



HAL
open science

USEFUL BOUNDS ON THE EXTREME EIGENVALUES AND VECTORS OF MATRICES FOR HARPER'S OPERATORS

Daniel Bump, Persi Diaconis, Angela Hicks, Laurent Miclo, Harold Widom

► **To cite this version:**

Daniel Bump, Persi Diaconis, Angela Hicks, Laurent Miclo, Harold Widom. USEFUL BOUNDS ON THE EXTREME EIGENVALUES AND VECTORS OF MATRICES FOR HARPER'S OPERATORS. Large truncated Toeplitz matrices, Toeplitz operators, and related topics, 2017. hal-02022492

HAL Id: hal-02022492

<https://hal.science/hal-02022492>

Submitted on 18 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Here $\lambda_1(C) = \lambda_1(D) = \frac{1}{2}$ giving $\lambda_1(M) \leq 1$. This was not useful in our application; in particular, we need $\lambda_1(M) \leq 1 - \frac{\text{const}}{n}$. This paper presents three different approaches to proving such bounds. The first approach uses the geometry of the eigenvectors and a discrete version of the Heisenberg uncertainty principle. It works for general Hermitian circulant:

Theorem 1. *Let C be an $n \times n$ Hermitian circulant with eigenvalues $\lambda_1(C) \geq \dots \geq \lambda_n(C)$. Let D be an $n \times n$ real diagonal matrix with eigenvalues $\lambda_1(D) \geq \dots \geq \lambda_n(D)$. If k, k' satisfy $1 \leq k, k' \leq n, kk' < n$, then*

$$\lambda_1(C + D) \leq \lambda_1(C) + \lambda_1(D) - \frac{1}{2} \min\{\lambda_1(D) - \lambda_k(D), \lambda_1(C) - \lambda_{k'}(C)\} \left(1 - \sqrt{\frac{kk'}{n}}\right)^2.$$

Example 1. For the matrix M_n in (1), the eigenvalues of C and D are real and equal to $\{\frac{1}{2} \cos(\frac{2\pi j}{n})\}_{0 \leq j \leq n-1}$. For simplicity, take n odd. Then, writing $\lambda_j = \lambda_j(C) = \lambda_j(D)$, $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \lambda_3 = \frac{1}{2} \cos(\frac{2\pi}{n})$, and $\lambda_{2j+1} = \lambda_{2j} = \frac{1}{2} \cos(\frac{2\pi j}{n})$ for $1 \leq j \leq \frac{n-1}{2}$. Choose $k = k' = \lfloor c\sqrt{n} \rfloor$ for a fixed $0 < c < 1$. Then

$$\lambda_k = \frac{1}{2} \left(1 - \frac{1}{2} \left(\frac{\pi c}{n^{1/2}}\right)^2 + O\left(\frac{1}{n^{3/2}}\right)\right), \quad \left(1 - \sqrt{\frac{kk'}{n}}\right)^2 \geq (1-c)^2$$

and the bound in Theorem 1 becomes

$$\lambda_1(M_n) \leq 1 - \frac{\pi^2 c^2 (1-c)^2}{8n} + O\left(\frac{1}{n^{3/2}}\right).$$

The choice $c = \frac{1}{2}$ gives the best result. Very sharp inequalities for the largest and smallest eigenvalues of M_n follow from [3]. They get better constants than we have in this example. Their techniques make sustained careful use of the exact form of the matrix entries while the techniques in Theorem 1 work for general circulants.

The second approach passes to the large n limit, showing that the largest eigenvalues of M_n from (1) tend to suitably scaled eigenvalues of the harmonic oscillator $L = -\frac{1}{4} \frac{d^2}{dx^2} + \pi^2 x^2$.

Theorem 2. *For a fixed $k \geq 1$, the k th largest eigenvalue of M_n equals*

$$1 - \frac{\mu_k}{n} + o\left(\frac{1}{n}\right)$$

with $\mu_k = \frac{(2k-1)\pi}{2}$, the k th smallest eigenvalue of L .

Theorem 2 gets higher eigenvalues with sharp constants for a restricted family of matrices. The argument also gives a useful approximation to the k th eigenvector. Similar results (with very different proofs) are in [28].

There are many techniques available for bounding the eigenvalues of stochastic matrices ([24], [13], and [7]). We initially thought that some of these would adapt to M_n . However, M_n is far from stochastic: the row sums of M_n are not constant and the entries are sometimes negative. Our third approach is to let $M'_n = \frac{1}{3}I + \frac{2}{3}M_n$. This is substochastic (having non-negative entries and row sums at most 1). If $a_i = 1 - \sum_j M'_n(i, j)$, consider the $(n+1) \times (n+1)$ stochastic matrix:

$$(2) \quad M_n'' = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a_1 & & & & \\ a_2 & & M_n' & & \\ \vdots & & & & \\ a_n & & & & \cdot \end{pmatrix}$$

This has the interpretation of an absorbing Markov chain (0 is the absorbing state) and the Dirichlet eigenvalues of M_n'' (namely those whose eigenvalues vanish at 0) are the eigenvalues of M_n' . In [5] path and other geometric techniques are used to bound these Dirichlet eigenvalues. This results in bounds of the form $1 - \frac{\text{const.}}{n^{4/3}}$ for $\lambda_1(M_n)$. While sufficient for the application, it is natural to want an improvement that gets the right order. Our third approach introduces a purely probabilistic technique which works to give bounds of the right order for a variety of similar matrices.

Theorem 3. *There is a $c > 0$ such that, for all $n \geq 1$ and M_n defined at (1), the largest eigenvalue satisfies $\lambda_1(M_n) \leq 1 - \frac{c}{n}$.*

Section 2 gives background and motivation. Theorems 1, 2, and 3 are proved in Sections 3, 4, and 5. Section 6 treats a simple random walk on the affine group mod p . It uses the analytic bounds to show that order p^2 steps are necessary and sufficient for convergence. It may be consulted now for further motivation. The final section gives the limiting distribution of the bulk of the spectrum of $M_n(a)$ using the Kac-Murdock-Szegö theorem.

2. BACKGROUND

Our work in this area starts with the finite Heisenberg group:

$$H_1(n) = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{Z}/n\mathbb{Z} \right\}.$$

Write such a matrix as (x, y, z) , so

$$(x, y, z)(x', y', z') = (x + x', y + y', z + z' + xy').$$

Let

$$(3) \quad S = \{(1, 0, 0), (-1, 0, 0), (0, 1, 0), (0, -1, 0)\} \text{ and}$$

$$(4) \quad Q(g) = \begin{cases} \frac{1}{4} & g \in S \\ 0 & \text{otherwise} \end{cases}.$$

Thus S is a minimal symmetric generating set for $H_1(n)$ and Q is the probability measure associated with ‘pick an element in S at random and multiply.’ Repeated steps of this walk correspond to convolution. For $(x, y, z) \in H_1(n)$,

$$Q^{*k}(x, y, z) = \sum_{(x', y', z') \in H_1(n)} Q(x', y', z') Q^{*(k-1)}((x, y, z)(x', y', z')^{-1}).$$

When k is large, Q^{*k} converges to the uniform distribution $U(x, y, z) = \frac{1}{n^3}$. The rate of convergence of Q^{*k} to U can be measured by the chi-squared distance:

$$(5) \quad \sum_{(x,y,z) \in H} |Q^{*k}(x, y, z) - U(x, y, z)|^2 / (U(x, y, z)) = \sum_{\substack{\rho \in \hat{H}_1 \\ \rho \neq 1}} d_\rho \|\hat{Q}(\rho)^k\|^2,$$

On the right, the sum is over nontrivial irreducible representations of $H_1(n)$ with ρ of dimension d_ρ and $\hat{Q}(\rho)^k = \sum_{(x,y,z)} Q^{*k}(x, y, z) \rho(x, y, z)$. For background on the Fourier analysis approach to bounded convergence see [8], Chapter 3.

For simplicity, (see [5] for the general case) take $n = p$ a prime. Then $H_1(p)$ has p^2 1-dimensional representations $\rho_{a,b}(x, y, z) = e^{\frac{2\pi i}{p}(ax+by)}$ for a, b in \mathbb{Z}_p . It has $p-1$ p -dimensional representations. These act on $V = \{f : \mathbb{Z}_p \rightarrow \mathbb{C}\}$ via

$$\rho_a(x, y, z)f(w) = e^{\frac{2\pi ia}{p}(yw+z)} f(x+w), \quad 0 \leq a \leq p-1.$$

The Fourier transform of Q at ρ_a is the matrix $M_n(a)$ as in (1) with $\cos\left(\frac{2\pi j}{p}\right)$ replaced by $\cos\left(\frac{2\pi aj}{p}\right)$ for $0 \leq j \leq p-1$.

The chi-squared norm in (5) is the sum of the $(2k)$ th power of the eigenvalues so proceeding needs bounds on these. The details are carried out in [5]. The main results show that k of order n^2 steps are necessary and sufficient for convergence. That paper also summarizes other appearances of the matrices $M_n(a)$. They occur in discrete approximations of the ‘almost Mathieu’ operator in solid state physics. In particular, see [31], [3], and [1]. If \mathcal{F}_n is the discrete Fourier transform matrix $\left((\mathcal{F}_n)_{jk} = \frac{1}{\sqrt{n}} e^{\frac{2\pi ijk}{n}}\right)$; it is easy to see that $\mathcal{F}_n M_n(1) = M_n(1) \mathcal{F}_n$. Diagonalizing \mathcal{F}_n has engineering applications and having a ‘nice’ commuting matrix should help. For this reason, there is engineering interest in the eigenvalues and vectors of $M_n(1)$. See [14] and [25].

3. PROOF OF THEOREM 1

Throughout this section C is an $n \times n$ Hermitian circulant with eigenvalues $\lambda_1(C) \geq \lambda_2(C) \geq \dots \geq \lambda_n(C)$ and D is a real diagonal matrix with eigenvalues $\lambda_1(D) \geq \lambda_2(D) \geq \dots \geq \lambda_n(D)$. Let x be an eigenvector of $C + D$ corresponding to $\lambda_1(C + D)$. Recall that $\left((\mathcal{F}_n)_{jk} = \frac{1}{\sqrt{n}} e^{\frac{2\pi ijk}{n}}\right)$ for $j, k \in \mathbb{Z}/n\mathbb{Z}$. This has rows or columns which simultaneously diagonalize all circulants. Write $\hat{x} = \mathcal{F}_n x$ and x^h for the conjugate transpose. We use $\|x\|^2 = x^h x$.

Our aim is to prove that for $kk' < n$,

$$(6) \quad \lambda_1(C + D) \leq \lambda_1(C) + \lambda_1(D) - \frac{1}{2} \min\{\lambda_1(D) - \lambda_{k+1}(D), \lambda_1(C) - \lambda_{k'+1}(C)\} \left(1 - \sqrt{\frac{kk'}{n}}\right)^2.$$

The first step is to write $x^h C x$ in terms of a Fourier transform pair $\hat{x} = \mathcal{F}_n x$. A subtle point is that although \mathcal{F}_n diagonalizes C , the resulting diagonal matrix does not necessarily have entries in decreasing order, necessitating a permutation indexing in the following lemma.

