J. Eur. Math. Soc. (Online first) DOI 10.4171/JEMS/1327

Published by EMS Press

Jonathan DeWitt

Simultaneous linearization of diffeomorphisms of isotropic manifolds

Received December 3, 2020; revised June 23, 2022

Abstract. Suppose that M is a closed isotropic Riemannian manifold and that $R_1, \dots, R_m$ generate the isometry group of M. Let $f_1, \dots, f_m$ be smooth perturbations of these isometries. We show that the f_i are simultaneously conjugate to isometries if and only if their associated uniform Bernoulli random walk has all Lyapunov exponents zero. This extends a linearization result of Dolgopyat and Krikorian [Duke Math. J. 136, 475–505 (2007)] from S^n to real, complex, and quaternionic projective spaces. In addition, we identify and remedy an oversight in that earlier work.

Keywords. Linearization, Diophantine, isotropic, KAM, Lyapunov exponents, symmetric space

1. Introduction

A basic problem in dynamics is determining whether two dynamical systems are equivalent. A standard notion of equivalence is conjugacy: if f and g are two diffeomorphisms of a manifold M, then f and g are *conjugate* if there exists a homeomorphism h of M such that $hfh^{-1} = g$. Some classes of dynamical systems are distinguished up to conjugacy by a small amount of dynamical information. One of the most basic examples of this is Denjoy's theorem: a C^2 orientation preserving circle diffeomorphism with irrational rotation number is conjugate to a rotation [17, §12.1]. In the case of Denjoy's theorem, the rotation number is all the information needed to determine the topological equivalence class of the diffeomorphism under conjugacy.

Rigidity theory focuses on identifying dynamics that are distinguished up to conjugacy by particular kinds of dynamical information such as the rotation number. There are finer dynamical invariants than rotation number which require a finer notion of equivalence to study. For instance, one obtains a finer notion of equivalence if one insists that the conjugacy be a C^1 or even C^∞ diffeomorphism. A smoother conjugacy allows one to consider invariants such as Lyapunov exponents, which may not be preserved under conjugacy by homeomorphisms. For a single volume preserving Anosov diffeomorphism, the Lyapunov exponents with respect to volume are invariant under conjugation by C^1 vol-

Jonathan DeWitt: Department of Mathematics, University of Chicago, Chicago, IL 60637, USA; dewitt@uchicago.edu

Mathematics Subject Classification (2020): Primary 37H15; Secondary 37J40

ume preserving maps. Consequently, one is naturally led to ask, “If two volume preserving Anosov diffeomorphisms have the same Lyapunov exponents, are the two C^1 conjugate?” In some circumstances the answer is “yes”. Such situations where knowledge about Lyapunov exponents implies systems are conjugate by a C^1 diffeomorphism are instances of a phenomenon called “Lyapunov spectrum rigidity”. See [13] for examples and discussion of this type of rigidity. For recent examples, see [4, 8, 14, 15, 27].

In rigidity problems related to isometries, it is often natural to consider a family of isometries. A collection of isometries may have strong rigidity properties even if the individual elements of the collection do not. For example, Fayad and Khanin [11] proved that a collection of commuting diffeomorphisms of the circle whose rotation numbers satisfy a simultaneous Diophantine condition are smoothly simultaneously conjugate to rotations. Their result is a strengthening of an earlier result of Moser [25]. A single diffeomorphism in such a collection might not satisfy the Diophantine condition on its own.

Although the two types of rigidity described above occur in the dissimilar hyperbolic and elliptic settings, a result of Dolgopyat and Krikorian combines the two. They introduce a notion of a Diophantine set of rotations of a sphere and use this notion to prove that certain random dynamical systems with all Lyapunov exponents zero are conjugate to isometric systems [10]. Our result is a generalization of this result to the setting of isotropic manifolds. We now develop the language to state both precisely.

Let $(f_1, \dots, f_m)$ be a tuple of diffeomorphisms of a manifold M. Let $(\omega_i)_{i\in\mathbb{N}}$ be a sequence of independent and identically distributed random variables with uniform distribution on $\{1, \dots, m\}$. Given an initial point $x_0 \in M$, define $x_n = f_{\omega_n} x_{n-1}$. This defines a Markov process on M. We refer to this process as the random dynamical system associated to the tuple $(f_1, \dots, f_m)$. Let f^n_ω be defined to equal $f_{\omega_n} \circ \cdots \circ f_{\omega_1}$. We say that a probability measure μ on M is a *stationary measure* for this process if $m^{-1} \sum_{i=1}^m (f_i)_* \mu = \mu$. A stationary measure is *ergodic* if it is not a non-trivial convex combination of two distinct stationary measures. Fix an ergodic stationary measure μ. For μ-almost every x, almost surely for any $v \in T_x M \setminus \{0\}$, the limit

$$\lim_{n\to\infty} \frac{1}{n} \ln \|D_x f^n_\omega v\| \tag{1}$$

exists and takes its value in a fixed finite list of numbers depending only on μ:

$$\lambda_1(\mu) \geq \cdots \geq \lambda_{\dim M}(\mu). \tag{2}$$

These are the *Lyapunov exponents* with respect to μ. In fact, for almost every ω and μ-a.e. x there exists a flag $V_1 \subset \cdots \subset V_j$ inside $T_x M$ such that if $v \in V_i \setminus V_{i-1}$ then the limit in (2) is equal to $\lambda_{\dim M - \dim V_i}$. The number of times a particular exponent appears in (2) is given by $\dim V_i - \dim V_{i-1}$; this number is referred to as the multiplicity of the exponent. For more information, see [20].

Our result holds for isotropic manifolds. By definition, an *isotropic manifold* is a Riemannian manifold whose isometry group acts transitively on its unit tangent bundle. The closed isotropic manifolds are S^n, $\mathbb{R}\mathrm{P}^n$, $\mathbb{C}\mathrm{P}^n$, $\mathbb{H}\mathrm{P}^n$, and the Cayley projective plane. In the following we write G° for the identity component of a Lie group G.

Theorem 1. *Let M^d be a closed isotropic Riemannian manifold other than S^1. There exists k_0 such that if $(R_1, \ldots, R_m)$ is a tuple of isometries of M such that the subgroup of* $\operatorname{Isom}(M)$ *generated by this tuple contains* $\operatorname{Isom}(M)^\circ$*, then there exists $\epsilon_{k_0} > 0$ such that the following holds. Let $(f_1, \ldots, f_m)$ be a tuple of C^∞ diffeomorphisms satisfying* $\max_i d_{C^{k_0}}(f_i, R_i) < \epsilon_{k_0}$*. Suppose that there exists a sequence of ergodic stationary measures μ_n for the random dynamical system generated by $(f_1, \ldots, f_m)$ such that* $|\lambda_d(\mu_n)| \to 0$*. Then there exists* $\psi \in \operatorname{Diff}^\infty(M)$ *such that for each i the map $\psi f_i \psi^{-1}$ is an isometry of M and lies in the subgroup of* $\operatorname{Isom}(M)$ *generated by* $(R_1, \ldots, R_m)$*.*

Dolgopyat and Krikorian proved Theorem 1 in the case of S^n [10, Thm. 1].

Dolgopyat and Krikorian also obtained a Taylor expansion of the Lyapunov exponents of the stationary measure of the perturbed system [10, Thm. 2]. Fix $(R_1, \ldots, R_m)$ generating $\operatorname{Isom}(S^n)^\circ$. Let $(f_1, \ldots, f_m)$ be a C^{k_0} small perturbation of $(R_1, \ldots, R_m)$ and μ be any ergodic stationary measure for $(f_1, \ldots, f_m)$. Let $\Lambda_r = \lambda_1 + \cdots + \lambda_r$ denote the sum of the top r Lyapunov exponents. In [10, Thm. 2], it is shown that the Lyapunov exponents of μ satisfy

$$\lambda_r(\mu) = \frac{\Lambda_d}{d} + \frac{d-2r+1}{d-1}\left(\lambda_1 - \frac{\Lambda_d}{d}\right) + o(1)|\lambda_d(\mu)|, \tag{3}$$

where $o(1)$ goes to zero as $\max_i d_{C^{k_0}}(f_i, R_i) \to 0$. Using this formula Dolgopyat and Krikorian obtain an even stronger dichotomy for systems on even-dimensional spheres: either $(f_1, \ldots, f_m)$ is simultaneously conjugate to isometries or the Lyapunov exponents of every ergodic stationary measure of the perturbation are uniformly bounded away from zero. By using this result they show if $(R_1, \ldots, R_m)$ generates $\operatorname{Isom}(S^{2n})^\circ$ and $(f_1, \ldots, f_m)$ is a C^{k_0} small perturbation such that each f_i preserves volume, then volume is an ergodic stationary measure for $(f_1, \ldots, f_m)$ [10, Cor. 2].

It is natural to ask if a similar Taylor expansion can be obtained in the setting of isotropic manifolds. Proposition 26 shows that Λ_r may be Taylor expanded assuming that $(R_1, \ldots, R_m)$ generates $\operatorname{Isom}(M)^\circ$ and the induced action of $\operatorname{Isom}(M)^\circ$ on $\operatorname{Gr}_r(M)$, the Grassmannian bundle of r-planes in TM, is transitive.

In Theorem 40, we give a Taylor expansion relating λ_1 and λ_d which holds for isotropic manifolds. However, we cannot Taylor expand every Lyapunov exponent as in equation (3) because if a manifold does not have constant curvature then its isometry group cannot act transitively on the 2-planes in its tangent spaces. The argument of Dolgopyat and Krikorian requires that the isometry group act transitively on the space of k-planes in TM for $0 \le k \le d$.

It is natural to ask why the proof of Theorem 1 does not work in the case of S^1 even though S^1 is isotropic. As Proposition 13 shows, for a tuple $(R_1, \ldots, R_m)$ as in the theorem, uniformly small perturbations of $(R_1, \ldots, R_m)$ are uniformly Diophantine in a sense explained below. This uniformity is used crucially in the proof when we change the tuple of isometries that we are working with. The same uniformity of Diophantineness does not hold for tuples of isometries of S^1: a small perturbation may lose all Diophantine properties. The reason that the proof of Proposition 13 does not work for S^1 is that the isometry group of S^1 is not semisimple.

There are not many other results like Theorem 1. In addition to the aforementioned result of Dolgopyat and Krikorian, there are some results of Malicet. In [24], a similar linearization result is obtained that applies to a particular type of map of $\mathbb{T}^2$ that fibers over a rotation on S^1. In a recent work, Malicet obtained a Taylor expansion of the Lyapunov exponent for a perturbation of a Diophantine random dynamical system on the circle [23].

1.1. Outline

The proof of Theorem 1 follows the general argument of [10]. For readability, the argument in this paper is self-contained. While a number of the results below appear in [10], we have substantially reformulated many of them and in many places offer a different proof. Doing so is not merely a courtesy to the reader: the results in [10] are stated in too narrow a setting for us to use. Simply stating more general reformulations would unduly burden the reader's trust. In addition, as will be discussed below, there are some oversights in [10] which we explain in Section 1.2 and that we have remedied in Section 5. We have also stated intermediate results and lemmas in more generality than is needed for the proof of Theorem 1 so that they may be used by others. Below we sketch the general argument of the paper and emphasize some differences with the approach in [10].

The proof of Theorem 1 is by an iterative KAM convergence scheme. Fix a closed isotropic manifold M. We start with a tuple $(f_1, \dots, f_m)$ of diffeomorphisms close to a tuple $(R_1, \dots, R_m)$ of isometries. We must find some smooth diffeomorphism ψ such that $\widetilde{f_i} := \psi f_i \psi^{-1} \in \mathrm{Isom}(M)$. To do this we produce a conjugacy ψ that brings each f_i closer to being an isometry. To judge the distance from being an isometry, we define a strain tensor that vanishes precisely when a diffeomorphism is an isometry. By solving a particular coboundary equation and using the fact that the Lyapunov exponents are zero, we can construct ψ so that $\widetilde{f_i}$ has small strain tensor. In our setting, a diffeomorphism with small strain is near to an isometry, so $(\widetilde{f_1}, \dots, \widetilde{f_m})$ is near to a tuple $(R_1', \dots, R_m')$ of isometries. We then repeat the procedure using these new tuples as our starting point. The results of performing a single step of this procedure comprise Lemma 39. Once Lemma 39 is proved, the rest of the proof of Theorem 1 is bookkeeping that checks that the procedure converges. Most of the paper is in service of the proof Lemma 39, which gives the result of a single step in the convergence scheme.

Proofs of technical and basic facts are relegated to a significant number of appendices. This has been done to focus the main exposition on the important ideas in the proof of Theorem 1 and not on the technical details. The appendices that might be most beneficial to look at before they are referenced in the text are Appendices A and B, concerning C^k calculus and interpolation inequalities. Both contain estimates that are common in KAM arguments. The organization of the main body of the paper reflects the order of the steps in the proof of Lemma 39. There are several important results in the proof of Lemma 39, which we now describe.

The first part of the proof of Lemma 39 requires that a particular coboundary equation can be tamely solved. The solution to this equation is one of the main subjects of Section 2. The equation is solved in Proposition 16. This proposition is essential in the work of Dolgopyat and Krikorian [10] and its proof follows from the appendix to [9]; it relies on a Diophantine property of the tuple $(R_1, \ldots, R_m)$ of isometries. This property is formulated in Section 2.2. The stability of this property under perturbations is crucial in the proof and an essential feature of our setting. In addition, the argument in Section 2 is different from Dolgopyat's earlier argument because we use the Solovay–Kitaev algorithm (Theorem 2), which is more efficient than the procedure used in the appendix to [9].

Section 3 considers stationary measures for perturbations of $(R_1, \ldots, R_m)$. Suppose M is a quotient of its isometry group, its isometry group is semisimple, and $\{R_1, \ldots, R_m\}$ is a Diophantine subset of $\text{Isom}(M)$. Suppose $(f_1, \ldots, f_m)$ is a small smooth perturbation of $(R_1, \ldots, R_m)$. There is a relation between a stationary measure μ for the perturbed system and the Haar measure. Proposition 23 relates integration against μ to integration against the Haar measure. Lyapunov exponents are calculated by integrating the ln Jacobian against a stationary measure of an extended dynamical system on a Grassmannian bundle over M. Consequently, this proposition relates stationary measures and their Lyapunov exponents to the volume on a Grassmannian bundle.

The relationship between Lyapunov exponents and stationary measures is explained in Section 4. Proposition 26 provides a Taylor expansion of the sum of the top r Lyapunov exponents of a stationary measure μ. Three terms appear in the Taylor expansion. The first two have a direct geometric meaning, which we interpret in terms of strain tensors introduced in Section 4.2. The final term in the Taylor expansion depends on a quantity $\mathcal{U}(\psi)$. This quantity does not have a direct geometric interpretation. However, in the proof of Lemma 39, we show that by solving the coboundary equation from Proposition 16 the quantity $\mathcal{U}(\psi)$ can be made to vanish. Once $\mathcal{U}(\psi)$ vanishes, we have an equation directly relating Lyapunov exponents to the strain. This equation then allows us to conclude that a diffeomorphism with small Lyapunov exponents also has small strain. We reformulate in a Riemannian geometric setting some arguments of [10] by using the strain tensor. This gives coordinate-free expressions that are easier to interpret.

Section 5 contains the most important connection between the strain tensor and isometries: diffeomorphisms of small strain on isotropic manifolds are near to isometries. The basic geometric fact proved in Section 5 is Theorem 27, which is true on any manifold. Theorem 27 is then used to prove Proposition 28, which is a more technical result adapted for use in the KAM scheme. Proposition 28 then allows us to prove that our conjugate tuple is near to a new tuple of isometries, which allows us to repeat the process.

All of the previous sections combine in Section 6 to prove Lemma 39. We then obtain the main theorem, Theorem 1, and prove an additional theorem that relates the top and bottom Lyapunov exponents of a perturbation, Theorem 40.

1.2. An oversight and its remedy

Section 5 is entirely new and different from anything appearing in [10]. Consequently, the reader may wonder why it is needed. Section 5 provides a method of finding a tuple $(R_1', \ldots, R_m')$ of isometries near to the tuple $(\widetilde{f}_1, \ldots, \widetilde{f}_m)$ of diffeomorphisms. In [10], the new diffeomorphisms R_m are found in the following manner. As in (10), one may find vector fields Y_i such that

$$\exp_{R_i(x)} Y_i(x) = f_i(x).$$

If Z is a vector field on M, we define ψ_Z, as in (11) in Section 3.1, to be the map $x \mapsto \exp_x Z(x)$. There is a certain operator, the Casimir Laplacian, which acts on vector fields. This operator is defined and discussed in more detail in Section 2.2. Dolgopyat and Krikorian then project the vector fields Y_i onto the kernel of the Casimir Laplacian, to obtain a vector field Y_i'. They then define R_i' to equal $\psi_{Y_i'} \circ R_i$. This happens in the line immediately below (19) in [10].

One difficulty is establishing that the maps $(R_1', \dots, R_m')$ are close to the $(\widetilde{f}_i, \dots, \widetilde{f}_m)$. The argument for their nearness hinges on part (d) of Proposition 3 in [10], which essentially says that, up to a third order error, the magnitude of the smallest Lyapunov exponent is a bound on the distance. As written, the argument in [10] suggests that part (d) is an easy consequence of part (c) of [10, Prop. 3]. However, part (d) does not follow. Here is a simplification of the problem. Suppose that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism. Pick a point $x \in \mathbb{R}^n$ and write $D_x f = A + B + C$, where A is a multiple of the identity, B is symmetric with trace zero, and C is skew-symmetric. The results in part (c) imply that A and B are small, but they offer no information about C.[1] Concluding that the norm of Df is small requires that C be small as well. As C is skew-symmetric it is natural to think of it as the germ of an isometry. Our modification to the argument is designed to accommodate the term C by recognizing it as the "isometric" part of the differential. Pursuing this perspective leads to the strain tensor and our Proposition 28. Conversation with Dmitry Dolgopyat confirmed that there is a problem in the paper on this point and that part (d) of Proposition 3 does not follow from part (c).

2. A Diophantine property and spectral gap

Fix a compact connected semisimple Lie group G and let $\mathfrak{g}$ denote its Lie algebra. Endow G with the bi-invariant metric arising from the negative of the Killing form on $\mathfrak{g}$. We denote this metric on G by d. We endow a subgroup H of G with the pullback of the Riemannian metric from G and denote the distance on H with respect to the pullback metric by d_H. We use the manifold topology on G unless explicitly stated otherwise. Consequently, whenever we say that a subset of G is dense, we mean this with respect to the manifold topology on G. We say that a subset S of G *generates* G if the smallest closed subgroup of G containing S is G. In other words, if $\langle S \rangle$ denotes the smallest subgroup of G containing S, then S generates if $\overline{\langle S \rangle} = G$.

Suppose that $S \subset G$ generates G. We begin this section by discussing how long a word in the elements of S is needed to approximate an element of G. Then using this

[1] For those comparing with the original paper, A and B correspond to the terms q_1 and q_2, respectively, which appear in part (c) of [10, Prop. 3].

approximation we obtain quantitative estimates for the spectral gap of certain operators associated to S. Finally, those spectral gap estimates allow us to obtain a "tameness" estimate for a particular operator that arises from S. This final estimate, Proposition 16, will be crucial in the KAM scheme that we use to prove Theorem 1.

The content of this section is broadly analogous to Appendix A in [9]. However, our development follows a different approach and in some places we are able to obtain stronger estimates.

2.1. The Solovay–Kitaev algorithm

Suppose that S is a subset of G. We say that S is *symmetric* if $s \in S$ implies $s^{-1} \in S$. For a natural number n, let S^n denote the n-fold product of S with itself. Let S^{-1} be $\{s^{-1} : s \in S\}$. For $n < 0$, define S^n to equal $(S^{-1})^{-n}$. The following theorem says that any sufficiently dense symmetric subset S of a compact semisimple Lie group is a generating set. More importantly, it also gives an estimate on how long a word in the generating set S is needed to approximate an element of G to within error ϵ. If $w = s_1 \cdots s_n$ is a word in the elements of the set S, then we say that w is *balanced* if for each $s \in S$, s appears the same number of times in w as s^{-1} does.

Theorem 2 (Solovay–Kitaev algorithm [7, Thm. 1]). *Suppose that G is a compact semisimple Lie group. There exist $\epsilon_0(G) > 0$, $\alpha > 0$ and $C > 0$ such that if S is any symmetric ϵ_0-dense subset of G then the following holds. For any $g \in G$ and any $\epsilon > 0$, there exists a natural number l_ϵ such that $d(g, S^{l_\epsilon}) < \epsilon$. Moreover, $l_\epsilon \le C \ln^\alpha(1/\epsilon)$. Further, there is a balanced word of length l_ϵ within distance ϵ of g.*

Later, we use a version of this result that does not require that the set S be symmetric. Using a non-symmetric generating set significantly increases the word length obtained in the conclusion of the theorem. It is unknown if there exists a version of the Solovay–Kitaev algorithm that does not require a symmetric generating set and keeps the $O(\ln^\alpha(1/\epsilon))$ word length. See [3] for a partial result in this direction.

Proposition 3. *Suppose that G is a compact semisimple Lie group endowed with a bi-invariant metric. There exist $\epsilon_0(G) > 0$, $\alpha > 0$, and $C \ge 0$ such that if S is any ϵ_0-dense subset of G then the following holds. For any $g \in G$ and any $\epsilon > 0$, there exists a natural number l_ϵ such that $d(g, S^{l_\epsilon}) < \epsilon$. Moreover, $l_\epsilon \le C\epsilon^{-\alpha}$.*

Our weakened version of the Solovay–Kitaev algorithm relies on the following lemma, which allows us to approximate the inverse of an element h by some positive power of h.

Lemma 4. *Suppose that G is a compact d-dimensional Lie group with a fixed bi-invariant metric. Then there exists a constant C such that for all $\epsilon > 0$ and any $h \in G$ there exists a natural number $n < C/\epsilon^d$ such that $d(h^{-1}, h^n) < \epsilon$.*

Proof. This follows from a straightforward pigeonhole argument. We cover G with sets of diameter ϵ. There exists a constant C such that we can cover G with at most $C\,\mathrm{vol}(G)/\epsilon^d$

such sets, where d is the dimension of G. Consider now the first $\lceil C\ \mathrm{vol}(G)/\epsilon^d \rceil$ iterates of h^2. By the pigeonhole principle, two of these must fall into the same set in the covering, and so there exist natural numbers n_i and n_j such that $0 < n_i < n_j < \lceil C\ \mathrm{vol}(G)/(\epsilon^d)\rceil$ and h^{2n_i} and h^{2n_j} lie in the same set in the covering. Thus $d(h^{2n_i}, h^{2n_j}) < \epsilon$. As h is an isometry it follows that $d(e, h^{2n_j-2n_i}) < \epsilon$ and hence $d(h^{-1}, h^{2n_j-2n_i-1}) < \epsilon$ as well. This finishes the proof. ■

We now prove the proposition.

Proof of Proposition 3. Let $\hat{S} = S \cup S^{-1}$. As $\hat{S}$ is a symmetric generating set of G, by Theorem 2 for any $\epsilon > 0$ there exists a number $l_{\epsilon/2} = O(\ln^\alpha(1/\epsilon))$ such that for any $g \in G$ there exists an element h in $\hat{S}^{l_{\epsilon/2}}$ such that $d(h, g) < \epsilon/2$. Further, by the statement of Theorem 2, we know that h is represented by a balanced word w in $\hat{S}^{l_{\epsilon/2}}$.

To finish the proof, we replace each element of w that is in S^{-1} by a word in S^j for some uniform $j > 0$. To do this we show that there exists a fixed j such that the elements of S^j approximate well the inverses of the elements of S. Write $S = \{s_1, \ldots, s_m\}$ and consider the element $(s_1, \ldots, s_m)$ in the group $G \times \cdots \times G$ (m terms). By applying Lemma 4 to the group $G \times \cdots \times G$ and the element $(s_1, \ldots, s_m)$, we find that there exists a uniform constant C' and $j < C' 2^{dm} l_{\epsilon/2}^{dm}/\epsilon^{dm}$ such that any $s \in S^{-1}$ may be approximated to distance $\epsilon/(2l_{\epsilon/2})$ by an element in S^j.

We now replace each element of S^{-1} appearing in w with a word in S^j that is at distance $\epsilon/(2l_{\epsilon/2})$ from it. Call this new word w'. Because w is balanced, we replace exactly half of the terms in w. Thus w' is a word of length $j l_{\epsilon/2}/2 + l_{\epsilon/2}/2$ as we have replaced half the entries of w, which has length $l_{\epsilon/2}$, with words of length j. Let h' be the element of G obtained by multiplying together the terms in w'.

Note that multiplication of any number of elements of G is 1-Lipschitz in each argument. Hence as we have modified the expression for h in exactly $l_{\epsilon/2}/2$ terms and each modification is of size $\epsilon/(2l_{\epsilon/2})$, h' is at most $\epsilon/2$ away from h and hence at most ϵ away from g. Thus $S^{j l_{\epsilon/2}/2 + l_{\epsilon/2}/2}$ is ϵ-dense in G and

$$j l_{\epsilon/2}/2 + l_{\epsilon/2}/2 < C'' l_{\epsilon/2}^{dm+1}/\epsilon^{dm} = O(\ln^{(dm+1)\alpha}(1/\epsilon)\epsilon^{-dm}),$$

which establishes the proposition as m depends only on $|S|$. ■

We record one final result that asserts that if $S \subseteq G$ generates, then the powers of S individually become dense in G.

Proposition 5. *Suppose that G is a compact connected Lie group. Suppose that $S \subseteq G$ generates G. Then for all $\epsilon > 0$ there exists a natural number n_ϵ such that S^{n_ϵ} is ϵ-dense in G.*

Proof. Let $\{g_1, \ldots, g_m\}$ be an $\epsilon/2$-dense subset of G. Because S generates, for each g_i there exist n_i and $w_i \in S^{n_i}$ such that $d(g_i, w_i) < \epsilon/2$. By a pigeonhole argument similar to the proof of Lemma 4, for all $\epsilon > 0$ there exists a natural number N such that for all $n \geq N$, $d(S^n, e) < \epsilon$. Thus there exists N such that for all $n \geq N$, S^n contains elements within distance $\epsilon/2$ of the identity. Thus $S^{N + \max_i n_i}$ is ϵ-dense in G. ■

2.2. *Diophantine sets*

We will now introduce a notion of a Diophantine subset of a compact connected semisimple Lie group G. Write $\mathfrak{g}$ for the Lie algebra of G. We recall the definition of the standard quadratic Casimir element in $U(\mathfrak{g})$, the universal enveloping algebra of $\mathfrak{g}$. Write B for the Killing form on $\mathfrak{g}$ and let X_i be an orthonormal basis for $\mathfrak{g}$ with respect to B. We will also denote the inner product arising from the Killing form by $\langle \cdot, \cdot \rangle$. Then the *Casimir* element Ω is the element of $U(\mathfrak{g})$ defined by

$$\Omega = \sum_i X_i^2.$$

The element Ω is well-defined and central in $U(\mathfrak{g})$. Elements of $U(\mathfrak{g})$ act on the smooth vectors of representations of G. Consequently, as Ω is central and every vector in an irreducible representation (π, V) is smooth, $\pi(\Omega)$ acts by a multiple of the identity. Given an irreducible unitary representation (π, V), define $c(\pi)$ by

$$c(\pi)\,\mathrm{Id} = -\pi(\Omega). \tag{4}$$

The quantity $c(\pi)$ is positive in non-trivial representations. Further, as π ranges over all non-trivial representations, $c(\pi)$ is uniformly bounded away from 0. For further information see [29, §5.6].

Definition 6. Let G be a compact, connected, semisimple Lie group. We say that a subset $S \subset G$ is (C, α)-*Diophantine* if the following holds for each non-trivial, irreducible, finite dimensional unitary representation (π, V) of G. For all non-zero $v \in V$ there exists $g \in S$ such that

$$\|v - \pi(g)v\| \geq C c(\pi)^{-\alpha} \|v\|,$$

where $c(\pi)$ is defined in (4). We say that S is *Diophantine* if S is (C, α)-Diophantine for some $C, \alpha > 0$. If $(g_1, \dots, g_m)$ is a tuple of elements of G, then we say that this tuple is (C, α)-Diophantine if the underlying set is (C, α)-Diophantine.

Our definition of Diophantine is slightly different from the definition in [9] as we refer directly to irreducible representations. We choose this definition because it allows for a unified analysis of the action of Ω in diverse representations of G.

It is useful to compare Definition 6 with the simultaneous Diophantine condition used when studying translations on tori, as considered in [6] or [26]. The condition for tori is a generalization of the simultaneous Diophantine condition considered by Moser [25] for circle diffeomorphisms. Denote by $\langle \cdot, \cdot \rangle$ the standard inner product in $\mathbb{R}^d$. A tuple $(\theta_1, \dots, \theta_m)$ of vectors in $\mathbb{R}^d$ defines a tuple of translations of $\mathbb{T}^d$. We say that this tuple is (C, α)-Diophantine if for every non-zero $k \in \mathbb{Z}^d$,

$$\max_{1 \leq i \leq m} \min_{l \in \mathbb{Z}} |\langle \theta_i, k \rangle - l| \geq \frac{C}{\|k\|^\alpha}. \tag{5}$$

One can see the relationship between this definition and the one for compact semisimple groups when we think of $\mathbb{Z}^d$ as indexing the unitary representations of $\mathbb{T}^d$. Although these definitions apply to different types of groups, one can check that the estimates at their core are equivalent: for a given unitary representation defined by $k \in \mathbb{Z}^d$, use the θ_i that achieves the maximum in (5) to act on the representation defined by k.

We now give a useful characterization of Diophantine subsets of compact semisimple groups.

Proposition 7 ([9, Thm. A.3]). *Suppose that S is a finite subset of a compact connected semisimple Lie group G. Then S is Diophantine if and only if $\overline{\langle S\rangle} = G$. Moreover, there exists $\epsilon_0(G)$ such that any ϵ_0-dense subset of G is Diophantine.*

Before proceeding to the proof we will show two preliminary results.

Lemma 8. *Suppose that G is a compact connected semisimple Lie group. Suppose that (π, V) is an irreducible unitary representation of G. Then for any $v \in V$ of unit length, any $X \in \mathfrak{g}$ of unit length, and $t \geq 0$,*

$$\|\pi(\exp(tX))v - v\| \leq t\sqrt{c(\pi)}.$$

Proof. A similar argument to the following appears in [29, §5.7.13]. There exists an orthonormal basis $\{X_1, \dots, X_n\}$ of $\mathfrak{g}$ such that $X_1 = X$. Observe that

$$\pi(\exp(tX))v - v = t\,d\pi(X)v + O(t^2).$$

The transformation $d\pi(X)$ is skew-symmetric with respect to the inner product. Thus $d\pi(X)^2$ is positive semidefinite. Consequently,

$$\langle d\pi(X)v, d\pi(X)v\rangle = -\langle d\pi(X)^2 v, v\rangle \leq -\langle \pi(\Omega)v, v\rangle = c(\pi)\|v\|^2.$$

Hence

$$\|\pi(\exp(tX)v) - v\| \leq t\sqrt{c(\pi)} + O(t^2).$$

For $0 \leq i \leq n$, let $t_i = \frac{i}{n}t$. Then

$$\begin{aligned}\|\pi(\exp(tX))v - v\| &\leq \sum_{i=1}^{n} \|\pi(\exp(t_i X))v - \pi(\exp(t_{i-1}X))v\| \\ &\leq \sum_{i=1}^{n} \|\pi(\exp(tX/n))v - v\| \leq n\left(\frac{t}{n}\sqrt{c(\pi)} + O((t/n)^2)\right).\end{aligned}$$

Taking the liminf of the right hand side as $n \to \infty$ gives the result. ∎

The following lemma will be of use in the proof of Proposition 10.

