The Elliptic Curve Method and Other Integer Factorization Algorithms

John Wright

April 12, 2012

Contents

Chapter 1

Introduction

The *Fundamental Theorem of Arithmetic*, first proved by Gauss [2], states that every positive integer has a unique factorization into primes. That is, for every positive integer N,

$$N = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$$

where the p_i's are distinct primes and each a_i is a positive integer. This paper is motivated by a computational question: given an arbitrary integer N, how might we find a non-trivial factor of N? That is, a factor of N other than $\pm N$ and ± 1.

While it is computationally easy to multiply numbers together, factoring a general integer is "generally believed to be hard" [4]. The security of RSA, the "most widely used public key cryptosystem in the world" [20], relies on this difficulty. Therefore, the study of integer factorization is of great practical importance.

The goal of this paper is to describe the *Elliptic Curve Method*(ECM), an integer factorization algorithm first proposed by Lenstra in [21]. Before describing this algorithm, we first discuss some preliminaries. In Chapter 2, we prove some elementary results from number theory and explicitly describe the RSA algorithm. In Chapter 3, we describe several factorization algorithms, including Pollard's Rho algorithm and the Quadratic Sieve. In Chapter 4, we discuss the algebraic properties of elliptic curves and the ECM. In Chapter 5, we discuss how performing the ECM with a curve in *Montgomery form* reduces computations. We conclude Chapter 5 by discussing some recent developments of the ECM.

Chapter 2

Preliminaries

2.1 Greatest common divisors and modular arithmetic

Definition 2.1. *Let a and b be positive integers. If $\gcd(a, b) = 1$ then a is* ***relatively prime*** *to b.*

Definition 2.2. *If $a - b$ is a multiple of m, then we write $a \equiv b \pmod{m}$.*

The following property is a basic property of division in the integers. For positive integers a and b with $b < a$ and $b \neq 0$, there is a non-negative integer r such that $a = bk + r$, where k is some positive integer and $r < b$. We will denote r as $a \mod b$.

We will need a way to compute greatest common divisors. The following algorithm, called the *Euclidean Algorithm*, finds $\gcd(x, y)$, where x and y are positive integers.

Suppose $y < x$. Express x as $x = m_1 y + r_1$, where m_1 and r_1 are integers with $0 \leq r_1 < y$. If $r_1 = 0$, then y divides x and $\gcd(x, y) = y$. If $r_1 \neq 0$, dividing y by r_1 yields $y = m_2 r_1 + r_2$ with $0 \leq r_2 < r_1$. If $r_2 \neq 0$, we divide r_1 by r_2 to obtain $r_1 = m_3 r_2 + r_3$ where $0 \leq r_3 < r_2$. Eventually some r_i will be 0, as each r_i is a non-negative integer strictly smaller than r_{i-1}. The process terminates when $r_{k+1} = 0$, for some k. Our last three equations are

$$\begin{aligned} r_{k-3} &= m_{k-1} r_{k-2} + r_{k-1} \\ r_{k-2} &= m_k r_{k-1} + r_k \\ r_{k-1} &= m_{k+1} r_k + 0 \end{aligned}$$

with $0 < r_k < r_{k-1}$. Since $r_{k-1} = m_{k+1}r_k$, r_k divides r_{k-1}. Since $r_{k-2} = m_k r_{k-1} + r_k$, r_k divides r_{k-2}. Since r_k divides r_{k-2} and r_{k-1}, r_k divides r_{k-3}. By working our way backwards through the equations in this manner we see that r_k divides both x and y. We will now show that $r_k = \gcd(x, y)$.

Suppose d is a divisor of both x and y. Since $x = m_1 y + r_1$, d must divide r_1. Since $y = m_2 r_1 + r_2$ and d divides y, if d divides r_1 then d also divides r_2. Continuing on, d must divide r_k, making d less than or equal to r_k. This proves that $r_k = \gcd(a, b)$.

Example 2.3. *We use the Euclidean Algorithm to calculate* $\gcd(45, 12)$.

$$\begin{aligned} 45 &= (3)(12) + 9 \\ 12 &= (1)(9) + 3 \\ 9 &= (3)(3) + 0 \end{aligned}$$

Thus, $\gcd(45, 12) = 3$.

We will use big O notation to describe the number of arithmetic in the Euclidean Algorithm. Later, we will use big O notation to describe other integer factorization algorithms.

Definition 2.4. *Let $N(x)$ be the number of arithmetic operations it takes to factor the integer x using some factorization algorithm. We will write*

$$N(x) = O(g(x))$$

if for some function $g(x)$ there exists a real number x_0 and a positive real number M such that

$$N(x) \leq M|g(x)| \text{ for all } x > x_0.$$

In [17](p.339), Knuth proves that the number of steps required to compute $\gcd(x, y)$ using the Euclidean Algorithm is $O(\ln N)$, where $N = \max(x, y)$. More precisely, Knuth proves that the number of steps is

$$\left\lceil \frac{\ln(\sqrt{5}N)}{\ln((1+\sqrt{5})/2)} \right\rceil - 2 \approx 2.078 \ln N + 1.672$$

However, each step in the Euclidean Algorithm requires a long division. In [18](p.13), Cohen states that this long division takes $O(\ln^2 N)$ operations.

In total, Cohen states that finding a greatest common divisor of a and b using the Euclidean Algorithm takes $O(\ln^2 N)$ operations, including the long division.

Extending the Euclidean Algorithm allows us to compute the multiplicative inverse of x modulo y, provided the inverse exists. In order for this inverse to exist, we require x and y to be relatively prime. Suppose we have integers a and b such that

$$ax + by = gcd(x, y).$$

If x and y are relatively prime, then

$$ax + by = 1,$$

and by considering both sides of this equation $\mod y$, we obtain

$$ax \equiv 1 \pmod{y}.$$

Thus, $a = x^{-1} \mod y$.

The *Extended Euclidean Algorithm* finds the integers a and b in the relation $ax + by = gcd(x, y)$, and hence can be used to calculate the multiplicative inverse of x modulo y, provided x and y are relatively prime. The Extended Euclidean Algorithm operates as follows. We first perform the Euclidean Algorithm to find the greatest common divisor of x and y, storing the values r_i, m_i, r_{i-1} at each step. By substituting r_i into r_{i-1} we can work in reverse to find a solution (a, b) to $ax + by = gcd(x, y)$.

Example 2.5. *We calculate a solution to the equation* $49a+15b = \gcd(49, 15)$. *In order to do this, we first calculate* $\gcd(49, 15)$ *using the Euclidean Algorithm*

$$\begin{aligned} 49 &= (3)(15) + 4 \\ 15 &= (3)(4) + 3 \\ 4 &= (1)(3) + 1 \\ 3 &= (3)(1) \end{aligned}$$

this shows that $\gcd(49, 15) = 1$. *We now work in reverse to calculate a solution to* $49a + 15b = 1$.

$$\begin{aligned} 1 &= 4 - (1)(3) \\ &= 4 - (1)\left(15 - (3)(4)\right) \\ &= (4)(4) - 15 \\ &= (4)(49 - (3)(15)) - 15 \\ &= (15)(-13) + (49)(4) \end{aligned}$$

Thus, $a = -13, b = 4$ *is a solution to our equation. Taking both sides of* $1 = (15)(-13) + (49)(4)$ *modulo 49 shows that* $-13 \equiv 26 \pmod{49}$ *is* $(15)^{-1} \mod 49$*.*

In [18](p.16), Cohen states that the Extended Euclidean Algorithm has the same run time as the Euclidean Algorithm.

2.2 Basic definitions and theorems

Before discussing integer factorization, more preliminaries need to be discussed.

Proposition 2.6. *If* x *and* m *are relatively prime and* $ax \equiv bx \pmod{m}$ *then* $a \equiv b \pmod{m}$*.*

Proof. If $ax \equiv bx \pmod{m}$ then m divides $ax - bx = (a-b)x$. We will now show that m divides $a - b$, which implies that $a \equiv b \pmod{m}$. Since x and m are relatively prime, we can use the Extended Euclidean Algorithm on x and m to find a c and an e such that

$$cx + em = 1.$$

Multiplying both sides of the above equation by $(a - b)$ yields

$$cx(a - b) + em(a - b) = (a - b).$$

Since m divides both $cx(a - b)$ and $em(a - b)$, m divides $(a - b)$. □

Definition 2.7. *The value of the* ***Euler Phi function****, denoted by* $\phi(n)$*, for* n *a positive integer, equals the number of positive integers less than or equal to* n *that are relatively prime to* n*.*

For any prime p, every positive integer less than p is relatively prime to p. Thus, $\phi(p) = p - 1$.

Theorem 2.8. *If n and b are positive, relatively prime integers then*

$$b^{\phi(n)} \equiv 1 \pmod{n}.$$

We will prove the above theorem momentarily.

Theorem 2.9. *(The Chinese Remainder Theorem) Suppose $m_1, m_2, \ldots, m_r$ are positive, pairwise relatively prime integers. Let $M = m_1 m_2 \ldots m_r$. Let $a_1, a_2, \ldots, a_r$ be arbitrary integers. Then there exists a unique integer a modulo M such that $a \equiv a_i \pmod{m_i}$ for every $i = 1, \ldots, r$.*

Proof. We prove the Chinese Remainder Theorem by constructing such an a and showing that a is unique modulo M. Let i and j be positive integers less than or equal to r and let $Q_i := \frac{M}{m_i}$. Observe that Q_i is relatively prime to m_i.

Define $M_i := Q_i^{\phi(m_i)}$ and $a := \sum_{i=1}^{r} a_i M_i$. By Theorem 2.7, $M_i \equiv 1 \pmod{m_i}$. If $j \neq i$ then m_i divides M_j and $M_j \equiv 0 \pmod{m_i}$. Thus, $a \equiv a_i \pmod{m_i}$ for every $i = 1, \ldots, r$. It remains to show that a is unique modulo M.

Suppose b satisfies the same r congruences as a. Then for each m_i, $a \equiv b \pmod{m_i}$. Thus, every m_i divides $a - b$, making M divide $a - b$ and $a \equiv b \pmod{M}$. Therefore, a is unique modulo M. □

Theorem 2.10. *If m and n are relatively prime, then*

$$\phi(mn) = \phi(m)\phi(n). \tag{2.1}$$

Proof. Let a be a positive integer less than and relatively prime to mn, that is, an integer counted by $\phi(mn)$. Since a is relatively prime to mn, a must be relatively prime to each of m and n as well. It follows that $a \bmod m$ and $a \bmod n$ are relatively prime to m and n respectively. From the uniqueness statement in the Chinese Remainder Theorem, distinct $a's$ that are less than and relatively prime to mn will produce distinct pairs of integers, one counted by $\phi(n)$ and the other counted by $\phi(m)$. Therefore, $\phi(mn) \leq \phi(m)\phi(n)$.

Conversely, suppose b is an integer counted by $\phi(m)$ and c is an integer counted by $\phi(n)$. Since m and n are relatively prime, by the Chinese Remainder Theorem there is a unique integer a with $1 \leq a < mn$ such that $a \equiv b \pmod{m}$ and $a \equiv c \pmod{n}$. Thus, the number of pairs (b, c) is at most $\phi(mn)$. The inequality $\phi(m)\phi(n) \leq \phi(mn)$ follows.

□

The Prime Number Theorem, a substantial result regarding the distribution of prime numbers among the positive integers, is stated below. A proof of the Prime Number Theorem can be found in Chapter 6 of [15].

Definition 2.11. *The **prime counting function**, denoted by $\pi(x)$, equals the number of prime numbers less than or equal to x, for any positive real x.*

Theorem 2.12. *(The Prime Number Theorem) The prime counting function, $\pi(x)$, has the same asymptotic behavior as $\frac{x}{\ln(x)}$. That is,*

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\ln(x)} = 1.$$

Finally, some elementary results about group theory will be used. A proof of Lagrange's Theorem can be found in [12](p.89).

Definition 2.13. *Let G be a finite group. The **order** of G is the number of elements in G. Let $a \in G$. The **order of** a **in** G is the smallest positive integer k such that $a^k = e$, where e is the identity element of G.*

Theorem 2.14. *(Lagrange's Theorem) If G is a group of finite order, then the order of every subgroup of G divides the order of G. As an immediate consequence of this, the order of every element of G divides the order of G.*

Theorem 2.7 is a consequence of Lagrange's Theorem, which we will now show. Suppose N is a positive integer. The set of all non-negative integers less than and relatively prime to N form a group under multiplication modulo N. This group will be denoted as $(\mathbb{Z}/N\mathbb{Z})^\times$. By definition, the order of $(\mathbb{Z}/N\mathbb{Z})^\times$ is $\phi(N)$. If a is an element of $(\mathbb{Z}/N\mathbb{Z})^\times$ with order k then by Lagrange's Theorem, k divides $\phi(N)$. Let B be a positive integer such that $\phi(N) = Bk$. It follows that

$$a^{\phi(N)} = a^{Bk} = (a^k)^B \equiv 1^B \equiv 1 \pmod{N},$$

proving Theorem 2.7.

