Random Cayley Graphs and Expanders

Noga Alon* Yuval Roichman †

February 22, 2002

Abstract

For every $1 > \delta > 0$ there exists a $c = c(\delta) > 0$ such that for every group G of order n, and for a set S of $c(\delta) \log n$ random elements in the group, the expected value of the second largest eigenvalue of the normalized adjacency matrix of the Cayley graph $X(G, S)$ is at most $(1-\delta)$. This implies that almost every such a graph is an $\varepsilon(\delta)$-expander. For Abelian groups this is essentially tight, and explicit constructions can be given in some cases.

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel. Research supported in part by a U.S.A.-Israeli BSF grant.

†Department of Mathematics, Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel

1. INTRODUCTION.

The (undirected) Cayley Graph $X(G,S)$ of a group G with respect to the set S of elements in the group is the graph whose set of vertices is G and whose set of edges is the set of all (unordered) pairs $\{\{g,gs\}|g \in G, s \in S\}$. Obviously, this is a regular graph of degree $|S \cup S^{-1}| \leq 2|S|$ and hence its diameter is at least $log_{2|S|}|G|$. Babai and Erdös [BE] proved that every group of order n has $\log_2 n + O(\log\log n)$ elements $x_1, \ldots, x_t$ such that every group element is a product of the form $x_1^{\varepsilon_1} \ldots x_t^{\varepsilon_t}, \varepsilon_i \in \{0,1\}$. It follows that G has a set of $log_2 n + O(\log\log n)$ generators such that the resulting Cayley graph has a logarithmic diameter. In [ABM] it is proved that the diameter is almost surely logarithmic with respect to $c\log n$ random elements, for an appropriate absolute constant c. For more details see [BE] and [ABM]. For a general survey on Cayley graphs with small diameters see [BHKLS].

Definition. A graph H is called a c-expander if for every set of vertices S, $|\Gamma(S)| > c|S|(1-|S|/|H|))$, where $\Gamma(S)$ is the set of all neighbours of S.

It is obvious that any c-expander (for some fixed $c > 0$) has a logarithmic diameter, but the converse does not hold, as shown, for example, by the Cayley graph of the symmetric group with respect to the set of all transpositions. For a graph H, let $\mu_1[H]$ denote the second largest eigenvalue in absolute value of the adjacency matrix A_H of H. If H is d-regular, the normalized adjacency matrix A_H^* of H is the doubly stochastic matrix $\frac{1}{d}A_H$. Let $\mu_1^*[H]$ denote the second largest eigenvalue in absolute value of A_H^*. Clearly $\mu_1^*[H] = \frac{1}{d}\mu_1[H]$. Here we prove the following:

Theorem 1: For every $1 > \delta > 0$ there is a $c(\delta) > 0$ such that the following holds. Let G be a group of order n and let S be a set of $c(\delta)\log_2 n$ random elements of G. Then

$$E(|\mu_1^*[X(G,S)]|) < 1 - \delta.$$

Note that the weaker statement that asserts that the expectation of the second normalized eigenvalue of a graph as above is at most $1 - O((\log n)^{-3})$ can be derived easily from the result of Babai in [Ba] which bounds the second eigenvalue of a vertex transitive graph as a function of its diameter.

By considering an appropriate martingale as done in [BS] and by the easy direction of the well known connection between the second eigenvalue of a graph and its expansion properties (see [AM] and [Ta]), this implies the following.

Corollary 1: For every $1 > \varepsilon > 0$ there exists a $c(\varepsilon) > 0$ such that the following holds. Let G be a group of order n, and let S be a random set of

$c(\varepsilon)\log_2 n$ elements of G, then the Cayley graph $X(G,S)$ is an ε-expander almost surely. (I.e., the probability it is such an expander tends to 1 as n tends to infinity.)

Simple examples (together with some techniques from the theory of error correcting codes) show that the best possible constant $c(\delta)$ in Theorem 1 is at least $1+\Omega(\delta\log(1/\delta)))$ for any small positive δ. In our proof of Theorem 1 we make no attempt to optimize the upper bound for $c(\delta)$ and the proof gives that for $0<\delta<1-\frac{1}{e}$, $c(\delta)\leq(1+o(1))2e^4\ln 2/[(1-\delta)e-1]^2$. We give sharper bounds for Abelian groups and for some related groups. We also discuss some explicit constructions for the groups Z_2^m by applying techniques from the theory of error correcting codes in order to estimate the relevant eigenvalues.

The rest of the paper is organized as follows. In Section 2 we consider random Cayley graphs of arbitrary finite groups and present the proofs of Theorem 1 and Corollary 1. The special case of Abelian groups is considered in Section 3.

2. RANDOM CAYLEY GRAPHS OF GENERAL GROUPS.

2.1.Proof of Theorem 1.

