Lenstra's Factoring Method with Elliptic Curves

Xun He

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science at

Concordia University

Montréal, Québec, Canada

June 2005

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-10213-6
Our file Notre référence
ISBN: 0-494-10213-6

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Xun He**

Entitled: **Lenstra's Factoring Method with Elliptic Curves**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science in Mathematics

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

________________________________ Chair

________________________________ Examiner

________________________________ Examiner

________________________________ Supervisor

Approved by ____
Chair of Department or Graduate Program Director

_______ 20 ___ __________________________________
Dean of Faculty

Abstract

Lenstra's Factoring Method
with Elliptic Curves

Xun He

Suppose that we want to factorize an integer N. We can use Lenstra's method, which is based on elliptic curves over finite fields, to find the smallest non-trivial prime factor p of N. The success of Lenstra's algorithm depends on the probability to find an elliptic curve over the finite field with p elements such that the number of points on the curve doesn't have large prime factor. One advantage of Lenstra's algorithm is that we can try different curves to increase the success probability. Lenstra's algorithm has sub-exponential running time.

In this thesis, we study Lenstra's algorithm and an implementation due to Brent, which has reduced the theoretical running time, under certain circumstances. We state their success conditions, success probabilities and running times, and discuss the relevant proofs. We also use PARI to implement this algorithm with Lenstra's and Brent's methods, do some tests, and collect some data which verify the theoretical results.

Acknowledgments

I wish to thank my supervisor Dr. Chantal David for her suggestion and comments throughout the preparation of this thesis, and for her invaluable guidance and teaching during my life at Concordia University.

I also want to thank all the committee members for reading my thesis and giving so many constructive suggestions.

Finally, I thank all my friends in the Department of Mathematics and Statistics, for their encouragement and help.

Contents

Chapter 1

Introduction

The problem of factoring an integer N is generally changed to the problem of finding a non-trivial prime factor p of N. The simplest method is to try to divide N by all numbers up to $\sqrt{N}$, but this approach is impractical since it has exponential running time. The best algorithms known today are the Quadratic Sieve and the Number Field Sieve, and they have sub-exponential running time. The Quadratic Sieve has running time $L(N)^{1+o(1)}$ where $L(x) = \exp\left(\sqrt{\log x \log\log x}\right)$. It is believed (but still unproved) that there exists no polynomial time algorithms on classical computer [1] , and that the sub-exponential running time algorithms are in some sense the best possible.

We study in this thesis a factorization method due to H. Lenstra, which uses elliptic curves. It has running time $L(p)^{\sqrt{2}+o(1)}$ where p is the smallest prime factor of N. Then, this algorithm is faster where N has a "small" prime factor, a feature which is unique to Lenstra's algorithm.

[1] If we use quantum computer to replace classical computer, we can factorize an integer N in polynomial running time with a certain algorithm [14].

Let x be an integer. If x has only small prime factors, then x is called a smooth number (see Section 4.1 for a more precise definition). If G_p is certain Abelian group related to p, then the smoothness of the order of G_p can be used in factorizing N. Number theorists invented several such factoring methods.

Pollard's "$p-1$" method [12] uses the multiplicative group $\mathbf{Z}_p^* = \{1, \ldots, p-1\}$. If N has a prime factor p such that $p-1$ (the order of $\mathbf{Z}_p^*$) only has small prime factors, then N can be factorized. If there doesn't exist such a prime factor p of N, then Pollard's "$p-1$" method will fail.

Lenstra's method [10], which uses $E(\mathbf{Z}_p)$, the group of points of an elliptic curve over $\mathbf{Z}_p$, is an improvement following Pollard's idea. $E(\mathbf{Z}_p)$ is also an Abelian group with order $N_p = \#E(\mathbf{Z}_p)$. Let p be a prime factor of N. If we can find a certain elliptic curve such that N_p has only small prime factors, then we can factorize N (under some additional mild conditions). The success probability of Lenstra's method is higher than Pollard's "$p-1$" method, since there are approximately $4[\sqrt{p}]+1$ different values of N_p for a given p, and we can try many different groups $E(\mathbf{Z}_p)$.

Lenstra defined a pseudo-group law on $E(\mathbf{Z}_N)$ ("Pseudo" means the addition cannot always give an element in the set. See Section 4.2 for a detail definition). Unlike the group law on $E(\mathbf{Z}_p)$, the pseudo-addition with two points on $E(\mathbf{Z}_N)$ is not always successful. When it fails, it will give a non-trivial prime factor of N. At the beginning, we fix a bound w and a number k which does not have prime factor $> w$. And then we perform a *TRIAL*, which means a choice of a random curve $E(\mathbf{Z}_N)$ with a point P on E and the computation of kP (pseudo-adding P for k times). If N_p does not have prime factor larger than w, and $N_p|k$, then the computation of kP is very likely to fail and consequently N will be factorized. Otherwise the algorithm has to choose another different curve $E(\mathbf{Z}_N)$ and to do another trial.

Brent's method [1] is a variation on the implementation of Lenstra's algorithm. In Lenstra's algorithm, a trial fails when computing $Q = kP$ is successful. However that computation is regarded as the first phase of a *TRIAL* in Brent's implementation. If it fails, the algorithm will try to find two multiples of Q, $Q_1 = (x_1, y_1)$, $Q_2 = (x_2, y_2)$, such that $x_1 - x_2$ has a non-trivial common factor with N. That is regarded as the second phase of a trial in Brent's implementation. If it still fails, then we have to go to next trial. Brent's implementation can be faster than Lenstra's method under certain circumstances (see Chapter 5).

In this thesis, we study Lenstra's algorithm and Brent's method. Some background knowledge on elliptic curves is provided in Chapter 2. And all the theoretical tools about factoring methods are assembled in Chapter 3. Chapter 4 is concerned with Lenstra's method, and Chapter 5 with Brent's implementation of Lenstra's method. We state their success conditions, success probabilities and running times, and give the relevant proofs. Finally we used PARI to implement this algorithm with Lenstra's and Brent's methods, did some tests, and collected some data. The results, which verify the theory, are in Chapter 6.

Chapter 2

Background

2.1 Important symbols

This section is just a list of all important symbols in this paper.

$\ll$, O, $\gg$: If there exists $C > 0$ and $n_0 > 0$ such that $0 \leq f(n) \leq Cg(n)$ when $n \geq n_0$, then we say $f(n) \ll g(n)$, $f(n) =$O$(g(n))$, or $g(n) \gg f(n)$.

$\asymp$: If $f(n) \ll g(n)$ and $f(n) \gg g(n)$ at the same time, then we say $f(n) \asymp g(n)$.

o: If $\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0$, then we say $f(n) =$o$(g(n))$.

R: **R** is a ring.

$\mathbb{P}^2(\mathbf{R})$: $\mathbb{P}^2(\mathbf{R})$ is the projective plane over **R**.

K, **K***: **K** is a field. $\mathbf{K}^*$ is the set of units in **K**. $\mathbf{K}^* = \mathbf{K} - \{0\}$.

$\bar{\mathbf{K}}$: $\bar{\mathbf{K}}$ is the algebraic closure of **K**.

$\mathbb{P}^2(\bar{\mathbf{K}})$: $\mathbb{P}^2(\bar{\mathbf{K}})$ is the projective plane over $\bar{\mathbf{K}}$.

P and $[X, Y, Z]$: P is a point on $\mathbb{P}^2(\bar{\mathbf{K}})$. $[X, Y, Z]$ are the coordinates of P in $\bar{\mathbf{K}}$.

E and O: E is an elliptic curve over $\mathbf{K}$. If $\text{Char}(\mathbf{K}) \neq 2, 3$, then E has the form $y^2z = x^3 + axz^2 + bz^3$ with $a, b \in \mathbf{K}$, and it can be represented by an ordered pair (a, b). O is the point at infinity on an elliptic curve.

$\mathbf{Z}_p$: $\mathbf{Z}_p$ is the field with p elements.

$E(\mathbf{Z}_p)$ and N_p: $E(\mathbf{Z}_p)$ is the set of all rational points on an elliptic curve over $\mathbf{Z}_p$. If p is a prime number > 3, then the curve can be represented by an ordered pair $(a, b) \in \mathbf{Z}_p^2$. N_p denotes the number of elements in the group $E(\mathbf{Z}_p)$.

$\{E : E(\mathbf{Z}_p)\}$: $\{E : E(\mathbf{Z}_p)\}$ is the set of all elliptic curves E over $\mathbf{Z}_p$.

$\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$: $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$ is the set of all isomorphic classes of elliptic curves E over $\mathbf{Z}_p$.

$\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$: $\mathcal{S}$ is a set of integers. $\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$ is the set of those isomorphic classes with $N_p \in \mathcal{S}$.

$\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$: $\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$ is the weighted cardinality of the set $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$ with weight $(\#\text{Aut}E)^{-1}$.

N: N is the positive integer that we want to factorize.

p and q: p, q are two distinct prime factors of N. p is the smallest one (We assume that p is at least 5).

$\mathbf{Z}_N$: $\mathbf{Z}_N$ is the ring of integers modulo N.

$E(\mathbf{Z}_N)$ and V_N: $E(\mathbf{Z}_N)$ is the set of all rational points on an elliptic curve over $\mathbf{Z}_N$. The curve can be represented by an ordered pair $(a, b) \in \mathbf{Z}_N^2$. V_N is a subset of $E(\mathbf{Z}_N)$.

$E(\mathbf{Z}_p)$ and P_p: $E(\mathbf{Z}_p)$ is the reduced curve $E(\mathbf{Z}_N)$ modulo p. If (a, b) represents $E(\mathbf{Z}_N)$, then $E(\mathbf{Z}_p)$ is represented by $(\bar{a}, \bar{b})$ with $\bar{a} = a \pmod{p}$, and $\bar{b} = b \pmod{p}$.

P_p is the reduction of the point P modulo p. If $P = [X, Y, Z]$, then $P_p = [X \pmod p, Y \pmod p, Z \pmod p]$.

$E(\mathbf{Z}_q)$ and P_q: Similarly, $E(\mathbf{Z}_q)$ is the reduced curve $E(\mathbf{Z}_N)$ modulo q, and P_q is the reduction of P modulo q.

2.2 Elliptic curves

Definition. Weierstrass equation.

Let $\mathbf{K}$ be a field. The equation

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with all coefficients $\in \mathbf{K}$, is called a Weierstrass equation over $\mathbf{K}$. If the characteristic of $\mathbf{K}$ is not 2 or 3, then this equation can be changed into the canonical form (See [15] , p.10-11)

$$y^2z = x^3 + axz^2 + bz^3.$$

In this article, we only discuss the situation with Char($\mathbf{K}$) $\neq 2$ or 3.

Definition. Elliptic curve.

A Weierstrass equation $y^2z = x^3 + axz^2 + bz^3$ $(4a^3 + 27b^2 \neq 0)$ over $\mathbf{K}$, determines an elliptic curve E in the projective plane $\mathbb{P}^2(\bar{\mathbf{K}})$, which contains all the points $[X, Y, Z]$ where X, Y, Z is a solution of the equation. E can then be represented by an ordered pair (a, b). Suppose that $P = [X, Y, Z]$ where $Z \neq 0$ is a point on E, then P can be also represented by an ordered pair (x, y) with $x = X/Z$ and $y = Y/Z$.

Definition. The point at infinity.

If $Z = 0$ then $X^3 = 0$ and hence $X = 0$. So $O = [0, 1, 0]$ is the only point on E with $Z = 0$. We call it the point at infinity.

Definition. Rational points.

Let $P = [X, Y, Z]$ be a point on $\mathbb{P}^2(\bar{\mathbf{K}})$. Let $\bar{\mathbf{K}}^* = \bar{\mathbf{K}} - \{0\}$. If there exists $\lambda \in \bar{\mathbf{K}}^*$ such that $\lambda X, \lambda Y, \lambda Z \in \mathbf{K}$, then we say P is a rational point over $\mathbf{K}$. $E(\mathbf{K})$ denotes the set of all rational points (over $\mathbf{K}$) on E.

