Compositio Mathematica **122:** 151–164, 2000.

A Formula for the Number of Elliptic Curves with Exceptional Primes

DAVID GRANT
Department of Mathematics, University of Colorado at Boulder, Boulder, CO 80309–0395, U.S.A. e-mail: grant@boulder.colorado.edu

(Received: 4 August 1998; in final form: 9 March 1999)

Abstract. We prove a conjecture of Duke on the number of elliptic curves over the rationals of bounded height which have exceptional primes.

Mathematics Subject Classification (2000): 11G05.

Key words: elliptic curves, Galois representations.

Let E be an elliptic curve defined over $\mathbb{Q}$. Let p be a prime and $E[p]$ be the points of E of order dividing p. Let $\overline{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$. Then $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on $E[p]$, and picking a basis for $E[p]$ as a 2-dimensional vector space over $\mathbb{Z}/p\mathbb{Z}$ gives a representation

$$\rho_p : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Z}/\mathrm{p}\mathbb{Z}),$$

whose image we denote by $G(p)$. We will call p an exceptional prime for E if ρ_p is not surjective. A theorem of Serre [S2] states that if E does not have complex multiplication, then E has only finitely many exceptional primes. Masser and Wüstholz have given a bound for the largest such in terms of the height of E [MW].

Recently, Duke proved that 'almost all' elliptic curves over $\mathbb{Q}$ have no exceptional primes [D]. More precisely, every elliptic curve E over $\mathbb{Q}$ has a unique model of the form

$$y^2 = x^3 + Ax + B,$$

with $A, B \in \mathbb{Z}$, which is minimal in the sense that the greatest common divisor (A^3, B^2) is twelfth-power free. For such a minimal model, we define the naive height $H(E)$ to be $\max(|\mathrm{A}^3|, |\mathrm{B}^2|)$. Let $\mathcal{C}(X)$ be the set of elliptic curves E with $H(E) \leqslant X^6$. If $\mathcal{E}(X)$ is the subset of $\mathcal{C}(X)$ consisting of curves with at least one exceptional prime, Duke showed that

$$\lim_{X\to\infty} |\mathcal{E}(X)|/|\mathcal{C}(X)| = 0.$$

In more detail, we know $|\mathcal{C}(X)| \asymp X^5$ [B], while Duke showed that for some constant β,

$$|\mathcal{E}(X)| = \mathrm{O}(X^4 \log^{\beta}(X)).$$

At the same time, Duke conjectured that

$$|\mathcal{E}(X)| \sim CX^3, \tag{1}$$

for some constant C. The purpose of this paper is to prove this conjecture. For any prime p, let $\mathcal{E}_p(X)$ denote the curves of $\mathcal{C}(X)$ which are exceptional at p. We prove the following.

PROPOSITION 1. *Let $\varepsilon_{\pm}$ be the real roots of $x^3 \pm x - 1 = 0$, and ζ the Riemann ζ-function. Let $C_2 = (4\varepsilon_+ + 4\varepsilon_- + 6\log(\varepsilon_-/\varepsilon_+))/3\zeta(6)$. Then*

$$|\mathcal{E}_2(X)| = C_2X^3 + \mathrm{O}(X^2 \log(X)).$$

Duke had shown Proposition 1 with an error term of $X^2 \log^5(X)$ [D].

PROPOSITION 2. *Let $C_3 = 2/\zeta(6)$. Then*

$$|\mathcal{E}_3(X)| = C_3X^3 + \mathrm{O}(X^2 \log^2(X)).$$

THEOREM. *For any $\varepsilon > 0$,*

$$|\mathcal{E}(X)| = (C_2 + C_3)X^3 + \mathrm{O}(X^{2+\varepsilon}),$$

so the conjecture (1) holds with constant $C = C_2 + C_3$.

The main tools come from earlier work of Serre [S1], where he proved his theorem for curves with non-integral j-invariant, and Mazur's work on the possible rational isogenies of E [M1], [M2], [M3]. We proceed by covering $\mathcal{E}(X)$ by sets whose orders we can bound.

Using modular curves, in the first section we address $\mathcal{E}_p(X)$ for $p > 3$. Propositions 1 and 2 are proved by more hands-on methods in Sections 2 and 3, and then in the final section we prove the theorem.

1. Bounds for Primes Greater than 3

We first recall that if E is the elliptic curve $y^2 = x^3 + Ax + B$, then its discriminant $\Delta(E) = -16(4A^3 + 27B^2) \neq 0$, and its j-invariant $j(E)$ is given by $-2^{12} \cdot 3^3 \cdot A^3/\Delta(E)$. Hence, if $E \in \mathcal{C}(X)$, then $\Delta(E) = \mathrm{O}(X^6)$ and $H(j(E)) = \mathrm{O}(X^6)$, where if a, b are relatively prime integers, the height $H(a/b) = \max(|a|, |b|)$. Recall that $j(E)$ determines the $\overline{\mathbb{Q}}$-isomorphism class of E, but all the twists of E over $\mathbb{Q}$ have the same j-invariant. If $j(E) \neq 0, 1728$, then E only has quadratic twists of the form $y^2 = x^3 + At^2x + Bt^3$, for some $t \in \mathbb{Q}^{\times}$. Therefore, if $E \in \mathcal{C}(X)$ is the curve in its

$\overline{\mathbb{Q}}$-isomorphism class of smallest height, and $j(E) \neq 0, 1728$, then E has at most $2X/H(E)^{1/6}$ twists in $\mathcal{C}(X)$. So it will be convenient to separately consider $\mathcal{C}^0(X) = \{E \in \mathcal{C}(X) | \, j(E) = 0\}$, and $\mathcal{C}^{1728}(X) = \{E \in \mathcal{C}(X) | \, j(E) = 1728\}$.

