1. Introduction

In 1999 Alan Weinstein [We] presented a procedure to average a family \( \{ N_g \} \) of submanifolds of a Riemannian manifold \( M \); if the submanifolds are close to each other in a \( C^1 \) sense, one can produce a canonical \( \text{average} N \) which is close to each member of the family \( \{ N_g \} \). The main property of this averaging procedure is that it is equivariant

---

1. The construction is canonical because it does not involve any arbitrary choice but uses only the Riemannian metric on \( M \).
with respect to isometries of $M$, and therefore if the family $\{N_g\}$ is obtained by applying the isometric action of a compact group $G$ to some submanifold $N_0$ of $M$, the resulting average will be invariant under the $G$-action. This generalizes results about fixed points of group actions [We].

In the first part of this paper we will exhibit a result by Hermann Karcher and the author which improves Weinstein’s theorem.

In the main body of the paper we specialize Weinstein’s averaging to the setting of symplectic geometry: given a family $\{N_g\}$ of isotropic submanifolds of a symplectic manifold $M$, we obtain an isotropic average $L$. We achieve this in two steps: first we introduce a compatible Riemannian metric on $M$ and apply Weinstein’s averaging to obtain a submanifold $N$. This submanifold will be “nearly isotropic” because it is $C^1$-close to isotropic ones, and using the family $\{N_g\}$ we will deform $N$ to an isotropic submanifold $L$. Our construction depends only on the symplectic structure of $M$ and on the choice of compatible metric. Therefore applying our construction to the case of compact group actions by isometric symplectomorphisms we can obtain isotropic submanifolds which are invariant under the action.

As a simple application we show that the image of an almost invariant isotropic submanifold under a compact Hamiltonian action is “small”.

Another application is the following: given a symplectic action of a compact group $G$ on two symplectic manifolds $M_1$ and $M_2$ together with an almost equivariant symplectomorphism $\phi : M_1 \to M_2$, apply the averaging procedure to $\text{graph}(\phi) \subset M_1 \times M_2$. If the resulting $G$-invariant submanifold $L$ is a graph, then it will be the graph of a $G$-equivariant symplectomorphism. This means that we would be able to deform almost equivariant symplectomorphisms to equivariant ones. To ensure that $L$ is again a graph one needs to improve Weinstein’s averaging procedure, this is the subject of work in progress.

We would like to extend our averaging procedure to coisotropic submanifolds too: indeed, if one could average any two coisotropic submanifolds $N_0$ and $N_1$ which are close to each other, then by “shifting weights” in the parameter space $G = \{0, 1\}$ one would produce a continuous path of coisotropic submanifolds connecting $N_0$ to $N_1$. This would show that the space of coisotropic submanifolds is locally path connected.

In the remainder of the introduction we will recall the averaging procedure in the Riemannian setting by Weinstein (see [We]), we will state our results, and we will outline our construction of averaging isotropic submanifolds.

1.1. Averaging of Riemannian submanifolds

The starting point for our isotropic averaging construction is the statement of Theorem 2.3 in [We]. We first recall some definitions from [We] in order to state the theorem.

---

2 It would be interesting to find a way to deform any given “nearly isotropic” submanifold to an honest isotropic one in a canonical fashion.

3 We need to improve Weinstein’s theorem in order to ensure that $\text{graph}(\phi)$ and $L$ be $C^1$-close; see Remark 1 in Section 8.
If $M$ is a Riemannian manifold and $N$ a submanifold, $(M, N)$ is called a gentle pair if (i) the normal injectivity radius of $N$ is at least 1; (ii) the sectional curvatures of $M$ in the tubular neighborhood of radius 1 about $N$ are bounded in absolute value by 1; (iii) the injectivity radius at each point of the above neighborhood is at least 1.

The distance between two subspaces $F, F'$ of the same dimension of a Euclidean vector space $E$, denoted by $d(F, F')$, is equal to the $C^0$-distance between the unit spheres of $F$ and $F'$ considered as Riemannian submanifolds of the unit sphere of $E$. This distance is symmetric and satisfies $d(F, F') = d(F^\perp, F'^\perp)$. It is always smaller than or equal to $\pi/2$, and it is equal to $\pi/2$ iff $F$ and $F'^\perp$ are not transversal.

One can define a $C^1$-distance between two submanifolds $N, N'$ of a Riemannian manifold if $N'$ lies in the tubular neighborhood of $N$ and is the image under the normal exponential map of $N$ of a section of $\nu N$ (so $N$ and $N'$ are necessarily diffeomorphic). This is done by assigning two numbers to each $x' \in N'$: the length of the geodesic segment from $x'$ to the nearest point $x$ in $N$ and the distance between $T_{x'} N'$ and the parallel translate of $T_x N$ along the above geodesic segment. The $C^1$-distance is defined as the supremum of those numbers as $x'$ ranges over $N'$ and is denoted by $d_1(N, N')$.

Note that this distance is not symmetric, but if $(M, N)$ and $(M, N')$ are both gentle pairs with $d_1(N, N') < 1/4$, then $d_1(N', N) < 250d_1(N, N')$ (see Remark 3.18 in [We]).

The improvement of Theorem 2.3 in [We] by Karcher and the author is our Theorem 4 and reads:

**Theorem (Weinstein).** Let $M$ be a Riemannian manifold and $\{N_g\}$ a family of submanifolds of $M$ parametrized in a measurable way by elements of a probability space $G$, such that all the pairs $(M, N_g)$ are gentle. If $d_1(N_g, N_h) < \epsilon < 1/20000$ for all $g$ and $h$ in $G$, there is a well defined center of mass submanifold $N$ with $d_1(N_g, N) < 2500\epsilon$ for all $g$ in $G$. The center of mass construction is equivariant with respect to isometries of $M$ and measure preserving automorphisms of $G$.

**Remark.** For any $g \in G$ the center of mass $N$ is the image under the exponential map of a section of $\nu N_g$ and $d_0(N_g, N) < 100\epsilon$.

From this one gets immediately a statement about invariant submanifolds under compact group actions (cf. Theorem 2.2 of [We]).

### 1.2. Averaging of isotropic submanifolds

Recall that for any symplectic manifold $(M, \omega)$ we can choose a compatible Riemannian metric $g$, i.e. a metric such that the endomorphism $I$ of $TM$ determined by $\omega(\cdot, I\cdot) = g(\cdot, \cdot)$ satisfies $I^2 = -\text{Id}_M$. The tuple $(M, g, \omega, I)$ is called an almost-Kähler manifold. To prove our Main Theorem we need to assume a bound on the $C^0$-norm of $\nabla \omega$ (here $\nabla$ is the Levi-Civita connection given by $g$), which measures how far our almost-Kähler

---

4 We omit the compactness assumption on the $N_g$'s stated there since it is superfluous.
manifold is from being Kähler. We state the theorem choosing the bound to be 1 (but see Remark (i) below).

**Theorem 1** *(Main Theorem).* Let \((M^m, g, \omega, I)\) be an almost-Kähler manifold satisfying \(|\nabla \omega| < 1\) and \(\{N^g_n\}\) a family of isotropic submanifolds of \(M\) parametrized in a measurable way by elements of a probability space \(G\), such that all the pairs \((M, N^g_*)\) are gentle. If \(d_1(N^g_*, N^h_*) < \epsilon < 1/70000\) for all \(g\) and \(h\) in \(G\), there is a well defined isotropic center of mass submanifold \(L^0\) with \(d_0(N^g_*, L) < 1000\epsilon\) for all \(g\) in \(G\). This construction is equivariant with respect to isometric symplectomorphisms of \(M\) and measure preserving automorphisms of \(G\).

**Remark.**

(i) The theorem still holds if we assume higher bounds on \(|\nabla \omega|\), but in this case the bound 1/70000 for \(\epsilon\) would have to be chosen smaller. See the remark in Section 7.4.

(ii) Notice that we are not longer able to give estimates on the \(C^1\)-distance of the isotropic center of mass from the \(N^g_*\)‘s. Such an estimate could possibly be given provided we have more information about the extrinsic geometry of Weinstein’s center of mass submanifold; see Remark 1 in Section 8. Instead we can only give estimates on the \(C^0\)-distances \(d_0(N^g_*, L) = \sup \{d(x, N^g_*) : x \in L\}\).

An easy consequence of our Main Theorem is a statement about group actions. Recall that, given any action of a compact Lie group \(G\) on a symplectic manifold \((M, \omega)\) by symplectomorphisms, by averaging over the compact group one can always find some invariant metric \(\tilde{g}\). Using \(\omega\) and \(\tilde{g}\) one can canonically construct a metric \(g\) which is compatible with \(\omega\) (see [Ca]), and since \(g\) is constructed canonically out of objects that are \(G\)-invariant, it will be \(G\)-invariant too. Therefore the group \(G\) acts respecting the structure of the almost-Kähler manifold \((M, g, \omega)\). In general it does not seem possible to give any bound on \(|\nabla \omega|\), where \(\nabla\) is the Levi-Civita connection corresponding to \(g\).

**Theorem 2.** Let \((M, g, \omega, I)\) be an almost-Kähler manifold satisfying \(|\nabla \omega| < 1\) and let \(G\) be a compact Lie group acting on \(M\) by isometric symplectomorphisms. Let \(N^0_0\) be an isotropic submanifold of \(M\) such that \((M, N^0_0)\) is a gentle pair and \(d_1(N^0_0, gN^0_0) < \epsilon < 1/70000\) for all \(g \in G\). Then there is a \(G\)-invariant isotropic submanifold \(L\) with \(d_0(N^0_0, L) < 1000\epsilon\).

The invariant isotropic submanifold \(L\) as above is constructed by endowing \(G\) with the bi-invariant probability measure and applying Theorem 1 to the family \(\{gN^0_0\}_{g \in G}\). The resulting isotropic average \(L\) is \(G\)-invariant because of the equivariance properties of the averaging procedure.

### 1.3. Outline of the proof of the Main Theorem

This is the main subsection of this paper. We will try to convince the reader that the construction we use to prove Theorem 1 works if only one chooses \(\epsilon\) small enough. Let us begin by requiring \(\epsilon < 1/20000\).

---

5 Recall that an almost-Kähler manifold is Kähler if the almost complex structure \(I\) is integrable, or equivalently if \(\nabla I = 0\) or \(\nabla \omega = 0\).
Part I. We start by considering the average of the submanifolds $N_g$ as in Theorem 2.3 of [We], which we will denote $N$. We will use the notation $\exp_N$ to indicate the restriction of the exponential map to $TM\vert_N$, and similarly for any of the $N_g$’s. For any $g$ in $G$, the average $N$ lies in a tubular neighborhood of $N_g$ and is the image under $\exp_{N_g}$ of a section $\sigma$ of $\nu N_g$ (see [We]). Therefore for any point $p$ of $N_g$ there is a canonical path $\gamma_q(t) = \exp_p(t \cdot \sigma(p))$ from $p$ to the unique point $q$ of $N$ lying in the normal slice of $N_g$ through $p$. Here, writing $(\nu N_g)^1$ for the open unit disk bundle in $\nu N_g$, we use the term “normal slice” for the submanifold $\exp_{N_g}(\nu p N_g)^1$. We define the following map:

$$\phi_g : \exp_{N_g}(\nu p N_g)^1 \to M, \quad \exp_p(v) \mapsto \exp_q(\gamma_q(v)),$$

Here $p$, $q$, and $\gamma_q$ are as above, $v \in (\nu p N_g)^1$, and $\gamma_q$ denotes parallel translation along $\gamma_q$. So $\phi_g$ takes the normal slice $\exp_p(\nu p N_g)^1$ to $\exp_q(\nu N_g)^1$, where $\nu N_g \subset T_q M$ is the parallel translation along $\gamma_q$ of $\nu p N_g \subset T_p M$.

We have $d(\nu N_g, \nu q N) < d_1(N_g, N) < 2500\epsilon < \pi/2$, so $\nu q N$ is transversal. Therefore $\phi_g$ is a local diffeomorphism at all points of $N_g$, and it is clearly injective there. Using the geometry of $N_g$, $N$ and $M$, in Proposition 6.1 we will show that $\phi_g$ is a diffeomorphism onto if restricted to the tubular neighborhood $\exp_{N_g}(\nu N_g)^{0.05}$ of $N_g$.

We restrict our map to this neighborhood and we also restrict the target space so as to obtain a diffeomorphism, which we will still denote by $\phi_g$.

Part II. Now we introduce the symplectic form $\omega_g := (\phi_g^{-1})^* \omega$ on $\exp_{N_g}(\nu N_g)^{0.05}$. Notice that $N$ is isotropic with respect to $\omega_g$ by construction, hence also with respect to the 2-form $\int_g \omega_g$, which is defined on $\bigcap_{g \in G} \exp_N(\nu N_g)^{0.05}$. We would like to apply Moser’s trick $\phi$ (see [Ca, Chapter III]) to $\omega$ and $\int_g \omega_g$. To do so we

---

\[\text{Recall that Moser’s Theorem states the following: if } \Omega_t \ (t \in [0, 1]) \text{ is a smooth family of symplectic forms lying in the same cohomology class in a compact manifold then there is a family of diffeomorphisms } \rho_t \text{ with } \rho_0 = \text{Id satisfying } \rho_t^* \Omega_t = \Omega_0.\]
first restrict our forms to a smaller tubular neighborhood $tub^\epsilon$ of $N$, which we define in Section 7.1. To apply Moser’s trick we have to check:

1. On $tub^\epsilon$ the convex linear combination $\omega_t = \omega + t(\int g \omega g - \omega)$ is a symplectic form for each $t \in [0, 1]$

Indeed, we will show that on $tub^\epsilon$ the differential of $\psi^{-1}_g$ is “close” to the parallel translation $\nabla$ along certain “canonical” geodesics that will be specified at the beginning of Section 3. This and the bound on $|\nabla \omega|$ imply that for any $q \in tub^\epsilon$ and nonzero $X, Y \in T_q M$,

$$(\omega\phi^1_q(X, Y) = \omega_{\psi^{-1}_g(q)}((\psi^{-1}_g)_* X, (\psi^{-1}_g)_* Y) \approx \omega_{\phi^{-1}_g(q)}(\nabla X, \nabla Y) \approx \omega_q(X, Y),$$

i.e. $\omega_g$ and $\omega$ are very close to each other. So $\omega_t(X, IX) \approx \omega(X, IX) = |X|^2 > 0$. Therefore each $\omega_t$ is nondegenerate, and clearly it is also closed.

2. On $tub^\epsilon$ the forms $\omega$ and $\int g \omega_g$ belong to the same cohomology class.

Fix $g \in G$. The inclusion $i : tub^\epsilon \hookrightarrow \exp N_g(vN_g)_1$ is homotopic to $\psi^{-1}_g : tub^\epsilon \to \exp N_g(vN_g)_1$. A homotopy is given by thinking of $N$ as a section of $vN_g$ and “sliding along the fibers” to the zero section. Therefore these two maps induce the same map in cohomology, and pulling back $\omega$ we have

$$[\omega]_{tub^\epsilon} = i^*[\omega] = (\psi^{-1}_g)^*[\omega] = [\omega_g].$$

Integrating over $G$ finishes the argument.

Now we can apply Moser’s trick: if $\alpha$ is a 1-form on $tub^\epsilon$ such that $d\alpha$ is equal to $\frac{d}{dt} \omega_t = \int g \omega_g - \omega$, then the flow $\rho_t$ of the time-dependent vector field $v_t := -\psi^{-1}_g(\alpha)$ has the property $\rho_t^* \omega_t = \omega$ (and in particular $\rho_t^* (\int g \omega_g) = \omega$) where it is defined. Therefore if $L := \rho_t^{-1}(N)$ is a well defined submanifold of $tub^\epsilon$, then it will be isotropic with respect to $\omega$ since $N$ is isotropic with respect to $\psi^{-1}_g$.

We will construct canonically a primitive $\alpha$ as above in Section 7.2. Using the fact that the distance between the $N_g$’s and $N$ is small, we will show that $\alpha$ has small maximum norm. So, if $\epsilon$ is small enough, the time-1 flow of the time-dependent vector field $\{-v_t\}_{t=1}$ will not take $N$ out of $tub^\epsilon$ and $L$ will be well defined.

Since our construction is canonical after fixing the almost-Kähler structure $(g, \omega, I)$ of $M$ and the probability space $G$, the construction of $L$ is equivariant with respect to isometric symplectomorphisms of $M$ and measure preserving automorphisms of $G$.

1.4. Structure of the paper and acknowledgements

This paper is organized as follows: In Section 2 we present the improvement of Theorem 2.3 of [We] obtained by Karcher and the author. In Section 3 we will start the proof of the Main Theorem by studying the map $\psi_g$. In Section 4 we will state a proposition
about geodesic triangles, and in Section 5 we will apply it in our setup. This will allow us to show in Section 6 that each $\psi_g$ is injective on $\exp_{N_g}(v)N_g)_{0,05}$. The proofs of some estimates of Sections 4 and 5 are rather involved, and we present them in the three appendices. This will conclude the proof of the first part of the theorem.

In Section 7 we will make use for the first time of the symplectic structure of $M$. We will show that the $\omega_t$‘s are symplectic forms and that the 1-form $\alpha$, and therefore the Moser vector field $v_t$, are small in the maximum norm. Comparison with the results of Section 6 will end the proof of the Main Theorem.

Section 8 will be devoted to remarks about the Main Theorem, and in Section 9 we will present a simple application to Hamiltonian group actions.

At this point I would like to thank everyone who helped me and supported me in the preparation of this paper. In particular I would like to thank Alan Weinstein for helpful discussions during the preparation of this paper, the referee for his careful review of the manuscript, his interest and for suggesting improvements, Yael Karshon for proposing the application in Section 9 and River Chiang for simplifying the arguments used there. Further I thank Hermann Karcher for sharing the ideas involved in Section 2 and for the collaboration.

2. Improved error estimates for the shape operators of parallel tubes with application to Weinstein’s submanifold averaging

In this section we will present the improvement of Theorem 2.3 of [We] obtained by Hermann Karcher and the author. In the first subsection we will improve Proposition 3.11 of [We]. Then using this result we will follow Weinstein’s proof and present the statement of the improved theorem.

2.1. Estimates for the shape operators of parallel tubes

In Proposition 3.11 of [We] one has the setup we are going to describe now. $M$ is a Riemannian manifold, $N$ is a submanifold of $M$ such that $(M, N)$ form a gentle pair (so the second fundamental form $B$ of $N$ satisfies $|B| \leq 3/2$, see [We, Cor. 3.2]). In the tubular neighborhood of radius 1 about $N$ let $\rho_N$ be the distance function from $N$, and $P_N = \frac{1}{2} \rho_N^2$. We are interested in estimating the Hessian of $P_N$, i.e. the symmetric endomorphism of each tangent space of the tubular neighborhood given by $H_N(v) = \nabla_v \text{grad } P_N$. Differentiating the relation $\text{grad } P_N = \rho_N \cdot \text{grad } \rho_N$ we see that

$$H_N(v) = \langle U_N, v \rangle U_N + \rho_N \cdot S_N(\text{pr}(v))$$

where $U_N = \text{grad } \rho_N$ is the radial unit vector (pointing away from $N$), $\text{pr}$ denotes orthogonal projection onto $U_N^\perp$, and $S_N$ is the second fundamental form of the tube given by a level set $\tau(t)$ of $\rho_N$ in the direction of the normal vector $U_N$.