Lemma 9. *Suppose that (π, V) is a non-trivial, irreducible, finite-dimensional, unitary representation of a compact connected semisimple group G. Then for any $v \in V$, there exists g such that $\langle \pi(g)v, v\rangle = 0$.*

Proof. If such a g does not exist, then for all $g \in G$, $\pi(g)v$ lies in the same half-space as v. But then $\int_G \pi(g)v\,dg \neq 0$ and is a G-invariant vector, which contradicts the irreducibility of π. ∎

Proposition 10. *Suppose that G is a compact connected semisimple Lie group. Then there exist $\epsilon_0, C, \alpha > 0$ such that any ϵ_0-dense subset of G is (C, α)-Diophantine. If S is a subset of G such that S^{n_0} is ϵ_0-dense in G, then S is $(C/n_0, \alpha)$-Diophantine.*

Proof. Let ϵ_0 equal the $\epsilon_0(G)$ in Theorem 2, the Solovay–Kitaev algorithm. If S is already ϵ_0-dense, let $n_0 = 1$. By Theorem 2, there exist C and α such that for each ϵ there exists $l_\epsilon \leq C \ln^\alpha(\epsilon^{-1})$ such that $S^{n_0 l_\epsilon}$ is ϵ-dense in G. Suppose that (π, V) is a non-trivial irreducible unitary representation of G and suppose that $v \in V$ is a unit vector. By Lemma 9 there exists $g \in G$ such that $\langle \pi(g)v, v\rangle = 0$. Now fix $\epsilon = 1/(100\sqrt{c(\pi)})$. Then there exists an element $w \in S^{n_0 l_\epsilon}$ such that $d(g, w) < \epsilon$. Thus by Lemma 8,

$$\|\pi(g)v - \pi(w)v\| \leq \epsilon\sqrt{c(\pi)} < \frac{1}{100}.$$

By the triangle inequality, this implies that

$$\|\pi(w)v - v\| \geq 1.$$

Write $w = g_1^{\sigma_1} \cdots g_{n_0 l_\epsilon}^{\sigma_{n_0 l_\epsilon}}$ where each $\sigma_i \in \{\pm 1\}$ and each $g_i \in S$. Let $w_i = g_1^{\sigma_1} \cdots g_i^{\sigma_i}$. Let $w_0 = e$. By applying the triangle inequality $n_0 l_\epsilon$ times, we see that

$$\sum_{i=0}^{n_0 l_\epsilon - 1} \|\pi(w_i)v - \pi(w_{i+1})v\| \geq \|v - \pi(w)v\| \geq 1.$$

Thus there exists some i such that

$$\|\pi(w_i)v - \pi(w_{i+1})v\| \geq \frac{1}{n_0 l_\epsilon}.$$

Applying $\pi(w_i^{-1})$ and noting by our choice of ϵ that $l_\epsilon \leq C \ln^\alpha(c(\pi))$, we obtain

$$\|v - \pi(g_i^{\sigma_i})v\| \geq \frac{1}{n_0 C' \ln^\alpha(c(\pi))}. \tag{6}$$

Thus we are done as we have obtained an estimate that is stronger than the required lower bound of $C/c(\pi)^\alpha$. ∎

We now prove the equivalence of the Diophantine property appearing in Proposition 10 with that in Definition 6.

Proof of Proposition 7. To begin, suppose that S is Diophantine. For the sake of contradiction, suppose that $H := \overline{\langle S\rangle} \neq G$. Consider the action of G on $L^2(G/H)$ by left translation. Note that H acts trivially. However, $L^2(G/H)$ contains non-trivial representations of G. Thus $S \subset H$ cannot be Diophantine, which is a contradiction.

For the other direction, suppose that $\overline{\langle S\rangle} = G$. Then by Proposition 5 there exists n such that S^n is $\epsilon_0(G)$-dense, and hence S is Diophantine by Proposition 10. ∎

The stronger bound in (6) gives an equivalent characterization of Diophantineness.

Corollary 11. *Let G be a compact connected semisimple Lie group. A subset S of G is Diophantine if and only if there exist $C, \alpha > 0$ such that the following holds for each non-trivial, irreducible, finite-dimensional, unitary representation (π, V) of G. For all $v \in V$ there exists $g \in S$ such that*

$$\|v - \pi(g)v\| \geq \frac{\|v\|}{C \ln^{\alpha}(c(\pi))}.$$

Diophantine subsets of a group are typical in the following sense.

Proposition 12. *Suppose that G is a compact connected semisimple Lie group. Let $U \subset G \times G$ be the set of ordered pairs (u_1, u_2) such that $\{u_1, u_2\}$ is a Diophantine subset of G. Then U is Zariski open and hence open and dense in the manifold topology on $G \times G$.*

Proof. Let $U \subset G \times G$ be the set of points (u_1, u_2) such that $\{u_1, u_2\}$ generates a dense subset of G. Theorem 1.1 in [12] shows that U is Zariski open and non-empty. By Proposition 7, this implies that $\{u_1, u_2\}$ is Diophantine. As U is non-empty, the final claim follows. ■

2.3. Polylogarithmic spectral gap

In this subsection, we study spectral properties of an averaging operator associated to a tuple of elements of G. Consider a tuple $(g_1, \ldots, g_m)$ of elements of G. Let $\mathbb{R}[G]$ denote the group ring of G over $\mathbb{R}$. From this tuple we form $\mathcal{L} := (g_1 + \cdots + g_m)/m \in \mathbb{R}[G]$. The element $\mathcal{L}$ acts in representations of G in the natural way. If (π, V) is a representation of G, then we write $\mathcal{L}_\pi$ for the action of $\mathcal{L}$ on V. The main result of this subsection is the following proposition, which gives some spectral properties of $\mathcal{L}_\pi$ under the assumption that $\{g_1, \ldots, g_m\}$ is Diophantine.

Proposition 13. *Let G be a compact connected semisimple Lie group, $(g_1, \ldots, g_m)$ a tuple of elements of G, and suppose that $\{g_1, \ldots, g_m\}$ generates G. Then there exists a neighborhood N of $(g_1, \ldots, g_m)$ in $G \times \cdots \times G$ and constants $D_1, D_2, \alpha > 0$ such that if $(g'_1, \ldots, g'_m) \in N$, then $\{g'_1, \ldots, g'_m\}$ is Diophantine and its associated averaging operator $\mathcal{L}$ satisfies*

$$\|\mathcal{L}_\pi^n\| \leq D_1 \left(1 - \frac{1}{D_2 \ln^{\alpha}(c(\pi))}\right)^n$$

for each non-trivial irreducible unitary representation (π, V).

The proof of Proposition 13 uses the following lemma, which is a sharpening the triangle inequality for vectors that are not colinear.

Lemma 14. *Suppose that v, w are two vectors in an inner product space. Suppose that $\|v\| \leq \|w\|$ and let $\hat{v} = v/\|v\|$ and $\hat{w} = w/\|w\|$. If*

$$\|\hat{v} - \hat{w}\| \geq \epsilon,$$

then

$$\|v + w\| \leq (1 - \epsilon^2/10)\|v\| + \|w\|.$$

Proof. We begin by considering the following estimate for unit vectors.

Claim 1. *Suppose that the angle between two unit vectors $\hat{v}$ and $\hat{w}$ is $\theta \in [0, \pi]$. Then*

$$\|\hat{v} + w\| \leq \|\hat{v}\| + (1 - \theta^2/10)\|\hat{w}\|.$$

Proof. It suffices to consider the two vectors $\hat{v} = (1, 0)$ and $\hat{w} = (\cos\theta, \sin\theta)$ in $\mathbb{R}^2$. We have to show

$$\|\hat{v} + \hat{w}\|^2 \leq \left(\|\hat{v}\| + \left(1 - \frac{\theta^2}{10}\right)\|\hat{w}\|\right)^2.$$

From the definitions,

$$\|\hat{v} + \hat{w}\|^2 = 2 + 2\cos\theta$$

and

$$\left(\|\hat{v}\| + \left(1 - \frac{\theta^2}{10}\right)\|\hat{w}\|\right)^2 = 4 - 4\frac{\theta^2}{10} + \frac{\theta^4}{100} \geq 4 - 4\frac{\theta^2}{10}.$$

Thus it suffices to show for $\theta \in [0, \pi]$ that

$$2 + 2\cos\theta \leq 4 - 4\frac{\theta^2}{10},$$

which follows because for $\theta \in [0, \pi]$ we have the estimate $\cos\theta \leq 1 - \theta^2/5$. ∎

We may prove the lemma once we have one more observation. Note that if $\hat{v}$ and $\hat{w}$ are two unit vectors, then $\|\hat{v} - \hat{w}\| = \epsilon$ is less than the angle θ between $\hat{v}$ and $\hat{w}$ because the distance between $\hat{v}$ and $\hat{w}$ along a unit circle they lie on is precisely θ. Thus we see that $\epsilon \leq \theta$ for $0 \leq \theta \leq \pi$.

We now compute. Note that without loss of generality we may assume that $\|w\| = 1$, which we do in the following. By the triangle inequality,

$$\|v + w\| \leq \|v\| \, \|\hat{v} + \hat{w}\| + (1 - \|v\|)\|\hat{w}\|.$$

By the claim it then follows that

$$\|v + w\| \leq \|v\|((1 - \theta^2)\|\hat{v}\| + \|\hat{w}\|) + (1 - \|v\|)\|\hat{w}\|.$$

Noting that $0 \leq \epsilon \leq \theta$ for $\theta \in [0, \pi]$, we then conclude:

$$\begin{aligned}\|v + w\| &\leq \|v\|((1 - \epsilon^2/10)\|\hat{v}\| + \|\hat{w}\|) + (1 - \|v\|)\|\hat{w}\| \\ &= (1 - \epsilon^2/10)\|v\| + \|w\|.\end{aligned}$$ ∎

Proof of Proposition 13. For convenience, let $W = (g_1, \ldots, g_m)$ and let $S = \{g_1, \ldots, g_m\}$. Let $\epsilon_0(G)$ be as in Proposition 10. By Proposition 5, because $\overline{\langle S \rangle} = G$ there exists some n_0 such that S^{n_0} is $\epsilon_0/2$-dense in G. Then let N be the neighborhood of $(g_1, \ldots, g_m)$ in $G \times \cdots \times G$ such that if $p = (g_1', \ldots, g_m') \in N$ then $\{g_1', \ldots, g_m'\}^{n_0}$ is at least ϵ_0-dense in G. It now suffices to obtain the given estimate for the set $W = (g_1, \ldots, g_m)$ using only the assumption that S^{n_0} is ϵ_0-dense. Below, W^{n_0} is the tuple of the m^{n_0} words of length n_0 with entries in W.

By Proposition 10, there exist (C,α) such that any ϵ_0-dense set is (C,α)-Diophantine. As S^{n_0} is ϵ_0-dense, so is $S^{n_0}S^{-n_0}$, and hence $S^{n_0}S^{-n_0}$ is (C,α)-Diophantine.

Consider now a non-trivial irreducible finite-dimensional unitary representation (π, V) of G. Since $S^{n_0}S^{-n_0}$ is (C,α)-Diophantine, Corollary 11 implies that for any unit length $v \in V$ there exist $w_1, w_2 \in S^{n_0}$ such that

$$\|v - \pi(w_1^{-1}w_2)v\| \geq \frac{1}{C\ln^{\alpha}(c(\pi))},$$

and so

$$\|\pi(w_1)v - \pi(w_2)v\| \geq \frac{1}{C\ln^{\alpha}(c(\pi))}.$$

Hence by Lemma 14, since π is unitary

$$\begin{aligned}\|\pi(w_1)v + \pi(w_2)v\| &\leq \left(1 - \frac{1}{10C^2\ln^{2\alpha}(c(\pi))}\right)\|\pi(w_1)v\| + \|\pi(w_2)v\| \\ &\leq \left(2 - \frac{1}{10C^2\ln^{2\alpha}(c(\pi))}\right)\|v\|.\end{aligned}$$

Then by the triangle inequality,

$$\begin{aligned}\|\mathcal{L}_\pi^{n_0}v\| &= \left\|\frac{1}{|W|^{n_0}}\sum_{w\in W^{n_0}}\pi(w)v\right\| \\ &\leq \frac{1}{|W|^{n_0}}\Big(\|\pi(w_1)v + \pi(w_2)v\| + \sum_{w\in W^{n_0}\setminus\{w_1,w_2\}}\|\pi(w)v\|\Big) \\ &\leq \frac{1}{|W|^{n_0}}\left(2 - \frac{1}{10C^2\ln^{2\alpha}(c(\pi))}\right)\|v\| + \frac{|W|^{n_0}-2}{|W|^{n_0}}\|v\| \\ &\leq \left(1 - \frac{1}{10C^2|W^{n_0}|\ln^{2\alpha}(c(\pi))}\right)\|v\|.\end{aligned}$$

Interpolating shows that for all $n \geq 0$,

$$\|\mathcal{L}_\pi^n\| \leq \left(1 - \frac{1}{10C^2|W^{n_0}|\ln^{2\alpha}(c(\pi))}\right)^{-1}\left(1 - \frac{1}{10C^2|W^{n_0}|\ln^{2\alpha}(c(\pi))}\right)^{n/n_0}.$$

As (π, V) ranges over all non-trivial representations, $c(\pi)$ is uniformly bounded away from 0; see [29, 5.6.7]. This implies that the first term above is uniformly bounded by some $D > 0$ independent of π. Applying the estimate $(1+x)^{\epsilon} \leq 1 + \epsilon x$ to the second term then gives the proposition. ■

Notice that in Proposition 13 we obtain an entire neighborhood of our initial set S on which we have the same estimates for $\mathcal{L}_\pi$. Consequently, because these estimates remain true under small perturbations, we think of them as being stable. We will use the term "stable" in the following precise sense.

Definition 15. Suppose that T is some property of a tuple $W = (g_1, \dots, g_m)$ with elements in a Lie group G. We say that T is *stable* at $W = (g_1, \dots, g_m)$ if there exists a neighborhood N of $(g_1, \dots, g_m)$ in $G \times \cdots \times G$ such that if $(g'_1, \dots, g'_m) \in N$ then T holds for $(g'_1, \dots, g'_m)$. We will also say that T is *stable* without reference to a subset when the relevant tuples that T is stable on are evident.

A crucial aspect of the Diophantine property in compact semisimple Lie groups is that by Proposition 10 there is a stable lower bound on (C, α). This stability will be essential during the KAM scheme.

2.4. Diophantine sets and tameness

Consider a smooth vector bundle E over a closed manifold M. We may consider the space $C^\infty(M, E)$ of smooth sections of E. Consider a linear map $L\colon C^\infty(M, E) \to C^\infty(M, E)$. We say that L is *tame* if there exists α such that for all k there exists C_k such that for all $s \in C^\infty(M, E)$,

$$\|Ls\|_{C^k} \le C_k \|s\|_{C^{k+\alpha}}.$$

See [16, §II.2.1] for more about tameness. The main result of this section is to show such estimates for certain operators related to $\mathcal{L}$.

Though $\mathcal{L}$ acts in any representation of G, we are most interested in the action of G on the sections of certain vector bundles, which we now describe. Suppose that K is a closed subgroup of G and that E is a smooth vector bundle over G/K. We say that E is a *homogeneous vector bundle* over G/K if G acts on E by bundle maps and this action projects to the action of G on G/K by left translation. We now give an explicit description of all homogeneous vector bundles over G/K via the Borel construction. See [29, Ch. 5] for more details about this topic and what follows. Suppose that (τ, E_0) is a finite-dimensional unitary representation of K. Form the trivial bundle $G \times E_0$. Then K acts on this bundle by $(g, v) \mapsto (gk, \tau(k)^{-1}v)$. Then $(G \times E_0)/K$ is a vector bundle over G/K that we denote by $G \times_\tau E_0$. Note, for instance, that $C^\infty(G, \mathbb{R})$ is the space of sections of the homogeneous vector bundle obtained from the trivial representation of $\{e\} < G$. The left action of G on $G \times E_0$ descends to $G \times_\tau E_0$, and hence this is a homogeneous vector bundle.

In order to do analysis in a homogeneous vector bundle, we must introduce some additional structures. Suppose that $E = G \times_\tau E_0$ is a homogeneous vector bundle. The base G/K comes equipped with the projection of the Haar measure on G. As the action of K on $G \times E_0$ is isometric on fibers, the fibers of E are naturally endowed with an inner product. We may then consider the space $L^2(E)$ of all L^2 sections of E. In addition, we will write $C^\infty(E)$ for the space of all smooth sections of E. The action of G on E preserves $L^2(E)$ and $C^\infty(E)$.

We recall briefly how one may do harmonic analysis on sections of such bundles. As before, let Ω be the Casimir operator, which is an element of $U(\mathfrak{g})$. Then Ω acts on the C^∞ vectors of any representation of G. Denote by Δ the differential operator obtained by the action of $-\Omega$ on $C^\infty(E)$. Then Δ is a hypoelliptic differential operator on E. We then use the spectrum of Δ to define for any $s \geq 0$ the Sobolev norm H^s in the following manner. $L^2(E)$ may be decomposed as the Hilbert space direct sum of finite-dimensional irreducible unitary representations V_π. Write $\phi = \sum_\pi \phi_\pi$ for the decomposition of an element $\phi \in L^2(E)$. Then the s-Sobolev norm is defined by

$$\|\phi\|_{H^s}^2 = \sum_\pi (1 + c(\pi))^s \|\phi_\pi\|_{L^2}^2.$$

We write $\|f\|_{C^s}$ for the usual C^s norm of a function or section of a vector bundle. It is not always necessary to work with the decomposition of $L^2(E)$ into irreducible subspaces; instead one can use a coarser decomposition as follows. We let H_λ denote the subspace of $L^2(E)$ on which Δ acts by multiplication by $\lambda > 0$. There are countably many such subspaces H_λ and each is finite-dimensional. In what follows, those functions that are orthogonal to the trivial representations in $L^2(E)$ will be of particular importance. We denote by $L_0^2(E)$ the orthogonal complement of the trivial representations in $L^2(E)$, and $C_0^\infty(E)$ the subspace $L_0^2(E) \cap C^\infty(E)$.

We now consider the action of $\mathcal{L}$ on the sections of a homogeneous vector bundle.

Proposition 16 (Tameness [10, Prop. 1]). *Suppose that $(g_1, \dots, g_m)$ is a Diophantine tuple with elements in a compact connected semisimple Lie group G. Suppose that E is a homogeneous vector bundle that G acts on. Then there exist constants $C_1, \alpha_1, \alpha_2 > 0$ such that for any $s \geq 0$ there exists C_s such that for any non-zero $\phi \in C_0^\infty(G/K, E)$,*

$$\|(\mathrm{Id} - \mathcal{L})^{-1}\phi\|_{H^s} \leq C_1 \|\phi\|_{H^{s+\alpha_1}}, \quad \|(\mathrm{Id} - \mathcal{L})^{-1}\phi\|_{C^s} \leq C_s \|\phi\|_{C^{s+\alpha_2}}.$$

Moreover, these estimates are stable.

Proof. As before, let H_λ be the λ-eigenspace of Δ acting on sections of E. Let $\mathcal{L}_\lambda$ denote the action of $\mathcal{L}$ on H_λ. From Proposition 13, we see that there exist D_1, D_2 and α_3 such that for all $\lambda > 0$, $\|\mathcal{L}_\lambda^n\|_{H^0} \leq D_1(1 - 1/(D_2 \ln^{\alpha_3}(\lambda\nu)))^n$. Thus there exists C_3 such that $\|(\mathrm{Id} - \mathcal{L}_\lambda)^{-1}\|_{H^0} \leq C_3 \ln^{\alpha_3}(\lambda)$. Now observe that in the following sum $\lambda \neq 0$ by our assumption that ϕ is orthogonal to the trivial representations contained in $L^2(E)$:

$$\begin{aligned}
\|(\mathrm{Id} - \mathcal{L})^{-1}\phi\|_{H^s}^2 &= \sum_{\lambda>0} (1+\lambda)^s \|(\mathrm{Id} - \mathcal{L}_\lambda)^{-1}\phi_\lambda\|_{L^2}^2 \\
&\leq \sum_{\lambda>0} (1+\lambda)^s \|(\mathrm{Id} - \mathcal{L}_\lambda)^{-1}\|^2 \|\phi_\lambda\|_{L^2}^2 \\
&\leq \sum_{\lambda>0} C_3^2 \ln^{2\alpha_3}(\lambda)(1+\lambda)^s \|\phi_\lambda\|_{L^2}^2 \\
&\leq \sum_{\lambda>0} C_4^2 (1+\lambda)^{s+\alpha_1} \|\phi_\lambda\|_{L^2}^2 \\
&\leq C_4^2 \|\phi\|_{H^{s+\alpha_1}}^2,
\end{aligned}$$

for any $\alpha_1 > 0$ and sufficiently large C_4. The second estimate in the proposition then follows from the first by applying the Sobolev embedding theorem. ∎

2.5. Application to isotropic manifolds

We now introduce the class of isotropic manifolds, which are the subject of this paper and whose isometry groups may be studied along the above lines. We say that M is *isotropic* if $\operatorname{Isom}(M)$ acts transitively on the unit tangent bundle of M, T^1M. This is equivalent to $\operatorname{Isom}(M)^\circ$ acting transitively on T^1M. There are not many isotropic manifolds. In fact, all are globally symmetric spaces. The following is the complete list of all compact isotropic manifolds:

(1) $S^n = \mathrm{SO}(n+1)/\,\mathrm{SO}(n)$, sphere,

(2) $\mathbb{R}\,\mathrm{P}^n = \mathrm{SO}(n+1)/\mathrm{O}(n)$, real projective space,

(3) $\mathbb{C}\,\mathrm{P}^n = \mathrm{SU}(n+1)/\mathrm{U}(n)$, complex projective space,

(4) $\mathbb{H}\,\mathrm{P}^n = \mathrm{Sp}(n+1)/(\mathrm{Sp}(n)\times\mathrm{Sp}(1))$, quaternionic projective space,

(5) $F_4/\mathrm{Spin}(9)$, Cayley projective plane.

A proof of this classification may be found in [30, Thm. 8.12.2].

Though S^1 is an isotropic manifold, we will exclude it in all future statements because its isometry group is not semisimple. The reason that we study isotropic manifolds is that if M is an isotropic manifold other than S^1, then $\operatorname{Isom}(M)$ is semisimple.

Lemma 17. *Suppose that M is a compact connected isotropic manifold other than S^1. Then* $\operatorname{Isom}(M)$ *is semisimple. The same is true for* $\operatorname{Isom}^0(M)$*, the connected component of the identity.*

For a proof of this lemma, see [28], which computes the isometry groups for each of these spaces explicitly. In fact, these isometry groups all have simple Lie algebras.

One minor issue with applying what we have developed so far to isotropic manifolds is that $\operatorname{Isom}(M)$ need not be connected. Even in the case of S^2, $\operatorname{Isom}(M)$ is disconnected. In fact, Dolgopyat and Krikorian assume that the isometries in their theorem all lie in the identity component of $\operatorname{Isom}(M)$ and hence are rotations. Here, we consider the full isometry group. Hence Theorem 1 is a generalization even in the case of S^n. That said, the generalization is minor: the identity component is index 2 in the full isometry group.

Although connectedness of $\operatorname{Isom}(M)$ has not been the crux of previous arguments, if $\operatorname{Isom}(M)\neq\operatorname{Isom}(M)^\circ$, then there are "extra" representations of $\operatorname{Isom}(M)$ that appear in the definition of Diophantineness that would need to be dealt with slightly differently. For this reason we give the following definition, which is adapted to the case where $\operatorname{Isom}(M)$ is not connected.

Definition 18. We say that a tuple $(g_1,\dots,g_m)$ with each $g_i\in\operatorname{Isom}(M)$ is *Diophantine* if there exists n such that if $S=\{g_1,\dots,g_m\}$ then $S^n\cap\operatorname{Isom}(M)^\circ$ is (C,α)-Diophantine for some $C,\alpha>0$. We say that such a tuple is (C,α,n)-*Diophantine*.

It follows from Proposition 7 that if a tuple is Diophantine, then there exists a neighborhood of that tuple such that the constants C,α,n may be taken to be uniform over that neighborhood. Thus Diophantineness in this more general sense is a stable property. The following analogue of Proposition 19 is then immediate.

Proposition 19. *Let M be a closed isotropic manifold of dimension at least 2 and S be a finite subset of* $\operatorname{Isom}(M)$. *The set S is Diophantine if and only if* $\operatorname{Isom}(M)^\circ \subseteq \overline{\langle S\rangle}$. *Moreover, there exist $\epsilon_0(M), C, \alpha, n > 0$ such that any subset of* $\operatorname{Isom}(M)$ *that is ϵ_0-dense in* $\operatorname{Isom}(M)^\circ$ *is stably (C, α, n)-Diophantine.*

We will show a tameness result in this setting. The important point is that $\operatorname{Isom}(M)^\circ$ is a semisimple connected Lie group and TM is a homogeneous vector bundle that $\operatorname{Isom}(M)^\circ$ acts on. Further, due to M being isotropic, $L^2(M, TM)$ contains no trivial representations of $\operatorname{Isom}(M)^\circ$. Thus we are almost in a position where we can apply Proposition 16. There is one small issue: there may be representations of $\operatorname{Isom}(M)$ that are trivial on $\operatorname{Isom}(M)^\circ$ and hence the previous arguments do not apply directly to these representations. However, for the purpose of studying sections of TM, studying representations of $\operatorname{Isom}(M)^\circ$ suffices. The following proposition explains how one may get around this issue to recover the appropriate analog of Proposition 13. It is important to note that there are many choices of a "Laplacian" acting on vector fields over a manifold, and they may not all be the same. In our case, we are choosing to work with the Casimir Laplacian, which arises from viewing TM as a homogeneous vector bundle. Given a tuple $(g_1, \dots, g_m)$ of isometries of M, the associated operator $\mathcal{L}$ that acts on $L^2(M, TM)$ is defined for a vector field V by $V \mapsto m^{-1} \sum_{i=1}^m (Dg_i)_* V$.

Proposition 20. *Suppose that M is a closed isotropic manifold with* $\dim M \geq 2$. *Suppose that $(g_1, \dots, g_m)$ is a Diophantine tuple with elements in* $\operatorname{Isom}(M)$. *There exists a neighborhood $\mathcal{N}$ of $(g_1, \dots, g_m)$ in* $\operatorname{Isom}(M) \times \cdots \times \operatorname{Isom}(M)$ *and constants $D_1, D_2, \alpha > 0$ such that if $(g_1', \dots, g_m') \in \mathcal{N}$, then $\{g_1', \dots, g_m'\}$ is Diophantine. Let H_λ denote the λ-eigenspace of Δ acting on sections of TM. For any tuple in this neighborhood, the associated operator $\mathcal{L}$ acts on $L^2(M, TM)$ and preserves the H_λ-eigenspaces. In fact, writing $\mathcal{L}_\lambda$ for this induced action we have*

$$\|\mathcal{L}_\lambda^n\| \leq D_1 \left(1 - \frac{1}{D_2 \ln^\alpha(\lambda)}\right)^n.$$

The same holds for the eigenspaces H_λ of Δ acting on other bundles over M assuming that $\operatorname{Isom}(M)$ *acts isometrically on the space of sections of those bundles. In cases where there is a trivial representation, we must also assume $\lambda > 0$. Examples of such bundles are $L^2(M, \mathbb{R})$ as well as $L^2(\operatorname{Gr}_r(M), \mathbb{R})$ when* $\operatorname{Isom}(M)^\circ$ *acts transitively on the r-planes in TM.*

Proof. The key steps in the proof are substantially similar to those in Proposition 13, once we show that the elements of $\operatorname{Isom}(M)$ all preserve the spaces H_λ. Let Γ be a bundle as in the statement of the proposition that $\operatorname{Isom}(M)$ acts on isometrically.

Claim 2. *Suppose that $V \subset \Gamma$ is an irreducible representation of* $\operatorname{Isom}(M)^\circ$ *isomorphic to (π, W). Then for any $k \in \operatorname{Isom}(M)^\circ$, kV is an irreducible representation of V isomorphic to $(\pi \circ \alpha, W)$ for some automorphism α of* $\operatorname{Isom}(M)^\circ$. *In particular, $c(\pi \circ \alpha) = c(\pi)$.*

Proof. Let $g^k = k^{-1}gk$ as usual. We claim that for any $k \in \mathrm{Isom}(M)$, kV is a representation of $\mathrm{Isom}(M)^\circ$. To see this, note that for $v \in V$, we have $gkv = kg^k v$, but $g^k \in \mathrm{Isom}(M)^\circ$, so $kg^k v \in kV$. Moreover, it is straightforward to see that the representation of $\mathrm{Isom}(M)^\circ$ on kV is isomorphic to the representation $(\pi \circ \alpha, W)$ where α is the automorphism $g \mapsto g^k$.

We now claim that $c(\pi \circ \alpha) = c(\pi)$. Because α is an automorphism, it preserves the Killing form, and hence we can write the Casimir element as $\sum_i (d\alpha^{-1}(X_i))^2$. Now note that if one traces through the computation of the value $c(\pi \circ \alpha)$ for the representation $\pi \circ \alpha$, then the α^{-1} we have introduced cancels with the α. Thus the computation reduces to the computation of $c(\pi)$ with the original expression $\sum_i X_i^2$. Hence $c(\pi \circ \alpha) = c(\pi)$. ∎

To conclude from this point, one uses the same argument as in Proposition 13, except we start with the set S^{n_0} and only make use of the elements in $S^{n_0} \cap \mathrm{Isom}(M)^\circ$. No issues arise because any terms that do not lie in $\mathrm{Isom}(M)^\circ$ are isometries of H_λ as we have now shown. ∎

Having established the previous proposition the following is immediate and may be shown by repeating the argument of Proposition 16.