We now want to study $|\mathcal{E}_p(X)|$ for each prime p. For $p \geqslant 5$, we will proceed in a crude fashion (which nonetheless suffices for our theorem), first counting rational points on modular curves of bounded j-invariant, and then accounting for twists. Hence it is easier to consider $\mathcal{E}'_p(X) = \mathcal{E}_p(X) - \mathcal{C}^0(X) - \mathcal{C}^{1728}(X)$.

Let $X(p)$ be the complete modular curve of level p, which parameterizes elliptic curves together with chosen bases of $E[p]$. Recall the following from [S3, p. 194], [M1], [M2]. The group $\mathrm{GL}_2(\mathbb{Z}/\mathrm{p}\mathbb{Z})$ acts on $X(p)$, and if L is a subgroup of $\mathrm{GL}_2(\mathbb{Z}/\mathrm{p}\mathbb{Z})$ such that the determinant map $L \to (\mathbb{Z}/p\mathbb{Z})^*$ is surjective, then $X(p)/L$ is a curve defined over $\mathbb{Q}$, whose non-cuspidal $\mathbb{Q}$-rational points parameterize elliptic curves E over $\mathbb{Q}$ with $G(p)$ contained in a conjugate of L. Furthermore, the function

$$j\colon X(p)/L \to X(1) = \mathbb{P}^1,$$

where j is the j-invariant, is a morphism over $\mathbb{Q}$, which is of degree $|\mathrm{GL}_2(\mathbb{Z}/\mathrm{p}\mathbb{Z})|/|\mathrm{L}|$ if $-I \in L$.

Recall that if E is an elliptic curve over $\mathbb{Q}$, that by the non-degeneracy of the Weil pairing, the image of $G(p)$ under the determinant map is all of $(\mathbb{Z}/p\mathbb{Z})^*$. Then [S1, p. IV-20] shows that if E is an elliptic curve over $\mathbb{Q}$ such that p is exceptional, then either $E[p]$ is reducible over $\mathbb{Q}$, or $G(p)$ does not contain a transvection, i.e., an element of the form $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ with respect to some basis. Indeed, for $p \geqslant 5$, it is shown in [S2] that either p is not exceptional, or $G(p)$ is contained in a Borel subgroup of $\mathrm{GL}_2(\mathbb{Z}/\mathrm{p}\mathbb{Z})$, in a normalizer of a split or non-split Cartan subgroup, or projects to a copy of the symmetric group $\mathcal{S}_4$ in $\mathrm{PGL}_2(\mathbb{Z}/\mathrm{p}\mathbb{Z})$. So if p is exceptional, it gives rise to a rational non-cuspidal point of the corresponding curves $X_0(p)$, $X_{\mathrm{split}}(p)$, $X_{\mathrm{non-split}}(p)$, and $X_{\mathcal{S}_4}(p)$. These are of degree $p+1$, $p(p+1)/2$, $p(p-1)/2$, and $(p^2-1)p/24$ over $X(1)$. For information about what is known about rational points of these curves, see [M1], [M2]. All we will use is a result of Mazur ([M3], Corollary 4.4), which shows that for $p \geqslant 17$, an elliptic curve over $\mathbb{Q}$ with $E[p]$ reducible over $\mathbb{Q}$ has potentially good reduction at all primes other than 2.

For $p \geqslant 5$, we can now bound $|\mathcal{E}'_p(X)|$ in a sequence of lemmas.

LEMMA 1. [S3, pp. 132–133]. *Let C be a curve of genus g over $\mathbb{Q}$, and let $f : C \to \mathbb{P}^1$ be a morphism defined over $\mathbb{Q}$ of degree d. Let $B(X)$ denote the $\mathbb{Q}$-rational points P of C such that $H(f(P)) \leqslant X$. Then, if $g = 0$, $|B(X)| = \mathrm{O}(X^{2/d})$; if $g = 1$, $|B(X)| = \mathrm{O}(\log(X)^{\rho/2})$, where ρ is the Mordell–Weil rank of the Jacobian of C; and if $g \geqslant 2$, then $|B(X)| = \mathrm{O}(1)$.*

LEMMA 2. *Suppose S is a set of elliptic curves over* $\mathbb{Q}$ *with j-invariant not 0 or 1728. Set* $S(X) = S \cap \mathcal{C}(X)$. *If there is an* $a \geqslant 0$, *such that for all* X, *the number of* $\overline{\mathbb{Q}}$-*isomorphism classes in* $S(X)$ *is* $O(X^a)$, *then for any* $\varepsilon > 0$,

$$|S(X)| = O(X^{\max(1,a)+\varepsilon}).$$

Proof. Let k be a positive integer such that $1/k < \varepsilon$. Let $S_i(X)$ contain those $E \in S(X)$ such that $H(E)$ is minimal in its $\overline{\mathbb{Q}}$-isomorphism class, and such that

$$X^{6i/k} \leqslant H(E) \leqslant X^{6(i+1)/k},$$

for $0 \leqslant i < k$. Each curve in $S_i(X)$ has at most $2X/X^{i/k} = O(X^{1-i/k})$ twists in $\mathcal{C}(X)$. But since $|S_i(X)| = O(X^{a(i+1)/k})$, the total number of curves in $S(X)$ whose $\overline{\mathbb{Q}}$-isomorphism class is represented by a curve of minimal height in $S_i(X)$ is $O(X^{(a-1)i/k+a/k+1})$, so $|S(X)|$ is $O(X^e)$, where

$$e = \max_{0 \leqslant i<k} ((a-1)i/k + a/k + 1).$$

If $a \geqslant 1$, then the maximum occurs at $i = k-1$, giving $e = a + 1/k < a + \varepsilon$. However, if $a < 1$, the maximum is achieved at $i = 0$, giving $e = 1 + a/k < 1 + 1/k < 1 + \varepsilon$.

