Research Report

KSTS/RR-97/010
Sep. 25, 1997

Maximal and Pointwise Ergodic Theorems for Word-Hyperbolic Groups

by

K. Fujiwara and A. Nevo

Koji Fujiwara	Amos Nevo
Department of Mathematics	University of Chicago
Keio University	and the Technion, Israel

Department of Mathematics
Keio University

Hiyoshi 3-14-1, Kohoku-ku, Yokohama, 223 Japan

KSTS/R-97/010
September 25, 1997

MAXIMAL AND POINTWISE ERGODIC THEOREMS FOR WORD-HYPERBOLIC GROUPS

KOJI FUJIWARA† AND AMOS NEVO‡

† Department of Mathematics, Keio University, Yokohama, 223 Japan
‡ University of Chicago and the Technion, Israel

To appear in Ergodic Theory and Dynamical Systems

ABSTRACT. Let Γ denote a word-hyperbolic group, and let $S = S^{-1}$ denote a finite symmetric set of generators. Let $S_n = \{w : |w| = n\}$ denote the sphere of radius n, where $|\cdot|$ denotes the word length on Γ induced by S. Define $\sigma_n \overset{d}{=} \frac{1}{\#S_n}\sum_{w\in S_n} w$, and $\mu_n = \frac{1}{n+1}\sum_{k=0}^{n}\sigma_k$. Let $(X, \mathcal{B}, m)$ be a probability space on which Γ acts ergodically by measure preserving transformations. We prove a strong maximal inequality in L^2 for the maximal operator $f_\mu^* = \sup_{n\geq 0}|\mu_n f(x)|$. The maximal inequality is applied to prove a pointwise ergodic theorem in L^2 for exponentially mixing actions of Γ, of the following form : $\mu_n f(x) \to \int_X f dm$ almost everywhere and in the L^2-norm, for every $f \in L^2(X)$. As a corollary, for a uniform lattice $\Gamma \subset G$, where G is a simple Lie group of real rank one, we obtain a pointwise ergodic theorem for the action of Γ on an arbitrary ergodic G-space. In particular, this result holds when $X = G/\Lambda$ is a compact homogeneous space, and yields an equidistribution result for sets of lattice points of the form Γg, for almost every $g \in G$.

§1 DEFINITIONS AND STATEMENTS OF RESULTS

1.1 Definition of ergodic sequences.

Let Γ 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 π of Γ in a Hilbert space $\mathcal{H}$, extend π 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 1.1. Given a unitary representation $(\pi, \mathcal{H})$ of Γ, a sequence $\nu_n \in \ell^1(\Gamma)$ is a *mean ergodic sequence in* $\mathcal{H}$ if $\|\pi(\nu_n)f - E_1 f\| \xrightarrow[n\to\infty]{} 0$ for all $f \in \mathcal{H}$.

1991 *Mathematics Subject Classification.* 22D40, 28D15, 43A20, 43A62.
Key words and phrases. Word-hyperbolic Groups, Word Metric, Maximal Inequality, Pointwise Ergodic Theorems, Convolution Estimates, Rank-one Lie Groups, Lattice Subgroups..
† Partially supported by the Sumitomo Foundation. ‡ Partially supported by the fund for promotion of research at the Technion.

KSTS/R-97/010
September 25, 1997

Let $(X, \mathcal{B}, m)$ be a standard Lebesgue measure space, namely a measure space whose σ-algebra is countably generated and countably separated. Assume Γ acts on X by measurable automorphisms preserving the probability measure m. The action $(\gamma, x) \mapsto \gamma x$ induces a representation of Γ 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 $\mathcal{B}_1 = \{A \in \mathcal{B} | m(\gamma A \triangle A) = 0 \quad \forall \gamma \in \Gamma\}$ is a sub-σ-algebra, and denote by E_1 the conditional expectation operator on $L^1(X)$ which is associated with $\mathcal{B}_1$.

Definition 1.2. Given an action of Γ on a standard Lebesgue space X which preserves a probability measure, a sequence $\nu_n \in \ell^1(\Gamma)$ is called a *pointwise ergodic sequence in* $L^p(X)$ if, 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 $L^p(X)$.

Definition 1.3. The maximal operator associated with the sequence ν_n is given by : $f^*_\nu(x) = \sup_{n \geq 0} |\nu_n f(x)|$, for $f \in L^p(X)$.

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

Definition 1.4.

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

If Γ is finite, then $S_n = \emptyset$ for $n > |\Gamma|$. Since the ergodic theory of finite groups is well known, we will assume from now on that Γ is infinite, without mentioning this condition explicitly. We can now state :

Theorem 1. *Let (Γ, S) be a word-hyperbolic group. Then the sequence μ_n satisfies the strong maximal inequality in L^2, i.e. $\|f^*_\mu\|_2 \leq C(\Gamma)\|f\|_2$ for every $f \in L^2(X)$.*

Let us note that the case where the group Γ is an elementary hyperbolic group has been of course well known for a long time. Indeed by definition, Γ has then a finite index subgroup isomorphic to $\mathbb{Z}$ (see §2). Furthermore, the averages μ_n defined above coincide with the usual ergodic averages when $\Gamma = \mathbb{Z}$ and $S = \{\pm 1\}$ is the set of standard generators. It is easily seen that the maximal inequality for the averages μ_n when Γ is an elementary word-hyperbolic group can be proved using the argument of the Wiener- Hopf maximal inequality for $\mathbb{Z}$-actions.

It is therefore natural to regard the sequence μ_n associated with a general word-hyperbolic group (Γ, S) as the analogue of the familiar ergodic averages on $\mathbb{Z}$. This point of view is bolstered by the fact that the proof we give for the maximal inequality of Theorem 1 applies equally well to elementary and non-elementary word-hyperbolic groups. Indeed, the essential element in the proof of Theorem 1 will be the second of the two convolution estimates given in :

KSTS/R-97/010
September 25, 1997

Theorem 2. *Let (Γ, S) be a non-elementary word-hyperbolic group. Then there exist constants $1 < q < \infty$ and $0 < C < \infty$, depending only on (Γ, S), such that the following inequalities hold :*

a) $\sigma_t * \sigma_s \leq C \sum_{j=0}^{2s} q^{-(s-\frac{1}{2}j)} \sigma_{t-s+j}$ *if* $t \geq s$.
b) $\mu_n * \mu_m \leq C(\mu_{2n} + \mu_{2m})$.

For an elementary word-hyperbolic group, the subadditivity of μ_n is easily verified, and in fact, it has been pointed out long ago by E. M. Stein [S1, S2] that it can be used to prove the L^2-maximal inequality in the case $\Gamma = \mathbb{Z}$ (see also the discussion in §1.2). The first inequality in Theorem 2 also holds, with $q = 1$.

To establish a pointwise ergodic theorem, we will use the maximal inequality, together with some spectral information. Consider first the following :

Definition 1.5.

