Counting elliptic curves of bounded Faltings height

by

RUTHI HORTSCH (Cambridge, MA)

1. Introduction. Let $\mathcal{E}$ be the set of isomorphism classes of elliptic curves over $\mathbb{Q}$. There are a number of invariants of elliptic curves we have yet to understand well (in particular, the rank of the rational points of an elliptic curve, or the size of its n-Selmer group). To discuss what value these take "on average", we need to define a measure on $\mathcal{E}$. This is often done by defining a *height*, a function $H : \mathcal{E} \to \mathbb{R}$ such that if $\mathcal{E}_{H<X} = \{E \in \mathcal{E} \mid H(E) < X\}$, then $\#\mathcal{E}_{H<X}$ is finite. We can then make sense of what an average value is on a finite set, and hope a limit exists as $X \to \infty$. To use this 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 over $\mathbb{Q}$ can be written uniquely as $y^2 = x^3 + Ax + B$ where $A, B \in \mathbb{Z}$ are such that there is no prime p with $p^4 \mid A$ and $p^6 \mid B$. If we call this curve $E_{A,B}$, then its naive height is

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

Using sieve methods or Möbius inversion (see [2]), one can show that

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

(for related calculations over number fields, see [1]). This formula comes from showing that the number of lattice points of a region is roughly its area with an error coming from its perimeter. The region where the naive height is less than X is the rectangle of length $2X^{1/3}$ and height $2X^{1/2}$, which has exactly area $4X^{5/6}$ and perimeter $O(X^{1/2})$ (the $\zeta(10)$ comes from excluding lattice points corresponding to nonminimal models).

Counting elliptic curves using other heights is more difficult. For example, counting elliptic curves of bounded discriminant is difficult because the region of points with bounded discriminant has *cusps*, by which is meant

[239]

there are points with large A, B and small discriminant. Controlling these cusps is difficult, even if the ABC conjecture is assumed.

Brumer and McGuinness [3] have given a heuristic for the number of elliptic curves with positive (respectively negative) discriminant up to a bound, and Watkins [13] 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.

In this paper, we show that if h_F is the Faltings height, then

$$\#\mathcal{E}_{h_F < Y} = 12\sigma\zeta(10)^{-1}e^{10Y} + O(e^{6Y}Y^3)$$

where σ is an absolute constant that we express as a specific integral given in Section 2. Note that if we rewrite this in terms of $H_F = e^{12h_F}$, this looks similar to the equation for the naive height:

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

As in the naive height case, this comes from approximating lattice points in a region by the area with an error from the perimeter: $\sigma X^{5/6}$ is the area, while $12\zeta(10)^{-1}$ adjusts it to consider only curves up to isomorphism (which is controlled by only considering certain residue classes of integral points). This is more difficult than the naive height case because, like the discriminant, the relevant region has a cusp with unbounded points, but these cusps can be controlled more easily than in the discriminant case.

2. Faltings height. In his proof of the Mordell Conjecture, Faltings introduces a height on the set of abelian varieties over a number field K, referred to as the *Faltings height* (see [4] for a longer overview, or the original paper [6]). Silverman [10] has given a formulation for elliptic curves, which gives the Faltings height as logarithmic. For the sake of simplicity we will use this as a definition.

If $h_F(E)$ denotes the original Faltings height for an elliptic curve E defined over a number field K, let $H_F(E) = e^{12h_F(E)}$. If v is an infinite place of K, define

- $\varepsilon_v = 1, 2$ if v is real, resp. complex,
- $\Delta_E^{\min}$ the minimal discriminant,
- $\tau_v \in \mathcal{H}$ (the complex upper half-plane) such that

$$E(\overline{K_v}) \cong \mathbb{C}/(\mathbb{Z} + \tau_v\mathbb{Z}),$$

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

Using Silverman's reformulation, we define

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

which simplifies to

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

in the case where $K = \mathbb{Q}$. Since $\Delta(\tau)$ is a modular form of weight 12, it is easy to check that this is independent of the choice of τ. (Silverman [10], Faltings [6], and Deligne [5] all use different normalizations of the Faltings height; this normalization agrees with that of Faltings.) See also [9] for a good background and summary.

Silverman uses the above to show that for all $\varepsilon > 0$ there exist $C_1, C_2(\varepsilon)$ such that for all elliptic curves E over $\mathbb{Q}$,

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

This inequality tells us that the Faltings height is very similar to the naive height, but not 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.

