The sphere eversion project

Patrick Massot Oliver Nash Floris van Doorn

December 15, 2022

Introduction

This project had two goals. First we wanted to check whether a proof assistant can do differential topology. Many people still think that formal mathematics are mostly suitable for algebra, combinatorics, or foundational studies. So we chose one of the most famous examples of geometric topology theorems associated to tricky geometric intuition: the existence of sphere eversions. Note however that we won't focus on any of the many videos of explicit sphere eversions. We will prove a general theorem which immediately implies the existence of sphere eversions.

The second goal of this project was to experiment using a formalization blueprint that evolves with the project until we get a proof that has very closely related formal and informal presentations. A full proof (by normal pen and paper standards) was written before the formalization effort began. This proof evolved a lot during the formalization. In particular, the chapter on the global theory required a lot of work during the formalization in order to ensure that its technical lemmas are both fully correct and actually sufficient for our purposes.

In this introduction, we will describe the mathematical context of this project, the main definitions and statements, and outline the proof strategy.

Gromov observed that it's often fruitful to distinguish two kinds of geometric construction problems. He says that a geometric construction problem satisfies the h-principle if the only obstructions to the existence of a solution come from algebraic topology. In this case, the construction is called flexible, otherwise it is called rigid. This definition is purposely vague. We will see a rather general way to give it a precise meaning, but one must keep in mind that such a precise meaning will fail to encompass a number of situations that can be illuminated by the h-principle dichotomy point of view.

The easiest example of a flexible construction problem which is not totally trivial and is algebraically obstructed is the deformation of immersions of circles into planes. Let f_0 and f_1 be two maps from $\mathbb{S}^1$ to $\mathbb{R}^2$ that are immersions. Since $\mathbb{S}^1$ has dimension one, this mean that both derivatives f_0' and f_1' are nowhere vanishing maps from $\mathbb{S}^1$ to $\mathbb{R}^2$. The geometric object we want to construct is a (smooth) homotopy of immersions from f_0 to f_1, ie a smooth map $F\colon \mathbb{S}^1 \times [0,1] \to \mathbb{R}^2$ such that $F|_{\mathbb{S}^1\times\{0\}} = f_0$, $F|_{\mathbb{S}^1\times\{1\}} = f_1$, and each $f_p := F|_{\mathbb{S}^1\times\{p\}}$ is an immersion. If such a homotopy exists then, $(t,p) \mapsto f_p'(t)$ is a homotopy from f_0' to f_1' among maps from $\mathbb{S}^1$ to $\mathbb{R}^2 \setminus \{0\}$. Such maps have a well defined winding number $w(f_i') \in \mathbb{Z}$ around the origin, the degree of the normalized map $f_i'/\|f_i'\|\colon \mathbb{S}^1 \to \mathbb{S}^1$. So $w(f_0') = w(f_1')$ is a necessary condition for the existence of F, which comes from algebraic topology. The Whitney–Graustein theorem states that this necessary condition is also sufficient. Hence this geometric construction problem is flexible. One can give a direct proof of this result, but it also follows from general results proved in this project (although we haven't formalized this consequence of our work).

An important lesson from the above example is that algebraic topology can give us more than a necessary condition. Indeed the (one-dimensional) Hopf degree theorem ensures that, provided $w(f_0') = w(f_1')$, there exists a homotopy g_p of nowhere vanishing maps relating f_0' and f_1'. We also know from the topology of $\mathbb{R}^2$ that f_0 and f_1 are homotopic, say using the straight-line homotopy $p \mapsto f_p = (1-p)f_0 + pf_0$. But there is no a priori relation between g_p and the derivative of f_p for $p \notin \{0, 1\}$. So we can restate the crucial part of the Whitney–Graustein theorem as: there is a homotopy of immersion from f_0 to f_1 as soon as there is (a homotopy from f_0 to f_1) and a homotopy from f_0' to f_1' among nowhere vanishing maps. The parenthesis in the previous sentence indicated that this condition is always satisfied, but it is important to keep in mind for generalizations. Gromov says that such a homotopy of uncoupled pairs (f, g) is a formal solution of the original problem.

One can generalize this discussion of uncoupled maps replacing a map and its derivative for maps from a manifold M to a manifold N. The so called 1-jet space $J^1(M, N)$ is the space of triples (m, n, φ) with $m \in M$, $n \in N$, and $\varphi \in \mathrm{Hom}(T_mM, T_nN)$, the space of linear maps from T_mM to T_nN. One can define a smooth manifold structure on $J^1(M, N)$, of dimension $\dim(M)+\dim(N)+\dim(M)\dim(N)$ which fibers over M, N and their product $J^0(M, N) := M \times N$. Beware that the notation (m, n, φ) does not mean that $J^1(M, N)$ is a product of three manifolds, the space where φ lives depends on m and n. Any smooth map $f \colon M \to N$ gives rise to a section $j^1 f$ of $J^1(M, N) \to M$ defined by $j^1 f(m) = (m, f(m), T_m f)$. Such a section is called a *holonomic section* of $J^1(M, N)$. In the Whitney–Graustein example, we use the canonical trivialization of $T\mathbb{S}^1$ and $T\mathbb{R}^2$ to represent $j^1 f$ has a pair of maps (f, f'). The role played by (f, g) in this example is played in general by sections of $J^1(M, N) \to M$ which are not necessarily holonomic.

One can generalize this discussion to $J^r(M, N)$ which remembers derivatives of maps up to order r for some given $r \geq 0$. One can also consider sections of an arbitrary bundle $E \to M$ instead of functions from M to N, which are sections of the trivial bundle $M \times N \to N$. But the case of $J^1(M, N)$ is sufficient for our purposes.

Definition. *A first order differential relation $\mathcal{R}$ for maps from M to N is a subset of $J^1(M, N)$. A solution of $\mathcal{R}$ is a function $f \colon M \to N$ such that $j^1 f(m)$ is in $\mathcal{R}$ for all m. A* formal solution *of $\mathcal{R}$ is a non-necessarily holonomic section of $J^1(M, N) \to M$ which takes value in $\mathcal{R}$.*

The partial differential relation $\mathcal{R}$ satisfies the h-principle if any formal solution σ of $\mathcal{R}$ is homotopic, among formal solutions, to some holonomic one $j^1 f$.

For instance, an immersion of M into N is a solution of

$$\mathcal{R} = \{(m, n, \varphi) \in J^1(M, N) \mid \varphi \text{ is injective}\}.$$

As we saw with the Whitney–Graustein problem, we are not only interested to individual solutions, but also in families of solutions. In differential topology, a smooth family of maps between manifolds X and Y is a smooth map $h \colon P \times X \to Y$ seen as the collection of maps $h_p : x \mapsto h(p, x)$. Here P stands for "parameter space". A smooth family of sections of $E \to X$ is a smooth family of maps $\sigma \colon P \times X \to E$ such that each σ_p is a section.

In such a case it is important that we start with a family of formal solutions that is holonomic for some values of the parameter and we don't modify it for those parameters. In the curve example $P = [0, 1]$, the formal solution is holonomic for parameters 0 and 1, and we want to keep the start and end curves. In this work we don't use manifolds with boundary when it is not necessary so we rather use $\mathbb{R}$ as a parameter space.

More generally it can also happen that a family of formal solutions $\sigma : P \times M \to J^1(M, N)$ has the property that σ_p is holonomic at $m \in M$ for some values of p and m and we want to preserve σ near the corresponding set in $P \times M$. This leads to the following definition.

Definition. *A partial differential relation $\mathcal{R} \subset J^1(M,N)$ satisfies the relative and parametric h-principle if every family of formal solutions $\sigma\colon M \times P \to J^1(M,N)$ which are holonomic for (p,m) near some closed set $C \subset P \times M$, is homotopic to a family of holonomic sections, and this homotopy can be chosen constant near C.*

One can also insist on the deformed solution to be C^0-close to the original one. In this case one talks about a C^0-dense h-principle.

Using this vocabulary, we can state the Smale-Hirsch immersion theorem as saying that the relation of immersions satisfies all forms of the h-principle provided the dimension of the target manifold is larger than the dimension of the source.

This theorem covers the Whitney–Graustein theorem (in its second form, assuming the existence of a homotopy between derivatives). But there are much less intuitive applications. The most famous one is the existence of sphere eversions: one can "turn $\mathbb{S}^2$ inside-out among immersions of $\mathbb{S}^2$ into $\mathbb{R}^3$).

Corollary (Smale 1958)**.** *There is a homotopy of immersion of $\mathbb{S}^2$ into $\mathbb{R}^3$ from the inclusion map to the antipodal map $a\colon q \mapsto -q$.*

The reason why this is turning the sphere inside-out is that a extends as a map from $\mathbb{R}^3 \setminus \{0\} \to \mathbb{R}^3 \setminus \{0\}$ by

$$\hat{a}\colon q \mapsto -\frac{1}{\|q\|^2} q$$

which exchanges the interior and exterior of $\mathbb{S}^2$. More abstractly, one can say the normal bundle of $\mathbb{S}^2$ is trivial, hence one can extend a to a tubular neighborhood of $\mathbb{S}^2$ as an orientation preserving map. Since a is orientation reversing, any such extension will be reversing coorientation.

Proof of the sphere eversion corollary. We denote by ι the inclusion of $\mathbb{S}^2$ into $\mathbb{R}^3$. We set $j_t = (1-t)\iota + ta$. This is a homotopy from ι to a (but not an immersion for $t = 1/2$). We need to check there is no obstruction to building a homotopy of formal solutions above those maps. One could show that the relevant homotopy group (replacing $\pi_1(\mathbb{S}^1)$ from the Whitney–Graustein example) is $\pi_2(\mathrm{SO}_3(\mathbb{R}))$. This group is trivial, hence there is no obstruction. But actually we can write an explicit homotopy here, without computing $\pi_2(\mathrm{SO}_3(\mathbb{R}))$. Using the canonical trivialization of the tangent bundle of $\mathbb{R}^3$, we can set, for $(q,v) \in T\mathbb{S}^2$, $G_t(q,v) = \mathrm{Rot}^{\pi t}_{Oq}(v)$, the rotation around axis Oq with angle πt. The family $\sigma\colon t \mapsto (j_t, G_t)$ is a homotopy of formal immersions relating $j^1\iota$ to $j^1 a$. The above theorem ensures this family is homotopic, relative to $t = 0$ and $t = 1$, to a family of holonomic formal immersions, ie a family $t \mapsto j^1 f_t$ with $f_0 = \iota$, $f_1 = a$, and each f_t is an immersion. □

The Smale-Hirsch theorem and its above corollary follows from a more general theorem: the h-principle for open and ample first order differential relations (see below). We will prove this theorem using a technique which is even more general: convex integration. For instance this technique also underlies the constructions of paradoxical isometric embeddings, which could be a nice follow-up project.

We'll end this introduction by describing the key construction of convex integration, since it is very nice and elementary. Convex integration was invented by Gromov around 1970, inspired in particular by the C^1 isometric embedding work of Nash and the original proof of flexibility of immersions. This term is pretty vague however, and there are several different implementations. The newest one, and by far the most efficient one, is Mélanie Theillière's corrugation process from 2018. And this is what we will use.

