Ergod. Th. & Dynam. Sys. (2019), **39**, 2689–2716 © *Cambridge University Press, 2017*
doi:10.1017/etds.2017.128

Hyperbolic geometry and pointwise ergodic theorems

LEWIS BOWEN† and AMOS NEVO‡

† *Mathematics Department, 1 University Station C1200, University of Texas at Austin, Austin, TX 78712, USA*
(e-mail: lpbowen@math.utexas.edu)

‡ *Mathematics Department, Technion – Israel Institute of Technology, Haifa 32000, Israel*
(e-mail: anevo@tx.technion.ac.il)

(*Received* 2 *February* 2017 *and accepted in revised form* 24 *October* 2017)

Abstract. We establish pointwise ergodic theorems for a large class of natural averages on simple Lie groups of real rank one, going well beyond the radial case considered previously. The proof is based on a new approach to pointwise ergodic theorems, which is independent of spectral theory. Instead, the main new ingredient is the use of direct geometric arguments in hyperbolic space.

Contents

1	Introduction	2689
2	Preliminaries	2692
3	The classical method of rotations: geodesic polar coordinates	2697
4	Ergodic theorems for the isometry group of the hyperbolic plane	2700
5	Ergodic theorems for general real rank-one groups	2708
Acknowledgements		2715
References		2715

1. *Introduction*

1.1. *Ergodic subgroups and ergodic theorems.*

Let G be a connected simple real Lie group of real rank one with finite center. Our purpose in the present paper is to generalize the existing pointwise and maximal ergodic theorems for the ball and shell averages on G well beyond the case of radial averages, using an entirely new approach.

The ingredients our method utilizes are elementary hyperbolic geometry, the classical pointwise and maximal ergodic theorems for one-dimensional flows, the Howe–Moore ergodicity theorem, and some variations on the classical 'method of rotation'. In particular, our proof is independent of any spectral estimates associated with spherical functions on

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

the group G. Refined and detailed estimates of spherical functions formed the basis of the only previous proof of pointwise ergodic theorems for radial averages on G [**Ne94**, **Ne97**, **NS97**], but reliance on such estimates necessarily restricts the averages under study to be radial. We remark that our approach in fact extends the range of validity of the radial pointwise ergodic theorems to the space $L \log L$, which is not readily accessible by spectral methods, but the main point in our analysis is the use of geometric ideas to dispense with the assumption of radiality in the pointwise ergodic theorems on G.

To be more precise, let (X, μ) be a standard probability space and let $\mathrm{Aut}(X, \mu)$ be the group of all measure-preserving automorphisms of (X, μ) in which two automorphisms are identified if they agree on a conull set. The group $\mathrm{Aut}(X, \mu)$ is equipped with the weak topology under which is it separable and completely metrizable (see [**Ke10**], for example). A *measure-preserving action of G on (X, μ)* is a continuous homomorphism from G into $\mathrm{Aut}(X, \mu)$. Given such an action, a probability measure η on G and a function $f \in L^1(X, \mu)$ on X, define $\eta(f) \in L^1(X, \mu)$ by

$$\eta(f)(x) = \int_G f(g^{-1}x)\, d\eta(g).$$

Also let $\mathbb{E}[f|G] \in L^1(X, \mu)$ denote the conditional expectation of f on the sigma-algebra of G-invariant Borel sets. This is well defined up to a measure-zero set.

A one-parameter family of probability measures $\{\eta_r\}_{r>0}$ on G is *pointwise ergodic in L^p* if, for any measure-preserving action $G \curvearrowright (X, \mu)$ and any $f \in L^p(X, \mu)$, the averages $\eta_r(f)$ converge pointwise almost everywhere (a.e.) as $r \to \infty$ to conditional expectation $\mathbb{E}[f|G]$. Such a family is not required to be a semigroup.

The basis of our approach to proving ergodic theorems is the following simple and natural idea. Suppose that G is a locally compact second countable (lcsc) group and $H < G$ is a closed subgroup. We say that H has the *automatic ergodicity property* if whenever G acts on a probability space (X, μ) by measure-preserving transformations ergodically then the action restricted to H is also ergodic. In this case, any pointwise ergodic family of probability measures η_r supported on H is a pointwise ergodic family for G. It follows that, for any $g, g' \in G$, the averages $\delta_g * \eta_r * \delta_{g'}$ satisfy the same conclusion. Given any parametrized family $\delta_{g_b} * \eta_r * \delta_{g'_b}$, with b ranging over some lcsc space B, the corresponding parametrized pointwise ergodic families can be averaged with respect to a probability measure on B. Under suitable natural conditions this gives rise to a host of additional pointwise ergodic families supported on G.

A most significant case where this method can be employed is when G is a simple non-compact algebraic group. Indeed then by the Howe–Moore theorem any closed non-compact subgroup $H < G$ has the automatic ergodicity property.

Of course, one natural possibility is to choose H as an amenable subgroup of G. Then we can use the classical theory of amenable groups to find ergodic sequences in H, whose translates $\delta_{g_b} * \eta_r * \delta_{g'_b}$, $b \in B$ can then be averaged further on B. For example, when G is a simple non-compact real Lie group, this raises the possibility of proving pointwise ergodic theorems for G by averaging on translates of probability measures on a unipotent subgroup, for example one which is isomorphic to $\mathbb{R}$. Below we will develop and utilize this approach extensively for the group $\mathrm{SL}_2(\mathbb{R})$ and a unipotent subgroup N.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Furthermore, let us note that parametrized families of translated averages on general, not necessarily amenable subgroups also occur naturally, and we will use the principle stated above in that case too. For example, we will consider the case of parametrized translates of averages on $\mathrm{SO}^0(2, 1)$ embedded in $\mathrm{SO}^0(n, 1)$, corresponding to embeddings of totally geodesic hyperbolic planes in the n-dimensional hyperbolic space. This will allow us to generalize ergodic theorems established for $\mathrm{SO}^0(2, 1)$ to isometry groups of higher-dimensional (real, complex and quaternionic) hyperbolic spaces.

Thus this approach may be viewed as a generalization of the familiar 'method of rotation' used extensively in classical analysis and singular integral theory.

We remark that the approach used in the present paper to prove ergodic theorems for simple real rank-one Lie groups was motivated by the method used to prove ergodic theorems for free groups in [**BN13**]. There the approach is based on considering an appropriately chosen amenable 'measurable subgroup' of $\mathbb{F}$. This 'subgroup' is a sub-equivalence relation $\mathcal{R}$ of the orbit equivalence relation of $\mathbb{F}$ acting on its boundary. Whenever $\mathbb{F}$ acts on a probability space $\mathbb{F}\curvearrowright(X, \lambda)$, there is a natural extension $\mathbb{F}\curvearrowright(X \times \partial\mathbb{F}, \lambda \times \nu)$ and a sub-equivalence relation $\mathcal{R}^X$ of the orbit relation on $X \times \partial\mathbb{F}$. It was shown in [**BN13**] that if the action $\mathbb{F}\curvearrowright(X, \lambda)$ is ergodic then the sub-equivalence relation $\mathcal{R}^X$ has at most two ergodic components, which is an analog of the Howe–Moore theorem in this case. Moreover, the sub-relation $\mathcal{R}^X$ is amenable (indeed, it is hyperfinite), and admits ergodic sequences. The radial ergodic theorems for the free groups are then proved by first averaging over finite sub-equivalence relations of the relation $\mathcal{R}^X$ and then averaging the result over the boundary. Note also that this method allows much more general types of averaging sequences to be analyzed similarly, since we can average with respect to a variety of measures on the boundary.

1.2. *Main results.* As above, let G be a connected simple real Lie group of real rank one with finite center and $G\curvearrowright(X, \mu)$ be a measure-preserving action on a standard Borel probability space. Let $KAK = G$ be a Cartan decomposition of G and $A = \{a_t\}_{t\in\mathbb{R}}$ be the Cartan subgroup. We allow any linear parametrization of the Cartan subgroup A, and denote the finite center of G by Z. We use these coordinates to define the following natural averages on G.

Definition 1. Let $U, V \subset K$ be sets of positive measure which are Z-invariant, namely $ZU = U$ and $ZV = V$. For $r, \epsilon > 0$, let

$$B_r^{U,V} = \{k_1 a_t k_2 : k_1 \in U, t \in [0, r], k_2 \in V\},$$
$$\Sigma_{r,\epsilon}^{U,V} = \{k_1 a_t k_2 : k_1 \in U, t \in [r, r + \epsilon], k_2 \in V\}.$$

Let $\sigma_{r,\epsilon}^{U,V}$, $\beta_r^{U,V}$ denote the probability measures on G obtained by restricting Haar measure to $\Sigma_{r,\epsilon}^{U,V}$ and $B_r^{U,V}$ respectively and normalizing to have mass one.

The following theorem is our main result.

THEOREM 1.1. *Let G be a connected simple real Lie group of real rank one with finite center. For any $\epsilon > 0$:*

(1) *the families $\{\sigma_{r,\epsilon}^{U,V}\}_{r>0}$ and $\{\beta_r^{U,V}\}_{r>0}$ are pointwise and mean ergodic in L^p $(1 < p < \infty)$ and in $L \log L$;*

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

(2) *the families* $\{\sigma_{r,\epsilon}^{U,V}\}_{r>0}$ *and* $\{\beta_r^{U,V}\}_{r>0}$ *satisfy the strong* (p, p) *type maximal inequality (for all* $p > 1$*) and the* $L \log L$ *maximal inequality.*

The terminology is explained in §2.

Remark 1.1.

(1) By taking $U = V = K$ we recover the fact that spherical shell averages are pointwise ergodic in L^p for all $p > 1$. This was first proven in [**Ne94**, **Ne97**, **NS97**] by spectral methods, and the fact that these averages are also pointwise ergodic in $L \log L$ is new.
(2) The main novelty occurs when U or V is not equal to K. In this case, the averages are referred to as 'bi-sector averages'. Special cases have been proven previously only under the very restrictive hypothesis that $X = G/\Gamma$ is a homogeneous action.
(3) We prove a more general result (Theorem 5.6) in which U and V are replaced with arbitrary bounded Z-invariant probability densities on K.
(4) Theorem 1.1 holds for the balls and shells defined by any choice of G-invariant Riemannian metric on the symmetric space G/K.
(5) The assumption that U and V are Z-invariant will be convenient in our arguments, but is not strictly necessary. We will explain this comment further in Remark 5.1.
(6) A general maximal inequality for not necessarily radial averages on connected semisimple Lie groups of any real rank was proved in [**GN10**]. Pointwise ergodic theorems for averages which are K-invariant on one side only were proved under extra spectral assumptions in [**Ne98**].

Organization of the paper. In §2 we introduce the necessary definitions and notation associated with maximal inequalities and ergodic theorems, and also list some basic standard analytic facts that will be used repeatedly in many of the arguments later on. Section 3 contains a brief exposition of the classical method of rotations associated with geodesic polar coordinates in Euclidean and hyperbolic space. Section 4 is devoted to proving Theorem 1.1 in the special case $G = \mathrm{PSL}_2(\mathbb{R})$. In §5 we prove Theorem 1.1 by using the $\mathrm{PSL}_2(\mathbb{R})$ case and the fact that any connected simple adjoint real Lie group G of real rank one contains an embedded subgroup L isomorphic to $\mathrm{PSL}_2(\mathbb{R})$, conveniently located in G. We will then treat the case of finite covers of the adjoint group at the end of §5, and in Remark 5.1 comment on the assumption of Z-invariance of the sets U and V.

2. *Preliminaries*

2.1. *Averaging operators, maximal inequalities and ergodic families.* Let G be an lcsc group acting by measure-preserving transformations on a standard Borel probability space (X, μ). If ν is any probability measure on G then we also consider ν to be an operator from $L^1(X, \mu)$ to $L^1(X, \mu)$ via the formula

$$\nu(f)(x) = \int_G f(g^{-1}x)\, d\nu(g).$$

Maximal functions and maximal inequalities. Let $r \mapsto \nu_r$, $r > 0$, be a one-parameter family of compactly supported probability measures on G. We do not require it to be a

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

semigroup or to consist of absolutely continuous measures on G. However, we do require it to be a w^*-continuous map from $\mathbb{R}_+$ to the space of probability measures $\mathcal{P}(G)$ on G, namely that for any continuous function F on G, $r \mapsto \nu_r(F)$ is continuous. We will make this assumption on every one-parameter family of probability measures without saying so explicitly. The reason why this assumption will be useful is as follows.

Let $\mathbb{M}_\nu$ denote the associated maximal operator defined by

$$\mathbb{M}_\nu[f] = \sup_{r \geq 1} \nu_r(|f|).$$

Note that $r \geq 1$ in the definition above instead of $r > 0$. This is because we are only interested in limits as $r \to \infty$ and will not be concerned with limits as $r \searrow 0$.

For a general family of averages ν_r, it need not be the case that the maximal function $\mathbb{M}_\nu[f]$ associated with a Borel function f is measurable. However, for an lcsc group G there is a topological model for the action $G \curvearrowright X$. In other words, we may assume without loss of generality that X is a compact metrizable space on which G acts jointly continuously and μ is a standard Borel probability measure on X [**BK96**, Theorem 2.6.6]. So there exists a subspace $\mathcal{C}(X) \subset L^\infty(X)$ which is norm dense in every $L^p(X)$, $1 \leq p < \infty$, such that for every $f \in \mathcal{C}(X)$ the map $g \mapsto f(g^{-1}x)$ is continuous in g for almost every $x \in X$. Namely, $\mathcal{C}(X)$ is the subspace of continuous functions on X. Under the w^*-continuity assumption for ν_r, for such f the maximal function $\mathbb{M}_\nu[f](x)$ is equal to the supremum of $\nu_q(|f|)(x)$ where $q \in \mathbb{Q} \cap \mathbb{R}_+$, so that it is indeed measurable.

