LOCAL COLLAPSING THEORY

F.T. FARRELL AND L.E. JONES

The purpose of this paper is to introduce a variant of the geometric collapsing theory for Riemannian manifolds contained in the joint work of J. Cheeger, K. Fukaya and M. Gromov (henceforth called the “C-F-G-theory”). The authors make important use of this variant geometric collapsing theory in Farrell and Jones, 1998(2), to develop a theory for collapsing foliated Riemannian manifolds and then to prove topological rigidity for certain aspherical manifolds in Farrell and Jones, 1998(1).

0. Introduction.

Cheeger-Fukaya-Gromov have developed in [4] and in their earlier papers [5], [6], [10] and [11] a theory of “collapsing” Riemannian manifolds (henceforth referred to as the C-F-G-theory). The purpose of this paper is to introduce a variant of the C-F-G-theory which is needed by the authors in [9] for developing their theory of “collapsing” foliated Riemannian manifolds, and in [8] for proving topological rigidity for $A$-regular complete non-positively curved Riemannian manifolds of dimension greater than four.

Recall that a Riemannian manifold $(M, g)$ is said to be $A$-regular for some sequence of positive numbers $A = \{A_i\}$ if we have that

$$(0.0) \quad |\nabla^i R| \leq A_i$$

for all $i$, where the indices $i$ vary over all the natural numbers and $\nabla^i R$ is the $i$-th covariant derivative of the curvature tensor (cf. [4, p. 334]). Note that the 0-th condition means that the sectional curvature values are pinched; i.e., bounded away from $\pm \infty$. The C-F-G-theory, as well as its variant developed in this paper, both apply to any complete $A$-regular Riemannian manifold $(M, g)$. In the rest of this paper we shall always assume that $(M, g)$ is a complete $A$-regular Riemannian manifold.

The main results of C-F-G-theory are reviewed in Appendix 2 below (cf. A.2.2 and A.2.3). These results tell us that each sufficiently small piece of a complete $A$-regular Riemannian manifold $(M, g)$ has a neighborhood in $M$ denoted by $V$ which is equal to the orbit space $\Lambda \backslash \hat{V}$ of a free and properly discontinuous group action $\Lambda \times \hat{V} \to \hat{V}$ by isometries; the Riemannian manifold $\hat{V}$ (which is just a covering space of $V$) has injectivity radius at all
points greater than \( \delta \), where \( \delta > 0 \) depends only on the \( \{A_i\} \) and \( \dim M \); and the action \( \Lambda \times \hat{V} \to \hat{V} \) extends to an effective group action \( H \times \hat{V} \to \hat{V} \) by a Lie group \( H \) which is both virtually connected and virtually nilpotent, and which contains \( \Lambda \) as a co-compact discrete subgroup.

In the remainder of this section we will formulate the main results of this paper (cf. Theorems 0.3 and 0.5 below); and give a brief outline of the paper. In 0.1-0.3 below let \( r : U \to B \) denote a smooth mapping from a submanifold \( U \subset M \) (with \( \partial U = \emptyset \)) onto a Riemannian manifold \( (B, g_B) \).

0.1. The curvature \( K(r; M) \). We let \( K(r; M) \) denote the glb (greatest lower bound) of all numbers \( \sigma > 0 \) which satisfy the following properties for any smooth path \( f : [0, 1] \to U 
\)

\( (a) \ \Theta(TU_f(1), P_f(TU_f(0))) < \sigma(\text{length}(f)); \\
(b) \ |Dr_f(1) - P_f(Dr_f(0))| < \sigma(\text{length}(f)). 
\)

Here \( \Theta(V, W) \) denotes the angular distance between planes \( V \) and \( W \) (i.e., the maximum of the angular distance from each vector of \( V \) to \( W \) and from each vector of \( W \) to \( V \)); and \( P_f \) denotes parallel translation along \( f \) in \( (M, g) \) in Part (a). In Part (b) \( P_f(Dr_f(0)) \) is defined to be the composite map
\[
P_{rof} \circ Dr_f(0) \circ \pi \circ P_f : TU_f(1) \to TB_{rof(1)},
\]
where \( P_f \) is a parallel translation in \( (M, g) \) along the path \( \bar{f}(t) = f(1-t), \) and \( \pi : TM_f(0) \to TU_f(0) \) is orthogonal projection, and \( P_{rof} \) denotes parallel translation along \( r \circ f \) in \( (B, g_B) \). Note that \( K(r; M) \) depends on \( U \) and \( (B, g) \) as well as on \( r \) and \( (M, g) \).

0.2. Infranil cores. An infranil core for \( (M, g) \) consists of a smooth submanifold \( U \subset M \) with empty boundary, Riemannian manifold \( (B, g_B) \), and a smooth bundle projection \( r : U \to B \), such that the following properties hold:

0.2.1. (a) The fibers of \( r \) are infranil manifolds; in particular they are closed aspherical manifolds with \( \pi_1 \) an infranil group. (Recall that an “infranil manifold” is a double coset space \( \Gamma \backslash G/K \) where \( G \) is a Lie group which is both virtually connected and virtually nilpotent, \( K \) is a maximal compact subgroup and \( \Gamma \) is a torsion free, co-compact, discrete subgroup of \( G \).)

(b) \( B \) is an open ball centered at the origin of some Euclidean space \( \mathbb{R}^n \); and \( g_B \) is the Euclidean metric.

(c) \( U \subset M \) has a well-defined tubular neighborhood \( E \) of radius equal \( \text{radius}(B) \). That is the exponential map \( \exp : T_{\delta}^\perp(U) \to M \) is a smooth embedding with image equal \( E \), where \( \delta = \text{radius}(B) \) and where \( T_{\delta}^\perp U \) denotes the set of all \( v \in TM_{|U} \) which are perpendicular to \( U \) and which satisfy \( |v| < \delta \).

The radius of an infranil core \( r \) is defined to be the radius of \( B \) if \( B \neq \) point. If \( B = \) point, then the radius of \( r \) is a fixed number \( \delta > 0 \) such that
$U$ has a well-defined tubular neighborhood in $M$ of radius $\delta$ (as described in 0.2.1(c)).

**Remark 0.2.1.1.** In the preceding definition of “infranil core”, by the boundary of $U$ we mean the manifold boundary. So typically $U$ is a non-compact manifold without manifold boundary. The closure of $U$ in $M$, denoted $C(U)$, is typically a compact manifold with boundary; and the map $r: U \to B$ typically extends to a fiber bundle map $C(U) \to C(B)$ where $C(B)$ denotes the closure of $B$ in $\mathbb{R}^n$.

**Example 0.2.1.2.** Let the double coset space $\Gamma\backslash G/K$ denote an infranil manifold. Let $h: \Gamma \to O(k)$ denote a representation for $\Gamma$ into the group of orthogonal transformations of $\mathbb{R}^k$; and let $\Gamma \times (G/K) \times \mathbb{R}^k \to (G/K) \times \mathbb{R}^k$ denote the diagonal action. Set

$$M = (\Gamma\backslash (G/K) \times \mathbb{R}^k) \times \mathbb{R}^l$$

and

$$U = (\Gamma\backslash (G/K) \times 0) \times B^l$$

where $B^l$ denotes the open ball of radius $\delta > 0$ centered at the origin of $\mathbb{R}^l$. Define

$$r: U \to B$$

to be the standard projection $(\Gamma\backslash (G/K) \times 0) \times B^l \to B^l$. Then $r: U \to B$ is an infranil core for $(M, g)$ having radius $\delta$ for any Riemannian metric $g$ on $M$ with respect to which $U$ has a tubular neighborhood of radius $\delta$ in $(M, g)$ (cf. 0.2.1(c)). (In Remark 0.2.2.1 below we construct such a metric $g$.)

Now let $r: U \to B$ denote an arbitrary infranil core for $(M, g)$ of radius $\delta$. We shall say that $r: U \to B$ is $(\epsilon, \theta)$-rigid, for numbers $\epsilon, \theta > 0$, if the following properties hold:

**0.2.2.** (a) $K(r; M) \leq \epsilon(\delta^{-1})$.

(b) $\text{diameter}(r^{-1}(x)) < \epsilon \delta$, for all $x \in B$. (This refers to the diameter of the manifold $r^{-1}(x)$ with respect to its Riemannian metric inherited from $(M, g)$.)

(c) For any $w \in TU$ which is perpendicular to the fibers of $r$ we have that

$$1 - \epsilon|w| \leq |Dr(w)| \leq (1 + \epsilon)|w|.$$ 

(d) For each $v \in TM|_U$ which is perpendicular to $U$ there is a smooth path $f: [0, 1] \to U$, which starts and ends at the foot of $v$, and which satisfies

$$\text{length}(f) < \epsilon \delta \quad \text{and} \quad \theta < \Theta(v, Pf(v)).$$

Here $\Theta(v, Pf(v))$ denotes the angular distance between vectors, and $P_f$ denotes parallel translation in $(M, g)$ along $f$. 
Remark 0.2.2.1. The question arises as to when the example constructed in 0.2.1.2 is \((\epsilon,\theta)\)-rigid? To answer this we first construct a specific Riemannian metric \(g\) on the manifold \(M = (\Gamma \setminus (G/K) \times \mathbb{R}^k) \times \mathbb{R}^l\) of 0.2.1.2. Let \(\langle \ , \ \rangle_{G/K}\) denote any Riemannian metric on \(G/K\) with respect to which \(\Gamma \times (G/K) \to G/K\) is an action by isometries; and let \(\langle \ , \ \rangle_{\mathbb{R}^k}\) and \(\langle \ , \ \rangle_{\mathbb{R}^l}\) denote the Euclidean metrics on \(\mathbb{R}^k\) and \(\mathbb{R}^l\). Define a Riemannian metric \(g\) on \(M\) to be the quotient of the product of \(\langle \ , \ \rangle_{G/K}\) and \(\langle \ , \ \rangle_{\mathbb{R}^k}\) and \(\langle \ , \ \rangle_{\mathbb{R}^l}\).

Note that (with respect to the metric \(g\) just constructed) we have that \(K(r; M) = 0\); so Property 0.2.2(a) holds. Property 0.2.2(b) will hold iff
diameter\((\Gamma \setminus G/K)\) < \(\epsilon\delta\)
with respect to the metric induced on \(\Gamma \setminus G/K\) by \(\langle \ , \ \rangle_{G/K}\). By construction we have that
\[|Dr(w)| = |w|\]
for any vector \(w \in TU\) which is perpendicular to the fibers of \(r\); so Property 0.2.2(c) holds. Finally, we note that Property 0.2.2(d) holds iff for each \(v \in \mathbb{R}^k\) and each \(x \in G/K\) there is \(g \in \Gamma\), and a smooth path \(p : [0,1] \to G/K\) which starts at \(x\) and ends at \(g(x)\), which satisfy
\[\text{length}(p) < \epsilon\delta\] and \(\theta < \Theta(v, h(g)(v))\).
Here \(\text{length}(p)\) is measured with respect to \(\langle \ , \ \rangle_{G/K}\), \(h : \Gamma \to O(k)\) is the representation of \(\Gamma\) given in 0.2.1.2, and \(\Theta(\ , \ )\) is the angular distance in \(\mathbb{R}^k\).

Existence Theorem 0.3. There is an integer \(\eta > 0\) and a number \(\theta \in (0,1)\) which depend only on \(\dim M\). For any given \(\epsilon \in (0,1)\) there is an arbitrarily small decreasing sequence of positive numbers \(\delta_1 > \delta_2 > \delta_3 > \ldots\) having arbitrarily small quotients \(\delta_j/\epsilon\delta_{j-1}\); each \(\delta_j\) depends only on \(j, \epsilon, \dim M, \{\delta_1, \delta_2, \ldots, \delta_{j-1}\}\) and the \(A = \{A_i\}\) of (0.0). There is, for each integer \(n \geq 0\) and for each \(p \in M\), an infranil core \(r : U \to B\) for \((M,g)\) and a point \(p' \in U\) which satisfy the following properties:

(a) The radius of \(r\) is equal \(\delta_c\) for some \(c \in (n, n + \eta)\); \(c\) depends on \(p\) as well as on \(n\).
(b) \(r\) is \((\epsilon, \theta)\)-rigid.
(c) \(d(p,p') < \epsilon \delta_c\) and \(|r(p')| = 0\).

Remark 0.3.1. \(U\) may be thought of as a stratum of lowest dimension provided by the C-F-G-theory for collapsing a small piece of \(M\); \(U\) is “collapsed” along the fibers of \(r : U \to B\). This idea will be made precise when Theorem 0.3 is proven in §3 below. Note that for some \(p \in M\) there might not be any collapsing at all: If the radius of injectivity at \(p\) is greater than \(\delta_1\) then we may choose \(U\) to be the open ball of radius \(\delta_n\) centered at \(p\) in
Let $M$, $B$ to be the open ball of radius $\delta_n$ centered at the origin in $TM_p$ and $r : U \to B$ to be the inverse of the exponential map.

**Remark 0.3.2.** The role of the integer $n$ in Theorem 0.3 is not very apparent; it becomes more apparent when the authors apply Theorem 0.3 in reference [9]. The reader should try to understand Theorem 0.3 and its proof for the special case $n = 0$; the general case $n \geq 0$ is handled in the same way.

**0.4. Thickened infranil cores.** Let $r : U \to B$ denote an $(\epsilon, \theta)$-rigid infranil core of radius $\delta > 0$. Note that Property 0.2.1(c) assures us that $U$ has a well-defined open tubular neighborhood $E$ of radius $\delta$ in $(M, g)$ and that the orthogonal projection $\rho : E \to U$ is a well-defined bundle projection map. (The orthogonal projection $\rho : E \to U$ is just the composite of the usual fiber bundle projection $T\delta^+(U) \to U$ with $\exp^{-1} : E \to T\delta^+(U)$ of 0.2.1(c).) Define $s : E \to U$ to be the composite $r \circ \rho$; and define $t : E \to \mathbb{R}$ by $t(x) = d(x, \rho(x))$ for all $x \in E$. The pair of maps $(s, t)$ represent a thickened infranil core of diameter $\delta$ which is the “thickening” for the infranil core $r$. For each $\delta' \in (0, \delta]$ we let $B(\delta')$ denote the open ball of radius $\delta'$ centered at the origin of $B$, and we set

$$E(\delta') = s^{-1}(B(\delta')) \cap t^{-1}((0, \delta')) \quad \text{and} \quad U(\delta') = r^{-1}(B(\delta')).$$

**Comparison Theorem 0.5.** Given $\theta > 0$ we let $\epsilon, \delta > 0$ denote sufficiently small numbers, where how small is sufficient depends only on $\theta$, $\dim M$ and on the numbers $A = \{A_i\}$ of (0.0) above. Let $r_i : U_i \to B_i$, $i = 1, 2$, denote $(\epsilon, \theta)$-rigid infranil cores both of radius $\delta$. If $E_1(\delta/9) \cap E_2(\delta/9) \neq \emptyset$ then there is an isometry $I : \mathbb{R}^k \to \mathbb{R}^k$ (where $k = \dim B_1$) such that the following properties hold for each $x \in E_1 \cap E_2$:

(a) $\dim U_1 = \dim U_2$ and $B_1 = B_2$.
(b) $|t_1(x) - t_2(x)| < O(\epsilon)\delta$.
(c) $|I \circ s_1(x) - s_2(x)| < O(\epsilon)\delta$.

**Remark 0.5.1.** The notation “$O(\epsilon)$” appearing in the inequalities of 0.5(b) and (c) means that there is a $C^\infty$-function $g : \mathbb{R} \to \mathbb{R}$ with $g(0) = 0$, which is independent of the infranil cores $r_1$, $r_2$ and of the numbers $\epsilon, \delta$, such that when $O(\epsilon)$ is replaced by the number $|g(\epsilon)|$ then the resulting inequality is actually true.

**Remark 0.5.2.** Suppose that all the hypotheses of Theorem 0.5 are still in effect. Then Properties 0.5(a)-(c) are equivalent to the following three properties. For each $x \in U_2 \cap E_1$ let $f_x : [0, 1] \to \rho^{-1}_1(\rho_1(x))$ denote the geodesic with $f_x(0) = x$, $f_x(1) = \rho_1(x)$; and let $\mathcal{G}_i$ denote the foliation of $U_i$ by the fibers of $r_i$.

(a) $\text{length}(f_x) < O(\epsilon)\delta$.
(b) $\Theta(P_{f_x}(TU_2(x)), TU_1(\rho_1(x))) < O(\epsilon)$. 

**LOCAL COLLAPSING THEORY 5**
\( (c) \Theta(P_{f_{x}}(T\mathcal{G}_{2|x}), T\mathcal{G}_{1|\rho_{1}(x)}) < \mathcal{O}(\epsilon). \)

This remark will be proven in Section 1.

This completes the statement of our main results of this paper.

The paper is organized as follows:

§ 1: In this section we focus on the Comparison Theorem proving Theorem 0.5 and Remark 0.5.2.

§ 2: In this section we formulate and prove a number of geometric lemmas which will be used in the proof of Theorem 0.3. These lemmas concern the isometric actions of connected nilpotent Lie groups which have “local angle control” (cf. 2.1). Lemma 2.8 describes the properties of 1-parameter subgroups of Lie groups which act with local control on themselves.

§ 3: In this section we complete the proof of Theorem 0.3. We begin with a local collapsing structure for \( M \) near the point \( p \) (of 0.3) which is provided by C-F-G-theory [4]. We then use the lemmas of §2 to modify this collapsing structure so that we have better geometric control over the modified one; now the desired infranil core \( U \) of 0.3 is taken to be the “thrice repeating layer” of this modified collapsing structure (cf. 3.5.1).

We have also included two appendices at the end of this paper. In Appendix 1 we discuss \( A \)-regularity of a Riemannian manifold \( (M, g) \) in terms of local (normal) coordinate systems for \( M \); these results are used both in §1 and §3 of this paper. It is recommended that this appendix be consulted before reading Section 1. In Appendix 2 we formulate the results from C-F-G-theory (cf. [4]) which are needed to prove Theorem 0.3 in §3 of this paper. In particular we prove Theorems A.2.3 and A.2.5 which describe the group action on the fibration of orthonormal frames. This appendix is best read after reading §1 and §2 and before reading §3.

1. Proof of Theorem 0.5 and Remark 0.5.2.

It is recommended that the reader consult Appendix 1 before reading this section.

Proof of Remark 0.5.2. We leave as an exercise for the reader to check that Properties 0.5(a)-(c) (together with \( (\epsilon, \theta) \)-rigidity for \( r_{1}, r_{2} \), and the inequality \( E_{1}(\delta/9) \cap E_{2}(\delta/9) \neq \phi \) imply Properties 0.5.2(a)-(c).

To see that Properties 0.5.2(a)-(c) (together with \( (\epsilon, \theta) \)-rigidity for \( r_{1}, r_{2} \), and the inequality \( E_{1}(\delta/9) \cap E_{2}(\delta/9) \neq \phi \) imply Properties 0.5(a)-(c) we argue as follows: First note that 0.5(b), and the first equality of 0.5(a), follow immediately from 0.5.2(a) and (b). The second equality of 0.5(a) is equivalent to \( \dim B_{1} = \dim B_{2} \) (since \( r_{1} \) and \( r_{2} \) are assumed to have the same radius); and this last equality follows from 0.5.2(a)-(c) applied at any \( x \in r_{1}^{-1}(0) \).
Towards verifying Property 0.5(c) we first need a candidate for the isometry \( I : \mathbb{R}^k \to \mathbb{R}^k \) of 0.5(c). Choose \( q \in r_1^{-1}(0) \), and let \( V \subset T(U_1)_q \) denote the subspace of all vectors which are perpendicular to \( r_1^{-1}(0) \) at \( q \). Note that it follows from 0.2.2 (as applied to \( r_1 \)) that \( Dr_1 : V \to T(B_1)_0 \) is a linear isomorphism which satisfies

(i) \( (1 - \varepsilon)|v| \leq |Dr_1(v)| \leq (1 + \varepsilon)|v| \)

for all \( v \in V \). It follows from 0.5.2(a)-(c), and from 0.2.2 (as applied to \( r_2 \)), and from the inequality \( E_1(\delta/9) \cap E_2(\delta/9) \neq \phi \), that \( Ds_2 : V \to T(B_2)_{s_2(q)} \) is a linear isomorphism which satisfies

(ii) \( (1 - O(\varepsilon))|v| \leq |Ds_2(v)| \leq (1 + O(\varepsilon))|v| \)

for all \( v \in V \). (In verifying this last inequality the reader should recall that \( \delta > 0 \) is chosen small relative to the \( \{A_i\} \) of (0.0).) Recall that \( B_1 = B_2 \) (cf. 0.5(a)), and that \( B_1 \) is an open ball of radius \( \delta \) centered at the origin of \( \mathbb{R}^k \); in what follows we will identify \( T(B_1) \) with \( \mathbb{R}^k \) via the Euclidean exponential map, and we shall identify \( T(B_2)_{s_2(q)} \) with \( T(B_1)_0 \) via Euclidean parallel translation. Then it follows from (i) and (ii) that we may choose a linear isometry \( L : \mathbb{R}^k \to \mathbb{R}^k \) so that

(iii) \( \|L - Ds_2 \circ (Dr_1 \mid V)^{-1}\| < O(\varepsilon) \).

Now we define an isometry \( I : \mathbb{R}^k \to \mathbb{R}^k \) as follows:

(iv) \( I(0) = s_2(q) \) and \( DI = L \).

Note that it follows from (i)-(iv) (see also 0.5.2(a)-(c)) that:

(v) \( I \circ s_1(q) = s_2(q) \) and \( \|D(I \circ s_1)_q - D(s_2)_q\| < O(\varepsilon) \).

Recall that \( K(\mathcal{X} : M, \delta) \) is defined preceding the statement of Theorem A.1.6 in Appendix 1. By applying Theorem A.1.6 (from the Appendix) to both \( s_1 \) and \( s_2 \) we may deduce that

(vi) \( K(s_i; M, \delta) < O(\varepsilon)\delta^{-1} \) for \( i = 1, 2 \).

Now Property 1.5(c) follows directly from (v) and (vi).

This completes the Proof of Remark 0.5.2.

\begin{proof}[Proof of Theorem 0.5] Let \( f : \mathbb{R}^m \to M \) denote a map which is just the composition of a linear isometry \( \mathbb{R}^m \to TM_{q_1} \) with the exponential map \( \exp : TM \to M \), where \( q_1 \in r_1^{-1}(0) \). We note that for \( \delta > 0 \) sufficiently small the restricted map \( f : B_{10\delta}^m \to M \) is a smooth immersion, where for any \( x > 0 \) we let \( B_x^m \) denote the open ball of radius \( x \) centered at the origin of \( \mathbb{R}^m \). (How small is sufficient for \( \delta \) here depends only on the dim \( M \), \( A = \{A_i\}; \) see Corollary A.1.2 in Appendix 1 for further details.) For each \( i = 1, 2 \) we set

\[ \hat{U}_i = f^{-1}(U_i) \cap B_{10\delta}^m \quad \text{and} \quad \hat{G}_i = f^{-1}(G_i) \cap B_{10\delta}^m. \]

Note that Properties 0.5(a)-(c) follow immediately from 0.2.2(a)-(c) (as applied to each of \( r_1 \) and \( r_2 \)), from 0.5.2, and from the following claim: In the
following claim the manifold $B^m_{108}$ is equipped with the Riemannian structure $f^*(g)$, and we denote by $P_u(\cdot)$ the parallel translation along a path $u : [0, 1] \to B^m_{108}$ with respect to $f^*(g)$, and we denote by $\Theta(\cdot, \cdot)$ the angular distance with respect to $f^*(g)$.

Claim 1.1. For each $x \in \tilde{U}_2 \cap B^m_{28}$ there is a smooth path $u : [0, 1] \to B^m_{108}$, with $u(1) \in \tilde{U}_1$ and $u(0) = x$, so that (a) (b) and (c) hold:

(a) length($u$) $< \mathcal{O}(\epsilon)\delta$.
(b) $\Theta(T\tilde{U}_{1[u(1)]}, P_u(T\tilde{U}_{2[u(0)]})) < \mathcal{O}(\epsilon)$.
(c) $\Theta(T\tilde{G}_{1[u(1)]}, P_u(T\tilde{G}_{2[u(0)]})) < \mathcal{O}(\epsilon)$.

By the hypothesis of 0.5 (that $E_1(\delta/9) \cap E_2(\delta/9) \neq \phi$), and since the Euclidean metric on $B^m_{108}$ is a close approximation to $f^*(g)$ (cf. A.1.2), we may choose $x_i \in \tilde{U}_i \cap B^m_{28/3}$ for $i = 1, 2$. Let $V_i \subset \mathbb{R}^m$ and $H_i \subset \mathbb{R}^m$ denote the vector subspaces obtained by parallel translating (in the Euclidean metric) the tangent planes $T\tilde{U}_{ij}$ and $T\tilde{G}_{ij}$ to the origin of $\mathbb{R}^m$, and let $\Theta_e(\cdot, \cdot)$ denote the Euclidean angular distance between subspaces of $\mathbb{R}^m$. Now we may deduce from results in Appendix 1 (cf. A.1.2 and A.1.4), and from the curvature restriction placed on $r$ by 0.2.2(a), that the following linearized version of 1.1 is actually equivalent to 1.1.

Claim 1.2.

(a) There is $y \in V_1 + x_1$ such that $|x_2 - y| < \mathcal{O}(\epsilon)\delta$.
(b) $\Theta_e(V_1, V_2) < \mathcal{O}(\epsilon)$.
(c) $\Theta_e(H_1, H_2) < \mathcal{O}(\epsilon)$.

Thus to complete the proof of 0.5 it will suffice to verify Claim 1.2. Towards this end we first wish to “linearize” Properties 0.2.2(a)-(d). Note that for each $x, y \in B^m_{28/6}$ with $f(x) = f(y)$ there is a smooth embedding $h : B^m_{28/6} \to B^m_{28/3}$ and an isometry $\tilde{h} : \mathbb{R}^m \to \mathbb{R}^m$ which satisfy the following properties (see Corollary A.1.3 in Appendix 1 for further details):

1.3. (a) $h(x) = y$; and for all $z \in B^m_{28/6}$ we have that $f(z) = f \circ h(z)$.
(b) $h(0) = \tilde{h}(0)$, and for all $z \in B^m_{28/6}$ we have that $|D(h - \tilde{h})|_z < \mathcal{O}(\epsilon)$.
(c) $h(\tilde{U}_i \cap B^m_{28/6}) \subset \tilde{U}_i \cap B^m_{28/3}$.

Now the “linearized” versions of 0.2.2(a), (c) and (d) as applied to $r_1, r_2$ (which follow from 0.2.2(a)-(d) as applied to $r_1, r_2$, and from 1.3, and from A.1.1-A.1.3) can be formulated as follows. For each isometry $\tilde{h}$ from 1.3 let $\tilde{h}_r$ denote its rotational part.

1.4. (a) For each $v \in V_i$ which is perpendicular to $H_i$, and for each $h$ in 1.3, we have that

$$\Theta_e(v, \tilde{h}_r(v)) < \mathcal{O}(\epsilon).$$
(b) For each \( w \in \mathbb{R}^m \) which is perpendicular to \( V_i \) there is an \( h \) from 1.3 such that

\[
\frac{\theta}{4} < \Theta_e(w, \tilde{h}_r(w)).
\]

The linearized version of 0.2.2(b) (which follows from 0.2.2(a)-(d) as applied to \( r_1, r_2, \) and from 1.3, and from A.1.1-A.1.3) can be formulated as follows. Set \( Q_i = H_i + x_i \) and let \( d_e(\, \cdot \, , \cdot) \) denote the Euclidean distance in the following:

1.5. (a) For each \( y \in Q_i \cap B^m(\delta) \) and each \( h \) of 1.3 we have that

\[
d_e(h(y), Q_i) < O(\epsilon)\delta.
\]

(b) For each \( y \in Q_i \cap B^m(\delta) \) there is a map \( h_{y,i} \) from 1.3 such that

\[
d_e(y, h_{y,i}(x_i)) < O(\epsilon)\delta.
\]

Now we can use 1.4 and 1.5 to verify 1.2.

We begin by verifying 1.2(c). Let \( u \in H_1 \) be a unit vector and set \( g(t) = x_1 + tu \) for \( t \in \mathbb{R} \). It will suffice to verify the following inequality for some fixed \( b \in (0, \delta) \) and all \( t \in [-\delta/9, \delta/9] \):

1.6. \( b - O(\epsilon)\delta < d_e(g(t), Q_2) < b + O(\epsilon)\delta. \)

For any \( t \in [-\delta/9, \delta/9] \) it follows from the fact \( x_1 \in B^m_{\delta/7} \) that \( g(t) \in B^m_{\delta/3} \). Thus we may apply 1.5(b) (with \( y = g(t) \) and \( i = 1 \)) to get a map \( h_{y,1} \) satisfying

\[
d_e(g(t), h_{y,1}(x_1)) < O(\epsilon)\delta.
\]

Let \( x \in Q_2 \) denote the orthogonal projection of \( x_1 \) onto \( Q_2 \). By applying 1.5(a) to \( x \in Q_2 \) with \( h = h_{y,1} \) we get that

\[
d_e(h_{y,1}(x), Q_2) < O(\epsilon)\delta;
\]

and it follows from this last inequality, from the fact that \( x - x_1 \) is perpendicular to \( Q_2 \), and from the fact that \( h_{y,1} \) is an isometry satisfying 1.5(a) for \( i = 2 \), that we have

\[
|x - x_1| - O(\epsilon)\delta < d_e(h_{y,1}(x_1), Q_2) < |x - x_1| + O(\epsilon)\delta.
\]

Now 1.6 follows from the first and third of the last three inequalities for \( b = |x - x_1| \).

Now we can verify 1.2(b). Let \( H_i^\perp \) denote the orthogonal complement for \( H_i \) in \( V_i \), and let \( V_i^\perp \) denote the orthogonal complement of \( V_i \) in \( \mathbb{R}^m \). For each unit length vector \( u \in H_i^\perp \) choose \( v \in V_2 \) and \( w \in V_2^\perp \) such that \( u = v + w \); and choose \( v_1 \in H_2 \) and \( v_2 \in H_2^\perp \) such that \( v = v_1 + v_2 \). Since 1.2(c) has already been verified, to complete the verification of 1.2(b) it will
suffice to show that $|w| < \mathcal{O}(\epsilon)$. Use 1.4(b) (as applied to $V_2$) to choose $\tilde{h}_r$ such that

$$\Theta_e(w, \tilde{h}_r(w)) > \theta/4.$$ 

By applying 1.4(a) to $u \in H_1^+$ we get that

$$\Theta_e(u, \tilde{h}_r(u)) < \mathcal{O}(\epsilon).$$

We deduce from 1.2(c), and from 1.4(a) as applied to $v_2 \in H_2^+$, that

$$|v_1| < \mathcal{O}(\epsilon) \quad \text{and} \quad \Theta_e(v_2, \tilde{h}_r(v_2)) < \mathcal{O}(\epsilon).$$

Finally, by using these last few inequalities, and the triangle inequality, and the fact that

$$|x| \Theta_e(x, y)/4 \leq |x - y| \leq 4|x| \Theta_e(x, y)$$

holds for any two vectors $x, y \in \mathbb{R}^m$ of the same length, we get that

$$\mathcal{O}(\epsilon) > |u - \tilde{h}_r(u)| \geq |w - \tilde{h}_r(w)| - |v - \tilde{h}_r(v)| > \theta|w|/16 - \mathcal{O}(\epsilon).$$

From which we deduce that $|w| < \mathcal{O}(\epsilon)$ as desired.

Before verifying 1.2(a) we first remark that 1.4(a) and 1.5(a) together imply that:

1.7. $\Theta_e(V_i, \tilde{h}_r(V_i)) < \mathcal{O}(\epsilon)$

for all $\tilde{h}$ from 1.3 and for $i = 1, 2$. Now to verify 1.2(a) we set $x_2 - x_1 = v + w$ where $v \in V_1$ and $w \in V_1^\perp$. It will suffice (by 1.2(b)) to show that $|w| < \mathcal{O}(\epsilon)\delta$. Use 1.4(b) (as applied to $V_1$) to choose $h$ as in 1.3 so that

$$\Theta_e(w, \tilde{h}_r(w)) > \theta/4;$$

and use 1.5(a) (as applied to $x_1 + v \in Q_1$) to get that

$$d_e(\tilde{h}(x_1 + v), Q_1) < \mathcal{O}(\epsilon)\delta.$$ 

Note that 1.7, and the fact that $\tilde{h}_r$ is an isometry, together imply that

$$\Theta_e(\tilde{h}_r(w), V_1^\perp) < \mathcal{O}(\epsilon).$$

Now by combining these last three inequalities with 1.2(b), and with the fact that $\tilde{h}(x_2) = \tilde{h}(x_1 + v) + \tilde{h}_r(w)$, we get that

$$d_e(\tilde{h}(x_2), Q_2) > \theta|w|/20 - \mathcal{O}(\epsilon)\delta.$$ 

On the other hand by applying 1.5(a) to $x_2 \in Q_2$ we get that

$$d_e(\tilde{h}(x_2), Q_2) < \mathcal{O}(\epsilon)\delta.$$ 

These last two inequalities imply that $|w| < \mathcal{O}(\epsilon)\delta$ as desired.

This completes the Proof of Theorem 0.5. \hfill \Box
2. Preliminaries to the proof of Theorem 0.3.

In this section we formulate several lemmas which will be needed for the proof of Theorem 0.3 given in the next section. These lemmas are concerned with isometric actions of a connected Lie group on a Riemannian manifold which have “local angle control” (cf. 2.1).

Let $N, g_N$ denote a connected Riemannian manifold (not necessarily complete), and we let $G \times N \to N$ denote an isometric effective action of the connected nilpotent Lie group $G$. For any numbers $\alpha, \beta > 0$ we shall say that this action has local angle control equal $(\alpha, \beta)$ at $x \in N$ if any path $f : [0, 1] \to N$ which satisfies 2.1(a) and (b) also satisfies 2.1(c).

2.1. (a) There is a path $\phi : [0, 1] \to G$ starting at the identity such that $f(t) = \phi(t)(x)$.

(b) $\text{length}(f) \leq \beta$.

(c) $\Theta(Dg(v), P_f|_{[0,1]}(v)) \leq \alpha$ for all $v \in TN_x$, where $g = \phi(1)$ and where $P_f$ denotes parallel translation along the path $f$.

We shall say that the action $G \times N \to N$ has local angle control equal $(\alpha, \beta)$ if it has such angle control at each of its points $x \in N$.

Remark 2.2. Let $H \times \hat{V} \to \hat{V}$ denote a smooth action of the Lie group $H$ by isometries of an $A$-regular Riemannian manifold $\hat{V}$. Let $\rho : E \to \hat{V}$ denote the bundle of orthonormal frames over $\hat{V}$ equipped with the canonical metric and let $H \times E \to E$ denote the canonical lifting of the action $H \times \hat{V} \to \hat{V}$ (see Appendix 2 for more details). We shall prove in Appendix 2 (cf. Theorem A.2.5) that the action $H \times E \to E$ has local angle control equal $(\lambda t, t)$ for any $t > 0$, where $\lambda > 1$ depends only on $\dim E, A_0$.

Now we will formulate six lemmas which are the main results of this section. The proofs for these lemmas are also given in this section. The first two of these lemmas are just refinements of 2.1. Recall that in Definition 2.1 $\phi$ is not necessarily a one-parameter subgroup of $G$.

Lemma 2.3. Let $\alpha \in (0, \pi)$ and $\beta > 0$. Suppose that $G \times N \to N$ does not have local angle control equal $(\alpha, \beta)$ at $x \in N$. Then there is a path $f : \mathbb{R} \to N$ and a unit vector $v \in TN_x$ which satisfy the following properties:

(a) There is a one-parameter subgroup $\phi : \mathbb{R} \to G$ such that $f(t) = \phi(t)(x)$ for all $t \in \mathbb{R}$.

(b) $\text{length}(f|_{[0,1]}) \leq \beta$.

(c) $\Theta(Dg(v), P_f|_{[0,1]}(v)) > \alpha$ where $g = \phi(1)$.

Lemma 2.4. Let $\alpha \in (0, \pi)$ and $\beta > 0$. Suppose $G \times N \to N$ has local angle control equal $(\alpha, \beta)$, and let $f : [0, 1] \to N$ and $\phi : [0, 1] \to G$ be any smooth paths which satisfy Properties 2.1(a) and (b). Then for all $v \in TN_{f(0)}$, and
for $g = \phi(1)$, we have that
\[ \Theta(Dg(v), P_f(v)) < (\alpha/\beta) \text{length}(f). \]

Recall that the isotropy subgroup $G_x \subset G$ for the action $G \times N \to N$ at $x \in N$ is defined by $G_x = \{ g \in G : g(x) = x \}$.

**Lemma 2.5.** Let $\alpha \in (0, \pi)$ and $\beta > 0$, and suppose that $G \times N \to N$ has local angle control equal $(\alpha, \beta)$. Then for each $x \in N$ the isotropy group $G_x$ in $G$ is a discrete subgroup of $G$.

Note that 2.5 implies that the orbits of $G \times N \to N$ foliate $N$. In the following remarks and lemma we denote this foliation of $N$ by $\mathcal{G}$.

**Remark 2.6(i).** We shall say that $\mathcal{G}$ is a strongly Riemannian foliation if for any path $f : [0, 1] \to L$ in a leaf $L \in \mathcal{G}$ there is a $\delta > 0$ and a continuous map $F : [0, 1] \times \mathbb{R}^k \to N$, where $k = \dim N - \dim \mathcal{G}$, which satisfies:

(a) $F(t, 0) = f(t)$ for all $t$; and $F \mid [0, 1] \times v$ is path in a leaf $L_v \in \mathcal{G}$ for each $v \in \mathbb{R}^k$.

(b) Let $V_i$ denote the open $\delta$-ball centered at the origin of $TG_{f(i)}^\perp$, for $i = 0, 1$, where $TG^\perp$ denotes the orthogonal complement in $TN$ for $TG$. Then the exponential map $\exp : TN \to N$ maps $V_i$ homeomorphically onto a submanifold $W_i \subset N$; and $F \mid i \times \mathbb{R}^k$ maps $i \times \mathbb{R}^k$ homeomorphically onto $W_i$.

(c) Let $h : 0 \times \mathbb{R}^k \to 1 \times \mathbb{R}^k$ be defined by $h(0, v) = (1, v)$. Then the composite map $(F \mid 1 \times \mathbb{R}^k) \circ h \circ (F \mid 0 \times \mathbb{R}^k)^{-1} : W_0 \to W_1$ is an isometry.

**Remark 2.6(ii).** The curvature $K(\mathcal{G}; N)$ for $\mathcal{G}$ in $(N, g_N)$ is defined to be the glb of all numbers $\sigma > 0$ which satisfy the following inequality for any smooth path $f : [0, 1] \to N$:
\[ \Theta(TG_{f(i)}^\perp, P_f(TG_{f(0)}^\perp)) < \sigma(\text{length}(f)). \]

**Lemma 2.6.** Let $\alpha \in (0, \pi)$ and $\beta > 0$, and suppose that $G \times N \to N$ has local angle control equal $(\alpha, \beta)$. Then $\mathcal{G}$ is a strongly Riemannian foliation of $(N, g_N)$; and the curvature $K(\mathcal{G}; N)$ for $\mathcal{G}$ in $(N, g_N)$ satisfies $K(\mathcal{G}; N) < 10^6 \alpha/\beta$.

In our next lemma we assume that the Riemannian metric $g_N$ on $N$ is $A$-regular for some collection $A = \{ A_i : i = 1, 2, 3, \ldots \}$ of positive numbers. We let $\Gamma \subset G$ denote a torsion free discrete subgroup such that the restricted action $\Gamma \times N \to N$ is fixed point free and properly discontinuous. For each $t > 0$ and $x \in N$ we let $\Gamma_{t,x} \subset \Gamma$ denote the subset of all $g \in \Gamma$ such that $d_N(x, g(x)) < t$, and we let $\Gamma_{t,x}$ denote the subgroup of $\Gamma$ generated by $\Gamma_{t,x}$.

**Lemma 2.7.** Let $B > 0$ be as in A.2.3; there is a number $\tau > 1$ which depends only on $B$ and on $\dim N$. Suppose that $G \times N \to N$ has local angle
control equal \((\alpha, \beta)\), where \(\alpha\) and \(\beta\) are sufficiently small and \(\beta < \alpha\) (how small is sufficient depends only on \(A = \{A_i\}\) and on \(\dim N\)). For a given \(x \in N\) we assume that \((N, g_N)\) has radius of injectivity greater than \(10^4 \tau B \beta\) at \(x\); we also assume that for each \(t \in (0, \beta)\) and each \(g \in \Gamma_{t,x}\) there is a one-parameter subgroup \(h_g : \mathbb{R} \to G\) with \(g \in h_g([0, 1])\) and \(d_{N}(x, h_g(s)(x)) < Bt\) for all \(s \in [0, 1]\). (And \(h_e\) equal the trivial subgroup.) Then for any \(t \in (0, \beta)\) there is a closed subgroup \(G_{t,x} \subset G\) such that all of the following properties hold:

(a) If \(h : \mathbb{R} \to G\) is any other one-parameter subgroup of \(G\) which satisfies \(g \in h([0, 1])\) and \(d_{N}(x, h(s)(x)) < Bt\) for all \(s \in [0, 1]\), then there is a number \(a \in \mathbb{R}\) such that \(h(as) = h(s)\) for all \(s \in \mathbb{R}\).
(b) \(G_{t,x}\) is the smallest connected Lie subgroup of \(G\) which contains each one-parameter subgroup \(h_g : \mathbb{R} \to G\), \(g \in \Gamma_{t,x}\).
(c) Set \(L_{t,x} = \{g(x) : g \in G_{t,x}\}\). Then the quotient space \(L_{t,x}/\Gamma_{t,x}\) is compact and has diameter less than \(\tau t\). (The diameter refers to the metric induced on \(L_{t,x}\) by the restricted Riemannian structure \(g_N | T(L_{t,x})\).)
(d) Suppose that \(G_{t,x} = G_{2\tau t, x}\) for \(t \in (0, \beta / 2\tau)\); and suppose that for \(g \in G_{t,x}\) we have that \(d_{N}(x, g(x)) < t\). Then there is \(g' \in G_{t,x}\) with \(gx = g'x\), and there is a one-parameter subgroup \(h_{g'} : \mathbb{R} \to G_{t,x}\) with \(g' \in h_{g'}([0, 2\tau])\) such that the path \(f_{g'} : \mathbb{R} \to N\) defined by \(f_{g'}(s) = h_{g'}(s)(x)\) has unit speed.

In our final lemma we let \(g_G\) denote a left invariant Riemannian metric on \(G\) with respect to which the (left multiplication) group action \(G \times G \to G\) has local angle control equal \((\alpha, \beta)\).

**Lemma 2.8.** Given any sufficiently small \(\alpha > 0\) (how small is sufficient depends only on \(\dim G\)) and given any \(\beta > 0\), suppose that \(G \times G \to G\) has local angle control equal to \((\alpha, \beta)\). If \(\phi_1 : \mathbb{R} \to G\) denotes a one-parameter subgroup which satisfies (a) then there is another one-parameter subgroup \(\phi_2 : \mathbb{R} \to G\) which satisfies (b)-(c).

(a) There are numbers \(s, t > 0\), with \(t < \beta\), such that \(d_G(\phi_1(0), \phi_1(s)) < t\).
(b) We have \(\phi_2(s) = \phi_2(0)\).
(c) \(d_G(\phi_1(u), \phi_2(u)) < (u/s)4^k t\) for all \(u \in [0, s]\), where \(k = \dim G\).

**Proof of Lemma 2.3.** Set \(G_x = \{g \in G : g(x) = x\}\); note that \(G_x\) is a closed subgroup of \(G\) and is thus a Lie group.

If \(\dim G_x > 0\) then choose any one-parameter subgroup \(\psi : \mathbb{R} \to G_x\). For some \(r > 0\) we have that \(\psi(r)\) rotates some unit vector \(v \in TN_x\) by an angle greater than \(\alpha\). We define \(\phi : \mathbb{R} \to G\) by \(\phi(t) = \psi(tr)\). Note that \(\phi\) satisfies Properties 2.3(a) and (b).

If \(\dim G_x = 0\), then the map \(h : G \to N\) defined by \(h(g) = g(x)\) is an immersion. Hence it induces a left invariant Riemannian metric \(g_G\) on
For each \( v \in TG_x \) with \(|v| = 1\), let \( \gamma_v : \mathbb{R} \to G \) be the one-parameter subgroup such that \( \gamma_v(0) = v \). Define a curve

\[
A_v : \mathbb{R} \to \text{Iso}(TN_x)
\]

by setting \( A_v(t) \) equal to the composition of parallel translation along \( \gamma_v \) from \( TN_x \) to \( TN_{\gamma_v(t)x} \) and \( D(\gamma_v(-t)) : TN_{\gamma_v(t)x} \to TN_x \). It is easily seen that \( A_v(t) \) is a one-parameter subgroup of \( \text{Iso}(TN_x) \cong O(n) \) where \( n = \dim N \).

Hence, there is an orthonormal basis for \( TN_x \) so that the matrix representing each \( A_v(t), t \in \mathbb{R} \), is the blocked sum of the identity matrix of a certain size and \( 2 \times 2 \) matrices of the form

\[
\begin{bmatrix}
\cos(\theta t) & \sin(-\theta t) \\
\sin(\theta t) & \cos(\theta t)
\end{bmatrix}
\]

Assuming that the conclusion of 2.3 is false, we see that \( \theta \leq \alpha/\beta \) for each of these matrices (provided \( \alpha \in (0, \pi) \)). We consequently get the following estimate valid for each \( g \in G \), for each one-parameter subgroup \( \gamma_v \) where \(|v| = 1\), for each \( u \in TN_{gx} \), and for each \( t \in [0, \infty) \):

\[
2.3.1. \quad \Theta(D(g\gamma_v(t))g^{-1})(u), P(u)) \leq \alpha t/\beta,
\]

where \( P(u) \) denotes the parallel transport of \( u \) along \( g\gamma_vx \) to \( TN_{g\gamma_v(t)x} \). Since each smooth path \( \phi : [0, 1] \to G \) can be approximated (as closely as we want) by a piecewise smooth path \( \psi : [0, 1] \to G \) where each piece of \( \psi \) has the form

\[
t \mapsto g\gamma_v(t - c)
\]

restricted to an interval of the form \([c, a]\), estimate 2.3.1 yields the following more general estimate where \( f(t) = \phi(t)x \):

\[
2.3.2. \quad \Theta(Dg(v), P_f(v)) \leq (\alpha/\beta)\text{length}(f).
\]

In 2.3.2 we have that \( v \in TN_x \) with \( v \neq 0 \) and \( g = \phi(1) \). Now clearly 2.3.2 implies that \( G \times N \to N \) has local angle control equal \((\alpha, \beta)\) at \( x \), in direct contradiction of the assumption of 2.3.

This completes the Proof of Lemma 2.3. \( \square \)

**Proof of Lemma 2.4.** Note that the conclusion of this lemma has been verified in 2.3.2 of the preceding lemma under the hypothesis that the conclusion of 2.3 is false. Certainly the conclusion of 2.3 is false if the hypotheses of 2.4 hold.

This completes the Proof of Lemma 2.4. \( \square \)

**Proof of Lemma 2.5.** Note that each isotropy subgroup \( G_x \subset G \), \( x \in N \), is a closed Lie subgroup of \( G \). If \( G_x \) is not a discrete subgroup of \( G \) for some \( x \in N \) then the identity component \( G_{x,e} \) of \( G_x \) has dimension \( \geq 1 \).
Choose a nontrivial one-parameter subgroup $\phi : \mathbb{R} \to G_{x,e}$, and note that the following hold:

**2.5.1.** (a) $\phi(t)(x) = x$ for all $t \in \mathbb{R}$.

(b) For some $s > 0$ and some $v \in TN_x$ we have that

$$\Theta(D(\phi(s))(v),v) = \pi.$$

Now define $\phi : [0,1] \to G$ by $\phi(t) = \phi(ts)$, where $s$ comes from 2.5.1(b).

It follows from 2.5.1, and from the definition for $\phi$ just given, that $\phi$ does satisfy Properties 2.1(a) and (b) but does not satisfy Property 2.1(c) for any $\alpha < \pi$ in 2.1(c). This contradicts the hypothesis for 2.5.

This completes the proof for Lemma 2.5. $\square$

*Proof of Lemma 2.6.* First we will verify that $G$ is a strongly Riemannian foliation. Let $f : [0,1] \to L$ be a path in a leaf $L \in \mathcal{G}$. We may choose a path $\phi : [0,1] \to G$ which is related to $f$ as follows (cf. 2.5):

**2.6.1.** (a) $\phi(0)$ is the identity of the group $G$.

(b) $f(t) = \phi(t)(f(0))$ for all $t \in [0,1]$.

Let $V_i$ denote the open $\delta$-ball centered at the origin of $T^G_{f(i)}, i = 0,1$, where $T^G$ denotes the orthogonal complement in $TN$ to $T\mathcal{G}$, and where $\delta > 0$ is chosen sufficiently small to assure that the exponential map $\exp : TN \to N$ is well-defined on $V_i$ and maps $V_i$ diffeomorphically onto a smooth submanifold $W_i \subset N$. Note it follows from 2.6.1(a) and (b), and from the fact that $G$ acts by isometries on $N$, that:

**2.6.1.** (c) $\phi(1)(W_0) = W_1$.

Choose a homeomorphism $F : 0 \times \mathbb{R}^k \to W_0$ which sends $(0,0) \in 0 \times \mathbb{R}^k$ to $\exp(0) \in W_0$, where $k = \dim W_0$. Extend $F | 0 \times \mathbb{R}^k$ to a map $F : [0,1] \times \mathbb{R}^k \to N$ by setting:

**2.6.1.** (d) $F(t,v) = \phi(t)(F(0,v))$

for each $t \in [0,1]$. Note that Properties (a), (b) and (c) in Remark 2.6(i) are an immediate consequence of 2.6.1(a)-(d).

Now we shall verify the curvature condition for $G$ stated in 2.6. Because $G$ is a strongly Riemannian foliation, it will suffice (cf. Remark 2.6.4 at the end of this proof) to show that the following property holds:

**2.6.2.** Let $f : [0,1] \to L$ denote a smooth path with length $(f) < \beta$ in a leaf $L \in \mathcal{G}$; and let $P_f : TN_{f(0)} \to TN_{f(1)}$ denote parallel translation along
Then we have that

$$\Theta(TG_{f(1)}, P_f(TG_{f(0)})) < (\alpha/\beta)\text{length}(f).$$

Using Lemma 2.5 we may choose \( \phi : [0, 1] \to G \) starting at the identity such that \( \phi(t)(x) = f(t) \) for all \( t \in [0, 1] \) and set \( x = f(0) \). Now since \( \phi, f \) satisfy Properties 2.1(a) and (b), and since \( G \times N \to N \) has local angle control equal \((\alpha, \beta)\), it follows from Lemma 2.4 that \( f : [0, 1] \to N \) satisfies:

### 2.6.3. \( \Theta(Dg(v), P_f(v)) < (\alpha/\beta)\text{length}(f) \)

for any \( v \in TN_x \) where \( g = \phi(1) \). Since \( Dg(TG_{f(0)}) = TG_{f(1)} \), the inequality of 2.6.2 is a direct consequence of the inequality in 2.6.3.

This completes the proof for Lemma 2.6 (modulo the contents of Remark 2.6.4 below).

**Remark 2.6.4.** We will show in this remark that if the inequality of 2.6.2 holds for all smooth paths \( f : [0, 1] \to N \) which are tangent to \( G \), then the inequality:

(i) \( \Theta(TG_{f(1)}, P_f(TG_{f(0)})) < 10^6(\alpha/\beta)\text{length}(f) \)

holds for all smooth paths in \( N \). To accomplish this it will suffice to verify that for any compact subset \( C \subset N \), and for any sufficiently small number \( \epsilon > 0 \) (how small is sufficient depends on \( C \)), a smooth path \( f : [0, 1] \to N \) satisfies (i) provided it satisfies:

(ii) \( f(0) = p \) for some \( p \in C \), and \( \text{length}(f) = \epsilon \).

Let \( x_1, x_2, \ldots, x_n \) denote the normal coordinates near \( p \) with \( p \sim (0, 0, \ldots, 0) \), and let \( g_{i,j}dx_idx_j \) denote the coordinate expression for the Riemannian metric on \( N \). Then \( g_{i,j}(0, \ldots, 0) = \delta_{ij}^k \partial g_{i,j}/\partial x_k(0, \ldots, 0) = 0 \), and \( \Gamma_{i,j}^k(0, \ldots, 0) = 0 \) for all \( i, j, k \). Thus, for sufficiently small \( \epsilon > 0 \) there is a number \( \tau > 0 \), independent of \( \epsilon \) and of \( p \in C \), such that:

(iii) \( |g_{i,j}(X) - \delta_{ij}^k| < \tau\epsilon^2 \) for all \( X \in B^n_{10\epsilon} \),

(iv) \( |\Gamma_{i,j}^k(X)| < \tau \epsilon \) for all \( X \in B^n_{10\epsilon} \),

where for any number \( t > 0 \) \( B^n_t \) denotes the set of all points \( X = (x_1, \ldots, x_n) \) in \( \mathbb{R}^n \) with Euclidean length less than \( t \). Let \( d(\ ,\ ) \) and \( \tilde{d}(\ ,\ ) \) denote the metrics on \( B^n_{10\epsilon} \) induced by \( g_{i,j}dx_idx_j \) and \( \delta_{ij}^kdx_idx_j \) respectively; and for any smooth path \( h : [0, 1] \to B^n_{10\epsilon} \) we let \( P_h(\ ) \) and \( \tilde{P}_h(\ ) \) denote parallel translation along \( f \) induced by \( g_{i,j}dx_idx_j \) and \( \delta_{ij}^kdx_idx_j \) respectively. We can deduce the following properties from (iii) and (iv), provided length\((h) < 10\epsilon\):

(v) \( |d(X,Y) - \tilde{d}(X,Y)| < 100n^2\tau\epsilon^3 \) for all \( X, Y \in B^n_{10\epsilon} \),

(vi) \( |P_h(v) - \tilde{P}_h(v)| < 1000n^3\tau\epsilon^2 \) for all unit vectors \( v \in T(B^n_{10\epsilon}) \).

Note that the path \( f : [0, 1] \to N \) of (ii) maps into \( B^n_{10\epsilon} \) (cf. (iii), and recall that \( \text{length}(f) = \epsilon \)); so we may apply (vi) to \( f \) to help us verify that \( f \)}
satisfies Property (i). Let \( L_0 \) and \( L_1 \) denote the leaves of \( G \) which contain \( f(0) \) and \( f(1) \) respectively; and let \( V_0 \) and \( V_1 \) denote the planes in \( \mathbb{R}^n \) which are tangent to \( L_0 \) and \( L_1 \) at \( f(0) \) and \( f(1) \) respectively. One deduces from 2.6.2 (as applied to \( L_0 \) near \( f(0) \), and to \( L_1 \) near \( f(1) \)), and from Properties (iii) and (vi) in the preceding paragraph, that the following holds for sufficiently small \( \epsilon \) > 0:

(vii) \( \tilde{d}(X, V_i \cap B_{10\epsilon}^n) < (\alpha/\beta)10^3\epsilon^2 \) and \( \tilde{d}(Y, L_i \cap B_{10\epsilon}^n) < (\alpha/\beta)10^3\epsilon^2 \)

for \( i = 0, 1 \) and for any \( X \in L_i \cap B_{10\epsilon}^n \) and any \( Y \in V_i \cap B_{10\epsilon}^n \). On the other hand, we may deduce from 2.6(i), and from the fact that \( \text{length}(f) = \epsilon \), that there is a positive number \( t \in (0, 2\epsilon) \) such that:

(viii) \( d(X, L_0 \cap B_{10\epsilon}^n) = t \)

for all \( X \in L_1 \cap B_{9\epsilon}^n \). Now we deduce from Properties (vii) and (viii) above, and from Property (v) in the preceding paragraph, that:

(ix) \( t - (\alpha/\beta)10^4\epsilon^2 < \tilde{d}(Y, V_0 \cap B_{10\epsilon}^n) < t + (\alpha/\beta)10^4\epsilon^2 \)

for all \( Y \in V_1 \cap B_{8\epsilon}^n \). It follows immediately from Property (ix) above, that:

(x) \( \Theta(V_i, \tilde{P}_f(V_0)) < (\alpha/\beta)10^5\epsilon \),

where \( \Theta(\ , \ ) \) denotes the Euclidean angular distance. Finally, since \( V_i = T\mathcal{G}_f(i) \) for \( i = 0, 1 \), and \( \text{length}(f) = \epsilon \), we may deduce Property (i) above from Properties (iii) and (x) above.

Proof of Lemma 2.7. Let \( L \in \mathcal{G} \) denote the leaf which contains \( x \). Towards verifying 6.2.7(a) we will first show that there is an open subset \( U \) of \( L \) which satisfies the following properties. For any \( t > 0 \) let \( B(x; L; t) \) denote the open ball in \( L \) of radius \( t \) centered at \( x \) (where the distance \( d_L(\ , \ ) \) on \( L \) is gotten by restricting \( g_N \) to \( TL \)).

2.7.0. (a) \( U \) is homeomorphic to \( \mathbb{R}^k \) (\( k = \dim L \)).

(b) \( B(x; L; 3B\beta) \subset U \subset B(x; L; 5B\beta) \).

Moreover for any \( y, z \in U \) we have that the distances \( d_L(y, z) \) and \( d_N(y, z) \) are related by:

2.7.0. (c) \( 2d_L(y, z)/3 < d_N(y, z) \leq d_L(y, z) \).

To verify 2.7.0 we refer to results in the appendix to §1 above, and use the curvature bound for \( G \) given in 2.6, and also use the hypothesis of 2.7 that \( (N, g_N) \) has radius of injectivity greater than \( 10^4\tau B\beta \) at \( x \). In more detail we let \( f_x : B_{10B\beta}^n \rightarrow N \) denote the map given in A.1.1 and A.1.2 (where \( \epsilon, p, M, B_\epsilon^n \) are replaced in A.1.1, A.1.2 by \( 10\beta, x, N, B_{10B\beta}^n \) with \( n = \dim N \)). Set \( \hat{G} = f_x^{-1}(\mathcal{G}) \); and let \( K(\hat{G}; B_{10B\beta}^n) \) denote the curvature for \( \hat{G} \) in \( B_{10B\beta}^n \) with respect to the Euclidean metric (cf. 2.6(ii)). The proof of Theorem A.1.4 also works to show the following (cf. A.1.4.1):
2.7.0.1. $K(\hat{G}; B^n_{10B\beta}) - O(10B\beta) < K(G; N) < K(\hat{G}; B^n_{10B\beta}) + O(10B\beta)$.

Note that 2.6 and 2.7.0.1 together imply that:

2.7.0.2. $K(\hat{G}; B^n_{10B\beta}) < 10^6\alpha/\beta + O(10B\beta)$.

We let $\hat{L}$ denote the leaf of $\hat{G}$ through the origin; there is no loss of generality in assuming that the subspace $R^k \subset R^n$ spanned by the first $k$ standard directions in $R^n$ is tangent to $\hat{L}$ at the origin. Now the curvature bounds placed on $\hat{L}$ by 2.7.0.2 assures that $\hat{L}$ is the graph of a smooth map $h : V \to R^{n-k}$ from an open subset $V \subset R^k$ which satisfies the following properties:

2.7.0.3. (a) $B^k_{4B\beta} \subset V$, where $B^k_t$ denotes the open ball of radius $t > 0$ centered at the origin of $R^k$.
(b) $h(0) = 0$ and $Dh_0 = [0]$.
(c) $\|D^2h_p\| < O(\alpha/\beta + \beta)$ for all $p \in V$.

Now we define the subset $U \subset L$ of 2.7.0 by:

2.7.0.4. $U = f_x(\text{graph}(h | B^k_{4B\beta}))$.

Since $N$ has injectivity radius greater than $10^4\tau B\beta$ at $x$ for $\tau > 1$, it follows that $f_x$ is a smooth embedding; thus $U$ is homeomorphic to $B^k_{4B\beta}$, which verifies Property 2.7.0(a). Note that Properties 2.7.0(b) and (c) can be deduced from 2.7.0.3, 2.7.0.4, and from the relation between the Euclidean metric on $B^n_{10B\beta}$ and the pulled back metric $g_{i,j}$ (pulled back from $g_N$ by $f_x$) given in A.1.1, provided $\alpha$ and $\beta$ are sufficiently small.

Now we will complete the proof for 2.7(a). Let $h, h_g : R \to G$ be as in 2.7(a). Note that it follows from 2.7.0(b) and (c), and from the inequalities

$$d_N(x, h(s)(x)) < B\beta \quad \text{and} \quad d_N(x, h_g(s)(x)) < B\beta$$

for all $s \in [0, 1]$ given in the hypotheses of 2.7, that the following property holds:

2.7.1. (a) $h(s)(x), h_g(s)(x) \in U$ for all $s \in [0, 1]$.

Since $g \in h([0, 1]) \cap h_g([0, 1])$ we may choose numbers $b, c \in [0, 1]$ such that:

2.7.1. (b) $h(b) = g$ and $h_g(c) = g$.

Now we may deduce from 2.7.1(a) and (b), from the fact that $U$ is simply connected (cf. 2.7.0(a)), and from the fact that the algebraic exponential map $\exp : T\hat{G}_e \to \hat{G}$ is a diffeomorphism (where $e \in \hat{G}$ denotes the identity element of the simply connected covering $\hat{G}$ for $G$), that:

2.7.1. (c) $h(bs/c) = h_g(s)$.
for all \( s \in \mathbb{R} \). Now setting \( a = b/c \) in 2.7(a), we see that Property 2.7(a) is a consequence of 2.7.1(c). (If \( g \) is the identity element of \( \Gamma_{t,x} \), set \( a = 0 \).)

Towards verifying 2.7(b) we let \( \pi : \hat{G} \to G \) denote the simply connected covering group of \( G \), and let \( \hat{h}_g : \mathbb{R} \to \hat{G} \) denote the lifting of the \( h_g : \mathbb{R} \to G \) to a one-parameter subgroup of \( \hat{G} \) (for each \( g \in \Gamma_{t,x} \)). By the hypotheses of 2.7 there is for each \( g \in \Gamma_{t,x} \) a number \( r_g \in [0,1] \) such that \( h_g(r_g) = g \); let \( \hat{\Gamma}_{t,x} \subset \hat{G} \) denote the subgroup generated by all the \( \{ \hat{h}_g(r_g) : g \in \Gamma_{t,x} \} \). Note that \( \hat{\Gamma}_{t,x} \) is a torsion free discrete subgroup of \( \hat{G} \); thus \( \hat{\Gamma}_{t,x} \) is a lattice in a closed connected subgroup \( \hat{G}_{t,x} \subset \hat{G} \) (cf. \([17, p. 31, Prop. 2.5]\)). We define \( G_{t,x} \) to be the image of \( \hat{G}_{t,x} \) under the covering projection \( \pi : \hat{G} \to G \). Note that \( G_{t,x} \subset G \) is a closed connected subgroup since \( \hat{G}_{t,x} \subset \hat{G} \) is a closed connected subgroup, \( \hat{\Gamma}_{t,x} \) is co-compact in \( \hat{G}_{t,x} \), and \( \pi(\hat{\Gamma}_{t,x}) \) is discrete. We note (since the algebraic exponential map \( TG_e \to \hat{G} \) is a diffeomorphism) that \( \hat{G}_{t,x} \) is the smallest connected Lie subgroup of \( \hat{G} \) which contains each of the one-parameter subgroups \( \hat{h}_g : \mathbb{R} \to \hat{G} \), \( g \in \Gamma_{t,x} \). Thus \( G_{t,x} \) is the smallest connected Lie subgroup of \( G \) which contains each of the one-parameter subgroups \( h_g : \mathbb{R} \to G \), \( g \in \Gamma_{t,x} \), as claimed in 2.7(b). This completes the verification for 2.7(b).

Towards verifying Property 2.7(c) we note that there is an action \( \hat{G}_{t,x} \times N \to N \) defined to be the composition

\[
\hat{G}_{t,x} \times N \subset \hat{G} \times N \xrightarrow{\times 1_N} G \times N \xrightarrow{\pi} N,
\]

where \( \pi : \hat{G} \to G \) denotes the covering projection and where \( G \times N \to N \) is the given action. Note that \( \hat{G}_{t,x} \times N \to N \) acts thru isometries of \( N \) and has local angle control equal \((\alpha, \beta)\). Now let the sequence of Lie subgroups

\[
e \subset G_2 \subset G_3 \subset \cdots \subset G_t = \hat{G}_{t,x} \text{ be defined by } G_{i-1} = [G_t, G_i].
\]

Note that each \( G_i \times N \to N \) acts thru isometries and has local angle control equal to \((\alpha, \beta)\); we let \( \hat{G}_i \) denote the foliation of \( N \) by the orbits of the action \( \hat{G}_i \times N \to N \) (cf. 2.5). We let \( L_i \subset \hat{G}_i \) denote the leaf containing \( x \); and we let \( d_{L_i}(\cdot,\cdot) \) denote the distance on \( L_i \) induced by the restriction of \( g_N \) to \( TL_i \). Since \( G_i \times N \to N \) also has local angle control equal to \((\alpha, \beta)\) it must also have local angle control equal to \((\alpha_i, \beta_i)\), where \( (\alpha_i, \beta_i) = (10^{2l/\alpha_i} \beta_i, 10^{2l/\beta_i}) \) (cf. 2.4). Thus, we may (by repeating the proof of 2.7.0 with \( G, L, d_L(\cdot,\cdot), (\alpha, \beta) \) replaced by \( G_i, L_i, d_{L_i}(\cdot,\cdot), (\alpha_i, \beta_i) \)) choose an open subset \( U_i \subset L_i \) such that the following properties hold. [Note that the hypothesis of 2.7 that \( N \) has radius of injectivity greater than \( 10^4 \tau B_\beta \) assures us that \( f_x : B_{10^4 \tau B_\beta} \to N \) is an imbedding, provided that \( \tau > 10^{2l} \); this is an important part of the proof of 2.7.0 as applied in the present context.]

2.7.2. (a) \( U_i \) is homeomorphic to a Euclidean space.
(b) \( B(x; L_i; 3B\beta_i) \subset U_i \subset B(x; L_i; 5B\beta_i) \), where \( B(x; L_i; s) \) denotes the open ball of radius \( s > 0 \) (with respect to \( d_{L_i} \)) centered at \( x \) in \( L_i \).

(c) \( 2d_{L_i}(y, z)/3 < d_N(y, z) \leq d_{L_i}(y, z) \) for all \( y, z \in U_i \).

We define subsets \( H_i \subset G_i \) as follows:

\[
H_l = \{ \hat{h}_g(r_g) : g \in \Gamma_t, x \} \;
\]

\( H_{i-1} \) denotes the set of all commutators \( aba^{-1}b^{-1} \) with \( a \in H_l \) and \( b \in H_i \). Now define subsets \( \Gamma_i \subset G_i \) by

\[
\Gamma_i = \bigcup_{j=1}^{i} H_j.
\]

Let \( \bar{\Gamma}_i \) denote the subgroup of \( G_i \) generated by \( \Gamma_i \). Note that \( \bar{\Gamma}_l = \hat{\Gamma}_{t, x} \) and that \( G_l = G_{t, x} \); since \( \bar{\Gamma}_i \) is a lattice in \( \hat{\Gamma}_{t, x} \) (see the verification of 2.7(b) above), we have that \( \bar{\Gamma}_i \) is a lattice in \( G_i \). Thus it follows that for each \( i = 1, 2, \ldots, l \) we have:

2.7.2. (d) \( \bar{\Gamma}_i \) is a lattice for the simply connected nilpotent Lie group \( G_i \), and \( \bar{\Gamma}_i \) is a generating set of \( \Gamma_i \). Moreover \( \bar{\Gamma}_i \) [defined by \( \bar{\Gamma}_i = \bar{\Gamma}_i \cap G_i \)] is a lattice in \( G_i \) which contains \( \Gamma_i \) as a subgroup of finite index; and \( \bar{\Gamma}_i/\bar{\Gamma}_{i-1} \) is a lattice in the abelian Lie group \( G_i/G_{i-1} \).

Note that it follows from the hypothesis of 2.7 that

\[
d_N(x, h_g(s)(x)) < Bt
\]

for all \( g \in \Gamma_{t, x} \) and all \( s \in [0, 1] \); so in particular we have

\[
d_N(x, \hat{h}_g(r_g)(x)) < Bt
\]

for all \( g \in \Gamma_{t, x} \). We deduce from these last two inequalities, and from the definition of \( (\alpha_i, \beta_i) \) and \( \Gamma_i \) given above, and from 2.7.2(a), (b) and (c), that:

2.7.2. (e) \( d_{G_i}(x, h(x)) < B\beta_i \) for all \( i = 1, 2, \ldots, l \) and all \( h \in \Gamma_i \),

where \( d_{G_i}(\ , \ ) \) is the distance function induced on \( G_i \) by the pull back of \( g_N \) under the composite map \( G_i = \hat{G}_i \times \{ x \} \subset G_i \times N \to N \); for future reference we denote this pull back of \( g_N \) by \( g_{G_i} \). Note that \( g_{G_i} \) is a left invariant Riemannian metric on the Lie group \( G_i \).

Now we can use 2.7.2 to verify 2.7(c). We will use an induction argument, and 2.7.2, to verify that the following holds for all \( i = 1, 2, \ldots, l \):

2.7.3. diameter\( (G_i/\bar{\Gamma}_i) < \tau_i\beta_i \),

where \( \tau_i > 1 \) is a number which depends only on \( i \), \( \dim G_i \) and \( B \), and where the diameter is computed with respect to the distance on \( G_i/\bar{\Gamma}_i \) induced by \( g_{G_i} \). Note that Property 2.7(c) is implied by 2.7.3 (with \( i = l \) in 2.7.3),
provided $\tau$ of 2.7(c) satisfies

$$\tau \geq 10^{2l+2} t_1.$$  

Towards verifying 2.7.3 we first note that if $G_i$ is abelian then 2.7.3 follows immediately from 2.7.2(d) and (e) with $\tau_i = B \dim G_i$. In particular 2.7.3 is satisfied for $i = 2$ with $\tau_2 = B \dim G_2$, which is the beginning of our induction argument. Suppose now that 2.7.3 holds for all $i \leq r$. Note that $G_{r+1}/G_r$ is an abelian Lie group equipped with a left invariant Riemannian metric $g_{r+1}$ uniquely determined (from $g_{G_{r+1}}$) by the requirement that the quotient map $G_{r+1} \to G_{r+1}/G_r$ be a Riemannian submersion. Thus it follows immediately from 2.7.2(d) and (e) (with $i = r + 1$ in 2.7.2(d) and (e)), that

$$\text{diameter}((G_{r+1}/G_r)/(\Gamma_{r+1}/\tilde{\Gamma}_r)) < B \beta_{r+1} \dim(G_{r+1}/G_r).$$

Now 2.7.3 (for $i = r + 1$) follows from this last inequality, and from 2.7.3 (for $i = r$), and from the equality $\beta_r = \beta_{r+1}$, provided we define $\tau_{r+1}$ by

$$\tau_{r+1} = \tau_r + 2B \dim (G_{r+1}/G_r).$$

This completes the verification of 2.7(c).

Now we will complete the proof for 2.7 by verifying 2.7(d). Towards this end we first note that for any $g \in G_{t,x}$ as in 2.7(d) the following property holds:

2.7.4. $d_{L_{t,x}}(x, g(x)) < 3t/2$,

where $d_{L_{t,x}}(\cdot, \cdot)$ is the distance induced by the restricted Riemannian metric $g_N|\{L_{t,x}\}$.

To verify 2.7.4 we will assume it is not true and derive a contradiction. Recall that

$$d_N(x, g(x)) < t$$

is a hypothesis of 2.7(d). Recall also that for $L_i$ of 2.7.2 we have that $L_i = L_{t,x}$ when $i = l$. Thus if 2.7.4 doesn’t hold then we deduce from the preceding inequality, and from 2.7.2(b) and (c), that:

2.7.5. (a) $d_{L_{t,x}}(x, g(x)) \geq 3B\beta_l$.

The number $\beta_l$ of 2.7.2 was defined to be $10^{2l+2} t$; however 2.7.2(a)-(c) can also be verified as before if $(\alpha_l, \beta_l)$ is chosen to be

$$(\alpha_l, \beta_l) = (100 \tau t \beta, 100 \tau t),$$

where $\tau > 1$ comes from 2.7(c) and (d). (Note that the hypothesis of 2.7, that $N$ has radius of injectivity greater than $10^4 \tau B\beta$, at $x$, implies that the map $f_x : B_{10B\beta}^\infty \to N$ is an embedding for $\beta_1 = 100 \tau t$; this is an important part of the verification of 2.7.2(a)-(c) for our present choice of $(\alpha_l, \beta_l)$.) This last equality, and Properties 2.7.5(a) and 2.7.2(c), together imply:
2.7.5. (b) $d_{L_t,x}(x, g(x)) > 100\tau Bt$.

Using 2.7(c) we may choose $g \in \Gamma_{t,x}$ such that:

2.7.5. (c) $d_{L_t,x}(g(x), \bar{g}(x)) < \tau t$.

Now combining 2.7.5(b) and (c), and recalling that $B > 1$ (cf. A.2.3), we deduce that:

2.7.5. (d) $d_{L_t,x}(x, \bar{g}(x)) > 90\tau Bt$.

Now this last inequality leads to a contradiction (which will complete the verification of 2.7.4) as follows: Since $d_N(x, g(x)) < t$, and $\tau > 1$, it follows from 2.7.5(c) that $d_N(x, \bar{g}(x)) < 2\tau t$, and so by definition of $\Gamma_{2\tau t,x}$ we have that $g \in \Gamma_{2\tau t,x}$. Then the hypothesis of 2.7(d) (that $G_{t,x} = G_{2\tau t,x}$), together with 2.7(b), assures us that $\text{Image}(h_\bar{g}) \subset G_{t,x}$ (where $h_\bar{g}$ comes from the hypothesis of 2.7). Thus the equation $f_{g}(s) = h_{\bar{g}}(s)x$ defines a path $f_{\bar{g}} : [0, 1] \to L_{t,x}$ which starts at $x$ and contains $\bar{g}x$ and which (by the hypothesis of 2.7 applied to $h_{\bar{g}}$) satisfies:

2.7.5. (e) $d_N(x, f_{\bar{g}}(s)) < 2\tau Bt$

for all $s \in [0, 1]$. Now 2.7.5(e) and 2.7.2(b) and (c), together with the continuity of $f_{\bar{g}}$ imply:

2.7.5. (f) $d_{L_t,x}(x, f_{\bar{g}}(s)) < 4\tau Bt$

for all $s \in [0, 1]$. [When applying 2.7.2(c) to verify 2.7.5(f) we set $i = l$ and $\beta_l = 100\tau t$ in 2.7(b) and (c), as in the derivation of 2.7.5(b) above.] Since $\bar{g}x \in \text{image}(f_{\bar{g}})$, we see that 2.7.5(f) contradicts 2.7.5(d). This contradiction completes the verification of 2.7.4.

In our verification of 2.7(d) we need also consider the composite map

$$TG_e \xrightarrow{\exp} G \xrightarrow{h} N,$$

where $TG_e$ is the tangent space for $G$ at the identity $e \in G$, $\exp : TG_e \to G$ is the algebraic exponential map, and $h : G \to N$ is given by $h(g) = g(x)$. Note that $h \circ \exp$ is an immersion (cf. 2.5). $TG_e$ is equipped with an inner product gotten by pulling $g_N|_{T(L_{t,x})}$ back along the linear map $D(\exp \circ h) : TG_e \to T(L_{t,x})$. For each unit vector $v \in TG_e$ we have the one-parameter subgroup $\rho_v : R \to G$ defined by $\rho_v(s) = \exp(sv)$. Define the action $\gamma_v : R \times N \to N$ by $\gamma_v(s, y) = \rho_v(s)(y)$. In order to complete the proof for 2.7(d) we shall need that the $\gamma_v$ satisfy the following property:

2.7.6. $|d_N(x, \gamma_v(s, x)) - s| < O(\alpha)s$ for all $s \in [-4\beta, 4\beta]$.

Towards verifying 2.7.6 we first note that each action $\gamma_v : R \times N \to N$ has local angle control equal $(\alpha, \beta)$, because the action $G \times N \to N$ has local
angle control equal \((\alpha, \beta)\). So (by 2.5 applied to \(\gamma_v\)) the orbits of the action \(\gamma_v : \mathbf{R} \times N \to N\) are the leaves of a smooth foliation \(\mathcal{G}_v\) of \(N\); and we may apply 2.6 to each action \(\gamma_v : \mathbf{R} \times N \to N\) to conclude that:

2.7.7. (a) \(K(\mathcal{G}_v; N) < 10^6 \alpha/\beta\).

Next we consider again the map \(f_x : B^n_{10\beta} \to N\) given in A.1.1 and A.1.2 (where \(\epsilon, p, M, B^n_n\) are replaced in A.1.1, A.1.2 by \(10\beta, x, N, B^n_{10\beta}\) with \(n = \dim N\)). Set \(\hat{\mathcal{G}}_v = f_x^{-1}(\mathcal{G}_v)\); and let \(K(\hat{\mathcal{G}}_v; B^n_{10\beta})\) denote the curvature for \(\hat{\mathcal{G}}_v\) in \(B^n_{10\beta}\) with respect to the Euclidean metric (cf. 2.6(ii)). The proof of Theorem A.1.4 also works to show the following (cf. A.1.4.1):

2.7.7. (b) \(K(\hat{\mathcal{G}}_v; B^n_{10\beta}) - O(10\beta) < K(\mathcal{G}_v; N) < K(\hat{\mathcal{G}}_v; B^n_{10\beta}) + O(10\beta)\).