---

So $S_N v = \text{pr}(\nabla_v U_N)$ for all vectors $v$ tangent to $\tau(t)$, where $\nabla$ is the Levi-Civita connection on $M$. 

---

7
Proposition 3.11 of \[\text{We}\] states that, at a point \(p\) of distance \(t \leq 1/4\) from \(N\), the following estimate holds for the decompositions into vertical and horizontal parts\(\text{[N]}\) of \(T_p M\):

\[
[0.64 \cdot I \ 0 \ 0 \ -3t \cdot I] < H_N < [1.32 \cdot I \ 0 \ 0 \ 3t \cdot I],
\]

where for two symmetric matrices \(P\) and \(Q\) the inequality \(P < Q\) means that \(Q - P\) is positive definite.

The above proposition is proved using the Riccati equation. An immediate consequence is Corollary 3.13 in \[\text{We}\], which states that, if \(v\) is a horizontal vector and \(w\) a vertical vector at \(p\), then

\[
|\langle H_N(v), w \rangle| \leq 3 \sqrt{t} |v| |w|.
\]

This square root is responsible for the presence of upper bounds proportional to \(\sqrt{\epsilon}\) rather than \(\epsilon\) in Theorems 2.2 and 2.3 of \[\text{We}\].

We will improve the estimate of Corollary 3.13 of \[\text{We}\], determining \(S_N\) by means of Jacobi field estimates rather than by the Riccati equation. More precisely, we will make use of this simple observation:

**Lemma 2.1.** Let \(N\) be a submanifold of the Riemannian manifold \(M\), and fix \(t \leq \text{normal injectivity radius of } N\). Let \(p\) lie in the tube \(\tau(t) = \rho - 1\) and \(S_N : T_p \tau(t) \to T_p \tau(t)\) be the second fundamental form in the direction of \(U_N\). For any \(v \in T_p \tau(t)\) consider the Jacobi field \(\tilde{J}(r)\) arising from the variation \(r \mapsto \exp(c(s)) r U_N(c(s))\), where \(c(s)\) is any curve in \(\tau(t)\) tangent to \(v\).

Then

\[
S_N v = \tilde{J}'(0).
\]

**Proof.** Denoting by \(f(s, r)\) the above variation and by \(\nabla\) the Levi-Civita connection on \(M\) we have

\[
\tilde{J}'(0) = \frac{\nabla}{dr} \left|_0 \frac{d}{ds} \right|_0 f(s, r) = \frac{\nabla}{ds} \left|_0 \frac{d}{dr} \right|_0 f(s, r) = \frac{\nabla}{ds} U_N(c(s)) = \nabla_v U_N = \text{pr}(\nabla_v U_N) = S_N v.
\]

Using the above lemma we will be able to prove this improvement of Proposition 3.11 of \[\text{We}\], for which we do not require \((M, N)\) to be a gentle pair but only a bound on \(|B|\) and the curvature assumption \(|K| \leq 1\):

**Theorem 3.** Let \(N\) be a submanifold of the Riemannian manifold \(M\) with second fundamental form \(B\), and fix \(t \leq \text{normal injectivity radius of } N\). Let \(\gamma\) be a unit-speed geodesic emanating normally from \(N\). Assume \(|K| \leq 1\) in the radius \(t\) tubular neighborhood of \(N\). Let \(\tau(t)\) be the \(t\)-tube about \(N\), and let \(S_N(t)\) denote the second fundamental form of \(\tau(t)\) in direction \(\hat{\gamma}(t)\) at \(\gamma(t)\). Then with respect to the splitting into vertical and horizontal spaces of \(T_{\gamma(t)} \tau(t)\), as long as \(t \leq \min\{1/2, 1/2|B|\}\), we have

\[
t \cdot S_N(t) \leq \begin{pmatrix} I & 0 \\ 0 & tB \end{pmatrix} + \begin{pmatrix} 16t^2 & 16t^2 \\ 16t^2 & (22 + 2|B|^2)t^2 \end{pmatrix}.
\]

See our Section 3 or Section 2.1 in \[\text{We}\] for the definition of vertical and horizontal bundle at \(p\).
Remark. We adopt the following unconventional notation: If $M, \bar{M}$ are matrices and $c$ a real number, $M \leq \bar{M} + c$ means that $M - \bar{M}$ has operator norm $\leq c$. Generalizing to the case where we consider also vertical-horizontal decompositions of matrices,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq \begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{bmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

means that the above convention holds for each endomorphism between horizontal/vertical spaces, i.e. $A - \bar{A}$ has operator norm $\leq a$ and so on.

Proof of Theorem 3. Choose an orthonormal basis $\{E_1, \ldots, E_{n-1}\}$ of $\hat{\gamma}(0)^{\perp} \subset T_{\gamma(0)}M$ such that $E_1, \ldots, E_k$ lie in the normal space to $N$ and $E_{k+1}, \ldots, E_{n-1}$ lie in the tangent space to $N$. (Here $\dim(M) = n$.) Now we define Jacobi fields $J_i$ along $\gamma$ with the following initial conditions:

$$\begin{cases} J_i(0) = 0, & J_i'(0) = E_i, \quad \text{if } i \leq k \text{ (vertical Jacobi fields)}, \\ J_i(0) = E_i, & J_i'(0) = B_{\gamma(0)}E_i, \quad \text{if } i \geq k + 1 \text{ (horizontal Jacobi fields)}. \end{cases}$$

Notice that, among all $N$-Jacobi fields (see Section 3 for their definition) satisfying $J_i(0) = E_i$, our $J_i$ are those having smallest derivative at time zero. Also notice that all $J_i$ and their derivatives are perpendicular to $\hat{\gamma}(0)$, therefore, as long as the $J_i(t)$ are linearly independent, they form a basis of $\hat{\gamma}(t)^{\perp} = T_{\gamma(t)}\tau(t)$. Also, the $J_i$’s are $N$-Jacobi fields, i.e. Jacobi fields for which $J_i(0)$ is tangent to $N$ and $J_i'(0) - B_{\gamma(0)}J_i(0)$ is normal to $N$, or equivalently Jacobi fields that arise from variations of geodesics emanating normally from $N$ (see [Wa, p. 342]). Moreover the $J_i$’s are a basis of the space of $N$-Jacobi fields along $\gamma$ which are orthogonal to $\hat{\gamma}$, and this space coincides with the space of $N$-Jacobi fields arising from a variation of unit-speed $\gamma$-geodesics normal to $N$. The velocity vectors of such variations at time $t$ coincide with $\tau(t)N$. Therefore applying Lemma 2.1 with $v = J_i(t)$ we conclude that $S_N(t)J_i(t) = J_i'(t)$ for all $i$.

Now consider the maps

$$J(t) : \mathbb{R}^{n-1} \rightarrow T_{\gamma(t)}\tau(t), \quad e_i \mapsto J_i(t),$$

and

$$J'(t) : \mathbb{R}^{n-1} \rightarrow T_{\gamma(t)}\tau(t), \quad e_i \mapsto J_i'(t),$$

where $\{e_i\}$ is the standard basis of $\mathbb{R}^{n-1}$. As long as the $J_i(t)$’s are linearly independent, we clearly have

$$S_N(t) = J'(t) \cdot J(t)^{-1}.$$

Propagating the $E_i$’s along $\gamma$ by parallel translation we obtain an orthonormal basis $\{E_i(t)\}$ of $T_{\gamma(t)}\tau(t)$. Furthermore, $\{E_1(t), \ldots, E_k(t)\}$ together with $\hat{\gamma}(t)$ span the vertical space at $\gamma(t)$ and $\{E_{k+1}(t), \ldots, E_{n-1}(t)\}$ span the horizontal space there. We will

---

9 Indeed, connecting the points of an integral curve in $\tau(t)$ of some $J_i(t)$ to $N$ by unit-speed shortest geodesics we obtain such a variation, and the Jacobi field arising from this variation must be $J_i$ since it is an $N$-Jacobi field orthogonal to $\hat{\gamma}$ which coincides with $J_i(t)$ at time $t$. 

represent the maps $J(t)$, $J'(t)$ and $S_N(t)$ by matrices with respect to the bases $\{e_i\}$ for $\mathbb{R}^{n-1}$ and $\{E_i(t)\}$ for $T_{\gamma(t)}\mathbb{R}$. 

Now we use Jacobi field estimates as in [BK 6.3.8iii] to determine the operator norm of $J(t)$, or rather of the endomorphisms $J(t)_{V V}$, $J(t)_{V H}$ and $J(t)_{H H}$ that $J(t)$ induces on horizontal and vertical subspaces.\(^{10}\) This will allow us to obtain corresponding estimates for $J^{-1}(t)$ and $J'(t)$, and therefore for $S_N(t)$.

For all $i$ let us define the vector fields $A_i(t) = \langle J_i(0) + t \cdot J_i'(0) \rangle$, where $\langle \rangle$ denotes parallel translation along $\gamma$. The map $\mathbb{R}^{n-1} \to T_{\gamma(t)}\mathbb{R}$, $e_i \mapsto A_i(t)$, in matrix form reads

$$A(t) = \begin{bmatrix} tI & 0 \\ 0 & I + tB \end{bmatrix}.$$ 

For $i \leq k$ we have $J_i(0) = 0$ and $|J_i'(0)|$ is an orthonormal set. If $(c_1, \ldots, c_k, 0, \ldots, 0)$ is a unit vector in $\mathbb{R}^{n-1}$, we have $|(\sum c_i J_i)'(0)| = 1$, so applying [BK 6.3.8iii] we obtain $|\sum c_i(J_i(t) - A_i(t))| \leq \sinh(t) - t$.

Similarly, for $i \geq k + 1$, the set $\{J_i(0)\}$ is an orthonormal set and $J_i'(0) = B(J_i(0))$.

Again, if $(0, \ldots, 0, c_{k+1}, \ldots, c_{n-1})$ is a unit vector in $\mathbb{R}^{n-1}$, since $|(\sum c_i J_i)'(0)| = |B(\sum c_i J_i(0))| \leq |B|$, we have $|\sum c_i(J_i(t) - A_i(t))| \leq \cosh(t) - 1 + |B|\sinh(t) - t$.

Therefore we have

$$J(t) - A(t) =: F_1(t) \leq \begin{bmatrix} \sinh(t) - t & \cosh(t) - 1 + |B|\sinh(t) - t \\ \sinh(t) - t & \cosh(t) - 1 + |B|\sinh(t) - t \end{bmatrix} \leq \begin{bmatrix} \frac{3}{2} t^2 & \frac{3}{2} t^2 \\ \frac{3}{2} t^2 & \frac{3}{2} t^2 \end{bmatrix}.$$ 

Now we want to estimate $t J^{-1}(t)$. Notice that, suppressing the $t$-dependence in the notation, we have $J = A \cdot [I + A^{-1} F_1]$, so that

$$t J^{-1} = [I + A^{-1} F_1]^{-1} \cdot t A^{-1}.$$ 

Clearly $A$ is invertible and

$$t A^{-1} = \begin{bmatrix} I & 0 \\ 0 & t \cdot (I + t B)^{-1} \end{bmatrix} \leq \begin{bmatrix} 1 & 0 \\ 0 & 2t \end{bmatrix}$$

since we assume $t \leq 1/2|B|$. We have

$$A^{-1} F_1 \leq \begin{bmatrix} \frac{3}{2} t^2 & \frac{3}{2} t^2 \\ \frac{3}{2} t^2 & \frac{3}{2} t^2 \end{bmatrix}.$$ 

Clearly its norm is less than $\sqrt{\frac{3}{2} t^2 + 4 t^2} \leq \frac{3}{2} t < 1$ since $t \leq \frac{1}{2}$. Therefore $I + A^{-1} F_1$ is invertible and $|I + A^{-1} F_1|^{-1} = \sum_{j=0}^{\infty} (-A^{-1} F_1)^j$. Using the above estimate

\(^{10}\) To be more precise: $J(t)_{H V} : \mathbb{R}^k \times \{0\} \to \text{Hor}(t)$ is given by restricting $J(t)$ and then composing with the orthogonal projection onto the horizontal space at $\gamma(t)$.

\(^{11}\) If $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then the full operator norm of the matrix is bounded by $\sqrt{\max\{||a||, ||b||\} + ab + cd} \leq \sqrt{2} \max\{||a||, ||b||\}$.
for $A^{-1}F_1$ we have 

$$[A^{-1}F_1]^2 \leq \left( \frac{1}{2}t^4 + \frac{3t^3}{2} \right).$$

Using the coarse estimate $A^{-1}F_1 \leq \frac{3}{2}t$ and $t \leq \frac{1}{2}$ we have $\sum_{j=3}^{\infty} [A^{-1}F_1]^j \leq 14t^3$.

Putting together these estimates we obtain 

$$[I + A^{-1}F_1]^{-1} = I + F_2 \quad \text{where} \quad F_2 \leq \left( \frac{15t^2}{15t^3} + \frac{5t}{2t^2} \right).$$

To estimate $J'(t)$ we first estimate $|J''(t) - A''(t)|$ and then integrate. For all $i$ we have 

$$|J''_i(t) - A''_i(t)| = |J'_i(t)| \leq |J_i(t)|$$

by the Jacobi equation using the bound on curvature, and an analogous estimate holds for linear combinations $\sum c_i J_i(t)$.

If $(c_1, \ldots, c_k, 0, \ldots, 0)$ is a unit vector in $\mathbb{R}^{n-1}$ we have $|\sum c_i J_i(t)| \leq \sinh(t)$ by Rauch’s theorem.

Similarly, if $(0, \ldots, 0, c_{k+1}, \ldots, c_{n-1})$ is a unit vector in $\mathbb{R}^{n-1}$ we have $|\sum c_i J_i(0)| = 1$ and $|\sum c_i J'_i(0)| \leq |B|$, so by Berger’s extension of Rauch’s theorem (see Lemma 2.7.9 in [Kl]) we have $|\sum c_i J_i(t)| \leq \cosh(t) + |B| \sinh(t)$.

In both cases integration yields 

$$\left| \sum c_i (J'_i(t) - A'_i(t)) \right| \leq \int_0^t \left| \sum c_i (J''_i(\tau) - A''_i(\tau)) \right| d\tau$$

$$\leq \begin{cases} \cosh(t) - 1 \leq \frac{3}{4}t^2 & \text{if } i \leq k, \\ \sinh(t) + |B|(|\cosh(t) - 1) \leq \frac{3}{2}t & \text{if } i \geq k + 1. \end{cases}$$

So altogether we obtain 

$$J'(t) - A'(t) =: F_3(t) \quad \text{where} \quad F_3(t) \leq \left( \frac{3t^2}{4t^2} + \frac{3}{2t} \right).$$

Now finally we can estimate 

$$t S_N(t) = t JJ^{-1} = (A' + F_3) \cdot (I + F_2) \cdot t A^{-1}$$

$$\leq \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix} + \begin{pmatrix} \frac{1}{3}t^2 & \frac{3}{2}t \\ \frac{3}{2t} & \frac{3}{2}t \end{pmatrix} + \begin{pmatrix} \frac{15t^2}{15t^3} & 5t \frac{19}{3B|t|^2} \\ 15B|t|^3 & \frac{19}{2B|t|^2} \end{pmatrix} + \begin{pmatrix} \frac{3t^2}{2} & 3t^2 \\ \frac{3}{2}t^2 & 3t^2 \end{pmatrix} + \begin{pmatrix} \frac{15t^2}{2} & 10t^2 \\ 15B|t|^3 & 19B|t|^3 \end{pmatrix}$$

$$+ \begin{pmatrix} \frac{30t^4}{30t^4} & 36t^4 \\ \frac{36t^4}{36t^4} & \frac{36t^4}{36t^4} \end{pmatrix}.$$
Here we used
\[ tA^{-1} \leq \begin{bmatrix} 1 & 0 \\ 0 & tI \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 2|B|t^2 \end{bmatrix} \leq \begin{bmatrix} 1 & 0 \\ 0 & 2t \end{bmatrix} \]
in the last inequality. In view of our bounds on \( t \) and the fact that \( S_N(t) \) is a symmetric operator this gives the claimed estimate. \( \square \)

Returning to the case when \((M, N)\) is a gentle pair, so that \(|B| \leq 3/2\), we obtain our improvement of Corollary 3.13 in \cite{We}. Now we can achieve an upper bound proportional to \( t^2 \), versus the bound proportional to \( \sqrt{t} \) in Corollary 3.13 of \cite{We}.

**Corollary 2.1.** Let \( M \) be a Riemannian manifold, and \( N \) a submanifold so that \((M, N)\) form a gentle pair. If \( v \) is a horizontal vector and \( w \) a vertical vector at some point of distance \( t \leq 1/3 \) from \( N \), then \(|\langle H_N(v), w \rangle| \leq 16t^2||v||w||\).

2.2. Improvement of Weinstein’s averaging theorem

Now we use Corollary 2.1 to replace some estimates in \cite{We} that were originally derived using Corollary 3.13 there. We will improve only estimates contained in Lemmata 4.7 and 4.8 of \cite{We}, where the author considers the covariant derivative of a certain vector field \( \mathcal{V} \) on \( M \) in directions which are almost vertical or almost horizontal \(^{12}\) with respect to a fixed submanifold \( N_\epsilon \). (The zero set of \( \mathcal{V} \) is the average \( N_\epsilon \) of the family \( \{N_\epsilon \} \).) As in \cite{We} all estimates will hold for \( \epsilon < 1/20000 \), and we set \( t = 100\epsilon \).

We will replace the constant 89/200 in Lemma 4.7 of \cite{We} by 4/5 as follows:

**Lemma 2.2.** For any almost vertical vector \( v \) at any point of \( N \),
\[ \langle D_v \mathcal{V}, v \rangle \geq \frac{4}{5} ||v||^2. \]

*Proof.* By Theorem 3 (applied to the gentle pair \((M, N_\epsilon)\)) for the operator norm of \( H_\epsilon \) we have \( 1 - 16\epsilon^2 \leq ||H_\epsilon|| \), so that one obtains \( \langle H_\epsilon(\mathcal{P}_\mathcal{V} v, \mathcal{P}_\mathcal{V} v) \rangle > 19/20 \) in the proof of Lemma 4.7 in \cite{We}. Similarly, Theorem 3 together with footnote 11 implies that \( ||H_\epsilon|| < 1.01 \). Using these estimates in the proof of Lemma 4.7 in \cite{We} gives the claim. \( \square \)

Similarly, we will replace the term 60\( \sqrt{\epsilon} \) in Lemma 4.8 of \cite{We} by 1950\( \epsilon \).

