Counting points on elliptic curves

Thomas Oliver and Chris Wuthrich

Abstract
Using relatively elementary terminology, we will discuss a natural question on the number of rational points on an elliptic curve. This will lead us to questions that are linked to the conjecture of Birch and Swinnerton-Dyer.

Let A and B be two integers such that $\Delta = -16 \cdot (4A^3 + 27B^2) \neq 0$. Then the equation

$$y^2 = x^3 + Ax + B \tag{1}$$

is called an elliptic curve. More precisely,

$$E: \qquad Y^2Z = X^3 + AXZ^2 + BZ^3$$

is a smooth projective cubic curve defined over $\mathbb{Q}$ in the projective plane $\mathbb{P}^2$.

Together with its unique point $O = (0:1:0)$ at infinity, it is an elliptic curve. The arithmetic of elliptic curve has attracted lots of interest, partly due to the famous conjecture by Birch and Swinnerton-Dyer. The aim of this text is to present some conjectures and questions formulated with as little technical terminology as possible. This is a comparable to Zagier's article [9], but we develop things in a different direction.

1. Points of bounded height

The set $E(\mathbb{Q})$ of points on the elliptic curve with rational coordinates $(X:Y:Z)$ consists of the one point O and of those of the form $(x:y:1)$ with $(x,y) \in \mathbb{Q}^2$ satisfying the equation (1). Since projective coordinates can be scaled, we can write any $P \in E(\mathbb{Q})$ as $(X:Y:Z)$ with integer X, Y, Z such that no $m > 1$ divides all three. Up to sign this representation is unique. Therefore the quantity

$$H(P) = \max\bigl\{|X|, |Y|, |Z|\bigr\}$$

is a well-defined integer for each $P \in E(\mathbb{Q})$, called the **height** of P.

Let T be a large integer. We define

$$\mathscr{N}(T) = \#\Bigl\{P \in E(\mathbb{Q}) \Bigm| H(P) < T\Bigr\}, \tag{2}$$

which is a finite number.

2. Points modulo integers

Let $Q > 1$ be any integer. We will denote by $\mathscr{M}(Q)$ the number of solutions $(X:Y:Z)$ to

$$Y^2Z \equiv X^3 + AXZ^2 + BZ^3 \pmod{Q}. \tag{3}$$

We count points in the projective plane over ${}^{\mathbb{Z}}/_{Q\mathbb{Z}}$, which means the triples (X,Y,Z) where there is no divisor $m > 1$ of Q dividing all three and scalar multiplication by an integer coprime to Q does not alter the point. We could instead count the number of solutions $(x,y) \in {}^{\mathbb{Z}}/_{Q\mathbb{Z}}$ to the equation

$$y^2 \equiv x^3 + Ax + B \pmod{Q} \tag{4}$$

but we would miss not only one point, but $Q/\prod_{p|Q} p$ points. By the Chinese remainder theorem, $\mathscr{M}$ is a multiplicative function: if Q and Q' are coprime then $\mathscr{M}(Q \cdot Q') = \mathscr{M}(Q) \cdot \mathscr{M}(Q')$.

3. An initial conjecture

We formulate a first conjecture:

Conjecture 1. For each $T > 1$ set $Q = T!$. The sequence

$$\frac{Q \cdot \mathscr{N}(T)^2}{\mathscr{M}(Q)} \tag{5}$$

converges to a positive real number as $T \to \infty$.

This seemingly harmless conjecture is actually a very strong statement and, to be honest, maybe even too much to hope for. We can already note that it would not be true if $\Delta = 0$.

Lemma 1. Let $T > 1$. The fraction

$$\frac{Q}{\mathscr{M}(Q)}$$

is independent of Q as long as $\Delta^2 \mid Q$ and the prime divisors of Q are exactly all primes below T.

