# Rational Points on Curves of Genus 2

Michael Stoll
Jacobs University Bremen

Warwick, January 25, 2008

# Curves of Genus 2

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 $C(\mathbb{Q})$.

**Theorem** (Mordell's Conjecture, proved by Faltings). $C(\mathbb{Q})$ is finite.

# The Questions

Consider curves of genus 2 over $\mathbb{Q}$:

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

with $f_j \in \mathbb{Z}$, $|f_j| \leq N$.

**Question.**
What can we say about $C(\mathbb{Q})$, the set of rational points on $C$, as $N$ grows?

- How do we determine $C(\mathbb{Q})$?
- How large is $C(\mathbb{Q})$ on average?
- How large can the points get?
- How many curves have many rational points?
- How are the sizes of the points distributed?

# Heuristics (1)

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 $2N$.

We can estimate the size of this set by the volume
of the corresponding slice of the cube,
divided by the covolume of the lattice.

We obtain for the average number of points with $x = \frac{a}{b}$:

$$\mathbb{E}_N\Big(\#C(\mathbb{Q})_x\Big) \sim \frac{\gamma(x)}{\sqrt{N}} \qquad \text{as } N \to \infty.$$

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

# Heuristics (2)

We let

$$\gamma(H) = \sum_{x \in \mathbb{P}^1(\mathbb{Q}), H(x) \le H} \gamma(x) = \gamma - O\Big(\frac{1}{H}\Big)$$

where $\gamma = \lim_{H \to \infty} \gamma(H) < \infty$.

If $C(\mathbb{Q})_H = \{(x, y) \in C(\mathbb{Q}) : H(x) \le H\}$, this gives

$$\mathbb{E}_N\Big(\#C(\mathbb{Q})_H\Big) \sim \frac{\gamma(H)}{\sqrt{N}} \qquad \text{as } N \to \infty.$$

**Corollary.** $$\liminf_{N \to \infty} \sqrt{N} \cdot \mathbb{E}_N\Big(\#C(\mathbb{Q})\Big) \ge \gamma\,.$$

**Conjecture.** $$\lim_{N \to \infty} \sqrt{N} \cdot \mathbb{E}_N\Big(\#C(\mathbb{Q})\Big) = \gamma\,.$$

# Heuristics (3)

Let $\pi : C \to \mathbb{P}^1$ be the $x$-coordinate map.
For $P \in C(\mathbb{Q})$, we write $H(P) = H(\pi(P))$.

The linear conditions for up to seven $x$-coordinates are independent, so we expect (at least for $k \leq 7$)

$$\text{Prob}_N\Big(\#\pi(C(\mathbb{Q})) \geq k\Big) \sim \frac{c_k}{N^{k/2}}\,.$$

We also expect that (using $\gamma - \gamma(H) \approx \beta/H$)

$$\#\Big\{C : \exists P \in C(\mathbb{Q}), H(P) \geq H\Big\} \approx \beta\,\frac{N^{13/2}}{H}\,,$$

which leads to the

**Conjecture.** $\quad \max\Big\{H(P) : P \in \bigcup_C C(\mathbb{Q})\Big\} \ll N^{13/2+\varepsilon}\,.$

(Note: This is polynomial in $N$, in contrast to curves of genus 1.)

# How To Find the Points

In order to test these heuristics, we need to find $C(\mathbb{Q})$ for all curves $C$ (with $|f_j| \le N$).

We can search for points on $C$;
this is feasible for heights up to $10^4$ or maybe $10^5$
(using `ratpoints`, or `Points()` in MAGMA; complexity of order $H^2$).

If we do not find any points, we can try to prove that $C(\mathbb{Q}) = \emptyset$:

- local obstruction: $C(\mathbb{R}) = \emptyset$ or $C(\mathbb{Q}_p) = \emptyset$ for some $p$
- 2-cover descent: $\mathrm{Sel}^{(2)}(C/\mathbb{Q}) = \emptyset$
- Mordell-Weil Sieve (see below)

**Example.**
For $N = 3$, there are 196 171 isomorphism classes of curves,
of which 137 490 have rational points, and 58 681 don't. (Bruin-Stoll)

# The Jacobian Variety

The points on an elliptic curve form a group in a natural way.
This helps tremendously when studying such curves.

The points on a curve of genus 2 do not form a group in a natural way.

However, we can embed $C$ into a 2-dimensional variety $J$
whose points do form a group in a natural way.
This variety $J$ is called the Jacobian variety of $C$.