For $f \in L^1(X, \mu)$ and $k \geq 1$, let

$$\|f\|_{L(\log L)^k} = \int_X |f|(\log(\max(|f|, 1)))^k \, d\mu$$

and let $L(\log L)^k(X, \mu) \subset L^1(X, \mu)$ be the set of all functions with $\|f\|_{L(\log L)^k} < \infty$. This is a vector subspace, although $\|\cdot\|_{L(\log L)^k}$ is not a norm.

We say that a family $\{\nu_r\}_{r>0}$ of Borel probability measures on G satisfies:

- the *weak-type* $(1, 1)$ *maximal inequality* if there is a constant $C_1 > 0$ such that
$$\mu(\{x \in X : \mathbb{M}_\nu[f] \geq t\}) \leq \frac{C_1 \|f\|_1}{t} \quad \text{for all } f \in L^1(X, \mu),\ t > 0;$$
- the *strong-type* (p, p) *maximal inequality* if there is a constant $C_p > 0$ such that
$$\|\mathbb{M}_\nu[f]\|_p \leq C_p \|f\|_p \quad \text{for all } f \in L^p(X, \mu);$$
- the *strong-type* $L(\log L)^k$ *maximal inequality* if there is a constant $C_{L(\log L)^k} > 0$ such that
$$\|\mathbb{M}_\nu[f]\|_{L^1} \leq C_{L(\log L)^k} \|f\|_{L(\log L)^k} \quad \text{for all } f \in L(\log L)^k(X, \mu).$$

Mean and pointwise convergence. We let $\mathbb{E}[f|G]$ denote the conditional expectation of f on the sigma-algebra of G-invariant measurable subsets. We say that a family $\{\nu_r\}_{r>0}$ of Borel probability measures on G is:

- *mean ergodic in* L^p if $\nu_r(f)$ converges in L^p-norm to $\mathbb{E}[f|G]$ as $r \to \infty$ for every $f \in L^p(X, \mu)$;
- *mean ergodic in* $L(\log L)^k$ if $\nu_r(f)$ converges in L^1-norm to $\mathbb{E}[f|G]$ as $r \to \infty$ for every $f \in L(\log L)^k(X, \mu)$;

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

- *pointwise convergent in* L^p (or $L(\log L)^k$) if $\nu_r(f)$ converges pointwise a.e. as $r \to \infty$ for every $f \in L^p(X, \mu)$ (or $L(\log L)^k$);
- *pointwise ergodic in* L^p (or $L(\log L)^k$) if $\nu_r(f)$ converges pointwise a.e. to $\mathbb{E}[f|G]$ as $r \to \infty$ for every $f \in L^p(X, \mu)$ (or $L(\log L)^k$).

Finally, we say that $\{\nu_r\}_{r>0}$ is a *good averaging family in* L^p, if it satisfies the strong-type (p, p) maximal inequality, is mean ergodic for functions in L^p and, in addition, the family is pointwise ergodic in L^p. We define good averaging families in $L(\log L)^k$ similarly.

When ν_r is a good averaging family in every L^p, $1 < p < \infty$, and also in $L \log L$ we will abbreviate and say that it is a *good averaging family*. If, in addition, the family satisfies the weak-type (1, 1) maximal inequality then we will say that it is an L^1*-good averaging family*. When this holds, it follows that the family is in fact pointwise and mean ergodic in L^1. This is one of several useful facts which we will use repeatedly, and which we now state.

2.2. *Standard arguments.* We list the following standard results which will be used frequently below. We start with the following elementary fact.

LEMMA 2.1. (Domination lemma) *Suppose* $\{\eta_r\}_{r>0}$ *and* $\{\nu_r\}_{r>0}$ *are w*-continuous families of probability measures on G and there is a constant* $C > 0$ *such that* $\eta_r \leq C\nu_r$ *for all r. If* $\{\nu_r\}_{r>0}$ *satisfies either a weak-type* (1, 1)*, strong-type* (p, p) *or* $L(\log L)^k$ *maximal inequality then* $\{\eta_r\}_{r>0}$ *satisfies the same type of maximal inequality.*

We will have occasion to average parametrized families of probability measures on G, and thus state the following fact, which is a straightforward consequence of the definitions.

LEMMA 2.2. (Averaging strong maximal inequalities) *Let* (Z, ζ) *be a standard probability space and let* $(z, r) \mapsto \tau_{z,r}$ *be a measurable map from* $Z \times (0, \infty)$ *into the space of Borel probability measures on G. Suppose that for each* $z \in Z$ *the family of averages* $\{\tau_{z,r}\}_{r>0}$ *is w*-continuous in r and satisfies a strong-type* (p, p) *(or* $L(\log L)^k$*) maximal inequality with constants* $C_{z,p}$ *and, moreover, the constants* $C_{z,p}$ *are uniformly bounded for* $z \in Z$*. Let* $\tau_r = \int_{z\in Z} \tau_{z,r}\, d\zeta(z)$*. Then* $\{\tau_r\}_{r>0}$ *satisfies the strong-type* (p, p) *(or* $L(\log L)^k$*) maximal inequality.*

We recall that given two bounded Borel measures ν and λ on G, their convolution is defined as the functional

$$\nu * \lambda(y) = \int_G \int_G y(gh)\, d\nu(g)\, d\lambda(h) \quad \text{for all } y \in C_c(G)$$

where $C_c(G)$ denotes the space of compactly supported continuous functions on G. Clearly, the support of $\nu * \lambda$ is contained in the closure of the product of the supports of ν and of λ, and if ν and λ are probability measures, then so is $\nu * \lambda$. Convolution is associative and therefore convolution of triples is well defined by

$$\nu * \alpha * \lambda(y) = \int_G \int_G \int_G y(gag')\, d\nu(g)\, d\alpha(a)\, d\lambda(g') \quad \text{for all } y \in C_c(G).$$

Below we will often consider maximal inequalities for a family of measures arising as convolutions of probability measures on G. We thus state the following proposition.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

PROPOSITION 2.3. *Let $\{\eta_r\}_{r>0}$, $\{\nu_r\}_{r>0}$, $\{\lambda_r\}_{r>0}$ be w* continuous families of compactly supported probability measures on G.*

(1) *If $\{\eta_r\}_{r>0}$ satisfies the strong-type (p, p) or $L(\log L)^k$ maximal inequality and ν and λ are fixed (but arbitrary) probability measures, then $\{\nu * \eta_r * \lambda\}_{r>0}$ satisfies the same maximal inequality.*

(2) *If $\{\eta_r\}_{r>0}$, $\{\nu_r\}_{r>0}$ and $\{\lambda_r\}_{r>0}$ each satisfy the strong-type (p, p) maximal inequality then so does $\{\nu_r * \eta_r * \lambda_r\}_{r>0}$.*

(3) *If $\{\eta_r\}_{r>0}$, $\{\nu_r\}_{r>0}$ and $\{\lambda_r\}_{r>0}$ each satisfy the weak-type $(1, 1)$ maximal inequality then $\{\nu_r * \eta_r * \lambda_r\}_{r>0}$ satisfies the $L(\log L)^3$ maximal inequality.*

We note that part (1) follows from Lemma 2.2, applied to the measure space $(G \times G, \nu \times \lambda)$ and the family defined by $\tau_{(g,g'),r} = \delta_g * \eta_r * \delta_{g'}$. Part (2) is elementary, since the maximal functions of each of the families ν_r, η_r and λ_r are themselves in L^p. Part (3) is proved in [**Fav72**] (see Theorem 1(ii) and its proof).

Finally, we recall the following well-known version of the classical Banach principle (see, for example, [**Ne06**] for complete details).

THEOREM 2.4. *Suppose there exists a norm-dense subset $\mathcal{D} \subset L^1(X, \mu)$ of functions such that for every $f \in \mathcal{D}$, $\eta_r(f)$ converges pointwise a.e. as $r \to \infty$.*

- *If $\{\eta_r\}_{r>0}$ satisfies the weak-type $(1, 1)$ maximal inequality, then $\{\eta_r\}_{r>0}$ is pointwise and mean convergent in L^1.*
- *If $\{\eta_r\}_{r>0}$ satisfies the strong-type (p, p) (or $L(\log L)^k$) maximal inequality, then $\{\eta_r\}_{r>0}$ is pointwise and mean convergent in L^p (or $L(\log L)^k$).*

If, for every $f \in \mathcal{D}$, $\eta_r(f)$ converges pointwise to the ergodic mean $\mathbb{E}[f|G]$ a.e. as $r \to \infty$, then we can replace 'pointwise and mean convergent' in the two conclusions above with 'pointwise and mean ergodic'.

Our discussion below will utilize certain polynomially weighted versions of the Birkhoff pointwise ergodic theorem. The results we require undoubtedly follow from a suitable weighted ergodic theorem on the real line already in existence, but we have not located a convenient reference, so we include a short self-contained proof of the simple special case we will use in the following two results.

THEOREM 2.5. (Polynomially weighted Birkhoff ergodic theorem) *Let $\psi : (0, \infty) \to (0, \infty)$ be a continuous function satisfying*

$$\psi(t) = Ct^{\kappa} + O(t^{\kappa'})$$

for some constants $C > 0$ and $\kappa > \kappa' > 0$. Let η be the measure on $\mathbb{R}$ defined by $\eta(E) = \int_E \psi(t)\, dt$ for Borel $E \subset \mathbb{R}$. Finally, for $T > 0$, let η_T denote η restricted to $[0, T]$ and normalized to have mass 1:

$$\eta_T(E) = \frac{\eta(E \cap [0, T])}{\eta([0, T])}, \quad E \subset \mathbb{R}.$$

Then $\{\eta_T\}_{T>0}$ is an L^1-good averaging family for $\mathbb{R}$ (as an additive group).

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Proof. Let $h_t \in \mathrm{Aut}(X, \mu)$ be an $\mathbb{R}$-flow, which we can assume to be ergodic. Let

$$\mathbb{A}_T^\eta f = \frac{1}{\eta[0, T]} \int_0^T f \circ h_t^{-1} \, d\eta(t), \quad \mathbb{M}^\eta[f](x) = \sup_{T>0} \mathbb{A}_T^\eta |f|(x).$$

There exists a dense subset $\mathcal{D} \subset L^1(X, \mu)$ satisfying the pointwise ergodic theorem, consisting of all functions of the form $f - f \circ h_t + c$ where $f \in L^\infty(X, \mu)$, $t \in \mathbb{R}$ and c is a constant. A standard computation shows that $\mathbb{A}_T^\eta(f - f \circ h_t + c)$ converges pointwise a.e. to c as $T \to \infty$, and a standard argument shows that $\mathcal{D}$ is dense in $L^2(X, \mu)$ and therefore dense in $L^1(X, \mu)$.

Let λ_T denote the uniform probability measure on $[0, T]$. Then there is a constant $C' > 0$ such that as measures on $\mathbb{R}$, $\eta_T \le (\kappa + 1)C'\lambda_T$, where C' depends only on the C and the implicit constant in the error term in the formula $\psi(t) = Ct^\kappa + O(t^{\kappa'})$. So the domination lemma and the well-known weak-type (1, 1) maximal inequality for $\{\lambda_T\}_{T>0}$ imply the weak-type (1, 1) maximal inequality for $\{\mathbb{A}_T^\eta\}_{T>0}$. Similarly, $\mathbb{M}^\eta[f]$ satisfies the strong-type (p, p) maximal inequality for all $p > 1$. So Theorem 2.4 concludes the proof for $\{\mathbb{A}_T^\eta\}_{T>0}$. □

We now apply the previous theorem to intervals of exponentially increasing size, as follows. Assume $\epsilon > 0$, $r > 0$, $b > 0$ and define the measures $\eta_{r,\epsilon}$ on $\mathbb{R}$ by

$$\eta_{r,\epsilon}(E) = \frac{\eta(E \cap [2 \sinh br, 2 \sinh b(r + \epsilon)])}{\eta[2 \sinh br, 2 \sinh b(r + \epsilon)]}, \quad E \subset \mathbb{R}.$$

PROPOSITION 2.6. *For any measure-preserving $\mathbb{R}$-action on a probability space, the measures $\{\eta_{r,\epsilon}\}_{r>0}$ constitute an L^1-good averaging family, for each fixed $\epsilon > 0$ and $b > 0$.*

Proof. Let $h_t \in \mathrm{Aut}(X, \mu)$ be an $\mathbb{R}$-flow, which we can assume to be ergodic. Define operators $\mathcal{A}_{r,\epsilon}[f]$ for $f \in L^p(X, \mu)$ by

$$\mathcal{A}_{r,\epsilon}^\eta[f](x) = \eta_{r,\epsilon}(f) = \frac{1}{\eta[2 \sinh br, 2 \sinh b(r + \epsilon)]} \int_{2 \sinh br}^{2 \sinh b(r+\epsilon)} f \circ h_t^{-1} \, d\eta(t).$$

Also define

$$\mathcal{M}_\epsilon^\eta[f](x) = \sup_{r>0} \mathcal{A}_{r,\epsilon}^\eta[|f|](x).$$

Let us first note the general fact that for any $\delta > 0$ and for $T > 0$, the averages on $\mathbb{R}$ defined by the normalized restriction of η to the set $[0, (1 + \delta)T) \setminus [0, T) = [T, (1 + \delta)T)$ satisfy

$$\begin{aligned}
\mathbb{A}_{(1+\delta)T}^\eta f(x) - \mathbb{A}_T^\eta f(x) &= \frac{1}{\eta([0, (1 + \delta)T))} \int_0^{(1+\delta)T} f(h_t^{-1}x) \, d\eta(t) \\
&\quad - \frac{1}{\eta([0, T))} \int_0^T f(h_t^{-1}x) \, d\eta(t) \\
&= \frac{\eta([0, T)) - \eta([0, (1 + \delta)T))}{\eta([0, (1 + \delta)T))} \frac{1}{\eta([0, T))} \int_0^T f(h_t^{-1}x) \, d\eta(t) \\
&\quad + \frac{\eta([T, (1 + \delta)T))}{\eta([0, (1 + \delta)T))} \frac{1}{\eta([T, (1 + \delta)T))} \int_T^{(1+\delta)T} f(h_t^{-1}x) \, d\eta(t),
\end{aligned}$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

so that we have the identity

$$\frac{1}{\eta([T,(1+\delta)T))}\int_T^{(1+\delta)T} f(h_t^{-1}x)\,d\eta(t) = \frac{\eta([0,T))}{\eta([T,(1+\delta)T))}(\mathbb{A}^\eta_{(1+\delta)T}f(x)-\mathbb{A}^\eta_T f(x))+\mathbb{A}^\eta_T f(x).$$

This identity implies immediately that the strong maximal inequalities which are valid for the family $\mathbb{A}^\eta_T$ are valid also for the left-hand side, as long as $\eta([0,T))/\eta([T,(1+\delta)T))$ remains bounded. The same argument also establishes the case of the weak-type $(1,1)$ maximal inequality.

