Some Diophantine problems

Victor Wang

Institute of Mathematics, Academia Sinica

FOUVRY 73, Bernoulli Center, EPFL
31 August 2026

Critical varieties surveyed in the talk

An affine variety $X \subset \mathbb{A}^n$ is *critical* if probability predicts that it should have $B^{o(1)}$ integral points in $[-B, B]^n$ as $B \to \infty$.

Geometrically, this typically means $K_{\widetilde{X}} + D_\infty$ is trivial for some compactification of X.

If in addition the pair $(\widetilde{X}, D_\infty)$ is log canonical then X is called *log Calabi–Yau*. See the recent heuristics of Browning–Wilsch for more details when $n = 3$ (log K3 surfaces).

examples	natural compactification	status
Norm and toric loci: Pell/unit-distance conics; $ab = k$, $xyz = k$, where $k \neq 0$	conic or toric pair	**lc**
Quadratic products: $z^2 = \prod_{i=1}^r P(x_i)$; $\prod_{i=1}^r P(x_i) = \prod_{i=1}^r P(y_i)$; $\deg P = 2$, $r \geq 2$	essentially toric, but with non-equivariant boundary	**lc**
Markoff: $x^2 + y^2 + z^2 - xyz = k$	cubic surface; triangle at infinity	**lc**
Critical diagonal surfaces: $(2, 4, 4), (2, 3, 6), (3, 3, 3)$	weighted-projective closure; anticanonical boundary	**lc**
Critical cubic outlaws: $x(y^2 + xz) = k$; $P(n) = a^2 b$, $\deg P = 3$	cubic closure; anticanonical boundary, non-lc	**non-lc**

For diagonal equations $\sum_i x_i^{d_i} = k$, the critical condition is $\sum_i d_i^{-1} = 1$.

Factoring

How do we quickly factor a large integer $N = pq$? That is, solve the critical quadratic equation

$$ab = N.$$

It is easy to do this in $\ll N^{1/2}$ steps. In fact $N^{o(1)}$ steps typically suffice.

Factoring

How do we quickly factor a large integer $N = pq$? That is, solve the critical quadratic equation

$$ab = N.$$

It is easy to do this in $\ll N^{1/2}$ steps. In fact $N^{o(1)}$ steps typically suffice. Most of the best (publicly available) factoring algorithms use difference of squares:

$$X^2 \equiv Y^2 \bmod N,\ X \not\equiv \pm Y \bmod N \Rightarrow 1 < \gcd(X - Y, N) < N.$$

The General Number Field Sieve (GNFS) uses an auxiliary number field $K = \mathbb{Q}(\theta)$ of *growing* degree

$$d \sim \left(\frac{3 \log N}{\log \log N}\right)^{1/3} \to \infty$$

to construct many potentially nontrivial solutions to the congruence $X^2 \equiv Y^2 \bmod N$, by the pigeonhole principle on friable numbers.

Ahead of the Diophantine curve

GNFS uses an auxiliary field $K = \mathbb{Q}(\theta)$ of degree

$$d \sim \left(\frac{3 \log N}{\log \log N} \right)^{1/3} \to \infty.$$

Friable norm relations eventually give many solutions to

$$X^2 \equiv Y^2 \bmod N.$$

Creators of the GNFS (Buhler, Lenstra, Pomerance, et al.) were ahead of the curve: The recent counterexample[1] to Erdős's unit-distance conjecture also uses growing degree. Frobenius cutting of Hajir–Maire–Ramakrishna gives a totally real Golod–Shafarevich tower with split primes. For suitable constants $D, c > 1$ and fields F in the tower,

$$\#\{(x, y) \in \mathcal{O}_F^2 : x^2 + y^2 = 4D^2\} \geq c^{[F:\mathbb{Q}]}.$$

[1] Trophy first claimed by a team of mathematicians at OpenAI, and later by Anthropic's system "too dangerous" for public use.

The first solution of Pell's equation

Pell is also critical. For a non-square $D > 0$, write

$$\varepsilon_D = t_0 + u_0\sqrt{D} > 1$$

for the least solution $(t_0, u_0) \in \mathbb{Z}^2_{\geq 1}$ of

$$t^2 - Du^2 = 1.$$

The size of ε_D varies enormously with D.

