Mixing of 3-Term Progressions in Quasirandom Groups

Amey Bhangale
University of California, Riverside, CA, USA

Prahladh Harsha
Tata Institute of Fundamental Research, Mumbai, India

Sourya Roy
University of California, Riverside, CA, USA

Abstract

In this paper, we show the mixing of three-term progressions (x, xg, xg^2) in every finite quasirandom group, fully answering a question of Gowers. More precisely, we show that for any D-quasirandom group G and any three sets $A_1, A_2, A_3 \subset G$, we have

$$\left| \Pr_{x,y\sim G}\left[x \in A_1, xy \in A_2, xy^2 \in A_3\right] - \prod_{i=1}^{3} \Pr_{x\sim G}[x \in A_i] \right| \leq \left(\frac{2}{\sqrt{D}}\right)^{1/4}.$$

Prior to this, Tao answered this question when the underlying quasirandom group is $\mathrm{SL}_d(\mathbb{F}_q)$. Subsequently, Peluse extended the result to all non-abelian finite *simple* groups. In this work, we show that a slight modification of Peluse's argument is sufficient to fully resolve Gowers' quasirandom conjecture for 3-term progressions. Surprisingly, unlike the proofs of Tao and Peluse, our proof is elementary and only uses basic facts from non-abelian Fourier analysis.

2012 ACM Subject Classification Mathematics of computing → Combinatoric problems

Keywords and phrases Quasirandom groups, 3-term arithmetic progressions

Digital Object Identifier 10.4230/LIPIcs.ITCS.2022.20

Funding *Prahladh Harsha*: Research supported by the Department of Atomic Energy, Government of India, under project 12-R&D-TFR-5.01-0500 and the Swarnajayanti Fellowship.

Acknowledgements We thank Shachar Lovett and Terence Tao for helpful comments on an earlier draft of this paper. We are also thankful to the anonymous reviewers whose comments helped in improving the presentation of the paper.

1 Introduction

In this note, we revisit a conjecture by Gowers [7] about mixing of three term progressions in quasirandom finite groups. Gowers initiated the study of quasirandom groups while refuting a conjecture of Babai and Sós [2] regarding the size of the largest product-free set in a given finite group. A finite group is said to be D-quasirandom for a positive integer D if all its non-trivial irreducible representations are at least D-dimensional. The quasirandomness property of groups can be used to show that certain "objects" related to the group "mix" well. For instance, the quasirandomness of the group $\mathrm{PSL}_2(\mathbb{F}_q)$ can be used to give an alternate (and weaker) proof [5] that the Ramanujan graphs of Lubotzky, Philips and Sarnak [10] are expanders. Bourgain and Gamburd [4] used quasirandomness to prove that certain other Cayley graphs are expanders.

Gowers proved that for any D-quasirandom group G and any three subsets $A, B, C \subset G$ satisfying $|A| \cdot |B| \cdot |C| \geq |G|^3/D$, there exist $x \in A, y \in B, z \in C$ such that $x \cdot y = z$. More generally, he proved that the number of such triples $(x, y, z) \in A \times B \times C$ such that $x \cdot y = z$

13th Innovations in Theoretical Computer Science Conference (ITCS 2022).
Editor: Mark Braverman; Article No. 20; pp. 20:1–20:9
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

is at least $(1-\eta)|A|\cdot|B|\cdot|C|/|G|$ provided $|A|\cdot|B|\cdot|C| \geq |G|^3/\eta^2 D$. In other words the set of triples of the form (x, y, xy) mix well in a quasirandom group. Gowers' proof of this result was the inspiration and the first step towards the recent optimal inapproximability result for satisfiable kLIN over non-Abelian groups [3]. After proving the well-mixing of triples of the form (x, y, xy) in quasirandom groups, Gowers conjectured a similar statement for triples of the form (x, xy, xy^2). More precisely, he conjectured the following statement: Let G be a D-quasirandom group and $f_1, f_2, f_3 : G \to \mathbb{C}$ such that $\|f_i\|_\infty \leq 1$, then

$$\left| \mathop{\mathbb{E}}_{x,y\sim G}\left[f_1(x)f_2(xy)f_3(xy^2)\right] - \prod_{i=1,2,3} \mathop{\mathbb{E}}_{x\sim G}\left[f_i(x)\right]\right| = o_D(1)\ , \tag{1}$$

where the expression $o_D(1)$ goes to zero as D increases.