Lemma 1. *Define a permutation σ such that*

$$\frac{e^{\frac{2\pi i}{n}\sigma_j^{-1}b}}{\sqrt{n}}, \quad 0 \leq b \leq n-1$$

is the eigenvector corresponding to $\lambda_j(C)$. Then

$$(7) \quad x^h C x = \hat{x}^h D' \hat{x}$$

with $D' = \text{diag}(\lambda_{\sigma_1}(C), \dots, \lambda_{\sigma_n}(C))$.

Proof. Since C is diagonalized by \mathcal{F}_n^h , $\mathcal{F}_n C \mathcal{F}_n^h = D'$. Thus

$$x^h C x = x^h \mathcal{F}_n^h \mathcal{F}_n C \mathcal{F}_n^h \mathcal{F}_n x = \hat{x}^h D' \hat{x}.$$

□

A key tool is the Donoho-Stark [15] version of the Heisenberg Uncertainty Principle. For this, call a vector y ' ϵ -concentrated on a set $S \subset [n]$ ' if $|x_i| < \epsilon$ for $i \notin S$.

Theorem 4 (Donoho-Stark). *Let y, \hat{y} be a unit norm Fourier Transform pair with $y \in_S$ -concentrated on S and $\hat{y} \in_T$ -concentrated on T . Then*

$$(8) \quad |S||T| \geq n(1 - (\epsilon_S + \epsilon_T))^2.$$

Let $(y)_S$ be the projection onto the subspace vanishing off S :

$$((y)_S)_i = \begin{cases} y_i & i \in S \\ 0 & \text{otherwise} \end{cases}.$$

A simple consequence of the bound (8) is

Corollary 1. *If $kk' < n$, z, \hat{z} a unit norm Fourier transform pair and S and T are sets of size k, k' , then*

$$\|(z)_{S^c}\|^2 + \|(\hat{z})_{T^c}\|^2 \geq \frac{1}{2} \left(1 - \sqrt{\frac{kk'}{n}}\right)^2.$$

Proof. Let $\epsilon_S = \|(z)_{S^c}\|$ and $\epsilon_T = \|(\hat{z})_{T^c}\|$. Then $\|z - (z)_S\| = \|(z)_{S^c}\|$ and $\|\hat{z} - (\hat{z})_T\| = \|(\hat{z})_{T^c}\|$. Thus z is ϵ_S concentrated on S and \hat{z} is ϵ_T concentrated on T . Thus (8) gives

$$\frac{|S||T|}{n} = \frac{kk'}{n} \geq (1 - (\epsilon_S + \epsilon_T))^2 \text{ or } (\epsilon_S + \epsilon_T) \geq 1 - \sqrt{\frac{kk'}{n}},$$

so if $kk' \leq n$

$$\|(z)_{S^c}\|^2 + \|(\hat{z})_{T^c}\|^2 = (\epsilon_S^2 + \epsilon_T^2) \geq \frac{1}{2}(\epsilon_S + \epsilon_T)^2 \geq \frac{1}{2} \left(1 - \sqrt{\frac{kk'}{n}}\right)^2$$

□

Proof of Theorem 1. With notation as above,

$$\lambda_1(C + D) = x^h(C + D)x = x^h C x + x^h D x = \hat{x}^h D' \hat{x} + x^h D x =: *.$$

Let $\bar{D} = D - \lambda_1(D)I$ and $\bar{D}' = D' - \lambda_1(C)I$. Then

$$\begin{aligned} * &= \hat{x}^h \lambda_1(C)I \hat{x} + \hat{x}^h \bar{D}' \hat{x} + x^h \lambda_1(D)I x + x^h \bar{D} x \\ &= \lambda_1(C) + \hat{x}^h \bar{D}' \hat{x} + \lambda_1(D) + x^h \bar{D} x \end{aligned}$$

Now \overline{D} and \overline{D}' have non-positive eigenvalues so our improvement over the Weyl bounds will follow by showing that x or \hat{x} have support on suitably negative entries of \overline{D}' or \overline{D} .

Let S and T correspond to the largest k, k' entries of $\overline{D}, \overline{D}'$, respectively. Then $x = (x)_S + (x)_{S^c}$, $\hat{x} = (\hat{x})_T + (\hat{x})_{T^c}$. Each of those decomposition is into orthogonal pieces. Multiplying any of the four pieces by an arbitrary diagonal matrix preserves this orthogonality. Thus

$$* = \lambda_1(C) + \lambda_1(D) + (\hat{x})_T^h \overline{D}' (\hat{x})_T + (\hat{x})_{T^c}^h \overline{D}' (\hat{x})_{T^c} + (x)_S^h \overline{D} (x)_S + (x)_{S^c}^h \overline{D} (x)_{S^c}.$$

For the last four terms on the right, terms 1 and 3 are bounded above by zero and 2 and 4 contribute with the following bounds:

$$\begin{aligned} * &\leq \lambda_1(C) + \lambda_1(D) + (\lambda_{k+1}(D) - \lambda_1(D)) \|(x)_{S^c}\|^2 + (\lambda_{k'+1}(C) - \lambda_1(C)) \|(\hat{x})_{T^c}\|^2 \\ &\leq \lambda_1(C) + \lambda_1(D) \\ &\quad + \min\{(\lambda_{k+1}(D) - \lambda_1(D)), (\lambda_{k'+1}(C) - \lambda_1(C))\} \left(\frac{1}{2} \left(1 - \sqrt{\frac{kk'}{n}} \right)^2 \right) \end{aligned}$$

where the last line follows from the corollary. \square

Remarks. (1) These arguments work to give the smallest eigenvalue as well, so in fact we also have for $ll' < n$:

$$(9) \quad \lambda_n(C + D) \geq \lambda_n(C) + \lambda_n(D)$$

$$+ \frac{1}{2} \min\{\lambda_{n-l}(D) - \lambda_n(D), \lambda_{n-l'}(C) - \lambda_n(C)\} \left(1 - \sqrt{\frac{ll'}{n}} \right)^2.$$

(2) Our thanks to a thoughtful anonymous reviewer, who pointed out that Corollary 1 can be improved using Cauchy- Schwartz to show that for $0 < a, b \leq 1$,

$$\left(1 - \sqrt{\frac{kk'}{n}} \right)^2 \leq (\epsilon_s + \epsilon_T)^2 \leq (a\epsilon_S^2 + b\epsilon_T^2) \left(\frac{1}{a} + \frac{1}{b} \right).$$

Setting $a = \lambda_{k+1}(D) - \lambda_1(D)$ and $b = \lambda_{k'+1}(C) - \lambda_1(C)$, one can improve the previous theorem:

$$\lambda_1(C + D) \leq \lambda_1(C) + \lambda_1(D)$$

$$- \frac{(\lambda_1(D) - \lambda_{k+1}(D))(\lambda_1(C) - \lambda_{k'+1}(C))}{(\lambda_1(D) - \lambda_{k+1}(D)) + (\lambda_1(C) - \lambda_{k'+1}(C))} \left(1 - \sqrt{\frac{kk'}{n}} \right)^2.$$

In our case, the result is the same, since the eigenvalues of C and D are identical.

- (3) Donoho and Stark [15] give many variations on their uncertainty principle suitable for other transforms. The techniques above should generalize, at least to the G -circulants of [9].
- (4) There should be similar theorems with C and D replaced by general Hermitian matrices and perhaps extensions to higher Weyl and Horn inequalities (see [2] and [18]).
- (5) Further applications/examples are in Section 6.

4. THE HARMONIC OSCILLATOR AS A LIMIT.

We prove Theorem 2, that for $k \geq 1$ the k th largest eigenvalue of M_n is equal to $1 - \mu_k/n + o(1/n)$ and the k th smallest eigenvalue of M_n is equal to $-1 + \mu_k/n + o(1/n)$, where μ_k is the k th smallest eigenvalue of

$$L = -\frac{1}{4} \frac{d^2}{dx^2} + \pi^2 x^2$$

on $(-\infty, \infty)$. By a classical computation (see [19]), $\mu_k = (2k - 1)/(2\pi)$.

The $n \times n$ matrix M_n has j, k -entry

$$\frac{1}{4} [\delta(j - k - 1) + \delta(j - k + 1)] + \frac{1}{2} \cos(2\pi k/n) \delta(j - k),$$

where $j, k \in \mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.

We define

$$\tilde{M}_n = n(I - M_n).$$

This has j, k entry $m_1(j, k) + m_2(j, k)$, where

$$(10) \quad m_1(j, k) = \frac{n}{2} \left(\delta(j - k) - \frac{1}{2} [\delta(j - k - 1) + \delta(j - k + 1)] \right)$$

$$(11) \quad m_2(j, k) = \frac{n}{2} (1 - \cos(2\pi k/n)) \delta(j - k).$$

We will show first that if μ is any limit of eigenvalues of \tilde{M}_n then μ is an eigenvalue of L ; and, second, that any eigenvalue μ of L has a neighborhood that contains exactly one eigenvalue, counting multiplicity, of \tilde{M}_n for n sufficiently large. These imply the stated result.

These will be accomplished as follows. Give each point of \mathbb{Z}_n measure $1/\sqrt{n}$, so the total measure equals \sqrt{n} . We then define an isometry T from $L^2(\mathbb{Z}_n)$ to $L^2(-\sqrt{n}/2, \sqrt{n}/2)$ (thought of as a subspace of $L^2(\mathbb{R})$ with Lebesgue measure) for which the following hold:

Proposition 1. *Suppose $\{u_n\}$ is a sequence of functions of norm one in $L^2(\mathbb{Z}_n)$ such that the sequence $\{(\tilde{M}_n u_n, u_n)\}$ of inner products is bounded. Then $\{T u_n\}$ has a strongly (i.e., in norm) convergent subsequence.*

Proposition 2. *If ϕ is a Schwartz function on \mathbb{R} then $T \tilde{M}_n T^* \phi \rightarrow L \phi$ strongly.¹*

These will easily give the desired results. (See Propositions 3 and 4 near the end.) The final section 4.2 treats the smallest eigenvalues.

4.1. Proofs for the largest eigenvalues. We use two transforms (with, confusingly, the same notation). First, for ϕ in $L^2(-\sqrt{n}/2, \sqrt{n}/2)$ we define

$$\hat{\phi}(\ell) = \int_{-\sqrt{n}/2}^{\sqrt{n}/2} e^{-2\pi i \ell x / \sqrt{n}} \phi(x) dx, \quad (\ell \in \mathbb{Z}),$$

and we have by Parseval (after making the substitution $x \mapsto x\sqrt{n}$ in the integral)

$$(12) \quad \|\hat{\phi}\| = n^{1/4} \|\phi\|.$$

Here $\|\hat{\phi}\|^2 = \sum_{\ell \in \mathbb{Z}} |\hat{\phi}(\ell)|^2$.

