DISPARITY IN THE STATISTICS FOR QUADRATIC TWIST FAMILIES

BARRY MAZUR

Very rough notes for a lecture delivered at the MSRI Workshop in Arithmetic Statistics, April 13, 2011. This represents joint work with Karl Rubin and Zev Klagsbrun.

1. The basic question

The type of question we will examine has it roots in a famous result of Heath-Brown on the statistics of 2-Selmer ranks of a specific family of CM elliptic curves over $\mathbf{Q}$ related to the congruent number problem[1]. This is the family

$$E_D : \quad Dy^2 = x^3 - x$$

for positive square-free integers D. The arithmetic of this family answers the question of whether or not D can be the common difference of an arithmetic progressions of squares of rational numbers.

This talk will present some on-going work joint with Karl Rubin and Zev Klagsbrun. The three of us are interested in *rank statistics* for twists of E an elliptic curve over a number field K[2]. We consider arbitrary elliptic curves and arbitrary number fields. I will try to focus on the contrast between statistics in this general context and statistics over $\mathbf{Q}$.

Before we begin in earnest, let me give a sense of what is meant by "disparity" in the title of this lecture. By "twists" we are referring to the quadratic twist family

$$\{E^\chi\}_\chi$$

[1]D.R. Heath-Brown, *The size of Selmer groups for the congruent number problem*, Inv. Math. **111** (1993), 171-195; see also *The size of Selmer groups for the congruent number problem, II.*

[2]We are working on this, even for twists by characters of order p where p is a general prime number despite the fact that this fascinating general question has quite a different flavor, and less immediate application, than the restricted question when $p = 2$. This hour I'll talk only of $p = 2$.

where χ ranges through all quadratic characters of K. Let $|\chi|$ denote the absolute value of the norm (to $\mathbf{Q}$) of the conductor of χ.

We shall be dealing with Selmer ranks, which—for the moment—can just be thought of as useful numbers. More specifically, It is convenient to define something that might be called the *reduced Selmer rank*.

Definition 1.1. If E is an elliptic curve over K, by $r(E;K)$, the **reduced 2-Selmer rank** of E over K, we mean:

$$r(E;K) := \{\text{the } 2-\text{Selmer rank of } E \text{ over } K\} \quad - \quad \dim_{\mathbf{F}_2} E(K)[2].$$

Among the many uses of this number $r(E,K)$ is that it is computable, it is an upper bound for the Mordell-Weil rank of E over K, and conjecturally it has the same parity as that Mordell-Weil rank.

Theorem 1.2. The ratio

$$\frac{|\{|\chi| < X;\ r(E^\chi;K) \ \text{ is odd}\}|}{|\{|\chi| < X\}|}$$

is constant for large enough X.

Note: Here is the format of how this is proved: Let Σ be the set of all places of K dividing $2 \cdot \infty$ or the conductor of E. Let $C(K)$ be the group of quadratic characters of K, and consider the set-theoretic mapping:

$$C(K) \longrightarrow \{\text{even},\ \text{odd}\}$$

which says whether the reduced 2-Selmer rank of E^χ over K is even or odd. This mapping is constant on cosets of the kernel of the homomorphism

$$h : C(K) \longrightarrow \Gamma := \prod_{v \in \Sigma} C(K_v)$$

that sends χ to the product of its local restrictions χ_v for $v \in \Sigma$.

More specifically, given E over K, one can define a function

$$C(K_v) \xrightarrow{f_v} \{\pm 1\}$$

(for $v \in \Sigma$) which is a slightly modified "arithmetic ratio of epsilon-factors" whose definition I omit to give here, but which has the effect that for every quadratic character χ of K, the ranks of the 2-Selmer groups of E^χ and E have the same parity if and only if

$$\prod_{v \in \Sigma} f_v(\chi_v) \ = \ 1 \ \in \{\pm 1\}.$$

Define

$$f : \Gamma \to \{\pm 1\}$$

to be the product:

$$f(\gamma) := \prod_{v \in \Sigma} f_v(\gamma_v)$$

where $\gamma = (\dots, \gamma_v, \dots)$.

