# LOGARITHMIC IMPROVEMENTS IN THE WEYL LAW AND EXPONENTIAL BOUNDS ON THE NUMBER OF CLOSED GEODESICS ARE PREDOMINANT

YAIZA CANZANI AND JEFFREY GALKOWSKI

ABSTRACT. Let $M$ be a smooth, compact manifold of dimension $d$ without boundary. We introduce the concept of predominance for Riemannian metrics on $M$, a notion analogous to full Lebesgue measure which, in particular, implies density. We show that there is $\Omega > 1$ such that, for a predominant $\mathcal{C}^\nu$ metric, the number, $\mathfrak{c}(T, g)$, of closed geodesics of length $\leq T$ satisfies

$$\log \mathfrak{c}(T, g) = O(T^\Omega).$$

In addition, for $g$ a Riemannian metric on $M$, let $0 = \lambda_1^2(g) < \lambda_2^2(g) \leq \lambda_3^2(g) \dots$ be the eigenvalues of $-\Delta_g$. The Weyl law states that there is $c_d > 0$ such that

$$\#\{j : \lambda_j(g) \leq \lambda\} = c_d \operatorname{vol}_g(M)\lambda^d + E(\lambda, g)$$

with $E(\lambda, g) = O(\lambda^{d-1})$ as $\lambda \to \infty$. We show that for $\nu > 0$ large enough there is $\Omega > 1$ such that for a predominant $\mathcal{C}^\nu$ metric

$$E(\lambda, g) = O(\lambda^{d-1}/(\log \lambda)^{1/\Omega}).$$

After an application of recent results of the authors in the case of the Weyl law [CG20], these estimates follow from a study of the non-degeneracy properties of nearly closed orbits for predominant sets of metrics.

## 1. INTRODUCTION

We study properties of the geodesic flow and remainders in the Weyl law for 'typical' metrics on a compact manifold without boundary. Since the space of Riemannian metrics, $\mathscr{G}$, on a manifold cannot be endowed with a non-trivial, translation invariant Borel measure, we introduce an analog of full Lebesgue measure in infinite dimensions called *predominance*. We then study properties of the geodesic flow and remainders in the Weyl law for predominant sets of metrics.

The notion of predominance has three important properties: 1) any predominant set is dense, 2) a finite intersection of predominant sets is predominant, and 3) in finite dimensions, a predominant set has full Lebesgue measure. Heuristically, a set $G \subset \mathscr{G}$ is *predominant* if there is a family of submanifolds endowed with finite Borel measures $\{(\Gamma_g, \mu_g)\}_{g \in \mathscr{G}}$ such that $g \in \Gamma_g$, $\mu_g$ assigns a positive measure to any neighborhood of $g$, the map $g \mapsto \Gamma_g$ is $\mathcal{C}^1$, and $G \cap \Gamma_g$ has full $\mu_g$ measure for every $g \in \mathscr{G}$. For the careful definition of this concept, see Definition 2.4 and Remark 2.5.

**Remainders in the Weyl law.** Let $\nu \geq 0$ and $M$ be a compact $\mathcal{C}^\nu$ manifold without boundary, of dimension $d$. Let $\mathscr{G}^\nu$ denote the space of $\mathcal{C}^\nu$ Riemannian metrics on $M$ with the topology induced from the $\mathcal{C}^\nu$ norm on symmetric tensors. For $g \in \mathscr{G}^\nu$, let $-\Delta_g$ denote the (positive) Laplace–Beltrami operator with eigenvalues $0 = \lambda_1^2(g) < \lambda_2^2(g) \leq \lambda_3^2(g) \leq \dots$. Then, for $\lambda > 0$

define the eigenvalue counting function

$$N(\lambda, g) := \#\{j \,:\, \lambda_j(g) \le \lambda\}.$$

Our first theorem shows that the Weyl law has a *logarithmic* improvement for a predominant set of metrics. In what follows, $B_1$ denotes the ball of radius 1 in $\mathbb{R}^d$, $\mathrm{vol}_{\mathbb{R}^d}(B_1)$ denotes its volume, and $\mathrm{vol}_g(M)$ denotes the volume of $M$ as measured by the metric $g$.

**Theorem 1.1.** *Let $d \ge 2$. There is $\nu_0 > 0$ and for all $\nu > 0$ there is $\Omega_\nu > 0$ such that the following holds. If $\nu \ge \nu_0$, $M$ is a compact $\mathcal{C}^\nu$-manifold of dimension $d$ without boundary, and $\Omega > \Omega_\nu$, then there is a predominant set $G_\Omega \subset \mathscr{G}^\nu$ (see Definition 2.4) such that for every $g \in G_\Omega$*

$$N(\lambda, g) = (2\pi)^{-d} \mathrm{vol}_g(M) \, \mathrm{vol}_{\mathbb{R}^d}(B_1)\lambda^d + O_g\big(\lambda^{d-1}/(\log \lambda)^{\frac{1}{\Omega}}\big), \qquad \lambda \to \infty.$$

*In particular, $G_\Omega$ is dense in $\mathscr{G}^\nu$.*

The constant $\Omega_\nu$ in Theorem 1.1 is explicit, and we can take

$$\Omega_\nu := 1 + \log_2(2(d-1)(2d+1)(\max(\nu, 6) + 3d - 1)) - \log_2(2d-1). \tag{1.1}$$

**Remark 1.1.** The authors wish to stress that, although our original motivation was to study typical properties of remainders in the Weyl law, the analysis in this article is dynamical in nature and studies predominant properties of the geodesic flow. Once these dynamical properties are established in Theorem 1.3 below, a direct application of the authors' work [CG20, Theorem 2] produces the Weyl remainder estimate of Theorem 1.1 as a corollary (See Section 3.3).

Defining $E(\lambda, g) := N(\lambda, g) - (2\pi)^{-d} \mathrm{vol}_g(M) \, \mathrm{vol}_{\mathbb{R}^d}(B_1)\lambda^d$, the Weyl law states that, for sufficiently smooth metrics, $E(\lambda, g) = O_g(\lambda^{d-1})$. This estimate is sharp on the round sphere and has a long history dating back to the work of Weyl [Wey12], who proved (in a slightly different context) that $E(\lambda, g) = o(\lambda^d)$. The estimate $E(\lambda, g) = O_g(\lambda^{d-1})$ was proved by Levitan [Lev52] and Avakumović [Ava56] after which Hörmander [Hör68] provided a general framework for studying such remainders, reproving this estimate and making far reaching generalizations.

Using this framework, Bérard [Bér77] showed that $E(\lambda, g) = O_g(\lambda^{d-1}/\log \lambda)$ on both surfaces without conjugate points and non-positively curved manifolds of any dimension. Duistermaat–Guillemin [DG75] showed that $E(\lambda, g) = o(\lambda^{d-1})$ provided that the measure of the set of closed geodesics in $S^*M$ is 0. Fifteen years later, Volovoy [Vol90a] provided estimates under dynamical conditions guaranteeing that $E(\lambda, g) = O_g(\lambda^{d-1}/\log \lambda)$ and verified these conditions for certain specific examples in [Vol90b]. The recent work of Bonthenneau [Bon17] improved a geometric estimate in Bérard's work, thus generalizing his result to manifolds without conjugate points of any dimension. Finally, in [CG20], the authors provided estimates on $E(\lambda, g)$ under assumptions on the volume of nearly closed geodesics which improve the results of [DG75, Bér77, Vol90a]. For manifolds with boundary, the analog of [DG75] was proved by Ivrii [Ivr80]. (For a more comprehensive history of the Weyl law, see [Ivr16].)

Manifolds where there are known polynomial improvements of the form $E(\lambda, g) = O_g(\lambda^{d-1-\varepsilon})$ are very rare. For instance, such estimates hold on the torus [Hux03, BW17], products of spheres [IW21], and other special integrable systems [Vol90b]. Nevertheless, it has long been expected that, for a 'typical' metric $g$, there exists $\varepsilon > 0$ such that $E(\lambda, g) = O_g(\lambda^{d-1-\varepsilon})$. However, until now, the best available result is that $E(\lambda, g) = o(\lambda^{d-1})$ for a Baire-generic set of $g$. This can

be recovered from the work of Sogge–Zelditch [SZ02] or can be seen by combining the remainder estimates in [DG75] with the bumpy metric theorem of Anosov and Abraham [Ano82, Abr70].

Theorem 1.1 improves on these bounds in two important ways. First, Baire genericity is replaced by the concept of predominance which is an analog of full Lebesgue measure in infinite dimensions. Just as in finite dimensions a full Lebesgue measure set is much more 'typical' than a Baire generic one (indeed, a Baire generic set may have measure 0), a predominant set in infinite dimensions is much more 'typical' than a Baire generic set. Second, although the change from $o(\lambda^{d-1})$ to $O(\lambda^{d-1}/(\log \lambda)^{1/\Omega})$ may seem small, this improvement requires the development of new ideas and requires subtle dynamical estimates. In addition, it is the only quantitative remainder estimate available for typical metrics.

**Remark 1.2.** We have not attempted to make the value of $\nu_0$ from in Theorem 1.1 explicit. However, it is likely that, following the arguments in [CG20] carefully, $\nu_0$ can be taken to be $\nu_0 = Cd$ for some $C > 0$.

**Remark 1.3.** Although we have kept careful track of the constant $\Omega_\nu$ in (1.1), we do not expect it to be optimal. Indeed, we conjecture that $\Omega_\nu$ could be replaced by $1 + \varepsilon$ for any $\varepsilon > 0$. This would also allow us to obtain the same estimate for predominant sets in $\mathscr{G}^\infty$. At the moment, when working in $\mathscr{G}^\infty$, we obtain weaker remainder estimates for predominant sets of metrics (see Remark 1.8).

**Growth of the number of periodic geodesics.** We next discuss the growth of the number of periodic geodesics of a given length. We say that a geodesic $\gamma \subset M$ is a primitive periodic geodesic with length $T > 0$, if there is a diffeomorphism $h : \mathbb{R}/T\mathbb{Z} \to \gamma$, such that $|\dot{h}|_{g(h(t))} = 1$,

$$(h(0), \dot{h}(0)) = (h(T), \dot{h}(T)), \qquad (h(0), \dot{h}(0)) \neq (h(t), \dot{h}(t)) \text{ for } t \in (0, T).$$

That is, $\gamma$ is a periodic geodesic and $T$ is its minimal period. For $T > 0$ and $g$ a Riemannian metric on $M$ let

$$\mathfrak{c}(T, g) := \#\{\gamma : \gamma \text{ is a primitive periodic geodesic for } g \text{ with length } \leq T\}.$$

We obtain the following theorem on the growth of $\mathfrak{c}(T, g)$.

**Theorem 1.2.** *Let $\nu \geq 5$, $M$ be a $\mathcal{C}^\nu$- compact manifold of dimension $d$ without boundary, and $\Omega_\nu$ as in (1.1). Then, for each $\Omega > \Omega_\nu$ there is a predominant set $G_\Omega \subset \mathscr{G}^\nu$ such that for $g \in G_\Omega$ there exists $C > 0$ such that for all $T > 0$*

$$\mathfrak{c}(T, g) \leq \exp\left(CT^\Omega\right).$$

*In particular, $G_\Omega$ is dense in $\mathscr{G}^\nu$.*

Bounds on $\mathfrak{c}(T, g)$ have a long history in the literature. For Baire generic metrics, a complete picture of the *non-quantitative* behavior of $\mathfrak{c}(T, g)$ is available. The bumpy metric theorem [Ano82, Abr70] can be used to show that for Baire generic smooth metrics, $g$, and all $T > 0$, $\mathfrak{c}(T, g) < \infty$ (see also [KT72]). Furthermore, Hingston [Hin84] showed that $\lim_{T \to \infty} \mathfrak{c}(T, g) = \infty$ for a Baire-generic set of metrics. Petkov and Stojanov have studied similar properties of closed billiards in generic domains in $\mathbb{R}^d$ [Sto87, PS87a, PS87b].

As with remainders in the Weyl law, quantitative estimates on $\mathfrak{c}(T, g)$ are much more subtle. Only recently, Contreras [Con10] showed that for $g$ in an open dense subset of $\mathscr{G}^2$ there is $c > 0$ such that

$$\log \mathfrak{c}(T, g) \geq cT - c. \tag{1.2}$$

One can check that this lower bound is optimal for any dense set of metrics (so certainly for any reasonable notion of a typical metric, including predominance). Indeed, in the case of manifolds with negative curvature, the works of Margulis [Mar69] and Bowen [Bow72] show that, for such a metric $g$, there is $\alpha > 0$ such that

$$\lim_{T \to \infty} \frac{1}{T} \log \mathfrak{c}(T, g) = \alpha.$$

In particular, this shows that there are open sets of metrics such that $\mathfrak{c}(T, g)$ grows exactly exponentially and hence that [Con10] gives a complete picture for lower bounds on $\mathfrak{c}(T, g)$ for typical $\mathcal{C}^2$ metrics. (See also [Kni98] for the case of compact rank 1 manifolds.)

On the other hand, as far as the authors are aware, Theorem 1.2 is the first quantitative upper bound on $\mathfrak{c}(T, g)$ for 'typical' metrics. One reason that lower bounds on $\mathfrak{c}(T, g)$ are well understood, but upper bounds are not is that one can find a structure, called a hyperbolic basic set (see [Con10]), which is stable under perturbation and guarantees the lower bound (1.2). Moreover, the existence of such a set can be guaranteed by studying the Poincaré maps associated to genuinely periodic orbits.

Unfortunately, no such structure for upper bounds exists and one must understand not only Poincaré maps of periodic orbits but also those of near periodic orbits. Indeed, although there is no rigorous proof at present, there is strong evidence that there is no quantitative upper bound on $\mathfrak{c}(T, g)$ which is Baire generic. For instance, in the case of diffeomorphims, Kaloshin [Kal00] showed that there is no growth rate for the number of periodic points that is Baire generic. (See Section 2 for an additional heuristic discussion for lack of Baire genericity.)

**Non-degeneracy of nearly periodic orbits.** Our main theorem, which implies both Theorems 1.1 and 1.2, controls how close two periodic orbits may be for a predominant set of metrics and can be used, for instance, to control the volume of nearly periodic orbits (see Section 3.1). The result will be stated in terms of how close $d\varphi_t^g$ may be to the identity for this set of metrics, where $\varphi_t^g$ is the geodesic flow for the metric $g$ at time $t$ acting on $S^*M$.

To understand how this is connected to the distance between periodic orbits, let $H_{|\xi|_g}$ denote the Hamiltonian vector field associated to $|\xi|_g$, and $\varphi_t^g := \exp(tH_{|\xi|_g}) : S^*M \to S^*M$. Observe that if $\rho$ is a $t$ periodic point (i.e. $\varphi_t^g(\rho) = \rho$ for some $t > 0$), and $v \in T_\rho S^*M/\mathbb{R}H_{|\xi|_g}(\rho)$ is not in the kernel of $I - d\varphi_t^g$, then any perturbation of the initial point $\rho$ in the direction of $v$ will not be periodic with period near $t$. (See Figure 2 for a schematic of such an orbit.) In particular, if the map

$$I - d\varphi_t^g : T_\rho S^*M/\mathbb{R}H_{|\xi|_g}(\rho) \to T_\rho S^*M/\mathbb{R}H_{|\xi|_g}(\rho)$$

is invertible, then there are no vectors in this kernel and hence every small perturbation of $\rho$ other than those along $H_{|\xi|_g}(\rho)$ will produce a point which is not periodic with period near $t$.

Motivated by this, we say that a $t$ periodic point, $\rho$, is *non-degenerate* if

$$I - d\varphi_t^g : T_\rho S^* M / \mathbb{R} H_{|\xi|_g}(\rho) \to T_\rho S^* M / \mathbb{R} H_{|\xi|_g}(\rho)$$

is invertible. The famous bumpy metric theorem [Abr70, Ano82] states that, for a Baire generic set of smooth metrics, every periodic trajectory is non-degenerate. In particular, this implies the finiteness of the set of closed geodesics of any bounded length.

Since we are interested in a quantitative version of non-degeneracy for both periodic geodesics and nearly periodic geodesics, we need to introduce a few concepts to make precise statements about non-degeneracy.

**Definition 1.4** (returning points)**.** Let $g \in \mathscr{G}^\nu$ and $\beta > 0$. For $t \in \mathbb{R}$ and $\rho \in S^* M$ we write

$$\rho \in \mathscr{R}(t, \beta, g) \qquad \text{if} \qquad d(\varphi_t^g(\rho), \rho) < \beta.$$

In this case, we say that $\rho$ *is $\beta$-returning for $g$ at time $t$.*

We also recall that the geodesic flow is a contact flow on $S^* M$ and thus there is a natural smooth decomposition of $TS^* M$ preserved by the geodesic flow (see e.g. [Pat99]). In particular, $\xi dx|_{TS^* M}$ is a contact form for $S^* M$ with the geodesic flow as its Reeb flow. Thus, for all $\rho \in S^* M$,

$$T_\rho S^* M = \mathscr{H}(\rho) \oplus \mathbb{R} H_{|\xi|_g}(\rho),$$

where $\mathscr{H}(\rho) := \ker(\xi dx|_{T_\rho S^* M})$ and $\oplus$ denotes the direct sum, and we have

$$d\varphi_t^g(\mathscr{H}(\rho)) = \mathscr{H}(\varphi_t^g(\rho)).$$

Since we work with nearly periodic orbits, we need to identify the tangent spaces at $\rho$ and $\varphi_t^g(\rho)$ when they are close. Let $g \in \mathscr{G}^\nu$ and $U \subset S^* M \times S^* M$. We say that $\mathcal{W}^U = \{\mathcal{W}_{\rho_2, \rho_1} : (\rho_2, \rho_1) \in U\}$ is *a family of transition maps for $g$ on $U$* if for each $(\rho_2, \rho_1) \in U$ the map $\mathcal{W}_{\rho_2, \rho_1}$ is an invertible linear transformation,

$$\mathcal{W}_{\rho_2, \rho_1} : T_{\rho_1} S^* M \to T_{\rho_2} S^* M, \qquad (\rho_2, \rho_1) \mapsto \mathcal{W}_{\rho_2, \rho_1} \text{ is Lipschitz}$$
$$\mathcal{W}_{\rho_1, \rho_1} = I, \qquad \mathcal{W}_{\rho_2, \rho_1} H_{|\xi|_g}(\rho_1) = H_{|\xi|_g}(\rho_2), \qquad \mathcal{W}_{\rho_2, \rho_1} \mathscr{H}(\rho_1) = \mathscr{H}(\rho_2). \tag{1.3}$$

Here, by asking that $(\rho_2, \rho_1) \mapsto \mathcal{W}_{\rho_2, \rho_1}$ be Lipschitz, we mean that for any choice of coordinates $\psi_i : W_i \to V_i \subset \mathbb{R}^{2d-1}$ near $\rho_i$ the map

$$W_1 \times W_2 \ni (x_1, x_2) \mapsto d\psi_2|_{\rho = \psi_2^{-1}(x_2)} \circ \mathcal{W}_{\psi_2^{-1}(x_2), \psi_1(x_1)} \circ d(\psi_1^{-1}(x))|_{x=x_1} \in \mathbb{GL}(2d-1)$$

is a Lipschitz family of matrices.

We say that a collection $\mathcal{W} = \{\mathcal{W}^{U_i}\}_{i=1}^N$ *is a family of transition maps for $g$* if $U_i \subset S^* M \times S^* M$ is open,

$$\{(\rho, \rho) : \rho \in S^* M\} \subset \bigcup_{i=1}^N U_i,$$

and, for each $i$, $\mathcal{W}^{U_i}$ is a family of transition maps for $g$ on $U_i$. We say $\mathcal{W}$ is a $\beta_0$-*family of transitions maps* if for each pair $(\rho_2, \rho_1)$ with $d(\rho_2, \rho_1) < \beta_0$, there is $i$ such that $(\rho_2, \rho_1) \in U_i$.

**Remark 1.5.** It will be convenient throughout the text to have a fixed reference metric on $M$. For this, we choose some $g_f \in \mathscr{G}^\nu$ and whenever we refer to a norm $|\cdot|$, on $T^* M$ or $TM$, it is the one induced by $g_f$. We will, in particular, use this metric to define $\mathcal{C}^{\nu'}$ norms for $\nu' \leq \nu$.

We are now in a position to define non-degeneracy of a nearly periodic orbit.

**Definition 1.6** (non-degenerate points). Let $g \in \mathscr{G}^\nu$, $\beta_0 > 0$ and $\mathcal{W} = \{\mathcal{W}^{U_i}\}_{i=1}^N$ be a $\beta_0$-family of transition maps for $g$. Let $0 < \beta < \beta_0$, $t \in \mathbb{R}$, $\rho \in \mathscr{R}(t, \beta, g)$, and $\alpha > 0$. We write $\rho \in \mathscr{N}(t, \alpha, (g, \mathcal{W}))$ if for every $i \in \{1, \dots N\}$ such that $(\varphi_t^g(\rho), \rho) \in U_i$,

$$|\Pi_{\varphi_t^g(\rho)}(\mathcal{W}^{U_i}_{\varphi_t^g(\rho), \rho} - d\varphi_t^g)v| \geq \alpha|\Pi_\rho v|, \quad v \in T_\rho S^*M.$$

Here, $\Pi_\rho : T_\rho S^*M \to T_\rho S^*M/\mathbb{R}H_p(\rho)$ denotes the natural projection map and, by an abuse of notation, $|\cdot|$ denotes the norm induced by the metric $g_f$ fixed above in Remark 1.5. In this case we say $\rho$ *is $\alpha$ non-degenerate for $(g, \mathcal{W})$ at time $t$.* (See Figure 2 for an example of a non-degenerate orbit.)

**Remark 1.7.** Although Definition 1.6 depends on the choice of the metric $g_f$, note that the norm induced by any other $g_f'$ is comparable to that induce by $g_f$.

The main theorem of this article shows that there is a predominant set of metrics such that *every* sufficiently returning geodesic is non-degenerate with the degree of non-degeneracy depending explicitly on the length of the trajectory. As far as the authors are aware, this theorem is the first quantitative estimate on non-degeneracy of orbits for typical metrics.

**Theorem 1.3.** *Let $\nu \geq 5$, $M$ be a compact $\mathcal{C}^\nu$ manifold of dimension $d$ without boundary, and $\Omega_\nu$ as in (1.1). Then for every $\Omega > \Omega_\nu$ there is a predominant set $G_\Omega \subset \mathscr{G}^\nu$ such that for all $g \in G_\Omega$ and every family of transition maps $\mathcal{W}$ for $g$ there are $C, c > 0$ such that*

$$\mathscr{R}(t, \boldsymbol{\beta}(t), g) \subset \mathscr{N}(t, \boldsymbol{\beta}(t), (g, \mathcal{W})) \quad \text{for } t > c$$

*where $\boldsymbol{\beta}(t) := C^{-C(t+1)^\Omega - 1}$, and*

$$d(\varphi_t^g(\rho), \rho) \geq c|t| \quad \text{for } |t| \leq c.$$

*In particular, $G_\Omega$ is dense in $\mathscr{G}^\nu$.*

**Remark 1.8.** The reason for the growth of $\Omega_\nu$ as $\nu \to \infty$ in Theorems 1.1, 1.2, and 1.3 comes from the fact that we make perturbations to the metric at increasingly small scales as the length of trajectories goes to infinity. Because of this, the size of these perturbations in $\mathcal{C}^\nu$ grows as $\nu \to \infty$ and this in turn results in weaker non-degeneracy statements. Moreover, with $f(t) : [0, \infty) \to [0, \infty)$ growing faster than any polynomial in $t$, if one replaces $\boldsymbol{\beta}(t)$ by $Ce^{Cf(t)}$ in Theorem 1.3, then one can work in $\mathcal{C}^\infty$. That is, we obtain predominance in the $\mathscr{G}^\infty$ topology.

**Remark 1.9.** The notion of predominance (see Section 2) involves using certain families of perturbations to probe the space of metrics. In this article, the predominance in Theorem 1.3 involves probing with the families of perturbations described in Section 9. However, one may wonder whether one can probe with other families of perturbations (e.g. conformal perturbations) and still obtain Theorem 1.3 for that family of probes. In fact, in Sections 7 and 8, we prove Proposition 8.9 which gives a result analogous to Theorem 1.3 under the assumption that a family of perturbations of metrics satisfies some abstract assumptions (see Definitions 7.2 and 7.3). The type of perturbations used to probe the space of metrics is then tied to the family of perturbations satisfying our abstract assumptions. It is only in Section 9 that we construct such a family of metric perturbations and it is likely that many other families of perturbations suffice. However, we do not pursue this here.

**Outline of the paper.** In Section 2, we define predominance and discuss the reasons for introducing this notion: predominance is in some sense more 'typical' than Baire generic and our results are unlikely to hold for a Baire generic set of metrics. Then, in Section 3, we use Theorem 1.3 to prove Theorems 1.1 and 1.2. Before starting the proof of Theorem 1.3, we give a detailed description of the ideas used in Section 4.

To begin the proof of Theorem 1.3, we study volumes of relevant sets of symplectic matrices in Section 5. In Section 6, we review some basic estimates for returning points and introduce the notion of a chain of symplectomorphisms associated to a flow, as well as that of a well-separated set. The notion of a well-separated set replaces that of a Poincaré section when a global section for the flow is not available. Although this requires some technical work, it does not substantially change the proof of the main result and the reader may wish to first assume that there is a global section and replace chains of symplectomorphisms by the standard Poincaré map for that section.

Section 7 defines sufficient assumptions on a family of metric perturbations to guarantee Theorem 1.3. Under these assumptions, we study the volume of perturbations that produce degenerate periodic points of a given length. Section 8 then proves an analog of Theorem 1.3 by implementing a delicate induction argument.

Finally, in Sections 9 and 10, we construct a family of metric perturbations which satisfies our technical assumptions and we prove Theorem 1.3 in Section 11.

Appendix A contains some elementary control estimates from ODE theory used to construct the perturbations of metrics in Section 9.

**Index of Notation.**

| | | | | | |
|---|---|---|---|---|---|
| $\mathscr{R}(t,\beta,g)$ | Def. 1.4 | $\mathcal{W}_{\rho_2,\rho_1}$ | (1.3) | $\mathscr{N}(t,\alpha(g,\mathcal{W}))$ | Def. 1.6 |
| probing map | Def. 2.2 | predominant | Def. 2.4 | $(\beta,q)$ non-degenerate | (4.1) |
| $\mathcal{M}_Y(V,s)$ | Def. 5.4 | $\mathscr{R}_F(n,\delta,g)$ | Def. 6.1 | $\mathscr{S}_F(n,\alpha,g)$ | Def. 6.2 |
| well-separated | Def. 6.5 | $\mathcal{P}_{\mathcal{I}}^{(n)}[g]$ | Def 6.8 | $\boldsymbol{T}_{\mathcal{I}}^{(n)}[g](\rho)$ | Def. 6.9 |
| $\mathscr{N}_F(n,\beta,g)$ | Def. 6.14 | $\mathscr{N}_{q,F}(n,\beta,g)$ | Def. 6.15 | $\mathfrak{L}$ | (7.1) |
| Good perturbation | Def. 7.2 | Admissible pairs | Def. 7.3 | $F_J^{\mathfrak{R}_\varepsilon,\delta_\varepsilon}$ | (8.8) |
| $F_\infty^{\mathfrak{R}_\varepsilon,\delta_\varepsilon}$ | (8.9) | $\mathfrak{b}$ | (8.18) | $\gamma_j$ | (8.19) |
| $\beta_{j,j}$ | (8.21) | $\alpha_j, \alpha_{j,\ell}, \beta_{i,j,\ell}, \tilde{\beta}_{j,\ell}, s_\ell$ | (8.22) | $\Phi_\rho^{g_\star}$ | (9.1) |
| $g_\sigma$ | (9.4) | $\Delta_\sigma, \tilde{\Delta}_\sigma$ | (9.9) | $\Psi_{\zeta_0}^{g_0}(\sigma), \Xi_i(\sigma,\zeta_0,g_0)$ | (9.10) |

## 2. Predominance in Banach Spaces

The main goals of this article are to give upper bounds on the number of closed geodesics of length $T$ and upper bounds for remainders in the Weyl law for a predominant set of metrics on a compact manifold $M$. Before proceeding to define our notion of predominance on Banach spaces in Section 2.2, we discuss several other available notions and motivate our choice of definition in Section 2.1. Finally, in Section 2.3 we explain why the notion of Baire genericity is not well-suited for our purposes.

### 2.1. **Existing notions of 'full measure'.**
The main difficulty in defining a concept analogous full measure in an infinite dimensional space, like the space of Riemannian metrics over a given manifold, is that there are no non-trivial, translation invariant, Borel measures. Several possible fixes for this problem have been introduced in the literature. We mention here the concepts of prevalence [HSY92] which uses an underlying linear structure and metric prevalence [Kal97] which does not require such a structure. These two notions have three important properties

(1) A prevalent set is dense.

(2) The intersection of prevalent sets is prevalent.                           (2.1)

(3) If $G \subset \mathbb{R}^n$ is prevalent, then $G$ has full Lebesgue measure.

Although quite flexible, as far as we are aware, the notion of metric prevalence has not proved useful in studying *quantitative* statements such as the growth of the number of periodic orbits of length $T$. Because of this, we focus on the notion of prevalence from [HSY92] which has appeared before in this type of application. In fact, for $\nu \geq 1$, [KH07] proves that there is a prevalent set $G \subset \mathcal{C}^\nu(I; I)$ of diffeomorphisms on the interval $I = [0, 1]$, such that for all $\varepsilon > 0$ and $f \in G$ there is $C > 0$ such that

$$\#\{x \in I : f^n(x) = x\} \leq Ce^{Cn^{1+\varepsilon}}.$$

Given a Banach space $\mathscr{G}$, a Borel set $G \subset \mathscr{G}$ is said to be *prevalent* if there is a Borel measure $\mu$ and a compact set $K \subset \mathscr{G}$ such that

$$0 < \mu(K) < \infty, \qquad G + g \text{ has full } \mu \text{ measure for all } g \in \mathscr{G}.$$

Usually, when one shows that a set $G$ is prevalent, it is convenient to construct a probe $\Sigma \subset \mathscr{G}$ which carries the measure $\mu$. In other words, if $G$ is prevalent, there are $\Sigma \subset \mathscr{G}$, a smooth map

$$F : \mathscr{G} \times \Sigma \to \mathscr{G}, \qquad F(g, \sigma) := g + \sigma \tag{2.2}$$

and a Borel measure, $\mu_\Sigma$, on $\Sigma$ such that for all $g \in \mathscr{G}$

$$\mu_\Sigma(\sigma \in \Sigma : F(g, \sigma) \in G^c) = 0. \tag{2.3}$$

Because we will be working in an open subset of a Banach space (the space metrics inside the space of symmetric 2-tensors) we would like a notion which does not rely on the fact that the space is linear. To do this, we generalize the type of functions allowed in (2.2) and slightly weaken (2.3). We now define the notion of predominance on an open subset of a Banach space $\mathscr{G}$.

2.2. **Predominant sets.** In this section we introduce the notion of predominance. Let $\mathscr{G}$ and $\mathscr{G}'$ be open subsets of the Banach spaces $(\mathscr{B}, \| \ \|_{\mathscr{B}})$ and $(\mathscr{B}', \| \ \|_{\mathscr{B}'})$ respectively,

$$\mathscr{G} \subset \mathscr{B}, \qquad \mathscr{G}' \subset \mathscr{B}', \tag{2.4}$$

and such that $\mathscr{G} \subset \mathscr{G}'$ and $\mathscr{B} \subset \mathscr{B}'$ via a continuous embedding, with $\mathscr{B}$ dense in $\mathscr{B}'$ and $\mathscr{G}$ dense in $\mathscr{G}'$. Let $\iota : \mathscr{G} \to \mathscr{G}'$ be the natural inclusion map.

**Remark 2.1.** When $\mathscr{O} \subset \mathscr{B}$ is open, we will say that a subset $K \subset \mathscr{O}$ is *bounded* if it is bounded as a subset of $\mathscr{B}$ and $d_{\mathscr{B}}(K, \partial\mathscr{O}) > 0$, where $d_{\mathscr{B}}$ is the distance induced by $\| \cdot \|_{\mathscr{B}}$.

The space $\mathscr{G}$ will be probed by perturbations indexed by a parameter $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon}$ for $\boldsymbol{\Sigma}_{N_\varepsilon}$ as follows. Given $\mathfrak{L} \in \mathbb{N}$ and $\{N_\varepsilon\}_{\varepsilon > 0} \subset \mathbb{N} \cup \{\infty\}$ let

$$\boldsymbol{\Sigma}_{N_\varepsilon} := \prod_{j=1}^{N_\varepsilon} B_{\mathbb{R}^{\mathfrak{L}}}(0, 1). \tag{2.5}$$

We endow this space with the sup-norm $\| \cdot \|_{\ell^\infty}$.

**Definition 2.2** (probing maps). We say that a collection $\mathscr{F} = \{(F_\varepsilon, N_\varepsilon)\}_\varepsilon$, with $\varepsilon \in (0, 1)$, is a family of $\mathscr{G}'$-*probing maps for* $\mathscr{G}$ if $N_\varepsilon \leq \dim \mathscr{B}$, there is a collection of closed bounded sets $\mathscr{G}_\varepsilon \subset \mathscr{G}$ with $\cup_\varepsilon \mathscr{G}_\varepsilon = \mathscr{G}$, $\mathscr{G}_{\varepsilon_2} \subset \mathscr{G}_{\varepsilon_1}$ for $\varepsilon_1 < \varepsilon_2$, and such that $F_\varepsilon : \mathscr{G}_\varepsilon \times \boldsymbol{\Sigma}_{N_\varepsilon} \to \mathscr{G}$ is a continuous map with

$$F_\varepsilon(g, 0) = g, \qquad g \in \mathscr{G}_\varepsilon, \tag{2.6}$$

and the following hold. For all $K \subset \mathscr{G}$ bounded,

$$\lim_{\varepsilon \to 0^+} \sup_{g \in K} \sup_{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon}} \|F_\varepsilon(g, \boldsymbol{\sigma}) - g\|_{\mathscr{B}} = 0, \tag{2.7}$$

$$\lim_{\varepsilon \to 0^+} \sup_{g \in K} \sup_{\substack{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \in \boldsymbol{\Sigma}_{N_\varepsilon} \\ \boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2}} \frac{\|F_\varepsilon(g, \boldsymbol{\sigma}_1) - F_\varepsilon(g, \boldsymbol{\sigma}_2)\|_{\mathscr{B}'}}{\|\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2\|_{\ell^\infty}} = 0. \tag{2.8}$$

In addition, the map $\tilde{F}_\varepsilon := \iota \circ F_\varepsilon : \mathscr{G}_\varepsilon \times \boldsymbol{\Sigma}_{N_\varepsilon} \to \mathscr{G}'$ is Lipschitz, and satisfies that for all $g \in \mathscr{G}_\varepsilon$ the Frechet derivative, $D_g \tilde{F}_\varepsilon$, of $\tilde{F}_\varepsilon$ in $g$ exists and for all $K \subset \mathscr{B}$ bounded

$$\lim_{\varepsilon \to 0^+} \sup_{g \in K, \boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon}} \|D_g \tilde{F}_\varepsilon|_{(g, \boldsymbol{\sigma})} - I\|_{\mathscr{B}' \to \mathscr{B}'} = 0, \tag{2.9}$$

**Remark 2.3.** We typically imagine that $N_\varepsilon$ is non-decreasing as $\varepsilon \to 0^+$ so that probing maps become more dispersed as $\varepsilon \to 0$. In fact, in our applications, $N_\varepsilon$ will often be identically equal to $\infty$. However, when $\mathscr{G}$ is finite dimensional it is reasonable to assume that $N_\varepsilon \leq \dim \mathscr{G}$ since, otherwise, one would be putting a very diffuse measure on a finite dimensional space.

In addition, the spaces $\mathscr{G}_\varepsilon$ allow us to define probes only in bounded subsets of $\mathscr{G}$ provided that, as $\varepsilon \to 0^+$, these subsets exhaust $\mathscr{G}$.

In our treatment, $\mathscr{G}$ and $\mathscr{G}'$ will be the spaces of $\mathcal{C}^\nu$ and $\mathcal{C}^{\nu-1}$ Riemannian metrics on a given manifold, while $\mathscr{B}$ and $\mathscr{B}'$ will be the spaces of $\mathcal{C}^\nu$ and $\mathcal{C}^{\nu-1}$ symmetric two-tensors. We need to refer to $\mathscr{G}'$ because our probing maps will typically *not* be Frechet differentiable as a map from
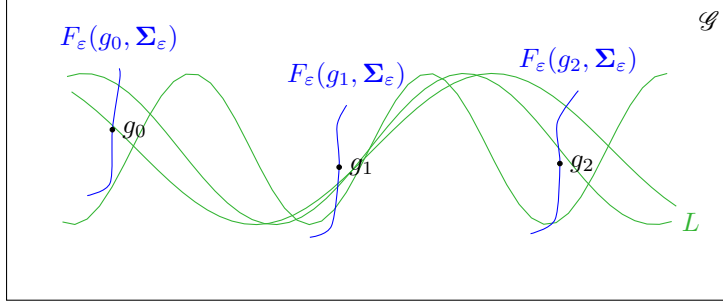
FIGURE 1. An example of a probing family $F_\varepsilon$ at a fixed $\varepsilon$ and a thin set $L$. One way to think of a probing map is that, to each point $g \in \mathscr{G}$, we attach a probe $F_\varepsilon(g, \mathbf{\Sigma}_{N_\varepsilon}^{\aleph})$. These probes are asymptotically translates of one another in the limit $\varepsilon \to 0$ and, moreover, are contained in a small ball around $g$. A set $L$ is thin if its intersection with each of these probes has vanishing measure in the limit $\varepsilon \to 0$.

$\mathcal{C}^\nu$ to $\mathcal{C}^\nu$ and instead, the Frechet derivative of the map will make sense as a map from $\mathcal{C}^\nu$ to $\mathcal{C}^{\nu-1}$ and will extend as in (2.9) to a map from $\mathcal{C}^{\nu-1}$ to $\mathcal{C}^{\nu-1}$.

We next discuss briefly the roles of each piece of the definition of a family of probing maps. The assumption (2.6) is crucial to know that $F_\varepsilon$ probes all of $\mathscr{G}_\varepsilon$ and does not avoid any open sets. To understand (2.9) and (2.8), recall [Mil97] that one can construct an example of a foliation of the unit square by *analytic* leaves $\{W_g\}_{g \in [0,1]}$, such that the map $g \mapsto W_g$ is continuous (but not differentiable), and there is a set $E \subset [0,1]^2$ with full measure such that $\#\{W_\sigma \cap E\} \le 1$. Thus, assumptions (2.9) and (2.8), which imply that the map $F_\varepsilon$ has reasonable regularity properties (both as a function of $g$ and $\varepsilon$), are crucial in proving that item (3) in list (2.1) holds for predominant sets.

In what follows we work with the measure $m_{\mathbf{\Sigma}_{N_\varepsilon}}$ on $\mathbf{\Sigma}_{N_\varepsilon}$ defined to be the product measure

$$m_{\mathbf{\Sigma}_{N_\varepsilon}} := \otimes_{j=1}^{N_\varepsilon} m, \qquad m := \frac{m_{\mathbb{R}^\mathfrak{L}}|_{B(0,1)}}{m_{\mathbb{R}^\mathfrak{L}}(B(0,1))},$$

where $m_{\mathbb{R}^\mathfrak{L}}$ denotes the Lebesgue measure on $\mathbb{R}^\mathfrak{L}$. Note that $m_{\mathbf{\Sigma}_\varepsilon}(\mathbf{\Sigma}_\varepsilon) = 1$.

**Definition 2.4** (predominant sets). Let $\mathscr{G}$ be an open subset of a Banach space $\mathscr{B}$, and $\mathscr{F} := \{(F_\varepsilon, N_\varepsilon)\}_{\varepsilon > 0}$ be a family of $\mathscr{G}'$ probing maps for $\mathscr{G}$. We say a set $L \subset \mathscr{G}$ is $\mathscr{F}$-thin if for all $K \subset \mathscr{G}$ bounded there is a Borel subset $L_0 \subset \mathscr{G}$ such that $L \subset L_0$ and for every $g \in K$, and $\varepsilon > 0$, there exists an $m_{\mathbf{\Sigma}_\varepsilon}$-measurable set $S_{g,\varepsilon} \subset \mathbf{\Sigma}_{N_\varepsilon}$ such that

$$\{\boldsymbol{\sigma} \in \mathbf{\Sigma}_{N_\varepsilon} : F_\varepsilon(g, \boldsymbol{\sigma}) \in L_0\} \subset S_{g,\varepsilon}, \qquad \limsup_{\substack{\varepsilon \to 0^+ \\ g \in K}} m_{\mathbf{\Sigma}_{N_\varepsilon}}(S_{g,\varepsilon}) = 0. \qquad (2.10)$$

We say $G \subset \mathscr{G}$ is $\mathscr{F}$-*predominant* if $\mathscr{G} \backslash G$ is $\mathscr{F}$-thin. We say $G \subset \mathscr{G}$ is respectively *predominant* or *thin* if there exists $\mathscr{G}'$ as above and a family of $\mathscr{G}'$ probing maps for $\mathscr{G}$ such that $G$ is $\mathscr{F}$-predominant or $\mathscr{F}$-thin respectively. Figure 1 shows a schematic of a family of probing maps and a thin set.

**Remark 2.5.** We note that if a set $G \subset \mathscr{G}$ is predominant the family of submanifolds endowed with Borel measures, $\{(\Gamma_{\varepsilon,g}, \mu_{\varepsilon,g})\}_{g \in \mathscr{G}}$, with $\Gamma_{\varepsilon,g} := F_\varepsilon(g, \boldsymbol{\Sigma}_{N_\varepsilon})$ and $\mu_{\varepsilon,g} := (F_\varepsilon(g, \cdot))_* m_{\boldsymbol{\Sigma}_\varepsilon}$ satisfies that $g \in \Gamma_{\varepsilon,g}$ (by (2.6)), $\mu_{\varepsilon,g}$ assigns a positive measure to any neighborhood of $g$ (by (2.7)), the map $g \mapsto \Gamma_{\varepsilon,g}$ is a $\mathcal{C}^1$ family of Lipschitz submanifolds (by (2.8), (2.9)), and $\mu_{\varepsilon,g}(G^c \cap \Gamma_{\varepsilon,g}) \le \varepsilon$ for every $g \in \mathscr{G}$ (by (2.10)). Indeed, by (2.9), we have that locally $\Gamma_{\varepsilon,g}$ is almost a translate of $\Gamma_{\varepsilon,g_0}$.

The direct analogy to (2.3) would replace the condition (2.10) with

$$\sup_{g \in K} m_{\boldsymbol{\Sigma}_{N_\varepsilon}}(S_{g,\varepsilon}) = 0.$$

We are, however, not able to show this in our applications and, instead, relax the condition to (2.10).

2.2.1. *Verification of properties* (2.1) *for predominance.* We now check that the notion of $\mathscr{F}$-predominance satisfies the properties listed in (2.1). We first prove that predominant sets are dense.

**Lemma 2.6.** *Suppose $\mathscr{F}$ is a family of $\mathscr{B}'$ probing maps for $\mathscr{G}$ and that $G \subset \mathscr{G}$ is $\mathscr{F}$-predominant. Then, $G$ is dense in $\mathscr{G}$.*

*Proof.* Let $G$ be $\mathscr{F}$-predominant. Fix $g_0 \in \mathscr{G}$ and $\delta > 0$. We will prove that $B_{\mathscr{G}}(g_0, \delta) \cap G \ne \emptyset$. Let $\mathscr{F} := \{(F_\varepsilon, N_\varepsilon)\}_{\varepsilon > 0}$ with $F_\varepsilon : \mathscr{G}_\varepsilon \times \boldsymbol{\Sigma}_{N_\varepsilon} \to \mathscr{G}$, $\cup_\varepsilon \mathscr{G}_\varepsilon = \mathscr{G}$, and $\mathscr{G}_{\varepsilon_2} \subset \mathscr{G}_{\varepsilon_1}$ for $\varepsilon_1 < \varepsilon_2$. By (2.7), there is $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$, $g_0 \in \mathscr{G}_\varepsilon$ and

$$\sup_{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon}} \|F_\varepsilon(g_0, \boldsymbol{\sigma}) - g_0\|_{\mathscr{B}} < \delta. \tag{2.11}$$

Let $G^c \subset L_0$ with $L_0$ Borel and satisfying (2.10). Then, since $m_{\boldsymbol{\Sigma}_\varepsilon}(\boldsymbol{\Sigma}_\varepsilon) = 1$, there is $\varepsilon_1 > 0$ such that $\{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon} : F_\varepsilon(g_0, \boldsymbol{\sigma}) \in L_0\} \ne \boldsymbol{\Sigma}_\varepsilon$ for $0 < \varepsilon < \min(\varepsilon_0, \varepsilon_1)$. In particular, there is $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\varepsilon$ such that $F_\varepsilon(g_0, \boldsymbol{\sigma}) \in \mathscr{G} \setminus L_0 \subset G$. By (2.11), this implies that $B_{\mathscr{G}}(g_0, \delta) \cap G \ne \emptyset$. $\qquad\square$

Next, we check that finite intersections of predominant sets are predominant.

**Lemma 2.7.** *Suppose $\mathscr{F}$ is a family of $\mathscr{B}'$ probing maps for $\mathscr{G}$ and that $G_j \subset \mathscr{G}$, $j = 1, 2, \ldots, J$ are $\mathscr{F}$-predominant. Then, $\bigcap_{j=1}^J G_j$ is $\mathscr{F}$-predominant.*

*Proof.* Let $K \subset \mathscr{G}$ bounded, let $\{L_{j,0}\}_{j=1}^J \subset \mathscr{G}$ be a collection of Borel sets with $G_j^c \subset L_{j,0}$, and for $\varepsilon > 0$ and $g \in K$ let $S_{j,g,\varepsilon} \subset \boldsymbol{\Sigma}_{N_\varepsilon}$ be $m_{\boldsymbol{\Sigma}_\varepsilon}$-measurable sets satisfying (2.10). Let $L_0 := \bigcup_{j=1}^J L_{j,0}$ and for each $\varepsilon > 0$ and $g \in K$ set $S_{g,\varepsilon} := \bigcup_{j=1}^J S_{j,g,\varepsilon}$. Then, for all $g \in K$ and $\varepsilon > 0$

$$\{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon} : F_\varepsilon(g, \boldsymbol{\sigma}) \in L_0\} = \bigcup_{j=1}^J \{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N_\varepsilon} : F_\varepsilon(g, \boldsymbol{\sigma}) \in L_{j,0}\} \subset S_{g,\varepsilon}.$$

Finally, $\lim_{\varepsilon \to 0^+} \sup_{g \in K} m_{\boldsymbol{\Sigma}_{N_\varepsilon}}(S_{g,\varepsilon}) \le \sum_{j=1}^J \lim_{\varepsilon \to 0^+} \sup_{g \in K} m_{\boldsymbol{\Sigma}_{N_\varepsilon}}(S_{j,g,\varepsilon}) = 0$, as claimed. $\qquad\square$

We end this section by checking that in finite dimensions predominant sets have full measure.

Since we are working in finite dimensions, we assume that $N_\varepsilon < \infty$ and $\mathscr{G} \subset \mathbb{R}^n$ for some $n < \infty$. We also take $\mathscr{B} = \mathbb{R}^n$. Now, the topology induced on $\mathbb{R}^n$ from $\mathscr{B}'$ for any Banach space $\mathscr{B}'$ is identical. Therefore, we may assume, without loss of generality, that $\mathscr{B}' = \mathscr{B} = \mathbb{R}^n$.

**Lemma 2.8.** *Let $n < \infty$, $N_0 < \infty$, and suppose $\mathscr{F} = \{(F_\varepsilon, N_\varepsilon)\}_{\varepsilon > 0}$ is a family of $\mathbb{R}^n$ probing maps for $\mathscr{G} := \mathbb{R}^n$ with $\sup_{\varepsilon > 0} N_\varepsilon < N_0$. If $G \subset \mathscr{G}$ is $\mathscr{F}$-predominant, then $G$ has full Lebesgue measure.*

*Proof.* We will show that if $L$ is $\mathscr{F}$-thin, then $L$ has zero Lebesgue measure. Let $K_0 \subset \mathscr{G} = \mathbb{R}^n$ be closed and bounded. Next, observe that by (2.7), for $\varepsilon > 0$ small enough,

$$\pi_{\mathscr{G}}(F_\varepsilon^{-1}(K_0)) \subset \{g \in \mathscr{G} \,:\, d(g, K_0) \le 1\} =: K.$$

Since $K \subset \mathscr{G}$ is bounded, we let $L_0 \subset \mathscr{G}$ be a Borel set with $L \subset L_0$ and $L_0$ satisfying (2.10) and we take $\varepsilon$ small enough such that $K \subset \mathscr{G}_\varepsilon$. Note that $F_\varepsilon^{-1}(L_0 \cap K_0)$ is measurable for all $\varepsilon > 0$. Fubini's theorem yields

$$\int_{\mathscr{G} \times \boldsymbol{\Sigma}_{N_\varepsilon}} 1_{F_\varepsilon^{-1}(L_0 \cap K_0)} d(m_{\mathbb{R}^n} \times m_{\boldsymbol{\Sigma}_{N_\varepsilon}}) = \int_K m_{\boldsymbol{\Sigma}_{N_\varepsilon}}\big(\{\boldsymbol{\sigma} : F_\varepsilon(g, \boldsymbol{\sigma}) \in L_0 \cap K_0\}\big) dm_{\mathbb{R}^n}(g). \qquad (2.12)$$

Next, by (2.9), $DF_\varepsilon|_{(g,\boldsymbol{\sigma})} : T_{(g,\boldsymbol{\sigma})}(\mathbb{R}^m \times \boldsymbol{\Sigma}_\varepsilon) \to T_g \mathbb{R}^m$ is surjective for $\varepsilon > 0$ small enough, $g \in K$, and $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\varepsilon$. Therefore, by the coarea formula,

$$\int_{\mathscr{G} \times \boldsymbol{\Sigma}_\varepsilon} 1_{F_\varepsilon^{-1}(L_0 \cap K_0)} d(m_{\mathbb{R}^n} \times m_{\boldsymbol{\Sigma}_{N_\varepsilon}}) = \int_{K_0} 1_{L_0 \cap F_\varepsilon(\mathscr{G} \times \boldsymbol{\Sigma}_{N_\varepsilon})} \Big(\int_{F_\varepsilon^{-1}(g)} \frac{1}{|JF_\varepsilon|} d\mathcal{H}_{g,\varepsilon}\Big) dm_{\mathbb{R}^n}(g), \quad (2.13)$$

where $\mathcal{H}_{g,\varepsilon}$ denotes the $\dim(\boldsymbol{\Sigma}_{N_\varepsilon})$-Hausdorff measure on $F_\varepsilon^{-1}(g)$ and $|JF_\varepsilon| := \sqrt{\det DF_\varepsilon(DF_\varepsilon)^*}$.

Next, we will prove that there is $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ and $g_0 \in K_0$

$$\pi_{\boldsymbol{\Sigma}_{N_\varepsilon}}\big(F_\varepsilon^{-1}(g_0)\big) = \boldsymbol{\Sigma}_{N_\varepsilon}. \qquad (2.14)$$

Once we have (2.14), observe that

$$\mathcal{H}_{g,\varepsilon}(F_\varepsilon^{-1}(g)) \ge \mathcal{H}_{g,\varepsilon}(\pi_{\boldsymbol{\Sigma}_{N_\varepsilon}}(F_\varepsilon^{-1}(g)) \ge c > 0.$$

Hence, since $|JF_\varepsilon| \le C < \infty$, we have for all $g$

$$\int_{F_\varepsilon^{-1}(g)} \frac{1}{|JF_\varepsilon|} d\mathcal{H}_{g,\varepsilon} > c > 0. \qquad (2.15)$$

To see (2.14), fix $(g_0, \boldsymbol{\sigma}) \in K_0 \times \boldsymbol{\Sigma}_{N_\varepsilon}$, and define $\Psi_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\Psi_\varepsilon(g_1) := g_1 - (D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})})^{-1}(F_\varepsilon(g_1 + g_0, \boldsymbol{\sigma}) - g_0).$$

We claim that there is $\varepsilon_0 > 0$ such that $\Psi_\varepsilon : \bar{B}(0,1) \to \bar{B}(0,1)$ is a contraction for $0 < \varepsilon < \varepsilon_0$, where $\bar{B}(0,1)$ is the closed unit ball. To prove the claim note that by (2.7), (2.8) and (2.9) there is $\varepsilon_0 > 0$ small enough such that for $0 < \varepsilon \le \varepsilon_0$, $g_0 \in K_0$, and $g_1 \in \bar{B}(0,1)$,

$$\|(D_g F|_{(g_0,\boldsymbol{\sigma})})^{-1}(F_\varepsilon(g_0, \boldsymbol{\sigma}) - g_0)\| < \tfrac{1}{2}, \qquad \|(D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})})^{-1}\big(D_g F_\varepsilon|_{(g_1 + g_0,\boldsymbol{\sigma})} - D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})}\big)\| < \tfrac{1}{4}. \qquad (2.16)$$

Now, for $g_1, g_2 \in \bar{B}(0,1)$,

$$\left\| (D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})})^{-1} \left[ F_\varepsilon(g_0 + g_1, \boldsymbol{\sigma}) - F_\varepsilon(g_0 + g_2, \boldsymbol{\sigma}) - D_g F_\varepsilon|_{(g_0+g_2,\boldsymbol{\sigma})}(g_1 - g_2) \right] \right\|$$

$$= \left\| \int_0^1 (D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})})^{-1} \left( D_g F_\varepsilon|_{(g_0+(1-t)g_2+tg_1,\boldsymbol{\sigma})} - D_g F_\varepsilon|_{(g_0+g_2,\boldsymbol{\sigma})} \right)(g_1 - g_2)dt \right\| \leq \tfrac{1}{2}\|g_1 - g_2\|. \quad (2.17)$$

Therefore, using (2.16) and letting $g_2 = 0$ in (2.17), we have that for $g_1 \in \bar{B}(0,1)$

$$\|\Psi_\varepsilon(g_1)\| = \|(D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})})^{-1} \left( D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})} g_1 - F_\varepsilon(g_0 + g_1, \boldsymbol{\sigma}) + F_\varepsilon(g_0, \boldsymbol{\sigma}) - F_\varepsilon(g_0, \boldsymbol{\sigma}) + g_0 \right)\| < 1,$$

and so $\Psi_\varepsilon : \bar{B}(0,1) \to \bar{B}(0,1)$ for $\varepsilon < \varepsilon_0$. Next, again by (2.16) and (2.17), for $g_1, g_2 \in \bar{B}(0,1)$

$$\|\Psi_\varepsilon(g_1) - \Psi_\varepsilon(g_2)\| = \|(D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})})^{-1} \left( D_g F_\varepsilon|_{(g_0,\boldsymbol{\sigma})}(g_1 - g_2) - F_\varepsilon(g_1 + g_0, \boldsymbol{\sigma}) + F_\varepsilon(g_2 + g_0, \boldsymbol{\sigma}) \right)\|$$

$$\leq \tfrac{3}{4}\|g_1 - g_2\|.$$

Hence, $\Psi_\varepsilon : \overline{B(0,1)} \to \overline{B(0,1)}$ is a contraction. For each $(g_0, \boldsymbol{\sigma}) \in \mathbb{R}^n \times \boldsymbol{\Sigma}_{N_\varepsilon}$ let $g_1 = g_1(g_0, \boldsymbol{\sigma})$ be the fixed point of $\Psi_\varepsilon$. Then, for each $(g_0, \boldsymbol{\sigma}) \in \mathbb{R}^n \times \boldsymbol{\Sigma}_{N_\varepsilon}$ there is $g_1$ with $F_\varepsilon(g_1 + g_0, \boldsymbol{\sigma}) = g_0$ as claimed in (2.14).

Finally, by (2.8) there is $C > 0$ with $\sup_{\varepsilon>0} \sup_{(g,\boldsymbol{\sigma})\in K \times \boldsymbol{\Sigma}_\varepsilon} |JF_\varepsilon(g,\boldsymbol{\sigma})| < C$, (2.12), (2.13), (2.10), (2.15) yield

$$0 = \lim_{\varepsilon\to 0^+} \int_{\mathscr{G}\times\boldsymbol{\Sigma}_{N_\varepsilon}} \mathbb{1}_{F_\varepsilon^{-1}(L_0\cap K_0)} d(m_{\mathbb{R}^n} \times m_{\boldsymbol{\Sigma}}) \geq \tfrac{1}{C} \int_{K_0} \mathbb{1}_{L_0\cap F_\varepsilon(\mathscr{G}\times\boldsymbol{\Sigma}_{N_\varepsilon})} dm_{\mathbb{R}^n} = \tfrac{1}{C} m_{\mathbb{R}^n}(L_0 \cap K_0).$$

Hence, $m_{\mathbb{R}^n}(L_0 \cap K_0) = 0$ and, since $K_0$ is an arbitrary closed bounded set, $m_{\mathbb{R}^n}(L_0) = 0$. In particular, by the completeness of the Lebesgue measure, $L$ is Lebesgue measurable with $m_{\mathbb{R}^n}(L) = 0$ as claimed.                                                                $\square$

## 2.3. Heuristic explanation for lack of Baire-genericity.

A set is Baire generic if it contains the intersection of countably many open dense sets. In order to explain why we do not pursue this notion of genericity, we discuss one of the key features we require at periodic points for the geodesic flow. Let $\Gamma \subset S^*M$ be a Poincaré section through the point $\rho$ and $\mathscr{P} : \Gamma \to \Gamma$ the corresponding Poincaré map. For simplicity, we will assume that $\mathscr{P}(\rho) = \rho$. Theorem 1.3 has consequences for the eigenvalues of $d\mathscr{P}|_\rho : T_\rho\Gamma \to T_\rho\Gamma$. Indeed, we have that there is $C > 0$ such that for all $n$

$$\|(d\mathscr{P}|_\rho)^n - I)^{-1}\| = \|(d(\mathscr{P}^n)|_\rho - I)^{-1}\| \leq (Cn)^{Cn^{\alpha\nu+\varepsilon}+1}.$$

Therefore, the eigenvalues, $\{\lambda_j\}_{j=1}^{2d-2}$ of $d\mathscr{P}|_\rho$ must satisfy

$$\inf_{j,p} \left( 1 + |\lambda_j| \right)^{n-1} |\lambda_j - e^{i\frac{2\pi p}{n}}| \geq \inf_j |\lambda_j^n - 1| \geq (Cn)^{-Cn^{\alpha\nu+\varepsilon}-1}, \qquad n \geq 1.$$

Since $d\mathscr{P}$ is a symplectic transformation, eigenvalues may be confined to the unit circle (see Section 5.1) and this becomes analogous to understanding the structure of the set real numbers which are poorly approximable by rational numbers.

To discuss the issues of genericity in a simpler setting, we forget now about Poincaré maps and instead discuss them in the context of approximation of real numbers. As explained above,

letting $h_\ell(q) = (\ell q)^{-\ell q^\alpha - 1}$ we want to investigate the set

$$\mathfrak{D} := \bigcup_{\ell > 0} \mathfrak{D}_\ell, \qquad \mathfrak{D}_\ell := \{s \in [0,1] : |s - \tfrac{p}{q}| > h_\ell(q), \text{ for all } p, q \in \mathbb{N}\}. \qquad (2.18)$$

To study $\mathfrak{D}$ we consider

$$\mathfrak{U} := \bigcap_{q_0=1}^{\infty} \bigcup_{q=q_0}^{\infty} \mathfrak{U}_q, \qquad \mathfrak{U}_q := \bigcup_{p=0}^{q-1} (\tfrac{p}{q} - h_q(q), \tfrac{p}{q} + h_q(q)).$$

Observe that $\mathfrak{U}$ is Baire generic. However, $\mathfrak{U} \cap \mathfrak{D}_\ell = \emptyset$ for all $\ell > 0$ which implies that $\mathfrak{D} \cap \mathfrak{U} = \emptyset$ and hence that $\mathfrak{D}$ is *not* Baire generic. Since even the property we want for the eigenvalues of $d\mathscr{P}$ is non-generic, it is unlikely that the set of metrics which produce such $\mathscr{P}$ is generic.

For further evidence of lack of Baire genericity, we consider the space $\mathrm{Diff}^\nu(\Gamma)$ of $\mathcal{C}^\nu$ diffeomorphisms on a smooth manifold $\Gamma$ for $\nu \geq 2$. Kaloshin [Kal00] showed that for any $\{a_n\}_{n=1}^\infty \subset [1, \infty)$, the set

$$\left\{ f \in \mathrm{Diff}^\nu(\Gamma) : \limsup_n \frac{P_n(f)}{a_n} < \infty \right\}$$

is *not* Baire generic, where $P_n(f) := \#\{\rho \in \Gamma \text{ isolated} : f^n(\rho) = \rho\}$. Indeed, when $\Gamma$ is a 3-manifold, this set is *not* Baire generic even in the space of volume preserving maps [KS06].

While $\mathfrak{D}$ in (2.18) is not Baire generic, one can see that, since $\lim_{n \to \infty} |[0,1] \setminus \mathfrak{D}_n| = 0$, $\mathfrak{D}$ has full Lebesgue measure and hence is, in a much stronger sense, typical. Indeed, for many purposes, the notion of full Lebesgue measure is a better version of 'typical' than Baire genericity. For example, a randomly chosen element of $[0,1]$ is almost surely in $\mathfrak{D}$ but may not be in a given Baire generic set. Also, a full Lebesgue measure set in $\mathbb{R}$ has Hausdorff dimension 1, while a Baire generic set in $\mathbb{R}$ may have 0-Hausdorff dimension. Motivated by this discussion, the notion of 'typicality' that we use, i.e. that of predominance (see Definition 2.4), is an analog of a full Lebesgue measure set of metrics $g \in \mathscr{G}^\nu$.

## 3. COUNTING CLOSED GEODESICS AND IMPROVEMENTS FOR WEYL LAWS: PROOF OF THEOREMS 1.1 AND 1.2

In this Section, we use Theorem 1.3 to prove Theorems 1.1 and 1.2. Both of these theorems rely on volume estimates on the set of nearly closed geodesics. We obtain these estimates in Section 3.1 and prove Theorems 1.1 and 1.2 in Sections 3.2 and 3.3 respectively.

We start by letting $M$ be a $\mathcal{C}^\nu$ manifold, $g \in \mathscr{G}^\nu$, $\beta_0 > 0$, and $\mathcal{W}$ be a $\beta_0$-family of transition maps for $g$. Let $\boldsymbol{\beta} : [0, +\infty) \to (0, +\infty)$ be a continuous, decreasing function. Throughout the section we will suppose that

$$\mathscr{R}(t, \boldsymbol{\beta}(t), g) \subset \mathscr{N}(t, \boldsymbol{\beta}(t), (g, \mathcal{W})), \quad \text{for } t > c, \qquad d(\varphi_t^g(\rho), \rho) \geq c|t|, \quad \text{for } |t| \leq c. \qquad (3.1)$$

In Theorem 1.3, we show that there is a predominant set of metrics $G \subset \mathscr{G}^\nu$ such that (3.1) holds for $g \in G$ with $\boldsymbol{\beta}(t) = C^{-C(t+1)^\Omega - 1}$ and some $C > 0$.

