

Economical elimination of cycles in the torus

Noga Alon*

Abstract

Let $m > 2$ be an integer, let C_{2m} denote the cycle of length $2m$ on the set of vertices $[-m, m) = \{-m, -m+1, \dots, m-2, m-1\}$, and let $G = G(m, d)$ denote the graph on the set of vertices $[-m, m)^d$, in which two vertices are adjacent iff they are adjacent in one coordinate of C_{2m} and equal in all others. This graph can be viewed as the graph of the d -dimensional torus. We prove that one can delete a fraction of at most $O(\frac{\log d}{m})$ of the vertices of G so that no topologically nontrivial cycles remain. This is tight up to the $\log d$ factor and improves earlier estimates by various researchers.

1 Introduction

Let $G = G(m, d)$ denote the graph on the set of vertices $[-m, m)^d$ in which two vertices are adjacent iff they are equal in all coordinates but one, in which they are adjacent in the cycle C_{2m} . This graph can be viewed as the graph of the d -dimensional torus. A cycle in it is called *nontrivial* if it wraps around the torus. In particular, the projection of each such cycle along at least one of the coordinates contains all vertices of C_{2m} . A (*vertex*) *spine* is a set of vertices that intersects every nontrivial cycle. It is easy to see that there is a spine containing a fraction of $O(d/m)$ of the vertices. This has been improved in [4], where it is shown that there is a vertex spine of size at most a fraction $O(\frac{d^{\log_2(3/2)}}{m}) \approx O(\frac{d^{0.6}}{m})$ of the vertices, and further improved to $O(\frac{\sqrt{d}}{m})$ in [2]. Here we show that the \sqrt{d} -term can be reduced significantly.

Theorem 1 *There is an absolute constant c so that for every m and d the graph $G(m, d)$ contains a vertex spine consisting of at most a fraction of $c\frac{\log d}{m}$ of the number of vertices.*

Note that a trivial lower bound for the size of any vertex spine in $G(m, d)$ is a $\frac{1}{2m}$ -fraction of the number of vertices, as, for example, for each fixed $(x_2, x_3, \dots, x_d) \in [-m, m)^{d-1}$, the spine must contain at least one vertex in the cycle $\{(x, x_2, \dots, x_d) : x \in [-m, m)\}$. Therefore, the best known earlier bound supplies an upper estimate exceeding the trivial lower bound by a factor of $O(\sqrt{d})$, while the new bound improves this factor to $O(\log d)$.

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel and IAS, Princeton, NJ 08540, USA. Email: nogaa@tau.ac.il. Research supported in part by the Israel Science Foundation, by a USA-Israel BSF grant, by NSF grant CCF 0832797 and by the Ambrose Monell Foundation.

The main part of the proof of Theorem 1 is the construction of an induced subgraph of G that contains no nontrivial cycles, and has a small vertex boundary. While in [2] the existence of such a subgraph is derived from a version of the vertex isoperimetric inequality of [1], here it is obtained by an explicit construction, which supplies a much better upper bound for the relative size of the boundary. The required spine is then obtained, as in [2] or [5], by taking pieces of boundaries of random shifts of this body.

The rest of this short paper is organized as follows. In Section 2 we state and prove the main lemma: the existence of an induced subgraph of G with no nontrivial cycles and a small vertex boundary. Section 3 contains the short derivation of Theorem 1 from this lemma. In Section 4 we describe a vertex isoperimetric inequality for graphs, with Dirichlet boundary condition, that may be helpful in tackling similar problems. The final section, Section 5, contains a few brief concluding remarks.

2 The main lemma

For a set of vertices X in a graph G , let $N(X)$ be the set of all neighbors of vertices in X . The *vertex boundary* $N'(X)$ of X is defined by $N'(X) = N(X) - X$. The main part of the proof is the following.

Lemma 2 *There exists an absolute constant c so that for all m and d there is a set of vertices B of $G = G(m, d)$ that contains no nontrivial cycles so that its vertex boundary in G is of size $|N'(B)| = h|B|$, where $\frac{h}{1+h} = \frac{|N'(B)|}{|B|+|N'(B)|} \leq c \frac{\log d}{m}$.*

Proof: The proof is by an explicit construction, using the intersection of an appropriate discrete ℓ_1 -ball in Z^d with the set $(-m, m)^d$. Throughout the proof, we make no attempt to optimize the absolute constants, and prove the theorem with $c = 30$, although this can certainly be improved. To simplify the presentation, we omit all floor and ceiling signs, whenever these are not crucial. All logarithms are in the natural base.