To simplify the notation we omit all floor and ceiling signs whenever these are not crucial. We use Wigner's method of proving the semi-circle-law, see [Wi], [Wi1]. We bound the traces of the powers of the adjacency matrix, by estimating the probability of a random walk to be closed. Broder and Shamir ([BS]) were the first to apply this method to estimate the eigenvalues of sparse random graphs. See also [Fr] and [FK]. Let A be a real matrix, with eigenvalues $|\mu_0|\geq|\mu_1|\ldots\geq|\mu_{n-1}|$. Then for every natural number m:

$$|\mu_1|\leq(Tr(A^{2m})-\mu_0^{2m})^{1/2m}$$

The probability space considered is that consisting of all the normalized adjacency matrices of the Cayley graphs of a given group G with respect to $c\log_2 n$ elements chosen randomly in the group. By Jensen's Inequality, for every positive integer m:

$$E(|\mu_1^*|)\leq(E(Tr(A^{2m}))-1)^{1/2m}$$

Denote by n the order of G and by P_{2m} the probability of a walk of length $2m$ in our Cayley graph to be closed. Since a Cayley graph is vertex transitive $E(Tr(A^{2m}))=nE(P_{2m})$, and hence:

$$E(|\mu_1^*|)\leq(nE(P_{2m})-1)^{1/2m} \tag{1}$$

Lemma 1.

$$E(P_{2m}) \leq 2^{2m}(2/c\log_2 n)^{m\ln\ln 2m/\ln 2m} + \\ +2^{l(w)}(m/c\log_2 n)^{l(w)/2} + \\ +1/n + O(m/n^2)$$

where

$$l(w) = 2m(1 - \ln\ln 2m/\ln 2m)$$

Proof: As in [BS] we consider a dynamic process for choosing the random set S and the random walk on $X(G,S)$. This is done as follows.

a We choose in the free monoid $M_{2c\log_2 n}$ (generated by $c\log_2 n$ distinct letters and their inverses) a random word of length $2m$.

b We reduce this word in the free group $F_{c\log_2 n}$.

c We assign to each letter an element of the group G at random.

It is easy to see that this process is equivalent to the one in which a random Cayley graph $X(G,S)$ with $|S| = c\log_2 n$ is chosen first and a random walk of length $2m$ in it is chosen afterwards.

In order to obtain an upper bound for $E(P_{2m})$ we estimate the probabilities of the following three events whose union includes the event that our walk of length $2m$ is closed.

A The reduced word in the free group has length $l(W)$ and

$$l(W) < l(w) = 2m(1 - \ln\ln 2m/\ln 2m)$$

B The reduced word has length at least $2m(1 - \ln\ln 2m/\ln 2m)$, and there is no letter which appeares exactly once in this word.

C (A) and (B) do not hold, but after the assignment of the chosen elements in the group G to the corresponding letters the word is reduced to the unity.

Obviously

$$E(P_{2m}) \leq Pr(A) + Pr(B) + Pr(C) \tag{2}$$

In order to estimate $Pr(A)$ we need the following simple fact whose proof can be found in [BS] or in [Lu] §4.5.

Fact: Let U be a random word of length $2t$ in the free monoid M_{2d}, then

$$(2/d)^t \geq Pr(U \text{ is an identity sequence}).$$

The number of possibilities to place the subset which is reduced to the identity is at most 2^{2m}. The length of this subset is at least

$$2m \ln \ln 2m / \ln 2m$$

and $d = c \log_2 n$. Hence, by the fact above,

$$Pr(A) \leq 2^{2m} (2/c \log_2 n)^{m \ln \ln 2m / \ln 2m} \tag{3}$$

In order to obtain an upper bound to $Pr(B)$ we have to estimate the probability of a reduced word of length $l(W)$ in $F_{c \log_2 n}$ to satisfy the condition of (B), i.e., to have no letter which appears exactly once. For our purposes a rough estimate suffices. Assume that the word W satisfies the conditions of (B). Obviously, the number of distinct letters that appear in W is at most $l(W)/2$. We expose the letters of W in the following order: First, we expose the subset consisting of the first occurrence of each letter that appears in the word. Second, we expose the other letters. The probability of each letter of the latter to be one which has appeared in the first subset is at most $l(W)/2c \log_2 n$. The number of possiblities to place the first subset is at most $2^{l(W)}$. Hence

$$Pr(B) \leq 2^{l(W)} (l(W)/2c \log_2 n)^{l(W)/2} \leq 2^{l(W)} (m/c \log_2 n)^{l(W)/2}, \tag{4}$$