(1) A unitary representation $(\pi, \mathcal{H})$ of a finitely-generated group (Γ, S) will be called exponentially mixing if there is a dense subspace $\mathcal{H}_0 \subset \mathcal{H}$, such that for every $f \in \mathcal{H}_0$ there are positive constants C and c (depending on f) with : $|\langle \pi(w)f, f\rangle| \leq C\exp(-c|w|)$ for every $w \in \Gamma$.
(2) An action of (Γ, S) on a probability space $(X, \mathcal{B}, m)$ will be called exponentially mixing if the unitary representation of (Γ, S) on $L_0^2(X) = \{f \in L^2(X) \mid \int_X f\,dm = 0\}$ is exponentially mixing.

Theorem 3. *Let (Γ, S) be a word-hyperbolic group. If Γ acts on $(X, \mathcal{B}, m)$ preserving the probability measure m, and the action is exponentially mixing, then the sequence μ_n is a pointwise ergodic sequence in $L^2(X)$.*

The foregoing result will be applied to obtain:

Theorem 4.

(1) *Let G be a connected finite-center simple Lie group of real rank one, and $\Gamma \subset G$ be a uniform lattice subgroup. Let $(X, \mathcal{B}, m)$ be a G-space with an ergodic probability measure m. Then, for any symmetric set of generators of Γ the corresponding sequence μ_n satisfies $\mu_n f(x) \to \int_X f\,dm$ almost everywhere, for every $f \in L^2(X)$.*
(2) *The same conclusion holds for every uniform lattice of the group of automorphisms $G_{n,m} = \mathrm{Aut}(T_{n,m})$ of the semi-homogeneous tree $T_{n,m}$, and uniform lattices of simple algebraic groups of split rank one over local fields.*

To formulate the next result we first recall the following :

Definition 1.6. Let X be a compact metric space, and let ν_n and m be probability measures on X. The sequence of measures ν_n is said to become *equidistributed with respect to m* if $\int_X f(x)d\nu_n \to \int_X f(x)dm$ for every continuous function f on X.

We can now state

Theorem 5. *Let G be as in Theorem 4, and $\Gamma \subset G$ a uniform lattice subgroup. Let (X, m) be compact metric space with a continuous G-action and an ergodic G-invariant measure*

KSTS/R-97/010
September 25, 1997

m. Then

(1) *The sequence of atomic measures* $\nu_n = \frac{1}{n+1}\sum_{k=0}^{n}\frac{1}{\sharp S_k}\sum_{|w|=k}\delta_{wx}$ *becomes equidistributed with respect to* m*, for almost all* $x \in X$.
(2) *In particular, if* $\Lambda \subset G$ *is another uniform lattice, and* $X = G/\Lambda$*, then the set of points* $\Gamma g\Lambda$ *becomes equidistributed in* G/Λ*, when taken with the weights* μ_n*, namely :* $\mu_n f(g\Lambda) \to \int_{G/\Lambda} f dm$*, for all* $f \in C(G/\Lambda)$*, and almost every* $g \in G$.

1.2 Method of proof, remarks and relevant references.

To prove the maximal inequality of Theorem 1 for the sequence μ_n, we will first prove Theorem 2 and establish the subadditivity of μ_n. We can then appeal to a general maximal inequality that applies to any subadditive sequence of self-adjoint Markov operators in L^2, in the sense of the following definition :

Definition 1.7. A sequence T_n of operators on $L^2(X)$ will be called a *subadditive sequence of self adjoint Markov operators* if it satisfies the following :
i) $T_n = T_n^*$, $\|T_n\| \le 1$.
ii) $T_n f \ge 0$ if $f \ge 0$, $T_n 1 = 1$.
iii) There exist a constant $C_0 > 0$, a positive integer k, and a fixed non-negative bounded operator B on $L^2(X)$ such that :

$$T_n T_m f(x) \le C_0(T_{kn}f(x) + T_{km}f(x)) + Bf(x)$$

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

We can now state the maximal inequality alluded to above, which was proved by J. Barrionuevo [B] and independently in [N1] :

Theorem 1.8 : Subadditive maximal inequality [B, N1]. *Let* T_n *be a subadditive sequence of self adjoint Markov operators. Define* $f^*(x) = \sup_{n\ge 0}|T_n f(x)|$*. Then* $\|f^*\|_2 \le C\|f\|_2$ *for all* $f \in L^2(X)$*. We can take* $C = 2C_0 + \|B\|$.

We note that the method of using the subadditive maximal inequality in the group algebra to prove a maximal ergodic theorem in L^2 was developed in [N1] and applied to actions of the free group $\mathbb{F}_k$ and certain other word-hyperbolic groups. This method and the subadditive maximal inequality above generalize similar results due to E. M. Stein [S1, S2] and B. Weiss [W]. In particular, it was applied by E. M. Stein [S1,S2] to prove a pointwise ergodic theorem for the even powers of positivity-preserving self-adjoint contractions on L^2. Also, it is noted in [S1, S2, W] that it implies the pointwise convergence of martingales in L^2, as well as Birkhoff's pointwise ergodic theorem in L^2. Other applications have been considered in [B]. The origin of this maximal inequality is attributed in [S1,W] to A. Kolmogoroff and G. Seliverstoff [K-S], and to R. E. A. C. Paley [P].

It is reasonable to expect that the subadditive maximal inequality for μ_n holds for a class of discrete groups much wider than the class of word-hyperbolic groups. The inequality also makes sense for semi-simple Lie groups for example, and it does in fact hold for the groups $G = SO^0(n,1)$, $n > 2$.

The averages σ_n and β_n seem more directly connected to the word metric than the averages μ_n. The problem of establishing an ergodic theorem for them was introduced by

KSTS/R-97/010
September 25, 1997

V. I. Arnold and A. L. Krylov already in [A-K]. In order to gain some perspective on the possibility of proving maximal and ergodic theorems for σ_n and β_n, we recall some results from [N1]. Consider the following non-elementary word-hyperbolic groups :

(1) F_k = the free group on k generators, $1 < k < \infty$, with $S = \{x_i, x_i^{-1}\}_{i=1}^k$, where $x_1, \dots, x_k$ are free generators.
(2) $\Gamma(r,h) = G_1 * G_2 \cdots * G_r$ = the free product of r finite groups each of order h, where $r \geq 2$, $h \geq 2$, $r + h > 4$, with $S = \bigcup_{i=1}^r G_i \setminus \{e\}$.

Define $q(\Gamma(r,h)) \stackrel{d}{=} (r-1)(h-1)$, and also $q(F_k) \stackrel{d}{=} 2k-1$, $r(\mathbb{F}_k) = 2k$, and $h(F_k) \stackrel{d}{=} 2$.

As to the maximal inequality for ball averages, note the following formula for the size of a sphere of radius n in the groups Γ above : $\#S_n = r(h-1)q^{n-1}$. Therefore $\sigma_n \leq b(\Gamma)\beta_n$ for some constant $b(\Gamma)$ depending only on Γ. It follows that a maximal inequality for the sequence β_n implies one for the sequence σ_n. The converse is also true, since β_n is a convex combination of σ_k, $0 \leq k \leq n$, so the maximal inequalities for spheres and balls in Γ are equivalent. This property holds in fact in every non-elementary word-hyperbolic group, as follows from Theorem 2.15 in §2. Note also that it is in marked contrast to the case of lattices in Euclidean groups, where ball averages satisfy the maximal inequality, but sphere averages do not (consider e.g. the spheres of radius n in $\mathbb{Z}^n$ w.r.t. the standard generators).

