On regular hypergraphs of high girth

David Ellis*

Nathan Linial

School of Mathematical Sciences Queen Mary, University of London U.K. School of Computer Science and Engineering Hebrew University of Jerusalem Israel

d.ellis@qmul.ac.uk

nati@cs.huji.ac.il

Submitted: Nov 1, 2013; Accepted: Feb 23, 2014; Published: Mar 10, 2014 Mathematics Subject Classifications: 05C65

Abstract

We give lower bounds on the maximum possible girth of an r-uniform, d-regular hypergraph with at most n vertices, using the definition of a hypergraph cycle due to Berge. These differ from the trivial upper bound by an absolute constant factor (viz., by a factor of between 3/2 + o(1) and 2 + o(1)). We also define a random r-uniform 'Cayley' hypergraph on the symmetric group S_n which has girth $\Omega(\sqrt{\log |S_n|})$ with high probability, in contrast to random regular r-uniform hypergraphs, which have constant girth with positive probability.

1 Introduction

The girth of a finite graph G is the shortest length of a cycle in G. (If G is acyclic, we define its girth to be ∞ .) The girth problem asks for the minimum possible number of vertices n(g,d) in a d-regular graph of girth at least g, for each pair of integers $d,g \ge 3$. Equivalently, for each pair of integers $n,d \ge 3$ with nd even, it asks for a determination of the largest possible girth $g_d(n)$ of a d-regular graph on at most n vertices.

The girth problem has received much attention for more than half a century, starting with Erdős and Sachs [11]. A fairly easy probabilistic argument shows that for any integers $d, g \ge 3$, there exist d-regular graphs with girth at least g. An extremal argument due to Erdős and Sachs [11] then shows that there exists such a graph with at most

$$2\frac{(d-1)^{g-1}-1}{d-2}$$

vertices. This implies that

$$g_d(n) \geqslant (1 - o(1)) \log_{d-1} n.$$
 (1)

^{*}Research supported in part by a Feinberg Visiting Fellowship from the Weizmann Institute of Science.

(Here, and below, o(1) stands for a function of n that tends to zero as $n \to \infty$.)

On the other hand, if G is a d-regular graph of girth at least g, then counting the number of vertices of G of distance less than g/2 from a fixed vertex of G (when g is odd), or from a fixed edge of G (when G is even), immediately shows that

$$|G| \geqslant n_0(g,d) := \begin{cases} 1 + d \sum_{i=0}^{k-1} (d-1)^i &= 1 + d \frac{(d-1)^k - 1}{d-2} & \text{if } g = 2k+1; \\ 2 \sum_{i=0}^{k-1} (d-1)^i &= 2 \frac{(d-1)^k - 1}{d-2} & \text{if } g = 2k. \end{cases}$$

This is known as the *Moore bound*. Graphs for which the Moore bound holds with equality are known as *Moore graphs* (for odd g), or *generalized polygons* (for even g). It is known that Moore graphs only exist when g = 3 or 5, and generalized polygons only exist when g = 4, 6, 8 or 12. It was proved in [1, 5, 17] that if $d \ge 3$, then

$$n(g,d) \ge n_0(g,d) + 2$$
 for all $g \notin \{3,4,5,6,8,12\};$

even for large values of q and d, no improvement on this is known.

A related problem is to give an explicit construction of a d-regular graph of girth g, with as few vertices as possible. The celebrated Ramanujan graphs constructed by Lubotzsky, Phillips and Sarnak [22], Margulis [26] and Morgenstern [27] constituted a breakthrough on both problems, implying that

$$g_d(n) \geqslant (4/3 - o(1)) \log_{d-1} n$$
 (2)

via an explicit (algebraic) construction, whenever d = q + 1 for some odd prime power q. One can obtain from this a lower bound on $g_d(n)$ for arbitrary $d \ge 3$, by choosing the minimum $d' \ge d$ such that d' - 1 is an odd prime power, taking a d'-regular Ramanujan graph with girth achieving (2), and removing d' - d perfect matchings in succession. This yields

$$g_d(n) \geqslant (4/3 - o(1)) \frac{\log(d-1)}{\log(d'-1)} \log_{d-1} n.$$
 (3)

In [19] and [20], Lazebik, Ustimenko and Woldar give different explicit constructions (also algebraic), which imply that

$$g_d(n) \geqslant (4/3 - o(1)) \log_d n$$

whenever d is an odd prime power, implying (3) whenever d-1 is not an odd prime power. (In fact, their constructions provide the best known upper bound on n(g,d) for many pairs of values (g,d).) Combining (3) with the Moore bound gives

$$(4/3 - o(1)) \frac{\log(d-1)}{\log(d'-1)} \log_{d-1} n \leqslant g_d(n) \leqslant (2 + o(1)) \log_{d-1} n. \tag{4}$$

Improving the constants in (4) seems to be a very hard problem.

In this paper, we investigate an analogue of the girth problem for r-uniform hypergraphs, where $r \ge 3$. There are several natural notions of a cycle in a hypergraph. We

refer the reader to Section 4 for a brief discussion of some other interesting notions of girth in hypergraphs, and to [9] for a detailed treatise. Here, we consider the least restrictive notion, originally due to Berge (see for example [3] and [4]).

A hypergraph H is a pair of finite sets (V(H), E(H)), where E(H) is a family of subsets of V(H). The elements of V(H) are called the vertices of H, and the elements of E(H) are called the edges of H. A hypergraph is said to be r-uniform if all its edges have size r. It is said to be d-regular if each of its vertices is contained in exactly d edges. It is said to be linear if any two of its edges share at most one vertex.

Let u and v be distinct vertices in a hypergraph H. A u-v path of length l in H is a sequence of distinct edges (e_1, \ldots, e_l) of H, such that $u \in e_1$, $v \in e_l$, $e_i \cap e_{i+1} \neq \emptyset$ for all $i \in \{1, 2, \ldots, l-1\}$, and $e_i \cap e_j = \emptyset$ whenever j > i+1 (Note that some authors call this a geodesic path, and use the term path when non-consecutive edges are allowed to intersect.) The distance from u to v in H, denoted dist(u, v), is the shortest length of a u-v path in H. (We define dist(v, v) = 0.) The ball of radius R and centre u in H is the set of vertices of H with distance at most R from u. The diameter of a hypergraph H is defined by

$$diam(H) = \max_{u,v \in V(H)} dist(u,v).$$

A hypergraph is said to be a *cycle* if it has at least two edges, and there is a cyclic ordering of its edges, (e_1, \ldots, e_l) say, such that there exist distinct vertices v_1, \ldots, v_l with $v_i \in e_i \cap e_{i+1}$ for all i (where we define $e_{l+1} := e_1$). This notion of a hypergraph cycle is originally due to Berge, and is sometimes called a *Berge-cycle*. The *length* of a cycle is the number of edges in it. The *girth* of a hypergraph is the length of the shortest cycle it contains.

