Forum of Mathematics, Sigma (2016), Vol. 4, e24, 35 pages
doi:10.1017/fms.2016.8

QUASIRANDOM GROUP ACTIONS

NICK GILL
Department of Mathematics, University of South Wales, Treforest, CF37 1DL, UK;
email: nick.gill@southwales.ac.uk

Received 15 September 2015; accepted 21 April 2016

Abstract

Let G be a finite group acting transitively on a set Ω. We study what it means for this action to be *quasirandom*, thereby generalizing Gowers' study of quasirandomness in groups. We connect this notion of quasirandomness to an upper bound for the convolution of functions associated with the action of G on Ω. This convolution bound allows us to give sufficient conditions such that sets $S \subseteq G$ and $\Delta_1, \Delta_2 \subseteq \Omega$ contain elements $s \in S, \omega_1 \in \Delta_1, \omega_2 \in \Delta_2$ such that $s(\omega_1) = \omega_2$. Other consequences include an analogue of 'the Gowers trick' of Nikolov and Pyber for general group actions, a sum-product type theorem for large subsets of a finite field, as well as applications to expanders and to the study of the diameter and width of a finite simple group.

2010 Mathematics Subject Classification: 60B15 (primary); 20P05, 20D60, 20F70 (secondary)

In his seminal 2008 paper entitled 'Quasirandom groups', Gowers introduced the notion of a *d-quasirandom group*. He gives a number of formulations of this idea but, for our purposes, it is easiest to define a group G to be *d-quasirandom* (for some $d \in \mathbb{R}^+$) if every nontrivial irreducible representation of G has dimension at least d. Gowers related this definition of quasirandomness to notions of quasirandomness for functions $G \to \mathbb{R}$, and for particular graphs related to G ('directed Cayley graphs'). These connections allowed him to prove the following fundamental result:

THEOREM 1. *Let G be a finite d-quasirandom group of order n. Let A, B and C be three subsets of Γ such that $|A|\cdot|B|\cdot|C| > n^3/d$. Then there exist $a \in A, b \in B$ and $c \in C$ with $ab = c$.*

Some time after Gowers proved this result, Babai, Nikolov and Pyber were able to give a different proof. They proved a bound for the convolution of probability

measures on G and showed that Theorem 1 followed directly. What is more, the convolution bound had a number of other important applications, most notably to the theory of expander graphs.

In this paper, we generalize Theorem 1. We show that it is a particular case of a result concerning arbitrary transitive actions G on a set Ω. Our method involves a careful study of the original arguments of Gowers, and of Babai–Nikolov–Pyber. We are able to adapt both arguments to give different bounds on the convolution of functions related to the action, and these bounds imply the mentioned generalization of Theorem 1, as well as a number of other significant results.

1. Main results

In order to state our main results we must establish some notation which will hold throughout the paper. First, we set G to be a finite group acting transitively on a finite set Ω.

Consider two functions $X : G \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$. We define the *convolution* $X *_c Y$ of X and Y to be the following function on Ω:

$$(X *_c Y)(\omega) = \sum_{g \in G} X(g) Y(g^{-1}\omega). \tag{1.1}$$

This definition, which has appeared in various places in the literature, is a generalization of the definition of convolution given in [2]. Observe that if X and Y are probability distributions then $X *_c Y$ is also a probability distribution; on the other hand, if either X or Y sum to 0 then $X *_c Y$ sums to 0.

We write $H = \mathrm{Stab}_G(\omega)$, the stabilizer in G of some element $\omega \in \Omega$. If χ is a representation of H, then we write χ_H^G for the representation of G induced from χ. The representation 1_H^G is the *permutation representation* of G on the (left) cosets of H. We set d_H to be the minimum degree of a nontrivial irreducible component of the representation 1_H^G; similarly, m_H is the minimum multiplicity of a nontrivial irreducible component of the representation 1_H^G.

We are now able to state two theorems about convolutions, both of which are proved in Section 2. The first is a generalization of [2, Theorem 2.1] and is couched in terms of probability distributions; all norms, here and elsewhere, are ℓ_2-norms. (L. Pyber has pointed out to me that a result similar to Theorem 2 has been proved by Szegedy [28, Corollary 1.3]. Szegedy's result applies not just to finite groups, but more generally to compact Hausdorff topological groups.)

THEOREM 2. *Let G be a finite group acting transitively on a set Ω and let X be a probability distribution over G, Y a probability distribution over Ω. Then*

$$\|X *_c Y - U_\Omega\| \leqslant \sqrt{|G|/d_H} \cdot \|X - U_G\| \cdot \|Y - U_\Omega\|, \tag{1.2}$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

where U_G (respectively U_Ω) is the uniform probability distribution on G (respectively on Ω).

The second convolution theorem is a generalization of [13, Lemma 3.2]. (Gowers did not state his original result in terms of convolution of functions, but he could have if he had wanted to.)

THEOREM 3. *Let G be a finite group acting transitively on a set Ω, let S be any subset of G, let $\chi_S : G \to \mathbb{R}$ be the characteristic function of S and let $f : \Omega \to \mathbb{R}$ be a function that satisfies $\sum_{x\in G} f(x) = 0$. Then*

$$\|\chi_S *_c f\| \leqslant \sqrt{\ell_S|\Omega|/m_H} \cdot \|\chi_S\| \cdot \|f\| \tag{1.3}$$

where $\ell_S = \max\{|g_1 H g_2 \cap S| \mid g_1, g_2 \in G\}$.

The results of Gowers and Babai–Nikolov–Pyber that these two theorems generalize both pertain to the (left) regular action of G on itself. For this action the distinction between d_H and m_H is lost, as both are equal to the minimum dimension of a nontrivial irreducible representation of G. The two theorems, then, highlight one of the main differences between the approach of Gowers (where bounds involve multiplicity, and dimension enters only by virtue of its connection to multiplicity) and the approach of Babai–Nikolov–Pyber (where bounds involve dimension directly).

1.1. General consequences. Theorems 2 and 3 have a number of general consequences for subsets connected to group actions. The first of these is an analogue of the main result of [2] which is itself a variant on the original 'Gowers Trick'.

THEOREM 4. *Let G be a finite group acting transitively on a set Ω, let $S \subseteq G$ and let $\Gamma \subseteq \Omega$. Then the following two inequalities hold:*

$$|S(\Gamma)| > \frac{|\Omega|}{1 + |G||\Omega|/d_H|S||\Gamma|} \geqslant \min\left\{\frac{|\Omega|}{2}, \frac{d_H|S||\Gamma|}{2|G|}\right\}; \tag{1.4}$$

$$|S(\Gamma)| > \frac{|\Omega|}{1 + \ell_S|\Omega|^2/m_H|S||\Gamma|} \geqslant \min\left\{\frac{|\Omega|}{2}, \frac{m_H|S||\Gamma|}{2\ell_S|\Omega|}\right\}. \tag{1.5}$$

In particular, if k is a positive number and $|S| \geqslant \min\{k(|G|/d_H), k(\ell_S|\Omega|/m_H)\}$, then $|S(\Gamma)| > \frac{1}{2}\min\{|\Omega|, k|\Gamma|\}$.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Note that, here and elsewhere, we write group actions on the left. In particular $S(\Gamma) = \{s\gamma \mid s \in S, \gamma \in \Gamma\}$. Recall that ℓ_S was defined in the statement of Theorem 3.

Theorem 4 has a number of consequences. The first is the generalization of Theorem 1 that we mentioned at the start of this paper.

COROLLARY 1.1. *Let G be a finite group acting transitively on a set Ω. Suppose that any nontrivial irreducible component of the corresponding permutation representation has degree at least d_H. Let S be a subset of G and Δ_1, Δ_2 subsets of Ω such that $|S||\Delta_1||\Delta_2| \geqslant |\Omega|^2|G|/d_H$. Then there exist $g \in S$, $\omega_1 \in \Delta_1$ and $\omega_2 \in \Delta_2$ such that $g(\omega_1) = \omega_2$.*

Proof. Write n for $|\Omega|$. The inequality $|S||\Delta_1||\Delta_2| \geqslant n^2|G|/d_H$, combined with the inequality (1.4)—setting $\Gamma = \Delta_1$—implies that

$$|S(\Delta_1)| > \frac{n^2}{n + |\Delta_2|} > n - |\Delta_2|.$$

Now the pigeonhole principle implies that $S(\Delta_1) \cap \Delta_2 \neq \emptyset$ and the result follows. □

The results stated so far take on a particularly interesting aspect when the group H is the centralizer of an element $g \in G$. In this case the action of G on Ω is isomorphic to the action of G on the conjugacy class C which contains g. In this context we have the following corollary, the proof of which is given in Section 4.1.

COROLLARY 1.2. *Let G be a finite group, let C be a conjugacy class of G and let H be the centralizer of an element of C. Suppose that A is a subset of C such that:*

(1) $|A| \geqslant |C|/2$*; and*

(2) $d_H > (8/k)|H|\ell_C$ *for some positive integer k.*

Then $(A \cup A^{-1})^{5+10k} \supseteq C$.

Note that, since C is invariant under conjugation,

$$\ell_C = \max\{|C \cap g_1 H g_2| \mid g_1, g_2 \in G\} = \max\{|C \cap gH| \mid g \in G\}$$

and note that for the rest of this paper we tend to use the symbol A (rather than S) for subsets of G that lie wholly inside a conjugacy class C.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Observe that Corollary 1.2 applies only to very large sets in C—sets that are at least half the size of C. In contrast, Theorem 4 can be applied to much smaller sets. In general, our method will be to apply Theorem 4 first, to obtain expansion results for sets up to half the size of C, and then to use Corollary 1.2 to obtain all of C.

Effectively, then, we use Corollary 1.2 much as the original 'Gowers Trick' of Nikolov and Pyber [23] is used; moreover, our proof of the result is a direct adaptation of that found in [23]. We have not attempted to optimize the value $5 + 10k$; a more involved analysis would substantially decrease this value.

1.2. Consequences for expanders. Let $X = (V, E)$ be a (directed) graph and $\epsilon > 0$ a real number. For a set of vertices $W \subseteq V$, define ∂W to be the number of edges of form (w, y) where $w \in W$ and $y \in V \setminus W$. Now recall that X is called an *ϵ-expander* if

$$\min\left\{\frac{|\partial W|}{|W|} \geqslant \epsilon \mid W \subset V, |W| \leqslant \frac{1}{2}|V|\right\}.$$

Consider a group G acting transitively on a set Ω and let S be a subset of G. Define the Schreier graph $\mathrm{Sch}(G, \Omega, S)$ to be the graph whose vertices are elements of Ω and whose edges are $(\omega, s\omega)$ for every $\omega \in \Omega$ and every $s \in S$.