Definition. Group structure on $E(\mathbf{K})$.

Let the point O be the zero, "$-$" be the opposite operation, and "$+$" be the addition operation. We can define an Abelian group structure on $E(\mathbf{K})$ as follows:

Rule. **Group law on E.**

1. *opposite.* Let P be a point on E. Its opposite "$-P$" is defined as:

(i) If $P = O$, let "$-P$" be the point O.

(ii) If $P \neq O$, take a line L passing through P and O. If L is the tangent line at P, then let "$-P$" be the point P. If L intersects E at the third point P', then let "$-P$" be the point P'.

Suppose that $P = (x, y)$, then the formula of "$-P$" is $-P = (x, -y)$.

2. *addition.* Let P, Q be two points on E. The addition "$P + Q$" is defined as:

(i) If $P = O$, let "$P + Q$" be the point Q. If $Q = O$, let "$P + Q$" be the point P.

(ii) If $P \neq O$, $Q \neq O$, take a line L passing through P and Q. If L passes through O, then let "$P + Q$" be the point O. If L intersects E at the third point R, then let "$P + Q$" be the point "$-R$", the opposite of R.

Suppose that $P = (x_1, y_1) \neq O$, $Q = (x_2, y_2) \neq O$, $P \neq -Q$, $P + Q = (x_3, y_3)$, then (x_3, y_3) can be calculated as follows:

$x_3 = \lambda^2 - x_1 - x_2$, $y_3 = \lambda(x_1 - x_3) - y_1$, where

$\lambda = (y_2 - y_1)/(x_2 - x_1)$, when $x_1 \neq x_2$,

$\lambda = (3x_1^2 + a)/(2y_1)$, when $x_1 = x_2$.

Remark. **Elliptic curves over R.**

Let $\mathbf{R}$ be a ring. Since a non-zero element in $\mathbf{R}$ is not necessarily invertible, a point $[X, Y, Z] \in \mathbb{P}^2(\mathbf{R})$ cannot be always changed into an ordered pair (x, y) with $x = X/Z$ and $y = Y/Z$. Hence, we cannot put such an Abelian group structure on $E(\mathbf{R})$ as what we defined on $E(\mathbf{K})$ above.

Remark. **Elliptic curves over finite fields.**

If $\mathbf{K}$ is a finite field, then $E(\mathbf{K})$ is a finite group. The number of curves over $\mathbf{K}$ is also finite. When $\mathbf{K} = \mathbf{Z}_p$, we have:

Proposition 2.1. *The number of all elliptic curves E over $\mathbf{Z}_p$ is*

$$\#\{E : E(\mathbf{Z}_p)\} = p^2 - p \, .$$

Proof. There are exactly p^2 pairs $(a, b) \in \mathbf{Z}_p^2$. And the elliptic curves over $\mathbf{Z}_p$ are the pairs $(a, b) \in \mathbf{Z}_p^2$ with $4a^3 + 27b^2 \neq 0$. So, it is enough to prove that there are exactly p pairs (a, b) with $4a^3 + 27b^2 = 0$.

For every $c \in \mathbf{Z}_p$, let $a = -3c^2$ and $b = 2c^3$. Then $4a^3 + 27b^2 = 0$. Conversely, for every pair $(a, b) \in \mathbf{Z}_p^2$ with $4a^3 + 27b^2 = 0$, we have:

(i) If $a = 0$, then $b = 0$. Let $c = 0$, then $a = -3c^2$, $b = 2c^3$.

(ii) If $a \neq 0$, then $b \neq 0$, $4a^3 = -27b^2$, $a = -3(-3b/(2a))^2$, and $b = 2(-3b/(2a))^3$. Let $c = -3b/(2a)$, then we also have $a = -3c^2$, $b = 2c^3$.

So a, b can be written as $a = -3c^2$, $b = 2c^3$ for some $c \in \mathbf{Z}_p$ in all cases.

Since c can take exactly p distinct values, there are exactly p pairs (a, b) with $4a^3 + 27b^2 = 0$. Hence, $\#\{E : E(\mathbf{Z}_p)\} = p^2 - p$. □

Chapter 3

Counting elliptic curves

3.1 Isomorphism classes

Definition. Isomorphism of elliptic curves.

Let $E_{a,b}$ and $E'_{a',b'}$ be two elliptic curves. We use an order pair (x, y) to represent a point on an elliptic curve. Let $\mathbf{K}^* = \mathbf{K} - \{0\}$. If there exists a certain unit $u \in \mathbf{K}^*$ such that $a' = u^4 a$, $b' = u^6 b$, and for each point (x, y) on E, the point $(u^2 x, u^3 y)$ is a point on E', then the map $\phi : (x, y) \mapsto (u^2 x, u^3 y)$ is called a $\mathbf{K}$-isomorphism from E to E'.

If E' is $\mathbf{K}$-isomorphic to E, and E' is the same curve as E ($a' = a$, $b' = b$), then we call the map $\phi : E \mapsto E$ an automorphism. The set of all $\mathbf{K}$-automorphisms from E to E is denoted by $\mathrm{Aut}_{\mathbf{K}}E$.

Remark. In the following, we will use a unit $u \in \mathbf{K}^*$ to represent a $\mathbf{K}$-isomorphisms mapping $E_{a,b}$ to some curve $E'_{a',b'}$ with $a' = u^4a$ and $b' = u^6b$, since the representation is unique. If $\mathbf{K}$ is finite, then the number of $\mathbf{K}$-isomorphisms mapping E to a certain curve is $\#\mathbf{K}^*$.

Since we only discuss the situation with $\mathbf{K} = \mathbf{Z}_p$ (p is prime) in this chapter, we just denote $\mathrm{Aut}_{\mathbf{Z}_p}E$ by $\mathrm{Aut}E$.

Lemma 3.1. *AutE can be calculated as follows:*

(i) *If $a, b \neq 0$, then $AutE = \{1, -1\}$ and $\#AutE = 2$.*

(ii) *When $a = 0$, $b \neq 0$, if $p \equiv 1 \pmod 3$, then there exists $\varrho \in \mathbf{Z}_p$ generating AutE and $\#AutE = 6$, else $AutE = \{1, -1\}$ and $\#AutE = 2$.*

(iii) *When $b = 0$, $a \neq 0$, if $p \equiv 1 \pmod 4$, then there exists $i \in \mathbf{Z}_p$ generating AutE and $\#AutE = 4$, else $AutE = \{1, -1\}$ and $\#AutE = 2$.*

Proof. (i) If $a, b \neq 0$, let $u \in \mathrm{Aut}E$. Then $a = u^4a$, $b = u^6b$. And $u^4 = 1$, $u^6 = 1$, since $a, b \neq 0$. So $u^2 = u^6/u^4 = 1$, and $u = \pm 1$.

(ii) If $a = 0$, $b \neq 0$, let g be a generator of $\mathbf{Z}_p^*$. Then $g^{p-1} \equiv 1 \pmod p$.
If $p \equiv 1 \pmod 3$, then $g^{(p-1)/6}$ will be an element of order 6 in $\mathbf{Z}_p^*$. Let $\varrho = g^{(p-1)/6}$, then $a = \varrho^4 a = 0$, $b = \varrho^6 b$. So $\varrho \in \mathrm{Aut}E$, and ϱ, ϱ^2, ϱ^3, ϱ^4, ϱ^5, $\varrho^6 (= 1)$ are all elements in $\mathrm{Aut}E$. Hence ϱ generates $\mathrm{Aut}E$ and $\#\mathrm{Aut}E = 6$.
If $p \equiv 2 \pmod 3$, then $6 \nmid (p-1)$. So there doesn't exist an element of order 6 or 3 in $\mathbf{Z}_p^*$, and $u^6 = 1$ means $u^2 = 1$. Hence $\mathrm{Aut}E = \{1, -1\}$ and $\#\mathrm{Aut}E = 2$.

(iii) If $a \neq 0$, $b = 0$, the proof is similar to (ii). □

Proposition 3.2. *For an elliptic curve E over $\mathbf{Z}_p$, there are $\dfrac{p-1}{\#AutE}$ distinct curves over $\mathbf{Z}_p$ which are $\mathbf{Z}_p$-isomorphic to E.*

Proof. Let $E'_{a',b'}$ be a curve $\mathbf{Z}_p$-isomorphic to $E_{a,b}$, and $\{v_1, \ldots, v_m\}$ are all different $\mathbf{Z}_p$-isomorphisms from E to E'. Then $a = (v_i/v_1)^4 a$, $b = (v_i/v_1)^6 b$, for $i = 1 \ldots m$. So $(v_i/v_1) \in \mathrm{Aut}E$, and $(v_i/v_1) \neq (v_j/v_1)$ if $i \neq j$. That means $m \leq \#\mathrm{Aut}E$.

Let's take any $u \in \mathrm{Aut}E$, then $v_1 u$ must be in $\{v_1, \ldots, v_m\}$. So $\#\mathrm{Aut}E \leq m$. That means $\#\mathrm{Aut}E = m$. A distinct curve E' which is $\mathbf{Z}_p$-isomorphic to E corresponds to $\#\mathrm{Aut}E$ different $\mathbf{Z}_p$-isomorphisms.

Since there are $\#\mathbf{Z}_p^* = p - 1$ different $\mathbf{Z}_p$-isomorphisms mapping E to a certain curve, the number of distinct curves over $\mathbf{Z}_p$ which are $\mathbf{Z}_p$-isomorphic to E is $\dfrac{p-1}{\#\mathrm{Aut}E}$. □

Definition. Isomorphism classes.

It is not difficult to see that isomorphism is an equivalence relation on the set $\{E : E(\mathbf{Z}_p)\}$. Then, we can divide this set into isomorphism classes. The set of $\mathbf{Z}_p$-elliptic curves up to $\mathbf{Z}_p$-isomorphism is denoted by $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$.

Notation. The weighted cardinality.

We denote by $\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$ the weighted cardinality of the set $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$ with weight $(\#\mathrm{Aut}E)^{-1}$, i.e.

$$\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p} = \sum_E \frac{1}{\#\mathrm{Aut}E} ,$$

where E ranges over all elements in the set $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$.

Proposition 3.3. $\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p} = p$.

Proof. By Proposition 3.2, there are $\dfrac{p-1}{\#\mathrm{Aut}E}$ distinct curves $\mathbf{Z}_p$-isomorphic to a curve E. Summing that over $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$, we have

$$\sum_E \frac{p-1}{\#\mathrm{Aut}E} = \#\{E : E(\mathbf{Z}_p)\} .$$

So, by Proposition 2.1,

$$\begin{aligned}\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p} &= \sum_E \frac{1}{\#\mathrm{Aut}E} \\ &= \frac{\#\{E : E(\mathbf{Z}_p)\}}{p-1} = \frac{p^2-p}{p-1} = p.\end{aligned}$$

□

Proposition 3.4. *Suppose that* $p \geq 5$. *Let* $t = \#\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$, *then* $t = 2p+6$, $2p+2$, $2p+4$, $2p$, *for* $p \equiv 1, 5, 7, 11 \pmod{12}$, *respectively.*

Proof. p can be only congruent to $1, 5, 7$ or $11 \pmod{12}$, since p is an odd prime number ≥ 5.

If $p \equiv 11 \pmod{12}$, then $p \not\equiv 1 \pmod 3$, $p \not\equiv 1 \pmod 4$. By Lemma 3.1, we have $\#\mathrm{Aut}E = 2$, for each class in $\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p}$. So,

$$\frac{t}{2} = \sum_E \frac{1}{\#\mathrm{Aut}E} = p \iff t = 2p\ .$$

If $p \equiv 1 \pmod{12}$, then $p \equiv 1 \pmod 3$, $p \equiv 1 \pmod 4$. By Lemma 3.1, we have $p-1$ curves with $a = 0$, $b \neq 0$ and $\#\mathrm{Aut}E = 6$. By Proposition 3.2, we know there are $\frac{p-1}{6}$ distinct curves isomorphic to a curve with $\#\mathrm{Aut}E = 6$. Hence, there are 6 classes with $\#\mathrm{Aut}E = 6$. Similarly, there are 4 classes with $\#\mathrm{Aut}E = 4$. That means there are $t - 10$ classes with $\#\mathrm{Aut}E = 2$. So,

$$\frac{t-10}{2} + \frac{4}{4} + \frac{6}{6} = \sum_E \frac{1}{\#\mathrm{Aut}E} = p \iff t = 2p + 6\ .$$

By the same arguments, we can also prove $t = 2p + 4$ when $p \equiv 7 \pmod{12}$, and $t = 2p + 2$ when $p \equiv 5 \pmod{12}$. □

3.2 Kronecker class numbers

Definition. Discriminant and fundamental discriminant.

