A HEURISTIC FOR BOUNDEDNESS OF RANKS OF ELLIPTIC CURVES

JENNIFER PARK, BJORN POONEN, JOHN VOIGHT, AND MELANIE MATCHETT WOOD

ABSTRACT. We present a heuristic that suggests that ranks of elliptic curves E over $\mathbb{Q}$ are bounded. In fact, it suggests that there are only finitely many E of rank greater than 21. Our heuristic is based on modeling the ranks and Shafarevich–Tate groups of elliptic curves simultaneously, and relies on a theorem counting alternating integer matrices of specified rank. We also discuss analogues for elliptic curves over other global fields.

CONTENTS

1. INTRODUCTION

1.1. **A new model.** The set $E(\mathbb{Q})$ of rational points of an elliptic curve E over $\mathbb{Q}$ has the structure of an abelian group. Mordell [Mor22] proved in 1922 that $E(\mathbb{Q})$ is finitely generated, so its rank $\operatorname{rk} E(\mathbb{Q})$ is finite. Even before this, in 1901, Poincaré [Poi01, p. 173] essentially asked for the possibilities for $\operatorname{rk} E(\mathbb{Q})$ as E varies. Implicit in this is the question of boundedness: Does there exist $B \in \mathbb{Z}_{\geq 0}$ such that for every elliptic curve E over $\mathbb{Q}$ one has $\operatorname{rk} E(\mathbb{Q}) \leq B$?

In this article, we present a probabilistic model providing a heuristic for the arithmetic of elliptic curves, and we prove theorems about the model that suggest that $\operatorname{rk} E(\mathbb{Q}) \leq 21$ for all but finitely many elliptic curves E.

Our model is inspired in part by the Cohen–Lenstra heuristics for class groups [CL84], as reinterpreted by Friedman and Washington [FW89]. These heuristics predict that for a fixed odd prime p, the distribution of the p-primary part of the class group of a varying imaginary quadratic field is equal to the limit as $n \to \infty$ of the distribution of the cokernel of the homomorphism $\mathbb{Z}_p^n \xrightarrow{A} \mathbb{Z}_p^n$ given by a random matrix $A \in \mathrm{M}_n(\mathbb{Z}_p)$; see Section 4 for the precise conjecture. In analogy, and in agreement with conjectures of Delaunay [Del01, Del07, DJ14], Bhargava, Kane, Lenstra, Poonen, and Rains [BKLPR15] predicted that for a fixed prime p and $r \in \mathbb{Z}_{\geq 0}$, the distribution of the p-primary part of the Shafarevich–Tate group $\mathrm{III}(E)$ as E varies over rank r elliptic curves over $\mathbb{Q}$ ordered by height equals the limit as $n \to \infty$ (through integers of the same parity as r) of the distribution of $\mathrm{coker}\, A$ for a random alternating matrix $A \in \mathrm{M}_n(\mathbb{Z}_p)$ subject to the condition $\mathrm{rk}_{\mathbb{Z}_p}(\ker A) = r$; see Section 5 for the precise conjecture and the evidence for it.

If imposing the condition $\mathrm{rk}_{\mathbb{Z}_p}(\ker A) = r$ yields a distribution conjecturally associated to the curves of rank r, then naturally we guess that if we choose A at random from the space $\mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}}$ of all alternating matrices *without* imposing such a condition, then the distribution of $\mathrm{rk}_{\mathbb{Z}_p}(\ker A)$ tends as $n \to \infty$ to the distribution of the rank of an elliptic curve. This cannot be quite right, however: since an alternating matrix always has even rank, the parity of n dictates the parity of $\mathrm{rk}_{\mathbb{Z}_p}(\ker A)$. But if we choose n uniformly at random from $\{\lceil \eta \rceil, \lceil \eta \rceil + 1\}$ (with $\eta \to \infty$), then we find that $\mathrm{rk}_{\mathbb{Z}_p}(\ker A)$ equals 0 or 1 with probability 50% each, and $\mathrm{rk}_{\mathbb{Z}_p}(\ker A) \geq 2$ with probability 0%; for example, when n is even, we have $\mathrm{rk}_{\mathbb{Z}_p}(\ker A) = 0$ unless $\det A = 0$, and $\det A = 0$ holds only when A lies on a (measure 0) hypersurface in the space $\mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}}$ of all alternating matrices. This 50%–50%–0% conclusion matches the elliptic curve rank behavior conjectured for quadratic twist families by Goldfeld [Gol79, Conjecture B] and Katz and Sarnak [KS99a, KS99b].

So far, however, this model does not predict anything about the number of curves of each rank ≥ 2 except to say that asymptotically they should amount to 0% of curves. Instead of sampling from $\mathrm{M}_n(\mathbb{Z}_p)$, we could sample from the set $\mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}, \leq X}$ of alternating *integer* matrices whose entries have absolute values bounded by X, and study

$$\lim_{X \to \infty} \mathrm{Prob}\left(\mathrm{rk}(\ker A) = r \mid A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}, \leq X}\right),$$

but this again would be 0 for each $r \geq 2$. To obtain finer information, instead of taking the limit as $X \to \infty$, we let X *depend on the height H of the elliptic curve being modeled*; similarly, we let η grow with H so that the random integer n grows too. Now for each $r \geq 2$, the event $\mathrm{rk}(\ker A) = r$ occurs with positive probability depending on H, and we can estimate for how many elliptic curves of height up to H the event occurs.

To specify the model completely, we must specify the functions $\eta(H)$ and $X(H)$; actually, it will turn out that specifying $X(H)^{\eta(H)}$ is enough for the conclusions we want to draw. We calibrate $X(H)^{\eta(H)}$ so that the resulting prediction for the expected size of $\mathrm{III}(E)$ for curves of height up to H agrees with theorems and conjectures about this expected size; this suggests requiring $X(H)^{\eta(H)} = H^{1/12 + o(1)}$.

Our model is summarized as follows. Fix increasing functions $\eta(H)$ and $X(H)$ such that $X(H)^{\eta(H)} = H^{1/12 + o(1)}$ as $H \to \infty$. (For technical reasons, we also require $\eta(H)$ to grow sufficiently slowly.) To model an elliptic curve E of height H:

1. Choose n uniformly at random from the pair $\{\lceil \eta(H) \rceil, \lceil \eta(H) \rceil + 1\}$.

2. Choose $A_E \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}}$ with entries bounded by $X(H)$ in absolute value, uniformly at random.

Then $(\mathrm{coker}\, A_E)_{\mathrm{tors}}$ models $\mathrm{III}(E)$, and $\mathrm{rk}(\ker A_E)$ models $\mathrm{rk}\, E(\mathbb{Q})$.

Thus, heuristically, for an elliptic curve E of height H, the "probability" that $\mathrm{rk}\, E(\mathbb{Q}) \geq r$ should be $\mathrm{Prob}(\mathrm{rk}(\ker A_E) \geq r)$. We prove that for any fixed $r \geq 1$, the latter probability is $H^{-(r-1)/24+o(1)}$ as $H \to \infty$ (Theorem 9.1.1). In other words, for each increase in rank beyond 1, the probability of attaining that rank drops by a factor of about $H^{1/24}$. Summing the probabilities $H^{-(r-1)/24+o(1)}$ over all elliptic curves E over $\mathbb{Q}$ yields a prediction for the expected number of curves of rank $\geq r$. It turns out that the sum diverges for $r < 21$ and converges for $r > 21$. The latter suggests that there are only finitely many E over $\mathbb{Q}$ with $\mathrm{rk}\, E(\mathbb{Q}) > 21$.[1] Summing instead over elliptic curves of height up to H leads to the prediction that for $1 \leq r \leq 20$, the number of E of height up to H satisfying $\mathrm{rk}\, E(\mathbb{Q}) \geq r$ is $H^{(21-r)/24+o(1)}$ as $H \to \infty$.

In order to separate as much as possible what is proved from what is conjectured, we express the model in terms of random variables serving as proxies for the rank and III of each elliptic curve, and prove unconditional theorems about these random variables before conjecturing that the conclusions of these theorems are valid also for the *actual* ranks and III. (This methodology is analogous to that of the Cramér model, which models the set of prime numbers by a random set P that includes each $n > 2$ independently with probability $1/\log n$; see, e.g., the exposition by Granville [Gra95].)

For example, we prove the following unconditional result (Theorem 7.3.3).

Theorem 1.1.1. *For each elliptic curve E over $\mathbb{Q}$, independently choose a random matrix A_E according to the model defined above, and let rk'_E denote the random variable $\mathrm{rk}(\ker A_E)$. Then the following hold with probability 1:*

(a) *All but finitely many E satisfy $\mathrm{rk}'_E \leq 21$.*
(b) *For $1 \leq r \leq 20$, we have $\#\{E : \mathrm{ht}\, E \leq H \text{ and } \mathrm{rk}'_E \geq r\} = H^{(21-r)/24+o(1)}$.*
(c) *We have $\#\{E : \mathrm{ht}\, E \leq H \text{ and } \mathrm{rk}'_E \geq 21\} \leq H^{o(1)}$.*

Remark 1.1.2. Our heuristic explains what should be expected if there are no significant phenomena in the arithmetic of elliptic curves beyond those incorporated in the model. It still could be, however, that there are special families of elliptic curves that behave differently for arithmetic reasons, just as there can be special subvarieties in the Batyrev–Manin conjectures on the number of rational points of bounded height on varieties [BM90]. When we generalize to global fields in Section 12, we *will* need to exclude some families of curves.

Remark 1.1.3. In fact, the known constructions of elliptic curves over $\mathbb{Q}$ of high rank proceed by starting with a parametric family with high rank generically, and then finding specializations of even higher rank. As Elkies points out, one cannot say that our heuristic for boundedness (let alone 21) is convincing until one refines the model to predict the rank distribution in such parametric families. One plausible heuristic is that for a family with generic rank r_0 and varying root number, the probability that a curve of height about H in the family has rank $r_0 + s$ is comparable (up to a factor $H^{o(1)}$) to the probability that an arbitrary curve of height about H has rank s. Although we cannot justify this directly, we can argue by analogy: the distribution of p-Selmer rank in certain families with generic rank r_0 is

---

[1]On the other hand, Elkies [Elk06] proved that there exist infinitely many E of rank at least 19.

conjecturally obtained simply by shifting the Selmer rank distribution for all elliptic curves by r_0 [PR12, Remark 4.17].

Remark 1.1.4. Venkatesh and Ellenberg [VE10, Section 4.1] observed that from the arithmetic of an imaginary quadratic field one can naturally construct an integer square matrix whose cokernel is the class group; see Section 4.1. In contrast, we do not know of any structure in the arithmetic of elliptic curves that suggests the model above for $\mathrm{rk}\, E(\mathbb{Q})$; in particular, we do not yet see a natural alternating matrix in the arithmetic of an elliptic curve. Our reason for using an alternating matrix is instead in the spirit of Occam's razor: the model of alternating matrices over $\mathbb{Z}_p$ proposed in [BKLPR15] is the simplest model we know of that models simultaneously the rank of an elliptic curve E and Ш (more precisely, $Ш(E)[p^\infty]$).

Remark 1.1.5. Deninger too has conjectured that $\mathrm{rk}\, E(\mathbb{Q})$ is naturally the dimension of the kernel of an alternating linear map [Den10, Example 5]. Specifically, in an attempt to explain the Riemann hypothesis for $L(E, s)$, he conjectured the existence of an infinite-dimensional $\mathbb{R}$-vector space H_E and an endomorphism $\theta \in \mathrm{End}\, H_E$ such that

- for any $\rho \in \mathbb{C}$, the endomorphism $\theta - \rho \in \mathrm{End}(H_E \otimes_\mathbb{R} \mathbb{C})$ satisfies $\dim_\mathbb{C} \ker(\theta - \rho) = \mathrm{ord}_{s=\rho} L(E, s)$, and
- the endomorphism $\theta - 1$ is alternating with respect to an inner product on H_E.

If these exist and the Birch and Swinnerton-Dyer conjecture is true, then $\mathrm{rk}\, E(\mathbb{Q}) = \dim \ker(\theta - 1)$.

1.2. **Outline of the paper.** Section 2 introduces some notation that will be used throughout the rest of the paper. Section 3 surveys some of the history regarding ranks of elliptic curves. Sections 4 and 5 discuss heuristics for class groups and Shafarevich–Tate groups, respectively, in terms of cokernels of matrices; the former heuristics are not logically necessary for our arguments, but they serve as the basis for an analogy. In Section 6 we prove theorems to help us predict the average size of Ш; the idea, due to Lang [Lan83], is to solve for this size in the Birch and Swinnerton-Dyer conjecture. These theorems will guide the setting of parameters in our model. Section 7 presents the model itself, and proves unconditional theorems about the random variables in it, while Section 8 conjectures that the conclusions of these theorems are valid also for the actual ranks and Ш. One of the statements in Section 7 depends on Theorem 9.1.1, whose proof is postponed to Section 9 so as not to interrupt the flow leading to the main conclusions and conjectures in Sections 7 and 8. Section 10 presents some computational evidence for our heuristic. Section 11 discusses some further questions. Finally, in Section 12 we discuss analogues of our heuristic for number fields K larger than $\mathbb{Q}$ and for global function fields such as $\mathbb{F}_p(t)$. In particular, we investigate whether our heuristic predicts a value for $B_K := \limsup_{E/K} \mathrm{rk}\, E(K)$. Also, using either Heegner points in anticyclotomic extensions of imaginary quadratic fields, or recent work of Bhargava, Skinner, and Zhang [BSZ14] combined with the multidimensional density Hales–Jewett theorem, one can prove the existence of number fields K for which B_K grows at least linearly in $[K : \mathbb{Q}]$.

2. NOTATION AND CONVENTIONS

We make many estimates of functions of several variables. If $x = (x_1, \ldots, x_m)$ and $a = (a_1, \ldots, a_n)$, we write $f(x, a) \ll_a g(x, a)$ to mean that there exists a positive-valued function $C(a)$ such that $f(x, a) \leq C(a)g(x, a)$ for all values of (x, a) we consider. We write $f(x, a) \asymp_a g(x, a)$ to mean $f(x, a) \ll_a g(x, a)$ and $g(x, a) \ll_a f(x, a)$. When "$o(1)$" appears in a sentence with a variable H going to infinity, our interpretation is that there exists a function $f(H)$, tending to 0 as $H \to \infty$, such that replacing the $o(1)$ by $f(H)$ makes the entire sentence true.

Let G be an abelian group. For $n \in \mathbb{Z}_{\geq 1}$, let $G[n] := \{x \in G : nx = 0\}$. For p prime, define $G[p^\infty] := \bigcup_{m \geq 1} G[p^m]$, and define the p-rank of G to be $\dim_{\mathbb{F}_p} G[p]$.

Let R be a commutative ring. For $n \in \mathbb{Z}_{\geq 0}$, let $\mathrm{M}_n(R)$ be the set of $n \times n$ matrices with entries in R. For $X \in \mathbb{R}_{>0}$, let $\mathrm{M}_n(\mathbb{Z})_{\leq X} \subseteq \mathrm{M}_n(\mathbb{Z})$ be the subset of matrices whose entries have absolute value less than or equal to X. A matrix $A \in \mathrm{M}_n(R)$ is alternating if $A^T = -A$ and all the diagonal entries are 0 (if 2 is not a zero divisor in R, then the skew-symmetry condition $A^T = -A$ suffices). Let $\mathrm{M}_n(R)_{\mathrm{alt}}$ be the set of alternating matrices, and let $\mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}, \leq X} := \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}} \cap \mathrm{M}_n(\mathbb{Z})_{\leq X}$.

For a subset $S \subseteq \mathrm{M}_n(\mathbb{Z}_p)$, define $\mathrm{Prob}(S) = \mathrm{Prob}(S \mid A \in \mathrm{M}_n(\mathbb{Z}_p))$ as the probability of S with respect to the normalized Haar measure on the compact group $\mathrm{M}_n(\mathbb{Z}_p)$.

Let R be an integral domain, and let $K := \mathrm{Frac}\, R$ be the field of fractions. If M is a finitely generated module over R, define $\mathrm{rk}\, M := \dim_K(M \otimes_R K)$. For $A \in \mathrm{M}_n(R)$, let $\mathrm{rank}\, A$ denote the rank of the matrix, so $\mathrm{rank}\, A = n - \mathrm{rk}(\ker A)$.

Finally, because both proven statements and conjectured statements play an important role in this paper, in order to distinguish the two, any unproven or conjectural (in)equality in a displayed equation comes with a question mark over the symbol, as in $\overset{?}{=}$.

3. HISTORY

3.1. **Brief history of boundedness guesses.** Many authors have proposed guesses as to whether ranks of elliptic curves over $\mathbb{Q}$ are bounded, and the consensus seems to have shifted over time.

Early researchers guessed that ranks were bounded. In 1950, Néron wrote "L'existence de cette borne est ... considérée comme probable" [Poi50, p. 495, end of footnote (^{3})], even though he himself proved the existence of elliptic curves of rank ≥ 11 [Nér56]. Honda conjectured in 1960 that for any abelian variety A over $\mathbb{Q}$, there is a constant c_A such that

rk $A(K) \leq c_A[K : \mathbb{Q}]$ for every number field K [Hon60, p. 98][2]; this would imply that ranks are bounded in the family of quadratic twists of any elliptic curve over $\mathbb{Q}$.

But from the mid-1960s to the present, it seems that most experts conjectured unboundedness. Cassels in a 1966 survey article [Cas66, p. 257] wrote "it has been widely conjectured that there is an upper bound for the rank depending only on the groundfield. This seems to me implausible because the theory makes it clear that an abelian variety can only have high rank if it is defined by equations with very large coefficients." Tate [Tat74, p. 194] wrote "I would guess that there is no bound on the rank." Mestre, who developed a method for finding elliptic curves of high rank, wrote "Au vu de cette méthode, il semble que l'on puisse sérieusement conjecturer que le rang des courbes elliptiques définies sur $\mathbb{Q}$ n'est pas borné" [Mes82], and proved a rank bound depending on the conductor N of E, namely $O(\log N)$ unconditionally [Mes86, II.1.1], and $O(\log N/\log \log N)$ conditionally on the Riemann hypothesis for $L(E, s)$ [Mes86, II.1.2]. Silverman [Sil09, Conjecture 10.1] wrote that it is a "folklore conjecture" that ranks are unbounded. In 1992 Brumer [Bru92, Section 1] wrote "Today, it is believed that the rank is unbounded," and noted that the available numerical data was "not incompatible with the possibility that, for each r, some positive proportion of all curves might have rank at least r."

Here are two possible reasons for this opinion shift towards unboundedness:

1. Tate and Shafarevich [TŠ67] and Ulmer [Ulm02] constructed families of elliptic curves over $\mathbb{F}_p(t)$ in which the rank is unbounded.
2. Every few years, the proved lower bound on the maximum rank of an elliptic curve over $\mathbb{Q}$ increased: see [RS02, Section 3] for the history up to 2002. The current record is held by Elkies [Elk06], who found an elliptic curve E over $\mathbb{Q}$ of rank ≥ 28, and an infinite family of elliptic curves over $\mathbb{Q}$ of rank ≥ 19.

Some authors have even proposed a rate at which rank can grow relative to the conductor N:

- Ulmer's examples over $\mathbb{F}_p(t)$ attained Brumer's (unconditional) function field analogue [Bru92, Proposition 6.9] of Mestre's conditional upper bound $O(\log N/\log \log N)$. This led Ulmer [Ulm02, Conjecture 10.5] to conjecture that Mestre's conditional bound would be attained over $\mathbb{Q}$, that is, that

$$\limsup_{N\to\infty} \frac{\operatorname{rk} E(\mathbb{Q})}{\log N/\log \log N} \overset{?}{>} 0.$$

- On the other hand, Farmer, Gonek, and Hughes [FGH07, (5.20)], based on conjectures for the maximal size of critical values and the error term in the number of zeros up to a given height for families of L-functions, suggest that

$$\limsup_{N\to\infty} \frac{\operatorname{rk} E(\mathbb{Q})}{\sqrt{\log N \log \log N}} \overset{?}{=} 1, \tag{3.1.1}$$

in contradiction to Ulmer's conjecture.

3.2. Previous heuristics for boundedness.

(a) Rubin and Silverberg [RS00, Remarks 5.1 and 5.2] gave a heuristic based on the expected size of squarefree parts of binary quartic forms. They showed that if certain lattices they define were randomly distributed, then ranks in a family of quadratic twists of a fixed

---

[2]Honda wrote = instead of $\leq$, but almost certainly $\leq$ was intended.

elliptic curve E would be bounded by 8. As they knew, however, the conclusion is wrong for some curves E, e.g., any E of rank greater than 8. Presumably this explains why they did not conjecture boundedness of rank based on this heuristic.

