PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 42, Fasc. 1 (2022), pp. 97–108
Published online 20.5.2022
doi:10.37190/0208-4147.00032

NO CUTOFF FOR CIRCULANTS: AN ELEMENTARY PROOF

BY

AARON ABRAMS (LEXINGTON, VA), ERIC BABSON (DAVIS, CA),
HENRY LANDAU (MURRAY HILL, NJ),
ZEPH LANDAU (BERKELEY, CA), AND JAMIE POMMERSHEIM (PORTLAND, OR)

Abstract. We give an elementary proof of a result due to Diaconis and Saloff-Coste (1994) that families of symmetric simple random walks on Cayley graphs of abelian groups with a bound on the number of generators never have sharp cutoff. Here convergence to the stationary distribution is measured in the total variation norm. This is a situation of bounded degree and no expansion; sharp cutoff (or the cutoff phenomenon) has been shown to occur in families such as random walks on a hypercube (Diaconis, 1996) in which the degree is unbounded as well as on a random regular graph where the degree is fixed, but there is expansion (Diaconis and Saloff-Coste, 1993).

2020 Mathematics Subject Classification: 60B15, 05C81, 20K01.

Key words and phrases: cutoff, mixing time, random walk, abelian groups.

1. INTRODUCTION

The notion of cutoff was introduced by Aldous and Diaconis [1], [2] to describe a phenomenon that is sometimes observed in the convergence behavior of random walks. Informally, a family of walks exhibits sharp cutoff if the evolving distributions of the walks spend a relatively long time far from the stationary distribution before quickly converging to it.

In the same work, Aldous and Diaconis conjectured that sharp cutoff should occur for random walks on Cayley graphs of groups whenever the number of generators is sufficiently large (see also [9] for a discussion of this conjecture). This conjecture has been shown to hold in a number of settings. Interesting recent work of Hermon and Olesker-Taylor [8] considers abelian groups, showing that if the number of generators becomes unbounded, then the cutoff phenomenon typically does occur. In a different paper [10], the same authors examined groups of upper triangular matrices, showing that cutoff also typically occurs for walks on these nonabelian groups when the number of generators grows. In a more general con-

text, Salez [14] recently showed that cutoff is a consequence of a certain non-negative curvature condition, which is often satisfied in the case of Cayley graphs of abelian groups. See the paper [8] and the references therein for a further summary of existing related results.

What about the behavior for walks with a small number of generators? The important work of Diaconis and Saloff-Coste [7] established that families of walks on groups exhibiting so-called *moderate growth* do not have sharp cutoff. Families of nilpotent groups with a bounded number of generators were shown to have moderate growth and therefore there is no cutoff for the corresponding families of walks. Later, Hough [11] re-proved this in the special case of cyclic groups of prime order, but with a quantitatively sharper result about the transition to stationarity.

In this paper, we give a new elementary proof that there is no cutoff for abelian groups with a bounded number of generators. The setup is as follows. We consider a finite abelian group G equipped with a generating set $\{a_i\}_{1 \leqslant i \leqslant r}$ of size r. We analyze the cutoff (or lack thereof) behavior of the random walk given by applying one of the elements $\{0, \pm a_i\}$ with equal probability; we will call this a *type* r walk.

THEOREM 1 ([7], [15]). *No family of walks all of the same type has sharp cutoff.*

What is new here is the proof method. The more general method of [7] in the case of moderate growth rests on a relation between the growth rate of the set of words of a fixed length in the generators and the rate at which the walk approaches the uniform distribution. Showing that abelian (and more generally nilpotent) groups satisfy moderate growth uses the same sort of "doubling" property employed by Gromov in his polynomial growth theorem.

In contrast, our proof focuses on the Fourier domain, connecting cutoff to eigenvalues of the walk and then analyzing the distribution of these eigenvalues. Before going into the details, we informally outline the approach below. See also [15, Section 8] for more on the convergence rate of type r walks.

2. FOURIER APPROACH TO CUTOFF

Sharp cutoff is defined as follows. If A is an irreducible symmetric Markov matrix with unique stationary distribution $\mathbf{v}_0$ (so that $A\mathbf{v}_0 = \mathbf{v}_0$ and $|\mathbf{v}_0|_1 = 1$) and $\mathbf{x}_0 = (1, 0, \ldots, 0)$, we write

$$d_A(t) = |A^t \mathbf{x}_0 - \mathbf{v}_0|_1$$

for the distance to the stationary distribution at time t, and

$$t_A(d) = \max\{t \mid d_A(t) \geqslant d\}$$

for the time it takes to get within distance d of the stationary distribution.