2.3 RSA Cryptosystem

The RSA cryptosystem is a public key cryptosystem first proposed by Rivest, Shamir and Adleman in [28]. Let p and q be two distinct primes and let $pq = N$. Let d and e be two integers, representing decryption and encryption, respectively, such that $de \equiv 1 \pmod{\phi(N)}$. In order for such a d to exist, we require that e is relatively prime to $\phi(N)$. The values N and e are made public, while p,q, and d are kept private.

Suppose M is a message that is to be sent. We consider each letter of the alphabet as a distinct positive integer and convert the letters in M to integers. By concatenating these integers we can consider M as a positive integer. If M is larger than N then we break M into smaller "blocks" of integers where the number of digits in each block is less than N. We then transmit each block separately. We assume that M is less than N and require that M and N are relatively prime. If M is less than p and q then M is automatically relatively prime to N. If M is larger than N, the probability that M is divisible by p or q is "negligible" [7] (p.44). To encode M, we compute and send $E = M^e \mod N$.

To decode E, we use equation (2.8) and the fact that $de = k\phi(N) + 1$ for some positive integer k. We then compute

$$E^d \equiv (M^e)^d \equiv M^{ed} \equiv M^{k\phi(N)+1} \equiv M \pmod{N}. \tag{2.2}$$

Since M and $E^d \mod N$ are positive and less than N, they must be equal.

Using (2.1), if we know the factorization of N, that is, the values of p and q, then we know

$$\phi(N) = \phi(p)\phi(q) = (p-1)(q-1)$$

as p and q are relatively prime. We can then compute d using the relation $de \equiv 1 \pmod{\phi(N)}$ and the Extended Euclidean Algorithm with the integers e and $\phi(N)$. If we know d, we can use (2.2) to recover the message M. Therefore, if the factorization of N is known then M can be recovered.

In order to keep M from being recovered, we need to choose values of p and q that are computationally difficult to find, given their product N. RSA Laboratories, the security company formed by the inventors of the RSA algorithm, suggest choosing primes p and q of roughly equal length. Here,

length refers to the number of bits required to express p and q. Thus, p and q should be roughly half the length of N. As for the size of N, they recommend using an N of length "1024 bits for corporate use and 2048 bits for extremely valuable keys" [19].

However, history has shown that RSA is not invulnerable to attacks. In 1977, the authors of RSA published [13], which contained a message that was encrypted using a 129 digit number for N. They claimed that it would take "40 quadrillion years" to factor this N. It took less than 20 years. As described in [10], in 1994 Lenstra used an algorithm called the *Quadratic Sieve*, performed on 1,600 computers over the course of eight months to factor this integer, displaying the message "The Magic Words Are Squeamish Ossifrage." This effort won Lenstra a $100 prize, and made mathematicians around the world wonder what an Ossifrage is. (An Ossifrage is a species of vulture that are not particularly squeamish.)

Clearly the authors of RSA underestimated the development of both the theory of integer factorization and the increasing computational power of computers. As Lenstra has shown, an N that makes RSA computationally secure now may not be a secure choice in the future. Thus, despite the fact that RSA is secure against any integer factorization algorithm in practice, the understanding of these algorithms is crucial to the security of RSA.

Chapter 3

Factorization Algorithms

3.1 The Sieve of Eratosthenes

Before discussing our first factorization algorithm, we discuss one way to find prime numbers. We use the Sieve of Eratosthenes, an algorithm that finds all the prime numbers up to some bound, N. It is attributed to Eratosthenes, an ancient Greek mathematician.[7](p.19) The algorithm operates as follows:

1. Write the integers from 2 to N: $\{2, 3, 4, 5, \ldots, N\}$
2. Let $p = 2$, the first prime number.
3. Starting at p, mark all multiples of p greater than p, up to N.
4. Go to the next unmarked number, let p equal this number.
5. If no more unmarked numbers exist greater than p, stop. Otherwise, repeat step 3.

We will never mark a prime number, as every number we mark is a multiple of a prime, and hence, composite. In step 4, this p is guaranteed to be prime, as if it were not, it would be divisible by some prime less than it, and would have been marked in a previous iteration. Therefore, once the algorithm has terminated, the list consists of all the prime numbers from 2 to N.

In step 3, it suffices to mark p^2 and every multiple of p larger than p^2. This is because any multiple of p less than p^2 is a composite integer with a factor less than p and thus, has already been marked. In step 5, we can stop the algorithm when p^2 is larger than N.

Arithmetically, step 3 is accomplished by computing $2p, 3p, 4p \ldots$ for each

prime p and stopping when a number larger than N is computed. These multiplications are the only arithmetic operation in the sieve. For each prime $p < N$, this corresponds to $\left\lfloor \frac{N}{p} \right\rfloor$ additions. From Theorem 427 in [14] we have

$$\sum_{p \leq N,\ \text{p prime}} \frac{1}{p} = \ln \ln N + C + O\left(\frac{1}{\ln N}\right)$$

for some constant C. It follows that

$$\sum_{p \leq N,\ \text{p prime}} \left\lfloor \frac{N}{p} \right\rfloor \leq \sum_{p \leq N,\ \text{p prime}} \frac{N}{p} = N \ln \ln N + O(N) = O(N \ln \ln N)$$

Thus, the the Sieve of Eratosthenes will take $O(N \ln \ln N)$ additions to compute the primes less than N.

3.2 Trial Division

The most straightforward factorization algorithm is Trial Division. Suppose we wish to find a factor of N by Trial Division. We use the Sieve of Eratothanes to generate a list of prime numbers. Upon determining a number to be prime we check if this prime divides N. If it does, N is composite, and the algorithm stops. Note that if prime factor p of N is larger than $\left\lfloor \sqrt{N} \right\rfloor$ then $\frac{N}{p}$ must be less than $\left\lfloor \sqrt{N} \right\rfloor$. Thus, if Trial Division reaches a prime p with $p^2 > N$ then N is prime and the algorithm terminates.

In practice, Trial Division is performed up to some bound before any other factorization algorithm is performed, with the primes up to this bound precomputed. If we wish to prove the primality of N, Trial Division requires at most $\pi(\sqrt{N})$ divisions. By the Prime Number Theorem, this is approximately $\frac{2\sqrt{N}}{\ln(N)}$ divisions, provided the primes have been precomputed. Assuming N to be odd, if we wish to avoid generating a list of primes, we can simply divide N by every odd number less than $\sqrt{N}$, resulting in $\frac{\sqrt{N}}{2}$ divisions. As stated in [9], for a workstation in 2005, "numbers that can be proved prime via trial division in one minute do not exceed 13 decimal digits. In one day of current workstation time, perhaps a 19-digit number can be resolved."(p. 119)

Trial Division is ineffective for factoring composite numbers with only large factors. However if N has a small factor, Trial Division can be quite successful. In fact, for most numbers it is quite effective, as 88% of all positive integers have a factor less than 100 and almost 92% have a factor less than 1000. A proof is given below.

Proposition 3.1. *Approximately* 88% *of all positive integers have a factor less than* 100 *and appoximately* 92% *of all positive integers have a factor less than* 1000*.*

Proof. The probability that a natural number is not even is $(1 - \frac{1}{2})$. The probability a natural number is not divisible by 3 is $(1 - \frac{1}{3})$. For a prime p the probability that a natural number is not divisible by p is $(1 - \frac{1}{p})$. These events are independent, so the probability that a natural number is not divisible by all the primes less than P is

$$\prod_{p \leq P, \text{ p prime}} (1 - \frac{1}{p})$$

Thus, the probability that a number is divisible by a prime less than P is

$$1 - \prod_{p \leq P, \text{ p prime}} (1 - \frac{1}{p})$$

For $P = 100$, this is approximately .88 and for $P = 1000$, this is approximately .92. □

3.3 Fermat's Little Theorem

We will see that many factorization algorithms use Fermat's Little Theorem.

Theorem 3.2. *(Fermat's Little Theorem) If p is prime and p does not divide b then*

$$b^{p-1} \equiv 1 \pmod{p}. \tag{3.1}$$

Proof. For prime p, $\phi(p) = p-1$. Since b is relatively prime to p, the theorem follows from an application of Theorem 2.7. □

The following example shows that the converse of Fermat's Little Theorem is not true.

Example 3.3. *Let $P = 341 = (11)(31)$. Note that $2^{341-1} = 2^{340} \equiv 1 \pmod{341}$. P is an example of a number that satisfies Fermat's Little Theorem, yet is not prime.*

Definition 3.4. *A composite number N is a* ***pseudoprime to base b*** *if $b^{N-1} \equiv 1 \pmod{N}$, with b relatively prime to N. If $b = 2$, we refer to N simply as a* ***pseudoprime****.*

In [7](p. 32), it is stated that there are only 245 pseudoprimes below a million.

3.4 Pseudoprime Test

We can sometimes use Fermat's Little Theorem to show that a number is composite. By the contrapositive of Fermat's Little Theorem, if b is a positive integer less than N and $b^{N-1} \not\equiv 1 \pmod{N}$, then N cannot be prime. The pseudoprime test operates by checking if $b^{N-1} \not\equiv 1 \pmod{N}$ for a $b < N$ with b relatively prime to N.

A pseudoprime test cannot prove that a number is prime, as there are composite numbers that satisfy Fermat's Little Theorem for every base b, with $\gcd(b, N) = 1$. These numbers are called Carmichael numbers.

Definition 3.5. *A composite number N is a* ***Carmichael number*** *if $b^{N-1} \equiv 1 \pmod{N}$ for every integer b relatively prime to N.*

Carmichael numbers must be odd, as if N is an even integer greater than 2 then

$$(N-1)^{N-1} \equiv (-1)^{N-1} \equiv -1 \pmod{N}.$$

Korselt gives a criteria for recognizing Carmichael numbers based on their prime factorization. Oddly enough, this was proven by Korselt in 1899, 11 years before Carmichael produced the first example of a Carmichael number. Pomerance, in [9] (p.134) hypothesizes that Korselt thought Carmichael numbers did not exist and hoped that his criterion would serve as a first step towards proving this. In order to prove Korselt's Criterion for recognizing Carmichael numbers, we first prove a result from number theory.

Definition 3.6. *A number is* ***square-free*** *if it not divisible by the square of an integer except 1.*

Theorem 3.7. *(Korselt's Criterion) A positive composite integer N is a Carmichael Number if and only if N is square-free and if a prime p divides N then $p - 1$ divides $N - 1$.*

Proof. Suppose N is a Carmichael Number and let p be an odd prime that divides N. Let $N = p^e q$ where q is a positive integer relatively prime to p. The set of all non-negative integers less than p^e that are relatively prime to p^e form a group under multiplication modulo p^e. This group is denoted $(\mathbb{Z}/p^e\mathbb{Z})^\times$, and it is cyclic. Let b be a generator of $(\mathbb{Z}/p^e\mathbb{Z})^\times$. Since q is relatively prime to p, q is relatively prime to p^e. By the Chinese Remainder Theorem, there exists an a such that

$$a \equiv b \pmod{p^e} \text{ and } a \equiv 1 \pmod{q}.$$

Since p^e and q are relatively prime, and a is relatively prime to p^e and q, a is relatively prime to N. Since N is Carmichael we have
$a^{N-1} \equiv 1 \pmod{N}$, $a^{N-1} \equiv 1 \pmod{p^e}$ and

$$b^{N-1} \equiv 1 \pmod{p^e}.$$

By definition, the order of $(\mathbb{Z}/p^e\mathbb{Z})^\times$ is $\phi(p^e)$. The only positive integers less than or equal to p^e that are not relatively prime to p^e are multiples of p. These multiples are $p, 2p, 3p, \ldots p^{e-1}p$. There are p^{e-1} of these multiples, so $\phi(p^e) = p^e - p^{e-1} = p^{e-1}(p-1)$. Since b is a generator of $(\mathbb{Z}/p^e\mathbb{Z})^\times$, the order of b is $p^{e-1}(p-1)$.

It follows that $p^{e-1}(p-1)$ divides $N-1$. If $e > 1$, then p divides $N-1$. But p divides N, forcing $e = 1$. Thus, $p-1$ divides $N-1$ and $N = pq$. Since this holds for an arbitrary prime p dividing N, N must be square-free.

Now suppose N is square-free, forcing N to be of the form $N = p_1 p_2 \ldots p_m$, where each p_i is a distinct prime divisor of N. If b is relatively prime to N then b is relatively prime to every p_i. Furthermore, suppose that $p_i - 1$ divides N for every i. By Fermat's Little Theorem, $b^{p_i-1} \equiv 1 \pmod{p_i}$. Since $p_i - 1$ divides $N-1$, by raising both sides of the congruence $b^{p_i-1} \equiv 1 \pmod{p_i}$ to the $\frac{N-1}{p_i-1}$-th power, we obtain $b^{N-1} \equiv 1 \pmod{p_i}$. Thus, for every prime divisor p_i of N, $b^{N-1} - 1 = k_i p_i$, where k_i is some positive integer and $i = 1, \ldots, m$.

