Submanifold averaging in Riemannian, symplectic and contact geometry

by

Marco Zambon

B.S. (University of Cologne) 1999

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor Alan Weinstein, Chair
Professor Charles Pugh
Professor Robert Littlejohn

Spring 2004

The dissertation of Marco Zambon is approved:

Chair Date

Date

Date

University of California, Berkeley

Spring 2004

Submanifold averaging in Riemannian, symplectic and contact geometry

Abstract

Submanifold averaging in Riemannian, symplectic and contact geometry

by

Marco Zambon
Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Alan Weinstein, Chair

In 1999 Alan Weinstein presented a procedure to average a family of submanifolds of a Riemannian manifold. We give an improvement of Weinstein's averaging procedure and further adapt it to the settings of symplectic and contact geometry. More precisely, we develop a construction to average isotropic submanifolds of sympectic manifolds and Legendrian submanifolds of contact manifolds. As one of several applications we show that nearby every isotropic (Legendrian) submanifold which is "almost invariant" under a compact group action we can find one which is really invariant.

Professor Alan Weinstein
Dissertation Committee Chair

Acknowledgements

This is the place to thank teachers, relatives and friends who supported me in many ways during the years spent in Berkeley. First of all I would like to thank Prof. Alan Weinstein for many encouraging and insightful discussions, and for being a model of how to live happily both in and outside of mathematics. I am also indebted to Prof. Hermann Karcher for the substantial advice given on the material of Chapter 2 and a very pleasant visit to Bonn, and to Prof. Gudlaugur Thorbergsson for giving me the foundations of differential geometry.

I thank the fellow students and friends whose advice and collaboration enriched my mathematical experience at Berkeley: among them Marius Crainic, Alfonso Gracia-Saz, Megumi Harada, Tara Holm, Tamás Kálmán, Eli Lebow, Raj Mehta, Xiang Tang, Vish Sankaran, and in particular Chenchang Zhu for her collaboration and Henrique Bursztyn for helping me not lose direction in the world of mathematics.

I would like to thank my relatives, both on this and that side of the Pond, in particular my parents Anne and Antonio, Andrea, and my grandmother Nonna. It is through their encouragement and support – both moral and in concrete terms – that I could write this thesis.

Of course I can not forget the wonderful friends I made in Berkeley: in addition to those mentioned above Márta Abrusán, Yelena Baraz, Kristine, Tasha, Tessie and Willis Carraway, Hermann Cuntz, Ishai Dan-Cohen, Sachin Deshmukh, Yanlei Diao, Polina Dimova, András Ferencz, Virginia Flanagin, Hannah Freed, Piotr Gibas, Alan Hammond, Anne-Céline Lambotte, Rui Li, Kristóf Madarász, Medha Patak, Júlia Patkós, Gábor Pete, Joaquin Rosales, Lela Samniashvili, Elizabeth Scott, Hong Tieu, Dongni Wang, Anne and Von Bing Yap, Roman Yorick, Elizabeth Zacharias, Xiaoyue Zhao, and several others. A role of his own had my friend Alyosha Efros, for his unique friendship, undestanding, care, and for sharing with me so many faults and mistakes from which to learn.

Contents

Chapter 1

Introduction

The main result of [We] is 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 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 Chapter 2 we will present Weinstein's averaging procedure. We will also exhibit results on the shape operators of certain parallel tubes in M, obtained with the help of Hermann Karcher. These results allow us to improve the estimates in Weinstein's theorem and are interesting in their own right.

The main result of this thesis is the adaptation of Weinstein's averaging to the setting of symplectic geometry: given a family of *isotropic* submanifolds $\{N_g\}$ 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

[1]The construction is canonical because it does not involve any arbitrary choice but uses only the Riemannian metric on M.

submanifold L.[2] Our construction depends only on the 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 symplectomorphism we can obtain isotropic submanifolds which are invariant under the action.
In Chapter 3 we will present our isotropic averaging theorem and give an outline of the proof. The proof itself is presented in Chapter 4, with some very technical estimates deferred to the three appendices.

We give two simple applications of the isotropic averaging theorem in Chapter 5. First we show that the image of an almost invariant isotropic submanifold under a compact Hamiltonian action is "small". Then we show that, when a certain technical assumption is satisfied, we can deform almost equivariant symplectomorphisms to equivariant ones. More precisely: given a symplectic action of a compact group G on two symplectic manifolds M_1 and M_2 together with an almost equivariant symplectomorphism $\phi : M_1 \to M_2$, we can apply the averaging procedure to graph$(\phi) \subset M_1 \times M_2$, and *if* the resulting G-invariant submanifold L is a graph, then it will be the graph of a G-equivariant symplectomorphism.

Moving to the setting of contact geometry, in Chapter 6 we give a procedure to average C^1-close Legendrian submanifolds of contact manifolds. As in the Riemannian and symplectic set-ups, we obtain that, whenever a compact group action leaves a Legendrian submanifold almost invariant, there is an invariant Legendrian submanifold nearby.

The proof of our isotropic averaging theorem breaks down if we try to apply it to coisotropic submanifolds of a symplectic manifold. The reason is essentially this: if a submanifold is coisotropic with respect to two distinct symplectic forms, then it is not necessarily coisotropic with respect to their average. 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. Therefore if the space of coisotropic submanifolds were not locally path connected, we would have an obstruction to the averaging

[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.

problem for coisotropic submanifolds. This leads us to analyze the set of coisotropic submanifolds nearby a fixed one in Chapter 7. We can only prove that this set does not have a nice manifold structure. A much deeper analysis of the geometry of coisotropic submanifolds was developed in [OP] by Oh and Park, but the question of whether the space of coisotropic submanifolds is locally path connected or not still remains unanswered.

Chapter 2

Weinstein's averaging and improved error estimates for the shape operators of parallel tubes

In this chapter we will present the improvement of Weinstein's averaging procedure (Theorem 2.3 in [We]) obtained with the help of Hermann Karcher. In the first section we will state the improved theorem (Theorem 2). In Section 2.2 we will give estimates for the shape operator (extrinsic curvature) of tubes about certain submanifolds (Theorem 1), thus improving Proposition 3.11 in [We]. Then in Section 2.3 using these results we will follow Weinstein's proof to obtain the improved theorem.

2.1 The Riemannian averaging theorem

We first recall some definitions from [We] in order to state the averaging theorem in the Riemannian setting.

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 one about N are bounded in absolute value by 1; (iii) the injectivity radius of each point of the above neighborhood is at least 1.

The C^0*-distance* between two submanifolds N, N' of a Riemannian manifold is $d_0(N, N') = \sup\{d(x, N') : x \in N\}$. The *distance between two subspaces* of the same dimension F, F' 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 less than or equal to $\frac{\pi}{2}$, and it is equal to $\frac{\pi}{2}$ 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. (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_xN along the above geodesic segment. The C^1-distance is defined as the supremum of these 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 $d_1(N, N') < \frac{1}{4}$, then $d_1(N', N) < 250 d_1(N, N')$ (see Remark 3.18 in [We]).

Our improvement of Theorem 2.3 in [We] reads[1]:

Theorem 1 *Let M be a Riemannian manifold and $\{N_g\}$ a family of submanifolds of M of the same dimension parametrized in a measurable way by elements of a probability space G, such that all the pairs (M, N_g) are gentle. If $d_1(N_g, N_h) < \epsilon < \frac{1}{20000}$ 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 1: This theorem differs from Theorem 2.3 in [We] only in that there we have the bound $136\sqrt{\epsilon}$ for $d_1(N_g, N)$, whereas here we have a bound linear in ϵ.
Remark 2: 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 Theorem 1 one gets immediately a statement about invariant submanifolds under compact group actions (see Theorem 2.2 of [We]).

[1] We omit the compactness assumption on the N_g's stated in [We] since it is superfluous.

2.2 Estimates for the shape operators of parallel tubes

In this section we will improve Proposition 3.11 and Corollary 3.13 of [We] using Jacobi fields techniques.

Let us start by recalling Weinstein's construction, assuming the set-up of Theorem 1 above. Given the family of C^1-close submanifolds $\{N_g\}$, consider the functions $P_{N_g} = \frac{1}{2}\rho_{N_g}^2$ on M, where ρ_{N_g} is the distance function from the submanifold N_g. Consider the gradient grad $\int_G P_{N_g}$ of the average of the P_{N_g}'s and project it onto the "averaged vertical bundle" obtained by parallel translation from the normal bundles νN_g. Weinstein shows that, near the N_g's, the zero set of the projected gradient forms a smooth submanifold, which he defines to be the average of the family $\{N_g\}$.
The C^1-closeness of the N_g's ensures that all objects used in the construction are well-defined, and the average submanifold is C^1-close to each N_g because of the gentleness of the pairs (M, N_g), which allows us to give bounds on the Hessian of P_{N_g}. Our aim is to improve the estimates on this Hessian using Jacobi fields. Until the end of this section we will fix a submanifold N_g and drop the index "g" in the notation.

The Hessian of P_N is the symmetric endomorphism of each tangent space of the tubular neighborhood given by $H_N(v) = \nabla_v \mathrm{grad} P_N$. Differentiating the relation $\mathrm{grad} P_N = \rho_N \cdot \mathrm{grad}\rho_N$ we see that

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

where $U_N = \mathrm{grad}\rho_N$ 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[2] of the tube given by a level set $\tau(t)$ of ρ_N in direction of the normal vector U_N .
Proposition 3.11 of [We] states that, at a point p of distance $t \leq \frac{1}{4}$ from N, the following estimate holds for the decompositions into vertical and horizontal parts[3] of T_pM:

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

[2]So $S_N v = \mathrm{pr}(\nabla_v U_N)$ for all vectors v tangent to $\tau(t)$, where ∇ is the Levi-Civita connection on M.
[3]See our Section 4.1 or Section 2.1 in [We] for the definition of vertical and horizontal bundle at p.

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 estimates 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 2.2.1 *Let M be any Riemannian manifold, N a submanifold, 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

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

□

Using the above lemma we will be able to prove the following improvement of Proposition 3.11 in [We], for which we don't require (M, N) to be a gentle pair, but only bounds on the second fundamental form[4] B of N and on the curvature of M.

Theorem 2 *Let M be a Riemannian manifold with curvature $|K| \leq 1$, N a submanifold 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. 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 w.r.t. the splitting into vertical and horizontal spaces of $T_{\gamma(t)}\tau(t)$, as long as $t \leq \min\{\frac{1}{2}, \frac{1}{2|B|}\}$, we have*

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

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: 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 4.1 for their definition) satisfying $J_i(0) = E_i$, our J_i are those having smallest derivative at time zero. Also notice that

[4] We adopt the following convention for the second fundamental form: $B_\xi v := (\nabla_v \xi)^T$ for tangent vectors v of N and normal vector fields ξ, where $(\cdot)^T$ denotes projecting to the component tangent to N and ∇ is the Levi-Civita connection on M. In the case that (M, N) is a gentle pair one has $|B| < \frac{3}{2}$, see [We, Cor 3.2].

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[5] geodesics normal to N. The velocity vectors of such variations at time t coincide with U_N. Therefore applying Lemma 2.2.1 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, $\{E_1(t), \dots, E_k(t)\}$ together with $\dot\gamma(t)$ span the vertical space at $\gamma(t)$ and $\{E_{k+1}(t), \dots, E_{n-1}(t)\}$ span the horizontal space there. We will represent the maps $J(t)$,$J'(t)$ and $S_N(t)$ by matrices w.r.t. 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][6] 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.[7] This will allow us to obtain corresponding

[5]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.

[6]See also our Lemma 4.1.1

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

estimates for $J^{-1}(t)$ and $J'(t)$, and therefore for $S_N(t)$.

For all i let us define the vector fields $A_i(t) = \\(J_i(0)+t\cdot J_i'(0))$, where $\\$ 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 $(c_1, \cdots, c_k, 0, \cdots, 0)$ 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 $(0, \cdots, 0, c_{k+1}, \cdots, c_{n-1})$ 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, $J = A \cdot [I + A^{-1}F_1]$, 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 $t \leq \frac{1}{2|B|}$. 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[8] 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 for $A^{-1}F_1$

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

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

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

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 $(c_1, \cdots, c_k, 0, \cdots, 0)$ 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 $(0, \cdots, 0, c_{k+1}, \cdots, c_{n-1})$ 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 delivers

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

So altogether we obtain

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

Now finally we can estimate

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

Here we used

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

in the last inequality. Using 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 gentle pair, so that $|B| \le \frac{3}{2}$ by [We, Cor 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 $\sqrt{t}$ of Corollary 3.13 in [We].

Corollary 2.2.1 *Let M be a Riemannian manifold, 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 $t \le \frac{1}{3}$ from N, then $|\langle H_N(v), w\rangle| \le 16t^2|v||w|$.*

2.3 Improvement of Weinstein's averaging theorem

In this section we use our Corollary 2.2.1 to replace some estimates in [We] that were originally derived using Corollary 3.13 in that paper. We will improve only estimates contained in Lemma 4.7 and Lemma 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[9] with respect to a fixed submanifold N_e. ($\mathcal{V}$ is obtained by a certain

[9] See our Section 4.1 or Section 3.2 in [We] for the definitions of almost horizontal and almost vertical bundle.

projection of grad $\int_G P_{N_g}$). As in [We] all estimates will hold for $\epsilon < \frac{1}{20000}$, and we set $t = 100\epsilon$.

We will replace the constant "$\frac{89}{200}$" in Lemma 4.7 by "$\frac{4}{5}$" as follows:

Lemma 2.3.1 *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 2 (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 $H_g(\mathbb{P}_{\Gamma_g} v, \mathbb{P}_{\Gamma_g} v) > \frac{19}{20}$ in the proof of Lemma 4.7 in [We]. Similarly, Theorem 2 together with footnote 8 imply 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 by "1950ϵ".

Lemma 2.3.2 *For any almost horizontal vector v at any point of N*

$$||D\mathcal{V}(v)|| \geq 1950\epsilon||v||.$$

Proof: By Corollary 2.2.1 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.[10] This delivers the improved estimate $H_g(v, \mathbb{P}_{\bar{\Gamma}} w) \leq 850\epsilon||v|| \cdot ||w||$ and simple arithmetic concludes the proof. □

From these two lemmas it follows that the operator from $(aVert^e)^{\perp}$ 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 and obtain Theorem 1.

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

Chapter 3

Averaging of isotropic submanifolds

In this chapter we adapt Weinstein's averaging procedure (see Chaper 2) to isotropic submanifolds of symplectic manifolds. In Section 3.1 we will state our results (Theorems 3 and 4), and in Section 3.2 we will outline the proof, which will be carried out in detail in the next chapter.

3.1 The isotropic averaging theorem

Recall that a symplectic manifold is a manifold M endowed with a closed, non-degenerate 2 form ω. A submanifold of (M, ω) is called isotropic if the pullback of ω to the submanifold vanishes. 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 $I^2 = -Id_{TM}$. The tuple (M, g, ω, I) is called an *almost-Kähler manifold*. To prove our Theorem 3 we need 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 manifold is from being Kähler[1]. We state the theorem with the bound choosen to be 1 (but see Remark 1 below).

[1] Recall that an almost-Kähler manifold is Kähler if the almost complex structure I is integrable, or equivalently if $\nabla I = 0$ or $\nabla\omega = 0$.

Theorem 3 *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 $d_1(N_g, N_h) < \epsilon < \frac{1}{70000}$ 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 1: The theorem still holds if we assume higher bounds on $|\nabla\omega|$, but in that case the bound $\frac{1}{70000}$ for ϵ would have to be chosen smaller. See the remark in Section 4.5.4.

Remark 2: Notice that we are no 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 4.6. 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 Theorem 3 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]); since g is constructed canonically out of objects that are G-invariant, it will be G-invariant too. Therefore G acts on the almost Kähler manifold (M, g, ω). But it does not seem possible to give any a priori bound on $|\nabla\omega|$, where ∇ is the Levi-Civita connection corresponding to the averaged metric.

Theorem 4 *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 $d_1(N_0, gN_0) < \epsilon < \frac{1}{70000}$ 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 3 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.

3.2 Outline of the proof of Theorem 3

We will try to convince the reader that the construction we use to prove Theorem 3 works if only one chooses ϵ small enough. Let us begin by requiring $\epsilon < \frac{1}{20000}$.

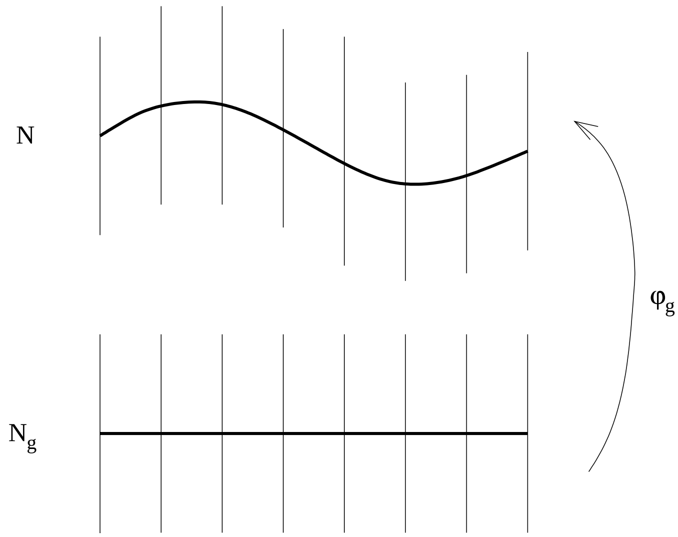
• **PART I** We start by considering the Riemannian average N of the submanifolds $\{N_g\}$ as in Theorem 1. We will use the notation $\exp_N$ to indicate the restriction of the exponential map to $TM|_N$, and similarly for any of the N_g's. For any g in G, the average N lies in a tubular neighborhood 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, using the notation $(\nu N_g)_1$ for the open unit disk bundle in νN_g, we denote by the term "normal slice" the submanifold $\exp_{N_g}(\nu_p N_g)_1$. We define the following map (see also the figure below):

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

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 $\exp_q(Vert^g_q)_1$, where $Vert^g_q \subset T_q M$ is the parallel translation along γ_q of $\nu_p N_g \subset T_p M$.

