SYMMETRIC RANDOM WALKS ON GROUPS[1]

BY
HARRY KESTEN

1. **Introduction.** Let G be a countable group and let $A=\{a_1, a_1, \cdots\}$ ($a_i \in G$) generate G. Consider the random walk on G in which every step consists of right multiplication by a_i or its inverse a_i^{-1}, each with probability p_i ($p_i \geqq 0$, $2\sum_i p_i = 1$). This does not mean that p_i is the *total* probability of multiplying by any element which equals a_i in G. It may be, for instance, that $a_i = a_j$ with $j \neq i$ (or $a_i = a_i^{-1}$). In this case the total probability of multiplying by a_i is at least $p_i + p_j$ (resp. $2p_i$). We say that $P=\{p_1, p_2, \cdots\}$ is a probability distribution on the set of generators A. This random walk defines a Markov chain whose possible states are the elements of G. The transition probability from g_1 to g_2 ($g_1 \in G$, $g_2 \in G$) is given by the probability that g_2 is reached in one step from g_1. Since G is countable we can number the possible states $1, 2, \cdots$ and represent the Markov chain by its matrix of transition-probabilities, $M(G, A, P)$, say (cf. [1] for terminology).

Several connections are derived between the spectrum of the matrix $M(G, A, P)$ and the structure of the group G. Some results deal with conditions on the spectrum to contain the value 1. Theorem 3 gives an interesting characterization of finitely generated free groups in terms of the upper bound of the spectrum of $M(G, A, P)$.

Since at every step the probability of right multiplication by a_i equals the probability of right multiplication by a_i^{-1}, the transition probabilities from g_1 to g_2 and from g_2 to g_1 are equal and $M(G, A, P)$ is symmetric. Furthermore, the entries of M are nonnegative and in every row the sum of all entries is 1. Denote the dimension of M by $r(M)$ (or r when no confusion is possible); r is a positive integer or $+\infty$. M represents a linear operator on the r dimensional Hilbert space H of vectors $y=\{y_1, y_2, \cdots\}$ (y_i complex numbers) with $\|y\| = (\sum_{i=1}^{r} |y_i|^2)^{1/2} < \infty$. As usual we define the norm of an r-dimensional matrix X (with complex valued entries) by

$$\text{(1.1)} \qquad \operatorname{norm}(X) = \sup_{\|y\|=1;\, y\in H} \|Xy'\|$$

(y' is the transposed vector of y). For hermitian matrices $X=\|x_{ij}\|$

$$\text{(1.2)} \qquad \operatorname{norm}(X) \leqq \sup_i \sum_j |x_{ij}| \qquad [6].$$

The *spectrum* of X is the set of all complex numbers such that $X-\lambda I$ does

Presented to the Society August 30, 1957; received by the editors March 20, 1958.

(1) This paper was submitted as a doctoral dissertation to the Graduate School of Cornell University.

not have an inverse with finite norm, where I is the identity matrix of the same dimension as X. The spectral radius of X is defined as

$$\sup_{\lambda \in \text{spectrum of } X} |\lambda| .$$

The spectrum is always a compact set [2, p. 52]. Since M is symmetric its spectrum is real and one can put

$$\lambda(G, A, P) = \max_{\lambda \in \text{spectrum of } M(G,A,P)} \lambda. \tag{1.3}$$

We shall first give some formulae and analytical properties of $\lambda(G, A, P)$ and then connect it with the structure of G.

2. **Analytic properties of** $\lambda(G, A, P)$. Because M is hermitian with norm $(M) \leqq 1$ (by (1.2)), one can introduce the spectral measure or spectral matrix of M in the usual way (cf. [2; 5 and 7]). There exists therefore a matrix $\sigma(\mu) = \|\sigma_{ij}(\mu)\|$ of functions of the real variable μ such that $\sigma_{ij}(\mu)$ is continuous from the right and the total variation of $\sigma_{ij}(\mu)$ on $(-\infty, +\infty)$ is at most equal to 1 for all i and j. Furthermore the spectrum of M is the set of all real values λ where at least one of the functions $\sigma_{ij}(\mu)$ is not constant, that is

$$\begin{aligned} \lambda \in \text{spectrum of } M \rightleftarrows \ & \text{there does not exist an } \epsilon > 0 \\ & \text{such that } \sigma_{ij}(\mu) \text{ is constant on } [\lambda - \epsilon, \lambda + \epsilon]\text{(2) for all } i \text{ and } j. \end{aligned} \tag{2.1}$$

The operator $\sigma(\mu_1) - \sigma(\mu_2)$ $(\mu_2 \leqq \mu_1)$ is a projection on a subspace of H and thus the matrix $\|\sigma_{ij}(\mu_1) - \sigma_{ij}(\mu_2)\|$ is hermitian and idempotent. From this one easily concludes that the spectrum of M is already determined by the diagonal elements of $\|\sigma_{ij}(\mu)\|$, i.e.

Lemma 2.1. *$\lambda \in$ spectrum of $M \rightleftarrows$ there does not exist an $\epsilon > 0$ such that $\sigma_{ii}(\mu)$ is constant on $[\lambda - \epsilon, \lambda + \epsilon]$ for all i.*

Proof. The sufficiency of the condition follows immediately from (2.1). For the necessity it suffices (again by (2.1)) to prove that $\sigma_{ij}(\mu_1) - \sigma_{ij}(\mu_2) \neq 0$ implies $\sigma_{ii}(\mu_1) - \sigma_{ii}(\mu_2) \neq 0$. But since $\|\sigma_{ij}(\mu_2) - \sigma_{ij}(\mu_2)\|$ is hermitian and idempotent if $\mu_2 \leqq \mu_1$, one has

$$\begin{aligned} \sigma_{ii}(\mu_1) - \sigma_{ii}(\mu_2) &= i, i \text{ entry of } [\sigma(\mu_1) - \sigma(\mu_2)]^2 \\ &= \sum_k [\sigma_{ik}(\mu_1) - \sigma_{ik}(\mu_2)][\sigma_{ki}(\mu_1) - \sigma_{ki}(\mu_2)] \\ &= \sum_k [\sigma_{ik}(\mu_1) - \sigma_{ik}(\mu_2)][\bar{\sigma}_{ik}(\mu_1) - \bar{\sigma}_{ik}(\mu_2)] \\ &\geqq | \sigma_{ij}(\mu_1) - \sigma_{ij}(\mu_2) |^2 \end{aligned} \tag{2.2}$$

and the lemma follows. $\sigma(\mu) = \|\sigma_{ij}(\mu)\|$ is called the *spectral matrix* of M. It is well known that

(2) Square brackets denote closed intervals, round ones denote open intervals.

