

Rational Points on Hypersurfaces in Projective Space

Jörg Jahnel

Mathematisches Institut der Universität Göttingen

25.07.2006

joint work with Andreas-Stephan Elsenhans

Problem (Diophantine equation)

Given $f \in \mathbb{Z}[X_0, \dots, X_n]$, describe the set

$$\{(x_0, \dots, x_n) \in \mathbb{Z}^{n+1} \mid f(x_0, \dots, x_n) = 0\},$$

explicitly.

The Fundamental Problem

More realistic from the computational point of view:

Problem (Diophantine equation—search for solutions)

Given $f \in \mathbb{Z}[X_0, \dots, X_n]$ and $B > 0$, describe the set

$$\{(x_0, \dots, x_n) \in \mathbb{Z}^{n+1} \mid f(x_0, \dots, x_n) = 0, |x_j| \leq B\},$$

explicitly.

B is usually called the *search limit*.

- Integral points on an n -dimensional hypersurface in \mathbf{A}^{n+1} .

- Integral points on an n -dimensional hypersurface in \mathbf{A}^{n+1} .
- If f is homogeneous: Rational points on an $(n - 1)$ -dimensional hypersurface V_f in \mathbf{P}^n .

A statistical forecast

$$Q(B) := \{(x_0, \dots, x_n) \in \mathbb{Z}^{n+1} \mid |x_i| \leq B\}$$

Thus,

$$\#Q(B) = (2B + 1)^{n+1} \sim C_1 \cdot B^{n+1}.$$

On the other hand,

$$\max_{(x_0, \dots, x_n) \in Q(B)} |f(x_0, \dots, x_n)| \sim C_2 \cdot B^{\deg f}.$$

Assuming equidistribution of the values of f on $Q(B)$, we are therefore led to expect the asymptotics

$$\#\{(x_0, \dots, x_n) \in V_f(\mathbb{Q}) \mid |x_0|, \dots, |x_n| \leq B\} \sim C \cdot B^{n+1 - \deg f}$$

for the number of solutions.

Examples

The statistical forecast explains the following well-known examples.

- $n + 1 - \deg f < 0$: Very few solutions.

Example: $x^k + y^k = z^k$ for $k \geq 4$.

The statistical forecast explains the following well-known examples.

- $n + 1 - \deg f < 0$: Very few solutions.

Example: $x^k + y^k = z^k$ for $k \geq 4$.

- $n + 1 - \deg f = 0$: A few solutions.

Example: $y^2z = x^3 + 8xz^2$.

Elliptic curves.

Another Example: $x^4 + 2y^4 = z^4 + 4w^4$.

More generally, surfaces of type $K3$.

The statistical forecast explains the following well-known examples.

- $n + 1 - \deg f < 0$: Very few solutions.

Example: $x^k + y^k = z^k$ for $k \geq 4$.

- $n + 1 - \deg f = 0$: A few solutions.

Example: $y^2z = x^3 + 8xz^2$.

Elliptic curves.

Another Example: $x^4 + 2y^4 = z^4 + 4w^4$.

More generally, surfaces of type $K3$.

- $n + 1 - \deg f > 0$: Many solutions.

Example: $x^2 + y^2 = z^2$.

Conics.

Another Example: $x^3 + y^3 + z^3 + w^3 = 0$.

Cubic surfaces.

- Unsolvability

- Unsolvability in reals,

$$x^2 + y^2 + z^2 = 0.$$

- p -adic unsolvability,

$$u^3 + 2v^3 + 7w^3 + 14x^3 + 49y^3 + 98z^3 = 0.$$

- Obstructions against the Hasse principle
(Brauer-Manin obstruction, unknown obstructions?).

A few complications

- Unsolvability

- Unsolvability in reals,

$$x^2 + y^2 + z^2 = 0.$$

- p -adic unsolvability,

$$u^3 + 2v^3 + 7w^3 + 14x^3 + 49y^3 + 98z^3 = 0.$$

- Obstructions against the Hasse principle

(Brauer-Manin obstruction, unknown obstructions?).