We have $d(Vert^g_q, \nu_q N) < d_1(N_g, N) < 2500\epsilon < \frac{\pi}{2}$, so $Vert^g_q$ 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 if N_g, N and M in Proposition 4.4.1 we will show that φ_g is an open embedding 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 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 $\exp_N(Vert^g)_{0.05}$. Notice that N is isotropic with respect to ω_g by construction, therefore it is also with respect to the two-form $\int_g \omega_g := \int_G \omega_g dg$ which is defined on $\cap_{g \in G} \exp_N(Vert^g)_{0.05}$ and which is the average of the ω_g's. We would like to apply Moser's trick[2] (see [Ca, Chapter III]) to ω and $\int_g \omega_g$. To do so we first restrict our forms to a smaller tubular neighborhood tub^ϵ of N, which we will define in Section 4.5.1. To apply Moser's trick we have to check:

1. *On* tub^ϵ *the convex linear combination* $\omega_t = \omega + t(\int_g \omega_g - \omega)$ *consists of symplectic forms.*

   Indeed we will show that on tub^ϵ the differential of φ_g^{-1} is "close" to the parallel translation $\mathbin{\!/\mkern-5mu/\!}$ along certain "canonical" geodesics that will be specified at the beginning of Section 4.1. This and the bound on $|\nabla \omega|$ imply that for any $q \in tub^\epsilon$

[2]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 on a compact manifold then there is a family of diffeomorphisms ρ_t with $\rho_0 = Id$ satisfying $\rho_t^* \Omega_t = \Omega_0$.

and nonzero $X, Y \in T_qM$:

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

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 non-degenerate, and it is clearly also closed.

2. *On tub^ϵ the forms ω and $\int_g \omega_g$ belong to the same cohomology class (the zero class).* Fix $g \in G$. The inclusion $i : tub^\epsilon \hookrightarrow \exp_{N_g}(\nu N_g)_1$ is homotopic to $\varphi_g^{-1} : tub^\epsilon \to \exp_{N_g}(\nu N_g)_1$. 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

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

Integrating over G finishes the argument. (Since N is isotropic w.r.t. $\int_g \omega_g$, the cohomology class of ω and $\int_g \omega_g$ is actually zero.)

Now we can apply Moser's trick: if α is a one 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 4.5.2. 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.

Chapter 4

The proof of the isotropic averaging theorem

The purpose of this chapter is to give the proof of Theorem 3, as outlined in Section 3.2. Using the notation introduced there, we can summarize the chapter as follows: in Section 4.1 we will study the map φ_g. In Section 4.2 we will state a proposition about geodesic triangles, and in Section 4.3 we will apply it in our set-up. This will allow us to show in Section 4.4 that each φ_g is injective on $\exp_{N_g}(\nu N_g)_{0.05}$. The proofs of some estimates of Sections 4.2 and 4.3 are rather involved, and we present them in three appendices at the end of this thesis. This will conclude the proof of the first part of the theorem. In Section 4.5 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 4.4 will end the proof of Theorem 3. Finally we will make a few remarks about the theorem.

4.1 Estimates on the map φ_g

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 $\mathbin{/\mkern-6mu/}X$ (this will be achieved in Proposition 4.1.4) where "$\mathbin{/\mkern-6mu/}X$" denotes the following parallel translation of

X. First we parallel translate X along the shortest geodesic from p to $\pi_{N_g}(p)$ (where π_{N_g} is the projection onto N_g along the normal slices), 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 "$\\$" as a canonical way to associate a vector in $T_{\varphi_g(p)}M$ to $X \in T_pM$.

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 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 $d(Vert^g_p, aVert^g_p) < \frac{1}{4}d(p, N_g)^2$ for any p in $\exp_{N_g}(\nu N_g)_1$, and similarly for Hor^g and $aHor^g$.
Since $\frac{1}{4}d(p, N_g)^2 < \frac{\pi}{2}$, $Vert^g$ and $aHor^g$ are always transversal (and clearly the same holds for Hor^g and $aVert^g$). As we have seen in Section 3.2, $Vert^g$ and TN are transversal along N, and $aVert^g$ and TN are also transversal since N corresponds to a section of νN_g and $aVert^g = \mathrm{Ker}(\pi_{N_g})_*$.

Now we are ready to give our estimates on the map φ_g. Recall from Section 3.2 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 4.1 to 4.4 all estimates will hold for $\epsilon < \frac{1}{20000}$.

4.1.1 Case 1: p is a point of N_g

Proposition 4.1.1 *If $p \in N_g$ and $X \in T_pN_g$ is a unit vector, then*

$$\left|\varphi_{g_*}(X) - \\X\right| \leq 3200\epsilon.$$

Remark: Notice that if X is a vector normal to N_g by definition of φ_g and $\|$ we have $\varphi_{g*}(X) = \|X$. Therefore in this subsection we will assume that X is tangent to N_g.

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

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

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

Lemma 4.1.1 *Let J be the Jacobi-field along the geodesic γ_q such that $J(0) = X$ and $J'(0) = B_{\dot\gamma_q(0)}X$, where B is the second fundamental form of N_g. Then*

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

Proof: This is an immediate consequence of [BK,6.3.8.iii][1] 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

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

where ${}_t^0\|$ denotes parallel translation to the starting point of the geodesic. Since (M, N_g) is a gentle pair, the second fundamental form B of N_g satisfies $|B_{\dot\gamma_q(0)}X| \leq \frac{3}{2}$ by [We, Cor 3.2], so the above estimate gives $|J(E) - \|X| \leq (\cosh(E) - 1) + \frac{3}{2}\sinh(E)$. Alternatively, this Lemma can be proven using the methods of [We, Prop 3.7]. □

Before proceeding we need a lemma about projections:

Lemma 4.1.2 *If $Y \in T_qM$ is a vertical unit vector, write $Y = Y_{av} + Y_h$ for the splitting into its almost vertical and horizontal components. Then*

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

[1] [BK,6.3.8] assumes that $J(0)$ and $J'(0)$ be 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 $|J|'(0) \leq |J'(0)|$.

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

Lemma 4.1.3 *If H is the unique horizontal vector at q such that $\pi_{N_g *}(H) = X$, then*

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

Proof: Let J be the Jacobi-field of Lemma 4.1.1. Write $J(E) = W + Y$ for the splitting into horizontal and vertical components. Then, using the notation of Lemma 4.1.2, we have $J(E)_h = W + Y_h$ and $J(E)_{av} = Y_{av}$. Notice that the Jacobi field J arises from a variation of geodesics orthogonal to N_g (see the Remark in Section 4.1.2), so $(\pi_{N_g})_* J(E) = X = (\pi_{N_g})_* H$. Using $aVert^g = \ker(\pi_{N_g})_*$ it follows that $H = J(E)_h$. So

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

where we used Lemma 4.1.2 and $|Y| \leq |J(E) - \backslash\backslash X|$ together with Lemma 4.1.1. □

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

Proof of Proposition 4.1.1: We have

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

The first and second terms are bounded by the estimates of Lemmas 4.1.1 and 4.1.3. For the third term we proceed analogously to Lemma 4.1.3: since $\varphi_{g*}(X)$ and H are both mapped to X via π_{N_g}, one has $(\varphi_{g*}(X))_{av} = \varphi_{g*}(X) - H$. As earlier, if $\varphi_{g*}(X) = \tilde{W} + \tilde{Y}$ is the splitting into horizontal and vertical components, we have $(\varphi_{g*}(X))_{av} = \tilde{Y}_{av}$. Therefore

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

Here we also used Lemma 4.1.2 and the fact that the angle enclosed by $\varphi_{g_*}(X)$ and its orthogonal projection onto Hor_g^q is at most $d(Hor_q^g, T_qN) \leq 2500\epsilon$ by Theorem 1. Altogether we have

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

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. □

4.1.2 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 $-B_{\dot{\gamma}_p(0)}\bar{J}(0) + \bar{J}'(0)$ 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

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

is an isomorphism. The N_g Jacobi fields that map to $aVert_p^g$ are exactly those with the property $J(0) = 0$, $J'(0) \in \nu_{\pi_{N_g}(p)}N_g$. 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 $B_{\dot{\gamma}_p(0)}J(0) = J'(0)$ - 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 4.1.2 *If $p \in \partial \exp_{N_g}(\nu N_g)_L$ and $X \in T_pM$ is an almost vertical unit vector, then*

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

We begin proving

Lemma 4.1.4 *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 such that $|J(L)| = 1$. Then*

$$\left|J(L) - L \cdot_{\gamma_p} \backslash\backslash J'(0)\right| \leq \frac{\sinh(L) - L}{\sin(L)}.$$

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 $|J'(0)| \leq \frac{1}{\sin(L)}$ 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 4.1.4, 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

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

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 4.1.5

$$\left|\tilde{J}(L) - L \cdot_{\varphi_g \circ \gamma_p} \backslash\backslash \tilde{J}'(0)\right| \leq \frac{\sinh(L) - L}{\sin(L)}.$$

Proof: Exactly as for Lemma 4.1.4 since $\tilde{J}(0) = 0$ and $|\tilde{J}'(0)| = |J'(0)|$. □

Proof of Proposition 4.1.2: 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 Lemma 4.1.4 and Lemma 4.1.5 respectively (recall $X = J(L)$ and $\varphi_{g*}(X) = \tilde{J}(L)$), and the equality holds because $\tilde{J}'(0) = \backslash\!\!\backslash J'(0)$.□

4.1.3 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 4.1.3 *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*

$$\left|\varphi_{g*}(X) - \backslash\!\!\backslash X\right| \leq \frac{18}{5}L + 3700\epsilon.$$

We proceed analogously to Case 2.

Lemma 4.1.6 *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 4.1.1 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

$$\tilde{f}_s(t) = \varphi_g\left(f_s(t)\right) = \exp_{\tilde{\sigma}(s)}\left(t\backslash\!\!\backslash v(s)\right),$$

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 4.1.7

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

Proof: Using [BK 6.3.8iii] as in Lemma 4.1.1 we obtain

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

so that we just have to estimate the norms of $\tilde{J}(0)$ and $\tilde{J}'(0)$.
Since $\tilde{J}(0) = \varphi_{g*}J(0)$, applying Proposition 4.1.1 gives $|\tilde{J}(0) - _{\gamma_0}\!\backslash\!\backslash J(0)| \leq 3200\epsilon|J(0)|$.
Using the bound for $|J(0)|$ given in Lemma 4.1.6 we obtain $|\tilde{J}(0)| \leq \frac{1+3200\epsilon}{1-\frac{3}{2}(e^L-1)}$.
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

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

and

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

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 $d(\sigma(0), \sigma(s)) \leq s|J(0)|$ and using Proposition 4.1.1 to estimate $|\dot{\tilde{\sigma}}(s)| = |\varphi_{g*}\dot{\sigma}(s)|$ we obtain $d(\tilde{\sigma}(0), \tilde{\sigma}(s)) \leq s(1+3200\epsilon)|J(0)|$. Using $d(\tilde{\sigma}(s), \sigma(s)) \leq 100\epsilon$ and Lemma 4.1.6 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

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

To bound $_{\gamma_0}\!\backslash\!\backslash J'(0)$ notice that $|J'(0)| \leq \frac{3}{2}|J(0)|$ using the fact that J is a strong Jacobi-field and [We, Cor. 3.2], so $|J'(0)| \leq \frac{3}{2}\frac{1}{1-\frac{3}{2}(e^L-1)}$.
Substituting our estimates for $|\tilde{J}(0)|$ and $|\tilde{J}'(0)|$ in $(*)$ we obtain a function which, for

$\epsilon < \frac{1}{20000}$ and $L < 0.08$, is bounded above by $\frac{9}{5}L$. □

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

Lemma 4.1.8

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

Proof: Since $\tilde{J}(0) = \varphi_{g_*}J(0)$, Proposition 4.1.1 gives

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

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

Proof of Proposition 4.1.3: 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. □

4.1.4 The general case

This proposition summarizes the three cases considered up to now:

Proposition 4.1.4 *Assume $\epsilon < \frac{1}{20000}$ and $L < 0.08$. Let $p \in \partial \exp_{N_g}(\nu N_g)_L$ and $X \in T_pM$ a unit vector. Then*

$$\left|\varphi_{g_*}(X) - \\X\right| \leq 4L + 4100\epsilon.$$

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 4.1.9 *Let J be a 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 such 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

$$\begin{aligned}|\langle J(L), K(L)\rangle| &= \left|\langle J(L), K(L)\rangle - \langle LJ'(0), K(0)\rangle\right| \\ &\leq |\langle J(L), K(L) - K(0)\rangle| + \left|\langle J(L) - LJ'(0), K(0)\rangle\right| \\ &\leq |K(L) - K(0)| + |K(0)| \cdot \left|J(L) - LJ'(0)\right|,\end{aligned}$$

which can be estimated using Lemma 4.1.6 and Lemma 4.1.4 □

Lemma 4.1.10 *Let $X \in T_pM$ be a unit vector such that $X = J(L) + K(L)$ where J, K are Jacobi-fields as in the Lemma 4.1.9 (up to normalization). Then*

$$|J(L)|, |K(L)| \leq \frac{1}{\sqrt{1 - \frac{9}{5}L}} \leq 1.1.$$

Proof: Let $c := \langle \frac{J(L)}{|J(L)|}, \frac{K(L)}{|K(L)|}\rangle$, so $|c| \leq \frac{9}{5}L$. 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 $1 = |J(L) + K(L)|^2 \geq (1 - |c|)(|J(L)|^2 + |K(L)|^2)$, from which the lemma easily follows. □

Proof of Proposition 4.1.4: 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 4.1.10. So, using Lemma 4.1.10, Proposition 4.1.2 and Proposition 4.1.3

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

□

4.2 Proposition 4.2.1 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 Section 3.2. 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 the 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 let us choose points A, B, C in M and let us 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 4.2.1 *Let M be a Riemannian manifold and ABC a geodesic triangle as above. Let P_C, 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:*

$$\measuredangle(P_A, \dot{AC}) \geq \frac{\pi}{2} - \delta$$

and

$$\theta := d(P_A, {}_{CA}\!\mathbin{\|}\!P_C) \leq \mathcal{C} d(A, C)$$

for some constants $\delta, \mathcal{C}$. We assume $\mathcal{C} \leq 2$. Then

$$d(C, B) \geq \frac{10}{11} \frac{1}{(\mathcal{C} + 6)\sqrt{1 + \tan^2 \delta}}.$$

Remark 1: Here ${}_{CA}\!\mathbin{\|}$ 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 non-oriented angle $\measuredangle(v, \dot{AC}) \in [0, \pi]$. Then we have

$$\measuredangle(P_A, \dot{AC}) := \text{Min } \left\{ \measuredangle(v, \dot{AC}) : v \in P_A \text{ nonzero } \right\} \in \left[0, \frac{\pi}{2}\right].$$

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

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 4.2.1 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 Proposition 4.2.1).

Proof: Using the chart $\exp_A$ we can lift B and C to the points $\tilde{B}$ and $\tilde{C}$ of T_AM. 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 $\measuredangle(P_A, \dot{AC}) \geq \frac{\pi}{2} - \delta$ 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 $[\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta]$. Since $\tilde{C} - 0$ and $Q - \tilde{B}$ are parallel, the angle between any vector of P_A and $Q - \tilde{B}$ lies in $[\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta]$. Since $P - \tilde{B} \in P_A$ we have

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

The triangle $\tilde{B}PQ$ has a right angle at P, so the angle $\measuredangle(P - Q, \tilde{B} - Q)$ is less of equal than δ, and claim 1 follows.

Claim 2:

$$|Q - P| \leq \sin\left[(6 + \mathcal{C}) \cdot d(C, A)\right] \cdot |Q - 0|.$$

$(\exp_A^{-1})_* \dot{CB}$ does not coincide with $\tilde{B} - \tilde{C} \in T_{\tilde{C}}(T_AM)$. Estimating - see Corollary A.1.2

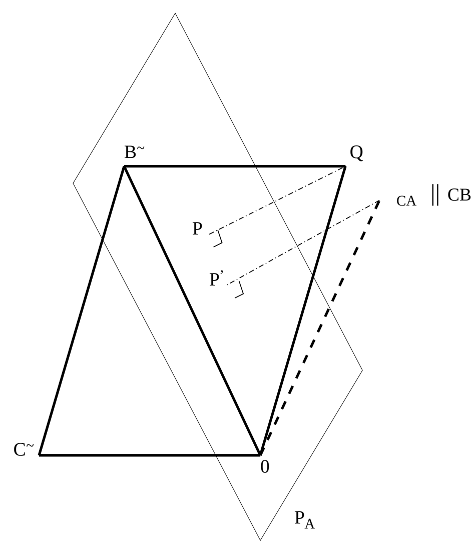
in Appendix A - the angle between these two vectors we will be able to estimate the angle between $\tilde{B} - \tilde{C} = Q - 0 \in T_A M$ and ${}_{CA}\backslash\!\!\backslash \dot{CB} \in T_A M$, i.e. the parallel translation in M of $\dot{CB}$ along the geodesic from C to A (see Corollary A.1.3). Our estimate will be $\measuredangle\left({}_{CA}\backslash\!\!\backslash \dot{CB}, Q - 0\right) < 4d(A, C)$.

Now let P' be the closest point to ${}_{CA}\backslash\!\!\backslash \dot{CB}$ in P_A. Since $P' - 0 \in P_A$ 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.1.1 - that $\measuredangle(P - 0, P' - 0) \leq 2d(C, A)$.

Combining the last three estimates we get $\measuredangle(P - 0, Q - 0) < (6 + \mathcal{C})d(C, A)$, which is less than $\frac{\pi}{2}$.

Claim 2 follows since $0PQ$ is a right triangle at P.



