IMMERSION THEORY ETC. FOR HOMOTOPY THEORISTS

MICHAEL WEISS

1. Intro: What is immersion theory ?

Definition 1.1. Suppose that M^m and N^n are smooth manifolds. A smooth map $f\colon M \to N$ is an *immersion* if for all $x \in M$, the differential $df(x)\colon T_xM \longrightarrow T_{f(x)}N$ is an *injective* linear map.

Exercise 1.2. Let $f\colon M \to N$ be an immersion, $x \in M$. Show that there exist smooth local coordinates near $x \in M$ and $f(x) \in N$ such that, in these coordinates, f has the form $(z_1, z_2, \dots, z_m) \mapsto (z_1, z_2, \dots, z_m, 0, 0, 0, \dots, 0)$.

Theorem 1.3. (Main theorem of immersion theory; [9], [4], [3].) *Let M^m, N^n be smooth manifolds where $m < n$. Suppose M compact, possibly with boundary, but N can be noncompact, and must be without boundary. Let $\mathbf{imm}(M, N)$ be the space of smooth immersions from M to N (details later). Let $\mathbf{fimm}(M, N)$ be the space of "formal" immersions, i.e. the space of pairs $(f, \delta f)$ where $f\colon M \to N$ is continuous (not necessarily smooth) and δf is* some *vector bundle map $TM \to f^*TN$ which is injective on each fiber. (See remark below.) Then the obvious map*

$$\mathbf{imm}(M, N) \longrightarrow \mathbf{fimm}(M, N)$$

given by $f \mapsto (f, df)$ is a homotopy equivalence.

Remark. The vector bundle map $\delta f\colon TM \to f^*TN$ amounts to a choice of a linear injection $T_xM \to T_{f(x)}M$ for each $x \in M$, depending continuously on x. You can think of δf as a "formal" total derivative for the continuous map f, but it need not agree with the honest derivative df of f and f may not even have an honest derivative !

Remark. There is a slightly stronger version of the main theorem which includes codimension zero situations. We will come to that later.

Exercise 1.4. Show that the obvious map $\mathbf{imm}(M, N) \to \mathbf{fimm}(M, N)$ is always injective, but almost never surjective. (Describe the exceptions.)

Exercise 1.5. Let $j\colon \mathbb{S}^1 \to \mathbb{R}^2$ be the standard inclusion. Let $f\colon \mathbb{S}^1 \to \mathbb{S}^1$ be any orientation–reversing diffeomorphism (for example, $z \mapsto z^{-1}$ in complex number notation). Using the main theorem, show that the immersions j and jf are not regularly homotopic, i.e., not in the same path component of $\mathbf{imm}(\mathbb{S}^1, \mathbb{R}^2)$. More generally, use the main theorem to make a bijection from $\pi_0(\mathbf{imm}(\mathbb{S}^1, \mathbb{R}^2))$ to $\mathbb{Z}$. Describe the geometric meaning of this bijection. Draw representing immersions for each $z \in \mathbb{Z}$.

Date: October–December 2004.

Exercise 1.6. Let $j: \mathbb{S}^2 \to \mathbb{R}^3$ be the standard inclusion. Let $f: \mathbb{S}^2 \to \mathbb{S}^2$ be any orientation–reversing diffeomorphism (for example, reflection at the equator). Using the main theorem, show that the immersions j and jf are regularly homotopic. How big is $\pi_0(\mathbf{imm}(\mathbb{S}^2, \mathbb{R}^3))$?

Exercise 1.7. Let $j: \mathbb{S}^{n-1} \to \mathbb{R}^n$ be the standard inclusion. Let $f: \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ be any orientation–reversing diffeomorphism (for example, reflection at the equator). *Suppose* that the immersions j and jf are regularly homotopic. What can you say about n ?

2. Fibrations and related notions

Immersion theory turns out to be "applied fibration theory". In this section and the next, the main results from fibration theory which we will need are collected.

Definition 2.1. A map $p: E \to B$ has the *homotopy lifting property* (also *covering homotopy property*) if the following holds. Given any space X and (continuous) maps

$$f: X \times [0,1] \to B, \qquad \bar{f}_0: X \to E$$

such that $p\bar{f}_0(x) = f(x,0)$ for all $x \in X$, there exists a (continuous) $\bar{f}: X \times [0,1] \to B$ such that $p\bar{f} = f$ and $\bar{f}\iota_0(x,0) = \bar{f}_0(x)$ for all $x \in X$. If p has the HLP, it is called a fibration.

More vocabulary. It is also common to say *Hurewicz fibration* in the above circumstances. If p satisfies the above whenever X is a *CW*–space, then p is a *Serre fibration.*

Theorem 2.2. *If $p: E \to B$ is a fibre bundle and B is paracompact, then p is a fibration.*

This is well known, but the proof is not easy. Not given here. Most reasonable spaces are paracompact — in particular every metric space is paracompact.

Exercise 2.3. Let $p: E \to B$ be a fibration, where B is path connected. Let $x, y \in B$. Show that the spaces $p^{-1}(x)$ and $p^{-1}(y)$ are homotopy equivalent.

Example 2.4. The map $p: \mathbb{S}^1 \to \mathbb{S}^1$ given by $z \mapsto z^2$ (in complex number notation) is a fiber bundle, therefore a fibration. The evaluation map $\mathrm{O}(n) \to \mathbb{S}^n$ given by $A \mapsto A\mathbf{e}_1$ is a fiber bundle (where $\mathbf{e}_1$ is the first standard basis vector). The projection from the triangle $\big\{(x,y) \in \mathbb{R}^2 \mid x, y \geq 0, x + y \leq 1\big\}$ to the interval $[0,1]$ given by $(x,y) \mapsto x$ is a fibration, but not a fiber bundle.

Definition 2.5. A map $p: E \to B$ is also called a *space over* B. Given two spaces over B, say $p_1: E_1 \to B$ and $p_2: E_2 \to B$, a map over B from p_1 to p_2 is a map $f: E_1 \to E_2$ such that $p_2 f = p_1$. To be honest, the standard expression is: "... a map from E_1 to E_2 over B ...". Given two maps $f, g: E_1 \to E_2$, both over B, and a homotopy $h: E_1 \times [0,1] \to E_2$ from f to g, we say that h is a homotopy *over* B if $p_2 h(x,t) = p_1(x)$ for all $x \in E_1$ and $t \in [0,1]$. In this situation we also say that h is a *vertical homotopy.*

Definition 2.6. Suppose given two spaces over B, say $p_1: E_1 \to B$ and $p_2: E_2 \to B$. They are *fiberwise homotopy equivalent* if there exist maps $u: E_1 \to E_2$, $v: E_2 \to E_1$ and homotopies α from vu to id_{E_1}, and β from uv to id_{E_2}, such that u, v and α, β are all over B.

Example 2.7. Let $B = \mathbb{R}$, $E_1 = \mathbb{R}$, $E_2 = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$, $p_1 = \mathrm{id}_\mathbb{R}$ and let $p_2 : E_2 \to \mathbb{R}$ be given by $(x, y) \mapsto x$. Then p_1 and p_2 are fiberwise homotopy equivalent. Namely, define $u : E_1 \to E_2$ by $u(x) = (x, 0)$ and $v : E_2 \to E_1$ by $v(x, y) = x$ and $\alpha(x, t) = x$ for $x \in E_1$, and $\beta((x, y), t) = (x, (1-t)y)$ for $(x, y) \in E_2$. However: p_1 is a fibration and p_2 isn't. (The path $[0, 1] \to B$ given by $t \mapsto t$ cannot be lifted to a path in E_2 with prescribed initial position $(0, 1)$, for example.)

Corollary 2.8. *The HLP is not a fibrewise homotopy invariant.*

Definition 2.9. [2]. Let $p : E \to B$ be a map. We say that p has the *weak homotopy lifting property*, WHLP, if for every space X and maps

$$f : X \times [0, 1] \to B, \qquad \bar{f}_0 : X \to E$$

such that $p\bar{f}_0(x) = f(x, 0)$ for all $x \in X$, there exists a map $\bar{f} : X \times [0, 1] \to B$ such that $p\bar{f} = f$ and the map $x \mapsto \bar{f}(x, 0)$ from X to E is *vertically* homotopic to $\bar{f}_0$. In that situation, the map p is called a *weak fibration.*

Proposition 2.10. *Suppose that $p_1 : E_1 \to B$ and $p_2 : E_2 \to B$ are fiberwise homotopy equivalent. If p_1 has the WHLP, then so does p_2.*

Proof. For each $s \in [0, 1]$ let $\iota_s : X \to X \times [0, 1]$ be given by $\iota_s(x) = (x, s)$. Let u, v, α, β be maps and homotopies as in definition 2.6. Let $f : X \times [0, 1] \to B$ and $\bar{f}_0 : X \to E_2$ be given such that $p_2\bar{f}_0(x) = f(x, 0)$ for all $x \in X$. Together, f and $\bar{f}_0$ make up a homotopy lifting problem for p_2. Then f together with $v\bar{f}_0$ constitute a homotopy lifting problem for p_1. Since p_1 has the WHLP, there exists $\bar{f} : X \times [0, 1] \to E_1$ such that $p_1\bar{f} = f$ and $\bar{f} \circ \iota_0$ is vertically homotopic to $v\bar{f}_0$. Then $u\bar{f} : X \times [0, 1] \to E_2$ is a homotopy such that $p_2 \circ (u\bar{f}) = f$ and $u\bar{f} \circ \iota_0$ is vertically homotopic to $u \circ (v\bar{f}_0) = (uv) \circ \bar{f}_0$, which is vertically homotopic to $\bar{f}_0$. This homotopy "solves" the homotopy lifting problem for p_2 that we started with, in the weak sense of the WHLP. □