Observe that two distinct edges e, f with $|e \cap f| \ge 2$ form a cycle of length 2 under this definition, so when considering hypergraphs of high girth, we may restrict our attention to linear hypergraphs.

We use the Landau notation for functions: if $F,G:\mathbb{N}\to\mathbb{R}^+$, we write F=o(G) if $F(n)/G(n)\to 0$ as $n\to\infty$. We write F=O(G) if there exists C>0 such that $F(n)\leqslant CG(n)$ for all n. We write $F=\Omega(G)$ if there exists c>0 such that $F(n)\geqslant cG(n)$ for all n. Finally, we write $F=\Theta(G)$ if F=O(G) and $F=\Omega(G)$.

Extremal questions concerning Berge-cycles in hypergraphs have been studied by several authors. For example, in [7], Bollobás and Győri prove that an n-vertex, 3-uniform hypergraph with no 5-cycle has at most $\sqrt{2}n^{3/2} + \frac{9}{2}n$ edges, and they give a construction showing that this is best possible up to a constant factor. In [18], Lazebnik and Verstraëte prove that a 3-uniform, n-vertex hypergraph of girth at least 5 has at most

$$\frac{1}{6}n\sqrt{n-\frac{3}{4}} + \frac{1}{12}n$$

edges, and give a beautiful construction (based on the so-called 'polarity graph' of the projective plane PG(2,q)) showing that this is sharp whenever $n=q^2$ for an odd prime power $q \ge 27$. Interestingly, neither of these two constructions are regular.

In [14] and [21], Györi and Lemons consider the problem of excluding a cycle of length exactly k, for general $k \in \mathbb{N}$. In [14], they prove that an n-vertex, 3-uniform hypergraph

with no (2k+1)-cycle has at most $4k^2n^{1+1/k} + O(n)$ edges. In [21], they prove that an n-vertex, r-uniform hypergraph with no (2k+1)-cycle has at most $C_{k,r}(n^{1+1/k})$ edges, and furthermore that an n-vertex, r-uniform hypergraph with no (2k)-cycle has at most $C'_{k,r}(n^{1+1/k})$ edges, where $C_{k,r}, C'_{k,r}$ depend upon k and r alone.

In this paper, we will investigate the maximum possible girth of an r-uniform, d-regular hypergraph on n vertices, for r and d fixed and n large. If $r \ge 3$ and $d \ge 2$, we let $g_{r,d}(n)$ denote the maximum possible girth of an r-uniform, d-regular hypergraph on at most n vertices. Similarly, if $d \ge 2$ and $r, g \ge 3$, we let $n_r(g, d)$ denote the minimum possible number of vertices in an r-uniform, d-regular hypergraph with girth at least g. Since a non-linear hypergraph has girth 2, we may replace 'hypergraph' with 'linear hypergraph' in these two definitions.

In section 2, we will state upper and lower bounds on the function $g_{r,d}(n)$, which differ by an absolute constant factor. The upper bound is a simple analogue of the Moore bound for graphs, and follows immediately from known results. The lower bound is a hypergraph extension of a similar argument for graphs, due to Erdős and Sachs [11] not a particularly difficult extension, but still, in our opinion, worth recording.

In section 3, we consider the girth of certain kinds of random r-uniform hypergraph. We define a random r-uniform 'Cayley' hypergraph on S_n which has girth $\Omega(\sqrt{\log |S_n|})$ with high probability, in contrast to random regular r-uniform hypergraphs, which have constant girth with positive probability. We conjecture that, in fact, our 'Cayley' hypergraph has girth $\Omega(\log |S_n|)$ with high probability. We believe it may find other applications.

2 Upper and lower bounds

In this section, we state upper and lower bounds on the function $g_{r,d}(n)$, which differ by an absolute constant factor.

We first state a very simple analogue of the Moore bound for linear hypergraphs. For completeness, we give the proof, although the result follows immediately from known results, e.g. from Theorem 1 of Hoory [16].

Lemma 1. Let r, d and g be integers with $d \ge 2$ and $r, g \ge 3$. Let H be an r-uniform, d-regular, n-vertex hypergraph with girth g. If g = 2k + 1 is odd, then

$$n \ge 1 + d(r-1) \sum_{i=0}^{k-1} ((d-1)(r-1))^i = 1 + d(r-1) \frac{(d-1)^k (r-1)^k - 1}{(d-1)(r-1) - 1}, \tag{5}$$

and if g = 2k is even, then

$$n \geqslant r \sum_{i=0}^{k-1} ((d-1)(r-1))^i = r \frac{(d-1)^k (r-1)^k - 1}{(d-1)(r-1) - 1}.$$
 (6)

Proof. The right-hand side of (5) is the number of vertices in any ball of radius k. The right-hand side of (6) is the number of vertices of distance at most k-1 from any fixed edge $e \in H$.

The following corollary is immediate.

Corollary 2. Let r, d and g be integers with $d \ge 2$ and $r, g \ge 3$. Let H be an r-uniform, d-regular hypergraph with n vertices and girth g. Then

$$g \leqslant \frac{2\log n}{\log(r-1) + \log(d-1)} + 2.$$

Hence,

$$g_{r,d}(n) \leqslant \frac{2\log n}{\log(r-1) + \log(d-1)} + 2.$$

Our aim is now to obtain a hypergraph analogue of the non-constructive lower bound (1). We first prove the following existence lemma.

Lemma 3. For all integers $d \ge 2$ and $r, g \ge 3$, there exists a finite, r-uniform, d-regular hypergraph with girth at least g.

