1. INTRODUCTION, DEFINITIONS, AND STATEMENT OF RESULTS

1.1. Definition of ergodic sequences. Let $\Gamma$ be a countable group, and let $\ell^1(\Gamma) = \{\mu = \sum_{\gamma \in \Gamma} \mu(\gamma) \gamma : \sum_{\gamma \in \Gamma} |\mu(\gamma)| < \infty\}$ denote the group algebra. Given any unitary representation $\pi$ of $\Gamma$ in a Hilbert space $\mathcal{H}$, extend $\pi$ to the group algebra by: $\pi(\mu) = \sum_{\gamma \in \Gamma} \mu(\gamma) \pi(\gamma)$. Denote by $\mathcal{H}_1$ the space of vectors invariant under every $\pi(\gamma)$, $\gamma \in \Gamma$, and by $E_1$ the orthogonal projection on $\mathcal{H}_1$.

**Definition 1.** A sequence $\nu_n \in \ell^1(\Gamma)$ is a mean ergodic sequence if for every unitary representation $\pi$ of $\Gamma$, $\|\pi(\nu_n) f - E_1 f\| \xrightarrow{n \to \infty} 0$ for all $f \in \mathcal{H}$.

Let $(X, \mathcal{B}, m)$ be a standard Lebesgue measure space, namely a measure space whose $\sigma$-algebra is countably generated and countably separated. Assume $\Gamma$ acts on $X$ by measurable automorphisms preserving the probability measure $m$. The action $(\gamma, x) \mapsto \gamma x$ induces a representation of $\Gamma$ by isometries on the $L^p(X)$ spaces, $1 \leq p \leq \infty$, and this representation can be extended to the group algebra by: $(\mu f)(x) = \sum_{\gamma \in \Gamma} \mu(\gamma) f(\gamma^{-1} x)$.

The set $\mathcal{B}_1 = \{A \in \mathcal{B} \mid m(\gamma A \Delta A) = 0 \quad \forall \gamma \in \Gamma\}$ is a sub-$\sigma$-algebra, and denote by $E_1$ the conditional expectation operator on $L^1(X)$ which is associated with $\mathcal{B}_1$.

**Definition 2.** A sequence $\nu_n \in \ell^1(\Gamma)$ is called a pointwise ergodic sequence in $L^p$ if, for any action of $\Gamma$ on a standard Lebesgue space $X$ which preserves a probability measure and for every $f \in L^p(X)$, $\nu_n f(x) \to E_1 f(x)$ for almost all $x \in X$, and in the norm of $L^p(X)$.

**Remark.** Let us note that ergodic sequences, both mean and pointwise, always exist. Indeed, fix a probability measure $\mu \in \ell^1(\Gamma)$ s.t. $\text{supp} \, \mu \overset{d}{=} \{\gamma \mid \mu(\gamma) \neq 0\}$ generates $\Gamma$ as a group. Then $\nu_n = \frac{1}{n+1} \sum_{k=0}^{n} \mu^k$ is a mean ergodic sequence by von Neumann's theorem for the operator $\pi(\mu)$, and also a pointwise ergodic sequence in $L^p$, $1 \leq p < \infty$, by the Dunford-Schwartz theorem for Markov.
operators [D-S, Vol. I, Chapter VIII, §6.5]. For a discussion of sequences of the form $\nu_n = \mu^n$, where $\mu$ is a probability measure, we refer to [Os, S] or [J-R-T].

1.2. Introduction. It is natural to consider sequences in $\ell^1(\Gamma)$ which are given in an explicit geometric form. To that end, assume $\Gamma$ is finitely generated, and let $S$ be a finite generating set which is symmetric: $S = S^{-1}$ (we will assume $e \notin S$). $S$ induces a length function $|\gamma|_S = \min\{|\gamma| = s_1 \cdots s_n, s_i \in S\}, |e| = 0$. Define the following sequences:

(i) $\sigma_n = \frac{1}{\# S_n} \sum_{w \in S_n} w$, where $S_n = \{w : |w| = n\}$ is the sphere of radius $n$, with center $e$. Define also: $\sigma'_n = \frac{1}{2}(\sigma_n + \sigma_{n+1}).$

(ii) $\mu_n = \frac{1}{n+1} \sum_{k=0}^n \sigma_k$, the average of the first $n+1$ spheres, $\mu_0 = \sigma_0 = e$.

(iii) $\beta_n = \frac{1}{\# B_n} \sum_{w \in B_n} w$, where $B_n = \{w : |w| \leq n\}$ denotes the ball of radius $n$ with center $e$.

The fundamental observation underlying the discussion below is that the convolution algebra $A_c(\Gamma)$ generated by the elements $\sigma_n$ in $\ell^1(\Gamma)$ is, in favourable cases, much simpler than $\ell^1(\Gamma)$ and admits an explicit spectral theory which can be used to prove ergodic theorems in much the same way as in the classical set-up. This observation is due originally to V. I. Arnold and A. L. Krylov [A-K], and we take one of the cases they considered, that of the free group $F_k$ on $k$ generators, $1 < k < \infty$, to illustrate the point. Indeed, if we take $S$ as a set of free generators and their inverses, then the algebra alluded to above, which we denote by $A_c(F_k)$, is in fact a cyclic algebra and is generated by the element $\sigma_1$. This fact is an immediate consequence of the identity:

$$\sigma_n * \sigma_1 = \frac{1}{2k} \sigma_{n-1} + \left(1 - \frac{1}{2k}\right) \sigma_{n+1}, \quad n \geq 1.$$ 

The identity, in turn, follows from the obvious fact that multiplying all words of length $n$ by all words of length 1 each word of length $n+1$ is obtained once and each word of length $n-1$ is obtained $2k-1$ times. In [A-K] this identity was utilized to prove an equidistribution theorem for words of length $n$ in dense free subgroups of the group of (orientation preserving) isometries of the Euclidean unit sphere $S^d$ when $d \geq 2$. This result generalizes Weyl's equidistribution theorem for dense free (i.e., infinite cyclic) subgroups of the group of (orientation preserving) isometries of the unit circle. Equidistribution of words of length $n$ implies a mean ergodic theorem for the sequence $\sigma_n$ when the unitary representation taken is the one in $L^2(S^d)$.

The fact that the sequence $\sigma'_n$ is a mean ergodic sequence for any unitary representation is due to Y. Guivarc'h [G], and his proof utilizes the spectral theorem in much the same way as von Neumann's original proof of the mean ergodic theorem. We indicate the argument briefly: Given a unitary representation $\pi$ of $F_k$ in a Hilbert space $H$, extend $\pi$ to a $\ast$-representation of the algebra $A_c(F_k)$ generated by $\sigma_1$ in $\ell^1(F_k)$. This algebra contains each $\sigma'_n$, and, as usual, for the mean ergodic theorem it is enough to show that $\|\pi(\sigma'_n)v\| \to 0$.
if \( \psi \) is orthogonal to the space of invariants. However, \( \pi(\sigma_1) \) is a selfadjoint operator and we can identify the points in its spectrum with the norm-continuous \(*\)-homomorphisms of the algebra it generates in \( \text{End}(\mathcal{A}) \). The spectral theorem for \( \pi(\sigma_1) \) implies that it is enough to show that \( \phi_0(\pi(\sigma'_n)) \to 0 \) for the set of homomorphisms \( \phi_0 \) just described. Consequently, if \( \phi(\sigma'_n) \to 0 \) for all norm-continuous nontrivial \(*\)-homomorphisms \( \phi \) of \( A_c(F_k) \) which take a real value at \( \sigma_1 \), then the mean ergodic theorem follows.

This discussion shows that we should focus our attention on the algebra \( A_c(F_k) \) and its spectral theory. We remark that the algebras \( A_c(F_k) \) introduced into ergodic theory in [A-K] are isomorphic when \( 2k = p^t + 1 \), \( p \) prime, to certain \( p \)-adic Hecke rings introduced by Satake [Sa], and it will be found useful in what follows to keep both realizations in mind. The structure, representation theory, and spectrum of these algebras are well studied [e.g., Sa, I-M, Ma, Mc, C, FT-P]. The most important fact about the (norm-continuous) real characters \( \phi(\sigma_n) \) of the algebra \( A_c(F_k) \) is that they all decay to zero exponentially as \( n \to \infty \), with two exceptions. One is the trivial character, and the other is the special character given by: \( \epsilon(\sigma_n) = (-1)^n \), whose existence makes it necessary to consider \( \sigma'_n \) rather than \( \sigma_n \).

In the following we propose to extend this approach to ergodic theorems in three directions:

(1) First consider the problem of pointwise ergodic theorems for actions of the free group \( F_k \), at least for \( L^2 \)-functions. We will start by proving a pointwise ergodic theorem for the sequence \( \mu_n \) introduced above, and the main tool in the proof will be a strong maximal inequality in \( L^2 \). The maximal inequality generalizes similar ones due to E. M. Stein [S, S1] and B. Weiss [W], and was suggested in [W]. It applies to any sequence of selfadjoint Markov operators in \( L^2 \), provided that the sequence is subadditive, in a sense which will be made precise below. We remark that the subadditive maximal inequality is a very useful one. It was applied by E. M. Stein [S, S1] to prove a pointwise ergodic theorem for the even powers of positivity-preserving selfadjoint contractions on \( L^2 \). Also, it is noted in [S, S1, W] that it implies the pointwise convergence of martingales in \( L^2 \), and, moreover, it implies Birkhoff's pointwise ergodic theorem in \( L^2 \). The origin of this maximal inequality is attributed in [S, W] to A. Kolmogoroff and G. Seliverstoff [K-S], and to R. E. A. C. Paley [P].

Having proved a maximal inequality for \( \mu_n \), which is defined as the average of the first \( n + 1 \) spheres, we will use it to prove a maximal inequality for the average \( \sigma'_n \) of two spheres. The argument here is modelled on a square-function argument due to E. M. Stein [S], and uses, unlike the previous maximal inequality, detailed information about the characters of \( A_c(F_k) \), and in particular the exponential decay of \( \phi(\sigma'_n) \) for nontrivial \( \phi \). We will also show that \( \sigma_{2n} \) and \( \beta_{2n} \) converge strongly and pointwise almost everywhere (for each \( f \in L^2(X) \)) to a limit operator. In the case of \( \sigma_{2n} \), the limit is a conditional expectation operator associated with a two-point factor, if the action is ergodic. Note that \( \beta_{2n} \) is not a subadditive sequence, and the proof of the maximal inequality for \( \beta_{2n} \) is obtained as a consequence of the maximal inequality for \( \sigma'_n \).
The ergodic theorems stated above depend only on the convergence of certain sequences of operators in a (commutative) Banach *-algebra, and therefore it is natural to look for groups which admit an embedding of the algebra \( A_c(F_k) \) into their group algebra. Where to look becomes evident when we consider the following natural interpretation of the algebra \( A_c(F_k) \): The elements of \( F_k \) correspond bijectively to the vertices of a regular tree \( T_{2k} \) of valency \( 2k \) (the Cayley graph w.r.t. \( S \)), and \( F_k \) acts isometrically on this tree by left translation. The operator of right convolution by the element \( \sigma_n \) on \( \ell^2(F_k) \) corresponds to the operator of averaging a function in \( \ell^2(T_{2k}) \) on spheres of radius \( n \). The algebra \( A_{2k} \) generated by these averaging operators in \( \text{End}(\ell^2(T_{2k})) \) is isomorphic to \( A_c(F_k) \). Note that the averaging operators commute not only with left translations by elements of \( F_k \) but with the whole group \( \text{Aut}(T_{2k}) \) of isometries of the tree. Furthermore, it is not hard to see that the algebra \( A_{2k} \) is isomorphic to a subalgebra of the group algebra of \( \text{Aut}(T_{2k}) \). The latter subalgebra consists of double cosets of a maximal compact subgroup of \( \text{Aut}(T_{2k}) \) which fixes a vertex in the tree, and constitutes a Gelfand pair algebra (e.g., [Fa]). Therefore, since \( L^1(\text{Aut}(T_{2k})) \) contains a copy of \( A(F_k) \), it follows that \( \text{Aut}(T_{2k}) \) also satisfies the ergodic theorems above.

The same conclusion will be shown to hold for two classes of subgroups of \( \text{Aut}(T_{2k}) \). One is the class of closed noncompact subgroups acting transitively on the boundary of the tree, for example, \( \text{PGL}_2(\mathbb{Q}_p) \). The other is the class of complemented lattices, namely lattice subgroups acting transitively on the vertices of the tree. We refer to §5 for a discussion of these matters.

Finally, it is possible to extend the class of algebras under discussion, and include the Gelfand pairs associated with any of the semi-homogenous trees, namely the bipartite trees \( T(r_1, r_2) \) with at most two (finite) valencies \( r_1 \geq 2, r_2 \geq 2, r_1 + r_2 > 4 \) (we exclude the case of the tree \( \mathbb{Z} \)). In this context, it is natural to consider also the double cosets algebra associated with the compact open subgroup stabilizing a geometric edge. This algebra includes as a proper subalgebra the Gelfand pair associated with the stabilizer of a vertex, but is no longer commutative. Its spectral theory, however, is analogous to the cases already considered. Indeed, the algebra has the property that all of its irreducible *-representations are of bounded degree, in fact all but two of them have degree 2. Moreover, the corresponding positive-definite functions satisfy analogous exponential decay estimates. We will use the spectral results obtained (mostly due to [Ma]) to prove ergodic theorems for another class of lattices in \( G(r_1, r_2) \), which will be described below. This class includes, for example, \( \text{PSL}_2(\mathbb{Z}) \).