where

$$l(W) \geq 2m(1 - \ln \ln 2m / \ln 2m).$$

It remains to bound $Pr(C)$. Let τ be a letter that appears only once in the reduced word. We expose the assignments of all the letters except that of τ . Denote by $x(\tau)$ the assignment of τ. The event whose probability we wish to estimate is now the event $gx(\tau)h = 1$ where g, h are some known elements in G. The probability that $x(\tau)$ solves this equation is at most $1/(n - 2m) = 1/n + O(m/n^2)$. Therefore

$$Pr(C) \leq 1/n + O(m/n^2) \tag{5}$$

The assertion of Lemma 1 now follows by substituting (3), (4) and (5) in (2). (Observe that in case the right hand side of (4) is at most 1, it is a monotone decreasing function of $l(W)$ and hence its maximum is obtained for the minimum possible value of $l(W)$ which is $l(w)$. Otherwise, the estimate in Lemma 1 is trivial.) $\Box$

Combining Lemma 1 (and its proof) with (1) we obtain

$$E(|\mu_1^*|) \le (nE(P_{2m}) - 1)^{1/2m} \le (n(Pr(A) + Pr(B) + Pr(C)) - 1)^{1/2m}$$

$$\le n^{1/2m}(Pr(A)^{1/2m} + Pr(B)^{1/2m}) + (nPr(C) - 1)^{1/2m}$$

$$\le n^{1/2m}(2(\frac{2}{c\log_2 n})^{\frac{1}{2}\ln\ln 2m/\ln 2m} + 2^{1-o(1)}(\frac{m}{c\log_2 n})^{\frac{1}{2}-o(1)}) + (O(m/n))^{1/2m}.$$

Setting $2m \sim \frac{\ln n}{b}$, where b is an absolute constant, we conclude that for large n the above upper bound for $E(|\mu_1^*|)$ is

$$(1 + o(1))(e^b(\frac{\ln 4}{cb})^{1/2} + \frac{1}{e^b}).$$

For every fixed $1 > \delta > 0$ one can obviously choose b and c so that the last quantity is smaller than $1 - \delta$. In particular, if $\delta < 1 - \frac{1}{e}$ then $b = 1$ and $c \ge (1 + o(1))2e^4 \ln 2/[(1 - \delta)e - 1]^2$ imply $E|\mu_1^*| \le 1 - \delta$. This completes the proof of Theorem 1. □

Remark The group structure is not essential in the last proof, and it is easy to modify it and obtain the following result, whose detailed proof is omitted.

Proposition 1. For every $1 > \delta > 0$ there is a $c(\delta) > 0$ such that the following holds. Let K be a complete graph on n vertices whose edges are colored by $n - 1$ colors so that each color class forms a perfect matching. Let H be the subgraph of K consisting of all edges of K whose colors belong to a randomly chosen set of $c(\delta)\log_2 n$ colors. Then

$$E(|\mu_1[H]|) < (1 - \delta)c(\delta)\log_2 n.$$

2.2 Proof of Corollary 1.

Broder and Shamir [BS] proved, by considering an appropriate martingale, that the second eigenvalue (in absolute value) of the adjacency matrix of a random d regular graph is concentrated around its mean. Their method easily implies the following result.

Lemma 2. Let G be a group and let S be a random set of d elements in G. Put $\mu_1 = \mu_1[X(G, S)]$. Then:

$$Pr(|\mu_1 - E\mu_1| \ge 2cd^{1/2}) \le 4e^{-c^2/2}$$

Proof. It is convenient to choose the elements $x_1, \ldots, x_d$ of the random set S one by one, in the order of their indices. Let $X(G, S)$ denote the

Cayley graph of G with respect to S and let $d = \lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_{n-1}$ denote its eigenvalues. Fix j, $1 \leq j \leq n-1$ and let $F_i^j = F_i^j(x_1, \ldots, x_i)$ denote the expectation of $\lambda_j(X(G,S))$, given the values of $x_1, \ldots, x_i$. It is not too difficult to check that for every fixed j the sequence $(F_0^j, \ldots, F_d^j)$ is a martingale. Moreover, by the variational definition of the jth eigenvalue it is easy to see that $|F_{i+1}^j - F_i^j| \leq 2$. Therefore, by Azuma's Inequality, (cf. e.g., [AS]), for every admissible fixed j

$$Pr(|\lambda_j - E\lambda_j| \geq 2cd^{1/2}) = Pr(|F_d^j - F_0^j| \geq 2cd^{1/2}) \leq 2e^{-c^2/2}.$$

Since the second largest eigenvalue in absolute value is either λ_1 or λ_{n-1} the desired result follows. □

