Designs, Codes and Cryptography, 20, 41–64 (2000)

# Analysis of the Xedni Calculus Attack

MICHAEL J. JACOBSON mjjacobs@cacr.math.uwaterloo.ca

NEAL KOBLITZ koblitz@math.washington.edu

JOSEPH H. SILVERMAN jhs@math.brown.edu

ANDREAS STEIN astein@cacr.math.uwaterloo.ca

EDLYN TESKE eteske@cacr.math.uwaterloo.ca
*Centre for Applied Cryptographic Research, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

**Abstract.** The xedni calculus attack on the elliptic curve discrete logarithm problem (ECDLP) involves lifting points from the finite field $\mathbb{F}_p$ to the rational numbers $\mathbb{Q}$ and then constructing an elliptic curve over $\mathbb{Q}$ that passes through them. If the lifted points are linearly dependent, then the ECDLP is solved. Our purpose is to analyze the practicality of this algorithm. We find that asymptotically the algorithm is virtually certain to fail, because of an absolute bound on the size of the coefficients of a relation satisfied by the lifted points. Moreover, even for smaller values of $p$ experiments show that the odds against finding a suitable lifting are prohibitively high.

**Keywords:** elliptic curve, discrete logarithm, Xedni calculus

## 1. Introduction

At the Second Elliptic Curve Cryptography Workshop (University of Waterloo, September 14–16, 1998), Joseph Silverman announced a new attack on the elliptic curve discrete logarithm problem (ECDLP) over a prime field $\mathbb{F}_p$. He called his method “xedni calculus” because it “stands index calculus on its head.”[1]

Recall that the ECDLP is the problem given two points $P, Q$ on an elliptic curve over $\mathbb{F}_p$, of finding an integer $w$ such that $Q = wP$. Very briefly, Silverman’s idea was to take $r$ random linear combinations of the two points $P, Q$, where $2 \le r \le 9$ (most likely $r = 4$, 5 or 6), and then consider points $P_i$ with rational coordinates that reduce modulo $p$ to these $r$ points and elliptic curves $E$ over the rational number field $\mathbb{Q}$ that pass through all of the $P_i$ and reduce mod $p$ to the original curve over $\mathbb{F}_p$. If those “lifted” points $P_i$ are linearly dependent, then the ECDLP is solved. The probability of dependence is almost certainly very low, but Silverman had an idea of how to increase this probability, possibly by a dramatic amount. Namely, he imposes on the $P_i$ and $E$ a set of auxiliary conditions modulo $l$ for several small primes $l$. These conditions guarantee that the elliptic curves will have fewer-than-expected points modulo $l$, and this presumably decreases the likelihood that the $r$ $\mathbb{Q}$-points $P_i$ will be independent. (More details will be given in §3 below.)

Silverman's algorithm, which had been circulating in manuscript form for about two weeks before the conference, created a stir for several reasons. In the first place, this was the first time in about seven years that a serious attack had been proposed on an important class of elliptic curve cryptosystems. In the second place, Silverman's approach involved some sophisticated ideas of arithmetic algebraic geometry—most notably, the heuristics of the Birch–Swinnerton-Dyer Conjecture—that had never before had any practical application. In the third place, because of the subtlety of the mathematics being used, even people who had computational experience with elliptic curves were completely baffled in their initial attempts to estimate the running time of the xedni calculus. No one, for example, could say with absolute certainty that it would not turn out to give a polynomial-time algorithm!

If it were practical, the xedni calculus would not only break elliptic curve cryptosystems (ECC). As Koblitz showed, it can easily be modified to attack (1) the Digital Signature Standard (i.e., the discrete logarithm problem in the multiplicative group of $\mathbb{F}_p$), and (2) RSA (i.e., the integer factorization problem). Thus, essentially all public-key cryptography that's in widespread use in the real world was threatened.

Of course, most people, including Silverman himself, thought that it was highly unlikely that the algorithm would turn out to be so efficient that it would render ECC, DSS, and RSA insecure. However, it is not enough to have a "gut feeling" about such matters. One needs to find solid mathematical arguments that enable one to evaluate the efficiency of the xedni calculus. That is the purpose of this paper.

## 2. Background

### *2.1. The Hasse–Weil L-Function*

Let $E$ be an elliptic curve defined over the field $\mathbb{Q}$ of rational numbers, and let $N_l = l+1-a_l$ denote the number of points on the reduction of E modulo $l$.[2] For each $l$ we have the associated quadratic polynomial $1 - a_l T + lT^2 = (1 - \alpha_l T)(1 - \overline{\alpha}_l T)$ whose value at $T = 1$ is $N_l$; this polynomial is the numerator of the zeta-function of $E \bmod l$. By Hasse's Theorem, $\alpha_l$ is a complex number of absolute value $\sqrt{l}$.

The Hasse–Weil $L$-function of the curve $E$ is defined by analogy with the Riemann zeta-function $\zeta(s) = \Pi_{\text{primes } l} \frac{1}{1-l^{-s}}$. Namely, we take $L(E, s)$ to be the product over $l$ of the following "Euler factor":

$$\frac{1}{(1 - \alpha_l \cdot l^{-s})(1 - \overline{\alpha}_l \cdot l^{-s})} = \frac{1}{1 - a_l \cdot l^{-s} + l \cdot l^{-2s}}.$$

It is easy to verify that the infinite product converges for $\mathrm{Re}(s) > 3/2$ (just as the Euler product for the Riemann zeta-function converges for $\mathrm{Re}(s) > 1$). By the "critical value" we mean the value of $L(E, s)$ at $s = 1$. Just as one has to analytically continue the Riemann zeta-function a distance $1/2$ to the left in order to react the "critical line," similarly one has to analytically continue $L(E, s)$ a distance $1/2$ to the left in order to reach the critical value.

However, analytic continuation of $L(E, s)$ is not nearly so simple as in the case of $\zeta(s)$; and, in fact, it has been proven only in the case when $E$ is "modular" in the following sense. If we expand the Euler product, we can write $L(E, s)$ in the form $\sum a_n \cdot n^{-s}$.[3] We now introduce a new complex variable $z$, and in each term we replace $n^{-s}$ by $e^{2\pi i n z}$. The result is a Fourier series $\sum a_n e^{2\pi i n z}$ that converges in the complex upper half-plane. We say that $E$ is "modular" if this Fourier series is a modular form, that is, if it satisfies a simple transformation rule when $z$ is replaced by $(az + b)/(cz + d)$ for any integer matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ of determinant 1 with $c \equiv 0 \pmod{N}$. Here $N$ is the "conductor" of the curve $E$; it is closely related to the curve's discriminant $D$.[4] In order to know unconditionally that analytic continuation is possible and the critical value $L(E, 1)$ is defined, we need the curve $E$ to be modular.

### 2.2. The Taniyama Conjecture

The Taniyama Conjecture is the assertion that all elliptic curves $E$ over $\mathbb{Q}$ are modular. One reason for its importance is that it guarantees that the Hasse–Weil $L$-function of $E$ can be analytically continued, and its behaviour near $s = 1$ can be studied.

It is for a different reason that most people have heard of this conjecture, namely, its connection to Fermat's Last Theorem. In 1985 Gerhard Frey suggested that if $A^p + B^p = C^p$ were a counterexample to Fermat's Last Theorem, then the elliptic curve

$$Y^2 = X(X - A^p)(X + B^p)$$

would have a very surprising property. Its discriminant would be

$$16(A^p B^p (A^p + B^p))^2 = 16(ABD)^{2p},$$

so every prime factor in this discriminant would occur to a very large power. Frey thought that it would then have to violate the Taniyama Conjecture. K. Ribet was able to prove that Frey's hunch was correct [24]; then, working intensively for many year, A. Wiles (partly in joint work with R. Taylor) [37, 36] proved that no such curve can violate the Taniyama Conjecture, and hence there can be no counterexample to Fermat's Last Theorem.

Wile proved the Taniyama Conjecture for a broad class of curves—the "semi-stable" ones, i.e., the ones whose conductor $N$ is squarefree—but not for all curves. What he proved was enough for Fermat's Last Theorem. The full conjecture was subsequently proven by Breuil, Conrad, Diamond, and Taylor.

### 2.3. The Conjecture of Birch and Swinnerton-Dyer

As before, let $E$ be an elliptic curve defined over $\mathbb{Q}$, and let $N_l$ denote the number of mod-$l$ points. As $l$ increases, suppose that we want to get an idea of whether or not $N_l$ tends to be toward the right end of the Hasse interval $[l+1-2\sqrt{l}, l+1+2\sqrt{l}]$, that is, whether or not there tend to be more-than average points on the curve. We might expect that if our original curve over $\mathbb{Q}$ has infinitely many points—that is, if its rank $r$ is positive—then these rational points would be a plentiful source of mod-$l$ points, and $N_l$ would tend to be large; whereas if $r = 0$, then $N_l$ would straddle both sides of $l+1$ equally. This is the intuitive idea of the (weak) Birch–Swinnerton-Dyer Conjecture [1, 2, 3].

To measure the relative size of $N_l$ and $l$ as $l$ varies, let us form the product $\prod_l \frac{l}{N_l}$. Because $N_l = l - \alpha_l + 1$, we can write this as

$$\prod_l \frac{l}{l-\alpha_l+1} = \prod_l \frac{1}{1-\alpha_l \cdot l^{-1} + l \cdot l^{-2}},$$

which is formally equal to the value at $s = 1$ of the Euler product for $L(E, s)$. We say "formally," because that product diverges, and the critical value is found by analytic continuation, not by evaluating an infinite product.

