Electronic Colloquium on Computational Complexity, Report No. 1 (2001)

Essentially every unimodular matrix defines an expander

Jin-Yi Cai [*]
Department of Computer Science and Engineering
State University of New York
Buffalo, NY 14260
and
Computer Sciences Department
University of Wisconsin
Madison, WI 53706
jyc@cs.wisc.edu

Abstract

We generalize the construction of Gabber and Galil to essentially every unimodular matrix in $SL_2(\mathbf{Z})$. It is shown that every parabolic or hyperbolic fractional linear transformation explicitly defines an expander of bounded degree and constant expansion. Thus all but a vanishingly small fraction of unimodular matrices define expanders.

1 Introduction

It has been recognized in the last 25 years that certain combinatorial objects called expanders are extremely useful in a number of computer science applications. These include sorting networks, superconcentrators and sparse connection networks in general, pseudorandom generators and amplifications and deterministic simulations, to name just a few.

An (n, k, d) *expander* is a bipartite graph $G = (L, R, E)$, with $|L| = |R| = n$ and at most kn edges, such that for every subset X of L, the neighbor set in R has $|\Gamma(X)| \geq [1 + d(1 - |X|/n)]|X|$. Thus, for every subset of input vertices of cardinality at most, say, $n/2$, its neighbor set *expands*, having cardinality at least a constant multiple more than $|X|$. It is generally desired to have k and d fixed and n grows to infinity.

The first existence theorems on expanders were provided by probabilistic counting argument [11][33]. Roughly speaking, such a proof starts by defining a certain probability space of graphs, and then one shows that the probability of such graphs is non-zero. In fact it is usually shown that such probability tends to 1. Thus not only such graphs exist, but they exist in *abundance*. The weakness of such a proof is that it is not explicit.

Margulis [29] was the first to give an explicit construction of a sequence of graphs $\{G_n\}$. This major achievement uses group representation theory. However, while his construction is explicit, the constant of expansion was not explicitly known. Gabber and Galil [20] in a beautiful paper

[*] Research supported in part by NSF grant CCR-9634665, and by a Guggenheim Fellowship. A preliminary version appeared in *The 11th International Symposium on Algorithm and Computation* (ISAAC) 2000, Taipei, Taiwan. Springer-Verlag Lecture Notes in Computer Science, D. T. Lee and Shang-Hua Teng (Eds.) Vol **1969** (2000) 2–22.

ISSN 1433-8092

gave an explicit construction of graphs $\{G_n\}$ with an explicitly stated constant of expansion. The Gabber-Galil proof also has the added advantage of being relatively elementary. We will follow the proofs of [20] closely. There is an extensive literature on expanders and their applications to the theory of computing, the reference section contains an incomplete list of important works. It was realized that expansion properties are closely related to the second largest eigenvalues of the graph $\lambda(G)$ (see [35, 7]), and for d-regular graphs the gap between d and $\lambda(G)$ provides estimates for both upper and lower bound for the expansion constant. The best construction was given by Lubotsky, Phillip and Sarnak [28] and by Margulis [30], where asymptotically optimal $\lambda(G)$ was achieved. The proofs in [28] use deep results from number theory, especially results of Eichler and Igusa concerning the Ramanujan conjecture.

We also mention the interesting construction of Ajtai, Komlós and Szemerédi [4], where they showed a randomly chosen transposition and a full cycle over the group S_n also supply an expander. If the original probabilistic constructions are one extreme of showing the "abundance" of expander graphs, the proof in [4] can be viewed as a construction with reduced randomness. The other extreme is of course the explicit constructions mentioned above. Recently, Reingold et. al. [34] considered a new construction technique called zig-zag graph product.

In this paper, we generalize the construction of Gabber and Galil [20] to essentially every unimodular matrix in $SL_2(\mathbf{Z})$. Our proofs are relatively elementary. They do provide a certain "abundance" as well as being explicit, with the same expansion constant $1 - \sqrt{3}/2$ as in [20]. It is shown that *every* parabolic or hyperbolic fractional linear transformation explicitly defines an expander of bounded degree and constant expansion.

2 Preliminary Remarks

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be an integral unimodular matrix, i.e., $A \in SL_2(\mathbf{Z})$, where $a, b, c, d \in \mathbf{Z}$ and $\det A = ad - bc = 1$.

We define a companion matrix $\tilde{A}$ to be $\begin{pmatrix} d & c \\ b & a \end{pmatrix}$. Note that in terms of the mappings they define on $\mathbf{R}^2$, $\tilde{A}$ is merely an exchange of the x and y coordinates. More formally, let $R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then $R = R^{-1}$ is the matrix form of the permutation (12). Thus $\tilde{A} = RAR$.

We are going to consider the set $\Sigma = \{A, \tilde{A}, A^{-1}, \tilde{A}^{-1}\}$. We will use this set to define a constant degree expander. To this end we want all 4 matrices in Σ to be distinct.

Lemma 1 *$A = \tilde{A}$ iff $A = \pm I$.*
$A = A^{-1}$ iff $A = \pm I$.
$A = \tilde{A}^{-1}$ iff $b + c = 0$.

For the other $\binom{4}{2}$ possibilities, we note that $\tilde{A} = RAR$, and thus

Lemma 2 *$\tilde{A} = A^{-1}$ iff $A = \tilde{A}^{-1}$ iff $b + c = 0$.*
$\tilde{A} = \tilde{A}^{-1}$ iff $A = A^{-1}$ iff $A = \pm I$.
$A^{-1} = \tilde{A}^{-1}$ iff $A = \tilde{A}$ iff $A = \pm I$.

There is also the possibility of choosing the transpose $A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ as the companion matrix. However there are examples where Theorem 3 is not valid for this choice.

We will henceforth assume $A \neq \pm I$ and $b + c \neq 0$.

3 One less, three more

We will assume none of a, b, c, d is zero, and deal with the case where $abcd = 0$ just prior to Theorem 3.

Let $p = (x, y)$. Define the max (or ∞-) norm $||p|| = \max\{|x|, |y|\}$. The goal in this section is to show that, under a mild condition, if one of the norms

$$\{||Ap||, ||\tilde{A}p||, ||A^{-1}p||, ||\tilde{A}^{-1}p||\}$$

is strictly less than the corresponding norm $||p||$, then the three other norms are all strictly greater than $||p||$. The proof involves an examination of all the cases with reductions using suitable symmetries.

Let us start with the following Lemma:

Lemma 3 $||Ap|| < ||p|| \Longrightarrow ||\tilde{A}p|| > ||p||$.