3.1. **Volume of nearly closed geodesics.** We start by following [DZ16, Appendix A], replacing their assumption of hyperbolicity of the flow with (3.1) to estimate the volume of nearly closed geodesics. Recall that there are $\tilde{C}, L > 0$ such that for all $f \in \mathcal{C}^2(S^*M)$ and $t \in \mathbb{R}$

$$\|f \circ \varphi_t^g\|_{\mathcal{C}^2(S^*M)} \leq \tilde{C} e^{L|t|} \|f\|_{\mathcal{C}^2(S^*M)}. \tag{3.2}$$

In particular, $d(\varphi_t^g(\rho), \varphi_t^g(\rho')) \leq \tilde{C} e^{L|t|} d(\rho, \rho')$ for all $\rho, \rho' \in S^*M$ and $t \in \mathbb{R}$.

The following lemma shows that two nearby orbits that return to their starting point after similar times are almost iterates of one another.

**Lemma 3.1.** *Let* $M, g, \beta_0, \mathcal{W}$ *and* $\boldsymbol{\beta}$ *be such that* (3.1) *and* (1.3) *hold. Given* $t_0 > 0$ *there are* $C, \delta > 0$ *such that the following holds. Let* $t \geq t_0$, $t' \geq t_0$ *with* $|t - t'| \leq \delta$, *and for* $0 \leq \varepsilon < \boldsymbol{\beta}(t)$ *let* $\rho, \rho' \in S^*M$ *be such that*

$$d(\rho, \varphi_t^g(\rho)) \leq \varepsilon, \qquad d(\rho', \varphi_{t'}^g(\rho')) \leq \varepsilon, \qquad d(\rho, \rho') \leq \delta \boldsymbol{\beta}(t) e^{-Lt}.$$

*Then,* $|t - t'| \leq C\varepsilon$ *and there exists* $s \in [-1, 1]$ *such that* $\boldsymbol{\beta}(t) d(\rho, \varphi_s^g(\rho')) \leq C\varepsilon$.

*Proof.* First, observe that for any $\varepsilon_0 > 0$, we may increase $C$ enough so that the statement becomes trivial for $\varepsilon \geq \varepsilon_0$. Therefore, we need only work with $\varepsilon < \varepsilon_0 < \min(\frac{1}{4}\beta_0, \frac{1}{2})$ small enough and we do so from now on. Next, observe that we may shrink $\delta$ so that $d(\varphi_t^g(\rho), \varphi_t^g(\rho')) \leq \frac{1}{4}\beta_0$ whenever $d(\rho, \rho') \leq \delta e^{-Lt}$, $t \geq t_0$. We may also assume that all of the relevant points are contained in a single coordinate chart and hence we may work in a small ball in $\mathbb{R}^{2d-1}$.

We also assume that $\delta$ is small enough such that whenever $d(\rho, \rho') \leq \delta$ there exists $|s| \leq 1$ with $\rho - \varphi_s^g(\rho') \in \mathscr{H}(\rho)$ and such that $d(\rho, \varphi_s^g(\rho')) \leq 2d(\rho, \rho')$. Set $\rho_0 := \varphi_s^g(\rho')$ and note

$$\rho - \rho_0 \in (H_p(\rho))^{\perp}, \qquad d(\rho, \rho_0) \leq 2\delta \boldsymbol{\beta}(t) e^{-Lt}, \qquad d(\varphi_{t'}^g(\rho_0), \rho_0) \leq \tilde{C} e^L \varepsilon. \tag{3.3}$$

By Taylor expanding in $\rho$ and using (3.2), we have

$$\|\varphi_t^g(\rho_0) - \varphi_t^g(\rho) - d\varphi_t^g(\rho)(\rho_0 - \rho)\| \leq \tilde{C} e^{Lt} d(\rho, \rho_0)^2 \leq 2\tilde{C}\delta \boldsymbol{\beta}(t) d(\rho, \rho_0).$$

Next, Taylor expanding in $t$ and using that there is $C > 0$ such that $\|\partial_t^2 \varphi_t^g\| \leq C$,

$$\|\varphi_{t'}^g(\rho_0) - \varphi_t^g(\rho_0) - H_p(\varphi_t^g(\rho_0))(t' - t)\| \leq C|t' - t|^2 \leq C\delta|t' - t|.$$

Thus, we have that there is $C > 0$ with

$$\|\varphi_{t'}^g(\rho_0) - \varphi_t^g(\rho) - d\varphi_t^g(\rho)(\rho_0 - \rho) - H_p(\varphi_t^g(\rho_0))(t' - t)\| \leq C\delta\Big(\boldsymbol{\beta}(t) d(\rho_0, \rho) + |t - t'|\Big).$$

Next, since $d(\varphi_t(\rho), \rho) \leq \varepsilon$, by (3.3) there is $C > 0$ such that

$$\|[I - d\varphi_t^g(\rho)](\rho_0 - \rho) - H_p(\varphi_t^g(\rho_0))(t' - t)\| \leq C\delta\Big(\boldsymbol{\beta}(t) d(\rho_0, \rho) + |t - t'|\Big) + C\varepsilon.$$

Now, by (1.3) there is $C > 0$ such that $\|\mathcal{W}_{\varphi_t^g(\rho), \rho} - I\| \leq Cd(\varphi_t^g(\rho), \rho) \leq C\varepsilon$. Therefore, there is $C > 0$ such that

$$\|[\mathcal{W}_{\varphi_t^g(\rho), \rho} - d\varphi_t^g(\rho)](\rho_0 - \rho) - H_p(\varphi_t^g(\rho))(t' - t)\| \leq C\delta\Big(\boldsymbol{\beta}(t) d(\rho_0, \rho) + |t - t'|\Big) + C\varepsilon. \tag{3.4}$$

Now, since $\rho_0 - \rho \in \mathscr{H}(\rho)$, we have by (3.1), Definition 1.6, and the fact that that $\varepsilon \le \boldsymbol{\beta}(t)$, that there is $C > 0$ such that

$$\boldsymbol{\beta}(t) d(\rho_0, \rho) \le \|\Pi_{\varphi_t(\rho)} \big[ \mathcal{W}_{\rho, \varphi_t^g(\rho)} - d\varphi_t^g(\rho) \big] (\rho_0 - \rho) \|$$
$$\le C \| \big[ \mathcal{W}_{\rho, \varphi_t^g(\rho)} - d\varphi_t^g(\rho) \big] (\rho_0 - \rho) - H_p(\varphi_t^g(\rho))(t' - t) \|. \tag{3.5}$$

On the other hand, since $\mathcal{W}_{\varphi_t^g(\rho), \rho}(H_p(\rho)) = H_p(\varphi_t^g(\rho))$, $\mathcal{W}_{\varphi_t(\rho), \rho}(\mathscr{H}(\rho)) = \mathscr{H}(\varphi_t^g(\rho))$, and $\mathscr{H}(\rho)$ is transverse to $H_p(\rho)$ (uniformly in $\rho$), we have

$$|t - t'| \le C \| \big[ \mathcal{W}_{\rho, \varphi_t^g(\rho)} - d\varphi_t^g(\rho) \big] (\rho_0 - \rho) - H_p(\varphi_t^g(\rho))(t' - t) \|. \tag{3.6}$$

Combining (3.4), (3.5), (3.6), and choosing $\delta \ll 1$ and $\varepsilon_0 < \frac{1}{2C}$, we have $\boldsymbol{\beta}(t) d(\rho, \rho_0) + |t - t'| \le C\varepsilon$ as claimed. $\qquad \square$

We proceed to bound the volume of nearly closed trajectories in $S^*M$.

**Lemma 3.2.** *Let $M, g, \beta_0, \mathcal{W},$ and $\boldsymbol{\beta}$ be such that (3.1) holds. Let $t_0 > 0$, then there is $C > 0$ such that for $T > t_0$ and $0 < \varepsilon \le \boldsymbol{\beta}(T)$, we have*

$$\mathrm{vol}_{S^*M} \Big( \big\{ \rho \, : \, \exists t \in [t_0, T] \text{ such that } d(\varphi_t^g(\rho), \rho) \le \varepsilon \big\} \Big) \le C\varepsilon^{2d-2} e^{CT} \boldsymbol{\beta}(T)^{-(4d-3)}.$$

*Proof.* Let $t_0 > 0$ and $T > t_0$. Then, let $\delta$ as in Lemma 3.1 and fix $0 < \varepsilon \le \boldsymbol{\beta}(T)$. Next divide $[t_0, T]$ into intervals, $\{I_i\}_{i=1}^{N_T}$ with right endpoints at $\{T_i\}_{i=1}^{N_T}$ satisfying $|I_i| \le \frac{\delta}{2}$. Next, for each $i = 1, \dots, N_T$ let $\{\rho_j\}_{j=1}^{K_{T_i}} \subset S^*M$ be a maximal $\frac{\delta}{2} e^{-LT_i} \boldsymbol{\beta}(T_i)$ separated set. Note that

$$\big\{ \rho \, : \, \exists t \in [t_0, T] \text{ such that } d(\varphi_t^g(\rho), \rho) \le \varepsilon \big\} \subset \bigcup_{i=1}^{N_T} \bigcup_{j=1}^{K_{T_i}} \pi_{S^*M}(\mathbf{P}_{i,\rho_j}),$$

$$\mathbf{P}_{i,\rho_j} := \Big\{ (\rho, t) \, : \, t \in I_i, \, d(\varphi_t^g(\rho), \rho) \le \varepsilon, \, d(\rho, \rho_j) \le \frac{\delta}{2} e^{-LT_i} \boldsymbol{\beta}(T_i) \Big\}. \tag{3.7}$$

To estimate $\mathrm{vol}_{S^*M}(\pi_{S^*M}(\mathbf{P}_{i,\rho_j}))$, fix $(\rho', t') \in \mathbf{P}_{i,\rho_j}$ and let $(\rho, t) \in \mathbf{P}_{i,\rho_j}$ with $(\rho, t) \ne (\rho', t')$. By Lemma 3.1,

$$|t - t'| \le C\varepsilon, \qquad d\Big(\rho \, , \, \bigcup_{|s| \le 1} \varphi_s^g(\rho')\Big) \le C\varepsilon \boldsymbol{\beta}(T_i)^{-1}.$$

Then, $\mathrm{vol}_{S^*M}(\pi_{S^*M}(\mathbf{P}_{i,\rho_j})) \le C\varepsilon^{2d-2} \boldsymbol{\beta}(T_i)^{-(2d-2)}$, and

$$\mathrm{vol}_{S^*M} \Big( \pi_{S^*M} \Big( \bigcup_{j=1}^{K_{T_i}} \pi_{S^*M}(\mathbf{P}_{i,\rho_j}) \Big) \Big) \le K_{T_i} C\varepsilon^{2d-2} \boldsymbol{\beta}(T_i)^{-(2d-2)} \le C\varepsilon^{2d-2} e^{(2d-1)LT_i} \boldsymbol{\beta}(T_i)^{-(4d-3)}, \tag{3.8}$$

since there is $C_\delta > 0$ such that $K_{T_i} \le C_\delta [e^{-LT_i} \boldsymbol{\beta}(T_i)]^{-(2d-1)}$. The claim follows from combining (3.8) with (3.7), and using that $\boldsymbol{\beta}$ is decreasing and $N_T \le CT$. $\qquad \square$

**Remark 3.3.** We note that if one is only interested in obtaining a volume estimate of the form

$$\mathrm{vol}_{S^*M} \Big( \big\{ \rho \, : \, \exists t \in [t_0, T] \text{ such that } d(\varphi_t^g(\rho), \rho) \le \varepsilon \big\} \Big) \le C\varepsilon^{2d-2-k} e^{CT} \boldsymbol{\beta}(T)^{-(4d-3-k)},$$

for some $k \geq 1$, then it is possible to use a weaker notion of non-degeneracy. Indeed, one can replace the definition of $\mathcal{N}(t, \rho, (g, \mathcal{W}))$ by saying that $\rho \in \mathcal{N}_k(t, \rho, (g, \mathcal{W}))$ if there is a $2d - 2 - k$ dimensional subspace of $v \in T_\rho S^*M$ such that $|\Pi_{\varphi_t^g(\rho)}(\mathcal{W}_{\rho, \varphi_t^g(\rho)} - d\varphi_t^g)v| > \alpha |\Pi_\rho v|$. If implemented throughout the article, this would in turn lead to a smaller value of $\alpha_\nu$ in Theorem 1.1 but would not be sufficient for counting closed geodesics.

3.2. **Proof of Theorem 1.2.** Let $\Omega > \Omega_\nu$. By Theorem 1.3 there is a predominant set $G_\Omega \subset \mathscr{G}^\nu$ such that for all $g \in G$ and $\mathcal{W}$ a family of transition maps for $g$ there is $C > 0$ such that (3.1) holds with $\boldsymbol{\beta}(t) = C^{-C(t+1)^\Omega - 1}$.

Let $g \in G_\Omega$ and $0 < t_0 < \text{inj}_g(M)$ so that there are no closed geodesics with length $\leq t_0$. Let $\delta$ be as in Lemma 3.1. First, we claim that $\mathfrak{c}(T, g) < \infty$ for each $T > 0$. Indeed, suppose that $\{\rho_i\}_{i=1}^\infty \subset S^*M$ and $\{t_i\}_{i=1}^\infty \subset [t_0, T]$ are sequences satisfying $\varphi_{t_i}^g(\rho_i) = \rho_i$ and such that $\rho_j \notin \{\varphi_s^g(\rho_i) : |s| \leq 4\}$ for all $i \neq j$. Then, we may assume that there are $\rho \in S^*M$ and $t \in [t_0, T]$ such that $\rho_i \to \rho$ and $t_i \to t$. By continuity, we conclude $\varphi_t^g(\rho) = \rho$. Next, let $i_0 > 0$ be such that $d(\rho_i, \rho) \leq \delta \boldsymbol{\beta}(T) e^{-LT}$ and $|t_i - t| \leq \delta$ for $i \geq i_0$. Then, by Lemma 3.1 we obtain that for $i > i_0$ there is $s_i \in [-1, 1]$ such that $\varphi_{s_i}^g(\rho_i) = \rho$ and $t_i = t$. In particular, for $i, j > i_0$, we have $\rho_j = \varphi_{s_i - s_j}^g(\rho_i)$ which is a contradiction since $|s_i - s_j| \leq 3$.

Since $\mathfrak{c}(T, g) < \infty$, we let $\{\gamma_i\}_{i=1}^{\mathfrak{c}(T, g)}$ be the finite collection of primitive closed geodesics of length $\leq T$. Then, there is $\delta_0 = \delta_0(T) > 0$ such that if $\gamma_i(\delta) \subset S^*M$ denotes the $\delta$ neighborhood of $\gamma_i$,

$$\gamma_i(\delta) \cap \gamma_j(\delta) = \emptyset, \qquad i \neq j, \qquad 0 \leq \delta < \delta_0. \tag{3.9}$$

Letting $T_i$ be the length of $\gamma_i$, we have by (3.2) that $\sup_{\rho \in \gamma_i(\delta)} d(\rho, \varphi_{T_i}^g(\rho)) \leq \tilde{C} \delta e^{LT_i}$. Therefore, using Lemma 3.2 with $\tilde{C} \delta e^{LT}$ in place of $\varepsilon$, there is $C > 0$ such that,

$$\sum_{1 \leq i \leq \mathfrak{c}(T, g)} \text{vol}_{S^*M}(\gamma_i(\delta)) = \text{vol}_{S^*M}\left(\bigcup_{1 \leq i \leq \mathfrak{c}(T, g)} \gamma_i(\delta)\right) \leq C \delta^{2d-2} e^{CT + (2d-2)LT} \boldsymbol{\beta}(T)^{-(4d-3)}. \tag{3.10}$$

In addition, since there is $c > 0$ such that $\text{vol}_{S^*M}(\gamma_i(\delta)) \geq c \delta^{2d-2} t_0$ for all $1 \leq i \leq \mathfrak{c}(T, g)$, we obtain, using (3.9) and (3.10) that there is $C > 0$ such that

$$\mathfrak{c}(T, g) t_0 \leq C e^{CT + (2d-2)LT} \boldsymbol{\beta}(T)^{-(4d-3)}.$$

Together with the fact that $\mathfrak{c}(T, g) = 0$ for $0 < T < t_0$, this implies Theorem 1.2. □

3.3. **Proof of Theorem 1.1.** Let $g$ be a $\mathcal{C}^\nu$-Riemannian metric on a manifold $M$ of dimension $d$. For $R > 0$ let $\mathcal{P}^R(t_0, T) \subset S^*M$ be the set of directions that yield trajectories that are $R$ close to being periodic at some time with $t_0 \leq |t| \leq T$. That is,

$$\mathcal{P}^R(t_0, T) := \left\{\rho \in S^*M : \bigcup_{t_0 \leq |t| \leq T} \varphi_t^g(B_{S^*M}(\rho, R)) \cap B_{S^*M}(\rho, R) \neq \emptyset\right\}. \tag{3.11}$$

In [CG20, Theorem 2] it is proved that if $\nu \geq \nu_0$ and $\mathbf{T}(R)$ is a sub-logarithmic resolution function and there exist $C > 0$ and $t_0$ such that

$$\limsup_{R \to 0^+} \text{vol}_{S^*M}\left(B(\mathcal{P}^R(t_0, \mathbf{T}(R)), R)\right) \mathbf{T}(R) \leq C, \tag{3.12}$$

then, as $\lambda \to \infty$,

$$\#\{j : \lambda_j(g) \le \lambda\} = \frac{\mathrm{vol}_{\mathbb{R}^d}(B^d)\,\mathrm{vol}_g(M)}{(2\pi)^d}\lambda^d + O\Big(\frac{\lambda^{d-1}}{\mathbf{T}(\lambda^{-1})}\Big). \tag{3.13}$$

**Remark 3.4.** The application of [CG20, Theorem 2] deserves a brief comment. Note that, as stated, the theorem is valid for $\mathcal{C}^\infty$ metrics. However, it is clear that all of the proofs rely only on the $\mathcal{C}^\nu$ norm of the metric for some sufficiently large $\nu$. Therefore, provided that we work with sufficiently smooth metrics, we may apply [CG20, Theorem 2].

The proof of Theorem 1.1, given in subsection 3.4, is then reduced to proving that for any $\Omega > \Omega_\nu$ and some $C_0 > 0$

$$\mathbf{T}(R) = f^{-1}(R^\gamma), \tag{3.14}$$

with

$$\gamma := \frac{2d-2}{4d-3}, \qquad f(t) = \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(0)}e^{-C_0 t}, \qquad \boldsymbol{\beta}(t) := C^{-C(t+1)^\Omega - 1}$$

is a sub-logarithmic resolution function and that (3.12) holds. Indeed, Theorem 1.1 follows since

$$C_0(\log\lambda)^{\frac{1}{\Omega}} \ge \mathbf{T}(\lambda^{-1}) = f^{-1}(\lambda^{-\gamma}) \ge \frac{1}{C_0}(\log\lambda)^{\frac{1}{\Omega}}, \qquad \lambda \ge \lambda_0 \tag{3.15}$$

for some $C_0 > 0$ and $\lambda_0 > 0$ large enough.

Before we proceed to the proof of Theorem 1.1, we recall the notion of a sub-logarithmic resolution function from [CG20, Definition 1.1]. We say that $\mathbf{T} : (0,1) \to (0,\infty)$ is a *resolution function* if it is continuous and decreasing. We say that $\mathbf{T}$ is *sub-logarithmic* if it is differentiable and

$$(\log\mathbf{T}(R))' \ge -\frac{1}{R\log R^{-1}} \qquad 0 < R < 1. \tag{3.16}$$

Next, we introduce a convenient class of sub-logarithmic rate functions.

**Lemma 3.5.** *Suppose that $A(t) : \mathbb{R} \to \mathbb{R}$ is twice differentiable and satisfies $A(0) = 0$, $A'(t) > 0$, $A''(t) \ge 0$ for $t \ge 0$. Then, with $f(t) = e^{-A(t)}$, we have that $\mathbf{T}(R) := f^{-1}(R^\gamma)$ is a sub-logarithmic rate function for $\gamma > 0$.*

*Proof.* First, observe that $\mathbf{T}(R)$ is decreasing and differentiable. Moreover,

$$(\log\mathbf{T})'(R) = \frac{\gamma R^\gamma}{Rf'(f^{-1}(R^\gamma))f^{-1}(R^\gamma)} = -\frac{\gamma}{RA'(f^{-1}(R^\gamma))f^{-1}(R^\gamma)}.$$

Next, note that

$$A'(t)t = \int_0^t A''(s)s + A'(s)ds \ge \int_0^t A'(s)ds = A(t).$$

Therefore, $A'(f^{-1}R^\gamma)f^{-1}(R^\gamma) \ge A(f^{-1}(R^\gamma)) = \gamma\log R^{-1}$, and hence (3.16) holds as claimed. $\square$

**Remark 3.6.** By Lemma 3.5, the function $\mathbf{T}(R)$ as defined in (3.14) is a sub-logarithmic resolution function. Indeed, $f(t) = e^{-A(t)}$ with $A(t) = C_0 t + C((t+1)^\Omega - 1)$. Note that $A(0) = 0$, and, since $\Omega \ge 1$, it is easy to check that $A'(t) > 0$, and $A''(t) \ge 0$ for $t \ge 0$ as claimed.

3.4. **Proof of Theorem 1.1.** Let $\Omega > \Omega_\nu$. By Theorem 1.3 there is a predominant set $G_\Omega \subset \mathscr{G}^\nu$ such that for all $g \in G_\Omega$ and $\mathcal{W}$ a family of transition maps for $g$ there is $C > 0$ such that (3.1) holds with $\boldsymbol{\beta}(t) = C^{-C(t+1)^\Omega-1}$. Without loss of generality we assume $C > 1$. In Lemma 3.5 we proved that $\mathbf{T}$ as defined in (3.14) is a sub-logarithmic resolution function. Therefore, as explained above, the proof of Theorem 1.1 would follow from combining (3.13) and (3.15) once we prove (3.12).

Let $0 < t_0 < T$. In order to prove (3.12), we first claim that there is $C > 0$ such that

$$B(\mathcal{P}^R(t_0, T), R) \subset \left\{ q \in S^*M : \exists t \in (t_0, T) \cup (-T, -t_0) \text{ s.t. } d(\varphi_t(q), q) \leq Ce^{CT}R \right\}. \quad (3.17)$$

Indeed, for $q \in B(\mathcal{P}^R(t_0, T), R)$, there is $\rho \in \mathcal{P}^R(t_0, T)$ such that $d(\rho, q) < R$. Then, there are $t \in (t_0, T) \cup (-T, -t_0)$ and $\rho_1 \in S^*M$ such that $d(\rho, \rho_1) < R$ and $d(\varphi_t(\rho_1), \rho) < R$. Since $\varphi_t((x, \xi)) = \varphi_{-t}(x, -\xi)$, we may assume without loss of generality that $t > 0$. Therefore, by (3.2) there is $C > 0$ such that

$$d(\varphi_t(q), q) \leq d(\varphi_t(q), \varphi_t(\rho_1)) + d(\varphi_t(\rho_1), \rho) + d(\rho, q) \leq \tilde{C}e^{Lt}d(q, \rho_1) + R + R \leq Ce^{Ct}R,$$

proving the claim in (3.17).

By (3.17) and Lemma 3.2, for $R \leq \boldsymbol{\beta}(T)(Ce^{CT})^{-1}$, we have

$$\mathrm{vol}_{S^*M}\left( B(\mathcal{P}^R(t_0, T), R) \right)T \leq CT(Ce^{CT}R)^{2d-2}e^{CT}\boldsymbol{\beta}(T)^{-(4d-3)} \leq Ce^{C_1T}R^{2d-2}\beta(T)^{-(4d-3)}.$$

Therefore, we may choose $C_0 = \frac{C_1}{4d-3}$ and use the definition of $\mathbf{T}(R)$ to obtain that $R \leq \beta(\mathbf{T}(R))(Ce^{C\mathbf{T}(R)})^{-1}$ for $R$ small enough and $\mathrm{vol}_{S^*M}\left( B(\mathcal{P}^R(t_0, \mathbf{T}(R)), R) \right)\mathbf{T}(R) \leq C$ as claimed. $\qquad \square$

## 4. Outline of the proof of Theorem 1.3

To simplify the exposition in this outline, we will first imagine that it is only necessary to understand exactly periodic points rather than returning points. We also assume that one can construct a global Poincaré section (see Section 4.5). In addition, rather than providing the details to prove the result for a predominant set of metrics, we will outline a proof of the theorem for a dense set of metrics. See the end of the section for remarks on how to drop these assumptions and prove predominance.

4.1. **The perturbation.** The first key point to understand when proving Theorem 1.3 is how to perturb away the degeneracy of a single periodic orbit with quantitative control on how large a perturbation is needed to produce non-degeneracy. When doing this, it is important to use a perturbation of the metric which interacts with the periodic geodesic exactly once. That is, the perturbation must be supported in a ball, $B \subset M$, over which the geodesic passes exactly once. Because of this, we will only be able to directly perturb away degeneracy for primitive closed geodesics. The construction of one such the perturbation is done in Section 9.

**Remark 4.1.** Our main inductive argument is actually proved under some general assumptions on the perturbation (see Definitions 7.2 and 7.3). Because of this, we postpone this construction until after our main inductive argument, which appears in Section 8.

Restricting our attention to a single geodesic, $\gamma_{g_0}$, for a metric $g_0$, we find a finite dimensional family of perturbations $g_\sigma$ of the metric $g_0$. When $\gamma_{g_0}$ is a primitive periodic geodesic, these perturbations should produce non-degeneracy of the orbit $\gamma_{g_\sigma}$. The perturbations are a combination of those considered by Anosov [Ano82] and Klingenberg [Kli78, §3.3] (see also Klingenberg–Takens [KT72]). In our case, when $\gamma_{g_0}$ is primitive, we give careful estimates on how these perturbations affect the Poincaré map, $\mathscr{P}_{\gamma_{g_\sigma}}$, associated to $\gamma_{g_\sigma}$ and its derivative, $d\mathscr{P}_{\gamma_{g_\sigma}}$. In particular, estimating the size of the inverse of the derivative of the map $\sigma \mapsto (\mathscr{P}_{\gamma_{g_\sigma}}, d\mathscr{P}_{\gamma_{g_\sigma}})$.

Once we have a family of perturbations for any given geodesic, we cover $S^*M$ by finitely many small balls, $B_i$, of radius $r$ and center $\rho_i$ and associate to each ball a family of perturbations modelled on those above; replacing the geodesic $\gamma_{g_0}$ by the geodesic, $\gamma_{g_0, \rho_i}$ for $g_0$ through $\rho_i$. After an approximation argument, this gives a finite (albeit very large) dimensional family of perturbations which can be used to perturb away degeneracy for primitive geodesics of a given length. In particular, for every such geodesic $\gamma_{g_0}$, we can find a ball, $B_i$, such that, with $g_{i,\sigma}$ the perturbation of $g_0$ associated to $\gamma_{g_0, \rho_i}$, the derivative of the map $\sigma \mapsto (\mathscr{P}_{\gamma_{g_{i,\sigma}, \rho_i}}, d\mathscr{P}_{\gamma_{g_{i,\sigma}, \rho_i}})$ is invertible at $\sigma$ for all $\sigma \in B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$. Once we have this in place, it will be possible to control the volume of the set of perturbations for which there is a degenerate primitive closed geodesic of some length, with some quantitative control on how degenerate such an orbit may be. In particular, once this volume is small enough, there is at least one perturbation of $g_0$ for which all primitive orbits of a certain length are non-degenerate. Since $g_0$ is arbitrary, and we make our perturbations arbitrarily small, we will eventually obtain density after an induction on the length of the closed geodesics.

**Remark 4.2.** In order to obtain predominance in our main theorem, we use the control on the volume of bad perturbations more seriously. In particular, making it smaller than $\varepsilon > 0$ for any chosen $\varepsilon$.

4.2. **The Induction.** The proof of Theorem 1.3 relies on an induction on the length of an orbit. In order to do this, we follow a strategy motivated by that of Yomdin [Yom85]. The work in [Yom85] shows that every diffeomorphism $f_0$ on $M$ can be perturbed to a diffeomorphism $f_\varepsilon$ in which every $n$-periodic point $x$ is $\gamma_\varepsilon$-hyperbolic in the sense that the eigenvalues $\lambda_j$ of $df_\varepsilon^n(x)$ satisfy $||\lambda_j| - 1| \geq \gamma_\varepsilon$ where $\gamma_\varepsilon$ is a power of $\varepsilon$ depending on $n, \varepsilon$ and the regularity of $f_0$. The idea there is to first use a perturbation to make primitive orbits hyperbolic and then to use the fact that hyperbolicity of an orbit passes to its multiples and, moreover, that every multiple is more hyperbolic than the previous multiple.

In contrast to what happens when working with diffeomorphisms on $M$, one of the main difficulties to overcome in the case of geodesic flows (or indeed symplectic maps) is that, in general, it is not possible to perturb a closed orbit to turn it into a hyperbolic closed orbit. Indeed, eigenvalues of the Poincaré map may be 'trapped' on the unit circle (see Section 5.1) and hence there are so-called stable closed orbits which cannot be perturbed away. Because of this structure, one must find a different property which guarantees non-degeneracy and can be passed from a primitive closed orbit to its multiples.

Given a metric $g_0$, our objective is to find a metric $g_\infty$, that is arbitrarily close to $g_0$, with the property that for each $\ell$ every periodic trajectory of length $\leq 2^\ell$ is $\beta_\ell$ non-degenerate in the sense of Definition 1.6. Here, $\{\beta_\ell\}_\ell$ is a decreasing sequence of numbers.

We build $g_\infty$ by induction on the parameter $\ell$ which represents (the logarithm of) the maximum length of the orbits up to which non-degeneracy is controlled. Note that we start with the most naive possible version of the induction and gradually add the details necessary to handle our situation.

**A.** *Hypothesis.* Assume $g_\ell$ is a metric such that for $k \leq \ell$, all closed trajectories of length $n \in (2^{k-1}, 2^k]$ are $\beta_k$ non-degenerate in the sense of Definition 1.6.

**B.** *Perturbation:* Find a perturbation, $g_{\ell+1}$ of $g_\ell$ which satisfies
- primitive closed trajectories of length $n \in (2^\ell, 2^{\ell+1}]$ are $\beta_{\ell+1}$ non-degenerate,
- closed trajectories of length $n \in (2^{k-1}, 2^k]$ with $k \leq \ell$ remain $\beta_k$ non-degenerate.

**C.** *Deal with non-primitive closed trajectories:* A non-primitive closed trajectory $\gamma$ for $g^{\ell+1}$ of length $n \in (2^\ell, 2^{\ell+1}]$ is a multiple of a primitive closed trajectory $\tilde{\gamma}$ of length $m \in (2^{k-1}, 2^k]$ with $k \leq \ell$. By the inductive hypothesis $\tilde{\gamma}$ is $\beta_k$ non-degenerate. Use this to show that the $\beta_k$ non-degeneracy of $\tilde{\gamma}$ implies $\beta_{\ell+1}$ non-degeneracy of $\gamma$.

*Why it fails:* The $\beta_k$ non-degeneracy of the orbit $\tilde{\gamma}$ does not imply $\beta_{\ell+1}$ non-degeneracy of its multiple $\gamma$. Indeed, the Poincaré map associated to $\gamma$ may have a root of unity as an eigenvalue. That is, Step C cannot be completed.

*Solution:* We define the notion of $(\beta, q)$-nondegeneracy. For $\aleph > 0$ and $\beta, q > 0$ we say a matrix $\boldsymbol{A} \in \mathbb{M}(\aleph)$, the set of $\aleph \times \aleph$ matrices, is $(\beta, q)$ non-degenerate if

$$\left\| (I - \boldsymbol{A}^q)^{-1} \right\| \leq (\tfrac{5}{2} q^2 \beta^{-1})^{\aleph} (1 + \|\boldsymbol{A}\|^q)^{\aleph - 1}. \tag{4.1}$$

The key observation here is that, if the derivative, $\boldsymbol{A} = d\mathscr{P}_\gamma$, of the Poincaré map associated to a closed orbit $\gamma$ is $(\beta, q)$-nondegenerate, then the $q$ iterate of the orbit is $(\tfrac{5}{2} q^2 \beta^{-1})^{-\aleph}$ non-degenerate (see Figure 2 for a schematic of a $(\beta, 2)$ non-degenerate orbit). Therefore, our goal will be to make perturbations of the metric, $g_\ell$, so that primitive orbits become $(\beta(q), q)$-non-degenerate for *all* $q$ and some sequence $\{\beta(q)\}_{q=1}^\infty$. This will be possible provided that $\beta(q) = O(q^{-2-0})$ as $q \to \infty$. The main difficulty with the notion of $(\beta(q), q)$ non-degeneracy is that, although for each fixed $(\beta(q), q)$ the property of $(\beta(q), q)$ non-degeneracy is stable under small perturbations, the property of being $(\beta(q), q)$ non-degenerate for all $q$ is not. Therefore, we only look for non-degeneracy for $q < Q$.

We next present the modified induction argument. The induction is done on the parameter $\ell$. One can find sequences of numbers $\{\beta_k\}_k$, $\{\beta_{k,\ell}\}_{k,\ell}$, and $\{Q_{k,\ell}\}_{k,\ell}$ so that the induction below can be completed.

**A.** *Hypothesis.* Assume $g_\ell$ is a metric such that for $k \leq \ell$
- closed trajectories of length $n \in (2^{k-1}, 2^k]$ are $\beta_k$ non-degenerate,
- primitive closed trajectories of length $n \in (2^{k-1}, 2^k]$ are $(\beta_{k,\ell} q^{-3}, q)$ non-degenerate for all $0 < q < Q_{k,\ell}$.

**B.** *Perturbation:* Find a perturbation, $g_{\ell+1}$, of $g_\ell$ which satisfies
- primitive closed trajectories of length $n \in (2^\ell, 2^{\ell+1}]$ for $g_{\ell+1}$ are $\beta_{\ell+1}$ non-degenerate
- closed trajectories of length $n \in (2^{k-1}, 2^k]$ for $g_{\ell+1}$ remain $\beta_k$ non-degenerate for all $k \leq \ell$.
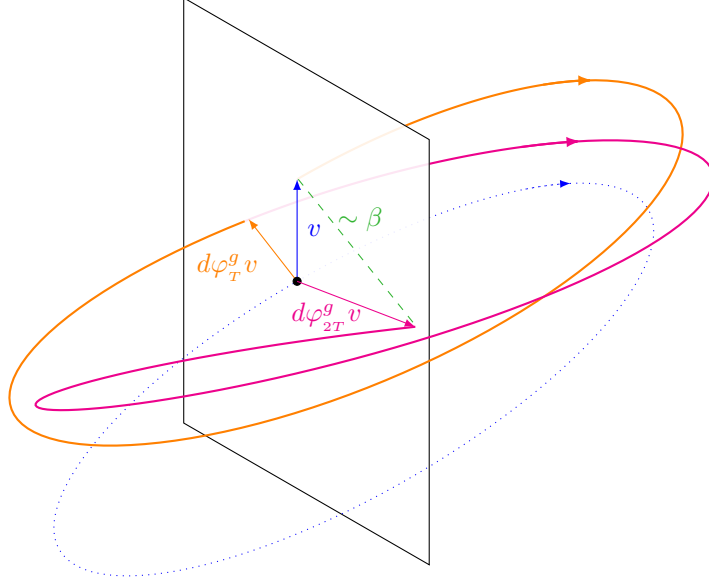
FIGURE 2. In the figure, we show a primitive periodic orbit of length $T$ such that the derivative $d\mathscr{P}$ of the Poincaré map is $(\beta, 2)$ non-degenerate. The primitive orbit is shown in the blue dotted line. Note that $d\mathscr{P}$ satisfies $(d\mathscr{P})^k = d\varphi^g_{kT}$. Thus, the property of $(\beta, 2)$ non-degeneracy implies that $\|d\varphi^g_{2T}v - v\| \sim \beta$.

– primitive closed trajectories of length $n \in (2^{k-1}, 2^k]$ for $g_{\ell+1}$ are $(\beta_{k,\ell+1}q^{-3}, q)$ non-degenerate for $0 < q < Q_{k,\ell+1}$ and $k \leq \ell + 1$,

**C.** *Deal with non-primitive closed trajectories:* Let $\gamma$ be a closed trajectory for $g_{\ell+1}$ of length $n \in (2^\ell, 2^{\ell+1}]$ that is non-primitive. Then $\gamma$ is a multiple of a primitive closed trajectory $\tilde{\gamma}$ of length $m \in (2^{k-1}, 2^k]$ with $k \leq \ell$. Note that then $\tilde{\gamma}$ satisfies the hypothesis in Step A. In particular, to show that $\gamma$ is $\beta_{\ell+1}$ non-degenerate, we would like to use that $\tilde{\gamma}$ is $(\beta_{k,\ell}q^{-3}, q)$ non degenerate with $q = n/m$, and for that we need $n/m < Q_{k,\ell}$. We then show that $2^{\ell+1-k} < Q_{k,\ell}$ and that by the inductive hypothesis this implies $\beta_{\ell+1}$ non-degeneracy of $\gamma$.

After the steps A-C above (modulo the fact that we have only consider exactly closed orbits), the inductive argument yields a metric $g_\infty$ which is arbitrarily close to $g_0$ and has the desired non-degeneracy property that, for each $\ell$, every closed trajectory for $g_\infty$ of length $\leq 2^\ell$ is $\beta_\ell$ non-degenerate. This yields a discretized version of the statement in Theorem 1.3.

It is important to note that one cannot do a simplified version of this induction in which a sequence $\{Q_k\}_k$ is used in place of $\{Q_{k,\ell}\}_{k,\ell}$ and $\{\beta_k\}_k$ is used in place of $\{\beta_{k,\ell}\}_{k,\ell}$. The problem would be that, when applying the inductive hypothesis to carry out Step C, one would need $n/m < Q_k$ so that the non-degeneracy of $\tilde{\gamma}$ may pass to that of $\gamma$. However, this may not be possible to arrange. For example, for a closed trajectory of length $n$ that is a multiple of a primitive trajectory of length $m = 1$. One might wonder, at this point, why we do not simply take $Q_{k,\ell} = Q_\ell = 2^{\ell+2}$. We will see below, in sections 4.3 and 4.4, that this is indeed not possible.

4.3. **How to deal with almost closed trajectories.** Next, we explain how to modify the argument to deal with trajectories that are not necessarily closed and primitive but are very close to being such. First of all, instead of working with closed trajectories, one works with $(n, \alpha)$ returning trajectories. In this discussion a point $\rho$ is said to be $(n, \alpha)$ returning if $d(\rho, \mathscr{P}^n(\rho)) \leq \alpha$, where we are still pretending that we can work with a globally defined Poincaré section $\mathscr{P}$ (see Section 4.5).

There is one serious difficulty which is hidden by imagining that we only deal with exactly closed trajectories. This difficulty arises from the fact that the notion of primitiveness for non-closed trajectories is not a well-defined condition. One should instead think of a degree of simplicity. We will say that a point $\rho$ is $(n, \alpha)$ simple if $d(\rho, \mathscr{P}^k(\rho)) > \alpha$ for $0 < k < n$. Instead of working with primitive closed trajectories of length $n$ one works with trajectories that are $(n, \alpha)$ simple for some $\alpha$.

It is in fact only possible to make an $(n, \alpha)$ simple point be $(\beta q^{-3}, q)$ non-degenerate whenever $\beta$ is small depending on $\alpha$. Therefore, there are sequences $\{\beta_{k,\ell}\}_{k,\ell}$ and $\{\alpha_{k,\ell}\}_{k,\ell}$ such that in the $\ell$-th step of the induction argument, orbits of length $n \in (2^{k-1}, 2^k)$ which are $\beta_{k,\ell}$ returning can only be made to be $(\beta_{k,\ell} q^{-3}, q)$ non-degenerate for $0 < q < Q_{k,\ell}$ provided they are $(n, \alpha_{k,\ell})$-simple and $\beta_{k,\ell}$ is small in terms of $\alpha_{k,\ell}$.

The induction argument is the same as the one outlined before but with each instance of the word 'closed' replaced by $(n, \beta)$ returning for some parameter $\beta$ depending on $(k, \ell)$ and each instance of the word 'primitive' replaced by $(n, \alpha)$ simple for some parameter $\alpha$ depending on $(k, \ell)$.

    **A.** *Hypothesis.* Assume $g_\ell$ is a metric such that for $k \leq \ell$

       – $\beta_{k,\ell}$ returning trajectories of length $n \in (2^{k-1}, 2^k]$ are $\beta_k$ non-degenerate,

       – $\beta_{k,\ell}$ returning trajectories that are $\alpha_{k,\ell}$ simple and have length $n \in (2^{k-1}, 2^k]$ are $(\beta_{k,\ell} q^{-3}, q)$ non-degenerate for all $0 < q < Q_{k,\ell}$.

    **B.** *Perturbation:* Find a perturbation, $g_{\ell+1}$, of $g_\ell$ which satisfies

       – $\beta_{\ell+1,\ell+1}$ returning trajectories that are $\alpha_{\ell+1,\ell+1}$ simple and have length $n \in (2^\ell, 2^{\ell+1}]$ are $\beta_{\ell+1}$ non-degenerate

       – $\beta_{k,\ell+1}$ returning trajectories of length $n \in (2^{k-1}, 2^k]$ will remain $\beta_k$ non-degenerate for all $k \leq \ell$.

       – $\beta_{k,\ell+1}$ returning trajectories that are $\alpha_{k,\ell+1}$ simple and have length $n \in (2^{k-1}, 2^k]$ will be $(\beta_{k,\ell+1} q^{-3}, q)$ non-degenerate for $0 < q < Q_{k,\ell+1}$ and $k \leq \ell + 1$.

    **C.** *Deal with non-simple trajectories:* This step becomes substantially more complicated as a result of needing to handle non-closed trajectories. One consequence of the new difficulties, is that, instead of using 2-adic intervals, we must use $\mathfrak{a}$-adic intervals for some $1 < \mathfrak{a} < 2$. We include a sketch of how to handle Step C below in Section 4.4, in order to present a more or less complete picture of the induction. However, the reader may wish to skip the next section on first reading.

4.4. **Detailed sketch of Step C.** Suppose that $\rho$ yields a trajectory of length $n \in (2^\ell, 2^{\ell+1}]$ that is $(n, \beta_{\ell+1,\ell+1})$ returning but not $(n, \alpha_{\ell+1,\ell+1})$ simple.

It is in general not possible to find an $m \leq n/2$ such that $\rho$ is $(m, \alpha_{\ell+1,\ell+1})$ returning and $(m, \alpha_{\ell+1,\ell+1})$ simple. Instead, one only obtains $(m_1, \tilde{\alpha}_1)$ returning with $\tilde{\alpha}_1 \sim C^n \alpha_{\ell+1,\ell+1}$ and $m_1 \leq 2^{-1}n$. Therefore, in Step C of the induction, if $m_1 \in (2^{k_1-1}, 2^{k_1})$, we would need $\tilde{\alpha}_1 < \beta_{k_1,\ell}$ *and* that the trajectory be at least $\alpha_{k_1,\ell}$ simple to apply the inductive hypothesis and obtain the $(\beta_{k_1,\ell} q^{-3}, q)$ non-degeneracy of the orbit for $q < Q_{k_1,\ell}$. In fact, the first condition that we impose on $\alpha_{\ell+1,\ell+1}$ is that $\tilde{\alpha}_1 < \beta_{k_1,\ell}$.

Then, we need to ask whether the trajectory is $(m, \alpha_{k_1,\ell})$ simple. If it is, we can apply the inductive hypothesis to obtain non-degeneracy. However, if it is not, we need to repeat this process, obtaining that $\rho$ is $(m_2, \tilde{\alpha}_2)$ returning, with $\tilde{\alpha}_2 \leq C^{m_1} \alpha_{k_1,\ell}$ and $m_2 \leq m_1/2$ with $m_2 \in (2^{k_2}, 2^{k_2+1}]$. As in the case of treating $\rho$ as $(m_1, \tilde{\alpha}_1)$ returning, we compare $\tilde{\alpha}_2$ and $\beta_{k_2,\ell-1}$ and ask whether $\rho$ is $(m_2, \alpha_{k_2,\ell-1})$ simple. The requirement that $\tilde{\alpha}_2 \leq \beta_{k_2,\ell-1}$ puts an additional upper bound on $\alpha_{k_1,\ell}$. If the trajectory is not simple enough, we repeat once again finding $(m_i, k_i, \tilde{\alpha}_i)$ until $\rho$ is $(m_i, \alpha_{k_i,\ell-i+1})$ simple. Note that this *will* happen, since every 1 returning trajectory is simple and hence, if $m_i = 1$, then, provided we have chosen the $\alpha_{k,\ell}$ correctly, $\rho$ is $\beta_{0,0}$ returning and $\alpha_{0,0}$ simple and hence we may apply the inductive hypothesis.

Since $m_i \leq m_{i-1}/2$, we have $m_i \leq 2^{-i}n \leq 2^{\ell+1-i}$. In some circumstances, we may be working with $n = 2^{\ell+1}$ and the iteration may not terminate until $m_i = 1$. If this is the case, then we want to use the inductive hypothesis on $(1, \beta_{1,1})$ returning points to obtain non-degeneracy of $\gamma$. However, this requires that every $(1, \beta_{0,0})$ returning point be $(\beta_{0,0} 2^{-3(\ell+1)}, 2^{\ell+1})$ non-degenerate. This will eventually fail, since $Q_{1,1} < \infty$ and $\ell$ is unbounded above. This can be remedied by putting $1 < \mathfrak{a} < 2$ and replacing 2-adic intervals by $\mathfrak{a}$-adic intervals. By doing this, we guarantee that $1 \leq m_i \leq 2^{-i} \mathfrak{a}^\ell$ and hence, the above iteration terminates in $i < c\ell$ steps for some $c < 1$.

### 4.5. Comments on the lack of a global Poincaré section.

In the outline above, we have worked as though one can find a single compact symplectic manifold without boundary, $\Gamma \subset S^*M$, to serve as a Poincaré section. In general, this may not be the case and one may need to work with Poincaré sections with boundaries. This leads to a number of technical complications in the argument which are handled by introducing the notion of chains of symplectomorphisms and well-separated sets (see Section 6.2). However, these technical complications do not lead to any fundamental change in the proof and the reader may wish to imagine that there *is* a global Poincaré section; in particular, the reader can then safely ignore the discussion about chains of symplectomorphisms and well-separated sets, instead thinking of the Poincaré map and its iterates.

### 4.6. Comments on proving predominance.

In order to prove predominance instead of density, at each step of the induction, rather than finding a single metric, $g_\ell$, we produce a family $g_{\ell,\boldsymbol{\sigma}}$ for $\boldsymbol{\sigma} \in \prod_{j=1}^\infty B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$ in a set with almost full measure. This is possible since we actually showed that, at each step, all but a small measure of our perturbations have the required properties.

## 5. VOLUME OF THE SET OF $(\beta, q)$-DEGENERATE SYMPLECTIC MATRICES

In this section, we do some preliminary work on volumes and covering numbers (see Definition 5.4) in the space of symplectic matrices with real coefficients in $2\aleph$ dimensions, $\mathrm{Sp}(2\aleph)$. As explained in Section 4, it is necessary to work with $(\beta, q)$-non-degenerate matrices (see (4.1) for

the definition). The goal of this section is to understand the covering number of the complement of these matrices in $\mathrm{Sp}(2\aleph)$.

## 5.1. Properties of eigenvalues of symplectic matrices.

We recall here a few properties of the spectrum of a symplectic matrix $\boldsymbol{A} \in \mathrm{Sp}(2\aleph)$. A $2\aleph \times 2\aleph$ matrix $\boldsymbol{A}$ with real coefficients is in $\mathrm{Sp}(2\aleph)$ provided

$$\boldsymbol{A}^T \boldsymbol{J} \boldsymbol{A} = \boldsymbol{J}, \qquad \boldsymbol{J} = \begin{pmatrix} 0 & I_\aleph \\ -I_\aleph & 0 \end{pmatrix}, \tag{5.1}$$

where $I_\aleph \in \mathbb{M}(\aleph)$ is the identity matrix. Symplectic matrices are invertible with $\|\boldsymbol{A}\| = \|\boldsymbol{A}^{-1}\|$. They also have $\det \boldsymbol{A} = 1$.

We note that if $\lambda$ is in the spectrum of $\boldsymbol{A}$ and has multiplicity $m$, then $\lambda^{-1}$, $\bar{\lambda}$ and $\bar{\lambda}^{-1}$ are also in the spectrum and have multiplicity $m$. However, $\lambda$, $\lambda^{-1}$, $\bar{\lambda}$ and $\bar{\lambda}^{-1}$ may not be distinct. First, if $\lambda \notin \mathbb{R}$ and $|\lambda| \neq 1$, then all four are distinct. Thus, each non-real eigenvalue off the unit circle comes in a quadruplet $\{\lambda, \lambda^{-1}, \bar{\lambda}, \bar{\lambda}^{-1}\}$. Next, if $\lambda \notin \mathbb{R}$ with $|\lambda| = 1$, then $\lambda^{-1} = \bar{\lambda}$ and such eigenvalues come in pairs of the form $\{\lambda, \lambda^{-1}\}$. Third, if $\lambda \in \mathbb{R}$ with $|\lambda| \neq 1$, then $\lambda = \bar{\lambda}$ and the eigenvalues again come in pairs of the form $\{\lambda, \lambda^{-1}\}$. Finally, if $\lambda = \pm 1$, then $\lambda = \bar{\lambda} = \lambda^{-1} = \bar{\lambda}^{-1}$, and it is known that $\pm 1$ always occurs with even multiplicity.

We say that an invertible matrix is hyperbolic if none of its eigenvalues lie on the unit circle. Note that if $\boldsymbol{A}$ is hyperbolic then the same is true for $\boldsymbol{A}^q$ for all $q$. Moreover, the distance from the eigenvalues to the unit circle is increasing with $q$.

One important consequence of the facts we recalled about eigenvalues of symplectic matrices is that every small enough perturbation of a symplectic matrix with an eigenvalue $\lambda_0 \in \{\lambda : |\lambda| = 1, \lambda \neq \pm 1\}$ having odd multiplicity must have an eigenvalue on the unit circle since eigenvalues off the unit circle come in quadruplets. In other words, certain eigenvalues are 'trapped' on the unit circle. Thus, unlike general matrices, symplectic matrices *cannot* be perturbed to be hyperbolic. This makes perturbing a matrix to one which is $(\beta(q), q)$-non-degenerate for all $q$ much more delicate in the symplectic category than in the category of matrices and requires the delicate induction argument carried out in Section 8.

## 5.2. $(\beta, q)$-non-degeneracy and the location of eigenvalues.

We start by discussing the relationship between $(\beta, q)$-non-degeneracy and the location of the eigenvalues $\{\lambda_j(\mathbf{A})\}_{j=1}^{2\aleph}$ of a matrix $\mathbf{A}$.

**Lemma 5.1.** *For all $\mathbf{A} \in \mathbb{M}(2\aleph)$, $q \in \mathbb{N}$, $s \in (0, 1)$,*

$$if \quad \min_{\substack{1 \leq j \leq 2\aleph \\ 0 \leq p \leq q-1}} |\lambda_j(\mathbf{A}) - e^{2\pi i p/q}| \geq s, \quad then \quad |\det(I - \mathbf{A}^q)| > \left(\min(\tfrac{2}{5}qs, \tfrac{1}{5})\right)^{2\aleph}.$$

*Proof.* First, note that $\lambda_j(I - \mathbf{A}^q) = 1 - \lambda_j(\mathbf{A})^q$. Therefore, $|\det(I - \mathbf{A}^q)| = \prod_{j=1}^{2\aleph} |(1 - \lambda_j(\mathbf{A})^q)|$. The lemma will follow once we prove that for all $q \in \mathbb{N}$, $s \in (0, 1)$,

$$\min_{j,p} |\lambda_j(\mathbf{A}) - e^{2\pi i p/q}| \geq s, \quad \Rightarrow \quad \min_j |1 - \lambda_j(\mathbf{A})^q| \geq \min(\tfrac{2}{5}qs, \tfrac{1}{5}). \tag{5.2}$$

Fix $1 \leq j \leq 2\aleph$, and set $\lambda_j(\mathbf{A}) = re^{i\theta}$ for some $r > 0$ and $\theta \in [0, 2\pi)$. First, notice that if $|r - 1| \geq \frac{1}{2}$, then $|1 - \lambda_j(\mathbf{A})^q| \geq \frac{1}{2}$ for all $q \in \mathbb{N}$, and hence (5.2) holds in this case. Therefore, we assume that $|r - 1| < \frac{1}{2}$ from now on.

Next, fix $q \in \mathbb{N}$, $s \in (0,1)$, and for $p \in \mathbb{N}$ let $\alpha_p = \theta - 2\pi\frac{p}{q}$. We claim that

$$\text{either} \quad |\alpha_p| \geq \tfrac{1}{2}s \quad \text{for all } p, \qquad \text{or} \qquad |r-1| \geq \tfrac{1}{2}s. \tag{5.3}$$

To prove the claim suppose there exists $p$ such that $|\alpha_p| < \frac{s}{2}$. First, note that by the assumption in (5.2)

$$s^2 \leq |re^{i\theta} - e^{2\pi ip/q}|^2 = |1 - re^{i\alpha_p}|^2 = (r - \cos\alpha_p)^2 + (\sin\alpha_p)^2.$$

Therefore, $\frac{3s^2}{4} \leq (r - \cos\alpha_p)^2$ since $|\sin\alpha_p| \leq |\alpha_p| < s/2$. In particular, since $|r-1| < \frac{1}{2}$ and $|1 - \cos\alpha_p| = 2|\sin^2(\frac{\alpha_p}{2})| \leq \frac{s^2}{8}$,

$$\tfrac{3s^2}{4} \leq \left(r - 1 + (1 - \cos\alpha_p)\right)^2 \leq (r-1)^2 + \tfrac{s^2}{4},$$

and hence $|r-1| \geq \frac{1}{\sqrt{2}}s \geq \frac{1}{2}s$ as claimed in (5.3).

With (5.3) in place, we first consider the case $|r-1| \geq \frac{s}{2}$ and divide it into two sub-cases.

**Case 1:** Suppose that $r \geq 1 + \frac{s}{2}$. Then, $|1 - \lambda_j(\mathbf{A})^q| \geq r^q - 1 \geq q\frac{s}{2}$ as claimed in (5.2).

**Case 2:** Suppose $r \leq 1 - \frac{s}{2}$. Then, $r^{-q} \geq 1 + q\frac{s}{2}(1 - \frac{s}{2})^{-1}$ since $s \in (0,1)$. Thus, $r^q \leq 1 - q\frac{s}{2}\frac{1}{1+(q-1)\frac{s}{2}}$. In particular, as claimed in (5.2),

$$|1 - \lambda_j(\mathbf{A})^q| \geq 1 - r^q \geq qs(2 + (q-1)s)^{-1} \geq \min(qs\tfrac{2}{5}, \tfrac{1}{5}).$$

Finally, suppose that $|r-1| < \frac{s}{2}$. Then, $|\alpha_p| \geq \frac{s}{2}$ for all $p$ by (5.3) and hence

$$|1 - \lambda_j(\mathbf{A})^q|^2 = |1 - r^q e^{iq\theta}|^2 \geq \inf_{a \geq 0} (1 - a\cos(q\theta))^2 + a^2 \sin^2(q\theta) = \begin{cases} \sin^2(q\theta) & \inf_{k \in \mathbb{Z}} |q\theta - 2\pi k| < \frac{\pi}{2}, \\ 1 & \text{else.} \end{cases}$$

In particular, since $\frac{s}{2} \leq \min_p |\alpha_p| \leq \frac{\pi}{q}$, we have $|1 - \lambda_j(\mathbf{A})^q| \geq |\sin(q\frac{s}{2})| \geq \min(\frac{2}{5}qs, \frac{1}{5})$.

$\square$

In what follows, it will be necessary to avoid symplectic matrices which have eigenvalues close to a given number. To this end, for $\theta \in \mathbb{R}$ and $s > 0$ let

$$\mathcal{D}(\theta, s) := \{\mathbf{A} \in \mathrm{Sp}(2\aleph) : B_{\mathbb{C}}(e^{i\theta}, s) \cap \mathrm{spec}(\mathbf{A}) \neq \emptyset\}. \tag{5.4}$$

We now record a corollary of Lemma 5.1 which we use in the next section. It controls the size of the inverse of $\mathbf{I} - \mathbf{A}^q$ in terms of the distance between the spectrum of $\mathbf{A}$ and the roots of unity $\{e^{2\pi ip/q}\}_{0 \leq p < q}$.

**Lemma 5.2.** *Let* $q \in \mathbb{N}$, $0 < s \leq \frac{1}{2q}$, *and* $\mathbf{A} \in \mathbb{M}(2\aleph)$. *If* $\mathbf{A} \notin \bigcup_{p=0}^{q-1} \mathcal{D}(\frac{2\pi p}{q}, s)$, *then,*

$$\|(\mathbf{I} - \mathbf{A}^q)^{-1}\| \leq (\tfrac{5}{2})^{2\aleph}(qs)^{-2\aleph}(1 + \|\mathbf{A}\|^q)^{2\aleph - 1}.$$

*Proof.* Note that by assumption $\min_{j,p} |\lambda_j(\mathbf{A}) - e^{2\pi ip/q}| \geq s$ and hence $|\det(\mathbf{I} - \mathbf{A}^q)| \geq (\frac{2}{5}qs)^{2\aleph}$ by Lemma 5.1 together with $s \leq \frac{1}{2q}$. In particular, the corollary follows from the fact that $\|\mathbf{B}^{-1}\| \leq |\det(\mathbf{B})|^{-1}\|\mathbf{B}\|^{2\aleph - 1}$ for all $\mathbf{B} \in \mathbb{M}(2\aleph)$. Indeed, let $\sigma_1^2 \leq \sigma_2^2 \leq \cdots \leq \sigma_{2\aleph}^2$ be the

eigenvalues of $\mathbf{B}^*\mathbf{B}$. Then, $\det(\mathbf{B}^*\mathbf{B}) = |\det(\mathbf{B})|^2 = \prod_{j=1}^{2\aleph} \sigma_j^2$, $\|\mathbf{B}^{-1}\|^2 = \sigma_1^{-2}$, and $\|\mathbf{B}\|^2 = \sigma_{2\aleph}^2$. Therefore, as claimed,

$$\|\mathbf{B}^{-1}\|^2 = \sigma_1^{-2} \leq \sigma_{2\aleph}^{2(2\aleph-1)} \Big(\prod_{j=1}^{2\aleph} \sigma_j^2\Big)^{-1} = \|\mathbf{B}\|^{2(2\aleph-1)} |\det(\mathbf{B})|^{-2}.$$

$\square$

An immediate consequence of the previous lemma is the following.

**Corollary 5.3.** *Let $q \in \mathbb{N}$ and $\beta > 0$ with $2\beta < q^2$. If $\mathbf{A}$ is not $(\beta, q)$ non-degenerate, then $\mathbf{A} \in \bigcup_{p=0}^{q-1} \mathcal{D}(\frac{2\pi p}{q}, \beta q^{-3})$.*

Corollary 5.3 will be used to estimate the volume of the set of matrices which are not $(\beta, q)$ non-degenerate using covering numbers.

### 5.3. Covering numbers and $(\beta, q)$ non-degeneracy. Throughout the article, the notion of the covering number of a set will play an important role. Let $(Y, d)$ be a metric space.

**Definition 5.4.** The *covering number of a set $V \subset Y$ and radius $s$* is defined by

$$\mathcal{M}_Y(V, s) := \inf \Big\{ N \in \mathbb{N} : \exists \{y_i\}_{i=1}^N \subset Y \text{ such that } V \subset \bigcup_{i=1}^N B(y_i, s) \Big\}.$$

5.3.1. *Covering numbers in $\mathbb{M}(2\aleph)$ and $\mathrm{Sp}(2\aleph)$.* In this subsection, we will relate covering numbers in $\mathbb{M}(2\aleph)$ to those in $\mathrm{Sp}(2\aleph)$. The goal is to compare covering numbers of a given set as measured by balls in $\mathrm{Sp}(2\aleph)$ and in $\mathbb{M}(2\aleph)$. The following lemma controls how $\mathrm{Sp}(2\aleph)$ sits inside $\mathbb{M}(2\aleph)$.

**Lemma 5.5.** *There exists $\varepsilon_0 > 0$ such that for all $\mathbf{A} \in \mathrm{Sp}(2\aleph)$ and $0 < r \leq \varepsilon_0 \|\mathbf{A}\|$,*

$$B_{\mathbb{M}(2\aleph)}\big(\mathbf{A}, \tfrac{r}{2\|\mathbf{A}\|^2}\big) \cap \mathrm{Sp}(2\aleph) \subset B_{\mathrm{Sp}(2\aleph)}(\mathbf{A}, r).$$

*Proof.* Let $\mathbf{A} \in \mathrm{Sp}(2\aleph)$ and set $\mathcal{V} := \mathbf{A}^{-1}(B_{\mathrm{Sp}(2\aleph)}(\mathbf{A}, r)) \subset \mathrm{Sp}(2\aleph)$. We have that $\mathcal{V}$ is a neighborhood of $\mathbf{I}$ in $\mathrm{Sp}(2\aleph)$ and $B_{\mathrm{Sp}(2\aleph)}(\mathbf{I}, \|\mathbf{A}\|^{-1}r) \subset \mathcal{V}$. Let $\varepsilon_0 = \varepsilon_0(\aleph)$ be such that for $0 < \tilde{r} \leq \varepsilon_0$

$$B_{\mathrm{Sp}(2\aleph)}(\mathbf{I}, \tfrac{1}{4}\tilde{r}) \subset B_{\mathbb{M}(2\aleph)}(\mathbf{I}, \tfrac{1}{2}\tilde{r}) \cap \mathrm{Sp}(2\aleph) \subset B_{\mathrm{Sp}(2\aleph)}(\mathbf{I}, \tilde{r})$$

Therefore, $B_{\mathbb{M}(2\aleph)}(\mathbf{I}, \tfrac{r}{2\|\mathbf{A}\|}) \cap \mathrm{Sp}(2\aleph) \subset \mathcal{V}$ for $r \leq \varepsilon_0 \|\mathbf{A}\|$. Hence,

$$\Big(B_{\mathbb{M}(2\aleph)}\Big(\mathbf{A}, \tfrac{r}{2\|\mathbf{A}\|\|\mathbf{A}^{-1}\|}\Big) \cap \mathrm{Sp}(2\aleph)\Big) \subset \mathbf{A}\Big(B_{\mathbb{M}(2\aleph)}(\mathbf{I}, \tfrac{r}{2\|\mathbf{A}\|}) \cap \mathrm{Sp}(2\aleph)\Big) \subset \mathbf{A}\mathcal{V} = B_{\mathrm{Sp}(2\aleph)}(\mathbf{A}, r).$$

Since $\mathbf{A} \in \mathrm{Sp}(2\aleph)$, we have $\|\mathbf{A}\| = \|\mathbf{A}^{-1}\|$ and the result follows. $\square$

With Lemma 5.5 in place, we are ready to study the relationship between the covering numbers of a given set as measured by balls in $\mathrm{Sp}(2\aleph)$ and in $\mathbb{M}(2\aleph)$. To ease notation, for $r > 0$ we also introduce the ball

$$B_{\mathbf{A_0}}(r) := \{\mathbf{A} \in \mathbb{M}(2\aleph) : \|\mathbf{A} - \mathbf{A_0}\| \leq r\}.$$

**Lemma 5.6.** *There exists $\varepsilon_0 > 0$ such that the following holds. If $\mathcal{Z} \subset B_{\mathbf{0}}(r) \cap \mathrm{Sp}(2\aleph)$ for some $r > 0$, $\mathbf{A_0} \in \mathrm{Sp}(2\aleph)$, $0 < s \le \varepsilon_0 r$, and $r_0 > 0$, then*

$$\mathcal{M}_{\mathrm{Sp}(2\aleph)}(\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0), s) \le \mathcal{M}_{\mathbb{M}(2\aleph)}(\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0), \tfrac{1}{4} s r^{-2}).$$

*Proof.* Fix $r_0 > 0$. Let $\tilde{s} > 0$ and $\{\mathbf{A_i}\}_{i=1}^N \subset \mathbb{M}$ be such that $\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0) \subset \bigcup_{i=1}^N B_{\mathbf{A_i}}(\tilde{s})$. Without loss of generality, we may assume that $B_{\mathbf{A_i}}(\tilde{s}) \cap \mathcal{Z} \ne \emptyset$. Let $\{\mathbf{Z_i}\}_{i=1}^N \subset \mathcal{Z}$ be such that $\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0) \subset \bigcup_{i=1}^N B_{\mathbb{M}(2\aleph)}(\mathbf{Z}_i, 2\tilde{s})$. Then, since $\mathcal{Z} \subset B_{\mathbf{0}}(r) \cap \mathrm{Sp}(2\aleph)$, by Lemma 5.5 there is $\varepsilon_0 > 0$ such that

$$\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0) \subset \bigcup_{i=1}^N B_{\mathrm{Sp}(2\aleph)}(\mathbf{Z}_i, 4\tilde{s}\|\mathbf{Z_i}\|^2) \subset \bigcup_{i=1}^N B_{\mathrm{Sp}(2\aleph)}(\mathbf{Z}_i, 4\tilde{s}r^2),$$

provided $4\tilde{s}r \le \varepsilon_0$. In particular, $\mathcal{M}_{\mathrm{Sp}(2\aleph)}(\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0), 4\tilde{s}r^2) \le \mathcal{M}_{\mathbb{M}(2\aleph)}(\mathcal{Z} \cap B_{\mathbf{A_0}}(r_0), \tilde{s})$. The Lemma follows by putting $s = 4\tilde{s}r^2$. $\qquad\square$

5.3.2. *Covering numbers for matrices that are not $(\beta, q)$ non-degenerate.* We now study the covering numbers of the set of matrices in $\mathrm{Sp}(2\aleph)$ with a given eigenvalue of unit length: for $\theta \in \mathbb{R}$ let

$$\mathcal{D}(\theta) := \{\mathbf{A} \in \mathrm{Sp}(2\aleph) : e^{i\theta} \in \mathrm{spec}(\mathbf{A})\}. \tag{5.5}$$

In particular, $\mathcal{D}(\theta) = \cap_s \mathcal{D}(\theta, s)$ with $\mathcal{D}(\theta, s)$ as in (5.4). We start by bounding the covering number of $\mathcal{D}(\theta)$ in $\mathbb{M}(2\aleph)$.

