The Annals of Probability
1995, Vol. 23, No. 4, 1961–1981

EIGENVALUES OF RANDOM WALKS ON GROUPS

BY RICHARD STONG

Rice University

In this paper we discuss and apply a novel method for bounding the eigenvalues of a random walk on a group G (or equivalently on its Cayley graph). This method works by looking at the action of an Abelian normal subgroup H of G on G. We may then choose eigenvectors which fall into representations of H. One is then left with a large number (one for each representation of H) of easier problems to analyze. This analysis is carried out by new geometric methods. This method allows us to give bounds on the second largest eigenvalue of random walks on nilpotent groups with low class number. The method also lets us treat certain very easy solvable groups and to give better bounds for certain nice nilpotent groups with large class number. For example, we will give sharp bounds for two natural random walks on groups of upper triangular matrices.

1. Introduction. Random processes have been an active field of mathematics over the last 10 years. A large number of interesting random processes can be described in the following manner. Let G be a finite group and let S be a symmetric generating set for G. This generating set induces a random walk on G. That is, start at the identity element and at each step multiply on the left by a (uniformly) randomly chosen element of S. Alternately let Γ be the Cayley graph of G with generating set S; that is, the vertex set of Γ is G and there is an edge joining g to g' if and only if $g'g^{-1} \in S$. Then the random walk is the usual random walk on the regular graph Γ. Start at the identity and at each step follow a randomly chosen edge out of the current vertex. Let P be the transition probability matrix for this random walk. We will find it convenient to regard P as the linear map $P\colon L^2(G) \to L^2(G)$ defined by

$$(P\psi)(g) = \frac{1}{|S|} \sum_{x \in S} \psi(xg).$$

Let $1 = \beta_0 > \beta_1 \geq \beta_2 \geq \cdots \geq \beta_{N-1} \geq -1$ be the eigenvalues of P, where $N = |G|$.

We are interested in giving an upper bound on β_1 and a lower bound on β_{N-1}. Such bounds give a bound on the rate of convergence of P^k to the uniform distribution, (e.g., [4], Proposition 3). A number of methods have been used to give such bounds. If S is a union of conjugacy classes, one can get the eigenvalues and eigenvectors from the representation theory of G [1]. More recently, techniques have been developed to give bounds on the eigen-

Received April 1994; revised January 1995.
AMS 1991 *subject classification*. 60J10.
Key words and phrases. Eigenvalues, random walks, nilpotent groups.

values using geometric properties of a collection of paths in the graph. Jerrum and Sinclair [6] give a path argument for bounding the Cheeger constant of a graph. This in turn gives bounds on β_1. Diaconis and Stroock ([4], Proposition 1) give a path argument for bounding β_1 directly and a cycle argument ([4], Proposition 2) for supplying the lower bound on β_{N-1}. We will present here a novel method which uses the group structure of G to reduce the problem to a number of easier bounds which are then determined by geometric methods. This method gives nice bounds on the second largest eigenvalue β_1. Typically they are of the form $\beta_1 \leq 1 - C/(e^2 n)$, where n is the number of generators and e is some measure of the exponent of elements of G. We also get weak bounds on the smallest eigenvalue β_{N-1} in these cases. These bounds will be of the same order as the bounds on the second largest, which is ample for most applications. Usually stronger bounds on β_{N-1} will be available by other means, for example, [4], Proposition 2, or its variant [9], Proposition 2.

We will be able to give nice bounds for a number of classes of groups. In particular, we will derive results for random walks on nilpotent groups with low class number. This result complements work of Diaconis and Saloff-Coste [2, 3]. These two papers give bounds on the eigenvalues and rates of convergence of random walks on groups satisfying certain growth conditions, including nilpotent groups. In particular, suppose G is a nilpotent group with a symmetric set of n generators (containing the identity), class number l and diameter γ. Let P be the transition probability matrix for the associated random walk on G, and let U be the uniform distribution on G. Diaconis and Saloff-Coste show ([3], Corollary 5.3) that there are constants $B = B(l, n)$ and $C = C(l, n)$ such that

$$\|P^k - U\|_{\mathrm{Var}} \leq Be^{-c} \quad \text{if } k = (1 + c)\gamma^2 n \text{ and } c > 0,$$

and

$$\|P^k - U\|_{\mathrm{Var}} \geq \tfrac{1}{2}e^{-c} \quad \text{if } k = c\gamma^2/C.$$

These results show that for many families of nilpotent groups (ones with n and l fixed) order γ^2 steps are necessary and sufficient to achieve randomness. However, for families of groups with n or l increasing, the bounds on the constants B and C in [3] are insufficient to give good bounds on the convergence. The results of this paper will include reasonable bounds on the second largest eigenvalue for families of nilpotent groups with l small and fixed and n increasing.

2. The group theoretic argument. For this method, we need the following observation. Let $H \subset G$ be an Abelian normal subgroup of G. We chose the random walk to be given by left multiplication by generators; therefore, H acts on the Cayley graph Γ by right multiplication. Hence the eigenfunctions of P may be chosen to lie in irreducible representations of H, which are of course one-dimensional since H is Abelian. (This method might still work in special cases even if H is non-Abelian. One would have to know the representation theory of H in detail and make other adjustments.) Pick one such representation $\rho\colon H \to \mathbf{C}$ and let ψ be an eigenfunction of P which lies in that representation; that is, $\psi(gh) = \psi(g)\rho(h)$ for all $g \in G$ and $h \in H$. If we fix coset representatives for G/H, then ψ is determined by its values on the coset representatives. That is, we can view ψ as a function on G/H.

Explicitly, let $[g]$ denote the coset containing g and let $t_{[g]}$ be the coset representative. Define $\bar{\psi}\colon G/H \to \mathbf{C}$ by $\bar{\psi}([g]) = \psi(t_{[g]})$. We can recover ψ from $\bar{\psi}$ by the identity $\psi(g) = \bar{\psi}([g])\rho(t_{[g]}^{-1}g)$. In terms of $\bar{\psi}$, the equation $P\psi = \lambda\psi$ becomes

$$\begin{aligned}\lambda\bar{\psi}([g]) &= \lambda\psi(t_{[g]}) = \frac{1}{|S|}\sum_{x\in S}\psi(xt_{[g]})\\ &= \frac{1}{|S|}\sum_{x\in S}\psi(t_{[xg]})\rho\left(t_{[xg]}^{-1}xt_{[g]}\right) = \frac{1}{|S|}\sum_{x\in S}\bar{\psi}([xg])\rho\left(t_{[xg]}^{-1}xt_{[g]}\right).\end{aligned} \tag{1}$$

That is, $\bar{\psi}$ is an eigenfunction of a "random walk on G/H with phases on the edges." Let $\bar{P}$ denote the corresponding matrix. Note that the phases are all values of ρ. If ρ is the trivial representation of H, then all the phases are 1. Thus in this case $\bar{\psi}$ is an eigenfunction of the usual random walk on G/H as encountered above. In particular, these include the unique eigenfunction with eigenvalue 1. In other words, if we wish to bound β_1 and β_{N-1} it suffices to bound the second largest and the smallest eigenvalues of the induced random walk on G/H and to bound the eigenvalues of $\bar{P}$ away from ± 1 for all nontrivial representations ρ of H. We have reduced the original problem to a large number of smaller problems. This result is summarized by the following theorem.

THEOREM 1. *Let G be a finite group and let S be a symmetric generating set for G. Let P be the transition probability matrix for the associated random walk on G and let $1 = \beta_0 > \beta_1 \geq \beta_2 \geq \cdots \geq \beta_{N-1} \geq -1$ be the eigenvalues of P, where $N = |G|$. Suppose $H \subset G$ is an Abelian normal subgroup and $t_{[g]} \in G$ are coset representatives for H. Then the eigenvalues of P are exactly the eigenvalues of the induced random walk on G/H and the eigenvalues of the matrices $\bar{P}$ on $L^2(G/H)$ given by*

$$\overline{P\psi}([g]) = \frac{1}{|S|}\sum_{x\in S}\bar{\psi}([xg])\rho\left(t_{[xg]}^{-1}xt_{[g]}\right) \tag{2}$$

for ρ some nontrivial character of H. In particular, β_1 and β_{N-1} must arise in one of these two ways.

3. The geometric argument. Theorem 1 constitutes the group-theoretic part of this paper. Techniques for bounding the second largest eigenvalue of the induced random walk on G/H are given in [1], [6], [4] and this paper, and are briefly described in the Introduction. The bounds on the eigenvalues of $\bar{P}$ are somewhat simpler because we are bounding the largest eigenvalue and not the second largest. To determine these bounds for the nontrivial representations, it is convenient to have some terminology for graphs with phases on the edges.