We will study the number of isomorphism classes of elliptic curves E with $H_F(E) < X$ for large X. Say that (A, B) is *weakly minimal with respect to a prime p* provided that either $p^4 \nmid A$ or $p^6 \nmid B$; if this holds true for all primes, simply say (A, B) is *weakly minimal*. Let $E_{A,B}$ indicate the elliptic curve corresponding to $y^2 = x^3 + Ax + B$, and

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

The elliptic curves $E_{A,B}$ such that $(A, B) \in S_X$ are representatives of the isomorphism classes of elliptic curves with Faltings height less than X.

For τ in the upper complex half-plane $\mathcal{H}$, let $j(\tau)$ denote its j-invariant, and $\Delta(\tau)$ the modular discriminant as defined above. By the *traditional fundamental domain* (with respect to the j-invariant) we mean $|\tau| \geq 1$ and $-1/2 < \operatorname{Re}\tau \leq 1/2$. In this paper we prove the following:

THEOREM 2.1. *For $t \in \mathbb{R}$, let $\tau_t \in \mathcal{H}$ be such that $j(\tau_t) = 6912t/(4t+27)$ and τ_t is in the traditional fundamental domain. Let*

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

Then

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

Roughly, the intuition for where these numbers come from is as follows: $\sigma X^{5/6}$ is the area of the two-dimensional region in $\mathbb{R}^2$ consisting of (A, B)

such that the corresponding elliptic curve $E_{A,B}$ has Faltings height less than X, $12\zeta(10)^{-1}$ corrects for the fact that we only want to take weakly minimal (A, B), and the error comes from the fact that estimating lattice point counts by area will have an error related to the boundary (in particular, the boundary of a bounded version of this region). When the integral defining σ is evaluated, we get approximately $\sigma \approx 29089$, which means the constant of the leading term is approximately 348716.

This result is more challenging to prove than one might expect since H_F is not simply bounded below by a constant times H_N, so we cannot simply apply results known about the naive height. Additionally, the calculation involves counting lattice points in an unbounded region whose boundary is given by a *transcendental* equation, which rules out many standard approaches.

3. Defining the region of interest. Assume that we have A, B such that $4A^3 + 27B^2 \neq 0$. For $A, B \in \mathbb{Q}$, let $\Delta_{E_{A,B}}^{\min}$ denote the minimal discriminant (note that over $\mathbb{Q}$ there is a global minimal model, and we could take $\Delta_{E_{A,B}}^{\min}$ to just be the polynomial discriminant of this minimal model). Let $\tau_{A,B} \in \mathcal{H}$ be such that

$$|\tau_{A,B}| \geq 1, \quad -1/2 < \operatorname{Re}\tau_{A,B} \leq 1/2, \quad E_{A,B}(\mathbb{C}) \cong \mathbb{C}/(\mathbb{Z} + \tau_{A,B}\mathbb{Z}),$$

and let

$$\Delta_{A,B} = -16(4A^3 + 27B^2), \quad j_{A,B} = -1728\frac{(4A)^3}{\Delta_{A,B}}, \quad \lambda_{A,B} = \frac{|\Delta_{E_{A,B}}^{\min}|}{|\Delta_{A,B}|}.$$

Note that the first three of these make sense for $A, B \in \mathbb{R}$, while the last requires $A, B \in \mathbb{Q}$. Furthermore, $\lambda_{A,B}$ will be one of four values depending on whether the model given for $E_{A,B}$ is minimal at 2 and 3 (we will discuss this in Section 7). To clarify things later, note we have three values in this paper which have a delta in their notation, meaning slightly different things: $\Delta_E^{\min}$ (the minimal discriminant), $\Delta_{A,B}$ (the polynomial discriminant), and $\Delta(\tau)$ (the modular discriminant).

Silverman's theorem tells us that

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

This motivates us to define the function

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

which is well-defined for all $A, B \in \mathbb{R}$ where $\Delta_{A,B} \neq 0$ (we invert and take a square root to make later calculations easier). Fixing a $\lambda > 0$, we would

like to know how many weakly minimal integer points there are in

$$\{(A, B) \in \mathbb{R}^2 \mid \lambda f(A, B)^{-2} < X\}.$$

To ease this, we define

$$R_{X,\lambda} = \{(A, B) \in \mathbb{R}^2 \mid \lambda f(A, B)^{-2} < X\} \cup \{(A, B) \in \mathbb{R}^2 \mid \Delta_{A,B} = 0\}.$$

Note that since $R_{X,\lambda} = R_{X/\lambda,1}$, we may as well just study $R_X = R_{X,1}$. Furthermore, f is a weighted homogeneous function: replacing (A, B) with $(X^{1/3}A, X^{1/2}B)$ will scale $f(A, B)$ by $X^{-1/2}$. So R_X is R_1 scaled by $X^{1/3}$ in the A-direction and $X^{1/2}$ in the B-direction.