Claim 3:

$$d(C, B) \geq \frac{20}{21} \frac{1}{(\mathcal{C} + 6)\sqrt{1 + \tan^2 \delta}}.$$

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

$$\begin{aligned} |Q - \tilde{B}|^2 =& |\tilde{B} - P|^2 + |Q - P|^2 \\ &\leq \left(1 + \tan^2 \delta\right) \cdot |Q - P|^2 \\ &\leq \left(1 + \tan^2 \delta\right) \cdot (6 + \mathcal{C})^2 \cdot d(C, A)^2 \cdot |Q - 0|^2. \end{aligned}$$

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

$$d(A,C) \leq \sqrt{(1+\tan^2\delta)}(6+\mathcal{C}) \cdot d(C,A)|\tilde{B} - \tilde{C}|,$$

and

$$\frac{1}{(6+\mathcal{C})\sqrt{1+\tan^2\delta}} \leq |\tilde{B} - \tilde{C}|.$$

Using standard estimates - see Corollary A.1.4 - we obtain $|\tilde{B} - \tilde{C}| \leq \frac{11}{10} d(C, B)$, and the proposition follows. □

4.3 Application of Proposition 4.2.1 to $Vert^g$.

Fix g in G. Let C and A be points on Weinstein's average N with $d(C, A) < 0.15$ joint by a minimizing geodesic γ in M. Suppose that $\exp_C(v) = \exp_A(w) =: B$ for vertical vectors $v \in Vert^g_C$ and $w \in Vert^g_A$ of lengths less than 0.5. In this section we will apply Proposition 4.2.1 to the geodesic triangle given by the above three points of M and $P_A = Vert^g_A, P_C = Vert^g_C$. We will do so in Proposition 4.3.5

To this aim, first we will estimate the constants δ and $\mathcal{C}$ of Proposition 4.2.1 in this specific case. As always our estimates will hold for $\epsilon < \frac{1}{20000}$.
Roughly speaking, the constant δ - which measures how much the angle between $\dot{CA} = \dot{\gamma}(0)$ and $Vert^g_C$ deviates from $\frac{\pi}{2}$ - 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 $Vert^g_A$ and $Vert^g_C$ depends on $d(A, C)$ - will be easier, noticing that both spaces are parallel translations of normal spaces to N_g, which is a submanifold with bounded second fundamental form.

Since we have

$$\measuredangle\left(\dot{\gamma}(0), Vert^g_C\right) = \frac{\pi}{2} - \measuredangle\left(\dot{\gamma}(0), Hor^g_C\right),$$

to determine δ we just have to estimate the angle

$$\alpha := \measuredangle\left(\dot\gamma(0), Hor^g_C\right).$$

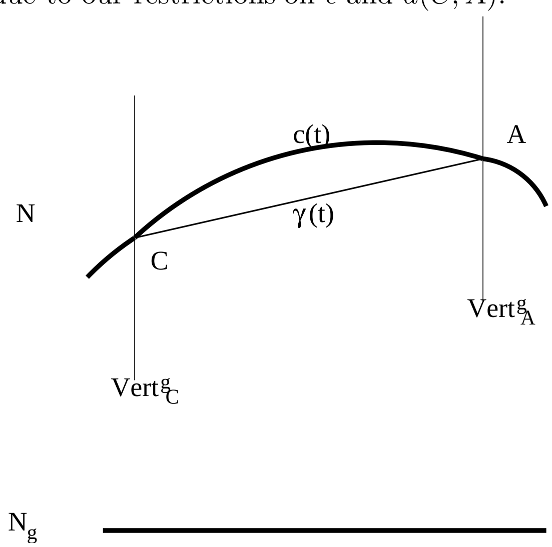
We already introduced the geodesic $\gamma(t)$ from C to A, which we assume to be parametrized by arc length. We now consider the curve $\pi(t) := \pi_{N_g} \circ \gamma(t)$ in N_g. We can lift the curve π to a curve $\varphi_g \circ \pi$ in N connecting C and A; we will call $c(t)$ the parametrization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

$\nabla^\perp$ will denote the connection induced on νN_g by the Levi-Civita connection ∇ of M, and $\frac{\perp}{\pi^b}\backslash\!\backslash$ 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

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

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

Proposition 4.3.1

$$\left|\exp^{-1}_{N_g} C - \frac{\perp}{\pi^b}\backslash\!\backslash(\exp^{-1}_{N_g} A)\right| \leq L(\gamma)\frac{3150\epsilon}{f(r)}.$$

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 in α. We will obtain:

Proposition 4.3.2

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

Comparison of Proposition 4.3.1 and Proposition 4.3.2 gives

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

Recall that $r = 100\epsilon + \frac{L(\gamma)}{2}$. If $L(\gamma)$ and ϵ are small enough one can solve the above inequality for α. With our restriction $\epsilon < \frac{1}{20000}$ this can be done whenever $L(\gamma) < 0.1$. One obtains

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

We can now state the main results of this section. First we determine the constant δ of Proposition 4.2.1 in our setting.

Proposition 4.3.3 *Let C, A be points in N and γ the shortest geodesic in M from C to A. Assume $\epsilon < \frac{1}{20000}$ and $L(\gamma) < 0.1$. Then $\delta(\epsilon, L(\gamma))$ is well defined and*

$$\measuredangle\left(Hor_C^g, \dot{\gamma}(0)\right) = \alpha < \delta\left(\epsilon, L(\gamma)\right).$$

Therefore

$$\measuredangle\left(Vert_C^g, \dot{\gamma}(0)\right) \geq \frac{\pi}{2} - \delta\left(\epsilon, L(\gamma)\right)$$

and for symmetry reasons

$$\measuredangle\left(Vert_A^g, -\dot{\gamma}(L(\gamma))\right) \geq \frac{\pi}{2} - \delta\left(\epsilon, L(\gamma)\right).$$

To determine the constant $\mathcal{C}$ we only need Lemma C.1.3:

Proposition 4.3.4 *Let C, A and γ be as above and assume $L(\gamma) < 0.1$. Then*

$$d(Vert^g_{C}, {}_{\gamma^b} \backslash\backslash Vert^g_A) \leq 2L(\gamma).$$

Proof: By Lemma C.1.3 we have $d(Vert^g_C, {}_{\gamma^b}\backslash\backslash Vert^g_A) \leq \arcsin[L(\gamma)(r + \frac{r+\frac{3}{2}}{f(r)})]$, where $r = 100\epsilon + \frac{L(\gamma)}{2}$. 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 4.2.1.

Proposition 4.3.5 *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 $v \in Vert^g_C, w \in Vert^g_A$. Then*

$$|v|, |w| \geq \frac{1}{9} \frac{1}{\sqrt{1 + \tan^2\left(\delta(\epsilon, d(A, C)\right)}}.$$

Proof: If $|v| \geq 0.5$ than the estimate for $|v|$ clearly holds, as the right hand side is $\leq \frac{1}{9}$. 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 4.2.1. Setting $P_C = Vert^g_C$ and $P_A = Vert^g_A$ in the statement of Proposition 4.2.1, Proposition 4.3.3 and Proposition 4.3.4 allow us to choose:

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

Therefore, since $\frac{10}{11}(\frac{1}{2+6}) > \frac{1}{9}$, we obtain

$$|v| \geq \frac{1}{9} \frac{1}{\sqrt{1 + \tan^2\left(\delta(\epsilon, d(A, C)\right)}}.$$

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

4.4 Estimates on tubular neighborhoods of N_g on which φ_g is injective

In this section we will finally apply the results of Section 4.2 and Section 4.3, which were summarized in Proposition 4.3.5, 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 $\cap_{g\in G}\exp_N(Vert^g)_{0.05}$ (where the 2-form $\int_g \omega_g$ is defined).

Proposition 4.4.1 *If* $\epsilon < \frac{1}{20000}$ *the map*

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

is a diffeomorphism.

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

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

Let $A, C \in N$ and $v \in Vert^g_C, w \in Vert^g_A$ be vectors of length strictly less than 0.05. We suppose that $\exp_C(v) = \exp_A(w)$ and argue by contradiction. Clearly $d(A,C) < 0.1$. We can apply Proposition 4.3.5, which implies $|v|, |w| \geq \frac{1}{9}\frac{1}{\sqrt{1+\tan^2(\delta(\epsilon,d(A,C))}}$. Since the function $\delta(\epsilon, L)$ increases with L we have

$$|v|, |w| \geq \frac{1}{9}\frac{1}{\sqrt{1+\tan^2\left(\delta(\epsilon, 0.1)\right)}}.$$

For $\epsilon < \frac{1}{20000}$ the above function is larger than 0.05, so we have a contradiction.
So $\exp_C(v) \neq \exp_A(w)$ 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 $\cap_{g\in G}\exp_N(Vert^g)_L$. This will be used in Section 4.5 to determine where $\int \omega_g$ is non-degenerate, 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 4.4.2 *For $L \leq 0.05$ and $\epsilon < \frac{1}{20000}$, using the notation*

$$R_L^\epsilon := \sin(L)\cos\left(\delta(\epsilon, 2L) + 2L^2\right)$$

we have

$$\exp_N(\nu N)_{R_L^\epsilon} \subset \cap_{g\in G} \exp_N(Vert^g)_L.$$

Remark: The function $R_{0.05}^\epsilon$ decreases with ϵ and assumes the value 0.039... at $\epsilon = 0$ and the value 0.027... when $\epsilon = \frac{1}{20000}$.

To prove the proposition we will consider again geodesic triangles:

Lemma 4.4.1 *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 that $\gamma \in [\frac{\pi}{2} - \tilde{\delta}, \frac{\pi}{2} + \tilde{\delta}]$. Then*

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

Proof: Denote by α, β the angles at A and B respectively, and denote further by α', β', γ' the angles of the Alexandrov 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]) $\gamma' \geq \gamma$. On the other hand, using the bound L^2 for the area of the Alexandrov triangles in S^2 and H^2 corresponding to ABC, we have $\gamma' - \gamma \leq 2L^2$ (see proof of Lemma A.1.1). So $\gamma' \in [\frac{\pi}{2} - \tilde{\delta}, \frac{\pi}{2} + \tilde{\delta} + 2L^2]$. Altogether this gives

$$d(A, B) \geq \sin\left(d(A, B)\right) \geq \sin(d(C, B))\sin(\gamma') \geq \sin(L)\sin\left(\frac{\pi}{2} + \tilde{\delta} + 2L^2\right).$$

□

Now we want to apply the Lemma 4.4.1 to our case of interest:

Lemma 4.4.2 *Let $C \in N$ and $B = \exp_C(w)$ for some $w \in Vert_C^g$ of length $L < 0.05$, and assume as usual $\epsilon < \frac{1}{20000}$. Then*

$$d(B, N) \geq \sin(L)\cos\left(\delta(\epsilon, 2L) + 2L^2\right) = R_L^\epsilon.$$

Here the function δ is as in Section 4.3.

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 4.3.3 we have

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

So, since $w \in Vert^g_C$,

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

Using the fact that, for any $g \in G$, the triangle ABC lies in $\exp_{N_g}(\nu N_g)_1$, the lemma follows using Lemma 4.4.1 with $\tilde\delta = \delta(\epsilon, 2L)$. □

Proof of Proposition 4.4.2: For any $g \in G$ and positive number $L < 0.05$, by Lemma 4.4.2 each point $B \in \partial(\exp_N(Vert^g)_L)$ has distance at least $\sin(L)\cos(\delta(\epsilon, 2L) + 2L^2) = R^\epsilon_L$ from N. Therefore $tub(R^\epsilon_L)$ lies in $\exp_N(Vert^g)_L$, and since this holds for all g we are done. □

4.5 Conclusion of the proof of Theorem 3

In sections 4.1-4.4, making use of the Riemannian structure of M, we showed that the two-form $\int_g \omega_g$ is well-defined in the neighborhood $\cap_{g\in G}\exp_N(Vert^g)_{0.05}$ of N (Recall that $\omega_g := ({\varphi_g}^{-1})^*\omega$ was defined in Section 3.2). In this section we will focus on the symplectic structure of M and conclude the proof of Theorem 3, as outlined in Part II of Section 3.2.

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 $\omega_t := \omega + t(\int_g \omega_g - \omega)$ consists of symplectic forms.
As we saw in Section 3.2, $[\omega] = [\int_g \omega_g] \in H^2(tub^\epsilon, \mathbb{R})$, so we can apply Moser's trick. The main step consists of constructing canonically a primitive α of small maximum norm for the two-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 4.1 hold for such L. We start by requiring $\epsilon < \frac{1}{20000}$ and introduce the abbreviation

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

for the upper bound obtained in Proposition 4.1.4 on $\exp_N(Vert^g)_L$.

4.5.1 Symplectic forms in tub^ϵ

In Section 4.1 we estimated the difference between $(\varphi_g)_*X$ and $\backslash\!\!\backslash X$. This lemma does the same for φ_g^{-1}.

Lemma 4.5.1 *Let $q \in \partial \exp_N(Vert^g)_L$ and $X \in T_qM$ a unit vector. Then*

$$\left|(\varphi_g^{-1})_*X - \backslash\!\!\backslash X\right| \leq \frac{D_L^\epsilon}{1 - D_L^\epsilon}.$$

Furthermore,

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

Proof: Let $p := \varphi_g^{-1}(q)$. By Proposition 4.1.4, for any vector $Z \in T_pM$ we have

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

The second statement of this lemma follows setting $Z = (\varphi_g^{-1})_*X$.
Choosing instead $Z = (\varphi_g^{-1})_*X - \backslash\!\!\backslash X \in T_pM$ and applying once more Proposition 4.1.4 gives

$$\left|(\varphi_g^{-1})_*X - \backslash\!\!\backslash X\right| \leq \frac{|X - (\varphi_g)_*\backslash\!\!\backslash X|}{1 - D_L^\epsilon} \leq \frac{D_L^\epsilon}{1 - D_L^\epsilon}.$$

□

Since $(\varphi_g^{-1})_*X$ is close to $\backslash\!\!\backslash X$ and since our assumption on $\nabla\omega$ allows us to control to which 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 4.5.2 *Let X, Y be unit tangent vectors at $q \in \exp_N(Vert^g)_L$. 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: Denoting $p := \varphi_g^{-1}(q)$ we have

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

Now since "$\backslash\!\backslash$" is the parallel translation along a curve of length $< 2L + 100\epsilon$ (see Section 4.1) and $|\nabla\omega| < 1$ we have $\omega_p(\backslash\!\backslash X, \backslash\!\backslash Y) - \omega_q(X, Y) < 2L + 100\epsilon$ and using Lemma 4.5.1 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 non-degeneracy of ω_t for L and ϵ small enough.

Corollary 4.5.1 *Let X be a unit tangent vector at $q \in \cap_{g\in G} \exp_N(Vert^g)_L$. 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 4.5.2. □

Remark: The right hand side of Corollary 4.5.1 is surely positive if $D_L^\epsilon \leq 0.1$. We set[2]

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

and require $\epsilon < \frac{1}{70000}$. We obtain immediately:

[2]This choice of L^ϵ will allow us to obtain good numerical estimates in Section 4.5.4.

Proposition 4.5.1 *On*

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

the convex linear combination $\omega_t := \omega + t(\int_g \omega_g - \omega)$ *consists of symplectic forms.*

Remark: Recall that the function R^ϵ_L was defined in Proposition 4.4.2. See Section 4.5.4 for a graph of $R^\epsilon_{L^\epsilon}$, a function of ϵ.

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

We want to construct canonically a primitive α of $\frac{d}{dt}\omega_t = \int \omega_g - \omega$ on $\cap_{g\in G} \exp_N(Vert^g)_{0.05}$. 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 $\exp : \tilde{U} \to U \subset M$ be 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

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

gives an operator with the above property.

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

$$\begin{aligned} |(Q_t\omega)_p X| &= \left|\omega_p \left(w_t|_{\rho_t(p)}, \rho_{t*}(X)\right)\right| \\ &\leq |\tilde{\omega}_p|_{\mathrm{op}} \cdot d(p, \pi(p)) \cdot |\rho_{t*}(X)| \qquad (\bigstar) \end{aligned}$$

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 $\exp_N(Vert^g)_{0.05}$ We do so in two steps:

Step I We apply the above procedure to the vector bundle $Vert^g \to N$ to obtain an operator Q_N^g such that

$$\mathrm{Id} - (\pi_N^g)^* \circ (i_N)^* = dQ_N^g + Q_N^g d$$

for all differential forms on $\exp_N(Vert^g)_{0.05}$. 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 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 $\exp_N(Vert^g)_{0.05}$.

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

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

In this subsection we will estimate the C^0-norm of the one-form α constructed in Section 4.5.2.

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

$$\left|(\rho_{N_g})_{t_*}X\right| \leq \frac{5}{4}.$$

Proof: Let $L := d(p, N_g) < 100\epsilon < \frac{1}{700}$ 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 Subsection 4.1.2).

$J(t)$ 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

$$(\rho_{N_g})_{t_*}J(L) = \frac{d}{ds}\Big|_0 \left[(\rho_{N_g})_t \circ \exp_{p'}(Lv(s))\right] = \frac{d}{ds}\Big|_0 \left[\exp_{p'}(tLv(s))\right] = J(tL).$$

Using Lemma 4.1.4 we have on 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 $\sigma(x) := \frac{\sinh(x) - x}{\sin(x)}$. 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 4.1.6 we have $|K(0)| \leq (1 + \frac{9}{5}L)|K(L)|$ and $|K(tL)| \leq \frac{|K(0)|}{1 - \frac{9}{5}tL}$, 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

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

where in the first and third inequality we used Lemma 4.1.9 and in the fourth in addition we used Lemma 4.1.10. □

Corollary 4.5.2 *The one-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 ($\bigstar$) , $|\tilde{\omega}|_{\mathrm{op}} = 1$ and Lemma 4.5.3, we have $|(Q_{N_g}\omega)_p| < 125\epsilon$. Clearly $|(Q_{N_g}\omega)_p| \geq |i_N^*(Q_{N_g}\omega)_p| = |\beta_p^g|$. □

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 4.5.4 *If $q \in \exp_N(Vert^g)_L$ 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 4.5.1 we have

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

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 4.5.3 requiring p to lie in $\exp_{N_g}(\nu N_g)_L$ we obtain[3] $|(\rho_{N_g})_{t_*} Y| < 1.5$ for unit vectors Y at p. Using this and Lemma 4.5.1 respectively we have

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

Altogether this proves the lemma. □

Corollary 4.5.3 *On $\exp_N(Vert^g)_L$ we have*

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

[3]Since $L < 0.05$ now we have to replace the constant $\frac{21}{20}$ in that proof by the constant $\frac{6}{5}$.

