ECM USING EDWARDS CURVES

DANIEL J. BERNSTEIN, PETER BIRKNER, TANJA LANGE, AND CHRISTIANE PETERS

ABSTRACT. This paper introduces EECM-MPFQ, a fast implementation of the elliptic-curve method of factoring integers. EECM-MPFQ uses fewer modular multiplications than the well-known GMP-ECM software, takes less time than GMP-ECM, and finds more primes than GMP-ECM. The main improvements above the modular-arithmetic level are as follows: (1) use Edwards curves instead of Montgomery curves; (2) use extended Edwards coordinates; (3) use signed-sliding-window addition-subtraction chains; (4) batch primes to increase the window size; (5) choose curves with small parameters and base points; (6) choose curves with large torsion.

1. INTRODUCTION

Factorization of integers is one of the most studied problems in algorithmic number theory and cryptology. One of the best general factorization methods available is the Elliptic-Curve Method (ECM), introduced in the 1987 paper [33] by Hendrik W. Lenstra, Jr. ECM plays an important role in factoring the "random" integers of interest to number theorists: it is not as fast as trial division and Pollard's rho method for finding tiny prime factors, but it is the method of choice for finding medium-size prime factors. ECM also plays an important role in factoring the "hard" integers of interest to cryptologists: those integers are attacked by sieving methods, which use ECM to find medium-size prime factors of auxiliary integers. ECM can also be used directly to find "large" prime factors; the current record (see [46]) is a 241-bit factor of the 1181-bit number $2^{1181} - 1$.

Implementations of ECM are available in most computer-algebra packages and have been the subject of many papers. The state-of-the-art implementation is GMP-ECM, described in detail in the paper [47] by Zimmermann and Dodson.

We have built a new ECM implementation, "EECM-MPFQ", that uses fewer modular multiplications than GMP-ECM, takes less time than GMP-ECM, and finds more primes than GMP-ECM. Our first prototype of EECM-MPFQ was "GMP-EECM", a program that added various improvements to GMP-ECM; we thank Zimmermann et al. for making their software freely available!

In this paper we present the background and speed results for EECM-MPFQ. To simplify verification and reuse of our results we have published the EECM-MPFQ software at http://eecm.cr.yp.to and placed it into the public domain. We are continuing to update http://eecm.cr.yp.to with new information about EECM: for example, the followup paper [11] reports a very fast parallel implementation of EECM using graphics cards as a computational platform.

1.1. **Representations of elliptic curves.** Elliptic curves can be expressed in many forms, and elliptic-curve computations can be carried out in many ways. Two fast options reigned supreme for twenty years of elliptic-curve factoring, elliptic-curve primality proving, and (in large characteristic) elliptic-curve cryptography:

- Short Weierstrass curves $y^2 = x^3 + a_4 x + a_6$, with Jacobian coordinates $(X : Y : Z)$ representing $(X/Z^2, Y/Z^3)$, were the representation of choice for most computations.
- Montgomery curves $By^2 = x^3 + Ax^2 + x$, with Montgomery coordinates $(X : Z)$ representing two points $(X/Z, \pm \cdots)$, were the representation of choice for single-scalar multiplication, and in particular for stage 1 of ECM.

The picture changed in 2007 with the advent of Edwards curves. A sequence of papers [13], [10], [14], [15], and [29] showed that, for cryptographic applications, Edwards curves involve significantly fewer multiplications than short Weierstrass curves in Jacobian coordinates, and — for sufficiently large scalar multiplications — fewer multiplications than Montgomery curves in Montgomery coordinates. Note that larger scalars benefit from larger windows, reducing the number of additions per bit for Edwards coordinates but not for Montgomery coordinates.

1.2. **Contributions of this paper.** In this paper we analyze the impact of Edwards curves on ECM, not just in multiplication counts but also in real-world software speeds.

Section 2 discusses the group law on Edwards curves and twisted Edwards curves, and reviews various coordinate systems for Edwards curves. Our prototype GMP-EECM used twisted inverted Edwards coordinates, and EECM-MPFQ uses extended Edwards coordinates. Section 3 analyzes points of small order on Edwards curves. Sections 4 and 5 discuss the use of Edwards curves inside ECM. Our announcement of GMP-EECM in January 2008 marked the first time that Edwards curves had been demonstrated to achieve software speed records.

A large portion of this paper is devoted to explaining which curves we use in EECM-MPFQ. Curves having 12 or 16 torsion points over **Q** are guaranteed to have 12 or 16 as divisors of their group orders modulo primes (of good reduction), improving the smoothness chance of the group orders and thus improving the success chance of ECM. We show how to use analogous improvements for Edwards curves; even better, we find new curves with large torsion group, small curve parameters, and small nontorsion points.

Section 6 explains how to construct Edwards curves having torsion group $\mathbf{Z}/12\mathbf{Z}$ or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ over **Q**; the symmetry of Edwards curves simplifies the constructions. Section 6 also shows that twisted Edwards curves cannot have torsion group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$ or $\mathbf{Z}/10\mathbf{Z}$ over **Q**, and that twisted Edwards curves with torsion group $\mathbf{Z}/12\mathbf{Z}$ or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ over **Q** cannot have curve parameter $a = -1$. Section 7, adapting a construction of Atkin and Morain from [3] to the Edwards context, explains how to construct an infinite family of Edwards curves having torsion group

$\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and (as required for ECM) an explicit nontorsion point; and similarly adapts a construction of Montgomery for $\mathbf{Z}/12\mathbf{Z}$. Section 8 describes how we found better choices of Edwards curves to use in EECM-MPFQ; each of these curves has torsion group $\mathbf{Z}/12\mathbf{Z}$ or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$, an explicit nontorsion point, and small (i.e., fast) parameters.

Section 9 reports measurements of ECM success probabilities, demonstrating the importance of a large torsion group. Section 10 reports the overall effectiveness of EECM-MPFQ when parameters are chosen sensibly; for example, it shows that one curve finds 13.414% of all 30-bit primes in just 3065 modular multiplications.

See [9] for followup work exploring the price-performance ratio of the $a = -1$ case of twisted Edwards curves in ECM.

2. EDWARDS CURVES

This section reviews twisted Edwards curves, and Edwards curves as a special case; the set of points on a twisted Edwards curve in affine, projective, inverted, extended, and completed forms; the Edwards addition law and a dual addition law, together turning the completed twisted Edwards curve into a group; and the speeds of addition and doubling in various representations.

For a collection of explicit formulas and operation counts for elliptic curves in various representations we refer to the Explicit-Formulas Database [12].

2.1. **Edwards curves and twisted Edwards curves.** Let k be a field in which $2 \neq 0$, and let a, d be distinct nonzero elements of k. The *twisted Edwards curve* $\mathrm{E}_{\mathrm{E},a,d}$ is given by

$$\mathrm{E}_{\mathrm{E},a,d} : ax^2 + y^2 = 1 + dx^2y^2.$$

An *Edwards curve* is a twisted Edwards curve in which $a = 1$; i.e., a curve of the form $x^2 + y^2 = 1 + dx^2y^2$ where $d \in k \setminus \{0, 1\}$.

If $a\bar{d} = \bar{a}d$, then the two curves $\mathrm{E}_{\mathrm{E},a,d}$ and $\mathrm{E}_{\mathrm{E},\bar{a},\bar{d}}$ are isomorphic over $k(\sqrt{a/\bar{a}})$ and therefore are quadratic twists over k. An isomorphism is given by $(x, y) \mapsto (\bar{x}, \bar{y}) = (\sqrt{a/\bar{a}}x, y)$. In particular, the twisted Edwards curve $\mathrm{E}_{\mathrm{E},a,d}$ is a quadratic twist of the Edwards curve $\mathrm{E}_{\mathrm{E},1,d/a}$.

Five slightly different ways to build a set of points from an Edwards curve, or more generally a twisted Edwards curve, have appeared in the literature. The simplest is the set of affine points $\{(x, y) \in \mathbf{A}^2 : ax^2 + y^2 = 1 + dx^2y^2\}$. Four others, with various theoretical and computational advantages, are the projective, inverted, extended, and completed sets discussed below.

2.2. **The Edwards addition law.** The Edwards addition law on $\mathrm{E}_{\mathrm{E},a,d}$ is given in affine coordinates by

$$(x_1, y_1), (x_2, y_2) \mapsto \left(\frac{x_1 y_2 + y_1 x_2}{1 + dx_1 x_2 y_1 y_2}, \frac{y_1 y_2 - ax_1 x_2}{1 - dx_1 x_2 y_1 y_2} \right).$$

By inserting appropriate denominators one obtains the Edwards addition law in projective coordinates, inverted coordinates, extended coordinates, and completed coordinates.

The Edwards addition law is strongly unified; i.e., the same formulas can also be used for generic doublings. The point $(0, 1)$ is the neutral element of the addition law. The negative of a point (x_1, y_1) is $(-x_1, y_1)$.

The Edwards addition law for $E_{E,a,d}$ was studied by Bernstein, Birkner, Joye, Lange, and Peters in [8], generalizing from the case $a = 1$ studied by Bernstein and Lange in [13], generalizing from the case $a = 1$, $d = c^4$ studied by Edwards in [25], generalizing from the case $a = 1$, $d = -1$ studied by Euler and Gauss.

Edwards actually used the form $x^2 + y^2 = c^2(1 + x^2y^2)$. Edwards showed that every elliptic curve over $\mathbf{Q}$ can be written in this normal form over an extension of $\mathbf{Q}$. Replacing (x, y) with (cx, cy) produces the curve $E_{E,1,c^4}$; this scaling turns out to save time in computations. The further generalizations to $E_{E,1,d}$ and to $E_{E,a,d}$ allow more curves over $\mathbf{Q}$ to be handled at similar speeds.

2.3. The dual addition law.
Hisil, Wong, Carter, and Dawson in [29] introduced the addition law

$$(x_1, y_1), (x_2, y_2) \mapsto \left(\frac{x_1 y_1 + x_2 y_2}{y_1 y_2 + a x_1 x_2}, \frac{x_1 y_1 - x_2 y_2}{x_1 y_2 - y_1 x_2} \right)$$

on $E_{E,a,d}$. This dual addition law produces the same output as the Edwards addition law when both are defined, but the exceptional cases are different. In particular, the dual addition law never works for doublings: if $(x_1, y_1) = (x_2, y_2)$, then the second output coordinate $(x_1 y_1 - x_2 y_2)/(x_1 y_2 - y_1 x_2)$ is $0/0$. The dual addition law nevertheless has some important advantages, as discussed below.

2.4. Projective points.
The *projective* twisted Edwards curve is

$$\left\{ (X : Y : Z) \in \mathbf{P}^2 : a X^2 Z^2 + Y^2 Z^2 = Z^4 + d X^2 Y^2 \right\}.$$

The projective points are the affine points (x_1, y_1), embedded as usual into $\mathbf{P}^2$ by $(x_1, y_1) \mapsto (x_1 : y_1 : 1)$, and two extra singular points at infinity, namely $(0 : 1 : 0)$ and $(1 : 0 : 0)$.

Fast projective addition and doubling formulas, starting from the Edwards addition law and eliminating multiplications in various ways, were introduced for Edwards curves in [13] and were generalized to twisted Edwards curves in [8]. Adding a generic pair of points uses just $10\mathbf{M} + 1\mathbf{S} + 1\mathbf{A} + 1\mathbf{D}$: i.e., 10 field multiplications, 1 field squaring, 1 multiplication by a, and 1 multiplication by d. Doubling takes just $3\mathbf{M} + 4\mathbf{S} + 1\mathbf{A}$ with the following formulas:

$$\begin{aligned} B &= (X_1 + Y_1)^2; \ C = X_1^2; \ D = Y_1^2; \ E = aC; \ F = E + D; \ H = Z_1^2; \\ J &= F - 2H; \ X_3 = (B - C - D) \cdot J; \ Y_3 = F \cdot (E - D); \ Z_3 = F \cdot J. \end{aligned}$$

These doubling formulas are used in EECM-MPFQ.

2.5. Inverted points.
The *inverted* twisted Edwards curve is

$$\left\{ (X : Y : Z) \in \mathbf{P}^2 : a Y^2 Z^2 + X^2 Z^2 = X^2 Y^2 + d Z^4 \right\}.$$

The inverted points are the affine points (x_1, y_1) other than $(0, \pm 1)$ and $(\pm 1, 0)$, embedded into $\mathbf{P}^2$ by $(x_1, y_1) \mapsto (1/x_1 : 1/y_1 : 1)$; two extra points if d is a square, namely $(\pm\sqrt{d} : 0 : 1)$; two extra points if d/a is a square, namely $(0 : \pm\sqrt{d/a} : 1)$; and two singular points at infinity, namely $(0 : 1 : 0)$ and $(1 : 0 : 0)$. Note that a generic inverted point $(X_1 : Y_1 : Z_1)$ corresponds to the affine point $(Z_1/X_1, Z_1/Y_1)$.

Fast inverted addition and doubling formulas were introduced for Edwards curves in [14] and for twisted Edwards curves in [8]. Adding a generic pair of points costs only $9\mathbf{M} + 1\mathbf{S} + 1\mathbf{A} + 1\mathbf{D}$, saving $1\mathbf{M}$ compared to projective Edwards coordinates. A doubling costs $3\mathbf{M} + 4\mathbf{S} + 1\mathbf{A} + 1\mathbf{D}$, losing $1\mathbf{D}$ compared to projective Edwards coordinates.

These formulas were used in the prototype GMP-EECM.

2.6. Extended points.

The *extended* twisted Edwards curve is

$$\left\{ (X : Y : Z : T) \in \mathbf{P}^3 : aX^2 + Y^2 = Z^2 + dT^2 \text{ and } XY = ZT \right\}.$$

The extended points are the affine points (x_1, y_1), embedded into $\mathbf{P}^3$ by $(x_1, y_1) \mapsto (x_1 : y_1 : 1 : x_1 y_1)$; two extra points at infinity if d is a square, namely $(0 : \pm\sqrt{d} : 0 : 1)$; and two extra points at infinity if d/a is a square, namely $(1 : 0 : 0 : \pm\sqrt{a/d})$.

Hisil, Wong, Carter, and Dawson in [29] introduced extended addition formulas costing only $9\mathbf{M} + 1\mathbf{A}$:

$$
\begin{aligned}
A &= X_1 \cdot X_2, \ B = Y_1 \cdot Y_2, \ C = Z_1 \cdot T_2, \ D = T_1 \cdot Z_2, \\
E &= D + C, \ F = (X_1 - Y_1) \cdot (X_2 + Y_2) + B - A, \ G = B + aA, \\
H &= D - C, \ X_3 = E \cdot F, \ Y_3 = G \cdot H, \ Z_3 = F \cdot G, \ T_3 = E \cdot H.
\end{aligned}
$$

These formulas save $1\mathbf{S}$ by switching from inverted coordinates to extended coordinates, and an extra $1\mathbf{D}$ by switching from the Edwards addition law to the dual addition law. These formulas are used in EECM-MPFQ. Our main concern in this paper is with the case $a = 1$, where these formulas cost only $9\mathbf{M}$. Hisil et al. also introduced addition formulas costing only $8\mathbf{M}$ for the case $a = -1$; but we show in Section 6 that the case $a = -1$ sacrifices torsion.

A doubling in extended coordinates loses $1\mathbf{M}$ for computing the extended output coordinate T_3. However, the doubling formulas make no use of the extended input coordinate T_1, so if the input is not used for anything else then the operation *producing* that input can skip the computation of T_1, saving $1\mathbf{M}$.

Scalar multiplication can be carried out as a series of operations on an accumulator P: doublings replace P by $2P$, and double-and-add operations replace P by $2P + Q$. If P is in projective coordinates and the precomputed points Q are in extended coordinates, then doubling costs $3\mathbf{M} + 4\mathbf{S} + 1\mathbf{A}$ and double-and-add costs $(3\mathbf{M} + 4\mathbf{S} + 1\mathbf{A}) + (9\mathbf{M} + 1\mathbf{A})$, with the $1\mathbf{M}$ loss in doubling cancelled by the $1\mathbf{M}$ savings in addition. This mixture of projective coordinates and extended coordinates was suggested in [29] and is used in EECM-MPFQ.

2.7. Completed points.

The *completed* twisted Edwards curve is

$$\overline{\mathrm{E}}_{\mathrm{E},a,d} = \left\{ ((X : Z), (Y : T)) \in \mathbf{P}^1 \times \mathbf{P}^1 : aX^2 T^2 + Y^2 Z^2 = Z^2 T^2 + dX^2 Y^2 \right\}.$$

The completed points are the affine points (x_1, y_1), embedded as usual into $\mathbf{P}^1 \times \mathbf{P}^1$ by $(x_1, y_1) \mapsto ((x_1 : 1), (y_1 : 1))$; two extra points at infinity if d is a square, namely $((1 : \pm\sqrt{d}), (1 : 0))$; and two extra points at infinity if d/a is a square, namely $((1 : 0), (\pm\sqrt{a/d} : 1))$. Starting in Section 6 we freely abbreviate $((x_1 : 1), (y_1 : 1))$ as (x_1, y_1).

The completed curve maps isomorphically to the extended curve via the Segre embedding $((X : Z), (Y : T)) \mapsto (XT : YZ : ZT : XY)$ of $\mathbf{P}^1 \times \mathbf{P}^1$ into $\mathbf{P}^3$. It maps onto the projective curve via $((X : Z), (Y : T)) \mapsto (XT : YZ : ZT)$, but this map is not an isomorphism: it sends the two points $((1 : \pm\sqrt{d}), (1 : 0))$ to $(0 : 1 : 0)$, and sends the two points $((1 : 0), (\pm\sqrt{a/d} : 1))$ to $(1 : 0 : 0)$. The completed curve also maps onto the inverted curve via $((X : Z), (Y : T)) \mapsto (YZ : XT : XY)$, but this map sends the two points $((0 : 1), (\pm 1 : 1))$ to $(1 : 0 : 0)$, and sends the two points $((\pm 1 : 1), (0 : 1))$ to $(0 : 1 : 0)$.

EECM-MPFQ uses the completed curve as an intermediate output of doublings (costing $4\mathbf{S} + 1\mathbf{A}$) and additions (costing $5\mathbf{M} + 1\mathbf{A}$); it then maps the completed point to a projective point (costing $3\mathbf{M}$) or to an extended point (costing $4\mathbf{M}$) as desired. One should not think that *all* addition formulas in the literature naturally factor through the completed curve: in particular, a detour through the completed curve would sacrifice $1\mathbf{M}$ in the inverted Edwards addition law and in the projective dual addition law.

2.8. **Addition with small inputs.** There are two compatible ways to choose "small" curves that save more time in scalar multiplication. First, choosing small curve parameters a, d speeds up any multiplications by those parameters inside addition formulas and doubling formulas. Second, choosing a small base point P_1 for scalar multiplication speeds up multiplications by the coordinates of P_1, and to some extent speeds up multiplications by the coordinates of $[3]P_1$ etc.

Let $P_1 = (x_1, y_1)$ be a rational point on the Edwards curve $\mathrm{E}_{\mathrm{E},1,\bar{d}}$, and assume that $x_1, y_1, \bar{d}$ have small height, i.e., small numerators and denominators. Then $\bar{d}$ can be written in the form d/a, where a is a small square and d is a small integer. Now the point $(x_1/\sqrt{a}, y_1)$ is on the isomorphic curve $\mathrm{E}_{\mathrm{E},a,d}$ and can be written with small integer coordinates on the inverted curve, the extended curve, etc., saving time in addition. A small inverted point $(X_1 : Y_1 : Z_1)$ replaces $4\mathbf{M}$ by 4 small multiplications, specifically 1 multiplication by each of the small integers $X_1, Y_1, X_1 + Y_1$, and Z_1; similarly, a small extended point replaces $5\mathbf{M}$ by 5 small multiplications.

2.9. **The Edwards group.** If $a = 1$ and d is not a square then, by [13, Theorem 3.3], the affine Edwards addition law is complete: the denominators $1 + dx_1x_2y_1y_2$ and $1 - dx_1x_2y_1y_2$ are always nonzero, and the affine points (x_1, y_1) on the curve form a group.

However, if d is a square, then the addition law is not necessarily a group law: there can be pairs (x_1, y_1) and (x_2, y_2) where $1 + dx_1x_2y_1y_2 = 0$ or $1 - dx_1x_2y_1y_2 = 0$. Even worse, there can be pairs (x_1, y_1) and (x_2, y_2) for which $1 + dx_1x_2y_1y_2 = 0 = x_1y_2 + y_1x_2$ or $1 - dx_1x_2y_1y_2 = 0 = y_1y_2 - ax_1x_2$. Switching from affine coordinates to projective or inverted or extended or completed coordinates does not allow the Edwards addition law to add such points.

Fortunately, on the completed curve $\overline{\mathrm{E}}_{\mathrm{E},a,d}$ in $\mathbf{P}^1 \times \mathbf{P}^1$, any pair of input points that cannot be added by the Edwards addition law can be added by the dual addition law. Gluing together the two addition laws produces a group law $+ : \overline{\mathrm{E}}_{\mathrm{E},a,d} \times \overline{\mathrm{E}}_{\mathrm{E},a,d} \to \overline{\mathrm{E}}_{\mathrm{E},a,d}$. This is the main result of [16]: the Edwards addition law and the dual addition law form a complete system of addition laws for $\overline{\mathrm{E}}_{\mathrm{E},a,d}$.

