c, то существуют точки порядка c (и делителей c), такие, что $cQ' = O$. Ясно, что при этом точка $S_{AB}' = e_A Q_B'$ тоже имеет порядок c, и все возможные

секретные ключи легко определяются перебором c различных x-координат точки S_{AB}'. Простая защита от этой угрозы может быть обеспечена дополнительным скалярным умножением точки S_{AB} на $c = N_E/n$. В этом случае при подтасовке ключа Q' малого порядка пользователь получает точку на бесконечности $S_{AB}' = ce_AQ_B' = O$ и ключевой обмен не состоится.

По сути, умножением на c пользователи верифицируют действительность открытого ключа. Очевидно, что если $cQ' = O$, то $nQ' \neq O$, т.е. п.5 проверки открытого ключа (см. Часть 1 стандарта) не выполняется.

Наряду с этим по взаимной договоренности пользователи могут ввести еще один дополнительный кофактор v, совместное использование которого обеспечивает их взаимную аутентификацию. Неинтерактивный протокол, описанный в стандарте, рекомендует следующие действия:

1. Пользователь А с помощью своего секретного ключа e_A и открытого ключа Q_B пользователя В вычисляет точку $S_{AB} = cv\, e_AQ_B$.
2. Пользователь В с помощью своего секретного ключа e_B и открытого ключа Q_A пользователя А вычисляет точку $S_{AB} = cv\, e_BQ_A$.
3. Если $S_{AB} \neq O$, в качестве ключа симметричного шифрования пользователи могут принять x-координату точки S_{AB}.

Этот протокол, как видим, не требует предварительного сеанса связи между пользователями.

B. Интерактивное распределение ключей

Здесь рассматриваются несколько протоколов: Диффи-Хэллмана, Эль-Гамаля, MQV и др.

B.1. Распределение ключей по схеме Эль-Гамаля

Эта схема обеспечивает ключевой обмен с помощью одной пересылки сообщения от пользователя А к пользователю В. Предполагается, что пользователь А располагает авторизованной копией открытого ключа Q_B пользователя В. Для разделения секрета используется разовый секретный ключ k_A пользователя А и долговременный ключ e_B пользователя В. С

включением дополнительных кофакторов c и v протокол регламентирует следующие действия:

1. Пользователь A генерирует случайное число $k_A \in [1, \cdots, n-1]$, вычисляет точку $R = k_A G$ и отправляет ее пользователю B.
2. Пользователь B вычисляет точку $S_{AB} = cve_B R$.
3. Пользователь A вычисляет точку $S_{AB} = cvk_A Q_B$.
4. Если $S_{AB} \neq O$, в качестве ключа симметричного шифрования пользователи могут принять x-координату точки S_{AB}.

Как отсюда видно, реализация протокола сводится к одной передаче данных от A к B.

B2. Распределение ключей по схеме Диффи-Хэллмана

Это классическая схема с обменом данными между A и B (т.е. двумя передачами сообщений). Разделение секрета в схеме достигается с помощью разовых секретных ключей k_A и k_B пользователей. С учетом дополнительных кофакторов c и v выполняются следующие действия:

1 Пользователь A генерирует случайное число $k_A \in [1, \cdots, n-1]$, вычисляет точку $R_A = k_A G$ и отправляет ее пользователю B.

2 Пользователь B генерирует случайное число $k_B \in [1, \cdots, n-1]$, вычисляет точку $R_B = k_B G$ и отправляет ее пользователю A.

4. Пользователь A вычисляет точку $S_{AB} = cvk_A R_B$.
5. Пользователь B вычисляет точку $S_{AB} = cvk_B R_A$.
6. Если $S_{AB} \neq O$, в качестве ключа симметричного шифрования пользователи могут принять x-координату точки S_{AB}.

В этой схеме долговременные ключи пользователей не используются, что оставляет открытой проблему аутентификации ключей пользователей.

B3. Распределение ключей по схеме Диффи-Хэллмана с двумя ключевыми парами

Здесь разделение секрета достигается с помощью разовых и долговременных ключей пользователей. При наличии открытых ключей друг друга и с учетом кофакторов c и v пользователи выполняют следующие действия:

1. Пользователь A генерирует случайное число $k_A \in [1, \cdots, n-1]$, вычисляет точку $R_A = k_A G$ и отправляет ее пользователю B.
2. Пользователь B генерирует случайное число $k_B \in [1, \cdots, n-1]$, вычисляет точку $R_B = k_B G$ и отправляет ее пользователю A.
3. Пользователь A вычисляет точки $S_{AB} = cvk_A R_B$, $Q_{AB} = cve_A Q_B$.
4. Пользователь B вычисляет точки $S_{AB} = cvk_B R_A$, $Q_{AB} = cve_B Q_A$
5. Если $S_{AB} \neq O$, $Q_{AB} \neq O$, в качестве ключа симметричного шифрования пользователи могут принять конкатенацию x-координат точек S_{AB} и Q_{AB}.

Эта схема, в отличие от предыдущей, обеспечивает взаимную аутентификацию ключей. Она, однако, не аутентифицирует самих пользователей, что можно обеспечить с помощью механизма цифровой подписи.

B4. Распределение ключей по схеме Диффи-Хэллмана с двумя подписями

Эта схема отличается наличием подписей в передаваемых данных и тремя сеансами связи. Для осуществления протокола пользователи имеют следующую информацию:

1. Каждый пользователь имеет ключевую пару (секретный и открытый ключи) подписи с взаимно согласованным алгоритмом формирования S_u и проверки V_u подписи.
2. Каждый пользователь имеет аутентифицированную копию открытого ключа проверки подписи своего партнера.
3. Каждый пользователь имеет все необходимые параметры процедуры формирования и проверки подписи.
4. Известна криптографическая функция $f_Z(X)$ с ключом Z.

При наличии открытых ключей друг друга и с учетом кофакторов c и v пользователи выполняют следующие действия:

1. Пользователь А генерирует случайное число $k_A \in [1, \cdots, n-1]$, вычисляет точку $R_A = k_A G$ и отправляет ее пользователю В.

2. Пользователь В генерирует случайное число $k_B \in [1, \cdots, n-1]$, и вычисляет точки $R_B = k_B G$ и $K_{AB} = cvk_B R_A$.

Далее он вычисляет подпись $KT_B = S_B(DB_1) \parallel f_Z(DB_1)$,
где

$$DB_1 = k_B G \parallel k_A G \parallel A, \quad Z = K_{AB},$$

и отправляет ее пользователю А (*KT – Key Token Construction*).

3. Пользователь А проверяет подпись S_B с помощью открытого ключа подписи пользователя В, проверяет правильность идентификатора А и значения $k_A G$. Если проверки прошли успешно, он определяет из блока DB_1 точку $R_B = k_B G$ и вычисляет точку $K_{AB} = cvk_A R_B$. В качестве ключа симметричного шифрования для связи с А он может использовать x-координату точки K_{AB}

Далее он вычисляет подпись $KT_A = S_A(DB_2) \parallel f_Z(DB_2)$,
где

$$DB_2 = k_A G \parallel k_B G \parallel B, \quad Z = K_{AB},$$

и отправляет ее пользователю В.

4. Пользователь В проверяет подпись S_A с помощью открытого ключа подписи пользователя А, проверяет правильность идентификатора В и значения $k_B G$. Если проверки прошли успешно, в качестве ключа симметричного шифрования для связи с А он использует x-координату точки K_{AB}

Отметим, что в подписи КТ функция $S(\cdot)$ предназначена для аутентификации пользователей, а функция $f_Z(\cdot)$ – для аутентификации ключа K_{AB}.

Кроме рассмотренных в стандарте приведены другие, не включенные в данный стандарт, протоколы. В частности, дано описание протокола MQV (см. п.6.1) и ряд других [95].

6.4.2. Национальный стандарт Украины ДСТУ ISO/IEC 14888-2014

Это международный стандарт [83] 2008 года, принятый как национальный стандарт Украины в декабре 2014 года [84,85]. Он включает 3 части: 1. Общие положения; 2. Механизмы, основанные на факторизации целюх чисел; 3. Механизмы, основанные на дискретном логарифме. Мы здесь приведем лишь некоторые сведения из 3-й части этого стандарта, в которой рассмотрены алгоритмы цифровой подписи (DSA – Digital Signature Algorithm).

Всего в стандарте приведено 12 алгоритмов DS, из них 10 алгоритмов онованы на сертификатах, и 2 алгоритма – на идентичности. В частности, приведены следующие алгоритмы DS, использующие арифметику эллиптических кривых:

- EC-DSA;
- EC-KCDSA (Корейский алгоритм цифровой подписи);
- EC-GDSA (Немецкий алгоритм цифровой подписи);
- EC-RDSA (Российский алгоритм цифровой подписи);
- EC-SDSA (Алгоритм цифровой подписи Шнорра);
- EC-FSDSA (Полный алгоритм цифровой подписи Шнорра – Elliptic Curve Full Schnorr Digital Signature Algorithm).

В качестве стандарта хэширования рекомендуется один из стандартов SHA-1, SHA-224, SHA-256, SHA-384 та SHA-512, описанных в ДСТУ ISO/IEC 10118-3:2005.

Ряд принятых в стандарте обозначений отличаются от принятых в данной книге, в частности: $(r,s)\to(R,S)$, $n\to q$, $h\to H$, $k\to K$, $R\to П$, $kP\to$[K]P, $e\to$ X, $Q\to$ Y. Сразу заметим, что эти обозначения очень неудачны, так как разные математические объекты обозначаются одинаково (например, X – секретный ключ, число, Y – открытый ключ, точка).