Given $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, for a contradiction assume $||Ap|| < ||p||$ and $||\tilde{A}p|| \leq ||p||$, where $p = (x, y)$. First let's assume $|y| \geq |x|$, thus $||p|| = |y|$. We have

$$\begin{aligned} |ax + by| &< |y| \\ |cx + dy| &< |y| \\ |dx + cy| &\leq |y| \\ |bx + ay| &\leq |y| \end{aligned}$$

Let $\xi = -\frac{x}{y}$. We note that since the strict inequality $||Ap|| < ||p||$ holds, $y \neq 0$. Dividing through by y and a, b, c, d respectively, we get the rational approximations of ξ

$$\begin{aligned} |\xi - \frac{b}{a}| &< \frac{1}{|a|} \\ |\xi - \frac{d}{c}| &< \frac{1}{|c|} \\ |\xi - \frac{a}{b}| &\leq \frac{1}{|b|} \\ |\xi - \frac{c}{d}| &\leq \frac{1}{|d|} \end{aligned}$$

(We recall that none of a, b, c, d is zero as assumed.) It follows that

$$\begin{aligned} \left||\xi| - |\frac{b}{a}|\right| &< \frac{1}{|a|} \\ \left||\xi| - |\frac{d}{c}|\right| &< \frac{1}{|c|} \\ \left||\xi| - |\frac{a}{b}|\right| &\leq \frac{1}{|b|} \\ \left||\xi| - |\frac{c}{d}|\right| &\leq \frac{1}{|d|} \end{aligned}$$

Then

$$\frac{|b|-1}{|a|} < |\xi| < \frac{|b|+1}{|a|}$$
$$\frac{|a|-1}{|b|} \leq |\xi| \leq \frac{|a|+1}{|b|}$$

Thus,

$$\frac{|a|-1}{|b|} < \frac{|b|+1}{|a|}$$

and

$$\frac{|b|-1}{|a|} < \frac{|a|+1}{|b|}.$$

If $|b| < |a|$ then, being integral, we get $|b|+1 \leq |a|$ and $|b| \leq |a|-1$, and so the following contradiction follows

$$1 \leq \frac{|a|-1}{|b|} < \frac{|b|+1}{|a|} \leq 1.$$

If $|a| < |b|$ then $|a|+1 \leq |b|$, $|a| \leq |b|-1$, and the following contradiction arises

$$1 \leq \frac{|b|-1}{|a|} < \frac{|a|+1}{|b|} \leq 1.$$

Hence it follows that $|a| = |b|$. Being a row of a unimodular matrix A, the gcd of (a, b) is 1. Thus $|a| = |b| = 1$.

The exact same argument can be made for the pair (c, d). We conclude that $|c| = |d| = 1$ as well. Hence

$$a, b, c, d = 1 \pmod 2.$$

However, taken modulo 2 in $\det A = 1$, we arrive at the contradiction

$$ad - bc = 0 \pmod 2.$$

Next we consider the case $|x| \geq |y|$. This is essentially symmetric. We have

$$\begin{aligned} |ax+by| &< |x| \\ |cx+dy| &< |x| \\ |dx+cy| &\leq |x| \\ |bx+ay| &\leq |x| \end{aligned}$$

Let $\eta = -\frac{y}{x}$. Since $x \neq 0$ in this case, η is well defined. Dividing through by x and a, b, c, d respectively, we get the rational approximations of η

$$\begin{aligned} |\eta - \frac{a}{b}| &< \frac{1}{|b|} \\ |\eta - \frac{c}{d}| &< \frac{1}{|d|} \\ |\eta - \frac{b}{a}| &\leq \frac{1}{|a|} \\ |\eta - \frac{d}{c}| &\leq \frac{1}{|c|} \end{aligned}$$

Then

$$\frac{|a|-1}{|b|} < |\eta| < \frac{|a|+1}{|b|}$$
$$\frac{|b|-1}{|a|} \le |\eta| \le \frac{|b|+1}{|a|}$$

and thus

$$\frac{|b|-1}{|a|} < \frac{|a|+1}{|b|}$$
$$\frac{|a|-1}{|b|} < \frac{|b|+1}{|a|}$$

The rest is the same.

This concludes the proof of Lemma 3.

By the symmetry of $a \leftrightarrow d$ and $b \leftrightarrow c$, which effects $A \leftrightarrow \tilde{A}$ we also have the following Lemma,

Lemma 4 $||\tilde{A}p|| < ||p|| \Longrightarrow ||Ap|| > ||p||$.

We next consider the pair $(||Ap||, ||A^{-1}p||)$.

Lemma 5 *Suppose* $|\mathrm{tr}(A)| = |a+d| \ge 2$, *then*

$$||Ap|| < ||p|| \Longrightarrow ||A^{-1}p|| > ||p||.$$

Before we give the proof of this lemma, we shall discuss briefly the condition on the trace.

The elements in $SL_2(\mathbf{Z})$ with trace $|a+d| < 2$ are called elliptic elements, $|a+d| = 2$ parabolic elements, and $|a+d| > 2$ hyperbolic elements. (A final class called loxodromic elements for complex linear fractional transformations $z \mapsto \frac{az+b}{cz+d}$ do not occur here since our matrix A is real.) We note that for integral matrix A, these classes are more simply stated as

- Elliptic elements: $a+d = 0, \pm 1$.
- Parabolic elements: $|a+d| = 2$.
- Hyperbolic elements: $|a+d| > 2$.

In view of the mapping properties of these classes, it is not surprising that we needed, for the construction of expanders, the condition that the mappings be parabolic or hyperbolic, and not elliptic. Using Cayley-Hamilton Theorem, it is easy to verify that for every elliptic $A \in SL_2(\mathbf{Z})$, $A^{12} = I$. We also note that except for a vanishingly small fraction, virtually all elements are hyperbolic.

We now turn to the proof of Lemma 5.

Assume for a contradiction that

$$||Ap|| < ||p|| \text{ and yet } ||A^{-1}p|| \le ||p||.$$

First let's assume that $|y| \ge |x|$. Then we have the inequalities

$$\begin{aligned} |ax+by| &< |y| \\ |cx+dy| &< |y| \\ |dx-by| &\le |y| \\ |-cx+ay| &\le |y|. \end{aligned}$$

With the second and the fourth inequalities we get

$$|(a+d)y| \leq |cx+dy| + |-cx+ay| < 2|y|,$$

and thus

$$|a+d| < 2,$$

where we have also used the fact that $y \neq 0$ as implied by the strict inequality $||Ap|| < ||p|| = |y|$. This is a contradiction to the assumption that A is not elliptic.

The remaining case for Lemma 5 is when $|x| \geq |y|$. Then

$$\begin{array}{rcl} |ax+by| & < & |x| \\ |cx+dy| & < & |x| \\ |dx-by| & \leq & |x| \\ |-cx+ay| & \leq & |x|. \end{array}$$

This time with the first and the third inequalities we again get

$$|a+d| < 2.$$

The proof of Lemma 5 is complete.

Exactly the same argument gives us the following

Lemma 6 *Suppose* $|\mathrm{tr}(A)| = |a+d| \geq 2$, *then* $||A^{-1}p|| < ||p|| \Longrightarrow ||Ap|| > ||p||$.

We next consider the pair $(||Ap||, ||\tilde{A}^{-1}p||)$. We now require the condition $|b+c| \geq 2$. This condition is the same as requiring the trace of the permuted matrix RA to be at least 2 in absolute value: $|\mathrm{tr}(RA)| = |b+c| \geq 2$. In terms of the symmetry involved for x and y, this is quite natural.

Lemma 7 *Suppose* $|\mathrm{tr}(RA)| = |b+c| \geq 2$, *then*

$$||Ap|| < ||p|| \Longrightarrow ||\tilde{A}^{-1}p|| > ||p||.$$

For the proof of Lemma 7, again we assume for a contradiction that

$$||Ap|| < ||p|| \text{ and yet } ||\tilde{A}^{-1}p|| \leq ||p||.$$

First assume that $|y| \geq |x|$. Then

$$\begin{array}{rcl} |ax+by| & < & |y| \\ |cx+dy| & < & |y| \\ |ax-cy| & \leq & |y| \\ |-bx+dy| & \leq & |y|. \end{array}$$

With the first and the third inequalities we get

$$|(b+c)y| = |(ax+by)-(ax-cy)| \leq |ax+by| + |ax-cy| < 2|y|,$$

and thus

$$|b+c| < 2,$$

just as before.