(b) Granville gave a heuristic, discussed in [Wat+14, Section 11] and further developed in [Wat15], based on estimating the number of integer solutions of bounded height to the equation defining a family of elliptic curves. His observation was that a single elliptic curve of high rank would by itself contribute more integer solutions than should be expected for the whole family. Watkins [Wat+14, Section 11.4] writes that similar ideas would lead to the conclusion that all but finitely many elliptic curves E satisfy $\mathrm{rk}\, E(\mathbb{Q}) \leq 21$. See also comments by Conrey, Rubinstein, Snaith, and Watkins [CRSW07, Section 1.3].

These two approaches seem completely unrelated to ours.

3.3. **Conjectures for rank 2 asymptotics.** For each elliptic curve E over $\mathbb{Q}$, let $L(E, s)$ be the L-function of E, and let $w(E) \in \{1, -1\}$ be the sign of its functional equation, or equivalently, the global root number. The Birch and Swinnerton-Dyer conjecture would imply the parity conjecture, that $w(E) \stackrel{?}{=} (-1)^{\mathrm{rk}\, E(\mathbb{Q})}$.

Much of the literature on the distribution of ranks of elliptic curves focuses on quadratic twist families. Fix an elliptic curve E_1 over $\mathbb{Q}$. Let d range over fundamental discriminants in $\mathbb{Z}$. For each d, let E_d denote the twist of E_1 by $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$. Given $r \in \mathbb{Z}_{\geq 0}$ and $D > 0$, define

$$N_{\geq r}(D) := \#\{d : |d| \leq D,\ \mathrm{rk}\, E_d(\mathbb{Q}) \geq r\}$$
$$N_{\geq r,\mathrm{even}}(D) := \#\{d : |d| \leq D,\ \mathrm{rk}\, E_d(\mathbb{Q}) \geq r,\ \text{and}\ w(E_d) = +1\}$$
$$N_{\geq r,\mathrm{odd}}(D) := \#\{d : |d| \leq D,\ \mathrm{rk}\, E_d(\mathbb{Q}) \geq r,\ \text{and}\ w(E_d) = -1\}.$$

There are many different approaches for estimating $N_{\geq 2,\mathrm{even}}(D)$, listed below, but they all lead to the conjecture that

$$N_{\geq 2,\mathrm{even}}(D) \stackrel{?}{=} D^{3/4+o(1)}. \tag{3.3.1}$$

In other words, the prediction is that for d such that $w(E_d) = +1$, the probability that $\mathrm{rk}\, E_d(\mathbb{Q}) \geq 2$ should be about $d^{-1/4}$. Since $\mathrm{ht}\, E_d \asymp d^6$, this prediction corresponds to a probability of $H^{-1/24}$ for an elliptic curve of height H.

(a) Let E be an elliptic curve over $\mathbb{Q}$ with $w(E) = +1$. Then Waldspurger's work [Wal81, Corollaire 2, p. 379] combined with the modularity of E, yields a weight $3/2$ cusp form $f = \sum a_n q^n$ such that for all odd fundamental discriminants $d < 0$ coprime to the conductor of E, we have $a_{|d|} = 0$ if and only if $L(E_d, 1) = 0$ (see also Ono and Skinner [OS98, Section 2, Proof of (2a,b)] and Gross [Gro87, Proposition 13.5]). When $w(E_d) = +1$, the condition $L(E_d, 1) = 0$ is equivalent to $\mathrm{ord}_{s=1} L(E_d, s) \geq 2$, which is equivalent to $\mathrm{rk}\, E_d(\mathbb{Q}) \geq 2$ if the Birch and Swinnerton-Dyer conjecture holds. The Ramanujan conjecture [Sar90, Conjecture 1.3.4] predicts that $a_{|d|}$ is an integer satisfying $|a_{|d|}| \leq |d|^{1/4+o(1)}$, so one might expect that $a_{|d|} = 0$ occurs with "probability" $|d|^{-1/4+o(1)}$. If we ignore the conditions on the sign, parity, and coprimality of d, then summing over $|d| \leq D$ suggests the guess $N_{\geq 2,\mathrm{even}}(D) \stackrel{?}{=} D^{3/4+o(1)}$. This heuristic argument has been attributed to Sarnak [CKRS02, p. 302].

(b) Conrey, Keating, Rubinstein, and Snaith [CKRS02] used random matrix theory to obtain a conjecture more precise than (3.3.1), namely that there exist $c_E, e_E \in \mathbb{R}$ such that

$$N_{\geq 2,\mathrm{even}}(D) \stackrel{?}{=} (c_E + o(1))D^{3/4}(\log D)^{e_E};$$

later, Delaunay and Watkins [DW07] explained how to predict e_E in terms of the 2-torsion of E. The starting point is the Katz–Sarnak philosophy [KS99a, KS99b], based on a function field analogy, that $L(E_d, s)$ should be modeled by the characteristic polynomial of a random matrix from $\mathrm{SO}_{2N}(\mathbb{R})$ for large N (these random matrices seem unrelated to the p-adic and integral matrices in our heuristics). Moment calculations of Keating and Snaith determined the distribution of the values at 1 of the characteristic polynomials [KS00, Section 3.2]. Conrey, Keating, Rubinstein, and Snaith obtained their conjecture by combining this with a discretization heuristic (interpreting sufficiently small L-values as 0).

Watkins [Wat08a] developed a variant for the family of *all* elliptic curves over $\mathbb{Q}$: he conjectured that there exists $c > 0$ such that

$$\#\{E : \operatorname{ht} E \leq H,\ w(E) = +1,\ \text{and}\ \operatorname{rk} E(\mathbb{Q}) \geq 2\} \overset{?}{=} (c + o(1))H^{19/24}(\log H)^{3/8},$$

which is a refined version of what our heuristic predicts. (Watkins counts by discriminant instead of height, but one of his assumptions is that the two counts are comparable [Wat08a, Section 3.4].)
(c) Watkins [Wat08a, Section 4.5] also gave another argument for $H^{19/24+o(1)}$: since the number $Ш_0(E)$ defined in Section 6.4 is expected to be a square integer of size at most $H^{1/12+o(1)}$ (see Theorem 6.4.2(b)), one can guess it is 0 about $H^{-1/24+o(1)}$ of the time, and there are $\asymp H^{20/24}$ elliptic curves in total.
(d) Granville's heuristic (see Section 3.2) would again suggest $H^{19/24+o(1)}$, according to Watkins [Wat15, Section 6] (see also [Wat08a, Section 4.5]).

Our model introduced in Section 1.1, based on yet another approach, again predicts (3.3.1).

3.4. Conjectures for rank 3 asymptotics.
While the conjectures for $N_{\geq 2,\mathrm{even}}(D)$ are in agreement, the conjectures in the literature for $N_{\geq 3,\mathrm{odd}}(D)$ are not.

(a) Rubin and Silverberg [RS01, Theorem 5.4], building on work of Stewart and Top [ST95], showed that the parity conjecture implies the lower bound $N_{\geq 3,\mathrm{odd}}(D) \gg D^{1/3}$ for many E.
(b) Conrey, Rubinstein, Snaith, and Watkins [CRSW07] used random matrix theory as in [CKRS02], but the discretization depends on a lower bound $L'(E_d, 1) \gg d^{-\theta}$ for analytic rank 1 twists E_d, and it is not clear what the best θ is. In fact, they proposed three approaches to suggest a value for θ:
   (1) The Birch and Swinnerton-Dyer conjecture implies a lower bound with $\theta = 1/2$, which leads to $N_{\geq 3,\mathrm{odd}}(D)$ being only about $D^{1/4}$, contradicting the conditional theorem of Rubin and Silverberg above [CRSW07, p. 3].
   (2) An analogy with the class number problem suggests that the lower bound is valid for any $\theta > 0$ [CRSW07, p. 2]; this leads to $N_{\geq 3,\mathrm{odd}}(D) = D^{1-o(1)}$, more than what is conjectured for $N_{\geq 2,\mathrm{even}}(D)$!
   (3) A model involving Heegner points (attributed "largely to Birch" [CRSW07, Section 1.2]) again suggests that any $\theta > 0$ is valid, and hence again that $N_{\geq 3,\mathrm{odd}}(D) = D^{1-o(1)}$.
(c) Conrey, Rubinstein, Snaith, and Watkins suggest another heuristic at the beginning of [CRSW07, Section 1.3], namely that the connection between rank 1 and rank 3 twists

should be the same as between rank 0 and rank 2, at least to first approximation; this
suggests $N_{\geq 3,\mathrm{odd}}(D) = D^{3/4+o(1)}$.

(d) Granville's heuristic, discussed at the end of [CRSW07, Section 1.3], suggests that
$N_{\geq 3}(D) \ll D^{2/3+o(1)}$.

(e) Delaunay and Roblot give heuristics on the moments of regulators that suggest $N_{\geq 3,\mathrm{odd}}(D) = D^{1-o(1)}$ [DR08, p. 608]. (See also [Del05] for related conjectures on the regulators.)

There is also numerical data [Elk02, DD03, Wat08b]. According to Rubin and Silverberg [RS02, p. 466], the numerical data of Elkies suggests that $N_{\geq 3,\mathrm{odd}}(D)$ is about $D^{3/4}$. Watkins writes in [Wat08b, Section 3.2], however, that fitting more extensive data suggests an exponent for $N_{\geq 3,\mathrm{odd}}(D)$ noticeably smaller than the $3/4$ exponent for $N_{\geq 2,\mathrm{even}}(D)$.

Our model, using a single approach that also reproduces the well-known rank 2 conjecture, predicts that $N_{\geq 3}(D) = D^{1/2+o(1)}$ and $N_{\geq 3,\mathrm{odd}}(D) = D^{1/2+o(1)}$. This prediction is different from all those above, but it is consistent with the conditional lower bound of Rubin and Silverberg and with the heuristic upper bound of Granville.

4. Cohen–Lenstra heuristics for class groups

In this section, we give a brief exposition of heuristics for class groups, to motivate Section 5 by analogy. The conjectures are due originally to Cohen and Lenstra [CL84] (with extensions by Cohen and Martinet [CM90]). Following Friedman and Washington [FW89] and Venkatesh and Ellenberg [VE10, Section 4.1], we reinterpret these conjectures in terms of random integer matrices.

4.1. **Class groups as cokernels of integer matrices.** Let K be a number field. Let I be the group of nonzero fractional ideals of K. Let P be the subgroup of I consisting of principal fractional ideals. The class group $\mathrm{Cl}\,K := I/P$ is a finite abelian group.

Let $\mathcal{O}_K$ be the ring of integers of K. Let S_∞ be the set of *all* archimedean places of K. The Dirichlet unit theorem states that the unit group $\mathcal{O}_K^\times$ is a finitely generated abelian group of rank $u := \#S_\infty - 1$.

Let S be a finite set of places of K containing S_∞. Let $n := \#(S - S_\infty)$. Let $\mathcal{O}_{K,S}$ be the ring of S-integers of K. By the Dirichlet S-unit theorem, $\mathcal{O}_{K,S}^\times$ is a finitely generated abelian group of rank $\#S - 1 = n + u$.

Let I_S be the group of fractional ideals generated by the (nonarchimedean) primes in S. Let P_S be the subgroup of I_S consisting of principal fractional ideals, so we obtain an injective homomorphism $I_S/P_S \hookrightarrow I/P = \mathrm{Cl}\,K$. If S is chosen so that its primes generate the finite group $\mathrm{Cl}\,K$, then $I_S/P_S \simeq I/P = \mathrm{Cl}\,K$.

The group I_S is a free abelian group of rank n. Since P_S is the image of the homomorphism $\mathcal{O}_{K,S}^\times \to I_S$, whose kernel is the torsion subgroup of $\mathcal{O}_{K,S}^\times$, the group P_S is a free abelian group of rank $n + u$. If we choose bases, then we represent $\mathrm{Cl}\,K$ as the cokernel of a homomorphism $\mathbb{Z}^{n+u} \to \mathbb{Z}^n$. We write this cokernel as $\mathrm{coker}\,A$ for some $n \times (n + u)$ matrix A over $\mathbb{Z}$. If we view this same A as a matrix over $\mathbb{Z}_p$, then $\mathrm{coker}(A\colon \mathbb{Z}_p^{n+u} \to \mathbb{Z}_p^n) = (\mathrm{Cl}\,K)[p^\infty]$.

Remark 4.1.1. Friedman and Washington [FW89] were the first to model (the Sylow p-subgroups of) class groups as cokernels of matrices, but they arrived at such a model via a different argument. Specifically, they considered the function field analogue, in which case $(\mathrm{Cl}\,K)[p^\infty]$ for $p \neq \mathrm{char}\,K$ can be understood in terms of the action of Frobenius on the Tate module T_pJ of the Jacobian J of a curve over a finite field. It was only later that Venkatesh

and Ellenberg [VE10, Section 4.1] noticed the connection with the presentation of the class group given above.

4.2. Heuristics for class groups.
Let $\mathscr{K}$ be the family of all imaginary quadratic fields up to isomorphism. What is the distribution of $\mathrm{Cl}\,K$ as K varies over $\mathscr{K}$? To formulate this question precisely, we order the fields by their discriminant $D := \mathrm{disc}\,K$. For $X > 0$, let $\mathscr{K}_{\leq X} := \{K \in \mathscr{K} : |\mathrm{disc}\,K| \leq X\}$. Define the **density** of a subset $S \subset \mathscr{K}$ by

$$\mu(S) = \mu(S \mid K \in \mathscr{K}) := \lim_{X \to \infty} \frac{\#(S \cap \mathscr{K}_{\leq X})}{\#\mathscr{K}_{\leq X}}$$

when this limit exists.

Hecke, Deuring, and Heilbronn proved that $\#\,\mathrm{Cl}\,K \to \infty$ as $|D| \to \infty$, and soon thereafter Siegel proved $\#\,\mathrm{Cl}\,K = |D|^{1/2+o(1)}$; see the appendix to Serre [Ser97] for the history. Therefore, for any finite abelian group G, the set $\{K \in \mathscr{K} : \mathrm{Cl}\,K \simeq G\}$ is finite, so $\mu(\mathrm{Cl}\,K \simeq G) = 0$.

To get subsets of positive density, we instead examine the p-Sylow subgroup $(\mathrm{Cl}\,K)[p^\infty]$ for a fixed prime $p \neq 2$ (The case $p = 2$ is different because of genus theory; Gerth [Ger87] formulated analogous conjectures by considering $(\mathrm{Cl}\,K)^2[2^\infty]$ instead, and Fouvry and Klüners [FK07] proved that $(\mathrm{Cl}\,K)^2[2]$ is distributed as Gerth conjectured.) For each finite abelian p-group G, the density $\mu((\mathrm{Cl}\,K)[p^\infty] \simeq G)$ is conjecturally positive, and there are two conjectures for its value, as follows.

(4.2.1) The density is inversely proportional to $\#\,\mathrm{Aut}\,G$:

$$\mu((\mathrm{Cl}\,K)[p^\infty] \simeq G) \overset{?}{=} \frac{(\#\,\mathrm{Aut}\,G)^{-1}}{\eta(p)},$$

where the normalization constant $\eta(p)$ needed for a probability distribution is given by Hall [Hal38] as

$$\eta(p) := \sum_{\substack{\text{finite abelian} \\ p\text{-groups } G}} (\#\,\mathrm{Aut}\,G)^{-1} = \prod_{i=1}^{\infty}(1 - p^{-i})^{-1}.$$

(4.2.2) Inspired by Section 4.1, with unit rank $u = \#S_\infty - 1 = 0$, one models $(\mathrm{Cl}\,K)[p^\infty]$ as $(\mathrm{coker}\,A)[p^\infty]$ for a "random" $n \times n$ matrix A over $\mathbb{Z}$ or $\mathbb{Z}_p$:

$$\mu((\mathrm{Cl}\,K)[p^\infty] \simeq G) \overset{?}{=} \lim_{n \to \infty} \lim_{X \to \infty} \frac{\#\{A \in \mathrm{M}_n(\mathbb{Z})_{\leq X} : (\mathrm{coker}\,A)[p^\infty] \simeq G\}}{\#\,\mathrm{M}_n(\mathbb{Z})_{\leq X}}$$
$$= \lim_{n \to \infty} \mathrm{Prob}(\mathrm{coker}\,A \simeq G \mid A \in \mathrm{M}_n(\mathbb{Z}_p)).$$

(Recall our conventions in Section 2 for these probabilities; the equality of the probabilities in the last two expressions follows from the asymptotic equidistribution of $\mathbb{Z}$ in $\mathbb{Z}_p$. The equality of the limits in the last two expressions is very robust; it holds when when we replace $A \in \mathrm{M}_n(\mathbb{Z})$ by drawing A from much more general distributions of integral matrices [Woo19].)

Conjecture (4.2.1) is due to Cohen and Lenstra [CL84]; they were motivated by numerical data and the general principle that an object should be counted with weight inversely proportional to the size of its automorphism group. Conjecture (4.2.2) in the second form

$$\mu((\mathrm{Cl}\,K)[p^\infty] \simeq G) \overset{?}{=} \lim_{n \to \infty} \mathrm{Prob}(\mathrm{coker}\,A \simeq G \mid A \in \mathrm{M}_n(\mathbb{Z}_p))$$

is due to Friedman and Washington [FW89].

In fact, Conjectures 4.2.1 and 4.2.2 are equivalent:

Theorem 4.2.3 (Friedman and Washington [FW89]). *For every finite abelian p-group G,*

$$\lim_{n \to \infty} \operatorname{Prob}(\operatorname{coker} A \simeq G \mid A \in \operatorname{M}_n(\mathbb{Z}_p)) = \frac{(\# \operatorname{Aut} G)^{-1}}{\eta(p)} = \frac{1}{\# \operatorname{Aut} G} \prod_{i=1}^{\infty} (1 - p^{-i}).$$

If instead we consider the family of *real* quadratic fields, then the unit rank u is 1, so Section 4.1 suggests that $\operatorname{Cl} K$ should be modeled by the cokernel of an $n \times (n+1)$ matrix, in which case there is a similar story to that above.

5. Heuristics for Shafarevich–Tate groups

In this section, we consider heuristics for the Shafarevich–Tate group of an elliptic curve over $\mathbb{Q}$, analogous to the heuristics for class groups in the previous section.

5.1. **Elliptic curves.** An elliptic curve E over $\mathbb{Q}$ is isomorphic to the projective closure of a curve $y^2 = x^3 + Ax + B$ for a unique pair of integers (A, B) such that there is no prime p such that $p^4 \mid A$ and $p^6 \mid B$. Conversely, any such pair (A, B) with $4A^3 + 27B^2 \neq 0$ defines an elliptic curve over $\mathbb{Q}$. Let $\mathscr{E}$ be the set of elliptic curves of this form, one in each $\mathbb{Q}$-isomorphism class.

Define the (naive) height of $E \in \mathscr{E}$ by

$$\operatorname{ht} E := \max(|4A^3|, |27B^2|).$$

Let $\mathscr{E}_{\leq H} := \{E \in \mathscr{E} : \operatorname{ht} E \leq H\}$. An elementary sieve argument [Bru92, Lemma 4.3] shows that

$$\# \mathscr{E}_{\leq H} = (\kappa + o(1)) H^{5/6}, \tag{5.1.1}$$

where $\kappa := 2^{4/3} 3^{-3/2} \zeta(10)^{-1}$.

For a subset $S \subseteq \mathscr{E}$, we define densities

$$\mu(S) := \lim_{H \to \infty} \frac{\#(S \cap \mathscr{E}_{\leq H})}{\# \mathscr{E}_{\leq H}}$$

$$\mu(S \mid \operatorname{rk} E(\mathbb{Q}) = r) := \lim_{H \to \infty} \frac{\#\{E \in S \cap \mathscr{E}_{\leq H} \mid \operatorname{rk} E(\mathbb{Q}) = r\}}{\#\{E \in \mathscr{E}_{\leq H} \mid \operatorname{rk} E(\mathbb{Q}) = r\}},$$

when the limits exist.

Remark 5.1.2. If for some r, there are no $E \in \mathscr{E}$ such that $\operatorname{rk} E(\mathbb{Q}) = r$, then the density $\mu(S \mid \operatorname{rk} E(\mathbb{Q}) = r)$ does not exist!

Remark 5.1.3. Elliptic curves can be ordered in other ways, such as by minimal discriminant or conductor. It is still true that the set of $E \in \mathscr{E}$ of minimal discriminant or conductor up to X is finite, but there is no unconditional estimate for its size, even though for most E (ordered by height), the minimal discriminant and conductor are of the same order of magnitude as the height. See Watkins [Wat08a, Section 4] for further discussion. Hortsch [Hor15] recently succeeded in counting elliptic curves of bounded *Faltings height*, however.