We aim to construct infinite families of Schreier graphs, $(X_n) = \mathrm{Sch}(G_n, \Omega_n, S_n)$ (where n varies over $\mathbb{N}$) such that each graph in the family is an ϵ-expander, for some absolute constant ϵ. In this case we say that $(X_n)_{n\in\mathbb{N}}$ is an *ϵ-expander family*. We restrict, first of all, to the case where our family consists of graphs which have constant degree d as this is the most interesting (and most difficult).

There are several methods for proving that a given family of Schreier graphs is an ϵ-expander family. The one that interests us here makes use of the product theorems of Helfgott [16, 17] and its generalizations [7, 24]. It was developed, first of all, by Bourgain and Gamburd [5, 6] using (*inter alia*) ideas of Sarnak and Xue [27].

Yehudayoff [31] gives a beautiful explanation of how the Bourgain–Gamburd method works: he breaks this method down into three stages, and it is the last of these, 'the end game' that is of concern to us here. In order to show that $(X_n) = \mathrm{Sch}(G_n, \Omega_n, S_n)$ is a family of ϵ-expanders for $n \in \mathbb{N}$, one needs to prove a lemma of the following form [31, Lemma 4]:

LEMMA 1.3. *There exists a universal constant $c > 0$ so that for every $n \in \mathbb{N}$, for every probability distribution μ_n on G_n and for every function $f_n : \Omega_n \to \mathbb{R}$ that satisfies $\sum_{x\in G_n} f_n(x) = 0$,*

$$\|\mu_n *_c f_n\|^2 \leqslant |G_n|^{1-c} \cdot \|\mu_n\| \cdot \|f_n\|. \tag{1.6}$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

To prove a result of this kind we use Lemma 2.3 to adjust Theorem 2 so that it is stated in terms of 'functions that sum to 0'.

PROPOSITION 1.4. *Let μ be a probability distribution on G and let $f : \Omega \to \mathbb{R}$ be a function that satisfies $\sum_{x\in G} f(x) = 0$. Then*

$$\|\mu *_c f\|^2 \leqslant |G|/d_H \cdot \|\mu\| \cdot \|f\|.$$

The proposition has the following immediate corollary.

COROLLARY 1.5. *Suppose that there exists a constant $c > 0$ and a family $(X_n)_{n\in\mathbb{N}} = \mathrm{Sch}(G_n, \Omega_n, S_n)$ of Schreier graphs such that the minimal dimension of an irreducible component of the permutation representation for the action of G_n on Ω_n is at least $|G_n|^c$. Then (1.6) holds.*

This corollary applies to many of the known constructions of ϵ-expander families:

- *The (left) regular action of G on itself*: Here $\Omega_n = G_n$ and the Schreier graph is actually a Cayley graph. This is the original setting of Bourgain and Gamburd. Note that once one knows that a Cayley graph is an ϵ-expander, then one can use standard results on eigenvalues of adjacency matrices (including, for instance, [**18**, Proposition 11.17]) to prove expansion on other Schreier graphs.

- *The action is 2-transitive*: In this case $1_G^H = 1 + \chi$ where χ is an irreducible representation, and thus $d_H = |\Omega| - 1$. This situation has been studied by Bourgain and Yehudayoff [**8**] and used to construct a *monotone* expander family. Yehudayoff refers to this work in the survey mentioned above, where he also states a special (and weaker) case of Corollary 1.5 [**31**, Lemma 14].

- *Margulis' original family of expanders*: These are expanders corresponding to a family of Schreier graphs $(X_p)_{p \text{ a prime}} = \mathrm{Sch}(\mathrm{AGL}_2(p), (\mathbb{Z}/p\mathbb{Z})^2, S_p)$ where S_p is a particular subset of size 8 in $\mathrm{AGL}_2(p)$. Again, since $\mathrm{AGL}_2(p)$ acts 2-transitively on $(\mathbb{Z}/p\mathbb{Z})^2$, Corollary 1.5 applies.

Thus, of the known ϵ-expander families, the only ones where Corollary 1.5 does not (obviously) apply are those constructed using the zig-zag product pioneered by Reingold *et al.* [**26**].

If one relaxes the condition that the family of graphs be d-regular, then the following result can be used (along with lower bounds for d_H given by [**20**]) to obtain infinite families of ϵ-expander families for (say) any given family of simple groups of Lie type.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

COROLLARY 1.6. *Let G be a finite group acting transitively on a set Ω and let H be the stabilizer of an element of Ω. Let $\delta > 0$ and let S be a subset of G satisfying*

$$|S| \geqslant \min\left\{\frac{(2+\delta)|G|}{d_H}, \frac{(2+\delta)\ell_S|\Omega|}{m_H}\right\}.$$

Then $\mathrm{Sch}(G, S, \Omega)$ *is an ϵ-expander where $\epsilon = \delta/(4+\delta)$.*

Proof. Let Γ be a subset of Ω of size at most $\frac{1}{2}|\Omega|$. The lower bounds on the order of S imply, by Theorem 4, that

$$\begin{aligned}|S(\Gamma)| &> \frac{|\Omega|}{1+|\Omega|/(2+\delta)|\Gamma|} = \frac{(2+\delta)|\Omega||\Gamma|}{(2+\delta)|\Gamma|+|\Omega|} \geqslant \frac{(2+\delta)|\Gamma|}{\frac{1}{2}(4+\delta)}\\ &= \left(1+\frac{\delta}{4+\delta}\right)|\Gamma|.\end{aligned}$$

Now $|\partial\Gamma| \geqslant |S(\Gamma)| - |\Gamma| > (\delta/(4+\delta))|\Gamma|$ and the result follows. □

1.3. Sum-product. We remarked in the previous section that our results are particularly effective when we consider a 2-transitive action of a finite group G. We study a particular instance of such an action in order to prove the following *sum-product result for large sets in finite fields.*

PROPOSITION 1.7. *Let A be a subset of $\mathbb{F}_q\backslash\{0\}$.*

(1) *If $|A| \geqslant q^{2/3}$ then $|A + AA| > q/2$.*

(2) *If $|A| = q^{1/2+\delta}$ for some $\delta \in (0, \frac{1}{6})$, then $|A + AA| > \frac{1}{2}q^{1/2+3\delta}$.*

Proof. We apply Theorem 4 to the following situation: $G = (\mathbb{F}_q, +) \rtimes (\mathbb{F}_q^*, \cdot)$ acting as a 1-dimensional affine group on $\Omega = (\mathbb{F}_q, +)$. The group G here is isomorphic to $E_q \rtimes C_{q-1}$, a semidirect product of an elementary-abelian group of order q with a cyclic group of order $q-1$. Observe that, for $(a, b) \in G$, $c \in \Omega$,

$$(a, b)(c) = a + bc. \tag{1.7}$$

The action of G on Ω is 2-transitive hence, as we observed in the previous section, $d_H = |\Omega| - 1 = q - 1$.

Next define sets

$$S = \{(a_1, a_2) \mid a_1, a_2 \in A\} \quad \text{and} \quad \Gamma = A,$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

and observe that (1.7) implies that $S(\Gamma) = A + AA$. Now Theorem 4 can be applied and (1.4) yields that

$$|A + AA| = |S(\Gamma)| > \frac{q}{1 + q(q-1)q/(q-1)|A|^3} = \frac{q|A|^3}{|A|^3 + q^2}. \tag{1.8}$$

Suppose first that $|A| \geqslant q^{2/3}$. Then (1.8) implies that $|A + AA| > q/2$ as required. On the other hand, if $|A| = q^{1/2+\delta}$ for some $\delta \in (0, \frac{1}{6})$, then (1.8) implies that

$$|A + AA| > \frac{q|A|^3}{2q^2} = \frac{1}{2}q^{1/2+3\delta}. \qquad \square$$

Note that the condition that $0 \notin A$ is included only to facilitate the cleanest statement possible. There are a number of comparable sum-product results for large subsets of finite fields; we refer particularly to [11] and to [14, 15]. (M. Rudnev has pointed out to me that Proposition 1.7 can be proved in an alternative way, as a consequence of a Szemerédi–Trotter type theorem (for instance [30, Theorem 3]). The proof goes as follows: for each $x \in A$, $y \in A+AA$, one defines a line l_{xy} in $(\mathbb{F}_q)^2$ as the set of (a, b) such that $a + bx = y$ (cf. (1.7)). Define $\mathcal{L}$ to be the set of all such lines and define $\mathcal{P}$ to be the set $A \times A \subset (\mathbb{F}_q)^2$. Observe that the set of incidences of $\mathcal{L}$ with $\mathcal{P}$ is at least $|A|^3$ (since every triple $(a, b, x) \in A^3$ yields a value $y \in A$). Then, since $|\mathcal{L}| = |A| \cdot |A + AA|$ and $|\mathcal{P}| = |A|^2$, [30, Theorem 3] yields the result. Analogous methods yield similar results in the Euclidean plane.)

1.4. Diameter and width. Our original motivation for this paper was to try and solve two outstanding conjectures in group theory. The first posits an upper bound on the *diameter* of a Cayley graph of a finite nonabelian simple group.

CONJECTURE 1.8 [3, Conjecture 1.7]. (Babai's conjecture) *There exists an absolute constant c such that, if G is a finite nonabelian simple group and S is a generating subset of G, we have $G = A^k$ where $k \leqslant (\log |G|)^c$.*

The second posits an upper bound on a *width* of a finite nonabelian simple group.

CONJECTURE 1.9 [21]. (The product decomposition conjecture) *There exists an absolute constant c such that if G is a finite nonabelian simple group and S is a subset of G of size at least two, then G is a product of N conjugates of S for some $N \leqslant c \log |G|/\log |S|$.*

Both of these conjectures are proved for groups of Lie type of bounded rank [7, 12, 24]. We are able to give partial results for groups of Lie type of unbounded

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

rank that complement those already in the literature due to the original Gowers trick.

PROPOSITION 1.10. *Fix α a positive real number, let n be odd and let $G = A_n$, the alternating group on n letters. Let C be a conjugacy class of n-cycles and suppose that $S \subset G$ such that $S \cap C \neq \emptyset$ and so that*

$$|S| \geqslant \left(\frac{1}{(1/2)n(n-3)}\right)^{1-\alpha}|G|.$$

Then there exists a positive integer k, depending only on α, such that $G = (S \cup S^{-1})^k$.

Elements in the conjugacy class C here can be characterized as regular semisimple elements whose centralizer is a 'maximally nonsplit torus' (or, in other language, whose centralizer is a *Singer cycle*).

PROPOSITION 1.11. *Fix α a positive real number, let $G = \mathrm{SL}_n(2)$ and let C be a conjugacy class of elements whose eigenvalues lie in no proper subfield of $\mathbb{F}_{2^n}$. Suppose that $S \subset G$ such that $S \cap C \neq \emptyset$ and so that*

$$|S| \geqslant \left(\frac{3}{(2^n-1)(2^n-4)}\right)^{1-\alpha}|G|.$$

Then there exists a positive integer k, depending only on α, such that $G = (S \cup S^{-1})^k$.