Idea of the proof. Let p be a prime smaller than T. Because $\mathscr{M}$ is multiplicative, it is enough to show that $\mathscr{M}(p^k)/p^k$ is independent of k as long as k is large enough for all $p \mid \Delta$. First if $p \nmid \Delta$, then using the multi-dimensional version of Hensel's lemma, one can show that $\mathscr{M}(p^k) = \mathscr{M}(p) \cdot p^{k-1}$ for all $k \geqslant 1$. For $p \mid \Delta$, one can show that $\mathscr{M}(p^k) = c \cdot p^{k-1}$ for some integer c when k is large enough. □

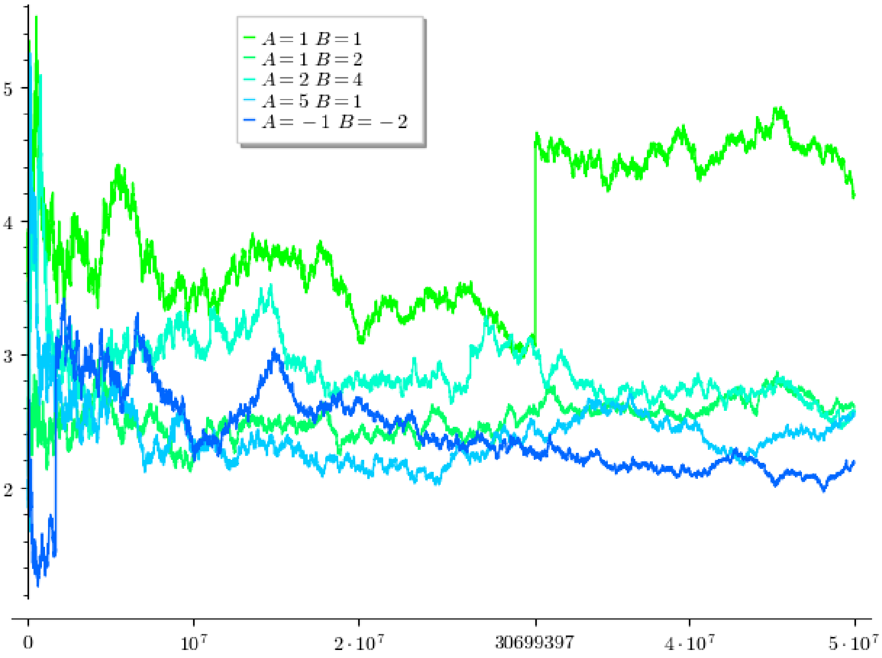


Figure 1. The sequence (5) for a few curves

Certainly the convergence of sequence (5) is very slow for some curves E, but it does not look implausible either. Some examples for a few curves are illustrated in Figure 1. The jump in the graph for the curve $y^2 = x^3 + x + 1$ is not an error. This is due to two points $(x,y) = \left(\frac{43992}{82369}, \pm\frac{30699397}{23639903}\right)$ of height 30699397, which increases $\mathcal{N}(T)$ from 9 to 11 at this T.

We learn from this that the function $\mathcal{N}(T)$ grows slowly and its jumps cause the convergence in Conjecture 1 to be too slow.

4. Counting points of bounded height

In Figure 2, there are the plots of the function $\mathcal{N}(T)$ for some curves. We observe that $\mathcal{N}(T)$ grows like $C \cdot \log^{r/2}(T)$ for some C and some integer r. We will describe a concrete incarnation of this integer r.

At the heart of the reason why elliptic curves stand out among algebraic curves (and why they are so useful in applications like cryptography) is the fact that the set of points $E(\mathbb{Q})$ forms an abelian group with identity element O. The group law is constructed geometrically using what is called the chord and tangent principle. About 100 years ago, Louis J. Mordell proved that for any elliptic curve defined over the rational numbers, a finite set of points suffices to obtain all points in $E(\mathbb{Q})$ using this group operation. The rank $r = \operatorname{rank} E(\mathbb{Q})$ is the minimal number of points needed to create a subgroup of finite index or, equivalently, that $E(\mathbb{Q})$ is a direct product of a finite abelian group, called the torsion subgroup $E(\mathbb{Q})_{\text{tors}}$, and a group isomorphic to $\mathbb{Z}^r$.

Our observation based on the graphs of $\mathcal{N}(T)$ in Figure 2 are confirmed by the following proposition. It is also the first Proposition in [9].