Lemma 2.11. *Suppose that $p_1 : E_1 \to B$ and $p_2 : E_2 \to B$ are both weak fibrations. Let $u : E_1 \to E_2$ be a map over B which is an ordinary homotopy equivalence. Then u is a fiberwise homotopy equivalence.*

Proof. Exercise. The corresponding statement with "HLP" instead of "WHLP" is well known and a proof can be found in many/some books on homotopy theory. Try to adapt this proof. □

Corollary 2.12. *Let $p : E \to B$ be a weak fibration. Let $b \in B$ and let $F = p^{-1}(b) \subset E$. Then for any choice of base point $c \in F$, and any $n \geq 0$, we have $\pi_n(E, F, c) \cong \pi_n(B, b)$ (isomorphism induced by p).*

Proof. This was originally proved by Serre for Serre fibrations. You can find the proof in Spanier's book [10], but be warned that Spanier uses the expression "weak fibration" for a Serre fibration !! This proof works just as well for weak fibrations in our sense (i.e., maps having the WHLP). In fact this argument shows that "weak Serre fibration" is enough — see remark 2.21. □

Theorem 2.13. [2]. *Let $p : E \to B$ be any map, where B is paracompact. Suppose that each $x \in B$ has an open neighbourhood U_x in B such that $p^{-1}(U_x) \to U_x$ (the restriction of p) is a weak fibration. Then p itself is a weak fibration.* □

Corollary 2.14. *Suppose that B is paracompact. If $p : E \to B$ is any map which is "locally fiber homotopy trivial", then p is a weak fibration.*

Explanation. The "locally fiber homotopy trivial" assumption means: each x in B has a nhood U_x such that $p^{-1}(U_x) \to U_x$ is fiberwise homotopy equivalent to a "trivial" fibration of the form $F \times U_x \to U_x$ (projection of a product onto a factor).

Proof of corollary. Since a trivial fibration certainly has the WHLP, and since the WHLP is fiberwise homotopy invariant, we see that each $p^{-1}(U_x) \to U_x$ is a weak fibration. By the theorem, this implies that p is a weak fibration. □

Definition 2.15. A map $p\colon E \to B$ is a *microfibration* (has the micro-HLP) if, for every $f\colon X \times [0,1] \to B$ and $\bar{f}_0\colon X \to E$ with $p\bar{f}_0 = f\iota_0$, there exist a neighborhood U of $X \times \{0\}$ in $X \times [0,1]$ and $\bar{f}\colon U \to E$ such that $p\bar{f} = f|U$ and $\bar{f}\iota_0 = \bar{f}_0$.

Exercise 2.16. Show that if $p\colon E \to B$ is a fibration and $V \subset E$ is open, then $p|V$ is a microfibration.

Example 2.17. Let $p\colon \mathbb{R}^2 \to \mathbb{R}$ be the the standard projection given by $(x,y) \mapsto x$. This is obviously a (trivial) fibration. Let $V = \mathbb{R}^2 \smallsetminus \{(0,0)\}$. This is open in $\mathbb{R}^2$, so $p|V$ is a microfibration. But $p|V$ is not a fibration, since the fiber of $p|V$ over $0 \in \mathbb{R}$ is not homotopy equivalent to the fiber of $p|V$ over $1 \in \mathbb{R}$.

Exercise 2.18. (A digression which I started after the discussion of the example just above. It comes from Dugundji's topology book.) Let $p\colon M \to X$ be a fibration, where M is a closed nonempty manifold and X is any path–connected space having more than one point. Show that p is not nullhomotopic. (*Hint:* Suppose for a contradiction that it is ! Make a homotopy lifting problem out of a nullhomotopy for p and the "initial lift" given by the identity $M \to M$.)

Exercise 2.19. The following question was raised. Let M, N be smooth manifolds and let $f\colon M \to N$ be a smooth submersion (i.e., for every $x \in M$ the differential $df(x)\colon T_xM \to T_{f(x)}M$ is surjective). Is f a microfibration ? (I'm not sure. I can show that f is a Serre microfibration — see exercise 2.22 below.)

Exercise 2.20. Show that a map $p\colon E \to B$ which is a weak fibration and a microfibration is a fibration.

Remark 2.21. You can of course modify the definitions of *fibration, weak fibration, microfibration* by only asking for the HLP, WHLP, micro–HLP when the "test space" X is a CW–space. This gives the notions of *Serre fibration, weak Serre fibration, Serre microfibration.*

Exercise 2.22. Show that $p\colon E \to B$ is a Serre fibration (Serre microfibration) if the lifted homotopies (lifted microhomotopies) exist in all the cases where the test space X is a disk $\mathbb{D}^i$ for some $i \geq 0$.
(Does this reduction also work for weak Serre fibrations ?)

A useful extension of the HLP and micro–HLP is called HELP (homotopy extension lifting property), respectively micro–HELP. Let X be any space, $A \subset X$ a closed subset. We say that the inclusion $A \to X$ is a *cofibration* if $C = \{(x,t) \in X \times [0,1] \mid t = 0 \text{ or } x \in A\}$ is a *retract* of $X \times [0,1]$; that is, there exists a continuous map $r\colon X \times [0,1] \to C$ such that $r(x,t) = (x,t)$ whenever $(x,t) \in C$.

Proposition 2.23. *Adopt the notation of definition 2.1. Suppose that $p\colon E \to B$ has the HLP. Let $A \subset X$ be closed and suppose the inclusion $A \to X$ is a cofibration. Let g be the restriction of f to $A \times [0,1]$. Suppose that a lift $\bar{g}\colon A \times [0,1] \to E$ of g*

is given ($p\bar{g} = g$) which agrees with $\bar{f}_0$ on $A \times \{0\}$. Then $\bar{f}$ can be constructed so that it extends $\bar{g}$.

The proof is harder than I thought ! See proposition 6.44 in James [6]. There is a micro–version of that with essentially the same proof (and obvious changes in the statement). The book by James is actually a very thorough reference for fibrations and related matters.

3. Composition structures and fibrations

Definition 3.1. Let E, B be spaces and let $\sigma, \tau \colon E \to B$ be maps. Let $E\,{}_\sigma\!\times_\tau E := \{(x, y) \in E \times E \mid \sigma(x) = \tau(y)\}$. Both E and $E\,{}_\sigma\!\times_\tau E$ will be regarded as spaces over $B \times B$ using $x \mapsto (\tau(x), \sigma(x))$ for E, and $(x, y) \mapsto (\tau(x), \sigma(y))$ for $E\,{}_\sigma\!\times_\tau E$. A *composition structure* on (E, B, σ, τ) consists of

- a map $\kappa \colon E\,{}_\sigma\!\times_\tau E \longrightarrow E$ over $B \times B$,
- a map $\iota \colon B \to E$ such that $\sigma\iota = \mathrm{id}_B = \tau\iota$.

These are subject to the condition that the maps $E \to E$ given by $x \mapsto \kappa(\iota\tau(x), x)$ and $x \mapsto \kappa(x, \iota\sigma(x))$ are both homotopic to the identity, over $B \times B$.

Remark 3.2. This is easy to understand if you know a few things about categories. Think of points $a, b, c, \dots$ in B as "objects" and think of points $v, w, x, \dots$ in E as "morphisms" with sources and targets in B. The sources of $x \in E$ is $\sigma(x)$, the target is $\tau(x)$. The map κ tells you how to compose morphisms. The map ι singles out the identity morphisms (that is, $\iota(b)$ can be regarded as the identity morphism for b). What we do not ask for is associativity of the composition. Also, the composition is not required to be "strictly" unital: the identity morphisms are only required to behave like identity homomorphisms up to some vertical homotopies.

Definition 3.3. Keep the notation of definition 3.1. Let $q \colon Z \to B$ be some map. Let $E\,{}_\sigma\!\times_q Z = \{(x, z) \in E \times Z \mid \sigma(x) = q(z)\}$ and make this into a space over B by $(x, z) \mapsto \tau(x)$. — An *action* of E on Z is a map $\alpha \colon E\,{}_\sigma\!\times_q Z \to Z$ over B, subject to the following conditions.

- the maps $E\,{}_\sigma\!\times_\tau E\,{}_\sigma\!\times_q Z \to Z$ given by $(x, y, z) \mapsto \alpha(\kappa(x, y), z)$ and $(x, y, z) \mapsto \alpha(x, \alpha(y, z))$ are fiberwise homotopic over B, where the reference map from $E\,{}_\sigma\!\times_\tau E\,{}_\sigma\!\times_q Z$ to B is $(x, y, z) \mapsto \tau(x)$;
- the map $Z \to Z$ given by $z \mapsto \alpha(\iota(q(z)), z)$ for $z \in Z$ is homotopic over B to the identity id_Z.

Remark 3.4. This is easy to understand if you know a few things about categories and functors. Think of points $a, b, c, \dots$ in B as "objects" and think of the rule which takes $b \in B$ to the space $q^{-1}(b)$ as a "functor". Indeed any "morphism" in E, given by a point $x \in E$ with "source" $\sigma(x) = b$ and "target" $\tau(x) = c$, induces a map $q^{-1}(b) \to q^{-1}(c)$ given by evaluating α on elements of the form (x, z), with the fixed x and any $z \in q^{-1}(b) \subset Z$.

For the following lemma, keep the notation of definition 3.1. Suppose that B is a disk $\mathbb{D}^i$, and that $(\tau, \sigma) \colon E \to B \times B$ is a Serre microfibration.