Similarly if $|x| \geq |y|$, then we use the second and the fourth inequalities to get the same contradiction

$$|b+c| < 2.$$

This completes the proof of Lemma 7.

Exactly the same argument gives us the following

Lemma 8 *Suppose* $|\mathrm{tr}(RA)| = |b+c| \geq 2$*, then*

$$||\tilde{A}^{-1}p|| < ||p|| \Longrightarrow ||Ap|| > ||p||.$$

Combining the 6 Lemmata above (Lemma 3 to Lemma 8), we conclude that under the condition $|\mathrm{tr}(A)| = |a+d| \geq 2$ and $|\mathrm{tr}(RA)| = |b+c| \geq 2$, for each of the 3 pairs

$$(||Ap||, ||\tilde{A}p||), (||Ap||, ||A^{-1}p||), (||Ap||, ||\tilde{A}^{-1}p||),$$

involving $||Ap||$ from the following set

$$\{||Ap||, ||\tilde{A}p||, ||A^{-1}p||, ||\tilde{A}^{-1}p||\}$$

there can be at most one of the entry to be strictly less than $||p||$, and in that case the other entry of the pair is strictly greater than $||p||$.

This is not quite enough for the goal of this section as stated, which include the remaining 3 pairs not involving $||Ap||$ (and corresponding 6 Lemmata above). However we will handle the remaining proof by symmetry.

For the pair $(||\tilde{A}p||, ||A^{-1}p||)$ we apply the symmetry $a \leftrightarrow d$, $b \leftrightarrow c$, thus $A \leftrightarrow \tilde{A}$. This reduces the pair $(||\tilde{A}p||, ||A^{-1}p||)$ to the pair $(||Ap||, ||\tilde{A}^{-1}p||)$ and Lemma 7, Lemma 8 give us respectively

Lemma 9 *Suppose* $|\mathrm{tr}(RA)| = |b+c| \geq 2$*, then* $||\tilde{A}p|| < ||p|| \Longrightarrow ||A^{-1}p|| > ||p||$.

and

Lemma 10 *Suppose* $|\mathrm{tr}(RA)| = |b+c| \geq 2$*, then* $||A^{-1}p|| < ||p|| \Longrightarrow ||\tilde{A}p|| > ||p||$.

For the pair $(||\tilde{A}p||, ||\tilde{A}^{-1}p||)$ we apply the symmetry $b \leftrightarrow -c$, (and $c \leftrightarrow -b$, $a \leftrightarrow a$, and $d \leftrightarrow d$), thus, $A \leftrightarrow \tilde{A}^{-1}$ and $\tilde{A} \leftrightarrow A^{-1}$. Thus this reduces the pair $(||\tilde{A}p||, ||\tilde{A}^{-1}p||)$ to the pair $(||A^{-1}p||, ||Ap||)$. Now Lemma 6, Lemma 5 give us respectively

Lemma 11 *Suppose* $|\mathrm{tr}(A)| = |a+d| \geq 2$*, then* $||\tilde{A}p|| < ||p|| \Longrightarrow ||\tilde{A}^{-1}p|| > ||p||$.

and

Lemma 12 *Suppose* $|\mathrm{tr}(A)| = |a+d| \geq 2$*, then* $||\tilde{A}^{-1}p|| < ||p|| \Longrightarrow ||\tilde{A}p|| > ||p||$.

Finally for the pair $(||A^{-1}p||, ||\tilde{A}^{-1}p||)$ we apply the same symmetry $b \leftrightarrow -c$ as above, which transforms it to the pair $(||\tilde{A}p||, ||Ap||)$. Then we apply Lemma 4, Lemma 3 respectively,

Lemma 13 $||A^{-1}p|| < ||p|| \Longrightarrow ||\tilde{A}^{-1}p|| > ||p||$.

and

Lemma 14 $||\tilde{A}^{-1}p|| < ||p|| \Longrightarrow ||A^{-1}p|| > ||p||$.

Combining Lemma 3 to Lemma 14 we have

Theorem 1 *For any $A \in SL_2(\mathbf{Z})$, where $abcd \neq 0$ and A, RA not elliptic, then if any one of the following 4 entries*

$$\{||Ap||, ||\tilde{A}p||, ||A^{-1}p||, ||\tilde{A}^{-1}p||\}$$

is strictly less than the corresponding norm $||p||$, then the three other norms are all strictly greater than $||p||$.

We note that the condition that none of a, b, c, d is zero is only technical, and will be handled later. Only the conditions on the trace are real restrictions.

4 At most two equalities

As shown in Section 3 if there is any one among

$$\{||Ap||, ||\tilde{A}p||, ||A^{-1}p||, ||\tilde{A}^{-1}p||\}$$

to be strictly less than $||p||$, then the three other norms are all strictly greater than $||p||$. In particular there are no equalities in this case. Suppose now, for this section, that there are no one among the four to be strictly less than $||p||$, i.e.,

$$\begin{aligned} ||Ap|| &\geq ||p|| \\ ||\tilde{A}p|| &\geq ||p|| \\ ||A^{-1}p|| &\geq ||p|| \\ ||\tilde{A}^{-1}p|| &\geq ||p|| \end{aligned}$$

We count the number of equalities among these 4. The goal in this section is to show that, for $p \neq 0$, there can be at most two among the four to be equalities. It follows that the other terms, at least 2 among 4, are all strictly greater than $||p||$. Clearly the condition that $p \neq 0$ is necessary for handling the equalities.

We prove this by contradiction. Suppose there are at least three among the four are equalities. Then there are the following *two alternatives. EITHER*

$$\begin{aligned} ||Ap|| &= ||p|| \\ ||\tilde{A}p|| &= ||p|| \end{aligned}$$

both hold and at least one of the following holds

$$\begin{aligned} ||A^{-1}p|| &= ||p|| \\ ||\tilde{A}^{-1}p|| &= ||p|| \end{aligned}$$

OR vice versa.

Without loss of generality (wolog) we also assume that $|y| \geq |x|$. We note that the symmetry $x \leftrightarrow y$ exchanges and permutes the equalities

$$\begin{aligned} ||Ap|| = ||p|| &\leftrightarrow ||\tilde{A}p|| = ||p|| \\ ||A^{-1}p|| = ||p|| &\leftrightarrow ||\tilde{A}^{-1}p|| = ||p|| \end{aligned}$$

respectively, and thus the assumption $|y| \geq |x|$ is indeed without loss of generality.

We will assume the first alternative. Since $p \neq 0$, $y \neq 0$. Since there are no strict inequalities $<$ by assumption, the first alternative leads to

$$\begin{aligned} |ax+by| &\leq |y| \\ |cx+dy| &\leq |y| \\ |dx+cy| &\leq |y| \\ |bx+ay| &\leq |y| \end{aligned}$$

and at least one of the following holds

$$\begin{aligned} |dx-by| &\leq |y| \\ |-cx+ay| &\leq |y| \end{aligned}$$

or

$$\begin{aligned} |ax-cy| &\leq |y| \\ |-bx+dy| &\leq |y|. \end{aligned}$$

As in the proof of Lemma 3, denoting $\xi = -\frac{x}{y}$, and dividing through by y and a, b, c, d respectively, we get the rational approximations of ξ

$$|\xi - \frac{b}{a}| \leq \frac{1}{|a|} \quad (1)$$

$$|\xi - \frac{d}{c}| \leq \frac{1}{|c|} \quad (2)$$

$$|\xi - \frac{a}{b}| \leq \frac{1}{|b|} \quad (3)$$

$$|\xi - \frac{c}{d}| \leq \frac{1}{|d|} \quad (4)$$

Lemma 15 *Either $|a| \neq |b|$ or $|c| \neq |d|$.*

To prove this Lemma, we assume instead both equalities hold $|a| = |b|$ and $|c| = |d|$. Since they form the rows of a unimodular matrix, the gcd of both (a, b) and (c, d) are 1. Thus

$$|a| = |b| = |c| = |d| = 1,$$

and taken modulo 2

$$a = b = c = d = 1 \pmod 2.$$

However this leads to

$$\det(A) = ad - bc = 0 \pmod 2$$

which contradicts the unimodularity again. Lemma 15 is proved.