Following [9], consider the following data: a graph $\Gamma = (V, E)$, a positive real number d, a function w from the directed edges to $\mathbf{C}$ with $w([y, x]) = \overline{w([x, y])}$ and a function s from the vertices to $[0, d]$. Call such a collection of data $\mathbf{\Gamma} = (\Gamma, d, w, s)$ a twisted graph. We will interpret this data as follows. The graph Γ is the underlying graph for our "random walk with phases on the edges." For now suppose Γ is simple graph. If x and y are vertices of Γ, let $[x, y]$ denote the edge from x to y. This simplifies the notation and, as discussed below, we can always reduce to this case. We will regard d as roughly the degree of the graph. The function w we will regard as giving the weight on the edge. Our weights will generally be in $\{z \in \mathbf{C}\colon |z| = 1\}$ and in this case we will also refer to them as the phases on the edge. We regard $s(x)$ as giving the weight for remaining stationary at x and $s(x)/d$ as the probability of remaining stationary. In our examples, we will always have $s(x)$ in the range $[0, d - \Sigma_y |w([x, y])|]$. We may regard $a(x) = d - \Sigma_y |w([x, y])| - s(x)$ as the weight for a particle at x to disappear, or be absorbed, and $a(x)/d$ is the probability a particle at x will be absorbed.

To any twisted graph we associate a Hermitian matrix A by taking the x, y entry of A to be $w([x, y])/d$ if Γ contains an edge from x to y and 0 otherwise, and by taking the x, x entry of A to be $s(x)/d$. If Γ is a d-regular graph, $w = 1$ and $s = 0$, then A is just the transition probability matrix for the usual random walk on Γ. We will call the eigenvalues of A the eigenvalues of the twisted graph as well. We will say a Hermitian matrix A' dominates another A if $v^\dagger A v \le v^\dagger A' v$ for all v. Similarly we will say one twisted graph dominates another if their vertex sets agree and the corresponding Hermitian matrix of the first dominates that of the second.

We will use these definitions to give eigenvalue bounds as follows. Call a twisted graph diagonal if the underlying graph has no edges (or each edge has weight 0). Then we have the following obvious lemma.

LEMMA 2. *If the twisted graph* $\mathbf{\Gamma}$ *is dominated by the diagonal twisted graph* $\mathbf{\Gamma}' = (\Gamma', d', w', s')$ *and* λ *is any eigenvalue of* $\mathbf{\Gamma}$, *then* $\lambda \le \max_{v \in \Gamma'} s'(v)/d'$.

If Γ is not a simple graph, the definitions above need to be adapted slightly and the notation becomes more cumbersome. For example, in the definition of $a(x)$ the sum over neighbors of x must be replaced by a sum over edges out of x. In computing the associated Hermitian matrix A, the x, y entry of A would be the sum over all edges e from x to y of the contribution $w(e)/d$. Similarly the x, x entry of A would be the sum of the contribution of $s(x)$ and the contributions of all loops at x. Note that we may always reduce to the case where the underlying graph is simple. Explicitly, suppose (Γ, d, w, s) is a twisted graph with Γ not simple. Let Γ' be the simple graph one gets by removing any loops of Γ and collapsing any multiple edges of Γ to a single edge. Define new weights on edges and vertices of Γ' by

$$w'([x, y]) = \sum_{\substack{e \text{ an edge} \\ \text{from } x \text{ to } y}} w(e)$$

and

$$s'(x) = s(x) + \sum_{\substack{l \text{ a loop} \\ \text{at } x}} w(l).$$

Then (Γ', d, w', s') is a twisted graph with the same associated matrix A (and the same absorption probabilities) as (Γ, d, w, s). For our purposes, a twisted graph is really just a geometric way of encoding the matrix A. Therefore, we may replace (Γ, d, w, s) by (Γ', d, w', s') without loss. Hence, below we will always assume our underlying graph is simple and use the resulting easier notation.

Suppose now that we have data as in Section 2. That is, G is a finite group with symmetric generating set S, H is an Abelian normal subgroup of G, $t_{[g]}$ are coset representatives for G/H and ρ is a nontrivial representation of H. Then there is an obvious twisted graph (Γ, d, w, s) associated with this data. Let Γ be the graph whose vertex set is G/H. For every $[g] \in G/H$ and every element $x \in S$, let Γ have an edge from $[g]$ to $[xg]$, and assign the weight $\rho(t_{[xg]}^{-1} x t_{[g]})$ to this edge. Let $d = |S|$ and $s = 0$. Note that Γ will have loops if $H \cap S \neq \varnothing$ and multiple edges if two elements of S are in the same coset of G/H. If this occurs, reduce to a simple graph as described above. The important fact to notice is that the Hermitian matrix A associated with the twisted graph (Γ, d, w, s) is the matrix $\overline{P}$ associated with the data G, H, S, $t_{[g]}$ and ρ by (2). To apply Theorem 1 and Lemma 2, we wish to bound the eigenvalues of the associated matrix $A = \overline{P}$ by finding a diagonal twisted graph which dominates this one. Hence we need to understand how to build dominating twisted graphs.

Note that we can build twisted graphs which dominate others by the following basic construction. If $\mathbf{X}$ is a sub(twisted graph) of Γ, $\mathbf{X}'$ dominates $\mathbf{X}$ and $\mathbf{\Gamma}'$ is the twisted graph we get by replacing $\mathbf{X}$ in $\mathbf{\Gamma}$ by $\mathbf{X}'$, then $\mathbf{\Gamma}'$ dominates $\mathbf{\Gamma}$. This basic construction allows us to build up dominating twisted graphs step by step using a few basic examples. It is not yet clear what all the basic examples should be, but the ones below are certainly among them.

The easiest example of this basic construction is the Cauchy–Schwarz inequality applied to an edge. View an edge as the twisted graph

$$\bullet \overset{1}{-\!-} \bullet \qquad (s = 0, d = 1).$$

This twisted graph corresponds to the 2×2 Hermitian matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ or equivalently the quadratic form $v^\dagger A v = 2\,\mathrm{Re}(v_1 \bar{v}_2)$. By Cauchy–Schwarz or a number of other standard inequalities, one has $2\,\mathrm{Re}(v_1 \bar{v}_2) \le (1 - \varepsilon)|v_1|^2 + |v_2|^2/(1 - \varepsilon)$ for any $0 \le \varepsilon < 1$. The right-hand side of this inequality is the quadratic form associated with the twisted graph

$$\overset{s=1-\varepsilon}{\bullet} \quad \overset{s=1/(1-\varepsilon)}{\bullet} \qquad (d=1).$$

Therefore, we see that $\bullet \overset{1}{--} \bullet$ $(s = 0,\ d = 1)$ is dominated by the twisted graph

$$\overset{s=1-\varepsilon}{\bullet} \quad \overset{s=1/(1-\varepsilon)}{\bullet} \qquad (d=1)$$

for any $0 \le \varepsilon < 1$. Also note that, for any phase ω, the quadratic form $2\,\mathrm{Re}(\omega v_1 \bar{v}_2)$ associated with an edge with weight ω satisfies the same upper bound. Therefore, we may replace an edge (with weight any phase) by stationary probabilities of $1-\varepsilon$ and $1/(1-\varepsilon)$ at its endpoints to get a dominating twisted graph. Notice that removing an edge symmetrically, that is, with $\varepsilon = 0$, with Cauchy–Schwarz leaves all the absorption weights $a(x)$ unchanged. This observation will be used extensively below (in fact, this observation, if not this terminology, was used extensively in [8].)

In this paper we will also make fundamental use of the following basic example. Suppose the twisted graph $\mathbf{\Gamma} = (\Gamma, 2, w, 0)$ has as its underlying graph Γ a cycle of length L $(L \ge 3)$, $d = 2$, $s = 0$ and all weights of norm 1. Suppose v is any vertex of Γ and η is any phase. Consider the twisted graph $\mathbf{\Gamma}' = (\Gamma, 2, w', 0)$, where $w'([x,y]) = w([x,y])$ if neither x nor y is v, $w'([x,v]) = \eta w([x,v])$ and $w'([v,y]) = \eta^{-1} w([v,y])$. This is equivalent to conjugating the associated Hermitian matrix A by the diagonal matrix with one η on the diagonal and all other diagonal entries 1. In particular the eigenvalues of A and A' agree. Using this operation repeatedly, we see that the eigenvalues of A, or equivalently of Γ, are determined by L and the product of the phases as one goes around the cycle (forward). Let $e^{i\theta(\mathbf{\Gamma})}$ denote this phase, where $-\pi < \theta(\mathbf{\Gamma}) \le \pi$. Since the eigenvalues only depend on the product, we may assume each forward edge has weight $e^{i\theta(\mathbf{\Gamma})/L}$. Therefore, the resulting matrix A is cyclic and the eigenvalues are $\cos((\theta(\mathbf{\Gamma}) + 2\pi k)/L)$. The largest of these is $\cos(\theta(\mathbf{\Gamma})/L)$. Therefore, we have proven the following lemma.