Nevertheless, let us suppose that it makes sense to talk about this infinite product as if it converged. One might expect that it would converge to zero if $N_l$ has a tendency to be significantly larger than $l$, and would converge to a nonzero value if $N_l$ is equally likely to be above or below $l$. And, indeed, the Birch–Swinnerton-Dyer Conjecture states that $L(E, s)$ vanishes at $s = 1$ if and only if the rank $r$ of the group of $E$ over $\mathbb{Q}$ is greater than zero, and that, moreover, its order of vanishing at $s = 1$ is equal to $r$. The conjecture further says that the leading coefficient in the Taylor expansion at $s = 1$ can be expressed in terms of certain number-theoretic invariants of $E$. Starting in 1977, a series of important partial results have been proved in support of this fundamental conjecture (see [5, 6, 25]), but in its most general form it remains a very difficult unsolved problem.

### *2.4. Heights*

Let $E$ be an elliptic curve (in Weierstrass form) over the field $\mathbb{Q}$ of rational numbers. Let $P = (x, y)$ be a rational point on $E$ (not the point at infinity). The *logarithmic height* of $P$ is defined by the formula $h(P) = \log \max(|a|, |b|)$, where $x = a/b$ is written as a fraction in lowest terms. The logarithmic height is closely related to the point's size in the computer-science sense (i.e., its bit-length).[5]

It can be shown that, if $P$ is a point of infinite order, then $h(nP)$ grows *quadratically* with $n$. That is, if you write out a list of the multiples of $P$, one on each line, the lengths of the lines will increase proportionately to $n^2$ and so form a parabola. (For a picture of this in the case of the elliptic curve $Y^2 + Y = X^3 - X$ and the point $P = (0, 0)$, see page 143 of [11].)

The logarithmic height, which has a roughly quadratic behavior, can be modified (this was done by Néron [23] and later simplified by Tate) in such a way that the resulting *canonical logarithmic height* $\hat{h}(P)$ is precisely a quadratic form. Namely, define

$$\hat{h}(P) = \frac{1}{2} \lim_{n\to\infty} \frac{1}{n^2} h(nP).$$

The values of $\hat{h}(P)$ and $\frac{1}{2}h(P)$ are close to one another—in fact, it can be shown (see p. 229 of [29]) that their difference is bounded by a constant depending only on $E$—but it is the function $\hat{h}$ rather than $h$ that has the nicer properties.

Suppose that the group $E(\mathbb{Q})$ has rank $r$, i.e., the quotient group $E(\mathbb{Q})/E(\mathbb{Q})_{\text{tors}}$ is isomorphic to $\mathbb{Z}^r$. Let $P_1, \ldots, P_r$ be a set of generators. The formula

$$\langle P, Q\rangle = \frac{1}{2}(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q))$$

defines a positive inner product on the $r$-dimensional real vector space $V$ obtained from $E(\mathbb{Q})/E(\mathbb{Q})_{\text{tors}}$ by formally allowing the $P_i$ to have real (rather than just integer) coefficients. This vector space can also be defined using the tensor product notation: $V = E(\mathbb{Q}) \otimes \mathbb{R}$. Note that $E(\mathbb{Q})/E(\mathbb{Q})_{\text{tors}}$ is a full lattice in $V$.

The *regulator* of $E$ is defined as follows:

$$\mathrm{R} = \det(\langle P_i, P_j\rangle)_{1\le i,j\le r}.$$

It is the square of the volume of a fundamental parallelepiped of the lattice $E(\mathbb{Q})/E(\mathbb{Q})_{\text{tors}}$ with respect to our inner product. The real number $R$ is an important constant attached to the elliptic curve. In the Birch–Swinnerton-Dyer Conjecture, it appears as one of the factors in the first non-zero Taylor coefficient of the expansion of $L(E, s)$ at $s = 1$.

## 3. Summary of the Algorithm

### 3.1. Simplified Version

We want to find an integer $w$ such that $Q = wP$ in $E(\mathbb{F}_p)$.

Working in projective coordinates, we choose two points $\tilde{P}$ and $\tilde{Q}$ with integer coordinates whose residues modulo $p$ are our points $P, Q \in E(\mathbb{F}_p)$. We also choose an elliptic curve $E(\mathbb{Q})$ that passes through $\tilde{P}$ and $\tilde{Q}$ and that reduces modulo $p$ to the curve $E(\mathbb{F}_p)$.

Now suppose that $\tilde{P}$ and $\tilde{Q}$ turn out to be dependent in $E(\mathbb{Q})$, that is,

$$n_1\tilde{P} + n_2\tilde{Q} = O,$$

in which case $n_1$ and $n_2$ can easily be found. If that happens, working modulo $p$ we get

$$n_1 + n_2 w \equiv 0$$

modulo the order of $P$ in $E(\mathbb{F}_p)$; from this we can easily find $w$.

However, in general the probability that $\tilde{P}$ and $\tilde{Q}$ are dependent is very, very small. Silverman's idea is to increase this probability by imposing some conditions of the following type:

$$\#E(\mathbb{F}_l) \approx l + 1 - 2\sqrt{l}$$

—that is, the reduction modulo $l$ of $E(\mathbb{Q})$ has relatively few points for all primes $l$, $L_0 \le l \le L_1$ (where $L_0 \approx 7$, $L_1 \approx 100$).

This idea was suggested by J.F. Mestre's success in obtaining curves of ***higher*** than expected rank by imposing conditions in the ***opposite*** direction, i.e.,

$$\#E(\mathbb{F}_l) \approx l + 1 + 2\sqrt{l}.$$

Both strategies (for obtaining either higher-than-expected or lower-than-expected rank) are based on the heuristic argument for the conjecture of Birch and Swinnerton-Dyer (see Section 2.3), which says that the rank of $E(\mathbb{Q})$ is equal to the order of vanishing of $L(E, s)$ at $s = 1$. Mestre's method is to force the first several terms in the formal infinite product for $L(E, 1)$ to be as small as possible, whereas Silverman wants them to be as large as possible.

### *3.2. The Algorithm*

We now describe the steps in the xedni algorithm [33].

*Step 1.* Choose an integer $r$ with $2 \leq r \leq 9$ (most likely $4 \leq r \leq 6$), and integers $L_0 \approx 7$ and $L_1 \leq 100$. Set

$$M = \prod_{l \text{ prime}, L_0 \leq l \leq L_1} l.$$

Also, decide whether you will be working with elliptic curves in general cubic form or in Weierstrass form. In the first case, for any $r$-tuple of projective points $P_i = (X_i, Y_i, Z_i)$ over a field, let $B(P_1, \ldots, P_r)$ denote the $(r \times 10)$-matrix whose $i$-th row is

$$(X_i^3 \;\; X_i^2Y_i \;\; X_iY_i^2 \;\; Y_i^3 \;\; X_i^2Z_i \;\; X_iY_iZ_i \;\; Y_iZ_i^2 \;\; Y_iZ_i^2 \;\; Z_i^3).$$

Then the $r$ points lie on a given cubic curve with coefficients $u_i$, $i = 1, \ldots, 10$, if and only if the column-vector $\overline{u}$ is in the kernel of the matrix $B(P_1, \ldots, P_r)$. If, on the other hand, the elliptic curve is given in the Weierstrass form[6]

$$a_0Y^2Z + a_1XYZ + a_3YZ^2 = a_0'X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3,$$

then we take $B(P_1, \ldots, P_4)$ to be the $(r \times 7)$-matrix whose $i$-th row is

$$(Y_i^2Z_i \;\; X_iY_iZ_i \;\; Y_iZ_i^2 \;\; X_i^3 \;\; X_i^2Z_i \;\; X_iZ_i^2 \;\; Z_i^3).$$

In this case the $r$ points lie on the curve if and only if the vector $(a_0 \; a_1 \; a_3 \; -a_0' \; -a_2 \; -a_4 \; -a_6)^T$ is in the kernel of $B(P_1, \ldots, P_r)$.

*Step 2.* For each $l|M$, choose $r$ points $P_{l,i}$ in the projective plane over $\mathbb{F}_l$ such that the matrix $B(P_{l,1}, \ldots, P_{l,r})$ has rank $r$. Let $P_{M,i}$ denote a point modulo $M$ that reduces to $P_{l,i}$ modulo $l$ for each $l|M$; such a point can be found by the Chinese Remainder Theorem. If $r \geq 4$ and you're working with the general form of a cubic (rather than Weierstrass form), for convenience and slightly greater efficiency choose the first four points to be $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, and $(1, 1, 1)$.

Also choose a mod-$M$ coefficient vector $(u_{M,1},\dots,u_{M,10})$ (or, if you're using Weierstrass form, $(a_{M,0}, a_{M,1}, a_{M,3}, -a'_{M,0}, -a_{M,2}, -a_{M,4}, -a_{M,6})$) that is in the kernel of the $B$-matrix for each $l|M$. Choose the coefficient vector so that for each $l|M$ the resulting cubic curve is an elliptic curve (i.e., the discriminant is nonzero) with the fewest possible $\mathbb{F}_l$-points:

$$N_l = \#E(\mathbb{F}_l) = l + 1 - [2\sqrt{l}],$$

which is the smallest integer in the Hasse interval. This equality is called the "reverse-Mestre condition" at $l$.

*Remark 1.* In some circumstances it might be better to allow a weaker reverse-Mestre condition, and instead require only that

$$N_l = \#E(\mathbb{F}_l) = l + 1 + \varepsilon - [2\sqrt{l}],$$

where $\varepsilon = 1$ or 2.

