Федеральное агентство по образованию
Государственное образовательное учреждение высшего профессионального образования
САРАТОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО
Кафедра теоретических основ компьютерной безопасности и криптографии
КУРСОВАЯ РАБОТА
Криптографические протоколы на эллиптических кривых
студента 4 курса факультета компьютерных наук
и информационных технологий
Платонова Артема Сергеевича
Научный руководитель
доцент, канд. физ.-мат. наук
А.Н. Гамова
Зав. кафедрой
профессор, канд. физ.-мат. наук
В.Н. Салий
Саратов 2012
Содержание

Введение

Средства и системы криптографической защиты информации играют важную роль в современных компьютерных информационных системах, используемых в сфере финансовой и коммерческой деятельности. Интерес к ним обусловлен не только возрастающими общественными потребностями в переводе экономических и государственно-правовых отношений на «электронную основу», но и сильно расширившимися возможностями передачи, обработки и хранения информации в распределенных вычислительных системах. Применение специальных криптографических протоколов и криптосистем позволяет осуществлять многообразные экономические отношения «дистанционно», исключая необходимость личной встречи участников этих отношений, а также поддерживать при этом должную финансовую и правовую дисциплину. К криптографическим протоколам относят протоколы шифрования, электронной цифровой подписи (ЭЦП), идентификации и протоколы аутентифицированного распределения ключей.

В 1985 году Нил Коблиц и Виктор Миллер независимо предложили использовать в криптографии некоторые алгебраические свойства эллиптических кривых. С этого момента началось бурное развитие нового направления в криптографии, для которого используется термин криптография на эллиптических кривых (Elliptic Curve Cryptography, сокращенно ECC). Криптосистемы с открытым ключом на эллиптических кривых обеспечивают такую же функциональность, как и алгоритм RSA. Однако их криптостойкость основана на другой проблеме, а именно на проблеме дискретного логарифма в группе точек эллиптической кривой (Elliptic Curve Discrete Logarithm Problem, сокращенно ECDLP). В настоящее время лучшие алгоритмы для решения ECDLP имеют экспоненциальное время работы, в отличие от алгоритмов для решения проблемы простого дискретного логарифма и проблемы факторизации целого числа, которые имеют субэкспоненциальное время работы. Это означает, что в системах на эллиптических кривых желаемый уровень безопасности может быть достигнут при значительно меньшей длине ключа, чем, например, в схеме RSA. Например, 160-битный ключ в ECC обеспечивает тот же уровень безопасности, что и 1024-битный ключ в RSA. В этой работе подробно рассматриваются способы и преимущества реализации криптографических протоколов с использованием теории эллиптических кривых и в качестве примера реализован алгоритм цифровой подписи на эллиптических кривых (Elliptic Curve Digital Signature Algorithm, сокращенно ECDSA) на языке Java.

алгоритм точка эллиптическая кривая протокол

Эллиптические кривые

В этом разделе будут изложены основы теории эллиптических кривых, даны основные определения, которые понадобятся в дальнейшем при описании алгоритмов арифметики эллиптических кривых.

Алгебраические кривые

Алгебраической кривой порядка n над полем F называется множество точек (x,y): x,y ?F, удовлетворяющих уравнению F(X,Y) = 0, где F(X,Y) -многочлен степени n с коэффициентами из F. Пары (x,y) ?F2, удовлетворяющие уравнению кривой, называются ее точками.
Точка (x,y) кривой F(X,Y) = 0 называется неособой, если в ней не равны нулю обе частные производные многочлена F(X,Y). Кривая называется неособой, или гладкой, если все ее точки - неособые.
Эллиптической кривой E над полем F называется гладкая кривая, задаваемая уравнением вида
Y2 + a1 XY + a3 Y = X3 + a2 X2 + a4 X + a6 , ai F.(1)
Будем обозначать E(F) множество точек (x,y) F2, удовлетворяющих этому уравнению и содержащее кроме того бесконечно удаленную точку, обозначаемую.
Две кривые E и E' над полем F называются изоморфными, если они переходят друг в друга при допустимой замене координат
X := u2x + r, Y := u3Y + u2sX + t.
В зависимости от характеристики поля F общее уравнение эллиптической кривой может быть упрощено. Далее рассмотрены стандартные формы записи эллиптических кривых для полей характеристики 2, 3 и для полей больших характеристик.
Поля больших характеристик. Если поле F не является полем характеристики 2 или 3, то заменив координаты
(x, y) (,),
можно привести кривую к виду
Y2 = X3 + aX + b, a,b F, char F ? 2, 3(2)