$$m_{ii}^{(n)} = \int_{-\infty}^{+\infty} \mu^n d\sigma_{ii}(\mu) \tag{2.3}$$

where

$$M^n = \left\| m_{ij}^{(n)} \right\| = \text{the } n\text{th power of the matrix } M. \tag{2.4}$$

The integral in (2.3) is a Lebesgue-Stieltjes integral and by Lemma 2.1 may be written as

$$m_{ii}^{(n)} = \int_{\lambda'-0}^{\lambda+0} \mu^n d\sigma_{ii}(\mu) \tag{2.5}$$

where λ' and λ are the lower and upper bound of the spectrum of M respectively. Since in our Markov chain

(2·6)
$m_{ii}^{(n)}$ = Prob. of returning to state i at the nth step, given that one starts in state i
= Prob. of returning to the group identity at the nth step, given that one starts at the group identity,

$m_{ii}^{(n)}$ is independent of i. Consequently, $\sigma_{ii}(\mu)$ is independent of i [7, pp. 179, 97] and we shall write

$$\sigma_0(\mu) = \sigma_{ii}(\mu) \qquad \text{for all } i. \tag{2.7}$$

$\sigma_0(\mu)$ is a real and nondecreasing function of μ (cf. (2.2)) and from Lemma 2.1 it follows that for every $\epsilon > 0$

$$\sigma_0(\lambda + \epsilon) - \sigma_0(\lambda - \epsilon) > 0 \quad \text{and} \quad \sigma_0(\lambda' + \epsilon) - \sigma_0(\lambda' - \epsilon) > 0. \tag{2.8}$$

It then follows that

$$\max(|\lambda'|, |\lambda|) = \limsup_n \left[m_{ii}^{(n)}\right]^{1/n} = [\text{radius of convergence of } m(x)]^{-1} \tag{2.9}$$

where

$$m(x) = \sum_{n=0}^{\infty} m_{ii}^{(n)} x^n \quad \text{with} \quad m_{ii}^{(0)} = 1. \tag{2.10}$$

From (2.5) and the fact that $m_{ii}^{(n)} \geqq 0$ for all n, it follows that

$$\lambda = \max(|\lambda'|, |\lambda|). \tag{2.11}$$

LEMMA 2.2.

$$\begin{aligned} \lambda(G, A, P) &= \limsup \left[m_{ii}^{(n)}\right]^{1/n} \\ &= [\textit{radius of convergence of } m(x)]^{-1} = \sup_{\|y\|=1; y \in H} yM(G, A, P)\bar{y}' \\ &= \textit{norm } M(G, A, P). \end{aligned}$$

Proof. The first two equalities follow from (2.9) and (2.11). The other equalities are well known [2, pp. 41 and 55].

Combining this lemma with (1.2) we get immediately

$$\lambda(G, A, P) \leqq 1. \tag{2.12}$$

LEMMA 2.3. *If A and B are two hermitian matrices of bounded norm and the same dimension, then*

$$\lambda_\xi = \textit{spectral radius of } \xi A + (1-\xi)B$$

is a convex function of ξ [4].

Proof. Using the triangle inequality for the norm one has

$$\begin{aligned}\lambda_{\eta\xi_1+(1-\eta)\xi_2} &= \text{norm } (\eta\xi_1 A + \eta(1-\xi_1)B + (1-\eta)\xi_2 A + (1-\eta)(1-\xi_2)B) \\ &\leqq |\eta| \text{ norm } (\xi_1 A + (1-\xi_1)B) \\ &\quad + |1-\eta| \text{ norm } (\xi_2 A + (1-\xi_2)B) = |\eta|\lambda_{\xi_1} + |1-\eta|\lambda_{\xi_2}.\end{aligned}$$

If one takes $0 \leqq \eta \leqq 1$ the lemma follows.

3. **Connections between $\lambda(G, A, P)$ and the structure of G.** Unless otherwise stated we assume that G is countable and generated by $A = \{a_1, a_2, \cdots\}$; if N is a normal subgroup of G then G/N can be generated by the cosets $a_1N, a_2N, \cdots$. We usually say, a bit loosely, that G/N is also generated by A. For $P = \{p_1, p_2, \cdots\}$, we put

(3.1) $P\{a_i | P\} = p_i =$ probability of right multiplying with a_i according to P, at any given step.

We only consider symmetric random walks, that is, the probabilities of right multiplication with a_i and a_i^{-1} are always equal. e denotes the identity element of G.

LEMMA 3.1. *Let N be a normal subgroup of G and consider G as well as G/N as generated by A. Then $\lambda(G, A, P) \leqq \lambda(G/N, A, P)$.*

Proof.(³)

$$\begin{aligned}\lambda(G) &= \limsup_n [\text{Probability of returning to } e \text{ at the } n\text{th step given that one starts at } e]^{1/n} \\ &\leqq \limsup_n [\text{Probability of reaching some element of } N \text{ at the } n\text{th step, given that one starts at } e]^{1/n} = \lambda(G/N).\end{aligned} \tag{3.2}$$

The probabilities above are the probabilities corresponding to the random-

(³) We drop some of the arguments in $M(G, A, P)$ and $\lambda(G, A, P)$ if no confusion is to be expected.

walk on G, defined by A and P, and (3.2) is an immediate application of Lemma 2.2.

If one considers the group G as defined by a set of relations between the elements of A (cf. [3, vol. I, p. 129]) then Lemma 3.1 may be expressed as "The introduction of new relations does not decrease $\lambda(G)$."

LEMMA 3.2. *Put $A'=\{e, a_1, a_2, \cdots\}$ and let for $0\leqq\xi\leqq 1$ $P'(\xi)$ be the probability distribution on A' defined by*

$$\begin{aligned} P\{e \mid P'\} &= \frac{\xi}{2}, \\ P\{a_i \mid P'\} &= (1-\xi)P\{a_i \mid P\} = (1-\xi)p_i. \end{aligned} \tag{3.3}$$

Then

$$\lambda(G, A', P') = \xi + (1-\xi)\lambda(G, A, P). \tag{3.4}$$

Proof. Since we multiply according to P' by e or $e^{-1}=e$ each with probability $\xi/2$(4) at every step, and with a_i or a_i^{-1} with probability $(1-\xi)p_i = (1-\xi)P\{a_i|P\}$, one has $M(G, A', P')=\xi I+(1-\xi)M(G, A, P)$ where I is the identity matrix with the same dimension as $M(G, A, P)$.