An integer Δ is a discriminant, if there exists an order $\mathcal{O}$ of an imaginary quadratic field $\mathbf{K}$ such that $\Delta =$Disc($\mathcal{O}$). If $\Delta =$Disc($\mathcal{O}_\mathbf{K}$) and $\mathcal{O}_\mathbf{K}$ is the maximal order of $\mathbf{K}$, then Δ is called a fundamental discriminant.

Theorem 3.5. *If $\Delta < 0$ is a discriminant, then $\Delta \equiv 0$ or $1 \pmod 4$. If it is also a fundamental discriminant, then $\Delta \equiv 1 \pmod 4$ and Δ is square free, or $\Delta \equiv 0 \pmod 4$ and $\Delta/4$ is square free.* (See [11] , p.32-33)

Definition. Class number.

Suppose that $\mathcal{O}$ is an order of an imaginary quadratic field $\mathbf{K}$. Let A, B be two ideals of $\mathcal{O}$. If there exists $a, b \in \mathcal{O}$ such that $aA = bB$, then we can define an equivalence relation: $A \sim B$. There are only finitely many equivalence classes of ideals under $\sim$, and the number of equivalence classes is called the class number of an order $\mathcal{O}$.

Suppose that Δ is the discriminant of a certain $\mathcal{O}$ which is an order of an imaginary quadratic field $\mathbf{K}$. We denote the class number of $\mathcal{O}$ by $h(\Delta)$, and the number of units in $\mathcal{O}$ by $w(\Delta)$.

Definition. Kronecker class number.

Kronecker class number $H(\Delta)$ is defined by

$$H(\Delta) = \sum_{d} \frac{h(\Delta/d^2)}{w(\Delta/d^2)} ,$$

where the summation ranges over those positive integers d where Δ/d^2 is a discriminant.

There exists f such that $\Delta_0 = \Delta/f^2$ is a fundamental discriminant. f is called the conductor of Δ. The d's are exactly the positive divisors of f.

Example. Suppose that $\Delta = -36$, calculate $H(\Delta)$.

Δ is the discriminant of $\mathcal{O} = \mathbf{Z}[3i] \subseteq \mathbf{Q}[i] = \mathbf{K}$. Then,

$$H(-36) = \frac{h(-36)}{w(-36)} + \frac{h(-4)}{w(-4)} .$$

$\Delta_0 = -4$ is the discriminant of $\mathcal{O}_{\mathbf{K}} = \mathbf{Z}[i]$ (fundamental discriminant).

The conductor of Δ is $f = 3$.

Definition. Legendre symbol.

Suppose that p is a prime number > 2. The Legendre symbol $\left(\frac{d}{p}\right)$ is defined by:

(1) If $p|d$, then $\left(\frac{d}{p}\right) = 0$.

(2) If $p \nmid d$, then

(a) $\left(\frac{d}{p}\right) = 1$, when d is a quadratic residue modulo p;

(b) $\left(\frac{d}{p}\right) = -1$, when d is not a quadratic residue modulo p.

Definition. Jacobi symbol.

Suppose that n is an odd number > 1 and $n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$ with $p_1, \ldots, p_r$ primes. The Jacobi symbol $\left(\frac{d}{n}\right)$ is defined by $\left(\frac{d}{n}\right) = \left(\frac{d}{p_1}\right)^{e_1} \cdots \left(\frac{d}{p_r}\right)^{e_r}$ where $\left(\frac{d}{p_i}\right)$ $(i = 1 \ldots r)$ are Legendre symbols.

Definition. Kronecker symbol.

Suppose that Δ is a discriminant and n is a positive integer. We define the Kronecker symbol $\left(\frac{\Delta}{n}\right)$ by:

(1) If $\gcd(\Delta, n) > 1$, then $\left(\frac{\Delta}{n}\right) = 0$.

(2) If $n = 1$, then $\left(\frac{\Delta}{1}\right) = 1$.

(3) If $n = 2$ and Δ is odd, then $\left(\frac{\Delta}{2}\right)$ is equal to Jacobi symbol $\left(\frac{2}{|\Delta|}\right)$.

(4) If n is odd prime, then $\left(\frac{\Delta}{n}\right)$ is the Legendre symbol.

(5) If $n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$ with $p_1, \ldots, p_r$ prime, then $\left(\frac{\Delta}{n}\right)$ is defined by $\left(\frac{\Delta}{n}\right) = \left(\frac{\Delta}{p_1}\right)^{e_1} \cdots \left(\frac{\Delta}{p_r}\right)^{e_r}$ where $\left(\frac{\Delta}{p_i}\right)$ $(i = 1 \ldots r)$ are Kronecker symbols.

Theorem 3.6. *Kronecker symbol is multiplicative.* $\left(\frac{\Delta}{mn}\right) = \left(\frac{\Delta}{m}\right)\left(\frac{\Delta}{n}\right)$. (See [3] , p.37-40)

The following functions are also multiplicative.

$\chi(n)$: $\chi(n) = \left(\frac{\Delta}{n}\right)$ is the quadratic character associated to Δ.

$\chi_0(n)$: $\chi_0(n) = \left(\frac{\Delta_0}{n}\right)$ is the quadratic character associated to Δ_0.

$\chi_d(n)$: $\chi_d(n) = \left(\frac{\Delta/d^2}{n}\right)$ is the quadratic character associated to Δ/d^2.

Definition. L-series.

Suppose that Δ is a fixed discriminant and χ is associated to Δ. Then the function $L(s, \chi) = \sum_{n=1}^{\infty} n^{-s}\chi(n)$ is called an L-series. Using properties of $\chi(n)$, we can show it converges for $\mathrm{Re}(s) > 1$.

Suppose that n is factorized as $n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$. Since $\chi(n)$ is multiplicative, a term $n^{-s}\chi(n)$ in the L-series can be written as

$$n^{-s}\chi(n) = \left(p_1^{-s}\chi(p_1)\right)^{e_1} \cdot \left(p_2^{-s}\chi(p_2)\right)^{e_2} \cdots \left(p_r^{-s}\chi(p_r)\right)^{e_r} .$$

And it is also a term in the expanded sum formula of the following product

$$\prod_p \left[1 + (p^{-s}\chi(p)) + (p^{-s}\chi(p))^2 + \cdots\right] .$$

Formalizing this argument, we can prove:

Theorem 3.7. *If s is a real number > 1, then L-series $L(s, \chi)$ converges, and it has a form of product (Euler product):*

$$L(s, \chi) = \prod_p \left[1 - p^{-s}\chi(p)\right]^{-1} ,$$

where p ranges over all prime numbers.

Lemma 3.8. *L-series $L(1, \chi)$ can be written as*

$$L(1, \chi) = L(1, \chi_0) \cdot \prod_{l|f} [1 - l^{-1}\chi_0(l)] ,$$

where l ranges over the primes diving f.

Proof. Let s be a real number > 1, Then we have $L(s, \chi) = \prod_p \left[1 - p^{-s}\chi(p)\right]^{-1}$ (by Theorem 3.7).

Suppose that l is a prime number, f is the conductor of Δ, and $\Delta_0 = \Delta/f^2$, then $\chi(l)$ can be calculated as

(i) If $l \nmid f$, then $\chi(l) = \chi_0(l)$, since $\left(\frac{\Delta}{l}\right) = \left(\frac{f^2}{l}\right)\left(\frac{\Delta_0}{l}\right)$, and $\left(\frac{f^2}{l}\right) = 1$;

(ii) If $l|f$, then $\chi(l) = 0$.

Hence we have

$$\begin{aligned} L(s,\chi) &= \prod_{l\nmid f} \left[1 - l^{-s}\chi_0(l)\right]^{-1} \cdot \prod_{l\mid f} \left[1 - l^{-s}\chi(l)\right]^{-1} \\ &= \prod_{l\nmid f} \left[1 - l^{-s}\chi_0(l)\right]^{-1} \\ &= L(s,\chi_0) \cdot \prod_{l\mid f} \left[1 - l^{-s}\chi_0(l)\right] . \end{aligned}$$

Take limit $s \rightarrow 1$ on both sides, then we finish the proof. □

The following theorem is called Dirichlet's class number formula (See [3] , p.49).

Theorem 3.9. *Suppose that $\Delta < 0$ is a discriminant of a certain $\mathcal{O}$ which is an order of an imaginary quadratic field $\mathbf{K}$. Let $h(\Delta)$ be the class number of $\mathcal{O}$, and $w(\Delta)$ be the number of units in $\mathcal{O}$, then*

$$\frac{h(\Delta)}{w(\Delta)} = \frac{\sqrt{-\Delta}}{2\pi} L(1,\chi) .$$

We need the above theorem to prove Lemma 3.10.

Lemma 3.10. *Kronecker class number $H(\Delta)$ has an L-series formula as*

$$H(\Delta) = \frac{\sqrt{-\Delta}}{2\pi} \cdot L(1,\chi_0) \cdot \psi(f) ,$$

where

$$\psi(f) = \sum_{d\mid f} \frac{1}{d} \left[\prod_{l\mid \frac{f}{d}} \left(1 - l^{-1}\chi_0(l)\right) \right] .$$

Proof. Using Dirichlet's class number formula and Lemma 3.8, we have

$$\begin{aligned} H(\Delta) = \sum_{d|f} \frac{h(\Delta/d^2)}{w(\Delta/d^2)} &= \sum_{d|f} \frac{\sqrt{-\Delta/d^2}}{2\pi} \cdot L(1, \chi_d) \\ &= \frac{\sqrt{-\Delta}}{2\pi} \sum_{d|f} \frac{1}{d} \left[L(1, \chi_0) \cdot \prod_{l|\frac{f}{d}} \left(1 - l^{-1}\chi_0(l)\right) \right] \\ &= \frac{\sqrt{-\Delta}}{2\pi} \cdot L(1, \chi_0) \cdot \psi(f) \ . \end{aligned}$$

□

Lemma 3.11. *Suppose that $\psi(f)$ is defined as in previous lemma. If* $\gcd(m, n) = 1$, *then $\psi(mn) = \psi(m)\psi(n)$.*

Proof. Take a term t in the expanded sum formula of $\psi(mn)$. Suppose that

$$t = \frac{1}{d} \prod_{l|\frac{mn}{d}} \left[1 - l^{-1}\chi_0(l)\right] \ ,$$

then there exists d_1, d_2 such that $d = d_1 d_2$, $\gcd(d_1, n) = 1$, $\gcd(m, d_2) = 1$. Let

$$t_1 = \frac{1}{d_1} \prod_{l|\frac{m}{d_1}} \left[1 - l^{-1}\chi_0(l)\right] , \text{ and } t_2 = \frac{1}{d_2} \prod_{l|\frac{n}{d_2}} \left[1 - l^{-1}\chi_0(l)\right] \ ,$$

then t_1 is a term in $\psi(m)$, t_2 is a term in $\psi(n)$, and $t = t_1 t_2$.

Conversely, take a term t_1 in the expanded sum formula of $\psi(m)$, and a term t_2 in the expanded sum formula of $\psi(n)$. Suppose that

$$t_1 = \frac{1}{d_1} \prod_{l|\frac{m}{d_1}} \left[1 - l^{-1}\chi_0(l)\right] , \text{ and } t_2 = \frac{1}{d_2} \prod_{l|\frac{n}{d_2}} \left[1 - l^{-1}\chi_0(l)\right] \ .$$

Let $d = d_1 d_2$, $t = t_1 t_2$, then

$$t = \frac{1}{d} \prod_{l | \frac{mn}{d}} \left[1 - l^{-1}\chi_0(l)\right] ,$$

and t is a term in $\psi(mn)$.