DEFINITION 2. A family $\{A_i\}$ of irreducible symmetric Markov matrices has *sharp cutoff* if

$$\lim_{n\to\infty} \frac{t_{A_n}(\epsilon)}{t_{A_n}(1-\epsilon)} = 1$$

for every $\epsilon \in (0, 1/2)$.

(See also [15, Definition 3.3].)

Our reasoning about the notion of sharp cutoff is inspired by the following two extreme scenarios. Consider on the one hand a family $\{A_n\}$ of Markov matrices with n eigenvalues $1, 1-\epsilon, 0, \dots 0$, and on the other hand a family $\{B_n\}$ with n eigenvalues $1, 1-\epsilon, \dots, 1-\epsilon$. For fixed n, beginning with a vector $\mathbf{x}_0 = (1, 0, \dots, 0)$ we are interested in how quickly the vectors $A_n^i\mathbf{x}_0$ and $B_n^i\mathbf{x}_0$ approach the stationary distribution $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ when measured in ℓ^1 norm. For the moment, let us imagine that we can take $\sqrt{n}$ times the ℓ^2 norm as a proxy for the ℓ^1 norm; in general this substitution is not rigorous but it does hold quite tightly in many cases (see [5, Chapter 3]), and doing it here allows us to change basis and analyze our scenarios in the diagonal basis of the Markov matrix (for a treatment of the ℓ^2 version of the problem see [3]).

Denote the image of $\mathbf{x}_0$ in the diagonal basis by $\mathbf{w}$; the image of the stationary distribution $\left(\frac{1}{n}, \dots \frac{1}{n}\right)$ is the first eigenvector $\left(\frac{1}{\sqrt{n}}, 0, \dots, 0\right)$.

Then we have for $i \geqslant 1$,

$$a^i \coloneqq A_n^i(\mathbf{w}) = (w_1, (1-\epsilon)^i w_2, 0, \dots, 0).$$

Under the assumption that the $|w_i|$ are each on the order of $\frac{1}{\sqrt{n}}$, the ℓ^2 distance to the stationary distribution is about $(1-\epsilon)^i w_2$, and our proxy measure is already within a constant of $\left(\frac{1}{\sqrt{n}}, 0, \dots, 0\right)$ at time $i = 1$. Further iteration moves closer to stationary at basically the constant multiplicative rate $1-\epsilon$ per time step. This does not display sharp cutoff, since A_n moves quickly to within a constant distance of the stationary. (See Definition 2.)

In contrast, for B_n, we see that

$$b^i \coloneqq B_n^i(\mathbf{w}) = (w_1, (1-\epsilon)^i w_2, (1-\epsilon)^i w_3, \dots, (1-\epsilon)^i w_n)$$

has ℓ^2 distance $(1-\epsilon)^i \sqrt{\sum_{i=2}^n |w_i|^2} \approx (1-\epsilon)^i$ which only gets within a constant of stationary (in our proxy measure) at a time i for which $(1-\epsilon)^i\sqrt{n}$ is constant, i.e., for $i = O(\log n)/\log \frac{1}{1-\epsilon}$. Once i exceeds this time, b^i moves towards the stationary distribution at the same rate per step as a^i. Thus B_n spends a long time ($O(\log n)$) getting close to the stationary relative to the time spent improving that closeness, meaning that B_n does exhibit sharp cutoff.

These examples illustrate the perspective that sharp cutoff is a criterion that captures those scenarios where the eigenvalues of the process drop off "slowly enough." This perspective is made rigorous in our context via Lemma 3, which

uses a standard argument to relate sharp cutoff to the decay of eigenvalues. As we will then show, the eigenvalues for the type r processes we consider here drop off too quickly for sharp cutoff to occur.

Our argument rests on one key idea which we now describe. The fact that our groups G are abelian allows us to describe the eigenvalues of the process explicitly via the one-dimensional representations of G. Rather than analyzing directly the distribution of the sizes of these eigenvalues, we observe that these eigenvalues correspond in a nice way to certain vectors in an r-dimensional lattice. We show that the sizes of the eigenvalues fall off more quickly than the set $e^{-c|x|^2}$ for x ranging over the lattice, which despite being an infinite set is more readily summable. The lengths of lattice vectors shrink quickly enough to rule out sharp cutoff via this comparison.