Let f be a map from $\mathbb{R}^n$ to $\mathbb{R}^m$. Say we want to turn f into a solution of some partial differential relation. For instance if we are interested in immersions, we want to make sure its differential is everywhere injective. We will ensure this by tackling each partial derivative in turn. In the immersion example, we first make sure $\partial_1 f(x) := \partial f(x)/\partial x_1$ is non-zero for all x. Then we make sure $\partial_2 f(x)$ is not colinear to $\partial_1 f(x)$. Then we make sure $\partial_3 f(x)$ is not in the plane spanned by the two previous derivatives, etc… until all n partial derivatives are everywhere linearly independent.

In general, what happens is that, for each number j between 1 and n, we wish $\partial_j f(x)$ could live in some open subset $\Omega_x \subset \mathbb{R}^m$. Assume there is a smooth family of loops $\gamma \colon \mathbb{R}^n \times \mathbb{S}^1 \to \mathbb{R}^m$ such that each γ_x takes values in Ω_x, and has average value $\fint_{\mathbb{S}^1} \gamma_x = \partial_j f(x)$. Obviously such loops can exist only if $\partial_j f(x)$ is in the convex hull of Ω_x, hence the name convex integration, and we will see this condition is almost sufficient. In the immersion case, this convex hull condition will always be met because, from the above description, we see that Ω_x will always be the complement of a linear subspace with codimension at least two.

For some large positive N, we replace f by the new map

$$x \mapsto f(x) + \frac{1}{N} \int_0^{Nx_j} \left[\gamma_x(s) - \partial_j f(x)\right] ds.$$

A wonderfully easy exercise shows that, provided N is large enough, we have achieved $\partial_j f(x) \in \Omega_x$, almost without modifying derivatives $\partial_i f(x)$ for $i \neq j$, and almost without moving $f(x)$. See 2.3 for a precise statement.

In addition, if we assume that γ_x is constant (necessarily with value $\partial_j f(x)$) for x near some subset K where $\partial_j f(x)$ was already good, then nothing changed on K since the integrand vanishes there. It is also easy to damp out this modification by multiplying the integral by a cut-off function. So this is a very local construction, and it isn't obvious how the absence of homotopical obstruction, embodied by the existence of a formal solution, should enter the discussion. The answer is that is essentially provides a way to coherently choose base points for the γ_x loops.

Now that we've seen how convex hulls enter the discussion we can provide one last definition and state the actual main theorem that we formalized.

Definition. *A relation $\mathcal{R} \subset J^1(M, N)$ is ample if, for every $(x, y, \varphi) \in \mathcal{R}$ and every hyperplane $H \subset T_x M$, the convex hull of the connected component of φ in*

$$\left\{\psi \in \operatorname{Hom}(T_x M, T_y N) \mid \psi|_H = \varphi|_H \text{ and } (x, y, \psi) \in \mathcal{R}\right\}$$

in the whole set of ψ such that $\psi|_H = \varphi|_H$.

We can now state our goal in its full glory.

Theorem (Gromov)**.** *For any manifolds M and N, any relation $\mathcal{R} \subset J^1(M, N)$ that is open and ample satisfies the full h-principle (relative, parametric and C^0-dense).*

Chapter 1 provides the loops supply. Chapter 2 then discusses the local theory, including the key construction above, and Chapter 3 finally moves to manifolds, and proves the main theorem and its sphere eversion corollary. Appendix A explains how the first two chapters are already enough to derive Smale's theorem, although in a slightly less natural way than using the manifold case. This served as an intermediate target in the formalization, and can be used for elementary teaching since it does not require any theory of manifolds.

Chapter 1

Loops

1.1 Introduction

In this chapter, we explain how to construct families of loops to feed into the corrugation process explained at the end of the introduction.

Throughout this document, E and F will denote finite-dimensional real vector spaces.

Definition 1.1. *A loop is a map defined on the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ with values in a finite-dimensional vector space. It can also freely be seen as 1-periodic maps defined on $\mathbb{R}$.*

The average of a loop γ is $\bar{\gamma} := \int_{\mathbb{S}^1} \gamma(s)\, ds$.

The support *of a family γ of loops in F parametrized by E is the closure of the set of x in E such that γ_x is not a constant loop.*

All of this chapter is devoted to proving the following proposition.

Proposition 1.2. *Let K a compact set in E. Let Ω be an open set in $E \times F$.*

Let β and g be smooth maps from E to F. Write $\Omega_x := \{y \in F \mid (x, y) \in \Omega\}$, assume that $\beta(x) \in \Omega_x$ for all x, and that $g(x) = \beta(x)$ near K.

If, for every x, $g(x)$ is in the convex hull of the connected component of Ω_x containing $\beta(x)$, then there exists a smooth family of loops

$$\gamma\colon E \times [0, 1] \times \mathbb{S}^1 \to F, (x, t, s) \mapsto \gamma_x^t(s)$$

such that, for all $x \in E$, all $t \in \mathbb{R}$ and all $s \in \mathbb{S}^1$,

- $\gamma_x^t(s) \in \Omega_x$
- $\gamma_x^0(s) = \gamma_x^t(1) = \beta(x)$
- $\bar{\gamma}_x^1 = g(x)$
- $\gamma_x^t(s) = \beta(x)$ *if x is near K.*

Let us briefly sketch the geometric idea behind the above proposition if we pretend there is only one point x, and drop it from the notation, and also focus only on γ^1. By assumption, there is a finite collection of points p_i in Ω and $\lambda_i \in [0, 1]$ such that g is the barycenter $\sum \lambda_i p_i$. Since Ω is open and connected, there is a smooth loop γ_0 which goes through each p_i. The claim is that g is the average value of $\gamma = \gamma_0 \circ h$ for some self-diffeomorphism h of $\mathbb{S}^1$. The idea is to choose h such that γ rushes to p_1, stays there during a time roughly λ_1, rushes to p_2, etc. But, in order to achieve average exactly g, it seems like h needs to be a discontinuous piecewise constant map. The assumption that Ω is open means that the convex hull is open, which gives enough slack to get away with a smooth h.

In the previous proof sketch, there is a lot of freedom in constructing γ, which is problematic when trying to do it consistently when x varies.

1.2 Surrounding points

This section collects elementary results about convex sets in finite dimensional real vector spaces that will help to construct families of loops. In this section, E is a real vector space with (finite) dimension d. The discussion will center around the following definition which is tailored to our ulterior needs.

Definition 1.3. *A point x in E is surrounded by points p_0, …, p_d if those points are affinely independent and there exist weights $w_i \in (0, 1)$ with sum 1 such that $x = \sum_i w_i p_i$.*

Note that, in the above definition, the number of points p_i is fixed by the dimension d of E, and that the weights w_i are the barycentric coordinates of x with respect to the affine basis $p_0, \dots, p_d$.

The first important point in this definition is that surrounding is smoothly locally stable: if x is surrounded by a collection of points p then points that are close to y are surrounded by every collection of points q that is closed to p, and the relevant barycentric coordinates smoothly depend on y and q. The precise statement follows.

Lemma 1.4. *For every x in E and every collection of points $p \in E^{d+1}$ surrounding x, there is a function $w\colon E \times E^{d+1} \to \mathbb{R}^{d+1}$ such that, for every (y, q) in a neighborhood of (x, p),*

- *w is smooth at (y, q)*
- *$w(y, q) > 0$*
- *$\sum_{i=0}^{d} w_i(y, q) = 1$*
- *$y = \sum_{i=0}^{d} w_i(y, q) q_i$.*

Proof. Let:

$$A = E \times \{q \in E^{d+1} \mid q \text{ is an affine basis for } E\},$$

and define:

$$\begin{aligned} w\colon & A \to \mathbb{R}^{d+1} \\ & (y, q) \mapsto \text{barycentric coordinates of } y \text{ with respect to } q. \end{aligned}$$

If we fix an affine basis b of E, we may express w as a ratio of determinants in terms of coordinates relative to b. More precisely, by Cramer's rule, if $0 \leq i \leq d$ and w_i is the i^{th} component of w, then:

$$w_i(y, q) = \det M_i(y, q) / \det N(q)$$

where $N(q)$ is the $(d+1) \times (d+1)$ matrix whose columns are the barycentric coordinates of the components of q relative to b, and $M_i(y,q)$ is $N(q)$ except with column i replaced by the barycentric coordinates of y relative to b.

Since determinants are smooth functions and $(y,q) \mapsto \det N(q)$ is non-vanishing on A, w is smooth on A.

Finally define:

$$U = w^{-1}((0,\infty)^{d+1}),$$

and note that U is open in A, since it is the preimage of an open set under the continuous map w. In fact since A is open, U is open as a subset of $E \times E^{d+1}$. Note that $(x,p) \in U$ since p surrounds x.

We may extend w to $E \times E^{d+1}$ by giving it any values at all outside A. □

Then we need a criterion ensuring a point x is surrounded by a collection of points taken in a given subset P. The first temptation is to hope that x being in the interior of the convex hull of P is enough. But this is not true. For instance the center of a square in a plane is in the interior of the convex hull of the set P of vertices of the square, but it isn't surrounded by any set of vertices. This counter example also shows that the stability lemma above is slightly less trivial than it sounds.

The rest of this section is devoted to the following result that proves no such issue arises when P is open.

Proposition 1.5. *If a point x of E lies in the convex hull of an open set P, then it is surrounded by some collection of points belonging to P.*

This proposition will be proven at the end of this section. We'll first need the Carathéodory lemma:

Lemma 1.6 (Carathéodory's lemma)**.** *If a point x of E lies in the convex hull of a set P, then x belongs to the convex hull of a finite set of affinely independent points of P.*

Proof. By assumption, there is a finite set of points t_i in P and weights f_i such that $x = \sum f_i t_i$, each f_i is non-negative and $\sum f_i = 1$. Choose such a set of points of minimum cardinality. We argue by contradiction that such a set must be affinely independent.

Thus suppose that there is some vanishing combination $\sum g_i t_i$ with $\sum g_i = 0$ and not all g_i vanish. Let $S = \{i | g_i > 0\}$. Let i_0 in S be an index minimizing f_i/g_i. We shall obtain our contradiction by showing that x belongs to the convex hull of the set $\{t_i | i \neq i_0\}$, which has cardinality strictly smaller than $\{t_i\}$.