As any choice of B gives $\frac{h}{1+h} \leq 1$, and since we prove the theorem with $c = 30$, there is nothing to prove for $m < 30 \log d$ and we thus assume that $m \geq 30 \log d$. In addition, since the set $W = (-m, m)^d$ satisfies $W \cup N'(W) = [-m, m]^d$, and its vertex boundary is of size at most $d(2m)^{d-1}$, we have $\frac{|N'(W)|}{|W|+|N'(W)|} \leq \frac{d}{2m} < 30 \frac{\log d}{m}$ provided $d < 60 \log d$, and we thus may and will assume that $\frac{d}{\log d} \geq 60$.

For any integer y define:

$$A(y, d) = \{(x_1, x_2, \dots, x_d) \in Z^d : \sum_{i=1}^d |x_i| \leq y\}.$$

Put $x = \frac{md}{5 \log d}$ and define $A = A(x, d)$. Define also

$$C = C(x, d) = \{(x_1, x_2, \dots, x_d) \in A(x, d) : \exists i, |x_i| \geq m\},$$

and $B = A - C$.

It is easy to check that

$$|A(x, d)| = \sum_{i=0}^d \binom{d}{i} 2^{d-i} \binom{x}{d-i}. \quad (1)$$

Indeed, if the number of zeros in the vector (x_1, x_2, \dots, x_d) is i , then there are $\binom{d}{i}$ possibilities to choose their location, and 2^{d-i} possibilities to choose the signs of all other coordinates. The number of ways to choose the absolute values of all these other $d-i$ coordinates so that their sum is at most x and all of them are positive is then $\binom{x}{d-i}$. Thus (1) holds.

Note, next, that $|C| \leq 2d|A(x-m, d)|$. To see this, observe that the number of members (x_1, x_2, \dots, x_d) of A in which a specific coordinate, say x_1 , satisfies $|x_1| \geq m$, is at most $2|A(x-m, d)|$ as there are two ways to choose the sign of x_1 , and then at most $|A(x-m, d)|$ ways to select the vector $(|x_1| - m, x_2, \dots, x_d)$.

We next show that $|C|$ is much smaller than $|A|$, that is, most vectors in A actually lie in $B = A - C$. To do so we bound the quantity $\frac{|A(x-m, d)|}{|A(x, d)|}$. Note, first, that for each $i \leq d/2$ we have:

$$\frac{\binom{d}{i} 2^{d-i} \binom{x-m}{d-i}}{\binom{d}{i} 2^{d-i} \binom{x}{d-i}} \leq \left(\frac{x-m}{x}\right)^{d/2} = \left(1 - \frac{5 \log d}{d}\right)^{d/2} \leq \frac{1}{d^{5/2}}. \quad (2)$$

In addition, for each $i \geq d/2$,

$$\frac{\binom{d}{i} 2^{d-i} \binom{x-m}{d-i}}{\binom{d}{0} 2^d \binom{x-m}{d}} \leq 2^{d/2} \left(\frac{d}{x-m-d}\right)^i \leq 2^{-d/2} 4^{-(i-d/2)}, \quad (3)$$

where here we used the fact that by our assumptions $x \geq 12m$ and $x \geq 6d$, implying that $x - m - d \geq \frac{3}{4}x > 4d$.

By (3), and since $d > 60$, the sum of all terms in the formula for $|A(x-m, d)|$ with $i > d/2$ is (much) smaller than, say, 0.1 times the first term, and thus, by (2)

$$|A(x-m, d)| \leq 1.1 \sum_{i=0}^{d/2} \binom{d}{i} 2^{d-i} \binom{x-m}{d-i} \leq \frac{1.1}{d^{5/2}} \sum_{i=0}^{d/2} \binom{d}{i} 2^{d-i} \binom{x}{d-i} < \frac{1}{d^2} |A(x, d)|.$$

As $B = A - C$, $d \geq 60$ and $|C| \leq 2d|A| \leq \frac{2}{d}|A|$, we conclude that

$$|B| > 0.9|A|, \quad (4)$$

(with room to spare.) Note that B is a subset of $(-m, m)^d$, and thus, considered as a set of vertices of the torus $G(m, d)$, it contains no nontrivial cycles, as its projection on each coordinate misses a point of $[-m, m]$. We now estimate the size of the vertex boundary $N^1(B)$ of B . This boundary is contained in the union of two sets, that we denote by S_1 and S_2 , where

$$S_1 = \{(x_1, x_2, \dots, x_d) \in Z^d : \sum_{i=1}^d |x_i| = x + 1\},$$

and

$$S_2 = \{(x_1, x_2, \dots, x_d) \in Z^d : \text{there is a unique } i \text{ with } |x_i| = m \text{ and } \sum_{i=1}^d |x_i| \leq x + 1\}.$$