Рассмотрим описание EC-DSA в стандарте.

$H = h(M)$ – урезанный хэш-код сообщения M, преобразованного в целое число согласно правилу, приведенному Приложении В.

Функция-лазейка (первый параметр DS) определяется формулой

$$R = FE2I(\Pi_x) \bmod q.$$

Правило преобразования *FE2I*, приведено в Приложении В.

Уравнение подписи имеет вид

$$SK - RX - H \equiv 0 \pmod{q}.$$

1. Формирование ключа подписи и ключа проверки

Ключ подписи является секретным целым числом X, таким, что $0 < X < q$. Соответствующий ключ проверки вычисляется как точка

$$Y = [X]G.$$

Секретный ключ отправителя X и открытый ключ проверки Y фиксируются на определенный период времени. Ключ подписи X *должен* держаться в секрете.

2. Процесс подписывания

2.1. Генерирование рандомизатора

Подписывающий вычисляет случайное или псевдослучайное число K такое, что $0 < K < q$.

2.2. Формирование предподписи

Входом на этом етапе является рандомизатор K; Подписывающий вычисляет точку

$\Pi_K = [K]G$..

2.3. Подготовка сообщения к подписи

Принимаем M_1 - пустое, а M_2 – сообщение, т.е. $M_2 = M$.

2.4. Вычисление лазейки (первого параметра DS)

$R = FE2I(\Pi_x) \bmod q$.

2.5. Вычисление хеш-кода

$H = h(M_2)$.

2.6. Вычисление второй части подписи

$S = K^{-1}(XR + H) \bmod q$.

Подписью является пара (R, S). Если $R = 0$ или $S = 0$, возврат в 2.1.

2.7. Построение дополнения

Дополнение есть конкатенация пары (R, S) и необязательного текстового поля : $((R, S), text)$.

2.8. Построение подисанного сообщения

Подисанное сообщение есть конкатенация сообщения M и дополнения:

$M||((R, S), text)$

3. Процесс проверки

Проверяющий получает элементы данных, необходимых для проверки.

3.1. Выделение лазейки

Проверяющий выделяет лазейку R и вторую часть подписи S с дополнением. Он сначала убеждается что $0 < R < q$ и $0 < S < q$. Если хотя бы одно из условий не выполняется, подпись неверна**.**

3.2. Подготовка сообщения к проверке

Проверяющий выделяет M из подписанного сообщения и разбивает его на 2 части M_1 и M_2. M_1 – пустое $M_2 = M$.

3.3. Вычисление хэш-кода и вектора Т

Проверяющий вычисляет хэш-код $h(M_2) = H$ и определяет вектор $T = (T_1, T_2)$, где $T_1 = -R$, $T_2 = -H$.

3.4. Перерасчет предподписи

Входом на этом этапе являются генератор G, ключ проверки Y, вектор $T = (T_1, T_2) = (-R, -H)$. Проверяющий вычисляет пересчитанное значение точки $\overline{\Pi}$ предподписи с помошью формулы

$$\overline{\Pi} = [-S^{-1}T_1 \bmod q]Y + [-S^{-1}T_2 \bmod q]G.$$

3.5. Перерасчет лазейки

Проверяющий вычисляет функцию-лазейку $\overline{R}$ аналогично п.2.4. Входом является величина $\overline{\Pi}$ из п.3.4. Виходом является перерассчитанная лазейка $\overline{R}$.

3.6. Проверка подписи

Проверяющий сравнивает лазейку $\overline{R}$ п.3.5 с выделенной версией лазейки R из п. 3.1. Если $\overline{R} = R$, підпись верна.

В стандарте имеются численные примеры выполнения всех алгоритмов формирования и проверки подписи и сравнительный анализ быстродействия алгоритмов при их программной реализации. К недостаткам стандарта, кроме неудачных обозначений, следует отнести отсутствие рекомендуемых эллиптических кривых и их общесистемных параметров.

6.4.3. FIPS-186-2-2000

Как уже отмечалось, этот национальный стандарт DSS (Digital Signature Standard) правительства США принят NIST (National Institute of Standard and Technology) в 2000 году [78]. Он сменил предыдущий стандарт FIPS-186 [73], действующий с 1994 года, и рекомендует алгоритм цифровой подписи на эллиптической кривой (ECDSA) взамен DSA, построенного на арифметике простого поля Галуа. Сам алгоритм в стандарте не приведен со ссылкой на описание алгоритма в стандарте X9.62 ANSI [77].

FIPS-186-2 рекомендует к использованию 10 полей и 15 эллиптических кривых, 5 из которых определены над простыми полями $\mathbf{F}_p$ и 10 – над расширенными полями $\mathbf{F}_{2^m}$. Значения модулей простых полей:

$$p_{192} = 2^{192} - 2^{64} - 1,$$

$$p_{224} = 2^{224} - 2^{96} + 1,$$

$$p_{256} = 2^{256} - 2^{244} + 2^{192} + 2^{96} - 1,$$

$$p_{384} = 2^{384} - 2^{128} - 2^{96} + 2^{32} - 1,$$

$$p_{521} = 2^{521} - 1.$$

Расширения двоичного поля равны $m = 163,\ 233, 283, 409$ и 571.

Длина в битах порядка полей определена таким образом, чтобы она приблизительно вдвое превосходила размер ключа известных стандартов симметричных шифров. Это согласует по стойкости задачу обмена ключами по схеме Диффи-Хэллмана с помощью ECC со стандартами блочных симметричных шифров SKIPJACK, DES и AES, таблица 6.3. Считается, что криптостойкость ECC (или сложность решения ECDLP ρ- методом Полларда, оцениваемая величиной $O(\sqrt{\pi n}/2) \approx 2^{m/2}$), близка при этом сложности силовой атаки на ключ симметричной криптосистемы.

Выбор значений m расширения поля $\mathbf{F}_2$ в стандарте обусловлен существованием при последовательном расширении m кривых Коблица «почти простого» порядка, равного $2n$ или $4n$. Отметим, что все

рекомендованные в стандарте кривые имеют почти простой порядок с минимальным кофактором, т.е. $N_E = n$, или $N_E = 2n$, или $N_E = 4n$.

Из кривых над простыми полями $\mathbf{F}_p$ NIST рекомендует кривую

$$y^2 = x^3 - 3x + b \bmod p.$$

Таблица 6.3. Длина ключа симметричных криптосистем, эквивалентных по стойкости сложности решения ECDLP ρ- методом Полларда

Алгоритм	Длина ключа	Длина модуля p поля $\mathbf{F}_p$ (бит)	Расширение m поля $\mathbf{F}_2$ (бит)
SKIPJACK	80	192	163
Тройной DES	112	224	233
Малый AES	128	256	283
Средний AES	192	384	409
Большой AES	256	521	571

Коэффициент a = – 3 позволяет наиболее быстро вычислять параметр $\nu = 3(x_1^2 - 1)/2y_1$ при удвоении точки. Порядок кривой $N_E = cn$ с минимальным кофактором $c = 1$ и простым n определяется после этого лишь коэффициентом b и порядком поля p. Их шестнадцатеричные значения для пяти рекомендуемых кривых приведены в таблице 6.4. Наряду с этими данными в стандарте приведены координаты одного из возможных генераторов – точки $G = (x_G, y_G)$ порядка n (в стандарте эти координаты обозначены как G_x , G_y). Заметим также, что порядок генератора G криптосистемы в стандарте обозначен как r вместо n.

Из 10 кривых над расширенными полями $\mathbf{F}_{2^m}$ 5 несуперсингулярных кривых $y^2 + xy = x^3 + ax^2 + b$ с коэффициентами $a = 1$ и $b \in \mathbf{F}_{2^m}$, $b \neq 0, 1$,

имеют порядок $2n$ (т.е. минимальный кофактор $c = 2$). Они обозначены как *B-m* с расширениями *m,* приведенными в таблице 6.3. В таблице 6.5 даны значения коэффициентов a и b этих кривых, порядка генератора n и неприводимые полиномы $P(x)$ с минимальным весом (трином или пентаном) для операций в полиномиальном базисе поля. Хотя значения b и n , записанные в шестнадцатеричной системе, имеют одинаковый вид, следует помнить о принципиальном отличии их математического смысла: n представляет собой целое число, тогда как b является элементом поля $\mathbf{F}_2{}^m$ (двоичным вектором, легко получаемым из шестнадцатеричной последовательности).

Еще 5 несуперсингулярных кривых вида $y^2 + xy = x^3 + ax^2 + 1$ над полем $\mathbf{F}_2{}^m$ с коэффициентами $a = 1$ или $a = 0$ имеют порядок $2n$ или $4n$ соответственно (т.е. кофактор $c = 2$ или $c = 4$). Эти кривые называют кривыми Коблица, в стандарте они обозначены как *K-m.* Значения степеней m расширения поля $\mathbf{F}_2$ выбраны именно для этих кривых так, что одна из пары кривых кручения (с коэффициентом $a = 1$ или $a = 0$) имеет почти простой порядок $2n$ или $4n$. Порядки эих кривых приведены в примере 4.9 главы 4 [29]. Отметим, что в стандарте FIPS-186-2, кроме порядков кривых, даны возможные значения координат точек G порядка n как в полиномиальном, так и в нормальном базисе.