The following theorem summarizes the results from [16]. Section 3 uses this group law to characterize points of small order in $\overline{\mathrm{E}}_{\mathrm{E},a,d}$, and subsequent sections of this paper use this characterization to construct Edwards curves with large $\mathbf{Q}$-torsion groups.

Theorem 2.10. *Fix a field k with $\mathrm{char}(k) \neq 2$. Fix distinct nonzero elements $a, d \in k$. Fix $P_1, P_2 \in \overline{\mathrm{E}}_{\mathrm{E},a,d}(k)$. Write P_1 as $((X_1 : Z_1), (Y_1 : T_1))$ and write P_2 as*

$((X_2 : Z_2), (Y_2 : T_2))$. *Define*

$$X_3 = X_1 Y_2 Z_2 T_1 + X_2 Y_1 Z_1 T_2,$$
$$Z_3 = Z_1 Z_2 T_1 T_2 + d X_1 X_2 Y_1 Y_2,$$
$$Y_3 = Y_1 Y_2 Z_1 Z_2 - a X_1 X_2 T_1 T_2,$$
$$T_3 = Z_1 Z_2 T_1 T_2 - d X_1 X_2 Y_1 Y_2;$$

and

$$X_3' = X_1 Y_1 Z_2 T_2 + X_2 Y_2 Z_1 T_1,$$
$$Z_3' = a X_1 X_2 T_1 T_2 + Y_1 Y_2 Z_1 Z_2,$$
$$Y_3' = X_1 Y_1 Z_2 T_2 - X_2 Y_2 Z_1 T_1,$$
$$T_3' = X_1 Y_2 Z_2 T_1 - X_2 Y_1 Z_1 T_2.$$

Then $X_3 Z_3' = X_3' Z_3$ and $Y_3 T_3' = Y_3' T_3$. Furthermore, at least one of the following cases occurs:

- $(X_3, Z_3) \neq (0,0)$ *and* $(Y_3, T_3) \neq (0,0)$. *Then* $P_1 + P_2 = ((X_3 : Z_3), (Y_3 : T_3))$.
- $(X_3', Z_3') \neq (0,0)$ *and* $(Y_3', T_3') \neq (0,0)$. *Then* $P_1 + P_2 = ((X_3' : Z_3'), (Y_3' : T_3'))$.

If $P_1 = P_2$, then the first case occurs.

A completely different way to define the same group law on $\overline{\mathrm{E}}_{\mathrm{E},a,d}$ is through a correspondence to the traditional chord-and-tangent elliptic-curve group law on an isomorphic Weierstrass curve. An explicit group isomorphism appears in [16, Theorems 7.1 and 7.3]. This correspondence has the disadvantage of requiring more case distinctions and more computations, but the advantage of making clear that the Edwards group is an elliptic-curve group. We use this to apply Mazur's theorem starting in Section 6.

3. Points of small order on $\overline{\mathrm{E}}_{\mathrm{E}}, a, d$

The complete set of addition laws from [16] (presented in the previous section) enables us to investigate the order of any point. In particular, it has often been stated that the points at infinity on an Edwards curve blow up to two points of order 2 and two points of order 4, e.g. in [8] in the context of exceptional points of the map between a twisted Edwards curve and a Montgomery curve. With the complete set of addition laws we can prove all statements purely in the context of Edwards curves.

This section characterizes all points of order 2, 3, and 4, and states conditions on the parameters of the twisted Edwards curve for such points to exist. These results are used later to construct curves with large **Q**-torsion groups. This section also characterizes points of order 8 relevant to later sections.

The following theorem gives a complete study of points of order 2 and 4 in $\overline{\mathrm{E}}_{\mathrm{E},a,d}$.

Theorem 3.1. *Fix a field k with $\mathrm{char}(k) \neq 2$. Fix distinct nonzero elements $a, d \in k$. The following points are in $\overline{\mathrm{E}}_{\mathrm{E},a,d}(k)$ and have the stated orders.*

Points of order 2:

The point $((0:1), (-1:1))$ has order 2.

If a/d is a square in k, then the points $((1:0), (\pm\sqrt{a/d} : 1))$ have order 2.

There are no other points of order 2.

Points of order 4 **doubling to** $((0 : 1), (-1 : 1))$**:**

If a is a square in k, then the points $((1 : \pm\sqrt{a}), (0 : 1))$ have order 4 and double to $((0 : 1), (-1 : 1))$.

If d is a square in k, then the points $((1 : \pm\sqrt{d}), (1 : 0))$ have order 4 and double to $((0 : 1), (-1 : 1))$.

There are no other points doubling to $((0 : 1), (-1 : 1))$.

Points of order 4 **doubling to** $((1 : 0), (\pm\sqrt{a/d} : 1))$**:** *Assume that $s \in k$ satisfies $s^2 = a/d$.*

If s and $-s/a$ are squares in k, then the points $((\pm\sqrt{-s/a} : 1), (\pm\sqrt{s} : 1))$, where the signs may be chosen independently, have order 4 and double to $((1 : 0), (s : 1))$.

There are no other points doubling to $((1 : 0), (s : 1))$.

Proof. Doublings can always be computed by X_3, Z_3, Y_3, T_3 from Theorem 2.10: in other words, all curve points $((X : Z), (Y : T))$ have $(2XYZT, Z^2T^2 + dX^2Y^2) \neq (0, 0)$ and $(Y^2Z^2 - aX^2T^2, Z^2T^2 - dX^2Y^2) \neq (0, 0)$, so

$$[2]((X : Z), (Y : T))$$
$$= ((2XYZT : Z^2T^2 + dX^2Y^2), (Y^2Z^2 - aX^2T^2 : Z^2T^2 - dX^2Y^2)).$$

In particular:

- $[2]((0 : 1), (-1 : 1)) = ((0 : 1), (1 : 1))$.
- $[2]((1 : 0), (\pm\sqrt{a/d} : 1)) = ((0 : \ldots), (-a : -d(a/d))) = ((0 : 1), (1 : 1))$.
- $[2]((1 : \pm\sqrt{a}), (0 : 1)) = ((0 : \ldots), (-a : a)) = ((0 : 1), (-1 : 1))$.
- $[2]((1 : \pm\sqrt{d}), (1 : 0)) = ((0 : \ldots), (d : -d)) = ((0 : 1), (-1 : 1))$.
- $[2]((\pm\sqrt{-s/a} : 1), (\pm\sqrt{s} : 1)) = ((\ldots : 1 + d(-s/a)s), (s - a(-s/a) : 1 - d(-s/a)s)) = ((1 : 0), (s : 1))$ since $d(s/a)s = s^2d/a = 1$.

To see that there is no other point of order 2 or 4, observe first that every point $((X : Z), (Y : T))$ on $\overline{\mathrm{E}}_{\mathrm{E},a,d}$ with $X = 0$ or $Y = 0$ or $Z = 0$ or $T = 0$ is either $((0 : 1), (1 : 1))$ or one of the points doubled above. The only remaining points are affine points $((x : 1), (y : 1))$ with $x \neq 0$ and $y \neq 0$. The double of $((x : 1), (y : 1))$ is $((2xy : 1 + dx^2y^2), (y^2 - ax^2 : 1 - dx^2y^2))$; but $2xy \neq 0$, so this double cannot be $((0 : 1), (1 : 1))$, so $((x : 1), (y : 1))$ cannot have order 2. For the same reason, the double cannot be $((0 : 1), (-1 : 1))$. The only remaining case is that the double is $((1 : 0), (s : 1))$ where $s^2 = a/d$. Then $ax^2 + y^2 = 1 + dx^2y^2 = 0$ so $ax^2 = -y^2$; and $y^2 - ax^2 = s(1 - dx^2y^2)$, so $2y^2 = y^2 - ax^2 = s(1 - dx^2y^2) = 2s$, so $y = \pm\sqrt{s}$, and finally $ax^2 = -s$ so $x = \pm\sqrt{-s/a}$. □

Later we will study Edwards curves over the rationals **Q** for which $((1 : \pm\sqrt{a}), (0 : 1))$ is on the curve. In this case the only points of order 8 double to either these points or to $((1 : \pm\sqrt{d}), (1 : 0))$.

Theorem 3.2. *Fix a field k with* $\mathrm{char}(k) \neq 2$. *Fix distinct nonzero elements $a, d \in k$.*

Points of order 8 **doubling to** $((1 : \pm\sqrt{a}), (0 : 1))$**:** *If $r \in k$ satisfies $r^2 = a$, then any element of $\overline{\mathrm{E}}_{\mathrm{E},a,d}(k)$ doubling to $((1 : r), (0 : 1))$ can be written as $((x_8 : 1), (rx_8 : 1))$ for some $x_8 \in k$ satisfying $adx_8^4 - 2ax_8^2 + 1 = 0$.*

Conversely, if $r, x_8 \in k$ satisfy $r^2 = a$ and $adx_8^4 - 2ax_8^2 + 1 = 0$, then the two points $((\pm x_8 : 1), (\pm rx_8 : 1))$, with matching signs, have order 8 and double to $((1 : r), (0 : 1))$. If also d is a square in k, then the two points $((1 : \pm rx_8\sqrt{d}), (1 : \pm x_8\sqrt{d}))$,

with matching signs, have order 8, double to $((1:r),(0:1))$, and are different from $((\pm x_8:1),(\pm rx_8:1))$. There are no other points doubling to $((1:r),(0:1))$.

Points of order 8 doubling to $((1:\pm\sqrt{d}),(1:0))$**:** If $s\in k$ satisfies $s^2 = d$, then any element of $\overline{\mathrm{E}}_{\mathrm{E},a,d}(k)$ doubling to $((1:s),(1:0))$ can be written as $((\bar{x}_8:1),(1:s\bar{x}_8))$ for some $\bar{x}_8\in k$ satisfying $ad\bar{x}_8^4 - 2d\bar{x}_8^2 + 1 = 0$.

Conversely, if $s,\bar{x}_8\in k$ satisfy $s^2 = d$ and $ad\bar{x}_8^4 - 2d\bar{x}_8^2 + 1 = 0$, then the two points $((\pm\bar{x}_8:1),(1:\pm s\bar{x}_8))$, with matching signs, have order 8 and double to $((1:s),(1:0))$. If also a is a square in k, then the two points $((1:\pm sx_8\sqrt{a}),(\pm x_8\sqrt{a}:1))$, with matching signs, have order 8, double to $((1:s),(1:0))$, and are different from $((\pm\bar{x}_8:1),(1:\pm s\bar{x}_8))$. There are no other points doubling to $((1:s),(1:0))$.

Proof. Every point with a zero coordinate has order at most 4 by Theorem 3.1, so any point of order 8 has the form $((x_8:1),(y_8:1))$, with $x_8\neq 0$ and $y_8\neq 0$, and with double $((2x_8y_8:1+dx_8^2y_8^2),(y_8^2-ax_8^2:1-dx_8^2y_8^2))$.

Part 1: If the double is $((1:r),(0:1))$, then $y_8^2 - ax_8^2 = 0$ and $2x_8y_8r = 1 + dx_8^2y_8^2 = ax_8^2 + y_8^2 = 2ax_8^2 = 2r^2x_8^2$. Cancel $2x_8r$ to see that $y_8 = rx_8$. Hence $adx_8^4 - 2ax_8^2 + 1 = dx_8^2y_8^2 - (1+dx_8^2y_8^2) + 1 = 0$ and the original point is $((x_8:1),(rx_8:1))$.

Conversely, if $r,x_8\in k$ satisfy $r^2 = a$ and $adx_8^4 - 2ax_8^2 + 1 = 0$, then the point $((x_8:1),(rx_8:1))$ is on the curve since $ax_8^2 + (rx_8)^2 = 2ax_8^2 = adx_8^4 + 1 = 1 + dx_8^2(rx_8)^2$, and it doubles to $((2x_8rx_8:1+dx_8^2r^2x_8^2),(r^2x_8^2-ax_8^2:\ldots)) = ((2x_8rx_8:2ax_8^2),(0:\ldots)) = ((1:r),(0:1))$.

The other points doubling to $((1:r),(0:1))$ are $((x:1),(rx:1))$ for other $x\in k$ satisfying $adx^4 - 2ax^2 + 1 = 0$. If d is not a square in k, then $adx^4 - 2ax^2 + 1 = adx^4 - (adx_8^2 + 1/x_8^2)x^2 + 1 = (x-x_8)(x+x_8)(adx^2 - 1/x_8^2)$, with $adx^2 - 1/x_8^2$ irreducible, so the only points doubling to $((1:r),(0:1))$ are $((\pm x_8:1),(\pm rx_8:1))$. If d is a square in k, then $adx^4 - 2ax^2 + 1 = (x-x_8)(x+x_8)(rx\sqrt{d}-1/x_8)(rx\sqrt{d}+1/x_8)$ so the only points doubling to $((1:r),(0:1))$ are $((\pm x_8:1),(\pm rx_8:1))$ and $((1:\pm rx_8\sqrt{d}),(1:\pm x_8\sqrt{d}))$. These points are distinct: otherwise $\pm rx_8^2\sqrt{d} = 1$, so $adx_8^4 = 1$, so $2ax_8^2 = 2$, so $ax_8^2 = 1$, so $y_8 = 0$ from the curve equation, contradiction.

Part 2: If the double of $((\bar{x}_8:1),(\bar{y}_8:1))$ is $((1:s),(1:0))$, then $1 - d\bar{x}_8^2\bar{y}_8^2 = 0$ and $2\bar{x}_8\bar{y}_8s = 1 + d\bar{x}_8^2\bar{y}_8^2 = 2$ so $\bar{y}_8 = 1/(s\bar{x}_8)$. Hence $ad\bar{x}_8^4 - 2d\bar{x}_8^2 + 1 = (a\bar{x}_8^2 - 2 + \bar{y}_8^2)d\bar{x}_8^2 = 0$ and the original point is $((\bar{x}_8:1),(1:s\bar{x}_8))$.

Conversely, if $s,\bar{x}_8\in k$ satisfy $s^2 = d$ and $ad\bar{x}_8^4 - 2d\bar{x}_8^2 + 1 = 0$, then the point $((\bar{x}_8:1),(1:s\bar{x}_8))$ is on the curve since $d\bar{x}_8^2(a\bar{x}_8^2 + \bar{y}_8^2) = d\bar{x}_8^2(a\bar{x}_8^2 + 1/(s^2\bar{x}_8^2)) = ad\bar{x}_8^4 + 1 = 2d\bar{x}_8^2 = d\bar{x}_8^2 + d\bar{x}_8^4/\bar{x}_8^2 = d\bar{x}_8^2(1 + d\bar{x}_8^2/(s^2\bar{x}_8^2)) = d\bar{x}_8^2(1 + d\bar{x}_8^2\bar{y}_8^2)$. The point doubles to $((2s\bar{x}_8^2:s^2\bar{x}_8^2+d\bar{x}_8^2),(1-as^2\bar{x}_8^4:s^2\bar{x}_8^2-d\bar{x}_8^2)) = ((1:s),(1-ad\bar{x}_8^4:s^2\bar{x}_8^2-s^2\bar{x}_8^2)) = ((1:s),(1:0))$.

The other points doubling to $((1:s),(1:0))$ are $((x:1),(1:sx))$ for other $x\in k$ satisfying $adx^4 - 2dx^2 + 1 = 0$. If a is not a square in k, then $adx^4 - 2dx^2 + 1 = adx^4 - (ad\bar{x}_8^2 + 1/\bar{x}_8^2)x^2 + 1 = (x-\bar{x}_8)(x+\bar{x}_8)(adx^2 - 1/\bar{x}_8^2)$, with $adx^2 - 1/\bar{x}_8^2$ irreducible, so the only points doubling to $((1:s),(1:0))$ are $((\pm\bar{x}_8:1),(1:\pm s\bar{x}_8))$. If a is a square in k, then $adx^4 - 2dx^2 + 1 = (x-\bar{x}_8)(x+\bar{x}_8)(sx\sqrt{a}-1/\bar{x}_8)(sx\sqrt{a}+1/\bar{x}_8)$ so the only points doubling to $((1:s),(1:0))$ are $((\pm\bar{x}_8:1),(1:\pm s\bar{x}_8))$ and $((1:\pm s\bar{x}_8\sqrt{a}),(\pm\bar{x}_8\sqrt{a}:1))$. These points are distinct: otherwise $\pm s\bar{x}_8^2\sqrt{a} = 1$ so $ad\bar{x}_8^4 = 1$ so $2d\bar{x}_8^2 = 2$ so $d\bar{x}_8^2 = 1$ so $\bar{x}_8 = 0$ from the curve equation, contradiction. □

Theorem 3.3. *Fix a field k with $\mathrm{char}(k) \neq 2$. Fix distinct nonzero elements $a, d \in k$. If $x_3, y_3 \in k$ satisfy $ax_3^2 + y_3^2 = 1 + dx_3^2 y_3^2 = -2y_3$, then $((x_3 : 1), (y_3 : 1))$ is a point of order 3 on $\overline{\mathrm{E}}_{E,a,d}(k)$. Conversely, all points of order 3 on $\overline{\mathrm{E}}_{E,a,d}(k)$ arise in this way.*

Proof. Doublings can always be computed by X_3, Z_3, Y_3, T_3 from Theorem 2.10, as in the proof of Theorem 3.1.

Observe that $((x_3 : 1), (y_3 : 1)) \in \overline{\mathrm{E}}_{E,a,d}(k)$ since $ax_3^2 + y_3^2 = 1 + dx_3^2 y_3^2$. Now

$$[2]((x_3 : 1), (y_3 : 1)) = ((2x_3 y_3 : 1 + dx_3^2 y_3^2), (y_3^2 - ax_3^2 : 1 - dx_3^2 y_3^2))$$
$$= ((2x_3 y_3 : -2y_3), (2y_3^2 + 2y_3 : 2y_3 + 2))$$
$$= ((-x_3 : 1), (y_3 : 1))$$

so $((x_3 : 1), (y_3 : 1))$ has order dividing 3. It cannot have order 1 (since otherwise $x_3 = 0$ so $y_3^2 = 1 = -2y_3$), so it has order 3.

Conversely, consider any point $P = ((X_1 : Z_1), (Y_1 : T_1))$ of order 3 in $\overline{\mathrm{E}}_{E,a,d}(k)$. The equation $[2]P = -P$ then implies $(2X_1 Y_1 Z_1 T_1 : Z_1^2 T_1^2 + dX_1^2 Y_1^2) = (-X_1 : Z_1)$. Every point in $\overline{\mathrm{E}}_{E,a,d}$ with a zero coordinate has order 1, 2, or 4 by Theorem 3.1, so $X_1, Z_1, Y_1, T_1 \neq 0$. Define $x_3 = X_1/Z_1$ and $y_3 = Y_1/T_1$. Then $P = ((x_3 : 1), (y_3 : 1))$; furthermore $(2x_3 y_3 : 1 + dx_3^2 y_3^2) = (-x_3 : 1)$ and $x_3 \neq 0$ so $-2y_3 = 1 + dx_3^2 y_3^2 = ax_3^2 + y_3^2$. $\qquad \square$

4. Using Edwards curves in ECM stage 1

This section discusses "stage 1" of ECM. It begins by reviewing the general idea of stage 1 and the state-of-the-art strategies used in GMP-ECM to perform the elliptic-curve computations in stage 1. It then analyzes the speedups obtained from using Edwards curves.

4.1. **Overview of stage 1.** Stage 1 of ECM tries to factor a positive integer n as follows. Choose an elliptic curve E defined over $\mathbf{Q}$, and a nontorsion element $P \in E(\mathbf{Q})$. Choose a rational function $\phi : E \to \mathbf{Q}$ that has a pole at the neutral element of E; for example, choose ϕ as the Weierstrass x-coordinate. Choose a positive integer s with many small prime factors. Choose a sequence of additions, subtractions, multiplications, and divisions that, if carried out over $\mathbf{Q}$, would compute $\phi([s]P)$, where $[s]P$ denotes the sth multiple of P in $E(\mathbf{Q})$. Compute $\phi([s]P)$ modulo n by carrying out this sequence of additions, subtractions, multiplications, and divisions modulo n. Hope for an impossible division modulo n. An attempt to divide by a nonzero nonunit modulo n immediately reveals a factor of n; an attempt to divide by 0 modulo n is not quite as informative but usually allows a factor of n to be obtained without much extra work.

If n has a prime divisor q such that $[s]P$ is the neutral element of $E(\mathbf{Z}/q\mathbf{Z})$, then the stage-1 ECM computation will involve an impossible division modulo n, usually revealing a factor of n. This occurs, in particular, whenever s is a multiple of the group size $\#E(\mathbf{Z}/q\mathbf{Z})$. As E varies randomly, $\#E(\mathbf{Z}/q\mathbf{Z})$ varies randomly (with some subtleties in its distribution; see, e.g., [35]) in the Hasse interval $[q - 2\sqrt{q} + 1, q + 2\sqrt{q} + 1]$. What makes ECM useful is that a surprisingly small s, allowing a surprisingly fast computation of $[s]P$, is a multiple of a surprisingly large fraction of the integers in the Hasse interval, and is a multiple of the order of P modulo q with (conjecturally) an even larger probability. See Section 9 for detailed statistics.