С уравнением (2) эллиптической кривой E можно связать дискриминант
(E) = ?16(4a3 +27b2). (3)
Понятие дискриминанта в общем случае кривой (1) выглядит более громоздко. А именно,
(E) = - b22 b8 - 8b43 - 27b62 + 9b2b4b ,
где
b2 = a12 + 4a2 ,
b4 = 2a4 + a1a3 ,
b6 = a32 + 4a6 ,
b8 = a12a6 + 4a2a6 - a1a3a4 + a2a32 - a42;
Если (E) = 0, то указанный многочлен имеет кратные корни и в точке (x, 0) нарушается условие гладкости кривой. Кривая E является гладкой тогда и только тогда, когда ее дискриминант ненулевой.
Поля характеристики 2. Для полей характеристики 2 следует рассмотреть два случая.
Если a1 ? 0, то заменой
(x, y) (x + , y +),
эллиптическая кривая сводится к виду
Y2 + XY = X3 + a2X2 +a6, ai F(4)

Кривые вида (4) называются несуперсингулярными. Дискриминант несуперсингулярной кривой равен (E) = a6 .
Если a1 = 0, то можно провести замену (x, y) (x + a2, y) и кривая будет иметь вид
Y2 + a3 X = X3 + a4 X +a6 , ai F(5)
Кривые такого вида называются суперсингулярными, и их дискриминант имеет вид (E) = a34.
Поля характеристики 3. Для полей характеристики 3 также возможны две замены.
Если a12 ? -a2 , то заменой
(x, y) (x + , y + a1 x + a1 + a3),
где d2 = a12 + a2 , d4 = a4 - a1 a3 кривая преобразуется к виду
Y2 = X3 + aX2 + b(6)
Такие кривые называются несуперсингулярными и имеют дискриминант, равный (E) = -a3 b.
Если a12 = -a2 , то заменой (x, y) (x, y + a1 x +a3) кривая преобразуется к виду
Y2 = X3 + aX + b(7)
Такие кривые называются суперсингулярными и имеют дискриминант, равный (E) = -a3.

Группа точек эллиптической кривой

На множестве E(F), состоящем из точек эллиптической кривой (1) и еще одного элемента - бесконечно удаленной точки кривой O (формально пока не являющейся точкой кривой), можно определить операцию, обладающую свойствами операции абелевой группы.
Принято получающуюся при этом группу рассматривать как аддитивную группу, а операцию называть операцией сложения и обозначать, как обычно, знаком плюс.
Упомянутая дополнительная точка O играет роль нейтрального элемента (в аддитивной записи нуля) этой группы.
По определению, полагаем для любой точки (x,y) E(F)
(x,y) + O = O + (x,y) = (x,y);
O + O = O;

Чтобы определить в общем случае операцию сложения абелевой группы, сначала покажем, что каждой точке (x,y) эллиптической кривой можно сопоставить в определенном смысле симметричную точку (далее будет ясно, что такая точка и будет точкой -(x,y), противоположной к (x,y) точкой в группе данной кривой).
Заметим, что вместе с точкой (x,y) кривая имеет и точку

(x,y') = (x, -a1x - a3 - y);
Убедиться в этом можно, подставив X = x и Y = -a1x - a3 - y, и учитывая, что при X = x и Y = y имеет место равенство. Симметричность проявляется в том, что по тому же правилу точке (x, y') соответствует исходная точка (x, y), так как имеет место инволютивный закон:

(x, y) = (x, y'').
Если кривая E имеет вид (2), то
(x, y') = (x, -y)