Hence we have two possibilities:

1. $|a| \neq |b|$

   Suppose $ab > 0$, i.e., they are of the same sign, then $\frac{b}{a} = \frac{|b|}{|a|}$, and

$$\frac{|b|-1}{|a|} \leq \xi \leq \frac{|b|+1}{|a|},$$

and also

$$\frac{|a|-1}{|b|} \leq \xi \leq \frac{|a|+1}{|b|}.$$

Note that these two bounds on ξ are symmetric for a and b. Thus, without loss of generality $|a| > |b|$. Then, by being integral, $|a| \geq |b|+1$, it follows that

$$1 \leq \frac{|a|-1}{|b|} \leq \xi \leq \frac{|b|+1}{|a|} \leq 1,$$

which means that these inequalities are in fact all equalities, and $\xi = 1$. By definition of ξ, $x = -y$. This is true regardless $|a| > |b|$ or $|a| < |b|$, as long as $ab > 0$.

The case where a and b are of opposite signs, i.e., $ab < 0$, is handled similarly with $\frac{b}{a} = -\frac{|b|}{|a|}$, and the corresponding rational approximations of $-\xi$. So we obtain $-\xi = 1$. Hence $x = y$.

We conclude in this case that $|x| = |y|$.

2. $|c| \neq |d|$

   This case is handled by the symmetry $a \leftrightarrow d$ and $b \leftrightarrow c$. Note that the rational approximations in Eqn. (1) to Eqn. (4) is invariant under this substitution. Hence we also get $|x| = |y|$.

We now proceed to deal with the possibility $|x| = |y|$, which is $\neq 0$, under the assumption that

$$\begin{aligned} |ax+by| &\leq |y| \\ |cx+dy| &\leq |y| \\ |dx+cy| &\leq |y| \\ |bx+ay| &\leq |y| \end{aligned}$$

and at least one of the following holds

$$\begin{aligned} |dx-by| &\leq |y| \\ |-cx+ay| &\leq |y| \end{aligned}$$

or

$$\begin{aligned} |ax-cy| &\leq |y| \\ |-bx+dy| &\leq |y|. \end{aligned}$$

1. $x = -y$

   Dividing through by $|y|$ we have

$$\begin{aligned} |a-b| &\leq 1 \\ |c-d| &\leq 1 \end{aligned}$$

   and

$$\begin{aligned} |d+b| &\leq 1 \\ |c+a| &\leq 1. \end{aligned}$$

From these we obtain

$$\begin{aligned} |a+d| &\leq 2 \\ |b+c| &\leq 2. \end{aligned}$$

By our condition on the trace of A and RA, i.e., they are not elliptic, we get

$$|a+d| = |b+c| = 2.$$

Hence we get

$$\begin{aligned} |a-b| &= 1 \\ |c-d| &= 1 \\ |d+b| &= 1 \\ |c+a| &= 1 \end{aligned}$$

Thus we can write

$$\begin{pmatrix} b & b \\ c & c \end{pmatrix} = \begin{pmatrix} a & -d \\ d & -a \end{pmatrix} + \mathcal{E}, \tag{5}$$

where we let

$$\mathcal{E} = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{21} & \epsilon_{22} \end{pmatrix},$$

and $\epsilon_{ij} = \pm 1$ for $i, j = 1, 2$.

In $\mathcal{E}$ the top row cannot be of the same sign, otherwise $a + d = 0$. Similarly the bottom row cannot be of the same sign, otherwise $a + d = 0$ as well.

Furthermore, we observe that

$$a + d + (\epsilon_{11} - \epsilon_{12}) = 0$$

and

$$a + d + (\epsilon_{21} - \epsilon_{22}) = 0.$$

Thus the trace $a + d = -2$ iff

$$\mathcal{E} = \begin{pmatrix} +1 & -1 \\ +1 & -1 \end{pmatrix},$$

and the trace $a + d = +2$ iff

$$\mathcal{E} = \begin{pmatrix} -1 & +1 \\ -1 & +1 \end{pmatrix}.$$

However in either cases we obtain

$$b + c = 0,$$

by adding the diagonal entries in the matrix equation Eqn.(5).

So under $|a + d| \geq 2, |b + c| \geq 2$ we conclude that $x = -y$ is impossible.

2. $x = y$

   This case is handled by the symmetry $b \leftrightarrow -b$ and $c \leftrightarrow -c$ in the above argument for $x = -y$. Thus $x = y$ is also impossible.

Finally we consider the second alternative: $|y| \geq |x|$ and,

$$\begin{aligned} |dx - by| &\leq |y| \\ |-cx + ay| &\leq |y| \\ |ax - cy| &\leq |y| \\ |-bx + dy| &\leq |y| \end{aligned}$$

and at least one of the following holds

$$\begin{aligned} |ax + by| &\leq |y| \\ |cx + dy| &\leq |y| \end{aligned}$$

or

$$\begin{aligned} |dx + cy| &\leq |y| \\ |bx + ay| &\leq |y|. \end{aligned}$$

Use $\eta = -\xi = \frac{x}{y}$, and the symmetry $a \leftrightarrow d$, and $b \leftrightarrow -b$, $c \leftrightarrow -c$, we conclude that the second alternative is also impossible.

Theorem 2 *For any $A \in SL_2(\mathbf{Z})$, where $abcd \neq 0$ and A, RA not elliptic, then for $p \neq 0$, among*

$$\{||Ap||, ||\tilde{A}p||, ||A^{-1}p||, ||\tilde{A}^{-1}p||\}$$

there cannot be more than two of them equal to $||p||$.

We now briefly handle the case with $abcd = 0$. Suppose $a = 0$. Then $bc = -1$ by unimodularity. Being both integral, $b = -c = \pm 1$. Then $b + c = 0$. This is excluded.

By the symmetry $a \leftrightarrow d$ and $b \leftrightarrow c$, which effects $A \leftrightarrow \tilde{A}$, we see that $d = 0$ is the same.

Suppose $b = 0$, then $ad = 1$ and being integral, $a = d = \pm 1$. Thus the matrix we are dealing with is $A = \pm \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$.

The case of $c = \pm 1$ with $b = 0$ is *the* matrix dealt with by Gabber and Galil [20]. (They show that $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ does define an expander with a smaller expansion constant. However the condition in Theorem 3 of RA being non-elliptic technically excludes this case.) It is not difficult to see that the mapping properties stated in Theorem 3 are valid for $|c| \geq 2$ and $b = 0$. This is quite clear if we consider the mapping on the set of lattice points with $||(x, y)||_\infty = r$ for any $r \geq 1$. ([20] also contained a discussion of $c = \pm 2$.) By symmetry, the same is true for the case $|b| \geq 2$ and $c = 0$.