As to the ergodic theorems for ball and sphere averages, we recall that it has been proved by Y. Guivarc'h [G] that for $\Gamma = F_k$, the sequence $\sigma'_n = \frac{1}{2}(\sigma_n + \sigma_{n+1})$ is a mean ergodic sequence. This result was generalized in [N1], where the ergodic theorem that follows was proved. First define E to be the orthogonal projection on the subspace of $L^2(X)$ consisting of functions satisfying : $\pi(\sigma_n)f = (-1)^n f$. Let (X,m) be a Γ-space where $E \neq 0$. Then :

Theorem 1.9 : Ergodic theorems for radial averages [N1]. *For Γ and (X,m) as above :*

(1) *σ_n and β_n satisfy the maximal inequality in $L^2(X)$, but are not mean (and hence not pointwise) ergodic sequences in $L^2(X)$.*
(2) *The sequences σ'_n and μ_n are pointwise ergodic sequences in $L^2(X)$.*
(3) *σ_{2n} converges to $E_1 + c(\Gamma)E$, which is a conditional expectation operator w.r.t. a Γ-invariant sub-σ-algebra iff $\Gamma = \Gamma(r,2)$ or $\Gamma = F_k$.*
(4) *β_{2n} converges to $E_1 + c(\Gamma)\frac{q(\Gamma)-1}{q(\Gamma)+1}E$, which is not a conditional expectation operator on a Γ-invariant sub-σ-algebra.*

The convergence is for each function $f \in L^2(X)$, pointwise almost everywhere and in the norm of $L^2(X)$.

Note that the behaviour of ball averages described in part (1) of Theorem 1.9 is again in marked contrast to the case of lattices in Euclidean groups, where ball averages are mean and pointwise ergodic sequences. Here the subsequence β_{2n} does have a limit in norm and pointwise, in contrast with β_n, but the limit is not a projection. As to σ_n, in the Euclidean case sphere averages are not even mean ergodic (again, consider the spheres of radius n in $\mathbb{Z}^n$ w.r.t. the standard generators). Here although σ_n is not a pointwise (or even mean) ergodic sequence, σ'_n has both these properties. σ_{2n} also has a limit in norm and pointwise, which is usually not a conditional expectation.

KSTS/R-97/010
September 25, 1997

In particular, since the sequences σ_n and β_n are not, in general, ergodic sequences, they do not provide a suitable generalization of the ergodic averages on $\mathbb{Z}$. The analysis of the possible limits of their subsequences in a general non-elementary word hyperbolic group seem at this point out of reach.

Finally, we make the following remarks on some closely related problems :

(1) It is natural to consider the behaviour of the radial averages in L^p for $p \neq 2$. We note that in [N-S1] it is shown that the sequence $\mu_n \in \ell^1(\mathbb{F}_k)$, satisfies a weak type (1,1) maximal inequality in every measure-preserving action of $\mathbb{F}_k$. In particular, μ_n is a pointwise ergodic sequence in L^p, for $1 \leq p < \infty$.
(2) Spectral methods are applied in [N1, N2, N-S2] to prove pointwise ergodic theorems for ball and sphere averages on simple Lie groups of real rank one. For some ergodic theorems for discrete subgroups of semi-simple groups see [N4].

Acknowledgements. The authors would like to thank Zlil Sela for many interesting conversations, and to Marc Bourdon for bringing M. Coornaert's paper [Co] to their attention.

§2 Convolution estimates for word-hyperbolic groups

2.1 Preliminaries about word-hyperbolic groups.

The theory of hyperbolic groups was developed by M. Gromov in [Gr], and we review briefly some definitions and results relevant to our purposes without proofs. General reference for the theory are, for example, [Gr], [Sh], [Bo] and [Gh-Ha].

Let X be a complete metric space. Suppose X is geodesic, i.e. the distance between any two points is realized by a geodesic. Denote by $N_\delta(A)$ the δ-neighbourhood of a set $A \subset X$. Recall the following basic definitions.

Definition 2.1(δ-thin triangle). Let T be a triangle in X with geodesic sides α, β, γ, and let $\delta \geq 0$. If

$$\alpha \subset N_\delta(\beta \cup \gamma),\ \beta \subset N_\delta(\gamma \cup \alpha),\ \gamma \subset N_\delta(\alpha \cup \beta),$$

then the triangle is called *δ -thin.*

Definition 2.2(δ-hyperbolic space). If there exists $\delta \geq 0$ such that all geodesic triangles in X are δ-thin, then X is called a *δ-hyperbolic space.*

Definition 2.3(Gromov-hyperbolic space). If X is δ-hyperbolic for some $\delta \geq 0$, we say X is a *Gromov-hyperbolic* space.

Examples of Gromov-hyperbolic spaces.

(1) A tree, i.e., a simply connected, one dimensional simplicial complex, is Gromov-hyperbolic for $\delta = 0$.
(2) Let M be a complete, simply connected Riemannian manifold. If there exists $c < 0$ such that $K \leq c < 0$, where K is the sectional curvature, then M is Gromov-hyperbolic.
(3) In particular, the symmetric spaces of simple Lie groups of real rank one, $\mathbb{RH}^n$, $\mathbb{CH}^n$, $\mathbb{HH}^n$ and $\mathbb{O}_2$ are Gromov-hyperbolic.
(4) Euclidean space $\mathbb{R}^n$ is Gromov-hyperbolic iff $n = 1$.

The following facts are straightforward consequences of the definitions above :

KSTS/R-97/010
September 25, 1997

Lemma 2.4. *Let X be a δ-hyperbolic space. If α, β are geodesics with a common starting point and a common end point, then*

$$\alpha \subset N_\delta(\beta),\ \beta \subset N_\delta(\alpha).$$

Lemma 2.5. *Let X be a δ-hyperbolic space. Suppose geodesics α, β, γ form a triangle. Then there exist points A, B, C on α, β, γ, respectively, such that*

$$d(A, B),\ d(B, C),\ d(C, A) \leq 4\delta.$$

To obtain some further examples for Gromov-hyperbolic spaces, first define :

Definition 2.7(quasi-isometry). Let (X, d) and (X', d') be geodesic spaces. If there exists map $\phi : X \to X'$ satisfying, for some $\epsilon \geq 0$ $N_\epsilon(\phi(X)) = X'$, and

$$\frac{1}{K}(d(x, y) - \epsilon) \leq d'(\phi(x), \phi(y)) \leq Kd(x, y) + \epsilon,$$

for all $x, y \in X$, then ϕ is called (K, ϵ)-*quasi-isometry.* If there exists a (K, ϵ)-quasi-isometry from X to X' for some K and ϵ, we say X and X' are *quasi-isometric.*

The following facts are immediate consequences of the definitions :

Lemma 2.8. *Let X be a Gromov-hyperbolic space. If a complete geodesic space Y is quasi-isometric to X, then Y is Gromov-hyperbolic as well.*

Corollary 2.9.
Let Γ be a countable group of isometries of a Gromov-hyperbolic space X, acting properly discontinuously, freely and with a compact fundamental domain. Then, for any given $x \in X$, the orbit $\Gamma \cdot x \subset X$ (with the induced metric) is quasi-isometric to X, and hence it is a Gromov-hyperbolic space.