Lemma 3.5. *Then for every $b \in B$ there exists a nhood U_b in B and maps*

- $m_{\mathrm{out}} \colon U_b \to E$,
- $m_{\mathrm{in}} \colon U_b \to E$,
- $h_{\mathrm{gen}} \colon U_b \times [0, 1] \to E$,

- $h_{\rm spec}\colon U_b \times [0,1] \to E$

such that for any $c \in U_b$,

- *the "morphism"* $m_{\rm out}(c)$ *has "source"* b *and "target"* c,
- *the "morphism"* $m_{\rm in}(c)$ *has "source"* c *and "target"* b,
- *the path* $t \mapsto h_{\rm gen}(c,t)$ *begins at the "composition"* $\kappa(m_{\rm out}(c), m_{\rm in}(c))$*, runs in* $\{x \in E \mid \tau(x) = c = \sigma(x)\}$ *and ends at* $\iota(c)$*;*
- *the path* $t \mapsto h_{\rm spec}(c,t)$ *begins at the "composition"* $\kappa((m_{\rm in}(c), m_{\rm out}(c))$*, runs in* $\{x \in E \mid \tau(x) = b = \sigma(x)\}$ *and ends at* $\iota(b)$*.*

Proof. Fix b. We can choose a neighborhood C of b in B which is a "cone" on something else, say ∂C, so that b itself is identified with the apex of the cone. (If b is not in the boundary of a disk, then ∂C can be a sphere; otherwise, a hemisphere.) Then we can write

$$C \cong [0,1] \times \partial C/\sim$$

where $\sim$ identifies all points of the form $(0,c)$ with b. Thus the inclusion $C \to B$ can be regarded as a homotopy from a constant map $\partial C \to B$ with value b to the inclusion $\partial C \to B$. This homotopy gives us *two* homotopies,

$$(g_s\colon \partial C \to B \times B)_{s\in[0,1]} \quad \text{and} \quad (h_s\colon \partial C \to B \times B)_{s\in[0,1]}\ ,$$

by setting either the second or the first output coordinate equal to b. Thus (g_s) is a homotopy from the constant map with value (b,b) to the map $c \mapsto (c,b)$, and (h_s) is a homotopy from the constant map with value (b,b) to the map $c \mapsto (b,c)$. For both of these homotopies, we have preferred initial lifts $\bar g_0\colon \partial C \to E$ and $\bar h_0\colon \partial C \to E$ which are constant with value $\iota(b)$. ("Lift" means: if you compose with the map $(\tau,\sigma)\colon E \to B \times B$, you get g_0 and h_0, respectively.) Using the micro–HLP for $(\tau,\sigma)\colon E \to B\times B$, we therefore get the map $m_{\rm out}$ as a "micro–lift" of the homotopy (g_s) , and $m_{\rm in}$ as a "micro–lift" of the homotopy (h_s). The maps $m_{\rm out}$, $m_{\rm in}$ are defined on a (small) neighborhood of b in C. (We obtain this as a neighborhood of $\{0\} \times \partial C$ in $[0,1] \times \partial C$, which we then divide out by $\{0\} \times \partial C$.)
The construction of the homotopies $h_{\rm gen}$ and $h_{\rm spec}$ is very similar, except that we need the relative HLP (alias HELP). Let

$$D = C \times [0,1], \qquad \partial D = \partial C \times [0,1]\,,$$

so that we can identify D with a quotient of $[0,1]\times\partial D$. In particular the two maps from D to $B \times B$ given by $(c,t) \mapsto (c,c)$ and $(c,t) \mapsto (b,b)$ can then be regarded as two homotopies

$$(G_s\colon \partial D \longrightarrow B \times B)_{s\in[0,1]}\,, \quad (H_s\colon \partial D \longrightarrow B \times B)_{s\in[0,1]}$$

starting (at time $s = 0$) with the constant map $(c,t) \mapsto (b,b)$ from ∂D to $B \times B$, and ending (at time $s = 1$) with $(c,t) \mapsto (c,c)$ and $(c,t) \mapsto (b,b)$, respectively. (The second of these homotopies is "stationary".) We have a preferred initial lift for both, given by the map

$$\partial D \to E$$

taking $(c,t) \in \partial C \times [0,1] = D$ to $\omega(t)$, where ω is a path in E such that $\omega(0) = \kappa(\iota(b), \iota(b))$ and $\omega(1) = \iota(b)$. (This path is meant to run in the fiber of the map $(\tau,\sigma)\colon E \to B\times B$ over the point (b,b); its existence is guaranteed by definition 3.1.) We also have preferred lifts for the restricted homotopies $(G_s|\partial C \times \{0,1\})$ and

$(H_s|\partial C \times \{0,1\})$. These lifted homotopies can be defined as maps from $C \times \{0,1\}$ to E, by the formulae

$$(c,t) \mapsto \begin{cases} \kappa(m_{\text{out}}(c), m_{\text{in}}(c)) & t=0, \text{ case of } (G_s) \\ \iota(c) & t=1, \text{ case of } (G_s) \\ \kappa((m_{\text{in}}(c), m_{\text{out}}(c)) & t=0, \text{ case of } (H_s) \\ \iota(b) & t=1, \text{ case of } (H_s). \end{cases}$$

From the micro–HELP of proposition 2.23, we now obtain lifted micro–homotopies, micro–lifting (G_s) and (H_s). What we get is in fact two maps defined on open subsets of

$$[0,1] \times \partial C \times [0,1]$$

containing $\{0\} \times \partial C \times [0,1]$. Here the left–hand copy of $[0,1]$ serves as the timekeeper for the homotopy, but we can change that and we will. Remembering also that $C \cong [0,1] \times \partial C/\sim$, we we view these maps as being defined on open neighborhoods of $\{b\} \times [0,1]$ in $C \times [0,1]$. Without loss of generality, these open neighborhoods have the form $U_b \times [0,1]$ for the same U_b, and so our maps can be viewed as two "homotopies" between certain maps from U_b to E. □

Corollary 3.6. *Keep the notation and assumptions of definition 3.3. Suppose that $(\tau,\sigma): E \to B \times B$ and $q: Z \to B$ are Serre microfibrations. Then q is actually a Serre fibration.*

Proof. We begin with some easy reductions. Firstly, we can reduce to the case where B is a disk. Namely, suppose that $f: X \times [0,1] \to B$ is a homotopy (with B still arbitrary) which we want to lift across $q: Z \to B$, with an initial lift $\bar{f}_0: X \to Z$. Since we are going for the Serre fibration property, we may assume that X is a disk $\mathbb{D}^i$. But then $B' := X \times [0,1]$ is also (homeomorphic to) a disk $\mathbb{D}^{i+1}$. We can now use $f: B' \to B$ to pull the entire homotopy lifting problem and the action data back to B'. Thus we replace B by B' and Z by

$$Z' = f^*Z = \{(c,z) \in B' \times Z \mid f(c) = q(z)\}$$

and E by $E' = \{(e,c,d) \in E \times B' \times B' \mid \tau(e) = f(c), \sigma(e) = f(d)\}$. The homotopy f itself can be replaced by the identity $X \times [0,1] \to B'$ (but try to forget that it is an identity map) and the initial lift $\bar{f}_0$ can be replaced by the map

$$X \longrightarrow Z' \subset B' \times Z$$

whose second coordinate is $\bar{f}_0$ and whose first coordinate is the inclusion of $X \cong X \times \{0\}$ to $X \times [0,1]$. If this new homotopy lifting problem (with base space B') has a solution (a lifted homotopy), then that solution determines a solution for the old homotopy lifting problem (just compose with projection $Z' \to Z$).

Secondly, we only need to show that $q: Z \to B$ is a *weak* Serre fibration because of exercise 2.20 (modified by attaching the "Serre" prefix everywhere). That is what we will do, assuming that B is a disk. And in that case we will even show that q is a weak fibration (no "Serre" prefix).

Thirdly, to show that $q: Z \to B$ is a weak fibration when B is a disk, we only need to show that it is "locally fiber homotopy trivial", because of corollary 2.14.

But this last statement is almost obvious from lemma 3.5. Given $b \in B$ we choose U_b as in that lemma, along with the maps and homotopies m_{out}, m_{in}, h_{gen} and h_{spec}. Now fix some $c \in U_b$. Then we have a preferred choice of maps

$$q^{-1}(b) \longrightarrow q^{-1}(c)\,,\ q^{-1}(c) \longrightarrow q^{-1}(b)$$

given by acting on the left with $m_{\rm out}(c)$ and $m_{\rm in}(c)$, respectively (using the "action" α described in definition 3.3). These two maps are homotopy inverses of each other. Indeed, the composite maps $q^{-1}(c) \longrightarrow q^{-1}(c)$ and $q^{-1}(b) \longrightarrow q^{-1}(b)$ are homotopic to the respective identity maps by means of the homotopies given by the left actions of $h_{\rm gen}(t,c)$ on $q^{-1}(c)$, and $h_{\rm spec}(t,c)$ on $q^{-1}(b)$, respectively, for $t \in [0,1]$. Allowing c to vary, we have a homotopy equivalence

$$q^{-1}(b) \times U_b \longrightarrow q^{-1}(U_b)$$

which is, by construction, a fiberwise homotopy equivalence over U_b. This completes the proof. ☐

Remark. Revaz Kurdiani noted that definitions 3.1 and 3.3 can be formalized in the following manner. The category of spaces over $B \times B$ is a *monoidal category* with multiplication $\Box$ given by

$$E \Box E' := \{(x,y) \in E \times E' \mid \sigma(x) = \tau'(y)\}$$

for spaces E and E' over $B \times B$, with reference maps $(\tau, \sigma)\colon E \to B \times B$ and $(\tau', \sigma')\colon E' \to B \times B$. Note that $E \Box E'$ is again a space over $B \times B$ with reference map $(x,y) \mapsto (\tau(x), \sigma'(y))$. There is a *unit object* for the multiplication given by the diagonal $B \to B \times B$, viewed as a space over $B \times B$.
Next, the category of spaces over $B \times B$ with the above multiplication "acts" on the category of spaces over B by

$$E \Box Z := \{(x,z) \in E \times Z \mid \sigma(x) = q(z)\}$$

for spaces E over $B \times B$ and Z over B, with reference maps $(\tau,\sigma)\colon E \to B \times B$ and $q\colon Z \to B$. In both categories, we have notions of homotopy (fiberwise over $B \times B$ and over B, respectively). What we see in definitions 3.1 and 3.3 could be called a *Hopf object* in the monoidal category of spaces over $B \times B$, and a *Hopf action* of that on an object in the category of spaces over B.