LEMMA 3. *For any prime* $p \geqslant 7$, *and any* $\varepsilon > 0$,

$$|\mathcal{E}'_p(X)| = O(X^{\max(1,12/(p+1))+\varepsilon}).$$

Proof. Recall for $E \in \mathcal{C}(X)$, $H(j(E)) = O(X^6)$. So we can bound the number of $\overline{\mathbb{Q}}$-isomorphism classes in $\mathcal{E}_p(X)$ by counting the $\mathbb{Q}$-rational points on $X_0(p)$, $X_{\text{split}}(p)$, $X_{\text{non-split}}(p)$, and $X_{S_4}(p)$, with j-invariant $O(X^6)$. Since $p \geqslant 7$, the minimal degree of these curves over $X(1)$ is $p+1$, hence by Lemma 1 the number of $\overline{\mathbb{Q}}$-isomorphism classes of $E \in \mathcal{C}(X) - \mathcal{C}^0(X) - \mathcal{C}^{1728}(X)$ which are exceptional at p is $O(X^{12/(p+1)})$. By Lemma 2 we are done.

To tackle $\mathcal{E}_p(X)$ for $p = 2, 3, 5$, we need the following.

LEMMA 4. *Let K be a number field, I be the set of non-zero integral ideals of K, and N the norm from K to* $\mathbb{Q}$. *Then*

(a) *For any* $\alpha > 1$,

$$|\{\mathcal{I}, \mathcal{J} \in I | N(\mathcal{I})N(\mathcal{J})^\alpha \leqslant X\}| = O(X).$$

(b)

$$|\{\mathcal{I}, \mathcal{J} \in I | N(\mathcal{I})N(\mathcal{J}) \leqslant X\}| = \mathrm{O}(X \log(X)).$$

Proof. (a) Recall [L, p. 132], that the number of $\mathcal{I} \in I$ with $N(\mathcal{I}) \leqslant X$ is $\mathrm{O}(X)$, say with implied constant M. We are trying to bound

$$\sum_{N(\mathcal{J})^{\alpha} \leqslant X} \sum_{N(\mathcal{I}) \leqslant X/N(\mathcal{J})^{\alpha}} 1 \leqslant M \sum_{N(\mathcal{J}) \leqslant X^{1/\alpha}} \frac{X}{N(\mathcal{J})^{\alpha}}.$$

By the convergence of the Dedekind ζ-function ζ_K at α, the last sum is $\mathrm{O}(X)$.

(b) Here we are trying to bound

$$\sum_{N(\mathcal{J}) \leqslant X} \sum_{N(\mathcal{I}) \leqslant X/N(\mathcal{J})} 1 \leqslant M \sum_{N(\mathcal{J}) \leqslant X} \frac{X}{N(\mathcal{J})}.$$

and the lemma follows since a Tauberian theorem [P, p. 26], applied to ζ_K, gives $\sum_{N(\mathcal{J}) \leqslant X} \frac{1}{N(\mathcal{J})} = \mathrm{O}(\log(X))$.

LEMMA 5. *For any $\varepsilon > 0$,*

$$|\mathcal{E}_5'(X)| = \mathrm{O}(X^{2+\varepsilon}).$$

Proof. As in the proof of Lemma 3, $E \in \mathcal{E}_5'(X)$ gives rise to a rational point on $X_0(5)$, $X_{\mathrm{split}}(5)$, $X_{\mathrm{non-split}}(5)$, or $X_{\mathcal{S}_4}(5)$. In the first three cases, as in the proof of Lemma 3, there are only $\mathrm{O}(X^{2+\varepsilon})$ possible such curves in $\mathcal{C}(X)$. So we will assume from now on that $E \in \mathcal{E}_5(X)$ is a curve with $AB \neq 0$ such that $G(5)$ projects under $\pi\colon \mathrm{GL}_2(\mathbb{Z}/5\mathbb{Z}) \to \mathrm{PGL}_2(\mathbb{Z}/5\mathbb{Z})$ to a group $G = G(5)/(G(5) \cap (\mathbb{Z}/5\mathbb{Z})^{\times} \cdot I)$ that is isomorphic to $\mathcal{S}_4$. We will count such E in three steps. In Step I we produce a sextic (3) over $\mathbb{Q}$ whose splitting field is the fixed field K in $\mathbb{Q}(E[5])$ of $G(5) \cap (\mathbb{Z}/5\mathbb{Z})^{\times} \cdot I$. In Step II we derive a quintic polynomial $h(t)$ over $\mathbb{Q}$ which splits in K, and show that it must have a rational root. In Step III we show that the number of $E \in \mathcal{E}_5'(X)$ that give rise to an $h(t)$ with a rational root is $\mathrm{O}(X^{4/3+\varepsilon})$.

Step I. Let $E[5]'$ be the non-trivial points of $E[5]$. It is well-known that $L = \mathbb{Q}(\{x(u) | u \in E[5]'\})$ is the fixed field in $\mathbb{Q}(E[5])$ of $G(5) \cap \{\pm I\}$. We claim that L is also $\mathbb{Q}(\{x(u) - x([2]u) | u \in E[5]'\})$, where $[2]$ is the multiplication-by-2 map on E. It suffices to show that $x(u) - x([2]u)$ takes on 12 distinct values as u varies in $E[5]'$. Towards this end, recall that the x-coordinates of the points of $E[5]'$ are roots of a 12th degree polynomial, described, for instance, in [Si p. 105]. If $u \in E[5]'$, $x(u)$ is a root of this polynomial, and $x([2]u)$ is another. Using the duplication formula on E, one finds that $x(u) - x([2]u)$ is a root of

$$5t^{12} + 48At^{10} + 10\Delta t^6 + \Delta^2, \tag{2}$$

so $(x(u) - x([2]u))^2$ is a root of the so-called Jacobi sextic

$$5t^6 + 48At^5 + 10\Delta t^3 + \Delta^2. \tag{3}$$

Since $\Delta \neq 0$ and the discriminant of (3) is $-2^{20} \cdot 3^{12} \cdot 5^6 \Delta^8 B^4$, the roots of (2) are distinct.