When D is small, one hope to bound the left-hand side expression above by any meaningful quantity. Consider G to be the Abelian group $\mathbb{Z}/n\mathbb{Z}$ which is 1-quasirandom and set $f_i = \mathbf{1}_B$ for all $i \in [3]$ where $B = \{1, \ldots, \lfloor \delta n \rfloor\}$ for any $\delta \in (0, 1/3)$. It is easy to observe that the first term in the left-hand side of (1) is $\Omega(\delta^2)$ while the second term is δ^3. A more interesting example is when the group is S_n. In this case, let $f_i = \mathbf{1}_{B_i}$, where $B_1 = A_n, B_2 = S_n$and $B_3 = S_N \setminus A_n$. Now, the $f_i's$ have density $1/2, 1, 1/2$ respectively. Note that there in no 3-term progression in (B_1, B_2, B_3) and therefore the first term in the left-hand side of (1) is 0. Although S_n is a non-Abelian group, it does have a non-trivial representation of dimension 1. Thus the conjecture essentially asks if the group is very "non-Abelian" (more precisely, is D-quasirandom for large D), then do these counterexamples go away.The conjecture can be naturally extended to k-term progressions and product of k functions for $k > 3$. However, in this note we will focus on the three term case.

For the specific case of 3-term progressions, Tao [12] proved the conjecture for the group $\mathrm{SL}_d(\mathbb{F}_q)$ for bounded d using algebraic geometric machinery. In particular, he proved that the left-hand side expression in (1) can be bounded by $O(1/q^{1/8})$ when $d = 2$ and $O_d(1/q^{1/4})$ for larger d. Tao's approach relied on algebraic geometry and was not amenable to other quasirandom groups. Later, Peluse [11] proved the conjecture for all non-Abelian finite simple groups. She used basic facts from non-Abelian Fourier analysis to prove that the left-hand side expression in (1) can be bounded by $\sum_{1\neq\rho\in\hat{G}} 1/d_\rho$ where $\hat{G}$ represents the set of irreducible unitary representation of G and d_ρ the dimension of the irreducible representation ρ. This latter quantity is the *Witten zeta function* ζ_G of the group G minus one and can be bounded for *simple* finite quasirandom groups using a result due to Liebeck and Shalev [9, 8].

In this paper, we show that a slight variation of Peluse's argument can be used to prove the conjecture for *all quasirandom groups* with *better* error parameters. More surprisingly, the proof stays completely elementary and short. Specifically, we prove the following statement:

▶ **Theorem 1.** *Let G be a D-quasirandom finite group, i.e, its all non-trivial irreducible representations are at least D-dimensional. Let $f_1, f_2, f_3 : G \to \mathbb{C}$ such that $\|f_i\|_\infty \leq 1$ then*

$$\left| \mathop{\mathbb{E}}_{x,y\sim G}\left[f_1(x)f_2(xy)f_3(xy^2)\right] - \prod_{i=1,2,3} \mathop{\mathbb{E}}_{x\sim G}\left[f_i(x)\right]\right| \leq \left(\frac{2}{\sqrt{D}}\right)^{\frac{1}{4}}.$$

2 Preliminaries

We begin by recalling some basic representation theory and non-Abelian Fourier analysis. See the monograph by Diaconis [6, Chapter 2] for a more detailed treatment (with proofs).

We will be working with a finite group G and complex-valued functions $f\colon G \to \mathbb{C}$ on G. All expectations will be with respect to the uniform distribution on G. The *convolution* between two function $f, h\colon G \to \mathbb{C}$, denoted by $f * h$, is defined as follows:

$$(f * h)(x) := \mathop{\mathbb{E}}_{y}[f(xy^{-1})h(y)].$$

For any $p \geq 1$, the p-norm of any function $f\colon G \to \mathbb{C}$ is defined as

$$\|f\|_p^p := \mathop{\mathbb{E}}_{x}[|f(x)|^p].$$

For any element $g \in G$, the *conjugacy class of* g, denoted by $C(g)$, refers to the set $\{x^{-1}gx | x \in G\}$. Observe that the conjugacy classes form a partition of the group G. A function $f\colon G \to \mathbb{C}$ is said to be a *class function* if it is constant on conjugacy classes.