Combining Theorem 1, Theorem 2, and the above comments regarding $abcd \neq 0$, we have

Theorem 3 *For any $A \in SL_2(\mathbf{Z})$, where A, RA not elliptic, and $p \neq 0$, then among*

$$\{||Ap||, ||\tilde{A}p||, ||A^{-1}p||, ||\tilde{A}^{-1}p||\},$$

- *Either one is less than $||p||$ and three others are greater than $||p||$,*
- *Or no more than two are equal to $||p||$ and the rest are all greater than $||p||$.*

5 Analytic proof of expansion

In this section we prove some explicit estimates using Fourier analysis. We will follow [20] and adapt their proof for special matrices to general matrices.

Let $B = A$ or $\tilde{A}$ and let $U = [0,1)^2$. B defines a measure preserving automorphism $\beta = \beta_B$ of U as follows:

$$\beta : (x,y) \mapsto (x,y)B \bmod 1.$$

We will denote $\alpha = \beta_A$ and $\tilde{\alpha} = \beta_{\tilde{A}}$. It is easy to check that β is a bijection on U with inverse map $\beta^{-1}(x,y) = (x,y)B^{-1} \bmod 1$. That it is measure preserving follows from the fact that the Jacobi of the map is $\det B = 1$.

For any function ϕ on U, we can define the function

$$B^*(\phi)(x,y) = \phi(\beta^{-1}(x,y)).$$

We will restrict our discussion to square integrable functions ϕ. For such ϕ the Fourier coefficients are defined as follows

$$a_{\binom{m}{n}}(\phi) = \int_U \phi(x,y)e^{-2\pi i(mx+ny)}d\mu(x,y),$$

where $m,n \in \mathbf{Z}$. The next lemma relates the Fourier coefficients of ϕ with that of $B^*(\phi)$.

Lemma 16

$$a_{\binom{m}{n}}(B^*(\phi)) = a_{B\binom{m}{n}}(\phi).$$

Proof:

$$\begin{aligned} a_{\binom{m}{n}}(B^*(\phi)) &= \int_U \phi(\beta^{-1}(x,y))e^{-2\pi i(x,y)\cdot\binom{m}{n}}d\mu(x,y) \\ &= \int_U \phi(\beta^{-1}(x,y))e^{-2\pi i(x,y)B^{-1}B\binom{m}{n}}d\mu(x,y) \end{aligned}$$

We can replace $(x,y)B^{-1}$ by $\beta^{-1}(x,y)$ in the exponent since the function $\exp[-2\pi iX]$ has integral period 1. Hence, by a substitution of variables $(x',y') = \beta^{-1}(x,y)$, and note that the Jacobi is 1, we get

$$\begin{aligned} a_{\binom{m}{n}}(B^*(\phi)) &= \int_U \phi(x',y')e^{-2\pi i(x',y')B\binom{m}{n}}d\mu(x',y') \\ &= a_{B\binom{m}{n}}(\phi). \end{aligned}$$

Our goal is to obtain a non-trivial estimate for

$$\sum_q \left[|a_{Aq} - a_q|^2 + |a_{\tilde{A}q} - a_q|^2\right],$$

where q ranges over $\mathbf{Z}^2$, and $\{a_q\}$ is square summable $\sum_q |a_q|^2 < \infty$. Note that A and $\tilde{A}$ define permutations on $\mathbf{Z}^2 - \{0\}$ while $A0 = \tilde{A}0 = 0$. Thus the above sum can also range over $\mathbf{Z}^2 - \{0\}$.

Let f, g be any complex square summable functions on $\mathbf{Z}^2 - \{0\}$. The inner product is defined as

$$\langle f, g\rangle = \sum_{q\neq 0} f(q)\cdot\overline{g(q)},$$

and the norm is

$$||f|| = \langle f, f\rangle^{1/2} = \sum_{q\neq 0} |f(q)|^2.$$

It follows that

$$||f - f\circ A||^2 + ||f - f\circ \tilde{A}||^2 = 4||f||^2 - C,$$

where the *cross terms*

$$C = \langle f, f\circ A\rangle + \langle f\circ A, f\rangle + \langle f, f\circ \tilde{A}\rangle + \langle f\circ \tilde{A}, f\rangle,$$

thus $|C| \leq 2\ [\langle |f|, |f\circ A|\rangle + \langle |f|, |f\circ \tilde{A}|\rangle]$.

Lemma 17

$$||f - f\circ A||^2 + ||f - f\circ \tilde{A}||^2 \geq (4 - 2\sqrt{3})\ ||f||^2.$$

Proof: We only need to show an upper bound $|C| \leq 2\sqrt{3}\ ||f||^2$. Define

$$\lambda(p,q) = \begin{cases} \sqrt{3} & \text{if } ||q|| < ||p|| \\ 1 & \text{if } ||q|| = ||p|| \\ 1/\sqrt{3} & \text{if } ||q|| > ||p|| \end{cases}$$

By Cauchy-Schwarz, $2|XY| \leq \lambda |X|^2 + \frac{1}{\lambda}|Y|^2$. Note that $\lambda(p,q) = \lambda(q,p)^{-1}$, and thus for $\sigma = A$ or $\tilde{A}$,

$$\begin{aligned} 2\sum_{q\neq 0} |f(q)||f(\sigma(q))| &\leq \sum_{q\neq 0}\left[\lambda(q,\sigma(q))|f(q)|^2 + \lambda(\sigma(q),q)|f(\sigma(q))|^2\right] \\ &= \sum_{q\neq 0} |f(q)|^2\left[\lambda(q,\sigma(q)) + \lambda(q,\sigma^{-1}(q))\right]. \end{aligned}$$

Hence

$$|C| \leq \sum_{q\neq 0} |f(q)|^2 \left[\sum_{\sigma\in\Sigma} \lambda(q,\sigma(q))\right].$$

(Recall that $\Sigma = \{A, \tilde{A}, A^{-1}, \tilde{A}^{-1}\}$.) By Theorem 3, the sum of four terms $\sum_{\sigma\in\Sigma}\lambda(q,\sigma(q)) \leq 2\sqrt{3}$ in all cases (being either $\leq \sqrt{3} + 3/\sqrt{3}$, or $\leq 4/\sqrt{3}$, or $\leq 1 + 3/\sqrt{3}$, or $\leq 2 + 2/\sqrt{3}$.) It follows that $|C| \leq 2\sqrt{3}\ ||f||^2$.

Stated for $\{a_q\}$ we have

Lemma 18 *If $a_0 = 0$ and $\sum_{q\neq 0} |a_q|^2 < \infty$, then*

$$\sum_q \left[|a_{Aq} - a_q|^2 + |a_{\tilde{A}q} - a_q|^2\right] \geq (4 - 2\sqrt{3}) \sum_{q\neq 0} |a_q|^2.$$

We next translate this lemma to integrals via Parseval equality.

Lemma 19 *For square integrable function ϕ on U with $\int_U \phi = 0$,*

$$\int_U |A^*(\phi) - \phi|^2 + \int_U |\tilde{A}^*(\phi) - \phi|^2 \geq (4 - 2\sqrt{3})\ \int_U |\phi|^2.$$

Proof: By Parseval equality, for square integrable ψ,

$$\int_U |\psi|^2 = \sum_q |a_q(\psi)|^2,$$

where $a_q(\psi)$ are the Fourier coefficients. Note that $a_0(\phi) = \int_U \phi = 0$. By linearity and Lemma 16, $a_q(A^*(\phi) - \phi) = a_q(A^*(\phi)) - a_q(\phi) = a_{Aq}(\phi) - a_q(\phi)$. Lemma 19 follows from Lemma 18.

