AVERAGE RANKS OF ELLIPTIC CURVES (VERY ROUGH DRAFT)

BAUR BEKTEMIROV, BARRY MAZUR, WILLIAM STEIN AND MARK WATKINS

1. Introduction

Suppose you are given an algebraic curve C defined over the rational number field, defined, let us say, as the locus of zeroes of a polynomial in two variables, $f(x, y)$ with rational coefficients. Suppose you are told that C has at least one rational point, i.e., that there is a pair of rational numbers (a, b) such that

$$f(a, b) = 0.$$

How likely is it that C will have infinitely many rational points?

Such a question, on the one hand, clearly touches on a fundamental issue in diophantine geometry, and on the other, is somewhat meaningless until it is made more precise and appropriately organized. The question we have just asked has distinctly different features, when considered for each of the three basic "types" of algebraic curves: curves of (geometric) genus 0, 1, and > 1. Curves of genus 0 possessing a rational point *always* have infinitely many rational points (an easy fact); curves of genus > 1 never do (a hard fact: Faltings' theorem).

This leaves curves of genus 1 as the interesting case, since some of them, like

$$X^3 + Y^3 = 1,$$

only have finitely many rational points (two, in this instance) and others of them, like

$$Y^2 + Y = X^3 - X,$$

have infinitely many.

Of course, if we are to try to extract an actual number between 0 and 1 that will describe "the" probability that a curve of genus 1 possessing at least one rational point has infinitely many, we have to be precise about exactly which curves we want to count, and how we propose to "sort" them. Let us agree, then,

- to deal only with the smooth projective models of the curves of genus one possessing a rational point (these being precisely the *elliptic curves* defined over $\mathbb{Q}$),
- to count their isomorphism classes over $\mathbb{Q}$, and
- to list them in order of increasing conductors, banking on the theorem that tells us that there are only finitely many isomorphism classes of elliptic curves over $\mathbb{Q}$ with any given conductor.

Here, then, is our question. Does

$$P(X) := \frac{\#\{\text{elliptic curves of conductor } \leq X \text{ with infinitely many rational points}\}}{\#\{\text{elliptic curves of conductor} \leq X\}}$$

converge as X tends to ∞, and if so, what is the limit

$$P := \lim_{X \to \infty} P(X)?$$

In this way we have made our initial question precise:

INTENDED FOR PUBLICATION IN THE BULLETINS OF THE AMS

What is $P =$ the *probability that an elliptic curve has infinitely many rational points*?

It is extraordinary how much vacillation there has been in the past three decades, in the various guesses about the answer to this—clearly basic—question. The subject of this paper is to discuss aspects of this drama. Its see-saw history, involving a network of heuristics and conjectures, and massive data that seemed not to offer much comfort to the conjecturers comes in four parts.

(1) **The minimalist conjecture.** The "classical" Birch-Swinnerton-Dyer conjecture would tell us that the probability P described by our question is at least $1/2$. The reason for this is the phenomenon of *parity*: elliptic curves can be sorted into two classes, those of **even parity**, where the "sign in the functional equation of their L-function" is $+1$, and those of **odd parity**, where the "sign" is -1. The probability that an elliptic curve is of even parity is $1/2$, and the same–of course—for odd parity. A consequence of the Birch-Swinnerton-Dyer conjecture is that *all* elliptic curves of odd parity have infinitely many rational points. This is why no one doubts that the probability P described above is $\geq 1/2$.

It had long been something of a folk conjecture that P is *exactly* $1/2$—let us call this the **minimalist conjecture**. Given the Birch-Swinnerton-Dyer conjecture, an equivalent, and cleaner, way of stating it is as follows:
Minimalist Conjecture: The probability that an elliptic curve of even parity has infinitely many rational points is 0.

This minimalist conjecture might seem appealing purely on the grounds that rational points of elliptic curves are accidental gems of mathematics, and it is hard to imagine that there could be bulk occurrence of these precious accidents—or at least substantially more bulk than is already predicted.

It seems that one cannot find such a minimalist conjecture explicitly in the literature until very recently. Nevertheless, for some particular families of elliptic curves (the "quadratic twist" families; see below) the conjecture that the probability that the elliptic curves in those families have infinitely many rational point is $1/2$ was made by Dorian Goldfeld a quarter of a century ago.

(2) **Contrary numerical data.** The next phase of our story involves the accumulation of numerical data regarding this probability taken over the entirety of elliptic curves, and also over various selected families of elliptic curves. The short description of this data (but see the detailed discussion in the body of our article) is the following. Over every data set accumulated so far, the probability that the elliptic curves in the families being consider have infinitely many rational points is roughly $2/3$, and rather flatly so over the range of conductors involved in the computations; these now include elliptic curves of conductor $< 10^8$.