Now by combining 2.7.7(a) and (b) we conclude that:

2.7.7. (c) \(K(\hat{\mathcal{G}}_v; B^n_{10\beta}) < 10^6 \alpha/\beta + O(10\beta)\).

Let \(g_{i,j}dx_idx_j\) denote the Riemannian metric on \(B^n_{10\beta}\) gotten by pulling back \(g_N\) along \(f_x : B^n_{10\beta} \to N\); let \(\hat{d}(\ ,\ )\) denote the distance on \(B^n_{10\beta}\) induced by \(g_{i,j}dx_idx_j\), and let \(\hat{d}(\ ,\ )\) denote the Euclidean distance on \(B^n_{10\beta}\). Let \(\hat{L}_v\) denote the (connected) leaf of \(\hat{\mathcal{G}}_v\) which contains the origin. Denote by \(\hat{d}_v(\ ,\ )\) the distance induced on \(\hat{L}_v\) by the restriction to \(T\hat{L}_v\) of \(g_{i,j}dx_idx_j\); and denote by \(d_v(\ ,\ )\) the distance on \(\hat{L}_v\) induced by the restriction to \(T\hat{L}_v\) of the Euclidean structure \(\delta_j^i dx_idx_j\). Note that we can deduce from 2.7.7(c) that:

2.7.7. (d) \(|d(0, y) - d_v(0, y)| < O(\alpha)d_v(0, y)|\),

for all \(y \in \hat{L}_v\), provided \(\alpha, \beta\) are sufficiently small (as assumed in the hypothesis of 2.7). [In more detail we could use 2.7.7(c) (in place of 2.7.0.2 above) to verify that \(\hat{\mathcal{G}}_v\) satisfies Properties 2.7.0.3(a) and (b) for \(k = 1\); where in 2.7.0.3 \(h : V \to \mathbf{R}^{n-1}\) is a smooth map from an open subset \(V \subset \mathbf{R}\) such that \(\hat{L}_v\) is equal to the graph of \(h\). Then we could deduce 2.7.7(d) from this version of 2.7.0.3.] Next we note it follows from the relation between the metrics \(g_{i,j}dx_idx_j\) and \(\delta^i_j dx_idx_j\) given in Theorem A.1.1, and also given in A.1.3.3.1, A.1.3.4.1, that:

2.7.7. (e) \(|\hat{d}(0, y) - \hat{d}(0, y)| + |d_v(0, y) - \hat{d}_v(0, y)| < O(d_v(0, y))d_v(0, y)|\)

for all \(y \in \hat{L}_v\), provided \(\beta\) is sufficiently small. (Note that \(d(0, y) = \hat{d}(0, y)\).) Now by combining 2.7.7(d) and (e) we can deduce that:

2.7.7. (f) \(|\hat{d}(0, y) - \hat{d}_v(0, y)| < O(\alpha)d_v(0, y)|\)
for all \( y \in \hat{L}_v \) with \( \hat{d}(0, y) \leq 4\beta \), provided \( \alpha \) and \( \beta \) are sufficiently small and \( \beta < \alpha \) (as hypothesized in 2.7). Finally we recall that \( f_x : B_{10\beta}^n \to N \) is an embedding, since one of the hypotheses of 6.2.7 is that \( N \) has injectivity radius greater than \( 10^4 \tau B \beta \) at \( x \) with \( \tau > 1 \) and \( B > 1 \). Thus, the desired inequality 2.7.6 is equivalent to the inequality of 2.7.7(f).

Now we can complete the proof of 2.7(d) by applying 2.7.4, 2.7.6, and 2.7.2(a)-(c) (with \( i = l \) in 2.7.2(a)-(c)). For any \( s > 0 \) let \( B_e(s) \) denote the closed ball of radius \( s \) centered at the origin of \( TG_e \), and let \( L_l \) and \( B(x; L_l, s) \) and \( \beta_l > 0 \) be as in 2.7.2(a), (b) and (c) (recall \( \beta_l = 10^{2l+2} \tau \)).

Recall that \( \exp : TG_e \to G \) denotes the algebraic exponential map and \( h : G \to N \) is given by \( h(g) = g(x) \). Note that 2.7(d) may be deduced directly from Property 2.7.4 and the following property. (Also recall that \( G_l \) of 2.7.2(a)-(c) is equal \( \hat{G}_{t,x} \).)

2.7.8. \( B(x; L_l; 3t/2) \subseteq h \circ \exp(B_e(2t) \cap TG_{t,x}) \).

Towards verifying 2.7.8 we note that it follows from 2.7.2(a) and (b), and from 2.7.6, and from the equality \( G_l = \hat{G}_{t,x} \), that \( h \circ \exp(B_e(2t) \cap TG_{t,x}) \subseteq U_l \) and \( h \circ \exp : B_e(2t) \to U_l \) is a smooth embedding, provided \( \alpha > 0 \) is sufficiently small. Now since \( B_e(2t) \cap TG_{t,x} \) is a closed ball, and \( U_l \) is homeomorphic to Euclidean space (cf. 2.7.2(a)), and \( \dim(B_e(2t) \cap TG_{t,x}) = \dim U_l \), it follows that the complement in \( U_l \) of the sphere \( h \circ \exp(\partial(B_e(2t) \cap TG_{t,x})) \) consists of two open connected sets \( U^+_l \) and \( U^-_l \), with

\[
U^+_l = h \circ \exp((B_e(2t) \cap TG_{t,x}) - \partial(B_e(2t) \cap TG_{t,x})).
\]

It follows from 2.7.6 that

\[
B(x; L_l; 3t/2) \cap h \circ \exp(\partial(B_e(2t) \cap TG_{t,x})) = \phi
\]

provided \( \alpha > 0 \) is sufficiently small; so the connected set \( B(x; L_l; 3t/2) \) must be contained in one of \( U^+_l \) or \( U^-_l \). Since both \( U^+_l \) and \( B(x; L_l; 3t/2) \) contain the point \( x \), we have that

\[
B(x; L_l; 3t/2) \subseteq U^+_l.
\]

This completes the verification of 2.7.8. \( \square \)

**Proof of Lemma 2.8.** The proof is carried out in the following six steps:

**Step I.** In this step we let \( G \) be a connected nilpotent Lie group with Lie algebra \( g \); \( G \) is equipped with a left invariant Riemannian metric. We will say that \( G \) has \((\alpha, \beta)\)-angular control when the left multiplication group action \( G \times G \to G \) has local angle control equal \((\alpha, \beta)\).

The purpose of this step is to verify the following claim. The notation \( \hat{u} \) will denote the left invariant vector field on \( G \) determined by a vector \( u \in g \).
Claim 2.8.1. Assume $\alpha < \pi$. Then, $G$ has $(\alpha, \beta)$-angular control iff for all $u, v \in \mathfrak{g}$ we have

$$|D\hat{v}\hat{u}| \leq (\alpha/\beta)|u||v|.$$ 

Towards verifying 2.8.1 we re-look at the Proof of Lemma 2.3 in the special case when $N = G$, $G \times G \to G$ is the group multiplication, and $x = e$. And observe that

$$d/dt (A - v)_{t=0}(u) = D\hat{v}\hat{u}(e)$$

for each pair $u, v \in \mathfrak{g}$ where $|v| = 1$. But, $d/dt (A - v)_{t=0}$ is the blocked sum of the zero matrix of a certain size and $2 \times 2$ matrices of the form

$$\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix},$$

where $0 \leq \theta \leq \alpha/\beta$. Now 2.8.1 is easily seen to follow from these observations.

Step II. In this step we let $G$ be a connected nilpotent Lie group with Lie algebra $\mathfrak{g}$; $G$ is equipped with a left invariant Riemannian metric. The purpose of this step is to verify the following claim. In this claim we set $k = \dim G$.

Claim 2.8.2. If $G$ has $(\alpha, \beta)$-angular control with $\alpha < \pi$, then for any pair $x, y \in \mathfrak{g}$ we have

$$|[x, y]| \leq (2k^2 \alpha/\beta)|x||y|.$$ 

Since $\mathfrak{g}$ is nilpotent, we can fix an orthonormal basis $x_1, x_2, \ldots, x_k$ for $\mathfrak{g}$ (with $k = \dim \mathfrak{g}$) such that

$$[x_i, x_j] \in \text{Span}\{x_0, x_1, \ldots, x_{i-1}\}$$

for every pair of indices $i, j \geq 1$. (We define $x_0 = 0$.) Recall the following standard formula (cf. [2, Prop. 7.7.1]). Let $X, Y$, and $Z$ be left invariant vector fields on $G$, then

$$(D_X Y) \cdot Z = 1/2\{-[X,Y] \cdot Z + [Y, Z] \cdot X + [X, Z] \cdot Y\}$$

where the “dot” indicates the inner product of vector fields. Applying this formula where $X \in \text{Span}\{\hat{x}_i\}$, $Y \in \text{span}\{\hat{x}_j\}$, and $Z = [X, Y]$, yields

$$(D_X Y) \cdot [X, Y] = -1/2[X, Y] \cdot [X, Y].$$

And combining this equation with Claim 2.8.1, yields:

2.8.2.1. $|[X, Y]| \leq (2\alpha/\beta)|X||Y|$

provided $G$ has $(\alpha, \beta)$-angular control and $\alpha < \pi$. 
Now we will deduce Claim 2.8.2 from 2.8.2.1 as follows: Set \( \hat{x} = \sum_{i=1}^{k} X_i \)
and \( \hat{y} = \sum_{i=1}^{k} Y_i \) where \( X_i, Y_i \in \text{Span}(\hat{x}_i) \). Then, by 2.8.2.1, we have
\[
|\langle x, y \rangle| \leq \sum_{i,j=1}^{k} |\langle X_i, Y_j \rangle| \leq \sum_{i,j=1}^{k} (2\alpha/\beta) |X_i| |Y_j|.
\]
Since \( \{x_1, x_2, \ldots, x_k\} \) is an orthonormal basis we also have that
\[
\sum_{i,j=1}^{k} (2\alpha/\beta) |X_i| |Y_j| = (2\alpha/\beta) \left( \sum_{i=1}^{k} |X_i| \right) \left( \sum_{i=1}^{k} |Y_i| \right) \leq (2\alpha/\beta)(k|x|)(k|y|).
\]
Combining these last two sets of inequalities yields 2.8.2.

**Step III.** We now let \( G \) be a **simply connected** nilpotent Lie group with Lie algebra \( g \); \( G \) is equipped with a left invariant Riemannian metric. The purpose of this step is to verify the following claim:

**Claim 2.8.3.** Let \( h \) be an ideal in \( g \) and let \( H \) be the corresponding (closed and normal) subgroup of \( G \). If \( G \) has \((\alpha, \beta)\)-angular control, then \( G/H \) also has \((\alpha, \beta)\)-angular control provided we equip \( G/H \) with the (unique) left invariant Riemannian metric which makes the quotient homomorphism \( q : G \to G/H \) a Riemannian submersion.

Note that \( H \) is closed because \( G \) is simply connected.

Let \( m = h^\perp \); i.e., \( m \) is the orthogonal complement to \( h \) in \( g \). Identify \( m \) with \( g/h \) via the quotient map \( g \to g/h \). And, for each \( u \in m \), let \( \bar{u} \) denote the left invariant vector field on \( G/H \) determined by \( u \). Denote the covariant derivative operators for \( G \) and \( G/H \) by \( \bar{D} \) and \( D \) respectively. Since the quotient map \( q : G \to G/H \) is a Riemannian submersion, we may apply [16, p. 212, Lemma 45] to get that
\[
|\bar{D}_v \bar{u}| \geq |D_v u|
\]
for all \( u, v \in m \). From this inequality, and also from Claim 2.8.1, we conclude that \( G/H \) has \((\alpha, \beta)\)-angular control. (Note that we may assume that \( \alpha < \pi \) in 2.8.1 since Claim 2.8.3 is otherwise trivially true.) This completes the verification for Claim 2.8.3.

**Step IV.** In this step we let \( G \) be a simply connected nilpotent Lie group with Lie algebra \( g \); \( G \) is equipped with a left invariant Riemannian metric. Let \( \gamma_i : \mathbb{R} \to G \) be the one-parameter subgroup of \( G \) such that \( \hat{\gamma}(0) = x_i \) (where the \( \{x_i\} \) come from Step II). Define a map \( f : \mathbb{R}^k \to G \) by
\[
f(t_1, t_2, \ldots, t_k) = \gamma_1(t_1)\gamma_2(t_2)\ldots\gamma_k(t_k).
\]
Then it is a classical result for simply connected nilpotent Lie groups that $f$ is a diffeomorphism. Thus, for each $g \in G$, we may define coordinates $\{g_i\}$ for $g$ by

$$(g_1, g_2, \ldots, g_k) = f^{-1}(g).$$

The purpose of this step is to verify the following claim:

**Claim 2.8.4.** Suppose that $G$ has $(\alpha, \beta)$-angular control where $\alpha$ is sufficiently small. (How small is sufficient depends only on $k = \dim G$.) The following inequality holds for each $g \in G$ with $d_G(g, e) \leq \beta$:

$$|g_1| \leq 2d_G(g, e).$$

To verify this claim we proceed by induction on $k = \dim G$. (We may assume that $k > 1$ since 2.8.4 is clearly true when $k = 1$.) Consider the function $\psi : G \to \mathbb{R}$ defined by

$$\psi(h) = h_1$$

for $h \in G$, and let the exact 1-form $\omega$ be defined by

$$\omega = d\psi.$$

Let $\gamma : [0, 1] \to G$ be a constant speed smooth path such that:

1. $(a) \; \gamma(0) = e$ and $\gamma(1) = g$,
2. $(b) \; \text{length}(\gamma) = d_G(g, e)$.

Since $g_1 = \int_0^1 \omega(\dot{\gamma}(t))dt$, we also have that:

**2.8.4.1.** $(c) \; |g_1| \leq \int_0^1 |\omega(\dot{\gamma}(t))|dt$.

Now, by 2.8.4.1, in order to complete the verification for Claim 2.8.4 it will suffice to verify the following property:

**2.8.4.2.** There exists a positive real number $C_k$ such that the following is true if $\alpha \leq C_k$. For each $h \in G$ with $d_G(h, e) \leq \beta$ and for each $v \in TG_h$ we have that $|\omega(v)| \leq 2|v|$.

To verify 2.8.4.2 we let $H$ denote the image of the one-parameter subgroup $\gamma_1$ and set $\mathfrak{h} = \text{span}\{x_1\}$. Then, $G/H$ is a simply connected nilpotent Lie group whose Lie algebra is $\mathfrak{g}/\mathfrak{h}$ and the quotient homomorphism $q : G \to G/H$ does not increase distances. (We equip $G/H$ with the left invariant Riemannian metric so that $q$ is a Riemannian submersion.) Also, $G/H$ has $(\alpha, \beta)$-angular control because of Claim 2.8.3. For each index $i \geq 0$ let $\bar{x}_i$
denote the coset \( x_{i+1} + h \). Then \( x_1, x_2, \ldots, x_{k-1} \) is an orthonormal basis for \( g/h \) such that
\[
[x_i, x_j] \in \text{Span}\{x_0, x_1, \ldots, x_{i-1}\}
\]
for all \( i, j \geq 1 \). Let \( f : \mathbb{R}^{k-1} \to G/H \) be the diffeomorphism corresponding to this basis (defined analogously to \( f : \mathbb{R}^k \to G \) above). And notice that \( h_2 = [q(h)]_1 \)
where \( h \) comes from 2.8.4.2 and where \( [q(h)]_1 \) is the first coordinate of \( \overline{f}^{-1}(q(h)) \). Hence our inductive assumption yields that
\[
|h_2| \leq 2d_{G/H}(q(h), e) \leq 2d_G(h, e).
\]
We see continuing in this way that:

**2.8.4.3.** (a) \( |h_i| \leq 2d_G(h, e) \leq 2\beta \)
for all \( i \geq 2 \). Associate to \( h \) a finite sequence of group elements \( h(1), h(2), \ldots, h(k) \in G \) defined by
\[
h(i) = f(0, 0, \ldots, 0, h_{i+1}, h_{i+2}, \ldots, h_k).
\]
And define a finite sequence of vectors \( y_1, y_2, \ldots, y_k \in g \) by:

**2.8.4.3.** (b) \( y_i = \text{Ad}(h(i)^{-1})(x_i) \)
where \( \text{Ad} : G \to GL(g) \) is the adjoint representation. Then note that:

**2.8.4.3.** (c) \( \{\hat{y}_2(h), \hat{y}_3(h), \ldots, \hat{y}_k(h)\} \) is a basis for \( \text{Ker}(\omega_h) \), and \( y_1 = x_1 \),
where \( \omega_h : TG_h \to \mathbb{R} \) is the restriction of the 1-form \( \omega \) to the tangent space to \( G \) at \( h \). (Recall \( \hat{x} \) is the left invariant vector field determined by each \( x \in G \).) A continuity argument, together with 2.8.4.3(c), shows that there exists a positive number \( \overline{C}_k \) such that if:

**2.8.4.3.** (d) \( |y_i - x_i| \leq \overline{C}_k \)
for all \( i \geq 2 \), then \( |\omega(v)| \leq 2|v| \) (as desired in 2.8.4.2). We proceed to show that there exists a positive number \( C_k \) such that if \( \alpha \leq C_k \), then 2.8.4.3(d) holds for all indices \( i \). (This will complete the verification of 2.8.4.2, and hence also of Claim 2.8.4.)

We note that by the definitions of \( h(i) \) and of \( f : \mathbb{R}^k \to G \) we have that:

**2.8.4.4.** (a) \( h(i) = \gamma_{i+1}(h_{i+1})\gamma_{i+2}(h_{i+2}) \cdots \gamma_k(h_k) \).
Substituting 2.8.4.4(a) into 2.8.4.3(b) yields:

**2.8.4.4.** (b) \( y_i = \text{Ad}(\gamma_k(-h_k)) \cdots \text{Ad}(\gamma_{i+1}(-h_{i+1}))(x_i) \).
Note now that, for all indices \( j \), we have:
2.8.4.4. (c) \( \text{Ad}(\gamma_j(-h_j)) = e^{-h_j \text{ad}(x_j)} \)

where \( \text{ad}(x_j) : g \to g \) is defined by \( \text{ad}(x_j)(u) = [x_j, u] \) for all \( u \in g \). (The equation in 2.8.4.4(c) follows from [12, formula (5), p. 118], and from the fact that \( \gamma_j(-h_j) = \text{Exp}(-h_j x_j) \) where \( \text{Exp} : g \to G \) is the algebraic exponential map.) Now, by combining 2.8.4.4(c) with Claim 2.8.2 and 2.8.4.3(a) and \( |x_i| = 1 \), we obtain:

\[
2.8.4.4. \quad (d) \quad |\text{Ad}(\gamma_j(-h_j))(u) - u| \leq |u|(e^{4k^2\alpha} - 1)
\]

for all indices \( j \) and every \( u \in g \). In particular, we have that:

\[
2.8.4.4. \quad (e) \quad |\text{Ad}(\gamma_j(-h_j))(u)| \leq |u|e^{4k^2\alpha}.
\]

It now follows from 2.8.4.4(b), (d) and (e), and from the fact that \( |x_i| = 1 \) for all \( i \), that:

\[
2.8.4.4. \quad (f) \quad |y_i - x_i| < e^{4k^3\alpha} - 1
\]

for all indices \( i \). This last inequality shows that the positive number \( C_k \) (posed above) clearly exists since \( e^x \) is continuous and \( e^0 = 1 \).

This completes the verification of Claim 2.8.4.

Step V. In this step we let \( G \) be a simply connected nilpotent Lie group with Lie algebra \( g \); \( G \) is equipped with a left invariant Riemannian metric. Recall that the algebraic exponential map \( \text{Exp} : g \to G \) is defined by \( \text{Exp}(v) = \gamma_v(1) \) where \( \gamma_v \) is the one parameter subgroup of \( G \) such that \( \gamma_v(0) = v \). It is a classical result for simply connected nilpotent Lie groups that \( \text{Exp} \) is a diffeomorphism. The purpose of this step is to verify the following claim. In this claim we set \( k = \dim G \) and let \( e \in G \) denote the identity element.

Claim 2.8.5. Suppose that \( G \) has \((\alpha, \beta)\)-angular control where \( \alpha \) is sufficiently small. (How small is sufficient depends only on \( \dim G \).) Then, for each \( g \in G \) such that \( d_G(g, e) \leq \beta \), we have that

\[
|\text{Exp}^{-1}(g)| \leq 4^k d_G(g, e).
\]

One easily observes that the following statement is a consequence of the proof given above for Lemma 2.3:

2.8.5.1. For every \( \lambda \geq 1 \), \((\alpha, \beta)\)-angular control implies \((\lambda \alpha, \lambda \beta)\)-angular control.

Now to verify Claim 2.8.5 we proceed by induction on \( k = \dim G \). (We may assume \( k > 1 \) since 2.8.5 is clearly true when \( k = 1 \).) Let \( H \) denote the image of the one-parameter subgroup \( \gamma_1 \); it is a closed subgroup contained in the center of \( G \). Hence, \( G/H \) is also a simply connected nilpotent Lie
group whose Lie algebra is \( \mathfrak{g}/\mathfrak{h} \) where \( \mathfrak{h} = \text{Span}\{x_1\} \) is the Lie algebra of \( H \).

Put the (unique) left invariant Riemannian metric on \( G/H \) which makes the quotient homomorphism \( q : G \to G/H \) a Riemannian submersion. Note that \( G/H \) also has \((\alpha, \beta)\)-angular control (by Claim 2.8.3) and 2.8.5 is valid for \( G/H \) since \( \dim(G/H) = k - 1 \). Set

\[
\text{Exp}^{-1}(g) = sx_1 + u,
\]

where \( u \in \text{Span}\{x_2, x_3, \ldots, x_k\} \), and let

\[
b = \text{Exp}(u) \quad \text{and} \quad c = \text{Exp}(sx_1) = \gamma_1(s).
\]

Note that:

\[2.8.5.2.\]

(a) \( g = cb \)

since \( c \in \text{Center}(G) \). Also let \( \bar{g} \) denote \( q(g) \). Then:

\[2.8.5.2.\]

(b) \( d_{G/H}(\bar{g}, e) \leq d_G(g, e) \leq \beta \)

since \( q : G \to G/H \) does not increase distances. Also note that \( \text{Exp}_{G/H}(u + \mathfrak{h}) = \bar{g} \) where \( \text{Exp}_{G/H} : \mathfrak{g}/\mathfrak{h} \to G/H \) denotes the algebraic exponential map for \( G/H \). Applying 2.8.5 to \( \bar{g} \), and using 2.8.5.2(b), we obtain that:

\[2.8.5.2.\]

(c) \(|u| \leq 4^{k-1}d_G(g, e)\).

Of course, this last inequality implies directly that:

\[2.8.5.2.\]

(d) \( d_G(b, e) \leq 4^{k-1}d_G(g, e) \).

Now, 2.8.5.2(a), and the fact that \( d_G(\ , \ ) \) is left invariant, imply that:

\[2.8.5.2.\]

(e) \( d_G(c, e) \leq d_G(g, e) + d_G(e, b) \).

Combining 2.8.5.2(d) and (e) yields:

\[2.8.5.2.\]

(f) \( d_G(c, e) \leq (4^{k-1} + 1)d_G(g, e) \).

Since \( \gamma_1(s) = c \), we also have that:

\[2.8.5.2.\]

(g) \( c_1 = s \)

where \( c_1 \) is the first coordinate of \( f^{-1}(c) \).

Now let \( C_k \) be the positive real number such that the conclusion of Claim 2.8.4 is true (in dimension \( k \)) provided \( \alpha \leq C_k \), and let \( \bar{C}_{k-1} \) be the corresponding number such that the conclusion of Claim 2.8.5 is true (in dimension \( k - 1 \)) provided \( \alpha \leq \bar{C}_{k-1} \). Set

\[
\bar{C}_k = \min\{\bar{C}_{k-1}, (4^{k-1} + 1)^{-1}C_k\}.
\]
In particular, we assume that $G$ has $(\alpha, \beta)$-angular control where $\alpha \leq C_k$. And $G$ also has $(C_k, \bar{\beta})$-angular control where $\bar{\beta} = (4^{k-1} + 1)\beta$

because of 2.8.5.1. Now note that $d_G(c, e) \leq \beta$ because of 2.8.5.2(f) and our assumption that $d_G(g, e) < \beta$. We can consequently apply 2.8.4 to $c$ and, using again 2.8.5.2(f), conclude that:

2.8.5.3. (a) $|c_1| \leq 2d_G(c, e) \leq 2(4^{k-1} + 1)d_G(g, e)$.

But the triangle inequality, and 2.8.5.2(g), and $|x_1| = 1$, together yield that:

2.8.5.3. (b) $|\text{Exp}^{-1}(g)| = |sx_1 + u| \leq |c_1| + |u|$.

Now, by combining the inequalities of 2.8.5.2(c) and 2.8.5.3(a) and (b) we get that

$$|\text{Exp}^{-1}(g)| \leq 4^{k}d_G(g, e)$$

as required in 2.8.5.

This completes the proof for Claim 2.8.5.

Step VI. In this step we shall complete the proof for Lemma 2.8. We assume $s > 0$ since the conclusion of 2.8 is trivially true when $s = 0$. Let $\rho: \tilde{G} \to G$ denote the universal covering group of $G$ and put the induced left invariant Riemannian metric on $\tilde{G}$. Clearly, $\tilde{G}$ also has $(\alpha, \beta)$-angular control. Let the one parameter subgroup $\tilde{\phi}_1 : \mathbb{R} \to \tilde{G}$ be the unique lift of $\phi_1$. Furthermore, let $\mu : [0, 1] \to G$ be a smooth path such that:

2.8.6. (a) $\text{length}(\mu) < t$,

(b) $\mu(0) = e$ and $\mu(1) = \phi_1(s)$.

And, let $\tilde{\mu} : [0, 1] \to \tilde{G}$ be the unique lift of $\mu$ such that $\tilde{\mu}(0) = e$. Set $g = \tilde{\mu}(1)$ and observe that $d_{\tilde{G}}(g, e) \leq \beta$ (cf. 2.8.6(a) and recall $t < \beta$). Thus we may apply Claim 2.8.5 to $\tilde{G}$ to get a (unique) one-parameter subgroup $\gamma : [0, s] \to \tilde{G}$ such that:

2.8.6. (c) $\gamma(s) = g$,

(d) $|\dot{\gamma}(0)| \leq s^{-1}4^{k}d_{\tilde{G}}(g, e)$.

The following useful inequality is a direct consequence of 2.8.6(d) and of the fact that $\gamma$ has constant speed:

2.8.6. (e) $d_{\tilde{G}}(\gamma(u), e) \leq (u/s)4^{k}t$ for all $u \in [0, s]$.

We denote by $C \subset G$ and $\tilde{C} \subset \tilde{G}$ the centers of $G$ and $\tilde{G}$, respectively. Note that $\tilde{C} = p^{-1}(C)$; consequently, $G/C = \tilde{G}/\tilde{C}$ is a simply connected
nilpotent Lie group. Let \( q : G \to G/C \) denote the quotient homomorphism. Then, the two one-parameter subgroups \( q \circ \rho \circ \gamma \) and \( q \circ \phi_1 \) are the same because the algebraic exponential map for \( G/C \) is a diffeomorphism and because of the equation

\[
q \circ \rho \circ \gamma(s) = q \circ \phi_1(s).
\]

Consequently, there is a one-parameter subgroup \( \psi : \mathbb{R} \to \tilde{C} \) such that \( \tilde{\phi}_1 \) factors as

\[
\tilde{\phi}_1(u) = \psi(u)\gamma(u)
\]

for all \( u \in \mathbb{R} \). Now, set

\[
\phi_2 = \rho \circ \psi.
\]

Then, \( \phi_1 \) factors as

\[
\phi_1(u) = (\phi_2(u))(\rho \circ \gamma(u))
\]

for all \( u \in \mathbb{R} \). Evaluating this factorization at \( u = s \) yields that \( \phi_2(s) = e \) (since \( \rho \circ \gamma(s) = \phi_1(s) \)). Thus Statement (b) of 2.8 is verified.

Since the metric on \( G \) is left invariant, the last factorization also yields that

\[
d_G(\phi_1(u), \phi_2(u)) = d_G(\rho \circ \gamma(u), e)
\]

for all \( u \in [0, s] \). And observe that

\[
d_G(\rho \circ \gamma(u), e) \leq d_{\tilde{G}}(\gamma(u), e)
\]

because \( \rho : \tilde{G} \to G \) does not increase distances. Combining these last two facts with 2.8.6(e) yields that statement (c) of Lemma 2.8 is true.

This completes the proof for Lemma 2.8. \( \square \)

### 3. Proof of Theorem 0.3.

It is recommended that the reader consult Appendix 2 before reading this section.

Let \( \eta, \epsilon, \delta_i \) be as in 1.3. Let \( \epsilon, \delta \) be as in Appendix 2 (cf. A.2.1-A.2.3); so as not to confuse the \( \epsilon \) of 1.3 with the \( \epsilon \) of Appendix 2, we shall henceforth denote the latter by \( \tilde{\epsilon} \). Recall that \( E \) denotes the total space of the frame bundle \( \rho : E \to \tilde{V} \) defined in Appendix 2. We choose \( \eta, \epsilon, \delta_i \) of 1.3 and \( \tilde{\epsilon}, \delta \) of Appendix 2 as follows, where \( m = \dim M \):

**3.0.** (a) \( \eta = (m + 4)^8 \).
(b) \( \delta_{i+1} \ll \epsilon \delta_i^{\dim E + 2} \).
(c) \( \delta_1 \ll \delta < \tilde{\epsilon} \ll \epsilon < 1 \).
(To satisfy inequalities (b) and (c) first choose $\delta$ and $\epsilon$ and then choose the $\delta_1, \delta_2, \ldots$.)

We shall prove Theorem 0.3 when $n = 0$ in 0.3: The proof for $n > 0$ is essentially the same. With this in mind it will prove convenient to introduce the following notation for some of the first $(m + 4)^8$ of the $\{\delta_i\}$. [Note that the following numbers do not exhaust all of the first $(m + 4)^8$ of the $\{\delta_i\}$.]

3.0. (d) For any integers $i, j, k, l \in \{1, 2, \ldots, (m + 4)^2 - 1\}$ we set

\[ t_i = \delta_i(m + 4)^6, \]
\[ t_{i, j} = \delta_i(m + 4)^6 + j(m + 4)^4, \]
\[ t_{i, j, k} = \delta_i(m + 4)^6 + j(m + 4)^4 + k(m + 4)^2, \]
\[ t_{i, j, k, l} = \delta_i(m + 4)^6 + j(m + 4)^4 + k(m + 4)^2 + l. \]

For given $p \in M$ we use Theorems A.2.2 and A.2.3 to choose a smooth group action $H \times \hat{V} \to \hat{V}$ which satisfies A.2.1(a)-(c) such that the lifted action $H \times E \to E$ satisfies A.2.3(a) and (b). Let $H_e \subset H$ denote the identity component, and set $\Lambda_e = \Lambda \cap H_e$ (see A.2.1 for $\Lambda \subset H$). The remainder of this proof is carried out in Subsections 3.1-3.5 below.

3.1. The groups $G_{t, x}$, $C_{i, j, x}$, and $\Lambda_{i, j, k, x}$. In this subsection we identify the $G \times N \to N$ and $g_N$ and $G$ of §2 with $H_e \times E \to E$ and $g'_E$, and $\Lambda_e \subset H_e$ of Appendix 2. We choose $x$ of 2.7 to lie in the pre-image of $p \in V$ under the composition of projection maps $E \overset{\rho}{\longrightarrow} \hat{V} \overset{\pi}{\longrightarrow} V$. Then we have that the conclusions of Lemmas 2.4, 2.6 and 2.7 hold for $\beta = \delta$ and $\alpha = \lambda\delta$ for any $\delta > 0$ (cf. 2.2 and A.2.5).

The closed connected Lie subgroups $G_{t, x} \subset G$ are as given in 2.7. We note that $G_{t_{i+1}, x} \subset G_{t, x}$ for all $1 \leq i < (m + 4)^2 - 1$. Thus a dimension argument gives us the following. (Recall that dim $M = m$ and dim $E \leq m^2 + m$; so dim $G \leq m^2 + m$ (cf. A.2.3).)

3.1.1. For some $1 \leq i < (m + 4)^2 - 1$ we have that

\[ G_{t, x} = G_{t_{i+1}, x}. \]

For each $1 \leq j \leq (m + 4)^2 - 1$, and for $i$ as in 3.1.1, we choose a subgroup $C_{i, j, x} \subset G_{t, x}$ which is maximal in $G_{t, x}$ with respect to the following properties:

3.1.2. (a) $d'_p(h\rho(x), \rho(x)) \leq t_{i, j}$ for each $h \in C_{i, j, x}$.

(b) $C_{i, j, x}$ is compact.

The existence of such a maximal subgroup of $G_{t, x}$ can be seen as follows: Since $G_{t, x}$ is a connected nilpotent Lie group (cf. 2.7) it must contain a single maximal compact subgroup $C \subset G_{t, x}$. Now if $C_1 \subset C_2 \subset C_3 \subset \ldots$ is a
sequence of compact subgroups satisfying 3.1.2(a) and (b) with \( C_l \neq C_{l+1} \) for all \( l \), then \( D = \text{closure} \left( \bigcup_{l=1}^{\infty} C_l \right) \) is a subgroup of \( C \) which satisfies Properties 3.1.2(a) and (b) and also satisfies \( \dim(C_1) < \dim(D) \). Thus to construct \( C_{i,j,x} \) we can first choose a compact subgroup \( C_1 \) of maximal dimension that satisfies 3.1.2(a) and (b); then by the preceding remarks \( C_n \) (one of the groups in the sequence \( C_1 \subset C_2 \subset C_3 \subset \ldots \)) may be chosen to be \( C_{i,j,x} \).

Note that these compact subgroups can be chosen to satisfy \( C_{i,j+1,x} \subset C_{i,j,x} \) for all \( 1 \leq j < (m+4)^2 - 1 \). We denote by \( F_{i,j,x} \subset \hat{V} \) the connected component of the fixed point set for \( C_{i,j,x} \times \hat{V} \to \hat{V} \) which is a distance less than \( 2t_{i,j} \) from \( \rho(x) \) (cf. Remark 3.1.3.1 below to see that \( F_{i,j,x} \neq \phi \)). \( F_{i,j,x} \) is a nonempty totally geodesic submanifold of \( \hat{V} \). It follows from 3.1.2 that each \( F_{i,j,x} \) is a nonempty closed subset of \( \hat{V} \) and that the equality \( C_{i,j,x} = C_{i,j+1,x} \) would be implied by the equality \( F_{i,j,x} = F_{i,j+1,x} \). On the other hand, it follows from the inclusion \( C_{i,j+1,x} \subset C_{i,j,x} \) that \( F_{i,j,x} \subset F_{i,j+1,x} \). Thus a dimension argument (over the dimension of the fixed point sets) can be used to verify the following:

3.1.3. For some \( 1 \leq j < (m+4)^2 - 1 \) we have that

\[ C_{i,j,x} = C_{i,j+1,x}. \]

Remark 3.1.3.1. We show in this remark that \( F_{i,j,x} \neq \phi \); i.e., there is a fixed point \( Y \in \hat{V} \) for the action \( C_{i,j,x} \times \hat{V} \to \hat{V} \) with \( d'_{\hat{V}}(\rho(x), Y) < 2t_{i,j} \).

Since \( \hat{V}, g'_{\hat{V}} \) is \( A' \)-regular with radius of injectivity at \( \rho(x) \) much greater than \( t_{i,j} \) (cf. 3.0 and Appendix 2), we may choose an imbedding \( f_p : B^m_{\epsilon} \to \hat{V} \) as in A.1.1 and A.1.2 with \( p = \rho(x) \) and \( \epsilon \geq t_{i,j} \). The action of \( C_{i,j,x} \times \hat{V} \to \hat{V} \) pulls back along \( f_p \) to give a smooth action \( C_{i,j,x} \times B^m_{\epsilon} \to B^m_{\epsilon} \) defined by the requirement that \( f_p(hq) = h f_p(q) \) for all \( h \in C_{i,j,x} \) and all \( q \in B^m_{\epsilon} \) (cf. 3.1.2(a)). In light of A.1.1(a) and (b), Property 3.1.2(a) implies that the pulled back action satisfies:

(i) \( |h0| < (3/2)t_{i,j} \)

for all \( h \in C_{i,j,x} \), where \( h0 \) indicates the action of \( h \) on the origin \( 0 \in B^m_{\epsilon} \).

Now if \( \hat{V} \) were a Riemannian flat space \( g_{i,j} = \delta^i_j \) everywhere in A.1.1(a) and (b) we could obtain a fixed point element for the action \( C_{i,j,x} \times B^m_{\epsilon} \to B^m_{\epsilon} \) by taking the geometric average:

(ii) \( X = \left( \int_{C_{i,j,x} \times B^m_{\epsilon}} h0 \right) / \text{Vol}(C_{i,j,x}) \)

of all the values in (i). Note that (i) would assure that \( X < (3/2)t_{i,j} \); so setting \( Y = f_p(X) \) we get the desired fixed point for the action \( C_{i,j,x} \times \hat{V} \to \hat{V} \). In the general situation (when \( \hat{V} \) is not flat) we
obtain a fixed point set for the action $C_{i,j,x} \times B^m_{\varepsilon/2} \to B^m_{\varepsilon}$ as the limit of points $X_n$, $n = 1, 2, 3, \ldots$, where:

(iii) $X_1 = X$; and $X_{n+1} = \left( \int_{C_{i,j,x}} hX_n \right) / \text{Vol}(C_{i,j,x})$ for $n \geq 1$.

Note that (i) and (iii), together with Properties A.1.1(a) and (b) and A.1.3.3.1 in Appendix 1, imply that:

(iv) $X_\infty = \lim_{n \to \infty} X_n$
exists and satisfies:

(v) $|X_\infty| < \left( \frac{7}{4} \right) t_{i,j}$ and $hX_\infty = X_\infty$

for all $h \in C_{i,j,x}$. Setting $Y = f_p(X_\infty)$ we get the desired fixed point for the action $C_{i,j,x} \times \hat{V} \to \hat{V}$. [Note that (v) above, and Properties A.1.1(a) and (b) in Appendix 1, together imply that $d'_\hat{V}(\rho(x), Y) < 2t_{i,j}$.]

For each $1 \leq k \leq (m + 4)^2 - 1$, and for $i$ and $j$ as in 3.1.1 and 3.1.3, we let $\Lambda_{i,j,k,x} \subset \Lambda$ denote the subgroup generated by all $g \in \Lambda$ which satisfy:

3.1.4. $d'_\hat{V}(\rho(x), gp(x)) < t_{i,j,k}$.

We use the remainder of this subsection to verify the following claim:

Claim 3.1.5. The following relations exist between the groups $G_{t_i,x}, C_{i,j,x}, \Lambda_{i,j,k,x}$, for $i$ and $j$ as in 3.1.1 and 3.1.3 and each $1 \leq k < (m + 4)^2 - 1$:

(a) $g(G_{t_i,x})g^{-1} = G_{t_i,x}$ for each $g \in \Lambda_{i,j,k,x}$.
(b) $g(C_{i,j,x})g^{-1} = C_{i,j,x}$ for each $g \in \Lambda_{i,j,k,x} \cup G_{t_i,x}$.
(c) Set $\Lambda'_{i,j,k,x} = \Lambda_{i,j,k,x} \cap G_{t_i,x}$. Then $\Lambda'_{i,j,k,x}$ is a normal subgroup of $\Lambda_{i,j,k,x}$ of index $\ll t_i^{-\text{dim } E}$.
(d) $\Lambda'_{i,j,1,x} = \Lambda'_{i,j,k,x}$.

Verification of Claim 3.1.5. It will suffice to verify 3.1.5(a) for any $g \in \Lambda_{i,j,k,x}$ satisfying 3.1.4. Note that any $g' \in \Gamma_{t_{i+1}} \Lambda$ satisfies:

3.1.6. (a) $d'_E(x, h_{g'}(s)x) < Bt_{i+1}$

for all $s \in [0, 1]$, where $h_{g'}$ is the one-parameter subgroup hypothesized in 2.7 for $t = t_{i+1}$, and where $B$ comes from A.2.3 and 2.7. And, since $\rho : E \to \hat{V}$ is a Riemannian submersion, 3.1.6(a) immediately implies:

3.1.6. (b) $d'_\hat{V}(\rho(x), h_{g'}(s)\rho(x)) < Bt_{i+1}$

for all $s \in [0, 1]$. Combining 3.1.4 with 3.1.6(b) we get:

3.1.6. (c) $d'_\hat{V}(\rho(x), gh_{g'}(s)g^{-1}\rho(x)) < 2t_{i,j,k} + Bt_{i+1}$
for all $s \in [0, 1]$. Finally, by combining 3.1.6(a) and (c) with the inequality $2t_{i,j,k} + Bt_{i+1} \ll t_i$ (cf. 3.0), we may deduce that:

3.1.7. (a) $d'_E(x, g(h'_g(s))g^{-1}(x)) \ll t_i$

for all $s \in [0, 1]$. Recall that $g' = h'_g(r'g')$ for some $r'g' \in [0, 1]$; so when $s = r'g'$ in 3.1.7(a) we get:

3.1.7. (b) $d'_E(x, gg'g^{-1}(x)) \ll t_i$.

We also have that:

3.1.7. (c) $gg'g^{-1} \in \Gamma$.

To see this first note that $gg'g^{-1} \in \Lambda$; and, since $g' \in H_e$ and $gH_eg^{-1} = H_e$, we also have that $gg'g^{-1} \in H_e$. Now 3.1.7(c) follows from these remarks and the equality $\Gamma = \Lambda \cap H_e$. An immediate consequence of Properties 3.1.7(b) and (c), and of the definition of $\Gamma_{t,x}$ given just prior to 2.7 above, is that:

3.1.8. (a) $gg'g^{-1} \in \Gamma_{t,x}$.

Thus, there is the one-parameter subgroup $h_{gg'g^{-1}} : \mathbb{R} \rightarrow G_{t_{i,x}}$ hypothesized in 2.7. Consider also the one-parameter subgroup $h : \mathbb{R} \rightarrow H_e$ defined by $h(s) = gh'_g(s)g^{-1}$. We note that 3.1.7(a) assures that 2.7(a) may be applied to conclude that $h$ is just a reparameterization for $h_{gg'g^{-1}}$; in particular we have that:

3.1.8. (b) $gh'_g(s)g^{-1} \in G_{t_{i,x}}$ for all $s$.

Since the one-parameter subgroups $h'_g : \mathbb{R} \rightarrow G_{t_{i,x}}, g' \in \Gamma_{t_{i,x}}$, generate $G_{t_{i+1,x}}$ (cf. 2.7(b)), it follows from 3.1.8(b) that:

3.1.8. (c) $gG_{t_{i,x}}g^{-1} \subset G_{t_{i,x}}$.

Now Property 3.1.5(a) is a consequence of Properties 3.1.1 and 3.1.8(c).

In verifying Claim 3.1.5(b), since $C_{i,j,x}$ is in the center of $G_{t_{i,x}}$, we only need to consider the case where $g \in \Lambda_{i,j,k,x}$. It also suffices to consider $g$ which satisfies 3.1.4 since $\Lambda_{i,j,k,x}$ is generated by such elements. It follows from 3.1.2 and from 3.1.4 that:

3.1.9. $d'_V(\rho(x), ghg^{-1}\rho(x)) < 2t_{i,j,k} + t_{i,j+1}$, for all $h \in C_{i,j+1,x}$.

Let $C \subset G$ denote the minimal compact subgroup which contains both $C_{i,j+1,x}$ and $g(C_{i,j+1,x})g^{-1}$. Since $C$ is in the center of $G$ it follows that any element in $C$ can be written as a product $g_1g_2$ of two elements in $C_{i,j+1,x} \cup g(C_{i,j+1,x})g^{-1}$. This last fact, together with 3.1.2(a) (as applied to $C_{i,j+1,x}$) and 3.1.9, implies that:
3.1.10. (a) \( d'_V(\rho(x), h\rho(x)) < \lambda(2t_{i,j,k} + t_{i,j+1}) \) for all \( h \in C \).

It follows from 3.0 that
\[ \lambda(2t_{i,j,k} + t_{i,j+1}) \ll t_{i,j}. \]

Combining this last inequality with 3.1.10(a) we get that:

3.1.10. (b) \( d'_V(\rho(x), h\rho(x)) < t_{i,j} \) for all \( h \in C \).

If 3.1.5(b) doesn’t hold, then \( C_{i,j+1,x} \subset C \) but \( C_{i,j+1,x} \neq C \) (cf. 3.1.3). These last properties, together with 3.1.2, 3.1.3 and 3.1.10(b), contradict the maximality property for \( C_{i,j,x} \). This completes the verification of Claim 3.1.5(b).

Now we verify Claim 3.1.5(c). It follows directly from 3.1.5(a) that \( \Lambda'_{i,j,k,x} \) is a normal subgroup of \( \Lambda_{i,j,k,x} \). To estimate its index choose a covering of the fiber \( \rho^{-1}(\rho(x)) \) by a finite number of open subsets \( X_1, \ldots, X_q \subset \rho^{-1}(\rho(x)) \) which satisfy the following properties:

3.1.11. (a) Each \( X_s \) is an open ball in \( \rho^{-1}(\rho(x)) \) of radius \( t_i/8 \). (Here the radius is measured with respect to the restriction of \( g'_E \) to \( \rho^{-1}(\rho(x)) \).)

(b) \( q \ll t_i^{-\dim E} \).

Let \( B(\rho(x), t_i/8) \subset \hat{V} \) denote the open ball in \( \hat{V} \) of radius \( t_i/8 \) centered at \( \rho(x) \in \hat{V} \), and define an open covering \( Y_1, \ldots, Y_q \) for \( \rho^{-1}(B(\rho(x), t_i/8)) \) as follows:

3.1.11. (c) Each \( Y_s \) is gotten by parallel translating the set \( X_s \) along all the geodesics in \( B(\rho(x), t_i/8) \) which start at \( \rho(x) \).

If for \( h_1, h_2 \in \Lambda_{i,j,k,x} \) we have that \( h_1(x), h_2(x) \in Y_s \) (for some \( s \in \{1, 2, \ldots, q\} \)), it follows from 3.1.11(a) and (c) that \( d'_E(x, h_2^{-1}h_1(x)) < t_i \). It follows from this last inequality, and from 2.3 and 3.0, that \( h_2^{-1}h_1 \in \Gamma = \Lambda \cap H_\epsilon \). Thus \( h_2^{-1}h_1 \in \Gamma_{t_i,x} \) (cf. the definition of \( \Gamma_{t_i,x} \) which precedes 2.7, and recall that \( \Gamma = \Lambda \cap H_\epsilon \)); so \( h_2^{-1}h_1 \in G_{t_i,x} \) by 2.7(b). It follows that there at most \( q \) distinct left cosets of \( \Lambda'_{i,j,k,x} \) in \( \Lambda_{i,j,k,x} \). This completes the verification of Claim 3.1.5(c).

Finally we will verify 3.1.5(d). First we note that:

3.1.12. \( \Gamma_{t_i,x} \subset \Lambda'_{i,j,k,x} \).

To see this we recall that \( \Gamma = \Lambda \cap H_\epsilon \), and thus (by the definition of \( \Gamma_{t_i,x} \) preceding 2.7 and the definition of \( \Lambda_{i,j,k,x} \) in 3.1.4) it follows that \( \Gamma_{t_i,x} \subset \Lambda_{i,j,k,x} \). We conclude from 2.7(b) and 3.0 that \( \Gamma_{t_i,x} \subset G_{t_i,x} \). Now these
last two inclusions, together with the definition given for \( \Lambda'_{i,j,k,x} \) in 3.1.5(c), imply 3.1.12. Next we note that it follows from 2.7(c) (with \( t = t_{i+1} \) in 2.7) and 3.1.1 that:

3.1.13. \( \text{diameter} \left( \frac{L_{t_{i,j,k,x}}}{\Gamma_{t_{i,j,k,x}}} \right) < \tau_{t_{i+1}} \)

for all \( 1 \leq k \leq (m+4)^2 - 1 \). From the definition given for \( \Lambda'_{i,j,k,x} \) in 3.1.5(c), and from 3.0 and 3.1.1, it follows that \( \Lambda'_{i,j,k,x} \) acts on the space \( L_{t_{i,j,k,x}} \); and thus there is the orbit space \( L_{t_{i,j,k,x}} / \Lambda'_{i,j,k,x} \). Now combining 3.1.12 and 3.1.13 we conclude (even when \( k = (m+4)^2 - 1 \)) that:

3.1.14. \( \text{diameter} \left( \frac{L_{t_{i,j,k,x}}}{\Lambda'_{i,j,k,x}} \right) < \tau_{t_{i+1}} \).

Setting \( k = 1 \) in 3.1.14 we see that:

3.1.15. \( \text{diameter} \left( \frac{L_{t_{i,j,1,x}}}{\Lambda'_{i,j,1,x}} \right) < \tau_{t_{i,j,k}} \), for \( 1 \leq k \leq (m+4)^2 - 1 \).

It follows immediately from 3.1.15 that there is a generating set \( S \subset \Lambda'_{i,j,1,x} \) for the group \( \Lambda'_{i,j,1,x} \) such that

\[
d'(E, gE) < 2\tau_{t_{i,j,k}}
\]

for all \( g \in S \); and, since \( \rho : E \to \hat{V} \) is a Riemannian submersion, this last inequality implies

\[
d'(\rho(x), g\rho(x)) < 2\tau_{t_{i,j,k}}
\]

for all \( g \in S \). It follows from 3.0 that

\[
2\tau_{t_{i,j,k}} \ll t_{i,j,k-1}.
\]

Combining these last two inequalities with the definition given for \( \Lambda_{i,j,k-1,x} \) in 3.1.4, and with the definitions of \( \Lambda'_{i,j,1,x} \) and \( \Lambda'_{i,j,k-1,x} \) given in 3.1.5(c), we see that:

3.1.16. \( \Lambda'_{i,j,1,x} \subset \Lambda'_{k,j,k-1,x} \) for \( 1 < k \leq (m+4)^2 - 1 \).

On the other hand it is a clear consequence of 3.1.4 and 3.1.5(c), and of the fact that \( t_{i,j,k} \leq t_{i,j,1} \) if \( k \geq 1 \) (cf. 3.0), that:

3.1.17. \( \Lambda'_{i,j,k,x} \subset \Lambda'_{i,j,1,x} \) for \( 1 \leq k \leq (m+4)^2 - 1 \).

Now 3.1.16 and 3.1.17 together imply 3.1.5(d).

3.2. **The group action** \( G_{i,j,x} \times F_{i,j,k,y} \to F_{i,j,k,y} \). Let the positive integers \( i, j \) be as in 3.1.1 and 3.1.3. We have referred to the subset \( F_{i,j,x} \subset \hat{V} \) in 3.1: It is the connected component of the fixed point set for the action \( C_{i,j,x} \times \hat{V} \to \hat{V} \) which is a distance less than \( 2t_{i,j} \) from \( \rho(x) \). Note that it follows from 3.1.2, 3.1.3, 3.1.5 that \( F_{i,j,x} \) satisfies the following properties:
3.2.1. (a) $d'_V(\rho(x), y) < 2t_{i,j+1}$, for some $y \in F_{i,j,x}$.

(b) $F_{i,j,x}$ is invariant under the action $G_{t_{i,x}} : \hat{V} \to \hat{V}$.

(c) $F_{i,j,x}$ is invariant under the action $\Lambda_{i,j,k,x} : \hat{V} \to \hat{V}$ for all $1 \leq k \leq (m+4)^2 - 1$.

Property 3.2.1(a) follows from 3.1.1-3.1.3. To verify Property 3.2.1(c) it will suffice to show that $gF_{i,j,x} = F_{i,j,x}$ for any $g \in \Lambda_{i,j,k,x}$ which satisfies 3.1.4. Note that by 3.1.5(b) we have that $gF_{i,j,x}$ is a connected component of the fixed point set for $C_{i,j,x} \times V \to \hat{V}$. Using Appendix 2, 3.0, 3.1.4, 3.2.1(a) we have that $y \in F_{i,j,x}$ and $gy \in gF_{i,j,x}$ are a distance apart less than $t_{i,j}$, and that $t_{i,j}$ is much less than the radius of injectivity for $\hat{V}$ at $y$. Thus there is a unique shortest geodesic arc connecting $y$ to $gy$, which must be in the fixed point set for $C_{i,j,x} \times V \to \hat{V}$ because $y$ and $gy$ are fixed points. It follows that $F_{i,j,x} = gF_{i,j,x}$. This completes the verification for 3.2.1(c).

The verification of 3.2.1(b) is the same as that for 3.2.1(c); note that since $G_{t_{i,x}}$ is connected it has a generating set of elements $g \in G_{t_{i,x}}$ which satisfy Property 3.1.4.

We see from 3.2.1(b) (see also 3.1.5(b)) that there is an action $G_{i,j,x} \times F_{i,j,x} \to F_{i,j,x}$ by the quotient group $G_{i,j,x} = G_{t_{i,x}}/C_{i,j,x}$; by 3.2.1(c) we see that there is the action $\Lambda_{i,j,k,x} \times F_{i,j,x} \to F_{i,j,x}$ for all $1 \leq k \leq (m+4)^2 - 1$. For each $1 \leq k \leq (m+4)^2 - 1$, we let $F_{i,j,k,y} \subset F_{i,j,x}$ denote the union of all the orbits of the action $G_{i,j,x} \times F_{i,j,x} \to F_{i,j,x}$ which intersect with $\bigcup_{g \in \Lambda_{i,j,k,x}} g(B(y; t_{i,j,k}))$ where $B(y; t_{i,j,k})$ denotes the open ball in $F_{i,j,x}$ of radius $t_{i,j,k}$ centered at $y$, and $y \in F_{i,j,x}$ comes from 3.2.1(a). Since $F_{i,j,k,y}$ is the union of some $G_{i,j,x}$-orbits, it is therefore left invariant by the action $G_{i,j,x} \times F_{i,j,x} \to F_{i,j,x}$. To see that $F_{i,j,k,y}$ is left invariant by the action $\Lambda_{i,j,k,x} \times F_{i,j,x} \to F_{i,j,x}$ we must appeal to 3.1.5(a). Thus we have the following:

3.2.2. $F_{i,j,k,y}$ is left invariant by each of the actions $G_{i,j,x} \times F_{i,j,x} \to F_{i,j,x}$ and $\Lambda_{i,j,k,x} \times F_{i,j,x} \to F_{i,j,x}$.

We use the remainder of this subsection for verifying the following two claims. In both these claims we let $i,j$ be as in 3.1.1, 3.1.3.
Claim 3.2.3. For any $2 \leq k < (m + 4)^2 - 1$ the following hold:

(a) The action $G_{i,j,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$ has local angle control equal $(\epsilon, t_{i,j,2})$.
(b) $G_{i,j,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$ is a free action.
(c) Let $G_{i,j,k,y}$ denote the foliation of $F_{i,j,k,y}$ by the orbits of this action. Then $G_{i,j,k,y}$ is a strongly Riemannian foliation which satisfies

$$K(G_{i,j,k,y}; F_{i,j,k,y}) < 10^6 e t_{i,j,2}^{-1}.$$  

(See Remarks 2.6(i) and (ii) for terminology and notation.)

For each $t > 0$ let $A^\perp(y; t)$ denote the set of all vectors $v \in T(F_{i,j,2,y})$ which are perpendicular to $G_{i,j,2,y}$ and have length $< t$. Note that the exponential map $\exp : A^\perp(y; t) \to V$ is a well-defined smooth embedding for all $t \in (0, t_{i,j,2}]$ (cf. A.2.1, A.2.2, 3.0); denote its image by $B^\perp(y; t)$.

Claim 3.2.4. For any $4 \leq k < (m + 4)^2 - 1$ we have that $B^\perp(y; t_{i,j,k-1})$ intersects each leaf of $G_{i,j,k,y}$ at most once.

Verification of Claim 3.2.3.

Note that Lemma 2.6 and 3.2.3(a) together imply Property 3.2.3(c). Towards verifying Property 3.2.3(b) we denote by $G_{i,j,x,z}$ the isotropy group for the action $G_{i,j,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$ at any point $z \in F_{i,j,k,y}$. To verify 3.2.3(b) it will suffice to show that $|G_{i,j,x,z}| = 1$. And, since $G_{i,j,x}gz = g(G_{i,j,x,z})g^{-1}$ for any $g \in G_{i,j,x} \cup \Lambda_{i,j,k,x}$ (cf. 3.1.5(a) and (b)) and

$$F_{i,j,k,y} = \bigcup_{g_1 \in G_{i,j,x}, g_2 \in \Lambda_{i,j,k,x}} g_1 g_2 B(y; t_{i,j,k}),$$

it follows that we only have to show that $|G_{i,j,x,z}| = 1$ for $z \in B(y; t_{i,j,k})$. Note that Lemma 2.5 and 3.2.3(a), together with Corollary A.2.4, imply that $|G_{i,j,x,z}| < \infty$. So the pre-image of $G_{i,j,x,z}$ under the quotient map $G_{i,x} \to G_{i,j,x}$ is a finite extension $C_{i,j,x} \subseteq C_{i,j,x}$ of the compact group $C_{i,j,x}$. Note also that $z$ is in the fixed point set for the action $C_{i,j,x} \times V \to V$, and that $d_V(p(x), z) \ll t_{i,j}$ (use 3.0, 3.2.1(a), and $z \in B(y; t_{i,j,k})$). Thus, if $|G_{i,j,x,z}| > 1$, $C_{i,j,x}$ will be a larger compact group than $C_{i,j,x}$ which also satisfies 3.1.2. This would contradict the maximality property for $C_{i,j,x}$. So we must have that $|G_{i,j,x,z}| = 1$ as desired.

We have just argued that Properties 3.2.3(b) and (c) can be deduced from Property 3.2.3(a). Thus to complete the proof of Claim 3.2.3 it will suffice to verify 3.2.3(a). Note also that $F_{i,j,k,y} \subset F_{i,j,2,y}$ for all $k \geq 2$; so it will suffice to verify 3.2.3(a) for $k = 2$. To verify 3.2.3(a) for $k = 2$ it will suffice to show that the action $G_{i,j,x} \times F_{i,j,2,y} \to F_{i,j,2,y}$ has local angle control equal $(\epsilon, t_{i,j,2})$ at every $z \in B(y, t_{i,j,2})$ since $gG_{i,j,x}g^{-1} = G_{i,j,x}$ for
all \( g \in G_{i,j,x} \cup \Lambda_{i,j,2,x} \) by 3.1.5(a) and (b), and

\[
F_{i,j,2,y} = \bigcup_{g_1 \in G_{i,j,x}, g_2 \in \Lambda_{i,j,2,x}} g_1 g_2 B(y; t_{i,j,2}).
\]

If the desired local angle control does not hold at some \( z \in B(y, t_{i,j,2}) \) then (by Lemma 2.3) there must exist a unit speed path \( f: \mathbb{R} \to F_{i,j,2,y} \) and a unit vector \( v \in T(F_{i,j,2,y}) \) which satisfy the following properties:

3.2.5.  (a) There is a one-parameter subgroup \( \phi: \mathbb{R} \to G_{i,j,x} \) such that

\[
f(t) = \phi(t)(z).
\]

(b) \( \Theta(Dg(v), P_{f[[0,r]}(v)) > \epsilon, \) where \( g = \phi(r) \) for some \( 0 < r \leq t_{i,j,2} \).

Choose \( \hat{z} \in \rho^{-1}(z) \), and let \( \hat{\phi}: \mathbb{R} \to G_{t_i,x} \) denote a one-parameter subgroup such that the composition \( \mathbb{R} \xrightarrow{\hat{\phi}} G_{t_i,x} \xrightarrow{q} G_{i,j,x} \) (\( q = \) quotient map) is equal to \( \phi \) of 3.2.5. Define a map \( \hat{f}: \mathbb{R} \to E \) by \( \hat{f}(t) = \hat{\phi}(t)(\hat{z}) \) for all \( t \in \mathbb{R} \). Then \( \hat{f} \) satisfies the following properties. (Note that 3.2.6(b) is a consequence of 3.2.6(a), and Appendix 2, and the fact that \( f \) has unit speed.)

3.2.6.  (a) \( \rho \circ \hat{f} = f \).

(b) \( \hat{f} \) has constant speed \( \geq 1 \).

Note that Properties 3.2.5(b) and 3.2.6(a) together imply that:

3.2.7.  (a) \( d_E'(g(\hat{z}), \hat{f}(u)) \geq \epsilon/8, \) for some \( u \in (0, t_{i,j,2}) \) and all \( g \in C_{i,j,x} \).

Combining 3.2.6(b) with 3.2.7(a) (with \( g = \text{id} \) in 3.2.7(a)) we get that:

3.2.7.  (b) \( \hat{f} \) has constant speed \( \geq \epsilon/8 t_{i,j,2} \).

Now, combining 3.2.6 and 3.2.7(b), together with the relation:

3.2.7.  (c) \( t_{i,j,2} \ll \epsilon(t_{i,j,1})^{\dim E + 2} \)

(which follows from 3.0) and the fact that \( \rho \circ \hat{f} \) has unit speed, imply that the following is true:

3.2.8.  (a) There is \( s \in (u, t_{i,j,1}/2) \) so that \( d_E'(\hat{f}(0), \hat{f}(s)) \ll t_{i,j,1} \).

[In more detail we verify 3.2.8(a) as follows: Let \( B(z, t_{i,j,1}) \) denote the open ball in \( \hat{V} \) of radius \( t_{i,j,1} \) centered at \( z \). We may choose a covering \( \{ B_k : 1 \leq k \leq \kappa \} \) for \( \rho^{-1}(B(z, t_{i,j,1})) \) by open balls \( B_k \) in \( E \) satisfying:

(i) \( \text{radius}(B_k) \ll t_{i,j,1} \),

(ii) \( \kappa \ll t_{i,j,1}^{-\dim E} \).]
On the other hand it follows from 3.2.6, and from the fact that \( f \) has unit speed, that:

(iii) \( \hat{f}(a) \in \bigcup_{k=1}^{\kappa} B_k \) for all \( a \in [0, t_{i,j,1}] \).

Finally, by combining (ii) and (iii) above with Property 3.2.7(c), we deduce that there are numbers \( a, c \in [0, \frac{t_{i,j,1}}{2}] \) such that:

(iv) \( a + t_{i,j,2} < c \) and \( \hat{f}(a), \hat{f}(c) \in B_k \) for some \( 1 \leq k \leq \kappa \).

Now Property 3.2.8(a) is an immediate consequence of (i) and (iv), and of the inequality \( u < t_{i,j,2} \) (cf. 3.2.7(a)), provided the number \( s \) of 3.2.8(a) is defined by \( s = c - a \).

Note that Appendix 2, 2.2, and 3.1.1 assure that all the hypotheses for 2.7(d) are satisfied by the action \( G_{t_i,x} \times E \rightarrow E \). Thus we may apply 2.7(d), in conjunction with 3.2.8(a), to deduce that:

\[ 3.2.8. \quad (b) \quad d_{G_{t_i,x}}(\hat{\phi}(0), \hat{\phi}(s)) < t_{i,j,1}. \]

Now we can apply Lemma 2.8 with \( \phi_1, t, s, G \) of 2.8 set equal to \( \hat{\phi}, t_{i,j,1}, s \) (of 3.2.8), \( G_{t_i,x} \), and with \( g_G \) of 2.8 set equal to the Riemannian metric on \( G_{t_i,x} \) induced from \( g_E' \) by identifying \( G_{t_i,x} \) with the subset \( G_{t_i,x}(\hat{z}) = \{ g(\hat{z}) : g \in G_{t_i,x} \} \) of \( E \) via the map \( g \rightarrow g(\hat{z}) \). We choose the local angle control numbers \( (\alpha, \beta) \) (assumed in the hypothesis of 2.8) to be equal \( (\lambda \delta_1, \delta_1) \) where \( \lambda > 1 \) comes from Remark 2.2 and A.2.5; then it follows from 2.2 and A.2.5 that the left action \( G \times G \rightarrow G \) has local angle control equal \( (\alpha, \beta) \) as required in 2.8. Note that 3.2.8(b), and the preceding remarks, imply that these choices for \( \phi_1, t, s, G, g_G \) satisfy Property 2.8(a). Thus we may apply Lemma 2.8 to conclude that there is another one-parameter group \( \phi_2 : \mathbb{R} \rightarrow G_{t_i,x} \) which satisfies the following properties:

\[ 3.2.9. \quad (a) \quad \phi_2(s)(\hat{z}) = \hat{z}. \]

(b) \( d'_{E}(\phi_2(u)(\hat{z}), \phi_1(u)(\hat{z})) < (u/s)^{4 \dim E} t_{i,j,1} \) for all \( u \in [0, s] \).

We can deduce from 3.2.9(b), 3.2.7(a), and from the inequality \( t_{i,j,1} < \epsilon/8 \) (cf. 3.0), the following additional property for \( \phi_2 \):

\[ 3.2.9. \quad (c) \quad \text{Image(} \phi_2 \text{)} \text{ is not contained in } C_{i,j,x}. \]

Note that 3.2.9(a), together with the fact that \( G_{t_i,x} \times E \rightarrow E \) is a free action [the action \( H \times \hat{V} \rightarrow \hat{V} \) is effective (cf. Appendix 2) and thus the action \( H \times E \rightarrow E \) is free] implies that \( \text{Image}( \phi_2 ) \) is a compact subgroup of \( G_{t_i,x} \). Let \( C \) denote the minimal compact subgroup of \( G_{t_i,x} \) which contains \( C_{i,j+1,x} \cup \text{Image}( \phi_2 ) \). The following properties hold:

\[ 3.2.10. \quad (a) \quad C_{i,j,x} \subset C \subset G_{t_i,x} \text{ and } C_{i,j,x} \neq C. \]
(b) \( d_V'(\rho(x), h\rho(x)) < t_{i,j} \), for all \( h \in C \).

Note that 3.2.10(a) is a consequence of 3.1.3, 3.2.9(c) and of the definition of \( C \). Also note that each element \( h \in C \) can be written as product \( h = g_1g_2 \) of two elements in \( C_{i,j+1,x} \cup \text{Image}(\phi_2) \) (because \( G_{t_i,x} \) is connected and nilpotent). Thus Property 3.2.10(b) follows from 3.0, 3.1.2 (as applied to \( C_{i,j+1,x} \)), and from the following property:

**3.2.11.** \( d_V'(\rho(x), g\rho(x)) < 8\dim E(t_{i,j,1}) \) for all \( g \in \text{Image}(\phi_2) \).

To verify 3.2.11 we note first that \( \rho : E \to \hat{V} \) doesn’t increase distances; thus

\[
d_V'(\phi_2(u)z, \phi_1(u)z) < (u/s)4\dim E(t_{i,j,1}), \quad \text{for all } u \in (0, s),
\]

follows from 3.2.9(b). Now \( \phi_1(u)z \) has unit speed in the variable \( u \) (because it coincides with the path \( f(u) \)), and \( s \in (0, t_{i,j,1}) \) by 3.2.8(a), so

\[
d_V'(z, \phi_1(u)z) < t_{i,j,1} \quad \text{for all } u \in (0, s).
\]

On the other hand we deduce from 3.0, 3.2.1(a), and from the fact that \( z \in B(y, t_{i,j,2}) \), that

\[
d_V'(\rho(x), z) < t_{i,j,1}.
\]

Now Property 3.2.11 follows directly from 3.2.9(a), and from the three preceding inequalities.

Note that 3.2.10 contradicts the maximality of the compact subgroup \( C_{i,j,x} \). This contradiction can be traced back to our assumption that the action \( G_{i,j,x} \times F_{i,j,k,y} \to F_{i,j,k,y} \) does not have local control equal \((\epsilon, t_{i,j,2})\) at \( z \in B(y, t_{i,j,2}) \) (cf. 3.2.5).

This completes the verification of Claim 3.2.3.

**Verification of Claim 3.2.4.**

We begin by verifying the following property:

**3.2.12.** \( \text{diameter} (L/A'_{i,j,k,x}) < 2\tau t_{i+1}, \) for each leaf \( L \in \mathcal{G}_{i,j,k,y} \).

Here \( \tau > 0 \) comes from 2.7(c), and the diameter is computed with respect to the Riemannian metric which \( L \subset \hat{V} \) inherits from \( (\hat{V}, g_{\hat{V}}') \).

It will suffice to consider only those leaves \( L \in \mathcal{G}_{i,j,k,x} \) in 3.2.12 which intersect with the subset \( B(y; t_{i,j,k}) \) of \( F_{i,j,x} \) (cf. 3.1.5, and review the definition of \( F_{i,j,k,y} \)); note that this condition on \( L \) is equivalent to:

**3.2.13.** (a) \( d_{\hat{V}}'(y, L) < t_{i,j,k} \),

since \( F_{i,j,x} \) is a path component of the fixed point set of \( C_{i,j,x} \times \hat{V} \to \hat{V} \) and the injectivity radius of \( \hat{V} \) at \( y \) is \( \gg t_{i,j,k} \) (cf. 3.0 and Appendix 2). Let \( \mathcal{G}_{i,x} \)
denote the foliation of $E$ by the orbits of the group action $G_{t_i,x} \times E \to E$ (recall that this action is free). Note that the pre-image of $L$ under the composition map $E \xrightarrow{\rho} \hat{V}$ is a union of leaves in $G_{i,x}$. It follows from Appendix 2, 3.0, 3.2.1(a), and from 3.2.13(a) that there is a leaf $\hat{L} \in G_{i,x}$ such that $\rho(\hat{L}) = L$ and:

**3.2.13.** (b) $d'_{\hat{L}}(x, \hat{L}) < 2t_{i,j,k}$.

Remark 2.2 and A.2.5 assure us that we may apply Lemma 2.6, with the group action $G \times N \to N$ and the numbers $\alpha, \beta$ of 2.6 taken to be the group action $G_{t_i,x} \times E \to E$ and the numbers $\lambda, t$ (for any small $t > 0$ and $\lambda$ as in 2.2 and A.2.5), to conclude that:

**3.2.13.** (c) $K(G_{i,x}; E) < 10^6 \lambda$.

Now 3.2.13(b) and (c) and 3.1.14 together imply that:

**3.2.13.** (d) $\text{diameter}(\hat{L}/\Lambda'_{i,j,k,x}) < 2\pi t_{i+1}$.

[When deriving 3.2.13(d) see also 3.0, 3.1.1, and recall that the actions of $G_{t_i}$ and $\Lambda_{i,j,k,x}$ on $E$ are free.] Since $\rho : E \to \hat{V}$ is a Riemannian submersion which maps $\hat{L}$ onto $L$, and which commutes the action of $\Lambda'_{i,j,k,x}$ on $E$ and on $\hat{V}$, we may deduce 3.2.12 from 3.2.13(d).

It follows from A.2.2 and from 3.0 that the radius of injectivity for $F_{i,j,x}$ is greater than $t_{i,j,2}$ at $y \in F_{i,j,x}$. This fact, together with the curvature restrictions placed on $\mathcal{G}_{i,j,2,y}$ by 3.2.3, allow us to define a “local projection”

$$\pi : X(y; t_{i,j,2}) \to A^+(y; t_{i,j,2})$$

from an open subset $X(y; t_{i,j,2}) \subset F_{i,j,2,y}$ as follows: Recall that $B^+(y; t)$ and $A^+(y; t)$ were defined just prior to 3.2.4 above for any $t \in (0, t_{i,j,2})$; $A^+(y; t)$ is a subset of $T(F_{i,j,2,y})$ and $B^+(y; t) \subset F_{i,j,2}$ is the diffeomorphic image of $A^+(y; t)$ under the exponential map $\exp : A^+(y; t) \to F_{i,j,2}$. For each $q \in B^+(y; t)$ let $L_q$ denote the leaf of $\mathcal{G}_{i,j,2,y}$ containing $q$ and let $X(y; t)_q$ denote the open ball in $L_q$ of radius $t$ centered at $q$. Set

$$X(y; t) = \bigcup_{q \in B^+(y; t)} X(y; t)_q$$

for each $t \in (0, t_{i,j,2})$; and define $\pi : X(y; t_{i,j,2}) \to A^+(y; t_{i,j,2})$ by sending $X(y; t_{i,j,2})_q$ to $\exp^{-1}(q)$ for each $q \in B^+(y; t_{i,j,2})$. Then $\pi$ is a smooth fiber bundle projection having the open balls $\{X(y; t_{i,j,2})_q : q \in B^+(y; t)\}$ for fibers. Note that for each $t \in (0, t_{i,j,2})$ the restricted map $\pi : X(y, t) \to A^+(y; t)$ is a smooth bundle projection having the open balls $\{X(y; t)_q : q \in B^+(y; t)\}$ for fibers.
If Claim 3.2.4 does not hold then we may deduce from 3.2.12 that there is \( g \in \Lambda'_{t_{i,j,k},x} \) which for some \( z \in X(y; t_{i,j,k-1}) \) satisfies:

3.2.14. (a) \( d'_{F_{i,j,k}}(z, g(z)) < 4t_{i,j,k-1} \), and thus \( g(z) \in X(y; t_{i,j,2}) \);
(b) \( \pi(z) \neq \pi(g(z)) \).

We will complete the proof for 3.2.4 by deriving a contradiction to 3.2.14(b).

We can deduce from 3.0, 3.2.14(a), and from the definition of the \( \pi \), that for each positive integer \( s \leq t_{i,j,k-2}^{-\dim E} \) left multiplication by \( g^s \) is a smooth embedding

\[
g^s : X(y; t_{i,j,2}/8) \to X(y; t_{i,j,2})
\]

which maps each fiber for \( \pi \mid X(y; t_{i,j,2}/8) \) into a fiber for \( \pi \). Thus each \( g^s \) induces a smooth embedding \( h_s : A^+(y; t_{i,j,2}/8) \to A^+(y; t_{i,j,2}) \). We will verify the following properties for \( \{ h_s : 1 \leq s \leq t_{i,j,k-2}^{-\dim E} \} \):

3.2.15. (a) \( |h_s(0)| \ll t_{i,j,k-2} \).
(b) There is linear isometry \( I_s : T^\perp(G_{i,j,2,y}) \to T^\perp(G_{i,j,2,y}) \) such that \( |Dh_s - I_s| < \mathcal{O}(\epsilon) \). (Here \( T^\perp(G_{i,j,2,y}) \) denotes the orthogonal complement for \( T_{i,j,2,y} \) in \( TF_{i,j,2,y} \).)
(c) If \( h_s \mid A^+(y; t_{i,j,2}/8) \) has a fixed point then \( h_s \) is the inclusion map.
(d) Each composition \( h^*_q \) is well-defined and equal to \( h_s \) on all of \( A^+(y; t_{i,j,2}/8^2) \).

Properties 3.2.15(a) and (b) are a consequence of 3.0 and 3.2.14(a), and of the curvature condition and the “strongly Riemannian” condition placed on the foliation of \( X(y; t_{i,j,2}) \) by the fibers of \( \pi \) by Claim 3.2.3 (see also Appendix 1). And Property 3.2.15(d) is a consequence of 3.2.15(a) and (b). Towards verifying Property 3.2.15(c) we denote by \( q \in A^+(y; t_{i,j,2}/8^2) \) a fixed point for \( h_s \); we have then that

\[
g^s(\pi^{-1}(q) \cap X(y; t_{i,j,2}/8^2)) \subset \pi^{-1}(q);
\]

and then by appealing to 3.0 and 3.2.14(a) this last assertion can be improved to

\[
g^s(\pi^{-1}(q) \cap X(y; t_{i,j,2}/8^2)) \subset \pi^{-1}(q) \cap X(y; t_{i,j,2}/8).
\]

It follows from this last inclusion, and from the fact that \( G_{i,j,x} \times F_{i,j,2,y} \to F_{i,j,2,y} \) is a free action by isometries (cf. 3.2.3(b)), that there is a path \( f : [0, 1] \to G_{i,j,x} \) which satisfies:

3.2.16. (a) \( f(0) = \text{identity and } f(1) = g^s \);
(b) \( f(t)(\pi^{-1}(q) \cap X(y; t_{i,j,2}/8^2)) \subset \pi^{-1}(q) \cap X(y; t_{i,j,2}/4) \) for all \( t \in [0, 1] \).
Now it follows from 3.2.16(b), and from the curvature condition placed on the foliation of $X(y; t_{i,j,k-2})$ by the fibers of $\pi$ in Claim 3.2.3, that:

**3.2.16.** (c) $f(t)(\pi^{-1}(u) \cap X(y; t_{i,j,2}/8^2)) \subset \pi^{-1}(u)$ for all $u \in A^\perp(y; t_{i,j,2}/8^2)$ and all $t \in [0, 1]$.

Property 3.2.15(c) is a consequence of 3.2.16(a) and (c).

A contradiction to Property 3.2.14(b) is contained in the following statement:

**3.2.17.** All the maps $\{h_s : 1 \leq s \leq t_{i,j,k-2}^{\dim E}\}$ have a common fixed point in $A^\perp(y; t_{i,j,2}/8^2)$.

In fact 3.2.17 and 3.2.15(c) together imply that $h_1$ is the inclusion map; this is equivalent to the equalities $\pi(z) = \pi(g(z))$ for all $z \in X(y; t_{i,j,2}/8^2)$, which contradicts 3.2.14(b).

