COUNTING CONNECTIONS IN A LOCALLY SYMMETRIC SPACE

JEREMY KAHN AND ALEX WRIGHT

This is a preliminary draft provided for the purpose of verifying the reference in [KW18]*.*

1. Statement of result and reduction to a special case

1.1. **Preliminaries.** We will denote Lie groups by capital letters such as G, H, and their Lie algebras by $\mathfrak{g}$, $\mathfrak{h}$. We will denote elements of a Lie algebra by capital letters such as X, Y, and elements of a Lie group by g, h, etc. Suppose G is a Lie group with Lie algebra $\mathfrak{g}$. We let $\exp\colon \mathfrak{g} \to G$ be the exponential map; we have $\exp(0) = \mathbf{1}$, and exp is a local diffeomorphism at 0. Therefore we can define a diffeomorphism $\log\colon B \to \log(B) \subset \mathfrak{g}$, where B is a sufficiently small ball around $\mathbf{1}$, such that $\exp \circ \log$ is the identity on B. For $X, Y \in \mathfrak{g}$, then we let $\operatorname{ad}_X Y = [X, Y]$. For $g, h \in G$, we let $C_g h = ghg^{-1}$. We recall that

$$C_{\exp(A)} \exp(B) = \exp(e^{\operatorname{ad}_A} B).$$

1.2. **Haar measures and convolution.** Let Q be any Lie group. Recall that the convolution $\alpha * \beta$ of two measures α, β on Q is defined to be the pushforward of the product measure $\alpha \times \beta$ on $Q \times Q$ via the multiplication map $Q \times Q \to Q$. We observe that convolution is associative. We will always treat convolution as having lower precedence than pointwise multiplication (by a function) so $f\alpha * \beta$ means $(f\alpha) * \beta$ rather than $f \cdot (\alpha * \beta)$ (for a function f and measures α and β).

We will use δ_g to denote the point mass at g. We observe that $\delta_g * \delta_h = \delta_{gh}$. Moreover, for any measure α on Q, we have $\delta_g * \alpha = (L_g)_*\alpha$, where $L_g : Q \to Q$ denotes left multiplication by g. For any function $f\colon Q \to \mathbb{R}$ we let $\delta_g * f$ be a shorthand for $(L_g)_* f$, which of course is defined by $(L_g)_* f(h) = f(g^{-1}h)$. Likewise for $f * \delta_g$.

Now let $\mathfrak{q}$ denote the Lie algebra for Q. For any volume form on $\mathfrak{q}$, we have a unique left Haar measure and right Haar measure on Q.

Date: January 13, 2019.

The authors acknowledge support from U.S. National Science Foundation grants DMS 1107452, 1107263, 1107367 "RNMS: Geometric structures And Representation varieties" (the GEAR Network).

We say that Q is unimodular when the two Haar measures are equal and we recall that this holds, in particular, when Q is semi-simple (or reductive) or nilpotent. We will denote left Haar measure on Q (for some volume form which will be specified when it is important) by η_Q^L, and right Haar measure by η_Q^R. In the case where Q is unimodular we denote the bi-invariant Haar measure by η_Q. In all cases, when $f\colon Q \to \mathbb{R}$ is continuous with compact support, we let $\int_Q f$ be a shorthand for $\int_Q f\, d\eta_Q^L$. We observe that

$$\int \phi = \int \phi\, d\eta_Q^L = (1 + O(\delta)) \int \exp^* \phi$$

and

$$\int \phi\, d\eta_Q^R = (1 + O(\delta)) \int \exp^* \phi$$

when $\operatorname{supp} \phi \subset B_\delta(\mathbf{1})$ and δ sufficiently small.

For $g \in G$, we let

$$\Delta_Q(g) = \frac{|d\eta_Q^L|}{|d\eta_Q^R|}$$

(where we normalize η_Q^L and η_Q^R such that $\Delta_Q(\mathbf{1}) = 1$). Then

$$\eta_Q^L = \Delta_Q(g)\eta_Q^L * \delta_g$$

and

$$\delta_g * \eta_Q^R = \Delta_Q(g)\eta_Q^R.$$

We then have $\Delta_Q(gh) = \Delta_Q(g)\Delta_Q(h)$, and we call Δ_Q the modular homomorphism. We observe that $\Delta_Q(\exp(X)) = 1 + O(X)$ when X is small.

When α is a finite measure on Q and $f\colon Q \to \mathbb{R}$ is continuous with compact support (or more generally all left translates of f are α-integrable) we define $\alpha * f$ by

$$\alpha * f = \int \big(\delta_g * f\big) d\alpha(g) = \int \big((L_g)_* f\big) d\alpha(g)$$

or

$$(\alpha * f)(h) = \int f(g^{-1}h) d\alpha(g);$$

we can also write

$$(\alpha * f)\eta_Q^L = \alpha * (f\eta_Q^L).$$

We can likewise define $f * \beta$ (for a finite measure β) so that $f\eta_Q^R * \beta = (f * \beta)\eta_Q^R$ and observe that $\alpha * (\beta * f) = (\alpha * \beta) * f$ and $(f * \alpha) * \beta = f * (\alpha * \beta)$, and $(\alpha * f) * \beta = \alpha * (f * \beta)$.

Let $f, \phi \colon Q \to [0,\infty)$ be nonnegative continuous functions of compact support. When Q is unimodular, we have $f\eta_Q * \phi = f * \phi\eta_Q$. In the sequel, it will be useful to compare $f\eta^L_Q * \phi$ with $f * \phi\eta^L_Q$ in the case of a general Q. For ϕ a function of compact support, we let $\underline{\Delta}(\phi) = \inf_{g\in\operatorname{supp}\phi} \Delta_Q(g^{-1})$ and $\overline{\Delta}(\phi) = \sup_{g\in\operatorname{supp}\phi} \Delta_Q(g^{-1})$. Then we have

Lemma 1.1. *For $f, \phi \colon Q \to \mathbb{R}$ nonnegative of compact support,*

$$\underline{\Delta}(\phi) f * \phi\eta^L_Q \le f\eta^L_Q * \phi \le \overline{\Delta}(\phi) f * \phi\eta^L_Q. \tag{1}$$

Proof. We first observe that

$$f\eta^L_Q * \delta_g = \Delta_Q(g^{-1})(f * \delta_g)\eta^L_Q.$$

We then have $(f\eta^L_Q * \phi)\eta^L_Q = f\eta^L_Q * \phi\eta^L_Q$, and

$$\begin{aligned} f\eta^L_Q * \phi\eta^L_Q &= \int \Delta_Q(g^{-1})(f * \delta_g)\eta^L_Q \, d(\phi\eta^L_Q) \\ &\le \overline{\Delta}(\phi) \int (f * \delta_g)\eta^L_Q \, d(\phi\eta^L_Q) \\ &= \overline{\Delta}(\phi)(f * \phi\eta^L_Q)\eta^L_Q. \end{aligned}$$

We have thus shown the second inequality of (1) (after multiplying by η^L_Q). The first inequality follows in the same manner. □

In certain cases we can multiply or convolve functions (depending on your point of view) in such a way that the product associates with certain convolutions. In particular, suppose that R and S are Lie subgroups of Q, and $\mathfrak{r} \oplus \mathfrak{s} = \mathfrak{q}$ as vector spaces. Then the multiplication map $R \times S \to Q$ is a diffeomorphism near $(\mathbf{1}, \mathbf{1})$, and a local diffeomorphism on all of $R \times S$; let us suppose that it is injective. Then for $f \colon R \to \mathbb{R}$ and $g \colon S \to \mathbb{R}$ continuous functions of compact support, we can define $f \circledast g \colon Q \to \mathbb{R}$ by $(f \circledast g)(rs) = f(r)g(s)$. Then if α is a compactly supported measure on R and β is a compactly supported measure on S, we have $\alpha*(f\circledast g) = (\alpha*f)\circledast g$ and $(f\circledast g)*\beta = f\circledast(g*\beta)$. Moreover, if $a \in Q$ normalizes R and S, then we have

$$(f \circledast g) * \delta_a = \delta_a * (C^*_a f \circledast C^*_a g).$$

In the case where Q is unimodular, we can define $f\circledast g$ in terms of the convolution of measures. We observe that $(L_r R_s)_*(\eta^L_R*\eta^R_S) = (\eta^L_R*\eta^R_S)$. Since the action of $R \times S$ on $RS = \{rs \mid r \in R, s \in S\}$ is transitive, the measure $\eta^L_R * \eta^R_S$ must be a scalar multiple of η_Q; we assume that $\eta^L_R * \eta^R_S = \eta_Q$. Then we have $(f\eta^L_R) * (g\eta^R_S) = (f \circledast g)\eta_Q$.