LEMMA 3.3. *Let H be generated by the infinite set $B=\{b_1, b_2, \cdots\}$ and let $Q=\{q_1, q_2, \cdots\}$ be a probability distribution on B. Then for every $\epsilon>0$ there exists a finite k such that*

$$|\lambda(H, B, Q) - \lambda(H, B, Q_k)| \leqq \epsilon$$

where Q_k is defined by

$$P\{b_i \mid Q_k\} = q_i \Big/ 2\sum_{i=1}^{k} q_i \qquad \text{for } 1 \leqq i \leqq k, \tag{3.5a}$$

$$P\{b_i \mid Q_k\} = 0 \qquad \text{for } i > k. \tag{3.5b}$$

Proof. Put $M_k=M(H, B, Q)-M(H, B, Q_k)$. For every row, the sum of the absolute values of the entries of M_k in that row tends to zero if $k\to\infty$. In fact for every row this sum is

$$\sum_{i=1}^{k} 2q_i\left[1 \Big/ 2\sum_{i=1}^{k} q_i - 1\right] + \sum_{i=k+1}^{\infty} 2q_i.$$

The lemma follows now from (1.2) and Lemma 2.2. Note that $\lambda(H, B, Q_k) =\lambda(H_k, B_k, Q_k)$ where $B_k=\{b_1, b_2, \cdots, b_k\}$, H_k the subgroup of H generated by B_k and Q_k is defined by (3.5a).

Let G again be generated by A and let H be a subgroup of G generated

(4) We make the artificial distinction between e and e^{-1} to keep (3.3) in agreement with our conventions (cf. §1).

by $B=\{b_1, b_2, \cdots\}$ $(b_i \in G)$. Denote the smallest normal subgroup of G, containing H by N and G/N by K. We consider K also as generated by A. One has the following

THEOREM 1. *If $P=\{p_1, p_2, \cdots\}$ and $Q=\{q_1, q_2, \cdots\}$ are probability distributions on $A=\{a_1, a_2, \cdots\}$ and $B=\{b_1, b_2, \cdots\}$ respectively and P assigns positive probability to every element of A, i.e.*

$$p_i > 0 \tag{3.6}$$

and

$$\lambda(H, B, Q) < 1, \tag{3.7}$$

then

$$\lambda(K, A, P) > \lambda(G, A, P). \tag{3.8}$$

Proof. We may restrict ourselves to the case where B is a finite set, say $B=\{b_1, b_2, \cdots, b_k\}$. For, if B is infinite we can replace Q by a probability distribution Q_k such that $\lambda(H_k, B_k, Q_k)<1$ (cf. Lemma 3.3). If N_k is the smallest normal subgroup of G containing H_k, then $N_k \subseteq N$ and thus $G/N \cong G/N_k/N/N_k$ and so by Lemma 3.1 it would suffice to show

$$\lambda(G/N_k, A, P) > (G, A, P).$$

Assume therefore that B has k $(<\infty)$ elements. Choose now a fixed ξ $(0<\xi<1)$, put $A'=\{e, a_1, a_2, \cdots\}$ and define $P'=P'(\xi)$ as in Lemma 3.2. By Lemma 3.2, (3.8) is equivalent to

$$\lambda(K, A', P') > \lambda(G, A', P'). \tag{3.9}$$

Every b_j can be written as a product of the form $a_{i_1}^{\epsilon_1} \cdots a_{i_t}^{\epsilon_t}$ $(\epsilon_j=+1$ or $-1)$. Fix one such representation for every b_j, say w_j $(1 \leqq j \leqq k)$. Thus w_j stands for only one product of elements of A and their inverses and does *not* denote any word equal to b_j in G. Let w_j be the product of l_j elements of A or their inverses and put

$$l = \max_{1 \leqq j \leqq k} l_j, \tag{3.10}$$

$$\overline{M}(G) = [M(G, A', P')]^l, \tag{3.11}$$

$$\overline{M}(K) = [M(K, A', P')]^l. \tag{3.12}$$

Introduce the set C which generates G as well as K.

$$C = \left\{ e^m a_{i_1}^{\epsilon_1} \cdots a_{i_{l-m}}^{\epsilon_{l-m}} \,\middle|\, 0 \leqq m \leqq l,\ a_{i_j} \in A,\ \epsilon_j = +1 \text{ or } -1 \right\}. \tag{3.13}$$

Define the probability distribution R on C by[5]

[5] In accordance with footnote 4 we have to take $P\{e^l | R\}=\xi^l/2$.

$$(3.14)\qquad P\{e^{m}a_{i_l}^{\epsilon_l}\cdots a_{i_{l-m}}^{\epsilon_{l-m}}\mid R\} = \binom{l}{m}\xi^m \prod_{j=1}^{l-m} p'_{i_j}.$$

Since l consecutive steps in the randomwalk on G defined by P' on A' amount to right multiplication by an element of C or its inverse, with the probability assigned to it by R, one has

$$(3.15)\qquad \overline{M}(G) = M(G, C, R);\qquad \overline{M}(K) = M(K, C, R).$$

Therefore, if

$$(3.16)\qquad \bar{\lambda}(G) = \lambda(G, C, R);\qquad \bar{\lambda}(K) = \lambda(K, C, R)$$

it follows from Lemma 2.2 that

$$\bar{\lambda}(G) = [\lambda(G, A', P')]^l \quad\text{and}\quad \bar{\lambda}(K) = [\lambda(K, A', P')]^l.$$

Consequently it suffices to show

$$(3.17)\qquad \bar{\lambda}(K) > \bar{\lambda}(G).$$

The set $C' = \{e^{l-l_j}w_j \mid 1 \leqq j \leqq k\}$ is contained in C and so we can define a probability distribution S on C by

$$(3.18)\qquad \begin{aligned} &P\{e^{l-l_j}w_j \mid S\} = P\{b_j \mid Q\} = q_j, && 1 \leqq j \leqq k,\\ &P\{e^{m}a_{i_1}^{\epsilon_l}\cdots a_{i_{l-m}}^{\epsilon_{l-m}} \mid S\} = 0 \text{ for all elements } e^{m}a_{i_1}^{\epsilon_1}\cdots a_{i_{l-m}}^{\epsilon_{l-m}} \in C - C'. \end{aligned}$$

By (3.6) $P\{c \mid R\} > 0$ for every $c \in C$ so that an $\alpha > 0$ exists such that

$$(3.19)\qquad (1+\alpha)P\{c \mid R\} - \alpha P\{c \mid S\} \geqq 0 \text{ for every } c \in C.$$

Fix $\alpha > 0$ such that (3.19) is satisfied and define for $0 \leqq \eta \leqq 1$ the probability distributions $T(\eta)$ on C by

$$(3.20)\qquad \begin{aligned} P\{c \mid T(\eta)\} &= (1-\eta)[(1+\alpha)P\{c \mid R\} - \alpha P\{c \mid S\}]\\ &\quad + \eta P\{c \mid S\} = (1-\eta)(1+\alpha)P\{c \mid R\}\\ &\quad + (\eta - \alpha(1-\eta))P\{c \mid S\} \text{ for every } c \in C. \end{aligned}$$