Thus to complete the proof of Claim 3.2.4 it remains only to verify 3.2.17. First we will verify the following weakened version of 3.2.17:

**3.2.18.** For some positive integer $r < t_{i,j,k-2}^{\dim E}$ we have that $h_r(0) = 0$.

Towards this end we first note it follows from 3.2.14(a), 3.0, 3.2.1(a) that:

**3.2.19.** (a) $d'_{E}(\rho(x), g^s(\rho(x))) < 10st_{i,j,k-1}$ for all $s \in \{1, 2, \ldots, t_{i,j,k-2}^{\dim E}\}$, where $x \in E$ was selected at the outset of 3.1. Now it follows from 3.2.19(a) and 3.0 that there are two distinct positive integers $r_1 < r_2 \leq t_{i,j,k-2}^{\dim E}$ such that:

**3.2.19.** (b) $d'_{E}(g^{r_1}(x), g^{r_2}(x)) \ll t_{i,j,k-2}$.

[We use a covering argument to verify 3.2.19, similar to the argument of 3.1.11 used above in the proof of Claim 3.1.5(c).] Now, if we set $r = r_2 - r_1$, we have (by 3.2.19(b)) that:

**3.2.19.** (c) $d'_{E}(x, g^r(x)) \ll t_{i,j,k-2}$.

Combining this last property with Lemma 2.7(d) (where we set the $N, G, \Gamma, t$ of 2.7 equal to $E, Gt_{i,x}, \Lambda'_{i,j,k,x}$, maximum $\{d'_{E}(x, g^r(x)), t_{i,j,k}\}$) we can deduce the following:

**3.2.19.** (d) There is a one-parameter subgroup $f : \mathbb{R} \to Gt_{i,x}$ such that $g^r \in f([0,1])$ and such that $d'_{E}(G_{t_{i,x}}(x, f(t)(x)) \ll t_{i,j,k-2}$ for all $t \in [0, 1]$.

We can combine 3.2.19(d) with 3.2.1(a), and with the fact that $\rho : E \to \hat{V}$ is a Riemannian submersion, to conclude:
3.2.19. (e) \( d'_{L_y}(y, f(t)(y)) \ll t_{i,j,k-2} \) for all \( t \in [0,1] \).

Now this last inequality, together with the definition of \( X(y;t_q) \) as the open ball in \( L_y \) of radius \( t \) and center \( y \), implies that all the points \( \{ f(t)(y) : t \in [0,1] \} \) are contained in the fiber \( \pi^{-1}(0) \). Since \( g^r \in f([0,1]) \) it follows that \( g^r(y) \in \pi^{-1}(0) \); from which we immediately deduce 3.2.18.

Now to complete the verification for 3.2.17 we note that it follows from 3.0, 3.2.18, and 3.2.15(a)-(d), that there is a cyclic group action \( \psi : \mathbb{Z}_r \times W \to W \), where

\[
A^-(y; t_{i,j,2}/s^2) \subset W \subset A^-(y; t_{i,j,2}/s^2)
\]

and

\[
\psi(s, w) = h_s(w) \quad \text{for} \quad 1 \leq s \leq r.
\]

By appealing to 3.2.15(a) and (b) we can obtain a fixed point for the action \( \psi : \mathbb{Z}_r \times W \to W \) as the limit of a sequence of points \( z_1, z_2, z_3, \ldots \) defined inductively by

\[
z_1 = 0,
\]

\[
z_{n+1} = r^{-1} \sum_{s=1}^{r} \psi(s, z_n).
\]

We note that limit \( z_n \) is the desired common fixed point of 3.2.17.

This completes the verification for Claim 3.2.4.

3.3. The projection \( \pi_{i,j,k,y} : F_{i,j,k,y} \to B_{i,j,k,y} \) for any \( 4 \leq k < (m + 4)^2 - 1 \). Set

\[
\Lambda_{i,j,k,x} = \Lambda_{i,j,k,x}/\Lambda'_{i,j,k,x},
\]

\[
\nabla = \hat{\nabla}/\Lambda'_{i,j,k,x},
\]

\[
F_{i,j,k,y} = F_{i,j,k,y}/\Lambda'_{i,j,k,x},
\]

\[
G_{i,j,k,y} = G_{i,j,k,y}/\Lambda'_{i,j,k,x},
\]

\[
\overline{y} = q_{i,j,k,y}(y),
\]

where \( q_{i,j,k,y} : F_{i,j,k,y} \to F_{i,j,k,y} \) is the quotient map and \( y \) comes from 3.2.1(a). We will verify that the following properties hold:

3.3.1. For each \( 4 \leq k < (m + 4)^2 - 1 \) we have:

(a) \( \overline{G}_{i,j,k,y} \) is a strongly Riemannian foliation which satisfies

\[
K(\overline{G}_{i,j,k,y}; F_{i,j,k,y}) < 10^6 \epsilon_{i,j}^{-1}.
\]

(b) diameter \( (L) < 2 \pi t_{i,j,k-1} \), for each leaf \( L \in \overline{G}_{i,j,k,y} \).

(c) diameter \( (F_{i,j,k,y}) \ll t_{i,j,k-l,1} \), where \( l = (m + 4)^2 - 1 \).

(d) \( F_{i,j,k+1,y} \subset F_{i,j,k,y} \).
48 F.T. FARRELL AND L.E. JONES

\[ G_{i,j,k} | F_{i,j,k+1,y} = G_{i,j,k+1,y} \]

Property 3.3.1(a) is a direct consequence of 3.2.3(c). Property 3.3.1(b) is a direct consequence of 3.2.12. Properties 3.3.1(d) and (e) follow from 3.1.1 and 3.1.5(d), and from the fact that \( F_{i,j,k+1,y} \subset F_{i,j,k,y} \). Now we will verify 3.3.1(c). Since \( \Lambda'_{i,j,k,x} \) is a normal subgroup of \( \Lambda_{i,j,k,x} \) (cf. 3.1.5(c)) it follows that \( \Lambda_{i,j,k,x} \) defined above is a group. We may choose, by 3.0, 3.1.4, 3.1.5(c), 3.2.1(a), a subset of elements \( \{g_1, g_2, \ldots, g_\lambda\} \subset \Lambda_{i,j,k,x} \) which satisfy the following properties:

3.3.2. (a) \( \{g_1, \ldots, g_\lambda\} \) maps onto \( \Lambda_{i,j,k,x} \) under the quotient map \( \Lambda_{i,j,k,x} \rightarrow \Lambda_{i,j,k,x} \).

(b) \( d_{V}(y, g_i(y)) < (t_i^{-2 \dim E})(2t_{i,j,k}) \).

Let \( U_{i,j,k,y} \) denote the union of all leaves in \( G_{i,j,k,y} \) which intersect the ball \( B(y; t_{i,j,k}) \) (cf. 3.2); set

\[ \overline{U}_{i,j,k,y} = U_{i,j,k,y} / \Lambda'_{i,j,k,x} \]

We note it follows from 3.3.1(b), and from the fact that \( B(y; t_{i,j,k}) \) has radius equal \( t_{i,j,k} \), that the following holds:

3.3.2. (c) diameter(\( U_{i,j,k,y} \)) < \( (4 \tau + 2)t_{i,j,k} \).

Since we have that \( F_{i,j,k,y} = \bigcup_{g \in \Lambda_{i,j,k,x}} g(U_{i,j,k,y}) \) (cf. 3.2), we also have that:

3.3.2. (d) \( F_{i,j,k,y} = \bigcup_{t=1}^{\lambda} g_t(\overline{U}_{i,j,k,y}) \).

Now we deduce 3.3.1(c) from 3.3.2(b)-(d) and from 3.0. This completes the verification for 3.3.1.

Now we can use 3.3.1 to aid in the definition of the projections \( \pi_{i,j,k,x} : F_{i,j,k,y} \rightarrow B_{i,j,k,y} \) for any \( 4 \leq k < (m+4)^2 - 1 \). Let \( W \) denote the space of all \( v \in T(F_{i,j,k,y}) \) which are perpendicular to \( G_{i,j,k,y} \), and let \( \exp : W \rightarrow F_{i,j,x} \) and \( f_W : W \rightarrow \mathbb{R}^\beta \) denote the exponential map and a fixed linear isometry respectively, where \( F_{i,j,x} \) is defined by

\[ F_{i,j,x} = F_{i,j,x} / \Lambda'_{i,j,k,x} \]

and where \( \beta = \dim W \) (cf. 3.2.1(c) and 3.1.5(d)). Let \( B(2t_i) \subset W \) denote the ball of radius \( 2t_i \) centered at the origin of \( W \). Since the injectivity radius of \( F_{i,j,x} \) is much greater than \( 2t_i \) at every one of its points (cf. A.2.2, 3.0) it follows from 3.3.1(a)-(c) and from 3.2.4 that \( \exp : B(2t_i) \rightarrow F_{i,j,x} \) maps \( B(2t_{i,j,3},1) \) diffeomorphically onto a smooth submanifold \( B(2t_{i,j,3},1) \) of
$F_{i,j,x}$ which intersects each leaf $L \in \mathcal{G}_{i,j,k,x}$ in exactly one point (see 3.0, and recall that $\Lambda'_{i,j,k,x}$ leaves each leaf of $\mathcal{G}_{i,j,k,x}$ invariant). Define a map $\pi_{i,j,k,x} : F_{i,j,k,x} \to B_{i,j,k,x}$ by:

3.3.3. (a) $\pi_{i,j,k,x}(L) = f_{\Lambda} \circ \exp^{-1}(L \cap B(2t_{i,j,3,1}))$ for $L \in \mathcal{G}_{i,j,k,x}$.
(b) $B_{i,j,k,x} = \text{Image}(\pi_{i,j,k,x})$.

There is also a canonical cross section $s_{i,j,k,x} : B_{i,j,k,x} \to F_{i,j,k,x}$ defined by:

3.3.3. (c) $s_{i,j,k,x}(v) = \exp \circ f_{\Lambda}^{-1}(v)$ for $v \in B_{i,j,k,x}$.

Note it follows from 3.3.1(d) and (e) that $W$ (in the preceding construction) is independent of $k$. Thus we concluded (by 3.3.1(d) and (e)) that the following relations hold:

3.3.4. (a) $\pi_{i,j,k,x} | F_{i,j,k+1,x} = \pi_{i,j,k+1,x}$.
(b) $B_{i,j,k+1,x} \subset B_{i,j,k,x}$.
(c) $s_{i,j,k,x} | B_{i,j,k+1,x} = s_{i,j,k+1,x}$.

Also, since $F_{i,j,k,x}$ contains the open ball $B(y; t_{i,j,k})$, it follows that $F_{i,j,k,x}$ contains the subset $\mathcal{B}(t_{i,j,k})$, where $\mathcal{B}(t_{i,j,k}) = \exp(B(t_{i,j,k}))$ and where $B(t_{i,j,k})$ is the open ball of radius $t_{i,j,k}$ centered at the origin of $W$. The following property is a reinterpretation of the containment $\mathcal{B}(t_{i,j,k}) \subset F_{i,j,k,x}$ and of Property 3.3.1(c):

3.3.5. $B_{i,j,k,x}$ is an open subset of $\mathbb{R}^\beta$ which contains the open ball of radius $t_{i,j,k}$ centered at the origin of $\mathbb{R}^\beta$; moreover $\text{diameter}(B_{i,j,k,x}) \ll t_{i,j,k} - 1$.

We note that if $\overline{y} \in F_{i,j,k,x}$ were replaced by any other point $\overline{y}' \in F_{i,j,k,x}$ in the preceding construction for the maps $\pi_{i,j,k,x}$ and $s_{i,j,k,x}$ we would obtain a new projection $\pi'_{i,j,k,x} : F_{i,j,k,x} \to B'_{i,j,k,x}$ and a new cross section $s'_{i,j,k,x} : B'_{i,j,k,x} \to F_{i,j,k,x}$. We will verify the following properties:

3.3.6. Set $s = \pi_{i,j,k,x} \circ s'_{i,j,k,x}$; note that $s$ maps the open subset $B'_{i,j,k,x} \subset \mathbb{R}^\beta$ diffeomorphically onto the open subset $B_{i,j,k,x} \subset \mathbb{R}^\beta$. For each $z \in B'_{i,j,k,x}$ there is a linear isometry $L_z : \mathbb{R}^\beta \to \mathbb{R}^\beta$ such that $s$ and $L_z$ satisfy:
(a) $\|DS|z - L_z\| \ll \epsilon$;
(b) $\|D^2s|z\| < O(\epsilon t_{i,j,2}^{-1})$.

Towards this end we let $W'$ denote all $v \in T(F_{i,j,k,x})\overline{y}$ which are perpendicular to $\mathcal{G}_{i,j,k,x}$, and let $f_{W'} : W' \to \mathbb{R}^\beta$ denote a given linear isometry; we use $W', f_{W'}$ in the preceding construction (in place of $W, f_W$) to obtain
\[ \pi'_{i,j,k,y} \text{ and } s'_{i,j,k,y} \text{ (in place of } \pi_{i,j,k,y} \text{ and } s_{i,j,k,y}). \]

For any given \( z \in B'_{i,j,k,y} \) we set \( z_1 = s'_{i,j,k,y}(z) \) and set \( z_2 = \pi'_{i,j,k,y}(z) \cap \text{Image}(s_{i,j,k,y}) \); let \( W_i \) denote all vectors \( v \in T(F_{i,j,k,y})z_i \) which are perpendicular to \( G_{i,j,k,y} \); and choose linear isometries \( f_{W_i} : W_i \to R^\beta \). For sufficiently small neighborhoods \( U', U_1, U_2 \) of \( z, s(z), 0, 0 \) in \( R^\beta \) there are diffeomorphisms

\[ g_3 : U' \to U_1 \text{ and } g_2 : U_1 \to U_2 \text{ and } g_1 : U_2 \to U \]

defined as follows: Define \( g_1 \) to be the composition of three maps \( g_1 = g_{1,1} \circ g_{1,2} \circ g_{1,3} \). Here \( g_{1,1} = \exp(f_{W}^{-1} | U) \); \( g_{1,3} = \exp(f_{W_2}^{-1} | U_2) \); and 
\( g_{1,2} : \text{Image}(g_{1,3}) \to \text{Image}(g_{1,1}^{-1}) \) is the holonomy map for \( G_{i,j,k,y} \). Define \( g_2 \) to be the composition of three maps \( g_2 = g_{2,1} \circ g_{2,2} \circ g_{2,3} \) where \( g_{2,1} = g_{1,3}^{-1} \); \( g_{2,3} = \exp(f_{W}^{-1} | U_1) \); and \( g_{2,2} : \text{Image}(g_{2,3}) \to \text{Image}(g_{1,3}) \) is the holonomy map for \( G_{i,j,k,y} \). Define \( g_3 \) to be the composition of three maps \( g_3 = g_{3,1} \circ g_{3,2} \circ g_{3,3} \) where \( g_{3,1} = g_{2,3}^{-1} \); \( g_{3,3} = \exp(f_{W}^{-1} | U') \); and \( g_{3,2} : \text{Image}(g_{3,3}) \to \text{Image}(g_{2,3}) \) is just the holonomy map for \( G_{i,j,k,y} \). Note that:

3.3.7. (a) \( s | U' = g_1 \circ g_2 \circ g_3 \).

It follows from the fact that \( G_{i,j,k,y} \) is “strongly Riemannian” (cf. 3.3.1(a)) that:

3.3.7. (b) \( g_2 : U_1 \to U_2 \) is an isometry.

On the other hand we will show that there are linear isometries \( I_1, I_3 : R^\beta \to R^\beta \) which satisfy the following claim in which \( 1 = 0 \) and \( 3 = z \):

Claim 3.3.8. (a) \( \|D^1(g_b - I_b) \|_0 \| < \epsilon \) for \( b = 1, 3 \).

(b) \( \|D^2(g_b - I_b) \|_0 < O(\epsilon) t_{i,j,2}^{-1} \) for \( b = 1, 3 \).

Now Properties 3.3.6(a) and (b) are a consequence of Properties 3.3.7(a) and (b) and 3.3.8(a) and (b).

Verification of Claim 3.3.8.

We will verify the claim for \( b = 1 \); the same proof works for \( b = 3 \).

First we remark that \( F_{i,j,k,y} \) is an open subset of \( F_{i,j,x} \), and that \( F_{i,j,x} \) is (locally) a totally geodesic subset of \( V \) with respect to the Riemannian metric \( g'_{V} \). Thus, if \( h \) denotes the Riemannian metric inherited by \( F_{i,j,k,y} \) from \( g'_{V} \), we see (cf. Appendix 2) that \( F_{i,j,k,y}, h \) is an \( A' \)-regular Riemannian manifold, where \( A' \) comes from A.2.2. Note that the quotient \( \tilde{h} \) of the metric \( h \) by the action \( \Lambda'_{i,j,k,x} \times F_{i,j,k,y} \to F_{i,j,k,y} \) is also an \( A' \)-regular metric for \( F_{i,j,k,y} \). Thus, for sufficiently small \( \epsilon > 0 \), there is an imbedding \( f_p : B^m_{\epsilon} \to F_{i,j,k,y} \) which satisfies the conclusions of A.1.1 and A.1.2 (for \( p = z_2 \) and \( m = \dim F_{i,j,k,y} \).
Define a foliation \( \hat{G}_{i,j,k,y} \) on \( B_m \) by
\[
\hat{G}_{i,j,k,y} = f_p^{-1}(\hat{G}_{i,j,k,y}).
\]
Let \( \alpha \) denote the dimension of the leaves of \( \hat{G}_{i,j,k,y} \); note that \( \beta = m - \alpha \).

Without loss of generality we may assume that the first \( \alpha \)-coordinates of \( R^m \) are tangent to \( \hat{G}_{i,j,k,y} \) at the origin. Thus there is a smooth function \( f : R^m \rightarrow R^\beta \) so that the following properties hold near the origin of \( R^m \).

In what follows we identify \( R^m \) with \( R^\alpha \times R^\beta \), and denote a variable point in \( R^m \) by \((u,v)\).

3.3.8.1. (a) \( f(0,0) = 0 \) and \( D_u f(0,0) = 0 \); \( f(0,v) = v \) for all \( v \in R^\beta \).

(b) For each \( v \in R^\beta \) define \( f_v : R^\alpha \rightarrow R^\beta \) by \( f_v(x) = f(x,v) \). The leaf of \( \hat{G}_{i,j,k,y} \) thru the point \((0,v)\) coincides with the graph of \( f_v \).

We note that the curvature bounds placed on \( G_{i,j,k,y} \) by 3.3.1(a) lead (thru A.1.4) to curvature bounds for \( \hat{G}_{i,j,k,y} \), which can be expressed in terms of the function \( f(u,v) \) as follows: Let \( u = (u_1, u_2, \ldots, u_\alpha) \) and \( v = (v_1, v_2, \ldots, v_\beta) \) denote the standard coordinates for the variables \( u \) and \( v \) respectively.

3.3.8.1. (c) Near the origin of \( R^\alpha \times R^\beta \) we have that \( |\partial^2 f / \partial u_{a_1} \partial u_{a_2}| < O(\epsilon) t_{i,j,2}^{-1} \) and \( |\partial^2 f / \partial v_b \partial u_a| < O(\epsilon) t_{i,j,2}^{-1} \) for all \( a_1, a_2, a \in \{1,2,\ldots, \alpha\} \) and all \( b \in \{1,2,\ldots, \beta\} \).

Define a map \( \hat{s}_{i,j,k,y} : \hat{B}_{i,j,k,y} \rightarrow B^m_\epsilon \) by
\[
\hat{B}_{i,j,k,y} = \hat{s}_{i,j,k,y}^{-1} (\text{Image}(f_p))
\]
and
\[
\hat{s}_{i,j,k,y} = f_p^{-1} \circ s_{i,j,k,y}.
\]
It follows from 3.0, 3.3.3, 3.3.8.1(a)-(c), and from A.1.1-A.1.3 and A.1.3.3.1 (as applied to our present \( f_p \) and also to the exponential map used in 3.3.3(c)), and from the preceding definition, that \( \hat{s}_{i,j,k,y} \) satisfies the following properties:

3.3.8.2. (a) \( \|D\hat{s}_{i,j,k,y} - L\| \ll \epsilon \) for some linear distance preserving map \( L : R^\beta \rightarrow R^m \).

(b) \( \|D^2\hat{s}_{i,j,k,y}\| \ll B \), where \( B > 0 \) depends only on \( m \) and \( A' \) (cf. A.2.2).

(c) Let \( T_0 \) denote the tangent plane to \( \text{Image}(\hat{s}_{i,j,k,y}) \) at the origin of \( R^m \); we also have the subspace \( 0 \times R^\beta \) of \( R^m = R^\alpha \times R^\beta \). We have that \( \Theta(T_0, 0 \times R^\beta) \ll \epsilon \).

Now, by appealing to Properties 3.3.8.1 and 3.3.8.2 we can define a smooth embedding \( \hat{g}_1 : V \rightarrow R^\beta \), from a small open neighborhood \( V \) of the origin in \( R^\beta \), as follows:
3.3.8.3. (a) \( \hat{g}_1(v) = s_{i,j,k,y}^{-1}(u_v, f(u_v, v)) \)
where \( u_v \in \mathbb{R}^\alpha \) is uniquely determined by the requirement:

3.3.8.3. (b) \((u_v, f(u_v, v)) \in \text{Image}(\hat{s}_{i,j,k,y})\).

To complete the verification of Claim 3.3.8 it will suffice to show that \( \hat{g}_1 \) satisfies the following properties:

3.3.8.4. (a) \[ \|D \hat{g}_1|_0 - \hat{I}_1\| \ll \epsilon \] for some linear isometry \( \hat{I}_1 : \mathbb{R}^\beta \to \mathbb{R}^\beta \).

(b) \[ \|D^2 \hat{g}_1|_0\| < O(\epsilon t_{i,j,2}^{-1}) \]

To verify 3.3.8.4(a) we first define a map \( \tilde{g}_1 : V \to \text{Image}(s_{i,j,k,y}) \) defined by

\[ \tilde{g}_1(v) = (u_v, f(u_v, v)) \]

where \( u_v \) is as in 3.3.8.3(b). And we let \( p_2 : \mathbb{R}^m \to \mathbb{R}^\beta \) denote the standard projection of \( \mathbb{R}^m = \mathbb{R}^\alpha \times \mathbb{R}^\beta \) onto its second factor. Note that:

3.3.8.4.1. (a) \[ D(\tilde{g}_1)|_0 = (p_2 \mid T_0)^{-1}, \]

(b) \[ \hat{g}_1 = s_{i,j,k,y}^{-1} \circ \tilde{g}_1, \]

where \( T_0 \) comes from 3.3.8.2(c). Now Property 3.3.8.4(a) is a direct consequence of 3.3.8.4.1(a) and (b) and 3.3.8.2(a) and (c).

To verify 3.3.8.4(b) we need some more maps. For each \( v \in V \) we let \( u_v \in \mathbb{R}^\alpha \) be as in 3.3.8.3(b), and we define a map \( h_v : V \to u_v \times \mathbb{R}^\beta \subset \mathbb{R}^m \) by

\[ h_v(w) = (u_v, f(u_v, w)) \]

for all \( w \in V \). Let \( P_v \) denote the tangent plane to \( \hat{G}_{i,j,k,y} \) at \( (u_v, f(u_v, v)) \); define a map \( \rho_v : \mathbb{R}^m \to u_v \times \mathbb{R}^\beta \) by

\[ \rho_v(q) = (P_v + q) \cap (u_v \times \mathbb{R}^\beta) \]

for all \( q \in \mathbb{R}^m \). Let \( T_v \) denote the tangent plane to \( \text{Image}(\hat{s}_{i,j,k,y}) \) at \( (u_v, f(u_v, v)) \). Note that \( h_v \) and \( \rho_v \) satisfy the following property:

3.3.8.4.2. (a) \[ D\tilde{g}_1|_v = (\rho_v \mid T_v)^{-1} \circ (Dh_v|_v). \]

It also follows from Properties 3.3.8.1(a)-(c), and from 3.3.8.2(a)-(c), that:

3.3.8.4.2. (b) \[ \|Dh_v|_0 - Dh_0|_0\| < O(\epsilon)|v|t_{i,j,2}^{-1}, \]

(c) \[ \|D s_{i,j,k,y}^{-1} \circ (\rho_v \mid T_v)^{-1} - D s_{i,j,k,y}^{-1} \circ (\rho_0 \mid T_0)^{-1}\| \ll O(\epsilon)|v|t_{i,j,2}^{-1}, \]

(d) \[ \|D s_{i,j,k,y}^{-1} \circ (\rho_v \mid T_v)^{-1}\| < 2 \text{ and } \|h_v\| < 2 \text{ for all } v \in V, \]
where we use Euclidean parallel translation to identify the ranges of the two linear maps $Dh_v|v$ and also use Euclidean parallel translation to identify the domains and ranges of the two linear maps $Ds_{-1}^{-1} i,j,k,y \circ (\rho_v | T_v)^{-1}$ and $Ds_{-1}^{-1} i,j,k,y \circ (\rho_0 | T_0)^{-1}$. Now, by combining Properties 3.3.8.4.2(a)-(d) with 3.3.8.4.1(b), and recalling that $t_{i,j,2} \ll \epsilon$ (cf. 3.0), we conclude that

$$\|D\hat{g}_1|v - D\hat{g}_1|0\| < O(\epsilon)|v| t_{i,j,2}^{-1},$$

where the domains and ranges of the linear maps are identified under Euclidean parallel translation. Note that 3.3.8.4(b) is an immediate consequence of this last inequality.

3.4. $\Lambda_{i,j,k,x} \times B_{i,j,4,y} \to B_{i,j,4,y}$ and the fixed point set $C_{i,j,k,y}$. It follows from 3.1.5(a) and (b), and from the fact that the leaves of the foliation $G_{i,j,k,x}$ are just the orbits of the action $G_{t,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$, that the leaves of $G_{i,j,k,x}$ are permuted by the action $\Lambda_{i,j,k,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$ for all $4 \leq k \leq (m+4)^2 - 1$. It follows from 3.1.5(c), and from the definitions given for $F_{i,j,k,y}$ and $\Lambda_{i,j,k,x}$ in 3.3, that the action $\Lambda_{i,j,k,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$ induces an isometric action

$$\Lambda_{i,j,k,x} \times F_{i,j,k,y} \to F_{i,j,k,y}$$

which permutes the leaves of $G_{i,j,k,y}$. Now since $B_{i,j,k,y}$ is obtained from $F_{i,j,k,y}$ by collapsing each leaf of $G_{i,j,k,y}$ to a point we also have the action

$$\Lambda_{i,j,k,x} \times B_{i,j,k,y} \to B_{i,j,k,y},$$

for each $4 \leq k \leq (m+4)^2 - 1$. Note that $B_{i,j,k,y}$ is an open subset of $B_{i,j,4,y}$ (cf. 3.3.4(b), and recall that $F_{i,j,k,y}$ is an open subset of $F_{i,j,k,x}$ for all $k$ and $G_{i,j,k,y} = G_{i,j,k,y} | F_{i,j,k,y}$). Note also that $\Lambda_{i,j,k,x} \subset \Lambda_{i,j,4,x}$ for all $k \geq 4$ (cf. 3.1.1, 3.1.4, 3.1.5(c) and (d)). Thus the action of $\Lambda_{i,j,k,x}$ on $B_{i,j,k,y}$ extends to an action

$$\Lambda_{i,j,k,x} \times B_{i,j,4,y} \to B_{i,j,4,y}.$$

This last action is not an isometric action, but it is nearly so. In fact the following properties can be deduced from 3.3.6(a) and (b):

3.4.1. For any $g \in \Lambda_{i,j,k,x}$ and any $z \in B_{i,j,4,y}$ there is a linear isometry $L_z : \mathbb{R}^3 \to \mathbb{R}^3$ (where $\mathbb{R}^3$ comes from 3.3.6) which satisfies:

(a) $|Dg_{|z} - L_z| \ll \epsilon$;

(b) $|D^2g| < O(\epsilon)t_{i,j,2}^{-1}$.

Note that the following additional property can be deduced from 3.3.5, and from the fact that the action of $g$ on $B_{i,j,4,y}$ leaves invariant the subset $B_{i,j,k,y} \subset B_{i,j,4,y}$:
3.4.1. (c) $|g(0)| \ll t_{i,j,k-1,l}$ for $l = (m+4)^2 - 1$.

We denote by $C_{i,j,k,y} \subset B_{i,j,4,y}$ the fixed point set of the action $\overline{X}_{i,j,k,x} \times B_{i,j,4,y} \to B_{i,j,4,y}$. We will verify the following important properties of $C_{i,j,k,y}$ for all $5 \leq k \leq (m+4)^2 - 1$:

3.4.2. (a) There is $q \in C_{i,j,k,y}$ satisfying $|q| \ll t_{i,j,k-1,l}$ for $l = (m+4)^2 - 1$.
(b) $K(C_{i,j,k,y}; \mathbb{R}^d) \ll O(\epsilon) t_{i,j,2}^{-1}$, where $K(C_{i,j,k,y}; \mathbb{R}^d)$ is defined to be the curvature $K(r; M)$ of $0.1$ when in $0.1$ we set $M = \mathbb{R}^d$ and we let $r : C_{i,j,k,y} \to \{1\}$ denote the constant map.
(c) $C_{i,j,k,y}$ is a closed subset of $B_{i,j,4,y}$.

We define the point $q$ of 3.4.2(a) by $q = \lim_{n \to \infty} q_n$ where the sequence $q_1, q_2, q_3, \ldots$ is defined inductively as follows:

$$q_1 = 0; \quad q_{n+1} = \left( \sum_{g \in X_{i,j,k,x}} g(q_n) \right)^{-1}.$$

It follows from 3.4.1(a) and (c) that $q$ is well-defined and satisfies 3.4.2(a).

We note that 3.4.2(b) is a direct consequence of 3.4.1(a) and (b). Property 3.4.2(c) is immediate. We leave as an exercise for the reader to verify the following additional property for $C_{i,j,k,y}$ for all $5 \leq k < (m+4)^2 - 1$.

3.4.2. (d) $C_{i,j,k,y} \cap B(t_{i,j,4})$ is a path connected manifold.

Note that for any $5 \leq k < (m+4)^2 - 1$ we have $\overline{X}_{i,j,k+1,x} \subset \overline{X}_{i,j,k,x}$ by 3.1.1, 3.1.4, 3.1.5(c) and (d), and 3.3; thus we have that

$$C_{i,j,k,y} \subset C_{i,j,k+1,y}.$$

Now we may use an induction argument (induction over dim$(C_{i,j,k,y})$), in conjunction with the preceding inclusion and with Properties 3.4.2(c) and (d), to conclude that for some $5 \leq k < (m+4)^2 - 2$ we have that:

3.4.3. $C_{i,j,k,y} \cap B(t_{i,j,4}) = C_{i,j,k+1,y} \cap B(t_{i,j,4}) = C_{i,j,k+2,y} \cap B(t_{i,j,4})$.

We also have as a consequence of 3.4.3 that:

3.4.4. $\overline{X}_{i,j,k,x} = \overline{X}_{i,j,k+1,x}$

for $k$ as in 3.4.3. To verify 3.4.4 we use 3.4.2(a) (with $k$ replaced by $k+2$) to choose $q \in C_{i,j,k+2,y}$ which satisfies:

3.4.5. (a) $|q| \ll t_{i,j,k+1,l}$ for $l = (m+4)^2 - 1$. 
It follows from 3.4.3 that:

**3.4.5.** (b) $q \in C_{i,j,k,y}$.

Note it also follows from 3.3.1(b) (with $k$ replaced by $k+1$ in 3.3.1(b)), 3.3.3, 3.3.4(a), and from 3.4.5(b) and from the fact that $\pi_{i,j,k,y}^{-1}(q)$ is a covering space for $\pi_{i,j,k,y}^{-1}(q)/\Lambda_{i,j,k,x}$, that the following property holds:

**3.4.5.** (c) $\text{diameter}(\pi_{i,j,k,y}^{-1}(q)/\Lambda_{i,j,k,x}) < 2\tau t_{i,j,k+2}$.

Let $L \in G_{i,j,k,y}$ denote the leaf which covers the leaf $\pi_{i,j,k,y}^{-1}(q) \in G_{i,j,k,y}$, then we may reinterpret Properties 3.4.5(a) and (c) as follows:

**3.4.5.** (d) There is $\hat{q} \in L$ with $d_{\hat{V}}(y, \hat{q}) \ll t_{i,j,k+1}$ for $l = (m+4)^2 - 1$.

(e) $L$ is left invariant by the action $\Lambda_{i,j,k,x} \times \hat{V} \to \hat{V}$. Moreover there is a generating set $\{s_1, s_2, \ldots, s_t\}$ for $\Lambda_{i,j,k,x}$ such that for all $s_r$ we have $d_{\hat{V}}(\hat{q}, s_r(\hat{q})) < 8\tau t_{i,j,k+2}$.

Now 3.4.4 follows from 3.4.5(d) and (e), and from 3.0, 3.1.4, 3.2.1(a).

### 3.5. The infranil core $r : U \to B$

In this subsection we shall complete the proof of Theorem 0.3 by defining $r : U \to B$ and $\delta_c > 0$ of 0.3 which satisfy Properties 0.3(a)-(c).

First we define $\delta_c$ by:

**3.5.1.** (a) $\delta_c = t_{i,j,k,3}$,

where for the duration of this subsection the positive integers $i, j, k$ are as in 3.1.1, 3.1.3, 3.4.3 and 3.4.4, respectively. Now we let $q \in C_{i,j,k+1,y}$ be as in 3.4.5(a). Recall that $C_{i,j,k,y} \subset \mathbb{R}^3$; let $P \subset \mathbb{R}^3$ denote the plane which is tangent to $C_{i,j,k,y}$ at $q$; and let $\hat{P} : \mathbb{R}^3 \to P$ denote the orthogonal projection onto $P$. Set:

**3.5.1.** (b) $B = \{ u \in P : |u - q| < \delta_c \}$;

then $B$ is isometrically equivalent to an open ball of radius $\delta_c$ centered at the origin (i.e., at $q$) in some Euclidean space (i.e., in $P$), as is required in 0.2.1(b) and 0.3(a). Note that $(\hat{P} | C_{i,j,k,y})^{-1}$ is well-defined (cf. 3.4.2(b)-(d)) and maps $B$ diffeomorphically onto an open subset $\tilde{B} \subset C_{i,j,k,y}$. Define the map $r : U \to B$ by:

**3.5.1.** (c) $U = \pi_{i,j,k,y}^{-1}(\tilde{B})/\Lambda_{i,j,k,x}$.

(d) Note that $\pi_{i,j,k,y}$ induces a map $\pi_{i,j,k,y} : U \to \tilde{B}$; let $r$ denote the composition of $\pi_{i,j,k,y}$ with $\hat{P} | \tilde{B}$. 


First we must verify that \( r : U \to B \) is an infranil core for \( M \) as defined in 0.2. We note that for each \( z \in B \) there is a leaf \( L_z \in G_{i,j,k,y} \) which is left invariant by the action \( \Lambda_{i,j,k,x} \times \hat{V} \to \hat{V} \) such that \( r^{-1}(z) = L_z/\Lambda_{i,j,k,x} \). Since the leaves of \( G_{i,j,k,y} \) are just the orbits of the free action \( G_{i,j,x} \times F_{i,j,k,y} \to F_{i,j,k,y}, \) it follows that each fiber \( r^{-1}(z) \) is diffeomorphic to the infranil manifold \( G_{i,j,x}/\Lambda_{i,j,k,x} \); cf. 3.1.5 and 3.2. This shows that \( r \) satisfies 0.2.1(a). We note that 0.2.1(b) is immediate from 3.5.1(b). To see that \( r \) satisfies 0.2.1(c) we argue as follows. First we deduce from 3.4.2(b) and 3.3.1(a) that:

3.5.2. \( a \) \( K(\mathcal{G}; U) < \mathcal{O}(\epsilon) t_{i,j,2}^{-1} \)

where \( \mathcal{G} \) denotes the foliation for \( U \) by the fibers of \( r \). And we may deduce from 3.4.5(c) and 3.5.1(a) (see also 3.0) that:

3.5.2. \( b \) \( \text{diameter}(U) < \mathcal{O}(\delta_c) \).

Let \( \hat{q} \in \hat{V} \) denote a point in the preimage of \( r^{-1}(q \sim 0) \) under the covering space projection \( \pi : \hat{V} \to V \) (cf. Appendix 2); and set:

3.5.2. \( c \) \( \hat{U} = \pi^{-1}(U) \cap B(\hat{q}, t), \)

where \( t \) denotes the \( \mathcal{O}(\delta_c) \) of 3.5.2(b) and where \( B(\hat{q}, t) \) denotes the open ball of radius \( t \) centered at \( \hat{q} \) in \( \hat{V} \). Note that 3.4.2(b) and 3.3.1(a) (together with the fact that \( F_{i,j,k,y} \) is locally a totally geodesic submanifold of \( \hat{V} \)) also imply:

3.5.2. \( d \) \( K(\hat{U}; \hat{V}) < \mathcal{O}(\epsilon) t_{i,j,2}^{-1} \)

where \( K(\hat{U}; \hat{V}) \) denotes the curvature of the constant map \( \hat{U} \to \{1\} \) (as described in 0.1) for the submanifold \( \hat{U} \subset \hat{V} \) with respect to the Riemannian metric \( g'_V \) of Appendix 2. We deduce from 3.5.2(c) and (d), and from A.2.1(b) and 3.0 and 3.5.1(a), that the following holds for \( \hat{U} \subset \hat{V} \):

3.5.2. \( e \) \( \hat{U} \) has a well-defined tubular neighborhood of radius \( \delta_c \) in \( \hat{V} \).

Finally, we may deduce that \( U \) satisfies Property 0.2.1(c) from 3.5.2(b), (c) and (e), and from 3.0, 3.1.1, 3.1.3, 3.1.4, 3.1.5(c) and (d), 3.4.4.

Now we must verify that \( r : U \to B \) satisfies Properties 0.3(a)-(c).

Towards verifying 0.3(a) we first note that \( B \) has radius equal \( \delta_c \) (cf. 3.5.1(b)). Moreover \( c \in (0, \eta) \) by 3.0 and 3.5.1(a). (Recall that we are proving 0.3 for the special case \( n = 0 \).) This completes the verification of 0.3(a).

Next we will verify 0.3(c). Let \( \hat{q} \in \hat{V} \) be as in 3.4.5(d), and define \( p' \in U \) of 0.3 to be the image of \( \hat{q} \) under the covering projection \( \pi : \hat{V} \to V \). Note that \( \hat{q} \) of 3.5.3 may be chosen equal to \( \hat{q} \) of 3.5.2(c); thus we have:
3.5.3. (a) $|r(p')| = 0$.

[Actually we have $r(p') = q \in P$; but we are identifying $(P, q)$ isometrically with $(\mathbb{R}^n, 0)$ for some integer $n$, in order that $B$ be an open ball centered at the origin of $\mathbb{R}^n$ as is required in 0.2.1.] Note also that it follows from 3.0, 3.2.1(a), 3.4.5(d), 3.5.1(a), and from the fact that $x \in E$ is mapped to $p \in M$ under the composition $E \xrightarrow{\rho} \hat{V} \xrightarrow{\pi} V$ (cf. 3.1), that:

3.5.3. (b) $d_M(p, p') < \epsilon \delta_c$,

where $d_M(\ , \ )$ denotes the distance in the manifold $M$ of 0.3. Now 0.3(c) is a consequence of 3.5.3(a) and (b).

It remains to verify Property 0.3(b); i.e., we must find a number $\theta \in (0, 1)$ which depends only on $\text{dim } M$ such that Properties 0.2.2(a)-(d) hold, for $\delta = \delta_c$ in 0.2.2.

Note that Property 0.2.2(c) can be deduced fairly directly from 3.0, 3.3.1(a), 3.3.3(a), and 3.5.1(a)-(d).

Towards verifying 0.2.2(b) we first note that

$$r^{-1}(0) = \pi_{i,j,k,y}^{-1}(q)/\Lambda_{i,j,k,x};$$

thus we may deduce from 3.0, 3.5.1(a), and from 3.4.5(c) that

$$\text{diameter}(r^{-1}(0)) \ll \epsilon \delta_c.$$

Now 0.2.2(b) follows from this last inequality and from Properties 3.5.2(a) and 0.2.2(c).

Next we verify 0.2.2(a). We have already verified part of this property by noting above that 3.4.2(b) and 3.3.1(a) together imply 3.5.2(a) and (d). Unfortunately the curvature bounds for $\mathcal{G}$ provided by 3.5.2(a) can not be translated into curvature bounds for $r$. [Although the converse is true: A bound on $K(r; M)$ can be translated into a bound for $K(\mathcal{G}; U)$ and for $K(\hat{U}, \hat{V})$.] What we need is a stronger version of 3.3.1(a) which we can combine with 3.4.2(b) to deduce 0.2.2(a), and this stronger version is provided by the following claim. Let $\pi_{i,j,k,y} : \bar{F}_{i,j,k,y} \to B_{i,j,k,y}$ denote the fiber bundle projection defined in 3.3, and $\bar{V}$ be as in 3.3.1(a).

**Claim 3.5.4.** $K(\pi_{i,j,k,y}; \bar{V}) < O(\epsilon)t_{i,j,2}^{-1}$.

The verification of this claim, which will be carried out at the end of this chapter, relies on 3.3.1(a) and 3.3.6. Note that 3.5.4 and 3.4.2(b), together with 3.3.3 and 3.5.1, imply that

$$K(r; M) < O(\epsilon)t_{i,j,2}^{-1}.$$
Finally we will verify 0.2.2(d). Let \( v \in TM|_U \) be perpendicular to \( U \); and denote by \( \hat{v} \in T\hat{V} \) a vector which is in the pre-image of under the covering projection \( \pi : \hat{V} \to V \). In order to verify that \( v \) satisfies the conclusions of 0.2.2(d) it will suffice to consider the following special cases:

(i) The foot of \( \hat{v} \) lies in \( F_{i,j,k,y} \), and \( \hat{v} \) is perpendicular to \( F_{i,j,4,y} \) in \( \hat{V} \);
(ii) the foot of \( \hat{v} \) lies in \( F_{i,j,k,y} \), and \( \hat{v} \) is tangent to \( F_{i,j,4,y} \).

If \( \hat{v} \) is perpendicular to \( F_{i,j,4,y} \) in \( \hat{V} \) then we argue as follows: Since \( F_{i,j,4,y} \) is an open subset of the fixed point set for the compact group action \( C_{i,j,x} \times \hat{V} \rightarrow \hat{V} \) (cf. 3.2), we can choose \( g \in C_{i,j,x} \) so that:

3.5.5. (a) \( \lambda_1 < \Theta(\hat{v}, g(\hat{v})) \)

where \( \lambda_1 > 0 \) depends only on \( \dim \hat{V} \). Recall that \( G_{i,x} \) is the foliation of \( E \) by the orbits of the free action \( \Gamma \times E \to E \); let \( u \) denote the foot of \( \hat{v} \), let \( \hat{u} \) denote a point in the pre-image of \( u \) under the map \( \rho : E \to \hat{V} \), and let \( \hat{L} \) denote the leaf of \( G_{i,x} \) which contains the point \( \hat{u} \). We have shown in 3.2.13(d) that:

3.5.5. (b) \( \text{diameter}(\hat{L}/\Lambda_{i,j,k,x}) < 2\tau t_{i+1} \).

[Note that there is no loss of generality in choosing the preimage \( \hat{u} \) of \( u \) so that our present choice of \( \hat{L} \) is equal the \( \hat{L} \) of 3.2.13.] Thus, for some \( g' \in \Lambda_{i,j,k+1,x} \) (cf. 3.1.5(d)), we have that:

3.5.5. (c) \( d_L'(g(\hat{u}), g'(\hat{u})) < \lambda_2 t_{i+1} \)

where \( \lambda_2 > 1 \) depends only on \( \dim E \). Since \( g(u) = u \), and since \( \rho : E \to \hat{V} \) is a Riemannian submersion, we may deduce from 3.5.5(c) that:

3.5.5. (d) \( d_L'(u, g'(u)) < \lambda_2 t_{i+1} \),

where the leaf \( L \in G_{i,j,k,y} \) contains \( u \). Hence we can find a smooth path \( \hat{f} : [0,1] \to \hat{L} \) which satisfies the following property:

3.5.5. (e) \( \hat{f}(0) = u \) and \( \hat{f}(1) = g'(u) \); \( \text{length}(\hat{f}) < 2\lambda_2 t_{i+1} \).

Now the desired path \( f : [0,1] \to U \) in 0.2.2(d) may be defined to be the composition of \( \hat{f} \) with the covering projection \( \pi : \hat{V} \to V \). That \( f \) satisfies the conclusions of 0.2.2(d) can be deduced from 3.0, 3.5.1(a), and 3.5.5(a), (c) and (e). This completes the verification of 0.2.2(d) for the case when \( \hat{v} \) is perpendicular to \( F_{i,j,4,y} \) in \( \hat{V} \).

If \( \hat{v} \) is tangent to \( F_{i,j,4,y} \) in \( \hat{V} \) then we verify 0.2.2(d) as follows: Let \( \overline{v} \) denote the image of \( \hat{v} \) under the composition of the covering map \( F_{i,j,4,y} \to \overline{F}_{i,j,4,y} \) (recall that \( \overline{F}_{i,j,4,y} = F_{i,j,4,y}/\Lambda_{i,j,4,x} \)) with the projection map \( \pi_{i,j,k,y} : \overline{F}_{i,j,4,y} \to \overline{V} \).
\( F_{i,j,4} \to B_{i,j,4,y} \). It follows from 3.0, 3.3.1, 3.3.3, 3.5.1, and from the fact that \( v \) is perpendicular to \( U \) in \( F_{i,j,4,x}/\Lambda_{i,j,k,x} \), that:

3.5.6. (a) \( \Theta(v, C_{i,j,k,x}) > \pi/2 - O(\epsilon) \),

where \( C_{i,j,k,x} \) is the fixed point set of the action \( \Lambda_{i,j,k,x} \times B_{i,j,4,x} \to B_{i,j,4,x} \).

Since the foot of \( v \) (denoted by \( u \)) is in \( C_{i,j,k,x} \), it follows from 3.5.6(a) (see also 3.4.1) that we can choose \( g \in \Lambda_{i,j,k,x} \) such that:

3.5.6. (b) \( \Theta(v, g(v)) > \lambda_3 \)

where \( \lambda_3 \) depends only on \( \dim B_{i,j,k,y} \). Note that it follows from 3.3.1(b), 3.4.3, 3.4.4, and from 3.5.1 that the fiber \( r^{-1}(u) \) has a finite covering space \( Y \to r^{-1}(u) \) which satisfies:

3.5.6. (c) \( \text{diameter}(Y) < 2\tau_t_{i,j,k+1} \);

(d) \( Y/\Lambda_{i,j,k,x} = r^{-1}(u) \).

Let \( u \in Y/\Lambda_{i,j,k,x} \) denote the foot of \( v \) and let \( u' \in Y \) denote a pre-image of \( u \) under the covering map \( Y \to r^{-1}(u) \). Using 3.5.6(c) and (d) we may choose a path \( \hat{f} : [0,1] \to Y \) which satisfies:

3.5.6. (e) \( \hat{f}(0) = u' \) and \( \hat{f}(1) = g(u') \); \( \text{length}(\hat{f}) < 2\tau_t_{i,j,k+1} \).

Now the desired path \( f : [0,1] \to U \) of 0.2.2(d) may be defined to be the composition of \( \hat{f} \) with the covering projection \( Y \to r^{-1}(u) \). That \( f \) and \( v \) satisfy the conclusions of 0.2.2(d) follows from 3.0, 3.5.1(a), 3.5.4, 3.3.6(a), and 3.5.6(b) and (e). This completes the verification for 0.2.2(d) when \( \hat{v} \) is tangent to \( F_{i,j,4} \).

**Verification of Claim 3.5.4.**

Since \( F_{i,j,k,y} \) is an open subset of a totally geodesic submanifold of \( V \), it will suffice to show that there is a positive number \( \kappa \), with \( \kappa < O(\epsilon)t_{i,j,2}^{-1} \), such that for any smooth path \( h : [0,1] \to F_{i,j,k,y} \) we have:

3.5.4.1. \( \|P_h(D\pi_{i,j,k,y}|h(0)), D\pi_{i,j,k,y}|h(1)\| < \kappa(\text{length}(h)) \).

Recall (cf. 0.1) that for any \( v \in T(F_{i,j,k,y})_{h(1)} \) we have that

\[
P_h(D\pi_{i,j,k,y}|h(0))v = P \circ D\pi_{i,j,k,y}|h(0) \circ \rho \circ P_{\hat{h}}(v)
\]

where \( P \) is Euclidean parallel translation in \( B_{i,j,k,y} \), and \( \rho : T\overrightarrow{V}_{|h(0)} \to T(F_{i,j,k,y})_{|h(0)} \) denotes orthogonal projection, and \( P_{\hat{h}} \) is parallel translation along \( \hat{h} \) in \( \overrightarrow{V} \) where \( \hat{h}(t) = h(1-t) \) for all \( t \in [0,1] \).

We will first verify 3.5.4.1 in the following two special cases:
Case I. We suppose in this case that $h$ of 3.5.4.1 satisfies Image($h$) $\subset L$ for some $L \in \tilde{G}_{i,j,k,y}$; and we shall deduce 3.5.4.1 from 3.3.1(a). In more detail, we may proceed as in the verification for Claim 3.3.8 given above. Set $p = h(0)$ and let $f_p : B_t^m \rightarrow F_{i,j,k,y}$ be a smooth embedding as in A.1.1 and A.1.2. Let $\hat{G}_{i,j,k,y}$ and $f : R^m \rightarrow R^\beta$ be as in 3.3.8.1. There will be no loss of generality in assuming that Image($h$) $\subset$ Image($f_p$); because if 3.5.4.1 holds for all curves of sufficiently small length, then it must hold for all curves. Thus we define $\hat{h} : [0,1] \rightarrow B_t^m$ by

$$\hat{h} = f_p^{-1} \circ h.$$

Recall that 3.3.8.1(c) is just a reformulation for the curvature bounds given in 3.3.1(a). Define a map $\hat{\pi}_{i,j,k,y} : B_t^m \rightarrow B_{i,j,k,y}$ by

$$\hat{\pi}_{i,j,k,y} = \pi_{i,j,k,y} \circ f_p.$$

Note that $\hat{\pi}_{i,j,k,y}$ is equal to the composition of the orthogonal projection of $B_t^m$ onto $R^\beta$ (cf. 3.3.8.1) with one of the maps $s = \pi_{i,j,k,y} \circ s'_{i,j,k,y}$ of 3.3.6. Thus we may apply 3.3.8.1(a)-(c) and 3.3.6(a) and (b) to deduce that:

**3.5.4.2.** $\|P_h(D\hat{\pi}_{i,j,k,y}|h(0)), D\hat{\pi}_{i,j,k,y}|h(1)\| < \hat{\kappa}(\text{length}(\hat{h}))$

for some positive number $\hat{\kappa}$ satisfying $\hat{\kappa} < O(\epsilon)t_{i,j}^{-1}$, where the $\Theta(\,)$ and $P_h$ and length($\hat{h}$) are all geometric constructions and measurements made with respect to the Euclidean metrics on $B_t^m$ and on $B_{i,j,k,y}$. Let $\delta_j^i dx_i dx_j$ denote the Euclidean metric on $B_t^m$, and let $g_{i,j} dx_i dx_j$ denote the metric $B_t^m$ pulled back along $f_p$ from $V$ (cf. A.1.1). Now, by comparing parallel translation in $B_t^m$ with respect to these two Riemannian metrics (see Appendix 1), we see that 3.5.4.2 implies 3.5.4.1.

Case II. We suppose in this case that $h$ of 3.5.4.1 satisfies Image($h$) $\subset$ Image($s'_{i,j,k,y}$) for some map $s'_{i,j,k,y} : B'_{i,j,k,y} \rightarrow F_{i,j,k,y}$ as in 3.3.6; and we shall deduce 3.5.4.1 from 3.3.1(a) and 3.3.6. Recall (cf. 3.3) that $s'_{i,j,k,y}$ is a composition

$$B'_{i,j,k,y} \subset R^\beta \xrightarrow{f_{W'}} W' \xrightarrow{\text{exp}} V,$$

where $B'_{i,j,k,y}$ is an open subset of the Euclidean space $R^\beta$, $f_{W'}$ is a linear isometry onto a subspace of $T(F_{i,j,k,y})|_{\overline{y}}$ for some $\overline{y} \in F_{i,j,k,y}$, and exp denotes the exponential map. Thus the Euclidean metric $\delta_j^i dx_i dx_j$ on $B'_{i,j,k,y}$ is a very close approximation to the Riemannian metric $g_{i,j} dx_i dx_j$ gotten by pulling the Riemannian metric on $V$ back along $s'_{i,j,k,y}$ (see Appendix 1). In light of these few preceding remarks we may deduce from Properties 3.3.6(a) and (b) that:
3.5.4.3. \( |P_h(D\pi_{i,j,k,y}|h(0))(v), D\pi_{i,j,k,y}|h(1)(v)| < \kappa'(\text{length}(h)) \)

for any \( v \in T(\text{image}(s'_{i,j,k,y}))|h(1)) \), where \( \kappa' \) is a positive number satisfying \( \kappa' < O(\epsilon)^{-1} \). Finally we may deduce 3.5.4.1 from 3.5.4.3 and 3.3.1(a).

To verify 3.5.4.1 in general we note that any smooth path \( h : [0,1] \rightarrow F_{i,j,k,y} \) may be point wise approximated by a piecewise smooth path \( g : [0,1] \rightarrow F_{i,j,k,y} \) which satisfies:

(i) \( g(0) = h(0) \) and \( g(1) = h(1) \);
(ii) every smooth piece of \( g \) is either a path as in Case I (reparametrized)
or a path as in Case II (reparametrized);
(iii) \( (1/3)\text{length}(h) < \text{length}(g) < 3(\text{length}(h)) \).

Then, since 3.5.4.1 has been verified for Cases I and II, it follows from (i) above that 3.5.4.1 holds also for \( g \). Also, since \( g \) point wise approximates \( h \), we have that parallel translation along \( g \) approximates parallel translation along \( h \). Thus, by (iii) and the two preceding remarks, we may conclude that \( h \) also satisfies 3.5.4.1.

Appendix 1.

Let \((M,g)\) denote a complete \(A\)-regular Riemannian manifold. In this appendix we reformulate the \(A\)-regular condition for \((M,g)\) in terms of the local coordinates associated with the exponential map \( \exp : TM \rightarrow M \) (the “normal” coordinates for \(M\)). We analyze to what extent the metric \(g\) is approximated by the Euclidean metric for a normal coordinate system (cf. A.1.1 and A.1.2), and to what extent germs of isometries of \(M\) are approximated by linear isometries for a normal coordinate system (cf. A.1.3). We also reexamine here the notion of “curvature” for a map (cf. 0.1): We analyze the relations between curvatures computed in \((M,g)\) and curvatures computed with respect to the Euclidean metric for a normal coordinate system for \(M\) (cf. A.1.4).

All of the results A.1.1-A.1.4 are referred to in the proof for Theorem 0.5 given above.

We begin by paraphrasing a result of J. Jost and H. Karcher [14, Satz 3.4 and 5.1], and of J. Bemelmans, M. Min-Oo, and E. Ruh [1]. In the next theorem we let \( B_\epsilon^m \) denote the open ball of radius \( \epsilon > 0 \) centered at the origin of \( \mathbb{R}^m \), and we let \( \partial/\partial x_i : \mathbb{R}^m \rightarrow T\mathbb{R}^m, i = 1,2,\ldots,m, \) denote the vector fields associated to the standard coordinates \( x_1,x_2,\ldots,x_m \) for the \( \mathbb{R}^m \).

**Theorem A.1.1.** There is \( \eta > 0 \) and a collection \( B = \{B_i\} \) of positive numbers \( B_1,B_2,B_3,\ldots \) which depend only \( A = \{A_i\} \) and \( m = \dim M \). For each \( p \in M \) and each \( \epsilon \in (0,\eta) \) there is a smooth immersion \( f : B_\epsilon^m \rightarrow M \)
which satisfies the following properties. Let \( g_{i,j} : B^m_\epsilon \to \mathbb{R} \) be defined by 
\[ g_{i,j}(x) = g(Df(\partial/\partial x_i(x)), Df(\partial/\partial x_j(x))) \] for all \( x \in B^m_\epsilon \).

(a) \( f(0) = p \) and \( g_{i,j}(0) = \delta^j_i \).

(b) \( |\partial^k g_{i,j}/\partial x_{s_1} \partial x_{s_2} \ldots \partial x_{s_k}(x)| \leq B_k \) for all \( k, i, j, \{s_1, \ldots, s_k\}, x \in B^m_\epsilon \).

**Remark A.1.1.** Actually Jost and Karcher [14] verify this theorem under the added hypothesis that there is a lower bound \( \tau > 0 \) to the injectivity radius of \( M \), and \( \epsilon > 0 \) is also dependent on \( \tau \). The following simple trick can be used to remove this hypothesis (cf. [1]). Since the curvature of \( M \) is pinched, there is \( \alpha > 0 \), which depends only on \( \text{dim} \ M \) and the bounds of the curvature of \( (M, g) \), so that the following is true: For any \( p \in M \) the exponential map \( \exp : B_\alpha \to M \) is an immersion, where \( B_\alpha \) denotes the open ball of radius \( \alpha \) centered at the origin in \( TM_p \); moreover the \( g^* \)-injectivity radius at each point \( p \in B_{\alpha/2} \) is greater than \( \tau > 0 \), where \( g^* \) is the pull back along \( \exp \) of the metric \( g \) on \( M \), and where \( \tau > 0 \) depends only on \( \text{dim} \ M \) and the curvature bounds for \( (M, g) \) (cf. Cheeger and Ebin [3]). Thus the results of [14, Satz 3.4 and 5.1] may be applied to \( (B_\alpha, g^*) \) near the origin of \( B_\alpha \) to obtain a desirable coordinate system \( h : B^m_\epsilon \to B_\alpha \) near the origin in \( (B_\alpha, g^*) \); now the immersion \( f : B^m_\epsilon \to M \) of Theorem A.1.2 can be defined to be the composition \( B^m_\epsilon \xrightarrow{h} B_\alpha \xrightarrow{\exp} M \).

The maps \( h : B^m_\epsilon \to B_\alpha \) referred to in the preceding remark are the harmonic coordinates for \( (B_\alpha, g^*) \) near the origin (cf. [14, Satz 3.4 and 5.1]). So, roughly speaking, the maps \( f : B^m_\epsilon \to M \) which Theorem A.1.1 provides are the “immersed” harmonic coordinates. The following corollary states Theorem A.1.1 is also true for the “immersed” normal coordinates. Both Corollaries A.1.2 and A.1.3 are proven below.

**Corollary A.1.2.** We may assume in A.1.1 that \( f = \exp \circ L \), where \( L : \mathbb{R}^m \to TM_p \) is a linear isometry and where \( \exp : TM_p \to M \) denotes the exponential map.

**Corollary A.1.3.** For given \( p, p' \in M \) we let \( f : B^m_\epsilon \to M \) and \( f' : B^m_\epsilon \to M \) denote the maps given by A.1.2 with \( f(0) = p \) and \( f'(0) = p' \). Suppose that for \( x, x' \in B^m_{\epsilon/9} \) we have that \( f(x) = f'(x') \). Then there is a smooth embedding \( h : B^m_{\epsilon/9} \to B^m_\epsilon \) uniquely determined by Property (a) below; and there is an isometry \( \bar{h} : \mathbb{R}^m \to \mathbb{R}^m \) of Euclidean space satisfying Properties (b) and (c) below:

(a) \( h(x) = x' \), and \( f \mid B^m_{\epsilon/9} = f' \circ h \).

(b) \( h(0) = \bar{h}(0) \), and \( \|D(h - \bar{h})z\| \leq O(\epsilon^2) \) at all \( z \in B^m_{\epsilon/9} \). [Here, and in Part (c) below, \( \|\| \) indicates the Euclidean norm of the appropriate derivative.]
There is a collection of positive numbers $\mathcal{B} = \{B_i : i = 1, 2, \ldots\}$ which depend only on the $B = \{B_i\}$ and $\dim M$. We have for all $i$ and all $z \in B_{\epsilon/9}^m$ that

$$\|D^i h_{|z}\| \leq B_i.$$

**Remark A.1.3.1.** Note that 1.3 (in the Proof of Theorem 0.5) can be deduced from A.1.3 in the special case when $p = p'$ and $f = f'$: Just identify $x, y, \epsilon$ of 1.3 with $x, x', \epsilon$ of Corollary A.1.3; and assume (with no loss of generality) that $100\delta < \epsilon$, where $\delta$ comes from 0.5.

In the next theorem we let $p \in M$ and $f : B_{\epsilon}^m \to M$ be as in Theorem A.1.1 and Corollary A.1.2. We let $r : U \to B$ denote an infranil core for $M$; recall that $B$ is a ball in some Euclidean space $\mathbb{R}^k$. We set

$$\hat{U} = f^{-1}(U),$$

$$\hat{r} = r \circ f.$$ So $\hat{r}$ maps $\hat{U}$ to $\hat{B} = B$. Let $K(r; M)$ denote the curvature for $r$ in $(M, g)$ as defined in 0.1; and let $K(\hat{r}; B_{\epsilon}^m)$ denote the curvatures for $\hat{r}$ in $(B_{\epsilon}^m, e)$ as defined in 0.1, where $e$ denotes the Euclidean metric.

**Theorem A.1.4.** Suppose that $U \subset f(B_{\epsilon}^m)$, $K(r; M) \leq \epsilon^{-1}$, and that $\|Dr_q\| < O(1)$ for all $q \in U$. Then we have that

$$K(\hat{r}; B_{\epsilon}^m) - O(\epsilon) < K(r; M) < K(\hat{r}; B_{\epsilon}^m) + O(\epsilon).$$

**Remark A.1.5.** The preceding theorem remains true if we replace $\hat{r}$ by $\hat{r} \mid X$, for any open subset $X \subset \hat{U}$ which satisfies $U \subset f(X)$. This theorem will be proven below.

In our next and last theorem of this appendix we let $r : U \to B$ denote an infranil core for $(M, g)$ of radius $\delta > 0$, and we let $s : E \to B$ denote the thickening for $r$ described in 0.4. For each positive number $t > 0$ we denote by $K(s; M, t)$ the glb for all numbers $\sigma > 0$ which satisfy

$$\|Ds_{f(1)} - P_{s \circ f} \circ Ds_{f(0)} \circ P_f^{-1}\| < \sigma t$$

for any smooth path $f : [0, 1] \to E$ having length less than or equal to $t$. Here $P_f$ denotes parallel translation along $f$ in $(M, g)$, and $P_{s \circ f}$ denotes Euclidean-parallel translation along $s \circ f$ in $B$ (recall that $B$ is an open ball in some Euclidean space). Note (cf. 0.1 as applied to $s$) that

$$K(s; M) = \text{lub}\{K(s; M, t) : 0 < t\}.$$
Theorem A.1.6. Let $\epsilon, \delta > 0$ denote sufficiently small numbers, where how small is sufficient for $\epsilon$ depends only on $A = \{A_i\}$ and $\dim M$, and how small is sufficient for $\delta$ depends only on $\epsilon, A = \{A_i\}$, and $\dim M$. If $r$ satisfies Properties 0.2.2(a)-(c) for our present choice of $\epsilon, \delta > 0$, then $s$ must satisfy

$$K(s; M, \delta) < O(\epsilon)\delta^{-1}.$$ 

The following lemma will be used in the proof for Theorem A.1.6. In this lemma we let $\hat{s} : \hat{E} \to \hat{B}$ denote the composition function $\hat{E} \xrightarrow{f_p} E \xrightarrow{\delta} B$, where $f_p$ now denotes the immersion $f : B^m_\epsilon \to M$ of A.1.1 and A.1.2, and $\hat{E} = f_p^{-1}(E)$; and we let $X \subset \hat{E}$ denote an open subset of $\hat{E}$. For any number $t > 0$ we denote by $K(\hat{s} \mid X; B^m_\epsilon, t)$ the glb of all numbers $\sigma > 0$ which satisfy

$$\|D\hat{s}_f(1) - P_{\hat{s}f(0)} \circ D\hat{s}_f^{-1}\| < \sigma t$$

for any smooth path $f : [0, 1] \to X$ which has Euclidean length less than or equal to $t$.

Lemma A.1.7. Suppose $E \subset f_p(B^m_\epsilon)$ and that for any smooth path $f : [0, 1] \to E$, having length less than or equal to $t$, there is another smooth path $\hat{f} : [0, 1] \to X$ such that $f_p \circ \hat{f} = f$. Suppose also that $\|Ds_q\| < O(1)$ for all $q \in E$. Then we have that

$$K(\hat{s} \mid X; B^m_\epsilon, t) = O(\epsilon) < K(s; M, t) < 3K(\hat{s} \mid X; B^m_\epsilon, t) + O(\epsilon)$$

provided $\epsilon$ is sufficiently small depending only on $A = \{A_i\}$ and $\dim M$.

The proof of Lemma A.1.7, which is analogous to the proof of Theorem A.1.4, is left as an exercise for the reader.

Proof of Corollary A.1.2. Recall Theorem A.1.1 which defines $\eta, \{B_i\}$ depending on $\{A_i\}$ and $\dim M$, such that for each $\epsilon \in (0, \eta)$ we have an immersion $f : B^m_\epsilon \to M$ and a pull back metric $\{g_{i,j}\}$ on $B^m_\epsilon$ satisfying Properties A.1.1(a) and (b). Let $\lambda_m$ be defined as below (A.1.2.1(c)); and choose $\epsilon$ as below (A.1.2.1(a)). Then Theorem A.1.1 implies that there is an immersion $f : B^m_\epsilon \to M$ such that the pulled back metric $\{g_{i,j}\}$ satisfies A.1.2.1(b). (Note that A.1.2.1(b) is true for any choice of $\lambda_m > 0$ just so long as $\epsilon, B_1, \eta, \lambda_m$ are related as in A.1.2.1(a); the restriction placed on $\lambda_m$ in A.1.2.1(c) will be relevant later in this section.)

A.1.2.1. (a) $\epsilon < \min\{\lambda_m/B_1, \eta\}$.
(b) $|g_{i,j}(x) - \delta_{i,j}| < \lambda_m$, for all $x \in B^m_\epsilon$.
(c) $\lambda_m = \min\{(10m)^{-m}, (10^{-m} + 1)^{1/m} - 1\}$.
Let $\Gamma^k_{i,j} : B_\epsilon^m \to \mathbb{R}$ denote the classical Christoffel functions for the Riemannian metric represented by the $\{g_{i,j}\}$ with respect to the standard coordinates $x_1, x_2, \ldots, x_m$. Using the equality

$$\Gamma^k_{i,j} = \frac{1}{2} \sum_{r=1}^{m} h_{k,r} \left( \frac{\partial g_{r,j}}{\partial x_i} + \frac{\partial g_{r,i}}{\partial x_j} + \frac{\partial g_{i,j}}{\partial x_r} \right)$$

where the matrix $[h_{i,j}]$ is the inverse of the matrix $[g_{i,j}]$, together with A.1.1(b) and A.1.2.1(b) and (c), we can deduce the following:

A.1.2.2. $|\partial^k \Gamma^k_{i,j}(x)| < C_n$ for all $n, i, j, k, \{s_1, \ldots, s_n\}, x \in B_\epsilon^m$.

Here $C = \{C_i\}$ is a collection of positive numbers which depend only on $B = \{B_i\}$ and $\dim M$.

The standard coordinates $x_1, x_2, \ldots, x_m$ for $\mathbb{R}^m$ induce coordinates on $T \mathbb{R}^m$ by $y_1, y_2, \ldots, y_m, y_{m+1}, \ldots, y_{2m}$ via the correspondence

$$\sum_{i=1}^{m} y_{m+i} \partial/\partial x_i(y_1, \ldots, y_m) \leftrightarrow (y_1, y_2, \ldots, y_m, y_{m+1}, \ldots, y_{2m}).$$

We define a vector field $V : TB_\epsilon^m \to T(B_\epsilon^m)$ on $TB_\epsilon^m$ by A.1.2.3(a); note that we may deduce from A.1.2.2 that $V$ satisfies Property A.1.2.3(b) and (c).

A.1.2.3. (a) $V = \sum_{i=1}^{2m} V_i \partial/\partial y_i$, where

$$V_i(y_1, \ldots, y_{2m}) = \begin{cases} y_{m+i} & \text{if } i \leq m \\ - \sum_{1 \leq j,k \leq m} y_{m+j} y_{m+k} \Gamma_{j,k}^i(y_1, \ldots, y_m) & \text{if } i > m. \end{cases}$$

(b) For all $k, i, \{s_1, \ldots, s_k\}, y \in B_\epsilon^m \times B_1^m$ we have that

$$|\partial^k V_i / \partial y_{s_1} \ldots \partial y_{s_k}(y)| < 3^k m^2 \left( \sum_{i=0}^{k} C_i \right) + 1.$$

(c) For all $i$ and all $y \in B_\epsilon^m \times B_1^m$ we have that

$$|V_i(y)| \leq \begin{cases} |y_{m+i}| & \text{if } i \leq m \\ m^2 C_0(y_{m+1}^2 + \ldots + y_{2m}^2) & \text{if } i > m. \end{cases}$$

Note that in Parts (b) and (c) above we have used the identification $T \mathbb{R}^m = \mathbb{R}^m \times \mathbb{R}^m$ — arising from the coordinates $(y_1, \ldots, y_m, y_{m+1}, \ldots, y_{2m})$ on $T \mathbb{R}^m$ — to make sense of “$y \in B_\epsilon^m \times B_1^m$.”
We integrate the vector field $V$ to get a partial flow $\psi : S \to T\mathbb{B}_{\epsilon}^m$, where $S \subset \mathbb{R} \times T\mathbb{B}_{\epsilon}^m$ is the maximal subset on which this partial flow makes sense. Note it follows from A.1.2.1(b) and (c) that the next two properties hold for $\psi$:

**A.1.2.4.**

(a) $[-2, 2] \times (B_{\epsilon/4}^m \times B_{\epsilon/4}^m) \subset S$.

(b) $\psi([-2, 2] \times (B_{\epsilon/4}^m \times B_{\epsilon/4}^m)) \subset B_{\epsilon}^m \times B_{\epsilon/3}^m$.

[To see this notice that $\psi$ is the geodesic flow on $T\mathbb{B}_{\epsilon}^m$ relative to the Riemannian metric $g_{i,j}$ on $\mathbb{B}_{\epsilon}^m$.]

Now we can deduce from A.1.2.3(b) and A.1.2.4(b) that the partial flow $\psi$: $[-2, 2] \times (B_{\epsilon/4}^m \times B_{\epsilon/4}^m) \to T\mathbb{B}_{\epsilon}^m$ satisfies:

**A.1.2.5.**

$|\partial^k \psi/\partial y_{s_1} \cdots \partial y_{s_k}(t, y)| \leq D_k$

for all $k$, $(t, y) \in [-2, 2] \times (B_{\epsilon/4}^m \times B_{\epsilon/4}^m)$, where $||$ denotes the Euclidean norm (with respect to the coordinates $y_1, y_2, \ldots, y_{2m}$). Here $D = \{D_i\}$ is collection of positive numbers which depend only on $C = \{C_i\}$ and $m = \dim M$.

To complete the proof for Corollary A.1.2 recall that $\psi : S \to T\mathbb{B}_{\epsilon}^m$ is the (partial) geodesic flow for the Riemannian metric $g_{i,j}$ on $\mathbb{B}_{\epsilon}^m$. Thus the composition $B_{\epsilon/4}^m = 0 \times B_{\epsilon/4}^m \subset B_{\epsilon/4}^m \times B_{\epsilon/4}^m \overset{\psi_1}{\to} T\mathbb{B}_{\epsilon}^m \overset{\text{proj.}}{\to} B_{\epsilon}^m \overset{f}{\to} M$ is equal to a composition $B_{\epsilon/4}^m \overset{L}{\to} TM_\rho \overset{\exp}{\to} M,$

where $L$ is a linear isometry and $\exp$ is the exponential map and $f$ is given in A.1.1 and $\psi_1$ is the composition $B_{\epsilon/4}^m \times B_{\epsilon/4}^m = 1 \times B_{\epsilon/4}^m \times B_{\epsilon/4}^m \subset S \overset{\psi}{\to} T\mathbb{B}_{\epsilon}^m$.

Thus it follows from A.1.1(b) and from A.1.2.5 that $\exp \circ L$ satisfies:

**A.1.2.6.**

$|\partial^k \bar{g}_{i,j}/\partial x_{s_1} \cdots \partial x_{s_k}(x)| \leq \mathcal{B}_k$

for all $k$ and all $x \in B_{\epsilon/4}^m$, where the $\mathcal{B} = \{\mathcal{B}_i\}$ is a set of positive numbers which depend only on the $B = \{B_i\}$ and $m = \dim M$, and the $\bar{g}_{i,j} : T\mathbb{B}_{\epsilon}^m \to \mathbb{R}$ are defined by

$\bar{g}_{i,j}(x) = g(D(\exp \circ L)(\partial/\partial x_i(x)), D(\exp \circ L)(\partial/\partial x_j(x))).$

[Recall that $B_{\epsilon/4}^m = 0 \times B_{\epsilon/4}^m \subset B_{\epsilon/4}^m \times \mathbb{R}^m$, and $B_{\epsilon/4}^m \times \mathbb{R}^m$ has coordinates $y_1, \ldots, y_{2m}$. Here we have renamed the coordinates $y_{m+1}, y_{m+2}, \ldots, y_{2m}$]
for $\mathbb{R}^m$ to be $x_1, x_2, \ldots, x_m$. If we replace (in A.1.1) the $f, B, g_{i,j}, \epsilon$ by $\tilde{f} = \exp \circ L, \tilde{B}, g_{i,j}, \epsilon = \epsilon/4$, then A.1.2.6 assures that Property A.1.1(b) is still satisfied for the “barred” quantities; Property A.1.1(a) is satisfied by the “barred” quantities because $\tilde{f}$ (restricted to a small neighborhood for $0 \in B_\epsilon^\mu$) is a normal coordinate system at $p$.

This completes the proof for Corollary A.1.2.

□

Proof of Corollary A.1.3. Note that all the steps in the proof for A.1.2 apply to the immersions $f, f' : B_\epsilon^m \to M$ given in the hypothesis of A.1.3. In particular Properties A.1.2.1(b) and (c) hold for the functions $\{g_{i,j}\}$ and $\{g'_{i,j}\}$ associated to $f$ and $f'$ respectively by A.1.1. For any $y \in B_\epsilon^m$ let $h_y : [0, 1] \to B_\epsilon^m$ be defined by $h_y(t) = (1 - t)x + ty$, $t \in [0, 1]$; note that $\text{length}(h_y) = |y - x| < 2\epsilon/9$, where the length and the norm are computed in Euclidean metric. Now it follows from A.1.2.1(b) and (c) as applied to both the $\{g_{i,j}\}$ and the $\{g'_{i,j}\}$, and from the fact that both $f$ and $f'$ are immersions, and from the fact that $x' \in B_{\epsilon/9}^\mu$, that there is a smooth path $h_y' : [0, 1] \to B_\epsilon^m$ which satisfies the following properties:

A.1.3.1. (a) $f \circ h_y(t) = f' \circ h_y'(t)$ for all $t \in [0, 1]$.

(b) $\text{length}(h_y') < \text{length}(h_y) + \epsilon^2 m^2 (10m)^{-m}$, where the lengths are computed in the Euclidean metric.

Now we can define $h : B_{\epsilon/9}^m \to B_\epsilon^m$ by:

A.1.3.1. (c) $h(y) = h_y'(1)$

for all $y \in B_{\epsilon/9}^m$. We leave as an exercise for the reader to determine that $h$ is a smooth embedding uniquely determined by the properties listed in A.1.3.1.

In order to verify Properties A.1.3(b) and (c) for $h : B_{\epsilon/9}^m \to B_\epsilon^m$ we will need another description of this map. Let $\psi : S \to TB_\epsilon^m$ and $\psi' : S' \to TB_\epsilon^m$ denote the (partial) geodesic flows for $g_{i,j}^\mu dx_i dx_j$ and $g'_{i,j}^\mu dx_i dx_j$ respectively. We have by A.1.2.4 that $[-2, 2] \times (B_{\epsilon/4}^m \times B_{\epsilon/4}^m) \subset S$ and $[-2, 2] \times (B_{\epsilon/4}^m \times B_{\epsilon/4}^m) \subset S'$. Define mappings $r : B_{\epsilon/4}^m \to B_\epsilon^m$ and $r' : B_{\epsilon/4}^m \to B_\epsilon^m$ to be the compositions:

A.1.3.2. (a)
respectively. Define the linear map $L : \mathbb{R}^m \to \mathbb{R}^m$ to be the composition:

\[ R^m = TB^m_{\epsilon/4|x} \xrightarrow{Df} TM_{f(x)}(Df')^{-1} \xrightarrow{T} B^m_{\epsilon/4|x'} = \mathbb{R}^m, \]

where the identifications $R^m = TB^m_{\epsilon/4|x}$ and $TB^m_{\epsilon/4|x'} = \mathbb{R}^m$ are induced from the standard coordinates $x_1, x_2, \ldots, x_m$ for $B^m_{\epsilon/4}$. Note that the diffeomorphism $h : B^m_{\epsilon/9} \to B^m_{\epsilon}$ defined in A.1.3.1(c) is also given by the following formula:

\[ h = r' \circ L \circ r^{-1} \mid B^m_{\epsilon/9}. \]

[Note that $r, r'$ are essentially the exponential maps for $g_{i,j}dx_idx_j$ and $g'_{i,j}dx_idx_j$ respectively. Thus Image $(r)$ is the ball $B_{\epsilon/4,x}$ of radius $\epsilon/4$ centered at $x$ in $B^m_{\epsilon/4}$ computed with respect to the metric $g_{i,j}dx_idx_j$. Note that A.1.2.1(b) and (c) (applied just to the $\{g_{i,j}\}$), together with the fact that $x \in B^m_{\epsilon/9}$, assure us that:

(i) $B^m_{\epsilon/9} \subset B_{\epsilon/4,x}$.