On the other hand, given $f \colon R \to \mathbb{R}$ and $g \colon S \to \mathbb{R}$, we let $(f \times g) \colon R \times S \to \mathbb{R}$ be defined by $(f \times g)(r, s) = f(r)g(s)$.

1.3. **An eigenspace factorization of a group.** Let G be a semisimple Lie group of non-compact type, and let A be a nonzero semisimple element of the Lie algebra $\mathfrak{g}$ such that ad_A has all real eigenvalues.

Define $\mathfrak{h}_-$ to be the subspace of $\mathfrak{g}$ spanned by eigenvectors of ad_A with negative eigenvalue. Similarly let $\mathfrak{h}_+$ be spanned by eigenvectors with positive eigenvalue, and $\mathfrak{h}_0 = \ker(\mathrm{ad}_A)$. Thus $\mathfrak{g}$ is the direct sum of $\mathfrak{h}_-$, $\mathfrak{h}_+$, and $\mathfrak{h}_0$. By the Jacobi identity, $\mathfrak{h}_-$, $\mathfrak{h}_+$, and $\mathfrak{h}_0$ are Lie sub-algebras (and $\mathfrak{h}_-$ and $\mathfrak{h}_+$ are nilpotent); let H_-, H_+ and H_0 be the corresponding Lie groups. Moreover, we observe that $\mathfrak{h}_{0+} \equiv \mathfrak{h}_0 \oplus \mathfrak{h}_+$ is a Lie sub-algebra, and that the corresponding Lie subgroup H_{0+} is equal to $\{h_0h_+ \mid h_0 \in H_0, h_+ \in H_+\}$. Likewise for $\mathfrak{h}_{0-}$ and H_{0-}.

We should also assume that H_{0-} is closed...***when can we assume this?***

Lemma 1.2. *The multiplication map $H_- \times H_0 \times H_+ \to G$ is an injective local diffeomorphism with dense image.*

Proof. We first observe that the exponential map $\exp\colon \mathfrak{h}_+ \to H_+$ is surjective. To show this, we consider the adjoint action of $\mathfrak{h}_+$ on $\mathfrak{h}_{0+}$. This representation is faithful, because $\mathrm{ad}_X A \neq 0$ for all $X \in \mathfrak{h}_+$. Moreover, every element of $\mathfrak{h}_+$ acts nilpotently in this representation, so the image of $\mathfrak{h}_+$ is conjugate to a Lie subalgebra of strictly upper-triangular matrices. Then we need only observe that every upper triangular matrix u with 1's on the diagonal can be written uniquely as e^S, where S is strictly upper triangular (with 0's on the diagonal).

We can then show the injectivity as follows. Let $H_{-0+} = H_{0-} \cap H_+$; we will show that $H_{-0+} = \{1\}$. Suppose that $x \in H_{-0+}$. Then $C_{\exp(tA)}x \in H_{-0+}$, and letting $x = \exp(X)$ (where $X \in \mathfrak{h}_+$), we have $C_{\exp(tA)}x = \exp(e^{t\,\mathrm{ad}_A}X)$, and $e^{t\,\mathrm{ad}_A}X \to 0$ as $t \to -\infty$. Let $X' = e^{t\,\mathrm{ad}_A}X$ for t large and negative. Then X' is small, $\exp(X') \in H_{0-}$, and H_{0-} is closed, so $X' \in \mathfrak{h}_{0-}$. Moreover, since $X \in \mathfrak{h}_+$, we have $X' \in \mathfrak{h}_+$. Then we must have $X' = 0$, and $x = 1$.

We haven't shown that the image is dense, but it appears that we never use this statement. □

We denote the image of the multiplication map by $H_-H_0H_+$. Let $K_A = \mathrm{tr}\,\mathrm{ad}_A|_{\mathfrak{g}_+}$.

1.4. **The assumption of exponential mixing.** Continuing the notation of the previous subsection, let Γ be a lattice in G. We assume that there are constants $C \equiv C(\Gamma)$, $k \equiv k(G)$, $q \equiv q(\Gamma)$ such that for all functions $f, g \in C^k(\Gamma\backslash G)$, and $t \in \mathbb{R}$,

(2)
$$\left|\int_{\Gamma\backslash G} 1 \int_{\Gamma\backslash G} (f * \delta_{\exp(tA)})g - \int_{\Gamma\backslash G} f \int_{\Gamma\backslash G} g\right| < Ce^{-q|t|} \|f\|_{C^k} \|g\|_{C^k}.$$

Here all the integrals are taken with respect to η_G.

1.5. **Summing connections over a lattice.** Continuing the notation from the previous two subsections, define

$$Z\colon H_{0-} \times H_+ \to G, \qquad (h_{0-}, h_+) \mapsto h_{0-}h_+^{-1}$$

and

$$Z_t\colon H_{0-} \times H_+ \to G, \qquad (h_{0-}, h_+) \mapsto h_{0-} \exp(tA)h_+^{-1}.$$

We observe that Z maps $\eta^L_{H_{0-}\times H_+}$ to η_G restricted to $H_{0-}H_+$, and Z_t maps $\eta^L_{H_{0-}\times H_+}$ to e^{tK_A} times the same restriction of η_G.

Define, for f a function on $H_{0-} \times H_+$ and $r, s \in G$,

$$\Sigma_t(f, r, s) = \sum_{\gamma\in\Gamma} ((Z_t)_* f)(r^{-1}\gamma s).$$

The meaning of Σ_t can be understood through the following example. Choose $A_- \subset H_-$, $A_0 \subset H_0$ and $A_+ \subset H_+$. Let $f(h_-h_0, h_+) = \chi_{A_-}(h_-)\chi_{A_0}(h_0)\chi_{A_+}(h_+)$. Then $\Sigma_t(f, r, s)$ counts the number of ways to start in rA_-, apply (right-multiply by) $\exp(tA)$, apply something in A_0, and end in γsA_+ for some $\gamma \in \Gamma$.

We can normalize η_G so that Γ has covolume 1, and we can then normalize $\eta^L_{H_0-\times H_+}$ accordingly. If we were to replace Γ with randomly chosen points in G with density 1, then the expected value of $\Sigma_t(f, r, s)$ would be

$$\int_G (Z_t)_* f = e^{tK_A} \int_{H_{0-}\times H_+} f.$$

We claim that this is approximately correct for an actual lattice Γ, a large t, and a reasonable f.

For any $f\colon G \to \mathbb{R}$ and $\delta > 0$, let $M_\delta(f)(p) = \sup_{B_\delta(p)} f$, and $m_\delta(f)(p) = \inf_{B_\delta(p)} f$. For $h \in G$, let $\epsilon_h = \min\left(\frac{1}{2} \inf_{\gamma\in\Gamma\backslash\{1\}} d(h, \gamma h), 1\right)$. The following is the main result of this paper.

Theorem 1.3. *We can find $a \equiv a(G, A)$ such that for all lattices $\Gamma < G$, $t > 0$, and $g, h \in G$ with $\epsilon_g, \epsilon_h > \delta$ (where $\delta = C(\Gamma)e^{-aqt}$), and $f\colon H_{0-} \times H_+ \to \mathbb{R}$ measurable, bounded, and compactly supported, we have*

$$(1 - \delta) \int_{H_{0-}\times H_+} m_\delta(f) \le e^{-tK_A}\Sigma_t(f, g, h) \le (1 + \delta) \int_{H_{0-}\times H_+} M_\delta(f).$$

(In the case where Γ is a uniform lattice, we can ignore the requirements on ϵ_g and ϵ_g, which will hold automatically).

Corollary 1.4. *With a, g, h, t, δ as above. Suppose $S \subset H_{0-} \times H_+$ is measurable and bounded. Then*

$$(1-\delta)\mathcal{N}_{-\delta}(S) < e^{-tK_A}\#(Z_t(S) \cap g\Gamma h) < (1+\delta)\mathcal{N}_\delta(S).$$

The following Proposition will be proven in Section 2; we will use it now to prove Theorem 1.3.

Proposition 1.5. *Let δ and Γ be as in Theorem 1.3. For all $t > 0$ there is $\psi^t \colon H_{0-} \times H_+ \to [0, \infty)$ with $\int \psi^t = 1$ and with support in a δ-neighborhood of the identity such that for all $g, h \in G$ with $\epsilon_g, \epsilon_h > \delta^{1/d}$,*

$$\left| e^{-tK_A}\Sigma_t(\psi^t, g, h) - \int \psi^t \right| \leq \delta.$$

The following Lemma will be used to prove Theorem 1.3 using Lemma 1.5.