Recall the definition of $\beta = \beta_B$ for $B \in \Sigma$, $\beta_B(\xi) = \xi B \bmod 1$.

Lemma 20 *For measurable set $Z \subseteq U$,*

$$\sum_{B=A,\tilde{A}} \mu[Z - \beta_B^{-1}(Z)] \geq (2 - \sqrt{3})\ \mu(Z)\mu(Z^c).$$

Proof: Define $\phi = \chi_Z - \mu(Z) = \begin{cases} \mu(Z^c) & \text{on } Z \\ -\mu(Z) & \text{on } Z^c \end{cases}$. Then $\int_U \phi = 0$, and

$$\int_U |\phi|^2 = \mu(Z)\mu(Z^c) < \infty.$$

Let $\xi \in U$, and denote β_A by α, i.e., $\alpha(\xi) = \xi A \bmod 1$. We observe that

$$\begin{aligned} A^*(\phi)(\xi) &= \phi(\alpha^{-1}(\xi)) \\ &= \begin{cases} \mu(Z^c) & \text{for } \xi \in \alpha(Z) \\ -\mu(Z) & \text{for } \xi \notin \alpha(Z) \end{cases} \\ &= \chi_{\alpha(Z)} - \mu(Z) \end{aligned}$$

It follows that

$$A^*(\phi) - \phi = \chi_{\alpha(Z)} - \chi_Z.$$

Hence for $\int_U |A^*(\phi) - \phi|^2$, the integrand is 1 on the symmetric difference $\alpha(Z)\Delta Z$, and 0 elsewhere. So

$$\int_U |A^*(\phi) - \phi|^2 = \mu[\alpha(Z)\Delta Z].$$

However, $\alpha(Z)\Delta Z = [\alpha(Z) - Z] \cup [Z - \alpha(Z)]$. Since α is bijective and measure preserving,

$$\begin{aligned} \mu[Z - \alpha(Z)] &= \mu[Z] - \mu[Z \cap \alpha(Z)] \\ &= \mu[\alpha(Z)] - \mu[Z \cap \alpha(Z)] \\ &= \mu[\alpha(Z) - Z] \\ &= \mu[Z - \alpha^{-1}(Z)] \end{aligned}$$

Thus

$$\int_U |A^*(\phi) - \phi|^2 = 2\mu[Z - \alpha^{-1}(Z)].$$

Similarly, denote $\tilde{\alpha} = \beta_{\tilde{A}}$, we have

$$\int_U |\tilde{A}^*(\phi) - \phi|^2 = 2\mu[Z - \tilde{\alpha}^{-1}(Z)].$$

Then by Lemma 19,

$$\sum_{B=A,\tilde{A}} \mu[Z - \beta_B^{-1}(Z)] = \frac{1}{2} \sum_{B=A,\tilde{A}} \int_U |B^*(\phi) - \phi|^2 \geq (2 - \sqrt{3})\ \int_U |\phi|^2 = (2 - \sqrt{3})\ \mu(Z)\mu(Z^c).$$

6 The graph

In this section we prove an explicit expansion constant for a family of bipartite graphs, constructed from every matrix A considered in Theorem 3.

We will first define the family of graphs. Denote the unit square by $U = [0,1)^2$. For $p = (i,j) \in \mathbf{Z}^2$, the translated square by p is denoted by $U_p = p + U$. We define a set of "neighborhood" points as follows: For $B = A, \tilde{A}$,

$$N_B = \{q \in \mathbf{Z}^2 \mid \mu[UB \cap U_q] \neq 0\},$$

where μ denotes the Lebesgue measure, and $UB = \{\xi B \mid \xi \in U\}$ is the image of U under B.

For $k \geq 1$, let the mod k "neighborhood" be $N_{B,k} = N_B \bmod k$. Note that the cardinality of $N_{B,k}$ is at most that of N_B for every k. In particular since $|N_B|$ is independent of k, $|N_{B,k}|$ is bounded in k. For any measurable set $V \subseteq \mathbf{R}^2$, denote its mod k fold in the torus $(\mathbf{R}/k\mathbf{Z})^2$ by $(V)_k = V \bmod k$. We claim that

$$N_{B,k} = \{q \in (\mathbf{Z}/k\mathbf{Z})^2 \mid \mu_k[(UB)_k \cap (U_q)_k] \neq 0\}.$$

where μ_k is the Lebesgue measure on the torus $(\mathbf{R}/k\mathbf{Z})^2$. This is fairly obvious. To carry out the detail, let $q \in N_{B,k}$. Then there exists an integral vector v such that $q + kv \in N_B$. Thus $\mu[UB \cap U_{q+kv}] \neq 0$. Note that $(U_{q+kv})_k = (U_q)_k$, it follows that $\mu_k[(UB)_k \cap (U_q)_k] \neq 0$. Conversely, if the above holds for q, then there exist integral vectors v and v' such that $\mu[\,[UB + kv] \cap [U + q + kv']\,] \neq 0$. So $\mu[UB \cap U_{q'}] \neq 0$, for $q' = q + k(v' - v)$. Hence $q' \in N_B$, and $q \equiv q' \bmod k$.

We now define the family of graphs. For every $k \geq 1$, the bipartite graph $G_k = (L, R, E)$ has $n = k^2$ vertices on both sides, $L = R = (\mathbf{Z}/k\mathbf{Z})^2$. The vertex $p \in L$ is connected to $p \in R$ and every $p' = pB + q \bmod k$, for $q \in N_{B,k}$, $B = A, \tilde{A}$. Thus, the maximum degree of G_k is bounded, being at most $d = 1 + |N_A| + |N_{\tilde{A}}|$. We note that the neighbors $p' = pB + q \bmod k$ of p are precisely those satisfying

$$\mu_k[(U_pB)_k \cap (U_{p'})_k] \neq 0.$$

We will denote by $\sigma_0 = \mathrm{id}$, σ_ℓ the permutations $p \mapsto pA + q \bmod k$, for $1 \leq \ell \leq |N_A|$, and $\tilde{\sigma}_\ell$ the permutations $p \mapsto p\tilde{A} + q \bmod k$, for $1 \leq \ell \leq |N_{\tilde{A}}|$. Thus for $p \in L$ the neighbor set of p in R is $\Gamma(p) = \{p, \sigma_i(p), \tilde{\sigma}_j(p) \mid 1 \leq i \leq |N_A|, 1 \leq j \leq |N_{\tilde{A}}|\}$. (If these neighbor points are not distinct, no multiple edges are drawn. Thus, in fact the degree is at most $d_k = 1 + |N_{A,k}| + |N_{\tilde{A},k}|$.)

The next Lemma discretizes Lemma 20.

Lemma 21 *Let $X \subseteq L$. There exists some $\tau = \sigma_\ell$ or $\tilde{\sigma}_\ell$, for some $\ell \geq 1$, such that*

$$|\tau(X) - X| \geq (1 - \sqrt{3}/2)|X||X^c|/n,$$

where $n = k^2$.

Proof: For X, define a subset of the torus $(\mathbf{R}/k\mathbf{Z})^2$ by $Y = \bigcup_{p \in X} U_p$. Thus each point $p = (i,j) \in X$ is replaced by the translated square U_p. Clearly $\mu_k(Y) = |X|$ and $\mu_k(Y^c) = |X^c|$. If we shrink Y by a factor of k, we may consider $Z = \frac{1}{k}Y \subseteq U$, in which we can identify U with the unit torus $(\mathbf{R}/\mathbf{Z})^2$. Clearly $\mu(Z) = \frac{|X|}{n}$ and $\mu(Z^c) = \frac{|X^c|}{n}$.