So we have $\psi(mn) = \psi(m)\psi(n)$. □

We denote by $\phi(n)$ the Euler function, which is defined as the number of positive integers s where $s \leq n$ and gcd(s, n)= 1. Then we can give a bound for $\psi(f)$.

Lemma 3.12. $1 \leq \psi(f) \leq (f/\phi(f))^2$.

Proof. Suppose that l is prime, then we can calculate $\psi(l^k)$ $(k > 0)$ as

$$\begin{aligned}
\psi(l^k) &= \sum_{d|l^k} \frac{1}{d} \left[\prod_{p|\frac{l^k}{d}} \left(1 - p^{-1}\chi_0(p)\right) \right] \\
&= \sum_{i=0}^{k-1} \frac{1}{l^i} \left[1 - l^{-1}\chi_0(l)\right] + \frac{1}{l^k} \\
&= \sum_{i=0}^{k} \frac{1}{l^i} - \frac{\chi_0(l)}{l} \sum_{i=0}^{k-1} \frac{1}{l^i} \\
&= \frac{l - l^{-k}}{l-1} - \chi_0(l) \cdot \frac{1 - l^{-k}}{l-1} .
\end{aligned}$$

If $\chi_0(l) = 1$, then $\psi(l^k) = 1$.

If $\chi_0(l) = 0$, then $\psi(l^k) = (l - l^{-k})/(l-1)$.

If $\chi_0(l) = -1$, then $\psi(l^k) = (l - 2l^{-k} + 1)/(l-1)$.

In any case, $1 \leq \psi(l^k) \leq (l+1)/(l-1)$.

From the definition of Euler $\phi(n)$-function, it is easy to see that $\phi(l^k) = l^{k-1}(l-1)$.

Then we have

$$(l^k/\phi(l^k))^2 = l^2/(l-1)^2 ,$$

$$1 \leq \psi(l^k) \leq (l+1)/(l-1) \leq (l^k/\phi(l^k))^2 \,.$$

According to Lemma 3.11, it is also true for any conductor f:

$1 \leq \psi(f) \leq (f/\phi(f))^2$. □

To get an upper bound for $H(\Delta)$, we need the following two theorems. Theorem 3.13 can be found in [7] (Theorem 328), and Theorem 3.14 can be found in [13] (Kapitel IV, Lemma 8.1).

Theorem 3.13. $(f/\phi(f))^2 = O((\log\log f)^2)$.

Theorem 3.14. $L(1,\chi_0) = O(\log|\Delta_0|)$.

An upper bound for $H(\Delta)$ is:

Proposition 3.15. *There exists a positive constant c_1 such that*

$$H(\Delta) \leq \frac{\sqrt{-\Delta}}{2\pi} \cdot L(1,\chi_0) \cdot (f/\phi(f))^2 \leq c_1 \cdot \sqrt{-\Delta} \cdot \log|\Delta| \cdot (\log\log|\Delta|)^2 \,.$$

Proof. From Lemmas 3.10, 3.12, and Theorems 3.13, 3.14, the proposition immediately follows. □

To get a lower bound for $H(\Delta)$, we need the following theorem which can be found in [13] (Kapitel IV, Section 6, Satz 6.6 and the arguments following from Section 8, eq. (8.26)).

Theorem 3.16. *There exists a positive constant c_2 such that for any integer $z > 1$ there exists $\Delta^* < -4$ with the property that*

$$L(1,\chi_0) \geq \frac{c_2}{\log z}, \quad \textit{if } |\Delta_0| \leq z, \textit{ and } \Delta_0 \neq \Delta^* \,.$$

A lower bound for $H(\Delta)$ is:

Proposition 3.17. *There exists a positive constant c_3 such that for each integer $z > 1$ there exists $\Delta^* < -4$ such that*

$$H(\Delta) \geq \frac{\sqrt{-\Delta}}{2\pi} \cdot L(1, \chi_0) \cdot 1 \geq \frac{c_3\sqrt{-\Delta}}{\log z},$$

for all Δ with $|\Delta_0| \leq z$ and $\Delta_0 \neq \Delta^*$.

Proof. From Lemmas 3.10, 3.12, and Theorem 3.16, the proposition immediately follows. □

3.3 Counting isomorphic classes

3.3.1 Distribution of the cardinalities N_p

Theorem 3.18. (Hasse, see [15], p.91)

Let N_p denote the number of points on $E(\mathbf{Z}_p)$, then $p+1-2\sqrt{p} < N_p < p+1+2\sqrt{p}$.

Theorem 3.19. (Deuring, see [4])

Suppose that $p > 3$, p prime, t is an integer with $|t| \leq 2\sqrt{p}$, and $s = p+1-t$, then

$$\#'\{E : N_p = s\}/_{\cong \mathbf{Z}_p} = H(t^2 - 4p),$$

where $H(\delta)$ denotes the Kronecker class number.

Using the above theorems, we can calculate $\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$ as

Lemma 3.20. *Suppose that $\mathcal{S}$ is a set of integers satisfying the hypotheses of Deuring's theorem, then we have*

$$\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p} = \sum_t H(t^2 - 4p) \ ,$$

where the summation ranges over all numbers t with $p + 1 - t \in \mathcal{S}$.

Proof. It immediately follows from Hasse's theorem and Deuring's theorem. □

When $\mathcal{S}$ is a set of integers s with $|s - (p + 1)| \leq 2\sqrt{p}$, we can calculate an upper bound for $\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$ (Proposition 3.21).

Proposition 3.21. *There exists a positive constant c_4 such that*

$$\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p} \leq c_4 \cdot \#\mathcal{S} \cdot \sqrt{p} \cdot \log p \cdot (\log \log p)^2 \ ,$$

where $\mathcal{S}$ is a set of integers s with $|s - (p + 1)| \leq 2\sqrt{p}$.

Proof. Let $\Delta = t^2 - 4p$, then $|\Delta| \leq 4p$. By Proposition 3.15, we have:

$$\begin{aligned} H(t^2 - 4p) &\leq c_1 \cdot \sqrt{4p - t^2} \cdot \log |t^2 - 4p| \cdot (\log \log |t^2 - 4p|)^2 \\ &\leq c_1 \cdot \sqrt{4p} \cdot \log(4p) \cdot (\log \log(4p))^2 \ . \end{aligned}$$

From Lemma 3.20, the proposition immediately follows. □

When $\mathcal{S}$ is a set of integers s with $|s - (p + 1)| \leq \sqrt{p}$, we can calculate a lower bound for $\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$ (Proposition 3.22).

Proposition 3.22. *There exists a positive constant c_5 such that*

$$\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p} \geq c_5 \cdot (\#\mathcal{S} - 2) \cdot \sqrt{p}/(\log p) \ ,$$

where $\mathcal{S}$ is a set of integers s with $|s - (p+1)| \leq \sqrt{p}$.

Proof. Let $\Delta = t^2 - 4p$, then $|\Delta| \geq 3p$. Let $z = 4p$, then $|\Delta_0| \leq z$. By Proposition 3.17, we know that there exists $\Delta^* < -4$ such that

$$H(t^2 - 4p) \geq c_3\sqrt{4p - t^2}/(\log 4p) \geq c_3\sqrt{3p}/(\log 4p) \ ,$$

except when $\Delta_0 = \Delta^*$.

To make the inequality valid, we need to exclude the elements in $\mathcal{S}$ with $\Delta_0 = \Delta^*$. Suppose that s is such an element, and $t = p + 1 - s$, $\Delta = t^2 - 4p$, $\Delta_0 = \Delta^*$.

Let $\alpha, \bar{\alpha}$ be the 2 roots of equation $X^2 - tX + p = 0$, then they will be in the ring of integer of $Q[\sqrt{\Delta}] = Q[\sqrt{\Delta^*}]$. $\alpha, \bar{\alpha}$ should be irreducible since $p = \alpha\bar{\alpha}$. They can be determined up to conjugation and sign, since $\Delta^* < -4$ and the only units are $\{1, -1\}$. Hence $t = \alpha + \bar{\alpha}$ is determined up to sign. There are at most two integers t with $\Delta_0 = \Delta^*$. And there are at least $(\#\mathcal{S} - 2)$ elements in the set $\mathcal{S}$ which can make the inequality valid.

From Lemma 3.20, the proposition immediately follows. □

3.3.2 Divisibility of the cardinalities N_p

Let l be a fixed prime. We want to estimate the number of curves such that $N_p \not\equiv 0 \pmod{l}$. The following two theorems are proven in [10] using modular curves.

Theorem 3.23. $\#'\{E : N_p \equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} = p/(l-1) + O(l\sqrt{p})$, *when l prime and $p \not\equiv 1 \pmod{l}$.*

Theorem 3.24. $\#'\{E : N_p \equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} = p \cdot l/(l^2-1) + O(l\sqrt{p})$, *when l prime and $p \equiv 1 \pmod{l}$.*

Using the above theorems, we can give bounds for $\#'\{E : N_p \not\equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p}$ (Proposition 3.25).

Proposition 3.25. *There exists a positive constant c_6 such that*

$$\#'\{E : N_p \not\equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} \geq c_6 p\,.$$

Since $\#'\{E : E(\mathbf{Z}_p)\}/_{\cong \mathbf{Z}_p} = p$, we can also change the above formula into the following form: There exists a positive constant c_6' such that

$$\#'\{E : N_p \equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} \leq c_6' p\,.$$

Proof. We need to consider the following 3 situations:

(i) $l \leq c_7\sqrt{p}$ for a suitable positive constant c_7.

If $p \equiv 1 \pmod{l}$, then $l \geq 2$, and the coefficient of p is $l/(l^2-1) \leq 2/3$.

If $p \not\equiv 1 \pmod{l}$, then $l \geq 3$, and the coefficient of p is $1/(l-1) < 2/3$.

When l increases, the coefficient of p decreases. But l is bounded. So the coefficient of p should be bounded in a range. We can say, there exists a positive constant c_6' such that $\#'\{E : N_p \equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} \leq c_6' p$.

(ii) $p \geq c_8$ and $l \geq c_9(\log p)(\log\log p)^2$ for suitable positive constants c_8, c_9.

Let $\mathcal{S}$ be the set of integers s with $|s - (p+1)| \leq 2\sqrt{p}$ and $s \equiv 0 \pmod{l}$. Since $0 \in \mathcal{S}$, it contains at least one element. And the number of other elements should be $[4\sqrt{p}/l]$. So, the cardinality of $\mathcal{S}$ is $1 + [4\sqrt{p}/l]$. If $\sqrt{p}/l > O(1)$, then $\#\mathcal{S} = O(\sqrt{p}/l)$. If $\sqrt{p}/l = O(1)$, then $\#\mathcal{S} = O(1)$. So, in any situation, we have $\#\mathcal{S} = O(\sqrt{p}/l)$.

Applying Proposition 3.21, we have:

$\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p} \leq c_4 \cdot \#\mathcal{S} \cdot \sqrt{p}(\log p)(\log\log p)^2$

$\leq c_4 \cdot O(\sqrt{p}/l) \cdot \sqrt{p} \cdot l/c_9 \leq O(p) \cdot c_4/c_9$.

That means there exists a positive constant c_6' such that

$\#'\{E : N_p \equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} \leq c_6' p$.

(iii) the remaining cases, $p < c_8$ or $c_7\sqrt{p} < l < c_9(\log p)(\log\log p)^2$.

If p is fixed, we have:

$\#'\{E : N_p = p\}/_{\cong \mathbf{Z}_p} = H(1 - 4p) > 0$, and

$\#'\{E : N_p = p + 1\}/_{\cong \mathbf{Z}_p} = H(-4p) > 0$, by Deuring's formula.