(3) **A gross heuristic, for special families.** For this phase of our story, to get the most precise results at present we change the data set a bit, and restrict attention to the probability that a member of a *quadratic twist family* of elliptic curves of even parity have infinitely many rational points. In 19?? Peter Sarnak guessed that among the first X members of such a quadratic twist family (ranged in order of increasing conductors) the number of those with even parity and infinitely many rational points is caught between $X^{3/4-\epsilon}$ and $X^{3/4+\epsilon}$ for any positive ϵ and X sufficiently large. This guess, based on consideration of the size of Fourier coefficients of modular forms of half-integral weight, revived the spirits of the minimalist conjecture: if Sarnak's estimate is correct, we would indeed have that the probability

that an even-parity member of a quadratic twist family of elliptic curves has infinitely many rational points is zero.

At this point in our story, there is decided friction between accumulated data which suggests something like 2/3 as probability for the general member to have infinitely many rational points, and a reasoned theoretical expectation, which suggests exactly 1/2 for that probability. Generally, the least we would expect of our data is that it either support our conjectures, or overthrow them. Here there was a somewhat more surprising interrelation between data and conjecture: a kind of truce between them: we believed our guesses, we believed the data, and acknowledged the apparent gap between them.

(4) **A refined heuristic, for special families.** More recently, another twist to this story has developed. The technology of the Katz-Sarnak statistics can been used to give much more precise guess regarding the asymptotics of members of a *quadratic twist family* of elliptic curves of even parity that have infinitely many rational points. For example, for the quadratic twist family ****, the guess is that among the first X members of this family, the number of those with even parity and infinitely many rational points is asymptotic to

$$F(X) := c \cdot X^{3/4} \log(X)^{11/8}$$

with $c \sim \dots$

On the one hand, this is a sharpening of the prior heuristic, for $F(X)$ is comfortably sandwiched between $X^{3/4\pm\epsilon}$. On the other hand, one may be in for a surprise if one actually plots the graph of the function $F(X)$. Here it is:

PLOT OF $c \cdot X^{3/4} \log(X)^{11/8}$ IN THE RANGE $[10^2, 10^8]$

The striking aspect to this graph is how "linear" it looks over the wide range $[10^2, 10^8]$. Indeed, if $F(X)$ were replaced by a linear function with roughly the slope that appears on the page, it would be predicting something closer to 2/3 than 1/2 for the probability that the members of this family have infinitely many rational points.

Here, roughly speaking, is where the story is, at present, as we will explain in detail in the body of this article. The curious last phase of it, focussing on *special families*, makes it seem now that data for these families is (a) more closely adhering to the refined guess than one might expect, even for relatively small values of the conductor, and (b) our refined guess predicts an asymptotic behavior that is far from linear, but within the currently attainable range is so close to linear, that the numerical evidence elucidating these phenomena (even the very large data sets that computers have amassed) seem indecisive when it comes to convincingly distinguishing between such gross questions as: is the probability closer to 1/2 or to 2/3?

But, of course, our story will continue. We would hope for

- as precise a refined heuristic that covers the full range of elliptic curves, and not just quadratic families,
- an extension of the numerical computation to conductors $< 10^{10}$, a range where we may begin to see some subtle differences between the graph of $F(X)$ and a linear function,
- a conceptual understanding of how to obtain—by more unified means—this impressive bulk of rational points that we see occurring for even parity elliptic curves.

Acknowledgment. We thank Armand Brumer, Frank Calegari, Noam Elkies, Oisin McGuinness for stimulating conversations. We used PARI, Python, and

SAGE to compute and analyze the data. We thank Bob Guralnik for references to work on counting distribution of Galois groups.

2. Mordell-Weil Ranks

An *elliptic curve* E over the rational numbers $\mathbb{Q}$ is a projective nonsingular curve defined by an equation of the form

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{1}$$

Mordell proved in 192? that the *Mordell-Weil group* $E(\mathbb{Q})$ of rational points on E is a finitely generated abelian group. The *rank* of E is the number of free factors of the group $E(\mathbb{Q})$. For any $r = 0, 1, 2, \ldots$ the question we now may ask is: what percentage of elliptic curves (nested according to size of conductor) have rank r? More correctly, we should ask: do these percentages exist, and if so what are they?