Most of the work to prove Theorem 1 already shows up in the cyclic case, which we prove in Section 4. The general abelian case, which is notationally more complicated, is completed in Section 5.

3. RELATING RATE OF CONVERGENCE TO EIGENVALUES

We start with a general lemma relating the distance to the stationary distribution at time t to the eigenvalues of the Markov matrix A corresponding to a random walk on a Cayley graph.

We write $\{\lambda_k\} \subseteq (-1, 1]$ for the eigenvalues of A with

$$\lambda_0 = 1 \quad \text{and} \quad |\lambda_m| = \max_{k \neq 0} |\lambda_k| < 1.$$

LEMMA 3. *Given a transition matrix A for a random walk on a Cayley graph for the abelian group G, we have*

$$\lambda_m^{2t} \leqslant d_A^2(t) \leqslant \sum_{k \neq 0} \lambda_k^{2t}. \tag{3.1}$$

Proof. For the left inequality note that $A = A^*$ is self-adjoint and the stationary distribution is $\mathbf{v}_0 = \frac{1}{n}\mathbf{1}$. Write $\mathbf{v}_m$ for the eigenvector with $A\mathbf{v}_m = \lambda_m \mathbf{v}_m$ and $|\mathbf{v}_m|_1 = 1$. Since A is a transition matrix for a random walk on a Cayley graph for G, it commutes with rotation (action by G). Thus $\mathbf{v}_m$ is an eigenvector for rotation and hence all entries of $\mathbf{v}_m$ have the same norm, which by the normalization is $|\mathbf{v}_m|_\infty = \frac{1}{n}$. Since $d_A(t) = \max_{\mathbf{v}} \frac{\langle A^t \mathbf{x}_0 - \mathbf{v}_0, \mathbf{v}\rangle}{|\mathbf{v}|_\infty}$ this gives

$$d_A(t) \geqslant n|\langle A^t \mathbf{x}_0 - \mathbf{v}_0, \mathbf{v}_m\rangle| = |\lambda_m|^t.$$

For the right inequality in (3.1), if $\mathbf{v}_k$ is the eigenvector of A with eigenvalue λ_k and every entry having norm $\frac{1}{n}$ then $|\langle \mathbf{x}_0, \mathbf{v}_k\rangle| = \frac{1}{n}$ so that $d_A^2(t) = |A^t \mathbf{x}_0 - \mathbf{v}_0|_1^2 \leqslant n|A^t \mathbf{x}_0 - \mathbf{v}_0|_2^2 = \sum_{k \neq 0} \lambda_k^{2t}$. ∎

4. LATTICES AND THE CYCLIC CASE

With Lemma 3 in hand, we first prove Theorem 1 in the case where every group is cyclic. In this case the Markov matrices are circulants. Focusing on one walk from the family, we will denote by $\{\pm a_i\}$ $(1 \leqslant i \leqslant r)$ the r possible steps of the walk on $\mathbb{Z}/n\mathbb{Z}$ and by A the corresponding symmetric Markov transition matrix. The Fourier transform yields an explicit form of the eigenvalues of A for $k = 1, \ldots, n$:

$$\lambda_k = \sum_{i=1}^{r} \frac{2}{2r+1} \cos\left(2\pi \frac{k a_i}{n}\right) + \frac{1}{2r+1}.$$

Our strategy is to bound the λ_k and use these bounds in conjunction with Lemma 3 to get upper and lower bounds on $d_A^2(t)$ which will be tight enough to show that sharp cutoff does not occur. Specifically, we will associate to each λ_k an r-dimensional vector σ_k for which $e^{-c_1|\sigma_k|^2} \leqslant \lambda_k \leqslant e^{-c_2|\sigma_k|^2}$. Modulo a minor complication, the σ_k will all lie in an r-dimensional lattice, with σ_m being a minimal length vector in that lattice. This will allow us to upper bound the right hand side of (3.1) by the sum over the whole lattice of $e^{-c_2|\sigma_k|^2}$. Despite the inclusion of many extra terms, this bound will be tight enough for our needs.

The minor complication is that if $\lambda_k < 0$ we need to analyze a slightly different lattice which we do via a slightly different vector $\tilde{\sigma}_k$. However, the core of the argument remains the same.