Кривые Коблица – наиболее технологичные кривые над полем характеристики 2. Они обеспечивают наивысшую производительность вычислений в поле $\mathbf{F}_2{}^m$ [96, 97]. В то же время они относятся к классу аномальных кривых, что снижает их стойкость в $\sqrt{m}$ раз по сравнению с кривыми *B-m* с произвольным значением коэффициента b и таким же порядком [98]. Для нижней границы порядка криптосистемы $n > 2^{160}$ со сложностью случайного поиска коллизий $\sqrt{n} = 2^{80}$ потеря сложности $\sqrt{m} \approx 2^{3.7}$ составляет сравнительно небольшую величину, не превышающую 4-х бит ($\sqrt{n/m} \approx 2^{76}$). Однако это уже ниже рекомендованной стандартом сложности.

Достаточно большой диапазон размеров поля и порядков криптосистем позволяет реализовать системы с различной степенью безопасности, работающие совместно с симметричными блочными шифрами с разной

длиной ключа. Вместе с тем следует помнить, что с ростом размера модуля падает скорость криптопреобразований и, соответственно, растет время шифрования и обмена ключами, формирования и проверки подписи. Безопасность и эффективность являются противоречивыми требованиями к криптосистеме, при разработке которой следует искать компромиссное решение.

Таблица 6.4. Параметры кривых над полем $\mathbf{F}_p$, рекомендуемых FIPS-186-2

P-192: $p = 2^{192} - 2^{64} - 1$, $a = -3$, $c = 1$, b = 0x 64210519 E59C80E7 0FA7E9AB 72243049 FEB8DEEC C146B9B1 n = 0x FFFFFFFF FFFFFFFF FFFFFFFF 99DEF836 146BC9B1 B4D22831
P-224: $p = 2^{224} - 2^{96} + 1$, $a = -3$, $c = 1$, b= 0x B4050A86 0C04B3AB F5413256 5044D0D7 D7BFD8BA 270B3943 2355FFB4 n = 0x FFFFFFFF FFFFFFFF FFFFFFFF FFFF16A2 E0B8F03E 13DD2945 5C5C2A3D
P-256: $p = 2^{256} - 2^{244} + 2^{192} + 2^{96} - 1$, $a = -3$, $c = 1$, b=0x 5AC635D8 AA3A93E7 B3EBBD55 769886BC 651D06B0 CC53B0F6 3BCE3C3E 27D2604B n = 0x FFFFFFFF 00000000 FFFFFFFF FFFFFFFF BCE6FAAD A7179E84 F3B9CAC2 FC632551
P-384: $p = 2^{384} - 2^{128} - 2^{96} + 2^{32} - 1$, $a = -3$, $c = 1$, b= 0x B3312FA7 E23EE7E4 988E056B E3F82D19 181D9C6E FE814112 0314088F5013875A C656398D 8A2ED19D 2A85C8ED D3EC2AEF

n = 0x FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF C7634D81 F4372DDF 581A0DB2 4880A77A ECEC196A CCC52973

Таблица 6.5. Параметры кривых над полем $\mathbf{F}_2{}^m$, рекомендуемых FIPS-186-2

B-163: $a = 1,\ c = 2,\ P(x) = x^{163} + x^7 + x^6 + x^3 + 1,$ b = 0x 00000002 0A601907 B8C953CA 1481EB10 512F7874 4A3205FD n = 0x 00000004 00000000 00000000 000292FE 77E70C12 A4234C33
B-233: $a = 1,\ c = 2,\ P(x) = x^{233} + x^{74} + 1,$ b = 0x 00000066 647EDE6C 332C7F8C 0923BB58 213B333B 20E9CE42 81FE115F 7D8F90AD n = 0x 00000100 00000000 00000000 00000000 0013E974 E72F8A69 22031D26 03CFE0D7
B-283: $a = 1,\ c = 2,\ P(x) = x^{283} + x^{12} + x^7 + x^5 + 1,$ b = 0x 027B680A C888596D A5A4AF8A 19A0303F CA97FD76 45309FA2 A581485A F6263E31 3B79A2F5 n = 0x 03FFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFEF90 399660FC 938A9016 5B042A7C EFADB307
B-409: $a = 1,\ c = 2,\ P(x) = x^{409} + x^{87} + 1,$ b = 0x 0021A5C2 C8EE9FEB 5C4B9A75 3B7B476B 7FD6422E F1F3DD67 4761FA99 D6AC27C8 A9A197B2 72822F6C D57A55AA 4F50AE31 7B13545F n = 0x 01000000 00000000 00000000 00000000 00000000 00000000 000001E2 AAD6A612 F33307BE 5FA47C3C 9E052F83 8164CD37 D9A21173

B- 571: $a = 1$, c = 2, $P(x) = x^{571} + x^{10} + x^5 + x^2 + 1$,

b = 0x 02F40E7E 2221F295 DE297117 B7F3D62F 5C6A97FF CB8CEFF1 CD6BA8CE 4A9A18AD 84FFABBD 8EFA5933 2BE7AD67 56A66E29 4AFD185A 78FF12AA 520E4DE7 39BACA0C 7FFEFF7F 2955727A

n =0x 03FFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF E661CE18 FF559873 08059B18 6823851E C7DD9CA1 161DE93D 5174D66E 8382E9BB 2FE84E47

В работах [96-97] проведен содержательный анализ эффективности вычислений в поле и группе точек кривых, рекомендованных в FIPS-186-2, при программной реализации алгоритмов. Рассмотрены наиболее быстрые алгоритмы вычислений на 2000-й год. Интегральные характеристики по эффективности вычислений дает вычисление скалярного умножения kP, где P рассматривается либо как фиксированная точка, либо как заранее неизвестная. В первом случае целесообразно запоминать предвычисления точек удвоения P, $2P$, $4P$, …, $2^{m-1}P$, после чего остается лишь операция сложения точек из этого базиса, соответствующих позициям 1 в двоичной записи числа k. При заранее неизвестной точке P вычисления производятся методом последовательных сложений-удвоений, причем необходимо выполнить ровно $(m - 1)$ удвоений и в среднем $m/2$ сложений (если использовать и вычитания, то их число сокращается до $m/3$). В целом по данным работы [96], при одинаковых алгоритмах и программной реализации арифметика эллиптических кривых над простым полем F_p выполняется в 2 - 3 раза быстрее, чем в поле $F_2{}^m$.

В последней версии данного стандарта FIPS-186-4, утвержденной в 2013 году, рекомендуются те же самые эллиптические кривые, что и в 2000 году. Это, по мнению лидеров эллиптической криптографии Коблица и Менезеса [15], связано с политикой АНБ США и его стремлением направить основные финансовые ресурсы потребителей и исследования криптографов на стандартизацию алгоритмов постквантовой криптографии (PQC). Однако

весьма отдаленная перспектива PQC (с оценкой не менее 15 – 20 лет) вызывает острую критику этой политики со стороны ученых.

6.4.4. ГОСТ Р 34.10-2001

Этот ГОСТ принят в Российской Федерации в октябре 2001 года и вступил в силу с июля 2002 года [80]. Он регламентирует процедуры формирования и проверки электронной цифровой подписи (ЭЦП) на основе арифметики эллиптических кривых, определенных над простым полем Галуа с модулем 256 бит. Соответствующий алгоритм EC-RDSA вошел через 7 лет в международный стандарт ISO/IEC 14888-1,2,3:2008.

Обозначения, принятые в этом ГОСТе, адаптированы к обозначениям предыдущего ГОСТ Р34.10-94, принятого в 1994 году. В частности, порядок генератора криптосистемы в нем обозначен как простое число q (вместо принятого здесь обозначения n, тогда как q у нас обозначает порядок поля Галуа). Сам термин „электронная цифровая подпись" характерен именно для этого ГОСТа и не встречается в других стандартах. В данной книге мы будем сохранять раз принятые обозначения, чтобы не вносить путаницу. Различия в обозначениях будут подытожены в таблице 6.6.

Общесистемные параметры в ГОСТе определены как в начале данной главы, с ограничительными условиями:

$\mathbf{F}_p$ – простое поле Галуа порядка $p > 2^{255}$;

$\mathbf{E}$ – эллиптическая кривая $y^2 = x^3 + ax + b$ над полем $\mathbf{F}_p$ с числом точек $N_E = cn$ и дискриминантом $(4a^3 + 27b^2) \neq 0$;

G – генератор криптосистемы простого порядка n, $2^{254} < n < 2^{256}$;

h(M) – хэш-функция, отображающая двоичные сообщения M произвольной длины в двоичный вектор длины 256 бит. Хэш-функция определена в ГОСТ Р 34.11-94 [81].

Кроме этих ограничений общесистемные параметры криптосистемы проверяются на стойкость с помощью трех известных тестов:

1. MOV – атака. Порядок n криптосистемы не должен делить порядок мультипликативной группы расширенного поля Галуа $p^k - 1$ с расширением $k = 2, 3, \ldots, B$, $B \geq 31$.

2. Тест на аномальность кривой. Должно выполняться условие $N_E \neq p$. В противном случае кривая аномальна и не приемлема для криптографии.

3. Тест на инвариант (и суперсингулярность). J-инвариант кривой $j(E) \neq 0 \bmod 1728$. Это, в частности, требует выполнения условий $a \neq 0$ и $b \neq 0$ для уравнения кривой.

Параметры пользователя

Пользователь А:

- генерирует и хранит в секрете долговременный секретный ключ как целое число $0 < e_A < n$;
- вычисляет открытый ключ как точку кривой $Q_A = e_A G$. Открытый ключ доступен для всех пользователей системы.