¹The operator T^* acts on $L^2(-\sqrt{n}/2, \sqrt{n}/2)$, so ϕ is first to be restricted to $(-\sqrt{n}/2, \sqrt{n}/2)$.

For $u \in L^2(\mathbb{Z}_n)$ we have its finite Fourier transform

$$\widehat{u}(\ell) = \sum_k e^{-2\pi i \ell k/n} u(k), \quad (\ell \in \mathbb{Z}_n),$$

and we compute below that

$$(13) \quad \|\widehat{u}\| = n^{3/4} \|u\|.$$

Here

$$\|\widehat{u}\|^2 = \sum |\widehat{u}(\ell)|^2,$$

the sum over any integer interval of length n . To show (13), we have

$$|\widehat{u}(\ell)|^2 = \sum_{j,k} e^{-2\pi i \ell(j-k)/n} u(j) \overline{u(k)}.$$

Since $e^{-2\pi i \ell(j-k)/n}$ is an n th root of unity, equal to 1 only when $j = k$ in \mathbb{Z}_n , we get

$$\|\widehat{u}\|^2 = \sum_{\ell} |\widehat{u}(\ell)|^2 = n \sum_k |u(k)|^2 = n^{3/2} \|u\|^2.$$

Now we define the operator T . Let J be an interval of integers of length n (which later will be specified further) and set

$$D_n(x) = \sum_{\ell \in J} e^{2\pi i \ell x}.$$

Then T is defined by

$$(Tu)(x) = \frac{1}{n} \sum_k D_n\left(\frac{x}{\sqrt{n}} - \frac{k}{n}\right) u(k).$$

Thus T has kernel

$$T(x, k) = \frac{1}{n} D_n\left(\frac{x}{\sqrt{n}} - \frac{k}{n}\right).$$

By the definition of the inner product on $L^2(\mathbb{Z}_n)$ we find that

$$T^* : L^2(-\sqrt{n}/2, \sqrt{n}/2) \rightarrow L^2(\mathbb{Z}_n)$$

has kernel

$$T^*(k, x) = \frac{1}{\sqrt{n}} D_n\left(\frac{k}{n} - \frac{x}{\sqrt{n}}\right).$$

In terms of the transforms we have the following:

Lemma 2. (a) For $u \in L^2(\mathbb{Z}_n)$,

$$\widehat{Tu}(\ell) = \begin{cases} \frac{1}{\sqrt{n}} \widehat{u}(\ell) & \text{if } \ell \in J, \\ 0 & \text{if } \ell \notin J. \end{cases}$$

(b) For $\phi \in L^2(-\sqrt{n}/2, \sqrt{n}/2)$,

$$\widehat{T^*\phi}(\ell) = \sqrt{n} \widehat{\phi}(\ell) \quad \text{when } \ell \in J.$$

Proof. For (a), we have

$$\begin{aligned}\widehat{Tu}(\ell) &= \frac{1}{n} \int \sum_{k, \ell' \in J} e^{-2\pi i \ell x / \sqrt{n}} e^{2\pi i \ell' (x / \sqrt{n} - k/n)} u(k) dx \\ &= \frac{1}{n} \int \sum_{\ell' \in J} e^{2\pi i (\ell' - \ell)x / \sqrt{n}} \widehat{u}(\ell') dx.\end{aligned}$$

The result follows.

For (b), we have when $\ell \in J$,

$$\begin{aligned}\widehat{T^* \phi}(\ell) &= \frac{1}{\sqrt{n}} \sum_{k, \ell' \in J} \int e^{-2\pi i k \ell / n} e^{2\pi i \ell' (k/n - x / \sqrt{n})} \phi(x) dx \\ &= \sqrt{n} \int e^{-2\pi i \ell x / \sqrt{n}} \phi(x) dx = \sqrt{n} \widehat{\phi}(\ell).\end{aligned}$$

□

We show two things about T . For the second we shall assume now and hereafter that the end-points of J are $\pm n/2 + O(1)$, although this is a lot stronger than necessary.

Lemma 3. (a) $T^* T = I$. (b) $T T^* \rightarrow I$ strongly as $n \rightarrow \infty$.

Proof. By Lemma 2b,

$$\widehat{T^* T u}(\ell) = \sqrt{n} \widehat{T u}(\ell'),$$

where $\ell' \in J$ and $\ell' - \ell \in n\mathbb{Z}$. By Lemma 2a this in turn equals $\widehat{u}(\ell')$, which equals $\widehat{u}(\ell)$ since \widehat{u} is n -periodic. This gives (a).

For (b) observe that $T T^*$ is self-adjoint. Since $(T T^*)^2 = T T^* T T^* = T T^*$, it is a (nonzero) projection and so has norm one. Therefore it suffices to show that if ϕ is a Schwartz function then $T T^* \phi \rightarrow \phi$. We have from Lemma 2a that

$$\widehat{T T^* \phi}(\ell) = \frac{1}{\sqrt{n}} \widehat{T^* \phi}(\ell)$$

if $\ell \in J$ and equals zero otherwise. If $\ell \in J$ then by Lemma 2b it equals $\widehat{\phi}(\ell)$. It follows that

$$\|\widehat{T T^* \phi} - \widehat{\phi}\|^2 = \sum_{\ell \notin J} |\widehat{\phi}(\ell)|^2.$$

Integrating by parts shows that

$$\widehat{\phi}(\ell) = O(\sqrt{n}/\ell),$$

and so, by our assumption on J , the sum on the right side is $O(1)$. Then by (12) we get

$$\|T T^* \phi - \phi\| = O(n^{-1/4}).$$

□

Now the work begins. First, an identity. We introduce the notations

$$\overline{C}(\xi) = 1 - \cos(2\pi\xi), \quad S(\xi) = \sin(\pi\xi),$$

and observe that $\overline{C}(\xi) = 2S(\xi)^2$.

Lemma 4. For $u \in L^2(\mathbb{Z}_n)$,

$$(\tilde{M}_n u, u) = n^{-1/2} \|S(\ell/n) \widehat{u}(\ell)\|^2 + n \|S(k/n) u(k)\|^2.$$

Note. Here and below we display “ k ” as the variable in the ambient space \mathbb{Z}_n and “ ℓ ” as the variable in the space \mathbb{Z}_n of the Fourier transform. We abuse notation and, for example, the “ $u(k)$ ” above denotes the function $k \rightarrow u(k)$.

Proof. We consider first the contribution of (10) to the inner product. If we define the operators A_\pm by

$$(A_\pm u)(k) = u(k \pm 1),$$

we see that the contribution to the inner product is

$$\frac{n}{2} (u - [A_+ u + A_- u]/2, u).$$

Now

$$\widehat{A_\pm u}(\ell) = e^{\pm 2\pi i \ell/n} \widehat{u}(\ell),$$

so if we use (13) we see that the above is equal to

$$\frac{1}{2} n^{-1/2} (\overline{C}(\ell/n) \widehat{u}, \widehat{u}(\ell)) = n^{-1/2} \|S(\ell/n) \widehat{u}\|^2.$$

To complete the proof of the lemma we note that the contribution to the inner product of (11) is clearly

$$\frac{n}{2} (\overline{C}(k/n) u(k), u(k)) = n \|S(k/n) u(k)\|^2.$$

□

Lemma 5. Suppose u_n satisfy $(\tilde{M}_n u_n, u_n) = O(1)$. Then (a) $\|x T u_n(x)\| = O(1)$, and (b) $\|(T u_n)'\| = O(1)$.

Proof of (a). We have

$$\widehat{u}_n(\ell) - \widehat{u}_n(\ell + 1) = \sum_k e^{-2\pi i \ell k/n} (1 - e^{-2\pi i k/n}) u_n(k),$$

the finite Fourier transform of $(1 - e^{-2\pi i k/n}) u_n(k)$. We have,

$$|(1 - e^{-2\pi i k/n}) u_n(k)| = 2|S(k/n) u_n(k)|.$$

Therefore from (13) and

$$\|S(k/n) u_n(k)\| = O(n^{-1/2}),$$

which follows from Lemma 4, we get

$$\|\widehat{u}_n(\ell) - \widehat{u}_n(\ell + 1)\| = O(n^{1/4}).$$

It follows from Lemma 2a that $\widehat{T u}_n(\ell) = \widehat{T u}_n(\ell + 1) = 0$ if both $\ell, \ell + 1 \notin J$ and

$$(14) \quad \sum_{\ell, \ell+1 \in J} |\widehat{T u}_n(\ell) - \widehat{T u}_n(\ell + 1)|^2 = O(n^{-1/2}).$$

If $\ell \in J$ but $\ell + 1 \notin J$ then $\widehat{T u}_n(\ell + 1) = 0$ and ℓ is the right end-point of J and therefore $n/2 + O(1)$. From

$$(15) \quad \|S(\ell/n) \widehat{u}_n(\ell)\| = O(n^{1/4}),$$

which also follows from Lemma 4, and that $|S(\ell/n)|$ is bounded below for $\ell = n/2 + O(1)$, we have in particular that $\widehat{u}_n(\ell) = O(n^{1/4})$, Therefore $\widehat{Tu}_n(\ell) = O(n^{-1/4})$.

So the bound in (14) holds when the sum is taken over all $\ell \in \mathbb{Z}$. Since

$$\widehat{Tu}_n(\ell) - \widehat{Tu}_n(\ell + 1) = \int e^{-2\pi i x \ell / \sqrt{n}} (1 - e^{-2\pi i x / \sqrt{n}}) Tu_n(x) dx,$$

it follows from (12) that

$$\|S(x/\sqrt{n}) Tu_n(x)\| = O(n^{-1/2}).$$

In the interval of integration $|x| < \sqrt{n}/2$, so $|S(x/\sqrt{n})|$ is bounded below by a constant times $|x|/\sqrt{n}$. This gives (a). \square

Proof of (b). We have

$$(Tu_n)'(x) = \frac{2\pi i}{n^{3/2}} \sum_{k, \ell \in J} \ell e^{2\pi i \ell (x/\sqrt{n} - k/n)} u_n(k) = \frac{2\pi i}{n^{3/2}} \sum_{\ell \in J} \ell e^{2\pi i \ell x / \sqrt{n}} \widehat{u}_n(\ell).$$

Thus $\widehat{Tu}_n'(\ell) = 2\pi i \ell \widehat{u}_n(\ell)/n$ for $\ell \in J$, and it follows from (12) that

$$\|(Tu_n)'\|^2 = \frac{4\pi^2}{n^{5/2}} \sum_{\ell \in J} \ell^2 |\widehat{u}_n(\ell)|^2.$$

Now $|S(\ell/n)|$ is bounded below by a constant times $|\ell/n|$ for $\ell \in J$, so (15) implies that

$$\sum_{\ell \in J} \ell^2 |\widehat{u}_n(\ell)|^2 = O(n^{5/2}),$$

which gives the result. \square

Proof of Proposition 1. Since T is an isometry each $\|Tu_n\| = 1$, and by passing to a subsequence we may assume $\{Tu_n\}$ converges weakly to some $f \in L^2(\mathbb{R})$. We use the fact that strong convergence will follow if we can show that $\|f\| \geq 1$. (In general, if $\|f_n\| = 1$ and $f_n \rightarrow f$ weakly, then $\|f\| \geq 1$ implies that $f_n \rightarrow f$ strongly. Here is the argument. We have that

$$\|f_n - f\|^2 = \|f_n\|^2 + \|f\|^2 - 2 \operatorname{Re}(f_n, f).$$

By weak convergence, $(f_n, f) \rightarrow \|f\|^2$. Therefore

$$\|f_n - f\|^2 \rightarrow 1 - \|f\|^2 \leq 0,$$

so $\|f_n - f\| \rightarrow 0$.)