Lemma 1.6. *For any measure α on $H_{0-} \times H_+$,*

$$\Sigma_t(\alpha * \psi, r, s) = \int \Sigma_t(\psi, rh_{0-}, sh_+)\alpha(h_{0-}, h_+). \tag{3}$$

Proof. It is enough to show (3) in the case where α is a point mass $\delta_{(h_{0-}, h_+)}$, and in this case the identity is straightforward to verify. □

As a corollary to this Lemma, we observe, letting $|\alpha|$ denote the total mass of α, and assuming $\operatorname{supp}\psi \in B_\delta(1)$,

$$|\alpha| \inf_{\substack{g \in B_\delta(r) \\ h \in B_\delta(s)}} \Sigma_t(\psi, g, h) \leq \Sigma_t(\alpha * \psi, r, s) \leq |\alpha| \sup_{\substack{g \in B_\delta(r) \\ h \in B_\delta(s)}} \Sigma_t(\psi, g, h).$$

We then observe that

$$\begin{aligned} f &\leq M_\delta f * \psi \eta^L_{H_{0-} \times H_+} \\ &\leq \overline{\Delta}(\psi)(M_\delta f)\eta^L_{H_{0-} \times H_+} * \psi \quad \text{(by Lemma 1.1)} \end{aligned}$$

and hence, by Lemma 1.6,

$$\Sigma_t(f, r, s) \leq \overline{\Delta}(\psi) \left(\int M_\delta(f)\eta^L_{H_{0-} \times H_+} \right) \sup_{\substack{g \in B_\delta(r) \\ h \in B_\delta(s)}} \Sigma_t(\psi, g, h) \tag{4}$$

and likewise

$$\Sigma_t(f, r, s) \geq \underline{\Delta}(\psi) \left(\int m_\delta(f)\eta^L_{H_{0-} \times H_+} \right) \inf_{\substack{g \in B_\delta(r) \\ h \in B_\delta(s)}} \Sigma_t(\psi, g, h). \tag{5}$$

Now we can prove Theorem 1.3.

Proof of Theorem 1.3 given Proposition 1.5. We observe that

$$\begin{aligned} e^{-tK_A}\Sigma_t(f,g,h) &\le e^{-tK_A}\overline{\Delta}(\psi^t)\left(\int M_\delta(f)\right)\sup_{\substack{g\in B_\delta(r)\\ h\in B_\delta(s)}}\Sigma_t(\psi^t,g,h) \\ &\le (1+O(\delta))\left(\int M_\delta(f)\right)\left(\int \psi^t+\delta\right) \\ &= (1+O(\delta))\left(\int M_\delta(f)\right), \end{aligned}$$

and we likewise use $m_\delta(f)$ to get the lower bound for $e^{-tK_A}\Sigma_t(f,g,h)$. $\square$

1.6. **Injectivity radius.** We have fixed a semi-simple Lie group G, a lattice $\Gamma\subset G$, and a left-invariant metric (determined by a left-invariant Riemannian metric) $d(\cdot,\cdot)$ on G. Recall $\epsilon_g=\min(\frac{1}{2}\inf_{\gamma\in\Gamma}d(g,\gamma g),1)$. We say that $f\colon G\to\mathbb{R}$ is coarsely Lipshitz if there is some K such that $|f(g)-f(g')|<K$ when $d(g,g')<1$. We then have

Lemma 1.7. *The function $g\mapsto\log(\epsilon_g)$ is coarsely Lipschitz on all of G.*

Proof. It's enough to show that there exist ϵ,K such that for all $g,\gamma\in G$,

$$d(gh,\gamma gh)<Kd(g,\gamma g) \tag{6}$$

when $d(h,\mathbf{1})<\epsilon$. Equation (6) is equivalent to

$$d(h^{-1}g^{-1}\gamma gh,\mathbf{1})<Kd(g^{-1}\gamma g,\mathbf{1}),$$

so letting $u=g^{-1}\gamma g$, we must show that

$$d(C_{h^{-1}}u,\mathbf{1})<Kd(u,\mathbf{1}) \tag{7}$$

(when h is close to $\mathbf{1}$). We have

$$d(C_{h^{-1}}u,\mathbf{1})<d(u,\mathbf{1})+2d(h,\mathbf{1}),$$

so it is clear that (7) holds except possibly when u is close to $\mathbf{1}$. So we can write $u=\exp(S)$, $h=\exp(H)$, and then we must show that

$$d(\exp(e^{-\operatorname{ad}_H}S),\mathbf{1})<Kd(\exp(S),\mathbf{1}),$$

which is tantamount to showing that $e^{-\operatorname{ad}_H}$ is bounded in norm when H is small. $\square$

2. The counting estimate for the test functions

2.1. **An *a priori* counting estimate.**

We begin in our setting of a Lie group G with a chosen $A \in \mathfrak{g}$ that in turn defines $H_-, H_0, H_+ < G$, and a lattice $\Gamma < G$. We will begin with the following volume estimate:

Lemma 2.1. *When B is a sufficiently small ball around $\mathbf{1}$, we have*

$$\eta_G(B \exp(tA) B) \leq C e^{tK_A}.$$

Proof. We recall that in our case that G and H_+ are unimodular. We let B_{0-}, B_+ be the unit balls around the identity in H_{0-} and H_+. We observe that

$$B \exp(tA) B \subset B_{0-} \exp(tA) B_+,$$

and

$$\eta_{H_+}(\exp(tA) B_+ \exp(-tA)) = e^{tK_A} \eta_{H_+}(B_+).$$

Then we have

$$\begin{aligned} \eta_G(B \exp(tA) B) &\leq \eta_G(B_{0-} \exp(tA) B_+) \\ &= \eta_G(B_{0-} \exp(tA) B_+ \exp(-tA)) \\ &= \eta^L_{H_{0-}}(B_{0-}) \eta_{H_+}(\exp(tA) B_+ \exp(-tA)) \\ &= e^{tK_A} \eta^L_{H_{0-}}(B_{0-}) \eta_{H_+}(B_+) \\ &= C e^{tK_A}. \end{aligned}$$

□

Let ϵ_G be half the radius of the ball B in Lemma 2.1. For $h \in G$, let $\epsilon_h = \min\left(\frac{1}{2} \inf_{\gamma \in \Gamma \setminus \{1\}} d(h, \gamma h), \epsilon_G\right)$, and let B_h be the ball of radius ϵ_h (around the identity), and let $v_h = \eta_G(B_h)$. We observe that $v_h \asymp \epsilon_h^d$. From the volume estimate of Lemma 2.1 we can prove the following counting estimate:

Lemma 2.2. *Take $B \equiv B_{\epsilon_G}(\mathbf{1})$. For all $g, h \in G$, we have*

$$\#(g\Gamma h \cap B \exp(tA) B) \leq C(\Gamma) e^{K_A t} / v_h.$$

Proof. We have that

$$\#(g\Gamma h \cap B \exp(tA) B) < \eta_G(N_{\epsilon_h}(B \exp(tA) B)) / v_h.$$

We observe that

$$N_{\epsilon_h}(B \exp(tA) B) \subset B \exp(tA) \hat{B}$$

where $\hat{B} \equiv B_{2\epsilon_G}(\mathbf{1})$. Moreover, by Lemma 2.1,

$$\eta_G(\hat{B} \exp(tA) \hat{B}) \leq C e^{tK_A}. \quad \square$$

2.2. **Estimates with linearly complementary subgroups.** In this subsection, we consider a more general situation where G is an arbitrary Lie group, A and B are Lie subgroups of G with Lie algebras $\mathfrak{a}$ and $\mathfrak{b}$, where $A \cap B = \{\mathbf{1}\}$ and $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ as vector spaces.

We assume that $\mathfrak{a}$ and $\mathfrak{b}$ are equipped with inner products; this determines an inner product on $\mathfrak{g}$, and left invariant metrics and left Haar measures on A, B and G.

Lemma 2.3. *Suppose $a_0, a_1 \in A$, $b_0, b_1 \in B$ are all sufficiently close to the identity and that $a_0 b_0 = b_1 a_1$. Let $D = \max(|\log b_0|, |\log a_1|)$. Then*

$$|\log a_0| \leq 2D \text{ and } |\log b_1| \leq 2D$$

Proof. We have

$$\log a_0 + \log b_0 + O(|\log a_0||\log b_0|) = \log b_1 + \log a_1 + O(|\log b_1||\log a_1|)$$

and hence

$$\log a_0 + \log b_0 + O(|\log a_0|D) = \log b_1 + \log a_1 + O(|\log b_1|D)$$

and therefore, because $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$,

$$\log a_1 = \log a_0 + O(ED) \tag{8}$$
$$\log b_1 = \log b_0 + O(ED) \tag{9}$$

where $E = |\log a_0| + |\log b_1|$. The Lemma follows because E is assumed to be small.