4. Examples of fibrations et al. in differential topology

Let M and N be smooth manifolds. Assume $\partial N = \emptyset$, but let's not assume that $\partial M = \emptyset$. We denote by $C^\infty(M,N)$ the set of all smooth maps from M to N. This comes with a preferred topology, the compact–open C^∞ topology. This is described in [5], but I want to make a little clearer why it is called "compact–open".
For an integer $k \geq 0$, a *k–jet* from M to N is an equivalence class of triples (x,f,y) where $x \in M$, $y \in N$, and f is a smooth map from an open neighborhood of x to N such that $f(x) = y$. Two such triples (x_1,f_1,y_1) and (x_2,f_2,y_2) are equivalent if $x_1 = x_2 =: x$ and $y_1 = y_2 =: y$, and in some local coordinate charts about $x \in M$ and $y \in N$, the maps f_1 and f_2 have the same k–th Taylor polynomial at x. (If this is the case for some choice of local charts, then it will also be the case for any other choice of local charts.) The set of equivalence classes is denoted by $J^k(M,N)$. There are obvious projections $J^k(M,N) \to M$ and $J^k(M,N) \to N$ given by $(x,f,y) \mapsto x$ and $(x,f,y) \mapsto y$ on representatives.
Example. Let $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ be open. Then $J^k(U,V)$ is easily identified with $U \times P \times V$ where P is the vector space of polynomial maps $\mathbb{R}^m \to \mathbb{R}^n$ of degree $\leq k$ taking 0 to 0. This identification takes the class of (x,f,y) to the k–th Taylor polynomial at 0 of the map

$$z \mapsto y + f(z - x)$$

(where $z \in \mathbb{R}^m$ is close to 0). In this case, clearly, $J^k(U,V)$ has the structure of a smooth manifold.
In the general case, we can also put a smooth manifold structure on $J^k(M,N)$ by noting that any pair of charts $\varphi\colon U \to M$, $\psi\colon V \to N$ with U open in $\mathbb{R}^m$ and V open in $\mathbb{R}^n$ determines an injection $J^k(U,V) \to J^k(M,N)$. These injections constitute a differentiable atlas for $J^k(M,N)$. There are obvious projections $J^k(M,N) \to M$ and $J^k(M,N) \to N$ given by
Another important thing to note is that every smooth $f\colon M \to N$ gives rise to a smooth map $j^kf\colon M \to J^k(M,N)$, the *$k$–jet prolongation* of f. This takes $x \in M$ to the jet defined by $(x, f, f(x))$. The composition

$$M \xrightarrow{j^kf} J^k(M,N) \xrightarrow{\text{proj}} M$$

is the identity (we say that j^kf is a *section* of the projection $J^k(M,N) \to M$) and the composition

$$M \xrightarrow{j^kf} J^k(M,N) \xrightarrow{\text{proj}} N$$

is equal to f. Hence f and j^kf determine each other. The information contained in $j^kf(x)$, for $x \in M$, is the k–th Taylor polynomial of f at x, but redefined in a coordinate free way. (Note that the main theorem of immersion theory, as formulated in the introduction, is precisely about the 1–jet prolongation $f \mapsto j^1f$, applied to immersions f.)

Now we are ready for the compact–open C^∞ topology on $C^\infty(M,N)$. Take an integer $k \geq 0$, a compact subset L of M and an open subset U of $J^k(M,N)$. Let $V(k,L,U) \subset C^\infty(M,N)$ consist of all smooth $f\colon M \to N$ such that $j^kf(L) \subset U$. We decree that a subset of $C^\infty(M,N)$ is open if it is a union (possibly an infinite union) of sets of the form $V(k,L,U)$. It is an exercise to check that the conditions for a topology are satisfied.

Example 4.1. If M is compact, then $\mathbf{imm}(M,N) \subset C^\infty(M,N)$ is an open subset. (If M is noncompact, this is usually not the case. Show that $\mathbf{imm}(\mathbb{R}^m,\mathbb{R}^n)$ is not open in $C^\infty(\mathbb{R}^m,\mathbb{R}^n)$ if $m \leq n$.)

Next we have a nasty lemma about constructing smooth functions with prescribed higher derivatives. This is essentially due to E. Borel.

Lemma 4.2. *Let L be a smooth compact manifold. For $i = 0, 1, 2, \dots$ let $f_i\colon L \to \mathbb{R}$ be smooth functions. Then there exists a smooth $F\colon L \times \mathbb{R} \to \mathbb{R}$ such that the i–th partial derivative of F in the $\mathbb{R}$ direction, evaluated along $L \times \{0\} \cong L$, equals f_i.*

Proof. (I reproduce the proof here from [7] and another source because it proves more than the lemma states, and we need that extra information.) There is no loss of generality in assuming that L is a codimension zero compact smooth submanifold of a euclidean space $\mathbb{R}^n$. (Otherwise, embed L in a euclidean space, and replace it by the total space of a normal disk bundle in the euclidean space.) The advantage which we have from that is that we can use standard notation for partial derivatives. To begin with fix a smooth function $\rho\colon \mathbb{R} \to \mathbb{R}$ such that $\rho(t) = 1$ for $|t| \leq \frac{1}{2}$ and $\rho(t) = 0$ for $|t| \geq 1$. Set

$$(*) \qquad F(x,t) = \sum_{i=0}^{\infty} \frac{t^i}{i!} \rho(\mu_i t) f_i(x)$$

where the (large) real numbers $\mu_i \geq 1$ are yet to be determined. We want to choose them in such a way that the series

$$(**) \qquad \sum_{i=0}^{\infty} D^{\alpha}\left(\frac{t^i}{i!}\rho(\mu_i t) f_i(x)\right)$$

is uniformly convergent for every multi–index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{n+1})$. If that can be achieved, then F is well defined and $(*)$ can be differentiated term by term, and F solves our problem.

To determine the numbers μ_i, write the i–th term in $(*)$ in the form

$$(\mu_i)^{-i} f_i(x) \cdot (i!)^{-1} (\mu_i t)^i \rho(\mu_i t) = (\mu_i)^{-i} f_i(x) \cdot \psi_i(\mu_i t)$$

where $\psi_i(t) = (i!)^{-1} t^i \rho(t)$. Next let

$$M_i = \max \left\{ D^{\alpha}\left(f_i(x)\psi_i(t)\right) \;\middle|\; (x,t) \in L \times \mathbb{R},\ |\alpha| < i \right\}.$$

(Because ψ_i vanishes outside $[-1,1]$, the maximum can be taken over (x,t) in the compact set $L \times [-1,1]$ and over the finitely many α which satisfy $|\alpha| < i$.) Since $\mu_i > 1$, it follows for $|\alpha| < i$ that

$$|i\text{–th element in } (**)| \leq (\mu_i)^{|\alpha|}(\mu_i)^{-i} M_i \leq M_i \mu_i^{-1}.$$

Now choose $\mu_i = \max\{1, 2^i M_i\}$. Then for fixed α and for any $i > |\alpha|$, the i–th term of $(**)$ is bounded by 2^{-i}. □

But let's not stop there. The construction of F in terms of the f_i is quite explicit. The only "random" choice which we made was the choice of the function ρ, which really should be made once and for all at the beginning. Then the construction amounts to a map of the form

$$\prod_{i=0}^{\infty} C^{\infty}(L, \mathbb{R}) \quad \longrightarrow \quad C^{\infty}(L \times \mathbb{R}, \mathbb{R}) \quad .$$

This map is *continuous* (with the product topology in the LHS). To verify this, let's first observe that the formula for each number $\mu_i = \mu_i(f_0, f_1, f_2, \ldots)$ is continuous as a function of the variables $f_0, f_1, f_2, \ldots$. In fact it depends only on f_i and that dependence can be expressed in terms of the values of f_i and the partial derivatives of f_i , of order $< i$. (The derivatives of ρ are also involved in that expression, but only those of order $< i$. Each of them has a maximum since ρ vanishes outside a compact interval.) Next we need to know that $(**)$ depends continuously on $(f_0, f_1, f_2, \ldots)$ for fixed α. Write $(**)$ as a sum

$$\sum_{i=0}^{k} D^{\alpha}\left(\frac{t^i}{i!}\rho(\mu_i t) f_i(x)\right) + \sum_{i=k+1}^{\infty} D^{\alpha}\left(\frac{t^i}{i!}\rho(\mu_i t) f_i(x)\right)$$

where k is larger than $|\alpha|$. The continuous dependence of each μ_i on $(f_0, f_1, f_2, \ldots)$ implies that the first of the two summands depends continuously on $(f_0, f_1, f_2, \ldots)$. For the other summand, we have the bound $2^{-k} + 2^{-k-1} + 2^{-k-2} + \cdots = 2^{1-k}$, which we can make as small as we like by choosing k large.