The hypothesis of Lemma 5 is satisfied. It follows from Lemma 5a that for each $\varepsilon > 0$ there is a bounded interval A such that

$$\|(1 - \chi_A) Tu_n\| \leq \varepsilon$$

for all n . So $\|\chi_A Tu_n\| \geq 1 - \varepsilon$. It follows from Lemma 5b that $\{Tu_n\}$ is equicontinuous, and this combined with $\|\chi_A Tu_n\| \leq 1$ shows that a subsequence of $\{Tu_n\}$ converges uniformly on A (to $\chi_A f$), and so $\|\chi_A f\| \geq 1 - \varepsilon$. Thus, $\|f\| \geq 1$. \square

Proof of Proposition 2. Consider first the operator corresponding to m_1 in (10). We call it $n\Delta_k^2$. (The subscript indicates that it acts on functions of k .) We show first that

$$n T \Delta_k^2 T^* \phi \rightarrow -\phi''/4$$

in $L^2(\mathbb{R})$. We have

$$\begin{aligned} n(T\Delta_k^2 T^* \phi)(x) &= \frac{1}{2n^{1/2}} \sum_{\ell, \ell', k} e^{2\pi i \ell((x/\sqrt{n}-k/n))} \int_{-\sqrt{n}/2}^{\sqrt{n}/2} \Delta_k^2 e^{2\pi i \ell'((k/n-y/\sqrt{n}))} \phi(y) dy \end{aligned}$$

The exponent in the integral is a function of $k/\sqrt{n} - y$. So taking the second difference Δ_k^2 in k is the same as taking the second difference Δ_y^2 in y as long as the differences in the y -variable are $1/\sqrt{n}$. With this understanding, the above equals

$$\frac{1}{2n^{1/2}} \sum_{\ell, \ell', k} e^{2\pi i \ell((x/\sqrt{n}-k/n))} \int_{-\sqrt{n}/2}^{\sqrt{n}/2} \Delta_y^2 e^{2\pi i \ell'((k/n-y/\sqrt{n}))} \phi(y) dy.$$

By changing variables in two of the three summands from Δ_y^2 we can put the Δ_y^2 in front of the $\phi(y)$. There is an error because of the little change of integration domains but (for ϕ a Schwartz function) this is a rapidly decreasing function of n , and so can be ignored. After this what we get is $nT T^* \Delta^2 \phi$. Taylor's theorem gives

$$n(\Delta_x^2 \phi)(x) = -\phi''(x)/4 + O\left(n^{-1/2} \max_{|y-x| < 1/\sqrt{n}} |\phi'''(y)|\right),$$

from which it follows that $n(\Delta_x^2 \phi)(x) \rightarrow -\phi''(x)/4$ strongly. Since $TT^* \rightarrow I$ strongly we deduce that $nT\Delta_k^2 T^* \phi \rightarrow -\phi''/2$ strongly.

Lastly, consider the operator corresponding to (11), which is multiplication by $n\overline{C}(k/n)/2 = n(1 - \cos(2\pi k/n))/2$. For convenience we call this operator $\overline{C}_n/2$.

By Lemma 2b we know that $\widehat{T^* \phi}(\ell) = \sqrt{n} \widehat{\phi}(\ell)$ when $\ell \in J$. In general,

$$\widehat{\overline{C}_n u}(\ell) = n \sum_k (1 - \cos(2\pi k/n)) e^{-2\pi i k \ell/n} u(k) = n(\Delta^2 \widehat{u})(\ell),$$

where here

$$(\Delta^2 \widehat{u})(\ell) = \widehat{u}(\ell) - [\widehat{u}(\ell-1) + \widehat{u}(\ell+1)]/2.$$

Applying this to $T^* \phi$ gives

$$\widehat{\overline{C}_n T^* \phi}(\ell) = n \Delta^2 \widehat{T^* \phi}(\ell) = n^{3/2} \Delta^2 \widehat{\phi}(\ell)$$

as long as $\ell \pm 1$ are also in J . If ℓ is in J but one of $\ell \pm 1$ is not in J then ℓ is near an end-point of J and the error committed will be rapidly decreasing as $n \rightarrow \infty$. For the right side is $n^{3/2}$ times a linear combination of integrals like

$$\int_{-\sqrt{n}/2}^{\sqrt{n}/2} e^{\pm \pi i x \sqrt{n}} \phi(x) dx,$$

and integration by parts many times shows this is rapidly decreasing. For the left side, if for example $\ell+1 \notin J$ then it is the same as the value at $\ell+1-n \in J$, which is rapidly decreasing. So we ignore this little error and use

$$\Delta^2 \widehat{\phi}(\ell) = \int_{-\sqrt{n}/2}^{\sqrt{n}/2} e^{-2\pi i x \ell/\sqrt{n}} \overline{C}(x/\sqrt{n}) \phi(x) dx.$$

We also have $\widehat{T^* x^2 \phi}(\ell) = \sqrt{n} \widehat{x^2 \phi}(\ell)$ for $\ell \in J$. Thus (ignoring the error),

$$\widehat{\overline{C}_n T^* \phi}(\ell) - 2\pi^2 \widehat{T^* x^2 \phi}(\ell) = \sqrt{n} \int_{-\sqrt{n}/2}^{\sqrt{n}/2} e^{-2\pi i x \ell/\sqrt{n}} (n\overline{C}(x/\sqrt{n}) - 2\pi^2 x^2) \phi(x) dx.$$

From this we get

$$\|\widehat{\overline{C}_n T^* \phi} - 2\pi^2 \widehat{T^* x^2 \phi}\| \leq n^{3/4} \|(n \overline{C}(x/\sqrt{n}) - 2\pi^2 x^2) \phi(x)\| = O(n^{-1/4}),$$

since $n \overline{C}(x/\sqrt{n}) - 2\pi^2 x^2 = O(x^4/n)$ and $x^4 \phi(x) \in L^2(\mathbb{R})$. Then we get from (13)

$$\|\overline{C}_n T^* \phi - 2\pi^2 T^* x^2 \phi\| = O(n^{-1}).$$

Equivalently,

$$\|T \overline{C}_n T^* \phi - 2\pi^2 T T^* x^2 \phi\| = O(n^{-1}).$$

Since $TT^* \rightarrow I$ strongly, this gives

$$\|T \overline{C}_n T^* \phi - 2\pi^2 x^2 \phi\| \rightarrow 0.$$

Since the operator corresponding to (11) is multiplication by $\overline{C}_n/2$, this completes the proof of Proposition 2. \square

Now we go back to the eigenvalues of \tilde{M}_n and easy consequences of Propositions 1 and 2.

Proposition 3. (a) If λ_n are eigenvalues of \tilde{M}_n and $\lambda_n \rightarrow \mu$, then μ is an eigenvalue of L . (b) Any eigenvalue μ of L has a neighborhood that contains at most one eigenvalue (counting multiplicity) of \tilde{M}_n for sufficiently large n .

Proof of (a). Suppose that u_n is an eigenvector of \tilde{M}_n of norm one with eigenvalue λ_n . In particular $(\tilde{M}_n u_n, u_n) = \lambda_n$. By Proposition 1 there is a subsequence of $\{Tu_n\}$ that converges strongly to some $f \in L^2(\mathbb{R})$. For a Schwartz function ϕ we have

$$\begin{aligned} (\mu f, \phi) &= \lim (\lambda_n Tu_n, \phi) = \lim (T\tilde{M}_n u_n, \phi) \\ &= \lim (u_n, \tilde{M}_n T^* \phi) = \lim (Tu_n, T\tilde{M}_n T^* \phi), \end{aligned}$$

since T is an isometry. By Proposition 2, $T\tilde{M}_n T^* \phi$ converges strongly to $L\phi$. Therefore² the limit equals $(f, L\phi)$, and we have shown

$$\mu(f, \phi) = (f, L\phi).$$

It follows that f is an eigenfunction of L with corresponding eigenvalue μ . Here is why. The eigenfunctions of L are the harmonic oscillator wave functions ϕ_i , and therefore Schwartz functions. If the corresponding eigenvalues are μ_i , then

$$(\mu - \mu_i)(f, \phi_i) = \mu(f, \phi_i) - (f, L\phi_i) = 0.$$

Since the ϕ_i are complete and $f \neq 0$, some $\mu - \mu_k = 0$. And f , being orthogonal to the ϕ_i with $i \neq k$, must be a multiple of ϕ_k and therefore a corresponding eigenfunction. \square

Proof of (b). Suppose the contrary were true. Then there would be sequences of eigenvalues $\{\lambda_n\}$ and $\{\lambda'_n\}$ of \tilde{M}_n , both converging to μ , and corresponding orthogonal (since \tilde{M}_n is self-adjoint) eigenfunctions u_n and u'_n . The strong (sub)limits f and f' of Tu_n and Tu'_n would be mutually orthogonal eigenfunctions of L corresponding to the same eigenvalue μ of L . Since the eigenvalues of L are simple, this cannot happen. \square

²For this we need only weak convergence of one and strong convergence of the other. But we also need that $f \neq 0$, which is no easier to show than strong convergence of Tu_n , and we shall need strong convergence of $T\tilde{M}_n T^* \phi$ for Proposition 4.

Proposition 4. *For each eigenvalue μ of L there is a sequence of eigenvalues λ_n of \tilde{M}_n that converges to μ .*

Proof. With corresponding eigenfunction ϕ of L we have, by Proposition 2,

$$\|T\tilde{M}_nT^*\phi - \mu\phi\| = \|T\tilde{M}_nT^*\phi - L\phi\| = o(1).$$

From this, and that $\|\phi - TT^*\phi\| = o(1)$ by Lemma 3b, we get

$$\|T(\tilde{M}_nT^*\phi - \mu T^*\phi)\| = o(1).$$

Since T is an isometry this is the same as

$$\|\tilde{M}_nT^*\phi - \mu T^*\phi\| = o(1).$$

Since $\|T^*\phi\| = \|TT^*\phi\| \rightarrow \|\phi\| \neq 0$, and the other eigenvalues of L are bounded away from μ , this implies that μ is within $o(1)$ of an eigenvalue of \tilde{M}_n and $T^*\phi$ within $o(1)$ of an eigenvector. See [11]. \square

Combining Propositions 3 and 4 shows that the k th largest eigenvalue of M_n equals $1 - \mu_k/n + o(1/n)$.