Let $C(K, X) \subset C(K)$ be the (finite) subgroup consisting of characters such that the absolute values of the norms of primes dividing their conductors are $< X$. So

$$C(K) = \cup_X C(K, X).$$

Since the target group Γ is finite, once X is large enough, $h(C(K, X)) = h(C(K))$. The limit stabilizes to the ratio

$$\frac{|\{\gamma \in \Gamma; f(\gamma) = \pm 1|}{|\{|\Gamma|}$$

for such values of X (where the sign ± 1 depends—in the evident way—on whether or not the rank of E over K is even or odd).

Define, then,

$$\delta(E, K, \text{odd}) := \frac{1}{2} - \lim_{X \to \infty} \frac{|\{|\chi| < X;\ r(E^\chi; K)\ \text{ is odd}\}|}{|\{|\chi| < X\}|}.$$

and its colleague:

$$\delta(E, K, \text{odd}) := \frac{1}{2} - \lim_{X \to \infty} \frac{|\{|\chi| < X;\ r(E^\chi; K)\ \text{ is even}\}|}{|\{|\chi| < X\}|}.$$

these being called the *odd* and *even* disparities of E over K. Of course:

$$\delta(E, K, \text{odd}) + \delta(E, K, \text{even}) = 0;$$

by the *disparity*,

$$0 \leq \delta(E, K) := |\delta(E, K, \text{odd})| = |\delta(E, K, \text{even})| \leq \frac{1}{2},$$

we mean the absolute value of either of the above. Whatever the disparity is—i.e., the relative frequency of odd to even ranks of the 2-Selmer groups of twists—if the Shafarevich-Tate Conjecture holds we would be getting exactly the same disparity relating odd to even ranks of the Mordell-Weil groups of twists.
If $\delta(E, K) = 0$ we "have parity" in the sense that there are statistically as many odd ranks as even; and if $\delta(E, K) = \frac{1}{2}$ all ranks are odd, or all

ranks are even. Either of these endpoints occur; for example, we show that if K has at least one real place, we "have parity." And it is not hard to find more interesting disparities[3].

Here is a random example of what Zev, Karl, and I show, regarding disparity, in the course of studying full rank statistics of 2-Selmer groups.

Let L be a finite number field extension of $\mathbf{Q}$ of degree d, in which 2 splits completely and 5 is unramified. Form the infinite sequence of number fields $K_n := L(\mu_{2^n})$ for $n = 3, 4, 5, \ldots$, and view the elliptic curve E

$$(50A1) \quad y^2 = x^3 - 675x - 79650$$

over each K_n.

Theorem 1.3.

$$\delta(E, K_n) = \frac{\left(1 - 2^{-(2^{n-1}+1)}\right)^d}{2}.$$

In particular, just dealing with these examples yields a set of achieved disparities that is dense in the full range of possibilities, $[0, \frac{1}{2}]$.

2. Density

Again, by way of introduction, let me formulate a general conjecture regarding the relative averages of Selmer ranks of twists of a general elliptic curve E over a general number field K.

Consider the function

$$\mathcal{D}(Z) := \sum_{n \geq 0} \mathcal{D}_n Z^n \ = \prod_{i=0}^{\infty} \frac{1 + 2^{-i} Z}{1 + 2^{-i}}$$

which has come up in the work of Heath-Brown, and later in that of Swinnerton-Dyer specifically as the stationary distribution for a certain Markov process, and has reappeared most recently as the basis of a heuristic regarding guesses for rank density averages over the range of all elliptic curves over a given number field, as formulated by Poonen and Rains. It also shows up in our work.

The coefficients $\mathcal{D}_n$ are all positive numbers and, setting $Z = 1$ we get that

$$\sum_n \mathcal{D}_n = 1$$

[3]For example we show that if K has no real place, and E is semistable over K then we never "have parity."

so $\mathcal{D}$ is a probability density (a positive measure with mass equal to 1) on the set of natural numbers. Setting $Z = -1$ we get $\sum_n (-1)^n \mathcal{D}_n = 0$ which gives us an equal balance of odd and even densities:

$$\sum_{n \text{ odd}} \mathcal{D}_n = \sum_{n \text{ even}} \mathcal{D}_n = \frac{1}{2}.$$