*Remark 2.* Note that the condition that $B$ have rank $r$ implies that the $P_{l,i}$ must be distinct points, and hence $N_l = \#E(\mathbb{F}_l) \geq r$. Thus, $L_0$ must be chosen large enough so that this inequality does not contradict the (weak) reverse-Mestre condition. For example, if $r = 4$ or 5, then one can choose $L_0 = 7$.

*Remark 3.* When constructing the $P_{l,i}$ and coefficient vectors for the different small primes $l$, some care has to be taken so as not to inadvertently cause the lifted points in Step 6 below to automatically be independent. In cases when $N_l$ and $N_{l'}$ have a common factor $\tau$, there has to be a certain compatibility between the images of the $P_{l,i}$ in the quotient group $E(\mathbb{F}_l)/\tau E(\mathbb{F}_l)$ and the images of $P_{l',i}$ in $E(\mathbb{F}_{l'})/\tau E(\mathbb{F}_{l'})$.

To illustrate in a simple situation, let us take $r = 2$ and suppose that $N_{13} = 7$ and $N_{31} = 21$ in accordance with the reverse-Mestre conditions. Suppose that $P_{13,2} = aP_{13,1}$ and $P_{31,2} = bP_{31,1}$, where $a$ and $b$ are integers modulo 7 and 21, respectively. (Here we are supposing that $P_{13,1}$ is not the point at infinity, and $P_{31,1}$ is not a point of order 3.) Unless $a \equiv b \pmod 7$,the lifted points $P_1$ and $P_2$ are forced to be independent. To see this, suppose that we had a nontrivial relation of the form $n_1P_1 + n_2P_2 = O$. Since our lifted curve will almost certainly have no torsion points (in particular, no points of order 7), we may suppose that 7 does not divide both $n_1$ and $n_2$. If we reduce this relation modulo 13 and 31, we obtain $(n_1 + n_2a)P_{13,1} = 0$ and $(n_1 + n_2b)P_{31,1} = 0$. Hence $n_1 + n_2a \equiv n_1 + n_2b \equiv 0 \pmod 7$, and so $a \equiv b \pmod 7$.

*Remark 4.* The reason for requiring that the $B$-matrix have rank $r$ for each $l|M$ is that this is precisely the condition that is needed in order to ensure that one can find coefficients for an elliptic curve over $\mathbb{Q}$ that both passes through the lifted points and reduces modulo the primes $l$ and $p$ to the curves $E(\mathbb{F}_l)$ (for $l|M$) and $E(\mathbb{F}_p)$ that we already have (see Step 7 below). This is proved in Appendix B of [33]. Here we shall motivate the rank-$r$ condition for the $B$-matrix by giving an example in a simpler setting.

Suppose that $r = 2$, and we're working with straight lines in the projective plane, rather than elliptic curves, so that the $B$-matrix is just $\begin{pmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{pmatrix}$. Let $l = 3$. Suppose that we have ignored the rank-$r$ condition and over $\mathbb{F}_3$ have chosen points $P_{3,1} = (1, 1, 1)$ and $P_{3,2} = (2, 2, 2)$ and the straight line $X - Y = 0$. Suppose that we have lifted the points to $\mathbb{Q}$ as follows: $P_1 = (1, 1, 1)$, $P_2 = (2, 5, -1)$. We now want to find a lifted line $(1 + 3a)X - (1 + 3b)Y + 3cZ = 0$ that reduces to $X - Y = 0$ modulo 3 and that passes through $P_1$ and $P_2$. A simple calculation shows that this is impossible.

*Step 3.* Let $P, Q \in E(\mathbb{F}_p)$ be the points in the discrete log problem; that is, $Q = wP$ for some unknown integer $w$. Choose $r$ random integer linear combinations of the two points $P, Q$:

$$P_{p,i} = s_i Q - t_i P \in E(\mathbb{F}_p).$$

Our entire purpose in the algorithm is to find a linear dependency among the $P_{p,i}$:

$$n_1 P_{p,1} + \cdots + n_r P_{p,r} = O.$$

If we succeed, then we immediately obtain the following congruence modulo the order of the point $P$:

$$(n_1 s_1 + \cdots + n_r s_r)w \equiv (n_1 t_1 + \cdots + n_r t_r) \bmod \operatorname{ord}(P).$$

From this we can almost certainly solve for $w$ (recall that in cryptographic applications the order of $P$ is usually a large prime).

*Step 4.* If $r \geq 4$, and if you want to look for a lifted elliptic curve in general cubic form (so that you have more coefficients to work with), then make a linear change of variables in the projective plane over $\mathbb{F}_p$ so that the first four points become $P_{p,1} = (1, 0, 0)$, $P_{p,2} = (0, 1, 0)$, $P_{p,3} = (0, 0, 1)$, $P_{p,4} = (1, 1, 1)$. In that case we let $u_{p,i}$, $i = 1, \ldots, 10$, denote the coefficients of the resulting equation for $E(\mathbb{F}_p)$.

*Step 5.* Use the Chinese Remainder Theorem to find coefficients $u'_i$ modulo $Mp$ that reduce to $u_{p,i}$ modulo $p$ and to $u_{M,i}$ modulo $M$, $i = 1, \ldots, 10$. (Do the analogous thing with the $a_i$ coefficients if you are working in Weierstrass form.)

*Step 6.* Lift the $r$ points to the projective plane over the rational numbers. That is, for $i = 1, \ldots, r$ choose points $P_i = (X_i, Y_i, Z_i)$ with integer coordinates that reduce to $P_{p,i}$ modulo $p$ and to $P_{M,i}$ modulo $M$. If $r \geq 4$ and you are working with the general form of a cubic, then take the first four points to be $P_1 = (1, 0, 0)$, $P_2 = (0, 1, 0)$, $P_3 = (0, 0, 1)$, $P_4 = (1, 1, 1)$.

*Step 7.* Using the $r$ points $P_i$ from Step 6, form the matrix $B(P_1, \ldots, P_r)$. Find an integer vector $\bar{u} = (u_1, \ldots, u_{10})$ such that $B\bar{u} = 0$ and $u_i \equiv u'_i \pmod{Mp}$ (or an analogous vector of $a_i$'s if you've been working with curves in Weierstrass form). The rank-$r$ condition on the mod-$l$ $B$-matrices ensure that we can do this. Try to find $\bar{u}$ so that the $u_i$ are as small as possible.

*Step 8.* If you've been working with the general equation of a cubic, make a linear change of variables to bring it into Weierstrass form.

*Steps 9–10 (optional).* Modify the solution $\bar{u}$ in Step 7 by adding or subtracting vectors of the form $Mp\bar{v}$, where the vectors $\bar{v}$ are chosen from a basis of solutions to $B\bar{v} = 0$ that have small coordinates. Choose a new solution $\bar{u}$ such that the discriminant of the curve with coefficients $u_1, \ldots, u_{10}$ is as small as possible. (Go through the analogous procedure with the $a_i$ if you've been working with curves in Weierstrass form).

Also, let $L$ be a constant of order about 200. For each curve compute the sum

$$\sum_{l \le L, l \nmid M} a_l \frac{\log l}{l}.$$

If this sum is smaller than a pre-determined quantity (that is arrived at experimentally), discard the curve and start over again with Step 2 or Step 3. Otherwise, continue to Step 11.

Step 10 is based on an analytic formula for the rank of a modular curve that was proved by Mestre [20]. (Notice that his formula can be used because of the Taniyama Conjecture, which says that all elliptic curves over $\mathbb{Q}$ are modular.) In Mestre's formula the above sum appears as a crucial term. Heuristically, it is plausible that the more negative this sum is, the more likely the curve is to have large rank. Since we want smaller-than-expected rank, we might want to throw out curves for which the sum is highly negative.

*Step 11.* Finally, test the points for dependence. There are at least two efficient methods of doing this (see [33]). If they are independent, return to Step 2 or Step 3. If they are dependent, it is not hard to find the coefficients of a relation. As explained in Step 3, it is then very easy to find the discrete logarithm $x$. This completes the description of the algorithm.

## 4. Asymptotic Failure of the Algorithm

The purpose of this section is to prove

THEOREM 4.1 *Under certain plausible assumptions (see the lemma below), there exists an absolute constant $C_0$ such that the probability of success of the xedni algorithm in finding a discrete logarithm on an elliptic curve over $\mathbb{F}_p$ is less than $C_0/p$.*

Unfortunately, $C_0$ is rather large, so this result does not immediately resolve the question of practicality of the algorithm. We address that question in the next section.

Recall the notion of the canonical logarithmic height $\hat{h}(P)$ (see §2.4). Given an elliptic curve $E$ over $\mathbb{Q}$ having infinitely many rational points, let $m$ denote the minimum of $\hat{h}(P)$ for all nontorsion points $P \in E(\mathbb{Q})$. Let $D$ denote the discriminant of $E$. Then a conjecture of Lang (see p. 92 of [12] or p. 233 of [29]) states that there exists a positive absolute constant $C_3$ such that $m > C_3 \log |D|$. This conjecture was proved for a large class of curves in [27, 8], but it has not yet been proved unconditionally for all curves over $\mathbb{Q}$.

LEMMA 4.1 *Assume that* $\log|D| \geq C_1 \max_{i=1,\ldots,r} \hat{h}(P_i)$ *for the lifted curves in the xedni algorithm, where D is the discriminant of the lifted curve,* $P_i$ *are the lifted points,* $\hat{h}$ *is the canonical logarithmic height, and* $C_1$ *is a positive absolute constant.*[7] *Then, under Lang's conjecture, if the lifted points are dependent, then they satisfy a nontrivial relation with coefficients bounded from above by an absolute constant* $C_2$.