Proof. We prove this by induction on g, for fixed r, d. When g = 3, all we need is a linear, r-uniform, d-regular hypergraph. Let H be the hypergraph on vertex-set \mathbb{Z}_r^d , whose edges are all the axis-parallel lines, i.e.

$$E(H) = \{ \{ \mathbf{x}, \mathbf{x} + \mathbf{e}_i, \mathbf{x} + 2\mathbf{e}_i, \dots, \mathbf{x} + (r-1)\mathbf{e}_i \} : \mathbf{x} \in \mathbb{Z}_r^d, i \in [d] \}.$$

(Here, \mathbf{e}_i denotes the *i*th standard basis vector in \mathbb{Z}_r^d , i.e. the vector with 1 in the *i*th coordinate and zero elsewhere. As usual, \mathbb{Z}_r denotes the ring of integers modulo r.) Clearly, H is linear and d-regular.

For $g \ge 4$ we do the induction step. We start from a finite, linear, r-uniform, d-regular hypergraph H of girth at least g-1. Of all such hypergraphs we consider one with the least possible number of (g-1)-cycles. Let M be the number of (g-1)-cycles in H. We shall prove that M=0. If M>0, we consider a random 2-lift H' of H, defined as follows. Its vertex set is $V(H')=V(H)\times\{0,1\}$, and its edges are defined as follows. For each edge $e\in E(H)$, choose an arbitrary ordering (v_1,\ldots,v_r) of the vertices in e, flip r-1 independent fair coins $c_e^{(1)},\ldots,c_e^{(r-1)}\in\{0,1\}$, and include in H' the two edges

$$\{(v_1, j), (v_2, j \oplus c_e^{(1)}), \dots, (v_r, j \oplus c_e^{(r-1)})\}$$
 for $j = 0, 1$.

(Here, \oplus denotes modulo 2 addition.) Do this independently for each edge. Note that H' is linear and d-regular, since H is.

Let $\pi: V(H') \to V(H)$ be the cover map, defined by $\pi((v, j)) = v$ for all $v \in V(H)$ and $j \in \{0, 1\}$. Since any cycle in H' is projected to a cycle in H of the same length, H' has girth at least g-1, and each (g-1)-cycle in H' projects to a (g-1)-cycle in H. Let C be a (g-1)-cycle in H. We claim that $\pi^{-1}(C)$ either consists of two vertex-disjoint (g-1)-cycles in H', or a single 2(g-1)-cycle in H', and that the probability of each is 1/2. To see this, let (e_1, \ldots, e_{g-1}) be any cyclic ordering of C; then $|e_i \cap e_{i+1}| = 1$ for all i (since H is linear). Let $e_i \cap e_{i+1} = \{w_i\}$ for all $i \in [g-1]$. For each i, consider the two

edges in $\pi^{-1}(e_i)$. Either one of the two edges contains $(w_{i-1}, 0)$ and $(w_i, 0)$ and the other contains $(w_{i-1}, 1)$ and $(w_i, 1)$, or one edge contains $(w_{i-1}, 0)$ and $(w_i, 1)$ and the other edge contains $(w_{i-1}, 1)$ and $(w_i, 0)$. Call these two events $S(e_i)$ and $D(e_i)$, for 'same' and 'different'. Observe that $S(e_i)$ and $D(e_i)$ each occur with probability 1/2, independently for each edge e_i in the cycle. Notice that $\pi^{-1}(C)$ consists of two disjoint (g-1)-cycles if and only if $D(e_i)$ occurs an even number of times, and the probability of this is 1/2, proving the claim.

It follows that the expected number of (g-1)-cycles in H' is M. Note that the trivial lift H_0 of H, which has $c_e^{(k)} = 0$ for all k and e, consists of two vertex-disjoint copies of H, and therefore has 2M (g-1)-cycles. It follows that there is at least one 2-lift of H with fewer than M (g-1)-cycles, contradicting the minimality of M. Therefore, M = 0, so in fact, H has girth at least g. This completes the proof of the induction step, proving the theorem.

Remark. Lemma 3 can also be proved by considering a random r-uniform, d-regular hypergraph on n vertices, for n large. In [8], Cooper, Frieze, Molloy and Reed analyse these using a generalisation of Bollobás' configuration model for d-regular graphs. It follows from Lemma 2 in [8] that if H is chosen uniformly at random from the set of all r-uniform, d-regular, n-vertex, linear hypergraphs (where r|n), then

$$Prob\{girth(H) \ge g\} = (1 + o(1)) \frac{\exp(-\sum_{l=1}^{g-1} \lambda_l)}{1 - \exp(-(\lambda_1 + \lambda_2))},$$
(7)

where

$$\lambda_i = \frac{(r-1)^i (d-1)^i}{2i} \quad (i \in \mathbb{N}),$$

so this event occurs with positive probability for sufficiently large n, giving an alternative proof of Lemma 3. (We note that the argument of [8] can easily be adapted to prove the same statement in the case where $r \mid dn$.)

By itself, the proof of Lemma 3 implies only that

$$n_r(g,d) \leqslant 2^{2^{2}} \cdot 2^{r^{Cd}},$$

where C is an absolute constant — i.e., tower-type dependence upon g. We now proceed to obtain an upper bound which is exponential in g.

Consider a d-regular graph with girth at least g, with the smallest possible number of vertices subject to these conditions. Erdős and Sachs [11] proved that the diameter of such a graph is at most g. But a d-regular graph with diameter D has at most

$$1 + d\sum_{i=0}^{D-1} (d-1)^i$$

vertices (since this is an upper bound on the number of vertices in a ball of radius D). This yielded the upper bound (1) on the number of vertices in a d-regular graph of girth at least g and minimal order.

We need an analogue of the Erdős-Sachs argument for hypergraphs.

Lemma 4. Let r, d and g be integers with $d \ge 2$ and $r, g \ge 3$. Let H be an r-uniform, d-regular hypergraph with girth at least g, with the smallest possible number of vertices subject to these conditions. Then H cannot contain r vertices every two of which are at distance greater than g from one another.

Proof. Let H be an r-uniform, d-regular hypergraph with girth at least g. Suppose that H contains r distinct vertices v_1, v_2, \ldots, v_r such that $\operatorname{dist}(v_i, v_j) > g$ for all $i \neq j$. We will show that it is then possible to construct an r-uniform, d-regular hypergraph with girth at least g, that has fewer vertices than H; this will prove the lemma.

