# Counting elliptic curves of bounded Faltings height
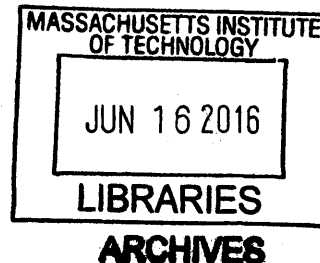
by

Ruthi Hortsch

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author . . . . . . . . . . . . . . . . **Signature redacted** . . . . . . . . . .
/　　　　　ν　　Department of Mathematics
April 29, 2016

Certified by . . . . . . . . . **Signature redacted** . . . . . . . . . . . .
ν
Bjorn Poonen
Claude Shannon Professor of Mathematics
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . **Signature redacted**
Alexei Borodin
Chairman, Department Committee on Graduate Theses

# Counting elliptic curves of bounded Faltings height

by

Ruthi Hortsch

Submitted to the Department of Mathematics
on April 29, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

## Abstract

Because many invariants and properties of elliptic curves are difficult to understand directly, the study of *arithmetic statistics* instead looks at what happens "on average", using heights to make this notion rigorous. Previous work has primarily used the naive height, which can be calculated easily but is not defined intrinsically.

We give an asymptotic formula for the number of elliptic curves over $\mathbb{Q}$ with bounded Faltings height. Silverman [34] has shown that the Faltings height for elliptic curves over number fields can be expressed in terms of the minimal discriminant and period of the elliptic curve. We use this to recast the problem as one of counting lattice points in an unbounded region in $\mathbb{R}^2$ defined by transcendental equations, and understand this region well enough to give a formula for the number of these lattice points.

Thesis Supervisor: Bjorn Poonen
Title: Claude Shannon Professor of Mathematics

# Acknowledgments

# Contents

# Chapter 1

# Introduction

Let $\mathcal{E}$ be the set of isomorphism classes of elliptic curves over $\mathbb{Q}$. There are many difficult open questions and conjectures about invariants of elliptic curves (such as the rank of the rational points of an elliptic curve or the size of its $n$-Selmer group). One approach to understanding these invariants better is to consider what value these take "on average".

We use different heights to make sense of what is meant by average here. A *height* is a function

$$H : \mathcal{E} \to \mathbb{R}$$

such that the set

$$\mathcal{E}_{H<X} = \{E \in \mathcal{E} \mid H(E) < X\}$$

is finite. For a given invariant $i$, it then makes sense to talk about the average value of $i$ over elements of $\mathcal{E}_{H<X}$; we can define the average value of $i$ to be the limit

$$\text{avg}(i) = \lim_{X \to \infty} \frac{\sum_{E \in \mathcal{E}_{H<X}} i(E)}{\#\mathcal{E}_{H<X}}$$

if this limit exists. To use heights to measure the invariants, we first need to understand just how quickly $\#\mathcal{E}_{H<X}$ grows.

Commonly used is the *naive height $H_N$*. Every elliptic curve $E$ over $\mathbb{Q}$ is isomorphic to an elliptic curve of the form $y^2 = x^3 + Ax + B$ with $A, B \in \mathbb{Z}$. If we require that

there is no prime $p$ with $p^4 \mid A$ and $p^6 \mid B$, this is still true and $(A, B)$ is unique. The naive height of $E$ is then

$$H_N(E) = \max(27B^2, 4|A|^3).$$

It is a standard result that

$$\#\mathcal{E}_{H_N < X} = 4^{2/3}3^{-3/2}\zeta(10)^{-1}X^{5/6} + O(X^{1/2}).$$

We have included a proof of this in Section 2.1 (Brumer [13] has also written out a proof; Bekyel [3] has done some related calculations over number fields). The key idea is to use the following standard result (see Section 8 of [26] Vol 2 for a proof).

**Theorem 1.1.** *Let $S \subset R^2$ be a region with a boundary that is a rectifiable curve. Let $L$ be the length of the boundary of $S$. Then*

$$|\#(S \cap \mathbb{Z}^2) - \mathrm{Area}(S)| \le 4(L + 1).$$

*In particular,*

$$\#(S \cap \mathbb{Z}^2) = \mathrm{Area}(S) + O(L).$$

The region where naive height is less than $X$ is a rectangle of length $2 \cdot (X/4)^{1/3}$ and height $2 \cdot (X/27)^{1/2}$, which has area $4^{2/3}3^{-3/2}X^{5/6}$ and boundary length $O(X^{1/2})$ (the $\zeta(10)$ corrects for overcounting isomorphic curves).

Approximating the number of lattice points by the area works better in situations where $S$ is "round looking", where there are no parts of the region that are long and skinny, like tentacles (since then the length of the boundary becomes large relative to the area). This is the difficulty one runs into when trying to apply such an argument when counting elliptic curves using discriminant as a height: the points with large $A, B$ and small discriminant form a tentacle-like region that is not easily controlled (this is discussed further in Sections 2.2 and 2.3). Similar tentacles show up for counting number fields by discriminant, discussed further in Section 2.4, although

10

unlike the equations we will deal with for the Faltings height, these are defined by algebraic equations.

Brumer and McGuinness [14] have given a heuristic for the number of elliptic curves with positive (respectively negative) discriminant up to a bound, and Watkins [39] has used this to give heuristics for the average rank counted this way, as well as heuristics for elliptic curves of bounded conductor. It is generally believed that the average rank should be the same for each of these heights. However, no proof has been given of these conjectures.

We will focus on using the Faltings height $h_F$, first introduced by Faltings [22] in the proof of the famous Mordell conjecture; Chapter 3 discusses the Faltings height more thoroughly. Silverman [17] reformulated the Faltings height over elliptic curves in terms of fundamental constants of an elliptic curve, and this formulation is the one we use. Let $\mathcal{H}$ denote the complex upper half plane. For $\tau \in \mathcal{H}$,

$$\Delta(\tau) = (2\pi)^{12} q_\tau \prod_{n=1}^{\infty} (1 - q_\tau^n)^{24} \text{ where } q_\tau = e^{2\pi i \tau}.$$

If $E$ is an elliptic curve over $\mathbb{Q}$, let

$\Delta_E^{\min}$ be the minimal discriminant of $E$, and

$\tau_E \in \mathcal{H}$ be a period of $E$, so that $E(\mathbb{C}) \cong \mathbb{C}/(\mathbb{Z} + \tau_E \mathbb{Z})$.

Note that there are three values used in this thesis which have a delta in their notation but are slightly different: $\Delta_E^{\min}$ (the minimal discriminant), $\Delta(\tau)$ (the modular discriminant), and $\Delta_{A,B}$ (the polynomial discriminant, defined in the next section).

Silverman's reformulation says that

$$h_F(E) = \frac{1}{12} \left( \log |\Delta_E^{\min}| - \log |\Delta(\tau_E) \operatorname{Im}(\tau_E)^6| \right).$$

Silverman also proved that for any $\varepsilon > 0$,

$$h_F(E) + O(1) \leq \frac{1}{12} \log(H_N(E)) \leq (1 + \varepsilon) h_F(E) + O_\varepsilon(1). \tag{1.1}$$

To simplify the notation, we instead study $H_F(E) = e^{12h_F(E)}$. The main result of this thesis, summarized below and proved in Chapter 4, counts $\#\mathcal{E}_{H_F < X}$.

## 1.1 Summary of the main result

For $\tau$ in the complex upper half plane $\mathcal{H}$, let $j(\tau)$ denote its $j$-invariant, and let $\Delta(\tau)$ be the modular discriminant as defined previously. For $t \in \mathbb{R}$ not equal to $-27/4$, let $\tau_t$ be a period for the elliptic curve given by the equation $y^2 = x^3 + tx + t$. Note then $j(\tau_t) = 6912t/(4t + 27)$; this $\tau_t$ is determined modulo the action of $\mathrm{SL}_2(\mathbb{Z})$, so $\Delta(\tau_t) \mathrm{Im}(\tau_t)^6$ is a well-defined function of $t$.

The main result in this thesis, proved in Chapter 4, is:

**Theorem 1.2.** *Let*
$$\sigma = \frac{2}{5} \int_{-\infty}^{\infty} t^{-2/3} \left| \frac{\Delta(\tau_t) \mathrm{Im}(\tau_t)^6}{16(4t + 27)} \right|^{5/6} dt$$

*(the indefinite integral converges even though the integrand tends to $\infty$ as $t$ approaches $-27/4$ or $0$). Then*

$$\#\mathcal{E}_{H_F < X} = 12\zeta(10)^{-1}\sigma X^{5/6} + O(X^{1/2}(\log X)^3).$$

When the integral defining $\sigma$ is evaluated, we get approximately $\sigma \approx 29089$, which means the constant of the leading term is approximately 348716.

This estimate does not follow from Silverman's result given by (1.1), since $H_F$ is not simply bounded below by a constant times $H_N$. To prove Theorem 1.2, we identify isomorphism classes of elliptic curves of Faltings height less than $X$ with lattice points (satisfying certain congruence conditions) in a region $R_X$ in the real $(A, B)$-plane (defined below), and (with some difficulty) apply Theorem 1.1.

For $(A, B) \in \mathbb{R}^2$, let $\Delta_{A,B} = -16(4A^3 + 27B^2)$, and if $\Delta_{A,B} \neq 0$, let $\tau_{A,B}$ be a period corresponding to the elliptic curve given by the equation $y^2 = x^3 + Ax + B$. Then let

$$R_X = \left\{ (A, B) \in \mathbb{R}^2 \ \middle| \ \frac{|\Delta_{A,B}|}{|\Delta(\tau_{A,B})| \operatorname{Im}(\tau_{A,B})^6} < X \right\}.$$

The lattice points in $R_X$ do not quite correspond to the set $\mathcal{E}_{H_F < X}$ that we want to count. For one thing, we want to count only one lattice point per isomorphism class; as in the naive height case, $\zeta(10)^{-1}$ corrects for this. Furthermore, $\Delta_{A,B}$ is not the exact minimal discriminant; it can differ by a factor having to do with behavior at the special fibers for $p = 2$ and $p = 3$ (this is discussed in Section 2.2). However, we can correct it with a constant that depends on $A, B$ mod $6^6$, which is where the 12 in Theorem 1.2 comes from; these can be calculated with Tate's algorithm, the summary of which is in Section 4.4.