Формирование ЭЦП

Пользователь А:

1. Вычисляет хэш-код сообщения М: $h = h(M) \bmod n$. Если $h = 0$, то принять $h = 1$.
2. Генерирует случайное целое число $0 < k_A < n$.
3. Вычисляет точку $R = k_A G = (x_1, y_1)$.
4. Вычисляет параметр $r = x_1 \bmod n$. При $r = 0$ возврат в п.2.
5. Вычисляет параметр $s = (k_A h + e_A r) \bmod n$. При $s = 0$ возврат в п.2.

6. Определяет цифровую подпись $DS = (r \| s)$ – в виде конкатенации двух двоичных векторов r и s .

Здесь, как и ранее, предполагается, что элемент x_1 поля $\mathbf{F}_q$ и хэш-функция h(M) представляются целыми числами с последующей редукцией по модулю n. В отличие от алгоритма ECDSA при формировании подписи здесь не требуется вычислять обратный элемент поля k_A^{-1} (см. п.5).

Для данной схемы можно записать ключевое уравнение

$$s = (k_A h + e_A r) \bmod n, \qquad (6.12)$$

совпадающее с п.5 алгоритма формирования. В отличие от (6.4) разовый ключ k_A здесь связан произведением не с параметром s, а с хэш-функцией h (т.е. перенесен из левой части уравнения в правую). Такая модификация DSA позволяет избежать трудоемкой инверсии элемента в поле и ускорить процедуру формирования. Данный алгоритм называют DSA-подобным.

Проверка ЭЦП

Пользователь B проверяет цифровую подпись пользователя A с помощью его открытого ключа Q_A, общесистемных параметров, алгоритма хэширования h(M) и подписанного сообщения (M, DS). Проверка заключается в вычислении на основе известных данных параметра r' и сравнении его с принятым значением r.

Умножив (6.12) на инверсию h^{-1} и учитывая, что $Q_A = e_A G$, для точек криптосистемы получим равенство

$$k_A G = h^{-1} s G - h^{-1} r\, e_A G = uG + vQ_A, \qquad (6.13)$$

$$u = h^{-1} s \bmod n, \quad v = -h^{-1} r \bmod n.$$

Итак, протокол проверки ЭЦП на основе (6.13) включает следующие вычисления.

Пользователь В:

1. По полученной подписи DS вычисляет целые числа r и s. Если $0 < r < n$, $0 < s < n$, то перейти к следующему шагу. В противном случае подпись неверна.

2. Вычисляет хэш-код полученного сообщения $H = h(M)$ как целое число.

3. Вычисляет $h = H \bmod n$. Если $h = 0$, то принять $h = 1$.

4. Вычисляет обратный элемент $h^{-1} \bmod n$ мультипликативной группы поля $\mathbf{F}_n$.

5. Вычисляет параметры $u = h^{-1}s \bmod n$, $v = -h^{-1}r \bmod n$.

6. Вычисляет точку $R' = uG + vQ_A = (x_1', y_1')$.

7. Вычисляет параметр $r' = x_1' \bmod n$.

8. Сравнивает вычисленное значение r' и принятое значение r. При равенстве $r = r'$ цифровая подпись принимается, в противном случае она неверна.

При проверке, как видим, требуется вычисление одной инверсии h^{-1} в поле.

Кроме алгоритмов формирования и проверки ЭЦП в ГОСТе приведены общие положения, математические определения, а также контрольный пример процессов формирования и проверки подписи для заданных параметров схемы цифровой подписи.

В таблице 6.6 приведен перечень различающихся обозначений, принятых в ГОСТ Р 34.10-2001 и в настоящей книге.

Таблица 6.6. Перечень обозначений, принятых в ГОСТ Р 34.10-2001 и в данной монографии

Параметр	Обозначение	ГОСТ Р 34.10-2001
Порядок группы точек эллиптической кривой	$N_E = cn$	$m = nq$
Генератор криптосистемы	G	P
Порядок точки G – генератора криптосистемы (простое число)	n	q
j-инвариант эллиптической кривой	j(E)	$J(E)$
Цифровая подпись	DS	ζ

В августе 2012 года в РФ были утверждены новые версии национальных стандартов ГОСТ Р34.10-2012 [87] и ГОСТ Р34.11-2012 [88], действующих с 2013 года. Первый регламентирует алгоритмы формирования и проверки ЭЦП, второй – алгоритм хэширования с хеш-кодом длиной 256 бит или 512 бит. Практически эти стандарты мало отличаются от предшествующих версий ГОСТ Р34.10-2001 и ГОСТ Р34.11-94. Здесь также отсутствуют сведения о рекомендуемых кривых, параметры которых «определяются на месте в зависимости от разрабатываемой системы». Существенное отличие новых стандартов – наряду с модулем поля p и размера хэш-кода длиной 256 бит утверждены те же параметры длиной 512 бит. Тем самым регламентируется 2 уровня стойкости систем ЭЦП, эквивалентных 128 и 256 битам симметричных шифров.

6.4.5. Государственный стандарт Украины ДСТУ 4145-2002

Этот стандарт утвержден приказом Госстандарта Украины в декабре 2002 года и вступил в действие с 2003 года [82]. Он рекомендует алгоритм цифровой подписи на эллиптических кривых, подобный алгоритму ECSS (с рядом модификаций), с хэшированием сообщения в соответствии с ГОСТ

34.311-95 [81] или другими стандартами. В настоящее время последний ГОСТ хэширования не используется.

Как и в предыдущем параграфе, сохраним принятые нами обозначения общесистемных параметров криптосистемы и цифровой подписи. Различия в обозначениях будут даны в таблице 6.11.

Общесистемные параметры в данном ГОСТе определены как:

$\mathbf{F}_2{}^m$ – основное поле Галуа как m-кратное расширение простого поля $\mathbf{F}_2$,

m – простое число в интервале [163, 509];

E: $y^2 + xy = x^3 + ax^2 + b$ – несуперсингулярная эллиптическая кривая над полем $\mathbf{F}_2{}^m$ с порядком $N_E = cn$ и коэффициентами $a = 0$ или $a = 1$ и $b \neq 0$.

G – генератор криптосистемы простого порядка n, $n > 2^{160}$;

$h(M)$ – хэш-функция, отображающая двоичные сообщения M произвольной длины в двоичный вектор длины $L(h)$ бит. Хэш-функция определена в ГОСТ 34.311 [81]. Допускается использование других стандартов хэш-функций.

Вычисления в поле производятся в полиномиальном или оптимальном нормальном базисах. В первом случае в качестве неприводимого полинома поля используются триномы или пентаномы в соответствии с таблицей 6.7. Оптимальный нормальный базис существует лишь при значениях m, приведенных в таблице 6.8. Кроме этих ограничений, общесистемные параметры криптосистемы проверяются на простоту порядка базовой точки n и на стойкость к MOV-атаке: порядок n криптосистемы не должен делить порядок мультипликативной группы расширенного поля Галуа $2^{mk} - 1$ (или $2^{mk} \neq 1 \bmod n$) с расширением $k = 2, 3, \ldots, 32$.

Пользователь А:

- генерирует и хранит в секрете долговременный секретный ключ как целое число $0 < e_A < n$;

- вычисляет открытый ключ как точку кривой $Q_A = -e_A G$. Открытый ключ доступен для всех пользователей системы.

Формирование ЦП

Пользователь А:

1. Вычисляет двоичный хэш-код сообщения M $h(\mathrm{M}) = (h_{255},\ldots,\ h_1,\ h_0)$ длиной 256 бит в соответствии с ГОСТ 34-311 (допускается использование других стандартов, которые определяются идентификатором iH).

2. Преобразует $h(M)$ в элемент поля $h = (h_{m-1},\ldots,h_1,\ h_0)$ с отсечением старших бит хэш-кода $h(M)$ при $m < 256$ и наращиванием в виде нулевых старших бит при $m > 256$ (если используются стандарты, отличные от ГОСТ 34.311, граница определяется размером хэш-функции). Если $h = 0$, то принимается $h = 1$.

3. Генерирует разовый секретный ключ как случайное целое число $0 < k_A < n$.

4. Вычисляет точку $R = k_A G = (x_\mathrm{R}, y_\mathrm{R})$. Параметр x_R называется *цифровой предподписью.* Если $x_R = 0$, возврат в п.3.

5. Вычисляет элемент поля как произведение $y = hx_R \in \mathbf{F}_2{}^m$.

6. Преобразует элемент поля y в целое двоичное число $r = \bar{y} < n$. При $r = 0$ возврат в п.3.

7. Вычисляет параметр $\mathrm{s} = (k_A + e_A r) \bmod n$. При $\mathrm{s} = 0$ возврат в п.3.

8. Определяет цифровую подпись $DS = (\mathbf{0} \,\|\, s \,\|\, \mathbf{0} \,\|\, r)$ длины, кратной 16 битам, в виде конкатенации наращенных нулями двоичных векторов s и r.

9. Направляет пользователю В подписанное сообщение ($iH,\ M,\ DS$), в котором iH – идентификатор хэш-функции, M – сообщение и DS – цифровая подпись.

Для данной схемы ключевое уравнение

$$s = (k_A + e_A r) \bmod n, \qquad (6.14)$$

совпадает с п.7 алгоритма формирования.

Проверка ЦП

Пользователь В проверяет цифровую подпись пользователя А с помощью его открытого ключа Q_A, общесистемных параметров, алгоритма хэширования h(М) и подписанного сообщения (М, DS). Проверка состоит в вычислении на основе известных данных параметра r' и сравнении его с принятым значением r.

Умножив (6.14) скалярно на G и учитывая, что $Q_A = -e_A G$, для точек криптосистемы получим равенство

$$k_A G = sG - r e_A G = sG + rQ_A, \qquad (6.15)$$

Итак, протокол проверки ЦП на основе (6.15) включает следующие вычисления.