Note that H is linear, since $g \geqslant 3$. For each $i \in [r]$, let $e_i^{(1)}, e_i^{(2)}, \ldots, e_i^{(d)}$ be the edges of H which contain v_i . Let

$$W_i = \bigcup_{k=1}^d (e_i^{(k)} \setminus \{v_i\})$$

for each $i \in [r]$. Notice that $|W_i| = d(r-1)$ for each i, since the edges $e_i^{(k)}$ $(k \in [d])$ are disjoint apart from the vertex v_i . Moreover, $W_i \cap W_j = \emptyset$ for all $i \neq j$, since $d(v_i, v_j) > 2$.

Define a new hypergraph H' by taking H, deleting v_1, v_2, \ldots, v_r and all the edges containing them, and adding d(r-1) pairwise disjoint edges, each of which contains exactly one vertex from W_i for each $i \in [r]$. (Note that none of these 'new' edges were in the original hypergraph H, otherwise some v_i and v_j would have been at distance at most 3 in H, a contradiction.) Clearly, H' is d-regular. We claim that it is linear. Indeed, if one of the 'new' edges shared two vertices with some edge $f \in H$ (say it shares $a \in W_i$ and $b \in W_j$, where $i \neq j$), then there would be a path of length 3 in H from v_i to v_j , a contradiction.

We now claim that H' has girth at least g. Suppose for a contradiction that H' has girth at most g-1. Let C be a cycle in H' of length $l \leq g-1$. Since H' is linear, we have $l \geq 3$. Let (f_1, \ldots, f_l) be a cyclic ordering of C. We split into two cases.

Case 1. Suppose that C contains exactly one of the 'new' edges (say f_i is a 'new' edge). Deleting f_i from C produces a path P of length at most g-2 in H. We have $|f_{i-1} \cap f_i| = |f_i \cap f_{i+1}| = 1$ (since H' is linear); let $f_{i-1} \cap f_i = \{a\}$, and let $f_i \cap f_{i+1} = \{b\}$. Note that $a \neq b$. Suppose that $a \in W_p$ and $b \in W_q$. Since $a \neq b$ and $a, b \in f_i$, we must have $p \neq q$, as each 'new' edge contains exactly one vertex from each W_k . Let e be the edge of H containing both v_p and a, and let e' be the edge of H containing both v_q and b; adding e and e' to the appropriate ends of the path P produces a path in H of length at most g from v_p to v_q , contradicting the assumption that $dist(v_p, v_q) > g$.

Case 2. Suppose instead that C contains more than one of the 'new' edges. Choose a minimal sub-path P of C which connects two 'new' edges. Suppose P connects the new edges f_i and f_j , so that $P = (f_i, f_{i+1}, \ldots, f_{j-1}, f_j)$. Note that $|i - j| \leq (g - 1)/2$, so P has length at most $(g + 1)/2 \leq g - 1$. Let $f_i \cap f_{i+1} = \{a\}$, and suppose $a \in W_p$;

let $f_{j-1} \cap f_j = \{b\}$, and suppose $b \in W_q$. Let e be the unique edge of H which contains both v_p and e, and let e' be the unique edge of e which contains both v_q and e. If e and then we can produce a path in e from e to e by taking e, and replacing e with e and e with e is a path has length at most e and e contradicting our assumption that e is a path in e suppose e and e in e which share the vertex e is a path at length at most e and adding the edges e and e which share the vertex e is this cycle has length at most e and e contradicting our assumption that e has girth at least e.

We may conclude that H' has girth at least g, as claimed. Clearly, H' has fewer vertices than H; this completes the proof.

This lemma quickly implies an upper bound on the minimal number of vertices in an r-uniform, d-regular hypergraph of girth at least g.

Theorem 5. Let r, d and g be integers with $d \ge 2$ and $r, g \ge 3$. There exists an r-uniform, d-regular hypergraph with girth at least g, and at most

$$(r-1)\left(1+d(r-1)\frac{(d-1)^g(r-1)^g-1}{(d-1)(r-1)-1}\right)<4((d-1)(r-1))^{g+1}$$

vertices. Hence,

$$n_r(g,d) < 4((d-1)(r-1))^{g+1}.$$

Proof. Let H be an r-uniform, d-regular hypergraph with girth at least g, with the smallest possible number of vertices subject to these conditions. Let $\{v_1, v_2, \ldots, v_k\}$ be a set of vertices of H whose pairwise distances are all greater than g, with k maximal subject to this condition. By the previous lemma, we have k < r. Any vertex of H must have distance at most g from one of the v_i 's. For each i, the number of vertices of H of distance at most g from v_i is at most

$$1 + d(r-1)\sum_{i=0}^{g-1} ((d-1)(r-1))^i = 1 + d(r-1)\frac{(d-1)^g(r-1)^g - 1}{(d-1)(r-1) - 1},$$

and therefore the number of vertices of H is at most

$$k\left(1+d(r-1)\frac{(d-1)^g(r-1)^g-1}{(d-1)(r-1)-1}\right)\leqslant (r-1)\left(1+d(r-1)\frac{(d-1)^g(r-1)^g-1}{(d-1)(r-1)-1}\right).$$

Crudely, we have

$$(r-1)\left(1+d(r-1)\frac{(d-1)^g(r-1)^g-1}{(d-1)(r-1)-1}\right)<4((d-1)(r-1))^{g+1}$$

for all integers r, d and g with $d \ge 2$ and $r, g \ge 3$, proving the theorem.

The following corollary is immediate.

Corollary 6. Let r, d and n be positive integers with $d \ge 2$ and $r \ge 3$. There exists an r-uniform, d-regular hypergraph on at most n vertices, with girth greater than

$$\frac{\log n - \log 4}{\log(d-1) + \log(r-1)} - 1.$$

Hence,

$$g_{r,d}(n) > \frac{\log n - \log 4}{\log(d-1) + \log(r-1)} - 1.$$

Observe that the lower bound in Corollary 6 differs from the upper bound in Corollary 2 by a factor of (approximately) 2.

For $r, d \ge 3$, we have not been able to improve upon the lower bound in Corollary 6 for large n. As mentioned in the Introduction, in the case of graphs, the bipartite Ramanujan graphs of Lubotzsky, Phillips and Sarnak [22], Margulis [26] and Morgenstern [27] provide d-regular, n-vertex graphs of girth at least

$$(1 - o(1))\frac{4}{3}\frac{\log n}{\log(d-1)},$$

for infinitely many n, whenever d-1 is a prime power. Recall that a finite, connected, d-regular graph is said to be Ramanujan if every eigenvalue λ of its adjacency matrix is either 'trivial' (i.e. $\lambda = \pm d$), or has $|\lambda| \leq 2\sqrt{d-1}$.

