Elliptic curves, their companions, and their statistics

Barry Mazur

What is the probability that a cubic plane curve with rational coefficients has infinitely many rational points?

Questions of this type (more exactly formulated, of course) are among the many statistical themes being pursued in the program *Arithmetic Statistics* at MSRI this semester. In this colloquium I'll give some background to help appreciate current work on such problems.

All theories in mathematics have their share of theorems, conjectures and heuristics. But Number Theory, more than the other branches of mathematics thrives on—even depends on—the accumulation of aggregates of numerical data that have to do with numbers themselves. (Of course, no surprise, given its name!) Our program *Arithmetic Statistics* then stands for those aspects of number theory—be it theory or computation—that connect closely with this.

Why study aggregates?

It is curious how *aggregates* rather than *single instances* creeps into our subject even when we aren't looking for statistical trouble.

Here is an example: the primes

$$3, 7, 11, 19, \ldots$$

all lie in an arithmetic progression (they're congruent to 3 mod 4) and we know that there are infinitely many primes in this arithmetic progression (and this can be shown in a Euclid-style way). In the Erdös spirit, I'll offer a \$5 prize for anyone who can manage to provide a proof of the fact that *every* linear form $aX + b$ with a, b relatively prime represents (for $X \mapsto x \in \mathbf{Z}$) at least *one* prime number and such that the proof doesn't actually show that it represents *infinitely many* primes. I think my \$5 is safe, but the point I want to make is that a certain amount of our work is—whether we want it or not—inescapably about "aggregates."

There is also the pleasure one gets in just working in the thick of "many numbers," as is vividly expressed in this letter of Gauss to one of his students (the *italics* are mine):

> Even before I had begun my more detailed investigations into higher arithmetic, one of my first projects was to turn my attention to the decreasing frequency of primes, to which end *I counted the primes in several chiliads* and recorded the results on the attached white pages. I soon recognized that behind all of its fluctuations, this frequency is on the average inversely proportional to the logarithm, so that the number of primes below a given bound n is approximately equal to
>
> $$\int dn/\log(n),$$
>
> where the logarithm is understood to be hyperbolic. Later on, when I became acquainted with the list in Vegas tables (1796) going up to 400031, I extended my computation further, confirming that estimate. *In 1811, the appearance of Chernaus cribrum gave me much pleasure and I have frequently (since I lack the patience for*

a continuous count) spent an idle quarter of an hour to count another chiliad here and there...

Often, in modern number theory, to actually sample a sufficient quantity of data that might allow you to guess even approximate qualitative behavior of the issue you are studying, you may have to go out to very high numbers. For example, there are basic questions about elliptic curves (E.g., *what is the frequency of those possessing two independent rational points of infinite order?*) and if you only test these questions for curves of conductor $< 10^8$, you might be tempted to make guesses that are not only wrong, but qualitatively wrong.

Also we sometimes find that the various members of any of the different aggregates we will be looking at (in this colloquium) tend to directly influence each other. So the most effective way—perhaps the only effective way— of studying them is as a single totality. In all branches of mathematics, we see the advantage to studying as a single ensemble *full collections of likeminded mathematical objects*—e.g., moduli in algebraic geometry, universal classifying spaces in algebraic topology, etc. And there is also the theme that many analytic number theorists allude to when they say (as Iwaniec has said) that the zeroes of different L-functions "know each other." The tension of computing statistics within an aggregate of instances that—in contrast to independent coin tosses—'know each other' is interesting!

All this was an introduction meant to lead to elliptic curves and to hint that we will be paying special attention their arithmetic statistics.

1. Elliptic curves

An **elliptic curve** E over a field K is a projective smooth curve of genus one *with a chosen (K-rational) point, called the "origin"*. Thinking of E as a locus of points

$$(x_0, x_1, \ldots, x_n)$$

in some projective space $\mathbf{P^n}$ we say that such a point is *rational over the field* K if all the (finite) ratios of these coordinates lie in the field K.

Denote by $E(K)$ the "pointed set" of K-rational points of E.

It is a theorem (essentially a corollary of the Riemann-Roch theorem) that allows you to represent any elliptic curve over K as a cubic plane curve (over K, of course) with its origin being its only point (even over the algebraic closure $\bar{K}$) at infinity.