Trivially $\varepsilon_D > 2\sqrt{D}$ and $\varepsilon_D^2 > 4D$. For $\alpha > 0$, let

$$S_f(x, \alpha) = \#\{2 \leq D \leq x : D \neq \square,\ \varepsilon_D \leq D^{1/2+\alpha}\}.$$

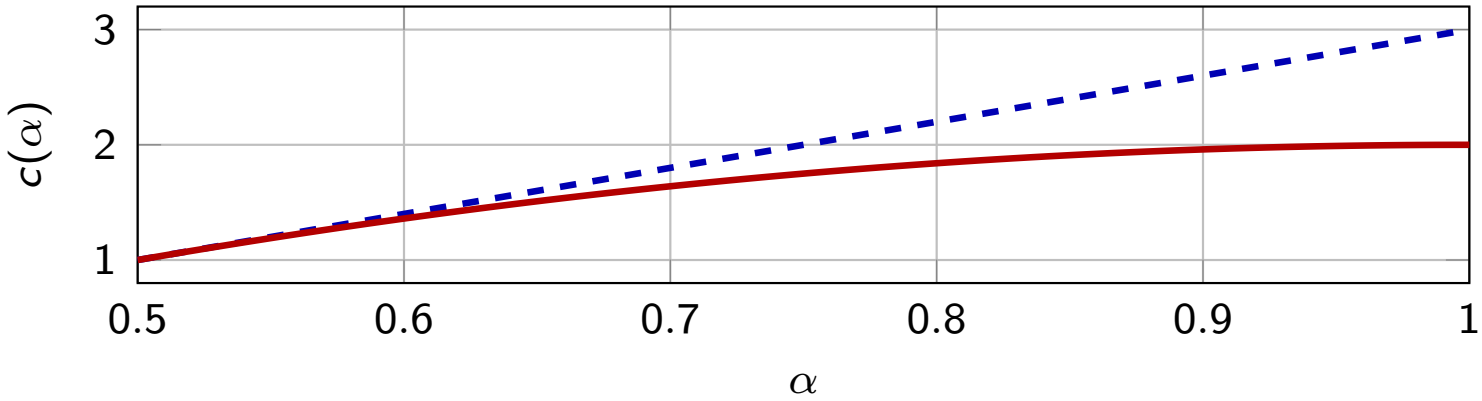
Question: estimate $S_f(x, \alpha)$. Hooley predicts phase transitions at $\alpha = \frac{1}{2}, 1, \frac{5}{2}$, via heuristics on the sizes of $\gcd(t \pm 1, u^2)$.

Hooley and Fouvry

$S_f(x,\alpha) = \#\{2 \le D \le x : D \ne \square,\ \varepsilon_D \le D^{1/2+\alpha}\}$. Hooley proved

$$S_f(x,\alpha) \sim \frac{4\alpha^2}{\pi^2}\sqrt{x}(\log x)^2 \qquad (0 < \alpha \le \tfrac{1}{2}).$$

For $\frac{1}{2} \le \alpha \le 1$, compare coefficients at the scale $\pi^{-2}\sqrt{x}(\log x)^2$:



$c_H(\alpha) = 4\alpha - 1$ (Hooley, conjectural),

$c_F(\alpha) = 4\alpha - 1 - (2\alpha - 1)^2$ (Fouvry, lower bound).

At $\alpha = 1$: $c_H(1) = 3$ and $c_F(1) = 2$.

Hooley (1984); Fouvry (2016)

Hooley's square-root threshold

Idea: Count ALL solutions to $t^2 - Du^2 = 1$ with $u \le D^\alpha$. Put

$$\rho(u) = \#\{\Omega \bmod u^2 : \Omega^2 \equiv 1 \bmod u^2\},$$

so the contribution of a fixed u is

$$\rho(u)\left\{\frac{\sqrt{x}}{u} + O\Big(u^{-1+1/(2\alpha)} + 1\Big)\right\}.$$

With G_2 an explicit 2-adic factor,

$$\sum_{u\ge 1} \frac{\rho(u)}{u^s} = G_2(s)\frac{\zeta(s)^2}{\zeta(2s)}.$$

Summing over $u \le x^\alpha$ gives

$$S_{\text{all}}(x,\alpha) = \frac{4\alpha^2}{\pi^2}\sqrt{x}\log^2 x + O\big(\sqrt{x}\log x + x^\alpha \log x\big).$$

This counts all Pell solutions, so it is an upper bound for $S_f(x,\alpha)$.

- If $\alpha \le \frac{1}{2}$, they are fundamental and the error is lower order.
- If $\alpha > \frac{1}{2}$, the error dominates: one gets only $S_f(x,\alpha) \ll_\alpha x^\alpha \log x$.

Fouvry beyond the threshold

For $\frac{1}{2} < \alpha \le 1$, we have $S_f(x,\alpha) = S_{\mathrm{all}}(x,\alpha) - S_{\mathrm{all}}(x, \frac{2\alpha-1}{4})$ after accounting for non-fundamental solutions ε_D^2. For odd u, split $u = u_1 u_2$, $(u_1, u_2) = 1$, and assign

$$t \equiv 1 \bmod u_1^2, \qquad t \equiv -1 \bmod u_2^2.$$

Fourier analysis of whether the CRT representative t mod u^2 lies in the required interval for $S_{\mathrm{all}}(x,\alpha)$ leads to the problem of obtaining cancellation in sums of the shape

$$\sum_{h \neq 0} \alpha_h \sum_{u_1} \beta_{u_1} \sum_{u_2} e\left(\frac{h \overline{u_2}^2}{u_1^2} \right),$$

where $\overline{u_2}$ is the inverse of u_2 mod u_1^2. Say $u_2 > u_1$. Completing the u_2-sum modulo u_1^2 leads to quadratic Gauss and Kloosterman-type sums; the former are handled using Heath–Brown's large sieve for Jacobi symbols. Some size ranges of u_1, u_2 remain uncontrolled, so one gets a lower bound rather than an asymptotic. Authors building on Fouvry's work include Bourgain, Xi, ...

