Sums of three squares: Beyond equidistribution

Zeev Rudnick, Tel Aviv University
joint with Jean Bourgain & Peter Sarnak (IAS)
Integers as sums of 3 squares

Legendre/Gauss: \( n = x^2 + y^2 + z^2 \iff n \not= 4^a (8b+7) \)

**Primitive representation**: \( \gcd(x, y, z) = 1 \)

\( n \) is primitively represented as a sum of 3 squares \( \iff n \not= 0, 4, 7 \mod 8 \)

**The number of representations**

\[
N_n := \# \left\{ (x, y, z) : x^2 + y^2 + z^2 = n \right\}
\]

**Example**: \( n = 5 \) then we have \( N_5 = 24 \)

\[
(\pm 2, \pm 1, 0), (\pm 1, \pm 2, 0), (\pm 2, 0, \pm 1),
(\pm 1, 0, \pm 2), (0, \pm 2, \pm 1), (0, \pm 1, \pm 2)
\]

**Exercise**: if \( n = 4^a \) then \( N_n = 6 \)
Growth of $N_n$ for large $n$

For squarefree $n$,  
$$N_n \approx n \cdot L(1, \chi_{-n})$$

Siegel:  
$$\frac{1}{n^\varepsilon} \ll L(1, \chi_{-n}) \ll n^\varepsilon, \quad \forall \varepsilon > 0$$

If $n$ is primitively representable as a sum of three squares then  
$$N_n = n^{1/2 + o(1)}$$
Spatial distribution of solutions

Project the different representations of $n$ to the unit sphere $S^2$:

$$(x, y, z) \mapsto \frac{1}{\sqrt{n}} (x, y, z) \in S^2$$

We get a set $E(n)$ of about $N_n$ points on $S^2$
- call them “Linnik points”

**Uniform distribution** (Linnik’s conjecture)

As $n \to \infty$, $n \neq 0, 4, 7$ mod 8, the set $E(n)$ becomes uniformly distributed on $S^2$.

Proved by Linnik assuming GRH, (1940),
Definition: The sets $E(n)$ become uniformly distributed in $S^2$ if for any nice set $B$ in $S^2$:

$$\frac{\#(E(n) \cap B)}{\# E(n)} \xrightarrow{n \to \infty} \frac{\text{area}(B)}{\text{area}(S^2)}$$

Equivalently, for any continuous function $f \in C(S^2)$,

$$\frac{1}{\# E(n)} \sum_{P \in E(n)} f(P) \xrightarrow{n \to \infty} \frac{1}{\text{area}(S^2)} \int_{S^2} f(x) \, dx$$
The electrostatic energy of \( N \) points on the sphere \( S^2 \) is

\[
\text{Energy}(P_1, \ldots, P_N) := \sum_{i=1}^{N} \sum_{j \neq i} \frac{1}{|P_i - P_j|}
\]

**Thompson’s question:** (1904): Find configurations of charges on the sphere which **minimize** \( E \) (stable configurations)

**Wagner (1992):** The energy of stable configurations is

\[
\text{Energy}_{\text{min}} = N^2 \int \int_{S^2 \times S^2} \frac{dx\,dy}{|x - y|} - O(N^{3/2})
\]

**Question:** What is the energy for Linnik points \( E(n) \)?

J.J. Thompson, Nobel prize 1903

The energy of Linnik points

**Theorem** (Bourgain, ZR, Sarnak): The energy of the Linnik points $E(n)$ is close to minimal

$$\text{Energy}(E(n)) = N^2 + O(N^{2-\delta})$$

Proof: equidistribution + control of # of close neighbours

$$\text{Energy}(E(n)) := \sum_{P \in E(n)} \left( \sum_{Q \neq P} \frac{1}{|P - Q|} \right)$$

Would like to use uniform distribution to claim that for each $P$

$$\sum_{Q \in E(n)} \frac{1}{|P - Q|} \sim N \int_{S^2} dx \frac{1}{|P - x|} = N \quad \text{Energy} \sim \sum_{P \in E(n)} N = N^2$$

Problem: The function $Q \rightarrow 1/|P-Q|$ is not continuous!

