An overview of arithmetic statistics

Peter Koymans
Utrecht University

Utrecht University

Bonn

19 January 2026

Arithmetic statistics

The aim of arithmetic statistics is to answer statistical questions of arithmetic objects (e.g. zeta functions, number fields, class groups) with many applications to other areas of mathematics.

Arithmetic statistics

The aim of arithmetic statistics is to answer statistical questions of arithmetic objects (e.g. zeta functions, number fields, class groups) with many applications to other areas of mathematics.

We will discuss several leading conjectures in arithmetic statistics in this talk and my recent work on them.

Arithmetic statistics

The aim of arithmetic statistics is to answer statistical questions of arithmetic objects (e.g. zeta functions, number fields, class groups) with many applications to other areas of mathematics.

We will discuss several leading conjectures in arithmetic statistics in this talk and my recent work on them.

Techniques from arithmetic statistics can also be used to prove results that do not have a statistical nature.

Hilbert's tenth problem

David Hilbert

At the 1900 mathematical conference in Paris, Hilbert introduced his famous list of 23 problems.

Hilbert's tenth problem

David Hilbert

At the 1900 mathematical conference in Paris, Hilbert introduced his famous list of 23 problems.

Question (Hilbert's tenth problem)

Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined in a finite number of operations whether the equation is solvable in rational integers.

Hilbert's tenth problem

David Hilbert

At the 1900 mathematical conference in Paris, Hilbert introduced his famous list of 23 problems.

Question (Hilbert's tenth problem)

Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined in a finite number of operations whether the equation is solvable in rational integers.

In modern terms: does there exist an algorithm such that:
Input: a polynomial $p \in \mathbb{Z}[x_1, \ldots, x_n]$.
Output: "YES" if there is an integer solution $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ with $p(a_1, \ldots, a_n) = 0$, "NO" otherwise.

Diophantine and listable sets

Definition (Diophantine set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is Diophantine if there exists a polynomial $p(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{Z}[\mathbf{x}, \mathbf{y}]$ such that

$$S = \{\mathbf{x} \in \mathbb{Z}^n : \exists \mathbf{y} \in \mathbb{Z}^m \text{ such that } p(\mathbf{x}, \mathbf{y}) = 0\}.$$

Diophantine and listable sets

Definition (Diophantine set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is Diophantine if there exists a polynomial $p(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{Z}[\mathbf{x}, \mathbf{y}]$ such that

$$S = \{\mathbf{x} \in \mathbb{Z}^n : \exists \mathbf{y} \in \mathbb{Z}^m \text{ such that } p(\mathbf{x}, \mathbf{y}) = 0\}.$$

Definition (Listable set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is listable (or recursively enumerable) if there is an algorithm that enumerates S when left running forever.

Diophantine and listable sets

Definition (Diophantine set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is Diophantine if there exists a polynomial $p(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{Z}[\mathbf{x}, \mathbf{y}]$ such that

$$S = \{\mathbf{x} \in \mathbb{Z}^n : \exists \mathbf{y} \in \mathbb{Z}^m \text{ such that } p(\mathbf{x}, \mathbf{y}) = 0\}.$$

Definition (Listable set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is listable (or recursively enumerable) if there is an algorithm that enumerates S when left running forever.

Theorem (MRDP, 1970)

A subset $S \subseteq \mathbb{Z}^n$ is Diophantine if and only if it is listable.

Diophantine and listable sets

Definition (Diophantine set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is Diophantine if there exists a polynomial $p(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{Z}[\mathbf{x}, \mathbf{y}]$ such that

$$S = \{\mathbf{x} \in \mathbb{Z}^n : \exists \mathbf{y} \in \mathbb{Z}^m \text{ such that } p(\mathbf{x}, \mathbf{y}) = 0\}.$$

Definition (Listable set)

We say that a subset $S \subseteq \mathbb{Z}^n$ is listable (or recursively enumerable) if there is an algorithm that enumerates S when left running forever.

Theorem (MRDP, 1970)

A subset $S \subseteq \mathbb{Z}^n$ is Diophantine if and only if it is listable.

Corollary (Hilbert's tenth problem)

Hilbert's tenth problem is undecidable, i.e. there is no algorithm that can decide whether a polynomial $p \in \mathbb{Z}[x_1, \ldots, x_n]$ has a zero or not.

Finitely generated rings

Matiyasevich asks in the 1970s: what about other rings?

Finitely generated rings

Matiyasevich asks in the 1970s: what about other rings?

Definition

For a finitely generated ring R, we have analogues of "Hilbert's tenth problem", "Diophantine set" and "listable set" by replacing all occurrences of $\mathbb{Z}$ by R.

Finitely generated rings

Matiyasevich asks in the 1970s: what about other rings?

Definition

For a finitely generated ring R, we have analogues of "Hilbert's tenth problem", "Diophantine set" and "listable set" by replacing all occurrences of $\mathbb{Z}$ by R.

Theorem (Mazur–Rubin, 2009)

Assume BSD. Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

Finitely generated rings

Matiyasevich asks in the 1970s: what about other rings?

Definition

For a finitely generated ring R, we have analogues of "Hilbert's tenth problem", "Diophantine set" and "listable set" by replacing all occurrences of $\mathbb{Z}$ by R.