- “Accumulating” subvarieties:

$x^3 + y^3 = z^3 + w^3$ defines a cubic surface V in \mathbf{P}^3 .

$$\#\{(x_0, \dots, x_n) \in V(\mathbb{Q}) \mid |x_0|, \dots, |x_n| \leq B\} \sim C \cdot B$$

is predicted.

However, V contains the line given by $x = z$, $y = w$, on which there is quadratic growth, already.

Manin's conjecture

Let us concentrate on the case $n + 1 - \deg f > 0$.

Conjecture (Manin)

Let V_f be a smooth hypersurface in \mathbf{P}^n . Assume $n + 1 - \deg f > 0$. Then,

$$\#\{(x_0, \dots, x_n) \in V^\circ(\mathbb{Q}) \mid |x_0|, \dots, |x_n| \leq B\} \sim C \cdot B^k \log^{r-1} B.$$

Here, $k := n + 1 - \deg f$, $r = \text{rk Pic } V$.

Remarks

- C is an explicit constant described by E. Peyre.
- The assumption $n + 1 - \deg f > 0$ implies V_f is a Fano variety.

What is known?

- Manin's conjecture is true for $n \gg 2^{\deg f}$ (circle method).
[Birch, B. J.: *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263]

What is known?

- Manin's conjecture is true for $n \gg 2^{\deg f}$ (circle method).
[Birch, B. J.: *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263]
- Manin's conjecture is established in many particular cases of low dimension, e.g.
 - generalized flag varieties (Franke, Manin, Tschinkel),
 - projective smooth toric varieties (Batyrev and Tschinkel),
 - certain toric fibrations over generalized flag varieties (Strauch and Tschinkel),
 - smooth equivariant compactifications of affine spaces (Chambert-Loir and Tschinkel),
 - $\mathbf{P}_{\mathbb{Q}}^2$ blown-up in up to four points in general position (Salberger, de la Bretèche).

What is known?

- Manin's conjecture is true for $n \gg 2^{\deg f}$ (circle method).
[Birch, B. J.: *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263]
- Manin's conjecture is established in many particular cases of low dimension, e.g.
 - generalized flag varieties (Franke, Manin, Tschinkel),
 - projective smooth toric varieties (Batyrev and Tschinkel),
 - certain toric fibrations over generalized flag varieties (Strauch and Tschinkel),
 - smooth equivariant compactifications of affine spaces (Chambert-Loir and Tschinkel),
 - $\mathbf{P}_{\mathbb{Q}}^2$ blown-up in up to four points in general position (Salberger, de la Bretèche).
- If Manin's conjecture is true for X and Y then for $X \times Y$, too (Franke, Manin, Tschinkel).

What is known?

- Manin's conjecture is true for $n \gg 2^{\deg f}$ (circle method).
[Birch, B. J.: *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263]
- Manin's conjecture is established in many particular cases of low dimension, e.g.
 - generalized flag varieties (Franke, Manin, Tschinkel),
 - projective smooth toric varieties (Batyrev and Tschinkel),
 - certain toric fibrations over generalized flag varieties (Strauch and Tschinkel),
 - smooth equivariant compactifications of affine spaces (Chambert-Loir and Tschinkel),
 - $\mathbf{P}_{\mathbb{Q}}^2$ blown-up in up to four points in general position (Salberger, de la Bretèche).
- If Manin's conjecture is true for X and Y then for $X \times Y$, too (Franke, Manin, Tschinkel).
- Manin's conjecture is *open* for smooth cubic surfaces. (There is, however, a lot of numerical evidence [Heath-Brown, Peyre-Tschinkel].)

Experimental Result (E.+J.)

There is numerical evidence for Manin's Conjecture in the case of the hypersurfaces in $\mathbf{P}_{\mathbb{Q}}^4$ given by $ax^e = by^e + z^e + v^e + w^e$ for $e = 3$ and 4 .

This requires algorithms to

- solve Diophantine equations,

Experimental Result (E.+J.)

There is numerical evidence for Manin's Conjecture in the case of the hypersurfaces in $\mathbf{P}_{\mathbb{Q}}^4$ given by $ax^e = by^e + z^e + v^e + w^e$ for $e = 3$ and 4 .