While we are on this topic, looking ahead, if you evaluate at $Z = 2$ and $Z = -2$ you get:

$$\sum_n 2^n \mathcal{D}_n = \prod_{i=0}^{\infty} \frac{1 + 2^{-i}2}{1 + 2^{-i}} = \prod_{i=0}^{\infty} \frac{1 + 2^{1-i}}{1 + 2^{-i}} = 3$$

and

$$\sum_n (-2)^n \mathcal{D}_n = \prod_{i=0}^{\infty} \frac{1 + 2^{-i}2}{1 + 2^{-i}} = \prod_{i=0}^{\infty} \frac{1 + 2^{1-i}}{1 + 2^{-i}} = 0,$$

respectively. This gives us that

$$\sum_{n \text{ odd}} 2^n \mathcal{D}_n = \sum_{n \text{ even}} 2^n \mathcal{D}_n = \frac{3}{2}$$

which eventually will be linked to "average sizes of 2-Selmer groups of odd and of even rank." The derivative of $\mathcal{D}(Z)$ evaluated at $Z = \pm 1$ will eventually be linked to the "average 2-Selmer (even and odd) rank."

Here is a conjectural statement that generalizes the work of Heath-Brown to arbitrary elliptic curves and number fields.

Conjecture 2.1. (1) Let $n \geq 0$, and let

$$\epsilon = \text{ "even," or "odd"}$$

according to the parity of n. Then the limit described the formula below exists and the formula holds:

$$\big(\frac{1}{2} - \delta(E, K; \epsilon)\big) \cdot \mathcal{D}_n \;=\; \lim_{X \to \infty} \frac{|\{|\chi| < X;\ r(E^\chi, K) = n\}|}{|\{|\chi| < X\}|}.$$

As corollaries of this conjecture (following the discussion above) one would have

Corollary 2.2. Let E be an elliptic curve over K. With the same ordering of χ's as in the statement of Conjecture 2.1 it follows—if that conjecture holds—that the average size of the reduced 2-Selmer groups of quadratic twists of E is 3 (independent of the disparity). Moreover, there is a finite upper bound to the average 2-Selmer rank, and Mordell-Weil rank, of quadratic twists of E.

The project we are currently working on is to write out a proof of a version of this general conjecture however

(1) we work only under the hypothesis that the image of the Galois group of K acting on 2-torsion in E is "full," i.e., the image is all of $\mathrm{GL}_2(\mathbf{F}_2)$, and, more significantly,
(2) we cannot yet manage to prove these limits arranging the quadratic twists χ in order of increasing absolute value of norm of conductor as described above, but rather—at the moment—in a less satisfactory way: in terms of certain increasing boxes, to be described below.

Here are some further qualitative comments about our general methods, before becoming specific.

(1) We use only standard methods: class field theory, global duality, an effective Cebotarev theorem (in either of the standard two strengths: the unconditionally proved theorem, but also if we want to improve some bounds, we formulate results using the conditional estimate based on GRH) and basic arithmetic of elliptic curves.
(2) More specifically, the actual densities we obtain all derive from an understanding of the relative densities of certain "Cebotarev classes" of places in various finite extension fields of K.
(3) For example, of use to us, in the context in which we work, are three distinct Cebotarev classes of "good" places of K related to the S_3-extension that is the splitting field of 2-torsion in E; we call these classes *types* $0, 1$, and 2 below according as $Frob_v$ is of order 3, 2, or 1.
(4) Now, averaging over *many* type 0 places has the effect of smoothing things out a lot, and this is a major piece of our machinery, thanks to which we avoided a certain interesting side-question[4].

    But since I also like the feel of this—no longer necessary—question, let me record what might be the simplest example of it here:

[4]Zev suggested this successful way of skirting such (side-)questions.