Associated to an elliptic curve $E \in \mathscr{E}$ are other invariants: the n-**Selmer group** $\mathrm{Sel}_n E$ for each $n \geq 1$ and the **Shafarevich–Tate group** $\Sha(E)$; see Silverman [Sil92, Chapter 10]. These invariants are related by an exact sequence

$$0 \to \frac{E(\mathbb{Q})}{nE(\mathbb{Q})} \to \mathrm{Sel}_n E \to \Sha(E)[n] \to 0$$

for each $n \geq 1$. Taking the direct limit as n ranges over powers of a prime p yields the exact sequence

$$0 \to E(\mathbb{Q}) \otimes \frac{\mathbb{Q}_p}{\mathbb{Z}_p} \to \mathrm{Sel}_{p^\infty} E \to \Sha(E)[p^\infty] \to 0. \qquad (5.1.4)$$

Instead of trying to predict a distribution for $\mathrm{rk}\, E(\mathbb{Q})$ in isolation, we model all three invariants at once. This lets us check our model against other theorems and conjectures in the literature.

5.2. Symplectic finite abelian groups.
We will soon focus on $\Sha(E)$, which is an abelian group with extra structure that we now describe.

Definition 5.2.1. A **symplectic finite abelian group** is a pair $(G, [\ ,\ ])$, where G is a finite abelian group and $[\ ,\ ]\colon G \times G \to \mathbb{Q}/\mathbb{Z}$ is a nondegenerate alternating pairing.

An isomorphism of symplectic finite abelian groups is an isomorphism of groups that respects the pairings. It turns out that if two symplectic finite abelian groups are isomorphic as abstract groups, there is automatically an isomorphism that respects the pairings. Let $\mathfrak{S}$ be a set of symplectic finite abelian groups containing exactly one from each isomorphism class. If J is a finite abelian group, then $J \times J^\vee$ equipped with a natural pairing is a symplectic finite abelian group, and every symplectic finite abelian group is isomorphic to one of this form. In particular, symplectic finite abelian groups have square order.

Let $\mathfrak{S}_p$ be the set of $G \in \mathfrak{S}$ such that $\#G$ is a power of p.

5.3. Distribution of the Shafarevich–Tate group.
It is widely conjectured that $\Sha(E)$ is finite. Cassels [Cas62] constructed a alternating pairing

$$\langle\ ,\ \rangle\colon \Sha(E) \times \Sha(E) \to \mathbb{Q}/\mathbb{Z}.$$

He proved also that if $\Sha(E)$ is finite, then $\langle\ ,\ \rangle$ is nondegenerate. In this case, $\Sha(E)$ equipped with $\langle\ ,\ \rangle$ is a symplectic finite abelian group, and in particular $\#\Sha(E)$ is a square. This already shows that the distribution of $\Sha(E)$ will be different from the conjectural distribution of class groups in Section 4.2.

The distribution of class groups conjecturally depended on the unit rank of the number field; analogously, the distribution of $\Sha(E)$ should depend on the rank of E.

Question 5.3.1. Fix a prime p. Given $r \geq 0$ and $G \in \mathfrak{S}_p$, what is the density

$$\mu(\Sha(E)[p^\infty] \simeq G \mid \mathrm{rk}\, E(\mathbb{Q}) = r)?$$

There are three conjectural answers to this question:

$(\mathscr{D}_r)$ Delaunay [Del01, Del07, DJ14], in analogy with the Cohen–Lenstra heuristics for class groups, made conjectures on the distribution of $\Sha(E)$ as E varies over elliptic curves of rank r. He ordered elliptic curves by conductor; but if we modify his conjectures

to order by height, they imply that the answer to Question 5.3.1 is given by the probability measure $\mathscr{D}_r = \mathscr{D}_{r,p}$ on $\mathfrak{S}_p$ defined by

$$\mathrm{Prob}_{\mathscr{D}_r}(G) := \frac{\#G^{1-r}}{\#\operatorname{Aut}G} \prod_{i \geq r+1} (1 - p^{1-2i}), \tag{5.3.2}$$

where $\operatorname{Aut}G$ denotes the group of automorphisms of G *that respect the pairing.*

($\mathscr{T}_r$) Work of Poonen and Rains [PR12] and Bhargava, Kane, Lenstra, Poonen, and Rains [BKLPR15] predicted the distribution of the isomorphism type of $\operatorname{Sel}_p E$ and the short exact sequence (5.1.4), respectively, and these were shown to be compatible with some known properties of the arithmetic of E. From these, one extracts a probability measure $\mathscr{T}_r$ on $\mathfrak{S}_p$ conjectured to model $\operatorname{III}(E)[p^\infty]$.

($\mathscr{A}_r$) The article [BKLPR15], in analogy with the Friedman–Washington interpretation of class group heuristics, proposed also another probability measure, $\mathscr{A}_r$, inspired by the observation that if $A \in \mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}}$, then $\operatorname{coker}(A : \mathbb{Z}_p^n \to \mathbb{Z}_p^n)_{\mathrm{tors}}$ is naturally a symplectic finite abelian p-group. Specifically, for $n \equiv r \pmod 2$, there is a canonical probability measure on the set

$$\{A \in \mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}} : \mathrm{rk}_{\mathbb{Z}_p}(\ker A) = r\},$$

and if $G \in \mathfrak{S}_p$, we let $\mathscr{A}_{n,r}(G)$ be the measure of

$$\{A \in \mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}} : \mathrm{rk}_{\mathbb{Z}_p}(\ker A) = r \text{ and } (\operatorname{coker} A)_{\mathrm{tors}} \simeq G\}.$$

Then the formula

$$\mathscr{A}_r(G) := \lim_{\substack{n \to \infty \\ n \equiv r \pmod 2}} \mathscr{A}_{n,r}(G)$$

defines a probability measure $\mathscr{A}_r$ on $\mathfrak{S}_p$.

Theorem 5.3.3 ([BKLPR15, Theorems 1.6(c) and 1.10(b)])**.** *The probability measures $\mathscr{D}_r$, $\mathscr{T}_r$, $\mathscr{A}_r$ coincide.*

Remark 5.3.4. Conjecturally, $\operatorname{III}(E)$ is large on average when $r = 0$ and small when $r \geq 1$, just as class groups of quadratic fields are large if the field is imaginary ($u = 0$) and conjecturally small if the field is real ($u = 1$). More precisely, it follows from Delaunay's conjectures on $\operatorname{III}(E)$ mentioned above that $\mu(\#\operatorname{III}(E) \leq B \mid \mathrm{rk}\,E(\mathbb{Q}) = 0) = 0$ for all $B > 0$, but for each $r \geq 1$ that $\mu(\#\operatorname{III}(E) \leq B \mid \mathrm{rk}\,E(\mathbb{Q}) = r) \to 1$ as $B \to \infty$. In fact, for fixed $r \geq 1$, Delaunay's conjectures predict for each $G \in \mathfrak{S}$ that $\mu(\operatorname{III}(E) \simeq G \mid \mathrm{rk}\,E(\mathbb{Q}) = r)$ is an explicit positive number, and these numbers define a measure on $\mathfrak{S}$ that agrees with the product over all primes of the measures $\mathscr{D}_r$. See also [BKLPR15, Section 5.6] for further discussion.

6. Average size of the Shafarevich–Tate group

Section 7 will propose a model for ranks and III. To set the parameters in that model, we will need to know the typical size of $\#\operatorname{III}(E)$ for a rank 0 elliptic curve of height about H. Our approach to estimating $\#\operatorname{III}(E)$ is similar to that in Lang [Lan83]; see also work of Goldfeld and Szpiro [GS95], de Weger [dW98], Hindry [Hin07], Watkins [Wat08a], and Hindry and Pacheco [HP16]. Although more precise results are known, we provide a streamlined version of these estimates that is sufficient for our purposes.

6.1. Size of the real period.

Lemma 6.1.1. *Let $A, B \in \mathbb{R}$ satisfy $4A^3 + 27B^2 \neq 0$, so that the equation $y^2 = x^3 + Ax + B$ defines an elliptic curve E over $\mathbb{R}$. Let $\Delta := -16(4A^3 + 27B^2)$, let $H := \max(|4A^3|, |27B^2|)$, and let $\Omega := \int_{E(\mathbb{R})} \left| \frac{dx}{2y} \right|$. Then*

$$H^{-1/12} \ll \Omega \ll H^{-1/12} \log(H/|\Delta|).$$

Proof. Changing (A, B) to $(\lambda^4 A, \lambda^6 B)$ with $\lambda \in \mathbb{R}^\times$ changes (H, Δ, Ω) to $(\lambda^{12} H, \lambda^{12} \Delta, \lambda^{-1} \Omega)$, so we may assume that (A, B) lies on the rectangle boundary where $H = 1$. By compactness, the bounds hold on this rectangle boundary except possibly as (A, B) approaches one of the two corners where $\Delta = 0$. Up to scaling by a λ bounded away from 0 and ∞, these are the curves $\pm y^2 = 4x^3 - g_4(\tau)x - g_6(\tau)$ for $\tau = it$ or $\tau = 1/2 + it$ as $t \to \infty$ (in the part of the fundamental domain outside a compact set, these are the τ such that $\mathbb{Z}\tau + \mathbb{Z}$ is homothetic to its complex conjugate). In these families, each of the Eisenstein series g_4 and g_6 tends to a finite nonzero limit, so H remains bounded, while $|\Delta| \asymp |q| = |e^{2\pi i \tau}| = e^{-2\pi \operatorname{im} \tau}$, and Ω is 1 or $\operatorname{im} \tau$ up to a bounded factor, so $1 \ll \Omega \ll \log(1/|\Delta|)$. $\square$

Corollary 6.1.2 ([Wat08a, Section 6.2]). *Under the hypotheses of Lemma 6.1.1, we have $\Omega \ll |\Delta|^{-1/12}$.*

Proof. We have $|\Delta| \ll H$. Then $H^{-1/12} \log(H/|\Delta|) \ll |\Delta|^{-1/12}$ since $x^{1/12} \log(1/x)$ remains bounded as $x \to 0^+$. $\square$

Corollary 6.1.3 (cf. [Hin07, Lemma 3.7]). *If $E \in \mathscr{E}$, then $H^{-1/12} \ll \Omega \ll H^{-1/12} \log H$.*

Proof. If $E \in \mathscr{E}$, then Δ is a nonzero integer, so $|\Delta| \geq 1$. Substitute this into Lemma 6.1.1. $\square$

Remark 6.1.4. Corollary 6.1.3 is similar to the theorem relating the naive height to the Faltings height [Sil86, second statement of Corollary 2.3], except that the Faltings height is defined using the covolume of the period lattice instead of just the real period.

Remark 6.1.5. The bounds in Corollary 6.1.3 are best possible, up to constants. For example, for large $a \in \mathbb{Z}_{>0}$, the curve $y^2 = (x - a)(x - a - 1)(x + 2a + 1)$ has $H \asymp a^6$ and $\Omega \asymp a^{-1/2} \log a \asymp H^{-1/12} \log H$; this shows that the upper bound is sharp.

Remark 6.1.6. If instead of a short Weierstrass model we use the minimal Weierstrass model $y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6$ and a Néron differential $\omega := \frac{dx}{2y + a_1 x + a_3}$, then ω differs from $\frac{dx}{2y}$ by bounded powers of 2 and 3. So if we define Ω using the Néron differential in place of $\frac{dx}{2y}$, the estimates in Corollary 6.1.3 are still valid. It is this Ω that appears in the Birch and Swinnerton-Dyer conjecture.

Remark 6.1.7. Some authors define the real period as the integral of a Néron differential over only *one component* of $E(\mathbb{R})$.

6.2. The product of the Tamagawa factors.
Consider $E \in \mathscr{E}$ of height about H. Let $\mathcal{E}$ be the Néron model of E over $\mathbb{Z}$. For each prime p, let Φ_p be the component group (scheme) of the special fiber $\mathcal{E}_{\mathbb{F}_p}$, and define the Tamagawa factor $c_p := \#\Phi_p(\mathbb{F}_p)$.

Lemma 6.2.1. *We have $\prod_p c_p = H^{o(1)}$.*

Compare this lemma with work of de Weger [dW98, Theorem 3], Hindry [Hin07, Lemma 3.5], and Watkins [Wat08a, pp. 114–115].

Proof. For $n \geq 1$, let $\sigma_0(n)$ denote the number of positive divisors of n. Factor the minimal discriminant Δ of E as $\prod_p p^{e_p}$. Whenever $e_p > 0$, Kodaira and Néron proved that $c_p \leq 4$ or $c_p = e_p$ [Sil09, Theorem VII.6.1], so in any case $c_p \leq (e_p + 1)^2 = \sigma_0(p^{e_p})^2$. Thus

$$\prod_p c_p \leq \sigma_0(\Delta)^2 = \left(\Delta^{o(1)}\right)^2 = H^{o(1)}. \qquad \square$$

Remark 6.2.2. If instead of $\sigma_0(n) = n^{o(1)}$ we used the more precise bound $\sigma_0(n) \leq n^{O(1/\log\log n)}$, we would get a direct proof of [dW98, Theorem 3], which states that $\prod_p c_p \leq \Delta^{O(1/\log\log \Delta)}$.

6.3. **Average size of $L(E, 1)$.** The Riemann hypothesis for the L-functions $L(E, s)$ would imply the corresponding Lindelöf hypothesis [IS00, p. 713], which in turn would imply

$$L(E, 1) \overset{?}{\leq} H^{o(1)}. \tag{6.3.1}$$

For our calibration, however, we need only estimate *averages* of $L(E, 1)$, so we conjecture the following.

Conjecture 6.3.2. We have $\underset{E \in \mathscr{E}_{\leq H}}{\text{Average }} L(E, 1) \overset{?}{=} H^{o(1)}$ as $H \to \infty$.

In quadratic twist families, the following stronger (unconditional) variant of Conjecture 6.3.2 is known.

Lemma 6.3.3. *Let E_1 be an elliptic curve over $\mathbb{Q}$. Let E_d be its twist by $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$. Given $D > 0$, let d range over fundamental discriminants satisfying $|d| \leq D$. Then* $\underset{|d|\leq D}{\text{Average }} L(E_d, 1) \asymp 1$ *as $D \to \infty$.*

Proof. This is a consequence of work of Kohnen and Zagier [KZ81, Corollaries 5 and 6]. $\square$

Remark 6.3.4. Lemma 6.3.3 makes plausible the conjecture that

$$\underset{E \in \mathscr{E}_{\leq H}}{\text{Average }} L(E, 1) \overset{?}{\asymp} 1 \tag{6.3.5}$$

as $H \to \infty$. This conjecture, slightly stronger than Conjecture 6.3.2, would not affect the calibration of our heuristic here, but is interesting in its own right. Young [You06] proved that (6.3.5) holds under the Riemann hypothesis for Dirichlet L-functions and the equidistribution of the root number of elliptic curves.

6.4. **Average size of the Shafarevich–Tate group.** Let $E \in \mathscr{E}$. Define

$$\text{Ш}_0(E) := \begin{cases} \#\text{Ш}(E), & \text{if rk } E(\mathbb{Q}) = 0; \\ 0, & \text{if rk } E(\mathbb{Q}) > 0. \end{cases}$$

Then the "rank 0 part" of the Birch and Swinnerton-Dyer conjecture states that

$$L(E, 1) \overset{?}{=} \frac{\text{Ш}_0(E)\, \Omega \prod_p c_p}{\#E(\mathbb{Q})^2_{\text{tors}}}; \tag{6.4.1}$$

see Wiles [Wil06] for an exposition and Stein and Wuthrich [SW13, Section 8] for a summary of some more recent advances towards it.

Theorem 6.4.2 (cf. [Lan83, Conjecture 1]). *Assume the Birch and Swinnerton-Dyer conjecture. Then the following hold.*

(a) *For $E \in \mathscr{E}$ of height H, we have $\text{Ш}_0(E) = H^{1/12+o(1)} L(E, 1)$.*

(b) *For $E \in \mathscr{E}$ of height H, if the Riemann hypothesis for $L(E, s)$ holds, then $\text{Ш}_0(E) \leq H^{1/12+o(1)}$.*

(c) *If Conjecture 6.3.2 holds, then $\underset{E \in \mathscr{E}_{\leq H}}{\text{Average}} \text{Ш}_0(E) = H^{1/12+o(1)}$ as $H \to \infty$.*

Proof. By Mazur [Maz77], we have $\#E(\mathbb{Q})_{\text{tors}} \leq 16$. By Corollary 6.1.3 and Remark 6.1.6, we have $\Omega = H^{-1/12+o(1)}$. By Lemma 6.2.1, we have $\prod c_p = H^{o(1)}$. Substitute all this into (6.4.1) to obtain (a). Combine (a) with (6.3.1) to obtain (b). Combine (a) with Conjecture 6.3.2 to obtain (c). $\qquad\square$

Remark 6.4.3. Theorem 6.4.2(c) agrees with a conjecture of Heath-Brown and with numerical investigations of Dąbrowski, Jędrzejak, and Szymaszkiewicz [DJS16, Section 7].

Remark 6.4.4. In a family of quadratic twists E_d, we have ht $E_d \asymp d^6$, so Theorem 6.4.2(b) would imply $\text{Ш}_0(E_d) \overset{?}{\leq} d^{1/2+o(1)}$ as $d \to +\infty$. This is consistent with the work of Waldspurger [Wal81, Corollaire 2, p. 379] relating $\sqrt{\text{Ш}_0(E_d)}$ to the dth coefficient a_d of a weight $3/2$ modular form, since for such a form we expect $|a_d| \leq d^{1/4+o(1)}$.

7. The basic model for ranks and Shafarevich–Tate groups

The construction of the measure $\mathscr{A}_r$ in Section 5.3 involved alternating matrices that modeled Shafarevich–Tate groups of elliptic curves of rank r. Specifically, the matrices were required to have corank r. Inspired by this model and interested in the distribution of ranks among all elliptic curves, we propose the following model for the arithmetic of an elliptic curve E over $\mathbb{Q}$ of height H. Informally, to each elliptic curve E we will associate a random matrix $A \in \mathrm{M}_n(\mathbb{Z})_{\text{alt}, \leq X}$ such that $\text{rk}(\ker A)$ models $\text{rk}\, E(\mathbb{Q})$ and $(\text{coker}\, A)_{\text{tors}}$ models $\text{Ш}(E)$. A more precise version of our model depends on increasing functions $\eta(H)$ and $X(H)$ to be calibrated later, with $\eta(H), X(H) \to \infty$ as $H \to \infty$.

7.1. The random model.
We now define a collection of independent random variables $(\text{rk}'_E, \text{Ш}'_E)_{E \in \mathscr{E}}$ taking values in $\mathbb{Z}_{\geq 0} \times \mathfrak{S}$. These random variables will be defined as functions of random matrices, and the only input from the elliptic curve E will be its height.

To define the random variable with index E, let $H := \text{ht}\, E$, choose n uniformly at random from $\mathbb{Z} \cap [\eta(H), \eta(H) + 2)$, choose $A \in \mathrm{M}_n(\mathbb{Z})_{\text{alt}, \leq X(H)}$ uniformly at random, define $\text{rk}'_E := \text{rk}(\ker A)$, and define $\text{Ш}'_E$ to be $(\text{coker}\, A)_{\text{tors}}$ equipped with its canonically defined nondegenerate alternating pairing [BKLPR15, Sections 3.4 and 3.5].

Remark 7.1.1. Replacing $[\eta(H), \eta(H)+2)$ with any other interval of length $o(\eta(H))$ containing $\eta(H)$ would not affect our results as long as the parity of n becomes equidistributed as $H \to \infty$.

In the rest of this section, we will prove unconditional theorems about random integral alternating matrices, in particular about the statistical behavior of rk'_E and $\text{Ш}'_E$ as E varies. These will inform our conjectures about $\text{rk}\, E(\mathbb{Q})$ and $\text{Ш}(E)$.

7.2. **First results on random matrices.** Define the random variable

$$\text{III}'_{0,E} := \begin{cases} \#\text{III}'_E, & \text{if } \text{rk}'_E = 0; \\ 0, & \text{if } \text{rk}'_E > 0. \end{cases}$$

We first prove a theorem about the individual random variables $(\text{rk}'_E, \text{III}'_E)$.