### 1.3. Statement of the ergodic theorems

In order to avoid repetitions and minimize the amount of computation, we will state and prove the results in the following way: We will discuss first discrete groups whose group algebra contains a commutative convolution subalgebra of a simple kind (which will turn out to be a Gelfand pair), and prove the ergodic theorems in this case. Then we will consider the connection between these algebras and groups acting on trees, and use it to produce some groups which admit an embedding of the algebra into their group algebra. We will then proceed to describe the spectrum of a
noncommutative convolution algebra (which is again a double coset algebra),
and finally derive the ergodic theorems in this case.

Accordingly, in the first half of the paper (§§2–4), we will consider the fol-
lowing groups (and use $\Gamma$ to denote any one of them):

(1) $F_k = \text{the free group on } k \text{ generators}, 1 < k < \infty$, with $S = \{x_i, x_i^{-1}\}_{i=1}^k$, where $x_1, \ldots, x_k$ are free generators.

(2) $\Gamma(r, h) = G_1 \ast G_2 \ast \cdots \ast G_r = \text{the free product of } r \text{ finite groups each of order } h$, where $r \geq 2, h \geq 2, r + h > 4$, with $S = \bigcup_{i=1}^r G_i \setminus \{e\}$.

We define $q(F_k) \overset{d}{=} 2k - 1$ and $q(\Gamma(r, h)) \overset{d}{=} (r - 1)(h - 1)$. Note that $q(F_k) = q(\Gamma(2k, 2))$. Associating the pair $(2k, 2)$ with $F_k$ we have, in all cases, $\#S_n = r(h - 1)q^{n-1}$, where we define: $h(\Gamma(r, h)) \overset{d}{=} h, h(F_k) \overset{d}{=} 2$.

As mentioned above, it has been proved by Y. Guivarc'h [G] that for $\Gamma = F_k$, the sequence $\sigma_n'$ is a mean ergodic sequence. This result has the following generalization:

**Theorem 1.** For $\Gamma$ as above, the sequences $\sigma_n'$ and $\mu_n$ are mean ergodic sequences, but $\sigma_n$ and $\beta_n$ are not. $\sigma_{2n}$ converges strongly to the operator $E_1 + c(\Gamma)E$, where $E$ is a projection disjoint from $E_1$. $\beta_{2n}$ converges strongly to the operator $E_1 + c(\Gamma)\frac{q(\Gamma)^{-1}}{q(\Gamma)+1}E$. Here $c(\Gamma) = \frac{q-(r-1)(h+2)}{r(h-1)(1-q^{-1})}$, and in particular $c(\Gamma) = 1$ iff $\Gamma = \Gamma(r, 2)$ or $\Gamma = F_k$.

Moreover, we have the following pointwise ergodic theorem:

**Theorem 2.** For $\Gamma$ as above, the sequences $\sigma_n'$ and $\mu_n$ are pointwise ergodic sequences in $L^2$. Similarly, the convergence of $\sigma_{2n}$ and $\beta_{2n}$ to the limit given in Theorem 1 is, for each function $f \in L^2(X)$, pointwise almost everywhere and in the norm of $L^2(X)$. The limit $E_1 + c(\Gamma)E$ of $\sigma_{2n}$ is a conditional expectation operator w.r.t. a $\Gamma$-invariant sub-$\sigma$-algebra if $\Gamma = \Gamma(r, 2)$ or $\Gamma = F_k$.

Our proof of the pointwise ergodic theorem for the sequences $\mu_n$, $\sigma_n'$ will have as its basic ingredient a strong $L^2$ maximal inequality. Let us first define the relevant maximal functions: $f^*(x) = \sup_{n \geq 0} |\mu_n f(x)|$ and $Mf(x) = \sup_{n \geq 0} |\sigma_n' f(x)|$. We then have:

**Theorem 3.** Let $\Gamma$ be as above. If $f \in L^2(X)$, then $f^*$ and $Mf$ are in $L^2(X)$ and

$$\|f^*\|_2 \leq 16h(\Gamma)\|f\|_2,$$

$$\|Mf\|_2 \leq B(\Gamma)\|f\|_2.$$  

**Remark.** It is interesting to note that for $\Gamma = F_k$, for example, the constant obtained in the maximal inequality for $\mu_n$ is independent of the number of generators $k$.

The proof of Theorem 3 depends on a subadditivity property satisfied by the sequence $\mu_n$, which we proceed to define:

**Definition.** A sequence $T_n$ of operators on $L^2(X)$ will be called a subadditive sequence of selfadjoint Markov operators if it satisfies the following:

(i) $T_n = T_n^*, \|T_n\| \leq 1$. 

(ii) $T_nf \geq 0$ if $f \geq 0$, $T_n1 = 1$.

(iii) There exists a constant $C > 0$ and a positive integer $k$, such that

$$T_nT_mf(x) \leq C(T_{nk}f(x) + T_{km}f(x))$$

for all bounded and nonnegative $f \in L^2(X)$.

We can now state the following result, which generalizes similar ones in [S, S1, W], and was suggested in [W]:

**Theorem 4.** Let $T_n$ be a subadditive sequence of selfadjoint Markov operators. Define $f^*(x) = \sup_{n \geq 0} |T_nf(x)|$. Then $\|f^*\|_2 \leq 2C\|f\|_2$ for all $f \in L^2(X)$, where $C$ is the constant implied by the definition of subadditivity.

The proof of Theorem 2 is as follows: We will show that the sequence of operators $\mu_n$ is subadditive, and so it has a maximal inequality. Then, following [S], the maximal inequality for $\sigma_n^*$ will be deduced from that of $\mu_n$, using the spectral theory of $A(\Gamma)$. Spectral estimates will then be applied to prove pointwise convergence on a dense subset, which, together with the maximal inequality, yields the pointwise ergodic theorem.

### 1.4. Double cosets algebras and length functions

As noted already, the commutative algebras $A_c(\Gamma)$ discussed so far admit a useful interpretation which points the way to a generalization of the discussion above. (We refer to §5 for terminology left unexplained here.) Associate with any finitely generated group $\Gamma$ and a finite symmetric generating set $S$ the (undirected) Cayley graph $X = X(\Gamma, S)$ with the edge path metric. Denote by $G = \text{Aut}(X)$ the group of isometries or equivalently the group of graph automorphisms. Let $C_c(G, K)$ denote the convolution algebra of bi-$K$-invariant locally constant compactly supported functions on $G$, where $K$ is the open compact subgroup fixing $e \in \Gamma$. The connection between the algebras is given by:

**Proposition 5.** The algebra $A_c(\Gamma)$ generated by $\{\sigma_n, n \geq 0\}$ is $*$-isometrically isomorphic to a subalgebra of $C_c(G, K)$.

We remark that in the cases discussed above, namely $\Gamma = \Gamma(r, h)$ or $\Gamma = F_k$, $G = \text{Aut}(X)$ is naturally isomorphic to the group $G(r_1, r_2)$ of automorphisms of the semi-homogeneous tree $T(r_1, r_2)$, for some $r_1, r_2$ depending on $\Gamma$ (see §5). Moreover, in these cases the subgroup $K$ is maximal compact, $C_c(G, K)$ is a Gelfand pair (namely it is a commutative algebra), $\Gamma$ is a lattice in $G$, and $A_c(\Gamma)$ is isometrically isomorphic to $C_c(G, K)$.

It is therefore natural to consider other lattices of $G = G(r_1, r_2)$ which have the same property, namely contain $C_c(G, K)$ as a subalgebra of their group algebra. Let us call two subgroups $\Gamma$ and $H$ of a group $G$ complementary if $G = \Gamma H$. A lattice of $G(r_1, r_2)$ will be called complemented if it is complementary to the stabilizer in $G$ of a vertex $v \in T(r_1, r_2)$. In that case $\Gamma$ is transitive on the orbit of $v$ under $G$. Define the following length function on $\Gamma$: $\ell(\gamma) = \frac{1}{2}d(v, v\gamma)$ if $r_1 \neq r_2$, and $\ell(\gamma) = d(v, v\gamma)$ if $r_1 = r_2$, where $d$ denotes the edge path metric in $T$. Consider the sequences $\sigma_n^*(\ell)$ and $\mu_n(\ell)$ defined with the length function $\ell$, namely: $\sigma_n(\ell) = \frac{1}{\#S_n(\ell)} \sum_{\ell(w) = n} w$, $\mu_n(\ell) =$
$\frac{1}{n+1} \sum_{k=0}^{n} \sigma_k(\ell) = \frac{1}{2}(\sigma_n(\ell) + \sigma_{n+1}(\ell))$, where $S_n(\ell) = \{ w \mid \ell(w) = n \}$.

As will be seen in §5, the results of Theorems 1 and 2 will be incorporated into the following framework:

**Theorem 6.** Let $\Gamma \subset G(r_1, r_2)$ be a complemented lattice. Then the sequences $\sigma'_n(\ell)$, $\mu_n(\ell)$ are mean ergodic sequences and pointwise ergodic sequences in $L^2$.

Clearly, the proof of the pointwise ergodic theorem can also be applied to nondiscrete groups provided their group algebra contains a convolution subalgebra isomorphic to one of the algebras discussed above. Examples of such groups are quite abundant: Consider a closed noncompact subgroup $H$ contained in $G(r_1, r_2)$, which acts transitively on the boundary of $T(r_1, r_2)$, namely on the space of ends. Let $\nu \in T(r_1, r_2)$ be a vertex, and let $H_v$ be the stabilizer in $H$ of $\nu$. $H$ has either one or two orbits of vertices in $T$ [N1, Proposition 3], and let $i \in \{1, 2\}$ denote the number of orbits of $H$ in the vertices of $T$. Consider the sets $\Sigma_n = \{ g \in H \mid d(\nu, gv) = in \}$, where $d$ is the edge path metric in $T$. $\Sigma_n$ is a double coset of $H_v$ in $H$ (see §6), and we define $\sigma_n(H)$ to be the (absolutely continuous) probability measure on $H$ obtained by dividing the characteristic function of $\Sigma_n$ by the Haar measure of $\Sigma_n$. Put also $\sigma'_n(H) = \frac{1}{2}(\sigma_n(H) + \sigma_{n+1}(H))$ and $\mu_n(H) = \frac{1}{n+1} \sum_{k=0}^{n} \sigma_k(H)$. We can now state:

**Theorem 7.** Let $H$ be a closed noncompact boundary-transitive subgroup of $G(r_1, r_2)$. Given a vertex $\nu \in T$, the associated sequences $\sigma'_n(H)$, $\mu_n(H)$ of probability measures defined above satisfy a mean ergodic theorem for continuous unitary representations of $H$, and a pointwise ergodic theorem in $L^2$ for measure-preserving (jointly measurable) actions of $H$. When $r_1 = r_2$ and $H$ has two orbits in $T$, the same conclusion holds already for $\sigma_n(H)$.

**Corollary.** The sequence of radial measures $\sigma_n$ on $\text{PSL}_2(\mathbb{Q}_p)$ is a pointwise ergodic sequence in $L^2$.

Finally, as noted above, it is natural to consider the double cosets algebra associated with the compact open subgroup stabilizing a geometric edge. It contains the algebra of double cosets of a maximal compact subgroup, but it is not commutative. We will once again use the language of discrete groups, and consider a convolution algebra which will be shown (Lemma 6.1) to be isomorphic to the double cosets algebra.

**Theorem 8.** Let $\Gamma = G_1 \ast G_2$, where $\#G_i = r_i$ and $r_1 > r_2 \geq 2$. Let $S = (G_1 \cup G_2) \setminus \{e\}$.

1. If $r_2 > 2$, then $\sigma'_n$ and $\mu_n$, defined using the length function $| \cdot |_S$, are mean ergodic sequences.

2. If $r_2 = 2$, then $\sigma''_n \overset{d}{=} \frac{1}{2}(\sigma_{2n} + \sigma_{2n+2})$ and $\mu_n$ are mean ergodic sequences. Moreover, $\sigma_{2n}$ converges strongly to a conditional expectation w.r.t a $\Gamma$-invariant $\sigma$-algebra in the case of an action, and to a projection of the form $E_1 + E$ in the case of a unitary representation.

**Corollary.** The sequence $\sigma''_n$ in the group algebra of $\text{PSL}_3(\mathbb{Z}) = \mathbb{Z}_2 \ast \mathbb{Z}_3$, defined w.r.t. the generating set $(\mathbb{Z}_2 \cup \mathbb{Z}_3) \setminus \{e\}$, is a mean ergodic sequence.
1.5. Remarks and some relevant references. (1) Consider a probability space $X$ and a sequence $\sigma_n$ of selfadjoint Markov operators on $L^2(X)$. The two essential ingredients that figure in the proof of the pointwise convergence of $\sigma_n$ are:

(i) The sequence generates an algebra of operators whose nontrivial irreducible $*$-representations $\pi$ have the property that $\|\pi(\sigma_n)\|$ decays exponentially in $n$.