Note that for any given $\delta>0$, as $T\to\infty$, $\mathbb{A}^\eta_{(1+\delta)T}f(x)-\mathbb{A}^\eta_T f(x)$ converges to zero a.e. and $\mathbb{A}^\eta_T f(x)$ converges a.e. to $\int_X f\,d\mu$ (in the ergodic case), as follows from Theorem 2.5. Therefore the identity shows that pointwise convergence of the left-hand side holds as well.

Let us apply this fact to the choice $T=2\sinh(br)$ and δ satisfying $(1+\delta)T=2\sinh(b(r+\epsilon))$. Then

$$\begin{aligned}\frac{\eta([0,T))}{\eta([T,(1+\delta)T))} &= \frac{\int_0^{2\sinh br}\psi(t)\,dt}{\int_{2\sinh br}^{2\sinh b(r+\epsilon)}\psi(t)\,dt}\\ &= \frac{(\sinh br)^{\kappa+1}+O((\sinh br)^{\kappa'+1})}{(\sinh b(r+\epsilon))^{\kappa+1}-(\sinh br)^{\kappa+1}+O((\sinh b(r+\epsilon))^{\kappa'+1})}\\ &= \frac{1}{\left(\dfrac{\sinh b(r+\epsilon)}{\sinh br}\right)^{\kappa+1}-1}(1+O((\sinh br)^{\kappa'-\kappa})).\end{aligned}$$

Now using

$$\frac{\sinh(b(r+\epsilon))}{\sinh(br)}=\cosh\frac{b\epsilon}{2}+\coth\frac{br}{2}\sinh\frac{b\epsilon}{2}\geq\cosh\frac{b\epsilon}{2}\geq 1+\frac{b^2\epsilon^2}{8},$$

we conclude that, given fixed $\epsilon>0$ and $b>0$, we can use the previous identity, since the ratio in question remains uniformly bounded.

We conclude that the family of operators $\mathcal{A}^\eta_{r,\epsilon}$ satisfies the strong-type (p,p) and $L\log L$ maximal inequalities and converges pointwise a.e. for f in these function spaces. In fact, by the same argument the weak-type $(1,1)$ maximal inequality and pointwise convergence for L^1-functions hold as well. □

Let us now turn to describe in more detail the method of rotations, which will play a significant role in the proof of radial ergodic theorems for simple groups of real rank one which will be established below.

3. *The classical method of rotations: geodesic polar coordinates*

3.1. *The method of rotations in Euclidean space.* Not long after Wiener's proof of the pointwise ergodic theorem for ball averages on multi-dimensional flows [**Wi39**], it was pointed out by Pitt [**Pi42**] that part (but not all) of Wiener's theorem can be established by an argument known as 'the method of rotation' in the context of

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Calderón–Zygmund theory. We summarize this approach to the pointwise ergodic theorem for Euclidean ball averages, since it includes several arguments and several facts which we will use repeatedly below.

The main idea is simply to view the normalized uniform measure $\beta_r^{(n)}$ on a ball of radius r in $\mathbb{R}^n$, $n \geq 2$, as a convex average of the normalized (weighted) measures on the intervals $[0, rv]$, with $v \in \mathbb{S}^{n-1}$ ranging over the unit sphere, taken with its unique rotation-invariant probability measure $m_{\mathbb{S}^{n-1}}$. Using polar coordinates on $\mathbb{R}^n$, namely representing a general point as (t, v) with $t \geq 0$ and $v \in \mathbb{S}^{n-1}$, this amounts to writing

$$\beta_r^{(n)} f(x) = \int_{v \in \mathbb{S}^{n-1}} \frac{n}{r^n} \int_0^r t^{n-1} f(T_{t,v}^{-1} x)\, dt\, dm_{\mathbb{S}^{n-1}}(v)$$

where $T_{t,v}^{-1} x = x - tv$. Now for each fixed v, the subgroup $\mathbb{R} \cdot v$ is isomorphic to $\mathbb{R}$, and the weighted one-dimensional operators $\mathcal{L}_r^v f(x) = (n/r) \int_0^r (\frac{t}{r})^{n-1} f(T_{t,v}^{-1} x)\, dt$ are supported on it. The polynomially weighted Birkhoff ergodic theorem (Theorem 2.5) implies these operators satisfy the weak-type (1, 1) maximal inequality and are pointwise convergent in L^1.

Now, the higher-dimensional ball average $\beta_r^{(n)}$ we are interested in is the average of $\mathcal{L}_r^v f(x)$ over $v \in \mathbb{S}^{n-1}$. As a result, norm convergence for the ball averages in L^p, $1 \leq p < \infty$, follows immediately from norm convergence of $\mathcal{L}_r^v f(x)$. Similarly, the strong-type maximal inequalities in L^p, $p > 1$, and $L \log L$ for the ball averages are immediate consequences of the fact that they hold for $\mathcal{L}_r^v f(x)$ with fixed uniform norm bounds, independent of v, using Lemma 2.2. As to pointwise convergence of the ball averages, it is immediate for bounded functions, for example by applying Lebesgue's dominated convergence theorem to the uniformly bounded family of functions $v \mapsto f(T_{t,v}^{-1}(x))$ in the measure space $L^p(\mathbb{S}^{n-1}, m_{\mathbb{S}^{n-1}})$. Pointwise convergence for general functions in L^p, $p > 1$ or $L \log L$ then follows using Theorem 2.4.

Finally, an important additional point is that we must identify the pointwise limit of $\beta_r^{(n)}(f)$, which differs, in general, from the limits of $\mathcal{L}_r^v f(x)$. However, the limit function of $\beta_r^{(n)} f(x)$ is in fact invariant under the $\mathbb{R}^n$-action (as noted in Wiener's original argument). Indeed the norm limit of $\beta_r^{(n)}(f \circ T_{t,v})$ is the same for any choice of $v \in \mathbb{R}^n$, as follows easily by comparing the two integrals, and using the asymptotic invariance (=Følner) property of Euclidean balls. Thus the limit is the conditional expectation of f with respect to the sigma-algebra of $\mathbb{R}^n$-invariant measurable sets, as stated in the ergodic theorem.

Note, however, that the previous argument fails to establish a crucial part of Wiener's ergodic theorem. Namely, it does not establish pointwise almost sure convergence for L^1-functions, and cannot be used to prove a weak-type (1, 1) maximal inequality. While the family of ball averages in $\mathbb{R}^n$ is the convex average of the one-dimensional operators $\mathcal{L}_r^v f(x)$ over $v \in \mathbb{S}^{n-1}$, and while each one-dimensional family satisfies the weak-type (1, 1) maximal inequality, this inequality does not average and the inequality for the convex average does not follow. This limitation will be present throughout our discussion below.

3.2. *The method of rotations in non-Euclidean space.* The hyperbolic space $\mathbb{H}^n$ also admits geodesic polar coordinates analogous to those on $\mathbb{R}^n$. To describe them more

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

explicitly, recall that the connected component G of the isometry group of $\mathbb{H}^n$ acts transitively and the stability group K of a point $p_0 \in \mathbb{H}^n$ acts transitively on the unit tangent sphere at p_0. Fix a geodesic line ℓ in $\mathbb{H}^n$ passing through p_0, which is an orbit of a one-parameter group $A = \{a_r, r \in \mathbb{R}\}$ isomorphic to $\mathbb{R}$, so that $\ell = A \cdot p_0$. Every other geodesic through p_0 is of the form $kA \cdot p_0$ for some $k \in K$, that is, it is the orbit of p_0 under the conjugate subgroup kAk^{-1}. It follows that the connected component G of the isometry group of hyperbolic space admits a decomposition of the form $G = \mathrm{Iso}^0(\mathbb{H}^n) = KAK$, and in fact $G = \{e_G\} \cup KA_+K$, where $A_+ = \{a_r \in A; r > 0\}$. Furthermore, the set Ka_rK is mapped under the map $g \mapsto gp_0$ to a sphere of radius $|r|$ with center p_0. We let m_K denote the unique Haar probability measure on K, and we let σ_r be the unique K-bi-invariant probability measure on the set Ka_rK. The measure σ_r coincides of course with the measure $m_K * \delta_{a_r} * m_K$, where the convolution is defined on G and δ_{a_r} is the Dirac probability measure at a_r.

PROPOSITION 3.1. *Let $G = \mathrm{Iso}^0(\mathbb{H}^n)$ act on (X, μ) by probability-measure-preserving transformations.*

(1) *The uniform average $\mu_r = (1/r) \int_0^r \sigma_t \, dt$ of the spherical measures σ_t is a good averaging family.*

(2) *Let ν and λ be any two Borel probability measures on the group K. Then*

$$r \mapsto \frac{1}{r} \int_0^r (\nu * \delta_{a_t} * \lambda) \, dt$$

is a good averaging family.

Proof. Part (2) implies part (1) upon taking $\nu = \lambda = m_K$. For part (2), first write, using bilinearity of convolution,

$$\frac{1}{r} \int_0^r (\nu * \delta_{a_t} * \lambda) \, dt = \nu * \left(\frac{1}{r} \int_0^r \delta_{a_t} dt \right) * \lambda.$$

Now the strong maximal inequalities in L^p, $p > 1$, and in $L \log L$ for $A \cong \mathbb{R}$-actions immediately imply the corresponding maximal inequalities for the averages under consideration, by Proposition 2.3(1). As to pointwise convergence for (say) bounded functions, applying the one-dimensional averages supported on $A \cong \mathbb{R}$ to λf, we can conclude pointwise convergence as $r \to \infty$. Now the averaging operator ν maps a pointwise convergent family of bounded functions to another pointwise convergent family of bounded functions, for example using Lebesgue's dominated convergence theorem as in the Euclidean case explained above. Using Theorem 2.4 again, pointwise almost sure and norm convergence for the desired averages follows. Finally, the identification of the limit requires an additional argument, since asymptotic invariance arguments are absent in our non-amenable group. In the present case, it is well known that $A \subset \mathrm{Iso}^0(\mathbb{H}^n)$ acts ergodically in every ergodic $\mathrm{Iso}^0(\mathbb{H}^n)$-space by the Howe–Moore theorem [**HM79**]. Thus $(1/r) \int_0^r (\lambda f) \circ a_t^{-1} dt$ converges pointwise a.e. to the ergodic mean $\mathbb{E}[f|G]$ by the one-dimensional pointwise ergodic theorem, and applying ν to this pointwise convergent family yields the desired conclusion. □

Let us note that the measure μ_r when projected to $\mathbb{H}^n$ is supported in the ball of radius r with center p_0. It gives a measure which is absolutely continuous with respect

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

to the Riemannian measure β_r on the ball, namely the measure which is the normalized restriction of the isometry-invariant measure on hyperbolic space to the ball. However, μ_r and β_r are radically different measures, since the radial density of β_r is given by

$$\beta_r = \frac{\int_0^r \sigma_t(\sinh t)^{n-1}\,dt}{\int_0^r (\sinh t)^{n-1}\,dt}. \tag{1}$$

Thus the application of the classical method of rotation using geodesic polar coordinates leaves much to be desired, in the case of $\mathrm{Iso}^0(\mathbb{H}^n)$. We are interested in establishing ergodic theorems for the ball measures β_r, which arise intrinsically from hyperbolic geometry, and not just for the uniform average of the sphere measures σ_t.

Much of the present paper is based on the following two observations. First, this goal can be still be achieved by the method of rotation applied to averages on horospheres, rather than geodesics. Second, averages on horospheres can be used to establish convergence even for natural non-radial averages as well, using more refined geometric arguments. We now demonstrate these observations and their consequences in the case of the hyperbolic plane, which is fundamental to the developments that follow.

4. *Ergodic theorems for the isometry group of the hyperbolic plane*

Let $\mathbb{H}^2$ denote the hyperbolic plane, equipped with a Riemannian metric of constant negative sectional curvature. We identify $\mathrm{PSL}_2(\mathbb{R})$ with the group of orientation-preserving isometries of $\mathbb{H}^2$ in the usual way. For $r > 0, \epsilon > 0$, define the annuli

$$\Sigma_{r,\epsilon} = \{g \in \mathrm{PSL}_2(\mathbb{R}) :\ d(gp_0,\ p_0) \in [r,\ r+\epsilon]\}$$

where $d(\cdot,\cdot)$ is the invariant Riemannian (=hyperbolic) distance in $\mathbb{H}^2$. Let $\sigma_{r,\epsilon}$ be the probability measure on $\Sigma_{r,\epsilon}$ obtained by normalizing the restriction of Haar measure. Also let β_r be the probability measure on $\{g \in \mathrm{PSL}_2(\mathbb{R}) :\ d(gp_0,\ p_0) \le r\}$ (for some $p_0 \in \mathbb{H}^2$) obtained by normalizing the restriction of Haar measure.