4.2. The bottom eigenvalues of M_n . We shall find a unitary operator U on $L^2(\mathbb{Z}_n)$ such that the quadratic form for $n(I + U_nM_nU_n^*)$ is the same as for $n(I - M_n)$ when n is even and close to it when n is odd. From that it will follow that the k th bottom eigenvalue of M_n equals $-1 + \mu_k/n + o(1/n)$.

Recall that Lemma 4 says that

$$(n(I - M_n)u, u) = n^{-1/2} \|S(\ell/n)\hat{u}\|^2 + n \|S(k/n)u(k)\|^2 =: Q(u),$$

where $S(\xi) = \sin(\pi\xi)$. For this we used the identity $1 - \cos\xi = 2\sin^2(\xi/2)$. For $n(I + M_n)$ this gets replaced by $1 + \cos\xi = 2\cos^2(\xi/2)$. So now we define

$$C(\xi) = \cos(\pi\xi),$$

and get

$$(n(I + M_n)u, u) = n^{-1/2} \|C(\ell/n)\hat{u}\|^2 + n \|C(k/n)u(k)\|^2.$$

We consider first the less straightforward case of n odd and define

$$v(k) = e^{2\pi i\alpha k} u(k - (n+1)/2),$$

where α (real) will be determined below. We have

$$\|C(k/n)u(k)\| = \|C(k/n + 1/2 + 1/2n)v(k)\| = \|S(k/n + 1/2n)v(k)\|.$$

Next,

$$\begin{aligned} \hat{v}(\ell) &= \sum_k e^{-2\pi i k \ell/n} e^{2\pi i \alpha k} u(k - (n+1)/2) \\ &= \sum_k e^{-2\pi i (k+(n+1)/2)\ell/n} e^{2\pi i \alpha (k+(n+1)/2)} u(k) \\ &= (-1)^\ell e^{-\pi i \ell/n} e^{\pi i \alpha (n+1)} \sum_k e^{-2\pi i k \ell/n} e^{2\pi i \alpha k} u(k) \end{aligned}$$

We choose $\alpha = -(n+1)/2n$. The factor outside the sum has absolute value 1 while the sum becomes $\hat{u}(\ell + (n+1)/2)$. Alternatively,

$$|\hat{u}(\ell)| = |\hat{v}(\ell - (n+1)/2)|.$$

Therefore

$$\|C(\ell/n)\widehat{u}(\ell)\| = \|C(\ell/n + 1/2 + 1/2n)\widehat{v}(\ell)\| = \|S(\ell/n + 1/2n)\widehat{v}(\ell)\|.$$

The map $U : u \rightarrow v$ is unitary and we have shown

Lemma 6. *For n odd we have,*

$$(n(I + UM_nU^*)v, v) = n^{-1/2} \|S(\ell/n + 1/2n)\widehat{v}(\ell)\|^2 + n \|S(k/n + 1/2n)v(k)\|^2.$$

Remark. . When n is even we replace the shift $(n+1)/2$ by $n/2$ and α by $-1/2$, and the extra $1/2n$'s do not appear in the arguments of the S 's. The quadratic form becomes $Q(v)$ exactly, so $U_n M_n U_n^* = -M_n$.³

Proposition 5. $(n(I + UM_nU^*)v, v) = (1 + O(n^{-1/2}))Q(v) + O(n^{-1/2}\|v\|^2)$.

Proof. Since $dS/d\xi$ is bounded,

$$\|S(k/n + 1/2n)v(k)\| \leq \|S(k/n)v(k)\| + O(n^{-1}\|v(k)\|).$$

It follows from the arithmetic-geometric mean inequality that for any $\varepsilon > 0$

$$|a + b|^2 \leq (1 + \varepsilon)|a|^2 + (1 + \varepsilon^{-1})|b|^2.$$

We will take $\varepsilon \rightarrow 0$ as $n \rightarrow \infty$, so we obtain

$$\|S(k/n + 1/2n)v(k)\|^2 \leq (1 + \varepsilon)\|S(k/n)v(k)\|^2 + O(\varepsilon^{-1}n^{-2}\|v(k)\|^2).$$

Similarly,

$$\begin{aligned} \|S(\ell/n + 1/2n)\widehat{v}\|^2 &\leq (1 + \varepsilon)\|S(\ell/n)\widehat{v}\|^2 + O(\varepsilon^{-1}n^{-2}\|\widehat{v}\|^2) \\ &= (1 + \varepsilon)\|S(\ell/n)\widehat{v}\|^2 + O(\varepsilon^{-1}n^{-1/2}\|v\|^2), \end{aligned}$$

where we used $\|\widehat{v}\| = n^{3/4}\|v\|$.

Thus,

$$(n(I + UM_nU^*)v, v) \leq (1 + \varepsilon)Q(v) + O(\varepsilon^{-1}n^{-1}\|v\|^2).$$

Similarly,

$$(n(I + UM_nU^*)v, v) \geq (1 + \varepsilon)^{-1}Q(v) - O(\varepsilon^{-1}n^{-1}\|v\|^2).$$

We set $\varepsilon = n^{-1/2}$ and put the inequalities together to get the statement of the proposition. \square

Recall that $Q(v) = (n(I - M_n)v, v)$. If in the statement of the proposition we take the minimum of both sides over all v with $\|v\| = 1$ we deduce that

$$n + n\lambda_n = (1 + O(n^{-1/2}))(n - n\lambda_1) + O(n^{-1/2}),$$

where λ_n is the bottom eigenvalue of M_n and λ_1 the top eigenvalue. Since $\lambda_1 = 1 - \mu_1/n + o(1/n)$, we have $n - n\lambda_1 = \mu_1 + o(1)$, and then $n + n\lambda_n = \mu_1 + o(1)$, and then $\lambda_n = -1 + \mu_1/n + o(1)$.

Using the minimax characterization of the eigenvalues we show similarly that $\lambda_{n-k+1} = -1 + \mu_k/n + o(1)$ for each k .

³This is easy to see directly.

5. A STOCHASTIC ARGUMENT

This section gives a bound on the largest eigenvalue of the matrix M_n using a probabilistic argument. By inspection,

$$M'_n = \frac{1}{3}I + \frac{2}{3}M_n$$

is a sub-stochastic matrix (with non-negative entries and row sums at most 1). Take M''_n as in (2), an $(n+1) \times (n+1)$ stochastic matrix corresponding to a Markov chain absorbing at 0. The first (Dirichlet) eigenvector has first entry 0 and its corresponding eigenvalue β^* is the top eigenvalue of M'_n . Thus

$$\beta = \frac{(\beta^* - \frac{1}{3})}{(2/3)} = \frac{3}{2}\beta^* - \frac{1}{2}$$

is the top eigenvector of M_n .

We will work in continuous time, thus, for any transition matrix M ,

$$M_t = \sum_{j=0}^{\infty} \frac{e^{-t} M^j t^j}{j!} = e^{t(M-I)}.$$

The matrix $L = I - M$, the opposite of the generator of the semigroup $\{M_t\}_{t \geq 0}$, has row sums zero, and non-positive off diagonal entries.⁴ If \mathbf{v} is a right eigenvector of M with eigenvalue β , then \mathbf{v} is an eigenvector of L with eigenvalue $1 - \beta$. A lower bound for the non trivial eigenvalues of L gives an upper bound for the eigenvalues of M . Throughout, we specialize to $L = I - M_n$, let λ^* be the lowest non-zero eigenvalue of L , and β the highest eigenvector of M_n .

Standard theory for absorbing Markov chains with all non-absorbing states connected shows that if τ is the time to first absorption, for any non absorbing state ξ , as t tends to infinity,

$$\lim_{t \rightarrow \infty} \frac{-\log P_{\xi}(\tau > t)}{t} = \lambda^*.$$

Thus an upper bound on β will follow from an upper bound on $P_{\xi}(\tau > t)$. Here is an outline of the proof. Begin by coupling the absorbing chain of interest with a simple random walk on $\mathcal{C}_n = \mathbb{Z}/(n\mathbb{Z})$. For a fixed b , let τ_b be the first time that the simple random walk travels $\pm b$ from its start. We derive the bound $P_{\xi}(\tau > \tau_b) \leq G_b$, where $G_b < 1$ is a particular constant described below. Define a sequence of stopping times τ_b^i as follows. $\tau_b^1 = \tau_b$, τ_b^2 is the first time following τ_b^1 that the walk travels $\pm b$, similarly define τ_b^m . By the strong law of large numbers, $\tau_b^m/m \rightarrow \mu_b = E(\tau_b^1)$ almost surely. Thus

$$P\{\tau > m\mu_b\} \cong P\{\tau > \tau_b^m\}.$$

Using the Markov property, $P\{\tau > \tau_b^m\} \leq G_b^m$. This implies there are positive c_1, c_2 with

$$P\{\tau > c_1 m \mu_b\} \leq G_b^m + e^{-c_2 m}.$$

⁴Since some of our readers (indeed some of our authors) may not be probabilists we insert the following note; given any matrix $L(x, y)$ with row sums zero and non-positive off diagonal entries one may construct a continuous time Markov process $W = (W_t)_{t \geq 0}$ as follows. Suppose $W_0 = w_0$ is fixed. The process stays at w_0 for an exponential time σ_0 with mean $1/|L(w_0, w_0)|$. (Thus $P\{\sigma_0 \geq t\} = e^{-tL(w_0, w_0)}$.) Then, choose $w_1 \neq w_0$ with probability $|L(w_0, w_1)|/L(w_0, w_0)$. Stay at w_1 for an exponential time σ_1 (with mean $1/L(w_1, w_1)$). Continue, choosing from $L(w_1, \cdot)/L(w_1, w_1)$.

In our problem, classical random walk estimates show $\mu_b \sim b^2$. We show, for $b = \sqrt{n}$, $G_{\sqrt{n}}$ is bounded away from one. Thus

$$P\{\tau > c_1 m \mu_b\} \leq 2 \max(G_b^m, e^{-c_2 m})$$

and

$$\frac{\log P\{\tau > c_1 m \mu_b\}}{m \mu_b} \leq \frac{c'}{\mu_b} = \frac{c''}{n}$$

for some $c', c'' < 0$. Backtracking gives the claimed bound in Theorem 3.

The argument is fairly robust—it works for a variety of diagonal entries. At the end of the proof, some additions are suggested which should give the right constant multiplying $\frac{1}{n}$.