(ii) The sequence of uniform averages $\mu_n = \frac{1}{n+1} \sum_{k=0}^{n} \sigma_k$ satisfies a strong $L^2$ maximal inequality, for example, it is subadditive.

As noted above, the basic case of the algebra $L^1(\mathbb{N})$, namely when $\sigma_n = \sigma^n$, $\sigma$ a selfadjoint Markov operator, was considered by E. M. Stein [S, SI], and forms the main motivation of the present paper.

(2) Let us compare (in $L^2$) the pointwise ergodic theorem obtained for the free group on $k > 1$ generators and the classical Birkhoff theorem for the free group on one generator, namely $\mathbb{Z}$. We see that for the non-Abelian free groups, $\sigma_n'$ as well as $\mu_n$ are mean and pointwise ergodic sequences, whereas for $\mathbb{Z}$ only the latter has these properties. The mean convergence of the sequence $\sigma_n'$ is a mixing property, and the pointwise result for $\sigma_n'$ can be interpreted as a more precise form of the Howe-Moore mixing theorem [H-M], when the averages considered are radial.

(3) In view of the previous remarks, it is natural to expect that similar radial pointwise ergodic theorems hold for closed noncompact boundary-transitive subgroups of the group of automorphisms of affine Bruhat-Tits buildings of higher split rank, as well as for their complemented lattices. Indeed, the theory of spherical functions associated with an Iwahori subgroup, as developed in [Sa, Ma, Mc], exhibit a similar exponential decay phenomenon. This subject will be discussed further elsewhere.

(4) It is natural to consider the behaviour of the radial averages in $L^p$ for $p \neq 2$. This problem is considered in a sequel to the present paper [N-S], where it is shown that the sequence $\sigma_n' \in L^1(F_k)$ is a pointwise ergodic sequence in $L^p$, for $p > 1$.

(5) A similar approach is applied in [N1, N2] to prove pointwise ergodic theorems for spherical averages on simple Lie groups of real rank one. We note that in the case of $\mathbb{R}^n$, $n \geq 3$, a maximal inequality for spherical averages was proved in [S2], and was subsequently used to prove radial pointwise ergodic theorems for $\mathbb{R}^n$-actions ($n \geq 3$) in [J]. We remark that radial mean ergodic theorems have numerous interesting applications, of which we mention [F-K-W] and [E-M].

(6) It is reasonable to expect that the subadditive maximal inequality for $\mu_n$ holds for a class of discrete groups much wider than the class discussed in the present paper. The inequality also makes sense for semi-simple Lie groups for example, and it does in fact hold for the groups $G = SO^0(n, 1)$.

(7) With regard to terminology, we have chosen to call all the algebras under discussion double cosets algebras. Some of them were variously referred to as $p$-adic Hecke rings (e.g., [Sa]), radial convolution algebras (e.g., [FT-P]), and affine Coxeter systems [Ma].
2. Proof of the mean ergodic theorem (Theorem 1)

2.1. The spectral approach. We start with the basic observation that the algebra $A_c(\Gamma)$ generated by $\sigma_1$ in $\ell^1(\Gamma(r, h))$ contains each $\sigma_n$, $n \geq 0$. This follows, by induction, from the easily verified formula:

$$\sigma_n * \sigma_1 = \frac{1}{r(h-1)} \sigma_{n-1} + \frac{h-2}{r(h-1)} \sigma_n + \frac{r-1}{r} \sigma_{n+1}. \tag{2.1.1}$$

We remark that, taking $h = 2$ and $r = 2k$, (2.1.1) gives the recurrence relations satisfied in the algebra generated by $\sigma_1$ in $\ell^1(F_k)$, and in fact $A_c(F_k) \cong A_c(\Gamma(2k, 2))$. Denote by $A(\Gamma)$ the norm-closure of the algebra generated by $\sigma_1$ in $\ell^1(\Gamma)$.

Let there be given a unitary representation $\pi$ of $\Gamma$ in a Hilbert space $H$. Using the spectral theorem for the selfadjoint operator $\pi(\sigma_1)$, if $\nu_n = p_n(\sigma_1)$ are polynomials in $\sigma_1$:

$$\int_{sp\pi(\sigma_1)} p_n(\varphi(\sigma_1)) d\mu_\nu(\varphi) = \int_{sp\pi(\sigma_1)} \varphi(\nu_n) d\mu_\nu(\varphi) \tag{2.1.2}$$

where $\varphi : A(\Gamma) \to \mathbb{C}$ is a continuous homomorphism and $\mu_\nu$ is the spectral measure associated with $\nu \in H$. $sp\pi(\sigma_1) \subset \mathbb{R}$, and hence $\varphi(\sigma_1) \neq \gamma(\varphi)$ is real. Suppose that $\nu_n$ are symmetric probability measures in $A(\Gamma)$, and $\varphi(\nu_n) \to 0$ for all nontrivial continuous homomorphisms of $A(\Gamma)$ that take a real value on $\sigma_1$. Then by Lebesgue dominated convergence, if $\nu \perp ker(\pi(\sigma_1) - I)$:

$$\int_{sp\pi(\sigma_1)} \varphi(\nu_n) d\mu_\nu(\varphi) \to m_\nu(1) = 0. \tag{2.1.3}$$

Moreover:

$$\langle \pi(\nu_n) v, \pi(\nu_n) v \rangle = \langle \pi(\nu_n^2) v, v \rangle = \int_{sp\pi(\sigma_1)} \varphi(\nu_n)^2 d\mu_\nu(\varphi) \to m_\nu(1) = 0. \tag{2.2.1}$$

Now note that if $E'$ denotes the projection onto $ker(\pi(\sigma_1) - I)$, then for any $v \in H$, $E'v$ is invariant under $\pi(\sigma_1)$, and therefore $\Gamma$-invariant. This follows from the fact that $\frac{1}{|S|} \sum_{s \in S} \pi(s)v = v$ implies $\pi(s)v = v$, $\forall s \in S$, by convexity. Therefore $E_1 = E'$ and $\|\pi(\nu_n)(v - E_1v)\| = \|\pi(\nu_n)v - E_1v\| \to 0$, as required.

We proceed to examine the characters of $A(\Gamma)$.

2.2. The spectral theory of $A(\Gamma)$. We now invoke the analysis of, e.g., [FT-P] (for $F_k$) and [I-P] (for $\Gamma(r, h)$) (see also, e.g., [Ma, Mc, C, BK]). We give a brief summary of the relevant facts about $A(\Gamma) = A$ which will be used later. Applying $\varphi$ to both sides of (2.1.1) we get, for $n \geq 1$:

$$\varphi(\sigma_1) \varphi(\sigma_n) = \frac{1}{r(h-1)} \varphi(\sigma_{n-1}) + \frac{h-2}{r(h-1)} \varphi(\sigma_n) + \frac{r-1}{r} \varphi(\sigma_{n+1}). \tag{2.2.1}$$
which is a second order recurrence relation on \( N \). Two linearly independent solutions of (2.2.1) are \( q^{-nz} \) and \( q^{-n(1-z)} \), where \( q = (r - 1)(h - 1) \), provided \( q^{-z} \neq q^{-1(1-z)} \). The equality \( q^{-z} = q^{-1(1-z)} \) is equivalent to \( z = \frac{1}{2} + \frac{ij\pi}{\log q} \), \( j \in \mathbb{Z} \), and for these values of \( z \), \( q^{-nz} \) and \( nq^{-nz} \) are the two linearly independent solutions. The two solutions are associated with the eigenvalue:

\[
(2.2.2) \quad \phi_z(\sigma_1) = \gamma(z) = \frac{1}{r(h-1)} \left( q^z + q^{1-z} + h - 2 \right).
\]

Any homomorphism \( \phi_z : A \to \mathbb{C} \) will be of the form

\[
(2.2.3) \quad \phi_z(\sigma_n) = c_z q^{-nz} + c_{1-z} q^{-n(1-z)}, \quad q^z \neq q^{1-z}
\]
or of the form

\[
(2.2.4) \quad \phi_z(\sigma_n) = \left[ a_z + a_z' n \right] (-1)^{n} q^{-\frac{n}{2}}, \quad z = \frac{1}{2} + \frac{ij\pi}{\log q}.
\]

The coefficients are uniquely determined by the two linear conditions \( \phi_z(\sigma_0) = 1 \) and \( \phi_z(\sigma_1) = \gamma(z) \). The results are, respectively:

\[
(2.2.5) \quad c_z = \frac{q^{1-z} - (r - 1)^{-1} q^z + h - 2}{r(h-1)(q^z - q^{z-1})}, \quad c_{1-z} = c_{1-z}
\]
and

\[
(2.2.6) \quad a_z = 1, \quad a_z' = \frac{2q + (-1)^{j} (h - 2) \sqrt{q} - r(h-1)}{r(h-1)}.
\]

To find the real spectrum, note that a continuous homomorphism of \( A \) is a bounded linear functional, and boundedness of \( \phi_z(\sigma_n) \), \( n \geq 0 \), is equivalent to \( 0 \leq \text{Re} z \leq 1 \). Now \( \phi_z(\sigma_1) = \gamma(z) \) must be real, which by (2.2.2) happens iff \( \text{Re} z = 1/2 \) or \( \text{Im} z = j\zeta \), \( j \in \mathbb{Z} \), where we put \( \zeta \frac{d}{d \log q} \). Note that since \( A \) is cyclic, the homomorphism \( \phi_z \) is determined by \( \phi_z(\sigma_1) = \gamma(z) \), and the set of values \( \gamma(z) \) takes on \( \frac{1}{2} + i(1 + t)\zeta \), \( 0 \leq t \leq 1 \), coincides with the set of values \( \gamma(z) \) takes on \( \frac{1}{2} + i\zeta t \), \( 0 \leq t \leq 1 \), as is easily checked. Furthermore, the functions \( z \mapsto \phi_z \) and \( z \mapsto \gamma(z) \) are periodic with period \( \frac{2\pi i}{\log q} \) and invariant under the transformation \( z \mapsto 1 - z \). Taking these facts into account, the real spectrum of \( A(G) \) is given by \( \{ \gamma(z) \mid \text{Re} z = 1/2 \text{ and } 0 \leq \text{Im} z \leq \zeta \text{, or } 0 \leq \text{Re} z \leq \frac{1}{2} \text{ and } \text{Im} z = j\zeta \text{, } j = 0, 1 \} \) [FT-P, I-P]. This set of representatives is depicted in Figure 1.

2.3. **Spectral estimates: Exponential decay of the characters.** We rewrite the real characters of \( A \), using (2.2.5) and (2.2.6), in a form which will be found useful later. We distinguish four cases:

(i) The principal series: \( z = \frac{1}{2} + it\zeta \) where \( 0 < t < 1 \). Here \( c_{1-z} = c_z \) and

\[
(2.3.1) \quad \phi_z(\sigma_n) = 2 \text{Re} c_z q^{-nz} = q^{-n/2} \left( \cos nt - \frac{r - 2}{r} \cos t + \frac{h - 2}{r(h-1)} \sin nt \right).
\]

(ii) The ends of the principal series: \( z = \frac{1}{2} + ij\zeta \), \( j = 0, 1 \):

\[
(2.3.2) \quad \phi_z(\sigma_n) = (1 + a_j n)(-1)^{jn} q^{-n/2}, \quad 0 < |a_j| < 3.
\]
Figure 1

(iii) The complementary series: 
\[ z = s + ij\zeta, \quad j = 0, 1, \quad 0 < s < \frac{1}{2} \] 
(2.3.3) \[ \varphi_{s+ij\zeta}(\sigma_n) = (-1)^{jn} \left[ c_{s+ij\zeta} q^{-ns} + c_{1-s-ij\zeta} q^{-n(1-s)} \right]. \]

Here \( c_{s+ij\zeta} = \frac{w(s)}{q^{-s} - q^{s+1}} \) where \( w(s) \) is analytic on \( \mathbb{R} \).

(iv) The ends of the complementary series: 
\[ z = ij\zeta, \quad j = 0, 1 \]
(2.3.4) \[ \varphi_{ij\zeta}(\sigma_n) = c_{ij\zeta}(-1)^{jn} + c_{1-ij\zeta}(-1)^{jn} q^{-n} \]

Here \( c_{ij\zeta} = \frac{(-1)^{j(q-(r-1)^{-1})+h-2}}{(-1)^{j(r(h-1)(1-\frac{h}{2}))}} \). Note that \( c_{ij\zeta} = 1 \) (and \( c_{1-ij\zeta} = 0 \)) iff \( h = 2 \).

Lemma 2.1. The following estimates hold (for some positive constant \( C(q) \)):

(i) For the principal strip given by \( \frac{1}{4} \leq \text{Re} \, z \leq \frac{1}{2} \):
\[ |\varphi_z(\sigma_n)| \leq C(q)(n+1) q^{-n/4}. \]

(ii) For the complementary series near the endpoints, \( 0 < \text{Re} \, z = s < \frac{1}{4} \):
\[ |\varphi_{s+ij\zeta}(\sigma_n)| \leq C(q)q^{-ns}. \]

Proof. The second estimate follows from (2.3.3) and the fact that \( |c_{s+ij\zeta}| \) is bounded for \( s \in [0, \frac{1}{4}] \).