We emphasize that in neither of these two propositions does the integer k depend on the variable n. Notice too that in neither proposition have we needed to assume that S generates G—this fact is implied by the suppositions on S. Significantly the lower bound on $|S|$ is not enough to guarantee generation in either case—one needs the extra supposition on the intersection with C. This also explains why the lower bounds that we require are weaker than those required by other versions of the 'Gowers trick' which apply to arbitrary sets in A_n and $\mathrm{SL}_n(2)$ [2, 23].

The two propositions (which are proved in Section 5) imply that Babai's conjecture and the Product Decomposition Conjecture hold for the set $S \cup S^{-1}$ and the group G in each case. Indeed, [1, Corollary 2.3] implies that Babai's conjecture holds for the set S and the group G in both cases.

1.5. Structure of the paper. Theorem 2 is proved in Section 2 using the linear algebra methods of Babai–Nikolov–Pyber. Theorem 3 is proved in Section 3 using

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

the graph-theoretic methods of Gowers. In Section 4, we derive Theorem 4 from Theorems 2 and 3; we also prove Corollary 1.2. Propositions 1.10 and 1.11 are proved in Section 5. Finally we conclude with Section 6 in which we discuss possible future directions for research.

2. The first convolution theorem

This section is devoted to a proof of Theorem 2. We use the notation established in the introduction without further comment. Note that, in this section, *all matrices are real*.

2.1. Circulants. If E is a matrix whose rows (respectively columns) are labelled by elements of a set $X = \{x_1, \ldots, x_m\}$ (respectively $Y = \{y_1, \ldots, y_n\}$) then we write $E(x_i, y_j)$ (or simply $E(i, j)$) for the entry in matrix E at row x_i, column y_j where $x_i \in X$, $y_j \in Y$.

A matrix E is said to be *biregular* if its row sums are all equal to a constant $s_r(E)$, and its column sums are all equal to a constant $s_c(E)$. Note that the product of biregular matrices (if defined) is biregular, and the quantities s_r and s_c are multiplicative.

LEMMA 2.1 [2, Proposition 5.2]. *If E is a nonnegative biregular $k \times n$ matrix, then*

$$\lambda_1(E^T E) = s_r(E)s_c(E)$$

and a corresponding eigenvector is $\mathbf{1}_n = (1, \ldots, 1)^T$.

Recall that a *G-circulant* of a group G is a $|G|$-by-$|G|$ matrix M, with rows labelled by elements of G and columns labelled by elements of G, and such that

$$M(g, h) = M(1, g^{-1}h). \tag{2.1}$$

We extend this idea: for a set Ω on which G acts we define a *$G\Omega$-circulant* to be a $|G|$-by-$|\Omega|$ matrix M, with rows labelled by elements of G and columns labelled by elements of Ω, and such that

$$M(g, \omega) = M(1, g^{-1}\omega). \tag{2.2}$$

Observe that a G-circulant is simply a $G\Omega$-circulant where we take $\Omega = G$ and consider the regular left action of G on itself.

LEMMA 2.2. *A $G\Omega$-circulant E is biregular, and $s_c(E) = (|G|/|\Omega|)s_r(E)$.*

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Proof. To see that row sums are constant, observe that, for $g \in G$,

$$\sum_{\omega\in\Omega} M(g,\omega) = \sum_{\omega\in\Omega} M(1,g^{-1}\omega) = \sum_{\omega\in g^{-1}(\Omega)} M(1,\omega) = \sum_{\omega\in\Omega} M(1,\omega).$$

To see that column sums are constant, observe that, for $\omega \in \Omega$,

$$\sum_{g\in G} M(g,\omega) = \sum_{g\in G} M(1,g^{-1}\omega) = |\mathrm{Stab}_G(\omega)| \sum_{\omega\in\Omega} M(1,g) = \frac{|G|}{|\Omega|} s_r(E).$$

This completes the proof. □

2.2. Functions. Let Λ be any set and $Z : \Lambda \to \mathbb{R}$ a function. We need some definitions:

If Z satisfies the property $\sum_{\lambda\in\Lambda} Z(\lambda) = 1$ then we call Z a *probability distribution*. The function Z is said to be *concentrated* on the subset Ξ of Ω if $Z(g) = 0$ whenever $g \in \Lambda \setminus \Xi$. We define the norm of Z as the positive square root of $\|Z\|^2 = \sum_{\lambda\in\Lambda} Z(\lambda)^2$.

2.2.1. Convolution. Consider two functions $X : G \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$. At (1.1) we defined the notion of *convolution* for X and Y, namely:

$$(X *_c Y)(\omega) = \sum_{g\in G} X(g)Y(g^{-1}\omega).$$

Observe that

$$\sum_{\omega\in\Omega} (X *_c Y)(\omega) = \left(\sum_{g\in G} X(g)\right)\left(\sum_{\omega\in\Omega} Y(\omega)\right).$$

In particular, if X and Y are probability distributions then $X *_c Y$ is also a probability distribution; on the other hand, if either X or Y sum to 0 then $X *_c Y$ sums to 0.

The key fact about convolutions is this: Suppose that $X : G \to \mathbb{R}$ is concentrated on $S \subset G$, and $Y : \Omega \to \mathbb{R}$ is concentrated on $\Gamma \subset \Omega$; it follows that $(X *_c Y)$ is concentrated on $S(\Gamma)$.

2.2.2. Norms. We close this section with a number of facts about norms.

LEMMA 2.3. *Let Z be a function on Ω that sums to 0, Y be a probability distribution over Ω, X a probability distribution over G, and U the uniform probability distribution over Ω. Then:*

(1) $\|Z + U\|^2 = \|Z\|^2 + 1/|\Omega|$;

(2) $\|Y - U\|^2 = \|Y\|^2 - 1/|\Omega|$;

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

(3) $\|X *_c (Y \pm U)\| = \|X *_c Y \pm U\|$;

(4) *for k a real number* $\|kY\| = k\|Y\|$.

Proof. For the first fact observe that

$$\|Z+U\|^2 = \sum_{\omega\in\Omega}\left(Z(\omega)+\frac{1}{|\Omega|}\right)^2 = \|Z\|^2+\frac{1}{|\Omega|}+\frac{2}{|\Omega|}\sum_{\omega\in\Omega}Z(\omega) = \|Z\|^2+\frac{1}{|\Omega|}.$$

For the second fact observe that

$$\|Y-U\|^2 = \sum_{\omega\in\Omega}\left(Y(\omega)-\frac{1}{|\Omega|}\right)^2 = \|Y\|^2+\frac{1}{|\Omega|}-\frac{2}{|\Omega|}\sum_{\omega\in\Omega}Y(\omega) = \|Y\|^2-\frac{1}{|\Omega|}.$$

For the third fact observe that

$$\begin{aligned}\|X *_c (Y-U)\|^2 &= \sum_{\omega\in\Omega}\left(\sum_{g\in G}X(g)\left(Y(g^{-1}\omega)\pm\frac{1}{|\Omega|}\right)\right)^2\\ &= \sum_{\omega\in\Omega}\left(\sum_{g\in G}X(g)Y(g^{-1}\omega)\pm\sum_{g\in G}X(g)\frac{1}{|\Omega|}\right)^2\\ &= \sum_{\omega\in\Omega}\left(\sum_{g\in G}X(g)Y(g^{-1}\omega)\pm\frac{1}{|\Omega|}\right)^2\\ &= \|X *_c Y \pm U\|^2.\end{aligned}$$

The final fact is immediate. □

2.3. Functions and circulants. Let us connect the concepts of the last two subsections. Throughout this subsection we consider functions $X : G \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$. We define the *$G\Omega$-circulant of Y* to be the $G\Omega$-circulant B such that $B(g,\omega) = Y(g^{-1}\omega)$.

We note a special case of this definition: we consider the natural left regular action of G on itself; in this case $G = \Omega$ and we have a $G\Omega$-circulant A for the function $X : G \to \mathbb{R}$. Now A is actually a G-circulant, since it satisfies $A(g, h) = X(g^{-1}h)$ and, so as not to confuse matters, we call A the *G-circulant of Y*.

Observe that if Y is a probability distribution then $s_r(B) = 1$, and hence $s_c(B) = |G|/|\Omega|$.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Note the following analogue of [2, (5.25)].

LEMMA 2.4. *Let B be a $G\Omega$-circulant of Y. Then*

$$\|Y\|^2 = \frac{1}{|G|}\mathrm{Tr}(BB^T).$$

Proof.

$$\begin{aligned}\mathrm{Tr}(BB^T) &= \sum_{g\in G}\left(\sum_{\omega\in\Omega} B(g,\omega)B^T(\omega,g)\right)\\ &= \sum_{g\in G}\left(\sum_{\omega\in\Omega}(B(g,\omega))^2\right)\\ &= \sum_{g\in G}\left(\sum_{\omega\in\Omega}(B(1,g^{-1}\omega))^2\right)\\ &= |G|\cdot\|Y\|^2.\end{aligned}$$

□

LEMMA 2.5. *Let A be the G-circulant for X, let B be the $G\Omega$-circulant for Y, and let D be the $G\Omega$-circulant for $X *_c Y$. Then $D = AB$.*

Proof. Observe that

$$\begin{aligned}AB(g,\omega) &= \sum_{h\in G} A(g,h)B(h,\omega)\\ &= \sum_{h\in G} X(g^{-1}h)Y(h^{-1}\omega)\\ &= \sum_{h\in G} X(h)Y(h^{-1}g^{-1}\omega)\\ &= (X *_c Y)(g^{-1}\omega)\\ &= D(g,\omega),\end{aligned}$$

as required. □

Lemmas 2.4 and 2.5 can be combined to yield an analogue of [2, Proposition 5.6].

PROPOSITION 2.6. *Let A be the G-circulant for X, and let B be the $G\Omega$-circulant for Y. Then*

$$\|X *_c Y\|^2 = \frac{1}{|G|}\mathrm{Tr}(ABB^TA^T).$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

2.4. Connection with representation dimension. Consider a vector space $\mathbb{R}^{|G|}$ (respectively $\mathbb{R}^{|\Omega|}$); we fix a basis and label each element of the basis with an element of G (respectively Ω). We consider three linear maps as follows.

2.4.1. A basis for $G\Omega$-circulants. For $\omega \in \Omega$ define a linear map

$$\rho_\omega : \mathbb{R}^{|G|} \to \mathbb{R}^{|\Omega|}, \quad g \mapsto g\omega.$$

Representing elements of $\mathbb{R}^{|G|}$ as row vectors, the corresponding matrix representation B_ω of ρ_ω (via postmultiplication) is

$$B_\omega(g, \gamma) = \begin{cases} 1 & \gamma = g\omega, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if we represent elements of $\mathbb{R}^{|G|}$ as column vectors, then the corresponding matrix representation ρ_ω (via premultiplication) is B_ω^T.

The key fact concerning the matrices B_ω is this: the $G\Omega$-circulant of a function $X : \Omega \to \mathbb{R}$ lies in the span of the set $\{B_\omega \mid \omega \in \Omega\}$.