We begin by defining for each $k = 1, \ldots, n-1$ the vector

$$\sigma_k = \left(\left\langle \frac{k a_1}{n} \right\rangle, \ldots, \left\langle \frac{k a_r}{n} \right\rangle\right) \in \mathbb{R}^r$$

where we use $\langle x \rangle \in \left(-\frac{1}{2}, \frac{1}{2}\right]$ for the smallest translate of x by an integer. The vector σ_k lives in the lattice

$$\Lambda = \mathbb{Z}^r + \mathbb{Z} \cdot \left(\frac{a_1}{n}, \ldots, \frac{a_r}{n}\right).$$

Note that if all the coordinates of σ_k are small, then all the cosines in the expression for λ_k are close to 1, so λ_k is close to 1. For λ_k to be close to -1, we are forced to have the coordinates of σ_k close to $\pm\frac{1}{2}$. For this reason, we also consider the vector

$$\tilde{\sigma}_k = \left(\left\langle \frac{k a_1}{n} + \frac{1}{2} \right\rangle, \ldots, \left\langle \frac{k a_r}{n} + \frac{1}{2} \right\rangle\right) \in \mathbb{R}^r,$$

which lives in the lattice

$$\tilde{\Lambda} = \mathbb{Z}^r + \mathbb{Z} \cdot \left(\frac{a_1}{n}, \ldots, \frac{a_r}{n}\right) + \mathbb{Z} \cdot \left(\frac{1}{2}, \ldots, \frac{1}{2}\right).$$

Let μ and $\tilde{\mu}$ denote the lengths of the shortest nonzero vectors of Λ and $\tilde{\Lambda}$ respectively. Since $\sigma \in \tilde{\Lambda}$ implies $2\sigma \in \Lambda$, we have

$$2\tilde{\mu} \geqslant \mu.$$

The following lemma gives a lower bound on $d(t)$ by bounding $|\lambda_m|$ in terms of the length μ of the shortest vector in Λ. Here we assume n is sufficiently large, which we may do since if any subfamily of walks has bounded size then the family cannot have sharp cutoff.

LEMMA 4. *Let $\beta = \frac{8\pi^2}{2r+1}$. Then if n is sufficiently large,*

$$|\lambda_m| \geqslant e^{-\beta\mu^2}.$$

Proof. First note that $\mathrm{Vol}[(\mathbb{R}^r/\Lambda)] = \frac{1}{n}$, which approaches 0. It follows that for large enough n, the shortest nonzero vector in Λ has length less than 1, and is therefore equal to σ_k for some $k = 1, \ldots, n-1$. For this shortest σ_k, we may also assume that the all the coordinates of σ_k have absolute value less than $\frac{1}{2\pi}$. We will now show

$$|\lambda_k| \geqslant e^{-\beta|\sigma_k|^2},$$

which will establish the lemma.

For any $x \in [-1, 1]$, we have $\cos(x) \geqslant e^{-x^2}$. Thus

$$\lambda_k \geqslant \frac{1}{2r+1}\Big(\sum_i 2e^{-(2\pi\langle\frac{ka_i}{n}\rangle)^2} + 1\Big).$$

Now consider the right hand side as an average of $2r+1$ values of the function e^{-x}. The convexity of this function yields

$$\lambda_k \geqslant e^{-\frac{2}{2r+1}\sum_i (2\pi\langle\frac{ka_i}{n}\rangle)^2} = e^{-\beta|\sigma_k|^2},$$

as desired. ∎

Next we establish an upper bound on $d(t)$ by bounding the sum of λ_k^{2t}. For each k, we bound $|\lambda_k|$ by an exponential of either $|\sigma_k|$ or $|\tilde{\sigma}_k|$. We then replace the sum of these exponentials with a sum over the entire lattice $\tilde{\Lambda}$.

