Refinements of the holonomic approximation lemma

DANIEL ÁLVAREZ-GAVELA

The holonomic approximation lemma of Eliashberg and Mishachev is a powerful tool in the philosophy of the h–principle. By carefully keeping track of the quantitative geometry behind the holonomic approximation process, we establish several refinements of this lemma. Gromov's idea from convex integration of working "one pure partial derivative at a time" is central to the discussion. We give applications of our results to flexible symplectic and contact topology.

1 Introduction and main results

1.1 Classical holonomic approximation

We begin by briefly recalling the holonomic approximation lemma, which is the starting point of our paper. Given a fiber bundle $p: X \to M$, where X and M are smooth manifolds, consider the bundle $p^r: X^{(r)} \to M$ of r–jets of p. The fiber of p^r over a point $x \in M$ consists of equivalence classes of germs of sections $h: \mathrm{Op}(x) \to X$ of p, where two germs are identified if they agree up to order r at the point x. Throughout we use Gromov's notation $\mathrm{Op}(A)$ to denote an arbitrarily small but unspecified open neighborhood of a subset $A \subset M$. If $X = M \times N$ is a trivial bundle, then sections of p are the same as maps from M to N, and we usually write $X^{(r)} = J^r(M, N)$.

Given a section $h: U \to X$ of p defined over an open subset $U \subset M$, we denote by $j^r(h): U \to X^{(r)}$ the section of p^r consisting of the r–jets of h. Sections of p^r of the form $j^r(h)$ are called holonomic. It is generally impossible to globally approximate an arbitrary section $s: M \to X^{(r)}$ of p^r by a holonomic section. Nevertheless, the approximation can always be achieved in a deformed neighborhood of any reasonable stratified subset of positive codimension. For simplicity we restrict our discussion to the following class of stratified subsets.

Definition 1.1 A closed subset $K \subset M$ is called a polyhedron if it is a subcomplex of some smooth triangulation of M.

The classical holonomic approximation lemma was first stated and proved in [3] by Eliashberg and Mishachev. They give numerous applications in their book [4]. Holonomic approximation is closely related to the method of flexible sheaves discovered by Gromov in his thesis [6] and further explored in his book [8]. Both of these techniques greatly generalize the immersion theory of Smale [12] and Hirsch [9]. The precise statement that we wish to recall reads as follows.

Theorem 1.2 (holonomic approximation lemma) *Let $s\colon M \to X^{(r)}$ be a section of the r–jet bundle of a fiber bundle $p\colon X \to M$, and let $K \subset M$ be a polyhedron of positive codimension. Then there exists an isotopy $F_t\colon M \to M$ and a holonomic section $\widehat{s}\colon \mathrm{Op}(F_1(K)) \to X^{(r)}$ such that the following properties hold:*

- *$\widehat{s}$ is C^0–close to s on $\mathrm{Op}(F_1(K))$.*
- *F_t is C^0–small.*

Remark 1.3 (1) More precisely, the C^0–closeness statement means that for any choice of Riemannian metric on X and for any $\varepsilon > 0$, there exist F_t and $\widehat{s}$ as in the statement of the theorem such that $\mathrm{dist}_{C^0}(\widehat{s}, s) < \varepsilon$ on $\mathrm{Op}(F_1(K))$ with respect to the choice of metric. Similar remarks apply below whenever we talk about the C^0–closeness of two maps.

(2) We say that an isotopy F_t is C^0–small if it is C^0–close to the identity.

(3) The holonomic approximation lemma also holds in relative and parametric form; see [4] for details.

1.2 Improved holonomic approximation

The holonomic approximation lemma (Theorem 1.2) is an extremely useful tool for proving h–principles. However, it turns out that for certain applications a stronger result is needed. The goal of this paper is to prove several refinements of the method of holonomic approximation, which we formulate in this section. In Section 1.3 we state the parametric versions. In Section 1.4 we explain how these refinements yield new flexibility results in symplectic and contact topology.

Recall that for $0 \le l < r$ there are maps $p^r_l\colon X^{(r)} \to X^{(l)}$ which forget higher-order information. The projection p^r_l gives $X^{(r)}$ the structure of a fiber bundle over $X^{(l)}$. Given a section $s\colon A \to X^{(r)}$ of p^r defined over any subset $A \subset M$, we call $s^{(l)} = p^r_l \circ s\colon A \to X^{(l)}$ the l–jet component of s.

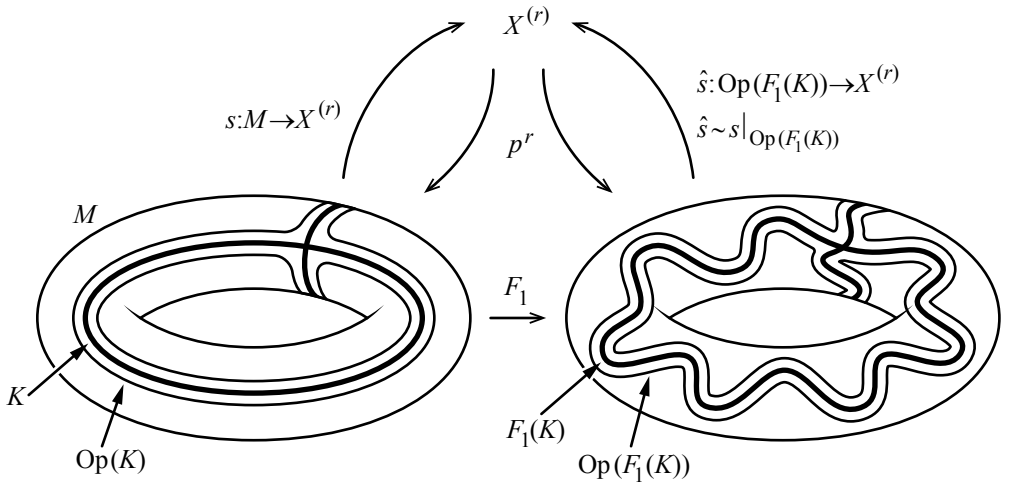Figure 1: The isotopy F_t is C^0–small but typically wiggles the subset K wildly inside M.

Definition 1.4 A section $s\colon U \to X^{(r)}$ of p^r defined over an open subset $U \subset M$ is called l–holonomic if $s^{(l)}$ is a holonomic section of p^l.

Our first refinement of the holonomic approximation lemma states that if we start with an l–holonomic section s of p^r, then it is possible to carry out the holonomic approximation process on s while ensuring global control of the l–jet component.

Theorem 1.5 (holonomic approximation lemma for l–holonomic sections) *Let $s\colon M \to X^{(r)}$ be a section of the r–jet bundle of a fiber bundle $p\colon X \to M$, and let $K \subset M$ be a polyhedron of positive codimension. Suppose that for some $l < r$ the section s is l–holonomic. Then there exists an isotopy $F_t\colon M \to M$ and a holonomic section $\hat{s}\colon M \to X^{(r)}$ such that the following properties hold:*

- $\hat{s}$ *is C^0–close to s on $\mathrm{Op}(F_1(K))$.*
- $\hat{s}^{(l)}$ *is C^0–close to $s^{(l)}$ on all of M.*
- F_t *is C^0–small.*
- $F_t = \mathrm{id}_M$ *and $\hat{s}^{(l)} = s^{(l)}$ outside of a slightly bigger neighborhood $\mathrm{Op}(K) \supset \mathrm{Op}(F_1(K))$.*

Remark 1.6 (1) Every section $s\colon M \to X^{(r)}$ of p^r is 0–holonomic. Therefore, taking $l = 0$ in Theorem 1.5 we recover the classical Theorem 1.2.

(2) Theorem 1.5 also holds in relative form. To be precise, if s is already holonomic on $\mathrm{Op}(A)$ for some closed subset $A \subset M$, then we can arrange it so that $F_t = \mathrm{id}_M$ and $\hat{s} = s$ on $\mathrm{Op}(A)$. The same comment applies to Theorem 1.8 below.

Our second refinement concerns a specific class of $(r-1)$–holonomic sections of p^r, which we call $\perp$–holonomic. Informally, we can describe a $\perp$–holonomic section as a section of p^r which differs from a holonomic section only by the formal analogue of a pure order r partial derivative. We show that when the holonomic approximation process is applied to a $\perp$–holonomic section it is not only possible to globally control the $(r-1)$–jet component, but it is also possible to globally control the order-r information complementary to this formal pure order-r partial derivative. The precise statement is most cleanly phrased in terms of the bundle $X^{\perp}$, which we define below. This bundle was first introduced by Gromov in [8] in the context of convex integration, where the language of $\perp$–jets is used to construct iterated convex hull extensions of partial differential relations. A thorough exposition of the theory of convex integration, including details on the geometry of $X^{\perp}$, can be found in Spring's book [13].

Let $\tau \subset TM$ be a hyperplane field on M. We associate to τ a bundle $p^{\perp} \colon X^{\perp} \to M$ in the following way. The fiber of $p^{\perp}$ over a point $x \in M$ consists of equivalence classes of germs of sections $h \colon \mathrm{Op}(x) \to X$ of p, where two germs are identified if their $(r-1)$–jets $y = j^{r-1}(h)(x) \in X^{(r-1)}$ at the point x are the same and, moreover, if the restrictions of their tangent maps $d(j^{r-1}(h)) \colon T_x M \to T_y X^{(r-1)}$ to the hyperplane $\tau_x \subset T_x M$ are the same. When $X = M \times N$ is a trivial bundle we also write $X^{\perp} = J^{\perp}(M, N)$. Observe that $X^{\perp}$ lies between $X^{(r)}$ and $X^{(r-1)}$ in the sense that the bundle $p^r_{r-1} \colon X^{(r)} \to X^{(r-1)}$ factors as the composition of $p^r_{\perp} \colon X^{(r)} \to X^{\perp}$ and $p^{\perp}_{r-1} \colon X^{\perp} \to X^{(r-1)}$. Given a section $s \colon A \to X^{(r)}$ of p^r defined over any subset $A \subset M$, we call $s^{\perp} = p^r_{\perp} \circ s \colon A \to X^{\perp}$ the $\perp$–component of s. Given a section $h \colon U \to X$ of p defined over an open subset $U \subset M$, we denote by $j^{\perp}(h) \colon U \to X^{\perp}$ the section of $p^{\perp}$ formed by the $\perp$–jets of h. Explicitly, $j^{\perp}(h) = p^r_{\perp} \circ j^r(h)$. Sections of $p^{\perp}$ of the form $j^{\perp}(h)$ are called holonomic.

Definition 1.7   A section $s \colon U \to X^{(r)}$ of p^r defined over an open subset $U \subset M$ is called $\perp$–holonomic with respect to a hyperplane field $\tau \subset TM$ if $s^{\perp}$ is a holonomic section of $p^{\perp}$.

Theorem 1.8 (holonomic approximation lemma for $\perp$–holonomic sections)   *Let $s \colon M \to X^{(r)}$ be a section of the r–jet bundle of a fiber bundle $p \colon X \to M$ and let $K \subset M$ be a polyhedron of positive codimension. Suppose that the section s*
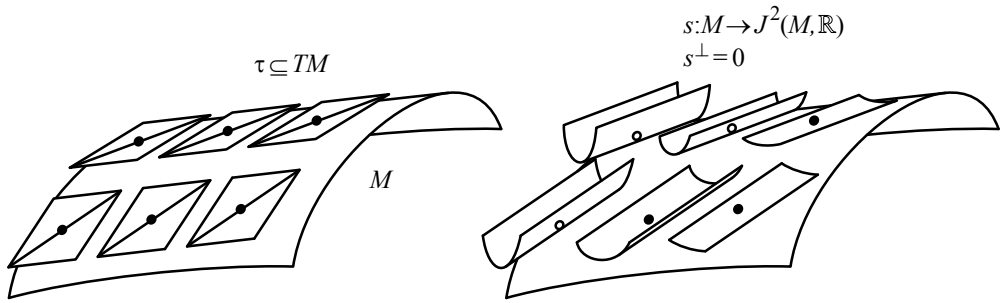
Figure 2: The section s of the 2–jet bundle $J^2(M, \mathbb{R})$ is $\perp$–holonomic with respect to the hyperplane field $\tau \subset TM$. Indeed, $s^\perp = 0$ is a holonomic section of $J^\perp(M, \mathbb{R})$.

is $\perp$–holonomic with respect to some hyperplane field $\tau \subset TM$. Then there exists an isotopy $F_t: M \to M$ and a holonomic section $\hat{s}: M \to X^{(r)}$ such that the following properties hold:

- $\hat{s}$ is C^0–close to s on $\mathrm{Op}(F_1(K))$.
- $\hat{s}^\perp$ is C^0–close to $s^\perp$ on all of M.
- F_t is C^0–small.
- $F_t = \mathrm{id}_M$ and $\hat{s}^\perp = s^\perp$ outside of a slightly bigger neighborhood $\mathrm{Op}(K) \supset \mathrm{Op}(F_1(K))$.

Remark 1.9 In fact, the proof of Theorem 1.8 will produce a very specific isotopy F_t. Informally, we can say that F_t wiggles K in such a way that the wiggles are parallel to the hyperplane field τ. More formally, we can arrange so that the pulled-back hyperplane field $F_t^*\tau$ is C^0–close to τ for all $t \in [0, 1]$; see Figure 3. An analogous comment applies in the parametric version Theorem 1.12 below.

1.3 Parametric versions

Our main results Theorem 1.5 and Theorem 1.8 remain true in families. We take our parameter space to be a compact manifold Z whose boundary ∂Z is possibly nonempty. We consider families of sections parametrized by Z. For example, a family of sections $s_z: M \to X^{(r)}$ depending on the parameter $z \in Z$ is a smooth mapping $s: Z \times M \to X^{(r)}$ such that for every $z \in Z$ the assignment $x \mapsto s(z, x)$ defines a smooth section s_z of p^r. Additionally, we allow the polyhedron K to vary with the parameter in the following way.
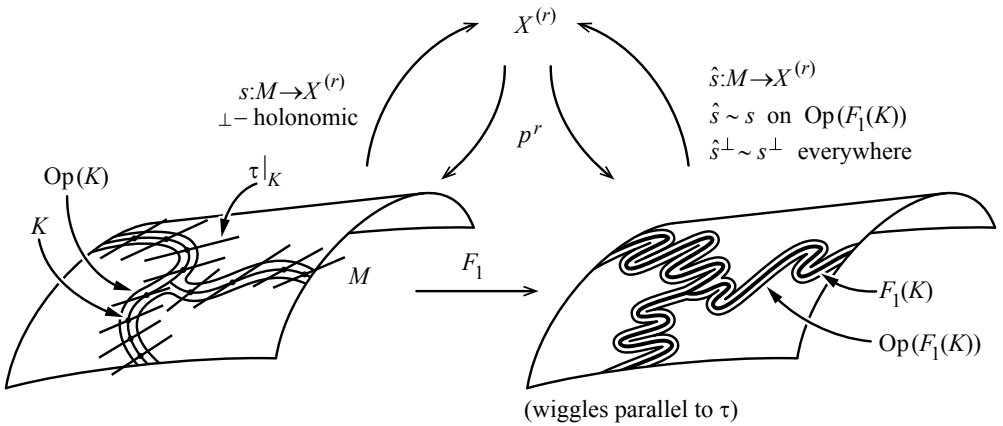
Figure 3: The wiggles of the deformed subset $F_1(K)$ are parallel to $\tau \subset TM$.