We have that

$$(b^{N-1} - 1)^m = p_1 k_1 p_2 k_2 \ldots p_m k_m = N k_1 k_2 \ldots k_m.$$

which shows that $(b^{N-1} - 1)^m \equiv 0 \pmod{N}$. By the Chinese Remainder Theorem, $b^{N-1} - 1$ is the unique integer modulo N that satisfies

$b^{N-1} \equiv 1 \pmod{p_i}$ every i. Since $b^{N-1} - 1 \equiv 0 \pmod{N}$ satisfies $(b^{N-1} - 1)^m \equiv 0 \pmod{N}$, it follows that $b^{N-1} - 1 \equiv 0 \pmod{N}$. This shows that N is a Carmichael number.

□

The smallest example of a Carmichael number is $561 = (3)(11)(17)$. This can be seen using Korselt's Criterion. Other examples of Carmichael numbers are $1105 = (5)(13)(17), 1729 = (7)(13)(19)$ and $2465 = (5)(17)(29)$.

In [32], Alford, Granville and Pomerance prove that there is an infinite number of Carmichael numbers. This is unfortunate, as there is an infinite set of numbers which Fermat's Little Theorem cannot prove are composite. Additionally, in [32], it is proven that for $x > N_0$, where N_0 is some positive integer, there are at least $x^{\frac{2}{7}}$ Carmichael numbers less than or equal to x. The number N_0 has not been explicitly calculated. However, in [9](p. 134), Crandall and Pomerance hypothesize N_0 to be the 96th Carmichael number, 8719309.

3.5 Strong Pseudoprime Test

After proving the following theorem, we will describe the Strong Pseudoprime Test.

Theorem 3.8. *Suppose N is an odd prime. Write $N - 1 = 2^s t$, where t is odd. If b is not divisible by N then either $b^t \equiv 1 \pmod{N}$ or $b^{2^i t} \equiv -1 \pmod{N}$ for some i with $0 \leq i \leq s - 1$.*

Proof. Observe that

$$\begin{aligned} b^{N-1} - 1 &= b^{2^s t} - 1 \\ &= (b^{2^{s-1} t} - 1)(b^{2^{s-1} t} + 1) \\ &\vdots \\ &= (b^{2^{s-2} t} - 1)(b^{2^{s-2} t} + 1)(b^{2^{s-1} t} + 1) \\ &= (b^t - 1)(b^t + 1)(b^{2t} + 1)(b^{4t} + 1) \dots (b^{(2^{s-1})(t)} + 1) \end{aligned}$$

By Fermat's Little Theorem, $b^{N-1} \equiv 1 \pmod{N}$. Thus, N divides at least one factor on the right side of the equation above. This completes the proof. □

We can use the contrapositive of the above theorem to check if a number is composite. Suppose N is an odd number, where we write N as $N = 1+2^s t$ with t odd. We choose some b relatively prime to N. If $b^t \not\equiv 1 \pmod{N}$ and $b^{2^i t} \not\equiv -1 \pmod{N}$ for all i with $0 \leq i \leq s-1$ then N is a composite number. This is known as the *Strong Pseudoprime Test.*

Just as in the Pseudoprime Test, the Strong Pseudoprime Test cannot prove primality. This is due to the existence of composite numbers that satisfy the conditions of the Strong Pseudoprime Test.

Definition 3.9. *Let N be a composite odd integer. Write $N-1 = 2^s t$ where t is odd. Let b be a positive integer prime to N. Then N is a* ***strong pseudoprime to the base*** *b if $b^t \equiv 1 \pmod{N}$ or $b^{2^i t} \equiv -1 \pmod{N}$ for some i with $0 \leq i \leq s-1$.*

There are only 13 strong pseudoprimes to the bases 2,3 and 5 less than 25×10^9. Furthermore, if we consider all of the bases 2,3,5 and 7, there is only one strong pseudoprime less than 25×10^9 [7](p. 32). This strong pseudoprime is $S_0 = 3215031751$. Thus, for $N < 25 \times 10^9$ with $N \neq S_0$, if a strong pseudoprime test to the bases 2,3,5 and 7 fails to show that N is composite then N is prime.

Before running a factorization algorithm, it is wise to verify, using a strong pseudoprime test, that the number we are attempting to factor is in fact composite. In our further discussions of factorization algorithms, we assume that N is composite.

3.6 A method of Fermat

Fermat observed that if $N = u^2 - v^2$ where u and v are positive integers, then $N = (u+v)(u-v)$. Thus, if we can write N as the difference of two squares, we can find a factorization of N. If $N = xy$ is odd and composite, then

$$N = xy = \left(\frac{x+y}{2}\right)^2 - \left(\frac{x-y}{2}\right)^2$$

for x and y positive integers greater than one. Since N is odd, both $x+y$ and $x-y$ are even, implying that every composite odd number can be written as a difference of two square integers. Fermat's method finds these two squares, and hence, a factorization of N.

Suppose N is an odd and composite positive integer. We wish to find a u and a v such that $N = u^2 - v^2$. Let $z = \lceil \sqrt{N} \rceil$, the smallest possible value of u. If $z^2 - N$ is a square, then we have found a v that works, and can find a factorization of N. We continue this method by checking if $(z+1)^2 - N, (z+2)^2 - N, \ldots$ are squares.

Example 3.10. *Suppose we wish to factor $N = 426749$. We have that $z = \lceil \sqrt{426749} \rceil = 654$. Observe that*

$$\begin{aligned} z^2 - N &= (654)^2 - N = 967 \\ (z+1)^2 - N &= (655)^2 - N = 2276 \\ (z+2)^2 - N &= (656)^2 - N = 3587 \\ (z+3)^2 - N &= (657)^2 - N = 4900 = 70^2 \end{aligned}$$

Thus, $N = (657)^2 - (70)^2 = (657+70)(657-70) = (727)(587)$.

Fermat's method finds the divisor of N that is closest to $\sqrt{N}$. If Fermat's method fails to find a divisor of N then N is prime. Proving primality in this way is very slow. However, if N has a factor close to $\sqrt{N}$, Fermat's method operates quickly. This is made precise below.

Proposition 3.11. *It takes $O(N)$ trial values of z for Fermat's method to prove primality. However, if N has a divisor within $(4N)^{\frac{1}{4}}$ of $\sqrt{N}$ then Fermat's method will be successful in one trial.*

Proof. Suppose $N = xy$, with x the smallest divisor of N greater than or equal to $\lceil \sqrt{N} \rceil$. Fermat's method terminates when $(\lceil \sqrt{N} \rceil + c)^2 = \left(\frac{x+y}{2}\right)^2$ for some positive integer c. This c represents the number of steps in Fermat's method. Observe that

$$\begin{aligned} c &= \frac{x+y}{2} - \lceil \sqrt{N} \rceil \\ &\leq \frac{x+y}{2} - \sqrt{N} \\ &= \frac{xy + y^2 - 2y\sqrt{N}}{2y} \end{aligned}$$

$$= \frac{(\sqrt{N} - y)^2}{2y} \tag{3.2}$$

If N is prime, then $y = 1$. Thus, it will take Fermat's method $O(N)$ steps to prove the primality of N.

From (3.2), Fermat's method will be successful after one step provided that y, the largest divisor of N less than or equal to $\sqrt{N}$ obeys $(\sqrt{N} - y)^2 \leq 2y$. Since $2y \leq 2\sqrt{N}$, we require that $(\sqrt{N} - y)^2 \leq 2\sqrt{N}$, or that $\sqrt{N} - y \leq (4N)^{\frac{1}{4}}$. Thus, Fermat's method is successful after one iteration provided that N has a factor within $(4N)^{\frac{1}{4}}$ of $\sqrt{N}$. □

Furthermore, by using congruences and quadratic residues we will not need to check if every $z^2 - N$ is a square.

Definition 3.12. *An integer q is a **quadratic residue modulo M** if it is congruent to a square modulo M. That is, if there exists an integer x such that $x^2 \equiv q \pmod{M}$.*

Fermat's method is successful when $z^2 - N$ is a square. If $z^2 - N$ is a square then $z^2 - N$ is a quadratic residue modulo M, for any positive integer M. This will reduce the number of z's that can be successful in Fermat's method. Two cases are described below.

Proposition 3.13. *Let N be a positive, odd and composite integer. Suppose we wish to find a z such that $z^2 - N$ is a square, as in Fermat's method. If $N \equiv 1 \pmod{4}$, then z must be odd.*

Proof. Suppose $N \equiv 1 \pmod{4}$. We require that $z^2 - N$ is a quadratic residue modulo 4. The quadratic residues modulo 4 are 0 and 1. If $z^2 - N \equiv 0 \pmod{4}$, then $u^2 \equiv 1 \pmod{4}$ and either $z \equiv 1 \pmod{4}$ or $z \equiv 3 \pmod{4}$. This forces z to be odd. If $z^2 - N \equiv 1 \pmod{4}$, then $z^2 \equiv 2 \pmod{4}$. Since 2 is not a quadratic residue modulo 4, this cannot occur.

□

The following proposition follows similarly.

Proposition 3.14. *Let N be a positive, odd and composite integer. Suppose we wish to find a z such that $z^2 - N$ is a square, as in Fermat's method. If $N \equiv 2 \pmod{3}$, then z is a multiple of 3.*

Thus, by examining N modulo 3 or 4 we can reduce the computations in Fermat's method. Fermat's method is only used if it is known that N

has a factor near $\sqrt{N}$. [7](p. 31) However, Fermat's method can be vastly improved by shifting our focus from expressing N as a difference of squares to finding two squares that are congruent modulo N. This notion is utilized by the Quadratic Sieve.

3.7 The Quadratic Sieve

Fermat's method finds an x and y that satisfy $x^2 - y^2 = N$. However, it suffices to find an x and y that satisfied $x^2 \equiv y^2 \pmod{N}$ with $x \not\equiv \pm y \pmod{N}$. The reason for this is if we find such a pair x and y then N will divide $x^2 - y^2 = (x-y)(x+y)$, and $\gcd(x-y, N)$ will produce a factor of N. This factor will not be N, as $x \not\equiv \pm y \pmod{N}$ forces $\gcd(x \pm y, N) \neq N$. In the following example, we show a method to find such a pair.

Example 3.15. *Suppose we wish to factor $N = 1649$. As in Fermat's method, we use $z = \lceil \sqrt{1649} \rceil = 41$. We calculate a few squares modulo N for integers greater than or equal to z.*

$$\begin{aligned} 41^2 &= 1681 \equiv 32 \pmod{1649} \\ 42^2 &= 1764 \equiv 115 \pmod{1649} \\ 43^2 &= 1849 \equiv 200 \pmod{1649} \end{aligned}$$

If we were to factor N using Fermat's method, we would need to continue to 57^2, and use the fact that $57^2 - N = 1600 = 40^2$. However, we can still factor N if we restrict ourselves to the three squares found above. Observe that $(32)(200) = 6400 = 80^2$ and

$$(41 \cdot 43)^2 \equiv 80^2 \pmod{1649}.$$

We have that $(41)(43) = 1764 \equiv 114 \pmod{1649}$ and $114 \not\equiv \pm 80 \pmod{1649}$. By calculating $\gcd(114 - 80, 1649) = 17$ we obtain a nontrivial factor of N.

In this example, we were able to find an x and y that satisfy $x^2 \equiv y^2 \pmod{N}$ with $x \not\equiv \pm y \pmod{N}$. We accomplished this by multiplying two quadratic residues modulo N. The Quadratic Sieve, invented by Pomerance in [26], attempts to find a set of ${x_i}^2 \mod N$ with $i = 1, \ldots, k$ whose product forms the relation

$$\prod_{i=1}^{k} {x_i}^2 \equiv y^2 \pmod{N} \text{ with } \prod_{i=1}^{k} x_i \not\equiv \pm y \pmod{N}. \tag{3.3}$$

If some ${x_i}^2 \mod N$ has a prime factor raised to an odd power, and we wish to use this ${x_i}^2 \mod N$ in our product we must find and include another ${x_i}^2 \mod N$ with this prime factor to an odd power. If this prime is large, finding another quadratic residue with this prime factor raised to an odd power can be difficult. Thus, we only consider quadratic residues with small prime factors. That is, when our quadratic residue is B-smooth modulo N, for some bound B.

Definition 3.16. *An integer is* ***B-smooth*** *if none of its prime factors is greater than B.*

If a positive integer r is B-smooth, then $r = p_1^{a_1} p_2^{a_2} \dots p_{\pi(B)}^{a_{\pi(B)}}$, where $p_1, p_2, \dots p_{\pi(B)}$ are the primes up to B and each a_i is a non-negative integer. The factorization of r can be denoted by the exponent vector

$$v(r) = (a_1, a_2, \dots, a_{\pi(B)}).$$

If $r_1, r_2, \dots r_k$ are all B-smooth then $\prod_{i=1}^{k} r_i$ is a square if and only if $\sum_{i=1}^{k} v(r_i)$ has all even coordinates.