(5) Let L/Q be, say, the cyclic (cubic) extension given by the (maximal) real subfield in $\mathbf{Q}(e^{2\pi i/7})$. Fix a generator $\sigma \in \mathrm{Gal}(L/K)$ and a congruence condition $m \subset \mathbf{Z}$ (not divisible by 7) such that every finite prime P of L of degree one with norm congruent to 1 mod m has a generator $\pi = \pi_P \equiv 1 \mod m$ such that π is *uniquely determined* modulo squares in $\mathcal{O}_L^*$ by that congruence condition[5]. Now let p be the primes in $\mathbf{Q}$ ranging through the arithmetic progression for which there is a P of the above sort lying above it and form the "Legendre symbol" $\binom{\sigma(\pi)}{\pi}$; this is dependent only on p and not on π. Taking those primes in the arithmetic progression such that distinguishing between primes such that $\binom{\sigma(\pi)}{\pi} = 1$, or $= -1$ breaks up this arithmetic progression into two classes. We'd like to know the density distribution: we think that it is 50/50. We also think that these classes are *not* Cebotarev classes (so there would not be a direct way of showing such a fact) but have not even been able to prove this. If anyone has any ideas about such questions, we're interested. We thank Heath-Brown for mentioning to us that this question is similar to the question—successfully treated by John Friedlander and Henryk Iwaniec[6]–of how often a prime p (congruent to 1 mod 4) expressible as $a2+b2$ with $a, b > 0$ and b even has the property that the Legendre symbol $\binom{a}{b}$ is 1 or -1. Friedlander and Iwaniec prove that the density distribution is 50/50, but even better, they show that

$$\sum_{p<X} \binom{a}{b} << X^{1-\epsilon}$$

for some small, but positive ϵ. This suggests, of course, that we may be dealing here with *non*-Cebotarev classes of primes, since such a fine upper bound for a Cebotarev class of primes is something we don't seem to have the technology to prove at present(it would follow, though, if one could show that a substrip of the appropriate critical strip for the relevant L-functions were free of zeroes).

[5]I haven't checked but think that $m = 4$ might be enough here.

[6]Friedlander, John; Iwaniec, Henryk (1997), "Using a parity-sensitive sieve to count prime values of a polynomial", PNAS **94** (4): 1054-1058

3. Our initial data

The essential issue has to do with quite finite data. Namely we give ourselves a (fixed) number field K with a continuous homomorphism of G_K to $\mathcal{H}$, the quaternionic group of order 8.

*We will show how this connects to elliptic curves endowed with, in effect, something *very close to* a level-4 structure*[7] *over K.*

If

$$(*) \quad 0 \to \mu_2 \to \mathcal{H} \to T \to 0$$

is the exact sequence with μ_2 the center of $\mathcal{H}$, we will be viewing the quotient $T := \mathcal{H}/C$ as a vector space of dimension two over F_2 with the inherited G_K action,

$$\pi : G_K \to \mathrm{Aut}(T) \simeq \mathrm{GL}_2(\mathbf{F}_2) \simeq S_3.$$

A fortiori, this representation to $\mathrm{GL}_2(\mathbf{F}_2)$ is self-dual.

4. Quadratic spaces

We will be interested in $H^1(K, T)$ and also $H^1(K_v, T)$ for the finite, or real places v of K, noting that there is a symmetric self-pairing

$$H^1(K, T) \times H^1(K, T) \to H^2(K, \mu_2)$$

induced from cup-product and the canonical map $T \otimes T \to \wedge^2 T = \mu_2$. Denote this pairing by angular brackets: $(a, b) \mapsto \langle a, b \rangle$, and note that it is compatible with the (corresponding) symmetric *nondegenerate* local pairings

$$H^1(K_v, T) \times H^1(K_v, T) \to H^2(K_v, \mu_2) = \mathbf{F}_2$$

for all (noncomplex) places v of K. There are a few more key ingredients here. Namely:

(1) Define $H^1_{\mathrm{unr}}(K_v, T) \subset H^1(K_v, T)$ by the exact sequence

$$0 \to H^1_{\mathrm{unr}}(K_v, T) \to H^1(K_v, T) \to H^1(L, T)$$

where L/K_v is the unqiue unramified quadratic extension. Call $H^1_{\mathrm{unr}}(K_v, T)$ the **unramified subspace** of $H^1(K_v, T)$; it is its own complement under the bilinear pairing $\langle\ ,\ \rangle_v$;

[7] Specifically, it determines a particular form over K of the elliptic modular curve attached to the congruence subgroup $\tilde{\Gamma}(4) := \ker\{\mathbf{SL_2}(\mathbf{Z}) \to \mathbf{PSL_2}(\mathbf{Z}/\mathbf{4Z})\}$, this being a curve of genus 0.