We begin by constructing two processes. For as long as possible, general absorption rates will be used. Let $X = (x_t)_{t \geq 0}$ be the standard continuous time random walk on \mathbb{Z} with jump rates 1 between neighbors. Take $x_0 = 0$. Fix $b \in \mathbb{Z}_+$ and let τ_b be the first hitting time of $\{-b-1, b+1\}$:

$$(16) \quad \tau_b = \inf\{t \geq 0 : |x_t| = b+1\}$$

Let $\{u_x\}_{x \in \mathbb{Z}}$ be killing rates, e.g. arbitrary non-negative real numbers. Add a cemetery state ∞ to \mathbb{Z} . An absorbed process $\bar{x} = (\bar{x}_t)_{t \geq 0}$, behaving as x until it is absorbed at ∞ with the rates $\{u_x\}_{x \in \mathbb{Z}}$ can be constructed as follows: Let \mathcal{E} be an independent exponential random variable with mean 1. Define an absorption time $\bar{\tau} \in [0, \infty]$ by

$$\bar{\tau} = \inf_{t \geq 0} \left\{ \int_0^t \mu_{x_s} ds \geq \mathcal{E} \right\}.$$

As soon as $\{u_x\}_{x \in \mathbb{Z}}$ does not vanish identically, $\bar{\tau}$ is characterized by

$$(17) \quad \int_0^{\bar{\tau}} u_{x_s} ds = \mathcal{E}.$$

More simply, $\bar{x}_t = \begin{cases} x_t & \text{if } t < \bar{\tau} \\ \infty & \text{otherwise} \end{cases}$ for $0 \leq t < \infty$.

The two processes are defined on the same probability space as are $\bar{\tau}$ and τ_b . The first goal is to estimate the chance that $\bar{\tau} > \tau_b$ in terms of the given rates. Our bounds are crude but suffice for Theorem 3.

Proposition 6. *With notation as above, for any $b \geq 1$,*

$$(18) \quad P\{\bar{\tau} > \tau_b\} \leq \left(\frac{1}{1 + (b+1)^2 v_0 / 2} \prod_{k=1}^b \frac{1}{1 + (b+1)(b+1-k)v_k} \right)^{\frac{1}{b+1}}$$

with $v_k = \min(u_{-k}, u_k)$.

Note that the bound is achievable; if all $v_y = 0$ then both sides equal 1.

Proof. For any $k \in \mathbb{Z}$, $v_k \leq u_k$. Thus if τ is the stopping time defined in (17) with u_k replaced by v_k , $\tau \geq \bar{\tau}$. Therefore it is sufficient to bound $P\{\tau > \tau_b\}$ from above. Now, everything is symmetric about zero. Consider the process $Y = (Y_t)_{t \geq 0} = (|x_t|)_{t \geq 0}$. This is Markov with jump rates:

$$J(y, y') = \begin{cases} 2 & \text{if } y = 0, y' = 1 \\ 1 & \text{if } y = \mathbb{Z}_+ \text{ and } |y' - y| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\tau_b = \inf_{t \geq 0} \{Y_t = b + 1\}$. Define the family of local times associated to Y :

$$L_y(t) = \int_0^t \delta_y(Y_s) ds \text{ for } y \in \mathbb{Z}_+, t \geq 0,$$

where δ_y is the indicator function of y . For any $t \geq 0$,

$$\int_0^t v_{x_s} ds = \sum_{0 \leq y \leq b} v_y L_y(t).$$

This gives

$$\{\tau > \tau_b\} = \left\{ \sum_{0 \leq y \leq b} v_y L_y(\tau_b) < \mathcal{E} \right\}.$$

Taking expectations of both sides with respect to \mathcal{E}

$$(19) \quad P\{\tau > \tau_b\} = E \left\{ \exp \left(- \sum_{0 \leq y \leq b} v_y L_y(\tau_b) \right) \right\}$$

$$(20) \quad \leq \prod_{y=1}^b E \{ \exp(-(b+1)v_y L_y(\tau_b)) \}^{\frac{1}{b+1}}.$$

The last bound follows from Hölder's inequality (with $b + 1$ functions).

It is well known (see [23] or Claim 2.4 of [26] for the discrete time version) that for any y , $1 \leq y \leq b$, $L_y(\tau_b)$ is distributed as an exponential random variable with mean $(b + 1 - y)$ and $L_0(\tau_b)$ is exponential with mean $\frac{b+1}{2}$. (The process leaves zero twice as fast as it leaves other points.) Thus, for $1 \leq y \leq b$,

$$E\{\exp(-(b+1)v_y L_y(\tau_b))\} = \frac{(b+1-y)^{-1}}{(b+1-y)^{-1} + (b+1)v_y}.$$

$$E\{\exp(-(b+1)v_0 L_0(\tau_b))\} = \frac{1}{1 + (b+1)^2 v_0 / 2}.$$

This completes the proof of Proposition 6. □

The bound of Proposition 6 suggests introducing functions F_b, G_b on \mathbb{R}_+^{b+1} . Given by

$$(21) \quad F_b(\mathbf{v}) = \left\{ \frac{1}{1 + (b+1)^2 v_0 / 2} \prod_{l=1}^b \frac{1}{1 + (b+1)(b+1-l)v_l} \right\}^{\frac{1}{b+1}}$$

$$(22) \quad G_b(\mathbf{v}) = \left\{ \prod_{k=0}^b \frac{1}{1 + (b+1)(b+1-k)v_k / 2} \right\}^{\frac{1}{(b+1)}}.$$

They have the following crucial monotonicity properties: say that $\mathbf{v}, \mathbf{v}' \in \mathbb{R}_+^{b+1}$ satisfies $\mathbf{v} \leq \mathbf{v}'$ if this is true coordinate-wise. For $\mathbf{v} \in \mathbb{R}_+^{b+1}$, let $\bar{\mathbf{v}}$ be the non-decreasing rearrangement of v . Then

$$(23) \quad F_b(\mathbf{v}) \leq G_b(\mathbf{v})$$

$$(24) \quad \mathbf{v} \leq \mathbf{v}' \Rightarrow G_b(\mathbf{v}) \geq G_b(\mathbf{v}')$$

$$(25) \quad G_b(\mathbf{v}) \leq G_b(\bar{\mathbf{v}})$$

Return now to the process underlying Theorem 3 (still keeping the extinction rates general.) Let $z = (z_t)_{t \geq 0}$ be defined on $\mathbb{Z}/n\mathbb{Z}$; it jumps to nearest neighbors at rate 1 and is killed with rates $u = (u_\xi)_{\xi \in \mathbb{Z}/n\mathbb{Z}}$. Suppose $z_0 = \xi$. Let v_p , $0 \leq p \leq n-1$ denote the non-decreasing rearrangement of \mathbf{u} . Let τ be the absorption time of z . Fix b , $0 \leq b \leq n/2 - 1$ and let

$$\tau_b = \inf_{t > 0} \{z_t \in \{\xi - b - 1, \xi + b + 1\}\}.$$

Proposition 6 in conjunction with properties (23), (24), and (25), imply that for any \mathbf{u} , with $G_b(\mathbf{u})$ depending only on the first b coordinates of \mathbf{u} ,

$$(26) \quad P_\xi[\tau > \tau_b] \leq G_b(\mathbf{u}).$$

Note that the upper bound is independent of ξ .

Introduce a sequence ξ_i of further stopping times: $\xi_1 = \xi_b$, and if ξ_m has been constructed,

$$(27) \quad \xi_{m+1} = \inf\{t > \xi_m : z_t \in \{z_{\xi_m} - b - 1, z_{\xi_m} + b + 1\}\}$$

Informally speaking these stopping times end up being good. Because they cannot be larger than τ , as in the previous treatment of a random walk on $\mathbb{Z}/n\mathbb{Z}$ coinciding with z_t up to the absorption time, then ξ_m are (almost surely) finite for all m and the strong law of large numbers gives:

$$\lim_{m \rightarrow \infty} \frac{\xi_m}{m} \rightarrow \mu_b = (b+1)^2$$

where $\mu_b = E\{\xi_1\} = (b+1)^2$ from the Classical Gambler's Ruin (see Chapter 14 of [16]).

This suggests that, for m large, the quantities

$$P_\xi[\tau > m\mu_b] \text{ and } P_\xi[\tau > \xi_m]$$

should behave similarly. Of course, care must be taken because τ and ξ_m are not independent. To proceed, we use a large deviations bound for ξ_m .

Proposition 7. *For ξ_m defined in (27), there are positive constants, c_1, c_2 , independent of b and n such that for all $m \geq 1$,*

$$P[\xi_m > c_1 m \mu_b] \leq e^{-c_2 m}.$$

Proof. Observe first that this is simply a large deviations bound for the first hitting time of the simple random walk (16) so that n does not enter. The law of ξ_1 is well known (see [22], [12], and [17]). It can be represented as a sum of $b+1$ independent exponential variables with means a_1, a_2, \dots, a_{b+1} given by:

$$a_k^{-1} = 2 \left(1 - \cos \left(\frac{\pi(2k-1)}{2(b+1)} \right) \right).$$

Thus for $\theta \in (0, a_1)$

$$E[e^{\theta \xi_1}] = \prod_{k=1}^{b+1} \frac{a_k}{a_k - \theta}.$$

By simple calculus, there is $c > 0$ such that for all $a \in (0, \frac{1}{2}]$, $-\log(1-a) \leq ca$. Thus, for $\theta \in (0, a_1/2]$,

$$E(e^{\theta \xi_1}) \leq e^{c\theta \sum_{n=1}^{b+1} \frac{1}{a_n}}.$$

Taking $\theta = \frac{a_1}{2}$,

$$E[e^{a_1 \xi_1 / 2}] \leq e^{c \sum_{n=0}^{b+1} \frac{a_1}{2a_n}}.$$

Note that a_n is of order $(n/b)^2$ so the right side of the last inequality is bounded uniformly in b , say by $k > 1$. Now, ξ_n is a sum of m i.i.d. random variables so for any $c_1 > 0$

$$P[\xi_m > c_1 m \mu_b] \leq e^{-\frac{c_1 a_1 m \mu_b}{2}} E[e^{a_1 \xi_1 / 2}]^m \leq e^{-m(c_1 a_1 \mu_b) / 2 - \log k}.$$

Since $\mu_b = (b+1)^2$, $a_1 \mu_b$ can be bounded below by a constant $\epsilon > 0$, uniformly in $b \in \mathbb{N}$. Thus if $c_1 = 4 \log k / \epsilon$, the claimed bound holds with $c_2 = \log(k)$. \square

We can now set up a bound for the top eigenvalue. Working on $\mathbb{Z}/n\mathbb{Z}$ but still with general absorption rates:

$$(28) \quad P_\xi[\tau > c_1 m \mu_b] = P_\xi[\tau > c_1 m \mu_b; \xi_m \leq c_1 m \mu_b] + P_\xi[\tau > c_1 m \mu_b, \xi_m > c_1 m \mu_b]$$

$$(29) \quad \leq P_\xi[\tau > \xi_m] + P_\xi[\xi_m > c_1 m \mu_b]$$

$$(30) \quad \leq G_b^m(\mathbf{v}) + e^{-c_2 m}.$$

It follows that

$$\lambda^* = - \lim_{m \rightarrow \infty} \frac{1}{c_1 m \tau_b} \log(P_\xi[\tau > c_1 m \tau_b]) \geq \frac{1}{c_1 \tau_b} \min\{c_2, -\log(G_b(x))\}.$$

Since $\mu_b = (b+1)^2$, proving that with b of order \sqrt{n} , $-\log(G_b(x))$ is bounded below by a positive constant, uniformly in n , will complete the proof.

