The Generalized Ramanujan Conjecture and the Sarnak–Xue Density Hypothesis

Henrik Gustafsson
hgustafsson.se

Department of Mathematics and Mathematical Statistics
Umeå University, Sweden

Algebraic geometry and number theory-seminar

Chalmers University of Technology and University of Gothenburg

November 18, 2025

The Ramanujan Conjecture

The Ramanujan conjecture for Δ

The discriminant modular form $\Delta : \mathscr{H} \to \mathbb{C}$ is the unique (up to normalization) cusp form of weight 12 with respect to $\mathrm{SL}_2(\mathbb{Z})$.

$$\Delta(\tau) = q \prod_{n>0} (1 - q^n)^{24} \qquad q = e^{2\pi i \tau}, \tau \in \mathscr{H} = \{\tau \in \mathbb{C} : \operatorname{Im} \tau > 0\}$$

$$= \sum_{n=1}^{\infty} a_n q^n = q - 24q^2 + 252q^3 - 1472q^4 + \ldots$$

Ramanujan conjectured in 1916 that for every $\epsilon > 0$: $|a_n| \ll_\epsilon n^{11/2+\epsilon}$

For $f, g : \mathbb{N} \to \mathbb{R}_{>0}$ we write $f \ll g$ if for any $\epsilon > 0$ there exists $c_\epsilon > 0$ such that $f(n) \leqslant c_\epsilon g(n)$ for all n.
Denote $f \asymp g$ if $f \ll g$ and $g \ll f$.

The Ramanujan–Petersson conjecture for modular forms

$a_0 = 0$

Any cusp form $f(\tau)$ of weight k w.r.t. any congruence subgroup Γ.

$$\mathrm{SL}_2(\mathbb{Z}) \supset \Gamma \supset \Gamma(N) = \ker(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}))$$

$$= \{g \in \mathrm{SL}_2(\mathbb{Z}) : g \equiv I \mod N\}$$

$$f(\tau) = \sum_{n=1}^{\infty} a_n q^n$$

Petersson generalized Ramanujan's conjecture: $|a_n| \ll_\epsilon n^{(k-1)/2+\epsilon}$

Deligne (1971) proved that this follows from the Weil conjectures.
3 out of 4 Weil conjectures had been proven by Dwork (1960) and Grothendieck (1965). The remaining conjecture, the Riemann hypothesis over finite fields was proved by Deligne (1974).

Representation theory formulation

There is a corresponding conjecture for Maass forms – the non-holomorphic cousins to the holomorphic modular forms – which remains open.

Modular forms and Maass forms can be embedded in the space of (adelic) automorphic forms on GL_2

Satake (1966) reformulated the conjecture in terms of representation theory which is easier to generalize to GL_n.

To state it we need to define the following:

- cuspidal automorphic representation
- tempered local representation

Setup

Let $\mathbb{Q}_p$ be the p-adic numbers (where p is a prime) and let

$$\mathbb{A} = \mathbb{Q}_\infty \times \prod_p{}' \mathbb{Q}_p = \mathbb{Q}_\infty \times \mathbb{A}_f$$ be the corresponding ring of adeles.

$\mathbb{Q}_\infty$: ↰ $\mathbb{R}$

$\mathbb{Q}$ embeds diagonally in $\mathbb{A}_\mathbb{Q}$ as a discrete and cocompact subring.

$\mathrm{GL}_n(\mathbb{A}) = \mathrm{GL}_n(\mathbb{R}) \times \mathrm{GL}_n(\mathbb{A}_f)$ which acts, untarily, with right multiplication on the arguments for functions in the Hilbert space $L^2(\mathrm{GL}_n(\mathbb{Q})\backslash \mathrm{GL}_n(\mathbb{A}))$.

An irreducible representation $\pi = \bigotimes_{v \le \infty} \pi_v$ of $\mathrm{GL}_n(\mathbb{A})$ is called automorphic if it "appears" in $L^2(\mathrm{GL}_n(\mathbb{Q})\backslash \mathrm{GL}_n(\mathbb{A}))$.

We are interested in the discrete part of the spectrum where "appears" means that it is an irreducible component.

Setup

Similar to in the setting of modular forms, cuspidal means that the constant terms should vanish (now integrating with respect to a more general set of unipotent subgroups).

For a unipotent local representation π_v of $G_v = \mathrm{GL}_n(\mathbb{Q}_v)$ where $v \leq \infty$ let $c^{\pi}_{u,v} : G_v \to \mathbb{C}$ be the matrix coefficient $c_{u,v}(g) = \langle \pi_v(g)u, v \rangle$.