We want to use Theorem 1.1 to say that the number of lattice points in $R_X$ is its area $\sigma X^{5/6}$ plus an error from the boundary, but there are some difficulties in the proof. As in the analogue where the discriminant is used as a height, this region has unbounded tentacles asymptotic to the curve $4A^3 + 27B^2 = 0$. A key argument in the proof is to show that all of the points being counted lie in a particular rectangle $R$, which importantly does not have too large a perimeter (this gets done in Section 4.3). Once that has been proved, we show in Section 4.5 that Theorem 1.1 can actually be applied to $R_X \cap R$ and the error, which comes from the boundary, is not too large. Section 4.6 pulls everything together to conclude the proof.

## 1.2 Further questions

There are natural questions to consider once one has this result. In particular, it should be simple to apply the same methods to count elliptic curves over function fields by the Faltings height (as defined and discussed in [2]).

Bhargava and Shankar [8] [9] using the naive height have proved a number of results including bounds on the average values of the 2-Selmer group and 3-Selmer group, which imply a bound on the average value of the rank (see [33] for a expository description of these results, [10] [11] for a generalization to all global fields). These methods combined with the work in this thesis might result in similar bounds for the

Faltings height, and give evidence for (or against) the hypothesis that the different heights should give the same averages. At the moment, there are still unresolved difficulties with combining these methods.

Harron and Snowden [24] also recently determined how often each of the possible torsion subgroups prescribed by Mazur's theorem occurs for elliptic curves over $\mathbb{Q}$ with bounded naive height. One should also be able to using the Faltings height to study this.

# Chapter 2

# Heights and arithmetic statistics

In this chapter, we discuss background on heights. Section 2.1 gives an introduction to the naive height, and explains how to count elliptic curves with bounded naive height. Section 2.2 introduces the minimal discriminant and discusses how one might attempt to count elliptic curves with bounded minimal discriminant, while Section 2.3 discusses what the abc conjecture gives us (and why it is still not enough). In Section 2.4, we mention work on counting number fields by discriminants, in which tentacles are successfully controlled, although the methods are not ones that apply to the case of the Faltings height.

## 2.1 The naive height

If $(A, B) \in \mathbb{Z}^2$, then call that point *weakly minimal* provided that there is no prime $p$ such that $p^4 \mid A$ and $p^6 \mid B$. Note that any elliptic curve over $\mathbb{Q}$ is isomorphic to a unique curve $E_{A,B}$ of the form $y^2 = x^3 + Ax + B$ with weakly minimal $(A, B) \in \mathbb{Z}^2$. The naive height of an elliptic curve $E$ is defined to be

$$H_N(E) = \max(4|A|^3, 27|B|^2)$$

for weakly minimal $(A, B) \in \mathbb{Z}$ such that $E$ is isomorphic to $E_{A,B}$.

**Theorem 2.1.**

$$\#\mathcal{E}_{H_N<X} = 4^{2/3}3^{-3/2}\zeta(10)^{-1}X^{5/6} + O(X^{1/2})$$

*Proof.* There is a bijection between

$$\left\{ \text{weakly minimal } (A,B) \in \mathbb{Z}^2 \,\middle|\, 4A^3+27B^2 \neq 0, |A| < (X/4)^{1/3}, \text{ and } |B| < (X/27)^{1/2} \right\}$$

and $\mathcal{E}_{H_N<X}$ given by $(A, B) \mapsto$ isomorphism class of $E_{A,B}$. So we can simply count the former. In fact, we can ignore the condition that $4A^3 + 27B^2 \neq 0$, since there are at most $O(X^{1/2})$ points that satisfy this condition in the rectangle (in fact, parametrizing, we can even show that it is $O(X^{1/6})$). Let

$$S_{H_N<X} = \left\{ (A,B) \in \mathbb{Z}^2 \,\middle|\, |A| < (X/4)^{1/3}, \text{ and } |B| < (X/27)^{1/2} \right\},$$

so we want to count the weakly minimal points in $S_{H_N<X}$.

Recall the statement of Theorem 1.1: given a region $S \subset \mathbb{R}^2$ with rectifiable boundary of length $L$, the number of lattice points contained in it is $\mathrm{Area}(S) + O(L)$. So it follows that

$$\#S_{H_N<X} = 4^{2/3}3^{-3/2}X^{5/6} + O(X^{1/2}) \tag{2.1}$$

since $S_{H_N<X}$ is just the lattice points in a rectangle with sides of lengths $2(X/4)^{1/3}$ and $2(X/27)^{1/2}$. Furthermore, for fixed $d$, there is a bijection between

$$\left\{ (A,B) \in S_{H_N<X} \,\middle|\, d^4 \mid A, d^6 \mid B \right\}$$

and

$$\left\{ (a,b) \in \mathbb{Z}^2 \,\middle|\, |a| < (X/4)^{1/3}d^{-4}, |b| < (X/27)^{1/2}d^{-6} \right\},$$

given by $(A, B) \mapsto (d^{-4}A, d^{-6}B)$. We count the second set using Theorem 1.1 and conclude that for $d < X^{1/12}$,

$$\#\left\{ (A,B) \in S_{H_N<X} \,\middle|\, d^4 \mid A, d^6 \mid B \right\} = 4^{2/3}3^{-3/2}d^{-10}X^{5/6} + O(d^{-6}X^{1/2}).$$

16

We will use a Möbius inversion argument to complete the proof. Note that if $\mu$ is the classical Möbius function, then for $(A, B) \in \mathbb{Z}^2$,

$$\sum_{d^4|A, d^6|B} \mu(d) = \begin{cases} 1 & \text{if } (A, B) \text{ weakly minimal} \\ 0 & \text{otherwise,} \end{cases}$$

so the number of weakly minimal $(A, B) \in S_{H_N < X}$ is

$$\sum_{\substack{(A,B) \in S_{H_N < X}}} \sum_{\substack{d^4|A \\ d^6|B}} \mu(d) = \sum_{d < (X/27)^{1/12}} \mu(d) \sum_{\substack{(A,B) \in S_{H_N < X} \\ d^4|A, d^6|B}} 1$$

$$= \sum_{d < (X/27)^{1/12}} \mu(d) \left( 4^{2/3} 3^{-3/2} d^{-10} X^{5/6} + O(d^{-6} X^{1/2}) \right)$$

$$= 4^{2/3} 3^{-3/2} X^{5/6} \sum_{d < (X/27)^{1/12}} \mu(d) d^{-10} + O\left( X^{1/2} \sum_{d < (X/27)^{1/12}} \mu(d) d^{-6} \right)$$

$$= 4^{2/3} 3^{-3/2} \zeta(10)^{-1} X^{5/6} + O(X^{1/2}).$$

The last step is a consequence of the statement

$$\sum_{d=1}^{Y} \mu(d) d^{-10} = \zeta(10)^{-1} + O(Y^{-9}),$$

which can be proved using the integral test since $|\mu(d)| \le 1$ for all $d$. $\qquad \square$

Counting by naive height has been a building block of many results about the distribution of invariants of elliptic curves, including the results of Bhargava and Shankar [8], and Harron and Snowden [24] mentioned in Section 1.2.

## 2.2 Using the minimal discriminant as a height

A natural height to use to count elliptic curves over $\mathbb{Q}$ is the *minimal discriminant* $\Delta_E^{\min}$. In this section we will give some background on what the minimal discriminant

is, and the difficulties encountered when counting elliptic curves by it. Section 2.3 will discuss why the abc conjecture is relevant but does not help much.

Given a Weierstrass equation of an elliptic curve $E/\mathbb{Q}$, there is a value called the discriminant $D$ that is an algebraic function of the coefficients (see Section III.1 of [36]). For a fixed prime $p$, we say that a Weierstrass equation for $E$ is a *minimal (Weierstrass) equation for $E$ at $p$* if the valuation $\mathrm{val}_p(D)$ of the discriminant is minimal over all Weierstrass equations with coefficients in $\mathbb{Z}$.

The *minimal discriminant* $\Delta_E^{\min} \in \mathbb{Z}_{>0}$ is the product of $p^{\mathrm{val}_p(D)}$ over all primes $p$ where $D$ is the discriminant of a minimal equation at $p$. Similarly, for an elliptic curve $E$ over a number field $K$, one can define $\Delta_K^{\min}$ as an integral $\mathcal{O}_K$-ideal. Every elliptic curve over $\mathbb{Q}$ has a Weierstrass equation that is minimal at all primes $p$, but this fails for general number fields. (See Sections VII.1 and VIII.8 of [36] for an introduction to the minimal discriminant.)

**Proposition 2.2.** *Let $E_{A,B}$ be an elliptic curve over $\mathbb{Q}$ of the form $y^2 = x^3 + Ax + B$, where $(A, B) \in \mathbb{Z}^2$ is weakly minimal. Let*

$$\Delta_{A,B} = -16(4A^3 + 27B^2).$$

*The minimal discriminant of $E_{A,B}/\mathbb{Q}$ is then*

$$\Delta_{E_{A,B}}^{\min} = \lambda \Delta_{A,B}$$

*for some $\lambda \in \{1, 2^{-12}, 3^{-12}, 6^{-12}\}$. The value of $\lambda$ depends only on $A \bmod 6^6$ and $B \bmod 6^6$.*

*Proof.* Note that $\Delta_{A,B} = -16(4A^3 + 27B^2)$ is the polynomial discriminant for the equation $y^2 = x^3 + Ax + B$. Remark 1.1 in [36] Chapter VII tell us that for prime $p$, if either $p^4 \nmid -48A$ or $p^{12} \nmid \Delta_{A,B}$ then $\mathrm{val}_p(\Delta_{A,B}) = \mathrm{val}_p(\Delta_{E_{A,B}}^{\min})$.

If $p \neq 2, 3$, $p^4 \mid -48A$ and $p^{12} \mid \Delta_{A,B}$ together imply that $p^4 \mid A$ and $p^6 \mid B$, so this would mean that $(A, B)$ is not weakly minimal. For $p = 2, 3$, if we start with a minimal Weierstrass equation not in the form of $y^2 = x^3 + Ax + B$, then completing

18

the square and cube requires dividing by 2 and 3, potentially resulting in non-integer coefficients. To correct this, scale $x$ by $u^2$ and $y$ by $u^3$, where $u$ is the minimal value of $1, 2, 3, 6$ that will make coefficients integral. Completing the square and cube doesn't change the discriminant, but scaling does, multiplying the discriminant by $u^{12}$.