Theorem 7.2.1. *If the function $X(H)$ grows sufficiently quickly relative to $\eta(H)$, then the following hold for $E \in \mathcal{E}$ as $H := \text{ht } E \to \infty$.*

(a) (0) *The probability that $\text{rk}'_E = 0$ is $1/2 - o(1)$.*
   (1) *The probability that $\text{rk}'_E = 1$ is $1/2 - o(1)$.*
   (2) *The probability that $\text{rk}'_E \geq 2$ is $o(1)$.*
(b) *Let $r \in \{0, 1\}$ and $G \in \mathfrak{S}_p$. Then*

$$\text{Prob}(\text{III}'_E[p^\infty] \simeq G \mid \text{rk}'_E = r) = \text{Prob}_{\mathscr{D}_{r,p}}(G) + o(1).$$

   (See (5.3.2) for an explicit formula for $\text{Prob}_{\mathscr{D}_{r,p}}(G)$.)
(c) (1) *Let $G \in \mathfrak{S}$. Then*

$$\text{Prob}(\text{III}'_E \simeq G \mid \text{rk}'_E = 1) = \prod_p \text{Prob}_{\mathscr{D}_{1,p}}(G[p^\infty]) + o(1).$$

   (2) *More generally, if $\mathcal{G} \subseteq \mathfrak{S}$, then*

$$\text{Prob}(\text{III}'_E \in \mathcal{G} \mid \text{rk}'_E = 1) = \sum_{G \in \mathcal{G}} \prod_p \text{Prob}_{\mathscr{D}_{1,p}}(G[p^\infty]) + o(1).$$

(d) (1) *Let $G \in \mathfrak{S}$. Then $\text{Prob}(\text{III}'_E \simeq G \mid \text{rk}'_E = 0) = o(1)$.*
   (2) *If $\mathcal{G}$ is the set of squares of cyclic groups, then*

$$\text{Prob}(\text{III}'_E \in \mathcal{G} \mid \text{rk}'_E = 0) = \prod_p \left(1 - \frac{1}{p^2} + \frac{1}{p^3}\right) + o(1).$$

(e) (1) *We have $\text{III}'_{0,E} \leq (X(H)^{\eta(H)})^{1+o(1)}$.*
   (2) *The probability that $\text{III}'_{0,E} \geq (X(H)^{\eta(H)})^{1-o(1)}$ is at least $1/3$.*
(f) *For fixed $r \geq 1$, we have $\text{Prob}(\text{rk}'_E \geq r) = (X(H)^{\eta(H)})^{-(r-1)/2+o(1)}$.*

The proof of (f) will require the main theorem of Section 9, while the proofs of (a)–(e) are comparatively straightforward (although some of them require the Ekedahl sieve). The constant $1/3$ in (e)(2) could be improved to any constant less than $1/2$, as will be clear from the proof.

Proof. Let $X := X(H)$ and $\eta := \eta(H)$. Any constant depending on n can be assumed to be $X^{o(1)}$ if X grows sufficiently quickly relative to η.

(a) Since we choose n uniformly in $\mathbb{Z} \cap [\eta, \eta + 2)$, it is even half of the time and odd half of the time. Any alternating matrix has even rank, and a generic alternating matrix of rank n has rank n or $n - 1$ according to whether n is even or odd. As $X \to \infty$ for fixed n, the probability that an *integer* matrix $A \in \text{M}_n(\mathbb{Z})_{\text{alt}, \leq X}$ has the generic rank tends to 1 (it fails on integer points in a proper Zariski-closed subset). It follows formally that the same holds if X tends to ∞ sufficiently quickly relative to η. Thus $\text{Prob}(\text{rk}'_E = 0) = 1/2 - o(1)$ as $H \to \infty$, and the other statements follow similarly.

17

(b) For a fixed k, and any n, the set $\mathrm{M}_n(\mathbb{Z})_{\mathrm{alt},\leq X}$ becomes equidistributed in $\mathrm{M}_n(\mathbb{Z}/p^k\mathbb{Z})$ as $X \to \infty$. When n is even, the same holds for the subset of $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt},\leq X}$ satisfying $\dim(\ker A) = 0$ (these are the ones in a nonempty Zariski-open subset). Given G, there exists a positive integer k such that the condition $\Sha'_E[p^\infty] \simeq G$ depends only on A modulo p^k. Thus, if X is sufficiently large relative to n, then

$$\mathrm{Prob}(\Sha'_E[p^\infty] \simeq G \mid \mathrm{rk}'_E = 0) = \mathrm{Prob}_{\mathscr{A}_0}(G) + o(1)$$

as $H \to \infty$. By Theorem 5.3.3, $\mathscr{A}_0$ coincides with $\mathscr{D}_0 = \mathscr{D}_{0,p}$. An analogous argument applies if we condition on rk'_E being 1.

(c) (1) We apply the Ekedahl sieve as adapted by Poonen and Stoll in [PS99, Section 9.3]. Consider a large odd integer n. Let $U_\infty = \mathrm{M}_n(\mathbb{R})_{\mathrm{alt}}$. For each prime p, let U_p be the set of $A \in \mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}}$ such that $(\mathrm{coker}\,A)_{\mathrm{tors}}[p^\infty] \not\simeq G[p^\infty]$. Let s_p be the Haar measure of U_p. The image of U_p in $\mathrm{M}_n(\mathbb{F}_p)_{\mathrm{alt}}$ is contained in the set of $\mathbb{F}_p$-points of the subscheme of $\mathbb{A}_{\mathbb{Z}}^{n^2}$ parametrizing matrices of corank ≥ 3, and this implies that $s_p = O(1/p^2)$ uniformly in n. Now, [PS99, Lemma 21] implies that hypothesis (10) in [PS99, Lemma 20] holds. The conclusion of [PS99, Lemma 20] for $S := \emptyset$ implies that the density of $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}}$ satisfying $(\mathrm{coker}\,A)_{\mathrm{tors}} \simeq G$ equals $\prod_p(1 - s_p)$. Because of the uniform estimate on s_p, we may take the limit as $n \to \infty$ *inside* the product, in which case $1 - s_p$ tends to $\mathrm{Prob}_{\mathscr{D}_{1,p}}(G[p^\infty])$ by Theorem 5.3.3, so (1) follows.
(2) This follows formally from (1) and the fact that $\sum_{G \in \mathfrak{G}} \prod_p \mathrm{Prob}_{\mathscr{D}_{1,p}}(G[p^\infty]) = 1$.

(d) (1) Since $\mathrm{rk}'_E = 0$, we have $\#\Sha'_E = |\det A|$. By the same reasoning as in the proof of (a), the probability that an integer matrix $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt},\leq X}$ has $|\det A|$ equal to a fixed value tends to 0 if X tends to ∞ sufficiently quickly relative to η.
(2) This follows from the Ekedahl sieve as in the proof of (c)(1) above, since the "square of cyclic" condition can be checked on the p-primary part one p at a time, and for each p the reductions modulo p of the $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}}$ such that $(\mathrm{coker}\,A)[p^\infty]$ is *not* cyclic lie in the $\mathbb{F}_p$-points of a subscheme of $\mathbb{A}_{\mathbb{Z}}^{n^2}$ of codimension ≥ 2.

(e) By (a), we have $\mathrm{Prob}(\mathrm{rk}'_E = 0) = 1/2 - o(1)$. If $\mathrm{rk}'_E > 0$, then $\Sha'_{0,E} = 0$. If $\mathrm{rk}'_E = 0$, then $\Sha'_{0,E}$ is the absolute value of the determinant of a random $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt},\leq X}$, which is the absolute value of a degree n polynomial evaluated on a box of dimensions very large relative to n. This implies that there are constants $m_n, M_n > 0$ depending only on n such that $\Sha'_{0,E} \leq M_n X^n$ and such that $\mathrm{Prob}(\Sha'_{0,E} \geq m_n X^n \mid \mathrm{rk}'_E = 0)$ is at least $9/10$. Since $(1/2 - o(1))(9/10) > 1/3$ and $X^{n+o(1)} = (X^\eta)^{1+o(1)}$, the results follow.

(f) We have

$$\mathrm{Prob}(\mathrm{rk}(\ker A) \geq r \mid n \equiv r \ (\mathrm{mod}\ 2)) = \frac{\#\{A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt},\leq X} : \mathrm{rk}(\ker A) \geq r\}}{\# \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt},\leq X}}$$

$$= \frac{X^{n(n-r)/2+o(1)}}{X^{n(n-1)/2+o(1)}} \quad \text{(by Theorem 9.1.1)}$$

$$= (X^n)^{-(r-1)/2+o(1)}$$

$$= (X^\eta)^{-(r-1)/2+o(1)} \ (\text{since } n = \eta + o(\eta) = \eta(1 + o(1))),$$

18

as $H \to \infty$. On the other hand,

$$\mathrm{Prob}(\mathrm{rk}(\ker A) \geq r \mid n \not\equiv r \pmod 2) = \mathrm{Prob}(\mathrm{rk}(\ker A) \geq r + 1 \mid n \not\equiv r \pmod 2)$$
$$= (X^{\eta})^{-r/2 + o(1)},$$

by a similar calculation. Combining these yields the result. □

Remark 7.2.2. See [WS17] for related work applying the Ekedahl sieve to study cokernels of not-necessarily-alternating integral matrices.

Remark 7.2.3. The conclusion of Theorem 7.2.1(b) is likely robust. The analogous conclusion for symmetric matrices is proved in [Woo17] without requiring any kind of uniform distribution of matrix entries. For instance, it holds even if $X(H)$ is always 1.

Remark 7.2.4. The proof of (b) implicitly used that the $\mathbb{Z}$-points on the moduli space of matrices (isomorphic to $\mathbb{A}^{n^2}$) are equidistributed in the $\mathbb{Z}_p$-points. For $r \geq 2$, this affine space gets replaced by a subvariety V defined by the vanishing of certain minors, and it is not clear that $V(\mathbb{Z})$ is equidistributed in $V(\mathbb{Z}_p)$. In fact, heuristics inspired by the circle method suggest that this might be false, and numerical experiments also suggest this. In this case, perhaps the three conjectural answers to Question 5.3.1 are wrong for $r \geq 2$. In particular, perhaps the "canonical probability measure" from [BKLPR15, Section 2] on the set

$$\{A \in \mathrm{M}_n(\mathbb{Z}_p)_{\mathrm{alt}} : \mathrm{rk}_{\mathbb{Z}_p}(\ker A) = r\}$$

used to define $\mathscr{A}_r$ (the measure proportional to p-adic volume) should be replaced by the measure that reflects the density of integer points.

Next we will pass from Theorem 7.2.1, which concerns the random variable associated to one E, to Corollary 7.2.6, which concerns the aggregate behavior of the random variables associated to all $E \in \mathscr{E}_{\leq H}$, as $H \to \infty$. To do this we will apply the following standard result, a version of the law of large numbers in which the random variables do not have to be identically distributed.

Lemma 7.2.5 (Theorem 2.3.8 of [Dur10]). *Let $B_1, B_2, \ldots$ be a sequence of independent events. For $i \geq 1$, let p_i be the probability of B_i. If $\sum p_m$ diverges, then with probability 1,*

$$\#\{i \leq m : B_i \text{ occurs}\} = (1 + o(1)) \sum_{i=1}^{m} p_i$$

as $m \to \infty$.

Corollary 7.2.6. *If $X(H)$ grows sufficiently quickly relative to $\eta(H)$, then the following hold with probability 1.*

(a) *We have*
$$\mu(\{E : \mathrm{rk}'_E = 0\}) = \mu(\{E : \mathrm{rk}'_E = 1\}) = 1/2 \ \text{and}$$
$$\mu(\{E : \mathrm{rk}'_E \geq 2\}) = 0.$$

(b) *For each $r \in \{0, 1\}$ and $G \in \mathfrak{S}_p$,*
$$\mu(\{E : \mathrm{III}'_E[p^{\infty}] \simeq G\} \mid \mathrm{rk}'_E = r) = \mathrm{Prob}_{\mathscr{D}_{r,p}}(G).$$

(c) (1) *For each $G \in \mathfrak{S}$, we have*
$$\mu(\{E : \mathrm{III}'_E \simeq G\} \mid \mathrm{rk}'_E = 1) = \prod_p \mathrm{Prob}_{\mathscr{D}_{1,p}}(G[p^{\infty}]).$$

(2) *More generally, for each $\mathcal{G} \subseteq \mathfrak{S}$, we have*

$$\mu(\{E : \text{III}'_E \in \mathcal{G}\} \mid \text{rk}'_E = 1) = \sum_{G \in \mathcal{G}} \prod_p \text{Prob}_{\mathscr{D}_{1,p}}(G[p^\infty]).$$

(d) (1) *For each $G \in \mathfrak{S}$, we have $\mu(\{E : \text{III}'_E \simeq G\} \mid \text{rk}'_E = 0) = 0$.*
(2) *If $\mathcal{G}$ is the set of squares of cyclic groups, then*

$$\mu(\{E : \text{III}'_E \in \mathcal{G}\} \mid \text{rk}'_E = 0) = \prod_p \left(1 - \frac{1}{p^2} + \frac{1}{p^3} \right).$$

(e) *We have*

$$\underset{E \in \mathscr{E}_{\leq H}}{\text{Average}} \, \text{III}'_{0,E} = (X(H)^{\eta(H)})^{1+o(1)}$$

as $H \to \infty$, assuming that the function $f(H) := X(H)^{\eta(H)}$ satisfies $f(2H) \leq f(H)^{1+o(1)}$.

Proof. For $E \in \mathscr{E}$, let B_E be the event $\text{rk}'_E = r$, and let C_E be the event that $\text{rk}'_E = r$ and $\text{III}'_E[p^\infty] \simeq G$.
(a) Apply Lemma 7.2.5 to (B_E), and use Theorem 7.2.1(a).
(b) Apply Lemma 7.2.5 to (B_E) and (C_E), and use Theorem 7.2.1(a,b) to compute the denominator and numerator in the definition of μ.
(c) Again apply Lemma 7.2.5, and use Theorem 7.2.1(c).
(d) (1) Apply Lemma 7.2.5 to the event $\text{rk}'_E = 0$ and $\text{III}'_E \not\simeq G$, and use Theorem 7.2.1(d)(1).
(2) Apply Lemma 7.2.5 to the event $\text{rk}'_E = 0$ and $\text{III}'_E \in \mathcal{G}$, and use Theorem 7.2.1(d)(2).
(e) By Theorem 7.2.1(e)(1),

$$\underset{E \in \mathscr{E}_{\leq H}}{\text{Average}} \, \text{III}'_{0,E} \leq \max_{E \in \mathscr{E}_{\leq H}} \, \text{III}'_{0,E} \leq (X(H)^{\eta(H)})^{1+o(1)}$$

as $H \to \infty$. By Theorem 7.2.1(e)(2) and the law of large numbers, with probability 1, as $H \to \infty$, at least $1/4$ of the elliptic curves with height in $(H/2, H]$ satisfy $\text{III}'_{0,E} \geq (X(\text{ht } E)^{\eta(\text{ht } E)})^{1-o(1)}$. We have $(X(\text{ht } E)^{\eta(\text{ht } E)})^{1-o(1)} = (X(H)^{\eta(H)})^{1-o(1)}$ by the growth hypothesis on $f(H)$. Since a positive fraction of the elliptic curves in $\mathscr{E}_{\leq H}$ have height in $(H/2, H]$, this implies $\text{Average}_{E \in \mathscr{E}_{\leq H}} \text{III}'_{0,E} \geq (X(H)^{\eta(H)})^{1-o(1)}$. $\qquad \square$

7.3. **Consequences for coranks of random matrices.** Comparing Theorem 6.4.2(c) and Corollary 7.2.6(e) suggests choosing $X(H)$ and $\eta(H)$ so that

$$X(H)^{\eta(H)} = H^{1/12 + o(1)} \tag{7.3.1}$$

as $H \to \infty$.

Remark 7.3.2. Alternatively, matching the conditional upper bound of Theorem 6.4.2(b) with an upper bound for $\det A$ would have also suggested (7.3.1).

We now prove a theorem about the asymptotic aggregate behavior of the rk'_E.

Theorem 7.3.3. *If $\eta(H)$ grows sufficiently slowly relative to H, and $X(H)^{\eta(H)} = H^{1/12 + o(1)}$, then the following hold with probability 1:*
(a) *All but finitely many $E \in \mathscr{E}$ satisfy $\text{rk}'_E \leq 21$.*
(b) *For $1 \leq r \leq 20$, we have $\#\{E \in \mathscr{E}_{\leq H} : \text{rk}'_E \geq r\} = H^{(21-r)/24 + o(1)}$.*
(c) *We have $\#\{E \in \mathscr{E}_{\leq H} : \text{rk}'_E \geq 21\} \leq H^{o(1)}$.*

20

Proof. Fix $r \geq 1$. For $E \in \mathscr{E}$, let $p_{E,r} := \mathrm{Prob}(\mathrm{rk}'_E \geq r)$. By Theorem 7.2.1(f), if E is of height H, then
$$p_{E,r} = (X(H)^{\eta(H)})^{-(r-1)/2+o(1)} = H^{-(r-1)/24+o(1)}.$$
It follows that
$$\sum_{E \in \mathscr{E}_{\leq H}} p_{E,r} = \sum_{E \in \mathscr{E}_{\leq H}} (\mathrm{ht}\,E)^{-(r-1)/24+o(1)}$$
$$= \begin{cases} H^{(21-r)/24+o(1)}, & \text{if } 1 \leq r \leq 21; \\ O(1), & \text{if } r > 21, \end{cases}$$

by summing over dyadic intervals, using the estimate $\#\mathscr{E}_{\leq H} = (\kappa + o(1))H^{20/24}$ from (5.1.1).

If $\sum_{E \in \mathscr{E}} p_{E,r}$ converges, as happens for $r > 21$ and possibly also for $r = 21$, then the Borel–Cantelli lemma implies that $\{E \in \mathscr{E} : \mathrm{rk}'_E \geq r\}$ is finite. If $\sum_{E \in \mathscr{E}} p_{E,r}$ diverges, as happens for $1 \leq r \leq 20$ and possibly also for $r = 21$, then Lemma 7.2.5 yields
$$\#\{E \in \mathscr{E}_{\leq H} : \mathrm{rk}'_E \geq r\} = (1 + o(1)) \sum_{E \in \mathscr{E}_{\leq H}} p_{E,r} = H^{(21-r)/24+o(1)}. \qquad \square$$

Remark 7.3.4. The conclusion that rk'_E is uniformly bounded with probability 1 is robust. For example, if the assumption $X(H)^{\eta(H)} = H^{1/12+o(1)}$ in (7.3.1) is replaced by $X(H)^{\eta(H)} = H^{c+o(1)}$ for a different positive constant c, then the same conclusion follows, but the bound beyond which there are only finitely many E might no longer be 21. Another example: taking our matrix coefficients in a sphere instead of a box (as we actually do in Section 9) does not change Theorem 9.1.1, so it does not change Theorem 7.3.3 either.

Remark 7.3.5. Although it would have been nice to have specifications for $X(H)$ and $\eta(H)$ individually, the specification of $X(H)^{\eta(H)}$ alone sufficed for Theorem 7.3.3.

8. Predictions for elliptic curves

The results of Section 7 are unconditional theorems about random matrices. We now conjecture that some statements about the statistics of $(\mathrm{rk}'_E, \text{Ш}'_E)$ as E varies are true also for the *actual* $(\mathrm{rk}\,E(\mathbb{Q}), \text{Ш}(E))$.

8.1. **Theoretical evidence.** Several consequences of this heuristic are widely believed conjectures for elliptic curves.

(i) Corollary 7.2.6(a) suggests the "minimalist conjecture" that ranks of elliptic curves are 0 half the time and 1 half the time asymptotically, as has been conjectured by others for quadratic twist families, including Goldfeld [Gol79, Conjecture B] and Katz and Sarnak [KS99a, KS99b].

(ii) Corollary 7.2.6(b) predicts the distribution of $\text{Ш}(E)[p^\infty]$ for elliptic curves E of rank r in the cases $r = 0, 1$. This distribution agrees with the three distributions in Section 5.3. More generally, for any finite set S of primes, our model predicts the joint distribution of $\text{Ш}(E)[p^\infty]$ for $p \in S$ as E varies among rank r curves in the cases $r = 0, 1$, and these predictions agree with the conjectures in [Del01, Del07, DJ14] and [BKLPR15, Section 5.6].

(iii) Corollary 7.2.6(c) predicts that $\text{Ш}(E)$ for E of rank 1 is distributed according to the conjectures in [Del01, Del07, DJ14].

(iv) Given $G \in \mathfrak{S}$, Corollary 7.2.6(d) predicts that the density of rank 0 curves E with $\text{Ш}(E) \simeq G$ is zero, in agreement with the conjectures discussed in Remark 5.3.4, and that the density of rank 0 curves E with $\text{Ш}(E)$ the square of a cyclic group is the density conjectured by Delaunay [Del01, Example E].

8.2. Predictions for ranks.

Our heuristic predicts also that the three conclusions of Theorem 7.3.3 hold if rk'_E is replaced by $\text{rk}\, E(\mathbb{Q})$. Specifically, it predicts

(a) All but finitely many $E \in \mathscr{E}$ satisfy $\text{rk}\, E(\mathbb{Q}) \leq 21$.

(b) For $1 \leq r \leq 20$, we have $\#\{E \in \mathscr{E}_H \mid \text{rk}\, E(\mathbb{Q}) \geq r\} \overset{?}{=} H^{(21-r)/24+o(1)}$.

(c) We have $\#\{E \in \mathscr{E}_{\leq H} : \text{rk}\, E(\mathbb{Q}) \geq 21\} \overset{?}{\leq} H^{o(1)}$.

In particular, (a) would imply that ranks of elliptic curves over $\mathbb{Q}$ are bounded.

The prediction (b) for $r = 1$ was discussed in (i) in Section 8.1. The prediction of (b) for $r = 2$ is consistent with all of the previous conjectures in Section 3.3. See Section 3.4 for a comparison of the prediction of (b) for $r = 3$ to the other conjectures for this asymptotic.

There are many other predictions made by the model of Section 7, though in some cases we are prevented from giving them explicitly because we do not know the corresponding fact about counting alternating matrices. We mention some of these in Section 11.