For any $b \in G$ we use $\Delta_b f(x) := f(x) \cdot f(xb)$. For any set $S \subset G$, $\mu_S\colon G \to \mathbb{R}$ denotes the scaled density function $\frac{|G|}{|S|}\mathbb{1}_S$. The scaling ensures that $\mathbb{E}_x[\mu_S(x)] = 1$.

Given a complex vector space V, we denote the vector space of linear operators on V by $\mathrm{End}(V)$. This space is endowed with the following inner product and norm (usually referred to as the *Hilbert-Schmidt* norm):

For $A, B \in \mathrm{End}(V)$, $\quad \langle A, B\rangle_{\mathrm{HS}} := \mathrm{Trace}(A^* B) \quad$ and $\quad \|A\|_{\mathrm{HS}}^2 := \langle A, A\rangle_{\mathrm{HS}} = \mathrm{Trace}(A^* A)$.

This norm is known to be submultiplicative (i.e, $\|AB\|_{\mathrm{HS}} \leq \|A\|_{\mathrm{HS}} \cdot \|B\|_{\mathrm{HS}}$).

Representations and Characters

A *representation* $\rho\colon G \to \mathrm{End}(V)$ is a homomorphism from G to the set of linear operators on V for some finite-dimensional vector space V over $\mathbb{C}$, i.e., for all $x, y \in G$, we have $\rho(xy) = \rho(x)\rho(y)$. The dimension of the representation ρ, denoted by d_ρ, is the dimension of the underlying $\mathbb{C}$-vector space V. The *character* of a representation ρ, denoted by $\chi_\rho\colon G \to \mathbb{C}$, is defined as $\chi_\rho(x) := \mathrm{Trace}(\rho(x))$.

The representation $1\colon G \to \mathbb{C}$ satisfying $1(x) = 1$ for all $x \in G$ is the *trivial* representation. A representation $\rho\colon G \to \mathrm{End}(V)$ is said to *reducible* if there exists a non-trivial subpsace $W \subset V$ such that for all $x \in G$, we have $\rho(x)W \subset W$. A representation is said to be *irreducible* otherwise. The set of all irreducible representations of G (upto equivalences) is denoted by $\hat{G}$.

For every representation $\rho\colon G \to \mathrm{End}(V)$, there exists an inner product $\langle\cdot,\cdot\rangle_V$ over V such that every $\rho(x)$ is unitary (i.e, $\langle \rho(x)u, \rho(x)v\rangle_V = \langle u, v\rangle_V$ for all $u, v \in V$ and $x \in G$). Hence, we might wlog. assume that all the representations we are considering are unitary.

The following are some well-known facts about representations and characters.

▶ **Proposition 2.**
1. *The group G is Abelian iff $d_\rho = 1$ for every irreducible representation ρ in $\hat{G}$.*
2. *For any finite group G, $\sum_{\rho \in \hat{G}} d_\rho^2 = |G|$.*
3. *[orthogonality of characters] For any $\rho, \rho' \in \hat{G}$ we have: $\mathbb{E}_x\left[\chi_\rho(x)\overline{\chi_{\rho'}(x)}\right] = \mathbb{1}[\rho = \rho']$.*

▶ **Definition 3** (quasirandom groups). *A non-Abelian group G is said to be D-*quasirandom *for some positive integer D if all its non-trivial irreducible representations ρ satisfy $d_\rho \geq D$.*

Any group G having a non-trivial Abelian subgroup is 1-quasirandom. For instance, the symmetric group S_n is 1-quasirandom, while the alternating group A_n is $\Omega(n)$-quasirandom. The special linear group $\mathrm{SL}_2(\mathbb{F}_p)$ for prime p is $(p-1)/2$-quasirandom. If G, G' are D-quasirandom, so is $G \times G'$.

Non-Abelian Fourier analysis

Given a function $f\colon G \to \mathbb{C}$ and an irreducible representation $\rho \in \hat{G}$, the Fourier transform is defined as follows:

$$\hat{f}(\rho) := \mathop{\mathbb{E}}_x[f(x)\rho(x)].$$

The following proposition summarizes the basic properties of Fourier transform that we will need.