Proposition 21. *Suppose that M is a closed isotropic manifold with* $\dim M \geq 2$. *Suppose that $(g_1, \dots, g_m)$ is a Diophantine tuple with elements in* $\mathrm{Isom}(M)$. *There exist constants $C_1, \alpha_1, \alpha_2 > 0$ such that for any $s \geq 0$ there exists C_s such that for any $\phi \in C^\infty(M, TM)$,*

$$\|(\mathrm{Id} - \mathcal{L})^{-1}\phi\|_{H^s} \leq C_1\|\phi\|_{H^{s+\alpha_1}}, \quad \|(\mathrm{Id} - \mathcal{L})^{-1}\phi\|_{C^s} \leq C_s\|\phi\|_{C^{s+\alpha_2}}.$$

Moreover, these estimates are stable. The same holds for the action of $\mathcal{L}$ on any of the sections of any of the bundles that Proposition 20 applies to.

3. Approximation of stationary measures

In this section, we introduce the notion of a stationary measure associated to a random dynamical system. We consider stationary measures of certain random dynamical systems associated to a Diophantine subset of a compact semisimple Lie group as well as perturbations of these systems. We begin by introducing these systems and some associated transfer operators. In Proposition 23, we give an asymptotic expansion of the stationary measures of a perturbation.

3.1. Random dynamical systems and their transfer operators

We now give some basic definitions concerning random dynamical systems. For general treatments of random dynamical systems and their basic properties, see [20] or [1]. If $(f_1, \dots, f_m)$ is a tuple of maps of a standard Borel space M, then these maps generate a uniform Bernoulli random dynamical system on M. This dynamical system is given by choosing an index $1 \le i \le m$ uniformly at random and then applying the function f_i to M. To iterate the system further, one chooses additional independent uniformly distributed indices and repeats. We always use the words *random dynamical system* to mean uniform Bernoulli random dynamical system in the sense just described.

Associated to this random dynamical system are two operators. The first operator is called the *averaged Koopman operator*. It acts on functions and is defined by

$$\mathcal{M}\phi := \frac{1}{m}\sum_{i=1}^{m} \phi \circ f_i. \tag{7}$$

The second operator is called the *averaged transfer operator*. It acts on measures and is defined by

$$\mathcal{M}^*\mu := \frac{1}{m}\sum_{i=1}^{m} (f_i)_*\mu. \tag{8}$$

Depending on the space M, we may restrict the domains of these operators to a suitable subset of the spaces of functions and measures on M. We say that a measure is *stationary* if $\mathcal{M}^*\mu = \mu$. We assume that stationary measures have unit mass.

In this paper, we take M to be a compact homogeneous space G/K. If $g \in G$, then left translation by g gives an isometry of G/K that we also call g. As before, a tuple $(g_1, \dots, g_m)$ with each g_i in G generates a random dynamical system on G/K. We will also consider perturbations of this random dynamical system. Consider a tuple $(f_1, \dots, f_m)$ where each f_i is in $\mathrm{Diff}^\infty(G/K)$. This collection also generates a random dynamical system on G/K. The indices $1, \dots, m$ give a natural way to compare the two systems. We refer to the initial system as *homogeneous* or *linear* and to the latter system as *non-homogeneous* or *non-linear*.

We will simultaneously work with a homogeneous and non-homogeneous systems, so we now introduce notation to distinguish the transfer operators of each. We write $\mathcal{M}$ for the averaged Koopman operator associated to the system generated by $(g_1, \dots, g_m)$, and we write $\mathcal{M}_\epsilon$ for the averaged Koopman operator associated to $(f_1, \dots, f_m)$. Analogously we use the notation $\mathcal{M}^*$ and $\mathcal{M}_\epsilon^*$.

Later we will compare the homogeneous system given by a tuple $(g_1, \dots, g_m)$ and a non-homogeneous perturbation $(f_1, \dots, f_m)$. We thus introduce the notation

$$\varepsilon_k := \max_i d_{C^k}(f_i, g_i), \tag{9}$$

to describe how large a perturbation is. In addition, it will be useful to have a linearization of the difference between f_i and g_i. The standard way to do this is via a chart on the Fréchet manifold $\mathrm{Diff}^\infty(G/K)$. If $d_{C^0}(f_i, g_i) < \mathrm{inj}(G/K)$, then we associate with f_i the vector field Y_i defined at $g_i(x) \in G/K$ by

$$Y_i(g_i(x)) := \exp_{g_i(x)}^{-1} f_i(x), \tag{10}$$

where we choose the minimum length preimage of $f_i(x)$ in $T_{g_i(x)}(G/K)$ under the map $\exp_{g_i(x)}^{-1}$. In addition, if Y is a vector field on M, then we define $\psi_Y : M \to M$ to be the map

$$\psi_Y : x \mapsto \exp_x(Y(x)). \tag{11}$$

The following theorem asserts the existence of Lyapunov exponents for random dynamical systems.

Theorem 22 ([20, Ch. 3, Thm. 1.1]). *Suppose that E is a measurable vector bundle over a Borel space M. Suppose that $F_1, F_2, \dots$ is a sequence of independent and identically distributed bundle maps of E with common distribution ν and suppose that ν has finite support. Suppose that μ is an ergodic ν-stationary measure on M for the random dynamics on M induced by those on E.*

Then there exists a list of numbers, the Lyapunov exponents,

$$-\infty < \lambda^s < \lambda^{s-1} < \dots < \lambda^1 < \infty,$$

such that for μ-a.e. $x \in M$ and almost every realization of the sequence, there exists a filtration of linear subspaces

$$0 \subset V^s \subset \dots \subset V^1 \subset E_x$$

such that, for that particular realization of the sequence, if $\xi \in V^{i+1} \setminus V^i$, where $V^i \equiv \{0\}$ for $i > s$, then

$$\lim_{n\to\infty} \frac{1}{n} \ln \|F^n \circ \dots \circ F^1 \xi\| = \lambda^i.$$

3.2. Approximation of stationary measures

Let dm denote the pushforward of Haar measure to G/K. Note that Haar measure is stationary for the homogeneous random dynamical system given by $(g_1, \dots, g_m)$. The following proposition compares the integral against a stationary measure μ for a perturbation $(f_1, \dots, f_m)$ and the Haar measure. Up to higher order terms, the difference between integrating against Haar and against μ is given by the integral of a particular function $\mathcal{U}(\phi)$. We obtain an explicit expression for $\mathcal{U}(\phi)$, which is useful because we can tell when $\mathcal{U}(\phi)$ vanishes and thus when μ is near to Haar. Compare the following with [10, Prop. 2].

Proposition 23. *Suppose that $S = (g_1, \dots, g_m)$ is a Diophantine tuple with elements in a compact connected semisimple group G or elements in* $\operatorname{Isom}(M)$ *for an isotropic manifold M with* $\dim M \geq 2$. *Let G/K be a quotient of G in the former case or a space* $\operatorname{Isom}(M)^\circ$ *acts transitively on in the latter. There exist constants k and C such that if $(f_1, \dots, f_m)$ is a tuple with elements in* $\operatorname{Diff}^\infty(G/K)$ *with $\varepsilon_0 = \max_i d_{C^0}(f_i, g_i) <$* $\operatorname{inj}(G/K)$*, then the following holds for each stationary measure μ for the uniform Bernoulli random dynamical system generated by the f_i. Let $Y_i = \exp^{-1}_{g_i(x)} f_i(x)$. Then for any $\phi \in C^\infty(G/K)$, we have*

$$\int_{G/K} \phi \, d\mu = \int_{G/K} \phi \, dm + \int_{G/K} \mathcal{U}(\phi) \, dm + O(\varepsilon_k^2 \|\phi\|_{C^k}), \tag{12}$$

where dm denotes the normalized pushforward of Haar measure to G/K and

$$\mathcal{U}(\phi) := \frac{1}{m}\sum_{i=1}^{m} \nabla_{Y_i} (\mathrm{Id} - \mathcal{M})^{-1}\Big(\phi - \int \phi \, dm\Big). \tag{13}$$

Moreover,

$$\left|\int \mathcal{U}(\phi)\, dm\right| \le C\,\|\phi\|_{C^k} \left\|\sum_{i=1}^{m} Y_i\right\|_{C^k}, \tag{14}$$

and the constants, including the constant in the big-O in (12), are stable in S.

Proof. The proof is similar to the proof of [24, Prop. 4]. We write the proof for the connected group G; the proof for $\mathrm{Isom}(M)$ is identical with the use of Proposition 21 instead of Proposition 16.

Note that a smooth real-valued function defined on G/K is naturally viewed as a section of the trivial bundle over G/K. If we view the averaged Koopman operator $\mathcal{M}$ associated to $(g_1, \dots, g_m)$ as acting on the sections of the trivial bundle $G/K \times \mathbb{R}$, then $\mathcal{M}$ satisfies the hypotheses of Proposition 16. Thus there exists α and constants C_s such that for any $\phi \in C_0^\infty(G/K)$, the space of integral 0 smooth functions on G/K,

$$\|(\mathrm{Id} - \mathcal{M})^{-1}\phi\|_{C^s} \le C_s \|\phi\|_{C^{s+\alpha}}. \tag{15}$$

Observe that for any i,

$$|\phi \circ f_i(x) - \phi \circ g_i(x)| \le \varepsilon_0 \|\phi\|_{C^1}.$$

Since μ is $\mathcal{M}_\epsilon^*$-invariant, this implies that

$$\left|\int (\phi - \mathcal{M}\phi)\, d\mu\right| = \left|\int (\mathcal{M}_\varepsilon \phi - \mathcal{M}\phi)\, d\mu\right| \le \varepsilon_0 \|\phi\|_{C^1}.$$

Substituting $(\mathrm{Id} - \mathcal{M})^{-1}(\phi - \int \phi\, dm)$ for the function ϕ in the previous line and using (15) yields a first order approximation:

$$\left|\int \phi\, d\mu - \int \phi\, dm\right| \le \varepsilon_0 C_1 \|\phi\|_{C^{1+\alpha}}. \tag{16}$$

We now use this first order approximation to obtain a better estimate. Note the Taylor expansion

$$\phi \circ f_i(x) - \phi \circ g_i(x) = (\nabla_{Y_i}\phi)(g_i(x)) + O(\varepsilon_0^2 \|\phi\|_{C^2}).$$

Integrating against μ yields

$$\begin{aligned}\int (\phi - \mathcal{M}\phi)\, d\mu &= \int (\mathcal{M}_\varepsilon\phi - \mathcal{M}\phi)\, d\mu \\ &= \int \frac{1}{m}\sum_{i=1}^{m} \nabla_{Y_i}\phi(g_i(x))\, d\mu + O(\varepsilon_0^2 \|\phi\|_{C^2}).\end{aligned}$$

We now plug in $(\mathrm{Id}-\mathcal{M})^{-1}(\phi-\int\phi\,dm)$ for ϕ in the previous line and use (15) to obtain

$$\int \phi\,d\mu - \int \phi\,dm = \int \frac{1}{m}\sum_{i=1}^{m}\Big(\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big)\Big)(g_i(x))\,d\mu + O(\varepsilon_0^2\|\phi\|_{C^{2+\alpha}}).$$

Using (16) on the first term on the right hand side above yields

$$\int \phi\,d\mu - \int \phi\,dm = \int \frac{1}{m}\sum_{i=1}^{m}\Big(\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big)\Big)(g_i(x))\,dm + O\Big(\varepsilon_0\Big\|\sum_{i=1}^{m}\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\phi\Big\|_{C^{1+\alpha}}\Big) + O(\varepsilon_0^2\|\phi\|_{C^{2+\alpha}}). \tag{17}$$

Note that

$$\Big\|\sum_{i=1}^{m}\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\phi\Big\|_{C^{1+\alpha}} = O(\varepsilon_{2+\alpha}\|(\mathrm{Id}-\mathcal{M})^{-1}\phi\|_{C^{2+\alpha}}).$$

The application of (15) to $\|(\mathrm{Id}-\mathcal{M})^{-1}\phi\|_{C^{2+\alpha}}$ then shows that the first big-O term in (17) is $O(\varepsilon_0\varepsilon_{2+\alpha}\|\phi\|_{C^{2+2\alpha}})$. Thus,

$$\int \phi\,d\mu - \int \phi\,dm = \int \frac{1}{m}\sum_{i=1}^{m}\Big(\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big)\Big)(g_i(x))\,dm + O(\varepsilon_{2+\alpha}^2\|\phi\|_{C^{2+2\alpha}}).$$

Now, by translation invariance of the Haar measure we may remove the g_i's:

$$\int \phi\,d\mu - \int \phi\,dm = \int \frac{1}{m}\sum_{i=1}^{m}\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big)\,dm + O(\varepsilon_{2+\alpha}^2\|\phi\|_{C^{2+2\alpha}}).$$

This proves everything except (14).

We now estimate the integral of

$$\begin{aligned}\mathcal{U}(\phi) &= \frac{1}{m}\sum_{i=1}^{m}\nabla_{Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big)\\ &= \nabla_{\frac{1}{m}\sum_{i=1}^{m}Y_i}(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big),\end{aligned}$$

against Haar. By (15) there exists C_1 such that

$$\Big\|(\mathrm{Id}-\mathcal{M})^{-1}\Big(\phi-\int\phi\,dm\Big)\Big\|_{C^1} \le C_1\|\phi\|_{C^{1+\alpha}},$$

which establishes equation (14) by a similar argument to the estimate of the big-O term occurring in the previous part of this proof. ∎

4. Strain and Lyapunov exponents

In this section we study the Lyapunov exponents of perturbations of isometric systems. The main result is Proposition 26, which gives a Taylor expansion of the Lyapunov exponents of a perturbation. The terms appearing in the Taylor expansion have a particular geometric meaning. We explain this meaning in terms of two "strain" tensors associated to a diffeomorphism. These tensors measure how far a diffeomorphism is from being an isometry. After introducing these tensors, we prove Proposition 26. The Lyapunov exponents of a random dynamical system may be calculated by integrating against a stationary measure of a certain extension of the original system. By using Proposition 23, we are able to approximate such stationary measures by the Haar measure and thereby obtain a Taylor expansion.

4.1. Norms on tensors

Throughout this paper we use the pointwise L^2 norm on tensors, which we now describe. For a more detailed treatment, see the discussion surrounding [22, Prop. 2.40]. If V is an inner product space with orthonormal basis $[e_1, \dots, e_n]$, then $V^{\otimes k}$ has a basis of tensors of the form

$$e_{i_1} \otimes \cdots \otimes e_{i_k}$$

where $1 \le i_j \le n$ for each $1 \le j \le k$. We declare the vectors of this basis to be orthonormal for the inner product on $V^{\otimes k}$. This norm is independent of the choice of orthonormal basis. For a continuous tensor field T on a closed Riemannian manifold M, we write $\|T\|$ for $\max_{x\in M} \|T(x)\|$. If T is a tensor on a Riemannian manifold M, we define its L^2 norm in the expected way by integrating the norm of $T(x)$ as a tensor on $T_x M$ over all points $x \in M$, i.e.

$$\|T\|_{L^2} = \left(\int_M \|T(x)\|^2 \, d\mathrm{vol}(x)\right)^{1/2}.$$

4.2. Strain

If a diffeomorphism of a Riemannian manifold is an isometry, then it pulls back the metric tensor to itself. Consequently, if we are interested in how near a diffeomorphism is to being an isometry, it is natural to consider the difference between the metric tensor and its pullback. This leads us to the following definition.

Definition 24. Suppose that f is a diffeomorphism of a Riemannian manifold (M, g). We define the *Lagrangian strain tensor* associated to f to be

$$E^f := \tfrac{1}{2}(f^* g - g).$$

This definition is consonant with the definition of the Lagrangian strain tensor that appears in continuum mechanics (cf. [21]).

The strain tensor will be useful for two reasons. First, it naturally appears in the Taylor expansion in Proposition 26, which will allow us to conclude that a random dynamical system with small Lyapunov exponents has small strain. Secondly, we prove in Theorem 27 that for certain manifolds, a diffeomorphism with small strain is near to an isometry. The combination of these two things will be essential in the proof of our main linearization result, Theorem 1, which shows that perturbations with all Lyapunov exponents zero are conjugate to isometric systems.

We now introduce two refinements of the strain tensor that will appear in the Taylor expansion in Proposition 26. Note that E^f is a $(0, 2)$-tensor. Consequently, we may take its trace with respect to the ambient metric g.

Definition 25. Suppose that f is a diffeomorphism of a Riemannian manifold (M, g). We define the *conformal strain tensor* by

$$E_C^f := \frac{\operatorname{Tr}(f^*g - g)}{2d} g.$$

We define the *non-conformal strain tensor* by

$$E_{NC}^f := E^f - E_C^f = \frac{1}{2}\left(f^*g - g - \frac{\operatorname{Tr}(f^*g - g)}{d} g\right).$$

4.3. *Taylor expansion of Lyapunov exponents*

Suppose that M is a manifold and that f is a diffeomorphism of M. Let $\operatorname{Gr}_r(M)$ denote the Grassmannian bundle of r-planes in TM. When working with $\operatorname{Gr}_r(M)$ we write a subspace of T_xM as E_x to emphasize the basepoint. Then f naturally induces a map $F\colon \operatorname{Gr}_r(M) \to \operatorname{Gr}_r(M)$ by sending a subspace $E_x \in \operatorname{Gr}_r(T_xM)$ to $D_xfE_x \in \operatorname{Gr}_r(T_{f(x)}M)$. If we have a random dynamical system on M, then by this construction we naturally obtain a random dynamical system on $\operatorname{Gr}_r(M)$. The following proposition should be compared with [10, Prop. 3].

Proposition 26. *Suppose that M is a compact connected Riemannian manifold such that* $\operatorname{Isom}(M)$ *is semisimple and that* $\operatorname{Isom}(M)^\circ$ *acts transitively on* $\operatorname{Gr}_r(M)$. *Suppose that $S = (g_1, \dots, g_m)$ is a Diophantine tuple of elements of* $\operatorname{Isom}(M)$. *Then there exist $\epsilon > 0$ and $k > 0$ such that if $(f_1, \dots, f_m)$ is a tuple with elements in* $\operatorname{Diff}^\infty(M)$ *such that $d_{C^k}(f_i, g_i) < \epsilon$, then the following holds. Suppose that μ is an ergodic stationary measure for the random dynamical system obtained from $(f_1, \dots, f_m)$. Let Λ_r be the sum of the top r Lyapunov exponents of μ. Then*

$$\Lambda_r(\mu) = -\frac{r}{2dm}\sum_{i=1}^m \int_M \|E_C^{f_i}\|^2 \, d\mathrm{vol} + \frac{r(d-r)}{(d+2)(d-1)m}\sum_{i=1}^m \int_M \|E_{NC}^{f_i}\|^2 \, d\mathrm{vol} + \int_{\operatorname{Gr}_r(M)} \mathcal{U}(\psi)\, d\mathrm{vol} + O(\varepsilon_k^3). \tag{18}$$

where $\psi = \frac{1}{m}\sum_{i=1}^m \ln\det(Df_i \mid E_x)$, $\varepsilon_k = \max_i d_{C^k}(f_i, g_i)$, $\mathcal{U}$ is defined as in Proposition 23, and det *is defined in Appendix D.*

Proof. Given the random dynamical system on M generated by the tuple $(f_1, \dots, f_m)$, there is the induced random dynamical system on $\mathrm{Gr}_r(M)$ generated by $(F_1, \dots, F_m)$ as described before the statement of the proposition. The Lyapunov exponents of the system on M may be obtained from the system on $\mathrm{Gr}_r(M)$ in the following way. By [20, Ch. III, Thm. 1.2], given an ergodic stationary measure μ on M, there exists a stationary measure $\overline{\mu}$ on $\mathrm{Gr}_r(M)$ such that

$$\Lambda_r(\mu) = \frac{1}{m}\sum_{i=1}^{m}\int_{\mathrm{Gr}_r(M)} \ln\det(Df_i \mid E_x)\, d\overline{\mu}(E_x).$$

Reversing the order of summation, this is equal to

$$\int_{\mathrm{Gr}_r(M)} \frac{1}{m}\sum_{i=1}^{m} \ln\det(Df_i \mid E_x)\, d\overline{\mu}(E_x). \tag{19}$$

As $\mathrm{Isom}(M)$ acts transitively on $\mathrm{Gr}_r(M)$, $\mathrm{Gr}_r(M)$ is a homogeneous space of $\mathrm{Isom}(M)$. Thus as $(g_1, \dots, g_m)$ is Diophantine, we may apply Proposition 23 to approximate the integral in (19). Letting $\mathcal{U}$ be as in that proposition, there exists k such that

$$\begin{aligned}\Lambda_r(\mu) = &\int_{\mathrm{Gr}_r(M)} \frac{1}{m}\sum_{i=1}^{m} \ln\det(Df_i \mid E_x)\, d\mathrm{vol}(E_x)\\ &+ \int_{\mathrm{Gr}_r(M)} \mathcal{U}\left(\frac{1}{m}\sum_{i=1}^{m} \ln\det(Df_i \mid E_x)\right) d\mathrm{vol}(E_x)\\ &+ O\left(\left(\max_i d_{C^k}(F_i, G_i)\right)^2 \left\|\sum_{i=1}^{m} \ln\det(Df_i \mid E_x)\right\|_{C^k}\right).\end{aligned} \tag{20}$$

We now estimate the error term. The following two estimates follow by working in a chart on $\mathrm{Gr}_r(M)$. If f, g are two maps of M and F, G are the induced maps on $\mathrm{Gr}_r(M)$, then $d_{C^k}(F, G) = O(d_{C^{k+1}}(f, g))$. In addition, by Lemma 58 in Appendix D we have

$$\left\|\sum_{i=1}^{m} \ln\det(Df_i \mid E_x)\right\|_{C^k} = O(\varepsilon_{k+1}). \tag{21}$$

Thus the error term in (20) is small enough to conclude (18).

To finish, we apply the Taylor expansion in Proposition 59 of Appendix E to

$$\int_{\mathrm{Gr}_r(M)} \ln\det(Df_i \mid E_x)\, d\mathrm{vol}(E_x),$$

which gives precisely the first two terms on the right hand side of (18) and an error that is $O(\varepsilon_1^3)$. ■

5. Diffeomorphisms of small strain: extracting an isometry in the KAM scheme

In this section we prove Proposition 28, which shows that a diffeomorphism of small strain on an isotropic manifold is near to an isometry. In the KAM scheme, we will see that diffeomorphisms with small Lyapunov exponents are low strain and hence conclude by

Proposition 28 that they are near to isometries. Proposition 28 follows from Theorem 27, which shows that certain diffeomorphisms with small strain of a closed Riemannian manifold are C^0 close to the identity.

Theorem 27. *Suppose that (M, g) is a closed Riemannian manifold. Then there exist $1 > r > 0$ and $C > 0$ such that if $f \in \mathrm{Diff}^2(M)$ and*

(1) *there exists $x \in M$ such that $f(x) = x$ and $\|D_x f - \mathrm{Id}\| = \theta < r$,*

(2) *$\|f^*g - g\| = \eta < r$,*

(3) *$d_{C^2}(f, \mathrm{Id}) = \kappa < r$,*

then for all $\gamma \in (0, r)$,

$$d_{C^0}(f, \mathrm{Id}) \le C(\theta + \kappa\gamma + \eta\gamma^{-1}).$$

Theorem 27 is the main ingredient in the proof of our central technical result.

Proposition 28. *Suppose that (M, g) is a closed isotropic Riemannian manifold. Then for all $\sigma > 0$ and all integers $\ell > 0$, there exist k and $C, r > 0$ such that for every $f \in \mathrm{Diff}^k(M)$, if there exists an isometry $I \in \mathrm{Isom}(M)$ such that*

(1) *$d_{C^k}(I, f) < r$,*

(2) *$\|f^*g - g\|_{H^0} < r$,*

then there exists an isometry $R \in \mathrm{Isom}(M)$ such that

$$d_{C^0}(R, I) < C(d_{C^2}(f, I) + \|f^*g - g\|_{H^0}^{1-\sigma}), \tag{22}$$

$$d_{C^\ell}(f, R) < C\big(\|f^*g - g\|_{H^0}^{1/2-\sigma} d_{C^2}(f, I)^{1/2-\sigma}\big). \tag{23}$$

Though the statement of Proposition 28 is technical, its use in the proof of Theorem 1 is fairly transparent: the proposition produces an isometry near to a diffeomorphism with small strain, which is the essence of the iterative step in the KAM scheme. This remedies the gap in [10].

5.1. Low strain diffeomorphisms on a general manifold: proof of Theorem 27

The main geometric idea in the proof of Theorem 27 is to study distances by intersecting spheres. In order to show that a diffeomorphism f is close to the identity, we must show that it does not move points far. As we shall show, a diffeomorphism of small strain distorts distances very little. Consequently, a diffeomorphism of small strain nearly carries spheres to spheres. If we have two points x and y that are fixed by f, then the unit spheres centered at x and y are carried near to themselves by f. Consequently, the intersection of those spheres will be nearly fixed by f. By considering the intersection of spheres in this way, we may take a small set on which f nearly fixes points and enlarge that set until it fills the whole manifold.

Before the proof of the theorem we prove several lemmas.

Lemma 29. *Let M be a closed Riemannian manifold. There exists $C > 0$ such that if $f \in \mathrm{Diff}^1(M)$ and $\|f^*g - g\| \leq \eta$ then for all $x, y \in M$,*

$$(1 - C\eta)d(x, y) \leq d(f(x), f(y)) \leq (1 + C\eta)d(x, y).$$

Proof. If γ is a path between x and y parametrized by arc length, then $f \circ \gamma$ is a path between $f(x)$ and $f(y)$. The length of $f \circ \gamma$ is equal to

$$\begin{aligned}\mathrm{len}(f \circ \gamma) &= \int_0^{\mathrm{len}(\gamma)} \sqrt{g(Df\dot{\gamma}, Df\dot{\gamma})}\, dt = \int_0^{\mathrm{len}(\gamma)} \sqrt{f^*g(\dot{\gamma}, \dot{\gamma})}\, dt \\ &= \int_0^{\mathrm{len}(\gamma)} \sqrt{g(\dot{\gamma}, \dot{\gamma}) + [f^*g - g](\dot{\gamma}, \dot{\gamma})}\, dt \\ &= \int_0^{\mathrm{len}(\gamma)} \sqrt{1 + [f^*g - g](\dot{\gamma}, \dot{\gamma})}\, dt.\end{aligned}$$

By our assumption on the norm of $f^*g - g$, there exists C such that $|[f^*g - g](\dot{\gamma}, \dot{\gamma})| \leq C\eta$. Then using $\sqrt{1 + x} \leq 1 + x$ for $x \geq 0$, we see that

$$\mathrm{len}(f \circ \gamma) \leq \int_0^{\mathrm{len}(\gamma)} 1 + |[f^*g - g](\dot{\gamma}, \dot{\gamma})|\, dt \leq \mathrm{len}(\gamma) + C\eta\, \mathrm{len}(\gamma).$$

The lower bound follows similarly by using $1 + x \leq \sqrt{1 + x}$ for $-1 \leq x \leq 0$. ■

Lemma 30. *Let M be a closed Riemannian manifold. Then there exist $r, C > 0$ such that for all $f \in \mathrm{Diff}^2(M)$, if*

(1) *there exists $x \in M$ such that $f(x) = x$ and $\|D_x f - \mathrm{Id}\| = \theta < r$,*

(2) *$d_{C^2}(f, \mathrm{Id}) = \kappa < r$,*

then for all $0 < \gamma < r$ and y such that $d(x, y) < \gamma$,

$$d(y, f(y)) \leq C(\gamma\theta + \gamma^2\kappa).$$

Proof. Let $r = \mathrm{inj}(M)/2$. We work in a fixed exponential chart centered at x, so that x is represented by 0 in the chart. Write

$$f(y) = 0 + D_0 f y + R(y) = y + (D_0 f - \mathrm{Id})y + R(y).$$

As the C^2 distance between f and the identity is at most κ, by Taylor's theorem $R(y)$ is bounded in size by $C\kappa|y|^2$ for a uniform constant C. Thus

$$|f(y) - y| \leq \theta|y| + C\kappa|y|^2.$$

In particular, for all y such that $|y| \leq \gamma < r$,

$$|f(y) - y| \leq C'(\gamma\theta + \gamma^2\kappa).$$

But the distance in such a chart is uniformly bi-Lipschitz with respect to the metric on M, so the lemma follows. ■

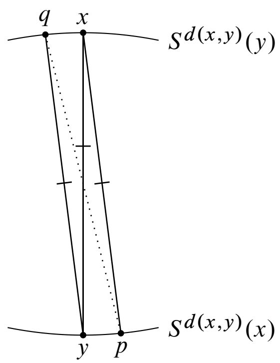

Fig. 1. The four points x, y, p, q appearing in Lemma 31. Given x, y, p, the lemma produces the point q and gives an estimate on the length of the dotted line, which is longer than $d(x, y)$.

The following geometric lemma produces points on two spheres in a Riemannian manifold that are further apart than the centers of the spheres.

Lemma 31. *Let M be a closed Riemannian manifold. There exist $C, r > 0$ such that for all $\beta \in (0, r)$, if $x, y \in M$ satisfy* $\operatorname{inj}(M)/3 < d(x, y) < \operatorname{inj}(M)/2$, *and there is a fixed $p \in M$ such that $d(x, p) = d(y, x)$ and $d(p, y) < r$, then there exists $q \in M$ depending on p such that*

(1) $d(q, y) = d(y, x)$,

(2) $d(q, x) < \beta$,

(3) $d(q, p) \geq d(x, y) + Cd(y, p)\beta$.

In order to prove Lemma 31, we recall the following form of the second variation of length formula. For a proof of this and related discussion, see [5, Ch. 1, §6].

Lemma 32. *Let M be a Riemannian manifold and γ be a unit speed geodesic. Let $\gamma_{v,w}$ be a two-parameter family of constant speed geodesics parametrized by $\gamma_{v,w}$: $[a, b] \times (-\epsilon, \epsilon) \times (-\epsilon, \epsilon) \to M$ such that $\gamma_{0,0} = \gamma$. Suppose that $\frac{\partial \gamma_{v,w}}{\partial v} = V$ and $\frac{\partial \gamma_{v,w}}{\partial w} = W$ are both normal to $\dot{\gamma}_{0,0}$, which we denote by T. Then*

$$\frac{\partial^2 \operatorname{len}(\gamma_{v,w})}{\partial v \partial w} = \langle \nabla_W V, T\rangle|_a^b + \langle V, \nabla_T W\rangle|_a^b.$$

Proof of Lemma 31. We will give a geometric construction using the points x and y and then explain how this construction may be applied to the particular point p to produce a point q.

Let Q be a unit tangent vector based at y that is tangent to $S^{d(x,y)}(x)$, the sphere of radius $d(x, y)$ centered at x. Let $\gamma_t : [a, b] \to M$ be a one-parameter family of geodesics parametrized by arc length so that γ_0 is the unit speed geodesic from x to y, $\partial_t \gamma_t(b)|_{t=0} = Q$, $\gamma_t(b)$ is a path in $S^{d(x,y)}(x)$, and $\gamma_t(a) = x$ for all t. The variation γ_t gives rise to a Jacobi field Y. Note that $Y(a) = 0$, $Y(b) = Q$, and Y is a normal Jacobi field.