We are therefore in a position to formulate the following astonishing corollary to (the proof of) E. Borel's lemma:

Corollary 4.3. *The map from $C^{\infty}(L \times \mathbb{R}, \mathbb{R})$ to $\prod_{i \geq 0} C^{\infty}(L, \mathbb{R})$ given by*

$$F \mapsto \left(\frac{\partial^i F}{\partial t^i}\Big|_{t=0} \right)_{i=0,1,2,\ldots}$$

has a continuous right inverse.

Let M and N be smooth as before, M compact. Let $M_0 \subset M$ be a compact codimension zero submanifold. (Boundary legislation is as follows: We require that there exist a smooth map $f\colon M \to \mathbb{R}$ such that f and $f|\partial M$ are transverse to 0, and such that $M_0 = \{x \in M \mid f(x) \leq 0\}$. This is equivalent to saying that M_0 looks "locally" like $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^{m-2}$ in $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^{m-2}$, where $\mathbb{R}_+ = \{z \in \mathbb{R} \mid z \geq 0\}$. In particular M_0 is a manifold with corners if $M_0 \cap \partial M$ is nonempty.)

Lemma 4.4. *The restriction map $\rho\colon C^\infty(M,N) \to C^\infty(M_0,N)$ is a Serre fibration.*

Proof. This proceeds by a few easy reductions.
Step 1. We show that M can be replaced by any compact codimension zero submanifold M_1 of M which is a neighbourhood of M_0. Indeed, suppose that we have a homotopy lifting problem for ρ, consisting of a map $f\colon \mathbb{D}^i \times [0,1] \longrightarrow C^\infty(M_0,N)$ an an "initial lift"

$$\bar{f}_0\colon \mathbb{D}^i \longrightarrow C^\infty(M,N)$$

such that $\rho\bar{f}_0(z) = f(z,0)$ for $z \in \mathbb{D}^i$. We obtain an easier homotopy lifting problem in which we keep f but replace $\bar{f}_0$ by the composition

$$\mathbb{D}^i \xrightarrow{\ \bar{f}_0\ } C^\infty(M,N) \xrightarrow{\ \text{res.}\ } C^\infty(M_1,N).$$

Suppose that this easier lifting problem has a solution

$$\bar{g}\colon \mathbb{D}^i \times [0,1] \longrightarrow C^\infty(M_1,N).$$

Now choose a smooth map $\psi\colon M \to [0,1]$ with support contained in M_1 such that $\psi(x) = 1$ for all $x \in M_0$. Then define

$$\bar{f}\colon \mathbb{D}^i \times [0,1] \longrightarrow C^\infty(M,N).$$

by the formula

$$\bar{f}(z,t)(x) := \begin{cases} \bar{g}(z,\psi(x)\cdot t)(x) & (x \in M_1) \\ \bar{f}_0(z)(x) & (x \in M \smallsetminus M_1). \end{cases}$$

Step 2. We show that N can be replaced by a euclidean space. The idea is to choose a smooth embedding $e\colon N \to \mathbb{R}^k$ with closed image and a smooth retraction $r\colon U \to e(N)$, where U is an open neighborhood of $e(N)$ in $\mathbb{R}^k$. Let

$$\rho'\colon C^\infty(M,\mathbb{R}^k) \longrightarrow C^\infty(M_0,\mathbb{R}^k)$$

be given by restriction. Suppose that we have a homotopy lifting problem for ρ, consisting of a map f and an initial lift $\bar{f}_0$ as in step 1. This determines a homotopy lifting problem for ρ' in the obvious way. Let's assume that this new homotopy lifting problem has a solution

$$\mathbb{D}^i \times [0,1] \longrightarrow C^\infty(M,\mathbb{R}^k).$$

We can assume that the image of this is contained in $C^\infty(M,U)$ (otherwise replace M by a sufficiently small open neighbourhood M_1 of M_0 as in step 1). Then we can compose with the map $C^\infty(M,U) \to C^\infty(M,N)$ induced by the projection $e^{-1}r\colon U \to N$.
Step 3. We note that, if $N = \mathbb{R}^k$, we may as well assume $N = \mathbb{R}$ since $\mathrm{map}(M,\mathbb{R}^k)$ is a product of k copies of $\mathrm{map}(M,\mathbb{R})$, and similarly for $\mathrm{map}(M_0,\mathbb{R}^k)$.
Step 4. Our homotopy lifting problems (and solutions) are unaffected if we delete from M_0 and M any subset of the interior of M_0. We may therefore assume that

$M_0 = L\times[-1,0]$ and $M = L\times[-1,1]$ where L is a smooth manifold with boundary.
Step 5. We take $M_0 = L \times [-1,0]$ and $M = L \times [-1,1]$, where L is smooth with boundary. We take $N = \mathbb{R}$. In this case the restriction map ρ is a linear map of topological vector spaces. Because of the linear structure, the lifting problem can be simplified. We start with a homotopy

$$f\colon \mathbb{D}^i \times [0,1] \longrightarrow C^\infty(L \times [-1,0], \mathbb{R})$$

and we look for *any* lift $\bar{f}\colon \mathbb{D}^i\times[0,1] \longrightarrow C^\infty(L\times[-1,1],\mathbb{R})$. What if the restriction of $\bar{f}$ to $\mathbb{D}^i \times \{0\}$ is not what we want it to be ? Let d be the difference between "what we want it to be" and "what it is". Then we can improve $\bar{f}$ by adding the composition

$$\mathbb{D}^i \times [0,1] \xrightarrow{\text{proj.}} \mathbb{D}^i \xrightarrow{\ d\ } C^\infty(L \times [-1,1], \mathbb{R}).$$

Step 6. The situation is now that $M_0 = L \times [-1,0]$ and $M = L \times [-1,1]$, where L is smooth with boundary, and $N = \mathbb{R}$. We start with a map

$$f\colon K \longrightarrow C^\infty(L \times [-1,0], \mathbb{R})$$

and we look for a lift $\bar{f}\colon K \to C^\infty(L\times[-1,1],\mathbb{R})$, where K is compact Hausdorff (a disk would be general enough). To get such a lift we note that the restriction map

$$C^\infty(L \times [-1,1], \mathbb{R}) \longrightarrow C^\infty(L \times [-1,0], \mathbb{R})$$

has a (continuous) right inverse. This follows easily from corollary 4.3. Note that a smooth function on $L\times[-1,+1]$ is as good as a pair of smooth functions, one on $L\times[-1,0]$ and the other on $L\times[0,1]$, subject to the condition that the values and higher partial derivatives in the interval direction agree on $L \times \{0\}$. □

Corollary 4.5. *The restriction map* $\mathbf{imm}(M,N) \to \mathbf{imm}(M_0,N)$ *is a Serre microfibration.*

Proof. See example 4.1. □

The next item is a re–arrangement and a more precise formulation of the main theorem of immersion theory. This relies on a lemma where we compare immersions of codimension > 0 with immersions of codimension 0.
We fix M and N as before, of dimension m and n, respectively. Let $V \to M$ be a vector bundle, of fiber dimension $n-m$, with a Riemannian metric. Let $\mathbb{D}(V)$ be the associated disk bundle. We are interested in (codimension zero) immersions $\mathbb{D}(V) \to N$. An immersion

$$f\colon \mathbb{D}(V) \to N$$

determines an immersion $g\colon M \to N$ by restriction to the zero section, and an isomorphism ι of the vector bundle $V \to M$ with the "normal bundle" of g. The normal bundle of the immersion g is $g^*TN/\mathrm{im}(dg)$, also known as the cokernel of the differential $dg\colon TM \to g^*TN$. The isomorphism ι is simply df, or more precisely, the thing you get if you "divide"

$$T\mathbb{D}(V)_{|M} \xrightarrow{\ df\ } f^*TN_{|M} = g^*TN$$

by appropriate vector subbundles (namely, the tangent bundle TM in the source, and $\mathrm{im}(dg)$ in the target).
For a fixed vector bundle V on M as above, let $\mathbf{imm}_V(M,N)$ be the space of pairs (g,ι) where g is *any* immersion $M \to N$ and ι is *any* isomorphism of V with the normal bundle of g.

Lemma 4.6. *The above map* $\mathbf{imm}(\mathbb{D}(V),N) \longrightarrow \mathbf{imm}_V(M,N)$ *is a weak homotopy equivalence.*

Proof sketch. Choose a Riemannian metric on N. Then every $(g,\iota) \in \mathbf{imm}_V(M,N)$ determines a map $g^{(\iota)}$ from a neighborhood of the zero section of V to N, as follows. Let's describe elements of V as pairs (x,v) with $x \in M$ and $v \in V$. For such an (x,v) let $\gamma_{(x,v)}$ be the unique geodesic in N whose initial position is $g(x)$ and whose initial velocity is $\iota(x,v)$, where $\iota(x,v) \in T_{g(x)}N$ should be regarded as a *normal* vector to $dg(T_xM)$. This makes sense since we have an inner product on $T_{g(x)}(N)$. We put

$$g^{(\iota)}(x,v) = \gamma_{(x,v)}(1)$$

if the right–hand side is defined. It is indeed defined for all (x,v) in some neighbourhood of the zero section. By inspection, the differential of $g^{(\iota)}$ at any point $(x,0)$ in the zero section of V is an invertible linear map

$$T_{(x,v)}M \cong T_xM \oplus V_x \longrightarrow T_{g(x)}(M)\,.$$

By continuity, the same will hold for all (x,v) in a suitable smaller open neighbourhood V' of the zero section in V, so that $g^{(\iota)}|V'$ is an immersion. Using radial shrinking, we can find an embedding e of $\mathbb{D}(V)$ in V' which is the identity on another (even smaller) neighborhood of the zero section in V. Then $g^{(\iota)}e$ is an immersion from $\mathbb{D}(V)$ to N.
The construction gives a map $\pi_0\mathbf{imm}_V(M,N) \longrightarrow \pi_0\mathbf{imm}(\mathbb{D}(V),N)$ which is seen to be inverse to the map on π_0 induced by $\mathbf{imm}(\mathbb{D}(V),N) \longrightarrow \mathbf{imm}_V(M,N)$. It is straightforward to generalize the construction to a parametrized setting, and to deduce that our map $\mathbf{imm}(\mathbb{D}(V),N) \longrightarrow \mathbf{imm}_V(M,N)$ induces a bijection on π_n for any $n \geq 0$ and any choice of base point in $\mathbf{imm}(\mathbb{D}(V),N)$. □