Proof: This is clear from $|((\pi_N^g)^*\beta^g)X| = |\beta^g((\pi_N^g)_*X)|$, Corollary 4.5.2 and Lemma 4.5.4. $\square$

Step I Now we estimate $|Q_N^g(\omega_g - \omega)|$. This is easily achieved using Lemma 4.5.2 and Lemma 4.5.4 to estimate the quantities involved in ($\bigstar$):

Corollary 4.5.4 *For $q \in \partial \exp_N(Vert^g)_L$ we have*

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

Remark: By Proposition 5.2 $d(q,N) \geq R_L^\epsilon$. Furthermore, when $\epsilon < \frac{1}{70000}$ 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 Corollary 4.5.3 and Corollary 4.5.4 we can estimate the norm of $\alpha := \int_g \alpha^g$:

Proposition 4.5.2 *Assuming $L < 0.05$ at $q \in \cap_g \exp_N(Vert^g)_L$ we have*

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

4.5.4 The end of the proof of Theorem 3

Proposition 4.5.1 showed that the Moser vector field $v_t := -\tilde{\omega}_t^{-1}\alpha$ is well-defined on $tub^\epsilon \subset \cap_{g\in G} \exp_N(Vert^g)_{L^\epsilon}$. Recalling that $D_{L^\epsilon}^\epsilon = 0.1$, Corollary 4.5.1 immediately implies

Corollary 4.5.5 *At $q \in tub^\epsilon \subset \cap_g \exp_N(Vert^g)_{L^\epsilon}$ we have*

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

From Corollary 4.5.5 and Proposition 4.5.2 we obtain:

Proposition 4.5.3 *For all $t \in [0,1]$ and $q \in tub^\epsilon$*

$$|(v_t)_q| \leq \left|(\widetilde{\omega_t})_q^{-1}\right|_{op} \cdot |\alpha_q| \leq 1.45\, d(q,N) + 374\epsilon.$$

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

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

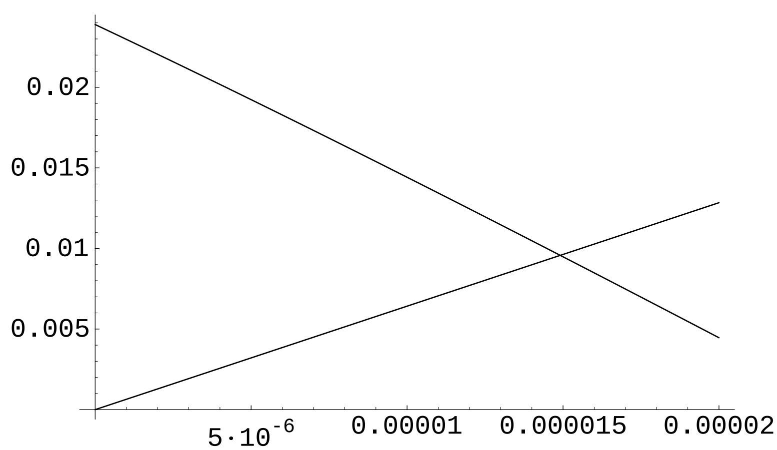
The solution of the ODE $\dot{s}(t) = As(t) + B$ satisfying $s(0) = 0$ is $\frac{B}{A}(e^{At} - 1)$. Hence, if the integral curve γ is well defined at time 1, we have

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

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 $tub^\epsilon := \exp_N(\nu N)_{R^\epsilon_{L^\epsilon}}$ 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 $\epsilon < \frac{1}{70000}$.



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

The estimate for $d_0(N_g, L)$ is obtained using $d_0(N, L) < 842\epsilon$ and $d_0(N_g, N) < 100\epsilon$. The proof of Theorem 3 is now complete.

Remark: In Theorem 3 we assumed that $|\nabla\omega| < 1$. When the bound on $|\nabla\omega|$ is bigger than one, the statement of Theorem 3 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 4.5.2; if $|\nabla\omega|$ is bigger than one, the right hand side of that lemma should read $|\,(\omega_g-\omega)(X,Y)| \leq \frac{D^\epsilon_L}{1-D^\epsilon_L}\left(\frac{D^\epsilon_L}{1-D^\epsilon_L}+2\right)+|\nabla\omega|(2L+100\epsilon)$ instead. Similarly, the quantity "$2L+100\epsilon$" appearing in Corollary 4.5.1, Corollary 4.5.4 and Proposition 4.5.2 should be multiplied by $|\nabla\omega|$. Now let us assume that $\epsilon < \frac{1}{|\nabla\omega|}\frac{1}{70000}$ and let us replace L^ϵ everywhere by $\tilde{L}^\epsilon := \frac{\frac{0.1}{|\nabla\omega|}-4100\epsilon}{4}$. Then the bounds on $\left|(\widetilde{\omega_t})^{-1}_q\right|_{\text{op}}$ and $|(v_t)_q|$ given in Corollary 4.5.5 and Proposition 4.5.3 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.

4.6 Remarks on Theorem 3

Remark 1: *Is the isotropic average L C^1-close to the N_g's?*

The main shortcoming in our Theorem 3 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 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 tangent spaces 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 so 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 4.5.2 always gives a one-form vanishing on points of N, Step II does not, even if N is isotropic for ω.

The procedure outlined in Remark 3 below, 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 the Riemannian averaging procedure (Theorem 1) .

Unfortunately our construction does not allow us to average 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. In the case that the N_g are not isotropic 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 $\int_g \omega_g$.

Chapter 5

Applications of the isotropic averaging theorem

In this chapter we will present two simple applications of Theorem 4 (see Chapter 3). The first one concerns isotropic submanifolds which are almost invariant under Hamiltonian actions, the second one concerns symplectomorphisms which are almost equivariant with respect to a compact group action.

5.1 Application to Hamiltonian actions

We apply Theorem 4 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 a Hamiltonian action.

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

Lemma 5.1.1 *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 $X \in T_xL$. 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 $G \cdot x_0 \subset L$ 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 5.1.1 *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) $d_1(L, g \cdot L) < \epsilon < \frac{1}{70000}$ 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 4 there exists an isotropic submanifold L' invariant under the G action with $d_0(L, L') < 1000\epsilon$. By Lemma 5.1.1 L lies in the 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

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

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. □

Remark: Notice that the bound $1000\epsilon \cdot C$ does not depend on L, unlike the bounds that one obtains if one replaces the assumption that L be "C^1 almost invariant" with the assumption that the vector fields v_M's be "almost tangent" to L.

5.2 Application to equivariant symplectomorphisms

In this section we present a preliminary result about almost equivariant symplectomorphisms.

Fact 5.2.1 *Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds, G a compact group acting on M_1 and M_2 by symplectomorphisms and $\phi : M_1 \to M_2$ a symplectomorphism. Suppose we can endow M_i with a G-invariant compatible metric g_i satisfying $|\nabla \omega_i| < 1$ and such that $\mathrm{graph}(\phi) \subset (M_1 \times M_2, g_1 \times g_2, \omega_1 - \omega_2)$ satisfies the following properties: $(M_1 \times M_2, \mathrm{graph}(\phi))$ is a gentle pair and $d_1(\mathrm{graph}(\phi), g \cdot \mathrm{graph}(\phi)) < \epsilon < \frac{1}{70000}$ for all g (where G acts on $M_1 \times M_2$ diagonally).*
If the isotropic average of the family $\{g \cdot \mathrm{graph}(\phi)\}$ is of the form $\mathrm{graph}(\psi)$, then $\psi : M_1 \to M_2$ is a G-equivariant symplectomorphism and $d_0(\mathrm{graph}(\phi), \mathrm{graph}(\psi)) < 1000\epsilon$.

Proof: First notice that $(M_1 \times M_2, g_1 \times g_2, \omega_1 - \omega_2)$ is an almost Kähler manifold satisfying $|\nabla(\omega_1 - \omega_2)| < 1$, that $\mathrm{graph}(\phi)$ is contained in it as a Lagrangian submanifold, and that the diagonal action of G on $M_1 \times M_2$ is by isometric symplectomorphism. Therefore we can apply Theorem 4 to obtain a G-invariant Lagrangian submanifold (which is exactly the isotropic average of the family $\{g \cdot \mathrm{graph}(\phi)\}$), and by assumption this submanifold is of the form $\mathrm{graph}(\psi)$. It is easy to check that a map from M_1 to M_2 is G-equivariant if and only if its graph is invariant under the diagonal G action on $M_1 \times M_2$, therefore ψ is indeed a G-equivariant symplectomorphism. $\square$

Remark 1: Fact 5.2.1 is a symplectic version of a theorem by Grove and Karcher [GK] stating that if two actions of a compact group G on a compact manifold M are C^1-close to each other, then there exists a diffeomorphism ψ intertwining them. Indeed, given an action ν_1 of G on a manifold M_1 and an action ν_2 of G on M_2 together with a "C^1-almost equivariant" diffeomorphism $\phi : M_1 \to M_2$, the action of G on M_1 given by $g \cdot m := \phi^{-1}(\nu_2(g, \phi(m))$ is C^1 close to ν_1, and if ψ is an intertwining diffeomorphism as in Karcher and Grove's theorem, then $\phi \circ \psi : M_1 \to M_2$ intertwines between ν_1 and ν_2. Grove and Karcher's theorem exhibits explicitly the intertwining diffeomorphism by means of a center of mass construction for maps.

Remark 2: In the above Fact we have to *assume* that the average of the family $\{g \cdot \mathrm{graph}(\phi)\}$ is a graph. An improvement of Theorem 3 to include a bound on the C^1-distance of the average submanifold from the original isotropic family (see also Remark 1 in Section 4.6) together with a mild assumption on the C^1-norm of ϕ (see the remark below) would allow us to remove this assumption.

Remark 3: It would be nice to express the assumptions of Fact 5.2.1 in terms of maps instead of graphs, however the formulation in terms of maps is more involved. Indeed, let us adopt the following definitions for the C^0 and C^1 distances (norms) of maps, where f and g are differentiable maps between Riemannian manifolds M and N:

a) $d_0(f, g) := \sup_{x \in M} d(f(x), g(x))$

b) $|f|_1 := \sup_{\{v \in TM \text{ unit vector}\}} |f_* v|$

c) $d_1(f, g) := \sup_{\{v \in TM \text{ unit vector}\}} |f_* v - \backslash\!\backslash g_* v|$,
defined whenever $d_0(f, g) <$ injectivity radius of N, where "$\backslash\!\backslash$" denotes parallel translation of $v \in T_x M$ from $g(x)$ to $f(x)$ along the shortest geodesic.

(When $M = N$ is compact and f is a diffeomorphism we have $|f|_1 \geq 1$, with equality iff f is an isometry). For the C^0 distances we have the simple estimates

$$d_0(\mathrm{graph}(f), \mathrm{graph}(g)) \leq d_0(f, g),$$

$$d_0(f, g) \leq (2 + |f|_1) d_0(\mathrm{graph}(f), \mathrm{graph}(g)),$$

however the corresponding C^1 estimates are more complicated and involve suitably defined C^2 norms of f and g.

Chapter 6

Averaging of Legendrian submanifolds

In this chapter we will consider the setting of contact geometry and give a construction to average Legendrian submanifolds. After presenting the results in Section 6.1 (Theorems 5 and 6), in Section 6.2 we will outline the proof, which consists of considering the symplectization of the given contact manifold and applying the isotropic averaging theorem there. The details of the proof are given in Section 6.3.

6.1 Results

Recall that a *contact manifold* is a manifold M^{2n+1} together with a hyperplane distribution $\mathcal{H}$ on M such that locally $\mathcal{H} = \ker\theta$ for some locally defined 1-form θ satisfying $(d\theta)^n \wedge \theta \neq 0$. A submanifold N of $(M^{2n+1}, \mathcal{H})$ is called Legendrian if it is tangent to $\mathcal{H}$ and it has maximal dimension among submanifolds with this property, i.e. $\dim(N) = n$. Also recall that a *contact one-form* on a manifold M^{2n+1} is a (global) 1-form θ such that $(d\theta)^n \wedge \theta$ is a volume form. The unique vector field E satisfying $\theta(E) = 1$, $d\theta(E, \cdot) = 0$ is called Reeb vector field. Any contact manifold $(M, \mathcal{H})$ for which the distribution $\mathcal{H}$ is co-orientable can be endowed with a contact form θ representing $\mathcal{H}$ (i.e. $\ker\theta = \mathcal{H}$). Now consider a manifold with a contact form (M, θ), and endow it with a "compatible"

Riemannian metric g as follows: for each fiber $\mathcal{H}_p$ of the vector bundle $\mathcal{H} = \ker\theta \to M$, $(\mathcal{H}_p, d\theta|_{\mathcal{H}_p})$ is a symplectic vector space, and we can choose a compatible positive inner product g (i.e. $d\theta(X, IY) = g(X, Y)$ determines an endomorphism I of $\mathcal{H}_p$ satisfying $I^2 = -Id$). We can do so in smooth way (see [Ca], Ch. 12). We extend g to a Riemannian metric on M by imposing that the Reeb vector field E have unit length and be orthogonal to $\mathcal{H}$.

We state our theorem for Legendrian submanifolds, even though it equally applies to submanifolds tangent to $\mathcal{H}$ of lower dimension. As a technical assumption we will require that the C^0-norms of the covariant derivatives of θ and $d\theta$ with respect to the Levi-Civita connection be bounded by 1 (but see the remark below).

Theorem 5 *Let (M, θ) be a manifold with a contact form, endowed with a Riemannian metric g as above so that $|\nabla\theta|, |\nabla d\theta| < 1$. Let $\{N_g\}$ be a family of Legendrian 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 $d_1(N_g, N_h) < \epsilon < \frac{1}{70000}$ for all g and h in G, there is a well defined* **Legendrian average** *submanifold L with $d_0(N_g, L) < 1000\epsilon$ for all g in G. This construction is equivariant with respect to isometric contactomorphisms of (M, θ) and invariant with respect to measure preserving automorphisms of G.*

Remark: The theorem holds even if the bound on $|\nabla\theta|$ and $|\nabla d\theta|$ is larger than 1, but in that case the bound on ϵ has to be chosen smaller.

A simple consequence, which we want to state in terms of contact manifolds, is the following:

Theorem 6 *Let $(M, \mathcal{H})$ be a contact manifold for which $\mathcal{H}$ is co-orientable, let G be a compact Lie group acting on M preserving $\mathcal{H}$, and let N_0 be a Legendrian submanifold. Endow $(M, \mathcal{H})$ with a contact form θ and a Riemannian metric g as above, both invariant under the G action. Suppose that $|\nabla\theta|, |\nabla d\theta| < 1$. Then if (M, N_0) is a gentle pair and $d_1(N_0, gN_0) < \epsilon < \frac{1}{70000}$ for all $g \in G$, there exists a G-invariant Legendrian submanifold L of $(M, \mathcal{H})$ with $d_0(N_0, L) < 1000\epsilon$.*

Indeed, we just need to endow G with its bi-invariant probability measure and apply Theorem 5 to the family $\{gN_0\}$: their average will be G-invariant by the equivariance properties of Theorem 5.

Remark We can always find θ and g which are G-invariant: by averaging over G we can obtain a G-invariant one-form θ representing $\mathcal{H}$, and using some G-invariant metric on the vector bundle $\mathcal{H} \to M$ as a tool we can construct a "compatible" metric g which is G-invariant (see Ch. 12 in [Ca]). However in general we can not give any a priori bounds on the covariant derivatives of θ and $d\theta$.

6.2 Idea of the proof of Theorem 5

An approach to prove Theorem 5 is to use the idea that worked in the symplectic setting (see Section 3.2). This would be carried out as follows: first construct the Weinstein average N of the Legendrian submanifolds $\{N_g\}$. For each $g \in G$, using the Riemannian metric construct a diffeomorphism φ_g from a neighborhood of N_g to a neighborhood of N and denote by θ_g the pullback of θ to the neighborhood of N. Since the submanifolds N_g are close to N, each form in the convex linear combination $\theta_t = \theta + t(\int \theta_g - \theta)$ is a contact form, say with contact distribution $\mathcal{H}_t$. Therefore we can apply the contact version of Moser's Theorem (see [Ca], Ch. 10). It states that if the vector field v_t is the inverse image of $-(\int \theta_g - \theta)|_{\mathcal{H}_t}$ by (the isomorphism induced by) $d\theta_t|_{\mathcal{H}_t}$, then the time-one flow ρ_1 of $\{v_t\}$ satisfies $(\rho_1)_* \mathcal{H} = \mathcal{H}_1$. Therefore, since N is tangent to $\mathcal{H}_1$, its pre-image L under ρ_1 is tangent to $\mathcal{H}$, i.e. it is a Legendrian submanifold of (M, θ).
This construction can indeed be carried out and satisfies the invariance properties stated in Theorem 5 since all steps are canonical. However it delivers a numerically quite unsatisfactory estimate for $d_0(N_g, L)$; therefore we choose *not* to use this approach but rather a different one, which we outline now.[1]

Recall that the *symplectization* of a manifold with contact form (M, θ) is the symplectic manifold $(M \times \mathbb{R}, d(e^s\theta))$, where s denotes the coordinate on the $\mathbb{R}$ factor. Here and in

[1] The estimates needed for our first approach are completely analogous to those needed for the second approach, i.e. those of Chapter 4. A difference though is that in the first approach we make use of a bound on the norm of the C^0-small one-form $\int_g \theta_g - \theta$, whereas in the second approach we will need the norm of a primitive of a certain C^0-small two-form. While passing to the primitive we will improve the C^0-norm by "one order of magnitude" (see equation (★) in Section 4.5.2), and this is responsible for the better estimates obtained using the second approach.

the following we abuse notation by writing θ in place of $\pi^*\theta$, where $\pi : M\times\mathbb{R} \to M$. The main observation is the following lemma, whose proof consists of a short computation and is omitted.

Lemma 6.2.1 *L is a Legendrian submanifold of M if and only if $L\times\mathbb{R}$ is a Lagrangian submanifold of $M \times \mathbb{R}$.*

Using this lemma and making use of the averaging procedure for Lagrangian submanifolds (Theorem 3) the strategy for our (second) approach is straightforward: given the family $\{N_g\}$ of Legendrian submanifolds of (M, θ) we consider the Lagrangian family $\{N_g \times \mathbb{R}\}$ in the symplectization of M, apply the Lagrangian averaging theorem, and *if* the average is a product $L \times \mathbb{R}$ then L will be our Legendrian average. The invariance properties stated in Theorem 5 are satisfied because this construction is canonical after we choose the contact form and metric on M.

6.3 The proof of Theorem 5

We endow the symplectization $M \times \mathbb{R}$ with the product metric obtained from $(\mathbb{R}, ds \otimes ds)$ and (M, g), and by abuse of notation denote this metric by g. Unfortunately $(M \times \mathbb{R}, d(e^s\theta), g)$ does *not* satisfy the assumptions of Theorem 3: indeed it is not an almost-Kähler manifold (however g is compatible with the non-degenerate 2-form $e^{-s}d(e^s\theta) = ds \wedge \theta + d\theta$). Furthermore the condition on the boundedness of $\nabla d(e^s\theta)$ is also violated. Therefore we can not just apply Theorem 3 but we have to follow the construction in the theorem and check that it applies to $(M \times \mathbb{R}, d(e^s\theta), g)$.
The remaining assumptions of Theorem 3 concerning the gentleness of the pairs $(M \times \mathbb{R}, N_g \times \mathbb{R})$ and the distances $d_1(N_g \times \mathbb{R}, N_h \times \mathbb{R})$ are satisfied, since the metric g on $M \times \mathbb{R}$ is a product metric.