Similarly, we claim that $K = \mathbb{Q}(\{(x(u) - x([2]u))^2 | u \in E[5]'\})$. Certainly $(x(u) - x([2]u))^2$ is fixed under any $\sigma \in (\mathbb{Z}/5\mathbb{Z})^\times \cdot I$, but by the above, any $\sigma \in G(5)$ that fixes $(x(u) - x([2]u))^2$ multiplies u by an element in $(\mathbb{Z}/5\mathbb{Z})^\times$. Since this is true for all $u \in E[5]'$, σ must be in $(\mathbb{Z}/5\mathbb{Z})^\times \cdot I$. Hence K is the splitting field of (3).

Step II: The Weil pairing forces a primitive fifth-root of unity μ to be in $\mathbb{Q}(E[5])$. Since the determinant of $-I$ is 1, μ is also in L. Likewise, since the determinant of $\pm\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ is -1, $\sqrt{5} = \mu - \mu^2 - \mu^3 + \mu^4$ is in K. However, μ cannot be in K, because there is no normal subgroup of $\mathcal{S}_\Delta$ of index 4. Hence all of $(\mathbb{Z}/5\mathbb{Z})^\times \cdot I$ is contained in $G(5)$.

Plugging $t = 1/s$ into (3), and multiplying by s^6/Δ^2 gives

$$g(s) = s^6 + \frac{10}{\Delta}s^3 + \frac{48A}{\Delta^2}s + \frac{5}{\Delta^2}. \tag{4}$$

Let $s_\infty = 1/(x(u) - x([2]u))^2$ be a chosen root of g, and let H be the subgroup of G that fixes $\mathbb{Q}(s_\infty)$. Since $[\mathbb{Q}(s_\infty) : \mathbb{Q}] \leqslant 6$, $|H| \geqslant 4$. So if $H' \in G(5)$ is the inverse image of H under π, then $|H'| \geqslant 16$. But if we take some $v \in E[5]'$ so that $\{v, u\}$ is an ordered basis for $E[5]$, then H' is contained in the upper triangular matrices, a group of order 80. Since $|H'|$ is prime to 5, it must be a group of order 16, so $|H| = 4$, and $[\mathbb{Q}(s_\infty) : \mathbb{Q}] = 6$. Furthermore, H' is a Sylow 2-subgroup of the upper triangular matrices, so taking a conjugate subgroup, and replacing v if necessary, we can assume H' is the subgroup of diagonal matrices. One sees then that H is a cyclic group of order 4.

The following manipulations are quite classical, related to $X(5)$ being the icosahedral cover of $X(1)$ [Kl]. We will follow the more recent text [Ki].

Let r be chosen so that

$$r^5 - \frac{1}{r^5} = -11 - \frac{125}{\Delta s_\infty^3}.$$

Setting $f = -1/\Delta, H = 48A/\Delta^2$, then $T = 3^3 \cdot 2^5 B/\Delta^3$ is a square root of $1728f^5 - H^3$. It is then shown in [Ki, p. 108] that the so-called Brioshi quintic

$$t^5 + \frac{10}{\Delta}t^3 + \frac{45}{\Delta^2}t - \frac{3^3 \cdot 2^5 B}{\Delta^3}, \tag{5}$$

has roots

$$t_k = ((1/\sqrt{5})(s_\infty - s_k)(s_{k+2} - s_{k+3})(s_{k+4} - s_{k+1}))^{1/2},$$

$0 \leqslant k \leqslant 4$, where the indices are taken mod 5, and where the other roots of the sextic (4) are

$$s_k = (s_\infty/5)\left(1 + r\mu^k - \frac{1}{r\mu^k}\right)^2, \tag{6}$$

for $0 \leqslant k \leqslant 4$. The quintic whose roots are the squares of those in (5) splits in K, and a calculation gives that quintic as

$$h(t) = t^5 + 20t^4/\Delta + 190t^3/\Delta^2 + 900t^2/\Delta^3 + 2025t/\Delta^4 - 2^{10} \cdot 3^6 B^2/\Delta^6.$$

We now want to show that h has a rational root. Since $[K : \mathbb{Q}]$ is prime to 5, h is not irreducible over $\mathbb{Q}$, so it suffices to show that it has an irreducible quartic factor over $\mathbb{Q}(s_\infty)$. To see this, we have to determine the action of $\mathrm{Gal}(\mathrm{K}/\mathbb{Q}(\mathrm{s}_\infty))$ on the roots of (4).

Note that $\mu^k r - 1/(\mu^k r)$, $0 \leqslant k \leqslant 4$, are the roots of the quintic

$$i(w) = w^5 + 5w^3 + 5w + 11 + 125/(\Delta s_\infty^3)$$

over $\mathbb{Q}(s_\infty)$. By (6), these roots are in a 2-power extension of K, so $i(w)$ can only have irreducible factors over $\mathbb{Q}(s_\infty)$ of degree a power of 2, so it must have a linear factor. Let r now be chosen so that $r - 1/r$ in in $\mathbb{Q}(s_\infty)$. Hence, $\mathbb{Q}(s_\infty, r)/\mathbb{Q}(s_\infty)$ is at most a quadratic extension. Note that (6) says K is a subfield of $\mathbb{Q}(s_\infty, r, \mu)$, and since μ is not in K, K must be a proper subfield. Likewise, $\sqrt{5} \notin \mathbb{Q}(s_\infty)$. So we have a sequence of fields, where each extension is quadratic:

$$\mathbb{Q}(s_\infty) \subset \mathbb{Q}(s_\infty, \sqrt{5}) \subset K \subset \mathbb{Q}(s_\infty, \mu, r).$$