8.3. Other families of elliptic curves.

Instead of taking all elliptic curves over $\mathbb{Q}$, one could restrict to other families of elliptic curves, such as a family of quadratic twists or a family with prescribed torsion subgroup, and prove analogues of Theorem 7.3.3. The predictions given by such an analogue are summarized in the following two tables; under "rank bound" is an integer such that our model predicts that the family contains only finitely many elliptic curves of strictly higher rank.

First, we consider a family of twists. In some cases, these predictions are stronger than the predictions coming from variants of Granville's heuristic [Wat$^+$14, Section 11].

	# curves of height $\leq H$	rank bound	Granville
quadratic twists of a fixed E_0	$H^{1/6} = H^{4/24}$	5	7
quartic twists of $y^2 = x^3 - x$	$H^{1/3} = H^{8/24}$	9	11
sextic twists of $y^2 = x^3 - 1$	$H^{1/2} = H^{12/24}$	13	13
all elliptic curves	$H^{5/6} = H^{20/24}$	21	21

Next, we consider a family with prescribed torsion. Harron and Snowden [HS17] prove that for each finite abelian group T that arises, $\#\{E \in \mathscr{E}_{\leq H} : E(\mathbb{Q})_{\text{tors}} \simeq T\} \asymp H^{1/d}$ for some $d \in \mathbb{Q}_{>0}$ depending on T. For such a family, our heuristic suggests that the expression $b_T := \limsup\{\text{rk}\, E(\mathbb{Q}) : E(\mathbb{Q})_{\text{tors}} \simeq T\}$ is bounded above by $1 + \lfloor 24/d \rfloor$. On the other hand, explicit families provide lower bounds on b_T [Duj]. These upper and lower bounds are given in the last two columns of the following table. Remarkably, for each T, the bounds are close and the conjectured upper bound is at least as large as the proven lower bound. Moreover, whenever the rank bound meets the known lower bound, our heuristic predicts that there should be $H^{o(1)}$ curves meeting this bound; and again remarkably, the known infinite families are over elliptic curves with positive rank, consistent with our prediction!

22

torsion subgroup	# curves of height $\leq H$	rank bound	known lower bound
—	$H^{5/6}$	21	19
$\mathbb{Z}/2\mathbb{Z}$	$H^{1/2}$	13	11
$\mathbb{Z}/3\mathbb{Z}$	$H^{1/3}$	9	7
$\mathbb{Z}/4\mathbb{Z}$	$H^{1/4}$	7	6
$\mathbb{Z}/5\mathbb{Z}$	$H^{1/6}$	5	4
$\mathbb{Z}/6\mathbb{Z}$	$H^{1/6}$	5	5
$\mathbb{Z}/7\mathbb{Z}$	$H^{1/12}$	3	2
$\mathbb{Z}/8\mathbb{Z}$	$H^{1/12}$	3	3
$\mathbb{Z}/9\mathbb{Z}$	$H^{1/18}$	2	1
$\mathbb{Z}/10\mathbb{Z}$	$H^{1/18}$	2	1
$\mathbb{Z}/12\mathbb{Z}$	$H^{1/24}$	2	1
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$	$H^{1/3}$	9	8
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$	$H^{1/6}$	5	5
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$	$H^{1/12}$	3	3
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z}$	$H^{1/24}$	2	1

9. Counting alternating matrices of prescribed rank

In this section we prove Theorem 9.1.1, which was used in the proof of Theorem 7.2.1(f).

9.1. Statement and overview of proof.

Theorem 9.1.1. *If $1 \leq r \leq n$ and $n - r$ is even, then*

$$\#\{A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}, \leq X} : \mathrm{rk}(\ker A) \geq r\} \asymp_n X^{n(n-r)/2}.$$

In fact, we will prove the same asymptotic for the count with $\geq r$ replaced by $= r$; then Theorem 9.1.1 follows by summing. Also, the ℓ^∞-norm of a matrix $A = (a_{ij})$ is bounded above and below by the ℓ^2-norm times constants depending on n, so we may instead use the ℓ^2-norm, which is defined by $|A|^2 = \sum_{i,j} a_{ij}^2$. Finally, we may use $\mathrm{rk}(\ker A) = n - \mathrm{rank}\, A$, and rename r as $n - r$. This leads us to define

$$N_{n,r}(T) := \#\{A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}} : \mathrm{rank}\, A = r \text{ and } |A| < T\}.$$

Now the result to be proved is as follows.

Theorem 9.1.2. *If $0 \leq r \leq n - 1$ and r is even, then $N_{n,r}(T) \asymp_n T^{nr/2}$.*

Theorem 9.1.2 is the analogue for alternating matrices of the Eskin–Katznelson theorem counting integral *symmetric* matrices of specified rank [EK95, Theorem 1.2]. Our proof of Theorem 9.1.2 follows the Eskin–Katznelson proof closely; indeed, the differences are almost entirely numerical (though we have incorporated a few simplifications, notably in the proofs of Lemma 9.3.5 and Theorem 9.5.1). The strategy is to count the matrices (the rank r matrices $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}}$ with $|A| < T$) by grouping them according to which primitive rank r sublattice $\Lambda \subseteq \mathbb{Z}^n$ contains the rows of A.

To obtain an upper bound on $N_{n,r}(T)$, first, we bound the determinant of the lattices Λ that arise in the previous sentence (Corollary 9.4.4). Second, a theorem of Schmidt (Theorem 9.2.5) bounds the number of Λ of bounded determinant. Third, the matrices A associated to *one* such Λ correspond to certain points in the lattice $\mathcal{A}(\Lambda)$ of matrices whose rows are contained

in Λ, so they can be counted by another result of Schmidt on lattice points in a growing ball (Lemma 9.2.4).

To obtain a matching lower bound on $N_{n,r}(T)$, it will turn out that it suffices to count lattices having a basis consisting of "almost orthogonal" vectors of roughly comparable length (Theorem 9.5.1).

9.2. **Lattices.** Fix $n \geq 0$, and let $(\ ,\ )$ and $|\ |$ denote the standard inner product and ℓ^2-norm on $\mathbb{R}^n$. By a lattice in $\mathbb{R}^n$, we mean a discrete subgroup $\Lambda \subset \mathbb{R}^n$; its rank r might be less than n. By convention, each $\mathbb{Z}$-basis $\{\ell_1, \ldots, \ell_r\}$ of Λ is ordered so that $|\ell_1| \leq \cdots \leq |\ell_r|$. A lattice $\Lambda \subset \mathbb{Z}^n$ is primitive if it is not properly contained in any other sublattice of $\mathbb{Z}^n$ of the same rank. The determinant $d(\Lambda) \in \mathbb{R}_{>0}$ of Λ is the r-dimensional volume (with respect to the metric induced by $|\ |$) of the parallelepiped spanned by any $\mathbb{Z}$-basis $\{\ell_1, \ldots, \ell_r\}$ of Λ; then $d(\Lambda)^2 = \det(\ell_i, \ell_j)_{1 \leq i,j \leq r}$. Among all bases $\{\ell_1, \ldots, \ell_r\}$ for Λ, any one that minimizes the product $|\ell_1| \cdots |\ell_r|$ is called a reduced basis. (This is equivalent to the usual definition of Minkowski reduced basis.)

Theorem 9.2.1 (Minkowski). *If $\{\ell_1, \ldots, \ell_r\}$ is a reduced basis for Λ, then $d(\Lambda) \asymp_r |\ell_1| \cdots |\ell_r|$.*

Theorem 9.2.1 can be interpreted as saying that a reduced basis is "almost orthogonal". The following lemma is essentially a reformulation of Minkowski's theorem [EK95, Lemma 2.1, comment after Lemma 2.2].

Lemma 9.2.2. *Fix r and a positive constant C. Let Λ be a lattice with basis $\{u_1, \ldots, u_r\}$ satisfying $d(\Lambda) \geq C|u_1| \cdots |u_r|$. Then for any $a_1, \ldots, a_r \in \mathbb{R}$,*

$$\left| \sum_{j=1}^{r} a_j u_j \right|^2 \gg_{r,C} \sum_{j=1}^{r} a_j^2 |u_j|^2.$$

The following lemma shows that different choices of reduced bases, or even bases within a constant factor of being reduced, have very similar lengths.

Lemma 9.2.3 (Lemma 2.2 of [EK95]). *Fix r and a positive constant C. Let Λ be a rank r lattice with reduced basis $\{\ell_1, \ldots, \ell_r\}$. If $\{u_1, \ldots, u_r\}$ is another basis of Λ, and $d(\Lambda) \geq C|u_1| \cdots |u_r|$, then $|u_j| \asymp_{r,C} |\ell_j|$ for all j.*

For $T \in \mathbb{R}_{>0}$, define the ball $B(T) := \{x \in \mathbb{R}^n : |x| < T\}$. For any lattice $\Lambda \subset \mathbb{R}^n$, let

$$N(T, \Lambda) := \#(\Lambda \cap B(T)).$$

Let $\mathcal{V}_r$ denote the volume of the r-dimensional unit ball.

Lemma 9.2.4 ([Sch68, Lemma 2]). *Let Λ be a rank r lattice with reduced basis $\{\ell_1, \ldots, \ell_r\}$. Then*

$$N(T, \Lambda) = \frac{\mathcal{V}_r T^r}{d(\Lambda)} + O_r \left(\sum_{j=0}^{r-1} \frac{T^j}{|\ell_1| \cdots |\ell_j|} \right).$$

Let $P_{n,r}(t)$ denote the number of primitive rank r sublattices of $\mathbb{Z}^n$ of determinant at most t. The following is a crude version of a more precise theorem of Schmidt.

Theorem 9.2.5 ([Sch68, Theorem 1]). *If $1 \leq r \leq n-1$, then $P_{n,r}(t) \asymp_n t^n$.*

9.3. Lattices of alternating matrices. From now on, Λ is a primitive rank r lattice in $\mathbb{Z}^n$, and $\{\ell_1, \ldots, \ell_r\}$ is a reduced basis of Λ. Define

$$\mathcal{A}(\Lambda) := \{A \in \mathrm{M}_n(\mathbb{R})_{\mathrm{alt}} : \text{every row of } A \text{ is in } \Lambda\}.$$

View $\mathcal{A}(\Lambda)$ as a lattice in the space $\mathrm{M}_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$. If $A \in \mathcal{A}(\Lambda)$, then rank $A \leq r$; we write

$$N(T, \mathcal{A}(\Lambda)) = N_1(T, \mathcal{A}(\Lambda)) + N_2(T, \mathcal{A}(\Lambda)),$$

where N_1 counts the matrices of rank exactly r and N_2 counts those of rank $< r$. If $A \in \mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}}$ and rank $A = r$, then there exists a unique primitive rank r lattice $\Lambda \subset \mathbb{Z}^n$ such that $A \in \mathcal{A}(\Lambda)$; namely, $\Lambda = (\text{row space of } A) \cap \mathbb{Z}^n$. Thus, as in [EK95, Proposition 1.1],

$$N_{n,r}(T) = \sideset{}{'}\sum_{\mathrm{rk}\,\Lambda = r} N_1(T, \mathcal{A}(\Lambda)), \tag{9.3.1}$$

where the prime indicates that Λ ranges over *primitive* rank r lattices in $\mathbb{Z}^n$. We will use this to estimate $N_{n,r}(T)$.

To apply Lemma 9.2.4 to count points in $\mathcal{A}(\Lambda)$, we need good estimates on the size of a reduced basis of $\mathcal{A}(\Lambda)$. Identify $\mathbb{R}^n \otimes \mathbb{R}^n$ with $\mathrm{M}_n(\mathbb{R})$ by mapping $u \otimes v$ to uv^T. For $1 \leq i < j \leq r$, define

$$R_{ij} := \ell_i \otimes \ell_j - \ell_j \otimes \ell_i \in \mathrm{M}_n(\mathbb{R}).$$

Lemma 9.3.2 (Analogue of Lemma 3.2 of [EK95]).
(a) *If $i < j$ and $s < t$, then $(R_{ij}, R_{st}) = 2(\ell_i, \ell_s)(\ell_j, \ell_t) - 2(\ell_i, \ell_t)(\ell_j, \ell_s)$.*
(b) *For $1 \leq i < j \leq r$, we have $|R_{ij}| \asymp_r |\ell_i||\ell_j|$.*

Proof.
(a) Distribute and use the identity $(u \otimes v, u' \otimes v') = (u, u')(v, v')$ four times.
(b) Taking $(i, j) = (s, t)$ in (a) yields

$$|R_{ij}|^2 = 2|\ell_i|^2|\ell_j|^2 - 2(\ell_i, \ell_j)^2 = 2|\ell_i|^2|\ell_j|^2 \sin^2 \theta,$$

where θ is the angle between ℓ_i and ℓ_j. By Theorem 9.2.1, we have $\sin \theta \asymp_r 1$. $\qquad \square$

Proposition 9.3.3 (Analogue of Proposition 3.3 of [EK95]). *The set $\{R_{ij} : 1 \leq i < j \leq r\}$ is a basis of $\mathcal{A}(\Lambda)$.*

Proof. The proof is analogous to that of [EK95, Proposition 3.3]. $\qquad \square$

For $g \in \mathrm{GL}_n(\mathbb{R})$, let $F_g \colon \mathrm{M}_n(\mathbb{R})_{\mathrm{alt}} \to \mathrm{M}_n(\mathbb{R})_{\mathrm{alt}}$ be the linear map $A \mapsto gAg^T$.

Proposition 9.3.4 (Analogue of Proposition 3.4 of [EK95]). *We have $\det F_g = (\det g)^{n-1}$ for all $g \in \mathrm{GL}_n(\mathbb{R})$.*

Proof. Since $g \mapsto \det F_g$ is an algebraic homomorphism $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$, the identity $\det F_g = (\det g)^\alpha$ holds for some $\alpha \in \mathbb{Z}$. Evaluating at $g = tI$ for any $t \in \mathbb{R}^\times$ yields $(t^2)^{n(n-1)/2} = (t^n)^\alpha$, so $\alpha = n - 1$. $\qquad \square$

Lemma 9.3.5 (Analogue of Lemma 3.5 of [EK95]). *We have $d(\mathcal{A}(\Lambda)) = 2^{r(r-1)/4} d(\Lambda)^{r-1}$.*

Proof. View $\mathcal{A}(\Lambda)$ as the lattice generated by the R_{ij} (Proposition 9.3.3). By Lemma 9.3.2(a), the square of the desired identity is a polynomial identity in the (ℓ_i, ℓ_j). Thus we may rotate $\ell_1, \ldots, \ell_r$ to vectors in $\mathbb{R}^r$ with the same inner products; now $n = r$.

The basis of $\mathcal{A}(\mathbb{Z}^r)$ provided by Proposition 9.3.3 consists of $r(r-1)/2$ orthogonal vectors of length $\sqrt{2}$, so for $\Lambda = \mathbb{Z}^r$, both sides of the identity equal $2^{r(r-1)/4}$. Now let Λ be any other rank r lattice in $\mathbb{Z}^r$. Let $g \in \mathrm{GL}_r(\mathbb{R})$ be the linear map taking $\mathbb{Z}^r$ to Λ. Then F_g takes $\mathcal{A}(\mathbb{Z}^r)$ to $\mathcal{A}(\Lambda)$. Replacing $\mathbb{Z}^r$ by Λ scales the two sides of the identity by $|\det F_g|$ and $|\det g|^{r-1}$, respectively; these factors are equal by Proposition 9.3.4, so the identity is preserved. $\square$

Combining Lemma 9.3.5, Theorem 9.2.1, and Lemma 9.3.2(b) yields

$$d(\mathcal{A}(\Lambda)) = 2^{r(r-1)/4} d(\Lambda)^{r-1} \gg_r \prod_i |\ell_i|^{r-1} \gg_r \prod_{i<j} |R_{ij}|. \tag{9.3.6}$$

9.4. Upper bound on $N_{n,r}(T)$.

Lemma 9.4.1. *The map*

$$\mathrm{M}_r(\mathbb{Z})_{\mathrm{alt}} \to \mathcal{A}(\Lambda)$$
$$(a_{ij}) \mapsto \sum_{i<j} a_{ij} R_{ij}$$

is a bijection that preserves ranks of matrices.

Proof. It is a bijection by Proposition 9.3.3. If $\ell_1, \ldots, \ell_r$ are the first r standard basis vectors, then the bijection simply extends a matrix in $\mathrm{M}_r(\mathbb{Z})_{\mathrm{alt}}$ by zeros to a matrix in $\mathrm{M}_n(\mathbb{Z})_{\mathrm{alt}}$; this preserves rank. The general case follows: if for some $g \in \mathrm{GL}_n(\mathbb{R})$ we replace $\ell_1, \ldots, \ell_r$ by $g\ell_1, \ldots, g\ell_r$, then $A := \sum_{i<j} a_{ij} R_{ij}$ is replaced by gAg^T, which has the same rank as A. $\square$

Lemma 9.4.2. *Let $B = (a_{ij}) \in \mathrm{M}_r(\mathbb{R})_{\mathrm{alt}}$. If there exist $i \le j$ with $i + j \le r + 1$ such that for every pair $s < t$ with $s \ge i$ and $t \ge j$, we have $a_{st} = 0$, then $\mathrm{rank}\, B \le r - 1$.*

Proof. The ith through the nth row are contained in the initial copy of $\mathbb{R}^{j-1}$ in $\mathbb{R}^r$, so they together with the first $i-1$ rows span a space of dimension at most $(i-1)+(j-1) \le r-1$. Thus $\mathrm{rank}\, B \le r - 1$. $\square$

Recall that $N_1(T, \mathcal{A}(\Lambda))$ counts the matrices in $\mathcal{A}(\Lambda)$ of rank exactly $r = \mathrm{rk}\, \Lambda$.

Lemma 9.4.3 (Analogue of Lemma 4.1 of [EK95]). *If $N_1(T, \mathcal{A}(\Lambda)) > 0$, then $|\ell_i||\ell_j| \ll_r T$ for all pairs (i,j) such that $i + j \le r + 1$.*

Proof. Suppose that $N_1(T, \mathcal{A}(\Lambda)) > 0$. Thus there exists $A \in \mathcal{A}(\Lambda)$ of rank r with $|A| \le T$. By Lemma 9.4.1, we may write $A = \sum_{i<j} a_{ij} R_{ij}$ for some $B = (a_{ij}) \in \mathrm{M}_r(\mathbb{Z})_{\mathrm{alt}}$ also of rank r. Lemma 9.4.2 yields a pair $s < t$ with $s \ge i$ and $t \ge j$ such that $a_{st} \ne 0$. By (9.3.6) and Lemma 9.2.2, $|A|^2 \gg_r a_{st}^2 |R_{st}|^2 \ge |R_{st}|^2$. Thus $|\ell_i||\ell_j| \le |\ell_s||\ell_t| \asymp_r |R_{st}| \ll_r |A| \le T$ (the second step is Lemma 9.3.2(b)). $\square$

Corollary 9.4.4 (Analogue of Corollary 4.2 of [EK95]). *If $N_1(T, \mathcal{A}(\Lambda)) > 0$, then $d(\Lambda) \ll_r T^{r/2}$.*

Proof. By Theorem 9.2.1 (at the left) and Lemma 9.4.3 (at the right),

$$d(\Lambda)^2 \asymp_r (|\ell_1| \cdots |\ell_r|)^2 = \prod_{i+j=r+1} |\ell_i||\ell_j| \ll_r T^r. \qquad \square$$

Theorem 9.4.5 (Analogue of Theorem 4.1 of [EK95]). *If $0 \le r \le n-1$ and r is even, then*

$$N_{n,r}(T) \ll_n T^{nr/2}.$$

Proof. Since $N_{n,0}(T) = 1$, we may assume that $r \geq 2$. Let E be the set of primitive rank r lattices $\Lambda \subset \mathbb{Z}^n$ for which $N_1(T, \mathcal{A}(\Lambda)) > 0$. Let c_r be the implied constant in Corollary 9.4.4, and let E^* be the set of primitive rank r sublattices of $\mathbb{Z}^n$ for which $d(\Lambda) \leq c_r T^{r/2}$. Thus $E \subseteq E^*$. By (9.3.1),

$$N_{n,r}(T) = \sum_{\Lambda \in E} N_1(T, \mathcal{A}(\Lambda)) \leq \sum_{\Lambda \in E} N(T, \mathcal{A}(\Lambda)).$$

By Lemmas 9.2.4 and 9.3.5,

$$N(T, \mathcal{A}(\Lambda)) = \frac{\mathcal{V}_{r(r-1)/2} T^{r(r-1)/2}}{2^{r(r-1)/4} d(\Lambda)^{r-1}} + O_r\left(\sum_{k=0}^{r(r-1)/2-1} \frac{T^k}{L_1 L_2 \cdots L_k}\right), \tag{9.4.6}$$

where the L_j are a reduced basis for $\mathcal{A}(\Lambda)$. By (9.3.6) and Lemma 9.2.3, we can order the tuples (i,j) with $1 \leq i < j \leq r$ as $(i_1, j_1), \ldots$ so that $L_k \asymp_r R_{i_k, j_k} \asymp_r |\ell_{i_k}||\ell_{j_k}|$, by Lemma 9.3.2(b), where the l_i are a reduced basis for Λ. Thus

$$\sum_{k=0}^{r(r-1)/2-1} \frac{T^k}{L_1 L_2 \cdots L_k} \ll_r \prod_{\substack{1 \leq i < j \leq r \\ (i,j) \neq (r-1,r)}} \max\left(1, \frac{T}{|l_i||l_j|}\right). \tag{9.4.7}$$

Let $\mathcal{E}_1$ be the sum over $\Lambda \in E$ of the right side of (9.4.7). If $4 \leq r \leq n-1$, then the proof of [EK95, Proposition 7.1] yields $\mathcal{E}_1 \ll_n T^{nr/2}$ (in [EK95, Proposition 7.1], the analogous sum is instead over Λ for which there is a *symmetric* rank r matrix of size $< T$ with rows in Λ, but the only property used of the Λ there is the conclusion of Lemma 9.4.3; also each summand in our $\mathcal{E}_1$ is at most the corresponding summand in [EK95, Proposition 7.1]). If $r = 2$, then the right side of (9.4.7) is 1, so $\mathcal{E}_1 = \#E \leq \#E^* \ll_n T^n$, by Schmidt's theorem (Theorem 9.2.5). Thus $\mathcal{E}_1 \ll_n T^{nr/2}$ in all cases.