Next, let X be the Jacobi field along γ_0 defined by $X(b) = 0$ and $\nabla_T X|_b = Y(b)$, where T denotes $\dot{\gamma}_0$, i.e. the tangent to the curve γ_0. Such a field exists and has uniformly bounded norms because γ_0 is shorter than the injectivity radius. Let $\eta_t\colon [a,b] \to M$ be a one-parameter family of geodesics tangent to the field X such that $\eta_t(b) = y$, η_t is arc length parametrized, and $\eta_0 = \gamma_0$. Note that each η_t has length $d(x,y)$. Let T now denote $\dot{\gamma}_{s,t}$, which gives the tangent direction to each curve $\gamma_{s,t}$ in the variation.

Define $\gamma_{s,t}\colon [a,b] \to M$ to be the arc length parametrized geodesic between $\eta_s(a)$ and $\gamma_t(b)$. The variation $\gamma_{s,t}$ is a two-parameter variation satisfying the hypotheses of Lemma 32. Consequently, we see that

$$\frac{d^2 \operatorname{len}(\gamma_{s,t})}{dsdt} = \langle \nabla_X Y, T\rangle|_a^b + \langle Y, \nabla_T X\rangle|_a^b. \tag{24}$$

The first term may be rewritten as

$$\langle \nabla_X Y, T\rangle|_a^b = \nabla_X \langle Y, T\rangle|_a^b - \langle Y, \nabla_X T\rangle|_a^b. \tag{25}$$

As $Y(a) = 0$ and $X(b) = 0$, the second term in (25) is zero. Similarly $\nabla_X \langle Y, T\rangle|_b = 0$. We claim that $\nabla_X \langle Y, T\rangle|_a = 0$ as well. To see this we claim that $Y = \partial_t \gamma_{s,t}|_a = 0$ for all s. This is because $\gamma_{s,t}(a)$ is constant in t as $\gamma_{s,t}(a)$ depends only on s. Thus $\langle Y, T\rangle|_a = 0$. When we differentiate by X, we are differentiating along the path $\gamma_{s,0}(a)$. Thus $\nabla_X \langle Y, T\rangle|_a = 0$ as $\langle Y, T\rangle$ is 0 along this path. Thus $\langle \nabla_X Y, T\rangle|_a^b = 0$. Noting in addition that $Y(a) = 0$, equation (24) simplifies to

$$\frac{d^2 \operatorname{len}(\gamma_{s,t})}{dsdt} = \langle Y, \nabla_T X\rangle|_b.$$

Hence as we defined X so that $\nabla_T X|_b = Y(b)$,

$$\frac{d^2 \operatorname{len}(\gamma_{s,t})}{dsdt} = \langle Y(b), Y(b)\rangle = \|Q\| = 1.$$

Note next that $\frac{d^2}{ds^2} \operatorname{len}(\gamma_{s,t}) = 0$ because the geodesics $\gamma_{s,0}$ all have the same length. Similarly, $\frac{d^2}{dt^2} \operatorname{len}(\gamma_{s,t}) = 0$. Thus we have the Taylor expansion

$$\frac{d^2}{dsdt} \operatorname{len}(\gamma_{s,t}) = d(x,y) + st + O(s^3, t^3). \tag{26}$$

There exist $r_0 > 0$ and $C > 0$ such that for all $0 \le s, t < r_0$,

$$\operatorname{len}(\gamma_{s,t}) \ge d(x,y) + Cst. \tag{27}$$

Consider now the pairs of points $\gamma_{s,0}(a)$ and $\gamma_{0,t}(b)$. We claim that if p is of the form $p = \gamma_{0,t}(b)$ for some small t then we may take $q = \gamma_{s,0}(a)$, where the choice of s will be dictated by β.

Note that

$$d(\gamma_{s,0}(a), x) = s\|X(a)\| + O(s^2) \quad \text{and} \quad d(\gamma_{0,t}(b), y) = t\|Y(b)\| + O(t^2).$$

Hence there exists s_0 such that for $0 < s, t < s_0$,

$$d(\gamma_{s,0}(a), x) < 2s\|X(a)\| \quad \text{and} \quad d(\gamma_{0,t}(b), y) < 2t\|Y(b)\|. \tag{28}$$

For any $\beta < \min\{2s_0\|X(a)\|, 2r_0\|X(a)\|\}$, by (27) taking $s = \beta/(2\|X(a)\|)$ we obtain

$$d(\gamma_{s,0}(a), \gamma_{0,t}(b)) \geq d(x, y) + t\beta C/(2\|X(a)\|),$$

which by (28) implies

$$d(\gamma_{s,0}(a), \gamma_{0,t}(b)) \geq d(x, y) + \frac{C}{4\|X(a)\|\,\|Y(b)\|}\beta d(\gamma_{0,t}(b), y).$$

By (28) and our choice of s,

$$d(\gamma_{s,0}(a), x) < \beta.$$

Finally, $d(\gamma_{s,0}(a), y) = d(x, y)$ by the construction of the variation. Thus the conclusion of the lemma holds for the points $p = \gamma_{0,t}(b)$ and $q = \gamma_{s,0}(a)$.

We claim that this gives the full result. First, note that for all pairs of points x and y and choices of vectors Q in our construction, the norms $\|X(a)\|$ and $\|Y(b)\|$ are bounded above and below. This is because the distance minimizing geodesic from X to Y does not cross the cut locus. Similarly, the constants C, r_0, and s_0 may be uniformly bounded below over all such choices of x and y by compactness. Thus as all these constants are uniformly bounded independent of x, y and Q, the above argument shows that for any pair x and y that there is a neighborhood N of y in $S^{d(x,y)}$ of uniformly bounded size, such that for any $p \in N$ there exists q satisfying the conclusion of the lemma. This gives the result as any p sufficiently close to y such that $d(x, p) = d(x, y)$ lies in such a neighborhood N. ∎

The following lemma shows that if a diffeomorphism with small strain nearly fixes a large region, then that diffeomorphism is close to the identity.

Lemma 33. *Let (M, g) be a closed Riemannian manifold. Then there exists $r_0 \in (0, 1)$ such that for any $r', \beta \in (0, r_0)$ there exists $C > 0$ such that if $f \in \mathrm{Diff}^1(M)$ and*

(1) $d_{C^0}(f, \mathrm{Id}) \leq r_0$,

(2) *there exists a point $x \in M$ such that all y with $d(x, y) < r'$ satisfy $d(y, f(y)) \leq \beta \leq r_0$,*

(3) $\|f^*g - g\| = \eta \leq r_0$,

then

$$d_{C^0}(f, \mathrm{Id}) < C(\beta + \eta). \tag{29}$$

Proof. Let r_1, C_1 denote the r and C in Lemma 31. Let C_2 be the constant in Lemma 29. There exists a constant r_2 such that for any $x, y \in M$ with $\mathrm{inj}(M)/3 < d(y, x) < \mathrm{inj}(M)/2$ and any z such that $d(y, z) < r_2$, we have $d(y, \hat{z}) < r_1$, where $\hat{z}$ is the radial projection of z onto $S^{d(x,y)}(x)$. Let $r_0 = \min\{r_1, r_2, \mathrm{inj}(M)/24\}$.

Suppose that $x \in M$ has the property that $d(x, z) < r$ implies $d(z, f(z)) \leq \beta$. Suppose that y is a point such that $\mathrm{inj}(M)/3 < d(y, x) < \mathrm{inj}(M)/2$. Let $\widehat{f(y)}$ be the radial projection of $f(y)$ onto $S^{d(x,y)}(x)$.

By choice of $r_0 \le r_2$, $d(y, f(y)) < r_2$ and so $d(y, \widehat{f(y)}) \le r_1$. Hence we may apply Lemma 31 with $\beta = r'$, $x = x$, $y = y$ and $p = \widehat{f(y)}$ to conclude that there exists a point $q \in M$ such that

$$d(q, y) = d(x, y), \tag{30}$$

$$d(q, x) < r', \tag{31}$$

$$d(q, \widehat{f(y)}) \ge d(x, y) + C_1 d(y, \widehat{f(y)}) r'. \tag{32}$$

Using the triangle inequality, we bound the left hand side of (32) to find

$$d(q, f(q)) + d(f(q), f(y)) + d(f(y), \widehat{f(y)}) \ge d(q, \widehat{f(y)}) \ge d(x, y) + C_1 d(y, \widehat{f(y)}) r'. \tag{33}$$

First, as $d(q, x) < r'$ and points within r' of x do not move more than β,

$$d(q, f(q)) \le \beta.$$

Second, by Lemma 29, as the distance between q and y is bounded above by $\mathrm{inj}(M)/2$, there exists C_3 such that

$$d(f(q), f(y)) \le d(q, y)(1 + C_2\eta) = d(x, y) + C_3\eta.$$

Similarly, as $\mathrm{inj}(M)/3 < d(x, y) < \mathrm{inj}(M)/2$, Lemma 29 implies the bounds

$$d(x, f(y)) \le d(x, f(x)) + d(f(x), f(y)) \le \beta + d(x, y) + C_3\eta \tag{34}$$

and similarly

$$d(x, f(y)) \ge d(x, y) - \beta - C_3\eta. \tag{35}$$

For w sufficiently close to $S^{d(x,y)}(x)$ we claim that the radial projection $\hat{w}$ is the point in $S^{d(x,y)}(x)$ that minimizes the distance to w. To see this we use the fact that below the injectivity radius, geodesics are the unique distance minimizing paths between two points. There are two cases: if $d(x, w) > d(x, y)$ and there is some other point $w' \in S^{d(x,y)}(x)$ with $d(w', w) \le d(\hat{w}, w)$, then the path from x to w' to w along geodesics must be strictly longer than the geodesic path from x directly to $\hat{w}$. If $d(x, w) < d(x, y)$ and $\hat{w} \ne w' \in S^{d(x,y)}(x)$, then one obtains two distance minimizing paths from x to $S^{d(x,y)}(x)$ passing through w: the first along a single geodesic and the second from x to w and then from w to w'. By the uniqueness of distance minimizing geodesics, the latter path must have length greater than $d(x, y)$ because it is not a geodesic. Thus $d(w, w') > d(w, \hat{w})$; a contradiction.

The estimates (34) and (35) imply that $|d(f(y), x) - d(x, y)| \le \beta + C_3\eta$. Thus the distance from $f(y)$ to $S^{d(x,y)}(x)$ is at most $\beta + C_3\eta$. By the previous paragraph, $\widehat{f(y)}$ is the point in $S^{d(x,y)}(x)$ that minimizes distance to $f(y)$. Thus

$$d(f(y), \widehat{f(y)}) \le \beta + C_3\eta. \tag{36}$$

Thus, from (33) we obtain

$$\beta + d(x, y) + C_3\eta + \beta + C_3\eta \ge d(x, y) + C_1 d(y, \widehat{f(y)}) r'.$$

Consequently,

$$\frac{2\beta + 2C_3\eta}{C_1 r'} \geq d(y, \widehat{f(y)}).$$

Hence

$$d(y, f(y)) \leq d(f(y), \widehat{f(y)}) + d(y, \widehat{f(y)}) \leq \frac{2\beta + 2C_3\eta}{C_1 r'} + \beta + C_3\eta.$$

Thus by introducing a new constant $C_4 \geq 1$, we see that for any y satisfying $\operatorname{inj}(M)/3 < d(y, x) < \operatorname{inj}(M)/2$, we have

$$d(y, f(y)) \leq C_4(\beta + \eta).$$

Note that the constant C_4 depends only on r' and (M, g).

Consider a point y where $(1/3 + 1/24)\operatorname{inj}(M) < d(x, y) < (1/2 - 1/24)\operatorname{inj}(M)$. Because $r' < \operatorname{inj}(M)/24$ such a point y has a neighborhood of size r' on which points are moved at most distance $C_4(\beta + \eta)$ by f. Hence we may repeat the procedure taking y as the new basepoint. Let x be the given point in the statement of the lemma. Any point $q \in M$ may be connected to x via a finite sequence of points $x = x_0, \ldots, x_n = q$ such that each consecutive pair of points in the sequence are a distance between $(1/3 + 1/24)\operatorname{inj}(M)$ and $(1/2 - 1/24)\operatorname{inj}(M)$ apart. As M is compact there is a uniform upper bound on the length of the shortest such sequence. If N is such a bound, the above argument shows that for all $q \in M$,

$$d(q, f(q)) \leq N C_4^N(\beta + \eta),$$

which gives the result. ∎

The proof of Theorem 27 consists of two steps. First a disk of uniform radius is produced on which f nearly fixes points. Then Lemma 33 is applied to this disk to conclude that f is near to the identity.

Proof of Theorem 27. Let r_1, C_1 be the r and C in Lemma 30, and let r_2, C_2 denote the r and c in Lemma 31. There will be a constant $r_3 > 0$ introduced later when it is needed. Let r_4 denote the constant r_0 appearing in Lemma 33. We let $r = \min\{1, r_1, r_2, r_3, r_4, \operatorname{inj}(M)/24\}$. Let C_3 be the constant in Lemma 29. Let $\gamma \in (0, r)$ be given.

By Lemma 30, for all z such that $d(x, z) < \gamma$,

$$d(z, f(z)) \leq C_1(\theta\gamma + \gamma^2\kappa). \tag{37}$$

Suppose that y satisfies $\operatorname{inj}(M)/3 < d(x, y) < \operatorname{inj}(M)/2$. Let $\widehat{f(y)}$ be the radial projection of $f(y)$ onto the sphere $S^{d(x,y)}(x)$.

By Lemma 29,

$$d(x, y)(1 - C_3\eta) \leq d(f(x), f(y)) \leq d(x, y)(1 + C_3\eta).$$

As $f(x) = x$, this implies

$$d(x, y)(1 - C_3\eta) \leq d(x, f(y)) \leq d(x, y)(1 + C_3\eta).$$

Hence as $d(x, y)$ is uniformly bounded above and below, there exists C_4 such that

$$d(f(y), \widehat{f(y)}) < C_4\eta. \tag{38}$$

There exists $r_3 > 0$ such that if $\eta < r_3$, then $C_4\eta < r_2$. Hence by our choice of r, $d(y, \widehat{f(y)}) < r_2$ and we may apply Lemma 31 with $\beta = \gamma$, $x = x$, $y = y$, $p = \widehat{f(y)}$ to deduce that there exists q such that

$$d(q, y) = d(x, y), \tag{39}$$

$$d(q, x) < \gamma, \tag{40}$$

$$d(q, \widehat{f(y)}) \geq d(x, y) + C_2 d(y, \widehat{f(y)})\gamma. \tag{41}$$

By Lemma 29, and using the fact that $d(x, y)$ is bounded by $\mathrm{inj}(M)/2$, there exists C_5 such that

$$d(f(q), f(y)) \leq d(q, y)(1 + C_3\eta) \leq d(x, y) + C_5\eta. \tag{42}$$

By (37), as $d(q, x) < \gamma$,

$$d(q, f(q)) < C_1(\theta\gamma + \kappa\gamma^2). \tag{43}$$

Using the triangle inequality with (38), (42), (43) to bound the left hand side of (41), we obtain

$$\begin{aligned} C_1(\theta\gamma + \kappa\gamma^2) + d(x, y) + C_5\eta + C_4\eta &\geq d(q, f(q)) + d(f(q), f(y)) + d(f(y), \widehat{f(y)}) \\ &\geq d(x, y) + C_2 d(y, \widehat{f(y)})\gamma. \end{aligned}$$

Moreover, (38) gives the lower bound $d(y, \widehat{f(y)}) > d(y, f(y)) - C_4\eta$. We then obtain

$$C_1(\theta\gamma + \kappa\gamma^2) + C_5\eta + C_4\eta \geq C_2 d(y, f(y))\gamma - C_2 C_4 \eta\gamma,$$

and so

$$\frac{C_1(\theta\gamma + \kappa\gamma^2) + C_5\eta + C_4\eta + C_2C_4\eta\gamma}{C_2\gamma} \geq d(y, f(y)).$$

The constants $C_1, \ldots, C_5$ are uniform over all y satisfying $\mathrm{inj}(M)/3 < d(x, y) < \mathrm{inj}(M)/2$. Thus there exists $C_6 > 0$ such that for all such y,

$$C_6(\eta\gamma^{-1} + \theta + \kappa\gamma) \geq d(y, f(y)). \tag{44}$$

Suppose that y is a point at distance $\frac{5}{12}\,\mathrm{inj}(M)$ from x. The above argument shows that if z satisfies $d(y, z) < \mathrm{inj}(M)/12$ then (44) holds with y replaced by z, i.e.

$$C_6(\eta\gamma^{-1} + \theta + \kappa\gamma) \geq d(z, f(z)).$$

Define

$$\alpha = C_6(\eta\gamma^{-1} + \theta + \kappa\gamma). \tag{45}$$

Assuming that $\alpha < r_4$, z satisfies hypothesis (2) of Lemma 33 with $\beta = \alpha$ and any $r' \leq \mathrm{inj}(M)/12$.

There are then two cases depending on whether $\alpha > r_4$ or $\alpha \le r_4$. If $\alpha \le r_4$, we apply Lemma 33 with $x_0 = z$, $r' = r/2$, and $\beta = \alpha$. This implies that there exists a C_7 depending only on $r/2$ such that

$$d_{C^0}(f, \mathrm{Id}) \le C_7(\eta\gamma^{-1} + \theta + \kappa\gamma).$$

If $\alpha > r_4$, then as $\kappa \le r_4$,

$$d_{C^0}(f, \mathrm{Id}) \le \kappa \le r_4 \le \alpha = C_6(\eta\gamma^{-1} + \theta + \kappa\gamma).$$

Thus letting $C_8 = \max\{C_6, C_7\}$, we have

$$d_{C^0}(f, \mathrm{Id}) \le C_8(\eta\gamma^{-1} + \theta + \kappa\gamma),$$

which gives the result. ∎

5.2. Application to isotropic spaces: proof of Proposition 28

We now prove Proposition 28, which is an application of Theorem 27 to isotropic spaces. The idea of the proof is geometric. We consider the diffeomorphism $I^{-1}f$. This diffeomorphism is small in C^0 norm, so there is an isometry R_1 that is close to the identity such that $R_1^{-1}I^{-1}f$ has a fixed point x. The differential of $R_1^{-1}I^{-1}f$ at x is very close to preserving both the metric tensor and the curvature tensor at x. We then use the following lemma to obtain an isometry R_2 that is close to $R_1^{-1}I^{-1}f$.

Lemma 34 ([18, Ch. IV, Ex. A.6]). *Let M be a simply connected Riemannian globally symmetric space or $\mathbb{R}\mathrm{P}^n$. If $x \in M$ and $L\colon T_xM \to T_xM$ is a linear map preserving both the metric tensor at x and the curvature tensor at x, then there exists $R \in \mathrm{Isom}(M)$ such that $R(x) = x$ and $D_xR = L$.*

We take the diffeomorphism in the conclusion of Proposition 28 to equal IR_1R_2. We then apply Theorem 27 to deduce that $R_2^{-1}R_1^{-1}I^{-1}f$ is near the identity diffeomorphism. It follows that IR_1R_2 is near to f. Before beginning the proof, we state some additional lemmas.

Lemma 35. *Suppose that V_1 and V_2 are two subspaces of a finite-dimensional inner product space W. Then there exists $C > 0$ such that if $x \in W$, then*

$$d(x, V_1 \cap V_2) < C(d(x, V_1) + d(x, V_2)).$$

Lemma 36. *Suppose that R is a tensor on $\mathbb{R}^n$. Let $\mathrm{stab}(R)$ be the subgroup of $\mathrm{GL}(\mathbb{R}^n)$ that stabilizes R under pullback. Then there exist $C, D > 0$ such that if $L\colon \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map and $\|L - \mathrm{Id}\| < D$, then*

$$d_{\mathrm{GL}(\mathbb{R}^n)}(L, \mathrm{stab}(R)) \le C\|L^*R - R\|.$$

Proof. Let $\mathfrak{s}$ be the Lie algebra of $\mathrm{stab}(R)$. Then consider the map ϕ from $\mathfrak{gl}$ to the tensor algebra on $\mathbb{R}^n$ given by

$$w \mapsto \exp(w)^*R - R.$$

We may write $w = v + v^{\perp}$, where $v \in \mathfrak{s}$ and $v \in \mathfrak{s}^{\perp}$. Because ϕ is smooth it has a Taylor expansion of the form

$$\phi(tv + tv^{\perp}) = 0 + tAv + tBv^{\perp} + O(t^2). \tag{46}$$

Note that A is zero because $v \in \mathfrak{s}$. We claim that B is injective. For the sake of contradiction, suppose $Bv^{\perp} = 0$ for some $v^{\perp} \in \mathfrak{s}^{\perp}$. Then $\exp(tv^{\perp})^* R - R = O(t^2)$. But then

$$\begin{aligned}\exp(v^{\perp})^* R - R &= \sum_{i=0}^{n-1} [\exp((i+1)v^{\perp}/n)^* R - \exp(iv^{\perp}/n)R] \\ &= \sum_{i=0}^{n-1} \exp(iv^{\perp}/n)^* (\exp(v^{\perp}/n)^* R - R) = O(1/n).\end{aligned}$$

Hence $\exp(v^{\perp})^* R - R = 0$, which contradicts $v^{\perp} \notin \mathfrak{s}$. Thus B is an injection and hence by Taylor's theorem for small $v^{\perp}$ there exists C_1 such that

$$\|\exp(v^{\perp})^* R - R\| \geq C_1 \|v^{\perp}\|. \tag{47}$$

By using the Taylor expansion (46) and noting that $A = 0$ there, we deduce from (47) that there exists $C_2 > 0$ such that

$$\|\exp(w)^* R - R\| \geq C_2 \|v^{\perp}\|. \tag{48}$$

It then follows there exists a neighborhood N of $\mathrm{Id} \in \mathrm{GL}(\mathbb{R}^n)$ such that $\mathrm{stab}(R) \cap N$ is the image of a disk $D \subset \mathfrak{s}$ under exp. Write $\mathfrak{gl} = \mathfrak{s} \oplus \mathfrak{s}^{\perp}$ as a vector space. Thus as exp is bi-Lipschitz in a neighborhood of $0 \in \mathfrak{gl}$ there exists C_3 such that if we write $w \in D$ as $w = v + v^{\perp}$, where $v \in \mathfrak{s}$ and $v^{\perp} \in \mathfrak{s}^{\perp}$, then

$$C_3^{-1} \|v^{\perp}\| \leq d_{\mathrm{GL}(\mathbb{R}^n)}(\exp(w), \exp(D)) \leq C_3 \|v^{\perp}\|. \tag{49}$$

As $\mathrm{stab}(R) \cap N = \exp(D)$, for all w in a smaller neighborhood $D' \subset D$ the middle term above is comparable to $d_{\mathrm{GL}(\mathbb{R}^n)}(\exp(w), \mathrm{stab}(R))$.

Thus combining (49) with (48), we obtain

$$d_{\mathrm{GL}(\mathbb{R}^n)}(\exp(w), \mathrm{stab}(R)) \leq C_2^{-1} C_3 \|\exp(w)^* R - R\|.$$

This gives the result as exp is a surjection onto a neighborhood of $\mathrm{Id} \in \mathrm{GL}(\mathbb{R}^n)$. ∎

The following lemma is immediate from [18, Thm. IV.3.3], which explicitly describes the isometries of globally symmetric spaces.

Lemma 37. *Suppose that M is a closed globally symmetric space. There exists $C > 0$ such that if $x, y \in M$, then there exists an isometry $I \in \mathrm{Isom}(M)^{\circ}$ such that $I(x) = y$ and $d_{C^0}(I, \mathrm{Id}) \leq C d(x, y)$. As $\mathrm{Isom}(M)^{\circ}$ is compact, it follows that for each k there exists a constant C_k such that one can choose I with $d_{C^k}(I, \mathrm{Id}) \leq C_k d(x, y)$.*

We also use the following lemma, which is the specialization of Lemma 36 to the metric tensor.

Lemma 38. *Suppose that V is a finite-dimensional inner product space with metric g. There exists a neighborhood U of* Id $\in$ GL(V) *and a constant C such that if $L \in U$ then*

$$d_{\mathrm{GL}(V)}(L, \mathrm{SO}(V)) \leq C\|L^*g - g\|,$$

where GL(V) *is endowed with the right-invariant Riemannian metric it inherits from the inner product space V.*

We now prove the proposition.

Proof of Proposition 28. Pick $0 < \lambda < 1$ and a small τ such that

$$\lambda/2 - \lambda\tau > 1/2 - \sigma \quad \text{and} \quad \sigma > \tau > 0. \tag{50}$$

We also assume without loss of generality that $\ell \geq 3$. By Lemma 55 there exist k_0 and $\epsilon_0 > 0$ such that if s is a smooth section of the bundle of symmetric 2-tensors over M with $\|s\|_{C^{k_0}} \leq 4$ and $\|s\|_{H^0} \leq \epsilon_0$, then $\|s\|_{C^\ell} \leq \|s\|_{H^0}^{1-\tau}$. Choose k such that

$$k > \max\left\{k_0, \frac{\ell}{1-\lambda}\right\}. \tag{51}$$

In addition, there are positive numbers $\epsilon_1, \ldots, \epsilon_7$ that will be introduced when needed in the proof below. We define

$$r = \min\{\epsilon_0, \epsilon_1^{1/(1-\tau)}, \epsilon_2, \ldots, \epsilon_7, 1\}.$$

Let $\epsilon_1 > 0$ be small enough that for any $x \in M$, if $L\colon T_xM \to T_xM$ is invertible and $\|L^*g - g\| \leq \epsilon_1$, then the conclusion of Lemma 38 holds for L.

Let $\eta = \|f^*g - g\|_{H^0}$ and $\varepsilon_2 = d_{C^2}(f, I)$. Consider the norm $\|f^*g - g\|_{C^{k_0}}$. As $d_{C^k}(I, f)$ is uniformly bounded, we see that $\|f^*g - g\|_{C^{k-1}}$ is uniformly bounded. In fact, there exists $\epsilon_2 > 0$ such that if $d_{C^k}(I, f) < \epsilon_2$, then $\|f^*g - g\|_{C^{k-1}} \leq 4$. As $r < \epsilon_0$, the discussion in the first paragraph of the proof implies that

$$\|f^*g - g\|_{C^3} \leq \eta^{1-\tau}. \tag{52}$$

Note that this is less than ϵ_1 by the choice of r.

For $x \in M$, we may consider the Lie group GL(T_xM) as well as its Lie algebra $\mathfrak{gl}$. There exists $\epsilon_3 > 0$ such that restricted to the ball of radius ϵ_3 about $0 \in \mathfrak{gl}$, the Lie exponential, which we denote by exp, is bi-Lipschitz with constant 2.

Let $x \in M$ be a point that is moved the maximum distance by $I^{-1}f$. By Lemma 37, there exists a constant $D_k > 0$ independent of x and an isometry R_1 such that $R_1(x) = I^{-1}f(x)$ and $d_{C^k}(R_1, \mathrm{Id}) < D_k d(x, I^{-1}f(x))$. Let $h = R_1^{-1}I^{-1}f$ and note that h fixes x. Note that there exists $\epsilon_4 > 0$ such that if $d_{C^k}(f, I) < \epsilon_4$, then by the previous sentence R_1 can be chosen so that $d_{C^k}(R_1, \mathrm{Id})$ is small enough that

$$\|D_xh - \mathrm{Id}\| \leq C_0\varepsilon_2. \tag{53}$$

We claim that $D_x h$ is near a linear map of $T_x M$ that preserves both the metric tensor and the curvature tensor. Let $\mathrm{SO}(T_x M)$ be the group of linear maps preserving the metric tensor on $T_x M$ and let G be the group of linear maps preserving the curvature tensor on $T_x M$. Both of these are subgroups of $\mathrm{GL}(T_x M)$. By the sentence after (52), $D_x h$ pulls back the metric on $T_x M$ to be within ϵ_1 of itself. Thus by Lemma 37, there exists a uniform constant C_1 such that $D_x h$ is within distance $C_1 \eta^{1-\tau}$ of $\mathrm{SO}(T_x M)$. Again by (52), we have $\|h^* g - g\|_{C^3} \leq \eta^{1-\tau}$. In particular, as the curvature tensor is defined by the second derivatives of the metric, this implies by Lemma 36 that there exists a constant C_2 such that $D_x h$ is within distance $C_2 \eta^{1-\tau}$ of G.

The previous paragraph shows that there exists C_3 such that $D_x h$ is within distance $C_3 \eta^{1-\tau}$ of both $\mathrm{SO}(T_x M)$ and G. Consider now the exponential map of $\mathrm{GL}(T_x M)$. As before, let $\mathfrak{gl}$ denote the Lie algebra of $\mathrm{GL}(T_x M)$. Let $H = \exp^{-1}(D_x h) \in \mathrm{GL}(T_x M)$. Note that this preimage is defined as $D_x h$ is near to the identity. Let $\mathfrak{so}$ be the Lie algebra of $\mathrm{SO}(T_x M)$ and let $\mathfrak{g}$ be the Lie algebra of G. As both $\mathrm{SO}(T_x M)$ and G are closed subgroups and exp is bi-Lipschitz, we conclude that the distance between H and each of $\mathfrak{so}$ and $\mathfrak{g}$ is bounded above by $2C_3 \eta^{1-\tau}$. Thus by Lemma 35, there exists C_4 such that H is at a distance at most $C_4 \eta^{1-\tau}$ from $\mathfrak{g} \cap \mathfrak{so}$. Let $X \in \mathrm{GL}(T_x M)$ be an element of $\mathfrak{g} \cap \mathfrak{so}$ minimizing the distance from H to $\mathfrak{g} \cap \mathfrak{so}$. There exists $\epsilon_5 > 0$ such that if $\eta \leq \epsilon_5$ then $C_4 \eta^{1-\tau} < \epsilon_3$. Hence as $r < \epsilon_5$, the same bi-Lipschitz estimate on the Lie exponential gives

$$d(\exp(X), D_x h) \leq 2C_4 \eta^{1-\tau}. \tag{54}$$

Note that $\exp(X) \in \mathrm{SO}(T_x M) \cap G$. By Lemma 34, there exists an isometry R_2 of M such that R_2 fixes x and $D_x R_2 = \exp(X)$. In fact, because of (53) and because X is within distance $C_4 \eta^{1-\tau}$ of H, we may bound the norm of X and hence deduce that there exists C_5 such that

$$d_{C^k}(R_2, \mathrm{Id}) \leq C_5(\varepsilon_2 + \eta^{1-\tau}). \tag{55}$$

The map R in the conclusion of the proposition will be $I R_1 R_2$. We must now check that $R = I R_1 R_2$ satisfies estimates (22) and (23). The former is straightforward from (55) combined with knowing that R_1 was constructed so that $d(R_1, \mathrm{Id}) \leq D' \varepsilon_2$ for some uniform $D' > 0$.