This already is a beautiful piece of mathematics and if you haven't seen it before here is a hint about how you get such a representation, each of these statements being directly obtainable from Riemann-Roch together with the sole fact that the curve we are dealing with has genus one:

- there is only one rational function on E (up to scalars) that has at worst a single pole at one point on E, namely the constant function 1;
- there are two independent rational functions on E having at worst a double pole at the origin and no poles elsewhere: call a choice of the 'new' (i.e., nonconstant) function x;
- there are three independent rational functions on E having at worst a triple pole at the origin and no poles elsewhere: call a choice of the 'new' function with an actual triple pole at the origin y;
- *and* there is a linear relation satisfied by the seven functions

$$1, x, y, x^2, xy, x^3, y^2,$$

all these having at worst poles of order six at the origin and none elsewhere.

In particular we get a mapping of our E onto a plane cubic in x and y (and this mapping turns out to be an isomorphism).

Even more explicitly, when K is a number field (our main focus here), letting $\mathcal{O}_K$ denote the ring of integers of K, we can choose our functions x and y judiciously so that any such E can be given in an affine plane by a cubic equation

$$(*) \qquad y^2 = x^3 + ax + b$$

for constant $a, b \in \mathcal{O}_K$, with its discriminant, $\Delta(a, b) = -4a^3 - 27b^2$, different from zero (this guarantees that E is a *smooth curve*).

Different pairs (a, b) may give rise to isomorphic elliptic curves; for instance, for any element $u \in \mathcal{O}_K$ setting $Y = u^3 y$ and $X = u^2 x$ gives, after clearing terms in the displayed equation, the new cubic equation

$$Y^2 = X^3 + Ax + B$$

where $(A, B) = (u^4 a, u^6 b)$. Here $\Delta(A, B) = u^{12}\Delta(a, b)$.

It is natural then to represent an elliptic curve E by such an affine model **(*)** with a and b *not divisible* by u^4 and u^6 respectively, for any nonunit $u \in \mathcal{O}_K$; equivalently, with minimal absolute value of the norm of its discriminant, among all affine models **(*)** representing E.

For number theory it is quite a good thing that we can represent elliptic curves over a number field K, i.e., *curves of genus one over K having a K-rational point*, in such a clean way. This is not it at all the case if you don't require the curve of genus one to have a K-rational point: it may well be that the only representation of such a curve that is rational over the field K in question is as a curve of very high degree in a projective space (and therefore any projection of such a curve to the plane will represent it only birationally as a curve of high degree with a large singularity locus). This issue will be what is behind the deep questions having to do with what I'll be calling the *companions* to elliptic curves—later in this lecture.

Here are two things one can deduce from this representation **(*)**:

2. Ordering the aggregate of elliptic curves:

We have a natural way of counting the curves!

Theorem 2.1. For any real number X there are only finitely many isomorphism classes (over K) of elliptic curves (over K) with a representation as above such that the absolute value of the norm of its discriminate is less than X; i.e.,

$$|N_{K/\mathbf{Q}}\Delta(a, b)| \ < \ X.$$

That is, we can order the collection of these mathematical objects, in terms of the size of the norms of the discriminants of their "smallest" representations as above[1].

[1]There are some slightly different, and competing, ways of ordering the array of elliptic curves. For example by *conductor*, or as in Manjul Bhargava's work by the size of a natural "height-type" function $H(E) := \max\{|N_{K/\mathbf{Q}}a|^3, |N_{K/\mathbf{Q}}b|^2\}$. The issue of whether the statistics compiled via these different ways of counting are

The proof of this finiteness exhibits the tendency of arithmetic results about elliptic curves to interleave with each other. Recall the formula for the discriminant:

$$\Delta = -4a^3 - 27b^2$$

and if we ask for all integral solutions with given discriminant, we are asking for all integral points on a certain affine model of a certain elliptic curve. This finiteness result already requires significant results in the arithmetic of elliptic curves: for each (rational) integer $N \neq 0$ let $\mathcal{N}$ denote a finite collection of integers in $\mathcal{O}_K$ such that every integer in $\mathcal{O}_K$ with norm equal to N is a twelfth power (of an integer in $\mathcal{O}_K$) times an element of $\mathcal{N}$. We are—in effect—counting the number of integral solutions to the following finite collection of diophantine equations in α and β:

$$-4\alpha^3 - 27\beta^2 = \nu$$

for ν running through the finite set $\mathcal{N}$. Each of these equations are again integral models of *elliptic curves* parametrized by the variables α and β. They have only finitely many integral solutions in $\mathcal{O}_K$.

Moral: the integral solutions over K of these particular elliptic curves "count" the totality of all elliptic curves over K. It's an example of elliptic curves "knowing" other elliptic curves.

That these affine models of elliptic curves have only finitely many integral solutions in $\mathcal{O}_K$, was shown by Siegel (using methods that were ineffective[2] ; effective solutions to this were provided later by Baker; and Faltings famous proof of Mordell-Conjecture also bears on this problem.