> "The opinion had been expressed that, in general, an elliptic curve might tend to have the smallest possible rank, namely 0 or 1, compatible with the rank parity predictions of Birch and Swinnerton-Dyer. We present evidence that this may not be the case. [...] This proportion of rank 2 curves seemed too large to conform to the conventional wisdom." – Brumer and McGuinness [BM90]

In [BM90], Brumer and McGuinness considered 310716 curves of prime conductor $\leq 10^8$. In this paper we discuss our extension of their data, and answer in the affirmative that there is a similar pattern for composite conductor $\leq 10^8$, and for prime conductor $\leq 10^{10}$. More precisely, we consider 136832795 curves of all conductors $\leq 10^8$, and 11378911 curves of prime conductor $\leq 10^{10}$. The results of our rank computation are similar to those of Brumer and McGuinness, which appear to suggest that if one orders all elliptic curves over $\mathbb{Q}$ by their conductor, then the average rank is bigger than 0.5. However, as discussed in the introduction, the standard conjectures predict that the average rank is 0.5.

3. Background

The equation (1) is equivalent to exactly one of the form $y^2 = x^3 - 27c_4x - 54c_6$, with $c_4, c_6, \Delta = (c_4^3 - c_6^2)/1728 \in \mathbb{Z}$ and for which there is no prime p with $p^4 \mid c_4$ and $p^{12} \mid \Delta$. We call Δ the *minimal discriminant* of E.

The *conductor* of an elliptic curve E over $\mathbb{Q}$ is a positive integer $N = N_E$ that is a measure of the nature of the reduction of the elliptic curve modulo the prime divisors of Δ. For example, a prime $p \geq 5$ divides the conductor N only if there is no way of modifying the defining equation above of E so that when reduced modulo p we obtain an equation over the field $\mathbb{F}_p$ without multiple roots; the maximal power of such a prime p dividing N is 2 and whether it is 1 or 2 is determined by the nature of the *best* reduction of E modulo p, i.e., whether its defining cubic polynomial has a double or a triple root. There is a slightly more involved, but elementary, recipe to give the maximal power of the prime 2 and of the prime 3 dividing the conductor.

The *L-function* $L(E, s)$ of E is a holomorphic function on $\mathbb{C}$ that is defined by counting points on E over $\mathbb{F}_p$ for primes p. See Silverman [Sil92, App. C §16] for more about the conductor and L-series.

4. Ranks of Elliptic Curves

One fairly firm anchor in the study of elliptic curves is a principle that goes under the heading of parity. This principle is still only conjectural, but is amply confirmed numerically in our accumulated data, and we also have theoretical reasons to believe it. The parity principle is that 50% of the members of any of the families of elliptic curves we will be considering, ordered by conductor, have even rank and 50% have odd rank.

In general terms, the minimalist principle proclaims that from the rough viewpoint of percentages, there are as few rational points on elliptic curves as is possible, given the constraint of the parity principle. That is, 50% of the members of any of the families of elliptic curves we will be considering have rank $r = 0$, and 50% have rank $r = 1$, and the remaining ranks $r \geq 2$ account for 0% of the family.

As one thing or another things comes to light in the subject, the minimalist position is sometimes favored, and sometimes not. Who knows? Tomorrow, a Bhargava-like surprise might change the landscape (see Section 7.1). For certain special families of elliptic curves this minimalist conjecture has long been in print, and has had a wild ride in terms of its being believed, and doubted. For example, Goldfeld considered families of quadratic twists of a single elliptic curve.

4.1. **Goldfeld's Conjecture.** Let E be an elliptic curve over $\mathbb{Q}$ defined by an equation $y^2 = x^3+ax+b$. The *quadratic twist* E^D of E by a square-free integer $D \neq 1$ is the elliptic curve defined by $y^2 = x^3+aD^2x+bD^3$. The twist E^D is isomorphic to E over the field $\mathbb{Q}(\sqrt{D})$.

Conjecture 4.1 (Goldfeld, [Gol79])**.** *The average rank of the curves E^D is $\frac{1}{2}$, in the sense that*

$$\lim_{X\to\infty} \frac{\sum_{|D|<X} \operatorname{rank}(E^D)}{\#\{D : |D| < X\}} = \frac{1}{2}.$$

(Here the D in the sum are squarefree.)

There are many conditional and unconditional results in the direction of Goldfeld's conjecture. For a survey of these results, see [RS02, Sil01].

Kramarz and Zagier [ZK87] considered cubic twists of $x^3 + y^3 = 1$ and found in their data that 23.3% of the curves with even rank have rank at least 2, and 2.2% of those with odd rank have rank at least 3. One of the authors [Wat04] and Fermigier have followed up on these computations. Also, Patricia Quattrini (Universidad de Buenos Aires) as part of her thesis work did some extensive calculations of the analytic Ш for the curves $y^2 = x^3 - nx$. As in the Kramarz-Zagier case the percentage of analytic rank ≥ 2 was in the 20% range but did seem to be going down.

4.2. **Brumer-McGuinness.** In [BM90], Brumer and McGuinness found, by thousands of hours of laborious computer search, 311219 curves of prime conductor $\leq 10^8$. For 310716 of these curves they computed the probable rank by a combination of point searches and computation of apparent order of vanishing of L-functions. The following table, which is taken from [BM90], summarizes the rank distribution that they found.