Definition 1.10 A closed subset $K \subset Z \times M$ is called a fibered polyhedron if it is a subcomplex of a smooth triangulation of $Z \times M$ which is in general position with respect to the fibers $z \times M$ for $z \in Z$.

A consequence of this definition is that for every $z \in Z$ the subset $K_z \subset M$ given by $K \cap (z \times M) = z \times K_z$ is a polyhedron in M; see Figure 4. If K has positive codimension in $Z \times M$, then K_z has positive codimension in M for all $z \in Z$. We are now ready to formulate the parametric analogues of Theorems 1.5 and 1.8.

Theorem 1.11 (parametric holonomic approximation lemma for l–holonomic sections) *Let $s_z \colon M \to X^{(r)}$ be a family of sections of the r–jet bundle associated to a fiber bundle $p \colon X \to M$ parametrized by a compact manifold Z. Let $K \subset Z \times M$ be a fibered polyhedron of positive codimension. Suppose that for some $l < r$ the sections s_z are l–holonomic for all $z \in Z$ and that they are holonomic for $z \in \mathrm{Op}(\partial Z)$. Then there exists a family of isotopies $F_t^z \colon M \to M$ and a family of holonomic sections $\hat{s}_z \colon M \to X^{(r)}$ such that the following properties hold:*

- $\hat{s}_z$ *is C^0–close to s_z on $\mathrm{Op}(F_1^z(K_z))$.*

- $\hat{s}_z^{(l)}$ *is C^0–close to $s_z^{(l)}$ on all of M.*

- F_t^z *is C^0–small.*

- $F_t^z = \mathrm{id}_M$ *and $\hat{s}_z^{(l)} = s_z^{(l)}$ outside of a slightly bigger neighborhood $\mathrm{Op}(K_z) \supset \mathrm{Op}(F_1^z(K_z))$.*

- $F_t^z = \mathrm{id}_M$ *and $\hat{s}_z = s_z$ for $z \in \mathrm{Op}(\partial Z)$.*
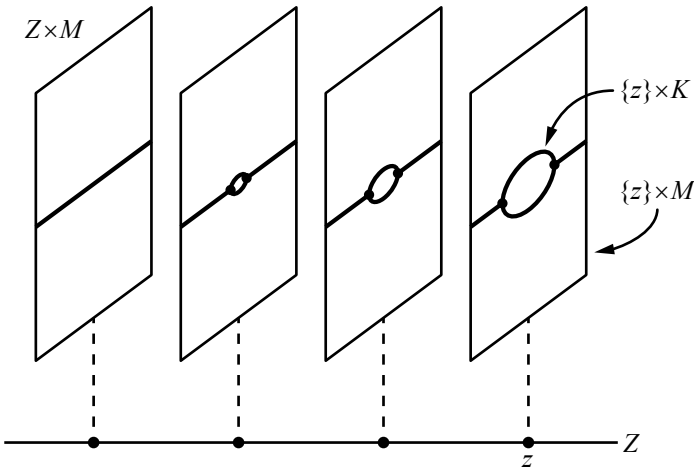
Figure 4: A typical fibered polyhedron $K \subset Z \times M$

Theorem 1.12 (parametric holonomic approximation lemma for $\perp$–holonomic sections) *Let $s_z: M \to X^{(r)}$ be a family of sections of the r–jet bundle associated to a fiber bundle $p: X \to M$ parametrized by a compact manifold Z. Let $K \subset Z \times M$ be a fibered polyhedron of positive codimension. Suppose that the sections s_z are $\perp$–holonomic with respect to some family of hyperplane fields $\tau_z \subset TM$ for all $z \in Z$ and that they are holonomic for $z \in \mathrm{Op}(\partial Z)$. Then there exists a family of isotopies $F_t^z: M \to M$ and a family of holonomic sections $\hat{s}_z: M \to X$ such that the following properties hold:*

- *$\hat{s}_z$ is C^0–close to s_z on $\mathrm{Op}(F_1^z(K_z))$.*
- *$\hat{s}_z^{\perp}$ is C^0–close to $s_z^{\perp}$ on all of M.*
- *F_t^z is C^0–small.*
- *$F_t^z = \mathrm{id}_M$ and $\hat{s}_z^{\perp} = s_z^{\perp}$ outside of a slightly bigger neighborhood $\mathrm{Op}(K_z) \supset \mathrm{Op}(F_1^z(K_z))$.*
- *$F_t^z = \mathrm{id}_M$ and $\hat{s}_z = s_z$ for $z \in \mathrm{Op}(\partial Z)$.*

Remark 1.13 (1) Observe that the formulation of the parametric Theorems 1.11 and 1.12 is relative with respect to the closed subset ∂Z of the parameter space Z. In typical applications, taking $(Z, \partial Z) = (D^m, S^{m-1})$ leads to a result about certain relative homotopy groups vanishing, which can be rephrased in terms of the existence of a homotopy equivalence (the h–principle).

(2)   Theorems 1.11 and 1.12 also hold in relative form with respect to a closed subset $A \subset M$. The statement is analogous to the one phrased in Remark 1.6.

(3)   Note that $s_z^{\perp}$ is a section of the bundle $X_z^{\perp} \to M$ associated to the hyperplane field $\tau_z \subset TI^m$, which varies with $z \in Z$. We can view the collection $s_z^{\perp}$ as a single section of the bundle $X_Z^{\perp} \to Z \times M$ whose fiber over (z, m) is $X_z^{\perp}$.

1.4   Applications to symplectic and contact topology

We begin with an example which illustrates the main point. For this section only we will use freely the elements of symplectic topology, which the reader can review in [5] by Eliashberg and Mishachev or in any other basic text on the subject. Suppose that $f: L^n \to W^{2n}$ is a Lagrangian embedding of a manifold L into a symplectic manifold (W, ω). The Weinstein neighborhood theorem states that a small tubular neighborhood $\mathcal{N}$ of $f(L)$ in W is symplectomorphic to a neighborhood of the zero section in the cotangent bundle $(T^*L, dp \wedge dq)$. We call such $\mathcal{N}$ a Weinstein neighborhood. Note that $J^1(L, \mathbb{R}) = T^*L \times \mathbb{R}$ (with the $\mathbb{R}$ factor corresponding to the 0–jet component) and that graphical exact Lagrangian submanifolds of T^*L correspond to holonomic sections of $J^1(L, \mathbb{R})$. To keep track of the tangential data attached to the embedding we must remember one more derivative, which we can do by considering holonomic sections of $J^2(L, \mathbb{R})$. We will now show how our refinements of the holonomic approximation lemma can be applied to this situation.

Denote by $\pi: \Lambda_n(W) \to W$ the Grassmannian bundle of Lagrangian planes in TW. The fiber of π over a point $x \in W$ consists of the linear Lagrangian subspaces of the symplectic vector space $(T_x W, \omega_x)$. The Gauss map $G(df): L \to \Lambda_n(W)$ of the embedding f is defined by $G(df)(q) = df(T_q M) \subset T_{f(q)} W$.

Definition 1.14   A tangential rotation of f is a compactly supported deformation $G_t: L \to \Lambda_n(W)$, $t \in [0, 1]$, of the Gauss map $G_0 = G(df)$ such that $\pi \circ G_t = f$.

By compactly supported we mean that $G_t = G(df)$ for all $t \in [0, 1]$ outside of a compact subset of L. The following approximation result is a simple corollary of our holonomic approximation lemma for l–holonomic sections.

Theorem 1.15   *Let $K \subset L$ be a polyhedron of positive codimension, and let $G_t: L \to \Lambda_n(M)$ be a tangential rotation of a Lagrangian embedding $f: L \to M$. Then there exists a compactly supported ambient Hamiltonian isotopy $\varphi_t: W \to W$ such that $G(d(\varphi_t \circ f))$ is C^0–close to G_t on $\mathrm{Op}(K)$.*
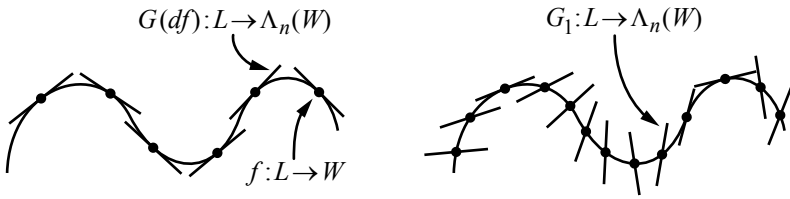
$$G(df): L \to \Lambda_n(W) \qquad G_1: L \to \Lambda_n(W)$$

$$f: L \to W$$

Figure 5: A tangential rotation G_t of a Lagrangian embedding f

Remark 1.16 (1) We can take φ_t to be C^0–close to the identity id_M on all of M.

(2) Moreover, we can also arrange it so that $\varphi_t = \mathrm{id}_M$ outside of an arbitrarily small neighborhood of $f(L)$ in M.

(3) The statement holds in relative form. Namely, if $G_t = G(df)$ on $\mathrm{Op}(A) \subset L$ for some closed subset $A \subset L$, then we can take φ_t so that $\varphi_t = \mathrm{id}_M$ on $\mathrm{Op}(f(A)) \subset M$.

(2) An analogous approximation result holds for tangential rotations of Legendrian embeddings into a contact manifold (W^{2n+1}, ξ). In this case $\pi: \Lambda_n(W) \to W$ is the Grassmannian bundle whose fiber over a point $x \in W$ consists of the linear Lagrangian subspaces of the symplectic vector space $(\xi_x^{2n}, d\alpha_x)$, where $\xi = \ker(\alpha)$ near the point x.

Proof Let $\mathcal{N} \subset W$ be a Weinstein neighborhood of $f(L)$. If we fix a Riemannian metric on L we can choose $\mathcal{N}$ symplectomorphic to $T_\delta^* L$ for some $\delta > 0$, where $T_\delta^* L$ consists of the cotangent vectors $p \in T^* L$ such that $\|p\| < \delta$.

For small time t we can think of the tangential rotation G_t as a family of sections $s_t: L \to J^2(L, \mathbb{R})$ such that $s_0 = 0$ and $s_t^{(1)} = 0$. In fact, by first subdividing the time interval finely enough, we can reduce to the case where s_t is defined for all $t \in [0, 1]$. The point here is that the planes $G_t(q) \subset T_{f(q)} W \simeq T_{(q,0)}(T^* L)$ must remain graphical over $T_q L \subset T_{(q,0)}(T^* L)$.

The parametric version of the holonomic approximation lemma for l–holonomic sections can be applied to produce an isotopy $F_t: L \to L$ and a family of functions $h_t: L \to \mathbb{R}$, $h_0 = 0$, such that $j^2(h_t)$ is C^0–close to s_t on $\mathrm{Op}(F_t(K))$ and such that $j^1(h_t)$ is C^0–small on all of L. In particular, we may assume that $\|dh_t\| < \delta$ for all $t \in [0, 1]$. Hence we can think of the composition $f_t = dh_t \circ F_t: L \to T_\delta^* L$ as a compactly supported exact homotopy of Lagrangian embeddings $f_t: L \to W$. Every such homotopy is induced by a compactly supported ambient Hamiltonian isotopy $\varphi_t: M \to M$ satisfying the required properties. $\qquad \square$
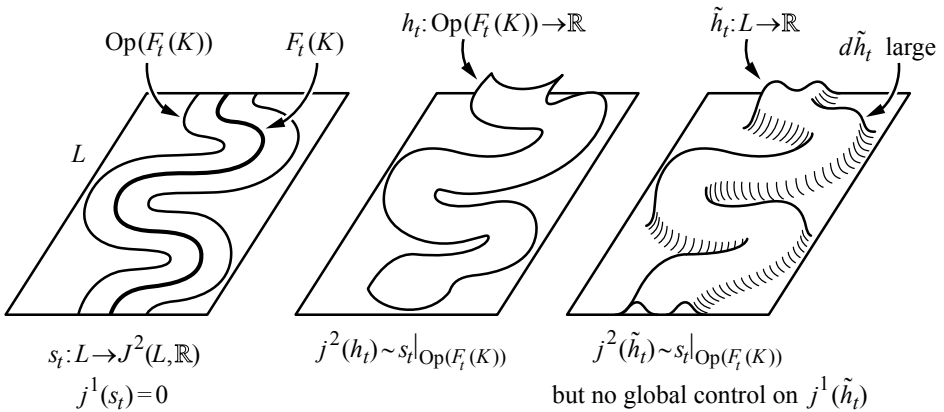
Figure 6: The potential problem with the cutting off of h_t after a naive application of the classical holonomic approximation lemma

If we attempt to prove Theorem 1.15 using the classical holonomic approximation lemma (Theorem 1.2) instead, we run into the following difficulty. The functions h_t produced by the holonomic approximation would a priori only be defined in open subsets $\operatorname{Op}(F_t(K)) \subset L$. We would therefore need to extend h_t to the whole of L by hand. The most straightforward way of doing so is to choose a family of cutoff functions $\psi_t \colon L \to \mathbb{R}$ supported on the domain of h_t such that $\psi_t = 1$ near $F_t(K)$. The product $\widetilde{h}_t = \psi_t \cdot h_t \colon L \to \mathbb{R}$ is then well defined on all of L. It follows that the composition $f_t = d\widetilde{h}_t \circ F_t \colon L \to T^*L$ is an exact homotopy of Lagrangian embeddings whose Gauss map provides the desired approximation near K.

If $\|d\widetilde{h}_t\| < \delta$ on all of L, then f_t has image contained in $T^*_\delta L$ and we can think of f_t as an exact homotopy of Lagrangian embeddings $f_t \colon L \to W$ as before. Observe, however, that there is no guarantee that $d\widetilde{h}_t = \psi_t dh_t + h_t d\psi_t$ has norm $< \delta$, because ψ_t will typically have a very large derivative. Indeed, the wiggling by F_t is quite dramatic; see Figure 6 for an illustration. Therefore, $f_t(L)$ might escape our Weinstein neighborhood $T^*_\delta L \simeq \mathcal{N}$. Hence f_t does not correspond to an exact homotopy of Lagrangian embeddings into W and our proof breaks down. Our holonomic approximation lemma for l–holonomic sections with $l = 1$ precisely provides the necessary global control on the 1–jet component so that this issue does not arise.

The parametric version of Theorem 1.15 also holds and is proved in the same way.