Proposition 2. There is an explicit constant C such that

$$\#\{P \in E(\mathbb{Q}) \mid H(P) < T\} \sim C \cdot \log(T)^{r/2}$$

as $T \to \infty$, where $r = \operatorname{rank} E(\mathbb{Q})$.

The symbol $\sim$ here means as usual that the fraction of the two sides tends to 1 as $T \to \infty$. More precisely, the difference of the two sides can be shown to be $\mathbf{O}\big(\log(T)^{(r-1)/2}\big)$. The constant C is also given in [9], though it missed an extra factor $\frac{2}{3}$:

$$C = \frac{\#E(\mathbb{Q})_{\text{tors}}}{\sqrt{R}} \cdot \left(\frac{2\pi}{3}\right)^{r/2} \cdot \frac{1}{(r/2)!}$$

where $R = \operatorname{Reg}(E) \in \mathbb{R}$ is the so-called regulator of E. If r is odd, we should interpret $(r/2)!$ as $\Gamma(r/2+1)$. In Figure 2, we plotted $\mathcal{N}(T)$ for some curves of rank 0, 1, 2, 3, and 4 against the prediction in this proposition. The names like “11a1” refer to their Cremona labels as in [3].

The proposition also implies that Conjecture 1 is equivalent to

$$\frac{Q \cdot C^2 \cdot \log(T)^r}{\mathcal{M}(Q)} \tag{6}$$

converging to a positive limit as $T \to \infty$ with Q as before. If we take $Q = T!$ or $Q = \prod_{p<T} p$, then Q is

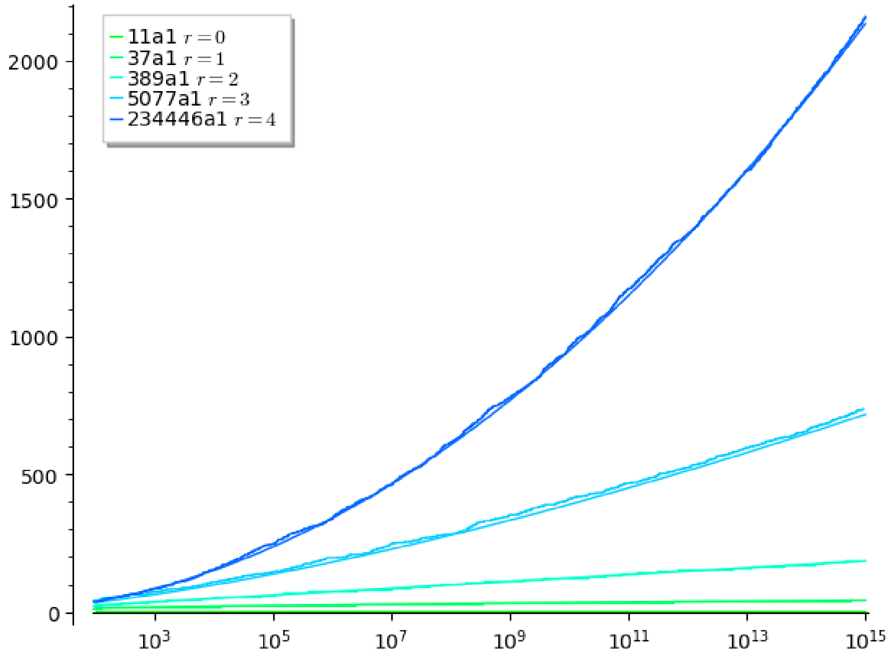


Figure 2. $\mathcal{N}(T)$ for some curves

exponential in T, while we compare it to $\mathcal{N}(T)$ which is logarithmic. This double exponential gap between the two illustrates again how fragile the behaviour of this quotient is. We plot in Figure 3 the graphs for the quotient (6) with the same curves as in Figure 1. In this figure, we also plot the conjectured limit that we are getting to in a moment.