In the remainder of the proof we will follow the construction of Theorem 3. We do so for two reasons: firstly, in order to make sure that the Lagrangian average of the $\{N_g \times \mathbb{R}\}$ is of the form $L \times \mathbb{R}$ for some submanifold L of M, and secondly to check that the construction applies to $(M \times \mathbb{R}, \omega := d(e^s\theta), g)$ even though the assumptions of Theorem 3 are not satisfied.

We refer the reader to Section 3.2 for the outline of the construction of Theorem 3 and adopt the notation used there. We divide the construction into 5 steps.

Step 1: *Construct the Weinstein average of the family* $\{N_g \times \mathbb{R}\}$.
Step 1 applies to our manifold $M \times \mathbb{R}$ because Step 1 involves only the Riemannian structure of $M \times \mathbb{R}$, which satisfies the assumptions of Weinstein's averaging theorem (see Chapter 3). Notice that the Weinstein average is of the form $N \times \mathbb{R}$ because the group $\mathbb{R}$ acts isometrically on $M \times \mathbb{R}$ by translation of the second factor and Weinstein's averaging procedure is equivariant w.r.t. isometries.

Step 2: *The restriction of*

$$\varphi_g : \text{ Neighborhood of } N_g \times \mathbb{R} \rightarrow \text{Neighborhood of } N \times \mathbb{R}$$

to $\exp_{N_g \times \mathbb{R}}(\nu(N_g \times \mathbb{R}))_{0.05}$ *is a diffeomorphism onto.*
Step 2 applies to $M \times \mathbb{R}$ because it involves only its Riemannian structure. Notice that φ_g preserves $M \times \{s\}$ for each $s \in \mathbb{R}$ and that the $\varphi_g|_{M \times \{s\}}$ coincide for all s (under the obvious identifications $M \times \{s\} \equiv M \times \{s'\}$), since the metric on $M \times \mathbb{R}$ is the product metric.

Step 3: *On* tub^ϵ *the family* $\omega_t := \omega + t(\int_g \omega_g - \omega)$ *consists of symplectic forms lying in the same cohomology class, where* $\omega_g := (\varphi_g^{-1})^* \omega$ *and* tub^ϵ *is a neighborhood of* N.
Define $\bar{\omega} := ds \wedge \theta + d\theta$, a non-degenerate 2 form compatible with the metric g on $M \times \mathbb{R}$. Below we will show that $|\nabla_X \bar{\omega}| \leq 2$ for "horizontal" unit vectors X. Since $\omega = e^s \bar{\omega}$ using this we see that the statements of Lemma 4.5.2 and Corollary 4.5.1 hold if one multiplies the right hand sides there by e^s and multiplies by 2 the term "$2L + 100\epsilon$"[2]. This shows that the ω_t are non-degenerate on tub^ϵ. One sees that the ω_t lie in the zero cohomology class exactly as in Section 3.2.
Now we derive the bound on $|\nabla_X \bar{\omega}|$, where $X \in T(M \times \{s\})$ is a unit vector:

$$\begin{aligned} |\nabla_X(ds \wedge \pi^*\theta + d\pi^*\theta)| &= |\nabla_X ds \wedge \pi^*\theta + ds \wedge \nabla_X \pi^*\theta + \nabla_X \pi^* d\theta| \\ &= |ds| \cdot |\nabla_X \pi^*\theta| + |\nabla_X \pi^* d\theta| \\ &\leq 1 \cdot 1 + 1. \end{aligned}$$

[2]This factor of 2 does not affect the numerical estimates that follow.

Step 4: *Construct canonically a primitive α of $\int_g \omega_g - \omega$.*
A primitive of $\omega_g - \omega$ is given in a canonical fashion by $e^s((\varphi_g^{-1})^*\theta - \theta)$, however we do *not* want to use this primitive since it would deliver bad numerical estimates, as it happens in our first approach. Instead we construct one using the (more involved) procedure of Section 4.5.2.
Because of the remark in Step 2 we can write $\omega_g = e^s\bar{\omega}_g$ where $\bar{\omega}_g := (\varphi_g^{-1})^*\bar{\omega}$, therefore the 2-form $\omega_g - \omega$ can be written as $e^s(\bar{\omega}_g - \bar{\omega})$. Notice that $\bar{\omega}$ is "constant in $\mathbb{R}$-direction" (i.e. the Lie derivative $L_{\frac{\partial}{\partial s}}\bar{\omega} = 0$), and by the remarks in Step 2 the same holds of $\bar{\omega}_g$. The construction in Section 4.5.2 commutes with multiplication of forms by functions of s and furthermore preserves the condition of being "constant in $\mathbb{R}$ direction" (this follows from the explicit formula for α^g in Section 4.5.2 and from the fact that the vector bundles used there are pullbacks of vector bundles over subsets of M via $\pi : M \times \mathbb{R} \to M$). Therefore that construction applied to $\int \omega_g - \omega$ delivers a primitive $\alpha = e^s\bar{\alpha}$ where $\bar{\alpha}$ is "constant in $\mathbb{R}$ direction". Furthermore $|\bar{\alpha}|$ is estimated as in Proposition 4.5.2, because $\bar{\omega}$ is compatible with the metric.

Step 5: *Obtain the Lagrangian average by following backwards the Moser vector field $v_t := -\omega_t^{-1}(\alpha)$ starting from $N \times \mathbb{R}$.*
The key observation is that

$$v_t = -\omega_t^{-1}(\alpha) = -\bar{\omega}_t^{-1}(\bar{\alpha}),$$

therefore the vector field v_t is independent of s (i.e. $L_{\frac{\partial}{\partial s}} v_t = 0$ for $t \in [0,1]$). This implies that the Lagrangian average is of the form $L \times \mathbb{R}$. Further, this implies that $|v_t|$ is bounded as in Proposition 4.5.3, so that the proof of the theorem can be concluded verbatim as in Chapter 4.

Chapter 7

Neighborhoods of coisotropic submanifolds

In Remark 3 of Section 4.6 we saw that the proof of the isotropic averaging theorem (Theorem 3) breaks down if we try to apply it to coisotropic submanifolds. If one could average coisotropic submanifolds, then the space of coisotropic submanifolds would necessarily be locally path connected. In this chapter we will consider the space of coisotropic submanifolds close to a given one, in an attempt to determine whether it is locally path connected. While we are not able to do so, we will display some of the properties of this space.

In the case of a compact Lagrangian submanifold L of a symplectic manifold (N, Ω), it is well known that a tubular neighborhood of L is symplectomorphic to (a neighborhood of the zero section of) the cotangent bundle T^*L endowed with its canonical symplectic structure. Furthermore, it is easy to see that Lagrangian submanifolds of (N, Ω) which are C^1-close to L correspond exactly to C^1-small closed one-forms on L [Ca, Chapter 3][1].

For the general case where M is a compact coisotropic submanifold of (N, Ω) — that

[1] We use the terms "C^1-close" and "C^1-small" in a loose way here. To make these notions precise one has to introduce a Riemannian metric on N. Then"C^1-close to L" refers to $d(L, \cdot)$ being small, where $d(\cdot, \cdot)$ is the C^1-distance introduced in Section 2.1. Similarly "C^1-small" refers to $|| \cdot ||_1$ being small, where $|| \cdot ||_1$ is the norm we introduce in Example 2 of Section 7.3.

is, $TM^{\Omega} \subset TM$ — we will proceed analogously: In Section 7.1, following Gotay [Go], we will see that a tubular neighborhood of M is symplectomorphic to (a neighborhood of the zero section of) a subbundle of T^*M, and in Section 7.2 we will determine which sections of this subbundle correspond to coisotropic submanifolds. The condition we obtain, unlike in the Lagrangian case, is not linear.
In Section 7.3 we will point out features of characteristic leaves which suggest that the coisotropic submanifolds C^1-close to M do not form a"nice" set and, in Theorem 7, we will make this statement precise in terms of Fréchet manifolds and prove it by means of a very simple counterexample.

7.1 Neighborhoods of coisotropic submanifolds

In analogy to the statement that the neighborhood of a Lagrangian submanifold is symplectomorphic to its cotangent bundle, in this section (which is based on Gotay's paper [Go]) we will give an explicit form for the neighborhood of a coisotropic submanifold.

Let M^n be a compact coisotropic submanifold of the symplectic manifold (N^{n+k}, Ω) and let us denote by ω the pullback of Ω to M. Notice that the kernel of ω has constant dimension equal to k. Modifying slightly the proof of the Existence Theorem in [Go], we will now construct another symplectic manifold (E^*, Ω_{E^*}) in which (M, ω) embeds coisotropically. According to the Local Uniqueness Theorem in [Go], any two coisotropic embeddings of the presymplectic manifold (M, ω) in symplectic manifolds N_1 and N_2 are neighborhood equivalent, i.e. there exists a symplectomorphism between suitable tubular neighborhoods of M in N_1 and N_2 fixing M. It follows that tubular neighborhoods of M in (N, Ω) and (E^*, Ω_{E^*}) are symplectomorphic.

Now we construct (E^*, Ω_{E^*}). Let E^k be the characteristic distribution of (M, ω), i.e. $E = \ker(\omega)$. Then $E^* := \cup_{x \in M} E^*_x$ is a vector bundle over M, whose projection map we denote by π. Let G be any fixed smooth distribution on M such that $TM = E \oplus G$. We have a decomposition $T^*M = E^* \oplus G^*$, which gives an embedding $i_{E^*,T^*M} : E^* \hookrightarrow T^*M$: if $(m, \xi) \in E^*$, then $i_{E^*,T^*M}(m, \xi)$ is the element of $T^*_m M$ that acts on E_m like ξ and annihilates G_m. Consider now

$$\Omega_{E^*} := \pi^*\omega + (i_{E^*,T^*M})^*(-d\alpha),$$

where α is the canonical one-form on T^*M, so that $-d\alpha$ is the canonical symplectic form on T^*M.

In the following we will use that at $m \in M$ we have a natural splitting $T_mE^* = T_mM \times E^*_m = E_m \times G_m \times E^*_m$.

Lemma 7.1.1 *Let $m \in M$, $X, Y \in T_mM$, and write $X = X_E + X_G + X_{E^*}$ according to the above splitting (similarly for Y). Then*

$$\Omega_{E^*}(X,Y) = \omega(X_G, Y_G) + \langle(X_E, Y_{E^*})\rangle - \langle(X_{E^*}, Y_E)\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the natural pairing of vectors and covectors.

From this lemma it is clear that Ω_{E^*} is non-degenerate at each $m \in M$, and therefore in a tubular neighborhood of M in E^*. Furthermore, on whole of E^*, the two-form Ω_{E^*} is closed, because ω and $d\alpha$ are closed. Therefore a neighborhood of M in E^* (which by abuse of notation we still denote by E^*), endowed with the above two form, is a symplectic manifold . It is clear that the pullback of Ω_{E^*} to M is exactly ω, and that the embedding of (M,ω) in (E^*, Ω_{E^*}) is coisotropic.

Proof of the Lemma: We only need to prove that at $m \in M$ we have $(i_{E^*,T^*M})^*(-d\alpha)(X,Y) = \langle(X_E, Y_{E^*})\rangle - \langle(X_{E^*}, Y_E)\rangle$.

We claim that $T_m i_{E^*,T^*M} : T_mE^* \to T_m(T^*M)$ is just the natural embedding of $E_m \times G_m \times E^*_m$ in $E_m \times G_m \times E^*_m \times G^*_m = T_m(T^*M)$. To this aim we just have to check that $T_m i_{E^*,T^*M}$ maps $(v, w, 0)$ to $(v, w, 0, 0)$ and $(0, 0, \xi)$ to $(0, 0, \xi, 0)$. Both statements are obvious. Now the lemma follows because

$$\begin{aligned}
(i_{E^*,T^*M})^*(-d\alpha)(X,Y) &= (-d\alpha)\big((X_E, X_G, X_{E^*}, 0), (Y_E, Y_G, Y_{E^*}, 0)\big) \\
&= \langle(X_E, X_G), (Y_{E^*}, 0)\rangle - \langle(Y_E, Y_G), (X_{E^*}, 0)\rangle \\
&= \langle X_E, Y_{E^*}\rangle - \langle Y_E, X_{E^*}\rangle.
\end{aligned}$$

□

Remark: When the distribution G is integrable, a computation in coordinates shows that Ω_{E^*} is non-degenerate on the whole vector bundle E^*, and not just in a neighborhood of the zero section.

7.2 The set $\mathcal{C}$ of coisotropic submanifolds C^1-close to M

We have already seen that a neighborhood of the coisotropic submanifold M in (N, Ω) is symplectomorphic to a neighborhood of M in $(E^*, \Omega_{E^*} = \pi^*\omega + (i_{E^*,T^*M})^*(-d\alpha))$. In analogy to the fact that Lagrangian submanifolds near a given one correspond to closed one forms, in this chapter we will characterize the sections of E^* that correspond to coisotropic submanifolds.
The inclusion $i_{E^*,T^*M} : E^* \hookrightarrow T^*M$ determined by the splitting $TM = E \oplus G$ induces an identification between sections of E^* and one forms on M which annihilate the distribution G, which we denote by $\Omega^1(M)_G$. So submanifolds of N C^1-close to M are identified with C^1-small elements of $\Gamma(E^*)$ and C^1-small elements of $\Omega^1(M)_G$.

Proposition 7.2.1 *Using the above identifications the coisotropic submanifolds of* (N, Ω) *which are* C^1*-close to* M *correspond to*

$$\mathcal{C} := \{\beta \in \Omega^1(M)_G \mid \beta \textit{ is } C^1\textit{-small and } \omega - d\beta \textit{ has the same rank as } \omega\}.$$

Remark 1: In Section 7.3 we will endow the set of smooth submanifolds of N which are C^1-close to M, $\Gamma(E^*)$ and $\Omega^1(M)_G$ with Fréchet manifold structures compatible with our identifications. Then the above statement can be made more precise by saying that the embedding of the Fréchet manifold of submanifolds of N C^1-close to M into the Fréchet manifold $\Omega^1(M)_G$ maps the coisotropic submanifolds onto $\mathcal{C}$.
Remark 2: When $M \subset (N, \Omega)$ is Lagrangian, i.e. when $\omega = 0$, we recover the well known description of Lagrangian submanifolds C^1-close to M as closed one-forms on M.

Proof: It is enough to show that under the above identification of $\Gamma(E^*)$ and $\Omega^1(M)_G$ the coisotropic submanifolds of (E^*, Ω_{E^*}) C^1-close to M are given by $\mathcal{C}$. So let $\tilde{\beta}$ be a section of the vector bundle $E^* \to M$ and let $i_{\mathrm{Im}(\tilde{\beta}),E^*}$ denote the inclusion into E^* of the image of $\tilde{\beta} : M \to E^*$. We will denote by β the corresponding one-form on M which annihilates $G \subset TM$. From dimension considerations we have

$$\text{Im}(\tilde{\beta}) \text{ is a coisotropic submanifold of } (E^*, \Omega_{E^*}) \Leftrightarrow$$
$$\text{the kernel of } (i_{\text{Im}(\tilde{\beta}),E^*})^*\Omega_{E^*} \text{ has constant dimension equal to } k.$$

We want to restate this condition in terms of the presymplectic manifold $(M, \tilde{\beta}^*(\Omega_{E^*}))$. We have

$$\begin{aligned} \tilde{\beta}^*(\Omega_{E^*}) &= \tilde{\beta}^*\big(\pi^*\omega + (i_{E^*,T^*M})^*(-d\alpha)\big) \\ &= \omega + (i_{E^*,T^*M} \circ \tilde{\beta})^*(-d\alpha) \\ &= \omega - d(\beta^*\alpha) \\ &= \omega - d\beta, \end{aligned}$$

where in the last equality we used that, for any manifold M and one-form β on M, the pullback (to M) of the canonical one-form on T^*M via β is exactly β (see [Ca, Chapter 3]). Therefore we have:

$$\text{Im}(\tilde{\beta}) \text{ is a coisotropic submanifold of } (E^*, \Omega_{E^*}) \Leftrightarrow$$
$$\text{the kernel of } (M, \omega - d\beta) \text{ has constant dimension equal to } k,$$

which finishes the proof. □

Now we consider a simple example that will be used again in the following sections.

Example 1: Consider the presymplectic manifold $(M, \omega) = (\mathbb{T}^4, dx_1 \wedge dx_2)$, where $(x_1, \cdots, x_4)$ are canonical coordinates on $\mathbb{T}^4$. The characteristic distribution is $E = \text{Span}\{\frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_4}\}$. We choose $G = \text{Span}\{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\}$. The symplectic manifold (E^*, Ω_{E^*}) as in Section 7.1 in which (M, ω) embeds coisotropically is easily seen to be $(N, \Omega) := (\mathbb{T}^4 \times \mathbb{R}^2, dx_1 \wedge dx_2 + dx_3 \wedge d\xi_3 + dx_4 \wedge d\xi_4)$ where (ξ_3, ξ_4) are canonical coordinates on $\mathbb{R}^2$.

Fact 7.2.1 shows that the set of coisotropic submanifolds of (N, Ω) which are C^1-close to of M is given exactly by

$$\mathcal{C} = \{f dx_3 + g dx_4 | f, g \in C^\infty(\mathbb{T}^4) \ C^1\text{-small} \ , \ dx_1 \wedge dx_2 - d(f dx_3 + g dx_4) \text{ has rank } 2\}.$$

Two-forms on $\mathbb{T}^4$ as above, at each point, can have rank 2 or rank 4. The first case occurs exactly when, at each point of $\mathbb{T}^4$, the determinant of the corresponding bilinear form is zero, i.e. when

$$f_4 - g_3 = f_1 g_2 - f_2 g_1. \qquad (\blacktriangleright)$$

Here the subscripts denote partial derivatives. Notice that this is a *non-linear* partial differential equation for f and g.

In this specific example, it is easy to determine directly the coisotropic submanifolds of (N, Ω) C^1-close to M without making use of Fact 7.2.1: any submanifold C^1-close to $\mathbb{T}^4$ is of the form $M_{f,g} = \text{graph}(f, g)$ where $(f, g) : \mathbb{T}^4 \to \mathbb{R}^2$. The condition $(TM_{f,g})^\Omega \subset TM_{f,g}$ translates exactly into the partial differential equation above.