2.4.2. The left regular representation. For $g \in G$ define two linear maps

$$\tau_g : \mathbb{R}^{|G|} \to \mathbb{R}^{|G|}, \quad h \mapsto g^{-1}h;$$
$$\tau_g^o : \mathbb{R}^{|G|} \to \mathbb{R}^{|G|}, \quad h \mapsto gh.$$

These two actions correspond to the *left regular representation* of G (written as a right (respectively left) action).

We represent elements of $\mathbb{R}^{|G|}$ as row vectors and write X_g for the matrix representation of τ_g via postmultiplication, so $\tau_g : h \mapsto hX_g$.

On the other hand, if we represent elements of $\mathbb{R}^{|G|}$ as column vectors then X_g is also the matrix representation of τ_g^o via premultiplication, so $\tau_g^o : h \mapsto X_g h$.

2.4.3. The permutation representation. For $g \in G$ define a linear map

$$\sigma_g : \mathbb{R}^{|\Omega|} \to \mathbb{R}^{|\Omega|}, \quad \omega \mapsto g^{-1}\omega;$$
$$\sigma_g^o : \mathbb{R}^{|\Omega|} \to \mathbb{R}^{|\Omega|}, \quad \omega \mapsto g\omega.$$

These two actions correspond to the *permutation representations* for G acting on Ω (written as a right (respectively left) action).

Now we represent elements of $\mathbb{R}^{|\Omega|}$ as row vectors and write Y_g for the matrix representation of σ_g via postmultiplication, so $\sigma_g : \omega \mapsto \omega Y_g$.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

On the other hand, if we represent elements of $\mathbb{R}^{|\Omega|}$ as column vectors then Y_g is also the matrix representation of σ_g^o via premultiplication, so $\sigma_g^o : \omega \mapsto Y_g\omega$.

We have already seen the notation 1_H^G for the representation σ_g^o. Note that, since the permutation action associated with 1_H^G is transitive, we have $\langle 1_H^G, 1_G\rangle = 1$ [19, (5.15)].

2.4.4. Commuting actions. The following lemma connects the three linear maps we have just defined. The fourth identity will be the one we use directly: it asserts that, for $g \in G$, the matrix $1_H^G(g)$ commutes with matrices of the form $B_\omega^T B_\omega$.

LEMMA 2.7. *For all $g \in G$ and $\omega \in \Omega$, the following hold:*

(1) $X_g B_\omega = B_\omega Y_g$;

(2) $Y_g B_\omega^T = B_\omega^T X_g$;

(3) $X_g B_\omega B_\omega^T = B_\omega B_\omega^T X_g$;

(4) $Y_g B_\omega^T B_\omega = B_\omega^T B_\omega Y_g$.

Proof. For the first identity, let $x \in G$ be represented as a row vector of length G. Then

$$xX_g B_\omega = (g^{-1}x)B_\omega = g^{-1}x\omega.$$

On the other hand,

$$xB_\omega Y_g = (x\omega)Y_g = g^{-1}x\omega.$$

The result follows.

For the second identity, let $x \in G$ be represented as a column vector of length G. Then

$$Y_g B_\omega^T x = Y_g(x\omega) = gx\omega.$$

On the other hand,

$$B_\omega^T X_g x = B_\omega^T(gx) = gx\omega.$$

The result follows.

Now the first two identities imply that

$$X_g B_\omega B_\omega^T = B_\omega Y_g B_\omega^T = B_\omega B_\omega^T X_g$$

and the third identity follows. Similarly, for the fourth identity, we have

$$Y_g B_\omega^T B_\omega = B_\omega^T X_g B_\omega = B_\omega^T B_\omega Y_g.$$

□

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

2.4.5. Symmetric matrices. Before we proceed to the proof of Theorem 2, we need a couple of easy results about symmetric matrices.

Observe first that if B is a real matrix, then B^TB is a symmetric matrix. Recall that every n-by-n real symmetric matrix U has n real eigenvalues, counting geometric multiplicities, and we denote them by

$$\lambda_1(U) \geqslant \lambda_2(U) \geqslant \cdots \geqslant \lambda_n(U).$$

Furthermore, B^TB is positive semidefinite, because

$$x^TEx = (x^TB^T)(Bx) = \|Bx\|^2 \geqslant 0,$$

which means that all eigenvalues of BB^T are real and nonnegative.

In the proof of the next lemma we use I to denote the n-by-n identity matrix, for any positive integer n.

LEMMA 2.8. *Suppose that B is a real matrix. Then BB^T and B^TB have the same nonzero eigenvalues, counting geometric multiplicities.*

Proof. Given a nonzero real number λ we can define a linear map from $\ker(B^TB - \lambda I)$ to $\ker(BB^T - \lambda I)$ by $v \mapsto Bv$. This is well defined, because $BB^T(Bv) = B(B^TB)(v) = B(\lambda v) = \lambda Bv$. It is injective, because if $Bv = 0$ then $\lambda v = B^TBv = 0$, which means $v = 0$. We can also define an injective linear map $v \mapsto B^Tv$ from $\ker(BB^T - \lambda I)$ to $\ker(B^TB - \lambda I)$. Therefore, both eigenspaces have the same dimension, as required. □

Note that, in particular, $\mathrm{Tr}(BB^T) = \mathrm{Tr}(B^TB)$.

2.5. Proof of Theorem 2. We are just about ready to give a proof of Theorem 2 using the methods of [2]. Recall that $X : \Omega \to \mathbb{R}$ and $Y : G \to \mathbb{R}$ are probability distributions; in particular, this means that the corresponding circulants are nonnegative real matrices. This is crucial in what follows (and will not apply when we come to prove Theorem 3).

In this section, we write U_Ω (respectively U_G) for the uniform probability distribution over the set Ω (respectively over G). We begin with an analogue of [2, Lemma 5.7].

PROPOSITION 2.9. *Let $H = \mathrm{Stab}_G(\omega)$ and let d_H be the minimum degree of an irreducible component of the representation 1_H^G. If B is a nonnegative $G\Omega$-circulant, then*

$$\lambda_2(BB^T) \leqslant \frac{\mathrm{Tr}(BB^T) - \lambda_1(BB^T)}{d_H}. \tag{2.3}$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Proof. Let $D = BB^T$ and $E = B^TB$. Since E is symmetric and positive semidefinite, all eigenvalues of E are real and nonnegative. We denote them by

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{|\Omega|}.$$

Lemma 2.8 implies that the eigenvalues of D are

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{|\Omega|} \geqslant 0 = 0 = \cdots = 0.$$

Observe that, since B is a $G\Omega$-circulant, it is biregular and so the same is true of B^T. Now Lemma 2.1 implies that $\lambda_1(E) = s_r(B)s_c(B)$, and a corresponding eigenvector is $\mathbf{1} = (1, \ldots, 1)$.

Observe that the representation 1_H^G preserves the one-dimensional subspace spanned by $\mathbf{1}$ (since it is a permutation representation). Then, since $\langle 1_H^G, 1_G\rangle = 1$, all other subspaces stabilized by 1_H^G have nontrivial irreducible components.

Now, since $1_H^G(g)$ commutes with E for every $g \in G$ (this is the fourth identity of Lemma 2.7) it follows that all eigenspaces of E are stabilized by 1_H^G. It follows that the multiplicity of every eigenvalue of the restriction of E to U is at least d_H. Lemma 2.8 implies that the same can be said for the multiplicity of every eigenvalue of the restriction of D to U; in particular, it is true of the eigenvalue $\lambda_2(D)$. Since the trace of D restricted to U is $\mathrm{Tr}(D) - \lambda_1(D)$ we conclude that

$$\mathrm{Tr}(D) - \lambda_1(D) \geqslant d_H\lambda_2(D). \qquad \square$$

LEMMA 2.10 [**2**, Lemma 5.8]. *Let A and B be nonnegative biregular matrices such that the product AB is defined. Then*

$$\mathrm{Tr}(B^TA^TAB) \leqslant \lambda_1(A^TA)\lambda_1(B^TB) + \lambda_2(A^TA)(\mathrm{Tr}(B^TB) - \lambda_1(B^TB)).$$

We are now ready to prove Theorem 2.

Proof of Theorem 2. Let A be the G-circulant for X and B be the $G\Omega$-circulant for Y. Proposition 2.6 and Lemma 2.10 imply that

$$\begin{aligned}\|X *_c Y\|^2 &= \frac{1}{|G|}\mathrm{Tr}(ABB^TA^T)\\ &\leqslant \frac{1}{|G|}\lambda_1(BB^T)\lambda_1(AA^T) + \frac{1}{|G|}\lambda_2(BB^T)(\mathrm{Tr}(AA^T) - \lambda_1(AA^T)).\end{aligned}$$

Because AA^T is nonnegative and biregular we see that $\lambda_1(AA^T) = s_r(A)s_c(A) = 1$, and using Lemma 2.2 we see that $\lambda_1(BB^T) = s_r(B)s_c(B) = |G|/|\Omega|$. Using Lemma 2.3 it follows that

$$\|X *_c Y - U_\Omega\|^2 = \|X *_c Y\|^2 - \frac{1}{|\Omega|} \leqslant \frac{1}{|G|}\lambda_2(BB^T)(\mathrm{Tr}(AA^T) - 1).$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Therefore, by Proposition 2.9,

$$\begin{aligned}\|X *_c Y - U_\Omega\|^2 &\leqslant \frac{1}{|G|d_H}\left(\mathrm{Tr}(BB^T) - \frac{|G|}{|\Omega|}\right)(\mathrm{Tr}(AA^T) - 1)\\ &= \frac{|G|}{d_H}\left(\|Y\|^2 - \frac{1}{|\Omega|}\right)\left(\|X\|^2 - \frac{1}{|G|}\right)\\ &= \frac{|G|}{d_H}\|Y - U_\Omega\|^2\|X - U_G\|^2.\end{aligned}$$

□

3. The second convolution theorem

In this section, we prove Theorem 3 using the methods of Gowers [13]. Although much of Gowers' work can be reframed without referring to his original graph-theoretic setting this would seem to be a mistake: it is difficult to retain intuition about what is going on once one has 'linearized' and written everything in terms of matrices. In addition, the geometry of the group action is nicely encapsulated by the graphs that Gowers considers and so we make use of them here.

3.0.1. Bipartite graphs. In what follows $\mathcal{G}$ is a bipartite graph with vertex sets X and Y. We write $\mathcal{A}$ for the adjacency matrix of $\mathcal{G}$. Note that, unlike for Gowers, our graph $\mathcal{G}$ is not necessarily simple, that is, we allow the possibility that there is more than one edge between two vertices. This implies, in particular, that the entries of $\mathcal{A}$ may exceed 1.