$\square$

Lemma 2.4. *Suppose that $a_0, a_1 \in A$, $b_0, b_1 \in B$, and a_0 and b_1 are close to the identity and $a_0 b_0 = b_1 a_1$. Then b_0 and a_1 are also close to the identity.*

Proof. We can write

$$b_0 = a_0^{-1} b_1 a_1 = b_1' a_0' a_1$$

for some $b_1' \in B, a_0' \in A$ close to the identity. But then $a_0' a_1 = b_1'^{-1} b_0 \in A \cap B = \{\mathbf{1}\}$. $\square$

Lemma 2.5. *Suppose we have $\hat{a}, \check{a} \in A$, and $\hat{b}, \check{b} \in B$, with $\check{a}, \check{b}$ sufficiently close to the identity. Suppose further we have*

$$\hat{a}\check{b} = \nu \hat{b} \check{a}$$

for some $\nu \in G$. Then we can write $\nu = \nu_a \nu_b$, with $\nu_a \in A$, $\nu_b \in B$.

Proof. We can find $a \in A, b \in B$ (close to the identity) such that $ab = \hat{b}\hat{a}^{-1}$. Then

$$\nu = \hat{a}\check{b}\check{a}^{-1}\hat{b}^{-1} = (\hat{a}a)(b\hat{b}^{-1}).$$

$\square$

Lemma 2.6. *Let $\hat{\psi}_A, \check{\psi}_A$ be functions on A, and $\hat{\psi}_B, \check{\psi}_B$ be functions on B, and let D be sufficiently small. Assume*

(1) $\operatorname{supp}\check{\psi}_A, \operatorname{supp}\check{\psi}_B$ *are supported in the D neighbourhood of the identity, and*
(2) $\check{\psi}_A$ *and* $\check{\psi}_B$ *are nonnegative on their domains, and*
(3) $\int \check{\psi}_A = \int \check{\psi}_B = 1$.

Let $E_A = \|\hat{\psi}_A\|_{C^1}$ (computed on the ball of radius $2D$ around the identity), and define E_B analogously. Then

$$\left| \int_G \left(\hat{\psi}_A \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) - \hat{\psi}_A(1)\hat{\psi}_B(1) \right| \leq C_{A,B} D E_A E_B.$$

Proof. By Lemmas 2.3 and 2.4, the integrand is supported on the product (in either order) of the balls of radius $2D$ (around $\mathbf{1}$) in A and B. Hence

$$\begin{aligned} &\left| \int_G \left(\hat{\psi}_A \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) - \int_G \left((\hat{\psi}_A(1)1_A) \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) \right| \\ \leq\ & \int_G \left((2DE_A 1_A) \circledast \check{\psi}_B\right) \cdot \left(\left|\hat{\psi}_B\right| \circledast \check{\psi}_A\right) \\ \leq\ & \int_G \left((2DE_A 1_A) \circledast \check{\psi}_B\right) \cdot \left(E_B 1_B \circledast \check{\psi}_A\right) \\ \leq\ & 2DE_A E_B \,\mathcal{S}, \end{aligned}$$

where $\mathcal{S} = \int_G (1_A \circledast \check{\psi}_B)(1_B \circledast \check{\psi}_A)$. Similarly

$$\begin{aligned} &\left| \int_G \left((\hat{\psi}_A(1)1_A) \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) - \int_G \left((\hat{\psi}_A(1)1_A) \circledast \check{\psi}_B\right) \cdot \left((\hat{\psi}_B(1)1_B) \circledast \check{\psi}_A\right) \right| \\ &\quad \leq 2DE_A E_B \,\mathcal{S}. \end{aligned}$$

Hence by the triangle inequality we get

$$\begin{aligned} &\left| \int_G \left(\hat{\psi}_A \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) - \hat{\psi}_A(1)\hat{\psi}_B(1)\,\mathcal{S} \right| \\ \leq\ & 2D(E_A E_B + E_B E_A)\,\mathcal{S}. \end{aligned}$$

It remains to estimate $\mathcal{S}$. Let $\mathbf{B}$ be the ball of radius $2D$ around the identity in $A \times B$. We define the map $\mathbf{B} \to G$ as follows.Given $(a, b) \in \mathbf{B}$, we solve $ab' = ba'$ for $a' \in A, b' \in B$ (by solving $b'a'^{-1} = a^{-1}b$), and then let $\rho(a, b) = ab'$.

Then

$$\mathcal{S} = \int_{A\times B} \check{\psi}_A \times \check{\psi}_B \, d\rho^*(\eta_G^L).$$

Moreover,

$$\operatorname{Jac}\rho \equiv \frac{|d\rho^*(\eta_G^L)|}{|d(\eta_A^L \times \eta_B^L)|}$$

satisfies $\operatorname{Jac}\rho(a,b) = 1 + O(|\log a| + |\log b|)$. Therefore

$$\begin{aligned}\int_{A\times B} \check{\psi}_A \times \check{\psi}_B \, d\rho^* \eta_G^L &= \int_{A\times B} \check{\psi}_A \times \check{\psi}_B (1 + O(D)) \, d(\eta_A^L \times \eta_B^L) \\ &= 1 + O(D).\end{aligned}$$

(In fact we can get $1+O(D^2)$, but we will not need this.) We conclude that

$$\left| \int_G \left(\hat{\psi}_A \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) - \hat{\psi}_A(1)\hat{\psi}_B(1)\mathcal{S} \right| < C_{A,B} D E_A E_B$$

when D is sufficiently small. □

Corollary 2.7. *Suppose that the conditions of Lemma 2.6 hold, except for assumption 3: the normalization of $\hat{\psi}_A$ and $\hat{\psi}_B$. Let $I_A = \int_G \hat{\psi}_A$, and $I_B = \int_G \hat{\psi}_B$. Then*

$$\left| \int_G \left(\hat{\psi}_A \circledast \check{\psi}_B\right) \cdot \left(\hat{\psi}_B \circledast \check{\psi}_A\right) - I_A I_B \hat{\psi}_A(1)\hat{\psi}_B(1) \right| \le C_{A,B} I_A I_B D E_A E_B.$$

Moveover, letting $I'_A = \int \exp^* \hat{\psi}_A$ and $I'_B = \int \exp^* \hat{\psi}_B$, the exact same statement holds with I_A and I_B replaced with I'_A and I'_B.

Proof. The Corollary is clear for I_A and I_B; let us prove it for I'_A and I'_B. We have $I'_A = (1+O(D))I_A$ and $I'_B = (1+O(D))I_B$ and therefore

$$\begin{aligned}\left| I_A I_B \hat{\psi}_A(\mathbf{1})\hat{\psi}_B(\mathbf{1}) - I'_A I'_B \hat{\psi}_A(\mathbf{1})\hat{\psi}_B(\mathbf{1}) \right| &\le C I'_A I'_B D \psi_A(\mathbf{1})\hat{\psi}_B(\mathbf{1}) \\ &\le C I'_A I'_B D E_A E_B,\end{aligned}$$

which is exactly what we require. □

2.3. **Defining the bump functions.** Let us fix a smooth function $g\colon [0,\infty) \to [0,\infty)$ such that all the derivatives of g at 0 are zero, $\|g\|_\infty = 1$, and $\operatorname{supp} g \subset [0,1)$. Let us then define Ξ_d on $\mathbb{R}^d$, for $d \in \mathbb{Z}^+$, by $\Xi_d(x) = C_d g(|x|)$, where C_d is such that $\int \Xi_d = 1$. For $t \ge 0$, let us then define Ξ_d^t by

$$\Xi_d^t(x) = e^{dt}\Xi_d(e^t x).$$

So Ξ_d^t has integral 1, is supported in the ball of radius e^{-t} around 0, has sup norm at most $C_d e^{dt}$, and $\|\Xi_d^t\|_{C^k} \le C_d e^{(d+k)t}$. Because

Ξ_d^t is rotationally symmetric, it is well-defined on any vector space of dimension d that has an inner product.

Let H be a Lie group equipped with a left-invariant metric, and let $\mathfrak{h}$ be its Lie algebra. We can define $\Xi_{\mathfrak{h}}^t$ on $\mathfrak{h}$ to be Ξ_d^t, and we then let ξ_H^t on H be defined by

$$\xi_H^t(\exp(X)) = \Xi_{\mathfrak{h}}^t(X); \tag{10}$$

this will certainly make sense when t is sufficiently large.

Returning now to the setting of Section 1, we let $m = \max(16(d + \max(k,1)), \lambda_1^{-1})$, where d is the dimension of G, k is as in equation (2), and λ_1 is the least positive eigenvalue for ad_A [or the negative of the least negative one?]. We then let $b = 1/m$ and $a = 1/m^2$. Letting q be the rate of mixing, we write

$$\Psi_+^t = \Xi_{\mathfrak{h}_+}^{aqt} \qquad \Psi_0^t = \Xi_{\mathfrak{h}_0}^{aqt}$$
$$\Psi_-^t = \Xi_{\mathfrak{h}_-}^{aqt} \qquad \tilde{\Psi}_0^t = \Xi_{\mathfrak{h}_0}^{4bqt}$$

and we let $\Psi_{0-}^t = \Psi_0^t \times \Psi_-^t$, and $\tilde{\Psi}_{0-}^t = \tilde{\Psi}_0^t \times \Psi_-^t$.