To see that the value of $\lambda$ is determined by $A$ mod $6^6$ and $B$ mod $6^6$, use Tate's algorithm (as described in Chapter IV Section 9 of [35]). $\qquad\square$

So given Proposition 2.2, here is an approach to counting the isomorphism classes of elliptic curves with minimal discriminant less than $X$ (these are the same steps we will use in Chapter 4 to prove Theorem 1.2):

1. Fix residue classes $A_0$ mod $6^6$ and $B_0$ mod $6^6$ (excluding those such that the pair would not be weakly minimal with respect to 2 or 3). Note that there is a well defined $\lambda \in \{1, 2^{-12}, 3^{-12}, 6^{-12}\}$ that depends only on $A_0$, $B_0$.

2. Count the number of weakly minimal $(A, B) \in \mathbb{Z}^2$ such that $4A^3 + 27B^2 \neq 0$, $16|4A^3 + 27B^2| < X/\lambda$ and $A \equiv A_0$ mod $6^6$, $B \equiv B_0$ mod $6^6$.

3. Sum this count over all the possible residue classes that could have been chosen in the first step.

The key step is to understand how to count lattice points such that $4A^3 + 27B^2 \neq 0$ in the region
$$R_{\Delta < X} = \left\{ (A, B) \in \mathbb{R}^2 \ \middle|\ 16|4A^3 + 27B^2| < X \right\}$$
(restricting to weakly minimal points and particular residue classes will change the answer only by a constant).

We would like to apply Theorem 1.1 to get the area of $R_{\Delta < X}$. Note that

$$R_{\Delta < X} = \{(X^{1/3}A, X^{1/2}B) \in \mathbb{R}^2 \mid (A, B) \in R_{\Delta < 1}\},$$

so $\mathrm{Area}(R_{\Delta < X}) = \mathrm{Area}(R_{\Delta < 1})X^{5/6}$.

Unfortunately, we cannot just apply Theorem 1.1 directly, since the boundary cannot possibly have finite length. (The region $R_{\Delta < X}$ is an open neighborhood of the
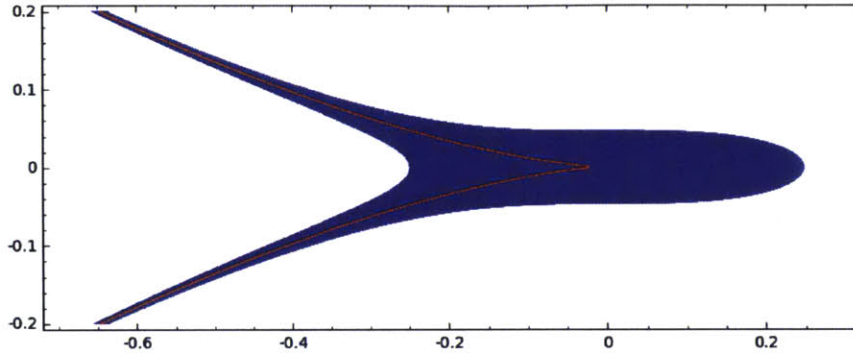
Figure 2-1: The region $16|4A^3 + 27B^2| < 1$ in blue, with the curve $4A^3 + 27B^2 = 0$ in red.

unbounded cubic $4A^3 + 27B^2 = 0$ — it has tentacles — so the boundary will clearly not have finite length. See Figure 2-1.) In the next section, we will show how the abc conjecture helps us restate the question in terms of a region with boundary that has finite length (it helps us cut off the tentacles), but this still will not help enough — we still will not be able to show that the length of the boundary is less than $O(X^3)$, which overwhelms the presumptive leading term of the area.

(For the Faltings height, we will have a similar region, but the tentacles will get narrower more quickly, and we will be able to cut them off at a point that allows us to more easily control what the length of their boundaries are.)

## 2.3   The abc conjecture

We begin by recalling the statement of the abc conjecture. For any nonzero $m \in \mathbb{Z}$, let

$$\mathrm{rad}(m) = \prod_{p|m} p.$$

**Conjecture 2.3** (abc conjecture). *Fix $\varepsilon > 0$. Then there exists a positive $\kappa_\varepsilon$ such that for all nonzero integers $a, b, c$ such that $\gcd(a, b, c) = 1$ and $a + b = c$,*

$$\mathrm{rad}(abc) > \kappa_\varepsilon \max(|a|, |b|, |c|)^{1-\varepsilon}.$$

This means that there is a limit to how small the radical of $abc$ can be relative to

the largest of $a, b, c$. We will apply to this conclude that the discriminant cannot be too small compared to the naive height, and hence not to small relative to $A$ and $B$. Let $a = 4A^3$, $b = 27B^2$, and $c = -\Delta_{A,B}/16$; then $a + b = c$. To make the argument simpler, assume $\gcd(a, b, c) = 1$.

Recall that the naive height $H_N(E_{A,B})$ is defined as $\max(4|A|^3, 27B^2)$. We have problems only where the discriminant is small, so if we restrict our attention to that situation, the naive height is $\max(|a|, |b|, |c|)$. So the abc conjecture tells us that for any $\varepsilon$ there is a $\kappa_\varepsilon$ such that

$$\mathrm{rad}((4A^3 + 27B^2) \cdot 4A^3 \cdot 27B^2) > \kappa_\varepsilon H_N(E_{A,B})^{1-\varepsilon}.$$

Furthermore, note

$$\mathrm{rad}((4A^3 + 27B^2) \cdot 4A^3 \cdot 27B^2) = \mathrm{rad}(3\Delta_{A,B}AB)$$
$$\leq 3|\Delta_{A,B}||A||B|$$
$$\leq 3|\Delta_{A,B}|H_N(E_{A,B})^{5/6}.$$

Putting the previous two equations together, we conclude that

$$|\Delta_{A,B}| > \frac{1}{3}\kappa_\varepsilon H_N(E_{A,B})^{1/6-\varepsilon}.$$

In particular, this tells us that $\Delta_{A,B}$ cannot get too small relative to $A$, $B$. More concretely, if $|\Delta_{A,B}| < X$, then (up to some constants) $|A| < X^{2+\varepsilon'}$ and $|B| < X^{3+\varepsilon''}$ (the $\varepsilon', \varepsilon''$ are not the same $\varepsilon$, but related small quantities). This tells us we can bound the region we need to consider, but counting the rest of the points in the "tentacles" is an open problem.


## 2.4   Counting number fields by discriminant

We will consider in this section some prominent results where lattice counting was successful despite tentacles. One fundamental invariant of a number field $K$ is its

discriminant $\text{Disc}(K)$. A fundamental theorem of Minkowski states that (up to isomorphism) there are only finitely many number fields of discriminant $D$. So it becomes natural to consider the distribution of number fields with respect to discriminants. This question then gets refined further to consider each degree and Galois group separately.

A degree $n$ number field $K$ is called an $S_n$-number field if its Galois closure has Galois group $S_n$. Let $N_n(X)$ denote the number of $S_n$-number fields of degree $n$ having absolute discriminant at most $X$. If $n = 1$, $N_n(X) = 1$ since $\mathbb{Q}$ is the only degree 1 number field. If $n = 2$,

$$\lim_{X \to \infty} \frac{N_n(X)}{X} = \frac{6}{\pi^2}$$

(because being a $S_2$-number field is equivalent to a square-free condition). These results led in the 1960s to the following conjecture (the origin of which is unknown, according to [5]).

**Conjecture 2.4.** *The limit*

$$c_n = \lim_{X \to \infty} \frac{N_n(X)}{X}$$

*exists and is positive for $n \geq 2$.*

Davenport and Heilbronn [20] calculated $c_3$ for cubic fields in 1970, while Bhargava calculated $c_4$ and $c_5$ (in [4], [7] respectively). Inspired by these results and heuristics of Serre, Bhargava [6] has also conjectured an explicit formula for $c_n$, in terms of partitions and the structure of $S_n$. Bhargava has also written a helpful expository summary of these results [5].

The method for calculating $c_3$, $c_4$, and $c_5$ is similar: counting number fields here is equivalent to counting orders, and these are then parametrized by appropriate forms and matrices, which correspond to lattice points in carefully defined regions in $\mathbb{R}^n$. (This summary is a vast simplification of the hard work that goes into these proofs!) In each of these cases, the relevant region has unbounded tentacles and the lattice points in the tentacles need to be properly controlled. One key element in each of

these proofs is a lemma of Davenport [18] [19] that relies on the fact that the region is a semialgebraic set. While these are remarkable proofs, the methods cannot be applied to the situation we will study regarding the Faltings height; in our situation, the region will be defined by transcendental inequalities.

# Chapter 3

# Faltings height

In this chapter we give the historical definition of the Faltings height in Section 3.1, discuss some of the research on it, and then describe the reformulation given by Silverman in Section 3.3, which will be used in the rest of the thesis.

## 3.1  Faltings' original formulation

When proving the Mordell conjecture, Faltings introduced a notion of height of a semiabelian variety over a number field $K$ (see [17] for a longer overview, or the original paper [22]).

Let $\mathcal{O}_K$ denote the ring of integers of $K$. For each infinite place $v$, let $K_v$ denote the completion of $K$ at $v$. A *metrized line bundle* is a projective $\mathcal{O}_K$-module $P$ of rank 1, together with norms $\| \ \|_v$ on $P \otimes_{\mathcal{O}_K} K_v$. The degree of the metrized line bundle is defined to be

$$\deg(P, (\| \ \|_v)_v) = \log(\#(P/p\mathcal{O}_K)) - \sum_{v \text{ infinite}} \varepsilon_v \log \|p\|_v,$$

where $p$ is any nonzero element of $P$ and $\varepsilon_v = 1, 2$ according to whether $v$ is real or complex.

For an abelian variety $A$ of dimension $g$ over $K$, take the metrized line bundle

given by the dualizing sheaf $\omega_{A/\mathcal{O}_K}$ and the norms

$$\|\alpha\|_v^2 = \left(\frac{i}{2}\right)^g \int_{A(\overline{K_v})} \alpha \wedge \overline{\alpha}$$

for $\alpha \in \omega_{A/K} \otimes K_v$. Then the *Faltings height* is defined as

$$h_F(A/K) = \frac{1}{[K:\mathbb{Q}]} \deg(\omega_{A/\mathcal{O}_K}).$$

Faltings used this to prove the Shafarevich conjecture, which says that there are only finitely many isomorphism classes of abelian varieties over a fixed number field of a fixed dimension, with polarization of a fixed degree and good reduction outside of some given finite set of places.