For example, one could try to factor n as follows. Choose the curve $E : y^2 = x^3 - 2$, the Weierstrass x-coordinate as ϕ, the point $(x, y) = (3, 5)$, and the integer $s = 420$. Choose the following strategy to compute the x-coordinate of $[420](3, 5)$: use the standard affine-coordinate doubling formulas to compute $[2](3, 5)$, then $[4](3, 5)$, then $[8](3, 5)$; use the standard affine-coordinate addition formulas to compute $[12](3, 5)$; continue similarly through $[2](3, 5)$, $[4](3, 5)$, $[8](3, 5)$, $[12](3, 5)$, $[24](3, 5)$, $[48](3, 5)$, $[96](3, 5)$, $[192](3, 5)$, $[384](3, 5)$, $[408](3, 5)$, $[420](3, 5)$. Carry out these computations modulo n, hoping for a division by a nonzero nonunit modulo n.

The denominator of the x-coordinate of $[420](3, 5)$ in $E(\mathbf{Q})$ has many small prime factors: 2, 3, 5, 7, 11, 19, 29, 31, 41, 43, 59, 67, 71, 83, 89, 109, 163, 179, 181, 211, 223, 241, 269, 283, 383, 409, 419, 433, 523, 739, 769, 811, 839, etc. If n shares any of these prime factors, then the computation of $[420](3, 5)$ will encounter an impossible division modulo n. To verify the presence of (e.g.) the primes 769, 811, and 839, one can observe that $[420](3, 5)$ is the neutral element in each of the groups $E(\mathbf{Z}/769\mathbf{Z})$, $E(\mathbf{Z}/811\mathbf{Z})$, $E(\mathbf{Z}/839\mathbf{Z})$; the order of $(3, 5)$ turns out to be 7, 42, 35, respectively. Note that the group orders are 819, 756, and 840, none of which divide 420.

4.2. The standard choice of s.
Pollard in [40, page 527] suggested choosing s as "the product of all the primes $p_i \leq L$ each to some power $c_i \geq 1$. There is some freedom in the choice of the c_i but the smallest primes should certainly occur to some power higher than the first."

Pollard's prime bound "L" is now called B_1. One possibility is to choose, for each prime $\pi \leq B_1$, the largest power of π in the interval $[1, n + 2\sqrt{n} + 1]$. Then $[s]P$ is the neutral element in $E(\mathbf{Z}/q\mathbf{Z})$ if and only if the order of P is "B_1-smooth", i.e., if and only if the order has no prime divisors larger than B_1. This possibility is theoretically pleasing but clearly suboptimal.

Brent in [18, Section 5] said that "in practice we choose" the largest power of π in the interval $[1, B_1]$ "because this significantly reduces the cost of a trial without significantly reducing the probability of success." In other words, $s = \mathrm{lcm}\{1, 2, \ldots, B_1\}$. GMP-ECM uses the same strategy; see [47, page 529].

4.3. The standard prime-by-prime strategy.
Pollard in [40, page 527] said that one "can choose between using the primes p_i in succession or computing P in advance and performing a single power operation." Pollard's "P" is s in the notation of this paper.

As far as we know, all ECM implementations use the first strategy, working with one prime at a time. Brent in [18, Section 5] wrote, "Actually, E [i.e., s in the notation of this paper] is not computed. Instead ... repeated operations of the form $P := P^k$ [i.e., $[k]P$ in the notation of this paper], where k ... is a prime power." Montgomery in [36, page 249] wrote, "It is unnecessary to compute R [i.e., s in the notation of this paper] explicitly." Zimmermann and Dodson in [47, page 529] wrote, "That big product is not computed as such" and presented the prime-by-prime loop used in GMP-ECM.

4.4. The standard elliptic-curve coordinate system.
Chudnovsky and Chudnovsky in [20, Section 4] wrote, "The crucial problem becomes the choice of the model of an algebraic group variety, where computations mod p are the least time

consuming." They presented explicit formulas for computations on several different shapes of elliptic curves.

Montgomery in [36, Section 10.3.1] introduced what are now called "Montgomery coordinates": a point (x_1, y_1) on the elliptic curve $By^2 = x^3 + Ax^2 + x$ is represented as a pair $(X_1 : Z_1)$ such that $X_1/Z_1 = x_1$. This representation does not distinguish (x_1, y_1) from $(x_1, -y_1)$, so it does not allow addition, but it does allow "differential addition," i.e., computation of $P + Q$ given $P, Q, P - Q$. In particular, Montgomery presented explicit formulas to compute $P, [2k]P, [(2k+1)]P$ from $P, [k]P, [k+1]P$ using $6\mathbf{M} + 4\mathbf{S} + 1\mathbf{C}$, or $5\mathbf{M} + 4\mathbf{S} + 1\mathbf{C}$ if P is given with $Z_1 = 1$, or $4\mathbf{M} + 4\mathbf{S} + 1\mathbf{C}$ if P is a very small point such as $(X_1 : Z_1) = (2 : 1)$. Here $\mathbf{C}$ stands for the cost of multiplication by $(A+2)/4$. One can find earlier formulas for the same computation in [20, formula (4.19)], but Montgomery's formulas are faster.

As far as we know, all subsequent ECM implementations have used Montgomery coordinates. In particular, GMP-ECM uses Montgomery coordinates for stage 1, with "PRAC," a particular differential addition chain introduced by Montgomery. Zimmermann and Dodson in [47, page 532, Figure 2] report a total cost of 2193683 differential additions to multiply an elliptic-curve point by $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdots 999983 \approx 2^{1440508.1677}$ in Montgomery coordinates. By adding a few counters to the source code of GMP-ECM we observed that GMP-ECM's stage 1, with $B_1 = 10^6$ and hence $s \approx 2^{1442098.6271}$, used 12982280 multiplications modulo n for 2196070 elliptic-curve differential additions, of which only 194155 were doublings.

4.5. Speedups in EECM-MPFQ. EECM-MPFQ breaks with stage-1 tradition in three ways:

- EECM-MPFQ uses Edwards curves $x^2 + y^2 = 1 + dx^2y^2$ with extended Edwards coordinates with $\phi = 1/x$ whereas GMP-ECM uses Montgomery curves with Montgomery coordinates. See below for performance results. Our prototype GMP-EECM used inverted twisted Edwards coordinates.
- EECM-MPFQ and GMP-EECM handle the prime factors π of s in batches, whereas GMP-ECM handles each prime factor separately. Specifically, GMP-EECM computed the product t of a batch, replaced P with $[t]P$, and then moved on to the next batch. EECM-MPFQ always uses a single batch: it computes the entire product s and then replaces P with $[s]P$. The large batches save time, as discussed below; the computation of s takes negligible time.
- EECM-MPFQ uses "signed sliding fractional window" addition-subtraction chains. Our prototype GMP-EECM used "signed sliding window" addition-subtraction chains. These chains compute $P \mapsto [s]P$ using only 1 doubling and ϵ additions for each bit of s. Here ϵ converges to 0 as s increases in length; this is why larger batches save time. The savings are amplified by the fact that an addition is somewhat more expensive than a doubling. Note that these chains are not compatible with Montgomery coordinates; they are shorter than any differential addition chain can be.

EECM-MPFQ follows tradition in its choice of s. Our experiments have not found significant speedups from other choices of s: for example, allowing prime powers in the larger interval $[1, B_1^{1.5}]$ has negligible extra cost when B_1 is large, but it also appears to have negligible benefit.

The addition-subtraction chains used in EECM-MPFQ are the chains $C_m(s)$ defined in [15, Section 3]. Given B_1, EECM-MPFQ computes s, computes $C_m(s)$ for various choices of the chain parameter m, and keeps the lowest-cost chain that it finds in a simple measure of cost. (Variations in the cost measure do not lead to noticeably better chains.) The total time spent on this computation is small: for example, under a second (on the CPU described below) for $B_1 = 1048576$. The resulting chain is reused for many curves and many inputs n.

Table 4.1 shows the actual number of elliptic-curve doublings and additions used by stage 1 of EECM-MPFQ. Table 4.1 also shows the actual number of field squarings, field multiplications, and field additions used by stage 1 of EECM-MPFQ. Recall that each doubling uses $3\mathbf{M} + 4\mathbf{S}$ while each addition uses $9\mathbf{M}$. The table shows that EECM-MPFQ uses only 8.84 multiplications per bit of s for $B_1 = 64$, only 8.42 multiplications per bit for $B_1 = 512$, only 7.91 multiplications per bit for $B_1 = 16384$, and only 7.61 multiplications per bit for $B_1 = 1048576$.

For comparison, GMP-ECM uses approximately 9 multiplications for each bit of s, as illustrated by the example with $B_1 = 10^6$ above. Zimmermann and Dodson explain in [47, Section 2] that one cannot expect differential addition chains to use fewer than $6/\log_2((1+\sqrt{5})/2) \approx 8.64$ multiplications per bit. Furthermore, only about one third of GMP-ECM's multiplications are squarings, while more than half of EECM-MPFQ's multiplications are squarings for $B_1 \geq 16384$. Even for the most carefully chosen curves, with extremely small parameters and extremely small base points, Montgomery's formulas use at least $4\mathbf{M} + 4\mathbf{S}$ per bit.

4.6. **Measurements of CPU cycles.** GMP-ECM relies primarily on the GMP integer-arithmetic library developed by Granlund et al., although for some CPUs it replaces portions of GMP with its own assembly-language subroutines for modular multiplication. EECM-MPFQ also uses GMP but performs almost all modular arithmetic using the MPFQ library introduced by Gaudry and Thomé in [27]. The tests described below used GMP 4.3.1 (released May 2009), GMP-ECM 6.2.3 (released April 2009), and MPFQ 1.0rc1 (released September 2008), all current at the time of testing (November 2009).

A 1000-curve test of EECM-MPFQ took 2.8 million cycles per curve on a single core of a 3.2GHz AMD Phenom II X4 (100f42) for a 240-bit n with $B_1 = 1024$ (and with $d_1 = 1$, disabling "stage 2"). For comparison, a 1000-curve test of GMP-ECM took 3.8 million cycles per curve on the same CPU for the same 240-bit n with the same B_1 (and with $B_2 = 1$).

The improvement in speed from GMP-ECM to EECM-MPFQ is even larger than what one would expect from comparing GMP-ECM's $8512\mathbf{M} + 4427\mathbf{S}$ to EECM-MPFQ's $6363\mathbf{M} + 5892\mathbf{S}$. The obvious explanation is that MPFQ's modular multiplications are faster than GMP's (and GMP-ECM's) modular multiplications; of course, the credit for this speedup belongs to Gaudry and Thomé.

Increasing B_1 to 16384 increased the EECM-MPFQ time to 40 million cycles per curve. There are 187307 modular multiplications per curve, specifically $92651\mathbf{M} + 94656\mathbf{S}$; evidently each modular multiplication took only about 220 cycles. For comparison, increasing B_1 to 16384 increased the GMP-ECM time to 60 million cycles per curve for 210307 modular multiplications, specifically $138884\mathbf{M}+71423\mathbf{S}$.

Increasing B_1 to 65536 increased the EECM-MPFQ time to 162 million cycles per curve. There are 735618 modular multiplications per curve, specifically $357866\mathbf{M}+ 377752\mathbf{S}$. For comparison, increasing B_1 to 65536 increased the GMP-ECM time

Table 4.1. Costs of computation of sP in EECM-MPFQ. The b column is the number of bits in $s = \operatorname{lcm}\{1, 2, \ldots, B_1\}$. #DBL and #ADD are the number of doublings and additions in the chain $C_m(s)$ selected by EECM-MPFQ. #**S**, #**M**, and #**a** are the number of field squarings, field multiplications, and field additions used by these elliptic-curve operations in extended Edwards coordinates. Per-curve setup costs and precomputation costs are included in the field-operation counts.

B_1	b	m	$\dfrac{\#\text{DBL}}{b}$	$\dfrac{\#\text{ADD}}{b}$	$\dfrac{\#\mathbf{S}+\#\mathbf{M}}{b}$	$\dfrac{\#\mathbf{S}}{b}$	$\dfrac{\#\mathbf{M}}{b}$	$\dfrac{\#\mathbf{a}}{b}$
8	10	5	0.800000	0.400000	10.100000	3.200000	6.900000	7.600000
12	15	3	0.933333	0.266667	9.533333	3.733333	5.800000	7.466667
16	20	5	0.900000	0.250000	9.000000	3.600000	5.400000	7.150000
24	33	3	0.939394	0.242424	9.030303	3.757576	5.272727	7.333333
32	48	11	0.916667	0.291667	9.229167	3.666667	5.562500	7.541667
48	69	7	0.956522	0.202899	8.652174	3.826087	4.826087	7.159420
64	90	9	0.977778	0.211111	8.844444	3.911111	4.933333	7.344444
96	130	15	0.969231	0.215385	8.792308	3.876923	4.915385	7.323077
128	184	15	0.978261	0.201087	8.706522	3.913043	4.793478	7.277174
192	275	29	0.985455	0.185455	8.600000	3.941818	4.658182	7.210909
256	363	15	0.988981	0.190083	8.658402	3.955923	4.702479	7.264463
384	557	27	0.991023	0.168761	8.472172	3.964093	4.508079	7.127469
512	743	27	0.993271	0.161507	8.418573	3.973082	4.445491	7.090175
768	1106	63	0.994575	0.150995	8.329114	3.978300	4.350814	7.024412
1024	1479	63	0.995943	0.145368	8.286004	3.983773	4.302231	6.993239
1536	2210	115	0.996833	0.138462	8.228054	3.987330	4.240724	6.950226
2048	2945	107	0.997623	0.131749	8.172156	3.990492	4.181664	6.907980
3072	4434	129	0.998647	0.124041	8.108029	3.994587	4.113442	6.860171
4096	5925	231	0.998650	0.120506	8.075949	3.994599	4.081350	6.835443
6144	8866	253	0.999098	0.114595	8.025603	3.996391	4.029213	6.796752
8192	11797	271	0.999322	0.111384	7.998135	3.997287	4.000848	6.775621
12288	17704	519	0.999492	0.105287	7.944306	3.997967	3.946340	6.733958
16384	23673	511	0.999620	0.101635	7.912263	3.998479	3.913784	6.709162
24576	35526	877	0.999719	0.097422	7.874965	3.998874	3.876091	6.680262
32768	47230	1019	0.999788	0.093966	7.844315	3.999153	3.845162	6.656490
49152	70828	1057	0.999859	0.090247	7.811303	3.999435	3.811868	6.630880
65536	94449	1847	0.999884	0.087698	7.788521	3.999534	3.788987	6.613188
98304	141805	2055	0.999922	0.084087	7.756278	3.999690	3.756588	6.588146
131072	189124	3079	0.999942	0.082057	7.738135	3.999767	3.738367	6.574052
196608	283651	4115	0.999958	0.078692	7.707947	3.999831	3.708117	6.550589
262144	378037	4639	0.999968	0.076815	7.691128	3.999873	3.691255	6.537516
393216	567462	8199	0.999977	0.073883	7.664799	3.999908	3.664890	6.517046
524288	756657	8187	0.999983	0.072121	7.648977	3.999931	3.649046	6.504745
786432	1134563	16383	0.999988	0.069733	7.627511	3.999951	3.627561	6.488054
1048576	1512566	16389	0.999991	0.067937	7.611370	3.999963	3.611407	6.475503

to 243 million cycles per curve for 842998 modular multiplications, specifically $557257\mathbf{M} + 285741\mathbf{S}$.

4.7. **EECM vs. HECM.** Chudnovsky and Chudnovsky in [20, Section 6] proposed a genus-2 hyperelliptic-curve method of factoring, using "simple forms of laws of addition on hyperelliptic surfaces, isogenous to the product of two elliptic curves."

Recently, in [21], Cosset reported that streamlined genus-2 formulas by Gaudry in [26] used only 189667 multiplications per elliptic curve (performed as 379334 multiplications per genus-2 curve) for $B_1 = 16384$, with the extra advantage that approximately 75% of the multiplications are squarings. Cosset quoted, for comparison, an earlier version of this paper that had reported 195111 multiplications per curve for GMP-EECM for $B_1 = 16384$.

A closer look shows that the formulas in [26] and [21] actually use, for each elliptic curve, 189667 multiplications *plus* approximately 189667 multiplications by small constants. EECM-MPFQ uses a *total* of only 187307 multiplications per elliptic curve, and the advantage grows as B_1 grows. Furthermore, the elliptic curves used in [21] are less effective than the elliptic curves used in EECM-MPFQ, and in fact are less effective than the elliptic curves used in GMP-ECM, according to the experiments described in [21, Section 3]. HECM is worth further investigation, but in its current form is clearly less efficient than EECM.

5. Using Edwards curves in ECM stage 2

This section discusses "stage 2" of ECM, and the benefit of switching to Edwards curves in stage 2.

5.1. **Overview of stage 2.**
Recall that stage 1 hopes for n to have a prime divisor q such that $[s]P$ is the neutral element of $E(\mathbf{Z}/q\mathbf{Z})$.

Stage 2 hopes for n to have a prime divisor q such that $[s]P$ has small prime order in $E(\mathbf{Z}/q\mathbf{Z})$: specifically, order ℓ for some prime ℓ between B_1 and B_2. Here B_1 is the stage-1 parameter described in the previous section, and B_2 is a new stage-2 parameter.

The most obvious way to check for a small order of $[s]P$ is a prime-by-prime approach, computing $[\ell s]P$ modulo n for each prime ℓ.

If ℓ' is the next prime after ℓ, then one can move from $[\ell s]P$ to $[\ell' s]P$ by adding a precomputed point $[(\ell' - \ell)s]P$. Computing all $[\ell s]P$ in this way takes about $B_2/\log B_2 - B_1/\log B_1$ elliptic-curve additions modulo n (where log without a subscript means the natural logarithm): there are about $B_2/\log B_2 - B_1/\log B_1$ primes ℓ, and the time for precomputation is quite small, since the differences $\ell' - \ell$ are generally quite small.

5.2. **Standard speedup: Baby steps and giant steps.**
A better way to check for a small order of $[s]P$ is with the following baby-step-giant-step computation. Fix a parameter $d_1 \in \{2, 4, 6, \ldots\}$, preferably a product of several different tiny primes. Choose a rational function $\psi : E \to \mathbf{Q}$ satisfying $\psi([js]P) = \psi([-js]P)$; for example, choose ψ as the Weierstrass x-coordinate or the Edwards y-coordinate. Compute

$$\gcd\left\{n, \prod_{\lceil B_1/d_1 - 1/2 \rceil \leq i \leq \lfloor B_2/d_1 + 1/2 \rfloor} \prod_{\substack{1 \leq j \leq d_1/2 \\ \gcd\{j,d_1\}=1}} (\psi([id_1 s]P) - \psi([js]P))\right\}.$$

The idea here is as follows. Assume that ℓ is a prime between B_1 and B_2 not dividing d_1. Write ℓ as $id_1 \pm j$ for some integers i, j with $j \in \{0, 1, \ldots, d_1/2\}$. Then i is between $B_1/d_1 - 1/2$ and $B_2/d_1 + 1/2$, and $\gcd\{j, d_1\} = \gcd\{\ell, d_1\} = 1$. If $[s]P$ has order ℓ in $E(\mathbf{Z}/q\mathbf{Z})$, then $[id_1 s]P = [\mp js]P$ in $E(\mathbf{Z}/q\mathbf{Z})$ so the numerator of $\psi([id_1 s]P) - \psi([js]P)$ is divisible by q.

In particular, the number of i's is balanced with the number of j's when $B_2 - B_1 \approx d_1\varphi(d_1)/2$, where φ is Euler's totient function. The baby steps $[js]P$ and the giant steps $[id_1s]P$ use about $\varphi(d_1)$ elliptic-curve additions, while the product of $\psi([id_1s]P) - \psi([js]P)$ uses about $\varphi(d_1)^2/4$ multiplications modulo n.

For comparison, the prime-by-prime approach uses roughly $d_1\varphi(d_1)/(4\log d_1)$ elliptic-curve additions. The baby-step-giant-step approach is an improvement whenever an elliptic-curve addition costs more than about $(\varphi(d_1)\log d_1)/d_1$ multiplications.

Asymptotically, $(\varphi(d_1)\log d_1)/d_1$ reaches ∞, even when d_1 is chosen as a product of tiny primes. However, in practice, $(\varphi(d_1)\log d_1)/d_1$ is always below 4; for example, if $d_1 = 510510 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17$, then $(\varphi(d_1)\log d_1)/d_1 \approx 2.37$. The baby-step-giant-step approach is therefore faster than the prime-by-prime approach.

5.3. **Standard speedup: Fast polynomial arithmetic.** If d_1 is large, then $\prod_i \prod_j (\psi([id_1s]P) - \psi([js]P))$ is more efficiently computed as $\prod_i F(\psi([id_1s]P))$ where $F = \prod_j (t - \psi([js]P)) \in (\mathbf{Z}/n\mathbf{Z})[t]$. Standard fast-arithmetic algorithms perform this computation in time $(\#\{i\} + \#\{j\})^{1+o(1)}$ rather than $\#\{i\}\#\{j\}$: first compute F via a "product tree"; then compute the values $F(\psi([id_1s]P))$ for all i via a "remainder tree" or a "scaled remainder tree"; then multiply the values. For details and further speedups see, e.g., [6]; [47, Section 3]; and [7, Sections 12, 18].

5.4. **Standard speedup: Higher-degree baby steps and giant steps.** One can replace $[js]P$ and $[id_1s]P$ by, e.g., $[j^6s]P$ and $[(id_1)^6s]P$. The advantage of this change is that one finds primes ℓ dividing $(id_1)^6 \pm j^6$, not just $id_1 \pm j$. If id_1 and j were uniformly distributed modulo ℓ, then $(id_1)^6 \pm j^6$ would be more than twice as likely as $id_1 \pm j$ to be divisible by ℓ. See [37, Section 5.3] for a probability analysis.