Theorem (Mazur–Rubin, 2009)

Assume BSD. Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

Theorem (K.–Pagano, 2024)

Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

Finitely generated rings

Matiyasevich asks in the 1970s: what about other rings?

Definition

For a finitely generated ring R, we have analogues of "Hilbert's tenth problem", "Diophantine set" and "listable set" by replacing all occurrences of $\mathbb{Z}$ by R.

Theorem (Mazur–Rubin, 2009)

Assume BSD. Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

Theorem (K.–Pagano, 2024)

Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

We do this by proving the following conjecture of Denef–Lipshitz (1978).

Finitely generated rings

Matiyasevich asks in the 1970s: what about other rings?

Definition

For a finitely generated ring R, we have analogues of "Hilbert's tenth problem", "Diophantine set" and "listable set" by replacing all occurrences of $\mathbb{Z}$ by R.

Theorem (Mazur–Rubin, 2009)

Assume BSD. Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

Theorem (K.–Pagano, 2024)

Let R be a finitely generated ring with $|R| = \infty$. Then Hilbert's tenth problem is undecidable over R.

We do this by proving the following conjecture of Denef–Lipshitz (1978).

Theorem (K.–Pagano, 2024)

Let K be a number field. Then $\mathbb{Z}$ is Diophantine over O_K.

Proof sketch

By work of Poonen and Shlapentokh, it suffices to construct for every quadratic extension L/K an elliptic curve E/K such that $\mathrm{rk}\, E(K) = \mathrm{rk}\, E(L) > 0$.

Proof sketch

By work of Poonen and Shlapentokh, it suffices to construct for every quadratic extension L/K an elliptic curve E/K such that $\mathrm{rk}\, E(K) = \mathrm{rk}\, E(L) > 0$.

We start with some $\widetilde{E}$ of the shape

$$y^2 = (x - a_1)(x - a_2)(x - a_3).$$

Proof sketch

By work of Poonen and Shlapentokh, it suffices to construct for every quadratic extension L/K an elliptic curve E/K such that $\mathrm{rk}\, E(K) = \mathrm{rk}\, E(L) > 0$.

We start with some $\widetilde{E}$ of the shape

$$y^2 = (x - a_1)(x - a_2)(x - a_3).$$

Then we take the following quadratic twist of $\widetilde{E}$

$$(n - a_1 d)(n - a_2 d)(n - a_3 d)dy^2 = (x - a_1)(x - a_2)(x - a_3),$$

which has the rational point $(x, y) = (n/d, 1/d^2)$ ensuring positivity of the rank.

Proof sketch

By work of Poonen and Shlapentokh, it suffices to construct for every quadratic extension L/K an elliptic curve E/K such that $\mathrm{rk}\, E(K) = \mathrm{rk}\, E(L) > 0$.

We start with some $\widetilde{E}$ of the shape

$$y^2 = (x - a_1)(x - a_2)(x - a_3).$$

Then we take the following quadratic twist of $\widetilde{E}$

$$(n - a_1 d)(n - a_2 d)(n - a_3 d)dy^2 = (x - a_1)(x - a_2)(x - a_3),$$

which has the rational point $(x, y) = (n/d, 1/d^2)$ ensuring positivity of the rank.

In order to make sure that the rank stays stable in L/K, we apply 2-descent.

Proof sketch

By work of Poonen and Shlapentokh, it suffices to construct for every quadratic extension L/K an elliptic curve E/K such that $\mathrm{rk}\, E(K) = \mathrm{rk}\, E(L) > 0$.

We start with some $\widetilde{E}$ of the shape

$$y^2 = (x - a_1)(x - a_2)(x - a_3).$$

Then we take the following quadratic twist of $\widetilde{E}$

$$(n - a_1 d)(n - a_2 d)(n - a_3 d)dy^2 = (x - a_1)(x - a_2)(x - a_3),$$

which has the rational point $(x, y) = (n/d, 1/d^2)$ ensuring positivity of the rank.

In order to make sure that the rank stays stable in L/K, we apply 2-descent.

Note that 2-descent involves the prime factors of $(n - a_1 d)(n - a_2 d)(n - a_3 d)d$; these are controlled via additive combinatorics.

Applications of new techniques

The combination of 2-descent with additive combinatorics has become a very active research area, and I envision many future applications.

Applications of new techniques

The combination of 2-descent with additive combinatorics has become a very active research area, and I envision many future applications.

Theorem (K.–Pagano, 2025)

Let K be a number field. Then there exists an elliptic curve E/K with $\mathrm{rk}\, E(K) = 1$.

Applications of new techniques

The combination of 2-descent with additive combinatorics has become a very active research area, and I envision many future applications.

Theorem (K.–Pagano, 2025)

Let K be a number field. Then there exists an elliptic curve E/K with $\mathrm{rk}\, E(K) = 1$.

Theorem (K.–Morgan, 2025)

Let K be a number field and let $g \geq 1$ be an integer. Then there exists a hyperelliptic curve C/K of genus g such that $\mathrm{rk}\, \mathrm{Jac}(C)(K) = 1$.

Pell's equation

Pell's equation is

$$x^2 - dy^2 = \pm 1 \text{ to be solved in } x, y \in \mathbb{Z}.$$

Pell's equation

Pell's equation is

$$x^2 - dy^2 = \pm 1 \text{ to be solved in } x, y \in \mathbb{Z}.$$

Studied by Archimedes, Pythagoras, Bhaskara II, Brahmagupta, Fermat, Brouncker, Wallis, Euler, Lagrange, Legendre, Gauss, Dirichlet ...