So there are elliptic curves E_1, E_2 over $\mathbf{Z}_p$ with $\#E_1(\mathbf{Z}_p) = p$, $\#E_2(\mathbf{Z}_p) = p+1$. Since p, $p+1$ cannot be multiples of l at the same time, we have

$\#'\{E : N_p \not\equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} > 0$.

But p is bounded. So, we can see that there exists a positive constant c_6 such that $\#'\{E : N_p \not\equiv 0 \pmod{l}\}/_{\cong \mathbf{Z}_p} \geq c_6 p$. □

3.4 Counting triples (a, x, y) in $\mathbf{Z}_p^3$

Definition. Suppose that $\mathcal{S}$ is a set of integers. We denote by $M_{\mathcal{S}}$ the number of triples (a, x, y) in $\mathbf{Z}_p^3$, where $b = y^2 - x^3 - ax$, $4a^3 + 27b^2 \neq 0$, and $N_p \in \mathcal{S}$.

Clearly, $M_{\mathcal{S}}$ is equal to the number of quadruples $(a, b, x, y) \in \mathbf{Z}_p^4$ where (x, y) is a point on elliptic curve $E_{a,b}(\mathbf{Z}_p)$ and $N_p \in \mathcal{S}$.

For each isomorphism class E in the set $\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$, it includes $\dfrac{p-1}{\#\mathrm{Aut}E}$ distinct curves. And each curve has $\#E - 1$ points distinct from the point at infinity. So, one class corresponds to $\dfrac{(p-1)(\#E-1)}{\#\mathrm{Aut}E}$ quadruples (a, b, x, y). Thus $M_{\mathcal{S}}$ can be estimated as

$$\begin{aligned} M_{\mathcal{S}} &= \sum_E \frac{(p-1)(\#E-1)}{\#\mathrm{Aut}E} \\ &\geq (p-1)(p-2\sqrt{p}) \sum_E \frac{1}{\#\mathrm{Aut}E} \\ &\gg p^2 \cdot \#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}\ , \end{aligned}$$

where the sum ranges over those isomorphic classes with $N_p \in \mathcal{S}$.

Proposition 3.26. *There exists a positive constant c_{10} such that*

$$M_{\mathcal{S}} \geq c_{10} \cdot (\#\mathcal{S} - 2) \cdot p^{5/2}/(\log p),$$

where $\mathcal{S}$ is a set of integers s with $|s - (p+1)| \leq \sqrt{p}$.

Proof. We have $M_{\mathcal{S}} \gg p^2 \cdot \#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$, and

$$\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p} \geq c_5 \cdot (\#\mathcal{S} - 2) \cdot \sqrt{p}/(\log p)$$

(by Proposition 3.22). So the proposition immediately follows. □

Proposition 3.27. *Let l be a fixed prime. There exists a positive constant c_{11} such that*

$$M_{\mathcal{S}} \geq c_{11} p^3,$$

where $\mathcal{S}$ is a set of integers s with $s \not\equiv 0 \pmod{l}$.

Proof. We have $M_{\mathcal{S}} \gg p^2 \cdot \#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p}$, and

$$\#'\{E : N_p \in \mathcal{S}\}/_{\cong \mathbf{Z}_p} \geq c_6 p$$

(by Proposition 3.25). So the proposition immediately follows. □

Chapter 4

Lenstra's algorithm

4.1 A conjecture about smooth numbers

Suppose that $n_i(x)$ denotes the i^{th} largest prime factor of integer x. Suppose that $\alpha > 1$. The symbol $\Psi_i(x, x^{1/\alpha})$ denotes the number of integers $s \in [1, x]$ such that $n_i(s) \leq x^{1/\alpha}$. We define

$$\rho_i(\alpha) = \lim_{x \to \infty} \frac{\Psi_i(x, x^{1/\alpha})}{x} \, ,$$

which is the probability that an integer between 1 and x has its i^{th} largest prime factor $\leq x^{1/\alpha}$.

Definition. $\langle w \rangle$-smooth numbers.

Let x be an integer. Fix a bound w. If x doesn't have prime factor larger than w, then we say x is a $\langle w \rangle$-smooth number.

Let the function $\rho(\alpha) = \rho_1(\alpha)$. Let $x > 0$ be a real number. Then the function

$\rho(\alpha)$ is concerned with the distribution of $\langle x^{1/\alpha}\rangle$-smooth integers in the interval $[1, x]$. The following theorem [2] gives an estimation of $\rho(\alpha)$.

Theorem 4.1. *If $\varepsilon > 0$ is arbitrary and $3 \leq \alpha \leq (1-\varepsilon)\log x/\log\log x$, then*

$$\rho(\alpha) = \exp\{-\alpha(\log\alpha + o(\log\alpha))\}$$

as $x \longrightarrow \infty$.

Inspired from the above theorem, Lenstra has given the following conjecture in [10].

Conjecture 4.2. *Let s be a random integer in the interval $(p+1-\sqrt{p}, p+1+\sqrt{p})$. Then the probability that s has all its prime factors $\leq p^{1/\alpha}$ is approximately $\rho(\alpha)$, when $p \longrightarrow \infty$.*

Suppose that $L(p) = \exp\left(\sqrt{\log p \log\log p}\right)$, and u is a positive number such that $L(p)^u = p^{1/\alpha}$. If we use $L(p)^u$ to replace $p^{1/\alpha}$ in the above formula, then we have

$$\alpha = \frac{1}{u}\sqrt{\frac{\log p}{\log\log p}}\,,$$

$$\alpha(\log\alpha + o(\log\alpha)) = (1/(2u) + o(1))\sqrt{\log p \log\log p}\,,$$

$$\rho(\alpha) = L(p)^{-1/(2u)+o(1)}\,.$$

From the above result, we can get the following conjecture:

Conjecture 4.3. *Let s be a random integer in the interval $(p+1-\sqrt{p}, p+1+\sqrt{p})$. Then the probability that s has all its prime factors $\leq L(p)^u$ is approximately $L(p)^{-1/(2u)+o(1)}$, when $p \longrightarrow \infty$.*

4.2 Pseudo-group V_N

Definition. Pseudo-group, pseudo-add.

Suppose that G is a set and "+" is the operation on G. Let P, Q are two elements in G. If $P+Q$ cannot always give an element $R = P+Q$ in G, then we say G is a pseudo-group and "+" is called pseudo-add.

Suppose that $[X, Y, Z]$ is a point on $E(\mathbf{Z}_N)$. If $Z = 0$, then $[X, Y, Z] = [0, Y, 0]$ cannot be always changed into $[0, 1, 0]$, since Y is not necessarily invertible. So $E(\mathbf{Z}_N)$ includes more than one points with $Z = 0$. For the same reason, if $Z \neq 0$, then $[X, Y, Z]$ cannot be always changed into $[x, y, 1]$, and $E(\mathbf{Z}_N)$ includes points with $Z \neq 0, Z \neq 1$. Hence we just take a subset of $E(\mathbf{Z}_N)$ as

$$V_N = \{[x, y, 1] \in E : x, y \in \mathbf{Z}_N, y^2 = x^3 + ax + b\} \cup \{O = [0, 1, 0]\} .$$

Remark. With the above definiton, we can use an ordered pair (x, y) to represent a point $[x, y, 1]$ in V_N, and give the following pseudo-group law on V_N, similar to the group law on $E(\mathbf{Z}_p)$.

Rule. **Pseudo-group law on V_N.**

Suppose that $P, Q \in V_N$, $R = P + Q$.

1. If $P = O$, then $R = Q$. If $Q = O$, then $R = P$.

2. If $P, Q \neq O$, then suppose that $P = (x_1, y_1)$, $Q = (x_2, y_2)$. We have the following possibilities (1)-(5):

(1) Compute $d_1 = \gcd(x_1 - x_2, N)$. If $1 < d_1 < N$, then stop and give a non-trivial factor of N.

(2) If $d_1 = 1$, let $R = (x_3, y_3)$, and (x_3, y_3) is calculated as

$x_3 = \lambda^2 - x_1 - x_2$, $y_3 = \lambda(x_1 - x_3) - y_1$, where $\lambda = (y_2 - y_1)/(x_2 - x_1)$.

(3) If $d_1 = N$, then compute $d_2 = \gcd(y_1 + y_2, N)$. If $1 < d_2 < N$, then stop and give a non-trivial factor of N.

(4) If $d_2 = 1$, let $R = (x_3, y_3)$, and (x_3, y_3) is calculated as

$x_3 = \lambda^2 - x_1 - x_2$, $y_3 = \lambda(x_1 - x_3) - y_1$, where $\lambda = (3x_1^2 + a)/(y_1 + y_2)$.

(5) If $d_2 = N$, let $R = O$.

Adding two points is not always successful. It either gives a non-trivial factor of N, or gives a point $R = P + Q \in V_N$. Therefore V_N is a pseudo-group and "+" is pseudo-add. Lenstra's method [10] just uses the property of the pseudo-addition on V_N to factorize N.

4.3 Basic idea

Let p be the smallest prime factor of N. Suppose that w and w' are two fixed bounds, and k is a $\langle w \rangle$-smooth number with the additional property

$$k = \prod p_i^{e_i} \text{ , with } p_i \text{ prime, } p_i \leq w \text{ , } p_i^{e_i} \leq w' = p + 1 + 2\sqrt{p} \, .$$

We randomly choose an elliptic curve $E(\mathbf{Z}_N)$: $y^2 z = x^3 + axz^2 + bz^3 (a, b \in \mathbf{Z}_N)$, take a random point $P = (x, y)$ on V_N, and compute kP (add P for k times). If N_p is also a $\langle w \rangle$-smooth number and $N_p | k$, then the computation is likely to fail and consequently N will be factorized (see the proposition below).

The above action is called a *TRIAL* in Lenstra's algorithm.

Notation. Suppose that point $P = [X, Y, Z]$, then $P_p = [X \pmod p), Y \pmod p), Z$ $\pmod p)]$ is called the reduction of P modulo p. And $\operatorname{ord}(P_p)$ is the order of the point P_p in the group $E(\mathbf{Z}_p)$.

The following proposition exhibits precise conditions under which the algorithm will be successful.

Proposition 4.4. *Let k be as before. Suppose that N, p, q, a, x, y, b, k satisfy conditions* (i) *and* (ii)*:*

(i) *Let p $(p \geq 5)$ be the smallest prime factor of N. Let $\bar{a} = a \pmod p$, and $\bar{b} = b \pmod p$ be such that $4\bar{a}^3 + 27\bar{b}^2 \not\equiv 0 \pmod p$. Suppose that N_p is a $\langle w \rangle$-smooth number, and that $N_p | k$.*

(ii) *Let q $(q \neq p)$ be another prime factor of N. Let $\hat{a} = a \pmod q$, and $\hat{b} = b \pmod q$ be such that $4\hat{a}^3 + 27\hat{b}^2 \not\equiv 0 \pmod q$. Suppose that l is the largest prime number dividing ord(P_p), which is the order of P_p, and that $l \nmid \#E(\mathbf{Z}_q)$.*

Then computing kP is successful in finding a non-trivial divisor of N.

Proof. Suppose that $e(l)$ is the largest integer such that $l^{e(l)} | \operatorname{ord}(P_p)$. Since $\operatorname{ord}(P_p) | N_p$, and $N_p | k$, we have that $l \leq w$, and $l^{e(l)} \leq w'$.

Put $k_0 = (\prod p_i^{e_i}) \cdot l^{e(l)-1}$, with p_i prime, $p_i \leq l - 1$, $p_i^{e_i} \leq w'$. From the definition of k_0 and k, we see that $k_0 | k$ and $k_0 l | k$. If kP is successfully calculated, then it is also true for $k_0 P$ and $k_0 l P$.

Now it is enough to prove that k_0P and k_0lP cannot both be defined. And consequently kP cannot be defined.