---

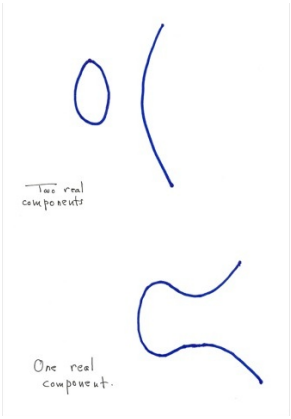
independent of the method used to order these elliptic curves we'll be referring to, in what follows, as the question of *robustness*.

[2]these methods being related to the Mordell-Weil rank of these elliptic curves, a notion which we'll discuss later

The rough number of such elliptic curves is—for X sufficiently large—squeezed between $X^{5/6-\epsilon}$ and $X^{5/6+\epsilon}$ (any $\epsilon > 0$ but presumably starting at larger and larger X).[3]

3. The algebraic group structure on elliptic curves:

What follows is the usual half-minute discussion about elliptic curves that many number theorists give as an introduction to their subject. Its virtue is conciseness, but its great drawback is that it avoids any mention of the dramatic historical evolution of the concept, that had multiple beginnings coming out of elliptic integrals in mechanics and other subjects. But here goes: these are two possible pictures of our plane cubic (over the reals).



The elliptic curve will have two or one real component according as it is representable by an equation

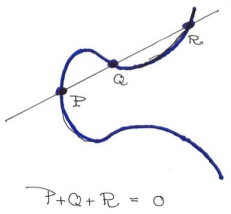
$$y^2 = x^3 + ax + b$$

with $a, b \in \mathbf{R}$ where the cubic polynomial on the right-hand-side of the equation has three real roots, or only one.

Now any smooth cubic hypersurface S in projective space of any dimension has the property that any line L is either contained in S, or else intersects it precisely in three points (if one counts intersections with proper multiplicity, and if one works over $\mathbf{C}$). This gives us a neat binary relationship among (lots of) points in a cubic hypersurface—the *collinear relation*: for any two points P, Q on the hypersurface draw the unique straight line connecting P to Q and and if this line is not

[3]More fun would be to get a precise asymptotic estimate with an error term; this is what Bhargava gets for the ordering of elliptic curves via the size of the function $E \mapsto H(E)$.

entirely contained in S, let R be its "third" intersection point with S. The beauty of this construction is that it is (canonical) and is defined over whatever field K the hypersurface S itself is defined. It is even more beautiful when our smooth cubic hypersurface is an elliptic curve E; i.e., E is of dimension one and for which we are given an "origin" over K. Then the collinear relation is well-defined for any two points[4] P, Q on E and will turn E into a commutative algebraic group by setting the "origin" to be 0, the origin of the group, and stipulating that the sum of any three collinear points be equal to 0. That this rule turns out to yield an associative law, and therefore renders E an algebraic group is one of the many miracles of projective geometry.



4. The Mordell-Weil Theorem

As we are doing number theory, one of the basic questions we would naturally ask about an elliptic curve E over a number field K is: *what is the set of its rational points?* Of course the intelligent thing to do here is to ask for more: since this set has a natural abelian group structure, we would also like to understand it together with this structure—or to put it in other terms: can we actually use this group structure to reduce our task of obtaining all the rational points? The answer is yes and the key result here is:

Theorem 4.1. Let K be a number field and E an elliptic curve defined over K. Then the (abelian) group of points of E that are rational over K—denoted $E(K)$ and called the **Mordell-Weil group of E over K**—is finitely generated. That is,

$$E(K) \quad \simeq \quad \mathbf{Z}^r \oplus \text{a finite abelian group.}$$

Here $r := r(E; K)$ is called the **Mordell-Weil rank** of the elliptic curve, E over K, and the finite group displayed above is called

[4] even if $P = Q$

Mordell-Weil torsion (of E over K). There is a fascinating story to tell about Mordell-Weil torsion. But one also has a (hard to prove) theorem that assures us that we can simply ignore Mordell-Weil torsion if we are considering the rough statistics of the full aggregate of elliptic curves over K. Namely, using an important result of Merel it is not hard to show:

Theorem 4.2. Let X be a real number. Let $\mathcal{N}(E; K, X)$ denote the number of elliptic curves over K that has a representation

$$y^2 = x^3 + ax + b$$

such that the absolute value of the norm of its discriminate is less than X. Let $\mathcal{N}_{\text{tor free}}(E; K, X)$ denote the number among those with torsion-free Mordell Weil group.

