# Rational Points on Curves of Genus 2: Experiments and Speculations

Michael Stoll

Mathematisches Institut, Universität Bayreuth
95440 Bayreuth, Germany
Michael.Stoll@uni-bayreuth.de

**Abstract.** We present results of computations providing statistics on rational points on (small) curves of genus 2 and use them to present several conjectures. Some of these conjectures are based on heuristic considerations, others are based on our experimental results.

**Keywords.** Rational points, genus 2

## 1 Introduction

A *curve of genus 2* over $\mathbb{Q}$ is given by an equation

$$C : y^2 = f_6x^6 + f_5x^5 + f_4x^4 + f_3x^3 + f_2x^2 + f_1x + f_0$$

with $f_j \in \mathbb{Z}$, such that $(f_6, f_5) \neq (0, 0)$ and the polynomial on the right does not have multiple roots.

A *rational point* on this curve $C$ is a pair of rational numbers $(\xi, \eta)$ satisfying the equation. In addition, there can be rational points "at infinity", corresponding to the square roots of $f_6$ in $\mathbb{Z}$. We denote the set of rational points on $C$ by $C(\mathbb{Q})$. It is known (Mordell's Conjecture, proved by Faltings) that $C(\mathbb{Q})$ is finite.

We consider curves of genus 2 over $\mathbb{Q}$ as above with $f_j \in \mathbb{Z}$, and of *size* $\max_j |f_j| \leq N$ for given $N$.

**Question.** What can we say about $C(\mathbb{Q})$ as $N$ grows?

- How many rational points are there on average?
- What is the distribution of the number of points?
- What is the largest number of rational points?
- How are the sizes of the points distributed?
- How large can the points get?

In the following, we present some experimental data and some conjectures relating to these questions.

## 2 Heuristics and a First Conjecture

The condition that the point $(\frac{a}{b}, \frac{c}{b^3})$ is on $C$ translates into a linear condition on the coefficients $f_j$:

$$a^6 f_6 + a^5 b f_5 + a^4 b^2 f_4 + a^3 b^3 f_3 + a^2 b^4 f_2 + a b^5 f_1 + b^6 f_0 = c^2$$

The curves satisfying this correspond to points in the intersection of a coset of a 6-dimensional lattice in $\mathbb{R}^7$ with a cube of side length $2N$. Standard estimates then give us the following asymptotic behavior of the average number of points with $x = \frac{a}{b}$:

$$\mathbb{E}_{(a:b)}(N) \sim \frac{\gamma(a:b)}{\sqrt{N}} \qquad \text{as } N \to \infty.$$

with $\gamma(a:b)$ of order $H(a:b)^{-3}$, where $H(a:b) = \max\{|a|, |b|\}$ is the *height* of $x$. We let

$$\gamma(H) = \sum_{\substack{(a:b)\in\mathbb{P}^1(\mathbb{Q}) \\ H(a:b)\le H}} \gamma(a:b) = \gamma - O\Big(\frac{1}{H}\Big)$$

where

$$\gamma = \sum_{(a:b)\in\mathbb{P}^1(\mathbb{Q})} H(a:b) = \lim_{H\to\infty} \gamma(H) \approx 4.79991\,.$$

We obtain for the average number of points with $H(x) \le H$:

$$\mathbb{E}_{\le H}(N) \sim \frac{\gamma(H)}{\sqrt{N}} \qquad \text{as } N \to \infty.$$

Let $\mathbb{E}(N)$ denote the average number of rational points on curves of size $\le N$.

**Corollary.**

$$\liminf_{N\to\infty} \sqrt{N} \cdot \mathbb{E}(N) \ge \gamma\,.$$

This leads to the following conjecture, which essentially says that there is not a large number of unexpectedly large points.

**Conjecture 1.**

$$\lim_{N\to\infty} \sqrt{N} \cdot \mathbb{E}(N) = \gamma\,.$$

## 3 Large Points

If we accept Conjecture 1, then we should expect about

$$(\gamma - \gamma(H)) 2^7 N^{13/2} = O\big(H^{-1} N^{13/2}\big)$$

curves of size $\le N$ that have points of height $\ge H$. So generically, we expect that rational points on a curve of size $N$ will have height $\ll N^{13/2+\varepsilon}$. A more precise version is as follows.