LEMMA 3. *Let $\mathbf{\Gamma} = (\Gamma, 2, w, 0)$ be a twisted graph with underlying graph Γ a cycle of length L and all weights of unit norm. Let $e^{i\theta(\mathbf{\Gamma})}$, $-\pi < \theta(\mathbf{\Gamma}) \le \pi$, be the product of the phases as one goes around the cycle. Then $\mathbf{\Gamma}$ is dominated by the twisted graph $\mathbf{\Gamma}' = (\Gamma', 2, 0, 2\cos(\theta(\Gamma)/L))$ whose underlying graph is L vertices with no edges and where each vertex has stationary weight $s = 2\cos(\theta(\mathbf{\Gamma})/L)$.*

Cycles are particularly relevant to giving bounds for the twisted graphs arising from Theorem 1. To see how cycles get involved suppose $\mathbf{\Gamma} = (\Gamma, |S|, w, 0)$ is the twisted graph associated with the data $H, G, S, t_{[g]}$, and ρ is a nontrivial representation of H. Suppose $y_1 y_2 \cdots y_k$ is a word in the elements of S which represents a nonidentity element of H. Then for each coset $[g]$ of G/H this word determines a cycle in Γ, namely, the cycle $[g], [y_k g], [y_{k-1} y_k g], \ldots, [y_2 \cdots y_k g], [y_1 y_2 \cdots y_k g] = [g]$. The product of the phases w as one goes around this cycle is

$$\begin{aligned} &\rho\left(t_{[g]}^{-1} y_1 t_{[y_2 \cdots y_k g]}\right) \rho\left(t_{[y_2 \cdots y_k g]}^{-1} y_2 t_{[y_3 \cdots y_k g]}\right) \cdots \rho\left(t_{[y_k g]}^{-1} y_k t_{[g]}\right) \\ &\quad = \rho\left(t_{[g]}^{-1} y_1 y_2 \cdots y_k t_{[g]}\right) \\ &\quad = \rho\left(g^{-1} y_1 y_2 \cdots y_k g\right). \end{aligned} \tag{3}$$

[The existence of such a cycle and the first equality in (3) are derived using only the assumption that H is a normal subgroup of G. Furthermore, under this assumption the last equality holds up to conjugation. This is one of the reasons it may be possible in certain circumstances to apply the methods of this paper to cases where H is non-Abelian.] If furthermore H is contained in the center $Z(G)$ of G, then (3) says that the phase around the cycle is independent of $[g]$. Each such cycle gives us a slight gain in bounding eigenvalues of $\overline{P}$; thus we have the following proposition.

PROPOSITION 4. *Suppose $H \subset Z(G)$ is a subgroup of the finite group G, S is an n element symmetric generating set for G, ρ: $H \to \mathbf{C}$ is a nontrivial representation of H and $\overline{P}$ is the associated matrix. Let λ be the largest eigenvalue of $\overline{P}$ and $\lambda_{\min}$ the smallest. Suppose $y_1 y_2 \cdots y_k$ is a word in the elements of S which represents an element of H with $\rho(y_1 y_2 \cdots y_k) = e^{i\theta} \neq 1$. Let m be the maximum over $x \in S$ of the number of pairs $(i, \pm 1)$ with $1 \leq i \leq k$ and $y_k = x^{\pm 1}$. Then $\lambda \leq 1 - 2k(1 - \cos(\theta/k))/(mn)$. If k is even, then $\lambda_{\min} \geq -1 + 2k(1 - \cos(\theta/k))/(mn)$ and if k is odd, then $\lambda_{\min} \geq -1 + 2k(1 - \cos((\pi - |\theta|)/k))/(mn)$.*

PROOF. Let Γ be the twisted graph associated with $\overline{P}$ and consider the cycles in Γ given by the word $y_1 y_2 \cdots y_k$. Every edge is in at most m such cycles. Therefore, if we take every cycle with weight $1/m$, we use each edge with total weight at most 1. By (3), as one goes around any of these cycles, the phase one gets is $\rho(g^{-1} y_1 y_2 \cdots y_k g)$. Since H is contained in the center of G, this is just $\rho(y_1 y_2 \cdots y_k) = e^{i\theta}$. Therefore, by Lemma 3 each cycle is dominated by the twisted graph with no edges and $s = 2\cos(\theta/k)$ on every vertex. Each vertex is in k/m such cycles counted with weights. Hence Γ is dominated by the twisted graph with these cycles removed and $s = 2k\cos(\theta/k)/m$ at every vertex. There are (counted with weights) $n - 2k/m$ edges remaining into every vertex. Remove them with Cauchy–Schwarz (symmetrically). Thus we see that Γ is dominated by the diagonal twisted graph with $d = n$ and $s = n - 2k(1 - \cos(\theta/k))/m$ at every vertex. Hence by Lemma 2 we have $\lambda \leq 1 - 2k(1 - \cos(\theta/k))/(mn)$.

For the lower bound on $\lambda_{\min}$ apply the above argument to the matrix $-\overline{P}$. If k is even, the phase around the cycle is unchanged and we get $-\lambda_{\min} \leq 1 - 2k(1 - \cos(\theta/k))/(mn)$, as claimed. If k is odd, then the phase around the cycle is $-e^{i\theta} = e^{\pm i(\pi - |\theta|)}$. Thus we get

$$-\lambda_{\min} \leq 1 - 2k\left(1 - \cos\left((\pi - |\theta|)/k\right)\right)/(mn). \qquad \square$$

REMARK. If $H \subset Z(G)$ as in Proposition 4 and one can find two or more words in S which represent elements of H with $\rho(\cdot) \neq 1$, then one can improve the bound. This bound is easy to derive but tedious to state since it depends heavily on how the cycles overlap. For example, if we have disjoint words (no generator or its inverse is used in two different words) with lengths k_i, maximum multiplicities m_i and phases θ_i, then

$$\lambda \le 1 - \frac{2}{n} \sum_i \frac{k_i}{m_i} \left(1 - \cos\left(\frac{\theta_i}{k_i}\right)\right).$$

If H is not central but H is cyclic, then one gets an interesting variant of this proposition.

PROPOSITION 5. *Suppose $H \subset G$ is a cyclic normal subgroup of order q of the finite group G, S is an n-element symmetric generating set for G, ρ: $H \to \mathbf{C}$ is a nontrivial representation of H and $\overline{P}$ is the associated matrix. Let λ be the largest eigenvalue of $\overline{P}$ and $\lambda_{\min}$ the smallest. Suppose $y_1 y_2 \cdots y_k$ is a word in the elements of S which represents a generator of H. Let m be the maximum over $x \in S$ of the number of pairs $(i, \pm 1)$ with $1 \le i \le k$ and $y_i = x^{\pm 1}$. Then $\lambda \le 1 - 2k(1 - \cos(2\pi/qk))/(mn)$. If k is even, then $\lambda_{\min} \ge -1 + 2k(1 - \cos(2\pi/qk))/(mn)$ and if k is odd and $q \neq 2$, then $\lambda_{\min} \ge -1 + 2k(1 - \cos(\pi/qk))/(mn)$.*

PROOF. Proceed as in the proof of Proposition 4. Take each of the cycles generated by $y_1 y_2 \cdots y_k$ with weight $1/m$. The phase as one goes around the cycle is $\rho(g^{-1} y_1 y_2 \cdots y_k g)$. Since H is cyclic and normal, $g^{-1} y_1 y_2 \cdots y_k g$ is also a generator of H. Since ρ is nontrivial, $\rho(g^{-1} y_1 y_2 \cdots y_k g) \neq 1$. Since H is of order q, the argument of $\rho(g^{-1} y_1 y_2 \cdots y_k g)$ must be at least $2\pi/q$. Therefore, by Lemma 3 each cycle is dominated by the twisted graph with no edges and $s = 2\cos(2\pi/qk)$ on every vertex. Each vertex is in k/m such cycles counted with weights. Hence $\mathbf{\Gamma}$ is dominated by the twisted graph with these cycles removed and $s = 2\cos(2\pi/qk)/m$ at every vertex. There are (counted with weights) $n - 2k/m$ edges remaining into every vertex. Remove them with Cauchy–Schwarz (symmetrically). Thus we see that $\mathbf{\Gamma}$ is dominated by the diagonal twisted graph with $d = n$ and $s = n - 2k(1 - \cos(2\pi/qk))/m$ at every vertex. Hence by Lemma 2 we have $\lambda \le 1 - 2k(1 - \cos(2\pi/qk))/(mn)$.

For the lower bound on $\lambda_{\min}$ apply the above argument to the matrix $-\overline{P}$. If k is even, the phase around the cycle is unchanged and we get $-\lambda_{\min} \le 1 - 2k(1 - \cos(2\pi/qk))/(mn)$, as claimed. If k is odd, then the phase around the cycle is $-\rho(g^{-1} y_1 y_2 \cdots y_k g)$. If $q \neq 2$, then the argument of this phase is at least π/q. Thus we get $-\lambda_{\min} \le 1 - 2k(1 - \cos(\pi/qk))/(mn)$. □

4. The noncentral case. In the above discussion, if H is neither central nor cyclic, then it is possible that only a fraction of the cycles will have $\rho(\cdot) \neq 1$. In this case one must work harder to give good bounds and the statements are more complicated. Very crudely, one needs to spread the benefits from the cycles with $\rho(\cdot) \neq 1$ to every vertex of the graph. In specific examples this trick of spreading absorption around has proved to be a powerful additional tool. For example, using the results of this section will allow us to give eigenvalue bounds for certain solvable groups; see Propositions 12 and 13 in the next section. An interesting variant of this idea is used in Section 6 for the specific examples of the upper triangular matrices and the upper triangular matrices with ones on the diagonal. Also a quite pleasant special case of this argument is given in [9] for the Burnside group $B(3, n)$. This section, therefore, contains a general discussion of this technique. Unfortunately this section is rather technical. The reader is advised to skip to Section 5 on the first pass.