Its rational points form a group $J(\mathbb{Q})$.
Weil generalised a theorem of Mordell's on elliptic curves and showed that
$J(\mathbb{Q})$ is a finitely generated abelian group.

We call $J(\mathbb{Q})$ the Mordell-Weil group of $J$.

# Determining the Mordell-Weil Group

We can use the Mordell-Weil group to get information on $C(\mathbb{Q})$.

For this, we need to compute generators of $J(\mathbb{Q})$;
in particular, we need to find its rank $\quad \operatorname{rank} J(\mathbb{Q}) = \dim_{\mathbb{Q}} J(\mathbb{Q}) \otimes_{\mathbb{Z}} \mathbb{Q}$.

- 2-descent on $J$ gives upper bound for the rank
- Search for points on $J$ gives lower bound for the rank
- Use canonical height $\hat{h}$ to saturate the known subgroup

(Algorithms are implemented in MAGMA.)

**Potential Problems.**

- Upper bound may not be tight
- Some generators may be too large to be found

# Using the Mordell-Weil Group

Let $\iota : C \to J$ be an embedding.
Then $\iota\big(C(\mathbb{Q})\big) = J(\mathbb{Q}) \cap \iota(C)$.

Let $r = \operatorname{rank} J(\mathbb{Q})$.

**Method 1.**

We have $\hat{h}\big(\iota(P)\big) \le \log H(P) + c$.

Enumerate all points $Q \in J(\mathbb{Q})$ with $\hat{h}(Q) \le \log H + c$
(working with the Mordell-Weil lattice $(J(\mathbb{Q})/\text{torsion}, \hat{h})$)
and check for each $Q$ if $Q \in \iota(C)$.

- Complexity of order $$\frac{\#J(\mathbb{Q})_{\text{tors}}}{\sqrt{\text{Reg}_J}} \frac{\pi^{r/2}(\log H + c)^{r/2}}{(r/2)!} = O\big((\log H)^{r/2}\big)$$
- Check $Q \in \iota(C)$ first modulo $p$ for many $p$.
- Can do $H = 10^{100}$ for $r = 3, 4, 5$.

# The Mordell-Weil Sieve (1)

**Method 2.**

Let $S$ be a finite set of primes of good reduction for $C$.
Consider the following diagram.

$$\begin{tikzcd}
C(\mathbb{Q}) \arrow[r, "\iota"] \arrow[d] & J(\mathbb{Q}) \arrow[r] \arrow[d] & J(\mathbb{Q})/nJ(\mathbb{Q}) \arrow[d, "\beta"] \\
\prod_{p \in S} C(\mathbb{F}_p) \arrow[r, "\iota"] \arrow[rr, bend right, "\alpha"'] & \prod_{p \in S} J(\mathbb{F}_p) \arrow[r] & \prod_{p \in S} J(\mathbb{F}_p)/nJ(\mathbb{F}_p)
\end{tikzcd}$$

We can compute the maps $\alpha$ and $\beta$.
If the image of the known rational points on $C$ coincides with $\beta^{-1}(\mathrm{im}(\alpha))$,
then any unknown point must have $\log H(P) \gg n^2$.

(If $\beta^{-1}(\mathrm{im}(\alpha)) = \emptyset$, i.e., $\mathrm{im}(\alpha) \cap \mathrm{im}(\beta) = \emptyset$, this proves that $C(\mathbb{Q}) = \emptyset$.)

# The Mordell-Weil Sieve (2)

A carefully optimized version of the Mordell-Weil sieve
works very well when $r = 2$
and allows us to reach $H = 10^{1000}$ (or more) in this case.

**Example (Bruin-Stoll).**
For the 1 492 curves $C$ for $N = 3$ without rational points
that do not have a local obstruction or a 2-cover obstruction,
a Mordell-Weil sieve computation proves that $C(\mathbb{Q}) = \emptyset$.
(For 42 curves,
we need to assume the Birch and Swinnerton-Dyer Conjecture.)

# A Refinement

Taking $n$ as a multiple of $N$,
the Mordell-Weil sieve gives us a way of proving
that a given coset of $NJ(\mathbb{Q})$ does not meet $\iota(C)$.

**Conjecture.**
If $(Q + NJ(\mathbb{Q})) \cap \iota(C) = \emptyset$, then there are $n \in N\mathbb{Z}$ and $S$ such that
the Mordell-Weil sieve with these parameters proves this fact.