Our first job is to analyse $\mathcal{A}$ and for this we will need some notation given on [13, page 7]. We let V and W be real vector spaces with the usual inner product. For $v \in V$, $w \in W$ define the linear map

$$w \otimes v : V \to W, \quad x \mapsto \langle x, v\rangle w.$$

We need the following result:

PROPOSITION 3.1 [13, Theorem 2.6]. *Let $\alpha : V \to W$ be a linear map. Then there exists a decomposition $\alpha = \sum_{i=1}^{k} \lambda_i w_i \otimes v_i$ where the sequences (w_i) (respectively (v_i)) are orthonormal in W (respectively V), the sequence (λ_i) is real, nonnegative and nonincreasing, and $k = \min\{\dim V, \dim W\}$.*

Note, in addition, that the sequence (λ_i) is uniquely determined, and that the vector v_1 can be taken to be any vector such that, for all $v \in V$,

$$\|\alpha(v_1)\|/\|v_1\| \geqslant \|\alpha(v)\|/\|v\|.$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

The last sentence of Proposition 3.1 does not appear in the statement of [13, Theorem 2.6] but is clear from the proof.

Our next result is an analogue of [13, Lemma 2.7] adjusted to hold for graphs which are not simple; in fact we will only need part of the original lemma.

LEMMA 3.2. *Let $\mathcal{G}$ be a bipartite graph with vertex sets X and Y and identify $\mathcal{G}$ with its bipartite adjacency matrix $\sum_{i=1}^{k} \lambda_i w_i \otimes v_i$, where (v_i) and (w_i) are orthonormal sequences. Then the number of edges in $\mathcal{G}$ is greater than or equal to*

$$\frac{1}{\ell} \sum_{i=1}^{k} \lambda_i^2$$

where ℓ is the maximum number of edges between any two vertices of $\mathcal{G}$.

Proof. Observe first that $\mathcal{A}^T$ is $\sum_i \lambda_i v_i \otimes w_i$ and that

$$(v_i \otimes w_i)(w_j \otimes v_j) = \begin{cases} v_i \otimes v_i & i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Now $\mathrm{Tr}(v_i \otimes v_i) = 1$ and thus $\mathrm{Tr}(\mathcal{A}^T \mathcal{A}) = \sum_i \lambda_i^2$.

But now $\frac{1}{\ell}\mathrm{Tr}(\mathcal{A}^T \mathcal{A})$ is less than or equal to the number of edges in $\mathcal{G}$. □

We use the graph $\mathcal{G}$ to define the following map:

$$\alpha : \mathbb{R}^X \to \mathbb{R}^Y, \quad f \mapsto \alpha f$$

where, for $f : X \to \mathbb{R}$ we have

$$\alpha f : Y \to \mathbb{R}, \quad y \mapsto \sum_{x \in X, xy \in E(\mathcal{G})} f(x). \tag{3.1}$$

Note that, if there is more than one edge between two vertices x and y, then our definition of $(\alpha)(f)(y)$ requires that the value $f(x)$ is added multiple times—once for each edge between x and y.

The map α will be central in what follows and we shall see in the next subsection that it is closely related to the idea of convolution.

The following lemma contains everything that we need to know about the map α. In the statement of the lemma, the graph $\mathcal{G}$ is assumed to be *regular*, that is, every vertex in X has the same degree and every vertex in Y has the same degree. We set λ_i, v_i, w_i and k to be as defined in Proposition 3.1.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

LEMMA 3.3. *Suppose that $\mathcal{G}$ is a regular bipartite graph. The following hold.*

(1) $\lambda_1 = \max\{\|\alpha(f)\|/\|f\| \mid f \in \mathbb{R}^{\Omega}\} = \|\alpha(v_1)\|/\|v_1\|$.

(2) *We can take v_1 to be the constant function*

$$X \to \mathbb{R}, \quad x \mapsto \frac{1}{\sqrt{|X|}}. \tag{3.2}$$

(3) *The set $\mathcal{F}$ of functions $X \to \mathbb{R}$ that sum to zero is a vector space of dimension $k-1$.*

(4) *For all $f \in \mathcal{F}$, $\|\alpha(f)\|/\|f\| \leqslant \lambda_2$.*

(5) *Let e be the positive integer such that*

$$\lambda_2 = \lambda_3 = \cdots = \lambda_e > \lambda_{e+1}.$$

Then the set $\mathcal{E}$ of functions in $\mathcal{F}$ such that $\|\alpha(f)\|/\|f\| = \lambda_2$ is a vector space (provided we include 0) of dimension e.

Proof. Observe first that

$$\alpha\left(\sum_{i=1}^{k} \mu_i v_i\right) = \sum_{i=1}^{k} \lambda_i \mu_i w_k. \tag{3.3}$$

In particular, (1) holds.

To prove (2) we set p to be the real number such that every vertex in X has degree $p|Y|$; observe that, since $\mathcal{G}$ is regular, every vertex in Y has degree $p|X|$. Now

$$\begin{aligned}
\|\alpha f\|^2 &= \sum_y \left| \sum_x f(x)\mathcal{A}(x,y) \right|^2 \\
&= \sum_{x,x'} f(x) f(x') \sum_y \mathcal{A}(x,y)\mathcal{A}(x',y) \\
&\leqslant \frac{1}{2} \sum_{x,x'} (f(x)^2 + f(x')^2) \sum_y \mathcal{A}(x,y)\mathcal{A}(x',y) \\
&= \sum_x f(x)^2 \sum_{x'} \sum_y \mathcal{A}(x,y)\mathcal{A}(x',y) \\
&= \sum_x f(x)^2 p^2 |X||Y| = p^2 |X||Y| \|f\|^2.
\end{aligned}$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

It follows that $\|\alpha(f)\|/\|f\| \leqslant p\sqrt{|X|\cdot|Y|}$. Now let $f = 1$, the function defined at (3.2) and we have $\|\alpha f\| = p|X|\sqrt{|Y|}$ and $\|f\| = \sqrt{|X|}$ in which case $\|\alpha(f)\|/\|f\| = p\sqrt{|X|\cdot|Y|}$ and (2) follows.

Item (3) is immediate once we observe that $\mathcal{F}$ is the orthogonal complement of the function (3.2). Taking v_1 to be this function (by (2)) we conclude that $\mathcal{F}$ is spanned by $\{v_2, \dots, v_k\}$ and the map $\alpha|_{\mathcal{F}}$ can be decomposed as $\sum_{i=2}^{k} \lambda_i w_i \otimes v_i$. Then (4) follows by applying (1) to this decomposition.

Applying (1) to the vector space $\mathcal{F}$ we observe that $\|\alpha(f)\|/\|f\| \leqslant \lambda_2$ for all $f \in \mathcal{E}$. Furthermore, (3.3) implies that $\|\alpha(f)\|/\|f\| = \lambda_2$ if and only if f is in the span of $\{v_2, \dots, v_e\}$. Now (5) is immediate. □

3.0.2. Graphs from groups. We return to the setting where G is a group acting transitively on a set Ω and S is a subset of G. We will work with the following bipartite graph, $\mathcal{G}$: the two vertex sets, X and Y, are copies of Ω and xy is an edge if and only if there exists $s \in S$ such that $s(x) = y$. Note that this graph is *regular*, that is, every vertex in X has the same degree and every vertex in Y has the same degree.

As before we write $\mathcal{A}$ for the adjacency matrix of $\mathcal{G}$. Observe that, for $x, y \in \Omega$, $\mathcal{A}(x, y)$ is the number of edges from x to y in $\mathcal{G}$.

If S is a subset of G we write χ_S for the characteristic function of S. Now, for this particular graph $\mathcal{G}$, we can use the more general definition of convolution given at (1.1) to describe the function α defined at (3.1) in a different way:

$$\alpha f(\omega) = \sum_{\nu\in\Omega} \mathcal{A}(\nu, \omega) f(\nu) = \sum_{g\in G} \chi_S(g) f(g^{-1}\omega) = (\chi_S *_c f)(\omega). \tag{3.4}$$

In other words $\alpha(f) = \chi_S *_c f$.

Note that the linear function $\alpha : \mathbb{R}^\Omega \to \mathbb{R}^\Omega$ has associated matrix $\mathcal{A}^T$. Note, moreover, that

$$\mathcal{A} = \sum_{g\in S} Y_{g^{-1}}$$

where, for $g \in G$, the matrix Y_g was defined in Section 2.4.3. With these observations in mind we are ready to prove Theorem 3. This is the analogue of [**13**, Lemma 3.2] and, in Gowers' language, asserts that the graph $\mathcal{G}$ is quasirandom.

Proof of Theorem 3. Let $\mathcal{G}$ be the bipartite Cayley graph defined above and observe that ℓ_S is equal to the maximum number of edges between vertices in $\mathcal{G}$. Observe too that $\mathcal{G}$ is regular and let α be the associated linear map (3.4).

By the observations above, the associated matrix for α (once we fix a basis) is equal to $\sum_{g\in S} Y^T_{g^{-1}}$. Since the matrices $Y_{g^{-1}}$ correspond to the permutation representation 1^G_H, these matrices then preserve a decomposition of $\mathbb{R}^\Omega$ into

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

subspaces, one for each irreducible component of the representation 1_H^G. Then the vectors $v_1, \ldots, v_{|\Omega|}$ can be chosen to lie inside these subspaces.

Suppose that the vector v_i lies inside a subspace W corresponding to an irreducible component χ of 1_H^G. It is easy to see that the corresponding real number λ_i will occur in the sequence $(\lambda_1, \ldots, \lambda_{|\Omega|})$ with multiplicity at least the multiplicity of the irreducible component χ, this multiplicity being $\langle \chi, 1_H^g \rangle \geqslant m_H$.

Let $\mathcal{E}$ and $\mathcal{F}$ be the vector spaces defined in Lemma 3.3. Referring to item (1) of that lemma we take v_1 to be the constant function (3.2). The subspace $\langle v_1 \rangle$ is preserved by the matrices $Y_{g^{-1}}$, as is $\mathcal{F}$, the orthogonal complement of v_1. Moreover, since $\langle 1_H^G, 1_G \rangle = 1$, the subspace $\langle v_1 \rangle$ is the unique 1-dimensional subspace of $\mathbb{R}^\Omega$ that is stabilized by Y_g for all $g \in G$. Hence, in particular, all of the subspaces of $\mathcal{F}$ stabilized by the matrices Y_g correspond to irreducible components of 1_H^G with multiplicity at least m_H. We conclude that the vector space $\mathcal{E}$ must have dimension at least m_H, that is, that the real number λ_2 occurs with multiplicity at least m_H.