A result of Besfort Shala

Besfort Shala (IMC 2025) proved

$$\lim_{N\to\infty} \frac{1}{N}\#\{1 \le a, b \le N : (a^2+a)(b^2+b) \text{ is a square}\} = 1.$$

Write

$$a(a+1) = dz^2, \qquad d \text{ squarefree.}$$

Then

$$(2a+1)^2 - d(2z)^2 = 1.$$

If $c_d(N) = \#\{a \le N : a(a+1) \text{ has squarefree core } d\}$, then

$$\sum_d c_d(N) = N.$$

Shala's result asserts the paucity of off-diagonal collisions:

$$\sum_d (c_d(N)^2 - c_d(N)) = o(N).$$

Square products of quadratic values

For a separable quadratic polynomial P, set

$$M_2(N;P) = \#\{n_1 \neq n_2 \leq N : P(n_1)P(n_2) = \square \neq 0\}.$$

Proposition [de la Bretèche–W.–Xu]

For every separable quadratic $P \in \mathbb{Z}[X]$,

$$M_2(N;P) \ll_P N^{1/2}.$$

This is sharp for $P = X^2 - 1$ by Hooley's aforementioned work.

Square products of quadratic values

For a separable quadratic polynomial P, set

$$M_2(N;P) = \#\{n_1 \neq n_2 \leq N : P(n_1)P(n_2) = \square \neq 0\}.$$

Proposition [de la Bretèche–W.–Xu]

For every separable quadratic $P \in \mathbb{Z}[X]$,

$$M_2(N;P) \ll_P N^{1/2}.$$

This is sharp for $P = X^2 - 1$ by Hooley's aforementioned work.

The variety $y^2 = P(x_1)P(x_2)$ is critical, cutting out a double cover of $\mathbb{A}^1 \times \mathbb{A}^1$ with nodal singularities at $P(x_1) = P(x_2) = 0$. This is a log K3 surface, whose compactification via $\mathbb{P}^1 \times \mathbb{P}^1$ is a toric dP4 surface of singularity type $4A_1$. (The recent toric technology of Santens does not apply, because the boundary divisor for integrality is not invariant under the torus action.)

Why the exponent is 1/2

Our proof of $O(N^{1/2})$: Say $P(X) = X^2 - \Delta$. A solution to $P(n_1)P(n_2) = \square$ induces two generalized Pell equations:

$$n_i^2 - dm_i^2 = \Delta, \qquad i = 1, 2.$$

$2 \le d \le \sqrt{N}$. Solutions lie in $O_\Delta(1)$ Pell orbits. Exponential growth of ε_d^k in each orbit gives

$$c_d(N) \ll_\Delta 1 + \frac{\log N}{\log d}.$$

Summing the fiber squares $c_d(N)^2$ costs $O(\sqrt{N})$.

$d > \sqrt{N}$. Compare two points in one orbit. If the smaller of the two satisfies $n \ll \sqrt{N} \ll d$, then count it like Hooley:

$$n^2 \equiv \Delta \bmod m^2.$$

Otherwise, the ratio of the two is $\ll N/\sqrt{N} = \sqrt{N} \ll d$ and is handled similarly.

The proof adapts Hooley's congruences-modulo-squares argument; it does not use a Hooley–Fouvry asymptotic as a black box.

Special curves on the variety $y^2 = P(x_1)P(x_2)$

For $P(X) = X^2 - 1$, the Chebyshev identity

$$T_k(t)^2 - 1 = (t^2 - 1)U_{k-1}(t)^2$$

gives curves on $y^2 = P(x_1)P(x_2)$. For $k = 2$,

$$x_2 = 2x_1^2 - 1, \qquad y = 2x_1(x_1^2 - 1),$$

which already contributes $\asymp N^{1/2}$ points.
More generally, primitive vectors $(r, s) \in \mathbb{Z}^2_{\geq 0}$ give curves

$$(x_1, x_2, y) = \Big(T_r(t), T_s(t), (t^2 - 1)U_{r-1}(t)U_{s-1}(t)\Big).$$

Every point for $M_2(N; X^2 - 1)$ lies on one of these curves: once the squarefree core d of $x_i^2 - 1$ for $i = 1, 2$ is fixed, both Pell points are powers $\varepsilon_d^{gr}, \varepsilon_d^{gs}$ of ε_d, for some $g \geq 1$ and primitive $(r, s) \in \mathbb{Z}^2_{\geq 0}$. We take $t = (\varepsilon_d^g + \varepsilon_d^{-g})/2$.
A similar story may hold whenever $\deg P = 2$?