Pell's equation

Pell's equation is

$$x^2 - dy^2 = \pm 1 \text{ to be solved in } x, y \in \mathbb{Z}.$$

Studied by Archimedes, Pythagoras, Bhaskara II, Brahmagupta, Fermat, Brouncker, Wallis, Euler, Lagrange, Legendre, Gauss, Dirichlet ...

Pythagoras proved that $\sqrt{2}$ is irrational, i.e. $x^2 - 2y^2 = 0$ has no solutions in $x, y \in \mathbb{Z}$ (except $x = y = 0$).

Pell's equation

Pell's equation is

$$x^2 - dy^2 = \pm 1 \text{ to be solved in } x, y \in \mathbb{Z}.$$

Studied by Archimedes, Pythagoras, Bhaskara II, Brahmagupta, Fermat, Brouncker, Wallis, Euler, Lagrange, Legendre, Gauss, Dirichlet ...

Pythagoras proved that $\sqrt{2}$ is irrational, i.e. $x^2 - 2y^2 = 0$ has no solutions in $x, y \in \mathbb{Z}$ (except $x = y = 0$).

The Pell equation is instead the "next best thing", namely $x^2 - 2y^2 = \pm 1$. It provides the best rational approximations to $\sqrt{2}$.

Pell's equation

Pell's equation is

$$x^2 - dy^2 = \pm 1 \text{ to be solved in } x, y \in \mathbb{Z}.$$

Studied by Archimedes, Pythagoras, Bhaskara II, Brahmagupta, Fermat, Brouncker, Wallis, Euler, Lagrange, Legendre, Gauss, Dirichlet ...

Pythagoras proved that $\sqrt{2}$ is irrational, i.e. $x^2 - 2y^2 = 0$ has no solutions in $x, y \in \mathbb{Z}$ (except $x = y = 0$).

The Pell equation is instead the "next best thing", namely $x^2 - 2y^2 = \pm 1$. It provides the best rational approximations to $\sqrt{2}$.

Solution x, y	Ratio x/y	Expansion of $\sqrt{2}$
$x = 1,\ y = 1$	1	1.4142135...
$x = 3,\ y = 2$	1.5	1.4142135...
$x = 7,\ y = 5$	1.4	1.4142135...
$x = 17,\ y = 12$	1.4166666...	1.4142135...
$x = 41,\ y = 29$	1.4137931...	1.4142135...
$x = 99,\ y = 70$	1.4142857...	1.4142135...

The positive Pell equation

Dirichlet proved that one can always non-trivially solve

$$x^2 - dy^2 = 1$$

in integers x, y (i.e. $x^2 = 1$ and $y^2 = 0$ being the trivial solution).

The positive Pell equation

Dirichlet proved that one can always non-trivially solve

$$x^2 - dy^2 = 1$$

in integers x, y (i.e. $x^2 = 1$ and $y^2 = 0$ being the trivial solution).

Archimedes seems to have already been aware of this, and the Indian mathematicians even provided an algorithm for solving this equation.

The positive Pell equation

Dirichlet proved that one can always non-trivially solve

$$x^2 - dy^2 = 1$$

in integers x, y (i.e. $x^2 = 1$ and $y^2 = 0$ being the trivial solution).

Archimedes seems to have already been aware of this, and the Indian mathematicians even provided an algorithm for solving this equation.

Example (Fermat's challenge to Brouncker and Wallis)

The smallest non-trivial solution to $x^2 - 61y^2 = 1$ *is*

The positive Pell equation

Dirichlet proved that one can always non-trivially solve

$$x^2 - dy^2 = 1$$

in integers x, y (i.e. $x^2 = 1$ and $y^2 = 0$ being the trivial solution).

Archimedes seems to have already been aware of this, and the Indian mathematicians even provided an algorithm for solving this equation.

Example (Fermat's challenge to Brouncker and Wallis)

The smallest non-trivial solution to $x^2 - 61y^2 = 1$ *is*

$$x = 1766319049, \quad y = 226153980.$$

The negative Pell equation

The negative Pell equation $x^2 - dy^2 = -1$ is more elusive.

The negative Pell equation

The negative Pell equation $x^2 - dy^2 = -1$ is more elusive.

To solve this equation, one certainly needs to be able to solve it modulo p for all primes p. But if p divides d, we get

$$x^2 \equiv -1 \bmod p.$$

The negative Pell equation

The negative Pell equation $x^2 - dy^2 = -1$ is more elusive.

To solve this equation, one certainly needs to be able to solve it modulo p for all primes p. But if p divides d, we get

$$x^2 \equiv -1 \bmod p.$$

This implies that $p \equiv 1 \bmod 4$ or $p = 2$. Define $\mathcal{D}$ to be the set of squarefree d for which $p \mid d$ implies $p \equiv 1 \bmod 4$ or $p = 2$.

The negative Pell equation

The negative Pell equation $x^2 - dy^2 = -1$ is more elusive.

To solve this equation, one certainly needs to be able to solve it modulo p for all primes p. But if p divides d, we get

$$x^2 \equiv -1 \bmod p.$$

This implies that $p \equiv 1 \bmod 4$ or $p = 2$. Define $\mathcal{D}$ to be the set of squarefree d for which $p \mid d$ implies $p \equiv 1 \bmod 4$ or $p = 2$.