Lemma 3.2 implies that $\frac{1}{\ell} m_H \lambda_2^2$ is less than or equal to the number of edges in $\mathcal{G}$. But $\mathcal{G}$ has $|S| \cdot |\Omega|$ edges and we conclude that

$$\lambda_2 \leqslant \sqrt{\ell_S |\Omega| / m_H} \cdot \sqrt{|S|}. \tag{3.5}$$

Lemma 3.3 part (4) implies that if $f : X \to \mathbb{R}$ is a function that sums to zero, then $\|(\alpha f)\| / \|f\| \leqslant \lambda_2$. Observing that $\|\chi_S\| = \sqrt{|S|}$ and substituting into (3.5) we obtain

$$\|\alpha f\| / \|f\| \leqslant \lambda_2 \leqslant \sqrt{\ell_S |\Omega| / m_H} \cdot \|\chi_S\|.$$

Now (3.4) gives the result. □

4. Large sets grow

Proof of Theorem 4. Let X be the probability distribution over G, Y the probability distribution over Ω given by the following definitions:

$$X(x) = \begin{cases} \dfrac{1}{|S|} & x \in S, \\ 0 & x \notin S, \end{cases} \qquad Y(x) = \begin{cases} \dfrac{1}{|\Gamma|} & x \in \Gamma, \\ 0 & x \notin \Gamma. \end{cases}$$

Observe that $\|X\| = 1/\sqrt{|S|}$ and $\|Y\| = 1/\sqrt{|\Gamma|}$. Recall that $X *_c Y$ is *concentrated* on $S(\Gamma)$, meaning that $(X *_c Y)(g) = 0$ whenever $g \in \Omega \setminus S(\Gamma)$.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

A simple application of the Cauchy–Schwarz inequality (or see [2, Observation 3.4]) gives

$$\frac{1}{|S(\Gamma)|} \leqslant \|X *_c Y\|^2. \tag{4.1}$$

This inequality, with Lemma 2.3 and Theorem 2, imply that

$$\begin{aligned}\frac{1}{|S(\Gamma)|} &\leqslant \|X *_c Y\|^2\\ &\leqslant \frac{1}{|\Omega|} + \|X *_c Y - U_\Omega\|^2\\ &\leqslant \frac{1}{|\Omega|} + \frac{|G|}{d_H}\|X - U_G\|^2\|Y - U_\Omega\|^2\\ &< \frac{1}{|\Omega|} + \frac{|G|}{d_H}\|X\|^2\|Y\|^2\\ &= \frac{1}{|\Omega|} + \frac{|G|}{d_H}\frac{1}{|S|}\frac{1}{|\Gamma|}.\end{aligned}$$

Rearranging we obtain

$$|S(\Gamma)| > \frac{|\Omega|}{1 + (|G||\Omega|/d_H|S||\Gamma|)},$$

which is the first inequality of (1.4). For the second inequality, observe that if $(|G||\Omega|/d_H|S||\Gamma|) \leqslant 1$ then

$$\frac{|\Omega|}{1 + (|G||\Omega|/d_H|S||\Gamma|)} \geqslant \frac{|\Omega|}{2}.$$

On the other hand, if $(|G||\Omega|/d_H|S||\Gamma|) > 1$ then

$$\frac{|\Omega|}{1 + (|G||\Omega|/d_H|S||\Gamma|)} = \frac{|\Omega|d_H|S||\Gamma|}{d_H|S||\Gamma| + |G||\Omega|} > \frac{|\Omega|d_H|S||\Gamma|}{2|G||\Omega|} = \frac{d_H|S||\Gamma|}{2|G|}.$$

In both cases, the second inequality holds.

Now we must prove (1.5). We begin by defining U_Ω to be the uniform probability distribution over Ω and observe that $f = Y - U_\Omega$ is a function on Ω that sums to 0. Observe too that $\chi_S = |S|p_S$.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Now we start with (4.1), apply Theorem 3 and make use of the identities in Lemma 2.3:

$$\begin{aligned}\frac{1}{|S(\Gamma)|} &\leqslant \|X *_c Y\|^2 \\ &= \|X *_c (f + U_\Omega)\|^2 \\ &= \|X *_c f + U_\Omega\|^2 \\ &= \|X *_c f\|^2 + \frac{1}{|\Omega|} \\ &= \frac{1}{|S|^2}\|\chi_S *_c f\|^2 + \frac{1}{|\Omega|} \\ &\leqslant \frac{1}{|S|^2} \cdot \frac{\ell_S|\Omega|}{m_H}\|\chi_S\|^2\|f\|^2 + \frac{1}{|\Omega|} \\ &= \frac{1}{|S|^2} \cdot \frac{\ell_S|\Omega|}{m_H} \cdot |S| \cdot \|p_\Gamma - U_\Omega\|^2 + \frac{1}{|\Omega|} \\ &< \frac{\ell_S|\Omega|}{m_H}\frac{1}{|S|}\frac{1}{|\Gamma|} + \frac{1}{|\Omega|}.\end{aligned}$$

Rearranging we obtain

$$|S(\Gamma)| > \frac{|\Omega|}{1 + (\ell_S|\Omega|^2/m_H|S||\Gamma|)},$$

which is the first inequality of (1.5). The second inequality follows just as for (1.4). □

4.1. Corollary 1.2. In this subsection, we prove Corollary 1.2. By way of introduction we state a weaker result, the proof of which illustrates our methods.

COROLLARY 4.1. *Let G be a finite group and let C be a conjugacy class of G. Let H be the centralizer of an element of C and let A be a subset of C. Suppose that*

(1) $|A| \geqslant |C|/2$*; and*

(2) $d_H > 8|H|\ell_C$.

Then

$$(A \cup A^{-1})^5 \supseteq AAAA^{-1}A^{-1} \supseteq C.$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Proof. Write n for $|C|$. We apply Corollary 1.1 with $S = \Delta_1 = A$ and Δ_2 the set of elements that are **not** in the set $S(\Delta_1)$, that is, are not of the form $a_1 a_2 a_1^{-1}$ for some $a_1, a_2 \in A_1$. We use the fact that $\ell_S = \ell_A \leqslant \ell_C$ and obtain that

$$|\Delta_2| \leqslant \frac{n^2|G|\ell_C}{m_H} \Big/ \left(\frac{n}{2}\right)^2 = \frac{4|G|}{d_H}.$$

Thus the set $A_2 = \bigcup_{a\in A} aAa^{-1}$ has size at least

$$n - |\Delta_2| \geqslant n - \frac{4|G|}{d_H}. \tag{4.2}$$

Now, for $g \in C$, define $B_g = \{a^{-1}ga \mid a \in A_1\}$ and observe that

$$|B_g| \geqslant \frac{|A|}{\ell_C} \geqslant \frac{n}{2\ell_C}.$$

Now, since $e_H > 8|H|\ell_C$, a little rearranging yields that

$$\frac{n}{2\ell_C} > \frac{4|G|}{d_H}.$$

Thus, by the pigeonhole principle $B_g \cap A_2$ is nonempty for every $g \in C$. We conclude, therefore, that

$$A_3 = \bigcup_{a\in A} aA_2a^{-1} = C.$$

Now

$$A_3 \subseteq AA_2A^{-1} \subseteq AAAA^{-1}A^{-1}$$

and the result follows. □

It turns out that the bound (2) needed for Corollary 4.1 is too strong for wide application, hence the need for the stronger statement given in Corollary 1.2.

Proof of Corollary 1.2. We define the sets A_2 and B_g as per the previous proof, and we recall (4.2):

$$|(A \cup A^{-1})^3 \cap C| \geqslant |A_2| \geqslant |C| - \frac{4|G|}{d_H}.$$

Using the fact that $m_H > (8/k)|H|\ell_C$ we observe that

$$|B_g| \geqslant \frac{|A|}{\ell_C} \geqslant \frac{|C|}{2\ell_C} > \frac{4|G|}{k \cdot d_H}.$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

The first step of our proof involves building a set X with particular properties; we begin by setting $X = \emptyset$. Now suppose that, for all $g_1 \in C \backslash A_2$, we have

$$B_{g_1} \cap \left(A_2 \cup \bigcup_{g \in X} B_g \right) = \emptyset.$$

In this case we add g_1 to our set X and repeat. Since $|B_g| \geqslant 4|G|/(k \cdot d_H)$ we can repeat this process until X has size at most k, at which point no such g_1 will exist. In this case we stop.

By way of comparison with the previous result note that if $X = \emptyset$ then we obtain immediately that $B_g \cap A_2 \neq \emptyset$ for every $g \in C$ and we obtain, as required that

$$(A \cup A^{-1})^5 \supseteq AA_2A^{-1} \supseteq C.$$

If X is not empty, we have a little more work to do. Observe first that

$$A_2 \cup \bigcup_{g \in X} AB_gA^{-1} \supseteq C.$$

Now AA_2A^{-1} is strictly larger than A_2 and hence intersects AB_gA^{-1} for some $g \in X$. Thus $A^{-1}AA_2A^{-1}A$ intersects B_g and thus

$$g \in AA^{-1}AA_2A^{-1}AA^{-1}.$$

Then $B_g \subset A^{-1}AA^{-1}AA_2A^{-1}AA^{-1}A$ and, finally,

$$AB_gA^{-1} \subseteq AA^{-1}AA^{-1}AA_2A^{-1}AA^{-1}AA^{-1}.$$

Since $A_2 \subseteq (A \cup A^{-1})^3$ we obtain that

$$AB_gA^{-1} \subseteq (A \cup A^{-1})^{13}.$$

Now we repeat the process with A_2 redefined to be $(A \cup A^{-1})^{13} \cap C$. We can repeat this at most k times at the end of which A_2 is the set $(A \cup A^{-1})^{3+10k} \cap C$ and it has the property that $B_g \cap A_2 \neq \emptyset$ for every $g \in C$. Now we conclude, as in the previous proof, that

$$(A \cup A^{-1})^{5+10k} \supseteq AA_2A^{-1} \supseteq C. \qquad \square$$

5. Simple groups

In this section we prove Propositions 1.10 and 1.11. We need a lemma.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

LEMMA 5.1. *Let G be a finite group and let C be a conjugacy class of G. Let H be the centralizer of an element of C and let S be a subset of G. Suppose that there exists a positive number α such that:*

(1) $|S| \geqslant (1/d_H)^{1-\alpha}|C|$;

(2) $|S \cap C| \geqslant (1/d_H)^3|C|$.

Then $|(S \cup S^{-1})^{2\lceil 3/\alpha\rceil - 1} \cap C| \geqslant |C|/2$.

Proof. Applying Theorem 4 with $\Gamma = S \cap C$ we conclude that

$$(S \cup S^{-1})^3 \geqslant \tfrac{1}{2}\min\{|C|, (d_h)^{\alpha-3}|C|\}.$$

Iterating we conclude that, for k a positive integer,

$$(S \cup S^{-1})^{2k-1} \geqslant \tfrac{1}{2}\min\{|C|, (d_h)^{k\alpha-3}|C|\}.$$

Taking $k = \lceil 3/\alpha\rceil$ the result follows. □

It will be convenient to use the following result of Liebeck and Shalev [22]. (This is a spectacular sledgehammer to crack a couple of rather tiny nuts. Nonetheless it saves us some tiresome computations in each case.) Note that a *normal subset* of a group is a union of conjugacy classes.