(2) We have the connecting map $q : H^1(K,T) \to H^2(K,\mu_2)$ coming from the (*nonabelian*) cohomology long exact sequence derived from the exact sequence **(*)** above. For each v we have the corresponding local maps $q_v : H^1(K_v,T) \to H^2(K_v,\mu_2) = \mathbf{F}_2$. The relation between q and $\langle\ ,\ \rangle$ is given by the formula:

$$\langle a,b\rangle = q(a+b) - q(a) - q(b);$$

i.e., q is the *quadratic function* that gives rise to the symmetric bilinear form $\langle\ ,\ \rangle$. And similarly for the q_v's.

(3) Such an object—a vector space with a quadratic function that gives rise to a quadratic form on it—is called a *quadratic space*.

The product of any finite number of quadratic spaces is again a quadratic space in a natural way. In particular, for any finite set $\mathcal{X}$ of places of K, the product $\prod_{v\in\mathcal{X}} H^1(K_v,T)$ with quadratic function $q_{\mathcal{X}}$ defined as

$$q_{\mathcal{X}}(\dots,h_v,\dots) = \sum_{v\in\mathcal{X}} q_v(h_v)$$

is again a quadratic space.

(4) We say that q is unramified at v if q_v maps the unramified subspace $H^1_{\mathrm{unr}}(K_v,T) \subset H^1(K_v,T)$ to the identity element in $H^1(K_v,\mu_2)$. Then q is unramified at all but finitely many v and (since a global cohomology class is also unramified at all but finitely many v) if $c \in H^1(K,T)$, the formula

$$\sum_v q_v(c) = 0$$

makes sense (since the left hand sum involves only finitely many nonzero elements) and moreover, the equation holds.

Definition 4.1. A subspace $V \subset H^1(K_v,T)$ is a **Lagrangian** subspace—relative to the quadratic form q_v— if V is equal to its own orthogonal complement under $\langle\ ,\ \rangle_v$ and if $q_v(V)$ is the identity element in μ_2.

Note that almost all v have the property , then, that the unramified subspace $H^1_{\mathrm{unr}}(K_v,T) \subset H^1(K_v,T)$ is Lagrangian. By convention (and, in fact, as literally following from the definition) if $H^1(K_v,T) = 0$ then we count 0 as a "Lagrangian subspace."

The basic starting data is the pair (T,q) where the G_K action on T cuts out an S_3-extension $K(T)/K$. If you wish, this is a study of S_3

extensions of number fields, together with a small bit of extra structure embodied in the quadratic map $q : H^1(K,T) \to \mathbf{F}_2$ *and its localizations* $q_v : H^1(K_v,T) \to \mathbf{F}_2$.

5. The full Selmer range for (T,q)

Let Σ be a finite set of places of K containing all places dividing $2 \cdot \infty$ or ramified under the Galois action on $\mathcal{H}$.

Definition 5.1. By Σ-**state** we mean a choice, for each $v \in \Sigma$ of a v-Lagrangian subspace in the corresponding $H^1(K_v,T)$.

Definition 5.2. A **Selmer structure** S on (T,q) is given by

- a choice of a finite set of places Σ_S (containing all places dividing $2 \cdot \infty$ or ramified under the Galois action on $\mathcal{H}$), and

- for every place v of K a choice of a v-Lagrangian subspace
$$H^1_{\mathrm{S}_v}(K_v,T) \subset H^1(K_v,T)$$
such that
  - if $v \notin \Sigma_S$ the v-Lagrangian subspace $H^1_{\mathrm{S}_v}(K_v,T) \subset H^1(K_v,T)$ is the unramified one, but
  - if $v \in \Sigma_S$ there is no restriction on which v-Lagrangian subspace it is.

  We'll call the choice at v the *v-Lagrangian* (or synonymously: the *local condition at* v) for the Selmer structure S. Therefore the set of Selmer structures S with $\Sigma_S = \Sigma$ is in one:one correspondence with the set of Σ-states.

Definition 5.3. The Selmer subgroup
$$H^1_S(K,T) \subset H^1(K,T)$$
attached to a Selmer structure S on (T,q) is the subgroup consisting of those cohomology classes $c \in H^1(K,T)$ that, under specialization to G_{K_v}-cohomology, project to an element in the v-Lagrangian subgroup $H^1_{\mathrm{S}_v}(K_v,T) \subset H^1(K_v,T)$ for every place v of K.