We start by proving the following radial ergodic theorem, which will be followed later on by a non-radial generalization.

THEOREM 4.1. *For $G = \mathrm{PSL}_2(\mathbb{R})$, the families $\{\beta_r\}_{r>0}$ and $\{\sigma_{r,\epsilon}\}_{r>0}$ are good averaging families (for any fixed $\epsilon > 0$).*

4.1. *The upper half plane model.*

For convenience we let $\mathbb{H}^2$ denote the upper half plane $\mathbb{H}^2 = \{x + iy \in \mathbb{C} :\ y > 0\}$ with the Riemannian metric $\mathcal{G}$ given by

$$\mathcal{G}(w, z) = \langle w, z\rangle / y$$

for any vectors w, z in the tangent space of $x + iy \in \mathbb{H}^2$, where the inner product on the right-hand side is the usual inner product in Euclidean space. With this metric, $\mathbb{H}^2$ is a model of the hyperbolic plane (so this is consistent with previous notation). It is a complete simply connected Riemannian manifold with constant sectional curvature -1. We denote the associated Riemannian distance in $\mathbb{H}^2$ by $d(\cdot,\cdot)$ and note the following well-known formula:

$$\cosh(d(x_1 + y_1 i,\ x_2 + y_2 i)) = 1 + \frac{|x_1 - x_2|^2 + |y_1 - y_2|^2}{2y_1 y_2}. \tag{2}$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

The group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{H}^2$ by fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d},$$

and this action preserves the Riemannian metric and the distance. Because the center $\{\pm I\} \leq \mathrm{SL}_2(\mathbb{R})$ acts trivially, this induces an action of $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{\pm I\}$. It is well known that this gives an isomorphism of $\mathrm{PSL}_2(\mathbb{R})$ with $\mathrm{Isom}^+(\mathbb{H}^2)$, the group of orientation-preserving isometries of the hyperbolic plane.

For $r, t \in \mathbb{R}$, $\theta \in [0, 2\pi)$, let

$$k_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad a_r = \begin{pmatrix} e^{r/2} & 0 \\ 0 & e^{-r/2} \end{pmatrix}, \quad n_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

and

$$\mathrm{SO}_2(\mathbb{R}) = \{k_\theta\}_{\theta \in \mathbb{R}}, \quad A = \{a_r\}_{r \in \mathbb{R}}, \quad N = \{n_t\}_{t \in \mathbb{R}}.$$

We note that the double cover $\mathrm{SL}_2(\mathbb{R}) \to \mathrm{PSL}_2(\mathbb{R})$ is injective when restricted to A and N, and is a double cover when restricted to $\mathrm{SO}_2(\mathbb{R})$. We denote by K the image of $\mathrm{SO}_2(\mathbb{R})$.

Because the isometry group acts transitively we may assume without loss of generality that $p_0 = i$. Note that K is the stabilizer of p_0 and $d(a_r p_0, p_0) = |r|$ according to the distance formula (2). The next lemma is central to our approach.

LEMMA 4.2. *Let $r, t > 0$. The following are equivalent:*

(1) $d(n_t p_0, p_0) = r$;
(2) $K a_r K = K n_t K$;
(3) $\cosh(r) = 1 + t^2/2$;
(4) $t = 2\sinh(r/2)$.

Proof. Since K is the stabilizer of p_0, $d(a_r p_0, p_0) = r$ and $\mathrm{PSL}_2(\mathbb{R})$ acts simply transitively on the unit tangent bundle of $\mathbb{H}^2$, it follows that $\{K a_r K\}_{r \geq 0}$ is a parametrization the space of double cosets of K in G. Using the distance formula (2),

$$\cosh(d(x_1 + y_1 i, x_2 + y_2 i)) = 1 + \frac{|x_1 - x_2|^2 + |y_1 - y_2|^2}{2y_1 y_2},$$

the equivalence of (1) and (3) follows upon substituting $x_1 = 0$, $x_2 = t$, and $y_1 = y_2 = 1$. The equivalence of (3) and (4) is elementary. □

Remark 4.1. We note that in the group $G = \mathrm{PSL}_2(\mathbb{R})$, the Cartan decomposition $g = k a_r k' = u a_s u'$ with $k, k', u, u' \in K \subset \mathrm{PSL}_2(\mathbb{R})$, and both r and s positive, is *unique*. Indeed, the absolute values of r and s are equal to $d(gi, i)$ (where K stabilizes i), hence $r = s$ if both are positive. Note that otherwise uniqueness fails, because $w_0 a_t w_0^{-1} = a_{-t}$ if w_0 is the Weyl group element. Now looking at gi, since $K \cdot i = i$ we have $k a_r i = u a_r i$. So $u^{-1}k$ stabilizes the point $a_r i \neq i$ and also the point i. Since it acts as a rotation with center i, it must be the identity. Then also $k' = u'$ and the representation is unique.

Remark 4.2. Let us briefly digress and note the following facts, which will be used in our discussion later on. Multiplying the Riemannian metric $\mathcal{G}$ used above by a positive scalar c gives rise to a different Riemannian manifold, namely the unique complete simply

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

connected Riemannian manifold of constant sectional curvature $-1/\sqrt{c}$. The distance between two points in the upper half plane in the associated metric d_c is $d_c(p, q) = cd(p, q)$. The geodesics passing through p_0 remain the same, and $d_c(a_r p_0, p_0) = c|r| = cd(a_r p_0, p_0)$. Thus the Riemannian metric $c\mathcal{G}$ gives rise to geodesics through p_0 which are parametrized by $d_c(a_{r/c} p_0, p_0) = r$. Thus changing the curvature on the hyperbolic plane amounts to reparametrizing the geodesics, and hence also the radii of balls and shells. It follows that to prove Theorem 4.1 it suffices to establish the result when the curvature is -1.

Furthermore, it follows that a version of Lemma 4.2 still holds for the metric d_c in the modified form $d_c(n_t p_0, p_0) = r \iff Ka_{r/c}K = Kn_tK \iff \cosh(r/c) = 1 + (t^2/2) \iff t = 2\sinh(r/2c)$.

Let η denote the measure on N given by

$$\eta(E) = \int 1_E(n_t)|t|\, dt.$$

LEMMA 4.3. (KNK decomposition) *Let m_K denote the Haar probability measure on K and m_G denote the Haar measure on $G = PSL_2(\mathbb{R})$ normalized so that*

$$m_G(\{g \in G : d(gp_0, p_0) \le r\}) = 2\pi(\cosh r - 1)$$

is the same as the area of the ball of radius r in $\mathbb{H}^2$. Then

$$\frac{1}{\pi} m_G = m_K * \eta * m_K.$$

Proof. Since both m_G and $m_K * \eta * m_K$ are bi-K-invariant, it suffices to prove

$$\frac{1}{\pi} m_G(E) = m_K * \eta * m_K(E)$$

for bi-K-invariant subsets $E \subset G$ (in other words, sets satisfying $E = KEK$). Because balls centered at p_0 generate such sets, it suffices to prove

$$\frac{1}{\pi} m_G(B_r(p_0)) = m_K * \eta * m_K(B_r(p_0))$$

where $B_r(p_0) = \{g \in G : d(gp_0, p_0) \le r\}$. From the previous lemma, and the fact that $d(n_t p_0, p_0) = d(n_{-t} p_0, p_0)$,

$$\begin{aligned} m_K * \eta * m_K(B_r(p_0)) &= \eta(\{n_t : |t| \le 2\sinh(r/2)\}) \\ &= 2\int_0^{2\sinh(r/2)} t\, dt = (2\sinh(r/2))^2 = 2\cosh(r) - 2 \\ &= \frac{1}{\pi} m_G(B_r(p_0)). \end{aligned}$$

□

Definition 2. For $r, \epsilon > 0$, let $\eta_{r,\epsilon}$ be the probability measure on N given by

$$\eta_{r,\epsilon}(E) = \frac{\eta(E \cap \{n_t : t \in [2\sinh(r/2), 2\sinh((r+\epsilon)/2)]\})}{\eta(\{n_t : t \in [2\sinh(r/2), 2\sinh((r+\epsilon)/2)]\})}.$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Lemmas 4.2 and 4.3 together imply

$$\sigma_{r,\epsilon} = m_K * \eta_{r,\epsilon} * m_K. \tag{3}$$

Proof of Theorem 4.1. By the polynomially weighted Birkhoff ergodic theorem , $\{\eta_{r,\epsilon}\}_{r>0}$ is an L^1-good averaging sequence for N. By the Howe–Moore theorem, N has the automatic ergodicity property as a subgroup of G. Therefore $\{\eta_{r,\epsilon}\}_{r>0}$ is an L^1-good averaging family for G. Proposition 2.3 implies that $\{\sigma_{r,\epsilon}\}_{r>0}$ satisfies the strong (p, p)-type maximal inequality and the $L \log L$ maximal inequality. Since $\{\eta_{r,\epsilon}\}_{r>0}$ is pointwise ergodic for bounded functions, the bounded convergence theorem implies $\{m_K * \eta_{r,\epsilon} * m_K\}_{r>0}$ is also pointwise ergodic for bounded functions. Since $L^\infty(X, \mu)$ is dense in $L^1(X, \mu)$ (for any probability space (X, μ)), Theorem 2.4 now implies $\{m_K * \eta_{r,\epsilon} * m_K\}_{r>0}$ is a good averaging family. Equation (3) finishes the proof. The proof that $\{\beta_r\}_{r>0}$ is a good averaging family is similar. By Remark 4.2, it suffices to consider the case of curvature -1. □

We now formulate the following generalization of Theorem 4.1 pertaining to non-radial averages. This result will be crucial to our discussion below of sector averages.

PROPOSITION 4.4.

(1) *Let ν and λ be arbitrary Borel probability measures on K. Then $\nu * \eta_{r,\epsilon} * \lambda$ is a good averaging family.*

(2) *Assume further that ν, λ, ν_r and λ_r are probability measures on K, each is absolutely continuous to the Haar measure and $d\nu_r/dm_K \to d\nu/dm_K$, and $d\lambda_r/dm_K \to d\lambda/dm_K$ in the $L^1(K, m_K)$-norm as $r \to \infty$. If the family $\{\nu_r * \eta_{r,\epsilon} * \lambda_r\}_{r>0}$ satisfies the strong maximal inequalities in L^p, $1 < p < \infty$, and in $L(\log L)$, then it is a good averaging family.*

Proof. By the polynomially weighted Birkhoff ergodic theorem, $\{\eta_{r,\epsilon}\}_{r>0}$ is an L^1-good averaging sequence for N. The Howe–Moore theorem implies $\{\eta_{r,\epsilon}\}_{r>0}$ is a good averaging family as a family of measures on $\mathrm{PSL}_2(\mathbb{R})$. Proposition 2.3 implies $\nu * \eta_{r,\epsilon} * \lambda$ is a good averaging family. This proves (1).

As to part (2), given the maximal inequalities assumed in it, by Theorem 2.4 it suffices to prove pointwise convergence for the dense subspace $L^\infty(X)$. For every bounded function f and for almost every $x \in X$, $|\lambda_r f(x) - \lambda f(x)| \le \|\lambda_r - \lambda\|_{L^1(K)} \|f\|_{L^\infty(X)}$, so that $\|\lambda_r f - \lambda f\|_{L^\infty(X)} \le \|\lambda_r - \lambda\|_{L^1(K)} \|f\|_{L^\infty(X)}$. A similar statement holds for $\nu_r - \nu$.

Therefore, for almost every $x \in X$,

$$\begin{aligned}
&|(\nu_r * \eta_{r,\epsilon} * \lambda_r) f(x) - (\nu * \eta_{r,\epsilon} * \lambda) f(x)| \\
&\quad \le |((\nu_r - \nu) * \eta_{r,\epsilon} * \lambda_r) f(x)| + |(\nu * \eta_{r,\epsilon} * (\lambda_r - \lambda)) f(x)| \\
&\quad \le \left\| \frac{d\nu_r}{dm_K} - \frac{d\nu}{dm_K} \right\|_{L^1(K)} \|(\eta_{r,\epsilon} * \lambda_r) f\|_{L^\infty(X)} \\
&\qquad + \|\nu * \eta_{r,\epsilon}\|_{L^\infty(X) \to L^\infty(X)} \|\lambda_r f - \lambda f\|_{L^\infty(X)} \\
&\quad \le \left\| \frac{d\nu_r}{dm_K} - \frac{d\nu}{dm_K} \right\|_{L^1(K)} \|f\|_{L^\infty(X)} + \left\| \frac{d\lambda_r}{dm_K} - \frac{d\lambda}{dm_K} \right\|_{L^1(K)} \|f\|_{L^\infty(X)}
\end{aligned}$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

and the limit of the latter expression as $r \to \infty$ is zero by assumption. By part (1), $(\nu * \eta_{r,\epsilon} * \lambda) f$ is pointwise convergent a.e. to the ergodic mean. So the computation above implies $(\nu_r * \eta_{r,\epsilon} * \lambda_r) f$ is also pointwise convergent a.e. to the ergodic mean. □

4.2. *From horocycle averages to bi-sector averages.* In the present subsection we consider $G = \mathrm{PSL}_2(\mathbb{R})$, whose maximal compact subgroup is denoted by K. Recall that $A = \{a_t\}_{t\in\mathbb{R}} \leq \mathrm{PSL}_2(\mathbb{R})$ is a one-parameter subgroup satisfying $d(a_t p_0, p_0) = |t|$. For $r, \epsilon > 0$, let $\alpha_{r,\epsilon}$ denote the probability measure on $A \subset G$ given by

$$\alpha_{r,\epsilon} = \frac{\int_r^{r+\epsilon} \sinh(t)\delta_{a_t}\, dt}{\int_r^{r+\epsilon} \sinh(t)\, dt}.$$

For example, note that $m_K * \alpha_{r,\epsilon} * m_K = \sigma_{r,\epsilon}$ where m_K denotes the Haar probability measure on K.