Theorem 7 (Lubotzsky-Phillips-Sarnak, Margulis, Morgenstern). For any odd prime power p, there exist infinitely many (bipartite) (p+1)-regular Ramanujan graphs $X^{p,q}$. The graph $X^{p,q}$ is a Cayley graph on the group PGL(2,q), so has order $q(q^2-1)$. Moreover, its girth satisfies

$$g(X^{p,q}) \geqslant \frac{4\log q}{\log p} - \frac{\log 4}{\log p}.$$

It is in place to remark that recently, Marcus, Spielman and Srivastava [24] proved the existence of infinitely many d-regular Ramanujan graphs for $every d \ge 3$. They did this by proving a weakening of a conjecture of Bilu and Linial [6] on 2-lifts of Ramanujan graphs, namely, that every d-regular Ramanujan graph has a 2-lift whose second-largest eigenvalue is at most $2\sqrt{d-1}$. Their proof uses a beautiful new technique for demonstrating the existence of combinatorial objects, which they call the 'method of interlacing polynomials'. (Even more spectacularly, they use this method to prove the Kadison-Singer conjecture, in [25].) Being non-constructive, however, their proof does not imply good bounds for the girth problem.

We are able to improve upon the lower bound in Corollary 6 when r=3 and d=2, using the following explicit construction, based upon the Ramanujan graphs of Theorem 7. Let G be an n-vertex, 3-regular graph of girth g. Take any drawing of G in the plane with straight-line edges, and for each edge $e \in E(G)$, let m(e) be its midpoint. Let H be the 3-uniform hypergraph with

$$V(H) = \{m(e) : e \in E(G)\},\$$

 $E(H) = \{\{m(e_1), m(e_2), m(e_3)\} : e_1, e_2, e_3 \text{ are incident to a common vertex of } G\}.$

Then the hypergraph H is 2-regular, and also has girth g. Taking $G = X^{2,q}$ (the Ramanujan graph of Theorem 7) yields a 3-uniform, 2-regular hypergraph H with

$$g(H) = g(X^{2,q})$$

$$\geqslant \frac{4 \log q}{\log 2} - 2$$

$$\geqslant \frac{4 \log n}{3 \log 2} - 2$$

improving upon the bound in Corollary 6 by a factor of $\frac{4}{3} - o(1)$.

The following explicit construction, also based on the Ramanujan graphs of Theorem 7, provides r-uniform, d-regular hypergraphs of girth approximately 2/3 of the bound in Corollary 6, whenever d is a multiple of r. (We thank an anonymous referee of an earlier version of this paper, for pointing out this construction.)

Suppose d = rs for some $s \in \mathbb{N}$. Let G be a 2(r-1)s-regular, n by n bipartite graph, with vertex-classes X and Y, and girth g. Then the edge-set of G may be partitioned into (r-1)-edge stars in such a way that each vertex of G is in exactly rs of the stars. (Indeed, by Hall's theorem, we may partition the edge-set of G into 2(r-1)s perfect matchings. First, choose r-1 of these matchings, and group the edges of these matchings into n (r-1)-edge stars with centres in X. Now choose r-1 of the remaining matchings, and group their edges into n (r-1)-edge stars with centres in Y. Repeat this process s times to produce the desired partition of E(G) into stars.)

Let H be the r-uniform hypergraph whose vertex-set is $X \cup Y$, and whose edge-set is the collection of vertex-sets of these stars; then H is (rs)-regular, and has girth at least g/2.

If 2(r-1)s-1 is a prime power, the bipartite Ramanujan graph $X^{p,q}$ (with p=2(r-1)s-1) can be used to supply the graph G. This yields a linear, r-uniform, (rs)-regular hypergraph with girth g(H) satisfying

$$\begin{split} g(H) \geqslant \frac{1}{2} \left(\frac{4 \log q}{\log(2rs - 2s - 1)} - \frac{\log 4}{\log(2rs - 2s - 1)} \right) \\ \geqslant \frac{1}{2} \left(\frac{4}{3} \frac{\log n}{\log(2rs - 2s - 1)} - \frac{\log 4}{\log(2rs - 2s - 1)} \right) \\ = \frac{2}{3} \frac{\log n}{\log(2d - 2d/r - 1)} - \frac{\log 2}{\log(2d - 2d/r - 1)}, \end{split}$$

where d = rs.

Unfortunately, this lower bound is asymptotically worse than that given by Corollary 6, for all values of r and d.

3 Random 'Cayley' hypergraphs

In this section, we give a construction of random 'Cayley' hypergraphs on the symmetric group S_n , which have girth $\Omega(\sqrt{\log |S_n|})$ with high probability. This is much higher than

the girth of a random regular hypergraph on the same number of vertices (which, by (7), has girth at most $C(\epsilon)$ with probability at least $1 - \epsilon$ for any $\epsilon > 0$, where $C(\epsilon)$ is a constant depending on ϵ alone), though it is still short of the optimal $\Theta(\log |V(H)|)$ in Corollary 6. The situation is analogous to the graph case, where random d-regular Cayley graphs on appropriate groups have much higher girth than random d-regular graphs of the same order (due to the dependency between cycles at different vertices of a Cayley graph).

First, we need some more definitions. If S is a set of symbols, a word in S is a string of the form

$$s_1^{a_1} s_2^{a_2} \dots s_l^{a_l}$$

where $s_1, \ldots, s_l \in S$ and $a_1, \ldots, a_l \in \mathbb{Z} \setminus \{0\}$. Such a word is said to be *cyclically irreducible* if $s_i \neq s_{i+1}$ for all $i \in [l]$, where we define $s_{l+1} := s_1$. Its *length* is $\sum_{i=1}^l |a_i|$.