We represent each B-smooth ${x_i}^2 \mod N$ as the exponent vector $v(r_i)$. We wish to find a set of quadratic residues whose product forms a square. That is, a set of $v(r_i)$'s that sum to the zero vector modulo 2. It suffices to find a linear dependency among the vectors $v(r_i)$ with entries taken modulo 2.

We consider each exponent vector modulo 2 as an element of the vector space $\mathbb{Z}_2^{\phi(B)}$ defined over $\mathbb{Z}_2$. There is a theorem from linear algebra that if the number of elements in a set of vectors is larger than the dimension of the space then the set is linearly dependent. Thus, if we can find more than $\pi(B)$ distinct $v(r_i)$'s that are B-smooth then we will be able to establish a linear dependency among this set. We can then use Gaussian elimination in the field $\mathbb{Z}/2\mathbb{Z}$ on a matrix of $v(r_i)$'s to find a linear dependency. With this linear dependency established, we will have a set of $x_i \mod N$'s of the form (3.3) and we will be able to find a factor of N.

The success of this algorithm depends on our ability to produce more than $\pi(B)$ distinct $v(r_i)$'s that are B-smooth. In order for this to occur, we need to find at least $\pi(B)+1$ distinct ${x_i}^2 \mod N$'s that are B-smooth. If we choose B to be small, we will not need as many ${x_i}^2 \mod N$'s to form a product that is a square. However, if B is too small, then we may not find enough

of these residues. An optimal choice of B must be made. Pomerance, in [27] conjectures that an optimal value for B is about $B = \exp\left(\frac{1}{2}\sqrt{\ln N \ln\ln N}\right)$. Additionally, he states that the Quadratic Sieve has a running time of about B^2. That is, a running time of about $\exp\left(\sqrt{\ln N \ln\ln N}\right)$.

3.8 Pollard Rho Factorization Method

Pollard's Rho Algorithm, first described in [25], utilizes an idea similar to the "birthday paradox". The birthday paradox is motivated by the following question: given a random set of people, what is the probability that two people have the same birthday? With 366 people the probability is 100%, while a 50% probability is reached with only 23 people.

Proposition 3.17. *Suppose we have a set of p distinct numbers. Form a sequence by choosing numbers at random from this set. Since the set is finite, our sequence must have a repeated element. When the sequence consists of more than $1.177\sqrt{p}$ numbers, the probability of there being a repeated element in this sequence exceeds 50%.*

A proof can be found in [4]. Based on the birthday paradox, if we have a finite set C, a $c \in C$ and a function f that takes an element from C and maps it to a random element in C, then the sequence

$$c, f(c), f(f(c)), \ldots$$

will have a 50% chance of having a repeated element after $1.177\sqrt{|C|}$ elements. When we encounter a repeated element our sequence becomes cyclic and resembles a circle with a tail, or a rho (ρ).

Suppose our set is the finite group $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ and $f(x) = x^2 + a \mod p$ for some fixed integer a. The hope is that f behaves in a random fashion. If so, then by the birthday paradox, the sequence mentioned above will have a repeated element after approximately $1.177\sqrt{p}$ elements. In [6], Brent states that f will behave in a random fashion provided $a \neq 0, -2$, based on empirical results. The Pollard Rho algorithm operates as follows.

Suppose p is a prime divisor of N. Let $f(x) = x^2 + a$ with $a \neq 0, -2$. We form the sequence

$$x_0, f(x_0) \mod N, f(f(x_0)) \mod N, \ldots$$

using some non-negative integer $x_0 < N$. This sequence corresponds to a sequence modulo p. The sequence modulo p eventually will have a repeated value as it is formed from values in a finite set. Thus, there exists distinct, positive integers i, j such that $f^{(i)}(x_0) \equiv f^{(j)}(x_0) \pmod{p}$ and $\gcd(f^{(i)}(x_0) - f^{(j)}(x_0) \mod N, N)$ is a multiple of p. If this multiple of p is not N then we have found a non-trivial factor of N.

We do not wish to calculate $\gcd(f^{(i)}(x_0) - f^{(j)}(x_0) \mod N, N)$ for every i and j. Instead, we use an algorithm called *Floyd's cycle finding algorithm* [29] to ease this computation. We compute two sequences,

$$x_0, f(x_0) \mod N, f(f(x_0)) \mod N, \ldots f^{(i)}(x_0) \mod N, \ldots$$

and

$$x_0, f(f(x_0)) \mod N, f^{(4)}(x_0) \mod N, \ldots, f^{(2i)}(x_0) \mod N \ldots$$

We then take the product of many $f^{(2i)}(x_0) - f^{(i)}(x_0) \mod N$ and then take the gcd of that product with N. If this gcd produces N then we can either take a gcd of a product with fewer
$f^{(2i)}(x_0) - f^{(i)}(x_0) \mod N$'s, or start over with a new $f(x)$.

By the birthday paradox, we expect to find a factor for $i > 1.177\sqrt{p}$, or simply after $O(\sqrt{p})$ steps. However, i "can be as large as the smallest prime divisor" [7](p. 63), as our sequence modulo p may take p terms to cycle. Because of this, the Pollard Rho algorithm, at its worst case, is comparable to Trial Division. Despite this shortcoming, the Pollard Rho Algorithm was used to factor the eighth Fermat number.

Definition 3.18. *The **n-th Fermat number** is* $F_n = 2^{2^n} + 1$.

In [5],R.P. Brent and Pollard used the Pollard Rho Algorithm to find the complete factorization of the eighth Fermat number $F_8 = 2^{2^8} + 1$, finding a prime factor that is 62 digits long. Furthermore, in [4], a 19 digit factor of $2^{2386} + 1$ was found using the Pollard Rho Algorithm.

Chapter 4

Elliptic curves

In our discussion of elliptic curves, we primarily refer to Chapters 2 and 4 in [31] to provide background information.

Definition 4.1. *An **elliptic curve over the field K** is a curve of the form $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, where each a_i is an element of the field K.*

Definition 4.2. *An elliptic curve over K of the form*

$$y^2 = x^3 + ax + b \tag{4.1}$$

*where a and b are elements of K that satisfy $4a^3 + 27b^2 \neq 0$, is said to be in **Weierstrass form**.*

Theorem 4.3. *If K is a field that is not of characteristic 2 or 3 then any elliptic curve defined over K can be written in Weierstrass form.*

Proof. We suppose that K is a field of characteristic 2 or 3. This allows us to divide by powers of 2 and powers of 3. Starting with an arbitrary elliptic

curve and then completing the square yields

$$\begin{aligned}
y^2 + a_1xy + a_3y &= x^3 + a_2x^2 + a_4x + a_6 \\
y^2 + a_1xy + a_3y + \frac{a_3a_1x}{2} + \frac{{a_1}^2x^2}{4} + \frac{{a_3}^2}{4} &= x^3 + \left(a_2 + \frac{{a_1}^2}{4}\right)x^2 \\
&\quad + \left(a_4 + \frac{a_1a_3}{2}\right)x + \left(\frac{{a_3}^2}{4} + a_6\right) \\
\left(y + \frac{a_1x}{2} + \frac{a_3}{2}\right)^2 &= x^3 + \left(a_2 + \frac{{a_1}^2}{4}\right)x^2 \\
&\quad + \left(a_4 + \frac{a_1a_3}{2}\right)x + \left(\frac{{a_3}^2}{4} + a_6\right) \\
{y_1}^2 &= x^3 + a_2'x^2 + a_4'x + a_6'
\end{aligned}$$

where $y_1 = y + \frac{a_1x}{2} + \frac{a_3}{2}$, $a_2' = a_2 + \frac{{a_1}^2}{4}$, $a_4' = a_4 + \frac{a_1a_3}{2}$ and $a_6' = \frac{{a_3}^2}{4} + a_6$. Furthermore, letting $x = x_1 - \frac{a_2'}{3}$ yields

$$\begin{aligned}
{y_1}^2 &= \left(x_1 - \frac{a_2'}{3}\right)^3 + a_2'\left(x_1 - \frac{a_2'}{3}\right)^2 + a_4'\left(x_1 - \frac{a_2'}{3}\right) + a_6' \\
&= \left({x_1}^3 - a_2'{x_1}^2 + \frac{1}{3}a_2'x_1 - \frac{1}{27}{a_1}^3\right) + a_2'\left({x_1}^2 - \frac{2a_2'x_1}{3} + \left(\frac{a_2'}{3}\right)^2\right) \\
&\quad + a_4'\left(x_1 - \frac{a_2'}{3}\right) + a_6' \\
&= {x_1}^3 + x_1\left(-\frac{1}{3}a_2' + a_4'\right) + \left(\left(\frac{a_2'}{3}\right)^2 - \frac{1}{27}{a_1'}^3 - \frac{a_x'a_4'}{3} + a_6'\right) \\
&= {x_1}^3 + Ax_1 + B.
\end{aligned}$$

□

Henceforth, we assume that our field is not of characteristic 2 or 3 and we will consider elliptic curves in Weierstrass form.

Definition 4.4. *The **discriminant** of the cubic polynomial x^3+ax^2+bx+c is $a^2b^2 - 4b^3 - 4a^3c - 27c^2 + 18abc$.*

For a curve in Weierstrass form, when we mention the discriminant of the curve we are referring to the discriminant of the cubic $x^3 + ax + b$. The discriminant of this cubic is $-(4a^3 + 27b^2)$. From [12] (p. 527), if the discriminant of an elliptic curve over $\mathbb{Q}$ is non-zero then the curve has distinct

roots in $\mathbb{R}$. If a cubic has positive discriminant then the cubic has 3 distinct real roots and if the discriminant is negative then the cubic has 1 real root. Elliptic curves over $\mathbb{Q}$ can be graphed. Two examples are given in figure 4.1 and figure 4.2.

We regard a tangent line to the elliptic curve at a point as intersecting the curve twice at that point. From [30] (p. 45), if the discriminant of (4.1) is non-zero then the curve has a well defined tangent line at every point. It will be shown below that for an elliptic curve defined over $\mathbb{Q}$ with a non-zero discriminant, every non-vertical line that intersects two rational points on the curve intersects a third rational point on the curve. This provides us with a way to define a binary operation on the rational points of an elliptic curve over defined over $\mathbb{Q}$.

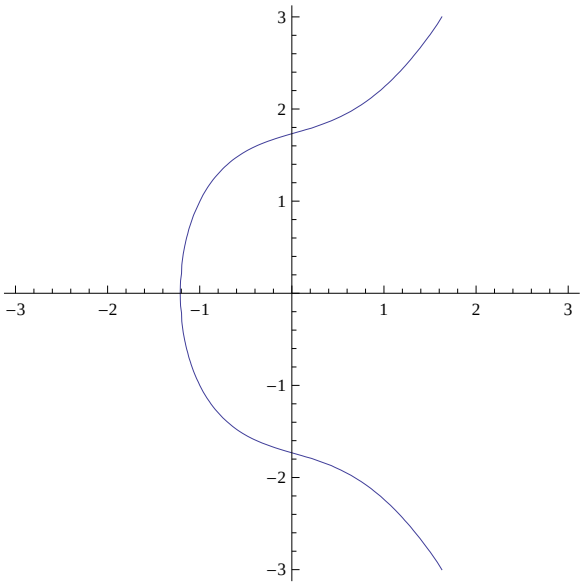

Figure 4.1: The elliptic curve $y^2 = x^3 + x + 3$.

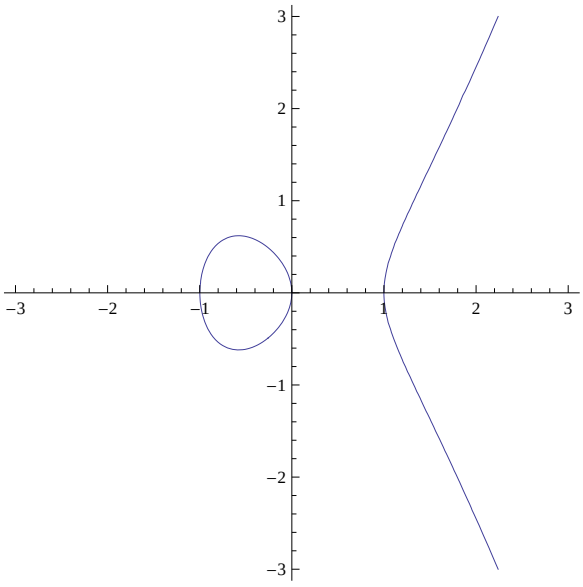

Figure 4.2: The elliptic curve $y^2 = x^3 - x$

4.1 Addition on an elliptic curve

The rational points on an elliptic curve over $\mathbb{Q}$ have an inherent group structure. We define the operation + below. We begin by defining the operation on curves over $\mathbb{Q}$ geometrically. We then examine the operation arithmetically. We will need to include the symbol ∞ in our group, which will serve as our identity element.

Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, with $x_1 \neq x_2$, $y_1 \neq y_2$ be rational points on an elliptic curve over $\mathbb{Q}$. It is proved below that the line between P_1 and P_2 intersects the curve at a third rational point. We will denote this point as (x, y). The sum of P_1 and P_2 is defined to be (x, y) reflected about the x-axis. That is, $P_1 + P_2 = (x, -y)$. This is illustrated in figure 4.3.