Recall that a regular graph Γ of constant degree is a δ-expander if and only if the second largest eigenvalue of its normalized adjacency matrix is at most $1 - \varepsilon(\delta)$. (The easy direction of this fact is proved in [Ta] and [AM]; the more difficult one is established in [Al]). The easy direction of this fact (which works for non-constant degrees as well), the last lemma and Theorem 1 imply Corollary 1. It is worth noting that as pointed out by one of the referees an alternative way to prove the corollary is by replacing Lemma 2 by the observation that the argument in Theorem 1 can be used to bound higher moments of the second eigenvalue and not only its first moment.□

3. CAYLEY GRAPHS OF ABELIAN GROUPS.

3.1 Random Cayley graphs.

It is trivial to see that the $log_2 n$ term appearing in Theorem 1 is best possible; in fact, for the special case of the groups $G = Z_2^m$, which are simply the vector spaces over Z_2, any connected Cayley graph $X(G,S)$ has at least $m = \log_2 |G|$ elements in S, and hence S is at least of that cardinality for every Cayley graph on such a group whose second eigenvalue differs from the first. A closer inspection of the eigenvalues of the Cayley graphs of these groups implies that the best possible constant $c(\delta)$ in Theorem 1 satisfies, for small $\delta > 0$;

$$c(\delta) \geq 1 + \Omega(\delta \log(\frac{1}{\delta})).$$

This is proved in the following proposition.

Proposition 2. Suppose $G = Z_2^m$ and let $X(G,S)$ be a Cayley graph of G whose second largest eigenvalue in absolute value is at most $(1-\delta)|S|$. Then

$$|S| \geq (1 + \Omega(\delta \log(\frac{1}{\delta})))m.$$

Proof. As is well known (see, e.g., [Lo]), the eigenvalues of $X(G,S)$ are simply all the quantities

$$\sum_{s \in S \cup S^{-1}} \chi(s),$$

where χ ranges over all characters of G. For the groups Z_2^m these can be interpreted in the following convenient way. Let $B = (b_{ij})$ be a binary matrix with m rows and $|S|$ columns whose columns are the elements of S. Put $l = |S|$ and let U be the subspace of Z_2^l generated by the rows of B. For each vector $u \in U$ let $e(u)$ denote the difference between the number of 0-entries and the number of 1-entries of u. Then $\{e(u) : u \in U\}$ is the set of all eigenvalues of the graph $X(G,S)$. Therefore, if we consider B as the generating matrix of a linear error correcting code it follows that for our graph $|\mu_1| < (1-\delta)l$ if and only if the Hamming weight of every nonzero code word is between $\delta l/2$ and $l - \delta l/2$. In particular, the distance of the code is at least $\delta l/2$. Therefore, by the sphere packing bound (see, e.g., [MS])

$$\sum_{j < \delta l/4} \binom{l}{j} 2^m < 2^l,$$

and this implies the desired estimate. □

Surprisingly, the above lower estimate for $c(\delta)$ that holds for Z_2^m is essentially sharp as an upper estimate for *every* Abelian group. This is proved in the following theorem. (In this section it is more convenient to consider the model in which the members of the generating set S are chosen with repetitions. This does not make any essential difference, since a random sequence of $O(\log n)$ elements in a group of order n consists, almost surely, of distinct members.)

Theorem 2. For each small $\delta > 0$ there exists an $\varepsilon = O(\delta \log(1/\delta))$ such that for any Abelian group A of order n the Cayley graph $X(A,S)$ of A with respect to a (multi-) subset S of $(1+\varepsilon)\log_2 n$ (not necessarily distinct) random elements satisfies $|\mu_1^*[X(A,S)]| \leq 1-\delta$ almost surely (i.e., with probability that tends to 1 as n tends to infinity).

Proof. Let A be an Abelian group. Then $A = Z_{p_1} \oplus Z_{p_2} \oplus \cdots \oplus Z_{p_k}$. Each nontrivial character of A is simply a product of characters of the cyclic groups Z_{p_i}, and at least one of these is non-trivial. It is therefore easy to see that for each fixed nontrivial character χ as above

$$Pr(\chi(\sigma + \sigma^{-1}) \geq 1.8) \leq 1/2$$

where σ is a random element of A. (Here 1.8 is not optimal; the essential fact is merely that it is an absolute constant which is strictly smaller than 2. Note that if σ is an element of order 2 then $\sigma = \sigma^{-1}$ and then the above inequality simply means that $\chi(\sigma)$ is -1 and not 1.)