We would like to identify which intermediate quadratic extension K is in the biquadratic extension $\mathbb{Q}(s_\infty, \mu, r)/\mathbb{Q}(s_\infty, \sqrt{5})$. Note that $\mathbb{Q}(s_\infty, r, \sqrt{5})$ must be a quartic extension of $\mathbb{Q}(s_\infty)$ since $\mu \notin K$, and since $\mathbb{Q}(s_\infty, r, \sqrt{5})/\mathbb{Q}(s_\infty)$ is a biquadratic extension, $\mathbb{Q}(s_\infty, r, \sqrt{5})$ is not K. Also K is not $\mathbb{Q}(s_\infty, \mu)$, so K must be the quadratic extension in $\mathbb{Q}(s_\infty, r, \mu)/\mathbb{Q}(s_\infty, \sqrt{5})$ which is the fixed field of the automorphism τ such that $\tau(\mu) = \mu^{-1}$ and $\tau(r) = -1/r$. Hence the Galois group of $K/\mathbb{Q}(s_\infty)$ can be identified with the Galois group of $\mathbb{Q}(s_\infty, r, \mu)/\mathbb{Q}(s_\infty)$ modded out by $< \tau >$, so is generated by an automorphism σ such that $\sigma(\mu) = \mu^2$ and $\sigma(r) = r$. Then we see from (6) that σ fixes s_∞, but $\sigma(s_k) = s_{2k}$, where the indices are taken mod 5.

We can check that (5) has distinct roots (its discriminant is $2^{24} \cdot 3^6 \cdot 5^5 A^6/\Delta^{12}$), so the action of σ on t_k^2 shows that $\sigma(t_k^2) = t_{2k}^2$, where the indices are taken mod 5. So $h(t)$ factors over $\mathbb{Q}(s_\infty)$ as a linear times a quartic factor, and $h(t)$ has a root in $\mathbb{Q}$.

Step III: Taking $t = z/\Delta$ in $h(t)$ and multiplying by Δ^5 gives

$$(z^2 + 10z + 45)^2 z - 2^{10} \cdot 3^6 B^2/\Delta. \tag{7}$$

So if a rational root to (7) is α/β with $(\alpha, \beta) = 1$, then $\beta^5\kappa = \Delta$ for some integer κ, where

$$\kappa = \pm(\Delta, 2^{10} \cdot 3^6 B^2), \tag{8}$$

and

$$(\alpha^2 + 10\alpha\beta + 45\beta^2)^2 \alpha\kappa = 2^{10} \cdot 3^6 B^2. \tag{9}$$

Writing $\Delta = -16(4A^3 + 27B^2)$, a calculation shows that

$$2^{12} \cdot 3^3 A^3 = -\kappa(\alpha + 3\beta)^3(\alpha^2 + 11\alpha\beta + 64\beta^2). \tag{10}$$

By assumption on E, $AB \neq 0$, and E has only quadratic twists. We will first assume that $E \in \mathcal{C}(X)$ is the curve in its $\overline{\mathbb{Q}}$-isomorphism class of minimal height with $\pi(G(5))$ isomorphic to $\mathcal{S}_4$, so that for no prime p does $p^6|(A^3, B^2)$. It is not hard to see from (8) that if p is a prime, and $p|\kappa$, then $p^2|\kappa$. Further, if $p \neq 2, 3$, then the minimality of E implies that $p^6 \nmid \kappa$, and that 2^{13} and 3^8 do not divide κ. Hence we can write $\kappa = \mu^2\lambda^3\nu^6$, with μ cube free and positive, λ squarefree, and $\nu|12$.

Since $\alpha^2 + 10\alpha\beta + 45\beta^2 \geqslant \frac{4}{9}\alpha^2$, from (9) we have

$$\alpha^5\kappa = \mathrm{O}(X^6). \tag{11}$$

Let $\delta = \alpha + 3\beta$. Then

$$\begin{aligned} &\alpha^2 + 11\alpha\beta + 64\beta^2 \\ &\quad = \delta^2 + 5\beta\delta + 40\beta^2 = \left(\delta + \left(\frac{5 + 3\sqrt{-15}}{2}\right)\beta\right)\left(\delta + \left(\frac{5 - 3\sqrt{-15}}{2}\right)\beta\right). \end{aligned} \tag{12}$$

Since $(\delta, \beta) = (\alpha, \beta) = 1$, in $\mathbb{Q}(\sqrt{-15})$ the two factors in (12) can only have the (ramified) primes over 3 and 5 as common prime factors. So if we take an ideal factorization

$$\left(\delta + \left(\frac{5 + 3\sqrt{-15}}{2}\right)\beta\right) = \mathcal{I}\mathcal{J}^3,$$

with $\mathcal{I}$ cube free, then $N(\mathcal{I})$ is cube free in $\mathbb{Z}$.

Hence by (10)

$$(2^4 \cdot 3A)^3 = \mu^2(-\lambda^3)(\nu^2)^3\delta^3 N(\mathcal{I})N(\mathcal{J})^3,$$

so $N(\mathcal{I}) = \mu$, and we get

$$2^4 \cdot 3A = -\lambda\nu^2\delta N(\mathcal{I})N(\mathcal{J}). \tag{13}$$

Now

$$\delta + \left(\frac{5+3\sqrt{-15}}{2}\right)\beta = \delta\left(\frac{11+3\sqrt{-15}}{6}\right) - \alpha\left(\frac{5+3\sqrt{-15}}{6}\right),$$

so by the triangle inequality, either

(i) $|\delta + \left(\frac{5+3\sqrt{-15}}{2}\right)\beta| \leqslant k_1|\delta|,$

or

(ii) $|\delta + \left(\frac{5+3\sqrt{-15}}{2}\right)\beta| \leqslant k_2|\alpha|,$

where

$$k_1 = 2|\frac{11+3\sqrt{-15}}{6}| \quad \text{and} \quad k_2 = 2|\frac{5+3\sqrt{-15}}{6}|.$$