It follows that

$$N_{n,r}(T) \ll_n T^{r(r-1)/2} \sum_{\Lambda \in E} d(\Lambda)^{-(r-1)} + T^{nr/2}. \tag{9.4.8}$$

We have

$$\sum_{\Lambda \in E} d(\Lambda)^{-(r-1)} \leq \sum_{\Lambda \in E^*} d(\Lambda)^{-(r-1)} = \sum_{d(\Lambda) \leq c_r T^{r/2}} d(\Lambda)^{-(r-1)} \ll_r T^{(n-r+1)r/2} \tag{9.4.9}$$

by summing over dyadic intervals since for each $t \in \mathbb{R}$,

$$\sum_{d(\Lambda) \in (t/2, t]} d(\Lambda)^{-(r-1)} \ll_r P_{n,r}(t) t^{-(r-1)} \ll_r t^{n-r+1},$$

by Theorem 9.2.5. Substituting (9.4.9) into (9.4.8) yields $N_{n,r}(T) \ll_n T^{nr/2}$. $\qquad \square$

9.5. **Lower bound on $N_{n,r}(T)$.**

Theorem 9.5.1 (Analogue of Theorem 4.2 of [EK95]). *If $0 \leq r \leq n-1$ and r is even, then*

$$N_{n,r}(T) \gg_n T^{nr/2}.$$

Proof. Since $N_{n,0}(T) = 1$, we may assume that $r \geq 2$. Let $c = c(n,r)$ be as in [EK95, Proposition 2.6]. Suppose that Λ is c-regular, i.e., has a reduced basis with $|\ell_1| \geq cd(\Lambda)^{1/r}$. By Theorem 9.2.1, $|\ell_i| \asymp_n d(\Lambda)^{1/r}$ for all i. By Lemma 9.3.2(b), $|R_{ij}| \asymp_n d(\Lambda)^{2/r}$ for all $i < j$. By Lemma 9.4.1, the matrix $A := \sum_{s=1}^{r/2} R_{2s-1,2s} \in \mathcal{A}(\Lambda)$ is of rank r, and $|A| \asymp_n d(\Lambda)^{2/r}$. Thus we can fix $\epsilon = \epsilon_{n,r} > 0$ so that $d(\Lambda) \leq \epsilon T^{r/2}$ implies $|A| \leq T$ and hence $N_1(T, \mathcal{A}(\Lambda)) \geq 1$. By [EK95, Proposition 2.6], the number of c-regular Λ with $d(\Lambda) \leq \epsilon T^{r/2}$ is $\asymp_n T^{nr/2}$, and each contributes at least 1 to $N_{n,r}(T)$. $\square$

Theorems 9.4.5 and 9.5.1 imply Theorem 9.1.2, and hence Theorem 9.1.1.

10. Computational evidence

10.1. **New evidence.** There is a long history of computational investigations on the ranks of elliptic curves: see [BMc90, Cre97, BMSW07] and the references therein. In this section, we provide some further experimental evidence for our conjecture on ranks, using the computer algebra system Magma [Magma]; specifically, we use its algorithms developed by Steve Donnelly and Mark Watkins for computing the 2-Selmer group, and by Tim Dokchitser and Mark Watkins for computing the analytic rank.

We sample random elliptic curves at various heights. For a height H, we compute uniformly random integers A, B with

$$A \in [-\sqrt[3]{H/4}, \sqrt[3]{H/4}], \quad B \in [-\sqrt{H/27}, \sqrt{H/27}],$$

and keep the pair (A, B) if $y^2 = x^3 + Ax + B$ defines an elliptic curve in $\mathcal{E}_{\leq H} - \mathcal{E}_{\leq H/2}$. By this procedure, we generate a uniformly random elliptic curve $E \in \mathcal{E}_{\leq H} - \mathcal{E}_{\leq H/2}$.

Next, we compute the 2-Selmer group $\mathrm{Sel}_2 E$ assuming the Riemann hypothesis for Dedekind zeta functions to speed up the computation of the relevant class groups and unit groups. We also compute $E(\mathbb{Q})_{\mathrm{tors}}$. Let $r := \mathrm{rk}_2 \mathrm{Sel}_2 E - \mathrm{rk}_2 E(\mathbb{Q})_{\mathrm{tors}}$. If $r \leq 1$ (and $\mathrm{III}(E)$ is finite), then $\mathrm{III}(E)[2] = \{0\}$ and $r = \mathrm{rk}\, E(\mathbb{Q})$. Otherwise, we attempt to compute the analytic rank $\mathrm{ord}_{s=1} L(E,s)$, which equals $\mathrm{rk}\, E(\mathbb{Q})$ (assuming the Birch and Swinnerton-Dyer conjecture). Computing the analytic rank is the most computationally intensive step.

The table below displays our results. For each H shown, N denotes the number of elliptic curves generated as above, with heights in $(H/2, H]$. The entries in the next columns show what percentage of these N have rank 0, have rank 1, etc.; we expect that each of these entries deviates from the true percentage (what we would see if we averaged over all elliptic curves with heights in $(H/2, H]$) by at most $O(1/\sqrt{N})$, by the central limit theorem, so we list $1/\sqrt{N}$ in the final column.

H	N	0	1	≥ 2	≥ 3	$1/\sqrt{N}$
10^{10}	279090	32.6%	47.3%	17.3%	2.7%	0.2%
10^{11}	66006	33.4	47.2	16.9	2.5	0.4
10^{12}	29844	34.0	47.2	16.4	2.5	0.6
10^{13}	20299	34.8	47.0	15.4	2.6	0.7
10^{14}	17836	35.0	47.4	15.3	2.2	0.7
10^{15}	5028	36.3	46.4	15.1	2.2	1.4

The proportion with rank ≥ 2 is slowly but steadily decreasing; this seems consistent with our model's predicted proportion of $H^{-1/24+o(1)}$. The data for rank ≥ 3 also seems consistent

28

with our prediction. The raw values predicted by our model using the functions in the last row below are as follows:

H	0	1	≥ 2	≥ 3
10^{10}	30.8%	42.7%	19.2%	7.3%
10^{11}	32.6	43.9	17.4	6.0
10^{12}	34.2	45.0	15.8	5.0
10^{13}	35.6	45.9	14.4	4.1
10^{14}	36.9	46.6	13.0	3.4
10^{15}	38.1	47.2	11.9	2.8
H	$\frac{1}{2}(1 - H^{-1/24})$	$\frac{1}{2}(1 - H^{-1/12})$	$\frac{1}{2}H^{-1/24}$	$\frac{1}{2}H^{-1/12}$

But these values should not be read too closely, since we are ignoring an implicit factor $H^{o(1)}$.

Further computations on the distribution of Selmer groups and ranks of elliptic curves grouped by order of magnitude of height have been carried out recently by Balakrishnan, Ho, Kaplan, Spicer, Stein, and Weigandt [BHK$^+$16]. Our more modest statistical sample is in good agreement with theirs.

10.2. Biases in the counts of higher rank curves depending on the reduction modulo p.

Fix a prime p and an elliptic curve E over $\mathbb{Q}$ with good reduction at p. Define $a_p \in \mathbb{Z}$ by $\#E(\mathbb{F}_p) = p + 1 - a_p$. Inspired by random matrix theory, Conrey, Keating, Rubinstein, and Snaith [CKRS02, Conjecture 2] conjectured that among even quadratic twists E_d, the ratio of the number of (analytic) rank ≥ 2 twists with $\left(\frac{d}{p}\right) = 1$ to the number of (analytic) rank ≥ 2 twists with $\left(\frac{d}{p}\right) = -1$ tends to $\sqrt{\frac{p+1-a_p}{p+1+a_p}}$.

We can give a different argument for this conjecture, based on the heuristic that the "probability" that $\sqrt{\text{Ш}_0} = 0$ should be inversely proportional to the width of the range in which the integer $\sqrt{\text{Ш}_0}$ lies. Specifically, when we solve for $\sqrt{\text{Ш}_0}$ in the Birch and Swinnerton-Dyer conjecture, the only systematic difference depending on $\left(\frac{d}{p}\right)$ we expect is in the local L-factor $L_p(E_d, s)$, which, for $\left(\frac{d}{p}\right) = \pm 1$, is $L_p^{\pm}(s) := (1 \mp a_p p^{-s} + p^{1-2s})^{-1}$. Since $L_p^+(1)/L_p^-(1) = \frac{p+1+a_p}{p+1-a_p}$, the probabilities that $\sqrt{\text{Ш}_0} = 0$ should be in the ratio $\sqrt{\frac{p+1-a_p}{p+1+a_p}}$. The ratio of the counts of rank ≥ 2 curves should equal this ratio of probabilities.

The fact that such a rank count bias has been observed experimentally [CKRS02] is further evidence that the methodology in Sections 6–8 of basing conclusions on the distribution of $\#\text{Ш}_0$ is reasonable.

11. FURTHER QUESTIONS

The model in Section 7.1 will yield more predictions about ranks and Shafarevich–Tate groups of elliptic curves if we can prove the corresponding statements about the functions $(\text{rk}'_E, \text{Ш}'_E)_{E \in \mathscr{E}}$ of random alternating integer matrices.

11.1. The density of rank 0 elliptic curves whose Shafarevich–Tate group belongs to a specified class.

For any subset $\mathcal{G} \subseteq \mathfrak{S}$, Corollary 7.2.6(c)(2) determines the density $\mu(\{E : \text{Ш}'_E \in \mathcal{G}\} \mid \text{rk}'_E = 1)$ (with probability 1). The analogous problem for $\text{rk}'_E = 0$ is of a different nature since each $G \in \mathfrak{S}$ arises with density 0 (Corollary 7.2.6(d)), so it may be

unreasonable to expect the density to exist for all of the uncountably many $\mathcal{G}$ in this case, but one can still ask about specific $\mathcal{G}$.

When $\mathcal{G}$ is the set of squares of cyclic groups, Corollary 7.2.6(d)(2) gives the density of E with $\Sha'_E \in \mathcal{G}$ (with probability 1). In contrast, consider $\mathcal{G} := \{G \in \mathfrak{S} : \sqrt{\#G} \text{ is squarefree}\}$. The squarefree condition can again be checked one p at a time, but this time the reductions modulo p lie on a codimension 1 subscheme unfortunately, and the condition instead is about *squarefree values* of the Pfaffian. This means that in order to apply the Ekedahl sieve, we would need results on squarefree values of a multivariable polynomial. There is a well-known heuristic, that the density of squarefree values is the product over p of the density of values not divisible by p^2, but making this rigorous in general seems to require the *abc* conjecture [Gra98, Poo03] (and in the multivariable case the results use a nonstandard way of counting, involving a nonsquare box). Thus, although it is clear what to predict for the density of E such that $\sqrt{\#\Sha(E)}$ is squarefree among all rank 0 elliptic curves, it is not clear that we can prove the corresponding statement about matrices.

11.2. Asymptotics for rank 0 elliptic curves with specified Shafarevich–Tate group.
It is conjectured that for each $G \in \mathfrak{S}$, the density of rank 0 curves E with $\Sha(E) \simeq G$ is 0: see Remark 5.3.4 and consequence (iv) of Section 8.1. But one can ask a more precise question:

Question 11.2.1. Given $G \in \mathfrak{S}$, what is the asymptotic growth rate of $\#\{E \in \mathcal{E}_{\leq H} : \operatorname{rk} E(\mathbb{Q}) = 0 \text{ and } \Sha(E) \simeq G\}$ as $H \to \infty$?

The literature contains contradictory conjectures even on whether the set of rank 0 curves with $\Sha(E) \simeq G$ is finite:

(a) Elkies [Elk02, Section 3.2] (in a family of quadratic twists) and Hindry [Hin07, Conjecture 5.4] (in general) made conjectures implying that $L(E,1) \gg H^{-o(1)}$ whenever $L(E,1) \neq 0$. Combining these with Theorem 6.4.2(a) would imply conditionally that $\#\Sha(E) \gg H^{1/12-o(1)}$ for all rank 0 curves, so only finitely many rank 0 curves would have $\Sha(E) \simeq G$. (Hindry, however, no longer believes his conjecture: see [HP16, Observations 1.15(b)].)

(b) Watkins [Wat08a, Section 4.5], on the other hand, considers it likely that among elliptic curves with root number $+1$, the outcome $\Sha_0 = 1$ is about as common as $\Sha_0 = 0$ (that is, $\operatorname{rk} E(\mathbb{Q}) \geq 2$). The numerical data in [DJS16, Section 11] for a family of quadratic twists supports this guess.

We suspect that (b) is the truth, and more precisely that for each G,

$$\#\{E \in \mathcal{E}_{\leq H} : \operatorname{rk} E(\mathbb{Q}) = 0 \text{ and } \Sha(E) \simeq G\} \overset{?}{=} H^{19/24+o(1)}. \tag{11.2.2}$$

Dąbrowski, Jędrzejak, and Szymaszkiewicz have formulated an analogous conjecture for their family of quadratic twists [DJS16, Conjecture 8].

This raises the question of whether the analogue of (11.2.2) for $(\operatorname{rk}'_E, \Sha'_E)$ can be proved. This has not been done, but [DRS93, Example 1.7] answer a closely related question by counting $A \in \operatorname{M}_n(\mathbb{Z})_{\mathrm{alt}, \leq X}$ with $\#(\operatorname{coker} A)$ equal to a given integer, instead of our finer question about $\operatorname{coker} A$ being a given group.

11.3. The distribution of the normalized size of the Shafarevich–Tate group of rank 0 curves.

Question 11.3.1. Does the uniform probability measure on the finite set
$$\left\{ \#\text{Ш}(E)/H^{1/12} : E \in \mathscr{E}_{\leq H}, \ \text{rk}\, E(\mathbb{Q}) = 0 \right\}$$
converge weakly to a limiting distribution on $\mathbb{R}_{\geq 0}$ as $H \to \infty$?

Our model would predict an answer if $X(H)$ and $\eta(H)$ were specified precisely instead of requiring only $X(H)^{\eta(H)} = H^{1/12+o(1)}$. But the limiting distribution, even if it existed, would not be robust: for example, it would probably change if we sampled integer matrices from a sphere instead of a box. We see no reason for favoring any particular shape, so we view our model as being insufficient for answering Question 11.3.1.

11.4. The distribution of the Shafarevich–Tate groups of curves of rank $r \geq 2$.
For $r \in \{0, 1\}$, Corollary 7.2.6(b) determined that the distribution of $\text{Ш}'_E[p^\infty]$ conditioned on $\text{rk}'_E = r$ agrees with Delaunay's predictions for the actual $\text{Ш}(E)[p^\infty]$. Can we refine the calculations of Section 9 to determine the distribution for $r \geq 2$? As mentioned in Remark 7.2.4, it may very well be that the distribution differs from Delaunay's prediction for $r \geq 2$.

We can also ask about the whole group $\text{Ш}'_E$ instead of only one p-primary part at a time. Is it true that for each $r \geq 1$ and $G \in \mathfrak{S}$,
$$\mu\left(\{E \in \mathscr{E} : \text{Ш}'_E \simeq G\} \mid \text{rk}'_E = r\right) = \prod_p \mu\left(\{E \in \mathscr{E} : \text{Ш}'_E[p^\infty] \simeq G[p^\infty]\} \mid \text{rk}'_E = r\right) > 0?$$

We proved this for $r = 1$: see Corollary 7.2.6(c)(1).

11.5. Higher rank calibration.

Question 11.5.1. Would using an upper bound on $\text{Ш}(E)$ for rank r curves for some fixed $r \geq 1$ instead of $r = 0$ have led to the same calibration of $X(H)^{\eta(H)}$?

The answer to this question is not necessary for our model, but a positive answer could be viewed as further support for it. Many of the necessary ingredients are in place:

1. The Riemann hypothesis for $L(E, s)$ implies $L^{(r)}(E, 1) \ll_r N^{o(1)} \leq H^{o(1)}$, where N is the conductor (use the proof of [CG06, Theorem 1] to bound $L(E, s)$ on a circle of radius $1/\log N$ containing 1, and apply the Cauchy integral formula for $L^{(r)}(E, 1)$).
2. Lang's conjectural lower bound for the canonical height of a non-torsion point [Lan90, Section 7] implies that the regulator is $\geq H^{o(1)}$. Hindry and Silverman [Sil81, HS88] have made significant progress towards this conjecture.
3. Bounds on the Tamagawa numbers, torsion, and real period are as in Section 6.
4. Substituting these into the Birch and Swinnerton-Dyer conjecture yields a conjectural upper bound $\#\text{Ш}(E) \overset{?}{\leq} H^{1/12+o(1)}$ for all E of rank r.
5. On the matrix side, for fixed n and r, all $A \in \text{M}_n(\mathbb{Z})_{\text{alt}, \leq X}$ of corank $\geq r$ satisfy $\#(\text{coker}\, A)_{\text{tors}} \ll X^{n-r}$ as $X \to \infty$.
6. Matching these upper bounds with $n \to \infty$ would yield $X(H)^{\eta(H)} = H^{1/12+o(1)}$ as before.

But $\text{Ш}(E)$ for $r \geq 1$ is usually small (at least conjecturally, as discussed in Remark 5.3.4), as is $\#(\text{coker}\, A)_{\text{tors}}$ for $r \geq 1$, so it is unclear whether the rare events of a large $\text{Ш}(E)$ or a large $\#(\text{coker}\, A)_{\text{tors}}$ occur frequently enough to make these upper bounds sharp. Hence the answer to Question 11.5.1 is not clear.

12. Generalizing to other global fields

Fix a global field K. Let $\mathscr{E}_K$ be a set of elliptic curves over K representing each isomorphism class once. Let $B_K := \limsup_{E \in \mathscr{E}_K} \operatorname{rk} E(K)$. Thus B_K, if finite, is the smallest integer such that $\{E \in \mathscr{E}_K : \operatorname{rk} E(K) > B_K\}$ is finite. Section 8.2 suggests that $20 \le B_{\mathbb{Q}} \le 21$ (in rank 21, the model could go either way depending on the sign of the function implicit in the $o(1)$ in the exponent). A naive generalization of our model (see Section 12.2) would suggest that $20 \le B_K \le 21$ for all global fields K.

12.1. Subfield issues.
But there is a problem. Some elliptic curves E over K may have extra structure that our model did not take into account. For example, if E is a base change of a curve over a subfield $K_0 \subsetneq K$ such that K/K_0 is Galois, then the group $G := \operatorname{Gal}(K/K_0)$ acts on $E(K)$ and $\text{Ш}(E)$. The exploitation of such curves E leads to the theorems of Section 12.4, which show in particular that there exist number fields K making B_K arbitrarily large.

This makes it clear that separate models are needed to describe such curves. Analogously, Cohen and Lenstra in [CL84, §9, III] observed that class groups of cyclic cubic extensions are not just abelian groups but $\mathbb{Z}[\zeta_3]$-modules and should be modeled as such. If E descends to a subfield K_0 such that K/K_0 is not Galois, the relevance of the extra structure is not as obvious, but it still may be that a separate model is needed as it is in the Cohen-Lenstra-Martinet heuristics for class groups of arbitrary fields [CM90].

We will not attempt here to construct a model for every possible situation. Instead we restrict attention to the set $\mathscr{E}_K^\circ$ consisting of $E \in \mathscr{E}_K$ such that E is not a base change of a curve from a proper subfield. Let $B_K^\circ := \limsup_{E \in \mathscr{E}_K^\circ} \operatorname{rk} E(K)$.