This requires algorithms to

- solve Diophantine equations,
- compute Peyre's constant,

Experimental Result (E.+J.)

There is numerical evidence for Manin's Conjecture in the case of the hypersurfaces in $\mathbf{P}_{\mathbb{Q}}^4$ given by $ax^e = by^e + z^e + v^e + w^e$ for $e = 3$ and 4 .

This requires algorithms to

- solve Diophantine equations,
- compute Peyre's constant,
- detect accumulating subvarieties.

An algorithm to solve Diophantine equations I

The following example was our starting point.

Example (Sir P. Swinnerton-Dyer, 2002)

The equation

$$x^4 + 2y^4 = z^4 + 4w^4$$

defines a $K3$ surface S in \mathbf{P}^3 .

$(1 : 0 : 1 : 0)$ and $(1 : 0 : (-1) : 0)$ are \mathbb{Q} -rational points on S , the two *obvious* points.

Is there another \mathbb{Q} -rational point on S ?

An algorithm to solve Diophantine equations II

How to find a solution of $x^4 + 2y^4 = z^4 + 4w^4$?

Algorithm (A naive algorithm)

Write $x^4 + 2y^4 - 4w^4 = z^4$.

Let x , y , and w run in a triple loop and test whether $x^4 + 2y^4 - 4w^4$ is a fourth power.

Complexity: $C \cdot B^3$.

Realistic search bound: 50 000.

(We did a trial run with search bound 10 000.)

An algorithm to solve Diophantine equations III

How to find a solution of $x^4 + 2y^4 = z^4 + 4w^4$?

Algorithm (A better algorithm)

The two sets $\{x^4 + 2y^4 \mid |x|, |y| \leq B\}$ and $\{z^4 + 4w^4 \mid |z|, |w| \leq B\}$ have $\sim B^2$ elements each. List them and form their intersection.

Facts

- *Complexity:* $O(B^2 \log B)$ (using sorting, D. Bernstein),
 $O(B^2)$ (assuming uniform hashing, E.+J.).
- *Memory Usage:* $O(B^2)$ (naively),
 $O(B)$ (D. Bernstein's Algorithm—
generates the sets in sorted order.)

Our method works for Diophantine equations of the form

$$f(x_1, \dots, x_k) = g(y_1, \dots, y_l).$$

Detection of solutions of Diophantine equations— Hashing II

Writing

We store the vectors (x_1, \dots, x_k) in a hash table (with space for up to 2^{27} entries).

The *hash function* $H: \mathbb{Z} \rightarrow [0, 2^{27} - 1]$ is given by a selection of bits, i.e. $H(z) :=$ a selection of bits of $(z \bmod 2^{64})$.

For each vector (x_1, \dots, x_k) , the expression $H(f(x_1, \dots, x_k))$ defines its position in the hash table.

Besides (x_1, \dots, x_k) , we also write a 31-bit *control value* $K(f(x_1, \dots, x_k))$, $K(z) :=$ a selection of the remaining bits of $(z \bmod 2^{64})$.

Reading

Then, we search for vectors (y_1, \dots, y_l) such that hash value and control value do fit.

Detection of solutions of Diophantine equations— Hashing III

Remarks

- 1 Assuming uniform hashing (which implies there are not too many solutions), the expected running time is $O(B^{\max(k,l)})$.

Congruence conditions might help to reduce the O -factor.

Detection of solutions of Diophantine equations— Hashing III

Remarks

- 1 Assuming uniform hashing (which implies there are not too many solutions), the expected running time is $O(B^{\max(k,l)})$.

Congruence conditions might help to reduce the O -factor.

- 2 The algorithm actually detects *pseudo-solutions* where a coincidence of the control values and an “almost coincidence” of the hash values occurs.

Some *post processing* with an exact multiprecision calculation is necessary (Aribas, GMP).

How to reduce memory usage when hashing?