The only disadvantage is that there are more elliptic-curve operations. GMP-ECM computes $[j^6s]P$ for each integer $j \in \{1, 2, \ldots, d_1/2\}$ by computing the differences $[((j+1)^6 - j^6)s]P$, the second differences $[((j+2)^6 - 2(j+1)^6 + j^6)s]P$, etc.; the sixth differences are constants $[6!s]P$, so GMP-ECM uses a total of approximately $6(d_1/2)$ elliptic-curve additions. More generally, GMP-ECM computes $[j^es]P$ for each integer $j \in \{1, 2, \ldots, d_1/2\}$ using approximately $e(d_1/2)$ elliptic-curve additions.

One consequence of this generalization is that elliptic-curve operations cannot be a negligible part of the time taken by a properly optimized stage 2, compared to the time needed for computing the final product. If they *were* negligible, then increasing e would find a considerable number of additional primes at negligible extra cost.

GMP-ECM actually uses $D_e(j)$ instead of j^e. Here D_e is the degree-e "Dickson polynomial" defined by $D_e(t-1/t) = t^e + (-1/t)^e$. The differences $D_e(id_1) \pm D_e(j)$ have the same chance as $(id_1)^e \pm j^e$ to be divisible by ℓ, but are less closely correlated than $(id_1)^e \pm j^e$ as (i, j) vary; see [37, Table 5.3.1].

5.5. **The standard elliptic-curve coordinate system.** GMP-ECM does not use Montgomery coordinates in stage 2. Montgomery coordinates allow efficient differential additions, but most of the additions involved in higher-degree steps are not differential additions: they are sums where the differences are unknown.

GMP-ECM instead switches to *affine* coordinates (x, y). Addition in affine coordinates involves $1\mathbf{I} + 2\mathbf{M} + 1\mathbf{S}$. For degree-$e$ steps there are e additions to perform

in parallel; GMP-ECM combines $e\mathbf{I}$ into $1\mathbf{I} + 3(e-1)\mathbf{M}$. GMP-ECM's total cost for baby steps is therefore $(d_1/2)\mathbf{I} + (5e-3)(d_1/2)\mathbf{M} + e(d_1/2)\mathbf{S}$.

5.6. **Speedups in EECM-MPFQ.** EECM-MPFQ uses higher-degree baby steps and giant steps, with the same Dickson polynomials D_e used in GMP-ECM, but changes the elliptic-curve computations in three ways:

- EECM-MPFQ skips the $d_1/2 - \varphi(d_1)/2$ values of $j \in \{1, 2, \ldots, d_1/2\}$ that have $\gcd\{j, d_1\} > 1$. It computes $[D_e(j)s]P$ for the $\varphi(d_1)/2$ values of j that have $\gcd\{j, d_1\} = 1$ (and $[D_e(id_1)s]P$ for consecutive integers i). GMP-ECM tries to do something similar, focusing on the $d_1/6$ values of j in the arithmetic progression $(1+6\mathbf{Z}) \cap [1, d_1]$; but $d_1/6$ is considerably larger than $\varphi(d_1)/2$.
- EECM-MPFQ delays all inversions until the elliptic-curve computations are finished. It computes the desired y-coordinates in one final batched division, costing $1\mathbf{I} + (4(\#\{i\} + \#\{j\}) - 3)\mathbf{M}$ in total for the baby steps and the giant steps.
- EECM-MPFQ performs each intermediate elliptic-curve addition in extended Edwards coordinates, costing $9\mathbf{M}$. Occasionally an addition is a doubling costing only $3\mathbf{M} + 4\mathbf{S}$.

EECM-MPFQ computes the desired multiples of $[s]P$ as a batch using a simple general-purpose multi-scalar-multiplication algorithm described in [22, Section 4] with credit to Bos and Coster. If $n_1 \geq n_2 \geq \cdots$, then the algorithm computes the multiples $[n_1 s]P, [n_2 s]P, \ldots$ by recursively computing $[(n_1 - n_2)s]P, [n_2 s]P, \ldots$ and then adding $[n_2 s]P$ to $[(n_1 - n_2)s]P$. Actually, the Bos–Coster algorithm recursively computes $[(n_1 \bmod n_2)s]P, [n_2 s]P, \ldots$ and then adds the appropriate multiple of $[n_2 s]P$ to $[(n_1 \bmod n_2)s]P$; but this refinement is irrelevant in the typical case that $n_1 < 2n_2$.

Table 5.1 reports the number of multiplications used inside elliptic-curve operations in EECM-MPFQ's stage 2, for various choices of d_1 and e. The number of multiplications is divided by $\#\{i\} + \#\{j\}$ to produce each "cost" column. The final batched division costs 4 in this measure; the remaining cost is 9 times the per-output length of the Bos–Coster addition chain.

One can see from the table that the Bos–Coster addition chain has per-output length approximately 1 for $e = 1$; 1.9 for $e = 2$; 3.3 for $e = 3$; 11 for $e = 6$; and 22 for $e = 12$. For comparison, the addition chain used in GMP-ECM has per-output length approximately $ed_1/\varphi(d_1)$: i.e., roughly $4e$ for the range of d_1 shown in the table. This does *not* imply that GMP-ECM would benefit from switching to the Bos–Coster addition chain: GMP-ECM's stage-2 time is determined not only by addition-chain length but also by the number of inversions that can be performed in parallel.

By default EECM-MPFQ uses MPFQ to compute the final product. However, the user can ask EECM-MPFQ to switch to product trees and scaled remainder trees, using Shoup's NTL library for fast polynomial arithmetic; this saves time when $\#\{i\}+\#\{j\}$ is sufficiently large. In theory, one can and should integrate these computations, using fast polynomial arithmetic to split the product computation into problems that are small enough to be handled efficiently by MPFQ; in practice, this approach is hampered by the difficulty of moving data between NTL and MPFQ.

TABLE 5.1. Cost of elliptic-curve operations in stage 2 of EECM-MPFQ. Cost means the number of multiplications divided by $\#\{i\}+\#\{j\}$. Baby steps and giant steps are included. Multiplications used for inversion are included. Multiplications for the final product are not included.

d_1	B_1	$\#\{j\}$	$\#\{i\}$	$d_1\#\{i\}$	Cost $e=1$	Cost $e=2$	Cost $e=3$	Cost $e=6$	Cost $e=12$
30	60	4	4	120	18.75000	38.00000	57.12500	114.50000	242.75000
42	84	6	6	252	16.75000	33.41667	48.33333	106.16667	214.16667
60	120	8	8	480	15.81250	30.56250	43.50000	104.25000	204.31250
90	180	12	12	1080	14.87500	26.95833	41.95833	101.20833	205.83333
120	240	16	16	1920	14.40625	26.56250	38.37500	98.56250	200.09375
150	300	20	20	3000	14.12500	24.97500	39.57500	101.02500	201.15000
180	360	24	24	4320	13.93750	24.47917	37.41667	100.04167	201.10417
210	420	24	24	5040	14.33333	25.22917	39.10417	102.29167	205.41667
330	660	40	40	13200	13.91250	23.60000	37.43750	102.35000	204.27500
390	780	48	48	18720	13.65625	22.76042	36.64583	101.61458	202.30208
420	840	48	48	20160	13.65625	22.66667	36.64583	102.64583	204.17708
510	1020	64	64	32640	13.49219	22.42969	35.79688	101.04688	202.01563
630	1260	72	72	45360	13.43750	22.31250	36.38889	101.57639	204.45139
660	1320	80	80	52800	13.45625	22.06875	35.28750	100.76250	203.02500
780	1560	96	96	74880	13.32813	21.63542	34.71354	100.94792	201.58854
840	1680	96	96	80640	13.32813	22.05729	35.41667	99.11458	203.97917
990	1980	120	120	118800	13.30417	21.78333	35.47083	100.45833	201.63333
1050	2100	120	120	126000	13.26250	21.70417	35.80833	101.13333	203.32083
1260	2520	144	144	181440	13.21875	21.63194	34.44444	100.85069	203.19444
1470	2940	168	168	246960	13.18750	21.52381	34.43452	101.63988	202.43452
1680	3360	192	192	322560	13.16406	21.48958	34.05208	97.09896	201.84115
1890	3780	216	216	408240	13.14583	21.52315	34.37963	100.83796	201.42130
2100	4200	240	240	504000	13.13125	21.42292	33.68542	100.58542	201.77917
2310	4620	240	240	554400	13.18958	21.53542	34.43542	101.74792	203.80417
2520	5040	288	288	725760	13.10938	21.23785	33.33160	97.80035	201.67535
2730	5460	288	288	786240	13.15799	21.36285	33.78299	101.20660	202.51910
2940	5880	336	336	987840	13.09375	21.11905	33.21280	99.97619	200.79762
3150	6300	360	360	1134000	13.08750	21.16528	33.37639	100.29028	201.01528
3360	6720	384	384	1290240	13.08203	21.12370	33.06510	96.26432	200.13932
3570	7140	384	384	1370880	13.11719	21.27474	33.45052	100.72917	202.09635
3780	7560	432	432	1632960	13.07292	21.16898	33.15856	99.90856	200.72107
3990	7980	432	432	1723680	13.10532	21.34606	33.53356	100.72106	201.30440
4200	8400	480	480	2016000	13.06563	21.17708	33.02708	97.48021	200.33333
4290	8580	480	480	2059200	13.09479	20.99896	33.44896	100.30208	200.83958
4620	9240	480	480	2217600	13.09479	21.32708	33.43021	100.70521	202.04896

5.7. **Measurements of CPU cycles.** A 300-curve test of EECM-MPFQ took 4.7 million cycles per curve on a single core of a 3.2GHz AMD Phenom II X4 (100f42) for a 240-bit n with $B_1 = 1024$, $d_1 = 630$, $\#\{i\} = 72$, $\#\{j\} = 72$, and $e = 1$. Here $B_1 + d_1\#\{i\} = 46384$. For comparison, a 300-curve test of GMP-ECM took 10.8 million cycles per curve on the same CPU for the same 240-bit n with $B_1 = 1024$, $B_2 = 41526$, and $e = 1$. GMP-ECM took 11.5 million cycles per curve with $B_2 = 50646$, the next B_2 supported by GMP-ECM after 41526.

Increasing e to 3 increased the EECM-MPFQ time with $B_1 + d_1\#\{i\} = 46384$ to 5.4 million cycles per curve, increased the GMP-ECM time with $B_2 = 41526$ to 13.4 million cycles per curve, and increased the GMP-ECM time with $B_2 = 50646$ to 14.7 million cycles per curve.

Increasing d_1 to 510510, $\#\{i\}$ to 46080, and e to 12 increased the EECM-MPFQ time to 34000 million cycles per curve. Here $d_1\#\{i\} = 23524300800$. For comparison, GMP-ECM took only 18000 million cycles per curve with $e = 12$ and $B_2 = 23412731170$; for such large sizes GMP-ECM includes, and benefits from, FFT-based-polynomial-arithmetic optimizations not present in NTL and consequently not present in EECM-MPFQ.

6. Edwards curves with large torsion

Mazur's theorem [34, Theorem 2] says that the torsion group $E_{\text{tor}}(\mathbf{Q})$ of any elliptic curve E is isomorphic to one of 15 finite groups: specifically,

$$E_{\text{tor}}(\mathbf{Q}) \cong \begin{cases} \mathbf{Z}/m\mathbf{Z}, & m \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12\}, \quad \text{or} \\ \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2m\mathbf{Z}, & m \in \{1, 2, 3, 4\}. \end{cases}$$

Any Edwards curve has a point $(1, 0)$ of order 4. It follows that the torsion group of an Edwards curve is isomorphic to either $\mathbf{Z}/4\mathbf{Z}$, $\mathbf{Z}/8\mathbf{Z}$, $\mathbf{Z}/12\mathbf{Z}$, $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$, or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

The most interesting cases for ECM are $\mathbf{Z}/12\mathbf{Z}$ and $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$, since they force the group orders of E modulo primes p (of good reduction) to be divisible by 12 and 16 respectively. In this section we show which conditions an Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$ must satisfy to have torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$ or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. We give parameterizations for both cases.

One could hope to force divisibility by 12 in a different way, namely by finding a twisted Edwards curve with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$. A twisted Edwards curve does not need to have a point of order 4. However, we will show that there are no twisted Edwards curves with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$. We will also show that there are no twisted Edwards curves with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/10\mathbf{Z}$.

Computations in extended Edwards coordinates would benefit from using twisted Edwards curves with $a = -1$. We show that such curves cannot have $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$ or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

We first present the constructions and then show the impossibility results.

6.1. Torsion group $\mathbf{Z}/12\mathbf{Z}$.
Theorem 6.2 states a genus-0 cover of the set of Edwards curves over $\mathbf{Q}$ with torsion group $\mathbf{Z}/12\mathbf{Z}$. Theorem 6.3 identifies all the points of finite order on such curves. Theorem 6.4 states a rational cover.

Theorem 6.2. *If $y_3 \in \mathbf{Q} \setminus \{-2, -1/2, 0, \pm 1\}$ and $x_3 \in \mathbf{Q}$ satisfy the equation $x_3^2 = -(y_3^2 + 2y_3)$, then the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$, where $d = -(2y_3 + 1)/(x_3^2 y_3^2)$, has (x_3, y_3) as a point of order 3 and has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$. Conversely, every Edwards curve over $\mathbf{Q}$ with a point of order 3 arises in this way.*

Proof. Assume that such a y_3 and x_3 exist. Then d is defined and not equal to 0 or 1, and $x_3^2 + y_3^2 = -2y_3 = 1 + dx_3^2 y_3^2$. By Theorem 3.3, (x_3, y_3) is a point of order 3 on $\overline{E}_{E,1,d}(\mathbf{Q})$. Since each Edwards curve has a point of order 4 the torsion group

must contain a copy of $\mathbf{Z}/12\mathbf{Z}$. By Mazur's theorem the torsion group cannot be larger.

Conversely, if $\overline{\mathbf{E}}_{\mathrm{E},1,d}(\mathbf{Q})$ has a point of order 3, then by Theorem 3.3 the point can be written as (x_3, y_3) where $x_3^2 + y_3^2 = 1 + dx_3^2 y_3^2 = -2y_3$. Hence $x_3^2 = -(y_3^2 + 2y_3)$. Note that $x_3 \neq 0$, since otherwise $y_3^2 = 1 = -2y_3$; and note that $y_3 \notin \{0, -2\}$ since otherwise $x_3 = 0$. Now $d = -(2y_3 + 1)/(x_3^2 y_3^2)$. Finally, note that $y_3 \notin \{-1/2, \pm 1\}$ since otherwise $d \in \{0, 1\}$, contradicting the definition of an Edwards curve. $\square$

Theorem 6.3. *Let* $x^2 + y^2 = 1 + dx^2 y^2$ *be an Edwards curve over* $\mathbf{Q}$ *with* $E_{\mathrm{tor}}(\mathbf{Q}) \cong \mathbf{Z}/12\mathbf{Z}$ *and let* $P_3 = (x_3, y_3)$ *be a point of order 3 on the curve.*

The 12 *torsion points on the curve and their respective orders are as follows:*

point	$(0, 1)$	$(0, -1)$	$(\pm x_3, y_3)$	$(\pm 1, 0)$	$(\pm x_3, -y_3)$	$(\pm y_3, \pm x_3)$
order	1	2	3	4	6	12

Proof. The points of order 6 are obtained as $(\pm x_3, y_3) + (0, -1)$, the points of order 12 by adding $(\pm 1, 0)$ to the points of order 3. $\square$

Theorem 6.4. *If* $u \in \mathbf{Q} \setminus \{0, \pm 1\}$, *then the Edwards curve* $x^2 + y^2 = 1 + dx^2 y^2$ *over* $\mathbf{Q}$, *where*

$$x_3 = \frac{u^2 - 1}{u^2 + 1}, \quad y_3 = -\frac{(u - 1)^2}{u^2 + 1}, \quad d = \frac{(u^2 + 1)^3 (u^2 - 4u + 1)}{(u - 1)^6 (u + 1)^2}$$

has (x_3, y_3) *as a point of order 3 and has* $\mathbf{Q}$-*torsion group isomorphic to* $\mathbf{Z}/12\mathbf{Z}$. *Conversely, every Edwards curve over* $\mathbf{Q}$ *with a point of order 3 arises in this way.*

The parameters u *and* $1/u$ *give the same value of* d.

Proof. Multiply the identity $(u + 1)^2 + (u - 1)^2 = 2(u^2 + 1)$ by $(u - 1)^2/(u^2 + 1)^2$ to see that $x_3^2 + y_3^2 = -2y_3$, and observe that

$$d = \frac{2(u - 1)^2 - (u^2 + 1)}{u^2 + 1} \cdot \frac{(u^2 + 1)^2}{(u^2 - 1)^2} \cdot \frac{(u^2 + 1)^2}{(u - 1)^4} = \frac{-2y_3 - 1}{x_3^2 y_3^2}.$$

Furthermore, $y_3 \notin \{-2, -1/2, 0, \pm 1\}$ since $u \in \mathbf{Q} \setminus \{0, \pm 1\}$. By Theorem 6.2, the Edwards curve $x^2 + y^2 = 1 + dx^2 y^2$ over $\mathbf{Q}$ has (x_3, y_3) as a point of order 3 and has torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$.

Conversely, assume that the Edwards curve $x^2 + y^2 = 1 + dx^2 y^2$ has a point of order 3. By Theorem 6.2, the curve has a point (x_3, y_3) of order 3 for some $y_3 \in \mathbf{Q} \setminus \{-2, -1/2, 0, \pm 1\}$ and $x_3 \in \mathbf{Q}$ satisfying $x_3^2 = -(y_3^2 + 2y_3)$ and $d = -(2y_3 + 1)/(x_3^2 y_3^2)$. Note that $(x_3, y_3 + 1)$ is a point on the unit circle.

If $x_3 = \pm 1$, then $y_3 + 1 = 0$ so $d = -(2y_3 + 1)/(x_3^2 y_3^2) = 1$; but Edwards curves have $d \neq 1$. Hence $x_3 \neq \pm 1$. Furthermore, $x_3 \neq 0$ since every point with x-coordinate 0 has order 1 or 2.

Define u as the slope of the line between $(1, 0)$ and $(x_3, -(y_3 + 1))$: i.e., $u = (y_3 + 1)/(1 - x_3)$. Substitute $y_3 + 1 = u(1 - x_3)$ into $(y_3 + 1)^2 = 1 - x_3^2$ to obtain $u^2(1 - x_3)^2 = 1 - x_3^2 = (1 + x_3)(1 - x_3)$, i.e., $u^2(1 - x_3) = 1 + x_3$, i.e., $x_3 = (u^2 - 1)/(u^2 + 1)$. Then $u \notin \{0, \pm 1\}$ since $x_3 \notin \{0, -1\}$. Furthermore, $y_3 = u(1 - x_3) - 1 = u(2/(u^2 + 1)) - 1 = -(u - 1)^2/(u^2 + 1)$ and as above $d = (2y_3 + 1)/(x_3^2 y_3^2) = (u^2 + 1)^3 (u^2 - 4u + 1)/((u - 1)^6 (u + 1)^2)$.

The value of d is invariant under the change $u \mapsto 1/u$ since

$$\frac{(1 + u^2)^3 (1 - 4u + u^2)}{(1 - u)^6 (1 + u)^2} = \frac{(u^2 + 1)^3 (u^2 - 4u + 1)}{(u - 1)^6 (u + 1)^2}.$$ $\square$

Solving the equation $d(u') = d(u)$ for u' in terms of u over the rationals shows that $u \mapsto 1/u$ is the only rational transformation leaving d invariant that works independently of u.

6.5. Torsion group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

Theorem 6.6 states a genus-0 cover of the set of Edwards curves over $\mathbf{Q}$ with torsion group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. Theorem 6.8 identifies all the affine points of finite order on such curves. Theorem 6.9 states a rational cover and identifies the degree of the cover.

There are actually two types of curves in Theorem 6.6: points of order 8 double to $(\pm 1, 0)$ on curves of the first type, or to $((1 : \pm\sqrt{d}), (1 : 0))$ on curves of the second type. Curves of the second type are birationally equivalent to curves of the first type by Theorem 6.7. Subsequent theorems consider only the first type.

Theorem 6.6. *If $x_8 \in \mathbf{Q} \setminus \{0, \pm 1\}$ and $d = (2x_8^2 - 1)/x_8^4$ is a square in $\mathbf{Q}$, then the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$ has $(x_8, \pm x_8)$ as points of order 8 doubling to $(\pm 1, 0)$, and has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. Conversely, every Edwards curve over $\mathbf{Q}$ with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and a point of order 8 doubling to $(\pm 1, 0)$ arises in this way.*

If $\bar{x}_8 \in \mathbf{Q} \setminus \{0, \pm 1\}$ and $d = 1/(\bar{x}_8^2(2 - \bar{x}_8^2))$ is a square in $\mathbf{Q}$, then the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$ has $(\bar{x}_8, \pm 1/(\bar{x}_8\sqrt{d}))$ as points of order 8 doubling to $((1 : \pm\sqrt{d}), (1 : 0))$, and has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. Conversely, every Edwards curve over $\mathbf{Q}$ with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and a point of order 8 doubling to $((1 : \pm\sqrt{d}), (1 : 0))$ arises in this way.

Every Edwards curve over $\mathbf{Q}$ with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ arises in one of these two ways.