**Conjecture 2.** *Given $\varepsilon > 0$, there is $B_\varepsilon > 0$ and a Zariski-open subset $\emptyset \neq U_\varepsilon \subset \mathbb{A}^7$ such that the rational points on every curve of size $\leq N$ whose coefficient vector is in $U_\varepsilon$ have height $\leq B_\varepsilon N^{13/2+\varepsilon}$.*

It is, however, likely that there are families of curves with larger points. A naive dimension count predicts a family with points of height $\gg N^9$ (maybe not over $\mathbb{Q}$). Still, we can hope that the following is true.

**Conjecture 3.** *There are $\kappa > 0$ and $B > 0$ such that every rational point on a genus 2 curve of size $N$ has height $\leq BN^\kappa$.*

Su-Ion Ih has shown [1] that Conjecture 3 follows from Vojta's Conjecture. It appears likely that a more careful study will show that Conjecture 2 follows from Vojta's Conjecture as well.

Let us compare with experimental data. Nils Bruin and I have found (very likely) *all* rational points on *all* genus 2 curves of size $N \leq 3$. We found the following records for the height of points.

| size of curves | $N = 1$ | $N = 2$ | $N = 3$ |
|---|---|---|---|
| max. $H(P)$ | 145 | 10 711 | 209 040 |
| max. $H(P)/N^{13/2}$ | 145.00 | 118.34 | 165.55 |

Using an efficient implementation of point search (`ratpoints`, see [5]), we have also found *all* rational points of height $< 16384$ on *all* genus 2 curves of size $N \leq 10$.

Comparing the counts for points in the height brackets $2^n \leq H < 2^{n+1}$ with the heuristic prediction, we observe an overall good agreement, but also that there seem to be "too many" relatively large points. This deviation might be related to the existence of families with "overly large" points as indicated above. We refer to the preprint [4] for more information, including graphics.

## 4 Number of Points

Under the assumption that the events "there is a rational point with $x$-coordinate $x_0$" are independent for all $x_0$, one arrives at a prediction for the number of curves of size $\leq N$ with at least a given number of point pairs (i.e., points with the same $x$-coordinate). Comparing these predictions with our experimental data, it is clear that the assumption must be *wrong*.

Rather, it appears that there is a fairly *constant* probability that a curve with at least $m$ point pairs actually has at least $m + 1$ such pairs. This leads to the following.

**Conjecture 4.** *There is some $B > 0$ such that*

$$\#C(\mathbb{Q}) \leq B \log(2N + 1)$$

*for curves of size $\leq N$.*

This conjecture is in good agreement with the known "best" curves in terms of number of point pairs versus size. See [4] for data and graphs. Some of these come from several families constructed recently by Noam Elkies using K3 surfaces. One of these curves represents the current record for the number of rational points on a genus 2 curve over $\mathbb{Q}$: The curve

$$y^2 = 82342800\,x^6 - 470135160\,x^5 + 52485681\,x^4 + 2396040466\,x^3 \\ + 567207969\,x^2 - 985905640\,x + 247747600$$

has at least **642** rational points! The previous record was 588 points, due to Keller and Kulesz [3].

From the data, it looks like we might have

$$\max\{\#C(\mathbb{Q}) : C \text{ of size} \le N\} \gg \log(2N+1)\,.$$

However, Caporaso, Harris and Mazur [2] show that the "weak Lang Conjecture" implies that $\#C(\mathbb{Q})$ is *bounded*.

We observe that the rank of the group of rational points on the Jacobian variety of the curve also grows with the number of rational points on the curve. So if we assume that the rank is bounded, our data and the CHM result can be reconciled. This is somewhat contrary to a folklore conjecture that the rank should be unbounded, but there seems to be no particularly good reason to assume this (other than there seemed to be no particularly good reason to believe the rank should be bounded, which one might want to reconsider, given the considerations above).

## References

1. Su-Ion Ih, *Height uniformity for algebraic points on curves,* Compositio Math. **134** (2002), no. 1, 35–57.
2. L. Caporaso, J. Harris and B. Mazur, *Uniformity of rational points,* J. Amer. Math. Soc. **10** (1997), 1–35.
3. W. Keller and L. Kulesz, *Courbes algébriques de genre 2 et 3 possédant de nombreux points rationnels* (French), C. R. Acad. Sci. Paris Sér. I Math. **321** (1995), 1469–1472.
4. M. Stoll, *On the average number of rational points on curves of genus 2,* Preprint (2009), arXiv:0902.4165v1 [math.NT].
5. M. Stoll, `ratpoints`, a program that searches for rational points on hyperelliptic curves.