В частности, для эллиптических кривых над полем действительных чисел, точки (x, y) и (x, -y) располагаются на прямой Y = x симметрично относительно оси абсцисс, как показано на рисунке.
Для суперсингулярных и несуперсингулярных кривых характеристики 2 симметричная точка (x, y') определяется соответственно уравнениями (x, y') = (x, y+1) и (x, y') = (x, x+y).
Полагается, что (x, y) + (x, y') = O и (x, y') обозначается -(x, y). Таким образом, множество E(F) удовлетворяет двум аксиомам группы (существует нулевой элемент и каждому элементу соответствует противоположный элемент).
Таким образом, операция сложения определена, когда одна из точек равна O или когда складываются противоположные точки.
Для двух точек (x1, y1), (x2, y2), таких, что x1 ? x2 или x1 = x2, y1 = y2 суммой двух этих точек объявляется точка
P + Q = -R = -(x3, y3) (в случае x1 ? x2)
P + P = 2P = -R = -(x3, y3) (в случае x1 = x2, y1 = y2)
Конкретные формулы для вычисления координат точки R в обоих случаях рассмотрены в разделе «Алгоритмы на эллиптических кривых». На рисунках изображено положение точки R для обоих случаев при рассмотрении эллиптической кривой над полем действительных чисел.
Логично было бы назвать результатом операции сложения саму точку R, но тогда не будет выполняться тождество
P + Q = R P = R - Q.
Операция сложения на множестве E(F) коммутативна и ассоциативна (это можно доказать, используя прямые формулы для вычисления R). Таким образом, множество E(F) (множество точек эллиптической кривой вместе с точкой O) с операцией сложения, описанной выше, является абелевой группой.
Порядком точки P эллиптической кривой E называется минимальное натуральное число n, такое, что nP = O. Если такого числа не существует, то точка имеет бесконечный порядок.
Эллиптические кривые над конечными полями

Эллиптические кривые над конечными полями имеют конечные группы точек. Порядок этой группы называется порядком эллиптической кривой. По теореме Лагранжа порядок точки делит порядок эллиптической кривой. Изоморфные кривые имеют одинаковые группы, а, следовательно, и порядки. Поэтому далее всегда можно ограничиться рассмотрением кривых с уравнениями специального вида (2), (4), (5), (6), (7).
Пользуясь символом Лежандра, легко указать формулу для числа точек на кривой Y2 = f(X) над полем GF(p), p > 2 (поля больших характеристик). Действительно,

сравнение Y2 = f(X) (mod p) относительно Y при фиксированном X имеет (при p > 2) 1 + решений (это верно и при f(x) = 0). Учитывая бесконечно удаленную точку, получаем формулу для порядка кривой над полем GF(p), p > 2 в виде

При малых простых p, пользуясь этой формулой и теорией квадратичных вычетов порядок кривой над полем GF(p) находится довольно легко. Но вычисление порядка эллиптической кривой не всегда просто и даже возможно. Общая формула для вычисления порядка произвольной кривой неизвестна. Неизвестно даже, можно ли за полиномиальное время найти кривую данного порядка. Тем не менее, известны способы выбора эллиптических кривых над конечными полями, допускающих простое определение порядка. Эти способы важны, потому что в криптографическом отношении полезными являются эллиптические кривые, порядок которых содержит большие простые множители. Для кривых, у которых порядок является гладким числом (т.е. разлагающимся только на малые простые) проблема дискретного логарифмирования может быть решена сравнительно быстро алгоритмом Полига-Хеллмана-Зильбера.
Алгоритмы на эллиптических кривых
В этом разделе представлены алгоритмы, необходимые для реализации криптографических приложений на эллиптических кривых.
Диаграмма, приведенная ниже, показывает, какие модули необходимо создать при реализации алгоритма цифровой подписи на эллиптических кривых (ECDSA). В приложении данной работы генерация случайных чисел, модульная арифметика и операции над большими числами осуществляются стандартными средствами языка Java. Арифметика эллиптических кривых основана на алгоритмах, приведенных в этой главе.

Сложение точек эллиптической кривой

В соответствии с определением операции сложения в группе точек эллиптической кривой общая схема алгоритма сложения точек P1 = (x1 , y1) и P2 = (x2 , y2) выглядит следующим образом:
Вход: коэффициенты эллиптической кривой, точки P1 и P2.
Выход: R = P1 + P2.
Алгоритм: если P1 = O, то R = P2 ,
если P2 = O, то R = P1 ,
если P2 = - P1, то R = O ,
если x2 ? x1, то R = P1 + P2 = -(x3 , y3) ,
иначе R = 2P1 = -(x3 , y3).
Вернуть: R.
Координаты x3 , y3 вычисляются по разным формулам в зависимости от вида эллиптической кривой и условия различия или совпадения точек.
Для эллиптических кривых над полем характеристики, большей 3 (т.е. для кривых,