THEOREM 5. *There exists an absolute positive constant a such that, if G is a finite simple group and S is a nontrivial normal subset of G, then $G = S^m$, where $m \leqslant a(\log|G|/\log|S|)$.*

5.1. Alternating groups. This section is devoted to a proof of Proposition 1.10. We write representations of S_n in the standard way: indexed by partitions of n. Then [25] implies:

LEMMA 5.2. *Suppose that $n \geqslant 15$. The first seven minimal character degrees d of S_n are given by representations S^λ as follows:*

(1) $d = 1$ *and* $\lambda \in \{(n), (1^n)\}$;

(2) $d = n-1$ *and* $\lambda \in \{(n-1, 1), (2, 1^{n-2})\}$;

(3) $d = \frac{1}{2}n(n-3)$ *and* $\lambda \in \{(n-2, 2), (2, 2, 1^{n-4})\}$;

(4) $d = \frac{1}{2}(n-1)(n-2)$ *and* $\lambda \in \{(n-2, 1, 1), (3, 1^{n-3})\}$;

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

(5) $d = \frac{1}{6}n(n-1)(n-5)$ *and* $\lambda \in \{(n-3,3),(2,2,2,1^{n-6})\}$;

(6) $d = \frac{1}{6}(n-1)(n-2)(n-3)$ *and* $\lambda \in \{(n-3,1^3),(4,1^{n-4})\}$;

(7) $d = \frac{1}{3}n(n-2)(n-4)$ *and* $\lambda \in \{(n-3,2,1),(3,2,1^{n-5})\}$.

Note that there are two representations in each case; they correspond to tensoring by the sign representation. In terms of partitions they correspond to reflecting in the diagonal.

Note that, for $n \geqslant 15$, the two partitions listed in each case are distinct, that is, tensoring by the sign representation yields a nonisomorphic representation. It follows by [9, Proposition 5.1] that the representations given in Lemma 5.2 stay irreducible when restricted to A_n. Furthermore any irreducible representation of A_n is obtained by restriction from an irreducible representation of S_n and it will either have the same degree (as in the above cases for $n \geqslant 15$) or half the original degree. Since $\frac{1}{3}n(n-2)(n-4)$ is more than double $\frac{1}{2}n(n-3)$ for $n \geqslant 15$ we conclude the following:

LEMMA 5.3. *Suppose that* $n \geqslant 15$. *The first three minimal character degrees* d *of* A_n *are given by representations* S^λ *as follows:*

(1) $d = 1$ *and* $\lambda \in \{(n),(1^n)\}$;

(2) $d = n-1$ *and* $\lambda \in \{(n-1,1),(2,1^{n-2})\}$;

(3) $d = \frac{1}{2}n(n-3)$ *and* $\lambda \in \{(n-2,2),(2,2,1^{n-4})\}$.

Note that although, again, there are two partitions listed in this case, the corresponding representations of A_n are isomorphic to each other. Thus there is really only one representation of A_n of the given degree.

Let H be a subgroup of $G = A_n$ and observe that $(1_H^G)_G^{S_n} = 1_H^{S_n}$. Consider θ to be a character of S_n. Then Frobenius reciprocity implies that

$$\langle 1_H^G, \theta|_G \rangle = \langle (1_H^G)_G^{S_n}, \theta \rangle = \langle 1_H^{S_n}, \theta \rangle \tag{5.1}$$

where θ_G is the restriction of θ to G.

If $n \geqslant 15$ and θ is one of the characters of S_n associated with the representations $S^{(n)}$, $S^{(n-1,1)}$, $S^{(n-2,2)}$, then θ corresponds to a partition which is not symmetric through the diagonal. Thus [9, Proposition 5.1] implies that θ_G is one of the three minimal characters listed in Lemma 5.3.

We need a result of Frobenius. From here on, for a partition λ of n we write θ^λ for the character of S_n associated with the representation S^λ; similarly we write χ^λ for the character of A_n associated with S^λ.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

LEMMA 5.4. *Let $H \leqslant S_n$ be a permutation group acting on $\{1, 2, \ldots, n\}$. Let $t_r(H)$ be the number of orbits of H on r-subsets of $\{1, 2, \ldots, n\}$. If $0 \leqslant r \leqslant n/2$ then*

$$\langle 1_H^{S_n}, \theta^{(n-r,r)} \rangle = \begin{cases} t_r(H) - t_{r-1}(H) & r \geqslant 1; \\ t_0(H) = 1 & r = 0. \end{cases} \tag{5.2}$$

Now (5.1) implies an immediate corollary of Lemma 5.4:

LEMMA 5.5. *Let $H \leqslant A_n$ be a permutation group acting on $\{1, 2, \ldots, n\}$. Let $t_r(H)$ be the number of orbits of H on r-subsets of $\{1, 2, \ldots, n\}$. If $0 \leqslant r \leqslant n/2$ then*

$$\langle 1_H^{G}, \chi^{(n-r,r)} \rangle = \begin{cases} t_r(H) - t_{r-1}(H) & r \geqslant 1; \\ t_0(H) = 1 & r = 0. \end{cases} \tag{5.3}$$

Recall that a permutation group H on $\{1, \ldots, n\}$ is called r-homogeneous (for $r > 1$ an integer) if H is transitive on the r-subsets of $\{1, \ldots, n\}$. We need one more lemma:

LEMMA 5.6. *Let n be odd, $G = A_n$ and g an n-cycle in G. Let H be transitive on $\{1, \ldots, n\}$ but not 2-homogeneous. The minimum degree of a nontrivial irreducible component of 1_H^G is equal to $\frac{1}{2}n(n-3)$.*

Proof. By Lemmas 5.3 and 5.5 we must show that $t_1(H) = 1$ and $t_2(H) \geqslant 1$. That $t_1(H) = 1$ follows from the fact that H is transitive; since H is not 2-homogeneous we conclude that $t_2(H) \geqslant 1$. □

If H is the centralizer of an n-cycle in G, then H is transitive but not 2-homogeneous. Thus we have the following:

COROLLARY 5.7. *If H is the centralizer of an n-cycle in G, then $d_H = \frac{1}{2}n(n-3)$.*

We can now prove Proposition 1.10.

Proof. Note first that the result is trivial for n less than any absolute constant. It will suit us to assume from here on that $n > 100$. Set $H = C_G(g)$ and observe that Corollary 5.7 implies that $d_H = \frac{1}{2}n(n-3)$. Since $|H| = n$, we apply the pigeonhole principle to the cosets of H to conclude that

$$|(S \cup S^{-1})^3| \geqslant \left| \bigcup_{s,g \in S} sgs^{-1} \right| \geqslant \frac{1}{n}\left(\frac{1}{d_H}\right)^{1-\alpha} |C| > \frac{1}{n}\left(\frac{1}{d_H}\right)|C| > \left(\frac{1}{d_H}\right)^3 |C|.$$

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Thus we can apply Lemma 5.1 to the set $(S \cup S^{-1})^3$ to conclude that

$$|(S \cup S^{-1})^{6\lceil 3/\alpha \rceil - 1} \cap C| \geqslant \frac{|C|}{2}.$$

Let A_1 be the set $(S \cup S^{-1})^{6\lceil 2/\alpha \rceil - 1} \cap C$ and observe that, since $n \geqslant 100$, we have $m_c > \frac{8}{20}|H|^2 \geqslant \frac{8}{20}\ell_C|H|$. Then Corollary 1.2, applied to A_1 with $k = 20$, implies that $(S \cup S^{-1})^{1230\lceil 2/\alpha \rceil}$ contains C.

Now Theorem 5 gives the result. □

5.2. $\mathbf{SL}_n(2)$. This section is devoted to a proof of Proposition 1.11. We first need a result of Tiep and Zalesskii [29].

LEMMA 5.8. *Let $G = \mathrm{SL}_n(2)$ with $n \geqslant 6$. Let χ_1 (respectively χ_2) be the nontrivial complex representation of smallest (respectively second smallest) degree. Then*

$$\chi_1(1) = 2^n - 2, \quad \chi_2(1) = \tfrac{1}{3}(2^n - 1)(2^{n-1} - 4).$$

COROLLARY 5.9. *Let H be a maximally split torus in G. Then*

$$d_H = \tfrac{1}{3}(2^n - 1)(2^{n-1} - 4).$$

Proof. Observe first that $|H| < \chi_2(1)$. Next we show that $\langle 1_H^G, \chi_1 \rangle = 0$.

Consider the action of G on nontrivial vectors in the natural module. The stabilizer of a point in this action is a parabolic subgroup $P = P_1$. Since this action is 2-transitive we conclude that $1_P^G = 1_G + \pi$ for some irreducible complex representation of degree $2^n - 2$. Thus $\pi = \chi_1$ and $1_P^G = 1_G + \chi_1$.

Let K be a subgroup of G and consider 1_K^G. Using Frobenius reciprocity we have

$$\langle 1_K^G, 1_P^G \rangle = \langle 1_K, (1_P^G)|_H \rangle = \langle 1_K, 1_{P\cap K}^K \rangle.$$

Now $\langle 1_K, 1_{P\cap K}^K \rangle$ is equal to the number of orbits of K on the nontrivial vectors in the natural module. Now consider the situation when $K = H$, a maximally split torus. Then H has a single orbit on nontrivial vectors and so

$$\langle 1_H^G, 1_P^G \rangle = 1.$$

Since $\langle 1_H^G, 1_G \rangle = 1$ we conclude that $\langle 1_H^G, \chi_1 \rangle = 0$. □

We are ready to prove Proposition 1.11.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Proof. Once again observe that the result is trivial for n less than any absolute constant and assume from here on that $n > 100$. Set $H = C_G(g)$, a maximally split torus, and observe that Corollary 5.9 implies that $d_H = \frac{1}{3}(2^n - 1)(2^{n-1} - 4)$.

Since $|H| = 2^n - 1$, we apply the pigeonhole principle to the cosets of H to conclude that $|(S \cup S^{-1})^3|$ exceeds

$$\left|\bigcup_{s,g\in S} sgs^{-1}\right| \geqslant \frac{1}{2^n - 1}\left(\frac{1}{d_H}\right)^{1-\alpha}|C| > \frac{1}{2^n - 1}\left(\frac{1}{d_H}\right)|C| > \left(\frac{1}{d_H}\right)^3|C|.$$

Thus we can apply Lemma 5.1 to the set $(S \cup S^{-1})^3$ to conclude that

$$|(S \cup S^{-1})^{6\lceil 2/\alpha\rceil - 1} \cap C| \geqslant \frac{|C|}{2}.$$

Let A_1 be the set $(S \cup S^{-1})^{6\lceil 2/\alpha\rceil - 1} \cap C$ and observe that, since $n \geqslant 100$, we have $m_c > \frac{8}{49}|H|^2$. Then Corollary 1.2, applied to A_1 with $k = 49$, implies that $(S \cup S^{-1})^{3000\lceil 2/\alpha\rceil}$ contains C.

Now Theorem 5 gives the result. □

6. Further work

There is plenty of scope for further work.