$T(1)$ equals S, so that the random walk defined by $T(1)$ and C on G is the same as the one defined by S and C, which in turn is the random walk defined by Q and B on H. Thus

$$(3.21)\qquad \lambda(G, C, S) = \lambda(H, B, Q) < 1.$$

Since $H \subseteq N$, multiplication by an element of H amounts in $K = G/N$ to multiplication by the identity. Therefore (using Lemma 2.2)

$$(3.22)\qquad \lambda(K, C, T(1)) = \lambda(K, C, S) = 1,$$

and (using Lemma 3.2)

$$(3.23)\qquad \lambda(K, C, T(\eta)) = \eta + (1-\eta)\lambda(K, C, T(0)).$$

But by Lemmas 2.3, 3.1, and (3.21) till (3.23) for $\eta>0$

$$\begin{aligned}\lambda(G, C, T(\eta)) &\leqq (1-\eta)\lambda(G, C, T(0)) + \eta\lambda(G, C, T(1))\\ &= (1-\eta)\lambda(G, C, T(0)) + \eta\lambda(G, C, S)\\ &< (1-\eta)\lambda(K, C, T(0)) + \eta = \lambda(K, C, T(\eta)).\end{aligned} \tag{3.24}$$

In particular, for $\eta=\alpha/(1+\alpha)$, $T(\eta)$ equals R so that

$$\bar{\lambda}(G) = \lambda(G, C, R) < \lambda(K, C, R) = \bar{\lambda}(K). \tag{3.25}$$

In fact

$$\begin{aligned}\lambda(K, C, R) - \lambda(G, C, R) &= \lambda\left(K, C, T\left(\frac{\alpha}{1+\alpha}\right)\right) - \lambda\left(G, C, T\left(\frac{\alpha}{1+\alpha}\right)\right)\\ &\geqq \frac{\alpha}{1+\alpha}\left[1-\lambda(G, C, S)\right]\\ &= \frac{\alpha}{1+\alpha}\left[1-\lambda(H, B, Q)\right].\end{aligned} \tag{3.26}$$

The theorem now follows.

Theorem 1 has a number of corollaries. Let us call probability distributions which assign a positive probability to every element of A, *strictly positive probability distributions.*

Corollary 1. *If the spectrum corresponding to the random walk on G, defined by a strictly positive probability distribution P on A contains the value* 1, *then the spectrum corresponding to a random walk on any subgroup $H \subseteq G$, defined by any probability distribution Q on any set B of generators of H contains* 1.

Proof. Using the notation of Theorem 1 the corollary says

$$\lambda(G, A, P) = 1 \quad \text{implies} \quad \lambda(H, B, Q) = 1$$

if P is a strictly positive probability distribution on A. But by Theorem 1 and (2.12) $\lambda(H, B, Q)<1$ would imply

$$\lambda(G, A, P) < \lambda(K, A, P) \leqq 1$$

which contradicts the assumptions.

In particular with $H=G$, we get

$$\lambda(G, A, P) = 1 \quad \text{implies} \quad \lambda(G, B, Q) = 1 \tag{3.27}$$

for a strictly positive probability distribution P and any B and Q. In view of this we shall often write $\lambda(G)=1$ or $\lambda(G)<1$ without further specification of the set of generators and the probability distribution. We should keep in mind though, that (3.27) is only valid if P is strictly positive. $\lambda(G)=1$ ($\lambda(G)<1$) means therefore: There exists (does not exist) a strictly positive probability distribution P on A such that $\lambda(G, A, P)=1$. By the above, *whether or not* $\lambda(G)=1$, *is solely determined by the structure of* G and it would be interesting to characterize all groups G for which $\lambda(G)=1$. Only some partial results in this direction are obtained here (Corollaries 3 and 4, and Theorem 5).

COROLLARY 2. *Let N be a normal subgroup of G and consider $K=G/N$ as generated by A. If $\lambda(N)=1$, then $\lambda(K, A, P)=\lambda(G, A, P)$.*

Proof. By Lemma 3.1

$$\lambda(G, A, P) \leqq \lambda(K, A, P). \tag{3.28}$$

On the other hand, it follows from (2.5), (2.8), and (2.11) that for every $\epsilon>0$ and sufficiently large n

(3.29) i, i entry of $[M(K, A, P)]^{2n}$ = Probability of returning to the identity at the $2n$th step, given that one starts at the identity, in the random walk on K defined by A and P = Probability of reaching some element of N at the $2n$th step, given that one starts at e, in the random walk on G defined by A and $P \geqq [(1-\epsilon)\lambda(K, A, P)]^{2n}$.

Given an $\epsilon>0$ choose n such that (3.29) is satisfied and put for $b\in N$.

(3.30) $2p_n(b)$ = Conditional probability of reaching b or b^{-1} at the $2n$th step, given that one starts at e and that one reached some element of N at the $2n$th step, in the random ralk on G, defined by A and P

$$= \frac{\text{Prob. of going from } e \text{ to } b \text{ or } b^{-1} \text{ in } 2n \text{ steps in the random walk on } G\text{, defined by } A \text{ and } P}{\text{Prob. of going from } e \text{ to some element of } N \text{ in } 2n \text{ steps in the random walk on } G\text{, defined by } A \text{ and } P}.$$

If $b\in N$ then also $b^{-1}\in N$. Select from every pair (b, b^{-1}) *one* element (if $b=b^{-1}$ we take that one element). Let B be the set of selected elements. Then

$$P\{b \mid P_n\} = p_n(b) \qquad \text{for } b \in B \tag{3.31}$$

defines a probability distribution P_n on B. Since $\lambda(N)=1$, one has by Corollary 1, $\lambda(N, B, P_n)=1$ and for sufficiently large m

(3.32) Probability of returning to e at the $2m$th step, given that one starts in e in the random walk on N, defined by B and $P_n \geqq (1-\epsilon)^{2m}$.

It is clear that

Probability of returning to e at the $(2n\cdot 2m)$th step, given that one starts in e in the random walk on G defined by A and $P \geqq$ [Probability of reaching some element of N at the $2n$th step, given that one starts at e in the random walk on G defined by A and $P]^{2m}$. Probability of returning to e at the $2m$th step given that one starts at e in the random walk on N defined by B and P_n.