Nagell (1930s) conjectured

$$\lim_{X \to \infty} \frac{\#\{d \leq X : d \in \mathcal{D}, x^2 - dy^2 = -1 \text{ sol.}\}}{\#\{d \leq X : d \in \mathcal{D}\}}$$

exists and lies in $(0, 1)$. The smallest $d \in \mathcal{D}$ for which the negative Pell equation is not soluble is $d = 34$.

The negative Pell equation

The negative Pell equation $x^2 - dy^2 = -1$ is more elusive.

To solve this equation, one certainly needs to be able to solve it modulo p for all primes p. But if p divides d, we get

$$x^2 \equiv -1 \bmod p.$$

This implies that $p \equiv 1 \bmod 4$ or $p = 2$. Define $\mathcal{D}$ to be the set of squarefree d for which $p \mid d$ implies $p \equiv 1 \bmod 4$ or $p = 2$.

Nagell (1930s) conjectured

$$\lim_{X \to \infty} \frac{\#\{d \leq X : d \in \mathcal{D}, x^2 - dy^2 = -1 \text{ sol.}\}}{\#\{d \leq X : d \in \mathcal{D}\}}$$

exists and lies in $(0, 1)$. The smallest $d \in \mathcal{D}$ for which the negative Pell equation is not soluble is $d = 34$. Stevenhagen (1995) predicted a precise value for the limit.

Theorem (K.–Pagano, 2022)

Nagell's and Stevenhagen's conjecture are true.

Translating to class groups

Consider the ring $\mathbb{Z}[\sqrt{d}]$. There is an automorphism $\sigma : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}[\sqrt{d}]$ given by $x + y\sqrt{d} \mapsto x - y\sqrt{d}$. Let $N(\alpha) = \alpha\sigma(\alpha)$. Note that

$$x^2 - dy^2 = \pm 1 \iff N(x + y\sqrt{d}) = \pm 1.$$

Translating to class groups

Consider the ring $\mathbb{Z}[\sqrt{d}]$. There is an automorphism $\sigma : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}[\sqrt{d}]$ given by $x + y\sqrt{d} \mapsto x - y\sqrt{d}$. Let $N(\alpha) = \alpha\sigma(\alpha)$. Note that

$$x^2 - dy^2 = \pm 1 \iff N(x + y\sqrt{d}) = \pm 1.$$

The norm map is multiplicative, i.e. $N(\alpha\beta) = N(\alpha)N(\beta)$, thus sends units to units. The only units of $\mathbb{Z}$ are ± 1.

Translating to class groups

Consider the ring $\mathbb{Z}[\sqrt{d}]$. There is an automorphism $\sigma : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}[\sqrt{d}]$ given by $x + y\sqrt{d} \mapsto x - y\sqrt{d}$. Let $N(\alpha) = \alpha\sigma(\alpha)$. Note that

$$x^2 - dy^2 = \pm 1 \iff N(x + y\sqrt{d}) = \pm 1.$$

The norm map is multiplicative, i.e. $N(\alpha\beta) = N(\alpha)N(\beta)$, thus sends units to units. The only units of $\mathbb{Z}$ are ± 1.

Conversely, if the norm is a unit, then the element itself is a unit. Thus negative Pell is soluble if and only if there is a unit of norm -1.

Translating to class groups II

Recall that the negative Pell equation is soluble if and only if there is a unit of norm -1.

Translating to class groups II

Recall that the negative Pell equation is soluble if and only if there is a unit of norm -1.

This is equivalent to the ideal $(\sqrt{d})$ admitting a totally positive generator, i.e. $(\sqrt{d}) = (\alpha)$ with $\alpha > 0$ and $\sigma(\alpha) > 0$.

Translating to class groups II

Recall that the negative Pell equation is soluble if and only if there is a unit of norm -1.

This is equivalent to the ideal $(\sqrt{d})$ admitting a totally positive generator, i.e. $(\sqrt{d}) = (\alpha)$ with $\alpha > 0$ and $\sigma(\alpha) > 0$.

Rephrase this as an equality between the narrow class group (ideals modulo principal ideals with a totally positive generator) and the ordinary class group (ideals modulo principal ideals).

Translating to class groups II

Recall that the negative Pell equation is soluble if and only if there is a unit of norm -1.

This is equivalent to the ideal $(\sqrt{d})$ admitting a totally positive generator, i.e. $(\sqrt{d}) = (\alpha)$ with $\alpha > 0$ and $\sigma(\alpha) > 0$.

Rephrase this as an equality between the narrow class group (ideals modulo principal ideals with a totally positive generator) and the ordinary class group (ideals modulo principal ideals).

Obtain the statistics of the joint distribution of the 2-Sylow subgroup of the narrow class group and ordinary class group (in the style of Cohen–Lenstra).

Translating to class groups II

Recall that the negative Pell equation is soluble if and only if there is a unit of norm -1.

This is equivalent to the ideal $(\sqrt{d})$ admitting a totally positive generator, i.e. $(\sqrt{d}) = (\alpha)$ with $\alpha > 0$ and $\sigma(\alpha) > 0$.

