ON SIMULTANEOUS LINEARIZATION OF DIFFEOMORPHISMS OF THE SPHERE.

DMITRY DOLGOPYAT AND RAPHAËL KRIKORIAN.

Abstract. Let $R_1, R_2 \dots R_m$ be rotations generating $\mathbb{SO}_{d+1}$, $d \geq 2$, and $f_1, f_2 \dots f_m$ be their small smooth perturbations. We show that $\{f_\alpha\}$ can be simultaneously linearized if and only if the associated random walk has zero Lyapunov exponents. As a consequence we obtain stable ergodicity of actions of random rotations in even dimensions.

1. Main results.

Let $f_1, f_2 \dots f_m$ be diffeomorphisms of $\mathbb{S}^d$, $d \geq 2$. Let $\omega = \{\omega_i\}_{i \in \mathbb{N}}$ be a sequence of independent random variables uniformly distributed on $\{1 \dots m\}$. Consider the Markov process on $\mathbb{S}^d$

$$x_n = f_{\omega_n} x_{n-1}. \tag{1}$$

If μ is an invariant measure for this process let

$$\lambda_1(\mu) \geq \lambda_2(\mu) \dots \geq \lambda_d(\mu)$$

be the Lyapunov exponents of μ. Denote

$$\Lambda_r = \sum_{j=1}^{r} \lambda_j.$$

Theorem 1. *Given d there exists a number k_0 such that for any m for any set of rotations $R_1 \dots R_m$ in $\mathbb{SO}_{d+1}$ such that $R_1 \dots R_m$ generate $\mathbb{SO}_{d+1}$ there exists a number $\varepsilon > 0$ such that if $\max_\alpha d_{C^{k_0}}(R_\alpha, f_\alpha) < \varepsilon$ then either*

(a) there exists $c > 0$ such that $\lambda_d(\mu) < -c$ for any invariant measure μ or

(b) f_α are simultaneously conjugated to rotations.

Remark. *Some analogies of our results for (non-measure preserving) diffeomorphisms of $\mathbb{S}^1$ can be found in* [24, 25] *(see also survey* [14]*).*

2000 *Mathematics Subject Classification.* 37C85, 37H15, 70H08.

Theorem 2. *Let $\{R_\alpha\}$ be as in Theorem 1.*

(a) In the category of volume preserving diffeomorphisms

$$\left|\lambda_r(\mu) - \frac{d-2r+1}{d-1}\lambda_1(\mu)\right| \ll |\lambda_d| \quad as \quad \max_\alpha d_{C^{k_0}}(f_\alpha, R_\alpha) \to 0$$

(here '$\ll$' means that either both the LHS and the RHS are 0 or their ratio can be made arbitrary close to 1 by taking $\max_\alpha d_{C^{k_0}}(f_\alpha, R_\alpha)$ small).

(b) In general (without the volume preservation assumption)

$$\left|\lambda_r - \frac{\Lambda_d}{d} - \frac{d-2r+1}{d-1}\left(\lambda_1 - \frac{\Lambda_d}{d}\right)\right| \ll |\lambda_d| \quad as \quad \max_\alpha d_{C^{k_0}}(f_\alpha, R_\alpha) \to 0.$$

Theorem 2 says that regardless of the dimension of the problem the Lyapunov exponents asymptotically depend only on two parameters. That is, knowing λ_d and Λ_d we can compute all exponents with high accuracy.

The relation

$$\lambda_r = \frac{\Lambda_d}{d} + \frac{d-2r+1}{d-1}\left(\lambda_1 - \frac{\Lambda_d}{d}\right)$$

is not new in the theory of random diffeomorphisms. Namely, it holds for isotropic Brownian flows on $\mathbb{R}^d$ (see [22]).

The asymptotic expressions for λ_r will be given in Section 5. For generic $\{f_\alpha\}$, λ_r are quadratic in $d(\{f_\alpha\}, \{R_\alpha\})$. More precisely we need to measure the distance between $\{f_\alpha\}$ and the systems obtained from rotations (maybe different from $\{R_\alpha\}$) by a change of variables.

We now state two consequences of our main results.

Corollary 1. *If $\{f_\alpha\}$ are C^0 conjugated to rotations then they are C^∞ conjugated to rotations.*

Remark. *Some sufficient conditions for $\{f_\alpha\}$ being C^0-conjugated to rotations are given in* [12].

Corollary 2. *If d is even and $R_1 \dots R_m$ generate $\mathbb{SO}_{d+1}$ then the system $\{R_\alpha\}_{\alpha=1}^m$ is stably ergodic. That is if f_α are sufficiently close to R_α and preserve volume then $\{f_\alpha\}$ is ergodic.*

Recently there was a significant progress in the study of stable ergodicity of a single diffeomorphism (see review [4]). Corollary 2 gives a first example of a stably ergodic system where each individual diffeomorphism is not stably ergodic. In fact, for one diffeomorphism it is known that some hyperbolicity is needed for stable ergodicity, since in the elliptic setting KAM theory applies (see e.g. [30]). By contrast our result shows that for several diffeomorphisms ellipticity does not contradict to stable ergodicity. Therefore the following conjecture is natural.

Conjecture. *Let M be a compact manifold and k be a sufficiently large number. Let $m \geq 2$. Then stable ergodicity is open and dense among m-tuples of C^k volume preserving diffeomorphisms of M.*

The proofs of the theorems occupy Sections 2–7. The proofs of the corollaries are given in Sections 8–11.

2. Notation and background.

2.1. Given a sequence ω we let $F_n(\omega) = f_{\omega_n} \circ \dots f_{\omega_2} \circ f_{\omega_1}$.
We shall write $d_k(f, R) = \max_\alpha d_{C^k}(f_\alpha, R_\alpha)$.

2.2. We denote by $G_{r,d}$ the bundle of r dimensional planes in $T\mathbb{S}^d$. Let $\mathbb{V}_d^s$ denote the space of C^s–vectorfields on $\mathbb{S}^d$. Given natural actions of $\mathbb{SO}_{d+1}$ on $C^s(\mathbb{S}^d)$, $C^s(G_{r,d})$, $C^s(\mathbb{SO}_{d+1})$, and $\mathbb{V}_d^s$ let $-\Delta$ be the image of the Casimir operator. (The properties of Δ used in this paper could be found for example in [13].) We let $H_\lambda(\mathbb{S}^d)$, $H_\lambda(G_{r,d})$, $H_\lambda(\mathbb{SO}_{d+1})$, $H_\lambda(\mathbb{V}_d^s)$ be the space of eigenvectors of Δ with eigenvalue λ. Let $\mathcal{M}$ denote the operator acting on the space of functions (the functions can be defined either on $\mathbb{S}^d$ or $\mathbb{SO}_{d+1}$) as follows

$$(\mathcal{M}A)(x) = \frac{1}{m} \sum_{\alpha=1}^{m} A(R_\alpha x). \tag{2}$$

Let $\mathcal{L}$ act on $\mathbb{V}_d^s$ as follows

$$(\mathcal{L}X)(x) = \frac{1}{m} \sum_{\alpha=1}^{m} dR_\alpha X(R_\alpha^{-1} x). \tag{3}$$

Let $\mathcal{L}_\lambda$ and $\mathcal{M}_\lambda$ be restrictions of $\mathcal{L}$ and $\mathcal{M}$ to H_λ. Below we discuss the spectrum of $\mathcal{L}$, the results for $\mathcal{M}$ are identical.

Proposition 1. *([8]) There exist constants $k_1(d, m), k_2(d, m)$, such that for any rotations $R_1 \dots R_m$ generating $\mathbb{SO}_{d+1}$ there exist constants $C_1(R_1 \dots R_m), C_2(R_1 \dots R_m)$ such that*

$$||\mathcal{L}_\lambda^n|| \leq C_1 \lambda^{k_1} \left(1 - \frac{1}{C_2 \lambda^{k_2}}\right)^n.$$

Moreover C_1 and C_2 can be chosen to depend continuously on $R_1 \dots R_m$.

2.3. We denote by $|| \quad ||_s$ the usual C^s-norms and by $|| \quad ||_{H^s}$ the Sobolev norms: $||A||_{H^s} =< (I+\Delta)^s A, A >^{1/2}$. By elliptic regularity of the Laplacian these norms satisfy

$$||A||_s \leq C_s ||A||_{H^{s+a}}, \qquad ||A||_{H^s} \leq C_s ||A||_{s+a} \tag{4}$$

for some fixed constant a (here and throughout the paper C_s denote some constants which value is not fixed).

The following estimates on products and compositions are true:

$$(I) \qquad ||AB||_s \leq C_s(||A||_s ||B||_0 + ||A||_0 ||B||_s)$$
$$(II) \qquad ||\phi \circ A||_s \leq C_s ||\phi||_s (1+||A||_0)^s (1+||A||_s)$$

and if ϕ is quadratic in A, B (that is $\phi(0,0)=0$, $D\phi(0,0)=0$) we have without expliciting the dependence in ϕ

$$(III) \qquad ||\phi(A,B)||_s \leq C_s(1+||A||_0+||B||_0)^{s+1}(||A||_0+||B||_0)(||A||_s+||B||_s).$$

Inequality (I) is proven in [16], Theorem A.7. Inequality (II) is proven by induction on s using the fact that

$$||A||_1 ||A||_s \leq C_s ||A||_0 ||A||_{s+1}$$

(this follows from the Hadamard inequalities $||A||_t \leq C_s ||A||_{t_0}^{a_0} ||A||_{t_1}^{a_1}$ if $t = a_0 t_0 + a_1 t_1$, $a_0 + a_1 = 1$, see [16], Lemma A.2) and the third inequality follows from the second since any quadratic ϕ can be written as

$$\phi(x,y) = q_1(x,x)\psi_1(x,y) + q_2(x,y)\psi_2(x,y) + q_3(y,y)\psi_3(x,y)$$

where q_1, q_2, q_3 are bilinear forms and $\psi_i(x,y)$, $i=1,2,3$ are smooth functions the C^s-norms of which are related to those of ϕ.

2.4. Let $\mathcal{T}_\lambda$ and $\mathcal{R}_\lambda$ denote the projections on

$$\oplus_{\bar\lambda \leq \lambda} H_{\bar\lambda} \quad \text{and} \quad \oplus_{\bar\lambda > \lambda} H_{\bar\lambda}.$$

We then have for $\bar s \geq s$

$$||\mathcal{T}_\lambda A||_{\bar s} \leq C_s \lambda^{k_3 + ((\bar s - s)/2)} ||A||_s, \tag{5}$$

$$||\mathcal{R}_\lambda A||_s \leq C_s \lambda^{k_3 - ((\bar s - s)/2)} ||A||_{\bar s}. \tag{6}$$

Indeed the above inequalities are obvious for Sobolev norms (with $k_3 = 0$). Observe that for Sobolev norms (5) is true even without the restriction $s \leq \bar s$.