Let $h_2 = R_2^{-1} h$. The map h_2 has x as a fixed point. There exists $C_6 > 0$ such that

$$\|D_x h_2 - \mathrm{Id}\| \leq C_6 \eta^{1-\tau}, \tag{56}$$

$$\|h_2^* g - g\|_{C^3} \leq \eta^{1-\tau}, \tag{57}$$

$$d_{C^2}(h_2, \mathrm{Id}) \leq C_6(\varepsilon_2 + \eta^{1-\tau}), \tag{58}$$

$$d_{C^k}(h_2, \mathrm{Id}) \leq C_6(\eta^{1-\tau} + d_{C^k}(I, f)). \tag{59}$$

The first two estimates above are immediate from (54) and (52), respectively. The third and fourth follow from an estimate on C^k compositions, Lemma 50, and (55).

Let r_0 be the cutoff r appearing in Theorem 27. Note that there exists $\epsilon_6 > 0$ such that if $d_{C^k}(f, I) < \epsilon_6$ and $\eta < \epsilon_6$, then the right hand side of each of inequalities (56) through (59) is bounded above by r_0. Hence as $r < \epsilon_6$ we apply Theorem 27 to h_2 to conclude that there exists C_7 such that for all $0 < \gamma < r_0$,

$$d_{C^0}(\mathrm{Id}, h_2) < C_7\big(\eta^{1-\tau} + C_6(\varepsilon_2 + \eta^{1-\tau})\gamma + \eta^{1-\tau}\gamma^{-1}\big).$$

But $h_2 = R_2^{-1}R_1^{-1}I^{-1}f$, so

$$d_{C^0}(R, f) < C_8\big(\eta^{1-\tau} + C_6(\varepsilon_2 + \eta^{1-\tau})\gamma + \eta^{1-\tau}\gamma^{-1}\big). \tag{60}$$

We now obtain the high regularity estimate (23), via interpolation. By similarly moving the isometries from one slot to the other, (59) gives

$$d_{C^k}(R, f) < C_9(\eta^{1-\tau} + d_{C^k}(I, f)). \tag{61}$$

There exists $\epsilon_7 > 0$ such that if $d_{C^k}(I, f) < \epsilon_7$ and $\eta < \epsilon_7$, then the right hand side of (61) is at most 1.

We now apply the interpolation inequality in Lemma 52 and interpolate between the C^0 and C^k distance to estimate $d_{C^\ell}(R, f)$. Write $\ell = (1 - \lambda')k$ for some λ' and note that $1 > \lambda' > \lambda$ by (51). We use the estimate (60) to estimate the C^0 norm and use 1 to estimate the C^k norm, which we may do because $r < \epsilon_7$. Thus there exists C_{10} such that for $0 < \gamma < r_0$,

$$d_{C^\ell}(R, f) < C_{10}(\eta^{1-\tau}\gamma^{-1} + \varepsilon_2\gamma)^{\lambda'}. \tag{62}$$

Note that there exists $C_{11} > 0$ such that $\|f^*g - g\|_{H^0} \leq C_{11}\varepsilon_2$. Consequently, there exists a constant C_{13} such that $C_{12}\sqrt{\eta/\varepsilon_2}$ is less than the cutoff r_0. We take γ to equal $C_{12}\sqrt{\eta/\varepsilon_2}$ in (62), which gives

$$d_{C^\ell}(R, f) < C_{13}(\eta^{1/2-\tau}\varepsilon_2^{1/2} + \eta^{1/2}\varepsilon_2^{1/2})^{\lambda'} < C_{14}(\eta^{\lambda/2-\lambda\tau}\varepsilon_2^{\lambda/2} + \eta^{\lambda/2}\varepsilon_2^{\lambda/2}). \tag{63}$$

Hence by our choice of λ and τ in (50) and because $\eta < r < 1$,

$$d_{C^\ell}(R, f) < C_{15}\eta^{1/2-\sigma}\varepsilon_2^{1/2-\sigma}, \tag{64}$$

which establishes (23) and finishes the proof. ■

6. KAM scheme

In this section we develop the KAM scheme and prove that it converges. A KAM scheme is an iterative approach to constructing a conjugacy between two systems in the C^∞ setting. We begin by discussing the smoothing operators that will be used in the scheme. Then we state a lemma, Lemma 39, that summarizes the results of performing a step in the scheme. We then prove in Theorem 1 that by iterating the single KAM step we obtain the convergence needed for this theorem. We conclude the section with a final corollary of the KAM scheme which gives an asymptotic relationship between the top exponent, the bottom exponent, and the sum of all the exponents.

6.1. One step in the KAM scheme

In the KAM scheme, we begin with a tuple $(R_1, \dots, R_m)$ of isometries and a nearby tuple $(f_1, \dots, f_m)$ of diffeomorphisms. We want to find a diffeomorphism ϕ such that for all i, $\phi^{-1} f_i \phi = R_i$. However, such a ϕ may not exist.

We will then attempt to construct a conjugacy ϕ that has the following property. Let $\widetilde{f_i}$ equal $\phi^{-1} f_i \phi$. If we consider the tuple $(\widetilde{f_1}, \dots, \widetilde{f_m})$ and $(R_1, \dots, R_m)$, we can arrange that the error term, $\mathcal{U}$, in Proposition 26, is small. Once we know that the error term is small, the estimate in Proposition 26 shows that small Lyapunov exponents imply that each $\widetilde{f_i}$ has small strain. Then by Proposition 28, small strain implies that there exist R_i' such that each $\widetilde{f_i}$ is near to that R_i'. We then apply the same process to the tuples $(\widetilde{f_1}, \dots, \widetilde{f_m})$ and $(R_1', \dots, R_m')$.

The previous paragraph contains the core idea of the KAM scheme. Following this scheme, one encounters a common technical difficulty inherent in KAM arguments: regularity. In our case, this problem is most crucial when we construct the conjugacy ϕ. There is not a single choice of ϕ, but rather a family depending on a parameter λ. The parameter λ controls how smooth ϕ is. Larger values of λ give less regular conjugacies. We refer to this as a *conjugation of cutoff* λ; the formal construction appears in the proof in Lemma 39, which also gives estimates following from this construction. The nth time we iterate this procedure we will use a particular value λ_n as our cutoff. The proof of Theorem 1 shows how to pick the sequence λ_n so that the procedure converges.

We now introduce the smoothing operators. Suppose that M is a closed Riemannian manifold. As before, let Δ denote the Casimir Laplacian on M as in Section 2.4. As Δ is self-adjoint, it decomposes the space of L^2 vector fields into subspaces depending on the particular eigenvalue associated to that subspace. We call these subspaces H_λ. For a vector field X, we may write $X = \sum_\lambda X_\lambda$, where $X_\lambda \in H_\lambda$ is the projection of X onto the λ-eigenspace of Δ. All of the eigenvalues of Δ are positive. By removing the components of X that lie in high eigenvalue subspaces, we are able to smooth X. Let $\mathcal{T}_\lambda X = \sum_{\lambda' < \lambda} X_{\lambda'}$ equal the projection onto the modes strictly less than λ in magnitude. Let $\mathcal{R}_\lambda X = \sum_{\lambda' \geq \lambda} X_{\lambda'}$ be the projection onto the modes of magnitude greater than or equal to λ. Then $X = \mathcal{T}_\lambda X + \mathcal{R}_\lambda X$.

We record two standard estimates which may be obtained by application of the Sobolev embedding theorem. For $s \geq 0$, there exists a constant $C_s > 0$ such that for any $\overline{s} \geq s$ and any C^∞ vector field X on M,

$$\|\mathcal{T}_\lambda X\|_{C^{\overline{s}}} \leq C_s \lambda^{k_3 + (\overline{s} - s)/2} \|X\|_{C^s}, \tag{65}$$

$$\|\mathcal{R}_\lambda X\|_{C^s} \leq C_s \lambda^{k_3 - (\overline{s} - s)/2} \|X\|_{C^{\overline{s}}}. \tag{66}$$

The smoothing operators and the above estimates on them are useful because without smoothing, certain estimates appearing in the KAM scheme become unusable. One may see this by considering what happens in the proof of Lemma 39 if one removes the smoothing operator $\mathcal{T}_\lambda$ from (73).

The proof of the following lemma should be compared with [10, Sec. 3.4]

Lemma 39. *Suppose that (M^d, g) is a closed isotropic Riemannian manifold other than S^1. There exists a natural number l_0 such that for any $\ell > l_0$, any (C, α, n_0) and any sufficiently small $\sigma > 0$, there exist a constant $r_\ell > 0$ and numbers k_0, k_1, k_2 such that for any $s > \ell$ and any m there exist constants $C_{s,\ell}, r_{s,\ell} > 0$ such that the following holds. Suppose that $(R_1, \dots, R_m)$ is a (C, α, n_0)-Diophantine tuple with entries in* $\operatorname{Isom}(M)$ *and $(f_1, \dots, f_m)$ is a collection of C^∞ diffeomorphisms of M. Suppose that the random dynamical system generated by $(f_1, \dots, f_m)$ has stationary measures with arbitrarily small (in magnitude) bottom exponent. Write ε_k for* $\max_i d_{C^k}(f_i, R_i)$. *If $\lambda \geq 1$ is a number such that*

$$\lambda^{k_0} \varepsilon_{l_0} \leq r_\ell, \tag{67}$$

$$\lambda^{k_1 - s/4} \varepsilon_s + \varepsilon_{l_0}^{3/2} < r_{s,\ell}, \tag{68}$$

then there exists a smooth diffeomorphism ϕ and a new tuple $(R_1', \dots, R_m')$ of isometries of M such that for all i, setting $\widetilde{f}_i = \phi f_i \phi^{-1}$ we have

$$d_{C^\ell}(\widetilde{f}_i, R_i') \leq C_{s,\ell}(\lambda^{k_1 - s/10} \varepsilon_s^{1-\sigma} + \varepsilon_{l_0}^{9/8}), \tag{69}$$

$$d_{C^0}(R_i, R_i') \leq C_{s,\ell}(\varepsilon_{l_0} + (\lambda^{k_1 - s/4} \varepsilon_s + \varepsilon_{l_0}^{3/2})^{1-\sigma}), \tag{70}$$

$$d_{C^s}(\widetilde{f}_i, R_i') \leq C_{s,\ell} \lambda^{k_2} \varepsilon_s, \tag{71}$$

$$d_{C^s}(\phi, \mathrm{Id}) \leq C_{s,\ell} \lambda^{k_2} \varepsilon_s. \tag{72}$$

The diffeomorphism ϕ is called a conjugation of cutoff λ.

Proof. As in (10) in Section 3.1, let Y_i be the smallest vector field on Y_i satisfying $\exp_{R(x)} Y_i(x) = f_i(x)$. Let $\mathcal{L}$ be the operator on vector fields defined by $\mathcal{L}(Z) = m^{-1} \sum_{i=1}^m (R_i)_* Z$ as in Proposition 21. Let

$$V := -(\mathrm{Id} - \mathcal{L})^{-1} \left(\frac{1}{m} \sum_i \mathcal{T}_\lambda Y_i \right) \tag{73}$$

and let $\widetilde{f}_i = \psi_V f_i \psi_V^{-1}$. Let $\widetilde{\varepsilon}_k = \max_i d_{C^k}(\widetilde{f}_i, R_i)$ and let $\widetilde{Y}_i$ be the pointwise smallest vector field such that $\exp_{R(x)} \widetilde{Y}_i(x) = \widetilde{f}_i(x)$. By Proposition 43, for a C^1 small vector field V,

$$\widetilde{Y}_i = Y_i + V - R_i V + Q(Y_i, V), \tag{74}$$

where Q is quadratic in the sense of Definition 42. By Proposition 16, we see that $\|V\|_{C^k} \leq C_k \varepsilon_{k+\alpha}$ for some fixed α. There exist β, D_1 such that $\|Q(Y_i, V)\|_{C^k} \leq D_k \varepsilon_{k+\beta}^2$. By estimating the terms in (74), it follows that for each $k > 0$, if $\varepsilon_{k+\alpha+\beta} < 1$ then there exists a constant $D_{2,k}$ such that

$$d_{C^k}(\widetilde{f}_i, R_i) < D_{2,k} \varepsilon_{k+\alpha+\beta}. \tag{75}$$

Let μ be an ergodic stationary measure on M for the tuple $(\widetilde{f}_1, \dots, \widetilde{f}_m)$ as in the statement of the lemma. We now apply Proposition 26 with $r = d - 1, d$ and recall why the hypotheses of that proposition are satisfied. First, by our assumption that M is isotropic, $\mathrm{Isom}(M)^\circ$ acts transitively on M and $\mathrm{Gr}_1(M)$. We have also assumed the tuple $(R_1, \dots, R_m)$ is Diophantine. The nearness of $(\widetilde{f}_1, \dots, \widetilde{f}_m)$ to $(R_1, \dots, R_m)$ is guaranteed by (75), a sufficiently small choice of r_ℓ, and sufficiently large choice of l_0 by (67) as $\lambda \geq 1$. Thus by applying Proposition 26 to the conjugate system, there exists k_1 such that, in the language of that proposition,

$$\Lambda_r(\mu) = \frac{-r}{2dm} \sum_{i=1}^{m} \int_M \|E_C^{\widetilde{f}_i}\|^2 + \frac{r(d-r)}{(d+2)(d-1)m} \sum_{i=1}^{m} \int_M \|E_{NC}^{\widetilde{f}_i}\|^2 \, d\mathrm{vol} + \int_{G_r(M)} \mathcal{U}(\psi_r) \, d\mathrm{vol} + O(\|\widetilde{Y}\|_{C^{k_1}}^3),$$

where $\psi_r(x) = \frac{1}{m} \sum_{i=1}^{m} \ln\det(D_x \widetilde{f}_i \mid E_x)$ and $\mathcal{U}$ is defined in Proposition 23.

Pick a sequence of ergodic stationary measures μ_n so that $|\lambda_d(\mu_n)| \to 0$. Subtracting the expression for $\Lambda_{d-1}(\mu_n)$ from the expression for $\Lambda_d(\mu_n)$, we obtain

$$\begin{aligned}\lambda_d(\mu_n) &= \Lambda_d(\mu_n) - \Lambda_{d-1}(\mu_n) \\ &= \frac{-1}{2dm} \sum_{i=1}^{m} \int_M \|E_C^{\widetilde{f}_i}\|^2 \, d\mathrm{vol} + \frac{-(d-1)}{(d+2)(d-1)m} \sum_{i=1}^{m} \int_M \|E_{NC}^{\widetilde{f}_i}\|^2 \, d\mathrm{vol} \\ &\quad - \int_{\mathrm{Gr}_{d-1}(M)} \mathcal{U}(\psi_{d-1}) \, d\mathrm{vol} + \int_{\mathrm{Gr}_d(M)} \mathcal{U}(\psi_d) \, d\mathrm{vol} + O(\|\widetilde{Y}\|_{C^{k_1}}^3). \quad (76)\end{aligned}$$

Write $\mathrm{Gr}_r(R)$ for the map on $\mathrm{Gr}_r(M)$ induced by R. Write $\mathbf{Y}_i$ for the shortest vector field on $\mathrm{Gr}_r(M)$ such that $\exp_{\mathrm{Gr}_r(R_i)(x)} \mathbf{Y}_i = \mathrm{Gr}_r(\widetilde{f}_i)(x)$. By Lemma 56, for each k there exists $C_{1,k}$ such that

$$\Big\|\sum_{i=1}^{m} \mathbf{Y}_i\Big\|_{C^k} \leq C_{1,k}\Big(\Big\|\sum_{i=1}^{m} \widetilde{Y}_i\Big\|_{C^{k+1}} + \widetilde{\varepsilon}_{k+1}^2\Big).$$

Hence by the above line and the final estimate in Proposition 23 there exists k_2 such that

$$\left|\int_{\mathrm{Gr}_r(M)} \mathcal{U}(\psi_r) \, d\mathrm{vol}\right| \leq C_2 \|\psi_r\|_{C^{k_2}} \left(\Big\|\frac{1}{m} \sum_{i=1}^{m} \widetilde{Y}_i\Big\|_{C^{k_2}} + \|\widetilde{Y}_i\|_{C^{k_2}}^2\right). \quad (77)$$

The term $\|\psi_r\|_{C^{k_2}}$ is bounded by a constant times $\widetilde{\varepsilon}_{k_2}$. By using (74) we may rewrite the second term appearing in the product in (77):

$$\begin{aligned}\frac{1}{m} \sum_{i=1}^{m} \widetilde{Y}_i &= \frac{1}{m} \sum_i Y_i - (\mathrm{Id} - \mathcal{L})^{-1}\Big(\frac{1}{m} \sum_i \mathcal{T}_\lambda Y_i\Big) - \frac{1}{m} \sum_i (R_i)_*(-(\mathrm{Id} - \mathcal{L})^{-1})(\mathcal{T}_\lambda Y_i) \\ &\quad + \frac{1}{m} \sum_i Q(Y_i, V) \\ &= \frac{1}{m} \sum_i \mathcal{R}_\lambda Y_i + \frac{1}{m} \sum_i \mathcal{T}_\lambda Y_i - (\mathrm{Id} - \mathcal{L})(\mathrm{Id} - \mathcal{L})^{-1}\Big(\frac{1}{m} \sum_i \mathcal{T}_\lambda Y_i\Big) \\ &\quad + \frac{1}{m} \sum_i Q(Y_i, V)\end{aligned}$$

$$= \frac{1}{m}\sum_i \mathcal{R}_\lambda Y_i + \frac{1}{m}\sum_i \mathcal{T}_\lambda Y_i - \frac{1}{m}\sum_i \mathcal{T}_\lambda Y_i + \frac{1}{m}\sum_i Q(Y_i, V)$$
$$= \frac{1}{m}\sum_i \mathcal{R}_\lambda Y_i + \frac{1}{m}\sum_i Q(Y_i, V).$$

By (66), there exists k_3 such that for all $s \geq 0$,

$$\|R_\lambda Y_i\|_{C^1} \leq C_{3,s}\lambda^{k_3 - s/2}\|Y_i\|_{C^s}.$$

As the Q term is quadratic, there exist ℓ_2, C_4 such that

$$\|Q(Y_i, V)\|_{C^{k_2}} \leq C_4\|Y_i\|_{C^{\ell_2}}\|V\|_{C^{\ell_2}} = C_4\|Y_i\|_{C^{\ell_2}}\|(\mathrm{Id} - \mathcal{L})^{-1}(\mathcal{T}_\lambda Y_i)\|_{C^{\ell_2}} \leq C_5\varepsilon_{\ell_3}^2$$

for some ℓ_3 by Proposition 21. Thus

$$\left\|\frac{1}{m}\sum_i \widetilde{Y}_i\right\|_{C^{k_2}} \leq C_{6,s}(\lambda^{k_3 - s/2}\varepsilon_s + \varepsilon_{\ell_3}^2).$$

Finally, by (75) we have $\|\widetilde{Y}_i\|_{C^{k_2}} \leq C_7\varepsilon_{\ell_3}$ as before. Let $\ell_4 = \max\{\ell_3, k_2 + \alpha + \beta\}$. Applying all of these estimates to (77) gives

$$\left|\int_{\mathrm{Gr}_r(M)} \mathcal{U}(\psi_r)\, d\mathrm{vol}\right| \leq C_{8,s}\varepsilon_{k_2}(\lambda^{k_3 - s/2}\varepsilon_s + \varepsilon_{\ell_4}^2). \tag{78}$$

By taking $\ell_5 > \max\{k_1 + \alpha + \beta, k_2, \ell_4\}$, using $\lambda_d(\mu_n) \to 0$,[2] and combining (78) and (76) we find for $s \geq 0$ that there exists $C_{9,s}$ such that

$$C_{9,s}(\lambda^{k_3 - s/2}\varepsilon_s\varepsilon_{\ell_5} + \varepsilon_{\ell_5}^3)$$
$$\geq \frac{1}{2dm}\sum_{i=1}^m \int_M \|E_C^{\widetilde{f}_i}\|^2\, d\mathrm{vol} + \frac{(d-1)}{(d+2)(d-1)m}\sum_{i=1}^m \int_M \|E_{NC}^{\widetilde{f}_i}\|^2\, d\mathrm{vol}. \tag{79}$$

Note that the coefficients of each of the strain terms are positive. If $s > \ell_5$, then by taking square roots, we see that there exist constants $C_{10,s}$ such that for each i,

$$C_{10,s}(\lambda^{k_3/2 - s/4}\varepsilon_s + \varepsilon_{\ell_5}^{3/2}) \geq \|\widetilde{f}_i^* g - g\|_{H^0}. \tag{80}$$

We now give a naive estimate on the higher C^s norms under the assumption that ε_1 is bounded by a constant $\epsilon_1 > 0$. To begin, by combining (65) and Proposition 16 we see that there exists $\alpha > 0$ such that for each s there exists $D_{3,s}$ such that $\|V\|_{C^s} \leq D_{3,s}\lambda^\alpha\varepsilon_s$. Hence by Lemma 48, both $d_{C^s}(\psi_V, \mathrm{Id})$ and $d_{C^s}(\psi_V^{-1}, \mathrm{Id})$ are bounded by $D_{4,s}\lambda^\alpha\varepsilon_s$. This establishes (72).

[2]Note that we do not need $\lambda(\mu_n) \to 0$ in order to deduce (79). It suffices to know that there is μ such that $\lambda_d(\mu)$ is comparable to the right hand side of (78). This observation is the essence of the proof of Theorem 40.

Now applying the composition estimate from Lemma 50, we find that assuming $\lambda \geq 1$,

$$\begin{aligned} d_{C^s}(f \circ \psi_V^{-1}, R) &\leq C_{11,s}\big(d_{C^s}(f, R) + d_{C^s}(\psi_V^{-1}, \mathrm{Id})\big) \\ &\leq C_{12,s}(\varepsilon_s + \lambda^\alpha \varepsilon_s) \leq C_{13,s}(\lambda^\alpha \varepsilon_s). \end{aligned}$$

We then apply the other estimate in Lemma 50 to find that

$$\begin{aligned} d_{C^s}(\psi_V \circ f \circ \psi_V^{-1}, R) &\leq C_{11,s}\big(d_{C^s}(\psi_V, \mathrm{Id}) + d_{C^s}(f \circ \psi_V^{-1}, R)\big) \\ &\leq C_{14,s}(\lambda^\alpha \varepsilon_s + \lambda^\alpha \varepsilon_s) \leq C_{15,s}\lambda^\alpha \varepsilon_s. \end{aligned}$$

Hence under an assumption of the type (67), namely $\varepsilon_1 < \epsilon_1$, we may conclude

$$d_{C^s}(\widetilde{f}_i, R) \leq C_{15,s}\lambda^\alpha \varepsilon_s, \tag{81}$$

which establishes (71).

We now apply Proposition 28 to this system. Let k_σ and r_σ be the k and r in Proposition 28 for a given choice of σ and our fixed ℓ. In preparation for the application of the lemma, we record some basic estimates:

(1) By combining (65) and Proposition 21 as before, we see that there exists ℓ_6 such that

$$d_{C^2}(\widetilde{f}_i, R_i) \leq \varepsilon_{\ell_6}. \tag{82}$$

(2) From the previous discussion we also have

$$\|\widetilde{f}_i^* g - g\|_{H^0} \leq C_{10,s}(\lambda^{k_3/2 - s/4}\varepsilon_s + \varepsilon_{\ell_5}^{3/2}).$$

(3) We also need the C^{k_σ} estimate

$$d_{C^{k_\sigma}}(\widetilde{f}_i, R) \leq C_{15,k_\sigma}\lambda^\alpha \varepsilon_{k_\sigma}.$$

Hence if

$$C_{15,k_\sigma}\lambda^\alpha \varepsilon_{k_\sigma} < r_\sigma, \tag{83}$$

$$C_{10,s}(\lambda^{k_3/2 - s/4}\varepsilon_s + \varepsilon_{\ell_5}^{3/2}) \leq r_\sigma, \tag{84}$$

then by Proposition 28 and the previous estimates there exist C_6 and isometries R_i' such that

$$d_{C^\ell}(\widetilde{f}_i, R_i') \leq C_{16,s}(\lambda^{k_3/2 - s/4}\varepsilon_s + \varepsilon_{\ell_5}^{3/2})^{1/2-\sigma}\varepsilon_{\ell_6}^{1/2-\sigma}, \tag{85}$$

$$d_{C^0}(R_i', R_i) < C_{17,s}(\varepsilon_{\ell_6} + (\lambda^{k_3/2 - s/4}\varepsilon_s + \varepsilon_{\ell_5}^{3/2})^{1-\sigma}). \tag{86}$$

Let $\ell_7 = \max\{\ell_5, \ell_6\}$. If $s > \ell_7$, then (85) implies

$$d_{C^\ell}(\widetilde{f}_i, R_i') \leq C_{16,s}(\lambda^{k_4 - s/9}\varepsilon_s^{1-2\sigma} + \varepsilon_{\ell_7}^{5/4 - (5/2)\sigma}),$$

which yields (69) under the assumption that $\sigma > 0$ is sufficiently small. Note that (86) establishes (70). Thus we are done as we have established these estimates assuming only bounds of the type appearing in (67) and (68). ■

Remark 1. In the above lemma, we could have assumed instead that there exist stationary measures for which both the top exponent and the sum of all exponents were arbitrarily small and concluded the same result. The reason is that if we had considered $\Lambda_1 - \Lambda_d$ in (76), the coefficients of the strain terms would still have the same sign and so we could conclude the same result. By related modifications, one can produce many other formulations of the main result in [10] that require other hypotheses on the Lyapunov exponents.

6.2. Convergence of the KAM scheme

In this section we prove the main linearization theorem. It is helpful to note that the approach to this theorem is somewhat different from the classical approach to KAM type results. In a classical argument, one might typically linearize the problem at a target isometric system and then find a solution to the linearized problem. In our case, while we are able to linearize the problem, the resulting linearized problem does not obviously have any solution. Consequently, we must give dynamical and geometric arguments that show that a related type of averaged linearized problem can be solved and that solving this averaged problem is indeed helpful. This then allows us to make progress in the KAM scheme by conjugating the system closer to an isometric one. In particular, note that in our case we do not know from the outset which isometric system our random system will ultimately be conjugate to.

Theorem 1. *Let M^d be a closed isotropic Riemannian manifold other than S^1. There exists k_0 such that if $(R_1, \dots, R_m)$ is a tuple of isometries of M such that the subgroup of* $\operatorname{Isom}(M)$ *generated by this tuple contains* $\operatorname{Isom}(M)^\circ$*, then there exists $\epsilon_{k_0} > 0$ such that the following holds. Let $(f_1, \dots, f_m)$ be a tuple of C^∞ diffeomorphisms satisfying* $\max_i d_{C^{k_0}}(f_i, R_i) < \epsilon_{k_0}$*. Suppose that there exists a sequence of ergodic stationary measures μ_n for the random dynamical system generated by $(f_1, \dots, f_m)$ such that* $|\lambda_d(\mu_n)| \to 0$*. Then there exists $\psi \in \operatorname{Diff}^\infty(M)$ such that for each i the map $\psi f_i \psi^{-1}$ is an isometry of M and lies in the subgroup of* $\operatorname{Isom}(M)$ *generated by* $(R_1, \dots, R_m)$.

Before giving the proof, we sketch briefly the argument, which is typical of arguments establishing the convergence of a KAM scheme. In a KAM scheme where one wishes to show that some sequence of objects h_n converges, there are often two parts. The first part of the proof is an inductive argument obtaining a sequence of estimates by the repeated application of the KAM step, which in our case is Lemma 39. The second half of the proof checks that the repeated application of the KAM step is valid by showing that we never leave the neighborhood of its validity and then checks that the procedure is converging in C^∞.

In the first part, one inductively produces a sequence of estimates by iterating a KAM step. The estimates produced usually come in two forms: a single good estimate in a low norm and bad estimates in high norms. The low regularity estimate probably looks like $\|h_n\|_{C^0} \le N^{-(1+\tau)^n}$ where $\tau > 0$, while for every s one has a high regularity estimate like $\|h_n\|_{C^s} \le N^{(1+\tau)^n}$. A priori, the h_n become superexponentially C^0 small, yet might be diverging in higher C^s norms. To remedy this situation one then interpolates between the low and high norms by using an equality derived from Lemma 52. In this case such an inequality for the objects h_n might assert something like

$$\|h_n\|_{C^{\lambda\cdot 0+(1-\lambda)s}} \leq C_s \|f\|_{C^0}^{\lambda} \|f\|_{C^s}^{1-\lambda}.$$

If λ is sufficiently close to 1 and s is sufficiently large, a brief calculation then implies that the $C^{(1-\lambda)s}$ norm is also superexponentially small. By changing s and λ one then obtains convergence in C^∞.

Proof of Theorem 1. The proof is by a KAM convergence scheme. To begin we introduce the Diophantine condition we will use. By Proposition 19, $(R_1, \ldots, R_m)$ is (C', α', n')-Diophantine for some $C', \alpha' > 0$ and is stably so. By stability, there exist (C, α, n) and a C^0 neighborhood $\mathcal{U}$ of $(R_1, \ldots, R_m)$ such that any tuple in $\mathcal{U}$ is also (C, α, n)-Diophantine. Hence if $(R_1', \ldots, R_m') \in \mathcal{U}$, then the coefficients $C_{i,s}$ appearing in Lemma 39 are uniform over all of these tuples. Assuming we do not leave the set $\mathcal{U}$, the constants appearing in Lemma 39 will be uniform. We check this at the end of the proof in the discussion surrounding (91).

We now show that there exists a sequence of cutoffs λ_n such that if we repeatedly apply Lemma 39 with the cutoff λ_n on the nth time we apply the lemma, then the resulting sequence of conjugates converges and the hypotheses of Lemma 39 remain satisfied. Given such a sequence λ_n the convergence scheme is run as follows. Let $(f_{1,1}, \ldots, f_{m,1}) = (f_1, \ldots, f_m)$ and let $(R_{1,1}, \ldots, R_{m,1}) = (R_1, \ldots, R_m)$. Given $(f_{1,n-1}, \ldots, f_{m,n-1})$ and $(R_{1,n-1}, \ldots, R_{m,n-1})$ we apply Lemma 39 with cutoff $\lambda = \lambda_n$ to produce a diffeomorphism ϕ_n and a tuple of isometries that we denote by $(R_{1,n}, \ldots, R_{m,n})$. We set $f_{i,n} = \phi_n f_{i,n-1} \phi_n^{-1}$ to obtain a new tuple $(f_{1,n}, \ldots, f_{m,n})$ of diffeomorphisms. We write ψ_n for $\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_1$, so that $f_{i,n} = \psi_n \circ f_i \circ \psi_n^{-1}$. Let $\varepsilon_{k,n} = \max_i d_{C^k}(f_{i,n}, R_{i,n})$.

We now show that such a sequence of cutoffs λ_n exists. Let σ be a small positive number and let l_0 and ϵ_{l_0} be as in Lemma 39. Let $k_0, k_1, k_2, r_\ell, C_{s,\ell}, r_{s,\ell}$ be as in Lemma 39 as well. To show that such a sequence of cutoffs λ_n exists we must also provide a fixed choice of s, ℓ for the application of Lemma 39. We will first show that the scheme converges in the C^{l_0} norm and then bootstrap to get C^∞ convergence. Fix some arbitrary $\ell > l_0$. The choice of ℓ does not matter in what follows because we will only consider estimates on the l_0 norm. We will choose s such that

$$s > \ell. \tag{87}$$

Further, if s is sufficiently large and τ is sufficiently small, then we can pick α such that

$$\frac{2+\tau}{s/4 - k_1} < \alpha < \min\{1/k_0, \tau/k_2\}. \tag{88}$$

So, we increase s if needed and choose such a τ satisfying

$$1/8 > \tau > 0. \tag{89}$$

Pick s, α, τ so that (87)–(89) are all satisfied.