Up to now, the kill rates \mathbf{u} have been general. Specialize now to the rates for the matrix M'' with any scrambling of its diagonal. The vector $\bar{\mathbf{v}}$ is given by the $b+1$ entries of:

$$0, \frac{1}{3} \left(1 - \cos\left(\frac{2\pi}{n}\right)\right), \frac{1}{3} \left(1 - \cos\left(\frac{2\pi}{n}\right)\right), \frac{1}{3} \left(1 - \cos\left(\frac{4\pi}{n}\right)\right), \\ \frac{1}{3} \left(1 - \cos\left(\frac{4\pi}{n}\right)\right), \dots, \frac{1}{3} \left(1 - \cos\left(\frac{2\pi \lfloor \frac{n}{2} \rfloor}{n}\right)\right)$$

From the definition of G_b at (21) with $b = \lfloor \sqrt{n} \rfloor$, a Riemann sum approximation gives

$$\lim_{n \rightarrow \infty} G_b(\bar{\mathbf{v}}) = e^{-\frac{\pi^2}{24}}.$$

Indeed,

$$\begin{aligned} -\log(G_b(\bar{\mathbf{v}})) &= \frac{1}{b+1} \sum_{y=0}^{b-1} \log\left(1 + (b+1)(b+1-y) \left(1 - \cos\left(\frac{2\pi}{n} \left\lfloor \frac{y+1}{2} \right\rfloor\right)\right)\right) \\ &\sim \frac{2\pi^2}{n^2} \frac{1}{b} \sum_{y=0}^{b-1} (b+1)(b+1-y) \left[\frac{y+1}{2}\right]^2 \\ &\sim \frac{\pi^2 b^4}{2n^2} \frac{1}{b} \sum_{y=0}^{b-1} \left(1 - \frac{y}{b}\right) \left(\frac{y}{b}\right)^2 \\ &\sim \frac{\pi^2 b^4}{2n^2} \int_0^1 (1-y)y^2 dy \\ &\sim \frac{\pi^2}{24}. \end{aligned}$$

Combining the pieces, we use β for the highest eigenvalue of M_n and thus $\beta = \frac{3}{2}\beta^* - \frac{1}{2}$. Using this notation, we have shown that $\frac{c}{n} \leq \lambda^* = 1 - \left(\frac{1}{3} + \frac{2}{3}\beta\right)$. Thus $\beta \leq 1 - \frac{3c}{2n}$. This completes the argument and ends the proof of Theorem 3.

Remarks. The above argument can be modified to handle quite general diagonal elements (in particular $\cos\left(\frac{2\pi aj}{n}\right)$, $0 \leq j \leq N-1$, needed for the application to the Heisenberg random walk). Indeed, for $a = o(n)$, the argument goes through with no essential change with $b = \sqrt{\frac{n}{a}}$ to show that with diagonal entries $\cos\left(\frac{2\pi aj}{n}\right)$, $0 \leq j \leq n-1$, the eigenvalue bound $1 - \frac{ca}{n}$ holds (with $c > 0$ independent of n and a).

The use of Hölder's inequality in (19) is crude. The joint distribution of the local times of birth and death processes is accessible (see [23]). We hope this can be used to give sharp results for the constant. Finally we note that the approach to bound β via an associated absorbing Markov chain was used in [5]. There, a geometric path argument was used to complete the analysis. This gave cruder bounds ($\beta \leq 1 - \frac{c}{n^{\frac{3}{4}}}$) but the argument worked for diagonal entries $\cos\left(\frac{2\pi a\xi}{n}\right)$ for any $1 \leq a \leq \frac{n}{2}$ as well as negative eigenvalues.

6. A RANDOM WALK ON THE AFFINE GROUP (MOD p)

Let \mathcal{A}_p be the affine group (mod p). Here, p is prime and elements of \mathcal{A}_p can be represented as pairs (a, b) , $1 \leq a \leq p-1$, $0 \leq b \leq p-1$

$$(a_1, b_1)(a_2, b_2) = (a_1 a_2, a_1 b_2 + b_1).$$

All entries are taken mod p . Fix a generator g of the multiplicative group. Let

$$S = \{(1, 0), (1, 1), (1, -1), (g, 0), (g^{-1}, 0)\}.$$

Set

$$(31) \quad Q(h) = \begin{cases} \frac{1}{5} & \text{if } h \in S \\ 0 & \text{else.} \end{cases}$$

Convolution powers of Q converge to the uniform distribution $U(h) = \frac{1}{p(p-1)}$. We use the representation theory of \mathcal{A}_p and the analytic results of previous sections to show that order p^2 steps are necessary and sufficient for convergence.

Theorem 5. *With definitions above, there are positive universal constants c_1, c_2 , and c_3 such that for all primes p and $k \geq 1$*

$$c_1 e^{-c_2 \frac{k}{p^2}} \leq \|Q^{*k} - U\|_{TV} \leq c_3 e^{-\frac{c_2 k}{p^2}}.$$

Proof. By the usual Upper Bound Lemma (see [8], Chapter 3):

$$4\|Q^{*k} - U\|_{TV} \leq \sum_{\rho \neq 1} d_\rho \|\hat{Q}(\rho)^k\|^2.$$

Here, the sum is over nontrivial irreducible representations ρ of \mathcal{A}_p , d_ρ is the dimension of ρ , $\hat{Q}(\rho) = \sum_h Q(h)\rho(h)$ and the norm on the right is the trace norm. There are $p-1$ one dimensional irreducible representations indexed by $\alpha \in \{1, 2, \dots, p-1\}$.

$$(32) \quad \rho_\alpha(a, b) = e^{2\pi i \alpha \sigma(a)/(p-1)}.$$

The methods above can be applied to other walks on other groups. While we won't carry out the details here, we briefly describe two further examples and point to our companion paper [4] for more.

Example 2 (Borel Subgroup of $SL_2(\mathbb{F}_p)$). Let G be the 2×2 matrices of the form:

$$\begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \quad a \in \mathbb{F}_p^*, b \in \mathbb{F}_p \leftrightarrow (a, b).$$

A minimal generating set (with the identity) is

$$S = \{\text{id}, (g, 0), (g^{-1}, 0), (1, 1), (1, -1)\}, \quad g \text{ a generator of } \mathbb{F}_p^*.$$

The group has order $p(p-1)$ with $p-1$ 1-dimensional representations and 4 representations of dimension $(p-1)/2$. They are explicitly described in [6] p. 67. The Fourier analysis of the measure Q supported on S is almost the same as the analysis for the affine group. The results are that order p^2 steps are necessary and sufficient for convergence to the uniform distribution.

Example 3 ($M(p^3)$). There are two nonabelian groups of order p^3 : the Heisenberg group discussed above and $M(p^3)$. See [29] Chapter 4 Section 4. One description of the latter is:

$$M(p^3) = \{(a, b) : a \in \mathbb{Z}_p, b \in \mathbb{Z}_p^2\}, \quad (a, b)(a', b') = (a + a', a * b' + b)$$

with $a * b = (1 + ap)b \pmod{p^2}$. This group has the same character table as $H_1(p)$. It thus has p^2 1-dimensional representations and $p-1$ representations of dimension p . A minimal generating set (for odd p the identity is not needed to take care of parity problems) is

$$S = \{(1, 0)(-1, 0)(0, 1), (0, -1)\}.$$

The Fourier transforms of the associated Q at the p dimensional representations have the same form as the matrices in (1) with diagonal elements

$$2 \cos \left(\frac{2\pi c}{p^2} (1 + jp) \right), \quad 0 \leq j \leq p-1$$

where $1 \leq c \leq p-1$ is fixed (for the c th representation). We have not carried out the details, but, as shown in [10], it is known that order p^2 steps are necessary and sufficient for convergence.

7. EIGENVALUES IN THE BULK

Consider the matrix $M_n(a)$ as in (1) with

$$\cos \left(\frac{2\pi ja}{m} \right), \quad 0 \leq j \leq n-1$$

as the diagonal elements. The sections above give bounds on the largest and smallest eigenvalues. It is natural to give bounds for the empirical measure of all the eigenvalues. This is straightforward, using a theorem of Kac-Murdock-Szegö from [21]. We use the elegant form of Trotter [30]. If $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of $M_n(a)$, let

$$\Lambda_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$$

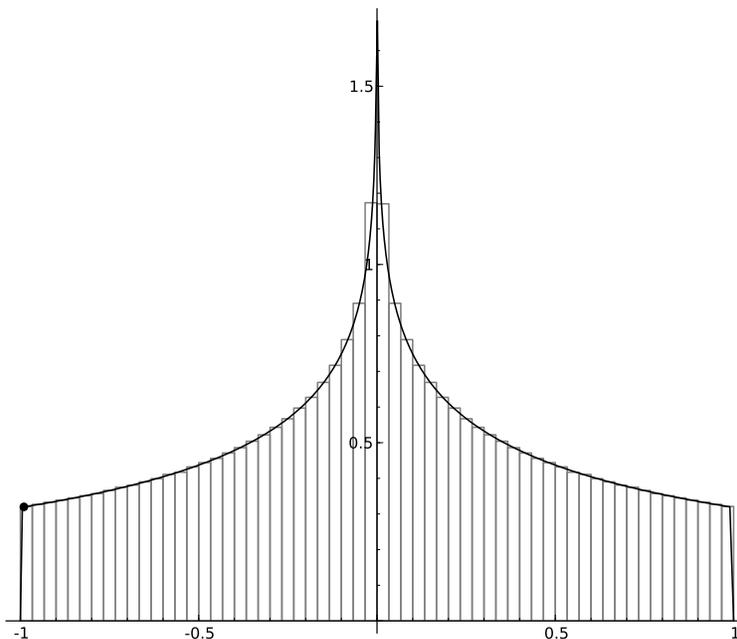


FIGURE 1. The curve shows the eigenvalues predicted by f_2 , while the histogram gives the distribution of the actual eigenvalues of $M_{10,000}(= M_{10,000}(1))$. Note that the curve has a very extreme, but finite slope around -1 and 1 . For example, although it is clear that $f_2(-1) = 0$, the small point on the left of the picture corresponds to $(-0.99, f_2(-0.99)) \approx .32$

be the associated empirical measure. To describe the limit let

$$(33) \quad f_2(x) = \begin{cases} \frac{2}{\pi(1+|x|)} F_{2,1} \left(\frac{1}{2}, \frac{1}{2}; 1; \left(\frac{1-|x|}{1+|x|} \right)^2 \right) & -1 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$$

where $F_{2,1}$ is the hypergeometric function. Let μ_2 be the associated measure. Distance between Λ_n and μ_2 is measured in the d_2 Wasserstein distance:

$$d_2^2(\Lambda_n, \mu_2) = \sup E|W - Z|^2 \text{ with } W \sim \Lambda_n, Z \sim \mu_2.$$

Theorem 6. *Let Λ_n be the empirical measure of the matrix $M_n(a)$ with $1 \leq a \leq n-1$. Let μ_2 be defined by (33). Then, with a fixed, as $n \rightarrow \infty$,*

$$d_2(\Lambda_n, \mu_2) \rightarrow 0.$$

See Figure 1 for an example.