Пользователь В:

1. По полученной подписи DS вычисляет целые числа s и r. Если $0 < s < n$, $0 < r < n$, то перейти к следующему шагу. В противном случае подпись неверна.

2. Вычисляет хэш-код полученного сообщения $h(M)$ как двоичный вектор стандартной разрядности.

3. Преобразует двоичный хэш-код $h(M)$ в элемент поля $h = (h_{m-1}, h_{m-2}, \ldots, h_1, h_0)$ разрядности m. Если $h = 0$, то принять $h = 1$.

4. Вычисляет точку $R' = sG + rQ_A = (x_R', y_R')$.

5. Вычисляет элемент поля $y' = hx_R' \in \mathbf{F}_2{}^m$.

6. Преобразует элемент поля y' в целое двоичное число $r' = \bar{y}' < n$.

7. Сравнивает вычисленное значение параметра r' и принятое значение r. При равенстве $r = r'$ цифровая подпись принимается, в противном случае она неверна.

В результате проверки пользователь В удостоверяется в подлинности отправителя А и целостности сообщения M.

Примечание 1. Для снижения временных затрат при формировании ЦП допускаются предвычисления серии секретных сеансовых ключей k_A и, соответственно, точек $R = k_A G$ (п.3 и 4 алгоритма формирования при этом заменяются селекцией одной из пар k_A и R). Они должны храниться в секретной базе данных, случайным образом выбираться для каждого сеанса с обязательным уничтожением использованной пары k_A и R.

Примечание 2. Преобразование элемента поля $\mathbf{F}_{2^m}$ в двоичное число в п.6 алгоритмов формирования и проверки подписи осуществляется путем отсечения старших разрядов двоичного вектора y так, чтобы разрядность $\bar{y}$ не превышала сниженной на 1 разрядности двоичного представления n ($L(\bar{y}) \le L(n) - 1$). В этом случае выполняется неравенство $\bar{y} < n$.

Примечание 3. Разрядность цифровой подписи $L(DS)$ должна быть кратной 16 (или $L(DS) = 0 \bmod 16$) с минимальной избыточностью. Поэтому каждый из двоичных векторов s и r в п.8 наращивается нулями в старших разрядах до разрядности $L(DS)/2$, после чего общая конкатенация расширенных параметров s и r образует цифровую подпись DS.

Примечание 4. Использование различных стандартов хэш-функций в одной группе пользователей делает необходимым включение идентификатора iH хэш-функции в пересылаемое подписанное сообщение. По согласованию сторон он может быть принят по умолчанию и опущен.

Примечание 5. Приведенные алгоритмы формирования и проверки включают основные вычислительные процедуры. В стандарте ДСТУ 4145-2002, кроме того, параметры подписи подвергаются проверкам на корректность. Эти проверки здесь не приведены в интересах более лаконичного изложения.

Таблица 6.7. Степени рекомендуемых в ДСТУ-4145 полей и неприводимых полиномов (триномов или пентаномов)

Степень поля m	Примитивний полином P(x)	Степень поля m	Примитивний полином P(x)
163	$x^{163}+x^7+x^6+x^3+1$	337	$x^{337}+x^{10}+x^6+x+1$
167	$x^{167}+x^6+1$	347	$x^{347}+x^{17}+x^6+x+1$
173	$x^{173}+x^{10}+x^2+x+1$	349	$x^{349}+x^6+x^5+x^2+1$
179	$x^{179}+x^4+x^2+x+1$	353	$x^{353}+x^{26}+x^7+x^3+1$
181	$x^{181}+x^7+x^6+x+1$	359	$x^{359}+x^{18}+x^4+x^2+1$
191	$x^{191}+x^9+1$	367	$x^{367}+x^{21}+1$
193	$x^{193}+x^{15}+1$	373	$x^{373}+x^9+x^6+x+1$
197	$x^{197}+x^{21}+x^2+x+1$	379	$x^{379}+x^{17}+x^6+x+1$
199	$x^{199}+x^{11}+x^2+x+1$	383	$x^{383}+x^9+x^5+x+1$
211	$x^{211}+x^{12}+x^6+x+1$	389	$x^{389}+x^{17}+x^{10}+x+1$
223	$x^{223}+x^{12}+x^2+x+1$	397	$x^{397}+x^{22}+x^3+x+1$
227	$x^{227}+x^{21}+x^2+x+1$	401	$x^{401}+x^{29}+x^4+x+1$
229	$x^{229}+x^{21}+x^2+x+1$	409	$x^{409}+x^{15}+x^6+x+1$
233	$x^{233}+x^9+x^4+x+1$	419	$x^{419}+x^{21}+x^{14}+x+1$
239	$x^{239}+x^{15}+x^2+x+1$	421	$x^{421}+x^7+x^4+x+1$
241	$x^{241}+x^{15}+x^4+x+1$	431	$x^{431}+x^5+x^3+x+1$
251	$x^{251}+x^{14}+x^4+x+1$	431	$x^{433}+x^{15}+x^5+x+1$
257	$x^{257}+x^{12}+1$	439	$x^{439}+x^8+x^3+x^2+1$
263	$x^{263}+x^{27}+x^2+x+1$	443	$x^{443}+x^{28}+x^3+x+1$

269	$x^{269}+x^7+x^6+x+1$	449	$x^{449}+x^{25}+x^5+x^3+1$
271	$x^{271}+x^{16}+x^3+x+1$	457	$x^{457}+x^{16}+1$
277	$x^{277}+x^{23}+x^3+x^2+1$	461	$x^{461}+x^{23}+x^4+x+1$
281	$x^{281}+x^9+x^4+x+1$	463	$x^{463}+x^{24}+x^3+x+1$
283	$x^{283}+x^{26}+x^9+x+1$	467	$x^{467}+x^{28}+x^3+x+1$
293	$x^{293}+x^{11}+x^6+x+1$	479	$x^{479}+x^{25}+x^6+x+1$
307	$x^{307}+x^8+x^4+x^2+1$	487	$x^{487}+x^{15}+x^2+x+1$
311	$x^{311}+x^{29}+x^4+x+1$	491	$x^{491}+x^{17}+x^6+x^2+1$
313	$x^{313}+x^7+x^3+x+1$	499	$x^{499}+x^{29}+x^6+x^2+1$
317	$x^{317}+x^9+x^5+x^2+1$	503	$x^{503}+x^3+1$
331	$x^{331}+x^{12}+x^5+x^2+1$	509	$x^{509}+x^{23}+x^3+x^2+1$

Таблица 6.8. Степени полей в списке рекомендуемых ДСТУ-4145, для которых существует ОНБ

Степень m	173	179	191	233	239	251	281
	293	359	419	431	443	491	509

Отметим некоторые отличия данного алгоритма от схемы ECSS (п.6.3.2). Во-первых, вместо целого числа $h(M)$ из хэш-функции в стандарте Украины определяется элемент поля $h \in \mathbf{F}_2{}^m$. Далее, операция сложения в п.5 алгоритма ECSS заменена умножением элементов поля $h \cdot x_R$. Операция умножения вводит нелинейность в уравнение, связывающее параметры подписи, и усложняет

криптоанализ. Наконец, открытый ключ пользователя A определен через обратную точку $Q_A = -e_A G$. Соответственно, изменился знак перед составляющей $e_A r$ в п.7 алгоритма.

Следует отметить, что стандарт ДСТУ 4145-2002 предлагает пользователям самый широкий диапазон из 60-ти простых значений расширения поля m = 163, 167, …, 509, что позволяет выполнять криптосистемы с самыми различными требованиями по стойкости как для правительственных служб, так и предприятий и физических лиц. В этом отношении российский стандарт с практически единственным размером поля в 256 бит (с 2013 года добавлен модуль длиной 512 бит) явно ограничивает требования и возможности заказчиков и производителей криптосистем. Он был ориентирован на использование единственного алгоритма хэш-функции, принятого в ГОСТ Р 34.11-94 (ныне – в ГОСТ Р 34.11-2012).

В приложениях Г.1 и Г.2 стандарта даны параметры 10 кривых в полиномиальном базисе и 5 кривых в оптимальном нормальном базисе, которые наряду с другими можно использовать в криптосистемах. Они приведены в таблицах 6.9 и 6.10 соответственно. В них, к сожалению, не указаны значения кофактора c в разложении порядка кривой, но можно быть уверенным, что $c = 2$ при $a = 1$ и $c = 4$ при $a = 0$. Кривые таблицы 6.10 совпадают с точностью до изоморфизма с пятью кривыми из таблицы 6.9, для степеней m которых существует ОНБ. Одна из кривых со степенью поля $m = 233$ изоморфна кривой B-233 стандарта FIPS 186-2 (см. таблицу 6.5). Эти кривые имеют одинаковый порядок, следовательно, их коэффициенты b принадлежат одному классу сопряженных элементов поля $\mathbf{F}_2{}^m$. Другие кривые, как отмечено в п.7.2 стандарта, «предоставляются в установленном порядке уполномоченным исполнительным органом государственной власти».

После выбора m и n случайным образом выбирается точка G кривой, и если nG = O, то она может быть принята за генератор криптосистемы. В противном случае тестируется на выполнение условия nG = O другая случайная точка.

Секретный ключ e_A пользователя A должен генерироваться случайно в интервале [1, n – 1], при этом условия его формирования и хранения должны

исключать несанкционированный доступ к ключу или его части, а также ко всем промежуточным данным, которые использовались при вычислении ключа. Условия хранения ключа должны исключить возможность его модификации, уничтожения или подмены.