We have:

$$\lim_{X \to \infty} \frac{\mathcal{N}_{\text{tor free}}(E; K, X)}{\mathcal{N}(E; K, X)} \;=\; 1.$$

So, if one is going to be considering *averages* over the range of all elliptic curves, one can ignore the ones with torsion, or not. In any event, the real mystery has to do with statistics regarding *Mordell-Weil rank* which is what we will concentrate on from now on. But before that, let us start with a single example:

5. Mordell's question: "What products of two consecutive integers are equal to a product of three consecutive integers?"

The answer to this question, by the way, known to Mordell half a century ago, is that the only such products are $0, 6$, and 210. Of course, the equation whose integral solutions "solves" Mordell's Question is

$$\mathcal{E}: \quad y^2 + y = x^3 - x$$

and this is an affine model, over $\mathbf{Z}$, of an elliptic curve over $\mathbf{Q}$. It has the clean virtue that the Mordell-Weil group of its associated elliptic curve $\mathcal{E}$ is torsion-free of rank one, i.e.,

$$\mathcal{E}(\mathbf{Q}) \simeq \mathbf{Z},$$

and a generator of its Mordell-Weil group is the point $(x, y) = (0, 0)$. (All rational points are generated by the "chord-and-tangent-process" from this "double zero.")

I mentioned at the end of the previous section that this is a "single example," but, following the theme of this colloquium, no single elliptic curve is isolated from the range of all other elliptic curves. For example, in a perfectly natural way, to every point P of our $\mathcal{E}$ above we can associate a pair of elliptic curves $E_P \leftrightarrow E'_P$ that are related by—of all things—a 37-isogeny. Moreover this association is rational over any field—i.e., if the point P is rational over K then so is the pair, and conversely. Even better: every such pair corresponds to a point (give or take the two cusps) on $\mathcal{E}$.

The famous modularity theorem of Wiles and Taylor, Breuil, Conrad, and Diamond (saying that all elliptic curves over $\mathbf{Q}$ are "modular") can be interpreted as saying a similar thing for absolutely *any* elliptic curve over $\mathbf{Q}$: its points are in (a specific) natural correspondence to certain finite subsets of elliptic curves, and as one ranges through all its points, all elliptic curves will occur within this correspondence.

6. A quick course in how one bounds Mordell-Weil rank

For any elliptic curve E over a number field K we have had, for the past eighty years or so a (proved) algorithm for providing an upper bound for the Mordell-Weil rank, $r(E, K)$ of E over K.

This algorithm focusses, more specifically, on finding an upper bound for the number

$$r_p = r_p(E; K) := \dim_{F_p}\{E(K)/pE(K)\}$$

(i.e., r_p is the dimension of the $\mathbf{F}_p$-vector space $E(K)/pE(K)$, or equivalently, $\log_p |E(K)/pE(K)|$.)

It has been known (by an elegant height argument) since the thirties that finiteness of r_p (for any one prime p) implies that $E(K)$ is finitely generated. Clearly, then, a finite bound for r_p will, in turn, bound $r(E, K)$ since

$$r_p(E; K) = r(E; K) + \epsilon_p(E; K)$$

where

$\epsilon_p(E; K) :=$ dimension of the $\mathbf{F}_p$ vector space of $p-$torsion of $E(K)$.

(Note that $\epsilon_p(E; K) \leq 2$.) So, how to bound these r_p's?

Happily, for each prime p an effective method of producing a certain number $s_p := s_p(E;K)$ called the p-**Selmer rank** such that two things are known to be true and a third important feature of $s_p(E;K)$ is conjectured:

(1) $r_p(E;K) \leq s_p(E;K)$;
(2) the difference, $s_p - r_p$, has an arithmetic algebraic geometric interpretation, important in its own right (this interpretation will be related to the existence of curves of genus one over K that we'll be calling *companions* to E and will be describing presently, and
(3) **Conjecture (derived from the Conjecture of Shafarevich-Tate)**
    - For all p the difference $s_p - r_p$ is even; moreover:
    - For all but finitely many p, this difference is zero; i.e.,

$$r_p(E;K) = s_p(E;K).$$

7. Having an algorithm that is conjectured to always work, but not yet proved to always work

Which is where we still are in the general problem of computing Mordell-Weil rank. I used to say that you should spend your days looking for rational points and your nights computing these s_p's for $p = 2, 3, 5, 7, 11, \ldots$ and eventually—the conjecture predicts—you'll hit an equality ($r_p = s_p$) and from then on it will be easy to get all the rational points. (Of course, I also hoped that no one would waste their golden days and nights that way.) But if you do manage to compute the Mordell-Weil rank, it still pays to try to compute the s_p's since the differences $s_p(E;K) - r_p(E;K)$ are telling you something interesting about the arithmetic of E; namely: if any of these differences are nonzero, then there are *companions* of E (not isomorphic to E) in the sense that we will discuss below.