**Lemma 5.7.** *There is $C > 0$ such that for all $\mathbf{A}_0 \in \mathbb{M}(2\aleph)$, $\theta \in \mathbb{R}$, $r_0 > 0$, and $0 < s < r_0$,*

$$\mathcal{M}_{\mathbb{M}(2\aleph)}(\mathcal{D}(\theta) \cap B_{\mathbf{A_0}}(r_0), \ s) \le C(r_0/s)^{\mathfrak{L}-2\aleph-1}, \qquad \mathfrak{L} = \aleph(2\aleph + 3).$$

*Proof.* Note that $\mathrm{Sp}(2\aleph) \subset \mathbb{M}(2\aleph)$ is defined by $\aleph(2\aleph - 1)$ algebraically independent polynomial equations. Now, having eigenvalue $e^{i\theta}$ is equivalent to asking both $\Re \det(\mathbf{A} - e^{i\theta}\mathbf{I}) = 0$ and $\Im \det(\mathbf{A} - e^{i\theta}\mathbf{I}) = 0$. In particular, $\mathcal{D}(\theta) \subset \mathcal{R}(\theta) := \{\mathbf{A} \in \mathrm{Sp}(2\aleph) : \Re \det(\mathbf{A} - e^{i\theta}\mathbf{I}) = 0\}$.

We claim that $\dim \mathcal{R}(\theta) = \dim \mathbb{M}(2\aleph) - (\aleph(2\aleph - 1) + 1) = \mathfrak{L} - 2\aleph - 1$. Once we prove this, we would have that $\mathcal{D}(\theta)$ is an algebraic set with $\dim \mathcal{D}(\theta) \le \mathfrak{L} - 2\aleph - 1$. Then, the bound on the covering number for $\mathcal{D}(\theta) \cap B_{\mathbf{A_0}}(r_0)$ will follow from the estimates on covering numbers for algebraic sets given in [YC04, Corollary 5.7].

To prove that $\dim \mathcal{R}(\theta) = \mathfrak{L} - 2\aleph - 1$, it suffices to show that $\Re \det(\mathbf{A} - e^{i\theta}\mathbf{I})$ does not just take the value 0 in $\mathrm{Sp}(2\aleph)$. Indeed, one may find matrices $\mathbf{A}$ in $\mathrm{Sp}(2\aleph)$ such that $\Re \det(\mathbf{A} - e^{i\theta}\mathbf{I})$ takes all values in an open set in $\mathbb{R}$. To see this, let $\mathbf{A}$ be the diagonal matrix with entries $(\lambda_1, \lambda_2, \ldots, \lambda_\aleph, \lambda_1^{-1}, \ldots, \lambda_\aleph^{-1})$ with $\lambda_i \in \mathbb{R} \setminus \{0\}$ and adjust $\lambda_i$ as needed. (In fact, it is enough to take $\lambda_3, \ldots, \lambda_\aleph = 1$ except when $\theta = 0$ in which case one can take $\lambda_3, \ldots, \lambda_\aleph = -1$.) $\qquad\square$

Our next goal is to bound the covering number of $\mathcal{D}(\theta, s)$. To do this, in Lemma 5.9 below, we first control how $\mathcal{D}(\theta)$ sits inside $\mathcal{D}(\theta, s)$. The proof of Lemma 5.9 hinges on the following result.

**Lemma 5.8.** *The set $\{A \in \mathrm{Sp}(2\aleph) : A \text{ has simple eigenvalues}\}$ is dense in $\mathrm{Sp}(2\aleph)$.*

*Proof.* Given $A \in \mathrm{Sp}(2\aleph)$ with eigenvalues $\{\lambda_i\}_{i=1}^{2\aleph}$ define the polynomial $p(A) = \prod_{i \neq j}(\lambda_i - \lambda_j)$ and notice that $A$ has simple eigenvalues if and only if $p(A) \neq 0$. Next, recall that any symmetric polynomial in the eigenvalues of a matrix can be expressed as a polynomial in the coefficients of that matrix [OCV11, Proposition 7.1.10]. In particular, since $\prod_{i \neq j}(\lambda_i - \lambda_j)$ is symmetric, there is $q : \mathrm{Sp}(2\aleph) \to \mathbb{C}$ such that $q(A)$ is a polynomial in the $(2\aleph)^2$ coefficients of the matrix with $q(A) \neq 0$ if and only if $A$ has simple eigenvalues (see also [dlCS21, Example ii]).

Let $A \in \mathrm{Sp}(2\aleph)$. Then $A$ can be written as $A = e^X e^Y$ for two elements $X, Y \in \mathfrak{sp}(2\aleph)$. (To see this, we combine [Hal15, Corollary 11.10], which states that the exponential map is surjective on compact, connected Lie groups, with the polar decomposition $A \in \mathrm{Sp}(2\aleph) = Be^Y$ with $Y \in \mathfrak{sp}(2\aleph)$ and $B \in \mathrm{SO}(2n) \cap \mathrm{Sp}(2n)$ [HN12, Proposition 4.3.3], noting that $\mathrm{SO}(2n) \cap \mathrm{Sp}(2n)$ is a compact Lie group.) Furthermore, it is easy to see that any diagonal matrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_\aleph, \lambda_1^{-1}, \ldots \lambda_\aleph^{-1})$, with $\lambda_i \in (0, \infty) \backslash \{1\}$, $\lambda_i \neq \lambda_j$, for $i \neq j$ has simple eigenvalues and lies in $\mathrm{Sp}(2\aleph)$. Fix one such $D$ and for $t \in \mathbb{R}$ let

$$F(t) = D^{1-t} e^{tX} e^{tY} \in \mathrm{Sp}(2\aleph).$$

We claim that there is $\{t_n\}_{n=1}^\infty \subset \mathbb{R}$ such that $F(t_n) \to F(1) = A$, and $F(t_n)$ has simple eigenvalues. This would prove the density statement. To prove the claim note that $q(F(t))$ extends to an analytic function of $t \in \mathbb{C}$ and $q(F(0)) = q(D) \neq 0$. In particular, $q(F(t))$ has a discrete set of zeros and hence, there is $t_n \in \mathbb{R}$ such that $t_n \to 1$ and $q(F(t_n)) \neq 0$. Since $F : \mathbb{R} \to \mathrm{Sp}(2\aleph)$ is continuous, $F(t_n) \to A$ as claimed. $\qquad\square$

**Lemma 5.9.** *Let $a > 0$ and suppose that $A \in \mathrm{Sp}(2\aleph)$ has an eigenvalue $\lambda$ with $|\lambda| > a$. Then, for all $\theta \in \mathbb{R}$ there is $B \in \mathcal{D}(\theta)$ and*

$$\|A - B\| \leq 3(1 + a^{-1})|\lambda - e^{i\theta}|. \tag{5.6}$$

*Proof.* Let $\theta \in \mathbb{R}$. Without loss of generality, we may assume that $\lambda \neq e^{i\theta}$ since otherwise the claim is trivial. By Lemma 5.8, for any $\varepsilon > 0$, there is $\tilde{A} \in \mathrm{Sp}(2\aleph)$ with simple eigenvalues, one of which we label $\mu$, such that

$$\|A - \tilde{A}\| \leq \varepsilon |\lambda - e^{i\theta}|, \qquad |\mu - \lambda| \leq \varepsilon \min(|\lambda - e^{i\theta}|, 1). \tag{5.7}$$

Without loss of generality we assume that $\varepsilon$ is small enough that

$$\varepsilon + 2(1 + \varepsilon)(1 + |a - \varepsilon|^{-1}) \leq 3(1 + a^{-1}). \tag{5.8}$$

Let $E_\mu(\tilde{A})$ denote the eigenspace of $\mu$ for $\tilde{A}$. Then we need to consider three cases $\mu \notin S^1 \cup \mathbb{R}$, $\mu \in S^1$, and $\mu \in \mathbb{R}$.

**Case 1: $\mu \notin S^1 \cup \mathbb{R}$.** In this case, one has that $\mu, \mu^{-1}, \bar{\mu}, \bar{\mu}^{-1}$ are distinct, and hence we have $E_\mu, E_{\bar{\mu}}, E_{\mu^{-1}}, E_{\bar{\mu}^{-1}}$ are distinct subspaces of dimension 1. Note that $\overline{E_\mu} = E_{\bar{\mu}}$ and $\overline{E_{\mu^{-1}}} = E_{\bar{\mu}^{-1}}$. Consider now the matrix

$$B = \tilde{A} + (e^{i\theta} - \mu)\Pi_\mu + (e^{-i\theta} - \bar{\mu})\Pi_{\bar{\mu}} + (e^{-i\theta} - \mu^{-1})\Pi_{\mu^{-1}} + (e^{i\theta} - \bar{\mu}^{-1})\Pi_{\bar{\mu}^{-1}},$$

where $\Pi_\mu$ denotes the projector which gives the component of a vector in the $E_\mu$ direction i.e. if $\{\mu_i\}_{i=1}^{2\aleph}$ are the eigenvalues of $\tilde{A}$, then $\Pi_{\mu_i} : \mathbb{C}^{2\aleph} \to E_{\mu_i}$ are the unique operators such that $v = \sum_i \Pi_{\mu_i} v$ for $v \in \mathbb{C}^{2\aleph}$.

Note that, $(B - e^{i\theta})\Pi_\mu = 0$ since $A\Pi_\mu = \mu\Pi_\mu$. That is, $B$ has eigenvalue $e^{i\theta}$. Then, by (5.7)

$$\|B - \tilde{A}\| \le |e^{i\theta} - \mu| + |e^{-i\theta} - \mu^{-1}| + |e^{-i\theta} - \bar{\mu}| + |e^{i\theta} - \bar{\mu}^{-1}|$$
$$\le 2(1 + |\mu|^{-1})|\mu - e^{i\theta}| \le 2(1 + |\mu|^{-1})(1 + \varepsilon)|\lambda - e^{i\theta}|$$
$$\le 2(1 + |a - \varepsilon|^{-1})(1 + \varepsilon)|\lambda - e^{i\theta}|.$$

The bound on (5.6) then follows from (5.8) and (5.7).

To see that $B \in \mathcal{D}(\theta)$, it remains to check that $B \in \mathrm{Sp}(2\aleph)$. First, to see that $B$ has real coefficients one must check that $Bv = \overline{Bv}$ for all $v$ real valued vectors. This claim follows from the facts that

$$\tilde{A}v = \overline{\tilde{A}v}, \quad \Pi_\mu \bar{v} = \overline{\Pi_{\bar{\mu}} v}, \quad \Pi_{\bar{\mu}} \bar{v} = \overline{\Pi_\mu v}, \quad \Pi_{\mu^{-1}} \bar{v} = \overline{\Pi_{\bar{\mu}^{-1}} v}, \quad \Pi_{\bar{\mu}^{-1}} \bar{v} = \overline{\Pi_{\mu^{-1}} v}.$$

To see that $B$ is symplectic, note that if $\mu_i \mu_j \ne 1$, then $\langle Jv_{\mu_i}, v_{\mu_j} \rangle = 0$ for $v_{\mu_i} \in E_{\mu_i}$ since

$$\langle Jv_{\mu_i}, v_{\mu_j} \rangle = \langle J\tilde{A}v_{\mu_i}, \tilde{A}v_{\mu_j} \rangle = \mu_i \mu_j \langle Jv_{\mu_i}, v_{\mu_j} \rangle. \tag{5.9}$$

Therefore, since $B$ preserves $E_{\mu_i}$ for all $i$ and $\mathbb{R}^{2\aleph} = \oplus_{i=1}^{2\aleph} E_{\mu_i}$ it suffices to note

$$\langle JBv_\mu, Bv_{\mu^{-1}} \rangle = \langle J(\tilde{A} + (e^{i\theta} - \mu))v_\mu, (\tilde{A} + (e^{-i\theta} - \mu^{-1}))v_{\mu^{-1}} \rangle = \langle Jv_\mu, v_{\mu^{-1}} \rangle.$$

**Case 2:** $\mu \in S^1$. Since $\mu \in S^1$, then $\mu = \bar{\mu}^{-1}$. We consider

$$B = \tilde{A} + (e^{i\theta} - \mu)\Pi_\mu + (e^{-i\theta} - \mu^{-1})\Pi_{\mu^{-1}}.$$

As before, $(B - e^{i\theta})\Pi_\mu = 0$, and $\|\tilde{A} - B\| \le (1 + \varepsilon)(1 + |a - \varepsilon|^{-1})|\lambda - e^{i\theta}|$. The bound on (5.6) then follows from (5.8) and (5.7). Therefore, to see that $B \in \mathcal{D}(\theta)$ we need to check that $B$ is real and symplectic. The computation is identical to that in Case 1.

**Case 3:** $\mu \in \mathbb{R}$. In the case $\mu \in \mathbb{R}$, we let $v_\mu \in \mathbb{R}^{2\aleph}$ with $\|v_\mu\| = 1$ and $(\tilde{A} - \mu)v_\mu = 0$, $v_{\mu^{-1}} \in \mathbb{R}^{2\aleph}$ with $\|v_{\mu^{-1}}\| = 1$, and $(\tilde{A} - \mu^{-1})v_{\mu^{-1}} = 0$. Observe that $\mathbb{R}^{2\aleph} = E_\mu \oplus E_{\mu^{-1}} \oplus_{i=3}^{2\aleph} E_{\mu_i}$, with $\mu_i \notin \{\mu, \mu^{-1}\}$, and $\dim E_{\mu_i} = 1$ for $i = 3, \dots, 2\aleph$. Note that by (5.9) we have $\langle Jv_{\mu_i}, v_\mu \rangle = 0$ and $\langle Jv_{\mu_i}, v_{\mu^{-1}} \rangle = 0$ for $v_{\mu_i} \in E_{\mu_i}$ and $i = 3, \dots, 2\aleph$. We then set

$$B = \tilde{A} + (\cos\theta - \mu)v_\mu v_\mu^t \Pi_\mu + (\cos\theta - \mu^{-1})v_{\mu^{-1}} v_{\mu^{-1}}^t \Pi_{\mu^{-1}} - \sin(\theta)v_\mu v_{\mu^{-1}}^t \Pi_{\mu^{-1}} + \sin(\theta)v_{\mu^{-1}} v_\mu^t \Pi_\mu.$$

One can check as before that $B \in \mathrm{Sp}(2\aleph)$. Then, $B \in \mathcal{D}(\theta)$ since $(B - e^{i\theta})(v_\mu - iv_{\mu^{-1}}) = 0$. Finally,

$$\|B - \tilde{A}\| \le 2\sqrt{|\cos\theta - \mu|^2 + \sin^2\theta} + 2\sqrt{|\cos\theta - \mu^{-1}|^2 + \sin^2\theta} = |\mu - e^{i\theta}| + |\mu^{-1} - e^{i\theta}|$$
$$\le 2(1 + \varepsilon)(1 + (a - \varepsilon)^{-1})|\lambda - e^{i\theta}|.$$

The bound on (5.6) then follows from (5.8) and (5.7). $\qquad\square$

We proceed to study the covering number for $\mathcal{D}(\theta, s)$.

**Lemma 5.10.** *There exists $\varepsilon_0 > 0$ so that if $r_0 > 0$, $r \ge 1$, $0 < s \le \varepsilon_0 r$, $\theta \in \mathbb{R}$, and $\mathbf{A_0} \in \mathrm{Sp}(2\aleph)$, then*

$$\mathcal{M}_{\mathrm{Sp}(2\aleph)}\big(\mathcal{D}(\theta, s) \cap B_\mathbf{0}(r) \cap B_{\mathbf{A_0}}(r_0), \; s\big) \le \mathcal{M}_{\mathbb{M}(2\aleph)}\big(\mathcal{D}(\theta) \cap B_\mathbf{0}(r) \cap B_{\mathbf{A_0}}(r_0), \; \tfrac{1}{12}sr^{-2}\big).$$

*Proof.* Let $r_0 > 0$, $r > 0$, $s > 0$, $\theta \in \mathbb{R}$, $\mathbf{A_0} \in \mathrm{Sp}(2\aleph)$. By Lemma 5.6 there is $\varepsilon_0 > 0$ such that

$$\mathcal{M}_{\mathrm{Sp}(2\aleph)}(\mathcal{D}(\theta, s) \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0), s) \leq \mathcal{M}_{\mathbb{M}(2\aleph)}(\mathcal{D}(\theta, s) \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0), \tfrac{s}{4r^2}),$$

for $0 < s \leq \varepsilon_0 r$. We claim that for all $\tilde{s} > 0$, $0 < s' < \frac{1}{2}$

$$\mathcal{M}_{\mathbb{M}(2\aleph)}(\mathcal{D}(\theta, s') \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0), 6s' + \tilde{s}) \leq \mathcal{M}_{\mathbb{M}(2\aleph)}(\mathcal{D}(\theta) \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0), \tilde{s}). \qquad (5.10)$$

Let $\{\mathbf{A_i}\}_{\mathbf{i=1}}^{\mathbf{N}} \subset \mathbb{M}(2\aleph)$ be so that $\mathcal{D}(\theta) \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0) \subset \bigcup_{i=1}^{N} B_{\mathbb{M}(2\aleph)}(\mathbf{A}_i, \tilde{s})$. Suppose $\mathbf{A} \in \mathcal{D}(\theta, s') \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0)$. Then, there exists $\lambda \in \mathrm{spec}\,\mathbf{A}$ with $|\lambda - e^{i\theta}| \leq s'$. In particular, by Lemma 5.9 $d(\mathbf{A}, \mathcal{D}(\theta)) \leq 6s'$, so

$$\left(\mathcal{D}(\theta, s') \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0)\right) \subset B_{\mathbb{M}(2\aleph)}(\mathcal{D}(\theta) \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0),\ 6s') \subset \bigcup_{i=1}^{N} B_{\mathbb{M}(2\aleph)}(\mathbf{A}_i, 6s' + \tilde{s}).$$

The result follows from (5.10) after setting $s' = \frac{s}{36r^2}$ and $\tilde{s} = \frac{s}{12r^2}$, which yield $\tilde{s} + 6s' = \frac{1}{4}sr^{-2}$, and observing that, since $r \geq 1$ and $s \leq \varepsilon_0 r$, shrinking $\varepsilon_0$ if necessary $s' \leq \frac{\varepsilon_0 r}{36r^2} \leq \frac{1}{2}$. $\qquad \square$

Next, we use the results above to control the covering numbers of $(\beta, q)$-degenerate matrices. Recall that by Corollary 5.3, if $\boldsymbol{A}$ is not $(\beta, q)$ non-degenerate, then $\mathbf{A} \in \bigcup_{p=0}^{q-1} \mathcal{D}(\frac{2\pi p}{q}, \beta q^{-3})$. In particular, the following result bounds the covering number for such matrices.

**Corollary 5.11.** *There exists $\varepsilon_0 > 0$ $C > 0$ such that for all $\mathbf{A_0} \in \mathrm{Sp}(2\aleph)$, $q \in \mathbb{N}$, $r_0 > 0$, $r \geq 1$, $0 < s \leq \min(\varepsilon_0 r, 12r^2 r_0)$,*

$$\mathcal{M}_{\mathrm{Sp}(2\aleph)}\left(\bigcup_{p=0}^{q-1} \mathcal{D}(\tfrac{2\pi p}{q}, s) \cap B_{\mathbf{0}}(r) \cap B_{\mathbf{A_0}}(r_0),\ s\right) \leq Cq\,(r^2 r_0/s)^{\mathfrak{L} - 2\aleph - 1}.$$

*Proof.* The Corollary follows from combining Lemmas 5.7 and 5.10. $\qquad \square$

## 6. Returning points, simple points, and their iterates

In Section 6.1, we introduce a version of returning points with discretized time as well as the notion of a simple point. Next, in Section 6.2 we introduce the concept of a well separated set for the geodesic flow and the corresponding chain of symplectomorphisms. These concepts are replacements for, respectively, a global Poincaré section and the Poincaré map associated to the global section and, while they require some technical work, do not substantially change the main idea of the proof. Because of this, the reader may first wish to replace the concepts of chains of symplectomorphisms and well separated sets by the simpler notions of Poincaré map and section.

In the next sections of this paper, we will be varying the metric $g$. Because of this, it will be useful to have a single space on which the geodesic flow for any $g$ is defined. This space will be canonically isomorphic to $S^*M$ for any $g$ and will be defined as follows. Let

$$\widetilde{S^*M} := (T^*M \setminus \{0\})/\sim, \qquad (x, \lambda\xi) \sim (x, \xi) \text{ for all } \lambda > 0.$$

Then, since the geodesic flow (the Hamiltonian flow for $p(x, \xi) = |\xi|_g$), is homogeneous of degree 0 in $\xi$, the flow $\varphi_t^g$ passes naturally to the quotient $\widetilde{S^*M}$. We also endow $\widetilde{S^*M}$ with the distance inherited from $S_{g_f}^* M$ for some fixed reference metric $g_f$.

6.1. **Returning points and simple points.** In this section, we define the notions of returning trajectories of length $n$ and simple trajectories of length $n$, and show that non-simple returning points can be seen as iterates of shorter returning trajectories. Because we will pass from the continuous time flow to a discrete time object below, these notions will be defined relative to a fixed number $F$ such that we can guarantee there are no periodic (or near periodic) trajectories of length $< F$. In Section 6.2 we will also insist that the flow can be effectively reduced to a discrete time map by cutting in time at intervals of length $\sim F$. We will be working with metrics in small balls $B \subset \mathscr{G}^3$ and will show in Section 10 that one can take $F$ to be a small multiple of the injectivity radius for a fixed metric in $B$.

The notion of a returning point generalizes that of periodicity.

**Definition 6.1.** For $F > 0$, $g \in \mathscr{G}^2$, $n \in \mathbb{N}$, and $\delta > 0$, we write

$$\rho \in \mathscr{R}_F(n, \delta, g) \qquad \text{if} \qquad \inf_{\substack{(n-1)F < t \leq nF \\ t > F/4}} d(\varphi_t^g(\rho), \rho) < \delta.$$

In this case we say that $\rho$ is $F\text{-}(n, \delta, g)$ returning.

It will be important to have a notion of a simple returning point of length $n$, i.e. one which is not returning for any smaller $n$.

**Definition 6.2.** For $F > 0$, $g \in \mathscr{G}^2$, $n \in \mathbb{N}$, and $\alpha > 0$, we write

$$\rho \in \mathscr{S}_F(n, \alpha, g) \qquad \text{if} \qquad \inf_{\frac{1}{2}F < t < (n-\frac{1}{2})F} d(\varphi_t^g(\rho), \rho) > \alpha.$$

In this case we say that $\rho$ is $F\text{-}(n, \alpha, g)$ simple.

The importance of simple points comes from the fact that the effect of perturbations of $g$ on simple trajectories can be understood. This is much harder to do when considering non-simple trajectories because the trajectory will interact with a given perturbation many times.

The following two lemmas are similar to [Yom85, Lemma 3.1] and show that non-simple returning trajectories are multiples of shorter returning trajectories.

**Lemma 6.3.** *Let $K \subset \mathscr{G}^3(\Gamma)$ be bounded and $F > 0$. Then there is $C > 0$ such that the following holds. Let $g \in K$, $\alpha > 0$, and $\rho \in \widetilde{S^*M}$ with $d(\varphi_t^g(\rho), \rho) < \alpha$. If $d(\varphi_s^g(\rho), \rho) < \alpha$ where $\frac{F}{2} \leq |s| < |t|$ and $q \in \mathbb{Z}$ such that $t = qs + r$ with $|r| < |s|$, then*

$$d(\varphi_r^g(\rho), \rho) < C^{|t|}\alpha.$$

*Proof.* By (3.2) there is $C = C(K) \geq 1$ such that, for all $m \in \mathbb{Z}$

$$d(\varphi_{ms}^g(\rho), \varphi_{(m-1)s}^g(\rho)) \leq C^{|s|} d(\varphi_{(m-1)s}^g(\rho), \varphi_{(m-2)s}^g(\rho)).$$

Using that $d(\varphi_s^g(\rho), \rho) < \alpha$ and $|s| \geq \frac{F}{2} > 0$, we have $d(\varphi_{ms}^g(\rho), \rho) \leq \sum_{j=0}^{m-1} C^{|js|}\alpha \leq 2C^{|ms|}\alpha$. Similarly, since $d(\varphi_t^g(\rho), \rho) \leq \alpha$, we have $d(\varphi_{t-qs}^g(\rho), \varphi_{-qs}^g(\rho)) \leq C^{|qs|}\alpha$. It follows that

$$d(\varphi_r^g(\rho), \rho) \leq d(\varphi_{t-qs}^g(\rho), \varphi_{-qs}^g(\rho)) + d(\varphi_{-qs}^g(\rho), \rho) \leq 3C^{|qs|}\alpha \leq C^{|t|}\alpha,$$

where in the last inequality we used that $|qs| = |t - r| \leq |t| + |r| \leq 2|t|$. $\qquad \square$

**Lemma 6.4.** *Let $K \subset \mathcal{G}^3$ be bounded and $F > 0$. Suppose there are $c > 0$, $C_F > 0$ such that for all $g \in K$ and $\rho \in \widetilde{S^*M}$ we have $d(\varphi_t^g(\rho), \rho) \geq c|t|$ if $|t| \leq C_F F$. Then there is $C > 0$ such that for all $g \in K$, $\delta > 0$, $t_0 > F$, $\alpha \geq \delta$, and $\rho \in \widetilde{S^*M}$ satisfying*

$$d(\varphi_{t_0}^g(\rho), \rho) < \delta,$$

*one of the following two statements hold:*

*(1) $d(\varphi_s^g(\rho), \rho) \geq \alpha$ for all $\frac{F}{2} \leq s \leq t_0 - \frac{F}{2}$.*

*(2) there is $s_0 \in [C_F F - C^{t_0}\alpha, \frac{t_0}{2}]$ such that $d(\varphi_{s_0}^g(\rho), \rho) \leq C^{t_0}\alpha$ and there is $q \in \mathbb{N}$, $q > 1$, such that $q s_0 = t_0$. Moreover, if $t_1 = q_1 s_0 + \mathfrak{r} \leq t_0$ for $q_1 \in \mathbb{N}$ and $0 \leq \mathfrak{r} < s_0$, then*

$$d(\varphi_{t_1}^g(\rho), \varphi_{\mathfrak{r}}^g(\rho)) \leq C^{t_0}\alpha.$$

*Proof.* Suppose that statement (1) is false. Then there is $s$ such that $\frac{F}{2} \leq s \leq t_0 - \frac{F}{2}$ and $d(\varphi_s^g(\rho), \rho) < \alpha$. Set $t_1 = s$ and let $q_1 \in \mathbb{Z}$ such that $t_0 = q_1 t_1 + t_2$ for $|t_2| \leq \frac{s}{2}$. Then, by Lemma 6.3, $d(\varphi_{t_2}^g(\rho), \rho) \leq C^{t_0}\alpha$. If $|t_2| \leq C_F F$, we put $\tilde{s}_0 = t_1$.

If $|t_2| > C_F F$, we continue the process and suppose we have found $\{q_i\}_{i=1}^{m+1} \subset \mathbb{Z}$, and $\{t_i\}_{i=0}^{m+2}$ such that for $i = 2, \ldots, m$,

$$t_i = q_{i+1}t_{i+1} + t_{i+2}, \qquad |t_{i+2}| \leq \tfrac{1}{2}|t_{i+1}|, \qquad |t_i| > C_F F, \qquad d(\varphi_{t_i}^g(\rho), \rho) \leq C^{\sum_{j=0}^{i-2} |t_j|}\alpha.$$

Then, letting $q_{m+2} \in \mathbb{Z}$, $|t_{m+3}| \leq \frac{1}{2}|t_{m+2}|$, such that $t_{m+1} = q_{m+2}t_{m+2} + t_{m+3}$, Lemma 6.3 yields $d(\varphi_{t_{m+3}}^g(\rho), \rho) \leq C^{\sum_{j=0}^{m+1} |t_j|}\alpha$. Then, if $|t_{m+3}| \leq C_F F$, we set $\tilde{s}_0 = t_{m+2}$. In particular, since $|t_j| \leq \frac{1}{2}|t_{j-1}|$ for $j = 2, \ldots$, we have $\sum_{j=0}^{\infty} |t_j| \leq 3|t_0|$. Therefore, we have found $\tilde{s}_0$ such that

$$d(\varphi_{\tilde{s}_0}^g(\rho), \rho) \leq C^{3|t_0|}\alpha.$$

Let this process terminate with $|t_{m+3}| \leq C_F F$. We claim that $|t_{m+3}| \leq C^{4|t_0|}\alpha$. Indeed, since $d(\varphi_t^g(\rho), \rho) \geq c|t|$ for $|t| \leq C_F F$ and $|t_{m+3}| \leq C_F F$, by Lemma 6.3 we know

$$c|t_{m+3}| \leq d(\varphi_{t_{m+3}}^g(\rho), \rho) \leq C^{3|t_0|}\alpha. \tag{6.1}$$

Finally, set $s_0 = \tilde{s}_0 + \frac{t_{m+3}}{q_{m+2}}$. Then, after possibly modifying $C$, and using the bound on $|t_{m+3}|$ from (6.1) we conclude

$$d(\varphi_{s_0}^g(\rho), \rho) \leq C^{|t_{m+3}|}d(\varphi_{\tilde{s}_0}^g(\rho), \rho) + d(\varphi_{\underbrace{\frac{t_{m+3}}{q_{m+2}}}}^g(\rho), \rho) \leq 2C^{4t_0}\alpha$$

as claimed, since $|q_{m+2}| \geq 1$. Note that, since $t_{m+1} = q_{m+2}t_{m+2} + t_{m+3} = q_{m+2}(t_{m+2} + \frac{t_{m+3}}{q_{m+2}}) = q_{m+2}s_0$, we have by construction $t_0 = qs_0$ for some $q \in \mathbb{Z}$. In addition,

$$|s_0| \geq |t_{m+2}| - |\tfrac{t_{m+3}}{q_{m+2}}| \geq C_F F - c^{-1}C^{4|t_0|}\alpha.$$

If $s_0 > 0$, we have the claimed properties of $s_0$. If, on the other hand, $s_0 < 0$, we have, modifying $C$ if necessary,

$$d(\varphi_{-s_0}^g(\rho), \rho) \leq C^{|s_0|}d(\rho, \varphi_{s_0}^g(\rho)) \leq C^{4|t_0|}\alpha.$$

Finally, observe that for $qs_0 \leq t_0$,

$$d(\varphi_{qs_0}^g(\rho), \rho) \leq \sum_{j=0}^{q-1} d(\varphi_{js_0}^g(\rho), \varphi_{(j+1)s_0}^g(\rho)) \leq \sum_{j=0}^{q-1} C^{j|s_0|} d(\varphi_{s_0}^g(\rho), \rho) \leq C^{|t_0|}\alpha.$$

Therefore, if $t_1 = qs_0 + \mathfrak{r} \leq t_0$ for some $q \in \mathbb{N}$, $0 \leq \mathfrak{r} < s_0$

$$d(\varphi_{t_1}^g(\rho), \varphi_{\mathfrak{r}}^g(\rho)) \leq C^{\mathfrak{r}} d(\varphi_{qs_0}^g(\rho), \rho) \leq C^{|t_0|}\alpha.$$

$\square$

6.2. **Non-degeneracy and chains of symplectomorphisms.** In general, it is not possible to find a global Poincaré section for the geodesic flow; i.e. a connected, closed submanifold, $\Gamma$, of co-dimension 1 in $\widetilde{S^*M}$ that is everywhere transverse to $H_{|\xi|_g}$ such that every geodesic passes through $\Gamma$. Because of this, we need to replace the notion of a global Poincaré section and its associated Poincaré map with a more complicated submanifold which captures all the dynamical information for $\varphi_t^g$. To this end, we introduce below, the notion of a well separated set and its associated chains of symplectomorpshims; replacing, respectively, the global Poincaré section and its associated Poincaré map. The idea of a well-separated set will be to find, $\tilde{\Gamma}$, a disjoint union of open $2d-2$ dimensional submanifolds, and a compactly embedded open subset, $\Gamma \subset \tilde{\Gamma}$, such that 1) every geodesic passes through $\Gamma$ in a controlled time and 2) no geodesic passes in a very short time from $\tilde{\Gamma}$ to itself. These properties will guarantee that understanding geodesics which pass from $\Gamma$ to itself allows one to understand the full dynamics for the geodesic flow.

To understand the need for $\tilde{\Gamma}$ note that since the connected components of $\Gamma$ have boundaries, there are points $\rho \in \Gamma$ such that small perturbations of $\rho$ will cause large jumps in the first impact point $\mathscr{P}_\Gamma(\rho)$ of the geodesic through $\rho$ with $\Gamma$; i.e there are sequences $\rho_n \to \rho$ such that $d(\mathscr{P}_\Gamma(\rho_n), \mathscr{P}_\Gamma(\rho)) \geq c > 0$. This forces us to include the slightly larger $\tilde{\Gamma}$ and introduce the notion of Poincaré chains below. (See Figure 3 for a schematic of a well-separated set and a Poincaré chain.)

Let $\mathcal{W}_i \subset \widetilde{S^*M}$, $i = 1, \ldots, N$ be disjoint, open symplectic submanifolds of dimension $2d-2 = 2\aleph$ and $\mathcal{V}_i \Subset \mathcal{W}_i$ open subsets. We write

$$\Gamma := \bigsqcup_{i=1}^N \mathcal{V}_i, \qquad \widetilde{\Gamma} := \bigsqcup_{i=1}^N \mathcal{W}_i. \tag{6.2}$$

**Definition 6.5.** For $F > 0$, and $G \subset \mathscr{G}^2$ bounded, we say that $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ are $F$-*well-separated for $G$* if the $\mathcal{W}_i$ are uniformly transverse to $H_{|\xi|_g}$ for all $g \in G$ and there are $c_\Gamma, C_\Gamma > 1$ such that

$$\sup_{g \in G} \sup_{\rho \in \widetilde{S^*M}} \inf\{t > 0 : \varphi_t^g(\rho) \in \Gamma\} < C_\Gamma F, \qquad \inf_{g \in G} \inf_{\rho \in \tilde{\Gamma}} \inf\{t > 0 : \varphi_t^g(\rho) \in \tilde{\Gamma}\} \geq c_\Gamma F,$$

$$\inf_{g \in G} \min_i \inf_{\rho \in \mathcal{W}_i} \inf\{t > 0 : \varphi_t^g(\rho) \in \overline{\mathcal{W}_i}\} > C_\Gamma F. \tag{6.3}$$

When it will not lead to confusion, we will say that $\Gamma$ *is $F$-well-separated*.

In words, if $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ is $F$-well-separated for $G$, it means that (1) every trajectory hits $\Gamma$ in time less than $C_\Gamma F$; (2) trajectories starting from $\tilde{\Gamma}$ always take at least $c_\Gamma F$ time to return to

$\tilde{\Gamma}$; (3) trajectories that start within some $\mathcal{W}_i$ always take at least $C_\Gamma F$ to return to $\overline{\mathcal{W}_i}$ no matter the choice of $i$.

Before we move on to the definition of chains of symplectomorphisms associated to well separated sets, we show that Lemma 6.4 applies when there is an $F$-well separated set.

**Lemma 6.6.** *Let $G \subset \mathscr{G}^2$ bounded. Suppose that $F > 0$ and $\Gamma$ is $F$-well separated for $G$. Then, for all $K \subset G$ bounded in $\mathscr{G}^3$, there is $c > 0$ such that for all $g \in K$ and $|t| \leq C_\Gamma F$,*

$$d(\varphi_t^g(\rho), \rho) \geq c|t|, \qquad \rho \in \widetilde{S^*M}.$$

*Proof.* First, notice that, since $K$ is bounded in $\mathscr{G}^3$, there are $c_1, c_2 > 0$ such that, $d(\varphi_t^g(\rho), \rho) \geq c_2|t|$ for all $g \in K$, $\rho \in \widetilde{S^*M}$, and $|t| < c_1$. Thus, it remains only to check that there is $c > 0$ such that $d(\varphi_t^g(\rho), \rho)) \geq c$ for all $g \in K$, $\rho \in \widetilde{S^*M}$, and $c_1 \leq |t| \leq C_F F$.

Suppose by contradiction there are $g_n \in K$, $\rho_n \in \widetilde{S^*M}$, and $t_n$ with $|t_n| \in [c_1, C_F F]$, such that

$$d(\varphi_{t_n}^{g_n}(\rho_n), \rho_n) \to 0.$$

Then, without loss of generality, we may assume $g_n \overset{\mathscr{G}^2}{\to} g \in G$, $\rho_n \to \rho \in \widetilde{S^*M}$, and $t_n \to t$ with $|t| \in [c_1, C_F F]$. Now,

$$d(\varphi_t^g(\rho), \rho) \leq d(\varphi_t^g(\rho), \varphi_t^{g_n}(\rho)) + d(\varphi_t^{g_n}(\rho_n), \varphi_t^{g_n}(\rho)) + d(\varphi_{t_n}^{g_n}(\rho_n), \varphi_t^{g_n}(\rho_n)) + d(\varphi_{t_n}^{g_n}(\rho_n), \rho_n) + d(\rho_n, \rho).$$

Since $g_n \to g$ in $\mathscr{G}^2$, the right hand side of the above inequality tends to 0 as $n \to \infty$, and we have $\varphi_t^g(\rho) = \rho$. Let $T_0 := \inf\{s > 0 \,:\, \varphi_s^g(\rho) \in \Gamma\}$ and $\rho_\Gamma := \varphi_{T_0}^g(\rho)$ (note that $0 < T_0 < \infty$ by the first inequality in (6.3)). Then, $\varphi_t^g(\rho_\Gamma) = \varphi_{t+T_0}^g(\rho) = \varphi_{T_0}^g(\rho) = \rho_\Gamma$, with $c_1 \leq |t| \leq C_F F$. This contradicts the last inequality in (6.3). $\qquad \square$

We next use well-separated sets to reduce the continuous flow $\varphi_t^g$ to a discrete time system. Let $\nu \geq 3$, $F > 0$, $G \subset \mathscr{G}^\nu$ bounded,, and $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ be $F$-well-separated for $G$. Then define

$$T_{j,i}^g : \mathcal{W}_i \to \mathbb{R}, \qquad T_{j,i}^g(\rho) := \inf\{t > 0 \,:\, \varphi_t^g(\rho) \in \mathcal{W}_j\}, \tag{6.4}$$

and, with $\mathcal{U}_{j,i}^g := (T_{j,i}^g)^{-1}((0, C_\Gamma F)) \subset \mathcal{W}_i$, let $\kappa_{j,i}^g : \mathcal{U}_{j,i}^g \to \mathcal{W}_j$ be the function

$$\kappa_{j,i}^g(\rho) = \varphi_{T_{j,i}^g(\rho)}^g(\rho).$$

**Lemma 6.7.** *For all $g \in G$, $j, i$, the map $\kappa_{j,i}^g : \mathcal{U}_{j,i}^g \to \mathcal{W}_j$ is a $\mathcal{C}^{\nu-2}$ symplectomorphism onto its image.*

*Proof.* Once we show that $\kappa_{j,i}^g$ is $\mathcal{C}^{\nu-2}$, the fact that $\kappa_{j,i}^g$ is a symplectomorphism onto its image is inherited from the facts that $\varphi_t^g$ is a symplectomorphism and that $\mathcal{W}_j$ is transverse to $H_{|\xi|_g}$.

We now show that $\kappa_{j,i}^g$ is $\mathcal{C}^{\nu-2}$. To do this, note that $\mathcal{W}_j$ is open and transverse to $H_{|\xi|_g}$. Therefore, since $(t, \rho) \mapsto \phi_t^g(\rho)$ is $\mathcal{C}^{\nu-2}$ with $\nu - 2 \geq 1$ and $\varphi_{T_{j,i}^g(\rho)}^g(\rho) \in \mathcal{W}_j$, the implicit function theorem implies that there is a neighborhood $V \subset \mathcal{W}_i$ of $\rho$ and a $\mathcal{C}^{\nu-2}$ function $\tilde{T}_{j,i}^g(\rho') : V \to \mathbb{R}$ such that $\varphi_{\tilde{T}_{j,i}^g(\rho')}^g(\rho') \in \mathcal{W}_j$ and $\tilde{T}_{j,i}^g(\rho) = T_{j,i}^g(\rho)$. Since $\tilde{T}_{j,i}^g$ is continuous and $T_{j,i}^g(\rho) < C_\Gamma F$, shrinking $V$ if necessary, we may assume $\tilde{T}_{j,i}^g(\rho') < C_\Gamma F$. Finally, we need to check that $\tilde{T}_{j,i}^g(\rho') =$

$T^g_{j,i}(\rho')$. For this, it suffices to show that $\varphi^g_s(\rho') \notin \mathcal{W}_j$ for $0 < s < \tilde{T}^g_{j,i}(\rho')$. Indeed, suppose that there is $0 < s < \tilde{T}^g_{j,i}(\rho')$ such that $\varphi^g_s(\rho') \in \mathcal{W}_j$. Then, $\varphi^g_{\tilde{T}^g_{j,i}(\rho')-s}(\varphi^g_s(\rho')) \in \mathcal{W}_j$, and $0 < \tilde{T}^g_{j,i}(\rho') - s < C_\Gamma F$. This contradicts the third part of (6.3). $\qquad\square$

**Definition 6.8.** Let $\nu \geq 3$. Let $G \subset \mathscr{G}^\nu$ be bounded in $\mathscr{G}^2$ and $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ be $F$-well separated for $G$. For $\mathcal{I} \in \{1, \ldots, N\}^{\mathbb{N}}$, $\mathcal{I} = (i_0, i_1, i_2, \ldots)$, $g \in G$, and $n \in \mathbb{N}$ we define the $C^{\nu-2}$ symplectomorphism

$$\mathcal{P}^{(n)}_{\mathcal{I}}[g] : \mathcal{D}^{(n)}_{\mathcal{I}}[g] \to \mathcal{W}_{i_n}, \qquad \mathcal{P}^{(n)}_{\mathcal{I}}[g] := \kappa^g_{i_n, i_{n-1}} \circ \ldots \kappa^g_{i_2, i_1} \circ \kappa^g_{i_1, i_0},$$

where the $\mathcal{D}^{(n)}_{\mathcal{I}}[g] \subset \mathcal{W}_{i_0}$ is the domain of the composition. We also define $\mathcal{P}^{(0)}_{\mathcal{I}}[g] = I$ and define

$$\mathcal{P}_{\mathcal{I}}[g] := (\mathcal{P}^{(0)}_{\mathcal{I}}[g], \mathcal{P}^{(1)}_{\mathcal{I}}[g], \ldots)$$

and call it the *Poincaré chain associated to $\mathcal{I}$*.

**Definition 6.9.** We define the time sequence $\boldsymbol{T}_{\mathcal{I}}$ associated to the chain $\mathcal{P}_{\mathcal{I}}$. For each $n \in \mathbb{N}$ and $g \in G$ let $\boldsymbol{T}^{(n)}_{\mathcal{I}}[g] : \mathcal{D}^{(n)}_{\mathcal{I}}[g] \to [0, \infty)$ be defined by $\boldsymbol{T}^{(0)}_{\mathcal{I}}[g] \equiv 0$ and for all $\rho \in \mathcal{D}^{(n)}_{\mathcal{I}}[g]$

$$\boldsymbol{T}^{(n)}_{\mathcal{I}}[g](\rho) = \boldsymbol{T}^{(n-1)}_{\mathcal{I}}[g](\rho) + T^g_{i_n, i_{n-1}}(\mathcal{P}^{(n-1)}_{\mathcal{I}}[g](\rho)).$$

for $T^g_{i_n, i_{n-1}}$ as defined in (6.4).

By construction, for $\rho \in \mathcal{D}^{(n)}_{\mathcal{I}}[g]$ we have

$$\mathcal{P}^{(n)}_{\mathcal{I}}[g](\rho) = \varphi^g_{\boldsymbol{T}^{(n)}_{\mathcal{I}}[g](\rho)}(\rho). \tag{6.5}$$

**Remark 6.10.** If $\rho \in \tilde{\Gamma}$ and $\varphi^g_{t_0}(\rho) \in \tilde{\Gamma}$, then there are $n$ and $\mathcal{I}$ such that $\rho \in \mathcal{D}^{(n)}_{\mathcal{I}}[g]$, $\mathcal{P}^{(n)}_{\mathcal{I}}[g](\rho) = \varphi^g_{t_0}(\rho)$ and $\boldsymbol{T}^{(n)}_{\mathcal{I}}[g](\rho) = t_0$.

Below, we will be varying the metric and need to control how Poincaré chains vary with the metric.

**Lemma 6.11.** *Let $\nu \geq 2$, $0 \leq k \leq \nu - 2$, $G \subset \mathscr{G}^\nu$ bounded, and $g : B_{\mathbb{R}^k}(0,1)_\sigma \to G$ a $\mathcal{C}^2$ map. Then there are $C > 0$, $C_k > 0$, such that*

$$d_\sigma \varphi^g_t : \mathbb{R}^k \to T\mathcal{C}^{\nu-3}(\widetilde{S^*M}; \widetilde{S^*M}), \qquad d^2_\sigma \varphi^g_t : \mathbb{R}^k \times \mathbb{R}^k \to T\mathcal{C}^{\nu-4}(\widetilde{S^*M}; \widetilde{S^*M})$$

*are well defined and satisfy that for all $t \in \mathbb{R}$, $v, w \in \mathbb{R}^k$, $f \in \mathcal{C}^k(S^*M; \mathbb{R})$,*

- $\|d_\sigma \varphi^g_t|_{\sigma=0} v\|_{\mathcal{C}^{\nu-3}} \leq C^{|t|} \|v\| \|\partial_\sigma g|_{\sigma=0}\|_{\mathcal{C}^{\nu-1}}$,      *if $\nu \geq 4$*
- $\|d^2_\sigma \varphi^g_t|_{\sigma=0}(v, w)\|_{\mathcal{C}^{\nu-4}} \leq C^{|t|} \|v\| \|w\| (\|\partial^2_\sigma g|_{\sigma=0}\|_{\mathcal{C}^{\nu-2}} + \|\partial_\sigma g|_{\sigma=0}\|^2_{\mathcal{C}^{\nu-1}})$,      *if $\nu \geq 5$*
- $\|f \circ \varphi^g_t\|_{\mathcal{C}^k} \leq C_k C^{|k||t|} \|f\|_{\mathcal{C}^k}$.

*Moreover, for all $n \in \mathbb{N}$,*

- $\|d_\sigma \mathcal{P}^{(n)}_{\mathcal{I}}[g_\sigma] v\|_{\mathcal{C}^{\nu-3}} \leq C^{n C_\Gamma F} \|v\| \|\partial_\sigma g|_{\sigma=0}\|_{\mathcal{C}^{\nu-1}}$,      *if $\nu \geq 4$*
- $\|d^2_\sigma \mathcal{P}^{(n)}_{\mathcal{I}}[g_\sigma](v, w)\|_{\mathcal{C}^{\nu-4}} \leq C^{n C_\Gamma F} \|v\| \|w\| (\|\partial^2_\sigma g|_{\sigma=0}\|_{\mathcal{C}^{\nu-2}} + \|\partial_\sigma g|_{\sigma=0}\|^2_{\mathcal{C}^{\nu-1}})$,      *if $\nu \geq 5$*
- $\|f \circ \mathcal{P}^{(n)}_{\mathcal{I}}[g]\|_{\mathcal{C}^k} \leq C_k C^{|k| n C_\Gamma F} \|f\|_{\mathcal{C}^k}$.

*Proof.* The proof of this lemma is a tedious calculation. We sketch here the proof.

Recall that the geodesics flow can be written as the solution $(x(t), \xi(t))$ to

$$\dot{x}^j(t) = \frac{1}{|\xi(t)|_{g(x(t))}} g^{ij}(x(t))\xi_i(t), \qquad \dot{\xi}_k(t) = -\frac{1}{2|\xi(t)|_{g(x(t))}} \partial_{x^k} g^{ij}(x(t))\xi_i(t)\xi_j(t).$$

Letting $g_\sigma$ be a family of metrics with $g_0 = g$ and differentiating with respect to $\sigma$, working always with initial data so that $|\xi_0(g_\sigma)|_{g_\sigma(x_0(g_\sigma))} = 1$, we obtain equations of the form

$$\partial_\sigma \dot{x} = (\partial_x g\xi)\partial_\sigma x + g\partial_\sigma \xi + \partial_\sigma g\xi, \qquad \partial_\sigma \dot{\xi} = (\partial_x^2 g)\xi^2 \partial_\sigma x + \partial_x g\xi \partial_\sigma \xi + (\partial_{x\sigma}^2 g)\xi^2.$$

In particular, provided that $\partial_\sigma g \in \mathcal{C}^{\nu-1}$, this is an equation with $\mathcal{C}^{\nu-2}$ coefficients (bounded by the $\mathcal{C}^\nu$ norm of $g$ and the $\mathcal{C}^{\nu-1}$ norm of $h$) and hence results in a $\mathcal{C}^{\nu-3}$ flow with the claimed bounds.

Differentiating again in $g$, we obtain equations of the form

$$\partial_\sigma^2 \dot{x} = (\partial_x^2 g\xi)(\partial_\sigma^2 x) + g\partial_\sigma^2 \xi + (\partial_x^2 g\xi)(\partial_\sigma x)^2 + \partial_x g\partial_\sigma x\partial_\sigma \xi + \partial_{x\sigma}^2 g\xi \partial_\sigma x + \partial_\sigma g\partial_\sigma \xi + \partial_\sigma^2 g\xi$$

$$\partial_\sigma^2 \dot{\xi} = (\partial_x^2 g)\xi^2 \partial_\sigma^2 x + \partial_x g\xi \partial_\sigma^2 \xi + (\partial_x^3 g)\xi^2(\partial_\sigma x)^2 + \partial_{\sigma xx}^3 g\xi^2 \partial_\sigma x + \partial_x^2 g\xi \partial_\sigma \xi \partial_\sigma x$$

$$+ \partial_{x\sigma}^2 g\xi \partial_\sigma \xi + \partial_x g(\partial_\sigma \xi)^2 + \partial_{x\sigma\sigma}^3 g\xi^2.$$

Now, since $\partial_\sigma(x, \xi) \in \mathcal{C}^{\nu-3}$ if $\partial_\sigma g \in \mathcal{C}^{\nu-1}$, we require $\partial_\sigma^2 g \in \mathcal{C}^{\nu-2}$, and $\partial_\sigma g \in \mathcal{C}^{\nu-1}$ to obtain a solution in $\mathcal{C}^{\nu-4}$.

The estimates on the derivative in time now follow from estimating the coefficients in these equations in $L^\infty$ by the relevant $\mathcal{C}^\nu$ norms of the metric and its derivatives. $\qquad \square$

The following lemma shows that if $\rho \in \Gamma$ is a returning point, say $d(\varphi_{t_0}^g(\rho), \rho) < \delta$, then we can associate to it a Poincaré chain under which $\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho) = \varphi_{t_0+s}^g(\rho)$ for some $m \in \mathbb{N}$ and $s$ small. Morover, $d(\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho), \rho) \lesssim \delta$.

**Lemma 6.12.** *Let $G \subset \mathscr{G}^\nu$ be bounded in $\mathscr{G}^2$ and $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ be $F$-well separated for $G$. Then there are $C_G > 0$ and $\delta_G > 0$ such that for all $g \in G$, $0 < \delta < \delta_G$, $j \in \{1, \ldots, N\}$, and $\rho \in \mathscr{R}_F(n, \delta, g) \cap \mathcal{V}_j$ the following holds. There exist $m \in \mathbb{N}$ and $\mathcal{I}$ such that*

$$\rho \in \mathcal{D}_{\mathcal{I}}^{(m)}[g], \qquad \boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho) \in \big[(n-1)F - C_G\delta, \, nF + C_G\delta\big], \tag{6.6}$$

*and $d(\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho), \rho) < C_G\delta$. Moreover, for any $m_1, m_2 \in \mathbb{N}$ and $\mathcal{I}_1, \mathcal{I}_2$ such that*

$$\rho \in \mathcal{D}_{\mathcal{I}_j}^{(m_j)}[g], \qquad \boldsymbol{T}_{\mathcal{I}_j}^{(m_j)}[g](\rho) \in \big[(n-1)F - C_G\delta_G, \, nF + C_G\delta_G\big], \qquad j = 1, 2, \tag{6.7}$$

$$\mathcal{P}_{\mathcal{I}_1}^{(m_1)}[g](\rho) = \mathcal{P}_{\mathcal{I}_2}^{(m_2)}[g](\rho), \qquad \boldsymbol{T}_{\mathcal{I}_1}^{(m_1)}[g](\rho) = \boldsymbol{T}_{\mathcal{I}_2}^{(m_2)}[g](\rho).$$

*Proof.* Since $\tilde{\Gamma}$ is uniformly transverse to $H_{|\xi|_g}$ for $g \in G$, and for all $j$ we have $\mathcal{V}_j \Subset \mathcal{W}_j$, there are $C_G > 0$ and $\delta_G > 0$ such that for all $g \in G$, $j \in \{1, \ldots, N\}$, $0 < \delta < \delta_G$, and $\rho \in \widetilde{S^*M}$ with $d(\rho, \mathcal{V}_j) < \delta$, there is $|s| \leq C_G\delta$ such that

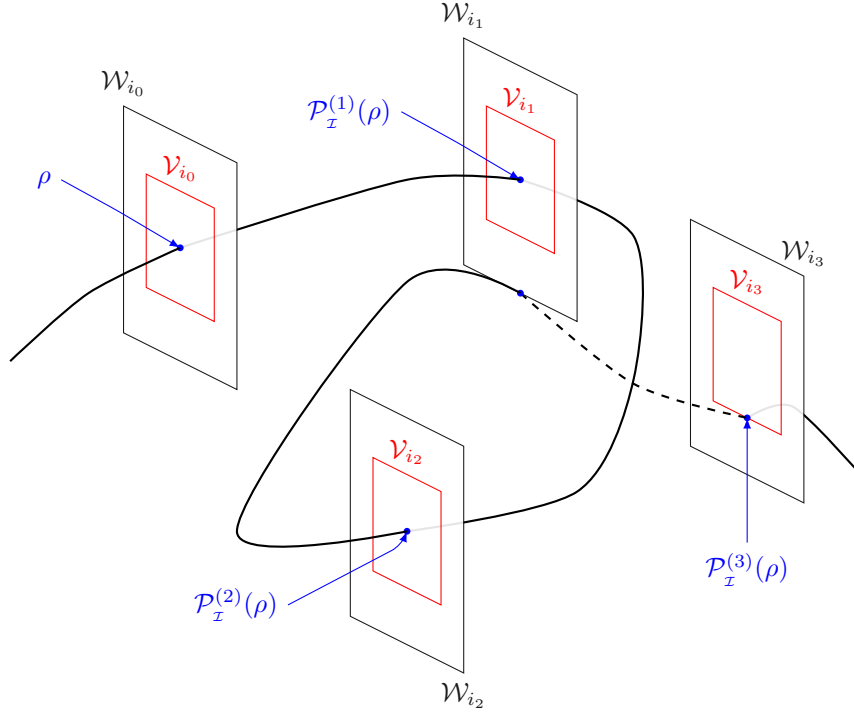$$\varphi_{t_0+s}^g(\rho) \in \mathcal{W}_j. \tag{6.8}$$

FIGURE 3. A portion of a well-separated set and the chain of symplectomorphisms evaluated at a point $\rho \in \Gamma$. The chain pictured begins $\mathcal{I} = \{i_0, i_1, i_2, i_3\}$. Notice that the geodesic encounters the boundary of $\mathcal{W}_{i_1}$ between $\mathcal{W}_{i_2}$ and $\mathcal{W}_{i_3}$. It is because of geodesics like this that we must define a Poincaré chain. Furthermore, this geodesic encounters the boundary of $\mathcal{V}_{i_3}$ and, to make the map $\mathcal{P}_{\mathcal{I}}^{(3)}[g]$ smooth near $\rho$, one must refer to the subset $\tilde{\Gamma}$ in addition to $\Gamma$.

Next, let $\rho \in \mathscr{R}_F(n, \delta, g) \cap \mathcal{V}_j$ with $0 < \delta < \delta_G$. Then, there is $(n-1)F < t_0 \leq nF$ such that $d(\varphi_{t_0}^g(\rho), \rho) < \delta$. Therefore, there is $|s| \leq C_G \delta$ such that $\varphi_{t_0+s}^g(\rho) \in \mathcal{W}_j$. In particular, by Remark 6.10 there are $m$ and $\mathcal{I}$ such that (6.6) holds and $\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho) = \varphi_{t_0+s}^g(\rho)$.

Note that, increasing $C_G$ if necessary (uniformly for $g \in G$),

$$d(\varphi_{t_0+s}^g(\rho), \rho) \leq d(\varphi_{t_0+s}^g(\rho), \varphi_s^g(\rho)) + d(\varphi_s^g(\rho), \rho) < C_G \delta. \tag{6.9}$$

Therefore, it only remains to check that if $(m_1, \mathcal{I}_1)$ are such that (6.7) holds, then $\boldsymbol{T}_{\mathcal{I}_1}^{(m_1)}[g](\rho) = t_0 + s$. Suppose this is not the case for some $(m_1, \mathcal{I}_1)$. Then, since $\varphi_{\boldsymbol{T}_{\mathcal{I}_1}^{(m_1)}[g](\rho)}^g(\rho) \in \tilde{\Gamma}$, by (6.5), the second equation in (6.3) yields $|\boldsymbol{T}_{\mathcal{I}_1}^{(m_1)}[g](\rho) - (t_0 + s)| \geq c_\Gamma F$. In particular, either $\boldsymbol{T}_{\mathcal{I}_1}^{(m_1)}[g](\rho) \geq (n-1)F - C_G \delta + c_\Gamma F$, or $\boldsymbol{T}_{\mathcal{I}_1}^{(m_1)}[g](\rho) \leq nF + C_G \delta - c_\Gamma F$. Since $\delta < \delta_G$, this contradicts (6.7) provided that $2C_G \delta_G < (c_\Gamma - 1)F$. Shrinking $\delta_G$ if necessary completes the proof. 

$\square$

Recall that the image of $\mathcal{P}_{\mathcal{I}}^{(m)}[g_0]$ lands in $\tilde{\Gamma}$ and that $\Gamma \Subset \tilde{\Gamma}$. The following lemma shows that if the iterates of $\rho$ land sufficiently close to $\Gamma$, then there is a neighborhood of $\rho$ that lies in the domain of the same Poincaré chain.