Suppose G is a finite group, S is an n-element symmetric generating set for G, $H \subset G$ is an Abelian normal subgroup of G, ρ: $H \to \mathbf{C}$ is a nontrivial representation of H, $\bar{P}$ is the associated matrix and $\boldsymbol{\Gamma} = (\Gamma, n, w, s)$ is the associated twisted graph. Let λ be the largest eigenvalue of $\bar{P}$. Suppose $y_1 y_2 \cdots y_k$ is a word in the elements of S which represents an element of H, and let m be the maximum over $x \in S$ of the number of pairs $(i, \pm 1)$ with $1 \leq i \leq k$ and $y_i = x^{\pm 1}$. Suppose $\{\Gamma_i\}$ is a cover of Γ by a vertex-disjoint collection of isomorphic vertex transitive subgraphs. Suppose each Γ_i has degree d. We will use the Γ_i's as follows. We will assume each Γ_i contains some vertices where we gain from the cycle construction above. We then spread that gain over all of Γ_i. The most common cases in applications would seem to be each Γ_i a vertex, an edge or a cycle.

Proceeding as in the proof of Proposition 4 take each of the cycles generated by the word $y_1 y_2 \cdots y_k$ with weight $1/(2m)$. (Thus each edge is used with total weight at most $1/2$.) For each such cycle C let $\pi \geq \theta(C) > -\pi$ be the argument of the phase one gets by going around the cycle. By the above discussion, if the cycle C starts at $[g]$, then $\rho(g^{-1} y_1 y_2 \cdots y_k g) = e^{i\theta(C)}$. By Lemma 3, replacing each of these cycles by a stationary weight of $\cos(\theta(C)/k)/m$ at every vertex on the cycle gives a dominating twisted graph. Thus we get an absorption weight at the vertex v of

$$a(v) = (1/m) \sum_{C \ni v} (1 - \cos(\theta(C)/k)), \tag{4}$$

where the sum runs over all cycles containing v. This will be the only gain we will exploit in this argument. Note that the largest eigenvalue of a Hermitian matrix cannot decrease if we replace every entry by its absolute value. Hence in our upper bound we may, without loss, replace the weight on every remaining edge by its absolute value. Thus we may assume each edge has real weight at least $1/2$. In particular, we have a copy of each subgraph Γ_i with weight $1/2$ among the remaining edges. Replace all other edges and any extra weight on the edges of the Γ_i by a stationary weight at the vertices using Cauchy–Schwarz symmetrically. This leaves the absorption at each vertex unchanged. Thus we see that the largest eigenvalue of $\bar{P}$ is at most the largest eigenvalue of the twisted graph $\boldsymbol{\Gamma}' = (\Gamma', n, w', s')$ whose underlying graph is the union of the Γ_i, $w' = 1/2$ on every edge and $s(v) = (n - d/2) - a(v)$, where $a(v)$ is given by (4) above. It suffices to find upper bounds for the eigenvalues of each component of $\mathbf{\Gamma}'$ separately.

For each subgraph Γ_i define the score of that subgraph to be

$$\mathrm{sc}(\Gamma_i) = \sum_{v \in \Gamma_i} a(v) = (1/m) \sum_{v \in \Gamma_i} \sum_{C \ni v} (1 - \cos(\theta(C)/k)).$$

That is, the score is the total absorption weight at all the vertices of Γ_i. For each vertex $v \in \Gamma_i$ consider the following twisted graph $\mathbf{\Gamma_i(v)}$. The underlying graph is Γ_i, the degree is $d' = da(v)/2\,\mathrm{sc}(\Gamma_i)$, the weight is $a(v)/2\,\mathrm{sc}(\Gamma_i)$ on each edge, $s(v) = -a(v)$ and $s(w) = 0$ for $w \neq v$. The union over all vertices in Γ_i of the twisted graphs $\mathbf{\Gamma_i(v)}$ and the twisted graph with no edges and a stationary weight of $(n - d/2)$ at every vertex of Γ_i is the component of $\mathbf{\Gamma}'$ with underlying graph Γ_i. Let A be the adjacency matrix of the graph Γ_i. The Hermitian matrix associated with $\mathbf{\Gamma_i(v)}$ is $(a(v)A/2\,\mathrm{sc}(\Gamma_i) - a(v)E_{vv})/d'$, where E_{vv} denotes the matrix whose only nonzero entry is a 1 in the v, v entry. The largest eigenvalue of this matrix is the largest root of

$$\det\bigl(tI - (a(v)A/2\,\mathrm{sc}(\Gamma_i) - a(v)E_{vv})/d'\bigr) = 0.$$

Thus the twisted graph $\mathbf{\Gamma_i(v)}$ is dominated by the diagonal twisted graph of degree d' with every stationary weight equal to the largest root of $\det(tI - a(v)A/2\,\mathrm{sc}(\Gamma_i) + a(v)E_{vv}) = 0$. Note that this root is

$$a(v)\bigl(d - \{\text{smallest root of } \det((d-t)I - A + 2\,\mathrm{sc}(\Gamma_i)E_{vv}) = 0\}\bigr)/2\,\mathrm{sc}(\Gamma_i).$$

Adding up all these contributions for all vertices $v \in \Gamma_i$ we see that $\mathbf{\Gamma}'$ is dominated by the diagonal twisted graph of degree n with stationary weight

$$s(v) = n - 1/2\{\text{the smallest root of } \det((d-t)I - A + 2\,\mathrm{sc}(\Gamma_i)E_{vv}) = 0\}.$$

Let $g(z) = \det((d - z)I - A)$ and N be the number of vertices in Γ_i. Then

$$\det((d-t)I - A - 2\,\mathrm{sc}(\Gamma_i)E_{vv}) = g(t) - 2\,\mathrm{sc}(\Gamma_i)g'(t)/N.$$

Note that in deriving this last formula we have used the fact that Γ_i is vertex-transitive for the first time. Thus we get the following proposition.

PROPOSITION 6. *Let G, S, H, $\rho: H \to \mathbf{C}$, $\overline{P}$, $\mathbf{\Gamma} = (\Gamma, n, w, s)$, λ, $y_1 y_2 \cdots y_k$, m, $\{\Gamma_i\}$, d, sc, A, N and $g(z)$ be as above. Let $\varepsilon = \min_i \mathrm{sc}(\Gamma_i)$. Then $\lambda \leq 1 - t/(2n)$, where t is the smallest root of $g(t) = 2\varepsilon g'(t)/N$.*

REMARK 1. One can get lower bounds on the smallest eigenvalue of $\overline{P}$ from the same data exactly as in Proposition 4. If k is even, then $\lambda_{\min} \geq -1 + t/(2n)$. If k is odd, then we define a modified score sc$'$ by

$$\mathrm{sc}'(\Gamma_i) = (1/m) \sum_{v \in \Gamma_i} \sum_{C \ni v} (1 - \cos((\pi - \theta(C))/k))$$

and let $\delta = \min_i \mathrm{sc}'(\Gamma_i)$. Then $\lambda_{\min} \geq -1 + t/(2n)$, where t is the smallest root of $g(t) = 2\delta g'(t)/N$.

REMARK 2. In many applications the cycle $y_1 y_2 \cdots y_k$ and the Γ_i are edge disjoint. In this case one can increase the weights in the proof by a factor of 2. This gives the slightly stronger bound $\lambda \le 1 - t/n$.

REMARK 3. More complicated statements are possible, for example, using several different types of Γ_i's or a collection of Γ_i's which cover every vertex several times. One of these will be used below. If we have a number of different disjoint words (as in Remark 1 after Proposition 4), then we may regard each as contributing separately to the score. That is,

$$\mathrm{sc}(\Gamma_j) = \sum_i (1/m) \sum_{v \in \Gamma_j} \sum_{C \ni v} (1 - \cos(\theta_i(C)/k_i)).$$

Proposition 6 reduces the upper bound to studying an equation involving the characteristic polynomial of the Laplacian on the graphs Γ_i. As such, it may be a little tedious to apply directly. For vertices, edges and cycles one has an explicit expression for $g(z)$ and one can give acceptable answers. If each Γ_i is a vertex, then $N = 1$ and $g(z) = -z$. Hence we get $t = 2\varepsilon$ and $\lambda \le 1 - \varepsilon/n$. If each Γ_i is an edge, then $N = 2$ and $g(z) = (z-1)^2 - 1$. Hence we get $t = 1 + \varepsilon - \sqrt{1+\varepsilon^2}$. If each Γ_i is a cycle of length N, then

$$g(2 - 2\cos\theta) = 2\cos N\theta - 2.$$

Hence we get $t = 2 - 2\cos\theta$, where $\cos N\theta + 2\varepsilon \sin N\theta/\sin\theta = 1$. These answers are still not transparent. The following weaker bounds may be more useful in practice.