So if we can find an $N$ that separates the rational points on $C$,
i.e., such that the composition $C(\mathbb{Q}) \xrightarrow{\iota} J(\mathbb{Q}) \to J(\mathbb{Q})/NJ(\mathbb{Q})$ is injective,
then we can effectively determine $C(\mathbb{Q})$ if the Conjecture holds for $C$:

For each coset of $NJ(\mathbb{Q})$, we either find a point on $C$ mapping into it,
or we prove that there is no such point.

# Chabauty's Method

Chabauty's method allows us to compute a separating $N$ when $r = 1$.

Let $p$ be a prime of good reduction for $C$. There is a pairing

$$\Omega^1_J(\mathbb{Q}_p) \times J(\mathbb{Q}_p) \longrightarrow \mathbb{Q}_p\,, \qquad (\omega, R) \longmapsto \int_0^R \omega\,.$$

Since $\operatorname{rank} J(\mathbb{Q}) = 1 < 2 = \dim_{\mathbb{Q}_p} \Omega^1_J(\mathbb{Q}_p)$, there is a differential

$0 \neq \omega_p \in \Omega_C(\mathbb{Q}_p) \cong \Omega^1_J(\mathbb{Q}_p)$ that kills $J(\mathbb{Q}) \subset J(\mathbb{Q}_p)$.

**Theorem.**
If the reduction $\bar{\omega}_p$ does not vanish on $C(\mathbb{F}_p)$ and $p > 2$,
then each residue class mod $p$ contains at most one rational point.

This implies that $N = \#J(\mathbb{F}_p)$ is separating.

# Chabauty + MW Sieve

We can easily compute $\bar{\omega}_p$.

Heuristically (at least if $J$ is simple),
we expect to find many $p$ satisfying the condition.

In practice, such $p$ are easily found;
the Mordell-Weil sieve computation then determines $C(\mathbb{Q})$ very quickly.

**Note.**
If $r = 0$, then $C(\mathbb{Q}) = \iota^{-1}\Big(J(\mathbb{Q})_{\text{tors}}\Big)$ is easily computed.

# Summary: Finding Points

For the 137 490 curves with rational points ($N = 3$):

- $r = 0$ (14 010 curves): $C(\mathbb{Q})$ is determined.
- $r = 1$ (46 575 curves): $C(\mathbb{Q})$ is determined.
- $r = 2$ (52 227 curves): $C(\mathbb{Q})_H$ is determined for $H = 10^{1000}$.
- $r = 3$ (22 343 curves): $C(\mathbb{Q})_H$ is determined for $H = 10^{100}$.
- $r = 4$ (2 318 curves): $C(\mathbb{Q})_H$ is determined for $H = 10^{100}$.
- $r = 5$ (17 curves): $C(\mathbb{Q})_H$ is determined for $H = 10^{100}$.

(Caveat: In some cases when $r$ is less than the Selmer rank, the rank is not yet proved to be correct.)

In addition, all points up to $H = 2^{14} - 1 = 16\,383$ were computed on all curves with $N \leq 10$. (This was done using `ratpoints`.)

# Some Records (1)

**Maximal Number of Points** ($H(P) \leq 16383$ for $N \geq 4$)

| size of curves $N$ | 1 | 2 | 3 | 4 | 5 | 6 | 7–10 |
|---|---|---|---|---|---|---|---|
| max. $\#C(\mathbb{Q})$ | 18 | 24 | 26 | 36 | 38 | 44 | 52 |

**Example.** $y^2 = x^6 - 2x^5 + x^4 - 5x^3 - 5x^2 + 7x + 4$

has 52 points, with $x$-coordinates

$$\infty,\ -3,\ -1,\ 0,\ 1,\ 3,\ 4,\ -\frac{5}{2},\ -\frac{1}{2},\ \frac{1}{2},\ \frac{5}{2},\ \frac{11}{2},\ -\frac{4}{3},\ -\frac{1}{3},\ -\frac{9}{4},\ -\frac{1}{4},$$

$$\frac{13}{5},\ \frac{5}{6},\ -\frac{3}{7},\ -\frac{63}{11},\ \frac{96}{11},\ -\frac{7}{15},\ \frac{19}{20},\ -\frac{265}{111},\ -\frac{7}{369},\ \frac{199}{504}.$$