Rephrase this as an equality between the narrow class group (ideals modulo principal ideals with a totally positive generator) and the ordinary class group (ideals modulo principal ideals).

Obtain the statistics of the joint distribution of the 2-Sylow subgroup of the narrow class group and ordinary class group (in the style of Cohen–Lenstra).

We only need to consider the 2-Sylow since $(\sqrt{d})$ has order 2 in the narrow class group. This is the only part of the class group that is well-understood by a recent breakthrough of A. Smith.

Sums of cubes

How many integers $|n| \leq X$ are such that

$$x^3 + y^3 = n$$

has a solution in rational numbers x, y?

Sums of cubes

How many integers $|n| \leq X$ are such that

$$x^3 + y^3 = n$$

has a solution in rational numbers x, y?

Note: it is not hard to show that there are $\leq CX^{2/3}$ integers n for which there is a solution $x, y \in \mathbb{Z}$, for some absolute constant $C > 0$.

Sums of cubes

How many integers $|n| \leq X$ are such that

$$x^3 + y^3 = n$$

has a solution in rational numbers x, y?

Note: it is not hard to show that there are $\leq CX^{2/3}$ integers n for which there is a solution $x, y \in \mathbb{Z}$, for some absolute constant $C > 0$.

Example

For $n = 6$ one can use the factorization $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$ to show that there are no integer solutions. However, we have

$$6 = \left(\frac{17}{21}\right)^3 + \left(\frac{37}{21}\right)^3.$$

Results

Alpöge–Bhargava–Shnidman (2022) showed that a proportion of at least $2/21$ of integers are sums of two rational cubes and at least $1/6$ are not sums of two rational cubes.

Results

Alpöge–Bhargava–Shnidman (2022) showed that a proportion of at least $2/21$ of integers are sums of two rational cubes and at least $1/6$ are not sums of two rational cubes.

Theorem (K.–Smith, 2024)

At least 31.4% *of the integers are not sums of two rational cubes. Assuming a parity conjecture, at least* 47.4% *of integers are sums of two rational cubes.*

Results

Alpöge–Bhargava–Shnidman (2022) showed that a proportion of at least $2/21$ of integers are sums of two rational cubes and at least $1/6$ are not sums of two rational cubes.

Theorem (K.–Smith, 2024)

At least 31.4% *of the integers are not sums of two rational cubes. Assuming a parity conjecture, at least* 47.4% *of integers are sums of two rational cubes.*

Key tool in our work: obtain distribution of 3-Selmer group of $x^3 + y^3 = n$ (ABS obtain average of 2-Selmer).

Results

Alpöge–Bhargava–Shnidman (2022) showed that a proportion of at least $2/21$ of integers are sums of two rational cubes and at least $1/6$ are not sums of two rational cubes.

Theorem (K.–Smith, 2024)

At least 31.4% *of the integers are not sums of two rational cubes. Assuming a parity conjecture, at least* 47.4% *of integers are sums of two rational cubes.*

Key tool in our work: obtain distribution of 3-Selmer group of $x^3 + y^3 = n$ (ABS obtain average of 2-Selmer).

Conjecturally, the limit should be $1/2$.

Theorem (K.–Smith, in progress)

For 100% *of the integers* n*, the rank of* $x^3 + y^3 = n$ *is* 0 *or* 1*. Thus assuming a parity conjecture, exactly* 50% *of integers are sums of two rational cubes.*

Proof overview

As already noted in ABS, it is natural to work with the $\sqrt{-3}$-Selmer group, but this does not have a distribution. This requires a new strong analytic tool. We first recall the bilinear (large) sieve.

Proof overview

As already noted in ABS, it is natural to work with the $\sqrt{-3}$-Selmer group, but this does not have a distribution. This requires a new strong analytic tool. We first recall the bilinear (large) sieve.

Theorem (Heath-Brown, 1995)

We have for all complex numbers $\alpha_m, \beta_n \in \mathbb{C}$ of magnitude at most 1

$$\sum_{m \leq M} \sum_{n \leq N} \alpha_m \beta_n \left(\frac{m}{n}\right) \ll_\epsilon (MN)^{1+\epsilon}(M^{-1/2} + N^{-1/2}).$$

Proof overview

As already noted in ABS, it is natural to work with the $\sqrt{-3}$-Selmer group, but this does not have a distribution. This requires a new strong analytic tool. We first recall the bilinear (large) sieve.

Theorem (Heath-Brown, 1995)

We have for all complex numbers $\alpha_m, \beta_n \in \mathbb{C}$ of magnitude at most 1

$$\sum_{m \leq M} \sum_{n \leq N} \alpha_m \beta_n \left(\frac{m}{n}\right) \ll_\epsilon (MN)^{1+\epsilon}(M^{-1/2} + N^{-1/2}).$$

The key to our work is a new trilinear large sieve. It involves the trilinear Rédei symbol $[a, b, c]$ that measures the splitting of c in a D_4-extension containing $\mathbb{Q}(\sqrt{a}, \sqrt{b})$ (this is an analogue of the "triple linking number" from knot theory).

Proof overview

As already noted in ABS, it is natural to work with the $\sqrt{-3}$-Selmer group, but this does not have a distribution. This requires a new strong analytic tool. We first recall the bilinear (large) sieve.