Theorem 1.17 Let $K \subset Z \times L$ be a fibered polyhedron of positive codimension and let $G^z_t \colon L \to \Lambda_n(M)$ be a family of tangential rotations of Lagrangian embeddings $f^z \colon L \to M$ parametrized by a compact manifold Z such that $G^z_t = G(df^z)$

for $z \in \mathrm{Op}(\partial Z)$. *Then there exists a family of compactly supported ambient Hamiltonian isotopies* $\varphi_t^z \colon W \to W$ *such that* $G(d(\varphi_t^z \circ f^z))$ *is* C^0–*close to* G_t^z *on* $\mathrm{Op}(K)$ *and such that* $\varphi_t^z = \mathrm{id}_M$ *for* $z \in \mathrm{Op}(\partial Z)$.

Remark 1.18 Analogous observations to the ones made in Remark 1.16 apply.

The following h–principle for directed embeddings follows immediately from the above approximation results. First, we recall the following definition of Gromov.

Definition 1.19 Given subsets $D \subset \Lambda_n(W)$ and $S \subset L$, we say that a Lagrangian embedding $f \colon L \to W$ is D–*directed along* S if $G(df)(S) \subset D$.

Theorem 1.20 *Let* $f \colon L \to W$ *be a Lagrangian embedding, let* $K \subset L$ *be a polyhedron of positive codimension and let* $D \subset \Lambda_n(W)$ *be an open subset. Suppose that there exists a tangential rotation* $G_t \colon L \to \Lambda_n(W)$ *of* f *such that* $G_1(K) \subset D$. *Then there exists a compactly supported Hamiltonian isotopy* $\varphi_t \colon W \to W$ *such that* $\varphi_1 \circ f$ *is* D–*directed along* $\mathrm{Op}(K)$.

Remark 1.21 (1) This h–principle also holds in C^0–close, relative and parametric versions. We leave it to the reader to formulate the appropriate statements.

(2) We can choose the Hamiltonian isotopy so that $\varphi_t = \mathrm{id}_W$ outside of an arbitrarily small neighborhood of $f(K)$ in W.

(3) An analogous h–principle holds for Legendrian embeddings into a contact manifold (W^{2n+1}, ξ) that are D–directed along a polyhedron of positive codimension.

Analogous problems in geometric topology have been studied by several authors. In [8], Gromov proved an h–principle for D–directed smooth embeddings of an open manifold into some ambient manifold which holds for any open subset D of the Grassmannian of the ambient manifold. See Spring [14] for a discussion of Gromov's argument. Rourke and Sanderson gave two independent proofs of this result in [10; 11]. Another proof was obtained by Eliashberg and Mishachev [3] using their holonomic approximation lemma. For embeddings of a closed manifold one cannot hope to prove an h–principle for D–directed embeddings when D is an arbitrary open subset of the Grassmannian. However, for certain special subsets D, called ample, Gromov proved in [7; 8] that an h–principle does hold. In a different direction, Eliashberg and Mishachev showed in [5] that an h–principle for D–directed embeddings of a closed manifold hold for an

arbitrary open D, but provided that we relax the notion of an embedding to that of a wrinkled embedding. In [1] we prove a symplectic and contact analogue of this last result using the tools developed in the present paper in a crucial way.

Example 1.22　Let $\eta \subset TW$ be a distribution of k–planes in a symplectic or contact manifold W. Set $D_x = \{P_x \in \Lambda_n(W)_x : P_x \pitchfork \eta_x\}$ for each $x \in W$. Then $D = \bigcup_{x \in W} D_x$ is an open subset of $\Lambda_n(W)$. We obtain a full h–principle for Lagrangian or Legendrian embeddings which are transverse to an ambient distribution near a given subset of positive codimension. In the particular case where $\eta = \ker(d\rho)$ for $\rho \colon W \to B$ a Lagrangian or Legendrian fibration, we can rephrase the result as an h–principle for Lagrangian or Legendrian embeddings whose front is nonsingular along a given subset of positive codimension.

The main application (and source of motivation) for our holonomic approximation lemma for $\perp$–holonomic sections, as well as its parametric version, is given in [1]. The control on the $\perp$–jet component is a key ingredient in the proof of the h–principle for the (global) simplification of singularities of Lagrangian and Legendrian fronts. Indeed, when attempting to apply a holonomic approximation argument near the $(n-1)$–skeleton of a Lagrangian or Legendrian submanifold, difficulties similar to the one illustrated in the proof of Theorem 1.15 above inevitably arise. The situation is in fact much more subtle because we need to respect a certain decomposition of a tangential rotation into so-called simple tangential rotations. Theorems 1.8 and 1.12 provide the precise control needed to make the proof work.

1.5　Idea of the proof

The strategy of proof is to carry out a sequence of reductions which simplify our refined holonomic approximation lemmas for l– and $\perp$–holonomic sections to a problem described by a concrete local model. We can then keep track of the geometry behind the holonomic approximation process in this carefully chosen model and establish the necessary estimates to achieve the desired global control. The outline of the paper is roughly as follows.

In Section 2 we reduce our global results to the local relative statements corresponding to the jet space $J^r(\mathbb{R}^m, \mathbb{R}^n)$ over the unit cube $I^m = [-1, 1]^m$. In Section 3 we study the space $J^r(\mathbb{R}^m, \mathbb{R}^n)$ and reduce the holonomic approximation lemma for l–holonomic sections to the holonomic approximation lemma for $\perp$–holonomic sections. For a section which is $\perp$–holonomic with respect to a hyperplane field τ, we construct in

Section 5 a holonomic approximation with controlled ⊥–component by wiggling the polyhedron K in a way such that the wiggles are parallel to the hyperplanes in τ. However, we cannot implement such a wiggling near the region where τ is almost tangent to K. A preliminary adjustment is therefore necessary in this region. This adjustment is performed in Section 4.

We should note that it is possible to prove the holonomic approximation lemma for l–holonomic sections in a more direct manner. One can extend by hand the holonomic approximation formulas written down in [4] or appeal to abstract extension results to reach the desired conclusion. However, we choose to deduce the holonomic approximation lemma for l–holonomic sections as a corollary of the holonomic approximation lemma for ⊥–holonomic sections.

The main reason for doing so is that such a reduction involves decompositions of $(r-1)$–holonomic sections into so-called primitive sections, which we define in Section 2.1. A similar strategy appears in Gromov's work [8] and is developed further in Spring's book [13], related to the construction of iterated convex hull extensions in the theory of convex integration. Moreover, primitive sections play a crucial role in the proof of our h–principle for the simplification of singularities of Lagrangian and Legendrian fronts [1], where they correspond to a particularly simple type of tangential rotation. We hope that the general idea of working one pure partial derivative at a time may have further applications to the philosophy of the h–principle and so have attempted to present the elements of the strategy as clearly as possible.

2 Localization of the problem

2.1 Holonomic trivialization

For a general fiber bundle $p\colon X^{m+n} \to M^m$, the bundle of r–jets $p^r\colon X^{(r)} \to M$ can be messy to work with globally. However, global h–principle type problems can often

be reduced to a local relative statement. In this section we explain how this reduction is accomplished for our refined holonomic approximation lemmas. We choose to work with the unit cube $I^m = [-1, 1]^m \subset \mathbb{R}^m$ as our local model. In what follows we use the language of l– and $\perp$–holonomic sections introduced in Section 1.2. We begin by recalling from [4] the following simple but crucial observation.

Remark 2.1 (holonomic trivialization) Let $\hat{s}\colon M \to X^{(r)}$ be a holonomic section of p^r, and let $Q \subset M$ be an embedded cube $Q \simeq I^m$. Then there exists a neighborhood $\mathcal{N} \subset X^{(r)}$ of the image $\mathrm{Im}(\hat{s})$ such that $(p^r)^{-1}(Q) \cap \mathcal{N} \simeq J^r(\mathbb{R}^m, \mathbb{R}^n)|_{I^m}$.

Proof Since the section $\hat{s}$ is holonomic, we have $\hat{s} = j^r(h)$ for some section $h\colon M \to X$ of p. Observe that the fibration $p\colon X \to M$ is trivial over the contractible subset Q. Hence a neighborhood of the image $h(Q) \subset X$ in $p^{-1}(Q)$ is diffeomorphic to $Q \times \mathbb{R}^n$. It follows that a neighborhood of the image $\hat{s}(Q) \subset X^{(r)}$ in $(p^r)^{-1}(Q)$ is diffeomorphic to $J^r(Q, \mathbb{R}^n) \simeq J^r(\mathbb{R}^m, \mathbb{R}^n)|_{I^m}$. $\qquad\square$

Under the above identification, sections $s\colon M \to \mathcal{N} \subset X^{(r)}$ such that $s = \hat{s}$ on $\mathrm{Op}(M \setminus \mathrm{int}(Q))$ correspond to sections $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$. The section s is holonomic if and only if the section σ is holonomic. The section $\hat{s}$ itself corresponds to the zero section $\sigma = 0$.

Similarly, l–holonomic sections $s\colon M \to \mathcal{N} \subset X^{(r)}$ such that $s = \hat{s}$ on $\mathrm{Op}(M \setminus \mathrm{int}(Q))$ and such that $s^{(l)} = \hat{s}^{(l)}$ on all of M correspond to sections $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$ and such that $\sigma^{(l)} = 0$ on all of I^m.

Fix a hyperplane field $\tau \subset TM$. Then $\perp$–holonomic sections $s\colon M \to \mathcal{N} \subset X^{(r)}$ such that $s = \hat{s}$ on $\mathrm{Op}(M \setminus \mathrm{int}(Q))$ and such that $s^\perp = \hat{s}^\perp$ on all of M (with respect to τ) correspond to sections $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$ and such that $\sigma^\perp = 0$ on all of I^m (with respect to the hyperplane field associated to τ under the identification $Q \simeq I^m$). This last remark motivates the following definition, which we will use repeatedly in what follows.

Definition 2.2 A section $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ is called primitive with respect to a hyperplane field $\tau \subset TI^m$ if $\sigma^\perp = 0$.

2.2 The local relative statements

The global versus local dictionary described in the previous subsection leads us to formulate the following local relative versions of our main results. We first state the nonparametric versions.

Theorem 2.3 (localized holonomic approximation lemma for l–holonomic sections)
Fix $k < m$. Let $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a section such that

- $\sigma = 0$ on $\operatorname{Op}(\partial I^m)$,
- $\sigma^{(l)} = 0$ on all of I^m for some $l < r$.

Then there exists an isotopy $F_t\colon I^m \to I^m$ and a holonomic section $\hat{\sigma}\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that the following properties hold:

- $\hat{\sigma}$ is C^0–close to σ on $\operatorname{Op}(F_1(I^k))$.
- $\hat{\sigma}^{(l)}$ is C^0–small on all of I^m.
- F_t is C^0–small.
- $F_t = \operatorname{id}_{I^m}$ and $\hat{\sigma} = 0$ on $\operatorname{Op}(\partial I^m)$.

Theorem 2.4 (localized holonomic approximation lemma for $\perp$–holonomic sections)
Fix $k < m$. Let $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a section such that

- $\sigma = 0$ on $\operatorname{Op}(\partial I^m)$,
- $\sigma^\perp = 0$ on all of I^m with respect to some hyperplane field $\tau \subset TI^m$.

Then there exists an isotopy $F_t\colon I^m \to I^m$ and a holonomic section $\hat{\sigma}\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that the following properties hold:

- $\hat{\sigma}$ is C^0–close to σ on $\operatorname{Op}(F_1(I^k))$.
- $\hat{\sigma}^\perp$ is C^0–small on all of I^m.
- F_t is C^0–small.
- $F_t = \operatorname{id}_{I^m}$ and $\hat{\sigma} = 0$ on $\operatorname{Op}(\partial I^m)$.

The global holonomic approximation lemmas for l– and $\perp$–holonomic sections, Theorems 1.5 and 1.8, follow from the local relative statements, Theorems 2.3 and 2.4, by induction over the skeleton of the polyhedron K, working one cube at a time. At each step we use the holonomic trivialization (Remark 2.1) to reduce the global problem to a local problem. Observe that the relative versions of the global holonomic approximation lemmas for l– and $\perp$–holonomic sections (see Remark 1.6) also follow from the above local relative statements.

Similarly, the parametric global holonomic approximation lemmas, Theorems 1.11 and 1.12, including the corresponding relative versions, follow from the parametric local relative statements phrased below. In this case we also localize with respect to the parameter space, setting $Z = I^q$, the unit q–dimensional cube $[-1, 1]^q$.

Theorem 2.5 (parametric localized holonomic approximation lemma for l–holonomic sections) *Fix $k < m$. Let $\sigma_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a family of sections parametrized by the unit cube I^q such that*

- $\sigma_z = 0$ *on* $\mathrm{Op}(\partial I^m)$,
- $\sigma_z^{(l)} = 0$ *on all of I^m for some $l < r$,*
- $\sigma_z = 0$ *on all of I^m for $z \in \mathrm{Op}(\partial I^q)$.*

Then there exists a family of isotopies $F_t^z \colon I^m \to I^m$ and a family of holonomic sections $\hat{\sigma}_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that the following properties hold:

- $\hat{\sigma}_z$ *is C^0–close to σ_z on $\mathrm{Op}(F_1^z(I^k))$.*
- $\hat{\sigma}_z^{(l)}$ *is C^0–small on all of I^m.*
- F_t^z *is C^0–small.*
- $F_t^z = \mathrm{id}_{I^m}$ *and $\hat{\sigma}_z = 0$ on $\mathrm{Op}(\partial I^m)$.*
- $F_t^z = \mathrm{id}_{I^m}$ *and $\hat{\sigma}_z = 0$ on all of I^m for $z \in \mathrm{Op}(\partial I^q)$.*

Theorem 2.6 (parametric localized holonomic approximation lemma for $\perp$–holonomic sections) *Fix $k < m$. Let $\sigma_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a family of sections parametrized by the unit cube I^q such that*

- $\sigma_z = 0$ *on* $\mathrm{Op}(\partial I^m)$,
- $\sigma_z^{\perp} = 0$ *on all of I^m with respect to some family of hyperplane fields $\tau_z \subset TI^m$,*
- $\sigma_z = 0$ *on all of I^m for $z \in \mathrm{Op}(\partial I^q)$.*

Then there exists a family of isotopies $F_t^z \colon I^m \to I^m$ and a family of holonomic sections $\hat{\sigma}_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that the following properties hold:

- $\hat{\sigma}_z$ *is C^0–close to σ_z on $\mathrm{Op}(F_1^z(I^k))$.*
- $\hat{\sigma}_z^{\perp}$ *is C^0–small on all of I^m.*
- F_t^z *is C^0–small.*
- $F_t^z = \mathrm{id}_{I^m}$ *and $\hat{\sigma}_z = 0$ on $\mathrm{Op}(\partial I^m)$.*
- $F_t^z = \mathrm{id}_{I^m}$ *and $\hat{\sigma}_z = 0$ on all of I^m for $z \in \mathrm{Op}(\partial I^q)$.*

The remainder of the paper is devoted to the proofs of the local relative theorems, Theorems 2.3, 2.4, 2.5 and 2.6.