Let $\lambda_n = N^{\alpha(1+\tau)^n}$ for some N we choose later. We will show that with this choice of cutoff at the nth step, the KAM scheme converges. In order to show this, we show that the following two estimates hold inductively given a choice of sufficiently large N:

$$\varepsilon_{l_0,n} \le N^{-(1+\tau)^n}, \tag{H1}$$

$$\varepsilon_{s,n} \le N^{(1+\tau)^n}, \tag{H2}$$

$$\max_i d_{C^0}(R_{i,n}, R_{i,1}) \le \sum_{i=1}^{n} N^{-\frac{1}{2}(1+\tau)^i}. \tag{H3}$$

This involves two arguments. The first argument shows that there is a sufficiently large N such that if we have these estimates for n, then the hypotheses of Lemma 39 are satisfied. The second argument is the actual induction, which checks that if (H1) and (H2) hold for n then they also hold for $n+1$, i.e. we apply Lemma 39 and then deduce (H1) and (H2) for $n+1$ from this.

We begin by checking that for all sufficiently large $N > 0$ and any $n \in \mathbb{N}$, if (H1)–(H3) are satisfied, then the hypotheses of Lemma 39 are satisfied as well. To begin, as the summation in (H3) converges superexponentially as $n \to \infty$, for all sufficiently large N, the tuple $(R_{1,n}, \dots, R_{m,n})$ lies in $\mathcal{U}$. The first numbered hypothesis of Lemma 39 is (67):

$$\lambda_n^{k_0} \varepsilon_{l_0,n} \le r_\ell.$$

Given the choice of λ_n, if (H1) and (H2) hold it suffices to have

$$N^{\alpha k_0 (1+\tau)^n} N^{-(1+\tau)^n} < r_\ell,$$

which holds for N sufficiently large and all n by our choice of α. The other hypothesis of Lemma 39, (68), requires that

$$\lambda_n^{k_1 - s/4} \varepsilon_{s,n} + \varepsilon_{l_0,n}^{3/2} < r_{s,\ell}.$$

Given (H1) and (H2) and our choice of λ_n it suffices to have

$$N^{\alpha(k_1 - s/4)(1+\tau)^n} N^{(1+\tau)^n} + N^{-\frac{3}{2}(1+\tau)^n} < r_{s,\ell}.$$

Our choice of s and α implies that $\alpha(k_1 - s/4) < -1$, hence the above inequality holds for sufficiently large N. Thus the two hypotheses of Lemma 39 follow from (H1) and (H2). Thus we may apply Lemma 39 given (H1)–(H3) and our choice of N.

We now proceed to the inductive argument. We will show that for all N sufficiently large, if we now require that our perturbation is small enough that (H1) and (H2) hold for $n = 1$ and our choice of N, we may continue applying Lemma 39, and these estimates as well as (H3) continue to hold. Note that (H3) is trivial when $n = 1$. We must then check that (H1)–(H3) are satisfied for $n+1$ given they hold for n. By the previous paragraph, we are free to apply the estimates from Lemma 39 as long as N is sufficiently large.

We now check that (H1) holds for $n+1$. By (69), we obtain

$$\varepsilon_{l_0,n+1} \le C_{s,\ell}(\lambda_n^{k_1 - s/10} \varepsilon_{s,n}^{1-\sigma} + \varepsilon_{l_0,n}^{9/8}).$$

By applying (H1) and (H2) to each term on the right it suffices to show

$$C_{s,\ell}(N^{\alpha(k_1-s/10)(1+\tau)^n}N^{(1-\sigma)(1+\tau)^n}+N^{-(9/8)(1+\tau)^n})<N^{-(1+\tau)^{n+1}}. \tag{90}$$

By our choice of s, α, and τ, the lower bound in (88) implies that

$$\alpha(k_1-s/10)+(1-\sigma)<-(1+\tau).$$

In addition, by (89), $-9/8<-(1+\tau)$. Thus for sufficiently large N the left hand side of (90) is bounded above by $N^{-(1+\tau)^{n+1}}$.

Next we check (H2) holds for $n+1$. By (71),

$$\varepsilon_{s,n+1}\le C_{s,\ell}\lambda_n^{k_2}\varepsilon_{s,n}.$$

Hence,

$$\varepsilon_s\le C_{s,\ell}N^{k_2\alpha(1+\tau)^n}N^{(1+\tau)^n}.$$

By (88), $1+k_2\alpha<1+\tau$ and hence, assuming N is sufficiently large, the right hand side is bounded by $N^{(1+\tau)^{n+1}}$, which shows that (H2) is satisfied.

We now check (H3). This follows easily by the application of (70), which gives

$$d_{C^0}(R_{i,n},R_{i,n+1})\le C_{s,\ell}\big(\varepsilon_{l_0,n}+(\lambda_n^{k_1-s/4}\varepsilon_{s,n}+\varepsilon_{l_0,n}^{3/2})^{1-\sigma}\big). \tag{91}$$

Applying (H1) and (H2) and the definition of λ_n to estimate the right hand side of (91), we find that for the γ given in (H3) and N sufficiently large,

$$d_{C^0}(R_{i,n},R_{i,n+1})\le N^{-\frac{1}{2}(1+\tau)^n}, \tag{92}$$

and (H3) holds for $n+1$.

We have now finished the induction but not the proof. We have shown that there exists a sequence λ_n and a choice s,α,ℓ,τ,N such that if the initial conditions of the scheme are satisfied then we may iterate indefinitely and be assured of the estimates in (H1)–(H3) at each step. We must now check that the conjugacies ψ_n are converging in C^∞ and that the tuples $(R_{1,n},\dots,R_{m,n})$ are converging. The latter is immediate because by (92) this is a Cauchy sequence. In fact, we chose N large enough that we never leave $\mathcal{U}$, hence the limit is in $\mathcal{U}$. As the group of isometries of M is C^0 closed and the distance from the tuples $(f_{1,n},\dots,f_{m,n})$ to a tuple of isometries is converging to 0, it follows that $(f_{1,n},\dots,f_{m,n})$ is converging to a tuple of isometries. To show that the ψ_n converge in C^∞, we obtain for every s an estimate on $d_{C^s}(\phi_n,\mathrm{Id})$. By a similar induction to that just performed, the estimate (72) implies

$$d_{C^s}(\phi_n,\mathrm{Id})\le C_sN^{(1+\tau)^n}.$$

Let $j>0$ be an integer. By Lemma 53, interpolating with $\lambda=1-1/10$ between the C^{l_0} distance and the C^{jl_0} distance of ϕ_n to the identity gives

$$d_{C^{.9l_0+(j/10)l_0}}(\phi_n,\mathrm{Id})\le C_jN^{-.9(1+\tau)^n}N^{.1(1+\tau)^n}=C_jN^{-.8(1+\tau)^n}.$$

Thus by increasing j, we see that there exists $\tau' > 0$ such that for each s,

$$d_{C^s}(\phi_n, \mathrm{Id}) < C_s' N^{-(1+\tau')^n}.$$

The previous line is summable in n. Hence we can apply Lemma 51 to obtain convergence of the $\psi_n = \phi_n \circ \cdots \circ \phi_1$ in the C^s norm for each s and thus C^∞ convergence.

Thus we see that we have simultaneously conjugated all f_i into $\mathrm{Isom}(M)$. In order to obtain the full theorem, we must be assured that $\psi^{-1} f_i \psi$ lies in the subgroup of $\mathrm{Isom}(M)$ generated by $(R_1, \dots, R_m)$. Note that $\mathrm{Isom}(M)/\mathrm{Isom}(M)^\circ$ is a finite group and that ψ is homotopic to the identity by construction. Thus we see that the image of the group generated by $(\psi^{-1} f_1 \psi, \dots, \psi^{-1} f_m \psi)$ in $\mathrm{Isom}(M)/\mathrm{Isom}(M)^\circ$ is the same as the image of the group generated by $(R_1, \dots, R_m)$. By our choice of N, the tuple $(\psi^{-1} f_1 \psi, \dots, \psi^{-1} f_m \psi)$ is in $\mathcal{U}$ and thus generates $\mathrm{Isom}(M)^\circ$. Thus the original tuple and the new one generate the same subgroup of $\mathrm{Isom}(M)$ and we are done. ■

6.3. Taylor expansion of Lyapunov exponents

In order to recover Dolgopyat and Krikorian's Taylor expansion in the setting of isotropic manifolds, we would need to apply Proposition 26 for each $0 \le r \le \dim M$. However, one of the hypotheses of Proposition 26 is that $\mathrm{Isom}(M)^\circ$ acts transitively on $\mathrm{Gr}_r(M)$. In Proposition 41, we see that unless M is S^n or $\mathbb{R}\mathrm{P}^n$, $\mathrm{Isom}(M)$ does not act transitively on $\mathrm{Gr}_r(M)$ for $r \neq 1$ or $d-1$. Despite Proposition 41, we are able to obtain a partial result: the greatest and least Lyapunov exponents are symmetric about the "average" Lyapunov exponent $\frac{1}{d}\Lambda_d(\mu)$.

Theorem 40. *Suppose that M^d is a closed isotropic manifold other than S^1 and that $(R_1, \dots, R_m)$ is a subset of $\mathrm{Isom}(M)$ that generates a subgroup of $\mathrm{Isom}(M)$ containing $\mathrm{Isom}(M)^\circ$. Suppose that $(f_1, \dots, f_m)$ is a collection of C^∞ diffeomorphisms of M. Then there exists k_0 such that if μ is an ergodic stationary measure of the random dynamical system generated by $(f_1, \dots, f_m)$, then*

$$\left| \lambda_1(\mu) - \left(-\lambda_d(\mu) + \frac{2}{d}\Lambda_d(\mu) \right) \right| \le o(1)|\lambda_d(\mu)|, \tag{93}$$

where the $o(1)$ term goes to 0 as $\max_i d_{C^{k_0}}(f_i, R_i) \to 0$. The $o(1)$ term depends only on $(R_1, \dots, R_m)$.

Proof. By Theorem 1, there are two cases: either $(f_1, \dots, f_m)$ is conjugate to isometries or it is not. In the isometric case (93) is immediate, so we may assume that there is an ergodic stationary measure μ with $\lambda_d(\mu)$ non-zero. The proof that follows is then essentially an observation about what happens when the KAM scheme is run on a system that has a measure with such a non-zero Lyapunov exponent. If we run the KAM scheme without assuming that $(f_1, \dots, f_m)$ has a measure with zero exponents, we can keep running the scheme until the non-trivial exponents prevent us from continuing. At a certain point in the procedure, the non-trivial exponents cause a certain inequality to fail. Using the failed inequality then gives the result.

We now give the details. Fix an ergodic stationary measure μ and consider equation (76) appearing in the KAM step:

$$\lambda_d(\mu) = \frac{-1}{2dm}\sum_{i=1}^{m}\int_M \|E_C^{\tilde{f}_i}\|^2\,d\mathrm{vol} + \frac{-(d-1)}{(d+2)(d-1)m}\sum_{i=1}^{m}\int_M \|E_{NC}^{\tilde{f}_i}\|^2\,d\mathrm{vol} - \int_{\mathrm{Gr}_{d-1}(M)} \mathcal{U}(\psi_{d-1})\,d\mathrm{vol} + \int_{\mathrm{Gr}_d(M)} \mathcal{U}(\psi_d)\,d\mathrm{vol} + O(\|\tilde{Y}\|^3_{C^{k_1}}). \tag{94}$$

This allows us to use the fact that the exponent λ_d is small in magnitude. In the KAM step, we proceed from this estimate by estimating the $\|\tilde{Y}\|^3_{C^{k_1}}$ term as well as the $\mathcal{U}$ terms. Inequality (78) and the choice of ℓ_5 imply that these terms satisfy

$$\left|\int_{\mathrm{Gr}_{d-1}(M)} \mathcal{U}(\psi_{d-1})\,d\mathrm{vol} - \int_{\mathrm{Gr}_d(M)} \mathcal{U}(\psi_d)\,d\mathrm{vol} + O(\|\tilde{Y}\|^3_{C^{k_1}})\right| \leq C_{8,s}\varepsilon_{\ell_5}(\lambda^{k_3-s/2}\varepsilon_s + \varepsilon_{\ell_5}^2). \tag{95}$$

Hence as long as

$$|\lambda_d(\mu)| < (C_{9,s} - C_{8,s})(\varepsilon_{\ell_5}(\lambda^{k_3-s/2}\varepsilon_s + \varepsilon_{\ell_5}^2)), \tag{96}$$

the proof of Lemma 39 may proceed to (79) even if there is not a sequence of measures μ_n such that $|\lambda_d(\mu_n)| \to 0$. Hence we may continue running the KAM scheme until inequality (96) fails to hold.

Suppose that we iterate the KAM scheme until (96) fails. We consider the estimates available in the KAM scheme at the step of failure. By applying Proposition 26 with r equal to 1, d, and $d-1$, we obtain

$$\begin{aligned}
\Lambda_1(\mu) &= \frac{-1}{2dm}\sum_{i=1}^{m}\int_M \|E_C^{\tilde{f}_i}\|^2\,d\mathrm{vol} + \frac{(d-1)}{(d+2)(d-1)m}\sum_{i=1}^{m}\int_M \|E_{NC}^{\tilde{f}_i}\|^2\,d\mathrm{vol} + \int_{G_1(M)} \mathcal{U}(\psi_1)\,d\mathrm{vol} + O(\|\tilde{Y}\|^3_{C^{k_1}}),\\
\Lambda_{d-1}(\mu) &= \frac{-(d-1)}{2dm}\sum_{i=1}^{m}\int_M \|E_C^{\tilde{f}_i}\|^2\,d\mathrm{vol} + \frac{(d-1)}{(d+2)(d-1)m}\sum_{i=1}^{m}\int_M \|E_{NC}^{\tilde{f}_i}\|^2\,d\mathrm{vol} + \int_{G_{d-1}(M)} \mathcal{U}(\psi_{d-1})\,d\mathrm{vol} + O(\|\tilde{Y}\|^3_{C^{k_1}}),\\
\Lambda_d(\mu) &= \frac{-d}{2dm}\sum_{i=1}^{m}\int_M \|E_C^{\tilde{f}_i}\|^2\,d\mathrm{vol} + \int_{G_d(M)} \mathcal{U}(\psi_d)\,d\mathrm{vol} + O(\|\tilde{Y}\|^3_{C^{k_1}}).
\end{aligned} \tag{97}$$

Write $\mathcal{U}_i$ as shorthand for $\int_{\mathrm{Gr}_i(M)} \mathcal{U}(\psi_i)\,d\mathrm{vol}$. Then

$$\begin{aligned}
\lambda_1(\mu) - \left(-\lambda_d(\mu) + \frac{2}{d}\Lambda_d(\mu)\right) &= \Lambda_1(\mu) - \Lambda_{d-1}(\mu) + \frac{d-2}{d}\Lambda_d(\mu)\\
&= \mathcal{U}_1 + \mathcal{U}_{d-1} + \mathcal{U}_d + O(\|\tilde{Y}\|^3_{C^{k_1}}).
\end{aligned} \tag{98}$$

Using (78), (75), and $\ell_5 > k_1 + \alpha$, we bound the right hand side of (98) to find

$$\left|\lambda_1(\mu) - \left(-\lambda_d(\mu) + \frac{2}{d}\Lambda_d(\mu)\right)\right| \le 4C_{8,s}(\lambda_n^{k_3-s/2}\varepsilon_s\varepsilon_{\ell_5} + \varepsilon_{\ell_5}^3).$$

But by the failure of (96), we may bound the right hand side to obtain

$$\left|\lambda_1(\mu) - \left(-\lambda_d(\mu) + \frac{2}{d}\Lambda_d(\mu)\right)\right| \le \frac{4}{C_{9,s} - C_{8,s}}|\lambda_d(\mu)|. \tag{99}$$

Note that in the above inequality the larger $C_{9,s}$, the smaller the left hand side. We can take $C_{9,s}$ as large as we like and still run the KAM scheme. Running the KAM scheme with a larger $C_{9,s}$ only requires that we assume our initial perturbation is closer to the original system of rotations in the C^{k_0} norm. Hence by assuming that the initial distance is arbitrarily small in the C^{k_0} norm, we may take $C_{9,s}$ as large as we like. Thus (93) follows from (99). ∎

We now check the claim about isotropic manifolds.

Proposition 41. *Suppose that M is a closed isotropic manifold other than $\mathbb{R}\,\mathrm{P}^n$ or S^n. Then $\mathrm{Isom}(M)$ does not act transitively on $\mathrm{Gr}_k(M)$ for $k \neq 0, 1, \dim M - 1, \dim M$.*

Proof. From Section 2.5, we have a list of all closed isotropic manifolds, so we may give an argument for each of the families, $\mathbb{C}\,\mathrm{P}^n$, $\mathbb{H}\,\mathrm{P}^n$, and $F_4/\mathrm{Spin}(9)$.

The isometry group of $\mathbb{C}\,\mathrm{P}^n$ is $\mathrm{PSU}(n+1)$. If we fix a point p in $\mathbb{C}\,\mathrm{P}^n$, then the isotropy group is naturally identified with $\mathrm{SU}(n)$. It is then immediate that the action of the isotropy group preserves complex subspaces of $\mathrm{Gr}_k(\mathbb{C}\,\mathrm{P}^n)$. Consequently, $\mathrm{Isom}(\mathbb{C}\,\mathrm{P}^n)$ does not act transitively on $\mathrm{Gr}_k(\mathbb{C}\,\mathrm{P}^n)$ as $\mathbb{C}\,\mathrm{P}^n$ has subspaces that are not complex. In the case of $\mathbb{H}\,\mathrm{P}^n$, which is constructed similarly to $\mathbb{C}\,\mathrm{P}^n$, a similar argument works where we use instead the fact that the isotropy group is $\mathrm{Sp}(k)$, the compact symplectic group.

We now turn to the Cayley plane, for which we give a dimension counting argument. The dimension of F_4 is 52 while $\dim F_4/\mathrm{Spin}(9) = 16$. Recall that if M is a manifold and $\dim M = d$ then $\dim \mathrm{Gr}_k(M) = (k+1)d + k(k+1)/2$. Hence $\dim \mathrm{Gr}_3(F_4/\mathrm{Spin}(9)) > 52$. If $\mathrm{Isom}(M)$ acts transitively on 2-planes then M must have constant sectional curvature and hence is a sphere. The Cayley plane does not have constant sectional curvature, hence $k = 2$ is ruled out. Similarly, a dimension count excludes the possibility that F_4 acts transitively on $\mathrm{Gr}_k(F_4/\mathrm{Spin}(0))$ when $k \neq 0, 1, 15, 16$. ∎

Appendix A. C^k estimates

In this section of the appendix, we collect some basic results concerning the calculus of C^k functions. Most of the estimates stated here are used to compare constructions coming from Riemannian geometry and constructions coming from a chart.

Most of the estimates we prove below involve the following definition, which is an appropriate form for a second order term in the C^k setting.

Definition 42. Suppose that X, Y, Z are vector fields and $Z = Z(X, Y)$ is a function of X and Y. We say that Z is *quadratic* in X and Y if there exists a fixed ℓ such that for each k there is a constant C_k depending only on Z such that

$$\|Z\|_{C^k} \le C_k(\|X\|^2_{C^{k+\ell}} + \|Y\|^2_{C^{k+\ell}}). \tag{100}$$

In addition to quadratic, we may also refer to Z as being *second order* in X and Y. When Z depends only on X, the definition is analogous.

One thinks of (100) as a quadratic tameness estimate. Our main use of this notion is the following proposition, which allows us to compose diffeomorphisms up to a quadratic error. As before, if Y is a vector field on M, we write ψ_Y for the map of M given by $x \mapsto \exp_x Y(x)$. To emphasize that ψ depends on a metric g, we may write ψ^g_Y.

The main result of this section is the following, which is used in the KAM scheme to see how the linearized error between f_i and R_i changes when f_i is conjugated by a diffeomorphism ψ.

Proposition 43 ([10, (8)]). *Suppose that (M, g) is a closed Riemannian manifold and R is an isometry of M. Suppose that f is a diffeomorphism of M that is C^1 close to R. Let $Y(x) = \exp^{-1}_{R(x)} f(x)$. If C is a C^1 small vector field on M, then the error field $\exp^{-1}_{R(x)} \psi_C f \psi_C^{-1}$ is equal to*

$$Y + C - R_*C + Q(C, Y),$$

where Q is quadratic in C and Y.

The proof of Proposition 43 is straightforward. It particularly relies on the following proposition, which simplifies working with diffeomorphisms of the form ψ_X.

Proposition 44. *Let M be a compact Riemannian manifold. If $X, Y \in \mathrm{Vect}^\infty(M)$ are sufficiently C^1 small and we define Z by*

$$\psi_Y \circ \psi_X = \psi_{X+Y+Z},$$

then there exists a fixed ℓ such that for each k there exists C_k such that

$$\|Z\|_{C^k} \le C_k(\|X\|^2_{C^{k+\ell}} + \|Y\|^2_{C^{k+\ell}}),$$

i.e. Z is quadratic in X and Y.

The proof of Proposition 44 uses the following two lemmas concerning maps of $\mathbb{R}^n$.

Lemma 45 ([19, Thm. A.7]). *Suppose that B is a compact convex domain in $\mathbb{R}^n$ with interior points. Then for $k \ge 0$, there exists C such if f, g are C^k maps from B to $\mathbb{R}$, then*

$$\|fg\|_{C^k} \le C_k(\|f\|_{C^k}\|g\|_{C^0} + \|f\|_{C^0}\|g\|_{C^k}).$$

Lemma 46 ([19, Thm. A.8]). *For $i \in \{1, 2, 3\}$, let B_i be a fixed compact convex domain in $\mathbb{R}^{n_i}$ with interior points. Let $k \ge 1$. There exists $C_k > 0$ such that if $f : B_1 \to B_2$ and $g : B_2 \to B_3$ are both C^k, then $f \circ g$ is C^k and*

$$\|f \circ g\|_{C^k} \le C_k(\|f\|_{C^k}\|g\|^k_{C^1} + \|f\|_{C^1}\|g\|_{C^k} + \|f \circ g\|_{C^0}).$$

Using the previous two lemmas, we prove the following.

Proposition 47. *Suppose that g is a metric on $\mathbb{R}^n$. For a smooth vector field Y such that $\|Y\|_{C^1} < 1$, and ψ_Y^g is defined, let*

$$Z(x) = \psi_Y^g(x) - Y(x) - x.$$

Let B be a compact convex domain in $\mathbb{R}^n$ with interior points. Then $Z|_B$ is quadratic in Y. In fact, for each k there exists C_k such that

$$\|Z|_B\|_{C^k} \leq C_k \|Y\|_{C^k}^2.$$

Proof. Set $\gamma(Y(x), t) = \exp_x tY(x) - x$, so that $\gamma(Y(x), 1) + x = \psi_Y^g$ and $\gamma(Y(x), 0) = 0$. We rewrite Z:

$$\begin{aligned} Z &= \psi_Y^g(x) - x - Y(x) = \gamma(Y(x), 1) - Y(x) \\ &= \int_0^1 [\dot\gamma(Y(x), t) - Y(x)]\, dt = \int_0^1 [\dot\gamma(Y(x), t) - \dot\gamma(Y(x), 0)]\, dt \\ &= \int_0^1 \int_0^1 t\ddot\gamma(Y(x), st)\, ds\, dt = \int_0^1 t \int_0^1 \ddot\gamma(Y(x), st)\, ds\, dt. \end{aligned}$$

By differentiating under the integral, we see that the nth derivatives of Z are controlled by the maximum of the nth derivatives of $\ddot\gamma(Y(x), t)$ for each fixed t. Hence it suffices to show for each $t \in [0, 1]$ that $\ddot\gamma(Y(x), t)$ is second order in Y.

Dropping the explicit dependence on x, we recall the coordinate expression of the geodesic equation. For a coordinate frame $[e_1, \dots, e_n]$ and indices $1 \leq \mu, \nu, \lambda \leq n$, we define the Christoffel symbols $\Gamma_{\mu\nu}^\lambda$ by $\langle \nabla_{e_\mu} e_\nu, e_\lambda\rangle$. In addition, we write $\dot\gamma^\nu$ for $\langle \dot\gamma, e_\nu\rangle$ and similarly for $\ddot\gamma$. The coordinate expression for the geodesic equation is then

$$\ddot\gamma^\lambda = -\Gamma_{\mu\nu}^\lambda \dot\gamma^\mu \dot\gamma^\nu.$$

We estimate the C^k norm of the right hand side. Write ϕ^t for the geodesic flow, and let TB denote the tangent vectors to $\mathbb{R}^n$ with basepoint in B. Note that as B is compact, for any tangent vector $v \in TB$, $\phi^t(v)$ is defined for some positive amount of time. For fixed $r > 0$ in TB, let $TB(r)$ be the set of vectors $v \in TB$ such that $\|v\| < r$. Note that $\|\phi^t|_{TB(r)}\|_{C^k}$ is bounded. Let π be the projection from a tangent vector in $T\mathbb{R}^n$ to its basepoint in $\mathbb{R}^n$. Then

$$\gamma(x, t) = \pi \circ \phi^t \circ Y(x).$$

Hence, writing $\dot\phi$ for the geodesic spray,

$$\dot\gamma(x, t) = D\pi \circ \dot\phi|_{\phi^t(Y(x))}. \tag{101}$$

$D\pi \circ \dot\phi^t|_{TB(r)}$ has its C^k norm uniformly bounded in t by some D_k. By Lemma 46 because $\|Y\|_{C^1} < 1$ it follows that $\|\phi^t(Y(x), t)\|_{C^k} \leq C_k \|Y\|_{C^k}$.

Hence by applying Lemma 46 to (101), and similarly using the fact that $\|Y\|_{C^1} < 1$ and $D\pi \circ \dot\phi$ is uniformly bounded, we find

$$\|(D\pi \circ \dot\phi^t) \circ Y\|_{C^k} \leq C_k'(D_k\|Y\|_{C^1} + D_1\|Y\|_{C^k} + \|Y\|_{C^0}).$$

Hence

$$\|\dot{\gamma}(x,t)\|_{C^k} = \|D\pi \circ \dot{\phi}|_{\phi^t(Y(x))}\|_{C^k} \le C_k \|Y\|_{C^k}.$$

The geodesic equation shows that at each point the coordinates of $\ddot{\gamma}$ are a quadratic polynomial in the coordinates of $\dot{\gamma}$. Hence by Lemma 45,

$$\|\ddot{\gamma}(x,t)\|_{C^k} \le C_k'' \|Y\|_{C^k}^2$$

for all $t \in [0,1]$. Thus we obtain a uniform estimate on Z. ∎

Proof of Proposition 44. As before, it suffices to prove the estimate in a chart. So, we are reduced to working in a neighborhood of $0 \in \mathbb{R}^n$. Fix some k; then by Proposition 47 we may write

$$\psi_Y(x) = x + Y(x) + Z_Y(x),$$

where $Z_Y(x)$ is quadratic in Y. Similarly define $Z_X(x)$ and $Z_{X+Y}(x)$. Then

$$\begin{aligned}\psi_Y \circ \psi_X &= \psi_Y(x + X(x) + Z_X(x)) \\ &= x + X(x) + Z_X(x) + Y(x + X(x) + Z_X(x)) + Z_Y(x + X(x) + Z_X(x)).\end{aligned}$$

To prove this proposition, we compare the previous line with

$$\psi_{X+Y} = x + X(x) + Y(x) + Z_{X+Y}(x).$$

The difference is

$$\begin{aligned}\psi_Y \circ \psi_X - \psi_{X+Y} = &\, Z_X(x) - Z_{X+Y}(x) + Y(x + X(x) + Z_X(x)) - Y(x) \\ &+ Z_Y(x + X(x) + Z_X(x)).\end{aligned}$$

The first and second terms satisfy the appropriate quadratic C^k estimate already. For the last term, we apply Lemma 46. Hence by assuming that $\|X\|_{C^1}$ is sufficiently small, we conclude that the Z_Y term is quadratic. We now turn to the Y terms:

$$Y(x + X(x) + Z_X(x)) - Y(x).$$

For this we apply the same trick as before. Write

$$Y(x + X(x) + Z_X(x)) - Y(x) = \int_0^1 Y'(x + t(X(x) + Z_X(x)))\|X(x) + Z_X(x)\|\, dt.$$

By differentiating under the integral, it suffices to show that the integrand is quadratic in X and Y. By Lemma 45, the integrand will be quadratic if there exists ℓ such that for each k there is a constant C_k such that both of $\|Y'(x + t(X(x) + Z_X(x)))\|_{C^k}$ and $\|X(x) + Z_X(x)\|_{C^k}$ are bounded by $C_k(\|X\|_{C^{k+\ell}} + \|Y\|_{C^{k+\ell}})$. This follows for both terms by the application of Lemma 46, so we are done. ∎

We now show another basic fact: near to the identity map a diffeomorphism and its inverse have comparable size.

Lemma 48. *Suppose that M is a closed Riemannian manifold. Then there exists $\epsilon > 0$ such that for all $k \geq 0$ there exists C_k such that if $f \in \mathrm{Diff}^k(M)$ and $d_{C^1}(f, \mathrm{Id}) < \epsilon$ then*

$$d_{C^k}(f^{-1}, \mathrm{Id}) \leq C_k d_{C^k}(f, \mathrm{Id}).$$

Proof. This proof follows the outline of the similar estimate in [16, Lem. 2.3.6]. For convenience, write $g = f^{-1}$. In a chart, we write $f(x) = x + X(x)$ where the C^k norm of X is bounded by $d_{C^k}(f, \mathrm{Id})$. Similarly write $g(x) = x + Y(x)$. We now apply the chain rule to differentiate $g \circ f$. The case where $n = 1$ is immediate by differentiating $g \circ f = x + X(x) + Y(x + X(x))$, which gives

$$DX + DY(\mathrm{Id} + DX) = 0.$$

Hence

$$DY = -DX(\mathrm{Id} + DX)^{-1},$$

which is uniformly comparable to $\|DX\|$ because $d_{C^1}(f, \mathrm{Id})$ is uniformly bounded.

For $k > 1$, we must estimate the higher order derivatives of Y. Note that for $k > 1$ we have $D^k g = D^k Y$ and $D^k f = D^k X$.