The same reasoning implies that for each nontrivial character χ

$$Pr(\chi(\sigma + \sigma^{-1}) \leq -1.8) \leq 1/2.$$

As A is an Abelian group, the eigenvalues of the adjacency matrix are the sums of the characters $\sum_{s \in S \cup S^{-1}} \chi(s)$. Since the members of S are chosen independently, the above two inequalities imply that for each fixed nontrivial character χ the probability that the corresponding normalized eigenvalue μ^* has absolute value strictly greater than $1 - \delta$ satisfies:

$$Pr(|\mu^*| \geq 1 - \delta) \leq 2 \left(\begin{array}{c} (1+\varepsilon)\log_2 n \\ \\ c\delta(1+\varepsilon)\log_2 n \end{array} \right) (1/2)^{(1-c\delta)(1+\varepsilon)\log_2 n},$$

where c is some absolute constant.

There are $n-1$ nontrivial characters and hence the probability that at least one of them gives a normalized eigenvalue whose absolute value exceeds $(1-\delta)$ is at most

$$2(n-1) \left(\begin{array}{c} (1+\varepsilon)\log_2 n \\ \\ c\delta(1+\varepsilon)\log_2 n \end{array} \right) (1/2)^{(1-c\delta)(1+\varepsilon)\log_2 n}.$$

The upper bound tends to zero when n tends to infinity provided ε is at least $c'\delta\log(1/\delta)$ for an appropriately chosen $c' = c'(c)$. This completes the proof. $\Box$

Remark. A simple modification of the proof above implies that for certain Abelian groups the multiplicative constant of the logarithm of the size of the group can be reduced considerably and still imply the conclusion of Theorem 2. In particular, it is easy see that for every small $\epsilon > 0$ there exists a $\delta = \delta(\epsilon) > 0$ such that the Cayley graph $X(Z_n, S)$ of the cyclic group Z_n with respect to a subset S of $\epsilon \log_2 n$ random elements satisfies $|\mu_1^*[X(A,S)]| \leq 1 - \delta$ almost surely (i.e., with probability that tends to 1 as n tends to infinity). We omit the details.

We conclude this subsection with the following simple observation showing that the $\log n$ term in Theorem 1 is necessary for every Abelian group.

Proposition 3. For every fixed $\delta > 0$ there is a constant $c = c(\delta) > 0$ such that the following holds; If A is an Abelian group of order n, and $X(A, S)$ is a δ-expander, then $|S| \geq c \log_2 n$.

Proof: Suppose $|S| = d$ and let the diameter of the Cayley graph $X(A, S)$ be D. Then every element of A is a product of of the form

$$s_1^{a_1} \cdot s_2^{a_2} \dots s_d^{a_d},$$

where $S = \{s_1, \dots, s_d\}$, each a_i is an integer and $\sum_{i=1}^{d} |a_i| \leq D$. It is easy to see that the total number of products of this form is at most

$$2^d \binom{D+d}{d},$$

since there are at most $\binom{D+d}{d}$ possibilities to choose the absolute values $|a_i|$, and for each such choice there are at most 2^d ways to assign signs to the non-zero numbers a_i. Therefore,

$$n \leq 2^d \binom{D+d}{d}. \tag{6}$$

If $X(A, S)$ is a δ-expander then its diameter D clearly satisfies $D \leq c' \log_2 n$ for some $c' = c'(\delta)$. This, together with (6) implies that $d \geq c \log_2 n$ for an appropriate $c = c(c')$, completing the proof. □

Remark. If G is a group, $S \subset G$ and H is a factor of G then trivialy the diameter of the Cayley graph of H with respect to the set S' of images of the elements of S is at most that of $X(G, S)$. Therfore, if G has an Abelian factor of order q the diameter D of $X(G, S)$ is at least that of $X(H, S')$ and hence

$$2^d \binom{D+d}{d} \geq q,$$

where $d = |S|$. It follows that if $X(G, S)$ is a δ-expander and $q \geq |G|^\epsilon$ for some fixed $\delta > 0$ and $\epsilon > 0$, then $|S| \geq c(\delta, \epsilon) \log |G|$. Thus, for example, since $Gl(n, q)/Sl(n, q) \cong Z_{q-1}$ the following claim holds.

Claim. For every fixed n and $\delta > 0$ there exists a constant $c = c(\delta, n)$ such that if $G_q = Gl(n, q)$, $S_q \subset G_q$ and the family of Cayley graphs $\{X(G_q, S_q)\}$ is a family of δ-expanders then $|S_q| \geq c \log |G_q|$.

