Notes on Simmons-Weiss Chapter 3

1 Preliminaries

Almost all of this is directly plagiarized from [SW]. We are keeping all the notation and theorem numbers unchanged.

Let G be a unimodular noncompact Lie group with finitely many connected components, $E \subset G$ compact, and μ a probability measure supported on E. We write $e \mapsto g_e$ for the inclusion $E \hookrightarrow G$, so in this context E is used as an indexing set for elements of g.

We denote $B = E^{\mathbb{N}}$ with Bernoulli measure $\beta = \mu^{\otimes \mathbb{N}}$ and left-shift map $T : B \to B$. If $b = (b_1, b_2, \dots) \in B$, let $b_1^n = (b_1, \dots, b_n) \in E^n$, and $g_{b_1^n} = g_{b_n} \cdots g_{b_1}$. For each $n \geq 1$, denote by μ^{*n} the pushforward to G of the product measure μ^n on E^n under the map $b_1^n \mapsto g_{b_1^n}$, i.e. for any integrable function f on G,

$$\int_G f(g)\, d\mu^{*n}(g) = \int_{E^n} f(g_{b_1^n})\, d\mu^n(b_1^n).$$

Let $V = \mathfrak{g}$ be the Lie algebra of G, and $\mathrm{Ad} : G \to \mathrm{SL}(V)$ the adjoint representation $\mathrm{Ad}(g) = (D\Psi_g)|_{g=e}$, where $\Psi_g(x) = gxg^{-1}$.

Theorem (*Oseledec*)**.** *There exist real numbers $\chi_1 > \cdots > \chi_k$ and for β-almost every $b \in B$ a descending chain of proper subspaces*

$$V = V_0 \supset V_1 \supset \cdots \supset V_{k-1} \supset V_k = 0$$

depending measurably on b, such that for all $v \in V_{i-1} \setminus V_i$,

$$\lim_{n \to \infty} \frac{1}{n} \log \|\rho(g_{b_1^n})v\| = \chi_i,$$

and the convergence is uniform as v ranges over any compact subset of $V_{i-1} \setminus V_i$. Each χ_i is called a Lyapunov exponent*, with multiplicity $d_i = \dim V_{i-1} - \dim V_i$. Furthermore, we have*

$$\sum_{i=1}^{k} d_i \chi_i = \int_G \log |\det \rho(g)|\, d\mu(g),$$

and for each i,

$$V_i(T(b)) = \rho(g_{b_1}) V_i(b).$$

We set $V_1(b) = V_b^{<\max}$, which we call the *Oseledec space of non-maximal expansion*, and $V_{j_0}(b) = V_b^{\leq 0}$ where $j_0 = \max\{j : \chi_j > 0\}$ is called the *Oseledec space of subexponential expansion.*

For each $d = 1, \dots, \dim G - 1$, let $V^{\wedge d}$ be the dth exterior power of V. Then Ad induces a representation $\rho_d : G \to \mathrm{SL}(V^{\wedge d})$. Assume that for each d there is a nontrivial proper subspace $W^{\wedge d} \subset V^{\wedge d}$ such that

(I) For every $g \in E$, $W^{\wedge d}$ is $\rho_d(g)$-invariant. For β-almost every $b \in B$, if $d = 1$ then $W^{\wedge d}$ is complementary to $V_b^{<\max}$ and if $d > 1$ then $V_b^{\leq 0} \cap W^{\wedge d} = 0$.

(II) For every $g \in E$, $\mathrm{Ad}(g)$ acts on $W = W^{\wedge 1}$ as a similarity map, and

$$\int_G \log \|\mathrm{Ad}(g)|_W\| \, d\mu(g) > 0.$$

(III) For any d, if a linear subspace $L \subset V^{\wedge d}$ has a finite orbit under the semigroup generated by E, then $L \cap W^{\wedge d} \neq 0$.

Fix an inner product on $V^{\wedge d}$ inducing a metric and an operator norm on $\mathrm{GL}(V^{\wedge d})$. Let $\mathbb{P}(V^{\wedge d})$ be the projective space of lines in $V^{\wedge d}$. If $x \in V^{\wedge d} \setminus \{0\}$, we write $[x]$ for its image under the quotient map. Similarly, if $W \subset V^{\wedge d}$ is a nonzero subspace, its image is $[W]$. The distance between a vector $v \in V^{\wedge d}$ and a subspace $W \subset V^{\wedge d}$ will be denoted $\mathrm{dist}(v, W)$, and the distances between their projectivizations will also be denoted $\mathrm{dist}([v], [W])$.