The *stable Faltings height* of an abelian variety $A$ is the Faltings height of $A$ over any field extension over which $A$ acquires semistable reduction. Some authors mean the stable Faltings height when they refer to the Faltings height.

Faltings [22], Silverman [34], and Deligne [21] all use different normalizations of the Faltings height on elliptic curves; in this thesis we consistently use the normalization that Faltings uses.

## 3.2   Review of research on the Faltings height

In this section we will briefly mention some existing results regarding the Faltings height.

Faltings' original proof of the Shafarevich conjecture gave a formula for the relationship between the Faltings height of isogenous semiabelian varieties. Szpiro and Ullmo [38] made this formula more explicit for elliptic curves without complex multiplication.

Colmez [15] [16] has conjectured a formula for the Faltings height of a CM abelian variety in terms of the logarithmic derivatives at $s = 0$ of Artin L-functions, referred to as the *Colmez conjecture*. He also put forward a motivic interpretation of this conjecture as a "product formula" for periods of CM abelian varieties, and proved it

in special cases. Other special cases have been shown in Obus [32] and Yang [40] [41]. Both Yuan and Zhang [42] (using [43] as a basis) and Andreatta, Goren, Howard, and Madapusi Pera [1] recently proved an averaged version.

A number of results bound the Faltings height from below. Some of these results are corollaries of the partial results towards the Colmez conjecture. Colmez [16] gave one of these bounds for elliptic curves with complex multiplication (a new proof of this is also in [25]). Bost [12] bounded general abelian varieties in terms of the dimension. Deligne [21] calculated the minimal value of the stable Faltings height of elliptic curves, and Löbrich [27] showed that while the Faltings height has no gap above this, the stable Faltings height does. Löbrich's paper also contains a nice introduction to the Faltings height.

Masser and Wüstholz [28] [29] [30] [31] have published a series of papers in which they work towards techniques giving an effective bound in Faltings' theorem. In the first paper, their approach is to study the isogenies between elliptic curve, and one of their results bounds the degree of an isogeny between two elliptic curves in terms of the Faltings height. The second and third papers then discuss applications and generalize to abelian varieties of arbitrary dimension.

## 3.3 Silverman's reformulation

Silverman [34] has showed that the Faltings height of an elliptic curve can be written in terms of other invariants of elliptic curves. We will use this in Chapter 4 to prove Theorem 1.2.

If $h_F(E)$ denotes the original Faltings height for an elliptic curve $E$ defined over a number field $K$, let $H_F(E) = e^{12h_F(E)}$. Recall that $\mathcal{H}$ denotes the complex upper half plane, and that for $\tau \in \mathcal{H}$,

$$\Delta(\tau) = (2\pi)^{12} q_\tau \prod_{n=1}^{\infty} (1 - q_\tau^n)^{24},$$

where $q_\tau = e^{2\pi i \tau}$. As before, let $\Delta_E^{\min}$ denote the minimal discriminant. If $v$ is an

infinite place of $K$, define

$$\varepsilon_v = \begin{cases} 1 & \text{if } v \text{ is real} \\ 2 & \text{if } v \text{ is complex} \end{cases}$$

and choose $\tau_v \in \mathcal{H}$ such that $E(\overline{K_v}) \cong \mathbb{C}/(\mathbb{Z} + \tau_v \mathbb{Z})$.

**Theorem 3.1** (Silverman [34] Proposition 1.1).

$$H_F(E) = e^{[K:\mathbb{Q}]} \frac{|N_{K/\mathbb{Q}} \Delta_E^{\min}|}{\prod_{v|\infty} |\Delta(\tau_v)|^{\varepsilon_v} (\operatorname{Im} \tau_v)^{6\varepsilon_v}}.$$

Since $\Delta(\tau)$ is a modular form of weight 12 for $\operatorname{SL}_2(\mathbb{Z})$, $|\Delta(\tau_v)|(\operatorname{Im} \tau_v)^6$ is an invariant of the action of $\operatorname{SL}_2(\mathbb{Z})$ on $\mathcal{H}$. Theorem 3.1 is normalized differently than in Silverman: we define $\Delta(\tau)$ with a leading term of $(2\pi)^{12}$, which agrees with the normalization of Faltings.

In the case where $K = \mathbb{Q}$, Theorem 3.1 simplifies to

$$H_F(E) = \frac{\Delta_E^{\min}}{|\Delta(\tau)|(\operatorname{Im} \tau)^6},$$

which is the formula for the Faltings height that we will be using in this thesis.

Silverman uses Theorem 3.1 to prove

**Theorem 3.2.** *For all* $\varepsilon > 0$, *there exist* $C_1, C_2(\varepsilon) > 0$ *such that for all elliptic curves* $E$ *over* $\mathbb{Q}$,

$$C_1 H_F(E) \le H_N(E) \le C_2(\varepsilon) H_F(E)^{1+\varepsilon}.$$

This inequality relates the Faltings height to the naive height, but does not show that it is within a bounded factor of the naive height. So counting elliptic curves by naive height does not give good bounds for counting curves by Faltings height.

# Chapter 4

# Counting elliptic curves bounded by the Faltings height

In this chapter we will give the proof of Theorem 1.2. It will bear similarities to parts of Chapter 2, in particular the proof counting the naive height and the discussion of using the minimal discriminant as a height.

Recall that we call $(A, B) \in \mathbb{Z}^2$ *weakly minimal* provided that there is no prime $p$ such that $p^4 \nmid A$ or $p^6 \nmid B$. Every elliptic curve over $\mathbb{Q}$ is isomorphic to a unique elliptic curve $E_{A,B}$ given by the equation $y^2 = x^3 + Ax + B$ where $(A, B) \in \mathbb{Z}^2$ are weakly minimal.

There is a bijection between the set

$$S_X = \{\text{weakly minimal } (A, B) \in \mathbb{Z}^2 \mid 4A^3 + 27B^2 \neq 0, \ H_F(E_{A,B}) < X\}$$

and $\mathcal{E}_{H_F < X}$, given by $(A, B) \mapsto E_{A,B}$. So in this chapter we will prove

$$\#S_X = 12\zeta(10)^{-1}\sigma X^{5/6} + O(X^{1/2}(\log X)^3),$$

which will suffice to prove Theorem 1.2.

Recall that

$$\tau_{A,B} \in \mathcal{H} \text{ such that } E_{A,B}(\mathbb{C}) \cong \mathbb{C}/(\mathbb{Z} + \tau_{A,B}\mathbb{Z})$$

$$\Delta_{A,B} = -16(4A^3 + 27B^2)$$

and let

$$\lambda_{A,B} = \frac{|\Delta_{E_{A,B}}^{\min}|}{|\Delta_{A,B}|}.$$

We will assume without loss of generality that $|\tau_{A,B}| \geq 1$ and $-\frac{1}{2} < \mathrm{Re}\,\tau_{A,B} \leq \frac{1}{2}$.

Note that the first two of these make sense for $A, B \in \mathbb{R}$, while the last requires $A, B \in \mathbb{Q}$. Furthermore, recall from Proposition 2.2 that $\lambda_{A,B}$ is in $\{1, 2^{-12}, 3^{-12}, 6^{-12}\}$ and depends on $A \bmod 6^6$ and $B \bmod 6^6$. In Section 4.4, we will count the number of residue classes giving rise to each possible $\lambda$-value.

By the definition of $\lambda_{A,B}$, Theorem 3.1 is equivalent to

$$H_F(E_{A,B}) = \lambda_{A,B} \frac{|\Delta_{A,B}|}{|\Delta(\tau_{A,B})| \,\mathrm{Im}(\tau_{A,B})^6}.$$

This motivates us to define the function

$$f(A, B) = \left| \frac{\Delta(\tau_{A,B}) \,\mathrm{Im}(\tau_{A,B})^6}{\Delta_{A,B}} \right|^{1/2}$$

for $(A, B) \in \mathbb{R}^2$ such that $\Delta_{A,B} \neq 0$, so that

$$H_F(E_{A,B}) = \frac{\lambda_{A,B}}{f(A, B)^2}.$$

(We inverted and took a square root when defining $f$ only to make later calculations easier.) So for each $\lambda \in \{1, 2^{-12}, 3^{-12}, 6^{-12}\}$ and each residue class with that $\lambda$-value, we want to count the weakly minimal integer points in the residue class that satisfy $\Delta_{A,B} \neq 0$ and $\lambda f(A, B)^{-2} < X$. Therefore define

$$R_{X,\lambda} = \left\{ (A, B) \in \mathbb{R}^2 \,\middle|\, \Delta_{A,B} \neq 0, \frac{\lambda}{f(A,B)^2} < X \right\} \cup \left\{ (A, B) \in \mathbb{R}^2 \mid \Delta_{A,B} = 0 \right\}.$$

Note that since $R_{X,\lambda} = R_{X/\lambda,1}$, we may as well just study $R_X = R_{X,1}$. Replacing $(A, B)$ by $(X^{1/3}A, X^{1/2}B)$ does not change $\tau_{A,B}$, so

$$f(X^{1/3}A, X^{1/2}B) = X^{-1/2}f(A, B).$$

So $R_X$ is $R_1$ scaled by $X^{1/3}$ in the $A$-direction and $X^{1/2}$ in the $B$-direction, and its area is just the area of $R_1$ times $X^{5/6}$.

Section 4.1 calculates the area of $R_1$ (the value denoted $\sigma$), while the rest of the chapter is focused on setting up the proof of Theorem 1.2. In particular:

- Section 4.2 formalizes what we claimed earlier about the minimal discriminant: that outside of a suitable rectangle, the points of $R_X$ are close to the cubic given by $4A^3 + 27B^2$ (the "tentacles" of the region).

- In Section 4.3, we show that any lattice point in $R_X$ with large enough $A, B$ satisfies $4A^3 + 27B^2 = 0$. That is, we need only consider the intersection of $R_X$ with a particular rectangle $R$ — we cut the tentacles short.

- Section 4.4 counts the residue classes for $A$ mod $6^6$ and $B$ mod $6^6$ that give the various $\lambda$.

- Section 4.5 counts for each fixed residue class, the corresponding lattice points in $R_X \cap R$. This is done by applying Theorem 1.1, and showing the length of the boundary of $R_X \cap R$ cannot be longer than perimeter of $R$.