The rate of decay of matrix coefficients is defined as
$r(\pi_v) := \inf\{r \geq 2 : \text{all } c^{\pi}_{u,v} \in L^r(G_v)\}$

A local representation π_v is called tempered if $r(\pi_v) = 2$

Generalized Ramanujan Conjecture for GL_n

The Generalized Ramanujan Conjecture (GRC) for GL_n:

Let $\pi = \bigoplus_{v \leq \infty} \pi_v$ be a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A})$.

Then, for each place v, π_v is tempered.

Has been proved for function fields via Langlands functoriality by Langlands (1970), Drinfeld (1970), Lafforgue (1998, 2002).

GRC is still open for number fields.

State-of-the-art: Clozel (2013) for cohomological representations π.

↑

There exists an algebraic representation E of G such that the $(\mathfrak{g}, K)$-cohomology of π_∞ with coefficients in E does not vanish.

GRC for reductive groups

The naive generalization from GL_n to any reductive group G is false:

$\pi = \bigoplus_{v \leq \infty} \pi_v$ a cuspidal automorphic representation of $G(\mathbb{A})$ then all π_v are tempered.

Counter-examples by Kurokawa (1978) and Howe–Pitatski-Shapiro (1979)

Current formulation for connected reductive group adds a condition that the representation should be generic, i.e. admits a Whittaker model

Cogdell–Kim–Piatetski-Shapiro–Shahidi (2004) showed that the full GRC for GL_n can be lifted to split classical groups SO and Sp using Langlands functoriality.

Applications & approximations

For applications it is often enough to approximately prove the GRC.

One approximation is to see how far from tempered the worst case is in some way. See for example Jacquet–Piatetskii-Shapiro–Shalika (1983), Luo–Rudnick–Sarnak (1999), Kim–Sarnak (2002), Blomer–Brumley (2011).

Applications of GRC and its approximations:

- Selberg's conjecture for eigenvalues of Maass forms
- Ramanujan graphs (in extremal graph theory)
- Subconvexity bounds for L-functions
- Cohomology on symmetric spaces, Betti numbers

Another type of approximation: non-tempered representations are sparse.
The more non-tempered, the less occurring (a density statement)

The Sarnak–Xue Density Hypothesis

Sarnak–Xue Density Hypothesis

Sarnak–Xue (1991) suggested a density hypothesis for representations which could replace GRC in (some) applications.

To understand what is meant by density, let us consider the case where the Ramanujan conjecture has been proven: modular forms, w.r.t

$\Gamma(N) = \ker(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})) = \{g \in \mathrm{SL}_2(\mathbb{Z}) : g \equiv I \bmod N\}$

From Diamond–Shurman (2005):

$$\dim \mathscr{M}_2(\Gamma(N)) = \frac{N+6}{12N}[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] \asymp [\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] \asymp \mathrm{vol}(X(N))$$

where $X(N) = \Gamma(N)\backslash\mathscr{H} \cong \Gamma(N)\backslash\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$

Sarnak–Xue Density Hypothesis

Setup: $G \subset \mathrm{GL}_n$ connected semisimple linear algebraic group / $\mathbb{Q}$ such that $G_\infty := G(\mathbb{R})$ is non-compact.

Level $q \in \mathbb{N}$ congruence subgroup $\Gamma(q) := G(\mathbb{Q}) \cap \ker(\mathrm{GL}_n(\mathbb{Z}) \to \mathrm{GL}_n(\mathbb{Z}/q\mathbb{Z}))$

π_∞ a unitary representation of G_∞ with rate of decay of matrix coeff's $r(\pi_\infty)$

Define multiplicity: $m(\pi_\infty; q) := \dim \mathrm{Hom}_{G_\infty}(\pi_\infty, L^2(\Gamma(q)\backslash G_\infty))$

If strong approximation: $= \dim \mathrm{Hom}_{G_\infty}(\pi_\infty, L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))^{K_f(q)})$

Two extremal cases:

- π_∞ trivial (or finite dim'l): $m(\pi_\infty; q) \asymp 1$ and $r(\pi_\infty) = \infty$
- π_∞ a discrete series: $m(\pi_\infty; q) \asymp \mathrm{vol}(X(q))$ and $r(\pi_\infty) = 2$ (tempered)

  DeGeorge-Wallach (1978-79) $X(q) = \Gamma(q)\backslash G_\infty$

Sarnak–Xue Density Hypothesis

π_∞ a unitary representation of G_∞ with rate of decay of matrix coeff's $r(\pi_\infty)$

Define multiplicity: $m(\pi_\infty; q) := \dim \mathrm{Hom}_{G_\infty}(\pi_\infty, L^2(\Gamma(q)\backslash G_\infty))$

If strong approximation: $= \dim \mathrm{Hom}_{G_\infty}(\pi_\infty, L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))^{K_f(q)})$