In fact a point $Q$ with $|P-Q|<1/N^{1-o(1)}$, gives a contribution bigger than main term $N$
Example: A close pair

\[ n = k^2 + (k + 1)^2, \quad P = \frac{1}{\sqrt{n}} (k + 1, k, 0) \quad Q = \frac{1}{\sqrt{n}} (k, k + 1, 0) \]

\[ |P - Q| = \frac{1}{\sqrt{n}} = \frac{L(1, \chi_n)}{N} = \frac{1}{N^{1+o(1)}} \]
Controlling close pairs

Counting pairs of points at a given distance

\[ A(n, h) := \# \left\{ x, y \in \mathbb{Z}^3 : |x|^2 = |y|^2 = n, \quad |x-y|^2 = h \right\} \]

Siegel’s mass formula:

\[ A(n, h) = 24 \cdot \alpha_2(n, h) \prod_{p \nmid h(2n-h), \ p \neq 2} \alpha_p(n, h) \]

Venkov, 1931, Pall, 1948: Explicit computation of local factors (crucial).

\[ A(n, h) \ll n^\varepsilon \gcd(n, h)^{1/2} \]

- allows to control contribution of “close” pairs and show Energy(E(n))~N^2. QED
Beyond equidistribution II: randomness on smaller scales

Uniform distribution means randomness on scale of $O(1)$ – subsets in $S^2$ of fixed size.

Question: randomness on smaller scales?
Analyzing point patterns

Fig. 1. 128 young pines in a 10 m x 10 m square plot in a Finnish forest. The trees are shown as circles, the diameters of which are proportional to the tree heights. This pattern is highly clustered; some trees are very close together.

Fig. 2. 134 spruces of age 60 years in a 56 m x 38 m stand of Tharandter Wald (Germany). The trees are shown as circles, where the diameters are proportional to the bdw's.
Point processes on $S^2$

Random point process on $S^2$ = random countable subset $X$ in $S^2$

Intensity of a point process: For any subset $A$ of $S^2$, $E(\#A \cap X) = \mu \cdot \sigma(A)$

$\sigma(A) := \text{area}(A)/\text{area}(S^2)$

Binomial process: $N$ independent points, each uniformly distributed on $S^2$:

$$\text{Prob}_{\text{Bin}(N)}(\# A \cap X = k) = \binom{N}{k} \sigma(A)^k (1 - \sigma(A))^{N-k}$$

Poisson process of intensity $\mu$: characterized by

1) $\# A \cap X$ is a Poisson variable with intensity $\mu \cdot \sigma(A)$

2) If $A, B$ are disjoint then $\# A \cap X$, $\# B \cap X$ are independent

$$\text{Prob}_{\text{Pois}(\mu)}(\# A \cap X = k) = e^{-\mu \cdot \sigma(A)} \frac{\mu \cdot \sigma(A)^k}{k!}$$
Tests for randomness: Minimal distances

Nearest neighbor distance:

$$\delta(x) := \min_{y \neq x} |x - y|^2$$

Average value:

$$\left< \delta(x) \right>_n := \frac{1}{N_n} \sum_{x \in E(n)} \delta(x)$$

For random points (Poisson model)

1) $$\langle \delta \rangle \sim 4/N$$

2) Distribution of normalized distances $$\delta(x)/\langle \delta \rangle$$ is exponential
**Distribution of minimal distances**

**Conjecture:** For the Linnik points $E(n)$

1) $\langle \delta \rangle_n \sim 4/N_n = \text{Poisson value}$

2) Distribution $P(s)$ of normalized distances $\delta(x)/\langle \delta \rangle$ is exponential: $P(s) = \exp(-s)$

**Proposition:** Any possible limit $P(s)$ is absolutely continuous (assuming GRH)
The least spacing statistic

Because the point sets \( E(n) \) come from integer points, the least spacing is \( >1/\sqrt{n} \)

\[
\delta_{\min}(n) := \min_{x \neq y} |x - y| > \frac{1}{\sqrt{n}} = N^{-1 + o(1)}
\]

Is this a non-random feature ?

Compare: For \( N \) random points, least spacing is a.s. \( N^{-1 + o(1)} \) (lower & upper bound) -- “birthday paradox”

**Theorem:** For almost all \( n \), \( \delta_{\min}(n) \approx n^{-1/2 + o(1)} = N^{-1 + o(1)} \)

Implied by: almost all \( n \) is a sum of two squares and a “mini-square”

\[
n = x^2 + y^2 + z^2, \quad |z| < n^\varepsilon
\]

Wooley: can make \( |z| < (\log N)^{1 + o(1)} \)
Ripley’s K-function: for a homogeneous point process with intensity $\mu$,

$\mu K(r) := \text{expected number of other points at distance } r \text{ from a typical point of the process}$

Alternatively,

$\mu^2 K(r) = \# \text{ distinct pairs (x,y) with } |x-y| < r$

For Poisson process, $E\{K(r)\} = \pi r^2 = \text{area}$

A smaller value of $K(r)$ is interpreted as repulsion (e.g. for determinantal processes).
The critical scaling regime

we look at values of $r$ where a ball of radius $r$ is expected to contain $O(1)$ points:

Define **critical radius** $\rho_n$ by

$$N_n \cdot \pi \rho_n^2 = \text{area}(S^2)$$

We conjecture that at the critical scale, the two-point statistic $K$ coincides with that of random points

Conjecture: Fix $\lambda>0$. For $r=\sqrt{\lambda} \cdot \rho_n$ then $K_n(r) \sim \pi r^2$ as $n \to \infty$

$$\lim_{n \to \infty} N_n K_n(\sqrt{\lambda} \rho_n) = \lambda$$
$K(r) \text{ at critical scaling}$

$$\lim_{n \to \infty} N_n K_n(\sqrt{\lambda} \rho_n) = \lambda$$

$n=10^7+1^{\text{th}} \text{prime, } N \approx 100,000 \text{ (effectively 2000 points)}$
bounds at critical scaling

Theorem (on GRH): For $N^{-1+o(1)} < r < 2$, $K_n(r) < 2010 \cdot r^2$

Theorem: For a positive density of square-free $n$, we have

$$N_n K_n(\sqrt{\lambda} \rho_n) > \frac{1}{2010} \lambda$$

Note: the upper bound holds for individual $n$

Thus we have lower & upper bounds on $K(r)$ consistent with random behavior.
Main tool: Venkov’s formula

Arithmetic ingredient: Counting pairs of vectors with length $^2=n$ and given distance:

$$K_n(r) = \frac{1}{N^2} \# \left\{ x, y : \| x \|^2 = \| y \|^2 = n, \quad \frac{x-y}{\sqrt{n}} \leq \lambda \frac{4}{N} \right\} = \frac{1}{N^2} \sum_{h<\frac{4n}{N}} A(n, h)$$

$$A(n, h) := \# \left\{ x, y \in \mathbb{Z}^3 : \| x \|^2 = \| y \|^2 = n, \quad \| x - y \|^2 = h \right\}$$

Siegel’s mass formula

$$A(n, h) = 24 \cdot \alpha_2(n, h) \prod_{p \mid h(2n-h)} \alpha_p(n, h)$$

$$\alpha_p(n, h) = \lim_{k \to \infty} \frac{1}{p^{3k}} \# \left\{ x, y \mod p^k : \| x \|^2 = \| y \|^2 = n, \quad \| x - y \|^2 = h \right\}$$

Venkov, 1931, Pall, 1948 : Explicit computation of local factors (crucial).
Approximating $A(n,h)$

$A(n,h)$ is essentially a multiplicative function along a quadratic progression

$$A(n,h) \leq 24 \cdot F_n \left( h(2n - h) \right)$$

$$F_n(a) \approx \sum_{d \mid a, \text{ odd}} \left( \frac{-n}{d} \right)$$

This allows us to use results about sums of multiplicative functions along polynomial progressions (Shiu, Nair) to get upper bound on $K(r)$
Nair’s theorem (1992)

F=multiplicative function: $F(1)=1$, $F(ab)=F(a)F(b)$ if a,b coprime,

F non-negative: $F\geq 0$, slowly growing: $F(n)\ll n^\varepsilon$

$P(t)\in \mathbb{Z}[t]$ polynomial

Then for $x^a < y < x$

$$\sum_{x-y < m < x} f(\lfloor P(m) \rfloor) \ll_{F,P} y \times \prod_{p \leq x} \left(1 - \frac{\rho(p)}{p}\right) \times \exp\left(\sum_{p \leq x} \frac{\rho(p)f(p)}{p}\right)$$

$$\rho(m) = \# \left\{ x \in \mathbb{Z} / m\mathbb{Z} : \ P(x) = 0 \mod m \right\}$$

In our case:

$$P(t) = t(2n-t) , \quad f(a) = F_n(a) \approx \sum_{d|a \atop d \text{ odd}} \left(\frac{-n}{d}\right) , \quad x = n, \quad y = \lambda \frac{4n}{N_n} \approx n^{1/2+o(1)}$$
Why do we need GRH?

Nair’s theorem gives a bound on $K(r)$ of

$$N_n K_n (\rho_n \sqrt{\lambda}) \ll \lambda \times \left( \frac{1}{L(1, \chi_{-n})} \exp \left( \sum_{p < n} \frac{\chi_{-n}(p)}{p} \right) \right)^2$$

The L-value arises because $N_n \approx \sqrt{n} \cdot L(1, \chi_n)$

- Need GRH to guarantee RHS is $\ll \lambda$
We studied properties of the sets $E(n)$ of points on the sphere arising from writing $n = x^2 + y^2 + z^2$ which go beyond uniform distribution:

• The electrostatic energy is close to minimal.
• Empirical agreement of various statistics with those of a random point process.
• Some theoretical results.