We then define ψ_+^t and its relatives by the direct analogue of Equation (10).

We further define

$$\check{\psi}_+^t = C^*_{\exp(tA/2)}\psi_+^t \qquad \check{\psi}_{0-}^t = C^*_{\exp(-tA/2)}\tilde{\psi}_{0-}^t$$
$$\hat{\psi}_+^t = C^*_{\exp(-tA/2)}\psi_+^t \qquad \hat{\psi}_{0-}^t = C^*_{\exp(tA/2)}\psi_{0-}^t.$$

Similarly we have $\check{\Psi}_+ = C^*_{\exp(t\,\mathrm{ad}_A/2)}\Psi_+$ etc. We let $\psi^t = \psi_{0-}^t \circledast \psi_+^t$.

We apply Corollary 2.7 to the setting of the ψ's.

Lemma 2.8. *With a, b taken as above, and C depending only on H_0, etc., we have*

$$\left| e^{K_A t}\int_G \left(\delta_{\mu_{0-}} * \hat{\psi}_{0-}^t \circledast \check{\psi}_+^t\right)\cdot\left(\delta_{\mu_+} * \hat{\psi}_+^t \circledast \check{\psi}_{0-}^t\right) - \hat{\psi}_{0-}^t(\mu_{0-}^{-1})\hat{\psi}_+^t(\mu_+^{-1})\right| < Ce^{-2bqt}.$$

Proof. We have $(\delta_{\mu_{0-}} * \hat{\psi}_{0-}^t)(1) = \hat{\psi}_0^t(\mu_{0-}^{-1})$ and

$$\|\delta_{\mu_{0-}} * \hat{\psi}_{0-}^t\|_{C^1} = \|\hat{\psi}_{0-}^t\|_{C^1} \leq \|\psi_{0-}^t\|_{C^1} \leq Ce^{(d+1)aqt} \leq Ce^{bqt}.$$

Likewise we have $\delta_{\mu_+} * \hat{\psi}_+^t = \hat{\psi}_-^t(\mu_+^{-1})$ and

$$\|\delta_{\mu_+} * \hat{\psi}_+^t\|_{C_1} = \|\hat{\psi}_+^t\|_{C_1} \leq \|\psi_+^t\|_{C_1} \leq Ce^{(d+1)aqt} \leq Ce^{bqt}.$$

Moreover, the radius (around the identity) of the support of ψ_+^t is at most $e^{-aqt} << 1$, and radius of support of $\check{\psi}_+^t$ is therefore at most $e^{-\lambda_1 t} \leq e^{-4bqt}$. The radius of support of $\check{\psi}_{0-}^t$ is at most e^{-4bqt}. Putting this all together and applying Corollary 2.7, we obtain the Lemma. $\square$

2.4. **Proving what must be proved.** We can now prove the following proposition, which immediately implies Proposition 1.5.

Proposition 2.9. *There exists C (depending only on Γ) such that for all $g, h \in G$ such that $\epsilon_g, \epsilon_h > e^{-aqt/d}$, we have*

$$\left|e^{-tK_A}\Sigma_t(\psi^t, g, h) - 1\right| \leq Ce^{-aqt}.$$

Proof. The idea is to relate the sum in $\Sigma_t(\psi^t, g, h)$ to a mixing integral. We consider the functions $\delta_g * \psi^t_{0-} \circledast \psi^t_+$ and $\delta_h * \psi^t_+ \circledast \tilde{\psi}^t_{0-}$ on G; they are supported in balls around g and h respectively, with radii $O(e^{-aqt})$ and $O(e^{-bqt})$. Our condition on ϵ_g and ϵ_h implies that the supports of these functions inject into $\Gamma\backslash G$, and hence we can think of them as functions on $\Gamma\backslash G$.

We then have, on the one hand, by exponential mixing in G,

(11)
$$\begin{aligned}&\left|\int_{\Gamma\backslash G} (\delta_g * \psi^t_{0-} \circledast \psi^t_+) \cdot (\delta_h * \psi^t_+ \circledast \tilde{\psi}^t_{0-} * \delta_{\exp(-tA)}) - \int_{\Gamma\backslash G} \psi^t_{0-} \circledast \psi^t_+ \int_{\Gamma\backslash G} \psi^t_+ \circledast \tilde{\psi}^t_{0-}\right| \\ &\quad < Ce^{-qt} \, \|\psi^t_{0-} \circledast \psi^t_+\|_{C^k} \, \|\psi^t_+ \circledast \tilde{\psi}^t_{0-}\|_{C^k} \\ &\quad < Ce^{-qt} e^{(d+k)aqt} e^{(d+k)bqt} < Ce^{-qt/2}.\end{aligned}$$

Moreover,

$$\int_{\Gamma\backslash G} \psi^t_{0-} \circledast \psi^t_+ = \int_G \psi^t_{0-} \circledast \psi^t_+ = (1 + O(e^{-bqt})) \int_{\mathfrak{g}} \exp^*(\psi^t_{0-} \circledast \psi^t_+) = 1 + O(e^{-bqt})$$

and likewise $\int_{\Gamma\backslash G} \psi^t_+ \circledast \tilde{\psi}^t_{0-} = 1 + O(e^{-aqt})$, so

$$\left|\int_{\Gamma\backslash G} \psi^t_{0-} \circledast \psi^t_+ \int_{\Gamma\backslash G} \psi^t_+ \circledast \tilde{\psi}^t_{0-} - 1\right| < Ce^{-aqt}. \tag{12}$$

On the other hand the first integral above is equal to

$$\sum_{\gamma \in \Gamma} \int_G (\delta_g * \psi^t_{0-} \circledast \psi^t_+) \cdot (\delta_\gamma * \delta_h * \psi^t_+ \circledast \tilde{\psi}^t_{0-} * \delta_{\exp(-tA)}).$$

We can rewrite each term in the sum as

$$\int_G (\psi^t_{0-} \circledast \psi^t_+) \cdot (\delta_{g^{-1}\gamma h} * \psi^t_+ \circledast \tilde{\psi}^t_{0-} * \delta_{\exp(-tA)})) \tag{13}$$

or

$$\int_G (\psi^t_{0-} \circledast \psi^t_+ * \delta_{\exp(tA/2)}) \cdot (\delta_{g^{-1}\gamma h} * \psi^t_+ \circledast \tilde{\psi}^t_{0-} * \delta_{\exp(-tA/2)}).$$

We then have, letting $\eta = g^{-1}\gamma h$ and $\nu = \exp(-tA/2)\eta\exp(-tA/2)$,

$$\begin{aligned}&\int_G \left(\psi_{0-}^t \circledast \psi_+^t * \delta_{\exp(tA/2)}\right)\cdot\left(\delta_{g^{-1}\gamma h} * \psi_+^t \circledast \tilde{\psi}_{0-}^t * \delta_{-\exp(tA/2)}\right)\\ &= \int_G \left(\delta_{\exp(tA/2)} * \hat{\psi}_{0-}^t \circledast \check{\psi}_+^t\right)\cdot\left(\delta_\eta * \delta_{\exp(-tA/2)} * \hat{\psi}_+^t \circledast \check{\psi}_{0-}^t\right)\\ &= \int_G \left(\hat{\psi}_{0-}^t \circledast \check{\psi}_+^t\right)\cdot\left(\delta_\nu * \hat{\psi}_+^t \circledast \check{\psi}_{0-}^t\right).\end{aligned}$$

It follows from Lemma 2.5 that if the above integrand is ever nonzero, we can write $\nu = \nu_{0-}\nu_+$ for $\nu_{0-} \in H_{0-}$, $\nu_+ \in H_+$. Then the above integral equals

$$\int_G \left(\delta_{\nu_{0-}^{-1}} * \hat{\psi}_{0-}^t \circledast \check{\psi}_+^t\right)\cdot\left(\delta_{\nu_+} * \hat{\psi}_+^t \circledast \check{\psi}_{0-}^t\right). \tag{14}$$

By Lemma 2.8,

$$e^{K_A t}\int_G \left(\delta_{\nu_{0-}^{-1}} * \hat{\psi}_{0-}^t \circledast \check{\psi}_+^t\right)\cdot\left(\delta_{\nu_+} * \hat{\psi}_+^t \circledast \check{\psi}_{0-}^t\right) \tag{15}$$

is approximately equal to

$$\hat{\psi}_{0-}^t(\nu_{0-})\hat{\psi}_+^t(\nu_+^{-1}) \tag{16}$$

which equals

$$\psi_{0-}^t(C_{\exp(tA/2)}\nu_{0-})\psi_+^t(C_{\exp(-tA/2)}\nu_+^{-1})$$

which in turn equals

$$(Z_t)_*(\psi_{0-}^t \times \psi_+^t)(\eta) = (Z_t)_*(\psi_{0-}^t \times \psi_+^t)(g^{-1}\gamma h). \tag{17}$$

In fact, by Lemma 2.8, (15) and (16) differ by at most Ce^{-2bqt}.