Thus $r^{-1} : B^m_{\epsilon/9} \to B^m_{\epsilon/4}$ is well-defined. In addition A.1.2.1(b) and (c) (applied to both the $\{g_{i,j}\}$ and the $\{g'_{i,j}\}$), assure us that:

(ii) $L \circ r^{-1}(B^m_{\epsilon/9}) \subset B^m_{\epsilon/4}$.

Thus (by (i) and (ii)) we have that $r' \circ L \circ r^{-1}$ is well defined on $B^m_{\epsilon/9}$ (the domain of $h$); this is a necessary condition for A.1.3.2 to hold.]

Note that Properties A.1.3(b) and (c) follow easily from A.1.3.2(d) and the following two claims:

**Claim A.1.3.3.** For every $v \in \mathbb{R}^m$ we have that

\[ (1 - O(\epsilon^2))|v| \leq |L(v)| \leq (1 + O(\epsilon^2))|v|, \]

where $| \cdot |$ denotes the Euclidean norm.
Claim A.1.3.4. Let \( \tau, \tau' : \mathbb{R}^m \to \mathbb{R}^m \) denote the translations of Euclidean space which satisfy \( \tau(0) = x \) and \( \tau'(0) = x' \). There is a collection of positive numbers \( E = \{ E_i \} \) which depend only on \( m = \dim M \) and on the \( B = \{ B_i \} \) of A.1.1, such that the following properties hold:

(a) \( \| D(\tau - \tau')_z \| < O(\epsilon^2) \) and \( \| D(\tau - \tau')_z \| < O(\epsilon^2) \) hold for all \( z \in B'_m \).

(b) \( \| D_{ij}(\tau - \tau')_z \| < E_i \) and \( \| D_{ij}(\tau - \tau')_z \| < E_i \) for all \( i, j \) and all \( z \in B'_m \).

Verification of Claim A.1.3.3.

The key idea here is to improve upon Properties A.1.2.1(b) and (c) for both the \( \{ g_{ij} \} \) and the \( \{ g'_{ij} \} \) (cf. A.1.3.3.2 below). Since \( f, f' : B_m \to M \) are (when restricted to a sufficiently small neighborhood for \( 0 \in B_m \)) normal coordinates for \( M \) near \( p, p' \), respectively, it follows that:

A.1.3.3.1. \( \partial g_{ij} / \partial x_k(0) = 0 = \partial g'_{ij} / \partial x_k(0) \) for all \( i, j, k \).

Property A.1.3.3.1, together with Properties A.1.1(a) and (b) for both the \( \{ g_{ij} \} \) and the \( \{ g'_{ij} \} \), imply that:

A.1.3.3.2. \( |g_{ij}(z) - \delta_{ij}^j| \leq O(\epsilon^2) \) and \( |g'_{ij}(z) - \delta_{ij}^j| \leq O(\epsilon^2) \)

for all \( i, j \) and all \( z \in B'_m \). By reviewing the definition given for the \( \{ g_{ij} \} \) and the \( \{ g'_{ij} \} \) just prior to A.1.1, we see it follows from A.1.3.2(c) that:

A.1.3.3.3. \( \sum_{i,j} g_{ij}(x)v_i v_j = \sum_{i,j} g'_{ij}(x')L(v)_i L(v)_j \)

for all \( v \in \mathbb{R}^m \), where \( (v_1, v_2, \ldots, v_m) \) and \( (L(v)_1, L(v)_2, \ldots, L(v)_m) \) denote the standard coordinates for \( v \) and \( L(v) \). Now Claim A.1.3.3 can be easily deduced from A.1.3.3.2 and A.1.3.3.3.

Verification of Claim A.1.3.4.

Note that A.1.3.4(b) can easily be deduced from A.1.2.5 (as applied to both \( \psi, \psi' \)) and from A.1.3.2(a) and (b).

Now we will verify A.1.3.4(a); the key idea here is to improve upon Property A.1.2.5 in the special case that \( k = 2 \) in A.1.2.5 (cf. A.1.3.4.4 below). Towards this end we first note it follows from A.1.3.3.1 that:

A.1.3.4.1. \( \Gamma^i_{j,k}(0) = 0 = \Gamma'^i_{j,k}(0) \)

for all \( i, j, k \), where \( \Gamma^i_{j,k} \) and \( \Gamma'^i_{j,k} \) are the Christoffel functions associated to the \( g_{ij}dx_idx_j \) and \( g'_{ij}dx_idx_j \). We deduce from A.1.2.2 (as applied to both \( \Gamma^i_{j,k} \) and \( \Gamma'^i_{j,k} \)) and from A.1.3.4.1, that:

A.1.3.4.2. \( |\Gamma^i_{j,k}(z)| \leq O(\epsilon) \) and \( |\Gamma'^i_{j,k}(z)| \leq O(\epsilon) \)
for all $i,j,k$ and all $z \in B^m$. Now let $V$ and $V'$ denote the vector fields on $TB^m_\epsilon$ defined in A.1.2.3(a) (where $\Gamma^{i}_{j,k}$ replaces $\Gamma^{i}_{j,k}$ in A.1.2.3(a) to define $V'$). Then it follows from A.1.2.3(a) and from A.1.3.4.2 that:

A.1.3.4.3. $|\partial^2 V/\partial y_{i_1} \partial y_{i_2}(y)| \leq O(\epsilon)$ and $|\partial^2 V'/\partial y_{i_1} \partial y_{i_2}(y)| \leq O(\epsilon)$

for all $(i_1,i_2)$ and all $(y_1,\ldots,y_{2m})$ are the standard coordinates for $TB^m_\epsilon$ referred to in A.1.2.3 and where $||$ indicates the Euclidean norm for vectors. Since we integrate $V$ and $V'$ to get the (partial) geodesic flows $\psi$, and $\psi'$, we may deduce the following inequalities from those in A.1.3.4.3:

A.1.3.4.4. $|\partial^2 \psi/\partial y_{i_1} \partial y_{i_2}(t,y)| \leq O(\epsilon)$ and $|\partial^2 \psi'/\partial y_{i_1} \partial y_{i_2}(t,y)| \leq O(\epsilon)$

for all $(i_1,i_2)$ and all $(t,y) \in [-2,2] \times (B^m_\epsilon \times B^m_\epsilon)$. To complete the verification for A.1.3.4(a) note that the following equalities are a consequence of the definition given for $r,r'$ in A.1.3.2(a) and (b):

A.1.3.4.5. $D(r-r')|_0 = 0 = D(r'-r')|_0$.

Now Claim A.1.3.4(a) is easily deduced from A.1.3.2(a) and (b), A.1.3.4.4, and A.1.3.4.5.

This completes the proof for Corollary A.1.3. □

Proof of Theorem A.1.4. For each smooth path $p : [0,1] \rightarrow B^m_\epsilon$ let $P_{p,g}$ and $P_{p,e}$ denote parallel translation along $p$ with respect to the metrics $g = g_{i,j} dx_i dx_j$ and $e = \delta_{i,j} dx_i dx_j$ respectively. The following relation between $P_{p,g}$ and $P_{p,e}$ is an easy consequence of Property A.1.3.4.2 (as applied to the Christoffel functions $\Gamma^{k}_{i,j}$ of $g_{i,j} dx_i dx_j$).

A.1.4.1. For each smooth path $p : [0,1] \rightarrow B^m_\epsilon$ and each vector $v \in TB^m_{\epsilon|p(0)}$ we have

$$|P_{p,g}(v) - P_{p,e}(v)|_e < \text{length}_e(p)O(\epsilon).$$

Now Theorem A.1.4 is a direct consequence of Property A.1.4.1, and of Property A.1.3.3.2 (as applied to the $g_{i,j}$), and of the hypothesis for A.1.4 that $\|Dr_q\| < O(1)$ for all $q \in \tilde{U}$.

This completes the proof for Theorem A.1.4. □

Proof of Theorem A.1.6. Note from the hypothesis of A.1.6 (that $r$ satisfies 0.2.2(a)-(c)) that:

A.1.6.1. (a) diameter$(U) < O(\delta)$.

Thus, by choosing $p \in U$ and choosing $\delta \ll \epsilon$, we may assume that:
A.1.6.1. (b) $U \subset f_p(B^m_{\epsilon/3})$.

In particular A.1.4 applies to $r$ and $\hat{r}$, where $\hat{r} : \hat{U} \to \hat{B}$ was defined just prior to A.1.4. We also want to apply A.1.7 to $s$ and $\hat{s}$, where $\hat{s} : \hat{E} \to \hat{B}$ was defined just prior to A.1.7; recall that $\hat{E} = f_p^{-1}(E)$.

The remainder of this proof is based upon the following claim, which we will verify at the end of this proof. In this claim we set $k = \dim \hat{U}$, and for any integer $i > 0$ and any number $t > 0$ we let $B^i_t$ denote the open ball of radius $t$ centered at the origin of $\mathbb{R}^i$.

**Claim A.1.6.2.** Let $\hat{Y} = \hat{U} \cap B^m_{\delta^3}$. There is a smooth embedding $h : Y \times B^m_{\delta^3} \to B^m_6$, where $Y$ is an open subset of $B^6_6$ containing 0, which satisfies the following properties:

(a) $h(0,0) = 0$, and $Dh(0,0)$ is an isometry with respect to the Euclidean metrics.

(b) $\|Dh(x,y) - Dh(x',y')\| < \mathcal{O}(\epsilon)$ for all $(x,y),(x',y') \in Y \times B^m_{\delta^3}$. (Here the norm $\|\|$ is computed with respect to the Euclidean metrics.)

(c) $h(Y,0) = Y$ and $h(z \times B^m_{\delta^3}) = \hat{\rho}^{-1}(h(z,0))$ for all $z \in Y$, where

$$\hat{\rho} : \hat{E} \cap B^m_6 \to \hat{U}$$

denotes the orthogonal projection with respect to the Riemannian metric $g_{i,j}$ of A.1.1.

To complete the proof for Theorem A.1.6 we first note we have as an immediate consequence of A.1.6.2(a) and (b) the following properties. Let $\rho : Y \times B^m_{\delta^3} \to Y$ denote the standard projection onto the first factor.

A.1.6.3. (a) $\|D(\rho \circ h^{-1})_q\| < 1 + \mathcal{O}(\epsilon)$ for all $q \in h(Y \times B^m_{\delta^3})$.

(b) $\|D(\rho \circ h^{-1})_q - D(\rho \circ h^{-1})_{q'}\| < \mathcal{O}(\epsilon)$ for all $q,q' \in h(Y \times B^m_{\delta^3})$.

Note also that it follows from A.1.6.2(a) and (b) as applied to $h|Y \times 0$ (see also A.1.6.2(c)), and from 0.2.2(a)-(c) as applied to $r$, and from A.1.1 and A.1.4 as applied to $r$ and $\hat{r}$, that:

A.1.6.3. (c) $\|D(\hat{r} \circ h \circ I)_z\| < 1 + \mathcal{O}(\epsilon)$ for all $z \in Y$,

(d) $\|D(\hat{r} \circ h \circ I)_z - D(\hat{r} \circ h \circ I)_{z'}\| < \mathcal{O}(\epsilon)$ for all $z,z' \in Y$,

where $I : Y \to Y \times 0$ sends $z$ to $(z,0)$. Finally we note that A.1.6.2(c) implies that:

A.1.6.3. (e) $(\hat{r} \circ h \circ I) \circ (\rho \circ h^{-1}) = \hat{s} | h(Y \times B^m_{\delta^3})$.

It follows from A.1.6.3(a)-(e) that:

A.1.6.4. (a) $\|D\hat{s}_q\| < 1 + \mathcal{O}(\epsilon)$ for all $q \in h(Y \times B^m_{\delta^3})$
(b) \[ \| D \hat{s}_q - D \hat{s}_{q'} \| < \mathcal{O}(\epsilon) \text{ for all } q, q' \in h(Y \times B^m_{\delta^-k}). \]

We may rephrase A.1.6.4(b) as follows. (See Remark A.1.7 for notation.)

**A.1.6.4.** (c) \[ K(\hat{s} \mid X; B^m_{\epsilon}, \delta) < \mathcal{O}(\epsilon)\delta^{-1} \text{ where } X = h(Y \times B^m_{\delta^-k}). \]

Note that it follows from A.1.6.2(a), (b) and (c), and from A.1.3.3.2 as applied to \( g_{i,j} \), and from 0.2.2(a)-(c) as applied to \( r \), that the subset \( X \subset \hat{E} \) of A.1.6.4(c) has the following desirable lifting property:

**A.1.6.4.** (d) For each smooth path \( f : [0, 1] \to E \) having length less than or equal to \( \delta \), there is a smooth path \( \hat{f} : [0, 1] \to X \) such that \( f \circ \hat{f} = f \).

Now the conclusion of A.1.6 follows immediately from Property A.1.6.4(c) and Lemma A.1.7: Note that Properties A.1.6.4(a) and (d), and Property A.1.3.3.2 as applied to \( g_{i,j} \) assure that the hypotheses of Lemma A.1.7 hold.

**Verification of Claim A.1.6.2.**

Let \( V \) denote the tangent plane to \( \hat{U} \) at \( 0 \in \hat{U} \), and choose an isometry \( g_1 : \mathbb{R}^k \to V \) with \( g_1(0) = 0 \) (here \( V \) has the metric inherited from the Euclidean metric on \( B^m_{\epsilon} \)). For sufficiently small \( \delta \) there is a unique smooth embedding \( g_2 : Y \to B^m_{\epsilon} \) which satisfies the following properties:

**A.1.6.2.1.**

(a) We have that \( g_2(Y) = \hat{Y} \) and, for all \( z \in Y \), \( g_2(z) - g_1(z) \) is perpendicular to \( V \). (To make sense of the difference \( g_2(z) - g_1(z) \) we identify \( g_1(z) \) with its image in \( B^m_{\epsilon} \) under the Euclidean-exponential map.)

(b) \( g_2(0) = 0 \); and \( Dg_2 \mid T(B^k_{\delta \epsilon}) = g_1 \) when \( T(B^k_{\delta \epsilon}) \) is identified with \( \mathbb{R}^k \) by the Euclidean exponential map.

(c) \( \| D^2 (g_2)_q \| < \mathcal{O}(\epsilon)\delta^{-1} \) for all \( q \in Y \).

We note that \( Y \) and \( g_2 \) are uniquely determined from \( g_1 \) by Property A.1.6.2.1(a); and Property A.1.6.2.1(b) is a direct consequence of A.1.6.2.1(a) and of the definition for \( g_1 \). Note that Property A.1.6.2.1(c) is equivalent to the inequality:

(i) \[ K(\text{Image}(g_2); B^m_{\epsilon}) < \mathcal{O}(\epsilon)\delta^{-1}; \]

and, since \( \text{Image}(g_2) \) is an open subset of \( \hat{U} \), this last inequality is implied by:

(ii) \[ K(\hat{U}; B^m_{\epsilon}) < \mathcal{O}(\epsilon)\delta^{-1}. \]

(Here for any smooth submanifold \( S \subset B^m_{\epsilon} \) we denote by \( K(S; B^m_{\epsilon}) \) the curvature of the constant map \( c : S \to \{1\} \) with respect to the Euclidean metric on \( B^m_{\epsilon} \) as described in 0.1.) Finally we note that inequality (ii) is a consequence of 0.2.2(a) and A.1.4. (Note that Properties 0.2.2(a)-(c) are
assumed in A.1.6; and A.1.4 may be applied here because its hypotheses are a consequence of 0.2.2(a)-(c) and A.1.3.3.2.)

Now we define smooth vector fields \(v_1, v_2, \ldots, v_{m-k}\) on \(\hat{Y} = g_2(Y)\) as follows: Let \(T \hat{U} \perp\) denote the orthogonal complement for \(T \hat{U}\) in \(T(B^m_\epsilon)\) with respect to the metric \(\{g_{i,j}\}\); and for each \(z \in Y\) let \(p_z : T(B^m_\epsilon)_{g_2(z)} \to T \hat{U} \perp_{g_2(z)}\) denote the orthogonal projection with respect to the metric \(\{g_{i,j}\}\). Choose the \(\{v_i(0)\}\) to be an orthonormal basis for \(T \hat{U} \perp_0\) with respect to the metric \(\{g_{i,j}\}\); and then define \(\{v_1(z), v_2(z), \ldots, v_{m-k}(z)\}\) to be the orthonormal subset of \(T \hat{U} \perp_{g_2(z)}\) obtained by applying the Gramm-Schmidt process to the set of vectors \(\{p_z \circ P_z(v_1(0)), p_z \circ P_z(v_2(0)), \ldots, p_z \circ P_z(v_{m-k}(0))\}\) with respect to the metric \(\{g_{i,j}\}\), where \(P_z\) denotes Euclidean parallel translation to \(T(B^m_\epsilon)_{g_2(z)}\). Note that \(T(B^m_\epsilon)\) is identified with \(B^m_\epsilon \times \mathbf{R}^m\) by using Euclidean parallel translation to identify each \(T(B^m_\epsilon)_{y}\), \(y \in B^m_\epsilon\), with \(T(B^m_\epsilon)_0\), and by identifying \(T(B^m_\epsilon)_0\) with \(\mathbf{R}^m\) via the exponential map \(\exp : T(B^m_\epsilon) \to \mathbf{R}^m\). Thus we may define maps \(w_i : Y \to \mathbf{R}^m\) (for \(1 \leq i \leq m-k\)) by \((g_2(z), w_i(z)) = v_i(g_2(z))\) for all \(z \in Y\). Note that it follows from the construction just given, and from Properties A.1.6.2.1(a)-(c), and from Properties A.1.1(a) and (b), that the maps \(\{w_i\}\) satisfy the following properties:

**A.1.6.2.2.** (a) For each \(z \in Y\) the vectors \(\{(g_2(z), w_i(z)) : 1 \leq i \leq m-k\}\) are an orthonormal basis for \(T \hat{U} \perp\) with respect to the metric \(\{g_{i,j}\}\).

(b) \(\|D(w_i)z\| < O(\epsilon)\delta^{-1}\) (where the norm \(\|\|\) is computed with respect to the Euclidean metrics on \(B^k_{6\delta}\) and on \(T(B^m_\epsilon)\)).

We can now define a map \(g_3 : Y \times B^m_{m-k} \to T(B^m_\epsilon) = B^m_\epsilon \times \mathbf{R}^m\) by:

**A.1.6.2.3.** \(g_3(z, a) = \left( g_2(z), \sum_{i=1}^{m-k} a_i w_i(z) \right) \)

for all \((z, a) \in Y \times B^m_{m-k}\), where \((a_1, \ldots, a_{m-k})\) are the standard coordinates for \(a \in B^m_{m-k}\). It follows from A.1.6.2.1-A.1.6.2.3, and from A.1.1(a) and (b), that \(g_3\) is a smooth embedding which satisfies the following properties. (See in particular A.1.6.2.1(b) and A.1.1(a) for Part (a) below.)

**A.1.6.2.4.** (a) \(g_3(0, 0) = (0, 0)\); and \(D(g_3)_{(0,0)}\) is a linear isometry with respect to the Euclidean metrics.

(b) \(\|D(g_3)(z,a) - D(g_3)(z',a')\| < O(\epsilon)\) for all \((z, a)\) and \((z', a')\) in \(Y \times B^m_{m-k}\). (Here the norm \(\|\|\) is computed with respect to the Euclidean metrics on \(B^k_{6\delta} \times B^m_{m-k}\) and \(T(B^m_\epsilon) = B^m_\epsilon \times \mathbf{R}^m\).)
(c) \( g_3(z,0) = (g_2(z),0) \), so \( g_3(z,0) \in \hat{U} \) for all \( z \in Y \); moreover \( g_3(z \times B^m_{\delta-k}) = \pi^{-1}(g_3(z,0)) \) for all \( z \in Y \), where \( \pi : (T\hat{U})^\delta \to \hat{U} \) denotes the usual bundle projection and \((T\hat{U})^\delta \) denotes all vectors \( v \in T\hat{U}^\delta \) whose \( \{g_{ij}\}\)-length is less than \( \delta \).

Now we can finally define the map \( h : Y \times B^m_{\delta-k} \to B^m_{\epsilon} \) of A.1.6.2 to be the composition:

\[
A.1.6.2.5. \quad Y \times B^m_{\delta-k} \xrightarrow{g_3} T(B^m_{\epsilon}) = B^m_{\epsilon} \times \mathbb{R}^m \supset B^m_{\epsilon/4} \times B^m_{\epsilon/4} \xrightarrow{\exp} B^m_{\epsilon},
\]

where “exp” denotes the exponential map with respect to the metric \( \{g_{ij}\} \) on \( B^m_{\epsilon} \). It follows from A.1.1(a) and (b) that “exp” is defined on \( B^m_{\epsilon/4} \times B^m_{\epsilon/4} \) (see also A.1.2.4, and note that \( \exp = \psi_1 \)). And, by choosing \( \delta \ll \epsilon \), it follows from the definition of \( g_3 \) in A.1.6.2.3 that \( \text{Image}(g_3) \subset B^m_{\epsilon/4} \times B^m_{\epsilon/4} \) (see also A.1.6.2.1). Thus \( h \) is well-defined by the preceding composition. Note that Property A.1.6.2(a) is implied by A.1.6.2.4(a) and A.1.6.2.3 and A.1.6.2.5. Note also that Property A.1.6.2(c) is implied by A.1.6.2.4(c) and A.1.6.2.5. Finally we note that Property A.1.6.2(b) is implied by A.1.6.2.4(a) and (b) and A.1.6.2.5.

This completes the verification of Claim A.1.6.2; thus also the proof of Theorem A.1.6 is completed.

\[
\square
\]

Appendix 2.

In this appendix we review the results for the Collapsing theory of Cheeger-Fukaya-Gromov (cf. [4]) which we need to carry out the proof of Theorem 0.3 in §3 above (cf. A.2.2-A.2.5 below).

Let \( M, g_M \) denote an \( A \)-regular complete Riemannian manifold, where \( A \) denotes a collection \( \{A_i : i = 0,1,2,\ldots\} \) of positive numbers. For any \( \delta > 0 \) we shall say that \( g_M \) is \( \delta \)-round at the point \( p \in M \) if there is an open neighborhood \( V \subset M \) for \( p \in M \), and a regular covering \( \pi : \hat{V} \to V \) for \( V \) with covering group \( \Lambda \), such that the following properties hold. Note that \( \hat{V} \) is equipped with a Riemannian metric \( g_{\hat{V}} \) gotten by pulling back \( g_M \) along \( \pi : \hat{V} \to V \).

A.2.1. (a) \( V \) contains the ball \( B(p,\delta) \) of radius \( \delta \) centered at \( p \in M \).

(b) The injectivity radius of \( \hat{V} \) at all points of \( \pi^{-1}(B(p,\delta)) \) is greater than \( \delta \).

(c) There is a virtually nilpotent Lie group \( H \) and an effective isometric action \( H \times \hat{V} \to \hat{V} \) extending that of \( \Lambda \times \hat{V} \to \hat{V} \). (Recall that \( H \times \hat{V} \to \hat{V} \) is effective if only the identity element of \( H \) fixes all points of \( \hat{V} \).)
Note that $\pi(Hq)$, for $q \in \hat{V}$, gives the orbits of the isometries in the collapsed directions.

We say that $g_M$ is $\delta$-round if it is $\delta$-round at every $p \in M$.

The next theorem is a direct consequence of [4, 1.3]. Let $\nabla$ denote the Levi-Civita connection for $g_M$ and let $\nabla'$ denote the Levi-Civita connection for the Riemannian metric $g'_M$ given in the following Theorem:

**Theorem A.2.2** (Cheeger-Fukaya-Gromov). Given any $\epsilon > 0$, any integer $m > 0$, and any collection of positive numbers $A = \{A_i \mid i = 0, 1, 2, \ldots\}$, there exists a positive number $\eta(\epsilon, m, A)$. For any $\delta \in (0, \eta(\epsilon, m, A))$, and for any $A$-regular complete Riemannian manifold $(M, g_M)$ of dimension $m$, there exists a $\delta$-round metric $g'_M$ such that the following hold:

(a) $e^{-\epsilon}g_M < g'_M < e^{\epsilon}g_M$.

(b) $|\nabla - \nabla'| < \epsilon$.

(c) $M, g'_M$ is $A'$-regular, where $A' = \{A'_i : i = 0, 1, 2, 3, \ldots\}$ denotes a collection of positive numbers which depends only on $\epsilon, m$, and $A = \{A_i : i = 0, 1, 2, 3, \ldots\}$. (In fact, $A'_i$ depends only on $A_0$, $i$, $\epsilon$ and $m$)

Note that for fixed $A = \{A_i\}$, there are sequences of $A$-regular manifolds whose injectivity radius approaches 0; so the cover of $V$ by $\hat{V}$ is nontrivial. Thus $\Lambda$ and $H$ are nontrivial as well and the collapsing can be seen.

In the next theorem we will need the following notation: Let $\rho : E \to \hat{V}$ denote the bundle of orthonormal frames over $\hat{V}$ with respect to the metric $g'_E$ (the pull back of $g'_M$ along $\pi : \hat{V} \to V$). $E$ is given a Riemannian metric structure $g'_E$ of bounded geometric type as follows: For any $v_i \in TE$, $i = 1, 2$, choose paths $\alpha_i : [-1, 1] \to E$, $i = 1, 2$, such that $\alpha_i(0) = v_i$ holds; each $\alpha_i$ may be regarded as a path $\beta_i : [-1, 1] \to \hat{V}$ together with an orthonormal frame $(w_{i,1}(t), w_{i,2}(t), \ldots, w_{i,m}(t))$ along $\beta_i(t)$, $t \in [-1, 1]$. Let $V_{i,j}$ denote the covariant derivative of $w_{i,j}(t)$ at $t = 0$, and set

$$
\langle g'_E(v_1, v_2) = g'_E(\beta_1(0), \beta_2(0)) + \sum_{j=1}^{m} g'_E(V_{1,j}, V_{2,j}).
$$

Let $d'_E(\cdot, \cdot)$ denote the metric on $E$ associated to $g'_E$. Note that the canonical lifting $H \times E \to E$ of the group action $H \times \hat{V} \to \hat{V}$ is a free isometric group action (cf. A.2.1(c)). The following theorem is not explicitly stated in [4], so we shall derive it at the end of this appendix.

**Theorem A.2.3.** There exists a real number $B > 1$, which depends only on $n = \dim M, \epsilon$, and on the sectional curvature bound $A_0$ for $M$, such that the following is true. Suppose for $x \in E$ and $g \in H$ ($g \neq e$) we have that $d'_E(x, g(x)) < \delta$. Then there is a one-parameter subgroup $\phi : R \to H$ of $H$ satisfying:
(a) $\phi(t) = g$, for some $t \in (0, B\delta)$.

(b) The path $f : \mathbb{R} \to E$, defined by $f(s) = \phi(s)(x)$, has unit speed.

For any $y \in \hat{V}$ we set $H_y = \{ h \in H : h(y) = y \}$, i.e., $H_y$ is the isotropy group at $y$ for the action $H \times \hat{V} \to \hat{V}$. Choose $\hat{y} \in \rho^{-1}(y)$. Since the lifted action $H \times E \to E$ is a free action, it follows from A.2.3 that the map $f : H \to E$ given by $f(g) = g(\hat{y})$ is an embedding onto a closed subset of $E$. Since $H_y$ is a closed subset of $H$ it follows that the restricted map $f : H_y \to \rho^{-1}(y)$ is an embedding onto a closed subset of the compact space $\rho^{-1}(y)$. Thus we have proven the following corollary of Theorem A.2.3:

**Corollary A.2.4.** For each $y \in \hat{V}$ the isotropy group $H_y$ is a compact subgroup of $H$.

We shall also need the following theorem, which among other uses will be used in proving Theorem A.2.3. Recall that the notion of “local angle control” was defined in 2.1 of §2 above.

**Theorem A.2.5.** The action $H \times E \to E$ has local angle control equal to $(\lambda t, t)$ for any $t > 0$, where $\lambda$ depends only on $\dim E, A_0, \epsilon$.

**Remark A.2.6.** In fact, we will show in proving A.2.5 that $\lambda$ depends only on $\dim \hat{V}$ and the sectional curvature bound $A'_0$ for $\hat{V}$. Recall from A.2.2 that $A'_0$ depends only on $A_0, \epsilon$ and $n = \dim \hat{V}$. Also, that $\dim E = n(n + 1)/2$. Moreover, we will use no special facts about $\hat{V}$; only that $H$ acts effectively on $\hat{V}$ by isometries.

**Proof of Theorem A.2.5.** Let $x \in E$, $u \in TE_x$ and $\alpha_v : \mathbb{R} \to H$ be a one-parameter subgroup with $\dot{\alpha}_v(0) = v \in g(H)$, where $g(H)$ is the Lie algebra of $H$. Define a curve $\gamma : \mathbb{R} \to E$ by

$$\gamma(t) = \alpha_v(t)x$$

and a vector field $\dot{u}$ along $\gamma$ by

$$\dot{u}(t) = d\alpha_v(t)(u).$$

And let $\bar{D}$ denote the Levi-Civita connection on $E$. The argument above proving Lemma 2.3 yields that it suffices to show that $|\bar{D}_{\dot{\gamma}(0)} \dot{u}|$ is bounded above by a constant multiplied by $|\dot{\gamma}(0)||u|$ and this constant depends only on $n = \dim \hat{V}$ and $A'_0$ (see Remark A.2.6). Since $\dot{\gamma}(0) \neq 0$ when $v \neq 0$, in bounding $|\bar{D}_{\dot{\gamma}(0)} \dot{u}|$, we may assume (and do assume) that $|u| = |\dot{\gamma}(0)| = 1$ and need only consider the two cases: $u$ vertical and $u$ horizontal with
respect to the Riemannian submersion
\[ O(n) \longrightarrow E \overset{\rho}{\longrightarrow} \hat{V} \]
where \( n = \dim \hat{V} \).

Let us first consider the case where \( u \) is vertical. A continuity argument then shows that we need only consider the subcase where \( \rho \circ \gamma : \mathbb{R} \to \hat{V} \) is an immersion. To verify this subcase, we need the following lemma. Consider the submanifold \( S \) of \( E \) defined by
\[ S = \rho^{-1}(\text{image}(\rho \circ \gamma)). \]

**Lemma A.2.5.1.** Let \( B(\ , \ ) \) denote the second fundamental form of \( S \) in \( E \). The set \( \{|B(v,u)| : |u|,|v| \leq 1 \text{ and } u \text{ vertical}\} \) is bounded above by a positive real number which depends only on \( n \) and \( A'_0 \).

Before proving Lemma A.2.5.1, we use it to bound \( |D_{\tilde{\gamma}(0)} \hat{u}| \) when \( u \) is vertical. Note that the map \( \zeta : \mathbb{R} \times O(n) \to S \) defined by
\[ (t,g) \to \alpha_v(t)xg^{-1} = \gamma(t)g^{-1} \]
is either a bijective immersion or a covering projection depending on whether \( \rho \circ \gamma \) is monic or not. Use this map to put a left invariant Riemannian metric on \( \mathbb{R} \times O(n) \); in particular, an inner product \( \langle \ , \ \rangle \) on the Lie algebra \( g(\mathbb{R} \times O(n)) \). Since \( \langle \ , \ \rangle \) restricted to \( g(O(n)) \) is \(-1\) times the Killing form, it satisfies the equation:

**A.2.5.2.** \( \langle [a,b],a \rangle = 0 \)
valid for all \( a,b \in g(O(n)) \).

**Remark A.2.5.2.1.** When \( n = 1 \) or \( 2 \), instead of the negative of the Killing form, we mean the unique bi-invariant metric such that \( \text{Vol}(O(n)) = 2 \) or \( 4\pi \), respectively.

Let \( U,V \in g(\mathbb{R} \times O(n)) \) be the vectors corresponding to \( u,\gamma(0) \in TE_x \) via \( d\zeta \). And let \( \hat{U},\hat{V} \) be left invariant vector fields on \( \mathbb{R} \times O(n) \) determined by \( U,V \). Then note that
\[ \bar{D}_{\tilde{\gamma}(0)} \hat{u} = d\zeta(D_V \hat{U}) + B(\gamma(0),u) \]
where \( D \) is the Levi-Civita connection for \( \mathbb{R} \times O(n) \) and \( B(\ , \ ) \) is the second fundamental form of \( S \) in \( E \). Because of this equation and Lemma A.2.5.1, it remains to show the following to verify A.2.5 when \( u \) is vertical:
Lemma A.2.5.3. Let $T$ and $X$ be unit length left invariant vector fields on $\mathbb{R} \times O(n)$ tangent to $\mathbb{R} \times *$ and $* \times O(n)$, respectively. Then $|D_TX|$ is bounded above by a constant which depends only on $n$.

Proof of Lemma A.2.5.3. Identify the left invariant vector fields on $\mathbb{R} \times O(n)$ with the elements in its Lie algebra $g(\mathbb{R} \times O(n))$. Then orthogonally decompose $T$ and $D_TX$ as:

A.2.5.3.1. $D_TX = Y + R$ and $T = Z + F$

where $Y, Z \in g(O(n))$ and $R, F \in g(O(n))^\perp$. Recall [2, Prop. 7.7.1] which states that:

A.2.5.3.2. $\langle [D_a b, c] \rangle = 1/2([[b, a], c] + ([a, c], b) + ([b, c], a))$

for all $a, b, c \in g(\mathbb{R} \times O(n))$. And note that:

A.2.5.3.3. $[T, a] = 0$

for all $a \in g(O(n))$. Setting $a = T$, $b = X$, and $c = Y$ in A.2.5.3.2 and applying A.2.5.3.3, we obtain:

A.2.5.3.4. $\langle D_TX, Y \rangle = 1/2([X, Y], T)$.

Note also that the norm of the Lie bracket $[ \ , \ ]$ restricted to $g(O(n))$ is bounded above by a constant $r \geq 0$ which depends only on $n$ since $\langle \ , \ \rangle$ restricted to $g(O(n))$ is the negative of the Killing form; cf. Remark A.2.5.2.1. Consequently, A.2.5.3.1 and A.2.5.3.4 yield that:

A.2.5.3.5. $|Y| \leq r/2$.

Next, setting $a = T$, $b = X$, and $c = T$ in A.2.5.3.2 and applying A.2.5.3.3, we obtain:

A.2.5.3.6. $\langle D_TX, T \rangle = 0$.

Hence A.2.5.3.1 and A.2.5.3.6 yield:

A.2.5.3.7. $\langle D_TX, F \rangle = -\langle D_TX, Z \rangle$.

Setting $a = T$, $b = X$, and $c = Z$ in A.2.5.3.2 and applying A.2.5.3.3 yields:

A.2.5.3.8. $\langle D_TX, Z \rangle = (1/2)\langle [X, Z], T \rangle$.

And substituting Equation A.2.5.3.1 for $T$ into A.2.5.3.8 yields:

A.2.5.3.9. $\langle D_TX, Z \rangle = (1/2)\langle [X, Z], F \rangle$

since $Z \perp [X, Z]$ by A.2.5.2 above. Combining A.2.5.3.7 and A.2.5.3.9 yields:
\[ \langle DT X, F \rangle = -(1/2)\langle [X, Z], F \rangle. \]

Note that \( F \neq 0 \) since we assumed that \( \rho \circ \gamma : \mathbb{R} \to \hat{V} \) is an immersion. Therefore, \( R \) is a scalar multiple of \( F \) and A.2.5.3.10 consequently yields:

\[ \langle DT X, R \rangle = -(1/2)\langle [X, Z], R \rangle. \]

Now A.2.5.3.11 implies:

A.2.5.3.12. \( |R| \leq r/2 \)

in the same way that A.2.5.3.4 implies A.2.5.3.5. Combining the inequalities of A.2.5.3.12 and A.2.5.3.5 yields that \( |DT X| \leq r \) proving Lemma A.2.5.3. \( \square \)

To prove Lemma A.2.5.1 we need to calculate \( \bar{D} \). And for this purpose we define special horizontal and vertical vector fields on \( E \) as follows: A horizontal vector field \( X \) on \( E \) is special if there exists a vector field \( \hat{X} \) on \( \hat{V} \) such that \( d\rho (X_x) = X_{\rho(x)} \) for all \( x \in E \). Let \( g(O(n)) \) denote the Lie algebra of \( O(n) \). Each \( v \in g(O(n)) \) determines a special vertical vector field \( \hat{v} \) on \( E \) by requiring

\[ \hat{v}_x = \hat{\delta}_x(0) \]

for each \( x \in E \). Here, \( \delta_x(t) = x \beta_v(t) \) where \( \beta_v \) is the one-parameter subgroup of \( O(n) \) such that \( \beta_v(0) = v \). To calculate \( \bar{D} \) it suffices to calculate \( (\bar{D}_a b) \cdot c \), where \( a, b, c \) are the special vector fields just defined. A routine application of the Koszul formula yields the following result:

**Lemma A.2.5.4.** Let \( u, v, w \in g(O(n)) \) and \( X, Y, Z \) be vector fields on \( \hat{V} \). Also let \( B(\ , \ ) \) denote the negative of the Killing form on \( g(O(n)) \) and \( D \) denote the Levi-Civita connection on \( \hat{V} \); cf. Remark A.2.5.2.1. Then the following are valid equations:

(a) \( (\bar{D}_a \hat{v}) \cdot \hat{w} = (1/2)B([u, v], w) \);
(b) \( (\bar{D}_X Y) \cdot Z = (DX Y) \cdot Z \);
(c) \( (\bar{D}_a X) \cdot \hat{Y} = (\bar{D}_X \hat{u}) \cdot \hat{Y} = -(\bar{D}_X \hat{Y}) \cdot \hat{u} = (1/2)[\hat{Y}, X] \cdot \hat{u} \);
(d) \( (\bar{D}_a \hat{v}) \cdot \hat{X} = (\bar{D}_X \hat{u}) \cdot \hat{v} = (\bar{D}_u X) \cdot \hat{v} = 0 \).

**Proof of Lemma A.2.5.1.** It suffices to prove the assertion in the two cases: \( v \) vertical and \( v \) horizontal. The equation of A.2.5.4(d) shows that the fibers of \( \rho \) are totally geodesic submanifolds of \( E \); i.e., \( B(v, u) = 0 \) when both \( u \) and \( v \) are vertical. Therefore, it remains to estimate

\[ B(X, u) \cdot \hat{Y} \]
where $\overline{X}$ and $\overline{Y}$ are special horizontal vector fields on $E$ induced by vector fields $X$ and $Y$, respectively, on $\hat{V}$. (Also, $\overline{X}$ is assumed tangent to $S$ and $\overline{Y}$ perpendicular to $S$ at points of $S$.) In this situation, the equation of A.2.5.4(c) yields that

$$B(\overline{X},u) \cdot \overline{Y} = 1/2[\overline{Y},\overline{X}] \cdot u = \Omega(\overline{X},\overline{Y}) \cdot w$$

where $\Omega(\ ,\ )$ denotes the curvature form for the frame bundle of $\hat{V}$ and $w$ is the unique vector in $\mathfrak{g}(O(n))$ such that $\hat{\omega}_x = u$. Since the norm of the curvature tensor for $\hat{V}$ is bounded by $A_0^0$, we also get the desired upper bound for $|B(v,u)|$ in this case. (See [15, Vol. 1, p. 133] for the standard relation between $\Omega(\ ,\ )$ and the curvature tensor on $\hat{V}$.)

This completes the proof for Lemma A.2.5.1.

It remains, in proving Theorem A.2.5, to consider the case where $u$ is horizontal with respect to $\rho$. Since we have assumed that $|\gamma(0)| = |u| = 1$, it suffices to show that $|D_{\gamma(0)} \dot{u}|$ is bounded above by a real number which depends only on $n$ and $A_0^0$. Let $z = \rho(x)$ and define a curve $\beta(t)$ in $\hat{V}$ by

$$\beta(t) = \alpha_v(t)z.$$ \hspace{1cm} (Note that $\alpha_v \gamma(0)$ is the special vertical vector field induced by $w$.)

And note that $\beta = \rho \circ \gamma$. Decompose

$$\gamma(0) = v^+ + v^-$$

where $v^+$ is vertical and $v^-$ is horizontal relative to $\rho$. A continuity argument yields that it suffices to consider the situation where $v^- \neq 0$; hence, we now make this assumption. Also, note that

$$|v^+|, |v^-| \leq 1.$$ \hspace{1cm} (Recall $\hat{\omega}$ is the special vertical vector field induced by $w$.)

Since $\beta(0) = dp(v^-) \neq 0$, there is a vector field $Y$ on $\hat{V}$ such that the induced special horizontal vector field $\overline{Y}$ on $E$ satisfies

$$\overline{Y}(\gamma(t)) = \dot{u}(t)$$

for all $t$ sufficiently close to 0. Hence, it suffices to show that $|D_{v^+} \overline{Y}|$ and $|D_{v^-} \overline{Y}|$ are both bounded above by a real number which depends only on $n$, $A_0$, and $\epsilon$. Therefore it remains to estimate

$$(D_{v^+} \overline{Y}) \cdot \overline{Z} \quad \text{and} \quad (D_{v^-} \overline{Y}) \cdot \hat{\omega}$$

where $Z$ is a vector field on $\hat{V}$ with $|Z_{\beta(0)}| = 1$ and $w \in \mathfrak{g}(O(n))$ with $|w| = 1$. (Recall $\hat{\omega}$ is the special vertical vector field induced by $w$.) We will use Lemma A.2.5.4 to do these estimates. For this purpose, we need to fix some notation. Let $X$ be a vector field on $\hat{V}$ such that $X_{\beta(0)} = dp(v^-)$. (Note that $\overline{X}(\gamma(0)) = v^-$. ) Also, let $\omega \in \mathfrak{g}(O(n))$ be the vector satisfying

$$\hat{\omega}(\gamma(0)) = v^+.$$
Let \( u_0 = d\rho(u) \) and \( \hat{u}_0 \) denote the vector field along the curve \( \beta \) in \( \hat{V} \) defined by
\[
\hat{u}_0(t) = d\alpha_v(t)(u_0).
\]
Recall that \( D \) is the covariant derivative in \( \hat{V} \). And \( \Omega( , ) \) denotes (as before) the curvature form for the frame bundle of \( \hat{V} \). With this notation, Lemma A.2.5.4 yields the following calculation:

A.2.5.5. (a) \( (D_v \hat{Y}) \cdot \hat{w} = 0; \)
(b) \( (D_v \hat{Y}) \cdot \hat{w} = (1/2)[\hat{X},\hat{Y}] \cdot \hat{w} = \Omega(\hat{Y},\hat{X}) \cdot \hat{w}; \)
(c) \( (D_v \hat{Y}) \cdot Z = (1/2)[\hat{Z},\hat{Y}] \cdot \hat{w} = \Omega(\hat{Y},\hat{Z}) \cdot \hat{w}; \)
(d) \( (D_v \hat{Y}) \cdot Z = (D_X Y) \cdot Z = D_{\beta(0)} \hat{u}_0 \cdot Z. \)

The four equations of A.2.5.5(a)-(d) combined with the fact that the norm of the curvature tensor for \( \hat{V} \) is bounded by \( A'0 \) yield the necessary estimates for \( |D_v \hat{Y}| \) and \( |D_v \hat{Y}| \) completing the proof for Theorem A.2.5 once we have verified the following lemma. (See again [15, Vol. 1, p. 133] for the standard relation between \( \Omega( , ) \) and the curvature tensor on \( \hat{V} \). It yields the bound for the norm of \( \Omega( , ) \) needed for using A.2.5.5(b) and (c).)

Lemma A.2.5.6. \( |D_{\beta(0)} \hat{u}_0 \cdot Z| \leq n, \) where \( n = \dim \hat{V}. \)

Proof of Lemma A.2.5.6. Recall that \( \gamma(0) \) is an orthonormal framing for the tangent space of \( \hat{V} \) at \( \beta(0) \); i.e., \( \gamma(0) = (u_1, u_2, \ldots, u_n) \) where each \( u_i \in T_{\beta(0)}\hat{V} \) and
\[
\begin{aligned}
  u_i \cdot u_j &= \begin{cases} 
    0 & \text{if } i \neq j \\
    1 & \text{if } i = j.
  \end{cases}
\end{aligned}
\]

And notice that \( \gamma(t) \) is the frame field \( (\hat{u}_1(t), \hat{u}_2(t), \ldots, \hat{u}_n(t)) \) along the curve \( \beta(t) \) in \( \hat{V} \); consequently,

\[
|\gamma(0)|^2 = |\beta(0)|^2 + |D_{\beta(0)} \hat{u}_1|^2 + \cdots + |D_{\beta(0)} \hat{u}_n|^2.
\]

We conclude from (1) and \( |\gamma(0)| = 1 \) that
\[
|D_{\beta(0)} \hat{u}_i| \leq 1 \text{ for } i = 1, 2, \ldots, n.
\]

Express \( u_0 \in T_{\beta(0)}\hat{V} \) in terms of the orthonormal frame \( (u_1, u_2, \ldots, u_n) \); i.e., let \( u_0 = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n \). Then for each \( i \) we have that
\[
|a_i| \leq 1
\]
since $|w_0| = 1$. Note also that $\hat{u}_0 = a_1\hat{u}_1 + a_2\hat{u}_2 + \ldots + a_n\hat{u}_n$. It follows therefore from inequalities (2) and (3) that

$$(4) \quad |D_{\beta(0)} \hat{u}_0| \leq \sum_{i=1}^{n} |a_i||D_{\beta(0)} \hat{u}_i| \leq n.$$  

The inequality posited in the lemma now follows immediately from (4) and the fact that $|Z| \leq 1$.

This completes the proof of Theorem A.2.5. □

**Proof of Theorem A.2.3.** We may identify $\hat{V}$ with $T_R(\hat{y})$: i.e., with the tubular neighborhood of radius $R > \delta$ where $y \in \hat{V}$ is a point such that the isotropy subgroup $C$ of the action of $H$ at $y$ is a maximal compact subgroup of the virtually nilpotent Lie group $H$; cf. [4, §8].

(We must distinguish between the orbit of $y$ under the action of $H$ denoted by $Hy$ and the isotropy subgroup of $H$ fixing $y$ denoted by $H_y$.)

**Remark 1.** Let $\bar{y} \in F\hat{V} = E$ lie over $y$; i.e., $\rho(\bar{y}) = y$ where $\rho : E \to \hat{V}$ is the canonical projection. Then the equation

$$c\bar{y} = \bar{y}c,$$

identifies $C$ with a subgroup $\bar{C}$ of $O(n)$, $n = \dim \hat{V}$. (Throughout the proof, $n$ denotes $\dim \hat{V} = \dim M$.) In fact, $\bar{C}$ is the isotropy group of the (right) action of $O(n)$ on $H\setminus E$ at $H\bar{y}$. Consequently, $\bar{C}$ is a maximal isotropy group for the action of $O(n)$ on $H\setminus E$.

**Fact 1.** There exists a positive integer $k = k(n, \epsilon, A_0)$ and closed subgroups $C_1, C_2, \ldots, C_k$ of $O(n)$, depending only on $n$, $\epsilon$ and $A_0$, such that $\bar{C}$ is orthogonally conjugate to some group $C_i$ in this list.

Fact 1 is verified at the end of this appendix after using it to prove Theorem A.2.3. There are the following ten steps in this Proof of Theorem A.2.3:

**Step 1.** Let $\alpha : [0,1] \to E = F\hat{V}$ be a piecewise smooth path connecting $x$ to $gx$ such that $\text{length}(\alpha) < \delta$. Let $\psi : \hat{V} = T_R(Hy) \to Hy$ be the orthogonal projection. It clearly suffices to consider the case where $\psi(\rho(x)) = y$. So we make this assumption. Let $\beta : [0,1] \to Hy$ be the composite of $\alpha$ and $\psi \circ \rho$; i.e.,

$$\beta = \psi \circ \rho \circ \alpha.$$

**Claim.** There exists a constant $C_1 = C_1(n, \epsilon, A_0) > 0$ such that $\text{length}(\beta) < C_1 \delta$ where $Hy$ has the Riemannian metric determined by being a submanifold of $\hat{V}$. 
Step 2. There exists a piecewise smooth path \( \beta : [0, 1] \to H \) and a constant \( C_2 = C_2(n, \epsilon, A_0) \) satisfying the following properties where \( \hat{\beta} : [0, 1] \to E \) is defined by \( \hat{\beta}(t) = \beta(t)\tilde{y} \) for all \( t \in [0, 1] \):
1. \( \beta(0) = e \). (Here \( e \) is the identity element of \( H \).)
2. \( \rho \circ \hat{\beta} = \beta \).
3. \( \text{length}(\hat{\beta}) < C_2 \delta \).

Step 3. Define a piecewise smooth path \( \gamma : [0, 1] \to E \) by \( \gamma(t) = \beta(t)x \) for all \( t \in [0, 1] \). It clearly satisfies the following properties:
1. \( \gamma(0) = x \).
2. \( \psi \circ \rho \circ \gamma = \beta \).

Claim. There exists a constant \( C_3 = C_3(n, \epsilon, A_0) > 0 \) such that \( \text{length}(\gamma) < C_3 \delta \).

Step 4. Let \( h = \beta(1) \in H \). Because of Step 3 (Property 2), we have \( hy = gy \) and consequently \( y = (h^{-1}g)y \). Thus setting \( c = h^{-1}g \), we have that \( c \in C \).

Also
\[
d_E(x, cx) \leq (C_3 + 1)\delta
\]
because of the following reasoning:
\[
d_E(x, cx) = d_E(x, h^{-1}gx) = d_E(hx, gx) \leq d_E(hx, x) + d_E(x, gx) \leq C_3\delta + \delta.
\]

Claim. There exists a constant \( C_4 = C_4(n, \epsilon, A_0) > 0 \) such that \( d_E(\bar{y}, c\bar{y}) < C_4 \delta \).

Remark. Let \( \Psi : E \to \rho^{-1}(HY) = HYO(n) \) be the orthogonal projection. Then \( \Psi \) is \( H, O(n) \) bi-equivariant and covers \( \psi : V \to HY \). Also fix \( \bar{y} = \Psi(x) \); i.e., this is a specific framing of \( y \). Then the above claim will result by establishing the inequality
\[
d_E(\Psi(x), \Psi(cx)) \leq C_4d_E(x, cx).
\]

Step 5. We make the following claim in this step:

Claim. There exists a constant \( C_5 = C_5(n, \epsilon, A_0) > 0 \) such that \( d_{\rho^{-1}(y)}(\bar{y}, c\bar{y}) < C_5 \delta \).

Remark. This is proved using the facts that:
1. \( \rho^{-1}(y) \) is a totally geodesic submanifold of \( E = FV \);
2. \( \rho : E \to V \) is a Riemannian submersion;
3. \( V \) has injectivity radius \( > \delta \) at \( y \).

Step 6. In this step we make the following claim:
Claim. Provided $\delta$ is sufficiently small, then $d_{C\bar{y}}(\bar{y}, c\bar{y}) < 2C_5\delta$.

Remark. Recall that $c\bar{y} = \bar{y}c$ where $c \in C \subset O(n)$. (See Remark 1 at the beginning of this proof.) Note also, for each $z \in E$, that the orbit map $g \rightarrow zg$ is an isometry between $O(n)$ with its standard metric and $zO(n)$ with its Riemannian metric induced by the inclusion $zO(n) \subset E$. Then the above claim is a consequence of Fact 1 (at the beginning of this proof) and the Claim made in Step 5.

Step 7. The following inequality is an immediate consequence of the Claim made in Step 6 and the inclusion $C\bar{y} \subset H\bar{y}$.

\[d_{H\bar{y}}(\bar{y}, c\bar{y}) < 2C_5\delta.\]

Step 8. We have the following claim:

Claim. $d_{Hx}(x, cx) < C_3(2C_5)\delta$.

This claim is a consequence of the inequality posited in Step 7 together with the argument which yields the Claim made in Step 3.

Step 9. The triangle inequality says that

\[d_{Hx}(x, gx) \leq d_{Hx}(x, hx) + d_{Hx}(hx, gx) = d_{Hx}(x, hx) + d_{Hx}(x, cx).\]

Combining this inequality with the claims made in Steps 3 and 8 yields that

\[d_{Hx}(x, gx) < C_3(1 + 2C_5)\delta.\]

Step 10. Put on $H$ the left invariant Riemannian metric induced from the monic immersion $H \rightarrow E$ defined by $k \rightarrow kx$ for all $k \in H$. Then the sectional curvatures of $H$ are bounded above by a constant which depends only on $n, \epsilon$ and $A_0$. This fact is a consequence of Theorem A.2.5. Hence, provided $\delta$ is sufficiently small (how small again depends only on $n, \epsilon$ and $A_0$) we can use Claim 2.8.5 (from §2) in conjunction with the inequality demonstrated in Step 9 to produce a 1-parameter subgroup $\phi : \mathbb{R} \rightarrow G$ satisfying Properties (a) and (b) posited in Theorem A.2.3 with $B = 4^nC_3(1 + 2C_5)$.

This completes the outline of the main steps in the Proof of Theorem A.2.3. Now in the following remarks we shall indicate some more details in the verifications of the Claims made in the preceding steps.

Remarks on Step 10.

To show that the hypotheses of Claim 2.8.5 are satisfied use the following Lemma which we proceed now to formulate and prove:

Let $G$ be a simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$ equipped with an inner product. Put on $G$ the corresponding left invariant
Riemannian metric. Let $a_1$ denote the norm of the Lie bracket on $\mathfrak{g}$; i.e.,

$$a_1 = \max \{|[u,v]| : u, v \in \mathfrak{g} \text{ and } |u| = |v| = 1\}.$$ 

Let $a_2$ denote the square root of the maximum of the absolute values of the sectional curvatures of $G$. And let $a_3$ denote the norm of $D\hat{v}\hat{u}$; i.e.,

$$a_3 = \max \{|D\hat{v}\hat{u}| : u, v \in \mathfrak{g} \text{ and } |u| = |v| = 1\},$$

where $\hat{u}, \hat{v}$ are the left invariant vector fields on $G$ with $\hat{u}(e)$ and $\hat{v}(e) = v$. (Here $e$ is the identity element of $G$.)

**Lemma.** There is a positive real number $B$ which depends only on $\dim G$, such that $a_i \leq B a_j$ for all $i, j \in \{1, 2, 3\}$.

**Proof of Lemma.** From [2, Proposition 7.7.1] one deduces the following two inequalities:

(a) $a_2 \leq \sqrt{6} a_1$;
(b) $a_3 \leq (3/2) a_1$.

Using [16, Theorem 47, p. 213], one can construct a positive real number $\overline{B}$, which depends only on $\dim G$, such that:

(c) $a_1 \leq \overline{B} a_2$.

Also Claims 2.8.1 and 2.8.2 (proven in §2 above) yield the following inequality:

(d) $a_1 \leq 2(\dim G)^2 a_3$.

Concatenating inequalities (a)-(d) proves the Lemma. \qed

**Remarks on Steps 1 and 4.**

Inequality (8.7) of [4] states that the norm of the second fundamental form of $H y$ in $\tilde{V}$ is bounded above by a positive constant depending only on $n, \epsilon$ and $A_0$. (This can also be independently verified using that the angle between $H \tilde{y}$ and $\tilde{y}O(n)$ is bounded below. We will show this in “Remarks on Step 2” given below.) Then Lemma A.2.5.4 can be used to verify the same fact for the second fundamental form of $H \tilde{y}O(n)$ in $E$. Consequently, both $\psi$ and $\Psi$ are $\Theta$-almost Riemannian submersions (cf. 2.5.1 of [4] for the definition) where $\Theta$ depends only on $n, \epsilon$ and $A_0$.

**Remarks on Steps 3 and 8.**

Note that $\Psi \circ \gamma = \beta$ and recall from “Remarks on Steps 1 and 4” that $\Psi$ is a $\Theta$-almost Riemannian submersion where $\Theta$ depends only on $n, \epsilon$ and $A_0$. Also note that $\gamma$ is tangent to the foliation of $E$ by the orbits of $H$. So it suffices to show that the angle between $H x$ and $\Psi^{-1}(y)$ at $x$ is larger than $C_6 = C_6(n, \epsilon, A_0) > 0$. But because of Theorem A.2.5, we can apply Lemma 2.6 to conclude this. Step 8 follows by the same argument.
Remarks on Step 2.

Let $\rho : H\bar{y} \to H y$ be the restriction of $\rho : F\hat{V} \to \hat{V}$. (Recall $\rho(\bar{y}) = y$ and $\rho$ is $H$-equivariant.) It suffices to show that $\rho$ is a $\Theta$-almost Riemannian submersion where $\Theta = \Theta(n, \epsilon, A_0) > 0$. To do this, it is enough to show that the angle between the tangent spaces to $H\bar{y}$ and $y O(n)$ at $y$ is bounded below by a number $\Theta = \Theta(n, \epsilon, A_0)$.

This is because $\rho : \rho^{-1}(H y) = H\bar{y}O(n) \to H y$ is a Riemannian submersion.

Let $E$ be the Riemannian manifold used in \cite[§8]{4} to construct $\hat{V}$. There is a virtually nilpotent Lie group $\hat{H}$ such that $\hat{H}$ and $O(n)$ both act freely (and properly) by isometries on $E$, these actions commute, and $\hat{V} = E/O(n)$.

($\hat{H}$ acts on the left and $O(n)$ on the right of $E$.)
Furthermore, there is a Lie group epimorphism $\phi : \hat{H} \to H$ with finite kernel such that $\phi|_{\hat{H}^0}$ is an isomorphism of $(\hat{H})^0$ to $H^0$ and such that the principal $O(n)$-bundle projection $\hat{\rho} : E \to \hat{V}$ is $\phi$-equivariant; \cite[pp. 365, 369]{4}.

Let $\hat{y} \in E$ be a point such that $\hat{\rho}(\hat{y}) = y$. To show the bound $\Theta$ exists, we need the following result:

**Theorem 1.** The angle between the tangent spaces to $H^0\hat{y}$ and $\hat{y}O(n)$ at $\hat{y}$ is bounded below by a number $\hat{\Theta} = \hat{\Theta}(n, \epsilon, A_0) > 0$.

To prove this result we need three lemmas.

**Lemma 1.** Given $\sigma > 0$, there exists $\mu = \mu(\sigma, n) > 1$ such that for any unit speed 1-parameter subgroup $\alpha : R \to O(n)$ there is a number $t \in [1, \mu]$ with $d(\alpha(t), \alpha(0)) < \sigma$.

**Proof.** There is a maximal torus $T$ in $O(n)$ containing the image of $\alpha$. The volume of $T$ depends only on $n$ since all maximal tori in $O(n)$ are conjugate. Consider the set of open balls in $T : B_{\sigma/2}(\alpha(1)), B_{\sigma/2}(\alpha(2)), B_{\sigma/2}(\alpha(3)), \ldots$. They all have the same volume $V = V(n, \sigma) > 0$. Let $\mu = 1 + [\text{Vol}(T)/V]$. Then clearly there are two integers $i$ and $j$ with $1 \leq i < j \leq \mu$ such that

$$B_{\sigma/2}(\alpha(i)) \cap B_{\sigma/2}(\alpha(j)) \neq \emptyset.$$

Let $t = j - i$, then we are done with the proof of Lemma 1.

Consider the orbit action of $\hat{H} \times O(n)$ on $E$ at $\hat{y}$ defined by

$$(a, b) \to a \hat{y} b^{-1}, \text{ where } (a, b) \in \hat{H} \times O(n).$$
and denote this action by \((a, g) \cdot \hat{y}\). Since the induced actions of \(\hat{H}\) and \(O(n)\), respectively, are faithful, they determine left invariant Riemannian metrics on \(\hat{H}\) and \(O(n)\). The metric determined thusly on \(O(n)\) is the same as that induced by the negative of the Killing form on \(\mathfrak{g}(O(n))\); cf. Remark A.2.5.2.1. For any \(v \in \mathfrak{g}(H^0 \times O(n))\), let \(\alpha_v : \mathbb{R} \to H^0 \times O(n)\) denote the corresponding 1-parameter subgroup. Let \(u \in \mathfrak{g}(O(n))\) and \(v \in \mathfrak{g}(H^0)\) be such that the two curves in \(E, \beta_u\) and \(\beta_v\), defined by
\[
t \to \hat{y} \alpha_u(t) \quad \text{and} \quad t \to \alpha_v(t) \hat{y}
\]
are both perpendicular to the orbit space \(C_0 \hat{y}\) at \(\hat{y}\). Let \(V \subset \mathfrak{g}(H^0 \times O(n))\) denote the subspace spanned by \(u\) and \(v\). Since \([u, v] = 0\), this subspace is an abelian Lie subalgebra. Furthermore, assume that
\[
|u| = |\beta_u(0)| = 1
\]
and that when \(\beta_v(0)\) is orthogonally projected onto \(T_{\hat{y}}(\hat{y}O(n))\) it hits the vector \(\beta_u(0)\). Denote the angle between \(\beta_u(0)\) and \(\beta_v(0)\) by \(\Theta\).

**Lemma 2.** Given \(\sigma > 0\) there exists \(t \geq 1\) such that
\[
d_E(\alpha_v(t) \hat{y}, \hat{y}) \leq \sigma/2 + \mu(\sigma/2, n) \tan(\Theta)
\]
and
\[
d_{H^0}(\alpha_v(t) \hat{y}, \hat{y}) \geq d = d(n, \epsilon, A_0) > 0.
\]

**Proof.** Let \(t \in [1, \mu(\sigma/2, n)]\) be the number posited in Lemma 1 when \(\alpha = \alpha_{-u}\). And let \(w = v - u\), then \(w \in V\). Hence we have that
\[
\alpha_v(t) = \alpha_u(t) \alpha_w(t)
\]
since \(V\) is abelian; consequently
\[
(1) \quad d_E(\alpha_v(t) \cdot \hat{y}, \hat{y}) \leq d_E((\alpha_u(t) \alpha_w(t)) \cdot \hat{y}, \alpha_u(t) \cdot \hat{y}) + d_E(\alpha_u(t) \cdot \hat{y}, \hat{y})
\]
by the triangle inequality. Substituting the equations
\[
d_E((\alpha_u(t) \alpha_w(t)) \cdot \hat{y}, \alpha_u(t) \cdot \hat{y}) = d_E(\alpha_w(t) \cdot \hat{y}, \hat{y}) \quad \text{and} \quad d_E(\alpha_u(t) \cdot \hat{y}, \hat{y}) = d_{O(n)}(\alpha_{-u}(t), \alpha_{-u}(0))
\]
into (1) and applying Lemma 1 to \(\alpha = \alpha_{-u}\), we obtain that
\[
(2) \quad d_E(\alpha_v(t) \cdot \hat{y}, \hat{y}) \leq \sigma/2 + d_E(\alpha_w(t) \cdot \hat{y}, \hat{y}).
\]
Let \(\beta_w\) denote the curve in \(E\) defined by
\[
s \to \alpha_w(s) \cdot \hat{y}
\]
and observe that
\[ |\dot{\beta}_w(0)| = \tan \Theta. \]

Note also that \( \beta_w \) is a constant speed curve; consequently,
\[ (3) \quad d_E(\alpha_w(t) \cdot \hat{y}, \hat{y}) < t(tan \Theta). \]

Substituting (3) into (2) and recalling that \( t \leq \mu(\sigma/2,n) \) yields the first inequality of Lemma 2.

To prove the second inequality, recall that the orbit action is an isometry between \( H^0 \) and \( H^0 \hat{y} \). Hence \( |v| = |\dot{\alpha}_v(0)| \geq 1 \) since
\[ (4) \quad |\dot{\alpha}_v(0)| = |\dot{\beta}_v(0)| \geq |\dot{\beta}_u(0)| = 1. \]

Next consider the simply connected nilpotent Lie group \( H^0/C^0 \) and its one-parameter subgroup \( \gamma \) defined by
\[ \gamma(s) = \alpha_v(s)C^0, \quad \text{for all } s \in \mathbb{R}. \]

Put the left invariant Riemannian metric on \( H^0/C^0 \) such that the quotient map \( H^0 \to H^0/C^0 \) is a Riemannian submersion; hence,
\[ (5) \quad |\dot{\gamma}(0)| \geq 1 \]
because of (4) and the fact that \( \dot{\alpha}_v(0) \) is perpendicular to \( T_eC^0 \). (We use \( e \) to denote the identity element of a group.) Also note that
\[ (6) \quad d_{H^0/C^0}(\alpha_v(t)\hat{y}, \hat{y}) \geq d_{H^0/C^0}(\gamma(t), C^0) \]
since \( d_{H^0}(\alpha_v(t)\hat{y}, \hat{y}) = d_{H}(\alpha_v(t), e) \) and \( H^0 \to H^0/C^0 \) is a Riemannian submersion. Next observe that the sectional curvatures of \( H^0/C^0 \) are bounded above by a positive real number \( k(n, \epsilon, A_0) \). This observation follows from Lemma in the above Remarks on Step 10 and from [4, 6.1.8, 7.21]. Therefore Claim 2.8.5, together with Lemma in Remarks on Step 10, and this bound on the sectional curvatures of \( H^0/C^0 \) yield that there exists a positive real number \( c(n, \epsilon, A_0) \) such that either
\[ (7) \quad d_{H^0/C^0}(\gamma(t), C^0) \geq c(n, \epsilon, A_0) \]
or
\[ d_{H^0/C^0}(\gamma(t), C^0) \geq 4^{-n}|\dot{\gamma}(0)|t \geq 4^{-n}. \]

Now define the number \( d(n, \epsilon, A_0) posited in Lemma 2 by
\[ d(n, \epsilon, A_0) = \min\{c(n, \epsilon, A_0), 4^{-n}\}. \]

Then combining the inequalities occurring in (6) and (7) proves the second assertion of Lemma 2.

This completes the proof of Lemma 2. \( \square \)
Lemma 3. There exist positive real numbers $\nu = \nu(n, \epsilon, A_0)$ and $\tau = \tau(n, \epsilon, A_0)$ such that, for any $z \in \hat{H}\hat{y}$,
\[
d_{H^0\hat{y}}(z, \hat{y}) \leq \tau d_E(z, \hat{y})
\]
whenever $d_E(z, \hat{y}) \leq \nu$.

Proof. This is a routine consequence of the following two facts:
1) The normal injectivity radius of $\hat{H}\hat{y}$ in $E$ is bounded below by a positive number which depends only on $n, \epsilon, A_0$.
2) The norm of the second fundamental form for $\hat{H}\hat{y}$ in $E$ is bounded above by a positive number which depends only on $n, \epsilon, A_0$.
These two facts are proven in [4] (cf. 7.21, 6.1.2, 6.1.4 and 6.1.8; also the proof of Proposition A.2.2). This completes the proof for Lemma 3. $\square$

Proof of Theorem 1. Set $\sigma = \min\{\nu, d/\tau\}$ and select $\hat{\Theta}$ to be any positive number such that
\[
\mu(\sigma/2, n) \tan \hat{\Theta} < \sigma/2.
\]
(Here $\nu, d, \tau$ and $\mu$ come from Lemmas 1, 2 and 3.) Now suppose the assertion in Theorem 1 is false. Then there exists vectors $u$ and $v$ as in the setup to Lemma 2 such that
\[
\tan \Theta < \tan \hat{\Theta}.
\]
Set $z = \alpha_v(t)\hat{y}$ in Lemma 3, where $t$ comes from Lemma 2. Then the first inequality in Lemma 2 yields that
\[
(1) \quad d_E(z, \hat{y}) < \sigma < \nu.
\]
Hence Lemma 3 applies to show that
\[
(2) \quad d_{H^0\hat{y}}(z, \hat{y}) \leq \tau d_E(z, \hat{y}).
\]
Combining inequalities (1) and (2) with the second inequality in Lemma 2 yields that
\[
(3) \quad d < \tau \sigma
\]
contradicting the fact that $\sigma \leq d/\tau$.
This completes the Proof of Theorem 1. $\square$

Continuing now with the verification of Step 2 we will need the following notation:

Notation. Let $G \times X \to X$ be a smooth action of a Lie group $G$ on a smooth manifold $X$ and $v \in g(G)$ where $g(G)$ denotes the Lie algebra of $G$. Then $\tilde{v}$ denotes the vector field on $X$ whose value at $x \in X$ is the tangent vector to the curve
\[
t \to \alpha_v(t)x
\]
at \( t = 0 \). (As usual \( \alpha_v \) denotes the one-parameter subgroup of \( G \) corresponding to \( v \).)

**Lemma 4.** There exists a positive number \( K_1 = K_1(n, \epsilon, A_0) \) such that

\[
|\hat{D}_X \hat{v}| \leq K_1 |X| \|\hat{v}| |
\]

for every pair of vectors \( \hat{v} \in \mathfrak{g}(H^0) \) and \( X \) tangent to \( E \). (Here \( \hat{D} \) denotes the Levi-Civita connection on \( E \).)

**Proof.** This follows from [4, 4.7 and 4.9].

Recall \( \hat{\rho} : E \rightarrow \hat{V} \) denotes the orbit map. (Also recall \( \hat{V} = E/O(n) \).) And let \( D \) denote the Levi-Civita connection on \( \hat{V} \). We say that \( \hat{v} \in \mathfrak{g}(H^0) \) is perpendicular to \( \mathfrak{g}(C^0) \) at \( \hat{y} \in E \) provided \( \hat{v} \) is perpendicular to \( \hat{u} \) at \( \hat{y} \) for every vector \( u \in \mathfrak{g}(C^0) \).

**Theorem 2.** There exists a positive number \( K_2 = K_2(n, \epsilon, A_0) \) such that

\[
|D_X \hat{v}| \leq K_2 |X| \|\hat{v}| |
\]

for every vector \( X \in T_y(\hat{V}) \) and every \( \hat{v} \in \mathfrak{g}(H^0) \) which is perpendicular to \( \mathfrak{g}(C^0) \) at \( \hat{y} \). (Recall \( y = \hat{\rho}(\hat{y}) \).)

**Proof.** Since \( \hat{\rho} : E \rightarrow \hat{V} \) is \( H^0 \)-equivariant, the vector field \( \hat{v} \) on \( E \) maps to the vector field \( \hat{v} \) on \( \hat{V} \) via \( d\hat{\rho} \); i.e., they are \( \hat{\rho} \)-related vector fields. Let \( \tilde{v} \) denote the horizontal lift of \( \hat{v} \) in \( \hat{V} \) to \( E \). Since \( \hat{v} \in E \) is \( \hat{\rho} \)-related to \( \hat{v} \) in \( \hat{V} \), \( \tilde{v} \) is the horizontal component of \( \hat{v} \) in \( E \); i.e.,

\[
\hat{v} = \tilde{v} + w
\]

where \( w \) is a vertical vector field in \( E \) relative to the Riemannian submersion \( \hat{\rho} : E \rightarrow \hat{V} \). Let \( \hat{X} \) denote the horizontal lift of \( X \) to \( E \), then

\[
\hat{D}_X \hat{v} = \hat{D}_X \tilde{v} - \hat{D}_X w
\]

because of (1). Consequently

\[
\mathcal{H}(\hat{D}_X \tilde{v}) = \mathcal{H}(\hat{D}_X \hat{v}) - \mathcal{H}(\hat{D}_X w)
\]

where \( \mathcal{H}(u) \) denotes the horizontal component of a vector \( u \) tangent to \( E \).

By a standard result (cf. [16, p. 212, Lemma 45])

\[
\mathcal{H}(\hat{D}_X \tilde{v}) = Y \quad \text{where} \quad Y = (D_X \hat{v}).
\]

Applying this fact and Lemma 4 to (3) yields

\[
|D_X \hat{v}| \leq K_1 |X| \|\hat{v}|| + |\mathcal{H}(\hat{D}_X w)|
\]

where \( \|\hat{v}| | \) denotes the length of \( \hat{v} \) in \( E \). Let \( \omega \) denote the angle between \( \hat{v} \) and \( \tilde{v} \) at \( \hat{y} \), then

\[
(\pi/2 - \omega) > \hat{\Theta}
\]
because of Theorem 1 and the fact that $v$ is perpendicular to $g(C^0)$ at $\hat{y}$. (Recall that $\hat{\Theta} = \hat{\Theta}(n, \epsilon, A_0) > 0$.) Consequently,

$$\|\hat{v}\| \leq (\csc \hat{\Theta})|\hat{v}|. \quad (6)$$

Combining inequalities (4) and (6) yields

$$|D_X\hat{v}| \leq (\csc \hat{\Theta})K_1|X||\hat{v}| + |\mathcal{H}(\hat{D}_Xw)|. \quad (7)$$

It remains to estimate $|\mathcal{H}(\hat{D}_Xw)|$. For this purpose, a routine application of the Koszul formula (as in the proof of Lemma A.2.5.4(c) above) yields that

$$\hat{D}_Xw \cdot \hat{Y} = \Omega(X, \hat{Y}) \cdot \hat{u} \quad (8)$$

where $Y$ is any vector tangent to $\hat{V}$ at $y$; $u \in g(O(n))$ is the unique vector such that $\hat{u} = w$ at $\hat{y}$, and $\Omega(\ ,\ )$ denotes the curvature form for the principal $O(n)$-bundle $\hat{\rho}: \mathcal{E} \rightarrow \hat{V}$ relative to its horizontal distribution.

**Claim 1.** The norm of $\Omega(\ ,\ )$ is bounded above by a positive number $K_0 = K_0(n, \epsilon, A_0)$.

Before verifying this claim, we use it to complete the proof of Theorem 2. (Its verification will come at the end of Remarks on Step 2.) Claim 1 applied to Equation (8) yields

$$|\mathcal{H}(\hat{D}_Xw)| \leq K_0|X||w|. \quad (9)$$

Now note that inequality (5) yields

$$|w| < (\cot \hat{\Theta})|\hat{v}|. \quad (10)$$

Combining inequalities (7), (9) and (10) yields that

$$|D_X\hat{v}| \leq (K_1 \csc \hat{\Theta} + K_0 \cot \hat{\Theta})|X||\hat{v}| \quad (11)$$

proving Theorem 2.

$\Box$

**Lemma 5.** Let $v \in g(H^0)$ and $u \in T_y(\hat{V})$, then

$$D_v\hat{u} = D_u\hat{v}. \quad \square$$

**Remark.** Recall $\hat{u}$ denotes the $H$-invariant vector field along $Hy$ whose value at $y$ is $u$.

**Proof.** This is because $[\hat{u}, \hat{v}] = 0$, which is seen by considering the parametrized surface $f: \mathbb{R}^2 \rightarrow \hat{V}$ defined by

$$f(s, t) = \alpha_v(t)\beta(s)$$
where $\beta : \mathbb{R} \to \hat{V}$ is any smooth curve such that $\beta(0) = u$. Then note that
\[
\frac{\partial f}{\partial s} = \hat{u}, \text{ while } \frac{\partial f}{\partial t} = \hat{v}.
\]
This completes the proof of Lemma 5. \qed

We are now ready to show that the angle between $Hy$ and $yO(n)$ at $y \in E$ is bounded by a positive number $\Theta = \Theta(n, \epsilon, A_0)$. To do this, let $v \in \mathfrak{g}(H^0)$ and define smooth curves $\beta : \mathbb{R} \to E$ and $\gamma : \mathbb{R} \to \hat{V}$ by the equations
\[
\beta(t) = \alpha_v(t)\bar{y} \text{ and } \gamma(t) = \alpha_v(t)y
\]
for all $t \in \mathbb{R}$. Note that $\gamma = \rho \circ \beta$ and that $\gamma(0) = \hat{v}$ in the notation fixed in Lemma 4. Assume also that $v$ is perpendicular to $\mathfrak{g}(O^0)$ at $\hat{y} \in \mathcal{E}$. Recall that $\bar{y}$ is an orthonormal framing of $T_y(\hat{V})$; i.e.,
\[
\bar{y} = (u_1, u_2, \ldots, u_n)
\]
where each $u_i \in T_y(\hat{V})$ and
\[
u_i \cdot u_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}
\]
Hence $\beta(0) \in T_{\bar{y}}(E)$ is represented by the following $(n+1)$-tuple of vectors in $T_{\bar{y}}(\hat{V})$:
\[
(\hat{v}, D_{\hat{v}}\hat{u}_1, D_{\hat{v}}\hat{u}_2, \ldots, D_{\hat{v}}\hat{u}_n)
\]
which equals
\[
(\hat{v}, D_{u_1}\hat{v}, D_{u_2}\hat{v}, \ldots, D_{u_n}\hat{v})
\]
because of Lemma 5. Applying Theorem 2, we see that
\[
|D_{u_i}\hat{v}| \leq K_2|\hat{v}| \text{ for } i = 1, 2, \ldots, n.
\]
Let $\omega$ denote the angle between $\beta(0)$ and the $O(n)$-orbit $\bar{y}O(n)$ at $\bar{y}$, then
\[
\tan \omega \geq 1/(nK_2)
\]
because of $(*)$ and $(**)$. It is now a routine exercise, using $(***)$, to show that the angle between the tangent spaces to $H\bar{y}$ and $\bar{y}O(n)$ at $\bar{y}$ is bounded from below by a number $\Theta = \Theta(n, \epsilon, A_0) > 0$. This completes the demonstration of Step 2 modulo the verification of Claim 1.