We thus define new weights $k_i = f_i - g_i f_{i_0}/g_{i_0}$. These weights sum to $\sum f_i - (\sum g_i) f_{i_0}/g_{i_0} = 1$ and $k_{i_0} = 0$. Each k_i is non-negative, thanks to the choice of i_0 if i is in S or using that f_i, $-g_i$ and f_{i_0}/g_{i_0} are all non-negative when i is not in S. It remain to compute

$$\begin{aligned}
\sum_{i \neq i_0} k_i t_i &= \sum_i k_i t_i \\
&= \sum_i (f_i - g_i f_{i_0}/g_{i_0}) t_i \\
&= \sum_i f_i t_i - \left(\sum_i g_i t_i \right) f_{i_0}/g_{i_0}) \\
&= x
\end{aligned}$$

where we use $k_{i_0} = 0$ in the first equality. □

Lemma 1.7. *Given an affine basis b of E, the interior of the convex hull of b is the set of points with strictly positive barycentric coordinates.*

Proof. For each i, let:

$$w_i \colon E \to \mathbb{R}$$

be the i^{th} barycentric coordinate with respect to the basis b. Since E is finite-dimensional, each w_i is a continuous open map. For such a map, the operation of taking interior commutes with preimage, and so we have:

$$\begin{aligned}
\operatorname{IntConv}(b) &= \operatorname{Int}\left(\bigcap_i w_i^{-1}([0,\infty))\right) \\
&= \bigcap_i \operatorname{Int}(w_i^{-1}([0,\infty)) \\
&= \bigcap_i w_i^{-1}(\operatorname{Int}([0,\infty)) \\
&= \bigcap_i w_i^{-1}((0,\infty))
\end{aligned}$$

as required. □

Lemma 1.8. *Given a point c of E and a real number t, let:*

$$h_t^c \colon E \to E$$

be the homothety which dilates about c by a scale of t.

Suppose c belongs to the interior of a convex subset C of E and $t > 1$, then

$$C \subseteq \operatorname{Int}(h_t^c(C))$$

Proof. Since h_t^c is a homeomorphism with inverse $h_{t^{-1}}^c$, taking $s = t^{-1}$, the required result is equivalent to showing:

$$h_s^c(C) \subseteq \operatorname{Int}(C)$$

where $s \in (0,1)$.

Let x be a point of C, we must show there exists an open neighborhood U of $h_s^c(x)$, contained in C. In fact we claim:

$$U = h_{1-s}^x(\operatorname{Int}(C))$$

is such a set. Indeed U is open since h_{1-s}^x is a homeomorphism and U contains $h_s^c(x)$ since:

$$h_s^c(x) = h_{1-s}^x(c) \in h_{1-s}^x(\operatorname{Int}(C))$$

since c belongs to $\operatorname{Int}(C)$. Finally:

$$h_{1-s}^x(\operatorname{Int}(C)) \subseteq h_{1-s}^x(C) \subseteq C$$

where the second inclusion follows since C is convex and contains x. □

We are now ready to come back to Proposition 1.5.

Proof of Proposition 1.5. It follows from Lemma 1.7 that we need only show that E has an affine basis b of points belonging to P such that x lies in the interior of the convex hull of b.

Carathéodory's lemma 1.6 provides affinely independent points $p_0, \dots, p_k$ in P such that x belongs to their convex hull. Since P is open, we may extend p_i to an affine basis

$$\hat{b} = \{p_0, \dots, p_d\},$$

where all points still belong to P. Note that x belongs to the convex hull of $\hat{b}$.

Now let c be a point in the interior of the convex hull of $\hat{b}$ (e.g., the centroid) and for each $\epsilon > 0$, consider the homothety

$$h_{1+\epsilon} \colon E \to E,$$

which dilates about c by a scale of $1 + \epsilon$.

Since $\hat{b}$ is finite and contained in P, and P is open, there exists $\epsilon > 0$ such that

$$h_{1+\epsilon}(\hat{b}) \subseteq P.$$

We claim the required basis is:

$$b = h_{1+\epsilon}(\hat{b})$$

for any such ϵ. Indeed, applying Lemma 1.8 to $\mathrm{Conv}(\hat{b})$ we see:

$$\begin{aligned} x \in \mathrm{Conv}(\hat{b}) &\subseteq \mathrm{Int}(h_{1+\epsilon}(\mathrm{Conv}(\hat{b}))) \\ &= \mathrm{Int}(\mathrm{Conv}(h_{1+\epsilon}(\hat{b}))) \end{aligned}$$

as required. □

1.3 Constructing loops

1.3.1 Surrounding families

It will be convenient to introduce some more vocabulary.

Definition 1.9. *We say a loop γ surrounds a vector v if v is surrounded by a collection of points belonging to the image of γ. Also, we fix a base point 0 in $\mathbb{S}^1$ and say a loop is based at some point b if 0 is sent to b.*

The first main task in proving Proposition 1.2 is to construct suitable families of loops γ_x surrounding $g(x)$, by assembling local families of loops. Those will then be reparametrized to get the correct average in the next section. In this section, we will work only with *continuous* loops. This will make constructions easier and we will smooth those loops in the end, taking advantage of the fact that Ω and the surrounding condition are open.

Thanks to Carathéodory's lemma, constructing *one* such loop with values in some open O is easy as soon as v belongs to the convex hull of O.

Lemma 1.10. *If a vector v is in the convex hull of a connected open subset O then, for every base point $b \in O$, there is a continuous family of loops $\gamma \colon [0,1] \times \mathbb{S}^1 \to E$, $(t, s) \mapsto \gamma^t(s)$ such that, for all t and s:*

- *γ^t is based at b*

- $\gamma^0(s) = b$
- $\gamma^t(s) \in O$
- γ^1 *surrounds* v

Proof. Since O is open, Proposition 1.5 gives points p_i in O surrounding x. Since O is open and connected, it is path connected. Let $\lambda\colon [0,1] \to \Omega_x$ be a continuous path starting at b and going through the points p_i. We can concatenate λ and its opposite to get γ^1. This is a round-trip loop: it back-tracks when it reaches $\lambda(1)$ at $s = 1/2$. We then define γ^t as the round-trip that stops at $s = t/2$, stays still until $s = 1 - t/2$ and then backtracks. $\square$

Definition 1.11. *A continuous family of loops* $\gamma : E \times [0,1] \times \mathbb{S}^1 \to F, (x,t,s) \mapsto \gamma_x^t(s)$ *surrounds a map* $g\colon E \to F$ *with base* $\beta\colon E \to F$ *on* $U \subset E$ *in* $\Omega \subset E \times F$ *if, for every* x *in* U*, every* $t \in [0,1]$ *and every* $s \in \mathbb{S}^1$*,*

- γ_x^t *is based at* $\beta(x)$
- $\gamma_x^0(s) = \beta(x)$
- γ_x^1 *surrounds* $g(x)$
- $(x, \gamma_x^t(s)) \in \Omega.$

The space of such families will be denoted by $\mathcal{L}(g, \beta, U, \Omega)$*.*

Families of surrounding loops are easy to construct locally.

Lemma 1.12. *Assume* Ω *is open over some neighborhood of* x_0*. If* $g(x_0)$ *is in the convex hull of the connected component of* Ω_{x_0} *containing* $\beta(x_0)$*, then there is a continuous family of loops defined near* x_0*, based at* β*, taking value in* Ω *and surrounding* g*.*

Proof. In this proof we don't mention the t parameter since it plays no role, but it is still there. Lemma 1.10 gives a loop γ based at $\beta(x_0)$, taking values in Ω_{x_0} and surrounding $g(x_0)$. We set $\gamma_x(s) = \beta(x) + (\gamma(s) - \beta(x_0))$. Each γ_x takes values in Ω_x because Ω is open over some neighborhood of x_0. Lemma 1.4 guarantees that this loop surrounds $g(x)$ for x close enough to x_0. $\square$

The difficulty in constructing global families of surrounding loops is that there are plenty of surrounding loops and we need to choose them consistently. The key feature of the above definition is that the t parameter not only allows us to cut out the corrugation process in the next chapter, but also brings a "satisfied or refund" guarantee, as explained in the next lemma.

Lemma 1.13. *For every set* $U \subset E$*,* $\mathcal{L}(g, \beta, U, \Omega)$ *is "path connected": for every* γ_0 *and* γ_1 *in* $\mathcal{L}(g, \beta, U, \Omega)$*, there is a continuous map* $\delta\colon [0,1] \times E \times [0,1] \times \mathbb{S}^1 \to F, (\tau, x, t, s) \mapsto \delta_{\tau,x}^t(s)$ *which interpolates between* γ_0 *and* γ_1 *in* $\mathcal{L}(g, \beta, U, \Omega)$*.*

The construction below morally proves that each $\mathcal{L}(g, \beta, U, \Omega)$ is contractible, but we will not even specify a topology on those spaces. The definition of "path connected" in quotation marks is the above specific statement, and only this statement will be used.

Proof. Let ρ be the piecewise affine map from $\mathbb{R}$ to $\mathbb{R}$ such that $\rho(\tau) = 1$ if $\tau \le 1/2$, ρ is affine on $[1/2, 1]$, $\rho(\tau) = 0$ if $\tau \ge 1$. We set

$$\delta^t_{\tau,x}(s) = \begin{cases} \gamma^{\rho(\tau)t}_{0,x}\left(\frac{1}{1-\tau}s\right) & \text{if } s \le 1-\tau \text{ and } \tau < 1 \\ \gamma^{\rho(1-\tau)t}_{1,x}\left(\frac{1}{\tau}(s-(1-\tau))\right) & \text{if } s \ge 1-\tau \text{ and } \tau > 0 \end{cases}$$

It is clear that if $s = 1-\tau$ then both branches agree and are equal to $\beta(x)$. Therefore it is easy to see that δ is continuous at (τ, x, t, s) except when $(\tau, s) = (1, 0)$ or $(\tau, s) = (0, 1)$.

To show the continuity for $(\tau, s) = (1, 0)$, let K be a compact neighborhood of x in E. Then γ_0 is uniformly continuous on the compact set $K \times [0, 1] \times \mathbb{S}^1$, which means that $\gamma^t_{0,x'}$ tends uniformly to the constant function $s \mapsto \beta(x)$ as (x', t) tends to $(x, 0)$. This means that $\gamma^{\rho(\tau)t'}_{0,x'}$ tends uniformly to the constant function $s \mapsto \beta(x)$ as (τ, x', t') tends to $(1, x, t)$. This means that δ is continuous at $(\tau, s) = (1, 0)$ (it is clear that the other branch also tends to $\beta(x)$). The continuity at $(\tau, s) = (0, 1)$ is entirely analogous.

The beautiful observation motivating the above formula is why each $\delta^1_{\tau,x}$ surrounds $g(x)$. The key is that the image of $\delta^1_{\tau,x}$ contains the image of $\gamma^1_{0,x}$ when $\tau \le 1/2$, and contains the image of $\gamma^1_{1,x}$ when $\tau \ge 1/2$. Hence $\delta^1_{\tau,x}$ always surrounds $g(x)$. □

Corollary 1.14. *Let U_0 and U_1 be open sets in E. Let $K_0 \subset U_0$ and $K_1 \subset U_1$ be compact subsets. For any $\gamma_0 \in \mathcal{L}(U_0, g, \beta, \Omega)$ and $\gamma_1 \in \mathcal{L}(U_1, g, \beta, \Omega)$, there exists a neighborhood U of $K_0 \cup K_1$ and there exists $\gamma \in \mathcal{L}(U, g, \beta, \Omega)$ which coincides with γ_0 near $K_0 \cup U_1^c$.*

Proof. Let $C_0 = K_0 \cup U_1^c$ and $C_1 := K_1 \setminus U_0$. Since C_0 and C_1 are disjoint closed sets, there is some continuous cut-off $\rho\colon E \to [0, 1]$ which vanishes on a neighborhood of C_0 and equals one on a neighborhood of C_1.