TABLE 1. Brumer-McGuinness Rank Distribution

Rank	0	1	2	3	4	5
$\Delta > 0$	31748	51871	24706	5267	377	0
$\Delta < 0$	61589	91321	36811	6594	427	5
Totals	93337	143192	61517	11861	804	5
Percents	30.04	46.08	19.80	3.82	0.26	

Let $r_\varepsilon(X)$ be the *average rank* of elliptic curves in [BM90] with conductor at most X and discriminant sign ε. They observe that in their data, r_+ steadily climbs to 1.04 and r_- climbs steadily to 0.94.

Finally [BM90, §5] contains a heuristic estimate for the number of discriminants of elliptic curves up to a given bound:

Conjecture 4.2 (Brumer-McGuinness)**.** *We have the following estimates for the number of positive or negative minimal discriminants of absolute value at most D (respectively):*

$$A_{\pm}(X) \sim \frac{\alpha_{\pm}}{\zeta(10)} X^{5/6}$$

where $\alpha_+ = 0.4206\ldots$ and $\alpha_- = \sqrt{3}\alpha_+ = 0.7285\ldots$ are given by the integral

$$\alpha_{\pm} = \frac{\sqrt{3}}{10} \int_{\pm 1}^{\infty} \frac{du}{\sqrt{u^3 \mp 1}}.$$

Brumer and McGuinness say little about where this heuristic comes from, but remark that it suggests a heuristic for prime discriminants that matches very well with their data (see [Wat05]).

Remark 4.3. In the course of our computations we discovered several errors in the Brumer-McGuinness data (available at [BM90]). The 11 curves in Table 2 have rank 0, not rank 2 as claimed [BM90]. There are floating point precision problems. For example, the computation of a certain square of an integer (the "analytic Ш") is often computed to insufficient precision (e.g., the first three computed values in Table 3 are negative). Also, the curve $y^2 + y = x^3 - 10000x + 384900$ inexplicably appears twice in their data.

Table 2. Incorrect Rank

Conductor	**Curve**
75909851	$y^2 + y = x^3 - x^2 - 358395x - 82463721$
83953699	$y^2 + y = x^3 - x^2 - 36x - 437$
84558059	$y^2 + y = x^3 - 133x - 738$
87978139	$y^2 + y = x^3 + x^2 - 24382x - 1473544$
89054453	$y^2 + y = x^3 + x^2 - 622x - 6166$
89662357	$y^2 + y = x^3 - 2842x - 58314$
92437159	$y^2 + xy = x^3 - 36x - 473$
95207909	$y^2 + y = x^3 + x^2 - 3312x - 74478$
96914827	$y^2 + y = x^3 + x^2 + 74x - 382$
99356003	$y^2 + y = x^3 + x^2 - 1442430x - 667272815$
99420619	$y^2 + y = x^3 + x^2 - 354x - 2730$

Table 3. Examples of Precision Problems

Conductor	Curve	Computed Ш
75047633	$y^2 + xy = x^3 - 1563492x - 752604497$	-0.884646
63473153	$y^2 + xy = x^3 - 1322357x - 585400328$	-0.18517
75255689	$y^2 + xy + y = x^3 - 1567827x - 755736175$	-0.084568
87516089	$y^2 + xy = x^3 + x^2 - 1823251x - 948344280$	287.757199

5. Our Tables

Brumer and McGuinness fixed the a_1, a_2, a_3 invariants (12 total possibilities) and then searched for a_4 and a_6 which made $|\Delta|$ small. Instead, we decided to break the c_4 and c_6 invariants into congruence classes, and then find small solutions to $c_4^3 - c_6^2 = 1728\Delta$. Write $c_4^{\times}$ for the least nonnegative residue of c_4 modulo 576, and $c_6^{\times}$ for the least nonnegative residue of c_6 modulo 1728. Connell [Con05, §5.2] has given necessary and sufficient conditions on c_4 and c_6 such that an elliptic curve with those invariants exists. We first need that $c_6 \equiv 3 \pmod 4$ (in which

case it follows that c_4 is odd), or $2^4 \mid c_4$ and $c_6 \equiv 0, 8 \pmod{32}$, and secondly we require a local condition at the prime 3, namely that $c_6 \not\equiv \pm 9 \pmod{27}$. Using this information and that $1728 \mid \left(c_4^3 - c_6^2\right)$, this leads to 288 possible $\left(c_4^\times, c_6^\times\right)$ pairs.