In case (i), $N(\mathcal{I})N(\mathcal{J})^3 \leqslant k_1^2\delta^2$, so (13) gives $N(\mathcal{I})^3N(\mathcal{J})^5 = \mathrm{O}(X^4)$, so by Lemma 4, there are only $\mathrm{O}(X^{4/3})$ such pairs of ideals. Note that $\mathcal{I}, \mathcal{J}$ determine δ, β up to sign, and hence determine α, β up to sign, and also determine μ since $N(\mathcal{I}) = \mu$. Also, from (9) we get that $\lambda\alpha$ is a square, and so $\lambda|\alpha \neq 0$. Since by (11), $\alpha = \mathrm{O}(X^{6/5})$, the number of such λ for each α is $\mathrm{O}(X^{\varepsilon/2})$, for any $\varepsilon > 0$ [HW, p. 260]. Since there are only finitely-many choices of v, we have that there are only $\mathrm{O}(X^{4/3+\varepsilon/2})$-many E satisfying (i) which are of minimal height in their $\overline{\mathbb{Q}}$-isomorphism class.

In case (ii), $N(\mathcal{I})N(\mathcal{J})^3 = \mathrm{O}(\alpha^2)$, so

$$N(\mathcal{I})^{9/5}N(\mathcal{J})^3 = \mathrm{O}(\alpha^2\mu^{4/5}) = \mathrm{O}(\alpha^2\kappa^{2/5}) = \mathrm{O}(X^{12/5}),$$

by (11). Hence by Lemma 4, the number of pairs of ideals $\mathcal{I}, \mathcal{J}$ is $\mathrm{O}(X^{4/3})$. Just as in case (i), we conclude that there are only $\mathrm{O}(X^{4/3+\varepsilon/2})$-many E satisfying (ii) which are of minimal height in their $\overline{\mathbb{Q}}$-isomorphism class. Together we see, by Lemma 2, that there are only $\mathrm{O}(X^{4/3+\varepsilon})$-many $E \in \mathcal{E}'_5(X)$ with $\pi(G(5))$ isomorphic to $\mathcal{S}_4$.

Remark. To count points on $X_0(5)$ one can search for rational points on (4), which would probably give a better bound than that in Lemma 5. Again, the crude bound suffices for our theorem.

2. Proof of Proposition 1

Since ρ_2 is not surjective for $E \in \mathcal{E}_2(X)$, either E has a rational 2-torsion point, or $G(2)$ is of index 2 in $\mathrm{GL}_2(\mathbb{Z}/2\mathbb{Z}) \cong \mathcal{S}_3$. In the latter case, as explained in §5.3 of [S2], $\Delta(E)$ is a square, say C^2. But then $(\alpha, \beta, \gamma) = (-4A, 12B, C)$ is an integral solution of

$$\alpha^3 = 3\beta^2 + \gamma^2.$$

We want to bound the number of such triples (α, β, γ). Since $\alpha = 0$ implies that $\gamma = \beta = 0$, we can assume $\alpha \neq 0$. If ω is a primitive third-root of unity, then $\mathbb{Z}[\omega]$ is a unique factorization domain, so it is not hard to see that there exist $\phi, \psi \in \mathbb{Z}[\omega]$ such that $\gamma + \sqrt{-3}\beta = \phi\bar{\phi}^2\psi^3$, where a bar denotes complex conjugation, and hence $|\alpha|$ is the norm of $\phi\psi$. Since $\alpha = O(X^2)$, Lemma 4 gives us that there are $O(X^2 \log(X))$-many pairs ϕ, ψ. Since ϕ and ψ uniquely determine γ and β, there are only $O(X^2 \log(X))$ such $E \in \mathcal{E}_2(X)$ where $\Delta(E)$ is a square.

So we need only count the number of $E \in \mathcal{E}_2(X)$ with a rational 2-torsion point. These are all of the form

$$\begin{aligned} y^2 &= x^3 + Ax + B = (x-a)(x^2 + ax + b) \\ &= x^3 + (b - a^2)x - ab, \end{aligned} \tag{14}$$

for some integers a and b. We want to count the number of pairs (a, b) which give rise to a minimal elliptic curve of height bounded by X^6. The only time two pairs give rise to the same minimal curve is when the curve has 3 rational 2-torsion points, and all these curves have a square as discriminant, so we have already seen that there are at most $O(X^2 \log(X))$ of these. Further, the cubic (14) is an elliptic curve unless $b = -2a^2$ or $b = a^2/4$, which only occurs for $O(X)$ pairs (a, b). So the main term in the proposition comes from determining the order of $P(X)$, the set of integer pairs (a, b) with $|b - a^2| \leqslant X^2$ and $|ab| \leqslant X^3$, and sieving out those pairs giving rise to non-minimal models. Let $A(X)$ be the area of the region in the (a, b)-plane bounded by the two parabolas $b = a^2 + X^2$ and $b = a^2 - X^2$ and the hyperbolas $ab = X^3$ and $ab = -X^3$. By a slight modification of the argument in [L, p. 128], the difference between $|P(X)|$ and $A(X) = X^3 A(1)$ is $O(X^2)$. The minimality of E is equivalent to the condition that for no prime p does p^2 divide a while simultaneously p^4 divides b. So for every prime p, we want to sieve out the pairs $(p^2 a', p^4 b') \in P(X)$ with (a', b') in $P(X/p^2)$. We get therefore that $C_2 = A(1)/\zeta(6)$, and the proposition follows from the computation of $A(1)$.

3. Proof of Proposition 2

Again, since ρ_3 is not surjective for $E \in \mathcal{E}_3(X)$, either $E[3]$ has a rational line, or $G(3)$ is of index a multiple of 3 in $\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$.