**Verification of Claim 1.**

We start by stating the following extension of Lemma A.2.5.4 which follows again from a routine application of the Koszul formula:

**Observation.** Let $p : E \to N$ be any principal $O(n)$-bundle where $O(n)$ acts by isometries on the Riemannian manifold $E$ and $p$ is a Riemannian submersion. Then formulas (b) and (c) of Lemma A.2.5.4 remain true in
this more general setting where $X, Y, Z$ are now vector fields on $N$ and special vector fields $\tilde{u}, \tilde{v}, \tilde{w}, \tilde{X}, \tilde{Y}, \tilde{Z}$ are defined as in the preamble to A.2.5.4. Furthermore, formulas (a) and (d) also hold true if the right invariant Riemannian metric induced on $O(n)$ by the immersion $g \rightarrow zg, g \in O(n)$, and $z$ a fixed point in $E$, is independent of $z$ and is identical to that determined by the negative of the Killing form for the Lie algebra $\mathfrak{g}(O(n))$; cf. Remark A.2.5.2.1.

Let $\Omega$ and $\hat{\Omega}$ denote the curvature forms for the principal $O(n)$-bundle $\hat{\rho} : E \rightarrow \hat{V}$ relative to the horizontal distributions determined by the $O(n)$-invariant Riemannian metrics $\tilde{h}_\epsilon$ and $\tilde{h}$ on $E$, respectively. Here $\tilde{h}_\epsilon$ and $\tilde{h}$ are as in Proposition 7.2.1 of [4] and $h$ is the Riemannian metric $S_{\epsilon/2}(g)$ of [4, p. 362, line-5]. Also $g_\epsilon$ denotes the Riemannian metric on $\hat{V}$ such that $\hat{\rho} : (E, \tilde{h}_\epsilon) \rightarrow (\hat{V}, g_\epsilon)$ is a Riemannian submersion.

Because the standard relation between $\hat{\Omega}$ and the curvature tensor on $\hat{V}$ for the Riemannian metric $h$, given in [15, Vol.1, p. 133], we see that the norm of $\hat{\Omega}$ is bounded above by a positive number $\hat{K}_0 = \hat{K}_0(n, \epsilon, A_0)$. So it suffices to show that the norms of $\Omega$ and $\hat{\Omega}$ are nicely related.

To do this, we use the notation $\tilde{X}$ and $\tilde{X}$ for the horizontal lifts of a vector (or vector field) $X$ in $\hat{V}$ to $E$ with reference to the horizontal distributions determined by $\tilde{h}_\epsilon$ and $\tilde{h}$, respectively. Let $X$ and $Y$ be vectors tangent to $\hat{V}$ based at a common point and such that $g_\epsilon(X, X) = g_\epsilon(Y, Y) = 1$ and let $u \in \mathfrak{g}(O(n))$ denote $\Omega(\tilde{X}, \tilde{Y})$ where the horizontal lifts $\tilde{X}, \tilde{Y}$ are also based at a common point in $E$. Recall that $u$ is the unique vector satisfying

$$\tilde{u} = 2^{-1} \tilde{[Y, X]}.$$  

We proceed to estimate $|u|$ where $||$ denotes the norm on $\mathfrak{g}(O(n))$ determined by the negative of its Killing form; cf. Remark A.2.5.2.1. First note that

$$|u|^2 \leq 2\tilde{h}_\epsilon(\tilde{u}, \tilde{u})$$  

since $|u|^2 = \tilde{h}(\tilde{u}, \tilde{u})$ and because of the inequality (7.21.1) of [4]. (Here and throughout the proof of Claim 1 we assume that $\epsilon$ is sufficiently small relative to the number 2.)

Let $\tilde{D}$ and $\bar{D}$ denote the Levi-Civita connections on $E$ with respect to $\tilde{h}_\epsilon$ and $\tilde{h}$, then

$$\tilde{h}_\epsilon(\tilde{u}, \tilde{u}) = \tilde{h}_\epsilon(\tilde{D}_{\tilde{X}}\tilde{u}, \tilde{Y}).$$
because of formula (c) of A.2.5.4 and Equation (1); cf. Observation. But note that
\[
\tilde{h}_v(D_{\tilde{X}}\tilde{u}, \tilde{Y}) \leq 2|\tilde{h}(D_{\tilde{X}}\tilde{u}, \tilde{Y})| + 2|u|
\]
because of inequality (7.21.1) of [4]. Let \( v, w \in g(O(n)) \) denote the unique vectors such that
\[
\tilde{X} = \tilde{X} + \tilde{v}, \\
\tilde{Y} = \tilde{Y} + \tilde{w}.
\]
It follows that
\[
|v|, |w|, h(X, X), h(Y, Y) \leq 2
\]
because of Equations (0) and inequality (7.21.1) of [4]. Expanding \( \tilde{h}(D_{\tilde{X}}\tilde{u}, \tilde{Y}) \) using Equations (5) yields that
\[
\tilde{h}(D_{\tilde{X}}\tilde{u}, \tilde{Y}) = \tilde{h}(D_{\tilde{X}}\tilde{u}, \tilde{Y}) + \tilde{h}(D_{\tilde{v}}\tilde{u}, \tilde{w}) + \tilde{h}(D_{\tilde{v}}\tilde{u}, \tilde{Y}) + \tilde{h}(D_{\tilde{v}}\tilde{u}, \tilde{Y}).
\]
And applying formulas (a), (c) and (d) of Lemma A.2.5.4 to Equation (7) yields
\[
\tilde{h}(D_{\tilde{X}}\tilde{u}, \tilde{Y}) = \hat{\Omega}(\hat{X}, \hat{Y}) \cdot u - 2^{-1}[u, v] \cdot w.
\]
(Formulas (a) and (d) hold because \( \tilde{h} \) is the naturally induced Riemannian metric on the principle tangent bundle of the Riemannian manifold \((\hat{V}, g)\); cf. Observation.) Combining formulas (2), (3), (4) and (8) and using the fact that the inner product on \( g(O(n)) \) is bi-invariant yields that
\[
|u|^2 \leq 4|\tilde{h}(\hat{X}, \hat{Y})||u| + 2|[v, w]||u| + 4|u|.
\]
Canceling \(|u|\) from inequality (9) and using that the norm of \( \hat{\Omega} \) is bounded above by \( \hat{K}_0 \) together with the inequalities (6) yields that
\[
|u| \leq 8\hat{K}_0 + 8\eta + 4
\]
where \( \eta \) is the norm of the Lie bracket on \( g(O(n)) \). This inequality (10) shows that we can set \( K_0 = 8\hat{K}_0 + 8\eta + 4 \) completing the verification of Claim 1.

We end these Remarks on Step 2 by fulfilling the promise made in “Remarks on Steps 1 and 4” to give an alternate verification of inequality (8.7) in [4]. This inequality follows directly from the following three results proven above: Theorem 2, Lemma 5, and the fact that the angle between the tangent spaces to \( Hg \) and \( gO(n) \) at \( g \) is bounded below by the positive number \( \Theta(n, \epsilon, A_0) \).

**Remarks on Fact 1.** To complete the Proof of Theorem A.2.3 it remains to verify Fact 1. We start our verification of Fact 1 with the following result:
Lemma 1. Let $T$ be a maximal torus in $O(n)$. Given $v > 0$, there exists only a finite number of closed connected subgroups $S$ of $T$ such that $\text{Vol}(S) \leq v$.

Proof. Let $m = \dim T$ and $s = \dim S$, then $s \in \{0, 1, 2, \ldots, m\}$. Hence in proving that the number of such subgroups $S$ is finite, we may assume that $s$ is fixed and $s \neq 0$. Let $M^*(m,s;\mathbb{Z})$ denote the set of all $m \times s$ matrices $A$ with integral entries and rank $A = s$. Also let $T^s$ denote the Lie group $S^1 \times S^1 \times \ldots \times S^1$ (i-factors) and identify $T$ with $T^m$. Then each $A \in M^*(m,s;\mathbb{Z})$ determines a homomorphism $f_A : T^s \to T$. Note that if $B \in M^*(s,s;\mathbb{Z})$, then $\text{image}(f_{AB}) = \text{image}(f_A)$. Furthermore for each closed connected subgroup $S$ of $T$, there is a matrix $A \in M^*(m,s;\mathbb{Z})$ such that both $\text{image}(f_A) = S$ and $f_A$ is monic. An elementary argument yields the following estimate for $\text{Vol}(S)$. There is a positive constant $\sigma$ (independent of $S$) such that

\[
\text{Vol}(S) \geq \sigma |\det| / |\ker(f_A)|
\]

where “$\det$” denotes the determinant of any $s \times s$ submatrix of $A \in M^*(m,s;\mathbb{Z})$ and $\text{image}(f_A) = S$. Since rank $A = s$, there exists an $s \times s$ submatrix $B$ of $A$ such that $\det B \neq 0$; i.e., $B \in M^*(s,s;\mathbb{Z})$. Let $D = A \text{adj}(B)$ where $\text{adj}(B)$ denotes the classical adjoint of $B$. Since $\text{image}(f_D) = \text{image}(f_A)$, to prove Lemma 1, it suffices to show that the absolute values of the entries in $D$ are all bounded above by a positive number $K = K(v,m)$, where $v$ is any upper bound for $\text{Vol}(f_A)$, when $f_A$ is monic. But this follows from inequality $(\ast)$, in which $D$ replaces $A$, by considering the different $s \times s$ submatrices of $D$. In fact, we can take

\[
K(v,m) = v/\sigma.
\]

This completes the proof of Lemma 1. \hfill \Box

Lemma 2. Given $v > 0$, there are only finitely many conjugacy classes of closed subgroups $C$ of $O(n)$ such that both $C^0$ is abelian and $\text{Vol}(C) \leq v$.

Proof. Fix a maximal torus $T$ in $O(n)$. Then $C^0$ is conjugate to a subgroup of $T$. Hence Lemma 1 shows that there exists a positive integer $\tau = \tau(n,v)$ such that the finite subgroup $C/C^0$ has order $\leq \tau$. Another consequence of Lemma 1 is that it suffices, in proving Lemma 2, to demonstrate the following weaker statement:

(\ast) Given $v > 0$ and a closed connected abelian subgroup $S$ of $O(n)$, there are only finitely many conjugacy classes of closed subgroups $C$ of $O(n)$ such that both $C^0 = S$ and $\text{Vol}(C) \leq v$.

To verify statement (\ast), consider the normalizer $N(S)$ of $S$ in $O(n)$ and note that $N(S)$ is a compact Lie group and that $C$ is a subgroup of $N(S)$. Consider the factor Lie group $N(S)/S$ and its finite subgroup $C/S$ whose
order is less than $\tau$. Then (*) is an easy consequence of the following observation:

(**) There are only finitely many conjugacy classes of finite subgroups of $N(S)/S$ of order $\leq \tau$.

Observation (**) is a consequence of the fact that representations of finite groups into Lie groups are (locally) rigid, because of Weil’s rigidity theorem, together with the fact that $N(S)/S$ is compact.

This completes the proof of Lemma 2. □

**Lemma 3.** There is a function $f : \mathbb{R}^+ \times \mathbb{Z}^+ \to \mathbb{R}^+$ for which the following statement is true. Let $C$ be any closed subgroup of $O(n)$ such that both $C^0$ is abelian and

\[ \text{Vol} (C \backslash S^{n-1}) \geq v > 0, \]

then

\[ \text{Vol} (C) \leq f(v, n). \]

**Remark.** The function $f$ in Lemma 3 can in fact be taken to be

\[ f(v, n) = \text{Vol} (T) \text{Vol} (T \backslash S^{n-1})/v \]

where $T$ is a maximal torus in $O(n)$. (Since all maximal tori in $O(n)$ are conjugate, this function depends only on $v$ and $n$.)

**Proof.** Let $P(C)$ denote the set of all points in $S^{n-1}$ at which $C$ acts freely. Then $P(C)$ is an open dense submanifold of $S^{n-1}$; in fact,

\[ \text{Vol} (P(C)) = \text{Vol} (S^{n-1}). \]

Put the Riemannian metric on $C \backslash P(C)$ such that the orbit map $P(C) \to C \backslash P(C)$ is a Riemannian submersion. Let $\mu_C$ be the measure on $C \backslash P(C)$ determined by this Riemannian metric, then

\[ \text{Vol} (C \backslash S^{n-1}) = \mu_C (C \backslash P(C)) \]

and

\[ (*) \quad \text{Vol} (S^{n-1}) = \int_{C \times C \backslash P(C)} \text{Vol} (C x) d\mu_C. \]

**Special Case.** We first verify Lemma 3 under the extra assumption that $C$ is connected; i.e., $C = C^0$. Then let $T$ be a maximal torus in $O(n)$ containing $C$. Note that $P(T) \subset P(C)$ and

\[ (1) \quad \text{Vol} (C \backslash P(C)) = \text{Vol} (C \backslash P(T)) \]

since $\text{Vol} (P(C)) = \text{Vol} (P(T))$. Consider the principal bundle

\[ (** \quad C \backslash T \to C \backslash P(T) \xrightarrow{p} T \backslash P(T) \]

and notice that $p$ is a Riemannian submersion since $C\backslash T$ acts by isometries on $C\backslash P(T)$. Our argument in this special case is based on the following assertion:

**Claim.** For each point $Cx \in C\backslash P(T)$, the following inequality is true:

$$\text{Vol} \left( C\backslash T(Cx) \right) \leq \text{Vol} \left( T \right) / \text{Vol} \left( C \right).$$

**Proof of Claim.** Note first that $C\backslash T(Cx) = q(Tx)$ where $q : P(T) \rightarrow C\backslash P(T)$ is the orbit map. Also note that

$$\text{Vol} \left( C(tx) \right) = \text{Vol} \left( Cx \right)$$

for each $t \in T$; since $T$ is abelian and $C \subset T$. Hence

$$\text{Vol} \left( Tx \right) = \text{Vol} \left( Cx \right) \text{Vol} \left( q(Tx) \right)$$

since $q$ is a Riemannian submersion; i.e.,

$$\text{(2)} \quad \text{Vol} \left( C\backslash T(Cx) \right) = \text{Vol} \left( Tx / \text{Vol} \left( Cx \right). \right.$$

Consider the Riemannian submersion $f : O(n) \rightarrow S^{n-1}$ defined by $f(g) = gx$, $g \in O(n)$. Note that $f$ maps both $T$ and $C$ diffeomorphically onto $Tx$ and $Cx$, respectively, since $x \in P(T)$. Let $\zeta$ denote $df : g(C) \rightarrow T_x(Cx)$ and $\xi$ denote $df : g(T) \rightarrow T_x(Tx)$. Then note that

$$\text{Vol} \left( Cx \right) = |\det \zeta| \text{Vol} \left( C \right) \text{ and } \text{Vol} \left( Tx \right) = |\det \xi| \text{Vol} \left( T \right)$$

where $\det \zeta$ and $\det \xi$ are computed using orthonormal bases for the relevant vector spaces; therefore, we obtain by dividing these equations that

$$\text{(3)} \quad \text{Vol} \left( Tx / \text{Vol} \left( Cx \right) \right) = |\det \xi| \text{Vol} \left( T \right) / |\det \zeta| \text{Vol} \left( C \right).$$

Concatenating Equations (2) and (3) establishes the Claim once we verify that

$$\text{(4)} \quad |\det \xi| \leq |\det \zeta|.$$  

But inequality (4) is true since $\xi$ is a weakly decreasing linear transformation; i.e.,

$$|\xi(v)| \leq |v|$$

for all $v \in \text{domain} \left( \xi \right)$. This completes the verification of Claim. \hfill \Box

We now complete the proof of the Special Case of Lemma 3. Recall that the projection $p$ in the principal bundle of $(\ast\ast)$ is a Riemannian submersion; hence,

$$\text{(5)} \quad \text{Vol} \left( C\backslash P(T) \right) = \int_{T \in T\backslash P(T)} \text{Vol} \left( C\backslash T(Cx) \right) d\mu_T.$$
Combining Equations (1) and (5) with the Claim and the facts that
\[ \text{Vol} (C \setminus S^{n-1}) = \text{Vol} (C \setminus P(C)) \quad \text{and} \quad \text{Vol} (T \setminus S^{n-1}) = \text{Vol} (T \setminus P(T)) \]
yield the following inequality:
\[ (6) \quad \text{Vol} (C \setminus S^{n-1}) \leq (\text{Vol} (T)/\text{Vol}(C)) \text{Vol} (T \setminus S^{n-1}). \]
Note that inequality (6) is equivalent to
\[ (7) \quad \text{Vol} (C) \leq \text{Vol} (T) \frac{\text{Vol} (T \setminus S^{n-1})}{\text{Vol} (C \setminus S^{n-1})}. \]
But inequality (7) shows that the function \( f(v,n) \) defined in Remark satisfies Lemma 3 in the special case.

We now prove Lemma 3 in general; i.e., without assuming \( C = C^0 \). Let \( T \) be a maximal torus in \( O(n) \) containing \( C^0 \) and \( G \) be the finite group \( C/C^0 \) whose order is denoted by \( |G| \). Note that inequality (7) is valid when we replace \( C \) by \( C^0 \); i.e.,
\[ (8) \quad \text{Vol} (C^0) \leq \text{Vol} (T) \frac{\text{Vol} (T \setminus S^{n-1})}{\text{Vol} (C \setminus S^{n-1})}. \]

Also note that
\[ (9) \quad \text{Vol} (C^0 \setminus S^{n-1}) = |G| \text{Vol} (C \setminus S^{n-1}). \]
This is because
\[ \text{Vol} (C^0 \setminus S^{n-1}) = \text{Vol} (C^0 \setminus P(C^0)) = \text{Vol} (C^0 \setminus P(C)) \]
and
\[ \text{Vol} (C^0 \setminus P(C)) = |G| \text{Vol} (C \setminus P(C)). \]
But we obviously have
\[ (10) \quad \text{Vol} (C) = |G| \text{Vol} (C^0). \]
Substituting Equations (9) and (10) into inequality (8) now yields that inequality (7) is true in general. This shows that the function \( f(v,n) \) defined in Remark satisfies Lemma 3 in general, completing the proof for Lemma 3. \( \square \)

We have the following corollary obtained by directly combining Lemmas 2 and 3:

**Corollary 4.** Given \( v > 0 \), there are only finitely many conjugacy classes of closed subgroups \( C \) of \( O(n) \) such that both \( C^0 \) is abelian and \( \text{Vol} (C \setminus S^{n-1}) \geq v \).

We have now completed discussing the preliminaries needed to verify Fact 1. We’ve shown that to verify Fact 1 it suffices to show that there exists a positive number \( b = b(n, \epsilon, A_0) \) such that
\[ \text{Vol} (\overline{C} \setminus O(n)) \geq b. \]
This assertion is an immediate consequence of Corollary 4. Let 

$$d : C \to \text{Iso} (T_y \hat{V})$$

denote the derivative map at \( y \in \hat{V} \). (Recall that \( C \) is the isotropy subgroup at \( y \) of the action of \( H \) on \( \hat{V} \).) The map \( d \) is a faithful representation. Recall that \( \bar{g} \in F\hat{V} \) is an orthonormal basis for \( T_y \hat{V} \) and that, for each \( c \in C \), \( \bar{c} \in O(n) \) is the matrix representing linear transformation 

$$d(c) : T_y \hat{V} \to T_y \hat{V}$$

with respect to this basis. Hence, to verify Fact 1 it suffices to show that 

$$\text{Vol} (d(C) \setminus S_y) \geq b$$

where \( S_y \) denotes the sphere of radius 1 in \( T_y \hat{V} \).

Recall again the Riemannian manifold \( E \) used in [4, §8] to construct \( \hat{V} \). It has an action of \( \hat{H} \times O(n) \) such that \( \hat{V} = E/O(n) \) and there is a point \( \hat{y} \in E \) such that \( y \) is the orbit \( \hat{y}O(n) \). There is also an isomorphism \( \alpha \) which maps \( \hat{C} \) onto a subgroup \( \tilde{C} \) of \( O(n) \) defined by the equation 

$$g \hat{y} = \hat{y} \alpha (g)$$

for each \( g \in \hat{C} \). (Recall that \( \hat{C} = \phi^{-1}(C) \) where \( \phi : \hat{H} \to H \) is an epimorphism such that the principal \( O(n) \)-bundle projection 

$$\hat{\rho} : E \to \hat{V}$$

is \( \phi \)-equivariant; cf. paragraph 2 of Remarks to Step 2.) Let \( z \) denote the orbit \( \hat{H} \hat{y} \) in \( \hat{H}\backslash E \) and \( S_z \) denote the sphere of radius 1 in \( T_z (\hat{H}\backslash E) \). Then it is easy to construct an \( \alpha \)-equivariant isometry from \( S_y \) to \( S_z \) where \( \hat{C} \) acts on \( S_y \) via the composite representation \( d \circ \phi \). Hence 

$$\text{Vol} (d(C) \setminus S_y) = \text{Vol} (d(\tilde{C}) \setminus S_z)$$

where again \( d : \tilde{C} \to \text{Iso} (T_z (\hat{H}\backslash E)) \) denotes the derivative map at \( z \in \hat{H}\backslash E \). But 

$$\text{Vol} (d(\tilde{C}) \setminus S_z) \geq b$$

because of [4, Lemma 8.5].

This completes the verification of Fact 1. \( \square \)

References


100  F.T. FARRELL AND L.E. JONES


Received December 15, 1997 and revised July 27, 2000. Both authors were supported in part by the National Science Foundation.

DEPARTMENT OF MATHEMATICS
SUNY AT BINGHAMTON
BINGHAMTON, NY 13902-6000
E-mail address: farrell@math.binghamton.edu

DEPARTMENT OF MATHEMATICS
STATE UNIVERSITY OF NEW YORK
STONY BROOK, NY 11794-3651
E-mail address: lejones@math.sunysb.edu
EXCEPTIONAL SURGERY CURVES IN TRIANGULATED 3-MANIFOLDS

Marc Lackenby

For the purposes of this paper, Dehn surgery along a curve $K$ in a 3-manifold $M$ with slope $\sigma$ is ‘exceptional’ if the resulting 3-manifold $M_K(\sigma)$ is reducible or a solid torus, or the core of the surgery solid torus has finite order in $\pi_1(M_K(\sigma))$. We show that, providing the exterior of $K$ is irreducible and atoroidal, and the distance between $\sigma$ and the meridian slope is more than one, and a homology condition is satisfied, then there is (up to ambient isotopy) only a finite number of such exceptional surgery curves in a given compact orientable 3-manifold $M$, with $\partial M$ a (possibly empty) union of tori. Moreover, there is a simple algorithm to find all these surgery curves, which involves inserting tangles into the 3-simplices of any given triangulation of $M$. As a consequence, we deduce some results about the finiteness of certain unknotting operations on knots in the 3-sphere.

1. Introduction.

Consider the following motivating problem from knot theory. Let $L$ be a nontrivial knot in $S^3$. If $K$ is an unknotted curve disjoint from $L$, then Dehn surgery along $K$ with slope $1/q$ has the effect of adding $|q|$ full twists to $L$, yielding a knot $L'$, say. (See Figure 1.2.) Suppose that $L'$ is the unknot, or (more generally) that $L'$ has smaller genus than that of $L$. Then, for a given knot $L$, are there only a finitely many possibilities for $q$ and $K$ (up to ambient isotopy keeping $L$ fixed)? The following theorem deals with this question.

**Theorem 1.1.** Let $L$ be a knot in $S^3$ which is not a nontrivial satellite knot. Let $K$ be an unknotted curve in $S^3$, disjoint from $L$ and having zero linking number with $L$. Let $q$ be an integer with $|q| > 1$. Suppose that $1/q$ surgery about $K$ yields a knot $L'$ with genus($L'$) < genus($L$). Then, for a given knot $L$, there are only finitely many possibilities for $K$ and $q$ up to ambient isotopy keeping $L$ fixed, and there is an algorithm to find them all.

Such ‘unknotting operations’ have been the object of considerable study. For example, the author in [8] dealt with the case where $K$ bounds a disc intersecting $L$ in two points of opposite sign, and proved that if such a
surgery reduces the genus of $L$, then there exists an upper bound on $|q|$ which depends only on $L$, not $K$. Theorem 1.1 gives a great deal more than numerical restrictions on $|q|$. It provides a classification of all such unknotted operations for a given knot $L$, when $|q| > 1$ and the linking number of $K$ and $L$ is zero.

Theorem 1.1 is an almost immediate corollary of new results on Dehn surgery. Let $M$ be an arbitrary compact orientable 3-manifold $M$, with $\partial M$ a (possibly empty) union of tori. (In Theorem 1.1, $M$ is the exterior of the knot $L$.) Our aim is to find the knots $K$ in $M$ with ‘exceptional’ or ‘norm-exceptional’ surgeries, which we define as follows.

**Definition 1.3.** Let $\sigma$ be a slope on $\partial \mathcal{N}(K)$ other than the meridional slope $\mu$. Let $M_K(\sigma)$ be the manifold obtained by Dehn surgery along $K$ via the slope $\sigma$. Then $\sigma$ is an exceptional slope and $K$ is an exceptional surgery curve if any of the following holds:

1. $M_K(\sigma)$ is reducible,
2. $M_K(\sigma)$ is a solid torus, or
3. the core of the surgery solid torus has finite order in $\pi_1(M_K(\sigma))$.

Also, $\sigma$ and $K$ are norm-exceptional if there is some $z \in H_2(M - \text{int}(\mathcal{N}(K)), \partial M)$ which maps to an element $z_\sigma \in H_2(M_K(\sigma), \partial M_K(\sigma))$, such that the Thurston norm of $z_\sigma$ less than the Thurston norm of $z$. (See Section 3 for a definition of the Thurston norm.)

If $K$ satisfies the conditions of Theorem 1.1, then it is a norm-exceptional surgery curve in $M = S^3 - \text{int}(\mathcal{N}(L))$. The reason for distinguishing norm-exceptional surgery curves from the exceptional case is that, in the former situation, our results will be slightly weaker. We restrict attention to knots $K$ with irreducible atoroidal exteriors. For technical reasons, we also have to
assume that $H_2(M - \text{int} (\mathcal{N}(K)), \partial M)$ is nontrivial. This implies in particular that the first Betti number of $M$ must be nonzero.

We will show that the problem of finding exceptional surgery curves in a given 3-manifold $M$ falls naturally into two cases, which depend on $\Delta(\sigma, \mu)$, where $\Delta(\sigma, \mu)$ is the intersection number on $\partial \mathcal{N}(K)$ between the surgery slope $\sigma$ and the meridian slope $\mu$. It is not hard to find examples of 3-manifolds $M$ as above containing an infinite number of pairwise non-isotopic surgery curves $K$ with exceptional or norm-exceptional surgery slopes $\sigma$ satisfying $\Delta(\sigma, \mu) = 1$. (We will do this in Section 12.) However, the main theorem of this paper asserts that, if $\Delta(\sigma, \mu) > 1$, then there is only a finite number of possibilities for $K$ and $\sigma$.

**Theorem 1.4.** Let $M$ be a compact connected orientable 3-manifold, with $\partial M$ a (possibly empty) union of tori. Let $K$ be a knot in $M$ such that $M - \text{int} (\mathcal{N}(K))$ is irreducible and atoroidal, and with $H_2(M - \text{int} (\mathcal{N}(K)), \partial M) \neq 0$. Let $\sigma$ be an exceptional or norm-exceptional slope on $\partial \mathcal{N}(K)$, such that $\Delta(\sigma, \mu) > 1$, where $\mu$ is the meridian slope on $\partial \mathcal{N}(K)$. Then, for a given $M$, there are at most finitely many possibilities for $K$ and $\sigma$ up to ambient isotopy, and there is an algorithm to find them all.

The algorithm is surprisingly simple. We describe it in Section 2. The input into the algorithm is any triangulation of $M$, or the following generalisation of a triangulation. Let $P$ be a (possibly empty) collection of components of $\partial M$. Then a generalised triangulation of $M$ is a representation of $M - P$ as a union of 3-simplices, with some or all of their faces identified in pairs and then possibly with some subcomplex removed. For example, an ideal triangulation is the case where $P = \partial M$ and where the subcomplex removed is the 0-cells. We will also refer to the case where $P = \emptyset$ as a genuine triangulation.

There is a yet simpler algorithm which deals with the $\sigma$-cable of $K$, which is defined to be the knot in $M$ lying on $\partial \mathcal{N}(K)$ having slope $\sigma$. Recall that a tangle is a (possibly empty) collection of disjoint arcs properly embedded in a 3-ball. Two tangles are identified if there is an isotopy of the 3-ball which is fixed on the boundary and which takes one tangle to the other.

**Theorem 1.5.** There is a finite collection of tangles, each lying in a 3-simplex and with the following property. Let $M$, $K$ and $\mu$ be as in Theorem 1.4, and let $\sigma$ be an exceptional slope on $\partial \mathcal{N}(K)$ with $\Delta(\sigma, \mu) > 1$. Pick any generalised triangulation of $M$. Then, we may insert a tangle from this finite collection into each 3-simplex, in such a way that the tangles join to form a knot which is ambient isotopic to the $\sigma$-cable of $K$. This finite collection of tangles is constructible and is independent of $M$, $K$ and $\sigma$.

Since these tangles are defined up to isotopy of the 3-simplex $\Delta^3$ which is fixed on $\partial \Delta^3$, Theorem 1.5 immediately gives that there are only finitely many possibilities in $M$ for the $\sigma$-cable of $K$. 
Theorem 1.5 is a very surprising result. If $\Delta(\sigma, \mu)$ is large, then one would expect the $\sigma$-cable of $K$ to intersect the triangulation of $M$ in a complicated way. But the above result asserts that one can control this complexity. It is also surprising that the same finite collection of tangles should work for all $M$ and all triangulations. Note that in Theorem 1.5 we did not assume that $K$ and $\sigma$ were norm-exceptional. In this case, we have the following slightly weaker result.

**Theorem 1.6.** Let $M$, $K$ and $\mu$ be as in Theorem 1.4, and let $\sigma$ be a norm-exceptional slope on $\partial N(K)$ with $\Delta(\sigma, \mu) > 1$. If $M$ is closed, pick any genuine triangulation of $M$. In the case where $M$ has nonempty boundary, pick any ideal triangulation of $M$. Then, we may insert into each 3-simplex a tangle from the finite collection of Theorem 1.5, in such a way that the tangles join to form a knot which is ambient isotopic to the $\sigma$-cable of $K$.

It is in fact possible to write down explicitly this list of tangles. We will give an algorithm in Section 11 to do this. We have not actually run this algorithm, since the task is fairly lengthy and is more suited to computer implementation.

## 2. The algorithm to find all possibilities for $K$ and $\sigma$.

In this section, we describe the algorithm for finding, in a given 3-manifold $M$, all surgery curves $K$ with exceptional or norm-exceptional surgery slopes $\sigma$, satisfying the conditions of Theorem 1.4. The first (and most important) step is to construct a finite list of possibilities for $K$ and $\sigma$, some of which may be neither norm-exceptional nor exceptional.

We will in Section 11 construct a finite collection of graphs, each embedded in a 3-simplex $\Delta^3$. Each graph $G$ meets $\partial \Delta^3$ in a collection of vertices. These vertices have valence one and lie in the interior of the 2-simplices of $\partial \Delta^3$. There is also a specified regular neighbourhood $N(G)$ and a collection of disjoint arcs properly embedded in $\Delta^3$, lying in $\partial N(G)$. Each arc is assigned one of two labels, $\gamma$ or $\tau$. Each graph $G$ (together with $N(G)$ and the arcs $\gamma$ and $\tau$) is defined up to isotopy of $\Delta^3$ which is fixed on $\partial \Delta^3$.

We will show during the course of the paper that it is possible to ambient isotope $K$ and $\sigma$, and to find a handle structure $H$ on $N(K)$ with the following properties. Each tetrahedron $\Delta^3$ of the generalised triangulation of $M$ intersects the 0-handles and 1-handles of $H$ in $N(G)$, where $G$ is one of the graphs mentioned above. The 2-handles of $H$ will be attached to $N(G)$ along the arcs $\tau$. There will also be a curve of slope $\sigma$ on $\partial N(K)$ which intersects $\Delta^3$ in the arcs $\gamma$.

The algorithm to find all possibilities for $K$ and $\sigma$ therefore proceeds as follows. We insert one of these graphs into each 3-simplex of the generalised triangulation of $M$. If $\Delta^2$ is any 2-simplex of $M$ adjacent to two 3-simplices, and $G_1$ and $G_2$ are the graphs inserted into these two 3-simplices, then we
insist that $\mathcal{N}(G_1) \cap \Delta^2 = \mathcal{N}(G_2) \cap \Delta^2$, and also that the endpoints in $\Delta^2$ of the arcs labelled $\gamma$ (respectively, $\tau$) in $G_1$ correspond precisely with the endpoints in $\Delta^2$ of the arcs labelled $\gamma$ (respectively, $\tau$) in $G_2$. Thus, the graphs $G$ combine to form a graph (which we also call $G$) embedded in $M$. We insist that $G$ is disjoint from $\partial M$. The collections of arcs combine to form a collection of curves $\gamma$ and $\tau$ properly embedded in $M$ and lying in $\partial \mathcal{N}(G)$. We insist that each component of $\gamma$ and $\tau$ is a simple closed curve, and that $\gamma$ is connected. Since there are only finitely many 3-simplices in the representation of $M$ and there are only finitely many possibilities for $\mathcal{N}(G) \cap \Delta^3$, $\tau \cap \Delta^3$ and $\gamma \cap \Delta^3$ for each 3-simplex $\Delta^3$ in $M$, there are only finitely many possibilities for $\mathcal{N}(G)$, $\tau$ and $\gamma$. The handlebody $\mathcal{N}(G)$ and curves $\tau$ specify a handle structure of a 3-manifold $M'$, which is a candidate for $\mathcal{N}(K)$. At this stage, $M'$ may be something other than a solid torus.

The algorithm proceeds by calculating $H_1(M')$ and $H_1(\partial M')$, and the map $H_1(\partial M') \to H_1(M')$ induced by inclusion. If this is not the map $\mathbb{Z} \oplus \mathbb{Z} \to \mathbb{Z}$ that is projection onto one factor, we stop. If it is, we can algorithmically find generators $\lambda$ and $\mu$ of $H_1(\partial \mathcal{N}(M'))$ such that $\lambda$ maps to $1 \in H_1(M')$, and $\mu$ maps to $0 \in H_1(M')$. We can construct a simple closed curve representative of $\lambda$ on $\partial \mathcal{N}(M')$ which avoids the 2-handles of $M'$. If $M'$ is $\mathcal{N}(K)$, then this curve is ambient isotopic in $M$ to $K$. The simple closed curve $\gamma$ has slope $\sigma$. Thus, we have constructed $K$ and $\sigma$. If we wish, we can also calculate $\Delta(\lambda, \sigma)$ and $\Delta(\mu, \sigma)$. If $K$ is homologically trivial, this (together with orientation information) gives the rational number $p/q$ associated with $\sigma$.

The above algorithm constructs a finite number of possibilities for $K$ and $\sigma$. We now wish to rule out the cases where $K$ and $\sigma$ are neither exceptional nor norm-exceptional. We construct the manifold $M_K(\sigma)$. There is an algorithm to determine whether $M_K(\sigma)$ is reducible ([5] and [11]), and there is an algorithm to determine whether $M_K(\sigma)$ is a solid torus ([5] and [11]). The assumption that $H_2(M - \text{int}(\mathcal{N}(K)), \partial M)$ is nontrivial implies that $H_1(M_K(\sigma))$ is infinite and hence that $\pi_1(M_K(\sigma))$ is infinite. If $M_K(\sigma)$ is irreducible, then according to Corollary 9.9 of [3], $\pi_1(M_K(\sigma))$ is torsion-free. Thus, if the core of the surgery solid torus in $M_K(\sigma)$ has finite order in $\pi_1(M_K(\sigma))$, then it is homotopically trivial. There is an algorithm to determine this, since the word problem is soluble for the fundamental groups of Haken 3-manifolds [13]. Finally, there is an algorithm to find the unit ball of the Thurston norm (Algorithm 5.9 of [12]), and so we can determine whether $\sigma$ is norm-exceptional.

3. The sutured manifold theory background.

The definition of an exceptional surgery was specifically designed so that sutured manifold theory can be applied. Sutured manifolds were defined
and studied by Gabai [1] who used them to construct taut foliations on certain 3-manifolds. In this section, we will outline a version of the theory due to Scharlemann [10]. Almost everything in this section can be found elsewhere, mostly in Scharlemann’s paper [10]. We include it here because it is absolutely central to our argument, but a reader familiar with the theory of sutured manifolds may safely skip this section.

Sutured manifold theory is intimately linked to the Thurston norm. Here, we give a definition of the Thurston norm and some related definitions of tautness.

Let \( S \) be a compact oriented surface embedded in a compact oriented 3-manifold \( M \). Let \( \chi(S) \) denote its Euler characteristic. If \( S \) is connected, define \( \chi_-(S) = \max\{0, -\chi(S)\} \). When \( S \) is not connected, define \( \chi_-(S) \) to be the sum of \( \chi_-(S_i) \) over all the components \( S_i \) of \( S \).

Let \( P \) be a subset of \( \partial M \), and let \( z \) be an element of \( H_2(M, P) \) which is represented by some embedded compact oriented surface. Then the Thurston norm of \( z \) is given by

\[
x(z) = \min\{\chi_-(S) : S \text{ is an embedded surface representing } z\}.
\]

Let \( (S, \partial S) \subset (M, \partial M) \) be an oriented compact surface embedded in \( M \). Let \( P \) be a subset of \( \partial M \) which contains \( \partial S \). Then \( S \) is norm-minimising in \( H_2(M, P) \) if \( x([S, \partial S]) = \chi_-(S) \). In the case where \( P = \partial S \), then \( S \) is taut if it is incompressible and norm minimising in \( H_2(M, P) \). This use of the word ‘taut’ is not entirely standard; some authors (for example, [12]) insist only that \( S \) be incompressible and norm-minimising in its class in \( H_2(M, \partial M) \). However, our definition is more suited to sutured manifold theory.

A sutured manifold \( (M, \gamma) \) is a compact oriented 3-manifold \( M \), with \( \partial M \) decomposed into two subsurfaces \( R_- \) and \( R_+ \), such that \( R_- \cup R_+ = \partial M \) and \( R_- \cap R_+ = \gamma \), where \( \gamma \) is a union of disjoint simple closed curves, known as the sutures. The subsurfaces \( R_- \) and \( R_+ \) are oriented so that the normal vectors of \( R_- \) (respectively, \( R_+ \)) point into (respectively, out of) \( M \). The symbol \( R_{\pm} \) will denote ‘\( R_- \) or \( R_+ \)’. When we wish to emphasise a particular sutured manifold, we will use the symbol \( R_{\pm}(M) \).

A sutured manifold \( (M, \gamma) \) is taut if \( M \) is irreducible, and \( R_- \) and \( R_+ \) are taut. For example it is not hard to show the following. Suppose that \( \partial M \) is a (possibly empty) union of tori, and that \( R_- = \partial M \) and \( R_+ = \emptyset \). Then \( (M, \emptyset) \) is taut if and only is \( M \) is neither reducible nor a solid torus.

One of the main techniques of the theory is to decompose a sutured manifold along a properly embedded surface. If \( (M, \gamma) \) is a sutured manifold, and \( S \) is an oriented surface properly embedded in \( M \), with \( \partial S \) and \( \gamma \) in general position, then \( M_S = M - \text{int}(N(S)) \) inherits a sutured manifold structure.
\((M_S, \gamma_S)\). This is written \( (M, \gamma) \xrightarrow{S} (M_S, \gamma_S) \). This decomposition is said to be taut if \((M, \gamma)\) and \((M_S, \gamma_S)\) are both taut.

If \((M, \gamma)\) is a connected taut sutured manifold and \(z\) is any nonzero homology class in \( H_2(M, \partial M) \), then (Theorem 2.6 of \([10]\)) there is a taut decomposition \( (M, \gamma) \xrightarrow{S} (M_S, \gamma_S) \) such that:

(i) No curve of \( \partial S \) bounds a disc in \( R^\pm(M) \),
(ii) no component \( X \) of \( M_S \) has \( \partial X \subset R^-(M_S) \) or \( \partial X \subset R^+(M_S) \), and
(iii) \([S, \partial S] = z \in H_2(M, \partial M)\).

When \( S \) satisfies (i), we will say that \( \partial S \) has essential intersection with \( R^\pm(M) \). It is not hard to show that if \( (M, \gamma) \xrightarrow{S} (M_S, \gamma_S) \) is a taut decomposition and \( \partial S \) has essential intersection with \( R^\pm(M) \), then \( S \) itself is taut.

Thus if \( H_2(M, \partial M) \neq 0 \), we may perform a taut sutured manifold decomposition along a taut surface having essential intersection with \( R^\pm(M) \). But if \( H_2(M, \partial M) \) is trivial, then it is a classical fact that \( \partial M \) is a (possibly empty) union of 2-spheres. If \( M \) is irreducible, then this implies that either \( \partial M = \emptyset \) or \( M \) is a collection of 3-balls. Using this argument, Gabai proved that, if \((M, \gamma)\) is a connected taut sutured manifold and \( H_2(M, \partial M) \neq 0 \), then there is a sequence of taut decompositions

\[(M, \gamma) = (M_1, \gamma_1) \xrightarrow{S_1} (M_2, \gamma_2) \xrightarrow{S_2} \ldots \xrightarrow{S_{n-1}} (M_n, \gamma_n), \]

with \( M_n \), a union of 3-balls.

An important step in Gabai’s argument is to show that this sequence of decompositions cannot continue indefinitely. This is not at all obvious. In the case where \( S_i \) is incompressible and \( \partial \)-incompressible in \( M_i \) and has no component parallel to a subsurface of \( \partial M_i \), it was proved by Haken \([2]\) that such a sequence of decompositions must eventually terminate. However, it is sometimes necessary to consider surfaces \( S_i \) which are \( \partial \)-compressible. Nevertheless, Gabai constructed a (complicated) argument which proved that this sequence of taut sutured manifold decompositions can be guaranteed to terminate. He did this by defining a complexity of a sutured manifold and then arguing by induction. In Section 5, we will offer a new definition of complexity for a sutured manifold with a given handle decomposition.

There is an extremely useful property of sutured manifold decompositions, which is summarised in the phrase ‘tautness usually pulls back’. It is this property which makes sutured manifold theory distinctly different from the theory of Haken manifolds.

**Theorem 3.1** (Theorem 3.6 of \([10]\)). Let \( (M, \gamma) \xrightarrow{S} (M_S, \gamma_S) \) be a decomposition, where \( \partial S \) has essential intersection with \( R^\pm(M) \), and where no component of \( S \) is a compression disc for a torus component of \( R^\pm(M) \). If \((M_S, \gamma_S)\) is taut, then so is \((M, \gamma)\).
There is a partial converse to this theorem which can be useful. If \( D \) is a disc properly embedded in \( M \) intersecting \( \gamma \) transversely in two points, then \( D \) is known as a \textit{product disc}. If \( A \) is an annulus properly embedded in \( M \) with one component of \( \partial A \) in \( R_- \) and one in \( R_+ \), then \( A \) is known as a \textit{product annulus}. These surfaces play a useful rôle, since if \((M, \gamma) \rightarrow (M_S, \gamma_S)\) is a decomposition along a product disc or an incompressible product annulus, then \((M, \gamma)\) is taut if and only if \((M_S, \gamma_S)\) is taut.

We have now described enough sutured manifold theory to explain the definition of an exceptional surgery curve, given in Section 1. The following argument is well-known, and is due to Gabai \cite{G}. Let \( K \) be a knot in a compact connected orientable 3-manifold \( M \), where \( \partial M \) is (possibly empty) union of tori. If \( M - \text{int}(\mathcal{N}(K)) \) is neither reducible nor a solid torus, then \((M - \text{int}(\mathcal{N}(K)), \emptyset)\) is a taut sutured manifold, with \( \mathcal{R}_- = \partial M \cup \partial \mathcal{N}(K) \). If \( H_2(M - \text{int}(\mathcal{N}(K)), \partial M) \neq 0 \), then we may perform a taut sutured manifold decomposition

\[
(M - \text{int}(\mathcal{N}(K))) \rightarrow (M_2 - \text{int}(\mathcal{N}(K))),
\]

such that:

- \( S_1 \) is disjoint from \( \partial \mathcal{N}(K) \),
- no simple closed curve of \( \partial S_1 \) bounds a disc in \( \mathcal{R}_+(M) \), and
- no component \( X \) of \( M_2 \) has \( \partial X \subset \mathcal{R}_-(M_2) \) or \( \partial X \subset \mathcal{R}_+(M_2) \).

If \( K \) is norm-exceptional, we insist that \([S_1, \partial S_1] = z \in H_2(M - \text{int}(\mathcal{N}(K)), \partial M)\), where \( z \) is the relevant homology class from Definition 1.3. Repeating this process, we construct a sequence of taut sutured manifold decompositions

\[
(M - \text{int}(\mathcal{N}(K)), \emptyset) \rightarrow \cdots \rightarrow (M_n - \text{int}(\mathcal{N}(K)), \gamma_n),
\]

satisfying the above conditions, and where \( H_2(M_n - \text{int}(\mathcal{N}(K)), \partial M_n) = 0 \).

If \( M - \text{int}(\mathcal{N}(K)) \) is atoroidal, then it is possible to show that this implies that \( M_n \) is a solid torus regular neighbourhood of \( K \) and possibly some 3-balls. No component \( X \) of \( M_n \) has \( \partial X \subset \mathcal{R}_-(M_n) \) or \( \partial X \subset \mathcal{R}_+(M_n) \). In particular, if \( X \) is the component of \( M_n \) containing \( K \), then \( \gamma_n \cap X \) is a collection of essential curves on \( \partial X \), parallel to some slope \( \rho \), say, on \( \partial \mathcal{N}(K) \). If we Dehn fill \( M - \text{int}(\mathcal{N}(K)) \) via any slope \( \tau \) on \( \partial \mathcal{N}(K) \), then \( M_n - \text{int}(\mathcal{N}(K)) \) is filled to become a 3-manifold \( M_n(\tau) \) which is a solid torus and some 3-balls. Now, \( M_n(\tau) \) inherits a sutured manifold structure \((M_n(\tau), \gamma_n)\) from \( M_n - \text{int}(\mathcal{N}(K)) \), which is taut if the surgery slope \( \tau \) is not the slope \( \rho \) of the sutures. Since tautness pulls back, this implies that

\[
(M_K(\tau), \emptyset) \rightarrow \cdots \rightarrow (M_n(\tau), \gamma_n)
\]

is a sequence of taut sutured manifolds, with each \( S_i \) taut in \( M_i(\tau) \). This implies that:

(i) \( M_K(\tau) \) is irreducible,
(ii) $M_K(\tau)$ is not a solid torus,
(iii) the core of the surgery solid torus in $M_K(\tau)$ has infinite order in $\pi_1(M_K(\tau))$, and
(iv) $S_1$ is taut in $M_K(\tau)$.

Now, if $\sigma$ is an exceptional or norm-exceptional surgery slope on $\partial N(K)$, then at least one of the above cannot be true for $M_K(\sigma)$. Thus, $\sigma$ must be the slope $\rho$ which is parallel to the sutures in $M_n$. We assume in Theorems 1.4, 1.5 and 1.6 that $\Delta(\sigma,\mu) > 1$ which implies in particular that $\sigma \neq \mu$. Thus, the facts (i)-(iv) above are true for $\tau = \mu$, and also

$$(M,\emptyset) \xrightarrow{S_1 \ldots S_{n-1}} (M_n,\gamma_n)$$

is a taut sutured manifold sequence. Each component of $\gamma_n$ lies inside $M$ as the $\sigma$-cable of $K$, or as an unknotted curve. The idea behind Theorems 1.4, 1.5 and 1.6 is (roughly speaking) inductively to find nice embeddings of $M_i$ in $M$. In particular, we will show that we can arrange that $\gamma_n \cap \Delta^3$ is one of a finite list of possibilities for each 3-simplex $\Delta^3$ of $M$. Since one component of $\gamma_n$ is the $\sigma$-cable of $K$, this will establish Theorem 1.5.

Thus, our definition of an exceptional surgery curve fits neatly into the sutured manifold setting. The sutured manifold theory which we have outlined above formed the basis for a theorem in [6] which will be an important technical tool in this paper. This result (Theorem 1.4 of [6]) deals with the interaction of exceptional surgery curves and embedded surfaces in a sutured manifold, and is given below.

**Theorem 3.2.** Let $(M,\gamma)$ be a taut sutured manifold, let $K$ be a knot in $M$ and let $\sigma$ be a slope on $\partial N(K)$. Suppose that at least one of the following is true:

(i) $\sigma$ is an exceptional surgery slope, or
(ii) $\sigma$ is a norm-exceptional surgery slope, $\partial M$ is a (possibly empty) union of tori and $\gamma = \emptyset$.

Suppose that $\Delta(\sigma,\mu) > 1$, where $\mu$ is the meridian slope on $\partial N(K)$. Suppose also that $M - \text{int}(N(K))$ is irreducible and atoroidal and that $H_2(M - \text{int}(N(K)),\partial M) \neq 0$. Let $F$ be a surface properly embedded in $M$, with components $F_1,\ldots,F_n$, none of which is a sphere or disc disjoint from $\gamma$. Then there is an ambient isotopy of $K$ in $M$, after which we have the following inequality for each $i$:

$$|K \cap F_i| \leq \frac{-2\chi(F_i) + |\gamma \cap F_i|}{2(\Delta(\sigma,\mu) - 1)}.$$

The numerator $-2\chi(F_i) + |\gamma \cap F_i|$ is known as the index $I(F_i)$ of $F_i$. Note in particular that a product disc and an annulus disjoint from $\gamma$ both have zero index. Thus, if $\Delta(\sigma,\mu) > 1$, Theorem 3.2 implies that we may ambient isotope $K$ off a collection of product discs and annuli disjoint from
Given that such surfaces play a useful technical rôle in sutured manifold theory, this will be very convenient. In fact, this is the only point in proof of Theorems 1.4, 1.5 and 1.6 where we use that $\Delta(\sigma, \mu) > 1$.


Recall that we are given a generalised triangulation of $M$. From this, we will construct the dual handle decomposition, which associates an $i$-handle with each $(3 - i)$-simplex of $M$ not lying entirely in $\partial M$. For this dual handle decomposition, the boundary of each 0-handle has at most four discs of intersection with the 1-handles, and each 1-handle has at most three discs of intersection with the 2-handles. An example is given below.

We will now give some definitions and conventions regarding handle decompositions. We will throughout this paper denote the $i$-handles of a handle decomposition by $H^i$. Henceforth, we will only consider handle decompositions of $n$-manifolds with the following properties:

- For $i > 0$, the $i$-handles are attached to $\bigcup_{j < i} H^j$.
- If $H_i = D^{n-i} \times D^i$ (respectively, $H_j = D^{n-j} \times D^j$) is an $i$-handle (respectively, $j$-handle) with $j < i$, then $H_i \cap H_j = E \times D^i = D^{n-i} \times F$ for some submanifold $E$ (respectively, $F$) of $\partial D^{n-j}$ (respectively, $\partial D^i$).

In words, the second condition requires that the attaching map of each handle respects the product structures of the handles to which it is attached. For a 3-manifold, this is relevant only for $j = 1$ and $i = 2$. In the case of a handle decomposition of a 3-manifold, we also insist that:

- No 2-handle is disjoint from $H^1$.

![Figure 4.1.](image-url)
We will use the term *handle structure* for a decomposition satisfying these conditions. Note in particular that the dual handle decomposition of a 3-manifold arising from a generalised triangulation has these properties. We will use \( \mathcal{H} \) to denote a handle structure, but occasionally, we will also write \( \mathcal{H}(M) \) when we wish to emphasise the manifold \( M \). Note also that a handle structure \( \mathcal{H} \) on a 3-manifold \( M \) induces a handle structure on \( \partial M \), which we will usually write as \( \mathcal{H}(\partial M) \).

We will fix a handle structure \( \mathcal{H} \) of \( M \), and then will consider embedded submanifolds of \( M \). We wish to ensure that each submanifold lies inside \( \mathcal{H} \) in a manageable way. The relevant notions are ‘vertical’ and ‘standard’ form, the first of which we now define.

**Definition 4.2.** Let \( M \) be an \( n \)-manifold with a handle structure \( \mathcal{H} \). Let \( S \) be an \((n - 1)\)-manifold properly embedded in \( M \). Then \( S \) is in vertical form if, for each \( i \)-handle \( D^{n-i} \times D^i \) of \( \mathcal{H} \), we have \( S \cap (D^{n-i} \times D^i) = E \times D^i \), where \( E \) is a properly embedded submanifold of \( D^{n-i} \). In particular, \( S \) is disjoint from \( \mathcal{H}^n \).

The only two cases which we will consider are where \( n = 2 \) or \( n = 3 \). Examples of 2-manifolds in vertical form in a 3-manifold are given in Fig. 4.3. The relevance of vertical form is its ubiquity.

**Lemma 4.4.** Let \( M \) be an \( n \)-manifold with a handle structure \( \mathcal{H} \), and let \( S \) be an \((n - 1)\)-manifold properly embedded in \( M \). Then there is an ambient isotopy which takes \( S \) into vertical form with respect to \( \mathcal{H} \).

*Proof.* We perform a sequence of ambient isotopies. The first pulls \( S \) off \( \mathcal{H}^n \). The second places \( S \) in vertical form in \( \mathcal{H}^{n-1} \), and so on. Let \( C_i \) be the cores of \( \mathcal{H}^i \); thus \( \mathcal{H}^i = C_i \times D^i \). We perform an ambient isotopy which makes \( S \) transverse to \( C_i \). By construction, \( S \) is already vertical in \( \mathcal{H}^j \) for \( j > i \), and so we may take this isotopy to be supported in \( \mathcal{H}^{i} - \partial \mathcal{H}^{i} \). After the isotopy, we may find a small disc \( D^j_0 \subset \text{int}(D^j) \), such that \( S \cap (C_i \times D^j_0) = E_i \times D^j_0 \),

---

**Figure 4.3.**
for some submanifold \( E_i \) of \( C_i \). Then we may use the product structure on \( D^i - \text{int}(D^i_0) \cong S^{i-1} \times I \) to ambient isotope \( C_i \times D^i_0 \) onto \( C_i \times D^i = \mathcal{H}^i \). We can take this isometry of \( M \) to be supported in an arbitrarily small neighbourhood of \( \mathcal{H}_i \), and also to leave \( S \cap \mathcal{H}_j \) invariant for \( j > i \). After performing these isotopies for \( i = n, n-1, \ldots, 0 \), we finish with \( S \) in vertical form. \( \square \)

If \( (M, \gamma) \) is a sutured manifold with a handle structure \( \mathcal{H} \), then by Lemma 4.4 there is an isotopy of \( \partial M \) which takes \( \gamma \) into vertical form (with respect to the induced handle decomposition \( \mathcal{H}(\partial M) \) on \( \partial M \)). This isotopy of \( \partial M \) extends to an isotopy of \( M \). We can therefore assume that \( \gamma \) is in vertical form in \( \mathcal{H}(\partial M) \), and we will henceforth make this assumption.

If \( S \) is any surface properly embedded in \( M \), we would like to ensure that we can place \( S \) in vertical form, and still keep \( \gamma \) vertical in \( \mathcal{H}(\partial M) \). This is the purpose of the following lemma.

**Lemma 4.5.** Let \( (M, \gamma) \) be a sutured manifold with a handle structure \( \mathcal{H} \), such that \( \gamma \) is vertical in \( \mathcal{H}(\partial M) \). If \( S \) is a surface properly embedded in \( M \), in general position with respect to \( \gamma \), then there is an ambient isotopy which leaves \( \gamma \) invariant and which moves \( S \) into vertical form.

**Proof.** The first two steps of the ambient isotopy in Lemma 4.4 are supported in a small neighbourhood of \( \mathcal{H}^3 \cup \mathcal{H}^2 \). Hence, we may assume that it leaves \( \gamma \) fixed. Since \( S \) and \( \gamma \) are in general position, we may pick the co-cores \( C_1 \) of the 1-handles so that \( C_1 \cap S \cap \gamma = \emptyset \). The ambient isotopy supported in a neighbourhood of \( \mathcal{H}^1 \) can then be taken to leave \( \gamma \) invariant. There is no restriction on \( S \cap \mathcal{H}^0 \), once \( S \) lies in the remaining handles in the correct way. Hence, we have ambient isotoped \( S \) into vertical form, leaving \( \gamma \) invariant. \( \square \)

For inductive purposes, we define a notion of complexity for surfaces in vertical form in a handle structure of a 3-manifold.

**Definition 4.6.** The **complexity** of a vertical surface \( S \) is the ordered pair of integers \((|S \cap \mathcal{H}_2|, |\partial S \cap \mathcal{H}_1|)\).

We order these pairs lexicographically. In other words, the pairs \((n_1, n_2)\) and \((m_1, m_2)\) satisfy \((n_1, n_2) > (m_1, m_2)\) precisely when:

- \( n_1 > m_1 \), or
- \( n_1 = m_1 \) and \( n_2 > m_2 \).

It is clear that this ordering is a well-ordering.

In the case of surfaces in 3-manifolds, there is a notion which is a little stronger than vertical form.

**Definition 4.7.** Let \( S \) be a vertical surface in a handle structure \( \mathcal{H} \) of a 3-manifold \( M \). Then \( S \) is **standard** if \( S \) intersects each handle of \( \mathcal{H} \) in a (possibly empty) collection of discs.
Examples of surfaces in standard form are given in Fig. 4.8. A general surface $S$ in $M$ might not have a representation in standard form, but if $S$ is incompressible and $M$ is irreducible, then we now show that it can be ambient isotoped into standard form.

**Lemma 4.9.** Let $(M, \gamma)$ be an irreducible sutured manifold with a handle structure $\mathcal{H}$. Let $S$ be a vertical incompressible surface properly embedded in $M$, with no component of $S$ a 2-sphere. Then there is an ambient isotopy of $S$ which leaves $\gamma$ fixed and which takes $S$ into standard form without increasing its complexity.

**Proof.** If $S$ is not in standard form, then it must differ from standard form in some 1-handle or some 0-handle of $\mathcal{H}$. Suppose first that, in some 1-handle $H_1 = D^2 \times D^1$, there is a component of $S \cap H_1$ which is $\alpha \times D^1$, for a simple closed curve $\alpha$. If both curves of $\alpha \times \partial D^1$ bound discs in $\mathcal{H}^0$, then $S$ has a 2-sphere component. Hence, we may assume that $S$ differs from standard form in some 0-handle $H_0$. That is, suppose that $S \cap H_0$ is not a union of discs. Then, since no component of $S$ is a 2-sphere, $S \cap H_0$ is compressible in $H_0$, via a compression disc $D$. Since $S$ is incompressible, $\partial D$ bounds a disc $D'$ in $S$. The disc $D'$ does not lie wholly in $H_0$, and so must intersect $\mathcal{H}^1$. As $M$ is irreducible, we may ambient isotope $S$, taking $D'$ onto $D$. This does not increase the complexity of $S$, and it reduces the number of components of $S \cap \mathcal{H}^1$. Hence, this process terminates with $S$ in standard form. The isotopy at each stage leaves $\partial M$ (and hence $\gamma$) fixed. \hfill $\square$

We may therefore assume that if $S$ and $M$ satisfy the conditions of Lemma 4.9, then $S$ is in standard form. We will now show that, if $(M_S, \gamma_S)$ is the sutured manifold resulting from the decomposition along $S$, then $M_S$ has an induced handle structure with $\gamma_S$ in vertical form in $\mathcal{H}(\partial M_S)$.

If $H$ is an $i$-handle $D^{3-i} \times D^i$ of $\mathcal{H}(M)$, then each component of $H - \text{int}(\mathcal{N}(S))$ inherits a structure $X \times D^i$, where $X$ is a $(3-i)$-submanifold of $D^{3-i}$. This is true because $S$ is vertical. Since $S$ is standard, then each component of $X$ is a copy of $D^{3-i}$, and so each component of $H - \text{int}(\mathcal{N}(S))$ has the structure of an $i$-handle. These handles combine to give a handle
structure on $M_S$. The curves $\gamma_S$ are a subgraph of the graph $\partial N(\partial S) \cup \gamma$. Since $\partial S$ and $\gamma$ are both vertical in $\mathcal{H}(\partial M)$, the curves $\gamma_S$ are then vertical in $\mathcal{H}(\partial M_S)$.

It is a very useful property that $(M_S, \gamma_S)$ inherits a handle structure from that of $(M, \gamma)$. It is the basis for an inductive proof of Theorems 1.4, 1.5 and 1.6. However, to construct such a proof, we need to define a ‘complexity’ for handle structures.

5. Complexity of handle structures of sutured manifolds.

We will now define a notion of complexity for a handle structure $\mathcal{H}$ of a sutured manifold $(M, \gamma)$. We will focus on the 2-spheres $\partial \mathcal{H}^0$. Lying in these 2-spheres, there is the surface $\partial \mathcal{H}^0 \cap (\mathcal{H}^1 \cup \mathcal{H}^2)$. We denote this surface by $\mathcal{F}(\mathcal{H})$, or sometimes simply $\mathcal{F}$.

![Figure 5.1.](image)

Recall from Section 4 that we insisted that no 2-handle of $\mathcal{H}$ is disjoint from $\mathcal{H}^1$. Therefore, $\mathcal{F}$ inherits a handle structure, with $\partial \mathcal{H}^0 \cap \mathcal{H}^1$ forming the 0-handles of $\mathcal{F}$ (which we denote by $\mathcal{F}^0$), and $\partial \mathcal{H}^0 \cap \mathcal{H}^2$ forming the 1-handles of $\mathcal{F}$ (which we denote by $\mathcal{F}^1$). Note that each component of the surface $\text{cl}(\partial \mathcal{H}^0 - \mathcal{F})$ lies either in $\partial M$ or in $\partial \mathcal{H}^3$, and the curves $\gamma \cap \partial \mathcal{H}^0$ are properly embedded in $\text{cl}(\partial \mathcal{H}^0 - \mathcal{F})$.

If $S$ is in standard form, then the simple closed curves $S \cap \partial \mathcal{H}^0$ satisfy the following (fairly weak) restrictions:

- $S \cap \partial \mathcal{H}^0$ is disjoint from $\mathcal{H}^3$, 

where repetitions are retained. If $X$ follows. If $X$ extends each of these sequences by concatenating with an infinite sequence of

$C_1(F) = |F \cap F^1| + 1,$
$C_2(F) = I(F),$
$C_3(F) = |\partial F|.$

The nature of $F$ will determine the complexity of $H$. One invariant of $F$ will be of particular importance, namely its index. Recall from Section 3 that the index $I(F)$ of a component $F$ of $F$ is defined to be

$$I(F) = -2\chi(F) + |F \cap \gamma|.$$  

If $V$ is a 0-handle of $F$, then the valence of $V$ is the number of arcs of $V \cap F^1$. We also define the index of $V$ to be

$$I(V) = |V \cap F^1| + |V \cap \gamma| - 2.$$  

The reason for this terminology is that

$$I(F) = \sum_{V \in F \cap \partial H^0} I(V).$$  

For each component $F$ of $F$, we define the following integers:

$C_1(F) = |F \cap F^1| + 1,$
$C_2(F) = I(F),$
$C_3(F) = |\partial F|.$

The $F$-complexity set $C_F(H)$ of $H$ is defined to be the set of ordered triples $C_F(H) = \{(C_1(F), C_2(F), C_3(F)) : F$ a component of $F$ with $I(F) > 0\}$, where repetitions are retained. If $X$ is a subset of $M$, with $X \cap F$ a nonempty collection of components of $F$, then we similarly define

$$C_F(X) = \{(C_1(F), C_2(F), C_3(F)) : F$$

a component of $X \cap F$ with $I(F) > 0\},$$

where again repetitions are retained.

An example is given in Fig. 5.2 of how $F$ and its complexity behave when $H$ is decomposed along a surface $S$.

We compare the triples $(C_1(F), C_2(F), C_3(F))$ and $(C_1(F'), C_2(F'), C_3(F'))$ by defining $(C_1(F), C_2(F), C_3(F)) > (C_1(F'), C_2(F'), C_3(F'))$ if:

- $C_1(F) > C_1(F')$, or
- $C_1(F) = C_1(F')$ and $C_2(F) > C_2(F')$, or
- $C_1(F) = C_1(F')$ and $C_2(F) = C_2(F')$ and $C_3(F) > C_3(F').$

It is clear that this is a total ordering and a well-ordering.

We define a total order on the $F$-complexity of handle structures, as follows. If $H$ and $H'$ are two handle structures, we order their $F$-complexity sets $C_F(H)$ and $C_F(H')$ into two non-increasing sequences of triples. We extend each of these sequences by concatenating with an infinite sequence of triples $(0, 0, 0)$. (Note that always $C_1(F) > 0$, and so $(C_1(F), C_2(F), C_3(F)) > (0, 0, 0)$.) Then, we compare the first (and hence largest) triple $(C_1(F),$
$C_2(F), C_3(F)$ of $C_F(H)$ with the first (and hence largest) triple $(C_1(F'), C_2(F'), C_3(F'))$ of $C_F(H')$. If $(C_1(F), C_2(F), C_3(F)) > (C_1(F'), C_2(F'), C_3(F'))$, say, then we define $C_F(H) > C_F(H')$. Otherwise, we pass to the second triples of $C_F(H)$ and $C_F(H')$. Continuing in this way, we can compare the $F$-complexities of $H$ and $H'$.

We now define the complexity $C(H)$ of a handle structure $H$ to be the ordered pair $(C_F(H), n(H))$, where $n(H)$ is the number of 0-handles of $H$ containing a component of $F(H)$ with positive index. We compare the complexity of handle structures $H$ and $H'$ by asserting that $C(H) > C(H')$ if one of the following holds:

- $C_F(H) > C_F(H')$, or
- $C_F(H) = C_F(H')$ and $n(H) < n(H')$.

**Lemma 5.3.** This ordering on complexity of handle structures is a well-ordering.

**Proof.** We need to show that there cannot exist an infinite strictly decreasing sequence \( \{C(H_i) : i \in \mathbb{N}\} \). Suppose that there is such a sequence. Then $C_F(H_i) \geq C_F(H_{i+1})$ for each $i$. Suppose first that this inequality is strict for only finitely many $i$. Then we may pass to a subsequence in which $C_F(H_i)$ is constant. Then the number of components of $F(H_i)$ with positive index
is constant. However, since \( C(\mathcal{H}_i) > C(\mathcal{H}_{i+1}) \) for each \( i \), \( n(\mathcal{H}_i) < n(\mathcal{H}_{i+1}) \) for each \( i \). This is impossible.

Therefore, we may suppose that \( C_F(\mathcal{H}_i) > C_F(\mathcal{H}_{i+1}) \) for infinitely many \( i \). Pass to this subsequence. Let \( T_i^n \) be the \( n \text{th} \) largest triple of \( C_F(\mathcal{H}_i) \). For each \( i \), there is a natural number \( N(i) \), such that:

- \( T_i^n = T_{i+1}^n \) for \( n < N(i) \), and
- \( T_i^{N(i)} > T_{i+1}^{N(i)} \).

Define \( M(i) = \min_{j \geq i} N(j) \). Then \( \{ M(i) : i \in \mathbb{N} \} \) is a non-decreasing sequence. For all \( i \), \( M(i) \leq N(i) \), and for infinitely many \( i \), this is an equality. Consider the sequence of triples \( \{ T_i^{M(i)} : i \in \mathbb{N} \} \). Then \( T_i^{M(i)} \geq T_{i+1}^{M(i)} \). For the infinitely many \( i \) when \( M(i) = N(i) \), we have

\[
T_i^{M(i)} = T_i^{N(i)} > T_{i+1}^{N(i)} = T_{i+1}^{M(i)} \geq T_{i+1}^{M(i+1)}.
\]

Thus, the infinite sequence of triples \( \{ T_i^{M(i)} : i \in \mathbb{N} \} \) contains an infinite strictly decreasing sequence. This is impossible, since the ordering on the triples is a well-ordering.

By the above lemma, we can use the complexity of handle structures as the basis for an inductive argument. We will start with a sutured manifold \((M, \gamma)\) with a handle structure \(\mathcal{H}\). If \(H_2(M, \partial M) \neq 0\), we will perform a taut decomposition \((M, \gamma) \xrightarrow{S} (M_S, \gamma_S)\). The manifold \(M_S\) will inherit a handle structure \(\mathcal{H}'\). We will try to ensure that the complexity of \(\mathcal{H}'\) is no more than that of \(\mathcal{H}\) (and preferably, strictly less than that of \(\mathcal{H}\)). The following lemma asserts that, to guarantee this, we need only restrict attention to smaller parts of \(\mathcal{H}\). For example, it shows that we need only check \(C(H_0 \cap \mathcal{H}') \leq C(H_0)\) for each 0-handle \(H_0\) of \(\mathcal{H}\).

**Lemma 5.4.** Let \(\mathcal{H}\) (respectively \(\mathcal{H}'\)) be a handle structure for a sutured manifold \((M, \gamma)\) (respectively \((M', \gamma')\)). Suppose that the 0-handles of \(\mathcal{H}\) (respectively \(\mathcal{H}'\)) have been partitioned into \(n\) subsets \(A_1, \ldots, A_n\) (respectively, \(A'_1, \ldots, A'_n\)). (For example, each \(A_i\) may be some 0-handle \(H_0\) of \(\mathcal{H}\), and \(A'_i\) is \(H_0 \cap \mathcal{H}'\)). Suppose that for each \(i\), \(C(A'_i) \leq C(A_i)\). Then \(C(\mathcal{H}') \leq C(\mathcal{H})\). Additionally, if \(C(A'_i) < C(A_i)\) for some \(i\), then \(C(\mathcal{H}') < C(\mathcal{H})\).

**Proof.** Arrange the triples of \(C_F(\mathcal{H})\) into a non-increasing sequence \(\{ T_j : j \in \mathbb{N} \}\). Consider the first integer \(j\) for which \(T_j > T_{j+1}\). Then the triples \(T_1, \ldots, T_j\) are all some fixed triple \(T\). The partitioning of \(\mathcal{H}'\) gives a partitioning of \(T_1, \ldots, T_j\) into \(n\) subsets (some of which may be empty). Say that \(k(i)\) of these lie in \(A_i\). Since \(C(A'_i) \leq C(A_i)\), we must have \(C_F(A'_i) \leq C_F(A_i)\). So, there are at most \(k(i)\) copies of \(T\) in \(A'_i\), and there are no larger triples. Hence, in \(C_F(\mathcal{H}')\), there are at most \(j\) copies of \(T\) and no larger triples. If there are fewer than \(j\) copies of \(T\) in \(C(\mathcal{H}')\), then \(C_F(\mathcal{H}') < C_F(\mathcal{H})\), and the lemma is proved. Otherwise, we can remove
each copy of \( T \) from \( C_F(\mathcal{H}) \) and \( C_F(\mathcal{H}') \), without affecting any ordering. Continuing in this fashion with the next largest triples of \( C(\mathcal{H}) \), and so on, we see that \( C_F(\mathcal{H}') \leq C_F(\mathcal{H}) \). Also, if we have equality, then we must have had \( C_F(A_i') = C_F(A_i) \) for each \( i \). Since \( C(A_i') \leq C(A_i) \), the number of 0-handles in \( A_i' \) containing components of \( F(\mathcal{H}') \) with positive index is at least the number of 0-handles in \( A_i \) containing components of \( F(\mathcal{H}) \) with positive index. Therefore, \( n(\mathcal{H}') \geq n(\mathcal{H}) \) and so \( C(\mathcal{H}') \leq C(\mathcal{H}) \). Also, if we have equality, then we must have had \( C(A_i') = C(A_i) \) for each \( i \).

To perform an inductive argument we need to ensure that the complexity of \( \mathcal{H}' \) is less than that of \( \mathcal{H} \), where \( \mathcal{H}' \) is the induced handle structure on \( (\mathcal{M}_S, \gamma_S) \). However, this is not in general true. To guarantee this, it is important that each 0-handle of \( F \) has positive index, and to ensure this, we may first need to decompose \( (\mathcal{M}, \gamma) \) along some product discs and annuli, and then simplify the handle decomposition of the resulting sutured manifold. Even then, to ensure that complexity is reduced by decomposition along \( S \), we may need to perform some modifications to \( S \).

We will give these procedures in Sections 7-10. But first we explain the idea behind the above definition of complexity. The surface \( S \) is in general \( \partial \)-compressible in \( \mathcal{M} \) and in [4] it was shown that there may exist infinitely long hierarchies of incompressible \( \partial \)-compressible surfaces in a 3-manifold. Thus, it is vital that we use the fact that \( \mathcal{M} \) has a sutured manifold structure. This is encoded in the quantity \( C_2(F) \) which was defined to be the index of a component \( F \) of \( \mathcal{F} \). We therefore study how index behaves under decomposition.

Let \( \mathcal{H} \) (respectively, \( \mathcal{H}' \)) be the handle decomposition of \((\mathcal{M}, \gamma)\) (respectively, \((\mathcal{M}_S, \gamma_S)\)). Let \( \mathcal{F} = \mathcal{F}(\mathcal{H}) \) and let \( \mathcal{F}' = \mathcal{F}(\mathcal{H}') \). Let \( V \) be a 0-handle of \( \mathcal{F} \) and let \( V_1', \ldots, V_k' \) be the 0-handles \( V \cap \mathcal{F}' \). Now, \( V_1', \ldots, V_k' \) are obtained from \( V \) by cutting along properly embedded arcs. The endpoint of each arc either lies in \( \mathcal{R}_\pm(\mathcal{M}) \) or in \( \mathcal{F}^1(\mathcal{H}) \). Therefore, an elementary counting argument shows that

\[
I(V) = \sum_{i=1}^{k} I(V_i').
\]

In particular, if \( F \) is a component of \( \mathcal{F} \) and \( F' = F \cap \mathcal{F}' \), then \( I(F) = I(F') \).

Hence, we can ensure that the quantity \( C_2 \) does not increase, as long as we create no discs of \( \mathcal{F}' \) with negative index. Thus, our goal is to alter \( S \) in order to remove these discs. But, in general, this does not seem to be possible. An example is given in Figure 5.2. There, a 0-handle of \( \mathcal{H} \) is decomposed into two 0-handles of \( \mathcal{H}' \). A negative index disc of \( \mathcal{F}' \) is created, but note that, nevertheless, the complexity of the handle structure has decreased.
It is fairly easy to show that, in general, under mild assumptions on $S$, neither $C_1$ nor $C_3$ can increase. Our aim is to show that, if $C_1$ is left unchanged, then in fact no negative index discs of $\mathcal{F}'$ are created, and so $C_2$ is not increased. Furthermore, if $C_1$, $C_2$ and $C_3$ are all left unchanged, then $[S, \partial S] = 0 \in H_2(M, \partial M)$.

6. Overview of the proof of the main theorems.

We have now developed enough machinery to outline the proofs of Theorems 1.4, 1.5 and 1.6. We start with a generalised triangulation of $M$, and using this, we construct the dual handle structure $\mathcal{H}(M)$ (which we sometimes abbreviate to $\mathcal{H}$). Roughly speaking, the idea is to decompose $M$ along surfaces until we end with a solid torus neighbourhood of $K$, plus perhaps some 3-balls. At each stage, we will be examining a 3-manifold $M'$ embedded in $M$. This manifold $M'$ will have a handle structure which respects $\mathcal{H}(M)$, in the following sense.

**Definition 6.1.** Let $(M, \gamma)$ be a sutured manifold with a handle structure $\mathcal{H}(M)$. Let $(M', \gamma')$ be a sutured manifold lying in $M$ with a handle structure $\mathcal{H}(M')$. Then $\mathcal{H}(M')$ respects $\mathcal{H}(M)$ if each of the following conditions holds:

- The 0-handles of $M'$ lie in the 0-handles of $M$.
- The 1-handles of $M'$ lie in the 1-handles of $M$ in a vertical fashion and inherit their product structure.
- The surface $\mathcal{F}(M')$ lies in $\mathcal{F}(M)$, with the intersection $\mathcal{F}^1(M') \cap \mathcal{F}^1(M)$ lying in $\mathcal{F}^1(M)$ in a vertical fashion.

Note that, if $\mathcal{H}(M')$ respects $\mathcal{H}(M)$, then automatically the arcs $\gamma' \cap \mathcal{H}^1(M')$ are vertical in $\mathcal{H}^1(M)$ and the discs $\mathcal{H}^2(M') \cap \mathcal{H}^1(M')$ are vertical in $\mathcal{H}^1(M)$. Thus, the only restriction on the 2-handles of $M'$ is a requirement on their attaching maps. The remainder of each 2-handle may lie inside $M$ in a complicated way.

Occasionally, the handle structure of $M'$ will resemble the handle structure of $M$ in some 0-handle, in the following sense.

**Definition 6.2.** Suppose that the handle structure $\mathcal{H}(M')$ of $(M', \gamma')$ respects the handle structure $\mathcal{H}(M)$ of $(M, \gamma)$. Let $H_0$ be a 0-handle of $M$. Then $\mathcal{H}(M') \cap H_0$ is obtained from $H_0$ by a trivial modification if each of the following conditions is satisfied:

(i) There is at most one 0-handle $H'_0$ of $\mathcal{H}(M') \cap H_0$ containing a component of $\mathcal{F}(M')$ with positive index;
(ii) $H_0 - H'_0$ consists of a parallelity region $R$ between $\partial H_0 - \partial H'_0$ and $\partial H'_0 - \partial H_0$;
(iii) for any component $F$ of $\mathcal{F}(M) \cap H_0$ with positive index, $F \cap R$ is either empty or a parallelity region between arcs and circles of $F \cap \partial H'_0$ and
arcs and circles in $\partial F$, the parallelity region respecting the handle structure of $F$;
(iv) the arcs $\partial H_0' \cap \gamma'$ are parallel in $R$ to the arcs of $\partial H_0 \cap \gamma$, possibly joined up by index zero discs of $F(M)$;
(v) $\partial H_0' \cap F(M)$ and $\partial H_0' \cap F(M')$ have the same handle structure.