Applying the chain rule to $f \circ g = \mathrm{Id}$ to calculate the kth derivative gives

$$0 = \sum_{l=1}^{k} \sum_{j_1+\cdots+j_l=k} C_{l,j_1,\ldots,j_l} D^l_{g(x)} f\{D^{j_1}_x g, \ldots, D^{j_l}_x g\},$$

and hence

$$D^k_x g = -(D_{g(x)} f)^{-1} \sum_{l=2}^{k} \sum_{j_1+\cdots+j_l=k} C_{l,j_1,\ldots,j_l} D^l_{g(x)} f\{D^{j_1}_x g(x), \ldots, D^{j_l}_x g(x)\}. \quad (102)$$

As $(Df)^{-1}$ has uniformly bounded norm, it suffices to show that each term in the sum has norm bounded by $\|X\|_{C^n}$.

We use the interpolation estimate in Lemma 52. If $j > 1$, then

$$\|D^j g\| = \|D^j Y\|.$$

By interpolation between the C^1 and C^{n-1} norms, for $1 \leq j \leq n-1$,

$$\|Y\|_{C^j} \leq C_{1,n-1} \|Y\|_{C^1}^{\frac{n-j-1}{n-2}} \|Y\|_{C^{n-1}}^{\frac{j-1}{n-2}}.$$

By interpolation between the C^1 and C^n norms, for $1 \leq j \leq n$,

$$\|X\|_{C^j} \leq C_{1,n} \|X\|_{C^1}^{\frac{n-j}{n-1}} \|X\|_{C^n}^{\frac{j-1}{n-1}}.$$

We now estimate the terms on the right hand side of (102). If $j_i = 1$ for some i, then $D^{j_i} g = \mathrm{Id} + DY$. Hence the right hand side of (102) may be rewritten as the sum of terms of the form

$$D^l_{g(x)} X\{A_1, \ldots, A_l\},$$

where each A_i is either Id or $D^{j_i}Y$ and the sum of the j_i is less than or equal to k. If $\|Y\|_{C^{k-1}} \leq 1$, then we are immediately done as the norm of this expression is at most $\|D^k f\|$. Otherwise, we may suppose that $\|Y\|_{C^{k-1}} \geq 1$. The C^1 norms of X and Y are uniformly bounded. Hence by interpolating between the C^1 and C^k norms to estimate the $D^l X$ term and between the C^1 and C^{k-1} norms to estimate the A_i terms, we find that

$$\|D^k_{g(x)}X\{A_1, \dots, A_k\}\| \leq C'\|X\|_{C^k}^{\frac{l-1}{k-1}}\|Y\|_{C^{k-1}}^{\frac{k-r}{n-2}},$$

where $r \geq l$. But as $\|Y\|_{C^{k-1}} > 1$, this is bounded above by

$$C'\|X\|_{C^k}^{\frac{l-1}{k-1}}\|Y\|_{C^{k-1}}^{\frac{k-l}{k-2}}.$$

Thus

$$\|D^k Y\|_{C^0} \leq C'' \sum_{l=2}^{k} \|X\|_{C^k}^{\frac{l-1}{k-1}}\|Y\|_{C^{k-1}}^{\frac{k-l}{k-2}}.$$

We may now proceed by induction on k. We already established the theorem for $k = 1$. Now, given that $\|Y\|_{C^{k-1}} \leq C_{k-1}\|X\|_{C^{k-1}}$, it follows that

$$\|D^k Y\|_{C^0} \leq C''' \sum_{l=2}^{k} \|X\|_{C^k}^{\frac{l-1}{k-1}}\|X\|_{C^{k-1}}^{\frac{k-l}{k-2}}.$$

By interpolation between the C^1 and C^k norms, and the uniform bound on the C^1 norm, we find that $\|X\|_{C^{k-1}} \leq D_k\|X\|_{C^k}^{\frac{k-2}{k-1}}$. This yields

$$\|D^k Y\|_{C^0} \leq D' \sum_{l=2}^{k} \|X\|_{C^k}^{\frac{l-1}{k-1}}\|X\|_{C^k}^{\frac{k-l}{k-1}} \leq D''\|X\|_{C^k},$$

which is the desired result. ∎

We now obtain the following corollary.

Corollary 49. *Suppose that M is a closed Riemannian manifold. For smooth C^1 small vector fields X on M, we may write*

$$\psi_X^{-1} = \psi_{-X+Z},$$

where Z is quadratic in X.

Proof. To begin we know by Proposition 44 that

$$\psi_X \circ \psi_{-X} = \psi_Z,$$

where Z is quadratic in X. Note that $\psi_{-X} \circ \psi_Z^{-1} = \psi_X^{-1}$. By Lemma 48, $\psi_Z^{-1} = \psi_{Z'}$ where Z' is quadratic in X. Hence $\psi_X^{-1} = \psi_{-X} \circ \psi_{Z'}$. By Proposition 44, this gives $\psi_X^{-1} = \psi_{-X+Z'+Q}$, where Q is quadratic in X and Z'. Hence as Z' is quadratic in X, the corollary follows. ∎

We can now complete the proof of the estimate on the error field of the conjugate system.

Proof of Proposition 43. To show this, we repeatedly apply Proposition 44 and Corollary 49. Writing Z for anything second order in C and Y, we find

$$\begin{aligned}\psi_C f \psi_C^{-1} &= \psi_C \psi_Y R \psi_C^{-1} = \psi_{C+Y+Z} R \psi_C^{-1} = \psi_{C+Y+Z} R \psi_{-C+Z} \\ &= \psi_{C+Y+Z+R_*(-C+Z)} R = \psi_{C+Y-R_*C+Z} R.\end{aligned}$$ ■

We now show two additional lemmas that we use in the KAM scheme.

Lemma 50. *Let M be a closed Riemannian manifold. Fix $k \geq 1$. There exist $C_k, \epsilon > 0$ such that if $R \in \mathrm{Isom}(M)$ and $f, g \in \mathrm{Diff}^k(M)$ satisfy $d_{C^1}(f, R) < \epsilon$, and $d_{C^1}(g, \mathrm{Id}) < \epsilon$, then*

$$\begin{aligned} d_{C^k}(f \circ g, R) &\leq C_k(d_{C^k}(f, R) + d_{C^k}(g, \mathrm{Id})), \\ d_{C^k}(g \circ f, R) &\leq C_k(d_{C^k}(f, R) + d_{C^k}(g, \mathrm{Id})). \end{aligned}$$

Proof. To prove the first inequality, in coordinates write $f(x) = R(x) + Y(x)$ and $g(x) = x + X(x)$. Then we just need to estimate

$$f \circ g(x) - R(x) = R(x + X(x)) - R(x) + Y(x + X(x)).$$

The last term is controlled by $d_{C^k}(f, R) + d_{C^k}(g, \mathrm{Id})$ by Lemma 46. So, it suffices to estimate the first term. The kth derivative of $R(x + X(x)) - R(x)$ is then

$$\sum_{l=1}^{k} \sum_{j_1+\cdots+j_l=k} [C_{l,j_1,\ldots,j_l} D^l_{x+X(x)} R\{D_x^{j_1} g, \ldots, D_x^{j_l} g\} - D_x^l R].$$

For all the terms with $l < k$, the same interpolation approach as in Lemma 48 gives the appropriate estimate, i.e. they are bounded by

$$C \sum_{l=1}^{k-1} \|X\|_{C^k}^{\frac{l-1}{k-1}} \|X\|_{C^{k-1}}^{\frac{k-l}{k-2}}.$$

There are two remaining terms which are unaccounted for: $D^k R_{x+X(x)} - D^k R_x$. This is bounded by a constant time $\|X\|_{C^0}$ and the result follows.

We now consider the second inequality. As before, we must estimate

$$g \circ f(x) - R(x) = X(x) + Y(R(x) + X(x)).$$

The important term is the second one. A similar argument to the one before then gives the result as all derivatives of R are uniformly bounded independent of R. ■

Lemma 51. *Let M be a closed Riemannian manifold and $k \geq 0$. If $g_n \in \mathrm{Diff}^k(M)$ is a sequence of diffeomorphisms and $\sum_n d_{C^k}(g_n, \mathrm{Id}) < \infty$, then the sequence of compositions of diffeomorphisms $h_n = g_n g_{n-1} \cdots g_2 g_1$ converges in C^k to a diffeomorphism.*

Proof. As before, we check in charts. Having fixed a chart, write $g_n(x) = x + X_n(x)$. Write $h_n(x) = 1 + Y_n(x)$. Let $a_n = \|X_n\|_{C^k}$ and let $b_n = \|Y_n\|_{C^k}$. Note that

$$h_n(x) = x + Y_{n-1}(x) + X_n(x + Y_n(x)). \tag{103}$$

Suppose for the moment that $\|Y_{n-1}\|_{C^k} \leq 1$. Using Lemma 46 and $\|Y_n\|_{C^k} \leq 1$, we have

$$\begin{aligned} \|X_n(x + Y_{n-1})\|_{C^k} &\leq C_k(\|X_n\|_{C^k}\|x + Y_{n-1}\|^k_{C^1} + \|X_n\|_{C^1}\|x + Y_{n-1}\|_{C^k} + \|X\|_{C^0}) \\ &\leq C'_k(a_n + a_n b_{n-1}) \end{aligned} \tag{104}$$

Hence it follows from (103) that there exists D_k such that if $b_{n-1} \leq 1$ then

$$b_n \leq b_{n-1} + D_k a_n(1 + b_{n-1}).$$

By induction, under the same assumption that $\|Y_j\|_{C^k} \leq 1$ for $j < n$, it follows that

$$b_n \leq -1 + \prod_{i=1}^{n}(1 + D_k a_i).$$

By noting that $\prod_{i=1}^{\infty}(1 + x_n) \leq \exp(\sum_{i=1}^{\infty} x_n)$ for $x_n \geq 0$, we can conclude that a tail of the sequence converges. Indeed, as $\sum_n a_n$ converges we can inductively check that these inequalities hold starting the argument from an index N satisfying $\exp(\sum_{i=N}^{\infty} D_k a_i) - 1 < 1$. Hence as a tail of the infinite composition converges, so does the whole composition. ∎

Appendix B. Interpolation inequalities

There is a basic C^k interpolation inequality, which may be found in the appendix of [19, Thm A.5]:

Lemma 52. *Suppose that M is a closed Riemannian manifold. For $0 \leq a \leq b < \infty$ and $0 < \lambda < 1$ there exists a constant $C(a, b, \lambda)$ such that for any real-valued C^b function f defined on M,*

$$\|f\|_{C^{\lambda a+(1-\lambda)b}} \leq C\|f\|^{\lambda}_{C^a}\|f\|^{1-\lambda}_{C^b}.$$

The following is an immediate consequence of Lemma 52.

Lemma 53. *Suppose that M is a closed Riemannian manifold. There exists $\epsilon > 0$ such that for $0 \leq a \leq b < \infty$ and $0 < \lambda < 1$ there exists a constant $C(a, b, \lambda)$ such that for any $f \in \mathrm{Diff}^{\infty}(M)$ such that $d_{C^0}(f, \mathrm{Id}) < \epsilon$,*

$$d_{C^{\lambda a+(1-\lambda)b}}(f, \mathrm{Id}) \leq C d_{C^a}(f, \mathrm{Id})^{\lambda} d_{C^b}(f, \mathrm{Id})^{1-\lambda}.$$

Lemma 54. *Consider the space $C^{\infty}(M, N)$ where M and N are Riemannian manifolds and M and N are closed. For all $j, \sigma > 0$, there exists a natural number k and a number $\epsilon_0 > 0$ such that if $f, g \in C^{\infty}(M, N)$, $\|f - g\|_{H^j} < \epsilon_0 < 1$, and $\|f - g\|_{C^k} \leq 1/2$ then $\|f - g\|_{C^j} \leq \|f - g\|^{1-\sigma}_{H^j}$.*

Proof. The proof is a relatively straightforward application of the Sobolev embedding theorem and interpolation inequalities. First, we recall an interpolation inequality for Sobolev norms [2, Thm. 6.5.4]: For each $0 < \theta < 1$, s_0, s_1, there exists a constant C such that if we let $s = (1-\theta)s_0 + \theta s_1$, then

$$\|f-g\|_{H^s} \leq C\|f-g\|_{H^{s_0}}^{1-\theta}\|f-g\|_{H^{s_1}}^{\theta}.$$

To begin the proof, note that it suffices to estimate $\|f-g\|_{C^{j+1}}$. Fix ℓ large enough that H^ℓ embeds compactly in C^{j+1} by a Sobolev embedding. Then pick k large enough that

$$\|f-g\|_{H^\ell} \leq C_{\lambda,\ell}\|f-g\|_{H^j}^{1-\theta}\|f-g\|_{H^k}^{\theta}$$

for some $0<\theta<\sigma$. The term $\|f-g\|_{H^k}^{\theta}$ is uniformly bounded by $C_k\|f-g\|_{C^k}^{\theta}$. Hence as H^ℓ compactly embeds in C^{j+1}, there exists $C'>0$ such that

$$\|f-g\|_{C^{j+1}} \leq C'\|f-g\|_{H^j}^{1-\theta} = C'\|f-g\|_{H^j}^{\sigma-\theta}\|f-g\|_{H^j}^{1-\sigma}.$$

If we choose ϵ_0 sufficiently small that $C'\|f-g\|_{H^j}^{\sigma-\theta} \leq 1$, then the result follows. ∎

A similar argument shows the following:

Lemma 55. *Suppose that E is a smooth Riemannian vector bundle over a closed Riemannian manifold M. For all choices $j, \ell, \sigma, D > 0$ there exist k, ϵ_0 such that if f is a smooth section of E and $\|f\|_{H^j} \leq \epsilon_0 < 1$ and $\|f\|_{C^k} \leq D$ then $\|f\|_{C^\ell} \leq \|f\|_{H^j}^{1-\sigma}$.*

Appendix C. Estimate on lifted error fields

The goal of this section is to prove a technical estimate on the error fields of a lifted system. The proof is a computation in charts.

Lemma 56. *Suppose that M is a closed Riemannian manifold. Fix numbers $m, k \geq 0$ and d such that $0 \leq d \leq \dim M$. There exists a constant C such that the following holds. For any tuple $(f_1, \dots, f_m)$ of diffeomorphisms of M and $(r_1, \dots, r_m)$ a C^1 close tuple of isometries of M, let Y_i be the shortest vector field such that $\exp_{r_i(x)} Y_i(x) = f_i(x)$. Let F_i be the lift of f_i to $\mathrm{Gr}_d(M)$ and R_i be the lift of r_i to $\mathrm{Gr}_d(M)$. Let $\widetilde{Y}_i$ be the shortest vector field on $\mathrm{Gr}_d(M)$ such that $\exp_{R_i(x)} \widetilde{Y}_i(x) = F_i(x)$. If $\|\sum_i Y_i\|_{C^k} = \epsilon$ and $\max_i \|Y_i\|_{C^k} = \eta$, then*

$$\left\|\sum_{i=1}^{m} \widetilde{Y}_i\right\|_{C^{k-1}} \leq C(\epsilon + \eta^2).$$

Proof. The proof is straightforward but tedious. We give the proof when each R_i is the identity. Removing this assumption both complicates the argument in purely technical ways and substantially obscures why the lemma is true. At the end of the argument, we indicate the modifications needed for the general proof.

For readability we redevelop some of the basic notions concerning Grassmannians. First we recall the charts on $\mathrm{Gr}_d(V)$, the Grassmannian of d-planes in a vector space V. Recall that given a vector space V and a pair of complementary subspaces P and Q of V, if $\dim P = d$ we obtain a chart on $\mathrm{Gr}_d(V)$ in the following manner. Let $L(P, Q)$ denote the space of linear maps from P to Q. For $A \in L(P, Q)$, we send A to the subspace $\{x + Ax : x \in P\} \in \mathrm{Gr}_d(V)$. This gives a smooth parametrization of a subset of $\mathrm{Gr}_d(V)$. Having fixed a complementary pair of subspaces P and Q, let π_P denote the projection to P along Q.

Suppose that U is a chart on M and let $\partial_1, \dots, \partial_n$ denote the coordinate vector fields. We use the usual coordinate framing of TU to give coordinates on the Grassmannian bundle $\mathrm{Gr}_d(M)$. The tangent bundle to U naturally splits into subbundles spanned by $\{\partial_1, \dots, \partial_d\}$ and $\{\partial_{d+1}, \dots, \partial_n\}$. Call these subbundles P and Q, respectively. Let $\mathrm{End}(P, Q)$ denote the bundle of maps from P to Q. We obtain a coordinate chart via associating to an element of $A \in \mathrm{End}(P, Q)$ and a point $x \in U$ the graph of A in the tangent space over x.

As we have assumed that each r_i is the identity, in charts we write $f_i(x) = x + X_i(x)$. As the f_i are C^1 small, we work in a single chart. It now suffices to prove the corresponding estimate on the field X_i because X_i and Y_i are equal up to an error that is quadratic in the sense of Definition 42. We now calculate the action of f on $\mathrm{Gr}_d(U)$. Suppose that $A \in \mathrm{End}(P, Q)$. Then $\{Df(v + Av)\}$ is a subspace of $T_{f(x)}M$. We must find the map A' whose graph gives the same subspace. Let I_A be the $n \times d$ matrix with top block I and bottom block A. Then Df sends A to

$$A' = Df I_A(\pi_P Df I_A)^{-1} - \mathrm{Id}\,.$$

To see this, we must check that $A'V \subseteq Q$ and $\{Dfv + DfAv : v \in V\} = \{v + A'v\}$. The second condition is evident from the definition of A'. If $v \in P$, then $(\pi_P Df I_A)^{-1} v = w$ is an element of P satisfying $\pi_P Df I_A w = v$. Thus $A'v = Df I_A(\pi_P Df I_A)^{-1} v - v \in Q$ and hence $A'V \subseteq Q$. Write F for the induced map on $\mathrm{Gr}_d(U)$. In coordinates, F is the map

$$(x, A) \mapsto (x, Df I_A(\pi_P Df I_A)^{-1} - \mathrm{Id}). \tag{105}$$

Write I_d for the $d \times d$ identity matrix. Let $\widehat{DX}_i$ be the matrix consisting of the first d rows of the matrix DX_i. In the estimates below, we will assume that the size of A is uniformly bounded. This does not restrict the generality as any subspace may be represented by such a uniformly bounded A. Then note that

$$\begin{aligned}\left(\pi_P D_f\left[\frac{I_d}{A}\right]\right)^{-1} &= \left(I_d + \widehat{DX}\left[\frac{I_d}{A}\right]\right)^{-1}\\ &= I_d - \widehat{DX}\left[\frac{I_d}{A}\right] + O(DX^2),\end{aligned}$$

where the $O(DX^2)$ is quadratic in the sense of Definition 42. Write X_A for the second term above.

We then have

$$\begin{aligned}
DfI_A(\pi_P DfI_A)^{-1} - \mathrm{Id} &= (\mathrm{Id} + DX)\left[\frac{I_d}{A}\right](I_d - X_A) - \left[\frac{I_d}{0}\right] + O(DX^2) \\
&= \left[\frac{I_d}{A}\right] - \left[\frac{I_d}{A}\right]X_A + DX\left[\frac{I_d}{A}\right] + DX\left[\frac{I_d}{A}\right]X_A - \left[\frac{I_d}{0}\right] + O(DX^2) \\
&= \left[\frac{0}{A}\right] - \left[\frac{I_d}{A}\right]X_A + DX\left[\frac{I_d}{A}\right] + O(DX^2) \\
&= \left[\frac{0}{A}\right] + H(A, DX) + O(DX^2),
\end{aligned}$$

where $H(A, DX)$ is the sum of the second and third terms two lines above. Note that H is linear in DX, and $\|H(A, DX)\| \le C\|DX\|$ given our uniform boundedness assumption on A.

Thus we see that in this chart on $\mathrm{Gr}_d(U)$,

$$F(x, A) - (x, A) = (f(x) - x, H(A, DX) + O(DX^2)). \tag{106}$$

In this chart, $\|\sum_i f_i(x) - x\|_{C^k} \le \epsilon$. Hence writing $f_i(x) = x + X_i(x)$ as before, we have $\|\sum_i DX_i(x)\|_{C^{k-1}} \le \epsilon$. Thus

$$\begin{aligned}
\Big\|\sum_i F_i(x, A) - (x, A)\Big\|_{C^{k-1}} &= \Big\|\sum_i (f_i(x) - x, H(A, DX_i) + O(DX^2))\Big\|_{C^{k-1}} \\
&\le C\Big(\Big\|\sum_i X_i\Big\|_{C^k} + \max_i \|X_i\|^2_{C^k}\Big)
\end{aligned}$$

by the linearity of H. This completes the proof in the special case where $r_i = \mathrm{Id}$ for each i.

In the general setting one follows the same sequence of steps. One writes $f_i(x) = r_i(x) + X_i(r_i(x))$. One then does the same computation to determine the action on the Grassmannian bundle. This is complicated by additional terms related to R. Having finished this computation, one finds a natural analog of $H(A, DX)$, which now comprises eight terms instead of two, and also depends on r_i. Recognizing the cancellation is then somewhat complicated because of the dependence on r_i. However, this dependence does not cause an issue because the terms that would potentially cause trouble satisfy some useful relations. These relations emerge when one keeps in mind the basepoints, which is crucial when the isometries are non-trivial. ■

Appendix D. Determinants

Suppose that V and W are finite-dimensional inner product spaces. Consider a linear map $L: V \to W$. The determinant of the map L is defined as follows. If $\{v_i\}$ is an orthonormal basis for V, one may measure the size of the tensor $Lv_1 \wedge \cdots \wedge Lv_n$ with respect to the norm on tensors induced by the metric on W. If $\{v_1, \dots, v_n\}$ is a basis for V, then we define

$$\det(L, g_1, g_2) := \sqrt{\frac{\mathrm{Det}(\langle Lv_i, Lv_j\rangle_{g_2})}{\mathrm{Det}(\langle v_i, v_j\rangle_{g_1})}},$$

where Det is the usual determinant of a square matrix. Sometimes we have a map $L\colon V \to W$ and a subspace $E \subset V$. We then define

$$\det(L, g_1, g_2 \mid E) = \det(L|_E, g_1|_E, g_2). \tag{107}$$

When the spaces V and W are understood, we may write $\det(L \mid E)$.

There are some properties of det that we will record for later use.

Lemma 57. *Fix a basis and suppose that $V = W$. Working with respect to this basis, the determinant has the following properties:*

$$\det(L, g_1, g_2) = \det(\mathrm{Id}, g_1, L^* g_2), \tag{108}$$

$$\det(\mathrm{Id}, \mathrm{Id}, A) = \sqrt{\det(A, \mathrm{Id}, \mathrm{Id})} = \sqrt{|\mathrm{Det}(A)|}. \tag{109}$$

Proof. For the first equality, let $\{v_i\}$ be a basis of (V, g_1). Then

$$\det(L, g_1, g_2) = \sqrt{\frac{\mathrm{Det}(\langle Lv_i, Lv_j\rangle_{g_2})}{\mathrm{Det}(\langle v_i, v_j\rangle_{g_1})}}.$$

But $\langle v_i, v_j\rangle_{L^* g_2} = \langle Lv_i, Lv_j\rangle_{g_2}$, so this is equal to

$$\sqrt{\frac{\mathrm{Det}(\langle v_i, v_j\rangle_{L^* g_2})}{\mathrm{Det}(\langle v_i, v_j\rangle_{g_1})}},$$

which is the definition of $\det(\mathrm{Id}, g_1, L^* g_2)$.

For the second equality, fix an orthonormal basis $\{e_i\}$. Then

$$\det(\mathrm{Id}, \mathrm{Id}, A) = \sqrt{\mathrm{Det}(\langle e_i, e_j\rangle_A)} = \sqrt{\mathrm{Det}(A_{ij})},$$

whereas

$$\det(A, \mathrm{Id}, \mathrm{Id}) = \sqrt{\mathrm{Det}(\langle Ae_i, Ae_j\rangle_{\mathrm{Id}})} = \sqrt{\mathrm{Det}(A^T A)} = \sqrt{|\mathrm{Det}(A)|^2} = |\mathrm{Det}(A)|. \quad \blacksquare$$

We record the following estimate which is used in the proof.

Lemma 58. *Let M be a closed manifold and let $0 \le r \le \dim M$. If g is an isometry of M, then $\ln\det(Df \mid E_x)$, which is defined on $\mathrm{Gr}_r(M)$, satisfies the following estimate:*

$$\|\ln\det(Df \mid E_x)\|_{C^k} = O(d_{C^{k+1}}(f, g)) \quad \text{as } f \to g \text{ in } C^{k+1}.$$

The big-O is uniform over all isometries g.

Proof. It suffices to show that this estimate holds in charts. So, fix a pair of charts U and V on M such that $f(U)$ has compact closure inside of V. We define a map $H : \mathrm{Gr}_d(U) \times U \times V \times \mathbb{R}^{n^2} \to \mathbb{R}$ by sending the point (E, x, y, A) to $\ln \det(A, g_x, g_y | E)$, where g_x and g_y denote the pullback metric from M. Using f we define a map $\widetilde{f} : \mathrm{Gr}_d(U) \times U \to \mathrm{Gr}_d(U) \times U \times V \times \mathbb{R}^{n^2}$ by

$$(E, x) \mapsto (E, x, f(x), Df),$$

where we are using the coordinates to express Df as a matrix. Then the quantity we wish to estimate the C^k norm of is $H \circ \widetilde{f}$. If we analogously define $\widetilde{g}$, then note that $H \circ \widetilde{g} \equiv 0$ because g is an isometry. By writing out the derivatives using the chain rule, and using the fact that f is uniformly close to g, one sees that $\|H \circ \widetilde{g} - H \circ \widetilde{f}\|_{C^k} = O(d_{C^{k+1}}(f, g))$, and the result follows. ∎

Appendix E. Taylor expansions

E.1. Taylor expansion of the log Jacobian

Proposition 59. *For C^1 small vector fields Y on a Riemannian manifold M, the following approximation holds:*

$$\int_{\mathrm{Gr}_r(M)} \ln \det(D_x \psi_Y, \mathrm{Id}, g_{\psi_{Y(x)}} \mid E_x)\, d\mathrm{vol}$$
$$= -\frac{r}{2d} \int_M \|E_C\|^2\, d\mathrm{vol} + \frac{r(d-r)}{(d+2)(d-1)} \int_M \|E_{NC}\|^2\, d\mathrm{vol} + O(\|Y\|^3_{C^1}),$$

where E_C and E_{NC} are the conformal and non-conformal strain tensors associated to ψ_Y as defined in Section 4.2. In addition, det *is defined in Appendix D and ψ_Y is defined in* (11)*.*

Proof. The proof is a lengthy computation with several subordinate lemmas.

In order to estimate the integral over M, we will first obtain a pointwise estimate on

$$\int_{\mathrm{Gr}_r(T_x M)} \ln \det(D_x \psi_Y \mid E)\, dE.$$

To estimate this we work in an exponential chart on M centered at x. In this chart, x is 0 and $\psi_Y(0) = Y(0)$. Then

$$\int_{\mathrm{Gr}_r(T_x M)} \ln \det(D_x \psi_Y \mid E)\, dE = \int_{\mathrm{Gr}_r(T_x M)} \ln \det(D_0 \psi_Y, \mathrm{Id}, g_{Y(0)} \mid E)\, dE.$$

We now rewrite the above line so that we can apply the Taylor approximation in Propositions 62 and 63 below.

Write the metric as $\mathrm{Id} + \hat{g}$. As we are in an exponential chart, $\|\hat{g}_{Y(0)}\| = O(\|Y\|^2_{C^0})$. Write $D\psi_Y = \mathrm{Id} + \hat{\psi}$. The integral we are calculating only involves $\hat{\psi}_0$ and $\hat{g}_{Y(0)}$, so below we drop the subscripts. Then

$$\int_{\mathrm{Gr}_r(T_xM)} \ln\det(D_x\psi_Y \mid E)\, dE = \int_{\mathrm{Gr}_r(T_xM)} \ln\det(\mathrm{Id} + \hat{\psi}, \mathrm{Id}, \mathrm{Id} + \hat{g} \mid E)\, dE.$$

Now applying the Taylor expansions in Propositions 62 and 63, we obtain the following expansion. For convenience let

$$K = \frac{\hat{\psi} + \hat{\psi}^T}{2} - \frac{\mathrm{Tr}(\hat{\psi})}{d}\,\mathrm{Id}. \tag{110}$$

Then

$$\begin{aligned}\int_{\mathrm{Gr}_r(T_xM)} &\ln\det(D\psi_Y, \mathrm{Id}, g_{Y(0)} \mid E)\, dE \\ &= \frac{r}{d}\mathrm{Tr}(\hat{\psi}) + \left[-\frac{r}{2d}\mathrm{Tr}(\hat{\psi}^2) + \frac{r(d-r)}{(d+2)(d-1)}\mathrm{Tr}(K^2)\right] + O(\|\hat{\psi}^3\|) \\ &\quad + \frac{r}{2d}\mathrm{Tr}(\hat{g}) + O(\|\hat{g}\|^2).\end{aligned} \tag{111}$$

Note that $\|\hat{\psi}\| = O(\|Y\|_{C^1})$ and $\|\hat{g}\| = O(\|Y\|^2_{C^0})$, hence the big-$O$ terms in the above expression are each $O(\|Y\|^3_{C^1})$.

We now eliminate the two trace terms that are not quadratic in their arguments. For this, we use a Taylor expansion of the determinant.[3] Thus

$$\det(D\psi, \mathrm{Id}, g_{Y(0)}) = 1 + \mathrm{Tr}(\hat{\psi}) + \frac{(\mathrm{Tr}(\hat{\psi}))^2 - \mathrm{Tr}(\hat{\psi}^2)}{2} + \frac{\mathrm{Tr}(\hat{g})}{2} + O(\|Y\|^3_{C^1})$$

The integral of the Jacobian is 1, so integrating the previous line over M against volume we obtain

$$1 = 1 + \int_M \mathrm{Tr}(\hat{\psi}) + \frac{(\mathrm{Tr}(\hat{\psi}))^2 - \mathrm{Tr}(\hat{\psi}^2)}{2} + \frac{\mathrm{Tr}(\hat{g})}{2}\, d\mathrm{vol} + O(\|Y\|^3_{C^1}).$$

Thus

$$\int_M \mathrm{Tr}(\hat{\psi}) + \frac{\mathrm{Tr}(\hat{g})}{2} - \frac{\mathrm{Tr}(\hat{\psi}^2)}{2}\, d\mathrm{vol} = -\int_M \frac{(\mathrm{Tr}(\hat{\psi}))^2}{2}\, d\mathrm{vol} + O(\|Y\|^3_{C^1}).$$

[3] Recall the usual Taylor expansion $\mathrm{Det}(\mathrm{Id} + A) = 1 + \mathrm{Tr}(A) + \frac{(\mathrm{Tr}(A))^2 - \mathrm{Tr}(A^2)}{2} + O(\|A\|^3)$. We combine this with the first order Taylor expansion

$$\det(\mathrm{Id}, \mathrm{Id}, \mathrm{Id} + G) = \sqrt{\mathrm{Det}(\mathrm{Id} + G)} = \sqrt{1 + \mathrm{Tr}(G) + O(\|G\|^2)} = 1 + \mathrm{Tr}(G)/2 + O(\|G\|^2).$$

Now, we integrate (111) over M and apply the previous line to eliminate the non-quadratic terms. This gives

$$\int_{\mathrm{Gr}_r(M)} \ln\det(D_x\psi_Y, \mathrm{Id}, g_{\psi_Y(x)} \mid E_x)\,dE_x$$
$$= \int_M -\frac{r}{2d}(\mathrm{Tr}(\hat{\psi}_x))^2 + \frac{r(d-r)}{(d+2)(d-1)}\mathrm{Tr}(K_x^2)\,d\mathrm{vol} + O(\|Y\|_{C^1}^3), \quad (112)$$

where we have written $\hat{\psi}_x$ and K_x to emphasize the basepoint. The formula above is not yet very usable as both K_x and $\hat{\psi}_x$ are defined in terms of exponential charts. We now obtain an intrinsic expression for these terms. Recall that pointwise we use the L^2 norm on tensors. Below we suppress the x in $\|E_C(x)\|$ and $\hat{\psi}_x$.