5. The link to the Birch and Swinnerton-Dyer conjecture

Those who are aware of the history of the conjecture made by Bryan Birch and Peter Swinnerton-Dyer will now have recognised the connection between our Conjecture 1 and the rank part of their famous conjecture. As explained in [1], their initial investigations concerned the behaviour of the product

$$\prod_{p<T} \frac{\mathcal{M}(p)}{p}$$

as T increases. They made an initial guess that it grows like $\log(T)^r$. Up to a factor linked to primes with $p \mid \Delta$, this product is equal to $\mathcal{M}(Q)/Q$. They made a better and, in many respects more interesting, conjecture involving the L-function $L(E,s)$ attached to E, which we will avoid in this exposition. Their conjecture says that $L(E,s)$ has a zero of order $r = \operatorname{rank} E(\mathbb{Q})$ at $s=1$ and they gave a precise formula for the leading term of its expansion at $s=1$. See [8].

Goldfeld [4] established the connection between the two versions and, as a consequence, we find the following.

Theorem 3. If Conjecture 1 holds then the rank part of the Birch and Swinnerton-Dyer conjecture holds.

However, it also turns out that much more is true. In [2], it is well explained that Conjecture 1 not only implies the Birch and Swinnerton-Dyer conjecture, it would also imply that the function $L(E,s)$ satisfies the analogue of the Riemann hypothesis in a very strong form. Maybe too strong to believe, but there is also no reason to disbelieve it currently.

6. The period

Number theorists like to split problems into global and local problems. Issues concerned with divisibility by one (or a few) primes, like the term $\mathcal{M}(Q)$, are local, while rational points and their heights, like $\mathcal{N}(T)$, are global. This terminology comes from the construction of the completions of $\mathbb{Q}$. For each prime p, there is a field $\mathbb{Q}_p$, called the field of p-adic numbers. It is the completion of $\mathbb{Q}$ with respect to the topology given by the distance

$$d(x,y) = |x-y|_p = p^{-k} \qquad \text{for } x,y \in \mathbb{Z},$$

where p^k is the largest power of p dividing $x-y$, extended to $\mathbb{Q}$ by $|a/b|_p = |a|_p/|b|_p$. Together with the more commonly known completion $\mathbb{R}$ with respect to the usual absolute value $|\cdot|$, they form all possible completions of $\mathbb{Q}$.

Let $T>1$ and take Q as before. Let $\varepsilon = p^{-k}$ where $p^k < Q \leqslant p^{k+1}$. To say that (4) holds for (x,y) with

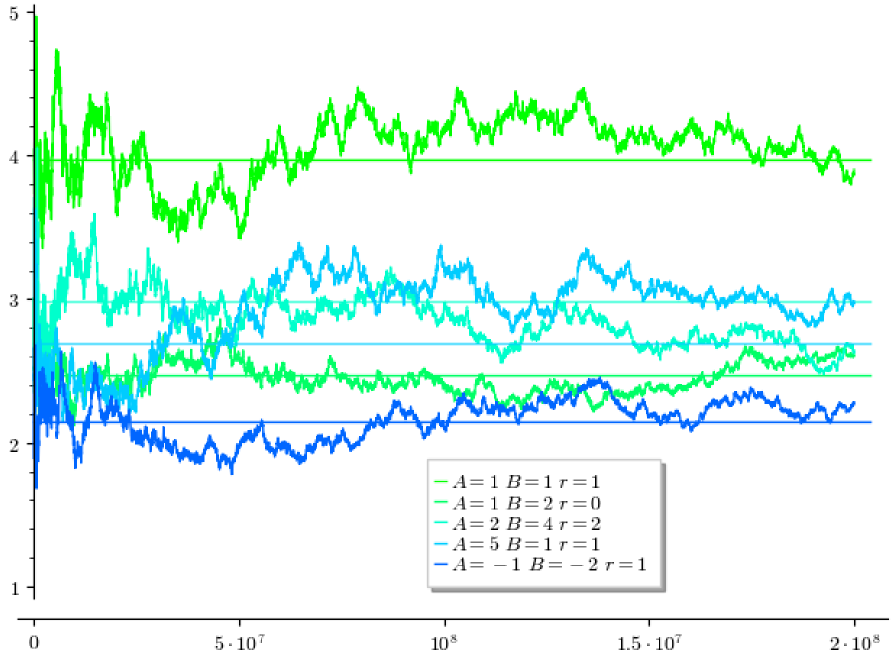


Figure 3. Convergence of the limit of (6) for some curves

integer x, y can now be formulated by asking that

$$\left|-y^2+x^3+Ax+B\right|_p<\varepsilon$$

for all $p<T$. The quantity $\mathscr{M}(p)$ is then linked to the area in $\mathbb{Q}_p\times\mathbb{Q}_p$ satisfying this inequality.