Now to get (5) use (4) to compare $||\mathcal{T}_\lambda A||_{\bar s}$ with $||\mathcal{T}_\lambda A||_{H^{\bar s + a}}$ and $||A||_s$ with $||A||_{H^{s-a}}$.

To get (6) consider two cases:

(I) $\bar s \geq s + 2a$. Then we can argue as for (5) comparing Sobolev and smooth norms.

(II) $\bar{s} < s+2a$. Then the result follows from the fact that $\mathcal{R}_\lambda = 1-\mathcal{T}_\lambda$ and (5).

2.5. Proposition 1 and the estimates of Sections 2.3, 2.4 imply that there exists constants b, γ such that

$$||\mathcal{L}^n X||_s = \mathcal{T}_0 X + O\left(\frac{||X||_{\bar{s}+b}}{n^{\gamma(\bar{s}-s)}}\right) \tag{7}$$

for all $s \leq \bar{s}$.

Observe that if we define $\mathcal{K}(X) = -\sum_{j=1}^{\infty} \mathcal{L}^j X$, we have

$$||\mathcal{K}(\mathcal{T}_\lambda X)||_s \leq C_s \lambda^{k_4} ||X||_s$$

for all X with $\mathcal{T}_0 X = 0$.

2.6. Let $X_\alpha(x) = [\exp_x^{-1}(R_\alpha^{-1} f_\alpha x)]$ and $\varepsilon_s = \max_\alpha(||X_\alpha||_s)$. For a vectorfield Y let $\psi_Y(x) = \exp_x(Y(x))$. We now make a change of variables $\tilde{x} = \psi_Y(x)$ where Y is a small vectorfield. Then $\tilde{f}_\alpha = \psi_Y f_\alpha \psi_Y^{-1}$ corresponds (up to higher order terms) to

$$\tilde{X}_\alpha = X_\alpha - Y + R_\alpha^{-1} Y \tag{8}$$

where $(R_\alpha^{-1} Y)$ is a shortcut for $dR_\alpha^{-1} Y(R_\alpha(x))$. Our goal is to find Y so that $\tilde{X}_\alpha$ has the simplest possible form.

3. Plan of the proof of Theorem 1.

3.1. **Invariant measures.** Our starting point is to observe that since R_α generate $\mathbb{SO}_{d+1}$ the Markov process where x moves to $R_\alpha x$ with probability $\frac{1}{m}$ has unique invariant measure (Haar). We shall use this observation to study the invariant measures for the process (1).

Let $\tilde{\mathbb{S}}^d$ and $\tilde{G}_{r,d}$ denote m disjoint copies of $\mathbb{S}^d$ and $G_{r,d}$ respectively and let $\tilde{\mathbb{V}}_d$ be the space of vectorfields on $\tilde{\mathbb{S}}^d$. Thus the point in $\tilde{G}_{r,d}$ is a triple (x, E, α) where α is an index of the sphere, $x \in \mathbb{S}^d$ and E is an r dimensional plane in $T_x\mathbb{S}^d$. $\tilde{\mathbb{V}}_d$ is the space of m-tuples of vectorfields on $\mathbb{S}^d$. In particular we can regard $\{X_\alpha\}$ as one vectorfield on $\tilde{\mathbb{S}}^d$ given by $X(x,\alpha) = X_\alpha(x)$. On $\tilde{G}_{r,d}$ we consider a Markov process

$$((x,E), w_1) \to (\hat{F}_n(w)(x,E), w_{n+1}) \tag{9}$$

where $\hat{F}(x,E) = (F(x), dF(x)E)$. In other words if our process is at state (x,E,α) then we apply f_α to x, df_α to E and choose the next symbol randomly from the uniform distribution on $\{1\dots m\}$. Observe that (9) and (1) are essentially the same processes but (9) is more convenient for bookkeeping if we want to consider observables which depend not only on x but also on the diffeomorphism we are applying each time.

Let $\tilde{\mathcal{M}}$ be the transition operator for the random rotations. That is

$$\left(\tilde{\mathcal{M}}A\right)(x,\alpha)=\frac{1}{m}\sum_{\beta=1}^{m}A(R_\alpha x,\beta)=\mathbb{E}_{(x,\alpha)}A(x_1,\alpha_1). \tag{10}$$

Then by induction

$$\left(\tilde{\mathcal{M}}^N A\right)(x,\alpha)=\frac{1}{m}\sum_{\beta=1}^{m}\left(\mathcal{M}^{N-1}A(\cdot,\beta)\right)(R_\alpha x)$$

where $\mathcal{M}$ acts by (2) where A is considered as a function of x with the second variable being fixed. Therefore the estimates of Sections 2.2 and 2.5 are valid for $\tilde{\mathcal{M}}$.

The following statement is proven in Section 4.

Proposition 2. *Given $\delta>0$ there exist constants C, k_5, k_6 and a bilinear form $\omega : C^{k_5}(\tilde{G}_{r,d})\times\tilde{\mathbb{V}}_d^{k_6}\to\mathbb{R}$ such that if μ is any invariant measure for the Markov process (9) then*

$$\left|\mu(A)-\int_{\tilde{G}_{r,d}}A(x,E)dxdE-\omega(A,X)\right|\leq C||A||_{k_5}\left(d_{k_6}(R,f)\right)^{2-\delta}.$$

Here $dxdE$ denotes the Haar measure on $\tilde{G}_{r,d}$ (the unique probability measure invariant under $\mathbb{SO}_{d+1}\times$ permutations).

Remark. *More information about the smoothness of invariant measures (in the context of deterministic dynamical systems) can be found in* [18, 27, 28]. *The time dependent case which is close to our setting is discussed in* [2].

3.2. Let $R^{-1},\tilde{\mathcal{L}}$ denote the operators on $\tilde{\mathbb{V}}^d$ given by

$$(R^{-1}X)(x,\alpha)=dR_\alpha^{-1}X(x,\alpha),\quad(\tilde{\mathcal{L}}X)(x,\alpha)=\frac{1}{m}\sum_{\beta=1}^{m}dR_\beta X_\beta(R_\beta^{-1}x).$$

The operators $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{M}}$ are adjoint in the following sense: if X and A are respectively a vector field and an observable on $\tilde{\mathbb{S}}^d$ we have

$$<X,\tilde{\mathcal{M}}A>=<\tilde{\mathcal{L}}X,A> \tag{11}$$

where $<Y,B>$ (Y and B being respectively vector field and observable on $\tilde{\mathbb{S}}^d$) denotes integration of $\partial_Y B$ on $\tilde{\mathbb{S}}^d$ with respect to Haar measure. (See Section 4 for a proof of (11).) Note that $\tilde{\mathcal{L}}$ preserves the space $\mathbb{V}^d$ of vectorfields which are the same on each copy of $\mathbb{S}^d$ and $\tilde{\mathcal{L}}(\tilde{\mathbb{V}}^d)=\mathbb{V}^d$, $\tilde{\mathcal{L}}|_{\mathbb{V}^d}=\mathcal{L}$. Thus the estimates of Sections 2.2 and 2.5 are valid for $\tilde{\mathcal{L}}$.

If $\mathcal{T}_0 X = 0$ let $Y = -\sum_{j=1}^{\infty} \tilde{\mathcal{L}}^j X$. Observe that Y does not depend on the second variable so it can be regarded as a vectorfield on $\mathbb{S}^d$. Let

$$\tilde{X} = X - Y + R^{-1} Y, \tag{12}$$

(see (8)). Then since $\tilde{\mathcal{L}}(R^{-1}Y) = Y$ we have

$$\tilde{\mathcal{L}}\tilde{X} = \tilde{\mathcal{L}}X - Y + \left(\sum_{j=2}^{\infty} \tilde{\mathcal{L}}^j X \right) = 0. \tag{13}$$

3.3. **Lyapunov exponents.** The following statement is proven in Section 5.

Proposition 3. *(a) There exist constants C, k_7 and quadratic form $q(r) : \tilde{\mathbb{V}}_d^{k_7} \to \mathbb{R}$ such that if μ is any invariant measure for $x \to F_n x$ then*

$$|\Lambda_r(\mu) - q(r)(X)| \leq C d_{k_7}^{3-\delta}(f, R),$$

(b) Given Y let $\tilde{X} = X - Y + R^{-1}Y$, then for all r we have

$$q(r)(\tilde{X}) = q(r)(X).$$

(c) If $\tilde{\mathcal{L}}(\mathcal{R}_0(X)) = 0$ then

$$|\Lambda_r(\mu) - (q_1(X) r + q_2(X) r(d-r))| \leq C d_{k_7}^{3-\delta}(f, R),$$

$$|\lambda_r(\mu) - (q_1(X) + q_2(X)(d - 2r + 1))| \leq C d_{k_7}^{3-\delta}(f, R),$$

where

$$q_1(X) = -\frac{1}{2dm} \sum_{\alpha} \int_{\mathbb{S}^d} (\mathrm{div} X_\alpha)^2 \, dx.$$

$$q_2(X) = \frac{1}{(d+2)(d-1)m} \sum_{\alpha} \int_{\mathbb{S}^d} \mathrm{Tr} \left[\frac{DX_\alpha + DX_\alpha^*}{2} - \frac{\mathrm{Tr} DX_\alpha}{d} \right]^2 dx.$$

(d) If $\tilde{\mathcal{L}}(\mathcal{R}_0(X)) = 0$ then

$$|\lambda_d| \geq \mathrm{Const} \sum_{\alpha} \int_{\mathbb{S}^d} < \Delta X_\alpha, X_\alpha > dx - \mathrm{Const} ||X||_{k_7}^{3-\delta}.$$

Remark. *The change of Lyapunov exponents for small perturbations of elliptic systems were studied in* [11, 5] *etc.*