3.2 Some explicit constructions.

Since expanders have numerous applications in theoretical computer science, the problem of constructing them explicitly received a considerable

amount of attention. Although the most useful expanders are the constant degree ones, expanders with higher degrees (that may be polylogarithmic in the number of vertices or even a small fixed power of this number) are also useful and have been applied extensively in the design of efficient parallel comparison algorithms (see, e.g., [Al1]) as well as in various other algorithmic applications. By Theorem 1, expanders with degrees greater than the logarithm of the number of vertices can be constructed from any finite group, and it is natural to try the simplest possible groups. The expanders in [Al1] are Cayley graphs of the cyclic groups Z_m. However, their degrees are at least the square root of the number of vertices. In [AIKPS] the authors construct explicit expanders with degrees $d = O(\log m(\log^* m)^{13\log^* m})$ (and second eigenvalue $O(d/\log^* m) = o(d)$) which are also Cayley graphs of the cyclic groups Z_m. Here we consider the groups Z_2^m where we can construct explicitly Cayley graphs that are expanders of any degree that exceeds some absolute constant times the logarithm of the order of the group. The technique can be easily extended to the more general case of F_q^m where F_q is any finite field but we prefer to discuss only the case $q = 2$ here.

It is well known (see [Al], [Ni], [Fr1], [Ka]) that the second largest eigenvalue of any regular graph with n vertices and degree d is at least

$$2\sqrt{d-1}(1 - O(\frac{\log^2 d}{\log^2 n}))$$

and this estimate is essentially optimal as shown by the Ramanujan graphs ([LPS], [Ma]) or by random regular graphs ([Fr],[FKS]). Here we describe, explicitly, for every fixed $k \geq 1$ a Cayley graph of degree $d = m^k$ of Z_2^m, every nontrivial eigenvalue of which is, in absolute value, at most $O(d^{0.5+1/k})$. Non explicitly this can be improved to $O(d^{0.5+1/2k})$. We further show that the second eigenvalue of any Cayley graph of degree $d = m^k$ of Z_2^m is at least $\Omega(d^{0.5+1/2k}/\sqrt{\log d})$. Thus the nontrivial eigenvalues of such graphs can be separated quite well from the first one, but not as well as the eigenvalues of random graphs or Ramanujan graphs. Similar results are proved in the end of the paper for general Abelian groups.

Most of the results in this subsection are all rather simple consequences of known results concerning error correcting codes. For a subset S of Z_2^m let $B = B(S)$ denote the m by $|S|$ binary matrix whose columns are the elements of S. Put $|S| = l$. As observed in the proof of Proposition 2, the Cayley graph $H = X(Z_2^m, S)$ satisfies $|\mu_1[H]| \leq \epsilon l$ if and only if the Hamming weight of every nonzero code word in the linear error correcting code generated by the rows of B is between $\frac{1-\epsilon}{2}l$ and $\frac{1+\epsilon}{2}l$. Although this

is not precisely the common notation used in Coding Theory, let us call a code an $[l, m, \epsilon]$-code if it is a linear code of length l, dimension m and the weight of each nontrivial codeword is between $\frac{1-\epsilon}{2}l$ and $\frac{1+\epsilon}{2}l$. By the above remark, explicit codes as above can be used to obtain explicit expanders. In [ABNNR] the basic idea in the well known construction of Justesen (see, e.g., [MS]) is combined with certain explicit Ramanujan graphs to obtain explicitly the generating matrices of $[l, m, \epsilon]$ codes for each fixed $1 > \epsilon > 0$ and for each m, where $l = O(\frac{1}{\epsilon^3}m)$. This gives the following.

Proposition 4. There exists an absolute constant $c > 0$ such that for every fixed $\epsilon > 0$ and for every m one can describe explicitly a Cayley graph $H = X(Z_2^m, S_m)$ where

$$|S_m| \leq c\frac{1}{\epsilon^3}m \quad (= c\frac{1}{\epsilon^3}\log_2 |Z_2^m|),$$

so that $|\mu_1[H]| \leq \epsilon |S_m|$.

Note that in this construction bounded degree expanders which are Cayley graphs of noncommutative groups are used to obtain expanders of logarithmic degrees which are Cayley graphs of commutative groups.

In [AGHP] there are three explicit constructions of $[l, m, \epsilon]$ codes (where now ϵ is not necessarily fixed), and all of them satisfy $l \leq O(m^2/\epsilon^2)$. (The best one in fact satisfies

$$l \leq O(\frac{m^2}{\epsilon^2(\log(m/\epsilon))^2}.)$$

Here is the description of the generating matrix B of one of these constructions, which is extremely simple to describe. Let l be a prime, $l \geq (\frac{2m}{\epsilon}+1)^2$, and define $B = (b_{ij})$ as follows. For $1 \leq i \leq m, 1 \leq j \leq l$, $b_{ij} = 0$ if $i - j$, considered as an element of the finite field Z_l , is a quadratic residue, and $b_{ij} = 1$ otherwise. By this construction (as well as by the other constructions given in [AGHP]) we obtain:

Proposition 5. There exists an absolute constant $c > 0$ such that for every $\epsilon > 0$ and for every m one can describe explicitly a Cayley graph $H = X(Z_2^m, S_m)$ where $l = |S_m| \leq cm^2/\epsilon^2$, so that $|\mu_1[H]| \leq \epsilon |S_m| = O(\sqrt{l}m)$.