As to the first, note that on the interval \( \frac{1}{2} + it\zeta, \quad 0 < t < 1 \), we can use the estimate \( |\sin nt| \leq n \), by (2.3.1). Similarly, on the interval \( s + ij\zeta, \quad \frac{1}{4} \leq s < \frac{1}{2} \), let us write:
\[ u(s) = q^{1-s} - (r-1)^{-1} q^s = 2\sqrt{h-1} \sinh \left( \frac{1}{2} - s \right) \log q + \frac{1}{2} \log(r-1) \],
so that \( c_{s+ij\zeta} \) takes the form:
\[ c_{s+ij\zeta} = \frac{h-2 + (-1)^{j} u(s)}{(-1)^{j} r(h-1) q^{-\frac{1}{2}} \sinh((\frac{1}{2} - s) \log q)}. \]

Expanding the expression (2.2.3) for \( \varphi_{s+ij\zeta} \) in a straightforward manner, the desired estimate follows from the fact that for all \( y > 0 \): \( \frac{\sinh ny}{\sinh y} \leq n \cosh ny \). \( \Box \)
Proof of the mean ergodic theorem. As noted in §2.1, to show that \( \sigma_n' \) and \( \mu_n \) are mean ergodic sequences it suffices to check that \( \varphi_z(\sigma_n') \) and \( \varphi_z(\mu_n) \) converge to zero for every nontrivial character \( \varphi_z \). This is evident for \( \varphi_z(\sigma_n') \) upon inspection of the preceding formulas, and it follows that it holds for \( \varphi_z(\mu_n) \) also.

Similarly, to check that \( \sigma_{2n} \) and \( \beta_{2n} \) converge strongly to the limit stated in Theorem 1, one first checks that \( \varphi_z(\sigma_{2n}) \) and \( \varphi_z(\beta_{2n}) \) converge to zero for \( z \neq ij\zeta, j = 0, 1 \), using Lemma 2.1. Evaluating \( \varphi_{i\zeta}(\sigma_{2n}) \) using (2.3.4), we see that \( \pi(\sigma_{2n}) \) does indeed converge to the operator \( E_1 + c_{i\zeta}E \), where \( E \) is the projection on the space \( \ker(\pi(\sigma_1) - \gamma(i\zeta)I) \). Here \( c_{i\zeta} = c(\Gamma) = \frac{q-(r-1)^{-1}-h+2}{r(h-1)(1-q^{-1})} \), so that \( c(\Gamma) = 1 \) iff \( \Gamma = \Gamma(r, 2) \) or \( \Gamma = F_k \).

As to the limit of \( \beta_{2n} \), we compute: \( \varphi_{i\zeta}(\beta_{2n}) = \frac{1}{\#B_n} \sum_{k=0}^{n} (\#S_k) \varphi_{i\zeta}(\sigma_k) \). Using (2.2.3) and \( \#S_n = (h-1)rq^{n-1} \) we obtain

\[
\varphi_{i\zeta}(\beta_n) = c_{i\zeta} \left( 1 + (h-1)r \sum_{k=0}^{n-1} (-1)^{k+1} q^k \right) \left( 1 + (h-1)r \sum_{k=0}^{n-1} q^k \right) + A(q)q^{-n}.
\]

The latter expression converges along the subsequence of even indices to

\[
\frac{q-1}{q+1} c(\Gamma),
\]

as stated.

Finally, since \( \varphi_{i\zeta}(\sigma_n) \) and \( \varphi_{i\zeta}(\beta_n) \) do not converge, these sequences are not mean ergodic sequences. \( \square \)

3. Proof of the pointwise ergodic theorem for \( \mu_n \)

3.1. Maximal inequalities and pointwise convergence. Let \( T_n \) be any sequence of operators on \( L^2(X) \). Define \( f^*(x) = \sup_{n \geq 0} |T_n f(x)| \). The following two lemmas are standard (e.g., [P]), and we state them for completeness only:

Lemma 3.1. Let \( V \) be a closed subspace of \( L^2(X) \), and suppose that for a dense set of functions \( f \in V \), \( T_n f(x) \) converges a.e., and that for all \( f \in V \), \( \|f^*\|_2 \leq B\|f\|_2 \). Then \( T_n f(x) \) converges a.e. for all \( f \in V \).

Proof. Fix \( f \in V \) and let \( f_k \to f \) in norm \( (f_k \in V) \), s.t. \( \{T_n f_k(x)\}_{n=0}^{\infty} \) converges a.e. for each \( k \). For any function \( g \in V \) and \( \epsilon > 0 \),

\[
\epsilon^2 m\{x : |T_n g(x)| > \epsilon\} \leq \epsilon^2 m\{x : g^*(x) > \epsilon\} \leq \|g^*\|_2^2 \leq B^2\|g\|_2^2.
\]

Fix \( \epsilon > 0 \), and for \( k \) fixed consider:

\[
|T_n f(x) - T_m f(x)| \leq |T_n(f - f_k)(x)| + |T_m(f - f_k)(x)| + |(T_n - T_m)f_k(x)| \\
\leq 2(f - f_k)^*(x) + |(T_n - T_m)f_k(x)|.
\]
Now \( \limsup_{m,n \to \infty} |(T_n - T_m)f_k(x)| = 0 \) a.e., so that
\[
m\{x : \limsup_{m,n \to \infty} |T_n f(x) - T_m f(x)| > 2\epsilon\} \leq m\{x : (f - f_k)^+ > \epsilon\} \leq \frac{B^2}{\epsilon^2} \|f - f_k\|_2^2 \xrightarrow{k \to \infty} 0.
\]
Hence \( \{T_n f(x)\} \) is a Cauchy sequence a.e.

\textbf{Lemma 3.2.} Suppose that for a dense set of \( f \in L^2(X) \), \( f^* \) is in \( L^2(X) \) and we have \( \|f^*\|_2 \leq B\|f\|_2 \). Then \( f^* \in L^2 \) for all \( f \in L^2(X) \) and the same bound holds.

\textit{Proof.} Let \( f \in L^2(X) \), and let \( f_n \to f \) in norm, where \( \|f_n^*\|_2 \leq B\|f_n\|_2 \). We can choose \( f_n \) so that \( \sum_0^\infty \|f_n - f\|_2 < \infty \). Then \( f_n(x) \to f(x) \) almost everywhere, and also \( T_k f_n(x) \to T_k f(x) \) a.e., for each \( k \). Now let \( M_L g(x) \equiv \sup_{0 \leq k \leq L} |T_k g(x)| \), and then we have \( M_L f_n(x) \to M_L f(x) \) a.e. as \( n \to \infty \), and by Fatou’s Lemma:
\[
\|M_L f^*\|_2 \leq \liminf_{n \to \infty} \|M_L f_n\|_2 \leq \liminf_{n \to \infty} \|f_n^*\|_2 \leq B\liminf_{n \to \infty} \|f_n\|_2 = B\|f\|_2.
\]
Now \( M_L f(x) \to f^*(x) \) a.e. as \( L \to \infty \), so by Fatou’s Lemma once again we get \( \|f^*\|_2 \leq \liminf_{L \to \infty} \|M_L f\|_2 \leq B\|f\|_2 \).

The dense subspace where the maximal inequality will be shown to hold initially is the space of bounded functions. To show that \( \|f^*\|_2 \leq B\|f\|_2 \) for every bounded \( f \) it is enough of course to show this bound for \( 0 \leq f \leq 1 \).

3.2. \textbf{The subadditive maximal inequality.} To obtain the bound, we now turn to

\textit{Proof of Theorem 4.} \( 0 \leq f \leq 1 \) implies \( 0 \leq T_n f \leq 1 \), so that \( 0 \leq f^* \leq 1 \), and therefore \( f^* \in L^2 \) and \( \|f^*\|_2 \leq 1 \). Let \( N : X \to \mathbb{N} \) be a measurable function with finite range, let \( \chi_n \) be the characteristic function of \( \{x \mid N(x) = n \} \), and denote by \( \epsilon_n \) the selfadjoint projection \( f \mapsto \chi_n f \). Denote \( a = ||N||_\infty \), and consider the operator \( (T_N f)(x) = \sum_0^a (\epsilon_n \circ T_n) f(x) \), which is a bounded operator on \( L^2(X) \), with adjoint \( T_N^* = \sum_0^a (T_n \circ \epsilon_n) \). Using the subadditivity of \( T_n \), we have:
\[
\|T_N^* f\|_2^2 = \left\langle T_N^* f , f \right\rangle
\leq \sum_0^a \sum_0^a \left\langle \epsilon_n T_n T_m (\epsilon_m f), f \right\rangle \leq \sum_0^a \sum_0^a C \left\langle \epsilon_n (T_n + T_m) \epsilon_m f, f \right\rangle
\leq C \sum_0^a \left\langle \epsilon_n T_n (\sum_0^a \epsilon_m f), f \right\rangle + C \sum_0^a \left\langle T_m \epsilon_m f, \sum_0^a \epsilon_n f \right\rangle
= C \left\langle T_N f, f \right\rangle + C \left\langle T_N^* f, f \right\rangle = 2C \left\langle T_N f, f \right\rangle \leq 2C \left\langle f^*, f \right\rangle.
\]
Therefore \( \|T_N^*f\|_2^2 \leq 2C \langle f, f^* \rangle \) for all simple functions \( N \). Now fix \( \delta > 0 \) and define:

\[
N'_\delta(x) = \min\{n \mid f^*(x) \leq T_n f(x) + \delta\},
\]

\[a_\delta = \min\{a \mid m\{x \mid N'_\delta(x) > a\} < \delta\}.
\]

Let \( A_\delta = \{x \mid N'_\delta(x) \leq a_\delta\} \), and define the simple function \( N_\delta(x) \) as \( N'_\delta(x) \) for \( x \in A_\delta \), and zero on \( X \setminus A_\delta \). Then \( \chi_{A_\delta}(x)f^*(x) \leq T_{N_\delta}(x)f(x) + \delta \) for almost all \( x \in X \). Note that \( 0 \leq f \leq 1 \) implies \( 0 \leq T_N f \leq 1 \) as well, and we can estimate:

\[
\langle f, f^* \rangle = \left[ \langle f, \chi_{A_\delta} f^* \rangle + \langle f, (1 - \chi_{A_\delta}) f^* \rangle \right]^2 \leq \left[ \langle f, T_{N_\delta} f + \delta \rangle + \delta \right]^2
\]

\[
\leq \left[ \langle f, T_{N_\delta} f \rangle + 2\delta \right]^2 \leq \left[ T_{N_\delta}^* f, f \right]^2 + 4\delta \langle f, 1 \rangle + 4\delta^2
\]

\[
\leq \left\| T_{N_\delta}^* f \right\|_2^2 \|f\|_2^2 + 4(\delta + \delta^2) \leq 2C \left\langle f, f^* \right\rangle \|f\|_2^2 + 4(\delta + \delta^2).
\]

Therefore letting \( \delta \to 0 \) we obtain: \( \langle f, f^* \rangle \leq 2C \|f\|_2^2 \).

Since \( 0 \leq f^* \leq 1 \), we can apply both inequalities to \( f^* \) and obtain (for all simple functions \( N \))

\[
\|T_N^* f^*\|_2^2 \leq 2C \langle f^*, f^{**} \rangle \leq 4C^2 \|f^*\|_2^2.
\]

Consequently, using the function \( N_\delta(x) \) defined above, we obtain

\[
\|f^*\|_2^4 = \langle f^*, f^* \rangle^2 = \left[ \langle f^*, \chi_{A_\delta} f^* \rangle + \langle f^*, (1 - \chi_{A_\delta}) f^* \rangle \right]^2
\]

\[
\leq \left[ \langle f^*, T_{N_\delta} f + \delta \rangle + \delta \right]^2 = \left[ T_{N_\delta}^* f, f \right]^2 + 4\delta \langle f^*, 1 \rangle + 4\delta^2
\]

\[
\leq \left\| T_{N_\delta}^* f \right\|_2^2 \|f\|_2^2 + 4(\delta + \delta^2) \leq 4C^2 \|f^*\|_2^2 \|f\|_2^2 + 4(\delta + \delta^2).
\]

Letting \( \delta \to 0 \), we finally obtain the maximal inequality: \( \|f^*\|_2 \leq 2C \|f\|_2 \). \( \square \)

**Remarks.** (1) We note that a simple modification of the foregoing proof yields the following stronger result. For \( j = 1, 2 \), let \( k_j : \mathbb{N} \to \mathbb{N} \) be two arbitrary functions, and let \( B \in \text{End} L^2(X) \) be an arbitrary bounded operator preserving the cone of nonnegative functions. Assume that for nonnegative bounded functions \( f \) the selfadjoint Markov operators \( T_n \) satisfy:

\[
T_n T_m f(x) \leq C_0 \left( T_{k_1(n)} f(x) + T_{k_2(m)} f(x) \right) + B f(x).
\]

Then \( \|f^*\|_2 \leq C \|f\|_2 \) for all \( f \in L^2 \), where \( C = \max\{2C_0 + 1, \|B\|\} \).