Consequently by (3.29) and (3.32)

$$m_{ii}^{(2n\cdot 2m)}(G, A, P) \geqq [(1-\epsilon)\lambda(K, A, P)]^{2n\cdot 2m}(1-\epsilon)^{2m}$$

and

$$\lambda(G, A, P) \geqq (1-\epsilon)^2\lambda(K, A, P). \tag{3.33}$$

Since (3.33) is valid for every $\epsilon > 0$, the corollary follows.

Let $L \subseteq G$ be a subgroup of G, generated by C and Q a probability distribution on C. As in Corollary 2 let N be a normal subgroup of G with $\lambda(N)=1$. Considering $L/L\cap N \cong LN/N$ also as generated by C, Corollary 2 implies $\lambda(L/L\cap N, C, Q)=\lambda(LN/N, C, Q)=\lambda(L, C, Q)$ since by Corollary 1, also $\lambda(L\cap N)=1$. If $K=G/N$, then we have from Theorem 1, Corollary 2 and the above

THEOREM 2. *For a strictly positive probability distribution P on A*

$$\lambda(K, A, P) > \lambda(G, A, P)$$

if and only if $\lambda(N)<1$[6]. *If L is a subgroup of G, generated by C and Q a probability distribution on C, then*

$$\lambda(L/L\cap N, C, Q) > \lambda(L, C, Q) \quad \textit{implies } \lambda(K, A, P) > \lambda(G, A, P).$$

Theorem 2 provides us with a necessary and sufficient condition for the upper bound of the spectrum to increase upon the introduction of new relations in the group (cf. remark after Lemma 3.1). This apparently depends on the structure of N.

Corollary 2 can be slightly generalized, e.g. "If G has a finite normal series

$$G = G_0 \supset G_1 \supset \cdots \supset G_k = (e)$$

with $\lambda(G_i/G_{i+1})=1$ $(0 \leqq i \leqq k-1)$, then $\lambda(G)=1$." This follows by induction, as $\lambda(G_{i-1})=1$ implies $\lambda(G_i)=1$ by Corollary 2. Similarly,

"If $N_1, N_2, \cdots, N_k$ are a finite number of normal subgroups of G with $\lambda(G/N_i)=1$ $(1 \leqq i \leqq k)$ then

(6) N is here considered as a subgroup of G. Whether $\lambda(N)<1$ or not is therefore determined by the relations valid in G.

$$\lambda(G/N) = 1, \quad \text{where} \quad N = \bigcap_{i=1}^{k} N_i."$$

This follows also by induction, once it is shown for $k=2$. But for $k=2$ $G/N_1 \cong G/N_1 \cap N_2/N_1/N_1 \cap N_2$ and $N_1/N_1 \cap N_2 \cong N_1N_2/N_2 \subseteq G/N_2$ so that $\lambda(N_1/N_1 \cap N_2)=1$.

Combinatorially the last statement can be formulated as

$\lim_{n\to\infty}$ [Probability of reaching some element of N_j at the $2n$th step, given that one starts in $e]^{1/2n}=1$

for $1 \leqq j \leqq k$ implies

$\lim_{n\to\infty}$ [Probability of reaching some element of N at the $2n$th step, given that one starts in $e]^{1/2n}=1$.

By induction one also proves:

"If $N_1, \cdots, N_k$ are normal subgroups of G and N is the smallest normal subgroup of G containing $N_1, \cdots, N_k$ then $\lambda(N_1)= \cdots =\lambda(N_k)=1$ implies $\lambda(N)=1$."

We only have to prove it for $k=2$. But then $N=N_1N_2$ and

$$\begin{aligned}\lambda(N_1N_2) &= \lambda(N_1N_2/N_2) \\ &= \lambda(N_1/N_1 \cap N_2) = 1.\end{aligned}$$

4. Computation of $\lambda(G)$ for some examples.

LEMMA 4.1. *Let G be a direct product $G_1 \otimes \cdots \otimes G_k$ (k finite). Suppose G_i is generated by $A_i = \{a_{i1}, a_{i2}, \cdots\}$. Define P on $A = \bigcup_{i=1}^{k} A_i$ by*

$$P\{a_{ij} \mid P\} = p_{ij} \left(2 \sum_{i=1}^{k} \sum_{j} p_{ij} = 1\right) \tag{4.1}$$

and P_i on A_i by

$$P\{a_{ij} \mid P_i\} = p_{ij}/p_i \tag{4.2}$$

where $p_i = 2\sum_j p_{ij}$. Then

$$\lambda(G, A, P) = \sum_{i=1}^{k} 2p_i\lambda(G_i, A_i, P_i). \tag{4.3}$$

Proof. Denote the diagonal element of the spectral matrices of $M(G_i, A_i, P_i)$ and $M(G, A, P)$ by $\sigma_i(\mu)$ and $\sigma(\mu)$ respectively. Since $\prod_{i=1}^{k} g_i = e$ ($g_i \in G$) if and only if g_i equals the identity for every i and since $g_i \in G_i$ and $g_j \in G_j$ commute for $i \neq j$, one has, using (2.5) for every G_i

Probability of returning to e at the nth step given that one starts at e, in the random walk on G, defined by A and P

$$(4.4)\quad = \sum_{x \geq 0; x_1 + \cdots + x_k = n} \frac{n!}{x_1! \cdots x_k!} (2p_1)^{x_1} \cdots (2p_k)^{x_k} \int \mu_1^{x_1} d\sigma_1(\mu_1) \cdots \int \mu_k^{x_k} d\sigma_k(\mu_k)$$

$$= \int \cdots \int (2p_1\mu_1 + \cdots + 2p_k\mu_k)^n d\sigma_1(\mu_1) \cdots d\sigma_k(\mu_k).$$

Clearly the lim sup of the $(1/n)$th power of the above probability is $2p_1\lambda(G_1, A_1, P_1) + \cdots + 2p_k\lambda(G_k, A_k, P_k)$ and the lemma follows.

LEMMA 4.2. *If $A = \{a\}$ and P assigns probability $1/2$ to a, then $\lambda(G, A, P) = 1$.*

Proof. Probability of returning to e at the $2n$th step given that one starts at e, in the random walk on G defined by A and $P \geqq$ Probability of multiplying n times by a and n times by a^{-1}

$$a^{-1} = C_{2n,n} 4^{-n} \sim \frac{1}{(\pi n)^{1/2}} \cdot$$

Since $\lim_{n \to \infty} (1/(\pi n)^{1/2})^{1/2n} = 1$, the lemma follows.