12.2. Heuristic for global fields.
If K is a number field, let S be the set of archimedean places. If K is a global function field, let S be any nonempty finite set of places. In asymptotic estimates below, we view K and S as fixed; e.g., $X \ll 1$ would mean that X is bounded by a constant depending on K and S. Let $\mathcal{O}_{K,S}$ be the ring of S-integers in K. For simplicity, we assume that $\operatorname{Cl} \mathcal{O}_{K,S}$ is trivial. (If one does not want to enlarge S to ensure this, one can use the finiteness of $\operatorname{Cl} \mathcal{O}_{K,S}$ to verify that the estimates remain valid up to bounded factors.) If v is a nonarchimedean place of K, let $\mathcal{O}_v \subset K_v$ be the valuation ring. For each place v of K, we fix a Haar measure μ_v on K_v: if v is archimedean, let μ_v be Lebesgue measure; if v is nonarchimedean, choose μ_v so that $\mu_v(\mathcal{O}_v) = 1$. For $a \in K_v$, let $|a|_v$ be the factor by which multiplication-by-a scales μ_v. For simplicity we assume $\operatorname{char} K \ne 2, 3$ from now on; minor modifications would be needed to handle the general case.

Each $E \in \mathscr{E}_K$ is represented by an equation $y^2 = x^3 + Ax + B$ with $A, B \in K$ uniquely determined up to replacing (A, B) by $(\lambda^4 A, \lambda^6 B)$ for $\lambda \in K^\times$. Choosing λ judiciously lets us assume that $A, B \in \mathcal{O}_{K,S}$. Since $\operatorname{Cl} \mathcal{O}_{K,S} = \{1\}$, we may assume also that for every $v \notin S$ we have $v(A) < 4$ or $v(B) < 6$. The only remaining freedom is to scale (A, B) using a λ in $\mathcal{O}_{K,S}^\times$. For $v \in S$, define $\operatorname{ht}_v E = H_v := \max\{|4A^3|_v, |27B^2|_v\}$. By the product formula, the height $\operatorname{ht} E = H := \prod_{v \in S} H_v \in \mathbb{R}_{>0}$ is independent of the scaling. Our model for $\operatorname{rk} E(K)$ is the same as for $\mathbb{Q}$, with X^n to be related to this new H.

Let $\Delta := -16(4A^3 + 27B^2) \in \mathcal{O}_{K,S}$. Let $\omega := \frac{dx}{2y}$. Let $\Omega_v := \int_{A(K_v)} |\omega_v| \mu_v \in \mathbb{R}_{>0}$. Let $\Omega := \prod_{v \in S} \Omega_v$. Essentially the same argument as in Section 6.1 shows that

$$H_v^{-1/12} \ll \Omega_v \ll H_v^{-1/12} \log(H_v/|\Delta|_v). \tag{12.2.1}$$

By the product formula, $\prod_{v \in S} |\Delta|_v \geq 1$, so $\sum_{v \in S} \log(H_v/|\Delta|_v) \leq \log H$. By the AM-GM inequality, $\prod_{v \in S} \log(H_v/|\Delta|_v) \ll (\log H)^{\#S}$. Therefore, taking the product of (12.2.1) over $v \in S$ yields $\Omega = H^{-1/12+o(1)}$.

For $v \notin S$, the Tamagawa factors c_v may be bounded in terms of $v(\Delta)$, so $\prod c_v \leq H^{o(1)}$ as before. By [Mer96], $\#E(K)_{\mathrm{tors}} \ll 1$.

Let $L_S(E, s)$ be the L-series of E with the Euler factors at $v \in S$ omitted. We assume that $L_S(E, s)$ admits an analytic continuation to $\mathbb{C}$ that $L_S(E, 1) \geq 0$, and that the average of $L_S(E, 1)$ over $E \in \mathscr{E}_K$ is $\overset{?}{\asymp} 1$. Define III_0 as before.

The Birch and Swinnerton-Dyer conjecture [Tat95, Conjecture (B)] implies that

$$L_S(E, 1) \overset{?}{=} \delta \frac{\mathrm{III}_0(E)\,\Omega \prod_{v \notin S} c_v}{\#E(K)_{\mathrm{tors}}^2},$$

where δ is a constant depending only on K (and our choice of measures μ_v). As before, we obtain $\mathrm{Average}_{E \in \mathscr{E}_{K, \leq H}}\, \mathrm{III}_0 \overset{?}{\asymp} H^{1/12+o(1)}$ as $H \to \infty$, which suggests the same calibration $X^{n/2} \sim H^{1/12}$ as for $\mathbb{Q}$.

Next we argue that $\#\mathscr{E}_{K, \leq H} \asymp H^{5/6}$ as $H \to \infty$. For $v \in S$, choose $C_v \in \mathbb{R}_{>0}$ such that $\prod_{v \in S} C_v = H$. Parallelotope estimates [Lan94, V.§2, Theorem 1] imply that the number of $A \in \mathcal{O}_{K,S}$ satisfying $|4A^3|_v \leq C_v$ for all $v \in S$ is $\asymp H^{1/3}$, and similarly for B; combining these estimates with an elementary sieve constructs $\asymp H^{5/6}$ curves, but some of them may be equivalent under scaling by $\lambda \in \mathcal{O}_{K,S}^\times$. If we fix suitably small constants $\epsilon_v > 0$ and remove the elliptic curves satisfying $\max\{|4A^3|_v, |27B^2|_v\} \leq \epsilon_v C_v$ for all $v \in S$, then the equivalence classes of those remaining are of bounded size; thus $\#\mathscr{E}_{K, \leq H} \gg H^{5/6}$. On the other hand, for suitably large constants $M_v > 0$, geometry of numbers shows that every $E \in \mathscr{E}_{K, \leq H}$ is represented by a pair $(A, B) \in \mathcal{O}_{K,S}^2$ such that $\max\{|4A^3|_v, |27B^2|_v\} \leq M_v C_v$ for all $v \in S$, so parallelotope estimates imply that $\#\mathscr{E}_{K, \leq H} \ll \prod_v (M_v C_v)^{1/3}(M_v C_v)^{1/2} \ll H^{5/6}$.

12.3. **Number fields.** Let $\mathscr{E}_{K, \leq H}^\circ := \{E \in \mathscr{E}_K^\circ : \mathrm{ht}\, E \leq H\}$. If $E \in \mathscr{E}$ is definable over some $K_0 \subsetneq K$, then $j(E) \in K_0$, and this implies nontrivial polynomial equations satisfied by the coefficients of A and B relative to a basis for K over $\mathbb{Q}$. The fraction of $E \in \mathscr{E}_{K, \leq H}$ for which these equations hold is asymptotically 0, so $\#\mathscr{E}_{K, \leq H}^\circ \asymp \#\mathscr{E}_{K, \leq H} \asymp H^{5/6}$.

Now the same arguments as for $\mathbb{Q}$ suggest $20 \overset{?}{\leq} B_K^\circ \leq 21$.

Remark 12.3.1. The upper bound $B_K^\circ \leq 21$ must fail for many number fields K, however, because of certain special families of elliptic curves, as we now explain. Shioda [Shi92] proves that the elliptic curve $E \colon y^2 = x^3 + t^{360} + 1$ over $\mathbb{C}(t)$ has rank 68. The coefficients involved in the coordinates of generators of $E(\mathbb{C}(t))$ are algebraic; let K be any number field containing them all. Then the rank of E over $K(t)$ is at least 68. Next, Néron's specialization result [Nér52, IV, Théorème 6] shows that for a in a density 1 subset of K, setting $t = a$ results in an elliptic curve $E_a \in \mathscr{E}_K^\circ$ such that $\mathrm{rk}\, E_a(K) \geq 68$. (Néron's result states only that infinitely many such a exist, but its proof, based on the Hilbert irreducibility theorem, gives a density 1 set of such a. In fact, by a refinement of Silverman [Sil83, Theorem C], the specialization map is injective with only finitely many exceptions.) These E_a fall into infinitely many isomorphism classes over K, so $B_K^\circ \geq 68$.

It still seems plausible that B_K° and B_K are finite for each number field K.

12.4. **Number fields with infinitely many elliptic curves of high rank.** Using the results of Cornut [Cor02] and Vatsal [Vat03] on Heegner points over anticyclotomic $\mathbb{Z}_p$-extensions of imaginary quadratic fields, one can prove the following.

Theorem 12.4.1. *There exists an effectively constructible sequence of number fields K in which $[K : \mathbb{Q}] \to \infty$ and $B_K \geq [K : \mathbb{Q}]/2$.*

Proof. Fix a prime $p \geq 5$. Choose an elliptic curve over $\mathbb{F}_p$ whose trace of Frobenius a_p is not 0, 1, or 2 (see [Maz84, p. 203] for the relevance of this condition), and lift it to an elliptic curve E over $\mathbb{Q}$ with good reduction at p. Let N be the conductor of E. By quadratic reciprocity and the Chinese remainder theorem, we can find an imaginary quadratic field K_0 satisfying the *Heegner hypothesis* that all prime factors of N split in K_0. Let K_∞ be the anticyclotomic $\mathbb{Z}_p$-extension of K_0, let K_n be the degree p^n subextension, and let $\Lambda := \mathbb{Z}_p[[\mathrm{Gal}(K_\infty/K)]]$ be the Iwasawa algebra [Maz84, Section 17]. The theorem on page 496 of [Cor02] implies Mazur's conjecture [Maz84, bottom of p. 203] that the Heegner module $\mathscr{E}(K_\infty)$ [Maz84, p. 203] is nonzero, and hence is a free Λ-module of rank 1. This implies that $\mathrm{rk}\, E(K_n) \geq p^n$. (The idea that Heegner points might yield unbounded rank in anticyclotomic towers is due to Kurčanov [Kur77].) For any prime ℓ splitting in $K_0/\mathbb{Q}$ such that $\left(\frac{\ell}{p}\right) = +1$, the twist E_ℓ satisfies the Heegner hypothesis and its reduction mod p has the same a_p, so $\mathrm{rk}\, E_\ell(K_n) \geq p^n$ too. The base extensions to K_n of these twists cover infinitely many K_n-isomorphism classes, so $B_{K_n} \geq p^n = [K_n : \mathbb{Q}]/2$. $\qquad\square$

Call a number field **multiquadratic** if it is a compositum of quadratic fields. We can obtain a faster rate of growth than in Theorem 12.4.1 by using multiquadratic fields instead of anticyclotomic fields:

Theorem 12.4.2. *For every $n \geq 0$, there exists a degree 2^n multiquadratic field K such that a positive proportion of $E \in \mathscr{E} = \mathscr{E}_\mathbb{Q}$ satisfy $\mathrm{rk}\, E(K) \geq 2^n$ and hence $B_K \geq 2^n = [K : \mathbb{Q}]$.*

Proof. For $E \in \mathscr{E}_\mathbb{Q}$, let $\Delta(E)$ be its minimal discriminant, and for $d \in \mathbb{Q}^\times$ (or $\mathbb{Q}^\times/\mathbb{Q}^{\times 2}$), let E_d be the corresponding quadratic twist of E. If G is a finite subgroup of $\mathbb{Q}^\times/\mathbb{Q}^{\times 2}$, and $K = \mathbb{Q}(\sqrt{G})$ is the corresponding multiquadratic field, then $\mathrm{rk}\, E(K) = \sum_{d \in G} \mathrm{rk}\, E_d(\mathbb{Q})$, as one sees by decomposing the $\mathrm{Gal}(K/\mathbb{Q})$-representation $E(K) \otimes \mathbb{Q}$.

Given n, the multidimensional density Hales–Jewett theorem of Furstenberg and Katznelson [FK91, Theorem 2.5] (reproved by D. H. J. Polymath [Pol12, Theorem 1.6]) shows that if m is sufficiently large, any subset of $\mathbb{F}_2^m$ of density 26% or more contains an affine n-plane. (The reason for using 26% will be explained later.) Fix such an m. Let $q_1, \ldots, q_m$ be primes congruent to 1 modulo 4 and to ± 2 modulo 5. Let $L := \mathbb{Q}(\sqrt{q_1}, \ldots, \sqrt{q_m})$.

Consider the following conditions on an elliptic curve $E \in \mathscr{E}$:

(a) E has good ordinary reduction at 5;
(b) E has good reduction at every q_i; and
(c) For every prime $\ell \equiv \pm 1 \pmod 5$ we have $\mathrm{ord}_\ell(\Delta(E)) \notin \{5, 10, 15, \ldots\}$.

Let S be the set of $E \in \mathscr{E}$ satisfying (a) and (b); this is a positive density subset of $\mathscr{E}$. Now

1. A method of Wong [Won01] (see [BSZ14, Theorem 16]) shows that S contains a family F' (a finite union of large families, in the sense of [BSZ14, Section 2.3]) of relative density $> 55.01\%$ in S in which the root number is equidistributed.

34

2. A squarefree sieve shows that at least 99.9% of the curves in F' satisfy (c) as well (cf. [BSZ14, Lemma 19]); these curves are contained in the set $S_1(5)$ of [BSZ14, Section 3.1].
3. The arguments of [BSZ14, Section 3.4] show that at least $19/40$ of the curves in this subfamily have rank 1.

Thus at least $(0.5501)(0.999)(19/40) > 26\%$ of the curves in S have rank 1.

Let $\mathcal{D}$ be the set of products formed by subsets of $\{q_1, \ldots, q_m\}$. For each $d \in \mathcal{D}$, let $S_d := \{E_d : E \in S\}$. The same arguments as above show that for each d, the subset S_d is of positive density in $\mathscr{E}$, and more than 26% of the curves in S_d have rank 1.

Choose $c > 0$ small enough that $E \in \mathscr{E}_{\leq cH}$ implies $E_d \in \mathscr{E}_{\leq H}$ for all $d \in \mathcal{D}$. For each $E \in S \cap \mathscr{E}_{\leq cH}$, call $\{E_d : d \in \mathcal{D}\}$ a hypercube. For a positive fraction (independent of H) of hypercubes $\mathcal{H}$, at least 26% of the 2^m curves in $\mathcal{H}$ have rank 1. In each such $\mathcal{H}$, the choice of m guarantees an affine n-plane consisting entirely of twists of rank ≥ 1; any one of these twists E has rank at least 2^n over the degree 2^n multiquadratic field $K \subseteq L$ corresponding to the orientation of the n-plane. But L has only finitely many such subfields, so one such K occurs for a positive fraction of hypercubes. These give a positive fraction of $E \in \mathscr{E}_{\leq H}$ for which $\mathrm{rk}\, E(K) \geq 2^n$. Thus $B_K \geq 2^n$. $\qquad\square$

Remark 12.4.3. The proof of Theorem 12.4.2 produces a finite list of degree 2^n multiquadratic fields K such that one of them satisfies $B_K \geq 2^n$, but it seems that we cannot determine effectively which K it is! We can, however, effectively construct their compositum L, a multiquadratic field of larger degree such that $B_L \geq 2^n$.

12.5. Function fields.
Let K be a global function field. Tate and Shafarevich [TŠ67] and Ulmer [Ulm02] constructed families of elliptic curves showing that $B_K = \infty$. But the elliptic curves of high rank constructed are always defined over proper subfields of K. (For example, in [TŠ67] the curves are isotrivial and not necessarily constant, but still they are defined over a proper *infinite* subfield of K.) Thus B_K° may still be finite.

REFERENCES

[BHK+16] Jennifer S. Balakrishnan, Wei Ho, Nathan Kaplan, Simon Spicer, William Stein, and James Weigandt, *Databases of elliptic curves ordered by height and distributions of Selmer groups and ranks*, LMS J. Comput. Math. **19** (2016), no. suppl. A, 351–370. MR3540965 ↑10.1

[BM90] V. V. Batyrev and Yu. I. Manin, *Sur le nombre des points rationnels de hauteur borné des variétés algébriques*, Math. Ann. **286** (1990), no. 1-3, 27–43, DOI 10.1007/BF01453564 (French). MR1032922 (91g:11069) ↑1.1.2

[BMSW07] Baur Bektemirov, Barry Mazur, William Stein, and Mark Watkins, *Average ranks of elliptic curves: tension between data and conjecture*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 2, 233–254, DOI 10.1090/S0273-0979-07-01138-X. MR2291676 (2009e:11107) ↑10.1

[BKLPR15] Manjul Bhargava, Daniel M. Kane, Hendrik W. Lenstra Jr., Bjorn Poonen, and Eric Rains, *Modeling the distribution of ranks, Selmer groups, and Shafarevich-Tate groups of elliptic curves*, Camb. J. Math. **3** (2015), no. 3, 275–321. MR3393023 ↑1.1, 1.1.4, 5.3, 5.3.3, 5.3.4, 7.1, 7.2.4, ii

[BSZ14] Manjul Bhargava, Christopher Skinner, and Wei Zhang, *A majority of elliptic curves over $\mathbb{Q}$ satisfy the Birch and Swinnerton-Dyer conjecture*, July 17, 2014. Preprint, `arXiv:1407.1826v2`. ↑1.2, 1, 2, 3

[Bru92] Armand Brumer, *The average rank of elliptic curves. I*, Invent. Math. **109** (1992), no. 3, 445–472, DOI 10.1007/BF01232033. MR1176198 (93g:11057) ↑3.1, 3.1, 5.1

[BMc90] Armand Brumer and Oisín McGuinness, *The behavior of the Mordell-Weil group of elliptic curves*, Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 2, 375–382, DOI 10.1090/S0273-0979-1990-15937-3. MR1044170 (91b:11076) ↑10.1

[Cas62] J. W. S. Cassels, *Arithmetic on curves of genus* 1. *IV. Proof of the Hauptvermutung*, J. reine angew. Math. **211** (1962), 95–112. MR0163915 (29 #1214) ↑5.3

[Cas66] J. W. S. Cassels, *Diophantine equations with special reference to elliptic curves*, J. London Math. Soc. **41** (1966), 193–291. MR0199150 (33 #7299) ↑3.1

[CL84] H. Cohen and H. W. Lenstra Jr., *Heuristics on class groups of number fields*, Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, pp. 33–62, DOI 10.1007/BFb0099440. MR756082 (85j:11144) ↑1.1, 4, 4.2, 12.1

[CM90] Henri Cohen and Jacques Martinet, *Étude heuristique des groupes de classes des corps de nombres*, J. reine angew. Math. **404** (1990), 39–76 (French). MR1037430 (91k:11097) ↑4, 12.1

[CG06] J. Brian Conrey and Amit Ghosh, *Remarks on the generalized Lindelöf hypothesis*, Funct. Approx. Comment. Math. **36** (2006), 71–78, DOI 10.7169/facm/1229616442. MR2296639 (2007k:11140) ↑1

[CKRS02] J. B. Conrey, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, *On the frequency of vanishing of quadratic twists of modular L-functions*, Number theory for the millennium, I (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, pp. 301–315. MR1956231 (2003m:11141) ↑a, b, b, 10.2

[CRSW07] J. Brian Conrey, Michael O. Rubinstein, Nina C. Snaith, and Mark Watkins, *Discretisation for odd quadratic twists*, Ranks of elliptic curves and random matrix theory, London Math. Soc. Lecture Note Ser., vol. 341, Cambridge Univ. Press, Cambridge, 2007, pp. 201–214, DOI 10.1017/CBO9780511735158.012. MR2322346 (2008h:11064) ↑b, b, b1, b2, b3, c, d

[Cor02] Christophe Cornut, *Mazur's conjecture on higher Heegner points*, Invent. Math. **148** (2002), no. 3, 495–523, DOI 10.1007/s002220100199. MR1908058 (2004e:11069a) ↑12.4, 12.4

[Cre97] J. E. Cremona, *Algorithms for modular elliptic curves*, 2nd ed., Cambridge University Press, Cambridge, 1997. MR1628193 (99e:11068) ↑10.1

[DJS16] Andrzej Dąbrowski, Tomasz Jędrzejak, and Lucjan Szymaszkiewicz, *Behaviour of the order of Tate-Shafarevich groups for the quadratic twists of* $(X_0)(49)$, Elliptic curves, modular forms and Iwasawa theory, Springer Proc. Math. Stat., vol. 188, Springer, Cham, 2016, pp. 125–159. MR3629650 ↑6.4.3, b, 11.2

[Del01] Christophe Delaunay, *Heuristics on Tate-Shafarevitch groups of elliptic curves defined over* $\mathbb{Q}$, Experiment. Math. **10** (2001), no. 2, 191–196. MR1837670 (2003a:11065) ↑1.1, 5.3, ii, iii, iv

[Del05] Christophe Delaunay, *Moments of the orders of Tate-Shafarevich groups*, Int. J. Number Theory **1** (2005), no. 2, 243–264, DOI 10.1142/S1793042105000133. MR2173383 (2006h:11077) ↑e

[Del07] Christophe Delaunay, *Heuristics on class groups and on Tate-Shafarevich groups: the magic of the Cohen-Lenstra heuristics*, Ranks of elliptic curves and random matrix theory, London Math. Soc. Lecture Note Ser., vol. 341, Cambridge Univ. Press, Cambridge, 2007, pp. 323–340. MR2322355 (2008i:11089) ↑1.1, 5.3, ii, iii

