J. Eur. Math. Soc. 8, 77–122

Marco Zambon

Submanifold averaging in Riemannian and symplectic geometry

Received October 8, 2002 and in revised form March 14, 2005

Abstract. We give a canonical construction of an "isotropic average" of given C^1-close isotropic submanifolds of a symplectic manifold. For this purpose we use an improvement (obtained in collaboration with H. Karcher) of Weinstein's submanifold averaging theorem and apply "Moser's trick". We also present an application to Hamiltonian group actions.

Keywords. Averaging, isotropic, Lagrangian, Legendrian, parallel tubes, shape operators

Contents

1. Introduction . . . 77
2. Improved error estimates for the shape operators of parallel tubes with application to Weinstein's submanifold averaging . . . 83
3. Estimates on the map φ_g . . . 89
4. Proposition 4.1 about geodesic triangles in M . . . 96
5. Application of Proposition 4.1 to Vert^g . . . 98
6. Estimates on tubular neighborhoods of N_g on which φ_g is injective . . . 101
7. Conclusion of the proof of the Main Theorem . . . 104
8. Remarks on the Main Theorem . . . 111
9. An application to Hamiltonian actions . . . 112
A. The estimates of Proposition 4.1 . . . 113
B. An upper bound for α using the curve c . . . 115
C. A lower bound for α using the curve γ . . . 118

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 *canonically*[1] an "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

M. Zambon: Department of Mathematics, UC Berkeley, CA 94720, USA;
e-mail: zambon@math.berkeley.edu

Mathematics Subject Classification (2000): 53C55, 53D12

[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.[2] 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 $\mathrm{graph}(\phi) \subset M_1 \times M_2$. 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;[3] 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 $\mathrm{graph}(\phi)$ 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 νN (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 $d_1(N', N) < 250 d_1(N, N')$ (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:[4]

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 νN_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, ω) 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 $I^2 = -\operatorname{Id}_{TM}$. The tuple (M, g, ω, 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 ∇ 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.[5] We state the theorem choosing the bound to be 1 (but see Remark (i) below).

Theorem 1 (Main Theorem)**.** *Let (M^m, g, ω, 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^n 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 ϵ 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, ω) by symplectomorphisms, by averaging over the compact group one can always find some invariant metric $\tilde{g}$. Using ω and $\tilde{g}$ one can canonically construct a metric g which is compatible with ω (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, ω). In general it does not seem possible to give any bound on $|\nabla\omega|$, where ∇ is the Levi-Civita connection corresponding to g.

Theorem 2. *Let (M, g, ω, 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 be an isotropic submanifold of M such that (M, N_0) is a gentle pair and $d_1(N_0, gN_0) < \epsilon < 1/70000$ for all $g \in G$. Then there is a G-invariant isotropic submanifold L with $d_0(N_0, L) < 1000\epsilon$.*

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 $\{gN_0\}_{g\in G}$. 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 ϵ 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 $\nabla\omega = 0$.

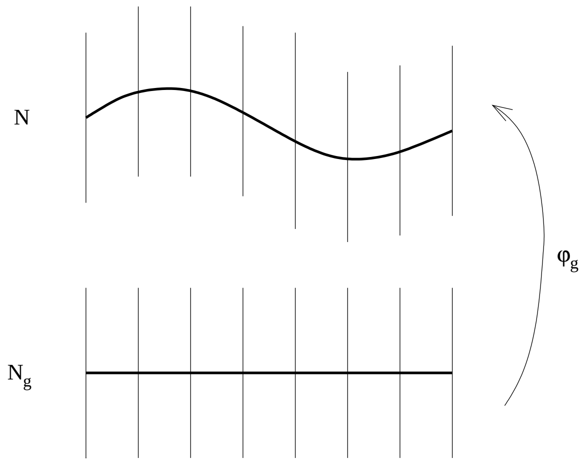
Part I. We start by considering the average of the submanifolds N_g as in Theorem 2.3 of [We], which we will denote N. We will use the notation $\exp_N$ to indicate the restriction of the exponential map to $TM|_N$, and similarly for any of the N_g's. For any g in G, the average N lies in a tubular neighborhood of N_g and is the image under $\exp_{N_g}$ of a section σ of νN_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 νN_g, we use the term "normal slice" for the submanifold $\exp_{N_g}(\nu_p N_g)_1$. We define the following map:

$$\varphi_g : \exp_{N_g}(\nu N_g)_1 \to M, \quad \exp_p(v) \mapsto \exp_q({}_{\gamma_q}\backslash\!\!\backslash v).$$

Here p, q, and γ_q are as above, $v \in (\nu_p N_g)_1$, and ${}_{\gamma_q}\backslash\!\!\backslash$ denotes parallel translation along γ_q. So φ_g takes the normal slice $\exp_p(\nu_p N_g)_1$ to $\exp_q(\mathrm{Vert}_q^g)_1$, where $\mathrm{Vert}_q^g \subset T_q M$ is the parallel translation along γ_q of $\nu_p N_g \subset T_p M$.

We have $d(\mathrm{Vert}_q^g, \nu_q N) < d_1(N_g, N) < 2500\epsilon < \pi/2$, so Vert_q^g and $T_q N$ are transversal. Therefore φ_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 φ_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 φ_g.



Part II. Now we introduce the symplectic form

$$\omega_g := (\varphi_g^{-1})^* \omega$$

on $\exp_N(\mathrm{Vert}^g)_{0.05}$. Notice that N is isotropic with respect to ω_g by construction, hence also with respect to the 2-form $\int_g \omega_g$ which is defined on $\bigcap_{g \in G} \exp_N(\mathrm{Vert}^g)_{0.05}$. We would like to apply Moser's trick[6] (see [Ca, Chapter III]) to ω and $\int_g \omega_g$. To do so we

[6] Recall that Moser's Theorem states the following: if Ω_t ($t \in [0, 1]$) is a smooth family of symplectic forms lying in the same cohomology class in a compact manifold then there is a family of diffeomorphisms ρ_t with $\rho_0 = \mathrm{Id}$ satisfying $\rho_t^* \Omega_t = \Omega_0$.

first restrict our forms to a smaller tubular neighborhood tub^ϵ of N, which we define in Section 7.1. To apply Moser's trick we have to check:

1. *On* tub^ϵ *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^ϵ the differential of φ_g^{-1} is "close" to the parallel translation $\backslash\!\backslash$ 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 $q \in \text{tub}^\epsilon$ and nonzero $X, Y \in T_q M$,

$$(\omega_g)_q(X,Y) = \omega_{\varphi_g^{-1}(q)}((\varphi_g^{-1})_*(X), (\varphi_g^{-1})_*(Y)) \approx \omega_{\varphi_g^{-1}(q)}(\backslash\!\backslash X, \backslash\!\backslash Y) \approx \omega_q(X,Y),$$

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

2. *On* tub^ϵ *the forms* ω *and* $\int_g \omega_g$ *belong to the same cohomology class.*

Fix $g \in G$. The inclusion $i : \text{tub}^\epsilon \hookrightarrow \exp_{N_g}(\nu N_g)_1$ is homotopic to $\varphi_g^{-1} : \text{tub}^\epsilon \to \exp_{N_g}(\nu N_g)_1$. A homotopy is given by thinking of N as a section of νN_g and "sliding along the fibers" to the zero section. Therefore these two maps induce the same map in cohomology, and pulling back ω we have

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

Integrating over G finishes the argument.

Now we can apply Moser's trick: if α is a 1-form on tub^ϵ such that $d\alpha$ is equal to $\frac{d}{dt}\omega_t = \int_g \omega_g - \omega$, then the flow ρ_t of the time-dependent vector field $v_t := -\tilde{\omega}_t^{-1}(\alpha)$ has the property $\rho_t^*\omega_t = \omega$ (and in particular $\rho_1^*(\int_g \omega_g) = \omega$) where it is defined. Therefore if $L := \rho_1^{-1}(N)$ is a well defined submanifold of tub^ϵ, then it will be isotropic with respect to ω since N is isotropic with respect to $\int_g \omega_g$.

We will construct canonically a primitive α 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 α has small maximum norm. So, if ϵ is small enough, the time-1 flow of the time-dependent vector field $\{-v_{1-t}\}$ will not take N out of tub^ϵ and L will be well defined.

Since our construction is canonical after fixing the almost-Kähler structure (g, ω, 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 φ_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 φ_g is injective on $\exp_{N_g}(\nu N_g)_{0.05}$. 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 ω_t's are symplectic forms and that the 1-form α, 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 ρ_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 $H_N(v) = \nabla_v \operatorname{grad} P_N$. Differentiating the relation $\operatorname{grad} P_N = \rho_N \cdot \operatorname{grad} \rho_N$ we see that

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

where $U_N = \operatorname{grad} \rho_N$ is the radial unit vector (pointing away from N), 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 ρ_N in the direction of the normal vector U_N.[7]

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

Proposition 3.11 of [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[8] of $T_p M$:

$$\begin{bmatrix} 0.64 \cdot I & 0 \\ 0 & -3t \cdot I \end{bmatrix} < H_N < \begin{bmatrix} 1.32 \cdot I & 0 \\ 0 & 3t \cdot I \end{bmatrix},$$

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 [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 ϵ in Theorems 2.2 and 2.3 of [We].

We will improve the estimate of Corollary 3.13 of [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 $t \leq$ normal injectivity radius of N. Let p lie in the tube $\tau(t) := \rho_N^{-1}(t)$, and let $S_N : T_p\tau(t) \to T_p\tau(t)$ 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 $r \mapsto \exp_{c(s)} rU_N(c(s))$, 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 ∇ the Levi-Civita connection on M we have

$$\begin{aligned} \tilde{J}'(0) &= \frac{\nabla}{dr}\bigg|_0 \frac{d}{ds}\bigg|_0 f(s,r) = \frac{\nabla}{ds}\bigg|_0 \frac{d}{dr}\bigg|_0 f(s,r) = \frac{\nabla}{ds}\bigg|_0 U_N(c(s)) \\ &= \nabla_v U_N = \mathrm{pr}(\nabla_v U_N) = S_N v. \end{aligned}$$ □

Using the above lemma we will be able to prove this improvement of Proposition 3.11 of [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 $t \leq$ normal injectivity radius of N. Let γ 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 $\dot{\gamma}(t)$ 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*

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

[8] See our Section 3 or Section 2.1 in [We] for the definition of vertical and horizontal bundle at p.

Remark. We adopt the following unconventional notation: If $M, \tilde{M}$ are matrices and c a real number, $M \leq \tilde{M} + c$ means that $M - \tilde{M}$ has operator norm $\leq c$. Generalizing to the case where we consider also vertical-horizontal decompositions of matrices,

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

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

Proof of Theorem 3. Choose an orthonormal basis $\{E_1, \ldots, E_{n-1}\}$ of $\dot{\gamma}(0)^\perp \subset T_{\gamma(0)}M$ 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 γ with the following initial conditions:

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

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 $\dot{\gamma}(0)$, therefore, as long as the $J_i(t)$ are linearly independent, they form a basis of $\dot{\gamma}(t)^\perp = T_{\gamma(t)}\tau(t)$. Also, the J_i's are N-Jacobi fields, i.e. Jacobi fields for which $J_i(0)$ is tangent to N and $J_i'(0) - B_{\dot{\gamma}(0)} J_i(0)$ 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 γ which are orthogonal to $\dot{\gamma}$, and this space coincides with the space of N-Jacobi fields arising from a variation of unit-speed[9] geodesics normal to N. The velocity vectors of such variations at time t coincide with U_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

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

and

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

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 γ 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 $\dot{\gamma}(t)$ 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 $\dot{\gamma}$ 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 $T_{\gamma(t)}\tau(t)$.