Theorem 5.4. The associated Selmer group, $H^1_S(K,T)$, of any Selmer structure S on (T,q) is a finite dimensional $\mathbf{F}_2$- vector space.

One might want to understand 2-Selmer rank statistics, i.e., the behavior of the function:

$$S \quad \mapsto r(S) := \dim H^1_S(K,T)$$

where S ranges through $\mathcal{S}(T,q) :=$ the set of *all* Selmer structures attached to (T,q).

But our actual interest is, for any specific elliptic curve E over K in the moduli problem attached to (T,q), to consider the 2-Selmer rank statistics for the subset

$$\mathcal{S}(E) \subset \mathcal{S}(T,q)$$

consisting of Selmer structures associated to the quadratic twists, E^χ of E, where χ ranges though all quadratic characters of K (see the discussion in Sections 7, 8 and 9 below).

6. How many choices are there for local conditions of a Selmer structure at v?

Suppose, for example, that v is a place of K not dividing 2 and is a place of good reduction for the elliptic curve E. The number of choices one has for v-Lagrangians depends directly on the dimension of T^{G_v}. For unramified v, dim T^{G_v}, in turn, simply depends on the order of the image of Frobenius at v in $\mathrm{GL}_2(\mathbf{F}_2)$. See Table 1 below as a summary of what we are about to discuss. Say that v (not dividing 2) is of "type" $0, 1$ or 2 depending upon whether dim T^{G_v} is $0, 1$ or 2. Each "type" of place forms a Cebotarev class among the allowed places of K, and under our assumption that the image of Galois is full in $\mathrm{GL}_2(\mathbf{F}_2)$ there are infinitely many places of each type. (That there are infinitely many "type 0" places is crucial for our methods.)

- For the places of "type 0" the local cohomology group $H^1(K_v,T)$ vanishes and therefore qualifies as its own Lagrangian subspace; hence the quotation-marks around the "1" in Table 1.
- For the places of type 1 there are only two Lagrangian, the *unramified* Lagrangian, and one other; hence the $1+1$ listed in the table.

- For places of type 2 (even though we are dealing with sets of very few elements) the structure deserves some discussion: In this case the dimension of $H^1(K_v, T)$ is 4. So the projectivization of this four-dimensional $\mathbf{F}_2$- vector space is $\mathbf{P}^3$ (over $\mathbf{F}_2$) in which the nondegenerate quadratic form q_v cuts out a smooth quadric surface V. Now, any such quadric surface is *bi-ruled*– i.e., there are two families (a priori, possibly conjugate over $\mathbf{F}_2$) of lines in V. Each line defined over $\mathbf{F}_2$ in the quadric V comprises a Lagrangian subspace. But, by hypothesis, the unramified maximal isotropic subspace *is* Lagrangian which implies that each of the families is defined over $\mathbf{F}_2$; consequently, there are six Lagrangian subspaces in all, three for each family. The unramified local condition consists of the unique unramified Lagrangian. Twisting, however, by a quadratic character only moves the local condition *within* the ruling containing the unramified Lagrangian as one of its members; more specifically, then, a v-ramified twist will move the local condition to one of the two "ramified Lagrangians" within the ruling containing the unramified Lagrangian.

To sum up:

- for primes v (of the above sort) of type 0—which we shall also be calling *the set of negligible places*—we have *only one* choice of local condition at v;
- for primes of type 1 once we stipulate whether the Lagrangian we wish to choose is unramified or ramified, the local condition is determined;
- for primes of type 2 there are two possible choices of ramified local conditions.

7. The Selmer structure attached to an elliptic curve

Let E be an elliptic curve over K; let $\mathcal{H}_E$ be the associated Heisenberg group[8] with G_K-action,; let

$$T := \mathcal{H}_E/\text{Center} = E[2];$$

and let q be the quadratic function associated to the G_K-"module" $\mathcal{H}_E$. Fix a finite set Σ of places containing all places of bad reduction for E, together with all places dividing $2 \cdot \infty$ or ramified under the Galois action on $\mathcal{H}$.