The next section calculates the area of R_1 (the value denoted σ), while the rest of the paper is focused on setting up the proof of the main result. We will show that for large but finite A, B, there are no more relevant lattice points in R_X. Scaling down to R_1, we can show that R_X has a finite boundary if contained in a finite box, but we want to measure the actual length of this boundary. We do this by showing that in a small enough box the boundary length does not exceed the box length, which follows from showing (asymptotic) uniformity in the cusps.

4. Area of R_1. To better calculate the area, we do a change of variables to the parameter $t = A^3/B^2$. Then $j_{A,B} = 6912t/(4t + 27)$, and we can also define a function

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

where τ_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

$$\operatorname{Area}(R_1) = \iint\limits_{R_1} dA \, dB$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{-f(t)}^{f(t)} \left(\frac{1}{3} t^{-2/3} B^{2/3} \right) dB \, dt$$

$$= \frac{1}{3} \int\limits_{-\infty}^{\infty} t^{-2/3} \left(\frac{3}{5} B^{5/3} \right) \Big|_{-f(t)}^{f(t)} dt$$

$$= \frac{2}{5} \int\limits_{-\infty}^{\infty} t^{-2/3} f(t)^{5/3} \, dt.$$

Using a computer to evaluate, we conclude $\mathrm{Area}(R_1) \approx 29089$. We will write $\sigma = \mathrm{Area}(R_1)$, and note that $\mathrm{Area}(R_X) = \sigma X^{5/6}$. This is the same σ that appears in Theorem 2.1. Note that in the cusp, near $\Delta_{A,B} = 0$, we have $t \approx -27/4$, and both τ_t and f grow large.

(If we wish to have a somewhat cleaner integral, we can also rewrite 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 taking into account that

$$\Delta(z) = (2\pi)^6 \frac{j'(z)^6}{j(z)^4(j(z) - 1728)^3},$$

we see the integral is

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

over a curve γ in the upper complex half-plane of all z with real $j(z)$.)

5. The rough shape of R_X

LEMMA 5.1. *The function $|\Delta(\tau)(\mathrm{Im}\,\tau)^6|$ is bounded on the complex upper half-plane $\mathcal{H}$.*

Proof. Since $\Delta(\tau)(\mathrm{Im}\,\tau)^6$ is invariant under $\mathrm{SL}_2(\mathbb{Z})$, assume that τ is in the standard fundamental domain, so that $|\tau| \geq 1$ and $|\mathrm{Re}\,\tau| \leq 1/2$. If we can show that $\Delta(\tau)(\mathrm{Im}\,\tau)^6$ is finite as $\mathrm{Im}\,\tau \to \infty$, then for any τ with $\mathrm{Im}\,\tau$ above some fixed constant, it is bounded. The section of the fundamental domain that has $\mathrm{Im}\,\tau$ not exceeding that constant is a compact domain, and thus the function must also be bounded on that domain.

It remains to show that the function is finite as $\mathrm{Im}\,\tau \to \infty$. By definition

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

where $q = e^{2\pi i \tau}$. So as $\mathrm{Im}\,\tau \to \infty$, we have $q \to 0$, and $|\Delta(\tau)| = O(q) = O(e^{-2\pi \mathrm{Im}\,\tau})$. Thus, as $\mathrm{Im}\,\tau$ grows large, $|\Delta(\tau)|(\mathrm{Im}\,\tau)^6$ tends to zero. ∎

COROLLARY 5.2. *Let C be a positive constant such that $|\Delta(\tau)|(\mathrm{Im}\,\tau)^6 < C$ for all $\tau \in \mathcal{H}$. If $(A, B) \in R_X$ then $|\Delta_{A,B}| < CX$.*

The above follows immediately from Lemma 5.1. The next lemma will show that points in R_X are either small compared to X, or close to the curve given by $\Delta_{A,B} = 0$.