3 Geometry of jet spaces

3.1 Jets as Taylor polynomials

The reduction carried out in Section 2 leads us to study the local space $J^r(\mathbb{R}^m, \mathbb{R}^n)$. We begin by giving an explicit description of this space in terms of Taylor polynomials. This description is useful both for intuition and for the explicit computations to be carried out later on.

Given a point $x \in \mathbb{R}^m$ and given n real polynomials $p_1(X), \ldots, p_n(X) \in \mathbb{R}[X]$ in m variables $X = (X_1, \ldots, X_m)$ of degree $\leq r$, set $s(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ to be the r–jet at x of the germ

$$y \mapsto (p_1(y-x), \ldots, p_n(y-x)) \in \mathbb{R}^n, \quad y \in \mathrm{Op}(x) \subset \mathbb{R}^m.$$

This assignment yields a trivialization $J^r(\mathbb{R}^m, \mathbb{R}^n) \simeq \mathbb{R}^m \times (\mathcal{P}_r)^n$ given by $s(x) \leftrightarrow (x, p_1(X), \ldots, p_n(X))$, where $\mathcal{P}_r = \{p(X) \in \mathbb{R}[X] : \deg(p(X)) \leq r\}$. Indeed, if $h \colon \mathrm{Op}(x) \subset \mathbb{R}^m \to \mathbb{R}^n$ is the germ of a smooth function at the point $x \in \mathbb{R}^m$ and $(p_1(X), \ldots, p_n(X))$ is its linear Taylor approximation of order r centered at x, then the above construction yields $s(x) = j^r(h)(x)$. In this way we think of an arbitrary section $s \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ as a family $s(x)$ of Taylor polynomials of degree $\leq r$ parametrized by the point $x \in I^m$.

Observe that for $l < r$ we obtain an induced trivialization $(p_l^r)^{-1}(0) \simeq \mathbb{R}^m \times (\mathcal{P}_{l,r})^n$, where we recall the fiber bundle $p_l^r \colon J^r(\mathbb{R}^m, \mathbb{R}^n) \to J^l(\mathbb{R}^m, \mathbb{R}^n)$ given by $\sigma \mapsto \sigma^{(l)}$, and we denote by $\mathcal{P}_{l,r} \subset \mathcal{P}_r$ the space of real polynomials in m variables $X = (X_1, \ldots, X_m)$ which are sums of monomials of degree strictly greater than l and at most r. If we further denote by $\mathcal{H}_j = \mathcal{P}_{j-1,j}$ the space of homogeneous polynomials in m variables $X = (X_1, \ldots, X_m)$ of degree exactly j, then from the degree splitting $\mathcal{P}_r = \mathcal{H}_0 \times \mathcal{H}_1 \times \cdots \times \mathcal{H}_r$, we obtain an induced decomposition $J^r(\mathbb{R}^m, \mathbb{R}^n) \simeq \mathbb{R}^m \times (\mathcal{H}_0)^n \times (\mathcal{H}_1)^n \times \cdots \times (\mathcal{H}_r)^n$ into homogeneous components. By the homogeneous component of order j of an r–jet $s(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ at $x \in \mathbb{R}^m$ we will mean the $(\mathcal{H}_j)^n$ entry corresponding to $s(x)$ under the above decomposition. The projection p_l^r simply forgets the homogeneous components of degree $> l$, and so we have a similar decomposition $(p_l^r)^{-1}(0) \simeq \mathbb{R}^m \times (\mathcal{H}_{l+1})^n \times \cdots \times (\mathcal{H}_r)^n$.

We next consider the trivialization $J^r(\mathbb{R}^m, \mathbb{R}^n) \simeq \mathbb{R}^m \times (\mathcal{P}_r)^n$ in the context of primitive sections, as defined in Section 2.1. Let $\tau \subset T\mathbb{R}^m$ be a hyperplane field. We can specify a coorientation of τ by choosing a family of unit vectors $u_x \in \mathbb{R}^m$ which are orthogonal to τ_x with respect to the Euclidean inner product $\langle \cdot, \cdot \rangle$. Set $l_x \colon \mathbb{R}^m \to \mathbb{R}$ to be the linear

function $l_x(\cdot) = \langle \cdot, u_x \rangle$. Given a point $x \in \mathbb{R}^m$ and a vector $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, define $s(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ to be the r–jet at the point $x \in \mathbb{R}^m$ of the germ

$$y \mapsto (l_x(y - x))^r \cdot (v_1, \ldots, v_n) \in \mathbb{R}^n, \quad y \in \mathrm{Op}(x) \subset \mathbb{R}^m.$$

In other words, we repeat the above construction $s(x) \leftrightarrow (x, p_1(X), \ldots, p_n(X))$ with the polynomials $p_j(X) = v_j \cdot (l_x(X))^r$, which are all multiples of the rth power of a linear function with kernel τ_x. Observe that the resulting r–jet satisfies $s(x)^\perp = 0$ with respect to the hyperplane field τ_x. In fact, the choice of coorientation u_x determines a trivialization of the space of sections which are primitive with respect to τ, namely $(p_\perp^r)^{-1}(0) \simeq \mathbb{R}^m \times \mathbb{R}^n$, $s(x) \leftrightarrow (x, v)$, where we recall the fibration $p_\perp^r \colon J^r(\mathbb{R}^m, \mathbb{R}^n) \to J^\perp(\mathbb{R}^m, \mathbb{R}^n)$, $\sigma \mapsto \sigma^\perp$.

3.2 Linear structure

For a general fiber bundle $p \colon X \to M$ we have fiber bundle structures $p_l^r \colon X^{(r)} \to X^{(l)}$ for each $l < r$, but in general we cannot invariantly define inclusions $X^{(l)} \subset X^{(r)}$. The reason is that the chain rule for derivatives of order r involves all derivatives of order $\leq r$ and therefore a change of coordinates will mix up the jet components of different orders.

Nevertheless, in the case $X^{(r)} = J^r(\mathbb{R}^m, \mathbb{R}^n)$ we have a canonical linear structure arising from the linear structure on $\mathbb{R}^n$. Explicitly, if $s(x), \sigma(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ are r–jets at the point $x \in \mathbb{R}^m$ corresponding to germs $h, g \colon \mathrm{Op}(x) \subset \mathbb{R}^m \to \mathbb{R}^n$ and if $a, b \in \mathbb{R}$ are any two real numbers, then we can define $a \cdot s(x) + b \cdot \sigma(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ to be the r–jet at the point $x \in \mathbb{R}^m$ corresponding to the germ $ah + bg \colon \mathrm{Op}(x) \subset \mathbb{R}^m \to \mathbb{R}^n$. We can therefore equip the r–jet bundle $p^r \colon J^r(\mathbb{R}^m, \mathbb{R}^n) \to \mathbb{R}^m$ with the structure of a vector bundle. Similarly, for all $l < r$ we can endow each of the projections $p_l^r \colon J^r(\mathbb{R}^m, \mathbb{R}^n) \to J^l(\mathbb{R}^m, \mathbb{R}^n)$ with vector bundle structures. However, we will reserve the addition sign to denote the linear structure on the bundle p^r. In terms of the trivialization $J^r(\mathbb{R}^m, \mathbb{R}^n) \simeq \mathbb{R}^m \times (\mathcal{P}_r)^n$, this linear structure corresponds to addition of polynomials in $\mathcal{P}_r$.

3.3 Reduction to the $l = r - 1$ case

In order to deduce the localized holonomic approximation lemma for l–holonomic sections (Theorem 2.3) as a corollary of the localized holonomic approximation lemma for $\perp$–holonomic sections (Theorem 2.4), it is useful to first reduce to the case $l = r - 1$. This reduction is accomplished by the following inductive argument.

Lemma 3.1 *Suppose that there exists $1 < j \leq r$ such that Theorem 2.3 holds for all r and l such that $r - l < j$. Then it also holds for all r and l such that $r - l < j + 1$.*

Remark 3.2 Before we dive into the proof, we recall the notion of pullbacks and pushforwards in jet spaces. Suppose that $F: \mathbb{R}^m \to \mathbb{R}^m$ is a diffeomorphism, $F(x) = y$ and $h: \mathrm{Op}(y) \to \mathbb{R}^n$ is a germ of a smooth function at the point $y \in \mathbb{R}^m$. Then $h \circ F: \mathrm{Op}(x) \to \mathbb{R}^n$ is a germ of a smooth function at $x \in \mathbb{R}^m$. This assignment defines a pullback map $F^*: J^r(\mathbb{R}^m, \mathbb{R}^n) \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ which covers F^{-1}. Similarly, we define the pushforward $F_* = (F^{-1})^*: J^r(\mathbb{R}^m, \mathbb{R}^n) \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ which covers F.

Proof of Lemma 3.1 Let $\sigma: I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a section such that $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$ and $\sigma^{(r-j)} = 0$ on all of I^m. Let $\mu = \sigma^{(r-1)}: I^m \to J^{r-1}(\mathbb{R}^m, \mathbb{R}^n)$ be its $(r-1)$–jet component. Then we also have $\mu = 0$ on $\mathrm{Op}(\partial I^m)$ and $\mu^{(r-j)} = 0$ on all of I^m. Observe that $r - j = (r-1) - (j-1)$ and therefore by assumption there exists a C^0–small isotopy $H_t: I^m \to I^m$ such that $H_t = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and a holonomic section $\widehat{\mu}: I^m \to J^{r-1}(\mathbb{R}^m, \mathbb{R}^n)$ such that $\widehat{\mu}$ is C^0–close to μ on $\mathrm{Op}(H_1(I^k))$, such that $\widehat{\mu}^{(r-j)}$ is C^0–small and such that $\mu = 0$ on $\mathrm{Op}(\partial I^m)$.

Since $\widehat{\mu}$ is holonomic, we have $\widehat{\mu} = j^{r-1}(h)$ for some function $h: I^m \to \mathbb{R}^m$. There exists a unique section $\nu: I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\nu^{(r-1)} = 0$ and such that the homogeneous order-r component of ν is equal to the homogeneous order-r component of the section $\sigma - j^r(h)$. Observe that the pullback $(H_1)^* \nu$ by the diffeomorphism H_1 also has zero $(r-1)$–jet component. We can therefore apply once again our inductive hypothesis (which includes the case $l = r - 1$ since $j > 1$) to ensure the existence of a C^0–small isotopy $\widetilde{H}_t: I^m \to I^m$ such that $\widetilde{H}_t = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and a holonomic section $\widehat{\nu}: I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\widehat{\nu}$ is C^0–close to $(H_1)^* \nu$ on $\mathrm{Op}(\widetilde{H}_1(I^k))$, such that $\widehat{\nu}^{(r-1)}$ is C^0–small and such that $\widehat{\nu} = 0$ on $\mathrm{Op}(\partial I^m)$.

Set $F_t = H_t \circ \widetilde{H}_t$ and $\widehat{\sigma} = j^r(h) + (H_1)_*(\widehat{\nu})$. Then we can rephrase our above conclusions by stating that F_t is C^0–small isotopy such that $F_t = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and that $\widehat{\sigma}$ is a holonomic section of $J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\widehat{\sigma}$ is C^0–close to σ on $\mathrm{Op}(F_1(I^m))$, such that $\widehat{\sigma}^{(r-j)}$ is C^0–small and such that $\widehat{\sigma} = 0$ on $\mathrm{Op}(\partial I^m)$. This is exactly what we wanted. $\square$

Lemma 3.1 also holds in families, with the same proof. The only difference is that one needs to add a parameter everywhere in the notation. Therefore, it also suffices to prove the parametric local relative Theorem 2.5 in the case $l = r - 1$.

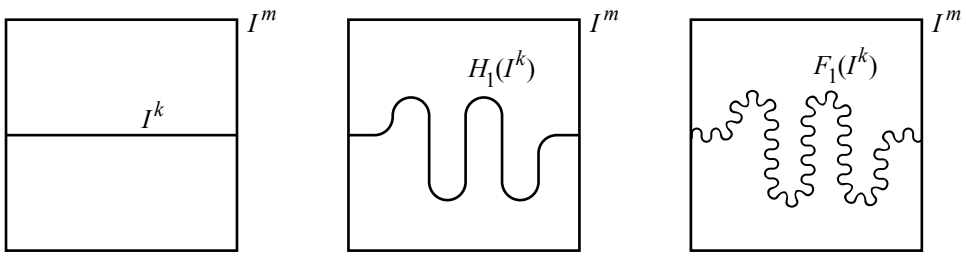Figure 7: Wiggling the wiggles with $F_t = H_t \circ \tilde{H}_t$

3.4 Decomposition into primitive sections

To reduce the $l = r - 1$ case of the localized holonomic approximation lemma for l–holonomic sections to the localized holonomic approximation lemma for $\perp$–holonomic sections we need to consider decompositions of r–jet sections with zero $(r-1)$–jet component into sums of primitive sections. The following discussion closely resembles the theory of principal decompositions in jet spaces invented by Gromov in [8] in the context of convex integration and further fleshed out by Spring in [13].

Given a fixed holonomic section $\hat{s} \colon M \to X^{(r)}$ of the r–jet bundle of a fiber bundle $p \colon X \to M$, recall that primitive sections are the local analogues of $\perp$–holonomic sections whose $\perp$–jet equals $\hat{s}^{\perp}$. We repeat the precise definition for convenience.

Definition 3.3 A section $\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ is called primitive with respect to a hyperplane field $\tau \subset TI^m$ if $\sigma^{\perp} = 0$.

We are particularly interested in sections which are primitive with respect to the hyperplane fields $\tau_\beta = \ker(dx_{\beta_1} + \cdots + dx_{\beta_k})$, where β ranges over all multi-indices such that $1 \le \beta_j \le m$ and such that $|\beta| = k \le r$, up to permutation. There are of course redundancies among the τ_β, but this is not important. We remark that the hyperplane fields τ_β are constant and hence integrable. Given a section $\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\sigma^{(r-1)} = 0$, our goal is to obtain a decomposition $\sigma = \sum_\beta \sigma_\beta$ where each section σ_β is primitive with respect to τ_β; see Figure 8. Moreover, we want this decomposition to be well behaved in a sense that is made precise below.