2 Preparatory lemmas

Proposition 3.5. *Under assumptions (I) and (II), for β-almost every $b \in B$, for any compact $C \subset V \setminus V_b^{<\max}$ there exists $c > 0$ such that for all $v \in C$ and $n \geq 1$,*

$$\|\mathrm{Ad}(g_{b_1^n})v\| \geq c\|\mathrm{Ad}(g_{b_1^n})\|.$$

The proof uses the following notation: $A \asymp_x B$ if there exists $c = c(x)$ with $c^{-1} \leq \frac{A}{B} \leq c$.

Proof. By Assumption (I) with $d = 1$, we can write $v = \pi_1(v) + \pi_W(v)$ with $\pi_1(v) \in V_b^{<\max}$ and $\pi_W(v) \in W \setminus 0$. By Oseledec's theorem,

$$\frac{\|\mathrm{Ad}(g_{b_1^n})\pi_1(v)\|}{\|\mathrm{Ad}(g_{b_1^n})\pi_W(v)\|} \to 0$$

as $n \to \infty$, so $\|\mathrm{Ad}(g_{b_1^n})v\| \asymp_{b,v} \|\mathrm{Ad}(g_{b_1^n})\pi_W(v)\|$, using the triangle inequality. By this and Assumption (II),

$$\|\mathrm{Ad}(g_{b_1^n})v\| \asymp_{b,v} \|\mathrm{Ad}(g_{b_1^n})\pi_W(v)\| \asymp_{b,v} \|\mathrm{Ad}(g_{b_1^n})|_W\|$$

By choosing a basis of V consisting of vectors that do not belong to $V_b^{<\max}$, we have

$$\|\mathrm{Ad}(g_{b_1^n})|_W\| \asymp_b \|\mathrm{Ad}(g_{b_1^n})\|.$$

By Oseledec's theorem, we can take the constant independent of v in a compact subset. □

Proposition 3.6. *Under assumptions (I) and (II), for β-almost every $b \in B$ and for all $v \in V \setminus V_b^{<\max}$, we have*

$$\frac{\mathrm{dist}(\mathrm{Ad}(g_{b_1^n})v, W)}{\|\mathrm{Ad}(g_{b_1^n})\|} \to 0$$

as $n \to \infty$, and hence

$$\mathrm{dist}([\mathrm{Ad}(g_{b_1^n})v], [W]) \to 0.$$

For fixed b the convergence is uniform for v in a compact subset of $V \setminus V_b^{<\max}$.

Proof. By Assumption (I), choose $w \in W$ such that $v - w \in V_b^{<\max}$. Again by (I) we have $\mathrm{Ad}(g_{b_1^n})w \in W$ for all n. Then

$$\begin{aligned}\mathrm{dist}(\mathrm{Ad}(g_{b_1^n})v, W) &= \inf_{w\in W} \|\mathrm{Ad}(g_{b_1^n})v - w\| \\ &\le \|\mathrm{Ad}(g_{b_1^n})v - \mathrm{Ad}(g_{b_1^n})w\| \\ &= \|\mathrm{Ad}(g_{b_1^n})(v-w)\| \\ &\sim e^{n\chi_2} = o(\|\mathrm{Ad}(g_{b_1^n})\|),\end{aligned}$$

which demonstrates the first statement. For the second statement, because $[v] = \frac{v}{\|v\|}$ up to a sign, we use Proposition 3.5. ☐

Proposition 3.7. *Assume that (I) and (III) hold. Fix $d \in \{1, \dots, \dim G - 1\}$ and $v \in V^{\wedge d}$. Then we have $v \notin V_b^{\le 0}$ for β-almost every $b \in B$, and if $d = 1$ then $v \notin V_b^{<\max}$.*

Proof. We will prove it for $d = 1$. The proof for $d > 1$ follows by substituting $V_b^{\le 0}$ for $V_b^{<\max}$ and ρ_d for Ad.