For each of the 288 pairs $\left(c_4^\times, c_6^\times\right)$, we loop over c_4 and c_6, hoping to find all curves with $|\Delta| \leq 10^{12}$. It is only under an effective ABC-conjecture that we would hope to have an upper bound on c_4 to ensure that we would have found all such curves, and even then the bound would be too large. We simply took $c_4 \leq 1.44 \cdot 10^{12}$ in this first step. We throw away all curves whose conductor is composite and $\geq 10^8$, or prime and $\geq 10^{10}$. We also include all curves that are isogenous to a curve with $c_4 \leq 1.44 \cdot 10^{12}$ and $|\Delta| \leq 10^{12}$, and all quadratic twists of these curves with conductor $\leq 10^8$.

Table 4 lists the number of our curves with various properties.

TABLE 4. Number of Curves in Our Tables

Type	Number
Curves with conductor $\leq 10^8$	136832795
Curves with square-free conductor $\leq 10^8$	21826791
Curves with prime conductor $\leq 10^{10}$	11378911
Curves with prime conductor $\leq 10^8$	312435

5.1. **Completeness of the Tables.** We found 312435 curves of prime conductor up to 10^8, which is 1216 more than Brumer and McGuiness found up to that conductor. Some of the extra curves we find (the "nonoptimal Neumann-Setzer curves") have prime conductor p, but discriminant $-p^2$, whereas all the Brumer-McGuinness curves have a discriminant whose absolute value is prime. Also, some curves were missed presumably because of the large size of the coefficients of the curves; e.g., the curve $y^2 + y = x^3 - 332306354x + 2331610926059$ of discriminant -517267 is missing from the Brumer-McGuiness data. That Brumer and McGuinness would miss some curves suggests that we may have as well. See the data of [Elk00] on Hall's conjecture.

Cremona [Cre] used the algorithms of [Cre97] and the modularity theorem of [BCDT01] to find *every* elliptic curve of conductor up to 120000. He found 347312 isogeny classes of elliptic curves of conductor up to 120000. In our computation, we find 279165 isogeny classes of curves of conductor up to 120000, so we miss 68147 isogeny classes. For example, the first conductor where Cremona has a curve and we do not is conductor 174; the curve $y^2 + xy + y = x^3 - 7705x + 1226492$ has discriminant -621261297432576, which is substantially larger than 10^{12}. The conductors up to 500 where we miss an isogeny class are

$$174, 222, 273, 291, 330, 354, 357, 390, 420, 442, 462, 493.$$

Figure 1 is the proportion of the number isogeny classes in our database to the number of isogeny classes in Cremona's database, as a function of log of the conductor. Thus, e.g., the graph indicates that we found about 83% of Cremona's curves of conductor up to 80000.

The rank distribution of Cremona's curves is given in Table 5. Note that there is a small but noticeable bias toward odd rank; this can be explained via heuristics for the relative sizes of plus-and-minus quotients of Jacobians of modular curves.

Andrei Jorza, Jennifer Balakrishna, and one of the authors (Stein) verified that our table of elliptic curves of prime conductor is complete for levels up to 234431 (see [JBS03]). They did this in order to prove that the smallest conductor of an elliptic curve of rank 4 is composite, in contrast to the case for rank 0 (conductor

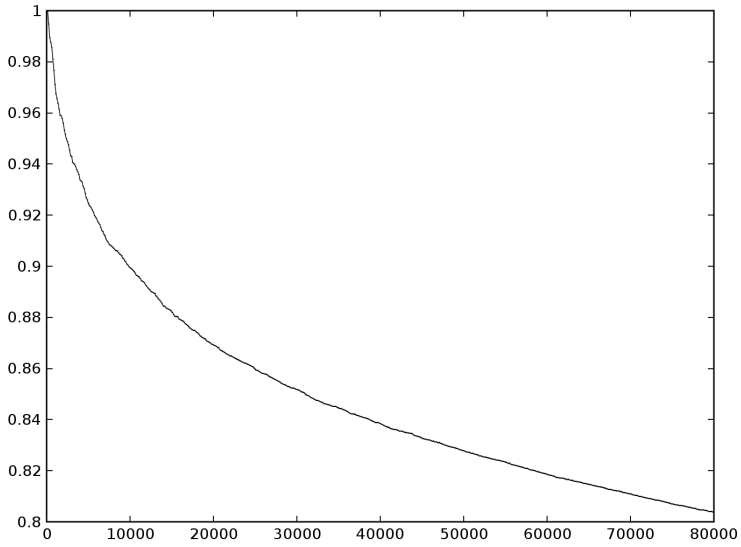

FIGURE 1. Our Proportion of Cremona's Curves of Conductor ≤ 80000

TABLE 5. Rank Distribution of Cremona's Curves

Rank	0	1	2	3
Proportion	0.410	0.507	0.082	0.001

11), rank 1 (conductor 37), rank 2 (conductor 389), and rank 3 (conductor 5077), where in each case the minimal conductor is prime.