As in the argument for computing (1),

$$|S_1| = \sum_{i=0}^d \binom{d}{i} 2^{d-i} \binom{x}{d-i-1}. \quad (5)$$

To estimate $\frac{|S_1|}{|A|}$ note that for every i ,

$$\frac{\binom{d}{i} 2^{d-i} \binom{x}{d-i-1}}{\binom{d}{i} 2^{d-i} \binom{x}{d-i}} = \frac{d-i}{x-d+i+1} \leq \frac{d}{x-d} \leq \frac{d}{5x/6} = \frac{6d}{5x} = \frac{6 \log d}{m}.$$

This implies that $\frac{|S_1|}{|A|} \leq \frac{6 \log d}{m}$, and hence, as $|B| > 0.9|A|$ by (4), we conclude that

$$|S_1| \leq \frac{7 \log d}{m} |B|. \quad (6)$$

It remains to bound the cardinality of S_2 . Clearly $|S_2| \leq d|A(x+1-m, d-1)|$, since there are d choices for the unique coordinate of $(x_1, x_2, \dots, x_d) \in S_2$ which is either m or $-m$, in the torus $G(m, d)$, m and $-m$ are identical, and then there are $|A(x+1-m, d-1)|$ ways to choose the rest of the coordinates of (x_1, \dots, x_d) . Clearly

$$|A(x+1-m, d-1)| = \sum_{i=0}^{d-1} \binom{d-1}{i} 2^{d-i-1} \binom{x+1-m}{d-i-1}.$$

We can thus bound $\frac{|A(x+1-m, d-1)|}{|A(x, d)|}$ as before. For each $i < d/2$,

$$\frac{\binom{d-1}{i} 2^{d-i-1} \binom{x+1-m}{d-i-1}}{\binom{d}{i} 2^{d-i} \binom{x}{d-i}} \leq \frac{1}{2} \frac{d-i}{x} \left(\frac{x+1-m}{x-1}\right)^{d-1-i} \leq \frac{1}{2} \frac{d}{x} \left(1 - \frac{m-1}{x-1}\right)^{(d-1)/2} < \frac{3 \log d}{m} \frac{1}{d^2}. \quad (7)$$

In addition, for each $i \geq d/2$,

$$\frac{\binom{d-1}{i} 2^{d-i-1} \binom{x+1-m}{d-i-1}}{\binom{d-1}{0} 2^{d-1} \binom{x+1-m}{d-1}} \leq 2^{d/2} \left(\frac{d}{x+1-m-d}\right)^i \leq 2^{-d/2} 4^{-(i-d/2)}. \quad (8)$$

By (8), and since $d > 60$, the sum of all terms in the formula for $|A(x-m+1, d-1)|$ with $i \geq d/2$ is (much) smaller than, say, 0.1 times the first term, and thus, by (7)

$$|A(x+1-m, d)| \leq 1.1 \sum_{i < d/2} \binom{d-1}{i} 2^{d-i-1} \binom{x+1-m}{d-i-1} < 1.1 \frac{3 \log d}{m} \frac{1}{d^2} |A(x, d)|.$$

Since $|S_2| \leq d|A(x+1-m, d-1)|$ and $|B| > 0.9|A(x, d)|$ we conclude that $|S_2| \leq \frac{4 \log d}{dm} |B|$, which, together with (6), completes the proof of the lemma. \blacksquare

3 The proof of the theorem

Proof of Theorem 1: Let v_1, v_2, \dots be a sequence of random, uniform, independent points in $[-m, m]^d$. For each $i \geq 1$, put $B_i = B + v_i$, where B is as in Lemma 2, and addition is in the torus. Note that the induced subgraph of $G = G(m, d)$ on each set B_i is isomorphic to the induced subgraph on B , and hence contains no nontrivial cycles. It is easy to see that with probability 1 there is a finite s so that $\cup_{i \leq s} (B_i \cup N'(B_i)) = [-m, m]^d$.