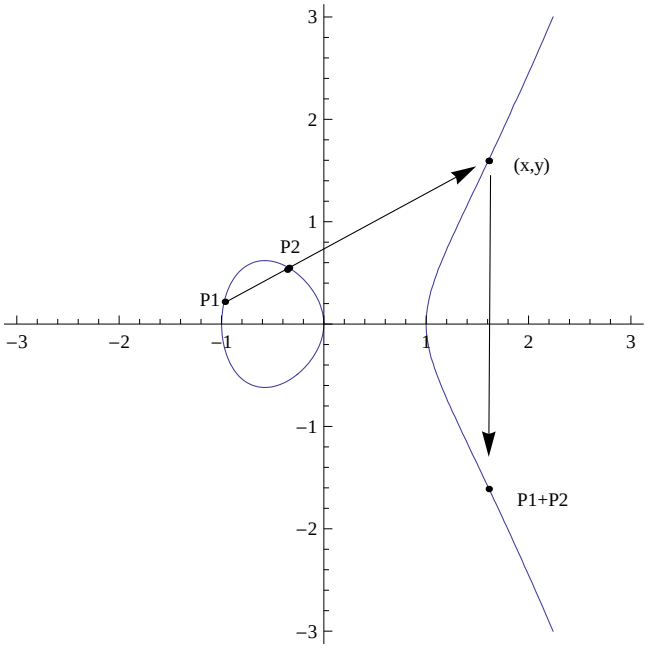

Figure 4.3: Elliptic curve addition on the curve $y^2 = x^3 - x$

We will now describe this operation arithmetically.

Theorem 4.5. *(Elliptic curve addition) For the rational points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ on the curve $y^2 = x^3 + ax + b$, where $4a^3 + 27b^2 \neq 0$, $x_1 \neq x_2$, and both a and b are rational, $+$ operates as follows:*

$$P_1 + P_2 = (x_3, y_3) = (m^2 - x_1 - x_2, m(x_1 - x_3) - y_1), \tag{4.2}$$

where

$$m = \frac{y_1 - y_2}{x_1 - x_2}.$$

Proof. The slope of the line through P_1 and P_2 is $m = \frac{y_2 - y_1}{x_2 - x_1}$. Since x_1, x_2, y_1 and y_2 are rational, m is also rational. The equation of the line intersecting P_1 and P_2 is $y = m(x - x_1) + y_1$. Substituting $y = m(x - x_1) + y_1$ into (4.1) yields:

$$(m(x - x_1) + y_1)^2 = x^3 + ax + b \tag{4.3}$$

$$m^2(x - x_1)^2 + 2my_1(x - x_1) + {y_1}^2 = x^3 + ax + b$$
$$m^2x^2 - 2xx_1m^2 + {x_1}^2m^2 + 2my_1x - 2my_1x_1 + {y_1}^2 = x^3 + ax + b$$

$$0 = x^3 - x^2m^2 + x(2x_1m^2 - 2my_1 + a) + (b - {x_1}^2m^2 + 2my_1x_1 - {y_1}^2) \tag{4.4}$$

Since $x = x_1, x_2$ satisfy (4.3), they satisfy (4.4). Since $x_1 \neq x_2$, we have identified two roots of the cubic (4.4). If (4.4) has two real roots it must have a third real root, which we will denote as x_3.

In general, note that if a monic cubic $f(x)$ has roots α, β and γ, then

$$f(x) = (x - \alpha)(x - \beta)(x - \gamma)$$

$$= x^3 - (\alpha + \beta + \gamma)x^2 + (\alpha\beta + \alpha\gamma - \beta\alpha)x - \alpha\beta\gamma \tag{4.5}$$

We will use (4.5) to write our third root of (4.4), x_3, in terms of the known roots x_1 and x_2. From (4.5), $m^2 = x_3 + x_2 + x_1$ and $x_3 = m^2 - x_2 - x_1$. Thus, the third point on both the line and the elliptic curve is

$$\left(m^2 - x_2 - x_1, m(x_3 - x_1) + y_1\right).$$

We reflect this point over the x-axis to find $P_1 + P_2$. Thus, $P_1 + P_2 = (x_3, y_3)$ where $x_3 = m^2 - x_1 - x_2$ and $y_3 = m(x_1 - x_3) - y_1$. Observe that since m, x_2, x_1 and y_1 are rational, both x_3 and y_3 are rational. □

Suppose now that $P_1 = (x_1, y_1)$ is a rational point with $y_1 \neq 0$. We will now define $P_1 + P_1$, which we will denote as $2P_1$, in the following theorem.

Theorem 4.6. *(Elliptic Curve Doubling) For the rational point $P_1 = (x_1, y_1)$ on the elliptic curve $y^2 = x^3 + ax + b$ with a and b rational and $y_1 \neq 0$,*

$$2P_1 = (x_3, y_3) = (m^2 - 2x_1, m(x_1 - x_3) - y_1), \tag{4.6}$$

where

$$m = \frac{3x_1^2 + a}{2y_1}.$$

Proof. Suppose $P_1 = P_2 = (x_1, y_1)$ with $y_1 \neq 0$. In order to derive the formula for $2P_1$, we use the line tangent to the elliptic curve at P_1. Our curve is assumed to have non-zero discriminant, and hence, P_1 has a well defined tangent line. We find the slope of this line by implicitly differentiating the elliptic curve equation. This results in $2y\frac{dy}{dx} = 3x^2 + a$. Solving for $\frac{dy}{dx}$ yields $m = \frac{dy}{dx} = \frac{3{x_1}^2+a}{2y_1}$. Since x_1, a and y_1 are rational, m is rational. Proceeding in the same manner as in the proof of Theorem 4.5 but using $x_1 = x_2$, $y_1 = y_2$ and $m = \frac{3{x_1}^2+a}{2y_1}$ yields

$$2P_1 = (m^2 - 2x_1, m(x_1 - x_3) - y_1).$$

Observe that since m, x_1 and y_1 are rational, $2P_1$ has rational coordinates. □

Equations (4.2) and (4.6) show that the sum of two rational points is a rational point. However, equation (4.2) does not consider addition of distinct points that lie on the same vertical line. Also, equation (4.6) does not account for doubling of a point with a vertical tangent line. As the graph of an elliptic curve is symmetric with respect to the x-axis, these vertical lines do not intersect the elliptic curve at a third point. We extend our notion of addition to include these points by introducing a symbol, ∞, and considering ∞ to be a point on every vertical line. We extend addition in this manner by defining the rules described below. Suppose $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ are rational points on a curve defined over $\mathbb{Q}$.

$$\begin{gathered}\text{Define } P_1 + \infty = P_1.\\ \text{Define } \infty + \infty = \infty.\\ \text{If } y_1 = 0, \text{ then } 2P_1 = \infty.\\ \text{If } x_1 = x_2 \text{ and } y_1 \neq y_2, \text{ then } P_1 + P_2 = \infty.\end{gathered} \tag{4.7}$$

The first two equations establish ∞ as the identity element of our addition. The last two equations show that if $P = (x, y)$ then $-P = (x, -y)$, establishing an inverse for our addition. We consider the combination of (4.2), (4.6) and (4.7) as the *elliptic curve group law.*

Using the group law as our binary operation, the set of all rational points on an elliptic curve with rational coefficients and the point ∞ form an abelian group. From the group law, it follows that our addition is commutative. It is difficult to show that the elliptic curve group law is associative. For a proof, see [31] (p. 20).

Theorem 4.7. *Let $E_{(a,b)}(\mathbb{Q}) = E \cup \{\infty\}$, where E is the set of all the rational points on the elliptic curve $y^2 = x^3 + ax + b$ with $4a^2 - 27b^2 \neq 0$, a and b rational and ∞ as defined above. Using the elliptic curve group law, $E_{(a,b)}(\mathbb{Q})$ is an abelian group with identity element ∞.*

Note that for some pair a, b, $E_{(a,b)}(\mathbb{Q})$ might just be $\{\infty\}$. the trivial group. That is, the equation $y^2 = x^3 + ax + b$ might have no rational solutions (x, y). One such example is the curve $y^2 = x^3 - 5$, which Cassels shows to have no rational solutions in [16]. When the coefficients a and b are arbitrary rational numbers we will refer to $E_{(a,b)}(\mathbb{Q})$ as $E(\mathbb{Q})$.

The Mordell-Weil Theorem provides additional structure for $E(\mathbb{Q})$, first proven by Louis Mordell in 1922. [24]

Theorem 4.8. *(Mordell-Weil) The group $E(\mathbb{Q})$ is a finitely generated abelian group.*

By the structure theorem for finitely generated abelian groups,

$$E(\mathbb{Q}) \cong T \oplus \mathbb{Z}^r$$

where T is a finite group and r is a non-negative integer. For details, see [12] (p.189). The finite group T is known as the *torsion subgroup* and r is called the *rank* of $E(\mathbb{Q})$. We now discuss curves defined over a finite field.

4.2 Reduction of elliptic curves defined modulo N

For a positive composite integer N, we can examine an elliptic curve defined modulo N, provided a few criteria are satisfied. Since a standing assumption is that we are working in a field of characteristic neither 2 nor 3, we require that N is relatively prime to 6. When we describe an elliptic curve modulo

N, we are referring to a curve of the form (4.1) with a and b non-negative integer coefficients that are less than N. Also, we require that $4a^3+27b^2 \neq 0 \mod N$. The points of this curve are of the form (x, y) where x and y are non-negative integers less than N that satisfy $y^2 = x^3 + ax + b$.

To produce a curve modulo N we first choose x_0, y_0 and a to be any positive integers less than N. This then defines $b = {y_0}^2 - ({x_0}^3 + ax_0) \mod N$. If $\gcd(4a^3 + 27b^2, N) = 1$, we can then consider the elliptic curve modulo N given by $y^2 = x^3 + ax + b$, with (x_0, y_0) a point on this curve.

The points on an elliptic curve defined modulo N do not form a group, as addition is not defined for every two points. We cannot compute elliptic curve addition or a doubling on the curve modulo N if $(x_1 - x_2)$ or $2y_1$ do not have multiplicative inverses modulo N, respectively. That is, we require both $(x_1 - x_2)$ and $2y_1$ to be relatively prime to N. If they are not, then addition modulo N will fail. Thus, if an addition on an elliptic curve fails, either $\gcd(x_1 - x_2, N)$ or $\gcd(2y_1, N)$ will produce a factor of N. This idea forms the foundation of Lenstra's Elliptic Curve Factorization Method (ECM).

For example, consider $N = 4453$ and the point $P = (1, 3)$ on the curve $y^2 = x^3 + 10x - 2 \mod 4453$. We will compute $3P$. First, we will compute $2P$. The slope of the tangent line at P is

$$\frac{3x^2 + 10}{2y} = \frac{13}{6} \equiv 3713 \pmod{4453}.$$

To find $6^{-1} \mod 4453$ we used the fact that 6 and 4453 are relatively prime. This yields $2P = (x, y)$ where

$$x = 3713^2 - 2 \equiv 4332, y = -3713(x - 1) - 3 \equiv 3230$$

To compute $3P$ we add $2P$ and P. The slope is

$$\frac{3230 - 3}{4332 - 1} = \frac{3227}{4331}$$

But 4331 and 4553 are not relatively prime, as $\gcd(4331, 4553) = 61$. We cannot compute $4331^{-1} \mod 4453$ and we cannot compute $3P$. We have however, found that 61 is a factor of 4453.

4.3 Reduction of curves modulo p

Suppose $p \neq 2, 3$ is a prime divisor of N. We can reduce a curve defined modulo N to a curve defined modulo p.

Definition 4.9. *Let $E_{(a,b)} : y^2 = x^3 + ax + b$ be an elliptic curve defined modulo N. Let $\tilde{a} = a \mod p$ and $\tilde{b} = b \mod p$. The curve $E_{(\tilde{a},\tilde{b})} : y^2 = x^3 + \tilde{a}x + \tilde{b}$ is the **reduction of** $E_{(a,b)}$ **modulo p**. We will refer to this curve as E modulo p.*

Definition 4.10. *If $4a^3 + 27b^2 \mod p \neq 0$ then $E_{(\tilde{a},\tilde{b})}$ has a **good reduction**. If $4a^3 + 27b^2 \mod p = 0$ then $E_{(\tilde{a},\tilde{b})}$ has a **bad reduction**.*

Henceforth, when discussing the reduction of a curve modulo p, we assume that it is a good reduction. The points (x, y) on the curve $E_{(\tilde{a},\tilde{b})}$ are the non-negative integers that are less than p and obey $y^2 = x^3 + \tilde{a}x + \tilde{b}$. From [31] (p.95), the set of all points on an elliptic curve modulo p with the element ∞ and the binary operation of addition forms an abelian group, as in Theorem 4.7. We will denote this group as $E_{(\tilde{a},\tilde{b})}(\mathbb{Z}/p\mathbb{Z})$, or just $E(\mathbb{Z}/p\mathbb{Z})$, depending on context. We denote the number of points in $E(\mathbb{Z}/p\mathbb{Z})$ as N_p. The following theorem, credited to Hasse, states that N_p is finite. Hasse's proof, which dates back to 1933, can be found in section V of [30].