It is therefore natural to look at the equivalent inequality over the real numbers $\mathbb{R}$. Consider $(x,y)\in\mathbb{R}^2$ such that

$$\left|-y^2+x^3+Ax+B\right|<\varepsilon$$

for some small $\varepsilon>0$. Assume that $\Delta<0$, which means that the graph of $E(\mathbb{R})$ is connected containing only (x,y) with x bigger than the unique solution e_1 of $x^3+Ax+B=0$ in $\mathbb{R}$. The area of this part of the plane $\mathbb{R}^2$ is given by

$$2\int_{e_1}^{\infty}\left(\sqrt{x^3+Ax+B+\varepsilon}-\sqrt{x^3+Ax+B-\varepsilon}\right)dx$$
$$=2\int_{e_1}^{\infty}\frac{2\varepsilon}{\sqrt{x^3+Ax+B+\varepsilon}+\sqrt{x^3+Ax+B-\varepsilon}}dx$$
$$=2\varepsilon\cdot\int_{E(\mathbb{R})}\frac{dx}{2|y|}+\mathbf{O}\left(\varepsilon^2\right)$$

The same result holds when $E(\mathbb{R})$ is formed of two connected components. The quantity $\Omega=\int_{E(\mathbb{R})}dx/(2|y|)$ is known as a period of E. One can consider $\Omega\cdot\mathscr{M}(Q)$ to be linked to the area of the subset of $\mathbb{R}^2\times\prod_{p<T}\mathbb{Q}_p^2$ cut out by the inequalities above.

7. Limit

Because Conjecture 1 is linked to the rank part of the Birch and Swinnerton-Dyer conjecture, the limit of the quotient (5) should have something to do with the leading term of the function $L(E,s)$ at $s=1$.

Theorem 4. Suppose Conjecture 1 holds. Then the Birch and Swinnerton-Dyer formula holds if and only if

$$\lim_{T\to\infty}\frac{Q\cdot\mathscr{N}(T)^2}{\mathscr{M}(Q)}=\frac{1}{\sqrt{2}\cdot((r/2)!)^2}\cdot\left(\frac{3e^{\gamma}}{2\pi}\right)^r\cdot\Omega\cdot S.$$

where S is a square integer, which is conjecturally the order of the mysterious Tate-Shafarevich group $Ш(E)$.

Oh, e and γ are both constants due to Euler. In Figure 3 we have plotted the expression (6) against this limit for the curves in question.

It remains to explain what the Tate-Shafarevich group $Ш(E)$ is. It is a torsion abelian group, which is conjectured to be finite. The group $Ш(E)$ can also be viewed as measuring a discrepancy between a local and a global questions. Its elements can be viewed as curves C defined over $\mathbb{Q}$ which become isomorphic to E when considered over any p-adic field as well as over $\mathbb{C}$. This group appears as an obstruction to effectively calculating $E(\mathbb{Q})$ by the method of infinite descent.

Refined counting of points modulo primes

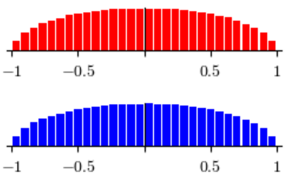
One of the vague arguments given initially for the conjecture is the following: If we have lots of points with integer $(X:Y:Z)$, then they produce lots of points modulo integers, and primes in particular. In other words the larger r, the more often $\mathscr{M}(p)$ should be above average. This is however too simplistic.