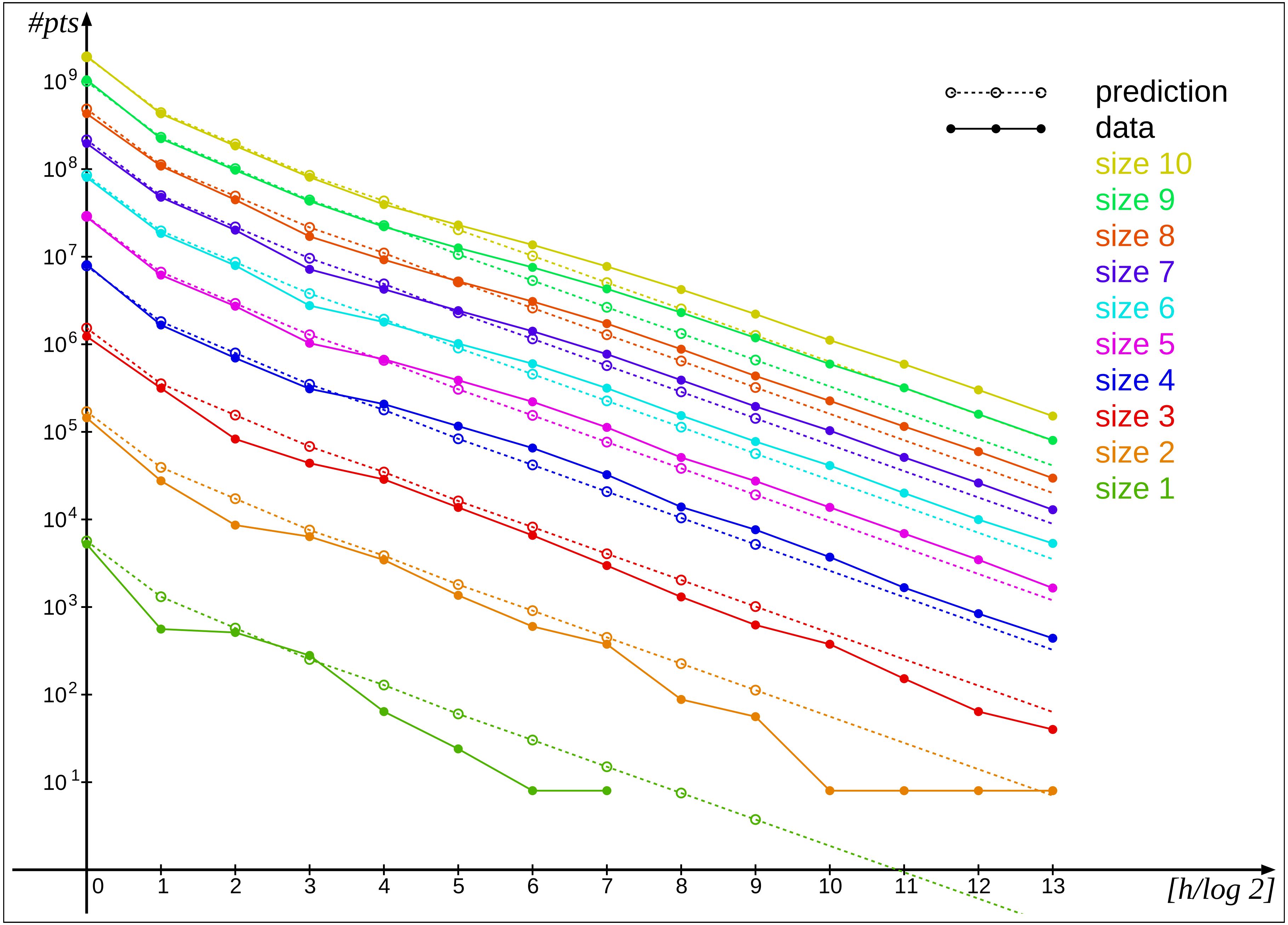
# Some Records (2)

**Maximal Height of Points** $(H(P) \leq 10^{100})$

| size of curves | $N = 1$ | $N = 2$ | $N = 3$ |
|---|---|---|---|
| max. $H(P)$ | 145 | 10711 | 209040 |
| max. $H(P)/N^{13/2}$ | 145.00 | 118.34 | 165.55 |