Remark. We have not seen a way to use this kind of asymptotics to bound the rate of convergence of a random walk. Indeed our limit theorem shows that the distribution of the bulk does not depend on a while previous results show the extreme eigenvalues crucially depend on a .

Proof. Trotter's version of the Kac-Murdock-Szegö theorem applies to M_n . If

$$\sigma(x, y) = \cos(2\pi ax) + \cos(2\pi y) \quad 0 \leq x, y \leq 1,$$

consider σ as a random variable on $[0, 1]^2$, endowed with the Lebesgue measure. This has distribution $\cos(2\pi aU_1) + \cos(2\pi U_2)$ where U_1 and U_2 are independent uniform on $[0, 1]$. An elementary calculation shows that $\cos(2\pi aU)$ has an arc-sine density $f(x)$ no matter what the integer a is.

$$(34) \quad f(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & -1 \leq x \leq 1 \\ 0 & \text{else} \end{cases}.$$

Trotter shows that the empirical measure is close to μ_2 , the distribution of σ . It follows that the empirical measure of the eigenvalues has limiting distribution the law of $(X + Y)/2$ where X and Y are independent with density $f(x)$. This convolution has density

$$(35) \quad f_2(x) = \begin{cases} \frac{2}{\pi^2} \int_{\max\{-1, 2x-1\}}^{\min\{1, 2x+1\}} \frac{1}{\sqrt{((2x-y)^2-1)(y^2-1)}} dy & -1 \leq x \leq 1 \\ 0 & \text{else} \end{cases}.$$

The argument below shows that this integral is in fact

$$(36) \quad f_2(x) = \begin{cases} \frac{2}{\pi(1+|x|)} F_{2,1} \left(\frac{1}{2}, \frac{1}{2}; 1; \left(\frac{1-|x|}{1+|x|} \right)^2 \right) & -1 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$$

The integral in (34) is in fact a well known integral in a different guise. Let $0 \leq k \leq 1$. Define

$$(37) \quad K(k) = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}.$$

This is a complete elliptic integral and equals

$$\frac{\pi}{2} F_{2,1} \left(\frac{1}{2}, \frac{1}{2}; 1, k^2 \right).$$

(See Section 22.301 of [32].) For ease of notation, we will evaluate

$$f_3(x) = \int_{\max(-1, x-1)}^{\min(1, x+1)} \frac{dt}{\sqrt{(1-t^2)(1-(x-t)^2)}}$$

for $|x| \leq 2$ Making the variable change $t \rightarrow t + \frac{x}{2}$ it becomes

$$\int_{\max(-\frac{x}{2}-1, \frac{x}{2}-1)}^{\min(-\frac{x}{2}+1, \frac{x}{2}+1)} \frac{dt}{\sqrt{h(t)}}$$

where

$$h(t) = (a^2 - t^2)(b^2 - t^2), \quad a = 1 + \frac{x}{2}, \quad b = 1 - \frac{x}{2}.$$

This is an even function of x so it is enough to consider when $x \geq 0$. Then we need to evaluate

$$\int_{-b}^b \frac{dt}{\sqrt{h(t)}}.$$

Make the variable change $t \rightarrow bt$ and the integral becomes

$$\int_{-1}^1 \frac{dt}{\sqrt{(1-t^2)(a^2-b^2t^2)}} = \frac{2}{a}K(k), \quad k = \frac{b}{a}.$$

The factor of 2 comes from the fact that we are integrating an even function from -1 to 1 , whereas in (37) the integral is from 0 to 1. Thus

$$f_3(x) = \frac{2\pi}{2+|x|} F_{2,1} \left(\frac{1}{2}, \frac{1}{2}; 1; \left(\frac{2-|x|}{2+|x|} \right)^2 \right).$$

Sending $x \rightarrow 2x$ and multiplying by the appropriate constant, we have that the integral in (34) is in fact

$$\frac{2}{\pi(1+|x|)} F_{2,1} \left(\frac{1}{2}, \frac{1}{2}; 1; \left(\frac{1-|x|}{1+|x|} \right)^2 \right)$$

□

Remark. [27] gives a similar expression for the sum of two general beta variables.

Acknowledgment. As this work progressed, we received useful comments from Florin Boca, Ben Bond, Bob Guralnick, Susan Holmes, Marty Isaacs, Evita Nestoridi, Jim Pitman, Laurent Saloff-Coste, and Thomas Strohmer. We offer thanks for the remarks of our very helpful reviewer.

REFERENCES

- [1] Cédric Béguin, Alain Valette, and Andrzej Zuk. On the spectrum of a random walk on the discrete Heisenberg group and the norm of Harper’s operator. *J. Geom. Phys.*, 21(4):337–356, 1997.
- [2] Rajendra Bhatia. Linear algebra to quantum cohomology: the story of Alfred Horn’s inequalities. *Amer. Math. Monthly*, 108(4):289–318, 2001.
- [3] Florin P. Boca and Alexandru Zaharescu. Norm estimates of almost Mathieu operators. *J. Funct. Anal.*, 220(1):76–96, 2005.
- [4] D. Bump, P. Diaconis, A. Hicks, L. Miclo, and H. Widom. Characters and super characters for step two nilpotent groups with applications to random walks. to appear.
- [5] D. Bump, P. Diaconis, A. Hicks, L. Miclo, and H. Widom. An Exercise (?) in Fourier Analysis on the Heisenberg Group. *ArXiv e-prints*, February 2015.
- [6] Charles W. Curtis and Irving Reiner. *Representation theory of finite groups and associative algebras*. AMS Chelsea Publishing, Providence, RI, 2006. Reprint of the 1962 original.
- [7] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.
- [8] Persi Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.
- [9] Persi Diaconis. Patterned matrices. In *Matrix theory and applications (Phoenix, AZ, 1989)*, volume 40 of *Proc. Sympos. Appl. Math.*, pages 37–58. Amer. Math. Soc., Providence, RI, 1990.
- [10] Persi Diaconis. Threads through group theory. In *Character theory of finite groups*, volume 524 of *Contemp. Math.*, pages 33–47. Amer. Math. Soc., Providence, RI, 2010.
- [11] Persi Diaconis, Sharad Goel, and Susan Holmes. Horseshoes in multidimensional scaling and local kernel methods. *Ann. Appl. Stat.*, 2(3):777–807, 2008.
- [12] Persi Diaconis and Laurent Miclo. On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.*, 22(3):558–586, 2009.
- [13] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.

- [14] Bradley W. Dickinson and Kenneth Steiglitz. Eigenvectors and functions of the discrete Fourier transform. *IEEE Trans. Acoust. Speech Signal Process.*, 30(1):25–31, 1982.
- [15] David L. Donoho and Philip B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, 49(3):906–931, 1989.
- [16] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.
- [17] James Allen Fill. The passage time distribution for a birth-and-death chain: strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.*, 22(3):543–557, 2009.
- [18] William Fulton. Eigenvalues, invariant factors, highest weights, and Schubert calculus. *Bull. Amer. Math. Soc. (N.S.)*, 37(3):209–249 (electronic), 2000.
- [19] D.J. Griffiths. *Introduction to Quantum Mechanics.* Pearson international edition. Pearson Prentice Hall, 2005.
- [20] Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, second edition, 2013. p. 181.
- [21] M. Kac, W. L. Murdock, and G. Szegő. On the eigenvalues of certain Hermitian forms. *J. Rational Mech. Anal.*, 2:767–800, 1953.
- [22] Julian Keilson. Log-concavity and log-convexity in passage time densities of diffusion and birth-death processes. *J. Appl. Probability*, 8:391–398, 1971.
- [23] John T. Kent. The appearance of a multivariate exponential distribution in sojourn times for birth-death and diffusion processes. In *Probability, statistics and analysis*, volume 79 of *London Math. Soc. Lecture Note Ser.*, pages 161–179. Cambridge Univ. Press, Cambridge, 1983.
- [24] Gregory F. Lawler and Alan D. Sokal. Bounds on the L_2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.*, 309(2):557–580, 1988.
- [25] M. L. Mehta. Eigenvalues and eigenvectors of the finite Fourier transform. *J. Math. Phys.*, 28(4):781–785, 1987.
- [26] Y. Peres and P. Sousi. Total variation cutoff in a tree. *ArXiv e-prints*, July 2013.
- [27] T.G. Pham and N. Turkkan. Reliability of a standby system with beta-distributed component lives. *Reliability, IEEE Transactions on*, 43(1):71–75, Mar 1994.
- [28] T. Strohmer and T. Wertz. Almost Eigenvalues and Eigenvectors of Almost Mathieu Operators. *ArXiv e-prints*, January 2015.
- [29] Michio Suzuki. *Group theory. II*, volume 248 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1986. Translated from the Japanese.
- [30] Hale F. Trotter. Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegő. *Adv. in Math.*, 54(1):67–82, 1984.
- [31] Wang, Y.Y., Pannetier, B., and Rammal, R. Quasiclassical approximations for almost-mathieu equations. *J. Phys. France*, 48(12):2067–2079, 1987.
- [32] E. T. Whittaker and G. N. Watson. *A course of modern analysis.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1996. An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions, Reprint of the fourth (1927) edition.

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, 450 SERRA MALL, BLDG. 380, STANFORD, CA 94305-2125, USA

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, 450 SERRA MALL, BLDG. 380, STANFORD, CA 94305-2125, USA

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, 450 SERRA MALL, BLDG. 380, STANFORD, CA 94305-2125, USA

INSTITUT DE MATHÉMATIQUES DE TOULOUSE, UNIVERSITÉ PAUL SABATIER, 118 ROUTE DE NARBONNE, F-31062 TOULOUSE CEDEX 9, FRANCE.

UC SANTA CRUZ, DEPARTMENT OF MATHEMATICS, SANTA CRUZ, CA 95064, USA