3.4. **Construction of the conjugation.** To prove our main result we assume that for each $\epsilon > 0$ there exists a measure μ_ϵ such that

$$\lambda_d(\mu_\epsilon) > -\epsilon \tag{14}$$

and show that f_α are simultaneously conjugated to rotations. The conjugation will be defined inductively. Let $f_{\alpha,0} = f_\alpha$, $\phi_0 = \mathrm{id}$, $R_{\alpha,0} = R_\alpha$. Assume that we have already constructed ϕ_p such that $f_{\alpha,p} = \phi_p f_\alpha \phi_p^{-1}$ satisfy $d_s(f_{\alpha,p}, R_{\alpha,p}) \leq \varepsilon_{p,s}$ for some rotations close to $R_{\alpha,0}$. For $N > 1$ big enough let

$$\lambda_p = N^{(1+\tau)^p} \tag{15}$$

where $0 < \tau < 1$. Let $\tilde{\mathcal{K}} = -\sum_{k=1}^{\infty} \tilde{\mathcal{L}}^k$. Define

$$X_{\alpha,p} = \exp^{-1}[(R_{\alpha,p}^{-1} f_{\alpha,p})],$$

$$\hat{Y}_p = \tilde{\mathcal{K}}(\mathcal{R}_0 X_p),$$

$$Y_p = \mathcal{T}_{\lambda_p}(\hat{Y}_p) = \tilde{\mathcal{K}}(\mathcal{T}_{\lambda_p} \mathcal{R}_0 X_p) = -\sum_{k=1}^{\infty} \tilde{\mathcal{L}}^k (\mathcal{T}_{\lambda_p} \mathcal{R}_0 X_p)$$

$$\phi_{p+1} = \psi_{Y_p} \phi_p.$$

Then $f_{p+1} = R_p \exp(Z_p)$ with

$$Z_p = X_p - Y_p + R_p^{-1} Y_p + O_2(X_p, Y_p) \tag{16}$$

$$= X_p - \mathcal{T}_{\lambda_p} X_p + \left(\mathcal{T}_{\lambda_p} X_p - Y_p + R_p^{-1} Y_p \right) + O_2(X_p, Y_p) \tag{17}$$

where $O_2(X_p, Y_p)$ denotes a quadratic expression in (X_p, Y_p) (in the sense of 2.3 (III)). Let us set

$$\hat{X}_p = X_p - \hat{Y}_p + R_p^{-1} \hat{Y}_p,$$

$$X_p^* = \mathcal{T}_{\lambda_p}(\hat{X}_p) = \mathcal{T}_{\lambda_p} X_p - Y_p + R_p^{-1} Y_p.$$

By construction $\tilde{\mathcal{L}}(\mathcal{R}_0 \hat{X}_p) = 0$ and since we are assuming (14) Proposition 3(d) enables us to conclude that

$$||X_p^* - \mathcal{T}_0 X_p^*||_{H^1} \leq ||\hat{X}_p - \mathcal{T}_0 \hat{X}_p||_{H^1} = O(\varepsilon_{p,k_7}^{(1+\sigma)}) \tag{18}$$

($1 + \sigma = ((3/2) - (\delta/2))$ and by Sections 2.3, 2.4 (remember that $X_p^* = T_{\lambda_p} X_p^*$!)

$$||X_p^* - \mathcal{T}_0 X_p^*||_s \leq C_s \lambda_p^{s/2 + k_3} \varepsilon_{k_7,p}^{(1+\sigma)}. \tag{19}$$

Observe also that $\Delta(\mathcal{T}_0 X_p^*) = 0$ so $\psi_{\mathcal{T}_0 X_p^*}$ is a rotation. Let

$$R_{\alpha,(p+1)} = \psi_{(\mathcal{T}_0 X_p^*)_\alpha} R_{\alpha,p}.$$

We can then write

$$f_{p+1} = R_{p+1} \exp(X_{p+1}),$$

with

$$X_{p+1} = Z_p - \mathcal{T}_0 X_p^* + O_2(Z_p, \mathcal{T}_0 X_p)$$
$$= (X_p^* - \mathcal{T}_0 X_p^*) + W_p + (\hat{X}_p - X_p^*)$$

where Z_p is given by (16) and W_p is quadratic in X_p. Combining (19) with the estimates of Sections 2.3,2.4 we get for any integers $\bar{s}, s$, such that $\bar{s} \geq s$

(20)

$$\varepsilon_{p+1,s} \leq C_{s,\bar{s}} \left(1 + \lambda_p^a \varepsilon_{p,k_8}\right)^{s+1} \left(\lambda_p^{a+s/2} \varepsilon_{p,k_8}^{1+\sigma} + \lambda_p^a \varepsilon_{p,s} \varepsilon_{p,0} + \lambda_p^{a-(\bar{s}-s)/2} \varepsilon_{p,\bar{s}}\right)$$

where a, k_8, σ are some positive constants. Namely, in the second factor in the RHS of (20) the first term comes from (19), the second term estimates W_p and the third term comes estimates $\hat{X}_p - X_p^* = \mathcal{R}_{\lambda_p} \hat{X}_p$. The first factor comes from 2.3(III). (Obviously (20) remains valid if all term in the RHS are multiplied by

$$\left(1 + \lambda_p^a \varepsilon_{p,k_8}\right)^{s+1}$$

not only W_p-part.) We shall choose $0 < \tau < \sigma$.

3.5. **Convergence of iterations.** The following statement is proven in Section 6.

Proposition 4. *There exists s_0, such that if $\max_\alpha d_{s_0}(f_\alpha, R_\alpha)$ is small enough then for any $m > 0$, $s \geq 0$ there exists a constant $C_{s,m}$ such that for any p*

$$\varepsilon_{p,s} \leq C_{s,m} \lambda_p^{-m}$$

(in that case we write $\varepsilon_{p,s} = O(\lambda_p^{-\infty})$)

Proposition 4 implies that ϕ_p C^∞-converge to a limit ϕ_∞ and that $R_{\alpha,\infty} = \phi_\infty f_\alpha \phi_\infty^{-1}$ are rotations. This proves Theorem 1.

Remark. *The iteration procedure we have just decribed is reminiscent of KAM theory. The first application of KAM techniques to hyperbolic dynamics is* [6]. *The idea to use it to establish stable ergodicity is due to* [17]. *Some further applications of KAM to hyperbolic dynamics can be found in* [7]. *In our paper the unavoidable use of a Nash-Moser type iteration procedure (due to the fact that the operator $\mathcal{L}$ displays loss of derivatives properties) has to be coupled at each step with perturbative computations of invariant measures and Lyapunov exponents. If at each step these perturbative formulas do not give valuable information on Lyapunov exponents, then a KAM step can be performed and* *we eventually get a conjugacy result (see* [21] *for an analogous situation where a KAM scheme and a renormalization scheme are run in parallel).*

4. Invariant measures.

Proof of Proposition 2. Our proof is similar to [9], however we are able to get better estimates since we deal with a more explicit situation.

We note that $G_{r,d} = \mathbb{SO}_{d+1}/(\mathbb{SO}_r \times \mathbb{SO}_{d-r})$, so it is enough to provide the asymptotic expansion for the invariant measures for the process on $\widetilde{\mathbb{SO}}_{d+1}$ given by $g_{\alpha+1} = \mathbf{f}_\alpha g_\alpha$ where $\mathbf{f}_\alpha$ are close to R_α. Observe that f_α can be lifted to $\mathbb{SO}_{d+1}$ because $\mathbb{SO}_{d+1}$ is a frame bundle of $\mathbb{S}^d$. The procedure to get the lift is the following. Define for any base $\mathcal{F} = (e_1, \dots, e_{d+1})$ of $\mathbb{R}^{d+1}$ the orthonormal base $\mathrm{Orth}(\mathcal{F}) = (e'_1, \cdots, e'_{d+1})$ obtained by applying the Gram-Schmidt orthonormalization procedure to $\mathcal{F}$ in such a way that $\mathrm{Span}(e'_1, \dots, e'_i) = \mathrm{Span}(e_1, \dots, e_i)$ $(1 \le i \le d+1)$. Next, any orthonormal base $e_1, e_2 \dots e_{d+1}$ can be considered as a orthonormal base in the tangent space $T_{e_1}\mathbb{S}^d$. Let now $\mathcal{F}_0 = (e_1, \dots, e_{d+1})$ be a fixed orthonormal base of $\mathbb{R}^{d+1}$. For $Q \in \mathbb{SO}_{d+1}$ let $\mathcal{F}(Q) = Q(\mathcal{F}_0)$. Then $Q \to \mathcal{F}(Q)$ is an diffeomorphism between $\mathbb{SO}_{d+1}$ and the space of frames in $\mathbb{R}^{d+1}$. If $f : \mathbb{S}^d \to \mathbb{S}^d$ is a diffeomorphism we can lift it to a diffeomorphism $\mathbf{f}$ of $\mathbb{SO}_{d+1}$ as follows. $\mathbf{f}(Q) = P$ iff $\mathcal{F}(P) = \mathrm{Orth}(\hat{f}\mathcal{F}(Q))$, where

$$\hat{f}((\bar{e}_1, \bar{e}_2, \dots \bar{e}_{d+1})) = (f(\bar{e}_1), df_{\bar{e}_1}\bar{e}_2, \dots df_{\bar{e}_1}\bar{e}_{d+1}).$$

It is clear that the lift of a rotation of the sphere is this rotation and that composition of maps commute with the lift procedure. We now make the following remark. If f is a diffeomorphism of the sphere $\mathbb{S}^d$ there is canonically defined a diffeomorphism $\tilde{f}$ on the Grassmann bundle $G_{r,d}$ such that $\pi_1 \circ \tilde{f} = f \circ \pi_1$ (where π_1 is the canonical projection from $G_{r,d}$ to $\mathbb{S}^d$). On the other hand the above procedure defines a diffeomorphism $\mathbf{f}$ on $\mathbb{SO}_{d+1}$ such that $f \circ \pi_2 = \pi_2 \circ \mathbf{f}$ (where π_2 is the canonical projection from $\mathbb{SO}_{d+1}$ to $\mathbb{S}^d$). From construction it is clear that $\pi_3 \circ \mathbf{f} = \tilde{f} \circ \pi_3$ where π_3 is the canonical projection $\mathbb{SO}_{d+1} \to G_{r,d}$.

We denote points of $\mathbb{SO}_{d+1}$ by z. Thus $z = (g, \alpha)$ where $g \in \mathbb{SO}_{d+1}$, $\alpha \in \{1 \dots m\}$. dz denotes the Haar measure on $\widetilde{\mathbb{SO}}_{d+1}$. Let $X_\alpha = \exp^{-1}(R_\alpha^{-1}\mathbf{f}_\alpha)$. We shall write $\varepsilon = \max_\alpha d_r(\mathbf{f}_\alpha, R_\alpha)$ for some sufficiently large r.

Let

$$N = (1/\varepsilon)^{\delta/3}$$

where δ is as in Proposition 2. Let $s_1 = 3/(\gamma\delta)$ where γ is the constant from (7). Thus $N^{-\gamma s_1} = \varepsilon^3$.