Theorem 8. Let r and n be positive integers with $r \ge 3$ and r|n. Let X(n,r) be the set of permutations in S_n that consist of $\frac{n}{r}$ disjoint r-cycles. Choose d permutations $\tau_1, \tau_2, \ldots, \tau_d$ uniformly at random and independently (with replacement) from X(n,r), and let H be the random hypergraph with vertex-set S_n and edge-set

$$\{\{\sigma, \sigma\tau_i, \sigma\tau_i^2, \dots, \sigma\tau_i^{r-1}\}: \sigma \in S_n, i \in [d]\}.$$

Then with high probability, H is a linear, r-uniform, d-regular hypergraph with girth at least

$$c_0 \sqrt{\frac{n \log n}{r(r-1)(\log(d-1) + \log(r-1))}},$$

for any absolute constant c_0 such that $0 < c_0 < 1/2$.

Remark. Here, 'with high probability' means 'with probability tending to 1 as $n \to \infty$ '.

Proof. Note that the edges of the form

$$\{\sigma, \sigma\tau_i, \sigma\tau_i^2, \dots, \sigma\tau_i^{r-1}\}\ (\sigma \in S_n)$$

are simply the left cosets of the cyclic group $\{\mathrm{Id}, \tau_i, \tau_i^2, \ldots, \tau_i^{r-1}\}$ in S_n , so they form a partition of S_n . We need two straightforward claims.

Claim 1. With high probability, the following condition holds.

$$\tau_1, \dots, \tau_d \text{ satisfy } \tau_i^k \neq \tau_j^l \text{ for all distinct } i, j \in [d] \text{ and all } k, l \in [r-1].$$
 (8)

Proof of claim: Let us fix $i, j \in [d]$ with i < j, and fix $k, l \in [r-1]$. We shall bound the probability that $\tau_j^l = \tau_i^k$. We regard τ_i as fixed, and allow τ_j to vary. Since τ_i is a product of n/r disjoint r-cycles, τ_i^k is a product of n/s disjoint s-cycles, for some integer $s \ge 2$ that is a divisor of r. The set X(n, s) of permutations which consist of n/s disjoint s-cycles has cardinality

$$\frac{n!}{(n/s)!s^{n/s}}\geqslant \frac{n!}{(n/2)!2^{n/2}}$$

(provided $n \ge 4$). Notice that τ_j^l is uniformly distributed over X(n, s'), for some s' that depends only on r and l. Therefore,

$$\text{Prob}\{\tau_i^k = \tau_j^l\} \leqslant \frac{(n/2)!2^{n/2}}{n!}.$$

By the union bound,

$$\operatorname{Prob}\{\tau_i^k = \tau_j^l \text{ for some } i \neq j \text{ and some } k, l \in [r-1]\} \leqslant (r-1)^2 \binom{d}{2} \frac{(n/2)! 2^{n/2}}{n!} \to 0 \quad \text{as } n \to \infty,$$

proving the claim.

Claim 2. If condition (8) holds, then for all $i \neq j$ and all $\sigma, \pi \in S_n$, the two cosets

$$\{\sigma, \sigma\tau_i, \sigma\tau_i^2, \dots, \sigma\tau_i^{r-1}\}$$
 and $\{\pi, \pi\tau_j, \pi\tau_j^2, \dots, \pi\tau_j^{r-1}\}$

have at most one element in common.

Proof of claim: Suppose for a contradiction that there are two distinct vertices v_1, v_2 with

$$v_1, v_2 \in \{\sigma, \sigma\tau_i, \sigma\tau_i^2, \dots, \sigma\tau_i^{r-1}\} \cap \{\pi, \pi\tau_j, \pi\tau_j^2, \dots, \pi\tau_j^{r-1}\}.$$

Then $v_1 = \sigma \tau_i^l = \pi \tau_j^m$ and $v_2 = \sigma \tau_i^{l'} = \pi \tau_j^{m'}$, where $l, m, l', m' \in \{0, 1, \dots, r-1\}$ with $l' \neq l$ and $m' \neq m$. Therefore,

$$v_1^{-1}v_2 = \tau_i^{l'-l} = \tau_i^{m'-m},$$

contradicting condition (8).

Claim 2 implies that H is a linear hypergraph, provided condition (8) is satisfied. Moreover, H is d-regular: every $\sigma \in S_n$ is contained in the edges (cosets)

$$(\{\sigma, \sigma\tau_i, \sigma\tau_i^2, \dots, \sigma\tau_i^{r-1}\} : i \in [d]),$$

and these d edges are distinct provided condition (8) is satisfied.

Finally, we make the following.

Claim 3. With high probability, H has girth at least

$$c_0 \sqrt{\frac{n \log n}{r(r-1)(\log(d-1) + \log(r-1))}},$$

where c_0 is any absolute constant such that $0 < c_0 < 1/2$.

Proof of claim: We may assume that condition (8) holds, so that H is a linear, d-regular hypergraph. Let C be a cycle in H of minimum length, and let (e_1, \ldots, e_l) be any cyclic ordering of its edges. Then we have $|e_i \cap e_{i+1}| = 1$ for all $i \in [l]$ (where we define $e_{l+1} := e_1$), and by minimality, we have $e_i \cap e_j = \emptyset$ whenever |i - j| > 1. Let $e_i \cap e_{i+1} = \{w_i\}$ for each $i \in [l]$. Suppose that e_i is an edge of the form

$$\{\sigma, \sigma\tau_{j_i}, \sigma\tau_{j_i}^2, \dots, \sigma\tau_{j_i}^{r-1}\}$$

for each $i \in [l]$. Since $e_i \cap e_{i+1} \neq \emptyset$ for each $i \in [l]$, we must have $j_i \neq j_{i+1}$ for all $i \in [l]$ (where we define $j_{l+1} := j_1$). For each $i \in [l]$, we have $w_i, w_{i+1} \in e_{i+1}$, so $w_i^{-1}w_{i+1} = \tau_{j_{i+1}}^{m_i}$ for some $m_i \in [r-1]$. Therefore,

$$\operatorname{Id} = (w_1^{-1} w_2)(w_2^{-1} w_3) \dots (w_{l-1}^{-1} w_l)(w_l^{-1} w_1) = \tau_{j_2}^{m_1} \tau_{j_3}^{m_2} \dots \tau_{j_l}^{m_{l-1}} \tau_{j_1}^{m_l}. \tag{9}$$

Since $j_i \neq j_{i+1}$ for all $i \in [l]$, the word on the right-hand side of (9) is cyclically irreducible. We therefore have a cyclically irreducible word in the symbols $\{\tau_j : j \in [d]\}$ with length $L := \sum_{j=1}^l m_i \leqslant (r-1)l$, which evaluates to the identity permutation. We must show that the probability of this tends to zero as $n \to \infty$, for an appropriate choice of l. We use an argument similar to that of [12], where it is proved that a random d-regular Cayley graph on S_n has girth at least $\Omega(\sqrt{\log_{d-1}(n!)})$.