имеющих вид Y2 = X3 + aX + b) противоположной точкой для точки P (x, y) будет являться -P = P (x , -y). Если P1 ? P2, то формулы для вычисления координат R выглядят так:
x3 = 2 -x1 - x2 ,
y3 = y1 + (x3 - x1) , где =
В случае P1 = P2 = (x, y) формулы имеют следующий вид:
x3 = (')2 - 2x,
 y3 = y + '(x3 - x) , где ' =

Для полей характеристики три (в общем виде Y2 = X3 + a2X2 + a4X + a6) при P1 ? P2 формулы имеют вид:
x3 = (2 -a2) -x1 - x2 ,
y3 = y1 + (x3 -x1), где =
А при P1 = P2

x3 = (')2 -a2) - 2x ,
y3 = y + '(x3 -x), где ' =

Для полей характеристики два случаи суперсингулярных и несуперсингулярных кривых рассматриваются отдельно. Точка кривой, противоположная точке (x,y) имеет координаты (x,x+y). Для несуперсингулярных кривых (в общем виде Y2 + XY = X3 + a2X2 +a6) при P1 ? P2 координаты R вычисляются по формулам:
x3 = 2 + + a2 + x1 + x2 ,
y3 = x3 + y1 + (x3 + x1), где =
А при P1 = P2
x3 = (')2 + (') + a2,
y3 = x2 + (' + 1)x3, где ' =
Для суперсингулярных кривых (в общем виде Y2 + a3 X = X3 + a4 X +a6) противоположной точкой для (x,y) будет (x,y + a3). При P1 ? P2

x3 = + x1 + x2
y3 = a3 + y1 + (x3 + x1), где =
А при P1 = P2
x3 = (')2
y3 ='(x + x3) + y + a3 , где ' =
Следует отметить, что при вычислении суммы двух точек описанными выше формулами, самая трудоемкая операция в арифметике конечного поля - мультипликативное обращение, выполняется однократно.
Реализация этого алгоритма для кривых характеристики, большей 3, находится в приложении 1 данной работы (метод pointAdd класса eCurve).

Алгоритм скалярного умножения

Алгоритмы умножения точки P эллиптической кривой на числовую константу k (кратко - алгоритмы вычисления kP), они же алгоритмы скалярного умножения точки, являются основными в арифметике эллиптических кривых. Существует большое число алгоритмов, разработанных для кривых специального вида (в том числе для конкретных кривых, рекомендованных стандартами шифрования и цифровой подписи), а также существуют алгоритмы вычисления kP при известном P (например, P может быть известно заранее в алгоритме цифровой подписи).
В приложении данной работы для вычисления скалярного произведения был использован аналог алгоритма быстрого возведения в степень для эллиптических кривых.
Вход: k = (kt?1, . . ., k1, k0)2, P Ґё E(Fq).
Выход: kP.
Алгоритм: 1. Q<O.
2. Для всех i от 0 до t ?1:
2.1 Если ki = 1 то Q<Q + P.
2.2 P<2P.
Вернуть: Q.
Число k представляется в двоичной форме записи. Точка Q на первом шаге равняется O, т.е. является нейтральным элементом сложения. Далее запускается цикл с количеством шагов, равным длине k в двоичной форме записи. На каждом шаге если ki = 1, то Q складывается с точкой P. В конце каждого шага P удваивается. Алгоритм состоит только из операций сложения двух точек и операций удвоения точки. Поэтому вычислительное время алгоритма зависит только от времени вычисления суммы двух точек и от времени удвоения точки.
Если точка P известна заранее, то можно произвести некоторые предварительные вычисления для повышения эффективности алгоритма. Например, вычислив все точки 2P, 22P … 2t-1P можно значительно ускорить вычисление скалярного произведения. Вычислительное время алгоритма будет зависеть только от времени вычисления суммы двух точек.
Алгоритм генерации случайных кривых
Алгоритм цифровой подписи с использованием эллиптических кривых (ECDSA) принят и описан в различных стандартах. Среди них ANSI X9.62, FIPS 186-2 (NIST), IEEE 1363-2000, ISO/IEC 14888-3, ISO/IEC 15946-3, SEC-1, SEC-2 и др.
Далее мы опишем основные рекомендации стандарта ANSI X9.62 ECDSA. К эллиптическим кривым предъявляются следующие требования:
1. Кривые рассматриваются или над простыми полями (порядок q которых равен простому числу p), или над полями характеристики два (у которых q = 2m).
2. Для представления элементов поля используется либо стандартный базис, порождаемый трехчленом или пятичленом, либо гауссов нормальный базис (GNB).
3. Кривая E задается выбором двух элементов a, b поля GF(q). В случае p > 2 она имеет вид y2 = x3+ ax + b, а в случае p = 2 вид y2+ xy = x3+ ax2+ b. Таким образом, стандарт рекомендует только несуперсингулярные кривые.
4. На кривой выбирается точка (xG, yG), xG, yG ? GF(q) простого порядка n > 2160, n > 4