Let $\tilde{\mathcal{M}}$ be the transition operator for random rotations acting on $\mathbb{SO}_{d+1}$ (see (10)). For each realization $\{z_n^{(\varepsilon)}\} = \{(g_n^{(\varepsilon)}, \alpha_n)\}$ of our Markov process starting from z let $\{z_n^{(0)}\} = \{(g_n^{(0)}, \alpha_n)\}$ be the corresponding unperturbed realization. Thus $z_0^{(\varepsilon)} = z_0^{(0)} = z$, $g_{n+1}^{(0)} = R_{\alpha_n} g_n^{(0)}$. It follows by induction that for all $n \leq N$

$$g_n^{(\varepsilon)} = \exp_{g_n^{(0)}} Y_n(\varepsilon), \tag{21}$$

where

$$Y_n(\varepsilon) = \sum_{p=0}^{n-1} R_{p,n} X(g_p^{(0)}) + O(\varepsilon^2 n^3) \tag{22}$$

and $R_{p,n} = R_{\alpha_{n-1}} \dots R_{\alpha_{p+1}} R_{\alpha_p}$. Indeed it is easy to see by induction that

$$||Y_j(\varepsilon)|| \leq \text{Const}\varepsilon j \tag{23}$$

and (23) implies that the error term in (22) is less than Const $\sum_{j=1}^{n} (\varepsilon j)^2$. Let $\tilde{\mathcal{M}}_\varepsilon$ be the transition operator for the perturbed process

$$\left(\tilde{\mathcal{M}}_\varepsilon A\right)(x, \alpha) = \frac{1}{m} \sum_{\beta=1}^{m} A(\mathbf{f}_\alpha x, \beta).$$

Then (21) implies that

$$\left(\tilde{\mathcal{M}}_\varepsilon^N A\right)(z) = \left(\tilde{\mathcal{M}}^N A\right)(z) + \mathbb{E}_z(\partial_Y A(z_n^{(0)})) + O\left(\varepsilon^2 N^3 ||A||_2\right). \tag{24}$$

By (7) the first term in (24) is

$$\tilde{\mathcal{M}}^N A = \int_{\widetilde{\mathbb{SO}}_{d+1}} A(z) dz + O(||A||_{s_1+b} \varepsilon^3).$$

Using (22) and (10) the second term in (24) can be rewritten as

$$\sum_{p=0}^{N-1} \tilde{\mathcal{M}}^p \left(\partial_X \left(\tilde{\mathcal{M}}^{N-p} A\right)\right).$$

Calling $q = N - p$ we see that the second term in (24) is $\sum_{q=1}^{N} \sigma_q$ where $\sigma_q = \tilde{\mathcal{M}}^{N-q} \left(\partial_X \left(\tilde{\mathcal{M}}^q A\right)\right)$. For σ_q we have two estimates.

(I) By (7)

$$\sigma_q = \int_{\widetilde{\mathbb{SO}}_{d+1}} \left(\partial_X \left(\tilde{\mathcal{M}}^q A\right)\right) dz + O\left(||\partial_X \left(\tilde{\mathcal{M}}^q A\right)||_{s_1+b} (N-q)^{-\gamma s_1}\right)$$

where b is a constant from (7). We shall use this estimate for $q \leq N/2$, then the second part is $O(||X||_{s_1+b} ||A||_{s_1+b+1} \varepsilon^3)$.

If $q > N/2$ we use

$$(II) \quad \sigma_q = O(||\partial_X \left(\tilde{\mathcal{M}}^q A\right) ||_0) = O\left(\frac{||X||_0||A||_{s_1+b+1}}{q^{\gamma s_1}}\right) = O(||A||_{s_1+b+1}\varepsilon^3)$$

because $\tilde{\mathcal{M}}^q A = \left(\int A(z)dz\right) 1 + \kappa_q$ where $||\kappa_q||_1 = O(||A||_{s_1+b+1}/q^{\gamma s_1})$ and $\partial_X 1 = 0$. Observe that (II) also implies that

$$\sum_{q=1}^{N/2} \int_{\widetilde{\mathbb{SO}}_{d+1}} \left(\partial_X \left(\tilde{\mathcal{M}}^q A\right)\right) dz = \sum_{q=1}^{\infty} \int_{\widetilde{\mathbb{SO}}_{d+1}} \left(\partial_X \left(\tilde{\mathcal{M}}^q A\right)\right) dz + O(\epsilon^3||A||_{s_1+b+1}).$$

Combining these bounds we get

$$\tilde{\mathcal{M}}_\varepsilon A = \int_{\widetilde{\mathbb{SO}}_{d+1}} A(z)dz + \sum_{q=1}^{\infty} \int_{\widetilde{\mathbb{SO}}_{d+1}} \left(\partial_X \left(\tilde{\mathcal{M}}^q A\right)\right) dz + O(\epsilon^2 N^3||A||_{s_1+b+1}).$$

If μ is an invariant measure then $\mu(\tilde{\mathcal{M}}_\varepsilon^N A) = \mu(A)$ and the result follows. □

Corollary 3. *If $\tilde{\mathcal{L}}(\mathcal{R}_0 X) = 0$ then $\omega(A, X) = 0$.*

Proof. Again it is enough to prove this result for the perturbed process on $\widetilde{\mathbb{SO}}_{d+1}$. We have an explicit formula

$$\omega(A, X) = \sum_{q=1}^{\infty} \int_{\widetilde{\mathbb{SO}}_{d+1}} \partial_X \left(\tilde{\mathcal{M}}^q A\right) dz.$$

Next if B is any function on $\widetilde{\mathbb{SO}}_{d+1}$ and Y is any vectorfield then we have the identities

$$\int_{\widetilde{\mathbb{SO}}_{d+1}} \partial_Y B dz = \int_{\widetilde{\mathbb{SO}}_{d+1}} \partial_{\mathcal{R}_0 Y}(\mathcal{R}_0 B) dz$$

(since $\mathcal{T}_0$ and $\mathcal{R}_0$ are orthogonal and $\mathcal{T}_0 B$ is piecewise constant) and

$$\int_{\widetilde{\mathbb{SO}}_{d+1}} \partial_Y (\tilde{\mathcal{M}} B) dz = \frac{1}{m} \int_{\mathbb{SO}_{d+1}} \sum_{\alpha=1}^{m} \partial_{Y(g,\alpha)} (\tilde{\mathcal{M}} B)(g, \alpha) dg$$

$$= \frac{1}{m^2} \int_{\mathbb{SO}_{d+1}} \sum_{\alpha=1}^{m} \sum_{\beta=1}^{m} \partial_{Y(g,\alpha)} B(R_\alpha g, \beta) dg$$

$$= \frac{1}{m^2} \int_{\mathbb{SO}_{d+1}} \sum_{\alpha=1}^{m} \sum_{\beta=1}^{m} \partial_{dR_\alpha Y(R_\alpha^{-1} h, \alpha)} B(h, \beta) dh$$

$$= \int_{\widetilde{\mathbb{SO}}_{d+1}} \partial_{\tilde{\mathcal{L}} Y} B dz.$$

Thus

$$\int_{\widetilde{\mathbb{SO}_{d+1}}} \partial_X \left(\tilde{\mathcal{M}}^q A\right) dz = \int_{\widetilde{\mathbb{SO}_{d+1}}} \partial_{\mathcal{R}_0 X} \left(\tilde{\mathcal{M}}^q \mathcal{R}_0 A\right) dz = \int_{\widetilde{\mathbb{SO}_{d+1}}} \partial_{\tilde{\mathcal{L}}^q(\mathcal{R}_0 X)} \mathcal{R}_0 A dz = 0.$$

□

5. Lyapunov exponents.

The proof of Proposition 3 relies on the following elementary formula (see Appendix A).

Lemma 1. *Let* $L(\varepsilon) = 1 + \varepsilon L_1 + \varepsilon^2 L_2 + O(\varepsilon^3)$. *Denote*

$$\Lambda_r = \int_{\mathbb{G}_{r,d}} \ln \det(L(\varepsilon)|E) dE, \quad \lambda_r = \Lambda_r - \Lambda_{r-1}.$$

Then

$$\Lambda_r = \varepsilon \frac{r}{d} \mathrm{Tr} L_1 + \varepsilon^2 \left[\frac{r}{d}\mathrm{Tr}L_2 - \frac{r}{2d}\mathrm{Tr}L_1^2 + \frac{r(d-r)}{(d+2)(d-1)}\mathrm{Tr}K^2\right] + O\left(\varepsilon^3\right).$$

$$\lambda_r = \varepsilon \frac{1}{d} \mathrm{Tr} L_1 + \varepsilon^2 \left[\frac{1}{d}\mathrm{Tr}L_2 - \frac{\mathrm{Tr}L_1^2}{2d} + \frac{d-2r+1}{(d+2)(d-1)}\mathrm{Tr}K^2\right] + O\left(\varepsilon^3\right).$$

where

$$K = \frac{L_1 + L_1^*}{2} - \frac{\mathrm{Tr}L_1}{d}.$$

Proof of Proposition 3. (a) Write

$$R_\alpha^{-1} df_\alpha = 1 + a_\alpha + b_\alpha + \dots$$

where a_α are linear in X_α and b_α are quadratic. It will be convinient to treat a and b as defined on $\tilde{G}_{r,d}$. Now

$$\Lambda_r(\mu) = \int_{\tilde{G}_{r,d}} \ln \det(df_\alpha|E)(x) d\bar{\mu}(x, E)$$

where $\bar{\mu}$ is an invariant measure on $\tilde{G}_{r,d}$ projecting to μ (see [19], page 94). By Proposition 2

(25)

$$\Lambda_r(\mu) = \frac{1}{m}\sum_\alpha \iint \ln\det(df_\alpha|E)(x) dx dE + \omega\left(\mathrm{Tr}(a|E), X\right) + O\left(||X||_{k_7}^{3-\delta}\right).$$

Now

$$\frac{1}{m}\sum_\alpha \iint \ln\det(df_\alpha|E)(x) dx dE =$$

$$\frac{1}{m}\int \ln\det\left((1+a_\alpha+b_\alpha)|E\right)(x)dxdE+O\left(||X||_{k_7}^{3-\delta}\right).$$

Next for fixed x Lemma 2 gives

$$\int \ln\det\left((1+a_\alpha+b_\alpha)|E\right)(x)dE= \tag{26}$$

$$\frac{r}{d}\left[\mathrm{Tr}a_\alpha(x)+\mathrm{Tr}b_\alpha(x)\right]-\frac{r}{2d}\mathrm{Tr}(a_\alpha^2)(x)+\frac{r(d-r)}{(d+2)(d-1)}\mathrm{Tr}(c_\alpha^2)+O\left(||X||_{k_7}^{3}\right)$$

where

$$c_\alpha=\frac{a_\alpha+a_\alpha^*}{2}-\frac{\mathrm{Tr}a_\alpha}{d}.$$

On the other hand

$$\det(df_\alpha)(x)=1+\mathrm{Tr}a_\alpha+\mathrm{Tr}b_\alpha+\frac{(\mathrm{Tr}a_\alpha)^2}{2}-\frac{\mathrm{Tr}a_\alpha^2}{2}+O\left(||X||_{k_7}^{3}\right).$$

Since

$$\int_{\mathbb{S}^d}\det(df_\alpha)(x)dx=1$$

we get

$$\frac{r}{d}\int_{\mathbb{S}^d}\left(\mathrm{Tr}a_\alpha+\mathrm{Tr}b_\alpha-\frac{\mathrm{Tr}a_\alpha^2}{2}\right)dx=-\frac{r}{2d}\int_{\mathbb{S}^d}(\mathrm{Tr}a_\alpha)^2dx \tag{27}$$

Combining (25), (26) and (27) we get (a).