**Lemma 6.13.** *Let $G \subset \mathscr{G}^3$ bounded and $\Gamma$ be $F$ well-separated for $G$. There are $c > 0$, $\delta_0 > 0$ such that the following holds. Let $g_0 \in G$, $\rho \in \Gamma$, $n \in \mathbb{N}$, and $\mathcal{I}$ be such that*

$$\sup_{m \leq n} d(\mathcal{P}_{\mathcal{I}}^{(m)}[g_0](\rho), \Gamma) < \delta_0.$$

*Then, $B_{\tilde{\Gamma}}(\rho, c^n) \subset \mathcal{D}_{\mathcal{I}}^{(m)}[g]$ for all $m \leq n$ provided $\|g - g_0\|_{\mathcal{C}^3} \leq c^n$.*

*Proof.* Let $\delta_0 > 0$ such that $\{\rho \in \tilde{\Gamma} : d(\rho, \Gamma) < 3\delta_0\} \Subset \tilde{\Gamma}$, and

$$\sup_{g \in G} \sup_{\rho \in \widetilde{S^*M}} \inf\{t > 0 : \varphi_t^g(\rho) \in \Gamma\} < C_\Gamma F - 3\delta_0.$$

First, note that there is $C > 0$ such that $d(\varphi_t^g(\rho), \varphi_t^{g_0}(\rho)) \leq \varepsilon C^{t+1}$, for all $g_0 \in G$, $\varepsilon > 0$, $t \in \mathbb{R}$, and $g \in \mathscr{G}^3$ with $\|g - g_0\|_{\mathcal{C}^3} \leq \varepsilon$. In particular, since $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_0](\rho) \leq mC_\Gamma F$ by (6.3),

$$d\left(\varphi_{\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_0](\rho)}^g(\rho), \varphi_{\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_0](\rho)}^{g_0}(\rho)\right) \leq \varepsilon C^{mC_\Gamma F + 1}.$$

Thus, since $d(\varphi_{\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_0](\rho)}^g(\rho), \Gamma) < \delta_0$ and $H_{|\xi|_g}$ is uniformly transverse to $\tilde{\Gamma}$, for $g \in G$, the implicit function theorem implies that for $\varepsilon$ small enough such that $\varepsilon C^{mC_\Gamma F + 3} \ll \delta_0$, there are $t_j$, $j = 1, \ldots, m$ such that $|t_j - \boldsymbol{T}_{\mathcal{I}}^{(j)}[g_0](\rho)| \leq \varepsilon C^{jC_\Gamma F + 2} \leq \delta_0$ and

$$\varphi_{t_j}^{g_0}(\rho) \in \tilde{\Gamma}, \qquad d(\varphi_{t_j}^{g_0}(\rho), \varphi_{\boldsymbol{T}_{\mathcal{I}}^{(j)}[g_0](\rho)}^g(\rho)) \leq \varepsilon C^{jC_\Gamma F + 3} \leq 2\delta_0.$$

In particular, since $\boldsymbol{T}_{\mathcal{I}}^{(j)}[g_0](\rho) - \boldsymbol{T}_{\mathcal{I}}^{(j-1)}[g_0](\rho) < C_\Gamma F - 3\delta_0$ for all $1 \leq j \leq m$, we have $|t_1| < C_\Gamma F$, and $|t_j - t_{j-1}| < C_\Gamma F$ for $1 \leq j \leq m$. Thus, $\rho \in \mathcal{D}_{\mathcal{I}}^{(m)}[g]$ and

$$|\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_0](\rho) - \boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho)| \ll 1,$$

provided $\varepsilon \ll C^{-mC_\Gamma F - 1}$. Thus, we have

$$d(\varphi_{\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho)}^g(\rho), \Gamma) < \delta_0 + \varepsilon C^{mC_\Gamma + 1} < 2\delta_0.$$

Arguing as above, since there is $C > 0$ such that for all $g$ with $d_{\mathcal{C}^3}(g, G) < c$,

$$d(\varphi_t^g(\rho), \varphi_t^g(\rho_1)) < C^t d(\rho, \rho_1),$$

we have for $d(\rho, \rho_1) < \varepsilon$,

$$|\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho_1) - \boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho_2)| \ll 1,$$

and hence $\rho_1 \in \mathcal{D}_{\mathcal{I}}^{(m)}[g]$. $\qquad\square$

We now define the notion of a non-degenerate returning point for a Poincaré chain $\mathcal{P}_{\mathcal{I}}$. Let $G \subset \mathscr{G}^\nu$ be bounded in $\mathscr{G}^3$ and $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ be $F$-well separated for $G$. To define non-degeneracy, we fix a finite atlas of coordinates charts on $\Gamma$

$$\mathcal{A} := \big\{(\psi_i, U_i) : \psi_i : U_i \to \psi_i(U_i), \ i = 1, \ldots, N\big\}, \tag{6.10}$$

with $U_i \subset \Gamma$ and $\psi_i(U_i) \subset \mathbb{R}^{2\aleph}$. Let $C_G$ and $\delta_G$ be as in Lemma 6.12 and let $\delta_0$ be as in Lemma 6.13. Without loss of generality we assume that $\delta_G$ is small enough that

$$\delta_G < \min(1, \delta_0),$$
$$\text{for every } \rho \in \Gamma \text{ there is } (\psi, U) \in \mathcal{A} \text{ such that } B(\rho, C_G \delta_G) \subset U. \tag{6.11}$$

Next, observe that by Lemma 6.12, if $\rho \in \Gamma \cap \mathscr{R}_F(n, \delta, g)$ for $0 < \delta < \delta_G$, then there is $(\psi, U) \in \mathcal{A}$ such that $\rho \in U$ for every $(m, \mathcal{I})$ satisfying (6.6), $\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho) \in U$, and we may work in $\mathbb{R}^{2\aleph}$ by setting

$$(\mathcal{P}_{\mathcal{I}}^{(m)}[g])_\psi := \psi \circ \mathcal{P}_{\mathcal{I}}^{(m)}[g] \circ \psi^{-1}. \tag{6.12}$$

Abusing notation slightly, we write $d((\mathcal{P}_{\mathcal{I}}^{(n)}[g])_\psi)(\rho) := d((\mathcal{P}_{\mathcal{I}}^{(n)}[g])_\psi)(\psi(\rho))$.

We now define the notion of non-degeneracy of a trajectory for a chain of symplectomorphisms.

**Definition 6.14.** Let $C_G$ and $\delta_G$ be as in Lemma 6.12 and (6.11). Let $n \in \mathbb{N}$, and $\beta > 0$. We say that $\rho \in \mathscr{R}_F(n, \delta_G, g)$ is $(n, \beta, g)$ non-degenerate, and write $\rho \in \mathscr{N}_F(n, \beta, g)$, if there exist $\mathcal{I}$ and $m$ such that $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho) \in \left[(n-1)F - C_G \delta_G, nF + C_G \delta_G\right]$ and

$$\sup_{\{(\psi, U) \in \mathcal{A} : B(\rho, C_G \delta_G) \subset U\}} \left\| \left(I - d((\mathcal{P}_{\mathcal{I}}^{(m)}[g])_\psi)(\rho)\right)^{-1} \right\| < \beta^{-1}. \tag{6.13}$$

We next strengthen the notion of non-degeneracy of a trajectory to one that can pass from a trajectory to its iterates by introducing a type of $(\beta, q)$-non-degeneracy for returning points.

**Definition 6.15.** Let $C_G$ and $\delta_G$ be as in Lemma 6.12 and (6.11). Let $q \in \mathbb{N}$, $n \in \mathbb{N}$ and $\beta \in (0, 1)$. We say that $\rho \in \mathscr{R}_F(n, \delta_G, g)$ is $(n, \beta, g)$ $q$-non-degenerate, and write $\rho \in \mathscr{N}_{q,F}(n, \beta, g)$, if there exist $\mathcal{I}$ and $m$ satisfying $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho) \in \left[(n-1)F - C_G \delta_G, nF + C_G \delta_G\right]$ and

$$\sup_{\{(\psi, U) \in \mathcal{A} : B(\rho, C_G \delta_G) \subset U\}} \frac{\left\| (I - [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g])_\psi)(\rho)]^q)^{-1} \right\|}{(1 + \|d((\mathcal{P}_{\mathcal{I}}^{(m)}[g])_\psi)(\rho)\|^q)^{2\aleph - 1}} \leq \mathfrak{f}(\beta, q), \qquad \mathfrak{f}(\beta, q) = (\tfrac{5}{2}q^2 \beta^{-1})^{2\aleph}. \tag{6.14}$$

That is, $\rho \in \mathscr{N}_{q,F}(n, \beta, \kappa)$ provided there are $\mathcal{I}$ and $m$ satisfying (6.7) such that the matrix $\mathbf{A} := d(\psi \circ \mathcal{P}_{\mathcal{I}}^{(m)}[g] \circ \psi^{-1})(\psi(\rho))$ is $(\beta, q)$ non-degenerate (as defined in (4.1)) for every chart $(\psi, U) \in \mathcal{A}$ with $B(\rho, C_G \delta_G) \subset U$.

**Remark 6.16.** The reason for the choice of $\mathfrak{f}$ comes from Corollary 5.3 and the fact that we need to be able to sum measures of bad sets over all $q$ in Lemma 7.6. We will make perturbations which guarantee certain properties of the eigenvalues of $d\mathcal{P}_{\mathcal{I}}^{(n)}[g]$ and will use Corollary 5.3 to bound the inverse of $I - (d\mathcal{P}_{\mathcal{I}}^{(n)}[g])^q$. In fact, $q^2$ could be replaced by $q^\gamma$ for any $\gamma > 1$.

**Remark 6.17.** In Definitions 6.14 and 6.15, the existence of $\mathcal{I}$ and $m$ satisfying (6.6) and, respectively (6.13) or (6.14) could be replaced by the requirement that for all $\mathcal{I}$ and $m$ satisfying (6.7), respectively, the estimates (6.13) or (6.14) hold. To see this, we observe that by the second part of Lemma 6.12, all $\mathcal{I}$ and $m$ satisfying (6.7) in fact satisfy (6.6) and result in the same map $\mathcal{P}_{\mathcal{I}}^{(m)}[g]$.

The next lemma demonstrates how the notion of $(n, \beta, g)$ $q$ non-degeneracy can be passed to $(nq, \beta, g)$ non-degeneracy with the trajectory of length $nq$ being a multiple of that of length $n$.

**Lemma 6.18.** *Let $G \subset \mathscr{G}^4(\Gamma)$ bounded and $\Gamma$ be $F$ well separated for $G$. let $C_G$ and $\delta_G$ be as in Lemma 6.12 and (6.11). There is $C > 0$ such that the following holds. Let $g \in G$, $k > 0$, $\beta > 0$, $q_0 \in \mathbb{N}$ and $0 < \delta \leq \min\{\delta_G, (2\mathfrak{f}(\beta, q_0)C^{kq_0})^{-1}\}$. Let*

$$\rho \in \Gamma \cap \mathscr{R}_F(k, \delta, g) \cap \mathscr{N}_{q_0, F}(k, \beta, g),$$

*and $\mathcal{I}$ and $m$ such that $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho) \in \left[(k-1)F - C_G\delta_G\,,\,kF + C_G\delta_G\right],$*

$$\{\mathcal{P}_{\mathcal{I}}^{(qm)}[g](\rho)\}_{q=0}^{q_0} \subset \mathcal{D}_{\mathcal{I}}^{(m)}[g], \qquad \max_{0 \leq q \leq q_0} d\left(\left(\mathcal{P}_{\mathcal{I}}^{(m)}[g_\ell]\right)^q(\rho), \rho\right) \leq \delta,$$

$$\mathcal{P}_{\mathcal{I}}^{(qm)}[g](\rho') = \left(\mathcal{P}_{\mathcal{I}}^{(m)}[g]\right)^q(\rho'), \qquad 0 \leq q \leq q_0, \qquad \rho' \text{ in a neighborhood of } \rho.$$

*Then,*

$$\rho \in \mathscr{N}_F\left(n, (2\mathfrak{f}(\beta, q_0)C^{kq_0})^{-1}, g\right)$$

*for all $n \in \mathbb{N}$ such that $\boldsymbol{T}_{\mathcal{I}}^{(q_0 m)}(\rho) \in [(n-1)F - C_G\delta_G, nF + C_G\delta_G].$*

*Proof.* For $0 \leq q \leq q_0$, set $\rho_q := \mathcal{P}_{\mathcal{I}}^{(qm)}[g](\rho)$ and $\Delta_q := d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho_q) - d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho)$. Then, since there is $C > 0$ such that $\|d^2\varphi_t^g\| \leq C^t$ for all $t \geq 0$, by assumption we conclude that there is $C > 0$ such that $\max_q \|\Delta_q\| \leq C^k\delta$. Therefore,

$$d\mathcal{P}_{\mathcal{I}}^{(mq_0)}[g](\rho) = (d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho) + \Delta_{q_0-1}) \circ \cdots \circ (d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho) + \Delta_1) \circ d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho)$$

$$= [d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho)]^{q_0} + \Delta'$$

with $\|\Delta'\| \leq 2^{q_0}\|d\mathcal{P}_{\mathcal{I}}[g]^{(m)}(\rho)\|^{q_0}C^k\delta$. We conclude that there is $C_0 > 0$ depending only on $G$ such that

$$\|d\mathcal{P}_{\mathcal{I}}^{(q_0 m)}[g](\rho) - [d\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho)]^{q_0}\| \leq C_0^{kq_0}\delta. \tag{6.15}$$

Since $\rho \in \mathscr{N}_{q_0, F}(k, \beta, g)$, there is $C > 0$ depending only on $G$ such that for all $(\psi, U) \in \mathcal{A}$ with $B(\rho, C_G\delta_G) \subset U$,

$$\left\|\left(I - \left[d(\mathcal{P}_{\mathcal{I}}^{(m)})_\psi(\rho)\right]^{q_0}\right)^{-1}\right\| \leq C^{kq_0}\mathfrak{f}(\beta, q_0).$$

In particular, provided that $C_0^{2kq_0}\delta\mathfrak{f}(\beta, q_0) < \frac{1}{2}$, (6.15) yields

$$\|[I - d(\mathcal{P}_{\mathcal{I}}^{(q_0 m)})_\psi(\rho)]^{-1}\| \leq C^{kq_0}\mathfrak{f}(\beta, q_0)(1 - C_0^{2kq_0}\delta\mathfrak{f}(\beta, q_0))^{-1} \leq 2C^{kq_0}\mathfrak{f}(\beta, q_0).$$

The claim now follows from the definition of non-degeneracy. $\qquad\square$

## 7. Perturbing away $q$-degeneracy for simple points

In this section, we first introduce general assumptions on a family of perturbations which will allow us to change the degeneracy properties of a given simple point. We then use a quantitative analog of Sard's theorem to show that the collection of perturbation parameters which produce degeneracy is small.

7.1. **General assumptions.** Let $\mathscr{G}^\nu$ denote the space of $\mathcal{C}^\nu$ metrics on $M$ with the topology induced by the $\mathcal{C}^\nu$ topology on symmetric tensors, and $\iota_{\mathcal{C}^\nu} : \mathcal{C}^\nu \to \mathcal{C}^{\nu-1}$ be the natural inclusion map.

Given a function $N : (0,1) \to \mathbb{N}$, $R > 0$, and $\delta > 0$, we will work with perturbation maps $Q^{R,\delta} : \mathscr{G}^\nu \times \mathbf{\Sigma}_{N(R)} \to \mathscr{G}^\nu$ where

$$\mathbf{\Sigma}_{N(R)} := \prod_{j=1}^{N(R)} B_{\mathbb{R}^{\mathfrak{L}}}(0,1), \qquad \mathfrak{L} := \aleph(2\aleph + 3), \tag{7.1}$$

**Remark 7.1.** Recall that $2\aleph = 2d - 2 = \dim\Gamma$, where $\Gamma$ is a section of $S^*M$. Thus, the dimension $\mathfrak{L} = \aleph(2\aleph + 3)$ appears because we want the parameter $\sigma \in B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$ to be in correspondence with points in the space $(\Gamma, \mathrm{Sp}(2\aleph))$, and the latter has dimension $\mathfrak{L}$. Here $\mathrm{Sp}(2\aleph)$ will parametrize the possible derivatives of a symplectic map at a point in $\Gamma$.

We write $\boldsymbol{\sigma} = (\sigma_j)_{j=1}^{N(R)}$ where each $\sigma_j$ is a component in $B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$, and we refer to $Q^{R,\delta}(g, \boldsymbol{\sigma})$ as the perturbation of $g$ by the parameter $\boldsymbol{\sigma}$. When we construct our family of perturbations in Section 9, $R$ will control the scale on which our metric perturbations take place and $\delta$ will control the size of the perturbation takes place. We ask that the perturbation maps satisfy the following assumptions.

**Definition 7.2** (Good perturbations). Let $\nu_0 \geq 0$ and $N : (0,1) \to \mathbb{N}$. For each $R > 0, \delta > 0$ consider a map

$$Q^{R,\delta} : \mathscr{G}^{\nu_0} \times \mathbf{\Sigma}_{N(R)} \to \mathscr{G}^{\nu_0}$$

such that for each $g \in \mathscr{G}^{\nu_0}$, the map $\boldsymbol{\sigma} \mapsto Q^{R,\delta}(g, \boldsymbol{\sigma})$ with $\boldsymbol{\sigma} \in \ell^\infty(B_{\mathbb{R}^{\mathfrak{L}}}(0,1))$ is Frechet differentiable, and for all $\boldsymbol{\sigma} \in \mathbf{\Sigma}_{N(R)}$ the map $g \mapsto \iota_{\mathcal{C}^{\nu_0-1}}(Q^{R,\delta}(g, \boldsymbol{\sigma}))$ is Frechet differentiable.

We say that $\{Q^{R,\delta}\}_{R,\delta}$ is a $(\nu_0, N)$-*good family of perturbations* if for each $K \subset \mathscr{G}^{\nu_0}$ bounded and each $0 \leq \nu \leq \nu_0$ there exist $\vartheta_\nu \in \mathbb{R}$, $\tilde{\vartheta}_\nu \in \mathbb{R}$ and $C > 0$ such that the following holds. For $\delta > 0$, $R > 0$,

- $Q^{R,\delta}(g, 0) = g, \qquad g \in \mathscr{G}^{\nu_0}.$ (7.2)
- $\|Q^{R,\delta}(g, \boldsymbol{\sigma}) - Q^{R,\delta}(g, \tilde{\boldsymbol{\sigma}})\|_{\mathcal{C}^\nu} \leq C\delta R^{-\vartheta_\nu} \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|_{\ell^\infty}, \qquad \boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}} \in \mathbf{\Sigma}_{N(R)}, g \in K.$ (7.3)
- $\|D_g Q^{R,\delta} - I\|_{\mathcal{C}^{\nu-1} \to \mathcal{C}^{\nu-1}} \leq C\delta R^{-\vartheta_\nu}, \qquad g \in K, \boldsymbol{\sigma} \in \mathbf{\Sigma}_{N(R)}.$ (7.4)

Furthermore, for all $\alpha \in \mathbb{N}^{\mathfrak{L}}$ there is $C_\alpha > 0$ such that for all $g \in K$, $\delta > 0$, and $R > 0$

$$\max_{1 \leq j \leq N(R)} \|\partial_{\sigma_j}^\alpha Q^{R,\delta}|_{(g,\boldsymbol{\sigma})}\|_{B_{\mathbb{R}^{\mathfrak{L}}}(0,1) \to \mathcal{C}^\nu} \leq C_\alpha R^{-|\alpha|\tilde{\vartheta}_\nu} \delta^{|\alpha|}, \qquad g \in K, \boldsymbol{\sigma} \in \mathbf{\Sigma}_{N(R)}. \tag{7.5}$$

We also need an assumption which guarantees that the perturbation $Q^{R,\delta}$ explores a sufficiently large set of possible metrics.

**Definition 7.3** (Admissible pairs). Let $\nu \geq 5$, $b > 0$ and $\mathfrak{y} \geq 1$. Let $N : (0,1) \to \mathbb{N}$ and $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations.

We say that $(\Gamma, G)$ is a $(F, b, \mathfrak{y})$-*admissible pair for* $\{Q^{R,\delta}\}_{R,\delta}$ if $G \subset \mathscr{G}^\nu$ is bounded, $\Gamma$ is $F$-well separated for $G$, and for all $K \subset G$ bounded in $\mathscr{G}^\nu$ there are $c > 0$ and $\varepsilon_0 > 0$ such that for all $R > 0$, $\delta > 0$, $g \in K$, $\alpha \in (0, c)$, $n \in \mathbb{N}$, and $\rho_0 \in \Gamma$ satisfying

$$R < c^n \alpha, \qquad\qquad \delta \leq R^b c^{n+1}, \qquad\qquad \rho_0 \in \mathscr{S}_F(n, \alpha, g) \cap \mathscr{R}_F(n, \varepsilon_0, g),$$

there exist $j_0 \in \{1, \ldots, N(R)\}$, $\mathcal{I}_0$, $m_0$ such that $B_\Gamma(\rho_0, R^\mathfrak{y}) \subset \mathcal{D}^{(m_0)}_{\mathcal{I}_0}[Q^{R,\delta}(g, \boldsymbol{\sigma})]$ for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)}$, $\boldsymbol{T}^{(m_0)}_{\mathcal{I}_0}[g](\rho) \in \big[(n-1)F - C_G \delta_G, \, nF + C_G \delta_G\big]$, and for all $\rho \in B_\Gamma(\rho_0, R^\mathfrak{y})$ and $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)}$,

$$|d_{\sigma_{j_0}} \Psi^{R,\delta}_{g, \mathcal{I}_0}(m_0, \boldsymbol{\sigma}, \rho) v| \geq \delta c^{n+1} |v|, \qquad \text{for all } v \in T_{\boldsymbol{\sigma}_{j_0}} B_{\mathbb{R}^\mathfrak{L}}(0,1), \tag{7.6}$$

where

$$\Psi^{R,\delta}_{g, \mathcal{I}_0}(m_0, \boldsymbol{\sigma}, \rho) := \Big(\mathcal{P}^{(m_0)}_{\mathcal{I}_0}[Q^{R,\delta}(g, \boldsymbol{\sigma})](\rho) \, , \, d_\rho\big(\mathcal{P}^{(m_0)}_{\mathcal{I}_0}[Q^{R,\delta}(g, \boldsymbol{\sigma})]\big)(\rho)\Big). \tag{7.7}$$

The constants $C_G$ and $\delta_G$ are the ones given in Lemma 6.12.

To understand why we are interested in the map $\Psi^{R,\delta}_g$, observe that knowing

$$\big(\mathcal{P}^{(m_0)}_{\mathcal{I}}[g](\rho), d_\rho \mathcal{P}^{(m_0)}_{\mathcal{I}}[g](\rho)\big)$$

is exactly the information need to understand (1) whether $\rho \in \mathscr{R}_F(n, \beta, g)$ and (2) if $\rho \in \mathscr{R}_F(n, \beta, g)$, whether $\rho \in \mathscr{N}_{q,F}(n, \beta, g)$. Thus, the dimension of $B_{\mathbb{R}^\mathfrak{L}}(0,1)$ is chosen so that we are able to use the $j_0$-th factor in $\boldsymbol{\Sigma}_{N(R)}$ to locally explore an open set in the image of the maps $\Psi^{R,\delta}_{g,\mathcal{I}}$ (whose dimension is $\mathfrak{L} = \aleph(2\aleph + 3)$) by varying $\sigma_{j_0} \in B_{\mathbb{R}^\mathfrak{L}}(0,1)$. See also Remark 7.1.

## 7.2. Perturbing away $(\beta, q)$-degeneracy for simple points.

In order to control the volume of the parameters in $\boldsymbol{\Sigma}_{N(R)}$ which may produce degeneracy we apply a quantitative version of Sard's theorem ([Yom85, Theorem 4.2]). We recall the theorem here for convenience. Below, $\mathcal{M}_{\mathbb{R}^k}(Y, s)$ denotes the covering number of $Y \subset \mathbb{R}^k$ by balls in $\mathbb{R}^k$ of radius $s$ (see Definition 5.4).

**Proposition 7.4** ([Yom85] Theorem 4.2)**.** *Let* $\Omega \subset \mathbb{R}^\ell$ *and* $L > 0$ *such that for all* $\rho_1, \rho_2 \in \Omega$, *there is a path in* $\Omega$ *of length smaller than* $L\|\rho_1 - \rho_2\|$ *connecting* $\rho_1$ *and* $\rho_2$. *Let* $\Psi : B_{\mathbb{R}^k}(0,1) \times \Omega \to \mathbb{R}^k$ *be a continuously differentiable mapping such that there exist* $c_1 > 0$, $c_2 > 0$ *so that for all* $\rho \in \Omega$, *and* $\sigma, \tilde{\sigma} \in B_{\mathbb{R}^k}(0,1)$

$$\|d_\rho \Psi(\sigma, \rho)\| \leq c_1, \qquad \|\Psi(\tilde{\sigma}, \rho) - \Psi(\sigma, \rho)\| \geq c_2 \|\sigma - \tilde{\sigma}\|.$$

*Then, for any* $s > 0$ *and any* $X \subset \Omega$, $Y \subset \mathbb{R}^k$,

$$\mathcal{M}_{\mathbb{R}^k}(\Delta_\Psi(X, Y), \, \tilde{s}) \leq \mathcal{M}_{\mathbb{R}^\ell}(X, s) \mathcal{M}_{\mathbb{R}^k}(Y, s)$$

*where* $\tilde{s} := 2(Lc_2^{-1}(1 + c_1) + 1)s$ *and*

$$\Delta_\Psi(X, Y) := \{\sigma \in B_{\mathbb{R}^k}(0,1) : \text{there is } \rho \in X \text{ such that } \Psi(\sigma, \rho) \in Y\}. \tag{7.8}$$

Let $\mathbf{m}_{\mathbb{R}^\mathfrak{L}}$ be the Lebesgue measure on $\mathbb{R}^\mathfrak{L}$ normalized so that

$$\mathbf{m}_{\mathbb{R}^\mathfrak{L}}(B_{\mathbb{R}^\mathfrak{L}}(0,1)) = 1 \tag{7.9}$$

and let $\mathbf{m}_{\boldsymbol{\Sigma}_{N(R)}}$ be the product measure on $\boldsymbol{\Sigma}_{N(R)}$. We write $\mathbf{m}_\Gamma$ for the volume measure on $\Gamma$.

Next, we show that there is $c > 0$ such that if $\rho_0 \in \Gamma$ is $(n, c)$ returning and $(n, \alpha)$ simple for some metric $g$, then there is only a small measure set of $\boldsymbol{\sigma}$ that yield a perturbed metric

$Q^{R,\delta}(g,\boldsymbol{\sigma})$ for which there are points $\rho \in B(\rho_0, R^{\mathfrak{y}'})$ that are $(n, \beta q^{-3})$ returning and are *not* $(n,\beta)$ $q$-non-degenerate for some $q \in \mathbb{N}$. This can be done for $\alpha < c$ and $\beta < c^n R^{\mathfrak{y}'}$ provided that we work with metrics that are at least $\mathcal{C}^5$.

**Remark 7.5.** Lemma 7.6 below is the only place where we require $g$ to have 5 derivatives.

**Lemma 7.6.** *Let $b > 0$, $\nu \geq 5$, $\mathfrak{y} \geq 1$, and $N : (0,1) \to \mathbb{N}$. Let $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations and $(\Gamma, G)$ be a $(F, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$.*

*Let $K \subset G$ bounded in $\mathscr{G}^5$ and $\mathfrak{y}' \geq \mathfrak{y}$. Let $\delta_G$ be as in Lemma 6.12 and (6.11). There exist $c \in (0,1)$, $C > 0$ such the following holds. Let $n \in \mathbb{N}$, $\alpha \in (0,c)$,*

$$0 < R < \min(c^n \alpha, \tfrac{1}{2} c \delta_G), \qquad 0 < \delta \leq \min(R^b c^{n+1}, R^{2\tilde{\vartheta}_4} c^{n+1}, cR^{\vartheta_5}).$$

*If $g \in K$, $\rho_0 \in \Gamma$, and $B_\Gamma(\rho_0, R^{\mathfrak{y}'}) \cap \mathscr{S}_F(n, \alpha, g) \cap \mathscr{R}_F(n, c, g) \neq \emptyset$, then for all $0 < \beta < \min\left(\tfrac{1}{2}\delta_G, c^n R^{\mathfrak{y}'}\right)$*

$$\mathbf{m}_{\boldsymbol{\Sigma}_{N(R)}}\left(\mathcal{S}_g^{\rho_0, R, \delta}(n, \alpha, \beta)\right) \leq \mathbf{m}_\Gamma\left(B_\Gamma(\rho_0, R^{\mathfrak{y}'})\right)(C^n R^{\mathfrak{y}'})^{\mathfrak{L} - 2\aleph - 1}\delta^{-\mathfrak{L}}\beta,$$

*where*

$$\mathcal{S}_g^{\rho_0, R, \delta}(n, \alpha, \beta) := \Big\{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)} \,:\, \text{there are } \rho \in B_\Gamma(\rho_0, R^{\mathfrak{y}'}) \text{ and } q \in \mathbb{N} \text{ such that}$$

$$\rho \in \mathscr{R}_F\left(n, \beta q^{-3}, Q^{R,\delta}(g,\boldsymbol{\sigma})\right) \backslash \mathscr{N}_{q,F}\left(n, \beta, Q^{R,\delta}(g,\boldsymbol{\sigma})\right)\Big\}.$$

*Proof.* Let $\varepsilon_0$ as in Definition 7.3. Using the estimate on $(\varphi_t^g)^* : \mathcal{C}^1 \to \mathcal{C}^1$ in Lemma 6.11, we first note that there exists $c = c(K) > 0$ such that if $g \in K$, $\rho_0 \in \Gamma$, $0 < \alpha < 1$, $n \in \mathbb{N}$, and $0 < R \leq c^n \alpha$, then $B(\rho_0, R^{\mathfrak{y}'}) \cap \mathscr{S}_F(n, \alpha, g) \cap \mathscr{R}_F(n, c, g) \neq \emptyset$ implies $\rho_0 \in \mathscr{S}_F(n, \tfrac{1}{3}\alpha, g) \cap \mathscr{R}_F(n, \varepsilon_0, g)$. By Definition 7.3, this allows us to work with indices $j_0$, $m_0$, and $\mathcal{I}_0$ such that the lower bound on the differential of $\Psi_{g, \mathcal{I}_0}^{R,\delta}$ in (7.6) holds.

The result will follow once we find $c, C$ such that, under the assumptions of the lemma, for all $\hat{\boldsymbol{\sigma}} \in \prod_{j=1}^{N(R)-1} B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$ and $0 < \beta < \min(\tfrac{1}{2}\delta_G, c^n R^{\mathfrak{y}'})$

$$\mathbf{m}_{\mathbb{R}^{\mathfrak{L}}}\left(\mathcal{S}_g^{\rho_0, R, \delta}(n, \alpha, \beta; j_0, \hat{\boldsymbol{\sigma}})\right) \leq \mathbf{m}_\Gamma\left(B_\Gamma(\rho_0, R^{\mathfrak{y}'})\right)C(C^n R^{\mathfrak{y}'})^{\mathfrak{L} - 2\aleph - 1}\delta^{-\mathfrak{L}}\beta, \tag{7.10}$$

where for $\hat{\boldsymbol{\sigma}} \in \prod_{j=1}^{N(R)-1} B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$, we let $\boldsymbol{\sigma}_{j_0}^- := (\hat{\boldsymbol{\sigma}}_i)_{i=1}^{j_0-1}$, $\boldsymbol{\sigma}_{j_0}^+ := (\hat{\boldsymbol{\sigma}}_i)_{i=j_0}^{N(R)-1}$, and

$$\mathcal{S}_g^{\rho_0, R, \delta}(n, \alpha, \beta; j_0, \hat{\boldsymbol{\sigma}}) = \Big\{\sigma \in B_{\mathbb{R}^{\mathfrak{L}}}(0,1) : \ (\boldsymbol{\sigma}_{j_0}^-, \sigma, \boldsymbol{\sigma}_{j_0}^+) \in \mathcal{S}_g^{\rho_0, R, \delta}(n, \alpha, \beta)\Big\}.$$

To prove the claim in (7.10) we restrict the map $\Psi_{g, \mathcal{I}_0}^{R,\delta}$ to its $j_0^{th}$-component in $\boldsymbol{\sigma}$. For $\hat{\boldsymbol{\sigma}} \in \prod_{j=1}^{N(R)-1} B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$ define the restricted map $\Psi_{g, \mathcal{I}_0, j_0, \hat{\boldsymbol{\sigma}}}^{R,\delta}(n, \sigma, \rho) = \Psi_{g, \mathcal{I}_0}^{R,\delta}\left(n, (\boldsymbol{\sigma}_{j_0}^-, \sigma, \boldsymbol{\sigma}_{j_0}^+), \rho\right)$, with domain $(n, \sigma, \rho) \in \mathbb{N} \times B_{\mathbb{R}^{\mathfrak{L}}}(0,1) \times B_\Gamma(\rho_0, R^{\mathfrak{y}'})$.

By assumption (7.3), after possibly shrinking $c$ in a way that depends only on $K$, the set $\{Q^{R,\delta}(g,\boldsymbol{\sigma}) : g \in K, \boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)}\} \subset \mathscr{G}^5$ is bounded provided $\delta < cR^{\vartheta_5}$. Then, combining (7.5) with Lemma 6.11 we obtain that there is $C > 0$ depending only on $(K, \Gamma, F)$ such that

$$\sup_{\rho, \hat{\boldsymbol{\sigma}}, \sigma, j} \sup_{\{\mathcal{I} : \rho \in \mathcal{D}_{\mathcal{I}}^{(n)}[Q^{R,\delta}(g,\boldsymbol{\sigma})]\}} \|d_\sigma^2 \Psi_{g, \mathcal{I}, j, \hat{\boldsymbol{\sigma}}}^{R,\delta}(n, \sigma, \rho)\| \leq C^n R^{-2\tilde{\vartheta}_4}\delta^2. \tag{7.11}$$

Therefore, since $\rho_0 \in \Gamma \cap \mathscr{S}_F(n, \frac{1}{3}\alpha, g) \cap \mathscr{R}_F(n, \varepsilon_0, g)$, assumption (7.6) and (7.11) yield that for each $\hat{\boldsymbol{\sigma}}$, and $\sigma, \tilde{\sigma} \in B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$,

$$\inf_{\rho \in B_\Gamma(\rho_0, R^{\mathfrak{y}})} \|\Psi_{g,\mathcal{I}_0,j_0,\hat{\boldsymbol{\sigma}}}^{R,\delta}(m_0,\sigma,\rho) - \Psi_{g,\mathcal{I}_0,j_0,\hat{\boldsymbol{\sigma}}}^{R,\delta}(m_0,\tilde{\sigma},\rho)\| \geq \delta c^{n+1}\|\sigma - \tilde{\sigma}\|, \tag{7.12}$$

provided that

$$R < c^n \alpha, \qquad \delta \leq R^b c^{n+1}, \qquad C^n R^{-2\tilde{\vartheta}_4}\delta^2 \leq \tfrac{1}{10}c^{n+1}\delta, \tag{7.13}$$

after possibly shrinking $c$ in a way that only depends on $(K, \Gamma, F)$.

Note that if $\sigma \in \mathcal{S}_g^{\rho_0,R,\delta}(n,\alpha,\beta;j,\hat{\boldsymbol{\sigma}})$, then by Lemma 6.12 there are $\rho \in B_\Gamma(\rho_0, R^{\mathfrak{y}'})$ and $q \in \mathbb{N}$ such that

$$d(\rho, \mathcal{P}_{\mathcal{I}_0}^{(m_0)}[Q^{R,\delta}(g,\boldsymbol{\sigma})](\rho)) < C_G \beta q^{-3}, \qquad \rho \notin \mathscr{N}_{q,F}(n,\beta,Q^{R,\delta}(g,\boldsymbol{\sigma})). \tag{7.14}$$

After requiring that $R^{\mathfrak{y}} < C_G \delta_G$ and $\beta < \frac{\delta_G}{2}$, there exists a coordinate chart $(U,\psi) \in \mathcal{A}$ (see (6.11)) with $B_\Gamma(\rho_0, R^{\mathfrak{y}'}) \subset \psi(U)$ such that (6.14) does not hold with $\mathcal{I} = \mathcal{I}_0$ and $m = m_0$. Furthermore, $\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[Q^{R,\delta}(g,\boldsymbol{\sigma}))](\rho) \in \psi(U)$ for all $\rho \in \mathscr{R}_F(n, \beta q^{-3}, Q^{R,\delta}(g,\boldsymbol{\sigma})) \cap B_\Gamma(\rho_0, R^{\mathfrak{y}'})$.

Next, consider the map $\tilde{\Psi}_g^{R,\delta} : \mathbb{N} \times \boldsymbol{\Sigma}_{N(R)} \times \psi(B_{\tilde{\Gamma}}(\rho_0, R^{\mathfrak{y}'})) \to \psi(U) \times \mathrm{Sp}(2\aleph)$

$$\tilde{\Psi}_g^{R,\delta}(n,\boldsymbol{\sigma},x) := \left( (\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[Q^{R,\delta}(g,\boldsymbol{\sigma})]_\psi(x) - x \, , \, d_x((\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[Q^{R,\delta}(g,\boldsymbol{\sigma})])_\psi(x)) \right).$$

and for each $n \in \mathbb{N}$ consider its restriction

$$\tilde{\Psi}_n : B_{\mathbb{R}^{\mathfrak{L}}}(0,1) \times \psi(B_\Gamma(\rho_0, R^{\mathfrak{y}})) \to \psi(U) \times \mathrm{Sp}(2\aleph), \qquad \tilde{\Psi}_n(\sigma,x) := \tilde{\Psi}_g^{R,\delta}\left(n, (\boldsymbol{\sigma}_{j_0}^-, \sigma, \boldsymbol{\sigma}_{j_0}^+), x\right).$$

It follows from (7.14), our choice of $\psi$, and the comment following Definition 6.15, that $\boldsymbol{A} := d_x(\psi \circ (\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[Q^{R,\delta}(g,\boldsymbol{\sigma})]) \circ \psi^{-1})(\psi(\rho))$ is not $(\beta, q)$ non-degenerate. Therefore, Corollary 5.3 applied to $\boldsymbol{A}$ combined with the first statement in (7.14) yield, that, with $s_q := \beta q^{-3}$, $\mathcal{D}(\frac{2\pi p}{q}, s_q)$ as in (5.4), and $\Delta_{\tilde{\Psi}_n}$ as in (7.8),

$$\mathcal{S}_g^{\rho_0,R,\delta}(n,\alpha,\beta;j,\hat{\boldsymbol{\sigma}}) \subset \Delta_{\tilde{\Psi}_n}(X, \cup_{q=1}^\infty Y_q), \tag{7.15}$$

$$X := \psi(B_\Gamma(\rho_0, R^{\mathfrak{y}'})), \qquad Y_q := B_{\mathbb{R}^{2\aleph}}(0, C_G s_q) \times \bigcup_{p=0}^{q-1} \mathcal{D}(\tfrac{2\pi p}{q}, s_q) \cap B_{\mathbf{A_0}}(r_0).$$

Here, we let $\mathbf{A_0} := d_x(\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[g])_\psi(x)$ and $r_0 := C_1^n R^{\mathfrak{y}'}$ with $C_1$ chosen so that the image of the second factor in $\tilde{\Psi}_n$ lies within $B_{\mathbf{A_0}}(r_0)$. This is possible to arrange since there is $C_0 > 0$ depending only on $(K, \psi, U)$ such that $\|\mathbf{A_0}\| \leq C_0^n$. Also, by Lemma 6.11 with $\nu = 4$ and Definition 7.2 with $\nu = 3$, there is $C_1 > 0$ depending only on $(K, \psi, U)$ such that

$$\|d_x((\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[Q^{R,\delta}(g,\boldsymbol{\sigma})])_\psi(x) - \mathbf{A_0}\| \leq (C_1 \delta R^{-\vartheta_3})^n R^{\mathfrak{y}'} \leq r_0.$$

Note that the last inequality holds provided $\delta < R^{\vartheta_3}$.

With (7.15) in place we proceed to check that the hypothesis in Proposition 7.4 hold. Note that by Lemma 6.11 there is $C > 0$, depending only on $(K, \psi, U)$, with $\sup_{x \in \psi(B_\Gamma(\rho_0, R^{\mathfrak{y}'}))} \|d_x \tilde{\Psi}_n(\boldsymbol{\sigma}, x)\| \leq$

$C^n$. In addition, by (7.12), there is $c > 0$ depending only on $(K, \psi, U)$ such that for $\sigma, \tilde{\sigma} \in B_{\mathbb{R}^{\mathfrak{L}}}(0, 1)$

$$\inf_{x \in \psi(B_{\Gamma}(\rho_0, R^{\mathfrak{y}}))} \|\tilde{\Psi}_n(\sigma, x) - \tilde{\Psi}_n(\tilde{\sigma}, x)\| \geq \delta c^{n+1} \|\sigma - \tilde{\sigma}\|.$$

It follows that, for each $n \in \mathbb{N}$, Proposition 7.4 applies to $\tilde{\Psi}_n$ with $k := \mathfrak{L} = \dim(\psi(U) \times \mathrm{Sp}(2\aleph))$, $\ell := 2\aleph$, $c_1 = C^n$, $c_2 = \delta c^{n+1}$, and some $L$ depending only on $(\psi, U)$. Note that we are abusing notation slightly here to identify $\mathrm{Sp}(2\aleph)$ with $\mathbb{R}^{\aleph(2\aleph+1)}$ in the codomain. Finally, since $\mathbf{m}_{\mathbb{R}^{\mathfrak{L}}}(\Delta_{\tilde{\Psi}_n}(X, Y_q)) \leq c_{\mathfrak{L}} \mathcal{M}_{\mathbb{R}^{\mathfrak{L}}}(\Delta_{\tilde{\Psi}_n}(X, Y_q), s) s^{\mathfrak{L}}$ for all $s > 0$, Proposition 7.4 and (7.15) yield that, with $\tilde{s}_q := 2(L\delta^{-1}c^{-n-1}(1 + C^n) + 1)s_q$,

$$\mathbf{m}_{\mathbb{R}^{\mathfrak{L}}}(\mathcal{S}_g^{\rho_0, \tilde{R}, \delta}(n, \alpha, \beta; j, \hat{\boldsymbol{\sigma}})) \leq \sum_{q=1}^{\infty} \mathbf{m}_{\mathbb{R}^{\mathfrak{L}}}(\Delta_{\tilde{\Psi}_n}(X, Y_q)) \leq c_{\mathfrak{L}} \sum_{q=1}^{\infty} \mathcal{M}_{\mathbb{R}^{2\aleph}}(X, s_q) \mathcal{M}_{\mathbb{R}^{\mathfrak{L}}}(Y_q, s_q) \tilde{s}_q^{\mathfrak{L}}. \quad (7.16)$$

To control the right hand side of (7.16) first observe that

$$\mathcal{M}_{\mathbb{R}^{2\aleph}}(X, s_q) \leq C\mathbf{m}_{\Gamma}(B_{\Gamma}(\rho_0, R^{\mathfrak{y}'}))s_q^{-2\aleph} \quad (7.17)$$

provided $0 < s_q \leq R^{\mathfrak{y}'}$, where $C > 0$ depends only on $(\Gamma, \psi, U)$. Next, by Corollary 5.11 with $r = C_0^n + r_0$, there are $C, \varepsilon_0 > 0$ such that if $0 < s_q \leq \min(\varepsilon_0(C_0^n + r_0), 12r_0(C_0^n + r_0)^2)$, then

$$\mathcal{M}_{\mathbb{R}^{\mathfrak{L}}}(Y_q, s_q) \leq Cq((C_0^n + r_0)^2 r_0 s_q^{-1})^{\mathfrak{L}-2\aleph-1}. \quad (7.18)$$

By possibly adjusting $c > 0$ one last time, asking $\beta \leq \min(\frac{\delta_G}{2}, c^n R^{\mathfrak{y}'})$ makes $s_q$ be sufficiently small, and so combining (7.16), (7.17) and (7.18) yields

$$\mathbf{m}_{\mathbb{R}^{\mathfrak{L}}}(\mathcal{S}_g^{\rho_0, \tilde{R}, \delta}(n, \alpha, \beta; j, \hat{\boldsymbol{\sigma}})) \leq \sum_{q=1}^{\infty} C\mathbf{m}_{\Gamma}(B_{\Gamma}(\rho_0, R^{\mathfrak{y}'}))s_q^{-2\aleph} Cq((C_0^n + r_0)^2 r_0 s_q^{-1})^{\mathfrak{L}-2\aleph-1} \tilde{s}_q^{\mathfrak{L}}$$

$$\leq C(C^n R^{\mathfrak{y}'})^{\mathfrak{L}-2\aleph-1}\mathbf{m}_{\Gamma}(B_{\Gamma}(\rho_0, R^{\mathfrak{y}'}))\delta^{-\mathfrak{L}}\beta \sum_{q=1}^{\infty} q^{-2}.$$

This proves the claim in (7.10) provided $R, \alpha, \delta, n$ satisfy (7.13), $\beta < \min(\frac{\delta_G}{2}, c^n R^{\eta'})$, and $\delta < cR^{\vartheta_5}$. $\qquad \square$

**Corollary 7.7.** *Let $b > 0$, $\nu \geq 5$, $\mathfrak{y} \geq 1$, and $N : (0, 1) \to \mathbb{N}$. Let $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations and $(\Gamma, G)$ be a $(F, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$.*

*Let $K \subset G$ bounded in $\mathscr{G}^5$ and $\mathfrak{y}' \geq \mathfrak{y}$. Let $\delta_G$ be as in Lemma 6.12 and (6.11). There exist $c \in (0, 1)$, $C > 0$ such the following holds. Let $n \in \mathbb{N}$, $\alpha \in (0, c)$,*

$$0 < R < \min(c^n \alpha, \tfrac{1}{2}c\delta_G), \qquad 0 < \delta \leq \min(R^b c^{n+1}, R^{2\tilde{\vartheta}_4}c^{n+1}, cR^{\vartheta_5+1}). \quad (7.19)$$

*If $g \in K$, then for all $0 < \beta < \min\left(\tfrac{1}{2}\delta_G, c^n R^{\mathfrak{y}'}\right)$*

$$\mathbf{m}_{\Sigma_{N(R)}}(\mathcal{S}_g^{R,\delta}(n, \alpha, \beta)) \leq C(C^n cR^{\mathfrak{y}'})^{\mathfrak{L}-2\aleph-1}\delta^{-\mathfrak{L}}\beta,$$

*where*

$$\mathcal{S}_g^{R,\delta}(n,\alpha,\beta) := \Big\{ \boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)} \ : \ \exists q \in \mathbb{N} \ \ s.t.$$

$$\Gamma \cap \mathscr{S}_{\mathsf{F}}\big(n,\alpha,Q^{R,\delta}(g,\boldsymbol{\sigma})\big) \cap \mathscr{R}_{\mathsf{F}}\big(n,\beta q^{-3},Q^{R,\delta}(g,\boldsymbol{\sigma}))\big) \backslash \mathscr{N}_{q,\mathsf{F}}\big(n,\beta,Q^{R,\delta}(g,\boldsymbol{\sigma})\big) \neq \emptyset \Big\}.$$

*Proof.* Let $\{\rho_i\}_i \subset \Gamma$ be an $R^{\mathfrak{y}'}$ maximal separated set in $\Gamma$. We claim that

$$\mathcal{S}_g^{R,\delta}(n,\alpha,\beta) \subset \mathcal{Z}_g^{R,\delta}(n,\alpha,\beta) \subset \bigcup_{i \in \mathcal{I}} \mathcal{S}_g^{\rho_i,R,\delta}(n,\alpha,\beta), \tag{7.20}$$

where $\mathcal{I} := \{i : B_\Gamma(\rho_i, R^{\mathfrak{y}'}) \cap \mathscr{S}_{\mathsf{F}}(n,\frac{1}{2}\alpha,g) \cap \mathscr{R}_{\mathsf{F}}(n,c,g) \neq \emptyset\}$ and

$$\mathcal{Z}_g^{R,\delta}(n,\alpha,\beta) := \Big\{ \boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)} \ : \ \exists q \in \mathbb{N} \ \ s.t.$$

$$\Gamma \cap \mathscr{S}_{\mathsf{F}}\big(n,\tfrac{1}{2}\alpha,g\big) \cap \mathscr{R}_{\mathsf{F}}\big(n,\beta q^{-3},Q^{R,\delta}(g,\boldsymbol{\sigma}))\big) \backslash \mathscr{N}_{q,\mathsf{F}}\big(n,\beta,Q^{R,\delta}(g,\boldsymbol{\sigma})\big) \neq \emptyset \Big\}.$$

The claim in (7.20) follows from observing that

$$\mathscr{S}_{\mathsf{F}}\big(n,\alpha,Q^{R,\delta}(g,\boldsymbol{\sigma})\big) \subset \mathscr{S}_{\mathsf{F}}\big(n,\tfrac{1}{2}\alpha,g\big), \qquad \mathscr{R}_{\mathsf{F}}\big(n,\beta q^{-3},Q^{R,\delta}(g,\boldsymbol{\sigma})\big) \subset \mathscr{R}_{\mathsf{F}}\big(n,c,g\big).$$

Indeed, these two inclusions follow from combining the conclusion from Lemma 6.11 that for all $t \in \mathbb{R}$

$$d(\varphi_t^g(\rho), \varphi_t^{Q^{R,\delta}(g,\boldsymbol{\sigma})}(\rho)) \leq C^{|t|} \| g - Q^{R,\delta}(g,\boldsymbol{\sigma}) \|_{C^3}$$

with (7.2), (7.3), and the requirement that $\delta R^{-\vartheta_3} \leq c^n \alpha$ (the latter holds since $\vartheta_5 \geq \vartheta_3$ and $\delta < cR^{\vartheta_5+1}$).

Finally, combining (7.20) with Lemma 7.6 applied to each ball $B_{\tilde{\Gamma}}(\rho_i, R^{\mathfrak{y}'})$ with $i \in \mathcal{I}$ yields the desired bound. $\qquad \square$

## 8. Induction on the length of orbits

Let $b > 0$. Let $\nu > 0$, $N : (0,1) \to \mathbb{N}$, $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations, and let $(\Gamma, G)$ be a $(\mathsf{F}, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$. In this section we will prove that there is a predominant set, $G$ of metrics such that for all $g$ in $G$, there is $C > 0$ such that

$$\Gamma \cap \mathscr{R}_{\mathsf{F}}(n,\beta_n,g) \subset \mathscr{N}_{\mathsf{F}}(n,\beta_n,g), \qquad \beta_n = C^{-Cn^{\Omega\nu}}, \tag{8.1}$$

We do this using a family of probing maps generated by the $(\nu, N)$-good family of perturbations. Once this is done (see Proposition 8.9), to prove Theorem 1.3 it will remain to show that there are $(\nu, N)$-good families of perturbations (we do this in Section 9). We continue to work in the setting of Section 7.

### 8.1. Construction of families of probing maps.
Let $\nu \geq 0$, a function $N : (0,1) \to \mathbb{N}$, and let $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations. Let $\{K_s\}_{s \in (0,1)} \subset \mathscr{G}^\nu$ be a family of nested closed bounded subsets $K_{s_1} \subset K_{s_2}$ for $s_1 > s_2$ such that $\mathscr{G}^\nu = \cup_{s \in (0,1)} K_s$. By Definition 7.2 for every $s \in (0,1)$ bounded there is $C_s > 0$ such that for all $\delta > 0$, $R > 0$,

$$\| Q^{R,\delta}(g,\boldsymbol{\sigma}) - Q^{R,\delta}(g,\tilde{\boldsymbol{\sigma}}) \|_{\mathcal{C}^\nu} \leq C_s \delta R^{-\vartheta_\nu} \| \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}} \|_{\ell^\infty}, \tag{8.2}$$

$$\| Q^{R,\delta}(\tilde{g},\boldsymbol{\sigma}) - Q^{R,\delta}(g,\boldsymbol{\sigma}) \|_{\mathcal{C}^{\nu-1}} \leq (1 + C_s \delta R^{-\vartheta_\nu}) \| g - \tilde{g} \|_{\mathcal{C}^{\nu-1}}. \tag{8.3}$$

for $\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}} \in \boldsymbol{\Sigma}_{N(R)}$, $g \in K_s$. Next, for $\varepsilon > 0$ let $s(\varepsilon) > 0$ be an increasing function such that $\lim_{\varepsilon \to 0^+} C_{s(\varepsilon)} \varepsilon = 0$ and $C_{s(\varepsilon)} \varepsilon$ is increasing. Next, define the closed and bounded sets

$$\mathscr{G}_\varepsilon^\nu := \{g \in K_{s(\varepsilon)} : d(g, \mathscr{G}^\nu \backslash K_{s(\varepsilon)}) \geq 2C_{s(\varepsilon)} \varepsilon\}. \tag{8.4}$$

Note that $\mathscr{G}_{\varepsilon_2}^\nu \subset \mathscr{G}_{\varepsilon_1}^\nu$ for $\varepsilon_1 < \varepsilon_2$ and $\mathscr{G}^\nu = \cup_{\varepsilon > 0} \mathscr{G}_\varepsilon^\nu$.

For $\varepsilon \in (0, 1)$ let $\boldsymbol{\delta}_\varepsilon = \{\delta_j(\varepsilon)\}_{j=0}^\infty \subset [0, 1)$ and $\mathfrak{R}_\varepsilon = \{R_j(\varepsilon)\}_{j=0}^\infty \subset (0, 1)$, satisfy

$$\sum_{j=0}^\infty \delta_j(\varepsilon)(R_j(\varepsilon))^{-\vartheta_\nu} \leq \varepsilon. \tag{8.5}$$

We now construct a family of probing maps associated to the perturbation maps $Q^{R_j(\varepsilon), \delta_j(\varepsilon)}$. Given $J = 0, \dots, N$ define the spaces

$$\boldsymbol{\Sigma}_J(\mathfrak{R}_\varepsilon) := \prod_{j=0}^J \boldsymbol{\Sigma}_{N(R_j(\varepsilon))}, \qquad \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon) := \prod_{j=0}^\infty \boldsymbol{\Sigma}_{N(R_j(\varepsilon))}. \tag{8.6}$$

Also, given a sequence $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3, \dots) \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon)$ we write $\hat{\boldsymbol{\sigma}}_J := (\boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_J) \in \boldsymbol{\Sigma}_J(\mathfrak{R}_\varepsilon)$.

We are now ready to define the family of probing maps (see Definition 2.2). We start by letting $F_0^{\mathfrak{R}, \boldsymbol{\delta}} : \mathscr{G}_\varepsilon^\nu \times \boldsymbol{\Sigma}_{N(R_0)} \to \mathscr{G}^\nu$ be the map

$$F_0^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) := Q^{R_0(\varepsilon), \delta_0(\varepsilon)}(g, \boldsymbol{\sigma}) \tag{8.7}$$

and $F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon} : \mathscr{G}_\varepsilon^\nu \times \boldsymbol{\Sigma}_J(\mathfrak{R}_\varepsilon) \to \mathscr{G}^\nu$ by

$$F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J) := Q^{R_J(\varepsilon), \delta_J(\varepsilon)}\big(F_{J-1}^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_{J-1}), \boldsymbol{\sigma}_J\big). \tag{8.8}$$

Finally, we define

$$F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon} : \mathscr{G}_\varepsilon^\nu \times \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon) \to \mathscr{G}^\nu, \qquad F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) = \lim_{J \to \infty} F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J). \tag{8.9}$$

Note that, apriori, the limit in (8.9) may not exist, but we prove in Lemma 8.1 that it does under our assumptions on $\mathfrak{R}_\varepsilon$ and $\boldsymbol{\delta}_\varepsilon$. When it will not lead to confusion, we will omit the $\mathfrak{R}, \boldsymbol{\delta}$ from the notation for $F_J^{\mathfrak{R}, \boldsymbol{\delta}}$.

When checking the probing maps definition we work with $\mathscr{G} := \mathscr{G}^\nu$ and $\mathscr{G}' := \mathscr{G}^{\nu-1}$, while $\mathscr{B}$ and $\mathscr{B}'$ will be the spaces of $\mathcal{C}^\nu$ and $\mathcal{C}^{\nu-1}$ symmetric two-tensors.

**Lemma 8.1.** *Let $\nu \geq 0$, $N : (0, 1) \to \mathbb{N}$ and $\{Q^{R, \delta}\}_{R, \delta}$ be a $(\nu, N)$-good family of perturbations. Then for all $\boldsymbol{\delta}_\varepsilon = \{\delta_j(\varepsilon)\}_{j=0}^\infty \subset (0, 1)$ and $\mathfrak{R}_\varepsilon = \{R_j(\varepsilon)\}_{j=0}^\infty \subset (0, 1)$ satisfying, (8.5). $\{(F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}, \infty)\}_{\varepsilon > 0}$ is a $\mathscr{G}^{\nu-1}$-family of probing maps for $\mathscr{G}^\nu$.*

*Moreover, for all $K \subset \mathscr{G}^\nu$ bounded and $\tilde{K} \subset \mathscr{G}^\nu$ bounded with $d(\tilde{K}^c, K) > 0$, there is $\varepsilon_1 > 0$ such that for all $0 < \varepsilon < \varepsilon_1$ and $\boldsymbol{\delta}_\varepsilon$ and $\mathfrak{R}_\varepsilon$ satisfying (8.5), $F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(K \times \boldsymbol{\Sigma}_J(\mathfrak{R}_\varepsilon)) \subset \tilde{K}$ for all $J \in \mathbb{N} \cup \{\infty\}$.*

*Proof.* Let $0 < \varepsilon < 1$ and assume (8.5) is satisfied. We first show that $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ is well defined.

Given $g \in \mathscr{G}^\nu_\varepsilon$ we claim that $F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J) \in K_{s(\varepsilon)}$ for all $J$ and $\hat{\boldsymbol{\sigma}}_J \in \boldsymbol{\Sigma}_J(\mathfrak{R}_\varepsilon)$. Indeed, we claim

$$d(F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J), \mathscr{G}^\nu \backslash K_{s(\varepsilon)}) \geq C_{s(\varepsilon)}\Big(2\varepsilon - \sum_{j=0}^J \delta_J(\varepsilon) R_J(\varepsilon)^{-\vartheta_\nu}\Big). \tag{8.10}$$

Note that when $J = 0$ this follows from the definition of $\mathscr{G}^\nu_\varepsilon$ since (8.2) and (7.2) yield $\|F_0^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) - g\|_{\mathcal{C}^\nu} \leq C_{s(\varepsilon)}\delta_0(\varepsilon) R_0(\varepsilon)^{-\vartheta_\nu}$. Next, assume the claim holds for all $j \leq J - 1$. Then, again by (8.2) we obtain

$$\|F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J) - F_{J-1}^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_{J-1})\|_{\mathcal{C}^\nu} = \|Q^{R_J(\varepsilon), \delta_J(\varepsilon)}(F_{J-1}^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_{J-1}), \boldsymbol{\sigma}_J) - F_{J-1}^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_{J-1})\|_{\mathcal{C}^\nu}$$
$$\leq C_{s(\varepsilon)}\delta_J(\varepsilon) R_J(\varepsilon)^{-\vartheta_\nu}\|\boldsymbol{\sigma}_J\|_{\ell^\infty}. \tag{8.11}$$

By (8.11) and (8.5) the claim in (8.10) holds and so $F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J) \in K_{s(\varepsilon)}$. Furthermore, (8.11) and (8.5) guarantee that $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma})$ is well defined and $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) \in \mathscr{G}^\nu$. Also, since $F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ is continuous, so is $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$.

Note that (8.11) implies the second statement in the Lemma and we need only check that $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ is a family of probing maps.

Next, note that (7.2) implies $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, 0) = g$ for all $g \in \mathscr{G}^\nu_\varepsilon$ so (2.6) holds. To check that (2.7), (2.8), (2.9) hold, let $K \subset \mathscr{G}^\nu$ be bounded. Then, there is $\varepsilon_0 > 0$ such that $K \subset \mathscr{G}^\nu_\varepsilon$ for all $0 < \varepsilon < \varepsilon_0$. In particular, (8.11) implies that for all $0 < \varepsilon < \varepsilon_0$, $g \in K$, and $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon)$,

$$\|F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) - g\|_{\mathcal{C}^\nu} \leq C_{s(\varepsilon)} \sum_{J=1}^\infty \delta_J R_J^{-\vartheta_\nu}\|\boldsymbol{\sigma}_J\|_{\ell^\infty} \leq C_{s(\varepsilon)}\varepsilon\|\boldsymbol{\sigma}\|_{\ell^\infty}.$$

Thus, since $\lim_{t \to 0^+} C_{s(\varepsilon)}\varepsilon = 0$, (2.7) holds.

We next prove that $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ satisfies (2.8). To ease notation, from now on we write $F_J$ in place of $F_J^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ and $\delta_J = \delta_J(\varepsilon)$, $R_J = R_J(\varepsilon)$. First, note that (8.2) implies that for $g \in K_{t(\varepsilon)}$ and $\boldsymbol{\sigma}, \boldsymbol{\mu} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R})$

$$\|F_0(g, \hat{\boldsymbol{\sigma}}_0) - F_0(g, \hat{\boldsymbol{\mu}}_0)\|_{\mathcal{C}^{\nu-1}} \leq C_{s(\varepsilon)}\delta_0 R_0^{-\vartheta_\nu}\|\boldsymbol{\sigma}_0 - \boldsymbol{\mu}_0\|_{\ell^\infty}.$$

Our induction hypothesis is that for some $J \geq 1$ and all $g \in \mathscr{G}^\nu_\varepsilon$, $\boldsymbol{\sigma}, \boldsymbol{\mu} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R})$,

$$\|F_{J-1}(g, \hat{\boldsymbol{\sigma}}_{J-1}) - F_{J-1}(g, \hat{\boldsymbol{\mu}}_{J-1})\|_{\mathcal{C}^{\nu-1}}$$
$$\leq C_{s(\varepsilon)} R_{J-1}^{-\vartheta_\nu-1}\delta_{J-1}\|\boldsymbol{\sigma}_{J-1} - \boldsymbol{\mu}_{J-1}\|_{\ell^\infty} + \sum_{j=0}^{J-2} C_{s(\varepsilon)} \prod_{k=j+1}^{J-1}(1 + C_{s(\varepsilon)}\delta_k R_k^{-\vartheta_\nu})R_j^{-\vartheta_\nu-1}\delta_j\|\boldsymbol{\sigma}_j - \boldsymbol{\mu}_j\|_{\ell^\infty}.$$

Next, note that by (8.2),

$$\|F_J(g, \hat{\boldsymbol{\sigma}}_J) - F_J(g, (\hat{\boldsymbol{\sigma}}_{J-1}, \boldsymbol{\mu}_J))\|_{\mathcal{C}^{\nu-1}} \leq C_{s(\varepsilon)} R_J^{\vartheta_\nu-1}\delta_J\|\boldsymbol{\sigma}_J - \boldsymbol{\mu}_J\|_{\ell^\infty}.$$

In addition, by (8.3) and the induction hypothesis

$$\|F_J(g, \hat{\boldsymbol{\mu}}_J) - F_J(g, (\hat{\boldsymbol{\sigma}}_{J-1}, \boldsymbol{\mu}_J))\|_{\mathcal{C}^{\nu-1}} \leq (1 + C_{s(\varepsilon)}\delta_J R_J^{-\vartheta_\nu})\|F_{J-1}(g, \hat{\boldsymbol{\sigma}}_{J-1}) - F_{J-1}(g, \hat{\boldsymbol{\mu}}_{J-1})\|_{\mathcal{C}^{\nu-1}}$$
$$\leq \sum_{j=0}^{J-1} \prod_{k=j+1}^{J} C_{s(\varepsilon)}(1 + C_{s(\varepsilon)}\delta_k R_k^{-\vartheta_\nu})R_j^{-\vartheta_\nu-1}\delta_j\|\boldsymbol{\sigma}_j - \boldsymbol{\mu}_j\|_{\ell^\infty}.$$

Therefore, for all $J = 0, 1, \dots,$ $g \in \mathscr{G}^\nu_\varepsilon$, and $\boldsymbol{\sigma}, \boldsymbol{\mu} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R})$,

$$\|F_J(g, \hat{\boldsymbol{\sigma}}_J) - F_J(g, \hat{\boldsymbol{\mu}}_J)\|_{\mathcal{C}^{\nu-1}}$$

$$\leq \|F_J(g, \hat{\boldsymbol{\sigma}}_J) - F_J(\kappa, (\hat{\boldsymbol{\sigma}}_{J-1}, \boldsymbol{\mu}_J))\|_{\mathcal{C}^{\nu-1}} + \|F_J(g, \hat{\boldsymbol{\mu}}_J) - F_J(g, (\hat{\boldsymbol{\sigma}}_{J-1}, \boldsymbol{\mu}_J))\|_{\mathcal{C}^{\nu-1}}$$

$$\leq C_{s(\varepsilon)} R_J^{-\vartheta_\nu - 1} \delta_J \|\boldsymbol{\sigma}_J - \boldsymbol{\mu}_J\|_{\ell^\infty} + \sum_{j=0}^{J-1} C_{s(\varepsilon)} \prod_{k=j+1}^{J} (1 + C_{s(\varepsilon)} \delta_k R_k^{-\vartheta_\nu}) R_j^{-\vartheta_\nu - 1} \delta_j \|\boldsymbol{\sigma}_j - \boldsymbol{\mu}_j\|_{\ell^\infty}.$$

Sending $J \to \infty$, we conclude that the inequality in (2.8) holds since $\lim_{t \to 0^+} C_{s(\varepsilon)}\varepsilon = 0$ and

$$\|F_\infty(g, \boldsymbol{\sigma}) - F_\infty(g, \boldsymbol{\mu})\|_{\mathcal{C}^{\nu-1}} \leq \sum_{j=0}^{\infty} C_{s(\varepsilon)} \prod_{k=j+1}^{\infty} (1 + C_{s(\varepsilon)} \delta_k R_k^{-\vartheta_\nu}) R_j^{-\vartheta_\nu - 1} \delta_j \|\boldsymbol{\sigma}_j - \boldsymbol{\mu}_j\|_{\ell^\infty}$$

$$\leq C_{s(\varepsilon)}\varepsilon \exp\left(C_{s(\varepsilon)}\varepsilon\right) \|\boldsymbol{\sigma} - \boldsymbol{\mu}\|_{\ell^\infty}.$$

The fact that $\tilde{F}^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}_\infty = \iota \circ F^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}_\infty : \mathscr{G}^\nu \times \boldsymbol{\Sigma}_\infty(\mathfrak{R}) \to \mathscr{G}^{\nu-1}$ is Lipschitz in both $\boldsymbol{\sigma}$ and $g$ follows after checking by induction that for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R})$

$$\|F_\infty(g, \boldsymbol{\sigma}) - F_\infty(\tilde{g}, \boldsymbol{\sigma})\|_{\mathcal{C}^{\nu-1}} \leq \prod_{j=0}^{\infty} (1 + C_{s(\varepsilon)} \delta_j R_j^{-\vartheta_\nu}) \|g - \tilde{g}\|_{\mathcal{C}^{\nu-1}} \leq \exp(C_{s(\varepsilon)}\varepsilon) \|g - \tilde{g}\|_{\mathcal{C}^{\nu-1}}.$$

We proceed to prove (2.9). We claim that

$$D_g F_\infty(g, \boldsymbol{\sigma}) = \lim_{J \to \infty} H_J(g, \hat{\boldsymbol{\sigma}}_J) \dots H_1(g, \hat{\boldsymbol{\sigma}}_1) H_0(g, \hat{\boldsymbol{\sigma}}_0), \qquad (8.12)$$

where we abbreviate $H_J(g, \hat{\boldsymbol{\sigma}}_J) := D_g Q^{R_J, \delta_J}(F_{J-1}(g, \hat{\boldsymbol{\sigma}}_{J-1}), \boldsymbol{\sigma}_J)$ for $J \geq 1$ and set $H_0(g, \hat{\boldsymbol{\sigma}}_0) = D_g Q^{R_0, \delta_0}(g, \hat{\boldsymbol{\sigma}}_0)$. Indeed, observe that for $J > K$

$$D_g F_J(g, \hat{\boldsymbol{\sigma}}_J) - D_g F_K(g, \hat{\boldsymbol{\sigma}}_K) = \left(H_J(g, \hat{\boldsymbol{\sigma}}_J) \dots H_{K+1}(g, \hat{\boldsymbol{\sigma}}_{K+1}) - I\right) D_g F_K(g, \hat{\boldsymbol{\sigma}}_K). \qquad (8.13)$$

Now, fix $\varepsilon_0 > 0$ and choose $K > 0$ such that $C_{s(\varepsilon)} \sum_{j=K+1}^{\infty} \delta_j R_j^{-\vartheta_\nu} < \varepsilon_0$. Then, using (7.4),

$$\|H_J(g, \hat{\boldsymbol{\sigma}}_J) \dots H_{K+1}(g, \hat{\boldsymbol{\sigma}}_{K+1}) - I\|_{\mathcal{C}^{\nu-1} \to \mathcal{C}^{\nu-1}}$$

$$\leq \sum_{j=K+1}^{J} \left\| H_J(g, \hat{\boldsymbol{\sigma}}_J) \dots H_{j+1}(g, \hat{\boldsymbol{\sigma}}_{j+1}) \left(H_j(g, \hat{\boldsymbol{\sigma}}_j) - I\right) \right\|_{\mathcal{C}^{\nu-1} \to \mathcal{C}^{\nu-1}}$$

$$\leq \sum_{j=K+1}^{J} \prod_{\ell=j+1}^{J} (1 + C_{s(\varepsilon)} \delta_\ell R_\ell^{-\vartheta_\nu}) C_{s(\varepsilon)} \delta_j R_j^{-\vartheta_\nu} \leq \exp(C_{s(\varepsilon)}\varepsilon)\varepsilon_0. \qquad (8.14)$$

Therefore, by (8.13), the sequence $\{D_g F_J(g, \hat{\boldsymbol{\sigma}}_J)\}_J$ is Cauchy. This implies that the limit in (8.12) exists and the convergence is uniform. Since $F_J \to F_\infty$, this implies the equality (8.12). Moreover, setting $K = -1$ in (8.14) and combining it with (8.12) we see that

$$\|D_g F_\infty - I\|_{\mathcal{C}^{\nu-1} \to \mathcal{C}^{\nu-1}} \leq C_{s(\varepsilon)} \exp(C_{s(\varepsilon)}\varepsilon) \sum_{j=0}^{\infty} \delta_j R_j^{-\vartheta_\nu} \leq C_{s(\varepsilon)}\varepsilon \exp(C_{s(\varepsilon)}\varepsilon) \to 0,$$

and hence the estimate in (2.9) holds.