*Proof.* Following [34], we estimate the number of points of $E(\mathbb{Q})$—more precisely, the number of points in the subgroup $E'$ spanned by the lifted points $P_1, \ldots, P_r$—whose canonical logarithmic height is bounded by a constant $B$. Suppose that the $P_i$ are independent, and let $r' \leq r-1$ denote the rank of $E'$. Let $T'$ denote the number of torsion points in $E'$. (In practice, almost certainly $T' = 1$; and by a famous theorem of Mazur [16] always $T' \leq 16$.) Let $V' = E' \otimes \mathbb{R}$, and let $R'$ denote the regulator of $E'$, i.e., $R' = \det(\langle P_i', P_j'\rangle)_{1\leq i,j\leq r'}$, where $P_1', \ldots, P_{r'}'$ are a basis for $E'/E'_{\text{tors}}$. Finally, we define $N(B) = \#\{P \in E' : \hat{h}(P) \leq B\}$.

To estimate $N(B)$, one uses standard results from the geometry of numbers. According to Theorem 7.4 of Chapter 5 of [13],

$$N(B) = T\frac{V_{\text{ball}}(r')}{\sqrt{R'}}B^{r'/2} + O(B^{(r'-1)/2}),$$

where $V_{\text{ball}}(r')$ is the volume of the $r'$-dimensional unit ball:

| $r'$ | $V_{\text{ball}}(r')$ |
|---|---|
| 1 | 2 |
| 2 | $\pi = 3.14\cdots$ |
| 3 | $\frac{4}{3}\pi = 4.18\cdots$ |
| 4 | $\frac{1}{2}\pi^2 = 4.93\cdots$ |
| 5 | $\frac{8}{15}\pi^2 = 5.26\cdots$ |
| 6 | $\frac{1}{6}\pi^3 = 5.16\cdots$ |
| 7 | $\frac{16}{105}\pi^3 = 4.72\cdots$ |
| 8 | $\frac{1}{24}\pi^4 = 4.05\cdots$ |

It follows from Corollary 7.8 of Chapter 5 of [13] that

$$R' \geq \left(\frac{1}{2}\sqrt{3}\right)^{r'(r'-1)} m^{r'},$$

where, as before, $m$ denotes the smallest positive value of $\hat{h}$ on $E(\mathbb{Q})$ (actually, we could replace $m$ by the smallest positive value of $\hat{h}$ on $E'$). If we combine these relations and denote

$$c(r) = V_{\text{ball}}(r)\left(\frac{\sqrt{3}}{2}\right)^{-r(r-1)/2},$$

we obtain

$$N(B) < Tc(r')\left(\frac{B}{m}\right)^{r'/2} + O(B^{(r'-1)/2}).$$

Now let $\mathcal{M}$ denote the maximum of $\hat{h}(P_i)$, $i = 1, \ldots, r$. Since $\sqrt{\hat{h}}$ is a metric, the height of any integer linear combination of the $P_i$ with coefficients $n_i$ bounded by $\frac{1}{2}C_2$ (where $C_2$ will be chosen later) is bounded as follows:

$$\hat{h}(n_1P_1 + \cdots + n_rP_r) \leq \left(\frac{r}{2}C_2\sqrt{\mathcal{M}}\right)^2 = \left(\frac{r}{2}\right)^2 C_2^2\mathcal{M}.$$

If we substitute $B = (\frac{1}{2}r)^2C_2^2\mathcal{M}$ in our inequality for $N(B)$, we find that the number of points that $\sum n_iP_i$ can be, i.e., the number of points that satisfy the above inequality for the height, is less than

$$Tc(r')\left(\frac{r}{2}\right)^{2'} C_2^{r'}\left(\frac{\mathcal{M}}{m}\right)^{r'/2}.$$

But the number of linear combinations $\sum n_iP_i$ with $|n_i| \leq \frac{1}{2}C_2$ is very close to $C_2^r$. If

$$C_2^r > Tc(r')\left(\frac{r}{2}\right)^{r'} C_2^{r'}\left(\frac{\mathcal{M}}{m}\right)^{r'/2},$$

then there must be two different linear combinations that are equal, and so the points $P_i$ satisfy a nontrivial linear relation with coefficients bounded by $C_2$.

We now use the assumptions in the lemma. By Lang's conjecture, $m \geq C_3 \log|D|$. Since we also assumed that $\log|D| \geq C_1\mathcal{M}$ for some positive absolute constant $C_1$, we have

$$\frac{\mathcal{M}}{m} \leq \frac{1}{C_1C_3}.$$

Dividing the previous inequality through by $C_2^{r'}$, and using the fact that $r' \leq r - 1$, we find that it suffices to have

$$C_2 \geq Tc(r-1)\left(\frac{r}{2}\right)^{r-1}(C_1C_3)^{-(r-1)/2}.$$

Since $T \leq 16$ and there are only finitely many possibilities for $r$, namely, $2 \leq r \leq 9$, this is an absolute constant. The lemma is proved. ■

We now show how the theorem follows from the lemma. The point is that any relation among the lifted points $P_i$ can be reduced modulo $p$ to get a relation with the same coefficients among the original $r$ points $P_{p,i}$ that were constructed at random in Step 3. However, it is extremely unlikely that $r$ random points on $E(\mathbb{F}_p)$ will satisfy a linear relation with coefficients less than a constant bound. In fact, using a pigeon-hole argument, one can show that the smallest value of $\max|n_i|$ that is likely to occur for the coefficients in a relation is of order $O(p^{1/r})$. If the points $P_{p,i}$ in Step 3 do not satisfy a relation with coefficients less than the bound in the lemma, then no amount of work with Mestre conditions is going to enable one to lift them to dependent points.

To make the argument more precise, consider the map from $r$-tuples of integers less than $C_2$ in absolute value to $E(\mathbb{F}_p)$ given by $(n_1, \ldots, n_r) \mapsto n_1 P_{1,p} + \cdots + n_r P_{r,p}$. The image is a set of $\approx (2C_2)^r$ randomly distributed points. The probability that the image contains 0 is approximately $(2C_2)^r/p$. This proves the theorem with $C_0 = (2C_2)^r$.

Unfortunately, the certain failure of the algorithm for large primes $p$ does not rule out its practicality for $p$ of an "intermediate" size, such as $p \approx 10^{50}$. After examining about 10000 curves, Silverman [27] was able to bound the constant $C_3$ in Lang's conjecture as follows: $C_3 < (2000)^{-1}$.[8] That circumstance alone contributes a factor of at least $2000^{(r-1)/2}$ to the constant $C_2$ in the lemma, and at least $2000^{r(r-1)/2}$ to the constant $C_0$ in Theorem 4.1. In any case, it is now clear that Silverman was correct to choose $r > 2$. If $r$ were equal to 2 (as in the "simplified version" in §3.1), then $C_2$ could be chosen much smaller, and our theorem would apply to $p$ of more moderate size.

This situation is very unusual. We know, subject to various reasonable conjectures, that for sufficiently large $p$ the xedni algorithm must be repeated at least $O(p)$ times (with different choices of $r$ points in Step 3) in order to find a discrete logarithm. In other words, asymptotically it is far slower than square-root attacks. However, because of the constants involved, this result does not necessarily imply that the algorithm is inefficient for $p$ in the range that arises in practical cryptography.

### *4.1. Estimate of the Constant in Theorem 4.1*

In order to get a very rough estimate for the constant $C_0$ in Theorem 4.1, we shall make the following assumptions:

- The constant $C_3$ in Lang's conjecture is no less than $1/10$ of the upper bound in [27], i.e., $C_3 \geq 0.00005$.

- For $r = 2, 3, 4, 5, 6$, one uses the Weierstrass form of the equation of the elliptic curve with 7 variable coefficients. We suppose that the ratio of length of the coefficients to length of the coordinates of the $r$ points is given by a formula derived from Siegel's Lemma, as in Appendix J of [33], namely, $1 + 3r/(7 - r)$. We further suppose that the length of the discriminant is 12 times the length of the coefficients.

- For $r = 7, 8, 9$, one uses the general equation of a cubic, which has 10 variable coefficients. We suppose that the ratio of lengths of coefficients to coordinates is now $1+3r/(10-r)$ (see Appendix J of [33]). In accordance with computations of Silverman (see Appendix C of [33]), we also assume that the length of the discriminant is 110 times the length of the coefficients.

- The curves over $\mathbb{Q}$ have no nontrivial torsion points, as one expects to happen in the vast majority of cases.

We now use the bound in the proof of Theorem 4.1:

$$C_0 = T^r \frac{2^{r(r-2)(r-3)/2}}{3^{r(r-1)(r-2)/4}} r^{r(r-1)} (V_{\mathrm{ball}}(r-1))^r (C_1 C_3)^{-r(r-1)/2},$$

where $T$, $C_1$, and $C_3$ are determined according to the four assumptions above, i.e., $T = 1$, $C_3 = 0.00005$ and $C_1 = 12(1 + 3R/(7 - R))$ or $110(1 + 3r/(10 - r))$ for $r = 2, 3, 4, 5, 6$ and $r = 7, 8, 9$, respectively. Here is the result:

| $r$ | very rough value for $C_0$ |
|---|---|
| 2 | $10^4$ |
| 3 | $10^{12}$ |
| 4 | $10^{23}$ |
| 5 | $10^{38}$ |
| 6 | $10^{54}$ |
| 7 | $10^{65}$ |
| 8 | $10^{84}$ |
| 9 | $10^{100}$ |

We conclude that for $p \approx 10^{50}$, Theorem 4.1 rules out the use of the algorithm with $r \leq 5$, but not necessarily with $r = 6, 7, 8, 9$. Nevertheless, in our experimental work, where the primes were much smaller, we took $r = 2, 3, 4$ in order to investigate the probability of dependence, the effect of reverse-Mestre conditions, and other issues.