**Lemma 2.3.** For any almost horizontal vector \( v \) at any point of \( N \),
\[ |D\mathcal{V}(v)| \geq 1950\epsilon ||v||. \]

*Proof.* By Corollary 2.1 we can replace 3\( \sqrt{\epsilon} \) by 16\( \epsilon^2 \) in the proof of Lemma 4.8 in \cite{We} and we can use 1.01 instead of 1.32 as an upper bound for \( ||H_\epsilon|| \). Furthermore, we replace the constant 1000 coming from Lemma 4.3 in \cite{We} by 525 \(^{13}\). This gives the improved estimate \( H_\epsilon(v, \mathcal{P}_\mathcal{T} w) \leq 850\epsilon ||v|| \cdot ||w|| \) and simple arithmetic concludes the proof. \( \square \)

\(^{12}\) See our Section 3 or Section 3.2 in \cite{W} for the definitions of almost horizontal and almost vertical bundle.

\(^{13}\) Lemma 4.3 of \cite{W} quotes incorrectly Proposition A.8 from its own appendix.
From these two lemmas it follows that the operator from \((a \text{Vert})^\perp\) to \(a \text{Vert}\) whose graph is \(T_x N\) has norm at most \(\frac{1}{3} 1950\epsilon\). Following to the end the proof of Theorem 2.3 in [We] allows us to replace the bound \(136\sqrt{\epsilon}\) by a bound linear in \(\epsilon\), so we obtain the following improved statement:

**Theorem 4.** Let \(M\) be a Riemannian manifold and \(\{N_g\}\) a family of submanifolds of \(M\) parametrized in a measurable way by elements of a probability space \(G\), such that all the pairs \((M, N_g)\) are gentle. If \(d_1(N_g, N_h) < \epsilon < 1/20000\) for all \(g\) and \(h\) in \(G\), there is a well defined center of mass submanifold \(N\) with \(d_1(N_g, N) < 2500\epsilon\) for all \(g\) in \(G\). The center of mass construction is equivariant with respect to isometries of \(M\) and measure preserving automorphisms of \(G\).

### 3. Estimates on the map \(\varphi_g\)

In Sections 3–7 we will prove the Main Theorem. The reader is referred to Section 1.3 for an outline of the proof and some of the notation introduced there. We will present Part I of the proof in Sections 3–6 and Part II in Section 7.

Fix \(g \in G\) and let \(p\) be a point in the tubular neighborhood of \(N_g\) and \(X \in T_p M\).

The aim of this section is to estimate the difference between \(\varphi_g^* X\) and \(\parallel X\). This will be achieved in Proposition 3.4.

Here we denote by \(\parallel X\) the following parallel translation of \(X\), where \(\pi_{N_g}\) is the projection onto \(N_g\) along the normal slices. First we parallel translate \(X\) along the shortest geodesic from \(\pi_{N_g}(p)\) to \(\varphi_g(p)\), and finally along the shortest geodesic to \(\varphi_g(p)\). We view \(\parallel\) as a canonical way to associate \(X\) to a vector in \(T_{\varphi_g(p)} M\).

Before we begin proving our estimates, following Section 2.1 of [We] we introduce two subbundles of \(T M_{\exp_{N_g}(\nu N_g)}\) and their orthogonal complements.

The **vertical bundle** \(\text{Vert}^s\) has fiber at \(p\) given by the parallel translation of \(\nu_{\pi_{N_g}(p)} N_g\) along the shortest geodesic from \(\pi_{N_g}(p)\) to \(p\).

The **almost vertical bundle** \(a \text{Vert}^s\) has fiber at \(p\) given by the tangent space at \(p\) of the normal slice to \(N_g\) through \(\pi_{N_g}(p)\).

The **horizontal bundle** \(\text{Hor}^s\) and the **almost horizontal bundle** \(a \text{Hor}^s\) are given by their orthogonal complements.

**Remark.** Notice that \(a \text{Vert}^s\) is the kernel of \(\pi_{N_g}\), and that according to Proposition 3.7 in [We] we have \(d(\text{Vert}^s, a \text{Vert}^s) < \frac{1}{4} d(p, N_g)^2\) for any \(p\) in \(\exp_{N_g}(\nu N_g)\), and similarly for \(\text{Hor}^s\) and \(a \text{Hor}^s\).

Since \(\frac{1}{4} d(p, N_g)^2 < \pi/2\), \(\text{Vert}^s\) and \(a \text{Hor}^s\) are always transversal (and clearly the same holds for \(\text{Hor}^s\) and \(a \text{Vert}^s\)). As seen in Section 1.3 \(\text{Vert}^s\) and \(TN\) are transversal along \(N\), and \(a \text{Vert}^s\) and \(TN\) are also transversal since \(N\) corresponds to a section of \(\nu N_g\) and \(a \text{Vert}^s = \text{Ker}(\pi_{N_g})\).

Now we are ready to give our estimates on the map \(\varphi_g\). Recall from the introduction that for any point \(q\) of the tubular neighborhood of \(N_g\) we denote by \(\gamma_q\) the geodesic from
\[ \pi_{N_g}(q) \in N_g \to q. \] Until the end of this section all geodesics will be parametrized by arc length.

In Sections 3 to 6 all estimates will hold for \( \epsilon < 1/20000. \)

### 3.1. Case 1: \( p \) is a point of \( N_g \)

**Proposition 3.1.** If \( p \in N_g \) and \( X \in T_pN_g \) is a unit vector, then

\[ |\varphi^\ast_g(X) - \|X\| \leq 3200 \epsilon. \]

**Remark.** Notice that if \( X \) is a vector normal to \( N_g \) by definition of \( \varphi_g \) and \( \| \) we have \( \varphi^\ast_g(X) = \|X\|. \) Therefore in this subsection we will assume that \( X \) is tangent to \( N_g \).

Also, we will denote by \( A \) the second fundamental form \(^{14}\) of \( N_g \), i.e. \( A \xi v := - (\nabla v \xi)^T \) for tangent vectors \( v \) of \( N_g \) and normal vector fields \( \xi \), where \( (\cdot)^T \) denotes projecting to the component tangent to \( N_g \) and \( \nabla \) is the Levi-Civita connection on \( M \). Since \( (M, N_g) \) is a gentle pair, the norm of \( A \) is bounded by \( 3/2 \), as shown in \([\text{We}, \text{Cor. 3.2}]\).

Now let \( p \in N_g, X \in T_pN_g \) a unit vector, and \( q := \varphi_g(p) \). We will denote by \( E \) the distance \( d(p, \varphi_g(p)) < 100 \epsilon \) (see end of Section 4 in \([\text{We}]\)). We will show that at \( q \),

\[ \|X \approx J(E) \approx H \approx \varphi^\ast_g(X) \]

where the Jacobi field \( J \) and the horizontal vector \( H \) will be specified below.

**Lemma 3.1.** Let \( J \) be the Jacobi field along the geodesic \( \gamma_q \) such that \( J(0) = X \) and \( J'(0) = -A\dot{\gamma}_q(0)X \). Then

\[ |J(E) - \|X\| \leq \frac{3}{2}(e^E - 1). \]

**Proof.** This is an immediate consequence of \([\text{BK}, 6.3.8]\) which will be used later again and which under the curvature assumption \( |K| \leq 1 \) states the following: if \( J \) is any Jacobi field along a unit-speed geodesic, then we have

\[ |J(t) - J(0)| \leq |t| \leq |J(0)||\cosh(t) - 1| + |J'(0)||\sinh(t) - t|, \]

where \( 0\| \) denotes parallel translation to the starting point of the geodesic. Using \( |A\dot{\gamma}_q(0)X| \leq 3/2 \) by \([\text{We}, \text{Cor. 3.2}]\) the above estimates gives \( |J(E) - \|X\| \leq (\cosh(E) - 1) + \frac{3}{2}\sinh(E) \). Alternatively, this lemma can be proven using the methods of \([\text{We}, \text{Prop. 3.7}]\). \( \square \)

Before proceeding we need a lemma about projections:

\(^{14}\) In Section 2 we adopted the sign convention of \([\text{We}]\) which differs from this.

\(^{15}\) \([\text{BK}, 6.3.8]\) assumes that \( J(0) \) and \( J'(0) \) are linearly dependent. However statement iii holds without this assumption, as one can always decompose \( J \) as \( J = J_1 + J_2 \), where \( J_1 \) and \( J_2 \) are Jacobi fields such that \( J_1(0) = J(0), J_1'(0) = 0 \) and \( J_2(0) = 0, J_2'(0) = J'(0) \) respectively. Furthermore we make use of \( |J'(0)| \leq |J'(0)| \).
Lemma 3.2. If $Y \in T_q M$ is a vertical unit vector, write $Y = Y_{av} + Y_h$ for the splitting into its almost vertical and horizontal components. Then

$$|Y_h| \leq \tan(E^2/4) \quad \text{and} \quad |Y_{av}| \leq \frac{1}{\cos(E^2/4)}.$$ 

Proof. By [We, Prop. 3.7] we have $d(\text{Vert}^E_q, \text{Vert}^E_q) \leq E^2/4 < \pi/2$, so the subspace $\text{Vert}^E_q$ of $T_q M$ is the graph of a linear map $\phi : \text{Vert}^E_q \to \text{Hor}^E_q$. So $Y_{av} = Y - \psi(Y)$ and $Y_h = -\phi(Y)$. Since the angle enclosed by $Y$ and $Y_{av}$ is at most $d(\text{Vert}^E_q, \text{Vert}^E_q) \leq E^2/4$, one obtains $|Y| \geq \cos(E^2/4)/|Y_{av}|$, which gives the second estimate of the lemma. From this, using $|Y_h|^2 = |Y_{av}|^2 - |Y|^2$ we obtain the first estimate. \hfill $\square$

Lemma 3.3. If $H$ is the unique horizontal vector at $q$ such that $(\pi_{N_q})_*(H) = X$, then

$$|J(E) - H| \leq \frac{3}{2}(e^E - 1)\frac{1}{\cos(E^2/4)}.$$ 

Proof. Let $J$ be the Jacobi field of Lemma 3.2. Write $J(E) = W + Y$ for the splitting into horizontal and vertical components. Then, using the notation of Lemma 3.2 we have $J(E)_{av} = W + Y_h$ and $J(E)_{av} = Y_{av}$. Notice that the Jacobi field $J$ arises from a variation of geodesics orthogonal to $N_q$ (see the Remark in Section 3.2), so $(\pi_{N_q})_*J(E) = X = (\pi_{N_q})_*H$. Using a variation $\phi$ of $(\pi_{N_q})_*$ it follows that $H = J(E)_{av}$. So

$$|J(E) - H| = |Y_{av}| \leq |Y| \frac{1}{\cos(E^2/4)} \leq \frac{3}{2}(e^E - 1)\frac{1}{\cos(E^2/4)}$$

where we used Lemma 3.2 and $|Y| \leq |J(E)| \leq \|X\|$ together with Lemma 3.1. \hfill $\square$

Now we will compare $H$ to $\psi_{gs}(X)$ and finish our proof.

Proof of Proposition 3.1. We have

$$\|X - \psi_{gs}(X)\| \leq \|X - J(E)\| + |J(E) - H| + |H - \psi_{gs}(X)|.$$ 

The first and second terms are bounded by the estimates of Lemmata 3.2 and 3.3. For the third term we proceed analogously to Lemma 3.2 since $\psi_{gs}(X)$ and $H$ are both mapped to $X$ via $\pi_{N_q}$, one has $(\psi_{gs}(X))_{av} = \psi_{gs}(X) - H$. As earlier, if $\psi_{gs}(X) = \tilde{W} + \tilde{Y}$ is the splitting into horizontal and vertical components, we have $(\psi_{gs}(X))_{av} = \tilde{Y}_{av}$. Therefore

$$|\psi_{gs}(X) - H| = |\tilde{Y}_{av}| \leq |\tilde{Y}| \frac{1}{\cos(E^2/4)} \leq |\psi_{gs}(X)| \frac{\sin(2500\epsilon)}{\cos(E^2/4)}.$$ 

Here we also used Lemma 3.2 and the fact that the angle enclosed by $\psi_{gs}(X)$ and its orthogonal projection onto $\text{Hor}^E_q$ is at most $d(H_{av}, T_q N) \leq 2500\epsilon$ by Theorem 4. Altogether we have

$$\|X - \psi_{gs}(X)\| \leq \frac{3}{2}(e^E - 1)\left[1 + \frac{1}{\cos(E^2/4)}\right] + |\psi_{gs}(X)| \frac{\sin(2500\epsilon)}{\cos(E^2/4)}.$$ 

Using this inequality we can bound $|\psi_{gs}(X)|$ from above in terms of $E$ and $\epsilon$. Substituting into the right hand side of the above inequality we obtain a function of $\epsilon$ (recall that $E = 100\epsilon$) which is increasing and bounded above by $3200\epsilon$. \hfill $\square$
3.2. Case 2: \( p \) is a point of \( \partial \exp_{N_p}(\nu N_p)_L \) and \( X \in T_p M \) is almost vertical

In this subsection we require \( L < 1 \), as in the definition of gentle pair.

**Remark.** Jacobi fields \( J \) along \( \gamma_p \) (the geodesic from \( \pi_{N_p}(p) \) to \( p \)) with \( J(0) \) tangent to \( N_g \) and \( A_{\gamma_p(0)} \bar{J}(0) + J'(0) \) normal to \( N_g \) are called \( N_g \)-Jacobi fields. They clearly form a vector space of dimension equal to \( \dim(M) \) and they are exactly the Jacobi fields that arise from variations of \( \gamma_p \) by geodesics that start on \( N_g \) and are normal to \( N_g \) there.

Since \((M, N_g)\) is a gentle pair, there are no focal points of \( \pi_{N_g}(p) \) along \( \gamma_p \), so the map

\[
\{ N_g \text{-Jacobi fields along } \gamma_p \} \to T_p M, \quad J \mapsto \bar{J}(L),
\]
is an isomorphism. The \( N_g \)-Jacobi fields that map to a\( \text{Vert}^g \) are exactly those with \( J(0) = 0 \) and \( J'(0) \in v_{\pi_{N_g}(p)} N_g \). Indeed, such a vector field is the variational vector field of the variation

\[
f_s(t) = \exp_{\pi_{N_g}(p)} t[\dot{\gamma}_p(0) + s J'(0)],
\]
so \( J(L) \) will be tangent to the normal slice of \( N_g \) at \( \pi_{N_g}(p) \). From dimension considerations it follows that the \( N_g \)-Jacobi fields that satisfy \( J(0) = 0 \) and \( J'(0) = 0 \) are exactly those with \( \gamma_p \) that satisfy \( J'(0) = 0 \); which are called strong \( N_g \)-Jacobi fields—map to a subspace of \( T_p M \) which is a complement of a\( \text{Vert}^g \). As pointed out in [Wa, p. 354], these two subspaces are in general not orthogonal.

**Proposition 3.2.** If \( p \in \partial \exp_{N_p}(\nu N_p)_L \) and \( X \in T_p M \) is an almost vertical unit vector, then

\[
|\varphi_g s(X) - \|X\| \leq \frac{2 \sinh(L) - L}{\sin(L)}.
\]

We begin by proving

**Lemma 3.4.** Let \( J \) be a Jacobi field along \( \gamma_p \) such that \( J(0) = 0 \) and \( J'(0) \in v_{\pi_{N_g}(p)} N_g \), normalized so that \( |J(L)| = 1 \). Then

\[
|J(L) - L \cdot \gamma_p \|J'(0)\| \leq \frac{\sinh(L) - L}{\sin(L)}.
\]

**Proof.** Again [BK] 6.3.8iii] shows that \( |J(L) - L \|J'(0)\| \leq |J'(0)| (\sinh(L) - L) \). Using the upper curvature bound \( K \leq 1 \) and Rauch’s theorem we obtain \( |J'(0)| \leq 1/\sin(L) \) and we are done. \( \square \)

We saw in the remark above that \( X \) is equal to \( J(L) \) for a Jacobi field \( J \) as in Lemma 3.4 and that \( J \) comes from a variation \( f_s(t) = \exp_{\pi_{N_g}(p)} t[\dot{\gamma}_p(0) + s J'(0)] \). So \( \varphi_g s(X) \) comes from the variation

\[
\varphi_g s(f_s(t)) = \exp_{\pi_{N_g}(p)} t[\|\dot{\gamma}_p(0) + s \|J'(0)\|]
\]
along the geodesic \( \varphi_g (\gamma_p(t)) \). More precisely, if we denote by \( \tilde{J}(t) \) the Jacobi field that arises from the above variation, we will have \( \varphi_g s(X) = \tilde{J}(L) \). Notice that \( \tilde{J}(0) = 0 \) and \( \tilde{J}'(0) = \|J'(0)\| \).
Lemma 3.5.

\[ |\tilde{J}(L) - L \cdot \varphi_g \gamma_p \| \dot{\tilde{J}}(0)\| | \leq \frac{\sinh(L) - L}{\sin(L)}. \]

Proof. Exactly as for Lemma 3.4 since \( \tilde{J}(0) = 0 \) and \( |\dot{J}(0)| = |J'(0)|. \)

Proof of Proposition 3.2. We have \( X \approx LJ'_{\tilde{J}(0)} = L \tilde{J}'_{\tilde{J}(0)} \approx \varphi_g^*(X) \). Here we identify tangent spaces to \( M \) parallel translating along \( \gamma_p \), along the geodesic \( \varphi_g(\pi_{N_g}(p)) \) from \( \pi_{N_g}(p) \) to its \( \varphi_g \)-image and along \( \varphi_g \circ \gamma_p \) respectively. Notice that these three geodesics are exactly those used in the definition of “\( \| \)”. The estimates for the two relations “\( \approx \)” are in Lemmata 3.4 and 3.5 respectively (recall \( X = J(L) \) and \( \varphi_g(X) = \tilde{J}(L) \)), and the equality holds because \( \tilde{J}'_{\tilde{J}(0)} = J'_{J(L)}. \)

3.3. Case 3: \( p \) is a point of \( \partial \exp_{N_g} (vN_g)_L \) and \( X = J(L) \) for some strong \( N_g \)-Jacobi field \( J \) along \( \gamma_p \).

From now on we have to assume \( L < 0.08. \)

Proposition 3.3. If \( p \in \partial \exp_{N_g}(vN_g)_L \) and \( X \) is a unit vector equal to \( J(L) \) for some strong \( N_g \)-Jacobi field \( J \) along \( \gamma_p \), then

\[ |\varphi_g(X) - \|X\| \leq \frac{18}{5}L + 3700\epsilon. \]

We proceed analogously to Case 2.

Lemma 3.6. For a vector field \( J \) as in the above proposition we have

\[ |J(L) - \varphi_g \gamma_p \| \dot{J}(0)\| | \leq \frac{3}{2}(e^L - 1) \leq \frac{9}{5}L. \]

Furthermore we have \( |\dot{J}(0)| \leq \frac{1}{1 - \frac{3}{2}(e^L - 1)}. \)

Proof. By Lemma 3.1 we have \( |J(L) - \varphi_g \gamma_p \| \dot{J}(0)\| | \leq \frac{3}{2}(e^L - 1)|J(0)| \), from which we obtain the estimate for \( |\dot{J}(0)| \) and then the first estimate of the lemma.