Idea (Paging)

Choose $m \in \mathbb{Z}$ sufficiently large. Form the sets

$$L_c := \{f(x_1, \dots, x_k) \mid |x_1|, \dots, |x_k| \leq B, f(x_1, \dots, x_k) \equiv c \pmod{m}\}$$

and

$$R_c := \{g(y_1, \dots, y_l) \mid |y_1|, \dots, |y_l| \leq B, g(y_1, \dots, y_l) \equiv c \pmod{m}\}$$

and work for each c separately.

- Memory usage: Reduced to $B^{\max(k,l)}/m$ (assuming equidistribution).
- One may work in parallel on several machines.

We choose m as a prime number, the *page prime*. Then, the Weil conjectures, proven by P. Deligne, imply an a-priori estimate for the load factor of the hash tables.

Equation $x^4 + 2y^4 = z^4 + 4w^4$ —

Optimization through congruence conditions I

x and z are odd. y and w are even.

- Case 1: $5|y, w \implies 5 \nmid x, z$.
Then, $x^4 \equiv z^4 \pmod{625}$.
We write pairs (x, z) and hash $x^4 - z^4$. We read $4w^4 - 2y^4$.
- Case 2: $5|x, y \implies 5 \nmid z, w$.
Then, $z^4 + 4w^4 \equiv 0 \pmod{625}$.
Here, we write pairs (z, w) and hash $z^4 + 4w^4$. We read $x^4 + 2y^4$.

These congruences are particularly strong. They reduce the number of writing steps to 0.512% and the number of reading steps to 4%.

Equation $x^4 + 2y^4 = z^4 + 4w^4$ —

Optimization through congruence conditions II

Further congruences:

- Some congruence modulo 8:

In Case 1, we always have $32|4w^4 - 2y^4$. But $32|x^4 - z^4$ implies $x \equiv \pm z \pmod{8}$. This saves on writing.

There is no such optimization for Case 2.

- Some congruences modulo 81:

In Case 1, $2y^4 - 4w^4$ represents $(0 \pmod{3})$ only trivially. Therefore, we do not need to write (x, z) when $x^4 \equiv z^4 \pmod{3}$ but $x^4 \not\equiv z^4 \pmod{81}$.

In Case 2, there is a similar congruence which saves on the reading step.

A new solution—

Answer to Sir P. Swinnerton-Dyer's question

Calculation

$\implies 1484801^{**4} + 2 * 1203120^{**4}.$

$-\text{: } 90509_10498_47564_80468_99201$

$\implies 1169407^{**4} + 4 * 1157520^{**4}.$

$-\text{: } 90509_10498_47564_80468_99201$

A new solution—

Answer to Sir P. Swinnerton-Dyer's question

Calculation

$\implies 1484801^{**4} + 2 * 1203120^{**4}.$

$-\colon 90509_10498_47564_80468_99201$

$\implies 1169407^{**4} + 4 * 1157520^{**4}.$

$-\colon 90509_10498_47564_80468_99201$

Theorem (E.+J.)

Up to changes of sign, $(1\ 484\ 801 : 1\ 203\ 120 : 1\ 169\ 407 : 1\ 157\ 520)$ is the only non-obvious \mathbb{Q} -rational point of height $\leq 10^8$ on Sir P. Swinnerton-Dyer's surface S .

This means, on S there exist precisely ten \mathbb{Q} -rational points of height $\leq 10^8$.

A new solution—

Answer to Sir P. Swinnerton-Dyer's question II

Remarks

- The new solution was found on December 2, 2004 by an intermediate version of our programs for search bound $2.5 \cdot 10^6$.
- The final version of the programs (for search bound 10^8) took almost exactly 100 days of CPU time on an AMD Opteron 248 processor. This time is composed almost equally of 50 days for Case 1 and 50 days for Case 2.
- We executed our computations at the Göttingen Gauß Laboratory for Scientific Computing. The main computation was run on a Sun Fire V20z Server in parallel on two processors during February and March, 2005.

Manin's Conjecture—Peyre's constant I

Recall, we consider the hypersurfaces in $\mathbf{P}_{\mathbb{Q}}^4$ given by

$$ax^e - by^e = z^e + v^e + w^e$$

for $e = 3$ and 4 .