(b) is clear since Lyapunov exponents are independent of the choice of coordinates.

(c) follows from (25), (26), (27) and Corollary 3.

To get (d) we rewrite

$$\lambda_d=$$

$$-\frac{1}{(d+2)m}\left(\sum_\alpha\frac{1}{2}\int_{\mathbb{S}^d}(\mathrm{div}X_\alpha)^2dx+\sum_\alpha\int_{\mathbb{S}^d}\mathrm{Tr}\left(\frac{DX_\alpha+DX_\alpha^*}{2}\right)^2dx\right)+O\left(||X||_{k_7}^{3-\delta}\right).$$

Thus

$$|\lambda_d|\geq \mathrm{Const}\sum_\alpha\int_{\mathbb{S}^d}\left(\mathrm{div}X_\alpha\right)^2dx-O\left(||X||_{k_7}^{3-\delta}\right)\geq$$

$$\mathrm{Const}\sum_\alpha\int_{\mathbb{S}^d}<\Delta X_\alpha,X_\alpha>dx-O\left(||X||_{k_7}^{3-\delta}\right).\quad\square$$

6. Convergence of iterations.

Proof of Proposition 4. By shifting the index s by k_8 and changing the value of a we can simplify (20) as follows

(28)
$$\varepsilon_{p+1,s} \leq C_{s,\bar{s}} \left(1 + \lambda_p^a \varepsilon_{p,0}\right)^s \left(\lambda_p^{a+s/2} \varepsilon_{p,0}^{1+\sigma} + \lambda_p^a \varepsilon_{p,s} \varepsilon_{p,0} + \lambda_p^{a-(\bar{s}-s)/2} \varepsilon_{p,\bar{s}}\right)$$

So it is enough to show that (28) implies that $\varepsilon_{p,s} = O(\lambda_p^{-\infty})$.

6.1. We first prove that if N in (15) is big enough then there exist positive real numbers $\gamma_0 > a, s_0, b$ such that for any $p \geq 0$

$$\varepsilon_{p,0} \leq \lambda_p^{-\gamma_0} \tag{29}$$

$$\varepsilon_{p,s_0} \leq \lambda_p^b \tag{30}$$

provided these estimates are true for $p = 0$. In view of (20) where we make $s = 0, \bar{s} = s_0$ and $s = s_0, \bar{s} = s_0$ we just have to check that

(31)
$$3 \times 2^{s_0} C_{s_0} \lambda_p^a \lambda_p^{-(1+\sigma)\gamma_0} \leq \lambda_p^{-\gamma_0(1+\tau)}, \qquad 3 \times 2^{s_0} C_{s_0} \lambda_p^{a+s_0/2} \lambda_p^{-(1+\sigma)\gamma_0} \leq \lambda_p^{b(1+\tau)},$$

(32)
$$3 \times 2^{s_0} C_{s_0} \lambda_p^a \lambda_p^{-2\gamma_0} \leq \lambda_p^{-\gamma_0(1+\tau)}, \qquad 3 \times 2^{s_0} C_{s_0} \lambda_p^a \lambda_p^{-\gamma_0} \lambda_p^b \leq \lambda_p^{b(1+\tau)},$$

(33)
$$3 \times 2^{s_0} C_{s_0} \lambda_p^{a-s_0/2} \lambda_p^b \leq \lambda_p^{-\gamma_0(1+\tau)}, \qquad 3 \times 2^{s_0} C_{s_0} \lambda_p^a \lambda_p^b \leq \lambda_p^{b(1+\tau)},$$

that is (provided N is big enough)

$$a < \gamma_0(\sigma - \tau), \qquad a + b < -\gamma_0(1+\tau) + s_0/2$$
$$a < \tau b, \qquad s_0/2 - (1+\sigma)\gamma_0 < b(1+\tau) - a$$

(We have four conditions here because the inequalities in the right column of (32) and (33) follow from the others.) If we take $b > a/\tau$, $s_0/2 = \gamma_0(1+\tau')$ with $\tau < \tau' < \sigma$, and if γ_0 is big enough (for further purpose we impose $(\tau' - \tau)\gamma_0 > (b+a)$; see subsection 6.3), this inequalities are satisfied. Then take N big enough so that the estimates (29) are satisfied for $p = 0$.

6.2. Next we show two lemmas.

Lemma 2. *If $a/(\sigma-\tau) < \gamma < c/(1+\tau)$ and if u_p is sequence of positive real numbers converging to zero and satisfying*

$$u_{p+1} \leq C(\lambda_p^a u_p^{1+\sigma} + \lambda_p^{-c})$$

then $u_p = O(\lambda_p^{-\gamma})$

Proof. We can assume $0 < u_p < 1$. Observe that

$$2C\lambda_p^a \lambda_p^{-\gamma(1+\sigma)} \leq \lambda_p^{-\gamma(1+\tau)}, \qquad 2C\lambda_p^{-c} \leq \lambda_p^{-\gamma(1+\tau)}$$

if p is big enough. Now

i) either for any p the inequality $u_p > \lambda_p^{-\gamma}$ is true and then

$$\lambda_p^{-c} < u_p^{c/\gamma} < u_p^{1+\tau'}$$

for some $\tau < \tau' < \sigma$. Hence

$$u_{p+1} \leq 2C\lambda_p^a u_p^{1+\tau'}$$

and since $\lim u_p = 0$ this implies $u_p = O(\lambda_p^{-\infty})$.

ii) or there exists $p_k \to \infty$ such that $u_{p_k} \leq \lambda_{p_k}^{-\gamma}$ and then the induction can be initiated. □

The next lemma is similar to but easier than Lemma 2 so we leave the proof to the reader.

Lemma 3. *Let sequence $u_p \geq 0$ satisfy*

$$u_{p+1} \leq C(\lambda_p^{-\gamma_1} u_p + \lambda_p^{-\gamma_2})$$

for some $\gamma_2 > 0$, $\gamma_1 \in \mathbb{R}$.

(a)If $\gamma_1 < 0$ then $u_p = O(\lambda_p^b)$ for any $b > |\gamma_1|/\tau$.

(b) If $\gamma_1 > 0$ then $u_p = O(\lambda_p^{-b})$ for any $b < \min(|\gamma_1|/\tau, \gamma_2/(1+\tau))$.

6.3. Let us choose τ'' such that $\tau' < \tau'' < \sigma$. We now prove by induction on k that the sequences γ_k, s_k such that

$$s_k = \frac{1+\tau''}{1+\tau'} s_{k-1}, \qquad \gamma_k = \frac{1}{1+\tau'}(s_k/2)$$

satisfy the following property: for any $p \in \mathbb{N}$

(P_k) $\quad \varepsilon_{p,0} = O(\lambda_p^{-\gamma_k}), \qquad \varepsilon_{p,s_k} = O(\lambda_p^b).$

Observe that

$$\varepsilon_{p+1,s_k} \leq C_{s_k}(\lambda_p^{(s_k/2)+a} \varepsilon_{p,0}^{1+\sigma} + \lambda_p^a \varepsilon_{p,0} \varepsilon_{p,s_k} + \lambda_p^a \varepsilon_{p,s_k})$$

and since (P_{k-1}) holds $\lambda_p^{(s_k/2)+a} \varepsilon_{p,0} = O(\lambda_p^{(s_k/2)+a-(1+\sigma)\gamma_{k-1}})$ with

$$(s_k/2) + a - (1+\sigma)\gamma_{k-1} = (\tau'' - \sigma)\gamma_{k-1} + a < 0.$$

Lemma 3(a) gives $\varepsilon_{p,s_k} = O(\lambda_p^b)$ with $b > a/\tau$.

Also

$$\varepsilon_{p+1,0} \leq C_{s_k}(\lambda_p^a \varepsilon_{p,0}^{1+\sigma} + \lambda_p^a \varepsilon_{p,0}^2 + \lambda^{a-(s_k/2)} \varepsilon_{p,s_k})$$

and

$$\frac{a}{\sigma - \tau} < \gamma_k < \frac{(s_k/2) - a - b}{1+\tau}$$

since $\gamma_k(1+\tau) < (1+\tau')\gamma_k - (b+a)$. Lemma 2 then applies.

6.4. Since $\gamma_k \to \infty$, we have proven that $\varepsilon_{p,0} = O(\lambda_p^{-\infty})$ and since for any s

$$\varepsilon_{p+1,s} \leq C_s(\lambda_p^{(s_k/2)+a}\varepsilon_{p,0}^{1+\sigma} + \lambda_p^a \varepsilon_{p,0}\varepsilon_{p,s} + \lambda_p^{a-[(s_k-s)/2]}\varepsilon_{p,s_k})$$

the fact that $s_k \to \infty$ and Lemma 3(b) imply that $\varepsilon_{p,s} = O(\lambda_p^{-\infty})$ for any $s \in \mathbb{N}$.

7. Ratio of the exponents.

Proof of Theorem 2. Choose a large constant K. If $|q(d)(X)| > K d_{k_7}^{3-\delta}(f,R)$ then the result follows from Proposition 3. In general consider ϕ_p and $f_{\alpha,p}$ constructed in Section 3. Then either for some p

$$|q(d)(X_p)| > K d_{k_7}^{3-\delta}(f_p, R_p) \tag{34}$$

and then the results holds by Proposition 3 applied to $\{f_{\alpha,p}\}$ or (18) holds for all p. Hence $\{f_\alpha\}$ are conjugated to rotations by the estimates of Section 3.4. □

8. Invariant manifolds.

Here we recall some facts about stable and unstable manifolds of random transformations. More detailed information can be found in [23], Chapter III (see also [3]). Given an infinite word w and $x \in \mathbb{S}^d$ let

$$W^s(x,w) = \{y : d(F_n(w)x, F_n(w)y) \to 0 \text{ exponentially fast, } n \to \infty\}.$$

$$W^u(x,w) = \{y : d(F_{-n}(w)x, F_{-n}(w)y) \to 0 \text{ exponentially fast, } n \to \infty\}.$$

Then for almost all x, w $W^s(x,w)$ and $W^u(x,w)$ are C^∞ manifolds. We endow $W^*(x,w)$ with induced Riemannian distance. Let $r(x,w)$ denote the injectivity radius of $W^s(x,w)$ and let $W^s_l(x,w)$ denote the l-ball in $W^s(x,w)$.