7.3 Some properties of the set $\mathcal{C}$.

As earlier, let (N, Ω) be a symplectic manifold, (M, ω) a compact coisotropic submanifold with characteristic distribution E, and G a fixed complement of E in TM. In this section we will investigate some properties of the set of coisotropic submanifolds C^1-close to M, which we can identify with the set $\mathcal{C}$ as in Fact 7.2.1, and show in Theorem 7 that it does not have a "nice" structure.

When $\dim(M) = \frac{1}{2}\dim(N)$, i.e. when M is Lagrangian, $\mathcal{C}$ consists of the C^1-small elements of a subspace of $\Omega^1(M)_G = \Omega^1(M)$, namely the closed one-forms on M. When $\dim(M) = \dim(N) - 1$ the set $\mathcal{C}$ consists of all C^1-small sections of $\Omega^1(M)_G$, because any codimension one submanifold is coisotropic .

In the general case, looking into the leaves of the characteristic distributions of coisotropic submanifolds seems to indicate that the subset $\mathcal{C}$ of $\Omega^1(M)_G$ does *not* have a "nice" structure. We will exhibit this by considering the following two lemmas.

Lemma 7.3.1 *Consider $(M, \omega) = (\mathbb{T}^4, dx_1 \wedge dx_2)$ as a coisotropic submanifold of $(N, \Omega) = (\mathbb{T}^4 \times \mathbb{R}^2, dx_1 \wedge dx_2 + dx_3 \wedge d\xi_3 + dx_4 \wedge d\xi_4)$. Arbitrarily C^1-close to M there exist coisotropic submanifolds of (N, Ω) with characteristic leaves not homeomorphic to those of M.*

Proof: Let $(f,g) : \mathbb{T}^4 \to \mathbb{R}^2$ such that $f_4 - g_3 = f_1 g_2 - f_2 g_1$. Then, as in Example 1, $M_{f,g} := \text{graph}(f,g)$ is a coisotropic submanifold of $\mathbb{T}^4 \times \mathbb{R}^2$. As in the proof of Fact 7.2.1, $(M_{f,g}, (i_{M_{f,g},N})^*\Omega)$ is symplectomorphic to $(\mathbb{T}^4, dx_1 \wedge dx_2 - d(f dx_3 + g dx_4))$, which has characteristic distribution $E_{f,g}$ spanned by $(-f_2, f_1, 1, 0)$ and $(-g_2, g_1, 0, 1)$. Applying [CL, Prop. 1, Ch. V.2] we see that the restriction of the projection $pr : \mathbb{T}^4 \to \mathbb{T}^2, (x_1, \cdots, x_4) \mapsto (x_3, x_4)$ to any leaf of $E_{f,g}$ is a covering map. In particular the leaves are homeomorphic to either $\mathbb{R}^2$, $S^1 \times \mathbb{R}$ or $\mathbb{T}^2$. When $f = g = 0$, i.e. when $M_{f,g} = M$, the characteristic leaves are all homeomorphic to $\mathbb{T}^2$.

Now, for any $\epsilon > 0$, let $f = \epsilon \sin(2\pi x_1)$, $g = 0$. The pair (f,g) clearly satisfies equation (▶) of Section 7.1. Let γ be any curve in $\mathbb{T}^4$ tangent to $E_{f,g}$ and $(\bar{x_1}, \bar{x_2}, \bar{x_3}, \bar{x_4}) = \gamma(0)$. Using the fact that $E_{f,g}$ is always orthogonal to $\frac{\partial}{\partial x_1}$ we have

$$\dot{\gamma}(t) = \alpha(t) \begin{pmatrix} 0 \\ 2\pi\epsilon\cos(2\pi\bar{x_1}) \\ 1 \\ 0 \end{pmatrix} + \beta(t) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

for some functions α, β. Now we have

$$\begin{aligned} & \gamma \text{ is a closed curve} \\ \Leftrightarrow \quad & \exists t_0 : \gamma(t_0) - \gamma(0) = \int_0^{t_0} \dot{\gamma}(t)dt \in \mathbb{Z}^4 \\ \Leftrightarrow \quad & \exists t_0 : \int_0^{t_0} \alpha(t)dt \in \mathbb{Z}, \int_0^{t_0} \beta(t)dt \in \mathbb{Z}, \left(\int_0^{t_0} \alpha(t)dt\right) \cdot 2\pi\epsilon\cos(2\pi\bar{x_1}) \in \mathbb{Z}. \end{aligned}$$

Now suppose that the characteristic leaf L in which γ lies is homeomorphic to $\mathbb{T}^2$. Then, since the covering map $pr : L \to \mathbb{T}^2$ induces an injection at the level of fundamental groups, we can find a loop in $\mathbb{T}^2$ through $(\bar{x_3}, \bar{x_4})$ whose class lies in the image of $\pi_1(L)$ and which "winds around in x_3-direction" a non-zero number of times. The lift of this curve is a loop in L with $\int_0^{t_0} \alpha(t)dt \neq 0$. So the above conditions imply that $2\pi\epsilon\cos(2\pi\bar{x_1}) \in \mathbb{Q}$.

Therefore leaves through points $\bar{x}$ of $\mathbb{T}^4$ with $2\pi\epsilon\cos(2\pi\bar{x_1}) \notin \mathbb{Q}$ must be homeomorphic to $S^1 \times \mathbb{R}$ (they cannot be homeomorphic to $\mathbb{R}^2$ because the curve $(\bar{x_1}, \bar{x_2}, \bar{x_3}, \bar{x_4}) +$

$t(0,0,0,1)$ in L is closed and not contractible). All other leaves are easily seen to be homeomorphic to $\mathbb{T}^2$. Making ϵ arbitrarily small finishes the argument. $\square$

Now we want to consider the "formal tangent space" at M to the set of coisotropic submanifolds of (N,Ω). Under our identifications it corresponds to

$$FT_0\mathcal{C} := \Big\{ \frac{d}{dt}\Big|_0 \beta^t : \beta : (-\epsilon,\epsilon) \to \Omega^1(M)_G \text{ is smooth, } \mathrm{Im}(\beta) \subset \mathcal{C}, \beta^0 = 0 \Big\}.$$

Since we have not specified the differentiable structure of $\Omega^1(M)_G$ (this will be done in Section 7.3) the term "smooth" above is not defined. The definition of $FT_0\mathcal{C}$ and the proof of Lemma 7.3.2 below will become rigorous only after introducing the appropriate definitions.

Lemma 7.3.2

$$FT_0\mathcal{C} \subset \{B \in \Omega^1(M)_G | d_E B = 0\},$$

where d_E denotes the leaf-wise differential along the leaves of the characteristic distribution E.

Proof: Let β^t be as in the definition of $TF_0\mathcal{C}$. For each t the one-form β^t satisfies $\beta^t|_G = 0$, and $\omega - d\beta^t$ has same rank as ω. The first condition immediately implies $(\frac{d}{dt}|_0\beta^t)|_G = 0$.
The second condition implies that $\ker(\omega - d\beta^t)$ is a smooth time-dependent distribution on M, so locally we can choose a smooth basis $\{e_1^t, \cdots, e_k^t\}$ for it. For each i,j we have

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_0 [(\omega - d\beta^t)(e_i^t, e_j^t)] \\ &= \Big[\frac{d}{dt}\Big|_0(\omega - d\beta^t)\Big](e_i^0, e_j^0) + \omega\Big(\frac{d}{dt}\Big|_0 e_i^t, e_j^0\Big) + \omega\Big(e_i^0, \frac{d}{dt}\Big|_0 e_j^t\Big). \end{aligned}$$

Now $\{e_1^0, \cdots, e_k^0\}$ is a basis for $E = \ker(\omega)$, and $\frac{d}{dt}|_0(\omega - d\beta^t) = -d(\frac{d}{dt}|_0\beta^t)$, so the above identity implies that $d(\frac{d}{dt}|_0\beta^t)$ is zero on $E \wedge E$, i.e. $d_E(\frac{d}{dt}|_0\beta^t) = 0$ $\square$

Now we combine the statements of the above two lemmas. Lemma 7.3.2 says that the "formal tangent space" of the set of coisotropic submanifolds at some point M is contained in a vector space which depends on the characteristic leaves of M. Lemma

7.3.1 says that an arbitrary small perturbation of M can change drastically the topology of the characteristic leaves. Altogether this suggests that at different points of the set of coisotropic submanifolds, even within the same connected component, the "formal tangent spaces" are not isomorphic. This would imply that the set itself does not have a"nice" manifold structure.

Notice that in the Lagrangian case the characteristic leaves are all diffeomorphic, and that in the codimension one case the characteristic leaves are one-dimensional, so the closedness condition of Lemma 7.3.2 is vacuous. Therefore, in these two special cases, the argument above does not apply.

Now we are ready to show that the set $\mathcal{C}$ defined in Fact 7.2.1 does not have a "nice" structure. More precisely, we will show that $\mathcal{C}$ is not a submanifold of the Fréchet manifold $\Omega^1(M)_G$. We first recall the notion of Fréchet manifold following Richard Hamilton's work [Ha].

Definition: ([Ha 1.1.1] or [Fo 5.4]) A *Fréchet space* is a complete topological vector space with topology induced by a countable family of seminorms.

Example 2: ([Ha 1.1.5]) Let V be any smooth vector bundle over a compact manifold M. Choose a vector bundle metric and a connection ∇ on V, as well as a Riemannian metric on M. For each $n \in \mathbb{N}$ we have a norm on $\Gamma(V)$, the smooth sections of V, given by

$$||f||_n^2 := \sum_{j=0}^{n} \max_{x \in M} |\nabla^j f(x)|^2$$

where $\nabla^j f$ is the j-th covariant derivative of f and its norm is defined using any choice of orthonormal basis of T_xM and the bundle metric on V. (This is not the norm used in [Ha 1.1.5] but us equivalent to it.)

Then $\Gamma(V)$ together with the topology induced by $\{|| \cdot ||_n\}_{n \in \mathbb{N}}$ is a Fréchet space. Furthermore the induced topology is independent of the choices we made, so that this construction is canonical after we fix the smooth vector bundle V.

Definition: ([Ha 4.1.1]) A *Fréchet manifold* is a Hausdorff topological space with an atlas of coordinate charts taking their values in Fréchet spaces such that the coordinate transition maps are smooth.

Example 3: ([Ha 4.1.7]) Let N be a finite dimensional manifold and let $\mathcal{S}(N)$ be the space of all compact submanifolds of N. For each $M \in \mathcal{S}(N)$, any diffeomorphism between a tubular neighborhood of M and the normal bundle $TN|_M/TM$ gives rise to a bijection between a neighborhood of M in $\mathcal{S}(N)$ and a neighborhood of 0 in $\Gamma(TN|_M/TM)$, which by the above example is canonically a Fréchet space. Since the transition maps between charts as above are smooth ([Ha 4.4.7]), $\mathcal{S}(N)$ is endowed with the structure of a Fréchet manifold . Notice that each connected component of $\mathcal{S}(N)$ consists of manifolds diffeomorphic to each other, and that the Fréchet manifold structure on $\mathcal{S}(N)$ is canonical after fixing the differentiable structure on N.

Now let us apply the above definition to the usual setting where (M, ω) is a compact coisotropic submanifold of the symplectic manifold (N, Ω), E is the characteristic distribution of (M, ω), and G is a fixed complement of E in TM. As shown in Example 3, the set $\mathcal{S}(N)$ of compact submanifolds of N has a canonical Fréchet manifold structure, and the same holds for $\mathcal{S}(E^*)$. In Example 2 we saw that $\Gamma(E^*)$ and $\Omega^1(M)_G = \Gamma(G^\circ)$ have canonical structures of Fréchet spaces and hence of Fréchet manifolds. Here $G^\circ \subset T^*M$ is the annihilator of G in TM. It is clear that the identifications we made at the beginning of Section 7.2 respect these structures, so the statement of Remark 1 in Section 7.2 follows at once.

Now we can make precise and prove that the set $\mathcal{C}$ does not have a"nice" structure:

Theorem 7 *Let (N, Ω) be a symplectic manifold and M a compact coisotropic submanifold. Then the set $\mathcal{C}$ as in Fact 7.2.1 is* not *a Fréchet submanifold of $\Omega^1(M)_G$.*
It follows that the set of compact coisotropic submanifolds of (N, Ω) is not *a Fréchet submanifold of $\mathcal{S}(N)$, the collection of all compact submanifolds of N.*

Proof: We consider again Example 1, i.e. $(N, \Omega) = (\mathbb{T}^4 \times \mathbb{R}^2, dx_1 \wedge dx_2 + dx_3 \wedge d\xi_3 + dx_4 \wedge d\xi_4)$, $M = \mathbb{T}^4 \subset N$, $G = \text{Span}\{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\}$. There we showed that

$$\mathcal{C} = \{f dx_3 + g dx_4 | f, g \text{ are } C^1\text{-small and } f_4 - g_3 = f_1 g_2 - f_2 g_1\}.$$

To prove the theorem it is enough to show that $FT_0\mathcal{C}$, as defined in Section 7.3, is not a vector subspace of $\Omega^1(M)_G = C^\infty(\mathbb{T}^4)^{\otimes 2}$. Let first $(f,g) : (-\epsilon,\epsilon) \to C^\infty(\mathbb{T}^4)^{\otimes 2}, t \mapsto (f^t, g^t)$ be any smooth curve with image lying in $\mathcal{C}$ such that $(f^0, g^0) = (0,0)$. For all t we have

$$0 = \int_0^1 \int_0^1 (f_4^t - g_3^t) dx_3 dx_4 = \int_0^1 \int_0^1 (f_1^t g_2^t - f_2^t g_1^t) dx_3 dx_4.$$

Applying $\frac{d^2}{dt^2}|_0$ and using the notation $F := \frac{d}{dt}|_0 f^t, G := \frac{d}{dt}|_0 g^t$ we obtain $\int_0^1 \int_0^1 (F_1 G_2 - F_2 G_1) dx_3 dx_4 = 0$. Therefore all elements (F,G) of $FT_0\mathcal{C}$ are subject to the above constraint, which is clearly *non-linear*.

Now consider the curves $(f^t, g^t) = (t\sin(2\pi x_1), 0)$ and $(\tilde{f}^t, \tilde{g}^t) = (0, t\sin(2\pi x_2))$. Both curves lie in $\mathcal{C}$, so their derivatives at time zero (F,G) and $(\tilde{F},\tilde{G})$ lie in $FT_0\mathcal{C}$. But their sum $(F+\tilde{F}, G+\tilde{G})$ does not, because it violates the above constraint:

$$\int_0^1 \int_0^1 ((F+\tilde{F})_1 (G+\tilde{G})_2 - (F+\tilde{F})_2 (G+\tilde{G})_1) dx_3 dx_4 =$$
$$\int_0^1 \int_0^1 \sin(2\pi x_1)\sin(2\pi x_2) dx_3 dx_4 \neq 0.$$

□

Remark: In the particular case of Example 1 we have

$$FT_0\mathcal{C} \subset \Big\{ F dx_3 + G dx_4 \mid F_4 - G_3 = 0 \;,\; \int_0^1 \int_0^1 (F_1 G_2 - F_2 G_1) dx_3 dx_4 = 0 \Big\}.$$

The first restriction is obtained by linearizing the equation $f_4 - g_3 = f_1 g_2 - f_2 g_1$ and is equivalent to the fact that the restriction of $F dx_3 \wedge G dx_4$ to the characteristic leaves of $\mathbb{T}^4$ be closed (compare to Lemma 7.3.2). The second restriction was derived in the proof of Theorem 7. These two restrictions can be explained in terms of a strong homotopy Lie Algebroid structure (see [OP]) on the vector bundle $E \to M$, and Corollary 10.5 and Example 10.3 of [OP] show that the right hand side of the above inclusion is exactly the space of infinitesimal *formal* deformations of $(\mathbb{T}^4, dx_1 \wedge dx_2)$.

Appendix A

The estimates of Proposition 4.2.1

Here we will prove the estimates used in the proof of Proposition 4.2.1 in Section 4.2. See Section 4.2 for the notation.

Applying a series of lemmas we will show first that

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

Then we will show

$$\measuredangle((\exp_A^{-1})_* \dot{CB}\,,\ \tilde{B} - \tilde{C}) < 3d(A, C),$$

from which it will follow very easily that

$$\measuredangle({}_{CA}\backslash\!\backslash \dot{CB}\,,\ Q - 0) < 4d(C, A).$$

This will conclude the proof of claim 2 in the proof of Proposition 4.2.1.

Finally using standard arguments we will obtain the estimate used in claim 3, namely

$$|\tilde{B} - \tilde{C}| \leq \frac{11}{10} d(C, B).$$

A.1 Estimates

We begin by stating an easy result without proof:

Fact *Let* $c : [0,1] \to \mathbb{R}^n$ *is a differentiable curve such that* $|\dot{c}(s) - \dot{c}(0)| \leq \varepsilon$ *for all* $s \in [0,1]$. *Then*

$$\big|[c(1) - c(0)] - \dot{c}(0)\big| \leq \varepsilon.$$

If in addition $\mathcal{C} \leq |\dot{c}(s)| \leq \mathcal{D}$ *for all* $s \in [0,1]$, *then*

$$\cos\Big(\measuredangle[c(1) - c(0)\,,\,\dot{c}(0)]\Big) \geq \frac{\mathcal{C}^2 - \frac{\varepsilon^2}{2}}{\mathcal{D}^2}.$$

Now we introduce the following notation: $\sigma(s)$ will be the shortest geodesic from C to B, defined on $[0,1]$, and $\tilde{\sigma}(s)$ will be its lift to T_AM via $\exp_A$. For any $s \in [0,1]$, $f_s(t) = \exp_A(t\tilde{\sigma}(s))$ will be the shortest geodesic from A to $\sigma(s)$. Notice that, since $d(A,C) + d(C,B) \leq 0.7$, we have $|\tilde{\sigma}(s)| \leq 0.7$.

For a fixed s, consider the geodesic triangle $A\sigma(s)C$ in M, and denote by α, β_s and γ the angles at $A, \sigma(s)$ and C respectively.

Lemma A.1.1

$$\textit{If } \frac{8}{7}\frac{d(C,A)}{\sin(70°)} \leq |\tilde{\sigma}(s)|, \textit{ then } \sin(\beta_s) \leq \frac{8}{7}\frac{d(C,A)}{|\tilde{\sigma}(s)|} \leq \sin(70°).$$

Remark: The lemma says that if $\sigma(s)$ is far enough from A, the direction of the geodesic σ there does not deviate too much from being radial with respect to A. This will allow us to estimate the radial and orthogonal-to-radial components of $\dot{\sigma}(s)$ in Lemma A.1.3 and Lemma A.1.5.