Theorem (Heath-Brown, 1995)

We have for all complex numbers $\alpha_m, \beta_n \in \mathbb{C}$ of magnitude at most 1

$$\sum_{m \leq M} \sum_{n \leq N} \alpha_m \beta_n \left(\frac{m}{n}\right) \ll_\epsilon (MN)^{1+\epsilon}(M^{-1/2} + N^{-1/2}).$$

The key to our work is a new trilinear large sieve. It involves the trilinear Rédei symbol $[a, b, c]$ that measures the splitting of c in a D_4-extension containing $\mathbb{Q}(\sqrt{a}, \sqrt{b})$ (this is an analogue of the "triple linking number" from knot theory).

Theorem (K.–Smith, 2024)

We have for all $H \geq 3$

$$\left| \sum_{|d_1|<H} \sum_{|d_2|<H} \sum_{|d_3|<H} [d_1, d_2, d_3] \right| \ll_\epsilon H^{3-1/512+\epsilon}.$$

Class groups

Definition

Let R be a commutative domain. Let $I, J \subseteq R$ be non-zero ideals. We write $I \sim J$ if there exist $\alpha, \beta \in R - \{0\}$ such that

$$I \cdot (\alpha) = J \cdot (\beta).$$

The class group $\mathrm{Cl}(R)$ of R is the set of equivalence classes under $\sim$.

Class groups

Definition

Let R be a commutative domain. Let $I, J \subseteq R$ be non-zero ideals. We write $I \sim J$ if there exist $\alpha, \beta \in R - \{0\}$ such that

$$I \cdot (\alpha) = J \cdot (\beta).$$

The class group $\mathrm{Cl}(R)$ of R is the set of equivalence classes under $\sim$.

For nice rings R (Dedekind domains), this is a commutative group. Furthermore, R is an UFD if and only if $\mathrm{Cl}(R)$ is trivial.

Class groups

Definition

Let R be a commutative domain. Let $I, J \subseteq R$ be non-zero ideals. We write $I \sim J$ if there exist $\alpha, \beta \in R - \{0\}$ such that

$$I \cdot (\alpha) = J \cdot (\beta).$$

The class group $\mathrm{Cl}(R)$ of R is the set of equivalence classes under $\sim$.

For nice rings R (Dedekind domains), this is a commutative group. Furthermore, R is an UFD if and only if $\mathrm{Cl}(R)$ is trivial.

Example

We have $\mathrm{Cl}(\mathbb{Z}) = \{0\}$ and $\mathrm{Cl}(\mathbb{Z}[\sqrt{-6}]) \cong \mathbb{Z}/2\mathbb{Z}$.

Class groups

Definition

Let R be a commutative domain. Let $I, J \subseteq R$ be non-zero ideals. We write $I \sim J$ if there exist $\alpha, \beta \in R - \{0\}$ such that

$$I \cdot (\alpha) = J \cdot (\beta).$$

The class group $\mathrm{Cl}(R)$ of R is the set of equivalence classes under $\sim$.

For nice rings R (Dedekind domains), this is a commutative group. Furthermore, R is an UFD if and only if $\mathrm{Cl}(R)$ is trivial.

Example

We have $\mathrm{Cl}(\mathbb{Z}) = \{0\}$ and $\mathrm{Cl}(\mathbb{Z}[\sqrt{-6}]) \cong \mathbb{Z}/2\mathbb{Z}$.

This definition also plays a key role in other areas of mathematics (Picard group, Jacobian etc.).

Why is the class group so important?

David Hilbert

Teiji Takagi

Number theorists are really interested in describing extensions (i.e. covers) of their favorite number ring (like $\mathbb{Z}$, $\mathbb{Z}[\zeta_n]$ or $\mathbb{Z}[\sqrt{-6}]$).

Why is the class group so important?

David Hilbert

Number theorists are really interested in describing extensions (i.e. covers) of their favorite number ring (like $\mathbb{Z}$, $\mathbb{Z}[\zeta_n]$ or $\mathbb{Z}[\sqrt{-6}]$).

The crowning achievement of early 20th century algebraic number theory (Hilbert, Takagi) was class field theory. It describes all abelian extensions of R in terms of $\mathrm{Cl}(R)$.

Teiji Takagi

Statistical questions

Statistical questions about class groups are *exceptionally difficult*.

Statistical questions

Statistical questions about class groups are *exceptionally difficult*.

Gauss already asked if there are infinitely many squarefree integers d such that $\mathrm{Cl}(\mathbb{Z}[\sqrt{d}]) = \{0\}$, i.e. $\mathbb{Z}[\sqrt{d}]$ is an UFD. Completely open!

Statistical questions

Statistical questions about class groups are *exceptionally difficult*.

Gauss already asked if there are infinitely many squarefree integers d such that $\mathrm{Cl}(\mathbb{Z}[\sqrt{d}]) = \{0\}$, i.e. $\mathbb{Z}[\sqrt{d}]$ is an UFD. Completely open!

If one numerically enumerates d such that 9 exactly divides $|\mathrm{Cl}(\mathbb{Z}[\sqrt{-d}])|$, then one sees that the group $\mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ is 8 times less likely than $\mathbb{Z}/9\mathbb{Z}$. Why?

Statistical questions

Statistical questions about class groups are *exceptionally difficult*.

Gauss already asked if there are infinitely many squarefree integers d such that $\mathrm{Cl}(\mathbb{Z}[\sqrt{d}]) = \{0\}$, i.e. $\mathbb{Z}[\sqrt{d}]$ is an UFD. Completely open!