Roughly speaking, a trivial modification leaves components of $F(M)$ with positive index relatively unaltered.

We make the following definition: If $\mathcal{H}(M)$ is a handle structure for $M$, we define the important $0$-handles $IH_0^0(M)$ to be the $0$-handles $H_0$ with $H_0 \cap F$ containing at least one component with positive index. In Sections 7 and 8, we will prove the following result, which gives a method of modifying a handle decomposition so that, afterwards, each $0$-handle of $F$ has positive index. Recall from Section 5 that the index of a component of $F$ is equal to the sum of the indices of its $0$-handles. So this implies that each component of $F$ has positive index. Therefore, each $0$-handle of $\mathcal{H}$ is either important or disjoint from the $1$-handles and $2$- handles.

**Proposition 6.3.** Let $\mathcal{H}(M)$ be a handle structure of a taut sutured manifold $(M, \gamma)$. Suppose that each component of $M$ has nonempty boundary, and that no component of $M$ is a solid torus. Suppose also that no component of $M$ is a Seifert fibre space disjoint from $\gamma$, with base space a disc and having two exceptional fibres. Then there is a (possibly empty) sequence of taut decompositions

$$(M, \gamma) \xrightarrow{P_1} \cdots \xrightarrow{P_m} (M', \gamma'),$$

where each $P_i$ is either a product disc or an incompressible annulus disjoint from the sutures. There is a handle structure $\mathcal{H}(M')$ of $(M', \gamma')$ and an embedding of $M'$ in $M$ isotopic to the embedding arising from the sutured manifold decomposition, with the following properties:

(i) $\mathcal{H}(M')$ respects $\mathcal{H}(M)$.

(ii) For each $0$-handle $H_0$ of $\mathcal{H}(M)$, the complexity of $H_0 \cap \mathcal{H}(M')$ is no more than that of $H_0$.

(iii) For each $0$-handle $H_0$ of $\mathcal{H}(M)$, the intersections

$$H_0 \cap IH_0^0(M')$$

$$H_0 \cap IH_0^0(M') \cap F(M')$$

$$H_0 \cap IH_0^0(M') \cap \gamma'$$

are each one of a finite number of possibilities (up to trivial modifications), which depend only on $F(M) \cap H_0$ and $H_0 \cap \gamma$, and are otherwise independent of $M$ and $M'$. 
(iv) If $H_0$ is a 0-handle of $\mathcal{H}(M)$ and the complexity of $H_0 \cap \mathcal{H}(M')$ is equal to that of $H_0$, then $H_0 \cap \mathcal{H}(M')$ is obtained from $H_0$ by a trivial modification.

(v) Each 0-handle of $\mathcal{F}(M')$ has positive index.

(vi) For each 0-handle $H_0'$ of $\mathcal{H}(M')$, $H_0' \cap (\mathcal{F}(M') \cup \gamma')$ is connected.

Once we have such a handle structure, we then perform a sutured manifold decomposition.

**Proposition 6.4.** Let $\mathcal{H}(M)$ be a handle structure of a taut sutured manifold $(M, \gamma)$. Suppose that each 0-handle of $\mathcal{F}(M)$ has positive index, and that, for each 0-handle $H_0$ of $\mathcal{H}(M), H_0 \cap (\mathcal{F}(M) \cup \gamma)$ is connected. Let $(M, \gamma) \xrightarrow{S} (M_S, \gamma_S)$ be a taut sutured manifold decomposition, where $\partial S$ has essential intersection with $\mathcal{R}_\pm(M)$ and $[S, \partial S] \neq 0 \in H_2(M, \partial M)$. Then there is a surface $S'$ properly embedded in $(M, \gamma)$ and a commutative diagram of sutured manifold decompositions and pull-backs

$$
(M, \gamma) \xrightarrow{S} (M_S, \gamma_S) = (\hat{M}_1, \hat{\gamma}_1) \xrightarrow{P_1} \ldots \xrightarrow{P_{r-1}} (\hat{M}_r, \hat{\gamma}_r)
$$

such that $S' 
\xrightarrow{S'} (\hat{M}_m, \hat{\gamma}_m) \xrightarrow{P_{m-1}} (\hat{M}_{m-1}, \hat{\gamma}_{m-1}) \xrightarrow{P_{m-2}} \ldots \xrightarrow{P_{r-1}} (\hat{M}_r, \hat{\gamma}_r)$.

Each $P_i$ is either a product disc, an incompressible product annulus or (for $i < r$) a surface parallel to a subsurface $F_i$ of $\mathcal{R}_\pm(M_{i+1})$, with the orientations of $P_i$ and $F_i$ disagreeing near $\partial P_i$. The induced handle structure $\mathcal{H}(M')$ on $(M', \gamma')$ satisfies properties (i), (ii), (iii) and (iv) of Proposition 6.3 and also the following:

(v) For some 0-handle $H_0$ of $\mathcal{H}(M), C(\mathcal{H}(M') \cap H_0) < C(H_0)$.

**Proof of Theorems 1.4, 1.5 and 1.6 using Propositions 6.3 and 6.4.** Let $\mathcal{H}$ be the dual handle structure for $M$, arising from the generalised triangulation of $M$. We give $M$ the trivial sutured manifold structure with $\mathcal{R}_- = \partial M$ and $\mathcal{R}_+ = \emptyset$. If this is not taut, then $M$ is either reducible or a solid torus. Hence, by Theorem 5.1 of [10], there are no exceptional or norm-exceptional surgery curves in $M$ satisfying the hypotheses of Theorem 1.4. Hence we may assume that $(M, \emptyset)$ is taut.

We will construct a sequence of taut sutured manifolds $(M_i, \gamma_i)$ where $1 \leq i \leq n$. The first sutured manifold $(M_1, \gamma_1)$ will be $(M, \emptyset)$. Each sutured manifold $(M_i, \gamma_i)$ will have a handle structure $\mathcal{H}_i$, and there will be an embedding of $M_i$ in $M_{i-1}$ having the following properties (some of which are only relevant for $i > 1$):

(i) $\mathcal{H}_i$ respects $\mathcal{H}_{i-1}$.

(ii) For each 0-handle $H_0$ of $\mathcal{H}_{i-1}$, the complexity of $H_0 \cap \mathcal{H}_i$ is no more than that of $H_0$. 


(iii) For each 0-handle $H_0$ of $\mathcal{H}_{i-1}$, the intersections
\[ H_0 \cap \mathcal{I}H_1^0, \]
\[ H_0 \cap \mathcal{I}H_1^0 \cap \mathcal{F}(\mathcal{H}_i), \]
\[ H_0 \cap \mathcal{I}H_1^0 \cap \gamma_i \]
are each one of a finite number of possibilities (up to trivial modification), which depend only on $\mathcal{F}(\mathcal{H}_{i-1}) \cap H_0$ and $H_0 \cap \gamma_{i-1}$.

(iv) For a 0-handle $H_0$ of $\mathcal{H}_{i-1}$, if the complexity of $H_0 \cap \mathcal{H}_i$ is equal to that of $H_0$, then $H_0 \cap \mathcal{H}_i$ is obtained from $H_0$ by a trivial modification.

(v) For some 0-handle $H_0$ of $\mathcal{H}_{i-1}$ ($i > 2$), we have $C(H_0 \cap \mathcal{H}_i) < C(H_0)$.

(vi) For $1 < i < n$, each 0-handle $H_0$ of $\mathcal{F}(\mathcal{H}_i)$ has positive index, and $H_0 \cap (\mathcal{F}(\mathcal{H}_i) \cup \gamma_i)$ is connected.

(vii) $K$ lies in $M_i$.

(viii) For $1 \leq i < n$, $H_2(M_i - \text{int}(\mathcal{N}(K)), \partial M_i) \neq 0$.

(ix) If $M_i(\sigma)$ is the manifold obtained from $M_i$ by Dehn surgery along $K$ with slope $\sigma$, then at least one of the following is true:
\[ \bullet \quad M_i = M, \]
\[ \bullet \quad (M_i(\sigma), \gamma_i) \text{ is not taut, or} \]
\[ \bullet \quad \text{the core of the surgery solid torus has finite order in } \pi_1(M_i(\sigma)). \]

The final manifold $M_n$ of the sequence is a solid torus neighbourhood of $K$, plus perhaps some 3-balls. The sequence is constructed using Propositions 6.3 and 6.4 in an alternating fashion.

We now show how to continue this sequence beyond $(M_1, \gamma_1)$. We would like to let $(M, \emptyset) = (M_2, \gamma_2)$, but (vi) above need not be satisfied in this case. Note, however, that each 0-handle of $\mathcal{F}(M)$ does indeed have positive index in either of the following cases:
\[ \bullet \quad M \text{ is closed (and so we have a genuine triangulation), or} \]
\[ \bullet \quad \partial M \neq \emptyset \text{ and we have an ideal triangulation.} \]

In the case where $K$ and $\sigma$ are norm-exceptional, we would like to ensure that one of the above is true. In Theorem 1.6, we explicitly make this assumption. In Theorem 1.4, we alter the given generalised triangulation of $M$ so that it is either genuine or ideal. This can be done algorithmically. Hence, in the case where $K$ and $\sigma$ are norm-exceptional, we let $(M, \emptyset) = (M_2, \gamma_2)$.

Suppose now that $K$ and $\sigma$ are exceptional and that some 0-handle of $\mathcal{F}$ has nonpositive index. Then we use Proposition 6.3 to decompose $(M, \emptyset)$ along product discs and incompressible annuli disjoint from the sutures, resulting in a taut sutured manifold $(M', \gamma')$ satisfying (i)-(vi) of 6.3. Since $\Delta(\sigma, \mu) > 1$, Theorem 3.2 gives that we may ambient isotope $K$ off each decomposing surface, and hence $K$ lies in $M'$.

To apply Proposition 6.3, we need to check that $M$ is not a Seifert fibre space, with base space a disc, having two exceptional fibres, and having $\gamma = \emptyset$. We will suppose it is, and then achieve a contradiction. Let $\alpha$ be
an arc properly embedded in the base space, separating the two exceptional points. Then \( \alpha \) lifts to an annulus \( A \) in \( M \). Using Theorem 3.2, we may ambient isotope \( K \) off \( A \). Let \( \hat{M} \) be the solid torus which is the closure of the component of \( M - A \) containing \( K \). Then, \( \partial \hat{M} \) is an incompressible torus in \( M - \text{int}(\mathcal{N}(K)) \). Since \( M - \text{int}(\mathcal{N}(K)) \) is atoroidal, we deduce that \( \partial \hat{M} \) must be parallel to \( \partial \mathcal{N}(K) \) and so \( K \) is isotopic to the exceptional fibre lying in \( \hat{M} \). But then \( H_2(M - \text{int}(\mathcal{N}(K)), \partial M) \) is trivial, contrary to assumption.

Hence, we may apply Proposition 6.3, and then we let \( (M, \partial M) \) be a sequence as far as \( \pi \) or the core of the surgery solid torus has finite order in \( \pi_1(M) \). Since \( (M, \partial M) \) is taut, we insist that \( \pi_1(M) \) is norm-exceptional and \( M \) are norm-exceptional and \( \sigma \) are exceptional. Hence, either \( (M, \partial M) \) is not taut or the core of the surgery solid torus has finite order in \( \pi_1(M) \). Suppose that \( (M, \partial M) \) is not taut.

We now verify (ix). We only applied Proposition 6.3 in the case where \( K \) and \( \sigma \) are exceptional. Hence, either \( (M, \partial M) \) is not taut or the core of the surgery solid torus has finite order in \( \pi_1(M) \). Suppose that \( (M, \partial M) \) is taut. Then, by 3.1, the sequence of decompositions

\[
(M, \partial M) \xrightarrow{P_1} \cdots \xrightarrow{P_m} (M', \partial M),
\]

is taut. In particular, \( (M, \partial M) \) is taut, and therefore (iii) of 1.3 holds. Also, since each \( P_i \) has essential intersection with \( \mathcal{R} \), it is therefore incompressible. Therefore, the map \( \pi_1(M, \partial M) \rightarrow \pi_1(M) \) induced by inclusion is an injection. Therefore, the core of the surgery solid torus has finite order in \( \pi_1(M) \).

Thus, we have now constructed \( (M, \partial M) \) and have verified that it has the correct properties. Suppose that we have constructed a sequence as far as \( (M_i, \gamma_i) \), satisfying (i)-(ix) above. If \( H_2(M_i - \text{int}(\mathcal{N}(K)), \partial M_i) \) is trivial, then we stop. If this homology group is nontrivial, then (see Section 3 or Theorem 2.6 of [10]) we may find a taut decomposition

\[
(M_i - \text{int}(\mathcal{N}(K)), \gamma_i) \xrightarrow{S} (M'_i - \text{int}(\mathcal{N}(K)), \gamma'_i),
\]

such that

- \( S \) is disjoint from \( \partial \mathcal{N}(K) \),
- no curve of \( \partial S \) bounds a disc in \( \mathcal{R}_\pm(M_i) \),
- no component \( X \) of \( M'_i \) has \( \partial X \subset \mathcal{R}_-(M'_i) \) or \( \partial X \subset \mathcal{R}_+(M'_i) \), and
- \([S, \partial S] \neq 0 \in H_2(M_i - \text{int}(\mathcal{N}(K)), \partial M_i) \).

This implies that \([S, \partial S] \neq 0 \in H_2(M_i, \partial M_i) \). In the case where \( K \) and \( \sigma \) are norm-exceptional and \( M_i = M \), we insist that \([S, \partial S] = z \in H_2(M - \text{int}(\mathcal{N}(K)), \partial M) \), where \( z \) is the homology class in Definition 1.3.

Let \( M'_i(\sigma) \) be the result of \( M'_i \) after Dehn surgery along \( K \) with slope \( \sigma \). Since \( (M_i(\sigma), \gamma_i) \) is not taut or the core of the surgery solid torus has finite order in \( \pi_1(M_i(\sigma)) \), the argument above gives that \( (M'_i(\sigma), \gamma_i) \) is not taut or the core of the surgery solid torus has finite order in \( \pi_1(M'_i(\sigma)) \). Using the argument in Section 3 (see also Theorem 1.8 of [1]), we deduce that the decomposition \( (M_i, \gamma_i) \xrightarrow{S} (M'_i, \gamma'_i) \) is taut. Since each 0-handle of \( \mathcal{F}(\mathcal{H}_i) \)
has positive index, we may apply Proposition 6.4 to \((M_i, \gamma_i)\), and so end with a sutured manifold \((M''_i, \gamma''_i)\) with a handle-decomposition \(H''_i\), satisfying (i)-(iv) of 6.3 and (v) of 6.4. Again using 3.2, we may isotope \(K\) off each product disc and incompressible product annulus, and so we may assume \(K\) lies in \(M''_i\). Also, using the commutative diagram in Proposition 6.4, we deduce that \((M''_i(\sigma), \gamma''_i)\) is not taut or the core of the surgery solid torus has finite order in \(\pi_1(M''_i(\sigma))\).

If any component of \(M''_i\) is a solid torus disjoint from \(K\), we decompose it along a meridian disc. If the component \(X\) of \(M''_i\) containing \(K\) is a solid torus, then the atoroidality of \(M - \text{int}(\mathcal{N}(K))\) implies that \(K\) is the core of \(X\). In this case, we set \((M_n, \gamma_n) = (X, \gamma''_i \cap X)\), together with some 3-balls obtained by decomposing \(M''_i - X\).

We may therefore assume that no component of \(M''_i\) is a solid torus, and so we can apply 6.3 to \((M''_i, \gamma''_i)\) to obtain a sutured manifold \((M_{i+1}, \gamma_{i+1})\) satisfying (i)-(ix) above. Note that each component of \(\mathcal{F}(H_{i+1})\) has positive index, and therefore, the only 0-handles of \(H_{i+1}\) not lying in \(IH^0_{i+1}\) are handles disjoint from \(\mathcal{F}(H_{i+1})\).

By (ii), (v) and Lemma 5.4, the complexity of \(H_{i+1}\) is strictly less than that of \(H_i\). Hence, eventually, the sequence terminates with a sutured manifold \((M_n, \gamma_n)\) such that \(H_2(M_n - \text{int}(\mathcal{N}(K)), \partial M_n) = 0\). Then \(M_n\) is a solid torus neighbourhood of \(K\), plus perhaps some 3-balls. By (ix), the sutures \(\gamma_n \cap \mathcal{N}(K)\) are parallel to \(\sigma\).

Note that, for each 0-handle \(H_0\) of \(\mathcal{H}(M)\), there are only finitely many possibilities for \(\mathcal{F}(M) \cap H_0\). Thus, by induction on complexity using (ii), (iii) and (iv) above, there is in each 0-handle \(H_0\) of \(\mathcal{H}(M)\), only a finite number of possibilities for

\[
\begin{align*}
H_0 \cap IH^0_n \\
H_0 \cap IH^0_n \cap \mathcal{F}(H_n) \\
H_0 \cap IH^0_n \cap \gamma_n.
\end{align*}
\]

Each possibility for \(H_0 \cap IH^0_n \cap \gamma_n\) gives a tangle in the associated 3-simplex of \(M\). These tangles join to form \(\gamma_n\) (with possibly some unknotted curves removed). Some component of \(\gamma_n\) is the \(\sigma\)-cable of \(K\), and hence the tangles required for Theorems 1.5 and 1.6 are constructed by taking all possible subtangles of \(H_0 \cap IH^0_n \cap \gamma_n\).

Each possibility for \(H_0 \cap IH^0_n\) and \(H_0 \cap IH^0_n \cap \mathcal{F}(H_n)\) gives a graph \(G\) in the associated 3-simplex of \(M\). When the collection of these graphs (one in each 3-simplex of \(M\)) are joined, they form the 0-handles and 1-handles of \(M_n\). The 2-handles of \(M_n\) are attached along the annuli \((\mathcal{H}^0_n \cup \mathcal{H}^1_n) \cap \mathcal{H}^2_n\), which are determined by \(\mathcal{H}^0_n \cap \mathcal{F}^1(H_n)\). Thus, we readily see that the algorithm given in Section 2 constructs all possibilities for \(K\) and \(\sigma\). Hence, Theorem 1.4 is established.
We end this section with the following:

Proof of Theorem 1.1.

This is an almost immediate corollary of Theorem 1.4, but there is one complication. If we set
\[ M = S^3 - \text{int}(\mathcal{N}(L)) \],
then it is not obvious that \( M - \text{int}(\mathcal{N}(K)) \) is atoroidal. To establish this, we will use a modified form of the argument in Proposition 2.3 of [6].

Suppose therefore that \( T \) is an incompressible torus in \( M - \text{int}(\mathcal{N}(K)) \) which is not parallel to \( \partial \mathcal{N}(K) \) or \( \partial \mathcal{N}(L) \). Since we are assuming that \( L \) is not a nontrivial satellite knot, then \( T \) must be compressible in \( M \) or be parallel to \( \partial \mathcal{N}(L) \). In the latter case, \( K \) lies in the collar between \( T \) and \( \partial \mathcal{N}(L) \), and then it is easy to see that \( L' \) is a winding number one satellite of \( L \). In particular, \( \text{genus}(L') \geq \text{genus}(L) \), which is contrary to hypothesis. Therefore, \( T \) must be compressible in \( M \). There are now two cases.

Case 1. \( T \) lies in a 3-ball in \( M \).

Then, \( T \) separates \( S^3 \) into a nontrivial knot exterior \( X \) disjoint from \( K \) and \( L \), and a solid torus containing \( K \) and \( L \). Let \( Y \) be the manifold obtained from \( M \) by removing the interior of \( X \). Let \( Y' \) be the manifold obtained from \( S^3 - \text{int}(\mathcal{N}(L')) \) by removing the corresponding knot exterior, which we also call \( X \). Note that \( H_2(Y - \text{int}(\mathcal{N}(K)), \partial Y) \) is nontrivial. Also, \( \partial Y \) has compressible boundary. So, by Theorem 5.1 of [10] (see also the argument in Section 3) the minimal genus of a Seifert surface for \( L' \) in the complement of \( X \) is at least the genus of \( L \) in the complement of \( X \), which is the genus of \( L \). Since \( X \) is a nontrivial knot exterior, the minimal genus of a Seifert surface for \( L' \) in the complement of \( X \) is just the genus of \( L' \). So, in this case, the genus of \( L' \) is at least the genus of \( L \), which is a contradiction.

Case 2. \( T \) bounds a solid torus \( V \) in \( M \) which contains \( K \).

Let \( V' \) be the manifold obtained from \( V \) by \( 1/q \) Dehn surgery along \( K \).

Case 2A. \( K \) has winding number zero in \( V \).

Then consider a minimal genus Seifert surface \( S \) for the knot \( L' \). We may assume that it intersects \( \partial V' \) in a collection of simple closed curves, which inherit an orientation from \( S \). The union of these curves is homologically trivial in \( V' \). Hence, by making annular modifications to \( S \), if necessary, which do not increase its genus, we may assume that each curve of \( S \cap \partial V' \) is homologically trivial in \( S' \). Since \( K \) has winding number zero in \( V \), these curves are also homologically trivial in \( V \). Hence, we may fill them in with meridian discs in \( V \). This gives a Seifert surface for \( L \) with genus at most that of \( S \), which is a contradiction.

Case 2B. \( K \) has nonzero winding number in \( V \).

Since \( K \) and \( L \) have zero linking number, so do \( L \) and the core of \( V \). Therefore, there exists a Seifert surface \( S' \) for \( L' \) which is disjoint from \( V' \).
If \( \partial V' \) is incompressible in \( V' \), then by Lemma A.16 of [7], there is a minimal genus Seifert surface for \( L' \) which is disjoint from \( V' \). This gives a Seifert surface for \( L \), and again we reach the contradiction that \( \text{genus}(L) \leq \text{genus}(L') \).

So, we may assume that \( \partial V' \) is compressible in \( V' \). This implies that \( V' \) is a solid torus, since it cannot be reducible, as it lies in \( S^3 \). Now, \( 1/q \) surgery along a knot \( K \) in the solid torus \( V \) never yields another solid torus if \( |q| > 1 \), unless \( K \) is a core of \( V \) or \( K \) lies in a 3-ball in \( V \) [9]. If \( K \) lies in a 3-ball in \( V \), then \( 1/q \) surgery along \( K \) does not alter \( L \), which is a contradiction. If \( K \) is a core of \( V \), then \( T \) is parallel to \( \partial N(K) \), contradicting the assumption that it is essential.

This proves then that \( M - \text{int}(N(K)) \) is atoroidal. Theorem 1.1 now follows directly from Theorem 1.4. \( \square \)

7. Simplifying handle structures.

In the next two sections, we will give a proof of Proposition 6.3. In particular, we will assume that each component of \( M \) has nonempty boundary, and that no component of \( M \) is a solid torus or a Seifert fibre space as in 6.3. We start by giving various elementary procedures for simplifying a handle structure \( \mathcal{H} \) of the taut sutured manifold \((M, \gamma)\). Our aim is to end with a handle structure in which each 0-handle of \( F \) has positive index. Each procedure will satisfy (i)-(iv) of 6.3. It will not increase the complexity of the handle structure, but it need not decrease it. To ensure that these procedures eventually terminate, we therefore introduce the following definition:

**Definition 7.1.** Let \( \mathcal{H} \) be a handle structure for a sutured manifold \((M, \gamma)\). Define the extended \( F \)-complexity for \( \mathcal{H} \) to be the set of triples

\[ C^+_F(\mathcal{H}) = \{(C_1(F), C_2(F), C_3(F)) : F \text{ a component of } \mathcal{F}\}, \]

where repetitions are retained. Here, \( C_1, C_2 \) and \( C_3 \) are the integers defined in Section 5. We also define the extended complexity \( C^+(\mathcal{H}) \) to be the ordered pair \((C^+_F(\mathcal{H}), n(\mathcal{H}))\).

The difference between the extended \( F \)-complexity and the \( F \)-complexity of a handle structure is that extended \( F \)-complexity also takes into account components of \( \mathcal{F} \) with nonpositive index.

We order the extended \( F \)-complexities and extended complexities as we do the \( F \)-complexities and complexities of handle structures (see Section 5). As in Lemma 5.3, this is a well-ordering. The procedures we give in the next two sections will all reduce the extended complexity of the handle structure, and so are guaranteed to terminate.

The following lemma will be useful in our verification that (i)-(iv) of 6.3 holds and that extended complexity is reduced:
Lemma 7.2. Let $(M, \gamma)$ be a sutured manifold with a handle structure $H$. Let $(M', \gamma')$ be embedded in $M$, with a handle structure $H'$ which respects $H$. Suppose that, for each component $F$ of $H$, either $C^+_F(F \cap F') < C^+_F(F)$ or $F \subset F'$. Suppose also that the former of the above two possibilities holds for at least one component $F$ of $H$. Suppose also that, if $I(F) \leq 0$, then the index of each component of $F \cap F'$ is nonpositive. Then, $C^+_F(H') < C^+_F(H)$, and (i), (ii) and (iv) of 6.3 are verified.

Proof. A version of the argument in Lemma 5.4 gives that $C^+_F(H') < C^+_F(H)$. We now check (ii) and (iv) of 6.3. (Note that (i) of 6.3 is part of the hypothesis of the lemma.) For each component $F$ of $F$, we have one of the following possibilities:

(i) $F \subset F'$ and so $F \cap F'$ is a copy of $F$, or
(ii) $I(F) > 0$ and $C_F(F \cap F') \leq C^+_F(F \cap F') < C^+_F(F) = C_F(F)$, or
(iii) $I(F) \leq 0$ and $C_F(F \cap F') = C_F(F)$.

In (iii), we are using that if $I(F) \leq 0$, then the index of each component of $F \cap F'$ is nonpositive, and so does not contribute to $F$-complexity. Therefore, for any 0-handle $H_0$ of $H$, $C_F(H' \cap H_0) \leq C_F(H_0)$. Also, if we have equality, then (ii) above cannot occur for any component $F$ of $F \cap H_0$, which implies that components $F$ of $F \cap H_0$ with positive index remain unchanged and hence that $n(H' \cap H_0) \geq n(H_0)$. So, $C(H' \cap H_0) \leq C(H_0)$. This verifies (ii) of 6.3. Also, if $C(H' \cap H_0) = C(H_0)$, then $H' \cap H_0$ is obtained from $H_0$ by a trivial modification, verifying (iv) of 6.3.

Before we describe the procedures in detail, we mention that many of them simply remove some handles of $H$. The following lemma will therefore be useful:

Lemma 7.3. Let $H'$ be a collection of handles of $H$ forming a 3-manifold $M'$ embedded in $M$, with $H - H'$ containing at least one $i$-handle for some $i \leq 2$. Suppose that $H'$ is a handle structure, that each handle of $H - H'$ is disjoint from $\gamma$ and that $(M', \gamma)$ is a sutured manifold structure. Then, (i)-(iv) of 6.3 are satisfied, and extended $F$-complexity is reduced.

Proof. It follows straight from the definition that $H'$ respects $H$. Let us now check that the hypotheses of 7.2 hold. Let $F$ be some component of $F$ and let $F' = F \cap F'$. The 1-handles of $F$ are either removed or divided up amongst the components of $F'$. In particular, each component $X$ of $F'$ has $C_1(X) \leq C_1(F)$. If this inequality is an equality for some $X$, then in fact $F' = F$. Hence, either $C^+_F(F') < C^+_F(F)$ or $F \subset F'$. Also, the former case holds for some component $F$ of $F$.

We now check that if $I(F) \leq 0$, then each component of $F'$ has nonpositive index. But $F'$ is obtained from $F$ by removing some 1-handles (or equivalently, cutting $F$ along properly embedded arcs), then removing some 0-handles disjoint from $\gamma$. Thus, each component $X$ of $F'$ has $\chi(X) \geq \chi(F)$.\[\boxdot\]
and $\gamma \cap X \leq \gamma \cap F$. Hence, $X$ does not have positive index if $F$ does not have positive index.

Thus, by Lemma 7.2, (i), (ii) and (iv) of 6.3 hold and extended $\mathcal{F}$-complexity is reduced. Also, (iii) of 6.3 is obvious. $\square$

**Procedure 1.** Slicing a 0-handle along a disc.

Suppose that there is a disc $D$ properly embedded in some 0-handle $H_0$ with $D \cap \gamma = D \cap \mathcal{F} = \emptyset$, and which separates $\mathcal{F} \cap H_0$. Then, $\partial D$ either lies in $\partial \mathcal{H}^3$ or in $\mathcal{R}_\pm$. In the former case, $\partial D$ bounds a disc $D'$ in $\partial \mathcal{H}^3$, since each component of $\partial \mathcal{H}^3$ is a sphere. In the case where $\partial D$ lies in $\mathcal{R}_\pm$, the incompressibility of $\mathcal{R}_\pm$ implies that $\partial D$ bounds a disc $D'$ in $\mathcal{R}_\pm$. The irreducibility of $M$ implies that $D \cup D'$ bounds a ball $B$ in $M$. Procedure 1 is the removal of all handles of $H$ which intersect int($B$), other than $H_0$. If $D' \subset \partial \mathcal{H}^3$, we extend $\mathcal{H}^3$ over $B$. By Lemma 7.3, (i)-(iv) of 6.3 hold, and extended $\mathcal{F}$-complexity is reduced. Thus, using this procedure, we eventually obtain a handle structure $\mathcal{H}(M')$ on the resulting sutured manifold $(M', \gamma')$, with $H'_0 \cap (\mathcal{F}(M') \cup \gamma')$ connected, for each 0-handle $H'_0$ of $\mathcal{H}(M')$. This is (vi) of 6.3.

**Procedure 2.** Collapsing a 2-handle and a 1-handle disjoint from $\gamma$.

Suppose now that $H_1$ is a 1-handle of $M$ which is disjoint from $\gamma$ and which intersects $\mathcal{H}^2$ in a single disc. Then this disc is contained in a single 2-handle $H_2$. We may remove $H_1$ and $H_2$, without changing the homeomorphism type of $(M, \gamma)$. Lemma 7.3 gives that (i)-(iv) of 6.3 are satisfied, and that extended $\mathcal{F}$-complexity is reduced.

**Procedure 3.** Collapsing a 2-handle and a 1-handle containing an arc of $\gamma$.

Let $H_1$ be a 1-handle of $M$ which intersects $\gamma$ in a single arc, and which intersects $\mathcal{H}^2$ in a single disc, lying in a 2-handle $H_2$. Procedure 3 is the collapsing of $H_1$ and $H_2$. This moves $\gamma \cap H_1$ onto an arc running along $\partial(\mathcal{H}^0 \cup \mathcal{H}^3 - H_1)$. Let $(M', \gamma')$ be the new sutured manifold, with handle structure $\mathcal{H}'$.

This procedure has the following effect on $\mathcal{F}$: Removing $H_2 \cap \partial \mathcal{H}^0$ (which is a collection of 1-handles of $\mathcal{F}$), removing $H_1 \cap \partial \mathcal{H}^0$ (which is precisely two 0-handles of $\mathcal{F}$) and then replacing each handle of $\mathcal{F}$ which we have removed with a sub-arc of $\gamma$. Thus, if $F$ is a component of $\mathcal{F}$, and $F' = F \cap \mathcal{F}'$, then each component of $X$ of $F'$ has $C_1(X) \leq C_1(F)$, and if we have equality for some component $X$, then in fact $F$ is unchanged by the procedure. This verifies one of the hypotheses of Lemma 7.2.

We now check that if $F$ has nonpositive index, then each component $X$ of $F'$ has nonpositive index. Suppose therefore $F$ has nonpositive index and that $F$ is changed by the procedure. It is simple to show that $I(F)$ is the
sum of the indices of $X$, as $X$ ranges over all components of $F'$. Therefore, the only way that a component $X$ of $F'$ can have positive index is if another component of $F'$ has negative index. However, since $F$ is changed by the procedure, then each component of $F'$ touches $\gamma'$, and hence $I(X) \geq 0$ for all components $X$ of $F'$. Lemma 7.2 now gives us that extended $F$-complexity decreases and that (i), (ii) and (iv) of 6.3 hold. It is straightforward to verify (iii) of 6.3.

**Procedure 4.** Decomposing along a product disc, then sliding $\gamma$.

Suppose that $H_1$ is a 1-handle of $\mathcal{H}$ which is disjoint from $\mathcal{H}^2$ and which has $|H_1 \cap \gamma| = 2$. Let $D$ be one of the two discs of $H_1 \cap \mathcal{H}^0$, lying in some 0-handle $H_0$. Push $D$ a little into $H_0$. Then $D$ is a product disc, which we decompose along. This decomposition creates a new handle decomposition which respects $\mathcal{H}$, and leaves both the complexity and extended complexity of $\mathcal{H}$ unchanged. But now the two arcs of $H_1 \cap \gamma$ are joined by an arc of $\mathcal{H}^0 \cap \gamma$. We may therefore perform an ambient isotopy which slides $\gamma$ off $H_1$. Then, using Procedure 1, we may remove $H_1$. Again, (i)-(iv) of 6.3 are satisfied and extended $F$-complexity is reduced.

**Procedure 5.** Collapsing a 3-ball disjoint from $\gamma$.

Suppose that a component of $M$ is a 3-ball disjoint from $\gamma$, comprised of two 0-handles joined by a 1-handle. Then we may collapse the 1-handle and one of the 0-handles. This reduces the extended $F$-complexity and (i)-(iv) of 6.3 are satisfied.

**Procedure 6.** Collapsing a 2-handle and a 3-handle.

Let $H_2 = D^1 \times D^2$ be a 2-handle, with one component of $\partial D^1 \times D^2$ in $\partial M$, and the other component touching a 3-handle $H_3$. Then we may remove $H_2$ and $H_3$ without changing the homeomorphism type of $M$. By Lemma 7.3, this procedure reduces extended $F$-complexity and satisfies (i)-(iv) of 6.3. Note that we are assuming in 6.3 that each component of $M$ has nonempty boundary. Hence, if $\mathcal{H}^3$ is nonempty, we may always apply this procedure somewhere. In this way, we remove all 3-handles from $M$.

The above six procedures are not enough to ensure that each 0-handle of $F$ has positive index. To deal with components of $F$ which are annuli disjoint from $\gamma$, we must clump collections of handles into groups, known as amalgams, which are defined as follows:

**Definition 7.4.** An **amalgam** $\mathcal{A}$ is a connected collection of handles with the following properties:

(i) $\mathcal{A}$ is disjoint from $\gamma$,
(ii) $\mathcal{A}$ is an $I$-bundle over a connected surface $G$,
(iii) the $(I - \partial I)$-bundle over $\partial G$ is disjoint from $\partial M \cup \partial \mathcal{H}^3$. 
(iv) the handles of $\mathcal{A}$ touching the $(I - \partial I)$-bundle over $\partial G$ are 1-handles and 2-handles,
(v) no 2-handle or 3-handle of $\mathcal{H} - \mathcal{A}$ touches $\mathcal{A}$,
(vi) the $\partial I$-bundle over $G$ lies in $\mathcal{R}_\pm$, and
(vii) $\text{cl}(\mathcal{H} - \mathcal{A})$ inherits a handle structure from $\mathcal{H}$.

An amalgam is \textit{trivial} if it is a single 2-handle; otherwise it is \textit{nontrivial}.

An amalgam $\mathcal{A}$ behaves in many ways just like a 2-handle. For example, it is attached onto the 0-handles and 1-handles of $\mathcal{H} - \mathcal{A}$ in a fashion that is very similar to the attachment of a 2-handle.

The main example of a nontrivial amalgam $\mathcal{A}$ is a connected collection of 2-handles and 1-handles disjoint from $\gamma$ and $\mathcal{H}^3$, such that each 1-handle of $\mathcal{A}$ intersects $\mathcal{H}^2$ in precisely two discs, and these discs lie in 2-handles of $\mathcal{A}$. For then the co-core $D^2$ of each 1-handle $H_1 = D^2 \times D^1$ in $\mathcal{A}$ has a product structure as $I \times I$, in which $H_1 \cap \mathcal{H}^2 = \partial I \times I \times D^1$. The product structures on the 1-handles combine with the product structures on the 2-handles to form an $I$-bundle structure on $\mathcal{A}$ with the required properties. An example is given in Figure 7.5.

![Figure 7.5.](image)

In Section 8, we will show how to remove all nontrivial amalgams. This, together with Procedures 1-6 is enough to ensure that each 0-handle of $\mathcal{F}$ has positive index.

\textbf{Lemma 7.6.} Let $\mathcal{H}$ be a handle-decomposition of a connected sutured manifold with nonempty boundary, containing no nontrivial amalgams. If some 0-handle of $\mathcal{F}$ has nonpositive index, then we may apply one of Procedures 1-6.
Proof. Let \( V \) be a 0-handle of \( \mathcal{F} \) with nonpositive index. Then, there are a number of cases.

1. \(|\mathcal{H}^2 \cap V| = 0\) and \(|\gamma \cap V| = 0\).
2. \(|\mathcal{H}^2 \cap V| = 1\) and \(|\gamma \cap V| = 0\).
3. \(|\mathcal{H}^2 \cap V| = 1\) and \(|\gamma \cap V| = 1\).
4. \(|\mathcal{H}^2 \cap V| = 0\) and \(|\gamma \cap V| = 2\).
5. \(|\mathcal{H}^2 \cap V| = 0\) and \(|\gamma \cap V| = 1\).
6. \(|\mathcal{H}^2 \cap V| = 2\) and \(|\gamma \cap V| = 0\).

For \( i = 2, 3 \) and 4, we may apply Procedure \( i \). In Case 1, let \( H_1 \) be the 1-handle containing \( V \), and let \( V' = \partial H_1 \cap \partial \mathcal{H}^0 - V \). If we cannot apply Procedure 1 to either of the discs \( V \) or \( V' \), then this component of \( M \) is a 3-ball disjoint from \( \gamma \), comprised of two 0-handles joined by a single 1-handle. We may therefore apply Procedure 5. Case 5 cannot arise since \( \gamma \) separates \( \partial M \) into \( \mathcal{R}_- \) and \( \mathcal{R}_+ \). By applying Procedure 6 if necessary, we may assume that \( M \) contains no 3-handles. In Case 6, the 1-handle of \( \mathcal{H} \) containing \( V \) is part of a nontrivial amalgam, contrary to assumption. \( \Box \)

8. Removing nontrivial amalgams.

We now give a procedure for removing all nontrivial amalgams, which will complete the proof of Proposition 6.3. Suppose that there is a nontrivial amalgam \( A \) in the handle structure \( \mathcal{H} \) of the taut sutured manifold \( (M, \gamma) \). We will assume that \( A \) is maximal, in the sense that if any other handles are added to \( A \), the resulting collection of handles does not form an amalgam. We will also assume that we cannot apply any of Procedures 1-6 in Section 7. In particular, due to Procedure 6, this implies that \( M \) has no 3-handles.

Recall that \( A \) has the structure of an \( I \)-bundle over a connected surface \( G \). The \( I \)-bundle over \( \partial G \) will be denoted by \( \partial_v A \).

Note that (iii) of 7.4 implies that \( \partial_v A \) is a union of intersections between handles of \( \mathcal{H} \). By (v) of 7.4, only 0-handles and 1-handles of \( \mathcal{H} - A \) touch \( \partial_v A \). By (iv) of 7.4, only 1-handles and 2-handles of \( A \) touch \( \partial_v A \). Hence, we may define a handle structure on \( \partial_v A \) as follows: The 0-handles of \( \partial_v A \) arise from the intersection of 1-handles of \( A \) with the 0-handles of \( \mathcal{H} - A \). The 1-handles of \( \partial_v A \) arise from the intersection of 2-handles of \( A \) with 0-handles and 1-handles of \( \mathcal{H} - A \).

**Lemma 8.1.** Suppose that we cannot apply Procedure 2 of Section 7. Then each 0-handle of \( \partial_v A \) abuts precisely two 1-handles of \( \partial_v A \).

**Proof.** If not, then some 0-handle of \( \partial_v A \) abuts precisely one 1-handle of \( \partial_v A \), since \( \partial_v A \) is a collection of annuli. This 0-handle \( D \) is a component of \( \mathcal{H}^1 \cap \mathcal{H}^0 \). Let \( H_0 \) (respectively, \( H_1 \)) be the 0-handle (respectively, the 1-handle) of \( \mathcal{H} \) containing \( D \). Then \( H_1 \) lies in \( A \), but \( H_0 \) does not. Since \( D \) abuts precisely one 1-handle of \( \partial_v A \), \( H_1 \) intersects \( A \cap \mathcal{H}^2 \) in a single disc,
lying in some 2-handle $H_2$. This in fact must the only intersection between $H_1$ and $H^2$, by (v) of 7.4. Hence, we may apply Procedure 2 of Section 7 to $H_1$ and $H_2$, which is a contradiction. □

The following lemma will also be useful:

**Lemma 8.2.** Let $\mathcal{H}$ be a handle structure for $(M, \gamma)$ to which we cannot apply any of Procedures 1-6. Let $A$ be a maximal amalgam in $\mathcal{H}$. Let $F$ be a component of $\mathcal{F}$ touching $\partial_vA$. Then $F$ has positive index.

**Proof.** Since we cannot apply any of Procedures 1-6, the only 0-handles of $F$ with nonpositive index have valence two and are disjoint from $\gamma$ (see the proof of Lemma 7.6). If $F$ has nonpositive index then each 0-handle of $F$ must be of this form. Hence, $F$ is an annulus disjoint from $\gamma$. Therefore, $F$ must be the only component of $\mathcal{F}$ lying in $H_0$, where $H_0$ is the 0-handle of $\mathcal{H}$ containing $F$. For, otherwise we could apply Procedure 1.

Consider a handle of $F$ lying in $\partial_vA$. This is a component of intersection between $H_0$ and some 1-handle or 2-handle of $\mathcal{H}$. By (iv) of 7.4, we must have $H_0 \not\in A$.

If $F$ lies entirely in $\partial_vA$, then each handle of $\mathcal{H}$ touching $H_0$ must be in $A$, and so we may extend $A$ over $H_0$, contradicting its maximality. Therefore, $F \cap \partial_vA$ is not the whole of $F$.

If $V$ is a 0-handle of $F$ lying in $\partial_vA$, then the 1-handle of $\mathcal{H}$ touching $V$ must lie in $A$. Hence, by (v) of 7.4, the 1-handles of $F$ touching $V$ also lie in $\partial_vA$. Hence, we may find a 0-handle $F_0$ of $F$ and a 1-handle $F_1$ of $F$ which are adjacent, with $F_1$ in $\partial_vA$, but $F_0$ not in $\partial_vA$. Let $H_1$ (respectively, $H_2$) be the 1-handle (respectively, 2-handle) of $\mathcal{H}$ containing $F_0$ (respectively, $F_1$). Then, we must have $H_0 \not\in A$, $H_1 \not\in A$ and $H_2 \in A$. Let $H'_2$ be the 2-handle other than $H_2$ which touches $H_1$. (If $H_2$ touches $H_1$ in two discs, then let $H'_2 = H_2$.) If $H'_2 \in A$, then we may extend $A$ over $H_1$. If $H'_2 \not\in A$, we may extend $A$ over $H_1 \cup H'_2$. In each case, the maximality of $A$ is contradicted. □

We now consider the various possibilities for $A$ case by case.

**Case 1.** $A$ is an $I$-bundle over a disc $G$.

In this case, we replace $A$ with a single 2-handle $H_2$. We attach $H_2$ to $H^0 \cup H^1 - A$ using the annulus $\partial_vA$. We now check that $\mathcal{H}' = (\mathcal{H} - A) \cup H_2$ is a handle structure. By (vii) of 7.4, the only requirement that is not immediately obvious is that $H_2$ touches some 1-handle. But, if not, then $\partial_vA$ would have been an annular component of $\mathcal{F}$, contradicting Lemma 8.2.

We now check that extended $\mathcal{F}$-complexity is decreased and that (i)-(iv) of 6.3 are satisfied. It is clear that $\mathcal{H}'$ respects $\mathcal{H}$. This is because (for $i = 0$ and 1) each $i$-handle of $\mathcal{H}'$ is an $i$-handle of $\mathcal{H}$ and inherits its product structure. Of course, $H_2$ need not lie in any 2-handle of $\mathcal{H}$, but this was not
a requirement of Definition 6.1. This explains why Definition 6.1 did not make more stringent requirements on 2-handles.

If \( F \) is any component of \( \mathcal{F} \) and \( F' = F \cap \mathcal{F}' \), then \( F' \) is either a copy of \( F \), or is completely removed, or is obtained by performing a sequence of the following operations: Remove a 0-handle of \( F \) which abuts precisely two 1-handles of \( F \), and amalgamate these two 1-handles into a single 1-handle of \( F' \). Hence, Lemma 7.2 ensures that Conditions (ii) and (iv) of 6.3 are satisfied and also that extended \( \mathcal{F} \)-complexity is reduced. Condition (iii) of 6.3 is clear.

We may therefore assume that \( \mathcal{A} \) is an \( I \)-bundle over a surface \( G \) other than a disc.

**Case 2.** \( \partial_v \mathcal{A} = \emptyset \).

In other words, \( G \) is a closed surface. If \( \partial \mathcal{A} \) is entirely contained in \( \mathcal{R}_- \) or entirely contained in \( \mathcal{R}_+ \), then it has zero Euler characteristic, since \( (M, \gamma) \) is taut and so \( G \) is a torus or Klein bottle. In either case, we pick a non-separating orientation-preserving curve in \( G \), and perform a decomposition along the \( I \)-bundle over this curve. This cuts \( \mathcal{A} \) into a solid torus. We perform one further taut decomposition along a product disc, ending with a single 3-ball. We let this be a 0-handle of \( \mathcal{H}' \). Suppose now that \( \partial \mathcal{A} \) intersects both \( \mathcal{R}_- \) and \( \mathcal{R}_+ \). Then \( \mathcal{A} \) must be a product \( G \times I \). We can then perform a sequence of decompositions along product annuli and product discs, ending with a 3-ball, which again we let be a single 0-handle of \( \mathcal{H}' \).

We therefore assume that \( \partial_v \mathcal{A} \) is nonempty.

**Case 3.** Each annulus of \( \partial_v \mathcal{A} \) is an incompressible product annulus.

Then by Lemma 4.2 of [10], the decomposition \( (M, \gamma) \xrightarrow{\partial_v \mathcal{A}} (M', \gamma') \) is taut. We perform this decomposition. In other words, we separate off \( \mathcal{A} \) from \( \mathcal{H} - \mathcal{A} \), and add sutures \( \gamma' \) as appropriate. By definition, \( \mathcal{H} - \mathcal{A} \) is a handle structure.

The amalgam \( \mathcal{A} \) does not inherit a handle structure (for example, 1-handles of \( \mathcal{A} \) need not be attached to 0-handles of \( \mathcal{A} \)). However, since \( \partial_v \mathcal{A} \) touches both \( \mathcal{R}_- \) and \( \mathcal{R}_+ \), the \( \partial I \)-bundle over \( G \) cannot be connected, and so \( \mathcal{A} \) must a product \( G \times I \). As in Case 2, we may perform some further decompositions along product discs, which reduce \( G \times I \) to a ball. We let this be a single 0-handle of \( \mathcal{H}' \).

This whole procedure has the effect of removing some components of \( \mathcal{F} \) and also replacing some 0-handles and 1-handles of \( \mathcal{F} \) with arcs of \( \gamma' \cap \mathcal{H}^0(M') \). An argument almost identical to that in Procedure 3 of Section 7 establishes that the hypotheses of Lemma 7.2 hold. Therefore, (i)-(iv) of 6.3 hold and extended \( \mathcal{F} \)-complexity has been reduced.

We therefore assume that some annulus of \( \partial_v \mathcal{A} \) is not an incompressible product annulus.
Case 4. Some annulus $A$ of $\partial_v A$ is compressible in $M$.

Then $A$ compresses in $M$ to two discs $D'_1$ and $D'_2$ with boundaries in $\mathcal{R}_\pm$. Since $\mathcal{R}_\pm$ is incompressible in $M$ and $M$ is irreducible, $D'_1$ and $D'_2$ are parallel in $M$ to discs $D_1$ and $D_2$ in $\mathcal{R}_\pm$. We pick $A$ so that the curve $\partial D_1$ is an innermost curve of $\partial_v A \cap \mathcal{R}_\pm$ in $\mathcal{R}_\pm$. Since $A$ is not an $I$-bundle over a disc, this implies that $\text{int}(D_1)$ is disjoint from $A$. The parallelity region between the discs $D_i$ and $D'_i$ is a ball $B_i$. Then, $B_1$ and $B_2$ are either disjoint or nested.

Case 4A. $B_1$ and $B_2$ are disjoint.

Then, $D_1$ and $D_2$ are disjoint, and the sphere $D_1 \cup D_2 \cup A$ bounds a ball $B$ in $M$. Since $A \cap B = A$, we can extend the $I$-bundle structure of $A$ over $B$. This contradicts the maximality of $A$.

Case 4B. $B_1$ and $B_2$ are nested.

Then $B_1 \subset B_2$ and $D_1 \subset D_2$. The component $V$ of $M - \text{int}(N(A))$ lying wholly within $B_2$ is homeomorphic to the exterior of a knot in $S^3$. The amalgam $A$ lies in $V$, and we may therefore remove $V$ from $M$ and still retain a handle structure. This does not change the homeomorphism type of $M$ and Lemma 7.3 gives that extended $\mathcal{F}$-complexity decreases and that (i)-(iv) of 6.3 hold.

Case 5. $\partial_v A$ is incompressible, and some component of $\partial_v A$ is not a product annulus.

Now, the $\partial I$-bundle over $G$ has at most two components. Therefore, if some component of $\partial_v A$ is not a product annulus, then no component of $\partial_v A$ is a product annulus. Let us suppose that $\partial_v A$ is disjoint from $\mathcal{R}_-$ (say).

Pick any component $A$ of $\partial_v A$. Then we let $A_1$ and $A_2$ be two parallel copies of $A$, incoherently oriented in such a way that the parallelity region $Y$ in $M' = M - \text{int}(N(A_1 \cup A_2))$ inherits four sutures. Isotope $A_1$ and $A_2$ a little so that they become standard surfaces. Consider the decomposition $(M, \gamma) \xrightarrow{A_1 \cup A_2} (M', \gamma')$.

Case 5A. $(M', \gamma')$ is taut.

Then we perform this decomposition. We now check the requirements of 6.3 and also that extended $\mathcal{F}$-complexity has been reduced. We will use Lemma 7.2 to do this.

Let $\mathcal{H}'$ be the handle structure which $(M', \gamma')$ inherits. Consider a component $F$ of $\mathcal{F}$ which is altered by this decomposition, and let $F'$ be $F \cap \mathcal{F}'$. By Lemma 8.2, $F$ must have had positive index.

Suppose that the extended $\mathcal{F}$-complexity of $F'$ is at least that of $F$; we aim to reach a contradiction. We must have $C_1(X) \geq C_1(F)$ for some component $X$ of $F'$. But each 1-handle of $F$ gives rise to precisely one 1-handle of $F'$. Hence, $X$ must have all the 1-handles of $F'$. 
Each component of $A \cap F$ yields three discs of $F'$. Two of these discs have no 1-handles and intersect $\gamma'$ four times. The remaining disc has at least one 1-handle, and has negative index. Since it has least one 1-handle, it must be $X$, and therefore $X$ has negative index. Therefore $C_2(X) < C_2(F)$. Hence, $C_2^+(F') < C_2^+(F)$. Lemma 7.2 now gives that (i), (ii) and (iv) of 6.3 hold, and that extended complexity has been reduced. Verifying (iii) of 6.3 is straightforward.

Case 5B. $(M', \gamma')$ is not taut.

Since $M$ is irreducible, so must $M'$ be. Also, $R_{\pm}(M')$ is norm-minimising in $H_2(M', \gamma')$. For if $S$ is any surface in $M'$ with $S \cap \partial M' = \gamma'$ and $[S, \partial S] = [R_{\pm}(M'), \gamma'] \in H_2(M', \gamma')$, then $[S - Y, \partial S - Y] = [R_{\pm}(M') - Y, \gamma' - Y] = [R_{\pm}(M), \gamma] \in H_2(M, \gamma)$. So, $\chi_-(S) \geq \chi_-(S - Y) \geq \chi_-(R_{\pm}(M)) = \chi_-(R_{\pm}(M'))$. Hence, the only way that $(M', \gamma')$ can fail to be taut is if $R_{\pm}(M')$ is compressible. This compression cannot reduce $\chi_-(R_{\pm}(M'))$, as $R_{\pm}(M')$ is norm-minimising. Hence, any compressible component of $R_{\pm}(M')$ is a torus or annulus. However, any circle in a compressible annulus is homotopically trivial in $M$. In particular, $A$ could not have been incompressible, contrary to assumption. Thus, if $R_{\pm}(M')$ is not taut, there are three cases to consider:

(i) Only one of $A_1$ and $A_2$ (say $A_1$) lies in a compressible torus component of $R_{\pm}(M')$ (called $T_1$, say) which disjoint from $\gamma'$, or
(ii) $A_1$ and $A_2$ both lie in the same compressible torus $T_1$ disjoint from $\gamma'$, or
(iii) $A_1$ and $A_2$ lie in distinct compressible tori $T_1$ and $T_2$ disjoint from $\gamma'$.

Since $T_i$ is compressible and $M'$ is irreducible, $T_i$ bounds a solid torus $V_i$ in $M'$.

In Case (iii), the component of $M$ containing $A$ is the union of two solid tori, glued along an essential annulus. Thus, it is a Seifert fibre space with base space a disc and having at most two exceptional fibres (which are the cores of the solid tori). Also, it is disjoint from $\gamma$. Recall that, in the statement of 6.3, we explicitly ruled out the case where it has two exceptional fibres. If it has at most one exceptional fibre, it is a solid torus, and again, we ruled this case out.

In Case (ii), we pick the $A_i$ which is closest to $A$. Then, the orientation of $A_i$ and $R_+ \cap A$ agree near $\partial A_i$. The decomposition $(M, \gamma) \xrightarrow{A_i} (M_1, \gamma_1)$ is taut. Exactly as in Case 5A, this reduces extended $\mathcal{F}$-complexity and (i)-(iv) of 6.3 are satisfied.

In Case (i), suppose first that $M - \text{int}(\mathcal{N}(V_1))$ contains $A$. As above, the decomposition $(M, \gamma) \xrightarrow{A_2} (M_1, \gamma_1)$ is taut, and, again, this reduces extended $\mathcal{F}$-complexity and (i)-(iv) of 6.3 are satisfied.
Suppose now that $V_1$ contains $A$. Then again the decomposition $(M, \gamma) \xrightarrow{A_2^+} (M_1, \gamma_1)$ is taut. This time one must work a little harder to verify that extended $F$-complexity decreases and that (i)-(iv) of 6.3 are satisfied. Let $F$ be a component of $\mathcal{F}$, let $F' = F \cap \mathcal{F}(M_1)$ and let $H_0$ be the 0-handle of $\mathcal{H}$ containing $F$. By Lemma 8.2, if $F$ is altered, then it must have positive index, and so it contributes towards $\mathcal{F}$-complexity. If $C_F^+(F') \geq C_F^+(F)$, then as in Case 5A, there must be a single component $X$ of $F'$ containing all the 1-handles of $F'$. Also, $X$ arises from a component of $F \cap A$. However, unlike in Case 5A, $X$ will not have negative index. In fact, it will be a disc which intersects $\gamma_1$ in four points. Hence, it has index two. Now, $F$ has positive index and therefore its index is at least two. Thus, if $X$ in Case 5A, then $C_F^+(F') > C_F^+(F)$, then we may assume that $F$ is a disc intersecting $\gamma$ in four points. In this case, all but two 0-handles of $F$ have valence two and are disjoint from $\gamma$. The two remaining 0-handles $D_1$ and $D_2$ each have valence one. These two handles contain a total of four points of $\gamma \cap F$. Since $\partial_+ A$ is disjoint from $\mathcal{R}_-$, each $D_i$ contains an even number of points of $\gamma \cap F$. If one of these 0-handles contains no points of $\gamma$, then it becomes a compression disc for $\mathcal{R}_+(M_1)$, which contradicts the fact that $(M_1, \gamma_1)$ is taut. Hence, each $D_i$ contains precisely two points of $\gamma \cap F$. If these two points are joined by an arc of $\gamma \cap \partial H_0$, then again $\mathcal{R}_+(M_1)$ is compressible, which is a contradiction. Therefore, for each $i$, the two points $\gamma \cap D_i$ are not joined by an arc of $\gamma \cap \partial H_0$.

Now, $F'$ is $X$, together with two index zero discs. We remove the two index zero discs using Procedure 4. The component $X$ is a copy of $F$, and so $C_{\mathcal{F}}(F') = C_{\mathcal{F}}(F)$. If $F$ was not the only component of $\mathcal{F}$ in $H_0$, then these components of $\mathcal{F}$ have positive index, since otherwise we can apply one of Procedures 1-5. Hence, $n(\mathcal{H} \cap H_0) = n(\mathcal{H}_0)$, where $n(\mathcal{H} \cap H_0)$ was defined in Section 5 to be the number of 0-handles of $\mathcal{H} \cap H_0$ containing a component of $\mathcal{F}$ of positive index. This implies that $C(H_0 \cap \mathcal{H}_0) < C(H_0)$. If $F$ was the only component of $\mathcal{F}$ in $H_0$, then this is a trivial modification. This verifies (i)-(iv) of 6.3.

We now need to check $C^+(\mathcal{H}') < C^+(\mathcal{H})$. But, if it is not, then the above must happen in every 0-handle of $\mathcal{H}$ which is altered by the decomposition. This implies that the component of $M - \text{int}(\mathcal{N}(V_1))$ containing $A$ is a solid torus $V_2$ with $A \cap V_2$ a single annulus in $\partial V_2$ having winding number one. But $V_1$ is a solid torus, and so this component of $M$ is a solid torus, contradicting one of the assumptions of 6.3. \qed
9. Modifications to a decomposing surface.

In the previous two sections, we performed a sequence of alterations to $\mathcal{H}$. We are now ready to tackle Proposition 6.4. Consider the taut decomposition $(M, \gamma) \rightarrow (M_S, \gamma_S)$, where $S$ is a compact oriented surface properly embedded in $M$, having essential intersection with $\mathcal{R}_\pm$. This implies that $S$ is taut and hence incompressible. Thus, by Lemmas 4.5 and 4.9, we can assume that $S$ is in standard form in $H$. But, as was remarked in Section 5, there is a great deal of freedom over the form of $S \cap \partial H^0$. The aim here is to perform a series of alterations to $S$, creating a new standard surface $S'$ which has a considerably more restricted intersection with $\partial H^0$. The sutured manifold obtained by decomposing $(M, \gamma)$ along $S'$ will be written as $(M_{S'}, \gamma_{S'})$.

**Modification 1. Tubing along an arc.**

Suppose that $\alpha$ is an arc in $\mathcal{R}_\pm$ with $\alpha \cap S = \partial \alpha$. Then there is an embedding of $\alpha \times [-1, 1]$ in $\mathcal{R}_\pm$ with $\alpha \times \{0\} = \alpha$ and $(\alpha \times [-1, 1]) \cap S = \partial \alpha \times [-1, 1]$. Suppose that the orientation that $\alpha \times [-1, 1]$ inherits from $\mathcal{R}_\pm$ agrees with the orientation of $S$ near $\partial \alpha \times [-1, 1]$. Then we call $\alpha$ a tubing arc. We construct a new surface $S'$ as follows: Embed $\alpha \times [-1, 1] \times [0, 1]$ in $M$ so that $\alpha \times [-1, 1] \times [0, 1]$ is a product disc in $(M_{S'}, \gamma_{S'})$. There is a commutative diagram:

![Diagram of tubing along an arc](image)

Figure 9.1.

We say that $S'$ is obtained from $S$ by tubing along the arc $\alpha$. Note that if $\{\ast\}$ is a point in $\alpha - \partial \alpha$, then $P = \{\ast\} \times [-1, 1] \times [0, 1]$ is a product disc in $(M_{S'}, \gamma_{S'})$. There is a commutative diagram:
\[ (M, \gamma) \xrightarrow{S'} (M_{S'}, \gamma_{S'}) \]
\[ \downarrow \quad \downarrow P \]
\[ (M_S, \gamma_S) = (M_S, \gamma_S). \]

Hence, \((M_S, \gamma_S)\) is taut if and only if \((M_{S'}, \gamma_{S'})\) is taut. However, \(S'\) need not have essential intersection with \(R_\pm(M)\).

**Modification 2.** Slicing under an incompressible annulus.

Suppose that there is an annulus \(A\) in \(R_\pm\) which is incompressible in \(M\) and has \(A \cap S = \partial A\). Let \(A \times [0, 1]\) be embedded in \(M\), so that \((A \times [0, 1]) \cap \partial M = A \times \{0\} = A\), and so that \((A \times [0, 1]) \cap S = \partial A \times [0, 1]\). If the orientation of \(A\) agrees with that of \(S\) near \(\partial A\), then we construct a new surface \(S' = S \cup (A \times \{1\}) - (\partial A \times [0, 1])\) by *slicing under the incompressible annulus* \(A\). Let \(C\) be a core circle of \(A\). If we give \(C \times [0, 1]\) any orientation, then we have a commutative diagram:

\[ (M, \gamma) \xrightarrow{S'} (M_{S'}, \gamma_{S'}) \]
\[ \downarrow S \quad \downarrow C \times [0, 1] \]
\[ (M_S, \gamma_S) = (M_S, \gamma_S). \]

Since \(A\) is incompressible in \(M\), the annulus \(C \times [0, 1]\) is incompressible in \((M_{S'}, \gamma_{S'})\), and so by 4.2 of [10], \((M_S, \gamma_S)\) is taut if and only if \((M_{S'}, \gamma_{S'})\) is taut.

![Figure 9.2](slide.png)

**Modification 3.** Sliding \(\partial S\) across \(\gamma\).

Suppose that \(D\) is a disc in \(\partial M\), such that \(D \cap S\) is an arc \(\alpha\) in \(\partial D\) and \(D \cap \gamma\) is an arc properly embedded in \(D\) disjoint from \(\alpha\). Suppose that the orientation of \(D\) near \(\partial S\) agrees with the orientation of \(S\). Let
$D \times [0, 1]$ be embedded in $M$ so that $(D \times [0, 1]) \cap \partial M = D \times \{0\} = D$ and $(D \times [0, 1]) \cap S = \alpha \times [0, 1]$. Let

$$S' = S \cup \partial(D \times [0, 1]) - (D - \partial D) \times \{0\} - (\alpha - \partial \alpha) \times [0, 1].$$

(See Figure 9.2.) Then $(M_S, \gamma_S)$ and $(M'_S, \gamma'_S)$ are homeomorphic.

**Modification 4. Slicing under a disc of contact.**

Suppose that there is a disc $D$ in $\mathbb{R}^+\mathbb{R}^-\mathbb{R}_0$ with $D \cap S = \partial D$, with the orientation of $D$ matching that of $S$ near $\partial D$. Then $D$ is known as a disc of contact. Embed $D \times [0, 1]$ in $M$, so that $(D \times [0, 1]) \cap \partial M = D \times \{0\} = D$, and so that $(D \times [0, 1]) \cap S = \partial D \times [0, 1]$. The surface $S' = S \cup (D \times \{1\}) - (\partial D \times [0, 1])$ is obtained from $S$ by slicing under the disc of contact.

![Figure 9.3.](image)

Unfortunately, in this case, we have no guarantee that $(M'_S, \gamma'_S)$ is taut. But the following sequence of lemmas circumvents this. We introduce a temporary definition.

**Definition 9.4.** The surface $S$ in $(M, \gamma)$ is *mountainous* if some curve of $\partial S$ bounds a disc $D$ in $\mathbb{R}^+\mathbb{R}^-\mathbb{R}_0$, such that the orientation of $D$ disagrees with that of $S$ near $\partial D$. The disc $D$ may also intersect $S$ away from $\partial D$.

**Lemma 9.5.** Suppose that $S$ is not mountainous. Let $S'$ be obtained from $S$ by a sequence of Modifications 1, 2 and 3. Then $S'$ is not mountainous.

**Proof.** It suffices to consider the case where $S'$ is obtained from $S$ by a single modification of Type 1, 2 or 3. Consider first a modification of Type 1. Suppose that some curve of $\partial S'$ bounds a disc $D$ in $\mathbb{R}^+\mathbb{R}^-\mathbb{R}_0$, with orientation on $D$ disagreeing with that on $S'$. Then at least one of the curves of $\alpha \times \{-1, 1\}$ must lie in $\partial D$, since $S$ was not mountainous.

If both curves of $\alpha \times \{-1, 1\}$ lie in $\partial D$, then the curves $\partial D \cup (\partial \alpha \times [-1, 1]) - ((\alpha - \partial \alpha) \times \{-1, 1\})$ would have bounded the discs $D - ((\alpha - \partial \alpha) \times [-1, 1])$. Then $S$ would have been mountainous.

Suppose now that only one curve of $\alpha \times \{-1, 1\}$ lies in $\partial D$, say $\alpha \times \{-1\}$. Then $\alpha \times \{1\}$ lies in $D - \partial D$, and so is part of a component of $\partial S'$ bounding
a disc $D'$ in $\mathcal{R}_\pm$. Then, $D - (D' - \partial D') - ((\alpha - \partial \alpha) \times [-1, 1])$ is a disc which would have made $S$ mountainous.

Now consider the case where $S'$ is obtained from $S$ by slicing under an incompressible annulus $A$. This has the effect of removing two curves of $\partial S$, neither of which bounded discs in $\mathcal{R}_\pm$. Hence, in this case, $S'$ is mountainous if and only if $S$ is mountainous.

Finally, consider the case where $S'$ is obtained from $S$ by sliding an arc of $\partial S$ across $\gamma$. Then, this only creates new intersection points between the surface and $\gamma$, and so a curve of $\partial S'$ disjoint from $\gamma$ is a copy of a curve of $\partial S$ disjoint from $\gamma$. Thus, if $S'$ is mountainous, then so is $S$. □

Lemma 9.6. Suppose that no component of $S$ is a disc disjoint from $\gamma$. Let $S'$ be a surface obtained from $S$ by Modifications 1, 2 and 3. Then no component of $S'$ is a disc disjoint from $\gamma$.

Proof. It suffices to consider a single modification of Type 1, 2 or 3. If a component of $S'$ which is a disc disjoint from $\gamma$ arises by tubing along an arc $\alpha$, then the components of $S$ containing $\partial \alpha$ were both discs disjoint from $\gamma$, contrary to assumption. If a disc component of $S'$ arises by slicing under an annulus, then that annulus could not have been incompressible in $M$. A Type 3 modification cannot create components of $S'$ disjoint from $\gamma$. □

Lemma 9.7. Let $S$ be a surface in a taut sutured manifold $(M, \gamma)$ having essential intersection with $\mathcal{R}_\pm$. Let $S_2$ be obtained from $S$ by a sequence of Modifications 1, 2 and 3, and let $S_3$ be obtained from $S_2$ by slicing under a disc of contact $D$. Then $S_3$ is in fact obtained from $S$ by a sequence of Modifications 1, 2 and 3.

Proof. We shall prove this by induction on the number $n$ of Type 1, 2 and 3 modifications from $S$ to $S_2$. For $n = 0$, the statement of the lemma is empty, since $S = S_2$ has essential intersection with $\mathcal{R}_\pm(M)$ and so has no discs of contact.

Suppose now the lemma is true for sequences of length at most $(n - 1)$. Let $S_1$ be the surface obtained from $S$ by the first $(n - 1)$ modifications. Then $S_2$ is obtained from $S_1$ by a modification of Type 1, 2 or 3, and $S_3$ is obtained from $S_2$ by slicing under a disc of contact $D$.

Suppose that $S_2$ is obtained from $S_1$ by tubing along an arc $\alpha$. Then, $D$ is disjoint from $\alpha \times (-1, 1)$. If the disc $D$ is disjoint from $\alpha \times \{-1, 1\}$, then we can obtain $S_3$ from $S_1$ by slicing under $D$, which gives a surface $S_4$ say, then tubing along $\alpha$. Inductively, $S_4$ is obtained from $S$ by Modifications 1, 2 and 3, and so the lemma is true in this case. Hence, we may assume that $D$ touches at least one of the arcs $\alpha \times \{-1, 1\}$. If it touches only one arc, then $S_3$ is ambient isotopic to $S_1$, and the lemma is true. If $D$ touches both arcs, then $D \cup (\alpha \times [-1, 1])$ is an annulus $A$ in $\mathcal{R}_\pm$. The two curves of $\partial A$ are boundary components of $S_1$. If $A$ is compressible in $M$, then both curves
of $\partial A$ bound discs in $\mathcal{R}_\pm$, since $(M, \gamma)$ is taut. One of these discs has an orientation disagreeing with that of $S_1$ near the boundary of the disc, and so $S_1$ is mountainous, contrary to Lemma 9.5. Hence, $A$ is incompressible in $M$. We may slice under $A$ to obtain $S_3$ from $S_1$. This proves the lemma in this case.

Suppose that $S_2$ is obtained from $S_1$ by slicing under an incompressible annulus $A$. Then $A$ cannot lie in $D$, since $A$ is incompressible. Hence, we can obtain $S_3$ from $S_1$ by slicing under $D$, then slicing under $A$. The inductive hypothesis proves the lemma.

Similarly, if $S_2$ is obtained from $S_1$ by sliding an arc of $\partial S_1$ across $\gamma$, then the relevant component of $\partial S_2$ is disjoint from $D$, and therefore we may obtain $S_3$ from $S_1$ by slicing under $D$, then performing the Type 3 modification. Again, the inductive hypothesis proves the lemma. □

**Corollary 9.8.** Let $S$ be a surface in the taut sutured manifold $(M, \gamma)$ having essential intersection with $\mathcal{R}_\pm$. Then any surface obtained from $S$ by Modifications 1, 2, 3 and 4 is in fact obtained from $S$ by Modifications 1, 2 and 3.

Unfortunately, if $S'$ is a surface created from $S$ by Modifications 1, 2 and 3, then $S'$ need not be incompressible. The incompressibility of $S$ was useful in showing that $S$ can be isotoped into standard form. We therefore need the following lemma:

**Lemma 9.9.** Let $S$ be a surface in $(M, \gamma)$, having essential intersection with $\mathcal{R}_\pm$. Suppose that $(M, \gamma) \xrightarrow{S} (M_S, \gamma_S)$ is taut. Let $S'$ be a surface obtained from $S$ by a sequence of Modifications 1, 2 and 3. If $S'$ is in vertical form with respect to some handle structure on $(M, \gamma)$, then we may perform an ambient isotopy of $S'$ and perhaps some Type 4 modifications, taking $S'$ into standard form, without increasing its complexity.

**Proof.** Consider again the proof of Lemma 4.9. The crucial property of $S$ was that if $D$ is any disc in $M - \partial M$ with $D \cap S = \partial D$, then $\partial D$ bounds a disc $D'$ in $S$, and we may ambient isotopy $D'$ onto $D$. In fact, we need only restrict attention to discs $D$ lying in a single 0-handle of $M$. In the case of $S'$, the circle $\partial D$ need not bound a disc in $S'$, since $S'$ might not be incompressible. But, since $(M_{S'}, \gamma_{S'})$ is taut, $\partial D$ bounds a disc $D'$ in $\mathcal{R}_\pm(M_{S'})$. Consider a circle $C$ of $D' \cap \partial S'$ innermost in $D'$. By Lemma 9.6, this cannot bound a disc of $S'$. Hence, it bounds a disc of contact in $\mathcal{R}_\pm(M)$. Slice under this disc of contact. By Corollary 9.8, the new surface (also called $S'$) is in fact obtained from $S$ by a sequence of Modifications 1, 2 and 3. Hence, the new $(M_{S'}, \gamma_{S'})$ is taut. So, we may repeat this process and, in this way, we may remove all circles of $D' \cap \partial S'$. But then $\partial D$ bounds a disc $D'$ in $S'$, and we may ambient isotop $D'$ onto $D$. The new surface $S'$ is obtained from the old $S'$ by removing $S' \cap (D' - \partial D')$ and gluing in $D$. Thus, the complexity
of the new $S'$ is no more than the complexity of the old $S'$. Continuing in this fashion, we may get $S'$ into standard form. □

We will alter $S$, using Modifications 1, 2, 3 and 4, until $S$ has become a standard surface satisfying each of the following three conditions:

1. Each curve of $S \cap \partial H^0$ meets any 1-handle of $F$ in at most one arc.
2. There exists no tubing arc in $\partial F^0 \cap R_{±}$.
3. Suppose that $D$ is a disc in $F^0$ with $\partial D$ the union of two arcs $\alpha$ and $\beta$, where $\alpha = S \cap \partial D$ and $\beta = D \cap \partial F^0$. Suppose that one endpoint of $\alpha$ lies in $R_{±}$ and one endpoint lies in $F^1$. Then at least one of the following must happen:
   • $\beta$ touches at least two components of $\partial F^0 \cap \partial F^1$,
   • $\beta$ touches $\gamma$ and the orientation of $\alpha$ and $\beta$ disagree locally near $\alpha \cap \beta \cap R_{±}$, or
   • $\beta$ touches $\gamma$ at least twice, and the orientation of $\alpha$ and $\beta$ agree locally near $\alpha \cap \beta \cap R_{±}$.

Diagrams clarifying Conditions 1, 2 and 3 are given in Figures 9.10, 9.11 and 9.12. The alterations to $S$ which ensure that Conditions 1, 2 and 3 hold will reduce its complexity, and so they are guaranteed to terminate. The above three conditions are not quite sufficient for our purposes. We also wish to ensure that the following two conditions hold:

4. Each curve of $S \cap \partial H^0$ meets any component of $R_{±} \cap \partial H^0$ in at most one arc.
5. If $\alpha$ is an arc of $S \cap F^0$ with both endpoints in $R_{±}$, then each of the two arcs in $\partial F^0$ joining $\partial \alpha$ must either touch $\partial F^1$ or hit $\gamma_S$ more than twice.

A diagram clarifying Condition 5 is given in Figure 9.17. To achieve Conditions 4 and 5, we will need two further types of modification to $S$, which we will describe later. We now show that Conditions 1-3 can be achieved. By Lemma 4.9, we may assume that $S$ is in standard form. Each alteration to $S$ leaves it in vertical form, but not necessarily standard form. However, we can then use Lemma 9.9 to get $S$ into standard form, since the alterations to $S$ used there result in the removal of some components of $S \cap \partial H^0$, and hence the new $S$ also satisfies Conditions 1-3.

**Condition 1.** Each curve of $S \cap \partial H^0$ meets any 1-handle of $F$ in at most one arc.

Suppose that this condition does not hold. We will construct a ball $B$ lying in $M - \partial M$, such that $B \cap S$ is a disc in $\partial B$. We will then ambient isotope $S$ across $B$, and in doing so, reduce the complexity of $S$.

By assumption, there is a curve $C$ of $S \cap \partial H^0$ containing two sub-arcs $\alpha_1$ and $\alpha_2$, which are both properly embedded in the same 1-handle $D_1$ of $F$. Pick $C$ to be a curve of $S \cap \partial H^0$ which is innermost in $\partial H^0$ amongst all
curves with this property. Since $S$ is standard, there is a disc $D_2$ of $S \cap \mathcal{H}^0$ with $\partial D_2 = C$. Let $H_0$ (respectively, $H_2$) be the 0-handle (respectively, 2-handle) containing $\alpha_1$ and $\alpha_2$, and let $E_1$ and $E_2$ be the discs of $H_2 \cap S$ containing $\alpha_1$ and $\alpha_2$. By the ‘innermost’ assumption on $C$, the two arcs $\alpha_1$ and $\alpha_2$ are adjacent in $D_1$, in the sense that no other arcs of $S \cap D_1$ lie between them. Let $B'$ be the closure of the component of $H_2 - (E_1 \cup E_2)$ lying between $E_1$ and $E_2$. The ball $B'$ will be part of $B$.

Let $\beta_1$ be an arc in the interior of $D_1$ which runs from $\alpha_1$ to $\alpha_2$, but which intersects $S$ in no other points. Let $\beta_2$ be an arc properly embedded in $D_2$, with $\partial \beta_2 = \partial \beta_1$. Then $\beta_1 \cup \beta_2$ bounds a disc $D_3$ in $H_0$. We can assume that $D_3 - \partial D_3$ misses $S$ and $\partial H_0$. Let $B$ be a small neighbourhood of $D_3 \cup B'$ in $M$. Then an ambient isotopy of $S$ across $B$ has the effect of reducing $|S \cap \mathcal{H}^2|$, by removing the discs $E_1$ and $E_2$. The new surface is a vertical surface with lower complexity than that of $S$.

**Condition 2.** There exists no tubing arc in $\partial \mathcal{F}^0 \cap \mathcal{R}_\pm$. 

---

**Figure 9.10.**

**Figure 9.11.**
Suppose that $\alpha$ is such an arc. We will tube $S$ along $\alpha$. Let $H_1$ be the 1-handle of $\mathcal{H}$ containing $\alpha$. We may pick $\alpha \times [-1, 1]$ so that $(\alpha \times [-1, 1]) \cap H_1$ is vertical in $H_1$, and so that $\text{cl}((\alpha \times [-1, 1]) - H_1)$ is two small discs in $\mathcal{H}^0$. Then the surface $S'$ constructed from $S$ by tubing along $\alpha$ has lower complexity than that of $S$, since $|S \cap \mathcal{H}^2| = |S' \cap \mathcal{H}^2|$, and $|\partial S' \cap \mathcal{H}^1| = |\partial S \cap \mathcal{H}^1| - 2$.