▶ **Proposition 4.** *For any $f, h\colon G \to \mathbb{C}$, we have the following*

1. *[Fourier transform of trivial representation]*
$$\hat{f}(1) = \mathop{\mathbb{E}}_x[f(x)].$$
2. *[Convolution]*
$$\widehat{f * h}(\rho) = \hat{f}(\rho) \cdot \hat{h}(\rho).$$
3. *[Fourier inversion formula]*
$$f(x) = \sum_{\rho \in \hat{G}} d_\rho \cdot \langle \hat{f}(\rho), \rho(x)\rangle_{\mathrm{HS}}.$$
4. *[Parseval's identity]*
$$\|f\|_2^2 = \sum_{\rho \in \hat{G}} d_\rho \cdot \|\hat{f}(\rho)\|_{\mathrm{HS}}^2.$$
5. *[Fourier transfrom of class functions] For any class function $f\colon G \to \mathbb{C}$, the Fourier transform satisfies*
$$\hat{f}(\rho) = c \cdot I_{d_\rho}$$
for some constant $c = c(f, \rho) \in \mathbb{C}$. In other words, the Fourier transform is a scaling of the Identity operator I_{d_ρ}.

The following claim (also used by Peluse [11]) observes that the scaled density function $\mu_{gC(g)}$ has a very simple Fourier transform since it is a translate of the class function $\mu_{C(g)}$

▷ Claim 5. For any $g \in G$ and $\rho \in \hat{G}$ we have:

$$\hat{\mu}_{gC(g)}(\rho) = \frac{\chi_\rho(g)}{d_\rho} \cdot \rho(g)$$

where $C(g)$ refers to the conjugacy class of g. Moreover, $\|\hat{\mu}_{gC(g)}\|_{\mathrm{HS}}^2 = \frac{|\chi_\rho(g)|^2}{d_\rho}$

Proof. We begin by observing that

$$\begin{aligned}
\hat{\mu}_{gC(g)}(\rho) &= \mathop{\mathbb{E}}_x\left[\mu_{gC(g)}(x) \cdot \rho(x)\right] \\
&= \mathop{\mathbb{E}}_x\left[\mu_{gC(g)}(gx) \cdot \rho(gx)\right] \\
&= \mathop{\mathbb{E}}_x\left[\mu_{gC(g)}(gx) \cdot \rho(g) \cdot \rho(x)\right] \\
&= \rho(g) \cdot \mathop{\mathbb{E}}_x\left[\mu_{C(g)}(x) \cdot \rho(x)\right] \\
&= \rho(g) \cdot \hat{\mu}_{C(g)}(\rho).
\end{aligned}$$

On the other hand, as $\mu_{C(g)}$ is a class function, we have $\hat{\mu}_{C(g)}(\rho) = c \cdot I_{d_\rho}$ for some constant $c \in \mathbb{C}$. The constant c can be determined by taking trace on either side of $c \cdot I_{d_\rho} = \hat{\mu}_{C(g)} = \mathbb{E}_x[\mu_{C(g)}(x) \cdot \rho(x)]$ and noting that $\text{Trace}(\rho(x)) = \chi_\rho(g)$ as follows:

$$c \cdot d_\rho = \mathop{\mathbb{E}}_x \left[\mu_{C(g)}(x) \cdot \chi_\rho(g)\right] = \mathop{\mathbb{E}}_x \left[\mu_{C(g)}(x)\right] \cdot \chi_\rho(g) = \chi_\rho(g).$$

Hence, $c = \frac{\chi_\rho(g)}{d_\rho}$ and $\hat{\mu}_{gC(g)} = \frac{\chi_\rho(g)}{d_\rho} \cdot \rho(g)$. Lastly we have,

$$\begin{aligned}
\|\hat{\mu}_{gC(g)}\|_{\text{HS}}^2 &= \left\| \frac{\chi_\rho(g)}{d_\rho} \cdot \rho(g) \right\|_{\text{HS}}^2 \\
&= \frac{|\chi_\rho(g)|^2}{d_\rho^2} \cdot \text{Trace}\left(\rho(g)^* \cdot \rho(g)\right) \\
&= \frac{|\chi_\rho(g)|^2}{d_\rho^2} \cdot d_\rho && \text{(By unitariness of } \rho(g)) \\
&= \frac{|\chi_\rho(g)|^2}{d_\rho}.
\end{aligned}$$ ◀

The key property of D-quasirandom groups that we will be using is the following inequality due to Babai, Nikolov and Pyber, the proof of which we provide for the sake of completeness.