THEOREM 3. *Let G be generated by $A = \{a_1, \cdots, a_h\}$ with $1 < h < \infty$, and P be defined on A by*

$$(4.5)\qquad P\{a_i \mid P\} = 1/2h \qquad (1 \leqq i \leqq h).$$

Then G is a free group with free generators $a_1, a_2, \cdots, a_h$ if and only if

$$(4.6)\qquad \lambda(G, A, P) = \left(\frac{2h-1}{h^2}\right)^{1/2}.$$

In this case the spectrum of $M(G, A, P)$ is the interval $[-(2h-1/h^2)^{1/2}, +(2h-1/h^2)^{1/2}]$ and

$$(4.7)\qquad \min_Q \lambda(G, A, Q) = \lambda(G, A, P) = \left(\frac{2h-1}{h^2}\right)^{1/2}$$

where Q runs through all probability distributions on A.

Proof. Let us first suppose that G is free and A is a set of free generators for G (cf. [3, vol. I, p. 124 ff.] for terminology). Put

$m^{(n)} =$ Probability of returning to e at the nth step given that one starts at e in the random walk on G defined by P and A.

$r^{(n)}$ = Probability of returning for the first time to e at the nth step given that one starts at e in the random walk on G defined by P and A.

$$m(x) = \sum_{n=0}^{\infty} m^{(n)}x^n \qquad (\text{taking } m^{(0)} = 1), \qquad r(x) = \sum_{n=1}^{\infty} r^{(n)}x^n.$$

Then $m(x)=1/(1-r(x))$ [1, p. 243]. By a word (cf. [3, vol. I, p. 124 ff.] for the terminology) of n letters we mean a product

$$a_{i_1}^{\epsilon_1} \cdots a_{i_n}^{\epsilon_n} (\epsilon_i = +1 \text{ or } -1;\ a_{i_j} \in A)$$

of elements of A or their inverses (the order of the factors is of course important). By a left segment of k letters of $a_{i_1}^{\epsilon_1} \cdots a_{i_n}^{\epsilon_n}$ $(k \leqq n)$, we mean the partial product $a_{i_1}^{\epsilon_1} \cdots a_{i_k}^{\epsilon_k}$. It may happen that in a word $w = a_{i_1}^{\epsilon_1} \cdots a_{i_n}^{\epsilon_n}$ 2 consecutive factors $a_{i_j}^{\epsilon_j} a_{i_{j+1}}^{\epsilon_{j+1}}$ occur with $a_{i_j} = a_{i_{j+1}} \epsilon_j = -\epsilon_{j+1}$. Then we can "reduce" w by cancelling these factors. A word which allows no further reductions is called reduced. Every word w is equal to exactly one reduced word $w' \cdot l(w)$, the length of w is the number of letters in w'. The empty word has length 0 and represents the identity e. A word w equals the identity if and only if, by successive reductions, it can be reduced to the empty word.

$r^{(n)} = (1/2h)^n \times$ number of words of n letters equal to e in the free group G, which have no left segment of less than n letters equal to e.

Obviously $r^{(n)}=0$ for odd n. Every word w of $2n$ letters, equal to e can be mapped on a path in the plane from (0, 0) to (n, n) along the lattice points, and not passing through any point (k, k) with $1 \leqq k \leqq n-1$. The mapping is constructed in the following way: For every letter we record a horizontal or vertical step of length 1. Let w_k denote the left segment of k letters of w. If $l(w_k) = l(w_{k-1})+1$ we record a horizontal step for the kth letter of w. If $l(w_k) = l(w_{k-1})-1$ we record a vertical step. The words of $2n$ letters equal to e with no left segment of less than $2n$ letters equal to e correspond to paths which are on the diagonal (the line through (0, 0) and (n, n) only in (0, 0) and (n, n)). There are $n^{-1}C_{2n-2,n-1}$ such paths [1, p. 246]. How many words are mapped on a fixed path? The first step is horizontal no matter which of the $2h$ possibilities for $a_{i_1}^{\epsilon_1}$ $(i=1, \cdots, h;\ \epsilon_1 = +1$ or $-1)$ is realized. If the kth step is horizontal and does not start on the diagonal it corresponds to $(2h-1)$ possibilities for the kth letter, namely every $a_{i_k}^{\epsilon_k}$ except the inverse of the last letter of the reduced form of the segment of $(k-1)$ letters. A vertical step corresponds to only one possibility, namely the inverse of the last factor in the reduced form of the segment of $(k-1)$ letters. Since every path from (0, 0) to (n, n) has n horizontal and n vertical steps

$$r^{(2n)} = \left(\frac{1}{2h}\right)^{2n} \frac{1}{n} C_{2n-2,n-1} 2h(2h-1)^{n-1}1^n$$

and

$$r(x) = \frac{h - (h^2 - (2h-1)x^2)^{1/2}}{2h-1}, \tag{4.8}$$

$$\begin{aligned} m(x) &= \frac{2h-1}{h-1+(h^2-(2h-1)x^2)^{1/2}} \\ &= \frac{(h^2-(2h-1)x^2)^{1/2}-(h-1)}{1-x^2}. \end{aligned} \tag{4.9}$$

(4.6) follows now from (4.9) and Lemma 2.2. Application of the inversion formula in [7, p. 96] gives for $\sigma_0(\mu)$, the diagonal element of the spectral matrix of $M(G, A, P)$,

$$\sigma_0(\mu) = \begin{cases} 0 \quad \text{for } \mu < -\left(\dfrac{2h-1}{h^2}\right)^{1/2}, \\ \dfrac{1}{\pi}\cdot\displaystyle\int_{-((2h-1))/h^2)^{1/2}}^{\mu} \frac{(2h-1-t^2h^2)^{1/2}}{1-t^2}\,dt \\ \qquad\qquad \text{for } -\left(\dfrac{2h-1}{h^2}\right)^{1/2} \leqq \mu \leqq +\left(\dfrac{2h-1}{h^2}\right)^{1/2}, \\ 1 \quad \text{for } \mu > +\left(\dfrac{2h-1}{h^2}\right)^{1/2}. \end{cases}$$

Therefore the spectrum is the whole interval

$$\left[-\left(\frac{2h-1}{h^2}\right)^{1/2}, \quad +\left(\frac{2h-1}{h^2}\right)^{1/2}\right].$$

Now let the probability distribution $Q(\xi)$ on A be defined by

$$\begin{aligned} P\{a_1 \mid Q(\xi)\} &= \xi q, \\ P\{a_2 \mid Q(\xi)\} &= (1-\xi)q, \\ P\{a_i \mid Q(\xi)\} &= q_i, \qquad 3 \leqq i \leqq h,\ 2q + 2\sum_{i=3}^{h} q_i = 1. \end{aligned} \tag{4.10}$$