Random signs and Diophantine equations

Let $P \in \mathbb{Z}[X]$. Even-integer moments of $\sum_{n \le N} f(P(n))$, for a random multiplicative function f, correspond to Diophantine counting problems. Write $P_i = P(n_i)$.

model for f	fourth moment	second moment
Steinhaus (Klurman–Shkredov–Xu)	$P_1P_2 = P_3P_4$	diagonal
Rademacher (Chinis–Shala)	$P_1P_2P_3P_4 = \square$ with each $\mu^2(P_i) = 1$	diagonal on the support $\mu^2 = 1$
extended Rademacher (de la Bretèche–W.–Xu)	$P_1P_2P_3P_4 = \square$	$P_1P_2 = \square$

Here $P_i = P(n_i)$. In the last model, even the variance is a nontrivial Diophantine counting problem, roughly

$$M_2(N; P) = \#\{n_1 \neq n_2 \le N : P(n_1)P(n_2) = \square \neq 0\}.$$

A central limit theorem

Let g be a completely multiplicative Rademacher function: the values $g(p) \in \{\pm 1\}$ are independent and uniform.

Theorem [de la Bretèche–W.–Xu]

Suppose $P \in \mathbb{Z}[X]$ is separable, has positive leading coefficient, $\deg P \geq 2$, and every irreducible factor of P has degree ≤ 3. Then

$$\frac{1}{\sqrt{N}} \sum_{n \leq N} g(P(n)) \implies \mathcal{N}(0,1).$$

For quadratic P, the main new fourth-moment count is

$$M_4(N; P) = \#\{n_1, n_2, n_3, n_4 \leq N \text{ distinct} : \prod_{i=1}^{4} P(n_i) = \square\}.$$

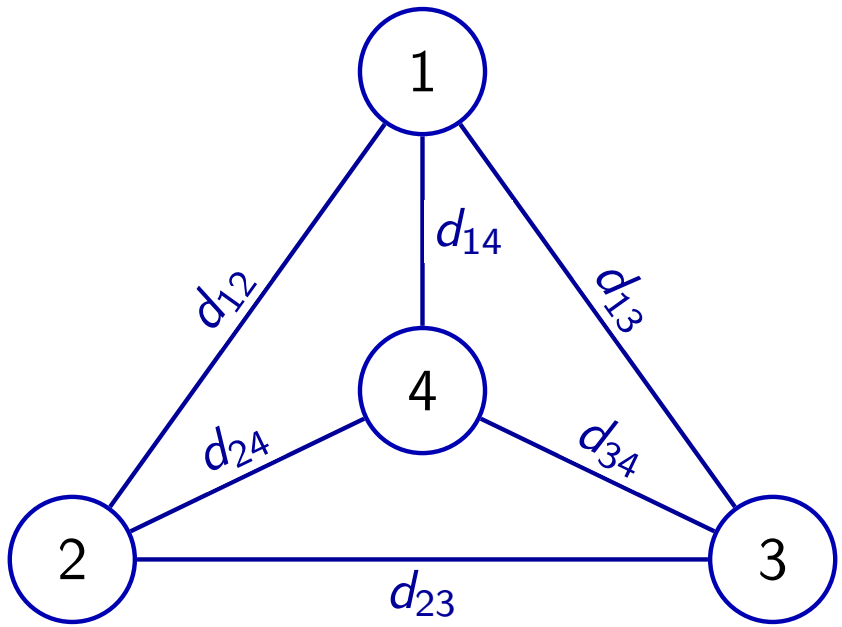
For cubic P, the full, unrestricted fourth moment remains open; we use a trick to first restrict to $P(n_i)/\operatorname{sq}(P(n_i)) \ll N^{\delta}$.

Parametrizing $q_1q_2q_3q_4 = \square$

Write $P(n_i) = q_i m_i^2$, with q_i squarefree. Since $q_1q_2q_3q_4$ is a square, every prime divides either two integers q_i or all four. Thus

$$q_i = d \prod_{j \neq i} d_{ij},$$

with the d, d_{ij} pairwise coprime and squarefree.



d_{ij} records the primes on the edge ij; d records those at all four vertices.