Let $\mu^{*i} * \delta_{[v]}$ be the pushforward of $\mu^{*i} \otimes \delta_{[v]}$ under the map $(g, [v]) \mapsto [\mathrm{Ad}(g)v]$; note that this is consistent with the usual convolution notation

$$\mu * \nu(f) = \int f(gx)\, d\mu(g)\, d\nu(x).$$

Using the shift-equivariance property from Oseledec's theorem and Fubini's theorem,

$$\begin{aligned}\int_B \mu * \delta_{[v]} \left([V_b^{<\max}]\right)\, d\beta(b) &= \int_B \mu \otimes \delta_{[v]} \left(\{(b_1, v) : \mathrm{Ad}(g_{b_1})v \in V_b^{<\max}\}\right)\, d\beta(b) \\ &= \int_B \mu \otimes \delta_{[v]}(\{(b_1, v) : v \in V_{b_1 b}^{<\max}\})\, d\beta(b) \\ &= \int_B \left(\int_E \mu \otimes \delta_{[v]}(\{(b_1, v) : v \in V_{b_1 b}^{<\max}\})\, d\mu(b_1)\right) d\beta(b) \\ &= \int_B \delta_{[v]}([V_b^{<\max}])\, d\beta(b).\end{aligned}$$

By induction on i, we then have

$$\int_B \mu^{*i} * \delta_{[v]} \left([V_b^{<\max}]\right)\, d\beta(b) = \int_B \delta_{[v]}([V_b^{<\max}])\, d\beta(b)$$

for each $i \ge 0$. For each $n \ge 1$ define

$$\nu_n = \frac{1}{n} \sum_{i=0}^{n-1} \mu^{*i} * \delta_{[v]},$$

a probability measure on the compact space $\mathbb{P}(V)$. Then by the above computation,

$$\beta\left(\{b \in B : v \in V_b^{<\max}\}\right) = \int_B \delta_{[v]}([V_b^{<\max}])\, d\beta(b) = \int_B \nu_n([V_b^{<\max}])\, d\beta(b)$$

for all $n \ge 1$. Thus it remains to show that the right-hand integral $\to 0$ as $n \to \infty$. By the Lebesgue dominated convergence theorem, it suffices to show that for β-almost every $b \in B$, $\nu_n([V_b^{<\max}]) \to 0$ as $n \to \infty$.

Suppose by contradiction there exists $\epsilon > 0$ and a set B_0 of positive β measure such that for all $b \in B_0$, there exists a subsequence $n_k \to \infty$ with $\nu_{n_k}([V_b^{<\max}]) \geq \epsilon$. Let $V' = V_{b_0}^{<\max}$ for some $b_0 \in B_0$, and let ν_∞ be a weak-* limit point of ν_{n_k}. Because $\mu * \nu_{n_k} - \nu_{n_k} \to 0$ by telescoping, ν_∞ is μ-stationary and satisfies $\nu_\infty([V']) \geq \epsilon$.

Some terminology: a stationary measure ν is said to be *μ-ergodic* if it cannot be written as a proper convex combination of two different μ-stationary Borel probability measures. Then by the Krein-Milman theorem we can decompose any μ-stationary measure into a convex combination of μ-ergodic measures called *ergodic components* of ν.

Thus we may assume that ν_∞ has an ergodic component ν'_∞ with $\nu'_\infty([V']) > 0$. Let $k \leq \dim V$ be minimal with respect to the property that some k-dimensional subspace has positive ν'_∞-measure. Then the intersection of any two distinct k-dimensional subspaces of V has measure zero, so ν'_∞ is an additive atomic measure on the set of all such subspaces. We have the following:

Lemma. If μ and ν are probability measures on G and X respectively, $G \curvearrowright X$, and ν is μ-ergodic (and μ-stationary by definition) and atomic, then it is G-invariant and finitely supported.

Proof. Let $x \in X$ be an atom with maximal ν-measure. Then the orbit Gx is G-invariant with positive measure, hence $\mu(Gx) = 1$ by ergodicity. By stationarity,

$$\nu(x) = \sum_g \mu(g)\nu(gx) \leq \max_g \nu(gx) \sum_g \mu(g) = \max_g \nu(gx),$$

and since $\nu(x)$ is maximal, $\nu(x) = \nu(gx)$ for all $g \in G$. Hence ν is just the uniform measure on the finite set Gx, and in particular it is G-invariant.