6.1. Quasirandom group actions. Clearly [13] is an El Dorado of a paper and we have mined but a small portion of it for our inspiration here. The latter parts of the paper (which we have neglected) put the notion of a *d-quasirandom group* on a firm footing, and present a number of different ways of characterizing such groups.

Our work suggests that these ideas belong more properly in the more general setting of *d-quasirandom group actions*. We have not defined this notion formally in the body of the paper, however, as it is not entirely clear what the 'correct' definition should be. Theorem 2 suggests that a transitive group action should be called *d-quasirandom* if $d_H \geqslant d$. This would interact well, for instance, with our treatment of expanders, as per Corollary 1.5.

The problem is that Theorem 3 also implies mixing properties for a different class of large set. The following lemma suggests that, since Theorem 3 is expressed in terms of m_H rather than d_H, one might suspect that the bound it specifies is weaker than Theorem 2.

LEMMA 6.1. *Let $J < H < G$ and let χ be an irreducible character of G. Then* $\langle 1_J^G, \chi\rangle \geqslant \langle 1_H^G, \chi\rangle$.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

Proof. We use Frobenius reciprocity:

$$\langle 1_J^G, \chi \rangle = \langle (1_J^H)_H^G, \chi \rangle = \langle 1_J^H, \chi|_H \rangle \geqslant \langle 1_H, \chi|_H \rangle = \langle 1_H^G, \chi \rangle. \tag{6.1}$$

□

(Note that, when $J = \{1\}$, 1_J^G is the (right) regular permutation character, and so $\langle 1_J^G, \chi \rangle = \dim(\chi)$. In particular, we obtain that $m_H \leqslant d_H$.)

Working in favour of Theorem 3, however, is the fact that $\ell_S \leqslant |G|/|\Omega|$. Thus, for particular sets S it is conceivable that Theorem 3 will be stronger than Theorem 2.

A further complicating factor is that, although Theorem 2 can be rewritten using Lemma 2.3, so that it is stated in terms of arbitrary functions that sum to 0 (see Proposition 1.4), the reverse process cannot be applied to Theorem 3. The method of proof for Theorem 3 uses specific properties of χ_S and does not admit (obvious) generalization to arbitrary measures on the set S.

6.2. The quantity ℓ_C. The considerations just discussed suggest that the size of the quantity ℓ_S should have a bearing in attempts to understand how quasirandomness interacts with arbitrary group actions.

Let us focus on the case when G acts by conjugation on a conjugacy class C. In this case ℓ_S is bounded above by the quantity

$$\ell_C = \max\{C \cap gH \mid g \in G\}.$$

(Here H is the centralizer of an element of C.)

The computation of ℓ_C would seem potentially more tractable than the computation of ℓ_S for arbitrary S. (Indeed to make use of our results one only needs an upper bound on ℓ_C.) However, we have been unable to make any general statements other than the obvious one: $\ell_C \leqslant |H|$.

There is reason to believe that better bounds hold. For instance, for the cases discussed in Sections 5.1 and 5.2, we have the following conjectures.

CONJECTURE 6.2. *Let $G = A_n$ with n odd and let C be a conjugacy class of n-cycles with H a centralizer of an element of C. Then*

$$\max\{|gH \cap C| \mid g \in G\} = |H \cap C| = |N_G(H) : H| \in \{\phi(n), \phi(n)/2\}.$$

Here ϕ is Euler's totient function.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

CONJECTURE 6.3. *Let $G = \mathrm{SL}_n(2)$ and let C be a conjugacy class of elements centralized by a maximally split torus, and let H be a centralizer of an element of C. Then*

$$\max\{|gH \cap C| \mid g \in G\} = |H \cap C| = |N_G(H) : H| = n.$$

In both conjectures the first equality is the difficult one. In both cases, too, the equality has been verified using GAP and MAGMA for small values of n [**4**, **10**]. A proof of these conjectures would immediately yield stronger versions of Propositions 1.10 and 1.11.

6.3. Minimally quasirandom actions. The results listed in Section 1.4 demonstrate that Theorem 4 can be applied to actions other than the (left) regular action of a group on itself. How many other such actions exist?

In order to answer this question we need to exclude some obvious redundancy. Observe first that Lemma 6.1 implies that if $H < N < G$, then $d_N \leqslant d_H$. Consider what happens when $d_N = d_H$: the bounds given in Theorem 4 apply equally to the action of G on cosets of H, as well as on cosets of N. However, in a sense, the growth in the action of N is simply a function of growth on the cosets of H, and is of its limited interest in its own right.

We propose, then, the following definition. We write $1 < d_1 < d_2 < \cdots$ for the degrees of the irreducible characters of G and, for i a positive integer, we say that (G, H) is an *i-minimal* QR-*action* if the following conditions are satisfied:

(1) the minimal degree of a nontrivial component of 1_H^G is at least d_i;

(2) if $F < H$ then the minimal degree of a nontrivial component of 1_H^G is strictly less than d_i;

(3) $d_i > |H|$.

If G is perfect, that is, $G = [G, G]$; then all nontrivial characters of G have degree strictly greater than 1 and we conclude that $(G, \{1\})$ is the only 1-minimal QR-action. This is the action to which the original Gowers trick applied. It is easy to check that the actions (G, H) discussed in Sections 5.1 and 5.2 are 2-minimal QR-actions. Now the question remains: can we classify all such actions for all simple groups, indeed for all perfect groups?

Acknowledgements

It is a pleasure to thank Mark Wildon, Ian Short, Jan Saxl, Misha Rudnev, Jeremy Rickard, Laci Pyber, Marty Isaacs, Jack Button and John Britnell (in

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

reverse alphabetical order!) for their generous help with various parts of this paper. In particular, the main idea of Section 5.2 is due to Jan Saxl.

References

[1] L. Babai, 'On the diameter of Eulerian orientations of graphs', in *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms* (ACM, New York, 2006), 822–831.

[2] L. Babai, N. Nikolov and L. Pyber, 'Product growth and mixing in finite groups', in *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (ACM, New York, 2008), 248–257.

[3] L. Babai and Á. Seress, 'On the diameter of permutation groups', *European J. Combin.* **13**(4) (1992), 231–243.

[4] W. Bosma, J. Cannon and C. Playoust, 'The Magma algebra system I. The user language', *J. Symbolic Comput.* **24**(3–4) (1997), 235–265. Computational algebra and number theory (London, 1993).

[5] J. Bourgain and A. Gamburd, 'On the spectral gap for finitely-generated subgroups of SU(2)', *Invent. Math.* **171**(1) (2008), 83–121.

[6] J. Bourgain and A. Gamburd, 'Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$', *Ann. of Math.* (2) **167**(2) (2008), 625–642.

[7] E. Breuillard, B. Green and T. Tao, 'Approximate subgroups of linear groups', *Geom. Funct. Anal.* **21**(4) (2011), 774–819.

[8] J. Bourgain and A. Yehudayoff, 'Monotone expansion', in *STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing* (ACM, New York, 2012), 1061–1078.

[9] W. Fulton and J. Harris, *Representation Theory*, *Graduate Texts in Mathematics, 129*, (Springer, New York, 1991), A first course, Readings in Mathematics.

[10] The GAP Group, *GAP–Groups, Algorithms, and Programming, Version 4412*, 2008.

[11] M. Z. Garaev, 'The sum-product estimate for large subsets of prime fields', *Proc. Amer. Math. Soc.* **136**(8) (2008), 2735–2739.

[12] N. Gill, L. Pyber, I. Short and E. Szabó, 'On the product decomposition conjecture for finite simple groups', *Groups Geom. Dyn.* **7**(4) (2013), 867–882.

[13] W. T. Gowers, 'Quasirandom groups', *Comb. Probab. Comp.* **17** (2008), 363–387.

[14] D. Hart and A. Iosevich, 'Sums and products in finite fields: an integral geometric viewpoint', *Radon transforms, geometry, and wavelets, Contemp. Math.* **464** (2008), 129–135.

[15] D. Hart, A. Iosevich, D. Koh and M. Rudnev, 'Averages over hyperplanes, sum-product theory in vector space over finite fields and the Erdös–Falconer distance conjecture', *Trans. Amer. Math. Soc.* **363**(6) (2011), 3255–3275.

[16] D. Hart, A. Iosevich, D. Koh and M. Rudnev, 'Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$', *J. Eur. Math. Soc.* **13**(3) (2011), 761–851.

[17] H. A. Helfgott, 'Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$', *Ann. of Math.* (2) **167**(2) (2008), 601–623.

[18] S. Hoory, N. Linial and A. Wigderson, 'Expander graphs and their applications', *Bull. Amer. Math. Soc. (N.S.)* **43**(4) (2006), 439–561. (electronic).

[19] I. M. Isaacs, *Character Theory of Finite Groups* (AMS Chelsea Publishing, Providence, RI, 2006), Corrected reprint of the 1976 original [Academic Press, New York].

[20] V. Landazuri and G. M. Seitz, 'On the minimal degrees of projective representations of the finite Chevalley groups', *J. Algebra* **32** (1974), 418–443.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press

[21] M. W. Liebeck, N. Nikolov and A. Shalev, 'Product decompositions in finite simple groups', *Bull. Lond. Math. Soc.* **44**(3) (2012), 469–472.
[22] M. W. Liebeck and A. Shalev, 'Diameters of finite simple groups: sharp bounds and applications', *Ann. of Math.* (2) **154**(2) (2001), 383–406.
[23] N. Nikolov and L. Pyber, 'Product decompositions of quasirandom groups and a Jordan-type theorem', *J. Eur. Math. Soc.* **13**(4) (2011), 1063–1077.
[24] L. Pyber and E. Szabó, 'Growth in finite simple groups of Lie type of bounded rank', *J. Amer. Math. Soc.* **29** (2016), 95–146.
[25] R. Rasala, 'On the minimal degrees of characters of S_n', *J. Algebra* **45**(1) (1977), 132–181.
[26] O. Reingold, S. Vadhan and A. Wigderson, 'Entropy waves, the zig-zag graph product, and new constant-degree expanders', *Ann. of Math.* (2) **155**(1) (2002), 157–187.
[27] P. Sarnak and X. X. Xue, 'Bounds for multiplicities of automorphic representations', *Duke Math. J.* **64**(1) (1991), 207–227.
[28] B. Szegedy, 'Limits of kernel operators and the spectral regularity lemma', *European J. Combin.* **32**(7) (2011), 1156–1167.
[29] P. H. Tiep and A. E. Zalesskii, 'Minimal characters of the finite classical groups', *Comm. Algebra* **24**(6) (1996), 2093–2167.
[30] L. A. Vinh, 'The Szemerédi–Trotter type theorem and the sum-product estimate in finite fields', *European J. Combin.* **32**(8) (2011), 1177–1181.
[31] A. Yehudayoff, 'Proving expansion in three steps', *ACM SIGACT News* **43**(3) (2012), 67–84.

https://doi.org/10.1017/fms.2016.8 Published online by Cambridge University Press