- Section 4.6 uses a Möbius inversion argument as in the proof of Theorem 2.1 to bring this together and conclude the proof of Theorem 1.2.

31

## 4.1 Area of $R_1$

To better calculate the area of $R_1$, we do a change of variables to the parameter $t = \dfrac{A^3}{B^2}$. Define a function

$$f(t) = \left| \frac{\Delta(\tau_t) \operatorname{Im}(\tau_t)^6}{16(4t + 27)} \right|^{1/2},$$

where $\tau_t$ is in the fundamental domain such that $j(\tau_t) = 6912t/(4t + 27)$. Then $R_1$ is by definition the area where

$$\frac{B^2}{f(t)^2} < 1$$

or alternatively, where $-f(t) < B < f(t)$. So with a change of variables we get that

$$\begin{aligned}
\operatorname{Area}(R_1) &= \iint_{R_1} dA \, dB \\
&= \int_{-\infty}^{\infty} \int_{-f(t)}^{f(t)} \left( \frac{1}{3} t^{-2/3} B^{2/3} \right) dB \, dt \\
&= \frac{1}{3} \int_{-\infty}^{\infty} t^{-2/3} \left( \frac{3}{5} B^{5/3} \right) \Bigg|_{-f(t)}^{f(t)} dt \\
&= \frac{2}{5} \int_{-\infty}^{\infty} t^{-2/3} f(t)^{5/3} dt.
\end{aligned}$$

Using a computer to evaluate, we conclude that $\operatorname{Area}(R_1) \approx 29089$. We will denote $\sigma = \operatorname{Area}(R_1)$, and note that $\operatorname{Area}(R_X) = \sigma X^{5/6}$. This is the same $\sigma$ that appears in Theorem 1.2. Note that in the tentacles, near $\Delta_{A,B} = 0$, we have $t \approx -27/4$, and both $\tau_t$ and $f$ grow large.

(If we wish to have a somewhat cleaner integral we can also rewrite it as follows: Use $j(z) = 6912t/(4t + 27)$ to do a change of variables to a complex number $z$ over an appropriate curve. Rearranging and applying that $\Delta(z) = (2\pi)^6 \frac{j'(z)^6}{j(z)^4(j(z)-1728)^3}$, the integral is

$$\int_\gamma |\Delta(z)| \operatorname{Im}(z)^5 \frac{j'(z)}{|j'(z)|} \, dz$$

where $\gamma$ consists of all $z \in \mathcal{F}$ such that $j(z)$ is real, where $\mathcal{F}$ is a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ acting on $\mathcal{H}$.)
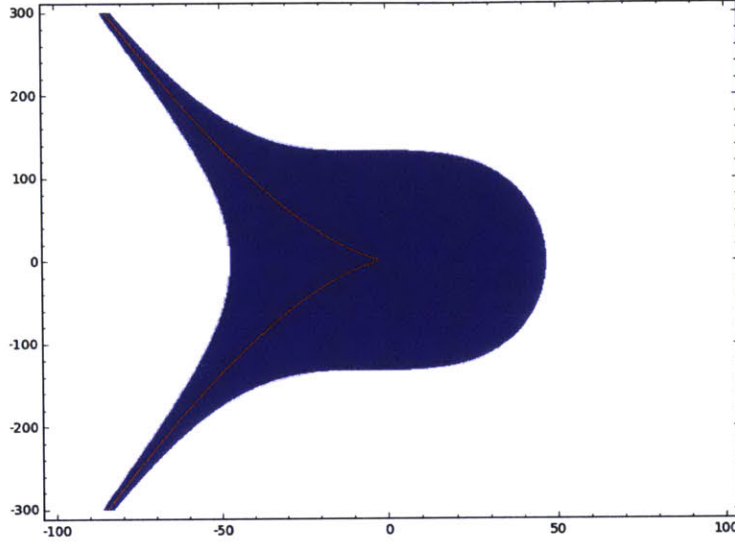
Figure 4-1: The region $R_1$ in blue, with the cubic $4A^3 + 27B^2$ in red

## 4.2 The rough shape of $R_X$

**Lemma 4.1.** *The function $|\Delta(\tau)\operatorname{Im}(\tau)^6|$ is bounded on the complex upper half-plane.*

*Proof.* Since $\Delta(\tau)\operatorname{Im}(\tau)^6$ is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$, assume that $\tau$ is in the standard fundamental domain, so that $|\tau| \geq 1$ and $|\operatorname{Re}(\tau)| \leq 1/2$. By definition

$$\Delta(\tau) = (2\pi)^{12}q\prod_{n=1}^{\infty}(1-q^n)^{24}$$

where $q = e^{2\pi i \tau}$. So as $\operatorname{Im}(\tau) \to \infty$, $q \to 0$, and $|\Delta(\tau)| = O(q) = O(e^{-2\pi \operatorname{Im}(\tau)})$. Thus, as $\operatorname{Im}(\tau)$ grows large, $|\Delta(\tau)|\operatorname{Im}(\tau)^6$ tends to zero. So for any constant $b_0 > 0$, there is a constant $b_1 > 0$ such that if $\operatorname{Im}(\tau) > b_1$, then $|\Delta(\tau)\operatorname{Im}(\tau)^6| < b_0$. Since the set of $\tau$ such that $\operatorname{Im}(\tau) \leq b_1$ is a compact set, there is a bound on that domain as well. $\square$

**Corollary 4.1.1.** *Let $C$ be a positive constant such that $|\Delta(\tau)|\operatorname{Im}(\tau)^6 < C$ for all $\tau \in \mathcal{H}$. If $(A, B) \in R_X$ then $|\Delta_{A,B}| < CX$.*

The above follows immediately from Lemma 4.1. The following lemma will show that points in $R_X$ are either small compared to $X$ or close to the curve given by $\Delta_{A,B} = 0$.

33

**Lemma 4.2.** *Let $(A, B) \in R_X$. If $|B|^2 < CX/27$, then $|A|^3 < CX/2$. Otherwise,*
$A = -cB^{2/3} + \varepsilon X B^{-4/3}$ *for some* $|\varepsilon| < C$ *and* $c = (27/4)^{1/3}$.

*Proof.* Suppose $|B|^2 < CX/27$ and $|A|^3 \geq CX/2$, then

$$
\begin{aligned}
|\Delta_{A,B}| &= 16|4A^3 + 27B^2| \\
&\geq 16(4|A^3| - 27|B|^2) \\
&\geq 16(2CX - CX) \\
&= 16CX \geq CX,
\end{aligned}
$$

and thus $(A, B) \notin R_X$.

Let $B \neq 0$ and $A = -cB^{2/3} + \varepsilon X B^{-4/3}$ with $|\varepsilon| \geq C$. Then

$$
|\Delta_{A,B}| = 64|\varepsilon|X(3c^2 - 3c\varepsilon X B^{-2} + \varepsilon^2 X^2 B^{-4}).
$$

Well the polynomial $x^2 - 3cx + 3c^2 \geq 1$ for all $x \in \mathbb{R}$ and by assumptions $|\varepsilon| > C$, so $|\Delta_{A,B}| > 64CX > CX$. This implies $(A, B) \notin R_X$. $\qquad\square$

Note that this tells us only in a very rough sense that points in $R_X$ are close to the cubic given by $\Delta_{A,B} = 0$. In the next section, we will give a stronger version of this statement which holds less generally. (And while Lemma 4.2 could also be shown when using the minimal discriminant as a height, the analogue in the next section will not.)

## 4.3   Bounding the size of lattice points

In this section we will show that all large enough integral $(A, B) \in R_X$ have $\Delta_{A,B} = 0$. This is done in two steps: first, we quantify how close a point has to be to curve $\Delta_{A,B} = 0$ to have the property $|\Delta_{A,B}| < 1$. Then we show that, for large enough $A$, a point farther from the curve cannot be in $R_X$.

**Lemma 4.3.** *Assume $B \neq 0$. Let $\varepsilon_0$ be the positive real root of $64x(3c^2 + x^2) - 3/4$, recalling that $c = (27/4)^{1/3}$ (this root is approximately 0.0011). For all $|\varepsilon| < \varepsilon_0$, if $A = -cB^{2/3} + \varepsilon B^{-4/3}$ then $|\Delta_{A,B}| < 1$.*

*Proof.* Under the assumptions, we have that

$$
\begin{aligned}
|\Delta_{A,B}| &= 64|\varepsilon||3c^2 - 3c\varepsilon B^{-2} + \varepsilon^2 B^{-4}| \\
&\leq 64|\varepsilon|(3c^2 + 3c|\varepsilon|B^{-2} + \varepsilon^2 B^{-4}) \\
&< 64\varepsilon_0(3c^2 + \varepsilon_0^2) + 192\varepsilon_0^2 c \\
&= 3/4 + 192\varepsilon_0^2 c \\
&< 1
\end{aligned}
$$

since $192\varepsilon_0^2 c$ is much smaller than $1/4$. $\qquad\square$

The following lemma is a strong version of Lemma 4.2 that only holds for large enough $X$ and $A$.

**Lemma 4.4.** *There exist positive constants $M$, $N$ such that for all $X \geq M$, if $|A| > NX^{1/3}(\log X)^2$, where $A = -cB^{2/3} + \varepsilon B^{-4/3}$ and $|\varepsilon| \geq \varepsilon_0$, then $(A, B) \notin R_X$.*

*Proof.* Note that

$$
\begin{aligned}
f(A, B)^{-2} &= \frac{|\Delta_{A,B}|}{|\Delta(\tau_{A,B})| \operatorname{Im}(\tau_{A,B})^6} \\
&= \frac{1728|4A|^3}{|\Delta(\tau_{A,B})||j(\tau_{A,B})| \operatorname{Im}(\tau_{A,B})^6}
\end{aligned}
$$

Recall that $\tau_{A,B}$ in the usual fundamental domain, so that $|\operatorname{Re}(\tau_{A,B})| \leq 1/2$ and $|\tau_{A,B}| \geq 1$. Since $|\Delta(\tau_{A,B})||j(\tau_{A,B})| \to 1$ as $\tau \to i\infty$, there is some positive constant bounding $|\Delta(\tau_{A,B})||j|$ from above, so

$$
f(A, B)^{-2} \geq c_1 \frac{|A|^3}{\operatorname{Im}(\tau_{A,B})^6}
$$

for some constant $c_1 > 0$.