8. Companions

Let us work over the rational field $K = \mathbf{Q}$. What do the following five homogeneous cubic equations have in common?

$$A: \quad 3X^3 + 4Y^3 + 5Z^3 = 0$$
$$B: \quad 12X^3 + Y^3 + 5Z^3 = 0$$

$$\begin{aligned} C: &\quad 15X^3 + 4Y^3 + Z^3 = 0 \\ D: &\quad 3X^3 + 20Y^3 + Z^3 = 0 \\ E: &\quad X^3 + Y^3 + 60Z^3 = 0. \end{aligned}$$

Well, the first thing to say is that these are all smooth genus one (cubic plane) curves over Q, and if you adjoin appropriate cube roots to the rational field, you can make any two, or all, of them equivalent over the larger field. But there is lots more to say.

The curve E has a rational point $(1, -1, 0)$ and if you take this as the "origin" you have an elliptic curve.

The curve

$$A: \quad 3X^3 + 4Y^3 + 5Z^3 = 0$$

is a famous curve, sometimes called the **Selmer Curve**. Selmer in the 1950's showed that A has a rational point over $\mathbf{Q}_p$ the p-adic completion of $\mathbf{Q}$—for *every* prime number p, and it also (visibly) has a real point , but A has no points rational over the number field $\mathbf{Q}$. This was a major moment, for it shows that cubic forms can behave in stark contrast to what happens with quadratic homogeneous forms, where the Hasse Principle (sometimes called the *local-to-global principle*) guarantees that if a quadratic form represents zero over every completion of a number field it represents zero over the number field itself.

This also happens for the three other curves B, C, D. In fact, all five curves are isomorphic to each other over each completion of $\mathbf{Q}$, yet they are all (isomorphically) distinct over $\mathbf{Q}$.

This leads us to the definition of *companion*:

Definition: Let E be an elliptic curve over a number field K. A **companion** to E is a curve C over K such that over every completion K_v of K the curve C (viewed as algebraic curve over K_v) is isomorphic to E over K_v.

So, if an elliptic curve E over K has a companion that isn't isomorphic to E over K we have a phenomenon that violates the spirit of the local-to-global principle (as in the theory of of quadratic forms). Shafarevich and Tate conjecture that any elliptic curve E over K has only finitely many distinct isomorphism classes of companions. This

conjecture was formulated in the sixties, and some two decades passed during which there was not even a single case where it was known to be true.

The first case of an elliptic curve over a number field where one **(a)** obtained the full list of isomorphism classes of its companions, and **(b)** proved the list to be complete, and **(c)** where the list contained more than one item, was the curve

$$E: \quad X^3 + Y^3 + 60Z^3 = 0$$

above over $K = \mathbf{Q}$, and this was achieved by Karl Rubin up in MSRI twenty-five years ago.

9. The Shafarevich-Tate group

I have avoided, till now, any mention of the so-called Shafarevich-Tate group $Sha(E;K)$ the elements of which are the companion curves to E over K with a tiny bit of extra structure (a choice of an isomorphism between the jacobian of the companion curve and E; you might call this an *orientation*[5] an abelian group. This choice of orientation might remind one of the way the ideal class group 'organizes' the set of equivalence classes of binary quadratic forms as an abelian group. The similarity between ideal class groups and Shafarevich-Tate groups,

[5]It is a little theorem that every companion, as defined above, actually *has* an orientation. Here is why: let E' be the jacobian of a companion of the elliptic curve E, so we now have two elliptic curves E, E' over the number field K that are isomorphic curves over every completion of K—and hence they are isomorphic as elliptic curves over every completion of K (i.e., the isomorphism can be made to preserve origins). It is easy to see that E and E' become isomorphic (as elliptic curves) over a finite Galois extension L of K, and hence can be viewed as *twists* of one another, via a cohomology class in $H^1(\mathrm{Gal}(L/K), \mathrm{Aut}(E))$ where $\mathrm{Aut}(E)$ is the group of automorphisms of the *elliptic curve* E over L, giving $\mathrm{Aut}(E)$ its natural $\mathrm{Gal}(L/K)$-action. It is also easy to see using Cebotarev that if the Galois action on $\mathrm{Aut}(E)$ is trivial, you win (since the field over which the cocycle twisting E to E' is a coboundary is cyclic, and the cocycle is a coboundary locally at every completion of K). So we are reduced to the case where the Galois action on $\mathrm{Aut}(E)$ is nontrivial, and by what we've just said, we see that E and E' become isomorphic over the splitting field of that Galois action. But this splitting field (if different from K) is quadratic (and can only be different from K in the cases where $j = 0$, or $j = 12^3$). Again using Cebotarev, together with the fact that E and E' are locally isomorphic at all completions tells us that the the cocycle twisting E to E' is a coboundary. QED.