LEMMA 7. (a) *Let g be as above and let K be the smallest root of $g'(K) = 0$. Then the smallest root t of $g(t) = 2\varepsilon g'(t)/N$ satisfies the bound $t \ge 2\varepsilon K/(KN + 2\varepsilon)$. In particular, with notation as in Proposition* 6, $\lambda \le 1 - \varepsilon K/((KN + 2\varepsilon)n)$.

(b) *If $\alpha_1 < d$ is the second largest eigenvalue of the adjacency matrix A of Γ_i, then*

$$t \ge \frac{2\varepsilon(d-\alpha_1)}{(d-\alpha_1+2\varepsilon)N} \quad \textit{and} \quad \lambda \le 1 - \frac{\varepsilon(d-\alpha_1)}{(d-\alpha_1+2\varepsilon)Nn}.$$

PROOF. Let $d = \alpha_0 > \alpha_1 \ge \cdots \ge \alpha_{N-1} \ge -d$ be the eigenvalues of the adjacency matrix A of Γ_i. Then $g(z) = \prod_{i=0}^{N-1}(d - z - \alpha_i)$ and t is the smallest root of

$$\frac{g'(t)}{g(t)} = \sum_{i=0}^{N-1} \frac{1}{t - (d - \alpha_i)} = \frac{N}{2\varepsilon},$$

or equivalently

$$(\#) \qquad \frac{1}{t} = \frac{N}{2\varepsilon} + \sum_{i=1}^{N-1} \frac{1}{(d-\alpha_i) - t}.$$

By continuity this root must occur for some t in the range $0 < t < K < d - \alpha_1$. The right-hand side of equation (#) will only increase if we replace t by the larger number K and since $g'(K) = 0$ at $t = K$, the rightmost summand is $1/K$. Therefore

$$\frac{1}{t} \leq \frac{N}{2\varepsilon} + \frac{1}{K},$$

which gives the bound claimed in part (a).

For part (b) note that for $0 < z < d - \alpha_1$ we have

$$\frac{g'(z)}{g(z)} \geq \frac{1}{z} + \frac{N-1}{z - (d - \alpha_1)}.$$

Hence $K \geq (d - \alpha_1)/N$. Thus we get the bound claimed in part (b). □

COROLLARY 7.1. (a) *If each* Γ_i *is a vertex, then* $t = 2\varepsilon$ *and* $\lambda \leq 1 - \varepsilon/n$.
(b) *If each* Γ_i *is an edge, then* $t \geq \varepsilon/(1 + \varepsilon)$ *and* $\lambda \leq 1 - \varepsilon/(2(1 + \varepsilon)n)$.
(c) *If each* Γ_i *is a cycle of length* $N \geq 3$, *then* $t \geq 2\varepsilon/(N + \varepsilon N^2)$ *and* $\lambda \leq 1 - \varepsilon/((N + \varepsilon N^2)n)$.

PROOF. For a vertex $K = \infty$, for an edge $K = 1$ and for a cycle of length $N \geq 3$,

$$K = 2(1 - \cos(\pi/N)) \geq 9(2 - \sqrt{3})N^{-2} \geq 2N^{-2}. \qquad \square$$

It is worth noting that the bounds in Lemma 7 say something important about the limits of Proposition 6. The bound in Lemma 7 is asymptotically correct if ε is large. In this case the lemma says Γ_i controls our upper bound and λ is approximately bounded by $1 - K/(2n) \leq 1 - (d - \alpha_1)/(2Nn)$. If ε is small compared to KN or $d - \alpha_1$, then λ is approximately bounded by $1 - \varepsilon/(Nn)$. Since ε/N is the average absorption weight on Γ_i, we have spread the gain over all vertices evenly with only negligible loss.

5. Applications. As a first application of the method above we have the following result for nilpotent groups. To fix notation, recall that if G is any finite group, we define a sequence of normal subgroups $\gamma_i(G)$ of G (the lower central series of G) by $\gamma_1(G) = G$, $\gamma_2(G) = [G, G]$ is the commutator subgroup of G and, inductively, $\gamma_{i+1}(G) = [G, \gamma_i(G)]$. We say G is nilpotent if $\gamma_i(G) = \{id\}$ for some i. The largest l for which $\gamma_l(G) \neq \{id\}$ is called the class number of G. If A is an Abelian group, let $\exp(A)$ be the least positive integer n such that $x^n = id$ for all $x \in A$. To simplify some of the bounds below we will use the following lemma, whose proof is left as an exercise.

LEMMA 8. (a) *If* $-\pi \leq x \leq \pi$, *then* $\cos x \leq 1 - 2x^2/\pi^2$.
(b) *If* $-\pi \leq x \leq \pi$, *then* $\cos x \leq 1 - x^2/(2 + 3x^2/\pi^2)$.
(c) *If* $-\pi/3 \leq x \leq \pi/3$, *then* $\cos x \leq 1 - 9x^2/(2\pi^2)$.

THEOREM 9. *Let G be a nilpotent group with class number l. Let S be an n-element symmetric generating set for G. Let P be the transition probability matrix for the associated random walk on G, and let $1 = \beta_0 > \beta_1 \geq \beta_2 \geq \cdots \geq \beta_{N-1} \geq -1$ be the eigenvalues of P. Let*

$$e = \max_{1 \leq i \leq l} \{2^{-2(l-i)} \exp(\gamma_i(G)/\gamma_{i+1}(G))\}.$$

Then $\beta_1 \leq 1 - 4\pi^2/(3e^2 2^{2l} n)$. If no element of S has order 2 in G, then $\beta_1 \leq 1 - 8\pi^2/(3e^2 2^{2l} n)$. If the underlying graph is not bipartite (i.e., $\beta_{N-1} \neq -1$), then $\beta_{N-1} \geq -1 + 4\pi^2/(3e^2 2^{2l} n)$, and if no element of S has order 2 in G, then $\beta_{N-1} \geq -1 + 8\pi^2/(3e^2 2^{2l} n)$.

PROOF. We will first prove the bounds on β_1. We proceed by induction on l. The case $l = 1$, that is, G Abelian, is straightforward. If G is Abelian and $S = \{x_1, x_2, \dots, x_n\}$, then the eigenvalues are all of the form $\lambda = (1/n)\sum_{i=1}^{n} \chi(x_i)$ for some character $\chi: G \to \mathbf{C}$. If χ is nontrivial, that is, $\lambda \neq 1$, then there must be some generator x_i with $\chi(x_i) = \omega \neq 1$. If x_i is not of order 2 in G, then $x_i^{-1} \in S$ and hence $\lambda \leq (n - 2 + 2\,\mathrm{Re}\,\omega)/n$. If $e = \exp(G)$, then

$$\mathrm{Re}\,\omega \leq \cos(2\pi/e).$$

Therefore, $\lambda \leq 1 - 2(1 - \cos(2\pi/e))/n$. By Lemma 8(a), $\lambda \leq 1 - 16/(e^2 n)$. If x_i is of order 2, then $\chi(x_i) = -1$ and one gets instead $\lambda \leq 1 - 2/n \leq 1 - 8/(e^2 n)$. In either event this completes the case $l = 1$.

Suppose now that G is of class l and $S = \{x_1, x_2, \dots, x_n\}$. Applying Theorem 1 to the Abelian normal subgroup $H = \gamma_l(G)$, we see that the inductive hypothesis provides the desired bound on the second largest eigenvalue of the associated random walk on G/H. Thus we need only show that the desired bound also holds for every matrix $\bar{P}$ coming from (2) for a nontrivial character $\rho: H = \gamma_l(G) \to \mathbf{C}$. Fix ρ a nontrivial representation of H and let λ be the largest eigenvalue of the corresponding matrix $\bar{P}$.

By induction on i one sees that $\gamma_i(G)$ is normally generated by elements of the form $[x_{j_1}, [x_{j_2}, [x_{j_3}, \dots, [x_{j_{i-1}}, x_{j_i}]] \cdots]]$ with $x_{j_{i-1}} \neq x_{j_i}$. In particular, $H = \gamma_l(G)$ is normally generated by elements of the form $[x_{j_1}, [x_{j_2}, [x_{j_3}, \dots, [x_{j_{l-1}}, x_{j_l}]] \cdots]]$. However, such l-fold commutators lie in the center of G since G is nilpotent of class l. Therefore, in fact H must be generated by these elements. In particular, since ρ is a nontrivial representation, there must be some choice of indices for which

$$\rho\Big(\big[x_{j_1}, \big[x_{j_2}, \big[x_{j_3}, \dots, \big[x_{j_{l-1}}, x_{j_l}\big]\big] \cdots\big]\big]\Big) = e^{i\theta} \neq 1.$$

Fix such a choice. This l-fold commutator is a word $x_{j_1} x_{j_2} \cdots (x_{j_1})^{-1} \cdots (x_{j_2})^{-1}$ of length $L = 3 \cdot 2^{l-1} - 2$ in the generators $\{x_1, x_2, \dots, x_n\}$ which lies in H. The most times any element of S or its inverse can occur in L is $m = 2^l - 2$ if no element of S has order 2 in G or $m = 2^{l+1} - 4$ otherwise. Assume no element of S has order 2 in G. By Proposition 4 above we have

$$\lambda \leq 1 - 2L(1 - \cos(\theta/L))/((2^l - 2)n).$$

Furthermore, since $\omega = e^{i\theta}$ is a value of ρ on an element of H and $e \geq \exp(H)$ we must have $|\theta| \geq 2\pi/e$. Hence $1 - \cos(\theta/L) \geq 1 - \cos(2\pi/(eL)) \geq 2\pi^2/(e^2L^2 + 6)$, where the last inequality follows by Lemma 8(b). Therefore, $\lambda \leq 1 - 4\pi^2 L/((e^2L^2 + 6)(2^l - 2)n)$. Combining this with the trivial bounds $2^l - 2 \leq 2^l$ and $e^2L^2 + 6 \leq e^2L(L+2) \leq 3 \cdot 2^{l-1}e^2L$ gives $\lambda \leq 1 - 8\pi/(3e^2 2^{2l} n)$, as desired. If some x_k has order 2 in G, then we lose a factor of 2 in the upper bound and get the stated value.