We next consider where does the small square $\frac{1}{k}U_p$ get mapped to under α; more specifically, which $\frac{1}{k}U_q$ contains images of $\frac{1}{k}U_p$ with non-zero measure. For $\xi = [(i,j) + (u,v)]/k$,

$$\alpha(\xi) = \xi A \bmod 1 = \frac{(i+u, j+v)A \bmod k}{k}.$$

So $\alpha(\xi) \in \frac{1}{k}U_q$ iff $(i+u, j+v)A \bmod k \in U_q$. Hence $\mu[\alpha(\frac{1}{k}U_p) \cap \frac{1}{k}U_q] \neq 0$ iff $\mu_k[(U_pA)_k \cap (U_q)_k] \neq 0$. Thus, q is a neighbor $\sigma_\ell(p)$ of p in the graph G_k for some $1 \leq \ell \leq |N_A|$. Similarly for $\tilde{\alpha}(\xi)$.

Let $w_\ell = \mu[\alpha(\frac{1}{k}U_p) \cap \frac{1}{k}U_{\sigma_\ell(p)}] > 0$ be the weight of intersection. (To be precise, the weight w_ℓ corresponds to the neighbor $p \mapsto pA + q$ *before* taking modulo k. Note that these weights correspond to disjoint slices of $\alpha(\frac{1}{k}U_p)$ even after taking modulo k.) Since α is measure preserving, $\sum_{1\leq\ell\leq|N_A|} w_\ell = 1/n$. Similarly one can define $\tilde{w}_\ell$ for $\tilde{A}$ and they also sum to $1/n$.

By definition, $Z = \bigcup_{p\in X} \frac{1}{k}U_p$. Within each $\frac{1}{k}U_p$, divide it according to

$$\left[\frac{1}{k}U_p\right] \cap \alpha^{-1}\left(\frac{1}{k}U_{\sigma_\ell(p)}\right),$$

each with weight w_ℓ.

$\xi \in Z - \alpha^{-1}(Z)$ iff $[\xi \in Z \ \& \ \alpha(\xi) \notin Z]$. For $\xi \in Z$, $\xi \in \frac{1}{k}U_p$ for a unique $p \in X$, and within $\frac{1}{k}U_p$ those $\xi \in (\frac{1}{k}U_p) \cap \alpha^{-1}(\frac{1}{k}U_{\sigma_\ell(p)})$ are mapped to $\frac{1}{k}U_{\sigma_\ell(p)}$. For those ξ, $\alpha(\xi) \notin Z$ iff $\sigma_\ell(p) \notin X$. It follows that

$$\begin{aligned}\mu[Z - \alpha^{-1}(Z)] &= \sum_{p\in X} \sum_{1\leq\ell\leq|N_A|} w_\ell \mathbf{1}_{[\sigma_\ell(p)\notin X]} \\ &= \sum_{1\leq\ell\leq|N_A|} w_\ell \sum_{p\in L} \mathbf{1}_{[p\in X \text{ and } \sigma_\ell(p)\notin X]} \\ &= \sum_{1\leq\ell\leq|N_A|} w_\ell |X - \sigma_\ell^{-1}(X)|.\end{aligned}$$

Similarly

$$\mu[Z - \tilde{\alpha}^{-1}(Z)] = \sum_{1\leq\ell\leq|N_{\tilde{A}}|} \tilde{w}_\ell |X - \tilde{\sigma}_\ell^{-1}(X)|.$$

By Lemma 20,

$$\mu[Z - \alpha^{-1}(Z)] + \mu[Z - \tilde{\alpha}^{-1}(Z)] \geq (2 - \sqrt{3})\frac{|X|}{n}\frac{|X^c|}{n}.$$

Hence,

$$\sum_{1\leq\ell\leq|N_A|} w_\ell |X - \sigma_\ell^{-1}(X)| + \sum_{1\leq\ell\leq|N_{\tilde{A}}|} \tilde{w}_\ell |X - \tilde{\sigma}_\ell^{-1}(X)| \geq (2 - \sqrt{3})|X||X^c|/n^2.$$

It follows that there exists ℓ_0, such that either

$$|X - \sigma_{\ell_0}^{-1}(X)| \geq (1 - \sqrt{3}/2)|X||X^c|/n,$$

or

$$|X - \tilde{\sigma}_{\ell_0}^{-1}(X)| \geq (1 - \sqrt{3}/2)|X||X^c|/n,$$

as $\sum_\ell w_\ell = \sum_\ell \tilde{w}_\ell = 1/n$.

In either cases, since $\tau = \sigma_{\ell_0}$ or $\tilde{\sigma}_{\ell_0}$ is a permutation, $|X - \tau^{-1}(X)| = |\tau(X) - X|$, and thus

$$|\tau(X) - X| \geq (1 - \sqrt{3}/2)|X||X^c|/n.$$

Lemma 21 is proved.

Now the neighbor set $\Gamma(X) \supseteq X \cup \tau(X)$, it follows that

$$\begin{aligned}|\Gamma(X)| &= |X| + |\Gamma(X) - X| \\ &\geq |X| + |\tau(X) - X| \\ &\geq \left[1 + (1 - \sqrt{3}/2)\left(1 - \frac{|X|}{n}\right)\right]|X|.\end{aligned}$$

Acknowledgements

I thank Pavan Aduri, Venkat Chakaravarthy, Bernard Chazelle, Charles Denis, Avi Wigderson and especially Samit Sengupta for valuable discussions, comments and feedbacks.

References

[1] M. Ajtai. Recursive construction for Γ-regular expanders. Combinatorica, 14(4):379-416, 1994.

[2] M. Ajtai, J. Komlos and E. Szemeredi, Sorting in $c \log n$ parallel steps, *Combinatorica*, **3**, (1983) 1–19.

[3] M. Ajtai, J. Komlos and E. Szemeredi, Deterministic simulation in LOGSPACE, Proc. of the *19th ACM STOC*, 132–140, 1987.

[4] M. Ajtai, J. Komlós and E. Szemerédi, Generating expanders from two permutations. In *A tribute to Paul Erdös*, edited by A. Baker, B. Bollobás & A. Hajnal. pp. 1–12. Cambridge University Press, 1990.

[5] N. Alon, Eigenvalues, geometric expanders, sorting in rounds and Ramsey Theory, *Combinatorica* **6**, 207–219.

[6] N. Alon, Eigenvalues and Expanders, *Combinatorica* **6** (2), 83–96. 1986.

[7] N. Alon and V. D. Milman, Eigenvalues, expanders and superconcentrators. Proc of *the 25th ACM STOC*, 320–322. 1984.

[8] N. Alon and V. D. Milman. Ω isoperimetric inequali ties for graphs, and superconcentrators. J. Combin. Theory Ser. B, 38(1):73-88, 1985.

[9] N. Alon, Z. Galil, and V. D. Milman. Better expanders an d superconcentrators. J. Algorithms, 8(3):337-347, 1987.

[10] N. Alon, O. Goldreich, J. Hasted and R. Peralta, Simple construction of almost k-wise independent random variables, The *31st FOCS*, 544–553.

[11] N. Alon and J. Spencer, with an appendix by P. Erdös, The Probabilistic Method. John Wiley and Sons, Inc.1992.

[12] D. Angulin, A note on a construction of Margulis, *Information Processing Letters*, **8**, pp 17–19, (1979).