Theorem 4.11. *(Hasse's Theorem) The number of points in $E(\mathbb{Z}/p\mathbb{Z})$ is finite. Furthermore,*

$$|N_p - (p + 1)| \leq 2\sqrt{p}$$

Hasse's Theorem implies that

$$\begin{aligned} N_p &\geq p + 1 - 2\sqrt{p} \\ &= (\sqrt{p} - 1)^2. \end{aligned}$$

as our p is a prime larger than 3, $E_{(\tilde{a},\tilde{b})}(\mathbb{Z}/p\mathbb{Z})$ is guaranteed to have more than one point.

Additionally, if $\gcd(4a^3 + 27b^2, N) = 1$, then $\gcd(4a^3 + 27b^2, p) = 1$ for all prime divisors of N. Thus, our curve modulo N can be reduced to a curve modulo p with coefficients $a \mod p$ and $b \mod p$.

4.4 Lenstra's Elliptic Curve Integer Factorization Method

Lenstra proposes the following integer factorization algorithm in [21]. Suppose we wish to factor the integer N where N is relatively prime to 6. In practice, N has no small factors, as these factors would have been found by Trial Division. As stated above, the ECM operates "by computing elliptic curve addition modulo N" and hoping for a failure. If an elliptic curve addition modulo N fails, then either $(x_1 - x_2)^{-1} \mod N$ or $(2y_1)^{-1} \mod N$ do not exist. That is, they are not relatively prime to N. We can then compute $\gcd(x_1 - x_2, N)$ or $\gcd(2y_1, N)$ to produce a factor of N.

We first examine when an addition modulo N fails. Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be two finite points on the elliptic curve modulo N. Suppose that $x_1 \neq x_2$ and $y_1 \neq y_2$. If a failure occurs when computing $P_1 + P_2$ then $(x_1 - x_2)^{-1}$ does not exist. That is, $x_1 - x_2$ is not relatively prime to N. Suppose $x_1 - x_2$ is a multiple of p, where p is a prime factor of N. We have $x_1 - x_2 \equiv 0 \pmod{p}$. If $x_1 \equiv x_2 \pmod{p}$, then by the elliptic curve group law, $P_1 + P_2 = \infty$ on the curve reduced modulo p.

Doubling P_1 on the curve modulo N will fail if $2y_1$ is a multiple of p, where p is some prime factor of N. This will occur when $2y_1 \equiv 0 \pmod{p}$. If $y_1 \equiv 0 \pmod{p}$, then $2P_1 = \infty$ on the curve modulo p. In general, an addition will fail on the curve modulo N when $P_1 + P_2 = \infty$ on the curve reduced modulo p.

Thus, to factor N we wish to perform elliptic curve additions and doublings modulo N until an addition produces ∞ on some curve modulo p, where p is a prime divisor of N. We use the following procedure which attempts generate to the element ∞ of $E(\mathbb{Z}/p\mathbb{Z})$. Let

$$k = \prod_{\text{r prime, } r \leq B} r^e$$

for some fixed positive integers e and B. Let P be a point on our curve modulo N. We then compute kP on the curve modulo N using the elliptic curve group law.

Recall N_p is the number of points on the elliptic curve modulo p. By Hasse's Theorem, N_p is finite. By Lagrange's Theorem, the order of the

point P on the curve reduced modulo p divides N_p. If k is a multiple of N_p then $kP = \infty$ on the curve reduced modulo p. Thus, in order for the ECM to succeed, we require that for some prime divisor p of N, N_p is B-smooth. Additionally, no prime powers of the form r^d with $d > e$ can divide N_p. If this is the case, then $kP = \infty$ on the curve reduced modulo p, and computing kP on the elliptic curve modulo N will fail. If $(x_1, y_1) + (x_2, y_2)$ is the addition that fails then computing either $\gcd(x_1 - x_2, N)$ or $\gcd(2y_1, N)$ will produce a factor of N.

The ECM may produce the trivial factor 1 or the trivial factor N. Suppose $N = pq$, where p and q are distinct primes. If both N_q and N_p are B-smooth for the same bound B and no prime powers of the form r^d with $d > e$ divide N_p or N_q then k will be a multiple of both N_q and N_p. In this case, kP will produce ∞ on both the curve modulo q and the curve modulo p. Thus, $x_1 - x_2$ and $2y_1$ are multiples of both p and q and computing either $\gcd(x_1 - x_2, N)$ or $\gcd(2y_1, N)$ will produce N.

In general, suppose $N = p_1^{a_1} p_2^{a_2} \dots p_l^{a_l}$, with each p_i prime. If every N_{p_i} divides k then the ECM will produce N. If no N_{p_i} divides k, the ECM will produce 1. When the ECM fails, we can simply try again with a different curve. In practice, the ECM is run on many curves.

In [21], Lenstra gives a heuristic estimate for the complexity of the ECM. This complexity estimate depends on the value of B, and Lenstra gives an optimal value of B as well. This optimal value of B is

$$B = \exp\left(\left(\frac{\sqrt{2}}{2} + o(1)\right)\sqrt{\ln p \ln\ln p}\right)$$

where p is the smallest prime factor of N. Since the value of p is unknown, a precise value of B cannot be used. Instead, Lenstra suggests using $B = 10000$ and arrives at the heuristic estimate of the ECM requiring

$$\exp\left((\sqrt{2} + o(1))\sqrt{\ln p \ln\ln p}\right)$$

arithmetic operations.

Unlike the Quadratic Sieve, the running time of the ECM depends on the size of the smallest prime factor p, not the number being factored, N. However, these are only the known estimates and N could be used in both

and get the same complexity estimate. The ECM has a worst case when N is a product of two equal primes. In this case, Lenstra states that the ECM will take $\exp\left(\sqrt{\ln N \ln\ln N}\right)$ arithmetic operations, the same as the Quadratic Sieve.

In practice, the ECM is used to find small factors of a very large integer with many factors. Once a divisor has been factored out, the remaining number, if it is composite, can be factored using other factorization techniques.

The largest integer factored using the ECM, as of now, has 73 digits and was discovered on 6 March 2010 by Joppe Bos, Thorsten Kleinjung, Arjen Lenstra and Peter Montgomery. For a current list of the largest integers factored using the ECM, see [1]. Additionally, in [3], Brent describes how he used the ECM to find a complete factorization of F_{10}, the tenth Fermat number.

4.5 The ECM in the projective plane

It will be advantageous for us to consider the points on an elliptic curve as points in the projective plane.

Definition 4.12. *Let K be a field. We define* ***the projective plane*** ${P_K}^2$ ***of K*** *to be the equivalence classes of triples $(x, y, x) \in K \times K \times K$ where the equivalence relation is $(x, y, z) \sim (cx, cy, cz)$ where c is a non-zero element of K.*

We denote an equivalence class as $(x : y : z)$. If $z \neq 0$ then $(x : y : z) = (\frac{x}{z}, \frac{y}{z}, 1)$ are the finite points in ${P_K}^2$. The class with $z = 0$ is called the infinity element of ${P_K}^2$.

Definition 4.13. *A polynomial is* ***homogeneous of degree n*** *if it is a sum of terms of the form $ax^i y^j z^k$ with $a \in K$ and $i + j + k = n$.*

If F is homogeneous of degree n then $F(cx, cy, cz) = c^n F(x, y, z)$. If F is homogeneous and $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ then $F(x_1, y_1, z_1) = 0$ if and only if $F(x_2, y_2, z_2) = 0$. Thus, a zero of F in ${P_K}^2$ does not depend on its equivalence class representative. This is our motivation for considering this equivalence relation.

Definition 4.14. *The corresponding homogeneous form of (4.1) is*

$$y^2z = x^3 + axz^2 + bz^3. \tag{4.7}$$

We will refer to this form of elliptic curve as the ***homogeneous Weierstrass form****.*

Recall that the ECM relies on producing ∞ on some elliptic curve modulo p. Our infinity element in projective coordinates is represented by the class $(x : y : 0)$. Observe that in the homogeneous Weierstrass form (4.7), our infinity element is represented by the class $(0 : y : 0) = (0 : 1 : 0)$, as $z = 0$ implies that $x = 0$. Thus, our infinity element in projective coordinates does not depend on the value of y.

When using the ECM to factor N, we hope that $kP = \infty$ on some curve modulo p, where p is a prime divisor of N. In projective coordinates, ∞ is represented by $(0 : 1 : 0)$. If $kP = (x :: z)$ when performed on a curve modulo N and $kP = (0 : 1 : 0)$ on a curve modulo p, where p is some prime factor of N then z must be a multiple of p, and $\gcd(z, N)$ will produce a factor of N.

We have that if $kP = [0 :: 0]$, then $nkP = [0 :: 0]$, for any integer n. Thus, we simply need to find a multiple of this k. In practice, we compute $kP = (x :: z)$ using a curve defined modulo N and then check $\gcd(z, N)$.

Chapter 5

Improving the Elliptic Curve Method

We now focus on improving the Elliptic Curve Method. We wish to reduce the amount and difficulty of computations in the algorithm while increasing the likelihood of finding a factor of N. One way of doing so is by using a curve in Montgomery Form and considering points in the projective plane, as described below. The idea of using Montgomery curves was first proposed in [22], which we loosely follow for the remainder of the paper.

5.1 Montgomery Curves

When we add points on an elliptic curve in Weierstrass form, we must compute an inverse element of our field, either $(2y_1)^{-1}$ or $(x_1 - x_2)^{-1}$. This is done by using the Extended Euclidean Algorithm, and is an expensive computation. However, in [22], Montgomery shows that using an elliptic curve of the form $by^2 = x^3 + ax^2 + x$ allows one to perform additions without having to compute any multiplicative inverses. We will now discuss some preliminaries regarding these types of elliptic curves.

Definition 5.1. *An elliptic curve of the form*

$$by^2 = x^3 + ax^2 + x \text{ with } b(a^2 - 4) \neq 0 \quad (5.1)$$

will be referred to as being in ***Montgomery form****. Every curve in Montgomery form has a corresponding* ***homogeneous Montgomery form****,*

$$by^2z = x^3 + ax^2z + xz^2 \quad (5.2)$$

Observe that we can go from the Montgomery formula to the homogeneous form by replacing x with $\frac{x}{z}$ and y with $\frac{y}{z}$. We assume that the points on our curve are finite, that is, $z \neq 0$. Similar to a Weierstrass Curve, our infinity element for a homogeneous Montgomery form is $(0 : 1 : 0)$. Two Montgomery curves defined over $\mathbb{R}$ are shown in figures 5.1 and 5.2.

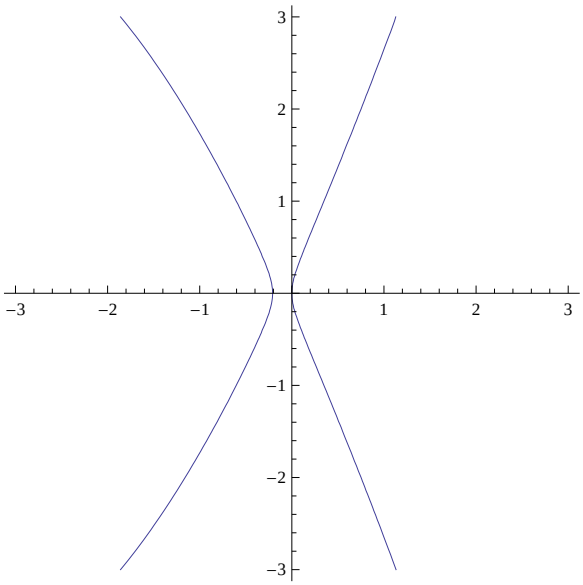

Figure 5.1: The Montgomery curve $y^2 = x^3 - 3x^2 + x$

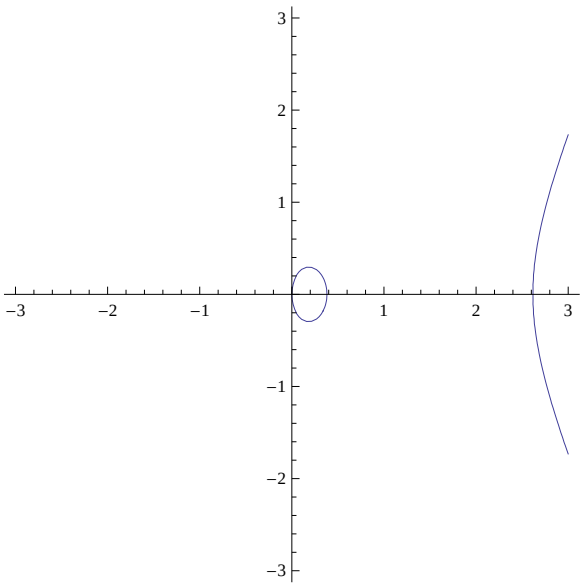

Figure 5.2: The Montgomery curve $y^2 = x^3 + 5x^2 + x$

Proposition 5.2. *Consider $Y^2 = X^3 + AX + B$, an elliptic curve in Weierstrass form. Suppose A can be written as $A = \frac{3-a^2}{3b^2}$ and B can be written as $B = \frac{2a^3-9a}{27b^3}$ for some a and $b \neq 0$ in K. This change of coordinates changes the elliptic curve in Weierstrass form to the curve in Montgomery form $by^2 = x^3 + ax^2 + x$ with $b(a^2 - 4) \neq 0$.*