\( J \) comes from a variation \( f_s(t) = \exp_{\sigma(t)} t \gamma \) for some curve \( \sigma \) in \( N_g \) with \( \dot{\sigma}(0) = J(0) \) and some normal vector field \( \gamma \) along \( \sigma \). We denote by \( \tilde{J} \) the Jacobi field along the geodesic \( \varphi_g(\gamma_p(t)) \) arising from the variation

\[ \tilde{f}_s(t) = \varphi_g(f_s(t)) = \exp_{\tilde{\sigma}(t)} (t \gamma_p(t)), \]

where \( \tilde{\sigma} = \varphi_g \circ \sigma \) is the lift of \( \sigma \) to \( N \). Then we have \( \tilde{J}(L) = \varphi_g(X) \). Notice that here \( \gamma_p(t) \) is just the parallel translation of \( \gamma(t) \) along \( \gamma_p(t) = \gamma \).

Lemma 3.7.

\[ |\tilde{J}(L) - \varphi_g \gamma_p \| \dot{\tilde{J}}(0)\| | \leq \frac{9}{5}L. \]
Proof. Using \[\textup{[BK] 6.3.8iii} \] as in Lemma 3.1 we obtain
\[|\dot{J}(L) - \varphi_{L} J(0)| \leq |\dot{J}(0)| \sinh(L) + |\dot{J}(0)| (\cosh(L) - 1), \tag{*} \]
so that we just have to estimate the norms of \(\dot{J}(0)\) and \(\ddot{J}(0)\).

Since \(\dot{J}(0) = \varphi_{L} J(0)\), Proposition 3.1 gives \(|\dot{J}(0) - \gamma_{0} J(0)| \leq 3200 \epsilon |J(0)|\). Using the bound for \(|J(0)|\) given in Lemma 3.6 we obtain
\[|\dot{J}(0)| \leq \frac{1 + 3200 \epsilon}{1 - \frac{1}{2} (e^L - 1)}. \]

To estimate \(\ddot{J}(0)\) notice that in the expression for \(f_{\epsilon}(t)\) we can choose \(v(s) = \sigma_{\gamma}(\gamma_{0}(0) + s J'(0))\), where \(\sigma_{\gamma}\) denotes parallel translation from \(0\) to \(s\) along \(\gamma\). So
\[\|v(s)\| = \gamma_{\gamma_{0}}(\gamma_{0}(0) + s J'(0)), \]
and
\[\ddot{J}(0) = \nabla_{s} \gamma_{\gamma_{0}}(\gamma_{0}(0) + s J'(0)), \]
where we used the Leibniz rule for covariant derivatives to obtain the second equality.

To estimate the first term note that the difference between the identity and the holonomy around a loop in a Riemannian manifold is bounded in the operator norm by the area of a surface spanned by the loop times a bound for the curvature (see \[\textup{[BK] 6.2.1} \]). Therefore we write \(\gamma_{\gamma_{0}}(\gamma_{0}(0) + s J'(0))\) as \(\gamma_{\gamma_{0}}(\gamma_{0}(0) + \epsilon(s))\) where \(\epsilon(s)\) is a vector field along \(\gamma(s)\) with norm bounded by the area of the polygon spanned by \(\gamma(0), \gamma(s), \gamma(s)\) and \(\gamma(0)\). Assuming that \(\gamma\) has constant speed \(|\dot{J}(0)|\) we can estimate \(d(\sigma(0), \sigma(s)) \leq |\dot{J}(0)|\), and using Proposition 3.1 to estimate \(|\dot{\gamma}(s)| = |\varphi_{\gamma_{0}} \dot{\gamma}(s)|\) we obtain \(d(\dot{\gamma}(0), \dot{\gamma}(s)) \leq s(1 + 3200 \epsilon) |\dot{J}(0)|\). Using \(d(\dot{\gamma}(s), \dot{\gamma}(s)) \leq 100 \epsilon\) and Lemma 3.6 we can bound the area of the polygon safely by
\[
\frac{100 \epsilon s (2 + 3200 \epsilon)}{1 - \frac{1}{2} (e^L - 1)}. \]
So we obtain
\[
\left| \frac{\nabla}{d s} \gamma_{\gamma_{0}}(\gamma_{0}(0)) \right| = \left| \frac{\nabla}{d s} \epsilon(s) \right| \leq \frac{100 \epsilon (2 + 3200 \epsilon)}{1 - \frac{1}{2} (e^L - 1)}.
\]
To bound \(\gamma_{\gamma_{0}}(J'(0))\) notice that \(|J'(0)| \leq \frac{1}{2} |J(0)|\) using the fact that \(J\) is a strong Jacobi field and \[\textup{[We, Cor. 3.2]} \], so
\[|J'(0)| \leq \frac{3}{2} \frac{1}{1 - \frac{1}{2} (e^L - 1)}. \]
Substituting our estimates for \(|\dot{J}(0)|\) and \(|\ddot{J}(0)|\) in \(\tag{*}\) we obtain a function which, for \(\epsilon < 1/20000\) and \(L < 0.08\), is bounded above by \(\frac{3}{2} L\).

\[\square\]

The vectors \(\dot{J}(0)\) and \(\ddot{J}(0)\) generally are not equal, so we need one more estimate that has no counterpart in Case 2.
Lemma 3.8.

\[ |\tilde{J}(0) - \langle J(0) \rangle| \leq \frac{3200\epsilon}{1 - 3(e^L - 1)} \leq 3700\epsilon. \]

Proof. Since \( \tilde{J}(0) = \varphi_{g*} J(0) \), Proposition 3.1 gives

\[ |\tilde{J}(0) - \langle J(0) \rangle| \leq 3200\epsilon |J(0)| \leq \frac{3200\epsilon}{1 - \frac{3}{2}(e^L - 1)}. \]

Since \( \frac{1}{1 - \frac{3}{2}(e^L - 1)} < 1.15 \) when \( L < 0.08 \) we are done.

Proof of Proposition 3.3 We have \( X \approx J(0) \approx \tilde{J}(0) \approx \varphi_{g*}(X) \) where we identify tangent spaces by parallel translation along \( \gamma_p, \gamma_0 \) and \( \varphi_g \circ \gamma_p \) respectively. Combining the last three lemmas and recalling \( X = J(L), \varphi_{g*}(X) = \tilde{J}(L) \) we finish the proof.

3.4. The general case

This proposition summarizes the three cases considered up to now:

Proposition 3.4. Assume \( \epsilon < 1/20000 \) and \( L < 0.08 \). Let \( p \in \partial \exp_{N_g} (\nu N_g)_L \) and \( X \in T_p M \) a unit vector. Then

\[ |\varphi_{g*}(X) - \langle X \rangle| \leq 4L + 4100\epsilon. \]

We will write the unit vector \( X \) as \( J(L) + K(L) \) where \( J \) and \( K \), up to normalization, are Jacobi fields as in the next lemma. We will need to estimate the norms of \( J(L) \) and \( K(L) \), so we begin by estimating the angle they enclose:

Lemma 3.9. Let \( J \) be an \( N_g \)-Jacobi field along \( \gamma_p \) with \( J(0) = 0 \), \( J'(0) \) normal to \( N_g \) (as in Case 2) and \( K \) a strong \( N_g \)-Jacobi field (as in Case 3), normalized so that \( J(L) \) and \( K(L) \) are unit vectors. Then

\[ |\langle J(L), K(L) \rangle| \leq \frac{\frac{3}{2}(e^L - 1)}{1 - \frac{3}{2}(e^L - 1)} + \frac{1}{1 - \frac{3}{2}(e^L - 1)} \cdot \frac{\sinh(L) - L}{\sin(L)} \leq \frac{9}{5} L. \]

Proof. Identifying tangent spaces along \( \gamma_p \) by parallel translation, we have

\[ |\langle J(L), K(L) \rangle| = |\langle J(L), K(L) \rangle - \langle L J'(0), K(0) \rangle| \]

\[ \leq |\langle J(L), K(L) - K(0) \rangle| + |\langle J(L) - L J'(0), K(0) \rangle| \]

\[ \leq |K(L) - K(0)| + |K(0)| \cdot |J(L) - L J'(0)|, \]

which can be estimated using Lemmata 3.6 and 3.4.

Lemma 3.10. Let \( X \in T_p M \) be a unit vector such that \( X = J(L) + K(L) \) where \( J, K \) are Jacobi fields as in Lemma 3.9 (up to normalization). Then

\[ |J(L)|, |K(L)| \leq \frac{1}{\sqrt{1 - \frac{9}{5} L}} \leq 1.1. \]
Proof. Let \( c := (J(L)/|J(L)|, K(L)/|K(L)|) \), so \(|c| \leq \frac{3}{4} L \). There is an orthonormal basis \( \{ e_1, e_2 \} \) of \( \text{span}(J(L), K(L)) \) such that \( J(L) = |J(L)| e_1 \) and \( K(L) = |K(L)|(ce_1 + \sqrt{1 - c^2} e_2) \). An elementary computation shows that \( 1 = |J(L) + K(L)|^2 \geq (1 - |c|)(|J(L)|^2 + |K(L)|^2) \), from which the lemma easily follows. \( \square \)

Proof of Proposition 3.10. The remark at the beginning of Case 2 implies that we can (uniquely) write \( X = J(L) + K(L) \) for \( N_g \)-Jacobi fields \( J \) and \( K \) as in Lemma 3.10. So, by Lemma 3.10, Proposition 3.2 and Proposition 3.3,

\[
|\varphi_g^*(X) - \|X| \leq |\varphi_g^* J(L) - \|J(L)| + |\varphi_g^* K(L) - \|K(L)|
\leq 1.1 \left( \frac{2 \sinh(L) - L}{\sin(L)} + \frac{18}{5} L + 3700 \epsilon \right) \leq 4L + 4100 \epsilon.
\]

\( \square \)

4. Proposition 4.1 about geodesic triangles in \( M \)

Fix \( g \) in \( G \) and let \( \varphi_g \) be the map from a tubular neighborhood of \( N_g \) to one of \( N \) defined in the introduction. Our aim in the next three sections is to show that \( \exp_{N_g}(vN_g)_{0.05} \) is a tubular neighborhood of \( N_g \) on which \( \varphi_g \) is injective.

We will begin by giving a lower bound on the length of edges of certain geodesic triangles in \( M \).

In this section we take \( M \) to be simply any Riemannian manifold with the following two properties:

(i) the sectional curvature lies between \(-1\) and \(1\),
(ii) the injectivity radius at any point is at least 1.

In our later applications we will work in the neighborhood of a submanifold that forms a gentle pair with \( M \), so these two conditions will be automatically satisfied.

Now choose points \( A, B, C \) in \( M \) and assume \( d(C, A) < 0.15 \) and \( d(C, B) < 0.5 \). Connecting the three points by the unique shortest geodesics defined on the interval \([0, 1]\), we obtain a geodesic triangle \( ABC \).

We will denote by the symbol \( \dot{CB} \) the initial velocity vector of the geodesic from \( C \) to \( B \), and similarly for the other edges of the triangle.

Proposition 4.1. Let \( M \) be a Riemannian manifold and \( ABC \) a geodesic triangle as above. Let \( P_C \) and \( P_A \) be subspaces of \( T_CM \) and \( T_AM \) respectively of equal dimensions such that \( \dot{CB} \in P_C \) and \( \dot{AB} \in P_A \). Assume that

\[
\angle(P_A, \dot{AC}) \geq \pi/2 - \delta
\]

and

\[
\theta := d(P_A, C) \leq C d(A, C)
\]

for some constants \( \delta, C \). Assume \( C \leq 2 \). Then

\[
d(C, B) \geq \frac{10}{11} \frac{1}{C + 1} \cos(\delta).
\]
Remark 1. Here \( C_M \parallel P_C \) denotes parallel translation of \( P_C \) along the geodesic from \( C \) to \( A \). The angle between the subspace \( P_A \) and the vector \( \dot{A}C \) is given as follows: for every nonzero \( v \in P_A \) we consider the nonoriented angle \( \angle(v, \dot{A}C) \in [0, \pi] \). Then we have

\[
\angle(P_A, \dot{A}C) := \min\{\angle(v, \dot{A}C) : v \in P_A \text{ non}\zero\} \in [0, \pi/2].
\]

Notice that \( \angle(P_A, \dot{A}C) \geq \pi/2 - \delta \) iff for all nonzero \( v \in P_A \) we have \( \angle(v, \dot{A}C) \in [\pi/2 - \delta, \pi/2 + \delta] \).

Remark 2. This proposition generalizes the following simple statement about triangles in the plane: if two edges \( CB \) and \( AB \) form an angle bounded by the length of the base edge \( AC \) times a constant \( C \), and if we assume that \( CB \) and \( AB \) are nearly perpendicular to \( AC \), then the lengths \( |CB| \) and \( |AB| \) will be bounded below by a constant depending on \( C \) (but not on \( |AC| \)).

In the general case of Proposition 4.1 however, we make assumptions on \( d(P_A, C_M \parallel P_C) \) from which we are not able to obtain easily bounds on the angle \( \angle(CB, AB) \) at \( B \) (such a bound together with the law of sines would immediately imply the statement of the proposition).

Proof of Proposition 4.1. Using the chart \( \exp_A \) we can lift \( B \) and \( C \) to the points \( \tilde{B} \) and \( \tilde{C} \) of \( TA_M \). We obtain a triangle \( 0\tilde{B}\tilde{C} \), which differs in one edge from the lift of the triangle \( ABC \). Denoting by \( Q \) the endpoint of the vector \( \tilde{B} - \tilde{C} \) translated to the origin, consider the triangle \( 0\tilde{B}Q \). Let \( P \) be the closest point to \( Q \) in \( P_A \).

Claim 1.

\[
|\tilde{B} - P| \leq \tan(\delta)|Q - P|.
\]

Using \( \angle(P_A, \dot{A}C) \geq \pi/2 - \delta \) and \( \dot{A}C = \tilde{C} - 0 \) we see that the angle between any vector in \( P_A \) and \( \tilde{C} - 0 \) lies in the interval \([\pi/2 - \delta, \pi/2 + \delta]\). Since \( \tilde{C} - 0 \) and \( Q - \tilde{B} \) are parallel, the angle between any vector of \( P_A \) and \( Q - \tilde{B} \) lies in \([\pi/2 - \delta, \pi/2 + \delta]\). Since \( P - \tilde{B} \in P_A \) we have

\[
\angle(P - \tilde{B}, Q - \tilde{B}) \in [\pi/2 - \delta, \pi/2 + \delta].
\]

The triangle \( \tilde{B}PQ \) has a right angle at \( P \), so \( \angle(P - \tilde{B}, Q - \tilde{B}) \leq \delta \), and Claim 1 follows.

Claim 2.

\[
|Q - P| \leq \sin[(1 + C) \cdot d(C, A)] \cdot |Q - 0|.
\]

In Corollary A.1 of Appendix A we will estimate the angle between \( \tilde{B} - \tilde{C} = Q - 0 \in TA_M \) and \( C_M \parallel CB \in TA_M \), i.e. the parallel translation in \( M \) of \( \dot{C}B \) along the geodesic from \( C \) to \( A \). Our estimate will be

\[
\angle(C_M \parallel \dot{C}B, Q - 0) \leq \frac{1}{2}d(A, C).
\]

Now let \( P' \) be the closest point to \( C_M \parallel \dot{C}B \) in \( P_A \). As \( P' - 0 \in P_A \) and \( \dot{C}B \in P_C \), using the definition of distance between subspaces we get

\[
\angle(C_M \parallel \dot{C}B, P' - 0) \leq d(P_A, C_M \parallel P_C) = \theta \leq C d(C, A).
\]

Finally, we will show (see Corollary A.2) that

\[
\angle(P - 0, P' - 0) \leq \frac{1}{2}d(C, A).
\]
Combining the last three estimates we get \( \angle (P - 0, Q - 0) \leq (1 + C)d(C, A) \), which is less than \( \pi/2 \). Claim 2 follows since 0\( PQ \) is a right triangle at \( P \).

**Claim 3.**

\[ d(C, B) \geq \frac{10}{11} \frac{1}{C + 1} \cos(\delta). \]

The triangle \( B \hat{P} Q \) is a right triangle at \( P \), so using Claims 1 and 2 we have

\[ |Q - \tilde{B}|^2 = |\tilde{B} - P|^2 + |Q - P|^2 \leq (1 + \tan^2 \delta) \cdot |Q - P|^2 \]
\[ \leq (1 + \tan^2 \delta) \cdot (1 + C)^2 \cdot d(C, A)^2 \cdot |Q - 0|^2. \]

The vector \( Q - \tilde{B} \) is just \( 0 - \tilde{C} \), the length of which is \( d(A, C) \), and the vector \( Q - 0 \) is \( B - \tilde{C} \). So

\[ d(A, C) \leq \sqrt{1 + \tan^2 \delta} (1 + C) \cdot d(C, A)|\tilde{B} - \tilde{C}|, \]

and

\[ \frac{1}{(1 + C)\sqrt{1 + \tan^2 \delta}} \leq |\tilde{B} - \tilde{C}|. \]

Using standard estimates (see Corollary \( \text{A.3} \)) we obtain \( |\tilde{B} - \tilde{C}| \leq \frac{11}{10}d(C, B) \), and since \( 1/\sqrt{1 + \tan^2 \delta} = |\cos(\delta)| \) the proposition follows. \( \square \)

**5. Application of Proposition 4.1 to \( \text{Vert}^g \)**

Fix \( g \) in \( G \). Let \( C \) and \( A \) be points on Weinstein’s average \( N \) with \( d(C, A) < 0.15 \) joined by a minimizing geodesic \( \gamma \) in \( M \). Suppose that \( \exp_C(v) = \exp_A(w) =: B \) for
vertical vectors $v \in \text{Vert}_v^g$ and $w \in \text{Vert}_w^g$ of lengths less than 0.5. In this section we will apply Proposition 4.1 to the geodesic triangle given by the above three points of $M$ and $P_A = \text{Vert}_A^g$, $P_C = \text{Vert}_C^g$. We will do so in Proposition 5.5.

To this end, first we will estimate the constants $\delta$ and $C$ of Proposition 4.1 in this specific case. As always our estimates will hold for $\epsilon < 1/20000$.

Roughly speaking, the constant $\delta$—which measures how much the angle between $\dot{C}A = \dot{\gamma}(0)$ and $\text{Vert}_C^g$ deviates from $\pi/2$—will be determined by using the fact that $N$ is $C^1$-close to $N_g$, so that the shortest geodesic $\gamma$ between $C$ and $A$ is “nearly tangent” to the distribution $\text{Hor}_C^g$.

Bounding the constant $C$—which measures how the angle between $\text{Vert}_A^g$ and $\text{Vert}_C^g$ depends on $d(A, C)$—will be easier, by noticing that both spaces are parallel translations of normal spaces to $N_g$, which is a submanifold with bounded second fundamental form.

Since

$$\angle(\dot{\gamma}(0), \text{Vert}_C^g) = \pi/2 - \angle(\dot{\gamma}(0), \text{Hor}_C^g),$$

to determine $\delta$ we just have to estimate the angle

$$\alpha := \angle(\dot{\gamma}(0), \text{Hor}_C^g).$$

We already introduced the geodesic $\gamma(t)$ from $C$ to $A$, which we assume to be parametrized by arc length. We now consider the curve $\pi(t) := \pi_{N_g} \circ \gamma(t)$ in $N_g$. We can lift the curve $\pi$ to a curve $\varphi_{\pi} \circ \pi$ in $N$ connecting $C$ and $A$; we will call $c(t)$ the parametrization by arc length of this lift.