LEMMA 5. *Let $\alpha = \frac{8}{(2r+1)^2}$. Then*

$$\sum_{k\neq 0} \lambda_k^{2t} \leqslant 2 \sum_{\sigma\in\tilde{\Lambda}\setminus\{0\}} e^{-\alpha|\sigma|^2\cdot 2t}. \tag{4.1}$$

We note that the above lemma gives a relation between the mixing time $t(\epsilon)$ and the smoothing parameter η of the dual lattice $\tilde{\Lambda}^*$ introduced by Micciancio

and Regev [13]. Specifically, using the definition of the smoothing parameter that appears in [13, Section 3], a computation shows that Lemma 5 implies

$$t(\epsilon) < \frac{\pi(2r+1)^2}{16}[\eta_{\epsilon^2/2}(\tilde{\Lambda}^*)]^2.$$

Proof of Lemma 5. Fix $k \in \{1, \ldots, n-1\}$. First assume $\lambda_k \geqslant 0$. As in the previous proof, we will replace cosines with exponentials, this time using the inequality $\cos(x) \leqslant e^{-\frac{1}{2\pi^2}x^2}$, which is valid for all $x \in \left[-\frac{3}{2}\pi, \frac{3}{2}\pi\right]$. This gives

$$\lambda_k \leqslant \frac{1}{2r+1}\Big(\sum_i 2e^{-2\langle \frac{ka_i}{n}\rangle^2} + 1\Big).$$

We now consider the right hand side as an average value of the function e^{-x^2} at $2r+1$ values of x, each of which lies in the interval $\left[0, \frac{1}{\sqrt{2}}\right]$. Since e^{-x^2} is concave down on this interval, we obtain

$$\lambda_k \leqslant e^{-\frac{8}{(2r+1)^2}[\sum |\langle \frac{ka_i}{n}\rangle|]^2} = e^{-\alpha|\sigma_k|_1^2}.$$

Since $|v|_1 \geqslant |v|$, we finally arrive at

$$\lambda_k \leqslant e^{-\alpha|\sigma_k|^2}.$$

Now consider the case in which $\lambda_k < 0$. In this case, $|\lambda_k| = -\lambda_k$ is the sum of negative cosines, so using $-\cos(2\pi x) = \cos\big(2\pi\big(x + \frac{1}{2}\big)\big)$, an argument similar to the first case shows that

$$|\lambda_k| \leqslant e^{-\alpha|\tilde{\sigma}_k|^2}.$$

Hence in both cases we see that $|\lambda_k|$ is bounded above by $e^{-\alpha|\sigma|^2}$ with $\sigma \in \tilde{\Lambda}$. To complete the proof of the lemma, we must show that no such σ appears for more than two different values of k. To see this, note if $\sigma_k = \sigma_{k'}$ for some $k, k' \in \{1, \ldots, n-1\}$, then $ka_i \equiv k'a_i \bmod n$ for all i. Since the a_i generate $\mathbb{Z}_n$, this forces $k = k'$. Similarly, $\tilde{\sigma}_k = \tilde{\sigma}_{k'}$ implies $k = k'$. Finally, if $\sigma_k = \tilde{\sigma}_{k'}$, then we get either $k = k'$, or $k = k' \pm \frac{n}{2}$ with n even. In any event, no more than two distinct choices of k lead to the same lattice element σ. ∎

The next two lemmas will further bound the right hand side of (4.1) in terms of r and Λ. This amounts to showing that the number of lattice points of length at most d is bounded above by a polynomial in d. Given two discrete (infinite) multisets $S, T \subset \mathbb{R}_{\geqslant 0}$, we say *$S$ dominates T* if there are orderings of $S = \{s_i : i \in \mathbb{Z}^+\}$ and $T = \{t_i : i \in \mathbb{Z}^+\}$ such that $s_i \geqslant t_i$ for all i. Note that dominance induces a partial order on such multisets.

For a lattice Λ, let $|\Lambda|$ denote the multiset of norms of vectors of Λ. So $|\Lambda|$ is discrete and contained in $\mathbb{R}_{\geqslant 0}$.

LEMMA 6. *Let r be a fixed integer. Let T be the multiset consisting of $\{0\}$ and $3^r(i+1)^r$ copies of $i \in \mathbb{N} = \{1, 2, \dots\}$. Then for every full rank lattice Λ in $\mathbb{R}^r$ with minimal nonzero vector of length 1, the set $|\Lambda|$ dominates the multiset T.*

Proof. Observe that for $d \geqslant 1$ the number $n(d)$ of lattice points of Λ of norm no greater than d is bounded above by $(3d)^r$. This is because the shortest vector of Λ has length 1, so radius $1/2$ balls around points of Λ are disjoint, so that $n(d) \cdot B(1/2) \leqslant B(d+1/2)$ where $B(t)$ is the volume of a ball in $\mathbb{R}^r$ of radius t. This gives $n(d) \leqslant B(d+1/2)/B(1/2) = 2^r(d+1/2)^r \leqslant 3^r d^r$.