Let W be a cyclically irreducible word in the τ_j 's, with length L. We must bound the probability that W fixes every element of [n]. Suppose

$$W = \tau_{j(1)}\tau_{j(2)}\dots\tau_{j(L)}.$$

Let $x_0 \in [n]$, and define $x_i = \tau_{j(i)}(x_{i-1})$ for each $i \in [L]$, producing a sequence of values $x_0, x_1, x_2, \ldots, x_L \in [n]$; then $W(x_0) = x_L$. We shall bound the probability that $x_L = x_0$. Let us work our way along the sequence, exposing the r-cycles of the permutations τ_1, \ldots, τ_d only as we need them, so that at stage i, the r-cycle of $\tau_{j(i)}$ containing the number x_{i-1} is exposed (if it has not already been exposed). If $x_L = x_0$, then (as $j(L) \neq j(1)$), there has to be a first time the sequence returns to x_0 via a permutation $\tau \neq \tau_{j(1)}$. Hence, at some stage, we must have exposed an r-cycle of τ containing x_0 . The probability that, at a stage i where $j(i) \neq j(1)$, we expose an r-cycle of $\tau_{j(i)}$ containing x_0 , is at most

$$\frac{r}{n-(i-2)r}\leqslant \frac{r}{n-(L-2)r},$$

since a total of at most i-2 r-cycles of τ have already been exposed, and the next r-cycle exposed is equally likely to be any r-element subset of the remaining n-(i-2)r numbers. There are at most L choices for the stage i, and therefore

$$Prob\{W(x_0) = x_0\} \le L \frac{r}{n - (L - 2)r}.$$

Suppose we have already verified that W fixes $y_1, y_2, \ldots, y_{m-1}$, by exposing the necessary r-cycles. Then we have exposed at most (m-1)L r-cycles. As long as (m-1)Lr < n,

we can choose a number $y_m \in [n]$ such that none of the previously exposed r-cycles contains y_m . Repeating the above argument yields an upper bound of

$$\frac{Lr}{n-mLr}$$

on the probability that W fixes y_m , even when conditioning on the (m-1)L previously exposed r-cycles. Therefore,

$$\operatorname{Prob}\{W = \operatorname{Id}\} \leqslant \left(\frac{Lr}{n - mLr}\right)^m,$$

as long as mLr < n. Substituting $m = \lceil n/(2Lr) \rceil$ yields the bound

$$\operatorname{Prob}\{W = \operatorname{Id}\} \leqslant \left(\frac{2Lr}{n}\right)^{n/(2Lr)}.$$

The number of choices for the word on the right-hand side of (9) is at most $(d-1)^l(r-1)^l$. (By taking a cyclic shift if necessary, we may assume that $j_2 \neq d$, so there are at most d-1 choices for j_2 , and at most d-1 choices for all subsequent j_i ; there are clearly at most r-1 choices for each m_i .) Hence, the probability that there exists such a word which evaluates to the identity permutation is at most

$$(d-1)^{l}(r-1)^{l}\left(\frac{2r(r-1)l}{n}\right)^{n/(2r(r-1)l)}$$
.

To bound the probability that H has a cycle of length less than g, we need only sum the above expression over all l < g:

$$Prob\{girth(H) < g\} \leqslant \sum_{l=3}^{g-1} (d-1)^l (r-1)^l \left(\frac{2r(r-1)l}{n}\right)^{n/(2r(r-1)l)}$$
$$< (d-1)^g (r-1)^g \left(\frac{2r(r-1)g}{n}\right)^{n/(2r(r-1)g)}.$$

In order for the right-hand side to tend to zero as $n \to \infty$, we must choose

$$g = c_0 \sqrt{\frac{n \log n}{r(r-1)(\log(d-1) + \log(r-1))}}$$

for some constant $c_0 < 1/2$; we then have

$$\operatorname{Prob}\{\operatorname{girth}(H) < g\} \leqslant \exp\left(-\Omega\left(\frac{1}{r}\sqrt{(\log(d-1) + \log(r-1))(n\log n)}\right)\right).$$

This completes the proof of Claim 3, and thus proves Theorem 8.

__ 14

4 Conclusion and open problems

Our best (general) upper and lower bounds on the function $g_{r,d}(n)$ differ approximately by a factor of 2:

$$(1+o(1))\frac{\log n}{\log(d-1)+\log(r-1)} \leqslant g_{r,d}(n) \leqslant (2+o(1))\frac{\log n}{\log(r-1)+\log(d-1)}.$$

It would be of interest to narrow the gap, possibly by means of an explicit algebraic construction \grave{a} la Ramanujan graphs.

In [12], Gamburd, Hoory, Shahshahani, Shalev and Virág conjecture that with high probability, a random d-regular Cayley graph on S_n has girth at least $\Omega(\log |S_n|)$, as opposed to the $\Omega(\sqrt{\log |S_n|})$ which they prove. We believe that the random hypergraph of Theorem 8 also has girth $\Omega(\log |S_n|)$.

In this paper, we considered a very simple and purely combinatorial notion of girth in hypergraphs, but other notions appear in the literature, for example using the language of simplicial topology, such as in [23, 13]. A different combinatorial definition was introduced by Erdős in [10]. Define the (-2)-girth of a 3-uniform hypergraph as the smallest integer $g \ge 4$ such that there is a set of g vertices spanning at least g-2 edges. Erdős conjectured in [10] that there exist Steiner Triple Systems with arbitrarily high (-2)-girth; this question remains wide open (see for example [2]), and seems very hard. In view of this, we raise the following.

Question 9. Is there a constant c > 0 such that there exist n-vertex 3-uniform hypergraphs with cn^2 edges and arbitrarily high (-2)-girth?

Note that Erdős' conjecture on Steiner Triple Systems, if true, would imply a positive answer for every $c < \frac{1}{6}$. This is clearly tight, since an *n*-vertex, 3-uniform hypergraph with at least $n^2/6$ edges cannot be linear, 1 and therefore has (-2)-girth 4.