Proof: Consider the Alexandrov triangle in S^2 corresponding to $A\sigma(s)C$, i.e. a triangle in S^2 with the same side lengths as $A\sigma(s)C$, and denote its angles by α', β'_s and γ'. Consider also the corresponding triangle in standard hyperbolic space H^2 and denote its angles by α'', β''_s and γ''. The sum of the angles of the triangle in S^2 deviates from 180° by the area of the triangle, which is bounded by $0.7 \cdot 0.15 < 10°$ (see [BK] 6.7 and 6.7.1.), so that the sum is at most 190°. For the same reason the sum of the angles of the triangle in H^2 is at least 170°. Since for the sectional curvature κ we have $|\kappa| \leq 1$, by Toponogov's Theorem (see Thm. 2.7.6 and Thm. 2.7.12 in [Kl]) we have $\beta''_s \leq \beta_s \leq \beta'_s$ (and similarly for α, γ) and by the above $\beta'_s - \beta''_s < 20°$.

Case 1: $\beta_s \leq 90°$.

By the law of sines (see [Kl, Remark 2.7.5]) we have

$$\sin\beta'_s = \sin(d(A,C)) \cdot \frac{\sin\gamma'}{\sin d(\sigma(s),A)} \leq \frac{d(A,C)}{\sin d(\sigma(s),A)} \leq \frac{8}{7}\frac{d(A,C)}{d(\sigma(s),A)},$$

where we used $\frac{sinx}{x} > \frac{7}{8}$ for $x \in [0, 0.8]$. By our assumption this quantity is bounded above by $\sin(70°)$, so $\beta'_s \in [0, 70°]$ or $\beta'_s \in [110°, 180°]$. Since $\beta'_s - \beta_s < 20°$ and $\beta_s \leq 90°$, we must have $\beta'_s \in [0, 70°]$. Therefore $\beta_s \leq \beta'_s$ implies $\sin\beta_s \leq \sin\beta'_s$, which we already bounded above.

Case 2: $\beta_s > 90°$.
One has to proceed analogously, but comparing with triangles in H^2. □

We now state a general fact about the exponential map:

Lemma A.1.2 *If* γ *is a geodesic parametrized by arc length and* $W \in T_{\gamma(0)}M$*, then for* $t < 0.7$

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

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 $J(0) = 0, J'(0) = W$ is given by $J(t) = (d_{t\dot\gamma(0)} \exp_{\gamma(0)})(tW)$ (see [Jo, Cor. 4.2.2]). The estimate $\frac{\sinh(t)-t}{t}$ follows from [BK 6.3.8iii]. This expression is bounded above by $\frac{t^2}{5}$ 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. □

Using Lemma A.1.1 we obtain a refinement of Lemma A.1.2 for $\dot\sigma(s) = (d_{\tilde\sigma(s)} \exp_A)\dot{\tilde\sigma}(s)$:

Lemma A.1.3 *For all* $s \in [0, 1]$

$$\left| {}_{f_s}\backslash\backslash \dot\sigma(s) - \dot{\tilde\sigma}(s) \right| \leq \frac{1}{5} d(C,A)|\dot\sigma(s)|.$$

Proof: By the Gauss Lemma, the decomposition into radial (denoted by ^{R}) and orthogonal-to-radial (denoted by $^\perp$) components of $\dot\sigma(s) = (d_{\tilde\sigma(s)}\exp_A)\dot{\tilde\sigma}(s) \in T_{\sigma(s)}M$ and $\dot{\tilde\sigma}(s) \in T_{\tilde\sigma(s)}(T_AM)$ is preserved by $d_{\tilde\sigma(s)}\exp_A$. Hence ${}_{f_s}\!\!\setminus\!\!\setminus\dot\sigma(s)^R = {}_{f_s}\!\!\setminus\!\!\setminus(d_{\tilde\sigma(s)}\exp_A)[\dot{\tilde\sigma}(s)^R] = \dot{\tilde\sigma}(s)^R \in T_AM$. Introducing the notation $D := |\tilde\sigma(s)| < 0.7$ this implies

$$\left|{}_{f_s}\!\!\setminus\!\!\setminus\dot\sigma(s) - \dot{\tilde\sigma}(s)\right| = \left|{}_{f_s}\!\!\setminus\!\!\setminus\dot\sigma(s)^\perp - \dot{\tilde\sigma}(s)^\perp\right| \leq \frac{D^2}{5}|\dot{\tilde\sigma}(s)^\perp|$$

by Lemma A.1.2. From this we obtain $|\dot{\tilde\sigma}(s)^\perp| \leq \frac{|\dot\sigma(s)^\perp|}{1-\frac{D^2}{5}} \leq \frac{5}{4}|\dot\sigma^\perp(s)|$. So $|{}_{f_s}\!\!\setminus\!\!\setminus\dot\sigma(s) - \dot{\tilde\sigma}(s)| \leq \frac{D^2}{4}|\dot\sigma^\perp(s)|$.

Case 1: $\frac{8}{7}\frac{d(C,A)}{\sin(70°)} \leq D$.
Using the notation of Lemma A.1.1 we have $|\dot\sigma(s)^R| = |\cos\beta_s|\cdot|\sigma(s)|$, so $|\dot\sigma(s)^\perp| = \sin(\beta_s)|\dot\sigma(s)|$. Applying the same lemma we obtain $|\dot\sigma(s)^\perp| \leq \frac{8}{7}\frac{d(C,A)}{D}|\dot\sigma(s)|$. Therefore $\frac{D^2}{4}|\dot\sigma(s)^\perp| \leq \frac{D^2}{4}\frac{8}{7}\frac{d(C,A)}{D}|\dot\sigma(s)| \leq \frac{1}{5}d(C,A)|\dot\sigma(s)|$.

Case 2: $\frac{8}{7}\frac{d(C,A)}{\sin(70°)} > D$.
We simply use $|\dot\sigma(s)^\perp| \leq |\dot\sigma(s)|$ to estimate

$$\frac{D^2}{5}|\dot\sigma(s)^\perp| \leq \left(\frac{8}{7}\frac{d(C,A)}{\sin(70°)}\right)^2\frac{1}{5}|\dot\sigma(s)| \leq \frac{3}{10}d(C,A)^2|\dot\sigma(s)|,$$

which is less than the estimate of *Case 1* since $d(C,A) \leq 0.15$. □

We want to apply the above Fact to the curve $\tilde\sigma : [0,1] \to T_AM$. We begin by determining the constant ε.

Lemma A.1.4 *For all* $s \in [0,1]$

$$|\dot{\tilde\sigma}(s) - \dot{\tilde\sigma}(0)| \leq d(C,A)\cdot d(C,B).$$

Proof: We will show that $\dot{\tilde\sigma}(s) \approx {}_{f_s}\!\!\setminus\!\!\setminus\dot\sigma(s) \approx {}_{f_0}\!\!\setminus\!\!\setminus\dot\sigma(0) \approx \dot{\tilde\sigma}(0)$, where in the first and last relation tangent spaces are identified by parallel translation along f_s and f_0 respectively. The estimates needed for the first and last relation are given by Lemma A.1.3 since $|\dot\sigma(s)| = d(C,B)$ for all s.
For the second relation we use the fact that σ is a geodesic and [BK, 6.2.1]. We see that $|{}_{f_s}\!\!\setminus\!\!\setminus\dot\sigma(s) - {}_{f_0}\!\!\setminus\!\!\setminus\dot\sigma(0)|$ is bounded above by the product of $|\dot\sigma(0)| = d(C,B)$ and of the area

spanned by the triangle $\sigma(s)AC$ - which is surely less than $0.6\, d(C,A)$ by [BK, 6.7.1] - and we are done. $\square$

Now we determine the constants $\mathcal{C}$ and $\mathcal{D}$. Recall that $|\dot{\sigma}(s)| = d(C,B)$, so that the constants we obtain are really independent of s.

Lemma A.1.5 *For all $s \in [0,1]$*

$$\frac{|\dot{\sigma}(s)|}{\sqrt{1+d(C,A)^2}} \leq |\dot{\tilde{\sigma}}(s)| \leq \frac{|\dot{\sigma}(s)|}{\sqrt{1-d(C,A)^2}}.$$

Remark: Since we need estimates involving $d(C,A)$, the classical Jacobi-field estimate $|\dot{\tilde{\sigma}}(s)| \leq \frac{D}{\sin(D)}|\dot{\sigma}(s)|$ or $|\dot{\tilde{\sigma}}(s)| \leq (1-\frac{D^2}{5})^{-1}|\dot{\sigma}(s)|$ (given by Lemma A.1.2) are not good enough for us. Here $D := d(\sigma(s),A)$. Furthermore, making use of Lemma A.1.1, in Lemma A.1.3 we bounded above $|_{f_s}\backslash\!\!\backslash\dot{\sigma}(s) - \dot{\tilde{\sigma}}(s)|$, which at once implies $|\dot{\tilde{\sigma}}(s)| \leq (1+\frac{d(C,A)}{5})|\dot{\sigma}(s)|$. This however is also not sufficient, because it would only allow us to bound the left hand side of Corollary A.1.2 by a multiple of $\sqrt{d(C,A)}$. Instead, we will use Lemma A.1.1 to bound directly $\frac{|\dot{\sigma}(s)|}{|\dot{\tilde{\sigma}}(s)|}$ in order to obtain a bound of the form $[1+o(d(C,A)^2)]|\dot{\sigma}(s)|$ in Lemma A.1.5.

Proof: We fix s and adopt the notation $D := |\tilde{\sigma}(s)| = d(\sigma(s),A) < 0.7$.
Case 1: $D < \frac{8}{7}\frac{d(C,A)}{\sin(70°)}$.
In Case 2 of the proof of Lemma A.1.3 we showed that $|_{f_s}\backslash\!\!\backslash\dot{\sigma}(s) - \dot{\tilde{\sigma}}(s)| \leq \frac{3}{10}d(C,A)^2|\dot{\sigma}(s)|$. So

$$|\dot{\sigma}(s)|\left(1-\frac{3}{10}d(C,A)^2\right) \leq |\dot{\tilde{\sigma}}(s)| \leq |\dot{\sigma}(s)|\left(1+\frac{3}{10}d(C,A)^2\right).$$

Case 2: $D \geq \frac{8}{7}\frac{d(C,A)}{\sin(70°)}$.
In view of the remark above we will use simple Jacobi field estimates only for the orthogonal-to-radial components of $\dot{\sigma}(s)$ and $\dot{\tilde{\sigma}}(s)$, which we can bound above using Lemma A.1.1, whereas for the radial components we just have to notice that they have the same length.
Recall that $f_s(t) = \exp_A(t\tilde{\sigma}(s))$ is a variation of geodesics emanating from A. We denote by $\tilde{f}_s(t) = t\tilde{\sigma}(s)$ its lift to T_AM. Fixing $s \in [0,1]$, we denote by $J(t)$ and

$\tilde{J}(t)$ the Jacobi-fields along f_s and $\tilde{f}_s$ arising from the variations $f_s(t)$ and $\tilde{f}_s(t)$. So $J(t) = (d_{t\tilde{\sigma}(s)} \exp_A) t\dot{\tilde{\sigma}}(s)$ and $\tilde{J}(t) = t\dot{\tilde{\sigma}}(s)$ have the same initial covariant derivative $E := \dot{\tilde{\sigma}}(s)$.

We decompose $E \in T_A M$ into its components parallel and orthogonal to $\tilde{\sigma}(s)$ as $E = E^R + E^\perp$. $J^R(t) := (d_{t\tilde{\sigma}(s)} \exp_A) tE^R$ and $J^\perp(t) := (d_{t\tilde{\sigma}(s)} \exp_A) tE^\perp$ are Jacobi fields, since $\exp_A$ maps lines through $0 \in T_A M$ to geodesics through $A \in M$. They both vanish at zero and their initial covariant derivatives are E^R and $E^\perp$ respectively, so by the uniqueness of Jacobi fields to given initial data we have $J = J^R + J^\perp$. The Gauss Lemma implies that this is the decomposition of J into radial and orthogonal-to-radial components.

To show the second inequality in the statement of the lemma we want to bound below

$$\frac{|\dot{\sigma}(s)|^2}{|\dot{\tilde{\sigma}}(s)|^2} = \frac{|J(1)|^2}{|\tilde{J}(1)|^2} = \frac{|E^R|^2 + |J^\perp(1)|^2}{|E|^2}.$$

Applying Lemma A.1.2 to $J(1)^\perp = (d_{\sigma(s)} \exp_A)\dot{\tilde{\sigma}}(s)^\perp$ we obtain $|J(1)^\perp| \geq (1 - \frac{D^2}{5})|E^\perp|$, so

$$\frac{|\dot{\sigma}(s)|^2}{|\dot{\tilde{\sigma}}(s)|^2} \geq \frac{|E^R|^2 + (1-\frac{D^2}{5})^2|E^\perp|^2}{|E|^2} = 1 + \left(-\frac{2}{5}\frac{D^2}{3} + \frac{D^4}{25}\right)\frac{|E^\perp|^2}{|E|^2}. \qquad (\blacktriangledown)$$

Now we bound $\frac{|E^\perp|^2}{|E|^2}$ from above. We already saw that $|E^\perp| \leq \frac{|J^\perp(1)|}{1-\frac{D^2}{5}}$, and by Lemma A.1.2 $|\dot{\sigma}(s)| \leq (1 + \frac{D^2}{5})|\dot{\tilde{\sigma}}(s)|$. Both $\frac{1}{1-\frac{D^2}{5}}$ and $1 + \frac{D^2}{5}$ are bounded by $\frac{8}{7}$ since $D < 0.7$. Now Lemma A.1.1 allows us to relate $|J^\perp(1)|$ and $|\dot{\sigma}(s)| \in T_{\sigma(s)}M$. Namely, we have $|J^\perp(1)| = \sin(\beta_s)|\dot{\sigma}(s)| \leq \frac{8}{7}\frac{d(C,A)}{D}|\dot{\sigma}(s)|$. The last three estimates give $|E^\perp| \leq (\frac{8}{7})^3\frac{d(C,A)}{D}|E|$. Substituting into ($\blacktriangledown$) gives

$$\frac{|\dot{\sigma}(s)|^2}{|\dot{\tilde{\sigma}}(s)|^2} \geq 1 + \left(\frac{8}{7}\right)^6\left(-\frac{2}{5}\frac{D^2}{3} + \frac{D^4}{25}\right)\frac{d(C,A)^2}{D^2} \geq 1 - \left(\frac{8}{7}\right)^6\frac{2}{5}d(C,A)^2.$$

To take care of the first inequality in the statement of the lemma we show that

$$\frac{|\dot{\sigma}(s)|^2}{|\dot{\tilde{\sigma}}(s)|^2} \leq 1 + \left(\frac{8}{7}\right)^6\frac{3}{7}d(C,A)^2$$

by repeating the above proof and using the estimate $|J^\perp(1)| \leq (1 + \frac{D^2}{5})|E^\perp|$.

To finish the proof we have to compare the estimates obtained in *Case 1* and *Case 2.*

To do so notice that $(\frac{8}{7})^6\frac{3}{7} < 1$, that $(1+\frac{3}{10}d(C,A)^2) \leq \frac{1}{\sqrt{1-d(A,C)^2}}$ and $\frac{1}{\sqrt{1+d(A,C)^2}} \leq (1-\frac{3}{10}d(C,A)^2)$. □

Now finally we can apply the Fact A.1 to the curve $\tilde{\sigma} : [0,1] \to T_A M$. The first statement of Fact A.1 allows us to prove

Corollary A.1.1

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

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

$$\begin{aligned}
\left|P-P'\right| \leq& \left|(Q-0) - {}_{CA}\!\!\setminus\!\!\setminus\dot{CB}\right| \\
\leq& \left|(Q-0) - (\exp_A^{-1})_*\dot{CB}\right| + \left|(\exp_A^{-1})_*\dot{CB} - {}_{CA}\!\!\setminus\!\!\setminus\dot{CB}\right| \\
\leq& d(C,A)|\dot{CB}| + \frac{d(C,A)^2}{5}|\dot{CB}|.
\end{aligned}$$

In the last inequality we used $Q-0 = \tilde{B}-\tilde{C} = \tilde{\sigma}(1)-\tilde{\sigma}(0)$ and $(\exp_A^{-1})_*\dot{CB} = \dot{\tilde{\sigma}}(0)$ to apply the first statement of Fact A.1 (with ε given by Lemma A.1.4) for the first term and Lemma A.1.2 for the second term.
On the other hand we have

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

Therefore we have

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

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

$$\measuredangle(P'-0, P'-P) \leq \frac{8}{7}\sin\left(\measuredangle(P'-0, P'-P)\right) \leq 2d(C,A).$$

□

The second statement of Fact A.1 delivers

Corollary A.1.2

$$\measuredangle((\exp_A^{-1})_* \dot{CB}, \tilde{B} - \tilde{C}) < 3d(C, A).$$

Proof: Applying again Fact A.1 to the curve $\tilde{\sigma} : [0, 1] \to T_A M$ we want to obtain an estimate for $\alpha := \measuredangle(\tilde{\sigma}(1) - \tilde{\sigma}(0) \ , \ \dot{\tilde{\sigma}}(0)) = \measuredangle(\tilde{B} - \tilde{C} \ , \ (\exp_A^{-1})_* \dot{CB})$. Lemma A.1.4 and Lemma A.1.5 deliver the estimates

$$\varepsilon = d(C, A) \cdot d(C, B), \quad \mathcal{C} = \frac{d(C, B)}{\sqrt{1 + d(C, A)^2}} \quad \text{and } \mathcal{D} = \frac{d(C, B)}{\sqrt{1 - d(C, A)^2}}.$$

Therefore, using the abbreviation $d := d(C, A)$, we have

$$\cos(\alpha) \geq \frac{\mathcal{C}^2 - \frac{\varepsilon^2}{2}}{\mathcal{D}^2} = 1 - d^2 \frac{5 - d^4}{2(1 + d^2)}.$$

So

$$\sin^2(\alpha) \leq d^2 \left[\frac{5 - d^4}{1 + d^2} - d^2 \left(\frac{5 - d^4}{2(1 + d^2)} \right)^2 \right].$$

Notice that due to the restriction $d < 0.15$ we have $\cos(\alpha) > \frac{\sqrt{3}}{2}$, so that $|\alpha| < \frac{\pi}{6} < 0.8$. So

$$|\alpha| \leq \frac{8}{7} |\sin(\alpha)| \leq \frac{8}{7} d \cdot \sqrt{\frac{5 - d^4}{1 + d^2} - d^2 \left(\frac{5 - d^4}{2(1 + d^2)} \right)^2} \leq 3d.$$

□

The above corollary estimates the angle of two vectors based at $\tilde{C} = \exp_A^{-1}(C)$. Now we will estimate the angle of certain vectors based at $0 \in T_A M$.