If (17) is nonzero (for a given $\gamma \in \Gamma$), then the integrand in (15) is not identically zero, and likewise for the integrand of (13). By Lemma 2.2, because $\psi_{0-}^t \circledast \psi_+^t$ and $\psi_+^t \circledast \tilde{\psi}_{0-}^t$ are both supported on the unit ball around the identity, the number of γ for which the integrand of (13) is nonzero is at most $Ce^{K_A t}/v_h$.

Therefore the sum of integrals (13) is approximately

$$e^{-K_A t}\Sigma_t(\psi_{0-}^t \times \psi_+^t, g, h),$$

and the difference is at most $Ce^{-2bqt}/v_h \leq Ce^{-bqt}$. □

3. Applications

3.1. **Haar measure as a volume form.** As before, we let η_G^L denote the left Haar measure on G. We let $d\eta_G^L$ denote the associated volume form, so that

$$\int f d\eta_G^L$$

can be interpreted as the integral of f with respect to the Haar measure, or with respect to the volume form, with identical results. Then $d\eta_G^L(\mathbf{1})$ is a top-dimensional multilinear form on $T_{\mathbf{1}}G$; it determines the normalization of η_G^L and $d\eta_G^L$.

3.2. **The Heteromodular homomorphism.** We recall that $[\mathfrak{h}_0, \mathfrak{h}_+] = \mathfrak{h}_+$, and therefore $[H_0, H_+] = H_+$. For any $h_0 \in H_0$, we have $(C_{h_0})_* \eta_{H_+} = \chi(h_0)\eta_{H_+}$. We call χ the heteromodular homomorphism. We **claim** that $(C_{h_0})_*\eta_{H_-} = \chi(h_0)^{-1}\eta_{H_-}$ for any $h_0 \in H_0$. Moreover, $\chi\colon H_0 \to \mathbb{R}^+$ is a homomorphism; we let H_{00} be its kernel. Then $H_0 = \exp(tA) \times H_{00}$, because $\exp(tA)$ commutes with H_{00}.

Moreover, the pullback of $\eta_{H_{0-}}$ to $H_- \times H_0$ by the multiplication map is $\chi(h_0)(\eta_{H_-} \times \eta_{H_0})$. Likewise the pullback of η_G to $H_- \times H_0 \times H_+$ is $\chi(h_0)(\eta_{H_-} \times \eta_{H_0} \times \eta_{H_+})$.

3.3. **Pullbacks of Haar Measure.** Suppose E_- and E_+ are Lie subgroups of G such that

$$\pi_{\mathfrak{h}_\pm}\colon \mathfrak{e}_\pm \to \mathfrak{h}_\pm$$

is an isomorphism. We define volume forms $d\eta_{E_\pm}$ on $\mathfrak{e}_\pm$ by

$$d\eta_{E_\pm} = (\pi_{\mathfrak{h}_\pm}|_{\mathfrak{e}_\pm})^* d\eta_{H_\pm}.$$

We also let $E_0 = H_0$, and keep its volume form. Now we also have maps

$$\Sigma_H\colon \bigoplus \mathfrak{h}_i \to \mathfrak{g}$$

and

$$\Sigma_E\colon \bigoplus \mathfrak{e}_i \to \mathfrak{g},$$

just given by

$$\Sigma_H(h_-, h_0, h_+) = h_- + h_0 + h_+,$$

and likewise for E. Moreover, Σ_H is invertible, and $\Sigma_H^*\eta_G = \bigwedge_i \eta_{H_i}$ on $\bigoplus \mathfrak{h}_i$. We want to compare $\Sigma_E^*\eta_G$ and $\bigwedge_i \eta_{E_i}$.

To this end, we let $\tau_i\colon \mathfrak{h}_i \to \mathfrak{e}_i$ be $(\pi_{\mathfrak{h}_i}|\mathfrak{e}_i)^{-1}$; $T_i\colon \mathfrak{h}_i \to \bigoplus \mathfrak{h}_i$ be $\Sigma_H^{-1} \circ \Sigma_E \circ \tau_i$, and $T\colon \mathfrak{h}_i \to \mathfrak{h}_i$ be $\bigoplus T_i$. Then

$$\frac{\Sigma_E^*\eta_G}{\bigwedge_i \eta_{E_i}} = \frac{T^* \bigwedge_i \eta_{H_i}}{\bigwedge_i \eta_{H_i}} = \det T \tag{18}$$

Letting $T_j^i = \pi_{\mathfrak{h}_i} \circ T_j$, we have that T_i^i is the identity for each i, and thus

$$T = \begin{pmatrix} 1 & 0 & T_+^- \\ T_-^0 & 1 & T_+^0 \\ T_-^+ & 0 & 1 \end{pmatrix}$$

and hence

$$\det T = \det \begin{pmatrix} 1 & T_+^- \\ T_-^+ & 1 \end{pmatrix} = \det(\mathbf{1} - T_-^+ T_+^-). \tag{19}$$

We let $m\colon E_- \times E_0 \times E_+ \to G$ be the multiplication map (so $m(a_-, a_0, a_+) = a_- a_0 a_+$).

Lemma 3.1. *We have*

$$m^* d\eta_G(a_-, a_0, a_+) = q(a_0) d\eta_{E_-}^L \wedge d\eta_{H_0} \wedge d\eta_{E_+}^R$$

where

$$q(a_0) = q(a_0; E_-, E_+) = \chi(a_0) \det(\mathbf{1}_{\mathfrak{h}_+} - T_+^- \circ \mathrm{Ad}_{a_0}^{-1}|_{\mathfrak{h}_+} \circ T_-^+ \circ \mathrm{Ad}_{a_0}|_{\mathfrak{h}_-}).$$

Proof. We first observe that $m^* d\eta_G$ must have the form given in the first line (for some q), because it is invariant under left multiplication in E_- and right multiplication in E_+. Then we observe that, for $u \in H_0$,

$$L_u \circ m = m \circ \big((a_-, a_0, a_+) \mapsto (C_u a_-, u a_0, a_+)\big)$$

(where on the left hand side m is $m\colon E_- \times H_0 \times E_+ \to G$, and the right hand side m is $m\colon C_u E_- \times H_0 \times E_+ \to G$). Since η_G is invariant under pullback by L_u, we obtain

$$q(h_0; E_-, E_+) = \frac{1}{\chi(u)} q(u h_0; C_u E_-, E_+),$$

and letting $u = h_0^{-1}$,

$$q(h_0; E_-, E_+) = \chi(h_0) q(\mathbf{1}; C_{h_0^{-1}} E_-, E_+). \tag{20}$$

When we replace $\mathfrak{e}_-$ with $\mathrm{Ad}_u \mathfrak{e}_-$, we replace T_-^+ with $\mathrm{Ad}_u \circ T_-^+ \circ \mathrm{Ad}_u^{-1}$. The Lemma then follows from (18), (19), and (20). □

3.4. **A more general setting.** Suppose now that that E_- and E_+ are subgroups such that

$$\ker \pi_{\mathfrak{h}_\pm}|_{\mathfrak{e}_\pm} \subset \mathfrak{h}_0. \tag{21}$$

We let $E_{0\pm} = E_\pm \cap H_0$, and we let E be the quotient of $E_- \times E_0 \times E_+$ by $(e_- e_{0-}, e_0, e_{0+} e_+) \sim (e_-, e_{0-}^{-1} e_0 e_{0+}^{-1}, e_+)$.