Assume that k_0P and k_0lP can both be defined. From the definition of k_0, it follows that $k_0 \not\equiv 0 \pmod{\mathrm{ord}(P_p)}$ and $k_0l \equiv 0 \pmod{\mathrm{ord}(P_p)}$, and then

$$k_0P_p \neq O_p, \text{ but } k_0lP_p = O_p.$$

We assume that k_0lP exists. So

$$(k_0lP)_p = O_p \Longrightarrow k_0lP = O \Longrightarrow (k_0lP)_q = O_q \Longrightarrow k_0l \cdot P_q = O_q\ .$$

Since $l \nmid \#E(\mathbf{Z}_q)$, we must have $k_0P_q = O_q$.

We assume that k_0P exists. So

$$(k_0P)_q = O_q \Longrightarrow k_0P = O \Longrightarrow (k_0P)_p = O_p \Longrightarrow k_0P_p = O_p\ .$$

However, it contradicts what we proved before. Therefore, k_0P and k_0lP cannot both be defined, and the calculation of kP fails. The algorithm will report success. □

4.4 Algorithm analysis

4.4.1 Success probability

Suppose that N, p, q, w are defined as in previous section.

Definition. The set $S(w)$.

$S(w)$ is the set of integers s where s is a $\langle w\rangle$-smooth number, and $|s-(p+1)| < \sqrt{p}$. If $p \geq 5$, then $\#S(w) \geq 3$.

Definition. (Probability $f(w)$.)

Let $f(w)$ be the probability that a random integer in the interval $(p+1-\sqrt{p}, p+1+\sqrt{p})$ has all its prime factors $\leq w$. Then $f(w) = \dfrac{\#S(w)}{2[\sqrt{p}]+1}$.

Definition. The set T_s.

Fix $s \in S(w)$. Then T_s is the set of triples $(\alpha, \xi, \eta) \in \mathbf{Z}_p^3$ with $\#E_{\alpha,\beta}(\mathbf{Z}_p) = s$, and $4\alpha^3 + 27\beta^2 \neq 0$, where $\beta = \eta^2 - \xi^3 - \alpha\xi$.

Definition. The set $U_{\alpha\xi\eta}$.

Fix $s \in S(w)$ and $(\alpha, \xi, \eta) \in T_s$. Let $l_{\alpha\xi\eta}$ denote the largest prime divisor of the order of the point $[\xi, \eta, 1] \in E_{\alpha,\beta}(\mathbf{Z}_p)$. Then $U_{\alpha\xi\eta}$ is the set of triples $(\alpha', \xi', \eta') \in \mathbf{Z}_q^3$ with $l_{\alpha\xi\eta} \nmid \#E_{\alpha',\beta'}(\mathbf{Z}_q)$, and $4\alpha'^3 + 27\beta'^2 \neq 0$, where $\beta' = \eta'^2 - \xi'^3 - \alpha'\xi'$.

Definition. The set $V_{\alpha\xi\eta\alpha'\xi'\eta'}$.

Fix $s \in S(w)$, $(\alpha, \xi, \eta) \in T_s$ and $(\alpha', \xi', \eta') \in U_{\alpha\xi\eta}$. Then $V_{\alpha\xi\eta\alpha'\xi'\eta'}$ is the set of triples $(a, x, y) \in \mathbf{Z}_N^3$ where $(a \pmod p), x \pmod p), y \pmod p)) = (\alpha, \xi, \eta)$, and $(a \pmod q), x \pmod q), y \pmod q)) = (\alpha', \xi', \eta')$.

In practice, we randomly choose a triple $(a, x, y) \in \mathbf{Z}_N^3$ in every trial, and calculate $b = y^2 - x^3 - ax$. This gives the curve $E_{a,b}(\mathbf{Z}_N)$ with a point $P = (x, y)$ on E. Suppose that M is the number of triples $(a, x, y) \in \mathbf{Z}_N^3$ which satisfy conditions (i) and (ii). The following proposition gives an estimation of M/N^3.

Proposition 4.5. *There exists a positive constant c_{12} such that*

$$\frac{M}{N^3} \geq \frac{c_{12} f(w)}{\log p} .$$

Proof. Take a triple $(a, x, y) \in \mathbf{Z}_N^3$ in a certain set $V_{\alpha\xi\eta\alpha'\xi'\eta'}$. Clearly (a, x, y) satisfies conditions (i) and (ii). Then we have:

$$M \geq \sum_{s \in S(w)} \sum_{(\alpha,\xi,\eta) \in T_s} \sum_{(\alpha',\xi',\eta') \in U_{\alpha\xi\eta}} \#V_{\alpha\xi\eta\alpha'\xi'\eta'} .$$

Since each set $V_{\alpha\xi\eta\alpha'\xi'\eta'}$ has cardinality $N^3/(pq)^3$, we have

$$\frac{M}{N^3} \geq \frac{1}{p^3 q^3} \sum_{s \in S(w)} \sum_{(\alpha,\xi,\eta) \in T_s} \#U_{\alpha\xi\eta} .$$

By Proposition 3.27 and the definition of $U_{\alpha\xi\eta}$, we have

$\#U_{\alpha\xi\eta} \geq c_{11} q^3$, and

$$\frac{M}{N^3} \geq \frac{c_{11}}{p^3} \sum_{s \in S(w)} \#T_s .$$

By Proposition 3.26 and the definition of T_s, we have

$\sum_{s \in S(w)} \#T_s \geq c_{10}(\#S(w) - 2)p^{5/2}/\log p$, and

$$\frac{M}{N^3} \geq \frac{c_{10}c_{11}}{\log p} \cdot \frac{\#S(w) - 2}{\sqrt{p}} \geq \frac{c_{10}c_{11}}{\log p} \cdot \frac{\#S(w) - 2}{2[\sqrt{p}] + 1} .$$

Since $\#S(w) \geq 3$, we have $\#S(w) - 2 \geq \#S(w)/3$.

Thus, the proposition immediately follows. □

From the above theorem, we know the success probability of performing one trial is $M/N^3 \gg f(w)/\log p$. Hence we just let $h = \dfrac{\log p}{f(w)}$ be the number of trials in Lenstra's algorithm.

4.4.2 The optimal running time

Definition. The unit of running time.

In this section, a *unit of running time* is the time of performing a pseudo-addition on V_N. An inverse operation modulo N ($T^{-1} \pmod{N}$) needs $\mathrm{O}(\log^3 N)$ bit operations (See [8] , p.18). Since the dominating time of a pseudo-addition on V_N is one or two inverse operations modulo N, a unit of running time corresponds to $\mathrm{O}(\log^3 N)$ bit operations.

Using the repeated squaring algorithm, we can compute kP in $\mathrm{O}(\log k)$ units of running time, and this is the dominating time of one trial. So, the total time of Lenstra's algorithm is $\mathrm{O}(h \log k)$.

Since $k = \prod p_i^{e_i}$, with p_i prime, $p_i \leq w$, $p_i^{e_i} \leq w' = p + 1 + 2\sqrt{p}$, we have

$$\log k \ll w \log (p_i^{e_i}) \ll w \log p \, .$$

Since $h = \dfrac{\log p}{f(w)}$, we have $h \log k \ll \dfrac{w \log^2 p}{f(w)}$. If we want to get the optimal running time, we can choose a proper value of w such that $w/f(w)$ is minimal. And we need the conjecture in section 4.1 .

Suppose that N, p are as defined in previous section. Replacing $w = L(p)^u$ into the formula of $f(w)$, we have $f(L(p)^u) = L(p)^{-1/(2u)+o(1)}$, for $p \to \infty$. Then,

$$\frac{w}{f(w)} = \frac{L(p)^u}{L(p)^{-1/(2u)+o(1)}} = L(p)^{1/(2u)+u+o(1)} \, .$$

If $w/f(w)$ is minimal, then $1/(2u) + u$ should be minimal, and

$$\left(\frac{1}{2u} + u\right)' = 0 \iff \frac{-1}{2u^2} + 1 = 0 \iff u = 1/\sqrt{2}.$$

Hence, the conjectural optimal choice for w and $1/f(w)$ are

$$w = L(p)^{1/\sqrt{2}+o(1)} \text{ , and } 1/f(w) = L(p)^{1/\sqrt{2}+o(1)} \text{ ,}$$

for $p \to \infty$. Let $w = p^{1/\alpha}$, then $\rho(\alpha) = L(p)^{-1/(2u)+o(1)} = L(p)^{-1/\sqrt{2}+o(1)}$. Since $h = \dfrac{\log p}{f(w)} = L(p)^{1/\sqrt{2}+o(1)}$, we can use

$$h = 1/\rho(\alpha)$$

as the parameter in our implementation.

Under the conjecture, we get the optimal running time of Lenstra's algorithm as

$$h \log k = \mathrm{O}\left(w \log p/\rho(\alpha)\right) = L(p)^{\sqrt{2}+o(1)}.$$

Chapter 5

Brent's implementation

5.1 A conjecture about smooth numbers

Suppose that $\alpha > 1$ and $\beta \in [1, \alpha]$. The symbol $\Psi(x, x^{1/\alpha}, x^{\beta/\alpha})$ denotes the number of integers $s \in [1, x]$ such that $n_1(s) \leq x^{\beta/\alpha}$ and $n_2(s) \leq x^{1/\alpha}$. We define

$$\mu(\alpha, \beta) = \lim_{x \to \infty} \frac{\Psi(x, x^{1/\alpha}, x^{\beta/\alpha})}{x},$$

which is the probability that an integer between 1 and x has its first largest prime factor $\leq x^{\beta/\alpha}$, and its second largest prime factor $\leq x^{1/\alpha}$.

Definition. $\langle w, w^{\beta} \rangle$-smooth numbers.

Let x be an integer. Fix a bound w. If $w = x^{1/\alpha}$ for some $\alpha > 1$, then let $\beta \in [1, \alpha]$. If the second largest prime factor of x is $\leq w$, and the first largest prime factor of x is $\leq w^{\beta}$, then x is called a $\langle w, w^{\beta} \rangle$-smooth number.

Let $x > 0$ be a real number. Then the function $\mu(\alpha, \beta)$ is concerned with the distribution of $\langle x^{1/\alpha}, x^{\beta/\alpha}\rangle$-smooth integers in the interval $[1, x]$. The following theorems, proven in [9], gives some relations between ρ and μ.

Theorem 5.1. *For $\beta = \alpha$, we have*

$$\mu(\alpha,\alpha) - \rho(\alpha) = \int_0^{\alpha-1} \frac{\rho(t)}{\alpha - t} dt ,$$

or equivalently under the change of variable $t = (1-u)\alpha$,

$$\mu(\alpha,\alpha) - \rho(\alpha) = \int_{1/\alpha}^{1} \frac{\rho((1-u)\alpha)}{u} du .$$

Theorem 5.2. *For $\beta = 2$, we have*

$$\mu(\alpha,2) - \rho(\alpha) = \int_{1/\alpha}^{2/\alpha} \frac{\rho((1-u)\alpha)}{u} du .$$

Theorem 5.3. *For $1 \le \beta \le \alpha$, we have*

$$\mu(\alpha,\beta) - \rho(\alpha) = \int_{1/\alpha}^{\beta/\alpha} \frac{\rho((1-u)\alpha)}{u} du = \int_{\alpha-\beta}^{\alpha-1} \frac{\rho(t)}{\alpha - t} dt .$$

From Theorem 5.3 , we get

Corollary 5.4. *Suppose that β is fixed, and $1 \le \beta \le 2$, then*

$$\mu(\alpha,\beta)/\rho(\alpha) = O(\alpha(\alpha \log \alpha)^{\beta-1}) \text{ , when } \alpha \to \infty .$$

Proof. Using Theorem 5.3 , we have

$$\begin{aligned} \mu(\alpha,\beta) &= \rho(\alpha) + \int_{\alpha-\beta}^{\alpha-1} \frac{\rho(t)}{\alpha-t} dt \\ &\ge \rho(\alpha) + \frac{1}{\beta} \int_{\alpha-\beta}^{\alpha-1} \rho(t) dt \\ &\ge \rho(\alpha) + \frac{1}{\beta} \int_{\alpha-\beta}^{\alpha-\beta+1} \rho(t) dt - \frac{1}{\beta} \int_{\alpha-1}^{\alpha} \rho(t) dt . \end{aligned}$$

Since $\alpha\rho(\alpha) = \int_{\alpha-1}^{\alpha} \rho(t)dt$, when $\alpha > 1$ (see [6]), we have

$$\mu(\alpha,\beta) \geq \rho(\alpha) + \frac{\alpha-\beta+1}{\beta}\rho(\alpha-\beta+1) - \frac{\alpha}{\beta}\rho(\alpha) .$$

Since $\dfrac{\rho(\alpha-\beta)}{\rho(\alpha)} = (\alpha\log\alpha)^{\beta}(1+o(1))$, when $0 \leq \beta \leq \alpha$ (see [6]), we have

$$\mu(\alpha,\beta) \geq \rho(\alpha) + \frac{\alpha-\beta+1}{\beta}(\alpha\log\alpha)^{\beta-1}\rho(\alpha) - \frac{\alpha}{\beta}\rho(\alpha) .$$

and the proposition immediately follows. □

As we did in chapter 4, we will use $\mu(\alpha,\beta)$ to measure the probability that N_p is $\langle p^{1/\alpha}, p^{\beta/\alpha}\rangle$-smooth. The following conjecture is a precise reformulation of the ideas of Brent [1].