Открытый ключ пользователя рассчитывается как точка кривой $Q = -eG$. Хранить и передавать ключ Q можно в сжатом виде [99], что вдвое сокращает его размер. Условия хранения открытого ключа не должны допускать возможности его модификации или подмены.

В Приложении А стандарта приведены требования к генератору случайных последовательностей, который используется для генерации секретного ключа цифровой подписи. Он реализуется на основе шифратора в режиме простой замены в соответствии с алгоритмом, приведенным в ГОСТ 28147-89.

Приложение Б включает примеры вычисления параметров (r, s) ЦП и проверки ЦП при операциях в полиномиальном и оптимальном нормальном базисе. Для этого используются первые кривые из таблиц 6.9 и 6.10 со степенями расширений поля $m = 163$ и $m = 173$.

В Приложении В рассмотрены основные математические понятия, используемые в стандарте: конечные группы и поля, операции в поле $\mathbf{F}_2{}^m$ в полиномиальном и нормальном базисах, операции в группе точек эллиптической кривой, порядок кривой и порядок точки и другие понятия.

В таблице 6.11 даны различающиеся с принятыми в книге обозначения параметров в стандарте ДСТУ 4145-2002.

Таблица 6.9. Параметры рекомендуемых в ДСТУ-4145 кривых при использовании полиномиального базиса поля

№ п/п	*m*	Параметры кривой (полиномиальный базис)
1	163	*a* = 1 *b* = 5FF6108462A2DC8210AB403925E638A19C1455D21 *n* = 400000000000000000002BEC12BE2262D39BCF14D
2	167	*a* = 1 *b* = EE3CEEB230811759F20518A0930F1A4315A827DAC *n* = FFFFFFFFFFFFFFFFFFFFB12EBCC7D7F29FF7701F
3	173	*a* = 0 *b*=108576C80499DB2FC16EDDF6853BBB278F6B6FB437D9 *n* = 800000000000000000000189B4E67606E3825BB2831
4	179	*a* = 1 *b*=4A6E0856526436F2F88DD07A341E32D04184572BEB710 *n*=3FFFFFFFFFFFFFFFFFFFFB981960435FE5AB64236EF
5	191	*a* = 1 *b*=BC86E2102902EC4D5890E8B6B4981ff27E0482750FEFC03 *n*=40000000000000000000000069A779CAC1DABC6788F7474F
6	233	*a* = 1 *b*=06973B15095675534C7CF7E64A21BD54EF5DD3B8A0326AA936ECE454D2C

		n=1000000000000000000000000000013E974E72F8A6922031D2603CFE0D7
7	257	*a* = 0 *b*=1CEF494720115657E18F938D7A7942394FF9425C1458C57861F9EEA6ADBE3BE10 *n*=800000000000000000000000000000006759213AF182E987D3E17714907D470D
8	307	*a* = 1 *b*=393C7F7D53666B5054B5E6C6D3DE94F4296C0C599E2E2E241050DF18B6090BDC90186904968BB *n*=FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFC079C2F3825DA70D390FBBA588D4604022B7B7
9	367	*a* = 1 *b*=43FC8AD242B0B7A6F3D1627AD5654447556B47BF6AA4A64B0C2AFE42CADAB8F93D92394C79A79755437B56995136 *n*=400000000000000000000000000000000000000000000009C300B75A3FA824F22428FD28CE8812245EF44049B2D49
10	431	*a* = 1 *b*=03CE10490F6A708FC26DFE8C3D27C4F94E690134D5BFF988D8D28AAEAEDE975936C66BAC536B18AE2DC312CA493117DA/A469C640CAF3 *n*=3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF/FFFFFBA3175458009A8C0A724F02F81AA8A1FCBAF80D90C7A9/5110504CF

Таблица 6.10. Параметры рекомендуемых в ДСТУ-4145 кривых при использовании нормального базиса поля

№ п/п	*m*	Параметры кривой (нормальный базис)
1	173	*a*=0 *b*=043D7E139319F43BA00944915740E1E6651B06E278C7 *n*=800000000000000000000189B4E67606E3825BB2831
2	179	*a*=1 *b*=0B151F20BCB9AC860B2F2CCE4F5E27EC1840C9E9D30BB *n*=3FFFFFFFFFFFFFFFFFFFFFB981960435FE5AB64236EF
3	191	*a*=1 *b*=09C38E4E94EB67753A07EABA227B97EDD72ACE09D0F18F7C *n*=4000000000000000000000069A779CAC1DABC6788F7474F
4	233	*a*=1 *b*=080F920952A702C75B704A424C018EEA55AA44664F3A003E0962D4F9A8E *n*=1000000000000000000000000000013E974E72F8A6922031D2603CFE0D7
5	431	*a*=1 *b*=19794FD03E5CFCBC1B17A03BF8ED0D946F57FE2C1AE97E4A669FDBD7B/ DA20380053791356B5D6941BC3235EC1F8FCA04085260CA9B77

		n=3FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF FFFFFFFFFFBA3175458009A8C0A724F02F81AA8A1FCBA F80D90C7A95110504CF

Таблица 6.11. Перечень обозначений, принятых в ДСТУ 4145-2002 и в данной книге

Параметр	Обозначение	ДСТУ 4145
Расширенное поле Галуа	$\mathbf{F}_2{}^m$	$GF(2^m)$
Примитивный полином, образующий полиномиальный базис	$P(x)$	$f(x)$
След элемента поля	$Tr(x)$	$tr(x)$
Коэффициенты эллиптической кривой	a, b	A, B
Порядок группы точек эллиптической кривой	$N_E = cn$	–
Генератор криптосистемы – точка простого порядка n (базовая точка)	G	P
Сообщение	M	T
Хэш-функция	$h(M)$	$H(T)$
Сеансовый секретный ключ	k_A	e
Предподпись	x_R	F_e
Цифровая подпись	DS	D

В заключение заметим, что данный стандарт не обновлялся с 2002 года и объективно устарел. Кривые над полем характеристики 2 сегодня рассматриваются как наиболее уязвимые по отношению к различным атакам [15] и не рекомендуются при разработке новых криптосистем. Лучшей альтернативой им, несомненно, являются быстрые кривые в форме Эдвардса.

СПИСОК ЛИТЕРАТУРЫ

1. Edwards H.M. A normal form for elliptic curves. Bulletin of the American Mathematical Society, Volume 44, Number 3, July 2007, PP. 393-422.
2. Bernstein D.J., Lange T. Faster Addition and Doubling on Elliptic Curves // Advancesin Cryptology—ASIACRYPT'2007 (Proc. 13th Int. Conf. On the Theory and Applicationof Cryptology and Information Security. Kuching, Malaysia. December 2–6, 2007). Lect. Notes Comp. Sci. V. 4833. Berlin: Springer, 2007. PP. 29–50.
3. Bernstein Daniel J., Birkner Peter, Joye Marc, Lange Tanja, Peters Christiane. Twisted Edwards Curves. //IST Programme under Contract IST–2002–507932 ECRYPT, and in part by the National Science Foundation under grant ITR–0716498, 2008, PP. 1-17.
4. Hisil Huseyin, Koon-Ho Wong Kenneth, Carter Gary, Dawson Ed. Twisted Edwards Curves Revisited // ASIACRYPT. – 5350. – New York: Springer, 2008. – PP. 326-343.
5. Bernstein D.J., Lange T. Inverted Edwards coordinates. National Science Foundation under grant ITR–0716498, 2007, 331 – 8 and in part by the European Commission through the IST Programme under Contract IST–2002–507932 ECRYPT.
6. Bernstein D. J., Lange T., Explicit-formulas database (2007). hyperelliptic.org/EFD.
7. Bernstein Daniel J., Lange Tanja. Farashahi Reza Rezaeian. Binary Edwards curves. Cryptographic hardware and embedded systems–CHES 2008, 10th international workshop, Washington, D.C., USA, August 10–13, 2008, PP. 224-256.
8. Bernstein Daniel J., Batch binary Edwards. Advances in cryptology–Crypto

2009, 29th annual international cryptology conference, Santa Barbara, CA, USA, August 16–20, 2009, PP. 317-336.

9. Bernstein D.J., Lange T. Failures in NIST's ECC standards. European Commission under Contract ICT-645421 ECRYPT-CSA; by the Netherlands Organisation for Scientific Research (NWO) under grant 639.073.005; and by the U.S. National Science Foundation under grant 1018836.

10. Moloney R., McGuire G. Two kinds of division polynomials for twisted Edwards curves. Applicable Algebrain Engineering, Communication and Computing, 2011, PP. 321-345.

11. Bernstein D.J., Birkner P., Lange T., Peters C. ECM using Edwards curves. European Commission through the ICT Programme under Contract ICT–2007–216676 ECRYPT-II, and in part by the National Science Foundation under grant ITR–0716498.

12. Morain F. Edwards curves and CM curves. ArXiv 0904/2243v1 [Math.NT] Apr.15, 2009.

13. Miller V.S. Use of Elliptic Curves in Cryptography. Advances in Cryptology – Proceedings of CRYPTO”85, Springer Verlag Lecture in Computer Science 218, 1986. – PP. 417-726.

14. Koblitz N. Elliptic Curve Cryptosystems.//Mathematics of. Computation, 48, 1987. – PP 203-209.

15. Koblitz N., Menezes A.J., A Riddle Wrapped in an Enigma. Technical Reports CACR-2015-14. Available: www.cacr.math.uwaterloo.ca.