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 $J(t)_{VV}$, $J(t)_{HV}$, $J(t)_{VH}$ and $J(t)_{HH}$ 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 $A_i(t) = \backslash\backslash(J_i(0) + t \cdot J_i'(0))$, where $\backslash\backslash$ denotes parallel translation along γ. The map $\mathbb{R}^{n-1} \to T_{\gamma(t)}\tau(t),\ e_i \mapsto A_i(t)$, 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

$$J(t) - A(t) =: F_1(t) \leq \begin{pmatrix} \sinh(t) - t & \cosh(t) - 1 + |B|(\sinh(t) - t) \\ \sinh(t) - t & \cosh(t) - 1 + |B|(\sinh(t) - t) \end{pmatrix} \leq \begin{pmatrix} \frac{1}{5}t^3 & \frac{3}{4}t^2 \\ \frac{1}{5}t^3 & \frac{3}{4}t^2 \end{pmatrix}.$$

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

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

Clearly A is invertible and

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

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

$$A^{-1}F_1 \leq \begin{pmatrix} \frac{1}{5}t^2 & \frac{3}{4}t \\ \frac{2}{5}t^3 & \frac{3}{2}t^2 \end{pmatrix}.$$

Clearly[11] its norm is less than $\sqrt{2}\frac{3}{4}t\sqrt{1+4t^2} \leq \frac{3}{2}t < 1$ 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: $J(t)_{HV} : \mathbb{R}^k \times \{0\} \to \mathrm{Hor}(t)$ is given by restricting $J(t)$ and then composing with the orthogonal projection onto the horizontal space at $\gamma(t)$.

[11] If $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \leq \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ then the full operator norm of the matrix is bounded by $\sqrt{\max\left\{\left|\begin{smallmatrix} a \\ c \end{smallmatrix}\right|, \left|\begin{smallmatrix} b \\ d \end{smallmatrix}\right|\right\} + ab + cd} \leq \sqrt{2}\max\left\{\left|\begin{smallmatrix} a \\ c \end{smallmatrix}\right|, \left|\begin{smallmatrix} b \\ d \end{smallmatrix}\right|\right\}$.

for $A^{-1}F_1$ we have

$$[A^{-1}F_1]^2 \leq \begin{pmatrix} \frac{1}{2}t^4 & \frac{3}{2}t^3 \\ t^5 & 3t^4 \end{pmatrix}.$$

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

$$[I + A^{-1}F_1]^{-1} = I + F_2 \quad \text{where} \quad F_2 \leq \begin{pmatrix} \frac{15}{2}t^2 & 5t \\ 15t^3 & \frac{19}{2}t^2 \end{pmatrix}.$$

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

$$\begin{aligned} \Big|\sum c_i(J_i'(t) - A_i'(t))\Big| &\leq \int_0^t \Big|\sum c_i(J_i''(\tau) - A_i''(t))\Big| d\tau \\ &\leq \begin{cases} \cosh(t) - 1 \leq \frac{3}{4}t^2 & \text{if } i \leq k, \\ \sinh(t) + |B|(\cosh(t) - 1) \leq \frac{3}{2}t & \text{if } i \geq k+1. \end{cases} \end{aligned}$$

So altogether we obtain

$$J'(t) - A'(t) =: F_3(t) \quad \text{where} \quad F_3(t) \leq \begin{pmatrix} \frac{3}{4}t^2 & \frac{3}{2}t \\ \frac{3}{4}t^2 & \frac{3}{2}t \end{pmatrix}.$$

Now finally we can estimate

$$\begin{aligned} tS_N(t) &= tJ'J^{-1} = (A' + F_3) \cdot (I + F_2) \cdot tA^{-1} \\ &\leq \left\{ \begin{bmatrix} I & 0 \\ 0 & B \end{bmatrix} + \begin{pmatrix} \frac{3}{4}t^2 & \frac{3}{2}t \\ \frac{3}{4}t^2 & \frac{3}{2}t \end{pmatrix} + \begin{pmatrix} \frac{15}{2}t^2 & 5t \\ 15|B|t^3 & \frac{19}{2}|B|t^2 \end{pmatrix} + \begin{pmatrix} 30t^4 & 18t^3 \\ 30t^4 & 18t^3 \end{pmatrix} \right\} \cdot tA^{-1} \\ &\leq \begin{bmatrix} I & 0 \\ 0 & tB \end{bmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 2|B|^2t^2 \end{pmatrix} + \begin{pmatrix} \frac{3}{4}t^2 & 3t^2 \\ \frac{3}{4}t^2 & 3t^2 \end{pmatrix} + \begin{pmatrix} \frac{15}{2}t^2 & 10t^2 \\ 15|B|t^3 & 19|B|t^3 \end{pmatrix} \\ &\quad + \begin{pmatrix} 30t^4 & 36t^4 \\ 30t^4 & 36t^4 \end{pmatrix}. \end{aligned}$$

Here we used

$$tA^{-1} \leq \begin{bmatrix} I & 0 \\ 0 & tI \end{bmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 2|B|t^2 \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 2t \end{pmatrix}$$

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. □

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 [We]. Now we can achieve an upper bound proportional to t^2, versus the bound proportional to $\sqrt{t}$ in Corollary 3.13 of [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 [We] that were originally derived using Corollary 3.13 there. We will improve only estimates contained in Lemmata 4.7 and 4.8 of [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_e. (The zero set of $\mathcal{V}$ is the average N of the family $\{N_g\}$.) As in [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 [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_g)) for the operator norm of H_g we have $1 - 16t^2 \leq |H_g|$, so that one obtains $H_g(\mathbb{P}_{\Gamma_g}v, \mathbb{P}_{\Gamma_g}v) > 19/20$ in the proof of Lemma 4.7 in [We]. Similarly, Theorem 3 together with footnote 11 implies that $|H_g| < 1.01$. Using these estimates in the proof of Lemma 4.7 in [We] gives the claim. □

Similarly, we will replace the term $60\sqrt{\epsilon}$ in Lemma 4.8 of [We] by 1950ϵ.

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 $3\sqrt{t}$ by $16t^2$ in the proof of Lemma 4.8 in [We] and we can use 1.01 instead of 1.32 as an upper bound for $|H_g|$. Furthermore, we replace the constant 1000 coming from Lemma 4.3 in [We] by 525.[13] This gives the improved estimate $H_g(v, \mathbb{P}_{\bar{\Gamma}}w) \leq 850\epsilon|v| \cdot |w|$ and simple arithmetic concludes the proof. □

[12] See our Section 3 or Section 3.2 in [We] for the definitions of almost horizontal and almost vertical bundle.

[13] Lemma 4.3 of [We] quotes incorrectly Proposition A.8 from its own appendix.

From these two lemmas it follows that the operator from $(\mathrm{aVert}^e)^\perp$ to aVert^e whose graph is T_xN has norm at most $\frac{5}{4}\cdot 1950\epsilon$. 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 ϵ, 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 φ_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 $X \in T_pM$. The aim of this section is to estimate the difference between $\varphi_{g*}X$ and $\backslash\!\backslash X$. This will be achieved in Proposition 3.4.

Here we denote by $\backslash\!\backslash X$ the following parallel translation of X, where π_{N_g} is the projection onto N_g along the normal slices. First we parallel translate X along the shortest geodesic from p to $\pi_{N_g}(p)$, then along the shortest geodesic from $\pi_{N_g}(p) \in N_g$ to its image under φ_g, and finally along the shortest geodesic to $\varphi_g(p)$. We view "$\backslash\!\backslash$" as a canonical way to associate $X \in T_pM$ 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 $TM|_{\exp_{N_g}(\nu N_g)_1}$ and their orthogonal complements.

The *vertical bundle* Vert^g 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* aVert^g 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* Hor^g and the *almost horizontal bundle* aHor^g are given by their orthogonal complements.

Remark. Notice that aVert^g is the kernel of $(\pi_{N_g})_*$, and that according to Proposition 3.7 in [We] we have $d(\mathrm{Vert}^g_p, \mathrm{aVert}^g_p) < \frac{1}{4}d(p, N_g)^2$ for any p in $\exp_{N_g}(\nu N_g)_1$, and similarly for Hor^g and aHor^g.

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

Now we are ready to give our estimates on the map φ_g. Recall from the introduction that for any point q of the tubular neighborhood of N_g we denote by γ_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 $X \in T_p N_g$ is a unit vector, then*

$$|\varphi_{g*}(X) - \backslash\backslash X| \leq 3200\epsilon.$$

Remark. Notice that if X is a vector normal to N_g by definition of φ_g and $\backslash\backslash$ we have $\varphi_{g*}(X) = \backslash\backslash X$. 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 ξ, where $(\cdot)^T$ denotes projecting to the component tangent to N_g and ∇ 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 [We, Cor. 3.2].

Now let $p \in N_g$, $X \in T_p N_g$ 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 [We]). We will show that at q,

$$\backslash\backslash X \approx J(E) \approx H \approx \varphi_{g*}(X)$$

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 γ_q such that $J(0) = X$ and $J'(0) = -A_{\dot\gamma_q(0)}X$. Then*

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

Proof. This is an immediate consequence of [BK, 6.3.8iii][15] 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

$$|J(t) - {}_t^0\backslash\backslash(J(0) + t \cdot J'(0))| \leq |J(0)|(\cosh(t) - 1) + |J'(0)|(\sinh(t) - t),$$

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

Before proceeding we need a lemma about projections:

---

[14] In Section 2 we adopted the sign convention of [We] which differs from this.

[15] [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 $J_1(0) = 0$, $J_1'(0) = J'(0)$ 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 $Y = Y_{\mathrm{av}} + Y_{\mathrm{h}}$ for the splitting into its almost vertical and horizontal components. Then*

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

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

Lemma 3.3. *If H is the unique horizontal vector at q such that $(\pi_{N_g})_*(H) = X$, 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 3.1. Write $J(E) = W + Y$ for the splitting into horizontal and vertical components. Then, using the notation of Lemma 3.2, we have $J(E)_{\mathrm{h}} = W + Y_{\mathrm{h}}$ and $J(E)_{\mathrm{av}} = Y_{\mathrm{av}}$. Notice that the Jacobi field J arises from a variation of geodesics orthogonal to N_g (see the Remark in Section 3.2), so $(\pi_{N_g})_* J(E) = X = (\pi_{N_g})_* H$. Using $\mathrm{aVert}^g = \ker(\pi_{N_g})_*$ it follows that $H = J(E)_{\mathrm{h}}$. So

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

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

Now we will compare H to $\varphi_{g*}(X)$ and finish our proof.

Proof of Proposition 3.1. We have

$$|\backslash\backslash X - \varphi_{g*}(X)| \leq |\backslash\backslash X - J(E)| + |J(E) - H| + |H - \varphi_{g*}(X)|.$$

The first and second terms are bounded by the estimates of Lemmata 3.1 and 3.3. For the third term we proceed analogously to Lemma 3.3: since $\varphi_{g*}(X)$ and H are both mapped to X via π_{N_g}, one has $(\varphi_{g*}(X))_{\mathrm{av}} = \varphi_{g*}(X) - H$. As earlier, if $\varphi_{g*}(X) = \tilde{W} + \tilde{Y}$ is the splitting into horizontal and vertical components, we have $(\varphi_{g*}(X))_{\mathrm{av}} = \tilde{Y}_{\mathrm{av}}$. Therefore

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

Here we also used Lemma 3.2 and the fact that the angle enclosed by $\varphi_{g*}(X)$ and its orthogonal projection onto Hor_g^q is at most $d(\mathrm{Hor}_q^g, T_q N) \leq 2500\epsilon$ by Theorem 4. Altogether we have

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

Using this inequality we can bound $|\varphi_{g*}(X)|$ from above in terms of E and ϵ. Substituting into the right hand side of the above inequality we obtain a function of ϵ (recall that $E = 100\epsilon$) which is increasing and bounded above by 3200ϵ. □

3.2. Case 2: p is a point of $\partial \exp_{N_g}(\nu N_g)_L$ and $X \in T_pM$ is almost vertical

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

Remark. Jacobi fields $\bar{J}$ along γ_p (the geodesic from $\pi_{N_g}(p)$ to p) with $\bar{J}(0)$ tangent to N_g and $A_{\dot{\gamma}_p(0)}\bar{J}(0) + \bar{J}'(0)$ 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 γ_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 γ_p, so the map

$$\{N_g\text{-Jacobi fields along } \gamma_p\} \to T_pM, \qquad \bar{J} \mapsto \bar{J}(L),$$

is an isomorphism. The N_g-Jacobi fields that map to aVert_p^g are exactly those with $J(0) = 0$ and $J'(0) \in \nu_{\pi_{N_g}(p)}N_g$. Indeed, such a vector field is the variational vector field of the variation

$$f_s(t) = \exp_{\pi_{N_g}(p)} t[\dot{\gamma_p}(0) + sJ'(0)],$$

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 $J(0) \in T_{\pi_{N_g}(p)}N_g$ and $A_{\dot{\gamma}_p(0)}J(0) + J'(0) = 0$—which are called *strong N_g-Jacobi fields*—map to a subspace of T_pM which is a complement of aVert_p^g. As pointed out in [Wa, p. 354], these two subspaces are in general not orthogonal.