и вычисляется кофактор h = |E(GF (q))|/n. В качестве кривых можно и удобно выбирать в случае p = 2 кривые, у которых a, b равны 0, 1, но стандарт рекомендует все же случайные кривые, т.е. кривые со случайно выбранными a, b.
При этом рекомендуется использовать следующие алгоритмы генерации случайных кривых:
1) Случай q = p. Положим
t = log2 p , s = (t ? 1)/160, v = t ? 160s.
1. Выбираем произвольную строчку битов seedE длиной g ? 160 бит, и полагаем z равным числу, двоичная запись которого совпадает с seedE.
2. Применяя к seedE стандартную хеш-функцию SHA1, вычисляем g-битовую строку H = SHA1(seedE). Выбирая в H v самых правых битов, получаем строку c0 длиной v битов.
3. Заменяя в c0 самый левый бит на 0, получаем строку W0.
4. Для i от 1 до s делаем следующее:
4.1 полагаем si равной g?битной строке, являющейся двоичной записью числа z + i mod 2g
4.2 вычисляем g-битовую строку Wi= SHA1(si).
5. Полагаем битовую строку W равной конкатенации (произведению) битовых строк Wi , i = 0, . . . , s, т.е. W = W0. . . Ws.
6. Полагаем r равным целому числу с двоичной записью W. Выполнение пункта 3 гарантирует, что r < p.
7. Если r = 0 или 4r + 27 ? 0(mod p), то возвращаемся к шагу 1.
8. Выбираем ненулевые a, b Ґё GF (p) так, чтобы rb2 ? a3(mod p). Например, можно взять a = b = r.
9. Полученная кривая есть E : y2 = x3 + ax + b.
Заметим, что условие невырожденности кривой 4a3 + 27b2 6 ? 0 (mod p) гарантировано выполняется, так как при b ? 0, r = a3/b2mod p удовлетворяет условиям r ? 0, 4r + 27 ? 0 (mod p). Имеется только две попарно неизоморфные кривые с одним и тем же r. Эти кривые являются скрученными и сумма их порядков равна 2 + 2p. Кривые с разными r неизоморфны друг другу. На шаге 8 поэтому есть по существу только еще одна возможность выбора a и b, кроме явно указанной.
2) Случай q = 2m. Положим, как и выше, s = (t ? 1)/160, v = t ? 160s.
1. Выбираем произвольную строчку битов seedE длиной g ? 160 бит, и полагаем z, равным числу, двоичная запись которого совпадает с seedE.
2. Вычисляем g-битовую строку H = SHA1(seedE). Выбирая в H v самых правых битов, получаем строку b0 длиной v битов.
3. Заменяя в b0 самый левый бит на 0, получаем строку W0.
4. Для i от 1 до s делаем следующее:
4.1 полагаем si равной g?битной строке, являющейся двоичной записью числа z + i mod 2g
4.2 вычисляем g-битовую строку bi = SHA1(si).
5. Вычисляем битовую строку b = b0. . . bs и полагаем b равным соответствующему элементу поля GF (q).