If one numerically enumerates d such that 9 exactly divides $|\mathrm{Cl}(\mathbb{Z}[\sqrt{-d}])|$, then one sees that the group $\mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ is 8 times less likely than $\mathbb{Z}/9\mathbb{Z}$. Why?

Conjecture (Cohen–Lenstra, 1984)

Let p be an odd prime. Let A be a finite abelian p-group. Then

$$\lim_{X\to\infty} \frac{|\{0 < d < X \text{ sqf.} : \mathrm{Cl}(\mathbb{Z}[\sqrt{-d}])[p^\infty] \cong A\}|}{|\{0 < d < X : d \text{ sqf.}\}|} = \frac{\prod_{i=1}^{\infty}\left(1 - \frac{1}{p^i}\right)}{|\mathrm{Aut}(A)|}$$

Statistical questions

Statistical questions about class groups are *exceptionally difficult*.

Gauss already asked if there are infinitely many squarefree integers d such that $\mathrm{Cl}(\mathbb{Z}[\sqrt{d}]) = \{0\}$, i.e. $\mathbb{Z}[\sqrt{d}]$ is an UFD. Completely open!

If one numerically enumerates d such that 9 exactly divides $|\mathrm{Cl}(\mathbb{Z}[\sqrt{-d}])|$, then one sees that the group $\mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ is 8 times less likely than $\mathbb{Z}/9\mathbb{Z}$. Why?

Conjecture (Cohen–Lenstra, 1984)

Let p be an odd prime. Let A be a finite abelian p-group. Then

$$\lim_{X\to\infty} \frac{|\{0 < d < X \text{ sqf.} : \mathrm{Cl}(\mathbb{Z}[\sqrt{-d}])[p^\infty] \cong A\}|}{|\{0 < d < X : d \text{ sqf.}\}|} = \frac{\prod_{i=1}^{\infty}\left(1-\frac{1}{p^i}\right)}{|\mathrm{Aut}(A)|}$$

I will now sample some of my results related to the Cohen–Lenstra heuristics.

Average class numbers

Write $h_\ell(d)$ for the size of the ℓ-torsion of the class group of $\mathbb{Q}(\sqrt{d})$.

Theorem (K.–Pagano–Sofos, 2024)

Let $n = 3 \cdot 2^k$ with $k \geq 1$. Then we have

$$X \log X \ll \sum_{|d| \leq X} h_n(d) \ll X \log X.$$

Average class numbers

Write $h_\ell(d)$ for the size of the ℓ-torsion of the class group of $\mathbb{Q}(\sqrt{d})$.

Theorem (K.–Pagano–Sofos, 2024)

Let $n = 3 \cdot 2^k$ with $k \geq 1$. Then we have

$$X \log X \ll \sum_{|d| \leq X} h_n(d) \ll X \log X.$$

Theorem (K.–Thorner, 2024)

Let ℓ be a prime. Then $\sum_{|d| \leq X} h_\ell(d) \ll_\epsilon X^{\frac{3}{2} - \frac{1}{\ell+1} + \epsilon}$.

Average class numbers

Write $h_\ell(d)$ for the size of the ℓ-torsion of the class group of $\mathbb{Q}(\sqrt{d})$.

Theorem (K.–Pagano–Sofos, 2024)

Let $n = 3 \cdot 2^k$ with $k \geq 1$. Then we have

$$X \log X \ll \sum_{|d| \leq X} h_n(d) \ll X \log X.$$

Theorem (K.–Thorner, 2024)

Let ℓ be a prime. Then $\sum_{|d| \leq X} h_\ell(d) \ll_\epsilon X^{\frac{3}{2} - \frac{1}{\ell+1} + \epsilon}$.

Theorem (K.–Lemke Oliver–Sofos–Thorne, 2025)

There is a constant $C > 0$ such that $\sum_{|d| \leq X} h_6(d) \sim CX \log X$.

Average class numbers

Write $h_\ell(d)$ for the size of the ℓ-torsion of the class group of $\mathbb{Q}(\sqrt{d})$.

Theorem (K.–Pagano–Sofos, 2024)

Let $n = 3 \cdot 2^k$ with $k \geq 1$. Then we have

$$X \log X \ll \sum_{|d| \leq X} h_n(d) \ll X \log X.$$

Theorem (K.–Thorner, 2024)

Let ℓ be a prime. Then $\sum_{|d| \leq X} h_\ell(d) \ll_\epsilon X^{\frac{3}{2} - \frac{1}{\ell+1} + \epsilon}$.

Theorem (K.–Lemke Oliver–Sofos–Thorne, 2025)

There is a constant $C > 0$ such that $\sum_{|d| \leq X} h_6(d) \sim CX \log X$.

Theorem (Chan–K., 2025)

We have $h_3(d) \ll |d|^{0.3194}$.

Bad primes

The Cohen–Lenstra heuristics exclude the prime $p = 2$ (since $\mathrm{Cl}(K)[2]$ is predictable by Gauss genus theory).

Bad primes

The Cohen–Lenstra heuristics exclude the prime $p = 2$ (since $\mathrm{Cl}(K)[2]$ is predictable by Gauss genus theory).