▶ **Lemma 6** ([1]). *If G is a D-quasirandom group and $f_1, f_2 \colon G \to \mathbb{C}$ such that either f_1 or f_2 is mean zero then*

$$\|f_1 * f_2\|_2 \le \frac{1}{\sqrt{D}} \cdot \|f_1\|_2 \cdot \|f_2\|_2.$$

Proof.

$$\begin{aligned}
\|f_1 * f_2\|^2 &= \sum_{\rho \in \hat{G}} d_\rho \|\widehat{f_1 * f_2}(\rho)\|_{\text{HS}}^2 \\
&= \sum_{\rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho) \cdot \hat{f}_2(\rho)\|_{\text{HS}}^2 \\
&\le \sum_{\rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho)\|_{\text{HS}}^2 \cdot \|\hat{f}_2(\rho)\|_{\text{HS}}^2 && \text{(By submultiplicativity of norm)} \\
&= \sum_{1 \ne \rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho)\|_{\text{HS}}^2 \cdot \|\hat{f}_2(\rho)\|_{\text{HS}}^2 && \text{(By mean zeroness)} \\
&\le \frac{1}{D} \cdot \sum_{1 \ne \rho \in \hat{G}} d_\rho^2 \|\hat{f}_1(\rho)\|_{\text{HS}}^2 \cdot \|\hat{f}_2(\rho)\|_{\text{HS}}^2 && \text{(By } D\text{-quasirandomness)} \\
&\le \frac{1}{D} \left(\sum_{1 \ne \rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho)\|_{\text{HS}}^2 \right) \cdot \left(\sum_{1 \ne \rho \in \hat{G}} d_\rho \|\hat{f}_2(\rho)\|_{\text{HS}}^2 \right) \\
&\le \frac{1}{D} \cdot \|f_1\|_2^2 \cdot \|f_2\|_2^2 .
\end{aligned}$$ ◀

The following is a simple corrollary of Lemma 6.

▶ **Corollary 7.** *If G is D-quasirandom; $f \colon G \to \mathbb{C}$ has zero mean and $\|f\|_\infty \le 1$ then*

$$\mathop{\mathbb{E}}_b \left[|\mathop{\mathbb{E}}_x \Delta_b f(x)| \right] \le \frac{1}{\sqrt{D}}.$$

Proof. Let $f'(x) := f(x^{-1})$. We have,

$$\begin{aligned}
\mathop{\mathbb{E}}_{b}\left[\left|\mathop{\mathbb{E}}_{x}\Delta_b f(x)\right|\right] &= \mathop{\mathbb{E}}_{b}\left[\left|\mathop{\mathbb{E}}_{x} f(x)f(xb)\right|\right] \\
&= \mathop{\mathbb{E}}_{b}\left[\left|\mathop{\mathbb{E}}_{x} f'(x^{-1})f(xb)\right|\right] \\
&= \mathop{\mathbb{E}}_{b}\left[\left|f' * f(b)\right|\right] \\
&\leq \mathop{\mathbb{E}}_{b}\left[\left|f' * f(b)\right|^2\right]^{1/2} && \text{(By Cauchy-Schwarz inequality)} \\
&= \|f' * f\|_2 \\
&\leq \frac{1}{\sqrt{D}} \cdot \|f'\|_2 \cdot \|f\|_2 && \text{(By Lemma 6)} \\
&\leq \frac{1}{\sqrt{D}}. && \text{(Since } \|f\|_2 \leq \|f\|_\infty \leq 1\text{).}
\end{aligned}$$

◀

3 Proof of Theorem 1

The following proposition is where we deviate from Peluse's proof [11]. We give an elementary proof for *every* quasirandom group while Peluse proved the same result for *simple* finite groups using the result of Liebeck and Shalev [9, 8] to bound the Witten zeta function ζ_G for *simple* finite groups.