¿From Corollary 2.9 we obtain
Further examples of Gromov-hyperbolic spaces :

(1) If Γ is a uniform lattice contained in the group of isometries $SO^0(n, 1)$ of hyperbolic space $\mathbb{RH}^n$, then the orbit $\Gamma \cdot o \subset \mathbb{RH}^n$ is a Gromov-hyperbolic space with respect to the Riemannian metric. The same applies to uniform lattices of any simple Lie group of real rank one.
(2) Let $T_{n.m}$ denote the semi-homogeneous tree, and let $G_{n,m} = \text{Aut}(T_{n.m})$ denote its group of automorphisms. Let $\Gamma \subset G_{n,m}$ be a uniform lattice. Then $\Gamma \cdot o \subset T_{n,m}$ is a Gromov-hyperbolic space w.r.t. the induced tree metric.

We have defined in §1 the word metric on Γ. This distance is associated with the following geometric object :

Definition 2.10(Cayley graph). Let Γ be a finitely generated group and let S be a finite generating set which is symmetric. Consider the geodesic space called the *Cayley graph* $G(\Gamma, S)$ of Γ with respect to S. This graph has a vertex for each element of Γ, and an edge connecting g to gs for each $g \in \Gamma$ and each $s \in S$. Since S generates Γ, G is

KSTS/R-97/010
September 25, 1997

connected. Γ acts on G by multiplication on the left: $g \in \Gamma$ sends a vertex $x \in \Gamma$ to a vertex $gx \in \Gamma$. Assigning unit length to each edge, we make G a geodesic space. The path metric $d(\cdot, \cdot)$ is called *word metric* w.r.t. S, and for $g \in \Gamma$, $|g| = d(1, g)$. The left-action of Γ on G is by isometries.

Examples of Cayley graphs.

(1) The Cayley graph of a free group w.r.t. a free basis is a regular tree.
(2) The Cayley graph of $\mathbb{Z}^2$ w.r.t. the standard generators $x^{\pm 1}, y^{\pm 1}$ is the square grid of horizontal and vertical lines in the plane.

Definition 2.11(word-hyperbolic group). Let (Γ, S) be a finitely generated group and a generating set. If the Cayley graph $G(\Gamma; S)$ is δ-hyperbolic for some $\delta \geq 0$, then we say Γ is a *δ-hyperbolic* group. If Γ is δ-hyperbolic for some δ, then Γ is called *word-hyperbolic.*

The following Lemma follows easily from the definitions, and shows that any choice of a generating set will do.

Lemma 2.12. *Suppose Γ is δ-hyperbolic w.r.t. some generating set S. Let S' be another generating set. Then (Γ, S) and (Γ', S') are quasi-isometric and there exists $\delta' \geq 0$ such that Γ is δ'-hyperbolic w.r.t. S'.*

Before giving some examples, let us note that finite groups and $\mathbb{Z}$ are word-hyperbolic groups, and introduce the following :

Definition 2.13(elementary hyperbolic group). Let Γ be a word-hyperbolic group. If Γ is either a finite group or contains $\mathbb{Z}$ as a finite index subgroup, then Γ is called *elementary.*

Examples of word-hyperbolic groups.

(1) $\mathbb{Z}$ is word-hyperbolic, but $\mathbb{Z}^n$ are not, if $n > 1$. Indeed, taking the standard generators, it is clear there is no δ for which *all* triangles are δ-thin, if $n > 1$. Moreover, any group that contains $\mathbb{Z}^2$ as a subgroup is not word-hyperbolic.
(2) If a countable group Γ acts isometrically, properly discontinuously, freely and with compact fundamental domain on a Gromov-hyperbolic space (X, d), then the orbit $(\Gamma \cdot x, d)$ is quasi-isometric to (Γ, S) (for any finite symmetric S and $x \in X$). By Lemma 2.8, any uniform lattice of the group $G_{n,m}$, or a simple Lie group of real rank one is word-hyperbolic.
(3) Let M be a closed, negatively curved Riemannian manifold. Then $\pi_1(M)$ is word-hyperbolic, by Corollary 2.9.

Another family of examples is given by :

Lemma 2.14. *Let Γ_1, Γ_2 be word-hyperbolic groups and let Γ_3 be a common subgroup. If Γ_3 is finite, then $\Gamma_1 *_{\Gamma_3} \Gamma_2$ is Gromov-hyperbolic.*

In particular, let A, B be finite groups and C a common subgroup. Then $A *_C B$ is word-hyperbolic. If $A \neq C, B \neq C$, then this group is non-elementary except for $\mathbb{Z}_2 * \mathbb{Z}_2$. For example, $SL_2(\mathbb{Z}) = \mathbb{Z}_4 *_{\mathbb{Z}_2} \mathbb{Z}_6$ and $PSL_2(\mathbb{Z}) = \mathbb{Z}_2 * \mathbb{Z}_3$ are non-elementary word-hyperbolic groups (except for $n = 2$, $A_1 = A_2 = \mathbb{Z}_2$). Also, if A_i, $1 \leq i \leq n$ are non-trivial finite groups, then $A_1 * A_2 * \cdots * A_n$ is a non-elementary word-hyperbolic group.

Let (Γ, S) be a finitely generated group, and let $s_n = \sharp S_n$, where as usual S_n is the sphere of radius n w.r.t. the word metric determined by S. For (Γ, S) word-hyperbolic, the sequence $\sharp S_n$ has exponential growth, a fact which is obvious in all the examples listed above : $\sharp S_n \geq CD^n$, where $D > 1$.

For the convolution estimates we intend to prove the following sharper estimate is needed.

Theorem 2.15 : The rate of growth of spheres[Co]. *Suppose (Γ, S) is a non-elementary word-hyperbolic group. Then there exist positive constants C_1, C_2 and $C_3 > 1$ depending only on (Γ, S), such that for all $n \geq 0$*

$$C_1 \cdot C_3^n \leq \sharp S_n \leq C_2 \cdot C_3^n.$$

This fact was proven by D. Sullivan [Su] for convex co-compact Kleinian groups, which are word-hyperbolic, in a slightly different form. The above theorem asserts the conclusion holds for all word-hyperbolic groups. For a proof we refer to [Co].

2.2 The convolution estimates : Proof of Theorem 2.

Let Γ be a Gromov-hyperbolic group. Let S be a generating set and and let δ be a hyperbolic constant of Γ w.r.t. S. Suppose Γ is non-elementary.