Lemma 60. *Let E_C be the conformal strain tensor associated to ψ_Y. Then*

$$\int_M (\mathrm{Tr}(\hat{\psi}_x))^2\,d\mathrm{vol} = \int \|E_C\|^2\,d\mathrm{vol} + O(\|Y\|_{C^1}^3).$$

Proof. We use an exponential chart and compute a coordinate expression for $\|E_C\|^2$ in the center of this chart. As before, write $D\psi_Y = \mathrm{Id} + \hat{\psi}$, where $\hat{\psi} = O(\|Y\|_{C^1})$. Then working in exponential coordinates, we obtain

$$\mathrm{Tr}(\psi_Y^* g - g) = \mathrm{Tr}((\mathrm{Id}+\hat{\psi})^T(\mathrm{Id}+O(\|Y\|_{C^0}^2))(\mathrm{Id}+\hat{\psi}) - \mathrm{Id})$$
$$= \mathrm{Tr}(\mathrm{Id}+\hat{\psi}^T+\hat{\psi}-\mathrm{Id}) + O(\|Y\|_{C^1}^2) = 2\,\mathrm{Tr}(\hat{\psi}) + O(\|Y\|_{C^1}^2).$$

Thus since $\hat{\psi} = O(\|Y\|_{C^1})$, by definition of E_C, we have

$$\|E_C\|^2 = \left\|\frac{\mathrm{Tr}(\psi_Y^* g - g)}{2d}\mathrm{Id}\right\|^2 = \left\|\frac{2\,\mathrm{Tr}(\hat{\psi})}{2d}\mathrm{Id}\right\|^2$$
$$= \frac{\mathrm{Tr}(\hat{\psi})}{d}\|\mathrm{Id}\|^2 = \mathrm{Tr}(\hat{\psi}).$$

Integrating over M, we obtain the result. ∎

Lemma 61. *Let E_{NC} be the non-conformal strain tensor associated to ψ_Y and let K_x be as in (110). Then*

$$\int_M \mathrm{Tr}(K_x^2)\,d\mathrm{vol} = \int_M \|E_{NC}\|^2\,d\mathrm{vol} + O(\|Y\|_{C^1}^3).$$

Proof. As before, we first compute a local expression for the integrand and check that this expression is comparable to the local expression for the non-conformal strain tensor. We compute at the center of an exponential chart. As before, write $D\psi_Y = \mathrm{Id} + \hat{\psi}$ where $\hat{\psi} = O(\|Y\|_{C^1})$. In this case

$$\psi_Y^* g = (\mathrm{Id}+\hat{\psi})^T(\mathrm{Id}+O(\|Y\|_{C^0}^2))(\mathrm{Id}+\hat{\psi}) = \mathrm{Id}+\hat{\psi}^T+\hat{\psi}+O(\|Y\|_{C^1}^2).$$

Using the above line and the definition of E_{NC} we then compute

$$\begin{aligned}
\|E_{NC}\|^2 &= \left\| \frac{1}{2}\left(\psi_Y^* g - g - \frac{\operatorname{Tr}(\psi_Y^* g - g)}{d} g \right) \right\|^2 \\
&= \frac{1}{4} \| (\mathrm{Id} + \hat{\psi})^T (\mathrm{Id} + O(\|Y\|_{C^0}^2)(\mathrm{Id} + \hat{\psi}) - \mathrm{Id} - 2\frac{\operatorname{Tr}(\hat{\psi})}{d}\mathrm{Id} + O(\|Y\|_{C^1}^2)\|^2 \\
&= \frac{1}{4} \left\| \hat{\psi}^T + \hat{\psi} - 2\frac{\operatorname{Tr}(\hat{\psi})}{d}\,\mathrm{Id} + O(\|Y\|_{C^1}^2) \right\|^2 \\
&= \frac{1}{4} \operatorname{Tr}\left(\left(\hat{\psi}^T + \hat{\psi} - 2\frac{\operatorname{Tr}(\hat{\psi})}{d}\,\mathrm{Id} + O(\|Y\|_{C^1}^2) \right)^2 \right) \\
&= \operatorname{Tr}\left(\left(\frac{\hat{\psi}^T + \hat{\psi}}{2} - \frac{\operatorname{Tr}(\hat{\psi})}{d}\,\mathrm{Id} \right)^2 \right) + O(\|Y\|_{C^1}^3) = \operatorname{Tr}(K^2) + O(\|Y\|_{C^1}^3).
\end{aligned}$$

By integrating the above equality over M, the result follows. ∎

Finally, Proposition 59 follows by applying Lemmas 60 and 61 to (112), which gives

$$\begin{aligned}
&\int_{\mathrm{Gr}_r(M)} \ln\det(D_x\psi_Y, \mathrm{Id}, g_{\psi_{Y(x)}} \mid E_x)\, dE_x \\
&\quad = -\frac{r}{2d}\int_M \|E_C\|^2\, d\mathrm{vol} + \frac{r(d-r)}{(d+2)(d-1)}\int_M \|E_{NC}\|^2\, d\mathrm{vol} + O(\|Y\|_{C^1}^3). \quad ∎
\end{aligned}$$

E.2. Approximation of integrals over Grassmannians

Let $\mathbb{G}_{r,d}$ be the Grassmannian of r-planes in $\mathbb{R}^d$. In this subsection, we prove the following simple estimate.

Proposition 62. *For $1 \le r \le d$, let $\Lambda_r\colon \mathrm{End}(\mathbb{R}^d) \to \mathbb{R}$ be defined by*

$$\Lambda_r(L) := \int_{\mathbb{G}_{r,d}} \ln\det(\mathrm{Id} + L, \mathrm{Id}, \mathrm{Id} \mid E)\, dE,$$

where dE denotes the Haar measure on $\mathbb{G}_{r,d}$. Then the second order Taylor approximation for Λ_r at 0 is

$$\Lambda_r(L) = \frac{r}{d}\operatorname{Tr}(L) + \left[-\frac{r}{2d}\operatorname{Tr}(L^2) + \frac{r(d-r)}{(d+2)(d-1)}\operatorname{Tr}(K^2) \right] + O(\|L\|^3),$$

where

$$K = \frac{L + L^T}{2} - \frac{\operatorname{Tr}(L)}{d}\,\mathrm{Id}.$$

Let $\lambda_r(L) = \Lambda_r(L) - \Lambda_{r-1}(L)$. Then the above expansion implies

$$\lambda_r(L) = \frac{1}{d}\operatorname{Tr}(L) + \left[-\frac{1}{2d}\operatorname{Tr}(L^2) + \frac{d-2r+1}{(d+2)(d-1)}\operatorname{Tr}(K^2) \right] + O(\|L\|^3).$$

Proof. Before beginning, note from the definition of Λ_r that if U is an orthogonal transformation, then $\Lambda_r(U^T LU) = \Lambda_r(L)$. Consequently, if α_i is the ith term in the Taylor expansion of Λ_r, then α_i is invariant under conjugation by isometries.

The map Λ_r is smooth, so it admits a Taylor expansion

$$\Lambda_r(L) = \alpha_1(L) + \alpha_2(L) + O(\|L\|^3),$$

where α_1 is linear in L and α_2 is quadratic in L. The rest of the proof is a calculation of α_1 and α_2. Before we begin this calculation we describe the approach. In each case, we reduce to the case of a symmetric matrix L. Then restricted to symmetric matrices, we diagonalize. There are few linear or quadratic maps from $\mathrm{End}(\mathbb{R}^n)$ to $\mathbb{R}$ that are invariant under conjugation by an orthogonal matrix. We then write α_i as a linear combination of such invariant maps from $\mathrm{End}(\mathbb{R}^n)$ to $\mathbb{R}$ and then solve for the coefficients of this linear combination.

We begin by calculating α_1.

Claim 3. *With notation as above,*

$$\alpha_1(L) = \frac{r}{d}\operatorname{Tr}(L).$$

Proof. Let $\widetilde{\Lambda}_r(\mathrm{Id} + L) = \Lambda_r(L)$. Then from the definition, note that if U is an isometry then $\widetilde{\Lambda}_r(U(\mathrm{Id} + L)) = \widetilde{\Lambda}_r((\mathrm{Id} + L)U)) = \Lambda_r(L)$. Suppose that O_t is some path tangent to $\mathrm{O}(n) \subset \mathrm{End}(\mathbb{R}^n)$ such that $O_0 = \mathrm{Id}$. Then $\widetilde{\Lambda}_r(O_t) = 0$. Write $O_t = \mathrm{Id} + tS + O(t^2)$ where S is skew-symmetric. Then we see that

$$\widetilde{\Lambda}_r(\mathrm{Id} + tS + O(t^2)) = O(t^2),$$

So, $\Lambda_r(tS) = O(t^2)$. Hence α_1 vanishes on skew-symmetric matrices.

Thus it suffices to evaluate α_1 restricted to symmetric matrices. Suppose that A is a symmetric matrix. Then there exists an orthogonal matrix U such that $U^T AU$ is diagonal. Restricted to the space of diagonal matrices, which we identify with $\mathbb{R}^d$ in the natural way, observe that $\alpha_1 \colon \mathbb{R}^d \to \mathbb{R}$ is invariant under permutation of the coordinates in $\mathbb{R}^d$ because it is invariant under conjugation by isometries. There is a one-dimensional space of maps having this property, and it is spanned by the trace, Tr. So, $\alpha_1(A) = \alpha_1(U^T AU) = a_1 \operatorname{Tr}(A)$ for some constant a_1. To compute a_1 it suffices to consider a specific matrix, e.g. $A = \mathrm{Id}$:

$$\begin{aligned}\alpha_1(\mathrm{Id}) &= \frac{d}{d\epsilon}\int \ln\det(\mathrm{Id} + \epsilon\,\mathrm{Id} \mid E)\, dE\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\int \ln(\mathrm{Id} + \epsilon)^r\, dE\bigg|_{\epsilon=0}\\ &= \frac{d}{d\epsilon} r\ln(1+\epsilon)\bigg|_{\epsilon=0} = r.\end{aligned}$$

So, $a_1 = r/d$. Thus for $L \in \mathrm{End}(\mathbb{R}^d)$, $\alpha_1(L) = \frac{r}{d}\operatorname{Tr}((L + L^T)/2) = \frac{r}{d}\operatorname{Tr}(L)$. ∎

We now compute α_2.

Claim 4. *With notation as in the statement of Proposition 62,*

$$\alpha_2(L) = -\frac{r}{2d}\operatorname{Tr}(L^2) + \frac{r(d-r)}{(d+2)(d-1)}\operatorname{Tr}(K^2).$$

Proof. Let $\widetilde{\Lambda}_r(\mathrm{Id}+L) = \Lambda_r(L)$. From the definition, note that for an isometry U, we have $\widetilde{\Lambda}_r((\mathrm{Id}+L)U) = \widetilde{\Lambda}_r(U)$. Fix L and let $J = (L - L^T)/2$. Observe that

$$(\mathrm{Id}+L)e^{-J} = \mathrm{Id} + (L-J) + (J^2/2 - LJ) + O(|L|^3).$$

Thus we see that

$$\begin{aligned}\Lambda_r(L) &= \widetilde{\Lambda}_r(\mathrm{Id}+L) = \widetilde{\Lambda}_r((L+\mathrm{Id})e^{-J})\\ &= \widetilde{\Lambda}_r(\mathrm{Id} + (L-J) + (J^2/2 - LJ) + O(|L|^3))\\ &= \Lambda_r((L-J) + (J^2/2 - LJ)) + O(|L|^3).\end{aligned}$$

Now comparing the two Taylor expansions of $\widetilde{\Lambda}_r(\mathrm{Id}+L)$, we find

$$\alpha_2(L) = \alpha_2(L-J) + \alpha_1(J^2/2 - LJ).$$

Thus, as we have already determined α_1,

$$\alpha_2(L) = \alpha_2((L+L^T)/2) + \frac{r}{d}\operatorname{Tr}(J^2/2 - LJ).$$

So, we are again reduced to the case of a symmetric matrix S. In fact, by invariance of α_2 under conjugation by isometries, we are reduced to determining α_2 on the space of diagonal matrices. Identify $\mathbb{R}^d$ with diagonal matrices as before. We see that α_2 is a symmetric polynomial of degree 2 in d variables. The space of such polynomials is spanned by $\sum x_i^2$ and $\sum_{i,j} x_i x_j$. It is convenient to observe that for a diagonal matrix D, $\operatorname{Tr}(D^2)$ and $(\operatorname{Tr}(D))^2$ span this space as well. Hence

$$\alpha_2(S) = b_1(\operatorname{Tr}(S))^2 + b_2\operatorname{Tr}(S^2).$$

Now in order to calculate b_1 and b_2 we will explicitly calculate $\alpha_2(\mathrm{Id})$ and $\alpha_2(P)$, where P is the orthogonal projection onto a coordinate axis.

In the first case,

$$\begin{aligned}2\alpha_2(\mathrm{Id}) &= \frac{d}{d\epsilon_1}\frac{d}{d\epsilon_2}\int_{\mathbb{G}_{r,d}} \ln\det((\mathrm{Id}+\epsilon_1+\epsilon_2) \mid E)\,dE\bigg|_{\epsilon_1=0,\epsilon_2=0}\\ &= \frac{d^2}{d\epsilon^2}\ln(1+\epsilon)^r\bigg|_{\epsilon=0} = -r.\end{aligned}$$

So, $\alpha_2(\mathrm{Id}) = -r/2$.

Next suppose that P is projection onto a fixed vector e. Suppose that $\angle(e,E) = \theta$. We now compute $\ln\det(\mathrm{Id}+\epsilon P \mid E)$. We fix a useful basis of E. Let v be a unit vector making angle $\angle(e, E)$ with e, and let $e_2, \dots, e_r$ be unit vectors in E that are orthogonal to e and v. Then using the basis $v, e_2, \dots, e_r$, we see that

$$\begin{aligned}\det(\mathrm{Id} + \epsilon P \mid E) &= \frac{\|(\mathrm{Id} + \epsilon P)v \wedge (\mathrm{Id} + \epsilon P)e_2 \wedge \cdots \wedge (\mathrm{Id} + \epsilon P)e_r\|}{\|v \wedge \cdots \wedge e_r\|} \\ &= \sqrt{\langle (\mathrm{Id} + \epsilon P)v, (\mathrm{Id} + \epsilon P)v \rangle},\end{aligned}$$

by considering the determinant defining the wedge product. But then as $Pv = \cos(\theta)e$,

$$\begin{aligned}\sqrt{\langle v + \epsilon\cos(\theta)e, v + \epsilon\cos(\theta)e \rangle} &= \sqrt{\langle v, v\rangle + 2\epsilon\cos(\theta)\langle v, e\rangle + \epsilon^2\langle Pv, Pv\rangle} \\ &= \sqrt{1 + 2\epsilon\cos^2(\theta) + \epsilon^2\cos^2\theta}.\end{aligned}$$

Now, the Taylor approximation for $\ln\sqrt{1+x}$ at $x = 0$ is $x/2 - x^2/4 + O(x^3)$, so

$$\ln\det(\mathrm{Id} + \epsilon P \mid E) = \epsilon\cos\angle(E, e) + \epsilon^2\left[\frac{\cos\angle(E, e)}{2} - \cos^4\angle(E, e)\right] + O(\epsilon^3).$$

Hence, as this estimate is uniform over E, by integrating we have

$$\begin{aligned}&\int_{\mathbb{G}_{r,d}} \ln\det(\mathrm{Id} + \epsilon P \mid E)\, dE \\ &\quad = \epsilon\int_{\mathbb{G}_{r,d}} \cos^2\angle(E, e)\, dE + \epsilon^2\int_{\mathbb{G}_{r,d}}\left[\frac{\cos^2\angle(E, e)}{2} - \cos^4\angle(E, e)\right] dE + O(\epsilon^3).\end{aligned}$$

So, we are reduced to calculating the coefficient of ϵ^2 in the above expression. One may rewrite the above integrals in the following manner, by definition of the Haar measure because $\mathbb{G}_{r,d}$ is a homogeneous space of $\mathrm{SO}(d)$. Write $x_1, \dots, x_d$ for the restriction of the Euclidean coordinates to the sphere. By fixing the coordinate plane $E_0 = \operatorname{span}\{e_1, \dots, e_r\}$, and letting $\theta = \angle((x_1, \dots, x_d), E)$ we then have $\cos(\theta) = \sqrt{\sum_{i=1}^r x_i^2}$. Thus

$$\begin{aligned}\int_{\mathbb{G}_{r,d}} \cos^2\angle(E, e)\, dE &= \int_{\mathrm{SO}_d} \cos^2\angle(gE_0, e)\, dg = \int_{\mathrm{SO}_d} \cos^2\angle(E_0, ge)\, dg \\ &= \int_{S^{d-1}} \cos^2\angle(E_0, x)\, dx = \int_{S^{d-1}} \sum_{i=1}^r x_i^2\, dx.\end{aligned}$$

Similarly, fixing the plane $E_0 = \operatorname{span}\{e_1, \dots, e_r\}$, we see that since $\cos^4\angle(E_0, x) = (\sum_{i=1}^r x_i^2)^2$,

$$\int_{\mathbb{G}_{r,d}} \cos^4\angle(E, e)\, dE = \int_{S^{d-1}} \Big(\sum_{i=1}^r x_i^2\Big)^2\, dx.$$

The evaluation of these integrals is immediate by using the following standard formulas:

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

Thus we see that

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

It follows that

$$\alpha_2(P) = \frac{r}{2d} - \frac{r(r+2)}{d(d+2)}.$$

Returning to b_1, b_2, the coefficients of $(\mathrm{Tr}(S))^2$ and $\mathrm{Tr}(S^2)$, respectively, combining the cases of Id and P gives

$$-\frac{r}{2} = b_1 d^2 + b_2 d, \quad \frac{r}{2d} - \frac{r(r+2)}{d(d+2)} = b_1 + b_2.$$

We can now solve for b_1 and b_2 with respect to this basis of the space of conjugation invariant quadratic functionals. However, the computation will be more direct if instead we use a different basis and write $\alpha_2(S)$ as

$$b_1(\mathrm{Tr}(S))^2 + b_2 \,\mathrm{Tr}\left(\left(S - \frac{\mathrm{Tr}(S)}{d} \right)^2 \right),$$

so that the second term is trace 0. Our computations above now show that

$$-\frac{r}{2} = b_1 d^2 + 0$$

and

$$\frac{r}{2d} - \frac{r(r+2)}{d(d+2)} = b_1 + \frac{d-1}{d} b_2 \left(= b_1(\mathrm{Tr}(P))^2 + b_2 \,\mathrm{Tr}\left(\left(P - \frac{\mathrm{Tr}(P)}{d} \,\mathrm{Id} \right)^2 \right) \right).$$

The first equation implies that

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

The left hand side of the second equation of the pair is equal to

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

This gives

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

So, for symmetric L, we have

$$\alpha_2(S) = \frac{-r}{2d^2}(\mathrm{Tr}(S))^2 + \left(\frac{r(d-r)}{(d-1)(d+2)} - \frac{r}{2d} \right) \mathrm{Tr}\left(\left(S - \frac{\mathrm{Tr}(S)}{d} \,\mathrm{Id} \right)^2 \right). \tag{113}$$

Recall that we specialized to the case of a symmetric matrix, and that for a non-symmetric matrix there is another term. For $L \in \mathrm{End}(\mathbb{R}^d)$, setting $J = (L - L^T)/2$ as before, we get

$$\alpha_2(L) = \alpha_2\left(\frac{L + L^T}{2} \right) + \frac{r}{d} \,\mathrm{Tr}\left(\frac{J^2}{2} - LJ \right).$$

To simplify this we compute that

$$\mathrm{Tr}\left(\frac{J^2}{2} - LJ\right) = \mathrm{Tr}\left(\frac{L^2 - LL^T - L^T L + (L^T)^2}{8} - L\frac{L - L^T}{2}\right) = \mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right).$$

Write

$$S = \frac{L + L^T}{2}.$$

Observe that for an arbitrary matrix X, $\mathrm{Tr}((X - \frac{\mathrm{Tr}(X)}{d}\,\mathrm{Id})^2) = \mathrm{Tr}(X^2) - (\mathrm{Tr}(X))^2/d$. Thus

$$\begin{aligned}
&-\frac{r}{2d^2}(\mathrm{Tr}(S))^2 - \frac{r}{2d}\,\mathrm{Tr}\left(\left(S - \frac{\mathrm{Tr}(S)}{d}\,\mathrm{Id}\right)^2\right) + \frac{r}{d}\,\mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right)\\
&\quad = -\frac{r}{2d^2}(\mathrm{Tr}(S))^2 - \frac{r}{2d}\,\mathrm{Tr}(S^2) - \frac{-r}{2d^2}(\mathrm{Tr}(S))^2 + \frac{r}{d}\,\mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right)\\
&\quad = -\frac{r}{2d}\,\mathrm{Tr}(S^2) + \frac{r}{d}\,\mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right)\\
&\quad = \frac{r}{d}\left[\frac{-1}{2}\,\mathrm{Tr}\left(\left(\frac{L + L^T}{2}\right)^2\right) + \mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right)\right]\\
&\quad = \frac{r}{d}\left[\frac{-1}{2}\,\mathrm{Tr}\left(\frac{L^2 + (L^T)^2 + 2LL^T}{4}\right) + \mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right)\right] = -\frac{r}{2d}\,\mathrm{Tr}(L^2).
\end{aligned}$$

From above, we have

$$\begin{aligned}
\alpha_2(L) = &-\frac{r}{2d^2}(\mathrm{Tr}(S))^2 + \left(\frac{r(d-r)}{(d-1)(d-2)} - \frac{r}{2d}\right)\mathrm{Tr}\left(\left(S - \frac{\mathrm{Tr}(S)}{d}\,\mathrm{Id}\right)^2\right)\\
&+ \frac{r}{d}\,\mathrm{Tr}\left(\frac{LL^T - L^2}{4}\right).
\end{aligned}$$

So substituting the previous calculation we obtain

$$\alpha_2(L) = -\frac{r}{2d}\,\mathrm{Tr}(L^2) + \frac{r(d-r)}{(d-1)(d-2)}\,\mathrm{Tr}\left(\left(\frac{L + L^T}{2} - \frac{\mathrm{Tr}(L)}{d}\,\mathrm{Id}\right)^2\right),$$

which is the desired formula. ∎

We have now calculated α_1 and α_2. This concludes the proof of Proposition 62. ∎

We will also use a first order Taylor expansion with respect to the metric.

Proposition 63. *Let $\Lambda_r(G)$ be defined for symmetric matrices G by*

$$\Lambda_r(G) := \int_{\mathbb{G}_{r,d}} \ln\det(\mathrm{Id}, \mathrm{Id}, \mathrm{Id} + G \mid E)\, dE.$$

Then $\Lambda_r(G)$ admits the following Taylor development:

$$\Lambda_r(G) = \frac{r}{2d}\,\mathrm{Tr}(G) + O(\|G\|^2).$$

Proof. The proof is substantially similar to that of the previous proposition. Let α_1 denote the first term in the Taylor expansion. Note that if U is an isometry then $\Lambda_r(U^T GU) = \Lambda_r(G)$. Thus α_1 is invariant under conjugation by isometries. Thus by conjugating by an orthogonal matrix, we are reduced to the case of G and diagonal matrix. As before, we see that $\alpha_1(D)$ is a multiple of $\mathrm{Tr}(D)$ as Tr spans the linear forms on $\mathbb{R}^d$ that are invariant under permutation of coordinates.

Thus it suffices to calculate the derivative for $D = \mathrm{Id}$. So, we see that

$$\alpha_1(\mathrm{Id}) = \frac{d}{d\epsilon} \int_E \ln \det(\mathrm{Id}, \mathrm{Id}, \mathrm{Id} + \epsilon\, \mathrm{Id} \mid E)\, dE.$$

Thus the integral is equal to $\ln \sqrt{(1+\epsilon)^r}$ on every plane E. Hence the derivative is $r/2$ and so

$$\alpha_1(\mathrm{Id}) = \frac{r}{2} = \frac{r}{2d} \mathrm{Tr}(\mathrm{Id}).$$

And so the result follows. ∎

Acknowledgments. The author thanks Aaron Brown and Amie Wilkinson for their critical comments during all parts of this project. The author also thanks Dmitry Dolgopyat for his generosity in answering the author's questions about [10]. The author is also grateful to the anonymous referees for carefully reading the manuscript and providing many useful comments and suggestions.

Funding. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1746045.

References

[1] Arnold, L.: Random Dynamical Systems. Springer Monogr. Math., Springer, Berlin (1998) Zbl 0906.34001 MR 1723992

[2] Bergh, J., Löfström, J.: Interpolation Spaces: an Introduction. Springer (1976) Zbl 0344.46071 MR 0482275

[3] Bouland, A., Ozols, M.: Trading inverses for an irrep in the Solovay–Kitaev theorem. In: 13th Conference on the Theory of Quantum Computation, Communication and Cryptography, LIPIcs. Leibniz Int. Proc. Inform. 111, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, art. 6, 15 pp. (2018) MR 3835813

[4] Butler, C.: Characterizing symmetric spaces by their Lyapunov spectra. arXiv:1709.08066 (2017)

[5] Cheeger, J., Ebin, D. G.: Comparison Theorems in Riemannian Geometry. North-Holland Math. Library 9, North-Holland, Amsterdam (1975) Zbl 0309.53035 MR 0458335

[6] Damjanović, D., Fayad, B.: On local rigidity of partially hyperbolic affine $\mathbb{Z}^k$ actions. J. Reine Angew. Math. **751**, 1–26 (2019) Zbl 1430.37031 MR 3956690

[7] Dawson, C. M., Nielsen, M. A.: The Solovay–Kitaev algorithm. Quantum Information Comput. **6**, 81–95 (2006) Zbl 1152.81706 MR 2212257

[8] DeWitt, J.: Local Lyapunov spectrum rigidity of nilmanifold automorphisms. J. Modern Dynam. **17**, 65–109 (2021) Zbl 1470.37046 MR 4251938

[9] Dolgopyat, D.: On mixing properties of compact group extensions of hyperbolic systems. Israel J. Math. **130**, 157–205 (2002) Zbl 1005.37005 MR 1919377

[10] Dolgopyat, D., Krikorian, R.: On simultaneous linearization of diffeomorphisms of the sphere. Duke Math. J. **136**, 475–505 (2007) Zbl 1113.37038 MR 2309172

[11] Fayad, B., Khanin, K.: Smooth linearization of commuting circle diffeomorphisms. Ann. of Math. (2) **170**, 961–980 (2009) Zbl 1177.37045 MR 2552115

[12] Field, M.: Generating sets for compact semisimple Lie groups. Proc. Amer. Math. Soc. **127**, 3361–3365 (1999) Zbl 0924.22013 MR 1618662

[13] Gogolev, A.: Rigidity lecture notes. https://people.math.osu.edu/gogolyev.1/index_files/CIRM_notes_all.pdf (2019)

[14] Gogolev, A., Rodriguez Hertz, F.: Smooth rigidity for very non-algebraic expanding maps. J. Eur. Math. Soc. (online, 2022)

[15] Gogolev, A., Kalinin, B., Sadovskaya, V.: Local rigidity of Lyapunov spectrum for toral automorphisms. Israel J. Math. **238**, 389–403 (2020) Zbl 1452.37022 MR 4145804

[16] Hamilton, R. S.: The inverse function theorem of Nash and Moser. Bull. Amer. Math. Soc. (N.S.) **7**, 65–222 (1982) Zbl 0499.58003 MR 656198

[17] Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems. Encyclopedia Math. Appl. 54, Cambridge Univ. Press (1997) Zbl 0878.58019 MR 1326374

[18] Helgason, S.: Differential Geometry, Lie Groups, and Symmetric Spaces. Grad. Stud. Math. 34, Amer. Math. Soc., Providence, RI (2001) Zbl 0993.53002 MR 1834454

[19] Hörmander, L.: The boundary problems of physical geodesy. Arch. Ration. Mech. Anal. **62**, 1–52 (1976) Zbl 0331.35020 MR 602181

[20] Kifer, Y.: Ergodic Theory of Random Transformations. Progr. Probab. Statist. 10, Birkhäuser Boston, Boston, MA (1986) Zbl 0604.28014 MR 884892

[21] Lai, W. M., Rubin, D., Krempl, E.: Introduction to Continuum Mechanics. 4th ed., Butterworth-Heinemann (2009)

[22] Lee, J. M.: Introduction to Riemannian Manifolds. 2nd ed., Grad. Texts in Math. 176, Springer, Cham (2018) Zbl 1409.53001 MR 3887684

[23] Malicet, D.: Lyapunov exponent of random dynamical systems on the circle. Ergodic Theory Dynam. Systems **42**, 2080–2107 (2022) Zbl 1500.37031 MR 4417346

[24] Malicet, D.: On simultaneous linearization of diffeomorphisms of $\mathbb{T}^2$. Preprint

[25] Moser, J.: On commuting circle mappings and simultaneous Diophantine approximations. Math. Z. **205**, 105–121 (1990) Zbl 0689.58031 MR 1069487

[26] Petković, B.: Classification of perturbations of Diophantine $\mathbb{Z}^m$ actions on tori of arbitrary dimension. Regul. Chaotic Dynam. **26**, 700–716 (2021) Zbl 1490.37085 MR 4348431

[27] Saghin, R., Yang, J.: Lyapunov exponents and rigidity of Anosov automorphisms and skew products. Adv. Math. **355**, art. 106764, 45 pp. (2019) Zbl 1461.37033 MR 3995224

[28] Shankar, K.: Isometry groups of homogeneous spaces with positive sectional curvature. Differential Geom. Appl. **14**, 57–78 (2001) Zbl 0991.53023 MR 1812525

[29] Wallach, N. R.: Harmonic Analysis on Homogeneous Spaces. Pure Appl. Math. 19, Dekker, New York (1973) Zbl 1406.22001 MR 0498996

[30] Wolf, J. A.: Spaces of Constant Curvature. 2nd ed., Department of Mathematics, Univ. of California, Berkeley, CA (1972) Zbl 1216.53003 MR 0343213