▶ **Proposition 8.** *Let G be a D-quasirandom group. Let $f\colon G \to \mathbb{C}$ such that $\|f\|_\infty \leq 1$, $\mathbb{E}[f] = 0$ and f_b is the mean zero component of the function $\Delta_b f$ (i.e., $f_b(x) = \Delta_b f(x) - \mathbb{E}_x[\Delta_b f(x)]$). Then*

$$\mathop{\mathbb{E}}_{g,b}\left[\left|\mathop{\mathbb{E}}_{x}\left[\Delta_b f(x)\cdot(f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})})(x)\right]\right|\right] \leq \frac{1}{\sqrt{D}}\,.$$

Proof. Let us denote the expression on the L.H.S. as Γ. We use simple manipulations and previously stated facts to simplify the expression.

$$\begin{aligned}
\Gamma^2 &\leq \mathop{\mathbb{E}}_{g,b}\left[\left(\|\Delta_b f\|_2\right)\cdot\left(\|f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_2\right)\right]^2 && \text{(By Cauchy-Schwarz inequality)} \\
&\leq \mathop{\mathbb{E}}_{g,b}\left[\|f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_2\right]^2 && \text{(Since } \|\Delta_b f\|_2 \leq 1\text{)} \\
&\leq \mathop{\mathbb{E}}_{g,b}\left[\|f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_2^2\right] && \text{(By Cauchy Schwarz inequality)} \\
&= \mathop{\mathbb{E}}_{g,b}\left[\sum_{1\neq\rho\in\hat{G}} d_\rho \cdot \|\hat{f}_{g^{-1}bg}(\rho)\cdot\hat{\mu}_{g^{-1}C(g^{-1})}(\rho)\|_{\mathrm{HS}}^2\right] \\
& && \text{(By Parseval's identity \& } \hat{f}_{g^{-1}bg}(1) = 0\text{)} \\
&\leq \mathop{\mathbb{E}}_{g,b}\left[\sum_{1\neq\rho\in\hat{G}} d_\rho \cdot \|\hat{f}_{gbg^{-1}}(\rho)\|_{\mathrm{HS}}^2\cdot\|\hat{\mu}_{g^{-1}C(g^{-1})}(\rho)\|_{\mathrm{HS}}^2\right] \\
& && \text{(By submultiplicativity of norm)} \\
&= \mathop{\mathbb{E}}_{g,b}\left[\sum_{1\neq\rho\in\hat{G}} \|\hat{f}_{g^{-1}bg}(\rho)\|_{\mathrm{HS}}^2\cdot|\chi_\rho(g)|^2\right] && \text{(By Claim 5)} \\
&= \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{g}\left[|\chi_\rho(g)|^2\cdot\mathop{\mathbb{E}}_{b}\left[\|\hat{f}_{gbg^{-1}}(\rho)\|_{\mathrm{HS}}^2\right]\right].
\end{aligned}$$

Now using the fact that gbg^{-1} is uniformly distributed in G for a fixed g and a uniformly random b in G, we can simplify the above expression as follows.

$$\begin{aligned}
\Gamma^2 &\le \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{g}\left[|\chi_\rho(g)|^2 \cdot \mathop{\mathbb{E}}_{b}\left[\|\hat{f}_b(\rho)\|_{\mathrm{HS}}^2\right]\right] \\
&= \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{b}\left[\|\hat{f}_b(\rho)\|_{\mathrm{HS}}^2\right] \cdot \mathop{\mathbb{E}}_{g}\left[|\chi_\rho(g)|^2\right] \\
&= \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{b}\left[\|\hat{f}_b(\rho)\|_{\mathrm{HS}}^2\right] && \text{(By orthogonality of } \chi_\rho) \\
&= \mathop{\mathbb{E}}_{b}\left[\sum_{1\neq\rho\in\hat{G}} \|\hat{f}_b(\rho)\|_{\mathrm{HS}}^2\right].
\end{aligned}$$

Finally, we use the fact that all the terms in the summation are non-negative and the group G is a D-quasirandom group.

$$\begin{aligned}
\Gamma^2 &\le \frac{1}{D} \cdot \mathop{\mathbb{E}}_{b}\left[\sum_{1\neq\rho\in\hat{G}} d_\rho \cdot \|\hat{f}_b(\rho)\|_{\mathrm{HS}}^2\right] \\
&= \frac{1}{D} \cdot \mathop{\mathbb{E}}_{b}\left[\|f_b\|_2^2\right] && \text{(By Parseval's identity)} \\
&\le \frac{1}{D}, && \text{(Because } \|f_b\|_2^2 \le 1).
\end{aligned}$$

The proof of this lemma is similar to the proof of the BNP inequality (Lemma 6). The key difference being that we have a complete characterization of the Fourier transform of $\mu_{gC(g)}$ from Claim 5 which we use to give a sharper bound. ◀

We are now ready to prove the main Theorem 1. This part of the proof is similar to the corresponding expression that appears in the paper of Peluse [11], which is in turn inspired by Tao's adaptation of Gowers' repeated Cauchy-Schwarzing trick to the nonebelian setting. We, however, present the entire proof for the sake of completeness.