Piecewise linear exponents R, S, and T

We have $P(n_i) = q_i m_i^2$ and $q_i = d \prod_{j \neq i} d_{ij}$. Suppose $d \asymp N^{\alpha}$ and $d_{ij} \asymp N^{\alpha_{ij}}$. Then $q_i \asymp N^{\beta_i}$ where $\beta_i = \alpha + \sum_{j \neq i} \alpha_{ij}$. Let

$$c(\beta, \delta) = \min\left\{\beta - \delta, \max\left(\frac{\beta}{2} - \delta, 1 - \frac{\beta}{2}\right)\right\}.$$

Let $R = (\beta_1 + \beta_2 + \beta_3 + \beta_4)/2$. For distinct indices i, j, k, let

$$S_{ijk} = c(\beta_i, 0) + c(\beta_j, \alpha + \alpha_{ij}) + c(\beta_k, \alpha + \alpha_{ik} + \alpha_{jk}),$$

$$T_{ij} = c(\beta_i, 0) + c(\beta_j, \alpha + \alpha_{ij}) + 2 - R + \frac{\beta_i + \beta_j}{2}.$$

Claim: If $\deg P = 2$, the dyadic contribution to $M_4(N; P)$ is

$$\ll N^{E+\varepsilon}, \qquad E = \min(\min S_{ijk}, \min T_{ij}).$$

- S_{ijk}: Choose q_i, q_j, q_k in order, keeping track of forced common divisors $d_{ij}, \ldots$ at each step. Either use q_i to fix m_i (Pell bounds), or vice versa (Hooley-type congruence).[2]
- T_{ij}: After fixing q_i, q_j, write $AP(n_k) = BP(n_\ell)$, $A \neq B$. Such a conic has $O(N^{\varepsilon})$ points (Hua, Vaughan–Wooley, ...).

[2] Fouvry–Jouve, "Size of regulators..." would improve some ranges.

The exponent 3/2

Exact feasibility check

The exponents are piecewise linear in seven parameters. Microsoft's Z3 Solver verifies over the rationals that

$$\max E = \frac{3}{2}.$$

Thus the off-diagonal fourth moment $M_4(N;P)$ is $O_{P,\varepsilon}(N^{3/2+\varepsilon})$.

In the pure linear program (LP), the equality case $\frac{3}{2}$ is obtained by assigning weights

$$\frac{5}{4}, \quad u, \quad \frac{1}{4} - u \qquad (0 \le u \le \tfrac{1}{4})$$

to the three perfect matchings of K_4; then every $\beta_i = 3/2$. It might be interesting to investigate whether the Diophantine estimates *leading* to the LP are sharp in this case.

Large square factors

Write

$$P(n) = dm^2, \qquad \mu^2(d) = 1.$$

- For small m, the K_4 and curve arguments still give a power saving. For curves $AP(n_k) = BP(n_\ell)$ with $A \neq B$, we use Bombieri–Pila if $\deg P \geq 3$.
- Medium m are controlled by Hooley-type congruences.
- Very large m require uniform point bounds on twists (Evertse–Silverman), together with a power-saving large-square estimate. The latter is unconditional when every irreducible factor of P has degree ≤ 3 (Reuss; see also Booker–Browning). For higher degree, it is an open squarefree-sieve problem.

Over $\mathbb{F}_q[t]$, Ramsay and Poonen control large square factors for every squarefree P, so our work might extend unconditionally to all P in that setting.

Reuss's approximate determinant method

Suppose $\deg P = 3$ (with P irreducible) and that we are in the critical case

$$P(n) = p^2 b, \quad n \asymp X, \quad p \asymp X, \quad b \asymp X.$$

Lift to a cubic field. Let θ be a root of P. Ignoring technicalities,

$$p^2 \mid P(n) \Longrightarrow n - \theta = \alpha^2 \beta, \qquad |\mathrm{N}\,\alpha| = p, \quad |\mathrm{N}\,\beta| = b,$$

after absorbing a unit into β. After unit rescaling, $|\alpha^\sigma| \asymp X^{1/3}\ \forall \sigma$.
Projectivize. Write $\alpha = \sum_{i=0}^{2} x_i \theta^i$, $\beta = \sum_{i=0}^{2} y_i \theta^i$ and $s_i = x_i/x_0$, $t_i = y_i/y_0$. Then $\beta = (n-\theta)/\alpha^2 \sim n/\alpha^2$. Extracting y_0, y_1, y_2 gives

$$t_j = y_j/y_0 = F_j(s_1, s_2) + O(X^{-1}) \qquad (j = 1, 2)$$

for some fixed rational functions F_1, F_2.
Force an exact equation. Box (s_1, s_2) at scale L^{-1}. For 9 points in one box, the 9×9 determinant Δ formed from $s_i t_j$ has denominator $\asymp X^6$, yet $\Delta \ll L^{-11} X^{-2}$. If $L \ggg X^{4/11}$, then $\Delta = 0$. The matrix kernel gives a bilinear equation $C(\boldsymbol{x}; \boldsymbol{y}) = 0$, leading to a saving if $L \ll X^{1/2-\delta}$.

Reuss, "Power-free values of polynomials" (2015), §§3–6.