For this purpose we invoke the following simple polynomial identity, which the author found in [2] but which may well be classical. Consider the formula

$$X_1 X_2 \cdots X_r = \frac{(-1)^r}{r!} \sum_U (-1)^{|U|} \left(\sum_{u \in U} X_u \right)^r,$$
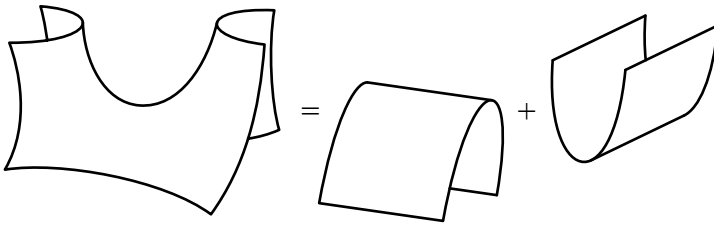
Figure 8: We can decompose the homogeneous degree-2 polynomial $x^2 - y^2$ into the sum of the monomials x^2 and $-y^2$.

where the sum ranges over all subsets $U \subset \{1, 2, \ldots, r\}$. Recall from Section 3.1 that we can think of an r–jet $\sigma(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ at the point $x \in I^m$ such that $\sigma^{(r-1)}(x) = 0$ as a homogeneous Taylor polynomial $(p_1(X), \ldots, p_n(X))$ of degree r centered at x. Each such polynomial $p_j(X)$ can be written uniquely as a sum of monomials: $p_j(X) = \sum_\alpha a_\alpha X_{\alpha_1} \cdots X_{\alpha_r}$, where α ranges through all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_r)$ such that $1 \leq \alpha_j \leq m$, up to permutation. Hence we can write

$$p_j(X) = \sum_\alpha a_\alpha X_{\alpha_1} \cdots X_{\alpha_r} = \sum_\alpha \frac{(-1)^r a_\alpha}{r!} \sum_U (-1)^{|U|} \left(\sum_{u \in U} X_{\alpha_u} \right)^r$$

$$= \sum_\beta \left(\sum_{(\alpha_u) = \beta} \frac{(-1)^{r+|\beta|} a_\alpha}{r!} \right) (X_{\beta_1} + \cdots + X_{\beta_k})^r,$$

where the inner sum ranges over all pairs (α, U) such that $(\alpha_u)_{u \in U} = \beta$. Observe that the homogeneous degree-r monomial $(X_{\beta_1} + \cdots + X_{\beta_k})^r$ corresponds to an r–jet which is primitive with respect to $\tau_\beta = \ker(dx_{\beta_1} + \cdots + dx_{\beta_k})$. We have therefore proved the following.

Lemma 3.4 *Given a section* $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ *such that* $\sigma^{(r-1)} = 0$, *we can write* $\sigma = \sum_\beta \sigma_\beta$ *for sections* $\sigma_\beta\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ *such that the following properties hold:*

- *Each section* σ_β *is primitive with respect to* τ_β.

- *Each section* σ_β *depends smoothly on* σ.

- *If* $\sigma = 0$ *on* $\mathrm{Op}(A)$ *for some closed subset* $A \subset I^m$, *then* $\sigma_\beta = 0$ *on* $\mathrm{Op}(A)$ *for all* β.

Remark 3.5 Observe that the number of indices β appearing in the sum only depends on m and r.

3.5 Reduction to the primitive case

We are now ready to reduce the localized holonomic approximation lemma for l–holonomic sections (Theorem 2.3) to the localized holonomic approximation lemma for $\perp$–holonomic sections (Theorem 2.4). Recall first that by the discussion of Section 3.3, we only need to prove Theorem 2.3 in the case $l = r - 1$. Let us therefore assume that Theorem 2.4 holds and let $\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a section such that $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$ and such that $\sigma^{(r-1)} = 0$ on all of I^m. Lemma 3.4 gives us a decomposition $\sigma = \sum_\beta \sigma_\beta$, where $\sigma_\beta \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ is a section such that $\sigma_\beta = 0$ on $\mathrm{Op}(\partial I^m)$ and such that $\sigma_\beta^\perp = 0$ with respect to the hyperplane field $\tau_\beta \subset T I^m$ defined in Section 3.4. We will inductively construct holonomic approximations for the partial sums of the decomposition $\sigma = \sum_\beta \sigma_\beta$. The main point in the following argument is that if an r–jet section has a C^0–small $\perp$–jet component, then in particular it also has a C^0–small $(r-1)$–jet component.

Let $\beta^1, \beta^2, \ldots, \beta^N$ be an ordering of the multi-indices β appearing in the decomposition $\sigma = \sum_\beta \sigma_\beta$ and denote by $\sigma_1, \ldots, \sigma_N$ and $\tau_1, \ldots, \tau_N$ the corresponding sections σ_β and hyperplane fields τ_β. We begin by applying Theorem 2.4 to the section σ_1. We obtain a C^0–small isotopy $F_t^1 \colon I^m \to I^m$ such that $F_t^1 = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and a holonomic section $\hat{\sigma}_1 \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\hat{\sigma}_1$ is C^0–close to σ_1 on $\mathrm{Op}(F_1(I^k))$, such that $\hat{\sigma}_1^{(r-1)}$ is C^0–small and such that $\hat{\sigma}_1 = 0$ on $\mathrm{Op}(\partial I^m)$. This concludes the base case of the induction.

Suppose that for some $j < N$ we have constructed a C^0–small isotopy $F_t^j \colon I^m \to I^m$ such that $F_t^j = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and a holonomic section $\hat{\sigma}_j \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\hat{\sigma}_j$ is C^0–close to $\sum_{i \leq j} \sigma_i$ on $\mathrm{Op}(F_1^j(I^k))$, such that $\hat{\sigma}_j^{(r-1)}$ is C^0–small and such that $\hat{\sigma}_j = 0$ on $\mathrm{Op}(\partial I^m)$. Apply Theorem 2.4 to the section $(F_1^j)^* \sigma_{j+1} \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$, which satisfies $(F_1^j)^* \sigma_{j+1} = 0$ on $\mathrm{Op}(\partial I^m)$ and $(F_1^j)^* \sigma_{j+1}^\perp = 0$ on all of I^m with respect to the hyperplane field $(F_1^j)^* \tau_{j+1}$. We obtain a C^0–small isotopy $\widetilde{F}_t \colon I^m \to I^m$ such that $\widetilde{F}_t = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and a holonomic section $\widetilde{\sigma} \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\widetilde{\sigma}$ is C^0–close to $(F_1^j)^* \sigma_{j+1}$ on $\mathrm{Op}(\widetilde{F}_1(K))$, such that $\widetilde{\sigma}^{(r-1)}$ is C^0–small and such that $\widetilde{\sigma} = 0$ on $\mathrm{Op}(\partial I^m)$. Set $F_t^{j+1} = F_t^j \circ \widetilde{F}_t$ and $\hat{\sigma}_{j+1} = \hat{\sigma}_j + (F_1^j)_* \widetilde{\sigma}$. This completes the inductive step.

At the last step we obtain a C^0–small isotopy $F_t = F_t^N \colon I^m \to I^m$ such that $F_t = \mathrm{id}_{I^m}$ on $\mathrm{Op}(\partial I^m)$ and a holonomic section $\hat{\sigma} = \hat{\sigma}_N$ such that $\hat{\sigma}$ is C^0–close to $\sigma = \sum_{i=1}^N \sigma_i$ on $\mathrm{Op}(F_1(I^k))$, such that $\hat{\sigma}^{(r-1)}$ is C^0–small and such that $\hat{\sigma} = 0$ on $\mathrm{Op}(\partial I^m)$.

This is exactly what we wanted. We have thus successfully reduced Theorem 2.3 to Theorem 2.4.

It remains to discuss the reduction of the parametric localized holonomic approximation lemma for l–holonomic sections (Theorem 2.5) to the parametric localized holonomic approximation lemma for $\perp$–holonomic sections (Theorem 2.6). However, the proof only differs in notation, namely one just needs to add a parameter everywhere. The key point here is that given a family $\sigma_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ of sections such that $\sigma_z^{(r-1)} = 0$, the decomposition given by Lemma 3.4 depends smoothly on the parameter z.

4 Transversality adjustment

4.1 The transversality condition

By the reductions carried out in Sections 2 and 3, we are left with the task of proving Theorems 2.4 and 2.6, the local relative holonomic approximation lemmas for $\perp$–holonomic sections. The strategy of proof, as in classical holonomic approximation, is to take advantage of the room provided by the positive codimension of I^k in I^m, where $k < m$. This room is used to interpolate between the Taylor polynomials determined by the nonholonomic section that we wish to approximate. More precisely, the idea is to wiggle the subset $I^k \subset I^m$ back and forth in the ambient space I^m and interpolate between Taylor polynomials along the wiggles. However, in order to obtain the fine estimates needed for the desired control on the $\perp$–jet component, our wiggles must be parallel to the hyperplane field τ under consideration. We therefore run into difficulties when τ is tangent to the subset I^k which we want to wiggle. In this section we will perform yet another reduction, so that we only need to consider hyperplane fields τ which are transverse to I^k.

The idea is to further localize the problem by subdividing the cube I^m into very small subcubes, on each of which the hyperplane field τ is almost constant. We show in Section 4.2 below that on the subcubes where the hyperplane field τ is almost tangent to I^k, the desired holonomic approximation can be explicitly constructed by hand. Moreover, in this case no wiggling is necessary. The accuracy of the approximation will depend on the extent to which τ is almost tangent to I^k, but given a fixed degree of accuracy desired we can always restrict our attention to those subcubes on which the angle between τ and I^k is sufficiently small. We explain precisely how to achieve this transversality adjustment in Section 4.3. On the remaining cubes, the hyperplane
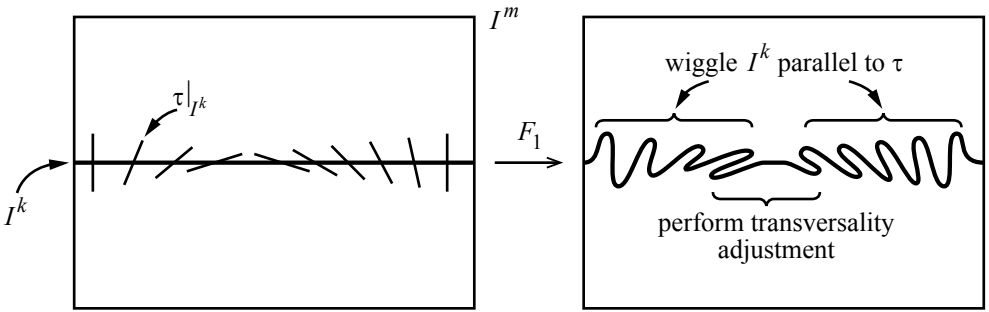
Figure 9: The two steps, a transversality adjustment and a wiggling parallel to $\tau \subset TI^m$

field τ is transverse to I^k and therefore we can perform the wiggling parallel to τ described in the previous paragraph. This last step is carried out in Section 5. We illustrate our strategy in Figure 9.

4.2 Almost tangent hyperplane fields

Let $V, W \subset \mathbb{R}^m$ be two linear subspaces of the same dimension. Recall that the angle $\measuredangle(V, W)$ between V and W is defined as $\measuredangle(V, W) = \sup_{w \in W \setminus 0}(\inf_{v \in V \setminus 0} \measuredangle(v, w))$. When $\dim V < \dim W$, we set $\measuredangle(V, W) = \inf_U \measuredangle(V, U)$, where the infimum is taken over all linear subspaces $U \subset W$ such that $\dim U = \dim V$. Equivalently, we have $\measuredangle(V, W) = \inf_U \measuredangle(U, W)$, where the infimum is taken over all subspaces $U \supset V$ such that $\dim U = \dim W$. For any two distributions $\tau, \eta \subset TI^m$ we define $\measuredangle(\tau, \eta) = \sup_{x \in I^m} \measuredangle(\tau_x, \eta_x)$. The main goal of this section is to establish the following local calculation, where we think of the hyperplane $H = \mathbb{R}^{m-1} \times 0 \subset \mathbb{R}^m$ as a constant hyperplane field on I^m.

Lemma 4.1 *Fix $k < m$. Let $\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a section and let $\tau \subset TI^m$ be a hyperplane field such that*

- $\sigma^\perp = 0$ *with respect to* τ,
- $\sigma = 0$ *on* $\mathrm{Op}(\partial I^m)$.

Then for every $\delta > 0$ there exists a holonomic section $\hat{\sigma} \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that the following properties hold:

- $\mathrm{dist}_{C^0}(\hat{\sigma}, \sigma) \leq C \|\sigma\|_{C^r}(\measuredangle(\tau, H) + \delta)$ *on* $\mathrm{Op}(I^k)$.
- $\|\hat{\sigma}^\perp\|_{C^0} \leq C \|\sigma\|_{C^r}(\measuredangle(\tau, H) + \delta)$ *on all of* I^m.
- $\hat{\sigma} = 0$ *on* $\mathrm{Op}(\partial I^m)$.

Remark 4.2 (1) The constant $C > 0$ only depends on m and r. We can extract an explicit upper bound for C from the proof if we so desire, but this is not important.

(2) To be more precise, for a function germ $h \colon \mathrm{Op}(x) \subset \mathbb{R}^m \to \mathbb{R}$ we define $\|j^r(h)(x)\| = \sup \|\partial_\alpha h(x)\|$, where the supremum is taken over all multi-indices α of order $|\alpha| \le r$. Similarly, we define $\|j^\perp(h)(x)\| = \sup \|\partial_\nu(\partial_\alpha h)(x)\|$, where the supremum is taken over all multi-indices α of order $|\alpha| < r$ and over all unit vectors $\nu \in \tau_x$. For a section $s \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ we set $\|s\|_{C^0} = \sup_{x \in I^m} \|s(x)\|$ and $\|s^\perp\|_{C^0} = \sup_{x \in I^m} \|s^\perp(x)\|$. These are the C^0 norms which appear in the statement of the lemma. There are of course many other equivalent definitions, but they all differ by a constant which can be absorbed into C.

(3) We can define the C^r norm in a similar way. Think of σ as a family of germs $y \mapsto h_x(y)$, $y \in \mathrm{Op}(x)$, parametrized by $x \in I^m$. We can differentiate the vector $h_x(y) \in \mathbb{R}^n$ with respect to x or with respect to y. Set

$$\|\sigma\|_{C^r} = \sup \left\| \frac{\partial}{\partial x_\beta} \frac{\partial}{\partial y_\alpha} h|_{(x,x)} \right\|,$$

where the supremum is taken over all multi-indices α, β of orders $|\alpha|, |\beta| \le r$ and all points $x \in I^m$.