5.2. **Future Plans for the Tables.** We intend to compute the full Mordell-Weil groups of all the curves in our tables of rank at least 2. We have almost completed this computation for the curves of prime conductor [CW]. We also plan to add families of elliptic curves to the tables using isogeny parametrizations, and by searching (c_4, c_6) congruence classes where large powers of 2 or 3 divide the discriminant.

6. GRAPHS

Graphs are an efficient way to convey some of the structure of our massive tables of ranks of elliptic curves. At a glance we see that the minimalist principle is contradicted by our data for curves of conductor $\leq 10^8$. For prime conductor $\leq 10^{10}$ the average ranks slightly drop, though only from 0.978 to 0.964. With some imagination, the distribution of rank for prime conductor might appear to support the conjecture that the average rank is 0.5.

(Everywhere below, when we write elliptic curves with property P, we mean elliptic curves in our tables with property P.)

6.1. **Curves Ordered By Conductor.** The average rank of all curves of conductor $\leq 10^8$ is about 0.87. Figure 2 gives average rank as a function of log of the conductor. We created this graph by computing the average rank of curves of conductor up to $n \cdot 10^5$ for $1 \leq n \leq 1000$. Figure 3 graphs the proportion of curves of rank each rank 0, 1, 2, and ≥ 3, as a function of the log of the conductor, all on a single graph. The overall rank proportions are in Table 6.

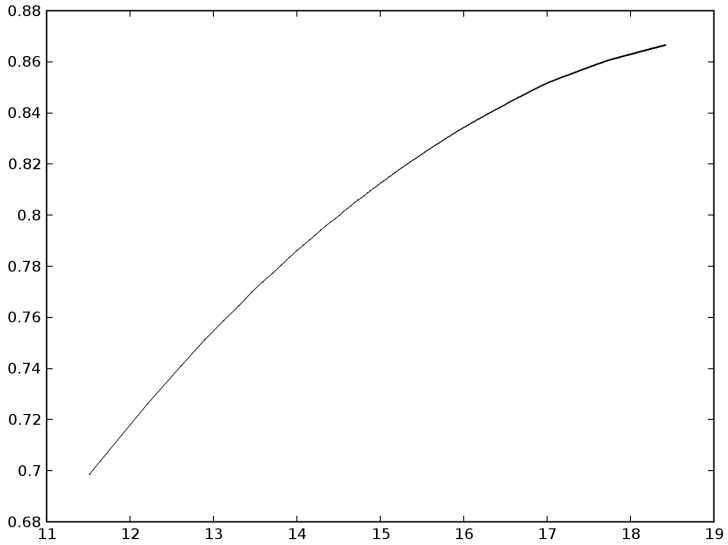

FIGURE 2. Average Rank of Curves of Conductor $\leq 10^8$

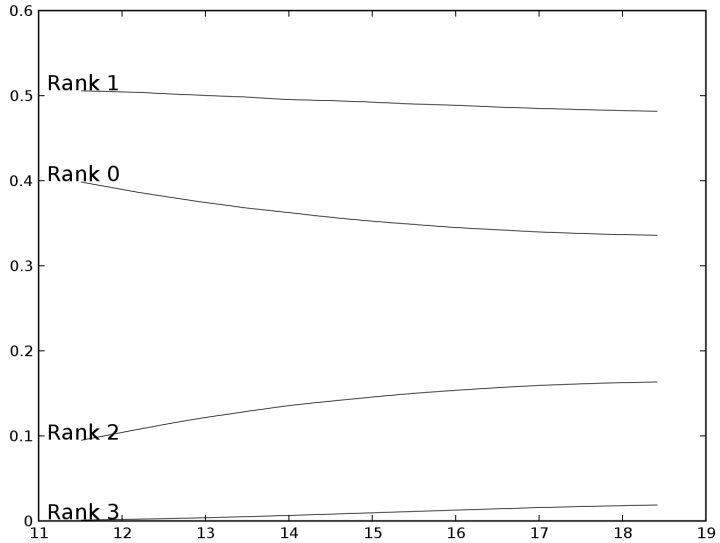

FIGURE 3. Rank Distribution of Curves of Conductor $\leq 10^8$

TABLE 6. Rank Distribution Among Curves of Conductor $\leq 10^8$

Rank	0	1	2	3	≥ 4
Proportion	0.336	0.482	0.163	0.019	0.000

6.2. **Prime Conductor Curves.** The average rank of curves of prime conductor $\leq 10^{10}$ is about 0.964; see Table 7 for the rank distribution. Figure 4 plots the

average rank of prime curves of conductor $\leq 10^{10}$ as a function of the logarithm of the conductor. The curve is solidly decreases, but only *slightly*, since the range of the vertical axis is very small!