Value sets and additive collisions

After multiplicative moments, how about additive moments? For a multivariate polynomial M and a finite region $\mathcal{R}_B$, let

$$r(k) = \#\{\boldsymbol{x} \in \mathcal{R}_B : M(\boldsymbol{x}) = k\}.$$

Here are 3 quantities of interest: Q_0, Q_1, Q_2.

$$Q_0 = \#\{k : r(k) > 0\}, \qquad Q_1 = \sum_k r(k) = \#\mathcal{R}_B,$$

$$Q_2 = \sum_k r(k)^2 = \#\{\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{R}_B : M(\boldsymbol{x}) = M(\boldsymbol{x}')\}.$$

By Cauchy–Schwarz,

$$Q_0 \geq \frac{Q_1^2}{Q_2}.$$

The choice of region $\mathcal{R}_B$ is part of the problem: a box for the multiplication table problem, a reduced fundamental region for Markoff surfaces, or a positive orthant in Waring's problem.

Example 1: The multiplication table (critical)

Let $r(k) = \#\{1 \le m, n \le B : mn = k\}$. Then

$$Q_1 = \sum_k r(k) = B^2$$

and there are many ways to show that

$$Q_2 = \sum_k r(k)^2 = \#\{mn = m'n'\} \asymp B^2 \log B.$$

Thus $Q_0 \gg B^2/\log B$. Ford proved the correct order

$$Q_0 \asymp \frac{B^2}{(\log B)^\delta (\log\log B)^{3/2}}, \qquad \delta = 1 - \frac{1 + \log\log 2}{\log 2}.$$

Green–Sawhney recently proved an asymptotic formula for Q_0.

Ford (2008); Green–Sawhney (2026)

Example 2: Markoff surfaces (critical)

Let $M = x^2 + y^2 + z^2 - xyz$, and let $h_M(k)$ be the number of Markoff-group orbits on $M = k$. It can be approached via

$$r(k) = \#\{\boldsymbol{x} \in \mathcal{R}_B : M(\boldsymbol{x}) = k\}$$

for well-chosen $\mathcal{R}_B$. The moment $Q_2 = \sum_k r(k)^2$ is decisive.

- Ghosh–Sarnak construct reduced representatives and prove a well-chosen variance estimate against truncated local densities. Thus $h_M(k) \to \infty$ for a density 1 of admissible k.
- Mishra moves into a wider $\mathbb{G}_m^2$-region and proves a denser version of the variance estimate. For any fixed $\varepsilon > 0$,

  $$(\log |k|)^{2-\varepsilon} \leq h_M(k) \leq (\log |k|)^{2+\varepsilon}$$

  for a density 1 of admissible k. The lower bound is new; the upper bound follows from $Q_1 = \sum_k r(k) = \#\mathcal{R}_B$.

Fixing z_1, z_2: the diagonal uses divisor estimates; the off-diagonal in both papers uses *Kloosterman's circle method*.

Ghosh–Sarnak (2017); Mishra (2024)

Example 3: Critical equations with squares

Recently, de la Bretèche and Tenenbaum made significant progress on critical equations with squares. Restricting to ternary examples (log K3 surfaces) for simplicity, this includes

$$y^2 + n_1^4 + n_2^4 = k \qquad \text{and} \qquad y^2 + n_1^3 + n_2^6 = k,$$

where $y, n_1, n_2 \geq 0$. Both are critical:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1, \qquad \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1.$$

The moment $Q_2 = \sum_k r(k)^2$ is again decisive. A collision gives

$$(y - y')(y + y') = (n_1')^a + (n_2')^b - n_1^a - n_2^b$$

where $(a, b) = (4, 4)$ or $(3, 6)$. Using progress on the Erdős–Hooley Δ-function, for both families, they show $Q_2 \ll x(\log\log x)^{5/2}$,

$$Q_0 = \#\{k \leq x : r(k) > 0\} \gg \frac{x}{(\log\log x)^{5/2}}.$$

de la Bretèche–Tenenbaum (2026)

Example 4: Ternary cubic forms (critical)

Let $Q_2(B)$ count solutions of $C(\boldsymbol{x}) = C(\boldsymbol{x}')$ in $[1, B]^6$.

$C(\boldsymbol{x})$	local geometry	global second moment
xyz	toric; focused	$Q_2(B) \sim cB^3(\log B)^4$ (classical)
$x(y^2 + xz)$	focused; the doubled form satisfies BTSC	An asymptotic for $Q_2(B)$ may be amenable to methods of Browning–Munshi–W.[3]
$x^3 + y^3 + z^3$	balanced; Fermat cubic fails BTSC	$Q_2(B) \ll \frac{B^{7/2}}{(\log B)^{5/2-\varepsilon}}$ (Vaughan), or $B^{3+o(1)}$ (Hooley, under GRH), or B^3 (W., under Ratios+SFSC)

[3] Using the circle method, we recently proved Manin–Peyre for the Perazzo-type cubic fourfold $\sum_{i=1}^{3} x_i y_i^2 = 0$, using a local "better than square-root cancellation" (BTSC) property, results on the Erdős–Hooley Δ-function as in de la Bretèche–Tenenbaum, and geometric insights on dual varieties. (This is a cubic hypersurface beyond the square-root barrier. For quadrics beyond the barrier we refer to recent work of Huang. For certain systems see Wooley, Brüdern–Wooley, Browning–Munshi, et al.)