$\nabla^\perp$ will denote the connection induced on $\nu N_g$ by the Levi-Civita connection $\nabla$ of $M$, and $\varphi_{\pi} \circ \pi$ applied to some $\xi \in \nu_{\pi(t)}N_g$ will denote its $\nabla^\perp$-parallel transport from $\pi(t)$ to $\pi(0)$ along $\pi$. (The superscript “b” stands for “backwards” and is a reminder that we are parallel translating to the initial point of the curve $\pi$.)

Further we will need

$$r := 100\epsilon + L(\gamma)/2 \geq \sup_{t} \{d(\gamma(t), N_g)\} \quad \text{and} \quad f(r) := \cos(r) - \frac{3}{2} \sin(r).$$

Notice that $r < 0.08$ due to our restrictions on $\epsilon$ and $d(C, A)$. 

---

![Diagram](image_url)
Using the fact that \( c \) is a curve in \( N \) and \( N \) is \( C^1 \)-close to \( N_g \), in Appendix B we will show that the section \( \tilde{c} := \exp^{-1}_{N_g}(c(t)) \) of \( vN_g \) along \( \pi \) is “approximately parallel”. This will allow us to bound from above the “distance” between its endpoints as follows:

**Proposition 5.1.**

\[
|\exp^{-1}_{N_g}(C) - \frac{1}{\pi} \exp^{-1}_{N_g}(A)| \leq L(\gamma) \frac{3150\epsilon}{f(r)}.
\]

Using the fact that \( \gamma \) is a geodesic and our bound on the extrinsic curvature of \( N_g \), in Appendix C we will show that the section \( \tilde{\gamma} := \exp^{-1}_{N_g}(\gamma(t)) \) of \( vN_g \) along \( \pi \) approximately “grows at a constant rate”. Since its covariant derivative at zero depends on \( \alpha \), we will be able to estimate the “distance” between its endpoints (which are also the endpoints of \( \tilde{c} \)) in terms of \( \alpha \). We will obtain:

**Proposition 5.2.**

\[
|\exp^{-1}_{N_g}(C) - \frac{1}{\pi} \exp^{-1}_{N_g}(A)| \geq L(\gamma) \left[ \frac{99}{100} \sin \left( \frac{\alpha - \epsilon}{4} \right) - 500\epsilon - 3r - \frac{8}{3} L(\gamma) \left( r + \frac{r + 3/2}{f(r)} \right) \right].
\]

Comparison of Propositions 5.1 and 5.2 gives

\[
\frac{3150\epsilon}{f(r)} \geq \frac{99}{100} \sin \left( \frac{\alpha - \epsilon}{4} \right) - 500\epsilon - 3r - \frac{8}{3} L(\gamma) \left( r + \frac{r + 3/2}{f(r)} \right).
\]

Recall that \( r = 100\epsilon + L(\gamma)/2 \). If \( L(\gamma) \) and \( \epsilon \) are small enough one can solve the above inequality for \( \alpha \). With our restriction \( \epsilon < 1/20000 \) this can be done whenever \( L(\gamma) < 0.1 \).

One obtains

\[
\delta(\epsilon, L(\gamma)) := \frac{\epsilon}{4} + \arcsin \left\{ \frac{100}{99} \left[ \frac{3150\epsilon}{f(r)} + 500\epsilon + 3r + \frac{8}{3} L(\gamma) \left( r + \frac{r + 3/2}{f(r)} \right) \right] \right\} > \alpha.
\]

We can now state the main results of this section. First we determine the constant \( \delta \) of Proposition 4.1 in our setting.

**Proposition 5.3.** Let \( C, A \) be points in \( N \) and \( \gamma \) the shortest geodesic in \( M \) from \( C \) to \( A \). Assume \( \epsilon < 1/20000 \) and \( L(\gamma) < 0.1 \). Then \( \delta(\epsilon, L(\gamma)) \) is well defined and

\[
\angle(\text{Hor}_C^g, \dot{\gamma}(0)) = \alpha < \delta(\epsilon, L(\gamma)).
\]

Therefore

\[
\angle(\text{Vert}_C^g, \dot{\gamma}(0)) \geq \frac{\pi}{2} - \delta(\epsilon, L(\gamma))
\]

and for symmetry reasons

\[
\angle(\text{Vert}_A^g, -\dot{\gamma}(L(\gamma))) \geq \frac{\pi}{2} - \delta(\epsilon, L(\gamma)).
\]

To determine the constant \( C \) we only need Lemma C.3.
Proposition 5.4. Let $C$, $A$ and $\gamma$ be as above and assume $L(\gamma) < 0$. Then

$$d(\text{Vert}_C, \gamma \| \text{Vert}_A) \leq 2L(\gamma).$$

Proof. By Lemma C.3 we have

$$d(\text{Vert}_C, \gamma \| \text{Vert}_A) \leq \arcsin \left[ L(\gamma) \left( \frac{r + \frac{3}{2}}{f(r)} \right) \right],$$

where $r = 100\epsilon + L(\gamma)/2$. For the above values of $\epsilon$ and $L(\gamma)$ this last expression is bounded above by $2L(\gamma)$. $\square$

Now making use of the estimates in the last two propositions we can apply Proposition 4.1.

Proposition 5.5. Fix $g \in G$. Let $C, A$ be points in $N$ such that $d(A, C) < 0.05$ and suppose that $\exp_C(v) = \exp_A(w) =: B$ for vertical vectors $v \in \text{Vert}_C, w \in \text{Vert}_A$. Then

$$|v|, |w| \geq \frac{3}{10} \cos(\delta(\epsilon, d(A, C))).$$

Proof. If $|v| \geq 0.5$ then the estimate for $|v|$ clearly holds, as the right hand side is $\leq 3/10$. So we assume $|v| = d(B, C) < 0.5$.

Since $d(B, N_g) < 0.5 + 100\epsilon < 1$ and $(M, N_g)$ is a gentle pair, the triangle $ABC$ lies in an open subset of $M$ with the properties

(i) the sectional curvature lies between $-1$ and $1$,

(ii) the injectivity radius at each point is at least 1.

Therefore we are in the situation of Proposition 4.1. Setting $P_C = \text{Vert}_C$ and $P_A = \text{Vert}_A$ in the statement of Proposition 4.1, Propositions 5.3 and 5.4 allow us to choose

$$\delta = \delta(\epsilon, d(A, C))$$

and $C = 2$.

Therefore, since $\frac{10}{11} \cdot \frac{1}{2} > \frac{3}{10}$, we obtain

$$|v| \geq \frac{3}{10} \cos(\delta(\epsilon, d(A, C))).$$

The statement for $|w|$ follows exactly in the same way. $\square$

6. Estimates on tubular neighborhoods of $N_g$ on which $\varphi_g$ is injective

In this section we will finally apply the results of Sections 4 and 5, which were summarized in Proposition 5.5, to show that $\exp_{N_g}(v N_g)_{0.05}$ is a tubular neighborhood of $N$ on which $\varphi_g$ is injective. We will also bound from below the size of $\bigcap_{g \in G} \exp_{N_g}(\text{Vert}^g)_{0.05}$ (where the 2-form $\int_g \omega_g$ is defined).
Proposition 6.1. If $\epsilon < 1/20000$ the map

$$\varphi_g : \exp_N(\nu N)^{0.05} \to \exp_N(\text{Vert}^g)^{0.05}$$

is a diffeomorphism.

Proof. From the definition of $\varphi_g$ it is clear that it is enough to show the injectivity of

$$\exp_N : (\text{Vert}^g)^{0.05} \to \exp_N(\text{Vert}^g)^{0.05}.$$

Let $A, C \in N$ and $v \in \text{Vert}^g_A$, $w \in \text{Vert}^g_A$ be vectors of length $< 0.05$. We argue by contradiction and suppose that $\exp_C(v) = \exp_A(w)$. Clearly $d(A, C) < 0.1$. We can apply Proposition 5.5, which implies $|v|, |w| \geq \frac{1}{10} \cos(\delta(\epsilon, d(A, C)))$. Since the function $\delta(\epsilon, L)$ increases with $L$ we have $|v|, |w| \geq \frac{3}{10} \cos(\delta(\epsilon, 0.1))$.

For $\epsilon < 1/20000$ the above function is larger than $0.05$, so we have a contradiction. Hence $\exp_C(v) \neq \exp_A(w)$ and the above map is injective. \qed

For each $L \leq 0.05$ we want to estimate the radius of a tubular neighborhood of $N$ contained in $\bigcap_{g \in G} \exp_N(\text{Vert}^g)_L$. This will be used in Section 7 to determine where $\int_g \omega_g$ is nondegenerate, so that one can apply Moser’s trick there. As a by-product, the proposition below will also give us an estimate of the size of the neighborhood in which $\int_g \omega_g$ is defined.

Proposition 6.2. For $L \leq 0.05$ and $\epsilon < 1/20000$, using the notation

$$R^g_L := \sin(L) \cos(\delta(\epsilon, 2L) + 2L^2)$$

we have

$$\exp_N(\nu N)^{R^g_L} \subset \bigcap_{g \in G} \exp_N(\text{Vert}^g)_L.$$

Remark. The function $R^g_{0.05}$ decreases with $\epsilon$ and assumes the value $0.039\ldots$ at $\epsilon = 0$ and the value $0.027\ldots$ when $\epsilon = 1/20000$.

To prove the proposition we will again consider geodesic triangles:

Lemma 6.1. Let $ABC$ be a geodesic triangle lying in $\exp_N(\nu N)^{0.05} \subset \exp_N(\text{Vert}^g)_L$ such that $d(A, B) \leq L < 0.05$. Let $\gamma$ denote the angle at $C$, and suppose $\gamma \in [\pi/2 - \delta, \pi/2 + \delta]$. Then

$$d(A, B) \geq \sin(L) \cos(\delta + 2L^2).$$
Proof. Denote by $\alpha, \beta$ the angles at $A$ and $B$ respectively, and denote further by $\alpha', \beta', \gamma'$ the angles of the Aleksandrov triangle in $S^2$ corresponding to $ABC$ (i.e. the triangle in $S^2$ having the same side lengths as $ABC$). By [Kl, Remark 2.7.5] we have
\[
\sin(d(A, B)) = \sin(d(C, B)) \frac{\sin(\gamma')}{\sin(\alpha')} \geq \sin(d(C, B)) \sin(\gamma').
\]
By Toponogov’s theorem (see [Kl]), $\gamma' \geq \gamma$. The sum of the angles of the triangle in $S^2$ deviates from $180^\circ$ by the area of the triangle, which is bounded above by $L^2$ (see [BK, 6.7.1]). The same holds for the corresponding triangle in standard hyperbolic space $H^2$. Hence, using [BK, 6.4.3], we obtain $\gamma' - \gamma \leq 2L^2$. So $\gamma' \in [\pi/2 - \tilde{\delta}, \pi/2 + \tilde{\delta} + 2L^2]$. Altogether this gives
\[
d(A, B) \geq \sin(d(A, B)) \geq \sin(d(C, B)) \sin(\gamma') \geq \sin(L) \sin(\pi/2 + \tilde{\delta} + 2L^2). \quad \Box
\]

Now we want to apply Lemma 6.1 to our case of interest:

**Lemma 6.2.** Let $C \in N$ and $B = \exp_C(w)$ for some $w \in \text{Vert}_G^C$ of length $L < 0.05$, and assume as usual $\epsilon < 1/20000$. Then
\[
d(B, N) \geq \sin(L) \cos(\delta(\epsilon, 2L)) + 2L^2) = R^\epsilon_L.
\]
Here the function $\delta$ is as in Section 5.

**Proof.** Let $A$ be the closest point in $N$ to $B$. Clearly $d(A, B) \leq d(C, B) = L$, so the shortest geodesic $\gamma$ from $C$ to $A$ has length $L(\gamma) \leq 2L$. By Proposition 5.3 we have
\[
\angle(\dot{\gamma}(0), \text{Vert}_G^C) \geq \pi/2 - \delta(\epsilon, L(\gamma)) \geq \pi/2 - \delta(\epsilon, 2L).
\]
So, since $w \in \text{Vert}_G^C$,
\[
\angle(\dot{\gamma}(0), w) \in [\pi/2 - \delta(\epsilon, 2L), \pi/2 + \delta(\epsilon, 2L)].
\]
If we use the fact that, for any $g \in G$, the triangle $ABC$ lies in $\exp_{N_g}(\nu N_g)$, the lemma follows from Lemma 6.1 with $\tilde{\delta} = \delta(\epsilon, 2L)$. $\Box$

**Proof of Proposition 6.2.** For any $g \in G$ and positive number $L < 0.05$, by Lemma 6.2 each point $B \in \partial \exp_N(\text{Vert}^G)_L$ has distance at least
\[
\sin(L) \cos(\delta(\epsilon, 2L)) + 2L^2) = R^\epsilon_L
\]
from $N$. Therefore tub($R^\epsilon_L$) lies in $\exp_N(\text{Vert}^G)_L$, and since this holds for all $g$ we are done. $\Box$
7. Conclusion of the proof of the Main Theorem

In Sections 3–6, making use of the Riemannian structure of $M$, we showed that the 2-form $\int g \omega_g$ is well defined in the neighborhood $\bigcap_{g \in G} \exp_N(Vert_g)_L$ of $N$ (recall that $\omega_g := (\varphi_g^{-1})^* \omega$ was defined in the introduction). In this section we will focus on the symplectic structure of $M$ and conclude the proof of the Main Theorem, as outlined in Part II of Section 1.3.

First we will show that $\int g \omega_g$ is a symplectic form on a suitably defined neighborhood of $N$. Then it will easily follow that the convex linear combination $\omega_t := \omega + t(\int g \omega_g - \omega)$ is a symplectic form for all $t \in [0, 1]$.

As we saw in the introduction, $[\omega] = [\int g \omega_g] \in H^2(\text{tub}^\epsilon, \mathbb{R})$, so we can apply Moser’s trick. The main step consists of constructing canonically a primitive $\alpha$ of small maximum norm for the 2-form $\frac{d}{dt} \omega_t$. Comparing the size of the resulting Moser vector field with the size of $\text{tub}^\epsilon$ we will determine an $\epsilon$ for which the existence of an isotropic average of the $N_g$’s is ensured.

In this section we require $L < 0.05$. Notice that the estimates of Section 3 hold for such $L$. We start by requiring $\epsilon < 1/20000$ and introduce the abbreviation

$$D^\epsilon_L := 4L + 4100\epsilon$$

for the upper bound obtained in Proposition 3.4 on $\exp_N(Vert^\epsilon)_L$.

7.1. Symplectic forms in $\text{tub}^\epsilon$

In Section 3 we estimated the difference between $\varphi_{g*}X$ and $\|X$. This lemma does the same for $\varphi_g^{-1}$.

**Lemma 7.1.** Let $q \in \partial \exp_N(Vert^\epsilon)_L$ and $X \in T_qM$ a unit vector. Then

$$|(\varphi_g^{-1})_* X - \|X| \leq \frac{D^\epsilon_L}{1 - D^\epsilon_L}.$$

Furthermore,

$$\frac{1}{1 + D^\epsilon_L} \leq |(\varphi_g^{-1})_* X| \leq \frac{1}{1 - D^\epsilon_L}.$$

**Proof.** Let $p := \varphi_g^{-1}(q)$. By Proposition 3.4 for any vector $Z \in T_pM$ we have

$$\frac{|\varphi_{g*}(Z)|}{1 + D^\epsilon_L} \leq |Z| \leq \frac{|\varphi_{g*}(Z)|}{1 - D^\epsilon_L}.$$

The second statement of the lemma follows by setting $Z = (\varphi_g^{-1})_* X$.

Choosing instead $Z = (\varphi_g^{-1})_* X - \|X \in T_pM$ and applying once more Proposition 3.4 gives

$$|(\varphi_g^{-1})_* X - \|X| \leq \frac{|X - \varphi_{g*}\|X|}{1 - D^\epsilon_L} \leq \frac{D^\epsilon_L}{1 - D^\epsilon_L}.$$  \[\square\]
Since \((\varphi_g^{-1})_*X\) is close to \(\|X\|\) and since our assumption on \(\nabla \omega\) allows us to control to what extent \(\omega\) is invariant under parallel translation we are able to show that \(\omega\) and \(\omega_g = (\varphi_g^{-1})^* \omega\) are close to each other:

**Lemma 7.2.** Let \(X, Y\) be unit tangent vectors at \(q \in \exp_N(\text{Vert}^g)_L\). Then

\[
|((\omega_g - \omega)(X, Y)| \leq \frac{D_L^e}{1 - D_L^e} \left( \frac{D_L^e}{1 - D_L^e} + 2 \right) + 2L + 100\epsilon.
\]

**Proof.** Setting \(p := \varphi_g^{-1}(q)\) we have

\[
(\omega_g - \omega)_q(X, Y) = \omega_p((\varphi_g^{-1})_*X, (\varphi_g^{-1})_*Y) - \omega_q(X, Y)
\]

\[
= \omega_p(\|X - [(\varphi_g^{-1})_*X - \|X\|], \|Y - [(\varphi_g^{-1})_*Y - \|Y\|)\| - \omega_q(X, Y)
\]

\[
= \omega_p((\varphi_g^{-1})_*X - \|X\|, (\varphi_g^{-1})_*Y - \|Y\|
\]

\[
+ \omega_p(\|X, (\varphi_g^{-1})_*Y - \|Y\|) + \omega_p((\varphi_g^{-1})_*X - \|X\|, \|Y\|
\]

\[
+ \omega_p(\|X, \|Y\| - \omega_q(X, Y).
\]

Now since \(\|\\cdot\|\) is the parallel translation along a curve of length < \(2L + 100\epsilon\) (see Section 3) and \(|\nabla \omega| < 1\) we have \(\omega_p(\|X, \|Y\| - \omega_q(X, Y) < 2L + 100\epsilon\) and using Lemma 7.1 we are done. \(\square\)

Since the symplectic form \(\omega\) is compatible with the metric and the \(\omega_g\)'s are close to \(\omega\) we obtain the nondegeneracy of \(\omega_t\) for \(L\) and \(\epsilon\) small enough.

**Corollary 7.1.** Let \(X\) be a unit tangent vector at \(q \in \bigcap_{g \in G} \exp_N(\text{Vert}^g)_L\). Then for all \(t \in [0, 1]\),

\[
\omega_t(X, IX) \geq 1 - \left[ \frac{D_L^e}{1 - D_L^e} \left( \frac{D_L^e}{1 - D_L^e} + 2 \right) + 2L + 100\epsilon \right].
\]

**Proof.** By definition

\[
\omega_t(X, IX) = \omega(X, IX) + t \cdot \int_g (\omega_g - \omega)(X, IX).
\]

The first term is equal to 1 because \(\omega\) is almost-Kähler, the norm of the second one is estimated using Lemma 7.2. \(\square\)

**Remark.** The right hand side of Corollary 7.1 is surely positive if \(D_L^e \leq 0.1\). We set

\[
L^e := \frac{0.1 - 4100\epsilon}{4}
\]

and require \(\epsilon < 1/70000\). We obtain immediately:

\[\text{This choice of } L^e \text{ will allow us to obtain good numerical estimates in Section 7.4}\]
Proposition 7.1. On
\[ \text{tub}^\varepsilon := \exp_N(\nu N)_{R^*_\varepsilon} \subset \bigcap_{g \in G} \exp_N(\text{Vert}^\varepsilon)_L \]
the convex linear combination \[ \omega_t := \omega + t(\int_g \omega_g - \omega) \]
is a symplectic form for all \( t \in [0, 1] \).