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

$$|\varphi_{g*}(X) - \backslash\backslash X| \leq 2\frac{\sinh(L) - L}{\sin(L)}.$$

We begin by proving

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

$$|J(L) - L \cdot {}_{\gamma_p}\backslash\backslash J'(0)| \leq \frac{\sinh(L) - L}{\sin(L)}.$$

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

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 $f_s(t) = \exp_{\pi_{N_g}(p)} t[\dot{\gamma_p}(0) + sJ'(0)]$. So $\varphi_{g*}(X)$ comes from the variation

$$\varphi_g(f_s(t)) = \exp_{\varphi_g(\pi_{N_g}(p))} t[\backslash\backslash\dot{\gamma_p}(0) + s\backslash\backslash J'(0)]$$

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 $\varphi_{g*}(X) = \tilde{J}(L)$. Notice that $\tilde{J}(0) = 0$ and $\tilde{J}'(0) = \backslash\backslash J'(0)$.

Lemma 3.5.
$$|\tilde{J}(L) - L \cdot {}_{\varphi_g \circ \gamma_p}\!\backslash\!\backslash \tilde{J}'(0)| \leq \frac{\sinh(L) - L}{\sin(L)}.$$

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

Proof of Proposition 3.2. We have $X \approx LJ'(0) = L\tilde{J}'(0) \approx \varphi_{g*}(X)$. Here we identify tangent spaces to M parallel translating along γ_p, along the geodesic $\gamma_{\varphi_g(\pi_{N_g}(p))}$ from $\pi_{N_g}(p)$ to its φ_g-image and along $\varphi_g \circ \gamma_p$ respectively. Notice that these three geodesics are exactly those used in the definition of "$\backslash\!\backslash$".

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 $\tilde{J}'(0) = \backslash\!\backslash J'(0)$. □

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

From now on we have to assume $L < 0.08$.

Proposition 3.3. *If $p \in \partial \exp_{N_g}(\nu N_g)_L$ and X is a unit vector equal to $J(L)$ for some strong N_g-Jacobi field J along γ_p, then*
$$|\varphi_{g*}(X) - \backslash\!\backslash X| \leq \frac{18}{5}L + 3700\epsilon.$$

We proceed analogously to Case 2.

Lemma 3.6. *For a vector field J as in the above proposition we have*
$$|J(L) - {}_{\gamma_p}\!\backslash\!\backslash J(0)| \leq \frac{\frac{3}{2}(e^L - 1)}{1 - \frac{3}{2}(e^L - 1)} \leq \frac{9}{5}L.$$
Furthermore we have $|J(0)| \leq \frac{1}{1 - \frac{3}{2}(e^L - 1)}$.

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

J comes from a variation $f_s(t) = \exp_{\sigma(s)} tv(s)$ for some curve σ in N_g with $\dot{\sigma}(0) = J(0)$ and some normal vector field v along σ. We denote by $\tilde{J}$ the Jacobi field along the geodesic $\varphi_g(\gamma_p(t))$ arising from the variation
$$\tilde{f}_s(t) = \varphi_g(f_s(t)) = \exp_{\tilde{\sigma}(s)}(t \backslash\!\backslash v(s)),$$
where $\tilde{\sigma} = \varphi_g \circ \sigma$ is the lift of σ to N. Then we have $\tilde{J}(L) = \varphi_{g*}(X)$. Notice that here $\backslash\!\backslash v(s)$ is just the parallel translation of $v(s)$ along $\gamma_{\tilde{\sigma}(s)} =: \gamma_s$.

Lemma 3.7.
$$|\tilde{J}(L) - {}_{\varphi_g \circ \gamma_p}\!\backslash\!\backslash \tilde{J}(0)| \leq \frac{9}{5}L.$$

Proof. Using [BK, 6.3.8iii] as in Lemma 3.1 we obtain

$$|\tilde{J}(L) - {}_{\varphi_g\circ\gamma_p}\backslash\!\backslash\tilde{J}(0)| \le |\tilde{J}'(0)|\sinh(L) + |\tilde{J}(0)|(\cosh(L) - 1), \qquad (*)$$

so that we just have to estimate the norms of $\tilde{J}(0)$ and $\tilde{J}'(0)$.

Since $\tilde{J}(0) = \varphi_{g*}J(0)$, Proposition 3.1 gives $|\tilde{J}(0) - {}_{\gamma_0}\backslash\!\backslash J(0)| \le 3200\epsilon|J(0)|$. Using the bound for $|J(0)|$ given in Lemma 3.6 we obtain

$$|\tilde{J}(0)| \le \frac{1 + 3200\epsilon}{1 - \frac{3}{2}(e^L - 1)}.$$

To estimate $\tilde{J}'(0)$ notice that in the expression for $f_s(t)$ we can choose $v(s) = {}_{\sigma_s}\backslash\!\backslash[\dot{\gamma}_0(0) + sJ'(0)]$, where ${}_{\sigma_s}\backslash\!\backslash$ denotes parallel translation from $\sigma(0)$ to $\sigma(s)$ along σ. So

$$\backslash\!\backslash v(s) = {}_{\gamma_s}\backslash\!\backslash\,{}_{\sigma_s}\backslash\!\backslash[\dot{\gamma}_0(0) + sJ'(0)],$$

and

$$\tilde{J}'(0) = \left.\frac{\nabla}{ds}\right|_0 (\backslash\!\backslash v(s)) = \left.\frac{\nabla}{ds}\right|_0 {}_{\gamma_s}\backslash\!\backslash\,{}_{\sigma_s}\backslash\!\backslash\dot{\gamma}_0(0) + {}_{\gamma_0}\backslash\!\backslash J'(0)$$

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 [BK, 6.2.1]). Therefore we write ${}_{\gamma_s}\backslash\!\backslash\,{}_{\sigma_s}\backslash\!\backslash\dot{\gamma}_0(0)$ as ${}_{\tilde{\sigma}_s}\backslash\!\backslash\,{}_{\gamma_0}\backslash\!\backslash\dot{\gamma}_0(0) + \varepsilon(s)$ where $\varepsilon(s)$ is a vector field along $\tilde{\sigma}(s)$ with norm bounded by the area of the polygon spanned by $\sigma(0), \sigma(s), \tilde{\sigma}(s)$ and $\tilde{\sigma}(0)$. Assuming that σ has constant speed $|J(0)|$ we can estimate $d(\sigma(0), \sigma(s)) \le s|J(0)|$, and using Proposition 3.1 to estimate $|\dot{\tilde{\sigma}}(s)| = |\varphi_{g*}\dot{\sigma}(s)|$ we obtain $d(\tilde{\sigma}(0), \tilde{\sigma}(s)) \le s(1 + 3200\epsilon)|J(0)|$. Using $d(\tilde{\sigma}(s), \sigma(s)) \le 100\epsilon$ and Lemma 3.6 we can bound the area of the polygon safely by

$$\frac{100\epsilon s(2 + 3200\epsilon)}{1 - \frac{3}{2}(e^L - 1)}.$$

So we obtain

$$\left|\left.\frac{\nabla}{ds}\right|_0 {}_{\gamma_s}\backslash\!\backslash\,{}_{\sigma_s}\backslash\!\backslash\dot{\gamma}_0(0)\right| = \left|\left.\frac{\nabla}{ds}\right|_0 \varepsilon(s)\right| \le \frac{100\epsilon(2 + 3200\epsilon)}{1 - \frac{3}{2}(e^L - 1)}.$$

To bound ${}_{\gamma_0}\backslash\!\backslash J'(0)$ notice that $|J'(0)| \le \frac{3}{2}|J(0)|$ using the fact that J is a strong Jacobi field and [We, Cor. 3.2], so

$$|J'(0)| \le \frac{3}{2}\frac{1}{1 - \frac{3}{2}(e^L - 1)}.$$

Substituting our estimates for $|\tilde{J}(0)|$ and $|\tilde{J}'(0)|$ in $(*)$ we obtain a function which, for $\epsilon < 1/20000$ and $L < 0.08$, is bounded above by $\frac{9}{5}L$. □

The vectors $\tilde{J}(0)$ and $\backslash\!\backslash J(0)$ generally are not equal, so we need one more estimate that has no counterpart in Case 2:

Lemma 3.8.
$$|\tilde{J}(0) - \backslash\backslash J(0)| \leq \frac{3200\epsilon}{1 - \frac{3}{2}(e^L - 1)} \leq 3700\epsilon.$$

Proof. Since $\tilde{J}(0) = \varphi_{g*}J(0)$, Proposition 3.1 gives
$$|\tilde{J}(0) - \backslash\backslash J(0)| \leq 3200\epsilon |J(0)| \leq \frac{3200\epsilon}{1 - \frac{3}{2}(e^L - 1)}.$$
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 γ_p, γ_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* $X \in T_pM$ *a unit vector. Then*
$$|\varphi_{g*}(X) - \backslash\backslash X| \leq 4L + 4100\epsilon.$$

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* γ_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 γ_p by parallel translation, we have
$$\begin{aligned}|\langle J(L), K(L)\rangle| &= |\langle J(L), K(L)\rangle - \langle LJ'(0), K(0)\rangle| \\ &\leq |\langle J(L), K(L) - K(0)\rangle| + |\langle J(L) - LJ'(0), K(0)\rangle| \\ &\leq |K(L) - K(0)| + |K(0)| \cdot |J(L) - LJ'(0)|,\end{aligned}$$
which can be estimated using Lemmata 3.6 and 3.4. □

Lemma 3.10. *Let* $X \in T_pM$ *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 $c := \langle J(L)/|J(L)|, K(L)/|K(L)|\rangle$, so $|c| \le \frac{9}{5}L$. There is an orthonormal basis $\{e_1, e_2\}$ of $\mathrm{span}\{J(L), K(L)\}$ 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 $1 = |J(L)+K(L)|^2 \ge (1-|c|)(|J(L)|^2 + |K(L)|^2)$, from which the lemma easily follows. □

Proof of Proposition 3.4. 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,

$$\begin{aligned}|\varphi_{g*}(X) - \backslash\backslash X| &\le |\varphi_{g*}J(L) - \backslash\backslash J(L)| + |\varphi_{g*}K(L) - \backslash\backslash K(L)| \\ &\le 1.1\left(2\frac{\sinh(L) - L}{\sin(L)} + \frac{18}{5}L + 3700\epsilon\right) \le 4L + 4100\epsilon.\end{aligned}$$ □

4. Proposition 4.1 about geodesic triangles in M

Fix g in G and let φ_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 $\exp_{N_g}(\nu N_g)_{0.05}$ is a tubular neighborhood of N_g on which φ_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*

$$\measuredangle(P_A, \dot{AC}) \ge \pi/2 - \delta$$

and

$$\theta := d(P_A, {}_{CA}\backslash\backslash P_C) \le \mathcal{C}d(A, C)$$

for some constants $\delta, \mathcal{C}$. Assume $\mathcal{C} \le 2$. Then

$$d(C, B) \ge \frac{10}{11}\frac{1}{\mathcal{C}+1}\cos(\delta).$$

Remark 1. Here ${}_{CA}\backslash\backslash P_C$ denotes parallel translation of P_C along the geodesic from C to A. The *angle* between the subspace P_A and the vector $\dot{AC}$ is given as follows: for every nonzero $v \in P_A$ we consider the nonoriented angle $\measuredangle(v, \dot{AC}) \in [0, \pi]$. Then we have

$$\measuredangle(P_A, \dot{AC}) := \min\{\measuredangle(v, \dot{AC}) : v \in P_A \text{ nonzero}\} \in [0, \pi/2].$$

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

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 $\mathcal{C}$, 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 $\mathcal{C}$ (but not on $|AC|$).