In our analysis we shall use the absolute continuity of W^s. The absolute continuity has three manifestations.

(AC1) For almost all w the following holds. Let $\Omega \subset \mathbb{S}^d$ be a set such that for almost all x the leafwise measure $\text{mes}(W^s(x,w) \bigcap \Omega) = 0$ then $\text{Leb}(\Omega) = 0$.

(AC2) Conversely, for almost all w the following holds. Let V be a submanifold of dimension $d - [d/2]$ ($[\dots]$ denotes the integer part). Let

$$K = \{x \in V : \quad W^s(x,w) \text{ is transversal to } V\}.$$

Let $\Omega \subset \mathbb{S}^d$ be a set such that there is a positive measure subset $\tilde{K} \subset K$ such that for $x \in \tilde{K}$ the leafwise measure $\text{mes}(W^s(x,w) \bigcap \Omega) > 0$ then $\text{Leb}(\Omega) > 0$.

(AC3) For almost all w the following holds. Let V_1 and V_2 be submanifolds of dimension $d-[d/2]$. Choose a number l and let

$$K_1 = \{x \in V_1 : W^s(x,w) \text{ is transversal to } V_1, \quad \text{Card}(W^s_l(x,w) \bigcap V_2) = 1$$

and this intersection is transversal$\}$.

Let $p : K_1 \to V_2$ be the holonomy map along the stable leaves and let $K_2 = p(K_1)$. Then p is absolutely continuous in the sense that it sends measure zero sets to measure zero sets.

Given a pair of numbers l, κ we say that the pair x, w is (l,κ)–standard if $r(x,w) > l$ and the sectional curvatures of $W^s_l(x,w)$ have absolute value at most κ.

Suppose now what f_α are volume preserving and d is even so that all Lyapunov exponents are non-zero. Let $\rho = \min_\mu \min_j |\lambda_j(\mu)|$. Given C, ϵ denote by $\Lambda_{C,\epsilon}$ the Pesin set

$$\Lambda_{C,\epsilon} = \{(x,\omega) \quad ||dF_{k,j+k}|E_s|| \leq Ce^{\epsilon k - \rho j} \tag{35}$$

$$||dF_{k,k-j}|E_u|| \leq Ce^{\epsilon k - \rho j} \tag{36}$$

$$\angle(E_s(F_k(x)), E_u(F_k(x))) \leq \left(Ce^{\epsilon k}\right)^{-1}\}. \tag{37}$$

Proposition 5. *(See* [3]*) Given C, ϵ there exist l, κ such that all points in $\Lambda_{C,\epsilon}$ are (l,κ)-standard.*

Finally we need the following estimates on the size of stable manifolds in case all exponents are negative.

Proposition 6. *Let g_j be a sequence of diffeomorphisms of a compact manifold N uniformly bounded in C^2, $G_j = g_j \circ \cdots \circ g_1$. Then given ρ there exists a constant K such that if $v \in N$ is such that*

$$||dG_j(v)|| \leq Ce^{-\rho j}$$

then

$$r(W^s(v)) \geq \frac{1}{CK}.$$

The proof of this proposition is very similar to the proof of Lemma 2.7 of [1] and we leave it to the reader.

9. Continuous conjugation.

Proof of Corollary 1. If $\{f_\alpha\}$ are not C^∞ conjugated to rotations then by Theorem 1 $\lambda_d < 0$ for all invariant measures. In particular, there exist x, w such that $W^s(x,w) \neq \{x\}$. Hence $\{f_\alpha\}$ can not be C^0 conjugated to rotations. □

10. Stable ergodicity.

10.1. In this section we prove Corollary 2. Let $\{f_\alpha\}$ be close to $\{R_\alpha\}$. Since the ergodicity is clear in case $\{f_\alpha\}$ can be simultaneously linearized we assume below that Markov process (1) has non-zero exponents.

10.2. **Large deviations.** In this section we assume that f_α preserve volume. The following result is proved in Section 11. Note that the fact that d is even is not used until the part (c) of Corollary 4.

Lemma 4. *Fix $\epsilon > 0$. Then if $\lambda_i \neq 0$ and $\{f_\alpha\}$ are sufficiently close to $\{R_\alpha\}$ then there exist constants $C, \theta < 1$ such that for any $x \in \mathbb{S}^d$ for any r for any r-dimensional $E \subset T_x\mathbb{S}^d$ we have*

$$\mathrm{Prob}\left(\left|\frac{\ln\det(dF_n(x)|E)}{n} - \frac{r(d-r)}{d-1}\lambda_1\right| > \epsilon\lambda_1\right) < C\theta^n$$

Corollary 4. *(a) There exist constants C_1, C_2 and $\theta < 1$ such that for any $x \in \mathbb{S}^d$ for any r for any r-dimensional $E \subset T_x\mathbb{S}^d$ we have*

$$\mathrm{Prob}\left(\forall v \in E \quad ||dF_n(x)v|| \geq C_1 \exp\left(\left[\frac{d-2r+1}{d-1} - \epsilon\right]\lambda_1 n\right)\right) \geq 1 - C_2\theta^n.$$

(b) Let $E_+^{(r)}$ and $E_-^{(r)}$ be the Lyapunov spaces generated by r largest and $d-r$ smallest exponents respectively. Then there exist constants C, β such that for any $x \in \mathbb{S}^d$ for any r for any r-dimensional space E for any ε

$$\mathrm{Prob}\left(\angle(E, E_-^{(r)}) \leq \varepsilon\right) \leq C\varepsilon^\beta. \tag{38}$$

(c) For each $\varepsilon > 0$ there exist constants l, κ, α such that for any x for any $d/2$-dimensional E the event
(x, w) is (l, κ)–standard and

$$\angle(W^s(x,w), E) > \alpha \tag{39}$$

has probability greater than $1 - \varepsilon$.

10.3. **Regular points.** By Birkhoff Ergodic Theorem for almost all x, w and for all continuous functions A there exists a limit

$$\nu^{x,w}(A) = \lim_{N\to\infty} \frac{1}{N}\sum_{j=0}^{N-1} A(F_j(w)x).$$

Also it is well known what for random systems $\nu^{x,w}$ does not depend on w, that is for almost all x there exists a measure ν^x such that for almost all w we have $\nu^{x,w} = \nu^x$. (One way to see this is to observe that by Birkhoff Ergodic Theorem the limits as $N \to \infty$ and as $N \to -\infty$ coincide almost surely. However the first limit depends only on w_j, $j \geq 0$ and the second limit depends only on $w_j, j < 0$.) Let $\mathbf{R}_0$ be the set of x such that there exists measure ν^x such that for almost all w

$$\forall A \in C^0(\mathbb{S}^d) \quad \frac{1}{N} \sum_{n=0}^{N-1} A(F_n(w)x) \to \nu^x(A) \tag{40}$$

Let $G = \{(x, w) : (40) \text{ holds.}\}$ Denote

$$\mathbf{R}_0(x) = \{y \in \mathbf{R}_0 : \quad \nu^y = \nu^x\}.$$

Define inductively $\mathbf{R}_{j+1} = \{x \in \mathbf{R}_j$ such that for almost all w $W^s(x, w) \bigcap \mathbf{R}_j(x)$ has the full measure in $W^s(x, w)\}$. Let

$$\mathbf{R}_{j+1}(x) = \{y \in \mathbf{R}_{j+1} : \quad \nu^y = \nu^x\}.$$

We claim that for all j, $\text{Leb}(\mathbf{R}_j) = 1$. This can be seen inductively. Indeed for almost all x, w

- $(x, w) \in G$,
- $\text{Leb}(y : (y, w) \notin G) = 0$,
- (AC1)–(AC3) hold.

Since $\text{Leb}(\mathbb{S}^d - \mathbf{R}_{j-1}) = 0$, (AC1) implies that for almost all x, w

$$\text{mes}\left(W^s(x, w) - \left(\mathbf{R}_{j-1} \bigcap G\right)\right) = 0.$$

However if $(x, w) \in G$, $(y, w) \in G$ and $y \in W^s(x, w)$ then $\nu^y = \nu^x$. Let

$$\mathbf{R}_\infty = \bigcap_j \mathbf{R}_j, \quad \mathbf{R}_\infty(x) = \{y \in \mathbf{R}_\infty : \quad \nu^y = \nu^x\}.$$

Then $\text{Leb}(\mathbf{R}_\infty) = 1$ and we want to show that for almost all x,

$$\text{Leb}(\mathbf{R}_\infty(x)) = 1.$$

10.4. **Positive measure.** Here we recall an argument of Hopf (see [26, 3]) showing that for almost all x the set $\mathbf{R}_\infty(x)$ has positive measure.

Let

$$\varepsilon = 0.1. \tag{41}$$

Let l, κ, α be such that Corollary 4(c) holds with this ε. Choose $x \in \mathbf{R}_\infty$. Choose a coordinate system near x. Take some $l_1 \ll l$. (More precisely we mean that l_1 should be so small that $W^s_{l_1}(x, w)$ is sufficiently close to a $d/2$ dimensional plane for the purposes of determining transverse intersections. So l_1 depends only on l and κ. The readers should have no difficulty of supplying the precise value of l_1 if they wish to do so.) Let w_1 be a word such that

- (x, w_1) is (l, κ)–standard,

- $\text{mes}(W^s_{l_1}(x, w_1) - \mathbf{R}_\infty(x)) = 0$.

Such words exist. Indeed the words satisfying the first property have positive probability by Corollary 4(c) and Fubini and among whose words almost all satisfy the second requirement by Section 10.3.

Let $V = W^s_{l_1}(x, w_1)$. By Corollary 4(c) and Fubini there exist a word w_2 and a subset $V_1 \subset V \bigcap \mathbf{R}_\infty(x)$ such that

- (AC1)–(AC3) hold
- $\text{mes}(V) > \text{mes}(V_1)/2$,
- for all $y \in V_1$ (y, w_2) is (l, κ)–standard and

$$\angle(E_s(y), TV) \geq \alpha/2, \tag{42}$$

- $\text{mes}(W^s_l(y, w_2) - \mathbf{R}_\infty(x)) = \text{mes}(W^s_l(y, w_2) - \mathbf{R}_\infty(y)) = 0$.