Proof Throughout the proof $C > 0$ will denote a constant, depending only on m and r, but which might be replaced with a bigger such constant whenever necessary. Assume without loss of generality that the angle between τ and H is small, say $\angle(\tau, H) < \frac{\pi}{3}$. Let $u_x \in \mathbb{R}^m$, $x \in I^m$, be the unique field of unit vectors such that for every $x \in I^m$ we have $u_x \perp \tau_x$ and $\angle(\tau_x, H_x) = \angle(u_x, \partial_m)$, where $\partial_m = (0, \ldots, 0, 1) \in \mathbb{R}^m$ (the first condition determines u_x up to sign and the second condition determines the sign). From Section 3.1, for every $x \in I^m$, we know that $\sigma(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ is the r–jet at x of a germ

$$y \mapsto (l_x(y - x))^r \cdot (v_1(x), \ldots, v_n(x)) \in \mathbb{R}^n, \quad y \in \mathrm{Op}(x) \subset \mathbb{R}^m,$$

where $l_x \colon \mathbb{R}^m \to \mathbb{R}$ is the linear function $l_x(\cdot) = \langle \cdot, u_x \rangle$ and $v = (v_1, \ldots, v_n) \colon I^m \to \mathbb{R}^n$ is a function such that $v = 0$ on $\mathrm{Op}(\partial I^m)$. Fix once and for all a cutoff function $\psi \colon \mathbb{R} \to \mathbb{R}$ such that $\psi(t) = 1$ when $|t| < \frac{1}{2}$ and $\psi(t) = 0$ when $|t| > 1$. For $\delta > 0$ small, set $\psi_\delta(t) = \psi(t/\delta)$. Define a holonomic section $\hat\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ by $\hat\sigma = j^r(h)$, where $h \colon I^m \to \mathbb{R}^n$ is the function

$$h(x) = \psi_\delta(x_m) \cdot x_m^r \cdot (v_1(x), \ldots, v_n(x)), \quad x = (x_1, \ldots, x_m) \in I^m.$$

To verify the desired properties for $\hat{\sigma}$, we introduce an auxiliary section $s\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ whose r–jet $s(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ at the point $x = (x_1, \ldots, x_m) \in I^m$ corresponds to the germ

$$y \mapsto (y_m - x_m)^r \cdot (v_1(x), \ldots, v_n(x)) \in \mathbb{R}^n, \quad y = (y_1, \ldots, y_m) \in \mathrm{Op}(x) \subset \mathbb{R}^m.$$

Indeed, $\angle(u, \partial_m) = \angle(\tau, H)$, and hence $\mathrm{dist}_{C^0}(s, \sigma) \le C \|\sigma\|_{C^r} \angle(\tau, H)$ on all of I^m. On the other hand, $\mathrm{dist}(\hat{\sigma}, s) \le C \|\sigma\|_{C^r} \delta$ on $\{|x_m| < \frac{1}{2}\delta\} \subset \mathbb{R}^m$ and we are free to choose δ as small as desired. This proves the first property stated in Lemma 4.1. The third property holds by inspection. It remains to prove the second property.

We compute explicitly the partial derivatives $\partial_\alpha h$ for a multi-index α of order $|\alpha| \le r$ at a point $x \in I^m$. Write $\alpha = (\beta, \gamma)$, where β consists of N indices $1 \le \beta_j < m$ and γ consists of M indices $\gamma_j = m$. Then we have the formula

$$\partial_\alpha h(x) = \sum_{j=0}^{M} \binom{M}{j} \left[\left(\sum_{i=0}^{j} \binom{j}{i} \psi_\delta^{(j-i)}(x_m) \cdot \frac{r!}{(r-i)!} x_m^{r-i} \right) \cdot (\partial_m^{M-j} \partial_\beta v)(x) \right]$$

with

$$|\psi_\delta^{(j-i)}(x_m)| = \left| \frac{1}{\delta^{j-i}} \psi^{(j-i)} \left(\frac{x_m}{\delta} \right) \right| \le \frac{1}{\delta^{j-i}} \|\psi\|_{C^r}.$$

For $M < r$ we can therefore bound $\|\partial_\alpha h\|_{C^0} \le C \|\sigma\|_{C^r} \delta$ on all of I^m. In particular this bound holds for all multi-indices α of order $|\alpha| < r$. For the multi-index $\alpha = (m, \ldots, m)$ of order $|\alpha| = r$ corresponding to the pure r^{th} derivative ∂_m^r we have $\|\partial_\alpha h\|_{C^0} \le C \|\sigma\|_{C^r}$. Hence the inequality $\|\hat{\sigma}\|_{C^0} \le C \|\sigma\|_{C^r}$ holds.

Fix an index α of order $|\alpha| < r$ and let $v \in \tau_x$ be a unit vector. Write $v = \sum_j a_j \partial_j$ in terms of the standard basis $\partial_1, \ldots, \partial_m$ of $\mathbb{R}^m$. Observe that

$$|a_m| = |\langle v, \partial_m \rangle| = |\langle v, \partial_m - u_x \rangle| \le \|\partial_m - u_x\| \le \angle(\tau, H).$$

It follows that

$$\|\partial_v(\partial_\alpha h)(x)\| \le \|a_m \partial_m(\partial_\alpha h)(x)\| + \left\| \sum_{j=1}^{m-1} a_j \partial_j(\partial_\alpha h)(x) \right\|$$

$$\le |a_m| \|\hat{\sigma}\|_{C^0} + \sqrt{\sum_{j=1}^{m-1} (\partial_j(\partial_\alpha h)(x))^2}$$

and therefore that $|\partial_v(\partial_\alpha h)(x)| \le C \|\sigma\|_{C^r}(\angle(\tau, H) + \delta)$. But the point $x \in I^m$, the multi-index α of order $|\alpha| < r$ and the unit vector $v \in \tau_x$ were all chosen arbitrarily, and therefore we have proved the remaining inequality $\|\hat{\sigma}^\perp\|_{C^0} \le C \|\sigma\|_{C^r}(\angle(\tau, H) + \delta)$. See F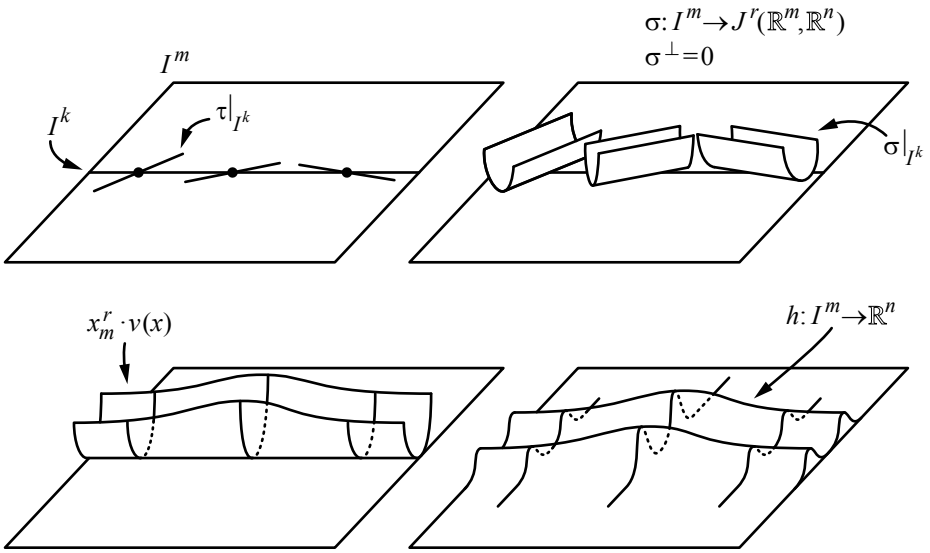igure 10 for an illustration of the argument.   □

Figure 10: The error in the transversality adjustment is proportional to the angle between τ and H.

By adding a subscript everywhere in the above proof we deduce the following parametric version of Lemma 4.1. Observations analogous to the ones made in Remark 4.2 apply.

Lemma 4.3 *Fix $k < m$. Let $\sigma_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a family of sections and let $\tau_z \subset TI^m$ be a family of hyperplane fields parametrized by the unit cube I^q such that*

- $\sigma_z^\perp = 0$ *with respect to τ_z,*
- $\sigma_z = 0$ *on $\mathrm{Op}(\partial I^k)$,*
- $\sigma_z = 0$ *on all of I^m for $z \in \mathrm{Op}(\partial I^q)$.*

Then for every $\delta > 0$ there exists a family of holonomic sections $\hat{\sigma}_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that the following properties hold:

- $\mathrm{dist}_{C^0}(\hat{\sigma}_z, \sigma_z) \leq C \|\sigma_z\|_{C^r}(\measuredangle(\tau_z, H) + \delta)$ *on $\mathrm{Op}(I^k)$.*
- $\|\hat{\sigma}_z^\perp\|_{C^0} \leq C \|\sigma_z\|_{C^r}(\measuredangle(\tau_z, H) + \delta)$ *on all of I^m.*
- $\hat{\sigma}_z = 0$ *on $\mathrm{Op}(\partial I^m)$.*
- $\hat{\sigma}_z = 0$ *on all of I^m for $z \in \mathrm{Op}(\partial I^q)$.*

Of course, there is nothing special about the hyperplane $\mathbb{R}^{m-1} \times 0 \subset \mathbb{R}^m$, which was only fixed for concreteness. In fact, Lemmas 4.1 and 4.3 hold in the following more

general form. Observe first that we may replace the constant hyperplane field H by any hyperplane field $\widetilde{H} \subset TI^m$ such that $\mathbb{R}^k \times 0^{m-k} \subset \widetilde{H}_x$ at all points $x \in I^m$. To see this it suffices to consider a local change of coordinates near the subset $I^k \subset I^m$ which fixes I^k pointwise and which sends H_x to $\widetilde{H}_x$ for all $x \in \mathrm{Op}(I^k) \setminus \mathrm{Op}(\partial I^m)$. It follows from this observation that in the statement of Lemma 4.1 we may replace the angle $\angle(\tau, H)$ with the angle $\angle(\tau, I^k)$ formed by the distributions $\tau|_{I^k}$ and TI^k along I^k. Indeed, $\angle(\tau, I^k) = \inf \angle(\tau|_{I^k}, \widetilde{H}|_{I^k})$, where the infimum is taken over all hyperplane fields $\widetilde{H} \subset TI^m$ such that $\mathbb{R}^k \times 0^{m-k} \subset \widetilde{H}_x$ at all points $x \in I^m$. In the parametric case, we also allow the hyperplane field $\widetilde{H}_z \subset TI^m$ to vary with the parameter $z \in I^q$. Therefore in the statement of Lemma 4.3 we may replace the angle $\angle(\tau_z, H)$ with the angle $\angle(\tau_z, I^k)$.

4.3 Reduction to the transverse case

We are ready to reduce Theorem 2.4 to the case where the hyperplane field is transverse to the subset $I^k \subset I^m$. Fix an arbitrary hyperplane field $\tau \subset TI^m$. Let $\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a section such that $\sigma^\perp = 0$ with respect to τ and such that $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$. Fix $\varepsilon > 0$ small, the desired accuracy for the C^0–approximation we must construct. Consider the cubical stratification of the subset $I^k \subset I^m$ in which the j–dimensional stratum consists of the union of the j–dimensional faces of the cubes

$$Q_N(j_1, \ldots, j_k) = \left[\frac{j_1}{N}, \frac{j_1 + 1}{N}\right] \times \cdots \times \left[\frac{j_k}{N}, \frac{j_k + 1}{N}\right] \subset I^k, \quad -N \le j_1, \ldots, j_k < N.$$

Let $\lambda > 0$ be small enough so that $C \|\sigma\|_{C^r} \lambda < \varepsilon$, where $C > 0$ is the constant which appears in the statement of Lemma 4.1. Choose $N \in \mathbb{N}$ big enough that for each cube $Q = Q_N(j_1, \ldots, j_k)$ we have $\angle(\tau_x, \tau_y) < \frac{1}{2}\lambda$ for all $x, y \in \mathrm{Op}(Q) \subset I^m$. Consider the polyhedron $R = \bigcup_{j=0}^k R^j \subset I^k$, where the stratum R^j consists of the union of all the j–dimensional faces F of the cubes $Q = Q_N(j_1, \ldots, j_k)$ such that $\angle(\tau_x, F) \le \lambda$ at all points $x \in F$. The hyperplane field τ is almost tangent to the faces F in R and transverse to all other faces F in the cubical stratification of I^k under consideration. Indeed, for all faces F not in R we have $\angle(\tau_x, F) \ge \frac{1}{2}\lambda$ at all points $x \in F$.

We proceed inductively to construct a holonomic approximation of σ over the cubical skeleton of R, working one face at a time as in Section 2. At each stage we apply Lemma 4.1. Since $\angle(\tau, F) \le \lambda$ for each face F under consideration, the resulting holonomic approximation has C^0–error $< \varepsilon$. On the remaining faces F which are not in R, we have $\tau \pitchfork F$. The following result will be proved in Section 5, where we think of the hyperplane $V = 0 \times \mathbb{R}^{m-1} \subset \mathbb{R}^m$ as a constant hyperplane field on $\mathbb{R}^m$.

Theorem 4.4 *Let* $\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ *be a section such that*

- $\sigma = 0$ *on* $\mathrm{Op}(\partial I^m)$,
- $\sigma^\perp = 0$ *on all of* I^m *with respect to* V.

Then there exists an isotopy $F_t\colon I^m \to I^m$ *and a holonomic section* $\hat\sigma\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ *such that the following properties hold:*

- $\hat\sigma$ *is* C^0*–close to* σ *on* $\mathrm{Op}(F_1(I^k))$.
- $\hat\sigma^\perp$ *is* C^0*–small on all of* I^m.
- F_t *is* C^0*–small.*
- $F_t = \mathrm{id}_{I^m}$ *and* $\hat\sigma = 0$ *on* $\mathrm{Op}(\partial I^m)$.
- V *is invariant under* F_t.

Remark 4.5 The more accurate the C^0–approximation desired, the bigger the derivative dF_t will need to be. However, for a fixed C^0–accuracy we can arrange it so that $\mathrm{dist}_{C^0}(F_t, \mathrm{id}_{I^m})$ is arbitrarily small and so that $F_t = \mathrm{id}_{I^m}$ outside of an arbitrarily small neighborhood of I^k in I^m while keeping $\|dF_t\|_{C^0}$ uniformly bounded. This scale invariance follows from the explicit construction of $\hat\sigma$ and F_t which is carried out in Section 5.

Assuming Theorem 4.4, we continue the inductive process over the rest of the skeleton of I^m to obtain a global holonomic ε–approximation of σ. Indeed, given a face F not in R, we can approximate the hyperplane field $\tau|_F$ by a constant hyperplane field along F; see Figure 11. We pay a price, of course, but the error can be made arbitrarily small by taking N sufficiently big. We can therefore reduce the problem at each face F to the local model considered in Theorem 4.4. Observe that the last property stated in Theorem 4.4 and the a priori bound on $\|dF_t\|_{C^0}$ provided by Remark 4.5 are needed to show that after each step of the inductive process the approximation of τ by a piecewise-constant hyperplane field has not been ruined by the corresponding isotopy. To be more precise, the distortion produced by each isotopy can be made arbitrarily small by taking N sufficiently big.