THEOREM 4.5. *If $\nu, \lambda \ll m_K$ are probability measures with densities $d\nu/dm_K, d\lambda/dm_K \in L^\infty(K, m_K)$ and $\epsilon > 0$ then $\{\nu * \alpha_{r,\epsilon} * \lambda\}_{r>0}$ is a good averaging family.*

A special case of this theorem pertains to *bi-sector averages*, defined as follows. For Borel subsets $U, V \subset K$ with positive Haar measure and $r, \epsilon > 0$, let $G_{r,\epsilon}^{U,V}$ be the set of all $g \in G$ such that $g = ua_t v$ for some $u \in U$, $t \in [r, r+\epsilon]$ and $v \in V$. Let $\sigma_{r,\epsilon}^{U,V}$ be the measure on G equal to the Haar measure restricted to $G_{r,\epsilon}^{U,V}$ and normalized to have total mass one.

COROLLARY 4.6. *For any Borel subsets $U, V \subset K$ with positive Haar measure and any $\epsilon > 0$, $\{\sigma_{r,\epsilon}^{U,V}\}_{r>0}$ is a good averaging family.*

Proof of Corollary 4.6 from Theorem 4.5. This follows immediately from Theorem 4.5 by setting $\nu = m_K(U)^{-1}\chi_U$ and $\lambda = m_K(V)^{-1}\chi_V$. Indeed, then $\nu * \alpha_{r,\epsilon} * \lambda = \sigma_{r,\epsilon}^{U,V}$. □

We will derive Theorem 4.5 from a special case of Corollary 4.6 which we prove after the next lemma.

LEMMA 4.7. *Suppose $\{\tau_r\}_{r>0}$ and $\{\tau'_r\}_{r>0}$ are families of probability measures on G and $C_r > 1$ satisfies:*

- *$\tau_r \leq C_r \tau'_r$;*
- *C_r is uniformly bounded, and $C_r \to 1$ as $r \to \infty$;*
- *$\{\tau'_r\}_{r>0}$ is a good averaging family.*

Then $\{\tau_r\}_{r>0}$ is also a good averaging family.

Proof. It follows from the domination lemma that $\{\tau_r\}_{r>0}$ satisfies the strong-type (p, p) maximal inequalities for $1 < p \leq \infty$ and the strong-type $L \log L$ maximal inequality. Let $f \in L^\infty(X, \mu)$ be non-negative. Then $\tau'_r(f)$ converges to $\mathbb{E}[f|G]$ pointwise a.e. as $r \to \infty$. Since $\tau_r(f) \leq C_r \tau'_r(f)$ and $C_r \to 1$ as $r \to \infty$, it follows that

$$\limsup_{r\to\infty} \tau_r(f)(x) \leq \mathbb{E}[f|G](x)$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

for a.e. x. Since τ_r preserves the L^1-norm of non-negative functions,

$$\int \tau_r(f)(x)\,d\mu(x) = \int \mathbb{E}[f|G](x)\,d\mu(x).$$

By Fatou's lemma,

$$\begin{aligned}\int \mathbb{E}[f|G](x)\,d\mu(x) &= \limsup_{r\to\infty} \int \tau_r(f)(x)\,d\mu(x)\\ &\le \int \limsup_{r\to\infty} \tau_r(f)(x)\,d\mu(x) \le \int \mathbb{E}[f|G](x)\,d\mu(x).\end{aligned}$$

Thus $\limsup_{r\to\infty} \tau_r(f) = \mathbb{E}[f|G]$ a.e. Since $\|f\|_\infty - f$ is also non-negative and bounded, the same argument gives $\limsup_{r\to\infty} \tau_r(\|f\|_\infty - f) = \mathbb{E}[(\|f\|_\infty - f)|G]$ a.e. Since $\|f\|_\infty$ is constant, this implies $\liminf_{r\to\infty} \tau_r(f) = \mathbb{E}[f|G]$ a.e. and therefore $\tau_r(f)$ converges pointwise a.e. to $\mathbb{E}[f|G]$ as $r \to \infty$. By decomposing an arbitrary $f \in L^\infty(X, \mu)$ into real and imaginary parts and then into positive and negative parts, we see that $\tau_r(f)$ converges pointwise a.e. to $\mathbb{E}[f|G]$ as $r \to \infty$. Since L^∞ is dense in $L \log L$ and in L^p $(1 < p < \infty)$ the lemma now follows from Theorem 2.4. □

THEOREM 4.8. *If U, $V \subset K$ are compact sets with positive Haar measure and $\epsilon > 0$, then the family $\{\sigma_{r,\epsilon}^{U,V}\}_{r>0}$ is a good averaging family.*

The proof of Theorem 4.8 is based on the following geometric lemma. As noted in Lemma 4.2, for every $r > 0$ there is a unique positive $t = t(r)$ with $Kn_tK = Ka_rK$, namely $t = 2\sinh(r/2)$. Using Remark 4.1, let $w_r, w_r' \in K$ be the unique elements with $n_t = w_r a_r w_r'$. We will utilize the following observation on the angular components in this decomposition.

LEMMA 4.9. *For $t, r > 0$, the polar coordinate decomposition $n_t = w_r a_r w_r'$ in $\mathrm{Isom}^+(\mathbb{H}^2) = \mathrm{PSL}_2(\mathbb{R})$ satisfies that the (unique) component w_r converges to the identity element as $r \to \infty$ while the (unique) component w_r' converges to the $180°$ rotation in $\mathbb{H}^2$ with center p_0. Furthermore, w_r and w_r' are continuous functions of t.*

Proof. We use notation as in §4.1; in particular, $\mathbb{H}^2$ denotes the upper half plane model and we identify $\mathrm{Isom}^+(\mathbb{H}^2)$ with $\mathrm{PSL}_2(\mathbb{R})$ through the latter's action on $\mathbb{H}^2$ by fractional linear transformations. Then the action of w_r is given by the matrix k_θ for some $\theta = \theta_r$ and the action of w_r' by the matrix $k_{\theta'}$ for some $\theta' = \theta_r'$, with $k_\theta, k_{\theta'} \in \mathrm{SO}(2, R) \subset \mathrm{SL}_2(\mathbb{R})$. Therefore,

$$\begin{aligned}n_t i = t + i = w_r a_r i &= \frac{e^{r/2} i\cos\theta - e^{-r/2}\sin\theta}{e^{r/2} i\sin\theta + e^{-r/2}\cos\theta}\\ &= \frac{(e^r - e^{-r})\sin\theta\cos\theta}{e^r\sin^2\theta + e^{-r}\cos^2\theta} + \frac{i}{e^r\sin^2\theta + e^{-r}\cos^2\theta}.\end{aligned}$$

Thus $e^r\sin^2\theta + e^{-r}\cos^2\theta = 1$ and since $r \to \infty$, we have $\sin^2\theta \to 0$ and so $\cos^2\theta \to 1$. Thus $w_r = \{\pm k_\theta\}$ converges to $\{\pm I\}$ in $\mathrm{SL}_2(\mathbb{R})/\{\pm I\} = \mathrm{PSL}_2(\mathbb{R})$ as $r \to \infty$.

For future reference we note that since $d(n_t i, i) = d(n_{-t} i, i)$ it is geometrically clear that $t = (e^r - e^{-r})\sin\theta_r\cos\theta_r$ can be solved uniquely for any $t \in \mathbb{R} \setminus \{0\}$, and for $\pm t$ the same value of $e^{r/2}$ is obtained, together with the values θ_r and $-\theta_r$.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Writing $n_{-t} = n_t^{-1} = (w_r')^{-1} a_r^{-1} w_r^{-1}$, we have $n_{-t} i = -t + i = (w_r')^{-1} a_r^{-1} i$. Substitution in the foregoing explicit formula shows that $e^{-r} \sin^2 \theta_r' + e^r \cos^2 \theta_r' = 1$, and thus $\cos^2 \theta_r' \to 0$ and $\sin^2 \theta_r' \to 1$, as $r \to \infty$. We conclude that $w_r' \to \{\pm \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\}$ in $\mathrm{SL}_2(\mathbb{R})/\{\pm 1\}$ as $r \to \infty$. Note that since $\mathrm{SO}_2(\mathbb{R}) \to K = \mathrm{SO}_2(\mathbb{R})/\{\pm I\}$ is a double cover, the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, which defines a 90° rotation in the Euclidean plane, is mapped to a 180° rotation of the non-Euclidean plane.

Finally, the continuity of w_r and w_r' (and, of course, a_r as well) is evident from the explicit formulas given for $\cos\theta$ and $\cos\theta'$ above. □

Remark 4.3. It is elementary to check that the map $N_+ \cong (0, \infty) \to (0, \infty) \cong A_+$ given by $t \mapsto 2 \sinh^{-1}(t/2) = s(t) = s$ maps the measure $\eta(t) = t\, dt$ to the measure $\sinh(s)\, ds$. Consider the map $J : K \times N_+ \times K \to K \times A_+ \times K$ given by $J(k, n_t, k') = (kw_s, a_s, w_s' k')$, where $n_t = w_s a_s w_s'$ is the unique Cartan coordinate representation of n_t with $s = s(t) = 2 \sinh^{-1}(t/2) > 0$ (using Lemma 4.2 and Remark 4.1). Given any compact sets of positive measure $U, V \subset K$, the measure $\nu = \chi_U dm_K \times t\, dt \times \chi_V dm_K$ satisfies

$$J_*(\chi_U dm_K \times t\, dt \times \chi_V dm_K) = \int_0^\infty \chi_{Uw_s}(k) dm_K(k) \times \delta_{a_s} \times \chi_{w_s' V}(k') dm_K(k') \sinh(s)\, ds.$$

To check this identity, it suffices to test it against product functions of the form $a_1(k) b(a_s) a_2(k')$, where it follows immediately from the definition of J.

Let $\mathcal{M}$ be the multiplication map into G, so that

$$\mathcal{M}(k, n_t, k') := k n_t k' = k w_s a_s w_s' k' = \mathcal{M} \circ J(k, n_t, k'),$$

namely $\mathcal{M} = \mathcal{M} \circ J$. For any measure ν on $K \times N_+ \times K$, we have $\mathcal{M}_*(\nu) = (\mathcal{M} \circ J)_*(\nu)$. It follows from the explicit expression for $J_*(\nu)$ above that for $f \in C_c(G)$,

$$\int_K \int_K \int_{n_t \in N_+} f(k n_t k') t\, dt\, \chi_U(k) dm_K(k) \chi_V(k') dm_K(k') = \int_K \int_K \int_{a_s \in A_+} f(k w_s a_s w_s' k') \sinh(s)\, ds\, \chi_U(k) dm_K(k) \chi_V(k') dm_K(k').$$

Proof of Theorem 4.8. Fix $\epsilon > 0$. As noted above, for $r, t > 0$ there are unique elements $w_r, w_r' \in K$ such that $n_t = w_r a_r w_r'$, with $t = 2 \sinh r/2$. Define $U_r = \bigcup_{r \le s < r+\epsilon} U w_s^{-1}$ and $V_r = \bigcup_{r \le s < r+\epsilon} (w_s')^{-1} V$. Let ν_r be the normalized restriction of m_K to U_r and λ_r be the normalized restriction of m_K to V_r. We will show that there is a constant $C_r > 1$ such that $\lim_{r \to \infty} C_r = 1$ and

$$\sigma_{r,\epsilon}^{U,V} \le C_r \nu_r * \eta_{r,\epsilon} * \lambda_r,$$

where $\eta_{r,\epsilon}$ is given in Definition 2 above. We prove the above inequality by comparing Radon–Nikodym derivatives of the two measures in question, for each given r. Using the formula for the Haar measure on G in polar coordinates, we have, for $f \in C_c(G)$,

$$\sigma_{r,\epsilon}^{U,V}(f) = \int_{k \in K} \int_{k' \in K} \int_{s \in [r, r+\epsilon)} f(k a_s k') \frac{\sinh s\, ds}{\cosh(r+\epsilon) - \cosh r} \times \frac{\chi_U(k)\, dm_K(k)}{m_K(U)} \frac{\chi_V(k') dm_K(k')}{m_K(V)}.$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

On the other hand, by definition of convolution,

$$\nu_r * \eta_{r,\epsilon} * \lambda_r(f) = \int_{k\in K}\int_{k'\in K}\int_{2\sinh(r/2)}^{2\sinh((r+\epsilon)/2)} f(kn_tk')\frac{t\,dt}{\cosh(r+\epsilon)-\cosh r} \times \frac{\chi_{U_r}(k)\,dm_K(k)}{m_K(U_r)}\frac{\chi_{V_r}(k')\,dm_K(k')}{m_K(V_r)},$$

and using Remark 4.3,

$$= \int_{k\in K}\int_{k'\in K}\int_{s\in[r,r+\epsilon)} f(kw_sa_sw_s'k')\frac{\sinh s\,ds}{\cosh(r+\epsilon)-\cosh r} \times \frac{\chi_{U_r}(k)dm_K(k)}{m_K(U_r)}\frac{\chi_{V_r}(k')dm_K(k')}{m_K(V_r)}.$$

Note that the support of $\sigma_{r,\epsilon}^{U,V}$ is contained in the support of the convolution above, by definition of U_r and V_r. Furthermore,

$$\frac{d\sigma_{r,\epsilon}^{U,V}}{d(\nu_r * \eta_{r,\epsilon} * \lambda_r)}(g) = \frac{m_K(U_r)m_K(V_r)}{m_K(U)m_K(V)} =: C_r,$$

and since $w_r \to 1$ and w_r' tends to the 180° rotation as $r \to \infty$ (by Lemma 4.9), it follows that $C_r \to 1$ as $r \to \infty$. Indeed, since U is compact and $s \mapsto w_s$ is continuous, the set

$$U_r' := \bigcup_{r\le s\le r+\epsilon} U w_s^{-1} w_r$$

is compact, $m_K(U) \le m_K(U_r) \le m_K(U_r')$. Moreover, $U \subset U_r'$ and U_r' is contained in the $\delta(r)$-neighborhood of U for some $\delta(r) > 0$ satisfying $\lim_{r\to\infty}\delta(r) = 0$ (by Lemma 4.9). Since the intersection of these neighborhoods is U, it follows that $m_K(U_r) \to m_K(U)$ as $r \to \infty$. Similarly, $m_K(V_r) \to m_K(V)$ as $r \to \infty$.