In the general case of Proposition 4.1, however, we make assumptions on $d(P_A, {}_{CA}\backslash\backslash P_C)$ from which we are not able to obtain easily bounds on the angle $\measuredangle(\dot{CB}, \dot{AB})$ 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 $T_A M$. 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 $\measuredangle(P_A, \dot{AC}) \geq \pi/2 - \delta$ and $\dot{AC} = \tilde{C} - 0$ 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

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

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

Claim 2.

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

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

$$\measuredangle({}_{CA}\backslash\backslash\dot{CB}, Q - 0) \leq \frac{1}{2} d(A, C).$$

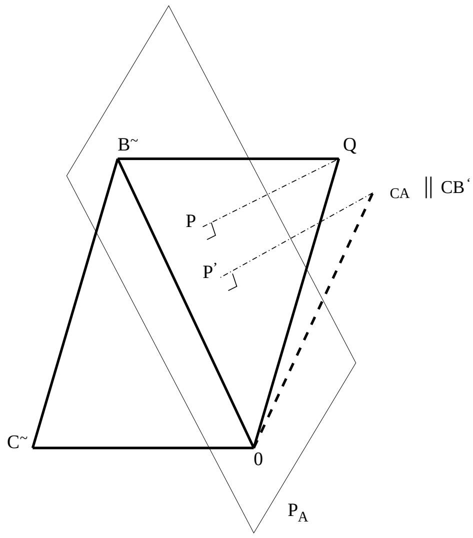
Now let P' be the closest point to ${}_{CA}\backslash\backslash\dot{CB}$ in P_A. As $P' - 0 \in P_A$ and $\dot{CB} \in P_C$, using the definition of distance between subspaces we get

$$\measuredangle({}_{CA}\backslash\backslash\dot{CB}, P' - 0) \leq d(P_A, {}_{CA}\backslash\backslash P_C) = \theta \leq \mathcal{C} d(C, A).$$

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

$$\measuredangle(P - 0, P' - 0) \leq \frac{1}{2} d(C, A).$$

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



Claim 3.

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

The triangle $\tilde{B}PQ$ is a right triangle at P, so using Claims 1 and 2 we have

$$\begin{aligned}|Q-\tilde{B}|^2 = |\tilde{B}-P|^2 + |Q-P|^2 &\leq (1+\tan^2\delta)\cdot|Q-P|^2 \\ &\leq (1+\tan^2\delta)\cdot(1+\mathcal{C})^2\cdot d(C, A)^2\cdot|Q-0|^2.\end{aligned}$$

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

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

and

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

Using standard estimates (see Corollary 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. □

5. Application of Proposition 4.1 to 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 γ in M. Suppose that $\exp_C(v) = \exp_A(w) =: B$ for vertical vectors $v \in \mathrm{Vert}_C^g$ and $w \in \mathrm{Vert}_A^g$ 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 $P_A = \mathrm{Vert}_A^g$, $P_C = \mathrm{Vert}_C^g$. We will do so in Proposition 5.5.

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

Roughly speaking, the constant δ—which measures how much the angle between $\dot{CA} = \dot{\gamma}(0)$ and 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 γ between C and A is "nearly tangent" to the distribution Hor^g.

Bounding the constant $\mathcal{C}$—which measures how the angle between Vert_A^g and 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

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

to determine δ we just have to estimate the angle

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

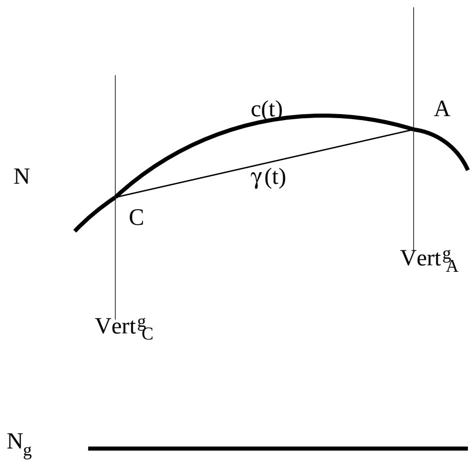
We already introduced the geodesic $\gamma(t)$ from C to A, which we assume to be parametrized by arc length. We now consider the curve $\pi(t) := \pi_{N_g} \circ \gamma(t)$ in N_g. We can lift the curve π to a curve $\varphi_g \circ \pi$ in N connecting C and A; we will call $c(t)$ the parametrization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

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

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 ϵ and $d(C, A)$.



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 $\tilde{c} := \exp_{N_g}^{-1}(c(t))$ of νN_g along π is "approximately parallel". This will allow us to bound from above the "distance" between its endpoints as follows:

Proposition 5.1.

$$|\exp_{N_g}^{-1}(C) - {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash \exp_{N_g}^{-1}(A)| \le L(\gamma)\frac{3150\epsilon}{f(r)}.$$

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

Proposition 5.2.

$$\begin{aligned}|\exp_{N_g}^{-1}(C) - {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash \exp_{N_g}^{-1}(A)|&\\ \ge L(\gamma)\left[\frac{99}{100}\sin\left(\alpha - \frac{\epsilon}{4}\right) - 500\epsilon - 3r - \frac{8}{3}L(\gamma)\left(r + \frac{r+3/2}{f(r)}\right)\right].&\end{aligned}$$

Comparison of Propositions 5.1 and 5.2 gives

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

Recall that $r = 100\epsilon + L(\gamma)/2$. If $L(\gamma)$ and ϵ are small enough one can solve the above inequality for α. 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 δ of Proposition 4.1 in our setting.

Proposition 5.3. *Let C, A be points in N and γ 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*

$$\measuredangle(\mathrm{Hor}^g_C, \dot{\gamma}(0)) = \alpha < \delta(\epsilon, L(\gamma)).$$

Therefore

$$\measuredangle(\mathrm{Vert}^g_C, \dot{\gamma}(0)) \ge \pi/2 - \delta(\epsilon, L(\gamma))$$

and for symmetry reasons

$$\measuredangle(\mathrm{Vert}^g_A, -\dot{\gamma}(L(\gamma))) \ge \pi/2 - \delta(\epsilon, L(\gamma)).$$

To determine the constant $\mathcal{C}$ we only need Lemma C.3:

Proposition 5.4. *Let C, A and γ be as above and assume $L(\gamma) < 0.1$. Then*

$$d(\mathrm{Vert}^g_C, {}_{\gamma^\flat}\backslash\backslash \mathrm{Vert}^g_A) \leq 2L(\gamma).$$

Proof. By Lemma C.3 we have

$$d(\mathrm{Vert}^g_C, {}_{\gamma^\flat}\backslash\backslash \mathrm{Vert}^g_A) \leq \arcsin\left[L(\gamma)\left(r + \frac{r + 3/2}{f(r)}\right)\right],$$

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

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 $d(A, C) < 0.1$ and suppose that $\exp_C(v) = \exp_A(w) =: B$ for vertical vectors $v \in \mathrm{Vert}^g_C$, $w \in \mathrm{Vert}^g_A$. 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 $P_C = \mathrm{Vert}^g_C$ and $P_A = \mathrm{Vert}^g_A$ in the statement of Proposition 4.1, Propositions 5.3 and 5.4 allow us to choose

$$\delta = \delta\,(\epsilon, d(A, C)) \quad \text{and} \quad \mathcal{C} = 2.$$

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

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

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

6. Estimates on tubular neighborhoods of N_g on which φ_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 $\exp_{N_g}(\nu N_g)_{0.05}$ is a tubular neighborhood of N on which φ_g is injective. We will also bound from below the size of $\bigcap_{g \in G} \exp_N(\mathrm{Vert}^g)_{0.05}$ (where the 2-form $\int_g \omega_g$ is defined).

Proposition 6.1. *If* $\epsilon < 1/20000$ *the map*

$$\varphi_g : \exp_{N_g}(\nu N_g)_{0.05} \to \exp_N(\mathrm{Vert}^g)_{0.05}$$

is a diffeomorphism.

Proof. From the definition of φ_g it is clear that it is enough to show the injectivity of

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

Let $A, C \in N$ and $v \in \mathrm{Vert}^g_C$, $w \in \mathrm{Vert}^g_A$ 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 $|v|, |w| \geq \frac{3}{10}\cos(\delta(\epsilon, d(A, C)))$. 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. □

For each $L \leq 0.05$ we want to estimate the radius of a tubular neighborhood of N contained in $\bigcap_{g\in G}\exp_N(\mathrm{Vert}^g)_L$. 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*

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

we have

$$\exp_N(\nu N)_{R^\epsilon_L} \subset \bigcap_{g\in G}\exp_N(\mathrm{Vert}^g)_L.$$

Remark. The function $R^\epsilon_{0.05}$ decreases with ϵ 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* $\exp_{N_g}(\nu N_g)_1$ *such that* $d(A, B) \leq d(C, B) =: L < 0.05$*. Let* γ *denote the angle at* C*, and suppose* $\gamma \in [\pi/2 - \tilde\delta, \pi/2 + \tilde\delta]$*. Then*

$$d(A, B) \geq \sin(L)\cos(\tilde\delta + 2L^2).$$

Proof. Denote by α, β the angles at A and B respectively, and denote further by α', β', γ' 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° 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

$$d(A,B) \geq \sin(d(A,B)) \geq \sin(d(C,B))\sin(\gamma') \geq \sin(L)\sin(\pi/2 + \tilde{\delta} + 2L^2). \quad \square$$

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 $w \in \mathrm{Vert}^g_C$ 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 δ 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 γ from C to A has length $L(\gamma) \leq 2L$. By Proposition 5.3 we have

$$\measuredangle(\dot{\gamma}(0), \mathrm{Vert}^g_C) \geq \pi/2 - \delta\,(\epsilon, L(\gamma)) \geq \pi/2 - \delta(\epsilon, 2L).$$

So, since $w \in \mathrm{Vert}^g_C$,

$$\measuredangle(\dot{\gamma}(0), w) \in [\pi/2 - \delta(\epsilon, 2L), \pi/2 + \delta(\epsilon, 2L)].$$

If we use the fact that, for any $g \in G$, the triangle ABC lies in $\exp_{N_g}(\nu N_g)_1$, the lemma follows from Lemma 6.1 with $\tilde{\delta} = \delta(\epsilon, 2L)$. $\square$

Proof of Proposition 6.2. For any $g \in G$ and positive number $L < 0.05$, by Lemma 6.2 each point $B \in \partial \exp_N(\mathrm{Vert}^g)_L$ has distance at least

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

from N. Therefore $\mathrm{tub}(R^\epsilon_L)$ lies in $\exp_N(\mathrm{Vert}^g)_L$, and since this holds for all g we are done. $\square$

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 $\bigcap_{g\in G} \exp_N(\mathrm{Vert}^g)_{0.05}$ 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 tub^ϵ 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, $[\omega] = [\int_g \omega_g] \in H^2(\mathrm{tub}^\epsilon, \mathbb{R})$, so we can apply Moser's trick. The main step consists of constructing canonically a primitive α 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 tub^ϵ we will determine an ϵ 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

$$D_L^\epsilon := 4L + 4100\epsilon$$

for the upper bound obtained in Proposition 3.4 on $\exp_N(\mathrm{Vert}^g)_L$.