Let t, s, j be integers satisfying $0 \leq s \leq t, 0 \leq j \leq 2s, 0 \leq t - s + j$, fixed throughout the following discussion. Suppose a word $w \in \Gamma$ is given such that $|w| = t - s + j$. We count the number of times we can write w as $w = uv$ with $|u| = t$ and $|v| = s$. In the following argument w is fixed while u, v are variable. (Actually, u and v determine each other, since $w = uv$).

Suppose then that $w = uv$ with $|u| = t$ and $|v| = s$. By Lemma 2.5 we can write

$$u = u_1 u_2, v = v_1 v_2, w = w_1 w_2,$$

such that

$$|u| = |u_1| + |u_2|, |v| = |v_1| + |v_2|, |w| = |w_1| + |w_2|,$$

and also (see Figure 1) :

$$|u_1^{-1} w_1|, |u_2 v_1|, |v_2 w_2^{-1}| \leq 4\delta.$$

The following lemma is our basic tool :

Lemma 2.16. *The following expressions :*

$$\left| |v_2| - \frac{j}{2} \right|, \left| |w_2| - \frac{j}{2} \right|, \left| |u_1| - \left(t - s + \frac{j}{2} \right) \right|,$$

$$\left| |w_1| - \left(t - s + \frac{j}{2} \right) \right|, \left| |v_1| - \left(s - \frac{j}{2} \right) \right|, \left| |u_2| - \left(s - \frac{j}{2} \right) \right|$$

are all less than or equal to 6δ.

Proof. Consider the case of v_2, and note that going over the sides of the triangle in Figure 1, the definition of u_i, v_i and w_i gives us the estimates that follow. First, $|w_2| - 4\delta \leq$

KSTS/R-97/010
September 25, 1997

$|v_2| \le |w_2| + 4\delta$. Since $|w_1| + |w_2| = t - s + j$, we can conclude that $t - s + j - 4\delta \le |v_2| + |w_1| \le t - s + j + 4\delta$. But $|u_1| - 4\delta \le |w_1| \le |u_1| + 4\delta$, and also $|u_1| = t - |u_2|$. Hence, substituting these estimates, we have $t - s + j - 8\delta \le |v_2| + t - |u_2| \le t - s + j + 8\delta$. Finally, $|v_1| - 4\delta \le |u_2| \le |v_1| + 4\delta$, and hence $-s + j - 12\delta \le |v_2| - |v_1| \le -s + j + 12\delta$. Together with the equation $|v_1| + |v_2| = s$, we get $j - 12\delta \le 2|v_2| \le j + 12\delta$, and hence $\left||v_2| - \frac{j}{2}\right| \le 6\delta$.

All the other inequalities are proved similarly.

□

Lemma 2.17. *There exists C_4 which depends only on (Γ, S) such that*

$$\sharp\{\textit{choices for } w_1\} \le C_4,$$

$$\sharp\{\textit{choices for } w_2\} \le C_4,$$

$$\sharp\{\textit{choices for } u_1\} \le C_4,$$

$$\sharp\{\textit{choices for } v_2\} \le C_4.$$

Proof. Take a geodesic α from 1 to w in G. Let $A \in \Gamma$ be the point on α such that $d(1, A) = t - s + \frac{j}{2}$(or $t - s + \frac{j+1}{2}$) if j is even(or odd, resp.). Since $|w| = |w_1| + |w_2|$, there exists a geodesic γ from 1 to w such that $w_1 \in \gamma$. By Lemma 2.4, there exists a point $B \in \gamma$ such that $d(A, B) \le \delta$. This implies

$$t - s + \frac{j}{2} - \delta \le d(1, B) \le t - s + \frac{j+1}{2} + \delta.$$

By Lemma 2.16,

$$t - s + \frac{j}{2} - 6\delta \le d(1, w_1) \le t - s + \frac{j}{2} + 6\delta.$$

Since B and w_1 are on the same geodesic γ, we find

$$d(w_1, B) < 7\delta + 1.$$

This shows

$$d(w_1, A) < 8\delta + 1,\ d(u_1, A) < 12\delta + 1.$$

Since $d(u_1, uv_1) \le 4\delta$ and $uv_1 = wv_2^{-1}$,

$$d(wv_2^{-1}, A) < 16\delta + 1.$$

Therefore, we obtain w_1, u_1, wv_2^{-1} from $A \in \Gamma$ by multiplying elements of Γ whose length are less than $16\delta + 1$. Clearly, w_2 is uniquely determined by w_1 (and w, which is fixed in this discussion). By Theorem 2.15, the number of elements whose length are less than $16\delta + 1$ is bounded by $C_2 \sum_{k=0}^{[16\delta+1]} C_3^k$. Letting $C_4 = C_2 \sum_{k=0}^{[16\delta+1]} C_3^k$, we get the desired inequalities. □

KSTS/R-97/010
September 25, 1997

Lemma 2.18. *There exists* C_5 *which depends only on* (Γ, S) *such that*

$$\sharp\{\textit{choices for } u_2\} \le C_5 \cdot C_3^{s-\frac{j}{2}},$$

$$\sharp\{\textit{choices for } v_1\} \le C_5 \cdot C_3^{s-\frac{j}{2}}.$$

Proof. By Theorem 2.15 and Lemma 2.16. □

Proof of Theorem 2 (a).

By Lemmas 2.17 and 2.18 :

$$\sharp\{\text{choices for } u\} \le \sharp\{\text{choices for } u_1\} \cdot \sharp\{\text{choices for } u_2\} \le C_4 \cdot C_5 \cdot C_3^{s-\frac{j}{2}}.$$

Since w is fixed and $w = uv$, clearly u uniquely determines v. Consequently, the number of ways we can write $w = uv$ with $|u| = t, |v| = s, |w| = t - s + j$, is bounded by $C_4 \cdot C_5 \cdot C_3^{s-\frac{j}{2}}$.

Using the estimates for $\sharp S_n$ given by Theorem 2.15, we obtain :

$$\frac{\sharp\{\text{choices for } u\} \cdot \sharp S_{t-s+j}}{\sharp S_t \cdot \sharp S_s} \le \frac{C_2 C_4 C_5}{C_1^2} \cdot C_3^{-(s-\frac{j}{2})}. \tag{2.2.0}$$

This inequality holds for all w satisfying $|w| = t - s + j$. Put $C_6 = \frac{C_2 C_4 C_5}{C_1^2} > 0$, $C_3 = q > 1$, and for each word w of length $t - s + j$, consider $\frac{1}{\sharp S_{t-s+j}} w \in \ell^1(\Gamma)$. By the foregoing, the weight attached to $\frac{1}{\sharp S_{t-s+j}} w$ in the product $\sigma_t * \sigma_s$ is bounded by $C_6 q^{-(s-\frac{j}{2})}$, for each $w \in S_{t-s+j}$. Summing over all words $w \in S_{t-s+j}$, and all possible values of j, namely $0 \le j \le 2s$, the desired inequality follows :

$$\sigma_t * \sigma_s \le C_6 \sum_{j=0}^{2s} q^{-(s-\frac{1}{2}j)} \sigma_{t-s+j}$$

This concludes the proof of Theorem 2(a). □

Proof of Theorem 2(b).