To complete the proof of the Theorem 4.8 it suffices, by Lemma 4.7 (setting $\tau_r = \sigma_{r,\epsilon}^{U,V}$ and $\tau_r' = \nu_r * \eta_{r,\epsilon} * \lambda_r$), to establish the conclusions for $\nu_r * \eta_{r,\epsilon} * \lambda_r$. By Proposition 4.4, $m_K * \eta_{r,\epsilon} * m_K$ is a good averaging family. Since, for all $r > 1$,

$$\nu_r * \eta_{r,\epsilon} * \lambda_r \le \frac{1}{m_K(U_r)m_K(V_r)} m_K * \eta_{r,\epsilon} * m_K \le \frac{C}{m_K(U)m_K(V)} m_K * \eta_{r,\epsilon} * m_K$$

for some $C > 0$, the domination lemma implies that $r \mapsto \nu_r * \eta_{r,\epsilon} * \lambda_r$ satisfies the strong-type (p, p), $1 < p < \infty$, and $L\log L$ maximal inequalities.

Let ν denote the normalized restriction of m_K to U and λ denote the normalized restriction of m_K to V. Then $d\nu_r/dm_K \to d\nu/dm_K$, $d\lambda_r/dm_K \to d\lambda/dm_K$ in $L^1(K)$ norm. So Proposition 4.4 implies $r \mapsto \nu_r * \eta_{r,\epsilon} * \lambda_r$ is a good averaging family. □

We now pass from $\sigma_{r,\epsilon}^{U,V}$ to averages defined by arbitrary densities on K, as follows.

LEMMA 4.10. *Suppose $\{\tau_r\}_{r>0}$ and $\{\tau_{n,r}'\}_{n\in\mathbb{N},r>0}$ are families of probability measures on G and $C_n > 1$ satisfies:*

- *$\tau_r \le C_n\tau_{n,r}'$ for all r, n;*
- *$C_n \to 1$ as $n \to \infty$;*
- *for each $n \in \mathbb{N}$, $\{\tau_{n,r}'\}_{r>0}$ is a good averaging family.*

Then $\{\tau_r\}_{r>0}$ is also a good averaging family.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Proof. It follows from the domination lemma that $\{\tau_r\}_{r>0}$ satisfies the strong-type $L\log L$ maximal inequality and the strong-type (p,p) maximal inequalities for $1<p<\infty$. Let $f\in L^\infty(X,\mu)$ be non-negative. Then $\tau'_{n,r}(f)$ converges to $\mathbb{E}[f|G]$ pointwise a.e. as $r\to\infty$. Since $\tau_r(f)\le C_n\tau'_{n,r}(f)$ it follows that $\limsup_{r\to\infty}\tau_r(f)(x)\le \limsup_{n\to\infty} C_n\cdot\mathbb{E}[f|G](x)$ for almost every x, and since $C_n\to 1$ as $n\to\infty$ we have $\limsup_{r\to\infty}\tau_r(f)(x)\le\mathbb{E}[f|G](x)$ for almost every x. The proof is now identical to the end of the proof of Lemma 4.7. □

Proof of Theorem 4.5. Let $A,B\subset K$ be open sets whose complements $U:=K-A$, $V:=K-B$ have positive measure. By Theorem 4.8, $\{\sigma^{U,K}_{r,\epsilon}\}_{r>0}$, $\{\sigma^{K,V}_{r,\epsilon}\}_{r>0}$ and $\{\sigma^{U,V}_{r,\epsilon}\}_{r>0}$ are good averaging families. Since

$$\sigma^{A,B}_{r,\epsilon}=\frac{\sigma_{r,\epsilon}-m_K(U)\sigma^{U,K}_{r,\epsilon}-m_K(V)\sigma^{K,V}_{r,\epsilon}+m_K(U)m_K(V)\sigma^{U,V}_{r,\epsilon}}{1-m_K(U)-m_K(V)+m_K(U)m_K(V)},$$

it follows that $\{\sigma^{A,B}_{r,\epsilon}\}_{r>0}$ is also a good averaging family.

Now let $A,B\subset K$ be Borel sets with positive measure. We will show that $\{\sigma^{A,B}_{r,\epsilon}\}_{r>0}$ is a good averaging family. For each $n>0$ there exist open sets $U_n\supset A$ and $V_n\supset B$ such that $m_K(U_n\setminus A)<1/n$ and $m_K(V\setminus B)<1/n$. By Lemma 4.8 $\{\sigma^{U_n,V_n}_{r,\epsilon}\}_{r>0}$ is a good averaging family. Since

$$\sigma^{A,B}_{r,\epsilon}\le\frac{m_K(U_n)m_K(V_n)}{m_K(A)m_K(B)}\sigma^{U_n,V_n}_{r,\epsilon},$$

it follows from Lemma 4.10 that $\{\sigma^{A,B}_{r,\epsilon}\}_{r>0}$ is a good averaging family.

Now let ν and λ be arbitrary probability measures on K_0 with bounded densities, namely $d\nu/dm_K, d\lambda/dm_K\in L^\infty(K)$. Recall that a simple function is a finite linear combination of characteristic functions of Borel subsets. Since $d\nu/dm_K, d\lambda/dm_K$ are essentially bounded, for any $n\in\mathbb{N}$ there exist simple functions $y_{\nu,n},\ y_{\lambda,n}\in L^\infty(K,m_K)$ such that $y_{\nu,n}\ge d\nu/dm_K$, $y_{\lambda,n}\ge d\lambda/dm_K$ and $\|y_{\nu,n}-d\nu/dm_K\|_\infty\le 1/n$, $\|y_{\lambda,n}-d\lambda/dm_K\|_\infty\le 1/n$. Then $y_{\nu,n}/\|y_{\nu,n}\|_1$ and $y_{\lambda,n}/\|y_{\lambda,n}\|_1$ are probability densities and simple functions. Denoting the probabilities they define by ν_n and λ_n, clearly $\nu\le(1+1/n)\nu_n$ and $\lambda\le(1+1/n)\lambda$. Because $y_{\nu,n}, y_{\lambda,n}$ are simple it follows from the previous paragraph and linearity that $\{\nu_n*\alpha_{r,\epsilon}*\lambda_n\}_{r>0}$ is a good averaging family for each n. Since $\nu\le(1+1/n)\nu_n$, $\lambda\le(1+1/n)\lambda_n$, it follows that

$$\nu*\alpha_{r,\epsilon}*\lambda\le(1+1/n)^2\nu_n*\alpha_{r,\epsilon}*\lambda_n.$$

So Lemma 4.10 implies $\{\nu*\alpha_{r,\epsilon}*\lambda\}_{r>0}$ is a good averaging family. □

5. *Ergodic theorems for general real rank-one groups*

5.1. *Structure theory for real rank-one groups.* In the present section we will extend Theorem 1.1 to general real rank-one groups using the method of rotations, applied to totally geodesic embeddings. We assume that G is a real rank-one connected non-compact simple Lie group with finite center. In the present section our notation will be different from the notation used thus far, where K, A and N denoted specific subgroups of $\mathrm{SL}_2(\mathbb{R})$. We now fix a maximal compact subgroup of G and denoted it by K, and a one-parameter

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

subgroup $A \cong \mathbb{R}$ of G such that $G = KAK$ is a Cartan decomposition. We let N be the horospherical subgroup of G associated with A, so that $G = KAN$ is an Iwasawa decomposition.

Let $\mathfrak{g}$ denote the Lie algebra of G. Fix a Cartan involution θ on G and $\mathfrak{g}$, and let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the associated Cartan decomposition of $\mathfrak{g}$ to the ± 1 eigenspaces of θ. Choose a maximal Abelian subalgebra $\mathfrak{a}$ contained in $\mathfrak{p}$. Because G has real rank one, $\dim_{\mathbb{R}} \mathfrak{a} = 1$. Let $\mathfrak{a}^* = \mathrm{Hom}(\mathfrak{a}, \mathbb{R})$ denote the real dual of $\mathfrak{a}$, and let $\Sigma = \Sigma(\mathfrak{a}, \mathfrak{g}) \subset \mathfrak{a}^*$ denote the set of non-zero roots of $\mathfrak{a}$ in $\mathfrak{g}$. Because G has real rank one, $\Sigma = \{\pm\alpha\}$ for some $\alpha \in \mathfrak{a}^*$, or $\Sigma = \{\pm\alpha, \pm 2\alpha\}$. The Weyl group $W = W(\mathfrak{a}, \mathfrak{g})$ is isomorphic to $\mathbb{Z}_2$ in both cases, and its non-trivial element acts as multiplication by -1 on $\mathfrak{a}$. The adjoint action of the Lie algebra $\mathfrak{a}$ on $\mathfrak{g}$ is diagonalizable, with the eigenspaces being $\mathfrak{g}_{\pm\alpha}$, $\mathfrak{g}_{\pm 2\alpha}$ (when non-empty), and $\mathfrak{g}_0$. $\mathfrak{g}$ is the direct sum of these subspaces, and $\mathfrak{g}_0 = \mathfrak{a} \oplus \mathfrak{m}$, where $\mathfrak{m}$ is the centralizer of $\mathfrak{a}$ in $\mathfrak{g}$. Denote $m_1 = \dim_{\mathbb{R}} \mathfrak{g}_\alpha$, $m_2 = \dim_{\mathbb{R}} \mathfrak{g}_{2\alpha}$. We fix an element $H_1 \in \mathfrak{a}$, satisfying $\alpha(H_1) = 1$, so that $\{e^{tH_1}\}_{t\in\mathbb{R}}$ is a parametrization of A.

LEMMA 5.1. (KAK decomposition) *Let m_K denote the Haar measure on K normalized to have total mass one. Let $m_1, m_2 \geq 0$ be as above and let m_G denote the measure on G defined by*

$$\int F(g)\, dm_G(g) = \int_K \int_0^\infty \int_K F(k_1 e^{tH_1} k_2) \sinh(t)^{m_1+m_2} \cosh(t)^{m_2}\, dm_K(k_1)\, dt dm_K(k_2).$$

Then m_G is a Haar measure on G.

Proof. For this well-known formula, see, for example, [**He94**] or [**Ko84**, equations (2.5), (4.8)]. □

We now choose the subgroup $L \subset G$ to which we will apply the method of rotations, using [**Kn96**, Proposition 6.52, p. 321] in our discussion. If $\mathfrak{g}_{2\alpha} = 0$, let $X_\alpha \in \mathfrak{g}_\alpha$ be any non-zero vector, and let $\mathfrak{l}$ be the Lie algebra spanned by X, $Y = \theta(X)$ and $H = [X, Y]$. Then $\mathfrak{l}$ is a Lie algebra isomorphic to $\mathfrak{sl}_2(\mathbb{R})$, and it is invariant under θ. The restriction of θ to $\mathfrak{l}$ is a Cartan involution of $\mathfrak{l}$, and $\mathfrak{a}$ is contained in $\mathfrak{l}$ and spanned by H. Multiplying X by a suitable multiple if necessary, we can assume that the map $E_{1,2} \mapsto X$, $E_{2,1} \mapsto Y$, $\mathrm{diag}(1/2, -1/2) \mapsto H_1$ is a Lie algebra isomorphism $\tau : \mathfrak{sl}_2(\mathbb{R}) \to \mathfrak{l}$. Here $E_{i,j}$ is the elementary 2×2 matrix with 1 as (i, j)th entry.

If $\mathfrak{g}_{2\alpha} \neq 0$, we choose any non-zero $X \in \mathfrak{g}_{2\alpha}$, and consider the Lie algebra $\mathfrak{l}$ spanned by X, $Y = \theta(X)$ and $H = [X, Y]$. Again $\mathfrak{l}$ is isomorphic with $\mathfrak{sl}_2(\mathbb{R})$ and contains $\mathfrak{a}$. Note, however, that the element $H_1 \in \mathfrak{a}$ we chose above to parametrize A now has the following property. When viewed as an element of the $\mathbb{R}$-split Cartan subalgebra $\mathfrak{a}$ of $\mathfrak{l}$, the evaluation of the unique root of $\mathfrak{a}$ (in $\mathfrak{l}$) on H_1 gives the value 2, and not 1. Thus, multiplying X by a suitable multiple if necessary, we can assume that the Lie algebra isomorphism $\tau : \mathfrak{sl}_2(\mathbb{R}) \to \mathfrak{l}$ is given now by $E_{1,2} \mapsto X$, $E_{2,1} \mapsto Y$, $\mathrm{diag}(1, -1) \mapsto H_1$.