[8]This should be given in an appendix ... actually: a pretty long appendix.

The Selmer structure $S_{E,\Sigma} \in \mathcal{S}(T,q)$ attached to E and Σ is given by the following prescription for its local conditions:

(1) We put $\Sigma_S = \Sigma$, and
(2) for all v we choose our Lagrangian subspace $H^1_{\mathrm{S}_v}(K_v,T)$ to be

$$H^1_{\mathrm{S}_v}(K_v,E[2]) = E(K_v)/2E(K_v) \quad \subset \quad H^1(K_v,E[2]) = H^1(K_v,T),$$

where the inclusion in the middle comes from the standard Kummer sequence.

8. Twisting

We now want to discuss twisting our Selmer structures by global quadratic characters χ of K—that is, given a Selmer structure S and a quadratic character χ, we will be interested in producing a new Selmer structure $S(\chi)$ that mimics the change in Selmer structures when we pass from that of some elliptic curve E to its twist E^χ.

The story here is different for each of the four classes of places: the finite collection in Σ_S, and the places outside Σ_S of each 'type" as discussed in the previous paragraph.

(1) For $v \notin \Sigma_S$ of type 0, there's absolutely nothing that can change:–the local condition, $H^1_{S(\chi)_v}(K_v,T)$, as well as the full $H^1(K_v,T)$ is 0.

It turns out to be quite an advantage for us that there is a set of places (of positive density among all places of K) of this sort: among other things we will be "averaging" over twists by characters that are ramified at those places,—noting that we haven't changed things there— to give us control of averages over the more difficult places.

(2) For $v \notin \Sigma_S$ of type 1, there are only two possible v-Lagrangians, the unramified Lagrangian, and a unique ramified one. Since v is not in Σ_S, $H^1_{S_v}(K_v,T)$ is the unramified v-Lagrangian. The recipe giving $H^1_{S(\chi)_v}(K_v,T)$ is as follows: if the character χ is unramified at v, then $H^1_{S(\chi)_v}(K_v,T) = H^1_{S_v}(K_v,T)$ is the unramified v-Lagrangian, and if χ is ramified at v, then $H^1_{S(\chi)_v}(K_v,T)$ is the unique ramified v-Lagrangian.

(3) For $v \notin \Sigma_S$ of type 2 and if χ is unramified at v, then, again, $H^1_{S(\chi)_v}(K_v,T) = H^1_{S_v}(K_v,T)$ is the unramified v-Lagrangian.

(4) For $v \notin \Sigma_S$ of type 2 and χ ramified at v then it will also be the case that $H^1_{S(\chi)_v}(K_v,T)$ is ramified. Since there are only two

ramified v-Lagrangians, to complete the recipe here we need only say which it is ...

(5) The final case, for the finitely many places $v \in \Sigma_S$ it is even a trickier business to say explicitly what $H^1_{S(\chi)_v}(K_v, T)$ is, but, again, given what we are averaging over, we need know nothing more than what we have discussed to obtain the statistics we're looking for.

9. "Arranging" the 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.

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. Fixing $a, b \in \mathcal{O}_K$ and varying $c \in \mathcal{O}_K - \{0\}$ consider the family

$$cy^2 = x^3 + ax + b,$$

or—tucking the c into the left-hand side of the equation, on gets the same elliptic curve from

$$y^2 = x^3 + ac^2x + bc^3.$$

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 c 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.

Let us define a **quadratic twist family of elliptic curves over** K to be given by an elliptic curve E over K together with all its twists $\chi \mapsto E^\chi$ indexed by quadratic Dirichlet characters χ over K.

Here we have various possible useful naturally arising choices of ordering this same collection of objects, and although sometimes one (e.g., Dan Kane) can 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.

Fix an elliptic curve E over a number field K, and Σ a finite subset of the set of places $\mathcal{P}_K$ of K (in practice it will be required to contain the archimedean places, and the places dividing p or the conductor of E).

By the **natural ordering** Let us mean that we arrange the members E^χ of our family by increasing absolute value of the norm (down to $\mathbf{Q}$) of the conductor of χ. There are a number of equivalent way of describing this, e.g., in terms of increasing absolute value of the norm of the discriminant, or the conductor, of E^χ.