According to Lemma 2.3 $\lambda(G, Q, Q(\xi))$ is a convex function of ξ and as long as a_1 and a_2 play the same rôle, $\lambda(G, A, Q(\xi))$ is symmetric around $\xi = 1/2$. A convex function of ξ, symmetric around $\xi = 1/2$, attains its minimum at $\xi = 1/2$. Therefore, if A is a set of free generators for the free group G, the probability distribution P which assigns equal probabilities to every generator minimizes $\lambda(G, A, Q)$ i.e.

$$\min_Q \lambda(G, A, Q) = \lambda(G, A, P) = \left(\frac{2h-1}{h^2}\right)^{1/2}.$$

We still have to prove that $\lambda(G, A, P)=((2h-1)/h^2)^{1/2}$ implies that A is a set of free generators for G. Let H be the free group generated by the set of h free generators $C=\{c_1, \cdots, c_h\}$ and define P' on C by

$$P\{c_i \mid P'\} = P\{a_i \mid P\} = 1/2h. \tag{4.11}$$

$G\cong H/N$ where N is a normal subgroup of H [3, vol. I, p. 128]. The isomorphism maps a_i onto the coset c_iN. If A is not a set of free generators for G, then N contains at least one reduced word $w=c_{i_1}^{\epsilon_1}\cdots c_{i_\rho}^{\epsilon_\rho}$ ($\epsilon_i=+1$ or -1, ρ a positive integer) which is not equal in H to the identity. Moreover we may assume

$$c_{i_1}^{\epsilon_1} \neq c_{i_\rho}^{\epsilon_\rho} \text{ i.e. } i_1 \neq i_\rho \text{ or } i_1 = i_\rho \text{ but } \epsilon_1 \neq -\epsilon_\rho \tag{4.12}$$

for, $w\in N$ implies $c_{i_2}^{\epsilon_2}\cdots c_{i_\rho}^{\epsilon_\rho}c_{i_1}^{\epsilon_1}\in N$ and if $c_{i_1}^{\epsilon_1}=c_{i_\rho}^{-\epsilon_\rho}$ we can replace w by $c_{i_2}^{\epsilon_2}\cdots c_{i_{\rho-1}}^{\epsilon_{\rho-1}}$. Let M be the smallest normal subgroup of H, containing w. Then

$$\lambda(G, A, P) = \lambda(H/N, C, P') \geqq \lambda(H/M, C, P'). \tag{4.13}$$

We can always choose $c_r, c_s\in C$ and ϵ, η (each ± 1), however, so that $c_r^{\epsilon}\neq c_{i_1}^{-\epsilon_1}$, $c_r^{\epsilon}\neq c_{i_\rho}^{\epsilon_\rho}$, $c_r^{\eta}\neq c_{i_1}^{-\epsilon_1}$, $c_s^{\eta}\neq c_{i_\rho}^{\epsilon_\rho}$ $c_r^{\epsilon}\neq c_s^{\eta}$ (all these inequalities are meant in the same sense as (4.12)). That is $w'=c_r^{\epsilon}wc_r^{-\epsilon}$ and $w''=c_s^{\eta}wc_s^{-\eta}$ are also reduced words in M and $c_r^{-\epsilon}c_s^{\eta}$ cannot be reduced. Thus w' and w'' are 2 free generators for a free subgroup $L\subseteq M$ of H and by what we proved already $\lambda(L)<1$. Since $L\subseteq M$, however, $\lambda(L/L\cap M)=1$ and by Theorem 2

$$\lambda(H/M, C, P') > \lambda(H, C, P') = \left(\frac{2h-1}{h^2}\right)^{1/2}. \tag{4.14}$$

From (4.13) and (4.14) it then follows that $\lambda(G, A, P)=((2h-1)/h^2)^{1/2}$ implies that A is a set of free generators for G. This completes the proof of Theorem 3.

Remark. In the case described in the second part of Theorem 3 one can give a lower bound for the increase of the spectral radius. In fact, one obtains readily from (3.26) that

$$\lambda(G, A, P) - \left(\frac{2h-1}{h^2}\right)^{1/2} \geqq \frac{4(1-(3/4)^{1/2})}{(\rho+2)(2h)^{\rho+2}}. \tag{4.15}$$

(ρ has the same meaning as above).

For, $\lambda(G, A, P)-((2h-1)/h^2)^{1/2}\geqq\lambda(H/M, C, P')-((2h-1)/h^2)^{1/2}$. For B in Theorem 1 we now take $\{w', w''\}$; further $k=2$, $w_1=w'$, $w_2=w''$, $l=\rho+2$. We take $P\{w'\mid Q\}=P\{w''\mid Q\}=1/4$, so that $\lambda(L, B, Q)=(3/4)^{1/2}$. Since w' and w'' can be written with the same number of letters ($\rho+2$), we can take

$$\xi = 0 \quad\text{and}\quad \alpha = \frac{4}{(2h)^{\rho+2}-4}.$$

Substitution in (3.26) gives

$$[\lambda(H/M, C, P')]^{\rho+2} - [((2h-1)/h^2)^{1/2}]^{\rho+2} \geqq \frac{4(1-(3/4)^{1/2})}{(2h)^{\rho+2}}$$

whence (4.15).

Theorem 3 states that the introduction of any relation in a free group H on h $(1<h<\infty)$ generators increases the upper bound of the spectrum corresponding to the random walk defined by C and P'. Practically the same proof shows that this is true for any strictly positive probability distribution Q on C, i.e.

$$\lambda(H/N, C, Q) > \lambda(H, C, Q)$$

if N is any normal subgroup of H, which does not consist of the identity only.

COROLLARY 3. *If* $\lambda(G)=1$, *then* G *has no free subgroups on more than one generator.*

Proof. If there is a free subgroup on more than one generator, there exists a free subgroup K on 2 free generators. By Theorem 3, $\lambda(K)<1$, which is impossible by Corollary 1.

THEOREM 4. *If* G *is a finite group or if* G *is a countable abelian group then* $\lambda(G)=1$.

Proof. If G is finite, $M(G, A, P)$ is a finite dimensional matrix and the sum of the entries in a row is 1 for every row. Thus the spectrum of $M(G, A, P)$ contains the value 1. (2.12) then shows $\lambda(G, A, P)=1$. If G is a free abelian group on a finite number of generators, then G is a direct product of a finite number of cyclic groups and $\lambda(G)=1$ as a consequence of Lemmas 4.1 and 4.2. If G is a free abelian group on a countable number of generators we also have to use Lemma 3.3. Since every abelian group is a factor group of a free abelian group [3, vol. I, p. 143], the general case follows by Lemma 3.1.

COROLLARY 4. *If* G *has a finite normal series* $G=G_0\supset G_1\supset \cdots \supset G_k=(e)$ *such that* G_i/G_{i+1} *is a finite group or a countable abelian group* $(0\leqq i\leqq k-1)$ *then* $\lambda(G)=1$.