Remark. Recall that the function \( R^*_\varepsilon \) was defined in Proposition 6.2. See Section 7.4 for the graph of \( R^*_\varepsilon \), a function of \( \varepsilon \).

7.2. The construction of the primitive of \( \frac{d}{dt} \omega_t \)

We want to construct canonically a primitive \( \alpha \) of
\[ \frac{d}{dt} \omega_t = \int_g \omega_g - \omega \]
on \( \bigcap_{g \in G} \exp_N(\text{Vert}^\varepsilon)_{0.05} \). We first recall the following fact, which is a slight modification of [Ca, Chapter III].

Let \( N \) be a submanifold of a Riemannian manifold \( M \), and let \( E \rightarrow N \) be a subbundle of \( TM|_N \rightarrow N \) such that \( E \oplus TN = TM|_N \). Furthermore let \( \bar{U} \) be a fiber-wise convex neighborhood of the zero section of \( E \rightarrow N \) such that \( \exp : \bar{U} \rightarrow U \subset M \) is a diffeomorphism. Denote by \( \pi : U \rightarrow N \) the projection along the slices given by exponentiating the fibers of \( E \), and by \( i : N \hookrightarrow M \) the inclusion. Then there is an operator \( Q : \Omega^*(U) \rightarrow \Omega^{*-1}(U) \) such that
\[ \text{Id} - (i \circ \pi)^* = dQ + Qd : \Omega^*(U) \rightarrow \Omega^*(U). \]
A concrete example is given by considering \( \rho_t : U \rightarrow U, \exp_q(v) \mapsto \exp_q(tv) \), and \( w_t|_{\rho_t(p)} := \frac{d}{dt}|_{t=0} \rho_t(p) \). Then
\[ Qf := \int_0^1 Q_t f \, dt, \quad Q_t f := \rho_t^*(i_{w_t} f), \]
gives an operator with the above property.

Note that for a 2-form \( \omega \) evaluated at \( X \in T_pM \) we have
\[ |(Q_t \omega)_p X| = |\omega_{\rho_t(w_t|_{\rho_t(p)})} \rho_t \omega_p(X)| \leq |\hat{\omega}_p|_{\text{op}} \cdot d(p, \pi(p)) \cdot |\rho_t \omega_p(X)| \quad (\star) \]
where \( |\hat{\omega}_p|_{\text{op}} \) is the operator norm of \( \hat{\omega}_p : T_pM \rightarrow T_p^*M \) and the inner product on \( T_p^*M \) is induced by the one on \( T_pM \).

For each \( g \) in \( G \) we want to construct a canonical primitive \( \alpha \) of \( \omega_g - \omega \) on \( \exp_N(\text{Vert}^\varepsilon)_{0.05} \). We do that in two steps:

Step I. We apply the above procedure to the vector bundle \( \text{Vert}^\varepsilon \rightarrow N \) to obtain an operator \( Q^\varepsilon_N \) such that
\[ \text{Id} - (\pi_N^\varepsilon)^* \circ (i_N)^* = dQ^\varepsilon_N + Q^\varepsilon_N d \]
for all differential forms on $\exp_N(\text{Vert}^g)_{0.05}$. Since $N$ is isotropic with respect to $\omega_g$ and $\omega_g - \omega$ is closed we have

$$\omega_g - \omega = dQ_N^g(\omega_g - \omega) + (\pi_N^g)^*(iN)^*(-\omega).$$

**Step II.** Now we apply the procedure to the vector bundle $vN_g \to N_g$ to get an operator $Q_N$ on differential forms on $\exp_N(vN_g)_{100e}$. Since $N_g$ is isotropic with respect to $\omega$ we have

$$\omega = dQ_N\omega,$$

so we have found a primitive of $\omega$ on $\exp_N(vN_g)_{100e}$. Since $N \subset \exp_N(vN_g)_{100e}$ the 1-form $\beta^g := i_N^g(Q_N\omega)$ on $N$ is a well defined primitive of $i_N^g\omega$.

Summing up these two steps we see that $\alpha^g := Q^g_N(\omega_g - \omega) - (\pi_N^g)^*\beta^g$ is a primitive of $\omega_g - \omega$ on $\exp_N(\text{Vert}^g)_{0.05}$. So clearly $\alpha := \int_N \alpha^g$ is a primitive of $d\int dt\omega_t$ on $\bigcap_{g \in G} \exp_N(\text{Vert}^g)_{0.05}$.

### 7.3. Estimates on the primitive of $d\int dt\omega_t$

In this section we will estimate the $C^0$-norm of the 1-form $\alpha$ constructed in Section 7.2.

**Step II.** We will first estimate the norm of $\beta^g := i_N^g(Q_N\omega)$ using (1) and then the norm of $(\pi_N^g)^*\beta^g$.

**Lemma 7.3.** If $p \in \exp_N(vN_g)_{100e}$, and $X \in T_pM$ is a unit vector, then for any $t \in [0, 1]$,

$$|\rho_{N_g}t_sX| \leq 5/4.$$

**Proof.** Let $L := d(p, N_g) < 100e < 1/700$ and write $X = J(t) + K(t)$, where $J$ and $K$ are $N_g$-Jacobi fields along the unit-speed geodesic $\gamma_p$ from $p' := \pi_{N_g}(p)$ to $p$ such that $J(0)$ vanishes, $J'(0)$ is normal to $N_g$, and $K$ is a strong $N_g$-Jacobi field (see the remark in Section 3.2). $J(t)$ is the variational vector field of a variation $f_X(t) = \exp_{p'}(tv(s))$ where the $v(s)$'s are unit normal vectors at $p'$. Therefore

$$(\rho_{N_g})_{\gamma_p}J(L) = \frac{d}{ds} \big|_{0}[(\rho_{N_g})_{\gamma_p} \exp_{p'}(Lv(s))\big] = \frac{d}{ds} \big|_{0}[\exp_{p'}(tLv(s))] = J(tL).$$

Using Lemma 3.4 we have on the one hand $|LJ'(0)| \leq (1 + \sigma(L)|J(L)|$ and on the other hand $|J(tL)|(1 - \sigma(tL)) \leq tL|J'(0)|$ where $\sigma(x) := (\sinh(x) - x)/\sin(x)$. So

$$|J(tL)| \leq \frac{1 + \sigma(L)}{1 - \sigma(tL)}|J(L)| \leq \frac{21}{20}t|J(L)|.$$
Similarly we have \((\rho_{N_g})_{\ast}K(L) = K(tL)\). Using Lemma 3.6 we deduce that \(|K(0)| \leq (1 + \frac{9}{5}L)|K(L)|\) and \(|K(tL)| \leq |K(0)|/(1 - \frac{9}{5}tL)\), therefore
\[
|K(tL)| \leq \frac{1 + \frac{9}{5}L}{1 - \frac{9}{5}tL}|K(L)| \leq \frac{21}{20}|K(L)|.
\]
Altogether we have
\[
|(\rho_{N_g})_{\ast}X| = |J(tL) + K(tL)|^2 \leq |J(tL)|^2 + |K(tL)|^2 + 2 \cdot \frac{9}{5}L|J(tL)||K(tL)|
\]
\[
\leq \left(\frac{21}{20}\right)^2 \left[|J(L)|^2 + |K(L)|^2 + \frac{18}{5}L|J(L)||K(L)|\right]
\]
\[
\leq \left(\frac{21}{20}\right)^2 \left[|J(L) + K(L)|^2 + \frac{36}{5}L|J(L)||K(L)|\right]
\]
\[
\leq \left(\frac{21}{20}\right)^2 \left[1 + \frac{36}{5} \cdot 1.1^2 L\right] \leq \frac{5}{4}
\]
where in the first and third inequalities we used Lemma 3.9 and in the fourth in addition Lemma 3.10.

Corollary 7.2. The 1-form \(\beta^g\) on \(N\) satisfies
\[
|\beta^g| < 125\epsilon.
\]

Proof. At any point \(p \in N \subset \exp_{N_g}(\nu_{N_g})_{100\epsilon}\), using (⋆), the fact that \(|\tilde{\omega}|_{op} = 1\) and Lemma 7.3 we have \(|(Q_{N_g}\omega)_p| < 125\epsilon\). Clearly
\[
|(Q_{N_g}\omega)_p| \geq |i^g_N(Q_{N_g}\omega)_p| = |\beta^g_p|.
\]

Now we would like to estimate \((\pi^g_N)_\ast X\) for a unit tangent vector \(X\). Since \(\pi^g_N = (\rho_{N_g})_0\) we prove a stronger statement that will be used again later. Recall that we assume \(L \leq 0.05\).

Lemma 7.4. If \(q \in \exp_{N_g}(\text{Vert}^g)_L\) and \(X \in T_qM\) is a unit vector then for any \(t \in [0, 1]\) we have
\[
|(\rho^g_{N_g})_{\ast}X| \leq 1.5 \frac{1 + D^g_L}{1 - D^g_L}.
\]

Proof. Using Lemma 7.1 we have
\[
|(\rho^g_{N_g})_{\ast}X| \leq (1 + D^g_L)|(\varphi^{-1})_\ast(\rho^g_{N_g})_{\ast}X|.
\]

Clearly \((\rho^g_{N_g})_\ast \varphi = \varphi \circ (\rho^g_{N_g})_\ast\), since—up to exponentiating—\(\varphi\) maps \(\nu_{N_g}\) to \(\text{Vert}^g\), and \((\rho^g_{N_g})_\ast\) and \((\rho^g_{N_g})_\ast\) are just rescaling of the respective fibers by a factor of \(t\).
If we reproduce the proof of Lemma 7.3 requiring \( p \) to lie in \( \exp_N(vN_p) \), we obtain: \(|(\rho_{N_p})_*Y| < 1.5\) for unit vectors \( Y \) at \( p \). Using this and Lemma 7.1 respectively we have
\[
|\rho_{N_p})_* (\varphi_g^{-1})_* X | \leq 1.5| (\varphi_g^{-1})_* X | \quad \text{and} \quad |(\varphi_g^{-1})_* X | \leq \frac{1}{1 - D_L^e}.
\]
Altogether this proves the lemma. \( \square \)

**Corollary 7.3.** On \( \exp_N(\text{Vert}^g) \) we have
\[
| (\pi^g_N)^* \beta | \leq 200 \epsilon + D_L^e \cdot 3 \frac{1}{1 - D_L^e - D_L^e} + 2L + 100 \epsilon.
\]

**Proof.** This is clear from the equality \( |((\pi^g_N)^* \beta)| = |\beta| \), Corollary 7.2 and Lemma 7.4. \( \square \)

**Step I.** Now we estimate \(|Q_N^g (\omega - \omega)|\). This is easily achieved using Lemmata 7.2 and 7.4 to estimate the quantities involved in (\( \spadesuit \)):

**Corollary 7.4.** For \( q \in \text{exp}_N(\text{Vert}^g) \) we have
\[
|Q_N^g (\omega - \omega)_q | \leq 1.5 \frac{1 + D_L^e}{1 - D_L^e} \cdot L \cdot \left[ \frac{D_L^e}{1 - D_L^e} \left( \frac{D_L^e}{1 - D_L^e} + 2 \right) + 2L + 100 \epsilon \right].
\]

**Remark.** By Proposition 5.2, \( d(q, N) \geq R \epsilon L \). Furthermore, when \( \epsilon < 1/70000 \) and \( L < 0.05 \), one can show that \( R_L^e \geq 2 \). So \( L \leq \frac{2}{3} d(q, N) \).

Now finally using Corollaries 7.3 and 7.4 we can estimate the norm of \( \alpha := \int_g \alpha^g \):

**Proposition 7.2.** Assuming \( L < 0.05 \) at \( q \in \text{exp}_N(\text{Vert}^g) \) we have
\[
|\alpha_q | \leq 1.5 \frac{1 + D_L^e}{1 - D_L^e} \cdot \frac{3}{2} d(q, N) \cdot \left[ \frac{D_L^e}{1 - D_L^e} \left( \frac{D_L^e}{1 - D_L^e} + 2 \right) + 2L + 100 \epsilon \right] + 200 \epsilon + \frac{1}{1 - D_L^e}.
\]

7.4. The end of the proof of the Main Theorem

Proposition 7.1 showed that the Moser vector field \( v_t := -\tilde{\omega}_t^{-1} \alpha \) is well defined on \( \text{tub}^g \subset \bigcap_{g \in G} \exp_N(\text{Vert}^g) \). Recalling that \( D_L^{e^g} = 0.1 \), Corollary 7.1 immediately implies

**Corollary 7.5.** At \( q \in \text{tub}^g \subset \bigcap_{g \in G} \exp_N(\text{Vert}^g) \) we have
\[
| (\tilde{\omega}_t)_q |^1_{\text{op}} \leq \frac{1}{1 - \left[ \frac{D_L^e}{1 - D_L^e} \left( \frac{D_L^e}{1 - D_L^e} + 2 \right) + 2L + 100 \epsilon \right]} \leq 1.53.
\]

\( ^1 \) Since \( L < 0.05 \) now we have to replace the constant 21/20 in that proof by the constant 6/5.
From Corollary 7.5 and Proposition 7.2 we obtain:

**Proposition 7.3.** For all \( t \in [0, 1] \) and \( q \in \text{tub}^\epsilon \),

\[
| (v_t)_q | \leq | (\tilde{\omega}_t)_q | \leq 1.45 d(q, N) + 374 \epsilon.
\]

Let \( \gamma(t) \) be an integral curve of the time-dependent vector field \( v_t \) on \( \text{tub}^\epsilon \) such that \( p := \gamma(0) \in N \). Where \( d(\cdot, p) \) is differentiable, its gradient has unit length. So \( \frac{d}{dt} d(\gamma(t), p) \leq \frac{d}{dt} | \dot{\gamma}(t) | \).

By Proposition 7.3 we have \( | \dot{\gamma}(t) | \leq 1.45 d(\gamma(t), p) + 374 \epsilon \). So altogether

\[
\frac{d}{dt} d(\gamma(t), p) \leq 1.45 d(\gamma(t), p) + 374 \epsilon.
\]

The solution of the ODE \( \dot{s}(t) = As(t) + B \) satisfying \( s(0) = 0 \) is \( \frac{B}{A} (e^{At} - 1) \). Hence, if the integral curve \( \gamma \) is well defined at time 1, we have

\[
d(\gamma(1), N) \leq 1.45 d(\gamma(1), p) \leq 374 \epsilon (e^{1.45} - 1) \leq 842 \epsilon.
\]

Let us denote by \( \rho_1 \) the time-1 flow of the time-dependent vector field \( v_t \), so that \( \rho_1^{-1} \) is the time-1 flow of \( -v_{1-t} \). Since by definition \( \text{tub}^\epsilon := \exp_N(vN)R_{\epsilon}^t \), the submanifold \( L := \rho_1^{-1}(N) \) will surely be well defined if

\[
842 \epsilon < R_{L,\epsilon}^t.
\]

This is always the case since \( \epsilon < 1/70000 \).

![Graphs of 842\epsilon (increasing) and R_{L,\epsilon}^t (decreasing).](image)

The estimate for \( d_0(N_g, L) \) is obtained by using \( d_0(N, L) < 842 \epsilon \) and \( d_0(N_g, N) < 100 \epsilon \). The proof of the Main Theorem is now complete.

**Remark.** In the Main Theorem we assumed that \(|\nabla \omega| < 1\). Let us now consider the case that \(|\nabla \omega| \geq 1\). Then the statement of the Main Theorem still holds verbatim if one makes...
the bound on $\epsilon$ smaller, as follows. The bound on $|\nabla \omega|$ enters our proof directly only in Lemma 7.2; if $|\nabla \omega| \geq 1$, the inequality of that lemma should read

$$|(\omega_g - \omega)(X, Y)| \leq \frac{D_L^\epsilon}{1 - D_L^\epsilon} \left( \frac{D_L^\epsilon}{1 - D_L^\epsilon} + 2 \right) + |\nabla \omega|(2L + 100\epsilon)$$

instead. Similarly, the quantity $2L + 100\epsilon$ appearing in Corollary 7.1, Corollary 7.4 and Proposition 7.2 should be multiplied by $|\nabla \omega|$. Now assume that $\epsilon < \frac{1}{|\nabla \omega| 70000}$ and replace $L^\epsilon$ everywhere by

$$\tilde{L}^\epsilon := \frac{0.1/|\nabla \omega| - 4100\epsilon}{4}.$$ 

Then the bounds on $|(\tilde{\omega}_t)^{-1}|_{op}$ and $|v_t|_g$ given in Corollary 7.5 and Proposition 7.3 still hold, and our isotropic average $L$ will be well defined if $842\epsilon < R^\epsilon_{\tilde{L}^\epsilon}$. This is satisfied for $\epsilon$ small enough, since $R^\epsilon_{\tilde{L}^\epsilon}$ is a continuous function and $R^0_{\tilde{L}^0}$ is positive.

8. Remarks on the Main Theorem

Remark 1 (Is the isotropic average $L$ $C^1$-close to the $N_g$’s?). The main shortcoming of our Main Theorem is surely the lack of an estimate on the $C^1$-distance $d_1(N_g, L)$.

To bound $d_1(N_g, L)$ it is enough to estimate the distance between the tangent spaces $T_pL$ and $T_{p_1(p)}N$. Indeed, this would allow us to estimate the distance between $T_pL$ and $T_{\pi_{N_g}(p)}N_g$, using which—when $\epsilon$ is small enough—one can conclude that $\pi_{N_g} : L \rightarrow N_g$ is a diffeomorphism and give the desired bound on the $C^1$-distance.

Using local coordinates and standard theorems about ODEs it is possible to estimate the distance between $T_pL$ and $T_{p_1(p)}N$ provided one has a bound on the covariant derivative of the Moser vector field, for which one would have to estimate $\nabla (\tilde{\omega}_t)^{-1}$. To do that one should be able to bound expressions like $\nabla_Y ((\phi_t)^{-1}X)$ for parallel vector fields $X$ along some curve.

This does not seem to be possible without more information on the extrinsic geometry of $N$. We recall that it is not known whether the average $N$ forms a gentle pair with $M$ (see Remark 6.1 in [We]). We are currently trying to improve Weinstein’s theorem so that one obtains a gentle average.

Remark 2 (The case of isotropic $N$). Unfortunately, if the Weinstein average $N$ happens to be already isotropic with respect to $\omega$, our construction will generally provide an isotropic average $L$ different from $N$. Indeed, while Step I of Section 7.2 always gives a 1-form vanishing at points of $N$, Step II does not, even if $N$ is isotropic for $\omega$.