though, goes only so far—since it has been known since the 19th century that ideal class groups are all finite, while it is still unknown, in general, whether all Shafarevich-Tate groups are finite[6]

Conjecture (Shafarevich and Tate): $Sha(E;K)$ is finite for any E over any number field K.

Note: In the example case of

$$E: \quad X^3 + Y^3 + 60Z^3 = 0$$

discussed in the previous section, Rubin showed that $Sha(E;\mathbf{Q}) \simeq \mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$. The five companions listed in that section correspond to the five pairs $\{x, -x\}$ of elements in $Sha(E;\mathbf{Q})$.

Having introduced this Shafarevich-Tate group, one can be more specific in explaining the nature of the p-Selmer ranks described above; namely, the relationship between $E(K)$, $Sha(E;K)$, and the p-Selmer rank $s_p(E;K)$ is given by the equation:

$$s_p(E;K) = r_p(E;K) + \dim_{F_p}\{Sha(E;K)/p \cdot Sha(E;K)\}.$$

10. Density questions having to do with rank

Let K be a fixed number field and consider the collection of all elliptic curves defined over K. The most natural 'first question' that is somewhat of a statistical nature that you might ask about Mordell-Weil rank is:

> Does $r(E;K)$ admit a finite upper bound (for fixed K and all elliptic curves over K)?

Here, far from actually having a resolution of this yes or no question, we don't even seem to enjoy a uniform consensus about guesses for what the truth is here, even for the field $\mathbf{Q}$. (There are number theorists who believe yes, and others who believe no.) The following chart, which I got off the web, tabulates world's record large ranks for elliptic curves over $\mathbf{Q}$—so far— with the year of their discovery and the winners.

[6]although it is known that $\dim_{F_p}\{Sha(E;K)/p \cdot Sha(E;K)\}$ is finite for all E, K and primes p.

rank $\geq$	year	Author(s)
3	1938	*Billing*
4	1945	*Wiman*
6	1974	*Penney − Pomerance*
7	1975	*Penney − Pomerance*
8	1977	*Grunewald − Zimmert*
9	1977	*Brumer − Kramer*
12	1982	*Mestre*
14	1986	*Mestre*
15	1992	*Mestre*
17	1992	*Nagao*
19	1992	*Fermigier*
20	1993	*Nagao*
21	1994	*Nagao − Kouya*
22	1997	*Fermigier*
23	1998	*Martin − McMillen*
24	2000	*Martin − McMillen*
28	2006	*Elkies*

Our knowledge, and the precision of our expectations, about densities, however, is somewhat more advanced.

Let us assume that you have chosen a size-function for your collection $E \mapsto \text{size}(E) \in \mathbf{R}$ where the main property you need is that for any real number X there are only finitely many isomorphism classes of E's of size $< X$.

Let $r(E;K)$ and $s_p(E;K)$ be the Mordell-Weil rank, and p-Selmer rank of an elliptic curve E over K.

Here are the basic density questions you might ask:

(1) For a given non-negative number r does

$$\rho(K,r) := \lim_{X\to\infty} \frac{|\{E \text{ with } r(E;K) = r \text{ and size}(E) < X\}|}{|\{E \text{ with size}(E) < X\}|}$$

exist, and, if so, what is it?

(2) For a given non-negative number s does

$$\sigma_p(K,s) := \lim_{X\to\infty} \frac{|\{E \text{ with } s_p(E;K) = s \text{ and size}(E) < X\}|}{|\{E \text{ with size}(E) < X\}|}$$

exist, and, if so, what is it?

Of course, even more fun is to predict the rates of convergence of these limits—i.e. to guess specific bounds on the error terms. This is where the subtle random matrix heuristics come in.