It follows that no more than $3^r(i+1)^r$ lattice points are between lengths i and $i+1$ and thus $|\Lambda|$ dominates the multiset T. ∎

LEMMA 7. *Fix r and Λ a full rank lattice in $\mathbb{R}^r$, and let $\mu = \min_{\mathbf{0} \neq \sigma \in \Lambda} |\sigma|_2$. Then*

$$\sum_{\sigma \in \Lambda} e^{-a|\sigma|_2^2} \leqslant 1 + 3^r r! \left(\frac{1}{(1-(e^{-a\mu^2}))^r} - 1 \right).$$

Proof. By Lemma 6 we have

$$\sum_{\sigma \in \Lambda} e^{-a|\sigma|_2^2} \leqslant 1 + \sum_{t \in T} e^{-a(\mu t)^2} \leqslant 1 + \sum_{i \in \mathbb{N}} 3^r (i+1)^r (e^{-a\mu^2})^{i^2}.$$

This last sum can be bounded as

$$\sum_{i \in \mathbb{N}} 3^r (i+1)^r (e^{-a\mu^2})^{i^2} \leqslant \sum_{j \in \mathbb{N}} 3^r (\sqrt{j}+1)^r (e^{-a\mu^2})^{j} \leqslant 3^r r! \sum_{j \in \mathbb{N}} \binom{j+r}{r} (e^{-a\mu^2})^{j}$$

where we have substituted $i^2 = j$ for the first inequality and $(\sqrt{j}+1)^r \leqslant r! \binom{j+r}{r}$ for the second. Finally, using the well known $\sum_{i=1}^{\infty} \binom{i+r}{i} x^i = \frac{1}{(1-x)^r} - 1$, we have

$$3^r r! \sum_{j \in \mathbb{N}} \binom{j+r}{r} (e^{-a\mu^2})^{j} = 3^r r! \left(\frac{1}{(1-(e^{-a\mu^2}))^r} - 1 \right). \qquad \blacksquare$$

We now prove Theorem 1 (in the cyclic case). Combining Lemma 5 with Lemma 7 yields the upper bound

$$\sum_i \lambda_i^{2t} \leqslant 2 \cdot 3^r r! \left(\frac{1}{(1-e^{-\alpha \tilde{\mu}^2 \cdot 2t})^r} - 1 \right).$$

Meanwhile, Lemma 4 together with $2\tilde{\mu} \geqslant \mu$ gives

$$e^{-\alpha \tilde{\mu}^2 \cdot 2t} \leqslant |\lambda_m|^{\frac{\alpha}{2\beta} t},$$

hence

$$\sum_i \lambda_i^{2t} \leqslant 2 \cdot 3^r r! \left(\frac{1}{(1-|\lambda_m|^{\frac{\alpha}{2\beta} t})^r} - 1 \right). \tag{4.2}$$

By the first inequality of Lemma 3 we have $t_A(\epsilon) \geqslant t_0$ where t_0 is the solution to $|\lambda_m^t| = \epsilon$. It follows that

$$t_A(\epsilon) \geqslant \log_{|\lambda_m|} \epsilon.$$

By the second inequality of Lemma 3 and (4.2), we have $t_A(1-\epsilon) \leqslant t_1$ with t_1 the solution to $2 \cdot 3^r r! \big(\frac{1}{(1-|\lambda_m|^{\frac{\alpha}{2\beta}t})^r} - 1 \big) = (1-\epsilon)^2$. Solving this yields

$$t_A(1-\epsilon) \leqslant \frac{2\beta}{\alpha} \log_{|\lambda_m|} C(\epsilon) \tag{4.3}$$

with $C(\epsilon) = 1 - \big(\frac{(1-\epsilon)^2}{2 \cdot 3^r r!} + 1 \big)^{-1/r}$.

Together we have, for each walk in the family,

$$\frac{t_A(\epsilon)}{t_A(1-\epsilon)} \geqslant \frac{\alpha}{2\beta} \log_{|\lambda_m|} \epsilon / \log_{|\lambda_m|} C(\epsilon) = \frac{1}{2\pi^2(2r+1)} \log_{C(\epsilon)} \epsilon.$$

Notice that as $\epsilon \to 0$ we have $C(\epsilon) \to 1 - \big(\frac{1}{2 \cdot 3^r r!} + 1 \big)^{-1/r}$, which is less than 1 and bounded away from zero, so there will be a choice of ϵ that will make the right hand side greater than 1 independent of A. This completes the proof in the cyclic case.