$\square$

8.2. **Induction on the length of orbits.** In this section, we construct the probing family for which (8.1) holds for a predominant set of metrics. The idea is to show that for any $1 < \mathfrak{a} < 2$, given non-degeneracy of returning points at time $n$ for all $n \leq \ell$ (together with some additional conditions on the eigenvalues of Poincaré maps), most perturbations have non-degeneracy of returning points for times $n$ with $n \leq \mathfrak{a}\ell$. An outline is given in Section 4.2.

We start by obtaining non-degeneracy of orbits of length 1. In what follows, given $\delta_0(\varepsilon)$ and $R_0(\varepsilon)$ we write $F_0(g, \boldsymbol{\sigma}) := Q^{R_0(\varepsilon), \delta_0(\varepsilon)}(g, \boldsymbol{\sigma})$ as in (8.7).

**Lemma 8.2.** *Let $\nu > 0$ and $b > 0$, a function $N : (0,1) \to \mathbb{N}$, and a $(\nu, N)$-good family of perturbations $\{Q^{R,\delta}\}_{R,\delta}$. For all $\varepsilon > 0$, there exist positive constants $\delta_0, R_0, \beta_{0,0}$ such that $\delta_0 R_0^{-\vartheta_\nu} < \frac{1}{2}\varepsilon$ and the following holds. For every $F > 0$ and $(\Gamma, G)$ that is a $(F, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$, and all $K \subset G$ bounded in $\mathscr{G}^\nu$, there are $C > 0$ and $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$*

$$\sup_{\boldsymbol{\sigma} \in \Sigma(R_0)} \|F_0(g, \boldsymbol{\sigma}) - g\|_{\mathcal{C}^\nu} < \tfrac{1}{4}C\varepsilon, \qquad \text{for all } g \in K, \tag{8.15}$$

*and*

$$m_{\boldsymbol{\Sigma}(R_0)}\big(\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(R_0) : \ F_0(g, \boldsymbol{\sigma}) \in \tilde{L}_{F,0}\big) \leq \tfrac{1}{2}\varepsilon, \qquad \text{for all } g \in K \tag{8.16}$$

*where the set $\tilde{L}_{F,0}$ is defined by*

$$\tilde{L}_{F,0} := \big\{g \in \mathscr{G}^\nu : \ \exists q \in \mathbb{N} \ s.t. \ \Gamma \cap \mathscr{R}_F\big(1, \beta_{0,0,0}q^{-3}, g\big) \backslash \mathscr{N}_{q,F}\big(1, \beta_{0,0,0}, g\big) \neq \emptyset\big\}, \tag{8.17}$$

*with $\beta_{0,0,0} := 2\beta_{0,0}$.*

*Proof.* Let $\tilde{\vartheta}_4, \vartheta_5$ be as in Definition 7.2, and set

$$R_0 = \tfrac{1}{2}\varepsilon, \qquad \delta_0 = \min\big(\varepsilon^2 R_0^{\max(b, 2\tilde{\vartheta}_4)}, \ \varepsilon R_0^{\max(\vartheta_\nu, \vartheta_5+1, 1)}\big), \qquad \beta_{0,0} = \tfrac{1}{4}R_0^{\mathfrak{y}+1}\delta_0^{\mathfrak{L}}.$$

Next, fix $K \subset G$ bounded in $\mathscr{G}^\nu$. By Definition 7.2 there is $C_0 > 0$ such that $\|g - F_0(g, \boldsymbol{\sigma})\|_{\mathcal{C}^\nu} \leq C_0 \delta_0 R_0^{-\vartheta_\nu} \leq \tfrac{1}{2}C_0\varepsilon$ for $g \in K$ implying (8.15). Let $c, C$ be as in Corollary 7.7 (with $\mathfrak{y}' = \mathfrak{y}$) and note that its hypothesis for $n = 1$ are satisfied with $\delta := \delta_0$ and $R := R_0$ provided $\varepsilon < c$. Note that since $\mathscr{S}_F(1, \alpha, g) = \widetilde{S^*M} \supset \Gamma$ for all $g \in G$ and $\alpha > 0$, then

$$\big\{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(R_0) : \ F_0(g, \boldsymbol{\sigma}) \in \tilde{L}_{F,0}\big\} = \mathcal{S}_g^{R_0, \delta_0}(1, \alpha, 2\beta_{0,0}).$$

Let $\alpha = \tfrac{1}{2}c$. Finally, since $\mathfrak{y} \geq 1$, if the conclusion follows from Corollary 7.7 after asking $\varepsilon = 2R_0 < \min(c^2, c\delta_G, C^{-1})$.

$\square$

Now that we have obtained non-degeneracy for returning points with orbits of length 1, we need to induct on the length of the returning orbit. In order to do this, we work with lengths between $\mathfrak{a}^\ell F$ and $\mathfrak{a}^{\ell+1}F$ for some $1 < \mathfrak{a} < 2$ and induct on $\ell$. Before stating our lemma, we set

$$\mathfrak{b} := 1 - \log_2 \mathfrak{a} \ \in (0,1), \tag{8.18}$$

and for $j \geq 0$ and $\varepsilon > 0$ set

$$\gamma_j(\varepsilon) := 5^{-2\aleph}\varepsilon^{(2\aleph+1)\mathfrak{a}^{(j+1)/\mathfrak{b}}}\mathfrak{a}^{-4\aleph(j+1)/\mathfrak{b}}. \tag{8.19}$$

We are now ready to do the induction on the length of the orbits. At step $\ell$ in the induction, we study returning trajectories of length $t \leq F\mathfrak{a}^{\ell+1}$, in particular, making perturbations so that the returning trajectories with $(\mathfrak{a}^\ell - 1)F < t \leq \mathfrak{a}^{\ell+1}F$ are non-degenerate *and* so that the non-degeneracy of shorter trajectories is maintained. See the outline of the proof in Section 4 for a more detailed explanation of the argument. In what follows, we continue to use the notation $\boldsymbol{\Sigma}_\ell(\mathfrak{R})$ introduced in (8.6).

**Proposition 8.3.** *Let $\nu \geq 5$, $\mathfrak{y} \geq 1$, $b > 0$, $N : (0,1) \to \mathbb{N}$, and $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations. There exists $\boldsymbol{d} > 0$ and for all $\varepsilon > 0$ there are $\boldsymbol{\delta}_\varepsilon := \{\delta_\ell\}_{\ell=0}^\infty$, $\mathfrak{R}_\varepsilon := \{R_\ell\}_{\ell=0}^\infty$, and $\beta_{0,0} > 0$, such that $\delta_\ell R_\ell^{-\vartheta_\nu} < 2^{-\ell-1}\varepsilon$ and the following holds. For every $F > 0$, $(\Gamma, G)$ that is a $(F, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$, and $K \subset G$ bounded in $\mathscr{G}^\nu$, there are $C > 0$ and $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$ and all $\ell = 0, 1, \ldots$*

$$\sup_{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\ell+1}(\mathfrak{R})} \|F_{\ell+1}(g, \boldsymbol{\sigma}) - F_\ell(g, \widehat{\boldsymbol{\sigma}}_\ell)\|_{\mathcal{C}^\nu} \leq C2^{-\ell-2}\varepsilon, \qquad g \in K. \tag{8.20}$$

*In addition, for $0 \leq i \leq j \leq \ell$, there exist constants $\beta_{i,j} \in (0,1]$, with $\beta_{i+1,j} \leq \beta_{i,j}$,*

$$\beta_{j+1,j+1} = \beta_{j,j}^{\mathfrak{c}} \varepsilon^{\boldsymbol{d}\mathfrak{a}^{j/\mathfrak{b}}}, \tag{8.21}$$

*with $\mathfrak{c} := \max(\frac{2\aleph(2\aleph+3)}{2\aleph+1}m_\nu, 2\aleph\mathfrak{y})$ and $m_\nu := \max(b, \vartheta_\nu, \vartheta_5 + 1, 2\tilde{\vartheta}_4, 1)$ such that the following holds. Let $\alpha_0 = 4\varepsilon^{-1}$ and for $0 \leq j \leq \ell$ set*

$$\begin{aligned}
\alpha_{j+1} &:= \gamma_j \beta_{j,j}^{2\aleph}, & \alpha_{j,\ell} &:= (1 - 2^{-(\ell-j)-1})\alpha_j, & \beta_{i,j,\ell} &:= (1 + 2^{-(\ell-j)})\beta_{i,j}, \\
\tilde{\beta}_{j,\ell+1} &:= s_\ell \tilde{\beta}_{j,\ell}, & \tilde{\beta}_{j,j} &:= \beta_{j,j}^{2\aleph}\varepsilon^{2\aleph-1}, & s_\ell &:= \frac{1+2^{-\ell-1}}{1+2^{-\ell}}
\end{aligned} \tag{8.22}$$

*with $\gamma_j$ as in (8.19). Then for all $g \in K$, $\ell \geq 0$, and $0 < \varepsilon < \varepsilon_0$*

$$m_{\boldsymbol{\Sigma}_\infty(\mathfrak{R})}\Big(\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}) : \exists j \leq \ell \ s.t. \ F_j(g, \widehat{\boldsymbol{\sigma}}_j) \in L_{F,j}\Big) \leq (1 - 2^{-\ell-1})\varepsilon, \tag{8.23}$$

*and, for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\ell-1}(\mathfrak{R})$ such that $F_{\ell-1}(g, \boldsymbol{\sigma}) \in \mathscr{G}^\nu \setminus L_{F,\ell-1}$*

$$m_{\Sigma(R_\ell)}\Big(\boldsymbol{\sigma}_\ell : F_\ell(g, (\boldsymbol{\sigma}, \boldsymbol{\sigma}_\ell)) \in L_{F,\ell}\Big) \leq 2^{-\ell-1}\varepsilon. \tag{8.24}$$

*Here, $L_{F,-1} = \emptyset$, $F_{-1} = \pi_{\mathscr{G}^\nu}$, and for $\ell \geq 0$ the set $L_{F,\ell}$ is defined as follows: $g \in \mathscr{G}^\nu \setminus L_{F,\ell}$*

(1) *For $0 \leq i \leq \ell$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$,*

$$\Gamma \cap \mathscr{R}_F(n, \tilde{\beta}_{i,\ell}, g) \subset \mathscr{N}_F(n, \tilde{\beta}_{i,\ell}, g). \tag{8.25}$$

(2) *For all $\mathfrak{b}\ell \leq j \leq \ell$, $0 \leq i \leq j$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$, and all $1 \leq q \leq \mathfrak{a}^{j/\mathfrak{b}-i+1}$,*

$$\Gamma \cap \mathscr{R}_F(n, \beta_{i,j,\ell}q^{-3}, g) \cap \mathscr{S}_F(n, \alpha_{j,\ell}, g) \subset \mathscr{N}_{q,F}(n, \beta_{i,j,\ell}, g). \tag{8.26}$$

**Remark 8.4.** There are a large number of parameters in Proposition 8.3. One should think of these parameters as follows. Those with a $\beta$ control the degree of $q$-non-degeneracy that trajectories have while those with a $\tilde{\beta}$ control simple non-degeneracy of trajectories. Parameters with an $\alpha$ control the degree of simplicity that a trajectory must have in order to apply the inductive hypothesis.

The most important constants will turn out to be $\mathfrak{c}$ and $\mathfrak{a}$. While $1 < \mathfrak{a} < 2$ is left free for the moment, and will be fixed in the proof of Proposition 8.9, $\mathfrak{c}$ controls how rapidly the

non-degeneracy of trajectories decays. This constant is determined in Step 3 of the induction (Section 8.2.4).

Notice that the parameter $\mathfrak{c}$ depends on $m_\nu$, which depends on the regularity $\nu$ in which we work. The reason for definition of $m_\nu$ is as follows: 1) $\vartheta_\nu$ appears because we want our perturbations to be close in $\mathscr{G}^\nu$ (see Lemma 8.7), 2) $b$, $2\tilde{\vartheta}_4$, and $\vartheta_5 + 1$ appear because we need to apply Corollary 7.7 (see Step 3 of the induction (Section 8.2.4)

Given $\ell$, the proposition yields that for all the perturbations $F_{\ell-1}^{\delta_\varepsilon, \mathfrak{R}_\varepsilon}(g, \boldsymbol{\sigma})$ that satisfy both (8.25) and (8.26) the set of $\boldsymbol{\sigma}_\ell$ that would yield a new perturbation $F_\ell^{\delta_\varepsilon, \mathfrak{R}_\varepsilon}(g, \boldsymbol{\sigma}, \boldsymbol{\sigma}_\ell)$ not satisfying either (8.25) or (8.26) is bounded in volume by $2^{-(\ell+1)}\varepsilon$. This allows us to show in the proof of Theorem 1.3 that, except for on a measure $\varepsilon$ set of $\boldsymbol{\sigma}$, the perturbations $F_\infty^{\delta_\varepsilon, \mathfrak{R}_\varepsilon}(g, \boldsymbol{\sigma})$ satisfy (8.25) for all $n$.

We next explain the reason why we need to include (8.26) in the statement (this is also explained carefully in Section 4). We prove the proposition by induction in $\ell$. Suppose we could perturb $F_{\ell-1}^{\delta_\varepsilon, \mathfrak{R}_\varepsilon}(g, \boldsymbol{\sigma})$ in such a way that orbits that start in $\Gamma$ and and are $(n, \beta_{\ell,\ell})$ returning with $n \in (\mathfrak{a}^{\ell-1}, \mathfrak{a}^\ell]$ will also be $(n, \beta_{\ell,\ell})$ non-degenerate. Then, applying (8.25) to get non-degeneracy of $(n, \beta_{\ell,\ell})$ returning orbits with $n \in (0, \mathfrak{a}^{\ell-1}]$ would finish the job. However, this cannot always be done. The issue is that we can only arrange for most of the perturbations to yield $(n, \beta_{\ell,\ell})$ non-degenerate orbits when the orbits are $(n, \alpha_{\ell,\ell})$ simple with $\beta_{\ell,\ell}$ small in terms of $\alpha_{\ell,\ell}$ (see Corollary 7.7). If all $(n, \beta_{\ell,\ell})$ returning orbits were indeed $(n, \alpha_{\ell,\ell})$ simple, then (8.25) applied to $\ell - 1$ would guarantee that orbits that start at $\Gamma$ and are $(n, \beta_{\ell,\ell})$ returning with $n \in [0, \mathfrak{a}^\ell]$ will be $(n, \beta_{\ell,\ell})$ non-degenerate as desired. In reality, some returning orbits will not be $(n, \alpha_{\ell,\ell})$ simple, and so those we will view as iterates of shorter 'loops' of length $n/q \in (\mathfrak{a}^{i-1}, \mathfrak{a}^{i-1}]$ with $q > 1$ that are $(n/q, C^n\alpha_{\ell,\ell})$ returning (see Lemma 6.4). If the shorter 'loop' is $(n/q, \alpha_{i,\ell})$ simple and $C^n\alpha_{\ell,\ell} < \beta_{i,\ell}q^{-3}$, then we can apply (8.26) to say that such an orbit is $(n/q, \beta_{i,\ell})$ $q$ non-degenerate and use that to infer that the long orbit is actually $(n, \beta_{\ell,\ell})$ non-degenerate (see Lemma 6.18). If the shorter loop is not $(n/q, \alpha_{i,\ell})$ simple, then we view it as an iterate of a shorter loop and repeat the process until it terminates. The reason for the third index in $\beta_{i,j,\ell}$ is that we will need some extra non-degeneracy which is 'used up' by making additional small perturbations in future inductive steps.

We divide the proof of Proposition 8.3 into steps. In Step 0 we show that the base case $\ell = 0$ holds. In Step 1 we deal with the returning orbits that are not simple enough to get non-degeneracy by perturbation, instead showing that they inherit the non-degeneracy by decomposing them into shorter loops as described above. In Step 2 we prove that the non-degeneracy created up to step $\ell$ is preserved by the perturbation performed in the $\ell + 1$ step. Finally, in Step 3, we prove that the volume of the set of perturbations that yield orbits whose degeneracy we cannot control is small.

8.2.1. *Step 0: Setting up the induction argument.* Let $(\Gamma, G)$ be a $(\mathcal{F}, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$, and $K \subset G$ bounded. Let $\tilde{K} \subset G$ be bounded and satisfy $K \subset \tilde{K}$ with $d(K, \tilde{K}^c) > 0$. By Lemma 8.1 there is $\varepsilon_1 > 0$ such that $\tilde{K} \subset \mathscr{G}_\varepsilon^\nu$ for $0 < \varepsilon < \varepsilon_1$, where $\mathscr{G}_\varepsilon^\nu$ is as in (8.4), and for

any choice of $\boldsymbol{\delta}_\varepsilon$ and $\mathfrak{R}_\varepsilon$ satisfying (8.5), we have

$$F_J^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(K \times \boldsymbol{\Sigma}_J(\mathfrak{R}_\varepsilon)) \subset \tilde{K}, \qquad 0 < \varepsilon < \varepsilon_1, \;\; J \in \mathbb{N} \cup \{\infty\}.$$

Since all the metrics we work with will be of the form $F_J^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g, \hat{\boldsymbol{\sigma}}_J)$, for some $g \in K$, they will lie inside $\tilde{K}$ and hence have uniform estimates.

Let $\tilde{\varepsilon}_0$ be the number $\varepsilon_0$ found in Lemma 8.2. From now on we assume that $\varepsilon_0 \leq \min\{\tilde{\varepsilon}_0, \varepsilon_1\}$. For convenience, in what follows we work with

$$C_0 = 2 \max\{1, C, C_1, C_2, C_3, C_4, \delta_G^{-1}, C_G\}, \tag{8.27}$$

where $C$ is the constant from Corollary 7.7, $C_1$ is the constant from Lemma 6.4, $C_2 > 0$ is such that $\|\varphi_t^g - \varphi_t^{\tilde{g}}\|_{\mathcal{C}^1} \leq C_2^t \|g - \tilde{g}\|_{\mathcal{C}^3}$ for all $g, \tilde{g} \in \tilde{K}$ and $t \in \mathbb{R}$, $C_3$ is as in Lemma 6.18, and $C_4$ is the maximum of the constants $C$ in Definition 7.2 for the choices of $\nu$ in $\{3, 5, \nu\}$, and $C_G, \delta_G$ are from Lemma 6.12 and (6.11). Note that $C_0$ depends only on $\nu, N, \mathfrak{y}, b, \mathfrak{a}, F, \Gamma, G, K$. During the induction argument we will ask that $\varepsilon_0$ be small in terms of powers of $C_0^{-1}$.

We prove the lemma by induction on $\ell$. We check the cases $0 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$. Since $\mathfrak{a}^{\lceil \log_\mathfrak{a} 2 \rceil - 1} < 2$, this amounts to considering $0 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$, $n = 1$, $0 \leq i \leq \ell$ in (8.25) and $\mathfrak{b}\ell \leq j \leq \ell$, $0 \leq i \leq j$, and $1 \leq q \leq \mathfrak{a}^{j/\mathfrak{b}-i+1}$ in (8.26).

By Lemma 8.2, for all $\varepsilon > 0$ there are $\beta_{0,0}$, $R_0$, and $\delta_0$ such that such that (8.15) holds and (8.16) holds for all $g \in K$ and $0 < \varepsilon < \varepsilon_0(\Gamma, K)$. ,Set $\delta_\ell(\varepsilon) = 0$, $R_\ell(\varepsilon) = 1$ for $1 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$. That is,

$$F_0^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g, \sigma_0) = F_\ell^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g, (\sigma_0, \sigma_1, \ldots, \sigma_\ell)), \qquad 1 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1. \tag{8.28}$$

Then, (8.15) implies (8.20) for $\ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$.

Set $\beta_{i,j} = \beta_{i,i}$ for $0 \leq i \leq j \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$. We note that by (8.28), it is enough to show that

$$m_{\Sigma(R_0)}\Big(\sigma \,:\, F_0(g,\sigma) \in \bigcup_{0 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1} L_{F,\ell}\Big) \leq 2^{-1}\varepsilon.$$

We next claim that for $\ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$, and $g \in K$,

$$\Big\{\sigma \,:\, F_0^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\sigma) \in \bigcup_{0 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1} L_{F,\ell}\Big\} \subset \Big\{\sigma \,:\, F_0^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\sigma) \in \tilde{L}_{F,0}\Big\}$$

where $\tilde{L}_{F,0}$ is defined in Lemma 8.2.

Indeed, since $\tilde{\beta}_{0,0} = \beta_{0,0}^{2\aleph}\varepsilon^{2\aleph-1}$ and $\beta_{0,0,0} = 2\beta_{0,0}$ for $\varepsilon > 0$ small enough depending on $K$ and all $g \in K$,

$$\mathcal{N}(1, \tilde{\beta}_{0,0}, F_0(g,\sigma)) \subset \mathcal{N}_{1,F}(1, \beta_{0,0,0}, F_0(g,\sigma)).$$

Therefore, by (8.17), if $F_0^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\sigma) \notin \tilde{L}_{F,0}$, then (8.25) holds for $0 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$ and $g$ replaced by $F_0^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\sigma)$.

Next, since $\beta_{i,j,\ell} \leq \beta_{0,0,0}$ for $i = 0$, $\mathfrak{b}\ell \leq j \leq \ell$, and $0 \leq \ell \leq \lceil \log_\mathfrak{a} 2 \rceil - 1$, we have that

$$\big\{g \in \mathscr{G}^\nu \,:\quad \Gamma \cap \mathscr{R}_F(1, \beta_{0,0,0}q^{-3}, g) \cap \mathscr{S}_F(n, \alpha_{j,\ell}, g) \backslash \mathscr{N}_{q,F}(1, \beta_{0,0,0}, g) \neq \emptyset\big\} \subset \tilde{L}_{F,0}$$

and hence, if $F_0^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \sigma) \notin \tilde{L}_{F,0}$, then (8.26) holds for $0 \leq \ell \leq \lceil \log_{\mathfrak{a}} 2 \rceil - 1$ and $g$ replaced by $F_0^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \sigma)$. The claim in (8.23) then follows from (8.16). This concludes the proof of the base case in the induction argument.

Note that with $\beta_{0,0}$ in place, the constants $\beta_{\ell,\ell}$ and $\alpha_\ell$ are defined as in (8.21) and (8.22) for all $\ell$. For $\varepsilon > 0$ and $\ell = \lceil \log_{\mathfrak{a}} 2 \rceil, \lceil \log_{\mathfrak{a}} 2 \rceil + 1, \ldots$ we set

$$R_\ell(\varepsilon) := \tfrac{1}{8} \varepsilon^{\mathfrak{a}^{\ell+1}} \alpha_\ell, \qquad \delta_\ell(\varepsilon) := \min \left( \varepsilon^{\mathfrak{a}^{\ell+1}} R_\ell^{\max(b, 2\tilde{\vartheta}_4)}, \, \varepsilon 2^{-\ell} R_\ell^{\max(\vartheta_\nu, \vartheta_5 + 1, 1)} \right). \tag{8.29}$$

We note that $\delta_\ell(\varepsilon) R_\ell(\varepsilon)^{-\vartheta_\nu} < 2^{-\ell-1} \varepsilon$ and so $\boldsymbol{\delta}_\varepsilon = \{\delta_\ell(\varepsilon)\}_{\ell=0}^\infty$ and $\mathfrak{R}_\varepsilon = \{R_\ell(\varepsilon)\}_{\ell=0}^\infty$ satisfy (8.5). We also note that with these definitions there is $\varepsilon_0$ small enough that the assumptions (7.19) in Corollary 7.7 are satisfied with $R = R_{\ell+1}(\varepsilon)$, $\delta = \delta_{\ell+1}(\varepsilon)$, $\alpha = \tfrac{1}{2} \alpha_{\ell+1}$, and $n \leq \mathfrak{a}^{\ell+1}$.

We now start the inductive step. Note that the base case $0 \leq \ell \leq \lceil \log_{\mathfrak{a}} 2 \rceil - 1$ only covers $n = 1$ since $\mathfrak{a}^\ell < n \leq \mathfrak{a}^{\ell+1}$ and $\mathfrak{a} < 2$. Since to get to $n = 2$ we need $2 \leq \mathfrak{a}^{\ell+1}$, we may assume that the inductive hypotheses hold for some $\ell \geq \lceil \log_{\mathfrak{a}} 2 \rceil - 1$. We split the proof of the inductive step with $\ell + 1$ into three steps.

### 8.2.2. *Step 1: Nondegeneracy of returning points that are not simple.* The goal of this step is to study the set of points in $\Gamma$ that under the perturbed metric $F_\ell(g, \sigma)$ generate orbits that return to $\Gamma$ at some 'discrete' time $n \in (\mathfrak{a}^\ell, \mathfrak{a}^{\ell+1}]$ but that are not simple enough that the $F_{\ell+1}$ perturbations would make them non-degenerate. As explained before, we decompose these returning orbits into shorter 'loops' and we use the $(\beta_q, q)$ non-degeneracy of shorter orbits given by (8.26) to show that the original orbits were already non-degenerate.

We note that for this step in the proof the exact powers in the definition of $\beta_{j,j}$ do not play a role. Instead, we only use that $\beta_{j,j} \geq \beta_{j+1,j+1}$ and $\beta_{i+1,j} \leq \beta_{i,j}$ for all $i$ and $j$. The precise definition of $\gamma_j$ as well as the definition of $\alpha_{j+1}$ in terms of $\gamma_j$ and $\beta_{j,j}$ do, however, play a role.

**Lemma 8.5.** *There is $\varepsilon_0 > 0$ depending only on $\nu, N, \mathfrak{y}, b, \mathfrak{a}, F, \Gamma, G, K$ so that the following holds. Suppose the conclusions of Proposition 8.3 hold up to the index $\ell$. Then, for all $0 < \varepsilon < \varepsilon_0$, $\mathfrak{a}^\ell < n \leq \mathfrak{a}^{\ell+1}$ and $(g, \sigma) \in K \times \boldsymbol{\Sigma}_\ell(\mathfrak{R}_\varepsilon)$ such that $g_\ell := F_\ell(g, \sigma) \in \mathscr{G}^\nu \backslash L_{T,\ell}$ we have*

$$\Gamma \cap \mathscr{R}_F(n, \alpha_{\ell+1}, g_\ell) \setminus \mathscr{S}_F(n, \alpha_{\ell+1}, g_\ell) \subset \mathscr{N}_F(n, \alpha_{\ell+1}, g_\ell). \tag{8.30}$$

*Proof.* Let $\mathfrak{a}^\ell < n \leq \mathfrak{a}^{\ell+1}$ and $(g, \sigma) \in K \times \boldsymbol{\Sigma}_\ell(\mathfrak{R})$ such that $g_\ell := F_\ell(g, \sigma) \in \mathscr{G}^\nu \backslash L_{F,\ell}$. Let

$$\rho \in \Gamma \cap \mathscr{R}_F(n, \alpha_{\ell+1}, g_\ell) \setminus \mathscr{S}_F(n, \alpha_{\ell+1}, g_\ell). \tag{8.31}$$

We divide the proof into fours steps. In Step A, we decompose the non-simple orbit associated to $\rho$ into shorter $(k_m, \delta_m(\ell))$ returning orbits that are $(k_m, \alpha_m(\ell))$ simple. The returning times $t_m$ associated to these orbits are such that $t_m$ divides $t_{m-1}$, and $k_m = \lceil \frac{t_m}{F} \rceil$ with $k_0 = n$. In Step B we show that $\rho \in \mathscr{N}_{q_m, F}(k_m, \beta_m(\ell), g_\ell)$ with $q_m := t_0/t_m \in \mathbb{Z}$. In Step C we prove that there are $\mathcal{I}$ and $p \in \mathbb{Z}$ such that for all $q \leq q_m$ there is a time $t_{q,m}$ sufficiently close to $q t_m$ with the property that $(\mathcal{P}_\mathcal{I}^p[g_\ell])^q(\rho) = \mathcal{P}_\mathcal{I}^{qp}[g_\ell](\rho) = \varphi_{t_{q,m}}^{g_\ell}(\rho)$. In particular, one can control the distance from $(\mathcal{P}_\mathcal{I}^p[g_\ell])^q(\rho)$ to $\rho$ for $q \leq q_m$. In Step D we check that the above guarantees that the hypothesis of Lemma 6.18 are satisfied and so $\rho$ is $(n, \alpha_{\ell+1})$ non-degenerate as claimed.

*Step A.* We start the proof by showing that there are $1 \leq m \leq \lfloor (\ell+1) \log_2 \mathfrak{a} \rfloor$, $\{t_i\}_{i=0}^m \subset \mathbb{R}$, and $\{k_i\}_{i=0}^m \subset \mathbb{Z}$, such that $k_0 = n$, $(k_i - 1)F < t_i \leq k_i F$, $t_i$ divides $t_{i-1}$ for $1 \leq i \leq m$, $t_i > F$,

$$\rho \in \Gamma \cap \mathscr{R}_F(k_m, \delta_m(\ell), g_\ell) \cap \mathscr{S}_F(k_m, \alpha_m(\ell), g_\ell), \qquad d(\varphi_{t_m}^{g_\ell}(\rho), \rho) \leq \alpha_m(\ell), \qquad (8.32)$$

and

$$d(\varphi_t^{g_\ell}(\rho), \varphi_{\mathfrak{r}}^{g_\ell}(\rho)) \leq \delta_i(\ell), \qquad 0 \leq t \leq t_{i-1}, \quad t = qt_i + \mathfrak{r}, \quad 0 \leq \mathfrak{r} < t_i. \qquad (8.33)$$

Here, for $i \geq 1$

$$\alpha_i(\ell) := \alpha_{\ell-i+1}, \qquad \delta_0(\ell) = \alpha_0(\ell), \qquad \delta_i(\ell) := C_0^{t_{i-1}} \alpha_{i-1}(\ell),$$

with $C_0$ as in (8.27).

Note that $\rho \in \Gamma \cap \mathscr{R}_F(k_0, \delta_0(\ell), g_\ell) \setminus \mathscr{S}_F(k_0, \alpha_0(\ell), g_\ell)$ by (8.31). We next show that, by splitting the orbit into shorter returning ones, we can find $\{t_i\}_{i=0}^{m-1}$ such that $t_{i+1}$ divides $t_i$ and $\rho \in \Gamma \cap \mathscr{R}_F(k_i, \delta_i(\ell), g_\ell) \setminus \mathscr{S}_F(k_i, \alpha_i(\ell), g_\ell)$ with $k_i := \lceil \frac{t_i}{F} \rceil$, and that both (8.32) and (8.33).

To do this, suppose we have found $\{t_i\}_{i=0}^{m-1}$ such that $t_{i+1}$ divides $t_i$, and, with $k_i := \lceil \frac{t_i}{F} \rceil$,

$$\rho \in \mathscr{R}_F(k_i, \delta_i(\ell), g_\ell), \qquad (8.34)$$
$$\rho \notin \mathscr{S}_F(k_i, \alpha_i(\ell), g_\ell), \qquad (8.35)$$

and (8.33) hold for $0 \leq i \leq m-1$. We prove by induction that there is $t_m$ dividing $t_{m-1}$ such that (8.34) holds with $i = m$ and (8.33) holds for $1 \leq i \leq m$. Note that the base case $i = 0$ follows from (8.31) after letting $t_0$ be such that $n = \lceil \frac{t_0}{F} \rceil$ since $k_0 = n$.

First, note that $\delta_i(\ell) \leq \alpha_i(\ell)$ for $1 \leq i \leq m-1$. Indeed, since $\{\beta_{j,j}\}_j$ is decreasing it is straightforward to check that $\delta_i(\ell) \leq \alpha_i(\ell)$ provided $C_0^{t_{i-1}} \gamma_{\ell-i-1} \leq \gamma_{\ell-i-2}$. Since $\mathfrak{a}^{-4\aleph/\mathfrak{b}} \leq 1$, this reduces to checking that $C_0^{t_{i-1}} \varepsilon^{(2\aleph+1)\mathfrak{a}^{(\ell-i+1)/\mathfrak{b}}(\mathfrak{a}^{1/\mathfrak{b}}-1)} \leq 1$. Since $t_{i-1} \leq k_{i-1} F \leq \mathfrak{a}^{\ell+1} 2^{-(i-1)} F$, the claim follows from letting $\varepsilon_0$ be small.

Next, since (8.34) and (8.35) hold with $i = m-1$ and $\delta_i(\ell) \leq \alpha_i(\ell)$ for $1 \leq i \leq m-1$, using Lemma 6.6 together with $C_\Gamma > 1$ we may apply Lemma 6.4 with $C_F := C_\Gamma > 1$, $\delta := \delta_{m-1}(\ell)$, $\alpha := \alpha_{m-1}(\ell)$, $t_0 := t_{m-1}$, to obtain the existence of $t_m \in [C_\Gamma F - C_0^{t_{m-1}} \alpha_{m-1}(\ell), \frac{1}{2}t_{m-1}]$ dividing $t_{m-1}$ such that $d(\varphi_{t_m}^{g_\ell}(\rho), \rho) \leq \delta_m(\ell)$ and for $0 \leq l < t_{m-1}$ with $l = qt_m + \mathfrak{r}$ and $0 \leq \mathfrak{r} < t_m$

$$d(\varphi_l^{g_\ell}(\rho), \varphi_{\mathfrak{r}}^{g_\ell}(\rho)) = \delta_m(\ell). \qquad (8.36)$$

Note that, provided $\varepsilon > 0$ was chosen small enough small enough (uniformly in $K$, $\ell$), and using that $C_\Gamma > 1$, we have $t_m > F$. If $\rho \notin \mathscr{S}_F(k_m, \alpha_m(\ell), g_\ell)$, the inductive step $i = m$ is complete. On the other hand, if $\rho \in \mathscr{S}_F(k_m, \alpha_m(\ell), g_\ell)$ then we have proved (8.32) and (8.33).

At this point, we have proved either that there is $m \leq \lfloor (\ell+1) \log_2 \mathfrak{a} \rfloor$ such that (8.32) and (8.33) hold or there are $\{t_i\}_{i=1}^N$, with $N > \lfloor (\ell+1) \log_2 \mathfrak{a} \rfloor$ such that $t_i$ divides $t_{i-1}$, $t_i > F$, and with $k_i = \lceil \frac{t_i}{F} \rceil$, (8.33), (8.34), (8.35) hold for all $i$. We claim that the second alternative is not possible. Indeed, note that since $t_i$ divides $t_{i-1}$,

$$F < t_i \leq 2^{-i} t_0 \leq 2^{-i} \mathfrak{a}^{\ell+1} F,$$

and this is not possible for $i > \lfloor (\ell+1) \log_2 \mathfrak{a} \rfloor$ since then $2^{-i} \mathfrak{a}^{\ell+1} \leq 1$. It follows that the claims (8.32) and (8.33) are true with $m$ as stated.

*Step B.* Let $q_m = t_0/t_m$ and $\beta_m(\ell) = \beta_{i_m,j_m}$ with $i_m = \lceil \log_{\mathfrak{a}} k_m \rceil$, $j_m = \ell - m + 1$. We claim

$$\rho \in \mathscr{N}_{q_m,F}(k_m, \beta_m(\ell), g_\ell). \tag{8.37}$$

The objective is to show that we can apply the induction hypothesis (8.26) with $(n, q, i, j, \ell) := (k_m, q_m, i_m, j_m, \ell)$. To this end, we first observe that by definition $\mathfrak{a}^{i_m-1} < k_m \le \mathfrak{a}^{i_m}$. In addition, $k_m \le 2^{-m}\mathfrak{a}^{\ell+1} \le \mathfrak{a}^{\ell-m+1}$ and so $i_m \le j_m$. Also, since $m \le \lfloor (\ell+1)(1-\mathfrak{b}) \rfloor$, we have $\mathfrak{b}(\ell+1) \le j_m \le \ell$. In particular, $\frac{t_0}{F} \le \mathfrak{a}^{\ell+1} \le \mathfrak{a}^{\frac{j_m}{\mathfrak{b}}}$. Since $\frac{t_m}{F} \ge \mathfrak{a}^{i_m-1}$, this also yields $q_m = \frac{t_0}{t_m} \le \mathfrak{a}^{j_m/\mathfrak{b}-i_m+1}$.

We first claim that

$$\delta_m(\ell) \le \beta_m(\ell) q_m^{-3}.$$

Indeed, using that $\beta_{j_m,j_m} \le \beta_{i_m,j_m}$ (since $i_m \le j_m$) and $q_m \le \mathfrak{a}^{\ell+1}$, this reduces to checking that $\mathfrak{a}^{3(\ell+1)} C_0^{t_{m-1}} \alpha_{j_m+1} \le \beta_{j_m,j_m}$. Furthermore, since $\beta_{j_m,j_m} \le 1$, the claim reduces to showing that $\mathfrak{a}^{3(\ell+1)} C_0^{t_{m-1}} \gamma_{j_m} \le 1$, and this is equivalent to

$$5^{-2\aleph} \le C_0^{-\mathfrak{a}^{\ell+1}2^{-m+1}} \mathfrak{a}^{4\aleph(\ell-m+2)/\mathfrak{b}-3(\ell+1)} \varepsilon^{-(2\aleph+1)\mathfrak{a}^{(\ell-m+2)/\mathfrak{b}}}.$$

The claim then follows from the facts that $\varepsilon_0 \le C_0^{-1}$ and $(\ell-m+2)/\mathfrak{b} \ge \ell + 1$.

Since $\delta_m(\ell) \le \beta_m(\ell) q_m^{-3}$ and $\beta_{i_m,j_m,\ell} = (1 + 2^{-m})\beta_m(\ell)$, we conclude that $\mathscr{R}_F(k_m, \delta_m(\ell), g_\ell) \subset \mathscr{R}_F(k_m, \beta_{i_m,j_m,\ell} q_m^{-3}, g_\ell)$. Therefore, the claim in (8.37) follows from combining (8.32) and (8.26) with $(n, i, j) = (k_m, i_m, j_m)$. This can be done since $\beta_{i_m,j_m,\ell} \ge \beta_m(\ell)$ and $\mathscr{S}_F(k_m, \alpha_m(\ell), g_\ell) \subset \mathscr{S}_F(k_m, \alpha_{j_m,\ell}, g_\ell)$ because $\alpha_{j_m,\ell} = (1 - 2^{-m})\alpha_m(\ell)$.

*Step C.* Recall that $k_m = \lceil \frac{t_m}{F} \rceil$ and $t_0 = q_m t_m$ with $k_0 = n$. Let $C_G$, $\delta_G$ be as in Lemma 6.12 and (6.11). Let

$$\delta_{m,\ell} := m C_0^{F \mathfrak{a}^{\ell+1}} \alpha_{m-1}(\ell). \tag{8.38}$$

We next show there are $\mathcal{I}$ and $p$ with $\boldsymbol{T}_{\mathcal{I}}^{(p)}[g_\ell](\rho) \in \big[ (k_m - 1)F - C_G \delta_G, \ k_m F + C_G \delta_G \big]$,

$$\big\{ \mathcal{P}_{\mathcal{I}}^{(qp)}[g_\ell](\rho) \big\}_{q=0}^{q_m} \subset \mathcal{D}_{\mathcal{I}}^{(p)}[g_\ell], \qquad \max_{0 \le q \le q_m} d\Big( \big( \mathcal{P}_{\mathcal{I}}^{(p)}[g_\ell] \big)^q(\rho), \rho \Big) \le C_G \delta_{m,\ell},$$

$$\mathcal{P}_{\mathcal{I}}^{(qp)}[g_\ell](\rho') = \big( \mathcal{P}_{\mathcal{I}}^{(p)}[g_\ell] \big)^q(\rho'), \qquad 0 \le q \le q_m, \qquad \rho' \text{ in a neighborhood of } \rho. \tag{8.39}$$

Let $0 \le q \le q_m$. Since $\{ \alpha_j(\ell) \}_j$ is decreasing, it follows from Lemma 8.6 below that

$$d(\varphi_{q t_m}^{g_\ell}(\rho), \rho) \le \delta_{m,\ell}. \tag{8.40}$$

Let $\varepsilon_0$ be small enough that $\delta_{m,\ell} \le \delta_G$ for any choice of $m, \ell$. Then, since $\rho \in \Gamma$, and $\delta_{m,\ell} \le \delta_G$ by (6.8) and (6.9) there exists $t_{q,m}$ with such that

$$|t_{q,m} - q t_m| < C_G \delta_{m,\ell}, \qquad \varphi_{t_{q,m}}^{g_\ell}(\rho) \in \tilde{\Gamma}, \qquad d(\varphi_{t_{q,m}}^{g_\ell}(\rho), \rho) < C_G \delta_{m,\ell}. \tag{8.41}$$

By Remark 6.10 there are $p$ and $\tilde{\mathcal{I}}$ such that $\rho \in \mathcal{D}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]$,

$$\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell](\rho) = \varphi_{t_{1,m}}^g(\rho), \qquad \boldsymbol{T}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell](\rho) = t_{1,m}, \qquad \max_{0 \le j \le p q_m} d(\mathcal{P}_{\tilde{\mathcal{I}}}^{(j)}[g_\ell](\rho), \Gamma) < \delta_G. \tag{8.42}$$

Let $c_*$ be the constants given in Lemma 6.13. Now, by Lemma 6.13, we have $B_{\tilde{\Gamma}}(\rho, c_*^p) \subset \mathcal{D}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]$. Note that, since $p \le C\mathfrak{a}^{\ell+1}$, one can choose $\varepsilon_0$ small so that $C_G \delta_{m,\ell} \le c_*^p$ for all $m, \ell$ and $\varepsilon < \varepsilon_0$. Thus, $[\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}]^q(\rho)$ exists for $0 \le q \le q_m$.

Next, we claim that for all $1 \leq q \leq q_m$

$$\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho) = \varphi_{t_{q,m}}^{g_\ell}(\rho). \tag{8.43}$$

To see this, set $t_{0,m} = 0$. Since the claim is true for $q = 1$, we may assume that for some $1 \leq q < q_m$, and every $1 \leq j \leq q$,

$$T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^{j-1}(\rho)\right) = t_{j,m} - t_{j-1,m}, \tag{8.44}$$

and hence that for $1 \leq j \leq q$

$$\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^j(\rho) = \varphi_{t_{j,m}-t_{j-1,m}}^{g_\ell}\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^{j-1}(\rho)\right) = \varphi_{t_{j,m}}^{g_\ell}(\rho). \tag{8.45}$$

Now, observe that there exists $C$ depending only on $G$ such that

$$|T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell](\rho) - T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell](\rho')| \leq C^{pF} d(\rho, \rho').$$

Hence, using (8.41), (8.42), and (8.45)

$$\left|t_{1,m} - T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho)\right)\right| \leq C^{pF} C_G \delta_{m,\ell}.$$

Thus, we have

$$\varphi_{t_{q+1,m}-t_{q,m}}^{g_\ell}\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho)\right) \in \tilde{\Gamma}, \qquad \varphi_{T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho)\right)}^{g_\ell}\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho)\right) \in \tilde{\Gamma},$$

and, choosing $\varepsilon_0$ small enough,

$$\left|t_{q+1,m} - t_{q,m} - T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho)\right)\right| < (C^{pF} + 3)C_G \delta_{m,\ell} < c_\Gamma F.$$

In particular, this implies that

$$t_{q+1,m} - t_{q,m} = T_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\left(\left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho)\right),$$

and hence (8.44) holds with $q$ replaced by $q+1$. This shows that (8.43) holds.

Let $\tilde{\mathcal{I}} := (\tilde{i}_0, \tilde{i}_1, \dots)$. Using again that $\varepsilon > 0$ can be chosen small enough, this implies that, with the chain

$$\mathcal{I} := (\tilde{i}_0, \tilde{i}_1, \dots, \tilde{i}_p, \tilde{i}_1, \dots, \tilde{i}_p, \dots),$$

we have

$$(\lceil qt_m \rceil - 1)F - C_G \delta_G \leq \boldsymbol{T}_{\mathcal{I}}^{(pq)}[g_\ell](\rho) \leq \lceil qt_m \rceil F + C_G \delta_G, \tag{8.46}$$

and hence

$$\mathcal{P}_{\mathcal{I}}^{(qp)}[g_\ell](\rho) = \varphi_{t_{q,m}}^{g_\ell}(\rho) = \left(\mathcal{P}_{\tilde{\mathcal{I}}}^{(p)}[g_\ell]\right)^q(\rho).$$

Moreover, by definition, using that the fact that $\mathcal{D}_{\mathcal{I}}^{(p)}[g_\ell]$ and $\mathcal{D}_{\mathcal{I}}^{(pq)}[g_\ell]$ are open, we have $\left(\mathcal{P}_{\mathcal{I}}^{(p)}[g_\ell]\right)^q(\rho') = \mathcal{P}_{\mathcal{I}}^{(qp)}[g_\ell](\rho')$ for $\rho'$ in a neighborhood of $\rho$ as claimed.

*Step D.* We now complete the proof of the lemma. It follows from (8.32) and (8.37) that

$$\rho \in \Gamma \cap \mathscr{R}_F(k_m, C_G \delta_{m,\ell}, g_\ell) \cap \mathscr{N}_{q_m,F}(k_m, \beta_m(\ell), g_\ell), \tag{8.47}$$

with $\delta_{m,\ell}$ as in (8.38). Here, we have used that $\delta_m(\ell) \leq C_G \delta_{m,\ell}$ since $t_{m-1} \leq F k_m \leq F \mathfrak{a}^{\ell+1}$. The goal is then to use Lemma 6.18 with $\delta := C_G \delta_{m,\ell}$, and $(\beta, q_0, k, m) := (\beta_m(\ell), q_m, k_m, p)$. Let $C_3$

be as in Lemma 6.18 and note that $C_0 > C_3$ by (8.27). Thus, to apply Lemma 6.18 we first need to check that

$$C_G \delta_{m,\ell} \leq \min\left\{\delta_G\,,\,(2\mathfrak{f}(\beta_m(\ell),q_m)C_0^{k_m q_m})^{-1}\right\}. \tag{8.48}$$

To see this, note that by the definitions (8.38) and (6.14)

$$C_G \delta_{m,\ell} 2\mathfrak{f}(\beta_m(\ell),q_m)C_0^{k_m q_m} = 2C_G m C_0^{F\mathfrak{a}^{\ell+1}+k_m q_m}\left(\tfrac{5}{2}\right)^{2\aleph} q_m^{4\aleph}\alpha_{m-1}(\ell)\beta_m(\ell)^{-2\aleph}.$$

In addition, $\alpha_{m-1}(\ell)\beta_m(\ell)^{-2\aleph} \leq \gamma_{\ell-m+1}(\varepsilon)$ since $\beta_{i_m,j_m} \geq \beta_{j_m,j_m}$. Therefore, the bound in (8.48) follows from (8.19) after noting that $k_m q_m \leq \mathfrak{a}^{\ell+1}$ and $(\ell-m+1)/\mathfrak{b} \leq \ell+1$, and choosing $\varepsilon$ small enough.

Combining (8.48), (8.47), and (8.39) we may apply Lemma 6.18 and obtain

$$\rho \in \mathscr{N}_F\left(n,(2\mathfrak{f}(\beta_m(\ell),q_m)C_0^n)^{-1},g_\ell\right).$$

Here, we have used that, by (8.46) and $(n-1)F \leq t_0 \leq nF$, we have

$$(n-1)F - C_G\delta_G \leq \boldsymbol{T}_{\mathcal{I}}^{(q_m p)}[g](\rho) \leq nF + C_G\delta_G.$$

Finally, using that $\beta_m(\ell) \geq \beta_{j_m,j_m} \geq \beta_{\ell,\ell}$ one can check that $C^{-k_0} \geq 2\gamma_\ell(5/2)^{2\aleph}q_m^{4\aleph}$. In particular, since $q_m \leq \mathfrak{a}^{\ell+1}$, $\mathfrak{b} \in [0,1]$, making $\varepsilon_0 < C_0^{-1}$ would imply $[2\mathfrak{f}(\beta_m,q_m)C_0^n]^{-1} \geq \alpha_{\ell+1}$, and the proof is complete. □

**Lemma 8.6.** *Let $\rho \in \Gamma$ satisfy (8.33) and (8.36). Then for all $0 \leq j \leq m-1$ and $0 \leq q \leq t_j/t_m$,*

$$d(\varphi_{qt_m}^{g_\ell}(\rho),\rho) \leq \sum_{i=j}^{m-1} C_0^{\sum_{l=i}^{m-1} t_l}\alpha_i(\ell). \tag{8.49}$$

*Proof.* First, note that for $j = m-1$, the statement follows from (8.36). Let $0 \leq j \leq m-2$ and assume (8.49) holds for $0 \leq q \leq t_{j+1}/t_m$. Next, let $q < t_j/t_m$ and write $qt_m = \sum_{i=j+1}^m q_i t_i$ with $q_j < t_{j-1}/t_j$. Then, by (8.33) with $i = j+1$, $t = q_{j+1}t_{j+1}$, and $\mathfrak{r} = 0$,

$$d(\varphi_{q_{j+1}t_{j+1}}^{g_\ell}(\rho),\rho) \leq C_0^{t_j}\alpha_j(\ell).$$

Next, note that $g_\ell = F_\ell(g,\boldsymbol{\sigma}) \in \tilde{K}$ since $g \in K$. In particular, $\|\varphi_F^{g_\ell}\|_{C^{\nu-2}} \leq C\|g_\ell\|_{C^\nu} \leq C_0$ for all $\ell$. It follows that

$$d\left(\varphi_{\sum_{i=j+1}^m q_i t_i}^{g_\ell}(\rho),\varphi_{\sum_{l=j+2}^m q_l t_l}^{g_\ell}(\rho)\right) \leq C_0^{t_j+\sum_{l=j+2}^m q_l t_l}\alpha_j(\ell) \leq C_0^{\sum_{l=j}^{m-1} t_l}\alpha_j(\ell).$$

By the induction hypothesis, $d\left(\varphi_{\sum_{i=j+2}^m q_i t_i}^{g_\ell}(\rho),\rho\right) \leq \sum_{i=j+1}^{m-1} C_0^{\sum_{l=i}^{m-1} t_l}\alpha_i(\ell)$. Thus,

$$d(\varphi_{\sum_{i=j+1}^{m+1} q_i t_i}^{g_\ell}(\rho),\rho) \leq C_0^{\sum_{l=j}^{m-1} t_l}\alpha_j(\ell) + \sum_{i=j+1}^{m-1} C_0^{\sum_{l=i}^{m-1} t_l}\alpha_i(\ell) = \sum_{i=j}^{m-1} C_0^{\sum_{l=i}^{m-1} t_l}\alpha_i(\ell)$$

and the claim follows by induction. □

8.2.3. *Step 2: Preserving non-degeneracy under perturbation.* In this section, we show that if the non-degeneracy properties (8.25) and (8.26) listed in Proposition 8.3 hold for $g_\ell := F_\ell(g, \boldsymbol{\sigma})$, then they also hold for the perturbed metric $(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta} := Q^{R,\delta}(g_\ell, \boldsymbol{\sigma})$ for appropriate $R, \delta$.

**Lemma 8.7.** *There is $\varepsilon_0 > 0$ depending only on $\nu, N, \mathfrak{y}, b, \mathfrak{a}, F, \Gamma, G, K$ so that the following holds. Suppose that the conclusions in Proposition 8.3 are valid up to the index $\ell$ and let $0 < \varepsilon < \varepsilon_0$. Then, for all $(g, \boldsymbol{\sigma}) \in K \times \boldsymbol{\Sigma}_\ell(\mathfrak{R}_\varepsilon)$ such that $g_\ell := F_\ell(g, \boldsymbol{\sigma}) \in \mathscr{G}^\nu \backslash L_{T,\ell}$ the following holds. Let $(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta} := Q^{R,\delta}(g_\ell, \boldsymbol{\sigma})$, with $R = R_{\ell+1}(\varepsilon)$ and $\delta = \delta_{\ell+1}(\varepsilon)$ as defined in (8.29). Then, with $C_0$ as in (8.27),*

$$\|(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta} - g_\ell\|_{\mathcal{C}^\nu} \leq C_0 2^{-\ell-2} \varepsilon, \tag{8.50}$$

*and*

- *for $0 \leq i \leq \ell$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$,*

$$\Gamma \cap \mathscr{R}_F(n, \tilde{\beta}_{i,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathscr{N}_F(n, \tilde{\beta}_{i,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}). \tag{8.51}$$

- *for $\mathfrak{a}^\ell < n \leq \mathfrak{a}^{\ell+1}$*

$$\Gamma \cap \mathscr{R}_F(n, \tfrac{1}{2}\alpha_{\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \backslash \mathscr{S}_F(n, \tfrac{1}{2}\alpha_{\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathscr{N}_F(n, \tfrac{1}{2}\alpha_{\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}). \tag{8.52}$$

- *for $\mathfrak{b}(\ell+1) \leq j \leq \ell$, $0 \leq i \leq j$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$, and $1 \leq q \leq \mathfrak{a}^{j/\mathfrak{b}-i}$,*

$$\Gamma \cap \mathscr{R}_F(n, \beta_{i,j,\ell+1}q^{-3}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \cap \mathscr{S}_F(n, \alpha_{j,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathscr{N}_{q,F}(n, \beta_{i,j,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}). \tag{8.53}$$

*Proof.* By Definition 7.2 and (8.27), the bound in (8.50) holds since

$$\|(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta} - g_\ell\|_{\mathcal{C}^\nu} = \|Q^{R,\delta}(g_\ell, \boldsymbol{\sigma}) - Q^{R,\delta}(g_\ell, 0)\|_{\mathcal{C}^\nu} \leq \tfrac{1}{2}C_0 \delta R^{-\vartheta_\nu} \tag{8.54}$$

and, by definition,

$$\delta \leq \varepsilon R^{\vartheta_\nu} 2^{-(\ell+1)}. \tag{8.55}$$

We next address (8.51). Let $0 \leq i \leq \ell$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$. We claim that

$$\mathscr{R}_F(n, \tilde{\beta}_{i,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathscr{R}_F(n, \tilde{\beta}_{i,\ell}, g_\ell),$$
$$\Gamma \cap \mathscr{R}_F(n, \tilde{\beta}_{i,\ell}, g_\ell) \cap \mathscr{N}_F(n, \tilde{\beta}_{i,\ell}, g_\ell) \subset \mathscr{N}_F(n, \tilde{\beta}_{i,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}), \tag{8.56}$$

provided

$$\delta \leq C_0^{-\mathfrak{a}^i F - 1} R^{\vartheta_3} \tilde{\beta}_{i,\ell}(1 - s_\ell), \qquad 0 \leq i \leq \ell. \tag{8.57}$$

Once we prove (8.56), the claim in (8.51) will follow by applying the inductive assumption (8.25) to $g_\ell$. One can check that (8.57) is valid after asking that $\varepsilon_0 < C_0^{-1}$.

To see the first part of (8.56) suppose that $\rho \in \mathscr{R}_F(n, \tilde{\beta}_{i,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta})$. Then, $\rho \in \mathscr{R}_F(n, \tilde{\beta}_{i,\ell}, g_\ell)$ since Definition 7.2, (8.57), and (8.27) imply

$$\left\|\varphi_t^{g_\ell} - \varphi_t^{(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}}\right\|_{\mathcal{C}^1} \leq C_0^t \|g_\ell - (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}\|_{\mathcal{C}^3} \leq C_0^{t+1} \delta R^{-\vartheta_3} < \tilde{\beta}_{i,\ell}(1 - s_\ell) = \tilde{\beta}_{i,\ell} - \tilde{\beta}_{i,\ell+1}, \tag{8.58}$$

provided $(n-1)F \leq t \leq nF$ and (8.57) holds.

For the second part of (8.56), note that with $\delta_G$ as in Lemma 6.12 and (6.11), if $\tilde{\beta}_{i,\ell} < \delta_G$, we may work in a single coordinate chart $(U, \psi)$. Suppose $\rho \in \Gamma \cap \mathscr{R}_F(n, \tilde{\beta}_{i,\ell}, g_\ell) \cap \mathscr{N}_F(n, \tilde{\beta}_{i,\ell}, g_\ell)$.

By Definition 6.14 and Lemma 6.12 there exist $\mathcal{I}$ and $m$ such that $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_\ell](\rho) \in \big[(n-1)F - C_G\tilde{\beta}_{i,\ell}\,,\, nF + C_G\tilde{\beta}_{i,\ell}\big]$ and

$$\sup_{\{(\psi,U)\in\mathcal{A}:\, B(\rho,C_G\delta_G)\subset U\}} \big\| \big(I - d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_\ell])_\psi)(\rho)\big)^{-1}\big\| < \tilde{\beta}_{i,\ell}^{-1}. \tag{8.59}$$

In addition, since otherwise we may modify the choice of chain, we may assume without loss of generality that

$$d(\mathcal{P}_{\mathcal{I}}^{(j)}[g_\ell](\rho),\Gamma) < \delta_G, \qquad 0 \le j \le m. \tag{8.60}$$

Then, by Lemma 6.13 there is $c > 0$ depending only on $G,\Gamma,F$ such that

$$\rho \in \mathcal{D}_{\mathcal{I}}^{(m)}[(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}],$$

as long as $\|(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta} - g_\ell\|_{\mathcal{C}^\nu} \le c^m$. Here, we use that (8.54) holds with $\nu = 3$ together with (8.29) and $\vartheta_5 + 1 - \vartheta_3 > 0$ to justify that $\varepsilon_0$ can be chosen small enough that $\|(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta} - g_\ell\|_{\mathcal{C}^\nu} \le c^m$. Next, since $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_\ell](\rho) \in \big[(n-1)F - C_G\tilde{\beta}_{i,\ell}\,,\, nF + C_G\tilde{\beta}_{i,\ell}\big]$, we have

$$\boldsymbol{T}_{\mathcal{I}}^{(m)}[(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}](\rho) \in [(n-1)F - C_G\tilde{\beta}_{i,\ell} - C_0^n \delta R^{-\vartheta_3}\,,\, nF + C_G\tilde{\beta}_{i,\ell} + C_0^{nF+2}\delta R^{-\vartheta_3}]. \tag{8.61}$$

Indeed, by (8.58), implicit function theorem, and that $H_{|\xi|_g}$ is uniformly transverse to $\Gamma$,

$$|\boldsymbol{T}_{\mathcal{I}}^{(m)}[(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}](\rho) - \boldsymbol{T}_{\mathcal{I}}^{(m)}[g_\ell](\rho)| \le C_G C_0^{nF+1}\delta R^{-\vartheta_3}$$

and (8.61) follows. Now, since for $\varepsilon$ small enough,

$$C_G C_0^{nF+1}\delta R^{-\vartheta_3} + C_G\tilde{\beta}_{i,\ell} < C_G\tilde{\beta}_{i,\ell+1} < C_G\delta_G,$$

we have

$$\boldsymbol{T}_{\mathcal{I}}^{(m)}[(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}](\rho) \in [(n-1)F - C_G\delta_G\,,\, nF + C_G\delta_G]. \tag{8.62}$$

Let $(\psi,U) \in \mathcal{A}$ be such that $B(\rho,C_G\delta_G) \subset U$. Abusing notation slightly, we work with $\mathcal{P}_{\mathcal{I}}^{(m)}[g_\ell]$, and $\mathcal{P}_{\mathcal{I}}^{(m)}[(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}]$ for the maps induced in the $(\psi,U)$ coordinates. Let $\boldsymbol{A} = I - d\mathcal{P}_{\mathcal{I}}^{(m)}[g_\ell](\rho)$ and $\boldsymbol{\Delta} = d\mathcal{P}_{\mathcal{I}}^{(m)}[g_\ell](\rho) - d\mathcal{P}_{\mathcal{I}}^{(m)}[(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}](\rho)$, and observe that (8.58) yields $\|\boldsymbol{\Delta}\| \le \tilde{\beta}_{i,\ell}(1 - s_\ell)$. In particular, by (8.59),

$$\|(\boldsymbol{A} + \boldsymbol{\Delta})^{-1}\| \le \frac{\|\boldsymbol{A}^{-1}\|}{1 - \|\boldsymbol{A}^{-1}\|\|\boldsymbol{\Delta}\|} \le \frac{\tilde{\beta}_{i,\ell}^{-1}}{1 - \tilde{\beta}_{i,\ell}^{-1}\tilde{\beta}_{i,\ell}(1 - s_\ell)} = \tilde{\beta}_{i,\ell}^{-1}s_\ell^{-1} = \tilde{\beta}_{i,\ell+1}^{-1}.$$

Combining this with (8.62) we conclude $\rho \in \mathcal{N}_F(n,\tilde{\beta}_{i,\ell+1},(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta})$, and hence the second part of (8.56). As noted earlier, this shows that (8.51) holds.