**Condition 3.** Suppose that $D$ is a disc in $\mathcal{F}^0$ with $\partial D$ the union of two arcs $\alpha$ and $\beta$, where $\alpha = S \cap \partial D$ and $\beta = D \cap \partial \mathcal{F}^0$. Suppose that one endpoint of $\alpha$ lies in $R_\pm$ and one endpoint lies in $\mathcal{F}^1$. Then at least one of the following must happen:

- $\beta$ touches at least two components of $\partial \mathcal{F}^0 \cap \partial \mathcal{F}^1$,
- $\beta$ touches $\gamma$ and the orientation of $\alpha$ and $\beta$ disagree locally near $\alpha \cap \beta \cap R_\pm$, or
- $\beta$ touches $\gamma$ at least twice, and the orientation of $\alpha$ and $\beta$ agree locally near $\alpha \cap \beta \cap R_\pm$.

![Figure 9.12.](image)

Suppose that $D$ is such a disc but that it fails to satisfy each of the three alternatives of Condition 3. In particular, $\beta$ touches only one component of $\partial \mathcal{F}^0 \cap \partial \mathcal{F}^1$ (which is therefore the component of $\partial \mathcal{F}^0 \cap \partial \mathcal{F}^1$ which contains an endpoint of $\alpha$). Suppose also that there is no sub-arc of $\beta$ which violates Condition 2. Let $H_1 = D^2 \times I$ be the 1-handle of $\mathcal{H}$ containing $D$.

There are a number of cases to consider. Suppose first that $\beta \cap \gamma = \emptyset$. Then, let $B$ be the vertical ball $D \times I$ in $H_1$. Let $D'$ be the disc of $S \cap \mathcal{H}^2$ which touches $\alpha$ at a single point. Let $B'$ be the closure of the component of $\mathcal{H}^2 - D'$ which has nonempty intersection with $\beta - \partial \beta$. There is an ambient isotopy of $S$ across the ball $B \cup B'$, leaving $S$ in vertical form, and reducing
This isotopy will also move parts of $S$ lying in $D \times I$, but this causes no problems.

Suppose now that $\beta \cap \gamma \neq \emptyset$. Then, by assumption, the orientations of $\alpha$ and $\beta$ agree near $\alpha \cap \beta \cap \mathcal{R}_\pm$, and also $\beta \cap \gamma$ is a single point. Suppose first that there is no arc of $S \cap \mathcal{F}^0$ other than $\alpha$ lying in $D$. Then we perform a Type 3 modification, supported in a small neighbourhood of $H_1$, which slides the vertical arc $(\beta \cap \alpha \cap \mathcal{R}_\pm) \times I$ in $H_1$ across the vertical arc $(\beta \cap \gamma) \times I$. Then, we can perform the ambient isotopy described above to reduce the complexity of $S$.

Suppose now that there exists some arc of $S \cap \partial \mathcal{F}^0$ other than $\alpha$ lying in $D$. Let $\alpha_1$ be the arc adjacent to $\alpha$. If the sub-arc $\beta_1$ of $\beta$ lying between $\alpha$ and $\alpha_1$ is a tubing arc, then Condition 2 is violated. If $\beta_1$ is not a tubing arc, then there are two possibilities: $\beta_1$ touches $\gamma$, or the orientations of $\alpha_1$ and $\beta_1$ disagree near $\alpha_1 \cap \beta_1$. Applying this argument to each arc of $S \cap D$ with an endpoint lying between $\beta \cap \alpha \cap \mathcal{R}_\pm$ and $\beta \cap \gamma$, we see that these arcs are all coherently oriented. In particular, we can slide each of these arcs across $\gamma$. Then, we can apply the ambient isotopy described above. Thus, we may ensure that $S$ satisfies Conditions 1, 2 and 3. We now ensure that $S$ also satisfies Condition 4 and 5. To do this, we will need two further modifications.

**Modification 5.** Surgery along a product disc.

This is defined to be the reverse of a Type 1 modification.

**Modification 6.** Removal of a product region.

Suppose that a component $F_1$ of $S$ is parallel to a surface $F_2$ in $\mathcal{R}_\pm$, and the orientations of $F_1$ and $F_2$ disagree near $\partial F_1 = \partial F_2$. Suppose also that the interior of the parallelity region between $F_1$ and $F_2$ is disjoint from $S$. If we remove $F_1$ from $S$, creating a new surface $S'$, then $(M_S, \gamma_S)$ is homeomorphic to $(M_{S'}, \gamma_{S'})$, plus a product component $(F_1 \times [0, 1], \partial F_1 \times \{0\})$. Hence, $(M_{S'}, \gamma_{S'})$ is taut if and only if $(M_S, \gamma_S)$ is taut.

We need some lemmas to ensure that Modifications 5 and 6 are well behaved.

**Lemma 9.13.** Let $S$ be a surface in $(M, \gamma)$, and let $S_2$ be a surface obtained from $S$ by a sequence of $n$ Type 5 modifications. Let $S_3$ be obtained from $S_2$ by slicing under a disc of contact. Then in fact, $S_3$ is obtained from $S$ by a sequence of Type 4 modifications, and then at most $n$ Type 5 modifications.

**Proof.** This is by induction on $n$. For $n = 0$, the lemma is trivial. Therefore, assume that the lemma is true for less than $n$ Type 5 modifications, and let $S_1$ be the surface obtained from $S$ by the first $(n - 1)$ Type 5 modifications. Let $\alpha$ be the tubing arc for $S_2$. If the disc of contact $D$ is disjoint from $\alpha$, then we may slice under $D$ before doing the Type 5 move, and so the lemma is true.
by induction. If $D$ is not disjoint from $\alpha$, then $D_1 \cup D_2 = D - (\alpha \times (-1,1))$ is two discs of contact for $S_1$, and $S_3$ is obtained by slicing under both $D_1$ and $D_2$. Applying the inductive hypothesis twice proves the lemma. □

**Lemma 9.14.** Let $S$ be a surface in $(M,\gamma)$. Let $S_2$ be a surface obtained from $S$ by a sequence of Type 5 modifications, and let $S_3$ be a surface obtained from $S_2$ by a Type 6 modification. Then $S_3$ is obtained from $S$ by at most one Type 6 modification, then perhaps some Type 5 modifications.

**Proof.** We will prove this by induction on the number $n$ of Type 5 modifications from $S$ to $S_2$. For $n = 0$, the statement of the lemma is empty. So, assume that the statement is true for sequences of length at most $(n - 1)$. Let $S_1$ be the surface obtained from $S$ by the first $(n - 1)$ Type 5 modifications. The surface $S_1$ is obtained from $S_2$ by tubing along an arc $\alpha$. Let $F_1$ be the component of $S_2$ which we remove in the Type 6 modification. If neither component of $\partial \alpha$ lies in $F_1$, then we may obtain $S_3$ from $S_1$ by performing the Type 6 modification, then the Type 5 modification. The lemma is then true by induction. If both components of $\partial \alpha$ lie in $F_1$, then a single Type 6 modification takes $S_1$ to $S_3$, and again the lemma is true by induction. If a single component of $\partial \alpha$ lies in $F_1$, then we find a (possibly empty) collection of properly embedded arcs in $F_1$ which cut it to a disc. These arcs define Type 5 moves which can be applied to $S_1$, at the end of which we obtain a surface ambient isotopic to $S_3$. Hence, in this case also, the lemma is true. □

**Lemma 9.15.** Let $S$ be a surface in $(M,\gamma)$ which is not mountainous. Let $S'$ be obtained from $S$ by a sequence of Type 6 modifications, then by slicing under a disc of contact $D$. Then $S'$ is obtained from $S$ by a Type 4 modification, then some Type 6 modifications.

**Proof.** Suppose that some component $F_1$ of $S$ has $\partial F_1$ lying in $\text{int}(D)$. Then it must be removed by some Type 6 modification. In particular, it must be parallel to a subsurface $F_2$ of $D$. The outermost component of $\partial F_2$ in $D$ is a component of $\partial S$ which makes $S$ mountainous, contrary to assumption. Hence, each component of $S$ is disjoint from $\text{int}(D)$, and we may therefore slice under $D$ before performing the Type 6 modifications. □

The above lemmas all give the following proposition:

**Proposition 9.16.** Let $S$ be a taut surface in the taut sutured manifold $(M,\gamma)$, with $S$ having essential intersection with $R_\pm$. Let $S'$ be obtained from $S$ firstly by a (possibly empty) sequence of Modifications 1, 2, 3 and 4, then by a (possibly empty) sequence of Modifications 3, 4, 5 and 6. Then $S'$ is in fact obtained from $S$ by a (possibly empty) sequence of Modifications 1, 2 and 3, then possibly by some Type 6 modifications, then possibly some Type 3 and 5 modifications. In particular, no Type 4 modifications are needed.
Proof. Consider the sequence of numbers from 1 to 6 which are the type of each modification. Ignore repetitions; for example, if we perform two Type 3 modifications in a row, then only write down one 3. Lemma 9.13 implies that if we write down 54, we may replace this with 45 or 4. Lemma 9.14 implies that if we write down 56, we may instead write down 65, 5, 6 or nothing. Also, we may replace 34 with 43, since each slide across $\gamma$ creates a component of $\partial S$ touching $\gamma$, whereas each slice under a disc of contact deals with a component of $\partial S$ disjoint from $\gamma$. Hence, we can perform the Type 4 modifications before the Type 3 modifications. Similarly, we can replace 36 with 63. Hence, in the sequence of 3, 4, 5 and 6 modifications, we can arrange to do all the Type 5 and 3 modifications last. Corollary 9.8 asserts that we may replace the initial sequence of 1, 2, 3 and 4 with just a sequence of 1, 2 and 3. Let $S_1$ be the surface obtained from $S$ after this initial sequence. Then, the sequence of numbers after $S_1$ is a (possibly empty) sequence of 4 and 6 (starting with 6), and then possibly a sequence of 5 and 3. If the sequence of 6 and 4 is empty or a single 6, the proposition is proved. Otherwise, the sequence of 6 and 4 starts with 64. By Lemma 9.5, $S_1$ is not mountainous and so, by Lemma 9.15, we may replace the 64 with 46. We may include the 4 in the initial sequence of 1, 2 and 3. Proceeding in this way, we prove the proposition. □

We have so far modified $S$ using Modifications 1, 2, 3 and 4, resulting in a surface satisfying Conditions 1, 2 and 3. We are now going to make some further alterations, using Modifications 3, 4, 5 and 6, resulting in a surface $S'$ which also satisfies Conditions 4 and 5. The point behind the above proposition is that we can in fact obtain $S'$ from $S$ without slicing under discs of contact.

Condition 4. Each curve of $S \cap \partial H^0$ meets any component of $R_{\pm} \cap \partial H^0$ in at most one arc.

Suppose that, on the contrary, there are two arcs $\alpha_1$ and $\alpha_2$ of $S \cap R_{\pm} \cap \partial H^0$ properly embedded in a component of $R_{\pm} \cap \partial H^0$, such that $\alpha_1$ and $\alpha_2$ belong to the same component of $S \cap \partial H^0$. We may assume that there is an arc $\beta$ in $R_{\pm} \cap \partial H^0$ with one endpoint in $\alpha_1$, one endpoint in $\alpha_2$, and the remainder of $\beta$ disjoint from $S$. There is also a disc $D$ in $H^0$ with $\partial D$ containing $\beta$, and $D \cap S = \text{cl}(\partial D - \beta)$.

Suppose first that $\beta$ is a tubing arc and hence that $D$ is disjoint from $\gamma_S$. Since $R_{\pm}(M_S)$ is incompressible, $\partial D$ bounds a disc $D'$ in $R_{\pm}(M_S)$. If $\partial S \cap D'$ is a single arc, then we may ambient isotope $S \cap D'$ onto $D$. It is straightforward to verify that the resulting surface $S'$ is standard and still satisfies Conditions 1-3. This procedure does not increase $|S \cap H^2|$ and it decreases $|\partial S \cap H^1|$. Hence, it decreases the complexity of $S$.

We must deal with the case where $\partial S \cap D'$ contains some circles. Pick one $C$ innermost in $D'$, bounding a disc $E$ in $D'$. Then $E$ is either a disc
of contact or a disc component of \( S \). In the former case, we slice under the disc of contact. We now give a procedure for dealing with the latter case. The curve \( C \) bounds a disc \( E' \) in \( \mathcal{R}_\pm(M) \), and \( E \cup E' \) bounds a ball \( B \) in \( M \). Pick a curve \( C' \) of \( E' \cap \partial S \) innermost in \( E' \). Then we may apply either Modification 4 or Modification 6 to the component of \( S \) containing \( C' \). In this way, we eventually remove all components of \( S \) lying in \( B \). We can then apply Modification 6 to \( E \). Continuing in this fashion, we eventually remove all circles of \( \partial S \cap D' \). Then we may ambient isotope \( S \cap D' \) onto \( D \).

Suppose now that \( \beta \) is not a tubing arc, in which case the disc \( D \) is a product disc in \( M \). Let \( D \times \{0\} \) be embedded in \( H \), so that

- \( D \times \{0\} = D \),
- \( (D \times [-1,1]) \cap \gamma = \emptyset \),
- \( (D \times [-1,1]) \cap \mathcal{R}_\pm(M) = \beta \times [-1,1] \),
- \( (D \times [-1,1]) \cap S = \text{cl}(\partial D - \beta) \times [-1,1] \).

Let \( S' \) be \( S - (\text{cl}(\partial D - \beta) \times (-1,1)) \cup (D \times \{-1,1\}) \). Let \( \{\ast\} \) be a point in \( \beta - \partial \beta \). Then \( S \) is obtained from \( S' \) by tubing along the arc \( \{\ast\} \times [-1,1] \). Hence, \( S' \) is obtained from \( S \) by a Type 5 modification. It is straightforward to check that \( S' \) still satisfies Conditions 1, 2 and 3. This procedure leaves the complexity of \( S \) unchanged. It also creates two discs of \( S' \cap H \) from one disc of \( S \cap H \). Each of the new discs either touches \( \gamma \) or touches \( \mathcal{H} \). But \( S \cap \gamma = S' \cap \gamma \) and \( S \cap \mathcal{H} = S' \cap \mathcal{H} \). Hence, eventually, this process terminates.

**Condition 5.** If \( \alpha \) is an arc of \( S \cap \mathcal{F} \) with both endpoints in \( \mathcal{R}_\pm \), then each of the two arcs in \( \partial \mathcal{F} \) joining \( \partial \alpha \) must either touch \( \partial \mathcal{F} \) or hit \( \gamma_S \) more than twice.

\[ \text{Figure 9.17.} \]

Let \( \alpha \) be such an arc. Suppose that there is an arc \( \beta \) in \( \partial \mathcal{F} \) joining the endpoints of \( \alpha \), such that \( \beta \) is disjoint from \( \mathcal{F} \) and touches \( \gamma_S \) at most twice. Let \( D \) be the disc of \( \mathcal{F} \) containing \( \alpha \). Then, since Condition 2 holds, we may take \( \alpha \) to be extrememost in \( D \), separating off a disc \( D' \) from \( D \) with \( D' \cap S = \alpha \). Then \( D' \) is a disc properly embedded in \( M_S \). Hence, \( \partial D' \) hits
\(\gamma_S\) either twice or not at all. If \(\partial D'\) is disjoint from \(\gamma_S\), then \(\beta\) is a tubing arc, contrary to Condition 2.

Suppose therefore that \(\partial D'\) hits \(\gamma_S\) at precisely two points. If both of these points lie at the endpoints of \(\alpha\), then we will apply Modification 5. If \(H_1 = D \times D^1\) is the 1-handle containing \(\alpha\), then we will take the tubing arc to be a slight extension of \(\{\ast\} \times D^1\), where \(\{\ast\}\) is a point in \(\beta\). The result of this Type 5 modification is to leave \(|S \cap \mathcal{H}^2|\) unchanged and to reduce \(|\partial S \cap \mathcal{H}^1|\) by 2. Hence, it reduces the complexity of \(S\).

If there is a point \(P\) of \(\gamma_S \cap \partial D'\) not lying at an endpoint of \(\alpha\), then \(P\) lies on \(\gamma\), and we can perform a Type 3 modification sliding \(\alpha \times D^1\) across \(P \times D^1\). This slide leaves the complexity of \(S\) unchanged. After possibly performing this operation once more, we end with a situation where both points of \(\gamma_S \cap \partial D'\) lie at the endpoints of \(\alpha\). Hence, we may apply Modification 5 to reduce the complexity of \(S\). It is clear that, in the above procedure, we have not violated Conditions 1-3.


The aim of this section is to complete the proof of Proposition 6.4. Recall that we are given a taut decomposition \((M, \gamma) \xrightarrow{S} (M_S, \gamma_S)\), where \(S\) has essential intersection with \(\mathcal{R}_\pm(M)\). Also, \((M, \gamma)\) is equipped with a handle structure \(\mathcal{H}\), for which each 0-handle of \(F = F(\mathcal{H})\) has positive index.

In the previous section, we performed a sequence of alterations to \(S\), resulting in a new standard surface (called \(S'\), say) satisfying Conditions 1-5. Let \((M, \gamma) \xrightarrow{S'} (M', \gamma')\) be the decomposition along \(S'\) and let \(\mathcal{H}'\) be the induced handle structure on \(M'\). Note that \([S, \partial S] = [S', \partial S'] \in H_2(M, \partial M)\).

Proposition 9.16 asserted that it sufficed to perform a sequence of Modifications 1, 2 and 3, then some Type 6 modifications, then some Type 3 and 5 modifications. If a Type 1 or 2 modification to \(S\) results in a surface \(S_1\), then there is a pull-back \((M_S, \gamma_S) \xleftarrow{P} (M_{S_1}, \gamma_{S_1})\), where \(P\) is a product disc or incompressible product annulus. Hence, the sequence of Modifications 1, 2, 3 and 6 gives rise to the sequence of pull-backs

\[(M, \gamma) = (\hat{M}_1, \hat{\gamma}_1) \xleftarrow{P_1} \cdots \xleftarrow{P_{r-1}} (\hat{M}_r, \hat{\gamma}_r)\]

which was mentioned in Proposition 6.4. Then Modifications 3 and 5 give the sequence of decompositions

\[(\hat{M}_r, \hat{\gamma}_r) \xleftarrow{P_r} \cdots \xleftarrow{P_{m-1}} (\hat{M}_m, \hat{\gamma}_m) = (M', \gamma').\]

In the case of a Type 5 modification, the relevant decomposing surface is a product disc.

We just have to check that Conditions (i)-(v) of 6.4 hold if \(S'\) satisfies Conditions 1-5. Since \(S'\) is standard, (i) is trivially true. We claim that
Conditions 1, 2 and 4 guarantee (iii). Conditions 1 and 4 ensure that there is only a finite number of possibilities for each curve of $S' \cap \partial \mathcal{H}^0$, up to ambient isotopy which keeps $\gamma$ and each handle invariant.

**Lemma 10.1.** Let $C$ and $C'$ be two disjoint simple closed curves of $S' \cap \partial \mathcal{H}^0$ where $S'$ satisfies Conditions 1, 2 and 4. If there is an isotopy of $\partial \mathcal{H}^0$ which leaves $\gamma$ and each handle of $\mathcal{F}$ invariant and which takes $C$ onto $C'$, then $C$ and $C'$ are parallel in a way which respects $\gamma$ and the handle structure on $\mathcal{F}$.

**Proof.** Let $\alpha$ be an arc of $C \cap \mathcal{F}^0$, $C \cap \mathcal{F}^1$ or $C \cap \mathcal{R}_\pm$, and let $\alpha'$ be the image of $\alpha$ after the isotopy taking $C$ to $C'$. It suffices to show that no arc of $C$ or $C'$ lies between $\alpha$ and $\alpha'$. Suppose that there is such an arc. If $\alpha$ lies in $\mathcal{F}^1$, then this means that Condition 1 is violated. If $\alpha$ lies in $\mathcal{R}_\pm$, then Condition 4 is violated. If $\alpha$ lies in $\mathcal{F}^0$, then by Condition 4, $C$ must intersect $\mathcal{F}^0$ in two arcs which are joined by two arcs in $\mathcal{R}_\pm$. If $V$ is the 0-handle of $\mathcal{F}$ containing $\alpha \cup \alpha'$, then some sub-arc of $\partial V$ is a tubing arc for $S'$, contrary to Condition 2. □

Hence, there is only a finite number of possible arrangements for $S' \cap \partial \mathcal{H}^0$, up to possibly taking multiple parallel copies of each curve and performing an ambient isotopy which leaves $H_0$, $H_0 \cap \mathcal{F}$ and $H_0 \cap \gamma$ invariant. Consider therefore a collection $C$ of $n$ parallel curves of $S' \cap \partial \mathcal{H}^0$, the parallelity regions respecting $\mathcal{F}$ and $\gamma$. Let $H'_1, \ldots, H'_{n-1}$ be the associated 0-handles of $\mathcal{H}'$ lying between them. There are three possibilities: Either each curve of $C$ lies entirely in $\mathcal{F}$, or each curve of $C$ is disjoint from $\mathcal{F}$, or each curve of $C$ hits $\partial \mathcal{F}$. In the first case, $\mathcal{F}' \cap H'_i$ is an annulus disjoint from $\gamma'$ for each $i$, and so none of the $H'_i$ lie in $\mathcal{I} \mathcal{H}^0(M')$, and hence can be ignored. Similarly, in the second case, none of the $H'_i$ lie in $\mathcal{I} \mathcal{H}^0(M')$. In the final case, we claim that at most one $H'_i$ lies in $\mathcal{I} \mathcal{H}^0(M')$. For if two adjacent curves of $C$ are coherently oriented, then $\mathcal{F}'$ intersects the 0-handle between them in a collection of product discs. Hence, this 0-handle does not lie in $\mathcal{I} \mathcal{H}^0(M')$. If two adjacent curves of $C$ are incoherently oriented (say they point towards each other), then the arcs of $\partial \mathcal{F}$ lying between them must all point out of $\mathcal{F}$. Otherwise, Condition 2 is violated. Hence, at most one pair of adjacent curves of $C$ can be incoherently oriented. In particular, at most one $H'_i$ can lie in $\mathcal{I} \mathcal{H}^0(M')$. Therefore, (iii) of 6.4 is established.

We will now focus on a component $F$ of $\mathcal{F}$, and will compare its complexity with the complexity of $F' = F \cap \mathcal{F}'$.

**Lemma 10.2.** Let $S'$ be a standard surface satisfying Condition 1. Then no simple closed curve of $S' \cap F$ bounds a disc in $F$.

**Proof.** We may pick such a simple closed curve to be innermost in $F$, bounding a disc $D$, which inherits a handle structure from $F$. Since $D$ is a disc,
there is a 0-handle of $D$ with valence at most one. If this 0-handle has valence zero, then $S'$ is not standard. If this 0-handle has valence one, then Condition 1 is violated. □

The following corollary of Lemma 10.2 is a simple property of planar surfaces:

**Corollary 10.3.** Let $S'$ be a standard surface satisfying Condition 1. Then one of the following holds:

- Each component $X$ of $F'$ has $|\partial X| < |\partial F|$, or
- $F'$ is obtained from $F$ by cutting along arcs and circles which are parallel to arcs and circles in $\partial F$.

Condition 1 also has the following implication:

**Lemma 10.4.** Let $S'$ be a standard surface satisfying Condition 1. Then any component of $F'$ meets any 1-handle of $F$ in at most one disc.

**Proof.** Suppose, on the contrary, that there is a component $X$ of $F'$ meeting a 1-handle $D$ of $F$ in more than one disc. Let $D_1$ and $D_2$ be two discs of $X \cap D$, and let $\alpha_1$ be an arc in $X$ joining $D_1$ to $D_2$. Let $\alpha_2$ be an arc in $D$ transverse to $S' \cap D$, joining the endpoints of $\alpha_1$, in such a way that $\alpha_1 \cup \alpha_2$ forms a simple closed curve, which bounds a disc $D'$ in $\partial H^0(M)$. There exists at least one circle $C$ of $S' \cap \partial H^0$ entering $D'$ through $\alpha_2$. This arc cannot leave $D'$ through $\alpha_1$ and so must leave $D'$ through $\alpha_2$. Hence, $C$ violates Condition 1. □

**Lemma 10.5.** Let $S'$ be a standard surface satisfying Conditions 1, 2 and 3. Let $D$ be a component of $F'$ with a negative index 0-handle. Then there is a 1-handle of $F$ which lies entirely in $D$.

**Proof.** The 0-handle $V$ of $D$ must be disjoint from $\gamma'$ and have valence at most one. The boundary of $V$ is divided into $\partial V \cap \partial F$, $\partial V \cap S'$ and at most one arc $V \cap F^1(M')$. If $\alpha$ is an arc of $\partial V \cap \partial F$ with both endpoints lying in $S'$, then $\alpha$ is a tubing arc, contrary to Condition 2.

Suppose first that $V$ has zero valence. Then, $\partial V$ is divided into $\partial V \cap \partial F$ and $\partial V \cap S'$. However, we cannot have an arc of $\partial V \cap \partial F$, since its endpoints would lie in $S'$ and so would be a tubing arc. Hence, $\partial V$ lies wholly in $S'$. But this violates the assumption that $S'$ is standard, which is a contradiction.

Now suppose that $V$ has valence one, with a single 1-handle $E$ of $F'$ attached to it. Let $\beta_1$ and $\beta_2$ be the two arcs of $\partial E \cap \partial D$. Then each $\beta_i$ originally came from $\partial F$ or from $S' \cap F$.

If both $\beta_1$ and $\beta_2$ lie inside $S' \cap F$, then the arc $\partial V \cap \partial E$ also lies inside $S' \cap F$, for otherwise there would be an arc of $\partial V \cap \partial F$ with endpoints in $S'$, violating Condition 2. But if $\partial V \cap \partial E$ lies wholly in $S'$, then Condition 1 is violated. Similarly, if $\beta_1$ lies inside $S' \cap F$ and $\beta_2$ lies inside $\partial F$, then
Condition 3 is violated. If both \( \beta_1 \) and \( \beta_2 \) lie inside \( \partial F \), then \( E \) is the required 1-handle of \( F \) lying solely in \( D \).

An example of a component \( D \) of \( F' \) with \( I(D) < 0 \) is given in Fig. 5.2.

**Proposition 10.6.** Let \( F \) be a component of \( \mathcal{F} \), and let \( F' = F \cap \mathcal{F}' \). Suppose that every 0-handle of \( F \) has positive index. If \( S \) is a standard surface satisfying Conditions 1-5, then \( C_{\mathcal{F}}(F') \leq C_{\mathcal{F}}(F) \). Also, if we have equality, then each component of \( S' \cap F \) is a circle parallel to a component of \( \partial F \) disjoint from \( \gamma \). The parallelity region inherits a handle structure from \( \mathcal{F} \) in which each 0-handle has valence two.

**Proof.** Suppose that \( C_{\mathcal{F}}(F') \geq C_{\mathcal{F}}(F) \). By Lemma 10.4, each component \( Y \) of \( F' \) has \( C_1(Y) \leq C_1(F) \). If some component \( D \) of \( F' \) has negative index, then some 0-handle of \( D \) has negative index and so by Lemma 10.5, all remaining components \( Y \) of \( F' \) have \( C_1(Y) < C_1(F) \). But by definition \( D \) does not contribute to the \( \mathcal{F} \)-complexity of \( F' \). Hence \( C_{\mathcal{F}}(F') < C_{\mathcal{F}}(F) \), which is contrary to assumption. Thus no component of \( F' \) has negative index. Hence, the index of \( F \) is shared among the components of \( F' \). Since \( C_{\mathcal{F}}(F') \geq C_{\mathcal{F}}(F) \), then one component \( X \) of \( F' \) has \( C_1(X) = C_1(F) \) and \( I(X) = I(F) \). All other components \( Y \) of \( F' \) have zero index, and so, by definition, they do not contribute to the \( \mathcal{F} \)-complexity of \( F' \). By Corollary 10.3, \( |\partial X| \leq |\partial F| \). Hence, \( C_{\mathcal{F}}(F') \leq C_{\mathcal{F}}(F) \), and so \( C_{\mathcal{F}}(F') = C_{\mathcal{F}}(F) \).

We now wish to examine further the case when \( C_{\mathcal{F}}(F') = C_{\mathcal{F}}(F) \). Since \( |\partial X| = |\partial F| \), Corollary 10.3 implies that \( F' \) is obtained from \( F \) by cutting along arcs and circles which are parallel to arcs and circles in \( \partial F \). Each component of \( F' - X \) has index zero. If \( V \) is a 0-handle of \( F' \) not lying in \( X \), then \( V \) cannot have negative index. For otherwise, Lemma 10.5 would imply that \( C_1(X) < C_1(F) \). Thus, each 0-handle \( V \) of \( F' - X \) must have zero index.

We will now show that in fact there are no index zero discs of \( F' \). If there is such a disc, then there is an arc of \( F \cap S' \) extrememost in \( F \), parallel to an arc in \( \partial F \) via a parallelity disc \( D \). If this disc \( D \) does not have zero index, then \( D = X \) and hence \( F \) is a disc. In this case, we may find an arc of \( F \cap S' \) which is extrememost in \( F \) and which does separate off an index zero disc. Thus, we may assume that \( D \) has zero index. Let \( V \) be a 0-handle of \( D \) with valence at most one. Since \( V \) has zero index, there are two cases to consider. If \( V \) has valence zero and hits \( \gamma' \) twice, then Condition 5 is violated. If \( V \) has valence one and hits \( \gamma' \) once, then Condition 3 is violated.

Hence, each component of \( F \cap S' \) is a simple closed curve parallel to a curve of \( \partial F \). An extrememost component of \( F \cap S' \) separates off an annulus \( A \). If \( A = X \), then \( F \) is an annulus, in which case we may find an arc of \( F \cap S' \) extrememost in \( F \) which is parallel to a component of \( \partial F \) via a component of \( F' \) other than \( X \). Hence, we may assume \( A \neq X \). Therefore, \( A \) has zero index, and so each 0-handle of \( A \) has valence two and is disjoint.
from $\gamma$. Repeating this argument for each component of $F \cap S'$ proves the proposition. \hfill $\square$

The following verifies (ii), (iv) and (v) of Proposition 6.4 and completes the proof of that proposition and hence of Theorems 1.4, 1.5 and 1.6:

**Proposition 10.7.** Suppose that every 0-handle of $\mathcal{F}$ has positive index. Suppose also that $H_0 \cap (\mathcal{F}(M) \cup \gamma)$ is connected for each 0-handle $H_0$ of $\mathcal{H}(M)$. Let $S'$ be a standard surface satisfying Conditions 1-5, with $[S', \partial S'] \neq 0 \in H_2(M, \partial M)$. Then $C(H_0 \cap \mathcal{H}') \leq C(H_0)$ for each 0-handle $H_0$ of $\mathcal{H}$. Also, this inequality is strict for some 0-handle $H_0$. If this inequality is an equality for some 0-handle $H_0$, then $H_0 \cap \mathcal{H}'$ is obtained from $H_0$ by a trivial modification.

**Proof.** Suppose that $C(H_0 \cap \mathcal{H}') \geq C(H_0)$ for some 0-handle $H_0$ of $\mathcal{H}$. Then $C_{\mathcal{F}}(H_0 \cap \mathcal{H}') \geq C_{\mathcal{F}}(H_0)$. But, by Proposition 10.6, each component $F$ of $\mathcal{F}$ has $C_{\mathcal{F}}(F \cap \mathcal{F}') \leq C_{\mathcal{F}}(F)$. Hence $C_{\mathcal{F}}(H_0 \cap \mathcal{H}') \leq C_{\mathcal{F}}(H_0)$. Therefore, for each component $F$ of $\mathcal{F}$ of $\mathcal{H}$, we must have $C_{\mathcal{F}}(F \cap \mathcal{F}') = C_{\mathcal{F}}(F)$. Proposition 10.6 then implies that each component of $S' \cap \mathcal{F} \cap H_0$ is a circle parallel to a component of $\partial \mathcal{F}$ disjoint from $\gamma$. Therefore, $n(\mathcal{H}' \cap H_0) \geq n(H_0) = 1$. Hence, $C(\mathcal{H}' \cap H_0) \leq C(H_0)$.

Suppose that this is an equality for some 0-handle $H_0$ of $\mathcal{H}$. Then, as above, this implies that $n(\mathcal{H}' \cap H_0) = n(H_0) = 1$. Also, the argument above gives that each component of $S' \cap \mathcal{F} \cap H_0$ is a circle parallel to a component of $\partial \mathcal{F}$ disjoint from $\gamma$. This component of $\partial \mathcal{F}$ bounds a disc in $\partial H_0$ with interior disjoint from $\mathcal{F}$, since $H_0 \cap (\mathcal{F}(M) \cup \gamma)$ is connected. Hence, $H_0 \cap \mathcal{H}'$ is obtained from $H_0$ by a trivial modification.

Suppose now that $C(H_0 \cap \mathcal{H}') = C(H_0)$ for every 0-handle $H_0$ of $\mathcal{H}$. We aim to achieve a contradiction. Let $C$ be the collection of circles of $S' \cap \mathcal{F}$ extrememost in $\mathcal{F}$. Then, there is a collection of annuli $A$ in $\mathcal{F}$ which is disjoint from $\gamma$ and with $A \cap S' = C$. Let $D$ be the collection of discs of $S' \cap \mathcal{H}_0$ which $C$ bounds. Then $A \cup D$ is a collection of discs properly embedded in $M$ which are parallel to discs in $\mathcal{R}_\pm$ via balls $B_0$. These balls lie in $\mathcal{H}_0$ since $H_0 \cap (\mathcal{F}(M) \cup \gamma)$ is connected for each 0-handle $H_0$ of $\mathcal{H}(M)$.

For each 1-handle $H_1 = D^2 \times [0, 1]$, the discs $D^2 \times \{0\}$ and $D^2 \times \{1\}$ are each divided up by the decomposition along $S'$. For $i = 0$ and 1, all but one 0-handle of $D^2 \times \{i\} - \text{int}(\mathcal{N}(S'))$ has index zero. The remaining component has index equal to the index of $D^2 \times \{i\}$. But the index of $D^2 \times \{i\}$ is positive, since we are assuming that the index of each 0-handle of $\mathcal{F}$ is positive. Hence, the product structure on $H_1$ matches $A \cap (D^2 \times \{0\})$ with $A \cap (D^2 \times \{1\})$. We may therefore unambiguously define $A \cap D^2$. Let $B_1$ be the union (over all 1-handles) of the balls $(A \cap D^2) \times [0, 1]$. Similarly, we may find a collection $B_2$ of components of $\mathcal{H}^2 - \text{int}(\mathcal{N}(S'))$, and such that $B_2 \cap \mathcal{H}^0 = \mathcal{H}^2 \cap A$. 


Then $B_0 \cup B_1 \cup B_2$ is a parallelity region between some closed components of $S'$ and a subsurface of $\mathcal{R}_\pm$. If we remove these components, we may repeat the argument, and show eventually that each component of $S'$ which touches $\mathcal{F}$ is closed and parallel to some component of $\mathcal{R}_\pm$. This does not quite show that $[S', \partial S'] = 0 \in H_2(M, \partial M)$, since there may be components of $S'$ which are disjoint from $\mathcal{F}$. Such a component $X$ lies entirely in a 0-handle $H_0$ of $\mathcal{H}$. But recall from above that $n(\mathcal{H}' \cap H_0) = n(H_0) = 1$. Hence, $\partial X$ cannot separate components of $H_0 \cap \mathcal{F}$. In particular, $X$ is parallel to a disc in $\partial M$.

Therefore, $[S', \partial S'] = 0 \in H_2(M, \partial M)$, contrary to assumption. \hfill \Box

11. The algorithm to construct the tangles.

In this section, we demonstrate how to construct the graphs $G$ required for algorithm of Theorem 1.4 which we outlined in Section 2. Recall that each graph $G$ is embedded in a 3-simplex $\Delta^3$ and comes with a regular neighbourhood $N(G)$ and arcs labelled $\gamma$ and $\tau$ in $\partial N(G)$. Recall that the arcs $\gamma$ form the tangles required for Theorems 1.5 and 1.6.

In line with the rest of this paper, we work with the handle structure $\mathcal{H}$ arising by dualising the given generalised triangulation of $M$. We will focus on a single 0-handle $H_0$ of $\mathcal{H}$. The algorithm starts with the 0-handle $H_0$ and the surface $\mathcal{F}(M) \cap H_0$. This surface is one of finitely many possibilities, but, for the moment, we will assume that $\mathcal{F}(M) \cap H_0$ is as in Figure 11.1. In general, it may be a subsurface of this; we will explain later how to cope with this eventuality.

At each stage $j \in \mathbb{N}$ of the algorithm, we will construct a finite list of possibilities for the following objects lying in $H_0$:

- A subset $\mathcal{H}_j^0$ of $H_0$, which is a union of 3-balls embedded in $H_0$,
- a subsurface $\mathcal{F}(\mathcal{H}_j)$ of $\mathcal{H}_j^0 \cap \mathcal{F}(M)$, and
- arcs $\gamma_j$ properly embedded in $\text{cl}(\partial \mathcal{H}_j^0 - \mathcal{F}(\mathcal{H}_j))$.

\begin{figure}[h]
\centering
\includegraphics[width=0.2\textwidth]{figure11.1.png}
\caption{Figure 11.1.}
\end{figure}
Each component of $\partial \mathcal{H}_j^0 - (\mathcal{F}(\mathcal{H}_j) \cup \gamma_j)$ will have a specified orientation, pointing into or out of $\mathcal{H}_j^0$. When we wish to refer to the above data, we will denote it simply by $\mathcal{H}_j$.

For $j = 1$, we take $\mathcal{H}_1^0 = H_0$, $\mathcal{F}(\mathcal{H}_1) = H_0 \cap \mathcal{F}(\mathcal{M})$, and $\gamma_1 = \emptyset$. We consider all possible orientations for $\partial \mathcal{H}_1^0 - \mathcal{F}(\mathcal{H}_1)$ and so there are 16 possible orientations.

The algorithm constructs the list of possibilities for $\mathcal{H}_{j+1}$ by considering each possibility for $\mathcal{H}_j$ in turn, and performing some modifications to it, which we describe below. These modifications have the property that, if $\mathcal{H}_j$ is some fixed possibility at the $j$th stage, then each possibility for $\mathcal{H}_{j+1}$ to which it gives rise satisfies one of the following:

- $C(\mathcal{H}_{j+1}) < C(\mathcal{H}_j)$ or
- $C(\mathcal{H}_{j+1}) = C(\mathcal{H}_j)$ and $C^+(\mathcal{H}_{j+1}) < C^+(\mathcal{H}_j)$

Thus, by Lemma 5.3, the algorithm will terminate at stage $m$, say. However, we do not know this value of $m$ until we run the algorithm.

It should be clear that this algorithm is modelling within the single 0-handle $H_0$ what is happening in the Proof of Theorem 1.4. Recall that, in that proof, we constructed a sequence of sutured manifolds embedded within $\mathcal{M}$, and examined how each sutured manifold $(\mathcal{M}_i, \gamma_i)$ intersected any given 0-handle $H_0$. However, the intersection $\mathcal{M}_i \cap H_0$ does not necessarily correspond precisely with the $i$th stage of the algorithm we are about to outline. This is because, when passing from a single possibility for $\mathcal{H}_j$ to several possibilities for $\mathcal{H}_{j+1}$, we insist that complexity or extended complexity strictly decreases within our given 0-handle. However, at each stage in the induction of Theorem 1.4, we merely insisted that complexity decreased within some 0-handle (not necessarily the one we are examining). Therefore, in order to determine how the final sutured manifold $\mathcal{M}_n$ lies in $H_0$, we must consider every possibility for $\mathcal{H}_j$, where $1 \leq j \leq m$. Given one such possibility $\mathcal{H}_j$, we construct the graph $G$ by associating a vertex of $G$ with each component of $\mathcal{H}_j^0$; we associate an edge of $G$ with each component of $\mathcal{F}^0(\mathcal{H}_j)$; the curves $\tau$ are specified by $\mathcal{F}^1(\mathcal{H}_j)$; the arcs $\gamma$ are formed by taking all possible subtangles of $\gamma_j$.

We now give the heart of the algorithm, namely the procedure which constructs each possibility for $\mathcal{H}_{j+1}$ arising from a single possibility for $\mathcal{H}_j$. We apply one of the following procedures to $\mathcal{H}_j$ (and we give the points in Sections 7-9 where we applied them):

1. Removal of a component of $\mathcal{H}_j^0$.

This can occur in Procedures 1 and 5 of Section 7. It can also occur in Cases 1, 2, 3 and 4B of Section 8. When a component of $\mathcal{H}_j^0$ is removed, so are the components of $\mathcal{F}(\mathcal{H}_j)$ and arcs of $\gamma_j$ which it contains.
2. Removing some handles of $\mathcal{F}(\mathcal{H}_j)$ disjoint from $\gamma_j$.

In order that the new surface inherits a handle structure, we insist that if a 0-handle of $\mathcal{F}(\mathcal{H}_j)$ is removed, then so are the 1-handles of $\mathcal{F}(\mathcal{H}_j)$ which abut it. Also, we may only perform this procedure if the components of $\partial H^0_j - (\mathcal{F}(\mathcal{H}_j) \cup \gamma_j)$ which touch any removed handle have orientations which agree. This operation can occur in Procedures 1, 2, 5 and 6 of Section 7, and Case 4B of Section 8.

Note that, in general, $\mathcal{F}(M) \cap H_0$ is obtained from the surface in Figure 11.1 by removing some handles. Therefore, by applying this procedure at the first stage $j = 1$, we can incorporate all possibilities for $\mathcal{F}(M) \cap H_0$ into this algorithm.

3. Replacing handles of $\mathcal{F}(\mathcal{H}_j)$ with a sub-arc of $\gamma_{j+1}$.

Here, we may replace a 1-handle of $\mathcal{F}(\mathcal{H}_j)$ with an arc of $\gamma_{j+1}$, providing that the components of $\partial H^0_j - (\mathcal{F}(\mathcal{H}_j) \cup \gamma_j)$ which touch this handle have orientations which disagree. We may also remove a 0-handle of $\mathcal{F}(\mathcal{H}_j)$ which has valence one and which intersects $\gamma_j$ in a single point, providing that we also remove the 1-handle of $\mathcal{F}(\mathcal{H}_j)$ which it abuts and we then replace these handles with a sub-arc of $\gamma_{j+1}$. This occurs in Procedure 3 of Section 7 and in Case 3 of Section 8.

4. Removal of a product disc component of $\mathcal{F}(\mathcal{H}_j)$.

If $F$ is a disc component of $\mathcal{F}(\mathcal{H}_j)$ intersecting $\gamma_j$ twice, we may replace $F$ with an arc of $\gamma_{j+1}$ joining the two points of $F \cap \gamma_j$. This occurs in Procedure 4 of Section 7.

5. Removal of a valence two 0-handle of $\mathcal{F}(\mathcal{H}_j)$.

If $V$ is a 0-handle of $\mathcal{F}(\mathcal{H}_j)$ which is disjoint from $\gamma_j$ and which abuts two distinct 1-handles of $\mathcal{F}(\mathcal{H}_j)$, then we may combine $V$ and the two 1-handles into a single 1-handle of $\mathcal{F}(\mathcal{H}_{j+1})$. This occurs in Case 1 of Section 8.

6. Decomposition along a surface.

This step models the sutured manifold decomposition outlined in Section 9. We only perform this operation providing each 0-handle of $\mathcal{F}(\mathcal{H}_j)$ has positive index and providing $H_0 \cap (\mathcal{F}(\mathcal{H}_j) \cup \gamma_j)$ is connected for each 0-handle $H_0$ of $\mathcal{H}_j$. We construct all possible oriented curves $C$ which satisfy Conditions 1-5 of Section 9 (viewing $C$ as a possibility for $S' \cap \partial H^0$). There is only a finite number of possibilities ($C_1, \ldots, C_t$, say) for $C$. We then let $C'$ be a collection of disjoint simple closed curves in $\partial H^0_j$, each curve being a copy of one of the $C_i$'s, and with no two components of $C'$ representing the same $C_i$ (although, two components of $C'$ may be the same underlying curve, but have opposite orientations). We insist that $C'$ also satisfies Condition 2 of Section 9. We then extend $C'$ to a collection of disjoint discs properly embedded in $H^0_j$. We then decompose $H^0_j$ along these discs, creating a new
collection of 0-handles $\mathcal{H}_{j+1}^0$, which naturally inherit $\mathcal{F}(\mathcal{H}_j)$ and sutures $\gamma_{j+1}$. By the argument of Propositions 10.6 and 10.7, $C(\mathcal{H}_{j+1}) < C(\mathcal{H}_j)$, unless each component of $C'$ is a curve lying entirely in $\mathcal{F}(\mathcal{H}_j)$ parallel to some component of $\partial\mathcal{F}(\mathcal{H}_j)$ disjoint from $\gamma_j$, the parallelity region respecting the handle structure of $\mathcal{F}(\mathcal{H}_j)$. In this case, the modification is trivial. We therefore do not include this case as a possibility for $\mathcal{H}_{j+1}$. However, the modification may alter the orientations of some components of $\partial\mathcal{H}_0^0 - \mathcal{F}(\mathcal{H}_j)$ disjoint from $\gamma_j$. We therefore have to consider all possible orientations for these components as giving distinct possibilities for $\mathcal{H}_j$.

7. Amalgam removal.

We only perform this operation when each component of $\mathcal{F}(\mathcal{H}_j)$ has positive index and each 0-handle of $\mathcal{F}(\mathcal{H}_j)$ with nonpositive index has valence two and is disjoint from $\gamma_j$. Suppose that $D$ is a union of handles of $\mathcal{F}(\mathcal{H}_j)$ which forms a disc disjoint from $\gamma$. Suppose also that if $D$ has any 0-handles, then each such 0-handle abuts precisely two 1-handles of $\mathcal{F}(\mathcal{H}_j)$, both of which lie in $D$. Suppose also that the two components of $\partial\mathcal{H}_0^0 - (\mathcal{F}(\mathcal{H}_j) \cup \gamma_j)$ which touch $D$ have the same orientation. We take one or two copies of $\partial D$ and move them a little, creating a curve $C_1$ (and possibly $C_2$) which intersect $\mathcal{F}(\mathcal{H}_j)$ in a collection of arcs lying in 0-handles of $\mathcal{F}(\mathcal{H}_j)$. Extend each $C_i$ to a disc $D_i$ properly embedded in $\mathcal{H}_j^0$. If we have both $D_1$ and $D_2$, we orient them inconsistently, in a way which gives the parallelity region between them some sutures. If we are just dealing with $D_1$, we consider both possible orientations. We then decompose $\mathcal{H}_j^0$ along $D_1$ (and possibly $D_2$), providing this is not a trivial modification, as outlined in Operation 6 above. This occurs in Case 5 of Section 8. Note that we cannot necessarily include this case here in Operation 6, since the curves $C_1$ and $C_2$ might fail Conditions 2, 4 or 5 of Section 9.

The Figure 11.2 gives a concrete example of some of the above operations. It is clear that these procedures may implemented algorithmically, although they may pose some challenges for a computer programmer.

12. Exceptional and norm-exceptional surgeries with $\Delta(\sigma, \mu) = 1$.

We now give examples which demonstrate that the restriction on $\Delta(\sigma, \mu)$ in Theorems 1.4, 1.5 and 1.6 is necessary. We give a method of constructing in a 3-manifold $M$ (satisfying certain conditions) an infinite number of surgery curves $K$ with exceptional or norm-exceptional surgery slopes $\sigma$ satisfying $\Delta(\sigma, \mu) = 1$, where $\mu$ is the meridian slope on $\partial N(K)$.

Let $M$ be a compact orientable 3-manifold with $\partial M$ a (possibly empty) union of tori. Suppose also that $M$ is irreducible, atoroidal and has incompressible boundary. Let $S$ be a connected oriented surface properly
embedded in $M$ with $[S, \partial S] \neq 0 \in H_2(M, \partial M)$, and so that $S$ is incompressible and norm-minimising in its homology class. Then $S$ is neither a sphere nor a disc. Let $K$ be any essential simple closed curve on $S$ disjoint from $\partial S$. Let $\sigma$ be the slope of the curves $\partial N(K) \cap S$, which is known as ‘surface framing’.

**Proposition 12.1.** The slope $\sigma$ is exceptional or norm-exceptional.

*Proof.* The surface $S$ determines a class $z \in H_2(M - \text{int}(N(K)), \partial M)$ as follows: The two curves $S \cap \partial N(K)$ divide $\partial N(K)$ into two annuli. Attach either of these annuli to $S - \text{int}(N(K))$ and let $S'$ be the resulting surface. Then $z = [S', \partial S'] \in H_2(M - \text{int}(N(K)), \partial M)$ is independent of the choice of annulus. In fact, $S'$ is norm-minimising in its class in $H_2(M - \text{int}(N(K)), \partial M)$, since $\chi_{-}(S') = \chi_{-}(S) = x([S, \partial S]) \leq x([S', \partial S'])$. Let $z_\sigma \in H_2(M_K(\sigma), \partial M_K(\sigma))$ be the image of $z$ under the map induced by inclusion.
We may construct a surface $S_\sigma$ in $M_K(\sigma)$ by starting with $S - \text{int}(N(K))$ and attaching a disc to each curve of $S \cap \partial N(K)$, the discs being meridian discs in the surgery solid torus. Then $[S_\sigma, \partial S_\sigma] = z_\sigma \in H_2(M_K(\sigma), \partial M_K(\sigma))$. Also, $-\chi(S_\sigma) = -\chi(S') - 2$. Since we assumed that $S$ was connected, there are two possibilities:

(i) $\chi_-(S_\sigma) < \chi_-(S')$, or

(ii) $\chi_-(S_\sigma) = \chi_-(S') = 0$.

In Case (i), $x(z_\sigma) < x(z)$, and therefore $K$ and $\sigma$ are norm-exceptional. In Case (ii), $S_\sigma$ is a non-separating sphere or two non-separating discs in $M_K(\sigma)$. If $S_\sigma$ is a sphere, then $M_K(\sigma)$ is reducible. If $S_\sigma$ is two discs, then $M_K(\sigma)$ has compressible boundary, which implies that either $M_K(\sigma)$ is a solid torus or it is reducible. Thus, in this case, $K$ and $\sigma$ are exceptional. \quad \Box

We now show that one may find an infinite number of such knots $K$ on a given $S$ (satisfying some conditions), such that no two knots in this collection are ambient isotopic to each other in $M$.

**Proposition 12.2.** Let $M$ be a compact 3-manifold with $\partial M$ a (possibly empty) union of tori and with $H_1(M)$ torsion free. Let $S$ be a compact connected oriented surface properly embedded in $M$ which has positive genus and which is norm-minimising in its class in $H_2(M, \partial M)$. Then, we may find an infinite collection of knots, each essential curves on $S$, no two of which are ambient isotopic in $M$.

**Proof.** We may find two simple closed curves $C_1$ and $C_2$ on $S$ which intersect each other precisely once. If one of these curves has infinite order in $H_1(M)$ ($C_1$, say), then, for any integer $n$, consider the curve $nC_1 + C_2$, which is constructed by taking $n$ (coherently oriented) parallel copies of $C_1$, together with $C_2$ and smoothing off the double-points. This is the required collection of knots on $S$.

Suppose therefore that $C_1$ and $C_2$ have finite order in $H_1(M)$. Since $H_1(M)$ is torsion free, this implies that $C_1$ and $C_2$ are homologically trivial. We will construct our collection of knots by analysing the ‘Seifert form’ on $S$. Given two disjoint homologically trivial closed curves $\alpha_1$ and $\alpha_2$ in $M$, define their linking number $\text{lk}(\alpha_1, \alpha_2)$ to be the signed intersection number between $\alpha_2$ and a (not necessarily embedded) Seifert surface for $\alpha_1$. This is independent of the choice of Seifert surface for $\alpha_1$, since any two Seifert surfaces can be glued to form a closed (not necessarily embedded) surface, with which $C_2$ has zero intersection, since it is homologically trivial. Also, it is symmetric: $\text{lk}(\alpha_1, \alpha_2) = \text{lk}(\alpha_2, \alpha_1)$. Given any curve $C$ on $S$, define $C^+$ to be the push-off of $C$ from $S$ in some specified normal direction. Define the framing $fr(C)$ of any curve $C$ on $S$ which is homologically trivial in $M$ as $\text{lk}(C^+, C)$. Now,

$$\text{lk}(C_1^+, C_2) - \text{lk}(C_1, C_2^+) = \pm 1,$$
since $C_1$ and $C_2$ intersect in one point on $S$. This implies that

$$\text{lk}(C_1^+, C_2) + \text{lk}(C_1, C_2^+) \neq 0.$$ 

Let $n_1$ be an arbitrary integer. Then

$$fr(n_1 C_1 + C_2) = \text{lk}((n_1 C_1 + C_2)^+, n_1 C_1 + C_2)
= n_1^2 \text{lk}(C_1^+, C_1) + n_1 (\text{lk}(C_2^+, C_1) + \text{lk}(C_1^+, C_2)) + \text{lk}(C_2^+, C_2)
= n_1^2 k_1 + n_1 k_2 + k_3,$$

for integers $k_1$, $k_2$ and $k_3$, where $k_2 = \text{lk}(C_1, C_2^+) + \text{lk}(C_1^+, C_2) \neq 0$. Hence, $fr(n_1 C_1 + C_2)$ takes infinitely many values.

We now claim that if $C$ and $C'$ on $S$ are two closed curves on $S$ which are homologically trivial in $M$ and freely homotopic in $M$, then $fr(C) = fr(C')$. A free homotopy is realised by a map $f : A \to M$, where $A$ is an annulus and where $f(\partial A) = C \cup C'$. We may ensure that $f$ respects the product structure on $N(S)$ and hence that $f^{-1}(S)$ is $\partial A$, together with some properly embedded simple closed curves. We may also ensure that no region of $A - \text{int}(N(f^{-1}(S)))$ is a disc, and hence that $f^{-1}(S)$ is a collection $\alpha_0, \alpha_1, \ldots, \alpha_n$ of disjoint essential simple closed curves in $A$, where $\partial A = \alpha_0 \cup \alpha_n$. Since the image of the annulus lying between $\alpha_i$ and $\alpha_{i+1}$ is disjoint from $S$, then $fr(\alpha_i) = fr(\alpha_{i+1})$. Therefore, $fr(C) = fr(C')$.

Hence, we have constructed the required infinite collection of knots. \[ \square \]

We now show that we may ensure that each knot $K$ in this infinite collection has $M - \text{int}(N(K))$ irreducible and atoroidal.

**Proposition 12.3.** Let $M$ and $S$ be as above. Then each essential simple closed curve $K$ on $S$ has $M - \text{int}(N(K))$ irreducible. Also, there are (up to ambient isotopy in $M$) at most finitely many knots $K$ on $S$ for which $M - \text{int}(N(K))$ is toroidal.

**Proof.** Let $K$ be an essential simple closed curve on $S$. If $M - \text{int}(N(K))$ contains a reducing sphere, then this bounds a ball in $M$. By assumption, $M$ is irreducible, and so the knot $K$ must lie in this 3-ball, and is therefore homotopically trivial in $M$. However, $K$ is essential on $S$ and $S$ is $\pi_1$-injective, since it is incompressible. This is a contradiction.

Suppose now that $M - \text{int}(N(K))$ is toroidal, and let $T$ be an essential torus in $M - \text{int}(N(K))$. Since $M$ is atoroidal, $T$ either is parallel in $M$ to a component of $\partial M$ or is compressible in $M$. Consider the former case, and let $T^2 \times I$ be the parallellity region between $T$ and a component $T'$ of $\partial M$. We may assume that the intersection $S \cap (T^2 \times I)$ is a collection of discs and annuli, with $K$ being a core of one of these annuli. Hence, $K$ is parallel to a curve on $T'$. It is not hard to show that if $K_1$ and $K_2$ are two curves on $S$ both parallel to curves in $T'$, then either $K_1$ and $K_2$ are ambient isotopic in
$M$ or $S$ contains a component parallel to $T'$. However, $S$ is connected and nontrivial in $H_2(M, \partial M)$, which gives a contradiction.

Hence, we may restrict attention to the case where $T$ is compressible in $M$. This implies that $T$ bounds a solid torus or lies in a 3-ball in $M$. We will show that the latter case cannot arise. For, $S = \text{int}(\mathcal{N}(K))$ and $T$ are both essential in $M - \text{int}(\mathcal{N}(K))$, and so we may isotope these surfaces in $M - \text{int}(\mathcal{N}(K))$, so that no curve of $S \cap T$ bounds a disc in $T$ or $S - \text{int}(\mathcal{N}(K))$. So no curve of $S \cap T$ bounds a disc in $S$. But if $T$ were to lie in a 3-ball in $M$, then each curve of $S \cap T$ would be homotopically trivial in $M$, and hence in $S$. Thus, $S \cap T$ would have to be empty. We now consider the intersection between $S$ and the compression disc for $T$. By an innermost curve argument, we may isotope $S$ and $K$, keeping them disjoint from $T$, so that afterwards they are disjoint from this disc. But then $T$ compresses in $M - \text{int}(\mathcal{N}(K))$, which is a contradiction.

Thus, we may assume that $T$ bounds a solid torus $V$ in $M$. We may also assume that the surface $V \cap S$ is incompressible in $V$ and so is a collection of discs and annuli. The knot $K$ lies on one such annulus $A$. If $A$ has winding number one in $V$, then $K$ is a core of $V$ and so $T$ is parallel to $\partial \mathcal{N}(A)$, contradicting the assumption that $T$ is essential in $M - \text{int}(\mathcal{N}(K))$. If $A$ has winding number greater than one in $V$, then a cabling annulus for $K$ is constructed by gluing $A - \text{int}(\mathcal{N}(K))$ to the closure of one of the components of $T - A$. It is now not hard to show that $K$ is ambient isotopic to the core $K'$ of an annular component $A'$ of $V \cap S$, where $K'$ has a cabling annulus disjoint from $S$. For the purposes of the proof of Proposition 12.3, we may consider the knot $K'$ instead of $K$. Suppose therefore that the cabling annulus for $K$ is disjoint from $S$.

**Claim.** Let $K_1$ and $K_2$ be two essential simple closed curves on $S$, each cabled with cabling annulus disjoint from $S$. Suppose that the cabling annuli both lie on the same side of $S$. Then there is an isotopy of $S$ which takes $K_1$ off $K_2$.

We may take the cabling annulus $A_i$ for $K_i$ to be properly embedded in $M - \text{int}(\mathcal{N}(S))$. Then a component of $M - \text{int}(\mathcal{N}(S \cup A_i))$ is a solid torus $V_i$. A simple examination of the intersection between the annulus $A_1$ and the solid torus $V_2$ establishes that we may isotope $A_1$ off $V_2$ unless the winding number of $A_2$ along is $V_2$ is two. Thus, the claim is proved unless the winding number of $A_i$ in $V_i$ is two, for both $i = 1$ and 2. In this case $V_i$ is an $I$-bundle over a Möbius band, with $V_i \cap \partial \mathcal{N}(S)$ being precisely the $\partial I$-bundle. Therefore, if $K_1$ and $K_2$ cannot be homotoped off each other, $V_1 \cup V_2$ is an $I$-bundle over a connected non-orientable surface $G$ other than a Möbius band. The $I$-bundle over $\partial G$ is a collection of annuli. If any of these annuli are compressible, we may extend the $I$-bundle. Thus, we may construct an $I$-bundle $X$ over a compact connected non-orientable surface
\[ G', \] such that the \( I \)-bundle over \( \partial G' \) is incompressible in \( M \), and so that \( X \cap \partial N(S) \) is the \( \partial I \)-bundle over \( G' \). If \( G' \) is a Möbius band, then there is an isotopy of \( S \) taking \( K_1 \) off \( K_2 \). Suppose therefore that \( G' \) has negative Euler characteristic. Expand the \( I \)-bundle a little, so that the \( \partial I \)-bundle lies in \( S \). If we remove the \( \partial I \)-bundle from \( S \), and attach the \( I \)-bundle over \( \partial G' \), we create a surface \( S' \) with \( [S', \partial S'] = [S, \partial S] \in H_2(M, \partial M) \), and with \( \chi(S') > \chi(S) \). This contradicts the assumption that \( S \) is norm-minimising and incompressible.

There are at most finitely many disjoint essential non-parallel simple closed curves on \( S \). This proves the proposition.

Propositions 12.1, 12.2 and 12.3 give the following result:

**Theorem 12.4.** Let \( M \) be a compact orientable 3-manifold with \( \partial M \) a (possibly empty) union of tori. Suppose that \( M \) is irreducible and atoroidal, and has incompressible boundary. Suppose also that \( H_1(M) \) is torsion free and that some nontrivial element of \( H_2(M, \partial M) \) is represented by a norm-minimising incompressible surface with positive genus. Then (up to ambient isotopy) there is an infinite number of surgery curves \( K \) in \( M \), with exceptional or norm-exceptional surgery slopes \( \sigma \) satisfying \( \Delta(\sigma, \mu) = 1 \), where \( \mu \) is the meridian slope on \( \partial N(K) \). We may ensure that each knot \( K \) in this collection has \( M - \text{int}(N(K)) \) irreducible and atoroidal, and has \( H_2(M - \text{int}(N(K)), \partial M) \neq 0 \).

**References**


Received July 20, 1999.

**Mathematical Institute**
**Oxford University**
**24-29 St Giles’**
**Oxford OX1 3LB**
**United Kingdom**
**E-mail address:** lackenby@maths.ox.ac.uk
NONLINEAR ALGEBRAIC ANALYSIS OF
DELTA SHOCK WAVE SOLUTIONS
TO BURGERS’ EQUATION

J.-A. Marti

By means of three fundamental structures we can define, in a general way, a sheaf \( \mathcal{A} \) of differential algebras containing most of the special cases met in the theory of generalized functions.

A convenient choice of these structures permits us to study Burgers’ equation with \( \delta \)-Dirac measure as initial data, and we can construct a generalized \( \delta \)-shock wave as an approximate solution, self-similar to the initial data.

1. Introduction.

It is not easy to pose and a fortiori to solve the Cauchy problem for Burgers’ equation

\[
\frac{\partial u}{\partial t} + f(u) \frac{\partial u}{\partial x} = 0, \quad t \geq 0
\]

with initial data as irregular as the \( \delta \)-Dirac measure.

When the data are smooth enough or have weak singularities, the problem has been studied in many classical or generalized ways, but when these singularities become stronger, we first have to describe them.

In this paper, we do not consider solutions of entropy type satisfying the Krushkov (or some other) criteria from the general theory of hyperbolic conservation laws (e.g., [7]).

We try here to study traveling wave solutions self-similar to initial data which can be as singular as \( \delta \) or even powers of \( \delta \) which do not exist in classical distribution theory. So, we begin by giving a special family \( \{ \delta_{\varepsilon} \} \) of \( \delta \)-approximations such that for each fixed \( \varepsilon > 0, p > 0 \), Burgers’ equation has an exact weak self-similar solution corresponding to \( \delta_{\varepsilon}^{p} \) as initial data. Then, some models of tsunami or soliton are studied in the same way. However we show that the distribution spaces are not convenient to solve our problem by the help of a limit process.

To get out of this situation we define some other technics of approximation thanks to association processes in \( (\mathcal{C}, \mathcal{E}, \mathcal{P}) \)-algebras (e.g., [11]) which contain most of the special cases met in the literature. Let us give an idea of their construction.
\( \mathbb{K} \) is the real or complex field and \( \Lambda \) a set of indices. \( C \) is the factor ring \( A/I \) where \( I \) is an ideal of \( A \), a given subring of \( \mathbb{K}^\Lambda \). \( (\mathcal{E}, \mathcal{P}) \) is a sheaf of topological \( \mathbb{K} \)-algebras on a topological space \( X \). A sheaf of \( (\mathcal{C}, \mathcal{E}, \mathcal{P}) \)-algebras on \( X \) is a sheaf \( \mathcal{A} = \mathcal{H}/\mathcal{J} \) of factor algebras where \( \mathcal{J} \) is a sheaf of ideals of \( \mathcal{H} \), a subsheaf of \( \mathcal{E}^\Lambda \). The sections of \( \mathcal{H} \) (resp. \( \mathcal{J} \)) have to verify some estimations given by means of \( \mathcal{P} \) and \( A \) (resp. \( I \)). In such algebras we have good tools to pose and solve many non-linear differential problems with irregular data.

The sketch of the procedure is the following: We begin by choosing \( \mathcal{E} \) and \( \mathcal{P} \) in relationship with the problem. Here, for \( \Omega = \mathbb{R} \times \mathbb{R}_+ \), we define \( \mathcal{E}(\Omega) \) as \( C^\infty(\Omega) \) with its usual topology. Then, information about data and equation are taken into account in the construction of \( \mathcal{C} \) and finally we can construct a \( (\mathcal{C}, \mathcal{E}, \mathcal{P}) \)-algebra adapted to our problem. It is solved by means of a two-parametric family of special mollifiers with an approximation depending itself upon data and equation peculiarities.

The same methods and technics solve the problem for our models of tsunami or soliton.

2. The weak form of Burgers’ problem.

The Cauchy problem for Burgers’ equation in the following non-conservative form equation:

\[
\frac{\partial u}{\partial t} + f(u) \frac{\partial u}{\partial x} = 0, \quad t \geq 0, \quad u_{|\{t=0\}} = u_0
\]

where \( f(u) = f \circ u \), and \( f \) a function of the real variable, has well-known solutions if \( f \) and \( u_0 \) are smooth enough and then (1) is equivalent to the conservative form

\[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} F(u) = 0, \quad t \geq 0, \quad u_{|\{t=0\}} = u_0
\]

when taking \( F(u) = F \circ u \), with \( F' = f \).

If \( v \) and \( F(v) \) belong to \( L^1_{\text{loc}}(\mathbb{R} \times \mathbb{R}_+) \), we define \( \tilde{v} \) as

\[
\tilde{v}(x, t) = v(x, t) \quad \text{when } t \geq 0 \quad \text{and} \quad \tilde{v}(x, t) = 0 \quad \text{when } t < 0
\]

and then it is clear that \( \tilde{v} \) belongs to \( L^1_{\text{loc}}(\mathbb{R}^2) \cap D'_t(\mathbb{R}^2) \) and \( F(\tilde{v}) \) belongs to \( L^1_{\text{loc}}(\mathbb{R}^2) \subset D'M(\mathbb{R}^2) \) with

\[
D'_t(\mathbb{R}^2) = \{ T \in D'(\mathbb{R}^2), \text{supp } T \subset \Gamma = \mathbb{R} \times \mathbb{R}_+ \}.
\]

If \( u_0 \) is in \( L^1_{\text{loc}}(\mathbb{R}) \), to say that \( v \) is a weak (distribution) solution of the Cauchy problem (2) means that this problem can be interpreted ([10]) as the equation

\[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} F(u) = u_0 \otimes \delta_t
\]
to be verified by \( \tilde{v} \) in the sense of distribution theory, or explicitly

\[
- \int_0^\infty \left( v \frac{\partial}{\partial t} \psi + \mathcal{F}(v) \frac{\partial}{\partial x} \psi \right) (x,t) dx dt - F(0) \int_0^\infty \frac{\partial}{\partial x} \psi(x,t) dx dt
= \int u_0(x) \psi(0,x) dx, \quad \psi \in C_0^\infty(\mathbb{R}^2).
\]

Then \( u_0 \) will still be called initial data for (3).

If \( u_0 \) has weak singularities, the problem has been studied in many classical or generalized ways ([1], [2], [4], [8], [18], [19] and many others).

For example, if \( \mathcal{F}(u) = \frac{1}{2} u^2 \) and \( u_0 = Y \), where \( Y \) is the Heaviside function, the search of a self-similar solution, that is to say a solution of the form

\[
u(x,t) = Y(t) u_0(x - \varphi(t)), \quad \varphi(0) = 0, \quad \varphi' > 0
\]

leads to the well-known Hugoniot-Renkin condition

\[
\varphi'(t) = \frac{1}{2}.
\]

But now, if \( u_0 = \delta \in \mathcal{D}'(\mathbb{R}) \), the research of a solution in the form (4) has no sense for the Cauchy problem (3).

However, if we choose \( u_0 = Y \), and if we search a self-similar solution \( u \) of (3) on the form (4) with an unknown strictly increasing function \( \varphi \in C^\infty(\mathbb{R}) \) verifying \( \varphi(0) = 0 \), we can see from

\[
F \circ Y = F(0) + [F(1) - F(0)] Y
\]

for any application \( F \) from \( \mathbb{R} \) to \( \mathbb{R} \), that we have also

\[
\mathcal{F}(u) = F(0) + [F(1) - F(0)] u.
\]

Then, with classical technics in distribution theory, it is easy to compute

\[
\frac{\partial u}{\partial t} = -(1_x \otimes \varphi') \delta_{\{x = \varphi(t), t \geq 0\}} + Y_x \otimes \delta_t,
\]

\[
\frac{\partial}{\partial x} \mathcal{F}(u) = [F(1) - F(0)] \delta_{\{x = \varphi(t), t \geq 0\}}.
\]

This leads to the necessary and sufficient condition for the existence and the uniqueness of the required solution on the form

\[
1_x \otimes \varphi' = F(1) - F(0)
\]

that is to say

\[
\varphi(t) = [F(1) - F(0)] t.
\]
2.1. Exact weak solutions of Heaviside type. We can try now to search a self-similar solution of (3) in $\mathcal{D}'_r(\mathbb{R}^2)$ of the form

\begin{equation}
\label{eq:6}
u_{h,l}(x,t) = Y(t) u_{0,h,l}(x - \varphi(t))
\end{equation}

by taking

\[ u_{0,h,l}(x) = h \left[ Y(x + l) - Y(x - l) \right] \]

where $h$ and $l$ are two strictly positive given constants, and then the following holds:

**Proposition 1.** A necessary and sufficient condition to have an unique solution of (3) in $\mathcal{D}'_r(\mathbb{R}^2)$ with the form (6) is

\begin{equation}
\label{eq:7}
\forall t \geq 0, \varphi(t) = F(h) - F(0)
\end{equation}

and then we can explicit the required solution in the form

\begin{equation}
\label{eq:8}
u_{h,l}(x,t) = Y(t) h \left[ Y\left(x + l - \frac{F(h) - F(0)}{h} t\right) - Y\left(x - l - \frac{F(h) - F(0)}{h} t\right) \right].
\end{equation}

**Proof.** We begin to compute

\[ \frac{\partial u_{h,l}}{\partial t} = -h (1_x \otimes \varphi') \left[ \delta_{\{x + t = \varphi(t), t \geq 0\}} - \delta_{\{x - t = \varphi(t), t \geq 0\}} \right] + u_{0,h,l} \otimes \delta_t \]

and we have, according to (5)

\[ \mathfrak{F}(u_{h,l}) = F(0) + [F(h) - F(0)] u_{h,l}. \]

This gives, in distributional sense

\[ \frac{\partial}{\partial x} \mathfrak{F}(u_{h,l}) = [F(h) - F(0)] \left[ \delta_{\{x + t = \varphi(t), t \geq 0\}} - \delta_{\{x - t = \varphi(t), t \geq 0\}} \right] \]

and then (6) is a solution of (3) if and only if

\[ h(1_x \otimes \varphi') = F(h) - F(0). \]

Hence we deduce the equalities (7) and (8) of the proposition. \hfill \Box

**Corollary 2.** Let

\begin{equation}
\label{eq:9}
\delta_\varepsilon(x) = u_{(0, \frac{1 - \varepsilon}{2\varepsilon})}(x) = \frac{1}{2\varepsilon} \left[ Y(x + \varepsilon) - Y(x - \varepsilon) \right],
\end{equation}

\[ \tau_\varepsilon(x) = u_{(0, \frac{1 - \varepsilon}{2\varepsilon})} = \frac{1}{2\varepsilon} \left[ Y(x + 1) - Y(x - 1) \right], \]

\[ \sigma_\varepsilon(x) = u_{(0, \frac{1 - \varepsilon}{2\varepsilon})} = Y(x + \varepsilon) - Y(x - \varepsilon) \]

and $p > 0$ a given real number. Then:

\begin{enumerate}
\item [(i)] Solution (6) of (3) corresponding to $\delta_\varepsilon^p$ as initial data is given by

\[ u_{p,\varepsilon}(x,t) = Y(t) \frac{1}{(2\varepsilon)^{p-1}} \delta_\varepsilon(x - c_{p,\varepsilon} t). \]
\end{enumerate}
(ii) Solution (6) of (3) corresponding to $\tau^p_\varepsilon$ as initial data is given by

$$u_{p,\varepsilon}(x,t) = Y(t) \frac{1}{(2\varepsilon)^{p-1}} \tau^p_\varepsilon(x - c_{p,\varepsilon} t).$$

(iii) Solution (6) of (3) corresponding to $\sigma^p_\varepsilon$ as initial data is given by

$$u_{p,\varepsilon}(x,t) = Y(t) \sigma^p_\varepsilon(x - ct)$$

with

$$c_{p,\varepsilon} = (2\varepsilon)^p \left[ F\left(\frac{1}{(2\varepsilon)^p}\right) - F(0) \right], \quad c = [F(1) - F(0)].$$

Proof. As $Y^p = Y$, we still have

$$\delta^p_\varepsilon(x) = \frac{1}{(2\varepsilon)^{p}} \left[ Y(x + \varepsilon) - Y(x - \varepsilon) \right] = \frac{1}{(2\varepsilon)^{p-1}} \delta_\varepsilon(x),$$

$$\tau^p_\varepsilon(x) = \frac{1}{(2\varepsilon)^{p}} \left[ Y(x + 1) - Y(x - 1) \right] = \frac{1}{(2\varepsilon)^{p-1}} \tau_\varepsilon(x),$$

$$\sigma^p_\varepsilon(x) = Y(x + \varepsilon) - Y(x - \varepsilon) = \sigma_\varepsilon(x),$$

and from (7) and (8), we obtain the results. \qed

2.2. Approximation in distribution spaces.

**Proposition 3.** For some real number $p > 0$, we give us $\delta^p_\varepsilon$ as initial data. Then the generalized sequence of initial data $(u_{p,\varepsilon}(t=0))_\varepsilon$ has no limit in $D'(\mathbb{R})$ except in the case $p = 1$ for which we have

$$\lim_{\varepsilon \to 0} \left( u_{1,\varepsilon}(t=0) \right) = \delta.$$

Moreover:

a) Suppose that we have: $\lim_{x \to -\infty} \frac{F(x)}{x} = L \geq 0$. If $p < 1$, then we have: $\lim_{\varepsilon \to 0} u_{p,\varepsilon} = 0$.

If $p = 1$, then we have: $\lim_{\varepsilon \to 0} u_{p,\varepsilon} = \delta_{\{x=L,t\geq 0\}}$.

If $p > 1$, then $u_{p,\varepsilon}$ has no limit in $D'(\mathbb{R}^2)$.

b) Suppose that we have: $\lim_{x \to -\infty} \frac{F(x)}{x} = \infty$.

If $p \leq 1$, then we have: $\lim_{\varepsilon \to 0} u_{p,\varepsilon} = 0$.

If $p > 1$, and if: $\lim_{x \to -\infty} \frac{x^p}{F(x^p)} \frac{1}{x^{1-p}} = M_p \geq 0$, then: $\lim_{\varepsilon \to 0} u_{p,\varepsilon} = M_p Y_x \otimes \delta_t.$
If \( p > 1 \), with: 
\[
\lim_{x \to \infty} \frac{x^p}{F(x^p)} = \infty,
\]
then \( u_{p,\varepsilon} \) has no limit in \( D'_1(\mathbb{R}^2) \).

**Proof.** The conclusions about the limit in \( D'(\mathbb{R}) \) of the sequence of initial data are obvious. Now, from Part (i) of Corollary 2 we have, for any test function \( \psi \in D(\mathbb{R}^2) \)
\[
\langle u_{p,\varepsilon}, \psi \rangle = \frac{1}{(2\varepsilon)^{p-1}} \int \int Y(t) \delta_{\varepsilon}(x-c_{p,\varepsilon}t) \psi(x,t) dx dt
= \frac{1}{(2\varepsilon)^{p-1}} \int \int Y(t) \delta_{\varepsilon}(u) \psi(u+c_{p,\varepsilon}t) dudt.
\]
From the hypothesis of Part a), we have: \( \lim_{\varepsilon \to 0} c_{p,\varepsilon} = L \), from which we deduce
\[
\lim_{\varepsilon \to 0} \int \int Y(t) \delta_{\varepsilon}(u) \psi(u+c_{p,\varepsilon}t) dudt = \int Y(t) \psi(Lt,t) dt
= \langle \delta_{\{x=L,t \geq 0\}}, \psi \rangle
\]
which gives the result of Part a).