Fourier second moments

Let $V \subset \mathbb{A}^n_{\mathbb{F}_p}$ be geometrically irreducible of dimension d. Put

$$T(\boldsymbol{b}) = p^{-d/2} \sum_{\boldsymbol{x} \in V(\mathbb{F}_p)} \mathrm{e}_p(\boldsymbol{b} \cdot \boldsymbol{x}),$$

so that probability predicts $T(\boldsymbol{b}) \ll 1$. By Plancherel,

$$p^{-n} \sum_{\boldsymbol{b} \in \mathbb{F}_p^n} |T(\boldsymbol{b})|^2 = p^{-d} |V(\mathbb{F}_p)| = 1 + O(p^{-1/2})$$

by Lang–Weil, uniformly over V of bounded complexity. This fixes the total ℓ^2-mass of $T(\boldsymbol{b})$, but not where it lies.

Heuristically there are two natural possibilities:

- $|T(\boldsymbol{b})| \asymp 1$ on a positive proportion of frequencies $\boldsymbol{b}$;
- large values $|T(\boldsymbol{b})| \asymp p^{c/2}$ concentrated on a thin algebraic set of frequencies $\boldsymbol{b}$ of codimension $c \geq 1$ in affine space.

Fouvry–Katz stratification

Fix $V \subset \mathbb{A}^n_{\mathbb{Z}}$ with $\dim V_{\mathbb{C}} = d$. Fouvry and Katz construct closed sets

$$\mathbb{A}^n_{\mathbb{Z}} = X_0 \supset X_1 \supset X_2 \supset \cdots, \qquad \operatorname{codim} X_j \geq j,$$

such that, away from X_j,

$$T(\boldsymbol{b}) = p^{-d/2} \sum_{\boldsymbol{x} \in V(\mathbb{F}_p)} \mathrm{e}_p(\boldsymbol{b} \cdot \boldsymbol{x}) \ll p^{(j-1)/2}.$$

This formalizes the heuristic on the last slide. On $X_{j-1} \setminus X_j$,

$$p^{-n} \sum_{\boldsymbol{b} \in (X_{j-1} \setminus X_j)(\mathbb{F}_p)} |T(\boldsymbol{b})|^2 \ll p^{-n} p^{n-j+1} p^{j-1} = 1.$$

In some cases, the only non-trivial geometric jump X_j is the dual variety of the original V, followed by the zero frequency. In other cases, more strata are genuinely involved.

Fouvry–Katz (2001), building on various works of Fouvry, Katz, and Laumon

Focused and balanced

For an absolutely irreducible cubic form F in 6 variables, let

$$T(\boldsymbol{b}) = p^{-5/2} \sum_{F(\boldsymbol{x}) \equiv 0 \bmod p} \mathrm{e}_p(\boldsymbol{b} \cdot \boldsymbol{x}),$$

so that $p^{-6} \sum_{\boldsymbol{b} \bmod p} |T(\boldsymbol{b})|^2 \asymp 1$.

Focused

A hypersurface $G(\boldsymbol{b}) = 0$ carries asymptotically all the ℓ^2-mass. The Perazzo-type cubic $F = \sum_{i=1}^{3} x_i y_i^2$ is an example, to be explained on a later slide.

Balanced

Every hypersurface carries a vanishing share of the ℓ^2-mass. The smooth six-variable Fermat cubic $F = \sum_{i=1}^{6} x_i^3$ is an example.

BTSC and the degree of the dual

$T(\boldsymbol{b}) = p^{-5/2} \sum_{F(\boldsymbol{x}) \equiv 0 \bmod p} \mathrm{e}_p(\boldsymbol{b} \cdot \boldsymbol{x})$. Say that BTSC holds for a cubic form F in 6 variables if some non-zero $G \in \mathbb{Z}[\boldsymbol{b}]$ satisfies

$$p \nmid G(\boldsymbol{b}) \quad \Longrightarrow \quad T(\boldsymbol{b}) \ll p^{-1/2}.$$

F	BTSC	dual degree
Perazzo primal: $xyz - uvw$	yes	3
$x(y^2 + xz) - u(v^2 + uw)$	yes	5
Perazzo-type: $\sum_{i=1}^{3} x_i y_i^2$	yes	6
$\sum_{i=1}^{3} x_i y_i^2 - c_1 x_1 x_2 x_3 - c_2 y_1 y_2 y_3$, for coefficients $c_1(c_1 c_2^2 - 4) \neq 0$	yes	6
smooth Fermat: $\sum_{i=1}^{6} x_i^3$	no	48

Dual degree correlates with BTSC, but is not a criterion.