Lemma 1.13 gives a homotopy of loops γ_τ from γ_0 to γ_1 on $U_0 \cap U_1$. Moreover, note that γ_τ is defined on all of E. On $U_0' \cup (U_0 \cap U_1) \cup U_1'$, which is a neighborhood of $K_0 \cup K_1$, we set

$$\gamma_x = \gamma_{\rho(x),x}$$

which has the required properties. □

Lemma 1.15. *In the setup of Proposition 1.2, assume we have a continuous family γ of loops defined near K which is based at β, surrounds g and such that each γ^t_x takes values in Ω_x. Then there such a family which is defined on all of E and agrees with γ near K.*

Proof. Let U_0 be an open set containing K such that γ forms a surrounding family of loops on U_0. Let U_i, $i \ge 1$ be a locally finite family of open sets with local surrounding families of loops γ^i and compact subsets $K_i \subset U_i$ covering E.

This is possible, since by Lemma 1.12 there is a local surrounding family of loops around each point in E. Since E is locally compact we may pick a compact neighborhood around each point in E with such a local family of loops. Since E is paracompact second countable, we can pick a countable refinement U_i which is locally finite and still covers E. By the shrinking lemma we can pick closed (and hence compact) subsets $K_i \subset U_i$ that also cover E.

Now we define a family $(\delta^i)_{i\in\mathbb{N}}$ by $\delta^0 = \gamma|_{U_0}$ and δ^{i+1} is obtained by extending δ^i using γ^i via Corollary 1.14. Since δ^{i+1} equals δ^i on U_i^c, this sequence is locally eventually constant. Therefore it has a well-defined limit δ which is defined and continuous on all E. Since δ^{i+1} equals δ^i on a neighborhood of $K \cup \bigcup_{j<i} K_i$, we know in particular that δ^i equals γ on a neighborhood of K. Since K is compact, it has some neighborhood O such that $\delta^i|_O$ is eventually constant with eventual value $\delta|_O$. Hence $\delta = \gamma$ on a neighborhood of K. □

1.3.2 The reparametrization lemma

The second ingredient needed to prove Proposition 1.2 is a parametric reparametrization lemma. Gromov's original proof of this lemma makes explicit use of a partition of unity. Motivated in particular by formalization purposes, we will first state more abstract versions whose statements do not involve any partition of unity but directly state a local-to-global property.

Lemma 1.16. *Let E and F be real normed vector spaces. Assume that E is finite dimensional. Let P be a predicate on $E \times F$ such that for all x in E, $\{y \mid P(x, y)\}$ is convex. Let n be a natural number or $+\infty$. Assume that every x has a neighbourhood U on which there exists a C^n function f such that $\forall x \in U, P(x, f(x))$. Then there is a global C^n function f such that $\forall x, P(x, f(x))$.*

Proof. The assumption give us an open cover $(U_i)_{i \in I}$ of E and functions $f_i : E \to F$ that are smooth on U_i and such that $P(x, f_i(x))$ for all x in U_i. Let ρ be a smooth partition of unity associated to this cover. The function $f = \sum \rho_i f_i$ is smooth on E and the convexity assumption on P ensures it satisfies $\forall x, P(x, f(x))$. Indeed each value $f(x)$ is a convex combination of finitely many values $f_i(x)$ where i satisfies that x is in U_i. □

We will also need a version where F is a space of smooth functions. Since there is no relevant norm to put on such a space, we cannot deduce this version from the above one.

Lemma 1.17. *Let E_1, E_2 and F be real vector spaces. Assume E_1 and E_2 are finite dimensional. Let n be a natural number or $+\infty$. Let P be a property of pairs (x, f) with $x \in E_1$ and $f : E_2 \to F$. Assume that, for every x, the space of functions f such that $P(x, f)$ holds is convex. Assume that for every x_0 in E_1 there is a neighborhood U of x_0 and a function $\varphi : E_1 \times E_2 \to F$ which is C^n on $U \times E_2$ and such that $P(x, \varphi(x, \cdot))$ holds for every x in U. There there is a global C^n function $\varphi\colon E_1 \times E_2 \to F$ such that $P(x, \varphi(x, \cdot))$ holds for every x.*

Proof. This is completely analogous to the previous proof. □

Lemma 1.18. *Let $\gamma\colon E \times \mathbb{S}^1 \to F$ be a smooth family of loops surrounding a map g. There is a smooth family $\varphi\colon E \times \mathbb{S}^1 \to \mathbb{S}^1$ such that each $\gamma_x \circ \varphi_x$ has average $g(x)$ and $\varphi_x(0) = 0$.*

Proof. Gromov's main idea in order to prove this result is to translate the problem of constructing a family of circle maps φ into the problem of constructing a family of smooth density functions f on the circle. We introduce some vocabulary in order to describe this reduction. Let $f\colon E \times \mathbb{R} \to \mathbb{R}$ be a smooth positive function that is 1-periodic in its second argument. We say that f is a centering density for (γ, g) at x if $f_x\colon \mathbb{R} \to \mathbb{R}$ has average value one when seen as a function on $\mathbb{S}^1$ and the average value of $f_x\gamma_x$ is $g(x)$. We claim that, in order to prove the lemma, it is sufficient to build such an f which is centering at every x. Indeed, assume we have such an f. We then get a smooth family of $\mathbb{Z}$-equivariant functions $\psi\colon E \times \mathbb{R} \to \mathbb{R}$ defined by $\psi_x(t) = \int_0^t f_x(s)ds$. Because ψ is smooth and each ψ_x is strictly monotone and $\mathbb{Z}$-equivariant, one can check there is a smooth map $\varphi : E \times \mathbb{R} \to \mathbb{R}$ which is $\mathbb{Z}$-equivariant and such that $\varphi_x \circ \psi_x = \mathrm{Id}$ for each x. Seen as a family of functions from $\mathbb{S}^1$ to $\mathbb{S}^1$, those functions are suitable since, for every x, the change of variable formula gives:

$$\int_{\mathbb{S}^1} \gamma_x \circ \varphi_x(s)ds = \int_{\mathbb{S}^1} \psi_x'(s)\gamma_x \circ \varphi_x(\psi_x(s))ds = \int_{\mathbb{S}^1} f_x(s)\gamma_x(s)ds = g(x).$$

We now prove the existence of a function which is a centering density at every point of x. For any given x, this constraint is clearly convex. Hence Lemma 1.17 ensures it is enough to prove existence of functions that are centering densities in a neighborhood of any given point x. So we fix some x in E.

Since γ_x strictly surrounds $g(x)$, there are points $s_1, ..., s_{n+1}$ in $\mathbb{S}^1$ such that $g(x)$ is surrounded by the corresponding points $\gamma_x(s_j)$.

Let $f_1, ..., f_{n+1}$ be smooth positive periodic maps from $\mathbb{R}$ to $\mathbb{R}$ which average value 1 on a period and such that the corresponding measures on $\mathbb{S}^1$ are very close to the Dirac measures on s_j, ie. for any function h, the average value of $f_j h$ is almost $h(s_j)$. We set $p_j = \int f_j \gamma_x \, ds$, which is almost $\gamma_x(s_j)$ so that $g(x) = \sum w_j p_j$ for some weights w_j in the open interval $(0, 1)$ according to Lemma 1.4.

If x' is in a sufficiently small neighborhood of x, Lemma 1.4 gives smooth weight functions w_j such that $g(x') = \sum w_j(x') p_j(x')$. Hence we can set $f_{x'}(s) = \sum w_j(x') f_j(s)$. □

1.3.3 Proof of the loop construction proposition

We finally assemble the ingredients from the previous two sections.

Proof of Proposition 1.2. Let γ^* be a family of loops surrounding the origin in $B_F(0, 1)$ the open unit ball in F, constructed using Lemma 1.12. For x in some neighborhood U^* of K where $g = \beta$, we set $\gamma_x = g(x) + \varepsilon\gamma^*$ where $\varepsilon > 0$ is sufficiently small to ensure that $B_{E\times F}((x, \beta(x)), 2\varepsilon) \subseteq \Omega$ (recall Ω is open and K is compact). Lemma 1.15 extends this family to a continuous family of surrounding loops γ_x for all x (this is not yet our final γ).

We then need to approximate this continuous family by a smooth one. Some care is needed to ensure that it stays based at β. We can first reparametrize γ on $[0, 1] \times \mathbb{S}^1$ to ensure that γ is constant in a neighborhood of $C = \{(t, s) \in [0, 1] \times \mathbb{S}^1 \mid t = 0 \text{ or } s = 0\}$. Using Lemma 1.16, we can find a smooth function that has distance at most ε from γ and coincides with γ on C (using the fact that γ is already smooth near C). Since all loops that are sufficiently close to γ still surround g, we can also ensure that the new smoothened γ is still surrounding.

Then Lemma 1.18 gives a family of circle diffeomorphisms h_x such that $\gamma^1_x \circ h_x$ has average $g(x)$.

Finally we choose a cut-off function function χ which vanishes near $E \setminus U^*$ and equals one near K. As our final family of loops, we choose $\chi(x)g(x) + (1 - \chi(x))(\gamma_x \circ h_x)$. This operation does not change the average values of these loops, because it rescales them around their average value, but makes them constant near K. Also, those loops stay in Ω, thanks to our choice of ε. □

Chapter 2

Local theory of convex integration

2.1 Key construction

The goal of this chapter is to explain the local aspects of (Theillière's implementation of) convex integration, the next chapter will cover global aspects.

The elementary step of convex integration modifies the derivative of a map in one direction. The precise meaning of "one direction" rely on the following definition.

Definition 2.1. *A dual pair on a vector space E is a pair (π, v) where π is a linear form on E and v a vector in E such that $\pi(v) = 1$.*

Let E and F be finite dimensional real normed vector spaces. Let $f\colon E \to F$ be a smooth map, and let (π, v) be a dual pair on E. We want to modify Df in the direction of v while almost preserving it on $\ker \pi$. Say we wish $Df(x)v$ could live in some open subset $\Omega_x \subset F$. Assume there is a smooth family of loops $\gamma\colon E \times \mathbb{S}^1 \to F$ such that each γ_x takes values in Ω_x, and its average value $\overline{\gamma}_x = \int_{\mathbb{S}^1} \gamma_x$ is $Df(x)v$ for all x. Obviously such loops can exist only if $Df(x)v$ is in the convex hull of Ω_x, and we saw in the previous chapter that this is almost sufficient (and we'll see this is sufficiently almost sufficient for our purposes). Then we can modify f to fulfil our wish using the following construction.

Definition 2.2 (Theillière 2018)**.** *The map obtained by corrugation of f in direction (π, v) using γ with oscillation number N is*

$$x \mapsto f(x) + \frac{1}{N} \int_0^{N\pi(x)} \left[\gamma_x(s) - \overline{\gamma}_x\right] ds.$$

In the above definition, we mostly think of N as a large natural number. But we don't actually require it, any positive real number will do.