We let $\hat{\mathfrak{e}}_\pm$ be a complement of $\mathfrak{e}_{0\pm}$ in $\mathfrak{e}_\pm$, and we let $\eta_{\hat{\mathfrak{e}}_\pm} = (\pi_{\mathfrak{h}_\pm}|_{\hat{\mathfrak{e}}_\pm})^{-1}$. Then $\eta_{\hat{\mathfrak{e}}_-} \wedge \eta_{H_0} \wedge \eta_{\hat{\mathfrak{e}}_-}$ effectively defines a volume form on $T_0 E$, and this form is independent of our choice of complements $\hat{\mathfrak{e}}_\pm$. What is more, we can define T as before with $\hat{\mathfrak{e}}_\pm$ in the place of $\mathfrak{e}_\pm$, and the $T_\pm^\mp$ will be independent of the choice of $\hat{\mathfrak{e}}_\pm$, and we will again have

$$\Sigma_E^*\eta_G = \det(\mathbf{1} - T_-^+T_+^-)\eta_E.$$

So far we have just defined η_E at the identity. We now suppose that $\eta_{E_-}^L$ is invariant under right multiplication by E_{0-}, and $\eta_{E_+}^R$ is invariant under left multiplication by E_{0+}. (This of course happens if both E_- and E_+ are unimodal). Then $\eta_{E_-}^L \times \eta_{E_0} \times \eta_{E_+}^R$ is invariant by the given action of $E_{0-} \times E_{0+}$, and we hence obtain a measure η_E (using our normalization on $\mathfrak{e}$) that is left-invariant by E_-, right-invariant by E_+, and bi-invariant by $E_0 = H_0$. We can then apply the same reasoning as in Lemma 3.1 to obtain (where $m\colon E \to G$ is the quotient of $m\colon E_- \times E_0 \times E_+ \to G$):

Lemma 3.2. *We have*

$$m^* d\eta_G(a_-, a_0, a_+) = q(a_0) d\eta_E.$$

where $m\colon E \to G$ *is the quotient of the multiplication map and*

$$q(a_0) = q(a_0; E_-, E_+) = \chi(a_0) \det(\mathbf{1}_{\mathfrak{h}_+} - T_+^- \circ \mathrm{Ad}_{a_0}^{-1}|_{\mathfrak{h}_+} \circ T_-^+ \circ \mathrm{Ad}_{a_0}|_{\mathfrak{h}_-}).$$

3.5. **Control of distance and measure for ζ_t and ζ.** We define $\zeta_t\colon G \times H_0 \times G \to G$ by $\zeta_t(e_-, h_0, e_+) = e_- h_0 \exp(tA) e_+$. Given $a_- \in G$ we can write $a_- = b_-^- b_0^- b_+^-$, and likewise for $a_+ \in G$. Then we have a map $\zeta\colon G \times H_0 \times G \to G$ defined by $\zeta(a_-, a_0, a_+) = b_-^- b_0^- a_0 b_0^+ b_+^+$.

Lemma 3.3. *For all compact* $K \subset G \times H_0 \times G$*, there exists* C *such that for all* $a \in K$*,*

$$d(Z^{-1}(\zeta(a)), Z_t^{-1}(\zeta_t(a))) < Ce^{-\lambda_1 t}.$$

Proof. Given $a = (a_-, a_0, a_+) \in G \times H_0 \times G$, we can write $a_- = b_-^- b_0^- b_+^-$ and likewise for a_+. We can find unique $\check{b}_- \in H_-$ and $\check{b}_+ \in H_+$ such that

$$C^{-1}_{\exp(tA/2)}(b_+^-) C_{a_0 \exp(tA/2)}(b_-^+) = \check{b}_- \check{b}_+.$$

We then obtain

$$\begin{aligned}
a_- \exp(tA) a_0 a_+ &= b_-^- b_0^- b_+^- \exp(tA) a_0 b_-^+ b_0^+ b_+^+ \\
&= b_-^- \exp(tA/2) b_0^- C^{-1}_{\exp(tA/2)}(b_+^-) C_{a_0 \exp(tA/2)}(b_-^+) a_0 b_0^+ \exp(tA/2) b_+^+ \\
&= b_-^- \exp(tA/2) b_0^- \check{b}_- \check{b}_+ a_0 b_0^+ \exp(tA/2) b_+^+ \\
&= b_-^- \check{\check{b}}_- \exp(tA/2) b_0^- a_0 b_0^+ \exp(tA/2) \check{\check{b}}_+ b_+^+ \\
&= b_-^- \check{\check{b}}_- b_0^- a_0 b_0^+ \exp(tA) \check{\check{b}}_+ b_+^+ \\
&= b_-^- b_0^- a_0 b_0^+ \check{\check{\check{b}}}_- \exp(tA) \check{\check{b}}_+ b_+^+.
\end{aligned}$$

Hence

$$Z_t^{-1}(\zeta_t(a)) = (b_-^- b_0^- a_0 b_0^+ \check{\check{\check{b}}}_-, (b_+^+)^{-1} \check{\check{b}}_+^{-1})$$

while

$$Z^{-1}(\zeta(a)) = (b_-^- b_0^- a_0 b_0^+, (b_+^+)^{-1}).$$

We observe that $\check{\check{\check{b}}}_-$ and $\check{\check{b}}_+$ lie in a $O(e^{-\lambda_1 t})$ neighborhood of $\mathbf{1}$. The Lemma follows. □

We observe that in the setting of Section 3.4, ζ and ζ_t descend to E, and we can restate Lemma 3.1 as

Lemma 3.4. *For all compact $K \subset E$, there exists C such that for all $a \in K$,*

$$d(Z^{-1}(\zeta(a)), Z_t^{-1}(\zeta_t(a))) < Ce^{-\lambda_1 t}.$$

Now (in the less general setting), let's restrict ζ and ζ_t to $E_- \times H_0 \times E_+$.

Lemma 3.5. *We have, on any compact $K \subset E_- \times H_0 \times E_+$,*

$$\left| \frac{e^{-tK_A} \zeta_t^* \eta_G}{\chi(h_0) \eta_{E_-}^L \times \eta_{H_0} \times \eta_{E_+}^R} - 1 \right| < C_K e^{-2\lambda_1 t}. \tag{22}$$

Proof. We let $M = \max(\|T_-\|, \|T_+\|)$. Then for all h_0 for which $\|\mathrm{Ad}_{h_0}\|, \|\mathrm{Ad}_{h_0}^{-1}\| < M'$, we have

$$q(h_0 \exp(tA), E_-, E_+) = e^{tK_A} \chi(h_0) \det(\mathbf{1}_{\mathfrak{h}_+} - T_+ \circ \mathrm{Ad}_{h_0 \exp(tA)}^{-1} |_{\mathfrak{h}_+} \circ T_- \circ \mathrm{Ad}_{h_0 \exp(tA)} |_{\mathfrak{h}_-}).$$

Now, for any linear transformation $T \colon V \to V$ with $\|T\| < 1$,

$$|1 - \det(\mathbf{1} - T)| < 2(\dim V) \|T\|.$$

Therefore, for t sufficiently large given M and M', we have

$$\left| 1 - \det(\mathbf{1}_{\mathfrak{h}_+} - T_+ \circ \mathrm{Ad}_{h_0 \exp(tA)}^{-1} |_{\mathfrak{h}_+} \circ T_- \circ \mathrm{Ad}_{h_0 \exp(tA)} |_{\mathfrak{h}_-}) \right| < 2(\dim H_+) M^2 M'^2 e^{-2\lambda_1 t}$$

when the right hand side is less than 1. □

We have the following remarkable corollary, which may or may not have a simpler proof:

Corollary 3.6.

$$\zeta^* d\eta_G = \chi(h_0) \eta_{E_-}^L \times \eta_{H_0} \times \eta_{E_+}^R \tag{23}$$

Proof. Let $d\eta_{E_\pm} = \zeta * d\eta_G = (\zeta \circ Z^{-1})^* d\eta_{H_{0-} \times H_+}^L$, and let $d\eta_{E_\pm^t} = e^{-K_A t} \zeta_t^* d\eta_G = (\zeta_t \circ Z_t^{-1})^* d\eta_{H_{0-} \times H_+}^L$. We let $\eta_{E_\pm}$ be the measure from integrating against $d\eta_{E_\pm}$, and likewise for $\eta_{E^t_\pm}$. By Lemmas 3.3 and 3.5, for any $A \subset E_- \times H_0 \times E_+$, and letting $t \to \infty$,

$$\begin{aligned}\eta_{E_\pm}(A) &\leq \eta_{E^t_\pm}(\mathcal{N}_{e^{-2\lambda_1 t}}(A)) \\ &\quad \to (\chi(h_0)\eta^L_{E_-} \times \eta_{H_0} \times \eta^R_{E_+})(A).\end{aligned}$$

We likewise obtain

$$\begin{aligned}\eta_{E_\pm}(A) &\geq \eta_{E^t_\pm}(\mathcal{N}_{-e^{-2\lambda_1 t}}(A) \\ &\quad \to (\chi(h_0)\eta^L_{E_-} \times \eta_{H_0} \times \eta^R_{E_+})(\mathrm{Int}A).\end{aligned}$$

As $\eta_{E_\pm}$ is a smooth measure, the Corollary follows. □

In the more general setting, we can similarly prove

Lemma 3.7.

$$\zeta^* d\eta_G(a_-, a_0, a_+) = q(a_0) d\eta_E.$$

3.6. **The application theorem.** Suppose E_- and E_+ are as in Section 3.3. We let $\eta_{E_\pm} = \chi(h_0)\eta^L_{E_-} \times \eta_{H_0} \times \eta^R_{E_+}$.