Note that when the degree l of the expander above is $m^k = (\log_2 |Z_2^m|)^k$ then $|\mu_1| \leq O(l^{0.5+1/k})$.

By the Gilbert-Varshamov bound (see, e.g., [MS]), or , in fact, by the obvious modification of it needed to obtain a code where the weight of each code-word is very close to half of its length, is is easy to obtain (non-constructively) the following result:

Proposition 6. There exists an absolute constant $c > 0$ such that for every $1 > \epsilon > 0$ and for every m there exists a Cayley graph $H = X(Z_2^m, S_m)$ where $l = |S_m| \leq cm/\epsilon^2$, so that $|\mu_1[H]| \leq \epsilon |S_m| = O(\sqrt{l}\sqrt{m})$.

Finally, we observe that by the McEliece- Rodemich- Rumsey- Welch bound (see [MS], page 559, where the proof for fixed ϵ is given, and note that this proof can be extended, as mentioned in [AGHP], for the case of much smaller ϵ as well), if $\epsilon \geq 2^{-\gamma m}$ for some small fixed $\gamma > 0$, and the weight of each nontrivial code word in a linear code of length l and dimension m is at least $\frac{1-\epsilon}{2}l$ then $l \geq \Omega(\frac{m}{\epsilon^2 \log(1/\epsilon)})$. This implies the following.

Proposition 7. There exists an absolute constant $\gamma > 0$ such that for every $m \leq l \leq 2^{\gamma m}$, the second largest eigenvalue of any Cayley graph $X(Z_2^m, S)$ with $|S| = l$ is at least $\Omega(\sqrt{lm}/\sqrt{\log l})$.

Note that this shows that for large m no such graph H in which the degree is at most some small power of the number of vertices can be a Ramanujan graph, i.e., satisfy $|\mu_1[H]| \leq 2\sqrt{l-1}$.

Proposition 6, as well as a slightly weaker version of Proposition 7, can be generalized to arbitrary Abelian groups. Indeed, if G is an arbitrary Abelian group of order n, χ is a fixed nontrivial character, and S is a sequence of l random members of G, then the value of the eigenvalue

$$\sum_{s \in S \cup S^{-1}} \chi(s)$$

of $X(G, S)$ can be estimated as a sum of l independent random variables each of which has a bounded absolute value. Thus, by the estimates in, e.g., [AS], the following assertion holds.

Proposition 6'. For every Abelian group G of order n and for every l, if S is a set of l random members of G then almost surely $|\mu_1[X(G, S)]| \leq O(\sqrt{l}\sqrt{\log n})$.

A lower bound for $|\mu_1[X(G, S)]|$ can be derived for Abelian groups as follows. Let G be an Abelian group of order n, and let S be an arbitrary set of l members of G. Suppose $l \geq 2m$ and consider the number of closed walks of length $2m$ in $X(G, S)$. This number is clearly at least the number of vertices times the number of words of length $2m$ consisting of a single appearance of x_i and a single appearance of x_i^{-1} for m distinct members x_i of S. Hence this number is at least

$$n \frac{(2m)!}{2^m m!} l(l-1) \cdots (l-m+1) \geq n(\frac{ml}{4})^m.$$

(Here n is the number of vertices and $\frac{(2m)!}{2^m m!}$ is the number of ways to split a sequence of length $2m$ into m subsequences of size 2 each, where each

subsequence will be that consisting of the occurrence of x_i and x_i^{-1}. The last term, $l(l-1)\cdots(l-m+1)$ is a lower bound on the number of ways to choose the actual elements x_i, where an extra factor of 2^m could be added in case no element is of order 2.) Therefore

$$|\mu_1[X(G,S)]| \geq (\frac{n(ml/4)^m - (2l)^{2m}}{n-1})^{1/2m}.$$

By taking $2m \sim \frac{logn}{\log 2l}$ (which is less than l as long as $l \geq c \log n / \log\log n$ for an appropriate $c > 0$) this gives the following.

Proposition 7'. For every Abelian group G of order n and for every $l = \Omega(\log n / \log\log n)$, if S is an arbitrary set of l members of G then

$$|\mu_1[X(G,S)]| \geq \Omega(\sqrt{l}\sqrt{\frac{\log n}{\log l}}).$$

Note that unlike Proposition 7 this only supplies a lower bound for the maximum absolute value of a nontrivial eigenvalue, and not for the second eigenvalue. Still, it shows that Cayley graphs of Abelian groups in which the degree is at most some small power of the number of vertices cannot be Ramanujan graphs. (As shown in [Al1] they can be Ramanujan graphs for degrees which are about the square root of the number of vertices.)