The expectations regarding the answers to these density questions seem to have more consensus than the question of boundedness of rank. Here, then, are the current guesses:

11. Current Guesses, and Theorems, about Mordell-Weil rank density

We expect 50% of the elliptic curves over K (ordered by any of the standard size-functions) to have Mordell-Weil rank 0 and 50% to have rank 1. This was first conjectured by Goldfeld in 1979 at least for families of quadratic twists over $\mathbf{Q}$, and in later years fit in with the various heuristic viewpoints of Katz-Sarnak, and also, with precise bounds on rates of convergence for quadratic twist families (Conrey, Keating, Rubinstein, and Snaith) coming from random matrix heuristics, and for all elliptic curves over $\mathbf{Q}$ (Mark Watkins). This has been referred to as the *minimalist conjecture*[7]. In the terminology of the previous section then, the "minimalist conjecture" is that $\rho(K,r) = 1/2$ if $r = 0, 1$ and $\rho(K,r) = 0$ if $r \geq 2$. As hinted in the introduction above, this conjecture is widely believed—and yet it is difficult to get numerical data that firmly support it! The reason for this is in the nature of the error term that is also predicted (coming from random matrix heuristics). The form that this type of error term takes (it will be slightly different in different contexts) if X is the number of instances counted) is

$$aX^b \log^c(X)$$

[7]The reason for the term "minimalist" is that—from the point of view of densities, these are the smallest possible densities that are compatible with the expected *parity* of Mordell-Weil ranks: i.e., 0 is the smallest even number and 1 the smallest odd.

for specific numbers $b < 1$ (but b close to 1). It is diabolic how the graphs of such functions are so very indistinguishable (to the eye) from the linear function aX, but —of course—from the point of view of densities the difference between $aX^b \log^c(X)$ (for any $b < 1$) and aX is major! This is one of the perils of prediction of qualitative behavior from too little data.

What can be proved?

If the minimalist conjecture is true, then the *average Mordell-Weil rank* when compiled for *all* elliptic curves would be 1/2. This, therefore, is the goal. In 1992 Armand Brumer showed (by analytic means, and conditional on standard conjectures) that the average rank of elliptic curves over $\mathbf{Q}$ is bounded by 2.3.

Recently, as I learned from Manjul Bhargava a few days ago, he and Aren Shankar have established (unconditionally) that the "average rank"[8] over $\mathbf{Q}$ is ≤ 0.99. The method here is via what might be called the "geometry of arithmetic orbits in linear representations of reductive groups." Manjul has hopes that these methods might work not only over $\mathbf{Q}$ but also over any fixed number field K.

But for now, over $\mathbf{Q}$ they prove that:

$$\rho(K;0) \geq 0.075$$

and

$$\rho(K;0) + \rho(K;1) \geq 0.80.$$

A striking further result that Bhargava obtained with Wei Ho is that among elliptic curves possessing one point of infinite order, a subset of positive density has Mordell-Weil rank one (at present this result is only for elliptic curves only over $\mathbf{Q}$).

12. Current Guesses, and Theorems, about statistics regarding p-Selmer ranks

Perhaps the most striking heuristic guess, that follows the spirit—on the one hand—of the Cohen-Lenstra heuristics that predict statistics regarding the size of ideal classes, and on the other hand—of the random matrix heuristics, is a recent idea of Poonen and Rains that would

[8]The quotation-marks here are meant to signal that the *in*-equalities regarding averages that we will be discussing will always mean the $\limsup_{X\to\infty}$ (for upper bounds) or the $\liminf_{X\to\infty}$ (for lower bounds) and if these upper and lower bounds are not equal, no claim is being made that the $\lim_{X\to\infty}$ actually exists.

give a guess about the relative frequency that elliptic curves—sampled from the collection of all elliptic curves—have a given p-Selmer rank. The idea behind this heuristic is surprisingly simple: p-Selmer groups are expressible as the intersection of two maximally isotropic subspaces in a certain vector space endowed with a specific nondegenerate bilinear form. Working this through, Poonen and Rains predict that the probability that their intersection is of dimension r is

$$D_p(r) := \frac{\prod_{j=1}^{r} p/(p^j - 1)}{\prod_{i=0}^{\infty}(1 + p^{-i})}$$

These being expected distributions of densities have the requiste property that their sum is 1.

We have, for example, if $p = 2$ or 101:

$$\begin{pmatrix} r & D_2(r) & D_{101}(r) \\ & & \\ 0 & 0.20971 & 0.49505 \\ 1 & 0.41942 & 0.50000 \\ 2 & 0.27961 & 0.0049510 \\ 3 & 0.079890 & 0.000000485 \\ 4 & 0.010652 & 4.7107E-13 \\ 5 & 0.00068723 & 4.5269E-21 \\ 6 & 0.000021817 & 4.3072E-31 \end{pmatrix}$$

Notice that the bulk of densities are distributed over the first two possible ranks 0 and 1.