To prove inequality (b), let $n \ge m$, and write

$$\left(\frac{1}{n+1}\sum_{k=0}^{n} \sigma_k\right)\left(\frac{1}{m+1}\sum_{k=0}^{m} \sigma_k\right) = \tag{2.2.1}$$

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

(i) Consider the first summand in the r.h.s of (2.2.1). Fix l with $0 \le l \le m$ and consider the expressions:

$$\sigma_{m+1} * \sigma_l, \sigma_{m+2} * \sigma_l, \cdots, \sigma_n * \sigma_l.$$

KSTS/R-97/010
September 25, 1997

For each i with $1 \leq i \leq n-m$, the set of lengths defined by

$$\{M \mid \text{some word of length } M \text{ has a non zero coefficient in } \sigma_{m+i} * \sigma_l\}$$

is a subset of the interval $[m+i-l, m+i+l]$. Fix a length $L, m-l \leq L \leq n+l$. Suppose L appears in the subset for some i, in other words, some word w of length L has a non-zero coefficient in $\sigma_{m+i} * \sigma_l$. Then the weight of the word w can be estimated according to the position of L in the interval $[m+i-l, m+i+l]$, as follows. The inequality of part (a) shows that if the distance between the point L and the starting point of the interval $m+i-l$ is j, then the weight of $\frac{1}{\sharp S_L} w$ in $\sigma_{m+i} * \sigma_l$ is bounded by $C_6 q^{-(l-\frac{1}{2}j)}$. When we vary i from 1 to $n-m$, j moves from 0 to $2l$, and each j occurs at most once. To see this, note that j uniquely determines i since m, l and L are fixed. Thus the weight of $\frac{1}{\sharp S_L} w$ in $\sum_{i=1}^{n-m} \sigma_{m+i} * \sigma_l$ is bounded by $C_6 \sum_{j=0}^{2l} q^{-(l-\frac{1}{2}j)} \leq C_6 \sum_{k=0}^{\infty} q^{-\frac{k}{2}} = \frac{C_6}{1-q^{-\frac{1}{2}}}$. Now we vary l from 0 to m, and find that for any given L and a word w of length L, the coefficient of $\frac{1}{\sharp S_L} w$ in $(\sum_{k=m}^{n} \sigma_k)(\sum_{k=0}^{m} \sigma_k)$ is bounded by $\frac{C_6(m+1)}{1-q^{-\frac{1}{2}}}$. Therefore, summing over all words w of a given length L, and all possible lengths $0 \leq L \leq 2n$, we conclude :

$$(\sum_{k=m}^{n} \sigma_k)(\sum_{k=0}^{m} \sigma_k) \leq \frac{C(m+1)}{1-q^{-\frac{1}{2}}} \sum_{k=0}^{2n} \sigma_k.$$

(ii) Consider the second summand in the r.h.s. of (2.2.1). Fix a length $L, 0 \leq L \leq 2m$. Then a word w of length L has a non-zero coefficient in $\sigma_t * \sigma_s$ only if $t-s+j=L$ for some j with $0 \leq j \leq 2s$. Fixing j, which is the position of L in the interval $[t-s, t+s]$, we consider the pairs $t \geq s$ such that $t-s+j=L$. Such a pair satisfies $\frac{j}{2} \leq s$ and $s \leq m$. As s ranges on this interval the sum of the weights of $\frac{1}{\sharp S_L} w$ in $\sum_{0 \leq s \leq t \leq m, t-s+j=L} \sigma_t * \sigma_s$ is bounded by $C_6 \sum_{s=\frac{1}{2}j}^{m} q^{-(s-\frac{1}{2}j)} \leq \frac{C_6}{1-q^{-\frac{1}{2}}}$, according to the inequality of part (a). To get the total weight we have to sum on all the possibilities for j but since $0 \leq j \leq 2m$ the weight of $\frac{1}{\sharp S_L} w$ in $\sum_{0 \leq s \leq t \leq m} \sigma_t * \sigma_s$ is bounded by $\frac{C_6(2m+1)}{1-q^{-\frac{1}{2}}}$. This bound holds for all possible lengths L. Putting (i) and (ii) together, we see that (2.2.1) can be estimated by

$$\frac{C_6}{1-q^{\frac{1}{2}}} \frac{1}{n+1} \frac{1}{m+1} \left[2(2m+1) \sum_{k=0}^{2m} \sigma_k + (m+1) \sum_{k=0}^{2n} \sigma_k \right] \leq \frac{8C_6}{1-q^{\frac{1}{2}}} (\mu_{2n} + \mu_{2m}).$$

Defining $C = \frac{8C_6}{1-q^{-\frac{1}{2}}}$ we obtain inequality (b), and this completes the proof of Theorem 2. □

§3 Maximal inequalities and pointwise convergence

3.1 Exponentially mixing actions : Proof of Theorem 3.

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 lemma is of course standard (e.g. [P, N1]) :

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 $f \in V$, $\|f^*\|_2 \le B\|f\|_2$. Then $T_n f(x)$ converges a.e. for all $f \in V$.*

Since Theorem 1 establishes the maximal inequality for all $f \in L^2(X)$, to complete the proof of Theorem 3, we need only show that in an exponentially mixing action there exists a dense set of functions in $V = L^2_0(X)$ for which $\pi(\mu_n)f(x)$ converges almost everywhere. This fact is an immediate consequence of the following

Lemma 3.2. *If $f \in L^2_0(X)$ satisfies (for some $C_f, c > 0$) $|\langle \pi(w)f, f\rangle| \le C_f \exp(-c|w|)$ for every $w \in \Gamma$, then*

(1) *There exist $B_f, b > 0$ satisfying $\|\pi(\sigma_n)f\| \le B_f \exp(-bn)$, for all $n \in \mathbb{N}$.*
(2) *$\pi(\sigma_n)f(x) \to 0$ for almost all $x \in X$.*
(3) *$\pi(\mu_n)f(x) \to 0$ for almost all $x \in X$.*

Proof. Given any function $U(u)$ on Γ, by definition of convolution in the group :

$$\sigma_n * \sigma_n(U) = \frac{1}{(\sharp S_n)^2} \sum_{|u|=n} \sum_{|v|=n} U(uv) \ .$$

By Theorem 2 (a), we have, for a non-negative function U :

$$\sigma_n * \sigma_n(U) \le C \sum_{j=0}^{2n} q^{-(n-\frac{1}{2}j)} \sigma_j(U).$$