We next prove (8.52). Let $\mathfrak{a}^\ell < n \le \mathfrak{a}^{\ell+1}$. We will show that

$$\mathcal{R}_F(n,\tfrac{1}{2}\alpha_{\ell+1},(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathcal{R}_F(n,\alpha_{\ell+1},g_\ell),$$
$$\mathcal{R}_F(n,\alpha_{\ell+1},g_\ell) \cap \mathcal{N}_F(n,\alpha_{\ell+1},g_\ell) \subset \mathcal{N}_F(n,\tfrac{1}{2}\alpha_{\ell+1},(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}), \tag{8.63}$$

$$\mathcal{S}_F(n,\alpha_{\ell+1},g_\ell) \subset \mathcal{S}_F(n,\tfrac{1}{2}\alpha_{\ell+1},(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}), \tag{8.64}$$

provided

$$\delta \le C_0^{-\mathfrak{a}^{\ell+1}F-1}R^{\vartheta_3}\tfrac{1}{2}\alpha_{\ell+1}. \tag{8.65}$$

Note that combining (8.63) and (8.64) with Lemma 8.5 yields the claim in (8.52). Also, it is immediate to check that (8.65) is valid after asking $\varepsilon_0 \leq C_0^{-1}$.

The claims in (8.63) follow from the same proof that yielded (8.56). To see that (8.64) holds, observe that if $\rho \notin \mathscr{S}_F(n, \frac{1}{2}\alpha_{\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}))$, then there is $t \in (\frac{1}{2}F, (n - \frac{1}{2})F)$ such that $d\big(\varphi_t^{(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}}(\rho), \rho\big) < \frac{1}{2}\alpha_{\ell+1}$. Thus, by (8.58) and (8.65), we know $\|\varphi_t^{g_\ell} - \varphi_t^{(g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}}\|_{\mathcal{C}^1} \leq \frac{1}{2}\alpha_{\ell+1}$. In particular, $d\big(\varphi_t^{g_\ell}(\rho), \rho\big) < \alpha_{\ell+1}$, and hence $\rho \notin \mathscr{S}_F(n, \alpha_{\ell+1}, \kappa_\ell)$ as needed.

Finally, we address (8.53). Let $0 \leq i \leq j$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$, $\mathfrak{b}(\ell+1) \leq j \leq \ell$, and $1 \leq q \leq \mathfrak{a}^{j/\mathfrak{b}-i}$. We claim that, provided (8.68) and (8.71) below hold,

$$\mathscr{R}_F(n, \beta_{i,j,\ell+1}q^{-3}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathscr{R}_F(n, \beta_{i,j,\ell}q^{-3}, g_\ell), \quad \mathscr{S}_F(n, \alpha_{j,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}) \subset \mathscr{S}_F(n, \alpha_{j,\ell}, g_\ell) \quad (8.66)$$

$$\Gamma \cap \mathscr{R}_F(n, \beta_{i,j,\ell}q^{-3}, g_\ell) \cap \mathscr{N}_{q,F}(n, \beta_{i,j,\ell}, g_\ell) \subset \mathscr{N}_{q,F}(n, \beta_{i,j,\ell+1}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}). \quad (8.67)$$

Note that combining (8.66) and (8.67) with (8.26) yields the claim in (8.53).

For (8.66) we argue as in the proofs of the first part of (8.56) and (8.64), and use that $\alpha_{j,\ell+1} \geq \frac{1}{2}\alpha_{j,\ell}$. We can do this provided

$$\delta < \min\left(C_0^{-\mathfrak{a}^\ell F - 1} R^{\vartheta_3}(\beta_{i,j,\ell} - \beta_{i,j,\ell+1})(\mathfrak{a}^{j/\mathfrak{b}-i})^{-3}, \ C_0^{-\mathfrak{a}^\ell F - 1} R^{\vartheta_3}\tfrac{1}{2}\alpha_{j,\ell}\right). \quad (8.68)$$

We note that (8.68) holds for $\varepsilon_0$ small enough since $\{\alpha_j\}_j$ is decreasing and $\beta_{\ell,\ell} \leq \beta_{i,j}$.

The inclusion in (8.67) will follow from Lemma 8.8 below with $\beta = \beta_{i,j,\ell}$. Indeed, provided

$$\|g_\ell - (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}\|_{\mathcal{C}^3} \leq \min(C_1^{-nq}\beta^{2\aleph}q^{-4\aleph-1}, \delta_G - \beta), \quad (8.69)$$

we have by Lemma 8.8

$$\begin{aligned}\Gamma \cap \mathscr{R}_F(n, \beta_{i,j,\ell}q^{-3}, g_\ell) &\cap \mathscr{N}_{q,F}(n, \beta_{i,j,\ell}, g_\ell) \\ &\subset \Gamma \cap \mathscr{R}_F(n, \beta_{i,j,\ell}, g_\ell) \cap \mathscr{N}_{q,F}(n, \beta_{i,j,\ell}, g_\ell) \\ &\subset \mathscr{N}_{q,F}(n, \beta_{i,j,\ell}(1 - \|g_\ell - (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}\|_{\mathcal{C}^3}\beta_{i,j,\ell}^{-2\aleph}C_1^{nq}q^{4\aleph+1})^{1/2\aleph}, (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}).\end{aligned} \quad (8.70)$$

Since,

$$\delta < R^{\vartheta_3}C_1^{-\mathfrak{a}^{j/\mathfrak{b}}-1}(\mathfrak{a}^{j/\mathfrak{b}-i})^{-4\aleph+1}\beta_{i,j,\ell}^{2\aleph}(1 - s_{\ell-j}^{2\aleph}), \quad (8.71)$$

the bound on $\|g_\ell - (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}\|_{\mathcal{C}^3}$ from (8.54) and (8.71) imply the first part of (8.69) and

$$\beta_{i,j,\ell+1} \leq \beta_{i,j,\ell}(1 - \|g_\ell - (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}\|_{\mathcal{C}^3}\beta_{i,j,\ell}^{-2\aleph}C^{nq}q^{4\aleph+1})^{1/2\aleph}. \quad (8.72)$$

Furthermore, for $\varepsilon < C_0^{-1}$, we have $\beta_{i,j,\ell} \leq \frac{\delta_G}{2}$, and $\|g_\ell - (g_\ell)_{\boldsymbol{\sigma}}^{R,\delta}\|_{\mathcal{C}^3} < \frac{\delta_G}{2}$. Therefore, the second part of (8.69) holds. Combining (8.70) with (8.72) then implies (8.67) as claimed. $\qquad \square$

**Lemma 8.8.** *Let $K \subset \mathscr{G}^3$ bounded and $C_1 > 0$ as above. Then for $g_1, g_2 \in K$, $\beta \in (0, \delta_G)$, $q \in \mathbb{N}$, and $\varepsilon > 0$ satisfying*

$$\|g_1 - g_2\|_{\mathcal{C}^3} \leq \varepsilon, \qquad \varepsilon \leq \min(C_1^{-nq}\beta^{2\aleph}q^{-4\aleph-1}, \delta_G - \beta)$$

*we have*

$$\Gamma \cap \mathscr{R}(n, \beta, g_1) \cap \mathscr{N}_{q,F}(n, \beta, g_1) \subset \mathscr{N}_{q,F}(n, \beta(1 - \varepsilon\beta^{-2\aleph}C^{nq}q^{4\aleph+1})^{1/2\aleph}, g_2).$$

*Proof.* Let $\rho \in \Gamma \cap \mathscr{R}(n, \beta, g_1) \cap \mathscr{N}_{q,F}(n, \beta, g_1)$ and $(\psi, U) \in \mathcal{A}$ such that $B(\rho, C_G \delta_G) \subset U$. Then, there are $\mathcal{I}$ and $m$ such that $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g_1](\rho) \in \left[(n-1)F - C_G\beta, \, nF + C_G\beta\right]$,

$$\sup_{m \leq n} d(\mathcal{P}_{\mathcal{I}}^{(m)}[g_1](\rho), \Gamma) < \delta_G,$$

and

$$\frac{\left\|(I - [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)(\rho)]^q)^{-1}\right\|}{(1 + \|d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)(\rho)\|^q)^{2\aleph - 1}} \leq (\tfrac{5}{2}q^2\beta^{-1})^{2\aleph}. \tag{8.73}$$

Note that $(1 + \|d((\mathcal{P}_{\mathcal{I}}^{(m)}[g])_\psi)(\rho')\|^q)^{2\aleph - 1} \leq C_1^{nq}$ for all $g \in K$ and $\rho' \in \mathcal{D}_{\mathcal{I}}^{(m)}[g]$.

Now, by Lemma 6.13, $\rho \in \mathcal{D}_{\mathcal{I}}^{(m)}[g_2]$. Therefore, defining

$$\mathbf{A}_q = I - [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)(\rho)]^q, \qquad \boldsymbol{\Delta}_q = [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)(\rho)]^q - [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_2])_\psi)(\rho)]^q,$$

(8.73) yields $\|\mathbf{A}_q^{-1}\| \leq \frac{5^{2\aleph}}{2^{2\aleph}} C_1^{nq} \beta^{-2\aleph} q^{4\aleph}$. Next, note that

$$\|\boldsymbol{\Delta}_q\| = \left\| \sum_{j=1}^{q} [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)(\rho)]^{q-j} [d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)(\rho) - (d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_1])_\psi)](\rho)[(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_2])_\psi)(\rho)]^j \right\|$$

$$\leq q C_1^{nq} \|\mathcal{P}_{\mathcal{I}}^{(m)}[g_1] - \mathcal{P}_{\mathcal{I}}^{(m)}[g_2]\|_{\mathcal{C}^1} \leq q C_1^{n(q+1)} \varepsilon,$$

where we take the $\mathcal{C}^1$ norm on the intersection of the two domains and the last inequality follows from (6.11). In particular,

$$\frac{\left\|(I - [(d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_2])_\psi)(\rho)]^q)^{-1}\right\|}{(1 + \|d((\mathcal{P}_{\mathcal{I}}^{(m)}[g_2])_\psi)(\rho)\|^q)^{2\aleph - 1}} \leq \|(\boldsymbol{A}_q + \boldsymbol{\Delta}_q)^{-1}\| \leq \frac{\|\boldsymbol{A}_q^{-1}\|}{1 - \|\boldsymbol{A}_q^{-1}\| q \varepsilon C_1^{n(q+1)}}$$

which yields the desired bound since $q \varepsilon C_1^{n(q+1)} \leq \frac{1}{2} \frac{2^{2\aleph}}{5^{2\aleph}} C_1^{-nq} \beta^{2\aleph} q^{-4\aleph} \leq \frac{1}{2} \|\boldsymbol{A}_q^{-1}\|^{-1}$.

Finally, we also note that

$$(n-1)F - C_G \delta_G \leq (n-1)F - C_G(\beta + \varepsilon) \leq \boldsymbol{T}_{\mathcal{I}}^{(m)}[g_2](\rho) \leq nF + C_G(\beta + \varepsilon) \leq nF + C_G\delta_G,$$

and hence, $\rho \in \mathscr{N}_{q,F}(n, \beta(1 - \varepsilon\beta^{-2\aleph}C_1^{nq}q^{4\aleph+1})^{1/2\aleph}, g_2)$ as claimed. $\qquad \square$

8.2.4. *Step 3: Controlling the volume of bad perturbations.* This step is dedicated to finishing the induction argument that yields the proof of Proposition 8.3. We still assume the conclusions of Proposition 8.3 hold up to some $\ell \geq \lceil \log_{\mathfrak{a}} 2 \rceil - 1$, and will show they also hold up to $\ell + 1$.

In what follows, we work with $R := R_{\ell+1}$, $\delta := \delta_{\ell+1}$, and for $0 \leq i \leq \ell + 1$ set

$$\beta_{i,\ell+1} = \tfrac{1}{2} \min_{\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i} \upsilon_n(\ell), \qquad \upsilon_n(\ell) := 2\beta_{\ell+1,\ell+1} \varepsilon^{(n-\mathfrak{a}^{\ell+1})\dim \mathfrak{sp}(2\aleph)}. \tag{8.74}$$

The first step is to show that for $(g, \boldsymbol{\sigma}) \in K \times \boldsymbol{\Sigma}_\ell(\mathfrak{R}_\varepsilon)$ such that $F_\ell(g, \boldsymbol{\sigma}) \in \mathscr{G}^\nu \backslash L_{T,\ell}$,

$$\left\{ \boldsymbol{\sigma}_{\ell+1} \in \Sigma(R_{\ell+1}) \,:\, F_{\ell+1}(g, (\boldsymbol{\sigma}, \boldsymbol{\sigma}_{\ell+1})) \in L_{T,\ell+1} \right\} \subset \bigcup_{n=0}^{\mathfrak{a}^{\ell+1}} \mathcal{S}_{F_\ell(g, \hat{\boldsymbol{\sigma}}_\ell)}^{R,\delta}(n, \tfrac{1}{2}\alpha_{\ell+1}, \upsilon_n(\ell)) \tag{8.75}$$

for $\mathcal{S}_{F_\ell(g, \hat{\boldsymbol{\sigma}}_\ell)}^{R,\delta}$ as defined in Corollary 7.7.

To prove (8.75), first note that if $\boldsymbol{\sigma}_{\ell+1}$ is such that

$$g_{\ell+1} := F_{\ell+1}(g,(\boldsymbol{\sigma},\boldsymbol{\sigma}_{\ell+1})) = Q^{R,\delta}(F_\ell(g,\boldsymbol{\sigma}),\boldsymbol{\sigma}_{\ell+1}) \in L_{T,\ell+1},$$

by the definition of $L_{T,\ell+1}$, then either (8.25) or (8.26) does not hold with $(g,\ell) = (g_{\ell+1},\ell+1)$. We claim that in this setting, one actually has that (8.26) does not hold with $(g,\ell,j) = (g_{\ell+1},\ell+1,\ell+1)$. That is, there exist $0 \leq i \leq \ell+1$, $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$, and $1 \leq q \leq \mathfrak{a}^{(\ell+1)/\mathfrak{b}-i+1}$, such that

$$\Gamma \cap \mathscr{R}_F(n,\beta_{i,\ell+1,\ell+1}q^{-3},g_{\ell+1}) \cap \mathscr{S}_F(n,\alpha_{\ell+1,\ell+1},g_{\ell+1}) \nsubseteq \mathscr{N}_{q,F}(n,\beta_{i,\ell+1,\ell+1},g_{\ell+1}). \tag{8.76}$$

Indeed, suppose (8.25) does not hold with $(g,\ell) = (g_{\ell+1},\ell+1)$. Then, since $F_\ell(g,\boldsymbol{\sigma}) \in \mathscr{G}^\nu \backslash L_{T,\ell}$, by Lemma 8.7 equation (8.51), (8.25) can only fail when $i = \ell+1$. Thus, there are $\mathfrak{a}^\ell < n \leq \mathfrak{a}^{\ell+1}$ and

$$\rho \in \Gamma \cap \mathscr{R}_F(n,\tilde{\beta}_{\ell+1,\ell+1},g_{\ell+1}) \backslash \mathscr{N}_F(n,\tilde{\beta}_{\ell+1,\ell+1},g_{\ell+1}).$$

Since $\tilde{\beta}_{\ell+1,\ell+1} \leq \alpha_{\ell+1,\ell+1} = \frac{1}{2}\alpha_{\ell+1}$, Lemma 8.7 equation (8.52) yields that $\rho \in \mathscr{S}_F(n,\alpha_{\ell+1,\ell+1},g_{\ell+1})$. Furthermore, since $\tilde{\beta}_{\ell+1,\ell+1} \leq 2\beta_{\ell+1,\ell+1} = \beta_{\ell+1,\ell+1,\ell+1}$, we also have $\rho \in \Gamma \cap \mathscr{R}_F(n,\beta_{\ell+1,\ell+1,\ell+1},g_{\ell+1})$. Using that $\mathfrak{f}(\beta_{\ell+1,\ell+1,\ell+1},1) \leq 1/\tilde{\beta}_{\ell+1,\ell+1}$, this implies (8.76) for $q = 1$ and $i = \ell+1$ since $\mathscr{N}_{1,F}(n,\beta_{\ell+1,\ell+1,\ell+1},g_{\ell+1}) \subset \mathscr{N}_F(n,\tilde{\beta}_{\ell+1,\ell+1},g_{\ell+1})$.

For any $\boldsymbol{\sigma}_{\ell+1}$ such that $g_{\ell+1} = F_{\ell+1}(g,(\boldsymbol{\sigma},\boldsymbol{\sigma}_{\ell+1})) \in L_{F,\ell+1}$, we have showed that (8.26) does not hold with $(g,\ell) = (g_{\ell+1},\ell+1)$. Since Lemma 8.7 equation (8.53) yields that (8.26) holds for all $j \leq \ell$, we conclude the claim in (8.76). In particular, since $\beta_{i,\ell+1,\ell+1} \leq \upsilon_n(\ell)$ for all $i \leq \ell+1$ and $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$, we conclude there exists $n \leq \mathfrak{a}^{\ell+1}$ such that $\boldsymbol{\sigma}_{\ell+1} \in \mathcal{S}^{R,\delta}_{F_\ell(g,\hat{\boldsymbol{\sigma}}_\ell),\delta}(n,\alpha_{\ell+1,\ell+1},\upsilon_\ell(n))$. This yields the claim in (8.75).

Next, we apply Corollary 7.7 with $R = R_{\ell+1}(\varepsilon)$, $\delta = \delta_{\ell+1}(\varepsilon)$, $\alpha = \frac{1}{2}\alpha_{\ell+1}$, and $\beta = \upsilon_n(\ell)$. Note that since $n \leq \mathfrak{a}^{\ell+1}$ one may choose $\varepsilon_0$ so that the conditions (7.19) on $R,\delta$ are satisfied. Furthermore, since $\beta_{\ell+1,\ell+1} = \varepsilon^{\boldsymbol{d}\sum_{j=1}^{\ell+1}\mathfrak{c}^{\ell+1-j}\mathfrak{a}^{j/\mathfrak{b}}}\beta_{0,0}^{\mathfrak{c}^{\ell+1}}$, one can check that there is $\boldsymbol{d}$ large enough (depending only on $\mathfrak{a},\mathfrak{b},\aleph$) such that

$$C_0 R^{\mathfrak{L}-2\aleph-1}\delta^{-\mathfrak{L}}\sum_{n=0}^\infty C_0^{n(\mathfrak{L}-2\aleph-1)}\upsilon_n(\ell) < \varepsilon 2^{-\ell-2}, \tag{8.77}$$

and $\mathfrak{c} \geq 2\aleph(m_\nu\mathfrak{L} + \mathfrak{y}'(1-\dim\mathfrak{sp}(2\aleph)))$. To see this, we use estimate the left hand side of (8.77) using the definitions of $\beta_{\ell,\ell}$ (8.21), $\upsilon_n$ (8.74), $R_\ell$, and $\delta_\ell$ (8.29), $\alpha_\ell$ (8.22) and $\gamma_\ell$ (8.19) to obtain that there is $C > 0$ such that

$$C_0 R^{\mathfrak{L}-2\aleph-1}\delta^{-\mathfrak{L}}\sum_{n=0}^\infty C_0^{n(\mathfrak{L}-2\aleph-1)}\upsilon_n(\ell) \leq C\beta_{\ell,\ell}^{\mathfrak{c}+2\aleph(\dim\mathrm{Sp}(2\aleph)-1)\mathfrak{y}'-m_\nu\mathfrak{L}}\varepsilon^{(\boldsymbol{d}-C)\mathfrak{a}^{\ell/\mathfrak{b}}}2^{C\ell}\mathfrak{a}^{C\ell/\mathfrak{b}}.$$

Choosing $\boldsymbol{d}$ large enough the estimate (8.77) follows provided $\mathfrak{c} \geq 2\aleph(m_\nu\mathfrak{L}+\mathfrak{y}'(1-\dim\mathfrak{sp}(2\aleph)))$. One can also check that $\beta = \upsilon_n(\ell) \leq c^n R_{\ell+1}^{\mathfrak{y}'}$ provided $\mathfrak{c} \geq 2\aleph\mathfrak{y}'$ and hence, we may apply Corollary 7.7. Optimizing in $\mathfrak{y}' \geq \mathfrak{y}$ we obtain $\mathfrak{c} = \max(\frac{2\aleph m_\nu\mathfrak{L}}{\mathfrak{L}-2\aleph},2\aleph\mathfrak{y})$.

Corollary 7.7 together with (8.77) and (8.75) then yield that (8.24) holds up to index $\ell+1$ in place of $\ell$.

We next show that (8.23) is a consequence of (8.24). Given $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\ell+1}(\mathfrak{R})$, if

$$\exists j \leq \ell + 1 \quad \text{s.t.} \quad F_j(g, \hat{\boldsymbol{\sigma}}_j) \in L_{F,j},$$

then either there exists $1 \leq j \leq \ell + 1$ such that $F_j(g, \hat{\boldsymbol{\sigma}}_j) \in L_{F,j}$ and $F_{j-1}(g, \hat{\boldsymbol{\sigma}}_{j-1}) \in \mathscr{G}^\nu \backslash L_{F,j-1}$ or $F_0(g, \hat{\boldsymbol{\sigma}}_0) \in L_{F,0}$. Therefore, by Fubini's theorem and the induction hypothesis (8.24)

$$m_{\boldsymbol{\Sigma}_{\ell+1}(\mathfrak{R})}\Big(\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\ell+1}(\mathfrak{R}) : \exists j \leq \ell + 1 \text{ s.t. } F_j(g, \hat{\boldsymbol{\sigma}}_j) \in L_{F,j}\Big) \leq \sum_{j=0}^{\ell+1} \varepsilon 2^{-j-1} \leq (1 - 2^{-(\ell+2)})\varepsilon.$$

Now, notice that

$$m_{\boldsymbol{\Sigma}_{\ell}(\mathfrak{R})}\Big(\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\ell}(\mathfrak{R}) : \exists j \leq \ell + 1 \text{ s.t. } F_j(g, \hat{\boldsymbol{\sigma}}_j) \in L_{F,j}\Big) =$$
$$m_{\boldsymbol{\Sigma}_{\infty}(\mathfrak{R})}\Big(\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\infty}(\mathfrak{R}) : \exists j \leq \ell \text{ s.t. } F_j(g, \hat{\boldsymbol{\sigma}}_j) \in L_{F,j}\Big)$$

and hence (8.23) holds. This finishes the proof of Proposition 8.3.

### 8.3. Predominance of quantitative non-degeneracy.

We now turn to the proof that (8.1) holds for a predominant set of $g$. The next proposition is the key technical result of the article. We will find sequences $\boldsymbol{\delta}_\varepsilon$ and $\mathfrak{R}_\varepsilon$ such that (8.1) holds for an $\mathscr{F} := \{(F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}, \infty)\}_\varepsilon$ predominant set of metrics (with $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ as in Section 8.1 built using any family perturbations $Q^{R,\delta}$ that are good in the sense of Definition 7.2). More precisely, provided we are able find a collection $\{(\Gamma, G)\}_{G \in \mathcal{G}}$ of $(F, b, \mathfrak{y})$ admissible pairs for $\{Q^{R,\delta}\}_{R,\delta}$ such that $\bigcup_{G \in \mathcal{G}} G = \mathscr{G}^\nu$, for each $\varepsilon > 0$ and $K \subset \mathscr{G}^\nu$ bounded we will find a $C > 0$ so that we can control the measure of the set of bad parameter values $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R})$ such that $F_\infty(g, \boldsymbol{\sigma})$ does not satisfy (8.1) with that $C$ and any $g \in K$. This result will be used in Section 11 to prove Theorem 1.3.

**Proposition 8.9.** *Let $\nu \geq 5$, $\mathfrak{y} \geq 1$, $b > 0$, $N : (0,1) \to \mathbb{N}$, and $\{Q^{R,\delta}\}_{R,\delta}$ be a $(\nu, N)$-good family of perturbations. For $0 < \varepsilon < 1$, there are $\boldsymbol{\delta}_\varepsilon = \{\delta_j(\varepsilon)\}_{j=0}^\infty$, $\mathfrak{R}_\varepsilon := \{R_j(\varepsilon)\}_{j=0}^\infty$, such that*

$$\sum_{j=0}^\infty \delta_j(\varepsilon) R_j(\varepsilon)^{-\vartheta_\nu} \leq \varepsilon,$$

*and, for all $(\Gamma, G)$ an $(F, b, \mathfrak{y})$-admissible pair for $\{Q^{R,\delta}\}_{R,\delta}$, $K \subset G$ bounded, there is $\varepsilon_0 > 0$ and $C > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$ and $g \in K$ there exists Borel set $S_{g,\varepsilon} \subset \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon)$ such that*

$$\sup_{g \in K} m_{\boldsymbol{\Sigma}_\infty(\mathfrak{R})}\big(S_{g,\varepsilon}\big) \leq \varepsilon,$$

*and for all $g \in K$*

$$\big\{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon) : F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) \in L_\infty(\varepsilon)\big\} \subset S_{g,\varepsilon}$$

*where, $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}$ is defined in Lemma 8.1,*

$$L_\infty(\varepsilon) = \Big\{g \in \mathscr{G}^\nu : \exists n \text{ such that } \Gamma \cap \mathscr{R}_F(n, \beta_n(\varepsilon), g) \not\subseteq \mathscr{N}_F(n, \beta_n(\varepsilon), g)\Big\}, \qquad (8.78)$$

*and*

$$\beta_n(\varepsilon) := \varepsilon^{Cn^\gamma} C^{-n^\gamma} n^{-C \log \varepsilon^{-1} n^\gamma}, \qquad (8.79)$$

*with $\gamma := 1 + \log_2\big[\max\big(\frac{2\aleph(2\aleph+3)}{2\aleph+1} m_\nu, 2\aleph\mathfrak{y}\big)\big]$ and $m_\nu = \max(b, \vartheta_\nu, \vartheta_5 + 1, 2\tilde{\vartheta}_4, 1)$. In addition, $L_\infty(\varepsilon)$ is Borel.*

*Proof.* From Proposition 8.3, for all $1 < \mathfrak{a} < 2$ there are $\{\delta_j(\varepsilon)\}_{j=0}^\infty$ and $\{R_j(\varepsilon)\}_{j=0}^\infty$ such that (8.5), (8.23) and (8.20) hold. Let $(\Gamma, G)$ satisfying Definition 7.3 and $K \subset G$ bounded. Suppose that $g \in K$, and $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R})$. Then, by Lemma 8.1, $F_\infty(g, \boldsymbol{\sigma}) \in \mathscr{G}^\nu$ is well defined. We claim that the statement holds with

$$S_{g,\varepsilon} := \bigcup_{\ell=1}^\infty S_{g,\varepsilon}(\ell), \qquad S_{g,\varepsilon}(\ell) := \{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}) : \exists j \leq \ell \text{ s.t. } F_j(g, \widehat{\boldsymbol{\sigma}}_j) \in L_{F,j}\},$$

with $L_{F,j}$ as defined in Proposition 8.3. Note that by Proposition 8.3, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ and $g \in K$, $m_{\boldsymbol{\Sigma}_\infty(\mathfrak{R})}\big(S_{g,\varepsilon}(\ell)\big) \leq (1 - 2^{-\ell-1})\varepsilon$, and so, since $S_{g,\varepsilon}(\ell+1) \subset S_{g,\varepsilon}(\ell)$, for all $0 < \varepsilon < \varepsilon_0$ and $g \in K$, $m_{\boldsymbol{\Sigma}_\infty(\mathfrak{R})}\big(S_{g,\varepsilon}\big) \leq \varepsilon$, as claimed.

Next, we claim that $\big\{\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon) : F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) \in L_\infty(\varepsilon)\big\} \subset S_{g,\varepsilon}$. For $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_\infty(\mathfrak{R}) \setminus S_{g,\varepsilon}$ we will show that $F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}(g, \boldsymbol{\sigma}) \notin L_\infty(\varepsilon)$. Writing $g_\ell = F_\ell(g, \hat{\boldsymbol{\sigma}}_\ell)$, we have that $g_\ell \in \mathscr{G}^\nu \backslash L_{F,\ell}$ for all $\ell$. In particular, for all $\ell$, $0 \leq i \leq \ell$, and $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$,

$$\Gamma \cap \mathscr{R}_F(n, \tilde{\beta}_{i,\ell}, g_\ell) \subset \mathscr{N}_F(n, \tilde{\beta}_{i,\ell}, g_\ell).$$

By definition, $\tilde{\beta}_{i,\ell} = \frac{1+2^{-\ell}}{1+2^{-i}} \tilde{\beta}_{i,i}$. Therefore, for $\ell$, $0 \leq i \leq \ell$, and $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$, we have

$$\Gamma \cap \mathscr{R}_F(n, \tfrac{1}{2}\tilde{\beta}_{i,i}, g_\ell) \subset \mathscr{N}_F(n, \tfrac{1}{2}\tilde{\beta}_{i,i}, g_\ell).$$

Now, by (8.20), we have $\|g_\ell - g_{\ell-1}\|_{\mathcal{C}^\nu} \leq C 2^{-\ell-1}\varepsilon$. Therefore, $g_\ell \to F_\infty(g, \boldsymbol{\sigma}) =: g_\infty \in \mathcal{C}^\nu$ with $\|g_\infty - g\|_{\mathcal{C}^\nu} \leq \varepsilon$.

Fix $\mathfrak{a}^{i-1} < n \leq \mathfrak{a}^i$. We claim that

$$\Gamma \cap \mathscr{R}_F(n, \tfrac{1}{4}\tilde{\beta}_{i,i}, g_\infty) \subset \mathscr{N}_F(n, \tfrac{1}{4}\tilde{\beta}_{i,i}, g_\infty). \tag{8.80}$$

To see this, suppose that $\rho \in \Gamma \cap \mathscr{R}_F(n, \tfrac{1}{4}\tilde{\beta}_{i,i}, g_\infty)$. Then, for $\ell$ large enough,

$$\rho \in \Gamma \cap \mathscr{R}_F(n, \tfrac{1}{2}\tilde{\beta}_{i,i}, g_\ell) \subset \mathscr{N}_F(n, \tfrac{1}{2}\tilde{\beta}_{i,i}, g_\ell) \subset \mathscr{N}_F(n, \beta_n(\varepsilon), g_\infty).$$

In particular, $\rho \in \mathscr{N}_F(n, \tfrac{1}{4}\tilde{\beta}_{i,i}, g_\infty)$.

Finally, we choose $1 < \mathfrak{a} < 2$ (and hence find $\{\mathfrak{R}_\varepsilon\}$ and $\{\boldsymbol{\delta}_\varepsilon\}$) such that (8.80) implies that for all $n \geq 1$,

$$\Gamma \cap \mathscr{R}_F(n, \beta_n(\varepsilon), g_\infty) \subset \mathscr{N}_F(n, \beta_n(\varepsilon), g_\infty), \tag{8.81}$$

and hence that $g_\infty \notin L_\infty$. To check (8.81), we need only show that for $\mathfrak{a}^{\ell-1} < n \leq \mathfrak{a}^\ell$, we have $\tfrac{1}{4}\tilde{\beta}_{\ell,\ell} \geq \beta_n(\varepsilon)$.

By definition, $\beta_{\ell,\ell} = \beta_{0,0}^{\mathfrak{c}^\ell} \varepsilon^{\boldsymbol{d}\mathfrak{c}^{\ell-1} \sum_{j=0}^{\ell-1} (\mathfrak{c}^{-1}\mathfrak{a}^{\frac{1}{\mathfrak{b}}})^j}$. In addition, $\lim_{\mathfrak{a} \to 2^-} \mathfrak{a}^{-\frac{1}{\mathfrak{b}}} = 2^{-\frac{\log_2 \mathfrak{a}}{1 - \log_2 \mathfrak{a}}} = 0$. Therefore, we may choose $\mathfrak{a}$ so that

$$\mathfrak{c}\mathfrak{a}^{-1/\mathfrak{b}} = 1, \tag{8.82}$$

where $\mathfrak{c} = \max\big(\frac{2\aleph(2\aleph+3)}{2\aleph+1} m_\nu, 2\aleph\mathfrak{y}\big) > 1$ and $m_\nu = \max(b, \vartheta_\nu, \vartheta_5 + 1, 2\tilde{\vartheta}_4, 1)$. Therefore,

$$\tilde{\beta}_{\ell,\ell} = \varepsilon^{2\aleph-1} \beta_{\ell,\ell}^{2\aleph} = \beta_{0,0}^{2\aleph\mathfrak{c}^\ell} \varepsilon^{2\aleph\boldsymbol{d}\mathfrak{c}^{\ell-1}\ell} \varepsilon^{2\aleph-1}$$

$$\geq \beta_{0,0}^{2\aleph\mathfrak{c}^\ell} \varepsilon^{2\aleph\boldsymbol{d}\mathfrak{c}^{\ell-1}\ell+2\aleph-1}.$$

Moreover, (8.82) implies $\log_2 \mathfrak{a} = \frac{\log_2 \mathfrak{c}}{1 + \log_2 \mathfrak{c}}$ and, in particular, $\frac{\log_2 \mathfrak{c}}{\log_2 \mathfrak{a}} = 1 + \log_2 \mathfrak{c} = \gamma$.

Notice that $\log_{\mathfrak{a}} n < \ell \leq 1 + \log_{\mathfrak{a}} n$ since $\mathfrak{a}^{\ell-1} < n \leq \mathfrak{a}^{\ell}$ and hence $\mathfrak{c}^{\ell} \leq cn^{\gamma}$ which, using that $\beta_{0,0}(\varepsilon) \geq c\varepsilon^{N}$ for some $c > 0$ and $N \geq 0$, yields, for $C > 0$ large enough (in the definition of $\beta_n(\varepsilon)$) depending on $\mathfrak{c}$, $\aleph$, and $\boldsymbol{d}$, yields

$$\tilde{\beta}_{\ell,\ell} \geq [\beta_{0,0}(\varepsilon)]^{2\aleph\mathfrak{c}n^{\gamma}} n^{-2\aleph\boldsymbol{d}\gamma\frac{\log_2 \varepsilon^{-1}}{\log_2 \mathfrak{c}} n^{\gamma}} \varepsilon^{2\aleph\boldsymbol{d}n^{\gamma}+2\aleph-1} \geq 4\beta_n(\varepsilon).$$

$\square$

## 9. Nice perturbations in the space of metrics

In this section, we define families of perturbations that satisfy Definition 7.2 and 7.3. The basic perturbation takes place on a small ball through which a geodesic passes exactly once and is inspired by those in [Ano82, Kli78]. We then build a large family of these perturbations in order to be able to handle any possible geodesic.

9.1. **The elementary family of perturbations.** Let $\nu \geq 0$. We fix a reference metric, $g_f \in \mathscr{G}^{\nu+2}$. Let $\rho_0 \in \widetilde{S^*M}$ and $e_1, \ldots, e_d \in T_{\pi_M(\rho_0)}M$ an orthonormal set of vectors for $g_f$.

For $g_\star \in \mathscr{G}^\nu$, consider the geodesic

$$\gamma_{\rho_0}^{g_\star}(t) := \pi_M(\varphi_t^{g_\star}(\rho_0)).$$

We choose a basis $E_1^{g_\star}, \ldots, E_d^{g_\star} \in T_{\gamma_{\rho_0}^{g_\star}(0)}M$, orthonormal with respect to $g_\star$ by applying Gram-Schmidt process to $(\dot{\gamma}_{\rho_0}(0), e_1, \ldots, e_d)$ so that, in particular, $E_d^{g_\star} = \dot{\gamma}_{\rho_0}(0)$. (Here, we only use $e_d$ if $\dot{\gamma}_{\rho_0}(0), e_1, \ldots, e_{d-1}$ are linearly dependent.) Let $E_i^{g_\star}(t) \in T_{\gamma_{\rho_0}^{g_\star}(t)}M$ be the parallel transport (again with respect to $g_\star$) of $E_i^{g_\star}$ along $\gamma_{\rho_0}^{g_\star}$. For $\varepsilon_f > 0$, define the map

$$\Phi_{\rho_0}^{g_\star} : B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times \mathbb{R} \to M, \qquad \Phi_{\rho_0}^{g_\star}(u, t) = \exp_{\gamma_{\rho_0}^{g_\star}(t)}^{g_f} \left( \sum_{i=1}^{d-1} u^i E_i^{g_\star}(t) \right), \qquad (9.1)$$

where $\exp_{\gamma_{\rho_0}(t)}^{g_f}$ denotes the exponential map for the metric $g_f$ with base point $\gamma_{\rho_0}(t)$. Then for $\varepsilon_f > 0$ small enough depending only on $g_f$, and any small enough (depending on $g_\star$ and $g_f$) interval $I \subset \mathbb{R}$, $\Phi_{\rho_0}^{g_\star}|_{I \times B(0, \varepsilon_f)}$ is a diffeomorphsim onto its image.

The reason for using $g_f$ above rather than $g_\star$ itself is that, in our perturbation argument, $g_\star$ will vary, remaining bounded in $\mathscr{G}^\nu$, but not necessarily in $\mathscr{G}^{\nu'}$ for $\nu' > \nu$. It will therefore be important for us to control the $\mathcal{C}^\nu$ norms of $\Phi_{\rho_0}^{g_\star}$ using only the $\mathscr{G}^\nu$ norms of $g_\star$ (See Lemma 9.4). This is *not* possible if one replaces $g_f$ by $g_\star$.

**Remark 9.1.** Although $\Phi_{\rho_0}^{g_\star}$ is not a global diffeomorphism, it is locally a diffeomorphism, and we will often ignore this fact in order to simplify the presentation. When defining our perturbations below, we will work in a $(u, t)$ neighborhood where the map is a genuine diffeomorphism.

**Lemma 9.2.** *In the $(u, t)$ coordinates $g_\star$ satisfies*

$$g_\star^{ij}(0, t) = \delta^{ij}, \qquad g_\star^{dd}(0, t) = 1, \qquad g_\star^{jd}(0, t) = 0, \qquad \partial_{u_j} g_\star^{dd}(0, t) = 0.$$

*where, here and below, we write $i, j, k, \ell$ for indices in $\{1, \ldots, d-1\}$.*

*Proof.* The first three equalities follow from the fact that $\{E_i^{g_\star}(t)\}_{i=1}^d$ is an orthonormal frame with respect to $g_\star$. Since $(0,t)$ is a geodesics for $g_\star$, letting $x(s) = (0,s) \in \mathbb{R}_u^{d-1} \times \mathbb{R}_t$, we have

$$\partial_s^2 x^\beta + \frac{1}{2} g_\star^{\mu\beta}(\partial_\alpha g_{\star,\mu\nu} + \partial_\nu g_{\star,\mu\alpha} - \partial_\mu g_{\star,\alpha\nu})\partial_s x^\alpha \partial_s x^\nu = 0.$$

Since $x^i = \partial_s x^i = \partial_s^2 x^i = 0$ for $i = 1, \ldots, d-1$, $\partial_s x^d = 1$, and $\partial_s^2 x^d = 0$, we have that for $i = 1, \ldots, d-1$

$$\frac{1}{2} g_\star^{\mu i}(\partial_d g_{\star,\mu d} + \partial_d g_{\star,\mu d} - \partial_\mu g_{\star,dd}) = -\frac{1}{2}\partial_{u_i} g_{\star,dd} = 0.$$

Since $g_{\star,ij}(0,t)$ is the identity, $\partial_{u_i} g_\star^{dd} = 0$. □

**Lemma 9.3.** *The map $F : \mathscr{G}^\nu \to \mathcal{C}^{\nu-1}(\mathbb{R}^d; M)$ given by $F(g) := \Phi_{\rho_0}^g$ is Frechet differentiable and, moreover, for any bounded subset $G \subset \mathscr{G}^\nu$ there is $C > 0$ such that for $g \in G$,*

$$\|DF|_g \delta_g\|_{\mathcal{C}^{\nu-1}} \leq C\|\delta_g\|_{\mathcal{C}^{\nu-1}}.$$

*Proof.* To prove the lemma, we work locally in a coordinate chart so that we may identify vectors along a curve with $\mathbb{R}^d$.

Define the map $\mathcal{E}_0 : \mathscr{G}^\nu \to (\mathbb{R}^d)^d$ by $\mathcal{E}_0(g) := (E_1^g(0), \ldots, E_d^g(0))$ with $E_i^g$ defined as above.

Let $\mathcal{E} : \mathscr{G}^\nu \times \mathcal{C}^{\nu+1}([0,1]; \mathbb{R}^d) \times (\mathbb{R}^d)^d \to C^\nu([0,1]; (\mathbb{R}^d)^d)$ be the map $\mathcal{E}(g, \gamma, (e_i)_{i=1}^d) := (e_i^g)_{i=1}^d$, where $e_i^g \in C^\nu([0,1]; \mathbb{R}^d)$ is a parallel frame along $\gamma$ (with respect to $g$) with initial conditions $(e_i^g(0))_{i=1}^d = (e_i)_{i=1}^d$.

In addition, define $\exp^{g_f} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ by $\exp^{g_f}(x,v) := \exp_x^{g_f}(v)$ and let $\Psi : \widetilde{S^*M} \times \mathscr{G}^\nu \to \mathcal{C}^{\nu+1}([0,1]; \mathbb{R}^d)$ be the map $\Psi(\rho, g) := \gamma_\rho^g$ where $\gamma_\rho^g$ is the unit speed geodesic for the metric $g$ with initial conditions given by $\rho$. Finally, let $Y : \mathcal{C}^{\nu+1}([0,1]; \mathbb{R}^d) \times C^\nu([0,1]; (\mathbb{R}^d)^d) \to C^\nu([0,1] \times \mathbb{R}^{d-1}; \mathbb{R}^d)$,

$$\left[Y(\gamma, (E_i)_{i=1}^d)\right](t,u) = \exp_{\gamma(t)}^{g_f}\left(\sum_{i=1}^{d-1} u^i E_i(t)\right).$$

We are interested in studying their composition

$$F(g) = Y\left(\Psi(\rho_0, g), \mathcal{E}\left(g, \Psi(\rho_0, g), \mathcal{E}_0(g)\right)\right) = \Phi_{\rho_0}^g.$$

To clarify the exposition, in what follows we write $\mathscr{S}^\nu$ for the space of symmetric tensors endowed with the $\mathcal{C}^\nu$ topology.

Note that $g \mapsto D_g \mathcal{E}_0$ is bounded and continuous in the $\mathscr{G}^\nu$ topology. By writing the ordinary differential equation satisfied by a parallel vector field and using that $\dot\gamma \in \mathcal{C}^{\nu-1}([0,1]; \mathbb{R}^d)$, $\partial_x g \in \mathscr{S}^{\nu-1}$, one can check that $D_g \mathcal{E} : \mathscr{S}^{\nu-1} \to \mathcal{C}^{\nu-1}([0,1]; (\mathbb{R}^d)^d)$ and the map $(g, \gamma) \mapsto D_g \mathcal{E}|_{(g,\gamma)}$ is continuous in the $\mathscr{G}^\nu \times \mathcal{C}^\nu$ topology with codomain the set of bounded linear maps from $\mathscr{S}^{\nu-1}$ to $\mathcal{C}^{\nu-1}([0,1] \times (\mathbb{R}^d)^d)$.

Next, again using the parallel transport equation and $\dot\gamma \in \mathcal{C}^{\nu-1}$, $g \in \mathscr{G}^\nu$, the map $D_\gamma \mathcal{E} : \mathscr{S}^{\nu-1} \to \mathcal{C}^{\nu-1}([0,1]; (\mathbb{R}^d)^d)$ is bounded and the map $(g, \gamma) \mapsto D_\gamma \mathcal{E}|_{(g,\gamma)}$ continuous in the $\mathscr{G}^\nu \times \mathcal{C}^\nu$ topology with codomain the set of bounded linear maps from $\mathscr{S}^{\nu-1}$ to $\mathcal{C}^{\nu-1}([0,1]; (\mathbb{R}^d)^d)$.

Furthermore, it is easy to see that

$$D_{(\gamma,E)}Y : \mathcal{C}^\nu([0,1];\mathbb{R}^d) \times \mathcal{C}^{\nu-1}([0,1];(\mathbb{R}^d)^d) \to \mathcal{C}^{\nu-1}([0,1] \times \mathbb{R}^{d-1};\mathbb{R}^d)$$

is bounded and $(\gamma, E) \mapsto D_{(\gamma,E)}Y$ is continuous in the $\mathcal{C}^\nu([0,1];\mathbb{R}^d) \times \mathcal{C}^{\nu-1}([0,1];(\mathbb{R}^d)^d)$ topology with codomain the set of bounded linear maps from $\mathcal{C}^\nu([0,1];\mathbb{R}^d) \times \mathcal{C}^{\nu-1}([0,1];(\mathbb{R}^d)^d)$ to $\mathcal{C}^{\nu-1}([0,1] \times \mathbb{R}^{d-1};\mathbb{R}^d)$.

Finally, using the geodesic equation together with $g \in \mathscr{G}^\nu$ and that geodesics for a $\mathscr{G}^\nu$ metric lie in $\mathcal{C}^{\nu+1}$, we have $D_g\Psi : \mathscr{S}^{\nu-1} \to \mathcal{C}^\nu([0,1];\mathbb{R}^d)$ is bounded and the map $g_0 \mapsto D_g\Psi|_{g_0}$ is continuous as a function of $g_0$ in the $\mathscr{G}^\nu$ topology with with codomain the set of bounded linear maps from $\mathscr{S}^{\nu-1}$ to $\mathcal{C}^\nu([0,1];\mathbb{R}^d)$.

With these estimates in place it is now easy to see that $D_gF : \mathscr{S}^{\nu-1} \to \mathcal{C}^{\nu-1}([0,1] \times \mathbb{R}^{d-1};\mathbb{R}^d)$ is bounded and $g_0 \mapsto D_gF|_{g_0}$ is continuous in the $\mathscr{G}^\nu$ topology with with codomain the set of bounded linear maps from $\mathscr{S}^{\nu-1}$ to $\mathcal{C}^{\nu-1}([0,1] \times \mathbb{R}^{d-1};\mathbb{R}^d)$. $\qquad\square$

**Lemma 9.4.** *Let $\nu \geq 2$ and $K \subset \mathscr{G}^\nu$ be bounded. Then there is $C > 0$ such that for all $g_\star \in K$ and $\rho_0 \in \widetilde{S^*M}$,*

$$\|\Phi_{\rho_0}^{g_\star}\|_{\mathcal{C}^\nu} \leq C.$$

*Proof.* First, recall that in any coordinate system $x$ on $M$ with dual coordinates $\xi$, if $\rho_0 = (0, (0,\ldots,0,1))$, then $\gamma_{\rho_0}(s)$ is given by the solution $x(s)$ to

$$\dot{x}^\alpha = \frac{g_\star^{\alpha\beta}\xi_\beta}{|\xi|_g}, \qquad \dot{\xi}_\alpha = -\frac{\partial_{x_\alpha} g_\star^{\mu\beta}\xi_\mu\xi_\beta}{|\xi|_g}, \qquad (x(0),\xi(0)) = (0,(0,\ldots,0,1)).$$

Therefore, standard regularity theory for ODEs shows that $\gamma_{\rho_0}(t) \in C^{\nu+1}(\mathbb{R};M)$ and there exists $C_K > 0$ with $\|\gamma_{\rho_0}\|_{\mathcal{C}^{\nu+1}} \leq C_K$. In particular, $\|\dot{\gamma}_{\rho_0}\|_{\mathcal{C}^\nu} \leq C_K$. Next, recall that for each $i$ the equations for parallel transport take the form $E_i(t) = e^\alpha(t)\partial_{x^\alpha}$ where $e^\alpha(t)$ solves an equation of the form $\dot{e}^\alpha(t) = \Gamma e(t)\dot{\gamma}_{\rho_0}^{g_\star}(t)$ with $\Gamma$ representing the Christoffel symbol for $g$. In particular, $\|e^\alpha\|_{\mathcal{C}^\nu} \leq C_K$ and hence $\|E_i\|_{\mathcal{C}^\nu} \leq C_K$.

Now, since $g_f \in \mathscr{G}^{\nu+2}$ is independent of $K$, the exponential map $(p,V) \mapsto \exp_p^{g_f}(V)$ is in $C^\nu(TM;M)$ with bounds independent of $K$. Therefore, $\Phi_{\rho_0}^{g_\star}$ satisfies the desired estimates. $\qquad\square$

For $s \in \mathbb{R}$ and $\varepsilon > 0$ small let

$$\mathcal{Z}_s := \Phi_{\rho_0}^{g_\star}(B(0,\varepsilon_f) \times \{s\}), \qquad \mathcal{H}_s := \widetilde{S_{\mathcal{Z}_s}^*M} \cap B_{\widetilde{S^*M}}(\varphi_s^{g_\star}(\rho_0),\varepsilon). \tag{9.2}$$

Here, we use the metric $g_f$ to define the ball in $\widetilde{S^*M}$.

Note that $\mathcal{H}_s$ is a smooth hypersurface such that $\rho_0 \in \mathcal{H}_0$, and $\mathcal{H}_s$ is transverse to $H_{|\xi|_{g_\star}}$; in particular, satisfying,

$$|H_{|\xi|_{g_\star}}(\rho)f| \geq \frac{1}{2}|df|, \qquad \text{for } f \text{ defining } \mathcal{H}_s \text{ and all } \rho \in \mathcal{H}_s.$$

Despite the fact that we define $\mathcal{H}_s$ as a subset of $\widetilde{S^*M}$, it will be convenient in the rest of this section to identify $\widetilde{S^*M}$ with $S_g^*M$ for each $g$ i.e. the unit sphere bundle with respect to the metric

$g$. To do this, we define the canonical isomorphism from $\widetilde{S^*M} \to S_g^*M$ to by $(x, [\xi]) \mapsto (x, \frac{\xi}{|\xi|_g})$.
Here, we have used that $[\xi] = \left[\frac{\xi}{|\xi|_g}\right]$ in $\widetilde{S^*M}$.

We now define a family of perturbations $g_\sigma \in \mathscr{G}^4$ of a metric $g_0 \in \mathscr{G}^4$ close to $g_\star$ as follows: Let

$$\Sigma_1 := \mathbb{R}^{d-1} \times \mathbb{R}^{d-1}, \qquad \Sigma_2 := \mathrm{Sym}(d-1) \times \mathbb{M}(d-1) \times \mathrm{Sym}(d-1), \qquad \Sigma := \Sigma_1 \times \Sigma_2$$

and denote the elements of $\Sigma$ by

$$\sigma := (\sigma_1, \sigma_2) = \big((\mathbf{A}, \mathbf{B}), (\mathbf{C}, \mathbf{D}, \mathbf{E})\big) \in \Sigma. \tag{9.3}$$

The closeness of $g_0$ to $g_\star$ will be crucial in Lemma 9.12.

Then, for $g_0 \in \mathscr{G}^g$, $\rho_0 \in \widetilde{S^*M}$, $\sigma \in \Sigma$, $R > 0$, and $t_\star \in \mathbb{R}$, the family of perturbations

$$g_\sigma = g_\sigma(\rho_0, t_\star, R, g_0, \sigma)$$

is defined in the $(u, t)$ coordinates introduced in (9.1) as follows

$$\begin{aligned}
g_\sigma^{dd}(u, t) &= g_0^{dd}(u, t) - \big(\langle \mathbf{A}, u \rangle + \tfrac{1}{2}\langle \mathbf{C}u, u \rangle\big)\chi_R(u, t - t_\star), \\
g_\sigma^{jd}(u, t) &= g_0^{jd}(u, t) + \tfrac{1}{2}\big(\mathbf{B}^j + (\mathbf{D}u)^j\big)\chi_R(u, t - t_\star), \\
g_\sigma^{ij}(u, t) &= g_0^{ij}(u, t) + \tfrac{1}{2}\mathbf{E}^{ij}\chi_R(u, t - t_\star),
\end{aligned} \tag{9.4}$$

$\chi_R(u, t) = \frac{1}{R}\chi(\frac{1}{R}t)\chi(\frac{1}{R}|u|)$ and $\chi \in C_c^\infty((-\sqrt{2}, \sqrt{2}); [0, 1])$ even with $\chi \equiv 1$ on $[-\frac{1}{4}, \frac{1}{4}]$ and $\int \chi = 1$. In order to identify $g_\sigma$ with a metric on $M$, we let $K \subset \mathscr{G}^3$ be bounded and work with $g_\star \in K$. In addition, we require that $0 < R < R_0$ where $R_0 = R_0(K)$ is chosen such that for all $0 < R < R_0$, $g_\star \in K$, and $t \in \mathbb{R}$, we have

$$\Phi_{\rho_0}^{g_\star}\Big(B_{\mathbb{R}^{d-1}}(0, \sqrt{2}R) \times [t - \sqrt{2}R, t + \sqrt{2}R]\Big) \subset B_{g_\star}(\gamma_{\rho_0}^{g_\star}(t), 3R) \tag{9.5}$$

with $\Phi_{\rho_0}^{g_\star}$ as in (9.1). This is possible since, by Lemma 9.2, $g_\star^{ij}(0, t) = \delta^{ij}$, for all $t \in \mathbb{R}$ and $i, j = 1, \ldots, d$. In addition, will fix $T_0 > 0$ below and require that

$$\Big((\Phi_{\rho_0}^{g_\star})^{-1}(B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R)) \cap B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times [0, T_0]\Big) \subset B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times [t_\star - 3R, T_\star + 3R].$$

That is, the geodesic $\gamma_{\rho_0}^{g_\star}$ passes near $\gamma_{\rho_0}^{g_\star}(t_\star)$ only once between time 0 and $T_0$.

**Remark 9.5.** Observe that $\chi_R$ is chosen so that the integral along $(0, t)$ is 1.

Define the map $\boldsymbol{\zeta} : B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times \mathscr{G}^4 \times \Sigma \to \mathcal{H}_0$

$$\boldsymbol{\zeta}(z_0, g_0, \sigma) := (u_0, 0, \omega_0, \tau_0(z_0, g_0, \sigma)), \qquad z_0 := (u_0, \omega_0) \tag{9.6}$$

where $\tau_0 > 0$ is defined so that $|(\omega_0, \tau_0)|_{g_\sigma(u_0, 0)} = 1$ and $\boldsymbol{\zeta}(z_0, g_0, \sigma) \in \mathcal{H}_0$. In particular,

$$\rho_0 = \boldsymbol{\zeta}(0, g_0, 0) = (0, 0, 0, 1) \in \mathcal{H}_0. \tag{9.7}$$

Let $\boldsymbol{u}, \boldsymbol{t}, \boldsymbol{\omega}, \boldsymbol{\tau}$ be functions defined on $\mathbb{R} \times B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times \mathscr{G}^4 \times \Sigma$ such that

$$\varphi_s^{g_\sigma}(\boldsymbol{\zeta}(z_0, g_0, \sigma)) = \big(\boldsymbol{u}, \boldsymbol{t}, \boldsymbol{\omega}, \boldsymbol{\tau}\big)(s, z_0, g_0, \sigma). \tag{9.8}$$

Note that by the definition of $\boldsymbol{\zeta}$, for all $s$ and $\sigma$

$$|(\boldsymbol{\omega}, \boldsymbol{\tau})|_{g_\sigma(\boldsymbol{u}, \boldsymbol{t})} = 1.$$

Also, $\boldsymbol{t}(s, 0, g_\star, \sigma) = s$. Next, define the matrices

$$\Delta_\sigma(s, z_0, g_0) := \begin{pmatrix} \partial_{u_0}\boldsymbol{u}(s, z_0, g_0, \sigma) & \partial_{\omega_0}\boldsymbol{u}(s, z_0, g_0, \sigma) \\ \partial_{u_0}\boldsymbol{\omega}(s, z_0, g_0, \sigma) & \partial_{\omega_0}\boldsymbol{\omega}(s, z_0, g_0, \sigma) \end{pmatrix},$$

$$\tilde{\Delta}_\sigma := \Delta_\sigma - [\partial_s\boldsymbol{t}]^{-1}\begin{pmatrix} \partial_s\boldsymbol{u} \\ \partial_s\boldsymbol{\omega} \end{pmatrix}\begin{pmatrix} \partial_{u_0}\boldsymbol{t} & \partial_{\omega_0}\boldsymbol{t} \end{pmatrix}. \tag{9.9}$$

We note that $\tilde{\Delta}_\sigma \in \mathrm{Sp}(2(d-1))$. Indeed, for each fixed $s$, observe that $\tilde{\Delta}_\sigma$ is the derivative of the following Poincaré map (a symplectomorphism)

$$\mathcal{P}_\sigma : \mathcal{H}_0 \to \mathcal{H}_{\boldsymbol{t}(s, z_0, g_0, \sigma)}, \qquad \mathcal{P}_\sigma(z_0) = (\boldsymbol{u}(S, z_0, g_0, \sigma), \boldsymbol{\omega}(S, z_0, g_0, \sigma)),$$

where $S = S(s, z_0, g_0, \sigma)$ is defined such that $\varphi_S^{g_\sigma}(\boldsymbol{\zeta}(z_0, \sigma)) \in \mathcal{H}_{\boldsymbol{t}(s, z_0, g_0, \sigma)}$. Here, we are identifying $\mathcal{H}_0$ with $B(0, \varepsilon_f) \times B(0, \varepsilon_f)$ via $z_0 = (u_0, \omega_0) \mapsto \boldsymbol{\zeta}(z_0, \sigma)$. Note also that the symplectic form induced on $\mathcal{H}_s$ from $T^*M$ is the standard symplectic form in $(u, \omega)$ coordinates.

Let $T_0 \in \mathbb{R}$, $g_0 \in \mathscr{G}^4$, and $z_0 \in B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times B_{\mathbb{R}^{d-1}}(0, \varepsilon_f)$. Define the map $\Psi_{z_0}^{g_0} : \Sigma \to \mathbb{R}^{d-1} \times \mathbb{R}^{d-1} \times \mathrm{Sp}(2(d-1))$ by

$$\Psi_{z_0}^{g_0, T_0}(\sigma) := (\Xi_1(z_0, g_0, \sigma), \Xi_2(z_0, g_0, \sigma)) \tag{9.10}$$

$$\Xi_1(z_0, g_0, \sigma) := (\boldsymbol{u}(T, z_0, g_0, \sigma), \boldsymbol{\omega}(T, z_0, g_0, \sigma)), \qquad \Xi_2(z_0, g_0, \sigma) := \tilde{\Delta}_\sigma\tilde{\Delta}_0^{-1}(T, z_0, g_0),$$

where $T = T(z_0, g_0, \sigma, T_0)$ is defined by

$$\varphi_{T(z_0, g_0, \sigma, T_0)}^{g_\sigma}(\boldsymbol{\zeta}(z_0, g_0, \sigma)) \in \mathcal{H}_{T_0}.$$

Note that $T(0, g_\star, 0, T_0) = T_0$, i.e., $\varphi_{T_0}^{g_\star}(\rho_0) \in \mathcal{H}_{T_0}$.

**Remark 9.6.** The choice of $\tilde{\Delta}_\sigma\tilde{\Delta}_0^{-1}$ as the last entry of $\Psi_{z_0}^{g_0}(\sigma)$ is motivated by the fact that we intend to write an ODE (in the $s$ variable) which $\partial_\delta\tilde{\Delta}_{\delta\sigma}|_{\delta=0}\tilde{\Delta}_0^{-1}(s)$ solves. Because $\tilde{\Delta}_{\delta\sigma}(s) \in \mathrm{Sp}(2d-1)$, $\partial_\delta\tilde{\Delta}_{\delta\sigma}(s)|_{\delta=0} \in T_{\tilde{\Delta}_0(s)}\mathrm{Sp}(2d-1)$, while $\partial_\delta\tilde{\Delta}_{\delta\sigma}(s)|_{\delta=0}\tilde{\Delta}_0^{-1} \in T_I\mathrm{Sp}(2d-1) = \mathfrak{sp}(2d-1)$. Because this is a linear subspace of $\mathbb{M}(2(d-1))$ which is independent of $s$, ODEs posed in this space are much simpler to work with.

For the purposes of the next calculations, we take $g_0 = g_\star$. In Lemma 9.10 we show how to handle $g_0 \neq g_\star$. Note that, with $g_0 = g_\star$, we have $(\boldsymbol{u}, \boldsymbol{t}, \boldsymbol{\omega}, \boldsymbol{\tau})|_{\sigma=0} = (0, s, 0, 1)$ when $z_0 = (0, 0)$. Let

$$\boldsymbol{u}_\sigma(s) := \boldsymbol{u}(s, 0, g_\star, \sigma), \qquad \boldsymbol{\omega}_\sigma(s) = \boldsymbol{\omega}(s, 0, g_\star, \sigma), \qquad \boldsymbol{t}_\sigma(s) = \boldsymbol{t}(s, 0, g_\star, \sigma). \tag{9.11}$$

We continue to write $\sigma = (\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E})$ and define

$$L_\sigma := \begin{pmatrix} \partial_{u^k}g_\sigma^{dj} & g_\sigma^{jk} \\ -\frac{1}{2}\partial_{u^ju^k}^2 g_\sigma^{dd} & -\partial_{u^j}g_\sigma^{dk} \end{pmatrix} \in \mathfrak{sp}(2(d-1)).$$

Next, we calculate how derivatives of $\Delta_{\delta\sigma}$ and $\Xi_1(z_0, g_0, \delta\sigma)$ with respect to $\delta$ behave.