Proof of Theorem 1. Let us denote the L.H.S. of the expression by Θ_{f_1,f_2,f_3}. Without loss of generality we assume $\mathbb{E}[f_3] = 0$. Now we have,

$$\begin{aligned}
\Theta_{f_1,f_2,f_3}^4 &= \left|\mathop{\mathbb{E}}_{x,y}\left[f_1(x)f_2(xy)f_3(xy^2)\right]\right|^4 \\
&= \left|\mathop{\mathbb{E}}_{x,z}\left[f_1(xz^{-1})f_2(x)f_3(xz)\right]\right|^4 && \text{(Change of variables: } x \leftarrow xy, z \leftarrow y) \\
&\le \left|\mathop{\mathbb{E}}_{x,z_1,z_2}\left[f_1(xz_1^{-1})f_1(xz_2^{-1})f_3(xz_1)f_3(xz_2)\right]\right|^2 \\
& && \text{(Cauchy-Schwarz over } x;\ \|f_2\|_\infty = 1 \text{ and expansion)} \\
&= \left|\mathop{\mathbb{E}}_{y,z,a}\left[f_1(y)f_1(ya)f_3(yz^2)f_3(yza^{-1}z)\right]\right|^2 \\
& && \text{(Change of variables: } y \leftarrow xz_1^{-1}, z \leftarrow z_1, a \leftarrow z_1z_2^{-1}) \\
&= \left|\mathop{\mathbb{E}}_{y,z,a}\left[\Delta_a f_1(y) \cdot \Delta_{z^{-1}a^{-1}z}\ f_3(yz^2)\right]\right|^2 \\
&\le \left|\mathop{\mathbb{E}}_{y,a,z_1,z_2}\left[\Delta_{z_1^{-1}a^{-1}z_1}\ f_3(yz_1^2) \cdot \Delta_{z_2^{-1}a^{-1}z_2}\ f_3(yz_2^2)\right]\right|, \\
& && \text{(Cauchy-Schwarz over } y, a;\ \|f_1\|_\infty \le 1\text{).}
\end{aligned}$$

Now, using the following change of variables, $z \leftarrow z_1,\ x \leftarrow yz_1^2,\ b \leftarrow z_1^{-1}a^{-1}z_1,\ g \leftarrow z_1^{-1}z_2$, we get

$$\begin{aligned}
\Theta^4_{f_1,f_2,f_3} &\le \left| \mathop{\mathbb{E}}_{x,b,z,g} \left[\Delta_b\ f_3(x) \cdot \Delta_{g^{-1}bg}\ f_3(xz^{-1}gzg)\right] \right| \\
&= \left| \mathop{\mathbb{E}}_{x,b,g} \left[\Delta_b\ f_3(x) \cdot \mathop{\mathbb{E}}_{z}[\Delta_{g^{-1}bg}\ f_3(xz^{-1}gzg)]\right] \right| \\
&= \left| \mathop{\mathbb{E}}_{x,b,g} \left[\Delta_b\ f_3(x) \cdot \mathop{\mathbb{E}}_{a}[\Delta_{g^{-1}bg}\ f_3(xa^{-1}) \cdot \frac{|G|}{|C(g^{-1})|} 1_{g^{-1}C(g^{-1})}(a)]\right] \right| \\
&= \left| \mathop{\mathbb{E}}_{x,b,g} \left[\Delta_b\ f_3(x) \cdot \mathop{\mathbb{E}}_{a}[\Delta_{g^{-1}bg}\ f_3(xa^{-1}) \cdot \mu_{g^{-1}C(g^{-1})}(a)]\right] \right| \\
&= \left| \mathop{\mathbb{E}}_{x,b,g} \left[\Delta_b\ f_3(x) \cdot \Delta_{g^{-1}bg}\ f_3 * \mu_{g^{-1}C(g^{-1})}(x)\right] \right|.
\end{aligned}$$