If $\text{Im}(\tau) \leq 1$, this means $f(A, B)^{-2} \geq c_1 |A|^3 > c_1 N^3 X (\log X)^6$. So as long as $M$ is large enough so $c_1 N^3 (\log M)^6 > 1$, this tells us that $f(A, B)^{-2} \geq X$ and thus $(A, B) \notin R_X$. (We will need to be careful to make sure this is compatible with how we choose $N$.)

Assume $\text{Im}(\tau) > 1$. Note that

$$|\Delta_{A,B}| = 64|\varepsilon||3c^2 - 3c\varepsilon B^{-2} + \varepsilon^2 B^{-4}|.$$

The polynomial $x^2 - 3cx + 3c^2$ is positive for all real numbers, with a minimal value of $f(3c/2) \geq 2$, which implies that $|\Delta_{A,B}| \geq 128\varepsilon_0 > 0.1$. Since $j = 1728(4A)^3 \Delta_{A,B}^{-1}$, this means that $|j| \leq c_2 |A|^3$ for some positive constant $c_2$.

But since $|q_\tau| = e^{-2\pi \text{Im}(\tau)}$, and for $\text{Im}(\tau) > 1$, $|\log(|j|) - \log(|q_\tau^{-1}|)|$ is bounded, it follows that

$$\text{Im}(\tau) < c_3 \log(|A|) + c_4$$

for some positive constants $c_3, c_4$.

So

$$f(A, B)^{-2} \geq c_1 \frac{|A|^3}{\text{Im}(\tau_{A,B})^6}$$
$$\geq c_1 \frac{|A|^3}{(c_3 \log(|A|) + c_4)^6}.$$

If we take $N$ such that $3c_1 N^3 > 2c_3$, then there is some constant $c_5$ such that if $X > c_5$ and $|A| > NX^{1/3}(\log X)^2$,

$$c_1 \frac{|A|^3}{(c_3 \log(|A|) + c_4)^6} > X.$$

This means that $(A, B) \notin R_X$.

So we can take any $N$ such that $3c_1 N^3 > 2c_3$ and $M$ such that $c_1 N^3 (\log M)^6 > 1$ and $M > c_5$, and the lemma holds for these. $\square$

Together Lemmas 4.3 and 4.4 tell us that for $X \geq M$, if $|A| > NX^{1/3}(\log X)^2$, then either $|\Delta_{A,B}| < 1$ or $(A, B) \notin R_X$. Thus, the only integer points in $R_X$ where

36

$|A| > NX^{1/3}(\log X)^2$ are those where $\Delta_{A,B} = 0$.

## 4.4  Weakly minimal curves not minimal at 2 or 3

Recall that we defined $R_X$ so that $H_F(E_{A,B}) < X$ is equivalent to $(A, B) \in R_{X/\lambda_{A,B}}$. In this section, we more carefully examine $\lambda_{A,B}$.

Recall that we defined $\lambda_{A,B}$ so that it relates the minimal and polynomial discriminants:

$$|\Delta_{E_{A,B}}^{\min}| = \lambda_{A,B}|\Delta_{A,B}|.$$

Since we are requiring that $(A, B)$ be weakly minimal, it follows from Proposition 2.2 that $\lambda_{A,B}$ is $1$, $2^{-12}, 3^{-12}$, or $6^{-12}$, depending respectively on whether the model $E_{A,B}$ is minimal everywhere, fails to be minimal only at 2, fails only at 3, or fails at both 2 and 3.

**Proposition 4.5.** *Let $E_{A,B}$ be the elliptic curve given by the equation*

$$y^2 = x^3 + Ax + B. \tag{4.1}$$

*The equation (4.1) fails to be a minimal equation for $E_{A,B}$ at 2 if and only if $(A, B)$ mod $2^6$ is one of*

| | | | | | |
|---|---|---|---|---|---|
| $(13, 14)$ | $(45, 46)$ | $(29, 30)$ | $(61, 62)$ | $(5, 22)$ | $(37, 54)$ |
| $(21, 38)$ | $(53, 6)$ | $(0, 16)$ | $(32, 16)$ | $(16, 16)$ | $(48, 16)$. |

*The equation (4.1) fails to be a minimal equation for $E_{A,B}$ at 3 if and only if $3^3 \mid A$, $3^3 \mid B$, and $(A/3^3 \bmod 3^2, B/3^3 \bmod 3^3)$ is*

| | | | | | |
|---|---|---|---|---|---|
| $(8, 2)$ | $(8, 25)$ | $(5, 11)$ | $(2, 7)$ | $(5, 16)$ | $(2, 20)$. |

*(Note we reduced $A/3^3$ only mod $3^2$.)*

*In particular, there are 12 (resp. 18) classes of $(A, B)$ mod $2^6$ (resp. $3^6$) for which*

*equation (4.1) fails to be a minimal equation at 2 (resp. 3).*

Noting that the situations for $p = 2, 3$ are independent, this is just an application of Tate's algorithm, as described in [35] (which we supplemented with sage [37]).

Let $\mathrm{Cl}_\lambda$ be the set of residue classes $(A_0, B_0) \in (\mathbb{Z}/6^6\mathbb{Z})^2$ such that if $(A, B) \in \mathbb{Z}^2$ reduces $(A_0, B_0$ then $(A, B)$ is weakly minimal at 2 and 3, and $\lambda_{A,B} = \lambda$. Proposition 4.5 tells us what $\# \mathrm{Cl}_\lambda$ is for each $\lambda$.

| Model is... | Factor $\lambda$ | Size of $\mathrm{Cl}_\lambda$ |
|---|---|---|
| Minimal everywhere | 1 | $(2^{12} - 12 - 2^2)(3^{12} - 18 - 3^2)$ |
| Minimal except at 2 | $2^{-12}$ | $12 \times (3^{12} - 18 - 3^2)$ |
| Minimal except at 3 | $3^{-12}$ | $(2^{12} - 12 - 2^2) \times 18$ |
| Minimal except at 2 and 3 | $6^{-12}$ | $12 \times 18$ |

Fix $(A_0, B_0) \in \mathrm{Cl}_\lambda$. For weakly minimal $(A, B)$ reducing to $(A_0, B_0)$ mod $6^6$, $H(E_{A,B}) < X$ is actually equivalent to $(A, B) \in R_{X/\lambda}$ (this is a consequence of Proposition 2.2). If we calculate

$$\#\{(A, B) \in \mathbb{Z}^2 \mid (A, B) \equiv (A_0, B_0) \bmod 6^6, (A, B) \text{ weakly minimal}, (A, B) \in R_{X/\lambda}\},$$

and sum over all $(A_0, B_0) \in \mathrm{Cl}_\lambda$ and $\lambda$, this will give the number of elliptic curves of Faltings height less than $X$ up to isomorphism. This is done in the next two sections.

## 4.5 Counting weakly minimal lattice points of a fixed residue class

**Proposition 4.6.** *Fix two residue classes $A_0, B_0$ mod $6^6$ such that $(A_0, B_0)$ is weakly minimal with respect to 2 and 3. Then there is a constant $M$ such that the number of lattice points in $R_X$ reducing mod $6^6$ to $(A_0, B_0)$ and $\Delta_{A,B} \neq 0$ is*

$$6^{-12}\sigma X^{5/6} + O(X_M^{1/2}(\log X_M)^3),$$

*where $X_M = \max(X, M)$ and $\sigma = \text{Area}(R_1)$.*

*Proof.* Since we are counting lattice points, we want to apply Theorem 1.1. Since the error term is dependent on the length of the boundary, we will need to use a region that has finite boundary and keep track of that boundary. (In the end, we will be doing this to the region scaled by $6^{-6}$ since counting all integral points in that rescaled region will be approximately the same as counting the points that satisfy the particular congruence. For simplicity, we ignore the rescaling for now.)

Let $X_M = \max(M, X)$. Let

$$R'_X = \left\{ (A, B) \in R_X \,\middle|\, |A| \leq N X_M^{1/3} (\log X_M)^2 \right\},$$

where $N, M$ are from Lemma 4.4. If $X \geq M$, then Lemmas 4.3 and 4.4 tells us that the lattice points $(A, B) \in R_X \setminus R'_X$ satisfy $\Delta_{A,B} = 0$. If $X < M$, then $R_X \subseteq R_M$, so $R_X \setminus R'_X \subseteq R_M \setminus R'_M$, and so also $\Delta_{A,B} = 0$.

Recall that $C$ is the bound on $|\Delta(\tau) \text{Im}(\tau)^6|$ from Lemma 4.1, and $c = (27/4)^{1/3}$. Using Lemma 4.2, there is a $\beta_0 > 0$ such that if $|A| \leq N X_M^{1/3} (\log X_M)^2$ and $(A, B) \in R_X$, $|B| \leq \beta_0 N^{3/2} X_M^{1/2} (\log X_M)^3$. Thus, we can think of $R'_X$ as the intersection of $R_X$ with a rectangle. The boundary of $R'_X$ is contained in the union of the boundary of that rectangle (and thus no worse than $O(X_M^{1/2} (\log X_M)^3)$) and the boundary of $R_X$ (that is, the boundary of the closure of $R_X$). So we need only show that the curve given by the boundary of $R_X$ inside that rectangle has length less than $O(X_M^{1/2} (\log X_M)^3)$. The lemmas that follow establish this.

**Lemma 4.7.** *For any positive $\alpha, \beta \in \mathbb{R}$, the boundary of $R_X$ in the rectangle $|A| \leq \alpha X^{1/3}$ and $|B| \leq \beta X^{1/2}$ is $O(X^{1/2})$.*

*Proof.* Since $R_X$ is $R_1$ scaled by $X^{1/3}$ in the $A$-axis and $X^{1/2}$ in the $B$-axis, it suffices to show that the boundary of $R_1$ in any rectangle is bounded. Since the scaling cannot make the boundary longer than the scaling itself, it will follow that the boundary in $R_X$ is $O(X^{1/2})$.

The boundary of $R_1$ is given by the zeros of

$$|\Delta_{A,B}| - |\Delta(\tau_{A,B})| \operatorname{Im}(\tau_{A,B})^6$$

This function is real analytic off $\Delta_{A,B} = 0$ (where it is not defined), and thus in any compact set, its zero set is rectifiable (see [23] 3.4.10). □

Note that [23] 3.4.10 even implies that the boundary of $R_X$ in any rectangle is finite, but since we want to know how big our error is, we want to calculate it more carefully.