Proof. Observe that the change of coordinates $X = \frac{3x+a}{3b}$ and $Y = \frac{y}{b}$, performed on $Y^2 = X^3 + AX + B$ produces $by^2 = x^3 + ax^2 + x$, a curve in Montgomery form since

$$\begin{aligned}
Y^2 &= X^3 + AX + B; \\
\left(\frac{y}{b}\right)^2 &= \left(\frac{3x+a}{3b}\right)^3 + \left(\frac{3-a^2}{3b^2}\right)\left(\frac{3x+a}{3b}\right) + \left(\frac{2a^3-9a}{27b^3}\right); \\
27by^2 &= (3x+a)^3 + 3(3-a^2)(3x+a) + 2a^3 - 9a \\
&= (a^3 + 9a^2x + 27ax^2 + 27x^3) + 3(-a^3 - 3a^2x + 3a + 9x) + 2a^3 - 9a \\
&= 27x^3 + 27ax^2 + 27x; \\
by^2 &= x^3 + ax^2 + x.
\end{aligned}$$

Recall that we require the coefficients of our Weierstrass curve to obey $4A^3 + 27B^2 \neq 0$. Through the same change of variables as above,

$$\begin{aligned}
4A^3 + 27B^2 &= 4\left(\frac{3-a^2}{3b^2}\right)^3 + 27\left(\frac{2a^3-9a}{27b^3}\right)^2 \\
&= \frac{4(3-a^2)^3 + a^2(2a^2-9)^2}{27b^6}. \\
&= \frac{(a^2-4)}{b^6}
\end{aligned}$$

Thus, if we require $4A^3 + 27B^2 \neq 0$ then we require $a^2 \neq 4$. We already have that $b \neq 0$. The condition $b(a^2 - 4) \neq 0$ follows. □

Proposition 5.3. *Using the change of variables $\frac{x}{B} = t - \frac{A}{3B}$ and $v = \frac{y}{B}$, the curve $By^2 = x^3 + Ax^2 + x$ in Montgomery form is transformed into a curve in Weierstrass form.*

Proof. We start with the Montgomery curve $By^2 = x^3 + Ax^2 + x$ with

$B(A^2-4) \neq 0$. We divide by B^3 and then perform the change of variables.

$$\begin{aligned} By^2 &= x^3 + Ax^2 + x \\ \frac{y^2}{B^2} &= \frac{x^3}{B^3} + A\frac{x^2}{B^3} + \frac{x}{B^3} \\ v^2 &= \left(t - \frac{A}{3B}\right)^3 + \frac{A}{B}\left(t - \frac{A}{3B}\right)^2 + \frac{1}{B^2}\left(t - \frac{A}{3B}\right) \\ &= t^3 + \left(\frac{3-A^2}{3B^2}\right)t + \frac{2A^3 - 9A}{27B^3} \\ &= t^3 + A't + B' \end{aligned}$$

which is now in Weierstrass form, with coefficients $A' = \frac{3-A^2}{3B^2}$ and $B' = \frac{2A^3-9A}{27B^3}$ provided that the discriminant is non-zero. That is, it must be shown that

$$4\left(\frac{3-A^2}{3B^2}\right)^3 + 27\left(\frac{2A^3-9A}{27B^3}\right)^2 \neq 0$$

In the previous proof, we saw that

$$4\left(\frac{3-A^2}{3B^2}\right)^3 + 27\left(\frac{2A^3-9A}{27B^3}\right)^2 = \frac{(A^2-4)}{B^6}.$$

Thus, the condition $B(A^2-4) \neq 0$ shows that the discriminant of $t^3 + A't + B'$ is non-zero.

□

5.2 Addition for elliptic curves in Montgomery form

We now describe the group law for the rational points on a Montgomery curve defined over $\mathbb{Q}$. These formulas are obtained in the same manner as equations (4.2) and (4.6).

Theorem 5.4. *(Montgomery curve addition) For the rational points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ on the Montgomery curve $By^2 = x^3 + Ax^2 + x$, with $x_1 \neq x_2$ and $y_1 \neq y_2$, and A and B rational, $+$ operates as follows:*

$$P_1 + P_2 = (x_3, y_3) = (Bm^2 - A - x_1 - x_2, m(2x_1 + x_2 + A) - Bm^3 - y_1), \quad (5.3)$$

where

$$m = \frac{y_1 - y_2}{x_1 - x_2}.$$

Theorem 5.5. *(Montgomery curve point doubling) For the rational point $P_1 = (x_1, y_1)$ on the Montgomery curve $By^2 = x^3 + Ax^2 + x$ with $y_1 \neq 0$, and A and B rational,*

$$2P_1 = (x_3, y_3) = (Bm^2 - A - 2x_1, m(3x_1 + A) - Bm^2 y_1), \tag{5.4}$$

where

$$m = \frac{3x_1^2 + 2Ax_1 + 1}{2By_1}.$$

Equations (5.3), (5.4) and the identities in (4.7) form the Montgomery curve group law. Just as with a Weierstrass curve, the set of all rational points on a Montgomery curve over $\mathbb{Q}$ and the ∞ element forms a group. Our motivation for considering this group comes from [22], where Montgomery shows that using a Montgomery curve with projective coordinates affords an addition formula that avoids the computation of multiplicative inverses. We will see that this addition formula does not require the y-coordinates of points, nor does it produce the y-coordinate of the sum. We will write $(x :: z)$ to denote a projective coordinate with the y-coordinate disregarded.

Theorem 5.6. *(Montgomery addition) Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be two rational points on the Montgomery curve $By^2 = x^3 + Ax^2 + x$ with $x_1 \neq x_2$, $y_1 \neq y_2$ and A and B rational. If $P_1 + P_2 = (X_3 :: Z_3)$ and $P_1 - P_2 = (X_4 :: Z_4)$ in projective coordinates then*

$$P_1 + P_2 = (X_3 :: Z_3) = \left(Z_4(X_1X_2 - Z_1Z_2)^2 :: X_4(X_1Z_2 - Z_1X_2)^2\right) \tag{5.5}$$

where $x_i = \frac{X_i}{Z_i}$.

Proof. Suppose $P_1 + P_2 = (x_3, y_3)$. From (5.3), we have that $x_3 = Bm^2 - A - x_1 - x_2$. Next, we substitute $m = \frac{y_1 - y_2}{x_1 - x_2}$ and proceed algebraically. In the following arithmetic, we use the substitutions $By_1{}^2 = x_1{}^3 + Ax_1{}^2 + x_1$ and $By_2{}^2 = x_2{}^3 + Ax_2{}^2 + x_2$, which follow from the fact that P_1 and P_2 are

points on our curve.

$$
\begin{aligned}
x_3 &= -A + B\left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2 - x_1 - x_2 \\
x_3(x_1 - x_2)^2 &= B(y_1 - y_2)^2 - (A + x_1 + x_2)(x_1 - x_2)^2 \\
&= B{y_1}^2 - 2By_1y_2 + B{y_2}^2 \\
&\qquad - (A{x_1}^2 - 2Ax_1x_2 + A{x_2}^2 + {x_1}^3 - 2{x_1}^2x_2 + x_1{x_2}^2 + {x_1}^2x_2 - 2x_1{x_2}^2 - {x_2}^3) \\
&= ({x_1}^3 + A{x_1}^2 + x_1) - 2By_1y_2 + ({x_2}^3 + A{x_2}^2 + x_2) \\
&\qquad - (A{x_1}^2 - 2Ax_1x_2 + A{x_2}^2 + {x_1}^3 - 2{x_1}^2x_2 + x_1{x_2}^2 + {x_1}^2x_2 - 2x_1{x_2}^2 - {x_2}^3) \\
&= -2By_1y_2 + x_1 + x_2 + x_1x_2(2A + x_2 + x_1) \\
&= -2By_1y_2 + {x_1}^2x_2 + Ax_1x_2 + x_2 + x_1{x_2}^2 + Ax_1x_2 + x_1 \\
&= ({x_1}^2 + Ax_1 + 1)x_2 - 2By_1y_2 + ({x_2}^2 + Ax_2 + 1)x_1 \\
&= \frac{({x_1}^3 + A{x_1}^2 + x_1)x_2}{x_1} - 2By_1y_2 + \frac{({x_2}^3 + A{x_2}^3 + x_2)x_1}{x_2} \\
&= \frac{(B{y_1}^2)x_2}{x_1} - 2By_1y_2 + \frac{(B{y_2}^2)x_1}{x_2} \\
&= \frac{B\left((x_2y_1)^2 - 2x_1y_2x_2y_1 + (x_1y_1^2)\right)}{x_1x_2} \\
&= \frac{B(x_2y_1 - x_1y_2)^2}{x_1y_1}
\end{aligned}
$$

We have found that

$$x_3(x_1 - x_2)^2 = \frac{B(x_2y_1 - x_1y_2)^2}{x_1y_1} \tag{5.6}$$

Let $P_1 - P_2 = (x_4, y_4)$. Since $-P_2 = (x_2, -y_2)$, we also have

$$x_4(x_1 - x_2)^2 = \frac{B(x_2y_1 + x_1y_2)^2}{x_1y_1} \tag{5.7}$$

Multiplying the two equations (5.6) and (5.7) together yields

$$\begin{aligned} x_3x_4(x_1-x_2)^4 &= \frac{B^2\left((y_1x_2)^2-(y_2x_1)^2\right)^2}{(x_1x_2)^2} \\ &= \frac{\left((x_1{}^3+Ax_1{}^2+x_1)x_2{}^2-(x_2{}^3+Ax_2{}^2+x_2)x_1{}^2\right)^2}{(x_1x_2)^2} \\ &= \frac{\left(x_1x_2(x_1^2x_2+x_2-x_1x_2{}^2-x_1)\right)^2}{(x_1x_2)^2} \\ &= (x_1^2x_2+x_2-x_1x_2^2-x_1)^2. \end{aligned}$$

Thus,

$$\begin{aligned} x_3x_4(x_1-x_2)^2 &= \left(\frac{x_1^2x_2+x_2-x_1x_2^2-x_1}{x_1-x_2}\right)^2 \\ &= \left(\frac{(x_1-x_2)(x_1x_2-1)}{x_1-x_2}\right)^2 \\ &= (x_1x_2-1)^2. \end{aligned}$$

For projective coordinates, we replace x_i with $\frac{X_i}{Z_i}$ for $i=1,2,3,4$ and then

$$\begin{aligned} \frac{X_3X_4}{Z_3Z_4} &= \left(\frac{\frac{X_1X_2}{Z_1Z_2}-1}{\frac{X_1}{Z_1}-\frac{X_2}{Z_2}}\right)^2 \\ \frac{X_3}{Z_3} &= \frac{Z_4(X_1X_2-Z_ZZ_2)^2}{X_4(X_1Z_2-X_2Z_1)^2}. \end{aligned}$$

Therefore, we now have a formula for addition of two points on a homogeneous Montgomery curve, provided we know their difference. That is, if $P_1+P_2=(x_3::z_3)$ and $P_1-P_2=(X_4::Z_4)$ then

$$X_3=Z_4(X_1X_2-Z_1Z_2)^2 \text{ and } Z_3=X_4(X_1Z_2-X_2Z_1)^2 \tag{5.8}$$

□

In practice, we set

$$t_1=(X_1-Z_1)(X_2+Z_2), t_2=(X_1+Z_1)(X_2-Z_2)$$

and compute

$$(X_3::Z_3)=(Z_4(t_1+t_2)^2::X_4(t_1-t_2)^2).$$

Using this method, a Montgomery addition can be computed with four additions, two squarings, four multiplications and no inversions, provided that $P_1 - P_2 = (X_4 :: Z_4)$ is known. We now consider doubling a point in Montgomery form.

Theorem 5.7. *(Montgomery Doubling)For the rational point $P_1 = (x_1, y_1)$ on the Montgomery curve $By^2 = x^3 + Ax^2 + x$ with A and B rational, $2P_1 = (X_3 :: Z_3)$ where*

$$X_3 = (X_1 + Z_1)^2(X_1 - Z_1)^2 \text{ and } Z_3 = 4Z_1X_1\left((X_1 - Z_1)^2 + X_1Z_1(A + 2)\right) \tag{5.9}$$

with $x = \frac{X_1}{Z_1}$.