We let L denote the closed subgroup of G with Lie algebra $\mathfrak{l}$, and then L is isomorphic to a finite covering group of $\mathrm{PSL}_2(\mathbb{R})$. We denote $K_L := K \cap L$, $N_L := N \cap L$, and then $L = K_L A N_L$ is an Iwasawa decomposition of L, and $L = K_L A K_L$ is a Cartan decomposition.

The restriction of (any multiple of) the Killing form on the Lie algebra $\mathfrak{g}$ to the Lie algebra $\mathfrak{l}$ is a non-degenerate invariant form, and hence a multiple of the Killing form on $\mathfrak{l}$.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

Pulling back this form to $\mathfrak{sl}_2(\mathbb{R})$ via the representation τ, we obtain a multiple of the Killing form on $\mathfrak{sl}_2(\mathbb{R})$, and upon restriction also a multiple $c\mathcal{G}$ of the Riemannian metric $\mathcal{G}$ on $\mathbb{H}^2$ used in §4. In the first case, when $\mathfrak{g}_{2\alpha} = 0$, the multiple is clearly $c = 1$, and in the second case, when $\mathfrak{g}_{2\alpha} \neq 0$, the multiple is clearly $c = 1/2$.

Consider now the case where G is an adjoint group, namely it is the unique group with trivial center in the class of groups with isomorphic Lie algebras. It then follows that L is in fact isomorphic to $\mathrm{PSL}_2(\mathbb{R})$ itself. This fact can be verified directly using the explicit formulas for the action of the isometry groups of hyperbolic spaces stated in [**BH99**, Ch. II.10.25].

LEMMA 5.2. *Let $\tau : PSL_2(\mathbb{R}) \to L \subset G$ be the representation constructed above, with G an adjoint group. Let $N_L = \{n_t^\tau = \tau(n_t)\}_{t\in\mathbb{R}}$, $A_L = \{a_r^\tau = \tau(a_r)\}_{r\in\mathbb{R}} = A$ where $n_t, a_r \in PSL_2(\mathbb{R})$ are the parametrizations indicated in §4.1. There exists a positive constant $c = c_\tau$ such that for all $t, r > 0$ the following are equivalent:*

(1) $K_L n_t^\tau K_L = K_L a_{r/c}^\tau K_L$;
(2) $t = 2\sinh(r/2c)$;
(3) $\cosh(r/c) = 1 + t^2/2$;
(4) $K n_t^\tau K = K a_{r/c}^\tau K$.

Proof. Let $p_0 \in G/K$ be the unique point in the symmetric space G/K with stability group K, so that the stability group of p_0 in L is $K_L = K \cap L$. We have, by definition, $d_{G/K}(\tau(y)p_0, p_0) = d_c(y \cdot o, o) = cd(y \cdot o, o)$ for all $y \in \mathrm{SL}_2(\mathbb{R})$, where $d_{G/K}$ is the invariant metric on G/K, d the metric on $\mathbb{H}^2_{-1}$ associated with constant curvature -1, and o a suitable reference point. The fact that (2) and (3) are equivalent to $K_L n_t^\tau K_L = K_L a_{r/c}^\tau K_L$ follows immediately from our discussion of the $\mathrm{PSL}_2(\mathbb{R})$ case in Lemma 4.2 and Remark 4.2. It remains to show that $K_L n_t^\tau K_L = K_L a_{r/c}^\tau K_L$ follows from $K n_t^\tau K = K a_{r/c}^\tau K$. It is well known (see, for example, [**He94**]) that the radial component of the Cartan decomposition in the real rank-one group G is determined uniquely. This is equivalent to the fact that in the symmetric space G/K we have $d_{G/K}(gp_0, p_0) = d_{G/K}(hp_0, p_0)$ if and only if $KgK = KhK$. Thus if $K n_t^\tau K = K a_{r/c}^\tau K$ then $d_{G/K}(n_t^\tau p_0, p_0) = d_{G/K}(a_{r/c}^\tau p_0, p_0)$. The distance $d_{G/K}$ restricts to a distance on the totally geodesic hyperbolic plane $L \cdot p_0 \cong L/K_L$. Using the parametrization of this plane via the representation τ of $\mathrm{SL}_2(\mathbb{R})$, by Remark 4.2 it follows that $K_L n_t^\tau K_L = K_L a_{r/c}^\tau K_L$. □

Thus $A \cong \mathbb{R}$ is parametrized by $\{a_r^\tau\}_{r\in\mathbb{R}}$, and also by $\{e^{rH_1}\}_{r\in\mathbb{R}}$. These parametrizations are identical when $\mathfrak{g}_{2\alpha} = 0$, but otherwise they are different and satisfy $\tau(a_{2t}) = a_{2t}^\tau = e^{tH_1}$.

LEMMA 5.3. (KN_LK decomposition) *Let G be a connected simple adjoint Lie group of real rank one and finite center, and $L \subset G$ chosen as above. Then $G = KN_LK$, and there exists a function ψ on $[0, \infty)$ satisfying, for any bounded measurable function F on G with compact support,*

$$\int_G F(g)\, dm_G(g) = \int_K \int_0^\infty \int_K F(k_1 n_t^\tau k_2)\psi(t)\, dm_K(k_1)\, dt\, dm_K(k_2),$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

and ψ has the following asymptotic form:

(1) *when $m_2 > 0$, namely when $\mathfrak{g}_{2\alpha} \neq 0$,*

$$\psi(T) = C_G T^{m_1+2m_2-1} + O(T^{m_1+2m_2-2}) \quad \text{for } T \geq 1,$$

(2) *when $m_2 = 0$, namely when $\mathfrak{g} \cong \mathfrak{so}(m_1+1, 1)$*

$$\psi(T) = C_G T^{2m_1-1} + O(T^{2m_1-2}) \quad \text{for } T \geq 1.$$

Proof. Let us consider first the case where $\mathfrak{g}_{2\alpha} \neq 0$. Define ψ by

$$\psi(T) = \frac{\sinh^{m_1+m_2}(R)\cosh^{m_2}(R)}{2\cosh(R)}$$

where $T = 2\sinh(R)$. ψ is well defined since $R \mapsto 2\sinh(R)$ is invertible on $[0, \infty)$. Now since

$$\sinh^{m_1+m_2}(R)\cosh^{m_2}(R) = 2\psi(T)\cosh(R)$$

we can conclude

$$\int_0^R \sinh^{m_1+m_2}(r)\cosh^{m_2}(r)\,dr = \int_0^{T(R)} \psi(t)\,dt$$

upon differentiating both sides with respect to R, and using $dT(R)/dR = 2\cosh(R)$.

Suppose χ_{B_R} is the characteristic function of a ball of radius R in G/K with center p_0. Then by Lemmas 5.1 and 5.2, since $d_{G/K}(n_t^\tau p_0,\, p_0) = r \iff t = 2\sinh(r)$,

$$\begin{aligned}\int_G \chi_{B_R}(g)\,dm_G(g) &= \int_K \int_0^R \int_K \chi_{B_R}(k_1 a_{2r}^\tau k_2)\sinh^{m_1+m_2}(r)\cosh^{m_2}(r)\,dk_1\,dr\,dk_2 \\ &= \int_K \int_0^{T(R)} \int_K \chi_{B_R}(k_1 n_t^\tau k_2)\psi(t)\,dm_K(k_1)\,dt\,dm_K(k_2) \\ &= \int_K \int_0^{\infty} \int_K \chi_{B_R}(k_1 n_t^\tau k_2)\psi(t)\,dm_K(k_1)\,dt\,dm_K(k_2).\end{aligned}$$

Therefore this formula holds for all radial functions. Because the measure on the right-hand side is bi-K-invariant, this formula must hold for all bounded measurable functions with compact support.

The formula $G = KN_LK$ is immediate from Lemma 5.2. It remains to prove the asymptotic formula for ψ. Clearly $\sinh^2(R) = \cosh^2(R) - 1 = T^2/4$ implies $\cosh(R) = \sqrt{T^2/4+1}$, so we obtain

$$\begin{aligned}\psi(T) &= 2^{-(m_1+m_2+1)} T^{m_1+m_2}(T^2/4+1)^{(m_2-1)/2} \\ &= C_G' T^{m_1+2m_2-1} + O(T^{m_1+2m_2-2}), \quad \text{for } T \geq 1,\end{aligned}$$

where $C_G > 0$ is a constant depending only on G.

The case where $\mathfrak{g}_{2\alpha} = 0$ is handled similarly, defining $\psi(T) = \sinh^{m_1}(R)/\cosh(R/2)$, with $T = 2\sinh R/2$. Then $\int_0^R \sinh^{m_1} r\,dr = \int_0^{T(R)} \psi(t)\,dt$, and, using $\sinh R = 2\sinh R/2\cosh R/2$, we have $\psi(T) = 2^{m_1} T^{m_1}(\sqrt{T^2/4+1})^{m_1-1}$ so that $\psi(T) = C_G T^{2m_1-1} + O(T^{2m_1-2})$ for $T \geq 1$. □

We note that for a finite cover $\tau' : L \to \mathrm{PSL}_2(\mathbb{R})$ the kernel is central, and so the inverse image of the subgroups A and N of $\mathrm{PSL}_2(\mathbb{R})$, denoted by $a_r^{\tau'}$ and $n_t^{\tau'}$, are isomorphic to A and N. Using these subgroups of L in the foregoing argument, we see that Theorem 5.3 holds for any finite cover group, not just the adjoint group.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

5.2. *Proof of the ergodic theorems for real rank-one groups.* Let us prove the ergodic theorems for averages on real rank-one groups, starting with the radial case. We will start by assuming G is the adjoint group, and complete the argument for the general case at the end of the section.

THEOREM 5.4. *Let G be a connected simple Lie group of real rank and finite center, and fix any invariant Riemannian metric on the symmetric space G/K. Then $\sigma_{r,\epsilon}$ and β_r are good averaging families, for every fixed $\epsilon > 0$.*

Proof. When G is the adjoint group, the proof is virtually the same as the proof of Theorem 4.1 when $\mathfrak{g}_{2\alpha} = 0$. For completeness we provide the details in the case $\mathfrak{g}_{2\alpha} \neq 0$. Let η be the measure on N_L defined by

$$\eta(E) = \int 1_E(n_t^\tau)\psi(t)\, dt$$

where ψ is the function defined in Lemma 5.3 and n_t^τ is as in Lemma 5.2. Let $\eta_{R,\epsilon}$ be the measure on N_L defined by

$$\eta_{R,\epsilon}(E) = \frac{\eta(E \cap \{n_t^\tau : t \in [2\sinh(R), 2\sinh(R+\epsilon)]\})}{\eta(\{n_t^\tau : t \in [2\sinh(R), 2\sinh(R+\epsilon)]\})}.$$

Theorem 2.5 implies $\{\eta_{R,\epsilon}\}_{R>0}$ is an L^1-good averaging family for N_L. By the Howe–Moore theorem, $\{\eta_{R,\epsilon}\}_{R>0}$ is an L^1-averaging family for G.

By Lemma 5.3, $m_K * \eta_{R,\epsilon} * m_K = \sigma_{R,\epsilon}$. So Proposition 2.3 now implies that $\{\sigma_{R,\epsilon}\}_{R>0}$ satisfies the strong (p, p) type maximal inequality $(p > 1)$ and the $L \log L$ maximal inequality. The bounded convergence theorem implies that $\{\sigma_{R,\epsilon}\}_{R>0}$ is pointwise ergodic in L^∞. So Theorem 2.4 implies $\{\sigma_{R,\epsilon}\}_{R>0}$ is pointwise ergodic in L^p for all $p > 1$ and in $L \log L$. The case of $\{\beta_r\}_{r>0}$ is handled similarly. For the case where G has finite center, see below. □

We now turn to the proof of Theorem 1.1 for non-radial averages. For $r, \epsilon > 0$, let $\alpha_{r,\epsilon}$ denote the probability measure on $A \subset G$ given by

$$\alpha_{r,\epsilon} = \frac{\int_r^{r+\epsilon} \sinh(t)^{m_1+m_2} \cosh(t)^{m_2} \delta_{e^{tH_1}}\, dt}{\int_r^{r+\epsilon} \sinh(t)^{m_1+m_2} \cosh(t)^{m_2}\, dt}.$$

For example, note that $m_K * \alpha_{r,\epsilon} * m_K = \sigma_{r,\epsilon}$ where m_K denotes the Haar probability measure on K. Recall the definition of $\sigma_{r,\epsilon}^{U,V}$ from §1.2. We first prove a special case of Theorem 1.1.

THEOREM 5.5. *For any compact Z-invariant subsets $U, V \subset K$ both with positive measure, the families $\{\sigma_{r,\epsilon}^{U,V}\}_{r>0}$ and $\{\beta_r^{U,V}\}_{r>0}$ are both good averaging families.*

Proof. We first assume that G is an adjoint group. The proof then is very similar to the proof of Theorem 4.8 when $\mathfrak{g}_{2\alpha} = 0$, and we provide the details when $\mathfrak{g}_{2\alpha} \neq 0$. If $t, r > 0$ are such that $Kn_t^\tau K = Ka_{2r}^\tau K$ then let $w_r, w_r' \in K \cap L$ be the elements satisfying $n_t^\tau = w_r a_{2r}^\tau w_r'$. Note the important fact that these identities hold in the subgroup $L \cong \mathrm{PSL}_2(\mathbb{R})$, and hence a_{2r}, w_r and w_r' are unique and form continuous functions of t. Define $U_r = \bigcup_{r \le s < r+\epsilon} U w_s^{-1}$ and $V_r = \bigcup_{r \le s < r+\epsilon} (w_s')^{-1} V$. Let ν_r be the normalized restriction of

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

m_K to U_r and λ_r be the normalized restriction of m_K to V_r. We will show that there is a constant $C_r > 1$ such that $\lim_{r\to\infty} C_r = 1$ and

$$\sigma_{r,\epsilon}^{U,V} \le C_r \nu_r * \eta_{r,\epsilon} * \lambda_r$$

where $\eta_{r,\epsilon}$ is as in the proof of Theorem 5.4.