6. Если b = 0, то возвращаемся к шагу 1.
7. Выбираем произвольный a Ґё GF (q).
8. Полученная кривая есть E: y2 + xy = x3 + ax2 + b.
Генерация криптографически надежных параметров кривых.
Стандартом рекомендуется определенный алгоритм генерации надежных параметров кривых.
1. Выбираем случайную кривую E(GF(q)) алгоритмом, указанным выше.
2. Вычисляем ее порядок N = |E(GF (q))|.
3. Проверяем, делится ли N на ранее выбранное простое n (n > 2160, n > 4). Если нет, то переходим к шагу 1.
4. Проверяем, что n не делит ни одно из чисел qk ? 1, k = 1, . . . 20. Если нет, то переходим к шагу 1.
5. Проверяем, что n ? q. Если нет, то переходим к шагу 1.
6. Выбираем произвольную точку G0 Ґё E(GF(q)) и полагаем G = (N/n) G0. Повторяем, пока не получим G ? O.
Генерация случайных кривых с подходящими криптографическими свойствами - чрезвычайно ресурсоемкий процесс. Кривую над полем GF(2m) при m примерно равным 200 можно сгенерировать за несколько часов. В некоторые стандарты ECDSA включен набор сгенерированных эллиптических кривых со специальными параметрами, повышающими эффективность операций с точками этих кривых.
Стандартом NIST к использованию рекомендуются кривые P-192, P-224, P-256, P-384, P-521 над полями больших характеристик (рассматриваются кривые вида y2 =x3-3x+b, то есть a=-3). Для полей характеристики два для каждого поля рекомендованы две эллиптических кривых - несуперсингулярная кривая (вида y2 + xy = x3 + x2 + b) и кривая Коблица (кривые вида y2 + xy = x3 + x2 + 1, где a = 0,1). Вот, к примеру, рекомендованная стандартом NIST кривая для поля большой характеристики P-192.
Curve P-192
p = 6277101735386680763835789423207666416083908700390324961279
r = 6277101735386680763835789423176059013767194773182842284081
s = 3045ae6f c8422f64 ed579528 d38120ea e12196d5
c = 3099d2bb bfcb2538 542dcd5f b078b6ef 5f3d6fe2 c745de65
b = 64210519 e59c80e7 0fa7e9ab 72243049 feb8deec c146b9b1
Gx = 188da80e b03090f6 7cbf20eb 43a18800 f4ff0afd 82ff1012
Gy = 07192b95 ffc8da78 631011ed 6b24cdd5 73f977a1 1e794811
Для кривой заданы:
Простое целое p - характеристика поля.
Порядок кривой r
Входное значение для хэш-функции SHA1 s = seedE
Выходное значение хэш-функции SHA1 c.
Коэффициент b
Координаты порождающей точки G (Gx, Gy)
Использование различных систем координат

Как видно из раздела «Сложение точек эллиптической кривой», при сложении двух точек эллиптической кривой над полем характеристики, большей 3 (т.е. для кривой, имеющих вид y2 = x3 + ax + b) требуется произвести два умножения, одно возведение в квадрат и одно обращение. Переход к другой системе координат позволяет полностью исключить операцию обращения за счет увеличения числа других операций. Поэтому если для данного поля операция обращения занимает значительно больше времени, чем операция умножения, то использование другой системы координат может значительно ускорить вычисления. В этом разделе будут рассмотрены наиболее часто используемые системы координат для кривых над полями больших характеристик (y2 = x3 + ax + b):

· стандартная проективная система координат
· система координат Якоби
· система координат Чудновского-Якоби
· модифицированная система координат Якоби

Для каждой системы координат будет получен алгоритм вычисления суммы двух точек эллиптической кривой. Это позволит определить, сколько в точности операций нужно затратить на сложение двух точек в каждой из рассмотренных систем.

Также будет рассмотрен метод, позволяющий использовать в одном алгоритме несколько различных систем координат. Будет приведен пример алгоритма, использующего преимущества данного подхода.

Проективные координаты

Рассмотрим эллиптическую кривую над полем K, и положим c и d положительными числами. Определим отношение эквивалентности на ненулевых тройках из K3 следующим образом:

(X1, Y1, Z1) ~ (X2, Y2, Z2) если X1 = c X2 , Y1= d Y2 , Z1 = Z2 для некоторого ненулевого K.

Класс эквивалентности, содержащий тройку (X, Y, Z) K3 \ {(0, 0, 0)}, обозначим следующим образом:

(X : Y : Z) = {(c X, d Y, Z): K }.

Класс эквивалентности (X : Y : Z) называется проективной точкой, а (X, Y, Z) - ее представителем. Множество всех проективных точек обозначается P(K) . Следует отметить, что если (X', Y', Z') (X : Y : Z), то (X' : Y' : Z') = (X : Y : Z). Таким образом, каждый элемент класса эквивалентности может служить его представителем. В частности, если положить Z ? 0, то (X / Zc , Y / Zd , 1) является представителем проективной точки (X : Y : Z), причем единственным представителем с координатой Z = 1. Таким образом, мы получаем однозначное соответствие между набором проективных точек

P(K) = {(X : Y : Z) : X, Y, Z K, Z ? 0}

и набором аффинных точек

A(K) = {(x, y) : x, y K}.