Remarks

- 1 Search for \mathbb{Q} -rational points is obviously of complexity $O(B^3)$.
- 2 When considering $O(B)$ varieties (differing only by a and b), simultaneously, then the running-time is *still* $O(B^3)$.

We considered the varieties with $a, b = 1, \dots, 100$ (i.e. 5 000 cubics and 10 000 quartics) with a search bound of 5 000 (cubics) and 100 000 (quartics), respectively.

Manin's Conjecture—Peyre's constant II

Conjecture (Manin's Conjecture—Version for hypersurfaces in \mathbf{P}^n)

Let the smooth variety $V_f \subset \mathbf{P}^n$ be given by $f = 0$. Then,

$$\#\{(x_0, \dots, x_n) \in V^\circ(\mathbb{Q}) \mid |x_0|, \dots, |x_n| \leq B\} \sim C \cdot B^k \log^{r-1} B,$$

for $k = n + 1 - \deg(f)$ and $r = rk \text{ Pic } V$.

Here, C is an explicit constant (due to E. Peyre),

[Peyre, E.: *Hauteurs et mesures de Tamagawa sur les variétés de Fano*, Duke Math. J. **79** (1995), 101–218, Définition 2.3].

Definition (Peyre's constant)

For $n \geq 4$, *Peyre's constant* is the Tamagawa-type number

$$C := \prod_{p \in \mathbb{P} \cup \{\infty\}} \left(1 - \frac{1}{p}\right) \tau_p$$

where

$$\tau_p = \lim_{m \rightarrow \infty} \frac{\#V(\mathbb{Z}/p^m\mathbb{Z})}{p^{m \dim(V)}} \quad \text{for } p \in \mathbb{P}$$

and

$$\tau_\infty = \frac{1}{2} \int_{\substack{f(x_0, \dots, x_n) = 0 \\ |x_0|, \dots, |x_n| \leq 1}} \frac{1}{\frac{\partial f}{\partial x_j}} dx_0 \dots \widehat{dx_j} \dots dx_n.$$

Remark

To compute Peyre's constant,

- the infinite place and primes of bad reduction require some programming efforts.
- But the main work to be done is on good primes where one has to *count* solutions of the same equation $f(x_0, \dots, x_n) = 0$ but over finite fields \mathbb{F}_p instead of \mathbb{Z} .

An algorithm to count solutions II

Consider an equation of the form

$$(*) \quad \sum_{i=0}^n f_i(x_i) = 0.$$

Denote by $d^{(i)}(k) := \#\{x \in \mathbb{F}_p \mid f_i(x) = k\}$ the numbers of representations. Then, the number of solutions of $(*)$ is equal to

$$(d^{(0)} * d^{(1)} * \dots * d^{(n)})(0).$$

Idea (of the algorithm to compute Peyre's constant)

Use FFT to compute the convolution $d^{(0)} * d^{(1)} * \dots * d^{(n)}$.

Remarks (Complexity)

- We need to compute n convolutions of vectors of length p .
- A convolution takes $O(p \log p)$ steps.

Investigation of the cubic threefolds I

We determined all \mathbb{Q} -rational points of height less than 5000 on the cubic threefolds $V_{a,b}^3$ given by

$$ax^3 = by^3 + z^3 + v^3 + w^3$$

for $a, b = 1, \dots, 100$ and $b \leq a$.

Points lying on a \mathbb{Q} -rational line in $V_{a,b}$ were excluded from the count. The smallest number of points found is 3930278 for $(a, b) = (98, 95)$. The largest numbers of points are 332137752 for $(a, b) = (7, 1)$ and 355689300 in the case that $a = 1$ and $b = 1$.

On the other hand, for each threefold $V_{a,b}^3$ where $a, b = 1, \dots, 100$ and $b + 3 \leq a$, we calculated the number of points expected (according to Manin-Peyre) and the quotients

$$\# \{ \text{points of height} < B \text{ found} \} / \# \{ \text{points of height} < B \text{ expected} \}.$$

Let us visualize the quotients by two histograms.