We turn briefly to some variants of Erdős' definition. The celebrated (6,3)-theorem of Ruzsa and Szemerédi [28] states that if H is an n-vertex, 3-uniform hypergraph in which no 6 vertices span 3 or more edges, then H has $o(n^2)$ edges. Therefore, if we define the (-3)-girth of a 3-uniform hypergraph to be the smallest integer $g \ge 6$ such that there exists a set of g vertices spanning at least g-3 edges, then an n-vertex, 3-uniform hypergraph with (-3)-girth at least 7 has $o(n^2)$ edges. Hence, the analogue of Question 9 for (-3)-girth has a negative answer. On the other hand, if we define the (-1)-girth of a 3-uniform hypergraph to be the smallest integer g such that there exists a set of g vertices spanning at least g-1 edges, it can be shown that the maximum number of edges in an n-vertex, 3-uniform hypergraph with (-1)-girth at least g, is $n^{2+\Theta(1/g)}$.

¹If H is a linear, n-vertex, 3-uniform hypergraph, then any pair of vertices is contained in at most one edge of H, so double-counting the number of times a pair of vertices in contained in an edge of H, we obtain $3e(H) \leq \binom{n}{2}$.

²The condition $g \ge 6$ is necessary to avoid triviality: if we replaced it with $g \ge 5$, then a 3-uniform hypergraph would have (-3)-girth 5 unless it consisted of isolated edges.

Acknowledgment

N. L. wishes to thank Shlomo Hoory for many years of joint pursuit of the girth problem [15]. Quite a few of the ideas of this present paper can be traced back to that joint research effort.

References

- [1] E. Bannai, T. Ito, Regular graphs with excess one, *Discrete Mathematics* 37 (1981), 147–158.
- [2] R. A. Beezer, The girth of a design, Journal of Combinatorial Mathematics and Combinatorial Computing 40 (2002), 97–113.
- [3] C. Berge, Hypergraphs, Combinatorics of Finite Sets. North-Holland, Amsterdam, 1989.
- [4] C. Berge, Hypergraphs, in Selected Topics in Graph Theory 3, eds. L. W. Beineke and R. J. Wilson, Academic Press, 1998, 189–207.
- [5] N. Biggs, T. Ito, Graphs with even girth and small excess, Mathematical Proceedings of the Cambridge Philosophical Society 88 (1980), 1–10.
- [6] Y. Bilu and N. Linial, Lifts, discrepancy and nearly optimal spectral gap, *Combinatorica* 26 (2006), 495–519.
- [7] B. Bollobás, E. Győri, Pentagons versus Triangles, *Discrete Mathematics* 308 (2008) 4332–4336.
- [8] C. Cooper, A. Frieze, M. Molloy, B. Reed, Perfect matchings in random r-regular s-uniform hypergraphs, Combinatorics, Probability and Computing 5 (1996), 1–14.
- [9] R. Duke, Types of cycles in hypergraphs. *Cycles in graphs*, 399–417, North-Holland Mathematical Studies, 115, North-Holland, Amsterdam, 1985.
- [10] P. Erdős, Problems and results in combinatorial analysis, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Atti dei Convegni Lincei, No. 17, pp. 3–17, Accad. Naz. Lincei, Rome, 1976.
- [11] P. Erdős, H. Sachs, Reguläre Graphen gegebener Taillenweite mit minimaler Knotenzahl. (Regular graphs with given girth and minimal number of knots.) Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg, Math.-Naturwiss. 12 (1963), 251–258.
- [12] A. Gamburd, S. Hoory, M. Shahshahani, A. Shalev, B. Virág, On the girth of random Cayley graphs, *Random Strutures and Algorithms* 35 (2009), 100–117.
- [13] M. Goff, Higher dimensional Moore bounds, *Graphs and Combinatorics* 27 (2011), 505–530.
- [14] E. Györi, N. Lemons, 3-uniform hypergraphs avoiding a given odd cycle. Preprint, available at http://www.renyi.hu/~nathan/.
- [15] S. Hoory, On graphs of high girth, PhD thesis, Hebrew University of Jerusalem, 2002.

- [16] S. Hoory, The size of bipartite graphs with a given girth, *Journal of Combinatorial Theory*, Series B 86 (2002), 215–220.
- [17] P. Kovács, The non-existence of certain regular graphs of girth 5, Journal of Combinatorial Theory, Series B, 30 (1981), 282–284.
- [18] F. Lazebnik, J. Verstraëte, On hypergraphs of girth five, *Electronic Journal of Combinatorics* 10 (2003), #R25.
- [19] F. Lazebnik, V. A. Ustimenko, A. J. Woldar, A New Series of Dense Graphs of High Girth, Bulletin of the American Mathematical Society 32 (1995), 73–79.
- [20] F. Lazebnik, V. A. Ustimenko, A. J. Woldar, New upper bounds on the order of cages, *Electronic Journal of Combinatorics* 14 (1997), #R13.
- [21] N. Lemons, Turán problems for hypergraphs. PhD dissertation, Central European University, 2008.

 Available at http://www.renyi.hu/~nathan/nathan-thesis.pdf.
- [22] A. Lubotzsky, R. Phillips, P. Sarnak, Ramanujan graphs, *Combinatorica* 8 (1988), 261–277.
- [23] A. Lubotzky, R. Meshulam, A Moore bound for simplicial complexes, *Bulletin of the London Mathematical Society* 39 (2007), 353–358.
- [24] A. Marcus, D. A. Spielman and N. Srivastava, Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees. To appear in *FOCS* 2013.
- [25] A. Marcus, D. A. Spielman and N. Srivastava, Interlacing Families II: Mixed Characteristic Polynomials and the Kadison-Singer Problem. Preprint available at http://arxiv.org/abs/1306.3969.
- [26] G. A. Margulis, Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators, *Problems of Information Transmission* 24 (1988), 39–46.
- [27] M. Morgenstern, Existence and explicit constructions of (q + 1)-regular Ramanujan graphs for every prime power q, Journal of Combinatorial Theory, Series B, 62 (1994), 44–62.
- [28] I. Z. Ruzsa, E. Szemerédi, Triple systems with no six points carrying three triangles, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, pp. 939– 945; also appears in Colloq. Math. Soc. János Bolyai 18, North-Holland, Amsterdam / New York, 1978.