Applying the lemma to ν'_∞, there exists a finite $\mathrm{supp}(\mu) = E$-invariant collection of subspaces $\{L_1, \ldots, L_r\}$ with $\mathrm{supp}(\nu'_\infty) \subset \bigcup L_i$. By Assumption (III), each $L_i \cap W \neq \emptyset$. But V' is complementary to W by Assumption (I), so $L_i \cap V'$ has dimension strictly less than k, so $\nu'_\infty([L_i \cap V']) = 0$, so $\nu'_\infty([V']) = 0$, a contradiction. $\square$

3 The main theorems

Proposition 3.3. *Under assumptions (I)-(III), there exists $n_0 \geq 1$ and $\epsilon > 0$ such that for all d, $v \in V^{\wedge d} \setminus \{0\}$ and $n \geq n_0$,*

$$\frac{1}{n} \int_G \log \frac{\|\rho_d(g)v\|}{\|v\|} \, d\mu^{*n}(g) > \epsilon.$$

Proof. Fix $\alpha > 0$ to be specified below. By Proposition 3.7, for each $v' \in V^{\wedge d}$ there exists $\epsilon_0 = \epsilon_0(v')$ and $B_0 = B_0(v') \subset B$ such that $\beta(B_0) \geq 1 - \alpha$ and for all $b \in B_0$, $\mathrm{dist}([v'], [V_b^{\leq 0}]) \geq \epsilon_0$. If $0 < \epsilon_1 < \epsilon_0$, there is a neighborhood $U = U(v')$ of $[v']$ in $\mathbb{P}(V^{\wedge d})$ such that for all $b \in B_0(v')$ and $[v] \in U$, $\mathrm{dist}([v], [V_b^{\leq 0}]) \geq \epsilon_1$. Now choose a finite cover $\{U_1, \ldots, U_k\}$ of $\mathbb{P}(V^{\wedge d})$ and a finite collection $\{B_1, \ldots, B_k\}$ of subsets of B such that $\beta(B_j) \geq 1 - \alpha$ for all j, and for all $b \in B_j$ and $[v] \in U_j$, $\mathrm{dist}([v], [V_b^{\leq 0}]) \geq \epsilon_1$. Now let $\chi > 0$ be strictly less than the smallest positive Lyapunov exponent of the action on $V^{\wedge d}$. By the uniformity on compact subsets from Oseledec's theorem, for each j there exists n_j such that for all $n \geq n_j$, $[v] \in U_j$, and $b \in B_j$ we have

$$\|\rho_d(g_{b_1^n})v\| \geq e^{n\chi}\|v\|.$$

Let $N = \max_j n_j$, and for each $v \in V^{\wedge d} \setminus \{0\}$ and $n \geq N$ let

$$S = S_{n,v} = \{b_1^n \in E^n : \|\rho_d(g_{b_1^n})v\| \geq e^{n\chi}\|v\|\},$$

so that if $[v] \in U_j$ and $b \in B_j$ then $b_1^n \in S_{n,v}$ for all $n \geq N$. Since $\beta(B_0(v')) \geq 1-\alpha$, we have $\mu^{\otimes n}(S) \geq 1-\alpha$, and thus

$$\begin{aligned}
\frac{1}{n}\int_G \log\frac{\|\rho_d(g)v\|}{\|v\|}\,d\mu^{*n}(g) &= \frac{1}{n}\int_{E^n} \log\frac{\|\rho_d(g_{b_1^n})v\|}{\|v\|}\,d\mu^{\otimes n}(b_1^n) \\
&\geq \frac{1}{n}\int_S \log e^{n\chi}\,d\mu^{\otimes n} + \frac{1}{n}\int_{E^n\setminus S} \log\|\rho_d(g_{b_1^n})^{-1}\|^{-1}\,d\mu^{\otimes n}(b_1^n) \\
&\geq \frac{1}{n}\big((1-\alpha)n\chi - \alpha n\log\max_{g\in\operatorname{supp}(\mu)}\|\phi_d(g)^{-1}\|\big) \\
&\geq (1-\alpha)\chi - \alpha\log\max_{g\in\operatorname{supp}(\mu)}\|\phi_d(g)^{-1}\|.
\end{aligned}$$

The second inequality comes from the fact that $\|v\| = \|A^{-1}Av\| \leq \|A^{-1}\|\|Av\|$. To complete the proof, choose α small enough so the last expression is positive and independent of v. □