(2) A similar result holds for a 1-parameter family \( \{T_t, t \in \mathbb{R}_+\} \) of selfadjoint Markov operators acting in \( L^2(X) \), provided the maximal function \( \sup_{t \geq 0} |T_t f(x)| \) is well defined, for example, if \( t \mapsto T_t f(x) \) is continuous in \( t \) for almost all \( x \in X \).
3.3. Subadditivity of $\mu_n$. We now show that the sequence $\mu_n$ is subadditive, which will complete the proof of the first half of Theorem 3. Recall $q(\Gamma(r, h)) = (r-1)(h-1)$ and $h(\Gamma(r, h)) = h$.

Lemma 3.3. As functions on $\Gamma(r, h)$:

(a) $\sigma_t * \sigma_s \leq h \sum_{j=0}^{2s} q^{-(s-\frac{j}{2})} \sigma_{t-s+j}$ if $t \geq s$.

(b) $\mu_n * \mu_m \leq 16h(\mu_{2n} + \mu_{2m})$

Proof. (a) Given a word $w$ of length $t-s+j$, $0 \leq j \leq 2s$, $t \geq s$, we count the number of times it can be written as $w = uv$, where $|u| = t$ and $|v| = s$. Assuming that writing $w = uv$ involves $p$ cancellations, and checking parity, we see that $t-p+s-p$ equals $t-s+j$ if $j$ is even, and equals $t-s+j+1$ if $j$ is odd. (Note that $j$ can be odd only if $h(\Gamma) > 2$.) Therefore the number of cancellations that will take place is $s - \frac{j}{2}$ if $j$ is even, and $s - \frac{j+1}{2}$ if $j$ is odd.

If $j$ is even, then the number of representations of $w = uv$ is clearly bounded by the number of words of length $s - \frac{j}{2}$. If $j$ is odd, then after cancelling the last $s - \frac{j+1}{2}$ letters of $u$ against the first $s - \frac{j+1}{2}$ letters of $v$, the last remaining letter of $u$ and the first remaining letter of $v$ are constrained by the requirement that they multiply to give the $\frac{j+1}{2}$-th letter of $w$. This can only happen if they are both taken from the same finite group $G \setminus \{e\} \subset S$.

Consequently, in all cases the number of representations of $w = uv$ in the form $w = uv$, where $|u| = t$ and $|v| = s$, is bounded by $h\#S_s - [\frac{j+1}{2}]$. Recalling that $\#S_n = r(h-1)q^{n-1}$, this bound satisfies:

$$h(\Gamma)\#S_{s-\lceil\frac{j+1}{2}\rceil} \leq hr(h-1)q^{\frac{s}{2}-1}.$$ 

Therefore the coefficient of $\sigma_{t-s+j}$ in $\sigma_t * \sigma_s$ is bounded by

$$\frac{\#S_{t-s+j}}{\#S_t \cdot \#S_s} \cdot hr(h-1)q^{s-\frac{j}{2}-1} \leq hq^{-(s-\frac{j}{2})}.$$ 

(b) Let $n \geq m$ and write:

$$(\sum_{k=0}^{n} \sigma_k) \left( \sum_{k=0}^{m} \sigma_k \right) = \frac{1}{(n+1)(m+1)} \left[ \sum_{k=m+1}^{n} \sigma_k \right]$$

(3.3.1)

$$= \frac{1}{(n+1)(m+1)} \left[ \left( \sum_{k=m+1}^{n} \sigma_k \right) \left( \sum_{k=0}^{m} \sigma_k \right) + 2 \sum_{0 \leq k \leq 0 \leq \ell \leq m} \sigma_t * \sigma_s \right].$$

(i) Consider the first summand in the r.h.s. of (3.3.1). Fix $0 \leq \ell \leq m$ and consider the expressions:

$$\sigma_{m+1} * \sigma_{\ell}, \quad \sigma_{m+2} * \sigma_{\ell}, \quad \sigma_{m+3} * \sigma_{\ell}, \quad \ldots, \quad \sigma_n * \sigma_{\ell}.$$ 

The set of lengths $\{L \mid \sigma_L \text{ has a nonzero coefficient in } \sigma_{m+i} * \sigma_{\ell}\}$ is a subset of the interval $[m+\ell \leq L, m+i+\ell]$. Therefore any given length $L$, $m-\ell \leq L \leq n+\ell$, will appear in at most $2\ell + 1$ intervals. The weight of $\sigma_L$, when it appears, depends on the position of $L$ in the interval, and by part (a) if the distance of
L to the endpoint \( m + i - \ell \) of the interval is \( j \), the weight of \( \sigma_L \) is bounded by \( hq^{-\frac{1}{2}j} \). Each position \( j \leq 2\ell \) occurs at most once, and therefore the weight of \( \sigma_L \) is bounded by \( h \sum_{j=0}^{2\ell} q^{-\frac{1}{2}j} \leq h \sum_{a=0}^{\infty} q^{-\frac{a}{2}} = \frac{h}{1-q^{-\frac{1}{2}}} \). Since \( \ell \) ranges only between 0 and \( m \), we have:

\[
\left( \sum_{k=0}^{n} \sigma_k \right) \left( \sum_{k=0}^{m} \sigma_k \right) \leq \frac{h(m+1)}{1-q^{-\frac{1}{2}}} \sum_{k=0}^{2n} \sigma_k.
\]

(ii) Consider the second summand in the r.h.s. of (3.3.1). Fix a length \( 0 \leq L \leq 2m \). \( \sigma_L \) has a nonzero coefficient in \( \sigma_i \sigma_j \) only if \( t - s + j = L \) for some \( j \). Fixing \( j \), which is the position of \( L \) in the interval \([t-s, t+s]\), we consider the pairs \( t \geq s \) s.t. \( t - s + j = L \). Such a pair necessarily has \( \frac{1}{2} \leq s \) and of course also \( s \leq m \). As \( s \) ranges on this interval the sum of the weights is bounded by \( 2h \sum_{s=\frac{1}{2}}^{m} q^{-\frac{1}{2}(s-j)} \leq \frac{2h}{1-q^{-\frac{1}{2}}} \), according to part (a). To get the total weight we have to sum on all the possibilities for \( j \), but since \( j \leq 2m \) the weight of \( \sigma_L \) is bounded by \( \frac{2h(2m+1)}{1-q^{-\frac{1}{2}}} \).

Putting (i) and (ii) together we see that (3.3.1) can be estimated by (using \( q \geq 2 \)):

\[
\frac{h}{1-q^{-\frac{1}{2}}} \left( \frac{1}{n+1} \right) \left( \frac{1}{m+1} \right) \left[ 2(m+1) \sum_{k=0}^{2m} \sigma_k + (m+1) \sum_{k=0}^{2n} \sigma_k \right] 
\leq 16h \left( \mu_{2m} + \mu_{2n} \right).
\]

Finally we note that since \( A_c(\Gamma(2k, 2)) \) and \( A_c(F_k) \) are isomorphic, we get the same conclusion for \( F_k \). \( \Box \)

3.4. Conclusion of the pointwise theorem for \( \mu_n \). To complete the pointwise ergodic theorem for \( \mu_n \), we need only show that there exists a dense set of functions in \( L^2(X) \) for which \( \pi(\mu_n) f(x) \) converges a.e., and then use Lemma 3.1 and Theorem 4. From now on we assume that the action of \( \Gamma \) on \( X \) is ergodic, and we note that a standard argument using the ergodic decomposition can then be used to give the general case. Now consider the spectrum of the selfadjoint operator \( \sigma_i \) acting on \( L^2(X) \) (where for notational convenience we suppress the explicit reference to the representation \( \pi \)). Let \( E(U) \) denote the spectral projection associated with \( U \subset \text{Sp} \sigma_i \), fix \( 0 < \epsilon < \frac{1}{2} \), and define:

\[
U_\epsilon = \left\{ \gamma(z) \mid \epsilon \leq \text{Re} z \leq \frac{1}{2} \right\}.
\]

Let \( \mathcal{H}_\epsilon = E(U_\epsilon) \mathcal{H} \) be the range of \( E(U_\epsilon) \).

Lemma 3.4. (a) For \( f \in \mathcal{H}_\epsilon \), \( \sum_{n=0}^{\infty} \| \sigma_n f \|_2^2 < \infty \). Therefore \( \sum_{n=0}^{\infty} |\sigma_n f|^2 \) is in \( L^1(X) \), \( \sigma_n f(x) \to 0 \) a.e., and hence also \( \mu_n f(x) \to 0 \) a.e.

(b) \( \sigma'_n f(x) \) and \( \mu_n f(x) \) converge a.e. for \( f \) ranging over a dense subspace of \( L^2(X) \).
Proof. (a) Consider the expression, for \( f \in \mathcal{H} \):

\[
\sum_{n=0}^{\infty} \|\sigma_n f\|_2^2 = \sum_{n=0}^{\infty} \int_{\mathbb{Z}/\mathbb{Z}} |\varphi_z(\sigma_n)|^2 \, dm_f(z).
\]

According to Lemma 2.1, if \( \epsilon \leq s = \text{Re} \, z \leq \frac{1}{2} \) and \( \varphi_z \) is a real character of \( A \), we can certainly estimate by \( |\varphi_z(\sigma_n)| \leq C(q)(n+1)q^{-q\epsilon} \), so that

\[
\sum_{n=0}^{\infty} |\varphi_z(\sigma_n)|^2 \leq C(q)^2 \sum_{n=1}^{\infty} (n+1)^2 q^{-2q\epsilon} \leq C < \infty.
\]

Therefore \( \sum_{n=0}^{\infty} \|\sigma_n f\|_2^2 < \infty \), as stated.

(b) Suppose \( f \perp \bigcup_{\epsilon > 0} \mathcal{H}_\epsilon \). Then clearly \( f \in \text{ker}(\sigma_1 - \gamma(0)I) \oplus \text{ker}(\sigma_1 - \gamma(i\zeta)I) \).

The first component consists of invariant functions, hence constants since we assume that the action is ergodic. Now if \( f \in \text{ker}(\sigma_1 - \gamma(i\zeta)I) \), then \( f \) is an eigenfunction of each \( \sigma_k \) and we compute

\[
\mu_n f = \frac{1}{n+1} \sum_{k=0}^{n} \sigma_k f = \left( \frac{1}{n+1} \sum_{k=0}^{n} \varphi_{i\zeta}(\sigma_k) \right) f.
\]

By (2.3.4)

\[
\frac{1}{n+1} \sum_{k=0}^{n} \varphi_{i\zeta}(\sigma_k) = \frac{1}{n+1} \left[ c_{i\zeta}((-1)^n + 1) + c_{1-i\zeta}(-1)^n \sum_{k=0}^{n} q^{-k} \right] \rightarrow 0 \quad n \rightarrow \infty.
\]

Therefore \( (\mu_n f)(x) = \varphi_{i\zeta}(\mu_n f)(x) \rightarrow 0 \), and so \( \mu_n f(x) \) converges for a dense set of functions in \( L^2(X) \). Clearly the same holds for \( \sigma_n' f(x) \). \( \square \)

To conclude the proof of the pointwise ergodic theorem for \( \mu_n \) we remark that since by \( \S 2 \), \( \mu_n f \rightarrow E_1 f \) in norm, \( E_1 f \) is necessarily the pointwise limit a.e. of \( \mu_n f(x) \). \( \square \)

4. The pointwise ergodic theorem for \( \sigma_n' \)

4.1. Reduction to a spectral problem. Recall that for \( f \in L^2(X) \), \( M f(x) = \sup_{n \geq 0} |\sigma_n' f(x)| \) and \( f'\prime(x) = \sup_{n \geq 0} |\mu_n f(x)| \). By Lemmas 3.1, 3.2, and 3.4 it is enough to show:

Lemma 4.1. Suppose \( f \), \( M f \in L^2(X) \). Then \( \|M f\|_2 \leq B(\Gamma)\|f\|_2 \).