Proof. Any such x_8 yields $d \neq 0, 1$, so $x^2 + y^2 = 1 + dx^2y^2$ is an Edwards curve. By Theorems 3.1 and 3.2, the curve has points $(0, -1)$ and $((1 : 0), (1 : \pm\sqrt{d}))$ of order 2, and points $(x_8, \pm x_8)$ of order 8 doubling to $(\pm 1, 0)$. Similarly, any such $\bar{x}_8$ yields an Edwards curve with points $(0, -1)$ and $((1 : 0), (1 : \pm\sqrt{d}))$ of order 2 and $(\bar{x}_8, \pm 1/(\bar{x}_8\sqrt{d}))$ of order 8 doubling to $((1 : \pm\sqrt{d}), (1 : 0))$.

In both cases the torsion group contains a copy of $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. By Mazur's theorem the torsion group cannot be larger.

Conversely, assume that $x^2 + y^2 = 1 + dx^2y^2$ is an Edwards curve with $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. There are four elements of order 4 in $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$, all doubling to the same element, so there are four order-4 points on the curve, all doubling to the same point.

The points $(\pm 1, 0)$ have order 4 and double to $(0, -1)$, so the other two points of order 4 also double to $(0, -1)$. By Theorem 3.1, those other two points must be $((1 : \pm\sqrt{d}), (1 : 0))$, and d must be a square.

Now any point of order 8 must double to $(\pm 1, 0)$ or to $((1 : \pm\sqrt{d}), (1 : 0))$. In the first case, by Theorem 3.2, the point is $(x_8, \pm x_8)$ for some root x_8 of $dx_8^4 - 2x_8^2 + 1$; hence $x_8 \notin \{0, \pm 1\}$ and $d = (2x_8^2 - 1)/x_8^4$. In the second case, by Theorem 3.2, the point is $(\bar{x}_8, \pm 1/(\bar{x}_8\sqrt{d}))$ for some root $\bar{x}_8$ of $d\bar{x}_8^4 - 2d\bar{x}_8^2 + 1$; hence $\bar{x}_8 \notin \{0, \pm 1\}$ and $d = 1/(\bar{x}_8^4 - 2\bar{x}_8^2)$. $\square$

Theorem 6.7. *Let d be a square. The Edwards curves $x^2 + y^2 = 1 + dx^2y^2$ and $\bar{x}^2 + \bar{y}^2 = 1 + (1/d)\bar{x}^2\bar{y}^2$ are birationally equivalent via the map $\bar{x} = x\sqrt{d}, \bar{y} = 1/y$ with inverse $x = \bar{x}/\sqrt{d}, y = 1/\bar{y}$. The map fixes $(0, \pm 1)$.*

Proof. Inserting $\bar{x} = x\sqrt{d}, \bar{y} = 1/y$ into $x^2 + y^2 = 1 + dx^2y^2$ gives $\bar{x}^2/d + 1/\bar{y}^2 = 1 + \bar{x}^2/\bar{y}^2$ which after multiplication by $\bar{y}^2$ gives $\bar{x}^2\bar{y}^2/d + 1 = \bar{y}^2 + \bar{x}^2$. The only exceptional points are $(\pm1, 0)$. The statement about $(0, \pm1)$ follows by direct inspection. $\qquad\square$

In particular, each curve of the second type in Theorem 6.6 is birationally equivalent to a curve of the first type. Indeed, assume that $\bar{x}_8 \in \mathbf{Q} \setminus \{0, \pm1\}$ and that $d = 1/(\bar{x}_8^2(2 - \bar{x}_8^2))$ is a square in $\mathbf{Q}$. Define $x_8 = \bar{x}_8\sqrt{d}$. Then $x_8^2 = 1/(2 - \bar{x}_8^2)$, so $(2x_8^2 - 1)/x_8^4 = (2/(2 - \bar{x}_8^2) - 1)(2 - \bar{x}_8^2)^2 = \bar{x}_8^2(2 - \bar{x}_8^2) = 1/d$, which is a square; furthermore, $x_8 \notin \{0, \pm1\}$ since $2 - \bar{x}_8^2 \neq 1$ since $\bar{x}_8 \notin \{\pm1\}$. Hence $x^2 + y^2 = 1 + (1/d)x^2y^2$ is a curve of the first type. The curve $x^2 + y^2 = 1 + dx^2y^2$ is birationally equivalent to $\bar{x}^2 + \bar{y}^2 = 1 + (1/d)\bar{x}^2\bar{y}^2$ by Theorem 6.7. Consequently, we can restrict attention to curves of the first type, i.e., curves on which the points of order 8 double to $(\pm1, 0)$.

Theorem 6.8. *Assume that $x_8 \in \mathbf{Q} \setminus \{0, \pm1\}$ and that $d = (2x_8^2 - 1)/x_8^4$ is a square in $\mathbf{Q}$. Then there are 16 points of finite order on $\overline{\mathrm{E}}_{\mathrm{E},1,d}$ over $\mathbf{Q}$. The affine points of finite order are as follows:*

point	$(0, 1)$	$(0, -1)$	$(\pm1, 0)$	$(\pm x_8, \pm x_8)$	$\left(\pm1/(x_8\sqrt{d}), \pm1/(x_8\sqrt{d})\right)$
order	1	2	4	8	8

where the signs are taken independently.

Proof. Theorem 3.1 (with $a = 1$) shows that the 4 affine points $(0, 1)$, $(0, -1)$, and $(\pm1, 0)$ are on $\overline{\mathrm{E}}_{\mathrm{E},1,d}$ and have the stated orders. It also shows that the 2 nonaffine points $((1 : 0), (1 : \pm\sqrt{d}))$ have order 2 and that the 2 nonaffine points $((1 : \pm\sqrt{d}), (1 : 0))$ have order 4. Theorem 3.2 shows that the other affine points listed are 8 distinct points on $\overline{\mathrm{E}}_{\mathrm{E},1,d}$ and have order 8. The torsion group has exactly 16 elements by Theorem 6.6. $\qquad\square$

Theorem 6.9. *If $u \in \mathbf{Q} \setminus \{0, -1, -2\}$, then the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$, where*

$$x_8 = \frac{u^2 + 2u + 2}{u^2 - 2}, \qquad d = \frac{2x_8^2 - 1}{x_8^4},$$

has (x_8, x_8) as a point of order 8 and has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

Conversely, every Edwards curve over $\mathbf{Q}$ with torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ on which the points of order 8 double to $(\pm1, 0)$ is expressible in this way.

The parameters $u, 2/u, -2(u+1)/(u+2), -(2+u)/(1+u), -(u+2), -2/(u+2), -u/(u+1)$, and $-2(u+1)/u$ give the same value of d and they are the only values giving this d.

Proof. Divide the identity $2(u^2 + 2u + 2)^2 - (u^2 - 2)^2 = (u^2 + 4u + 2)^2$ by $(u^2 - 2)^2$ to see that $2x_8^2 - 1 = (u^2 + 4u + 2)^2/(u^2 - 2)^2$. Hence d is a square. Furthermore, $x_8 \neq 0$ since $u^2 + 2u + 2 \neq 0$; $x_8 \neq 1$ since $u \neq -2$; and $x_8 \neq -1$ since $u \notin \{0, -1\}$. By Theorem 6.6, the curve $\overline{\mathrm{E}}_{\mathrm{E},1,d}$ has (x_8, x_8) as a point of order 8, and has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

Conversely, assume that an Edwards curve has torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and has a point of order 8 doubling to $(\pm1, 0)$. By Theorem 6.6, the

curve can be expressed as $\overline{E}_{E,1,d}$ for some $x_8 \in \mathbf{Q}\backslash\{0,\pm1\}$ such that $d = (2x_8^2-1)/x_8^4$ is a square in $\mathbf{Q}$; i.e., such that $2x_8^2 - 1$ is a square in $\mathbf{Q}$.

Choose $r \in \mathbf{Q}$ such that $2x_8^2 - 1 = r^2$. Define u as the slope of the line between $(1,-1)$ and (x_8,r): i.e., $u = (r+1)/(x_8-1)$. Substitute $r = u(x_8-1)-1$ into $2(x_8^2-1) = (r+1)(r-1)$ to obtain $2(x_8^2-1) = u(x_8-1)(u(x_8-1)-2)$, i.e., $2(x_8+1) = u(u(x_8-1)-2)$, i.e., $2x_8 + 2 = u^2x_8 - u^2 - 2u$; i.e., $x_8 = (u^2+2u+2)/(u^2-2)$. Finally, $u \notin \{0,-1\}$ since $x_8 \neq -1$, and $u \neq -2$ since $x_8 \neq 1$.

The identity

$$(d(u) - d(v))((u+1)^2 + 1)^4((v+1)^2 + 1)^4$$
$$= 16(u-v)(uv-2)((u+2)v + 2(u+1))(u+2+(u+1)v)$$
$$\cdot (u+v+2)((u+2)v + 2)(u+(u+1)v)(uv + 2(u+1))$$

immediately shows that if v is any of the listed values $u, 2/u, \ldots$ then $d(v) = d(u)$. Conversely, if v is not one of those values then none of the factors $u-v, uv-2, \ldots$ are 0 so $d(v) \neq d(u)$. $\qquad\square$

6.10. **Impossibility results.** The following theorem shows that the only way for a twisted Edwards curve to have exactly 12 torsion points is to have torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$. The next two theorems consider twisted Edwards curves with $a = -1$ and show that these cannot have $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$ or $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. The last theorem shows that a twisted Edwards curve cannot have exactly 10 torsion points.

Theorem 6.11. *There exists no twisted Edwards curve over $\mathbf{Q}$ with torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$.*

Proof. Let a, d be distinct nonzero elements of $\mathbf{Q}$. Suppose that the twisted Edwards curve $\overline{E}_{E,a,d} : ax^2 + y^2 = 1 + dx^2y^2$ has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$.

There are three elements of order 2 in $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$, so there are three points of order 2 in $\overline{E}_{E,a,d}(\mathbf{Q})$. By Theorem 3.1 the only possible points of order 2 are $(0,-1)$ and $((1:0),(\pm\sqrt{a/d}:1))$. Hence $\sqrt{a/d} \in \mathbf{Q}$.

There are also elements of order 3 in $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/6\mathbf{Z}$. Choose a point of order 3 in $\overline{E}_{E,a,d}(\mathbf{Q})$. By Theorem 3.3 this point can be expressed as (x_3, y_3) where $ax_3^2 + y_3^2 = 1 + dx_3^2y_3^2 = -2y_3$.

Write $u = 1 + y_3$. Then $1 - u^2 = -2y_3 - y_3^2 = ax_3^2$. Starting from $dx_3^2y_3^2 = ax_3^2 + y_3^2 - 1$, replace x_3^2 by $(1-u^2)/a$ and replace y_3 by $u-1$ to see that $(d/a)(1-u^2)(u-1)^2 = (1-u^2) + (u-1)^2 - 1 = 1 - 2u$. Hence $s^2 = 4(1-2u)(1-u^2)$ where $s = 2(1-u^2)(u-1)\sqrt{d/a} \in \mathbf{Q}$.

In other words, $(2u,s)$ is a $\mathbf{Q}$-rational point (σ,τ) on the elliptic curve $\tau^2 = \sigma^3 - \sigma^2 - 4\sigma + 4$. This elliptic curve has rank 0 over $\mathbf{Q}$, and has exactly 7 affine points over $\mathbf{Q}$, as one can verify by typing

```
E=EllipticCurve(QQ,[0,-1,0,-4,4])
print E.rank()
print E.torsion_points()
```

into the Sage computer-algebra system [43]. Specifically, (σ,τ) must be one of $(\pm2,0),(0,\pm2),(1,0),(4,\pm6)$. Hence $u \in \{\pm1,0,1/2,2\}$. In each case $(a:d) = ((1-u^2)(u-1)^2 : 1-2u) \in \{(1:1),(0:1),(1:0)\}$, contradicting the assumption that a, d are distinct nonzero elements of $\mathbf{Q}$. $\qquad\square$

Theorem 6.12. *There exists no twisted Edwards curve of the form $ax^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$ with $a = -1$ and torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$.*

Proof. Suppose that the twisted Edwards curve $\overline{\mathrm{E}}_{\mathrm{E},-1,d} : -x^2 + y^2 = 1 + dx^2y^2$ has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$.

There is a unique element of order 2 in $\mathbf{Z}/12\mathbf{Z}$, so $(0, -1)$ is the only point of order 2 on $\overline{\mathrm{E}}_{\mathrm{E},-1,d}(\mathbf{Q})$. Furthermore, there are elements of order 4 in $\mathbf{Z}/12\mathbf{Z}$, so there are points on $\overline{\mathrm{E}}_{\mathrm{E},-1,d}(\mathbf{Q})$ doubling to $(0, -1)$. By Theorem 3.1 the only possibilities for such points are $((1 : \pm\sqrt{a}), (0 : 1))$ or $((1 : \pm\sqrt{d}), (1 : 0))$. Hence a or d is a square in $\mathbf{Q}$; but $a = -1$ is not a square in $\mathbf{Q}$, so d is a square in $\mathbf{Q}$.

There are also elements of order 3 in $\mathbf{Z}/12\mathbf{Z}$. As in the proof of Theorem 6.11 there exists $u \in \mathbf{Q}$ such that $(d/a)(1 - u^2)(u - 1)^2 = 1 - 2u$. Here $a = -1$ so $s^2 = -4(1 - u^2)(1 - 2u)$ where $s = 2(1 - u^2)(u - 1)\sqrt{d} \in \mathbf{Q}$.

In other words, $(-2u, s)$ is a $\mathbf{Q}$-rational point on the elliptic curve $\tau^2 = \sigma^3 + \sigma^2 - 4\sigma - 4$. This elliptic curve has rank 0 over $\mathbf{Q}$, and has exactly 3 affine points over $\mathbf{Q}$: specifically, (σ, τ) must be one of $(\pm 2, 0), (-1, 0)$. Hence $u \in \{\pm 1, 1/2\}$. If $u \in \{\pm 1\}$ then $0 = (d/a)(1 - u^2)(u - 1)^2 = 1 - 2u \neq 0$, contradiction; if $u = 1/2$ then $0 = 1 - 2u = (d/a)(1 - u^2)(u - 1)^2 \neq 0$, contradiction. $\square$

Theorem 6.13. *There exists no twisted Edwards curve of the form $ax^2 + y^2 = 1 + dx^2y^2$ over $\mathbf{Q}$ with $a = -1$ and torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.*

Proof. Suppose that the twisted Edwards curve $\overline{\mathrm{E}}_{\mathrm{E},-1,d} : -x^2 + y^2 = 1 + dx^2y^2$ has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

The torsion group contains exactly three elements of order 2, so $\sqrt{a/d} \in \mathbf{Q}$ as in the proof of Theorem 6.11: i.e., $\sqrt{-d} \in \mathbf{Q}$. Consequently, d is not a square in $\mathbf{Q}$.

The torsion group also contains exactly 4 elements of order 4. These elements cannot double to $(0, -1)$: otherwise they would have the form $((1 : \pm\sqrt{-1}), (0 : 1))$ or $((1 : \pm\sqrt{d}), (1 : 0))$ by Theorem 3.1, but neither -1 nor d is a square in $\mathbf{Q}$. The elements of order 4 therefore double to $((1 : 0), (\pm\sqrt{-1/d} : 1))$.

If $s^2 = -1/d$, then the elements of order 4 doubling to $((1 : 0), (s : 1))$ are $(\pm\sqrt{s}, \pm\sqrt{s})$ by Theorem 3.1, where the $\pm$ signs are assumed independently. In particular, if such elements are defined over $\mathbf{Q}$, then $\pm\sqrt{s} \in \mathbf{Q}$, so s is a square in $\mathbf{Q}$, so $-1/d$ is a fourth power in $\mathbf{Q}$, say f^4. Now $(\pm f, \pm f)$ are points of order 4 doubling to $((1 : 0), (f^2 : 1))$, and there are no other points of order 4.

The torsion group contains a point P_8 of order 8. This point doubles to $(\pm f, \pm f)$. Assume, without loss of generality, that $[2]P_8 = (\pm f, f)$: otherwise replace f by $-f$. Further assume, without loss of generality, that $[2]P_8 = (f, f)$; otherwise replace P_8 by $-P_8$. Any point having a zero coordinate has order at most 4, so P_8 must be an affine point, say (x_8, y_8), with $x_8 \neq 0$ and $y_8 \neq 0$.

Now $[2]P_8 = (f, f)$ implies $(2x_8y_8)/(-x_8^2 + y_8^2) = f = (y_8^2 + x_8^2)/(2 + x_8^2 - y_8^2)$, with $-x_8^2 + y_8^2 \neq 0$ and $2 + x_8^2 - y_8^2 \neq 0$. In particular, $(y_8^2 + x_8^2)(-x_8^2 + y_8^2) = (2x_8y_8)(2 + x_8^2 - y_8^2)$, so $(y_8^2 - x_8^2)(x_8^2 + y_8^2 + 2x_8y_8) = 4x_8y_8$; i.e., $(y_8^2 - x_8^2)r^2 = 4x_8y_8$ where $r = x_8 + y_8$.

Define $s = 2(y_8^2 + x_8^2)/(y_8^2 - x_8^2)$. Then

$$s^2 - 4 = \frac{4((y_8^2 + x_8^2)^2 - (y_8^2 - x_8^2)^2)}{(y_8^2 - x_8^2)^2} = \frac{16y_8^2x_8^2}{(y_8^2 - x_8^2)^2} = r^4$$

so $(s + r^2)^2 - 4 = 2r^2(s + r^2)$; consequently, $((s + r^2)/2, r(s + r^2)/2)$ is a rational point on the elliptic curve $\tau^2 = \sigma^3 - \sigma$. This curve has rank 0 over $\mathbf{Q}$ and exactly 3

affine points over $\mathbf{Q}$, namely $(\pm 1, 0)$ and $(0, 0)$. Hence $r(s + r^2) = 0$; consequently, $0 = r(s + r^2)(s - r^2) = r(s^2 - r^4) = 4r$, so $r = 0$, so $x_8 + y_8 = 0$, contradicting $-x_8^2 + y_8^2 \neq 0$. □

Theorem 6.14. *There exists no twisted Edwards curve over $\mathbf{Q}$ with torsion group isomorphic to $\mathbf{Z}/10\mathbf{Z}$.*

Proof. Suppose that the twisted Edwards curve $\overline{\mathrm{E}}_{\mathrm{E},a,d} : ax^2 + y^2 = 1 + dx^2y^2$ has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/10\mathbf{Z}$. This means, in particular, that there exists a point $P_5 \in \overline{\mathrm{E}}_{\mathrm{E},a,d}(\mathbf{Q})$ of order 5. Points at infinity have order at most 4 by Theorem 3.1, so $P_5 = (x_5, y_5)$ for some $x_5, y_5 \in \mathbf{Q}$. Points with a zero coordinate also have order at most 4 by Theorem 3.1, so $x_5 \neq 0$ and $y_5 \neq 0$. Note also that $y_5 \notin \{-1, 1\}$ since $x_5 \neq 0$.

Apply the doubling formulas twice to see that the x-coordinate of $[4]P_5$ satisfies

$$x([4]P_5) - (-x_5) =$$
$$\frac{x_5 \left(ax_5^2 + y_5^2 - 2y_5\right) F}{a^4 x_5^8 + 4y_5^2 a^3 x_5^6 + (6y_5^4 - 16y_5^2)a^2 x_5^4 + (4y_5^6 - 16y_5^4 + 16y_5^2)ax_5^2 + y_5^8},$$

where $F = a^3 x_5^6 + (3y_5^2 + 6y_5)a^2 x_5^4 + (3y_5^4 + 4y_5^3 - 4y_5^2 - 8y_5)ax_5^2 + y_5^6 - 2y_5^5 - 4y_5^4$. The equation $[4]P_5 = -P_5$ implies $x([4]P_5) - (-x_5) = 0$, so $x_5(ax_5^2 + y_5^2 - 2y_5)F = 0$.

Case 1: $ax_5^2 + y_5^2 = 2y_5$. Then $(x_5, -y_5)$ is a curve point of order 3 by Theorem 3.3, contradicting the hypothesis that the torsion group is isomorphic to $\mathbf{Z}/10\mathbf{Z}$.

Case 2: $F = 0$. Define $q = (ax_5^2 + y_5^2 + 2y_5)/y_5$ and $r = q/(y_5 + 1)$. The identity $rq^2 - (r^2 + 8)q + 16 = F/(y_5^2(y_5 + 1)^2)$ then implies $rq^2 - (r^2 + 8)q + 16 = 0$.

Define $U = q - r$, $V = q - r - 4$, and $W = 4 - q - r$. Then $(U, V, W) \neq (0, 0, 0)$, and $V^2 W - U^3 - U^2 W + UW^2 = 4(rq^2 - (r^2 + 8)q + 16) = 0$, so $(U : V : W)$ is a rational point on the elliptic curve $\tau^2 = \sigma^3 + \sigma^2 - \sigma$. This curve has rank 0 over $\mathbf{Q}$ and exactly 6 points over $\mathbf{Q}$, namely $(\pm 1, \pm 1)$, $(0, 0)$, and $(0 : 1 : 0)$, so $(U : V : W)$ is one of those points.