The next proposition implies that, provided N is large enough, we have achieved $Df'(x)v \in \Omega_x$, almost without modifying derivatives in the directions of $\ker \pi$, and almost without moving $f(x)$.

Proposition 2.3 (Theillière 2018)**.** *Let f be a $\mathcal{C}^1$ function from E to F. Let (π, v) be a dual pair on E. Let $\gamma\colon E \times \mathbb{S}^1 \to F$ be a $\mathcal{C}^1$ family of loops such that $\overline{\gamma_x} = Df(x)v$ for all x.*

For any compact set $K \subset E$ and any positive ε, the function f' obtained by corrugation of f in direction (π, v) using γ with large enough oscillation number N satisfies:

1. $\forall x \in K, \|f'(x) - f(x)\| \leq \varepsilon$

2. $\forall x \in K, \|(Df'(x) - Df(x))_{|\ker \pi}\| \leq \varepsilon$.

3. $\forall x \in K, \|Df'(x)v - \gamma(x, N\pi(x))\| \leq \varepsilon$

In addition, all the differences estimated above vanish if x is outside the support of γ.

Proof. We set $\Gamma_x(t) = \int_0^t \left(\gamma_x(s) - \overline{\gamma}_x\right) ds$, so that $f'(x) = f(x) + \Gamma_x(N\pi(x))/N$. Because each Γ_x is 1-periodic, and everything has compact support in E, all derivatives of Γ are uniformly bounded. Item 1 in the statement is then obvious. Item 2 also follows since $\partial_i f'(x) = \partial_i f(x) + \partial_i \Gamma(x, N\pi(x))/N$. In order to prove Item 3, we compute:

$$\begin{aligned} Df'(x)v &= Df(x)v + \frac{1}{N}\partial_j \Gamma(x, N\pi(x)) + \frac{N}{N}\partial_t \Gamma(x, N\pi(x)) \\ &= Df(x)v + O\left(\frac{1}{N}\right) + \gamma(x, N\pi(x)) - Df(x)v \\ &= \gamma(x, N\pi(x)) + O\left(\frac{1}{N}\right). \end{aligned}$$

Outside the support of γ, Γ_x and its derivative with respect to x vanish identically (for the derivative computation, it is important that the support of γ is the *closure* of the set of x where γ_x is not constant). □

2.2 The main inductive step

Definition 2.4. *Let E' be a linear subspace of E. A map $\mathcal{F} = (f, \varphi) : E \to F \times \mathrm{Hom}(E, F)$ is E'–holonomic if, for every v in E' and every x, $Df(x)v = \varphi(x)v$.*

Definition 2.5. *A first order differential relation for maps from E to F is a subset $\mathcal{R}$ of $E \times F \times \mathrm{Hom}(E, F)$.*

Until the end of this section, $\mathcal{R}$ will always denote a first order differential relation for maps from E to F.

Definition 2.6. *A formal solution of a differential relation $\mathcal{R}$ is a map $\mathcal{F} = (f, \varphi) \colon E \to F \times \mathrm{Hom}(E, F)$ such that, for every x, $(x, f(x), \varphi(x))$ is in $\mathcal{R}$.*

The first component of a map $\mathcal{F} : E \to F \times \mathrm{Hom}(E, F)$ will sometimes be denoted by $\mathrm{bs}\,\mathcal{F} : E \to F$ and called the base map of $\mathcal{F}$.

Definition 2.7. *A* 1*-jet section from E to F is a function from E to $F \times \mathrm{Hom}(E, F)$. A homotopy of* 1*-jet sections is a smooth map $\mathcal{F} : \mathbb{R} \times E \to F \times \mathrm{Hom}(E, F)$.*

Typically, $x \mapsto \mathcal{F}(t, x)$ will be denoted by $\mathcal{F}_t$. It could seem more natural to take $[0, 1] \times E$ as the source of a homotopy but this would be less convenient for formalization and wouldn't change anything since any map from $\mathbb{R} \times E$ can be restricted to $[0, 1] \times E$ and every map from $[0, 1] \times E$ could be extended.

Definition 2.8. *For every $\sigma = (x, y, \varphi)$, the slice of $\mathcal{R}$ at σ with respect to (π, v) is:*

$$\mathcal{R}(\sigma, \pi, v) = \{w \in F \mid (x, y, \varphi + (w - \varphi(v)) \otimes \pi) \in \mathcal{R}\}.$$

Lemma 2.9. *The linear map $\varphi + (w - \varphi(v)) \otimes \pi)$ coincides with φ on $\ker \pi$ and sends v to w. If σ belongs to $\mathcal{R}$ then $\varphi(v)$ belongs to $\{w \in F, (x, y, \varphi + (w - \varphi(v)) \otimes \pi) \in \mathcal{R}\}$.*

Proof. These are direct checks. □

We'll use the notation $\mathrm{Conn}_w A$ to denote the connected component of A that contains w, or the empty set if w doesn't belong to A.

Definition 2.10. *A formal solution $\mathcal{F}$ of $\mathcal{R}$ is (π, v)–short if, for every x, $Df(x)v$ belongs to the interior of the convex hull of $\mathrm{Conn}_{\varphi(v)} \mathcal{R}((x, f(x), \varphi(x)), \pi, v)$.*

Lemma 2.11. *Let $\mathcal{F}$ be a formal solution of $\mathcal{R}$. Let $K_1 \subset E$ be a compact subset, and let K_0 be a compact subset of the interior of K_1. Let C be a closed subset of E. Let E' be a linear subspace of E contained in $\ker \pi$. Let ε be a positive real number.*

Assume $\mathcal{R}$ is open. Assume that $\mathcal{F}$ is E'–holonomic near K_0, (π, v)–short, and holonomic near C. Then there is a homotopy $\mathcal{F}_t$ such that:

1. *$\mathcal{F}_0 = \mathcal{F}$;*
2. *$\mathcal{F}_t$ is a formal solution of $\mathcal{R}$ for all t ;*
3. *$\mathcal{F}_t(x) = \mathcal{F}(x)$ for all t when x is near C or outside K_1 ;*
4. *$d(\mathrm{bs}\, \mathcal{F}_t(x), \mathrm{bs}\, \mathcal{F}(x)) \le \varepsilon$ for all t and all x ;*
5. *$\mathcal{F}_1$ is $E' \oplus \mathbb{R}v$–holonomic near K_0.*

Proof. We denote the components of $\mathcal{F}$ by f and φ. Since $\mathcal{F}$ is short, Proposition 1.2 applied to $g\colon x \mapsto Df(x)v$, $\beta\colon x \mapsto \varphi(x)v$, $\Omega_x = \mathcal{R}(\mathcal{F}(x), \pi, v)$, and $K = C \cap K_1$ gives us a smooth family of loops $\gamma\colon E \times [0,1] \times \mathbb{S}^1 \to F$ such that, for all x:

- $\forall t\, s,\ \gamma_x^t(s) \in \mathcal{R}(\mathcal{F}(x), \pi, v)$
- $\forall s,\ \gamma_x^0(s) = \varphi(x)v$
- $\bar{\gamma}_x^1 = Df(x)v$
- if x is near C, $\forall t\, s,\ \gamma_x^t(s) = \varphi(x)v$

Let $\rho : E \to \mathbb{R}$ be a smooth cut-off function which equals one on a neighborhood of K_0 and whose support is contained in K_1.

Let N be a positive real number. Let $\bar{f}$ be the corrugated map constructed from f, γ^1 and N. Proposition 2.3 ensures that, for all x,

$$D\bar{f}(x) = Df(x) + \left[\gamma_x^1(N\pi(x)) - Df(x)v\right] \otimes \pi + R_x$$

for some remainder term R which is ε-small and vanishes whenever γ_x is constant, hence vanishes near C.

We set $\mathcal{F}_t(x) = (f_t(x), \varphi_t(x))$ where:

$$f_t(x) = f(x) + \frac{t\rho(x)}{N} \int_0^{N\pi(x)} \left[\gamma_x^t(s) - Df(x)v\right] ds$$

and

$$\varphi_t(x) = \varphi(x) + \left[\gamma_x^{t\rho(x)}(N\pi(x)) - \varphi(x)v\right] \otimes \pi + \frac{t\rho(x)}{N} B_x.$$

We now prove that $\mathcal{F}_t$ has the announced properties, starting with he obvious ones. The fact that $\mathcal{F}_0 = \mathcal{F}$ is obvious since $\gamma_x^0(s) = \varphi(x)v$ for all s.

When x is near C, $Df(x) = \varphi(x)$ since $\mathcal{F}$ is holonomic near C. In addition, $\gamma_x^t(s) = \varphi(x)v$ for all s and t, hence B_x vanishes. Hence $\mathcal{F}_t(x) = \mathcal{F}(x)$ for all t when x is near C.

Outside of K_1, ρ vanishes. Hence $f_t(x) = f(x)$ for all t, and $\gamma_x^{t\rho(x)}(s) = \varphi(x)v$ for all s and t, and $\varphi_t(x) = \varphi(x)$.

The distance between $f(x)$ and $f_t(x)$ is zero outside of K_1 and ε-small everywhere.

We now turn to the interesting parts. The first one is that each $\mathcal{F}_t$ is a formal solution of $\mathcal{R}$. We already now that $\mathcal{F}_t$ coincides with $\mathcal{F}$, which is a formal solution, outside of the compact set K_1. We set

$$\mathcal{F}'_t(x) = \left(f(x), \varphi(x) + \left[\gamma_x^{t\rho(x)}(N\pi(x)) - \varphi(x)v\right] \otimes \pi\right).$$

Since $\mathcal{R}$ is open, and $K_1 \times [0,1]$ is compact and $\mathcal{F}_t$ is within $O(1/N)$ of $\mathcal{F}'_t$, it suffices to prove that $\mathcal{F}'_t$ is a formal solution for all t. This is guaranteed by the definition of the slice $\mathcal{R}(\mathcal{F}(x), \pi, v)$ to which $\gamma_x^{t\rho(x)}(N\pi(x))$ belongs.