Note that when $p \approx 10^{50}$ we can expect to be working with elliptic curves over $\mathbb{Q}$ whose discriminants have at least 10000 decimal digits when $r = 6$ and 150000 digits when $r = 9$. This obviously casts doubt on the feasibility of the computations in the algorithm. We shall explore the practicality question in more detail in the next section.

*Remark 5.* Our estimate for $C_1$ might be too high, because sometimes one can obtain smaller coefficients and discriminants using lattice-basis reduction and other methods. On the other hand, the value we are using for $C_3$ is almost certainly too low; so it is reasonable to hope that our value for the product $C_1 C_3$ is about right.

## 5. Empirical Analysis in the Practical Range

To get a practical estimate of the probability of success of the xedni algorithm, we did several experiments, including an implementation of the algorithm itself. All experiments were carried out using the computer algebra systems LiDIA [14] and SIMATH [38]. We began with a couple of preliminary computations. The purpose of this was to obtain some insight into which parameters have an impact on the probability of dependence. Our strategy and the size of parameters were chosen with the aim of producing a significant number of dependencies. We tried to keep the size of the curve coefficients, and hence the size of the discriminant, as small as possible. We worked with $r = 2, 3$ and 4 points through which the curve was made to pass, and we did not impose any reverse-Mestre conditions. The data obtained through these experiments already suggested that most likely the xedni algorithm has a negligible probability of success. However, to be more confident of this statement, we implemented the algorithm. It turned out that the probability of success was small even for 8-bit primes.

### *5.1. A First Approach*

#### *5.1.1. The Experiment*

For each value $r = 2, 3, 4$, 200000 curves were generated as follows. First, $r$ affine points $P_1, \ldots, P_r$, $P_i = (x_i, y_i)$ were randomly selected with integers $|x_i|$ and $|y_i|$ bounded by 40 when $r = 2$ and by 30 when $r = 3, 4$, such that the points had pairwise distinct $x$-coordinates and none of them was the point at infinity. The points were discarded if any three of the $r$ points ($r \geq 3$) were collinear. Note that if three points $P$, $Q$ and $R$ are collinear and $E$ is an elliptic curve passing through these points, then $P + Q + R = O \in E(\mathbb{Q})$, independently of $E$. Second, the five coefficients $a_i$ ($i = 1, 2, 3, 4, 6$) of a curve in standard Weierstrass form (with $a_0 = a_0' = 1$) were selected so that the curve passed through the $r$ points and the coefficients were small. If there was no solution with integer coefficients, the points were discarded. Third, the curve (and points) were discarded if the curve had the same $j$-invariant as an earlier curve. Fourth, the same was done if any of the $r$ points were torsion points or if the curve had nontrivial 2-torsion. Finally, in the cases $r = 3, 4$ if the discriminant was greater than $2^{80}$, that case was also discarded. The reason for this was that in preliminary experiments we were unable to find a single case of dependency with discriminant greater than $2^{72}$, and we wanted to avoid a lot of fruitless computation.

In all cases we computed the discriminant and the number of mod-$l$ points for $7 \leq l \leq 97$. A 2-descent (see [33], Appendix D) was used to check dependence. When the points were dependent, the dependency relation with smallest coefficients was determined.

#### *5.1.2. Results*

Among the 200000 examples considered for each $r = 2, 3, 4$, we found 2895, 21165 and 10698 dependent cases, respectively. For each value of $r = 2, 3, 4$ and each bit-length of the discriminant $D$, the proportion of dependent cases (i.e., the probability of dependence) was tabulated and compared with various fractional powers of the discriminant. The data suggest that when $r = 2, 3, 4$ the probability of dependence is bounded, respectively, by $5|D|^{-1/4}$, $66|D|^{-1/4}$, $322|D|^{-1/4}$. Some explicit results for $r = 4$ are given in Table 1. Here column $A$ is the bit-length of the discriminant; to keep the table small, we restrict ourselves to listing the data for discriminants of bit-length $5k$, $k \geq 1$, and for the largest discriminants. Column $B$ is the number of example curves having discriminant of bit-length A. Column $C$ is the number of these curves for which the four points are dependent. The fourth column is the proportion $C/B$ of dependencies. The last three columns show the values of $R_e = 2^{A/e}C/B$, where $e = 3, 4, 5$. Thus, $R_e$ is approximately equal to the $e$-th root of the discriminant times the fraction of examples where the points were dependent: $R_e \approx |D|^{1/e} \cdot C/B$.

*Table 1.* $r = 4$: probability of dependence.

| length($\|D\|$) | # curves | # curves w.dep.pts. | $\frac{\text{\#dep.}}{\text{total\#}}$ | $R_3$ | $R_4$ | $R_5$ |
|---|---|---|---|---|---|---|
| 10 | 4 | 4 | 1 | 10.08 | 5.66 | 3.17 |
| 15 | 10 | 10 | 1 | 32 | 13.45 | 5.66 |
| 20 | 112 | 106 | 0.946 | 96.15 | 30.29 | 9.54 |
| 25 | 446 | 352 | 0.789 | 254.56 | 60.07 | 14.17 |
| 30 | 991 | 535 | 0.54 | 552.82 | 97.72 | 17.28 |
| 35 | 1433 | 411 | 0.287 | 932.42 | 123.48 | 16.35 |
| 40 | 1879 | 299 | 0.159 | 1642.4 | 162.95 | 16.17 |
| 45 | 2409 | 194 | 0.081 | 2638.85 | 196.13 | 14.58 |
| 50 | 3079 | 133 | 0.043 | 4493.75 | 250.22 | 13.93 |
| 55 | 3725 | 65 | 0.017 | 5763.29 | 240.41 | 10.03 |
| 60 | 4417 | 17 | 0.004 | 4035.72 | 126.12 | 3.94 |
| 65 | 5100 | 4 | 0.001 | 2611 | 61.13 | 1.43 |
| 70 | 5754 | 2 | 0.0003 | 3673.61 | 64.43 | 1.13 |
| 71–72 | 12150 | 2 | 0.0002 | 2761.68 | 43.15 | 0.67 |
| 73–80 | 52156 | 0 | 0 | 0 | 0 | 0 |

The average value of $\sum_{7\leq l\leq 97} \frac{a_l \log l}{l}$ for $r = 2, 3, 4$ was, respectively, $-4.401$, $-6.163$, $-8.108$ for all curves and $-2.227$, $-4.336$, $-6.597$ for the dependent cases. In other words, very roughly it was equal on average to $-2$(rank of curve).

We also looked at the reverse-Mestre conditions for $7 \leq l \leq 97$. Of the 22 values of $l$, no curve satisfied more than 3 reverse-Mestre conditions. The dependent cases had significantly more likelihood than the independent cases of satisfying these conditions—but still not a large probability. When $r = 4$, for example, 17 out of the 10698 dependent cases (about 0.16%) satisfied 2 or 3 reverse-Mestre conditions, whereas only 156 out of the 189302 independent cases (about 0.08%) did. In both cases, this proportion was far less than one expects for a random curve. The reason is that, since the curves were constructed to pass through $r$ points, they generally had higher rank, and hence in most cases more mod-$l$ points, than an average curve. We also compiled statistics on the number of 'reverse-Mestre+1' and 'reverse-Mestre+2' conditions (i.e., $N_l$ is $l+1-[2\sqrt{l}]+1$ or $l+1-[2\sqrt{l}]+2$, respectively); the results were similar to what we found for the pure reverse-Mestre conditions. For example, when $r = 4$, out of the 10698 dependent cases there were 83 cases when 2 or 3 reverse-Mestre $+1$ conditions held (none with $> 3$), and there were 148 cases when 2 or 3 reverse-Mestre $+2$ conditions held (none with $> 3$). Out of the 189302 independent cases there were 703 cases with 2 or 3 reverse-Mestre $+1$ conditions (none with $> 3$) and 1555 with 2 or 3 reverse-Mestre $+2$ conditions (2 with $> 3$).

Most remarkably, the coefficients in the dependency relations were very small. When $r = 2$, over 98% of the coefficients were 4 or less in absolute value, and no coefficient was greater than 8. When $r = 3$, over 99.75% of the coefficients were 3 or less in absolute value, and no coefficient was greater than 13. When $r = 4$, over 99% of the coefficients were 2 or less in absolute value, and no coefficient was greater than 8. This is much less than the theoretical bound $C_2$ derived in the previous section.

### *5.2. A Second Approach*

While doing experiments similar to those described above, we found an interesting effect when we tried to mix our bounds on the coordinates. Namely, at one point (with $r = 3$) we tried to add to our sample 100000 examples for which the absolute values of the coordinates of the 3 points were between 31 and 50 (rather than between 0 and 30). The large proportion of cases that led to large discriminants were discarded, leaving only the examples with smaller-than-average discriminants. In that situation there was a significant increase in the probability of dependence (roughly by a factor of 4) for fixed bit-length of the discriminant. This suggests that the probability depends not only on the size of the discriminant, but also on how this size relates to the logarithmic heights $\hat{h}(P_i)$ of the lifted points. In particular, the probability of dependence seems to be significantly greater for cures whose discriminants are much smaller than the median.