7.1. Symplectic forms in tub^ϵ

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

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

$$|(\varphi_g^{-1})_*X - \backslash\backslash X| \le \frac{D_L^\epsilon}{1 - D_L^\epsilon}.$$

Furthermore,

$$\frac{1}{1 + D_L^\epsilon} \le |(\varphi_g^{-1})_*X| \le \frac{1}{1 - D_L^\epsilon}.$$

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

$$\frac{|\varphi_{g*}(Z)|}{1 + D_L^\epsilon} \le |Z| \le \frac{|\varphi_{g*}(Z)|}{1 - D_L^\epsilon}.$$

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

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

$$|(\varphi_g^{-1})_*X - \backslash\backslash X| \le \frac{|X - \varphi_{g*}\backslash\backslash X|}{1 - D_L^\epsilon} \le \frac{D_L^\epsilon}{1 - D_L^\epsilon}. \qquad \square$$

Since $(\varphi_g^{-1})_* X$ is close to $\backslash\backslash X$ and since our assumption on $\nabla\omega$ allows us to control to what extent ω is invariant under parallel translation we are able to show that ω and $\omega_g = (\varphi_g^{-1})^*\omega$ are close to each other:

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

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

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

$$\begin{aligned}(\omega_g - \omega)_q(X, Y) &= \omega_p((\varphi_g^{-1})_* X, (\varphi_g^{-1})_* Y) - \omega_q(X, Y)\\ &= \omega_p(\backslash\backslash X - [(\varphi_g^{-1})_* X - \backslash\backslash X], \backslash\backslash Y - [(\varphi_g^{-1})_* Y - \backslash\backslash Y]) - \omega_q(X, Y)\\ &= \omega_p((\varphi_g^{-1})_* X - \backslash\backslash X, (\varphi_g^{-1})_* Y - \backslash\backslash Y)\\ &\quad + \omega_p(\backslash\backslash X, (\varphi_g^{-1})_* Y - \backslash\backslash Y) + \omega_p((\varphi_g^{-1})_* X - \backslash\backslash X, \backslash\backslash Y)\\ &\quad + \omega_p(\backslash\backslash X, \backslash\backslash Y) - \omega_q(X, Y).\end{aligned}$$

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

Since the symplectic form ω is compatible with the metric and the ω_g's are close to ω we obtain the nondegeneracy of ω_t for L and ϵ small enough.

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

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

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 ω is almost-Kähler, the norm of the second one is estimated using Lemma 7.2. ◻

Remark. The right hand side of Corollary 7.1 is surely positive if $D_L^\epsilon \leq 0.1$. We set[16]

$$L^\epsilon := \frac{0.1 - 4100\epsilon}{4}$$

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

[16] This choice of L^ϵ will allow us to obtain good numerical estimates in Section 7.4.

Proposition 7.1. *On*

$$\mathrm{tub}^{\epsilon} := \exp_N(\nu N)_{R^{\epsilon}_{L^{\epsilon}}} \subset \bigcap_{g\in G} \exp_N(\mathrm{Vert}^g)_{L^{\epsilon}}$$

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^{ϵ}_L was defined in Proposition 6.2. See Section 7.4 for the graph of $R^{\epsilon}_{L^{\epsilon}}$ a function of ϵ.

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

We want to construct canonically a primitive α of

$$\frac{d}{dt}\omega_t = \int_g \omega_g - \omega$$

on $\bigcap_{g\in G} \exp_N(\mathrm{Vert}^g)_{0.05}$. 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 $E \to N$ be a subbundle of $TM|_N \to N$ such that $E \oplus TN = TM|_N$. Furthermore let $\tilde{U}$ be a fiber-wise convex neighborhood of the zero section of $E \to N$ such that $\exp : \tilde{U} \to U \subset M$ is a diffeomorphism. Denote by $\pi : U \to N$ 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 $Q : \Omega^{\bullet}(U) \to \Omega^{\bullet-1}(U)$ such that

$$\mathrm{Id} - (i\circ\pi)^* = dQ + Qd : \Omega^{\bullet}(U) \to \Omega^{\bullet}(U).$$

A concrete example is given by considering $\rho_t : U \to U,\ \exp_q(v) \mapsto \exp_q(tv)$, and $w_t|_{\rho_t(p)} := \frac{d}{ds}|_{s=t}\rho_s(p)$. Then

$$Qf := \int_0^1 Q_t f\, dt, \qquad Q_t f := \rho_t^*(i_{w_t} f),$$

gives an operator with the above property.

Note that for a 2-form ω evaluated at $X \in T_pM$ we have

$$|(Q_t\omega)_p X| = |\omega_p(w_t|_{\rho_t(p)}, \rho_{t*}(X))| \le |\tilde{\omega}_p|_{\mathrm{op}} \cdot d(p, \pi(p)) \cdot |\rho_{t*}(X)| \qquad (\bigstar)$$

where $|\tilde{\omega}_p|_{\mathrm{op}}$ is the operator norm of $\tilde{\omega}_p : T_pM \to T_p^*M$ 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 α of $\omega_g - \omega$ on $\exp_N(\mathrm{Vert}^g)_{0.05}$. We do that in two steps:

Step I. We apply the above procedure to the vector bundle $\mathrm{Vert}^g \to N$ to obtain an operator Q^g_N such that

$$\mathrm{Id} - (\pi^g_N)^* \circ (i_N)^* = dQ^g_N + Q^g_N d$$

for all differential forms on $\exp_N(\mathrm{Vert}^g)_{0.05}$. Since N is isotropic with respect to ω_g and $\omega_g - \omega$ is closed we have

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

Step II. Now we apply the procedure to the vector bundle $\nu N_g \to N_g$ to get an operator Q_{N_g} on differential forms on $\exp_{N_g}(\nu N_g)_{100\epsilon}$. Since N_g is isotropic with respect to ω we have

$$\omega = dQ_{N_g}\omega,$$

so we have found a primitive of ω on $\exp_{N_g}(\nu N_g)_{100\epsilon}$. Since $N \subset \exp_{N_g}(\nu N_g)_{100\epsilon}$ the 1-form $\beta^g := i_N^*(Q_{N_g}\omega)$ on N is a well defined primitive of $i_N^*\omega$.

Summing up these two steps we see that

$$\alpha^g := Q_N^g(\omega_g - \omega) - (\pi_N^g)^*\beta^g$$

is a primitive of $\omega_g - \omega$ on $\exp_N(\mathrm{Vert}^g)_{0.05}$. So clearly $\alpha := \int_g \alpha^g$ is a primitive of $\frac{d}{dt}\omega_t = \int_g \omega_g - \omega$ on $\bigcap_{g\in G} \exp_N(\mathrm{Vert}^g)_{0.05}$.

7.3. Estimates on the primitive of $\frac{d}{dt}\omega_t$

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

Step II. We will first estimate the norm of $\beta^g := i_N^*(Q_{N_g}\omega)$ using (★) and then the norm of $(\pi_N^g)^*\beta^g$.

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

$$|(\rho_{N_g})_{t*}X| \leq 5/4.$$

Proof. Let $L := d(p, N_g) < 100\epsilon < 1/700$ and write $X = J(L) + K(L)$, where J and K are N_g-Jacobi fields along the unit-speed geodesic γ_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 $f_s(t) = \exp_{p'}(tv(s))$ where the $v(s)$'s are unit normal vectors at p'. Therefore

$$(\rho_{N_g})_{t*}J(L) = \frac{d}{ds}\bigg|_0 [(\rho_{N_g})_t \circ \exp_{p'}(Lv(s))] = \frac{d}{ds}\bigg|_0 [\exp_{p'}(tLv(s))] = J(tL).$$

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

$$|J(tL)| \leq t\frac{1+\sigma(L)}{1-\sigma(tL)}|J(L)| \leq \frac{21}{20}t|J(L)|.$$

Similarly we have $(\rho_{N_g})_{t*}K(L) = K(tL)$. 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

$$\begin{aligned}
|(\rho_{N_g})_{t*}X|^2 &= |J(tL)+K(tL)|^2 \\
&\leq |J(tL)|^2 + |K(tL)|^2 + 2\cdot\frac{9}{5}tL|J(tL)||K(tL)| \\
&\leq \left(\frac{21}{20}\right)^2\left[|J(L)|^2 + |K(L)|^2 + \frac{18}{5}L|J(L)||K(L)|\right] \\
&\leq \left(\frac{21}{20}\right)^2\left[|J(L)+K(L)|^2 + \frac{36}{5}L|J(L)||K(L)|\right] \\
&\leq \left(\frac{21}{20}\right)^2\left[1+\frac{36}{5}\cdot 1.1^2 L\right] \leq \frac{5}{4}
\end{aligned}$$

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* β^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 $|\tilde{\omega}|_{\mathrm{op}} = 1$ and Lemma 7.3, we have $|(Q_{N_g}\omega)_p| < 125\epsilon$. Clearly

$$|(Q_{N_g}\omega)_p| \geq |i_N^*(Q_{N_g}\omega)_p| = |\beta_p^g|. \qquad ◻$$

Now we would like to estimate $(\pi_N^g)_*X$ for a unit tangent vector X. Since $\pi_N^g = (\rho_N^g)_0$ we prove a stronger statement that will be used again later. Recall that we assume $L \leq 0.05$.

Lemma 7.4. *If* $q \in \exp_N(\mathrm{Vert}^g)_L$ *and* $X \in T_qM$ *is a unit vector then for any* $t \in [0,1]$ *we have*

$$|(\rho_N^g)_{t*}X| \leq 1.5\frac{1+D_L^\epsilon}{1-D_L^\epsilon}.$$

Proof. Using Lemma 7.1 we have

$$|(\rho_N^g)_{t*}X| \leq (1+D_L^\epsilon)|(\varphi_g^{-1})_*(\rho_N^g)_{t*}X|.$$

Clearly $(\rho_N^g)_t \circ \varphi_g = \varphi_g \circ (\rho_{N_g})_t$, since—up to exponentiating—φ_g maps νN_g to Vert^g, and $(\rho_{N_g})_t$ and $(\rho_N^g)_t$ 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 $\exp_{N_g}(\nu N_g)_L$ we obtain[17] $|(\rho_{N_g})_{t*}Y| < 1.5$ for unit vectors Y at p. Using this and Lemma 7.1 respectively we have

$$|(\rho_{N_g})_{t*}(\varphi_g^{-1})_* X| \leq 1.5|(\varphi_g^{-1})_* X| \quad \text{and} \quad |(\varphi_g^{-1})_* X| \leq \frac{1}{1 - D_L^\epsilon}.$$

Altogether this proves the lemma. □

Corollary 7.3. *On* $\exp_N(\mathrm{Vert}^g)_L$ *we have*

$$|(\pi_N^g)^* \beta^g| \leq 200\epsilon \frac{1 + D_L^\epsilon}{1 - D_L^\epsilon}.$$

Proof. This is clear from the equality $|((\pi_N^g)^* \beta^g)X| = |\beta^g((\pi_N^g)_* X)|$, Corollary 7.2 and Lemma 7.4. □

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

Corollary 7.4. *For* $q \in \partial \exp_N(\mathrm{Vert}^g)_L$ *we have*

$$|Q_N^g(\omega_g - \omega)_q| \leq 1.5 \frac{1 + D_L^\epsilon}{1 - D_L^\epsilon} \cdot L \cdot \left[\frac{D_L^\epsilon}{1 - D_L^\epsilon} \left(\frac{D_L^\epsilon}{1 - D_L^\epsilon} + 2 \right) + 2L + 100\epsilon \right].$$

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

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* $q \in \bigcap_g \exp_N(\mathrm{Vert}^g)_L$ *we have*

$$|\alpha_q| \leq 1.5 \frac{1 + D_L^\epsilon}{1 - D_L^\epsilon} \cdot \frac{3}{2} d(q, N) \cdot \left[\frac{D_L^\epsilon}{1 - D_L^\epsilon} \left(\frac{D_L^\epsilon}{1 - D_L^\epsilon} + 2 \right) + 2L + 100\epsilon \right] + 200\epsilon \frac{1 + D_L^\epsilon}{1 - D_L^\epsilon}.$$

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 $\mathrm{tub}^\epsilon \subset \bigcap_{g \in G} \exp_N(\mathrm{Vert}^g)_{L^\epsilon}$. Recalling that $D_{L^\epsilon}^\epsilon = 0.1$, Corollary 7.1 immediately implies

Corollary 7.5. *At* $q \in \mathrm{tub}^\epsilon \subset \bigcap_g \exp_N(\mathrm{Vert}^g)_{L^\epsilon}$ *we have*

$$|(\tilde{\omega}_t)_q^{-1}|_{\mathrm{op}} \leq \frac{1}{1 - \left[\frac{D_{L^\epsilon}^\epsilon}{1 - D_{L^\epsilon}^\epsilon} \left(\frac{D_{L^\epsilon}^\epsilon}{1 - D_{L^\epsilon}^\epsilon} + 2 \right) + 2L^\epsilon + 100\epsilon \right]} \leq 1.53.$$

[17] 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* $q \in \text{tub}^{\epsilon}$,

$$|(v_t)_q| \leq |(\tilde{\omega}_t)_q^{-1}|_{\text{op}} \cdot |\alpha_q| \leq 1.45\, d(q, N) + 374\epsilon.$$

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

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

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

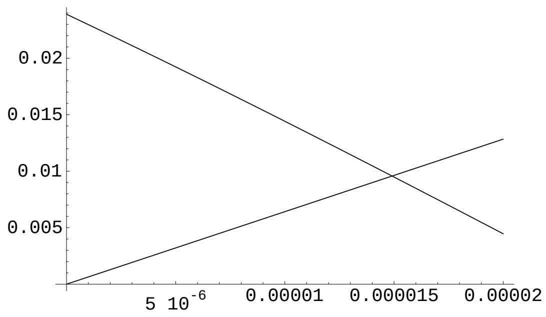
The solution of the ODE $\dot{s}(t) = As(t) + B$ satisfying $s(0) = 0$ is $\frac{B}{A}(e^{At} - 1)$. Hence, if the integral curve γ is well defined at time 1, we have

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

Let us denote by ρ_1 the time-1 flow of the time-dependent vector field v_t, so that ρ_1^{-1} is the time-1 flow of $-v_{1-t}$. Since by definition $\text{tub}^{\epsilon} := \exp_N(\nu N)_{R^{\epsilon}_{L^{\epsilon}}}$ the submanifold $L := \rho_1^{-1}(N)$ will surely be well defined if

$$842\epsilon < R^{\epsilon}_{L^{\epsilon}}.$$

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



Graphs of 842ϵ (increasing) and $R^{\epsilon}_{L^{\epsilon}}$ (decreasing).