The record point on $y^2 = x^6 - 3x^4 - x^3 + 3x^2 + 3$

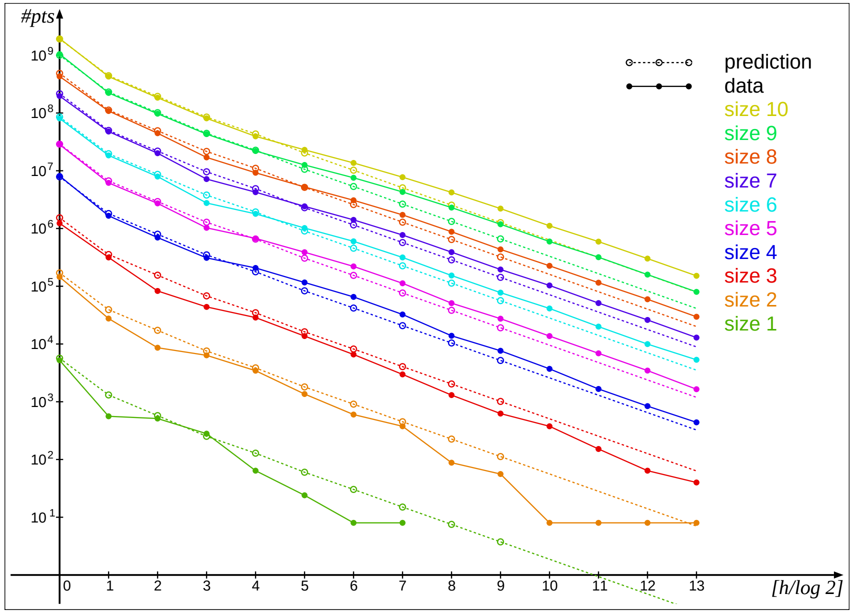
has $x = -\frac{58189}{209040}$

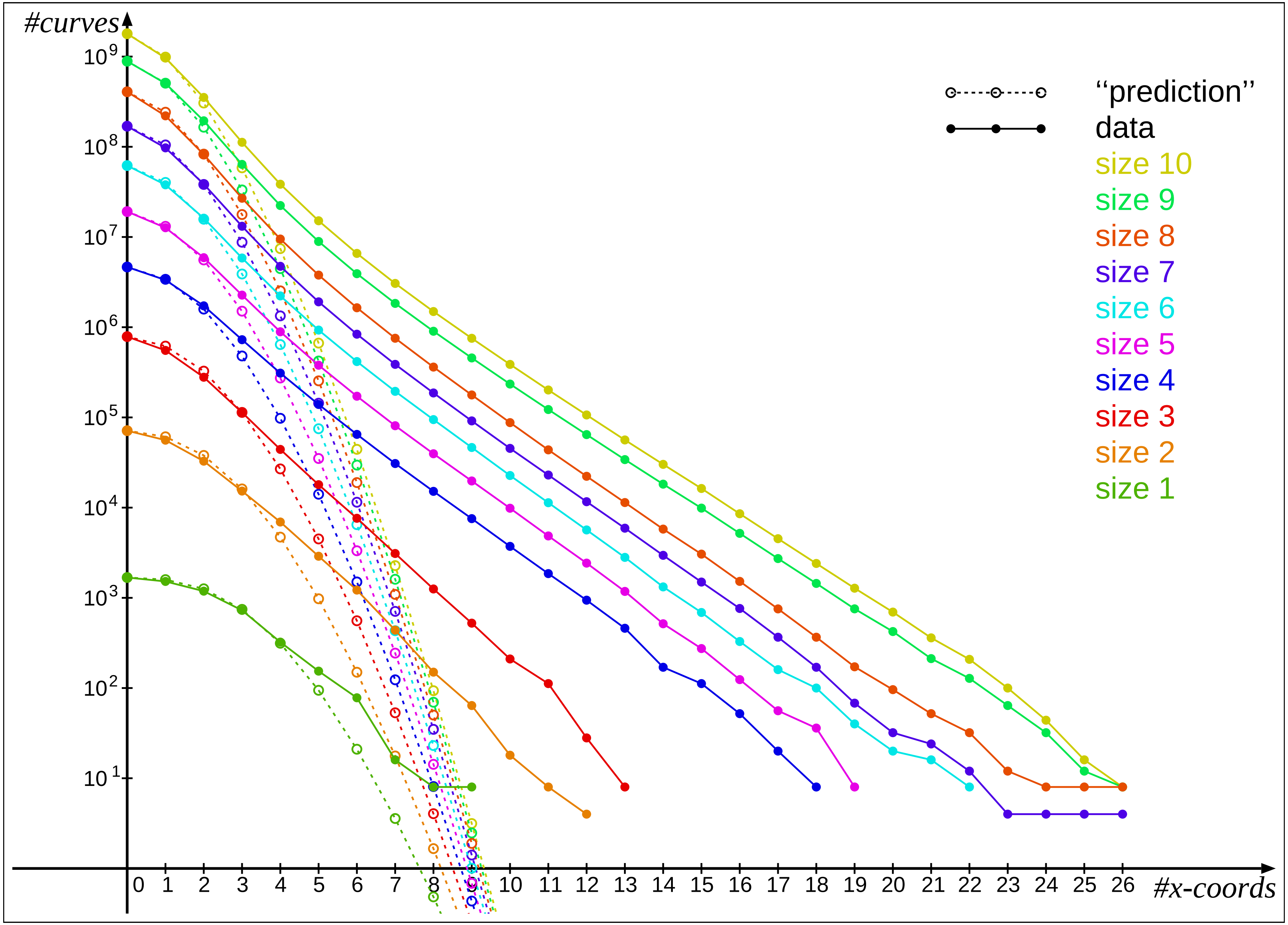


#pts
$10^9$
$10^8$
$10^7$
$10^6$
$10^5$
$10^4$
$10^3$
$10^2$
$10^1$
0
1
2
3
4
5
6
7
8
9
10
11
12
13
[h/log 2]
prediction
data
size 10
size 9
size 8
size 7
size 6
size 5
size 4
size 3
size 2
size 1

# Height Distribution of Points

- We count points $P$ with $2^n \le H(P) < 2^{n+1}$.
- We expect this number to be proportional to $2^{-n}$.
- We observe a reasonably good fit, but for $N \ge 4$, there are "too many" large points. Explanation?





#curves
$10^9$
$10^8$
$10^7$
$10^6$
$10^5$
$10^4$
$10^3$
$10^2$
$10^1$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
#x-coords
"prediction"
data
size 10
size 9
size 8
size 7
size 6
size 5
size 4
size 3
size 2
size 1

# Number of $x$-Coordinates

- We count curves $C$ with $\#\pi\Big(C(\mathbb{Q})\Big) \ge k$.
- We expect this number to be proportional to $N^{-k/2}$ (at least for $k \le 7$).
- We observe that the numbers drop much more slowly than expected: Given that there are $n$ point pairs already, there is a $\approx 50\%$ chance for another pair.
- This needs to be investigated!

# A Pre-Conjecture

Generalizing to larger $N$, this leads to the following

**Expectation.** $\max\left\{\#\pi(C(\mathbb{Q}))\right\} \approx c \log(2N+1)$.

Here is a table for $N \leq 10$.