**Lemma 9.7.** Let $K \subset \mathscr{G}^3$ bounded. Then there is $R_0 > 0$ such that for all $g_\star \in K$ the following holds. Let $\rho_0 \in \widetilde{S^*M}$, $T_0 \in \mathbb{R}$, $0 < R < R_0$ and $t_\star \in [3R, T_0 - 3R]$ be such that the set

$$\{t \in [0, T_0] : \gamma_{\rho_0}^{g_\star}(t) \cap B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R)\} \tag{9.12}$$
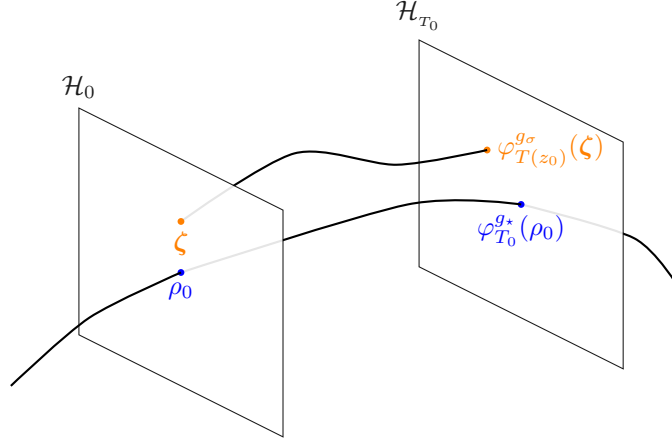
is connected.

FIGURE 4. The figure shows the setup for the definition of the functions $\boldsymbol{t}$ and $T$. Here, we abbreviate $\boldsymbol{\zeta} = \boldsymbol{\zeta}(z_0, g_0, \sigma)$ and $T(z_0) = T(z_0, g_0, \sigma)$. The hypersurface $\mathcal{H}_0$ parametrized by $z_0 = (u_0, \omega_0)$ via $z_0 \mapsto \boldsymbol{\zeta}(z_0, \sigma)$, while the corresponding points in $\mathcal{H}_{T_0}$ are $(\boldsymbol{u}, \boldsymbol{\omega}) = \Xi_1(z_0, g_0, \sigma)$.

Let $g_0 = g_\star$, $\sigma = (\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}) \in \Sigma$ and $\delta > 0$. Then, with $\boldsymbol{u}_\sigma$ and $\boldsymbol{\omega}_\sigma$ as defined in (9.11), for $s \in [0, T_0]$

$$\partial_s \begin{pmatrix} \partial_\delta \boldsymbol{u}_{\delta\sigma}|_{\delta=0} \\ \partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{\delta=0} \end{pmatrix} = L_0 \begin{pmatrix} \partial_\delta \boldsymbol{u}_{\delta\sigma}|_{\delta=0} \\ \partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{\delta=0} \end{pmatrix} + \tfrac{1}{2} \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{A} \end{pmatrix} \chi_R(0, s - t_\star), \tag{9.13}$$

and

$$\partial_s \left( \partial_\delta \Delta_{\delta\sigma}|_{\delta=0} \Delta_0^{-1} \right) = \tfrac{1}{2} \begin{pmatrix} \boldsymbol{D}_k^j & \boldsymbol{E}^{jk} \\ \boldsymbol{C}_{jk} & -\boldsymbol{D}_j^k \end{pmatrix} \chi_R(0, s - t_\star) + [L_0, \partial_\delta \Delta_{\delta\sigma}|_{\delta=0} \Delta_0^{-1}] + F_R(\boldsymbol{A}, \boldsymbol{B}, s), \tag{9.14}$$

where $L_\sigma$ is given by (9.18), $F_R(\boldsymbol{A}, \boldsymbol{B}, s) \in \mathfrak{sp}(2(d-1))$ and

$$|F_R(\boldsymbol{A}, \boldsymbol{B}, s)| \leq C e^{C|s - t_\star|} \|(\boldsymbol{A}, \boldsymbol{B})\|$$

with $C = C(\|g_\star\|_{\mathcal{C}^3})$.

*Proof.* There is $R_0 = R_0(K) > 0$ such that (9.5) holds and we can work with coordinates $(u, t)$ on the ball $B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R_0)$. Moreover, by the connectedness of the set in (9.12)

$$\left( (\Phi_{\rho_0}^{g_\star})^{-1}(B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R_0)) \cap B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times [0, T_0] \right) \subset B_{\mathbb{R}^{d-1}}(0, \varepsilon_f) \times [t_\star - 3R_0, T_\star + 3R_0].$$

Since the support of the perturbation is inside $B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R_0)$, this allows us to identify the perturbation $g_\sigma$ with a genuine metric perturbation on $M$ and to treat $(u, t)$ as though they were global coordinates for the purposes of the calculations in this lemma. By (9.8), the Hamiltonian flow, $\varphi_s^{g_\sigma}$, for $|\xi|_{g_\sigma}$ is defined by

$$\begin{aligned} \partial_s \boldsymbol{u}^j &= g_\sigma^{ij} \boldsymbol{\omega}_i + g_\sigma^{dj} \boldsymbol{\tau}, & \partial_s \boldsymbol{\omega}_j &= -\tfrac{1}{2} \partial_{u^j} g_\sigma^{i\ell} \boldsymbol{\omega}_i \boldsymbol{\omega}_\ell + \partial_{u^j} g_\sigma^{d\ell} \boldsymbol{\tau} \boldsymbol{\omega}_\ell + \tfrac{1}{2} \partial_{u^j} g_\sigma^{dd} \boldsymbol{\tau}^2, \\ \partial_s \boldsymbol{t} &= g_\sigma^{dj} \boldsymbol{\omega}_j + g_\sigma^{dd} \boldsymbol{\tau}, & \partial_s \boldsymbol{\tau} &= -\tfrac{1}{2} \partial_t g_\sigma^{i\ell} \boldsymbol{\omega}_i \boldsymbol{\omega}_\ell + \partial_t g_\sigma^{d\ell} \boldsymbol{\tau} \boldsymbol{\omega}_\ell + \tfrac{1}{2} \partial_t g_\sigma^{dd} \boldsymbol{\tau}^2, \end{aligned} \tag{9.15}$$

where $\xi = (\boldsymbol{\omega}, \boldsymbol{\tau})$.

Combining the Hamilton equations (9.15) together with (9.4) and Lemma 9.2 yields (9.13). Here, we are using that $\partial_t g_\star^{jd}(0,t) = g_\star^{jd}(0,t) = \partial_t \partial_{u^j} g_\star^{dd} = \partial_{u^j} g_\star^{dd} = 0$, $\partial_\delta \boldsymbol{u}_{\delta\sigma}|_{s=0} \equiv 0$, and $\partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{s=0} \equiv 0$. The last two equalities follow from the facts that $\partial_\delta \boldsymbol{u}_\delta|_{s=0} = \partial_\delta(\boldsymbol{u}_\delta|_{s=0})$ and $\boldsymbol{u}_\delta|_{s=0} \equiv u_0$, and the analogue for $\boldsymbol{\omega}_\delta$.

The same computation yields that at $z_0 = (0,0)$,

$$\partial_s \Delta_0 = L_0 \Delta_0, \qquad \Delta_0|_{s=0} = I. \tag{9.16}$$

Differentiating in $(u_0, \omega_0)$, then $\delta$, using that $\Delta_\sigma|_{s=0} = I$, and that $\partial_\delta \boldsymbol{u}_{\delta\sigma}|_{\delta=0}$ and $\partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{\delta=0}$ depend only on $(\boldsymbol{A}, \boldsymbol{B})$ and satisfy

$$\|(\partial_\delta \boldsymbol{u}_{\delta\sigma}|_{\delta=0}(s), \partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{\delta=0}(s))\| \leq C e^{C|s-t_\star|}, \tag{9.17}$$

we find there is $F_R(\boldsymbol{A}, \boldsymbol{B}, s)$ such that at $z_0 = (0,0)$,

$$\partial_s \partial_\delta \Delta_{\delta\sigma}|_{\delta=0} = L_0 \partial_\delta \Delta_{\delta\sigma}|_{\delta=0} + \partial_\delta L_{\delta\sigma}|_{\delta=0} \Delta_0 + F_R(\boldsymbol{A}, \boldsymbol{B}, s)\Delta_0, \qquad \partial_\delta \Delta_{\delta\sigma}|_{s=0} = 0.$$

Note that

$$\partial_\delta L_{\delta\sigma}|_{\delta=0} = \tfrac{1}{2} \begin{pmatrix} \boldsymbol{D}_k^j & \boldsymbol{E}^{jk} \\ \boldsymbol{C}_{jk} & -\boldsymbol{D}_j^k \end{pmatrix} \chi_R(0, s - t_\star). \tag{9.18}$$

Therefore, at $z_0 = (0,0)$, by (9.16) we have $(\partial_\delta \Delta_{\delta\sigma}|_{\delta=0}\Delta_0^{-1})|_{s=0} = 0$ and conclude (9.14). We claim that (9.14) implies

$$F_R(\boldsymbol{A}, \boldsymbol{B}, s) \in \mathfrak{sp}(2(d-1)).$$

Indeed, this follows from the fact that 1) $\Delta_0 \in \mathrm{Sp}(2(d-1))$ and 2) $\partial_\delta \Delta_{\delta\sigma}|_{\delta=0}\Delta_0^{-1} \in T_I \mathrm{Sp}(2(d-1)) = \mathfrak{sp}(2(d-1))$. Claim 1) follows from the fact that $\partial_s(\boldsymbol{u}, \boldsymbol{\omega})|_{(u_0,\omega_0)=0} = 0$ and hence that $\Delta_0 = \tilde{\Delta}_0 \in \mathrm{Sp}(2(d-1))$. Claim 2) The claim is not obvious since we do not know that $\Delta_\sigma \in \mathrm{Sp}(2(d-1))$ for all $\sigma \neq 0$. However, using that $\partial_{u_0}\boldsymbol{t}|_{\delta=0} = \partial_{\omega_0}\boldsymbol{t}|_{\delta=0} = 0$ and $\partial_s\boldsymbol{u}|_{\delta=0} = \partial_s\boldsymbol{\omega}|_{\delta=0} = 0$, we have $\partial_\delta \Delta_{\delta\sigma}|_{\delta=0} = \partial_\delta \tilde{\Delta}_{\delta\sigma}|_{\delta=0}$, and hence, since $\tilde{\Delta}_\sigma \in \mathrm{Sp}(2(d-1))$, $\partial_\delta(\Delta_{\delta\sigma}^T J \Delta_{\delta\sigma})|_{\delta=0} = 0$. In particular, this implies that

$$(\partial_\delta \Delta_{\delta\sigma}|_{\delta=0}\Delta_0^{-1})^T J \partial_\delta \Delta_{\delta\sigma}|_{\delta=0}\Delta_0^{-1} = (\Delta_0^T)^{-1}(\partial_\delta(\Delta_{\delta\sigma}^T J \Delta_{\delta\sigma})|_{\delta=0})\Delta_0^{-1} = 0,$$

and hence that $\partial_\delta \Delta_{\delta\sigma}|_{\delta=0}\Delta_0^{-1} \in \mathfrak{sp}(2(d-1))$.

The estimate on $F_R(\boldsymbol{A}, \boldsymbol{B}, s)$, now follows from (9.17). □

In what follows we prove that $\partial_\sigma \Psi_0^{g_\star}|_{\sigma=0}$ is bijective and bound its inverse. We write $\sigma = (\sigma_1, \sigma_2)$ as introduced in (9.3). We note that $\Psi_{z_0}^{g_0, T_0}(\sigma)$ is defined using the Hamiltonian flow for $g_\sigma = g_\sigma(\rho_0, t_\star, R, g_0, \sigma)$, where $g_\sigma$ is a perturbation of the metric $g_0$ done in a neighborhood of the point $\gamma_{\rho_0}^{g_\star}(t_\star)$. Indeed, in the $(u,t)$ coordinates, the perturbation is supported where $\chi_R(u, t - t_\star)$ is, i.e. for $|u| \leq \sqrt{2}R$ and $|t - t_\star| \leq \sqrt{2}R$.

**Lemma 9.8.** *Let $K \subset \mathscr{G}^3$ bounded. Then there is $C > 0$ and $R_0 > 0$ suc that for all $g_\star \in K$ the following holds. Let $\rho_0 \in \widetilde{S^*M}$, $T_0 \in \mathbb{R}$, $0 < R < R_0$ and $t_\star \in [3R, T_0 - 3R]$ be such that the set*

$$\{t \in [0, T_0] : \gamma_{\rho_0}^{g_\star}(t) \cap B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R)\} \tag{9.19}$$

*is connected. For $0 < R < R_0$ and $j = 1, 2$, let $\Xi_j(\sigma) := \Xi_j(0, g_\star, \sigma)$ be the maps $\Xi_j$ introduced in (9.10), which are defined using the perturbations $g_\sigma = g_\sigma(\rho_0, t_\star, R, g_\star, \sigma)$ introduced in (9.4).*

*Then, the map* $\partial_{\sigma_1}\Xi_1(0)|_{\Sigma_1} : \Sigma_1 \to \mathbb{R}^{2d-2}$ *is bijective with inverse bounded by* $Ce^{C|T_0-t_\star|}$ *and* $\partial_{\sigma_2}\Xi_1(0)|_{\Sigma_1} = 0$. *Also, for every fixed* $\sigma_1 \in \Sigma_1$, *the map* $\partial_{\sigma_2}\Xi_2(\sigma_1,0)|_{\Sigma_2} : \Sigma_2 \to \mathfrak{sp}(2(d-1))$ *is bijective with inverse bounded by* $Ce^{C|T_0-t_\star|}$, *and the map* $\partial_{\sigma_1}\Xi_2(0)|_{\Sigma_1} : \Sigma_1 \to \mathfrak{sp}(2(d-1))$ *is bounded by* $Ce^{C|T_0-t_\star|}$. *In particular,*

$$\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0} \text{ is bijective with inverse bounded by } Ce^{C|T_0-t_\star|}.$$

*Proof.* First, note that $\partial_\delta\Xi_1(\delta\sigma)|_{\delta=0} = (\partial_\delta \boldsymbol{u}_{\delta\sigma}(T_0), \partial_\delta \boldsymbol{\omega}_{\delta\sigma}(T_0))|_{\delta=0}$ since $\partial_s(\boldsymbol{u},\boldsymbol{\omega})(T_0, z_0, g_\star, 0) = 0$. Fix $\mathbf{v} \in \mathbb{R}^{2d-2}$. By the connectedness of the set (9.19), and the fact that $t_\star \in [3R, T_0 - 3R]$, the right hand side of the equation (9.13) is the same as that in Lemma A.2. In particular, taking $R_0$ small enough, Lemma A.2 together with (9.13), there are $R_0, C_1 > 0$ depending on $\|g\|_{C^2}$ such that for all $R < R_0$, there are $\sigma_1 = (\mathbf{A}, \mathbf{B}) \in \mathbb{R}^{2d-2}$ satisfying

$$(\partial_\delta \boldsymbol{u}_{\delta(\sigma_1,0)}(T_0), \partial_\delta \boldsymbol{\omega}_{\delta(\sigma_1,0)}(T_0)) = \mathbf{v}, \qquad \|\sigma_1\| \le C_1 e^{C_1|T_0-t_\star|}\|\mathbf{v}\|.$$

Thus, the map $\sigma_1 \mapsto (\partial_\delta \boldsymbol{u}_{\delta(\sigma_1,0)}(T_0), \partial_\delta \boldsymbol{\omega}_{\delta(\sigma_1,0)}(T_0))$, which equals $\partial_\delta\Xi_1(0)|_{\Sigma_1}$, is bijective with inverse bounded by $C_1 e^{C_1|T_0-t_\star|}$.

Checking that $\partial_\delta\Xi_1(\delta\sigma)|_{\delta=0} = 0$ when $\sigma_1 = 0$ is straightforward. It follows that $\partial_{\sigma_2}\Xi_1(0)|_{\Sigma_1} = 0$.

Next, since $\Xi_2(\delta\sigma) = [\tilde{\Delta}_{\delta\sigma}\tilde{\Delta}_0^{-1}](T(z_0, g_\star, \delta\sigma, T_0), z_0, g_\star)$, by (9.9)

$$\begin{aligned}
\partial_\delta\Xi_2(\delta\sigma)\Big|_{\delta=0} &= \left(\partial_\delta\Delta_{\delta\sigma} - \partial_\delta\Big[[\partial_s \boldsymbol{t}_{\delta\sigma}]^{-1}\begin{pmatrix}\partial_s \boldsymbol{u}_{\delta\sigma} \\ \partial_s \boldsymbol{\omega}_{\delta\sigma}\end{pmatrix}\begin{pmatrix}\partial_{u_0}\boldsymbol{t}_{\delta\sigma} & \partial_{\omega_0}\boldsymbol{t}_{\delta\sigma}\end{pmatrix}\Big]\right)\Big|_{\substack{s=T_0 \\ \delta=0}}\Delta_0^{-1}(T_0) \\
&= \left(\partial_\delta\Delta_{\delta\sigma}\big|_{\delta=0}\Delta_0^{-1}\right)\Big|_{s=T_0}.
\end{aligned} \tag{9.20}$$

Where to obtain the final equality, we observe that

$$\begin{pmatrix}\partial_{u_0}\boldsymbol{t} & \partial_{\omega_0}\boldsymbol{t}\end{pmatrix}\Big|_{\substack{s=T_0 \\ z_0=0}} = 0, \qquad \text{and} \qquad \begin{pmatrix}\partial_s \boldsymbol{u} \\ \partial_s \boldsymbol{\omega}\end{pmatrix}\Big|_{\substack{s=T_0 \\ z_0=0}} = 0.$$

We now write $Q_0$ for the solution of

$$\partial_s Q_0 = [L_0, Q_0] + F_R(\mathbf{A}, \mathbf{B}, s), \qquad Q_0(0) = 0. \tag{9.21}$$

Note that since $L_0(s)$, $F_R(\mathbf{A}, \mathbf{B}, s) \in \mathfrak{sp}(2(d-1))$ we have $Q_0(s) \in \mathfrak{sp}(2(d-1))$. Observe that, working with $\sigma = (\sigma_1, 0)$, we have $\partial_\delta\Xi_2(\delta\sigma)\Big|_{\delta=0} = \partial_{\sigma_1}\Xi_2(0)\sigma_1$ and so (9.14) yields

$$\big\|\partial_{\sigma_1}\Xi_2(0)\sigma_1\big\| = \|Q_0(T_0)\| \le Ce^{C|T_0-t_\star|}\|\sigma_1\|,$$

where $C = C(\|g\|_{\mathcal{C}^3})$, and hence $\|\partial_{\sigma_1}\Xi_2(0)|_{\Sigma_1}\| \le Ce^{C|T_0-t_\star|}$, as claimed.

Next, we show that $\partial_{\sigma_2}\Xi_2(0)|_{\Sigma_2}$ is invertible. Fix $\mathbf{L} \in \mathfrak{sp}(2(d-1))$. Then, we use Lemma A.2 to find $\sigma_2 = (\mathbf{C}, \mathbf{D}, \mathbf{E})$ such that the solution $Q_1$ to

$$\partial_s Q_1 = [L_0, Q_1] + \frac{1}{2}\begin{pmatrix}\mathbf{D}_k^j & \mathbf{E}^{jk} \\ \mathbf{C}_{jk} & -\mathbf{D}_j^k\end{pmatrix}\chi_R(0, s - t_\star), \qquad Q_1(0) = 0 \tag{9.22}$$

satisfies $Q_1(T_0) = \mathbf{L}$ and such that

$$\partial_{\sigma_2}\Xi_2(0)\sigma_2 = Q_1(T_0) = \mathbf{L}, \qquad \|\sigma_2\| \le Ce^{C|T_0-t_\star|}\|\mathbf{L}\|,$$

and hence $\partial_{\sigma_2}\Xi_2(0)$ is invertible as claimed.

The application Lemma A.2 requires a brief explanation. Observe that, again using the connectedness of (9.19), (9.22) is a differential equation on $\mathfrak{sp}(2(d-1))$ of the form

$$\partial_s Q_1 = \boldsymbol{R}Q_1 + \sum_i \boldsymbol{a}_i f_{i,t_\star}^R$$

where $\boldsymbol{R}(s)$ is a linear map on $\mathfrak{sp}(2(d-1))$ with $\|\boldsymbol{R}\|_{L^\infty} \leq C\|g\|_{\mathcal{C}^2}$ and the $\boldsymbol{a}_i$ are a basis for $\mathfrak{sp}(2(d-1))$.

Finally, we show that $\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0}$ has the required properties. Notice that, as a map from $\mathbb{R}^{2(d-1)} \times \mathfrak{sp}(2(d-1))$ to itself,

$$\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0} = \begin{pmatrix} \partial_{\sigma_1}\Xi_1 & \partial_{\sigma_2}\Xi_1 \\ \partial_{\sigma_1}\Xi_2 & \partial_{\sigma_2}\Xi_2 \end{pmatrix} = \begin{pmatrix} \partial_{\sigma_1}\Xi_1 & 0 \\ \partial_{\sigma_1}\Xi_2 & \partial_{\sigma_2}\Xi_2 \end{pmatrix}.$$

**Remark 9.9.** Here it is crucial that $\partial_{\sigma_1}\Xi_2 : \mathbb{R}^{2d-2} \to \mathfrak{sp}(2(d-1))$ so that we may think of the map $\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0}$ acting on $\mathbb{R}^{2d-2} \times \mathfrak{sp}(2(d-1))$. This follows from the fact that $F_R(\boldsymbol{A},\boldsymbol{B},s) \in \mathfrak{sp}(2(d-1))$.

In particular, we have

$$[\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0}]^{-1} = \begin{pmatrix} (\partial_{\sigma_1}\Xi_1)^{-1} & 0 \\ -[\partial_{\sigma_2}\Xi_2]^{-1}\partial_{\sigma_1}\Xi_2(\partial_{\sigma_1}\Xi_1)^{-1} & [\partial_{\sigma_2}\Xi_2]^{-1} \end{pmatrix},$$

from which the estimates on $[\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0}]^{-1}$ easily follow. $\qquad\square$

Before we prove that the estimates in Lemma 9.8 are stable under small changes of the metric or initial position, we need to control how much these changes affect $\partial_\sigma \Psi_0^{g_\star,T_0}|_{\sigma=0}$. This is done in our next lemma.

**Lemma 9.10.** *Given $K \subset \mathscr{G}^4$ bounded, there is $C > 0$ such that the following holds.*

*Let $g_\star \in K$ and $R_0$ be as in Lemma 9.8. Let $\rho_0 \in \widetilde{S^*M}$, $T_0 \in \mathbb{R}$, $0 < R < R_0$, and $t_\star \in [3R, T_0 - 3R]$ be such that the set $\{t \in [0,T_0] : \gamma_{\rho_0}^{g_\star}(t) \cap B(\gamma_{\rho_0}^{g_\star}(t_\star), 3R)\}$ is connected. Then, for all $g_0 \in \mathscr{G}^4$ with $\|g_0 - g_\star\|_{\mathcal{C}^4} \leq 1$, $z_0 \in \mathbb{R}^{d-1}(0,\varepsilon_f) \times B_{\mathbb{R}^{d-1}}(0,\varepsilon_f)$, and all $R > 0$,*

$$\|d_\sigma(\Psi_{z_0}^{g_0,T_0} - \Psi_0^{g_\star,T_0})|_{\sigma=0}\|_{\mathcal{C}^0} \leq C(1 + R^{-2})\big(\|g_0 - g_\star\|_{\mathcal{C}^1} + |z_0|\big)e^{C|T_0|}, \qquad (9.23)$$

*where $\Psi_{z_0}^{g_0}$ is the map introduced in (9.10), defined using the perturbations $g_\sigma = g_\sigma(\rho_0, t_\star, R, g_0, \sigma)$ introduced in (9.4).*

**Remark 9.11.** It is possible to replace $R^{-2}$ in (9.23) by $R^{-1}$ through a more careful study of the terms produced by perturbing $g_\star$ and $z_0$. In particular, using the fact that the $L^1$ norm of the worst terms produced gains a power of $R$. This would, in turn, lead to a slightly smaller power $\Omega_\nu$ in our main theorems: Theorems 1.1, 1.2, and 1.3. However, the form of the power would still be the same so we do not pursue this here.

*Proof.* First, observe that $\|H_{|\xi|_{g_0}} - H_{|\xi|_{g_\star}}\|_{\mathcal{C}^{k-1}} \leq \|g_0 - g_\star\|_{\mathcal{C}^k}$. In particular, defining $\rho_0(t)$ and $\rho_\star(t)$ such that $\partial_t\rho_0(t) = H_{|\xi|_{g_0}}(\rho_0(t))$, $\partial_t\rho_\star(t) = H_{|\xi|_{g_\star}}(\rho_\star(t))$, and $\rho_0(0) = \rho_\star(0)$, we obtain for

all $t \in \mathbb{R}$

$$|\rho_0(t) - \rho_\star(t)| \leq \int_0^t |H_{|\xi|_{g_0}}(\rho_0(s)) - H_{|\xi|_{g_\star}}(\rho_\star(s))|ds$$

$$\leq \int_0^t (\|g_0\|_{\mathcal{C}^2} + \|g_\star\|_{\mathcal{C}^2})|\rho_0(s) - \rho_\star(s)|ds + t\|g_0 - g_\star\|_{\mathcal{C}^1}.$$

Therefore, using Gronwall's inequality, there is $C = C(\|g_\star\|_{\mathcal{C}^2}, \|g_0\|_{\mathcal{C}^2})$ such that

$$d(\varphi_t^{g_0}(\rho), \varphi_t^{g_\star}(\rho)) \leq C\|g_0 - g_\star\|_{\mathcal{C}^1}e^{C|t|}, \qquad \text{for all } \rho \in T^*M \setminus \{0\}.$$

In addition, recall that for every metric $g$ there is $C = C(\|g\|_{\mathcal{C}^2})$ such that for all $\rho_1, \rho_2 \in T^*M \setminus \{0\}$

$$d(\varphi_t^g(\rho_1), \varphi_t^g(\rho_2)) \leq Cd(\rho_1, \rho_2)e^{C|t|}.$$

In particular, there is $C = C(\|g_\star\|_{\mathcal{C}^2})$ such that for $\|g_\star - g_0\|_{\mathcal{C}^2} \leq 1$, and $\rho_0, \rho_\star \in T^*M \setminus \{0\}$, $t \in \mathbb{R}$,

$$d(\varphi_t^{g_0}(\rho_0), \varphi_t^{g_\star}(\rho_\star)) \leq C(\|g_\star - g_0\|_{\mathcal{C}^1} + d(\rho_0, \rho_\star))e^{C|t|}. \tag{9.24}$$

Next, in analogy with (9.11), for $\sigma \in \Sigma$ let

$$\boldsymbol{u}_\sigma(s) := \boldsymbol{u}(s, z_0, g_0, \sigma), \qquad \boldsymbol{\omega}_\sigma(s) = \boldsymbol{\omega}(s, z_0, g_0, \sigma), \qquad \boldsymbol{t}_\sigma(s) = \boldsymbol{t}(s, z_0, g_0, \sigma).$$

and differentiate (9.15) to get

$$\partial_s \begin{pmatrix} \partial_\delta \boldsymbol{u}_{\delta\sigma}|_{\delta=0} \\ \partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{\delta=0} \end{pmatrix} = L_0 \begin{pmatrix} \partial_\delta \boldsymbol{u}_{\delta\sigma}|_{\delta=0} \\ \partial_\delta \boldsymbol{\omega}_{\delta\sigma}|_{\delta=0} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{A} \end{pmatrix}\chi_R(0, s - t_\star) + E_1, \qquad \partial_s\partial_\delta \boldsymbol{t}_{\delta\sigma}|_{\delta=0} = \tilde{E}_1, \tag{9.25}$$

where

$$\|\tilde{E}_1\| + \|E_1\| \leq C\Big((1 + R^{-2})(\|g_0 - g_\star\|_{\mathcal{C}^1} + |z_0|)e^{C|s|}\Big)\|\sigma\|. \tag{9.26}$$

Note that (9.25) reduces to (9.13) when $(z_0, g_0) = (0, g_\star)$. To obtain (9.25), we differentiate (9.15) and use (9.24) together with Lipschitz bounds on $g_0$, $g_\star$, and $\chi_R$ to estimate the change from (9.13).

The bound on (9.26) follows from combining (9.24) together with $|\partial_\delta \boldsymbol{u}| + |\partial_\delta \boldsymbol{\omega}| + |\partial_\delta \boldsymbol{\tau}| + |\partial_\delta \boldsymbol{t}| \leq Ce^{C|s|}\|\sigma\|$ where $C = C(\|g_0\|_{\mathcal{C}^2}, \|g_\star\|_{\mathcal{C}^3})$.

Similarly, differentiating again as to obtain (9.14),

$$\partial_s\big((\partial_\delta \Delta_{\delta\sigma}|_{\delta=0})\Delta_0^{-1}\big) = \frac{1}{2}\begin{pmatrix} \boldsymbol{D}_k^j & \boldsymbol{E}^{jk} \\ \boldsymbol{C}_{jk} & -\boldsymbol{D}_j^k \end{pmatrix}\chi_R(0, s - t_\star) + [L_0, (\partial_\delta \Delta_{\delta\sigma}|_{\delta=0})\Delta_0^{-1}] + F_R(\boldsymbol{A}, \boldsymbol{B}, s) + E_2,$$

$$\partial_s\partial_\delta\partial_{(u_0,\omega_0)}\boldsymbol{t}|_{\delta=0} = \tilde{E}_2$$

where

$$\|\tilde{E}_2\| + \|E_2\| \leq C(\|g_\star\|_{\mathcal{C}^4})\Big((1 + R^{-2})(\|g_0 - g_\star\|_{\mathcal{C}^1} + |z_0|)e^{Cs}\Big)\|\sigma\|.$$

By (9.13) and (9.14), the bounds on $E_1$ and $E_2$ directly imply that for any $s > 0$,

$$\|d_\sigma\big((\boldsymbol{u}, \boldsymbol{\omega}, \Delta_\sigma\Delta_0^{-1})(s, z_0, g_0, \sigma) - (\boldsymbol{u}, \boldsymbol{\omega}, \Delta_\sigma\Delta_0^{-1})(s, 0, g_\star, \sigma))|_{\sigma=0}\|$$
$$\leq C(\|g_\star\|_{\mathcal{C}^4})(1 + R^{-2})(\|g_0 - g_\star\|_{\mathcal{C}^1} + |z_0|)e^{Cs}.$$

In addition, computing as in (9.20), we have

$$\|d_\sigma\tilde{\Delta}_\sigma\tilde{\Delta}_0^{-1}(T_0, z_0, g_0, \sigma) - d_\sigma\Delta_\sigma\Delta_0^{-1}(T_0, z_0, g_0, \sigma)\| \leq C(\|g_\star\|_{\mathcal{C}^4})(1 + R^{-2})(\|g_0 - g_\star\|_{\mathcal{C}^1} + |z_0|)e^{CT_0}$$
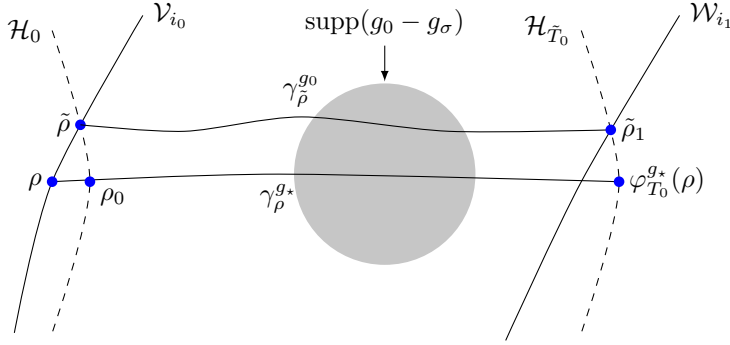
and hence, we obtain (9.23). $\qquad \square$

FIGURE 5. The figures shows the geometric setup used in Lemma 9.12 to pass from the special hypersurfaces defined using the coordinates $\Phi_\rho^{g_\star}$ to a given well separated set $(\Gamma, \tilde{\Gamma})$.

In what follows we use the notion of a well-separated set introduced in Definition 6.5.

**Lemma 9.12.** *Let $G \subset \mathscr{G}^4$ bounded, $F > 0$, and $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ be $F$-well separated for $G$. There exist $\delta > 0$, $R_0 > 0$, $C > 0$, and $C_1 > 0$, such that the following holds. Let $g_\star \in G$, $T_0 > 0$, $t_\star \in [3R_0, T_0 - 3R_0]$, $\rho \in \mathcal{V}_{i_0}$, and $i_1 \in \{1, \ldots, N\}$ such that*

$$\varphi_{T_0}^{g_\star}(\rho) \in \mathcal{W}_{i_1}, \qquad d(\varphi_{T_0}^{g_\star}(\rho), \mathcal{V}_{i_1}) < \delta,$$

*and the set*

$$\{t \in [0, T_0] : \gamma_\rho^{g_\star}(t) \cap B(\gamma_\rho^{g_\star}(t_\star), 3R_0) \neq \emptyset\} \tag{9.27}$$

*is connected. Then for all*

$$0 < R < R_0, \qquad 0 < \varepsilon \leq R^2 e^{-C_1|T_0|}/C_1,$$

*$g_0 \in \mathscr{G}^4$ and $\tilde{\rho} \in \mathcal{V}_{i_0}$ with*

$$\|g_\star - g_0\|_{\mathscr{G}^4} \leq \varepsilon, \qquad \|d(\tilde{\rho}, \rho)\| \leq \varepsilon,$$

*the map*

$$d_\sigma(\mathscr{P}_\sigma|_{q=\tilde{\rho}}, d_q\mathscr{P}_\sigma|_{q=\tilde{\rho}})|_{\sigma=0} : \Sigma \to T_{\mathscr{P}_0(\tilde{\rho})}\mathcal{W}_{i_0} \times T_{d_q\mathscr{P}_0|_{q=\tilde{\rho}}}\mathrm{Sp}(T_{\tilde{\rho}}\mathcal{W}_{i_0}; T_{\mathscr{P}_0(\tilde{\rho})}\mathcal{W}_{i_1}),$$

*is bijective with inverse bounded by $Ce^{C|T_0|}$, where $\mathscr{P}_\sigma : B_{\mathcal{V}_{i_0}}(\tilde{\rho}, e^{-C_1|T_0|}/C_1) \subset \mathcal{V}_{i_0} \to \mathcal{W}_{i_1}$ is the map*

$$\mathscr{P}_\sigma(q) := \varphi_{T^{g_\sigma}(q)}^{g_\sigma}(q), \qquad T^{g_\star}(\tilde{\rho}) = T_0, \qquad \varphi_{T^{g_\sigma}(q)}^{g_\sigma}(q) \in \mathcal{W}_{i_1},$$

*with $T^{g_\sigma}(q)$ smooth in $\sigma$ and $q$, $g_\sigma = g_\sigma(\rho, t_\star, R, g_0, \sigma)$ is the perturbed metric defined in (9.4), and $\mathrm{Sp}(T_{\tilde{\rho}}\mathcal{W}_{i_0}; T_{\mathscr{P}^0(\tilde{\rho})}\mathcal{W}_{i_1})$ denotes the symplectic linear maps from $T_{\tilde{\rho}}\mathcal{W}_i$ to $T_{\mathscr{P}_0(\tilde{\rho})}\mathcal{W}_{i_1}$.*

*Proof.* Let $(u_0, t_0)$ be such that $\Phi_\rho^{g_\star}(u_0, t_0) = \pi_M(\tilde{\rho})$. Next, let $\rho_0 := \gamma_\rho^{g_\star}(t_0)$. From now on we work with the system of coordinates $(u, t)$ induced by $(u, t) \mapsto \Phi_{\rho_0}^{g_\star}(u, t)$ and with $\mathcal{H}_s$ as defined in (9.2). Note that $\tilde{\rho} \in \mathcal{V}_{i_0} \cap \mathcal{H}_0$.

Let $\tilde{\rho}_1 := \mathscr{P}_0(\tilde{\rho}) = \varphi_{T^{g_0}(\tilde{\rho})}^{g_0}(\tilde{\rho}) \in \mathcal{W}_{i_1}$, and $\tilde{T}_0$ be such that $\tilde{\rho}_1 \in \mathcal{W}_{i_1} \cap \mathcal{H}_{\tilde{T}_0}$ Define the maps

$$\mathcal{Q}_{0,\sigma} : \mathcal{W}_{i_0} \to \mathcal{H}_0, \qquad \mathcal{Q}_{1,\sigma} : \mathcal{H}_{\tilde{T}_0} \to \mathcal{W}_{i_1},$$

by

$$\mathcal{Q}_{0,\sigma}(q) = \varphi_{S_0(q)}^{g_\sigma}(q), \qquad \mathcal{Q}_{1,\sigma}(q) = \varphi_{S_1(q)}^{g_\sigma}(q),$$

where $S_0(q)$ is the real number with smallest absolute value such that $\varphi_{S_0(q)}^{g_\sigma}(q) \in \mathcal{H}_{t_0}$, and $S_1(q)$ is the real number with smallest absolute value such that $\varphi_{S_1(q)}^{g_\sigma}(q) \in \mathcal{W}_{i_1}$.

Note that, by construction, the perturbation is supported on $(u,t)$ such that $|t - t_\star| \leq \sqrt{2}R$. Hence, since $t_\star > 3R_0$ and $R < R_0$, the perturbation is supported away from $\mathcal{W}_{i_0}$ and $\mathcal{W}_{i_1}$. Therefore, choosing $R_0$ small enough, $\mathcal{Q}_{0,\sigma}(q) = \mathcal{Q}_{0,0}(q)$ for all $q$ with $d(q,\rho) \ll R_0$ and $\mathcal{Q}_{1,\sigma}(q) = \mathcal{Q}_{1,0}(q)$ for all $q$ with $d(q,\rho_1) \ll R_0$.

In the $(u,t)$ coordinate system, with $\omega$ dual to $u$, we write $\mathcal{Z}_{0,\sigma}$ and $\mathcal{Z}_{1,\sigma}$ for the $(u,\omega)$ representation of $\mathcal{Q}_{0,\sigma}$ and $\mathcal{Q}_{1,\sigma}$ respectively. Next, consider the map

$$\Psi_{z_0}^{g_0,\tilde{T}_0}(\sigma) := (\Xi_1(z_0,g_0,\sigma), \Xi_2(z_0,g_0,\sigma))$$

defined in (9.10) with $\tilde{T}_0$ in place of $T_0$. Using that the perturbation is supported away from $\mathcal{W}_{i_0}$ and $\mathcal{W}_{i_1}$, we have $\mathcal{Z}_{j,\sigma}(z_0) = \mathcal{Z}_{j,0}(z_0)$ for $|z_0| \ll R_0$.

Now, abbreviating $\Xi_{j,\sigma}(z_0) := \Xi_j(z_0,g_0,\sigma)$, and putting $\tilde{\Xi}_{2,\sigma}(z_0) := \Xi_{2,\sigma}(z_0)\tilde{\Delta}_0(T,z_0,g_0)$, with $\tilde{\Delta}_0$ as in (9.9) and $T = T(z_0,g_0,\sigma,\tilde{T}_0)$ as in (9.10) with $T_0$ replaced by $\tilde{T}_0$, we have (in the $(u_0,\omega_0)$ coordinates),

$$\mathscr{P}_\sigma = \mathcal{Z}_{1,\sigma} \circ \Xi_{1,\sigma} \circ \mathcal{Z}_{0,\sigma}, \qquad d_\rho \mathscr{P}_\sigma = d_\rho \mathcal{Z}_{1,\sigma} \circ \tilde{\Xi}_{2,\sigma} \circ d_\rho \mathcal{Z}_{0,\sigma}. \tag{9.28}$$

Thus, since for $|z_0| < \varepsilon$ with $C_1$ to be determined later, and $\|\sigma\| \leq 1$, we have $|\Xi_{1,\sigma}(\mathcal{Z}_{0,\sigma}(z_0))| \ll R_0$ and $|z_0| \ll R_0$, and hence we may replace $\mathcal{Z}_{j,\sigma}$ by $\mathcal{Z}_{j,0}$ in (9.28) to obtain

$$d_\sigma((\mathscr{P}_\sigma(z_0), d_\rho \mathscr{P}_\sigma(z_0)))|_{\sigma=0}$$
$$= \Big( d_\rho \mathcal{Z}_{1,0}|_{\Xi_{1,0}(\mathcal{Z}_{0,0}(z_0))} d_\sigma \Xi_{1,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0} \, , \, d_\rho^2 \mathcal{Z}_{1,0}|_{\Xi_{1,0}(\mathcal{Z}_{0,0}(z_0))} \tilde{\Xi}_{2,0} d_\rho \mathcal{Z}_{0,0}$$
$$+ d_\rho \mathcal{Z}_{1,0}|_{\Xi_1(\mathcal{Z}_{0,0}(z_0))} d_\sigma \tilde{\Xi}_{2,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0} d_\rho \mathcal{Z}_{0,0}|_{z_0} \Big).$$

Next, observe that by Lemmas 9.8 and 9.10

$$\|d_{\sigma_2}\Xi_{1,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0}\| \leq C(1 + R^{-2})\varepsilon e^{C|\tilde{T}_0|},$$

$$\|d_{\sigma_1}\Xi_{1,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0}\| + \|d_{\sigma_1}\Xi_{2,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0}\| \leq C e^{C|\tilde{T}_0|},$$

and, provided that $C_1$ is chosen large enough,

$$\|(d_{\sigma_1}\Xi_{1,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0})^{-1}\| + \|(d_{\sigma_2}\Xi_{2,\sigma}(\mathcal{Z}_{0,0}(z_0))|_{\sigma=0})^{-1}\| \leq C e^{C|\tilde{T}_1|}.$$

Together with the fact that

$$\|d_\rho^2 \mathcal{Z}_{k,0}\| + \|d_\rho \mathcal{Z}_{k,0}\| + \|(d_\rho \mathcal{Z}_{k,0})^{-1}\| \leq C, \qquad \|\Delta_0\| \leq C e^{C|\tilde{T}_0|},$$

this implies the lemma. To see this, observe that, taking $C_1$ large enough, the estimates above imply that $d_\sigma((\mathscr{P}_\sigma(q), d_\rho \mathscr{P}_\sigma(q)))|_{\sigma=0}$ takes the form $\begin{pmatrix} A & R \\ B & D \end{pmatrix}$, where $\|BA^{-1}R\| < \frac{1}{2}\|D^{-1}\|$. $\qquad \square$

9.2. **Probing families of metrics.** Before defining our family of probing metrics, we need one more auxiliary lemma which will allow us to control the size of the perturbation, $g_\sigma - g_0$, in the $C^\nu$ norm on symmetric tensors. Here, we recall that, to make sense of $C^\nu$ norms, we use the fixed norm as in Remark 1.5. Note that, in the $(u,t)$ coordinates from (9.4), controlling the size of the perturbation is trivial. However, we need to estimate how the coordinate change, $\Phi_\rho^{g_\star}$ affects these norms. This is the purpose of our next lemma.

**Lemma 9.13.** *Let $\nu \geq 2$ and $K \subset \mathscr{G}^\nu$ be bounded. Then there are $C_\nu > 0$ and $R_0 > 0$ such that for all $g_\star \in K$, $t_0 \in \mathbb{R}$, $0 < R < R_0$ and $v \in C^\nu(\mathbb{R}^n)$ supported in the ball $B((0,t_0), 2R)$, we have*

$$\|[(\Phi_\rho^{g_\star})^{-1}]^* v\|_{C^\nu} \leq C_\nu \|v\|_{C^\nu},$$

*Proof.* Once we understand the $C^\nu$ norms of $(\Phi_\rho^{g_\star})^{-1}$, it will remain to apply the Faà di Bruno formula. To do this, we first choose $R_0 \ll 1$ small enough such that for all $g_\star$ in $K$, $\Phi_\rho^{g_\star}$ is a diffeomorphism on $B((0,t_0), 2R)$ for any $t_0$ and $R < R_0$.

To estimate the $C^\nu$ norms of $(\Phi_\rho^{g_\star})^{-1}$, it is enough to estimate $(d\Phi_\rho^{g_\star}|_{(0,t)})^{-1}$ from below and $\|\Phi_\rho^{g_\star}\|_{C^\nu}$ from above. To see that estimating the inverse at $u = 0$ is sufficient, observe that

$$\|d\Phi_\rho^{g_\star}|_{(0,t)} - d\Phi_\rho^{g_\star}|_{(\tilde{u},t)}\| \leq \|\Phi_\rho^{g_\star}\|_{C^2} |\tilde{u}| \leq R_0 \|\Phi_\rho^{g_\star}\|_{C^2},$$

for $(\tilde{u}, t) \in \operatorname{supp} v$. Therefore, provided $\|(d\Phi_\rho^{g_\star}|_{(0,t)})^{-1}\| < \frac{1}{2R_0 \|\Phi_\rho^{g_\star}\|_{C^2}}$, we have $d\Phi_\rho^{g_\star}$ is also invertible in the support of $v$.

Observe that there is $C > 0$ such that for all $g \in K$ and $V \in TM$

$$C^{-1} |V|_g \leq |V|_{\widetilde{S^*M}} \leq C |V|_g. \tag{9.29}$$

Therefore, since

$$d\Phi_\rho^{g_\star}|_{(0,t_0)}(E^i(0)\delta_{u^i} + E^d(0)\delta_t) = E^d(t_0)\delta_t + E^i(t_0)\delta_{u^i},$$

we have

$$C|d\Phi_\rho^{g_\star}|_{(0,t_0)}(E^i(0)\delta_{u^i} + E^d(0)\delta_t)|^2_{\widetilde{S^*M}} \geq |d\Phi_\rho^{g_\star}|_{(0,t_0)}(E^i(0)\delta_{u^i} + E^d(0)\delta_t)|^2_{g_\star}$$
$$= |(E^i(0)\delta_{u^i} + E^d(0)\delta_t|^2_{g_\star} \geq C^{-1} |E^d(0)\delta_t + E^i(0)\delta_{u^i})|^2_{\widetilde{S^*M}}.$$

and hence $\|(d\Phi_\rho^{g_\star}|_{(0,t)})^{-1}\| \leq C^2$.

Next, Lemma 9.4 implies that there is $C$, depending only on $K$, such that $\|\Phi_\rho^{g_\star}\|_{C^\nu} \leq C$. In particular, we obtain that there is $c_0 > 0$ depending on $K$ such that for $R \leq c_0$ there is $C_\nu$ depending on $K$ and $\nu$ such that $\|(\Phi_\rho^{g_\star})^{-1}\|_{C^\nu} \leq C_\nu$ on $\Phi_\rho^{g_\star}(\operatorname{supp} v)$. Using this in the Faà di Bruno formula then completes the proof of the lemma. $\square$

Let $r > 0$ and $\{\rho_i(r)\}_{i=1}^{N(r)}$ be a maximal $r$ separated set in $\widetilde{S^*M}$ so that

$$\widetilde{S^*M} \subset \bigcup_{i=1}^{N(r)} B(\rho_i, r). \tag{9.30}$$

We associate to each $\rho_i(r)$, a family of perturbations defined by iterative application of the construction (9.4). In particular, let

$$\boldsymbol{\Sigma}(r) := \left(B_{\mathbb{R}^{\mathfrak{L}}}(0,1)\right)^{N(r)}$$

We define for $\delta, R > 0$, $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(r)$, and $g_\star \in \mathscr{G}^\nu$,

$$g_{\boldsymbol{\sigma}, g_\star}^{r, R, \delta} = g_\star + \sum_{j=1}^{N(r)} (g_{\delta\sigma_j}(\rho_j, 0, R, g_\star, \delta\sigma_j) - g_\star),$$

where $g_\sigma = g_\sigma(\rho_j, 0, R, g_\star, \sigma)$ is the metric perturbation of $g_\star$ defined along $\gamma_{\rho_j}^{g_\star}(t)$ in $R$-neighborhood of $\rho_j = \gamma_{\rho_j}^{g_\star}(0)$ (see (9.4)).

In what follows, we write $\mathscr{S}^\nu$ for the space of $\mathcal{C}^\nu$ symmetric tensors on $M$.

For $R, r, \delta > 0$ define the map $Q : \mathscr{G}^\nu \times \boldsymbol{\Sigma}(R) \to \mathscr{S}^\nu$ by

$$Q^{r, R, \delta}(g, \boldsymbol{\sigma}) := g_{\boldsymbol{\sigma}, g}^{r, R, \delta}. \tag{9.31}$$

**Lemma 9.14.** *Let $\iota : \mathscr{G}^\nu \to \mathscr{G}^{\nu-1}$ be the natural embedding. Then, for all $K \subset \mathscr{G}^\nu$ bounded, there is $\varepsilon_0 > 0$ such that if*

$$\delta R^{-1} r^{-2d+1} \max(r, R)^d < \varepsilon_0,$$

*then*

$$Q^{r, R, \delta} : K \times \boldsymbol{\Sigma}(R) \to \mathscr{G}^\nu. \tag{9.32}$$

*Moreover, with $\tilde{Q}^{r, R, \delta} := \iota \circ Q^{r, R, \delta} : \mathscr{G}^\nu \times \boldsymbol{\Sigma}(R) \to \mathscr{G}^\nu$ is Frechet differentiable, the map $g \to D_{(g, \boldsymbol{\sigma})} \tilde{Q}^{r, R, \delta}$ is continuous, and*

$$D_g \tilde{Q}^{r, R, \delta} : \mathscr{S}^{\nu-1} \to \mathscr{S}^{\nu-1}$$

*is bijective, where we have extended $D_g \tilde{Q}^{r, R, \delta}$ to $\mathscr{S}^{\nu-1}$ by density. Moreover, for all $K \subset \mathscr{G}^\nu$ bounded, there is $C > 0$ such that for all $\delta, r, R > 0$, $g \in K$, and $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}(R)$,*

$$\|D_{\boldsymbol{\sigma}} Q^{r, R, \delta}\|_{\ell^\infty \to \mathcal{C}^\nu} \leq C\delta R^{-1-\nu} r^{-2d+1} \max(r, R)^d, \tag{9.33}$$

$$\|D_g Q^{r, R, \delta}(g, \boldsymbol{\sigma}) - I\|_{\mathcal{C}^{\nu-1} \to \mathcal{C}^{\nu-1}} \leq C_1 \delta R^{-\nu} r^{-2d+1} \max(r, R)^d, \tag{9.34}$$

$$\|\partial_{\sigma_j} Q^{r, R, \delta}(g, \boldsymbol{\sigma})\|_{B_{\mathbb{R}^{\mathfrak{L}}}(0,1) \to \mathcal{C}^\nu} \leq C\delta R^{-1-\nu}, \tag{9.35}$$

$$\partial_{\sigma_j}^\alpha Q^{r, R, \delta}(g, \boldsymbol{\sigma}) = 0, \qquad |\alpha| \geq 2. \tag{9.36}$$

*Proof.* First observe that (9.33) implies (9.32), so we only check the estimates (9.33) through (9.36). Let $\rho \in \{\rho_i\}_{i=1}^{N(r)}$ and $\sigma \in B_{\mathbb{R}^{\mathfrak{L}}}(0,1)$, and set $g_\sigma := g_{\delta\sigma}(\rho, 0, R, g, \sigma)$. Observe that $(\Phi_\rho^g)^* g_{\delta\sigma} = (\Phi_\rho^g)^* g + \delta h_\sigma$ where $h_\sigma$ does not depend on $g$ and

$$|\partial^\alpha h_\sigma| \leq C_\alpha R^{-|\alpha|} \|\sigma\|, \qquad |\partial^\alpha D_\sigma h_\sigma| \leq C_\alpha R^{-|\alpha|}, \qquad \partial_\sigma^\alpha h_\sigma = 0, \quad |\alpha| \geq 2$$

and $h_\sigma$ is supported in a ball of radius $2R$ around $(\Phi_\rho^g)^{-1} \pi_M(\rho)$. Therefore, using Lemma 9.13

$$\|D_\sigma g_{\delta\sigma}\|_{\mathcal{C}^\nu} = \|\delta D_\sigma((\Phi_\rho^g)^{-1})^* h_\sigma\|_{\mathcal{C}^\nu} \leq C_\nu R^{-1-\nu} \delta, \qquad D_\sigma^\alpha g_{\delta\sigma} = 0, \qquad |\alpha| \geq 2.$$

This proves (9.35) and (9.36).

Next, let $g_i \in \mathscr{G}^\nu$, $i = 0, 1$ and note

$$(g_0)_{\delta\sigma} - g_0 - [(g_1)_{\delta\sigma} - g_1] = \left[ ([\Phi_\rho^{g_0}]^{-1})^* - ([\Phi_\rho^{g_1}]^{-1})^* \right] \delta h_\sigma.$$

Therefore, by Lemmas 9.3 and 9.13,

$$\|D_g[(g)_{\delta\sigma} - g]\|_{\mathcal{C}^{\nu-1} \to \mathcal{C}^{\nu-1}} \leq C_\nu \delta \|\sigma\| R^{-\nu}, \tag{9.37}$$

the map $(g, \sigma) \mapsto D_g[g_{\delta\sigma} - g]$ is continuous in the $\mathscr{G}^\nu \times B_{\mathbb{R}^{\mathfrak{L}}}(0, 1)$ topology, and the range of the derivative is supported in a ball of radius $CR$ around $\pi_M(\rho)$ where $C > 0$ depends only on $K$.

Using these estimates together with the definition of $Q^{r,R,\delta}$, the lemma follows after a counting argument. In particular, it is enough to bound

$$\sup_j \#\{i \in \{1, \ldots, N(r)\} : \ B(\pi_M(\rho_j), 2R) \cap B(\pi_M(\rho_i), 2R) \neq \emptyset\}.$$

To do this, note that since $\rho_j$ are a maximal $r$ separated set, there is a constant $\mathfrak{D}$ depending only on $d$ and $C$ such that there are $\{\mathcal{J}_\ell\}_{\ell=1}^{\mathfrak{D}}$, $\mathcal{J}_\ell \subset \{1, \ldots N(r)\}$ such that

$$\{1, \ldots, N(r)\} = \bigcup_\ell \mathcal{J}_\ell, \qquad B(\rho_j, Cr) \cap B(\rho_i, Cr) = \emptyset, \quad i \neq j, \, i, j \in \mathcal{J}_\ell.$$

In particular, for each fixed $j$,

$$\begin{aligned}
\#\{i \in \{1, \ldots, N(r)\} : \ &B(\pi_M(\rho_j), 2R) \cap B(\pi_M(\rho_i), 2R) \neq \emptyset\} \\
&\leq \#\{i \in \{1, \ldots, N(r)\} : \pi_M(\rho_i) \in B_M(\pi_M(\rho_j), 4R)\} \\
&\leq \#\{i \in \{1, \ldots, N(r)\} : \pi_M(B_{\widetilde{S^*M}}(\rho_i, 2r)) \subset B_M(\pi_M(\rho_j), 4R + 2r)\} \\
&\leq C_d \mathfrak{D} \, \mathrm{vol}\left(\rho \in \widetilde{S^*M} : \ \pi_M(\rho) \in B_M(\pi_M(\rho_j), 2r + 4R)\right) r^{-2d+1} \\
&\leq C_d \tilde{\mathfrak{D}} (\max(r, R))^d r^{-2d+1}
\end{aligned}$$

Combining this bound with (9.35) implies (9.33), and with (9.37) implies (9.34). $\qquad\square$

## 10. Construction of well separated sets

In this section, we construct a well-separated set for a small neighborhood of a metric $g_\dagger \in \mathscr{G}^3$. We then show that these well-separated sets are admissible in the sense of Definition 7.3 for the perturbation from Section 9.

### 10.1. Construction of the well-separated set. Define for $g \in \mathscr{G}^3$

$$\Psi^g : \mathbb{R} \times \widetilde{S^*M} \to \widetilde{S^*M}, \qquad \Psi^g(t, q) = \varphi_t^g(q),$$

and for $T > 0$, $V \subset \widetilde{S^*M}$, consider the geodesic tube

$$\mathcal{T}_V^g(T) := \Psi^g\big((-T, T) \times V\big).$$

**Lemma 10.1.** *Let $g_\dagger \in \mathscr{G}^3$. There exist $N > 0$, $\delta > 0$, hypsersurfaces $\mathcal{Z}_i \subset M$, $i = 1, \ldots, N$, and open subsets $V_i \Subset W_i \subset \widetilde{S^*_{\mathcal{Z}_i}M}$ such that $H_{|\xi|_{g_\dagger}}$ is transverse to $W_i$ and there is a bounded neighborhood $G \subset \mathscr{G}^3$ of $g_\dagger$ such that for all $g \in G$,*

$$\mathcal{T}_{W_i}^g(\delta) \cap \mathcal{T}_{W_j}^g(\delta) = \emptyset, \quad i \neq j, \qquad \text{and} \qquad \widetilde{S^*M} \subset \bigcup_{i=1}^N \mathcal{T}_{V_i}^g(\tfrac{1}{3}T_\dagger), \qquad (10.1)$$

*where $T_\dagger = \mathrm{inj}_{g_\dagger}(M)$. Moreover, $\Psi^g : (-\frac{T_\dagger}{3}, \frac{T_\dagger}{3}) \times W_i \to \widetilde{S^*M}$ is a $\mathcal{C}^1$-diffeomorphism onto its image.*

*Proof.* For each $q \in \widetilde{S^*M}$, let $\mathcal{Z}_q \subset M$ be a hypersurface with $q \in N^*\mathcal{Z}_q$ and such that we can work with Fermi normal coordinates associated to $\mathcal{Z}_q$. That is, we assume that there is $\varepsilon_q > 0$ such that $(x_1, x')$ are well defined coordinates for $|(x_1, x')| < 2\varepsilon_q$ with the property that $(0,0) = \pi_M(q)$ and $\mathcal{Z}_q = \{(x_1, x') : |x'| < \varepsilon_q, \ x_1 = 0\}$. Let

$$\mathcal{H}_q^{\varepsilon_q} := \{(0, x', \xi_1, \xi') \in \widetilde{S^*M} \ : \ |\xi_1| > \tfrac{1}{2}|\xi|_{g_\dagger}, \ |x'| < \varepsilon_q\},$$

where we write $(\xi_1, \xi')$ for the dual coordinates so that we have $q = (0, 0, 1, 0)$. Shrinking $\varepsilon_q$ if necessary, we may further assume that $\Psi^{g_\dagger} : (-\tfrac{1}{2}T_\dagger, \tfrac{1}{2}T_\dagger) \times \mathcal{H}_q^{3\varepsilon_q} \to \widetilde{S^*M}$ is a $\mathcal{C}^1$-diffeomorphism onto its image.

Next, for each $q \in \widetilde{S^*M}$ and $x_1 \in (-\varepsilon_q, \varepsilon_q)$ let $\mathcal{Z}_q(x_1) := \{(x_1, x') : |x'| < 3\varepsilon_q\}$. Observe that there exist $0 < \delta_q < \varepsilon_q$ and $C_q > 0$ such that for $x_1 \in (0, \delta_q)$ and $\rho \in \mathcal{H}_q^{2\varepsilon_q}$,

$$T_{x_1}(\rho) := \inf\left\{t > 0 : \pi_M(\varphi_t^{g_\dagger}(\rho)) \in \mathcal{Z}_q(x_1)\right\} < C_q|x_1|,$$

and the map $\rho \mapsto \varphi_{T_{x_1}(\rho)}^{g_\dagger}(\rho)$ is a $\mathcal{C}^1$-diffeomorphism from $\mathcal{H}_q^{2\varepsilon_q}$ onto its image. In particular, choosing $\delta_q$ small enough, we have for $x_1 \in (0, \delta_q)$, $V \subset \mathcal{H}_q^{2\varepsilon_q}$ open, we have

$$\mathcal{T}_V^{g_\dagger}(\tfrac{1}{5}T_\dagger) \subset \mathcal{T}_{V[x_1]}^{g_\dagger}(\tfrac{1}{4}T_\dagger), \qquad V[x_1] := \mathcal{T}_V^{g_\dagger}(\tfrac{1}{2}T_\dagger) \cap \widetilde{S^*\mathcal{Z}_q(x_1)}, \tag{10.2}$$

and $V[x_1]$ is $\mathcal{C}^1$-diffeomorphic to $V$. Therefore, we can shift slightly the initial position of our tubes without changing covering properties. Moreover, after possibly shrinking $\delta_q$, we may assume that

$$\sup_{x_1 \in [0, \delta_q)} \operatorname{diam}\left(V[x_1]\right) \leq 4\operatorname{diam}(V). \tag{10.3}$$

By compactness, there are $\{q_i\}_{i=1}^L \subset \widetilde{S^*M}$ such that $\widetilde{S^*M} \subset \bigcup_{i=1}^L \mathcal{T}_{\mathcal{H}_{q_i}^{\varepsilon_{q_i}}}(\tfrac{1}{5}T_\dagger)$. Fix $r > 0$. For each $i = 1, \ldots, L$, we let $\{\rho_{ij}\}_{j=1}^{N_i(r)}$ be a maximal $r/2$ separated set on $\mathcal{H}_{q_i}^{2\varepsilon_{q_i}}$ so that

$$\mathcal{H}_{q_i}^{\varepsilon_{q_i}} \subset \bigcup_{j=1}^{N_i(r)} B_{\mathcal{H}_{q_i}^{2\varepsilon_{q_i}}}(\rho_{ij}, r/2),$$

and there is $\mathfrak{D} > 0$, independent of $r > 0$, and a partition $\{1, \ldots, N_i(r)\} = \cup_{k=1}^{\mathfrak{D}} \mathcal{J}_k$ such that for any $k \in 1, \ldots, \mathfrak{D}$ and any $j_1 \neq j_2$, $j_1, j_2 \in \mathcal{J}_k$

$$B_{\mathcal{H}_{q_i}^{2\varepsilon_{q_i}}}(\rho_{ij_1}, 4r) \cap B_{\mathcal{H}_{q_i}^{2\varepsilon_{q_i}}}(\rho_{ij_2}, 4r) = \emptyset.$$

In particular, since $L$ is independent of $r$, there is $\tilde{\mathfrak{D}} > 0$ independent of $r$ such that the tubes

$$\mathcal{T}_{ij}(4r, \tfrac{1}{2}T_\dagger) := \Psi^{g_\dagger}((-\tfrac{1}{2}T_\dagger, \tfrac{1}{2}T_\dagger) \times B_{\mathcal{H}_{q_i}^{2\varepsilon_{q_i}}}(\rho_{ij}, 4r)),$$

can be divided into $\tilde{\mathfrak{D}}$ collections of disjoint tubes, and

$$\widetilde{S^*M} \subset \bigcup_{i=1}^L \mathcal{T}_{\mathcal{H}_{q_i}^{\varepsilon_{q_i}}}(\tfrac{1}{5}T_\dagger) \subset \bigcup_{i=1}^L \bigcup_{j=1}^{N_i(r)} \mathcal{T}_{ij}(\tfrac{1}{2}r, \tfrac{1}{5}T_\dagger). \tag{10.4}$$

We now fix $0 < r < \frac{\delta_0}{100 \mathfrak{D}}$, where $\delta_0 = \min_i \delta_{q_i}$. By (10.3), for $|x_1| < \delta_0$,

$$\max_{i,j} \operatorname{diam}\left(\mathcal{B}_{H_{q_i}^{2\varepsilon q_i}}(\rho_{ij}, 4r)\right)[x_1] \leq 32r.$$

Therefore, we may find $|x_1(i,j)| < \delta_0$ for all $i, j$ such that for $(i,j) \neq (k,\ell)$,

$$W_{ij} \cap W_{kl} = \emptyset, \qquad W_{ij} := \left(\mathcal{B}_{H_{q_i}^{2\varepsilon q_i}}(\rho_{i,j}, 4r)\right)[x_1(i,j)]. \tag{10.5}$$

(This is possible, for instance, using a greedy algorithm where we simply select $x_1(i,j)$ iteratively for each intersecting tube.)