In contrast, however, to the natural ordering, our results require a slightly different type of ordering, and we give some hints about this in the next, and last section.

10. Skew-box ordering

By a **skew-box ordering** of our family we mean the following.

(1) First, for integers $1, 2, 3, \dots \nu, \dots$ we give positive-real-valued monotonically increasing functions $\alpha_\nu(X)$ of a positive real variable X; we assume further that for each ν $\alpha_\nu(X)$ tends to infinity with X.

(2) If $\chi \in C(K)$ let $d(\chi)$ be its conductor, and write it as follows:

$$d(\chi) = d_\Sigma(\chi) d_0(\chi) d_1(\chi) d_2(\chi),$$

where we have factored $d(\chi)$ into the part involving places in Σ and the places (outside Σ) of types $0, 1$ and 2.

Definition 10.1. For positive integers j, k define **the skew-box** $B_{j,k}(K, X)$ **with sides** $\{\alpha_\nu\}_\nu$ **and cuttoff** X to be the finite subset of the group $C(K)$ of quadratic characters where

(a) $d_1(\chi) = q_1 q_2 \dots q_{j'}$ is a product of j' places, where $j' \leq j$ and the absolute value of the norm of q_i is $< \alpha_i(X)$, for $i = 1, 2, \dots j'$, and where

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

(b) $d_2(\chi) = q_{j'+1}q_{j'+2}\dots q_{j'+k'}$ is a product of k' places, where $k' \leq k$ and the absolute value of the norm of q_i is $< \alpha_i(X)$, for

$$i = j'+1, j'+2, \dots j'+k',$$

(c) (in contrast to the requirement that we bound the norms of each of the places of types 1 and 2, and take account of how many places of those types there are) we require that the absolute value of the norm of $d_0(\chi)$ is $< \alpha_{j'+k'+1}(X)$.

Note that $C(K)$ is the union of the finite "skew-boxes" $B_{j,k}(K,X)$ as X, j, and k tend to infinity.

Here is our theorem:

Theorem 10.2. Let E be an elliptic curve over K with full Galois action on 2-torsion; that is, the natural homomorphism

$$\mathrm{Gal}(\bar{K}/K) \longrightarrow \mathrm{Aut}\big(E(\bar{K})[2]\big)$$

is surjective. For integers

$$1, 2, 3, \dots \nu, \dots$$

there are explicit positive-real-valued monotonically increasing functions[10] $\alpha_\nu(X)$ of a positive real variable X, each tending to infinity with X, such that defining skew-boxes $B_{j,k}(K,X)$ with sides given by those $\{\alpha_\nu\}_\nu$, we have:

(1) Let $n \geq 0$, and let

$$\epsilon = \text{ "even," or "odd"}$$

according to the parity of n. Then the limit described the formula below exists and the formula holds:

[10]These functions depend on E and K. I won't give the formulas here, but just mention that these are defined "recursively" and come from successively applying the effective Cebotarev Theorem; we have unconditional bounds, and also better bounds conditional on GRH.

$$\big(\frac{1}{2}-\delta(E,K;\epsilon)\big)\cdot\mathcal{D}_n \;=\; \lim_{j+k\to\infty}\lim_{X\to\infty}\frac{|\{\chi\in B_{j,k}(K,X);\; r(E^\chi,K)=n\}|}{|B_{j,k}(K,X)|}$$

where X, j, and k all go to infinity.

As discussed in the context of Conjecture 2.1 a series of corollaries follow:

Corollary 10.3. Let E be an elliptic curve over K with full Galois action on 2-torsion. With the same skew-box ordering of χ's as in the statement of Theorem 10.2 the average size of the reduced 2-Selmer groups of quadratic twists of E is 3 (independent of the disparity). Moreover, there is a finite upper bound to the average 2-Selmer rank, and Mordell-Weil rank, of quadratic twists of E.

TABLE 1. Basic Count

Type	order of $Frob_v$ in $\mathrm{Aut}(T)$	$\dim T^{G_v}$	$\dim H^1(K_v, T)$	# of Lagrangians in $H^1(K_v, T)$
0	3	0	0	"1"
1	2	1	2	1+1
2	1	2	4	1+2