Using the formula for the Haar measure on G in polar coordinates, we have, for any bounded measurable function f on G,

$$\sigma_{r,\epsilon}^{U,V}(f) = \int_{k\in K}\int_{k'\in K}\int_r^{r+\epsilon} f(ke^{sH_1}k')\frac{\sinh(s)^{m_1+m_2}\cosh(s)^{m_2}\,ds}{\int_r^{r+\epsilon}\sinh(s)^{m_1+m_2}\cosh(s)^{m_2}\,ds} \times \frac{\chi_U(k)\,dm_K(k)}{m_K(U)}\frac{\chi_V(k')\,dm_K(k')}{m_K(V)}.$$

On the other hand, by definition of convolution,

$$\nu_r * \eta_{r,\epsilon} * \lambda_r(f) = \int_{k\in K}\int_{k'\in K}\int_{2\sinh(r)}^{2\sinh(r+\epsilon)} f(kn_t^\tau k')\frac{\psi(t)\,dt}{\int_{2\sinh(r)}^{2\sinh(r+\epsilon)}\psi(t)\,dt} \times \frac{\chi_{U_r}(k)\,dm_K(k)}{m_K(U_r)}\frac{\chi_{V_r}(k')\,dm_K(k')}{m_K(V_r)},$$

and, using Lemma 5.3,

$$= \int_{k\in K}\int_{k'\in K}\int_r^{r+\epsilon} f(kw_s a_{2s}^\tau w_s' k')\frac{\sinh(s)^{m_1+m_2}\cosh(s)^{m_2}\,ds}{\int_r^{r+\epsilon}\sinh(s)^{m_1+m_2}\cosh(s)^{m_2}\,ds} \times \frac{\chi_{U_r}(k)\,dm_K(k)}{m_K(U_r)}\frac{\chi_{V_r}(k')\,dm_K(k')}{m_K(V_r)}.$$

Note that the support of $\sigma_{r,\epsilon}^{U,V}$ is contained in the support of the convolution above, by definition of U_r and V_r. Furthermore,

$$\frac{d\sigma_{r,\epsilon}^{U,V}}{d(\nu_r * \eta_{r,\epsilon} * \lambda_r)}(g) = \frac{m_K(U_r)m_K(V_r)}{m_K(U)m_K(V)} =: C_r,$$

and since $w_r \to 1$ and w_r' tends to the 180° rotation as $r \to \infty$ (by Lemma 4.9), it follows that $C_r \to 1$ as $r \to \infty$. Indeed, since U is compact and $s \mapsto w_s$ is continuous, the set

$$U_r' := \bigcup_{r\le s\le r+\epsilon} U w_s^{-1} w_r$$

is compact, $m_K(U) \le m_K(U_r) \le m_K(U_r')$. Moreover, $U \subset U_r'$ and U_r' is contained in the $\delta(r)$-neighborhood of U for some $\delta(r) > 0$ satisfying $\lim_{r\to\infty}\delta(r) = 0$ (by Lemma 4.9). Since the intersection of these neighborhoods is U, it follows that $m_K(U_r) \to m_K(U)$ as $r \to \infty$. Similarly, $m_K(V_r) \to m_K(V)$ as $r \to \infty$.

To complete the proof it suffices, by Lemma 4.7 (setting the averages τ_r and τ_r' that appear there as $\tau_r = \sigma_{r,\epsilon}^{U,V}$ and $\tau_r' = \nu_r * \eta_{r,\epsilon} * \lambda_r$), to establish the conclusions for $\nu_r * \eta_{r,\epsilon} * \lambda_r$. By Proposition 4.4 (which holds for general real rank-one groups with $\eta_{r,\epsilon}$ as in the proof of Theorem 5.4 by exactly the same argument), $\{m_K * \eta_{r,\epsilon} * m_K\}_{r>0}$ is a good averaging family. Since, for all $r > 1$,

$$\nu_r * \eta_{r,\epsilon} * \lambda_r \le \frac{1}{m_K(U_r)m_K(V_r)} m_K * \eta_{r,\epsilon} * m_K \le \frac{C}{m_K(U)m_K(V)} m_K * \eta_{r,\epsilon} * m_K$$

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

for some $C > 0$, the domination lemma implies that $r \mapsto \nu_r * \eta_{r,\epsilon} * \lambda_r$ satisfies the strong-type (p, p), $1 < p < \infty$, and $L \log L$ maximal inequalities.

Let ν denote the normalized restriction of m_K to U and λ denote the normalized restriction of m_K to V. Then $d\nu_r/dm_K \to d\nu/dm_K$, $d\lambda_r/dm_K \to d\lambda/dm_K$ in $L^1(K)$ norm. So Proposition 4.4 implies $r \mapsto \nu_r * \eta_{r,\epsilon} * \lambda_r$ is a good averaging family. □

THEOREM 5.6. *As above, let G be a connected non-compact simple Lie group of real rank one with finite center. Fix a maximal compact subgroup K. If $\nu, \lambda \ll m_K$ are Z-invariant probability measures with densities $d\nu/dm_K$, $d\lambda/dm_K \in L^\infty(K, m_K)$ and $\epsilon > 0$ then $\{\nu * \alpha_{r,\epsilon} * \lambda\}_{r>0}$ is a good averaging family.*

Proof. When G is adjoint, the proof is essentially the same as the proof of Theorem 4.5 using $\{n_t^\tau\}$ in place of $\{n_t\}$. □

Proof of Theorem 1.1. For adjoint groups Theorem 1.1 follows immediately from Theorem 5.6 by setting $\nu = m_K(U)^{-1}\chi_U$ and $\lambda = m_K(V)^{-1}\chi_V$. Indeed, then $\nu * \alpha_{r,\epsilon} * \lambda = \sigma_{r,\epsilon}^{U,V}$. We now finally turn to consider groups with finite center. Let G have a finite center Z, and let (X, μ) be an ergodic probability-measure-preserving action. The function space $L^2(X)$ decomposes into a finite direct sum of closed subspaces $\mathcal{V}_\chi$, with χ ranging over the characters of the finite Abelian group Z, and for a function $f \in \mathcal{V}_\chi$, we have $f(zx) = \chi(z)f(x)$. Z being central, each space $\mathcal{V}_\chi$ is G-invariant, and it remains to show the pointwise convergence of the averages in question for functions in each $\mathcal{V}_\chi$, $\chi \in Z^*$. But since U and V are Z-invariant sets and Z is central, it follows from the fact that $\sum_{z\in Z} \chi(z) = 0$ for a character $\chi \neq 1$ of Z, that the action of the corresponding averages annihilate each $\mathcal{V}_\chi$, unless χ is the trivial character, denoted 1. The functions in $\mathcal{V}_\chi$, $\chi \neq 1$, have zero integral on X, so the pointwise ergodic theorem holds in these subspaces.

The function space $\mathcal{V}_1$ consists of Z-invariant functions on X, and is naturally identified with the function space $L^2(X/Z)$, where X/Z is the space of orbits of Z in X. The representation of G on this function space is such that center acts trivially, and it is equivalent to the representation that arises from the action of G/Z on the space X/Z of Z-orbits in X. The desired convergence results then follow from our previous arguments for the adjoint group G/Z.

This concludes the proof of Theorem 1.1. □

Remark 5.1. Case of general averages. We note that the assumption that the sets $U, V \subset K$ are Z-invariant is not strictly necessary. Let us briefly outline a proof of the pointwise ergodic theorem for general sets $U, V \subset K \subset G$. Unlike the rest of our discussion throughout this paper, the argument we indicate here is not geometric, but rather spectral in nature.

Clearly, for any $f \in \mathcal{V}_\chi$ defined above, the absolute value $|f(x)|$ is a Z-invariant function. It follows immediately that the averages defined by U and V satisfy the strong maximal inequality in L^p, $p > 1$, and $L \log L$. Indeed, clearly

$$|\pi(\sigma_{r,\epsilon}^{U,V})f| \leq \pi(\sigma_{r,\epsilon}^{U,V})|f| = \pi(\bar{\sigma}_{r,\epsilon}^{U,V})|f|$$

where $\bar{\sigma}_{r,\epsilon}^{U,V}$ is the projection of the measure $\sigma_{r,\epsilon}^{U,V}$ on G to G/Z. Thus the maximal inequalities for the operators $\bar{\sigma}_{r,\epsilon}^{U,V}$ on G/Z, acting on X/Z, imply the desired result.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

It then remains to show that there is a dense subspace in $\mathcal{V}_\chi$ where pointwise almost sure convergence of our averages occurs. This fact can be deduced using the argument appearing in [**N**, §2.5]. There, pointwise convergence for a suitable spectrally defined dense set of K-finite functions is established for $\mathrm{SL}_2(\mathbb{R})$, based on derivative estimates for K-finite functions. Similar derivative estimates can be established for all real rank-one groups with finite center, using, for example, the results in [**Co79**] or [**CM82**]. This establishes pointwise almost sure convergence in a dense subspace of $L^2(X)$, and by a routine application of the maximal inequalities the pointwise ergodic theorem holds as stated.

Acknowledgements. The authors would like to thank the referee for several important and useful comments which lead to significant improvements in the presentation of the arguments in the paper. The first author was supported in part by NSF grant DMS-1500389 and the second author was supported by ISF Moked grant no. 2095/15.

References

[**BK96**] H. Becker and A. S. Kechris. *The Descriptive Set Theory of Polish Group Actions (London Mathematical Society Lecture Note Series, 232)*. Cambridge University Press, Cambridge, 1996.

[**BN13**] L. Bowen and A. Nevo. Geometric covering arguments and ergodic theorems for free groups. *Enseign. Math.* **59** (2013), 133–164.

[**BH99**] M. Bridson and A. Haefliger. *Metric Spaces of Non-Positive Curvature (Grundlehren der mathematischen Wissenschaften, 319)*. Springer, Berlin, 1999.

[**CM82**] W. Casselman and D. Milicic. Asymptotic behavior of matrix coefficients of admissible representations. *Duke Math. J.* **49**(4) (1982), 869–930.

[**Co79**] M. Cowling. Sur les coefficients des representations unitaires des groupes de Lie simples. *Analyse harmonique sur les groupes de Lie (Séminaire, Nancy-Strasbourg, 1976–1978) (Lecture Notes in Mathematics, 739)*. Springer, Berlin, 1979, pp. 132–178.

[**Fav72**] N. A. Fava. Weak type inequalities for product operators. *Studia Math.* **42** (1972), 271–288.

[**GN10**] A. Gorodnik and A. Nevo. *The Ergodic Theory of Lattice Subgroups (Annals of Mathematics Studies, 172)*. Princeton University Press, Princeton, NJ, 2010.

[**He84**] S. Helgason. *Groups and Geometric Analysis*. Elsevier, Burlington, MA, 1984.

[**He94**] S. Helgason. *Geometric Analysis on Symmetric Spaces (Mathematical Surveys and Monographs, 39)*. American Mathematical Society, Providence, RI, 1994.

[**HM79**] R. Howe and C. C. Moore. Asymptotic properties of unitary representations. *J. Funct. Anal.* **32**(1) (1979), 72–96.

[**Ke10**] A. S. Kechris. *Global Aspects of Ergodic Group Actions (Mathematical Surveys and Monographs, 160)*. American Mathematical Society, Providence, RI, 2010.

[**Kn96**] A. W. Knapp. *Lie Groups Beyond An Introduction (Progress in Mathematics, 140)*. Birkhäuser, Boston, 1996.

[**Ko84**] T. H. Koornwinder. Jacobi functions and analysis on noncompact semisimple Lie groups. *Special Functions: Group Theoretical Aspects and Applications, 185 (Mathematics and Its Applications, 18)*. Reidel, Dordrecht, 1984.

[**Ne94**] A. Nevo. Pointwise ergodic theorems for radial averages on simple Lie groups I. *Duke Math. J.* **76** (1994), 113–140.

[**Ne97**] A. Nevo. Pointwise ergodic theorems for radial averages on simple Lie groups II. *Duke Math. J.* **86** (1997), 239–259.

[**Ne98**] A. Nevo. Spectral transfer and pointwise ergodic theorems for semi-simple Kazhdan groups. *Math. Res. Lett.* **5**(3) (1998), 305–325.

[**Ne06**] A. Nevo. Pointwise ergodic theorems for actions of groups. *Handbook of Dynamical Systems*. Vol. 1B Eds. B. Hasselblatt and A. Katok. Elsevier, Amsterdam, 2006, pp. 871–982.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press

[**N**] A. Nevo. Equidistribution in measure-preserving actions of semisimple groups: case of $\mathrm{SL}_2(\mathbb{R})$. *Preprint*, 2017, arXiv:1708.03886.

[**NS97**] A Nevo and E. M. Stein. Analogs of Wiener's ergodic theorems for semi-simple Lie groups I. *Ann. of Math. (2)* **145** (1997), 565–595.

[**Pi42**] H. R. Pitt. Some generalizations of the ergodic theorem. *Math. Proc. Cambridge Philos. Soc.* **38** (1942), 325–343.

[**Wi39**] N. Wiener. The ergodic theorem. *Duke Math. J.* **5** (1939), 1–18.

https://doi.org/10.1017/etds.2017.128 Published online by Cambridge University Press