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 ϵ 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

$$\epsilon < \frac{1}{|\nabla\omega|}\frac{1}{70000}$$

and replace L^ϵ everywhere by

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

Then the bounds on $|(\tilde{\omega}_t)_q^{-1}|_{\text{op}}$ and $|(v_t)_q|$ 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 ϵ 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 $T_{\rho_1(p)}N$. Indeed, this would allow us to estimate the distance between T_pL and $T_{\pi_{N_g}(p)}N_g$, using which—when ϵ is small enough—one can conclude that $\pi_{N_g} : L \to N_g$ 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 $T_{\rho_1(p)}N$ 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 $\nabla_Y((\varphi_g^{-1})_*X)$ 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 ω, 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 ω.

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 ϵ would depend on the geometry of N.

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 ω. 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 $d^*(\triangle^{-1} i_N^*(\omega_g - \omega))$ 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 ω_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, ω) by symplectomorphisms. A *moment map* for the action is a map $J : M \to \mathfrak{g}^*$ such that for all $v \in \mathfrak{g}$ we have $\omega(v_M, \cdot) = d\langle J, v\rangle$ and which is equivariant with respect to the G-action on M and the coadjoint action of G on $\mathfrak{g}$. 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, ω) with moment map J. Let L be a connected isotropic submanifold of (M, ω) which is invariant under the group action. Then $L \subset J^{-1}(\mu)$ where μ is a fixed point of the coadjoint action.*

Proof. Let $X \in T_x L$. For each $v \in \mathfrak{g}$ we have

$$d_x\langle J, v\rangle X = \omega(v_M(x), X) = 0,$$

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 $\mu \in \mathfrak{g}^*$.

Now let $x_0 \in L$ and let $G \cdot x_0 \subset L$ 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 μ 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, ω) 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 μ of the coadjoint action. Here $\mathfrak{g}^*$ is endowed with any inner product and $C := \max\{|v_M| : v \in \mathfrak{g}$ 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 μ 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 γ 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

$$\langle J(p) - \mu, v\rangle = \int_0^1 \langle dJ(\gamma(t))\dot\gamma(t), v\rangle\, dt = \int_0^1 d\langle J, v\rangle\dot\gamma(t)\, dt$$
$$= \int_0^1 \omega(v_M, \dot\gamma(t))\, dt.$$

Since for all t we have $|\omega(v_M, \dot\gamma(t))| \leq |v_M| \cdot |\dot\gamma(t)| \leq 1000\epsilon \cdot C$ we are done. □

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 γ is a geodesic parametrized by arc length and $W \in T_{\gamma(0)}M$, then for $t < 0.7$,*

$$|{}_\gamma\backslash\backslash(d_{t\dot\gamma(0)} \exp_{\gamma(0)})W - W| \leq \frac{\sinh(t) - t}{t}|W| \leq \frac{t^2}{5}|W|$$

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 γ such that $J(0) = 0$ and $J'(0) = W$ is given by $J(t) = (d_{t\dot\gamma(0)} \exp_{\gamma(0)})(tW)$ (see [Jo, Cor. 4.2.2]). The bound $(\sinh(t) - t)/t$ follows from [BK, 6.3.8iii]. 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. □

Corollary A.1.

$$\measuredangle(Q - 0, {}_{CA}\backslash\backslash\dot{CB}) < \frac{1}{2} d(C, A).$$

Proof. By [BK, 6.6.1] (choosing $v = \tilde{C} - 0$ and $w = {}_{CA}\backslash\backslash\dot{CB}$) we get

$$\begin{aligned} & d(\exp_A((\tilde{C} - 0) + {}_{CA}\backslash\backslash\dot{CB}), \exp_C \dot{CB} = B) \\ & \quad \le \frac{1}{3} d(A, C) \cdot d(C, B) \cdot \sinh(d(A, C) + d(C, B)) \cdot \sin(\measuredangle(\tilde{C} - 0, {}_{CA}\backslash\backslash\dot{CB})) \\ & \quad \le \frac{1}{3} d(A, C) d(C, B) \end{aligned}$$

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

In order to estimate distances in $T_A M$ (instead of in M) we denote the shortest geodesic from $\exp_A((\tilde{C} - 0) + {}_{CA}\backslash\backslash\dot{CB})$ to B by τ, and by $\tilde{\tau}$ its image under $\exp_A^{-1}$. By Lemma A.1,

$$\frac{9}{10} |\dot{\tilde{\tau}}(s)| \le \left(1 - \frac{d(\tau(s), A)^2}{5}\right) |\dot{\tilde{\tau}}(s)| \le |\dot{\tau}(s)|$$

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

$$\begin{aligned} d(\exp_A((\tilde{C} - 0) + {}_{CA}\backslash\backslash\dot{CB}), B) &\ge \frac{9}{10} |(\tilde{C} - 0) + {}_{CA}\backslash\backslash\dot{CB} - (\tilde{B} - 0)| \\ &= \frac{9}{10} |{}_{CA}\backslash\backslash\dot{CB} - (\tilde{B} - \tilde{C})|. \end{aligned}$$

Altogether, since $\tilde{B} - \tilde{C} = Q - 0$, we obtain $|{}_{CA}\backslash\backslash\dot{CB} - (Q - 0)| \le \frac{2}{5} d(A, C) d(C, B)$, and using $\frac{\sin x}{x} \ge \frac{7}{8}$ for $x \in [0, 0.8]$, we obtain

$$\measuredangle({}_{CA}\backslash\backslash\dot{CB}, Q - 0) \le \frac{8}{7} \sin(\measuredangle({}_{CA}\backslash\backslash\dot{CB}, Q - 0)) \le \frac{1}{2} d(C, A). \qquad \square$$

Corollary A.2.

$$\measuredangle(P - 0, P' - 0) \le \frac{1}{2} d(C, A).$$

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 ${}_{CA}\backslash\backslash\dot{CB}$ and Q respectively,

$$|P - P'| \le |(Q - 0) - {}_{CA}\backslash\backslash\dot{CB}| \le \frac{2}{5} d(A, C) d(C, B),$$

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

On the other hand, we have

$$\begin{aligned} |P' - 0| &= |\dot{CB}| \cdot \cos(\measuredangle({}_{CA}\backslash\backslash\dot{CB}, P_A)) \ge |\dot{CB}| \cdot \cos(\theta) \\ &\ge |\dot{CB}| \sqrt{1 - \theta^2} \ge |\dot{CB}| \sqrt{1 - \mathcal{C}^2 d(C, A)^2}. \end{aligned}$$

Therefore we have

$$\sin(\measuredangle(P'-0,\,P-0)) \le \frac{|P'-P|}{|P'-0|} \le \frac{2}{5}\frac{1}{\sqrt{1-\mathcal{C}^2 d(C,A)^2}}d(C,A).$$

So, using the restrictions $\mathcal{C} \le 2$ and $d(C,A) < 0.15$, and using $\frac{\sin x}{x} \ge \frac{7}{8}$ for $x \in [0, 0.8]$, we obtain

$$\measuredangle(P'-0,\,P'-P) \le \frac{8}{7}\sin(\measuredangle(P'-0,\,P'-P)) \le \frac{1}{2}d(C,A). \qquad \square$$

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 α using the curve c

Here we will prove Proposition 5.1, namely the estimate

$$|\exp_{N_g}^{-1}(C) - {}^{\perp}_{\pi^b}\!\!\setminus\!\!\setminus \exp_{N_g}^{-1}(A)| \le L(\gamma)\frac{3150\epsilon}{f(r)}.$$

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 π_c is just a reparametrization of π. We will use exp as a short-hand notation for the normal exponential map $\exp_{N_g} : (\nu N_g)_1 \to \exp_{N_g}(\nu N_g)_1$. Therefore $\tilde{c}(t) := \exp^{-1}(c(t))$ will be a section of νN_g along π_c. The image under $\exp_*$ of the Ehresmann connection corresponding to $\nabla^{\perp}$ will be the subbundle LC^g of $TM|_{\exp_{N_g}(\nu N_g)_1}$. To simplify notation we will denote by $\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{Hor}^g$ the projection of $\dot\gamma(t) \in T_{\gamma(t)}M$ onto $\mathrm{Vert}^g_{\gamma(t)}$ along $\mathrm{Hor}^g_{\gamma(t)}$. We will also use $\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{aHor}^g$ and $\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{LC}^g$ to denote projections onto $\mathrm{aVert}^g_{\gamma(t)}$ along $\mathrm{aHor}^g_{\gamma(t)}$ and $\mathrm{LC}^g_{\gamma(t)}$ respectively.

Our strategy will be to bound above $|\frac{\nabla^{\perp}}{dt}\tilde{c}(t)| = |\exp_*^{-1}(\mathrm{pr}_{\dot c(t)}\,\mathrm{LC}^g)|$ (see Lemma B.3) using

$$TN \approx \mathrm{Hor}^g \approx \mathrm{aHor}^g \approx \mathrm{LC}^g.$$

Integration along π_c will give the desired estimate.

The estimates to make precise $TN \approx \mathrm{Hor}^g$ and $\mathrm{Hor}^g \approx \mathrm{aHor}^g$ were derived in [We]. In the next two lemmata we will do the same for $\mathrm{aHor}^g \approx \mathrm{LC}^g$.

Lemma B.1. *If* $L < 0.08$ *and* p *is a point in* $\partial \exp_{N_g}(\nu N_g)_L$, *then*

$$d(\mathrm{aHor}^g_p, \mathrm{LC}^g_p) \le \arcsin\left(\frac{9}{5}L\right).$$

Proof. It is enough to show that, if $Y \in \mathrm{LC}^g_p$ is a unit vector, then

$$|\mathrm{pr}_Y\, \mathrm{aHor}^g| \le \frac{9}{5}L.$$

Let $\beta(s)$ be a curve tangent to the distribution LC^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 ξ of νN_g along the curve $\gamma(s) := \pi_{N_g}(\beta(s))$. If we denote by K the N_g-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 ξ is a strong Jacobi field (see the remark in Section 3.2): we have $\frac{\partial}{\partial t}|_0 f_s(t) = \xi(s)$, so

$$K'(0) = \frac{\nabla}{dt}\bigg|_0 \frac{\partial}{\partial s}\bigg|_0 f_s(t) = \frac{\nabla}{ds}\bigg|_0 \xi(s) = \frac{\nabla^\perp}{ds}\bigg|_0 \xi(s) - A_{\xi(0)}\dot\gamma(0) = -A_{\xi(0)}K(0).$$

The claim follows since $\xi(0) = \dot\gamma_p(0)$, where γ_p denotes the unique geodesic parametrized by arc length connecting $\pi_{N_g}(p)$ to p.