Gerth found an adaptation to include "bad primes". This was proven for quadratic extensions by A. Smith (2017), and by K.–Pagano (2018) for $\mathrm{Cl}(K)[\ell^\infty]$ when K varies over degree ℓ cyclic extensions (conditional under GRH).

Bad primes

The Cohen–Lenstra heuristics exclude the prime $p = 2$ (since $\mathrm{Cl}(K)[2]$ is predictable by Gauss genus theory).

Gerth found an adaptation to include "bad primes". This was proven for quadratic extensions by A. Smith (2017), and by K.–Pagano (2018) for $\mathrm{Cl}(K)[\ell^\infty]$ when K varies over degree ℓ cyclic extensions (conditional under GRH).

It is also natural to study prime parameter families.

Bad primes

The Cohen–Lenstra heuristics exclude the prime $p = 2$ (since $\mathrm{Cl}(K)[2]$ is predictable by Gauss genus theory).

Gerth found an adaptation to include "bad primes". This was proven for quadratic extensions by A. Smith (2017), and by K.–Pagano (2018) for $\mathrm{Cl}(K)[\ell^\infty]$ when K varies over degree ℓ cyclic extensions (conditional under GRH).

It is also natural to study prime parameter families.

Theorem (K.–Milovic, 2017-2018)

The proportion of primes $p \equiv 1 \bmod 4$ *such that* $16 \mid h(-2p)$ *is* $1/8$, *and the proportion satisfying* $16 \mid h(-p)$ *is also* $1/8$.

Bad primes

The Cohen–Lenstra heuristics exclude the prime $p = 2$ (since $\mathrm{Cl}(K)[2]$ is predictable by Gauss genus theory).

Gerth found an adaptation to include "bad primes". This was proven for quadratic extensions by A. Smith (2017), and by K.–Pagano (2018) for $\mathrm{Cl}(K)[\ell^\infty]$ when K varies over degree ℓ cyclic extensions (conditional under GRH).

It is also natural to study prime parameter families.

Theorem (K.–Milovic, 2017-2018)

The proportion of primes $p \equiv 1 \bmod 4$ such that $16 \mid h(-2p)$ is $1/8$, and the proportion satisfying $16 \mid h(-p)$ is also $1/8$.

Moreover, assuming a short character sum conjecture, there does not exist a number field $K/\mathbb{Q}$ and a class function $f : \mathrm{Gal}(K/\mathbb{Q}) \to \{0, 1\}$ such that

$$\mathbf{1}_{16 \mid h(-p)} = f(\mathrm{Frob}_p) \quad \text{for all but finitely many } p.$$

Some future work

Conjecture (Chowla's conjecture)

We have $L(\frac{1}{2}, \chi) \neq 0$ *for all primitive Dirichlet characters* χ.

Some future work

Conjecture (Chowla's conjecture)

We have $L(\frac{1}{2}, \chi) \neq 0$ *for all primitive Dirichlet characters* χ.

There has also been interest in the function field case of this conjecture.

Some future work

Conjecture (Chowla's conjecture)

We have $L(\frac{1}{2}, \chi) \neq 0$ for all primitive Dirichlet characters χ.

There has also been interest in the function field case of this conjecture.

Theorem (Li, 2018)

Let q be an odd prime power. There are infinitely many monic, squarefree polynomials $D \in \mathbb{F}_q[t]$ such that $L(\frac{1}{2}, \chi_D) = 0$.

Some future work

Conjecture (Chowla's conjecture)

We have $L(\frac{1}{2}, \chi) \neq 0$ *for all primitive Dirichlet characters* χ*.*

There has also been interest in the function field case of this conjecture.

Theorem (Li, 2018)

Let q *be an odd prime power. There are infinitely many monic, squarefree polynomials* $D \in \mathbb{F}_q[t]$ *such that* $L(\frac{1}{2}, \chi_D) = 0$*.*

Theorem (K.–Pagano–Shusterman, in progress)

We have $L(\frac{1}{2}, \chi_D) \neq 0$ *for* 100% *of the monic squarefree polynomials* D*.*

Some future work

Conjecture (Chowla's conjecture)

We have $L(\frac{1}{2}, \chi) \neq 0$ for all primitive Dirichlet characters χ.

There has also been interest in the function field case of this conjecture.

Theorem (Li, 2018)

Let q be an odd prime power. There are infinitely many monic, squarefree polynomials $D \in \mathbb{F}_q[t]$ such that $L(\frac{1}{2}, \chi_D) = 0$.

Theorem (K.–Pagano–Shusterman, in progress)

We have $L(\frac{1}{2}, \chi_D) \neq 0$ for 100% of the monic squarefree polynomials D.

Theorem (K.–Smith, in progress)

We have for every finite abelian 2-group M

$$\lim_{X \to \infty} \frac{\sum_{K/\mathbb{Q}(\sqrt{-1}) \text{ quadratic}, D_K \leq X} \#\mathrm{Surj}(2\mathrm{Cl}(K)[2^\infty], M)}{\#\{K/\mathbb{Q}(\sqrt{-1}) \text{ quadratic} : D_K \leq X\}} = \frac{\# \wedge^2 (M)[2]}{\#M}.$$

Questions?

Thank you for your attention! Quick recap:

1. Introduction
2. Hilbert's tenth problem
3. The negative Pell equation
4. Sums of rational cubes
5. Class groups