For the bounds on β_{N-1} we use the same inductive argument. All the cycles used in the above proof were of even length; hence, the argument for the inductive step is exactly the same as the above. A new argument is required only for the case $l = 1$. If G is Abelian and $S = \{x_1, x_2, \dots, x_n\}$, then as above the eigenvalues are all of the form $\lambda = (1/n)\sum_{i=1}^{n} \chi(x_i)$ for some character $\chi: G \to \mathbf{C}$. If the underlying graph is not bipartite, the $\chi(x_i)$ cannot all be -1. Thus there must be some generator x_i with $\chi(x_i) = \omega \neq -1$. If x_i is not of order 2 in G, then $x_i^{-1} \in S$ and hence $\lambda \geq (-n + 2 + 2\operatorname{Re}\omega)/n$. If x_i has order $m \leq e = \exp(G)$, then

$$\operatorname{Re}\omega = \cos(2\pi k/m) = -\cos(\pi(2k - m)/m).$$

Since this is not -1, at least one must have $\operatorname{Re}\omega \geq -\cos(\pi/m) \geq -\cos(\pi/e)$. Therefore, $\lambda \geq -1 + 2(1 - \cos(\pi/e))/n$. By Lemma 8(c), $\lambda \geq -1 + 9/(e^2n)$. If x_i is of order 2, then $\chi(x_i) = 1$ and one gets instead $\lambda \geq -1 + 2/n \geq -1 + 8/(e^2n)$. □

The bound above relies on a cycle one has in any nilpotent group of class l. This cycle is quite long for l large and can contain one element inordinately often. For a specific nilpotent group one can hope to greatly improve this bound. For example, for p an odd prime, let $U_n(p)$ denote the group of upper triangular matrices over $\mathbf{Z}/p\mathbf{Z}$ with 1's on the diagonal. Let E_k, $1 \leq k \leq n-1$, be the matrix all of whose off-diagonal entries are zero except the $k, k+1$ entry, which is 1. Then $S = \{E_1^{\pm 1}, \dots, E_{n-1}^{\pm 1}\}$ is a generating set for $U_n(p)$. Sharp bounds on β_1 for the associated random walk on $U_n(p)$ are given in [8] and in Section 6 below. It is shown that there are constants c_1 and c_2 such that $1 - c_1/(p^2n) \geq \beta_1 \geq 1 - c_2/(p^2n)$. A direct application of Proposition 4 gives the following weaker result. It should be noted that even this weak result is stronger than can be obtained by path methods of [6] and [4].

PROPOSITION 10. *Let β_1 be the second largest eigenvalue for the random walk on $U_n(p)$ defined above. Then* $\beta_1 \leq \cos(\pi/(2p(n-1)))$.

PROOF. The argument proceeds almost exactly as in the proof of Theorem 9 above. However, instead of the cycle coming from an iterated commutator which we used in that proof, we use cycles of the form

$$E_{i+k-1}E_{i+k-2} \cdots E_{i+1}E_i^{-1}E_{i+1}^{-1} \cdots E_{i+k-1}^{-1}E_{i+k}E_{i+k-1} \cdots E_{i+1}E_iE_{i+1}^{-1}E_{i+2}^{-1} \cdots E_{i+k}^{-1},$$

for $1 \le k \le n-2$, $1 \le i \le n-k-1$. (These cycles were used by Ellenberg to study the diameter of $U_n(p)$ [5].) This cycle represents the matrix all of whose off-diagonal entries are zero except the $i, i+k+1$ entry, which is 1 [5]. For k fixed, these elements generate $\gamma_{k+1}(G)/\gamma_{k+2}(G)$. The inductive argument goes as above except that the relevant cycle now has length $L = 4k$ and no generator occurs more than $m = 4$ times in the cycle. Hence Proposition 4 gives $\lambda \le 1 - 2k(1 - \cos(\theta/(4k)))/(2n-2)$. The exponent of $\gamma_{k+1}(G)/\gamma_{k+2}(G)$ is p; hence, $\lambda \le 1 - 2k(1 - \cos(\pi/(2pk)))/(2n-2)$. This upper bound is weakest when $k = n-1$, in which case it becomes $\lambda \le \cos(\pi/(2p(n-1)))$. □

Thus one sees that Theorem 1 and Proposition 4 above give a pleasant method for bounding eigenvalues of random walks on nilpotent groups. Other examples can be done using Proposition 5 instead. Following [3], let p be an odd prime and let G_N be the set of polynomials with coefficients in the finite field $\mathbf{Z}/p\mathbf{Z}$ taken mod x^{N+1} of the form $a_1 x + a_2 x^2 + \cdots + a_N x^N$ with $a_1 \in (\mathbf{Z}/p\mathbf{Z})^*$ and $a_i \in \mathbf{Z}/p\mathbf{Z}$ for $2 \le i \le N$. Then G_N is a group under composition and if α is a multiplicative generator for $(\mathbf{Z}/p\mathbf{Z})^*$, then $S = \{\alpha x, \alpha^{-1}x, x + x^2, (x+x^2)^{-1}, x + x^3, \dots, x + x^N, (x + x^N)^{-1}\}$ is a natural symmetric generating set. See [7] for more about these groups.

PROPOSITION 11. *Let G_N and S be as above, and let $1 = \beta_0 > \beta_1 \ge \beta_2 \ge \cdots \ge \beta_{|G|-1} \ge -1$ be the eigenvalues of the associated transition probability matrix P. Then*

$$\beta_1 = 1 - (1 - \cos(2\pi/(p-1)))/N \le 1 - 8/(p^2 N).$$

PROOF. We proceed by induction on N using the exact sequence

$$0 \to H = \mathbf{Z}/p\mathbf{Z} \to G_N \to G_{N-1} \to 0.$$

If $N = 1$, then G_N is the group $(\mathbf{Z}/p\mathbf{Z})^* \simeq \mathbf{Z}/(p-1)\mathbf{Z}$ and the second largest eigenvalue is $\beta_1 = \cos(2\pi/(p-1)) \le 1 - 8/p^2$. If $N \ge 2$, then we are reduced by Theorem 1 to bounding the eigenvalues of an induced random walk on G_{N-1} and the twisted versions for all nontrivial representations of H. The induced random walk on G_{N-1} has the identity occurring twice in the generating set. Therefore, its second largest eigenvalue satisfies

$$\begin{aligned}\lambda &= (N-1)/N \cdot \beta_1(G_{N-1}) + 1/N = 1 - (1 - \cos(2\pi/(p-1)))/N \\ &\le 1 - 8/(p^2 N).\end{aligned}$$

If ρ is a nontrivial representation of H, then we apply Proposition 5 with the word $y_1 = (x + x^N)$, which has $k = m = 1$. We conclude that the second largest eigenvalue of $\overline{P}$ satisfies

$$\lambda \le 1 - (1 - \cos(2\pi/p))/N < 1 - (1 - \cos(2\pi/(p-1)))/N.$$

Therefore, the second largest eigenvalue for G_N comes from the trivial representation of H and is $1 - (1 - \cos(2\pi/(p-1)))/N$. □

It is also possible to give nice bounds using Proposition 6. The following two results are examples.

PROPOSITION 12. *Suppose $G = H \rtimes K$ for H and K Abelian and S is a union of a generating set for K and one for H. Let P be the associated transition probability matrix and let β_1 be its second largest eigenvalue. Let $e = \max(\exp(H), \exp(K))$ and $n = |S|$. Then $\beta_1 \leq 1 - 8/(e^2 n)$.*

PROOF. Let $S = S_K \cup S_H$. By Theorem 1 it suffices to bound the eigenvalues of the induced random walk on K and the eigenvalues of the matrix $\overline{P}$ associated with each nontrivial representation of H. The induced random walk on K is the one generated by the n-element set consisting of S_K together with the identity element with multiplicity $|S_H|$. Therefore, by Theorem 8 (for the trivial case $l = 1$) it has second largest eigenvalue bounded above by $1 - 8/(e^2 n)$. (Note that here we have used the slightly stronger bound contained in the proof of Theorem 9 for $l = 1$.)