Набор проективных точек

P(K)0 = {(X : Y : Z) : X, Y, Z Ґё K, Z = 0}
называется прямой в бесконечности, так как точки из этого набора не соответствуют никаким аффинным точкам.
Проективная форма уравнения эллиптической кривой E над полем K получается путем подстановки x = X / Zc , y = Y / Zd и избавления от знаменателей (то есть путем домножения на Z в некоторой степени). Если некоторый представитель класса (X, Y, Z) K3 \ {(0, 0, 0)} удовлетворяет полученному проективному уравнению, то ему удовлетворяет любой представитель класса (X:Y:Z). Таким образом можно сказать, что проективная точка (X : Y : Z) лежит на кривой E. Проективные точки из P(K)0, лежащие на кривой E, будут являться бесконечно удаленными точками.
Для удобства изложения дальнейшего материала, введем некоторые обозначения. Будем обозначать рассматриваемые системы координат прописными буквами латинского алфавита, с возможным добавлением верхнего индекса (например, аффинную систему координат будем обозначать A). Дадим также обозначение основным операциям, производящимся при сложении двух точек эллиптической кривой:
M - умножение двух чисел,
S - возведение числа в квадрат,
I - обращение числа.
Далее, число операций, требуемых для сложения двух точек (удвоения точки) в заданной системе координат будем обозначать следующим образом:
t(X1 + X2 = X3) = nM + kS + lI, где n - число требуемых умножений, k - число требуемых возведений в квадрат, l - число требуемы обращений, X1 и X2 - используемые системы координат для первой и второй точки соответственно, X3 - система координат результирующей точки (если X1 = X2 = X3 , то сократим запись до t(X + X)).
Например, в описанных обозначениях для аффинных координат можно записать:
t(A + A) = 2M + S + I, t(2A) = 2M + 2S + I.

Стандартная проективная система координат (P)
В стандартной проективной системе координат c = d = 1. В этом случае уравнение кривой y2 = x3 + ax + b преобразуется к виду:
Y 2Z = X 3 + aXZ 2 + bZ 3
Бесконечно удаленная точка имеет координаты (0, 1, 0). Обратной точкой для (X, Y, Z) будет точка (X, -Y, Z).
Возьмем две проективные точки P1 = (X1, Y1, Z1), P2 = (X2, Y2, Z2), принадлежащие нашей кривой. Тогда координаты точки P3:
P3 = P1 + P2 = (X3, Y3, Z3)
вычисляются по следующим формулам (эти формулы легко получаются при подстановке замен x, y в стандартные формулы сложения двух точек):
Если P1 ? ±P2 ,
u = Y2Z1 ? Y1Z2,
v = X2Z1 ? X1Z2,
w = u2Z1Z2 ? v3 ? 2v2X1Z2,

X3 = vw,
Y3 = u(v2X1Z2 ? w) ? v3Y1Z2,
Z3 = v3Z1Z2.
Введем несколько дополнительных переменных с целью уменьшения числа операций:
t1= Y1Z2 ; t2 = X1Z2 ;
t3 = Z1Z2 ; v2 = v2 ;
v3 = v3 = v2v ; t4 = v2t2 ;
Подставляя эти переменные в формулы для сложения, получим, что t(P+P) = 12M + 2S.
Рассмотрим теперь случай P1 = P2 (удвоение точки):
t = aZ12 + 3X12 ,
u = Y1Z1 ,
v = uX1Y1 ,
w = t2 ? 8v,
X3 = 2uw,
Y3 = t(4v ? w) ? 8Y12u2 ,
Z3 = 8u3.
Получаем, что t(2P) = 7M + 5S.
При P1 = ?P2 получаем P3 = O.
Теперь рассмотрим, какой выигрыш в производительности мы получим при использовании стандартных проективных координат. В последующих рассуждениях учитываются только операции умножения, возведения в квадрат и обращения. Операции сложения и вычитания производятся намного быстрее и при оценке скорости алгоритма их можно не учитывать.
Напомним, что при использовании обычного метода сложения двух точек нам требуется произвести 2 умножения, одно возведение в квадрат и одно обращение. Таким образом, при использовании стандартных проективных координат мы производим на 10 умножений и на 1 возведение в квадрат больше, но нам не требуется операция обращения элемента.
При удвоении точки в проективных координатах нам требуется провести на 5 умножений и на 3 возведения в квадрат больше. Обращение из алгоритма также исключается.
Таким образом, для того, чтобы использование проективных координат давало прирост производительности, необходимо, чтобы операция обращения была хотя бы в 11 раз медленнее операции умножениях в случае сложения двух точек и в 8 раз медленнее в случае удвоения точки. Для сравнения скорости работы этих операций была написана отдельная программа (test.java в приложении данной работы, метод compMultInv). Эта программа использует реализации операций умножения и обращения из библиотеки BigInteger языка java (именно эта библиотека используется в данной работе при реализации ECDSA). В результате работы данной программы было установлено, что обращение элемента поля K в среднем занимает приблизительно в 14-16 раз больше времени, чем умножение двух элементов