The second equality follows because after g, x, b have been fixed we only use z to compute $z^{-1}gz$ and the map that takes $z \in G$ to $z^{-1}gz \in C(g)$ is surjective where each member in the range has preimage of size $\frac{|G|}{|C(g^{-1})}| = |\text{Centralizer}(g)|$. We now separate the function $\Delta_{g^{-1}bg}\ f_3$ from its the mean zero part as follows: Let $\Delta_{g^{-1}bg}\ f_3 = f'_{g^{-1}bg} + f_{g^{-1}bg}$ where $f'_{g^{-1}bg} = \mathbb{E}_x[\Delta_{g^{-1}bg}\ f_3(x)]$ and $f_{g^{-1}bg}(x) = \Delta_{g^{-1}bg}\ f_3(x) - f'_{g^{-1}bg}$.

$$\begin{aligned}
\Theta^4_{f_1,f_2,f_3} &\le \left| \mathop{\mathbb{E}}_{x,b,g} \left[\Delta_b\ f_3(x) \cdot (f_{g^{-1}bg} + f'_{g^{-1}bg}) * \mu_{g^{-1}C(g^{-1})}(x)\right] \right| \\
&\le \mathop{\mathbb{E}}_{b,g} \left[\left| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x) \cdot f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}(x)\right] \right| \right] \\
&\qquad\qquad + \mathop{\mathbb{E}}_{b,g} \left[\left| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x) \cdot f'_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}(x)\right] \right| \right] \\
&\le \frac{1}{\sqrt{D}} + \mathop{\mathbb{E}}_{b,g} \left[\left| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x)\right] \right| \cdot \|f'_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_\infty \right] \\
&\qquad\qquad \text{(Using Proposition 8 to bound the first expectation)} \\
&= \frac{1}{\sqrt{D}} + \mathop{\mathbb{E}}_{b,g} \left[\left| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x)\right] \right| \cdot |f'_{g^{-1}bg}| \right] \\
&\le \frac{1}{\sqrt{D}} + \mathop{\mathbb{E}}_{b} \left[\left| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x)\right] \right| \right] && \text{(Using } |f'_{g^{-1}bg}| \le 1\text{)} \\
&\le \frac{2}{\sqrt{D}}\ , && \text{(By Corollary 7 and } \|f_3\|_\infty \le 1\text{).}
\end{aligned}$$

◀

References

1 László Babai, Nikolay Nikolov, and László Pyber. Product growth and mixing in finite groups. In *Proc. 19th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 248–257, 2008. doi:10.1145/1347082.1347110.

2 László Babai and Vera T. Sós. Sidon sets in groups and induced subgraphs of Cayley graphs. *Eur. J. Comb.*, 6(2):101–114, 1985. doi:10.1016/S0195-6698(85)80001-9.

3 Amey Bhangale and Subhash Khot. Optimal inapproximability of satisfiable k-LIN over non-abelian groups. In *Proc. 53rd ACM Symp. on Theory of Computing (STOC)*, pages 1615–1628, 2021. doi:10.1145/3406325.3451003.

4 Jean Bourgain and Alex Gamburd. Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$. *Ann. of Math.*, 167:625–642, 2008. doi:10.4007/annals.2008.167.625.

5. Giuliana Davidoff, Peter Sarnak, and Alain Valette. *Elementary Number Theory, Group Theory and Ramanujan Graphs*. London Mathematical Society Student Texts. Cambridge University Press, 2003. doi:10.1017/CBO9780511615825.
6. Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *IMS Lecture Notes Monogr. Ser.* Institute of Mathematical Statistics, 1998. doi:10.1214/lnms/1215467411.
7. William Timothy Gowers. Quasirandom groups. *Comb. Probab. Comput.*, 17(3):363–387, 2008. doi:10.1017/S0963548307008826.
8. Martin W Liebeck and Aner Shalev. Character degrees and random walks in finite groups of Lie type. *Proc. Amer. Math. Soc.*, 90(1):61–86, 2004. doi:10.1112/S0024611504014935.
9. Martin W Liebeck and Aner Shalev. Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks. *Journal of Algebra*, 276(2):552–601, 2004. doi:10.1016/S0021-8693(03)00515-5.
10. Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. doi:10.1007/BF02126799.
11. Sarah Peluse. Mixing for three-term progressions in finite simple groups. *Math. Proc. Cambridge Philos. Soc.*, 165(2):279–286, 2018. doi:10.1017/S0305004117000482.
12. Terrence Tao. Mixing for progressions in nonabelian groups. *Forum of Mathematics, Sigma*, 1:e2, 2013. doi:10.1017/fms.2013.2.