We have successfully reduced Theorem 2.4 to the transverse local model Theorem 4.4 above. The same argument also works in families, using Lemma 4.3 instead of Lemma 4.1. Thus we can also reduce the parametric Theorem 2.6 to a parametric transverse local model. The only difference in the reduction is that we must also subdivide the parameter space I^q, as well as the domain I^m, into small enough subcubes. The parametric version of Theorem 4.4 reads as follows.
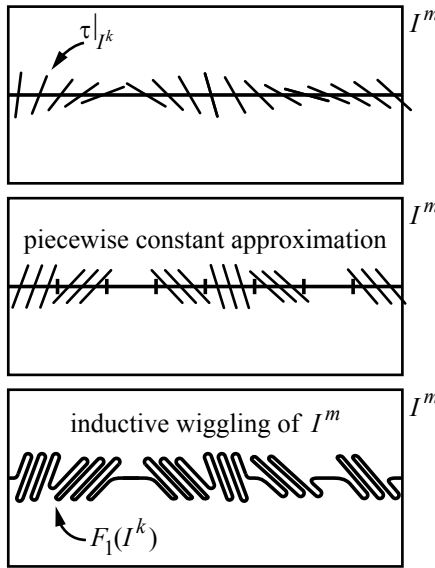
Figure 11: Approximation of τ by a piecewise-constant hyperplane field and the corresponding wiggling

Theorem 4.6 *Let* $\sigma_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ *be a family of sections parametrized by the unit cube* I^q *such that*

- $\sigma_z = 0$ *on* $\mathrm{Op}(\partial I^m)$,
- $\sigma_z^\perp = 0$ *on all of* I^m *with respect to* V,
- $\sigma_z = 0$ *on all of* I^m *for* $z \in \mathrm{Op}(\partial I^q)$.

Then there exists a family of isotopies $F_t^z \colon I^m \to I^m$ *and a family of holonomic sections* $\hat\sigma_z \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ *such that the following properties hold:*

- $\hat\sigma_z$ *is* C^0*–close to* σ_z *on* $\mathrm{Op}(F_1^z(I^k))$.
- $\hat\sigma_z^\perp$ *is* C^0*–small on all of* I^m.
- F_t^z *is* C^0*–small.*
- $F_t^z = \mathrm{id}_{I^m}$ *and* $\hat\sigma_z = 0$ *on* $\mathrm{Op}(\partial I^m)$.
- $F_t^z = \mathrm{id}_{I^m}$ *and* $\hat\sigma_z = 0$ *on all of* I^m *for* $z \in \mathrm{Op}(\partial I^q)$.
- V *is invariant under* F_t^z.

We note that there also is an priori bound on $\|dF_t^z\|_{C^0}$ depending on the desired accuracy of the C^0–approximation, just as in Remark 4.5. We have now completed all preparatory reductions.

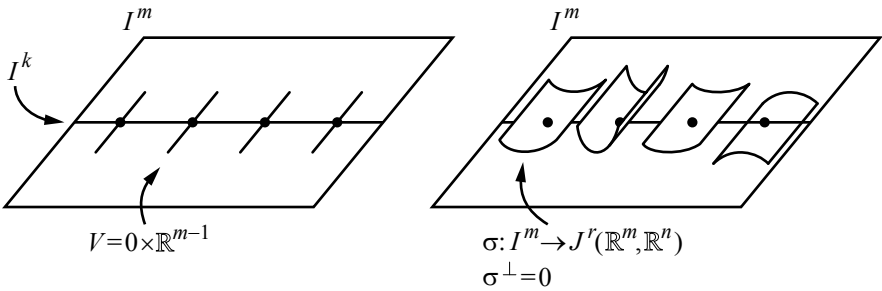Figure 12: A typical section σ which is primitive with respect to the hyperplane field V

5 Holonomic approximation with controlled cutoff

5.1 The transverse local model

We begin by establishing some simple estimates which will be crucial in the quantitative holonomic approximation process described below. We exploit the concreteness of the local models considered in Theorems 4.4 and 4.6 by writing down the main objects explicitly, differentiating them by hand and thereby deducing the necessary bounds. We once again use $C > 0$ to denote a constant, which only depends on m and r, but which will be replaced with a bigger such constant whenever necessary.

Consider a section $\sigma \colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $\sigma^\perp = 0$ with respect to the constant hyperplane field $V = 0 \times \mathbb{R}^{m-1}$. In the spirit of Section 3.1, we can give an explicit description of σ. Each r–jet $\sigma(x) \in J^r(\mathbb{R}^m, \mathbb{R}^n)$ at a point $x = (x_1, \ldots, x_m) \in I^m$ corresponds to a germ

$$y \mapsto h_x(y) = (y_1 - x_1)^r \cdot (v_1(x), \ldots, v_n(x)) \in \mathbb{R}^n, \quad y = (y_1, \ldots, y_m) \in \mathrm{Op}(x) \subset I^m.$$

If $\sigma = 0$ on $\mathrm{Op}(\partial I^m)$, then the function $v = (v_1, \ldots, v_n) \colon I^m \to \mathbb{R}^n$ also satisfies $v = 0$ on $\mathrm{Op}(\partial I^m)$. We must control the derivatives of the function $h(x, y) = h_x(y)$ with respect to both x and y. For this purpose, let α and β be multi-indices such that $|\alpha| + |\beta| \le r$. Write $\alpha = (\xi, \gamma)$, where ξ consists of M indices $\xi_j = 1$ and γ consists only of indices $1 < \gamma_j \le m$. Similarly, write $\beta = (\zeta, \mu)$, where ζ consists of N indices $\zeta_j = 1$ and μ consists only of indices $1 < \mu_j \le m$. We compute the following formula for the derivatives of $h(x, y)$:

$$\frac{\partial}{\partial x_\alpha} \frac{\partial}{\partial y_\beta} h|_{(x,y)}$$

$$= \sum_{j=1}^{M} \sum_{i=1}^{N} \binom{M}{j} \binom{N}{i} (-1)^j \frac{r!}{(r-j-i)!} (y_1 - x_1)^{r-j-i} \cdot (\partial_1^{M-j-i} \partial_\mu \partial_\gamma v)(x).$$

We therefore obtain the estimate

$$\left\| \frac{\partial}{\partial x_\alpha} \frac{\partial}{\partial y_\beta} h \right\| \leq C \|\sigma\|_{C^r} \delta^{r-(M+N)} \quad \text{for } y \in \{|y_1 - x_1| < \delta\} \subset I^m.$$

We next give a local model for the wiggling. Many other choices work just as well, of course, but we want to write down an explicit model for concreteness. Given $\varepsilon, \delta > 0$ small such that $\delta \ll \varepsilon$ to an extent which will be made precise later, consider the sinusoidal curve

$$w(u) = \frac{\varepsilon}{2} \sin\left(\frac{\pi u}{2\delta}\right), \quad u \in \mathbb{R}.$$

Fix a cutoff function $\psi \colon \mathbb{R} \to \mathbb{R}$ such that $\psi(u) = 0$ for $|u| < \frac{1}{2}$ and $\psi(u) = 1$ for $|u| > \frac{3}{4}$. For $\varepsilon > 0$ small enough that $\operatorname{supp}(\sigma) \subset [-1+\varepsilon, 1-\varepsilon]^m$, define an isotopy $F_t \colon I^m \to I^m$ by the formula

$$F_t(x_1, \ldots, x_m) = (x_1, \ldots, x_m + \varphi_t(x)),$$

$$\varphi_t(x) = t \, \psi\left(\frac{1-|x_1|}{\varepsilon}\right) \cdots \psi\left(\frac{1-|x_m|}{\varepsilon}\right) w(x_1).$$

Observe that $\operatorname{dist}_{C^0}(F_t, \operatorname{id}_{I^m}) < \varepsilon$. We also have the following estimate for the derivative dF_t of the isotopy F_t, where we note that the ratio ε/δ will typically be
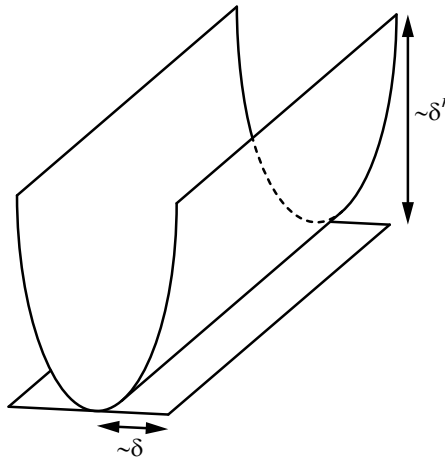

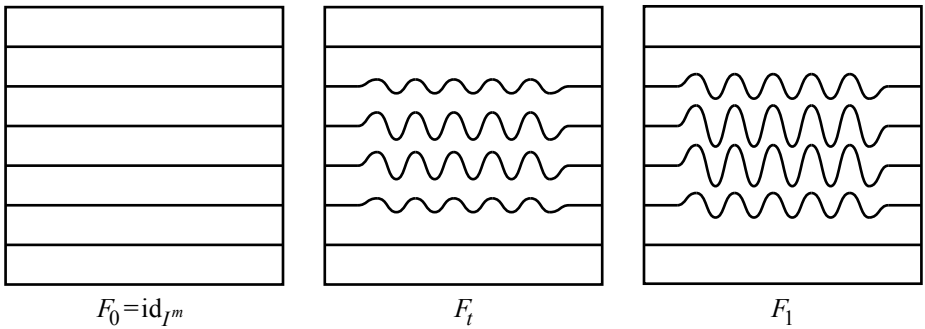Figure 13: The estimate $\|h\| \leq C \|\sigma\|_{C^r} \delta^r$

Figure 14: The isotopy F_t

very big but remains invariant by a simultaneous scaling of ε and δ:

$$\|dF_t\|_{C^0} \leq C \frac{\varepsilon}{\delta}.$$

In Section 5.2 we apply the same method of proof as Eliashberg and Mishachev in [4] to produce a holonomic approximation $j^r(g)$ of the section σ. The domain of definition of the function g is the wiggled neighborhood of $I^k \subset I^m$ given by $U = F_t(\{|x_m| < \frac{1}{4}\varepsilon\}) \subset I^m$. To extend our holonomic approximation to the whole of I^m, we multiply the function $g: U \to \mathbb{R}^n$ by a cutoff function supported in U. We must control the derivatives of such a cutoff function, so we now write down an explicit model together with the appropriate estimate.

In terms of the function ψ fixed above, define $\phi: I^m \to \mathbb{R}$ by $\phi(x) = 1 - \psi(4|y_m|/\varepsilon)$, where $F_1(y) = x$. Note that $\phi = 1$ near $F_1(I^k)$ and that $\operatorname{supp}(\phi) \subset U$; see Figure 16. The following bound holds for the derivatives of ϕ. Let α be a multi-index of order $|\alpha| \leq r$. Write $\alpha = (\beta, \gamma)$, where β consists of N indices $1 < \beta_j \leq m$ and γ consists of M indices $\gamma_j = 1$. Then we have

$$|\partial_\alpha \phi| \leq C \frac{1}{\varepsilon^N \delta^M}.$$

5.2 The holonomic approximation process

We are ready to construct the holonomic approximation $\hat{\sigma}$ of σ. We use the isotopy $F_t: I^m \to I^m$ and the cutoff function $\phi: I^m \to \mathbb{R}$ defined in Section 5.1, which depend on two parameters ε and δ. We will obtain an arbitrarily good C^0–approximation $\hat{\sigma}$ by choosing $\varepsilon, \delta > 0$ arbitrarily small such that the ratio δ/ε is also arbitrarily small. Fix a function $\eta: \mathbb{R} \to \mathbb{R}$ such that

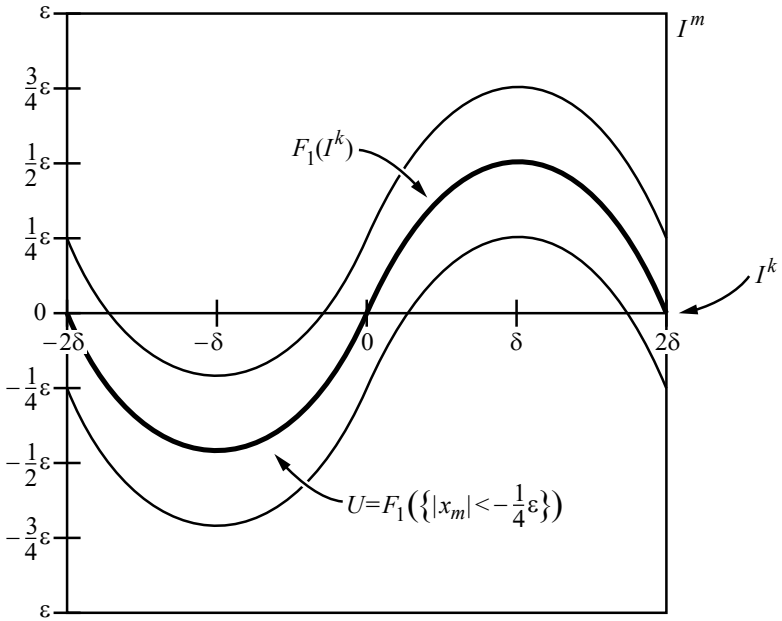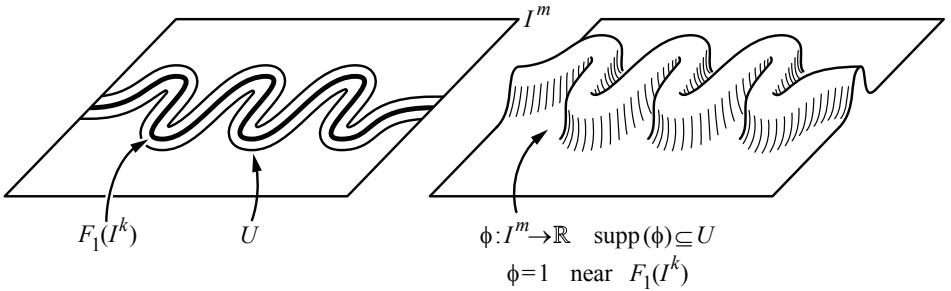Figure 15: The sinusoidal subset $U \subset I^m$ in a neighborhood of the origin

Figure 16: The cutoff function ϕ

- $\eta(u) = -1$ for $u < -1$,
- $1 \le \eta(u) \le 1$ for $-1 \le u \le 1$,
- $\eta(u) = 1$ for $u > 1$.