Theorem 3.8. *Let $K \subset E_- \times H_0 \times E_+$ be compact, and take $S \subset K$. For $t \geq t_0(E_-, E_+)$, let*

$$S_t = \{a_- \exp(tA) a_0 a_+ \mid (a_-, a_0, a_+) \in S\},$$

Then, letting $\delta = C_{K,\Gamma} e^{-aqt}$, for $q = q(\Gamma)$, $a = a(E_-, E_+)$, and assuming $\epsilon(g), \epsilon(h) > \delta$,

$$(1-\delta)\eta_{E_\pm}(\mathcal{N}_{-\delta}(S)) < e^{-tK_A} \#(S_t \cap g\Gamma h) < (1+\delta)\eta_{E_\pm}(\mathcal{N}_\delta(S)), \tag{24}$$

where we take inner and outer neighborhoods in $E_- \times H_0 \times E_+$.

Proof. We let $S_t = \zeta_t(S)$. By Theorem 1.3, we have

$$(1-\delta)\eta_{H_{0-}\times H_+}(\mathcal{N}_{-\delta}(Z_t^{-1}(S_t))) < e^{-tK_A}\#(S_t \cap g\Gamma h) < (1+\delta)\eta_{H_{0-}\times H_+}(\mathcal{N}_\delta(Z_t^{-1}(S_t))). \tag{25}$$

By Lemma 3.3, we have

$$Z_t^{-1}(S_t) \subset \mathcal{N}_\delta(Z^{-1}(\zeta(S)),$$

and hence

$$\mathcal{N}_\delta(Z_t^{-1}(S_t)) \subset \mathcal{N}_{2\delta}(Z^{-1}(\zeta(S))). \tag{26}$$

Taking $\zeta^{-1} \circ Z$ to be $C_K/2$-Lipschitz on K, we have

$$(\zeta^{-1} \circ Z)(\mathcal{N}_{2\delta}((Z^{-1} \circ \zeta)(S))) \subset \mathcal{N}_{C_K\delta}(S); \tag{27}$$

combining (26) and (27), we obtain

$$\mathcal{N}_\delta(Z_t^{-1}(S_t)) \subset (Z^{-1} \circ \zeta)(\mathcal{N}_{C_K\delta}(S)). \tag{28}$$

We likewise obtain

$$\mathcal{N}_{-\delta}(Z_t^{-1}(S_t)) \supset (Z^{-1} \circ \zeta)(\mathcal{N}_{-C_K\delta}(S)). \tag{29}$$

Finally, by (22),

$$\eta_{H_{0-} \times H_+}((Z^{-1} \circ \zeta)(\mathcal{N}_{C_K\delta}(S))) = \eta_{E_\pm}(\mathcal{N}_{C_K\delta}(S)). \tag{30}$$

Combining (25), (28), (29), and (30), we obtain the Theorem. □

We likewise have the following in our more general setting, where we compute the neighborhoods with respect to a given Riemannian metric ρ on E:

Theorem 3.9. *Let $K \subset E$ be compact, and take $S \subset K$. Let $t \geq t_0(E)$, and let*

$$S_t = \{a_- \exp(tA) a_0 a_+ \mid [(a_-, a_0, a_+)] \in S\}.$$

Then, letting $\delta = C_{K,\Gamma,\rho} e^{-aqt}$,

$$(1-\delta)\eta_{E_\pm}(\mathcal{N}_{-\delta}(S)) < e^{-tK_A} \#(S_t \cap g\Gamma h) < (1+\delta)\eta_{E_\pm}(\mathcal{N}_\delta(S)).$$

where we take inner and outer neighborhoods in E (and multiply δ by a constant), and the (implicit) constants depend on K.

3.7. **Examples.** Let us now discuss some actual examples of counting situations.

Orthogeodesic connections in $\mathbb{H}^3/\Gamma$. Suppose that $\Gamma < \mathrm{Isom}(\mathbb{H}^3)$ is a lattice (possibly nonuniform), and let $M = \Gamma\backslash\mathbb{H}^3$. Suppose that α and β are (oriented) geodesic segments in M. For each orthgeodesic connection η between α and β, we can record the feet of η on α and β, the length of η, and the monodromy of η (for example the angle that α, parallel translated along η, makes with β). We can even think of the real length of η and the monodromy of η as the complex length: it is the complex distance along η between α and β. In this way the set of such η is a set of points in $N^1(\alpha) \times N^1(\beta) \times \mathbb{C}/2\pi i\mathbb{Z}$.

In this example both E_- and E_+ are the centralizer of the orthogonal flow, which is just the centralizer of the geodesic flow, conjugated by a rotation by $\pi/2$. We have

$$\eta_{E_\pm} = q(a_0) d\eta^L_{E_-} \wedge d\eta_{H_0} \wedge d\eta^R_{E_+},$$

where $q(a_0) = C_0 e^{2a_0}$, and C_0 is a constant that I am currently too lazy to calculate. But $\eta^L_{E_-}$ and $\eta^R_{E_+}$ are just the natural measures on $N^1(\alpha)$ and $N^1(\beta)$, and η_{H_0} is the natural measure on $\mathbb{C}/2\pi\mathbb{Z}$. So taking $g, h \in \mathrm{Isom}(\mathbb{H}^3)$ to translate our base frame to ones in $N^1(\alpha)$ and $N^1(\beta)$ respectively, Theorem 3.8 becomes

Theorem 3.10. *The number of connections for a given subset* $A \subset K \subset N^1(\alpha) \times N^1(\beta) \times S^1 \times [L, \infty)$ *satisfies*

$$(1-\delta)\mathrm{Vol}(\mathcal{N}_{-\delta}(A)) < C(A)/(C_1 \mathrm{Vol}(M)) < (1+\delta)\mathrm{Vol}(\mathcal{N}_{\delta}(A))$$

where K *is compact and* $\delta = C_{K,\Gamma} e^{-qL}$, $q = q(\Gamma)$, *provided that the height of one of the* α *or* β *projections of* K *is at most* qL.

This theorem is sufficient for [KM12] and [KW18], but Theorems 3.8 and 3.9 have many other applications, such as counting connections (with specific monodromy) between points. For simplicity let us assume that M^n is hyperbolic, and let $x, y \in M$. We let σ_x be a section of the projection from frames at x to vectors at x, and likewise define σ_y. Then any subset of the natural quotient of $\mathcal{F}(x) \times H_0 \times \mathcal{F}(y)$ can be lifted to a subset of $T^1(x) \times H_0 \times T^1(y)$ via the sections σ_x and σ_y, and the measure on the quotient becomes the measure on $T^1(x) \times H_0 \times T^1(y)$.

Thus from Theorem 3.9 we obtain

Theorem 3.11. *The number of connections for a given subset* $A \subset K \subset T^1(x) \times H_0([L, \infty)) \times T^1(y)$ *satisfies*

$$(1-\delta)\mathrm{Vol}(\mathcal{N}_{-\delta}(A)) < C(A)/(C_1 \mathrm{Vol}(M)) < (1+\delta)\mathrm{Vol}(\mathcal{N}_{\delta}(A)) \tag{31}$$

where K *is compact and* $\delta = C_{K,\Gamma} e^{-qL}$, $q = q(\Gamma)$, *provided that the height* x *and* y *is at most* qL.

Here we should say a few words about the volume that appears in the upper and lower bounds of (31). It is $e^{\chi(a_0)}$ times the quotient of the product measure on $(a_0, a_0, a_+) \in E_- \times E_0 \times E_+$, and it is often natural and convenient to take a section of the quotient map, and use this to compute the measure.

For example, in the setting of Theorem 3.11, we can take sections of the projections $\mathcal{F}(x) \to T^1(x)$ and $\mathcal{F}(y) \to T^1(y)$. These give us a section σ of the projection $E_- \times E_0 \times E_+ \to E$. Hence, given $A \subset E$, we can think of it as $A \subset T^1(x) \times H_0 \times T^1(y)$, and $\eta_{E_\pm}(A)$ will just be $e^{\chi(a_0)}$ times the product measure of $T^1(x) \times H_0 \times T^1(Y)$. For a sufficiently smooth section, we can also use this latter product to compute our δ-neighborhood.

We can likewise count orthogeodesic connection in H^n, with $n > 3$, by again taking sections of the projection from the "aligned frame bundle" over a geodesic α to $N^1(\alpha)$, where a frame is aligned with α if its base point lies on α and its first vector is tangent to α.

Of course we can also make similar statements in other symmetric spaces, both rank 1 and higher rank.

References

[KM12] Jeremy Kahn and Vladimir Markovic. Immersing almost geodesic surfaces in a closed hyperbolic three manifold. *Ann. of Math. (2)*, 175(3):1127–1190, 2012.

[KW18] Jeremy Kahn and Alex Wright. Nearly Fuchsian surface subgroups of finite covolume Kleinian groups. *arXiv preprint arXiv:1809.07211*, 2018.