Next suppose ρ is a nontrivial representation of H and $\overline{P}$ is the associated matrix. Apply Proposition 6, where we use several disjoint words. If $S_H = \{x_1^{\pm 1}, x_2^{\pm 1}, \dots, x_r^{\pm 1}\}$, then take the words to be $x_1, x_2, \dots, x_r$. Suppose $g \in K$ is viewed as a vertex of the quotient. Since the conjugates $g^{-1}x_1 g, g^{-1}x_2 g, \dots, g^{-1}x_r g$ generate H, at least one of the phases $\rho(g^{-1}x_i g)$ must be nontrivial. Additionally, since H has exponent at most e, its argument must be at least $2\pi/e$. Apply Proposition 6 (as generalized in Remark 3) with each Γ_i a vertex. Since at least one of the terms in the score must have argument at least $2\pi/e$, we have $\mathrm{sc}(\Gamma_i) \geq (1 - \cos(2\pi/e))/2$ (the extra 2 since we may have an x_i of order 2, hence $m = 2$). Therefore, $\lambda \leq 1 - (1 - \cos(2\pi/e))/n \leq 1 - 8/(e^2 n)$ (where we have gained an extra factor of 2 since the Γ_i and the cycles are edge-disjoint). □

PROPOSITION 13. *Suppose $G = H \rtimes \mathbf{Z}/N\mathbf{Z}$ with H Abelian, $N \geq 3$ and $S = \{x_1, x_2, \dots, x_n\}$ is a symmetric generating set for G for which $x_1 = (x_2)^{-1}$ is a generator of $\mathbf{Z}/N\mathbf{Z}$ and $\{x_3, \dots, x_n\} \subset H$. (Note that $\{x_3, \dots, x_n\}$ are not assumed to generate H.) Let β_1 be the second largest eigenvalue of the associated transition probability matrix P. Let $e = \exp(H)$. Then $\beta_1 \leq 1 - 8/((e^2 N + 4N^2)n)$.*

PROOF. By Theorem 1 it suffices to bound the second largest eigenvalue for the induced random walk on $\mathbf{Z}/N\mathbf{Z}$ and the eigenvalues of the matrix $\overline{P}$ associated with each nontrivial representation of H. The induced random walk on $\mathbf{Z}/N\mathbf{Z}$ has the identity occurring with multiplicity $n - 2$; hence, its second largest eigenvalue is

$$(n - 2 + 2\cos(2\pi/N))/n \leq 1 - 27/(N^2 n).$$

Let ρ be a nontrivial representation of H and $\overline{P}$ be the associated matrix. To bound the largest eigenvalue λ of $\overline{P}$ we will apply Proposition 6 (as generalized in Remark 3). Take the cycles to be half of the x_i's, $3 \leq i \leq n$, one from each generator/inverse pair as in Proposition 12 and take Γ_1 to be all of $\mathbf{Z}/N\mathbf{Z}$. At the vertex $[g] \in G/H$ the phase on the cycle generated by x_i is $\rho(g^{-1}x_i g)$. As the $g^{-1}x_i g$ range over all g and all x_i, they give a generating set for H. Hence at least one of the phases $\rho(g^{-1}x_i g)$ must be nontrivial and since H has exponent e, it must have argument at least $2\pi/e$. Therefore, at least one of the vertices must have absorption of at least $(1 - \cos(2\pi/e))/2 \geq 4/e^2$ (the extra 2 since we may have an x_i of order 2). Thus the score of the one Γ_1 is at least $4/e^2$. By Corollary 7.1(c), $\lambda \leq 1 - 8/((e^2N + 4N^2)n)$. (An extra factor of 2 is added since the cycles and Γ_1 are edge-disjoint.) □

The bound in Proposition 13 is essentially sharp if G is a wreath product $(\mathbf{Z}/e\mathbf{Z})^N \rtimes \mathbf{Z}/N\mathbf{Z}$, with the obvious four-element symmetric generating set. However, since the argument assumes that the various conjugates of the x_i's are independent, it is probably not very good for most extensions. Better bounds in those cases will require more detailed information about the group.

6. More applications. In this section we give another argument for using a noncentral Abelian subgroup. This argument can be reformulated as using a cover by a collection of subgraphs as in Section 4. However, instead of using the crude bounds given in Section 4, one uses bounds which depend on which subset of the vertices has absorption. This argument is very closely related to the one given in [8]. As in Proposition 10 above, let p be an odd integer and let $U_n(p) \subset \mathrm{GL}(n, p)$ be the group of upper triangular $n \times n$ matrices over $\mathbf{Z}/p\mathbf{Z}$ with 1's on the diagonal. Let E_k, $1 \leq k \leq n - 1$, be the $n \times n$ matrix with 1's on the diagonal and in the $k, k + 1$ entry and 0's elsewhere. Then $S = \{E_1^{\pm 1}, \dots, E_{n-1}^{\pm 1}\}$ is a natural generating set for $U_n(p)$. If p is prime, let $T_n(p) \subset \mathrm{GL}(n, p)$ be the group of upper triangular $n \times n$ matrices over $\mathbf{Z}/p\mathbf{Z}$ with nonzero entries on the diagonal. Let α be a multiplicative generator of $(\mathbf{Z}/p\mathbf{Z})^*$. Let $D_i(\gamma)$ denote the diagonal matrix with a γ in the i, i entry and 1's in all other diagonal entries. Then $S' = \{D_1(\alpha), D_1(\alpha^{-1}), \dots, D_n(\alpha^{-1}), E_1^{\pm 1}, \dots, E_{n-1}^{\pm 1}\}$ is a natural generating set for $T_n(p)$. The goal of this section will be to prove the following bounds.

THEOREM 14. (a) *Let β_1 be the second largest eigenvalue of the transition probability matrix P_n for the random walk on $U_n(p)$ generated by S. Then there is a constant $C > 0$ such that $\beta_1 \leq 1 - C/(p^2 n)$.*

(b) *Let β_1 be the second largest eigenvalue of the transition probability matrix P_n for the random walk on $T_n(p)$ generated by S'. Then there is a constant $C' > 0$ such that $\beta_1 \leq 1 - C'/(p^2 n)$.*

PROOF. The proofs of (a) and (b) are almost identical. We will give the proof of (a) in its entirety and sketch the changes needed to extend it to $T_n(p)$. We will proceed by induction on n, the case $n = 2$ being the usual random walk on $\mathbf{Z}/p\mathbf{Z}$. Let H be the normal subgroup of $U_n(p)$ consisting of all matrices whose only nonzero entries are the ones on the diagonal or are in the nth column. Identify the quotient $U_n(p)/H$ with $U_{n-1}(p)$ which we view as being contained in $U_n(p)$ as the matrices with zeros above the diagonal in the nth column (in particular, this gives us our required coset representatives.) By Theorem 1 it suffices to bound the second largest eigenvalue of the induced random walk on $U_{n-1}(p)$ and the eigenvalues of the matrix $\overline{P}$ associated with each nontrivial representation of H.

The induced random walk on $U_{n-1}(p)$ is just the random walk given by the usual generating set for $U_{n-1}(p)$ together with the identity element with multiplicity 2. Therefore, it is given by the matrix $P' = ((n-2)P_{n-1} + I)/(n-1)$. The second largest eigenvalue of this matrix is

$$\lambda = \big((n-2)\beta_1(U_{n-1}(p)) + 1\big)/(n-1) \le 1 - c(n-2)/\big(p^2(n-1)^2\big)$$
$$\le 1 - c/(p^2 n).$$

Now we turn to the bounds on the eigenvalues of $\overline{P}$ for nontrivial representations of H. Choose any nontrivial character $\rho\colon H \to \mathbf{C}$ and view ρ as also being a map

$$\rho\colon \{(n-1) \times 1 \text{ column vectors over } \mathbf{Z}/p\mathbf{Z}\} \to \mathbf{C}.$$

If $A \in U_{n-1}(p)$ we may therefore evaluate ρ on any column of A. Let $A_{\cdot i}$ denote the ith column of A. Since all the generators other than $E_{n-1}^{\pm 1}$ lie in $U_{n-1}(p)$, they map coset representatives to coset representatives. Hence the edges they generate all have weight 1. Only the edges generated by $E_{n-1}^{\pm 1}$ (which are loops) get nontrivial phases. In fact, if $A \in U_{n-1}(p)$, then the loop generated by E_{n-1} at A gets a phase of $\rho(A^{-1}E_{n-1}A) = \rho((A^{-1})_{\cdot n-1})$. Therefore, the twisted graph $\boldsymbol{\Gamma}$ associated with $\overline{P}$ has $d = 2(n-1)$, all phases 1 and a stationary weight of $s(A) = 2\,\mathrm{Re}(\rho((A^{-1})_{\cdot n-1}))$ or equivalently an absorption weight of $a(A) = 2(1 - \mathrm{Re}(\rho((A^{-1})_{\cdot n-1})))$ at the vertex A.