Investigation of the cubic threefolds II

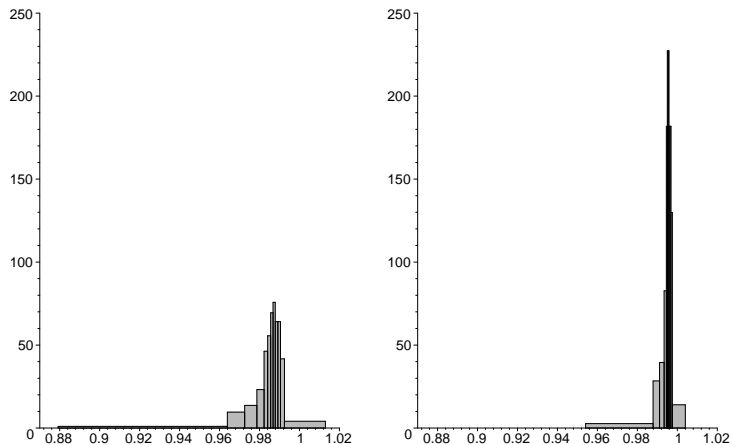


Figure: Distribution of the quotients for $B = 1000$ and $B = 5000$.

Investigation of the cubic threefolds III

Table: Parameters of the distribution in the cubic case

	$B = 1\,000$	$B = 2\,000$	$B = 5\,000$
mean value	0.981 79	0.988 54	0.993 83
standard deviation	0.012 74	0.008 23	0.004 55

Investigation of the quartic threefolds I

We determined all \mathbb{Q} -rational points of height less than 100 000 on the quartic threefolds $V_{a,b}^4$ given by

$$ax^4 = by^4 + z^4 + v^4 + w^4$$

for $a, b = 1, \dots, 100$.

It turns out that on 5015 of these varieties, there are no \mathbb{Q} -rational points occurring at all as the equation is unsolvable in \mathbb{Q}_p for $p = 2, 5, \text{ or } 29$. In this situation, Manin's conjecture is true, trivially.

For the remaining varieties, the points lying on a known \mathbb{Q} -rational conic in $V_{a,b}$ were excluded from the count.

Investigation of the quartic threefolds II

Table: Numbers of points of height $< 100\,000$ on the quartics.

a	b	# points	# not on conic	# expected (by Manin-Peyre)
29	29	2	2	135
58	87	288	288	272
58	58	290	290	388
87	87	386	386	357
\vdots	\vdots	\vdots	\vdots	\vdots
34	1	9 938 976	5 691 456	5 673 000
17	64	5 708 664	5 708 664	5 643 000
1	14	7 205 502	6 361 638	6 483 000
3	1	12 657 056	7 439 616	7 526 000

Investigation of the quartic threefolds III

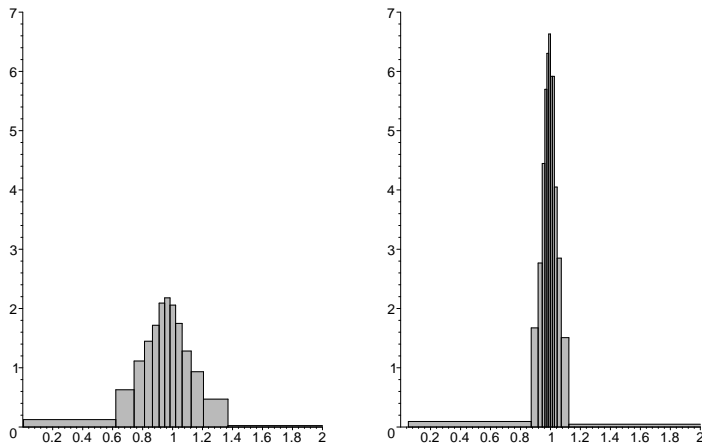


Figure: Distribution of the quotients for $B = 1000$ and $B = 10000$.