We construct $\hat\sigma$ by writing down an explicit formula on each of the rectangles

$$R_j = \big[(2j-1)\delta, (2j+1)\delta\big] \times I^{m-2} \times [-\varepsilon, \varepsilon] \subset I^m$$

such that R_j is contained in the support of σ. Suppose first that $j \in \mathbb{N}$ is even. Define a function

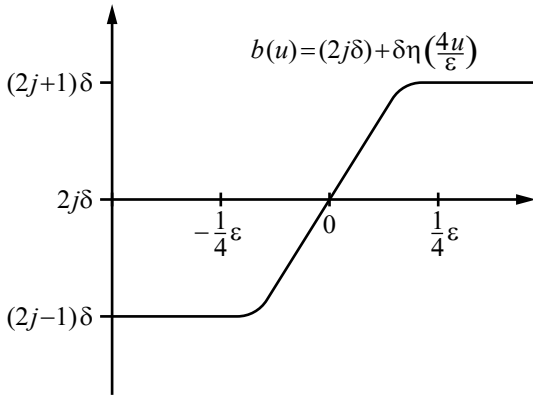$$g \colon R_j \to \mathbb{R}^n, \quad g(x) = h(p(x), x),$$

Figure 17: The basepoint-interpolating function b

where
$$p(x) = \left((2j\delta) + \delta\eta\left(\frac{4x_m}{\varepsilon}\right), x_2, \ldots, x_{m-1}, 0\right), \quad x = (x_1, \ldots, x_m) \in R_j.$$

Let $b(u) = (2j\delta) + \delta\eta(4u/\varepsilon)$, so that $p(x) = (b(x_m), x_2, \ldots, x_{m-1}, 0)$. We note for future reference the following bound on the derivatives of the function b:

$$|b^{(i)}| \le C\frac{\delta}{\varepsilon^i}.$$

Remark 5.1 Observe that:

(1) On $R_j \cap \{x_m < -\frac{1}{4}\varepsilon\}$ we have $p(x) = ((2j-1)\delta, x_2, \ldots, x_{m-1}, 0)$.

(2) On $R_j \cap \{x_m > \frac{1}{4}\varepsilon\}$ we have $p(x) = ((2j+1)\delta, x_2, \ldots, x_{m-1}, 0)$.

(3) On $R_j \cap \{|x_m| < \frac{1}{4}\varepsilon\}$ the point $p(x)$ interpolates between

$$((2j-1)\delta, x_2, \ldots, x_{m-1}, 0) \quad \text{and} \quad ((2j+1)\delta, x_2, \ldots, x_{m-1}, 0).$$

Similarly, if $j \in \mathbb{N}$ is odd, we define a function

$$g: R_j \to \mathbb{R}^n, \quad g(x) = h(p(x), x),$$

where

$$p(x) = \left((2j\delta) - \delta\eta\left(\frac{4x_m}{\varepsilon}\right), x_2, \ldots, x_{m-1}, 0\right), \quad x = (x_1, \ldots, x_m) \in R_j.$$

The difference in sign corresponds to the fact that on the interval $\left[(2j-1)\frac{\pi}{2}, (2j+1)\frac{\pi}{2}\right]$ the function $u \mapsto \sin(u)$ is increasing for j even and decreasing for j odd. Note that the locally defined functions $g: R_j \to \mathbb{R}^n$ do not glue together on $\{|x_m| < \varepsilon\} = \bigcup_j R_j \subset I^m$.
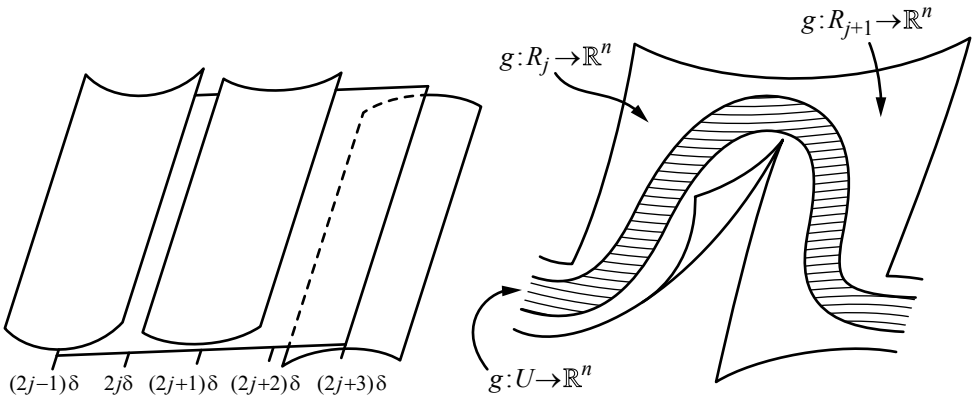
Figure 18: The glued-up function $g\colon U \to \mathbb{R}^n$

However, they do glue together on the wiggled neighborhood $U = F_1\big(\{|x_m| < \tfrac{1}{4}\varepsilon\}\big)$ of $F_1(I^k)$; see Figure 18. We therefore obtain a globally defined function $g\colon U \to \mathbb{R}^n$.

Set $f(x) = \phi(x) \cdot g(x)$. Since $\mathrm{supp}(\phi) \subset U$, the function $f\colon I^m \to \mathbb{R}$ is defined on all of I^m. The holonomic section $\hat{\sigma} = j^r(f)$ is the desired holonomic approximation to σ. It remains to prove that for an adequate choice of the parameters ε and δ all of the properties listed in Theorem 4.4 are satisfied. The last three properties can be verified by inspection. In the next section we carry out the calculation required to establish the other two.

5.3 Conclusion of the proof

Let α be a multi-index of order $|\alpha| \leq r$. Write $\alpha = (\beta, \gamma, \xi)$, where β consists of I indices $\beta_j = 1$, γ consists of J indices $1 < \gamma_j < m$ and ξ consists of K indices $\xi_j = m$. Then we compute

$$\frac{\partial}{\partial x_\alpha} g|_x = \sum \left(\frac{\partial^i}{\partial x_1^i} \frac{\partial}{\partial x_\gamma} \frac{\partial^I}{\partial y_1^I} \frac{\partial^j}{\partial y_m^j} h|_{(p(x),x)} \right) b^{(k_1)}(x_m) \cdots b^{(k_i)}(x_m),$$

where the sum is taken over all nonnegative integers such that $j + k_1 + \cdots + k_i = K$. From the estimates established in Section 5.1 we deduce that

$$\left\| \frac{\partial}{\partial x_\alpha} g \right\| \leq C \|\sigma\|_{C^r} \sum \delta^{r-(i+I)} \left(\frac{\delta}{\varepsilon^{k_1}} \right) \cdots \left(\frac{\delta}{\varepsilon^{k_i}} \right)$$

$$= C \|\sigma\|_{C^r} \sum \frac{\delta^{r-I}}{\varepsilon^{k_1 + \cdots + k_i}}$$

$$\leq C \|\sigma\|_{C^r} \frac{\delta^{r-I}}{\varepsilon^K}.$$

Suppose first that α is a multi-index such that $I < r$. Observe that $\delta^{r-I} \leq \delta^{K+J} \leq \delta^K$. It follows that we have

$$\left\| \frac{\partial}{\partial x_\alpha} g \right\| \leq C \|\sigma\|_{C^r} \frac{\delta}{\varepsilon},$$

and therefore we can make this derivative arbitrarily small by ensuring that the ratio δ/ε is arbitrarily small. When $I = r$, the multi-index $\alpha = (1, \ldots, 1)$ corresponds to the pure r^{th} derivative ∂_1^r. Observe that in this case the sum collapses to

$$\frac{\partial^r}{\partial x_1^r} g |_x = \frac{\partial^r}{\partial y_1^r} h |_{(p(x),x)}.$$

Since $\|x - p(x)\| \leq C\varepsilon$, it follows that

$$\left\| \frac{\partial^r}{\partial y_1^r} h |_{(p(x),x)} - \frac{\partial^r}{\partial y_1^r} h |_{(x,x)} \right\| \leq C \|\sigma\|_{C^r} \varepsilon.$$

Hence we deduce the inequality

$$\operatorname{dist}_{C^0}(\hat{\sigma}, \sigma) \leq C \|\sigma\|_{C^r} \left(\varepsilon + \frac{\delta}{\varepsilon} \right)$$

on the smaller neighborhood of $F_1(I^k)$ in U where $\phi = 1$, so $f = g$ and $\hat{\sigma} = j^r(g)$. This proves that the C^0–approximation can be made arbitrarily accurate by choosing $\varepsilon, \delta > 0$ arbitrarily small such that the ratio δ/ε is also arbitrarily small.

It remains to show that for such a choice of ε and δ we can also ensure that $\hat{\sigma}^\perp$ is small on all of I^m. We again must explicitly compute some derivatives. Let α be a multi-index of order $|\alpha| \leq r$. Then we have

$$\frac{\partial}{\partial x_\alpha} f = \sum_{\beta + \gamma = \alpha} \left(\frac{\partial}{\partial x_\beta} \phi \right) \left(\frac{\partial}{\partial x_\gamma} g \right).$$

Suppose that $\alpha_j = 1$ for exactly N indices, where $N \leq r$. Invoking the estimates established in [Section 5.1](#) we deduce that

$$\left\| \frac{\partial}{\partial x_\alpha} f \right\| \leq C \|\sigma\|_{C^r} \left(\frac{\delta}{\varepsilon} \right)^{r-N}.$$

It follows that with the exception of the case $N = r$ we have

$$\left\| \frac{\partial}{\partial x_\alpha} f \right\| \leq C \|\sigma\|_{C^r} \frac{\delta}{\varepsilon}.$$

Since the $\perp$–jet $\hat{\sigma}^\perp$ consists of all derivatives $\partial_\alpha f$ for multi-indices α of order $|\alpha| \leq r$ such that $N < r$, we obtain the inequality

$$\|\hat{\sigma}^\perp\|_{C^0} \leq C \|\sigma\|_{C^r} \frac{\delta}{\varepsilon}.$$

This concludes the proof of Theorem 4.4.

5.4 The parametric case

The above calculation also works in families, by adding a parameter everywhere. We spell out the details for completeness. Let $\sigma_z\colon I^m \to J^r(\mathbb{R}^m, \mathbb{R}^n)$ be a family of sections parametrized by the unit cube I^q such that $\sigma_z^\perp = 0$ with respect to $V = 0 \times \mathbb{R}^{m-1}$, such that $\sigma_z = 0$ on $\mathrm{Op}(\partial I^m)$ and such that $\sigma_z = 0$ on all of I^m for $z \in \mathrm{Op}(\partial I^q)$. We think of σ_z as a family of germs $(z, x, y) \mapsto h(z, x, y) \in \mathbb{R}^m$, where $x \in I^m$, $y \in \mathrm{Op}(x)$ and $z \in I^q$.

We can define a family of functions g_z as before by setting $g_z(x) = h(z, p(x), z)$ on each rectangle R_j. The domain of definition of g_z is $U_z = F_1^z\big(\{|x_m| < \frac{1}{4}\varepsilon\}\big) \subset I^m$, where for each $z = (z_1, \ldots, z_q) \in I^q$ we have an isotopy $F_t^z\colon I^m \to I^m$ given by

$$F_t^z(x_1, \ldots, x_m) = (x_1, \ldots, x_m + \varphi_t^z(x)),$$

where

$$\varphi_t^z(x) = t\,\psi\Big(\frac{1-|x_1|}{\varepsilon}\Big)\cdots\psi\Big(\frac{1-|x_m|}{\varepsilon}\Big)\psi\Big(\frac{1-|z_1|}{\varepsilon}\Big)\cdots\psi\Big(\frac{1-|z_q|}{\varepsilon}\Big)w(x_1).$$

We use the same functions ψ and w as in the nonparametric case. We also use the corresponding family of cutoff functions $\phi_z\colon I^m \to \mathbb{R}$, satisfying $\phi_z = 1$ near $F_1^z(I^k)$ and $\mathrm{supp}(\phi_z) \subset U_z$, which are given by

$$\phi_z(x) = 1 - \psi\Big(\frac{4|y_m|}{\varepsilon}\Big), \quad F_1^z(x) = y.$$

Set $f_z(x) = \phi_z(x) \cdot g_z(x)$ and $\hat{\sigma}_z = j^r(f_z)$ to obtain the desired holonomic approximation. The computation carried out in Section 5.2 shows that by taking ε and δ arbitrarily small such that the ratio δ/ε is also arbitrarily small, all of the properties stated in Theorem 4.6 are satisfied.

References

[1] **D Álvarez-Gavela**, *The simplification of singularities of Lagrangian and Legendrian fronts*, preprint (2016) arXiv

[2] **I Bogdanov**, *Decomposition of symmetric homogeneous polynomials*, MathOverflow comment (2014) Available at https://mathoverflow.net/questions/180978/decomposition-of-symmetric-homogeneous-polynomials

[3] **Y M Eliashberg**, **N M Mishachev**, *Holonomic approximation and Gromov's h–principle*, from "Essays on geometry and related topics, 1, 2" (E Ghys, P de la Harpe, V F R Jones, V Sergiescu, T Tsuboi, editors), Monogr. Enseign. Math. 38, Enseign. Math., Geneva (2001) 271–285 MR

[4] **Y M Eliashberg**, **N M Mishachev**, *Introduction to the h–principle*, Graduate Studies in Mathematics 48, Amer. Math. Soc., Providence, RI (2002) MR

[5] **Y M Eliashberg**, **N M Mishachev**, *Wrinkled embeddings*, from "Foliations, geometry, and topology" (N C Saldanha, L Conlon, R Langevin, T Tsuboi, P Walczak, editors), Contemp. Math. 498, Amer. Math. Soc., Providence, RI (2009) 207–232 MR

[6] **M L Gromov**, *Stable mappings of foliations into manifolds*, Izv. Akad. Nauk SSSR Ser. Mat. 33 (1969) 707–734 MR In Russian; translated in Math. USSR Izv. 3 (1969) 671–694

[7] **M L Gromov**, *Convex integration of differential relations, I*, Izv. Akad. Nauk SSSR Ser. Mat. 37 (1973) 329–343 MR In Russian; translated in Math. USSR Izv. 7 (1973) 329–343

[8] **M L Gromov**, *Partial differential relations*, Ergeb. Math. Grenzgeb. 9, Springer (1986) MR

[9] **M W Hirsch**, *Immersions of manifolds*, Trans. Amer. Math. Soc. 93 (1959) 242–276 MR

[10] **C Rourke**, **B Sanderson**, *The compression theorem, I*, Geom. Topol. 5 (2001) 399–429 MR

[11] **C Rourke**, **B Sanderson**, *The compression theorem, II: Directed embeddings*, Geom. Topol. 5 (2001) 431–440 MR

[12] **S Smale**, *The classification of immersions of spheres in Euclidean spaces*, Ann. of Math. 69 (1959) 327–344 MR

[13] **D Spring**, *Convex integration theory*, Monographs in Mathematics 92, Birkhäuser, Basel (1998) MR

[14] **D Spring**, *Directed embeddings and the simplification of singularities*, Commun. Contemp. Math. 4 (2002) 107–144 MR

Department of Mathematics, Stanford University
Stanford, CA, United States

dgavela@stanford.edu