By compactness of $\mathbb{G}_{d/2,d}$ and (42) for each sufficiently small δ_1 there exists a universal constant δ_2 and a direction E such that if $\mathcal{K} = \{E' : d(E, E') < \delta_1\}$ then for any $E' \in \mathcal{K}$

$$\angle(E', TV) > \alpha/4 \quad \text{and} \quad \text{mes}(y \in V_1 : d(E_s(y), E) \leq \delta_1) \geq \delta_2.$$

Let

$$Z_1 = \bigcup_{d(E_s(y),E)\leq\delta_1} W^s_l(y, w_2).$$

By (AC2), Z_1 has positive measure and by (AC1),

$$\text{Leb}(Z_1 - \mathbf{R}_\infty(x)) = 0. \tag{43}$$

10.5. **Large measure.** Now take $r \ll l$. By Corollary 4(c), Fubini and Section 10.3 there exists a word w_3 and a set Z_2 such that

- Z_2 has density $1 - 2\varepsilon$ in $B(x, r)$,
- for all $y \in Z_2$ the pair (y, w_3) is (l, κ)–standard,
- for all $y \in Z_2$ $\angle(W^s(y, w_3), E) \geq \alpha$,
- for all $y \in Z_2$ $(y, w_3) \in G$,
- $\text{mes}\,(W^s_l(y, w_3) - (\mathbf{R}_\infty(y) \bigcap G)) = 0$.

Now consider $y \in Z_2$. Recall that w_2 satisfies (AC3). Applying this with V_1 as above and $V_2 = W^s_l(y, w_3)$ we get that $W^s_l(y, w_3) \bigcap Z_1$ has positive measure. By the last property in the definition of Z_2 the set $W^s_l(y, w_3) \bigcap Z_1 \bigcap G$ has positive measure. Since $W^s_l(y, w_3) \bigcap Z_1 \bigcap G$ has positive measure and $(y, w_3) \in G$ (43) gives $\nu^y = \nu^x$. Thus $Z_2 \subset \mathbf{R}_\infty(x)$. Recall (41). We have proved

Proposition 7. *There exist $r > 0$ such that for all $x \in \mathbf{R}_\infty$, $\mathbf{R}_\infty(x)$ has density larger than 80 per cent in $B(x, r)$.*

Remark. *In fact any number greater than 50 percent would suffice for the proof.*

10.6. **Full measure.** *Proof of Corollary 2.* Let r_1 be so small that if x_1 and x_2 are with distance r_1 to each over than the ball of radius r centered at either point has density greater than 99 per cent inside the r–ball around the other point. Then by Proposition 7 $\mathbf{R}_\infty(x_1) = \mathbf{R}_\infty(x_2)$ that is $\nu^x = \nu^y$. Thus almost all points are at the distance more than r_1 from the boundary of their ergodic component. This imply that this boundary is empty. This proves Corollary 2. □

11. Large deviations.

Proof of Lemma 4. We show how to bound $\ln\det(dF_n(x)|E)$ from below. That is, we estimate

$$\text{Prob}\left(\ln\det(dF_n(x)|E) < \left[\frac{r(d-r)}{d-1} - \epsilon\right]\lambda_1\right).$$

The bound from above is similar. By Theorem 2 $\lambda_1 > 0$ implies that there are integer n_0 and $\rho > 0$ such that for any x, E

$$\mathbb{E}\left(\ln\det(dF_{n_0}(x)|E) - \left[\frac{r(d-r)}{d-1} - \frac{\epsilon}{2}\right]n_0\lambda_1\right) \geq \rho. \tag{44}$$

Indeed

$$\ln\det(dF_{n_0}(x)|E) = \sum_{j=0}^{n-1}\ln\det(dF(F_jx)|F_jE)$$

so if (44) failed for infinitely many n (for some points (x_n, E_n)) then taking a weak limit of $\mu_n(A) = (\frac{1}{n}\sum_{j=0}^{n-1}\mathcal{M}^j(A))(x_n, E_n)$ we would get an invariant measure on $G_{r,d}$ violating Theorem 2.

Now using the Taylor expansion

$$|\det(dF_{n_0}(x)|E)|^{-\sigma} = 1 - \sigma\,|\ln\det(dF_{n_0}(x)|E)| + O(\sigma^2)$$

we conclude that for small $\sigma > 0$ for all x, E we have

$$\mathbb{E}\left(\left[\frac{\det(dF_{n_0}(x)|E)}{\exp\left(n_0\lambda_1\left[\frac{r(d-r)}{d-1} - \frac{\epsilon}{2}\right]\right)}\right]^{-\sigma}\right) \leq \gamma(\sigma) < 1.$$

Iterating we obtain inductively

$$\mathbb{E}\left(\left[\frac{\det(dF_{kn_0}(x)|E)}{\exp\left(kn_0\lambda_1\left[\frac{r(d-r)}{d-1} - \frac{\epsilon}{2}\right]\right)}\right]^{-\sigma}\right) \leq \gamma(\sigma)^k$$

and Lemma 4 follows by Chebyshev inequality. □

Proof of Corollary 4. In this proof we let $P_{\tilde{E}}$ denote the orthogonal projection to $\tilde{E}$.

(a) Take x, E as in the statement and let $\{e_1, e_2 \dots e_r\}$ be an orthonormal frame in E. Denote E_j the span of $e_1, e_2 \dots e_j$. Then by Lemma 4 for large n for all $i \geq j$

(45)
$$\exp\left(\left[\frac{d-2j+1}{d-1}-\epsilon\right]\lambda_1 n\right) \leq ||P_{dF_n(E_{j-1})^\perp}\left(dF_n(e_i)\right)|| \leq \exp\left(\left[\frac{d-2j+1}{d-1}+\epsilon\right]\lambda_1 n\right)$$

except for the set of exponentially small probability. For arbitrary $v \in E$ decompose $v = \sum_j c_j e_j$. Take j such that $|c_m| \exp(3\epsilon m n \lambda_1)$ attains the maximal value at $m = j$. Considering orthogonal complement to $dF_n(E_{j-1})$ we obtain that (45) implies that

$$\|dF_n(v)\| \geq \text{Const} c_j \exp\left(\left[\frac{d-2j+1}{d-1}-\epsilon\right]\lambda_1 n\right)$$

(the main contribution comes from $c_j e_j$). On the other hand $|c_j| \geq \text{Const}\exp(-3nd\lambda_1\epsilon)$. Since ϵ is arbitrary (a) follows.

(b) We can restate (38) as follows. Given ε there exists $\tilde{\delta}$ such that for all x, E

$$\text{Prob}\left(\forall E' : d(E, E') < \varepsilon \quad E' \bigcap E_-^{(r)} = \{0\}\right) \geq 1 - \tilde{C}\varepsilon^\beta.$$

Now if $E \bigcap E_-^{(r)} = \{0\}$ then $E' \bigcap E_-^{(r)} = \{0\}$ iff

$$d(dF_n(x)E, dF_n(x)E') \to 0, \quad n \to \infty. \tag{46}$$

To show that (46) has large probability we apply Lemma 6 to the action of dF_n on r-dimensional Grassmanians. To apply this Lemma we need to check that the derivative of this action is a strong contraction (except for a set of exponentially small probability).

Now if E is an r-dimensional space then any space E' nearby is a graph of a map $L : E \to E^\perp$. Now if Q is a matrix then $Q(E')$ is a graph of the map $P_{(QE)^\perp} Q L Q^{-1}$. Thus we have to show what there exists $\gamma > 0$ such that for any $L : E \to E^\perp$ for all n

$$\left\|P_{(dF_n(E))^\perp}\left(dF_n L dF_n^{-1}\right)\right\| \leq C(w)e^{-\gamma n}$$

where the distribution of $C(w)$ has a power tail. On the other hand for all $u \in E$, $v \in E^\perp$ there exists L such that $Lu = v$. Since L is arbitrary the last inequality can be restated as follows

$$\forall v \in E^\perp \quad \forall n \in \mathbb{N} \quad \frac{\left\|P_{(dF_n(E))^\perp}\left(dF_n(v)\right)\right\| / ||v||}{\min_{u\in E} ||dF_n(u)||/||u||} \leq \text{Const} e^{-\gamma n}.$$

Take $\gamma = \frac{\lambda_1}{2(d-1)}$, so that $\gamma < \min_j(\lambda_j - \lambda_{j-1})$. Now given $N \geq 1$ let us estimate the probability that for all n and v

$$\frac{\left\|P_{(dF_n(E))^\perp}\left(dF_n(v)\right)\right\| / ||v||}{\min_{u\in E} ||dF_n(u)||/||u||} \geq Ne^{\gamma n}. \tag{47}$$

Since $\{f_\alpha\}$ are bounded in C^2 there exists a constant $c > 0$ such that (47) can only happen for $n > c \ln N$. However by Lemma 4 and part (a) of the present corollary for most trajectories the numerator is

$$O\left(\exp\left(\left[\frac{d-2r-1}{d-1} + \epsilon\right] \lambda_1 n\right)\right)$$

whereas the denominator is at least

$$\text{Const} \exp\left(\left[\frac{d-2r+1}{d-1} - \epsilon\right] \lambda_1 n\right)$$

and the exceptional set has measure at most $C\theta^n$. Therefore the probability that (47) holds for some n is less than

$$\text{Const} \sum_{n=c\ln N}^{\infty} \theta^n = \text{Const} N^{-c|\ln\theta|}$$

and (b) follows.

(c) In view of Proposition 5 and part (b) of the present corollary it remains to show that

$$\text{Prob}(x \notin \Lambda_{C,\epsilon}) \to 0$$

as $C \to \infty$. The fact that (36) fails on a small probability set follows from part (a) and Borel-Cantelli. Applying part (a) to $\{f_\alpha^{-1}\}$ we conclude that (35) fails on a small probability set. Finally, part (b) and the fact that E_u does not depend on the future imply that

$$\text{Prob}(\angle(E_u, E_s) < \varepsilon) \leq \text{Const}\varepsilon^\beta.$$

This shows that (37) fails on a small probability set. □

Appendix A. Linear Algebra.