[DD03] C. Delaunay and S. Duquesne, *Numerical investigations related to the derivatives of the L-series of certain elliptic curves*, Experiment. Math. **12** (2003), no. 3, 311–317. MR2034395 (2004m:11098) ↑3.4

[DJ14] Christophe Delaunay and Frédéric Jouhet, p^ℓ-*torsion points in finite abelian groups and combinatorial identities*, Adv. Math. **258** (2014), 13–45, DOI 10.1016/j.aim.2014.02.033. MR3190422 ↑1.1, 5.3, ii, iii

[DR08] Christophe Delaunay and Xavier-François Roblot, *Regulators of rank one quadratic twists*, J. Théor. Nombres Bordeaux **20** (2008), no. 3, 601–624 (English, with English and French summaries). MR2523310 (2011a:11128) ↑e

[DW07] Christophe Delaunay and Mark Watkins, *The powers of logarithm for quadratic twists*, Ranks of elliptic curves and random matrix theory, London Math. Soc. Lecture Note Ser., vol. 341, Cambridge Univ. Press, Cambridge, 2007, pp. 189–193, DOI 10.1017/CBO9780511735158.010. MR2322344 (2008c:11084) ↑b

[Den10] Christopher Deninger, *The Hilbert-Polya strategy and height pairings*, Casimir force, Casimir operators and the Riemann hypothesis, Walter de Gruyter, Berlin, 2010, pp. 275–283. MR2777722 ↑1.1.5

[Duj] Andrej Dujella, *Infinite families of elliptic curves with high rank and prescribed torsion.* `http://web.math.pmf.unizg.hr/~duje/tors/generic.html`, Accessed 2017-11-20. ↑8.3

[DRS93] W. Duke, Z. Rudnick, and P. Sarnak, *Density of integer points on affine homogeneous varieties*, Duke Math. J. **71** (1993), no. 1, 143–179, DOI 10.1215/S0012-7094-93-07107-4. MR1230289 (94k:11072) ↑11.2

[Dur10] Rick Durrett, *Probability: theory and examples*, 4th ed., Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010. MR2722836 (2011e:60001) ↑7.2.5

[Elk02] Noam D. Elkies, *Curves $Dy^2 = x^3 - x$ of odd analytic rank*, Algorithmic number theory (Sydney, 2002), Lecture Notes in Comput. Sci., vol. 2369, Springer, Berlin, 2002, pp. 244–251. MR2041088 (2005a:11075) ↑3.4, a

[Elk06] Noam D. Elkies, $\mathbb{Z}^{28}$ *in $E(\mathbb{Q})$, etc.*, May 3, 2006. Email to `NMBRTHRY@LISTSERV.NODAK.EDU` mailing list. ↑1, 2

[EK95] Alex Eskin and Yonatan R. Katznelson, *Singular symmetric matrices*, Duke Math. J. **79** (1995), no. 2, 515–547, DOI 10.1215/S0012-7094-95-07913-7. MR1344769 (96h:11099) ↑9.1, 9.2, 9.2.3, 9.3, 9.3.2, 9.3.3, 9.3, 9.3.4, 9.3.5, 9.4.3, 9.4.4, 9.4.5, 9.4, 9.5.1, 9.5

[FGH07] David W. Farmer, S. M. Gonek, and C. P. Hughes, *The maximum size of L-functions*, J. reine angew. Math. **609** (2007), 215–236, DOI 10.1515/CRELLE.2007.064. MR2350784 (2009b:11140) ↑3.1

[FK07] Étienne Fouvry and Jürgen Klüners, *On the 4-rank of class groups of quadratic number fields*, Invent. Math. **167** (2007), no. 3, 455–513, DOI 10.1007/s00222-006-0021-2. MR2276261 ↑4.2

[FW89] Eduardo Friedman and Lawrence C. Washington, *On the distribution of divisor class groups of curves over a finite field*, Théorie des nombres (Quebec, PQ, 1987), de Gruyter, Berlin, 1989, pp. 227–239. MR1024565 (91e:11138) ↑1.1, 4, 4.1.1, 4.2, 4.2.3

[FK91] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem*, J. Anal. Math. **57** (1991), 64–119, DOI 10.1007/BF03041066. MR1191743 (94f:28020) ↑12.4

[Ger87] Frank Gerth III, *Extension of conjectures of Cohen and Lenstra*, Exposition. Math. **5** (1987), no. 2, 181–184. MR887792 (88h:11081) ↑4.2

[Gol79] Dorian Goldfeld, *Conjectures on elliptic curves over quadratic fields*, Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), Lecture Notes in Math., vol. 751, Springer, Berlin, 1979, pp. 108–118. MR564926 (81i:12014) ↑1.1, i

[GS95] Dorian Goldfeld and Lucien Szpiro, *Bounds for the order of the Tate-Shafarevich group*, Compositio Math. **97** (1995), no. 1-2, 71–87. Special issue in honour of Frans Oort. MR1355118 (97a:11102) ↑6

[Gra95] Andrew Granville, *Harald Cramér and the distribution of prime numbers*, Scand. Actuar. J. **1** (1995), 12–28, DOI 10.1080/03461238.1995.10413946. Harald Cramér Symposium (Stockholm, 1993). MR1349149 (96g:01002) ↑1.1

[Gra98] Andrew Granville, *ABC allows us to count squarefrees*, Internat. Math. Res. Notices **19** (1998), 991–1009. MR1654759 (99j:11104) ↑11.1

[Gro87] Benedict H. Gross, *Heights and the special values of L-series*, Number theory (Montreal, Que., 1985), CMS Conf. Proc., vol. 7, Amer. Math. Soc., Providence, RI, 1987, pp. 115–187. MR894322 (89c:11082) ↑a

[Hal38] P. Hall, *A partition formula connected with Abelian groups*, Comment. Math. Helv. **11** (1938), no. 1, 126–129, DOI 10.1007/BF01199694. MR1509594 ↑4.2

[HS17] Robert Harron and Andrew Snowden, *Counting elliptic curves with prescribed torsion*, J. Reine Angew. Math. **729** (2017), 151–170, DOI 10.1515/crelle-2014-0107. MR3680373 ↑8.3

[Hin07] Marc Hindry, *Why is it difficult to compute the Mordell-Weil group?*, Diophantine geometry, CRM Series, vol. 4, Ed. Norm., Pisa, 2007, pp. 197–219. MR2349656 (2008i:11074) ↑6, 6.1.3, 6.2, a

[HP16] Marc Hindry and Amílcar Pacheco, *An analogue of the Brauer–Siegel theorem for abelian varieties in positive characteristic*, Mosc. Math. J. **16** (2016), no. 1, 45–93. MR3470576 ↑6, a

[HS88] M. Hindry and J. H. Silverman, *The canonical height and integral points on elliptic curves*, Invent. Math. **93** (1988), no. 2, 419–450, DOI 10.1007/BF01394340. MR948108 (89k:11044) ↑2

[Hon60] Taira Honda, *Isogenies, rational points and section points of group varieties*, Japan. J. Math. **30** (1960), 84–101. MR0155828 (27 #5762) ↑3.1

[Hor15] Ruthi Hortsch, *Counting elliptic curves of bounded Faltings height*, May 19, 2015. Preprint, `arXiv:1505.05112v1`. ↑5.1.3

[IS00] H. Iwaniec and P. Sarnak, *Perspectives on the analytic theory of L-functions*, Geom. Funct. Anal. **Special Volume** (2000), 705–741. GAFA 2000 (Tel Aviv, 1999). MR1826269 (2002b:11117) ↑6.3

[KS99a] Nicholas M. Katz and Peter Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications, vol. 45, American Mathematical Society, Providence, RI, 1999. MR2000b:11070 ↑1.1, b, i

[KS99b] Nicholas M. Katz and Peter Sarnak, *Zeroes of zeta functions and symmetry*, Bull. Amer. Math. Soc. (N.S.) **36** (1999), no. 1, 1–26, DOI 10.1090/S0273-0979-99-00766-1. MR1640151 (2000f:11114) ↑1.1, b, i

[KS00] J. P. Keating and N. C. Snaith, *Random matrix theory and L-functions at $s = 1/2$*, Comm. Math. Phys. **214** (2000), no. 1, 91–110, DOI 10.1007/s002200000262. MR1794267 (2002c:11108) ↑b

[KZ81] W. Kohnen and D. Zagier, *Values of L-series of modular forms at the center of the critical strip*, Invent. Math. **64** (1981), no. 2, 175–198, DOI 10.1007/BF01389166. MR629468 (83b:10029) ↑6.3

[Kur77] P. F. Kurčanov, *The zeta-function of elliptic curves over certain abelian extensions of imaginary quadratic fields*, Mat. Sb. (N.S.) **102(144)** (1977), no. 1, 56–70, 151 (Russian). MR0498391 (58 #16519) ↑12.4

[Lan83] Serge Lang, *Conjectured Diophantine estimates on elliptic curves*, Arithmetic and geometry, Vol. I, Progr. Math., vol. 35, Birkhäuser Boston, Boston, MA, 1983, pp. 155–171. MR717593 (85d:11024) ↑1.2, 6, 6.4.2

[Lan90] Serge Lang, *Old and new conjectured Diophantine inequalities*, Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 1, 37–75, DOI 10.1090/S0273-0979-1990-15899-9. MR1005184 (90k:11032) ↑2

[Lan94] Serge Lang, *Algebraic number theory*, 2nd ed., Graduate Texts in Mathematics, vol. 110, Springer-Verlag, New York, 1994. MR1282723 (95f:11085) ↑12.2

[Magma] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265. Computational algebra and number theory (London, 1993). Magma is available at `http://magma.maths.usyd.edu.au/magma/` . MR1484478 ↑10.1

[Maz77] B. Mazur, *Modular curves and the Eisenstein ideal*, Inst. Hautes Études Sci. Publ. Math. **47** (1977), 33–186 (1978). MR488287 (80c:14015) ↑6.4

[Maz84] B. Mazur, *Modular curves and arithmetic*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), PWN, Warsaw, 1984, pp. 185–211. MR804682 (87a:11054) ↑12.4

[Mer96] Loïc Merel, *Bornes pour la torsion des courbes elliptiques sur les corps de nombres*, Invent. Math. **124** (1996), no. 1-3, 437–449 (French). MR1369424 (96i:11057) ↑12.2

[Mes82] Jean-François Mestre, *Construction d'une courbe elliptique de rang ≥ 12*, C. R. Acad. Sci. Paris Sér. I Math. **295** (1982), no. 12, 643–644 (French, with English summary). MR688896 (84b:14019) ↑3.1

[Mes86] Jean-François Mestre, *Formules explicites et minorations de conducteurs de variétés algébriques*, Compositio Math. **58** (1986), no. 2, 209–232 (French). MR844410 (87j:11059) ↑3.1

[Mor22] L. J. Mordell, *On the rational solutions of the indeterminate equations of the third and fourth degrees*, Proc. Cambridge Phil. Soc. **21** (1922), 179–192. ↑1.1

[Nér52] André Néron, *Problèmes arithmétiques et géométriques rattachés à la notion de rang d'une courbe algébrique dans un corps*, Bull. Soc. Math. France **80** (1952), 101–166 (French). MR0056951 (15,151a) ↑12.3.1

[Nér56] A. Néron, *Propriétés arithmétiques de certaines familles de courbes algébriques*, Proceedings of the International Congress of Mathematicians, 1954, Amsterdam, vol. III, Erven P. Noordhoff N.V., Groningen; North-Holland Publishing Co., Amsterdam, 1956, pp. 481–488 (French). MR0087210 (19,321b) ↑3.1

[OS98] Ken Ono and Christopher Skinner, *Non-vanishing of quadratic twists of modular L-functions*, Invent. Math. **134** (1998), no. 3, 651–660, DOI 10.1007/s002220050275. MR1660945 (2000a:11077) ↑a

[Poi01] H. Poincaré, *Sur les propriétés arithmétiques des courbes algébriques*, J. Pures Appl. Math. (5) **7** (1901), 161–234. ↑1.1

[Poi50] Henri Poincaré, *Œuvres d'Henri Poincaré, Volume 5* (Albert Châtelet, ed.), Gauthier-Villars, Paris, 1950. ↑3.1

[Pol12] D. H. J. Polymath, *A new proof of the density Hales-Jewett theorem*, Ann. of Math. (2) **175** (2012), no. 3, 1283–1327, DOI 10.4007/annals.2012.175.3.6. MR2912706 ↑12.4

[Poo03] Bjorn Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J. **118** (2003), no. 2, 353–373. MR1980998 (2004d:11094) ↑11.1

[PR12] Bjorn Poonen and Eric Rains, *Random maximal isotropic subspaces and Selmer groups*, J. Amer. Math. Soc. **25** (2012), no. 1, 245–269, DOI 10.1090/S0894-0347-2011-00710-8. MR2833483 ↑1.1.3, 5.3

[PS99] Bjorn Poonen and Michael Stoll, *The Cassels-Tate pairing on polarized abelian varieties*, Ann. of Math. (2) **150** (1999), no. 3, 1109–1149. MR1740984 (2000m:11048) ↑c

[RS00] Karl Rubin and Alice Silverberg, *Ranks of elliptic curves in families of quadratic twists*, Experiment. Math. **9** (2000), no. 4, 583–590. MR1806293 (2001k:11105) ↑a

[RS01] Karl Rubin and Alice Silverberg, *Rank frequencies for quadratic twists of elliptic curves*, Experiment. Math. **10** (2001), no. 4, 559–569. MR1881757 (2002k:11081) ↑a

[RS02] Karl Rubin and Alice Silverberg, *Ranks of elliptic curves*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 4, 455–474 (electronic), DOI 10.1090/S0273-0979-02-00952-7. MR1920278 (2003f:11080) ↑2, 3.4

[Sar90] Peter Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, vol. 99, Cambridge University Press, Cambridge, 1990. MR1102679 (92k:11045) ↑a

[Sch68] Wolfgang M. Schmidt, *Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height*, Duke Math. J. **35** (1968), 327–339. MR0224562 (37 #161) ↑9.2.4, 9.2.5

[Ser97] Jean-Pierre Serre, *Lectures on the Mordell-Weil theorem*, 3rd ed., Aspects of Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1997. Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt; With a foreword by Brown and Serre. MR1757192 (2000m:11049) ↑4.2

[Shi92] Tetsuji Shioda, *Some remarks on elliptic curves over function fields*, Astérisque **209** (1992), 12, 99–114. Journées Arithmétiques, 1991 (Geneva). MR1211006 (94d:11046) ↑12.3.1

[Sil81] Joseph H. Silverman, *Lower bound for the canonical height on elliptic curves*, Duke Math. J. **48** (1981), no. 3, 633–648. MR630588 (82k:14043) ↑2

[Sil83] Joseph H. Silverman, *Heights and the specialization map for families of abelian varieties*, J. reine angew. Math. **342** (1983), 197–211, DOI 10.1515/crll.1983.342.197. MR703488 (84k:14033) ↑12.3.1

[Sil86] Joseph H. Silverman, *Heights and elliptic curves*, Arithmetic geometry (Storrs, Conn., 1984), Springer, New York, 1986, pp. 253–265. MR861979 ↑6.1.4

[Sil92] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1992. Corrected reprint of the 1986 original. MR95m:11054 ↑5.1

[Sil09] Joseph H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009. MR2514094 (2010i:11005) ↑3.1, 6.2

[SW13] William Stein and Christian Wuthrich, *Algorithms for the arithmetic of elliptic curves using Iwasawa theory*, Math. Comp. **82** (2013), no. 283, 1757–1792, DOI 10.1090/S0025-5718-2012-02649-4. MR3042584 ↑6.4

[ST95] C. L. Stewart and J. Top, *On ranks of twists of elliptic curves and power-free values of binary forms*, J. Amer. Math. Soc. **8** (1995), no. 4, 943–973, DOI 10.2307/2152834. MR1290234 (95m:11055) ↑a

[Tat74] John T. Tate, *The arithmetic of elliptic curves*, Invent. Math. **23** (1974), 179–206. MR0419359 (54 #7380) ↑3.1

[Tat95] John Tate, *On the conjectures of Birch and Swinnerton-Dyer and a geometric analog*, Séminaire Bourbaki, Vol. 9, Soc. Math. France, Paris, 1995, pp. Exp. No. 306, 415–440. MR1610977 ↑12.2

[TŠ67] D. T. Tĕĭt and I. R. Šafarevič, *The rank of elliptic curves*, Dokl. Akad. Nauk SSSR **175** (1967), 770–773 (Russian). MR0237508 (38 #5790) ↑1, 12.5

[Ulm02] Douglas Ulmer, *Elliptic curves with large rank over function fields*, Ann. of Math. (2) **155** (2002), no. 1, 295–315, DOI 10.2307/3062158. MR1888802 (2003b:11059) ↑1, 3.1, 12.5

[Vat03] V. Vatsal, *Special values of anticyclotomic L-functions*, Duke Math. J. **116** (2003), no. 2, 219–261, DOI 10.1215/S0012-7094-03-11622-1. MR1953292 (2004e:11069b) ↑12.4

[VE10] Akshay Venkatesh and Jordan S. Ellenberg, *Statistics of number fields and function fields*, Proceedings of the International Congress of Mathematicians. Volume II, Hindustan Book Agency, New Delhi, 2010, pp. 383–402. MR2827801 (2012h:11160) ↑1.1.4, 4, 4.1.1

[Wal81] J.-L. Waldspurger, *Sur les coefficients de Fourier des formes modulaires de poids demi-entier*, J. Math. Pures Appl. (9) **60** (1981), no. 4, 375–484 (French). MR646366 (83h:10061) ↑a, 6.4.4

[WS17] Yinghui Wang and Richard P. Stanley, *The Smith normal form distribution of a random integer matrix*, SIAM J. Discrete Math. **31** (2017), no. 3, 2247–2268. MR3706911 ↑7.2.2

[Wat08a] Mark Watkins, *Some heuristics about elliptic curves*, Experiment. Math. **17** (2008), no. 1, 105–125. MR2410120 (2009g:11076) ↑b, c, d, 5.1.3, 6, 6.1.2, 6.2, b

[Wat08b] Mark Watkins, *On elliptic curves and random matrix theory*, J. Théor. Nombres Bordeaux **20** (2008), no. 3, 829–845 (English, with English and French summaries). MR2523320 (2010f:60021) ↑3.4

[Wat⁺14] Mark Watkins, Stephen Donnelly, Noam D. Elkies, Tom Fisher, Andrew Granville, and Nicholas F. Rogers, *Ranks of quadratic twists of elliptic curves*, Publ. math. de Besançon **2014/2** (2014), 63–98 (English, with English and French summaries). ↑b, 8.3

[Wat15] Mark Watkins, *A discursus on* 21 *as a bound for ranks of elliptic curves over* **Q***, and sundry related topics*, August 20, 2015. Available at `http://magma.maths.usyd.edu.au/~watkins/papers/DISCURSUS.pdf`. ↑b, d

[dW98] Benjamin M. M. de Weger, *A + B = C and big* Ш*'s*, Quart. J. Math. Oxford Ser. (2) **49** (1998), no. 193, 105–128, DOI 10.1093/qjmath/49.193.105 (English). MR1617347 (99j:11065) ↑6, 6.2, 6.2.2

[Wil06] Andrew Wiles, *The Birch and Swinnerton-Dyer conjecture*, The millennium prize problems, Clay Math. Inst., Cambridge, MA, 2006, pp. 31–41. MR2238272 ↑6.4

[Won01] Siman Wong, *On the density of elliptic curves*, Compositio Math. **127** (2001), no. 1, 23–54, DOI 10.1023/A:1017514507447. MR1832985 (2002d:11066) ↑1

[Woo17] Melanie Matchett Wood, *The distribution of sandpile groups of random graphs*, J. Amer. Math. Soc. **30** (2017), no. 4, 915–958. MR3671933 ↑7.2.3

[Woo19] Melanie Matchett Wood, *Random integral matrices and the Cohen-Lenstra heuristics*, Amer. J. Math. **141** (2019), no. 2, 383–398. ↑4.2

[You06] Matthew P. Young, *On the non-vanishing of elliptic curve L-functions at the central point*, Proc. London Math. Soc. (3) **93** (2006), no. 1, 1–42, DOI 10.1017/S0024611506015760. MR2235480 (2007h:11075) ↑6.3.4

Department of Mathematics, University of Michigan, Ann Arbor, MI, USA
E-mail address: jmypark@umich.edu
URL: www-personal.umich.edu/~jmypark/

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA
E-mail address: poonen@math.mit.edu
URL: http://math.mit.edu/~poonen/

Department of Mathematics, Dartmouth College, 6188 Kemeny Hall, Hanover, NH 03755, USA
E-mail address: jvoight@gmail.com
URL: http://www.math.dartmouth.edu/~jvoight/

Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53705, USA
E-mail address: mmwood@math.wisc.edu
URL: http://www.math.wisc.edu/~mmwood/