So, we can write now
\[
\langle u_{p,\varepsilon}, \psi \rangle = \frac{1}{c_{p,\varepsilon}} \frac{1}{(2\varepsilon)^{p-1}} \int \int \delta_{\varepsilon}(u) Y(v) \psi \left( u + v, \frac{v}{c_{p,\varepsilon}} \right) dv du.
\]
From the hypothesis of Part b), we have, for each \( p > 0 \): \( \lim_{\varepsilon \to 0} c_{p,\varepsilon} = \infty \), from which we deduce
\[
\lim_{\varepsilon \to 0} \int \int \delta_{\varepsilon}(u) Y(v) \psi \left( u + v, \frac{v}{c_{p,\varepsilon}} \right) dv du = \int Y(v) \psi(v,0) dv
= \langle \delta_{\{t=0\}}, \psi \rangle = \langle Y_x \otimes \delta_t, \psi \rangle
\]
and we also have
\[
\lim_{\varepsilon \to 0} \frac{1}{c_{p,\varepsilon}} \frac{1}{(2\varepsilon)^{p-1}} = \lim_{x \to \infty} \frac{x^p}{F(x^p)} \frac{1}{x^{1-p}}
\]
which gives the result of Part b). \( \square \)

**Proposition 4.** For some real number \( p > 0 \), we give us \( \tau_{\varepsilon}^p \) as initial data. Then the generalized sequence of initial data \( (u_{p,\varepsilon})_{\tau_{\varepsilon}^p} \) has no limit in \( D'_1(\mathbb{R}) \).

Moreover:

a) Suppose that we have: \( \lim_{x \to \infty} \frac{F(x)}{x} = L \geq 0 \), then \( u_{p,\varepsilon} \) has no limit in \( D'_1(\mathbb{R}^2) \).

b) Suppose that we have: \( \lim_{x \to \infty} \frac{F(x)}{x} = \infty \).
If \( \lim_{x \to \infty} \frac{x^2}{F(x)} = M \geq 0 \), then:
\[
\lim_{\varepsilon \to 0} u_{p,\varepsilon} = M(\theta * Y)_x \otimes \delta_t.
\]

If \( \lim_{x \to \infty} \frac{x^2}{F(x)} = \infty \), then \( u_{p,\varepsilon} \) has no limit in \( \mathcal{D}'(\mathbb{R}^2) \).

**Proof.** We can write
\[
\tau_{\varepsilon}(x) = \frac{1}{2\varepsilon} \theta(x), \quad \text{with} \quad \theta(x) = Y(x + 1) - Y(x - 1)
\]
which proves the conclusion about the limit of the sequency of initial data. Then from Part (ii) of Corollary 2 we can compute the corresponding solution as
\[
u_{p,\varepsilon}(x, t) = \frac{1}{(2\varepsilon)^p} Y(t)\theta(x - c_{p,\varepsilon}t)
\]
for each \((x, t) \in \mathbb{R} \times \mathbb{R}_+\). That is to say that, for any test function \( \psi \in \mathcal{D}(\mathbb{R}^2) \) we have
\[
\langle u_{p,\varepsilon}, \psi \rangle = \frac{1}{(2\varepsilon)^p} \int \int Y(t)\theta(x - c_{p,\varepsilon}t)\psi(x, t)dxdt
\]
\[
= \frac{1}{(2\varepsilon)^p} \int \int Y(t)\theta(u)\psi(u + c_{p,\varepsilon}t, t)dudt.
\]
From the hypothesis of Part a), we have:
\[
\lim_{\varepsilon \to 0} c_{p,\varepsilon} = L,
\]
from which we deduce
\[
\lim_{\varepsilon \to 0} \int \int Y(t)\theta(u)\psi(u + c_{p,\varepsilon}t, t)dudt = \int \int Y(t)\theta(u)\psi(u + Lt, t)dudt.
\]
Then \( u_{p,\varepsilon} \) has no limit in \( \mathcal{D}'(\mathbb{R}^2) \).

Therefore we also have
\[
\langle u_{p,\varepsilon}, \psi \rangle = \frac{1}{c_{p,\varepsilon}} \frac{1}{(2\varepsilon)^p} \int \int Y(v)\theta(u)\psi(u + v, \frac{v}{c_{p,\varepsilon}})dudv.
\]
From the hypothesis of Part b), we have, for each \( p > 0 \)
\[
\lim_{\varepsilon \to 0} c_{p,\varepsilon} = \infty,
\]
from which we deduce
\[
\lim_{\varepsilon \to 0} \int \int Y(v)\theta(u)\psi(u + v, \frac{v}{c_{p,\varepsilon}})dudv = \int \int Y(v)\theta(u)\psi(u + v, 0)dudv
\]
\[
= \langle (\theta * Y)_x \otimes \delta_t, \psi \rangle
\]
and we also have
\[
\lim_{\varepsilon \to 0} \frac{1}{c_{p,\varepsilon}} \frac{1}{(2\varepsilon)^p} = \lim_{x \to \infty} \frac{x^2}{F(x)}
\]
which gives the result of Part b). \(\square\)
Proposition 5. For some real number \( p > 0 \), we give us \( \sigma^p_\varepsilon \) as initial data. Then the generalized sequence of initial data \( (u_{p,\varepsilon(t=0)})_\varepsilon \) tends to 0 in \( D'(\mathbb{R}) \).

Moreover, the generalized sequence \( (u_{p,\varepsilon})_\varepsilon \) tends to 0 in \( D'(\mathbb{R}^2) \).

Proof. We have here
\[
\sigma^p_\varepsilon(x) = \sigma_\varepsilon(x) = Y(x + \varepsilon) - Y(x - \varepsilon)
\]
which proves the conclusion about the limit of the sequence of initial data. From Part (iii) of Corollary 2 we can compute the corresponding solution as
\[
u_{p,\varepsilon}(x,t) = Y(t)\sigma_\varepsilon(x - ct) = 2\varepsilon Y(t)\delta_\varepsilon(x - ct)
\]
and we have, for any test function \( \psi \in D(\mathbb{R}^2) \)
\[
\langle u_{p,\varepsilon}, \psi \rangle = 2\varepsilon \int \int Y(t)\delta_\varepsilon(x - ct)\psi(x,t)dxdt = 2\varepsilon \int \int Y(t)\delta_\varepsilon(u)\psi(u + ct,t)dudt
\]
from which we deduce
\[
\lim_{\varepsilon \to 0} \int \int Y(t)\delta_\varepsilon(u)\psi(u + ct,t)dudt = \int Y(t)\psi(ct,t)dt = \langle \delta_{\{x=ct,t\geq 0\}}, \psi \rangle
\]
which gives the result. \( \Box \)

Remark 6. So, the distribution spaces are not convenient to describe the above solutions of Burgers’ equation for at least two reasons. First, some of these families of solutions have 0 as limit or no limit in \( D'(\mathbb{R}^2) \) and cannot be distinguished, and secondly, except for only two cases, the family of initial data has no limit in \( D'(\mathbb{R}) \).

3. The sheaves of \((C, E, P)\)-algebras.

In the theory of generalized functions, the following construction extends many points of view met in the literature (e.g., [3], [9], [20]).

We give here a more general definition of the \((C, E, P)\)-algebra than the first previous one [11], [13], [15]. In such algebras we have good tools to pose and solve many non-linear (and even linear) problem with irregular data [12], [14], [16], [17]. The topological aspects are studied in [5], [6]. And we will choose the algebraic structure adapted to our Burgers’ problem.

3.1. The algebraic structure.

a) It is given:
- A set \( \Lambda \) of indices,
- a subring \( A \) of the ring \( K^\Lambda \), \((K = \mathbb{R} \text{ or } \mathbb{C})\),
- \( A^+ = \{(r_\lambda)_\lambda \in A : r_\lambda \geq 0\} \),
the following stability by overestimation property for $A$: Whenever $(|s_\lambda|)_\lambda \leq (r_\lambda)_\lambda$ (that is to say: For each $\lambda$ we have: $|s_\lambda| \leq r_\lambda$) for any pair $((s_\lambda)_\lambda, (r_\lambda)_\lambda) \in \mathbb{K}^\lambda \times A^+$, it follows that we have: $(s_\lambda)_\lambda \in A$.

- an ideal $I_A$ of $A$ with the same stability by overestimation property,

- a sheaf $E$ of $\mathbb{K}$-algebra on a topological space $X$, such that, for each open set $\Omega$ in $X$, the algebra $E(\Omega)$ is endowed with the family $\mathcal{P}(\Omega) = (p_i)_{i \in I(\Omega)}$ of semi-norms verifying

$$\forall i \in I(\Omega), \exists (j, k, C) \in I(\Omega) \times I(\Omega) \times \mathbb{R}^+_\ast : p_i(fg) \leq Cp_j(f)p_k(g),$$

- the following property: For two open subsets of $X$ such that $\Omega_1 \subset \Omega_2$, we have $I(\Omega_1) \subset I(\Omega_2)$, and that for every $i \in I(\Omega_1)$ and $u \in E(\Omega_2)$, we have $p_i \left(\frac{u}{\Omega_1}\right) = p_i(u)$.

b) We put

$$H_{(A, E, \mathcal{P})}(\Omega) = \left\{(u_\lambda)_\lambda \in [E(\Omega)]^\lambda, \forall i \in I(\Omega) : (p_i(u_\lambda))_\lambda \in A^+\right\},$$

$$J_{(I_A, E, \mathcal{P})}(\Omega) = \left\{(u_\lambda)_\lambda \in [E(\Omega)]^\lambda, \forall i \in I(\Omega) : (p_i(u_\lambda))_\lambda \in I_A^+\right\},$$

$$C = A/I_A.$$

It may be easily seen that $A^+$ is not a subring of $A$, but is stable under addition and product. It is the same for $I_A^+$.

At the end, we suppose that the sets $|A| = \{(|r_\lambda|)_\lambda \in \mathbb{R}^+_\ast : (r_\lambda)_\lambda \in A\}$ and $|I_A| = \{(|r_\lambda|)_\lambda \in \mathbb{R}^+_\ast : (r_\lambda)_\lambda \in I_A\}$ are respectively subsets of $A$ and $I_A$. Then we have: $|A| = A^+$ and $|I_A| = I_A^+$.

c) From [11] it follows that under the above hypothesis, we obtain:

**Proposition 7.**

(i) $H_{(A, E, \mathcal{P})}$ is a sheaf of subalgebras of the sheaf $E^\lambda$;

(ii) $J_{(I_A, E, \mathcal{P})}$ is a sheaf of ideals of $H_{(A, E, \mathcal{P})}$;

(iii) the constant factor sheaf $H_{(A, E, \mathcal{P})}/J_{(I_A, E, \mathcal{P})}$ is exactly the factor ring $C = A/I_A$.

d) Now, we can give the following definition:

**Definition 8.** We call $(C, E, \mathcal{P})$-algebra every factor algebra

$$A = H_{(A, E, \mathcal{P})}/J_{(I_A, E, \mathcal{P})}$$

and we denote by $[u_\lambda]$ the class defined by the representative $(u_\lambda)_\lambda \in A$.

By a convenient choice of $C$, $E$ and $\mathcal{P}$ as parameters, we can describe many algebras of generalized functions [11] and define some other operations than algebraic ones such as differentiation, restriction and sheaf embeddings. We also can define local or microlocal analysis. The association or weak equality
previously defined by Colombeau [3] is a very useful process to study some differential equation in non-conservative form as Burgers’ one.

3.2. Operations, processes and properties.

3.2.1. Nonlinear functions. If \( f \) is a sheaf mapping: \( E \to E \), we can define \( \sim f : A \to A \) as a sheaf extension mapping by putting \( f(u) = [f(u_\lambda)] \), for any \( u \in A \), defined by the representative \( (u_\lambda)_{\lambda \in \Lambda} \in \mathcal{H}_{(A,E,P)}(\Omega) \). Of course, such an extension needs the following conditions:

- If \( (u_\lambda)_{\lambda \in \Lambda} \) belongs to \( \mathcal{H}_{(A,E,P)}(\Omega) \), then \( (f(u_\lambda))_{\lambda} \) belongs to \( \mathcal{H}_{(A,E,P)}(\Omega) \).
- If \( (i_\lambda)_{\lambda \in \Lambda} \) belongs to \( \mathcal{I}_{(A,E,P)}(\Omega) \), then \( (f(u_\lambda + i_\lambda) - f(u_\lambda))_{\lambda} \) belongs to \( \mathcal{I}_{(A,E,P)} \).

If \( P = \sum_{k=0}^{n} f_k U_k \) is a polynomial with given coefficients in \( \mathbb{K} \), or in \( \mathcal{E}(X) \), it is easy to see that the above conditions are fulfilled for the sheaf mapping denoted by \( P : \mathcal{E} \to \mathcal{E} \) and defined by

\[
P_\Omega(e) = \sum_{k=0}^{n} f_k e^k \quad \text{with} \quad e \in \mathcal{E}(\Omega).
\]

So, \( P \) has an extension \( \tilde{P} = \Psi : A \to A \) which is the sheaf mapping defined by

\[
\Psi_\Omega(u) = \left[ \sum_{k=0}^{n} f_k u^k_\lambda \right] \quad \text{with} \quad u = [u_\lambda] \in A(\Omega).
\]

Remark 9. Therefore we always have polynomials as non-linear functions in any \((\mathcal{C}, \mathcal{E}, \mathcal{P})\)-algebra.

3.2.2. Overgenerated rings. In view of applications, it is interesting to define rings generated by some given elements. More precisely, let \( B_p = \{(r_n)_\lambda \in (\mathbb{R}_+^\ast)^\Lambda, \ n = 1, 2, \ldots, p\} \) and \( B = \text{span} B_p \) be the set of elements of \((\mathbb{R}_+^\ast)^\Lambda\) obtained as products, quotients and linear combinations with coefficients in \( \mathbb{R}_+^\ast \), of elements in \( B_p \).

Define

\[
A = \{(a_\lambda)_\lambda \in \mathbb{K}^\Lambda, \exists (b_\lambda)_\lambda \in B : |a_\lambda| \leq b_\lambda\}.
\]

It is easy to see that \( A \) is a subring of \( \mathbb{K}^\Lambda \) with the stability by overestimation property. Then, we set the following definition:

Definition 10. \( A \) is overgenerated by \( B_p \). And if \( I_A \) is some ideal of \( A \) with the same stability by overestimation property, we can also say that \( \mathcal{C} = A/I_A \) is overgenerated by \( B_p \).
Example 11. As an ideal $I_A$ of $A$, we can take

$$I_A = \{ (a_\lambda) \in \mathbb{K}^A \forall (b_\lambda) \in B : |a_\lambda| \leq b_\lambda \}.$$

3.2.3. Relationship with distribution theory. We can adapt the $(\mathcal{C}, \mathcal{E}, \mathcal{P})$-algebras to the multiplication of the distributions in $A$. So we have first to embed $D'$ in $A$. If $(\varphi_\varepsilon)_{\varepsilon \in [0,1]}$ is some given family of mollifiers

$$\varphi_\varepsilon (x) = \frac{1}{\varepsilon^n} \varphi \left( \frac{x}{\varepsilon} \right), x \in \mathbb{R}^n, \int \varphi \, dx = 1,$$

we can prove that if one has: $T \in D' (\mathbb{R}^n)$, the convolution product family $(T * \varphi_\varepsilon)_{\varepsilon}$ is a slowly increasing in $1/\varepsilon$ family of smooth functions of $C^\infty (\mathbb{R}^n)$.

Example 12. If $B_p$ contains only that family, it is easy to prove that we have

$$A = \left\{ (r_\varepsilon) \in \mathbb{R}^{[0,1]}, \exists p \in \mathbb{N}, \exists C > 0 : (|r_\varepsilon|) \leq \left( \frac{C}{\varepsilon^p} \right) \right\},$$

and we can take

$$I_A = \left\{ (r_\varepsilon) \in \mathbb{R}^{[0,1]}, \forall q \in \mathbb{N}, \exists D > 0 : (|r_\varepsilon|) \leq (D\varepsilon^q) \right\}.$$

Example 13. Now, for each open set $\Omega$ in $X = \mathbb{R}^n$, let us take $\mathcal{E}(\Omega) = C^\infty (\Omega)$ with the usual $\mathcal{P}(\Omega)$ topology of uniform convergency of all the derivatives on the compact subset of $\Omega$. So, we have then:

$$\mathcal{H}_{(\mathcal{A}, \mathcal{E}, \mathcal{P})} (\Omega) = \left\{ (u_\varepsilon) \in [C^\infty (\Omega)]^{[0,1]}, \forall K \subset \subset \Omega, \forall \alpha \in \mathbb{N}^n, \exists p \in \mathbb{N}, \right.$$ 

$$\exists C > 0 : \left( \sup_{x \in K} |D^\alpha u_\varepsilon (x)| \right) \leq \left( \frac{C}{\varepsilon^p} \right), \left. \right\},$$

$$\mathcal{J}_{(\mathcal{A}, \mathcal{E}, \mathcal{P})} (\Omega) = \left\{ (u_\varepsilon) \in [C^\infty (\Omega)]^{[0,1]}, \forall K \subset \subset \Omega, \forall \alpha \in \mathbb{N}^n, \forall q \in \mathbb{N}, \right.$$ 

$$\exists D > 0 : \left( \sup_{x \in K} |D^\alpha u_\varepsilon (x)| \right) \leq (D\varepsilon^q), \left. \right\}.$$

Then, in this case, the algebra $\mathcal{A} (\Omega) = \mathcal{H}_{(\mathcal{A}, \mathcal{E}, \mathcal{P})} (\Omega) / \mathcal{J}_{(\mathcal{A}, \mathcal{E}, \mathcal{P})} (\Omega)$ is exactly the Colombeau’s simplified one, [3], and we can embed $D'(\mathbb{R}^n)$ into $\mathcal{A} (\mathbb{R}^n)$ by the mapping

$$(17) \quad T \rightarrow (T * \varphi_\varepsilon)$$

because $(T * \varphi_\varepsilon)\varepsilon$ belongs to $\mathcal{H}_{(\mathcal{A}, \mathcal{E}, \mathcal{P})} (\mathbb{R}^n)$.

In the same way, with the help of a cutoff function, we can define, for each open set $\Omega$ in $\mathbb{R}^n$, an embedding of $D'(\Omega)$ into $\mathcal{A} (\Omega)$, and finally a sheaf embedding: $D' \rightarrow \mathcal{A}$.  

This embedding depends on the choice of the mollifier $\varphi_\varepsilon$. And it is easy to define a canonical embedding from $C^\infty$ into $A$. But we can have a $C^k$ embedding only through the $D'$ one.

3.2.4. On some embeddings. In a more general way, if $E$ is a given sheaf of $K$-vector spaces with a linear sheaf embedding: $l : E \to E$, some problems about existence of embeddings from $E$ to $A$ and from $E$ to $A$ are solved by the following results:

**Lemma 14.**

(a) There exists a canonical sheaf morphism

$$j : E \to A,$$

i.e., $j_\Omega : E(\Omega) \to A(\Omega)$ for each open set $\Omega \subset X$

such that

$$\forall f \in E(\Omega) : j_\Omega(f) = [(f)_\lambda]$$

if and only if we have

$$\lambda \in A.$$ (1)

(b) There exists a linear sheaf morphism

$$k : E \to A,$$

i.e., $k_\Omega : E(\Omega) \to A(\Omega)$ for each open set $\Omega \subset X$

if and only if, for each $\lambda \in \Lambda$ and each open set $\Omega \subset X$, there exists a linear embedding $k_{\lambda,\Omega} : E(\Omega) \to E(\Omega)$, such that we have

$$\forall u \in E(\Omega) : (k_{\lambda,\Omega}(u))_\lambda \in H_{(A,E,P)}(\Omega).$$ (2)

**Proof.** (a) Suppose (18) holds, and suppose $f \neq 0$. There exists $i \in I(\Omega)$ such that $p_i(f) = a > 0$. Then $(p_i(f))_\lambda = (a)_\lambda \in A$. But we have also $(\frac{1}{a})_\lambda \in A$, and (19) holds. Reciprocally, if (19) holds, we have obviously: $(p_i(f))_\lambda \in A$ for each $f \in E(\Omega)$. So, (18) defines a mapping $j_\Omega$ from $E(\Omega)$ to $A(\Omega)$ which is obviously a morphism of algebras.

(b) Suppose (20) holds. For each $u \in E(\Omega)$, we have $k_\Omega(u) = [u_\lambda]$ for some $(u_\lambda)_\lambda \in H_{(A,E,P)}(\Omega)$. Then, for each $\lambda \in \Lambda$, we can put $k_{\lambda,\Omega}(u) = u_\lambda$ and (21) is fulfilled. Conversely, if there exists a linear embedding $k_{\lambda,\Omega} : E(\Omega) \to E(\Omega)$, such that (21) holds, when putting $k_\Omega(u) = [k_{\lambda,\Omega}(u)]$, for each $u \in E(\Omega)$, we define a linear sheaf morphism which verifies (20).

**Proposition 15.** We suppose that the mappings $j$ and $k$ verify the conditions of the previous lemma and moreover there exists a linear sheaf embedding

$$l : E \to E$$

i.e., $l_\Omega : E(\Omega) \to A(\Omega)$ for each open set $\Omega \subset X$.

Then the subsheaf $\text{Im} \ l$ can be canonically equipped with an algebraic structure for which $k$ is an algebra sheaf morphism from $\text{Im} \ l$ to $A$ if, for each open set $\Omega \subset X$

$$\forall u \in \text{Im} \ l_\Omega : ((k_{\lambda,\Omega} - l^{-1})(u))_\lambda \in I_{(A,E,P)}(\Omega).$$ (22)
Proof. For \( u = l^{-1}(f) \) and \( v = l^{-1}(g) \) in \( \text{Im} \ l_{\Omega} \), we can define the product \( uv \) by putting

\[
uv = l(fg)
\]

which gives \( \text{Im} \ l_{\Omega} \) an algebraic structure.

Now, if (22) is fulfilled, for \( u = l(f) \), we have

\[
\forall f \subset \mathcal{E}(\Omega) : (k_{\lambda\Omega}(l(f)) - f)_\lambda \in \mathcal{I}(\lambda) (\Omega),
\]

which is a necessary and sufficient condition to have

\[
j = k \circ l.
\]

As \( j \) and \( l \) are sheaf morphisms of algebras (from \( \mathcal{E} \) to \( A \) and \( \mathcal{E} \) to \( \text{Im} \ l_{\Omega} \)), it is the same for \( k \) (from \( \text{Im} \ l_{\Omega} \) to \( A \)). \( \Box \)

Remark 16. This result summarizes some questions posed in [20] about the construction of \( A \) from given \( E \) and \( E \). For example, let us put \( \lambda = (\varphi, \varepsilon) \in \Lambda = \{ \varphi \} \times [0, 1] \), with \( \mathcal{E} = C^\infty \) and \( E = \mathcal{D}' \), with a given special mollifier \( \varphi \in \mathcal{S}(\mathbb{R}^n) \) such that \( \int \varphi(x)x^\alpha dx = 0 \) (resp. 1) for \( \alpha \in \mathbb{N}^n \) (resp. \( \alpha = 0 \)). In this case, it is proved in [3] that for \( f = C^\infty(\Omega) \) and \( \langle l(f), \psi \rangle = \int f(x)\psi(x)dx \) when \( \psi \in \mathcal{D}(\Omega) \), we have

\[
(k_{\lambda\Omega}(l(f)) - f)_\lambda = (\varphi_\varepsilon \ast (\chi_\varepsilon f) - f)_\varepsilon \in \mathcal{I}(\lambda) (\Omega)
\]

with \( \varphi_\varepsilon(x) = \frac{1}{\varepsilon^n} \varphi \left( \frac{x}{\varepsilon} \right) \) and \( \chi_\varepsilon \in \mathcal{D}(\Omega) \) such that \( \chi_\varepsilon(x) = 1 \) if \( d(x, \partial \Omega) < \varepsilon \) and \( d(x, 0) > \frac{\varepsilon}{\kappa} \).

Then, the sheaf embedding \( \mathcal{D}' \to A \) defined, for \( T \in \mathcal{D}'(\Omega) \) by

\[
k_{\Omega}(T) = [\varphi_\varepsilon \ast (\chi_\varepsilon T)]
\]

is an algebra sheaf morphism from \( \text{Im} \ l = \mathcal{E} \to A \).

3.2.5. The restriction. When \( \mathcal{E} = C^\infty \), the restriction to the submanifold

\[
\{x = (x_1, x_2, \ldots, x_j, \ldots, x_n) : x_j = 0\} \subset \mathbb{R}^n
\]

of the generalized function \( u = [u_\lambda] \in A(\mathbb{R}^n) \) is the generalized one:

Definition 17.

\[
\text{u}_{\{x_j=0\}} = \left[x \mapsto u_\lambda \left(x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n\right)\right].
\]

That restriction belongs to a subalgebra of \( A(\mathbb{R}^n) \) which is canonically identified with \( A(\mathbb{R}^{n-1}) \) and agrees with the similar process met in the literature.
3.2.6. The derivation. When $\mathcal{E} = \mathcal{C}^\infty$, the derivation is defined for each $u = [u_\lambda] \in \mathcal{A}(\Omega)$, by:

$$D^\alpha u = [D^\alpha u_\lambda]$$

where $D^\alpha$ is the classical derivation. Naturally the restriction of $D^\alpha$ to $\mathcal{C}^\infty$ or $D'$ agrees with the classical derivation.

Let be $\mathcal{E}$ a sheaf on $X$, with a “derivation operator”, that is to say a sheaf endomorphism $D^\alpha$ such that:

- $D^\alpha + D^\beta = D^\alpha \circ D^\beta = D^\beta \circ D^\alpha$,
- $\forall f, g \in \mathcal{E}(\Omega): D^\alpha (fg) = \sum_{\beta \leq \alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!} D^\beta f D^{\alpha-\beta} g$.

Then it is possible to define a sheaf endomorphism $D^\alpha$ on $\mathcal{A}$, with the same properties as $D^\alpha$ under the only condition:

If $(u_\lambda)_\lambda \in J(I, \mathcal{A}, \mathcal{E}, \mathcal{P})(\Omega)$, then: $(D^\alpha u_\lambda)_\lambda \in J(I, \mathcal{A}, \mathcal{E}, \mathcal{P})$.

In this case, we can give as above the following definition:

**Definition 18.** If $u = [u_\lambda] \in \mathcal{A}(\Omega)$, then, the $D^\alpha$-derivation is given by

$$D^\alpha u = [D^\alpha u_\lambda] .$$

3.3. The association process. We suppose that $\Lambda$ is left-filtering for the given (partial) order relation $\prec$.

Let us denote by:

- $\Omega$ every open set in $X$,
- $\mathcal{E}$ a given sheaf of topological $\mathbb{K}$-vector spaces containing $\mathcal{E}$ as a subsheaf,
- $\Phi$ a given application from $\Lambda$ to $\mathbb{K}$ such that $(\Phi(\lambda))_\lambda = (\Phi_\lambda)_\lambda \in \mathcal{A}$.

We also suppose that we have

$$J(I, \mathcal{A}, \mathcal{E}, \mathcal{P})(\Omega) \subset \left\{ (u_\lambda)_\lambda \in \mathcal{H}(I, \mathcal{A}, \mathcal{E}, \mathcal{P})(\Omega) : \lim_{\mathcal{E}(\Omega)} u_\lambda = 0 \right\} . \quad (24)$$

Then, for $u = [u_\lambda]$ and $v = [v_\lambda] \in \mathcal{E}(\Omega)$, we define the $\Phi$-$\mathcal{E}$ association.

**Definition 19.** We denote by

$$u \underset{\mathcal{E}(\Omega)}{\approx} v$$

the $\Phi$-$\mathcal{E}$ association between $u$ and $v$ defined by

$$\lim_{\mathcal{E}(\Omega)} \Phi_\lambda (u_\lambda - v_\lambda) = 0. \quad (25)$$

That is to say that for each neighbourhood $V$ of 0 for the $\mathcal{E}$-topology, there exists $\lambda_0 \in \Lambda$ such that

$$\lambda \prec \lambda_0 \Rightarrow \Phi_\lambda(u_\lambda - v_\lambda) \in V.$$
To be sure that the above condition is independent of the representatives of \( u \) and \( v \), we have to verify that if \( \lim_{E(\Omega)} H_{(\Lambda, E, P)}(\Omega) \), then, for any \( (i_\lambda)_{\lambda} \in J_{(I_A, E, P)}(\Omega) \), \( \lim_{E(\Omega)} \Phi_{\lambda}(w_\lambda + i_\lambda) = 0 \) also holds. To prove the last condition, it is sufficient to show that

\[
(\Phi_{\lambda}i_\lambda)_{\lambda} \in J_{(I_A, E, P)}(\Omega).
\]

But for each \( i \in I(\Omega) \), we have \( p_i(\Phi_{\lambda}i_\lambda) = |\Phi_{\lambda}|p_i(i_\lambda) \). And, according to the definitions and the stability properties given in Section 3.1, we have \( (|\Phi_{\lambda}|)_{\lambda} \in A^+ \) and \( (p_i(i_\lambda))_{\lambda} \in I_A^+ \). Then we also have \( (|\Phi_{\lambda}|p_i(i_\lambda))_{\lambda} \), which proves the required above condition.

**Remark 20.** When we have \( \Phi = 1 \), it is clear that the above association is weaker than equality. And when taking \( E = D', E = C^\infty, \Lambda = [0,1] \), and \( \Phi = 1 \), we find again the association process defined by Colombeau [3] or Egorov [9] who works in the \((C, E, P)\)-algebra defined by choosing \( C = C^0/[C^0,1] \) where \( C^0,1 \) is the ring of complex numbers families \( (z_\varepsilon)_{\varepsilon} \) such that \( z_\varepsilon = 0 \) if \( \varepsilon \) is small enough, and where \( E \) is the sheaf of \( C^\infty \)-complex valued functions. However it is possible to give stronger forms of association, as in the following definition:

**Definition 21.** Let be \( A \) and \( I_A \) respectively given by (15) and (16).

The \( E \)-association \( u \approx v \), between two elements \( u \) and \( v \in A(\Omega) \) is defined by: For each \( p \in \mathbb{N} \) we have

\[
\lim_{E(\Omega)\varepsilon \to 0} \frac{1}{p} (u_\varepsilon - v_\varepsilon) = 0.
\]

Then, by taking \( E = E \), we obtain:

**Proposition 22.** The \( E \)-association is equivalent to the equality in \( A(\Omega) \).

**Proof.** If one has \( u \approx v \), then for each \( i \in I(\Omega) \) and \( p \in \mathbb{N}^m \), \( p_i \left( \frac{1}{p} (u_\varepsilon - v_\varepsilon) \right) \) is bounded, that is to say: \( (u_\varepsilon - v_\varepsilon)_\varepsilon \) is in the set:

\[
\left\{(u_\varepsilon)_\varepsilon \in [E(\Omega)]^{0,1}, \forall i \in I(\Omega), \forall \varepsilon \in [0,1]: (p_i(u_\varepsilon))_{\varepsilon} \in I_A^+ \right\} = J_{(I_A, E, P)}(\Omega).
\]

\( \square \)

**4. The Burgers problem in \((C, E, P)\)-algebras.**

The first step is to choose the algebra \( A(\mathbb{R}) \) and the initial data by means of convenient parameters. The second is to give an as good as possible approximation of corresponding self-similar solutions of the Cauchy problem (1) in a convenient algebra \( A(\Omega) \) with \( \Omega = \mathbb{R} \times \mathbb{R}^+ \).
4.1. The algebra $\mathcal{A}(\mathbb{R})$ and the initial data. We choose $\mathcal{E} = E = C^\infty$. We define $\mathcal{E}(\mathbb{R})$ as $C^\infty(\mathbb{R})$, with the usual $\mathcal{P}(\mathbb{R})$ topology of uniform convergence of all the derivatives on the compact subset of $\mathbb{R}$ and we define $E(\mathbb{R})$ as $\mathcal{D}'(\mathbb{R})$. According to (13), we have here

$$\mathcal{A}(\mathbb{R}) = \mathcal{H}_{(A,\mathcal{E},\mathcal{P})}/\mathcal{J}_{(I_{A,\mathcal{E},\mathcal{P}})}(\mathbb{R}).$$

We choose $\Lambda = [0,1] \times [0,1]$, left-filtering for the partial order $\prec$ defined by

$$(\varepsilon, \eta_1) \prec (\varepsilon_2, \eta_2) \text{ if } \varepsilon_1 \leq \varepsilon_2 \text{ and } \eta_1 \leq \eta_2.\quad (26)$$

Now, for a given $\varphi \in \mathcal{D}(\mathbb{R})$, with $\text{supp} \varphi = [-1,1]$, $0 \leq \varphi \leq 1$, $\varphi(0) = 1$, and $\varphi^{(k)}(0) = 0$ for each $k \in \mathbb{N}^*$, let us consider, for $x \in \mathbb{R}$

$$\varphi_{(\varepsilon,\eta)}(x) = \frac{1}{2\varepsilon} \varphi \left( \frac{x}{\eta} \right).$$

Then, we define $\delta_{(\varepsilon,\eta)}$, $\tau_{(\varepsilon,\eta)}$, $\sigma_{(\varepsilon,\eta)} \in \mathcal{D}(\mathbb{R})$ in the following way:

$$(\varepsilon, \eta) \delta_{(\varepsilon,\eta)}(x) = \begin{cases} \varphi_{(\varepsilon,\eta)}(x + \varepsilon) = \frac{1}{2 \varepsilon} \varphi \left( \frac{x + \varepsilon}{\eta} \right) & \text{for } x \in ]-\infty, -\varepsilon[ \\ \frac{1}{2 \varepsilon} & \text{for } x \in [-\varepsilon, \varepsilon] \\ \varphi_{(\varepsilon,\eta)}(x - \varepsilon) = \frac{1}{2 \varepsilon} \varphi \left( \frac{x - \varepsilon}{\eta} \right) & \text{for } x \in ]\varepsilon, +\infty[ \end{cases}\quad (27)$$

$$(\varepsilon, \eta) \tau_{(\varepsilon,\eta)}(x) = \begin{cases} \varphi_{(\varepsilon,\eta)}(x + 1) = \frac{1}{2 \varepsilon} \varphi \left( \frac{x + 1}{\eta} \right) & \text{for } x \in ]-\infty, -1[ \\ \frac{1}{2 \varepsilon} & \text{for } x \in [-1, 1] \\ \varphi_{(\varepsilon,\eta)}(x - 1) = \frac{1}{2 \varepsilon} \varphi \left( \frac{x - 1}{\eta} \right) & \text{for } x \in ]1, +\infty[ \end{cases}\quad (28)$$

$$(\varepsilon, \eta) \sigma_{(\varepsilon,\eta)}(x) = \begin{cases} \varphi_{(\varepsilon,\eta)}(x + \varepsilon) = \varphi \left( \frac{x + \varepsilon}{\eta} \right) & \text{for } x \in ]-\infty, -\varepsilon[ \\ 1 & \text{for } x \in [-\varepsilon, \varepsilon] \\ \varphi_{(\varepsilon,\eta)}(x - \varepsilon) = \varphi \left( \frac{x - \varepsilon}{\eta} \right) & \text{for } x \in ]\varepsilon, +\infty[ \end{cases}\quad (29)$$

It is easy to see that these functions belong to $\mathcal{D}(\mathbb{R})$ and we have:

$$\text{supp } \delta_{(\varepsilon,\eta)} = \text{supp } \sigma_{(\varepsilon,\eta)} = [-\eta - \varepsilon, \varepsilon + \eta] \text{ and supp } \tau_{(\varepsilon,\eta)} = [-\eta - 1, 1 + \eta].$$

Proposition 23. If $A$ is the subring of $\mathbb{R}^{[0,1] \times [0,1]}$ overgenered by the set

$$B_2 = \{ (\varepsilon)(\varepsilon,\eta), (\eta)(\varepsilon,\eta) \}$$

then, for any real number $p > 0$, $\Delta^p = \left[ \delta^p_{(\varepsilon,\eta)} \right]$, $\Upsilon^p = \left[ \tau^p_{(\varepsilon,\eta)} \right]$, and $\Xi^p = \left[ \sigma^p_{(\varepsilon,\eta)} \right]$ belong to $\mathcal{A}(\mathbb{R})$. 
Proof. First, we can see that, for each compact $K \subset \mathbb{R}$ and each $\alpha \in \mathbb{N}$

$$\sup_{x \in K} \left| D^{\alpha} \delta^{p}_{(\varepsilon, \eta)} (x) \right| \leq \frac{1}{(2\varepsilon)^{p}} \frac{1}{\eta^{\alpha}} \sup_{x \in [-1,1]} \left| D^{\alpha} \varphi^{p} (x) \right|$$

and then for each seminorm $p_{K,l}$ of the $\mathcal{P} (\mathbb{R})$ topology of $\mathcal{E} (\mathbb{R}) = C^{\infty} (\mathbb{R})$, one has: $p_{K,l} (\delta^{p}_{(\varepsilon, \eta)}) \in A^{+}$, that is to say: $\delta^{p}_{(\varepsilon, \eta)}$ belongs to $\mathcal{H} (A, \mathcal{E}, \mathcal{P}) (\mathbb{R})$. The same proof holds for $\tau^{p}_{(\varepsilon, \eta)}$ because we have the same estimation for $\sup_{x \in K} \left| D^{\alpha} \tau^{p}_{(\varepsilon, \eta)} (x) \right|$. Then, from

$$\sup_{x \in K} \left| D^{\alpha} \sigma^{p}_{(\varepsilon, \eta)} (x) \right| \leq \frac{1}{\eta^{\alpha}} \sup_{x \in [-1,1]} \left| D^{\alpha} \varphi^{p} (x) \right|$$

we can easily prove the result for $\Xi$. In this last case, the result holds if $A$ is overgenerated by \{ $(\eta)_{(\varepsilon, \eta)}$ \}, and a fortiori by $B_{2}$.

4.2. The algebra $\mathcal{A} (\overline{\Omega})$ and the delta shock wave solutions. For $\overline{\Omega} = \mathbb{R} \times \mathbb{R}_{+}$, we define $\mathcal{E} (\overline{\Omega})$ as $C^{\infty} (\overline{\Omega})$ with the usual $\mathcal{P} (\overline{\Omega})$ topology of uniform convergency of all the derivatives on the compact subset of $\overline{\Omega}$. And we define $E (\overline{\Omega})$ as $C^{\infty} (\overline{\Omega})$, with the $(p_{\psi})_{\psi \in D (\mathbb{R}^{2})}$-topology, such as

$$p_{\psi} (f) = \left| \int_{\overline{\Omega}} f (x) \psi (x) \, dx \right|$$

for each $f \in C^{\infty} (\overline{\Omega})$. It is clear that the $(p_{\psi})_{\psi \in D (\mathbb{R}^{2})}$-topology is weaker than the $\mathcal{P} (\overline{\Omega})$ one.

Then the algebra $\mathcal{A} (\overline{\Omega})$ is defined as

$$\mathcal{A} (\overline{\Omega}) = \mathcal{H} (A, \mathcal{E}, \mathcal{P}) / \mathcal{J} (I_{A}, \mathcal{E}, \mathcal{P}) (\overline{\Omega}) .$$

The main results are the two following propositions:

**Proposition 24.** We suppose that $F$ is a polynomial function with real coefficients and increasing with $F(0) = 0$, we suppose that $A$ is overgenerated by $B_{2}$, we take $\Phi_{p} (\varepsilon, \eta) = \frac{\varepsilon}{\eta F (\frac{1}{(2\varepsilon)^{p}})}$, and we put respectively $\Delta^{p}$ or $\Upsilon^{p}$ as initial data. Then,

$$c_{p} = \left[ c_{p, (\varepsilon, \eta)} \right]_{\Phi_{p} \in D^{\prime} (\mathbb{R})} \left[ (2\varepsilon)^{p} F \left( \frac{1}{(2\varepsilon)^{p}} \right) \right]$$

is a necessary and sufficient condition to have in $\mathcal{A} (\overline{\Omega})$ a $\Phi_{p}$-$E$ approximate solution of the Cauchy problem (1) with the respective self-similar form

$$\Delta^{p} (x - c_{p} t) = \left[ (x, t) \rightarrow \delta^{p}_{(\varepsilon, \eta)} (x - c_{p} (\varepsilon, \eta) t) \right]$$

or

$$\Upsilon^{p} (x - c_{p} t) = \left[ (x, t) \rightarrow \tau^{p}_{(\varepsilon, \eta)} (x - c_{p} (\varepsilon, \eta) t) \right] .$$
Proof. To simplify the proof, we suppose that \( p = 1 \). As \( F \) is polynomial, the \( \left( F \left( \frac{\varepsilon}{2\varepsilon} \right) \right)_{(\varepsilon,\eta)} \) and \( \left( \frac{\varepsilon}{\eta F \left( \frac{\varepsilon}{2\varepsilon} \right)} \right)_{(\varepsilon,\eta)} \) belong to \( A \). Let be \( \mathcal{F} \) the sheaf mapping \( \mathcal{E} \to \mathcal{E} \) induced by \( F \), and \( \hat{\mathcal{F}}: \mathcal{A} \to \mathcal{A} \) the sheaf mapping extension of \( \mathcal{F} \). As \( \mathcal{F} \left( \delta_{(\varepsilon,\eta)} \right) \) belongs to \( \mathcal{H}(A,\mathcal{E},\mathcal{P})(\mathbb{R}) \), according to the hypothesis on \( \mathcal{F} \), we can consider the difference
\[
\mathcal{F} \left( \delta_{(\varepsilon,\eta)} \right) - 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta_{(\varepsilon,\eta)}.
\]
Then, if \( x \in [-\varepsilon,\varepsilon] \), we have
\[
\left( \mathcal{F} \left( \delta_{(\varepsilon,\eta)} \right) \right)(x) = 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta_{(\varepsilon,\eta)}(x).
\]
For \( \psi \in \mathcal{D}(\mathbb{R}) \), we compute
\[
\left\langle \left( \mathcal{F} \left( \delta_{(\varepsilon,\eta)} \right) - 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta_{(\varepsilon,\eta)} \right), \psi \right\rangle = A_{(\varepsilon,\eta)} + B_{(\varepsilon,\eta)}
\]
with
\[
A_{(\varepsilon,\eta)} = \int_{-\varepsilon-\eta}^{-\varepsilon} \left( F \left[ \delta_{(\varepsilon,\eta)}(x) \right] - F \left( \frac{1}{2\varepsilon} \right) \varphi \left( \frac{x + \varepsilon}{\eta} \right) \right) \psi(x) dx,
\]
\[
B_{(\varepsilon,\eta)} = \int_{\varepsilon}^{\varepsilon + \eta} \left( F \left[ \delta_{(\varepsilon,\eta)}(x) \right] - F \left( \frac{1}{2\varepsilon} \right) \varphi \left( \frac{x - \varepsilon}{\eta} \right) \right) \psi(x) dx.
\]
As \( F \) is increasing, we have when \( x \in [-\varepsilon - \eta, -\varepsilon] \)
\[
\left| F \left( \frac{1}{2\varepsilon} \right) \varphi \left( \frac{x + \varepsilon}{\eta} \right) - F \left[ \delta_{(\varepsilon,\eta)}(x) \right] \right| \leq F \left( \frac{1}{2\varepsilon} \right)
\]
and then
\[
\left| A_{(\varepsilon,\eta)} \right| \leq F \left( \frac{1}{2\varepsilon} \right) \int_{-\varepsilon-\eta}^{-\varepsilon} |\psi(x)| dx.
\]
Therefore, we have finally
\[
\left| A_{(\varepsilon,\eta)} \right| \leq C\eta F \left( \frac{1}{2\varepsilon} \right)
\]
for some positive constant \( C \), with a similar estimation for \( B_{(\varepsilon,\eta)} \). Then we have
\[
\lim_{D'(\mathbb{R}) \to 0} \frac{\varepsilon}{\eta F \left( \frac{1}{2\varepsilon} \right)} \left( \mathcal{F} \left( \delta_{(\varepsilon,\eta)} \right) - 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta_{(\varepsilon,\eta)} \right) = 0.
\]
From the hypothesis on \( \mathcal{F} \) and \( A \) we can see that \( \left( \mathcal{F} \left( \delta_{(\varepsilon,\eta)} \right) \right)_{(\varepsilon,\eta)} \) and \( \left( 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta_{(\varepsilon,\eta)} \right)_{(\varepsilon,\eta)} \) belong to \( \mathcal{H}(A,\mathcal{E},\mathcal{P}) \) and \( \left( \frac{\varepsilon}{\eta F \left( \frac{1}{2\varepsilon} \right)} \right)_{(\varepsilon,\eta)} \) belongs to \( A \).
That is to say
\[
\mathcal{F} \left( \delta(\varepsilon, \eta) \right) \overset{\Phi}{\rightarrow} D' \left( \mathbb{R} \right) \left[ 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta(\varepsilon, \eta) \right]
\]
according to our definition of association process, with \( \Phi_1(\varepsilon, \eta) = \frac{\varepsilon}{\eta F \left( \frac{1}{2\varepsilon} \right)} \).

And that is a good approximation because if \( p \) is given as large as we want, we can have
\[
\Phi_1(\varepsilon, \eta) > \frac{1}{\varepsilon^p} \text{ by choosing } \eta < \frac{\varepsilon^{p+1}}{F \left( \frac{1}{2\varepsilon} \right)}.
\]

Now, if we put \( \Delta = [\delta(\varepsilon, \eta)] \in \mathcal{A} \left( \mathbb{R} \right) \), we can write
\[
\mathcal{F} \left( \delta(\varepsilon, \eta) \right) \overset{\Phi_1}{\rightarrow} D' \left( \mathbb{R} \right) \left[ 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \right] \Delta.
\]

Let be \( \mathfrak{f} : \mathcal{A} \rightarrow \mathcal{A} \) the sheaf mapping defined from \( \mathcal{F} : \mathcal{E} \rightarrow \mathcal{E} \) and \( F \) by the derivation \( F' = f \). It is easy to see that if one has: \( \mathfrak{f} (u) \overset{\Phi_1}{\rightarrow} v \), with \( u \) and \( v \) in \( \mathcal{A} \left( \mathbb{R} \right) \), then we have: \( \mathfrak{f} (u) u' \overset{\Phi_1}{\rightarrow} v' \) in the sense of generalized functions. And then we have also
\[
\mathfrak{f} (\Delta) \Delta' \overset{\Phi_1}{\rightarrow} D' \left( \mathbb{R} \right) \left[ 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \right] \Delta'.
\]

Now, for \( (x, t) \in \overline{\Omega} \), for each compact \( K \subseteq \mathbb{R} \) and each \((k_1, k_2) \in \mathbb{N} \times \mathbb{N}\) we have
\[
\sup_{x \in K} \left| \mathcal{D}^{(k_1, k_2)} \delta(\varepsilon, \eta) (x - c_{1, (\varepsilon, \eta)t}) \right| \leq \frac{1}{2\varepsilon} \left| \varepsilon \right|^{k_1+k_2} \left| c_{1, (\varepsilon, \eta)t} \right|^{k_2} \sup_{x \in [-1, 1]} \left| \mathcal{D}^{k_1+k_2} \varphi(x) \right|
\]
from which it is easy to prove that the family \( \left((x, t) \rightarrow \delta(\varepsilon, \eta)(x - c_{1, (\varepsilon, \eta)t})\right)_{(\varepsilon, \eta)} \) belongs to \( \mathcal{H} \left( \mathcal{A}, \mathcal{E}, \mathcal{P} \right) \left( \overline{\Omega} \right) \). Then, with \( c_1 = \left[ c_{1, (\varepsilon, \eta)} \right] \), we can put
\[
\Delta(x - c_1 t) = \left[(x, t) \rightarrow \delta(\varepsilon, \eta)(x - c_{1, (\varepsilon, \eta)t})\right]
\]
and prove, from (32) that we have
\[
\mathcal{F} \left( \delta(x - c_1 t) \right) \overset{\Phi_1}{\rightarrow} D' \left( \mathbb{R} \right) \left[ 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \right] \Delta(x - c_1 t).
\]

To do that, we first have to compute
\[
D(\varepsilon, \eta) = p_\psi \left( \Phi_1(\varepsilon, \eta) G(\varepsilon, \eta) \right)
\]
with
\[
G(\varepsilon, \eta)(x, t) = \left( F \left( \delta(\varepsilon, \eta)(x - c_{1, (\varepsilon, \eta)t}) \right) - 2\varepsilon F \left( \frac{1}{2\varepsilon} \right) \delta(\varepsilon, \eta)(x - c_{1, (\varepsilon, \eta)t}) \right).
\]
Thus we can write
\[ D(\varepsilon, \eta) = \left| \int \int Y(t) \Phi_1(\varepsilon, \eta) G(\varepsilon, \eta) (x, t) \psi(x, t) dx \right| dt = \int H(\varepsilon, \eta)(t) dt \]
where
\[ H(\varepsilon, \eta)(t) = \int Y(t) \Phi_1(\varepsilon, \eta) \left[ F(\delta(\varepsilon, \eta)) (u) - 2\varepsilon F\left(\frac{1}{2\varepsilon}\right) \delta(\varepsilon, \eta) (u) \right] \psi(u + c_1(\varepsilon, \eta) t, t) du. \]

From (31), we can see that for each \( t \in \mathbb{R} \):
\[ \lim_{(\varepsilon, \eta) \to 0} H(\varepsilon, \eta)(t) = 0. \]
Moreover, if \( A > 0 \) is chosen such that \( \text{supp} \psi \subset [-A, A] \times [-A, A] \), we can see from (34) that we have
\[ |H(\varepsilon, \eta)(t)| \leq C \chi_{[0, A]} \]
for some constant \( C \), \( \chi_{[0, A]} \) being the characteristic function of \([0, A]\), and so, by the Lebesgue majorant convergence theorem, we obtain:
\[ \lim_{(\varepsilon, \eta) \to 0} D(\varepsilon, \eta) = 0, \]
which proves (33).

Now we can compute
\[ \frac{\partial}{\partial t} \Delta(x - c_1 t) = \left[ (x, t) \rightarrow -c_1(\varepsilon, \eta) \delta'(\varepsilon, \eta) (x - c_1, \varepsilon, \eta) t) \right] \]
\[ = - [c_1(\varepsilon, \eta)] \Delta'(x - c_1 t) \]
and from (33)
\[ \frac{\partial}{\partial x} \tilde{\Delta} (\Delta(x - ct)) = \int (\Delta(x - c_1 t)) \Delta'(x - c_1 t) \left. \Phi_1 \right|_{E(\Omega)} \left[ 2\varepsilon F\left(\frac{1}{2\varepsilon}\right) \right] \Delta'(x - c_1 t). \]

So, we have proved that the condition is necessary. And it is obviously sufficient. Particularly, we can take \( c_1(\varepsilon, \eta) = 2\varepsilon F\left(\frac{1}{2\varepsilon}\right) \).

Instead of \( \mathcal{F} \left( \delta(\varepsilon, \eta) \right) - 2\varepsilon F\left(\frac{1}{2\varepsilon}\right) \delta(\varepsilon, \eta) = 0 \) on \([-\varepsilon, \varepsilon]\) we have: \( \mathcal{F} \left( \tau(\varepsilon, \eta) \right) - F\left(\frac{1}{2\varepsilon}\right) \tau(\varepsilon, \eta) = 0 \) on \([-1, 1]\). But on \([-\eta - 1, -1] \), \( \mathcal{F} \left( \tau(\varepsilon, \eta) \right) - F\left(\frac{1}{2\varepsilon}\right) \tau(\varepsilon, \eta) \) is deduced by translation from \( \mathcal{F} \left( \delta(\varepsilon, \eta) \right) - 2\varepsilon F\left(\frac{1}{2\varepsilon}\right) \delta(\varepsilon, \eta) \). So, the same estimations, computations and conclusions holds for \( \Upsilon(x - c_1 t) \) as well as for \( \Delta(x - c_1 t) \).

The case \( p > 0 \) follows with slight modifications. \( \square \)

**Proposition 25.** We suppose that \( F \) is a polynomial function with real coefficients and increasing with \( F(0) = 0 \), we suppose that \( A \) is overgenerated by
\[ B_2 = \{(\varepsilon), (\eta)\}. \]
we take $\Psi(\varepsilon, \eta) = \frac{\varepsilon}{\eta} F (1)$, and for any integer $p \geq 1$, we put $\Sigma^p$ as initial data. Then,

\[
d = \begin{bmatrix} d_{(\varepsilon, \eta)} \end{bmatrix}_{[0,1] \times [0,1]} [F (1)]
\]

is a necessary and sufficient condition to have in $A(\Omega)$ a $\Psi$-E approximate solution of the Cauchy problem (1) with the self-similar form

\[
\Sigma^p(x - dt) = \left[ (x, t) \rightarrow \sigma^p_{(\varepsilon, \eta)} (x - d_{(\varepsilon, \eta)} t) \right].
\]

Proof. We can consider the difference $F \left( \sigma^p_{(\varepsilon, \eta)} \right) - F (1) \sigma^p_{(\varepsilon, \eta)}$.

Then

\[
\left\langle \left( F \left( \sigma^p_{(\varepsilon, \eta)} \right) - F (1) \sigma^p_{(\varepsilon, \eta)} \right), \psi \right\rangle
= \int_{-\varepsilon-\eta}^{-\varepsilon} \left( F \left( \sigma^p_{(\varepsilon, \eta)} \right) - F (1) \sigma^p_{(\varepsilon, \eta)} \right) (x) \psi(x) dx
+ \int_{\varepsilon}^{\varepsilon+\eta} \left( F \left( \sigma^p_{(\varepsilon, \eta)} \right) - F (1) \sigma^p_{(\varepsilon, \eta)} \right) (x) \psi(x) dx
= A'_{(\varepsilon, \eta)} + B'_{(\varepsilon, \eta)}.
\]

We can compute

\[
A'_{(\varepsilon, \eta)} = \int_{-\varepsilon-\eta}^{-\varepsilon} \left( F \left[ \sigma^p_{(\varepsilon, \eta)} \right] (x) - F (1) \varphi^p \left( \frac{x + \varepsilon}{\eta} \right) \right) \psi(x) dx.
\]

As $F$ is increasing, we have when $x \in [-\varepsilon - \eta, -\varepsilon]$

\[
\left| F (1) \varphi^p \left( \frac{x + \varepsilon}{\eta} \right) - F \left[ \sigma^p_{(\varepsilon, \eta)} \right] (x) \right| \leq F (1)
\]

and then

\[
\left| A'_{(\varepsilon, \eta)} \right| \leq F (1) \int_{-\varepsilon-\eta}^{-\varepsilon} |\psi(x)| dx.
\]

Therefore, we have finally

\[
\left| A'_{(\varepsilon, \eta)} \right| \leq C \eta F (1)
\]

for some positive constant $C$, with a similar estimation for $B'_{(\varepsilon, \eta)}$. Then we have
From the hypothesis on $\mathcal{F}$ and $A$ it follows that \( \mathcal{F}\left(\sigma_p(\varepsilon,\eta)\right) \) and \( \left( F(1)\sigma_p(\varepsilon,\eta)\right) \) belong to $\mathcal{H}_{(A,\mathcal{E},\mathcal{P})}(\mathbb{R})$, with \( \varepsilon \eta \mathcal{F}\left(1\right) \) in $A$. That is to say

\[
\left[ \mathcal{F}\left(\sigma_p(\varepsilon,\eta)\right) \right] \overset{\mathcal{D}'(\mathbb{R})}{\approx} \left[ F(1)\sigma_p(\varepsilon,\eta) \right].
\]

The next steps can be easily deduced from the previous proof, with slight modifications, and we can choose $F(1)$ as the speed of propagation of the $\Psi$-approximate solution. □

**Conclusion 26.** For any $p > 0$, one can say that $\Delta^p(x - c_p t)$, $\Upsilon^p(x - c_p t)$, $\Sigma^p(x - dt)$ are some modelizations of $p$-power of delta-waves, tsunamis, or solitons and approximate self-similar solutions of Burgers’ equation. The generalized function $\Sigma^p(x - dt)$ propagates approximatively with a finite speed $d = F(1)$, but $\Delta^p(x - c_p t)$ and $\Upsilon^p(x - c_p t)$ have the same generalized speed $c_p \approx \left[ (2\varepsilon)^p F\left(\frac{1}{(2\varepsilon)^p}\right) \right]$ depending only upon the generalized height (and not thickness) of the initial data $\Delta^p$ and $\Upsilon^p$. And in each case, the solution lies in an (algebra) adapted to the problem by the choice of the algebra $\mathcal{E}(\Omega)$ and the ring $\mathcal{C}$ of generalized numbers. Indeed $\mathcal{E}[0,1] \times [0,1](\Omega)$ contains the representatives of our expected approximate solutions and $\mathcal{C}$ is overgenerated by some elements connected to equation and data singularities. Moreover, these singularities also are connected to the required ($\Phi_p-E$ or $\Psi-E$) approximation processes.

**References**


THE CROSSING NUMBER AND MAXIMAL BRIDGE LENGTH OF A KNOT DIAGRAM

A. STOIMENOW (WITH AN APPENDIX BY MARK KIDWELL)

We give examples showing that Kidwell’s inequality for the maximal degree of the Brandt-Lickorish-Millett-Ho polynomial is in general not sharp.

1. Introduction.

The $Q$ (or absolute) polynomial is a polynomial invariant in one variable $z$ of links (and in particular knots) in $S^3$ without orientation. It can be defined by being 1 on the unknot and the relation

$$A_1 + A_{-1} = z(A_0 + A_{\infty}).$$

Here $A_i$ denote the $Q$ polynomials of links $K_i$, such that $K_i$ ($i \in \mathbb{Z} \cup \{\infty\}$) possess diagrams equal except in one tangle site (or “room” in the terminology of [LM]), where an $i$-tangle (in the Conway [Co] sense) is inserted. See Figure 1.

![Figure 1. The Conway tangles.](image)

The polynomial was discovered in 1985 independently by Brandt, Lickorish and Millett [BLM] and Ho [Ho]. Several months after its discovery, Kauffman [Ka] found a 2-variable polynomial $F(a, z)$, specializing to $Q$ by setting $a = 1$.

In [Ki], Kidwell found a nice inequality for the maximal degree of the $Q$ polynomial:
Theorem 1.1 (Kidwell). Let $D$ be a diagram of a knot (or link) $K$. Then

\[
\max \deg Q(K) \leq c(D) - d(D),
\]

where $c(D)$ is the crossing number of $D$ and $d(D)$ its maximal bridge length, i.e., the maximal number of consecutive crossing over- or underpasses. Moreover, if $D$ is alternating (i.e., $d(D) = 1$) and prime, then equality holds in (2).

In [Mo, Problem 4, p. 560] he asked whether (2) always becomes equality when minimizing the r.h.s. over all diagrams $D$ of $K$. From the theorem it follows that this is true for alternating knots and also for those non-alternating knots $K$, where $\max \deg Q(K) = c(K) - 2$ (here $c(K)$ denotes the crossing number of $K$). All non-alternating knots in Rolfsen’s tables [Ro] have this property except for one – the Perko knot 10161 (and its obversed duplication 10162), where $\max \deg Q = 6$. Hence, as quoted by Kidwell, this knot became a promising candidate for strict inequality in (2). To express ourselves more easily, we define:

Definition 1.1. Call a knot $K$ $Q$-maximal, if (2) with the r.h.s. minimized over all diagrams $D$ of $K$ becomes equality.

The aim of this note is to show that indeed the Perko knot is not $Q$-maximal. We give several modifications of our arguments and examples showing how they can be applied to exhibit non-$Q$-maximality.

2. Plane curves.

We start with some discussion of plane curves.

Definition 2.1. A non-closed plane curve is a $C^1$ map $\gamma : [0, 1] \to \mathbb{R}^2$ with $\gamma(0) \neq \gamma(1)$ and only transverse self-intersections. $\gamma$ carries a natural orientation.

Example 2.1. Here are some plane curves:

\begin{center}
\includegraphics[width=0.5\textwidth]{plane_curves.png}
\end{center}

In the following, whenever talking of plane curves we mean non-closed ones with orientation, unless otherwise stated. However, in some cases it is possible to forget about orientation if it is irrelevant. It is convenient to identify $\gamma$ with $\gamma([0, 1])$ wherever this causes no confusion. Whenever we want to emphasize that a line segment in a local picture starts with an endpoint, the endpoint will be depicted as a thickened dot.

Definition 2.2. The crossing number $c(\gamma)$ of a curve $\gamma$ is the number of self-intersections (crossings). A curve $\gamma$ with $c(\gamma) = 0$ is called simple or trivial.
**Definition 2.3.** We call a non-closed curve $\tilde{\gamma}$ *transversely homotopic* to $\gamma$ (and denote it by $\tilde{\gamma} \sim \gamma$) if $\tilde{\gamma}(0) = \gamma(0)$, $\tilde{\gamma}(1) = \gamma(1)$ and $\tilde{\gamma}$ intersects $\gamma$ only transversely. Call the number $d(\gamma) = \min\{ \#(\gamma \cap \tilde{\gamma}) : \tilde{\gamma} \sim \gamma \} - 2$ the *endpoint distance* of $\gamma$. (The term ‘−2’ serves to discount the coincidence of endpoints.)

A curve $\tilde{\gamma}$ realizing this minimum is called *minimal transversely homotopic* to $\gamma$. Such $\tilde{\gamma}$ can be chosen to have no self-intersections.

**Example 2.2.** The curves ——— and ——— have $d = 0$, while $d(\text{??}) = 1$.

**Definition 2.4.** We call a plane curve $\gamma$ *composite*, if there is a *closed* plane curve $\gamma'$ (with no self-intersections) such that $\gamma'$ intersects $\gamma$ in exactly one point, transversely, and in both components of $\mathbb{R}^2 \setminus \gamma'$ there are crossings of $\gamma$. In this case $\gamma'$ separates $\gamma$ into two parts $\gamma_1$ and $\gamma_2$, which we call *components* of $\gamma$. We write $\gamma = \gamma_1 \# \gamma_2$. Conversely, this can be used to define the operation ‘$\#$’ (*connected sum*) of $\gamma_1$ and $\gamma_2$, wherever $\gamma_1(1)$ or $\gamma_2(0)$ are in the closure of the unbounded component of their complements.

We call $\gamma$ *prime*, if it is not composite.

**Example 2.3.**

These examples illustrate that the connected sum in general depends on the orientation of the summands and their order.

It is clear that the crossing number is additive under connected sum and it’s a little exercise to verify that the endpoint distance is as well.

The path we are going to follow starts with the following:
Exercise 2.1. Verify that the complement of a curve \( \gamma \) has \( c(\gamma) + 1 \) connected components, and conclude from that that \( d(\gamma) \leq c(\gamma) \).

Hint: One way to show that is to observe it when \( \gamma \) is trivial, to prove that you can obtain any \( \gamma \) from the trivial one by the four local moves

\[
\begin{align*}
\includegraphics{local_moves1} & \quad \leftrightarrow \quad \includegraphics{local_moves2} \\
\includegraphics{local_moves3} & \quad \leftrightarrow \quad \includegraphics{local_moves4}
\end{align*}
\]

and to trace how the number of components and \( c(\gamma) \) change under these moves.

For the Perko knot we need to work a little harder. We need to prove that \( d(\gamma) \leq \max(3, c(\gamma) - 3) \) if \( \gamma \) prime. This was originally achieved by the author by refining the argument in the exercise above. Subsequently, Kidwell further generalized and simplified the argument, showing the following stronger inequality.

**Theorem 2.1** (Kidwell). For any plane curve \( \gamma \) we have \( d(\gamma) \leq c(\gamma)/2 \).

A proof of this theorem will be given in the Appendix.

3. Non-\( Q \)-maximal knots.

Using Kidwell’s theorem, we are prepared to exhibit the Perko knot as non-\( Q \)-maximal.

**Theorem 3.1.** If \( D \) is a prime diagram of a knot \( K \) of \( c(D) \) crossings with a bridge of length \( l = c(D) - k \), and \( D \) has minimal crossing number among all such diagrams for fixed \( k \), then \( l \leq k/2 \), hence \( c(D) \leq 3/2 k \).

From this we have the desired example:

**Example 3.1.** If \( 10_{161} \) were \( Q \)-maximal, then we could pose \( k = 6 \) in the theorem and would obtain a 9 crossing diagram of the knot, which does not exist. Hence \( 10_{161} \) is not \( Q \)-maximal.

**Proof of Theorem 3.1.** This is basically Kidwell’s theorem. Consider \( \gamma' \) to be the part of \( D \) consisting of the maximal (length) bridge and \( \gamma \) consisting of the rest of (the solid line of) \( D \) with signs of all crossings ignored. Then the freedom to move the bridge corresponds to the freedom to move \( \gamma' \).

Clearly, for many phenomena Rolfsen’s tables up to 10 crossings are very limited. Scanning the list of non-alternating knots of at most 15 crossings
provided by Thistlethwaite (see [HTW]), I found 192 15 crossing knots for which \( \text{max deg } Q \leq 8 \), and hence for which we would be done showing non-\( Q \)-maximality already with Exercise 2.1. The most striking examples are the knots \( 15_{119574} \) and \( 15_{119873} \), where \( \text{max deg } Q = 4 \). Although for both knots \( \text{max deg}_z F(a, z) = 11 \), the coefficients of the 7 highest powers of \( z \) cancel when setting \( a = 1 \). This property exhibits non-\( Q \)-minimality in an alternative way, because Theorem 1.1 in fact can be analogously shown to hold with the maximal \( z \)-degree of the Kauffman polynomial \( F(a, z) \) in place of \( \text{max deg } Q \). Nevertheless, placing emphasis on \( \text{max deg } Q \) rather than \( \text{max deg}_z F(a, z) \) in the above discussion, beside by Kidwell’s original work, can be justified at least because for the main example, the Perko knot, \( \text{max deg } Q = \text{max deg}_z F(a, z) \), and so the geometric argument is needed.

There are several ways that the theorem can be modified.

**Theorem 3.2.** If \( D \) is a diagram of a knot \( K \) of \( c(D) \) crossings with a bridge of length \( l = c(D) - k \), then \( u(K) \leq \lfloor k/2 \rfloor \), where \( u(K) \) denotes the unknotting number of \( K \).

**Proof.** By switching at most half of the crossings in \( D \) not involved in the maximal bridge, the remaining part \( \gamma \) of the plane curve (this time with signs of the crossings) can be layered, i.e., any crossing is passed the first time as over- and then as under-crossing or vice versa. But reinstalling the bridge to a layered \( \gamma \) gives a layered, and hence unknotted, diagram. \( \Box \)

**Corollary 3.1.** If \( u(K) > \lfloor \text{max deg } Q(K)/2 \rfloor \), then \( K \) is not \( Q \)-maximal.

Unfortunately, this corollary does not work to show non-\( Q \)-maximality of Perko’s knot. Verifying both hand-sides of the inequality (using that the unknotting number of \( 10_{161} \) is 3, see [St, Km, Ta]), we find that we just have equality. And that equality does not suffice is seen, e.g., from all 8 closed positive braid knots in Rolfsen’s tables (see [Cr, Bu]) and more generally from the \((2, n)\)-torus knots for \( n \) odd.

For knots of > 10 crossings unknotting numbers are not tabulated (anywhere I know of) and a general machinery does not exist to compute them, hence when wanting to extend the search space for examples applicable to Corollary 3.1, it makes sense to replace the unknotting number by lower bounds for it, which can be computed straightforwardly. I tried two such bounds. First we have the signature \( \sigma \):

**Corollary 3.2.** If \( |\sigma(K)| > \text{max deg } Q(K) \), then \( K \) is not \( Q \)-maximal.

Clearly, replacing \( u(K) \) by lower bounds for it makes the condition more and more restrictive. However, when checking the above mentioned list of 192 knots, I found that at least one of them satisfied strict inequality. It is \( 15_{166028} \), where \( \sigma = 8 \) and \( \text{max deg } Q = 7 \).
Another possibility is to minorate \( u(K) \) by the bound coming from the \( Q \) polynomial itself.

**Corollary 3.3.** If \( 2 \log_{-3} Q(-1) > \max \deg Q(K) \), then \( K \) is not \( Q \)-maximal.

**Remark 3.1.** The negative logarithm base may disturb the reader because such logarithms are usually not defined. But by work of Sakuma, Murakami, Nakanishi (see Theorem 8.4.8 (2) of [Kw]) and Lickorish and Millett [LM] \( Q(-1) \) is always a(n integral) power of \(-3\) and this one it is referred to by this expression.

The inequality in Corollary 3.3 looks rather bizarre. First, the inequality \( u(K) \geq \log_{-3} Q(-1) \) is in general much less sharp than the one with the signature and secondly, the inequality in Corollary 3.3 requires the coefficients of \( Q \) to be of an average magnitude which grows exponentially with \( \max \deg Q \). Thus, non-surprisingly, my quest for applicable examples among the non-alternating 15 and 16 crossing knots ended with no success in this case.

**Question 3.1.** Is there a knot \( K \) with \( 2 \log_{-3} Q(-1) > \max \deg Q(K) \)?

I nevertheless gave the above inequality, because it is self-contained w.r.t. \( Q \) and would decide about non-\( Q \)-maximality from \( Q \) itself (without knowing anything else about the knot) and hence is, in some sense, also beautiful.

**References**


Appendix I. A proof of Theorem 2.1 (by M. Kidwell).

By the additivity of \( c \) and \( d \) under connected sum, it clearly suffices to prove the inequality for \( \gamma \) prime. We can also exclude another degeneracy of \( \gamma \).

**Definition I.1.** A *nugatory crossing* of \( \gamma \) is a crossing \( p \) such that there is a closed curve \( \gamma' \) with \( \gamma \cap \gamma' = \{ p \} \) and \( \gamma' \) intersects transversely both strands of \( \gamma \) intersecting at \( p \).

If \( \gamma \) has a nugatory crossing, then one of the components of \( \mathbb{R}^2 \setminus \gamma \) has both and the other one has no one of the endpoints of \( \gamma \). Removing the part of \( \gamma \) in latter component and smoothing \( \gamma \) near \( p \) reduces \( c(\gamma) \), but not \( d(\gamma) \), hence we may (say, by induction on \( c(\gamma) \)) also assume that \( \gamma \) has no nugatory crossing.

Let \( \gamma : [0, 1] \to \mathbb{R}^2 \) be a prime curve in the plane with no nugatory crossings. We will assume that \( \gamma \) has only transverse crossings. In fact, we will speak of “right angles” and “straight angles” as if all crossings are orthogonal. Without loss of generality, assume that \( \gamma(0) \) lies on the boundary of
the unbounded region $U$ of $\mathbb{R}^2 \setminus \gamma$. Define the index of any region $R$ of $\mathbb{R}^2 \setminus \gamma$ as the minimum intersection number of a curve with one endpoint in $R$ and the other in $U$ with $\gamma$. Thus the unbounded region has index 0. The endpoint distance $d(\gamma)$ of $\gamma$ is the index of the region with $\gamma(1)$ on its boundary. Call any portion of $\gamma$ between two crossing points an edge of $\gamma$. The indices of regions that share an edge along their boundary either differ by one or are equal. Call an edge neutral if its two incident regions have the same index. (The edges containing $\gamma(0)$ and $\gamma(1)$ are considered neutral. By the assumption that $\gamma$ is prime, these are the only edges of $\gamma$ that border only one region.) We define the index of a neutral edge to be the common index of its two incident regions. We call a curve $\gamma'$ efficient if its endpoints lie in regions of index $i$ and $j$ and $\gamma'$ intersects $\gamma$ in exactly $|i - j|$ points.

**Lemma I.1.** An efficient curve $\gamma'$ cannot intersect a neutral edge of $\gamma$.

**Proof.** Let the endpoints of $\gamma'$ lie in regions of index $i$ and $j$ with $i \leq j$. Suppose $\gamma'$ intersects a neutral edge of index $k$ which borders regions $R_1$ and $R_2$, and let $a$ and $b$ be points in $[0, 1]$ such that $\gamma'(a) \in R_1$ and $\gamma'(b) \in R_2$. Then $\gamma'([0, a])$ intersects $\gamma$ in at least $|k - i|$ points, $\gamma'([a, b])$ intersects $\gamma$ in at least one point and $\gamma'([b, 1])$ intersects $\gamma$ in at least $|j - k|$ points. The sum of these three numbers exceeds $j - i$. □

**Lemma I.2.** If a crossing point $C$ of $\gamma$ is incident to a neutral edge, then it is incident to at least two neutral edges.

**Proof.** Let the neutral edge be incident to two regions of index $i$. Then the other two regions incident to $C$ can only have index $i$, $i - 1$, or $i + 1$. Since these two regions are incident to a common edge, their indices cannot be $i - 1$ and $i + 1$. If one of these regions has index $i$, then two neutral edges form a right angle at $C$. If these regions have index $i \pm 1$, then two neutral edges form a straight angle at $C$. Figure 3 illustrates these two cases. □

![Figure 3](image_url)

**Figure 3.** Two types of crossings involving neutral edges (the thickened lines), together with the indices of their neighbored regions.

If three neutral edges are incident to one crossing point, then all four regions incident to that point must have the same index, and thus the fourth
edge incident to that point must also be neutral. We next show that this
case cannot occur.

**Lemma I.3.** _Four neutral edges cannot be incident to one crossing point_ $C$.

**Proof.** One edge incident to $C$ is part of a continuous chain of neutral edges
ending at $\gamma(0)$ and another is part of a continuous chain of neutral edges
ending at $\gamma(1)$. The other two edges at $C$ must be part of a continuous
chain of neutral edges that forms a loop. (All these assertions follow from
Lemma I.2.) Let $R$ be any region of $\mathbb{R}^2 \setminus \gamma$ in the interior of this loop.
By definition of index, there is an efficient path joining a point in $R$ to a
point in $U$. This path must cross the loop of neutral edges, contradicting
Lemma I.1. □

These lemmas taken together show that the union of all neutral edges of
$\gamma$ is a simple curve joining $\gamma(0)$ to $\gamma(1)$. We call this curve the _neutral curve_.

**Lemma I.4.** _For every number $i$ with $0 \leq i \leq d(\gamma)$, there is at least one
neutral edge of index $i$.*_

**Proof.** The first and last edges along the neutral curve of $\gamma$ have index 0
and $d(\gamma)$. Figure 3 shows that adjacent neutral edges have equal indices or
indices that differ by 1. □

More particularly, adjacent edges that form a straight angle at a crossing
point have indices that differ by 1 and edges that form a right angle have
equal indices.

**Lemma I.5.** _Unless $\gamma$ is a simple curve, there is at least one crossing point
$C$ of $\gamma$ incident to neutral edges which form a right angle._

**Proof.** If $\gamma$ consists entirely of neutral edges then, by Lemma I.3, $\gamma$ must be
a simple curve. The curve that starts at $\gamma(0)$ and goes “straight ahead” at
every crossing point traverses all of $\gamma$. Thus if the neutral curve is to be a
proper subset of $\gamma$, it must contain at least one right angle. □

We are now ready to count crossing points, edges and regions in the plane
and apply Euler’s Theorem.

**Theorem I.1.** _If $\gamma$ is prime, without nugatory crossings and has crossing
number $c(\gamma)$, then $d(\gamma) \leq \lfloor c(\gamma)/2 \rfloor$._

**Proof.** Without changing the Euler characteristic of our configuration, we
can eliminate $\gamma(0)$, $\gamma(1)$ and the edges incident to them. We are left with a
configuration with $c(\gamma)$ vertices, two with valence 3 and $c(\gamma) - 2$ with valence
4.
Figure 4. A plane curve and its neutral curve (the thickened part). The numbers in the regions denote their indices.

The edge count is

\[
\frac{4(c(\gamma) - 2) + 3 \cdot 2}{2} = 2c(\gamma) - 1.
\]

If \( F \) is the number of faces (regions of \( \mathbb{R}^2 \setminus \gamma \), including the unbounded region), then

\[
c(\gamma) - (2c(\gamma) - 1) + F = 2, \quad \text{and so} \quad F = c(\gamma) + 1.
\]

We now count regions that are incident to the (original) neutral curve. (The curve in Figure 4 may aid understanding.) There is the unbounded region of index 0 and the region of index \( d(\gamma) \) containing \( \gamma(1) \). By Lemma I.4, there are at least two regions of index \( i \) for \( 1 \leq i \leq d(\gamma) - 1 \), since we are assuming that \( \gamma \) is prime. However, at any crossing point where the neutral curve makes a right angle, there are three regions of the same index, distinct by our assumption that \( \gamma \) has no nugatory crossings. Thus we can compare the total number of regions to the lower bound on the number of regions along the neutral curve:

\[
c(\gamma) + 1 \geq 2 + 2(d(\gamma) - 1) + 1 = 2d(\gamma) + 1, \quad \text{or}
\]
\[
c(\gamma) \geq 2d(\gamma), \quad \text{or}
\]
\[
d(\gamma) \leq \frac{c(\gamma)}{2}.
\]

Since the numbers \( d(\gamma) \) and \( c(\gamma) \) are integers, we have our desired relation. \qed
The curve in Figure 4 can be generalized to a family of examples with $c(\gamma) = 2d(\gamma)$ for any even number $c(\gamma)$. Thus the inequality cannot be improved.

These curves are not all curves with $c = 2d$; Figure 5 shows two other ones. However, one can give an explicit description of all such curves. We leave this as a task to the reader.

**Exercise I.1.** Show that $\gamma$ satisfies $c(\gamma) = 2d(\gamma)$ if and only if its neutral curve has exactly one angle, and splicing this angle

one obtains a picture like

(with its obvious generalization).

Received May 30, 1999 and revised March 28, 2001. A. Stoimenow was supported by a DFG postdoc grant.

c/o Kunio Murasugi  
Department of Mathematics  
University of Toronto  
Canada M5S 3G3  
E-mail address: stoimeno@math.toronto.edu

Department of Mathematics  
572 Holloway Road  
U.S. Naval Academy  
Annapolis, MD 21402
Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.
Local collapsing theory 1
F. T. Farrell and L. E. Jones

Exceptional surgery curves in triangulated 3-manifolds 101
Marc Lackenby

Nonlinear algebraic analysis of delta shock wave solutions to Burgers’ equation 165
J.-A. Marti

The crossing number and maximal bridge length of a knot diagram 189
A. Stoimenow (with an appendix by Mark Kidwell)