Thus, by (10.2) and (10.4),

$$\widetilde{S^*M} \subset \bigcup_{i=1}^{L} \bigcup_{j=1}^{N_i(r)} \mathcal{T}_{V_{ij}}^{g_\dagger}(\tfrac{1}{4} T_\dagger), \qquad V_{ij} := \left(\mathcal{B}_{\mathcal{H}_{q_i}^{2\varepsilon q_i}}(\rho_{ij}, \tfrac{1}{2} r)\right)[x_1(i,j)].$$

Moreover, by (10.5) there is $\delta > 0$ such that for $(i,j) \neq (k,\ell)$,

$$\mathcal{T}_{W_{ij}}^{g_\dagger}(2\delta) \cap \mathcal{T}_{W_{kl}}^{g_\dagger}(2\delta) = \emptyset.$$

Finally, since the map $\mathscr{G}^3 \ni g \mapsto \Psi^g \in \mathcal{C}^1$ is continuous, $\Psi^{g_\dagger} : (-\tfrac{1}{2} T_\dagger, \tfrac{1}{2} T_\dagger) \times \mathcal{H}_q^{3\varepsilon q} \to \widetilde{S^*M}$ is a $\mathcal{C}^1$-diffeomorphism, which this proves the claim. $\square$

**Lemma 10.2.** *Let $g_\dagger \in \mathscr{G}^3$ and $\{(W_i, V_i)\}_{i=1}^{N}$ as in Lemma 10.1. Let $\mathcal{V}_i = V_i$, $V_i \Subset \mathcal{W}_i \Subset W_i$. Then, there is a bounded neighborhood $G \subset \mathscr{G}^3$ of $g_\dagger$ and $F > 0$ such that $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^{N}$ is a $F$-well separated set for $G$.*

*Proof.* Let $G$, $\delta$, and $T_\dagger$ as in Lemma 10.1. Shrinking $G$ if necessary, we may assume that $H_{|\xi|_g}$ is uniformly transverse to $\tilde{\Gamma}$ for $g \in G$.

We need to check the conditions (6.3). Let $F = \tfrac{1}{2}\delta$. Then, by the second condition in (10.1), the first condition in (6.3) follows with $C_\Gamma = \frac{4T_\dagger}{3\delta}$.

Next, by the first condition in (10.1), the second condition in (6.3) holds with $c_\Gamma = 2$. Finally, since for all $g \in G$, and $i$, $\Psi^g(-\frac{T_\dagger}{3}, \frac{T_\dagger}{3}) \times W_i \to \widetilde{S^*M}$ is a $\mathcal{C}^1$-diffeomorphism onto its image, and $\mathcal{W}_i \Subset W_i$,

$$\inf_{\rho \in \mathcal{W}_i} \inf\{t > 0 : \varphi_t^g(\rho) \in \overline{\mathcal{W}_i}\} \geq \frac{2T_\dagger}{3} = \frac{4T_\dagger}{3\delta} F \geq C_\Gamma F,$$

and hence, the last condition in (6.3) holds. $\square$

### 10.2. Admissibility of the family of perturbations from Section 9.

In order to show that the well separated-sets constructed above are admissible for the perturbations from Section 9, we study these perturbations near simple points.

We start by showing that one can find a ball through which the geodesic emanating from a simple point passes only once.

**Lemma 10.3.** *Let $K \subset \mathscr{G}^3$ bounded, $F < \inf_{g \in K} \mathrm{inj}_g(M)$. Then there are $c_0 > 0$ and $c_1 > 0$ such that if $0 < \alpha < c_0$, $n \in \mathbb{N}$, and $\rho \in \mathscr{S}_F(n, \alpha, g)$, then for every $R > 0$ and any maximal $R$ separated set $\{x_j\}_{j=1}^{J(R)}$ on $M$, and any $g \in K$*

$$\#\{j : I_g(x_j, R) \neq \emptyset \text{ and } I_g(x_j, 4R) \text{ is disconnected}\} \leq \frac{n^2 F^2}{\alpha c_1^n},$$

*where $I_g(x_j, R) := \{t \in [0, (n - \frac{1}{2})F] : \pi_M(\varphi_t^g(\rho)) \cap B(x_j, R) \neq \emptyset\}$.*

*Proof.* We claim that since $\rho \in \mathscr{S}_F(n, \alpha, g)$, there is $\varepsilon > 0$, $0 < c_1 < 1$ such that if $g \in K$, $\alpha < \varepsilon$, $0 \leq t, s \leq (n - \frac{1}{2})F$, and $|t - s| > c_1^{n-1}\alpha$, then

$$d(\varphi_s^g(\rho), \varphi_t^g(\rho)) \geq c_1^n \alpha. \tag{10.6}$$

To prove the claim in (10.6), recall that there is $C > 0$ such that $d(\varphi_t^g(\rho_1), \varphi_t^g(\rho_2)) \leq C^{|t|} d(\rho_1, \rho_2)$, for $g \in K$ and all $\rho_1, \rho_2 \in \widetilde{S^*M}$. In addition, by Lemma 6.6, for all $g \in K$ and $\rho \in \widetilde{S^*M}$ we have

$$d(\varphi_t^g(\rho), \rho) \geq |t|/C, \quad |t| \leq F. \tag{10.7}$$

By contradiction, suppose there are $0 \leq t \leq s - c_1^{n-1}\alpha$ and $s \leq (n - \frac{1}{2})F$ such that $d(\varphi_s^g(\rho), \varphi_t^g(\rho)) \leq c_1^n \alpha$. First, by (10.7), if $|t - s| \leq F$, we have

$$C^{-1}|s - t| \leq d(\varphi_{s-t}^g(\varphi_t^g(\rho)), \varphi_t^g(\rho)) \leq c_1^n \alpha.$$

This contradicts $t \leq s - c_1^{n-1}\alpha$ for $c_1$ chosen small enough. In particular, we have $|s - t| \geq F$. Next, applying the flow for time $-t$,

$$d(\varphi_{s-t}^g(\rho), \rho) \leq c_1^n C^t \alpha \leq c_1^n C^{nF} \alpha.$$

Since $|s - t| \geq \frac{1}{2}F$, this contradicts the fact that $\rho \in \mathscr{S}_F(n, \alpha, g)$ and hence proves the claim in (10.6).

We next claim that there is $c_0 > 0$ such that if $g \in K$, $\rho \in \mathscr{S}_F(n, \alpha, g)$, $t_0, s_0 \in [0, (n - \frac{1}{2})F]$, $|s_0 - t_0| \geq c_1^n \alpha$, and

$$d(\pi_M(\varphi_{t_0}^g(\rho)), \pi_M(\varphi_{s_0}^g(\rho))) < 4R,$$

then for $|t - t_0| < c_0$ and $|s - s_0| < c_0$,

$$d(\pi_M(\varphi_t^g(\rho)), \pi_M(\varphi_s^g(\rho))) > c_0(|t - t_0| + |s - s_0|)c_1^n \alpha - 8R. \tag{10.8}$$

We now prove (10.8). Without loss of generality, we may assume $c_1^n \alpha \geq 8R$ since otherwise (10.8) is trivial. Since we are working locally near $\varphi_{s_0}^g(\rho)$, we may work in geodesic normal coordinates such that $\varphi_{s_0}^g(\rho) = (0, e_1)$ where $e_1 := (1, 0, \ldots, 0)$ and $\varphi_{t_0}^g(\rho) = (x_1, \xi_1)$. Observe that by (10.6),

$$c_1^n \alpha \leq |(x_1, \xi_1) - (0, e_1)| \leq 4R + |\xi_1 - e_1|.$$

Thus, $|\xi_1 - e_1| \geq \frac{1}{2}c_1^n \alpha$. Then, define

$$(x(t, (x_0, \xi_0)), \xi(t, (x_0, \xi_0))) := \varphi_t^g(x_0, \xi_0).$$

We first claim that

$$|\xi(t_0 - s_0, (0, e_1)) + e_1| > \frac{1}{2}c_1^n \alpha. \tag{10.9}$$

Indeed, suppose not. Then,

$$|x(t_0 - s_0, (0, e_1))| + |\xi(t_0 - s_0, (0, e_1)) + e_1| \leq 4R + \tfrac{1}{2}c_1^n \alpha < c_1^n \alpha.$$

In particular, applying the backward flow, we have

$$d\Big( \big(x(t_0-s_0-t, (0, e_1)), \xi(t_0-s_0-t, (0, e_1))\big), \big(x(-t, (0, -e_1)), \xi(-t, (0, -e_1))\big) \Big) < C^{|t|}c_1^n \alpha. \quad (10.10)$$

Next, observe that

$$(x(-t, (0, -e_1)), \xi(-t, (0, -e_1))) = (x(t, (0, e_1)), -\xi(t, (0, e_1))).$$

Therefore, putting $t = \frac{t_0 - s_0}{2}$ in (10.10), we have

$$d\Big( \big(x(\tfrac{t_0-s_0}{2}, (0, e_1)), \xi(\tfrac{t_0-s_0}{2}, (0, e_1))\big), \big(x(\tfrac{t_0-s_0}{2}, (0, e_1)), -\xi(\tfrac{t_0-s_0}{2}, (0, e_1))\big) \Big) < C^{nF} c_1^n \alpha.$$

Choosing $c_1$ small enough (uniformly in $K$), this is a contradiction since $|\xi(\frac{t_0-s_0}{2}, (0, e_1))|_g = 1$. In particular, (10.9) holds.

Now, we prove (10.8). Since we work in geodesic normal coordinates at $x(s_0, (0, e_1))$, we have

$$\begin{aligned}
|x(s - s_0, (0, e_1)) &- x(t - t_0, (x_1, \xi_1))| \\
&\geq |x(s - s_0, (0, e_1)) - x(t - t_0, (0, \xi_1))| - |x(t - t_0, (0, \xi_1)) - x(t - t_0, (x_1, \xi_1))| \\
&\geq |(s - s_0)e_1 - (t - t_0)\xi_1| - (1 + C|t - t_0|)|x_1| \\
&\geq c_0 c_1^n \alpha(|s - s_0| + |t - t_0|) - 8R,
\end{aligned}$$

where in the last line we have used that $|\xi_1|_{g(x_1)} = |e_1| = 1$, $\min(|\xi_1 - e_1|, |\xi_1 + e_1|) > c_0 c_1^n \alpha$, and that $C|t - t_0| < Cc_0 \leq 1$. This proves the claim in (10.8).

We now use (10.8) to control the length of time for which the (projection to the base of the) segment of the geodesic near $t_0$ can be close to that near $s_0$. Let $\Omega := B\Big( \bigcup_{|s-s_0|<c_0} \pi_M(\varphi_s^g(\rho)), 8R \Big)$. We claim that

$$\left| \Big\{ t : \; |t - t_0| < c_0, \; B(\pi_M(\varphi_t^g(\rho)), 4R) \cap \Omega \neq \emptyset \Big\} \right| \leq \frac{CR}{c_0 c_1^n \alpha}, \qquad (10.11)$$

where $C$ depends only on $K$. To see (10.11), note that if $t$ lies in the left hand side, then there is $s$ such that $|s - s_0| < c_0$ and, by (10.8), we have

$$12R > d(\pi_M(\varphi_t(\rho)), \pi_M(\varphi_s(\rho))) \geq c_0 c_1^n \alpha |t - t_0| - 8R.$$

In particular, (10.11) holds.

Now, let $I_\ell := [\ell c_0, (\ell+2)c_0)$, $\ell = 1, \ldots, L-1$, $L = \lfloor \frac{F(n-\frac{1}{2})-2c_0}{c_0} \rfloor + 1$, $I_L := [Lc_0, F(n-\frac{1}{2})]$, and

$$\Omega_\ell := B\Big( \bigcup_{t \in I_\ell} \pi_M(\varphi_t^g(\rho)), 8R \Big).$$

Since for $|\ell - k| > 1$, $\inf_{t \in I_\ell, s \in I_k} |t - s| \geq 2c_0$, using (10.11) for each pair $\ell, k$ with $|\ell - k| > 1$,

$$\left| \{ t : \; \exists k, \ell \text{ s.t. } |k - \ell| > 1, \; t \in I_\ell, \; B(\pi_M(\varphi_t^g(\rho)), 4R) \cap \Omega_k \neq \emptyset \} \right| \leq \frac{CR}{c_0 c_1^n \alpha} \frac{n^2 F^2}{c_0^2}.$$

Next, note that there exists $c > 0$ such that the following holds. Let $g \in K$ and $j$ such that $I_g(x_j, R) \neq \emptyset$ and $I_g(x_j, 4R)$ is disconnected there is an interval of length $\geq cR$ inside the set

$$\{t : \exists k, \ell \text{ s.t. } |\ell - k| > 1, \, t \in I_\ell, \quad B(\pi_M(\varphi_t^g(\rho)), 4R) \cap \Omega_k \neq \emptyset\},$$

and therefore the lemma follows.

Indeed, to see the last claim, shrinking $c_0$ so that $4c_0 < \inf_{g \in K} \mathrm{inj}_g(M)$ if necessary, we can find $t_1 \in I_{k_1}$, $t_2 \in I_{k_2}$ such that $\pi_M(\varphi_{t_1}^g(\rho)) \in B(x_j, R)$ and $\pi_M(\varphi_{t_2}^g(\rho)) \in B(x_j, 4R)$, with $|t_1 - t_2| > 4c_0$ and hence $|k_1 - k_2| > 1$. Notice that $B(x_j, 4R) \subset \Omega_{k_2}$ and hence, since $\pi_M(\varphi_{t_1}^g(\rho)) \in B(x_j, R)$, we have $\pi_M(\varphi_t^g(\rho)) \in \Omega_{k_2}$ for $|t - t_1| \leq cR$ (here, again $c > 0$ is a constant depending only on $K$). $\qquad\square$

Let $Q^{r,R,\delta}$ be the map that defines the perturbation of the metric as introduced in (9.31). We now check that for any $\nu_0$ and $\mathfrak{y} > 2$, there is $b$ such that $(\Gamma, G)$ is a $(F, b, \mathfrak{y})$ admissible pair for $\{Q^{R^{\mathfrak{y}}, R, \delta}\}_{R, \delta}$. (see Definition 7.3).

**Lemma 10.4.** *Let $\nu_0 \geq 5$ and $g_\dagger \in \mathscr{G}^3$. Let $F > 0$, $G \subset \mathscr{G}^3$ be the bounded neighborhood $G \subset \mathscr{G}^3$ of $g_\dagger$, and $\{(\mathcal{W}_i, \mathcal{V}_i)\}_{i=1}^N$ be the $F$-well separated set for $G$ given by Lemma 10.2. Let $\Gamma := \bigsqcup_{i=1}^N \mathcal{V}_i$. Then, for $\mathfrak{y} > 2$,*

$$(\Gamma, G) \text{ is a } (F, b, \mathfrak{y}) - \text{admissible pair for } \{Q^{R, \delta}\}_{R, \delta},$$

*with $b := 7 - \mathfrak{y} + d(2\mathfrak{y} - 1)$ and $Q^{R, \delta} := Q^{R^{\mathfrak{y}}, R, \delta}$ as defined in (9.31).*

*Proof.* Let $K \subset G$ bounded in $\mathscr{G}^{\nu_0}$. We will show that there are $c > 0$, $\varepsilon > 0$ such that for all $\delta > 0$, $g \in K$, $\alpha \in (0, c)$, $n \in \mathbb{N}$, $0 < R < c^n \alpha$, $0 < \delta \leq R^b c^{n+1}$, and

$$\rho_0 \in \Gamma \cap \mathscr{S}_F(n, \alpha, g) \cap \mathscr{R}_F(n, \varepsilon, g),$$

then there are $i_0 \in \{1, \ldots, N(R)\}$, $\mathcal{I}_0$, and $m_0$ such that $B_\Gamma(\rho_0, R^{\mathfrak{y}}) \subset \mathcal{D}_{\mathcal{I}_0}^{(m_0)}[Q^{R, \delta}(g, \boldsymbol{\sigma})]$ for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)}$, $\boldsymbol{T}_{\mathcal{I}_0}^{(m_0)}[g](\rho) \in [(n-1)F - C_G \delta_G, \, nF + C_G \delta_G]$, and for all $\rho \in B_\Gamma(\rho_0, R^{\mathfrak{y}})$ and $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)}$, the lower bound in (7.6) holds.

By Lemma 6.12, for $c, R, \varepsilon$ small enough there are $\mathcal{I}$ and $m$ such that $\rho_0 \in \mathcal{D}_{\mathcal{I}}^{(m)}[g]$, $\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho) \in [(n-1)F - C_G \varepsilon, \, nF + C_G \varepsilon]$, and $d(\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho_0), \rho_0) < C_G \varepsilon$ holds.

Let $\delta_0$ be as in Lemma 6.13. Next, since

$$\sup_{g \in G} \sup_{\rho \in \tilde{\Gamma}} \inf\{t > 0 : \varphi_t^g(\rho) \in \Gamma\} < C_\Gamma F,$$

by removing elements from $\mathcal{I}$, there is $C > 0$ depending only on $G$ such that for any $\delta_0 > 0$, we find $(m_0, \mathcal{I}_0)$ such that $\boldsymbol{T}_{\mathcal{I}_0}^{(m_0)}[g](\rho_0) \in [(n-1)F - C_G \varepsilon, \, nF + C_G \varepsilon]$ and

$$\sup_{m \leq m_0} d(\mathcal{P}_{\mathcal{I}_0}^{(m)}[g](\rho_0), \Gamma) \leq \max(C_G \varepsilon, \delta_0).$$

In particular, for $c$ small enough (depending only on $G$) and

$$\delta < R^{1 - \mathfrak{y} + 3 + d(2\mathfrak{y} - 1)} c^{n+1}, \qquad 0 < \tilde{R} < c^{n+1}, \tag{10.12}$$

we have by Lemma 6.13 that $B(\rho_0, \tilde{R}) \subset \mathcal{D}_{\mathcal{I}_0}^{(m)}[Q^{R,\delta}(g, \boldsymbol{\sigma})]$, for all $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{N(R)}$ and $m \leq m_0$. Indeed, since

$$C_\Gamma F m_0 \geq \boldsymbol{T}_{\mathcal{I}_0}^{(m_0)}[g](\rho_0) \geq c_\Gamma F m_0 > F m_0,$$

we have $n \leq C_\Gamma m_0 + 2$ and $m_0 \leq (n+1)$. Thus, $\tilde{R} < c^n \leq c^{m_0}$ and the condition on $\delta$ yields $\|Q^{R,\delta}(g, \boldsymbol{\sigma}) - g\|_{\mathcal{C}^3} < c^{n+1} \leq c^{m_0}$ as needed, after applying Lemma 9.14 with $(R^{\mathfrak{y}}, R, \delta, 3)$ in place of $(r, R, \delta, \nu)$.

We now study the action of the perturbation $Q^{R,\delta}$ on $\Psi_{g,\mathcal{I}_0}^{R,\delta}$ as defined in (7.7). To do this, we will use Lemma 9.12 with $g_\star = g$, and some appropriate choice of $\rho$, $g_0$, and $T_0$.

Note that for any maximal $R$ separated set $\{x_j\}_{j=1}^{J(R)} \subset M$, since $F < \mathrm{inj}_g(M)$,

$$\#\{j : \text{ there exists } t \in [F/5, F/4] \text{ satisfying } \pi_M(\varphi_t^g(\rho_0)) \in B(x_j, R)\} \geq cF R^{-1}.$$

Therefore, by Lemma 10.3 if

$$\alpha < c_0 \qquad \text{and} \qquad \frac{n^2 F^2}{\alpha c_1^n} < cF R^{-1}, \tag{10.13}$$

then there is an index $j_0$ such that

$$\{t \in [0, (n - \tfrac{1}{2})F] : \pi_M(\varphi_t^g(\rho_0)) \cap B(x_{j_0}, 4R) \neq \emptyset\} \quad \text{is connected} \tag{10.14}$$

and there is $t \in [F/5, F/4]$ satisfying $\pi_M(\varphi_t^g(\rho_0)) \in B(x_{j_0}, R)$.

Moreover, there is $c = c(F) > 0$ such that, with $\{\rho_i\}_{i=1}^{N(R)}$ as in (9.30), if

$$R^{\mathfrak{y}} \leq c\tilde{R}, \tag{10.15}$$

then for some $i_0 = 1, \dots, N(R)$ satisfying $\pi_M(\rho_{i_0}) \in B(x_{j_0}, R + R^{\mathfrak{y}})$ there are $\rho_\star \in B(\rho_0, \tilde{R})$ and $t_\star \in (1/6F, F/3)$ such that $\varphi_{t_\star}^g(\rho_\star) = \rho_{i_0}$. In addition, if

$$\tilde{R} < c^n R, \tag{10.16}$$

then (10.14) implies

$$\{t \in [0, (n - \tfrac{1}{2})F] : \pi_M(\varphi_t^g(\rho_\star)) \cap B(x_{j_0}, \tfrac{7}{2}R) \neq \emptyset\} \quad \text{is connected.} \tag{10.17}$$

To obtain the lower bound in (7.6), we aim to apply Lemma 9.12. To do this, let $T_0 = \boldsymbol{T}_{\mathcal{I}_0}^{(m_0)}[g](\rho_\star)$, since $\boldsymbol{T}_{\mathcal{I}_0}^{(m_0)}[g](\rho_0) \in [(n-1)F - C_G\varepsilon, \ nF + C_G\varepsilon]$,

$$(n-1)F - C_G\varepsilon - C^n\tilde{R} \leq T_0 \leq nF + C_G\varepsilon + C^n\tilde{R}.$$

Now, we claim that

$$\{t \in [0, T_0] : \pi_M(\varphi_t^g(\rho_\star)) \cap B(\pi_M(\varphi_t^g(\rho_\star)), \tfrac{7}{2}R) \neq \emptyset\} \quad \text{is connected.} \tag{10.18}$$

Indeed, if $T_0 \leq (n - \tfrac{1}{2})F$, then this is true by (10.17). Suppose instead that $T_0 > (n - \tfrac{1}{2})F$. We need to show that there is no $t \in [(n - \tfrac{1}{2})F, T_0]$ such that $\pi_M(\varphi_t^g(\rho_\star)) \in B(x_j, \tfrac{7}{2}R)$. To do this, recall that $\rho_\star \in \mathcal{W}_l \subset \tilde{\Gamma}$, $\mathcal{W}_l \subset \widetilde{S_{\mathcal{Z}}^* M}$ for some hypersurface $\mathcal{Z} \subset M$, $H_{|\xi|_g}$ is transverse to $\mathcal{W}_l$ and $F$ is chosen small enough that there are local coordinates $\Omega \ni (y_1, y') \overset{\psi}{\mapsto} V \subset M$ with $\mathcal{Z} = \psi(\{y_1 = 0\} \cap \Omega)$, $\pi_M(\varphi_t(\mathcal{W}_l)) \subset V$ for $|t| \leq F$, and $H_{|\xi|_g}(\psi^{-1})^* y_1 > c > 0$ on $\varphi_t(\mathcal{W}_l)$ for $|t| \leq F$. Therefore, $y_1(\pi_M(\varphi_t(\rho_\star))) \geq ct$. In particular, since $\pi_M(\varphi_{t_\star}^g(\rho_\star)) \in B(x_{j_0}, R + R^{\mathfrak{y}})$

and $t_\star \in (1/6F, F/3)$, this implies for $R$ small enough (depending only on $G$ and $F$), we have $B(x_{j_0}, \frac{7}{2}R) \subset \{y_1 > 0\}$.

We claim that $\varphi_{T_0}^g(\rho_\star) \in \mathcal{W}_l$. To see this, note that if $\tilde{R}$ is small enough, then $\rho_0 \in \mathcal{W}_l$. This implies that $\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[g](\rho_0) \in \mathcal{W}_l$, and hence, since $B(\rho_0, \tilde{R}) \in \mathcal{D}_{\mathcal{I}_0}^{(m_0)}[g]$, $\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[g](\rho_\star) \in \mathcal{W}_l$. Therefore, the claim holds. Now, since $\varphi_{T_0}^g(\rho_\star) \in \mathcal{W}_l$, we have $y_1(\pi_M(\varphi_{T_0-s}^g(\rho_\star))) \leq -cs$ for $0 \leq s < F$. In particular, since $T_0 \leq nF + C_G\varepsilon + C^n\tilde{R} < (n+\frac{1}{2})F$ (for $\varepsilon$ and $\tilde{R}$ chosen small enough as above), for $t \in [(n-\frac{1}{2})F, T_0]$, $\pi_M(\varphi_t(\rho_\star)) \in \{y_1 \leq 0\}$, and hence $\pi_M(\varphi_t^g(\rho_\star)) \notin B(x_{j_0}, \frac{7}{2}R)$, as claimed. This proves (10.18).

Shrinking $\delta_0$ if necessary and collecting (10.15), (10.12),(10.13),(10.16), and that, by (9.33),

$$\|Q^{R,\delta}(g, \boldsymbol{\sigma}) - g\|_{\mathscr{G}^4} < C\delta R^{-5-\mathfrak{y}(2d-1)+d} < R^2\delta R^{-7-d(2\mathfrak{y}-1)+\mathfrak{y}}c^{n+1} < c^{n+1} < R^2 e^{-C_1(n+1)F}/C_1, \tag{10.19}$$

for $c$ chosen small enough, provided $\delta < R^b c^{n+1}$ with $b = 7-\mathfrak{y}+d(2\mathfrak{y}-1)$. we conclude that (10.18) yields that the hypotheses of Lemma 9.12 hold with $\rho = \rho_\star$, $g_\star = g$, $\varepsilon = \max(R^{\mathfrak{y}}, \|R^{\mathfrak{y}} - g_0\|_{\mathscr{G}^4})$, $g_0 = Q^{R,\delta}(g, \boldsymbol{\sigma})$, $R_0 = \frac{7}{6}R$, provided

$$R^{\mathfrak{y}} \leq c\tilde{R} \leq \min(c^{n+1}, c^n R^2), \qquad R \leq c^n \alpha, \qquad \delta < R^{7-\mathfrak{y}+d(2\mathfrak{y}-1)}c^{n+1}.$$

Next, noticing that, since $\varphi_{t_\star}^{g_\star}(\rho_\star) = \rho_{i_0}$,

$$Q^{R,\delta}(g, \boldsymbol{\sigma} + \delta\sigma_{i_0}) = g_{\delta\sigma_{i_0}}(\rho_{i_0}, 0, R, g_0, \delta\sigma_{i_0}) = g_{\delta\sigma_{i_0}}(\rho_\star, t_\star, R, g_0, \delta\sigma_{i_0}),$$

with $\mathscr{P}_\sigma : B_\Gamma(\rho_\star, R^{\mathfrak{y}}) \subset \mathcal{W}_{i_0} \to \mathcal{W}_{i_{m_0}}$ as defined in Lemma 9.12 with $T_0 = \boldsymbol{T}_{\mathcal{I}_0}^{(m_0)}[g](\rho_\star)$ we have

$$\Psi_{g,\mathcal{I}_0}^{R,\delta}(m_0, \boldsymbol{\sigma} + \delta\sigma_{i_0}, \rho) = \left( \mathcal{P}_{\mathcal{I}_0}^{(m_0)}[g_{\delta\sigma_{i_0}}](\rho), \, d_\rho\big(\mathcal{P}_{\mathcal{I}_0}^{(m_0)}[g_{\delta\sigma_{i_0}}]\big)(\rho) \right)$$

$$= (\mathscr{P}_{\delta\sigma_{i_0}}(\rho), d_\rho\mathscr{P}_{\delta\sigma_{i_0}}(\rho)).$$

Since $\mathfrak{y} > 2$

$$R^{\mathfrak{y}} < R^2 R^{\mathfrak{y}-2} < R^2 c^{n(\mathfrak{y}-2)} < R^2 e^{-C_1(n+1)F}/C_1, \tag{10.20}$$

for $c$ chosen small enough and $C_1$ as in Lemma 9.12. Note that we have used $\mathfrak{y} > 2$ and $d \geq 2$ to obtain the next to last inequality. In particular, (10.20) and (10.19) show that Lemma 9.12 implies (7.6): that for all $\rho \in B_\Gamma(\rho_\star, R^{\mathfrak{y}})$,

$$|d_{\sigma_{i_0}}\Psi_{g,\mathcal{I}_0}^{R,\delta}(m_0, \boldsymbol{\sigma}, \rho)v| \geq \delta c^n|v|, \qquad v \in T_{\sigma_{i_0}}B_{R^{\mathfrak{L}}(0,1)}.$$

From Definition 7.3 it follows that $(\Gamma, G)$ is $(F, b, \mathfrak{y})$ admissible for $Q^{R^{\mathfrak{y}}, R, \delta}$ with $b = 7-\mathfrak{y}+d(2\mathfrak{y}-1)$. $\qquad\square$

## 11. Proof of Theorem 1.3

We now use Proposition 8.9 to prove Theorem 1.3. Let $\nu \geq 5$ and $\Omega > \Omega_\nu$.

Let $\mathfrak{y} > 2$ to be chosen close to 2 later, and $\tilde{Q}^{\mathfrak{R},\delta}(g, \sigma) := Q^{\mathfrak{R}^{\mathfrak{y}}, \mathfrak{R}, \delta}(g, \sigma)$ with $Q$ as in (9.31). Let $\mathscr{F} := \{(F_\infty^{\mathfrak{R}_\varepsilon, \boldsymbol{\delta}_\varepsilon}, \infty)\}_\varepsilon$ be the $\mathscr{G}^{\nu-1}$ family of probing maps for $\mathscr{G}^\nu$ constructed as in Proposition 8.9 (see also Lemma 8.1).

Let $K_n \subset \mathscr{G}^\nu$ bounded such that $K_n \subset K_{n+1}$ and $\bigcup_n K_n = \mathscr{G}^\nu$. Then for each $g \in K_n$, by Lemmas 10.2 and 10.4, there are $r_g > 0$, $F_g > 0$, and a symplectic manifold $\Gamma_g$ such that $(\Gamma_g, G_g)$

is $F_g$ well separated with $G_g := B_{\mathscr{G}^3}(g, 2r_g) \cap \mathscr{G}^\nu$. Moreover, $(\Gamma_g, G_g)$ is $(F_g, b, \mathfrak{y})$ admissible for $\tilde{Q}^{\mathfrak{R}, \delta}$ with $b = 7 - \mathfrak{y} + d(2\mathfrak{y} - 1)$.

Since $\nu > 3$, $K_n$ is compact in $\mathscr{G}^3$, there are $\{g_{i,n}\}_{i=1}^{N_n}$ such that $K_n \subset \bigcup_{i=1}^{N_n} B_{\mathscr{G}^3}(g_{i,n}, r_{g_{i,n}})$. We relabel $\{g_i\}_{i=1}^\infty = \cup_n \cup_{i=1}^{N_n} \{g_{i,n}\}$. Next, fix $K \subset \mathscr{G}^\nu$ bounded and let $N_K$ be such that

$$K \subset \bigcup_{i=1}^{N_K} B_{\mathscr{G}^3}(g_i, r_{g_i}).$$

To each $i \in \{1, \ldots, N_K\}$ apply Proposition 8.9 to $(\Gamma_{g_i}, G_{g_i} \cap K)$ in place of $(\Gamma, K)$. Let $\varepsilon_i, C_i$ be the constants $\varepsilon_0, C$ given by the proposition. Let $\varepsilon_{K,1} = \min_{1 \leq i \leq N_K} \varepsilon_i$.

Next, let

$$\varepsilon_{K,2} = \sup\{\varepsilon > 0 \,:\, \|F_\infty^{\mathfrak{R}_\varepsilon, \delta_\varepsilon}(g, \boldsymbol{\sigma}) - g\|_{\mathscr{G}^\nu} < \min_{1 \leq i \leq N_K} r_{g_i} \text{ for all } g \in K, \boldsymbol{\sigma} \in \Sigma_\infty\},$$

and set $\varepsilon_K := \min(\varepsilon_{K,1}, \varepsilon_{K,2}) > 0$. By Proposition 8.9, for each $i \in \{1, \ldots, N_K\}$ and $0 < \varepsilon < \varepsilon_K$ there exists $S_{g,\varepsilon}(g_i) \subset \Sigma_\infty(\mathfrak{R}_\varepsilon)$ an $m_{\Sigma_\infty(\mathfrak{R})}$ measurable set such that for all $g \in K \cap G_{g_i}$

$$\{\boldsymbol{\sigma} \in \Sigma_\infty(\mathfrak{R}_\varepsilon) \,:\, F_\infty^{\mathfrak{R}_\varepsilon, \delta_\varepsilon}(g, \boldsymbol{\sigma}) \in L_\infty(\varepsilon, g_i)\} \subset S_{g,\varepsilon}(g_i), \qquad \sup_{g \in K \cap G_{g_i}} m_{\Sigma_\infty(\mathfrak{R})}\big(S_{g,\varepsilon}(g_i)\big) \leq \varepsilon, \quad (11.1)$$

where $L_\infty(\varepsilon, g_i) := \Big\{g \in \mathscr{G}^\nu \,:\, \exists n \text{ such that } \Gamma_{g_i} \cap \mathscr{R}_{F_{g_i}}(n, \beta_n(\varepsilon, g_i), g) \not\subseteq \mathscr{N}_{F_{g_i}}(n, \beta_n(\varepsilon, g_i), g)\Big\}$, with

$$\beta_n(\varepsilon, g_i) := \varepsilon^{C_i n^\gamma} C_i^{-n^\gamma} n^{-C_i \log \varepsilon^{-1} n^\gamma},$$

where, as in (8.79), $\gamma = \gamma_\nu(\mathfrak{y})$ is given by

$$\gamma_\nu(\mathfrak{y}) := 1 + \log_2\Big[\max\big(\tfrac{2\aleph(2\aleph+3)}{2\aleph+1} m_\nu(\mathfrak{y}), 2\aleph\mathfrak{y}\big)\Big], \qquad m_\nu(\mathfrak{y}) := 1 + \nu + (2d-1)\mathfrak{y} - d.$$

Here, we have used $m_\nu(\mathfrak{y}) = 1 + \max(\nu, 6) + (2d-1)\mathfrak{y} - d = \max(b, \vartheta_\nu)$ and that $\vartheta_\nu = 1 + \nu + (2d-1)\mathfrak{y} - d$ by (9.33) with $r = R^\mathfrak{y}$.

Then, let

$$\tilde{G}_i := \left\{g \in G_{g_i} \,\middle|\, \begin{array}{c} \text{there exists } C > 0 \text{ such that for all } n > 0 \\ \Gamma_{g_i} \cap \mathscr{R}_{F_{g_i}}(n, (Cn)^{-Cn^\gamma}, g) \subset \mathscr{N}_{F_{g_i}}(n, (Cn)^{-Cn^\gamma}, g) \end{array}\right\}. \quad (11.2)$$

We claim that $\tilde{G} := \bigcup_i \tilde{G}_i$ is $\mathscr{F}$-predominant. To see this, fix $\varepsilon > 0$. We claim that for all $g \in K$, there is an $m_{\Sigma_\infty(\mathfrak{R})}$-measurable set $S_{g,\varepsilon}$ such that

$$\{\boldsymbol{\sigma} \in \Sigma_\infty(\mathfrak{R}_\varepsilon) \,:\, F_\infty^{\mathfrak{R}_\varepsilon, \delta_\varepsilon}(g, \boldsymbol{\sigma}) \in \tilde{G}^c\} \subset S_{g,\varepsilon}, \qquad m_{\Sigma_\infty(\mathfrak{R})}(S_{g,\varepsilon}) \leq \varepsilon. \quad (11.3)$$

To obtain (11.3), fix $g \in K$. Then there is $i$ such that $g \in K \cap G_{\mathscr{G}^3}(g_i, r_{g_i})$ and $F_\infty^{\mathfrak{R}_\varepsilon, \delta_\varepsilon}(g, \boldsymbol{\sigma}) \in G_{g_i}$ for all $\boldsymbol{\sigma} \in \Sigma_\infty(\mathfrak{R}_\varepsilon)$ (by our choice of $\varepsilon_{K,2}$). Now, since for all $i$ and $\varepsilon > 0$ there is $C > 0$ such that $(Cn)^{-Cn^\gamma} \leq \beta_n(\varepsilon, g_i)$ for all $n$, we have

$$\tilde{G}^c \cap G_{g_i} \subset \bigcap_{s > 0} L_\infty(s, g_i) \cap G_{g_i}.$$

Therefore,

$$\big\{\boldsymbol{\sigma}\in\boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon):\ F_\infty^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\boldsymbol{\sigma})\in\tilde{G}^c\big\}\subset\big\{\boldsymbol{\sigma}\in\boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon):\ F_\infty^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\boldsymbol{\sigma})\in\bigcap_{s>0}L_\infty(s,g_i)\cap G_{g_i}\big\}$$

$$\subset\big\{\boldsymbol{\sigma}\in\boldsymbol{\Sigma}_\infty(\mathfrak{R}_\varepsilon):\ F_\infty^{\mathfrak{R}_\varepsilon,\boldsymbol{\delta}_\varepsilon}(g,\boldsymbol{\sigma})\in L_\infty(\varepsilon,g_i)\cap G_{g_i}\big\}\subset S_{g,\varepsilon}(g_i).$$

In particular, (11.3) follows from (11.1).

We next show that if $\Omega>\Omega_\nu$, we may choose $\mathfrak{y}>2$ such that if $g\in\tilde{G}$, and $\mathcal{W}=\{\mathcal{W}^{U_i}\}_i$ is a family of transition maps for $g$, then there is $C_\Omega>0$ such that for all $t>\mathsf{F}_{g_i}/2$

$$\mathscr{R}\Big(t,C_\Omega^{-C_\Omega(t+1)^\Omega-1},g\Big)\subset\mathscr{N}\Big(t,C_\Omega^{-C_\Omega(t+1)^\Omega-1},(g,\mathcal{W})\Big). \tag{11.4}$$

To see (11.4), let $g\in\tilde{G}$ and $i$ such that $g\in\tilde{G}_i$. Let $C_g>0$ be such that for all $n>0$

$$\Gamma_{g_i}\cap\mathscr{R}_{\mathsf{F}_{g_i}}\big(n,(C_gn)^{-C_gn^\gamma},g\big)\subset\mathscr{N}_{\mathsf{F}_{g_i}}\big(n,(C_gn)^{-C_gn^\gamma},g\big). \tag{11.5}$$

Suppose that $t>\mathsf{F}_{g_i}/2$ and $\rho\in\mathscr{R}(t,(\Lambda(t+1))^{-\Lambda(t+1)^\gamma},g)$ for some $\Lambda>0$. Define

$$s(\rho):=\inf\{s\geq0:\ \varphi_s^g(\rho)\in\Gamma_{g_i}\},\qquad\tilde{s}_\pm(\rho):=\inf\{s\geq0:\ \varphi_{\pm s}^g(\rho)\in\tilde{\Gamma}_{g_i}\}$$

where $\tilde{\Gamma}_{g_i}$ is defined as in (6.2). In what follows, $C,c$ are two positive constants that depend only on $K$. Note also that $\sup_{\rho\in S^*M}s(\rho)<C\,\mathrm{inj}_{g_i}(M)$. Define $\rho_+:=\varphi_{s(\rho)}^g(\rho)$ and observe that there is a choice of $\pm$ such that

$$d(\rho_+,\rho_t)<C(\Lambda(t+1))^{-\Lambda(t+1)^\gamma},\qquad\rho_t:=\varphi_{\pm\tilde{s}_\pm(\varphi_t^g(\rho))}(\varphi_t^g(\rho)),$$

$$|s(\rho)-\tilde{s}_\pm(\varphi_t^g(\rho))|<C(\Lambda(t+1))^{-\Lambda(t+1)^\gamma},\qquad\text{or}\qquad|s(\rho)+\tilde{s}_\pm(\varphi_t^g(\rho))|<C(\Lambda(t+1))^{-\Lambda(t+1)^\gamma}.$$

In particular, by Remark 6.10, there are $\mathcal{I}$ and $m$ such that

$$\mathcal{P}_{\mathcal{I}}^{(m)}[g](\rho_+)=\rho_t,\qquad|\boldsymbol{T}_{\mathcal{I}}^{(m)}[g](\rho_+)-t|<C(\Lambda(t+1))^{-\Lambda(t+1)^\gamma},$$

and so, choosing $\Lambda>0$ large enough, there is $n$ with $\frac{1}{4}<\frac{t}{\mathsf{F}_{g_i}}-\frac{1}{4}<n\leq\frac{t}{\mathsf{F}_{g_i}}+\frac{3}{4}$, such that $\rho_+\in\Gamma_{g_i}\cap\mathscr{R}_{\mathsf{F}_{g_i}}(n,(c\Lambda n)^{-c\Lambda n^\gamma},g)$. Furthermore, choosing $\Lambda$ large enough, we have $c\Lambda>C_g$ and hence, by (11.5),

$$\rho_+\in\mathscr{N}_{\mathsf{F}_{g_i}}\big(n,(C_gn)^{-C_gn^\gamma},g\big).$$

Now, let $v\in T_\rho(S^*M)_g$. Then, there is $w\in\mathbb{R}H_{|\xi|_g}(\rho)$ such that

$$d\varphi_{s(\rho)}\tilde{v}\in T_{\rho_+}\Gamma_{g_i},\qquad\tilde{v}:=w+v.$$

Therefore, choosing coordinates on $\Gamma_{g_i}$ near $\rho_+$ to identify $T_{\rho_+}\Gamma_{g_i}$ and $T_{\rho_t}\Gamma_{g_i}$, and writing the identification of tangent spaces as $\tilde{\mathcal{W}}_{\rho_t,\rho_+}:T_{\rho_+}\Gamma_{g_i}\to T_{\rho_t}\Gamma_{g_i}$,

$$\|(d\mathcal{P}_{\mathcal{I}}^{(m)}[g]-\tilde{\mathcal{W}}_{\rho_t,\rho_+})d\varphi_{s(\rho)}^g\tilde{v}\|\geq(C_gn)^{-C_gn^\gamma}\|\tilde{v}\|.$$

In particular,

$$\|(d\varphi_{-s(\rho)}^g d\mathcal{P}_{\mathcal{I}}^{(m)}[g]d\varphi_{s(\rho)}^g-d\varphi_{-s(\rho)}^g\tilde{\mathcal{W}}_{\rho_t,\rho_+}d\varphi_{s(\rho)}^g)\tilde{v}\|\geq c(C_gn)^{-C_gn^\gamma}\|\tilde{v}\|.$$

Now, there is $\tilde{t}$ with $|\tilde{t} - t| < C(\Lambda(t+1))^{-\Lambda(t+1)^\gamma}$, such that

$$d\varphi^g_{-s(\rho)} d\mathcal{P}_{\mathcal{I}}[g]^{(m)} d\varphi^g_{s(\rho)} = d\varphi^g_{\tilde{t}},$$

and hence

$$\|(d\varphi^g_{\tilde{t}} - d\varphi^g_{-s(\rho)} \tilde{\mathcal{W}}_{\rho_t,\rho_+} d\varphi^g_{s(\rho)}) \tilde{v}\| \geq c(C_g n)^{-C_g n^\gamma} \|\tilde{v}\|.$$

Applying $d\varphi^g_{t-\tilde{t}}$ on the left, we have

$$\|(d\varphi^g_t - d\varphi^g_{t-\tilde{t}} d\varphi^g_{-s(\rho)} \tilde{\mathcal{W}}_{\rho_t,\rho_+} d\varphi^g_{s(\rho)}) \tilde{v}\| \geq c(C_g n)^{-C_g n^\gamma} \|\tilde{v}\|. \tag{11.6}$$

Now, the map $d\varphi^g_{t-\tilde{t}} d\varphi^g_{-s(\rho)} \tilde{\mathcal{W}}_{\rho_t,\rho_+} d\varphi^g_{s(\rho)}$ identifies $T_\rho S^*M/\mathbb{R}H_p$ with $T_{\varphi^g_t(\rho)} S^*M/\mathbb{R}H_p$ and has uniformly bounded derivatives in $t$. Suppose $\mathcal{W} := \mathcal{W}^{U_k}$ is a transition map such that $(\rho, \varphi^g_t(\rho)) \in U_k$. Then, there is $C_{\mathcal{W}} > 0$

$$\|\mathcal{W}_{\varphi^g_t(\rho),\rho} - d\varphi^g_{t-\tilde{t}} d\varphi^g_{-s(\rho)} \tilde{\mathcal{W}}_{\rho_t,\rho_+} d\varphi^g_{s(\rho)}\| \leq C_{\mathcal{W}} d(\varphi^g_t(\rho), \rho) \leq C_{\mathcal{W}} (\Lambda(t+1))^{-\Lambda(t+1)^\gamma}. \tag{11.7}$$

Hence, letting $\Lambda$ be large enough, there is $C_{\mathcal{W},g} > 0$, depending on $g$ and $\mathcal{W}$, such that

$$\|(d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho}) \tilde{v}\| \geq (C_{\mathcal{W},g}(t+1))^{-C_{\mathcal{W},g}(t+1)^\gamma} \|\tilde{v}\|. \tag{11.8}$$

Next, we claim that

$$\|(d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho})v + \mathbb{R}H_p\| \geq (C_{\mathcal{W},g}(t+1))^{-C_{\mathcal{W},g}(t+1)^\gamma} \|v + \mathbb{R}H_p\|. \tag{11.9}$$

Note that then (11.9) implies $\rho \in \mathcal{N}(t, (C_{\mathcal{W},g}(t+1))^{-C_{\mathcal{W},g}(t+1)^\gamma}, g)$. Thus, increasing $\Lambda$ again if necessary, we would conclude that there is $C_{\mathcal{W},g} = C_{\mathcal{W},g}(g, \mathcal{W}) > 0$ such that, for $t > F_{g_i}/2$,

$$\mathcal{R}(t, (C_{\mathcal{W},g}(t+1))^{-C_{\mathcal{W},g}(t+1)^\gamma}, g) \subset \mathcal{N}(t, (C_{\mathcal{W},g}(t+1))^{-C_{\mathcal{W},g}(t+1)^\gamma}, g). \tag{11.10}$$

To prove (11.9) recall the smooth decomposition $T_\rho \widetilde{S^*M} = \mathcal{H}(\rho) \oplus \mathbb{R}H_p(\rho)$ and the fact that, by Definition (1.3), $\mathcal{W}$ preserves this splitting. Therefore, letting $v = v_{\mathcal{H}} + tH_p(\rho)$ with $v_{\mathcal{H}} \in \mathcal{H}(\rho)$,

$$(d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho})v = (d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho})\tilde{v} = (d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho})v_{\mathcal{H}} \in \mathcal{H}(\varphi^g_t(\rho)).$$

In particular, since $\mathcal{H}(\rho)$ is uniformly transverse to $H_p(\rho)$, there is $c > 0$ such that

$$\|(d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho})v + \mathbb{R}H_p(\varphi^g_t(\rho))\| = \inf_{s \in \mathbb{R}} \|(d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho})v + sH_p(\varphi^g_t(\rho))\|$$
$$\geq c\|(d\varphi^g_t - \mathcal{W}_{\varphi^g_t(\rho),\rho}\tilde{v}\|. \tag{11.11}$$

Next, since $v = \tilde{v} + w$, with $w \in \mathbb{R}H_p$, we have

$$\|\tilde{v}\| \geq \inf_{s \in \mathbb{R}} \|v + sH_p\| = \|v + \mathbb{R}H_p\|. \tag{11.12}$$

Combining (11.8), (11.11), and (11.12) implies (11.9).

Finally, to finish the proof of Theorem 1.3, we observe that $\mathfrak{y} \mapsto \gamma(\mathfrak{y})$ is an increasing function and $\Omega_\nu = \gamma_\nu(2)$. Therefore, for every $\Omega > \Omega_\nu$ there is $\mathfrak{y} > 2$ such that $\gamma(\mathfrak{y}) < \Omega$. In particular, there is $C_\Omega > 0$ such that for $t > F_{g_i}/2$,

$$C_\Omega^{-C_\Omega(t+1)^\Omega - 1} < (C(t+1))^{-C(t+1)^\gamma},$$

and hence (11.10) implies that (11.4) holds for $g \in \tilde{G}_i$ and $t > F_{g_i}/2$. Thus, since $\cup_i \tilde{G}_i$ is predominant, Theorem 1.3 follows after recalling that Lemma 6.6 implies that for all $i$ there is $c > 0$ such that for $g \in G_i$,

$$d(\varphi_t^g(\rho), \rho) \geq c|t|, \qquad |t| < F_{g_i}.$$

$\square$

## APPENDIX A. ELEMENTARY CONTROL THEORY FOR ODES

Let $\aleph \in \mathbb{N}$, $R > 0$, $t_\star \in \mathbb{R}$, and for $1 \leq i \leq \aleph$ and $t \in \mathbb{R}$ define

$$f_{i,t_\star}^R(t) := \tfrac{1}{R}\chi(\tfrac{1}{R}(t - t_\star))\mathbf{e}_i, \qquad \mathbf{e}_i := (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{\aleph-i})^t,$$

where $\chi \in C_c^\infty((-2, 2); [0, 1])$ with $\int \chi = 1$.

**Lemma A.1** (Observation). *Let $\mathbf{L} \in \mathbf{L}^\infty(\mathbb{R}; \mathbb{M}(\aleph))$ and $\mathbf{u_0} \in \mathbb{R}^\aleph$ and suppose that $\mathbf{u}$ solves*

$$\dot{\mathbf{u}} + \mathbf{L}^*\mathbf{u} = 0, \qquad \mathbf{u}(0) = \mathbf{u_0}.$$

*Then, there are $C, R_0 > 0$ depending on $\|\mathbf{L}\|_{\mathbf{L}^\infty}$ such that for any $t_0 \in \mathbb{R}$, $0 < R < R_0$, and $T \in \mathbb{R}$*

$$\|\mathbf{u}(T)\|^2 \leq Ce^{C(|T-t_\star|)} \sum_{i=1}^\aleph |\langle \mathbf{u}, f_{i,t_\star}^R \rangle_{L_t^2}|^2.$$

*Proof.* Note that

$$\langle \mathbf{u}, f_{i,t_\star}^R \rangle - \mathbf{u}_i(t_0) = \langle \mathbf{u}(t_0), f_{i,t_\star}^R \rangle - \mathbf{u}_i(t_0) - \iint_0^t \mathbf{L}^*(s)\mathbf{u}(s)\overline{f_{i,t_\star}^R(t)}dsdt$$

$$= -\iint_0^t \mathbf{L}^*(s)\mathbf{u}(s)\overline{f_{i,t_\star}^R(t)}dsdt.$$

Then, since there is $C > 0$ depending only on $\|\mathbf{L}^*\|_{L^\infty}$ such that $\|\mathbf{u}(s)\| \leq C\|\mathbf{u}(t_0)\|$ for $|s - t_0| \leq 1$, we obtain for $R < \frac{1}{2}$,

$$\left| \iint_0^t \mathbf{L}^*(s)\mathbf{u}(s)\overline{f_{i,t_\star}^R(t)}dsdt \right| \leq C\|\mathbf{u}(t_0)\| \int_{-2R}^{2R} \|\mathbf{L}^*(s)\|ds \leq CR\|\mathbf{u}(t_0)\|\|\mathbf{L}^*\|_{L^\infty}.$$

In particular, $\left|\langle \mathbf{u}, f_{i,t_\star}^R \rangle - \mathbf{u}_i(t_0)\right| \leq CR\|\mathbf{L}^*\|_{L^\infty}\|\mathbf{u}(t_0)\|$ for $R < \frac{1}{2}$. Therefore,

$$\sum_{i=1}^\aleph \left|\langle \mathbf{u}, f_{i,t_\star}^R \rangle\right|^2 \geq \|\mathbf{u}(t_0)\|^2(\tfrac{1}{2} - CR^2\|\mathbf{L}^*\|_{L^\infty}^2).$$

In particular, taking $R < \frac{1}{2C\|\mathbf{L}^*\|_{L^\infty}}$, we have

$$\sum_{i=1}^\aleph \left|\langle \mathbf{u}, f_{i,t_\star}^R \rangle\right|^2 \geq \tfrac{1}{4}\|\mathbf{u}(t_0)\|^2.$$

The claim now follows from standard estimates on first order ODE's. $\square$

**Lemma A.2** (Control). *Let $\mathbf{L} \in \mathbf{L}^\infty(\mathbb{R}; \mathbb{M}(\aleph))$. There are $C, R_0 > 0$, depending only on $\|\mathbf{L}\|_{\mathbf{L}^\infty}$, such that for all $0 < R < R_0$, $\mathbf{v} \in \mathbb{R}^\aleph$, $t_\star > 2R$ and $T > t_\star + 2R$, there is $\mathbf{a} \in \mathbb{R}^\aleph$ such that the solution, $\mathbf{w}$ to*

$$\dot{\mathbf{w}} - \mathbf{L}\mathbf{w} = \sum_{i=1}^{\aleph} \mathbf{a_i}\mathbf{f}^{\mathbf{R}}_{\mathbf{i},\mathbf{t_\star}}, \qquad \mathbf{w}(0) = \mathbf{0} \tag{A.1}$$

*satisfies $\mathbf{w}(T) = \mathbf{v}$ and, moreover, $\|\mathbf{a}\| \leq Ce^{C|T-t_\star|}\|\mathbf{v}\|$.*

*Proof.* Define a map $K : \mathbb{R}^\aleph \to \mathbb{R}^\aleph$ by

$$K(\mathbf{u_0}) = \sum_i \int \langle \mathbf{u}, f^R_{i,t_\star}\rangle(s)ds\,\mathbf{e}_i$$

where $\mathbf{u}$ solves

$$\dot{\mathbf{u}} + \mathbf{L}^*\mathbf{u} = 0, \qquad \mathbf{u}(0) = \mathbf{u_0}. \tag{A.2}$$

Next, fix $\mathbf{v} \in \mathbb{R}^\aleph$ and define the linear map $\ell_{\mathbf{v}} : K(\mathbb{R}^\aleph) \to \mathbb{R}$ by

$$\ell_{\mathbf{v}}(K(\mathbf{u_0})) = \langle \mathbf{u}(T), \mathbf{v}\rangle$$

where again $\mathbf{u}$ is as in (A.2). Then, by Lemma A.1, there are $R_0, C > 0$ such that for $0 < R < R_0$,

$$\|\mathbf{u}(T)\| \leq Ce^{C|T-t_\star|}\|K(\mathbf{u_0})\|.$$

In particular, $\ell_{\mathbf{v}}$ is a bounded linear function on the subspace $K(\mathbb{R}^\aleph)$ with $\|\ell_{\mathbf{v}}\| \leq Ce^{C|T-t_\star|}\|\mathbf{v}\|$. Thus, by the Hahn–Banach theorem, we can extend $\ell$ to a linear functional on $\mathbb{R}^\aleph$ with the same norm and there is $\mathbf{a} \in \mathbb{R}^\aleph$ with

$$\|\mathbf{a}\| \leq Ce^{C|T-t_\star|}\|\mathbf{v}\|$$

such that $\ell(\mathbf{b}) = \langle \mathbf{b}, \mathbf{a}\rangle$, for all $\mathbf{b} \in \mathbb{R}^\aleph$. Now, suppose that $\mathbf{w}$ solves

$$\dot{\mathbf{w}} - \mathbf{L}\mathbf{w} = \sum_i \mathbf{a_i}\mathbf{f}^{\mathbf{R}}_{\mathbf{i},\mathbf{t_\star}}, \qquad \mathbf{w}(\mathbf{T}) = \mathbf{v}.$$

Then, integrating by parts, we have for $\mathbf{u}$ solving (A.2),

$$\int_0^T -\langle \sum_i \mathbf{a}_i f^R_{i,t_\star}, \mathbf{u}\rangle(s)ds = \int_0^T \langle \mathbf{w}(s), \dot{\mathbf{u}} + \mathbf{L}^*\mathbf{u}\rangle(s)ds + \langle \mathbf{w}(0), \mathbf{u}(0)\rangle - \langle \mathbf{w}(T), \mathbf{u}(T)\rangle$$

$$= \langle \mathbf{w}(0), \mathbf{u_0}\rangle - \langle \mathbf{v}, \mathbf{u}(T)\rangle.$$

But, by construction of $\mathbf{a}$, and that $T > t_\star + 2R > 4R$,

$$\Big\langle \sum_i \int \langle \mathbf{u}, f^R_{i,t_\star}\rangle(s)ds\,\mathbf{e}_i, \mathbf{a}\Big\rangle = \langle \mathbf{u}(T), \mathbf{v}\rangle.$$

In particular, $\langle \mathbf{w}(0), \mathbf{u_0}\rangle = 0$ for all $\mathbf{u_0} \in \mathbb{R}^\aleph$ and hence $\mathbf{w}(0) = 0$ and the lemma is proved. $\qquad \square$

## References

[Abr70] R. Abraham. Bumpy metrics. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 1–3. Amer. Math. Soc., Providence, R.I., 1970.

[Ano82] D. V. Anosov. Generic properties of closed geodesics. *Izv. Akad. Nauk SSSR Ser. Mat.*, 46(4):675–709, 896, 1982.

[Ava56] Vojislav G. Avakumović. Über die Eigenfunktionen auf geschlossenen Riemannschen Mannigfaltigkeiten. *Math. Z.*, 65:327–344, 1956.

[Bér77] Pierre H. Bérard. On the wave equation on a compact Riemannian manifold without conjugate points. *Math. Z.*, 155(3):249–276, 1977.

[Bon17] Yannick Bonthonneau. The $\Theta$ function and the Weyl law on manifolds without conjugate points. *Doc. Math.*, 22:1275–1283, 2017.

[Bow72] Rufus Bowen. Periodic orbits for hyperbolic flows. *Amer. J. Math.*, 94:1–30, 1972.

[BW17] B. Bourgain and N. Watt. Mean square of zeta function, circle problem and divisor problem revisited. *arXiv:1709.04340*, 2017.

[CG20] Yaiza Canzani and Jeffrey Galkowski. Weyl remainders: an application of geodesic beams. *arXiv:2010.03969*, 2020.

[Con10] Gonzalo Contreras. Geodesic flows with positive topological entropy, twist maps and hyperbolicity. *Ann. of Math. (2)*, 172(2):761–808, 2010.

[DG75] J. J. Duistermaat and V. W. Guillemin. The spectrum of positive elliptic operators and periodic bicharacteristics. *Invent. Math.*, 29(1):39–79, 1975.

[dlCS21] Ralph J. de la Cruz and Philip Saltenberger. On the density of semisimple matrices in indefinite scalar product spaces. *Electron. J. Linear Algebra*, 37:387–401, 2021.

[DZ16] Semyon Dyatlov and Maciej Zworski. Dynamical zeta functions for Anosov flows via microlocal analysis. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(3):543–577, 2016.

[Hal15] Brian Hall. *Lie groups, Lie algebras, and representations*, volume 222 of *Graduate Texts in Mathematics*. Springer, Cham, second edition, 2015. An elementary introduction.

[Hin84] N. Hingston. Equivariant Morse theory and closed geodesics. *J. Differential Geom.*, 19(1):85–116, 1984.

[HN12] Joachim Hilgert and Karl-Hermann Neeb. *Structure and geometry of Lie groups*. Springer Monographs in Mathematics. Springer, New York, 2012.

[Hör68] Lars Hörmander. The spectral function of an elliptic operator. *Acta Math.*, 121:193–218, 1968.

[HSY92] Brian R. Hunt, Tim Sauer, and James A. Yorke. Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc. (N.S.)*, 27(2):217–238, 1992.

[Hux03] M. N. Huxley. Exponential sums and lattice points. III. *Proc. London Math. Soc. (3)*, 87(3):591–609, 2003.

[Ivr80] V. Ja. Ivriĭ. The second term of the spectral asymptotics for a Laplace-Beltrami operator on manifolds with boundary. *Funktsional. Anal. i Prilozhen.*, 14(2):25–34, 1980.

[Ivr16] Victor Ivrii. 100 years of Weyl's law. *Bull. Math. Sci.*, 6(3):379–452, 2016.

[IW21] A. Iosevich and E. Wyman. Weyl law improvement for products of spheres. *Analysis Mathematica*, 47(3):593–612, 2021.

[Kal97] V. Yu. Kaloshin. Prevalence in spaces of finitely smooth mappings. *Funktsional. Anal. i Prilozhen.*, 31(2):27–33, 95, 1997.

[Kal00] Vadim Yu. Kaloshin. Generic diffeomorphisms with superexponential growth of number of periodic orbits. *Comm. Math. Phys.*, 211(1):253–271, 2000.

[KH07] Vadim Yu. Kaloshin and Brian R. Hunt. Stretched exponential estimates on growth of the number of periodic points for prevalent diffeomorphisms. I. *Ann. of Math. (2)*, 165(1):89–170, 2007.

[Kli78] Wilhelm Klingenberg. *Lectures on closed geodesics*. Grundlehren der Mathematischen Wissenschaften, Vol. 230. Springer-Verlag, Berlin-New York, 1978.

[Kni98] Gerhard Knieper. The uniqueness of the measure of maximal entropy for geodesic flows on rank 1 manifolds. *Ann. of Math. (2)*, 148(1):291–314, 1998.

[KS06] Vadim Kaloshin and Maria Saprykina. Generic 3-dimensional volume-preserving diffeomorphisms with superexponential growth of number of periodic orbits. *Discrete Contin. Dyn. Syst.*, 15(2):611–640, 2006.

[KT72]   Wilhelm Klingenberg and Floris Takens. Generic properties of geodesic flows. *Math. Ann.*, 197:323–334, 1972.

[Lev52]  Boris M. Levitan. On the asymptotic behavior of the spectral function of a self-adjoint differential equation of the second order. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 16:325–352, 1952.

[Mar69]  G. A. Margulis. Certain applications of ergodic theory to the investigation of manifolds of negative curvature. *Funkcional. Anal. i Priložen.*, 3(4):89–90, 1969.

[Mil97]  John Milnor. Fubini foiled: Katok's paradoxical example in measure theory. *Math. Intelligencer*, 19(2):30–32, 1997.

[OCV11]  Kevin C. O'Meara, John Clark, and Charles I. Vinsonhaler. *Advanced topics in linear algebra*. Oxford University Press, Oxford, 2011. Weaving matrix problems through the Weyr form.

[Pat99]  Gabriel P. Paternain. *Geodesic flows*, volume 180 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999.

[PS87a]  Vesselin Petkov and Luchezar Stojanov. Periods of multiple reflecting geodesics and inverse spectral results. *Amer. J. Math.*, 109(4):619–668, 1987.

[PS87b]  Vesselin Petkov and Luchezar Stojanov. Spectrum of the Poincaré map for periodic reflecting rays in generic domains. *Math. Z.*, 194(4):505–518, 1987.

[Sto87]  Luchezar Stojanov. Generic properties of periodic reflecting rays. *Ergodic Theory Dynam. Systems*, 7(4):597–609, 1987.

[SZ02]   Christopher D. Sogge and Steve Zelditch. Riemannian manifolds with maximal eigenfunction growth. *Duke Math. J.*, 114(3):387–437, 2002.

[Vol90a] A. V. Volovoy. Improved two-term asymptotics for the eigenvalue distribution function of an elliptic operator on a compact manifold. *Comm. Partial Differential Equations*, 15(11):1509–1563, 1990.

[Vol90b] A. V. Volovoy. Verification of the Hamilton flow conditions associated with Weyl's conjecture. *Ann. Global Anal. Geom.*, 8(2):127–136, 1990.

[Wey12]  Hermann Weyl. Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). *Math. Ann.*, 71(4):441–479, 1912.

[YC04]   Yosef Yomdin and Georges Comte. *Tame geometry with application in smooth analysis*, volume 1834 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2004.

[Yom85]  Y. Yomdin. A quantitative version of the Kupka-Smale theorem. *Ergodic Theory Dynam. Systems*, 5(3):449–472, 1985.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL, CHAPEL HILL, NC, USA

*Email address*: `canzani@email.unc.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY COLLEGE LONDON, LONDON, UK

*Email address*: `j.galkowski@ucl.ac.uk`