The points $(1, 1)$ and $(-1, -1)$ and $(0, 0)$ are excluded since $U \neq V$. The point $(1, -1)$ implies $(q, r) = (2, 0)$, contradicting $r = q/(y_5 + 1)$. The point $(-1, 1)$ implies $(q, r) = (4, 2)$, again contradicting $r = q/(y_5 + 1)$ since $y_5 \neq 1$. Finally, the point $(0 : 1 : 0)$ implies $(q, r) = (2, 2)$, again contradicting $r = q/(y_5 + 1)$ since $y_5 \neq 0$. □

7. Edwards curves with large torsion and positive rank

Atkin and Morain in [3] found an infinite family of elliptic curves over $\mathbf{Q}$ with torsion group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and with explicit nontorsion points. Montgomery in [36, page 263] had earlier found an analogous family for $\mathbf{Z}/12\mathbf{Z}$. Suyama in [44] had earlier given an infinite sequence of Montgomery curves with explicit nontorsion points and with group order divisible by 12 over any prime field. GMP-ECM uses Suyama curves; see [47]. See [36, Section 6] for further $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ constructions.

Montgomery in [36, Section 6] selected various curves with torsion group $\mathbf{Z}/12\mathbf{Z}$, computed the group orders modulo primes p in the interval $[10^4, 10^5]$, and found that the average exponents of 2 and 3 in the group orders were close to $11/3$ and $5/3$ respectively. We performed an analogous computation for primes in $[10^6, 2 \cdot 10^6]$, using Edwards curves with torsion group $\mathbf{Z}/12\mathbf{Z}$ constructed as in Section 6, and found an even closer match to $11/3$ and $5/3$.

Most Suyama curves have $\mathbf{Q}$-torsion group only $\mathbf{Z}/6\mathbf{Z}$, and have averages only $10/3$ and $5/3$. However, Kruppa discovered in 2007 that the Suyama curve with $\sigma = 11$ has averages $11/3$ and $5/3$, despite having torsion group only $\mathbf{Z}/6\mathbf{Z}$; see [31, Section 5.2.3].

In this section we translate the Atkin–Morain and Montgomery constructions from Weierstrass curves to Edwards curves. We also translate the Suyama construction to twisted Edwards curves.

This section relies on the equivalence in [8] between Montgomery curves and twisted Edwards curves. The twisted Edwards curve $E_{E,a,d}$ is birationally equivalent to the Montgomery curve $E_{M,A,B} : Bv^2 = u^3 + Au^2 + u$, where $A = 2(a+d)/(a-d)$ and $B = 4/(a-d)$. The map $(x, y) \mapsto (u, v) = ((1+y)/(1-y),$ $(1+y)/((1-y)x))$ is a birational equivalence from $E_{E,a,d}$ to $E_{M,A,B}$, with inverse $(u, v) \mapsto (x, y) = (u/v, (u-1)/(u+1))$.

7.1. **The Atkin–Morain construction.** The Atkin–Morain family is parameterized by points (s, t) on a particular elliptic curve $T^2 = S^3 - 8S - 32$. Atkin and Morain suggest computing multiples (s, t) of $(12, 40)$, a nontorsion point on this curve. Beware that these points have rapidly increasing height.

Theorem 7.2 (Atkin, Morain). *Let (s, t) be a rational point on the curve $T^2 = S^3 - 8S - 32$. Define $\alpha = ((t+25)/(s-9) + 1)^{-1}$, $\beta = 2\alpha(4\alpha + 1)/(8\alpha^2 - 1)$, $c = (2\beta - 1)(\beta - 1)/\beta$, and $b = \beta c$. Then the elliptic curve*

$$E_\alpha :\ V^2 = U^3 + \frac{((c-1)^2 - 4b)}{4}U^2 + \frac{b(c-1)}{2}U + \frac{b^2}{4}$$

has torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and a point with U-coordinate $-(2\beta - 1)/4$.

Theorem 7.3. *Let (s, t) be a rational point on the curve $T^2 = S^3 - 8S - 32$. Define α and β as in Theorem 7.2. Define $d = (2(2\beta - 1)^2 - 1)/(2\beta - 1)^4$. Then the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ has torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and a point (x_1, y_1) with $x_1 = (2\beta - 1)(4\beta - 3)/(6\beta - 5)$ and $y_1 = (2\beta - 1)(t^2 + 50t - 2s^3 + 27s^2 - 104)/((t + 3s - 2)(t + s + 16))$.*

Proof. Put $x_8 = 2\beta - 1$. By construction x_8 satisfies $(2x_8^2 - 1)/x_8^4 = d$. Furthermore,

$$d = \frac{(8\alpha^2 - 1)^2(8\alpha^2 + 8\alpha + 1)^2}{(8\alpha^2 + 4\alpha + 1)^4},$$

so d is a square. By Theorem 6.6, the Edwards curve has torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$. Finally, a straightforward calculation shows that $x_1^2 + y_1^2 = 1 + dx_1^2y_1^2$. $\qquad\square$

The point with U-coordinate $-(2\beta - 1)/4$ in Theorem 7.2 is generically a nontorsion point. The V-coordinate of the point is not stated explicitly in [3]. The point (x_1, y_1) in Theorem 7.3 is the corresponding point on the Edwards curve.

7.4. **The Suyama construction.** The Suyama family has lower torsion but a simpler parameterization. We briefly review Suyama's family and present an analogous result for twisted Edwards curves.

Theorem 7.5 (Suyama). *Let $\sigma > 5$ be an integer. Define*

$$\alpha = \sigma^2 - 5, \quad \beta = 4\sigma, \quad W_1 = \beta^3, \quad A = \frac{(\beta - \alpha)^3(3\alpha + \beta)}{4\alpha^3\beta} - 2, \quad B = \frac{\alpha}{W_1}.$$

Then the $\mathbf{Q}$-torsion group of the elliptic curve $\mathrm{E}_{\mathrm{M},A,B} : Bv^2 = u^3 + Au^2 + u$ has a subgroup isomorphic to $\mathbf{Z}/6\mathbf{Z}$.

Define $V_1 = (\sigma^2 - 1)(\sigma^2 - 25)(\sigma^4 - 25)$. Then $(u_1, v_1) = (\alpha^3/W_1, V_1/W_1)$ is a nontorsion point on $\mathrm{E}_{\mathrm{M},A,B}$.

Theorem 7.6. *Let $\sigma > 5$ be an integer. Define α, β, V_1 as in Theorem 7.5. Define $a = (\beta - \alpha)^3(3\alpha + \beta)$ and $d = (\beta + \alpha)^3(\beta - 3\alpha)$. Then the $\mathbf{Q}$-torsion group of the twisted Edwards curve $ax^2 + y^2 = 1 + dx^2y^2$ has a subgroup isomorphic to $\mathbf{Z}/6\mathbf{Z}$, and $(x_1, y_1) = (\alpha\beta/(2V_1), (\alpha^3 - \beta^3)/(\alpha^3 + \beta^3))$ is a nontorsion point on the curve.*

Proof. Define W_1, A, B as in Theorem 7.5. Then $2(a+d)/(a-d) = A$ and $4/(a-d) = B\beta^2/(2\alpha^2)^2$. The twisted Edwards curve $ax^2 + y^2 = 1 + dx^2y^2$ is birationally equivalent to the Montgomery curve $(B\beta^2/(2\alpha^2)^2)v^2 = u^3 + Au^2 + u$, which in turn is isomorphic to the Montgomery curve $Bv^2 = u^3 + Au^2 + u$, so its $\mathbf{Q}$-torsion group has a subgroup isomorphic to $\mathbf{Z}/6\mathbf{Z}$ by Theorem 7.5.

Define u_1, v_1 as in Theorem 7.5. Then (u_1, v_1) is a nontorsion point on $Bv^2 = u^3 + Au^2 + u$, so $(u_1, v_1(2\alpha^2)/\beta)$ is a nontorsion point on $(B\beta^2/(2\alpha^2)^2)v^2 = u^3 + Au^2 + u$. Mapping this point to $\mathrm{E}_{\mathrm{E},a,d}$ yields exactly (x_1, y_1):

$$x_1 = \frac{u_1}{v_1(2\alpha^2)/\beta} = \frac{\alpha^3}{V_1(2\alpha^2)/\beta} = \frac{\alpha\beta}{2V_1} \quad \text{and} \quad y_1 = \frac{u_1 - 1}{u_1 + 1} = \frac{\alpha^3 - \beta^3}{\alpha^3 + \beta^3}.$$

Hence (x_1, y_1) is a nontorsion point on $ax^2 + y^2 = 1 + dx^2y^2$. $\square$

7.7. The Montgomery construction.

Montgomery's family is parameterized by points (s,t) on the curve $T^2 = S^3 - 12S$, except for the points $(0,0)$, $Q = (-2,-4)$, $R = (6,-12)$, $-Q$, and $-R$. Observe that this curve has rank 1, that Q is a nontorsion point on the curve, and that $R = Q + (0,0)$ is not in the cyclic group generated by Q, so one way to generate infinitely many suitable points on the curve is to compute $[2]Q, [3]Q$, etc. Montgomery does not state any particular generation mechanism but gives five examples in [36, page 263] that appear to have been generated as $[2]R, [3]R, [4]R, [5]R$, and $[6]R$.

Theorem 7.8. *Let $(s,t) \notin \{(0,0), (-2,\pm4), (6,\pm12)\}$ be a rational point on the curve $T^2 = S^3 - 12S$. Define*

$$d = \frac{-(s-2)^3(s+6)^3(s^2 - 12s - 12)}{1024s^2t^2}.$$

Then the Edwards curve $E : x^2 + y^2 = 1 + dx^2y^2$ has $\mathbf{Q}$-torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$ and has a nontorsion point (x_1, y_1) where

$$x_1 = \frac{8t(s^2 + 12)}{(s-2)(s+6)(s^2 + 12s - 12)} \quad \text{and} \quad y_1 = \frac{-4s(s^2 - 12s - 12)}{(s-2)(s+6)(s^2 - 12)}.$$

Proof. Note that the restrictions on (s,t) ensure that all denominators are nonzero and that $d \notin \{0,1\}$.

Define $x_3 = 4t/((s-2)(s+6))$ and $y_3 = -8s/((s-2)(s+6))$. A straightforward calculation shows that $x_3^2 + y_3^2 = 1 + dx_3^2y_3^2 = -2y_3$, so (x_3, y_3) is a point of order 3 on E by Theorem 3.3, so the torsion group of E has a subgroup isomorphic to $\mathbf{Z}/12\mathbf{Z}$. By Mazur's theorem, the torsion group of E is isomorphic to $\mathbf{Z}/12\mathbf{Z}$.

Another straightforward calculation shows that $x_1^2 + y_1^2 = 1 + dx_1^2 y_1^2$, so (x_1, y_1) is a curve point. If it were a torsion point, then the ratio $(x_1 : y_1) = (-2t(s^2 - 12)(s^2 + 12) : s(s^2 - 12s - 12)(s^2 + 12s - 12))$ would have to be equal to $(0 : 1)$ or $(1 : 0)$ or $(\pm x_3 : y_3) = (\pm t : 2s)$ or $(\pm y_3 : x_3) = (\pm 2t : s^2 - 12)$ by Theorem 6.3; this is also avoided by the restrictions on (s, t). □

8. EDWARDS CURVES WITH SMALL PARAMETERS, LARGE TORSION, AND POSITIVE RANK

One way to save time in computations on twisted Edwards curves is to choose small curve parameters a and d and a small-height nontorsion base point $(X_1 : Y_1 : Z_1)$; see Section 2.8. Another way to save time is to construct curves with large **Q**-torsion group and positive rank; see Section 7. Unfortunately, essentially all of the **Z**/12**Z** and **Z**/2**Z** × **Z**/8**Z** curves constructed in Section 7 have a, d, X_1, Y_1, Z_1 of large height.

Our aim in this section is to combine these two time-saving techniques, finding twisted Edwards curves that simultaneously have small parameters a, d, a small-height nontorsion point $(X_1 : Y_1 : Z_1)$, and large torsion over **Q**.

Overall we found more than 100 small Edwards curves having small-height nontorsion points and at least 12 torsion points over **Q**. For the complete list see `http://eecm.cr.yp.to/goodcurves.html`. The number of d's below height H appears to grow as roughly $\log H$; for comparison, the Atkin-Morain procedure discussed in Section 7 generates only about $\sqrt{\log H}$ examples below height H. Of course, one can easily write down many more small curves if one is willing to sacrifice some torsion, as illustrated by the Suyama construction.

8.1. Torsion group **Z**/12**Z**.
Theorem 6.4 gives a complete parameterization of all Edwards curves with torsion group isomorphic to **Z**/12**Z**. Any rational point $(u, x_3, y_3, d, x_1, y_1)$ on the surface described by

$$x_3 = \frac{u^2 - 1}{u^2 + 1}, \; y_3 = -\frac{(u-1)^2}{u^2 + 1}, \; d = \frac{(u^2 + 1)^3(u^2 - 4u + 1)}{(u-1)^6(u+1)^2}, \; x_1^2 + y_1^2 = 1 + dx_1^2 y_1^2$$

gives us a suitable curve for ECM if $u \notin \{0, \pm 1\}$ and (x_1, y_1) is not a torsion point. Theorem 6.3 lists all affine torsion points.

Assume, without loss of generality, that $|u| > 1$: otherwise replace u by $1/u$, obtaining the same d. Write u as a/b for integers a, b satisfying $0 < |b| < a$. Define $e = (a^2 - b^2)/x_1$ and $f = -(a-b)^2/y_1$, and assume, without loss of generality, that e, f are integers: otherwise scale a, b appropriately. The curve equation $x_1^2 + y_1^2 = 1 + dx_1^2 y_1^2$ now implies, after some simplification, the $(1, 1, 2, 2)$-weighted-homogeneous equation

$$(e^2 - (a^2 - b^2)^2)(f^2 - (a - b)^4) = 16a^3 b^3(a^2 - ab + b^2).$$

We found many small solutions to this equation, and thus many of the desired Edwards curves, as follows. We considered a range of positive integers a. For each a we enumerated integers b with $0 < |b| < a$. For each (a, b) we enumerated all divisors of $16a^3 b^3(a^2 - ab + b^2)$ and added $(a^2 - b^2)^2$ to each divisor. For each sum of the form e^2 we added $(a - b)^4$ to the complementary divisor, checked for a square, checked that the corresponding (x_1, y_1) was a nontorsion point, etc.

After about a week of computation on some computers at LORIA we had found 78 different values of d and checked that we had 78 different j-invariants. Here are two examples:

- the very small solution $(a, b, e, f) = (3, 2, 23, 7)$ produces the order-3 point $(x_3, y_3) = (5/13, -1/13)$ and the nontorsion point $(x_1, y_1) = (5/23, -1/7)$ on the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ where $d = -11 \cdot 13^3/5^2$;
- the solution $(a, b, e, f) = (15180, -7540, 265039550, 161866240)$ produces the nontorsion point $(3471616/5300791, -201640/63229)$ on the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ where $d = 931391 \cdot 359105^3/140003330048^2$.

8.2. Torsion group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$.

Theorem 6.9 gives a complete parameterization of all Edwards curves with torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ and with a point of order 8 doubling to $(\pm 1, 0)$. Any rational point (u, x_8, d, x_1, y_1) on the surface described by $x_8 = (u^2 + 2u + 2)/(u^2 - 2)$, $d = (2x_8^2 - 1)/x_8^4$, and $x_1^2 + y_1^2 = 1 + dx_1^2y_1^2$ gives us a suitable curve for ECM if $u \notin \{0, -1, -2\}$ and (x_1, y_1) is not a torsion point. Theorem 6.8 lists all affine torsion points.

We consider only $u > \sqrt{2}$. Various transformations of u listed in Theorem 6.9 show that this does not lose any generality: if $0 < u < \sqrt{2}$, then $2/u > \sqrt{2}$, and $2/u$ produces the same curve; if $u < -2$, then $-(u+2) > 0$, and $-(u+2)$ produces the same curve; if $-2 < u < -1$, then $-2(u+1)/(u+2) > 0$, and $-2(u+1)/(u+2)$ produces the same curve; if $-1 < u < 0$, then $-u/(u+1) > 0$, and $-u/(u+1)$ produces the same curve.

Write $u = a/b$, $x_1 = (a^2 + 2ab + 2b^2)/e$, and $y_1 = (a^2 + 2ab + 2b^2)/f$ where a, b, e, f are integers. Then a, b, e, f satisfy the $(1, 1, 2, 2)$-weighted-homogeneous equation

$$(e^2 - (a^2 + 2ab + 2b^2)^2)(f^2 - (a^2 + 2ab + 2b^2)^2) = (4ab(a + b)(a + 2b))^2.$$

We found many small solutions to this equation, and thus many of the desired Edwards curves, by a procedure similar to the procedure used for $\mathbf{Z}/12\mathbf{Z}$. We considered a range of positive integers a. For each a we enumerated integers b between 1 and $\lfloor a/\sqrt{2} \rfloor$. For each (a, b) we enumerated all divisors of $(4ab(a + b)(a + 2b))^2$, added $(a^2 + 2ab + 2b^2)^2$ to each divisor, and searched for squares.

After about a week of computation on some computers at LORIA, we had found 25 different values of d and checked that we had 25 different j-invariants. Here are two examples:

- the very small solution $(a, b, e, f) = (3, 1, 19, 33)$ produces the order-8 point $(x_8, x_8) = (17/7, 17/7)$ and the nontorsion point $(x_1, y_1) = (17/19, 17/33)$ on the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ where $d = 161^2/17^4$;
- the solution $(a, b, e, f) = (24882, 9009, 258492663, 580153002)$ produces the nontorsion point $(x_1, y_1) = (86866/18259, 8481/4001)$ on the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ where $d = 5657719^2/3341^4$.

9. The impact of large torsion

This section reports various measurements of the success probability of EECM-MPFQ. These measurements demonstrate the importance of choosing a curve with a large torsion group. They also demonstrate the inaccuracy of several common methods of estimating the success probability of ECM.

9.1. **Impact of torsion for 20-bit primes.** There are exactly 38635 primes between 2^{19} and 2^{20}. As an experiment we fed each of these primes to EECM-MPFQ with $B_1 = 256$ and $d_1 = 1$. It turned out that the first curve configured into EECM-MPFQ finds 12467, i.e., 32.2687%, of these primes. This curve is the Edwards curve $x^2 + y^2 = 1 - (24167/25)x^2y^2$, with base point $P = (5/23, -1/7)$; this curve has torsion group isomorphic to $\mathbf{Z}/12\mathbf{Z}$.

We then modified EECM-MPFQ to instead start with the curve $x^2 + y^2 = 1 + (25921/83521)x^2y^2$, with base point $P = (13/7, 289/49)$, and repeated the same experiment. This curve has torsion group isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$; it is one of the curves that EECM-MPFQ normally tries, although not the first in the list. This curve finds 32.8433% of the primes.

We then made a more drastic modification to EECM-MPFQ, trying two new curves with smaller torsion groups. The curve $x^2 + y^2 = 1 + (1/36)x^2y^2$, with base point $P = (8,9)$, has torsion group only $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$ and finds only 27.4854% of the primes, losing a factor 1.17 compared to the original $\mathbf{Z}/12\mathbf{Z}$ curve. The curve $x^2 + y^2 = 1 + (1/3)x^2y^2$, with base point $P = (2,3)$, has torsion group only $\mathbf{Z}/4\mathbf{Z}$ and finds only 23.4709% of the primes, losing a factor 1.37 compared to the original $\mathbf{Z}/12\mathbf{Z}$ curve.

9.2. **Impact of torsion for 30-bit primes.** As a larger experiment we replaced the 38635 20-bit primes by a random sample of 65536 distinct 30-bit primes and increased (B_1, d_1) from $(256, 1)$ to $(1024, 1)$. The same four curves again had remarkably different performance:

- 12.1597% of the primes were found by the $\mathbf{Z}/12\mathbf{Z}$ curve.
- 11.9751% of the primes were found by the $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ curve.
- 9.8465% of the primes were found by the $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$ curve.
- 9.0073% of the primes were found by the $\mathbf{Z}/4\mathbf{Z}$ curve.

For comparison, GMP-ECM with a typical Suyama curve (specifically $\sigma = 10$) finds 11.6837% of the same primes. We also tried GMP-ECM's Pollard $p - 1$ option; it found 6.3507% of the same primes. Normally the $p - 1$ method is assumed to be a helpful first step before ECM, because it uses fewer multiplications per bit than an elliptic curve, but we comment that this benefit is reduced by the $p - 1$ curve (a hyperbola) having torsion group only $\mathbf{Z}/2\mathbf{Z}$.

Figures 9.1 and 9.2 show the results of similar measurements for the same four EECM curves for many prime powers B_1: specifically, every prime power $B_1 \le 500$ for the 20-bit primes, and every prime power $B_1 \le 2000$ for the 30-bit primes. The figures show that $\mathbf{Z}/12\mathbf{Z}$ (black) and $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ (blue) are consistently better than $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$ (blue, lower) and $\mathbf{Z}/4\mathbf{Z}$ (black, lower).

The figures also include measurements for the same GMP-ECM Suyama curve (red) and $p - 1$ (red, lower). When B_1 is large, the EECM-MPFQ $\mathbf{Z}/12\mathbf{Z}$ and $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ curves find significantly more primes than the GMP-ECM Suyama curve.

9.3. **Review of methods of estimating the success probability.** Consider the fraction of primes $p \in [L, R]$ found by stage 1 of ECM with a particular curve E, point $P \in E(\mathbf{Q})$, and smoothness bound B_1. Assume that E is chosen to guarantee t as a divisor of $\#E(\mathbf{F}_p)$.