We wish to spread this absorption over all the vertices. Instead of a cover of $U_{n-1}(p)$ by graphs, we will use the following definition. Call a matrix $A \in U_{n-1}(p)$ of type k if k is the least positive integer with $\rho((A^{-1})_{\cdot n-k}) \ne 1$. Note that since ρ is nontrivial and the diagonal entries of A are all 1's, every A must have some type between 1 and $n-1$. The above says that a vertex A has nonzero absorption if and only if it is of type 1. Also note that any vertex of type 1 must have $a(A) \ge 2(1 - \cos(2\pi/p))$. Consider the p-cycles in $U_{n-1}(p)$ generated by E_{n-2}. Since $E_{n-2}A$ means add the $(n-1)$st row of A to the $(n-2)$nd row of A, $(E_{n-2}A)^{-1}$ means subtract the $(n-2)$nd column of A^{-1} from the $(n-1)$st column of A^{-1}. Suppose one of the p-cycles generated by E_{n-2} contains an element of type 2. Then it must contain exactly one element of type 2 and the remaining $p-1$ elements are of type 1. If a p-cycle generated by E_{n-2} does not contain an element of type 2, remove it by applying Cauchy–Schwarz symmetrically (leaving the absorption unchanged). If it does contain an element of type 2, we wish to remove it using Cauchy–Schwarz asymmetrically. For this purpose we have the following lemma whose proof is an easy exercise in applying Cauchy–Schwarz. For a complete proof see [8], Lemma 4. (Note that this lemma really says a certain asymmetric diagonal twisted graph dominates a cycle.)

LEMMA 15. *There is a function* $f(\varepsilon) > 0$ (*depending on* p) *such that, for any* $0 < \varepsilon < 1.5p^{-2}$ *and any real numbers* $\phi_0, \phi_1, \dots, \phi_{p-1}$,

$$\phi_0\phi_1 + \phi_1\phi_2 + \cdots + \phi_{p-2}\phi_{p-1} + \phi_{p-1}\phi_0 \le (1 - f(\varepsilon))\phi_0^2 + (1+\varepsilon)\sum_{i=1}^{p-1}\phi_i^2.$$

Furthermore $(p-1)\varepsilon \ge f(\varepsilon) \ge (p-1)\varepsilon - p^3\varepsilon^2/6$.

Applying this lemma to a p-cycle with one element of type 2 and $p-1$ elements of type 1, we see that we may remove the p-cycle, leaving an absorption weight of $a(A) \ge 2(1 - \cos(2\pi/p) - \varepsilon)$ at the elements of type 1 and an absorption weight of $2f(\varepsilon)$ at the vertex of type 2. Here ε will be fixed for the time being and will be determined later.

Repeat this argument looking at the p-cycles generated by E_{n-3}. If such a p-cycle contains a matrix of type 3, then the remaining $p-1$ matrices in the cycle must all be of type 2. Remove the p-cycles that do not contain an element of type 3 symmetrically with Cauchy–Schwarz (leaving absorptions unchanged). Remove the p-cycles containing an element of type 3 asymmetrically using Lemma 15. Doing so we may leave an absorption weight of $2(f(\varepsilon) - \varepsilon)$ at the vertices of type 2 and $2f(\varepsilon)$ at the vertex of type 3. Inductively one can remove all edges and spread the absorption over all the vertices, leaving each with an absorption weight of at least

$$a(A) \ge 2(1 - \cos(2\pi/p) - \varepsilon) \ge 27/p^2 - 2\varepsilon$$

at the elements of type 1 and at least

$$a(A) \ge 2(f(\varepsilon) - \varepsilon) \ge 2(p-2)\varepsilon - p^3\varepsilon^2/3$$

at all other vertices. Taking $\varepsilon = cp^{-3}$ for some small constant $c > 0$ gives $a(A) \ge c'p^{-2}$ at all vertices for some other constant $c' > 0$. Thus $\mathbf{G}$ is dominated by the diagonal twisted graph with $d = 2(n-1)$ and a stationary weight of at most $2(n-1) - c'p^{-2}$ at every vertex. Thus the largest eigenvalue is bounded by $1 - C/(p^2 n)$, for some $C > 0$.

The proof of (b) is virtually identical. We let H be the normal subgroup of $T_n(p)$ consisting of all matrices whose only nonzero entries are on the diagonal or are in the nth column and which have 1's on the diagonal except possibly in the last column. Again the quotient $T_n(p)/H$ can be identified with $T_{n-1}(p)$, which we view as being contained in $T_n(p)$. By Theorem 1 it again suffices to bound the second largest eigenvalue of the induced random walk on $T_{n-1}(p)$ and the eigenvalues of the matrix $\overline{P}$ associated to each nontrivial representation of H.

The induced random walk on $T_{n-1}(p)$ is just the random walk given by the usual generating set for $T_{n-1}(p)$ together with the identity element with multiplicity 4. Therefore, the bound goes through as before. Choose a nontrivial character $\rho\colon H \to \mathbf{C}$ and view ρ as (ρ_1, ρ_2) with ρ_1 being viewed as a map

$$\rho_1\colon \{(n-1) \times 1 \text{ column vectors over } \mathbf{Z}/p\mathbf{Z}\} \to \mathbf{C},$$

and $\rho_2\colon (\mathbf{Z}/p\mathbf{Z})^* \to \mathbf{C}$. If ρ_2 is nontrivial, then we get nontrivial absorption at every vertex from the generators $D_n(\alpha)$ and $D_n(\alpha^{-1})$. Thus we need only consider the case where ρ_2 is trivial. In this case we proceed exactly as for $U_n(p)$. Call a matrix $A \in T_{n-1}(p)$ of type k if k is the least positive integer with $\rho_1((A^{-1})_{\cdot n-k}) \neq 1$. Then a vertex A has nonzero absorption weight and of at least $2(1 - \cos(2\pi/p))$ if and only if it is of type 1. Remove the p-cycles inductively, pushing the absorption to higher types using Lemma 15 exactly as before. Then the bounds from before still apply but with different final constants since we have more generators. □

7. Concluding remarks. It is worth noting that there is something very differential geometric/topological about this method. In fact, geometric ideas motivated several of the steps in this paper and may assist in generalizing these results to non-Abelian H. The subgroup H acts on the Cayley graph Γ' of G and the map $\Gamma' \to \Gamma$ is a covering map. We may view Γ' as sitting over Γ with each point having preimage a copy of H. We want to exploit the way these copies of H twist as we go around loops to bound the eigenvalues. To detect this twist, we choose a representation $\rho\colon H \to S^1 \subset \mathbf{C}$ of H. With this representation fixed we get a complex line bundle E [or alternately a $U(1)$-bundle] over Γ. That is, over every point of Γ we have a copy of $\mathbf{C}$. The choice of coset representatives is a choice of local trivialization of this $U(1)$-bundle, that is, a concrete identification of this preimage with $\mathbf{C}$. The phases on the edges are then the gluing data describing the bundle. These data measure the twisting of the bundle. For example, this picture motivated the discussion that for a twisted graph with underlying graph a cycle only the product of all the phases as one went around mattered. In topological terminology, since E is induced by the bundle Γ', it is a flat bundle. Therefore, one gets a monodromy representation $\pi_1(\Gamma) \to U(1)$ as follows. If α is a cycle in Γ, that is, a product of elements of S representing an element of H, then it has a natural lift to a path in Γ', hence to a path in E. Since the final endpoint of this lifted path projects to the same element of Γ as the initial endpoint, the difference between the initial and final values is an element $e^{i\theta} \in U(1)$. This element is an image of α under the map $\pi_1(\Gamma) \to U(1)$. Thus the twist around cycles we used to bound eigenvalues is also the twist of a certain natural $U(1)$ bundle over Γ.

The final step which is not as well understood is to turn this twist into a global eigenvalue bound. There seems to be no way to do this other than the somewhat ad hoc bounds given above. One would like to be able to say that there is some global "characteristic class" that describes the twisting and that the nontriviality of this global number gives a global bound on the absorption and hence a bound on β_1. This does not seem to be the case.

REFERENCES

[1] Diaconis, P. (1988). *Group Representations in Probability and Statistics*. IMS, Hayward, CA.

[2] Diaconis, P. and Saloff-Coste, L. (1993). An application of Harnack inequalities to random walk on nilpotent quotients. Technical Report 434, Dept. Statist., Stanford Univ.

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

[4] DIACONIS, P. and STROOCK, D. (1991). Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.* **1** 36–61.

[5] ELLENBERG, J. (1993). A sharp diameter bound for upper triangular matrices. Senior honors thesis, Dept. Mathematics, Harvard Univ.

[6] JERRUM, M. and SINCLAIR, A. (1989). Approximating the permanent. *SIAM J. Comput.* **18** 1147–1178.

[7] JOHNSON, D. L. (1988). The groups of formal power series under substitutions. *J. Austral. Math. Soc.* **45** 296–302.

[8] STONG, R. (1995). Random walks on the groups of upper triangular matrices. *Ann. Probab.* **23** 1939–1949.

[9] STONG, R. (1995). Eigenvalues of the natural random walk on the Burnside group $B(3, n)$. *Ann. Probab.* **23** 1950–1960.

DEPARTMENT OF MATHEMATICS
RICE UNIVERSITY
P.O. BOX 1892
HOUSTON, TEXAS 77251-1892