данного поля (характеристика поля K выбиралась случайным образом, ее длина полагалось равной 200 битам). Из этих результатов можно сделать следующий вывод: в заданных условиях при использовании стандартных проективных координат сложение двух точек происходит в среднем в ~1.28 раз быстрее, а удвоение точки - в ~1.58 раз быстрее.
Следует отметить, что при использовании стандартных проективных координат на суперсингулярных кривых можно получить еще больший прирост производительности, так как для сложения двух точек там требуется лишь 9 операций умножения, а для удвоения точки - только 6 возведений в квадрат. Таким образом, переходя к проективным координатам, можно полностью избавиться от операции обращения, увеличив число умножений всего в 4.5 раза.
В результате получаем, что в реализации криптографических протоколов все операции над точками эллиптической кривой можно проводить в проективных координатах. Когда нужно будет совершить переход обратно к аффинным координатам, достаточно разделить координаты X, Y проективной точки на Z.

Система координат Якоби (J)

Рассмотрим теперь систему координат Якоби. В этой системе координат c = 2, d = 3, то есть (X : Y : Z) отображает аффинную точку (X / Z2, Y / Z3). Эллиптическая кривая y2 = x3 + ax + b преобразуется к виду:
Y 2 = X 3 + aXZ 4 + bZ 6.
Бесконечно удаленная точка имеет координаты (1 : 1 : 0). Обратной точкой для (X, Y, Z) будет точка (X, -Y, Z).
Возьмем две проективные точки P1 = (X1, Y1, Z1), P2 = (X2, Y2, Z2), принадлежащие нашей кривой. Тогда координаты точки P3:
P3 = P1 + P2 = (X3, Y3, Z3)
вычисляются по следующим формулам:
Если P1 ? ±P2 ,
r = X1Z22, s = X2Z12,
t = Y1Z23, u = Y2Z13,
v = s - r, w = u - t,
X3 = -v3 - 2rv2 + w2,
Y3 = -tv3 + (rv2 - X3)w,
Z3 = vZ1Z2.
Если P1 = P2 (удвоение точки):
v = 4X1Y12, w = 3X12 + aZ14,
X3 = -2v + w2,
Y3 = -8Y14 + (v - X3)w,
Z3 = 2Y1Z1.
При P1 = ?P2 получаем P3 = O.
В этой системе координат t(J + J) = 12M + 4S. Удвоение точки требует t(2J) = 3M + 6S. Операция обращения, как и в случае стандартных проективных координат, не

задействована. Следовательно, сложение точек в системе координат Якоби происходит чуть медленнее, чем в стандартной проективной системе координат. Однако в сравнении со стандартной проективной системой координат, удвоение точки требует на 4 умножения меньше и только на одно возведение в квадрат больше.

При использовании эллиптических кривых особого вида можно дополнительно ускорить процесс удвоения точки. Если в уравнении кривой a = -3, то

w = 3X12 + aZ14 = 3(X12 - Z14) = 3(X1 - Z12)(X1 + Z12).

В этом случае t(2J) = 4M + 4S. Именно по этой причине стандартом NIST над полями больших характеристик все рекомендуемые кривые имеют вид y2 = x3 - 3x + b.

Полный алгоритм удвоения точки в системе координат Якоби для случая a = -3 выглядит следующим образом:

Вход: точка P = (X1 : Y1 : Z1) в системе координат Якоби лежащая на кривой y2 = x3 - 3x + b.

Выход: 2P = (X3 : Y3 : Z3) в системе координат Якоби.