- π_∞ trivial (or finite dim'l): $m(\pi_\infty; q) \asymp 1$ and $r(\pi_\infty) = \infty$
- π_∞ a discrete series: $m(\pi_\infty; q) \asymp \mathrm{vol}(X(q))$ and $r(\pi_\infty) = 2$ (tempered)

$X(q) = \Gamma(q)\backslash G_\infty$

Sarnak–Xue Density Hypothesis (SXDH):

For any unitary representation π_∞ and $\epsilon > 0$

$$m(\pi_\infty; q) \ll_\varepsilon \mathrm{vol}(X(q))^{\frac{2}{r(\pi_\infty)}+\epsilon}$$

The more non-tempered (larger $r(\pi_\infty)$), the less occurring (smaller $m(\pi_\infty; q)$)

There is also a p-adic version for π_p

Sarnak–Xue Density Hypothesis

State-of-the-art:

Sarnak–Xue (1991) and Huntley-Katznelson (1993) proved the density hypothesis for $G_\infty = \mathrm{SL}_2(\mathbb{R})$ or $\mathrm{SL}_2(\mathbb{C})$.
Frączyk–Harcos–Maga–Milićević (2024) for $G_\infty = \mathrm{SL}_2(\mathbb{R})^m \times \mathrm{SL}_2(\mathbb{C})^n$.

Variant with Satake parameters instead of $r(\pi_\infty)$:
Blomer (2023) for SL_n for parahoric-level congruence subgroups $\Gamma_0(q)$
Man (2022), Assing (2023) for Sp_4 for Iwahori-level congruence subgroups.
Assing–Blomer (2024), Jana–Kamber (2024) for SL_n for $\Gamma(q)$ assuming q is square free.

Rely on a Kuznetsov-type relative trace formula and is more analytical than our work (on the next slide) which follows the work of Marshall and others relying on results from the Langlands program.

Sarnak–Xue Density Hypothesis

Joint with Shai Evra and Mathilde Gerbelli-Gauthier arXiv:2309.12413

Theorem: Let G be a Gross inner form of the split group SO_5 defined over a totally real number field. Then the Sarnak–Xue density hypothesis holds for any cohomological representation of G_∞.

We use Arthur's endoscopic classification of automorphic representations and Clozel's results for GRC for cohomological representations of GL_n

Conditionality:

- Twisted weighted fundamental lemma (Arthur's endoscopic classification)
- q runs exclusively over integers which are only divisible by primes larger than $10 \cdot [F : \mathbb{Q}] + 1$. (Best known results towards depth preservation in the local Langlands correspondence)

Strategy

- Decompose cohomological representations into Arthur *A*-packets
- The global *A*-parameters are essentially formal sums $\psi = \boxplus_i (\mu_i \boxtimes \nu(m_i))$ where μ_i a self-dual cusp'l autom. rep. of GL_{n_i} and $\nu(m_i)$ the m_i-dim'l irrep of $\mathrm{SL}_2(\mathbb{C})$ with $N = \sum_i m_i n_i$ the dimension of the standard representation of $\hat{G}$.
- The list of pairs $(n_i, m_i)_i$ is called the *A*-shape $\varsigma(\psi)$ of the *A*-parameter ψ.
- We prove a stronger, *A*-shape version of the cohomological SXDH: $m(\varsigma; q) \ll_\varepsilon \mathrm{vol}(X(q))^{\frac{2}{r(\varsigma)}+\epsilon}$ where $r(\varsigma)$ is the worst possible rate of decay of matrix coefficients for any place v and for any representation in the *A*-packets associated to the shape ς, and $m(\varsigma; q)$ essentially sums up multiplicities of representations in packets associated to ς but recast into dimensions of $(\mathfrak{g}, K)$-cohomologies using Matsushima's formula.

Strategy

- We prove a stronger, *A*-shape version of the cohomological SXDH: $m(\varsigma; q) \ll_\varepsilon \mathrm{vol}(X(q))^{\frac{2}{r(\varsigma)}+\epsilon}$ where $r(\varsigma)$ is the worst possible rate of decay of matrix coefficients for any place v and for any representation in the *A*-packets associated to the shape ς, and $m(\varsigma; q)$ essentially sums up multiplicities of representations in packets associated to ς but recast into dimensions of $(\mathfrak{g}, K)$-cohomologies using Matsushima's formula.

- Compute each $r(\varsigma)$ using Arthur's endoscopic classification and Clozel's GRC for cohomological representations.
- For $m(\varsigma; q)$ we use different methods for different shapes for counting chomology dimensions: endoscopic character relations, explicit description of *A*-packets for SO_5 by Schmidt (2019), and, for some shapes, naive bounds.

Applications

- Upper bounds on L^2-Betti numbers for $X(q)$
- density-Ramanujan complexes and cut-off phenomena (after small fix recently communicated by Simon Marshall)

Slides will be made available at hgustafsson.se