LEMMA 5.3. *Let $(A, B) \in R_X$. If $|B|^2 < CX/27$, then $|A|^3 < CX/2$. Otherwise, $A = -cB^{2/3} + \varepsilon XB^{-4/3}$ for some $|\varepsilon| < C$ and $c = (27/4)^{1/3}$.*
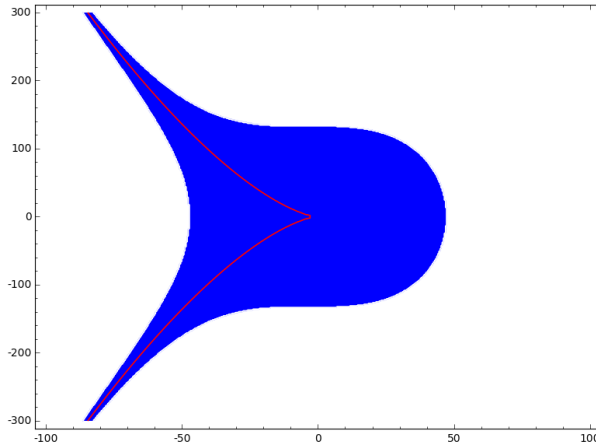
Fig. 1. The region R_1 (blue), with the cubic $4A^3 + 27B^2$ inside (red) (for colors see the online version)

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

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

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

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

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

Since $x^2 - 3cx + 3c^2 \geq 1$ for all $x \in \mathbb{R}$ and by assumptions $|\varepsilon| > C$, we have $|\Delta_{A,B}| > 64CX > CX$. This implies $(A, B) \notin R_X$. ∎

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 use it to count the lattice points in the region R_X. It is this calculation that differs from considering only elliptic curves of bounded discriminant.

6. 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 the 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 6.1. *Assume that $B \neq 0$. Let ε_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 find that

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

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

The following lemma is a strong version of Lemma 5.3 that only holds for large enough X and A.

LEMMA 6.2. *There are positive constants M and N such that for all $X \geq M$, if $|A| > NX^{1/3}(\log X)^2$ and $A = -cB^{2/3} + \varepsilon B^{-4/3}$ where $|\varepsilon| \geq \varepsilon_0$, then $(A, B) \notin R_X$.*

Proof. Assume that we have taken $\tau_{A,B}$ in the usual fundamental domain, so that $|\operatorname{Re}\tau_{A,B}| \leq 1/2$ and $|\tau_{A,B}| \geq 1$. We note that

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

However, 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 $\operatorname{Im}\tau \leq 1$, 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 to have $c_1 N^3 (\log M)^6 > 1$, 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 $\operatorname{Im}\tau > 1$. 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 $|q_\tau| = e^{-2\pi\operatorname{Im}\tau}$, and for $\operatorname{Im}\tau > 1$, $\left|\log|j| - \log|q_\tau^{-1}|\right|$ is bounded, it follows that

$$\operatorname{Im}\tau < c_3 \log|A| + c_4$$

for some positive constants c_3, c_4. So

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

If we take N such that $3c_1 N^3 > 2c_3$, then there is some constant c_5 such

that whenever $X > c_5$ and $|A| > NX^{1/3}(\log X)^2$,

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

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

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

Together Lemmas 6.1 and 6.2 tell us that for $X \geq M$, if we have $|A| > NX^{1/3}(\log X)^2$, then either $|\Delta_{A,B}| < 1$ or $(A, B) \notin R_X$. Thus, the only integer points in R_X where $|A| > NX^{1/3}(\log X)^2$ are those where $\Delta_{A,B} = 0$.

7. Weakly minimal curves not minimal at 2 or 3. Ultimately, what we want to count is

$$\#\{(A, B) \in \mathbb{Z}^2 \mid (A, B) \text{ is weakly minimal, } H_F(E_{A,B}) < X\}.$$

We defined f so that $H_F(E_{A,B}) < X$ is equivalent to $(A, B) \in R_{X,\lambda_{A,B}} = R_{X/\lambda_{A,B}}$, so to use our study of R_X, we need to examine $\lambda_{A,B}$ more carefully.

Recall that we defined $\lambda_{A,B} = \lambda_{E_{A,B}}$ so that $|\Delta_{E_{A,B}}^{\min}| = \lambda_{A,B}|\Delta_{A,B}|$. Since we are requiring that (A, B) be weakly minimal, it follows that the model $E_{A,B}$ is minimal everywhere except possibly at 2 or 3. Thus it will be the case 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. Let Cl_λ be the set of residue classes modulo 6^6 such that if $(A, B) \in \mathbb{Z}^2$ reduces to a class in Cl_λ, then (A, B) is weakly minimal at 2 and 3, and $\lambda_{A,B} = \lambda$. The table below summarizes the values of $\# \mathrm{Cl}_\lambda$, which were calculated using Tate's algorithm (found in [11]), supplemented by some calculations with `sage` [12].