In a second series of experiments we took advantage of this phenomenon. Here we also were interested in the distribution of the discriminants of curves forced to go through $r$ random points whose coordinates were chosen to lie within certain ranges.

#### *5.2.1. The Experiment*

In this series of experiments we worked with $r = 4$ points whose coordinates $x_i, y_i$ were chosen so that

$$B_k \leq |x_i| < B_{k+1} \quad \text{and} \quad B'_k \leq |y_i| < B'_{k+1}, \quad i = 1, 2, 3, 4,$$

where

$$B_k = 6k \quad \text{and} \quad B'_k = [(6k)^{3/2}].$$

Initially, we planned to take $k = 1, \ldots, 10$, but we ended up working with $k = 1$–45, 51–57, 101–110, 150–157, 200–204, 250, 251, 252, 1000, 2100, 3000. For each such value of $k$, 100000 curves were generated in the way described above. Besides the modified bounds on the coefficients, the only difference was that we used the homogeneous Weierstrass form with 7 coefficients, computed an LLL-reduced basis $\vec{v}_1, \ldots, \vec{v}_{7-r}$ of the kernel of the matrix $B(P_1, \ldots, P_r)$, and then chose a solution vector $\vec{u}$ from the set $\{e_1\vec{v}_1 + \cdots + e_{7-r}\vec{v}_{7-r} : e_i = 0, \pm 1\}$ such that the discriminant of the corresponding curve is minimal. For each $k$, out of the 100000 curves only the 1000 with smallest discriminant were examined for dependency. Thus, about 8 million curves were generated, and 1% of them were examined for dependency. For each $k$, we also looked at the distribution of the 100000 discriminants.

#### *5.2.2. Results*

The distribution of the bit-length of the discriminant was very similar for different ranges of $k$. It was not exactly a normal distribution—in particular, the mode was a few bits larger than the median, which was a few bits larger than the mean. The ratio of the standard

*Table 2.* The coefficients of the dependency relations.

| $k$ | # curves | # depend. | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1–30 | 30000 | 9893 | 7591 | 29536 | 2060 | 279 | 74 | 24 | 6 | 2 |
| 31–45 | 15000 | 2019 | 1084 | 6805 | 173 | 13 | 1 | — | — | — |
| 51–57 | 7000 | 773 | 385 | 2651 | 54 | 2 | — | — | — | — |
| 101–110 | 10000 | 909 | 296 | 3311 | 27 | 2 | — | — | — | — |
| 150–157 | 8000 | 625 | 174 | 2311 | 14 | 1 | — | — | — | — |
| 200–204 | 5000 | 623 | 87 | 2395 | 10 | — | — | — | — | — |
| 250–252 | 3000 | 169 | 42 | 631 | 3 | — | — | — | — | — |
| 1000 | 1000 | 27 | 2 | 106 | — | — | — | — | – | — |
| 2100 | 1000 | 41 | 3 | 161 | — | — | — | — | — | — |
| 3000 | 1000 | 36 | 1 | 143 | — | — | — | — | — | — |

deviation to the mean was 0.22 for all $k \geq 11$ and between 0.25 and 0.23 for $1 \leq k \leq 10$. As a function of $k$, the median was very close to $23 \log_2 k + 30$. The largest bit-length of discriminant for the bottom 1% was consistently 48% or 49% of the median bit-length, i.e., about $11.5 \log_2 k + 15$. For example, for $k = 101$ the smallest 1% of the curves had discriminants of bit-length between 22 and 92, while for $k = 51$ the range was 24 to 81 bits, and for $k = 3000$ the range was 63 to 151 bits.

In general, there was a much greater probability of dependence than in the previous experiment. For example, for $k = 101, \ldots, 110$, the probability of dependence was about 30% for discriminants of $\leq 40$ bits, it was about 5% for discriminants in the 60-bit range, and it dropped off gradually to about 1% for discriminants of $> 90$ bits. For the larger values of $k$, where most of the smallest 1% of discriminants had more than 100 bits, we also found many dependent cases. For example, for $k = 3000$ there were 35 dependent cases among the 998 curves with discriminants of $> 100$ bits, the largest of which was for a 151-bit discriminant. This contrasts dramatically with the earlier data, when the coordinates of the $P_i$ were much smaller and the discriminants of $> 60$ bits came from the middle and high range of discriminants; in that case we did not find a single dependency among the vast number of cases of discriminant $> 2^{72}$. Moreover, the probability of dependence was no longer bounded by $\mathit{const} \cdot |D|^{-1/4}$. Hence, having smaller than expected discriminant helps force the points to be dependent.

However, when we examined the sizes of the coefficients in the dependency relations, we realized that it was the very small size of these coefficients, rather than the small probability of dependence for large $|D|$, that would be the most serious obstacle to the xedni calculus. These coefficients tended to be as small or smaller than in the previous experiment. Moreover, the chance of finding a dependency coefficient other than 1, $-1$, 0 drops significantly as the discriminant grows. For example, for $k > 32$ we encountered no coefficients of absolute value greater than 3. In Table 2 we give the distribution of the dependency coefficients. The first column is the range of $k$-values; the second column is the number of curves examined (i.e., 1000 times the number of $k$-values in the range); the third column is the number of dependent cases. The column labeled $i$ is the number of dependency coefficients of absolute value $i$ (thus, the sum of all of these columns is equal to 4 times the third column).

Out of the 27 dependent cases for $k = 1000$, 17 relations were of the form $P_1 + P_2 = P_3 + P_4$. For $k = 2100$, 33 out of 41 relations were of this form, and for $k = 3000$ this was the case for 31 out of 36 relations. Note that the probability of getting this relation is simply the probability that, when one passes a curve through the four points, it also passes through the point of intersection of a line through two of the points with the line through the other two. Although there is a significant chance of this happening even when $k$ is large, this type of relation with coefficients $\pm 1$ is not useful for solving the ECDLP, where the coefficients will be large.

We also wanted to see if the data could have been affected by the particular way we generated the points (especially, the narrow range of $|x_i|$ and the fact that $|y_i|$ was so close to $|x_i|^{3/2}$). So we returned to a range roughly similar to $k = 250$, but this time with $37^2 = 1369 \leq |x_i| < 1600$, $37^3 = 50653 \leq |y_i| < 64000$ and also with $1369 \leq |x_i|$, $|y_i| < 1600$. In each case we generated 100000 examples and examined the bottom 1%. This time the discriminants were much larger than before (up to 162 bits in the first case and up to 125 bits in the second case), presumably because LLL had been able to find much smaller coefficients when $|x_i|^3$ was very close to $|y_i|^2$. Out of 1000 curves there were, respectively, 14 and 50 dependencies, of which ten and eight were of the form $P_1 + P_2 = P_3 + P_4$. Once again there were no coefficients other than 1, $-1$, 0.

### 5.3. Preliminary Conclusions

So far, our experiments showed the following. First, the probability of dependence drops off with increasing bit-length of the discriminant, but this drop-off depends on more than just the bit-length. Another factor is the ratio of the actual size of the discriminant to the expected size.

Second, reverse-Mestre conditions are more likely to be satisfied in the dependent cases than in the independent cases. What is the probability of dependence given that reverse-Mestre conditions hold for a few small primes? Such data cannot be extracted from our experiments. For example, in the first experiment (with 200000 curves) and in the second experiment (with 10000 curves of relatively small discriminant and $k = 101$–$110$) we checked for reverse-Mestre conditions and reverse-Mestre $+1$ conditions for $l = 7$, 11 and 13. We found that in none of the cases, dependent or independent, were any two such conditions satisfied simultaneously.

Third, the small sizes of the dependency coefficients seemed to cast doubt on the practicality of the xedni algorithm. At this point we did not yet have data reflecting the situation of ECDLP, where we deal with points whose smallest relation is necessarily fairly large. What is the probability that $P_1, \ldots, P_r$ are dependent, given that we know a priori that any relation they satisfy must have moderately large coefficients?

### 5.4. Experiments with the Xedni Algorithm

To answer the questions raised above, we implemented the xedni algorithm. The size of the parameters was chosen so that we had a reasonable chance of finding some dependent cases.

For $p = 17$ and for $p = 67$, we did three different experiments, which can be classified as follows: (A) no reverse-Mestre conditions imposed; (B) reverse-Mestre conditions imposed for two small primes whose product $M$ is of approximately the same size as $p$; (C) instead of $p$ work with $p' \approx Mp$, with no reverse-Mestre conditions imposed but with $p'$ taken to be of the same magnitude as the product $Mp$ in (B).

In the context of an actual ECDLP, this means that both Experiments A and B would be used to solve the same ECDLP but with different strategies. That is, the reverse-Mestre conditions in Experiment B would presumably contribute to a greater likelihood of dependency, but at the expense of much larger discriminants (which would work against dependency). Experiment C, on the other hand, would be used to solve an unrelated instance of ECDLP, but the discriminants in Experiment C are of similar size to those in Experiment B. Comparing Experiments A and C with Experiment B, we should be able to judge whether the reverse-Mestre conditions are helpful enough to compensate for the larger discriminants.