The procedure outlined in Remark 3, on the other hand, would produce $N$ as the isotropic average, but in that case the upper bound for $\epsilon$ would depend on the geometry of $N$. 

Submanifold averaging 111
Remark 3 (Averaging of symplectic and coisotropic submanifolds). The averaging of \( C^1 \)-close gentle symplectic submanifolds of an almost-Kähler manifold is a much simpler task than for isotropic submanifolds. The reason is that \( C^1 \)-small perturbations of symplectic manifolds are symplectic again and one can simply apply Weinstein’s averaging procedure ([We, Thm. 2.3]).

Unfortunately our construction does not allow averaging coisotropic submanifolds. In our proof we were able to canonically construct a primitive of \( \int g \omega_g - \omega \) using the fact that the \( N_g \)'s are isotropic with respect to \( \omega \). If they are not, it is still possible to construct canonically a primitive, following Step I of our construction and making use of the primitive \( d^* (\Delta^{-1} i^*_N (\omega_g - \omega)) \) of \( i^*_N (\omega_g - \omega) \) (but the upper bound on its norm would depend on the geometry of \( N \)).

Nevertheless, our construction fails in the coisotropic case, since the fact that \( N \) is coisotropic for all \( \omega_g \)'s does not imply that it is for their average \( \int g \omega_g \).

9. An application to Hamiltonian actions

As a simple application of our Main Theorem we apply Theorem 2 to almost invariant isotropic submanifolds of a Hamiltonian \( G \)-space and deduce some information about their images under the moment map.

We start by recalling some basic definitions (see [Ca]): consider an action of a Lie group \( G \) on a symplectic manifold \((M, \omega)\) by symplectomorphisms. A moment map for the action is a map \( J : M \to g^* \) such that for all \( v \in g \) we have \( \omega(v_M(v), \cdot) = d\langle J, v \rangle \) and which is equivariant with respect to the \( G \)-action on \( M \) and the coadjoint action of \( G \) on \( g \). Here \( v_M \) is the vector field on \( M \) given by \( v \) via the infinitesimal action. An action admitting a moment map is called the Hamiltonian action.

This simple lemma is a counterpart to [Ch, Prop. 1.3].

Lemma 9.1. Let the compact connected Lie group \( G \) act on the symplectic manifold \((M, \omega)\) with moment map \( J \). Let \( L \) be a connected isotropic submanifold of \((M, \omega)\) which is invariant under the group action. Then \( L \subset J^{-1}(\mu) \) where \( \mu \) is a fixed point of the coadjoint action.

Proof. Let \( X \in T_x L \). For each \( v \in g \) we have
\[
d_x (J, v) X = \omega(v_M(x), X) = 0,
\]
since both \( v_M(x) \) and \( X \) are tangent to the isotropic submanifold \( L \). Therefore every component of the moment map is constant along \( L \), so \( L \subset J^{-1}(\mu) \) for some \( \mu \in g^* \).

Now let \( x_0 \in L \) and let \( G \cdot x_0 \subseteq L \) be the orbit through \( x_0 \). Then from the equivariance of \( J \) it follows that for all \( g \) we have \( \mu = J(g \cdot x_0) = g \cdot J(x_0) = g \cdot \mu \), so \( \mu \) is a fixed point of the coadjoint action.

Now we apply the lemma above to the case where \( L \) is almost invariant.
Corollary 9.1. Let the compact Lie group $G$ act on the symplectic manifold $(M, \omega)$ with moment map $J : M \to \mathfrak{g}^*$. Suppose $M$ is endowed with a $G$-invariant compatible Riemannian metric so that the Levi-Civita connection satisfies $|\nabla \omega| < 1$. If a connected isotropic submanifold $L \subset M$ satisfies:

(i) $(M, L)$ is a gentle pair,
(ii) $d_1(L, g \cdot L) < \epsilon < 1/70000$ for all $g \in G$,

then $J(L)$ lies in the ball of radius $1000\epsilon \cdot C$ about a fixed point $\mu$ of the coadjoint action. Here $\mathfrak{g}^*$ is endowed with any inner product and $C := \max\{|v_M| : v \in \mathfrak{g} \text{ has unit length}\}$.

Proof. By Theorem 2 there exists an isotropic submanifold $L'$ invariant under the $G$-action with $d_0(L, L') < 1000\epsilon$. By Lemma 9.1, $L$ lies in some fiber $J^{-1}(\mu)$ of the moment map, where $\mu$ is a fixed point of the coadjoint action. We will show that $|J(p) - \mu| < 1000\epsilon \cdot C$ for all $p \in L$.

Let $p'$ a closest point to $p$ in $L'$. The shortest geodesic $\gamma$ from $p$ to $p'$, which we choose to be defined on the interval $[0, 1]$, has length $< 1000\epsilon$. Therefore for any unit-length $v \in \mathfrak{g}$ (with respect to the inner product induced on $\mathfrak{g}$ by its dual) we have

$$\langle J(p) - \mu, v \rangle = \int_0^1 \langle dJ(\gamma(t))\gamma'(t), v \rangle \, dt = \int_0^1 d\langle J, v \rangle \gamma'(t) \, dt = \int_0^1 \omega(v_M, \gamma'(t)) \, dt.$$ 

Since for all $t$ we have $|\omega(v_M, \gamma'(t))| \leq |v_M| \cdot |\gamma'(t)| \leq 1000\epsilon \cdot C$ we are done. \qed

A. The estimates of Proposition 4.1

Here we will prove the estimates used in the proof of Proposition 4.1. See Section 4 for the notation.

We first state a general fact about the exponential map:

Lemma A.1. If $\gamma$ is a geodesic parametrized by arc length and $W \in T_{\gamma(0)}M$, then for $t < 0.7$, $|\gamma'(d_t\gamma(0)) W - W| \leq \frac{\sinh(t) - t}{t} |W| \leq \frac{t^2}{5} |W|$ and $\left(1 - \frac{t^2}{5}\right)|W| \leq |\exp_W| \leq \left(1 + \frac{t^2}{5}\right)|W|$. 

Proof. The unique Jacobi field along $\gamma$ such that $J(0) = 0$ and $J'(0) = W$ is given by $J(t) = (d_t\gamma(0) \exp_{\gamma(0)})(tW)$ (see [Jo, Cor. 4.2.2]). The bound $(\sinh(t) - t)/t$ follows from [BK] 6.3.8(iii). This expression is bounded above by $t^2/5$ when $t < 0.7$. The second estimate follows trivially from the first one. We prefer to use these estimates rather than more standard ones (see [BK] 6.4.1) in order to keep the form of later estimates more concise. \qed
Corollary A.1.

\[ \angle(Q - 0, cA\dot{\|}C B) < \frac{1}{2}d(C, A). \]

**Proof.** By [BK, 6.6.1] (choosing \( v = \tilde{C} - 0 \) and \( w = cA\dot{\|}C B \)) we get

\[
d(\exp_A((\tilde{C} - 0) + cA\dot{\|}C B), \exp_C \dot{C} B = B) \leq \frac{1}{3}d(A, C) \cdot d(C, B) \cdot \sinh(d(A, C) + d(C, B)) \cdot \sin(\angle(\tilde{C} - 0, cA\dot{\|}C B))
\]

\[
\leq \frac{1}{3}d(A, C)d(C, B)
\]

using \( d(A, C) < 0.15 \) and \( d(C, B) < 0.5 \).

In order to estimate distances in \( T_{AM} \) (instead of in \( M \)) we denote the shortest geodesic from \( \exp_A((\tilde{C} - 0) + cA\dot{\|}C B) \) to \( B \) by \( \tau \), and by \( \tilde{\tau} \) its image under \( \exp^{-1}_A \).

By Lemma A.1,

\[
|\dot{\tilde{\tau}}(s)| \leq \left( 1 - \frac{d(\tau(s), A)^2}{5} \right) |\dot{\tau}(s)| \leq |\dot{\tau}(s)|
\]

(using \( d(\tau(s), A) < 0.7 \) in the first inequality), so

\[
d(\exp_A((\tilde{C} - 0) + cA\dot{\|}C B), B) \geq \frac{9}{10} |(\tilde{C} - 0) + cA\dot{\|}C B - (\tilde{B} - 0)|
\]

\[
= \frac{9}{10} |cA\dot{\|}C B - (\tilde{B} - \tilde{C})|.
\]

Altogether, since \( \tilde{B} - \tilde{C} = Q - 0 \), we obtain

\[
|cA\dot{\|}C B - (Q - 0)| \leq \frac{2}{5}d(A, C)d(C, B),
\]

and using \( \frac{\sin x}{x} \geq \frac{7}{8} \) for \( x \in [0, 0.8] \), we obtain

\[
\angle(cA\dot{\|}C B, Q - 0) \leq \frac{8}{7} \sin(\angle(cA\dot{\|}C B, Q - 0)) \leq \frac{1}{2}d(C, A).
\]

\( \square \)

Corollary A.2.

\[ \angle(P - 0, P' - 0) \leq \frac{1}{2}d(C, A). \]

**Proof.** We first want to bound \( |P' - P| \) from above and \( |P' - 0| \) from below. Since \( P' \) and \( P \) are the closest points in \( P_A \) to \( cA\dot{\|}C B \) and \( Q \) respectively,

\[
|P - P'| \leq |(Q - 0) - cA\dot{\|}C B| \leq \frac{2}{5}d(A, C)d(C, B),
\]

by the last estimate of the proof of Corollary A.1.

On the other hand, we have

\[
|P' - 0| = |cB| \cdot \cos(\angle(cA\dot{\|}C B, P_A)) \geq |cB| \cdot \cos(\theta)
\]

\[
\geq |cB| \sqrt{1 - \theta^2} \geq |cB| \sqrt{1 - C^2d(C, A)^2}.
\]
Therefore we have
\[
\sin(\angle(P' - 0, P - 0)) \leq \frac{|P' - P|}{|P' - 0|} \leq \frac{2}{5} \frac{1}{\sqrt{1 - C^2 d(C, A)^2}} d(C, A).
\]
So, using the restrictions $C \leq 2$ and $d(C, A) < 0.15$, and using $\sin x \geq \frac{x}{2}$ for $x \in [0, 0.8]$, we obtain
\[
\angle(P' - 0, P' - P) \leq \frac{8}{7} \sin(\angle(P' - 0, P' - P)) \leq \frac{1}{2} d(C, A).
\]

We conclude this appendix by deriving the estimate needed in Claim 3 of Proposition 4.1.

**Corollary A.3.**
\[
|\tilde{B} - \tilde{C}| < \frac{11}{10} d(C, B).
\]

**Proof.** This follows by choosing a shortest geodesic between $C$ and $B$ and using Lemma A.1, exactly as we did in the proof of Corollary A.1. \(\square\)

**B. An upper bound for $\alpha$ using the curve $c$**

Here we will prove Proposition 5.1, namely the estimate
\[
|\exp_{N^g}^{-1}(C) - \frac{1}{f(r)}\exp_{N^g}^{-1}(A)| \leq L(\gamma) \frac{3150\epsilon}{f(r)}.
\]
To do so we will use the fact that $N$ is $C^1$-close to $N^g$ (see Lemma B.3).

In addition to the notation introduced in Section 5 to state the proposition, we will use the following. We will denote by $\pi_c(t)$ the curve $\pi_{N^g} \circ c(t)$, so $\pi_c$ is just a reparametrization of $\pi$. We will use $\exp$ as a short-hand notation for the normal exponential map $\exp_{N^g}: (vN^g) \rightarrow \exp_{N^g}(vN^g)$. Therefore $\tilde{c}(t) := \exp^{-1}(c(t))$ will be a section of $vN^g$ along $\pi_c$. The image under $\exp_c$ of the Ehresmann connection corresponding to $\nabla^\perp$ will be the subbundle $LC^g$ of $TM|_{\exp_{N^g}(vN^g)}$. To simplify notation we will denote by $pr_{\gamma(t)} \Hor^g$ the projection of $\dot{\gamma}(t) \in T_{\gamma(t)}M$ onto $\Vert^\perp_{\gamma(t)}$ along $\Hor^g_{\gamma(t)}$. We will also use $pr_{\gamma(t)} \ahor^g$ and $pr_{\gamma(t)} LC^g$ to denote projections onto $\ahor^g_{\gamma(t)}$, along $\ahor^g_{\gamma(t)}$ and $LC^g_{\gamma(t)}$ respectively.

Our strategy will be to bound above $|\nabla^\perp_{\frac{d}{dt}} c(t)| = |\exp_{\gamma(t)}^{-1}(pr_{\gamma(t)} LC^g)|$ (see Lemma B.3) using
\[
TN \approx \Hor^g \approx \ahor^g \approx LC^g.
\]
Integration along $\pi_c$ will give the desired estimate.

The estimates to make precise $TN \approx \Hor^g$ and $\Hor^g \approx \ahor^g$ were derived in [We].

In the next two lemmata we will do the same for $\ahor^g \approx LC^g$. 
Lemma B.1. If \( L < 0.08 \) and \( p \) is a point in \( \partial \exp_{N_p}(vN_g)L \), then
\[
d(a_{\text{Hor}}^g, L \xi_p^g) \leq \arcsin \left( \frac{9}{5} L \right).
\]

Proof. It is enough to show that, if \( Y \in L \xi_p^g \) is a unit vector, then
\[
|pr_Y a_{\text{Hor}}^g| \leq \frac{9}{5} L.
\]

Let \( \beta(s) \) be a curve tangent to the distribution \( L \xi_p^g \) such that \( \beta(0) = p \) and \( \dot{\beta}(0) = Y \). Then \( \exp^{-1}(\beta(s)) = L \xi(s) \) for a unit-length parallel section \( \xi \) of \( vN_p \) along the curve \( \gamma(s) := \pi_{N_p}(\beta(s)) \). If we denote by \( K \) the \( N_p \)-Jacobi field arising from the variation \( f_s(t) = \exp(t\xi(s)) \), then clearly \( K(L) = Y \) and \( K(0) = \dot{\gamma}(0) \).

We claim that \( \xi \) is a strong Jacobi field (see the remark in Section 3.2): we have
\[
\frac{\partial}{\partial s} |f_s(0)|_0 = |\nabla_{ds} |f_s(0)|_0| = \frac{\partial}{\partial t} |pr_{\gamma \xi(0)}Y| = |\nabla_{ds} |pr_{\gamma \xi(0)}Y| | = \frac{\partial}{\partial t} |pr_{\gamma \xi(0)}Y| = \frac{\partial}{\partial t} |pr_{\gamma \xi(0)}Y|,
\]
so
\[
K'(0) = \frac{\partial}{\partial t} |f_s(0)|_0 = \frac{\partial}{\partial s} |f_s(0)|_0.
\]

The claim follows since \( \xi(0) = \dot{\gamma}_p(0) \), where \( \gamma_p \) denotes the unique geodesic parameterized by arc length connecting \( \pi_{N_p}(p) \) to \( p \).

Now let us denote by \( J \) the \( N_p \)-Jacobi field along \( \gamma_p \) vanishing at 0 such that \( J(L) = pr_Y a_{\text{Hor}}^g \in a_{\text{Vert}}^g \). By Lemma 3.9, using the fact that \( Y \) is a unit vector, we have
\[
|pr_Y a_{\text{Hor}}^g|^2 = (pr_Y a_{\text{Hor}}^g, Y) = |(J(L), K(L))| \leq \frac{9}{5} L \cdot |pr_Y a_{\text{Hor}}^g|
\]
and we are done.

\( \Box \)

Lemma B.2. Let \( L < 0.08 \). For any point \( p \) in \( \partial \exp_{N_p}(vN_g)L \) the projections \( T_p M \to a_{\text{Vert}}^g \) along \( a_{\text{Hor}}^g \) and \( L \xi_p^g \) differ at most by \( 2L \) in the operator norm.

Proof. Let \( \phi : a_{\text{Hor}}^g \to a_{\text{Vert}}^g \) be the linear map whose graph is \( L \xi_p^g \). Let \( X \in T_p M \) be a unit vector and write \( X = X_{\text{ah}} + X_{\text{av}} \) for the decomposition of \( X \) into almost horizontal and almost vertical vectors. Then \( X = (X_{\text{ah}} + \phi(X_{\text{ah}})) + (X_{\text{av}} - \phi(X_{\text{ah}})) \) is the decomposition with respect to the subspaces \( L \xi_p^g \) and \( a_{\text{Vert}}^g \). The difference of the two projections onto \( a_{\text{Vert}}^g \) maps \( X \) to \( \phi(X_{\text{ah}}) \).

\[
|\phi(X_{\text{ah}})| \leq |\phi|_{op} \leq \frac{9}{5} L \leq \frac{\arcsin \left( \frac{9}{5} L \right)}{\sqrt{1 - \left( \frac{9}{5} L \right)^2}} < 2L,
\]
where we used [We, Cor. A.5] in the second inequality and Lemma B.1 in the third one.

\( \Box \)

Now we are ready to bound the covariant derivative of \( \tilde{c}(t) \):
Lemma B.3. For all $t$, 
\[
\left| \frac{\nabla_{\pi} \hat{c}(t)}{dt} \right| \leq 2702 \epsilon.
\]

Proof. Let 
\[
\frac{\nabla_{\pi} \hat{c}(t)}{dt} \in \nu_{\pi(t)} N_g
\]
denote the projection of $\hat{c}(t)$ under the projection $T_{\pi(t)}(\nu N_g) \to T_{\pi(t)}(\nu_{\pi(t)} N_g)$ along the Ehresmann connection on $\nu N_g$ corresponding to $\nabla$. Therefore, since $\exp_\ast$ maps the Ehresmann connection to $LC^g$ and tangent spaces to the fibers of $\nu N_g$ to $aVert^g$, we have 
\[
\exp_\ast \left( \frac{\nabla_{\pi} \hat{c}(t)}{dt} \right) = \text{pr}_{\pi(t)} LC^g.
\]
Notice that here $\exp_\ast$ denotes $d_{\tilde{c}(t)} \exp_{N_g}$.

The fact that $N$ is $C^1$-close to $N_g$ (see Theorem 4) implies $\angle(\hat{c}(t), Hor_{\pi(t)}^g) \leq 2500 \epsilon$ since $\hat{c}(t) \in T_{\pi(t)} N$. By [We, Prop. 3.7], $d(Hor_{\pi(t)}^g \ast Hor_{\pi(t)}^g) \leq \epsilon/2$ since $d(c(t), N_g) \leq 100 \epsilon$. Therefore, $\angle(\hat{c}(t), aHor_{\pi(t)}^g) \leq 2501 \epsilon$ and $|\text{pr}_{\pi(t)} aHor^g| \leq \sin(2501 \epsilon) \leq 2501 \epsilon$.