Model	Factor λ	Size of Cl_λ
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}	12×18

Fix two residue classes $A_0 \bmod 6^6$ and $B_0 \bmod 6^6$ so that this choice corresponds to a particular value $\lambda = \lambda_{A_0,B_0} = 1$, 2^{-12}, 3^{-12} or 6^{-12}. Then for weakly minimal (A, B) corresponding to these classes, $H(E_{A,B}) < X$ is equivalent to $(A, B) \in R_{X/\lambda}$. We want to calculate

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

since summing these over all residue classes will give the number of elliptic curves of Faltings height less than X up to isomorphism. We do this in the next section.

8. Counting weakly minimal lattice points of a fixed residue class

PROPOSITION 8.1. *Fix two residue classes A_0, B_0 modulo 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 modulo 6^6 to (A_0, B_0) such that $\Delta_{A,B} \neq 0$ is $\sigma 6^{-12} X^{5/6} + O(X_M^{1/2} (\log X_M)^3)$ where $X_M = \max(X, M)$ and $\sigma = \mathrm{Area}(R_1)$.*

Proof. The general idea in this proof is the classical one of approximating the number of lattice points by the area of a region. Since the error term is dependent on the length of the boundary, we will need to use a region that has finite 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 $X_M = \max(M, X)$. Then

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

where N, M are from Lemma 6.2. If $X \geq M$, then Lemmas 6.1 and 6.2 tell 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 hence also $\Delta_{A,B} = 0$.

Recall that C is the bound on $|\Delta(\tau)(\mathrm{Im}\,\tau)^6|$ from Lemma 5.1, and $c = (27/4)^{1/3}$. Using Lemma 5.3, we find a $\beta_0 > 0$ such that if $|A| \leq N X_M^{1/3} (\log X_M)^2$ and $(A, B) \in R_X$, then $|B| \leq \beta_0 N^{3/2} X_M^{1/2} (\log X_M)^3$. 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 8.2. *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 away from $\Delta_{A,B} = 0$, and thus in any compact set, its zero set is rectifiable (see [7, 3.4.10]). ∎

Note that [7, 3.4.10] also 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 $\beta X^{1/2} < |B|$ (note that we can fix α, β 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 hence so is $\Delta(\tau_{A,B})$. In fact, it shares the same sign as $\Delta_{A,B}$. So

$$(8.1) \qquad F(A, B) = \Delta_{A,B} - \Delta(\tau_{A,B})(\operatorname{Im}\tau_{A,B})^6$$

is a real-valued function whose zero set is the boundary of R_1. We will show that there is a β such that for all $B > \beta$, each partial derivative 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 limits by how much a curve can change direction and thus bounds its length.)

LEMMA 8.3. *There is a $\beta > 0$ such that if $(A, B) \in \mathbb{R}^2$ satisfies $F(A, B) = 0$ and $|B| > \beta$, then the gradient ∇F at (A, B) is in the third quadrant if $B > 0$, and in the second 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 (8.1).

Note that Corollary 5.2 bounds the second term in the definition of F by C (since we are 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 hence so does $j_{A,B}$, from which it follows that $q_{A,B} \to 0$.

Define

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

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 $dq/dJ = 1 + O(J) = 1 + O(q)$. Similarly, since $\Delta(q)$ is a convergent power series in q with leading term q, we have $d\Delta(q)/dq = 1 + O(q)$. Using the fact that q is real, we get

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

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

Lastly, $\partial \Delta_{A,B}/\partial A$ and $\partial J/\partial A$ are obtained by straightforward calculations. Putting all this together, we get

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

and similarly

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

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

$$(8.2) \qquad \Delta_{A,B} = \Delta(q_{A,B})(\log|q_{A,B}|)^6.$$

Additionally, for small q, we have $\Delta(q) = q(1 + o(1)) = 1/j(1 + o(1))$; substituting this into (8.2) yields

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

Now, it follows from Lemma 5.3 (for R_1) that B^2/A^3 approaches $-4/27$. We put these two estimates into (8.1) to conclude that

$$\frac{\partial F}{\partial A} = -A^2\left(192 - \frac{3}{\pi^6}(1 + o(1))\right), \quad \frac{\partial F}{\partial B} = -B\left(864 - \frac{27}{2\pi^6}(1 + o(1))\right). \quad \blacksquare$$