Now take $U(u) = C_f \exp(-c|u|)$, and compute, using $\sigma_n^* = \sigma_n$:

$$\|\sigma_n f\|_2^2 = \langle \pi(\sigma_n * \sigma_n)f, f\rangle \le \frac{1}{(\sharp S_n)^2} \sum_{|u|=n} \sum_{|v|=n} |\langle \pi(uv)f, f\rangle|$$

$$\le \frac{1}{(\sharp S_n)^2} C_f \sum_{|u|=n} \sum_{|v|=n} \exp(-c|uv|) = \sigma_n * \sigma_n(U)$$

$$\le C \sum_{j=0}^{2n} q^{-(n-\frac{1}{2}j)} \sigma_j(U) = CC_f \sum_{j=0}^{2n} q^{-(n-\frac{1}{2}j)} \exp(-cj)$$

$$= CC_f q^{-n} \sum_{j=0}^{2n} \exp(\frac{\log q}{2} - c)j \le B_f \exp(-bn)$$

Therefore, $\sum_{n=0}^\infty \|\sigma_n f\|_2^2 < \infty$, and consequently $\sum_{n=0}^\infty |\sigma_n f|^2$ is in $L^1(X)$. It follows that $\sigma_n f(x) \to 0$ for almost all $x \in X$, and the same holds of course also for $\mu_n f(x)$. □

Combining Lemma 3.1, Theorem 1 and Lemma 3.2 completes the proof of Theorem 3. □

KSTS/R-97/010
September 25, 1997

3.3 Uniform lattices : Proofs of Theorems 4 and 5. The proof of Theorem 4(1) will be complete once we prove the following :

Lemma 3.3. *In any ergodic action of a connected finite-center simple Lie group of real rank one, the action of a uniform lattice is exponentially mixing. More generally, the restriction of any unitary representation of G that does not contain G-invariant vectors to a uniform lattice is exponentially mixing.*

We first recall the exponential decay estimates for the matrix coefficients of C^∞-vectors in irreducible unitary representations of connected finite-center simple non-compact Lie groups. The estimates are based on results of M. Cowling [C] and R. Howe [H], and have been recently reformulated in a convenient form by A. Katok and R. Spatzier [Ka-Sp], and D. Kleinbock and G. A. Margulis [K-M]. We formulate here a special case of Corollary 2.4.4 in [K-M] which will suffice for our purposes. First recall the following

Definition 3.4. A unitary representation $(\pi, \mathcal{H}_\pi)$ of a locally compact second countable group G is said to contain the trivial representation weakly, if there exists a sequence of unit vectors $f_j \in \mathcal{H}_\pi$, such that the matrix coefficients $\varphi_j(g) = \langle \pi(g)v_j, v_j \rangle$ converge to the constant function $\varphi(g) = 1$, uniformly on compact sets in G.

Now for G as above, fix a left G-invariant Riemannian metric d on G, which is also right K-invariant, where K is a maximal compact subgroup.

Theorem 3.5 : Exponential decay of matrix coefficients [K-M, Cor. 2.4.4]**.** *Let G be a connected finite center simple non compact Lie group, and let ρ be a unitary representation of G, which does not contain the trivial representation weakly. Then there exist a positive constant α_ρ, satisfying :*

$$|\langle \pi(g)f, h \rangle| \leq C(f,h)\exp(-\alpha_\rho d(g,e))$$

for any two C^∞-vectors $f, h \in \mathcal{H}_\rho$, and any $g \in G$.

Proof of Lemma 3.3. Given an arbitrary unitary representation π of G that does not contain the trivial representation, it is possible to find an increasing sequence of closed G-invariant subspaces $\mathcal{H}_n \subset \mathcal{H}$, with the following properties :

(1) The closure of $\bigcup_{n\geq 0} \mathcal{H}_n$ equals $\mathcal{H}$.
(2) The representation ρ_n of G in $\mathcal{H}_n$ does not contain the trivial representation weakly.

To construct $\mathcal{H}_n$, decompose π to a direct integral of irreducible representations. To each neighbourhood U of the trivial representation (in the Fell topology on the unitary dual of G) there corresponds a self adjoint projection $\mathcal{E}_U$ acting on $\mathcal{H}_\pi$, and commuting with the operators $\pi(g)$, $g \in G$. Define $\mathcal{H}_n$ as the kernel of the projection $\mathcal{E}_{U_n}$ corresponding to a set U_n, where the sequence U_n constitutes a neighbourhood basis of the trivial representation. The intersection of the decreasing sequence U_n contains only the trivial representation, and hence property (1) of $\mathcal{H}_n$ is evident. Property (2) of $\mathcal{H}_n$ holds by definition of the Fell topology. Since $\mathcal{H}_n$ does not contain the trivial representation weakly, the restriction π_n of π to $\mathcal{H}_n$ satisfies the conclusion of Theorem 3.5, with exponential decay rate α_{π_n} depending on n, in general.

Now note that in any unitary representation of G, C^∞-vectors are dense, since for every vector $v \in \mathcal{H}_\pi$, the vector $\int_G f(g)\pi(g)vdg$ is a C^∞-vector if $f \in C_c^\infty(G)$. Hence each $\mathcal{H}_n$, and therefore $\mathcal{H}$, admits a dense set of vectors with exponentially decaying matrix coefficients.

Consequently, it suffices to show that the restriction of a unitary representation ρ that does not contain the trivial representation weakly to a uniform lattice is exponentially mixing. Recall that for a uniform lattice, the distance on Γ inherited from the Riemannian distance on G and the word length metric on Γ are quasi-isometric, as noted in example (2) following Def. 2.13. Indeed this fact follows since Γ acts on G by left translations, isometrically with respect to the Riemannian distance, and with a compact fundamental domain. We therefore have the estimate $A^{-1}|w| \leq d(w,e) \leq A|w|$, where $|w|$ is the word length of $w \in \Gamma$ with respect to a given symmetric set of generators S, and $d(w,e)$ is the distance introduced above on G. Hence, for a unitary representation ρ of G that does not contain the trivial representation weakly :

$$|\langle \pi(\gamma)f, h\rangle| \leq C(f,h)\exp(-\alpha_\rho A^{-1}|\gamma|)$$

and the proof of Lemma 3.3 and Theorem 4(1) is complete. □

Theorem 4(2) is proved similarly, using similar decay estimates for the analog of C^∞-vectors in irreducible unitary representations of the group $G_{n,m}$ [Ol,F-N], and simple algebraic groups of split rank one over local fields.

Proof of Theorem 5.

Let $f_i, i \in \mathbb{N}$ be a sequence of continuous functions dense in $C(X)$ in the uniform norm. Then there exists a set $Y \subset X$ of measure one, such that for every $y \in Y$ and every f_i, $\mu_n f_i(y) \to \int_X f dm$, as $n \to \infty$. For an arbitrary $f \in C(X)$, and every $y \in Y$, we have :

$$\left|\mu_n f(y) - \int_X f dm\right| \leq |\mu_n f(y) - \mu_n f_i(y)| + \left|\mu_n f_i(y) - \int_X f_i dm\right| + \left|\int_X f_i dm - \int_X f dm\right|.$$

For a given $\epsilon > 0$ we choose f_i satisfying $\|f - f_i\| < \frac{\epsilon}{3}$, and then N s.t. for any $n > N$, $\left|\mu_n f_i(y) - \int_X f_i dm\right| < \frac{\epsilon}{3}$, and the result follows. □

References

[A-K] V.I. Arnold and A.L. Krylov *Uniform distribution of points on a sphere and some ergodic properties of solutions of linear ordinary differential equations in the complex plane.* Sov. Math. Dokl. Vol. 4, No. 1, pp 1-5 (1962).