Proof. Define \( \mu'_n = \frac{1}{n+1} \sum_{k=0}^{n} \sigma'_k \). Clearly \( \mu'_n \) satisfies the strong \( L^2 \) maximal inequality, since \( \mu'_n \leq 2\mu_{n+1} \). Now using the square-function method of E. M. Stein [S], the maximal inequality for \( \sigma_n' \) will be deduced from the maximal inequality for \( \mu_n' \), as follows: By Abel's summation formula,

\[
\sigma'_n - \frac{1}{n+1} \sum_{k=0}^{n} \sigma'_k = \frac{1}{n+1} \sum_{k=1}^{n} k(\sigma'_k - \sigma'_{k-1}).
\]
Therefore
\[ \left| \sigma'_n f(x) - \mu'_n f(x) \right|^2 \leq \frac{1}{(n+1)^2} \left( \sum_{k=1}^{n} \sqrt{k} \cdot \sqrt{k} (\sigma'_k - \sigma'_{k-1}) f(x) \right)^2 \]
\[ \leq \frac{1}{(n+1)^2} \left( \sum_{k=1}^{n} k \right) \left( \sum_{k=1}^{n} k (\sigma'_k - \sigma'_{k-1}) f(x) \right)^2. \]

Since \( \sum_{k=1}^{n} k \leq (n+1)^2 \), taking the square root we can write
\[ M f(x) = \sup_{n \geq 0} \left| \sigma'_n f(x) \right| \leq \sup_{n \geq 0} \left| \mu'_n f(x) \right| + R(f)(x), \]
where \( R(f)(x)^2 = \sum_{k=1}^{\infty} k \left| (\sigma'_k - \sigma'_{k-1}) f(x) \right|^2 \). Consequently \( \| M f \|_2 \leq 2 \| f \|_2 \) and to prove the maximal inequality it is enough to show that \( \| R(f) \|_2 \leq C \| f \|_2 \). By the spectral theorem,
\[ \| R(f) \|_2^2 = \sum_{k=1}^{\infty} k \| \sigma'_k f - \sigma'_{k-1} f \|_2^2 = \int_{Z} \sum_{k=1}^{\infty} k \left| \varphi_z(\sigma'_k) - \varphi_z(\sigma'_{k-1}) \right|^2 \, dm_f(z). \]

Therefore it suffices to prove that \( \psi(z) = \sum_{k=0}^{\infty} k \left| \varphi_z(\sigma'_k) - \varphi_z(\sigma'_{k-1}) \right|^2 \) is a bounded function on the real spectrum of \( A \), and then \( \| R(f) \|_2 \leq \| \psi \|_{1,\infty} \| f \|_2 \). For the sequel, note that \( \sigma'_k - \sigma'_{k-1} = \frac{1}{2} (\sigma_{k+1} - \sigma_{k-1}) \).

4.2. Spectral estimates. To bound \( \psi \) we consider the expression for \( \varphi_z \) given in §2.3. We distinguish three cases:

1. Principal strip: \( \frac{1}{4} \leq \text{Re} z \leq \frac{1}{2} : \)
\[ \psi(z) = \frac{1}{4} \sum_{k=1}^{\infty} k \left| \varphi_z(\sigma_{k+1}) - \varphi_z(\sigma_{k-1}) \right|^2 \leq \frac{1}{4} \sum_{k=1}^{\infty} k \left( \left| \varphi_z(\sigma_{k+1}) \right| + \left| \varphi_z(\sigma_{k-1}) \right| \right)^2 \]
\[ \leq C(q)^2 \sum_{k=1}^{\infty} (k+1)^3 q^{-k-1} < \infty. \]

2. Complementary series, excluding endpoints: \( z = s + ij\xi, \ 0 < s < \frac{1}{4}, \ j = 0, 1 \). By (2.3.3):
\[ \psi(z) = \frac{1}{4} \sum_{k=1}^{\infty} k \left| c_{s+i\xi}(q^{-s} - q^s) q^{-ks} + c_{1-s-i\xi}(q^{-(1-s)} - q^{(1-s)}) q^{-k(1-s)} \right|^2 \]
\[ \leq \max_{0 \leq s \leq \frac{1}{4}} \left( \left| c_{s+i\xi} \right| + \left| c_{1-s-i\xi} \right| \right) \sum_{k=1}^{\infty} k \left( (q^s - q^{-s}) q^{-ks} + (q + 1) q^{-k(1-s)} \right)^2. \]

Recalling that \( 0 < s < \frac{1}{4} \), so that \( q^{-k(1-s)} \leq q^{-\frac{3k}{4}} \), we estimate \( \psi(z) \) as follows:
\[ a_1(q)(q^s - q^{-s})^2 \sum_{k=1}^{\infty} k(q^{-2s})^{k-1} + a_2(q) \sum_{k=1}^{\infty} k q^{-ks} q^{-\frac{3k}{4}} + a_3(q)^2 \sum_{k=1}^{\infty} k q^{-\frac{k}{4}} \]
\[ \leq a_4(q) \frac{(q^s - q^{-s})^2}{(1 - q^{-2s})^2} + a_5(q). \]
This expression is bounded independently of $s$, $0 < s \leq \frac{1}{4}$.

(3) Endpoints of the complementary series, namely $z = ij\zeta$, $j = 0, 1$. By (2.3.4):

$$|\varphi_{ij}(\sigma_{k+1}) - \varphi_{ij}(\sigma_{k-1})| \leq 2|c_{1-i\zeta}|^{-(k-1)}.$$ 

Therefore $\psi(z)$ is bounded on the real spectrum of $A$, and the pointwise ergodic theorem for $\sigma_n'$ follows.

**Completion of the proof of Theorem 2.** First let us consider the groups $\Gamma(r, 2)$ and $F_k$ and establish the convergence of $\sigma_{2n}$ to a conditional expectation. In this case $\gamma(i\zeta) = -1$, $\varphi_{i\zeta}(\sigma_n) = (-1)^{\gamma}$, and we claim that:

$$\ker(\sigma_1 + I) = \left\{ f \in L^2(X) \mid f(\gamma x) = (-1)^{\gamma} f(x), \quad \forall \gamma \in \Gamma \right\}.$$

Indeed, $\sigma_1 f = -f$ implies

$$\|f\| \leq \frac{1}{\#S} \left\| \sum_{s \in S} \pi(s)f \right\|.$$ 

By convexity, it follows that $|f(x)| = |f(sx)|$ a.e. for $s \in S$, and so necessarily $f(sx) = -f(x)$ a.e. for $s \in S$. Assuming the action is ergodic, $f$ is bounded and so $\ker(\sigma_1 - I) \oplus \ker(\sigma_1 + I)$ is a $\Gamma$-invariant function algebra. Hence it is of the form $L^2(X, \mathcal{B}_0)$, where $\mathcal{B}_0$ is the $\sigma$-algebra corresponding to a two-point factor $\alpha : X \to \{\pm 1\}$. The action of $\Gamma$ on $\{\pm 1\}$ is via multiplication by $(\pm 1)^{\gamma}$, namely: $\gamma(\pm 1) = (\pm 1)^{\gamma}(\pm 1)$.

Denote by $\mathcal{H}_{\pm 1}$ the subspace that realizes the characters $\varphi_0, \varphi_{i\zeta}$. Note that $\sigma_n \leq 2\sigma_n'$ satisfies a maximal inequality for nonnegative $L^2$ functions, and hence on all of $L^2(X)$, by Lemma 3.2. Clearly $\sum_{n=0}^{\infty} \|\sigma_{2n}f\|^2 < \infty$ for $f$ in a dense subset of the orthogonal complement of $\mathcal{H}_{\pm 1}$, so that indeed $\sigma_{2n}f$ converges a.e. for $f$ in that subset. Therefore, since $\sigma_{2n}$ satisfies a maximal inequality, Lemma 3.1 implies pointwise convergence of $\sigma_{2n}f(x)$ for all $f \perp \mathcal{H}_{\pm 1}$. The convergence is to the right limit, namely zero, which is the projection on $\mathcal{H}_{\pm 1}$, since $\sigma_{2n}$ converges in norm to that limit (by §2.3). Noting that $\sigma_{2n}$ acts as the identity on $\mathcal{H}_{\pm 1}$, it follows that $\sigma_{2n}$ converges to the conditional expectation w.r.t. a $\Gamma$-invariant $\sigma$-algebra, which arises from a two-point factor if the action is ergodic.

Finally, we note that the existence of a maximal inequality for the sequence $\sigma_n$ in the space of nonnegative $L^2$ functions implies one for $\beta_n$, on all of $L^2(X)$ by Lemma 3.2. Moreover, $\beta_{2n}$ converges pointwise on a dense set of functions in the orthogonal complement of $\mathcal{H}_{\pm 1}$, by the spectral estimates of Lemma 2.1, and so it follows from Lemma 3.1 that $\beta_{2n}f$ converges almost everywhere for all $f \perp \mathcal{H}_{\pm 1}$. Using the description of the norm limit of $\beta_{2n}f$ contained in Theorem 1 concludes the proof of Theorem 2. □

**5. Convolution algebras and double cosets algebras**

**5.1. Cayley graphs.** Given a finitely generated group $\Gamma$ with a finite symmetric generating set $S$, let $X = X(\Gamma, S)$ denote the Cayley graph of $\Gamma$ determined by
\( S \). \( X \) has vertex set \( \Gamma \), and \( (\gamma', \gamma) \) are connected by an (undirected) edge iff \( \gamma^{-1}\gamma' \in S \) (we assume \( e \notin S \)). The distance \( d(\gamma, \gamma') = |\gamma^{-1}\gamma'|_S \) coincides with the edge path metric on \( X \), and each \( \gamma \in \Gamma \) acts, by left multiplication, as an isometry of \( X \) with the metric \( d \) or, what is the same, as a graph automorphism.

The group \( G = \text{Aut}(X) \) of all isometries/ggraph automorphisms is a Hausdorff locally compact second countable group with a neighbourhood basis at \( \iota \) given by the (open compact) subgroups \( K_n = \{ g \in G \mid gx = x \ \forall x \in B_n(e) \} \), where \( B_n(x) = \{ y \mid d(x, y) \leq n \} \). Denote by \( C_c(X) \) the space of finitely supported functions on (the vertex set \( \Gamma \) of the graph) \( X \). Define the two natural representations

\[
\lambda : G \to \text{End} \ C_c(X), \quad \rho : \Gamma \to \text{End} \ C_c(X)
\]

by \( \lambda(g)f(x) = f(g^{-1}x) \) and \( \rho(\gamma)f(x) = f(x\gamma) \). (Note that the notation \( g^{-1}x \) when \( g \in \Gamma \subset G \) denotes, indeed, left multiplication.) Extend \( \rho \) to a representation of \( \ell^1(\Gamma) \), and observe that \( \sigma_n \in \ell^1(\Gamma) \) is mapped to the operator given by:

\[
\rho(\sigma_n)f(x) = \frac{1}{|S_n|} \sum_{w \in S_n} f(xw) \quad \text{averaging } f \text{ on spheres of radius } n.
\]

The action of \( G \) on \( X \) is isometric, and therefore each operator \( \rho(\sigma_n) \) commutes with the operators \( \lambda(g), g \in G \). Recall \( A_\epsilon(\Gamma) = \{ \sigma_n, n \geq 0 \} \), and denote \( \text{End}_G C_c(X) = \{ R : C_c(X) \to C_c(X) \mid \lambda(g)R = R\lambda(g), \forall g \in G \} \). Then \( \rho : A_\epsilon(\Gamma) \to \text{End}_G C_c(X) \) is an algebra homomorphism, and it is injective. To see that, let \( \mu \mapsto \mu^* \) denote the involution \( \mu^* (\gamma) = |\mu(\gamma^{-1})| \) on \( \ell^1(\Gamma) \), and now:

\[
(\rho(\mu)\mu^*)(e) = \mu^* \mu(e) = \sum_{\gamma \in \Gamma} |\mu(\gamma)|^2.
\]

5.2. Operators associated with \( G \)-invariant random walks. There is another natural construction of operators on \( C_c(X) \) which commute with the operators \( \lambda(g), g \in G \). Denote by \( F_c(G) \) the space of compactly supported locally constant functions on \( G \). Note that \( G \supseteq \Gamma \), so that \( G \) is transitive on \( X \), and let \( K = G_e \) denote the subgroup of \( G \) stabilizing the point \( e \in X(\Gamma, S) \). Then \( j : ge \to gK \) implements a \( G \)-equivariant isometry \( j : X \to G/K \). Let \( x \in F_c(G) \) be left \( K \)-invariant: \( x(kg) = \chi(g), \forall k \in K \). Then we can define an operator \( R_x \) on \( C_c(G/K) \cong C_c(X) \) by the formula:

\[
(R_x f)(g_0K) = \int_G f(g_0gK)\chi(g) \, dm(g),
\]

where \( m \) is a fixed left Haar measure on \( G \). Normalize \( m \) by taking \( m(K) = 1 \), and note that any Haar measure is \( K \)-invariant, and therefore \( R_x \) is well defined:

\[
\int_G f(g_0kgK)\chi(g) \, dm(g) = \int_G f(g_0hK)\chi(k^{-1}h) \, dm(h) = (R_x f)(g_0K).
\]

Clearly \( R_x \) commutes with the operators \( \lambda(g) \) induced by the left translations \( g : g_0K \to g_0gK \) on \( G/K \). It is easily seen that if \( \chi_1 \) and \( \chi_2 \) are \( K \)-invariant from the left and the right, then so is their product \( \chi_1 \star \chi_2 \) in
HARMONIC ANALYSIS AND POINTWISE ERGODIC THEOREMS

\[ F_c(G). \] Hence \( C_c(G, K) = \{ f \in F_c(G) \mid f(kgk') = f(g), \forall k, k' \in K \} \) is a subalgebra of \( C_c(G) \), the convolution algebra of double cosets of \( K \). Clearly \( R_{x_1 \ast x_2} = R_{x_1} \circ R_{x_2} \), and therefore the map \( R : C_c(G, K) \to \text{End}_G C_c(G/K) \), given by \( \chi \mapsto R_{\chi} \), is a representation of the double cosets algebra. \( R \) is injective, since viewing \( \chi \in C_c(G, K) \) as a function on \( G/K \), we have: \( R_{\chi}(\gamma)(e) = \int_G \chi(g)\overline{\chi(g)} dm(g) \).