Conjecture 5.5. *Let s be a random integer in the interval $(p+1-\sqrt{p}, p+1+\sqrt{p})$. Then the probability that s is a $\langle p^{1/\alpha}, p^{\beta/\alpha}\rangle$-smooth integer is $\mu(\alpha,\beta)$, when $p \to \infty$.*

5.2 Birthday Paradox

What is the probability that at least two people in a group of r people will have the same birthday? This problem is called "Birthday Paradox". It can be rephrased as the probability of picking the same elements twice when choosing r elements in a set of t elements.

Notation. PE(r,t). Let PE(r,t) denote the probability that no element is picked twice, when picking r elements in a set of t elements.

Proposition 5.6. *$PE(r,t) = (1-t^{-1})(1-2t^{-1})\dots[1-(r-1)t^{-1}]$.*

Proof. Suppose that $A_1, \dots, A_r$ are the r elements and $A_i \neq A_j$ if $i \neq j$.

For the element A_1, there are t possible choices.

Fix A_1. For the element A_2, there are $t-1$ possible choices, since $A_2 \neq A_1$.

Similarly, if we fix $A_1, \dots, A_s$ $(s < r)$, then there are $t-s+1$ possible choices for the element A_{s+1}. Hence, for elements $A_1, \dots, A_r$, there are $t(t-1)(t-2)\dots(t-r+1)$ possible choices. Then

$$PE(r,t) = \frac{t(t-1)(t-2)\dots(t-r+1)}{t^r} \, .$$

The proposition immediately follows. □

Example. If $t = 365$, and $r = 23$, then the probability of at least 2 people among r people having the same birthday is $1-\text{PE}(23,365) \geq 1/2$.

Remark. $\text{PE}(r,t)$ has an approximation $e^{-r(r-1)/2t}$ (See [5], p.33).

Proposition 5.7. *Suppose that* $B = e^{-r(r-1)/2t}$, *then*

$$0 < \frac{B - PE(r,t)}{B} < \epsilon \, , \; with \, \epsilon < \frac{r^3}{6(t-r+1)^2} \, .$$

5.3 Basic idea

Brent's method [1] is based on Lenstra's algorithm, but it uses a two-phase way to implement it. In Lenstra's algorithm, a trial means choosing a random curve with a point on it and computing kP. But in Brent's implementation, a trial has two phases and that is the first phase. The second phase will be performed when the first fails. If both phases fail, then the trial is declared to fail.

Definition. Equivalence.

Suppose that $R, S \neq O$ are two points in V_N with $R = (x_1, y_1)$ and $S = (x_2, y_2)$. Let p be a prime dividing N, R_p be the reduction of R modulo p, and S_p be the reduction of S modulo p. If $R_p = \pm S_p$, then we say that R is p-equivalent to S, denoted by $R \simeq_p S$. If R and S are in the same equivalence class, then we can factorize N by computing $\gcd(N, (x_1 - x_2))$.

Suppose that the first phase of a trial fails in Brent's method, then we have a point $Q = kP$. Suppose that Q_p is the reduction of Q modulo p, and ord(Q_p) is the order of Q_p. We notice that ord(Q_p) =ord(P_p)/gcd(k,ord(P_p)). clearly, Q_p is likely a point of smaller order than average. That means it is relatively easy to find two equivalent points from some multiples of Q, since there are $t = [\text{ord}(Q_p)/2]$ equivalence classes in the set of all multiples of Q. The idea of Brent's method is just based on this fact.

The second phase of a trial in Brent's method is to randomly generate r multiples of Q, say $q_1Q, \ldots, q_rQ$ for some positive integers q_i. If there exists two points q_iQ and q_jQ such that $q_iQ \simeq q_jQ$ $(1 \leq i < j \leq r)$, then N can be factorized.

Using the answer of the "Birthday Paradox" problem, we have: the probability that there exists two equivalent points q_iQ and q_jQ $(1 \leq i < j \leq r)$ out of r randomly chosen points q_1Q, ..., q_rQ in all points with t equivalence classes, is equal to 1−PE(r, t).

Suppose that E is the curve which we randomly choose in the first phase, and w, k are fixed parameters. Suppose that $w = p^{1/\alpha}$ and $1 \leq \beta \leq \alpha$. Let $r = [(w^\beta \log 2)^{1/2}] + 1$. Then the success conditions are:

Theorem 5.8. *If N_p is a $\langle w, w^\beta \rangle$-smooth number, then Brent's implementation will be successful with a probability higher than $1/2$.*

Proof. Let $N'_p = \dfrac{N_p}{n_1(N_p)}$. Since N_p is $\langle w, w^\beta \rangle$-smooth, we have N'_p is $\langle w \rangle$-smooth. From the definition of k, we know $N'_p | k$.

Since $N_p \nmid k$, $Q = kP$ can be calculated, and the first phase fails. But $N_p | (k \cdot n_1(N_p))$, therefore $n_1(N_p)$ is the order of Q_p.

Since $2t \leq \mathrm{ord}(Q_p)$, we have $2t \leq n_1(N_p) \leq w^\beta$. Then $e^{-(r-1)^2/w^\beta} > e^{-r(r-1)/2t}$. But $e^{-(r-1)^2/w^\beta} = 1/2$, since $r = (w^\beta \log 2)^{1/2} + 1$. So

$$1/2 > e^{-r(r-1)/2t}.$$

Since $e^{-r(r-1)/2t}$ is an approximation of $\mathrm{PE}(r,t)$ and $e^{-r(r-1)/2t} > \mathrm{PE}(r,t)$, we have $1/2 > \mathrm{PE}(r,t)$. Hence $1 - \mathrm{PE}(r,t)$, the probability of finding two equivalent points, will be higher than $1/2$. □

Let $w = p^{1/\alpha}$, $w^\beta = p^{\beta/\alpha}$. Using the conjecture in section 5.1 , we can assume that the success probability of one trial is $\mu(\alpha, \beta)$, and the number of trials which the algorithm needs is $1/\mu(\alpha, \beta)$.

5.4 Algorithm analysis

In section 5.1, we mentioned that we need to find two equivalent points q_iQ and q_jQ. However, we don't know exactly the values of i, j. So, Brent suggested a way that we multiply all terms $(x_i - x_j)$ $(1 \leq i < j \leq r)$ together. We let D be

$$D = \prod_{i=1}^{r-1} \prod_{j=i+1}^{r} (x_i - x_j) \pmod N.$$

Then, computing $\gcd(N, D)$ can also factorize N.

The second phase can be divided into the following steps:

1. Generate r points, and get $x_1, \ldots, x_r$.

Let the running time of step1 be w_{21}, then w_{21} =O(r) units of running time.

2. Compute D.

It includes $O(r^2)$ multiplications modulo N. A multiplication modulo N ($AB \pmod N$) needs $O(\log^2 N)$ bit operations (See [8] , p.7). So, a multiplication modulo N needs $O(1/\log N)$ units of running time. Let the running time of step2 be w_{22}, then w_{22} =O$(r^2/\log N)$ units of running time.

3. Compute $\gcd(N, D)$.

Let the running time of step3 be w_{23}, then w_{23} =O$(\log^3 N)$ =O(1) units of running time.

Let the running time of the first phase be w_1. Since $w_{21} \ll w_{22}$, $w_{23} \ll w_{22}$, we just need to compare w_1 and w_{22}. From chapter 4, we know w_1 =O$(\log k)$ =O$(w \log p)$.

If $w_{22} \ll w_1$, then

$$\begin{aligned} r^2/\log N &\ll w \log p \\ w^\beta &\ll w \log p \log N \\ (\beta - 1)\log w &\ll \log\log p + \log\log N \\ \beta &\leq 1 + \frac{\log\log N + \log\log p}{\log w} \end{aligned}$$

So, when $1 < \beta \leq 1 + \dfrac{\log\log N + \log\log p}{\log w}$, we get a speedup of Brent's method over Lenstra's, which is

$$\frac{1/\mu(\alpha, \beta)}{1/\rho(\alpha)} = \mathrm{O}\left(\frac{\ln \alpha}{(\alpha \ln \alpha)^\beta}\right).$$

Chapter 6

Experiments

6.1 Parameters for implementation

We choose a prime number $p \simeq 10^{10}$, and another prime number $q \simeq 10^{20}$. Let $N = pq$ be the number used in our simulation. Then $N \simeq 10^{30}$. From chapter 4, we have the parameter k as $k = \prod p_i^{e_i}$, with p_i prime, $p_i \leq w$, $p_i^{e_i} \leq p+1+2\sqrt{p}$. However our computer cannot finish running the program in an acceptable time, since it is too large. So we change it to

$$k = \prod p_i^{e_i} \text{ , with } p_i \text{ prime, } p_i \leq w \text{ , } p_i^{e_i} \leq w \text{ .}$$

According to our experience, this change will not decrease the success probability too much, but the algorithm runs faster.

We don't know the exact size of p at the beginning. So we must choose a value v to estimate p. Since $w = L(p)^{1/\sqrt{2}+o(1)}$, we can just let

$$w = \exp\left(\sqrt{\log v \log\log v/2}\right).$$

Suppose that $w = p^{1/\alpha}$, then $h = 1/\rho(\alpha) = L(p)^{1/\sqrt{2}+o(1)}$. So we can let

$$h = w$$

in practice.

6.2 The size of v

From previous section, we see that the parameters w, w' and h all depend on v. This experiment is designed to verify the effect on the running efficiency of different sizes of v. It includes the following steps 1-3:

1. Use PARI to implement Lenstra's algorithm. (please see A.1.1)

2. Fix a certain value of v, run the algorithm for 10 times.

3. Select different v and repeat step2 again.

We set h as the maximal number of trials, and a variable t with initial value 1. If one trial fails, the value of t is increased by 1. When t becomes $> h$, the algorithm stops and declares that the factorization has failed. Therefore, t records the observed value of the number of trials needed for successfully factorizing N.

When v has lower order than p, the algorithm will not always succeed, since k and h are too small. Suppose that there are s times among 10 times such that the algorithm is successful. We record s for every choice of v. The result is as follows:

v	10^8	$2 \cdot 10^8$	$4 \cdot 10^8$	$6 \cdot 10^8$	$8 \cdot 10^8$	10^9	$3 \cdot 10^9$	$7 \cdot 10^9$
$s/10$	0.1	0.2	0.3	0.4	0.6	0.8	0.9	0.9

Let $s/10$ denote the success probability (under specified choice of v). From the above result, we see that v shouldn't be chosen too small, otherwise it may not keep the success probability at a reasonable level.