Investigation of the quartic threefolds IV

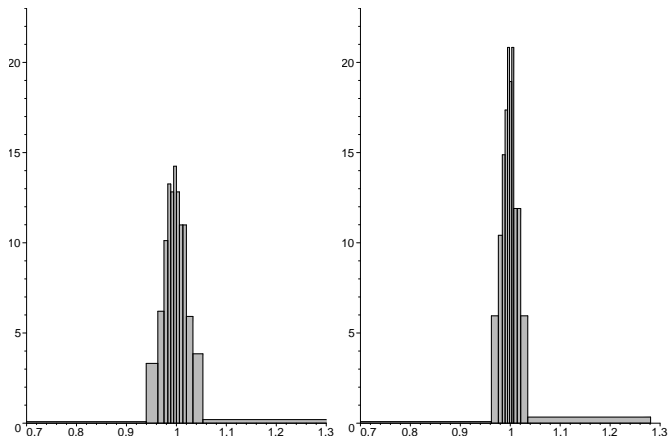


Figure: Distribution of the quotients for $B = 50\,000$ and $B = 100\,000$.

Investigation of the quartic threefolds V

Table: Parameters of the distribution in the quartic case

	$B = 1\,000$	$B = 10\,000$	$B = 100\,000$
mean value	0.9853	0.9957	0.9982
standard deviation	0.3159	0.1130	0.0372

Remark

In the cubic case, presented before, the standard deviation is by far smaller than in the case of the quartics.

This is not very surprising as on a cubic there tend to be much more rational points than on a quartic (quadratic growth versus linear growth). Thus, in the case of the cubic the sample is more reliable.

Investigation of the quartic threefolds VI

search limit 50000 - color represents quotient at limit 10000

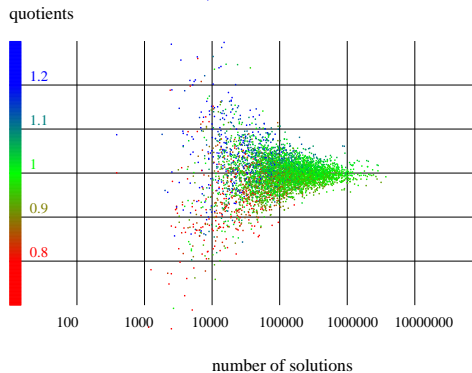


Figure: number of solutions and quotients for $B = 50\,000$.

Summary

- *To search systematically for solutions of Diophantine equations like $x^4 + 2y^4 = z^4 + 4w^4$ or $7x^3 = 11y^3 + z^3 + v^3 + w^3$ ($n \geq 4$ variables), there are ways faster than the obvious $(n - 1)$ -times iterated loop. (Essentially, we need $O(B^{\lceil n/2 \rceil})$ steps).*

Summary

- *To search systematically for solutions of Diophantine equations like $x^4 + 2y^4 = z^4 + 4w^4$ or $7x^3 = 11y^3 + z^3 + v^3 + w^3$ ($n \geq 4$ variables), there are ways faster than the obvious $(n - 1)$ -times iterated loop. (Essentially, we need $O(B^{\lceil n/2 \rceil})$ steps).*
- *To count solutions over \mathbb{F}_p (without determining all of them) is even faster ($O(np \log p)$ steps).*

Summary

- *To search systematically for solutions of Diophantine equations like $x^4 + 2y^4 = z^4 + 4w^4$ or $7x^3 = 11y^3 + z^3 + v^3 + w^3$ ($n \geq 4$ variables), there are ways faster than the obvious $(n - 1)$ -times iterated loop. (Essentially, we need $O(B^{\lceil n/2 \rceil})$ steps).*
- *To count solutions over \mathbb{F}_p (without determining all of them) is even faster ($O(np \log p)$ steps).*
- *These two observations together may be used to test Manin's conjecture, numerically.*

Summary

- *To search systematically for solutions of Diophantine equations like $x^4 + 2y^4 = z^4 + 4w^4$ or $7x^3 = 11y^3 + z^3 + v^3 + w^3$ ($n \geq 4$ variables), there are ways faster than the obvious $(n - 1)$ -times iterated loop. (Essentially, we need $O(B^{\lceil n/2 \rceil})$ steps).*
- *To count solutions over \mathbb{F}_p (without determining all of them) is even faster ($O(np \log p)$ steps).*
- *These two observations together may be used to test Manin's conjecture, numerically.*

Remark (Conclusion)

The results suggest that Manin's conjecture should be true for the two families of threefolds considered.