Corollary A.1.3

$$\measuredangle(Q - 0, {}_{CA}\backslash\backslash \dot{CB}) < 4d(C, A).$$

Proof: Since $\tilde{B} - \tilde{C} = Q - 0$ we just have to estimate

$$\measuredangle\big({}_{CA}\backslash\backslash \dot{CB} \ , \ (\exp_A^{-1})_* \dot{CB}\big) \quad = \quad \measuredangle\big({}_{f_0}\backslash\backslash \dot{\sigma}(0) \ , \ \dot{\tilde{\sigma}}(0)\big)$$

and apply the triangle inequality together with Corollary A.1.2.

Denoting by L the distance from ${}_{f_0}\|\dot{\sigma}(0) \in T_A M$ to the line spanned by $\dot{\tilde{\sigma}}(0)$, we have $L \leq |{}_{f_0}\|\dot{\sigma}(0) - \dot{\tilde{\sigma}}(0)| \leq \frac{1}{5} d(C,A)d(C,B)$, where we used Lemma A.1.3 in the second inequality. Hence

$$\frac{7}{8} \cdot \measuredangle\big({}_{f_0}\|\dot{\sigma}(0)\ ,\ \dot{\tilde{\sigma}}(0)\big) < \sin\big(\measuredangle({}_{f_0}\|\ \dot{\sigma}(0)\ ,\ \dot{\tilde{\sigma}}(0))\big) = \frac{L}{|\dot{\sigma}(0)|} \leq \frac{1}{5} d(C,A)$$

where we used $\frac{\sin(x)}{x} > \frac{7}{8}$ for $x \in [0, 0.8]$.

Combining this with Corollary A.1.2 gives

$$\measuredangle\left({}_{CA}\|\dot{CB}, \tilde{B} - \tilde{C}\right) < 3d(A,C) + \frac{8}{35} d(A,C) < 4d(A,C).$$

□

We conclude this appendix by deriving the estimate need in Claim 3 of Proposition 4.2.1.

Corollary A.1.4

$$|\tilde{B} - \tilde{C}| < \frac{11}{10} d(C,B).$$

Proof: This follows easily from Lemma A.1.2 since

$$|\tilde{B} - \tilde{C}| \leq \int_0^1 |\dot{\tilde{\sigma}}(s)| ds \leq \frac{11}{10} d(C,B).$$

□

Appendix B

The proof of Proposition 4.3.1

Here we will prove Proposition 4.3.1 of Section 4.3, namely the estimate

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

To do so we will use the fact that N is C^1-close to N_g, see Lemma B.1.3.

In addition to the notation introduced in Section 4.3 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 $pr_{\dot\gamma(t)}Hor^g$ the projection of $\dot\gamma(t) \in T_{\gamma(t)}M$ onto $Vert^g_{\gamma(t)}$ *along* $Hor^g_{\gamma(t)}$. We will also use $pr_{\dot\gamma(t)}aHor^g$ and $pr_{\dot\gamma(t)}LC^g$ to denote projections onto $aVert^g_{\gamma(t)}$ *along* $aHor^g_{\gamma(t)}$ and $LC^g_{\gamma(t)}$ respectively.

Our strategy will be to bound above $|\frac{\nabla^{\perp}}{dt}\tilde{c}(t)| = |\exp_*^{-1}(pr_{\dot c(t)}LC^g)|$ (see Lemma B.1.3) using

$$TN \approx Hor^g \approx aHor^g \approx LC^g.$$

Integration along π_c will deliver the desired estimate.

B.1 Estimates

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

Lemma B.1.1 *If $L < 0.08$ and p is a point in $\partial \exp_{N_g}(\nu N_g)_L$, then*

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

Proof: It is enough to show that, if $Y \in LC^g_p$ is a unit vector, $|pr_Y aHor^g| \leq \frac{9}{5}L$.
Let $\beta(s)$ be a curve tangent to the distribution LC^g such that $\beta(0) = p, \dot{\beta}(0) = Y$. 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 4.1.2): we have $\frac{\partial}{\partial t}|_0 f_s(t) = \xi(s)$, so

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

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 $J(L) = pr_Y aHor^g \in aVert^g_p$. By Lemma 4.1.9, using the fact that Y is a unit vector, we have

$$|pr_Y aHor^g|^2 = \langle pr_Y aHor^g, Y\rangle = |\langle J(L), K(L)\rangle| \leq \frac{9}{5}L \cdot |pr_Y aHor^g|$$

and we are done. □

Lemma B.1.2 *Let $L < 0.08$. For any point p in $\partial \exp_{N_g}(\nu N_g)_L$ the projections $T_pM \to aVert^g_p$ along $aHor^g_p$ and LC^g_p differ at most by $2L$ in the operator norm.*

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

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

where we used [We, Cor. A.5] in the second inequality and Lemma B.1.1 in the third one. □

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

Lemma B.1.3 *For all t*

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

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 this Ehresmann connection to LC^g and tangent spaces to the fibers of νN_g to $aVert^g$, we have

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

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

The fact that N is C^1-close the N_g (see Theorem 1), since $\dot{c}(t) \in T_{c(t)}N$ implies that $\measuredangle(\dot{c}(t), Hor^g_{c(t)}) \leq 2500\epsilon$. By [We, Prop. 3.7] $d(Hor^g_{c(t)}, aHor^g_{c(t)}) \leq \frac{\epsilon}{4}$ since $d(c(t), N_g) \leq 100\epsilon$.

Therefore $\measuredangle(\dot{c}(t), aHor^g_{c(t)}) \leq 2501\epsilon$ and $|pr_{\dot{c}(t)}aHor^g| \leq \sin(2501\epsilon) \leq 2501\epsilon$.

On the other hand, by Lemma B.1.2, $|pr_{\dot{c}(t)}aHor^g - pr_{\dot{c}(t)}LC^g| \leq 200\epsilon$. The triangle inequality therefore gives $|pr_{\dot{c}(t)}LC^g| \leq \sin(2701\epsilon)$. Therefore, using Lemma A.1.2 and $\epsilon < \frac{1}{20000}$,

$$|\exp_*^{-1}(pr_{\dot{c}(t)}LC^g)| \leq \frac{1}{1-\frac{\epsilon}{5}}\left|pr_{\dot{c}(t)}LC^g\right| \leq 2702\epsilon.$$

□

Lemma B.1.3 allows us to bound $\left|\exp^{-1}C - {}^{\perp}_{\pi^b}\backslash\!\backslash\left(\exp^{-1}A\right)\right|$ 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 $r := 100\epsilon + \frac{L(\gamma)}{2}$. Notice also that $r < 0.08$ due to our restrictions on ϵ and $d(C, A)$.

Lemma B.1.4

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

Proof: Since $\varphi_{g*}(\pi_{N_g*}\dot{c}(t)) = \dot{c}(t)$, by Proposition 4.1.1 we have $|\dot{c}(t) - \backslash\!\backslash(\pi_{N_g*}\dot{c}(t))| \leq 3200\epsilon|\pi_{N_g*}\dot{c}(t)|$, so

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

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

Proof of Proposition 4.3.1: We have

$$\begin{aligned}
\left|\exp^{-1}C - {}^{\perp}_{\pi^b}\backslash\!\backslash\left(\exp^{-1}A\right)\right| &= \left|\int_0^{L(c)} \frac{d}{dt}\, {}^{\perp}_{\pi^b_c}\backslash\!\backslash\tilde{c}(t)\mathrm{dt}\right| \\
&= \left|\int_0^{L(c)} {}^{\perp}_{\pi^b_c}\backslash\!\backslash\frac{\nabla^{\perp}}{dt}\tilde{c}(t)\mathrm{dt}\right| \\
&\leq 2702\epsilon L(c) \\
&\leq 2702\epsilon\frac{(1 + 3200\epsilon)}{f(r)}L(\gamma)
\end{aligned}$$

where we used Lemma B.1.3 and lemma B.1.4 in the last two inequalities. The proposition follows using the bound $\epsilon < \frac{1}{20000}$. □

Appendix C

The proof of Proposition 4.3.2

Here we will prove Proposition 4.3.2 of Section 4.3, i.e. the estimate

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

We will use the fact that N_g has bounded second fundamental form (see the first statement of Lemma C.1.3) and that γ is a geodesic (see the second statement of the same Lemma).

We will use the notation introduced in Section 4.3 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.

C.1 Estimates

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

Lemma C.1.1 *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.1.3 we have $\exp_*(\widehat{\frac{\nabla^{\perp}}{dt}\tilde{\gamma}(0)}) = pr_{\dot{\gamma}(0)}LC^g$, 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 $d(Hor^g_C\ ,\ aHor^g_C) \leq \frac{\epsilon}{4}$. So

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

Therefore $|pr_{\dot\gamma(0)} aHor^g| \geq \sin(\alpha - \frac{\epsilon}{4})$.

On the other hand, by Lemma B.1.2, $|pr_{\dot\gamma(0)} aHor^g - pr_{\dot\gamma(0)} LC^g| \leq 200\epsilon$. The inverse triangle inequality gives

$$|pr_{\dot\gamma(0)} LC^g| \geq \sin(\alpha - \frac{\epsilon}{4}) - 200\epsilon.$$

Applying $\exp_*^{-1}$, by Lemma A.1.2 we have $|\exp_*^{-1}(pr_{\dot\gamma(0)} LC^g)| \geq \frac{1}{1+\frac{\epsilon}{5}} |pr_{\dot\gamma(0)} LC^g|$, and since $\frac{1}{1+\frac{\epsilon}{5}} \geq \frac{99}{100}$ 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.1.3. Together with Lemma C.1.1 and integration along π this will deliver the estimate of Proposition 4.3.2.

The next two lemmas will be used to prove Corollary C.1.1, 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 aim we show that

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

where we identify tangent spaces by parallel translation along γ. The crucial step is the second "≈", where we use that fact that γ is a geodesic. Applying $\exp_*^{-1}$ will easily imply Corollary C.1.1 since $\exp_*^{-1}(\mathrm{pr}_{\dot\gamma(t)} LC^g) = \widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)}$.

Lemma C.1.2 *For any $L < 1$ and any point $p \in \exp_{N_g}(\nu N_g)_L$ the orthogonal projections $T_pM \to aHor_p^g$ and $T_pM \to Hor_p^g$ differ at most by $\frac{L^2}{5}$ in the operator norm.*

Proof: This follows immediately from [We, Prop. 3.7]. □

Lemma C.1.3 *For all t*

$$d(Vert^g_{C,\gamma^b} \backslash\backslash Vert^g_{\gamma(t)}) \leq \arcsin\left[t \left(r + \frac{r + \frac{3}{2}}{f(r)} \right) \right].$$

Furthermore,

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

Proof: We first want to estimate $d(Vert^g_C\ ,\ _{\gamma^b}\|Vert^g_{\gamma(t)})$. Let $v \in \nu_C N_g$ be a normal unit vector.

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

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

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

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

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

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, $L(\pi|_{[0,t]}) \leq \frac{t}{f(r)}$.

Together this gives

$$\begin{aligned}\left|_{\tau_0}\|v\ -\ _{\gamma^b}\|\circ\ _{\tau_t}\|\circ\ ^\perp_\pi\|v\right| \leq & \left|_{\tau_0}\|v\ -\ _{\gamma^b}\|\circ\ _{\tau_t}\|\circ\ _\pi\|v\right| + \left|_{\gamma^b}\|\ \circ\ _{\tau_t}\|\circ\ \left[_\pi\|v -^\perp_\pi\|v\right]\right| \\ \leq & t\left(r + \frac{r+\frac{3}{2}}{f(r)}\right).\end{aligned}$$

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

Corollary C.1.1 *For all t*

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

Proof: From Lemma C.1.2 and Lemma B.1.2 we have for all t

$$\begin{aligned}\left|pr_{\dot\gamma(t)}Hor^g - pr_{\dot\gamma(t)}LC^g\right| &\le \left|pr_{\dot\gamma(t)}Hor^g - pr_{\dot\gamma(t)}aHor^g\right| + \left|pr_{\dot\gamma(t)}aHor^g - pr_{\dot\gamma(t)}LC^g\right| \\ &\le \frac{r^2}{5} + 2r \\ &\le 2.1r.\end{aligned}$$

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

$$\left|pr_{\dot\gamma(0)}Hor^g - pr_{\dot\gamma(0)}LC^g\right| \le 210\epsilon.$$

Combining this with the second statement of Lemma C.1.3 gives

$$\left|pr_{\dot\gamma(0)}LC^g -_{\gamma^b}\!\!\backslash\!\!\backslash pr_{\dot\gamma(t)}LC^g\right| \le 2.1(100\epsilon + r) + t\left(r + \frac{r+\frac{3}{2}}{f(r)}\right).$$

Recall that $pr_{\dot\gamma(t)}LC^g = \exp_*(\widehat{\frac{\nabla^\perp}{dt}\tilde\gamma(t)})$, as in the proof of Lemma B.1.3. Also, for any vector $X \in T_CM$ we have $|\exp_*^{-1}X| \le \frac{|X|}{1-\frac{\epsilon}{5}}$ by Lemma A.1.2. So applying $(\exp^{-1})_*$ to $pr_{\dot\gamma(0)}LC^g -_{\gamma^b}\!\!\backslash\!\!\backslash pr_{\dot\gamma(t)}LC^g$ we get

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

□

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 ${}^{\perp}_{\pi^b}\!\!\backslash\!\!\backslash\xi$ and $\exp_*^{-1}\circ_{\gamma^b}\!\!\backslash\!\!\backslash \circ \exp_*\hat\xi \in T_CM$ 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.1.5). In Corollary C.1.2 we will apply this to the vector $\frac{\nabla^\perp}{dt}\tilde\gamma(t)$.

Lemma C.1.4 *Denoting by τ_t the unit speed geodesic from $\pi(t)$ to $\gamma(t)$,*

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

Proof: First let us notice that applying Lemma A.1.2 three times we get

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

Therefore, applying Lemma A.1.2 to ξ, the left hand side of the statement of this lemma is bounded above by

$$\begin{aligned}&\left|{}_{\tau_0^b}\backslash\!\!\backslash \circ\ {}_{\gamma^b}\backslash\!\!\backslash\left[{}_{\tau_t}\backslash\!\!\backslash\xi\right] - {}_{\tau_0^b}\backslash\!\!\backslash \circ\ {}_{\gamma^b}\backslash\!\!\backslash\left[\exp_* \hat{\xi}\right]\right| + \left|{}_{\tau_0^b}\backslash\!\!\backslash \circ \left[{}_{\gamma^b}\backslash\!\!\backslash \exp_* \hat{\xi}\right] - \exp_*^{-1}\left[{}_{\gamma^b}\backslash\!\!\backslash \exp_* \hat{\xi}\right]\right| \\ &\leq \frac{r^2}{5} + \frac{r^2}{5}\frac{1+\frac{r^2}{5}}{1-\frac{r^2}{5}} \\ &\leq r^2\frac{2}{5(1-\frac{r^2}{5})}.\end{aligned}$$

□

Lemma C.1.5

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

Proof: The left hand side in the statement of the lemma is bounded above by

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

The first term is estimated by Lemma C.1.4.

The second one is just the holonomy as once goes around the loop given by $\tau_t, \gamma^b, \tau_0^b$ and π, which was bounded above in the proof of Lemma C.1.3.

The third and last term is estimated in the proof of Lemma C.1.3 as well. □

Corollary C.1.2 *The section $\tilde{\gamma}$ satisfies*

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

Proof: We apply Lemma C.1.5 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 $|\widehat{\frac{\nabla^\perp}{dt}\tilde{\gamma}(t)}| = |\exp_*^{-1}(pr_{\dot{\gamma}(t)}LC^g)| \leq \frac{1}{1-\frac{r^2}{5}}|pr_{\dot{\gamma}(t)}LC^g|$ by Lemma A.1.2, and

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

by Lemma B.1.2. Since $\frac{2r+1}{1-\frac{r^2}{5}} \leq \frac{4}{3}$ for $r \leq 0.08$ we are done. □

Now Corollary C.1.1 and Corollary C.1.2 immediately imply that $\tilde{\gamma}(t)$ "grows at a nearly constant rate":

Corollary C.1.3 *The section $\tilde{\gamma}$ satisfies*

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

Proof of Proposition 4.3.2: The estimate of Proposition 4.3.2 follows from

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

using Lemma C.1.1 and Corollary C.1.3. □

Bibliography

[BK] Buser, P., Karcher, H.: *Gromov's Almost Flat Manifolds*, Astérisque **81**, Société Mathématique de France, Paris (1981)

[Ca] Cannas da Silva, A.: *Lectures on Symplectic Geometry*, Springer Lecture Notes, Vol. 1764. Berlin, Heidelberg, New York: Springer (2000)

[CL] Camacho, C., and Lins Neto, A., *Geometric Theory of Foliations*, Birkhäuser (1985).

[Ch] Chiang, R.: *Nonstandard Lagrangian submanifolds in* $\mathbb{C}P^n$, Preprint arXiv:math:SG/0303262

[Fo] Folland, G., *Real Analysis: Modern Techniques and their Applications*, second edition, Wiley-Interscience (1999)

[Go] Gotay, M., *On coisotropic imbeddings of presymplectic manifolds*, Proc. AMS, Volume 84, Number 1, 111-114 (1982)

[GK] Grove, K., and Karcher, H., *How to conjugate* C^1*-close group actions*, Math. Z. 132, 11-20 (1973)

[Ha] Hamilton, R., *The inverse function theorem of Nash and Moser*, Bull. AMS, Volume 7, Number 1, 65-222 (1982)

[Jo] Jost, J.: *Riemannian Geometry and Geometric Analysis*, second edition. Berlin, Heidelberg, New York: Springer (1998)

[Kl] Klingenberg, W.: *Riemannian Geometry*, second edition. De Gruyter (1995)

[La] Lang, S., *Fundamentals of Differential Geometry*, Springer Verlag (1999)

[OP] Oh, Y.G. and Park, J.S., *Deformations of coisotropic submanifolds and strongly homotopy Lie algebroids*, ArXiv:math.SG/0305292

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

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