When v has the same or higher order than p, basically it will succeed. But t has a different value in each of the 10 trials. (please see A.1.2) So we take the average value as $\bar{t} = (\sum t)/10$. The result is as follows:

v	10^{10}	10^{11}	10^{12}	10^{13}	10^{14}	10^{15}
s	10	10	10	10	10	10
w	408	601	873	1252	1776	2493
$\bar{t}$	74.3	51.5	31.7	31.7	27.1	20.5
$w\bar{t}$	30314.4	30951.5	27674.1	39688.4	48129.6	51106.5

We can use $w\bar{t}$ to estimate the observed running time. It seems to have a relatively stable value. When v increases, the observed running time also slowly increases.

6.3 The observed speedup

Let's review the conclusion in Chapter 5 in which we found that Brent's method has a speedup of $\rho(\alpha)/\mu(\alpha,\beta)$ over Lenstra's. Since $\rho(\alpha)/\mu(\alpha,\beta) = O\left(\ln\alpha/(\alpha\ln\alpha)^{\beta}\right)$, this experiment is designed to compare the observed value of the speedup and $\ln\alpha/(\alpha\ln\alpha)^{\beta}$. It includes the following steps 1-4:

1. Use PARI to program the algorithm using Brent's implementation. (please see A.2.1) We set value h as the maximal times of trial, and variables t_1, t_2 with initial values 1.

(1.1) If one trial fails, the values of t_1, t_2 are both increased by 1. When t_1 exceeds h, the algorithm stops and declares that the factorization has failed.

(1.2) If it succeeds in the second phase of a certain trial, we still let the program continue to run. But only the first phase will be executed in a trial after that time, and only the value of t_1 will be increased.

(1.3) If it succeeds in the first phase of a certain trial, then we just stop the program.

Clearly, t_1, t_2 record the observed values of the number of trials needed in Lenstra's algorithm (Chapter 4) and in Brent's implementation (Chapter 5), respectively.

2. Choose parameters.

(2.1) The values of N and p used in our simulation are the same as in previous section: $N \simeq 10^{30}$, $p \simeq 10^{10}$.

(2.2) v (the estimation of p) is fixed to 10^{10}. So we take $h = w = 408$ (please see the result in previous section).

(2.3) Since $w = p^{1/\alpha} = L(p)^{1/\sqrt{2}+o(1)}$, we can get $\alpha = O(\sqrt{2\log p/\log\log p})$. In practice, we let $\alpha = \sqrt{2\log v/\log\log v}$.

3. Fix a certain value of β, run the algorithm for 10 times.

4. Select different β and repeat step3 again.

Since t_1 and t_2 have different values in each of the 10 trials (please see A.2.2), we take the average value as $\bar{t_1} = (\sum t_1)/10$ and $\bar{t_2} = (\sum t_2)/10$. Then, for a fixed value of β, $s = \bar{t_2}/\bar{t_1}$ can be regarded as the observed value of the speedup of Brent's method over Lenstra's. And we calculate $d = \ln\alpha/(\alpha\ln\alpha)^{\beta}$ for every value of β. So we can compare s and d. The result is as follows:

β	1.9	1.8	1.7	1.6	1.5	1.4	1.3	1.2	1.1
$\bar{t_2}$	12.6	13.9	19.8	23.6	23.6	23.6	27.7	32.8	35.7
$\bar{t_1}$	74.3	74.3	74.3	74.3	74.3	74.3	74.3	74.3	74.3
s	0.170	0.187	0.266	0.318	0.318	0.318	0.373	0.441	0.480
d	0.060	0.070	0.083	0.098	0.115	0.136	0.160	0.188	0.222
s/d	2.833	2.671	3.205	3.245	2.765	2.338	2.331	2.346	2.162

The average value of s/d is 2.655. It looks approximately like a constant. Thus, we conjecture that s and d have the same order.

Bibliography

[1] R. P. Brent, *Some integer factorization algorithms using elliptic curves*, Australian Computer Science Communication, 8 (1986), 149-163.

[2] E. R. Canfield, P. Erdös, C. Pomerance, *On a problem of Oppenheim concerning "Factorisatio Numerorum"*, J. Number Theory 17 (1983), 1-28.

[3] H. Davenport, *Multiplicative number theory*, 3rd ed. Springer-Verlag, New York, 2000.

[4] M. Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*, Abh. Nath. Sem. Hansischen Univ. 14 (1941), 197-272.

[5] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed. Wiley, New York, 1968.

[6] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, Journal de Théorie des Nombres de Bordeaux 5 (1993), 1-74.

[7] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 4th Edition, 1960.

[8] N. Koblitz, *A course in number theory and cryptography*, Springer-Verlag, New York, 1987.

[9] D. E. Knuth and L. Trabb Pardo, *Analysis of a simple factorization algorithm*, Theoretical Computer Science, 3 (1976), 321 - 348.

[10] H. W. Lenstra, *Factoring integers with elliptic curves*, Ann. Math. 126 (1987), 649-673.

[11] D. A. Marcus, *Number fields*, Springer-Verlag, New York 1977.

[12] J. M. Pollard, *Theorems on factorization and primality testing*, Proc. Cambridge Philos. Soc. 76 (1974), 521-528.

[13] K. Prachar, *Primzahlverteilung*, Springer-Verlag, Berlin 1957.

[14] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM Journal on Computing, Volume 26, Issue 5 (October 1997), 1484 - 1509.

[15] L. C. Washington, *Elliptic curves: number theory and cryptography*, Chapman & Hall/CRC, 2003.

Appendix A

PARI Programs

A.1 Program 1

A.1.1 codes

These are the codes for generating n which is $\simeq 10^{30}$.

```
p0 = 10^10; p1 = precprime(p0); p2 = nextprime(p0);
if (p0-p1 <= p2-p0, p = p1, p = p2);
q0 = 10^20; q1 = precprime(q0); q2 = nextprime(q0);
if (q0-q1 <= q2-q0, q = q1, q = q2);
n = p*q;
```

These are the codes for $v = 10^{10}$. If v takes a different value, the codes are similar.

```
v = 10^10;
w = ceil(exp(sqrt( log(v)*log(log(v))/2 ))); h = w;
```

```
write(”c:\\record1.txt”, ”###,”);
EZn = vector(5,Vi,0);
P = vector(2,Vi,0);
Q = vector(2,Vi,0);
t = 1;
while (t <= h,
      a = Mod(random(n),n);
      x = Mod(random(n),n);
      y = Mod(random(n),n);
      b = y*y - x*x*x - a*x;
      EZn[4] = a; EZn[5] = b;
      P[1] = x; P[2] = y;
      write(”c:\\record1.txt”, t ”,”);
```

PARI function ”elladd” is for adding two points, and ”ellpow” is for adding a points for a given number of times. If they can’t be done, PARI system will exit from the program and display an error message which gives a non-trivial prime factor of n.

```
      forprime (p_i = 2, w,
            e_i = floor(log(w)/log(p_i));
            for (i=1, e_i,
                  Q = ellpow(EZn, P, p_i); P = Q
            )
      );
      t = t + 1
)
```

A.1.2 results

When $v = 10^{10}$, $t = \{71, 101, 33, 66, 205, 39, 10, 58, 136, 24\}$. So $\bar{t} = 74.3$.

When $v = 10^{11}$, $t = \{42, 29, 84, 17, 33, 50, 16, 171, 34, 39\}$. So $\bar{t} = 51.5$.

When $v = 10^{12}$, $t = \{42, 23, 6, 84, 17, 33, 50, 16, 9, 37\}$. So $\bar{t} = 31.7$.

When $v = 10^{13}$, $t = \{42, 23, 6, 84, 17, 33, 50, 16, 9, 37\}$. So $\bar{t} = 31.7$.

When $v = 10^{14}$, $t = \{42, 23, 6, 1, 83, 4, 13, 33, 50, 16\}$. So $\bar{t} = 27.1$.

When $v = 10^{15}$, $t = \{6, 36, 23, 4, 2, 1, 83, 4, 13, 33\}$. So $\bar{t} = 20.5$.

A.2 Program 2

A.2.1 codes

The codes for generating n which is $\simeq 10^{30}$ are the same as in Program 1.

These are the codes for $\beta = 1.5$. If β takes a different value, the codes are similar.

```
v = 10ˆ10;
w = ceil(exp(sqrt( log(v)*log(log(v))/2 ))); h = w;
r = ceil(sqrt( log(2)*wˆ1.5 ));
write(”c:\\record2.txt”, ”###,”);
EZn = vector(5,Vi,0);
P = vector(2,Vi,0);
Q = vector(2,Vi,0);
Q0 = vector(2,Vi,0);
X_Q = vector(r,Vi,0);
t = 1;
```

```
flag1 = 0;

while (t <= h,
          a = Mod(random(n),n);
          x = Mod(random(n),n);
          y = Mod(random(n),n);
          b = y*y - x*x*x - a*x;
          EZn[4] = a; EZn[5] = b;
          P[1] = x; P[2] = y;
          write(”c:\\record2.txt”, t ”,”);

          forprime (p_i = 2, w,
               e_i = floor(log(w)/log(p_i));
               for (i=1, e_i,
                    Q = ellpow(EZn, P, p_i); P = Q
               )
          );
```

For the second phase, a randomly generated array ”rc” of values 0 and 1 is set in advance. In each time the program will take a value from ”rc”. If the value is 0, $q_{i+1}Q = q_iQ + q_iQ$. If the value is 1, $q_{i+1}Q = q_iQ + q_iQ + Q$.

```
          if (flag1==0,
               Q0 = Q; X_Q[1] = Q0[1];
               d=Mod(1, n);
               write(”c:\\record2.txt”, ”**,”);
```

```
        for (i=2, r,
            l = component(Q[2] + Q[2],2);
            if (gcd(n,l)!=1, flag1 = 1;break );
            P = elladd(EZn, Q, Q);
            if (rc[i]==0, Q=P);
            if (rc[i]==1,
                if (P[1]==Q0[1], l = component(P[2] + Q0[2],2) );
                if (P[1]!=Q0[1], l = component(P[1] - Q0[1],2) );
                if (gcd(n,l)!=1, flag1 = 1;break );
                Q = elladd(EZn, P, Q0)
            );
            X_Q[i] = Q[1];
            for (j=1, i-1,
                l = component(X_Q[i] - X_Q[j],2);
                if (gcd(n,l)!=1, flag1 = 1;break )
            );
            if (flag==1, break)
        )
    );
    t = t + 1
)
```

A.2.2 results

Even if β is different, the values of t_1 remain the same, since v is fixed. $t_1 = \{71, 101, 33, 66, 205, 39, 10, 58, 136, 24\}$. So $\bar{t_1} = 74.3$.

When $\beta = 1.1$, $t_2 = \{71, 84, 33, 44, 9, 39, 10, 38, 5, 24\}$. So $\bar{t_2} = 35.7$.

When $\beta = 1.2$, $t_2 = \{42, 84, 33, 44, 9, 39, 10, 38, 5, 24\}$. So $\bar{t_2} = 32.8$.

When $\beta = 1.3$, $t_2 = \{42, 62, 33, 44, 9, 10, 10, 38, 5, 24\}$. So $\bar{t_2} = 27.7$.

When $\beta = 1.4$, $t_2 = \{4, 62, 33, 44, 9, 10, 10, 35, 5, 24\}$. So $\bar{t_2} = 23.6$.

When $\beta = 1.5$, $t_2 = \{4, 62, 33, 44, 9, 10, 10, 35, 5, 24\}$. So $\bar{t_2} = 23.6$.

When $\beta = 1.6$, $t_2 = \{4, 62, 33, 44, 9, 10, 10, 35, 5, 24\}$. So $\bar{t_2} = 23.6$.

When $\beta = 1.7$, $t_2 = \{4, 28, 33, 40, 9, 10, 10, 35, 5, 24\}$. So $\bar{t_2} = 19.8$.

When $\beta = 1.8$, $t_2 = \{4, 28, 33, 9, 9, 4, 10, 35, 5, 2\}$. So $\bar{t_2} = 13.9$.

When $\beta = 1.9$, $t_2 = \{4, 28, 20, 9, 9, 4, 10, 35, 5, 2\}$. So $\bar{t_2} = 12.6$.