Corollary 4.7. *The following is a (weak) homotopy pullback square:*

$$\begin{array}{ccc} \mathbf{imm}(\mathbb{D}(V),N) & \xrightarrow{\text{1-jet prolong.}} & \mathbf{fimm}(\mathbb{D}(V),N) \\ \text{restr.}\downarrow & & \downarrow\text{restr.} \\ \mathbf{imm}(M,N) & \xrightarrow{\text{1-jet prolong.}} & \mathbf{fimm}(M,N)\,. \end{array}$$

Proof. By the lemma, we may replace $\mathbf{imm}(\mathbb{D}(V),N)$ by $\mathbf{imm}_V(M,N)$. By inspection, we may also replace $\mathbf{fimm}(\mathbb{D}(V),N)$ by the space $\mathbf{fimm}_V(M,N)$ of triples $(f,\delta f,\iota)$ where $(f,\delta f) \in \mathbf{fimm}(M,N)$ and ι is a vector bundle isomorphism from V to $\mathrm{coker}(\delta f) = f^*TN/\mathrm{im}(\delta f)$. Then our diagram turns into

$$\begin{array}{ccc} \mathbf{imm}_V(M,N) & \xrightarrow{\text{1-jet prolong.}} & \mathbf{fimm}_V(M,N) \\ \text{restr.}\downarrow & & \downarrow\text{restr.} \\ \mathbf{imm}(M,N) & \xrightarrow{\text{1-jet prolong.}} & \mathbf{fimm}(M,N)\,. \end{array}$$

It is a strict pullback square and it is easy to verify that the vertical arrows are Serre fibrations. Hence it is a weak homotopy pullback square. □

We now come to a slight reformulation of the promised theorem 1.3 which, in some respects, looks weaker than the original. Using the corollary above, we will be able to show that it is stronger.

Theorem 4.8. *Let L and N be smooth manifolds of the same dimension n, where L is compact and N has empty boundary. Assume that L has no closed component. Then the jet prolongation map* $\mathbf{imm}(L,N) \to \mathbf{fimm}(L,N)$ *is a weak homotopy equivalence.*

Reduction of theorem 1.3 to theorem 4.8. Let M, N be as in theorem 1.3. Fix some point $(f,\delta f)$ in $\mathbf{fimm}(M,N)$. This determines a formal normal vector bundle $\mathrm{coker}(\delta f) = f^*TN/\mathrm{im}(\delta f)$. Apply corollary 4.7 with V equal or isomorphic to $\mathrm{coker}(\delta f)$. By theorem 4.8, which we are supposed to believe, the top row of the diagram in that corollary is a weak homotopy equivalence. Hence the bottom row is also a weak homotopy equivalence *provided* we restrict to the path component of $\mathbf{fimm}(M,N)$ which contains $(f,\delta f)$ and the path components of $\mathbf{imm}(M,N)$ mapping to that path component of $\mathbf{fimm}(M,N)$ under jet prolongation. But since $(f,\delta f)$ was arbitrary, we see that $\mathbf{imm}(M,N) \to \mathbf{fimm}(M,N)$ has to be a weak homotopy equivalence. □

We are getting closer to the proof of theorem 4.8 with the following lemma.

Lemma 4.9. *Let L be smooth, compact, without closed components, of dimension n. Then L admits a handle decomposition with no handles of index n. Equivalently, L admits a smooth Morse function $f\colon L \to [0,1]$ such that $f^{-1}(1) = \partial L$, df is nonzero on T_xL for all $x \in \partial L$, and f has no local maxima in $L \smallsetminus \partial L$.*

Vocabulary. For Morse functions, see [1], [5] or [7]. A *handle decomposition* of L is really a filtration

$$\emptyset = L_0 \subset L_1 \subset L_2 \subset L_3 \subset \cdots \subset L_k = L.$$

Here each L_i is a smooth compact n–manifold with boundary, and L_i can be obtained from L_{i-1} by attaching a handle of index m_i, which is a copy of

$$\mathbb{D}^{m_i} \times \mathbb{D}^{n-m_i}\ ,$$

using an embedding of $\mathbb{S}^{m_i-1} \times \mathbb{D}^{n-m_i}$ in ∂L_{i-1}. Some smoothing of corners is required or recommended. Note that m_1 has to be 0 so that $\mathbb{S}^{m_1-1} \times \mathbb{D}^{n-m_1}$ can be embedded in $\partial L_0 = \emptyset$. Hence L_1 has to be $\cong \mathbb{D}^n$. A manifold with a handle decomposition is also called a *handlebody*.

Suppose that $f\colon L \to [0,1]$ is a Morse function as in the lemma, with critical points $p_1, p_2, p_3, \dots, p_k$. We can assume $f(p_{i-1}) \le f(p_i)$ for $i = 2, \dots, k$ and after a small perturbation, if necessary, we can also assume $f(p_i) < f(p_{i+1})$. Then we can choose numbers $v_1 < v_2 < v_3 < \dots v_k = 1$ such that $v_{i-1} < f(p_i) < v_i$ for $i = 2, \dots, k$. Now put

$$L_i := \{x \in L \mid f(x) \le v_i\}$$

for $i \ge 1$. This filtration is a handle decomposition (except for the fact that, in order to obtain L_i from L_{i-1}, you first have to attach an outward collar $\cong \partial L_{i-1} \times [0,\varepsilon]$, which doesn't really change anything, and then a handle). The index of the handle which you have to attach to make L_i from L_{i-1} is equal to the Morse index of the critical point p_i. So a Morse function determines a handle decomposition. There is a converse, i.e., any handle decomposition as above can be "realized" by a Morse function. These assertions are part of Morse theory.

Using Morse theory is not always the best way to make or simplify handle decompositions. Another method is to use smooth triangulations. A triangulation of

L is a homeomorphism $X \to L$ where X is a simplicial complex (and we identify that with L in the following). The triangulation is *smooth* if, for each simplex in L, the characteristic embedding $\Delta^i \to L$ is a smooth embedding (i.e., it is smooth and its differential at any point of Δ^i is an injective linear map). Now suppose that L comes with a smooth triangulation $L \cong X$ (such a triangulation always exists). Choose a filtration of $L \cong X$ by subcomplexes,

$$\emptyset = X_0 \subset X_1 \subset X_2 \subset X_3 \subset \cdots \subset X_k = X$$

such that X_i has exactly one more simplex than X_{i-1}. (This is always possible.) Let L_i be a regular neighbourhood of X_i (that's a lot like a tubular neighbourhood). Then the filtration

$$\emptyset = L_0 \subset L_1 \subset L_2 \subset L_3 \subset \cdots \subset L_k = L$$

is a handle decomposition of L (with one handle of index m for each simplex of dimension m in the triangulation).

Proof of lemma 4.9. We can assume that L is connected. Choose a smooth triangulation for L. Associated with that, we have a handle decomposition, as explained above, but it will have many handles of index n since the triangulation has simplices of dimension n. Make an embedded graph $\Gamma \subset L$ as follows: one vertex for each n–simplex, situatued at the barycentre of that simplex; one edge for any two adjacent n–simplices, going from one barycenter to the other, straight through the barycentre of the $(n-1)$–simplex which is the unique common face of the two adjacent n–simplices. For the graph Γ, choose a maximal subtree $T \subset \Gamma$. In the tree T, choose a vertex p such that the corresponding n–simplex meets ∂L. For each vertex $q \in T$ there is a unique minimal path from p to q in T; let d_q be the number of edges involved and (for $q \neq p$) let $e(q)$ be the last edge in that path, the one that contains q. This gives us a way to pair off each simplex of dimension n with one of its faces of dimension $n-1$. Namely, an n–simplex corresponding to a vertex q in T, where $q \neq p$, is paired off with the unique $(n-1)$–simplex which meets the edge $e(q)$. For the n–simplex corresponding to p, we improvise by choosing some $(n-1)$–dimensional face which belongs to ∂L. We now "cancel" handles in pairs. First we cancel the n–handle corresponding to p against its selected partner $(n-1)$–handle, by removing both of them. It should be clear (from a suitable picture that you should have drawn by now !) that this does not change the diffeomorphism type of the handlebody under consideration. Then we cancel some n–handle corresponding to a vertex q in T with $d_q = 1$ against the $(n-1)$–handle corresponding to the edge $e(q)$, by removing both handles. Again, this does not change the diffeomorphism type of the handlebody. We repeat the procedure until all vertices q in T with $d_q = 1$ have been taken care of. Then we remove some n–handle corresponding to a vertex q with $d_q = 2$ and its partner $(n-1)$–handle corresponding to the edge $e(q)$. And so on. At the end, there are no n–handles left. □

The proof of theorem 4.8 will be given by induction on the number of handles in a handle decomposition of L, without any handles of index $n = \dim(L)$. Here comes the induction step. We consider a standard handle

$$\mathbb{D}^k \times \mathbb{D}^{n-k}$$

and a codimension zero subset of the form $A \times \mathbb{D}^{n-k}$ where A is an annulus. That is, A consists of all $z \in \mathbb{D}^i$ having distance $\geq \frac{1}{2}$ from the centre.