Now let us denote by J the N_g-Jacobi field along γ_p vanishing at 0 such that $J(L) = \mathrm{pr}_Y\, \mathrm{aHor}^g \in \mathrm{aVert}^g_p$. By Lemma 3.9, using the fact that Y is a unit vector, we have

$$|\mathrm{pr}_Y\, \mathrm{aHor}^g|^2 = \langle \mathrm{pr}_Y\, \mathrm{aHor}^g, Y\rangle = |\langle J(L), K(L)\rangle| \le \frac{9}{5}L \cdot |\mathrm{pr}_Y\, \mathrm{aHor}^g|$$

and we are done. □

Lemma B.2. *Let* $L < 0.08$. *For any point* p *in* $\partial \exp_{N_g}(\nu N_g)_L$ *the projections* $T_pM \to \mathrm{aVert}^g_p$ *along* aHor^g_p *and* LC^g_p *differ at most by* $2L$ *in the operator norm.*

Proof. Let $\phi : \mathrm{aHor}^g_p \to \mathrm{aVert}^g_p$ be the linear map whose graph is LC^g_p. Let $X \in T_pM$ be a unit vector and write $X = X_{\mathrm{ah}} + X_{\mathrm{av}}$ for the decomposition of X into almost horizontal and almost vertical vectors. Then $X = (X_{\mathrm{ah}} + \phi(X_{\mathrm{ah}})) + (X_{\mathrm{av}} - \phi(X_{\mathrm{ah}}))$ is the decomposition with respect to the subspaces LC^g_p and aVert^g_p. The difference of the two projections onto aVert^g_p maps X to $\phi(X_{\mathrm{ah}})$. Now

$$|\phi(X_{\mathrm{ah}})| \le |\phi|_{\mathrm{op}} \le \tan(d(\mathrm{aHor}^g_p, \mathrm{LC}^g_p)) \le \frac{\frac{9}{5}L}{\sqrt{1 - (\frac{9}{5}L)^2}} < 2L,$$

where we used [We, Cor. A.5] in the second inequality and Lemma B.1 in the third one. □

Now we are ready to bound the covariant derivative of $\tilde c(t)$:

Lemma B.3. *For all t,*

$$\left|\frac{\nabla^\perp}{dt}\tilde{c}(t)\right| \le 2702\epsilon.$$

Proof. Let

$$\widehat{\frac{\nabla^\perp}{dt}\tilde{c}(t)}$$

denote $\frac{\nabla^\perp}{dt}\tilde{c}(t) \in \nu_{\pi_c(t)}N_g$ but considered as an element of $T_{\tilde{c}(t)}(\nu_{\pi_c(t)}N_g)$. First notice that, by definition, $\widehat{\frac{\nabla^\perp}{dt}\tilde{c}(t)}$ is the image of $\dot{\tilde{c}}(t)$ under the projection $T_{\tilde{c}(t)}(\nu N_g) \to T_{\tilde{c}(t)}(\nu_{\pi_c(t)}N_g)$ along the Ehresmann connection on νN_g corresponding to $\nabla^\perp$. Therefore, since $\exp_*$ maps the Ehresmann connection to LC^g and tangent spaces to the fibers of νN_g to aVert^g, we have

$$\exp_*\left(\widehat{\frac{\nabla^\perp}{dt}\tilde{c}(t)}\right) = \mathrm{pr}_{\dot{c}(t)}\,\mathrm{LC}^g.$$

Notice that here $\exp_*$ denotes $d_{\tilde{c}(t)}\exp_{N_g}$.

The fact that N is C^1-close to N_g (see Theorem 4) implies $\measuredangle(\dot{c}(t), \mathrm{Hor}^g_{c(t)}) \le 2500\epsilon$ since $\dot{c}(t) \in T_{c(t)}N$. By [We, Prop. 3.7], $d(\mathrm{Hor}^g_{c(t)}, \mathrm{aHor}^g_{c(t)}) \le \epsilon/4$ since $d(c(t), N_g) \le 100\epsilon$. Therefore $\measuredangle(\dot{c}(t), \mathrm{aHor}^g_{c(t)}) \le 2501\epsilon$ and $|\mathrm{pr}_{\dot{c}(t)}\,\mathrm{aHor}^g| \le \sin(2501\epsilon) \le 2501\epsilon$.

On the other hand, by Lemma B.2, $|\mathrm{pr}_{\dot{c}(t)}\,\mathrm{aHor}^g - \mathrm{pr}_{\dot{c}(t)}\,\mathrm{LC}^g| \le 200\epsilon$. The triangle inequality therefore gives $|\mathrm{pr}_{\dot{c}(t)}\,\mathrm{LC}^g| \le \sin(2701\epsilon)$. Therefore, using Lemma A.1 and $\epsilon < 1/20000$,

$$|\exp_*^{-1}(\mathrm{pr}_{\dot{c}(t)}\,\mathrm{LC}^g)| \le \frac{1}{1-\epsilon/5}|\mathrm{pr}_{\dot{c}(t)}\,\mathrm{LC}^g| \le 2702\epsilon. \qquad \square$$

Lemma B.3 allows us to bound $|\exp^{-1}(C) - {}^{\perp}_{\pi^{\flat}}\backslash\backslash\exp^{-1}(A)|$ 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 ϵ and $d(C, A)$.

Lemma B.4.

$$L(c) \le \frac{1+3200\epsilon}{f(r)}L(\gamma).$$

Proof. Since $\varphi_{g*}(\tau_{N_{g*}}\dot{c}(t)) = \dot{c}(t)$, by Proposition 3.1 we have $|\dot{c}(t) - \backslash\backslash(\pi_{N_g})_*\dot{c}(t)| \le 3200\epsilon|(\pi_{N_g})_*\dot{c}(t)|$, so

$$|\dot{c}(t)| \le (1+3200\epsilon)|(\pi_{N_g})_*\dot{c}(t)|.$$

Since $L(\pi_{N_g} \circ c) = L(\pi)$, it follows that $L(c) \le (1+3200\epsilon)L(\pi)$. By [We, Lemma 3.3] we have $f(r)L(\pi) \le L(\gamma)$ and we are done. $\square$

Proof of Proposition 5.1. We have

$$|\exp^{-1}(C) - {}^{\perp}_{\pi^{\flat}}\!\!\setminus\!\!\setminus \exp^{-1}(A)| = \left| \int_0^{L(c)} \frac{d}{dt} {}^{\perp}_{\pi^{\flat}_c}\!\!\setminus\!\!\setminus \tilde{c}(t)\, dt \right| = \left| \int_0^{L(c)} {}^{\perp}_{\pi^{\flat}_c}\!\!\setminus\!\!\setminus \frac{\nabla^{\perp}}{dt} \tilde{c}(t)\, dt \right|$$
$$\leq 2702\epsilon L(c) \leq 2702\epsilon \frac{1+3200\epsilon}{f(r)} L(\gamma)$$

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

C. A lower bound for α using the curve γ

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

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

We will use the fact that N_g has bounded second fundamental form (see the first statement of Lemma C.3) and that γ 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 $\tilde{\gamma}(t) := \exp^{-1}_{N_g}(\gamma(t))$ is a section of νN_g along π.

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

Lemma C.1. *We have*

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

Proof. Analogously to the proof of Lemma B.3 we have $\exp_*(\widehat{\frac{\nabla^{\perp}}{dt}\tilde{\gamma}(0)}) = \mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{LC}^g$, where $\widehat{\frac{\nabla^{\perp}}{dt}\tilde{\gamma}(0)}$ is an element of $T_{\tilde{\gamma}(0)}\nu_{\pi(0)}N_g$.

By [We, Prop. 3.7] we have $d(\mathrm{Hor}^g_C, \mathrm{aHor}^g_C) \leq \epsilon/4$. So

$$\measuredangle(\dot{\gamma}(0),\ \mathrm{aHor}^g_C) \geq \measuredangle(\dot{\gamma}(0), \mathrm{Hor}^g_C) - d(\mathrm{Hor}^g_C, \mathrm{aHor}^g_C) \geq \alpha - \epsilon/4.$$

Therefore $|\mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{aHor}^g| \geq \sin(\alpha - \epsilon/4)$.

On the other hand, by Lemma B.2, $|\mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{aHor}^g - \mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{LC}^g| \leq 200\epsilon$. The inverse triangle inequality gives

$$|\mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{LC}^g| \geq \sin(\alpha - \epsilon/4) - 200\epsilon.$$

Applying $\exp^{-1}_*$, by Lemma A.1 we have

$$|\exp^{-1}_*(\mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{LC}^g)| \geq \frac{1}{1+\epsilon/5}|\mathrm{pr}_{\dot{\gamma}(0)}\,\mathrm{LC}^g|,$$

and since $\frac{1}{1+\epsilon/5} \geq \frac{99}{100}$ we are done. □

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 π 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 $\frac{\nabla^\perp}{dt}\tilde\gamma(0)$ and $\exp_*^{-1}\circ{}_{\gamma^b}\!\backslash\!\backslash\circ\exp_*(\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)})$, i.e. the parallel translate of $\frac{\nabla^\perp}{dt}\tilde\gamma(t)$ "along γ", are close for all t. Here $\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)}$ denotes the vector $\frac{\nabla^\perp}{dt}\tilde\gamma(t)$ regarded as an element of $T_{\tilde\gamma(t)}(\nu_{\pi(t)}N_g)$. To this end we show that

$$\mathrm{pr}_{\dot\gamma(0)}\,\mathrm{LC}^g\approx\mathrm{pr}_{\dot\gamma(0)}\,\mathrm{Hor}^g\approx\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{Hor}^g\approx\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{LC}^g,$$

where we identify tangent spaces by parallel translation along γ. The crucial step is the second "≈", where we use the fact that γ is a geodesic. Applying $\exp_*^{-1}$ will easily imply Corollary C.1 since $\exp_*^{-1}(\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{LC}^g)=\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)}$.

Lemma C.2. *For any $L<1$ and any point $p\in\exp_{N_g}(\nu N_g)_L$ the orthogonal projections $T_pM\to\mathrm{aHor}_p^g$ and $T_pM\to\mathrm{Hor}_p^g$ 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,*

$$d(\mathrm{Vert}_C^g,{}_{\gamma^b}\!\backslash\!\backslash\mathrm{Vert}_{\gamma(t)}^g)\le\arcsin\left[t\left(r+\frac{r+3/2}{f(r)}\right)\right].$$

Furthermore,

$$|\mathrm{pr}_{\dot\gamma(0)}\,\mathrm{Hor}^g-{}_{\gamma^b}\!\backslash\!\backslash\,\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{Hor}^g|\le t\left(r+\frac{r+3/2}{f(r)}\right).$$

Proof. We first want to estimate $d(\mathrm{Vert}_C^g,{}_{\gamma^b}\!\backslash\!\backslash\mathrm{Vert}_{\gamma(t)}^g)$. Let $v\in\nu_CN_g$ be a normal unit vector.

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

$$|{}_\pi\!\backslash\!\backslash v-{}^\perp_\pi\!\backslash\!\backslash v|\le\frac{3}{2}L(\pi|_{[0,t]})\le\frac{3}{2}\frac{t}{f(r)}.$$

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 $f(r)L(\pi|_{[0,t]})\le L(\gamma|_{[0,t]})$, which follows from [We, Lemma 3.3].

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

$$|{}_{\tau_0}\!\backslash\!\backslash v-{}_{\gamma^b}\!\backslash\!\backslash\circ{}_{\tau_t}\!\backslash\!\backslash\circ{}_\pi\!\backslash\!\backslash v|\le rt\left(1+\frac{1}{f(r)}\right).$$

Indeed, the above expression just measures the holonomy as one goes once around the polygonal loop given by the geodesics τ_0^b,π,τ_t and γ^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 $L(\tau_t)$, $L(\tau_0) \leq r$, $L(\gamma|_{[0,t]}) = t$ and, as we just saw, $L(\pi|_{[0,t]}) \leq t/f(r)$.