Acknowledgments. We would like to thank A. Lubotzky for suggesting some of the problems considered here. We also thank A. Frumkin for useful discussions and E. Shamir for helpful remarks.

References

[AIKPS] M. Ajtai, H. Iwaniec, J. Komlós, J. Pintz and E. Szemerédi, Construction of a thin set with small Fourier coefficients, Bull. London Math. Soc. 22 (1990), 583-590.

[Al] N. Alon, Eigenvalues and expanders, Combinatorica 6(1986), 83-96.

[Al1] N. Alon, Eigenvalues, geometric expanders, sorting in rounds and Ramsey Theory, Combinatorica 6(1986), 207-219.

[ABM] N. Alon, A. Barak and U. Manber, On disseminating information reliably without broadcasting, Proc. of the 7^{th} International Conference on Distributed Computing Systems (ICDS), Berlin, September 1987, pp. 74-81.

[ABNNR] N. Alon, J. Bruck, J. Naor, M. Naor and R. Roth, Construction of asymptotically good, low-rate error-correcting codes through pseudo-random graphs, IEEE Transactions on Information Theory 38 (1992), 509-516.

[AM] N. Alon and V. D. Milman, λ_1, isoperimetric inequalities for graphs and superconcentrators, J. Comb. Theory, Ser.B, 38 (1985), 73–88. See also: N. Alon and V. D. Milman, Eigenvalues, expanders and superconcentrators, Proc. 25^{th} IEEE FOCS, Singer Island, Florida, IEEE(1984), 320-322.

[AGHP] N. Alon, O. Goldreich, J. Hastad and R. Peralta, Simple constructions of almost k-wise independent random variables, Proc. 31^{st} IEEE FOCS, St. Louis, Missouri, IEEE (1990), 544-553. See also: Random Structures and Algorithms 3 (1992), 289-304.

[AS] N. Alon and J. H. Spencer, The Probabilistic Method, Wiley, 1991.

[Ba] L. Babai, Local expansion of vertex transitive graphs and random generation of finite groups, Proc. 23^{rd} Annual ACM STOC, ACM Press (1991), 164-174.

[BE] L. Babai and P. Erdös, Representation of group elements as short products, in: Theory and Practice of Combinatorics, G. Sabidussi, editor, Ann. Disc. Math. 12 (1982), 27–30.

[BHKLS] L. Babai, G. Hetyei, W. M. Kantor, A. Lubotzky and A. Seress, On the diameter of finite groups, Proc. 31^{st} IEEE FOCS, IEEE (1990), 857-865.

[BS] A. Broder and E. Shamir, On the second eigenvalue of random regular graphs, Proc. 28^{th} IEEE FOCS, IEEE (1987), 286–294.

[Fr] J. Friedman, On the second eigenvalue and random walks in random d-regular graphs, Combinatorica 11 (1991), 331-362.

[Fr1] J. Friedman, Some geometric aspects of graphs and their eigenfunctions, Duke Mathematical Journal 69 (1993), 487-525.

[FK] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, Combinatorica 1(1981), 233–241.

[FKS] J. Friedman, J. Kahn and E. Szemerédi, On the second eigenvalue in random regular graphs, Proc. 21^{st} ACM STOC, ACM Press (1989), 587–598.

[Ka] N. Kahale, On the second eigenvalue and linear expansion of regular graphs, Proc. 33^{rd} Annual IEEE FOCS, IEEE (1992), 296-303.

[Lo] L. Lovász, Combinatorial Problems and Exercises, North Holland, Amsterdam, 1979, Problem 11.8.

[Lu] A. Lubotzky, Discrete Groups, Expanding Graphs and Invariant Measures, to appear.

[LPS] A. Lubotzky, R. Phillips and P. Sarnak, Explicit expanders and the Ramanujan conjectures, Proc. 18^{th} ACM STOC (1986), 240-246. See also: A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan graphs, Combinatorica 8 (1988), 261-277.

[Ma] G. A. Margulis, Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and superconcentrators, Problemy Peredachi Informatsii 24(1988), 51-60 (in Russian). English translation in Problems of Information Transmission 24(1988), 39-46.

[MS] F. J. MacWilliams and N. J. A. Sloane, The Theory of Error-Correcting Codes, North Holland, Amsterdam, 1977.

[Ni] A. Nilli, On the second eigenvalue of a graph, Discrete Mathematics 91 (1991), 207-210.

[Ta] R. M. Tanner, Explicit construction of concentrators from generalized N-gons, SIAM J. Alg. Disc. Meth. 5 (1984), 287-293.

[Wi] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Ann. Math. 62 (1955), 548–564.

[Wi1] E. P. Wigner, On the distribution of the roots of certain symmetric matrices, Ann. Math. 67 (1958), 325-327.