16. L. C. Washington. Elliptic Curvres. Number Theory and Cryptography. Second Edition. CRC Press, 2008.

17. Handbook of elliptic and hyperelliptic curve cryptography / Scientific editors, Henri Cohen & Gerard Frey ; authors, Roberto M Avanzi Roberto Avanzi, Christophe Doche, Tanja Lange, Kim Nguyen, and Frederik Vercauteren. Chapman & Hall/CRC, Taylor & Francis Group, London, New York, Singapur, 2006,

18. Blahut R.E. Cryptography and Secure Communication. Cambridge University Press, 2014.

19. Мао, Венбо. Современная криптография: теория и практика: Пер. с англ – М: Издательский дом «Вильямс», 2005. – 768с.

20. Лидл Р., Нидеррайтер Г. Конечные поля / Р. Лидл, Г. Нидеррайтер // «Мир», Том 1 – 1988 – 273 стр.
21. Лидл Р., Нидеррайтер Г Конечные поля / Р. Лидл, Г. Нидеррайтер // «Мир», Том 2 – 1988 – 822 стр.

22. Коблитц Н. Введение в эллиптические кривые и модулярные формы. Пер. с англ. – М.: Мир, 1988. – 320 с.
23. Koblitz Neal. A course in number theory and cryptography. – New York, Springer Verlag, , 1993. – 179p.

24. Степанов С.А. Арифметика алгебраических кривых. – М.: Наука. 1991. – 368 с.
25. Прасолов В.В., Соловьев Ю.П. Эллиптические функции и алгебраические уравнения. – М.: Факториал, 1997. – 288 с.
26. Айерлэнд К., Роузен М. Классическое введение в современную теорию чисел. Пер. с англ. – М.: Мир, 1987. – 416 с.

27. Ковальчук Л.В. Рекурентні алгоритми обчислення кореню довільного степеню у кільці лишків. / Л.В. Ковальчук, О.Ю. Беспалов, П.В. Огнєв // Правове, нормативне та метрологічне забезпечення захисту інформації в Україні, випуск 25 – 2013 – С. 58-66.
28. Горбенко І.Д., Горбенко Ю.І. Прикладна криптологія. Теорія, практика, застосування: монографія. – Харків: Видавництво «Форт», 2012. – 870с.

29. Бессалов А.В., Телиженко А.Б. Криптосистемы на эллиптических кривых: Учеб. пособие. – К.: ІВЦ «Політехніка», 2004. – 224с.
30. Бессалов А.В. Число изоморфизмов и пар кручения кривых Эдвардса над простым полем. Радиотехника, вып. 167, 2011. С. 203-208.
31. Бессалов А.В., Дихтенко А.А.,Третьяков Д.Б. Сравнительная оценка быстродействия канонических эллиптических кривых и кривых в форме Эдвардса над конечным полем. Сучасний захист інформації, №4, 2011. С.33-36.
32. Бессалов А.В., Гурьянов А.И., Дихтенко А.А. Кривые Эдвардса почти простого порядка над расширениями малых простых полей. Прикладная радиоэлектроника, 2012, Том 11, №2. С. 225-227

33. Бессалов А.В., Дихтенко А.А., Третьяков Д.Б. Оценка реальной стойкости криптосистемы на кривой Эдвардса над расширениями малых полей. Сучасний захист інформації, №2, 2012. С.17-20.
34. Бессалов А.В. Деление точки на два для кривой Эдвардса над простым полем. Прикладная радиоэлектроника, 2013, Том 12, №2. – С. 278-279.
35. Бессалов А.В., Третьяков Д.Б. Удвоение точки и обратная задача для кривой Эдвардса над простым полем. Сучасний захист інформації, №3, 2013. – С.56-58.
36. Бессалов А.В., Діхтенко А.А., Яценко О.І. Параметри криптосистеми на кривій Едвардса над розширеннями малих простих полів. Прикладная радиоэлектроника, 2013, Том 12, №2. – С.273-277.
37. Бессалов А.В., Дихтенко А.А. Криптостойкие кривые Эдвардса над простыми полями. Прикладная радиоэлектроника, 2013, Том 12, №2. – С. 285-291.
38. Бессалов А.В., Діхтенко А.А. Параметри генератора криптосистеми на кривий Эдвардса над розширеннями простих полів. Спеціальні телекомунікаційні системи та захист інформації. Збірник наукових праць, випуск 1(23), 2013. – С.5 – 8.
39. Бессалов А.В., Дихтенко А.А., Цыганкова О.В. Алгоритм выбора канонической кривой, изоморфной кривой Эдвардса над простым полем. Радиотехника, №175, 2014.– С.195-198.
40. Бессалов А.В., Дихтенко А.А. Изоморфные канонической форме эллиптические кривые Эдвардса над расширенными полями характеристики 2. Радиотехника, №175, 2014. – С. 200-205.
41. Бессалов А.В., Дихтенко А.А., Цыганкова О.В. Мощность семейства эллиптических кривых, изоморфных кривым Эдвардса над простым полем. Захист інформації - Том 16, №1, січень-березень 2014. – С.23-28.
42. Бессалов А.В., Дихтенко А.А., Цыганкова О.В. Плотность канонических эллиптических кривых со свойством изоморфизма к форме Эдвардса. Известия ЮФУ. Технические науки.", вып. №4, 2014. – С.146-153. http://izv-tn.tti.sfedu.ru/?cat=412.
43. Бессалов А.В., Дихтенко А.А. Изоморфизм несуперсингулярных кривых над полями характеристики 2 и кривых Эдвардса с одним параметром. Радиотехника, №176, 2014. – С.88-93.

44. Бессалов А.В., Ковальчук Л.В. Точное число эллиптических кривых в канонической форме, изоморфных кривым Эдвардса над простым полем. Кибернетика и системный анализ, т.51, №2, 2015. – С.3-12.
45. Bessalov A.V., Kovalchuk L.V. Exact Number of Elliptic Curves in the Canonical Form, Which are Isomorphic to Edwards Curves Over Prime Field. Cybernetics and Systems Analysis: Volume 51, Issue 2 (2015), Pages 165-172. http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s10559-015-9709-x
46. Бессалов А.В. Построение кривой Эдвардса на базе изоморфной эллиптической кривой в канонической форме. Прикладная радиоэлектроника, 2014, Том 13, №3. – С.286-289.
47. Бессалов А.В., Третьяков Д.Б., Цыганкова О.В. Новый подход к определению точного числа кривых Эдвардса над простым полем. Сучасний захист інформації, №3, 2014. – С.11-15.
48. Бессалов А.В., Цыганкова О.В. Метод определения точек максимального порядка на кривой Эдвардса**.** Спеціальні телекомунікаційні системи та захист інформації. Збірник наукових праць, випуск 2(26), 2014. – С.18-21.
49. Бессалов А.В., Цыганкова О.В. Новые свойства эллиптической кривой в форме Эдвардса над простым полем. Радиотехника №180, 2015. – С.137-143.
50. Бессалов А.В., Цыганкова О.В. Взаимосвязь семейств точек больших порядков кривой Эдвардса над простым полем. Захист інформації –Том 17, №1, січень-березень 2015. –С.73-80.
51. Бессалов А.В., Цыганкова О.В. Взаимосвязь семейств точек больших порядков кривой Эдвардса над простым полем. Проблемы передачи информации, - Том 51, вып 4, 2015. – С.92-98.
52. Bessalov A.V., Tsygankova O.V. Interrelation of families of points of high order on the Edwards curve over a prime field. Problems of Information Transmission, 51(4), 2015. PP.391-397. http://link.springer.com/article/10.1134/S0032946015040080.
53. Бессалов А.В., Цыганкова О.В. Производительность групповых операций на скрученной кривой Эдвардса над простым полем. Радиотехника №181, 2015. – С.58-63.

54. Бессалов А.В., Цыганкова О.В. Классификация кривых в форме Эдвардса над простым полем. Прикладная радиоэлектроника: научно-техн. журнал. – 2015. – Том 14. – №4. – С.197 -203.
55. Ковальчук Л.В., Бессалов А.В. Беспалов О.Ю., Алгоритмы генерации базовой точки кривой Эдвардса с использованием критериев делимости точки кривой. Кибернетика и системный анализ, т.52, №5, 2016. – С.14-24.
56. Бессалов А.В., Цыганкова О.В. Число кривых в обобщенной форме Эдвардса с минимальным четным кофактором порядка кривой. Проблемы передачи информации, - Том 53, вып 1, 2017. – С.101-111.
57. Бессалов А.В., Трет′яков Д.Б., Циганкова О.В. Властивості точок малих порядків кривих в узагальненої формі Едвардса. Сучасний захист інформації, №2, 2016. – С.46-54.
58. Бессалов А.В. Метод нахождения порядка точки скрученной кривой Эдвардса. Радиотехника, вып.186, 2016. – С. 110-118.
59. Бессалов А.В. Метод решения проблемы дискретного логарифмирования на эллиптической кривой путем деления точек на два. Кибернетика и системный анализ, №6, 2001. – С.50-53.
60. Бессалов А.В., Діхтенко А.А., Яценко О.І. Загальносистемні параметри криптосистеми на кривій Едвардса над розширеннями малих простих полів. Збірник наукових праць міжнародної проблемно-наукової міжгалузевої конференції «ІНФОРМАЦІЙНІ ПРОБЛЕМИ КОМП'ЮТЕРНИХ СИСТЕМ, ЮРИСПРУДЕНЦІЇ, ЕНЕРГЕТИКИ, ЕКОНОМІКИ, МОДЕЛЮВАННЯ ТА УПРАВЛІННЯ», (ПНМК – 2012, 7 – 10 червня), Випуск №8, Бучач, 2012. – С.36-37.
61. Бессалов А.В., Дихтенко А.А. Определение параметров криптостойких кривых Эдвардса над простыми полями. XVI международная научно-практическая конференция «Безопасность информации в информационно-телекоммуникационных системах». 21 – 24 мая 2013 г. Тезисы докладов. – К.: ООО «ИП Эдельвейс», НИЦ «Тезис» НТУУ «КПИ», 2013. – С.36-37.
62. Бессалов А.В., Діхтенко А.А. Кривые Эдвардса над простыми полями с почти простым значением порядка. Матеріали Двадцятої всеукраїнської науково-практичної конференції "Інноваційний потенціал української науки -- XXI сторіччя". Квітень, 2013. http://nauka.zinet.info/konf_20.php.