But, what is this average? Hasse proved that $p+1-2\sqrt{p} < \mathscr{M}(p) < p+1+2\sqrt{p}$, which suggests $p+1$ as the average. We define

$$a_p = p+1-\mathscr{M}(p),$$

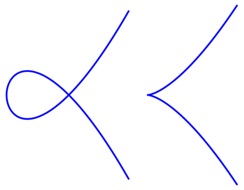
where negative values of a_p indicate an excess of points modulo p.

Originally conjectured by Sato and Tate, the values $a_p/(2\sqrt{p}) \in [-1,1]$ are known to satisfy a precise distribution, which is independent of r. See [6] for a good overview. In particular positive and negative values should appear with the same frequency. Except for the special elliptic curves with extra endomorphisms (complex multiplication), the distribution of $a_p/(2\sqrt{p})$ looks like this: The red histogram on top is for the curve $A = B = 1$ for which $E(\mathbb{Q})$ is infinite and the bottom in blue is for $A = 1$, $B = 2$ which has a finite $E(\mathbb{Q})$. For both we used all primes $p \leqslant 10^7$.



Murmurations

Finally, we present a recently discovered phenomenon related to counting points on elliptic curves. The reduction of an elliptic curve curve at a prime p can be bad in that the reduced curve has a singularity. There are two possibilities: On the left, we have the case of nodal reduction and, on the right, that of cuspidal reduction.



The conductor $N(E)$ of an elliptic curve E defined over $\mathbb{Q}$ is an integer divisible only by the primes of bad reduction. In other words, the conductor may be written as a product

$$N(E) = \prod_{p\text{: bad}} p^{e_p}.$$

More precisely, for a bad prime $p \notin \{2,3\}$, we have $e_p = 1$ (resp. $e_p = 2$) if E has a nodal reduction (resp. cuspidal reduction) modulo p. For $p = 2$ and 3, the recipe is known but more complicated.

The so-called **murmuration** phenomenon refers to the oscillating behaviour of the average value of $a_p(E)$, as a function of p, where E varies over a suitable finite set of elliptic curves [5]. More precisely, we set:

$$M_{\mathscr{E},r}(p) = \frac{1}{\#\mathscr{E}(r)} \sum_{E\in\mathscr{E}(r)} a_p(E),$$

where $r \in \mathbb{Z}_{\geqslant 0}$, $\mathscr{E}$ is a finite set of elliptic curves over $\mathbb{Q}$, and $\mathscr{E}(r)$ is the subset of $\mathscr{E}$ containing its curves of rank r. In Figure 4 below we plot $M_{\mathscr{E},r}(p)$ against p for even r (blue) and odd r (red), where we take $\mathscr{E}$ to be the set all elliptic curves defined over $\mathbb{Q}$ with conductor in the interval $2^{17} < N(E) < 2^{18}$. Most of these curves have either rank 0 or 1 and so the blue dots correspond to curves with a finite $E(\mathbb{Q})$ and the red for infinite $E(\mathbb{Q})$.

In forthcoming work of He, Lee, Oliver, Pozdnyakov, and Sutherland, this oscillating behaviour is shown to hold for elliptic curves with much larger conductor, and also for related arithmetic objects such as modular forms and genus 2 curves. There is some initial progress to the understanding as to why this pattern appears due to Zubrilina, but there is no comprehensive explanation for this phenomenon so far. Since most curves in the interval are of rank either 0 or 1, and because the total average should be 0 by the Sato-Tate distribution, we expect the two waves of murmuration to be complementary. The naive heuristic that curves of rank 1 have more points modulo prime, and hence a_p is more frequently negative, may be the reason that the red wave dips initially for small primes p.

References

[1] B. J. Birch, *Conjectures concerning elliptic curves*, Proc. Sympos. Pure Math., Vol. VIII, Amer. Math. Soc., Providence, R.I., 1965, pp. 106–112.

[2] Keith Conrad, *Partial Euler products on the critical line*, Canad. J. Math. **57** (2005), no. 2, 267–297.

[3] John E. Cremona, *Algorithms for modular elliptic curves*, second ed., Cambridge University Press, 1997, Now available at lmfdb.