Now fix \( n \geq 0 \), and suppose \( K \) has \( N_n \) orbits in the sphere \( S_n(e) \subset X \), denoted \( \{ \theta_i^{(n)} \mid 1 \leq i \leq N_n \} \). Let \( \chi_U \) denote the characteristic function of a set \( U \), and put \( U_i^{(n)} = \{ g \in G \mid g \cdot e \in \theta_i^{(n)} \} \). Note that \( U_i^{(n)} \) is the double coset \( KgK \), where \( g \cdot e \in \theta_i^{(n)} \). Recall \( m(K) = 1 \).

**Lemma 5.1.** (a) The Haar measure of the double coset \( KgK \) is the number of elements in the orbit of \( g \cdot e \) under \( K = G_e \).

(b) Let \( \chi_n = \frac{1}{\#S_n} \sum_{i=1}^{N_n} \chi_{U_i^{(n)}} \). Then \( R_{\chi_n} \) is the operator of averaging a function on a sphere of radius \( n \), namely \( R_{\chi_n} = \rho(\sigma_n) \).

**Proof.** (a) It is easy to check that \( KgK \) is the disjoint union of the sets \( kgK \) as \( k \) ranges over a set of representatives of the left cosets of \( K \cap gKg^{-1} \) in \( K \). Hence for \( g \cdot e \equiv x \in X = G/K \), we obtain \( m(KgK) = \#(K : K \cap gKg^{-1}) = \#K \cdot x \).

(b) \( \sum_{i=1}^{N_n} \int_G f(gK)\chi_{U_i^{(n)}}(g) dm(g) \) equals the sum of the values of \( f \) on the sphere of radius \( n \) in \( X(\Gamma, S) = G/K \) with center \( e = [K] \), by definition.

**Proof of Proposition 5.** \( R, \rho \) are injective and since by Lemma 5.1(b) the image under \( \rho \) of the algebra \( A_c(\Gamma) \) is contained in the image of \( C_c(G, K) \) under \( R \), it follows that the map given by

\[ \psi : R^{-1} \circ \rho : A_c(\Gamma) \to C_c(G, K) \]

is an injection. Using Lemma 5.1(a), we see that \( \psi \) is isometric in the \( \ell^1 \)-norm on both sides. Finally, \( \psi \) commutes with the natural involutions \( \mu^*(\gamma) = \overline{\mu(\gamma)} \) and \( f^*(KgK) = f(Kg^{-1}K) \), since \( R \) and \( \rho \) are \( \ast \)-representations.

5.3. Complemented lattices. Consider the group of automorphisms \( G = G(r_1, r_2) \) of the semi-homogeneous tree \( T = T(r_1, r_2) \) of valencies \( r_1, r_2 \), where \( r_i \geq 2 \), \( i = 1, 2 \), and \( r_1 + r_2 > 4 \). Denote the stabilizer of a vertex \( v \) by \( K = G_v \), and note that \( K \) is transitive on each sphere with centre \( v \), and so it follows easily that the algebra \( C_c(G, K) \) is isomorphic to the algebra generated in \( \text{End} \ell^2(T) \) by the operators of averaging a function on a sphere of radius \( n \geq 0 \). Now assume \( \Gamma \subset G \) is a lattice complementary to \( K \), namely \( G = \Gamma K \). Recall the length function on \( \Gamma \) defined by \( \ell(\gamma) = \frac{1}{2}d(v, \gamma v) \) if \( r_1 \neq r_2 \), and \( \ell(\gamma) = d(v, \gamma v) \) otherwise, where \( d \) is the distance in \( T \). Let \( \Gamma_0 \) denote the finite group \( \Gamma \cap K \). The sets \( S_\ell(\ell) = \{ \gamma \in \Gamma \mid \ell(\gamma) = n \} \) are double cosets of \( \Gamma_0 \) in \( \Gamma \). We denote by \( A_c(\Gamma, \ell) \) the algebra generated by the elements
\[ \sigma_n(\ell) = \frac{1}{\#S_n(\ell)} \sum_{w \in S_n(\ell)} w, \quad \text{where } n \geq 0. \]

Endow the algebra with the involution inherited from \( \ell^1(\Gamma) \). We can now state:

**Lemma 5.2.** The map \( \sigma_n(\ell) \mapsto \sigma_n \) extends uniquely to an isometric \(*\)-isomorphism between the algebras \( A_c(\Gamma, \ell) \) and \( C_c(G, K) \).

**Proof.** There is a natural identification \( \tau : G/K \to \Gamma/\Gamma_0 \), which is equivariant w.r.t. the action of \( \Gamma \) on both sides. Each double coset \( U \) of \( \Gamma_0 \) in \( \Gamma \) defines an operator \( P_U \) on \( \ell^2(\Gamma/\Gamma_0) \) commuting with the left action of \( \Gamma \), as we observed in the previous section. Moreover \( \tau : \Gamma/\Gamma_0 \mapsto P_U \) is a faithful representation of the algebra of double cosets of \( \Gamma_0 \), as before. The operator corresponding to \( \sigma_n(\ell) \) is the operator of averaging a function on spheres of radius \( n \) defined by the length function \( \ell \). Using the identification \( \tau \), we see that \( R^{-1} \circ \rho : A_c(\Gamma, \ell) \to C_c(G, K) \) is an isometric isomorphism, commuting with \(*\), since both algebras have a faithful \(*\)-representation onto the algebra generated in \( \text{End} \ell^2(T) \) by the operators of sphere averaging. \( \square \)

**Proof of Theorem 6.** According to Lemma 5.2, the proof of Theorem 6 amounts to showing that the algebra \( C_c(G, K) \), where \( G \) is the group of automorphisms of a semi-homogeneous tree, is \(*\)-isometrically isomorphic to one of the algebras \( A_c(\Gamma, \ell) \) that figured in Theorems 1 and 2. If the tree \( T \) is regular, then it is self-evident that \( C_c(G(r, r), K) \cong A_c(\Gamma(r, 2)) \). Generally, \( C_c(G(r, h), K) \cong A_c(\Gamma(r, h)) \), where \( K \) is the stabilizer of a vertex of valency \( r \). This fact is well known (e.g., [BK, I-P, N1]), and can be easily established as follows: Consider the Cayley graph \( X(\Gamma(r, h), S) \), in which each vertex is the intersection of \( r \) uniquely determined subgraphs of \( X \), each of them a complete graph on \( h \) vertices (without loops). These subgraphs correspond to the right cosets of the groups \( G_i, 1 \leq i \leq r \), and let us say that the vertex in question is related to these subgraphs. For each complete subgraph on \( h \) vertices, add a new vertex, and connect it by new edges to all the vertices of \( X \) which are related to the subgraph. Erasing all the original edges, we obtain a semi-homogeneous tree of valencies \( r \) and \( h \). Note that by construction, the vertices of \( X \) have a natural one-to-one correspondence with the set of vertices of \( T \) at an even distance from a fixed vertex \( v \in T \) of valency \( r \), which corresponds to \( e \in X \). The group of automorphisms of the Cayley graph is naturally identified with \( G(r, h) \). The stabilizer of the vertex \( e \in X \) is identified with the stabilizer \( K \) of a vertex \( v \) of valency \( r \) in \( T(r, h) \). According to Lemma 5.1, \( A_c(\Gamma(r, h)) \) injects \(*\)-isometrically into \( C_c(\text{Aut}(X), K) \), and hence into \( C_c(G(r, h), K) \). Recall that \( K \) acts transitively on each sphere of radius \( n \) with center \( v \). Therefore the orbits of \( K \) in \( G/K \) are in one-to-one correspondence with spheres of radius \( 2n \) in \( T \) centered at \( v \) (when \( r \neq h \)). Hence the injection is onto and the algebras are \(*\)-isometric. \( \square \)

**Remark.** A similar argument applies in the case that \( \Gamma K = G'(r, r) \), where \( G'(r, r) \) is the subgroup of \( \text{Aut}T(r, r) \) that acts without inversions.

**Proof of Theorem 7.** First we note that if \( H \subset G(r_1, r_2) \) is a closed noncompact subgroup which is transitive on the boundary of \( T(r_1, r_2) \) (namely the space of
ends), then $H$ has either one or two orbits in the vertices of $T$ [N1, Proposition 3]. Moreover, the stabilizer $H_v$ in $H$ of a vertex $v \in T$ is transitive on the boundary, and hence transitive on each sphere with center $v$ [N1, Proposition 3]. Therefore, the orbits of $K$ in $H/K$ are in one-to-one correspondence with spheres of radius $i$ with center $v$, where $i$ is the number of orbits of $H$ in $T$. Hence, the algebra $C_c(H, H_v)$ is naturally identified with $A_c(\Gamma(r_1, r_2))$, by the map $\sigma_n(H) \mapsto \sigma_n$. The only exception occurs when the tree is regular, and $H$ has two orbits of vertices. In this case, a proper subalgebra of $C_c(G, K)$ is obtained, generated by double cosets associated with spheres of even radius around $v$. If the valency of the tree is $r$, then this subalgebra is of course isomorphic to the subalgebra generated by $A_c(r(r_1, r_2))$, and the map above takes the form $\sigma_n(H) \mapsto \sigma_{2n}$.

All the continuous real characters of the subalgebra are inherited from $A(\Gamma(r, 2))$, as is readily verified using the method described in §2.2, or [BK, §2.9]. The special character of $A(\Gamma(r, 2))$ restricts to the trivial character on the subalgebra generated by $A_2$. However the space of vectors invariant under $\sigma_2$ coincides with the space of $H$-invariant vectors. This follows since $\sigma_2 f = f$ implies $hf = f$ for almost every $h$ in the support of $\sigma_2$, and such a set generates $H$. Consequently, $\sigma_n(H)$ does indeed converge to the projection on the space of $H$-invariant vectors. This concludes the proof of Theorem 7.  

6. THE NONCOMMUTATIVE ALGEBRA $C_c(G, B)$

6.1. Structure and representations of $C_c(G, B)$. In this section we will consider the convolution algebra of double cosets of $B$ in $G$, where $G = \text{Aut} T(r_1, r_2)$ is the group of automorphisms of a semi-homogeneous tree with $r_1 > r_2 \geq 2$, and $B$ is the compact subgroup stabilizing a geometric edge. As noted in §1.4, it will prove convenient to pass from $C_c(G, B)$ to an isomorphic algebra, which we now describe. Let $\Gamma = G_1 * G_2$, where $G_i$ are two finite groups with $|G_i| = r_i \geq 2$, $r_1 > r_2$. Fix the generating set $S = (G_1 \cup G_2) \setminus \{e\}$, and construct the Cayley graph $X(\Gamma, S)$. Let us note the following: Each word $e \neq w \in \Gamma$ has a unique reduced decomposition $w = u_1 u_2 \ldots u_n$, where $u_i \in G_i \setminus \{e\}$ and two consecutive letters do not belong to the same subgroup $G_i$. Let $S^{(i)}_n = \{w \in \Gamma : |w| = n, \text{ the first letter of } w \text{ is in } G_i\}$, $n \geq 1$, $i = 1, 2$. Put $\sigma^{(i)}_n = \frac{1}{\#S^{(i)}_n} \sum_{w \in S^{(i)}_n} w$, $\sigma^{(1)}_0 = \sigma^{(2)}_0 = e$. Let $A'_c(\Gamma)$ denote the convolution algebra generated in $\ell^1(\Gamma)$ by the elements $\sigma^{(i)}_n$, $n \geq 0$, $i = 1, 2$. Clearly $A'_c(\Gamma)$ contains the algebra $A_c(\Gamma)$ generated by the spheres $\sigma_n$, $n \geq 0$. We now have the following:

**Lemma 6.1.** The algebra $A'_c(\Gamma)$ is $\ast$-isometrically isomorphic to the double cosets algebra $C_c(G, B)$.

**Proof.** Consider the dual graph of $T(r_1, r_2)$: By definition, its vertices are the edges of $T$ and two such vertices are connected by an edge iff they have a vertex of $T$ in common. This graph is naturally isomorphic to the Cayley graph $X(\Gamma, S)$ and so $G$ is clearly isomorphic with $\text{Aut}(X(\Gamma, S))$ (e.g., [N1, §8]). The isomorphism identifies the stabilizer of a vertex in $X$ with the stabilizer
of an edge of $T$. Now fix an edge $y_0$ of $T$. The orbits $\theta(i, n)$ of $B = G_{y_0}$ in the set of edges are parametrized by $(i, n)$, $n \geq 0$, $i = 1, 2$. Here, given an edge $y$, $n + 1$ is the number of edges in the unique path whose first edge is $y_0$ and last edge is $y$, and $i = 1, 2$ according to whether the first vertex of the path has valency $r_1$ or $r_2$. (We define $\theta(1, 0) = \theta(2, 0) = \{y_0\}$.)