What is particularly curious about this collection of numbers

$$r \mapsto D_p(r)$$

is that it is *also* the equilibrium distribution of a simple Markov Process. The relevance of this Markov process to the heuristics we are interested in in this context is not too clear, but it becomes far clearer, if one considers the same p-Selmer rank statistical questions for collections of elliptic curves that are quadratic twists of a given elliptic curve, which we discuss below.

There are two combinatorial features of our numbers

$$r \mapsto D_p(r)$$

(1) (**Even versus Odd frequencies**) The sum of the even densities is equal to the sum of the odd ones; i.e.:

$$\sum_{r \text{ even}} D_p(r) = \sum_{r \text{ odd}} D_p(r) \quad = \quad 1/2.$$

(2) (**Even versus Odd "sizes"**) The sums

$$\sum_{r \text{ even}} D_p(r) \cdot p^r = \sum_{r \text{ odd}} D_p(r) \cdot p^r \quad = \quad p+1.$$

This would suggest the following amazing conjecture.

Conjecture: Let K be a number field. The average "size" of the p-Selmer groups in the collection of all elliptic curves over K is $p+1$, for any prime number p. More generally the average "size" of the N-Selmer group is $\sigma(N) := \sum_{d \mid N} 1$.

Even more amazing is that Bhargava and Shakar prove that this is the case for $K = \mathbf{Q}$ $N = 2, 3, 4, 5$.

13. Elliptic curves that are quadratic twists of a given elliptic curve

Recall that to do statistics on these mathematical objects we have to stipulate two things:

- the collection of objects to be counted, and
- the way in which they are ordered.

Before we get into the basic statistics, we should point out that there is some degree of freedom in the choice of range of our collections. The collection, for example, of elliptic curves given by families of *quadratic twists of a given elliptic curve* has some fascinating features, and deserves to be studied separately. That is, fixing $a, b \in \mathcal{O}_K$ and varying $d \in \mathcal{O}_K - \{0\}$ consider the family

$$dy^2 = x^3 + ax + b.$$

The elliptic curves in this family are all isomorphic over $\mathbf{C}$; they are quadratic twists of one another (in various senses, but most directly:) in the sense that any two of them become isomorphic over some quadratic extension of the base field K.

Note also that modifying d by multiplying by a square in $\mathcal{O}_K$ does change the isomorphism type of the elliptic curve so what is really at issue is a class of elliptic curves indexed by elements in $\mathcal{O}_K - \{0\}$ mod squares.

Here we have various possible useful naturally arising choices of ordering this same collection of objects, and although one (e.g., Dan Kane) can sometimes prove a kind of *robustness*; i.e., that the averages

that are computed via various different orderings are the same,[9] things are a bit delicate.

13.1. **Work of Heath-Brown and Swinnerton-Dyer.** The special case of elliptic curves over $\mathbf{Q}$ with "full 2-torsion" (i.e., the kernel of multiplication by 2 in $E(\mathbf{Q})$ is as large as it can be; namely, isomorphic to the Klein group of order 4) has been studied from the point of view of 2-Selmer statistics by Heath-Brown (1994) for the single family of elliptic curves related to the congruence number problem, and more recently (2008) by Swinnerton-Dyer for many families of elliptic curves over $\mathbf{Q}$. The 2-Selmer rank distributions achieved in this context follows Poonen-Rains heuristics (with a shift of 2 due to rational 2-torsion). Moreover, the nature of this work puts the Markov Process more naturally into this picture.

13.2. **Quadratic twist families of elliptic curves over a given (general) number field with** *"no"* **2-torsion.** Here the elliptic curves are required to be of the form

$$dy^2 = g(x)$$

over a number field K where the Galois group of $g(x)$ over K is S_3. Karl Rubin, Zev Klagsbrun, and I are developing an approach (which has a "Markov Process feel") to unconditionally prove the basic statistics for 2-Selmer ranks for such families—although we must order the members of our family in a certain box-like manner. There are surprises:

- *Disparity:* If E is semistable over K (and has full Galois action on 2-torsion) then the densities of twists of E having even 2-Selmer rank is equal to the density having odd 2-Selmer rank (i.e., we have "parity") *if and only if K has a real place.*
- *K-dependence:* Therefore things can change substantially with changes of base field K. However, the even, and the odd, rank statistics for 2-Selmer rank—separately—follow the Poonen-Rains heuristic.

[9]Of course, *naturally arising* is a key phrase here: one can perversely order infinite collections of objects to mess up things.