Put $S_1 = N'(B_1)$, and for each i , $1 \leq i \leq s$, let $S_i = N'(B_i) - \cup_{j < i} (B_j \cup N'(B_j))$. Finally, define $S = \cup_{i=1}^s S_i$. We claim that S is a spine. Indeed, if Q is a nontrivial cycle in G , then let i be the smallest integer so that there is a vertex of Q in $B_i \cup N'(B_i)$. As B_i contains no nontrivial cycles, Q must contain vertices outside B_i , and hence also vertices in $N'(B_i)$. By the minimality in the choice of i any such vertex must lie in $S_i \subset S$, as it cannot belong to any union $B_j \cup N'(B_j)$ for $j < i$. Thus S is indeed a (random) spine. We next estimate the expected size of S . Fix a vertex v of G , and let i be the smallest integer so that $v \in B_i \cup N'(B_i)$. Therefore, there is a vertex $u \in B \cup N'(B)$ so that (in the torus) $v = v_i + u$. The crucial observation is that u is a uniform random vertex of $B \cup N'(B)$, as for each point $u \in \cup N'(B)$, there is a unique choice of v_i so that $v = v_i + u$. Therefore, the probability that v lies in $N'(B_i)$ is precisely $\frac{|N'(B)|}{|B| + |N'(B)|}$, which is at most $c \frac{\log d}{m}$, by the choice of B from Lemma 2. We have therefore shown that any fixed vertex v of G lies in S with probability at most $c \frac{\log d}{m}$. Linearity of expectation thus shows that the expected number of vertices in the spine S is at most $(2m)^d \cdot \frac{\log d}{m}$, implying that there exists a spine of size at most that quantity, and completing the proof. \blacksquare

4 A vertex isoperimetric inequality

The existence of a set of vertices $B \subset [-m, m]^d$ with a small vertex boundary is proved in the previous section by a direct construction. In [2] the existence of such a set (with a weaker estimate for the relative size of its boundary) is derived from an appropriate vertex isoperimetric inequality (with Dirichlet boundary condition). Although the direct approach appears to supply better results here, the approach of [2] seems more powerful in general, and indeed it yields better results in the study of edge spines of a related graph, discussed in [2] (as well as in [6]). We therefore present here another vertex isoperimetric inequality, with Dirichlet boundary condition. Although we have not been able to apply it to derive any estimates that improve the statement of Lemma 2, we believe that it may be useful in the study of similar problems for other Cayley graphs. Our approach here follows closely the one of Bobkov, Houdré and Tetali in [3], which in turn borrows some ideas from [1]. The main difference between the inequality given here and that of [3] is that here we are interested in the inequality for graphs with boundary.

For a graph $G = (V, E)$ and a vertex $i \in V$, let $N(i)$ denote the set of all neighbors of i in G , and let $N^+(i) = N(i) \cup \{i\}$ denote the closed neighborhood of i .

Theorem 3 Let $G = (V, E)$ be a graph on the set of vertices $V = \{1, 2, \dots, n\}$, and let $U = \{r + 1, r + 2, \dots, n\} \subset V$ be a set of vertices of G , called the boundary. Define

$$c = \min_{W \subset V - U} \frac{|N'(W)|}{|W|}.$$

Let $x = (x_1, x_2, \dots, x_n)$ be a real non-negative vector assigning a real value $x_i \geq 0$ to each vertex i , where $x_i = 0$ for all $i \in U$. Then

$$\frac{c^2}{4(1+c)^2} \leq (\sqrt{1+c} - 1)^2 \leq \frac{\sum_{i \in V} [\max_{j \in N^+(i)} (x_j - x_i)]^2}{\sum_{i \in V} x_i^2}.$$

Note that in the above theorem, the quantity $[\max_{j \in N^+(i)} (x_j - x_i)]^2$ is not $\max_{j \in N^+(i)} (x_j - x_i)^2$, that is, we only maximize over $j \in N^+(i)$ satisfying $x_j \geq x_i$ (there is at least one such j , namely, $j = i$).

Proof: Without loss of generality, assume that $x_1 \geq x_2 \geq \dots \geq x_r \geq x_{r+1} = x_{r+2} = \dots = x_n = 0$.

We claim that

$$\sum_{i \in V} \max_{j \in N^+(i)} (x_j^2 - x_i^2) = \sum_{s=1}^r |N'(\{1, 2, \dots, s\})| (x_s^2 - x_{s+1}^2). \quad (9)$$

To see this, simply replace each difference $(x_j^2 - x_i^2)$ with $x_j^2 > x_i^2$ (and hence with $j < i$) in the expression in the left hand side by $(x_j^2 - x_{j+1}^2) + (x_{j+1}^2 - x_{j+2}^2) + \dots + (x_{i-1}^2 - x_i^2)$. Doing this, it is clear that the total number of times a nonzero difference of the form $(x_s^2 - x_{s+1}^2)$ is obtained is equal to the number of vertices $i > s$ for which $\max_{j \in N^+(i)} x_j \geq x_s$, that is, equal to $|N'(\{1, 2, \dots, s\})|$, as needed. (Since $x_s = 0$ for all $s > r$, there is no contribution from the terms $(x_s^2 - x_{s+1}^2)$ for $s > r$.) Since by the definition of c , $|N'(\{1, 2, \dots, i\})| \geq ci$ for all i , the equality (9) implies that

$$\sum_{i \in V} \max_{j \in N^+(i)} (x_j^2 - x_i^2) \geq \sum_{i=1}^r ci(x_i^2 - x_{i+1}^2) = c \sum_{i=1}^r x_i^2 = c \sum_{i \in V} x_i^2. \quad (10)$$