[B] J. Barrionuevo, *Estimates for some Kakeya-type maximal operators.* Trans. A. M. S., vol. 335, no. 2, pp. 667-682 (1993).

[Bo] B. Bowditch, *Notes on Gromov's hyperbolicity criterion*, in "Group Theory from a Geometric Viewpoint", ed. by E. Ghys, et al., World Sci. Pub., New Jersey, 1991, 64-167.

[Ca] J. Cannon, *The combinatorial structure of cocompact discrete hyperbolic groups*, Geom. Dedicata, 16, 1984, 123-148.

[C] M. Cowling, *Sur les coefficients des representations unitaires des groupes de Lie simples.* Lect. Notes in Math vol. 739, pp. 132 -178 (1979).

[Co] M. Coornaert, *Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov.* Pacific J. Math., vol. 159 no. 2, pp. 241–270 (1993).

[FT-N] A. Figa-Talamanca and C. Nebbia, *Harmonic analysis and representation theory for groups acting on homogeneous trees.* London Mathematical Society Lecture Notes Series vol.162, Cambridge Univ. Press (1992).

[Gr] M. Gromov *Hyperbolic groups*, in "Essays in group theory", MSRI Publ. 8, ed. by S.M. Gersten, Springer, 1987, 75-263.

[Gh-Ha] E. Ghys, P. de la Harpe, *Sur les groupes hyperboliques d'apres Mikhael Gromov*, edit. by E. Ghys and P. de la Harpe, Birkhauser, 1990.

[G] Y. Guivarc'h, *Généralisation d'un théorème de von Neumann.* C.R. Acad. Sc. Paris, t.268 pp. 1020-1023 (1969).

[H] R. Howe, *On a notion of rank for unitary representations of the classical groups.* in: Harmonic Analysis and Group Representations, C.I.M.E. , pp. 223 - 232 (1980).

[K-M] D. Y. Kleinbock and G. A. Margulis, *Bounded orbits of non-quasi-unipotent flows on homogeneous spaces.* In : Translations of the A.M.S., series 2, vol. 171, Eds. L. A. Bunimovich, B.M. Gurevich, Ya. B. Sinai, pp. 141-173 (1996).

[Ka-Sp] A. Katok and R. J. Spatzier, *First cohomology of Anosov actions of higher rank Abelian groups and applications to rigidity.* Publ. Math. I. H. E. S., vol. 79, pp. 131-156 (1994).

[K-S] A. Kolmogoroff and G. Seliverstoff, *Sur la convergence des sèries de Fourier.* Comptes Rendous,178, pp. 303-306 (1924).

[N1] A. Nevo, *Harmonic analysis and pointwise ergodic theorems for non-commuting transformations.* Journal of the A. M. S. vol. 7, no.4, pp. 875-902 (1994).

[N2] A. Nevo, *Pointwise ergodic theorems for radial averages on simple Lie groups I.* Duke Journal of Mathematics, vol. 76, no.1, pp.113-140 (1994).

[N3] A. Nevo, *Pointwise ergodic theorems for radial averages on simple Lie groups II.* Duke Journal of Mathematics, vol. 86, no. 2, pp. 239-259 (1997).

[N4] A. Nevo, *Spectral transfer and pointwise ergodic theorems for Kazhdan groups.* Preprint.

[N-S1] A. Nevo and E. M. Stein, *A generalization of Birkhoff's pointwise ergodic theorem.* Acta Mathematica, vol. 173, pp.135-154 (1994).

[N-S2] A. Nevo and E. M. Stein, *Analogs of Wiener's ergodic theorems for semisimple groups I.* Annals of Mathematics, vol. 145, pp. 565-595 (1997).

[Ol] G. I. Olshansky, *Classification of irreducible representations of groups of automorphisms of Bruhat-Tits trees.* Funct. Anal. and Appl. 11,26-34,(1977).

[P] R. E. A. C. Paley, *A proof of a theorem of averages.* Proc. London Math. Soc.,31,289-300 (1930).

[Pe] K. Petersen, ERGODIC THEORY Cambridge University Press, 1983.

[Sh] H. Short, *Notes on word hyperbolic groups*, in "Group Theory from a Geometric Viewpoint", ed. by E. Ghys, et al., World Sci. Pub., New Jersey, 1991, 3-63.

[S1] E. M. Stein, *On the maximal ergodic theorem.* Proc. Nat. Acad. Sci. U.S. 47, 1894-1897,(1961)

[S2] E. M. Stein, *Topics in harmonic analysis related to the Littlewood Paley theory.* Annals of Math. Studies, vol. 63, Princeton University Press (1970).

[S3] E. M. Stein, *Maximal functions : Spherical means.* Proc. Natl. Acad. Sci. U.S.A. vol. 73, no 7, pp 2174-2175 (1976).

[Su] D. Sullivan, The density at infinity of a discrete group of hyperbolic motions, Publ. Math. I. H. E. S., vol 50, pp 171–202 (1979).

KSTS/R-97/010
September 25, 1997

[W] B. Weiss, *Positive cones in Hilbert space and a maximal inequality.* INEQUALITIES III, Ed. O. Shisha, Academic Press, 1972.

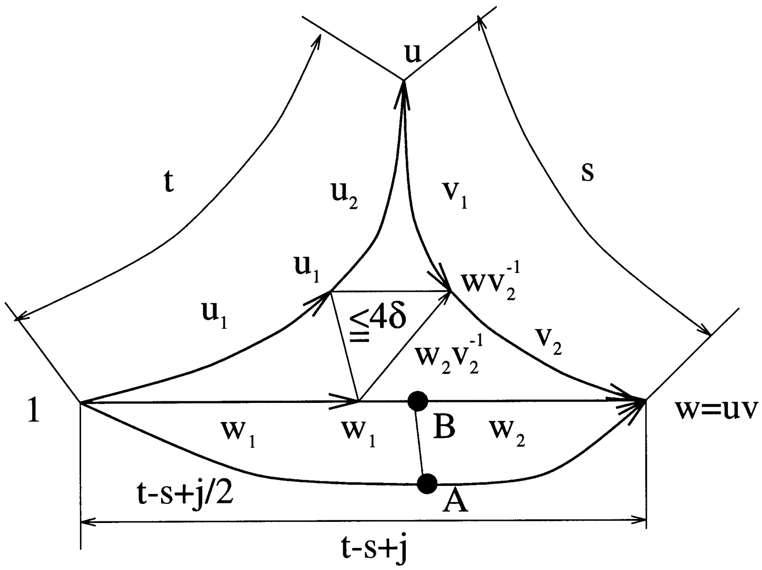

Figure 1