We first consider the latter case, in which case it follows from §5.3 of [S2] that $\Delta(E)$ is a cube, say C^3. But then $C = O(X^2)$, and $(a, b, c) = (-4A, 12B, C)$ is an integral solution to

$$-3b^2 = c^3 - a^3.$$

We claim that if c and a are integers with $c = O(X^2)$, $a = O(X^2)$, then the number of triplets (a, b, c) satisfying (15) is $O(X^2 \log(X))$.

Indeed, there are $O(X^2)$ such solutions with $a = c$, so without loss of generality we can assume $c - a < 0$. Suppose (a, b, c) is such a triple. Then we can factor $c - \omega a$ over $\mathbb{Z}[\omega]$, absorbing the square factors into some Υ^2, the remaining powers of $\lambda = 1 - \omega$ into λ^ρ where $\rho = 0$ or 1, the remaining second degree prime factors and norms of first degree primes factors into some $s \in \mathbb{Z}$, and the remaining first degree prime factors and units into some σ. Therefore, s, σ, and $\overline{\sigma}$ are all prime to each other over $\mathbb{Z}[\omega]$, are prime to λ, and are squarefree. Complex conjugation determines $c - \omega^2 a$, and then using (15) we have factorizations:

$$c - a = -3^{1-\rho} N(\sigma) T^2,\ c - \omega a = \lambda^\rho s \sigma \Upsilon^2,\ c - \omega^2 a = \overline{\lambda}^\rho s \overline{\sigma}\,\overline{\Upsilon}^2, \tag{16}$$

where N denotes the norm from $\mathbb{Z}[\omega]$ to $\mathbb{Z}$, and $T \in \mathbb{Z}$. But since $(c - a)+$ $\omega(c - \omega a) + \omega^2(c - \omega^2 a) = 0$, we have

$$3^{1-\rho} N(\sigma) T^2 = \omega \lambda^\rho s \sigma \Upsilon^2 + \omega^2 \overline{\lambda}^\rho s \overline{\sigma}\,\overline{\Upsilon}^2,$$

hence s divides T^2, so s divides T. Therefore

$$3^{1-\rho} N(\sigma) s (T/s)^2 = \omega \lambda^\rho \sigma \Upsilon^2 + \omega^2 \overline{\lambda}^\rho \overline{\sigma}\,\overline{\Upsilon}^2, \tag{17}$$

and for a given choice of ρ, σ and Υ determine s, and, hence, c and a. But (17) also shows that $\overline{\sigma}$ divides Υ^2, and hence Υ. So if $\Upsilon = \overline{\sigma}\tau$, then σ and τ determine Υ and by (16) we have

$$c - \omega a = \lambda^\rho s \sigma \overline{\sigma}^2 \tau^2.$$

Since $|c - \omega a| = O(X^2)$, we have that $|\sigma\tau| = O(X)$ so $N(\sigma\tau) = O(X^2)$. Again by Lemma 4, there are only $O(X^2 \log(X))$-many such σ and τ, and since there are only 2 choices of ρ, and 2 choices of b once a and c are determined, we have our claim.

So we are left with counting $E \in \mathcal{C}(X)$ with $E[3]$ having a rational line, i.e., E having a non-trivial 3-torsion point with a rational x-coordinate. The curve $y^2 = x^3 + Ax + B$ has a non-trivial 3-torsion point with rational x-coordinate if and only if the three-division polynomial [Si p. 105]

$$3x^4 + 6Ax^2 + 12Bx - A^2$$

has a rational (hence, integral) root. So A and B are such that there exist integers r, s, t with

$$(x - r)(3x^3 + 3rx^2 + sx + t) = 3x^4 + 6Ax^2 + 12Bx - A^2,$$

or

$$6A = s - 3r^2, \qquad 12B = t - rs, \qquad A^2 = rt. \tag{18}$$

If $A \neq 0$, then letting d be the squarefree part of r, we have from the last equation of

(18) that

$$A = duv, \qquad r = du^2, \qquad t = dv^2,$$

for some integers u, v. Hence by (18)

$$s = 6duv + 3d^2u^4,$$

so the choice of d, u, v determines r, s, t, hence A and B. But there are not many choices of d, u, v with $|duv| = |A| \leqslant X^2$. Indeed, the techniques of [Sh, §3.8] show that the number of positive integers α, β, γ with $\alpha\beta\gamma \leqslant M$ for some M is $\mathrm{O}(M \log^2(M))$, so there are only $\mathrm{O}(X^2 \log^2(X))$ such E.

So the main term of $|\mathcal{E}_3(X)|$ comes entirely from those curves with $A = 0$. These correspond precisely to those curves with $|B| \leqslant X^3$ and B sixth-power free. There are $2X^3/\zeta(6) + \mathrm{O}(X^{1/2})$ such ([Sh, p. 291]).

4. Proof of the Theorem

Recall that a positive integer is called r-full if for every prime p dividing it, p^r divides it. For a given $r \geqslant 1$, every positive integer can be factored uniquely as a product of relatively prime r-full and r-free numbers. If we let $\mathrm{Full}_r(X)$ denote the r-full numbers less than or equal to X, then $|\mathrm{Full}_r(X)| = k_r X^{1/r} + \mathrm{O}(X^{1/(r+1)})$, for some constant $k_r > 0$ [Sh, p. 297].