| size of curves $N$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| max. $\frac{\#\pi(C(\mathbb{Q}))}{\log(2N+1)}$ | 8.2 | 7.5 | 6.7 | 8.2 | 7.9 | 8.6 | 9.6 | 9.2 | 8.8 | 8.6 |

A few examples of curves with many points:

| Author | $N$ | $\#\pi(C(\mathbb{Q}))$ | $\frac{\#\pi(C(\mathbb{Q}))}{\log(2N+1)}$ |
|---|---|---|---|
| Stahlke | 282 | 63 | 9.94 |
| Stahlke | 249094440 | 183 | 9.14 |
| Keller and Kulesz | 22999624761 | 294 | 11.97 |

# A Couple of Remarks

**Remark 1.**
By work of Caporaso, Harris, and Mazur,
Lang's conjecture on rational points on varieties of general type implies that $\#C(\mathbb{Q})$ is uniformly bounded for curves $C/\mathbb{Q}$ of fixed genus.

**Remark 2.**
One can easily construct families of genus 2 curves with

- at least 14 point pairs and no extra automorphisms,
- at least 16 point pairs and an extra involution,
- at least 24 point pairs and a large automorphism group,

whereas the indpendence assumption for many points would predict that there are only finitely many curves in total with more than 14 point pairs.

# Summary

- For “small” curves $C$, we can decide if $C(\mathbb{Q})$ is empty or not, and we can find all points up to very large height.
- We can use heuristic coniderations that lead to expectations regarding the number and size of rational points.
- In a massive computation, we found all rational points of height $\leq 16383$ on all curves of size $\leq 10$.
- The experimental data confirm some of the expectations, but are in disagreement with others.
- It is an interesting challenge to explain the discrepancies!