Proof. Suppose that $2P_1 = (x_3, y_3)$. From (5.4), we have that $x_3 = -A + Bm^2 - 2x_1$, where $m = \frac{3x_1^2+2Ax_1+1}{2By_1}$. Substitution of this value, and of ${y_1}^2 = {x_1}^3 + A{x_1}^2 + x_1$ yields the following:

$$\begin{aligned}
x_3 &= -A + Bm^2 - x_1 - x_2 \\
x_3 &= -A + B\left(\frac{3{x_1}^2 + 2Ax_1 + 1}{2By_1}\right)^2 - 2x_1 \\
(x_3)(2By_1)^2 &= -A(2By_1)^2 + B(3{x_1}^2 + 2Ax_1 + 1)^2 - (2x_1)(2By_1)^2 \\
4B(x_3)(By_1)^2 &= -4AB(By_1)^2 + B(3{x_1}^2 + 2Ax_1 + 1)^2 - 8Bx_1(By_1)^2 \\
4x_3({x_1}^3 + A{x_1}^2 + x_1) &= -4A({x_1}^3 + A{x_1}^2 + x_1) + (3{x_1}^2 + 2Ax_1 + 1)^2 \\
&\quad - 8x_1({x_1}^3 + A{x_1}^2 + x_1) \\
4x_3x_1({x_1}^2 + Ax_1 + 1) &= ({x_1}^2 - 1)^2
\end{aligned}$$

Replacing x_i with $\frac{X_i}{Z_i}$ for $i = 1, 3$ results in the following:

$$4\frac{X_1X_3}{Z_1Z_3}\left(\left(\frac{X_1}{Z_1}\right)^2+A\frac{X_1}{Z_1}+1\right)=\left(\left(\frac{X_1}{Z_1}\right)^2-1\right)^2$$

$$\begin{aligned}\frac{X_3}{Z_3}&=\frac{Z_1\left(\left(\frac{X_1}{Z_1}\right)^2-1\right)^2}{\left(\left(\frac{X_1}{Z_1}\right)^2+A\frac{X_1}{Z_1}+1\right)(4X_1)}\\&=\frac{((X_1-Z_1)(X_1+Z_1))^2\,(\frac{1}{Z_1})}{(\frac{{X_1}^2+AX_1Z_1+{Z_1}^2}{Z_1^2})(4X_1)}\\&=\frac{((X_1-Z_1)(X_1+Z_1))^2}{({X_1}^2+AX_1Z_1+{Z_1}^2)(4X_1Z_1)}\\&=\frac{((X_1-Z_1)(X_1+Z_1))^2}{((X_1-Z_1)^2+X_1Z_1(A+2))(4X_1Z_1)}.\end{aligned}$$

Thus, $2P_1=(X_3::Z_3)$ where

$$X_3=(X_1+Z_1)^2(X_1-Z_1)^2 \text{ and } Z_3=4Z_1X_1\left((X_1-Z_1)^2+X_1Z_1(A+2)\right).$$

□

In practice, we perform a doubling as follows. Let

$$u=(X_1+Z_1)^2, v=(X_1-Z_1)^2, b=\frac{A+2}{4} \text{ and } t=b(u-v)+v$$

and compute

$$X_3=uv \text{ and } Z_3=(u-v)t.$$

This requires only 5 multiplications.

5.3 Montgomery multiplication

In the context of the ECM, we wish to calculate kP for many positive integers k and some point P on our curve modulo N. Despite the computational advantages of Montgomery addition, we are still limited in that we must know P_1-P_2 in order to compute P_1+P_2. We will overcome this limitation by using Lucas chains.

Definition 5.8. *A **Lucas chain** for an integer k is an increasing sequence of integers $(a_0, a_1, \ldots, a_l)$ with $a_0 = 1$, $a_1 = 2$ and $a_l = k$ such that for every a_n, there exists $0 \leq i, j, m \leq n$ such that $a_n = a_i + a_j$ where either $a_i = a_j$ or $|a_i - a_j| = a_m$ for some $0 \leq i, j, m < n$.*

Example 5.9. *The sequence of powers of 2, $(1, 2, 4, 8, 16, 32 \ldots)$ and the Fibonacci sequence with starting point 1, $(1, 2, 3, 5, 8, 13, \ldots)$ are Lucas chains.*

If a Lucas chain is known, it can be used to compute kP using the Montgomery addition formula (5.5). Suppose $(a_0, a_1, \ldots a_l)$ is a Lucas chain for k. First, we compute $2P$ using the Montgomery doubling formula (5.9). For $n \geq 2$, each a_n is a term in the Lucas Chain for k, so $a_nP = (a_j + a_i)P$ for some $0 \leq i, j < n$. If $a_i = a_j$ then $a_nP = (2a_i)P$ can be calculated using the Montgomery doubling formula (5.9). Suppose $|a_i - a_j| = a_m$. Since we are assuming we have calculated a_mP with $0 \leq m < n$, $|a_i - a_j|P$ has already been calculated. As $|a_i - a_j|P$ is known, we can use (5.5) to compute $a_nP = (a_j + a_i)P$. In this manner, we can successively compute a_nP for $n = 3, 4, \ldots, l$, eventually computing kP.

We need a way to generate a Lucas chain for k. If k is even, then $k = 2^b a$, where b is a positive integer and a is an odd integer. If aP is known, we can use the Montgomery doubling formula (5.9) b times to arrive at kP. Computing a doubling using (5.9) does not require a Lucas chain. Thus, we are only interested in finding Lucas chains for odd numbers. In [23], Montgomery describes the following method, known as the the binary method, for constructing a Lucas chain for a positive odd integer k.

Let $k = d_l + d_{l-1}2 + d_{l-2}2^2 + \ldots + d_0 2^l$ with $d_i \in \{0, 1\}$ and with l the largest integer such that $2^l < k$. Define $a_0 := d_0 = 1$ and inductively define $a_i = 2a_{i-1} + d_i$. We will show that the sequence

$$(a_0, a_0 + 1, a_1, a_1 + 1, \ldots, a_{l-1}, a_{l-1} + 1, a_l)$$

forms a Lucas chain for k.

Depending on d_i, the elements a_{i+1} and $a_{i+1} + 1$ can be written as either $2a_i, a_i + (a_i + 1)$ or $2(a_i + 1)$, all of which are a sum of two previous elements of our sequence. These two elements differ by either one or zero. Thus, the above sequence is a Lucas chain. It remains to show that $a_l = k$. By

definition,

$$\begin{aligned} a_l &= 2a_{l-1} + d_l \\ &= 2(2a_{l-2} + d_{l-1}) + d_l \\ &= 2^2 a_{l-2} + 2d_{l-1} + d_l \\ &= 2^2(2a_{l-3} + d_{l-2}) + 2d_{l-1} + d_l \\ &= 2^3 a_{l-3} + 2^2 d_{l-2} + 2d_{l-1} + d_l \\ &\vdots \\ &= d_0 2^l + d_1 2^{l-1} + \ldots + d_{l-1}2 + d_l \\ &= k. \end{aligned}$$

We can now use (5.5) to compute $a_n P$.

Example 5.10. *Suppose we wish to compute kP using (5.5), where k=23. We begin by constructing a Lucas chain for 23.*

Let $k = 23 = 2^4 + 2^2 + 2^1 + 2^0$. We have $d_0 = 1, d_1 = 1, d_2 = 1, d_3 = 0$ and $d_4 = 1$. Thus, $a_0 = 1, a_1 = 2, a_2 = 5, a_3 = 11$ and $a_4 = 23$. From this, a Lucas chain for 23 is

$$(1, 2, 3, 5, 6, 11, 12, 23).$$

We compute $2P$ using (5.9). Since $3P = (2+1)P$ and we know $(2-1)P = P$, we can use (5.5) to compute $3P$. Since $5P = (3+2)P$ and we know $(3-2)P$, we can compute $5P$ using (5.5). We continue this process to calculate $23P$.

5.4 Recent developments

Several extensions of the ECM have been proposed recently. In 2010, it is proposed [11] that a *twisted Edwards curve*, that is, a curve of the form

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

has an addition that is faster than addition on a Montgomery curve. Additionally, it is proposed that performing the ECM using a twisted Edwards curve will find more primes than a Montgomery Curve.

In 2010, Cossett [8] proposes using hyperelliptic curves. Hyperelliptic curves are curves of the form $y^2 = f(x)$, where $f(x)$ is a polynomial of degree 5 with distinct roots. Addition on a hyperelliptic curve is slower than on an elliptic curve. Furthermore, finding a prime factor using the "hyperelliptic curve method"(HECM) is less likely than with the ECM. However, Cossett shows that for "decomposable" hypercurves, one run of the HECM is comparable to two of the ECM. Cosset claims that using the HECM to factor integers is no worse than the ECM, but hopes that further developments could make HECM faster than ECM.

5.5 Conclusion

We have now seen several integer factorization algorithms. In Chapter 3, we described the Pollard's Rho algorithm and the Quadratic Sieve. We described the algebraic properties of elliptic curves and the ECM in Chapter 4. In Chapter 5, we saw that curves in Montgomery form reduce the computation in the ECM. This concludes our exposition of integer factorization algorithms.

Bibliography

[1] 50 Largest factors found by ECM. http://www.loria.fr/ zimmerma/records/top50.html.

[2] Richard P. Brent. Parallel Algorithms for Integer Factorization. *Number Theory and Cryptography*, pages 26–37, 1990.

[3] Richard P. Brent. Factorization of the Tenth and Eleventh Fermat Numbers. *Computer Sciences Laboratory, Australian National Univ.*, 1996.

[4] Richard P. Brent. Some Parallel Algorithms for Integer Factorization. *Euro-Par Parallel Processing*, 1685:1–22, 1999.

[5] Richard P. Brent and John M. Pollard. Factorization of the Eighth Fermat Number. *Mathematics of Computation*, 36, 1981.

[6] R.P. Brent. An improved Monte Carlo factorization algorithm. *BIT*, 20:176–184, 1980.

[7] David M. Bressoud. *Factorization and Primality Testing*. Springer, 1989.

[8] Romain Cosset. Factorization with genus 2 curves. *Mathematics of Computation*, 2010.

[9] Robert Crandall and Carl Pomerance. *Prime numbers: A Computational Perspective*. Springer, 2005.

[10] Arjen K. Lenstra Derek Atkins, Michael Graff and Paul C. Leyland. The Magic Words Are Squeamish Ossifrage, 1994. http://web.mit.edu/warlord/www/rsa129.ps.

[11] T.Lange D.J. Bernstein, P. Birkner and C.Peters. ECM using Edwards curves. *Mathematics of Computation, to appear*, 2010.

[12] David S Dummit and Richard M. Foote. *Abstract Algebra*. Prentice Hall, 1991.

[13] Martin Gardner. Mathematical Games: A new kind of cipher that would take millions of years to break, 1977. http://simson.net/ref/1977/GardnerRSA.pdf.

[14] G.H. Hardy and E.M. Wright. *An introduction to the theory of numbers*. Oxford University Press, 1979.

[15] Graham James and Oscar Jameson. *The Prime Number Theorem*. Cambridge University Press, 2003.

[16] J.W.S.Cassels. The rational solutions of the Diophantine equation $y^3 = x^3 - D$. *Acta Math.*, 82, 1950.

[17] Knuth. *Art of computer Programming*. Addison-Wesley, 1998.

[18] Knuth. *A course in computational algebraic number theory*. Springer, 2005.

[19] RSA Laboratories. How large a key should be used in the RSA cryptosystem? http://www.rsa.com/rsalabs/node.asp?id=2218.

[20] RSA Laboratories. Is the RSA cryptosystem currently in use? http://www.rsa.com/rsalabs/node.asp?id=2222.

[21] Hendrik W. Lenstra. Factoring Integers with Elliptic Curves. *Annals of Mathematics*, 126:649–673, 1987.

[22] Peter L. Montgomery. Speeding the Pollard and Elliptic Curve Methods of Factorization. *Mathematics of Computation*, 48:243–264, 1986.

[23] Peter L. Montgomery. An FFT extension of the Elliptic Curve Method of Factorization. 1992.

[24] L.J. Mordell. On the rational solutions of the indeterminate equations of the third and fourth degrees. *Proc Cam. Phil. Soc.*, 21, 1922.

[25] John M. Pollard. A Monte Carlo Method for Factorization. *BIT Numerical Mathematics*, 15:331334, 1975.

[26] Carl Pomerance. Analysis and Comparison of Some Integer Factoring Algorithms. *Computational Methods in Number Theory*, pages 89–139, 1982.

[27] Carl Pomerance. A tale of two sieves. *Notices of the AMS*, page 14731485, 1996.

[28] A. Shamir R.L. Rivest and L.Adleman. A Method for Obtaining Digital Signitures and Public-Key Cryptosystems. *Communications of the ACM*, 21, 1978.

[29] R.W.Floyd. Non-deterministic Algorithms. *Journal of the ACM (JACM)*, 14, 1967.

[30] Joseph H. Silverman. *Arithmetic of Elliptic Curves*. Springer, 2009.

[31] Lawrence C. Washington. *Elliptic Curves: Number Theory and Cryptography*. Chapman and Hall, 2008.

[32] Andrew Granville W.R. Alford and Carl Pomerance. There are Infinitely Many Carmichael Numbers. *Annals of Mathematics*, 140:703–722, 1994.