Let us describe Experiment B with $p = 67$ in detail. We chose $a_p = 0$, $b_p = 28$. Then the curve $y^2 = x^3 + a_p x + b_p$ over $\mathbb{F}_p$ has $N = 73$ points. We chose $P_0 = (1, 30)$ as a generator for $E(\mathbb{F}_p)$. Next, we chose $M = 77 = 7 \cdot 11$, and we chose $P_{M,i}$, $i = 1, 2, 3, 4$, to be the four points $(14, \pm15)$, $(9, \pm19)$ on the mod-$M$ curve $y^2 = x^3 + a_M x + b_M$ with $a_M = 1$, $b_M = 8$. Note that the numbers of points mod 7 and 11 are, respectively, 5 and 6. In each case the $B$-matrix has rank 4; and since the numbers of points for different $l$ are relatively prime there is no worry about incompatibility and forced independence. Using the Chinese Remainder Theorem, we then compute $a, b$ with $-77p/2 < a, b < 77p/2$ to be congruent to $a_p, b_p \bmod p$ and congruent to $a_M, b_M \bmod 77$. Hence $a = 1541$ and $b = 162$. Then the steps 1–4 below are repeated 100000 times.

1. For any vector $n \in \mathbb{F}_N^4$ define $\|n\|^2$ to be $\sum_i n_i^2$, where the coordinates $n_i$ of $n$ are taken in the interval $-N/2 < n_i < N/2$. For $i = 1, 2, 3, 4$ set $P_{p,i} = \mu_i P_0$, where the vector $\mu \in \mathbb{F}_N^4$ is chosen so that $\mu_1 = 1$, $\mu_i \not\equiv 0 \bmod N$, and $\|n\|^2 \geq 5$ for all nonzero vectors $n \in \mathbb{F}_N^4$ orthogonal to $\mu$. This means that we do not allow the $P_i$ to satisfy a relation with all coefficients $0, 1, -1$.

2. For each $i = 1, 2, 3, 4$ use the Chinese Remainder Theorem to choose $(x_i, y_i)$ to be congruent to the coordinates of $P_{p,i} \bmod p$ and to those of $P_{M,i} \bmod 77$. Now choose $P_i = (X_i, Y_i, Z_i)$ in projective coordinates by finding a short vector in the lattice generated by the columns of the matrix
$$\begin{pmatrix} x_i & 77p & 0 & 0 \\ y_i & 0 & 77p & 0 \\ 1 & 0 & 0 & 77p \end{pmatrix}$$
subject to the condition that $Z_i$ is not divisible by 7, 11, or $p$.

3. Solve for small integers $u_i$ such that the curve $E(\mathbb{Q})$ with equation
$$\begin{aligned}(1+77pu_1)Y^2Z+77pu_2XYZ+77pu_3YZ^2 &= (1+77pu_4)X^3+77pu_5X^2Z \\ &\quad + (a+77pu_6)XZ^2+(b+77pu_7)Z^3\end{aligned}$$

passes through $P_i = (X_i, Y_i, Z_i)$, $i = 1, 2, 3, 4$, and has minimal discriminant. Here we use the techniques described in Steps 7 and 9 and Appendix B of [33], including the Havas-Majewski-Matthews Hermite normal form algorithm [7].

4. Finally, check whether the $P_i$ are dependent. In case of dependency, compute the dependency relation with smallest coefficients.

Experiment A differs from Experiment B only in that $M = 1$. For the corresponding Experiment C, we chose $p = 5167$, $a_p = 2462$, $b_p = 1260$, and $P_0 = (2{,}946)$. The curve $y^2 = x^3 + a_p x + b_p$ over $\mathbb{F}_p$ has $N = 5153$ points.

For Experiments A and B with $p = 17$, we took $r = 2$ and chose $a_p = b_p = 2$. Then the curve $y^2 = x^3 + a_p x + b_p$ over $\mathbb{F}_p$ has $N = 19$ points. We chose $P_0 = (3, 1)$ as generator. We worked with $M = 15 = 3 \cdot 5$, and chose $P_{M,1}$ and $P_{M,2}$ to be the two points $(5, \pm 2)$ on the mod-$M$ curve $y^2 = x^3 + x^2 + x + 14$. The number of points is 3 both mod 3 and mod 5. The fact that $P_{l,1} = -P_{l,2}$ for both $l = 3, 5$ guarantees that we do not force the lifted points to be independent. Chinese Remaindering gives coefficients $-119$, 121 and 104. Hence, in Experiment A we work with the curve $y^2 = x^3 + 2x + 2 \bmod 17$, while in Experiment B we work with the curve $y^2 = x^3 - 119x^2 + 121x + 104 \bmod 255$. For Experiment C we chose $p = 257$, $a_p = 88$, $b_p = -41$, and $P_0 = (2, 20)$. The curve $y^2 = x^3 + a_p x + b_p$ over $\mathbb{F}_p$ has $N = 263$ points. Note that since we work with only two points, the vectors $n$ and $\mu$ of Step 1 above are vectors in $\mathbb{F}_N^2$. The only conditions imposed on $P_{p,i}$ $(i = 1, 2)$ are that they are not the point at infinity and $P_{p,1} \neq \pm P_{p,2}$.

### *5.4.1. Results*

Among the 6 series of 100000 executions of Steps 1–4 above, only in 3 series did we obtain any dependencies. This was in Experiment A with $p = 17$ and $p = 67$, and in Experiment B with $p = 17$. Details are shown in Table 3.

The data show that, given an instance of the ECDLP—i.e., a fixed value of $p$—we are more likely to produce dependent cases if we do not impose reverse-Mestre conditions. When we work with discriminants of approximately the same size—i.e., with variable $p$ but fixed size of $Mp$—the different outcomes of Experiment B when $Mp = 15 \cdot 17$ and Experiment C when $Mp = 1 \cdot 257$ might be interpreted as evidence that imposing reverse-Mestre conditions has a significant impact. However, the three relations in Experiment B are all of the form $P_1 = 2P_2$ or $P_2 = 2P_1$. Notice that once one of the two points mod $p$ is chosen, there are $N - 3$ possibilities for the other one, and the probability that the two points satisfy a dependency with coefficients $\leq 2$ is $2/(N - 3) = 1/8$ in Experiment B and $4/(N - 3) = 1/65$ in Experiment C. (Note that in Experiment B the coefficients $n_1, n_2$ must satisfy the congruence $n_1 - n_2 \equiv 0 \pmod 3$, because $P_{l,1} = -P_{l,2}$ and $N_l = 3$ for $l = 3, 5$; that is why the numerator above $N - 3$ is 2 rather than 4 in Experiment B.) Our experience has been that it is much more likely that a relation of the form $P_{p,1} \pm 2P_{p,2} = 0$ can be lifted than that a relation with larger coefficients can be lifted. Thus, the greater likelihood of dependency in Experiment B than in Experiment C might have little or nothing to do with the reverse-Mestre conditions.[9]

*Table 3.* Experiments A, B, C (100000 examples of each).

| | | $D_{\mathrm{min}}$ | $D_{\mathrm{max}}$ | $D_{\mathrm{av}}$ | # dep. | dependent cases | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | $D_{\mathrm{min}}$ | $D_{\mathrm{max}}$ | $D_{\mathrm{av}}$ |
| A | $p = 17$<br>$M = 1$ | 17 bits | 131 bits | 73 bits | 317 | 23 bits | 91 bits | 61 bits |
| B | $p = 17$<br>$M = 15$ | 41 bits | 273 bits | 182 bits | 3 | 140 bits | 151 bits | 144 bits |
| C | $p = 257$<br>$M = 1$ | 40 bits | 255 bits | 167 bits | 0 | — | — | — |
| A | $p = 67$<br>$M = 1$ | 57 bits | 257 bits | 148 bits | 153 | 59 bits | 170 bits | 114 bits |
| B | $p = 67$<br>$M = 77$ | 289 bits | 612 bits | 421 bits | 0 | — | — | — |
| C | $p = 5167$<br>$M = 1$ | 269 bits | 581 bits | 394 bits | 0 | — | — | — |

*Table 4.* Experiments A: coefficients.

| | # deps. | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|
| $p = 17, r = 2$ | 317 | — | 311 | 304 | 16 | 2 | — | — | 1 |
| $p = 67, r = 4$ | 153 | 232 | 221 | 155 | 2 | 1 | — | 1 | — |

Looking at the relations in the Experiments A, we find that the great majority have coefficients 0, $\pm 1$, $\pm 2$. The sizes of the coefficients are shown in Table 4. As in Table 2, the column labeled $i$ shows the number of dependency coefficients of absolute value $i$. We see that the coefficients are very small.

Furthermore, 301 out of the 317 relations for $p = 17$ were of the form $P_1 = \pm 2P_2$ or $2P_1 = \pm P_2$. Out of the remaining 16 relations, only 6 have both coefficients larger than 1. For $p = 67$, 97 out of the 153 relations were of the form $P_i = \pm 2P_j$, 33 were of the form $P_i \pm P_j = 2P_k$, and 9 were of the form $P_i \pm P_j \pm P_k = 2P_l$. Out of the remaining 14 relations, six have two coefficients larger than one.

## 6. Conclusion

Xedni calculus is impractical for $p$ in the range used in elliptic curve cryptography. In the first place, the basic properties of the canonical logarithmic height, along with a pigeonhole argument, show that the coefficients in a dependency relation among the lifted points are bounded by an absolute constant. This implies an asymptotic running time of at least $O(p)$. In a sense, xedni fails asymptotically for much the same reason that index calculus is infeasible (see [21, 34]). In the second place, even if liftings exist with dependency among the points, the probability of finding such a lifting decreases as the discriminant grows, and it becomes very low by the time $p$ reaches the practical range. In the third place, empirical data show that the theoretical bounds on the size of the dependency coefficients are far too generous compared to what happens in practice; and for high discriminants it is virtually impossible to find dependencies where the coefficients cannot be taken to be of trivial size (usually $\pm 1$). Finally, although, in the absence of other considerations, the reverse-Mestre conditions do increase the likelihood of dependency, they also cause the discriminant to increase substantially, and so most likely the net effect is to do more harm than good.