Lemma 4.10. *The restriction map* $\mathbf{imm}(\mathbb{D}^k \times \mathbb{D}^{n-k}, N) \longrightarrow \mathbf{imm}(A \times \mathbb{D}^{n-k})$ *is a Serre fibration provided* $n > k$.

Proof. This will be deduced from corollary 3.6. We let $Z = \mathbf{imm}(\mathbb{D}^k \times \mathbb{D}^{n-k}, N)$ and $B = \mathbf{imm}(A \times \mathbb{D}^{n-k}, N)$, with $q\colon Z \to B$ equal to the restriction map. Let

$$E \quad = \quad \mathbf{imm}(\mathbb{S}^{k-1} \times [0,3] \times \mathbb{D}^{n-k}, N)\,.$$

Let $\sigma\colon E \to B$ be given by composition with $(tz, z') \mapsto (z, 2t-1, z')$ for $z \in \mathbb{S}^{k-1}$, $t \in [\frac{1}{2}, 1]$ and $z' \in \mathbb{D}^{n-k}$. Similarly let $\tau\colon E \to B$ be given by composition with the map $(tz, z') \mapsto (z, 2t+1, z')$. To obtain $\kappa\colon E\,{}_\sigma\!\times_\tau E \longrightarrow E$ choose an identification of the colimit of

$$[0,3] \xleftarrow{t \mapsto 2t+1} [\tfrac{1}{2}, 1] \xrightarrow{t \mapsto 2t-1} [0,3]$$

with $[0,3]$ which extends the identity on the left–hand copy of $[0,1]$ and on the right–hand copy of $[2,3]$. Similarly, to obtain the action map $\alpha\colon E\,{}_\sigma\!\times_q Z \to Z$ choose an appropriate identification of the colimit of

$$\mathbb{D}^k \xleftarrow{\text{scal. mult.}} \mathbb{S}^{k-1} \times [\tfrac{1}{2}, 1] \xrightarrow{t \mapsto 2t-1} \mathbb{S}^{k-1} \times [0,3]$$

with $\mathbb{D}^k$. All that is straightforward. We will write $f * g$ instead of $\kappa(f,g)$, assuming $f, g \in E$ are composable, and we will also write $f * g$ for $\alpha(f,g)$, assuming $f \in E$ and $g \in Z$ are composable.

The real challenge consists in constructing $\iota\colon B \to E$. To do that, we first introduce certain regular neighborhoods of $A \times \mathbb{D}^{n-k}$ and $\mathbb{D}^k \times \mathbb{D}^{n-k}$, respectively. Let Q be the product of an annulus with radii $(\frac{1}{4}, \frac{5}{4})$ about $0 \in \mathbb{R}^k$ and a disk of radius 2 about $0 \in \mathbb{R}^{n-k}$. Let Q' be the product of a disk of radius $\frac{5}{4}$ about $0 \in \mathbb{R}^k$ and a disk of radius 2 about $0 \in \mathbb{R}^{n-k}$. It follows from corollary 4.3 that the restriction map

$$\mathbf{imm}(Q, N) \longrightarrow \mathbf{imm}(A \times \mathbb{D}^{n-k}, N)$$

admits a section

$$s\colon \mathbf{imm}(A \times \mathbb{D}^{n-k}, N) \longrightarrow \mathbf{imm}(Q, N).$$

[Without loss of generality, N is a smooth submanifold of $\mathbb{R}^i$. Let U be a tubular neighborhood of N. Construct a section s_1 of $\mathbf{imm}(Q, \mathbb{R}^i) \longrightarrow \mathbf{imm}(A \times \mathbb{D}^{n-k}, \mathbb{R}^i)$ using corollary 4.3. Restrict s_1 to $\mathbf{imm}(A \times \mathbb{D}^{n-k}, N)$ and, if necessary, modify the restriction so that its image is contained in $\mathbf{imm}(Q, U)$. Then compose with the map $\mathbf{imm}(Q, U) \to \mathbf{imm}(Q, N)$ given by composition with the retraction $U \to N$.]

The same reasoning shows that the restriction map

$$\mathbf{imm}(Q', N) \longrightarrow \mathbf{imm}(\mathbb{D}^k \times \mathbb{D}^{n-k}, N)$$

admits a section

$$\bar{s}\colon \mathbf{imm}(\mathbb{D}^m \times \mathbb{D}^{n-k}, N) \longrightarrow \mathbf{imm}(Q', N).$$

which covers s (so that $q\bar{s} = sq$).

For $f \in B$ we want to define $\iota(f)$ by a formula of type

$$\iota(f) = s(f) \circ u$$

where $u\colon \mathbb{S}^{k-1} \times [0,3] \times \mathbb{D}^{n-k} \longrightarrow C$ is an immersion yet to be determined, independent of f. (We have to ensure that $\tau(u) = \sigma(u) = j$, where $j\colon A \times \mathbb{D}^{n-k} \longrightarrow Q$ is the inclusion.) We want to define a vertical homotopy from $g \mapsto \iota(g) * g$ to the identity $g \mapsto g$ on Z by a formula of type

$$(g, t) \mapsto \bar{s}(g) \circ v_t$$

where $(g,t) \in Z \times [0,1]$ and $(v_t)_{t\in[0,1]}$ is a regular homotopy of immersions from $u * j$ to j, yet to be determined. It is enough to do the case $k=1$, provided we can do this so that u and each v_t are equivariant for the "antipodal" action of $\mathbb{Z}/2$. [Solutions for larger k can be obtained from a solution for $k=1$ by taking product with $\mathbb{S}^{k-1}$, and dividing out by the diagonal–antipodal action of $\mathbb{Z}/2$.] We can also take out a common factor $\mathbb{D}^{n-k-1}$ in the source and the target of u and/or v_t , which means that we can assume $n-k=1$ as well. Assuming now $k=1$ and $n-k=1$, we are looking first of all for an immersion

$$u\colon \mathbb{S}^0 \times [0,3] \times \mathbb{D}^1 \longrightarrow \mathbb{S}^0 \times [\tfrac{1}{4},\tfrac{5}{4}] \times \mathbb{D}^1$$

which is equivariant (for the antipodal action of $\mathbb{Z}/2$). The immersion u is prescribed on $\mathbb{S}^0 \times [0,1] \times \mathbb{D}^1$ and on $\mathbb{S}^0 \times [2,3] \times \mathbb{D}^1$, where it has to agree with j up to certain reparametrizations. If we make an identification $\mathbb{S}^1 \cong [0,3]/\sim$ where the relation identifies $t \in [0,1]$ with $t+2 \in [2,3]$, then we are looking for an (equivariant) immersion

$$u\colon \mathbb{S}^0 \times \mathbb{S}^1 \times \mathbb{D}^1 \longrightarrow \mathbb{S}^0 \times \mathbb{S}^0 \times [\tfrac{1}{4},\tfrac{5}{4}] \times \mathbb{D}^1$$

prescribed on $\mathbb{S}^0 \times J \times \mathbb{D}^1$ for a certain arc J in $\mathbb{S}^1$. Given that after the construction of u we still have to construct the regular homotopy (v_t), any choice of u which has winding number 0 on the core circles is a good choice. (Such an immersion u will not embed the core circles; in the simplest case, each core circle is mapped to a "figure eight" in the rectangle $[\frac{1}{4},\frac{5}{4}] \times \mathbb{D}^1$. Pictures would help at this point ...) With such a choice of u, the construction of (v_t) is an easy exercise. $\square$

Let P be smooth, compact, n–dimensional. Let an embedding of $\mathbb{S}^{k-1} \times \mathbb{D}^{n-k}$ in ∂P be given, with $k<n$ as before. We use it to attach a handle $H = \mathbb{D}^k \times \mathbb{D}^{n-k}$ to P. The resulting manifold $P \cup H$ can also be written as $L \cup H$, where L is the union of P and the thickened annulus $A \times \mathbb{D}^{n-k} \subset \mathbb{D}^k \times \mathbb{D}^{n-k} = H$. Note that L is diffeomorphic to P (after smoothing some corners).

Corollary 4.11. *Let N be smooth, n–dimensional, without boundary. The commutative square of restriction maps*

$$\begin{array}{ccc} \mathbf{imm}(L\cup H,N) & \longrightarrow & \mathbf{imm}(L,N) \\ \downarrow & & \downarrow \\ \mathbf{imm}(H,N) & \longrightarrow & \mathbf{imm}(L\cap H,N) \end{array}$$

is a (weak) homotopy pullback square.