[13] M. Blum. Independent unbiased coin flips from a correlat ed biased source–a finite state Markov chain. Combinatorica, 6(2):97-108, 1986.

[14] A. Broder and E. Shamir. On the second eigenvalue of random regular graphs. FOCS 1987.

[15] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. SIAM Journal on Computing, 17(2):230-261, 1988.

[16] F. R. K. Chung, On Concentrators, superconcentrators, generalized, and non-blocking networks, *Bell Sys. Tech J.* **58**, pp 1765–1777, (1978).

[17] A. Cohen and A. Wigderson. Dispersers, deterministic amplification, and weak random sources. In 30th Annual Symposium on Foundations of Computer Science, pages 14-19, 1989.

[18] J. Friedman. On the second eigenvalue and random walks in random regular graphs. Combinatorica, 11(4):331-362, 1991.

[19] J. Friedman, J. Kahn, and E. Szemer'edi. On the second eigenvalue in random regular graphs. STOC 1989.

[20] O. Gabber and Z. Galil, Explicit construction of linear size superconcentrators. *JCSS* **22**, pp 407–420 (1981).

[21] O. Goldreich, R. Impagliazzo, L. Levin, R. Venkatesan, and D. Zuckerman. Security preserving amplification of hardness. FOCS 1990.

[22] O. Goldreich and A. Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. Random Structures and Algorithms, 11(4):315-343, 1997.

[23] J. Hastad, R. Impagliazzo, L. Levin, and M. Luby. A pseudorandom generator from any one-way function. SIAM J. Comput., 28(4):1364-1396, 1999.

[24] R. Impagliazzo, N. Nisan, and A. Wigderson. Pseudorandomness for network algorithms. STOC 1994.

[25] R. Impagliazzo, R. Shaltiel, and A. Wigderson. Extractors and pseudo-random generators with optimal seed length. STOC 2000.

[26] R. Impagliazzo and D. Zuckerman. How to recycle random bits. FOCS 1989.

[27] S. Jimbo and A. Maruoka, Expanders obtained from affine transformations. *Combinatorica*, 7 (4): 343-355, 1987.

[28] A. Lubotsky, R. Phillip and P. Sarnak, Explicit expanders and the Ramanujan conjectures. Proceedings of the *18th ACM STOC*, 1986, 240–246. *Combinatorica*, **8**, 1988, 261–277.

[29] G. A. Margulis, Explicit construction of concentrators. *Problems Inform. Transmission* **9**, 1973, 325–332.

[30] G. A. Margulis, Explicit group-theoretic constructions for combinatorial designs with applications expanders and concentrators. *Problems Inform. Transmission* **24**, 1988, 39–46.

[31] J. Naor and M. Naor, Small bias probability spaces: efficient constructions and applications. Proc. of *22nd ACM STOC*, 1990. 213–223.

[32] N. Pippenger, Superconcentrators. *SIAM J. Computing* **6**, pp 298–304, (1972)

[33] M. Pinsker, On the complexity of a concentrator, The *7th International Teletraffic Conference*, Stockholm, 318/1–318/4, 1973.

[34] O. Reingold, S. Vadhan and A. Wigderson, Entropy Waves, The Zig-Zag Graph Product, and New Constant-Degree Expanders and Extractors, Proc. of *41st IEEE FOCS*, 2000, 3–13.

[35] M. Tanner, Explicit concentrators from generalized n-gons. *SIAM J. on Algebraic Discrete Methods*, 5 (3): 287–293, 1984.

[36] L. Valiant, Graph-theoretic properties in computational complexity, *JCSS*, **13**, 1976, 278–285.

Appendix

In this appendix we give some further concrete geometric description of the neighbor set

$$N_B = \{q \in \mathbf{Z}^2 \mid \mu[UB \cap U_q] \neq 0\}$$

used to define the expander graph.

Consider the parallelogram UB. We are to collect all lattice points $q \in \mathbf{Z}^2$ such that there is $\xi \in U$, with $z = q + \xi \in UB$. We can reverse this process and start with an arbitrary $z \in UB$, and "cover" with a square $z + (-U) = \{z - \xi \mid \xi \in U\}$. As z runs through UB, we get a region as the union

$$UB + (-U) = \{z - \xi \mid z \in UB, \xi \in U\}.$$

We look for all lattice points in this region. Since we are only interested in non-zero measure intersections, we can actually restrict the above to open interior sets $(U^o)B + (-U^o)$.

The more interesting claim in this appendix is the following: It suffices to trace the point z along the boundary of UB only, i.e., there is no need to place z in the interior of UB.

$$\begin{aligned} N_B &= \{q \in \mathbf{Z}^2 \mid \mu[(\partial U)B \cap U_q] \neq 0\} \\ &= \{q \in \mathbf{Z}^2 \mid q \in (\partial U)B + (-U^o)\} \end{aligned}$$

Let $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be an integral unimodular matrix. The vertices of UB are $(0,0)$, (a,b), (c,d) and $(a+c, b+d)$ respectively. First we note that being unimodular, there are no integral points in the interior of UB.

We claim that the interior of UB is entirely placed inside one of the four quadrants. Suppose not, say, (a,b) is in the first quadrant ($b > 0$), and (c,d) is in the fourth quadrant ($d < 0$). (The other cases are similar.) Wolog $c \geq a$. Then draw a vertical line from (a,b) to the x-axis, we get a lattice point $(a,0)$ in the interior of UB. By geometric symmetry, it is clear that wolog we may assume UB is entirely placed inside the first quadrant.

There are two cases: $a \leq c$ or $a > c$. We will only consider $a \leq c$; the other case is similar. Cut the parallelogram UB into 3 parts by drawing vertical lines at $x = a$ and $x = c$. (If $a = c$ the middle section is empty.) Thus UB consists of a triangle Δ from $x = 0$ to $x = a$, a parallelogram from $x = a$ to $x = c$ (possibly empty); and another triangle Δ' which is geometrically Δ rotated by π.

Since UB has no lattice point to its interior, the length of the vertical line segment at $x = a$ is at most 1, otherwise $(a, b+1)$ is an interior lattice point. Being a triangle, the length of the vertical line segment on Δ at any x, $0 \leq x \leq a$, is at most 1.

Start at an interior point $z = (x,y) \in \Delta$, and consider the square $z + (-U)$. If we slide this square vertically, there are points $z' = (x, y')$ and $z'' = (x, y'')$ both on the boundary of UB, where $y' < y < y''$, and $y'' - y' \leq 1$, such that $z + (-U)$ is covered by the union of the two corresponding squares based at z' and z'', namely,

$$z + (-U) \subset [z' + (-U)] \cup [z'' + (-U)].$$

Thus if a lattice point $q \in \mathbf{Z}^2$ is found in $z + (-U)$, it is also found in $z' + (-U)$ or $z'' + (-U)$. (Being of non-zero measure, we can also restrict them to $z' + (-U^o)$ and $z'' + (-U^o)$.)

For the parallelogram from $x = a$ to $x = c$ the line segments at $a \leq x \leq c$ are of the same length, all ≤ 1. Finally for Δ' the situation is the same as Δ by symmetry.

Thus to collect all lattice points in N_B, it suffices to trace the point z with an attached square $-U$ on the boundary of UB.

Still placing the parallelogram UB in the first quadrant, we can see that the region is the interior of the convex hull of

$$(-1,-1),(0,-1),(a,b-1),(a+c,b+d-1),(a+c,b+d),(a+c-1,b+d),(c-1,d),(-1,0).$$