Proposition 3.1. *Under assumptions (I)-(III) we have*

1. *For every $\alpha > 0$ there exists $c_0 > 0$, $q_0 \geq 1$ such that for any $v \in V\setminus\{0\}$ we have*

$$\beta\big(\{b\in B : \forall q \geq q_0, \|\mathrm{Ad}(g_{b_1^q})v\| \geq c_0\|\mathrm{Ad}(g_{b_1^q})\|\|v\|\}\big) \geq 1-\alpha.$$

2. *For every $\alpha,\eta > 0$ there exists $q_0 \geq 1$ such that for any $v \in V\setminus\{0\}$ we have*

$$\beta\big(\{b\in B : \forall q \geq q_0, \operatorname{dist}([\mathrm{Ad}(g_{b_1^q})v],[W]) \leq \eta\}\big) \geq 1-\alpha.$$

Proof. Fix $\alpha,\eta > 0$. By Proposition 3.7 and a compactness argument similar to the one used in the proof of Proposition 3.3, there exists $\epsilon > 0$ such that for all $v \in V\setminus\{0\}$,

$$\beta\big(\{b\in B : \forall q \geq q_0, \operatorname{dist}([v],[V_b^{<\max}]) \geq \epsilon\}\big) \geq 1-\alpha/2.$$

For each $b \in B$, define $N(b)$ as the smallest integer such that for all $v \in V$ with $\operatorname{dist}([v],[V_b^{<\max}]) \geq \epsilon$ and all $n \geq N(b)$, we have

$$\|\mathrm{Ad}(g_{b_1^n})v\| \geq \frac{1}{N(b)}\|\mathrm{Ad}(g_{b_1^n})\|\|v\|, \text{ and } \operatorname{dist}([\mathrm{Ad}(g_{b_1^q})v],[W]) \leq \eta, \tag{1}$$

so that $N(b) < \infty$ for β-almost every $b \in B$, by Propositions 3.5 and 3.6. Thus there exists N_0 such that

$$\beta(\{b\in B : N(b) \leq N_0\}) \geq 1-\alpha/2.$$

Now fix $v \in V\setminus\{0\}$. For all $b \in B$ such that $\operatorname{dist}([v],[V_b^{<\max}]) \geq \epsilon$ and $N(b) \leq N_0$, and for all $n \geq N_0$, Equation 1 holds, and we are done. □

Proposition 3.2. *Under assumptions (I) and (III), for each $d = 1,\ldots,\dim G-1$ the only μ-stationary probability measure on $V^{\wedge d}$ is the Dirac measure δ_0 centered at 0.*

Proof. Let ν be a μ-stationary measure on $V^{\wedge d}$ which is not equal to the Dirac measure δ_0, let $Z = B\times V^{\wedge d}$, let $\lambda = \beta\times\nu$, and let

$$Y = \{(b,z)\in Z : v \notin V_b^{\leq 0}\}.$$

By Proposition 3.7, $\lambda(Y) = 1$. Now define $\hat{T} : Z \to Z$ by $\hat{T}(b, v) = (Tb, \phi_d(g_{b_1})v)$. Because ν is μ-stationary, λ is $\hat{T}$-invariant:

$$\begin{aligned}
\int_Z f \circ \hat{T} \, d\lambda &= \int_Z f(Tb, \phi_d(g_{b_1})v) \, d\beta(b) \, d\nu(v) \\
&= \int_Z f((b_2, b_3, \ldots), \phi_d(g_{b_1})v) \, d\mu(b_1) \, d\beta(b_2, b_3, \ldots) \, d\nu(v) \\
&= \int_Z f((b_2, b_3, \ldots), v) \, d\beta(b_2, b_3, \ldots) \, d\nu(v) \\
&= \int_Z f(b, v) \, d\beta(b) \, d\nu(v) \\
&= \int_Z f \, d\lambda
\end{aligned}$$

Furthermore, $\|\rho_d(g_{b_1^n})v\| \to \infty$ for every $(b, v) \in Y$, by definition of Y. Let $t > 0$ be large enough so that $\lambda(Y_0) > 0$, where

$$Y_0 = \{(b, v) \in Y : \|v\| \leq t\}.$$

Then for all $(b, v) \in Y_0$, for all n large enough, $\hat{T}^n(b, v) \notin Y_0$, which contradicts Poincaré recurrence.

$\square$