The isometric isomorphism $\psi : A'_c(\Gamma) \to C_c(G, B)$ is implemented as follows:

Define $U(i, n) \triangleq \{g \in G \mid gy_0 \in \theta(i, n)\}$. Then $U(i, n)$ is a double coset of $B$, and we define: $\psi(\sigma_n^{(i)}) = \frac{1}{m(U(i, n))} \chi_{U(i, n)} \triangleq \chi_n^{(i)}$. Here we normalize by $m(B) = 1$ and $m(U(i, n)) = [B : B_{i,n}]$ the index of the stabilizer (in $B$) of $y \in \theta(i, n)$. The arguments of §5.2 show that $\psi$ is an isometric $*$-isomorphism $\psi : A'_c(\Gamma) \to C_c(G, B)$. \(\square\)

Let us denote by $C(G, B)$ and $A'_c(\Gamma)$ the closures, in the $\ell^1$-norm, of $C_c(G, B)$ and $A'_c(\Gamma)$, respectively. Now recall the following fact [Ma, §2.1.13]:

**Lemma 6.2.** All the algebras $C(G, B) = A'_c(\Gamma)$ are $*$-isometrically isomorphic to the algebra $\ell^1(\mathbb{Z}_2 * \mathbb{Z}_2)$.

**Proof.** Consider the following elements in $A'_c(\Gamma)$:

\[
(6.1.1) \quad \alpha_i = \frac{1}{r_i} \left( (r_i - 1) \sigma_1^{(i)} + \sigma_0 \right) = \frac{1}{\# G_i} \sum_{g \in G_i} g.
\]

The $\alpha_i$ are two (noncommuting) idempotents of norm 1, and clearly $e, \alpha_1, \alpha_2$ generate $A'_c(\Gamma)$. Now $\ell^1(\mathbb{Z}_2 * \mathbb{Z}_2)$ is the $\ell^1$-closure of the free algebra on the two noncommuting idempotents $\eta_i = \frac{1}{2}(e + x_i)$, $x_i^2 = e$. It is readily verified that the elements $e, \alpha_1 \alpha_2 \cdots \alpha_i$, $i_j \in \{1, 2\}, i_j \neq i_{j+1}, n \geq 1$, are all linearly independent in $\ell^1(\mathbb{Z}_2 * \mathbb{Z}_2)$, and their linear span coincides with $A'_c(\Gamma)$. The same is true of $\eta_1 \eta_2 \cdots \eta_n$, $i_j \in \{1, 2\}, i_j \neq i_{j+1}, n \geq 1$, and hence $\eta_i \mapsto \alpha_i$ extends to an isometric $*$-isomorphism $\ell^1(\mathbb{Z}_2 * \mathbb{Z}_2) \to A'_c(\Gamma) = C(G, B)$. \(\square\)

It is now immediate to classify the irreducible $*$-representations of $A'_c(\Gamma)$, since they are in one-to-one correspondence with the irreducible unitary representation of $\mathbb{Z}_2 * \mathbb{Z}_2$. These are classified as follows:

**Lemma 6.3.** The irreducible unitary representations of $\mathbb{Z}_2 * \mathbb{Z}_2 = \langle x_1 \rangle \ast \langle x_2 \rangle$ are given by:

(a) $x_1 \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $x_2 \mapsto \begin{pmatrix} 0 & z \\ z^{-1} & 0 \end{pmatrix}$, $z = e^{i\phi}$, $0 < \phi < \pi$.

(b) $x_1 \mapsto \pm 1$, $x_2 \mapsto \pm 1$.

**Proof.** Note that the subgroup generated by $x_1, x_2$ is infinite cyclic and of index 2. \(\square\)

6.2. **Proof of the mean ergodic theorem for $\Gamma = G_1 \ast G_2$.** Given a unitary representation $\pi$ of $\Gamma$ in $\mathcal{H}$ and a unit vector $v \in \mathcal{H}$, the function $\mu \mapsto \langle \pi(\mu)v, v \rangle$ is a normalised positive definite function on $A(\Gamma)$, hence a linear
combination of normalised extreme positive definite functions \( \langle \pi_z(\mu) v_z, v_z \rangle \) associated with irreducible *-representations of \( A(\Gamma) \), namely \( \langle \pi(\mu) v, v \rangle = \int \langle \pi_z(\mu) v_z, v_z \rangle \, dm_v(z) + \sum_\delta \langle \pi(\mu) E_\delta v, E_\delta v \rangle \). Here \( m_v \) is the measure determined by \( v \) on the continuous spectrum of \( A(\Gamma) \), \( \delta \) ranges over the four \( 1 \)-dimensional representations of \( A(\Gamma) \), and \( E_\delta \) is the projection on the subspace of \( \mathcal{H} \) that affords the representation \( \delta \). To establish the mean ergodic theorem for a sequence \( \nu_n \) it suffices to show that \( \|\pi(\nu_n)\| \to 0 \) for \( \pi \neq 1 \).

To analyze the sequence \( \sigma_n \) we note first that by (6.1.1):

\[
(6.2.1) \quad \sigma^{(i)}_1 = \frac{1}{r_i - 1}(r_i \alpha_i - \sigma_0), \quad \sigma^{(i)}_2 = \sigma^{(i)}_1 \sigma^{(3-i)}_1.
\]

Also, by definition, for \( n \geq 1 \):

\[
(6.2.2) \quad \sigma^{(i)}_{2n} = \left( \sigma^{(i)}_1 \sigma^{(3-i)}_1 \right)^n = \left( \sigma^{(i)}_2 \right)^n, \quad \sigma^{(i)}_{2n+1} = \sigma^{(i)}_{2n} \cdot \sigma^{(i)}_1 \cdot (3-i).
\]

Now note that \( \sigma_n, n \geq 0 \), has the following form:

\[
(6.2.3) \quad \sigma_{2n} = \frac{1}{2}(\sigma^{(1)}_{2n} + \sigma^{(2)}_{2n}), \quad \sigma_{2n+1} = \frac{(r_1 - 1)\sigma^{(1)}_{2n+1} + (r_2 - 1)\sigma^{(2)}_{2n+1}}{r_1 + r_2 - 2}.
\]

The last formula shows that in order to establish the mean ergodic theorem it is enough to show that the norm of each of the matrices \( \pi_z(\sigma_1^{(i)}), \pi_\delta(\sigma_2^{(i)}) \), where \( \delta \neq 1 \) and \( i = 1, 2 \), is strictly less than 1. Then it follows immediately from (6.2.3) that \( \|\pi(\sigma_n)\| \) converges to zero exponentially as \( n \to \infty \) for \( \pi \neq 1 \).

Using the isomorphism of Lemma 6.2, and evaluating the representations described in Lemma 6.3 (recalling that \( \alpha_i = \frac{1}{2}(e + x_i) \)), we obtain:

\[
(6.2.4) \quad \pi_z(\sigma^{(1)}_1) = \frac{1}{2} \left( \begin{array}{cc} r_1 - 2 & 0 \\ 1 & 1 \end{array} \right) I + \frac{1}{2} \left( \begin{array}{cc} r_1 & 1 \\ 1 & 0 \end{array} \right) \equiv A,
\]

\[
(6.2.5) \quad \pi_z(\sigma^{(2)}_1) = \frac{1}{2} \left( \begin{array}{cc} r_2 - 2 & 0 \\ 1 & z \end{array} \right) I + \frac{1}{2} \left( \begin{array}{cc} z & 0 \\ 0 & 1 \end{array} \right) \equiv A_z,
\]

\[
(6.2.6) \quad \pi_\delta(\sigma^{(i)}_1) = -\frac{1}{r_i - 1} \quad \text{iff} \quad \delta(x_i) = -1.
\]

The matrices whose norm should be estimated are \( AA_z \) and \( A_z A = (AA_z)^* \). We note the following:

(1) The matrix \( A_z \), \(|z| = 1\), has unit eigenvectors \( \frac{1}{\sqrt{2}}(\begin{array}{c} 1 \\ i \end{array}) = v_z \) and \( \frac{1}{\sqrt{2}}(\begin{array}{c} 1 \\ -i \end{array}) = w_z \). The corresponding eigenvalues are 1 and \(-\frac{1}{r_i - 1}\). The matrix \( A \) has the eigenvectors: \( \frac{1}{\sqrt{2}}(\begin{array}{c} 1 \\ i \end{array}) \), \( \frac{1}{\sqrt{2}}(\begin{array}{c} 1 \\ -i \end{array}) \) with eigenvalues 1 and \(-\frac{1}{r_i - 1}\). Each of the matrices \( A \) and \( A_z \) is selfadjoint, and hence each transforms every vector in \( \mathbb{C}^2 \) that is not proportional to its invariant vector to a vector of smaller length. If \( z = e^{i\phi}, 0 < \phi < \pi \), then the fact that \( A \) and \( A_z \) have distinct eigenspaces implies \( \|AA_z\| < 1 \).

(2) For \( z = 1 \), \( \|AA_1\| = 1 \) and the eigenspaces of \( A \) and of \( A_z \) coincide. Therefore, we see that the eigenvalues of \( AA_1 \) are 1 and \( [(r_1 - 1)(r_2 - 2)]^{-1} \). For \( z = -1 \), again the eigenspaces coincide, and we see similarly that the
eigenvalues of $AA_{-1}$ are $-(r_1 - 1)^{-1}$ and $-(r_2 - 1)^{-1}$. As a consequence, $\|AA_{-1}\| < 1$ iff $r_2 > 2$.

(3) Assume first that $r_2 > 2$. Taking (6.2.6) into account we see that in this case $\delta(\sigma_2^{(i)}) \leq \max\{\frac{1}{r_1-1}, \frac{1}{r_2-1}\} < 1$, if $\delta \neq 1$. It follows that $\sigma_n$ is a mean ergodic sequence, hence also $\mu_n$.

(4) When $r_2 = 2$, we claim that $\sigma''_n \stackrel{d}{=} \frac{1}{2}(\sigma_{2n} + \sigma_{2n+2})$ converges to the projection on the space of invariants, and $\sigma_4n$ converges to the projection on the sum of the space of invariants and a space orthogonal to it. The latter space affords the representation $\delta_2$ of $G_1 \ast Z_2$ that takes the value 1 on $G_1$ and $-1$ on the generator of $Z_2$. These statements follow easily by checking the values obtained in (6.2.6) when we apply the four characters of the algebra (using also the fact that $\|AA_z\| < 1$ for $0 < \varphi < \pi$).

This concludes the proof of the mean ergodic theorem in the present case. □

Acknowledgments

I wish to express my gratitude to Professor Hillel Furstenberg for suggesting the problem and for his advice and encouragement.

I would like to thank the Mathematical Science Research Institute in Berkeley, California, for financial support during the time that this paper was written.

References


ABSTRACT. Let \( F_k \) denote the free group on \( k \) generators, \( 1 < k < \infty \), and let \( S \) denote a set of free generators and their inverses. Define \( \sigma_n = \frac{1}{|S^n|} \sum_{w \in S^n} w \), where \( S^n = \{ w : |w| = n \} \), and \( |\cdot| \) denotes the word length on \( F_k \) induced by \( S \). Let \( (X, \mathcal{B}, m) \) be a probability space on which \( F_k \) acts ergodically by measure preserving transformations. We prove a pointwise ergodic theorem for the sequence of operators \( \sigma_n' \), where \( \sigma_n' = \frac{1}{2} (\sigma_n + \sigma_{n+1}) \) acting on \( L^2(X) \), namely: \( \sigma_n' f(x) \to \int_X f \, dm \) almost everywhere, for each \( f \in L^2(X) \). We also show that the sequence \( \sigma_n \) converges to a conditional expectation operator with respect to a \( \sigma \)-algebra which is invariant under \( F_k \). The proof is based on the spectral theory of the commutative convolution subalgebra of \( \ell^1(F_k) \) generated by the elements \( \sigma_n \), \( n \geq 0 \). We then generalize the discussion to algebras arising as a Gelfand pair associated with the group of automorphisms \( G(r_1, r_2) \) of a semi-homogeneous tree \( T(r_1, r_2) \), where \( r_1 \geq 2, r_2 \geq 2, r_1 + r_2 > 4 \).
(The case of $F_k$ corresponds to that of a homogeneous tree of valency $2k$.) We prove similar pointwise ergodic theorems for two classes of subgroups of $G(r_1, r_2)$. One is the class of closed noncompact boundary-transitive subgroups, including any simple algebraic group of split rank one over a local field, for example, $PSL_2(\mathbb{Q}_p)$. The second class is that of lattices complementing a maximal compact subgroup. We also prove a strong maximal inequality in $L^2(X)$ for the groups listed above, as well as a mean ergodic theorem for unitary representations of the groups (due to Y. Guivarc'h for $F_k$). Finally, we describe the structure and spectral theory of a noncommutative algebra which arises naturally in the present context, namely the double coset algebra associated with the subgroup of $G(r_1, r_2)$ stabilizing a geometric edge. The results are applied to prove mean ergodic theorems for a family of lattices in $G(r_1, r_2)$, which includes, for example, $PSL_2(\mathbb{Z})$.

Mathematical Science Research Institute, Berkeley, California
Current address: Department of Mathematics, Princeton University, Princeton, New Jersey 08544
E-mail address: nevo@math.princeton.edu