Proof. It is clearly a pullback square. Moreover, the lower horizontal arrow is a Serre fibration by lemma 4.10. $\square$

Proof of theorem 4.8. We begin by observing that the commutative square of restriction maps

$$\begin{array}{ccc} \mathbf{fimm}(L\cup H,N) & \longrightarrow & \mathbf{fimm}(L,N) \\ \downarrow & & \downarrow \\ \mathbf{fimm}(H,N) & \longrightarrow & \mathbf{fimm}(L\cap H,N) \end{array}$$

is a homotopy pullback square. [It is again a pullback square and all the maps in it are fibrations by inspection.] Combining that observation with corollary 4.11, we see that if the jet prolongation map

$$\mathbf{imm}(Y,N) \to \mathbf{fimm}(Y,N)$$

is a weak homotopy equivalence for $Y = H, L, L \cap H$, then it is also a weak homotopy equivalence for $Y = L \cup H$. This gives us an induction argument. In detail, we note first (to make an induction beginning) that

$$\mathbf{imm}(Y,N) \to \mathbf{fimm}(Y,N)$$

is indeed a (weak) homotopy equivalence if Y is a disk, alias "handle". This is already contained in corollary 4.7 [the special case where M is a single point]. Then we fix an $r > 1$ and a $k < n$, and suppose that $\mathbf{imm}(Y,N) \to \mathbf{fimm}(Y,N)$ is known to be a weak homotopy equivalence for all Y having a handle decomposition with fewer than r handles, and also for all Y having a handle decomposition with handles of index $< k$ only. In the situation where $Y = L \cup H$ as above, we may suppose that H is the "last" of altogether r handles in a handle decomposition of Y (and H has index k). Now L has a handle decomposition with $r-1$ handles, H has a handle decomposition with only one handle, and $H \cap L$ has a handle decomposition with handles of index $< k$ only. So L, H and $L \cap H$ are covered by our inductive assumption. □

Remark. The main ingredients of the above proof are clearly corollary 3.6 and lemma 4.10. In the booklet which accompanies the film *Outside In* about eversing the 2–sphere in $\mathbb{R}^3$, Bill Thurston writes about his "corrugation" approach to immersion theory (corrugation as in *corrugated cardboard*). It was one of my goals to make this transparent. Corrugation just "happens" at the end of the proof of lemma 4.10.

5. Submersion theory and Gromov's theorem

A *submersion* is a smooth map $f: L \to N$ with the property that, for each $x \in L$, the differential $T_xL \to T_{f(x)}M$ is surjective. Here, as before, N should be without boundary, L can have a nonempty boundary, but we pay no special attention to the tangent spaces $T_x\partial L$ for $x \in \partial L$. Hence the restriction of a submersion $L \to N$ to ∂L need not be a submersion.
Assuming that L is compact, let $\mathbf{subm}(L,N)$ be the space of smooth submersions from L to N, an open subspace of $C^\infty(L,N)$. Also let $\mathbf{fsubm}(L,N)$ be the space of formal submersions from L to N. An element in $\mathbf{fsubm}(L,N)$ is a pair $(f, \delta f)$ where $f: M \to N$ is a continuous map and $\delta f: TL \to f^*TN$ is a vector bundle surjection. There is a jet prolongation map

$$\mathbf{subm}(L,N) \to \mathbf{fsubm}(L,N)$$

given by $f \mapsto (f, df)$.

Theorem 5.1. *Let L and N be smooth manifolds, where L is compact and N has empty boundary. Assume that L has no closed component. Then the jet prolongation map* $\mathbf{subm}(L,N) \to \mathbf{fsubm}(L,N)$ *is a weak homotopy equivalence.*

Remark. This is uninteresting if $\dim(L) < \dim(N) = n$, because then both $\mathbf{subm}(L,N)$ and $\mathbf{fsubm}(L,N)$ are empty. If $\dim(L) = \dim(N)$, we recover theorem 4.8, because in that case

$$\mathbf{subm}(L,N) = \mathbf{imm}(L,N), \quad \mathbf{fsubm}(L,N) = \mathbf{fimm}(L,N).$$

The proof of theorem 5.1 is not just similar to the proof of theorem 4.8, but in fact identical. Nevertheless it is worth highlighting a few points.
First of all, if M is a compact smooth manifold, then $\mathbf{subm}(M,N)$ is an open subspace of $C^\infty(M,N)$. Hence the analogue of corollary 4.5 for submersions is valid. We need this, of course, in order to have access to corollary 3.6. Then we can prove the analogue of lemma 4.10 for submersions. The statement is that the restriction map

$$\mathbf{subm}(\mathbb{D}^k\times\mathbb{D}^{\ell-k},N)\longrightarrow\mathbf{subm}(A\times\mathbb{D}^{\ell-k})$$

is a Serre fibration provided $\ell>k$. Here ℓ need not be equal to $n=\dim(N)$, but we may as well assume $\ell\geq n$; otherwise there is nothing to prove. For the proof, let $Z=\mathbf{subm}(\mathbb{D}^k\times\mathbb{D}^{\ell-k},N)$ and $B=\mathbf{subm}(A\times\mathbb{D}^{\ell-k},N)$, with $q\colon Z\to B$ equal to the restriction map, and

$$E\quad=\quad\mathbf{subm}(\mathbb{S}^{k-1}\times[0,3]\times\mathbb{D}^{\ell-k},N)\,.$$

In short, replace *immersion* by *submersion* wherever the opportunity arises. However, one point should be emphasized. When you reach the passage ... *for $f\in B$ we want to define $\iota(f)$ by a formula of type*

$$\iota(f)=s(f)\circ u$$

where $u\colon\mathbb{S}^{k-1}\times[0,3]\times\mathbb{D}^{\ell-k}\longrightarrow C$ is an immersion yet to be determined ... then you may or may not replace the word *immersion* by *submersion*. It does not matter because we are talking about a codimension zero situation. Thus, the proof of lemma 4.10 in the setting $\ell>n$ involves not only submersions where the dimension drops from ℓ to n, but also submersions/immersions of codimension zero (dimension ℓ mapping to dimension ℓ). □

There is an even more general result which can be proved with exactly the same arguments. Fix a "target" manifold N and integers $m,k>0$ and a subset $\mathfrak{W}$ of the jet space $J^k(\mathbb{R}^m,N)$. Suppose that

(i) $\mathfrak{W}$ is open in $J^k(\mathbb{R}^m,N)$
(ii) $\mathfrak{W}$ is invariant under "local diffeomorphisms" of $\mathbb{R}^m$. [This means that if you have a diffeomorphism $\varphi\colon U\to V$ between open subsets of $\mathbb{R}^m$, and some jet $s\in J^k(V,N)\cap\mathfrak{W}$, then $s\circ\varphi\in J^k(U,N)\cap\mathfrak{W}$.]

Now suppose that M is any smooth m–manifold, possibly with boundary. Let $\mathfrak{W}_M\subset J^k(M,N)$ consist of all the jets which, in local coordinate charts about their source in M, belong to $\mathfrak{W}\subset J^k(\mathbb{R}^m,\mathbb{R}^n)$. [Because of the diffeomorphism invariance condition, it does not matter how you choose the coordinate charts. But see the remark just below.] Let $C^\infty_{\mathfrak{W}}(M,N)$ consist of all the smooth maps from M to N whose k–jets at any point $x\in M$ belong to $\mathfrak{W}_M$. Let $\Gamma(\mathfrak{W}_M)$ be the space of continuous sections (with the compact–open C^0 topology) of the bundle projection

$$\mathfrak{W}_M\longrightarrow M\,.$$

Remark. Officially, an element of $J^k(\mathbb{R}^m,\mathbb{R}^n)$ is represented by a triple (x,U,f) where $x\in\mathbb{R}^m$ and f is a smooth map from a neighborhood U of x to $\mathbb{R}^n$. Unofficially, however, an element of $J^k(\mathbb{R}^m,\mathbb{R}^n)$ is a collection of numbers, one for each possible "mixed partial derivative" $\partial^\alpha/\partial x_\alpha$ where $|\alpha|\leq k$. With the second definition, it is easy to represent elements of $J^k(\mathbb{R}^m,\mathbb{R}^n)$ by "slightly less" than the above — for example, a triple (x,U,f) where x belongs to a hyperplane $\mathbb{R}^{m-1}\times\{0\}\subset\mathbb{R}^m$

and U is a neighborhood of x in $\mathbb{R}^{m-1} \times [0, \infty[$, and $f : U \to \mathbb{R}^n$ is smooth. Hence a smooth map $M \to N$ has well defined k–jets at any point of M, even at points in ∂M.

Theorem 5.2 (Gromov)**.** *Suppose that M has no closed component. Then the jet prolongation map $C^\infty_{\mathfrak{W}}(M, N) \longrightarrow \Gamma(\mathfrak{W}_M)$ is a weak homotopy equivalence.*

Proof. Once you have unravelled the meaning, you will see that it can be proved exactly like the submersion theorem. The following is important: Given

$$f \in C^\infty_{\mathfrak{W}}(M, N)$$

and given any codimension zero immersion/submersion $u : L \to M$, the composition $f \circ u$ belongs to $C^\infty_{\mathfrak{W}}(L, N)$. The reason is that codimension zero immersions/submersions are locally diffeomorphic (inverse function theorem !) and our assumptions on $\mathfrak{W}$ include some diffeomorphism invariance. □

Gromov's proof of this theorem (about 1970) is not very different from the one I have given. But I have tried to give an argument specifically for "topologists", for example by exploiting Dold's basic results on fibrations, which Gromov doesn't do.

References

[1] T. Bröcker and K. Jänich, *Introduction to differential topology*, Camb. Univ. Press, 1982, German edition Springer-Verlag (1973).

[2] A. Dold, *Partitions of unity in the theory of fibrations*, Ann. of Math. **78** (1963), 223–255.

[3] A. Haefliger, *Lectures on the theorem of Gromov*, Proc. of 1969/70 Liverpool Singularities Symp., Lecture Notes in Math., vol. 209, Springer, 1971, pp. 128–141.

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

[5] ———, *Differential topology*, Grad. Texts in Math., Springer, 1976.

[6] I. M. James, *General topology and homotopy theory*, Springer, 1984.

[7] M.Golubitsky and V. Guillemin, *Stable mappings and their singularities*, Grad. Texts in Math., Springer, 1973.

[8] J. Milnor, *Morse theory*, Ann. of Math. Studies, Princeton Univ. Press, 1963 (1969).

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

[10] E. H. Spanier, *Algebraic topology*, McGraw–Hill, New York, 1966.