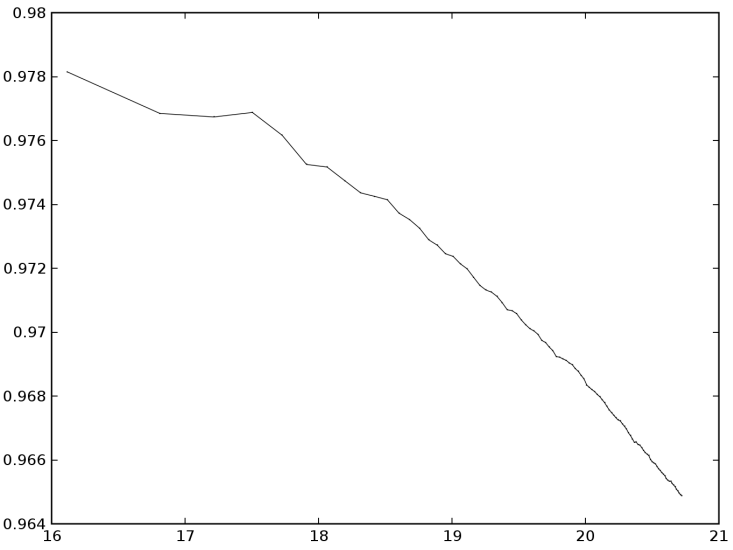

FIGURE 4. Average Rank of Prime Conductor Curves of Conductor $\leq 10^{10}$

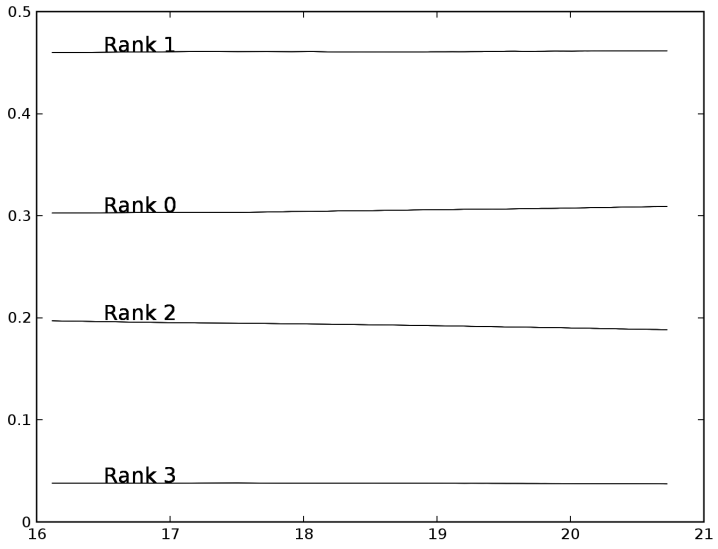

FIGURE 5. Rank Distribution of Prime Conductor Curves of Conductor $\leq 10^{10}$

We also computed 89913 random curves of prime conductor very close to 10^{14}. The average rank is 0.937 and the rank distribution is given in Table 8.

TABLE 7. Rank Distribution for Prime Conductor $\leq 10^{10}$

Rank	0	1	2	3	≥ 4
Proportion	0.309	0.462	0.188	0.037	0.004

TABLE 8. Rank Distribution For Prime Conductors Near 10^{14}

Rank	0	1	2	3	4
Proportion	0.319	0.467	0.176	0.034	0.004

7. Conjectures

See the preprint [Wat05] for more about the following conjectures, which are motivated by lattice point counting, ABC conjectures on average, and heuristics involving the real period.

Conjecture 7.1. *The number of rational elliptic curves whose conductor is less than X is bounded by the function $c_1 X^{5/6} \exp(c_2\sqrt{\log X})$ for some constants c_1 and c_2.*

Conjecture 7.2. *The number of analytic rank 2 elliptic curves with absolute discriminant less than X is bounded by $c_\epsilon X^{19/24+\epsilon}$ for every $\epsilon > 0$. The same is true if we replace absolute value of the discriminant by conductor, but possibly with a different c_ϵ. In particular, asymptotically almost all elliptic curves with even functional equation have analytic rank 0.*

7.1. **Quartic Number Fields.** One must be careful when choosing the coefficients of a fourth degree polynomial if you want a root of that polynomial to generate anything other than a field whose Galois group is S_4. Hilbert's irreducibility theorem provides corroboration of this with a proof that if you rank algebraic numbers of degree 4 by the size of the coefficients of their minimal polynomial (monic, over $\mathbb{Q}$) then 100% of them have Galois group S_4. But let us count quartic fields (rather than algebraic numbers that generate them) nested by the size (absolute value) of their discriminant. Counting field extensions of a given field with a fixed Galois group (i.e., Galois group of their Galois closure) has been the subject of a number of precise conjectures (initially: [CDyDO00], and then successively refined in [Mal02, Mal04]). Barghava's remarkable paper [Bha05], which is further evidence for these conjectures, proves that when you count quartic fields, nested by discriminant, you do not get 100% of them having Galois group S_4.