63. Бессалов А.В., Цыганкова О.В. Свойства точек больших порядков кривой Эдвардса. Матеріали XVII міжнародної науково-практичної конференції «Безпека інформації у інформаційно-телекомунікаційних системах», 26-28 травня 2015р., м. Київ. – С. 30-31.
64. Ковальчук Л.В., Бессалов А.В., Беспалов О.Ю. Порівняний аналіз алгоритмів генерації базової точки на кривій Эдвардса. Матеріали XVII міжнародної науково-практичної конференції «Безпека інформації у інформаційно-телекомунікаційних системах», 26-28 травня 2015р., м. Київ. –С. 32-33.
65. Бессалов А.В., Циганкова О.В. Класифікація кривих в узагальненої формі Эдвардса. Матеріали XVIII міжнародної науково-практичної конференції «Безпека інформації у інформаційно-телекомунікаційних системах», 25-26 травня 2016р., м. Київ. С.9.
66. Бессалов А.В., Олешко К.А., Поречная Д.Н., Цыганкова О.В., Черный О.Н. Криптостойкие скрученные кривые Эдвардса с минимальной сложностью групповых операций. Прикладная радиоэлектроника: научно-техн. журнал. – 2016. – Том 15. – №3. – С.141 – 150.
67. Дэвенпорт Г. Высшая арифметика: введение в теорию чисел // Пер. с англ. под редакцией Ю.В.Линника. – М: «Наука», 1965. – 176с.
68. Silverman J. H. The arithmetic of elliptic curves. – New York, Springer Verlag, 1986 – 868p.
69. Diffie W. and Hellman M.E. New Direction in Cryptography. //IEEE Transactions on Information Theory, V22, 1976. – PP. 644-654.
70. T. ElGamal. A Public Key Cryptosystems and a Signature Scheme Based on Discrete Logarithms.// IEEE Transactions on Information Theory, V31, 1985. – PP 469-472.
71. Current Public-Key Cryptographic Systems. A Certicom Whitepaper. Certicom, 1997. www.certicom.com.
72. Elliptic Curve and Cryptography. A Certicom Whitepaper. Certicom, 1998. www.certicom.com.
73. FIPS 186. Digital Signature Standard. National Institute of Standard and Technology. 1994.
74. IEEE P1363 STANDARD. Part 6: Elliptic Curve Systems (Draft 5). 1995.

75. IEEE P1363-2000. Standard Specifications for Public Key Cryptography. Institute of Electrical and Electronics Engineers, Inc.,2000.
76. ISO/IEC 15946:2002, Information technology — Security techniques — Specification of TTP services to support the application of digital signatures.
77. ANSI X9.62-1999. Public Key Cryptography for the Financial Services Industry: The Ellipticф Curve Digital Signature Algorithm (ECDSA). 1999.
78. FIPS 186-2. Digital Signature Standard (DSS). National Institute of Standard and Technology. January, 2000.
79. FIPS PUB 186-4. Federal Information Processsing Standards Publication. Digital Signature Standard (DSS). National Institute of Standard and Technology. 2013.
80. ГОСТ Р34.10-2001.Инормационная технология. Криптографическая защита информации. Процедуры формирования и проверки электронной цифровой подписи. – М.: Госстандарт России, 2001. – 20с.
81. ГОСТ Р34.11-1994. Информационная технология. Криптографическая защита информации. Функция хэширования. – М.: Госстандарт России, 1994. – 12с.
82. Державний стандарт України ДСТУ 4145-2002. Інформаційні технології. Криптографічний захист інформації. Цифровий підпис, що ґрунтується на еліптичних кривих.Формування та перевіряння. Київ, Держстандарт України, 2003. – 94с.
83. ISO/IEC 14888-1,2,3:2008.
84. Державний стандарт України ДСТУ ISO/IEC 14888-1:2014. Інформаційні технології. МЕТОДИ ЗАХИСТУ. ЦИФРОВІ ПІДПИСИ З ДОПОВНЕННЯМ. Частина 1. Загальні положення. 2014.
85. Державний стандарт України ДСТУ ISO/IEC 14888-3:2014. Інформаційні технології. МЕТОДИ ЗАХИСТУ. ЦИФРОВІ ПІДПИСИ З ДОПОВНЕННЯМ. Частина 3. Механізми, що грунтуються на дискретному логарифмі. 2014.
86. Закон України «Про електронний цифровий підпис» від 22.05.2003 №852-IX.
87. Национальный стандарт Российской Федерации ГОСТ Р34.10-2012. Инормационная технология. Криптографическая защита информации.

Процессы формирования и проверки электронной цифровой подписи. – М.: Стандартинформ, 2012. – 29с.
88. Национальный стандарт Российской Федерации ГОСТ Р34.11-2012. Информационная технология. Криптографическая защита информации. Функция хэширования. – М.: Стандартинформ, 2012. – 28с.
89. Brown D.R.L. The Exact Security of ECDSA. Certicom Research, CORR-2000-54, Canada. – PP. 1-5.
90. Brown D.R.L. Johnson D.B. Formal Security Proofs for a Signature Scheme with Partial Message Recovery. Certicom Research, CORR-2001-55, Canada.
91. Giuliani K.J. Attacks on Elliptic Curve Discrete Logarithm Problem. University Of Waterloo, Waterloo, Ontario, Canada, 1999. www.cacr.math.uwaterloo.ca.
92. Wiener M. J. And Zuccherato R.J. Faster Attacks On Elliptic Curve Cryptosystems. SELECTED AREAS IN CRYPTOGRAPHY, Lecture Notes in Computer Science, V1556, Springer-Verlag, 1999. – PP. 252-266.
93. Jacobson M., Menezes A.. Solving Elliptic Curve Discrete Logarithm Problems Using Weil Descent. University of Illinois, May, 2001. E-mail: amenezes@certicom.com.
94. Menezes A.,.Qu M. Analysis of the Weil Descent Attack of Gaudry, Hess and Smart. University of Waterloo, Canada, 2001. E-mail: amenezes@certicom.com.
95. ISO/IEC JTC 1/SC 27 n 2303, CD 15946-2. Information Technology- Sequrity Techniques – Cryptographic Techniques based on Elliptic Curves: Part 2- Digital Signatures. 1999 -05-26..
96. Brown M., Hankerson D., Lopez J.and Menezes A. Software Implementation of the NIST Elliptic Curves Over Prime Fields. Certicom Research, CORR-2000-56, Canada. www.cacr.math.uwaterloo.ca.
97. Hankerson D., Lopez J.and Menezes A. Software Implementation of Elliptic Curve Cryptography Over Binary Fields. Certicom Research, CORR-2000-42, Canada. www.cacr.math.uwaterloo.ca.
98. Wiener M. J. And Zuccherato R.J. Faster Attacks On Elliptic Curve Cryptosystems. SELECTED AREAS IN CRYPTOGRAPHY, Lecture Notes in Computer Science, V1556, Springer-Verlag, 1999. – PP. 252-266.
99. Seroussi G. Compact Representation of Elliptic Curve Points over $GF(2^m)$. Hewlitt-Packard Laboratories Technical Report No HPL-98-135. – 1998.

Оглавление

Науково-методичне видання

Бессалов Анатолий Владимирович

ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ

В ФОРМЕ ЭДВАРДСА

И КРИПТОГРАФИЯ

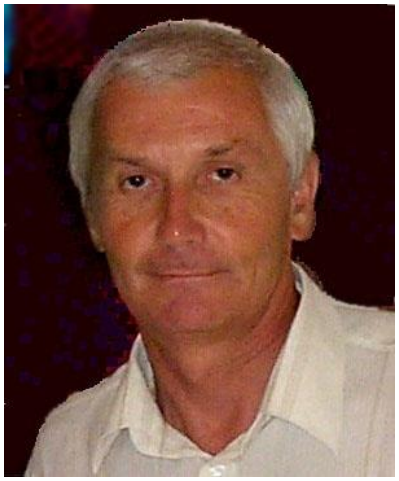
МОНОГРАФИЯ

Відповідальний за випуск

Підписано до друку

Формат. Папір

ІВЦ «Видавництво «Політехніка»»



Бессалов Анатолий Владимирович bessalov@ukr.net

Закончил Киевское высшее инженерно-авиационное военное училище ВВС (1968). Защитил кандидатскую (1974) и докторскую (1993) диссертации по специальности «Вооружение и военная техника». Профессор кафедры математических методов защиты информации НТУУ «КПИ», академик международной академии наук Прикладной радиоэлектроники. Область научных интересов: теория корректирующего кодирования, асимметричная криптография на эллиптических кривых. Автор 23-х учебных пособий, свыше 200 научных работ, 27 из них переведены за рубежом и индексированы в НМБ «Scopus». Подготовил к защите 2-х докторов и 8 кандидатов технических наук.