Proof. Apply the remark after Theorem 2 and Theorem 4.

In particular: "*If* G *is solvable* (cf. [3, vol. II] for definition) *then* $\lambda(G)=1$."

It is possible to give a sufficient condition for $\lambda(G)=1$ in terms of the expected length of the word reached after n steps. Let G again be generated by $A=\{a_1, a_2, \cdots\}$ and $P=\{p_1, p_2, \cdots\}$ a probability distribution on A. If $w=a_{i_1}^{\epsilon_1}\cdots a_{i_n}^{\epsilon_n}$ $(\epsilon_j=\pm 1, a_{i_j}\in A)$ is a word of n letters, we define its length, $l(w)$ say, as the smallest integer k for which there exists a word

$$w' = a_{j_k}^{\gamma_1}\cdots a_{j_k}^{\eta_k} \ (\gamma_i = \pm 1,\ a_{j_i}\in A)$$

of k letters such that $w=w'$. If w equals the identity, then $l(w)=0$. Suppose now that the random walk starts at the identity; after n steps some word w_n of n letters is reached. $l(w_n)$ is a random variable whose expectation we denote by

(4.16) $E_n(G, A, P)=$expected value of $l(w_n)$ in the random walk on G, defined by P on A.

THEOREM 5. *If G is generated by $A=\{a_1, \cdots, a_r\}$, r finite, and $P=\{p_1, \cdots, p_r\}$ is a strictly positive probability distribution on A and*

$$\liminf \frac{E_n(G, A, P)}{n} = 0 \tag{4.17}$$

then

$$\lambda(G, A, P) = 1$$

(*and consequently G has no free subgroups on more than* 1 *generator*).

Proof. Write

$$\epsilon_n = \frac{E_n(G, A, P)}{n} \tag{4.18}$$

and

$$p = \min_{1\leqq i\leqq r} p_i. \tag{4.19}$$

Since $0\leqq l(w_n)/n$ one has

$$\text{Prob.}\left\{\frac{l(w_n)}{n} \leqq 2\epsilon_n\right\} \geqq \frac{1}{2}\,. \tag{4.20}$$

Given that we reached a word of length l at the nth step, there is a probability of at least p^l to return to the identity at the $(n+l)$th step. Therefore, if $m_{ii}^{(k)}$ is the diagonal element of $[M(G, A, P)]^k$, it follows from (2.5) and (4.20) that

$$\frac{1}{2}\,p^{2\epsilon_n n} \leqq \sum_{k=n}^{[n(1+2\epsilon_n)]+1} m_{ii}^{(k)} \leqq \frac{[\lambda(G, A, P)]^n}{1-\lambda(G, A, P)}\,.$$

Since $\liminf \epsilon_n=0$, $\limsup\,[p^{2\epsilon_n n}/2]^{1/n}=1$. This proves that $\lambda(G, A, P)=1$.

It is clear that Theorem 5 is not valid if A is an infinite set, for we can take a group G with $\lambda(G)<1$ and let A be the set of all elements in G (so that $l(w_n)=0$ or 1). Also it is possible to construct a group G with $\lambda(G)=1$ but $\lim\,(E_n(G, A, P))/n>0$, so that (4.17) is not a necessary condition for $\lambda(G, A, P)=1$.

5. **Unsolved problems.** As mentioned in §3, it would be interesting to find all groups with $\lambda(G)=1$. Especially, since for every finite group, the spectrum contains 1. A weak form of the Burnside conjecture would be: "If G is finitely generated and every element has bounded (or more general, finite) order, then $\lambda(G)=1$." This would readily follow if one could prove the converse of Corollary 3, i.e., "If G has no free subgroups on more than 1 generator, then $\lambda(G)=1$." However, the author was unable to prove or disprove this. If this converse of Corollary 3 is not true, however, it might be possible to construct a group G in which every element has finite order but $\lambda(G)<1$. Such a group would disprove the generalized Burnside conjecture. In fact one may try the following. Let G be a free group generated by the free generators $a_1, \cdots, a_h$ and A and P as in Theorem 3. Then $\lambda(G, A, P) = ((2h-1)/h^2)^{1/2} = 1-2\alpha$, say. Order all possible words in G into a sequence, say $w_1, w_2, \cdots$. One can then introduce a relation $w_1^{n_1}=e$ and try to choose n_1 such that $\lambda(G)$ increases by less than $\alpha/2$. If such an n_1 is found one tries to add the relation $w_2^{n_2}=e$ with n_2 such that by this extra relation, the spectral radius increases by less than $\alpha/2^2$ etc. If for every i a proper n_i can be found, then in the group with $w_i^{n_i}=e$ $(i=1, 2, \cdots)$ the spectral radius will still be less than 1.

One can show that n can be chosen such that the relations $a_1^n = \cdots = a_h^n = e$ increase $\lambda(G, A, P)$ arbitrary little, but the author has been unable to do anything more along the above lines.

Even the statement: "If $\lambda(H)=1$ for every *proper* subgroup H of G, then $\lambda(G)=1$," which is still weaker than the converse of Corollary 3 is not yet proved or disproved.

6. **Acknowledgments.** The author wishes to thank Professor M. Kac for suggesting the thesis problem and his valuable help during its investigation. The author also had many helpful discussions with Professor W. Feit. Theorem 5 was "inspired" by a discussion with Professor M. P. Schützenberger. The research done during the summer of 1957 was supported by the Office of the Naval Research.

Note added in proof. Since this paper was submitted, the author proved that $\lambda(G)=1$ is equivalent to the existence of an invariant mean on G (cf. *Full Banach mean values on countable groups*, Math. Scand. vol. 7 (1959)).

It seems that the Burnside conjecture has been disproved recently in Russia.

References

1. W. Feller, *An introduction to probability theory and its applications*, New York, John Wiley and Sons, 1950.

2. Paul R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity*, New York, Chelsea Publishing Company, 1951.

3. A. G. Kurosch, *The theory of groups*, English translation, New York, Chelsea Publishing Company, 1955.

4. P. D. Lax, *The largest eigenvalue as a convex matrix function*, Bull. Amer. Math. Soc. Abstract 63-2-235.

5. B. v. Sz. Nagy, *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes*, Berlin, Springer Verlag, 1942.

6. J. Schur, *Beschränkte Bilinearformen unendlich vieler Veränderlicher*, (Crelle's) J. Reine Angew. Math. vol. 140 (1911) p. 6.

7. A. Wintner, *Spektraltheorie der unendlichen Matrizen*, Leipzig, S. Hirzel Verlag, 1929.

Cornell University,
Ithaca, N. Y.