## Acknowledgments

The authors wish to thank the faculty and staff of the Centre for Applied Cryptographic Research at the University of Waterloo for their help during the work reported on in this paper. We are especially grateful to Prof. Alfred Menezes for his support and encouragement.

## Notes

1. At about the same time, some similar ideas were developed independently in Korea [4].
2. We're assuming that $E$ has "good reduction" at $l$, i.e., that $l$ does not divide the denominators of the coefficients or the discriminant $D$ of the curve. For brevity, we shall not discuss the modifications needed for the "bad" primes $l$.
3. When $n = l$ is prime, then $a_n$ is our earlier $a_l$; for composite $n$ it is not hard to express $a_n$ in terms of $a_l$ for $l \mid n$.
4. In particular, $N|D$, and both $N$ and $D$ have the same prime divisors.
5. It would not make much difference if, instead, the logarithmic height were defined as $\log \max(|a|, |b|, |c|, |d|)$, where $y = c/d$ in lowest terms, or even as $\log_2|abcd|$, which really is (essentially) the number of bits needed to write down $P$.
6. Since it is customary to write $a_i$ for the coefficients of the general Weierstrass equation, we shall also adhere to this notation and hope that it does not lead to confusion with the use of $a_l$ ($l$ prime) to denote $l + 1 - \#E(\mathbb{F}_l)$, which is also customary. Also note that usually one takes $a_0 = a_0' = 1$ and $a_1, a_2, a_3, a_4, a_6 \in \mathbb{Q}$; however, we want integer rather than rational coefficients, so it is useful to introduce $a_0$ and $a_0'$.
7. Roughly speaking, this condition says that the discriminant of the lifted curve is greater than the $C_1$-th power of the maximum absolute value of the numerators and denominators of the coordinates of the lifted points, for some absolute constant $C_1 > 0$. This is a reasonable assumption, since the discriminant is a polynomial function of the coefficients of the curve, and the coefficients tend to grow proportionally to a power of the integer projective coordinates of the points through which the curve must pass.
8. On the other hand, it is known (see [8, 27]) that in order to get a very small value of $C_3$, it is necessary that the discriminant $D$ be divisible by many primes to fairly high powers. However, from the way they are constructed, the xedni curves tend to have discriminants that are square-free or almost square-free.
9. There is a reason unrelated to the heuristics of the Birch–Swinnerton-Dyer Conjecture why, among the conditions that one might impose modulo $l$, $l|M$, the reverse-Mestre conditions are the ones that are most likely to produce dependencies. Note that the mod-$l$ conditions lead to congruences that the dependency coefficients must satisfy. These congruences are likely to be more restrictive if $N_l = \#E(\mathbb{F}_l)$ is larger. For example, we saw that the reverse-Mestre conditions in Experiment B led only to the constraint that $n_1 - n_2 \equiv 0 \pmod 3$, which has a small nontrivial solution $n_1 = 1, n_2 = -2$. Suppose that we had instead chosen our mod-$l$ curves and points so that $N_3 = 5, N_5 = 7$ (which are "average" rather than reverse-Mestre values) and $P_{3,2} = 2P_{3,1}$, $P_{5,2} = -P_{5,1}$. Then any dependency coefficients must satisfy $n_1 + 2n_2 \equiv 0 \pmod 5$, $n_1 - n_2 \equiv 0 \pmod 7$. One can check that the smallest (in the sense of $\|n\|$) nonzero solution to these congruences is $n_1 = 3, n_2 = -4$. It is far, far harder to find dependencies with both $n_1, n_2 \geq 3$ than it is to find dependencies with $n_1 = 1$, $n_2 = -2$.

## References

1. B. Birch and H. P. F. Swinnerton-Dyer, Notes on elliptic curves I and II, *J. Reine Angew. Math.*, Vol. 212 (1963) pp. 7–25 and Vol. 218 (1965) pp. 79–108.
2. B. Birch and H. P. F. Swinnerton-Dyer, Elliptic curves and modular functions. In B. Birch and W. Kuyk (eds.), *Modular Functions of One Variable IV* (Lect. Notes in Math., Vol. 476), Springer-Verlag, 1975, pp. 2–32.
3. J. W. S. Cassels, Diophantine equations with special reference to elliptic curves, *J. London Math. Soc*., Vol. 41 (1966) 193–291.
4. J. H. Cheon, S. G. Hahn, and H. J. Kim, Analogue of the index calculus for elliptic discrete logarithm, preprint.
5. J. Coates and A. Wiles, On the conjecture of Birch and Swinnerton-Dyer, *Invent. Math*., Vol. 39 (1977) pp. 223–251.
6. R. Greenberg, On the Birch and Swinnerton-Dyer conjecture, *Invent. Math*., Vol. 72 (1983) pp. 241–265.
7. G. Havas, B. Majewski, and K. Matthews, Extended GCD and Hermite normal form algorithms via lattice basis reduction, *Experimental Math.*, Vol. 7 (1998) pp. 125–136.
8. M. Hindry and J. H. Silverman, The canonical height and integral points on elliptic curves, *Invent. Math*., Vol. 93 (1988), 419–450.
9. N. Koblitz, Elliptic curve cryptosystems, *Math. Comp*., Vol. 48 (1987) pp. 203–209.
10. N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, 2nd ed., Springer-Verlag, 1993.
11. N. Koblitz, *Algebraic Aspects of Cryptography*, Springer-Verlag, 1998.
12. S. Lang, *Elliptic Curves: Diophantine Analysis*, Springer-Verlag, 1978.
13. S. Lang, *Fundamental of Diophantine Geometry*, Springer-Verlag, 1983.
14. LiDIA Group, Technische Universität Darmstadt, Darmstadt, Germany, *LiDIA—A Library for Computational Number Theory, Version 1.3*, 1997.
15. D. W. Masser, Specializations of finitely generated subgroups of abelian varieties, *Trans. Amer. Math. Soc*., Vol. 311 (1989) pp. 413–424.
16. B. Mazur, Modular curves and the Eisenstein ideal, *Inst. Hautes Études Sci. Publ. Math*., Vol. 47 (1977) pp. 33–186.
17. A. Menezes, *Elliptic Curve Public Key Cryptosystems*, Kluwer Acad. Pub., 1993.
18. A. Menezes, P. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1996.
19. J. F. Mestre, Construction d'une courbe elliptique de rang $\geq$ 12, *C. R. Acad. Sci. Paris*, Vol. 295 (1982) pp. 643–644.
20. J. F. Mestre, Formules explicites et minoration de conducteurs de variétés algébriques, *Compos. Math*., Vol. 58 (1986) pp. 209–232.
21. V. S. Miller, Use of elliptic curves in cryptography, *Advances in Cryptology—Crypto '85* (Lect. Notes in Comp. Sci., Vol. 218), Springer-Verlag (1986), pp. 417–426.
22. A. Néron, Propriétés arithmétqiues et géométriques attachés à la notion de rang d'une courbe algébrique dans un corps, *Bull. Soc. Math. France*, Vol. 80 (1952) pp. 101–166.
23. A. Néron, Quasi-fonctions et hauteurs sur les variétés abéliennes, *Annals of Math*., Vol. 82 (1965) pp. 249–331.
24. K. Ribet, On modular representations of $\mathrm{Gal}(\bar{\mathbb{Q}}, \mathbb{Q})$ arising from modular forms, *Invent. Math*. Vol. 100 (1990) pp. 431–476.
25. K. Rubin, Elliptic curves with complex multiplication and the conjecture of Birch and Swinnerton-Dyer, *Invent. Math*., Vol. 64 (1981) pp. 455–470.
26. R. Schoof, Nonsingular plane cubic curves, *J. Combinatorial Theory, Ser. A*, Vol. 46 (1987) pp. 183–211.
27. J. H. Silverman, Lower bound for the canonical height on elliptic curves, *Duke Math. J*., Vol. 48 (1981) pp. 633–648.
28. J. H. Silverman, Divisibility of the specialization map for families of elliptic curves, *Amer. J. Math*., Vol. 107 (1985) pp. 555–565.
29. J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag (1986).
30. J. H. Silverman, Computing heights on elliptic curves, *Math. Comp*., Vol. 51 (1988) pp. 339–358.
31. J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag (1994).
32. J. H. Silverman, Computing canonical heights with little (or no) factorization, *Math. Comp*., Vol. 66 (1997) pp. 787–805.

33. J. H. Silverman, The xedni calculus and the elliptic curve discrete logarithm problem, *Designs, Codes and Cryptography*, Vol. 20 (2000), pp. 5–40.
34. J. H. Silverman and J. Suzuki, Elliptic curve discrete logarithms and the index calculus, *Advances in Cryptology—ASIACRYPT '98* (Lecture Notes in Comp. Sci. Vol.), Springer-Verlag (1998), pp. 110–125.
35. J. H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag (1992).
36. R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Annals of Math*., Vol. 141 (1995) pp. 553–572.
37. A. Wiles, Modular elliptic curves and Fermat's Last Theorem, *Annals of Math*., Vol. 141 (1995) pp. 443–551.
38. H. G. Zimmer et al., *SIMATH Manual*, University of Saarland, Saarbrücken, Germany (1997).