[4] Dorian Goldfeld, *Sur les produits partiels eulériens attachés aux courbes elliptiques*, C. R. Acad. Sci. Paris Sér. I Math. **294** (1982), no. 14, 471–474.

[5] Y.-H He, K.-H Lee, T. Oliver, and A. Pozdynakov, *Murmurations of elliptic curves*, arXiv: 2204.10140.

[6] Barry Mazur, *Finding meaning in error terms*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 2, 185–228.

[7] Louis J. Mordell, *On the rational solutions of the indeterminate equations of the third and fourth degree*, Math. Proc. Cambridge Philos. Soc. **21** (1922), 179–182.

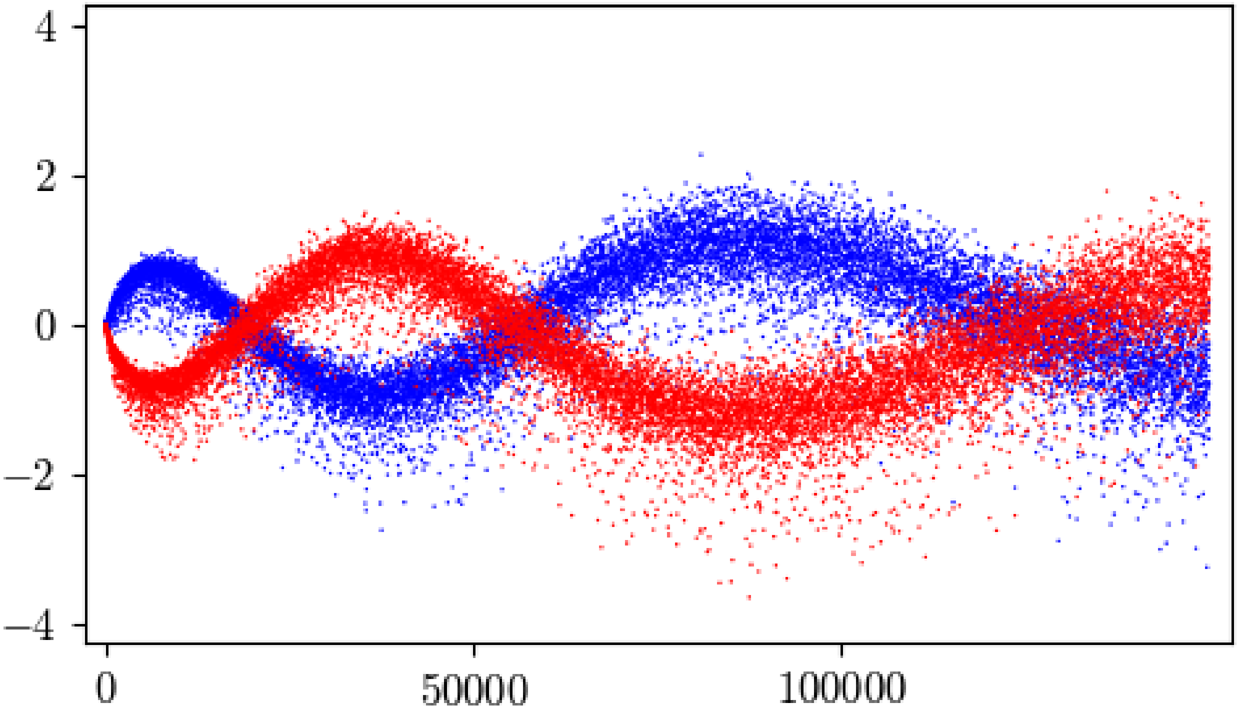


Figure 4. Figure 4: Murmuration with even rank curves in blue and odd rank curves in red.

[8] Andrew Wiles, *Birch and Swinnerton-Dyer Conjecture*, Clay Mathematics Institute.

[9] Don Zagier, *The Birch-Swinnerton-Dyer conjecture from a naive point of view*, Arithmetic algebraic geometry (Texel, 1989), Progr. Math., vol. 89, Birkhäuser Boston, Boston, MA, 1991, pp. 377–389.