We must then consider the boundary in the region $\alpha X^{1/3} < |A|$ or $\beta X^{1/2} < |B|$ (noting that we can fix $\alpha, \beta$ as we like). We will first consider $R_1$ and then generalize. For $(A, B)$ such that $\Delta_{A,B} \neq 0$, if we choose $\tau_{A,B}$ in the standard fundamental domain, then $q$ is real valued, and so is $\Delta(\tau_{A,B})$. In fact, it shares the same sign as $\Delta_{A,B}$. So

$$F(A, B) = \Delta_{A,B} - \Delta(\tau_{A,B}) \operatorname{Im}(\tau_{A,B})^6 \tag{4.2}$$

is a real-valued function the zero set of which is the boundary of $R_1$. We will show that there is a $\beta$ such that for all $B > \beta$, each partial of $F$ on the boundary of $R_1$ has a constant sign, and the same hold for all $B < -\beta$. First, this implies that since $F$ is defined everywhere but where $\Delta_{A,B} = 0$, there can be at most two components of the boundary (one for each connected region separated by the cubic). Second, it tells us that where $|B| > \beta$, a connected curve given by the zero set of $F$ (and thus also the corresponding boundary of $R_X$) is such that its length in any rectangle is less than or equal to the length of the boundary of the rectangle. (This follows from the triangle inequality. This controls how much a curve can change direction and thus bounds its length.)

**Lemma 4.8.** *There is a $\beta > 0$ such that if $(A, B) \in \mathbb{R}^2$ such that $F(A, B) = 0$ and $|B| > \beta$, then the gradient $\nabla F$ at $(A, B)$ is in the third quadrant if $B > 0$ and in the second quadrant if $B < 0$.*

*Proof.* We calculate the partial derivatives of $F$ on the curve where $F = 0$. The choice

of $\tau_{A,B}$ implies that $q_{A,B} = e^{2\pi i \tau_{A,B}}$ is real valued, so we can write $F$ and $\Delta(\tau_{A,B})$ in terms of $q_{A,B}$ using $2\pi \operatorname{Im}(\tau_{A,B}) = \log |q_{A,B}|$ in (4.2).

Note that Lemma 4.1.1 bounds the second term in the definition of $F$ by $C$ (since we currently considering $R_1$), so $|\Delta_{A,B}|$ is bounded on $F(A, B) = 0$, and thus as $|B|$ grows large we have $4A^3 \approx 27B^2$. So $A^3$ grows large and thus also $j(E_{A,B})$, from which it follows that $q_{A,B} \to 0$.

Define

$$J(\tau_{A,B}) = \frac{1}{j(E_{A,B})} = \frac{1}{1728} + \frac{B^2}{256A^3}.$$

By the chain rule,

$$\frac{\partial F}{\partial A} = \frac{\partial \Delta_{A,B}}{\partial A} - \frac{1}{(2\pi)^6} \left( \frac{d(\Delta(q)(\log |q|)^6)}{dq} \right) \left( \frac{dq}{dJ} \right) \left( \frac{\partial J}{\partial A} \right).$$

Since $q$ can be written as a convergent power series in $j^{-1} = J$ with leading term $J$, it follows that $\frac{dq}{dJ} = 1 + O(J) = 1 + O(q)$. Similarly, since $\Delta(q)$ is a convergent power series in $q$ with leading term $q$, we also have $\frac{d\Delta(q)}{dq} = 1 + O(q)$. Using that $q$ is real, we get that

$$\frac{d\Delta(q)(\log |q|)^6}{dq} = (1 + O(q))(\log |q|)^6 + 6(1 + O(q))(\log |q|)^5$$

$$= (\log |q|)^6 + O((\log |q|)^5).$$

Lastly, $\frac{\partial \Delta_{A,B}}{\partial A}$ and $\frac{\partial J}{\partial A}$ are straightforward calculations. Putting this together, we obtain

$$\frac{\partial F}{\partial A} = -192A^2 - \frac{1}{(2\pi)^6} (\log |q_{A,B}|)^6 \left( -\frac{3B^2}{256A^4} \right) (1 + o(1))$$

and a similar calculation gives

$$\frac{\partial F}{\partial B} = -864B - \frac{1}{(2\pi)^6} (\log |q_{A,B}|)^6 \left( \frac{B}{128A^3} \right) (1 + o(1)).$$

We want to know the value of these partial derivatives where $F(A, B) = 0$, which means that

$$\Delta_{A,B} = \Delta(q_{A,B})(\log |q_{A,B}|)^6. \tag{4.3}$$

41

Additionally, for small $q$, $\Delta(q) = q(1 + o(1)) = 1/j(1 + o(1))$; substituting this into (4.3) yields

$$(\log|q|)^6 = j\Delta_{A,B}(1 + o(1)) = -1728(4A)^3(1 + o(1)).$$

Lastly, it follow from Lemma 4.2 (for $R_1$), that $B^2/A^3$ approaches $-4/27$. We put these two estimates into (4.2) to conclude that

$$\frac{\partial F}{\partial A} = -A^2 \left(192 - \frac{3}{\pi^6}(1 + o(1))\right)$$

and

$$\frac{\partial F}{\partial B} = -B \left(864 - \frac{27}{2\pi^6}(1 + o(1))\right) \qquad \square$$

By Lemma 4.2, we can chose $\alpha$ such that if $(A, B) \in R_X$ and $|B| \leq \beta X^{1/2}$ then $|A| \leq \alpha X^{1/3}$. Thus, for points where $|B| < \beta X^{1/2}$, Lemma 4.7 applies, and for those where $|B| \geq \beta X^{1/2}$, the above argument applies, and applying it with the two rectangles where $\beta X^{1/2} < |B| < \beta_0 N^{3/2} X^{1/2}(\log X)^3$ and $|A| < NX^{1/3}(\log X)^2$, the boundary of $R_X$ in that region is $O(X^{1/2}(\log X)^3)$.

Having shown that the boundary of $R'_X$ is $O(X_M^{1/2}(\log X_M)^3)$, we turn now to the lattice points of $R'_X$. If we wanted all lattice points, we could just apply Theorem 1.1, which tells us there should be $\text{Area}(R'_X) + O(X_M^{1/2}(\log X_M)^3)$. The area of $R_X$ is $\sigma X^{5/6}$, and this differs from the area of $R'_X$ by the region where $|B| > \beta_0 N^{3/2} X_M^{1/2}(\log X_M)^3$. Lemma 4.4 tells us that for a fixed $B$ in that region, the width of the region is $O(B^{-4/3})$. So the area of the difference between $R_X$ and $R'_X$ (including both negative and positive $B$) must be bounded by

$$2\int_{\beta_0 N^{3/2} X_M^{1/2}(\log X_M)^3}^{\infty} O(B^{-4/3})dB = O(X_M^{-1/6}(\log X_M)^{-1})$$

so the area of $R'_X$ is $\sigma X^{5/6} + O(X_M^{-1/6}(\log X_M)^{-1})$. We conclude that the total number of lattice points in $R'_X$ is $\sigma X^{5/6} + O(X_M^{1/2}(\log X_M)^3)$.

Lastly, we want to apply the entirety of the above argument after rescaling in both directions. The number of lattice points of a particular residue class mod $m$ in

a region is the number of all lattice points in the same region scaled by $1/m$ in both directions and translated appropriately. So scaling the region will scale area by $6^{-12}$ and length by $6^{-6}$. Translating will change the number of lattice points by at most the boundary, so the final result must be $6^{-12}\sigma X^{5/6} + O(X_M^{1/2}(\log X_M)^3)$.  □


## 4.6   Concluding the proof of the main result

Recall that $R_X$ was defined so that what we want to count for Theorem 1.2 was

$$S_X = \{\text{weakly minimal } (A, B) \in \mathbb{Z}^2 \mid \Delta_{A,B} \neq 0, (A, B) \in R_{X/\lambda_{A,B}}\}.$$

Let

$$S_{X,\lambda} = \{(A, B) \in S_X \mid \lambda_{A,B} = \lambda\}$$
$$= \{\text{weakly minimal } (A, B) \in \mathbb{Z}^2 \cap R_{X/\lambda} \mid \Delta_{A,B} \neq 0, \ \lambda_{A,B} = \lambda\},$$

so that $S_X$ is a disjoint union of $S_{X,1}, S_{X,2^{-12}}, S_{X,3^{-12}}$, and $S_{X,6^{-12}}$.

We now go about calculating $\#S_{X,\lambda}$. Recall from Section 4.4 that $\mathrm{Cl}_\lambda$ is the set of residue classes mod $6^6$ such that if $(A, B)$ reduces to a class in $\mathrm{Cl}_\lambda$, then $(A, B)$ is weakly minimal at 2 and 3 and $\lambda_{A,B} = \lambda$. A Möbius inversion argument shows that

$$\sum_{\substack{d^4|A \\ d^6|B}} \mu(d) = \begin{cases} 1 & \text{if } (A, B) \text{ is weakly minimal} \\ 0 & \text{otherwise,} \end{cases}$$

so

$$\#S_{X,\lambda} = \sum_{\substack{(A,B)\in\mathbb{Z}^2\cap R_{X/\lambda} \\ (A,B) \bmod 6^6 \in \mathrm{Cl}_\lambda \\ \Delta_{A,B}\neq 0}} \ \sum_{\substack{d^4|A \\ d^6|B}} \mu(d).$$

We want to switch the sums so that we end up summing over $d$ first. To be able to describe the inner sum, let $d^*\mathrm{Cl}_\lambda$ be the set of residue classes $(\bar{a}, \bar{b}) \in (\mathbb{Z}/6^6\mathbb{Z})^2$ such that $(d^4\bar{a}, d^6\bar{b}) \in \mathrm{Cl}_\lambda$; note that $\#d^*\mathrm{Cl}_\lambda = \#\mathrm{Cl}_\lambda$. Note that if $(A, B) \bmod 6^6 \in \mathrm{Cl}_\lambda$,

43

it is also true that $2 \nmid d, 3 \nmid d$.