Finally, let's prove that $\mathcal{F}_1$ is $E' \oplus \mathbb{R}v$–holonomic near K_0. Since $\rho = 1$ near K_0, we have, for x near K_0,

$$Df_1(x) = Df(x) + \left[\gamma_x^1(N\pi(x)) - Df(x)v\right] \otimes \pi + \frac{1}{N}B_x,$$

and

$$\varphi_1(x) = \varphi(x) + \left[\gamma_x^1(N\pi(x)) - \varphi(x)v\right] \otimes \pi + \frac{1}{N}B_x.$$

Let p be the projection of E onto $\ker \pi$ along v, so that $\mathrm{Id}_E = p + v \otimes \pi$. We can rewrite the above formulas as

$$Df_1(x) = Df(x) \circ p + \gamma_x^1(N\pi(x)) \otimes \pi + \frac{1}{N}B_x,$$

and

$$\varphi_1(x) = \varphi(x) \circ p + \gamma_x^1(N\pi(x)) \otimes \pi + \frac{1}{N}B_x.$$

So we see the difference is $Df(x) \circ p - \varphi(x) \circ p$ which vanishes on E' since $\mathcal{F}$ is E'–holonomic near K_0, and vanishes on v since $p(v) = 0$. $\square$

2.3 Ample differential relations

Definition 2.12. *A subset Ω of a real vector space E is ample if the convex hull of each connected component of Ω is the whole E.*

Lemma 2.13. *The complement of a linear subspace of codimension at least 2 is ample.*

Proof. Let F be subspace of E with codimension at least 2. Let F' be a complement subspace. Its dimension is at least 2 since it is isomorphic to E/F and $\dim(E/F) = \mathrm{codim}(F) \geq 2$. First note the complement of F is path-connected. Indeed let x and y be points outside F. Decomposing on $F \oplus F'$, we get $x = u + u'$ and $y = v + v'$ with $u' \neq 0$ and $v' \neq 0$. The segments from x to u' and y to v' stay outside F, so it suffices to connect u' and v' in $F' \setminus \{0\}$. If the segment from u' to v' doesn't contains the origin then we are done. Otherwise $v' = \mu u'$ for some (negative) u'. Since $\dim(F') \geq 2$ and $u' \neq 0$, there exists

$f \in F'$ which is linearly independent from u', hence from v'. We can then connect both u' and v' to f by a segment away from zero.

We now turn to ampleness. The connectedness result reduces to prove that every e in E is in the convex hull of $E \setminus F$. If e is not in F then it is the convex combination of itself with coefficient 1 and we are done. Now assume e is in F. The codimension assumption guarantees the existence of a subspace G such that $\dim(G) = 2$ and $G \cap F = \{0\}$. Let (g_1, g_2) be a basis of G. We set $p_1 = e + g_1$, $p_2 = e + g_2$, $p_3 = e - g_1 - g_2$. All these points are in $E \setminus F$ and $e = p_1/3 + p_2/3 + p_3/3$. □

Definition 2.14. *A first order differential relation $\mathcal{R}$ is ample if all its slices are ample.*

Lemma 2.15. *Let $\mathcal{F}$ be a formal solution of $\mathcal{R}$. Let $K_1 \subset E$ be a compact subset, and let K_0 be a compact subset of the interior of K_1. Assume $\mathcal{F}$ is holonomic near a closed subset C of E. Let ε be a positive real number.*

If $\mathcal{R}$ is open and ample then there is a homotopy $\mathcal{F}_t$ such that:

1. $\mathcal{F}_0 = \mathcal{F}$
2. *$\mathcal{F}_t$ is a formal solution of $\mathcal{R}$ for all t ;*
3. *$\mathcal{F}_t(x) = \mathcal{F}(x)$ for all t when x is near C or outside K_1.*
4. *$d(\operatorname{bs}\mathcal{F}_t(x), \operatorname{bs}\mathcal{F}(x)) \le \varepsilon$ for all t and all x ;*
5. *$\mathcal{F}_1$ is holonomic near K_0 ;*
6. *$t \mapsto F_t$ is constant near 0 and 1.*

Proof. This is a straightforward induction using Lemma 2.11. Let $(e_1, \dots, e_n)$ be a basis of E, and let $(\pi_1, \dots, \pi_n)$ be the dual basis. Let E'_i be the linear subspace of E spanned by $(e_1, \dots, e_i)$, for $1 \le i \le n$, and let E'_0 be the zero subspace of E. Each (π_i, e_i) is a dual pair and the kernel of π_i contains E'_{i-1}.

Lemma 2.11 allows to build a sequence of homotopies of formal solutions, each homotopy relating a formal solution which is E'_i–holonomic to one which is E'_{i+1}–holonomic (always near K_0). The shortness condition is always satisfies because $\mathcal{R}$ is ample. Each homotopy starts where the previous one stopped, stay at C^0 distance at most ε/n, and is relative to C and the complement of K_1.

It then suffices to do a smooth concatenation of theses homotopies. We first pre-compose with a smooth map from $[0, 1]$ to itself that fixes 0 and 1 and has vanishing derivative to all orders at 0 and 1. Then we precompose by affine isomorphisms from $[0, 1]$ to $[i/n, (i + 1)/n]$ before joining them. □

For the purposes of the next chapter, we will need a slight variant of the above result where instead of improving a single formal solution we improve a homotopy of formal solutions.

Lemma 2.16. *Let $\mathcal{F}$ be a homotopy of formal solutions of $\mathcal{R}$. Let $K_1 \subset E$ be a compact subset, and let K_0 be a compact subset of the interior of K_1. Assume that near a closed set A in E, each $\mathcal{F}_t$ is holonomic and equal to $\mathcal{F}_0$, and $\mathcal{F}_1$ is holonomic near a closed subset C of E. Also assume that $t \mapsto F_t$ is constant near 0 and 1. Let ε be a positive real number.*

If $\mathcal{R}$ is open and ample then there is a homotopy $\mathcal{F}'$ of formal solutions of $\mathcal{R}$ such that:

1. $\mathcal{F}'_0 = \mathcal{F}_0$

2. $\mathcal{F}'_t(x) = \mathcal{F}_0(x)$ *for all* t *when* x *is near* A.
3. $\mathcal{F}'_t(x) = \mathcal{F}_t(x)$ *for all* t *when* x *is near* C *or outside* K_1.
4. *for all* t *and* x, *either* $\mathcal{F}'_t(x) = \mathcal{F}_{t'}(x)$ *for some* t' *or* $\| \operatorname{bs} \mathcal{F}'_t(x) - \operatorname{bs} \mathcal{F}_1(x) \| < \varepsilon$;
5. $\mathcal{F}'_1$ *is holonomic near* $K_0 \cup C \cup A$;
6. $t \mapsto F_t$ *is constant near* 0 *and* 1.

Proof. We apply Lemma 2.15 to $\mathcal{F}_1$ with closed subset $C \cup A$. This way we get a homotopy $t \mapsto \mathcal{G}_t$ starting at $\mathcal{G}_0 = \mathcal{F}_1$, relative to $C \cup A$ and the complement of K_1 and such that $\mathcal{G}_1$ is holonomic on K_0. Let $\rho \colon E \to [0,1]$ be a smooth cut-off function which equals two near K_0 and equals one outside K_1. Let $\chi \colon [0,1] \to [0,1]$ be a smooth function which vanishes near 0, equals 1 near 1 and coincides with the identity map away from 0 and 1.

We set

$$\mathcal{F}'_t(x) = \begin{cases} \mathcal{F}_{2\chi(t)/(2-\rho(x))}(x) & \text{if } \chi(t)\rho(x) \leq 1 \\ \mathcal{G}_{2\chi(t)/(2-\rho(x))-1}(x) & \text{if } \chi(t)\rho(x) \geq 1 \end{cases}$$

Near K_0, where ρ equals two, this $\mathcal{F}'$ is simply the concatenation of $\mathcal{F}$ and $\mathcal{G}$. Outside K_1 where ρ equals one, it is simply $\mathcal{F}$. $\square$

Chapter 3

Global theory of open and ample relations

3.1 Preliminaries

3.1.1 Localisation data

In order to conveniently globalize the theory of the previous chapter, we'll need a number of constructions and lemmas. By definition, manifolds are covered by open sets that are diffeomorphic to open sets of vector spaces. But for us it is slightly more convenient to work with smooth open embeddings of whole vector spaces. Here a smooth open embedding from a manifold X to a manifold Y is a smooth map $\varphi : X \to Y$ which is open and for which there is some smooth $\psi : \varphi(X) \to X$ such that $\psi \circ \varphi = \mathrm{Id}$ (hence also and $\varphi \circ \psi = \mathrm{Id}$). Remember that a family of sets V_i in a topological space X is locally finite if every point of X has a neighborhood that intersects only finitely many V_i. Note that in this whole text, every manifold is paracompact by definition. In particular their topology are metrizable and we will arbitrarily fix a compatible distance function on every manifold.

Definition 3.1. *Given smooth open embeddings $\varphi : X \to M$ and $\psi : Y \to N$, the update of a map $f : M \to N$, using a map $g : X \to Y$ is the map from M to N sending m to $\psi \circ g \circ \varphi^{-1}(m)$ if $m \in \varphi(X)$ and $f(m)$ otherwise.*

Lemma 3.2. *Let $\varphi : P \times X \to M$ and $\psi : P \times Y \to N$ be families of smooth open embeddings. Let K be a set in X whose image in M is closed. Let $f : P \times M \to N$ and $g : P \times X \to Y$ be smooth families of maps. If for each p and for every x not in K, $f_p(\varphi(x)) = \psi(g_p(x))$ then the family of maps f_p updated using g_p is smooth from $P \times M$ to N.*

Proof. Note that $P \times M = (P \times \varphi(X)) \cup (P \times \varphi(K)^c)$. Both those sets are open and the updated maps coincide with $(p, m) \mapsto \psi \circ g_p \circ \varphi^{-1}(m)$ on the first one and f on the second one. $\square$

Lemma 3.3. *Let $\varphi : P \times X \to M$ and $\psi : P \times Y \to N$ be families of smooth open embeddings. Let K_X and K_P be compact sets in X and P. Let $f : P \times M \to N$ be a continuous family of maps such that, for each p, $f_p(\varphi(X)) \subset \psi(Y)$. For every continuous function $\varepsilon : M \to \mathbb{R}_{>0}$, there is some positive number η such that, for every map $g : P \times X \to Y$ and every (p, p', x) in $K_P \times K_P \times K_X$, $d(g_{p'}(x), \psi^{-1} \circ f_p \circ \varphi(x)) < \eta$ implies $d(f'_{p'}(\varphi(x)), f_p(\varphi(x))) < \varepsilon(\varphi(x))$ where f' is obtained by updating f using g.*

Proof. Let ε be a positive continuous function on M. Since K_X is compact, we get a positive number ε_0 such that $\varepsilon(m) \geq \varepsilon_0$ for each m in K_X. We denote by K_1 the closed 1-thickening of the image of $K_P \times K_X$ under $(p, x) \mapsto \psi^{-1} \circ f_p \circ \varphi(x)$. This is a compact set so ψ is uniformly continuous on K_1 and we get a positive τ such that for all x and y in K_1, $d(x, y) < \tau \Rightarrow d(\psi(x), \psi(y)) < \varepsilon_0$.

We now prove that $\eta = \min(\tau, 1)$ is suitable. Fix (p, p', x) in $K_P \times K_P \times K_X$ such that $d(g_{p'}(x), \psi^{-1} \circ f_p \circ \varphi(x)) < \eta$. In particular $d(g_{p'}(x), \psi^{-1} \circ f_p \circ \varphi(x)) < 1$ hence $g_{p'}(x)$ is in K_1. Since $\psi^{-1} \circ f_p \circ \varphi(x)$ is also in K_1 and $d(g_{p'}(x), \psi^{-1} \circ f \circ \varphi(x)) < \tau$, we get $d(\psi \circ g_{p'}(x), \psi \circ \psi^{-1} \circ f_p \circ \varphi(x)) < \varepsilon_0$. This precisely means that $d(f'_{p'}(\varphi(x)), f_p(\varphi(x)) < \varepsilon_0$. Since (p, p', x) is in $K_P \times K_P \times K_X$, this is less than $\varepsilon(m)$. ☐

We now introduce a seemingly abstract definition, but its only goal is to treat uniformly the case of $\mathbb{N}$ and finite sets $\{0, \ldots, n\}$ so that we can treat uniformly atlases of compact and non-compact manifolds.