By Lemma 5.3, we can choose α such that if $(A, B) \in R_X$ and $|B| \le \beta X^{1/2}$ then $|A| \le \alpha X^{1/3}$. Thus, for points where $|B| < \beta X^{1/2}$, Lemma 8.2 works, and for those where $|B| \ge \beta X^{1/2}$, the above argument works, 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 $|A| < NX^{1/3}(\log X)^2$, we see that 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, it is a standard result (for a proof see [8, Vol. 2, p. 186]) that if the boundary of a region is rectifiable, the difference between its area and the number

of lattice points it contains is bounded by $4(L+1)$ where L is the length of the boundary, thus $O(X_M^{1/2}(\log X_M)^3)$. The area of the region R_X is $\sigma X^{5/6}$, and this differs from the area of R_X' by the region where $|B| > \beta_0 N^{3/2} X^{1/2}(\log X)^3$. Lemma 6.2 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

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

so we can 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 by $1/m$ in both directions. The number of lattice points of a particular residue class modulo 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 $\sigma 6^{-12} X^{5/6} + O(X_M^{1/2}(\log X_M)^3)$.

This finishes the proof of Proposition 8.1. ∎

9. Conclusion of the proof of Theorem 2.1. Recall that R_X was defined so that

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

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 the size of these. Recall from Section 7 that Cl_λ is the set of residue classes modulo 6^6 such that if (A,B) reduces to a class in Cl_λ, then $\lambda_{A,B} = \lambda$ and (A,B) is weakly minimal at 2 and 3. A Möbius inversion argument shows that

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

so

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

Note that none of the points we are counting have $|A| > NX^{1/3}(\log X)^2$ (proved in Lemmas 6.1 and 6.2), so $d \le N^{1/4}(X/\lambda)^{1/12}(\log(X/\lambda))^{1/2}$. Additionally, since $(A, B) \bmod 6^6 \in \mathrm{Cl}_\lambda$, it is also true that $2 \nmid d$ and $3 \nmid d$.

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 $\#d^*\mathrm{Cl}_\lambda = \#\mathrm{Cl}_\lambda$. If $X > M$, there is a bijection between

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

and

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

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

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

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

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

by Proposition 8.1. It is simple to show that

$$(9.1) \qquad \sum_{\substack{d < Y \\ 2 \nmid d,\, 3 \nmid d}} \mu(d)d^{-10} = \frac{1}{\zeta(10)(1 - 2^{-10})(1 - 3^{-10})} + O(Y^{-9}),$$

so

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

with an error of

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

(which dominates the error coming from applying (9.1)). Using basic algebra and applying (9.1) again, we can show that this error is

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

Summing over the four possible values of λ using the calculations of $\#\mathrm{Cl}_\lambda$ from Section 7, this gives us

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

References

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

[2]   A. Brumer, *The average rank of elliptic curves. I*, Invent. Math. 109 (1992), 445–472.

[3]   A. Brumer and O. McGuinness, *The behavior of the Mordell–Weil group of elliptic curves*, Bull. Amer. Math. Soc. 23 (1990), 375–382.

[4]   G. Cornell and J. H. Silverman (eds.), *Arithmetic Geometry* (Storrs, CT, 1984), Springer, New York, 1986.

[5]   P. Deligne, *Preuve des conjectures de Tate et de Shafarevitch (d'après G. Faltings)*, Astérisque 121–122 (1985), 25–41.

[6]   G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. 73 (1983), 349–366.

[7]   H. Federer, *Geometric Measure Theory*, Grundlehren Math. Wiss. 153, Springer, New York, 1969.

[8]   E. Landau, *Vorlesungen über Zahlentheorie*, Chelsea, New York, 1969.

[9]   S. Löbrich, *A gap in the spectrum of the Faltings height*, arXiv:1505.00602 (2015).

[10] J. H. Silverman, *Heights and elliptic curves*, in: [4], 253–265.

[11] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Grad. Texts in Math. 151, Springer, New York, 1994.

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

[13] M. Watkins, *Some heuristics about elliptic curves*, Experiment. Math. 17 (2008), 105–125.

Ruthi Hortsch
Department of Mathematics, 2-239A
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139, U.S.A.
E-mail: rhortsch@math.mit.edu