Together this gives

$$|{}_{\tau_0}\backslash\!\backslash v - {}_{\gamma^\flat}\backslash\!\backslash \circ {}_{\tau_t}\backslash\!\backslash \circ {}^{\perp}_{\pi}\backslash\!\backslash v| \leq |{}_{\tau_0}\backslash\!\backslash v - {}_{\gamma^\flat}\backslash\!\backslash \circ {}_{\tau_t}\backslash\!\backslash \circ {}_{\pi}\backslash\!\backslash v| + |{}_{\gamma^\flat}\backslash\!\backslash \circ {}_{\tau_t}\backslash\!\backslash \circ [{}_{\pi}\backslash\!\backslash v - {}^{\perp}_{\pi}\backslash\!\backslash v]| \leq t\left(r + \frac{r+3/2}{f(r)}\right).$$

So we obtain a bound on the distance from ${}_{\tau_0}\backslash\!\backslash v \in \mathrm{Vert}^g_C$ to a unit vector in ${}_{\gamma^\flat}\backslash\!\backslash \mathrm{Vert}^g_{\gamma(t)}$. This yields the first statement of the lemma. The second follows from [We, Prop. A.4], since ${}_{\gamma^\flat}\backslash\!\backslash\, \mathrm{pr}_{\dot\gamma(t)} \mathrm{Hor}^g = \mathrm{pr}_{\dot\gamma(0)}({}_{\gamma^\flat}\backslash\!\backslash\, \mathrm{Hor}^g_{\gamma(t)})$ because γ is a geodesic. □

Corollary C.1. *For all t,*

$$\left|\exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_*\left(\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)}\right) - \frac{\nabla^\perp}{dt}\tilde\gamma(0)\right| \leq \frac{51}{50}\left[2.1(100\epsilon + r) + t\left(r + \frac{r+3/2}{f(r)}\right)\right].$$

Proof. From Lemmata C.2 and B.2 we have, for all t,

$$|\mathrm{pr}_{\dot\gamma(t)} \mathrm{Hor}^g - \mathrm{pr}_{\dot\gamma(t)} \mathrm{LC}^g| \leq |\mathrm{pr}_{\dot\gamma(t)} \mathrm{Hor}^g - \mathrm{pr}_{\dot\gamma(t)} \mathrm{aHor}^g| + |\mathrm{pr}_{\dot\gamma(t)} \mathrm{aHor}^g - \mathrm{pr}_{\dot\gamma(t)} \mathrm{LC}^g| \leq r^2/5 + 2r \leq 2.1r.$$

For $t = 0$, since $d(C, N_g) < 100\epsilon$, we have the better estimate

$$|\mathrm{pr}_{\dot\gamma(0)} \mathrm{Hor}^g - \mathrm{pr}_{\dot\gamma(0)} \mathrm{LC}^g| \leq 210\epsilon.$$

Combining this with the second statement of Lemma C.3 gives

$$|\mathrm{pr}_{\dot\gamma(0)} \mathrm{LC}^g - {}_{\gamma^\flat}\backslash\!\backslash\, \mathrm{pr}_{\dot\gamma(t)} \mathrm{LC}^g| \leq 2.1(100\epsilon + r) + t\left(r + \frac{r+3/2}{f(r)}\right).$$

Recall that $\mathrm{pr}_{\dot\gamma(t)} \mathrm{LC}^g = \exp_*(\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)})$, as in the proof of Lemma B.3. Also, for any vector $X \in T_C M$ we have $|\exp_*^{-1} X| \leq |X|/(1-\epsilon/5)$ by Lemma A.1. So applying $(\exp^{-1})_*$ to $\mathrm{pr}_{\dot\gamma(0)} \mathrm{LC}^g - {}_{\gamma^\flat}\backslash\!\backslash\, \mathrm{pr}_{\dot\gamma(t)} \mathrm{LC}^g$ we get

$$\left|\frac{\nabla^\perp}{dt}\tilde\gamma(0) - \exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_*\left(\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)}\right)\right| \leq \left[2.1(100\epsilon + r) + t\left(r + \frac{r+3/2}{f(r)}\right)\right]\frac{1}{1-\epsilon/5}. \qquad \square$$

Now let ξ be a unit vector in $\nu_{\pi(t)} N_g$. Denote by $\hat\xi$ the same vector thought of as an element of $T_{\tilde\gamma(t)}(\nu_{\pi(t)} N_g)$. In the next two lemmas we want to show that ${}^{\perp}_{\pi^\flat}\backslash\!\backslash \xi$ and $\exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_* \hat\xi \in T_C M$ are close to each other, i.e. that under the identification by exp the $\nabla^\perp$-parallel translation along π and the ∇-parallel translation along γ 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 $\frac{\nabla^\perp}{dt}\tilde\gamma(t)$.

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

$$|{}_{\tau_0^\flat}\backslash\!\backslash \circ {}_{\gamma^\flat}\backslash\!\backslash \circ {}_{\tau_t}\backslash\!\backslash \xi - \exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_* \hat{\xi}| < \frac{r^2}{2}.$$

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

$$\begin{aligned} |{}_{\tau_0^\flat}\backslash\!\backslash[{}_{\gamma^\flat}\backslash\!\backslash \exp_* \hat{\xi}] - \exp_*^{-1}[{}_{\gamma^\flat}\backslash\!\backslash \exp_* \hat{\xi}]| &\le \frac{r^2}{5}|\exp_*^{-1}[{}_{\gamma^\flat}\backslash\!\backslash \exp_* \hat{\xi}]| \\ &\le \frac{r^2}{5}\frac{1}{1-r^2/5}|{}_{\gamma^\flat}\backslash\!\backslash \exp_* \hat{\xi}| \le \frac{r^2}{5}\frac{1+r^2/5}{1-r^2/5}. \end{aligned}$$

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

$$\begin{aligned} |{}_{\tau_0^\flat}\backslash\!\backslash \circ {}_{\gamma^\flat}\backslash\!\backslash[{}_{\tau_t}\backslash\!\backslash \xi] - {}_{\tau_0^\flat}\backslash\!\backslash \circ {}_{\gamma^\flat}\backslash\!\backslash[\exp_* \hat{\xi}]| + |{}_{\tau_0^\flat}\backslash\!\backslash \circ [{}_{\gamma^\flat}\backslash\!\backslash \exp_* \hat{\xi}] - \exp_*^{-1}[{}_{\gamma^\flat}\backslash\!\backslash \exp_* \hat{\xi}]| \\ \le \frac{r^2}{5} + \frac{r^2}{5}\frac{1+r^2/5}{1-r^2/5} \le r^2 \frac{2}{5(1-r^2/5)}. \end{aligned}$$ □

Lemma C.5.

$$|\exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_* \hat{\xi} - {}^{\perp}_{\pi^\flat}\backslash\!\backslash \xi| \le \frac{r^2}{2} + t\left(r + \frac{r+3/2}{f(r)}\right).$$

Proof. The left hand side is bounded above by

$$\begin{aligned} |\exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_* \hat{\xi} - {}_{\tau_0^\flat}\backslash\!\backslash \circ {}_{\gamma^\flat}\backslash\!\backslash \circ {}_{\tau_t}\backslash\!\backslash \xi| + |{}_{\tau_0^\flat}\backslash\!\backslash \circ {}_{\gamma^\flat}\backslash\!\backslash \circ {}_{\tau_t}\backslash\!\backslash \xi - {}_{\pi^\flat}\backslash\!\backslash \xi| + |{}_{\pi^\flat}\backslash\!\backslash \xi - {}^{\perp}_{\pi^\flat}\backslash\!\backslash \xi| \\ \le \frac{r^2}{2} + rt\left(1 + \frac{1}{f(r)}\right) + \frac{3}{2}\frac{t}{f(r)}. \end{aligned}$$

The first term is estimated by Lemma C.4. The second one is just the holonomy as one goes around the loop given by τ_t, $\gamma^\flat$, $\tau_0^\flat$ and π, 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. □

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

$$\left|\exp_*^{-1} \circ {}_{\gamma^\flat}\backslash\!\backslash \circ \exp_*\left(\widehat{\frac{\nabla^\perp}{dt}\tilde{\gamma}(t)}\right) - {}^{\perp}_{\pi^\flat}\backslash\!\backslash \frac{\nabla^\perp}{dt}\tilde{\gamma}(t)\right| \le \frac{2}{3}r^2 + \frac{4}{3}t\left(r + \frac{r+3/2}{f(r)}\right).$$

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

$$\left|\widehat{\frac{\nabla^\perp}{dt}\tilde{\gamma}(t)}\right| = |\exp_*^{-1}(\mathrm{pr}_{\dot{\gamma}(t)}\,\mathrm{LC}^g)| \le \frac{1}{1-r^2/5}|\mathrm{pr}_{\dot{\gamma}(t)}\,\mathrm{LC}^g|$$

by Lemma A.1, and

$$|\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{LC}^g| \leq |\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{LC}^g - \mathrm{pr}_{\dot\gamma(t)}\,\mathrm{aHor}^g| + |\mathrm{pr}_{\dot\gamma(t)}\,\mathrm{aHor}^g| \leq 2r+1$$

by Lemma B.2. Since $\frac{2r+1}{1-r^2/5} \leq \frac{4}{3}$ for $r \leq 0.08$ we are done. □

Now Corollaries C.1 and C.2 immediately imply that $\tilde\gamma(t)$ "grows at a nearly constant rate":

Corollary C.3. *The section $\tilde\gamma$ satisfies*

$$\left|\frac{\nabla^\perp}{dt}\tilde\gamma(0) - {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash\frac{\nabla^\perp}{dt}\tilde\gamma(t)\right| \leq 3(100\epsilon + r) + \frac{8}{3}t\left(r + \frac{r+3/2}{f(r)}\right).$$

Proof of Proposition 5.2. The estimate of Proposition 5.2 follows from

$$\begin{aligned}|\exp^{-1}(C) - {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash\exp^{-1}(A)| &= \left|\int_0^{L(\gamma)} {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash\frac{\nabla^\perp}{dt}\tilde\gamma(t)\,dt\right| \\ &\geq \left|\int_0^{L(\gamma)}\frac{\nabla^\perp}{dt}\tilde\gamma(0)dt\right| - \left|\int_0^{L(\gamma)}\left(\frac{\nabla^\perp}{dt}\tilde\gamma(0) - {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash\frac{\nabla^\perp}{dt}\tilde\gamma(t)\right)dt\right| \\ &\geq L(\gamma)\cdot\left|\frac{\nabla^\perp}{dt}\tilde\gamma(0)\right| - \int_0^{L(\gamma)}\left|\frac{\nabla^\perp}{dt}\tilde\gamma(0) - {}^{\perp}_{\pi^\flat}\!\backslash\!\backslash\frac{\nabla^\perp}{dt}\tilde\gamma(t)\right|dt\end{aligned}$$

by using Lemma C.1 and Corollary C.3. □

References

[BK] Buser, P., Karcher, H.: Gromov's almost flat manifolds. Astérisque **81** (1981) Zbl 0459.53031 MR 0619537

[Ca] Cannas da Silva, A.: Lectures on Symplectic Geometry. Lecture Notes in Math. 1764, Springer, Berlin (2001) Zbl 1016.53001 MR 1853077

[Ch] Chiang, R.: New Lagrangian submanifolds of $\mathbb{CP}^n$. Int. Math. Res. Not. **2004**, no. 45, 2437–2441 Zbl pre02150555 MR 2076100

[Jo] Jost, J.: Riemannian Geometry and Geometric Analysis. 2nd ed., Springer, Berlin (1998) Zbl 0997.53500 MR 1625976

[Kl] Klingenberg, W.: Riemannian Geometry. 2nd ed., de Gruyter (1995) Zbl 0911.53022 MR 1330918

[Wa] Warner, F.: Extensions of the Rauch comparison theorem to submanifolds. Trans. Amer. Math. Soc. **112**, 341–356 (1966) Zbl 0139.15601 MR 0200873

[We] Weinstein, A.: Almost invariant submanifolds for compact group actions. J. Eur. Math. Soc. **2**, 53–86 (1999) Zbl 0957.53021 MR 1750452