Now take any $\varepsilon > 0$. Pick a positive integer r large enough so that $6/r < \varepsilon$ and so that $r \geqslant 13$. Now let $\mathcal{E}^r_{\mathrm{int}}(X)$ be the set of $E \in \mathcal{C}(X) - \mathcal{C}^0(X)$ such that when $\Delta(E)$ is factored as

$$\Delta(E) = \pm 2^\alpha 3^\beta c_r d_r, \tag{19}$$

where $c_r > 0$ and $d_r > 0$ are prime to 6, c_r and d_r are prime to each other, c_r is an r-full number, and d_r is an r-free number, then d_r divides A^3. Since $\Delta(E) = \mathrm{O}(X^6)$, the number of possible such α and β are $\mathrm{O}(\log(X))$. As above, the number of such possible c_r is $\mathrm{O}(X^{6/r})$, and since $A \neq 0$, for each choice of A the number of such possible d_r is $\mathrm{O}(X^\delta)$ for any $\delta > 0$. Then writing $\log(X) = \mathrm{O}(X^\delta)$ and taking $\delta < \frac{1}{3}(\varepsilon - 6/r)$, since $A = \mathrm{O}(X^2)$, we have

$$|\mathcal{E}^r_{\mathrm{int}}(X)| = \mathrm{O}(X^{2+\varepsilon}), \tag{20}$$

since there are at most 2 curves for a given choice of A and Δ.

We next note that

$$\mathcal{E}(X) \subseteq \mathcal{C}^0(X) \cup \mathcal{C}^{1728}(X) \cup (\cup_{p \leqslant r} \mathcal{E}'_p(X)) \cup \mathcal{E}^r_{\mathrm{int}}(X). \tag{21}$$

Indeed, if $E \in \mathcal{E}(X)$, and $E \notin \mathcal{C}^0(X) \cup \mathcal{E}^r_{\mathrm{int}}(X)$, then by (19) there is a $p > 3$ such that $-r < \mathrm{ord}_p(j(E)) < 0$, hence E has multiplicative reduction at p. So there is an extension K of degree 1 or 2 over $\mathbb{Q}$, such that if π is a prime of K over p, then E over the local field K_π is isomorphic to a Tate curve of parameter q, with $\mathrm{ord}_\pi(q) = -\mathrm{ord}_\pi(j(E)) = -e(\mathrm{ord}_p(j(E)))$, where $e = 1$ or 2 (see [Si] p. 355 for properties of the Tate curve). So if $P > r$ is a prime, then $P \nmid \mathrm{ord}_\pi(q)$. Hence by properties of the Tate curve, for all $P > r$, $\mathrm{Gal}(K(E[P])/K)$ contains a transvection [S1, p. IV-20], hence so does $\mathrm{Gal}(\mathbb{Q}(E[P])/\mathbb{Q})$. By the theorem of Mazur quoted in Section 1, since $r \geqslant 13$, and E has multiplicative reduction at $p > 2$, $E[P]$ is irreducible for $P > r$. Therefore E is not exceptional for all $P > r$, and (21) holds.

Putting together (20), and Lemmas 3 and 5, since it is easy to see that $\mathcal{C}^0(X) \subseteq \mathcal{E}_3(X)$ and $\mathcal{C}^{1728}(X) \subseteq \mathcal{E}_2(X)$, we have from (21) that

$$|\mathcal{E}(X)| = |\mathcal{E}_2(X) \cup \mathcal{E}_3(X)| + \mathrm{O}(X^{2+\varepsilon}).$$

The proof now follows from Propositions 1 and 2, and the observation that

$$|\mathcal{E}_2(X) \cap \mathcal{E}_3(X))| = \mathrm{O}(X^2 \log^2(X)). \tag{22}$$

Indeed, we saw in the proofs of Propositions 1 and 2 that the only curves in $\mathcal{E}_2(X)$ and $\mathcal{E}_3(X)$ which contribute to the dominant terms in the statements of the propositions are those which have a rational two-torsion point and those of the form $y^2 = x^3 + B$. For $y^2 = x^3 + B$ to have a rational 2-torsion point forces B to be a cube, and there are only $\mathrm{O}(X)$ such of absolute value bounded by X^3. Therefore (22), and the theorem, follow.

Acknowledgements

I would like to thank Bill Duke, Sheldon Kamienny, and Eric Stade for helpful discussions on this material.

References

[B] Brumer, A.: The average rank of elliptic curves I, *Invent. Math.* **109** (1992), 445–472.
[D] Duke, W.: Elliptic curves with no exceptional primes, *CR Acad. Sci. Paris Série I.* **325** (1997), 813–818.
[HW] Hardy, G. H. and Wright, E. M.: *An Introduction to the Theory of Numbers*, 4th edn. Oxford Univ. Press, 1960.
[Ki] King, R.: *Beyond the Quartic Equation.* Birkhäuser, Boston, 1996.
[Kl] Klein, F.: *Lectures on the Icosahedron and the Solution of Equations of the Fifth Degree*, Dover, New York, 1956.
[L] Lang, S.: *Algebraic Number Theory*, Springer, New York, 1986.
[MW] Masser, D. and Wüstholz, G.: Galois properties of division fields of elliptic curves, *Bull. London Math. Soc.* **25** (1993), 247–254.
[M1] Mazur, B.: Rational points on modular curves, In: *Lecture Notes in Math.* 601, Springer, New York, 1977, pp. 107–148.
[M2] Mazur, B.: Modular curves and the Eisenstein ideal, *IHES* **47** (1977), 33–186.
[M3] Mazur, B.: Rational isogenies of prime degree, *Invent. Math.* **44** (1978), 129–162.
[P] Postnikov, A.: *Introduction to Analytic Number Theory*, Amer. Math. Soc., Providence, 1988.

[S1] Serre, J.-P.: *Abelian ℓ-adic Representations and Elliptic Curves*, Benjamin, New York, 1968.

[S2] Serre, J.-P.: Propriétés galoisiennes des points d'ordre fini des courbes elliptiques, *Invent. Math.* **15** (1972), 123–201 (= Collected Papers, III, 1–73).

[S3] Serre, J.-P.: *Lectures on the Mordell-Weil Theorem*, Vieweg, Braunschweig, 1989.

[Sh] Shapiro, H.: *Introduction to the Theory of Numbers*, Wiley, New York, 1983.

[Si] Silverman, J.: *The Arithmetic of Elliptic Curves*, Springer, New York, 1986.