Definition 3.4. *A convenient indexing set is a totally ordered set ι equipped with maps $\pi : \mathbb{N} \to \iota$ and $\sigma : \iota \to \mathbb{N}$ such that π is order preserving and $\pi \circ \sigma = \mathrm{Id}$ (in particular π is surjective).*

In the case $\iota = \mathbb{N}$ we will use $\pi = \sigma = \mathrm{Id}$ while in the case $\iota = \{0, \ldots, n-1\}$ we use the unique order-preserving retraction as π and the inclusion as σ.

Lemma 3.5. *Let M be a manifold modelled on the normed space E and $(V_j)_{j \in J}$ a cover of M by open sets. There exists a convenient indexing set ι and a family of smooth open embeddings $\varphi : \iota \times E \to M$ such that*

- *for each i there is some j such that $\varphi_i(E) \subseteq V_j$,*
- *$i \mapsto \varphi_i(E)$ is a locally-finite collection of sets in M,*
- *$\bigcup_i \varphi_i(B_E(0, 1)) = M$ where $B_E(0, 1)$ is the open unit ball in E.*

Proof. The proof is a standard compact-exhaustion argument. Let $K_0, K_1, K_2, \ldots$ be a compact exhaustion of M and define:

$$
\begin{aligned}
C_n &= K_{n+2} \setminus K^o_{n+1}, \\
U_n &= K^o_{n+3} \setminus K_n.
\end{aligned}
$$

Thus:

- C_n is compact,
- U_n is open,
- $C_n \subseteq U_n$,
- $\bigcup_n C_n = M$,
- $U_n \cap U_m = \varnothing$ if $|n - m| > 2$.

For any $y \in E$ and $r > 0$, fix a smooth diffeomorphism $f_{y,r} : E \simeq B_E(y, r)$ such that $f_{y,r}(0) = y$. For each n and $x \in C_n$, let ψ_x be a smooth chart mapping an open neighbourhood of x to an open set of the model space E. Writing $y = \psi_x(x) \in E$, let:

$$
\begin{aligned}
B_{n,x} &= \psi_x^{-1}(B_E(y, r)), \\
W_{n,x} &= \psi_x^{-1}(f_{y,r}(B_E(0, 1))),
\end{aligned}
$$

for some $r > 0$ (which may depend on n, x) sufficiently small that:

- $B_E(y, r)$ lies in the target of the chart ψ_x,
- $B_{n,x}$ is contained in U_n,
- $B_{n,x}$ is contained in V_j for some j.

Note that $x \in W_{n,x}$. For each n, choose a finite subcovering of C_n by $W_{n,x_1}, \dots, W_{n,x_{l_n}}$ and define $\iota \subseteq \mathbb{N} \times M$ by:

$$
\iota = \bigcup_n \{(n, x_1), \dots, (n, x_{l_n})\}.
$$

Note that ι is countable and furthermore:

- for each $i \in \iota$, there is some j such that $B_i \subseteq V_j$,
- $(B_i)_{i\in\iota}$ is locally-finite (indeed more is true: B_i meets only finitely-many $B_{i'}$ for $i, i' \in \iota$ since $B_{m,x} \cap B_{n,x'} = \varnothing$ if $|n - m| > 2$),
- $(W_i)_{i\in\iota}$ covers M.

Given $i = (n, x_j) \in \iota$, the required map $\phi_i : E \to M$ is just:

$$
E \simeq B_E(y_j, r) \simeq B_{n,j} \subseteq M.
$$

□

Definition 3.6. *Let $f : M \to N$ be a continuous map between manifolds. A localisation data for f is a tuple $(E, F, \iota, \iota', \varphi, \psi, j)$ where E and F are normed vector spaces, ι is a convenient indexing set, ι' is a set (that is morally also convenient but that will play no role), $\varphi : \iota \times E \to M$ and $\psi : \iota' \times F \to N$ are families of smooth open embeddings, and $j : \iota \to \iota'$ such that:*

- *$\bigcup_i \varphi_i(B_E) = M$ where B_E is the open unit ball in E,*
- *$\bigcup_i \psi_i(B_F) = N$ where B_F is the open unit ball in F,*
- *$\forall i, f(\varphi_i(E)) \subset \psi_{j(i)}(B_F)$ where B_F is the open unit ball in F,*
- *$i \mapsto \psi_i(F)$ is locally finite.*

Such a tuple will be denoted by (φ, ψ, j) for brevity.

Lemma 3.7. *Any continuous map between manifolds has some localisation data.*

Proof. The preceding lemma (applied to the trivial cover of N by itself) gives a family of $\psi : \iota' \times F \to N$ of open smooth embeddings that the images of B_F cover N. We then apply this lemma again to the cover of M given by all $f^{-1}(\psi_j(B_F))$. $\square$

The general idea will be to apply the results of the previous chapters to all the $\psi_{j(i)}^{-1} \circ f \circ \varphi_i : E \to F$ for some maps f. However we must be careful that doing this for some i does not ruin the setup for the next i. This is easier to control using a distance function on the target manifold as in Lemma 3.9 below. First we need a general lemma about a single metric space (actually the formalized statement is stronger, it assumes only closed sets instead of compact ones, but here we explain the easier proof which is sufficient for our purposes).

Lemma 3.8. *In a metric space X, let $U : \iota \to \mathcal{P}\, X$ be a family of open subsets of X and let $K : \iota \to \mathcal{P}\, X$ be a locally-finite family of closed subsets such that $K_i \subset U_i$ for all i. There exists a continuous function $\delta : X \to \mathbb{R}_{>0}$ such that:*

$$\forall x\, x',\ \forall i, [x \in K_i \text{ and } d(x, x') < \delta(x)] \Rightarrow x' \in U_i.$$

Proof. We first note that, for any given i, compactness of K and openness of V_i give a positive number δ_i such that the δ_i-neighborhood of K_i is contained in V_i. We now prove that solutions exist locally. Let x be any point in X. From the local finiteness assumption, we get a neighborhood U of x such that $\{i | U \cap V_i \neq \varnothing\}$ is finite. The constant function with value the minimum of the corresponding δ_i is a solution on U. Since the condition we put on δ is convex, we can glue those local solutions using a partition of unity. $\square$

Lemma 3.9. *Let $f : M \to N$ be a continuous map between manifolds, and let (φ, ψ, i) be some localisation data for f. There exists a continuous positive function $\varepsilon : M \to \mathbb{R}_{>0}$ such that:*

$$\forall g : M \to N, \left[\forall m,\ d(f(m), g(m)) < \varepsilon(m)\right] \Rightarrow \forall i,\ g(\varphi_i(E)) \subset \psi_{j(i)}(F).$$

Note that, in the preceding lemma, the conclusion $g(\varphi_i(E)) \subset \psi_{j(i)}(F)$ is weaker than the condition $f(\varphi_i(E)) \subset \psi_{j(i)}(B_F)$ that appears in the definition of localisation data.

The condition $\forall m,\ d(g(m), f(m)) < \varepsilon(m)$ will be abbreviated $d(g, f) < \varepsilon$.

Proof. The preceding lemma applied to the family of open sets $\psi_j(F)$ and the family of compact sets $\psi_j(\overline{B_F})$ give a positive continuous function $\delta : N \to \mathbb{R}$ such that $\varepsilon = \delta \circ f$ is suitable. Indeed, assume $g : M \to N$ satisfies $d(g, f) < \varepsilon$ and fix some i and some $m \in \varphi_i(E)$. We know $f(m) \in \psi_{j(i)}(\overline{B_F})$ and our assumption on g gives $d(g(m), f(m)) < \delta(f(m))$. So the property of δ ensures $g(m) \in \psi_{j(i)}(F)$. $\square$

Lemma 3.10. *Let M a topological space and Z a set. Let $f : \mathbb{N} \times M \to Z$ be a sequence of functions. If the family of sets $n \mapsto \{x \parallel f_{n+1}(x) \neq f_n(x)\}$ is locally finite then there exists $U : M \to \mathcal{P}\, M$ and $n_0 : M \to \mathbb{N}$ such that:*

$$\forall x,\ U_x \in \boldsymbol{\mathcal{N}}_x \text{ and } \forall n \geq n_0(x),\ f_n|_{U_x} = f_{n_0(x)}|_{U_x}.$$

Note that the conclusion of above lemma ensures that the sequence f_n converges pointwise and the limit inherits all local properties of the f_n (such as continuity or differentiability when applicable).

Proof. The local finiteness assumption provides a function $U : M \to \mathcal{P}\, M$ such that

$$\forall x,\ U_x \in \boldsymbol{\mathcal{N}}_x \text{ and } \{n \mid \{y \mid f_n(y) \neq f_{n+1}(y)\} \cap U_x \neq \varnothing\} \text{ is finite.}$$

For any x, we denote by N_x the largest n of the finite set mentioned above. By construction, $n_0 : x \mapsto N_x + 1$ is suitable. $\square$

3.1.2 Jets spaces

We now need to introduce the bundles that will replace the jet spaces $E \times F \times \mathrm{Hom}(E, F)$ from the previous chapter. We need a couple of fiber bundles constructions.

Definition 3.11. *For every bundle $p : E \to B$ and every map $f : B' \to B$, the pull-back bundle $f^*E \to B'$ is defined by $f^*E = \{(b', e) \in B' \times E \mid p(e) = f(b')\}$ with the obvious projection to B'.*

Definition 3.12. *Let $E \to B$ and $F \to B$ be two vector bundles over some smooth manifold B. The bundle $\mathrm{Hom}(E, F) \to B$ is the set of linear maps from E_b to F_b for some b in B, with the obvious projection map.*

Set-theoretically, one can define $\mathrm{Hom}(E, F)$ as the set of subsets S of $E \times F$ such that there exists b such that $S \subset E_b \times F_b$ and S is the graph of a linear map. But the type theory formalization will use other tricks here. The facts that really matter are listed in Lemma 3.15 below.

Definition 3.13. *Let M and N be smooth manifolds. Denote by p_1 and p_2 the projections of $M \times N$ to M and N respectively.*
The space $J^1(M, N)$ of 1-jets of maps from M to N is $Hom(p_1^*TM, p_2^*TN)$

We will use notations like (m, n, φ) to denote an element of $J^1(M, N)$, but one should keep in mind that $J^1(M, N)$ is not a product, since φ lives in $\mathrm{Hom}(T_mM, T_nN)$ which depends on m and n.

Definition 3.14. *The 1-jet of a smooth map $f : M \to N$ is the map from m to $J^1(M, N)$ defined by $j^1f(m) = (m, f(m), T_mf)$.*

The composition of a section $\mathcal{F} : M \to J^1(M, N)$ with the projection onto N will sometimes be denoted by $\mathrm{bs}\,\mathcal{F} : M \to N$ and called the base map of $\mathcal{F}$. For any m, $\mathcal{F}(m)_\varphi$ will denote the component of $\mathcal{F}(m)$ living in $\mathrm{Hom}(T_mM, T_{\mathrm{bs}\,\mathcal{F}(m)}N)$.