On the other hand, by Lemma B.2, $|\text{pr}_{\pi(t)} aHor^g - \text{pr}_{\pi(t)} LC^g| \leq 200 \epsilon$. The triangle inequality therefore gives $|\text{pr}_{\pi(t)} LC^g| \leq \sin(2701 \epsilon)$. Therefore, using Lemma A.1 and $\epsilon < 1/20000$, 
\[
|\exp_\ast^{-1}(\text{pr}_{\pi(t)} LC^g)| \leq \frac{1}{1 - \epsilon/5} |\text{pr}_{\pi(t)} LC^g| \leq 2702 \epsilon.
\]

Lemma B.3 allows us to bound $|\exp^{-1}(C) - \frac{1}{aN} \exp^{-1}(A)|$ in terms of $L(c)$. However, we want a bound in terms of $L(\gamma)$, so now we will compare the lengths of the two curves.

Recall that $f(x) = \cos(x) - \frac{3}{2} \sin(x)$ and $r := 100 \epsilon + L(\gamma)/2$. Notice also that $r < 0.08$ due to our restrictions on $\epsilon$ and $d(C, A)$.

Lemma B.4.

\[
L(c) \leq \frac{1 + 3200 \epsilon}{f(r)} L(\gamma).
\]

Proof. Since $\varphi_{\pi N_g}(\pi N_g \ast \hat{c}(t)) = \hat{c}(t)$, by Proposition 3.1 we have $|\hat{c}(t) - \| (\pi N_g) \ast \hat{c}(t) | \leq 3200 \epsilon |(\pi N_g) \ast \hat{c}(t)|$, so 
\[
|\hat{c}(t)| \leq (1 + 3200 \epsilon) |(\pi N_g) \ast \hat{c}(t)|.
\]

Since $L(\pi N_g \circ c) = L(\pi)$, it follows that $L(c) \leq (1 + 3200 \epsilon) L(\pi)$. By [We, Lemma 3.3] we have $f(r) L(\pi) \leq L(\gamma)$ and we are done. □
Proof of Proposition 5.1. We have

\[ |\exp^{-1}(C) - \exp^{-1}(A)| = \left| \int_0^{L(c)} \frac{d}{dt} \pi_b \tilde{c}(t) \, dt \right| \leq 2702\epsilon L(c) \leq 2702\epsilon \frac{1 + 3200\epsilon}{f(r)} L(\gamma) \]

where we used Lemmata B.3 and B.4 in the last two inequalities. The proposition follows by using the bound \( \epsilon < 1/20000 \). \( \square \)

C. A lower bound for \( \alpha \) using the curve \( \gamma \)

Here we will prove Proposition 5.2, i.e. the estimate

\[ |\exp^{-1}(C) - \exp^{-1}(A)| \geq L(\gamma) \left[ \frac{99}{100} \sin \left( \alpha - \frac{\epsilon}{4} \right) - 500\epsilon - 3 - \frac{8}{3} L(\gamma) \left( r + \frac{r + 3/2}{f(r)} \right) \right]. \]

We will use the fact that \( N_g \) has bounded second fundamental form (see the first statement of Lemma C.3) and that \( \gamma \) is a geodesic (see the second statement of the same lemma).

We will use the notation introduced in Section 5 and at the beginning of Appendix B. Recall that \( \tilde{\gamma}(t) := \exp_{\tilde{\gamma}(t)}^{-1}(\gamma(t)) \) is a section of \( vN_g \) along \( \pi \).

First we will set a lower bound on the initial derivative of \( \tilde{\gamma} \).

Lemma C.1. We have

\[ \left| \frac{\nabla}{dt} \tilde{\gamma}(0) \right| \geq \frac{99}{100} \left[ \sin \left( \alpha - \frac{\epsilon}{4} \right) - 200\epsilon \right]. \]

Proof. Analogously to the proof of Lemma B.3 we have \( \exp_{\tilde{\gamma}(0)}(\nabla_{\frac{d}{dt}} \tilde{\gamma}(0)) = \text{pr}_{\tilde{\gamma}(0)} LC^\pi \), where \( \frac{d}{dt} \tilde{\gamma}(0) \) is an element of \( T_{\tilde{\gamma}(0)} v\pi(0)N_g \).

By [WY Prop. 3.7] we have \( d(\text{Hor}_C^\pi, a\text{Hor}_C^\pi) \leq \epsilon/4 \). So

\[ \angle(\tilde{\gamma}(0), a\text{Hor}_C^\pi) \geq \angle(\tilde{\gamma}(0), \text{Hor}_C^\pi) - d(\text{Hor}_C^\pi, a\text{Hor}_C^\pi) \geq \alpha - \epsilon/4. \]

Therefore \( |\text{pr}_{\tilde{\gamma}(0)} a\text{Hor}_C^\pi| \geq \sin(\alpha - \epsilon/4) \).

On the other hand, by Lemma B.2 \( |\text{pr}_{\tilde{\gamma}(0)} a\text{Hor}_C^\pi - \text{pr}_{\tilde{\gamma}(0)} LC^\pi| \leq 200\epsilon \). The inverse triangle inequality gives

\[ |\text{pr}_{\tilde{\gamma}(0)} LC^\pi| \geq \sin(\alpha - \epsilon/4) - 200\epsilon. \]

Applying \( \exp_{\tilde{\gamma}(0)}^{-1} \), by Lemma A.1 we have

\[ |\exp_{\tilde{\gamma}(0)}^{-1}(\text{pr}_{\tilde{\gamma}(0)} LC^\pi)| \geq \frac{1}{1 + \epsilon/5} |\text{pr}_{\tilde{\gamma}(0)} LC^\pi|, \]

and since \( \frac{1}{1 + \epsilon/5} \geq \frac{99}{100} \) we are done. \( \square \)
Our next goal is to show that $\tilde{\gamma}(t)$ “grows at a nearly constant rate”. This will be achieved in Corollary C.3. Together with Lemma C.1 and integration along $\pi$ this will give the estimate of Proposition 5.2.

The next two lemmas will be used to prove Corollary C.1, where we will show that $\nabla_{\pi(0)} \tilde{\gamma}(t)$ and $\exp^{-1} \circ \exp_{\gamma}^{-1} \circ \nabla_{\pi(t)} \tilde{\gamma}(t)$, i.e. the parallel translate of $\nabla_{\pi(t)} \tilde{\gamma}(t)$ “along $\gamma$”, are close for all $t$. Here $\nabla_{\pi(t)} \tilde{\gamma}(t)$ denotes the vector $\nabla_{\pi(t)} \tilde{\gamma}(t)$ regarded as an element of $T_{\tilde{\gamma}(t)}(\nu\pi(t)N_g)$. To this end we show that $pr_{\dot{\gamma}(0)} LC \approx pr_{\dot{\gamma}(0)} Hor \approx pr_{\dot{\gamma}(t)} Hor \approx pr_{\dot{\gamma}(t)} LC$, where we identify tangent spaces by parallel translation along $\gamma$. The crucial step is the second “$\approx$”, where we use the fact that $\gamma$ is a geodesic. Applying $\exp^{-1}$ will easily imply Corollary C.1 since $\exp^{-1}(pr_{\dot{\gamma}(t)} LC) = \frac{\partial}{\partial t} \tilde{\gamma}(t)$.

Lemma C.2. For any $L < 1$ and any point $p \in \exp_{N_g}(\nu N_g)_L$ the orthogonal projections $T_p M \rightarrow aHor^p$ and $T_p M \rightarrow Hor^p$ differ at most by $L^2/5$ in the operator norm.

Proof. This follows immediately from [We, Prop. 3.7]. ⊓ ⊔

Lemma C.3. For all $t$,

$$d(Vert_C^g, \gamma^b \setminus Vert_C^g) \leq \arcsin \left[ t \left( r + \frac{r + 3/2}{f(r)} \right) \right].$$

Furthermore,

$$|pr_{\dot{\gamma}(0)} Hor - \gamma^b \circ pr_{\dot{\gamma}(t)} Hor| \leq t \left( r + \frac{r + 3/2}{f(r)} \right).$$

Proof. We first want to estimate $d(Vert_C^g, \gamma^b \setminus Vert_C^g)$. Let $v \in \nu C N_g$ be a normal unit vector.

First of all, for the $\nabla$ and $\nabla^\perp$ parallel translations along $\pi$ from $C$ to $\pi(t)$ we have

$$|\nabla v - \frac{1}{f} \nabla v| \leq \frac{3}{2} L(\pi\lli;0,t\rli) \leq \frac{3}{2} \frac{t}{f(r)}.$$ 

The first inequality follows from a simple computation involving the second fundamental form of $N_g$, which is bounded in norm by $3/2$ (see [We, Cor. 3.2]). The second inequality is due to $f(r)L(\pi\lli;0,t\rli) \leq L(\gamma\lli;0,t\rli)$, which follows from [We, Lemma 3.3].

Secondly, denoting by $\tau_t$ the unit-speed geodesic from $\pi(t)$ to $\gamma(t)$, we have

$$|\tau_t v - \gamma^b \circ \tau_t \circ v| \leq \tau_t \left( 1 + \frac{1}{f(r)} \right).$$

Indeed, the above expression just measures the holonomy as one goes once around the polygonal loop given by the geodesics $\gamma^b, \pi, \tau_t$ and $\gamma^b$. Using the bounds on curvature, we know that this is bounded by the area of a surface spanned by the polygon (see [BK]...
6.2.1]. The estimate given above surely holds since $L(\tau_t), L(\tau_0) \leq r, L(\gamma t |_{[0, t]} ) = t$ and, as we just saw, $L(\pi t |_{[0, t]} ) \leq t/f(r)$.

Together this gives
\[
| \| \tau_t \gamma \| _\pi \circ \tau_t \circ \frac{\gamma}{t} \|_\pi | \leq | \| \tau_t \gamma \| _\pi \circ \tau_t \circ \frac{\gamma}{t} \|_\pi | + | \| \frac{\gamma}{t} \|_\pi | \leq t \left( r + \frac{r + 3/2}{f(r)} \right).
\]

So we obtain a bound on the distance from $\tau_t \gamma \in \text{Vert}_{\gamma(t)}$ to a unit vector in $\nu_{\gamma(t)} \text{Vert}_{\gamma(t)}$. This yields the first statement of the lemma. The second follows from [We Prop. A.4], since $\nu_{\gamma(t)} \text{Hor} = \nu_{\gamma(t)} \text{Hor}_{\gamma(t)}$ because $\gamma$ is a geodesic.

**Corollary C.1.** For all $t$, 
\[
| \exp^{-1} \gamma \circ \exp \left( \frac{\nabla}{dt} \tilde{\gamma}(t) \right) - \frac{\nabla}{dt} \tilde{\gamma}(0) | \leq \frac{51}{50} \left[ 2.1(100\epsilon + r) + t \left( r + \frac{r + 3/2}{f(r)} \right) \right].
\]

**Proof.** From Lemmata C.2 and B.2 we have, for all $t$, 
\[
| \text{pr}_{\gamma(t)} \text{Hor} - \text{pr}_{\gamma(t)} \text{LC} | \leq | \text{pr}_{\gamma(t)} \text{Hor} - \text{pr}_{\gamma(t)} \text{aHor} | + | \text{pr}_{\gamma(t)} \text{aHor} - \text{pr}_{\gamma(t)} \text{LC} | \\
\leq r^2/5 + 2r \leq 2.1 r.
\]

For $t = 0$, since $d(C, N_\gamma) < 100\epsilon$, we have the better estimate 
\[
| \text{pr}_{\gamma(0)} \text{Hor} - \text{pr}_{\gamma(0)} \text{LC} | \leq 210\epsilon.
\]

Combining this with the second statement of Lemma C.3 gives 
\[
| \text{pr}_{\gamma(0)} \text{LC} - \nu_{\gamma(t)} \text{pr}_{\gamma(t)} \text{LC} | \leq 2.1(100\epsilon + r) + t \left( r + \frac{r + 3/2}{f(r)} \right).
\]

Recall that $\text{pr}_{\gamma(t)} \text{LC} = \exp_{\gamma(t)} \left( \frac{\nabla}{dt} \tilde{\gamma}(t) \right)$, as in the proof of Lemma B.3. Also, for any vector $X \in T_\gamma M$ we have $| \exp^{-1} \gamma X | \leq | X | / (1 - \epsilon/5)$ by Lemma A.1. So applying $(\exp^{-1})_\gamma$ to $\text{pr}_{\gamma(t)} \text{LC} - \nu_{\gamma(t)} \text{pr}_{\gamma(t)} \text{LC}$ we get 
\[
| \frac{\nabla}{dt} \tilde{\gamma}(0) - \exp^{-1} \gamma \circ \exp \left( \frac{\nabla}{dt} \tilde{\gamma}(t) \right) | \leq \left[ 2.1(100\epsilon + r) + t \left( r + \frac{r + 3/2}{f(r)} \right) \right] \frac{1}{1 - \epsilon/5}.
\]

Now let $\xi$ be a unit vector in $\nu_{\pi(t)} N_\gamma$. Denote by $\hat{\xi}$ the same vector thought of as an element of $T_{\gamma(t)} (\nu_{\pi(t)} N_\gamma)$. In the next two lemmas we want to show that $\exp_{\gamma(t)}^{-1} \gamma \circ \exp_{\gamma(t)} \hat{\xi} \in T_{\gamma(t)} M$ are close to each other, i.e. that under the identification by $\exp$ the $\nabla$-parallel translation along $\gamma$ and the $\nabla$-parallel translation along $\gamma$ do not differ too much. Here we also make use of the fact that $N$ has bounded second fundamental form (see Lemma C.5). In Corollary C.2 we will apply this to the vector $\frac{\nabla}{dt} \tilde{\gamma}(t)$. 

Marco Zambon
Lemma C.4. Denoting by $\tau_t$, the unit-speed geodesic from $\pi(t)$ to $\gamma(t)$, we have

$$|\tau_0\circ\gamma\circ\tau_1\circ\xi - \exp_{\gamma}\circ\exp_{\gamma}\circ\exp_{\gamma}\circ\dot{\xi}| < \frac{r^2}{2}.$$

**Proof.** First let us notice that applying Lemma A.1 three times we get

$$|\tau_0\circ\gamma\circ\exp_{\gamma}\circ\dot{\xi} - \exp_{\gamma}\circ\exp_{\gamma}\circ\dot{\xi}| \leq \frac{r^2}{5}|\exp_{\gamma}\circ\exp_{\gamma}\circ\dot{\xi}| \leq \frac{r^2}{5} + \frac{r^2 + r^2/5}{5(1 - r^2/5)}.$$

Therefore, by applying Lemma A.1 to $\xi$, the left hand side of the statement of this lemma is bounded above by

$$|\tau_0\circ\gamma\circ[I_{\gamma}\circ\xi] - \tau_0\circ\gamma\circ[I_{\gamma}\circ\exp_{\gamma}\circ\dot{\xi}]| + |\tau_0\circ[I_{\gamma}\circ\exp_{\gamma}\circ\dot{\xi}] - \exp_{\gamma}\circ[I_{\gamma}\circ\exp_{\gamma}\circ\dot{\xi}]|$$

$$\leq \frac{r^2}{5} + \frac{r^2 + r^2/5}{5(1 - r^2/5)} \leq \frac{2}{5(1 - r^2/5)}.$$  \hfill \Box

Lemma C.5.

$$|\exp^{-1}_{\gamma}\circ\gamma\circ\exp_{\gamma}\circ\dot{\xi} - \frac{1}{\pi^b}\circ\xi| \leq \frac{r^2}{2} + t \left(r + \frac{r + 3/2}{f(r)}\right).$$

**Proof.** The left hand side is bounded above by

$$|\exp^{-1}_{\gamma}\circ\gamma\circ\exp_{\gamma}\circ\dot{\xi} - \frac{1}{\pi^b}\circ\xi| + |\exp^{-1}_{\gamma}\circ\gamma\circ\xi - \pi^b\circ\xi| + |\pi^b\circ\xi - \frac{1}{\pi^b}\circ\xi|$$

$$\leq \frac{r^2}{2} + rt \left(1 + \frac{1}{f(r)}\right) + \frac{1}{2} \frac{t}{f(r)}.$$  \hfill \Box

The first term is estimated by Lemma C.4. The second one is just the holonomy as one goes around the loop given by $\tau_t$, $\gamma^b$, $\tau_{1/0}$ and $\pi$, which was bounded above in the proof of Lemma C.3. The third and last term is estimated in the proof of Lemma C.3 as well.  \hfill \Box

Corollary C.2. The section $\tilde{\gamma}$ satisfies

$$\left|\exp^{-1}_{\gamma}\circ\gamma\circ\exp_{\gamma}\left(\frac{1}{dt}\tilde{\gamma}(t)\right) - \frac{1}{\pi^b}\circ\exp_{\gamma}\left(\frac{1}{dt}\tilde{\gamma}(t)\right)\right| \leq \frac{2}{3} r^2 + \frac{4}{3} t \left(r + \frac{r + 3/2}{f(r)}\right).$$

**Proof.** We apply Lemma C.5 to $\frac{1}{dt}\tilde{\gamma}(t)$, where now we have to take into consideration the length of $\frac{1}{dt}\tilde{\gamma}(t)$ in our estimate. We have

$$\left|\frac{1}{dt}\tilde{\gamma}(t)\right| = |\exp^{-1}_{\gamma}(pr_{\gamma}(LC))| \leq \frac{1}{1 - r^2/5}|pr_{\gamma}(LC)|,$$
by Lemma A.1 and
\[ |pr_{\dot{\gamma}(t)} \mathcal{L} \gamma| \leq |pr_{\dot{\gamma}(t)} \mathcal{L} \gamma - pr_{\dot{\gamma}(t)} \mathcal{aHor}| + |pr_{\dot{\gamma}(t)} \mathcal{aHor}| \leq 2r + 1 \]
by Lemma B.2. Since \(\frac{2r+1}{1-t^2 r} \leq \frac{1}{4} \) for \(r \leq 0.08\) we are done. \(\square\)

Now Corollaries C.1 and C.2 immediately imply that \(\dot{\gamma}(t)\) “grows at a nearly constant rate”:

**Corollary C.3.** The section \(\dot{\gamma}\) satisfies
\[
\left| \frac{\nabla}{dt} \dot{\gamma}(0) - \frac{1}{\pi h} \frac{\nabla}{dt} \dot{\gamma}(t) \right| \leq 3(100r + r) + \frac{8}{3} t \left( r + \frac{3/2}{f(r)} \right).
\]

**Proof of Proposition 5.2.** The estimate of Proposition 5.2 follows from
\[
|\exp^{-1}(C) - \frac{1}{\pi h} \exp^{-1}(A)| = \left| \int_{0}^{L(y)} \frac{\nabla}{dt} \dot{\gamma}(t) dt \right|
\geq \left| \int_{0}^{L(y)} \frac{\nabla}{dt} \dot{\gamma}(0) dt \right| - \left| \int_{0}^{L(y)} \left( \frac{\nabla}{dt} \dot{\gamma}(0) - \frac{1}{\pi h} \frac{\nabla}{dt} \dot{\gamma}(t) \right) dt \right|
\geq L(y) \cdot \left| \frac{\nabla}{dt} \dot{\gamma}(0) \right| - \int_{0}^{L(y)} \left| \frac{\nabla}{dt} \dot{\gamma}(0) - \frac{1}{\pi h} \frac{\nabla}{dt} \dot{\gamma}(t) \right| dt
\]
by using Lemma C.1 and Corollary C.3. \(\square\)

**References**