Standard practice in the literature is to estimate this fraction through the following series of heuristic approximations, where $\langle P \bmod p \rangle$ means the subgroup of $E(\mathbf{F}_p)$ generated by P:

$$\Pr[\text{uniform random prime } p \in [L, R] \text{ has } B_1\text{-powersmooth } \#\langle P \bmod p\rangle]$$

$$\overset{?}{\approx} \Pr[\text{uniform random prime } p \in [L, R] \text{ has } B_1\text{-powersmooth } \#E(\mathbf{F}_p)]$$

$$\overset{?}{\approx} \Pr[\text{uniform random} \in t\mathbf{Z} \cap [(\sqrt{L} - 1)^2, (\sqrt{R} + 1)^2] \text{ is } B_1\text{-powersmooth}]$$

$$\overset{?}{\approx} \Pr[\text{uniform random} \in t\mathbf{Z} \cap [L, R] \text{ is } B_1\text{-powersmooth}]$$

$$\overset{?}{\approx} \Pr[\text{uniform random} \in t\mathbf{Z} \cap [1, R] \text{ is } B_1\text{-powersmooth}]$$

$$\overset{?}{\approx} \Pr[\text{uniform random} \in \mathbf{Z} \cap [1, R/t] \text{ is } B_1\text{-powersmooth}]$$

$$\overset{?}{\approx} \rho(u) \text{ where } B_1^u = R/t$$

$$\overset{?}{\approx} 1/u^u.$$

Here "B_1-powersmooth" means "having no prime-power divisors larger than B_1," and ρ is Dickman's rho function introduced in [24]. Similar comments apply to stage 2, with B_1-powersmoothness replaced by a more complicated notion of smoothness and with ρ replaced by a standard generalization.

For example, Montgomery in [37, Section 7] estimated the success chance of a curve with 16 torsion points over $\mathbf{Q}$ as the B_1-powersmoothness chance for a uniform random integer in $[1, p/16]$. Similarly, Silverman and Wagstaff in [41] estimated the success chance of a Suyama curve as the B_1-powersmoothness chance for a uniform random integer in $[1, p/12]$, following Brent's comment in [18, Section 9.3] that choosing a Suyama curve "effectively reduces p to $p/12$ in the analysis." (As mentioned in Section 7, a typical Suyama curve has only 6 torsion points over $\mathbf{Q}$, but a Suyama curve modulo p is guaranteed to have order in $12\mathbf{Z}$.) Brent, Montgomery, et al. used Dickman's rho function to estimate the B_1-powersmoothness chance for a uniform random integer.

9.4. Inaccuracy of the estimates.
There are many reasons to question the accuracy of the above approximations:

- Dickman's rho function ρ is asymptotically $1/u^u$ in the loose sense that $(\log \rho(u))/(-u \log u) \to 1$ as $u \to \infty$, but is not actually very close to $1/u^u$: for example, $\rho(2) \approx 1.23/2^2$, $\rho(3) \approx 1.31/3^3$, and $\rho(4) \approx 1.26/4^4$.
- For each $u \geq 0$, the B_1-smoothness probability for an integer in $[1, B_1^u]$ converges to $\rho(u)$ as $B_1 \to \infty$, and the same is true for B_1-powersmoothness, but the convergence is actually quite slow.
- Multiplying an element of $\mathbf{Z} \cap [1, R/16]$ by 16 never gains powersmoothness but can lose powersmoothness when the original exponent of 2 was large, not an uncommon event among powersmooth integers.
- The ratio of smoothness probabilities for (e.g.) $[1, B_1^u]$ and $[(1/2)B_1^u, B_1^u]$ converges to 1 as $B_1 \to \infty$, but the convergence is again quite slow.

- Lenstra commented in [33, page 660] that an elliptic curve has even order with probability approximately 2/3, not 1/2. Many subsequent reports (for example, by Brent in [18, Table 3] and McKee in [35, Section 2]) have lent support to the idea that elliptic-curve orders are somewhat more likely to be smooth than uniform random integers.
- The group order $\#E(\mathbf{F}_p)$ is a multiple of the point order $\#\langle P \text{ in } E(\mathbf{F}_p)\rangle$. The ratio is usually small but often enough to change powersmoothness, as illustrated by the $s = 420$ example in Section 4.1.

The overall error is not extremely large but can easily be large enough to interfere with optimization.

Recall that the curve $x^2 + y^2 = 1 - (24167/25)x^2y^2$, with 12 torsion points, finds 32.2687% of the primes in $[2^{19}, 2^{20}]$ with $B_1 = 256$ and $d_1 = 1$; and that changing to three other curves with 16, 8, and 4 torsion points changes 32.2687% to 32.8433%, 27.4854%, and 23.4709% respectively. We computed several of the standard estimates for these four success probabilities:

- A uniform random element of $12\mathbf{Z}\cap[2^{19}, 2^{20}]$ has a 23.6067% chance of being 256-powersmooth. Note that this probability drastically underestimates the actual ECM smoothness chance. Changing 12 to 16, 8, 4 changes 23.6067% to 24.8192%, 20.5777%, and 16.8006% respectively.
- A uniform random element of $12\mathbf{Z}\cap[1, 2^{20}]$ has a 30.0317% chance of being 256-powersmooth. Changing 12 to 16, 8, 4 changes 30.0317% to 31.3019%, 26.4328%, and 21.8632% respectively.
- A uniform random element of $\mathbf{Z}\cap[1, 2^{20}/12]$ has a 30.7652% chance of being 256-powersmooth. Changing 12 to 16, 8, 4 changes 30.7652% to 33.3694%, 27.3689%, and 22.2511% respectively.
- If $u = (\log(2^{20}/12))/\log 256$, then $\rho(u) \approx 28.1894\%$. Changing 12 to 16, 8, 4 changes 28.1894% to 30.6853%, 24.9832%, and 20.2442% respectively.
- If $u = (\log(2^{20}/12))/\log 256$, then $u^{-u} \approx 22.8824\%$. Changing 12 to 16, 8, 4 changes 22.8824% to 25%, 20.1540%, and 16.1283% respectively.

These approximations make 16 seem better than 12 by factors of 1.051, 1.042, 1.085, 1.089, and 1.093, when in fact 16 is better than 12 by a factor of only 1.018.

Figure 9.1 includes, for many prime powers B_1, the B_1-powersmoothness chance of a uniform random element of $t\mathbf{Z} \cap [2^{19}, 2^{20}]$ for four values of t (green and gray graphs, bumpy), and $\rho((\log(2^{20}/t))/\log B_1)$ for four values of t (green and gray graphs, smooth). Figure 9.2 includes analogous results for 30-bit primes. It is clear that the ρ value is a poor approximation to the powersmoothness chance, and that the powersmoothness chance is a poor approximation to the ECM success chance.

One can ask whether better approximations are possible. We comment that a fast algorithm to compute tight bounds on smoothness probabilities appeared in [5], and that the same algorithm can be adapted to handle powersmoothness, local conditions such as more frequent divisibility by 2, etc. However, one can also ask whether approximations are necessary in the first place. ECM is most frequently used to find rather small primes (for example, inside the number-field sieve), and for those primes one can simply measure ECM's performance by experiment.
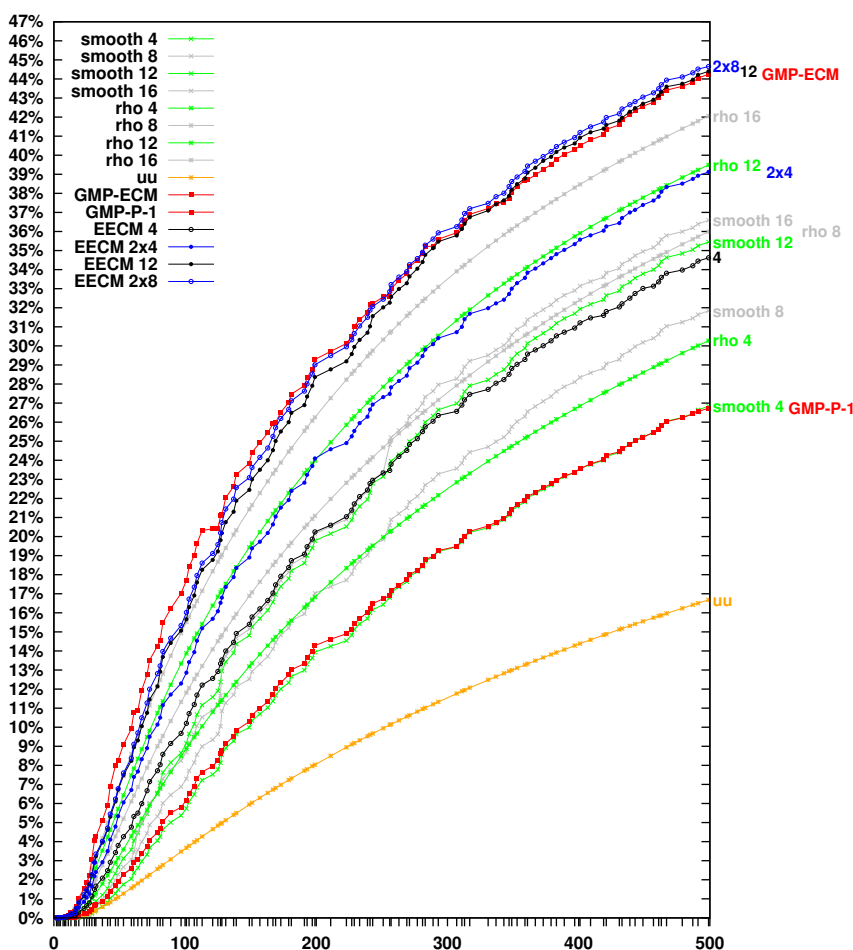
FIGURE 9.1. For the set of all 38635 20-bit primes: Measured stage-1 success probabilities for six curves, and nine estimates. Horizontal axis is B_1. Vertical axis is probability. Graphs from top to bottom on right side: (blue, bumpy) EECM-MPFQ with a $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ curve; (black, bumpy) EECM-MPFQ with a $\mathbf{Z}/12\mathbf{Z}$ curve; (red, bumpy) GMP-ECM with a Suyama curve; (gray, smooth) the ρ approximation to smoothness probability for $[1, 2^{20}/16]$; (green, smooth) the ρ approximation for $[1, 2^{20}/12]$; (blue, bumpy) EECM-MPFQ with a $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$ curve; (gray, bumpy) powersmoothness probability for $16\mathbf{Z} \cap [2^{19}, 2^{20}]$; (gray, smooth) the ρ approximation for $[1, 2^{20}/8]$; (green, bumpy) powersmoothness probability for $12\mathbf{Z} \cap [2^{19}, 2^{20}]$; (black, bumpy) EECM-MPFQ with a $\mathbf{Z}/4\mathbf{Z}$ curve; (gray, bumpy) powersmoothness probability for $8\mathbf{Z} \cap [2^{19}, 2^{20}]$; (green, smooth) the ρ approximation for $[1, 2^{20}/4]$; (green, bumpy) powersmoothness probability for $4\mathbf{Z} \cap [2^{19}, 2^{20}]$; (red, bumpy) GMP-ECM with $p-1$; (orange, smooth) the u^{-u} approximation for $[1, 2^{20}]$.
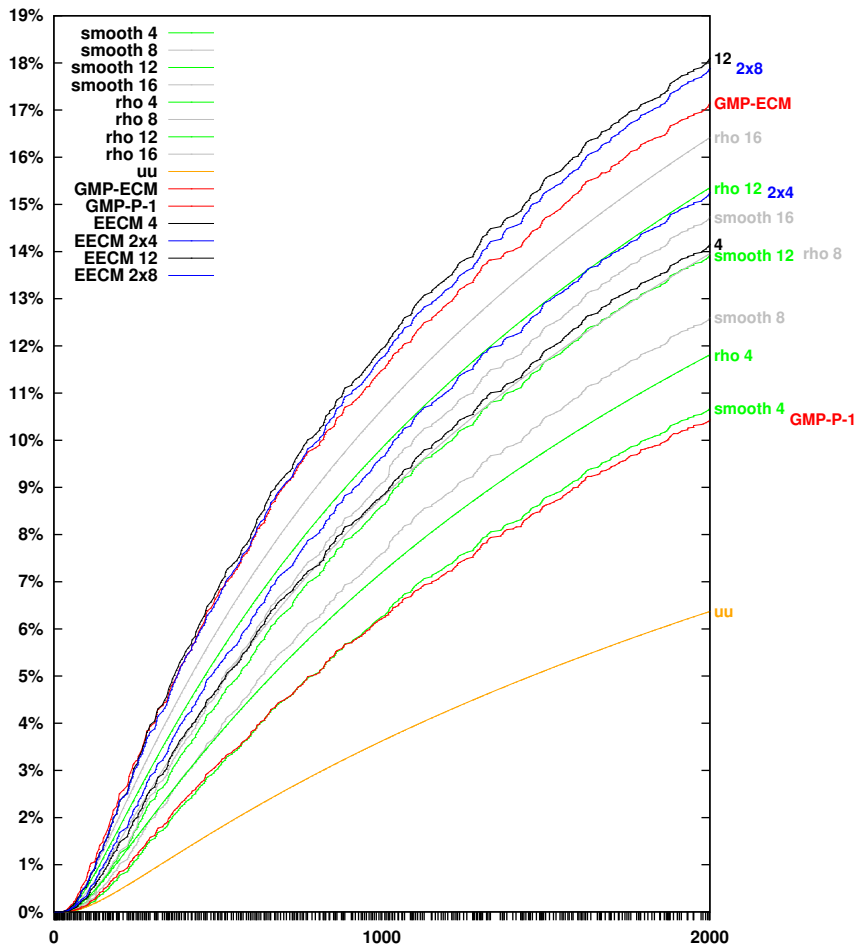
FIGURE 9.2. For a sample of 65536 30-bit primes: Measured stage-1 success probabilities for six curves, and nine estimates. Horizontal axis is B_1. Vertical axis is probability. Graphs from top to bottom on right side: (black, bumpy) EECM-MPFQ with a $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/8\mathbf{Z}$ curve; (blue, bumpy) EECM-MPFQ with a $\mathbf{Z}/12\mathbf{Z}$ curve; (red, bumpy) GMP-ECM with a Suyama curve; (gray, smooth) the ρ approximation to smoothness probability for $[1, 2^{30}/16]$; (green, smooth) the ρ approximation for $[1, 2^{30}/12]$; (blue, bumpy) EECM-MPFQ with a $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$ curve; (gray, bumpy) powersmoothness probability for $16\mathbf{Z} \cap [2^{29}, 2^{30}]$; (black, bumpy) EECM-MPFQ with a $\mathbf{Z}/4\mathbf{Z}$ curve; (gray, smooth) the ρ approximation for $[1, 2^{30}/8]$; (green, bumpy) powersmoothness probability for $12\mathbf{Z} \cap [2^{29}, 2^{30}]$; (gray, bumpy) powersmoothness probability for $8\mathbf{Z} \cap [2^{29}, 2^{30}]$; (green, smooth) the ρ approximation for $[1, 2^{30}/4]$; (green, bumpy) powersmoothness probability for $4\mathbf{Z} \cap [2^{29}, 2^{30}]$; (red, bumpy) GMP-ECM with $p - 1$; (orange, smooth) the u^{-u} approximation for $[1, 2^{30}]$.

10. Choosing parameters

This section reports EECM-MPFQ's overall performance at finding various sizes of primes, when the parameters B_1, d_1, etc. are chosen sensibly.

10.1. **Normalizing the success probability.** Stage 1 will almost never find any factors of n if B_1 is very small, and stage 2 will almost never find any factors of n if d_1 and e are very small. The success probability increases as the parameters increase, and eventually reaches 1 (for any particular size of prime); however, the costs of stage 1 and stage 2 then become enormously large. It is generally best to use intermediate parameters that balance the cost of each curve against the success probability of the curve, and to compensate for a low success probability by trying several curves.

Montgomery in [36, Table 7.4.1] computed an "expected time" obtained by multiplying an "expected number of curves" by an "estimated time per curve". The "expected number of curves" was the reciprocal of an estimate of the success probability per curve. The "estimated time per curve" was $5.5B_1$ milliseconds for stage 1, $105d_1$ milliseconds for initial elliptic-curve operations in stage 2, etc. Montgomery selected the constants 5.5, 105, etc. to approximately fit timings of his ECM implementation on a DEC 5000.

We instead report actual measurements of EECM-MPFQ's price-performance ratio. Specifically, we report the actual number of modular multiplications used by an EECM-MPFQ curve for both stage 1 and stage 2, divided by the actual success probability of that curve within a target set of primes. To simplify these reports we count $\mathbf{S}$ as $\mathbf{M}$, we count multiplications by small numbers (such as the coordinates of the base point) as $\mathbf{M}$, and we skip the fast-polynomial-arithmetic variant described in Section 5.3.

Our experiments actually used slightly fewer modular multiplications per prime, because primes found in stage 1 did not incur the costs of stage 2 (and primes found during the batched division in stage 2 did not incur the remaining costs of stage 2). This cost reduction is reported as "savings" in Table 10.1 below. An application that uses EECM with a similar distribution of primes within its inputs will see a similar savings. On the other hand, an application faced with a large pool of inputs, where primes of the desired size appear within relatively few inputs, will see smaller savings.

We also report, later in the section, the number of clock cycles used by EECM-MPFQ for both stage 1 and stage 2, again divided by success probability. The number of multiplications per prime found is a simpler measure than the number of cycles per prime found, and is an adequate measure for seeing most of this paper's improvements, but it is not adequate for seeing the speedup from GMP to MPFQ.

10.2. **Impact of B_1 and d_1 for 20-bit primes.** Recall from Section 9 that EECM-MPFQ's first curve finds 12467 of the 38635 20-bit primes using $B_1 = 256$ and $d_1 = 1$. This experiment used a total of $65900078\mathbf{M} + 55479860\mathbf{S}$; each successful prime therefore consumed 9736 modular multiplications.

We tried the same curve again using $B_1 = 37$, $d_1 = 90$, $e = 1$, and $\#\{i\} = \#\{j\}$. This time EECM-MPFQ found 1527 primes in stage 1 and an additional 14017 primes in stage 2 (1242 during conversion to affine and 12775 at the end of stage 2), for an overall success probability of $15544/38635 \approx 40.2329\%$. The cost of handling a worst-case input was $734\mathbf{M} + 212\mathbf{S}$, and if EECM-MPFQ had incurred
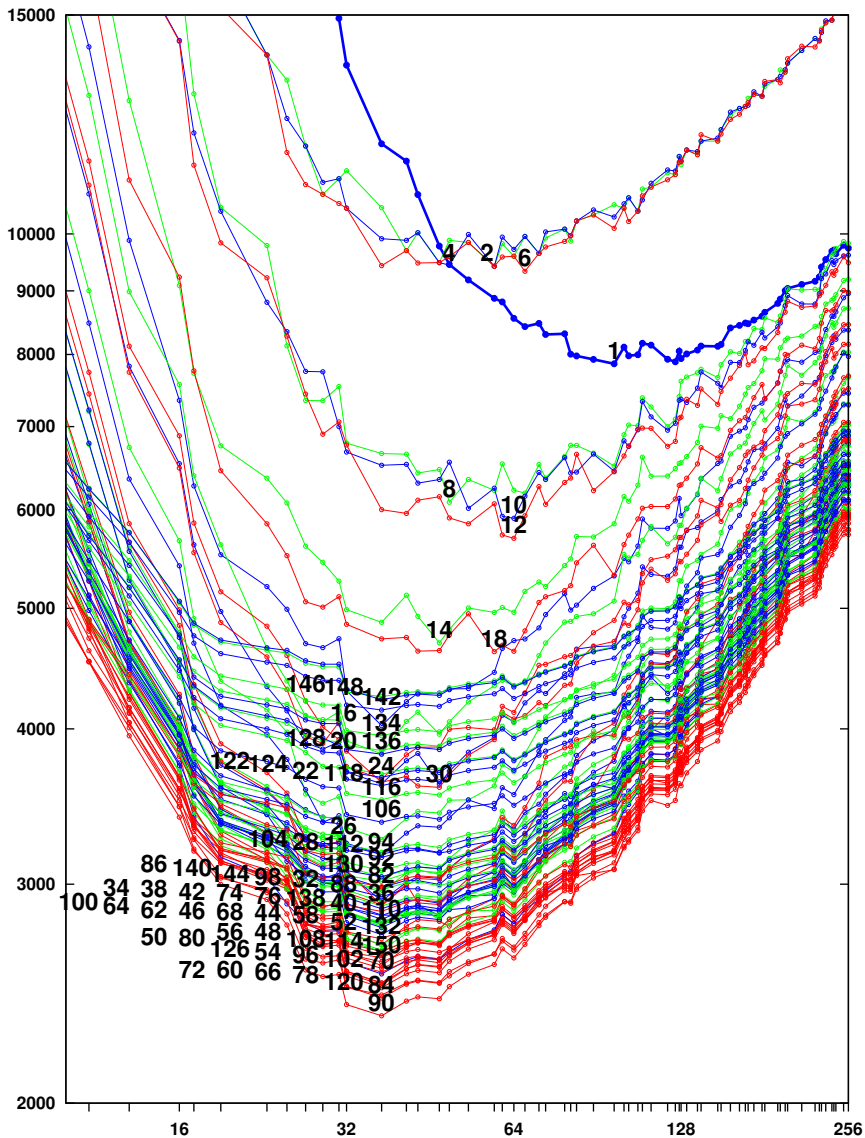
FIGURE 10.1. Cost ratio for the curve $x^2 + y^2 = 1 - (24167/25)x^2y^2$ with torsion group $\mathbf{Z}/12\mathbf{Z}$ and base point $(5/23, -1/7)$. The vertical axis is the number of modular multiplications used for both stage 1 and stage 2, times the number of primes in $[2^{19}, 2^{20}]$, divided by the number of primes found. The horizontal axis is B_1. The label inside the graph is d_1. Data points with the same d_1 are connected by line segments.