Dual-degree data: github.com/wangyangvictor/cubic_4folds

The Perazzo-type complete sum

$T(\boldsymbol{m},\boldsymbol{n}) = p^{-5/2}\sum_{F(\boldsymbol{x},\boldsymbol{y})\equiv 0 \bmod p} \mathrm{e}_p(\boldsymbol{m}\cdot\boldsymbol{x}+\boldsymbol{n}\cdot\boldsymbol{y})$. For

$$F(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^{3} x_i y_i^2,$$

the complete sum T roughly satisfies

$$T(\boldsymbol{m},\boldsymbol{n}) = -p^{-1/2}\left(1+\sum_{1\le i<j\le 3}\left(\frac{m_i m_j}{p}\right)\right) \ll p^{-1/2}$$

off the sextic dual hypersurface

$$D(\boldsymbol{m},\boldsymbol{n}) = 2\sum_{i=1}^{3} m_i^2 n_i^4 - \left(\sum_{i=1}^{3} m_i n_i^2\right)^2 = 0.$$

Thus BTSC holds for F. On $D=0$ roughly $|T(\boldsymbol{m},\boldsymbol{n})| \asymp p^{1/2}$, carrying the focused ℓ^2-mass.

A deformation of the Perazzo-type example

$T(\boldsymbol{b}) = p^{-5/2} \sum_{F(\boldsymbol{x}) \equiv 0 \bmod p} \mathrm{e}_p(\boldsymbol{b} \cdot \boldsymbol{x})$. Consider

$$F = \sum_{i=1}^{3} x_i y_i^2 - c_1 x_1 x_2 x_3 - c_2 y_1 y_2 y_3.$$

If $c_1(c_1 c_2^2 - 4) \neq 0$, the singular locus is a union of three conics, and a generic hyperplane section of $F = 0$ in $\mathbb{P}^5$ is a six-nodal cubic threefold. Formulas of Segre, Dolgachev, W., . . . make

$$T(\boldsymbol{b}) \ll p^{-1/2}$$

away from a fixed hypersurface in $\boldsymbol{b}$. Thus BTSC holds.[4] The dual of F is sextic. If $c_1 = 0$, cf. Browning–Munshi–W. If $c_1 c_2^2 = 4$, then $F = \det(\text{symmetric } 3 \times 3 \text{ matrix})$, which was studied by Schmidt, Derenthal, . . .; its dual is the degree-4 Veronese surface.

[4] A proof by Gauss and Salié sums is also possible, at least if $c_2 = 0$.

Big picture when applying Fouvry–Katz

There are often two key questions.

1. Where does the stratification stop?
   A separate uniform pointwise bound may be critical for avoiding deeper strata. Selected analogies:
   - 1.1 In the Bombieri–Vinogradov theorem, a pointwise Siegel–Walfisz/zero-free-region input handles small conductors; the large sieve inequality (analogous to Fouvry–Katz in spirit) treats the rest.
   - 1.2 In work of Kubilius, Maknys, Ricci, Duke, Thorner, Stucky, ... on norm-form primes in thin regions, a uniform zero-free region for certain families of Hecke L-functions near the line $\sigma = 1$ is needed to combat logarithmic losses from relevant large sieve inequalities.
2. How many integral frequencies lie on each stratum?
   Dimension growth bounds, determinant methods, divisor bounds, geometry of numbers, ... are often useful.

A general application of Fouvry–Katz

Let $V \subset \mathbb{A}^N_{\mathbb{Z}}$, with $\dim V_{\mathbb{C}} = D$. Suppose that, for a positive-density set $\mathcal{P}$ of primes and every $\mathbf{0} \neq \boldsymbol{c} \in \mathbb{F}_p^N$,

$$\sum_{\boldsymbol{x} \in V(\mathbb{F}_p)} e_p(\boldsymbol{c} \cdot \boldsymbol{x}) \ll_V p^{D-L} \qquad (p \in \mathcal{P}),$$

where $2L \in \mathbb{Z}$. Then

$$\#(V(\mathbb{Z}) \cap [-B, B]^N) \ll_V B^{D-L+\frac{L^2}{N-D+L}}.$$

If $N - D \geq 4$ and $L \geq \frac{3}{2}$, this improves the dimension-growth bound $B^{D-1+\varepsilon}$. In [Browning–Sawin–W.], we used this strategy to prove new bounds for $V =$ the commuting variety $XY = YX$ of $n \times n$ matrices X, Y. Many natural questions remain: Can we average over p or use composite moduli? Can we understand not just the cutoff $D - L$, but the structure of sums near the cutoff?

Browning–Sawin–W.; cf. Fujiwara, Shparlinski, Skorobogatov, and Heath-Brown

Joyeux anniversaire, Étienne !