Proof of Lemma 1. Observe that

$$\Lambda_r = \varepsilon\alpha_1(L_1) + \varepsilon^2\alpha_2(L_1) + \varepsilon^2\alpha_3(L_2)$$

where α_1 and α_3 are linear and α_2 is quadratic. To compute α_1 write

$$\alpha_1(L_1) = \alpha_1\left(\frac{L_1 + L_1^*}{2}\right) + \alpha_1\left(\frac{L_1 - L_1^*}{2}\right).$$

Now

$$\alpha_1\left(L_1 - L_1^*\right) = \frac{d}{d\varepsilon}\int_{\mathbb{G}_{r,d}} \ln\det\left(e^{(L_1-L_1^*)\varepsilon}|E\right)dE = 0$$

so we may assume that L_1 is symmetric. Since α_1 is invariant under conjugations, we get $\alpha_1 = a_1\mathrm{Tr}$. Substituting $L_1 = 1$ we get $a_1{=}r/d$.

Next, letting $L_1 = 0$ we obtain $\alpha_3 = \alpha_1$.

To compute α_2 observe that Λ_r does not change if we replace $L(\varepsilon)$ by $L(\varepsilon)e^{-\varepsilon J}$ for any skew symmetric J. Take $J = \frac{L_1 - L_1^*}{2}$. Then

$$L(\varepsilon)e^{-\varepsilon J} = 1 + \varepsilon(L_1 - J) + \varepsilon^2\left(L_2 + \frac{J^2}{2} - L_1 J\right) + \dots$$

It follows that

$$\alpha_2(L_1) = \alpha_2\left(\frac{L_1 + L_1^*}{2}\right) + \frac{r}{d}\mathrm{Tr}\left(\frac{J^2}{2} - L_1 J\right). \tag{48}$$

Now

$$\mathrm{Tr}\left(\frac{J^2}{2} - L_1 J\right) = \mathrm{Tr}\left(\frac{L_1^* L_1 - L_1^2}{4}\right). \tag{49}$$

Since α_2 is invariant under conjugations, we obtain

$$\alpha_2\left(\frac{L_1 + L_1^*}{2}\right) = b_1(\mathrm{Tr}L_1)^2 + b_2\mathrm{Tr}\left(\frac{L_1 + L_1^*}{2} - \frac{\mathrm{Tr}L_1}{d}\right)^2. \tag{50}$$

Substituting again $L_1 = 1$ we get

$$b_1 = -\frac{r}{2d^2}. \tag{51}$$

To compute b_2 consider the case then L_1 is a projection onto some vector e. Then

$$\det(L(\varepsilon)|E) = \sqrt{(1+\varepsilon)^2\cos^2\angle(E,e) + \sin^2\angle(E,e)}.$$

Hence

$$\ln\det(L(\varepsilon)|E) = \varepsilon\cos^2\angle(E,e) + \varepsilon^2\left[\frac{\cos^2\angle(E,e)}{2} - \cos^4\angle(E,e)\right].$$

So

$$\int_{\mathbb{G}_{r,d}} \ln(\det(L(\varepsilon)|E))dE =$$

$$\varepsilon\int_{\mathbb{G}_{r,d}} \cos^2\angle(E,e)dE + \varepsilon^2\int_{\mathbb{G}_{r,d}}\left[\frac{\cos^2\angle(E,e)}{2} - \cos^4\angle(E,e)\right]dE.$$

Let E_0 be the span of the first r coordinate vectors. We now use the following formulas.

$$\int_{\mathbb{S}^{d-1}} x_1^2 dx = \frac{1}{d}, \quad \int_{\mathbb{S}^{d-1}} x_1^4 dx = \frac{3}{d(d+2)}, \quad \int_{\mathbb{S}^{d-1}} x_1^2 x_2^2 dx = \frac{1}{d(d+2)}.$$

where

$$\mathbb{S}^{d-1} = \{x : \sum_{j=1}^{d} x_j^2 = 1\}.$$

Now

$$\int_{\mathbb{G}_{r,d}} \cos^2 \angle(E, e) dE = \int_{\mathbb{SO}_d} \cos^2 \angle(gE_0, e) dg = \int_{\mathbb{SO}_d} \cos^2 \angle(E_0, ge) dg =$$

$$\int_{\mathbb{S}^{d-1}} \cos^2 \angle(E_0, x) dx = \sum_{j=1}^{r} \int_{\mathbb{S}^{d-1}} x_i^2 dx = \frac{r}{d}.$$

Similarly

$$\int_{\mathbb{G}_{r,d}} \left[\frac{\cos^2 \angle(E, e)}{2} - \cos^4 \angle(E, e) \right] dE = \frac{r}{2d} - \frac{r(r+2)}{d(d+2)}.$$

On the other hand

$$(\mathrm{Tr} L_1)^2 = 1, \quad \mathrm{Tr} \left(L_1 - \frac{1}{d} \right)^2 = \frac{d-1}{d}.$$

This gives

$$\frac{r}{2d} - \frac{r(r+2)}{d(d+2)} = -\frac{r}{2d^2} + b_2 \frac{d-1}{d}.$$

The LHS of this equation equals

$$\frac{r(d-r)}{d(d+2)} - \frac{r}{2d}.$$

Hence

$$b_2 = \frac{r(d-r)}{(d-1)(d+2)} - \frac{r}{2d}.$$

Now if K is a matrix then

$$\mathrm{Tr} \left(K - \frac{\mathrm{Tr} K}{d} 1 \right)^2 = \mathrm{Tr} K^2 - \frac{(\mathrm{Tr} K)^2}{d}.$$

Recall (49), (51). Combine

$$\frac{r}{d} \left[-\frac{1}{2} \mathrm{Tr} \left(\frac{L_1 + L_1^*}{2} - \frac{\mathrm{Tr} L_1}{d} \right)^2 + \mathrm{Tr} \left(\frac{L_1^* L_1 - L_1^2}{4} \right) \right] - \frac{r}{2d^2} (\mathrm{Tr} L_1)^2 =$$

$$-\frac{r}{2d} \mathrm{Tr} L_1^2.$$

The lemma is proven. □

References

[1] Alves J., Bonnatti C.& Viana M. *SRB measures for partially hyperbolic systems: mostly expanding case*, Inv. Math. **140** (2000) 351–398.
[2] Bakhtin V. I. *Random processes generated by a hyperbolic sequence of mappings,* Russian Acad. Sci. Izv. Math. **44** (1995) I: 247–279, II: 617–627.
[3] Barreira L. & Pesin Ya. B. *Lyapunov exponents and smooth ergodic theory,* University Lecture Series, 23. AMS, Providence, RI, 2002. xii+151 pp.
[4] Burns K, Pugh C. C., Shub M. & Wilkinson A. *Recent results about stable ergodicity,* Proc. Symposia A.M.S., **69** (2001) 327-366.
[5] Chulaevsky V. & Spencer T. *Positive Lyapunov exponents for a class of deterministic potentials,* Comm. Math. Phys. **168** (1995) 455–466.
[6] de la Llave R., Marco J. M. & Moriyon R. *Canonical perturbation theory of Anosov systems and regularity results for Livsic cohomology equation* Ann. Math. **123** (1986) 537–611.
[7] Damjanovic D. & Katok A. *Local rigidity of actions of higher rank abelian groups and KAM method,* Electron. Res. Announc. AMS **10** (2004), 142–154.
[8] Dolgopyat D. *On mixing properties of compact group extensions of hyperbolic systems*, Israel J. Math. **130** (2002) 157–205.
[9] Dolgopyat D. *On differentiability of SRB states for partially hyperbolic systems,* Inv. Math. **155** (2004) 389–449.
[10] Field M., Melbourne I. & Torok A. *Decay of correlations, central limit theorem and approximation by Brownian motion for compact group extensions,* Erg. Th. Dyn. Sys. **23** (2003), 87–110.
[11] Figotin A. & Pastur L. *Spectra of random and almost-periodic operators,* Grundlehren der mathematischen Wissenschaften **297** (1992) Springer, Berlin-Heidelberg-New York.
[12] Fisher D. and Margulis G. A. *Almost isometric actions, property (T), and local rigidity,* to appear in Inv. Math.
[13] Fujiwara D. *Elliptic partial differential operators on a manifold,* appendix to Kodaira K. *Complex manifolds and deformations of complex structures,* (Grundlehren der Mathematischen Wissenschaften **283** (1986) Springer-Verlag, New York) 353–458.
[14] Ghys E. *Groups acting on the circle,* L'Enseignement Math. **47** (2001) 329–407.
[15] Herman M. R. *Sur les courbes invariantes par les difféomorphismes de l'anneau,* Astérisque **103-104** (1983) 1–221.
[16] Hörmander L. *The boundary problem of physical geodesy,* Arch. for Rat. Mech. Analysis **62** (1976) 1-52.
[17] Hertz F. *Stable ergodicity of certain automorphisms of torus,* Ann. Math. **162** (2005).
[18] Katok A., Knieper G. & Weiss H. *Formulas for the derivative and critical points of topological entropy for Anosov and geodesic flows,* Comm. Math. Phys. **138** (1991) 19–31.
[19] Kifer Y. *Ergodic theory of random transformations,* Progress Prob. Stat. **10** (1986) Birkhauser Boston, Inc., Boston.
[20] Krikorian R. *Réductibilité des systèmes produits-croisés réductibles à valeurs dans des groupes compacts,* Astérisque **259** (1999) 1–216

[21] Krikorian R. *Global density of reducible quasi-periodic cocycles on* $\mathbf{T}^1 \times \mathrm{SU}(2)$*. Ann. of Math. (2)* **154** *(2001), no. 2, 269–326.*
[22] Le Jan Y. *On isotropic Brownian motions,* Z. Wahrsch. Verw. Gebiete **70** (1985), 609–620.
[23] Liu P.-D. & Qian M. *Smooth ergodic theory of random dynamical systems,* Lect. Notes Math. **1606** (1995) Springer, Berlin–Heidelberg.
[24] Moser J. *On commuting circle mappings and simultaneous Diophantine approximations,* Math. Z. **205** (1990) 105–121.
[25] Navas A. *Sur les groupes de diffeomorphismes du circle engendrés par des éléments proches des rotations,* L'Enseignement Math. **50** (2004) 29–68.
[26] Pesin Ya. B. *Characteristic Ljapunov exponents, and smooth ergodic theory,* Uspehi Mat. Nauk **32** (1977), no. 4, 55–112.
[27] Ruelle D. *Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms,* Nonlinearity **11** (1998), 5–18.
[28] Ruelle D. *Perturbation theory for Lyapunov exponents: an extension of a result of Shub and Wilkinson,* Israel J. Math. **134** (2003) 345–361.
[29] Shub M. & Wilkinson A. *Pathological foliations and removable zero exponents,* Inv. Math. **139** (2000), 495–508.
[30] Yoccoz J.-C. *Travaux de Herman sur les tores invariants,* Astérisque **206** (1992) 311–344.