TABLE 10.1. Cost ratio for sample sets of b-bit primes for $b \in \{15, 16, \ldots, 50\}$. "Samples" is the size of the sample set. "Pr" is the success probability, within the sample set, of the curve $x^2 + y^2 = 1 - (24167/25)x^2y^2$ with base point $(5/23, -1/7)$ and torsion group $\mathbf{Z}/12\mathbf{Z}$, using EECM-MPFQ parameters B_1, d_1, e, and $\#\{i\}/\#\{j\} = R$. "Mults" is the number of modular multiplications used for both stage 1 and stage 2. "Ratio" is "mults" divided by "Pr"; i.e., the number of modular multiplications per prime found. The logarithm of "ratio" is "power" times $\sqrt{2 \log 2^b \log \log 2^b}$. "Savings" is the fraction of modular multiplications saved within the sample set by primes found before the end of stage 2. "Cycles" is the number of cycles used for both stage 1 and stage 2 on a 3.2GHz AMD Phenom II X4 (100f42) for n between 192 bits and 256 bits, divided by "Pr"; i.e., the number of cycles per prime found.

bits	B_1	d_1	e	R	samples	Pr	mults	ratio	power	savings	cycles
15	16	60	1	1	1612	65.4467%	475	725.8	0.9440	6.0743%	398383
16	16	60	1	1	3030	50.5941%	475	938.8	0.9369	3.7812%	519846
17	27	60	1	1	5709	54.6856%	632	1155.7	0.9250	4.4719%	542882
18	27	90	1	1	10749	53.7073%	816	1519.3	0.9231	4.2088%	655581
19	37	90	1	1	20390	50.4708%	946	1874.4	0.9145	4.0127%	753288
20	37	90	1	1	38635	40.2329%	946	2351.3	0.9092	2.7869%	940225
21	37	90	1	1	73586	30.9787%	946	3053.7	0.9088	1.8815%	1229364
22	47	120	1	1	140336	33.0086%	1292	3914.1	0.9075	2.1786%	1413109
23	64	120	1	1	268216	30.3744%	1491	4908.7	0.9045	1.9515%	1682287
24	81	210	1	1	513708	36.7985%	2276	6185.0	0.9026	2.5365%	1971371
25	97	210	1	1	985818	31.7403%	2427	7646.4	0.8994	2.1164%	2390659
26	97	210	1	1	1048576	25.4204%	2427	9547.5	0.8976	1.5147%	2973591
27	131	210	1	1	1048576	24.4857%	2904	11860.0	0.8959	1.4755%	3528747
28	131	210	1	1	1048576	19.7381%	2904	14712.7	0.8944	1.0864%	4371407
29	149	210	1	1	1048576	16.5716%	3065	18495.5	0.8945	0.8797%	5363333
30	149	210	1	1	1048576	13.1368%	3065	23331.5	0.8953	0.6386%	6769323
31	263	210	1	2	1048576	18.4570%	5376	29127.1	0.8953	1.0937%	7834148
32	263	210	1	2	1048576	15.0913%	5376	35623.1	0.8938	0.8394%	9615434
33	263	210	1	2	1048576	12.1644%	5376	44194.5	0.8939	0.6248%	11915678
34	343	330	1	1	1048576	12.3212%	6787	55084.0	0.8945	0.6643%	14534927
35	389	420	1	1	1048576	12.3528%	8384	67871.0	0.8944	0.6747%	17488151
36	433	420	1	1	1048576	10.6944%	8892	83146.3	0.8941	0.5658%	21345174
37	521	420	1	1	1048576	9.7486%	9909	101644.8	0.8937	0.4983%	25652386
38	521	420	1	1	1048576	7.9452%	9909	124717.5	0.8939	0.3825%	31436961
39	587	420	1	1	1048576	6.8847%	10621	154270.3	0.8948	0.3185%	38319718
40	587	420	1	1	1048576	5.6551%	10621	187812.8	0.8946	0.2510%	47190133
41	937	630	1	1	1048576	7.8935%	18236	231026.5	0.8954	0.4196%	56113371
42	1031	630	1	1	1048576	6.9196%	19386	280161.7	0.8953	0.3544%	67437743
43	1031	630	1	1	1048576	5.7678%	19386	336106.1	0.8945	0.2840%	81087478
44	1031	630	1	1	1048576	4.6908%	19386	413273.7	0.8957	0.2201%	99684763
45	1151	630	1	1	1048576	4.1508%	20833	501906.6	0.8960	0.1901%	121979006
46	1319	630	1	1	1048576	3.7610%	22884	608454.3	0.8964	0.1619%	144341609
47	1709	840	1	1	1048576	4.3684%	32129	735486.6	0.8966	0.2170%	175028834
48	1889	840	1	1	1048576	3.8442%	34195	889529.8	0.8970	0.1785%	211435752
49	2221	840	1	1	1048576	3.5111%	37877	1078765.6	0.8977	0.1548%	251473421
50	2521	840	2	1	1048576	3.3565%	42981	1280546.8	0.8971	0.1686%	296542182

this cost for every input then it would have used a total of $28358090\mathbf{M} + 8190620\mathbf{S}$, i.e., just 2351 modular multiplications per successful prime. EECM-MPFQ actually used only $27357827\mathbf{M} + 8172296\mathbf{S}$, saving 2.8%, because primes found in stage 1 did not incur the costs of stage 2.

Figure 10.1 shows the results of similar computations for many more pairs (B_1, d_1). The figure quantifies the well-known importance of stage 2: $d_1 = 1$ costs more than three times as many modular multiplications as the best d_1 for 20-bit primes. The figure also confirms the idea that d_1 should have several small prime factors. Each computation used $e = 1$ and used EECM-MPFQ's default ratio $\#\{i\}/\#\{j\} = 1$.

10.3. Other sizes of primes. Table 10.1 reports the effectiveness of good choices of $(B_1, d_1, e, \#\{i\}/\#\{j\})$ for 15-bit primes, 16-bit primes, 17-bit primes, and so on through 50-bit primes. The "power" column shows that EECM-MPFQ uses fewer than $\exp(0.9\sqrt{2\log 2^b \log\log 2^b})$ modular multiplications per b-bit prime found, for each $b \in \{25, 26, \ldots, 50\}$. See `http://eecm.cr.yp.to` for performance data for larger values of b.

The conventional wisdom (see, e.g., [18]) is that one should use Pollard's rho method for primes up to about 30 bits and then switch over to ECM. We present ECM performance data for much smaller sizes as a basis for comparison and for future improvements. Our guess is that improvements in ECM have drastically reduced the optimal rho-to-ECM cutoff.

Acknowledgments

References

[1] *SPEED: software performance enhancement for encryption and decryption*, 2007. `http://www.hyperelliptic.org/SPEED`. See [27].

[2] Michel Abdalla, Paulo S. L. M. Barreto (editors), *Progress in cryptology—LATINCRYPT 2010, first international conference on cryptology and information security in Latin America, Puebla, Mexico, August 8–11, 2010, proceedings*, Lecture Notes in Computer Science, 6212, Springer, 2010. See [9].

[3] A. O. L. Atkin, Francois Morain, *Finding suitable curves for the elliptic curve method of factorization*, Mathematics of Computation **60** (1993), 399–405. ISSN 0025-5718. `http://www.lix.polytechnique.fr/~morain/Articles/articles.english.html`. Citations in this document: §1.2, §7, §7.1. MR93k:11115

[4] Michael A. Bennett, Bruce C. Berndt, Nigel Boston, Harold G. Diamond, Adolf J. Hildebrand, Walter Philipp (editors), *Number theory for the millennium. I: papers from the conference held at the University of Illinois at Urbana-Champaign, Urbana, IL, May 21–26, 2000*, A. K. Peters, Natick, Massachusetts, 2002. ISBN 1-56881-126-8. MR2003h:11004. See [5].

[5] Daniel J. Bernstein, *Arbitrarily tight bounds on the distribution of smooth integers*, in Number Theory 2000 [4] (2002), 49–66. `http://cr.yp.to/papers.html#psi`. Citations in this document: §9.4. MR1956218 (2004b:11135)

[6] Daniel J. Bernstein, *Scaled remainder trees* (2004). `http://cr.yp.to/papers.html#scaledmod`. Citations in this document: §5.3.

[7] Daniel J. Bernstein, *Fast multiplication and its applications*, in Algorithmic Number Theory [19] (2008), 325–384. `http://cr.yp.to/papers.html#multapps`. Citations in this document: §5.3. MR2467550 (2010a:68186)

[8] Daniel J. Bernstein, Peter Birkner, Marc Joye, Tanja Lange, Christiane Peters, *Twisted Edwards curves*, in Africacrypt 2008 [45] (2008), 389–405. `http://eprint.iacr.org/2008/013`. Citations in this document: §2.2, §2.4, §2.5, §3, §7. MR2482341 (2010e:11057)

[9] Daniel J. Bernstein, Peter Birkner, Tanja Lange, *Starfish on strike*, in Latincrypt 2010, [2] (2010), 62–80. `http://eprint.iacr.org/2010/367`. Citations in this document: §1.2.

[10] Daniel J. Bernstein, Peter Birkner, Tanja Lange, Christiane Peters, *Optimizing double-base elliptic-curve single-scalar multiplication*, in Indocrypt 2007 [42] (2007), 167–182. `http://eprint.iacr.org/2007/414`. Citations in this document: §1.1. MR2570254 (2010k:94034)

[11] Daniel J. Bernstein, Tien-Ren Chen, Chen-Mou Cheng, Tanja Lange, Bo-Yin Yang, *ECM on graphics cards*, in Eurocrypt 2009 [30] (2009), 483–501. `http://eprint.iacr.org/2008/480`. Citations in this document: §1. MR2538444

[12] Daniel J. Bernstein, Tanja Lange, *Explicit-formulas database* (2007). `http://hyperelliptic.org/EFD`. Citations in this document: §2.

[13] Daniel J. Bernstein, Tanja Lange, *Faster addition and doubling on elliptic curves*, in Asiacrypt 2007 [32] (2007), 29–50. `http://eprint.iacr.org/2007/286`. Citations in this document: §1.1, §2.2, §2.4, §2.9. MR2565722 (2011d:11125)

[14] Daniel J. Bernstein, Tanja Lange, *Inverted Edwards coordinates*, in AAECC 2007 [17] (2007), 20–27. `http://eprint.iacr.org/2007/410`. Citations in this document: §1.1, §2.5.

[15] Daniel J. Bernstein, Tanja Lange, *Analysis and optimization of elliptic-curve single-scalar multiplication*, in Fq8 [38] (2008), 1–19. `http://eprint.iacr.org/2007/455`. Citations in this document: §1.1, §4.5. MR2436321 (2010a:94049)

[16] Daniel J. Bernstein, Tanja Lange, *A complete set of addition laws for incomplete Edwards curves*, Journal of Number Theory, **131**, 858–872. `http://eprint.iacr.org/2009/580`. Citations in this document: §2.9, §2.9, §2.9, §3. MR2772476

[17] Serdar Boztas, Hsiao-Feng Lu, *Applied algebra, algebraic algorithms and error-correcting codes, 17th international symposium*, AAECC-17, Bangalore, India, December 16–20, 2007, proceedings, Lecture Notes in Computer Science, 4851, Springer, 2007. ISBN 978-3-540-77223-1. See [14]. MR2640522 (2011a:94003)

[18] Richard P. Brent, *Some integer factorization algorithms using elliptic curves*, Australian Computer Science Communications **8** (1986), 149–163. ISSN 0157–3055. `http://maths.anu.edu.au/~brent/pub/pub102.html`. Citations in this document: §4.2, §4.3, §9.3, §9.4, §10.3. See [44].

[19] Joe Buhler, Peter Stevenhagen, *Algorithmic number theory: lattices, number fields, curves and cryptography*, Mathematical Sciences Research Institute Publications, 44, Cambridge University Press, 2008. ISBN 978-0521808545. MR2009h:11003. See [7].

[20] David V. Chudnovsky, Gregory V. Chudnovsky, *Sequences of numbers generated by addition in formal groups and new primality and factorization tests*, Advances in Applied Mathematics **7** (1986), 385–434. MR88h:11094. Citations in this document: §4.4, §4.4, §4.7.

[21] Romain Cosset, *Factorization with genus 2 curves*, Mathematics of Computation **79** (2010), 1191–1208. `http://arxiv.org/pdf/0905.2325`. Citations in this document: §4.7, §4.7, §4.7, §4.7. MR2600562 (2011d:11289)

[22] Peter de Rooij, *Efficient exponentiation using precomputation and vector addition chains*, in Eurocrypt 1994 [23] (1995), 389–399. MR1479665. Citations in this document: §5.6.

[23] Alfredo De Santis (editor), *Advances in cryptology—EUROCRYPT '94, workshop on the theory and application of cryptographic techniques, Perugia, Italy, May 9–12, 1994, proceedings*, Lecture Notes in Computer Science, 950, Springer, Berlin, 1995. ISBN 3-540-60176-7. MR98h:94001. See [22].

[24] Karl Dickman, *On the frequency of numbers containing primes of a certain relative magnitude*, Arkiv för Matematik, Astronomi och Fysik **22** (1930), 1–14. ISSN 0365-4133. Citations in this document: §9.3.

[25] Harold M. Edwards, *A normal form for elliptic curves*, Bulletin of the American Mathematical Society **44** (2007), 393–422. `http://www.ams.org/bull/2007-44-03/S0273-0979-07-01153-6/home.html`. Citations in this document: §2.2. MR2318157 (2008b:14052)

[26] Pierrick Gaudry, *Fast genus 2 arithmetic based on theta functions*, Journal of Mathematical Cryptology **1** (2007), 243–265. `http://www.loria.fr/~gaudry/publis/arithKsurf.pdf`. Citations in this document: §4.7, §4.7. MR2372155 (2009f:11156)

[27] Pierrick Gaudry, Emmanuel Thomé, *The mpFq library and implementing curve-based key exchanges*, in [1] (2007), 49–64. `http://www.loria.fr/~gaudry/papers.en.html`. Citations in this document: §4.6.

[28] Florian Hess, Sebastian Pauli, Michael E. Pohst (editors), *Algorithmic number theory, proceedings of the 7th international symposium (ANTS-VII) held at the Technische Universität Berlin, Berlin, July 23–28, 2006*, Lecture Notes in Computer Science, 4076, Springer, Berlin, 2006. ISBN 3-540-36075-1. MR2007h:11001. See [47].

[29] Huseyin Hisil, Kenneth Koon-Ho Wong, Gary Carter, Ed Dawson, *Twisted Edwards curves revisited*, in Asiacrypt 2008 [39] (2008). `http://eprint.iacr.org/2008/522`. Citations in this document: §1.1, §2.3, §2.6, §2.6. MR2546103

[30] Antoine Joux (editor), *Advances in cryptology—EUROCRYPT 2009, 28th annual international conference on the theory and applications of cryptographic techniques, Cologne, Germany, April 26–30, 2009, proceedings*, Lecture Notes in Computer Science, 5479, Springer, 2009. ISBN 978-3-642-01000-2. See [11]. MR2590599 (2010j:94007)

[31] Alexander Kruppa, *Améliorations de la multiplication et de la factorisation d'entier*, Ph.D. thesis, Université Henri Poincaré Nancy I, 2010. `http://tel.archives-ouvertes.fr/tel-00477005/en/`. Citations in this document: §7.

[32] Kaoru Kurosawa (editor), *Advances in cryptology—ASIACRYPT 2007, 13th international conference on the theory and application of cryptology and information security, Kuching, Malaysia, December 2–6, 2007, proceedings*, Lecture Notes in Computer Science, 4833, Springer, 2007. ISBN 978-3-540-76899-9. See [13]. MR2590581 (2010i:94001)

[33] Hendrik W. Lenstra, Jr., *Factoring integers with elliptic curves*, Annals of Mathematics **126** (1987), 649–673. ISSN 0003-486X. MR89g:11125. `https://openaccess.leidenuniv.nl/bitstream/1887/3826/1/346_086.pdf`. Citations in this document: §1, §9.4.

[34] Barry Mazur, *Rational isogenies of prime degree*, Inventiones Mathematicae **4** (1978), 129–162. Citations in this document: §6. MR482230 (80h:14022)

[35] James McKee, *Subtleties in the distribution of the numbers of points on elliptic curves over a finite prime field*, Journal of the London Mathematical Society **59** (1999), 448–460. Citations in this document: §4.1, §9.4. MR1709178 (2000g:11055)

[36] Peter L. Montgomery, *Speeding the Pollard and elliptic curve methods of factorization*, Mathematics of Computation **48** (1987), 243–264. ISSN 0025-5718. MR88e:11130. `http://www.ams.org/mcom/1987-48-177/S0025-5718-1987-0866113-7/S0025-5718-1987-0866113-7.pdf`. Citations in this document §4.3, §4.4, §7, §7.7. See [44].

[37] Peter L. Montgomery, *An FFT extension of the elliptic curve method of factorization*, Ph.D. thesis, University of California at Los Angeles, 1992. `ftp://ftp.cwi.nl/pub/pmontgom/ucladissertation.psl.gz`. Citations in this document: §5.4, §5.4, §7, §7, §9.3, §10.1. MR2688742

[38] Gary L. Mullen, Daniel Panario, Igor E. Shparlinski (editors), *Finite fields and applications: papers from the 8th international conference held in Melbourne, July 9–13, 2007*, Contemporary Mathematics, 461, American Mathematical Society, 2008. ISBN 978-0-8218-4309-3. MR2009h:11004. See [15].

[39] Josef Pieprzyk (editor), *Advances in cryptology — ASIACRYPT 2008, 14th international conference on the theory and application of cryptology and information security, Melbourne, Australia, December 7–11, 2008*, Lecture Notes in Computer Science, 5350, 2008. ISBN 978-3-540-89254-0. See [29]. MR2590580 (2010j:94005)

[40] John M. Pollard, *Theorems on factorization and primality testing*, Proceedings of the Cambridge Philosophical Society **76** (1974), 521–528. ISBN 0305–0041. MR50:6992. `http://cr.yp.to/bib/entries.html#1974/pollard`. Citations in this document: §4.2, §4.3.

[41] Robert D. Silverman, Samuel S. Wagstaff, Jr., *A practical analysis of the elliptic curve factoring algorithm*, Mathematics of Computation **61** (1993), 445–462. `http://www.ams.org/journals/mcom/1993-61-203/S0025-5718-1993-1122078-7/S0025-5718-1993-1122078-7.pdf`. Citations in this document: §9.3. MR1122078 (93k:11117)

[42] Kannan Srinathan, Chandrasekaran Pandu Rangan, Moti Yung (editors), *Progress in cryptology—INDOCRYPT 2007, 8th international conference on cryptology in India, Chennai, India, December 9–13, 2007, proceedings*, Lecture Notes in Computer Science, 4859, Springer, 2007. ISBN 978-3-540-77025-1. See [10]. MR2574217 (2010g:94137)

[43] William Stein (editor), *Sage Mathematics Software (Version 3.2.3)*, The Sage Group, 2009. `http://www.sagemath.org`. Citations in this document: §6.10.

[44] Hiromi Suyama, *Informal preliminary report (8)*, cited in [18] as personal communication and in [36] (1985). Citations in this document: §7.

[45] Serge Vaudenay, *Progress in cryptology — AFRICACRYPT 2008, First international conference on cryptology in Africa, Casablanca, Morocco, June 11-14, 2008, proceedings*, Lecture Notes in Computer Science, 5023, Springer, 2008. ISBN 978-3-540-68159-5. See [8]. MR2528441 (2009m:94064)

[46] Paul Zimmermann, *50 largest factors found by ECM*. `http://www.loria.fr/~zimmerma/records/top50.html`. Citations in this document: §1.

[47] Paul Zimmermann, Bruce Dodson, *20 years of ECM*, in ANTS VII [28] (2006), 525–542. `http://www.loria.fr/~zimmerma/papers/40760525.pdf`. Citations in this document: §1, §4.2, §4.3, §4.4, §4.5, §5.3, §7. MR2282947 (2007j:11172)

DEPARTMENT OF COMPUTER SCIENCE (MC 152), UNIVERSITY OF ILLINOIS AT CHICAGO, CHICAGO, ILLINOIS 60607–7053
  E-mail address: `djb@cr.yp.to`

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, TECHNISCHE UNIVERSITEIT EINDHOVEN, P.O. BOX 513, 5600 MB EINDHOVEN, NETHERLANDS
  E-mail address: `pbirkner@fastmail.fm`

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, TECHNISCHE UNIVERSITEIT EINDHOVEN, P.O. BOX 513, 5600 MB EINDHOVEN, NETHERLANDS
  E-mail address: `tanja@hyperelliptic.org`

DEPARTMENT OF MATHEMATICS, TECHNICAL UNIVERSITY OF DENMARK, 2800 KGS. LYNGBY, DENMARK
  E-mail address: `c.p.peters@mat.dtu.dk`