If $X > M$, there are no $(A, B) \in \mathbb{Z}^2 \cap R_{X/\lambda}$ such that $|A| > N(X/\lambda)^{1/3}(\log(X/\lambda))^2$ (proved in Lemmas 4.3 and 4.4), so $d \leq N^{1/4}(X/\lambda)^{1/12}(\log(X/\lambda))^{1/2}$. There is a bijection between

$$\{(A, B, d) \in \mathbb{Z}^3 \mid (A, B) \in R_{X/\lambda}, (A, B) \bmod 6^6 \in \mathrm{Cl}_\lambda, \Delta_{A,B} \neq 0, d^4 \mid A, d^6 \mid B\}$$

and

$$\{(a, b, d) \in \mathbb{Z}^3 \mid (a, b) \in R_{d^{-12}X/\lambda}, (a, b) \bmod 6^6 \in d^* \mathrm{Cl}_\lambda, \Delta_{a,b} \neq 0,$$
$$d \leq N^{1/4}(X/\lambda)^{1/12}(\log(X/\lambda))^{1/2}, 2 \nmid d, 3 \nmid d\}$$

given by $(A, B, d) \mapsto (Ad^{-4}, Bd^{-6}, d)$.

So

$$\#S_{X,\lambda} = \sum_{\substack{d < N^{1/4}(X/\lambda)^{1/12}(\log(X/\lambda))^{1/2} \\ 2 \nmid d, 3 \nmid d}} \mu(d) \sum_{\substack{(a,b) \in \mathbb{Z}^2 \cap R_{d^{-12}X/\lambda} \\ (a,b) \bmod 6^6 \in d^* \mathrm{Cl}_\lambda \\ \Delta_{a,b} \neq 0}} 1$$

$$= \sum_{\substack{d < N^{1/4}(X/\lambda)^{1/12}(\log(X/\lambda))^{1/2} \\ 2 \nmid d, 3 \nmid d}} \mu(d) \left( \frac{\# \mathrm{Cl}_\lambda}{6^{12}} \sigma(d^{-12}X/\lambda)^{5/6} + \right.$$

$$\left. O\left( \max(M, d^{-12}X/\lambda)^{1/2} \log(\max(M, d^{-12}X/\lambda))^3 \right) \right)$$

by Proposition 4.6. As in the proof of Theorem 2.1, it can be shown that

$$\sum_{\substack{d < Y \\ 2 \nmid d, 3 \nmid d}} \mu(d) d^{-10} = \frac{\zeta(10)^{-10}}{(1 - 2^{-10})(1 - 3^{-10})} + O(Y^{-9}), \tag{4.4}$$

so applying this to the leading term we get

$$\#S_{X,\lambda} = \frac{\# \mathrm{Cl}_\lambda}{\lambda^{5/6}6^{12}} \frac{\zeta(10)^{-10}}{(1 - 2^{-10})(1 - 3^{-10})} \sigma X^{5/6}$$

44

with an error of

$$\sum_{\substack{d < N^{1/4}(X/\lambda)^{1/12}(\log(X/\lambda))^{1/2} \\ 2 \nmid d,\, 3 \nmid d}} \mu(d) \max(M, d^{-12}X/\lambda)^{1/2} \log(\max(M, d^{-12}X/\lambda))^3$$

(which dominates the error coming from applying Equation 4.4). With some algebraic manipulation and another application of Equation 4.4, it can be shown that this error is

$$O(X^{1/2}(\log X)^3).$$

Summing over the four possible values of $\lambda$ using the calculations of $\#\operatorname{Cl}_\lambda$ from Section 4.4, this gives us

$$\#S_X = 12\zeta(10)^{-1}\sigma X^{5/6} + O(X^{1/2}(\log X)^3),$$

as desired.

# Bibliography

[1] Fabrizio Andreatta, Eyal Z. Goren, Benjamin Howard, and Keerthi Mada-
pusi Pera. Faltings heights of abelian varieties with complex multiplication.
Preprint. http://arxiv.org/abs/1508.00178v2.

[2] Andrea Bandini, Ignazio Longhi, and Stefano Vigni. Torsion points on elliptic
curves over function fields and a theorem of Igusa. *Expo. Math.*, 27(3):175–209,
2009.

[3] Ebru Bekyel. The density of elliptic curves having a global minimal Weierstrass
equation. *J. Number Theory*, 109(1):41–58, 2004.

[4] Manjul Bhargava. The density of discriminants of quartic rings and fields. *Ann.
of Math. (2)*, 162(2):1031–1063, 2005.

[5] Manjul Bhargava. Higher composition laws and applications. In *International
Congress of Mathematicians. Vol. II*, pages 271–294. Eur. Math. Soc., Zürich,
2006.

[6] Manjul Bhargava. Mass formulae for extensions of local fields, and conjectures on
the density of number field discriminants. *Int. Math. Res. Not. IMRN*, (17):Art.
ID rnm052, 20, 2007.

[7] Manjul Bhargava. The density of discriminants of quintic rings and fields. *Ann.
of Math. (2)*, 172(3):1559–1591, 2010.

[8] Manjul Bhargava and Arul Shankar. Binary quartic forms having bounded in-
variants, and the boundedness of the average rank of elliptic curves. *Ann. of
Math. (2)*, 181(1):191–242, 2015.

[9] Manjul Bhargava and Arul Shankar. Ternary cubic forms having bounded in-
variants, and the existence of a positive proportion of elliptic curves having rank
0. *Ann. of Math. (2)*, 181(2):587–621, 2015.

[10] Manjul Bhargava, Arul Shankar, and Xiaoheng Jerry Wang. Geometry-of-
numbers methods over global fields I: Prehomogenous vector spaces. Preprint.
http://arxiv.org/abs/1512.03035v1.

[11] Manjul Bhargava, Arul Shankar, and Xiaoheng Jerry Wang. Geometry-of-
numbers methods over global fields II: Coregular representations. Preprint.

[12] Jean-Benoît Bost. Intrinsic heights of stable varieties and abelian varieties. *Duke Math. J.*, 82(1):21–70, 1996.

[13] Armand Brumer. The average rank of elliptic curves. I. *Invent. Math.*, 109(3):445–472, 1992.

[14] Armand Brumer and Oisín McGuinness. The behavior of the Mordell–Weil group of elliptic curves. *Bulletin (New Series) of the American Mathematical Society*, 23(2):375–382, 10 1990.

[15] Pierre Colmez. Périodes des variétés abéliennes à multiplication complexe. *Ann. of Math. (2)*, 138(3):625–683, 1993.

[16] Pierre Colmez. Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe. *Compositio Math.*, 111(3):359–368, 1998.

[17] Gary Cornell and Joseph H. Silverman, editors. *Arithmetic geometry*. Springer-Verlag, New York, 1986. Papers from the conference held at the University of Connecticut, Storrs, Connecticut, July 30–August 10, 1984.

[18] Harold Davenport. On a principle of Lipschitz. *J. London Math. Soc.*, 26:179–183, 1951.

[19] Harold Davenport. Corrigendum: "On a principle of Lipschitz". *J. London Math. Soc.*, 39:580, 1964.

[20] Harold Davenport and Hans Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London Ser. A*, 322(1551):405–420, 1971.

[21] Pierre Deligne. Preuve des conjectures de Tate et de Shafarevitch (d'après G. Faltings). *Astérisque*, (121-122):25–41, 1985. Seminar Bourbaki, Vol. 1983/84.

[22] Gerd Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983.

[23] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[24] Robert Harron and Andrew Snowden. Counting elliptic curves with prescribed torsion. *J. Reine Angew. Math., to appear*. http://arxiv.org/abs/1311.4920v1.

[25] Kai Köhler and Damian Roessler. A fixed point formula of Lefschetz type in Arakelov geometry. IV. The modular height of C.M. abelian varieties. *J. Reine Angew. Math.*, 556:127–148, 2003.

[26] Edmund Landau. *Vorlesungen über Zahlentheorie*, volume 2. S. Hirzel, Leipzig, Germany, 1927.

[27] Steffen Löbrich. A gap in the spectrum of the Faltings height. *Journal de Théorie des Nombres de Bordeaux, to appear.* http://arxiv.org/abs/1505.00602v2.

[28] David Masser and Gisbert Wüstholz. Estimating isogenies on elliptic curves. *Invent. Math.*, 100(1):1–24, 1990.

[29] David Masser and Gisbert Wüstholz. Isogeny estimates for abelian varieties, and finiteness theorems. *Ann. of Math. (2)*, 137(3):459–472, 1993.

[30] David Masser and Gisbert Wüstholz. Factorization estimates for abelian varieties. *Inst. Hautes Études Sci. Publ. Math.*, (81):5–24, 1995.

[31] David Masser and Gisbert Wüstholz. Refinements of the Tate conjecture for abelian varieties. In *Abelian varieties (Egloffstein, 1993)*, pages 211–223. de Gruyter, Berlin, 1995.

[32] Andrew Obus. On Colmez's product formula for periods of CM-abelian varieties. *Math. Ann.*, 356(2):401–418, 2013.

[33] Bjorn Poonen. Average rank of elliptic curves [after Manjul Bhargava and Arul Shankar]. *Astérisque*, (352):Exp. No. 1049, viii, 187–204, 2013. Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.

[34] Joseph H. Silverman. Heights and elliptic curves. In *Arithmetic geometry (Storrs, Conn., 1984)*, pages 253–265. Springer, New York, 1986.

[35] Joseph H. Silverman. *Advanced topics in the arithmetic of elliptic curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[36] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[37] William A. Stein et al. *Sage Mathematics Software (Version 6.3)*. The Sage Development Team, 2014. http://www.sagemath.org.

[38] Lucien Szpiro and Emmanuel Ullmo. Variation de la hauteur de Faltings dans une classe de $\overline{\mathbf{Q}}$-isogénie de courbe elliptique. *Duke Math. J.*, 97(1):81–97, 1999.

[39] Mark Watkins. Some heuristics about elliptic curves. *Experiment. Math.*, 17(1):105–125, 2008.

[40] Tonghai Yang. The Chowla-Selberg formula and the Colmez conjecture. *Canad. J. Math.*, 62(2):456–472, 2010.

[41] Tonghai Yang. Arithmetic intersection on a Hilbert modular surface and the Faltings height. *Asian J. Math.*, 17(2):335–381, 2013.

[42] Xinyi Yuan and Shou-Wu Zhang. On the averaged Colmez conjecture. Preprint. http://arxiv.org/abs/1507.06903v1.

[43] Xinyi Yuan, Shou-Wu Zhang, and Wei Zhang. *The Gross–Zagier formula on Shimura curves*, volume 184 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2013.