Note that for every $i \in V$,

$$\max_{j \in N^+(i)} (x_j^2 - x_i^2) = [\max_{j \in N^+(i)} (x_j - x_i)]^2 + 2x_i [\max_{j \in N^+(i)} (x_j - x_i)], \quad (11)$$

where here we used the fact that since $x_j \geq 0$ for all j , then for each vertex i , a vertex $j \in N^+(i)$ maximizes the value of $x_j^2 - x_i^2$ among all $j \in N^+(i)$ iff it maximizes the value of $x_j - x_i$ among these vertices j .

Plugging (11) in (10), we conclude that

$$c \sum_{i \in V} x_i^2 \leq \sum_{i \in V} \max_{j \in N^+(i)} (x_j^2 - x_i^2) = \sum_{i \in V} [\max_{j \in N^+(i)} (x_j - x_i)]^2 + 2x_i [\max_{j \in N^+(i)} (x_j - x_i)]. \quad (12)$$

Define $X^2 = \sum_{i \in V} x_i^2$ and $Y^2 = \sum_{i \in V} [\max_{j \in N^+(i)} (x_j - x_i)]^2$. Since $\max_{j \in N^+(i)} (x_j - x_i) \geq 0$ for all i , we conclude, by Cauchy Schwartz, that

$$\sum_{i \in V} x_i [\max_{j \in N^+(i)} (x_j - x_i)] \leq XY$$

and thus (12) implies that $cX^2 \leq Y^2 + 2XY$. Therefore $(c+1)X^2 \leq (X+Y)^2$ and thus $(\sqrt{c+1} - 1)^2 \leq \frac{Y^2}{X^2}$.

Since it is easy to check that for all $c \geq 0$

$$(\sqrt{1+c} - 1)^2 = \frac{c^2}{(\sqrt{1+c} + 1)^2} = \frac{c^2}{2+c+2\sqrt{1+c}} \geq \frac{c^2}{4(1+c)^2},$$

this completes the proof of the theorem. ■

5 Concluding remarks

- As mentioned in the introduction, every vertex spine of $G(m, d)$ must trivially contain at least a fraction of $\frac{1}{2m}$ of the total number of vertices. The authors of [4] showed that for $d = 2$ the smallest vertex spine is of size precisely $3m - 1$, which is roughly $\frac{3}{2}$ times the trivial lower bound. Since for any $d > 2$, the torus $G(m, d)$ can be partitioned into $(2m)^{d-2}$ pairwise disjoint copies of $G(m, 2)$, a lower bound of $(3m - 1)(2m)^{d-2}$ for the size of any vertex spine in $G(m, d)$ follows. It will be interesting to decide if the size of the smallest spine in $G(m, d)$ is in fact upper bounded by an absolute constant times $(2m)^{d-1}$.
- The proof of Theorem 1 can be easily modified to yield a similar result for the torus obtained by starting from an odd cycle, that is, the d -dimensional torus on a set of $(2m+1)^d$ vertices.

Acknowledgment: I would like to thank Emanuel Milman and Amir Yehudayoff for helpful discussions.

References

- [1] N. Alon, Eigenvalues and expanders, *Combinatorica* 6(1986), 83-96.
- [2] N. Alon and B. Klartag, Economical toric spines via Cheeger's Inequality, to appear.
- [3] S. Bobkov, C. Houdré and P. Tetali, λ_∞ , vertex isoperimetry and concentration, *Combinatorica* 20 (2000), no. 2, 153-172.
- [4] B. Bollobás, G. Kindler, I. Leader and R. O'Donnell, Eliminating cycles in the discrete torus, *LATIN 2006: Theoretical informatics*, 202-210, *Lecture Notes in Comput. Sci.*, 3887, Springer, Berlin, 2006. Also: *Algorithmica* 50 (2008), no. 4, 446-454.

- [5] G. Kindler, R. O'Donnell, A. Rao and A. Wigderson, Rounding schemes and cubical tilings with sphere-like surface area, Proc. FOCS 2008, to appear.
- [6] R. Raz, A counterexample to strong parallel repetition, Proc. FOCS 2008, to appear.