5. THE GENERAL CASE

To prove the main theorem for arbitrary abelian groups, we closely follow the proof for cyclic groups given above.

We consider an arbitrary finite abelian group G, which we express as the product of s cyclic groups of orders $n_1, \ldots, n_s$. Let $n = |G| = n_1 \cdots n_s$. We suppose that we have r generators $a_1, \ldots, a_r$ of G, with

$$a_j = (a_{j1}, \ldots, a_{js}),$$

with each $a_{jh} \in \mathbb{Z}_{n_h}$, $h = 1, \ldots, s$. We will assume that we have chosen a product decomposition in which $s \leqslant r$, which is always possible for an abelian group generated by r elements.

It will be convenient to have

$$\eta_{jh} = \frac{a_{jh}}{n_h} \in [0, 1),$$

and define $\eta_j = (\eta_{j1}, \ldots, \eta_{js})$.

The eigenvalues of our Markov process are indexed by tuples $k = (k_1, \ldots, k_s)$ with $k_h \in \mathbb{Z}_{n_h}$ and are given by

$$\lambda_k = \frac{1}{2r+1} \Big(2 \sum_{j=1}^{r} \cos(2\pi k \cdot \eta_j) + 1 \Big).$$

As before, we identify the r-tuple appearing in the arguments of the cosines above, letting

$$\sigma_k = (\langle k \cdot \eta_1 \rangle, \ldots, \langle k \cdot \eta_r \rangle) \in \mathbb{R}^r.$$

The vector σ_k lives in the lattice

$$\Lambda = \mathbb{Z}^r + \sum_{h=1}^{s} \mathbb{Z}\theta_h, \quad \text{where} \quad \theta_h = (\eta_{1h}, \ldots, \eta_{rh}).$$

(The θ matrix is the transpose of the η matrix.)

As before, we also introduce

$$\tilde{\sigma}_k = \left(\left\langle k \cdot \eta_1 + \tfrac{1}{2} \right\rangle, \ldots, \left\langle k \cdot \eta_r + \tfrac{1}{2} \right\rangle\right) \in \mathbb{R}^r,$$

which lives in the lattice

$$\tilde{\Lambda} = \mathbb{Z}^r + \sum_{h=1}^{s} \mathbb{Z}\theta_h + \mathbb{Z} \cdot \left(\tfrac{1}{2}, \ldots, \tfrac{1}{2}\right).$$

As before, we let λ_m be the eigenvalue with largest absolute value, $m \neq 0$, and we have

$$\lambda_m^{2t} \leqslant d^2(t) \leqslant \sum_{k \neq 0} \lambda_k^{2t}.$$

With this setup, the rest of the argument is similar to the cyclic case. The conclusion of Lemma 4 holds exactly as before, while Lemma 5 must be modified as follows:

LEMMA 8. *Let* $\alpha = \frac{8}{(2r+1)^2}$. *Then*

$$\sum_{k \neq 0} \lambda_k^{2t} \leqslant 2^s \sum_{\sigma \in \tilde{\Lambda} \setminus \{0\}} e^{-\alpha |\sigma|^2 \cdot 2t}.$$

Proof. As in the proof of Lemma 5, we see that if $\lambda_k > 0$, then

$$\lambda_k \leqslant e^{-\alpha |\sigma_k|^2},$$

and if $\lambda_k < 0$, then

$$|\lambda_k| \leqslant e^{-\alpha |\tilde{\sigma}_k|^2}.$$

So in either case $|\lambda_k|$ is bounded above by $e^{-\alpha|\sigma|^2}$ with $\sigma \in \tilde{\Lambda}$. To complete the proof of the lemma, we must show that no such σ appears for more than 2^s different values of k. For this purpose, assume that $\sigma_k = \sigma_{k'}$. This implies that for $j = 1, \ldots, r$,

$$\langle k \cdot \eta_j \rangle = \langle k' \cdot \eta_j \rangle.$$

Equivalently, if we consider the $r \times s$ matrix $A = (a_{jh})$, and let x be the length s column vector with $x_h = \frac{k_h - k'_h}{n_h}$, then

$$Ax \in \mathbb{Z}^r.$$

Now we use the fact the a_j generate G. This implies that there exists an integer vector $c = (c_1, \ldots, c_r)$ such that $\sum c_j a_j = (1, 0, \ldots, 0)$ in G, i.e.,

$$cA = (1 + \gamma_1 n_1, \gamma_2 n_2, \ldots, \gamma_s n_s),$$

where the γ_h are integers. It follows that

$$cAx = x_1 + \sum_{h=1}^{s} \gamma_h (k_h - k'_h),$$

and hence $x_1 \in \mathbb{Z}$. Similarly, all x_h are in $\mathbb{Z}$, and we get $k_h \equiv k'_h \bmod n_h$, i.e. $k = k'$.