Bhargava thinks of the problem of counting quartic fields as a problem purely in the Geometry of Numbers, and *proves* the following theorem:

Theorem 7.3 (Bhargava)**.** *When ordered by absolute discriminant,* a positive proportion *(approximately 0.09356) of quartic fields have associated Galois group D_4. The remaining approximately 0.90644 of quartic fields have Galois group S_4.*

Bhargava's surprise theorem leaves open the possibility that there are also surprises in the asymptotic average rank of elliptic curves ordered by conductor.

But forget all questions of asymptotics. Consider only the curves of prime conductor up to 10^{10} in our data. Is there an argument other than just computing ranks for each of the elliptic curves in the tables—is there a pure thought heuristic—that explains why we are witnessing so much Mordell-Weil rank? In a sense, these rational points are both analogous, and not analogous, to the physicist's dark matter. This large mass of rational points for elliptic curves of conductor 10^{10} is palpably there. We aren't in the dark about that. We are merely in the dark about how to give a satisfactory account of their being there, other than computing instances, one after another.

We are, in a word, just at the very beginning of this story.

References

[BCDT01] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*: wild 3-adic exercises*, J. Amer. Math. Soc. **14** (2001), no. 4, 843–939 (electronic). MR **2002d:**11058

[Bha05] M. Bhargava, *The density of discriminants of quartic rings and fields*, Annals of Mathematics **162** (2005), no. 2.

[BM90] A. Brumer and O. McGuinness, *The behavior of the Mordell-Weil group of elliptic curves*, Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 2, 375–382, Data at `http://oisinmc.com/math/310716/`.

[CDyDO00] Henri Cohen, Francisco Diaz y Diaz, and Michel Olivier, *Counting discriminants of number fields of degree up to four*, Algorithmic number theory (Leiden, 2000), Lecture Notes in Comput. Sci., vol. 1838, Springer, Berlin, 2000, pp. 269–283. MR **MR1850611 (2002g:**11148)

[Con05] I. Connell, *The Elliptic Curves Handbook,* `http://www.math.mcgill.ca/connell/public/ech1/`.

[Cre] J. E. Cremona, *Tables of Elliptic Curves,* `http://www.maths.nott.ac.uk/personal/jec/ftp/data/`.

[Cre97] ———, *Algorithms for modular elliptic curves*, second ed., Cambridge University Press, Cambridge, 1997, `http://www.maths.nott.ac.uk/personal/jec/book/`.

[CW] J. Cremona and M. Watkins, *Lattice Distribution of Mordell-Weil Groups of Elliptic Curves*, In Preparation.

[Elk00] N. D. Elkies, *Rational points near curves and small nonzero* $|x^3 - y^2|$ *via lattice reduction*, Algorithmic number theory (Leiden, 2000), Lecture Notes in Comput. Sci., vol. 1838, Springer, Berlin, 2000, pp. 33–63.

[Gol79] D. 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. MR **MR564926 (81i:**12014)

[JBS03] A. Jorza, J. Balakrishna, and W. Stein, *The Smallest Conductor for an Elliptic Curve of Rank Four is Composite*, `http://modular.fas.harvard.edu/rank4/` (2003).

[Mal02] G. Malle, *On the distribution of Galois groups*, J. Number Theory **92** (2002), no. 2, 315–329. MR **MR1884706 (2002k:**12010)

[Mal04] ———, *On the distribution of Galois groups. II*, Experiment. Math. **13** (2004), no. 2, 129–135. MR **MR2068887 (2005g:**11216)

[RS02] K. Rubin and A. Silverberg, *Ranks of elliptic curves*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 4, 455–474 (electronic). MR **MR1920278 (2003f:**11080)

[Sil92] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original.

[Sil01] A. Silverberg, *Open questions in arithmetic algebraic geometry*, Arithmetic algebraic geometry (Park City, UT, 1999), IAS/Park City Math. Ser., vol. 9, Amer. Math. Soc., Providence, RI, 2001, pp. 83–142. MR **MR1860041 (2002g:**11073)

[Wat04] M. Watkins, *Rank distribution in a family of cubic twists*, `http://www.arxiv.org/abs/math.NT/0412427` (2004).

[Wat05] ———, *Some heuristics about elliptic curves*, Preprint (2005).

[ZK87] D. Zagier and G. Kramarz, *Numerical investigations related to the L-series of certain elliptic curves*, J. Indian Math. Soc. (N.S.) **52** (1987), 51–69 (1988). MR **MR989230 (90d:**11072)