We have now established that $\sigma_k = \sigma_{k'}$ implies $k = k'$. A similar argument shows that $\tilde{\sigma}_k = \tilde{\sigma}_{k'}$ implies $k = k'$. Finally, if $\sigma_k = \tilde{\sigma}_{k'}$, then for each h, we get either $k_h = k'_h$ or $k_h = k'_h \pm \frac{n_h}{2}$ with n_h even. It follows that no more than 2^s distinct choices of k lead to the same lattice element σ. ∎

Finally, because r is constant and $s \leqslant r$, the presence of the factor 2^s in the above lemma does not effect the rest of the proof given in the cyclic case, which goes through without further modification.

Acknowledgments. The authors thank Persi Diaconis for suggesting this problem and several anonymous reviewers for helping us to frame this contribution in a larger context. We also thank Julie Landau for her court coverage.

REFERENCES

[1] D. Aldous and P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly 93 (1986), 333–348.

[2] D. Aldous and P. Diaconis, *Strong uniform times and finite random walks*, Adv. Appl. Math. 8 (1987), 69–97.

[3] G.-Y. Chen and L. Saloff-Coste, *The cutoff phenomenon for ergodic Markov processes*, Electron. J. Probab. 13 (2008), 26–78.

[4] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. USA 93 (1996), 1659–1664.

[5] P. Diaconis, *Group Representations in Probability and Statistics*, Inst. Math. Statist., Hayward, CA, 1988.

[6] P. Diaconis and L. Saloff-Coste, *Comparison techniques for random walks on finite groups*, Ann. Probab. 21 (1993), 2131–2156.

[7] P. Diaconis and L. Saloff-Coste, *Moderate growth and random walk on finite groups*, Geom. Funct. Anal. 4 (1994), 1–36.

[8] J. Hermon and S. Olesker-Taylor, *Cutoff for almost all random walks on abelian groups*, arXiv:2102.02809v1 (2021).
[9] J. Hermon and S. Olesker-Taylor, *Geometry of random Cayley graphs of abelian groups*, arXiv:2102.02801 (2021).
[10] J. Hermon and S. Olesker-Taylor, *Cutoff for almost all random walks on upper triangular matrices*, arXiv:1911.02974v2 (2019).
[11] R. Hough, *Mixing and cut-off in cycle walks*, Electron. J. Probab. 22 (2017), art. 90, 49 pp.
[12] E. Lubetzky and A. Sly, *Cutoff phenomena for random walks on random regular graphs*, Duke Math. J. 153 (2010), 475–510.
[13] D. Micciancio and O. Regev, *Worst-case to average-case reductions based on Gaussian measures*, SIAM J. Comput. 37 (2007), 267–302.
[14] J. Salez, *Cutoff for non-negatively curved Markov chains*, arXiv:2102.02809v1 (2021).
[15] L. Saloff-Coste, *Random walks on finite groups*, in: Probability on Discrete Structures, Encyclopaedia Math. Sci. 110, Springer, Berlin, 2004, 263–346.

Aaron Abrams
Mathematics Department
Washington and Lee University
Lexington, VA, 24450, USA
E-mail: abramsa@wlu.edu

Eric Babson
Department of Mathematics
University of California, Davis
Davis, CA 95616, USA
E-mail: babson@math.ucdavis.edu

Henry Landau
AT&T Research
Murray Hill, NJ 07974, USA
E-mail: henry.j.landau@gmail.com

Zeph Landau
Department of Computer Science
University of California, Berkeley
Berkeley, CA 94706, USA
E-mail: zeph.landau@gmail.com

Jamie Pommersheim
Mathematics Department
Reed College
Portland, OR 97202, USA
E-mail: jamie@reed.edu

Received 19.7.2021;
accepted 22.1.2022