Lemma 3.15. *For every smooth map $f : M \to N$,*

1. *j^1f is smooth*
2. *j^1f is a section of $J^1(M, N) \to M$*

Proof. Points 2 and 3 are obvious by construction.

To show that j^1f is smooth, suppose that M is modelled over E with charts $C_x : M \to E$ and coordinate change functions $C_{x,x'} = C_{x'}C_x^{-1} : E \to E$ and similarly let C'_y be charts for N. By construction of the 1-jet bundle, we need to check that for each x_0 the map

$$x \mapsto TC'_{f(x),f(x_0)} \circ T(C'_{f(x)}fC_x^{-1}) \circ T_{C_{x_0}(x)}(C_{x_0,x}) : M \to L(E, E)$$

is smooth at x_0 (we occasionally omit the point where the tangent maps are taken). For x close to x_0 the coordinate changes are smooth, so we can write

$$\begin{aligned} TC'_{f(x),f(x_0)} \circ T(C'_{f(x)}fC_x^{-1}) \circ T(C_{x_0,x}) &= T_{C_{x_0}(x)}(C'_{f(x),f(x_0)}C'_{f(x)}fC_x^{-1}C_{x_0,x}) \\ &= T_{C_{x_0}(x)}(C'_{f(x_0)}fC_{x_0}) \end{aligned}$$

This is smooth since $C'_{f(x_0)}fC_{x_0}$ is smooth. ∎

Definition 3.16. *A section $\mathcal{F}$ of $J^1(M, N) \to M$ is called holonomic if it is the 1–jet of its base map. Equivalently, $\mathcal{F}$ is holonomic if there exists $f : M \to N$ such that $\mathcal{F} = j^1f$, since such a map is necessarily $\mathrm{bs}\,\mathcal{F}$.*

3.2 First order differential relations

Definition 3.17. *A first order differential relation for maps from M to N is a subset $\mathcal{R}$ of $J^1(M,N)$.*

Definition 3.18. *A formal solution of a differential relation $\mathcal{R} \subset J^1(M,N)$ is a section of $J^1(M,N) \to M$ taking values in $\mathcal{R}$.*

Definition 3.19. *A homotopy of formal solutions of $\mathcal{R}$ is a smooth family of sections $\mathcal{F} : \mathbb{R} \times M \to J^1(M,N)$ such that each $m \mapsto \mathcal{F}(t,m)$ is a formal solution.*

The next definition will be used in cases where X and Y are vector spaces, in order to relate the global theory to the local one.

Definition 3.20. *Given manifolds M, X, N and Y and smooth open embeddings $g : Y \to N$ and $h : X \to M$ we get a transfer map $\psi_{g,h} : J^1(X,Y) \to J^1(M,N)$ defined by*

$$\psi_{g,h}(x,y,\varphi) = (h(x), g(y), T_y g \circ \varphi \circ (T_x h)^{-1})$$

and an operator on sections which sends $\mathcal{F} : M \to J^1(M,N)$ to $\Psi_{g,h}\mathcal{F} : X \to J^1(X,Y)$ defined when $\operatorname{bs}\mathcal{F}(h(X)) \subset g(Y)$ by

$$\Psi_{g,h}\mathcal{F}(x) = (x, g^{-1} \circ \operatorname{bs}\mathcal{F} \circ h(x), (T_{g^{-1}\circ \operatorname{bs}\mathcal{F}\circ h(x)} g)^{-1} \circ \mathcal{F}(h(x))_\varphi \circ T_x h).$$

Given a relation $\mathcal{R} \subset J^1(M,N)$, the induced relation in $J^1(X,Y)$ is $\psi_{g,h}^{-1}\mathcal{R}$.

The following is a localization lemma needed to take advantage of all the work from the previous chapter.

Lemma 3.21. *In the situation of the previous definition, given a section $\mathcal{F} : M \to J^1(M,N)$:*

- *$\Psi_{g,h}(\mathcal{F})$ is a smooth section of $J^1(X,Y)$.*
- *$\mathcal{F}$ is holonomic on $s \subset h(X) \cap \operatorname{bs}\mathcal{F}^{-1}(g(Y))$ if and only if $\Psi_{g,h}(\mathcal{F})$ is holonomic on $h^{-1}(s)$.*
- *$\mathcal{F}$ is a formal solution of $\mathcal{R}$ on $h(X) \cap \operatorname{bs}\mathcal{F}^{-1}(g(Y)$ if and only if $\Psi_{g,h}(\mathcal{F})$ is a formal solution of the induced relation $\Psi_{g,h}^{-1}\mathcal{R}$.*

Proof. The first point is clear by composition. In order to prove the second point while keeping notations under control, we set $f(x) = g^{-1} \circ \operatorname{bs}\mathcal{F} \circ h$. Using this notation $\Psi_{g,h}\mathcal{F}(x) = (x, f(x), (T_{f(x)}g)^{-1} \circ \mathcal{F}(h(x))_\varphi \circ T_x h)$. We have

$$\begin{aligned} T_x f &= T_{\operatorname{bs}\mathcal{F}\circ h(x)}(g^{-1}) \circ T_{h(x)} \operatorname{bs}\mathcal{F} \circ T_x h \\ &= \left(T_{f(x)} g\right)^{-1} \circ T_{h(x)} \operatorname{bs}\mathcal{F} \circ T_x h \end{aligned}$$

hence $\Psi_{g,h}\mathcal{F}$ is holonomic at x if and only if $\left(T_{f(x)}g\right)^{-1} \circ \mathcal{F}(h(x))_\varphi \circ T_x h = \left(T_{f(x)}g\right)^{-1} \circ T_{h(x)} \operatorname{bs}\mathcal{F} \circ T_x h$ and this is equivalent to $\mathcal{F}(h(x))_\varphi = T_{h(x)} \operatorname{bs}\mathcal{F}$ which is the holonomy condition for $\mathcal{F}$ at $h(x)$.

The third point is a direct consequence of the easy formula $\psi_{g,h} \circ \Psi_{g,h}(\mathcal{F}) = F \circ h$. $\square$

Definition 3.22. *A first order differential relation $\mathcal{R} \subset J^1(M,N)$ satisfies the h-principle if every formal solution of $\mathcal{R}$ is homotopic to a holonomic one. It satisfies the parametric h-principle if, for every manifold P and every closed set C in $P \times M$, every family $\mathcal{F} : P \times M \to J^1(M,N)$ of formal solutions which are holonomic for (p,m) near C is homotopic to a family of holonomic ones relative to C.*

Parametricity for free

In many cases, relative parametric h-principles can be deduced from relative non-parametric ones with a larger source manifold. Let X, P and Y be manifolds, with P seen a parameter space. Denote by Ψ the map from $J^1(X \times P, Y)$ to $J^1(X, Y)$ sending (x, p, y, ψ) to $(x, y, \psi \circ \iota_{x,p})$ where $\iota_{x,p} : T_xX \to T_xX \times T_pP$ sends v to $(v, 0)$.

To any family of sections $F_p : x \mapsto (f_p(x), \varphi_{p,x})$ of $J^1(X, Y)$, we associate the section $\bar{F}$ of $J^1(X \times P, Y)$ sending (x, p) to $\bar{F}(x, p) := (f_p(x), \varphi_{p,x} \oplus \partial f/\partial p(x, p))$.

Lemma 3.23. *In the above setup, we have:*

- *$\bar{F}$ is holonomic at (x, p) if and only if F_p is holonomic at x.*
- *F is a family of formal solutions of some $\mathcal{R} \subset J^1(X, Y)$ if and only if $\bar{F}$ is a formal solution of $\mathcal{R}^P := \Psi^{-1}(\mathcal{R})$.*

Proof. For the first part, the derivative of $\bar{F}$ is $\partial f/\partial x(x, p) \oplus \partial f/\partial p(x, p)$, which is equal to $\bar{F}_\varphi$ iff $\partial f/\partial x(x, p) = f_\varphi$.

The second part follows from $\Psi \circ \bar{F}(x, p) = F_p(x)$. ☐

Lemma 3.24. *Let $\mathcal{R}$ be a first order differential relation for maps from M to N. If, for every manifold with boundary P, $\mathcal{R}^P$ satisfies the h-principle then $\mathcal{R}$ satisfies the parametric h-principle. Likewise, the C^0-dense and relative h-principle for all $\mathcal{R}^P$ imply the parametric C^0-dense and relative h-principle for $\mathcal{R}$.*

Proof. By Lemma 3.23 we can turn a formal solution of $\mathcal{R}$ into a formal solution of $\mathcal{R}^P$, so we get a homotopy to a holonomic formal solution. We can turn this homotopy back to a homotopy of the original formal solution. ☐

3.3 The h-principle for open and ample differential relations

In this chapter, X and Y are smooth manifolds and $\mathcal{R}$ is a first order differential relation on maps from X to Y: $\mathcal{R} \subset J^1(X, Y)$. For any $\sigma = (x, y, \varphi)$ in $\mathcal{R}$ and any dual pair $(\lambda, v) \in T_x^*X \times T_xX$, we set:

$$\mathcal{R}_{\sigma,\lambda,v} = \mathrm{Conn}_{\varphi(v)} \left\{ w \in T_yY \ ; \ (x,\ y,\ \varphi + (w - \varphi(v)) \otimes \lambda) \in \mathcal{R} \right\}$$

where $\mathrm{Conn}_a A$ is the connected component of A containing a. In order to decipher this definition, it suffices to notice that $\varphi + (w - \varphi(v)) \otimes \lambda$ is the unique linear map from T_xX to T_yY which coincides with φ on $\ker \lambda$ and sends v to w. In particular, $w = \varphi(v)$ gives back φ.

Of course we will want to deal with more than one point, so we will consider a vector field V and a 1–form λ such that $\lambda(V) = 1$ on some subset U of X, a formal solution F (defined at least on U), and get the corresponding $\mathcal{R}_{F,\lambda,v}$ over U.

One easily checks that $\mathcal{R}_{\sigma,\kappa^{-1}\lambda,\kappa v} = \kappa \mathcal{R}_{\sigma,\lambda,v}$ hence the above definition only depends on $\ker \lambda$ and the direction $\mathbb{R}V$.

Definition 3.25. *A relation $\mathcal{R}$ is ample if, for every $\sigma = (x, y, \varphi)$ in $\mathcal{R}$ and every (λ, v), the slice $\mathcal{R}_{\sigma,\lambda,v}$ is ample in T_yY.*

Lemma 3.26. *Given manifolds W, X, Y and Z and smooth open embeddings $g : Z \to Y$ and $h : W \to X$, the relation induced (in the sense of Definition 3.20) in $J^1(W, Z)$ by a ample relation in $J^1(X, Y)$ is ample.*

Proof. By definition, the relation induced by $\mathcal{R}$ is $\psi_{g,h}^{-1}\mathcal{R}$ where $\psi_{g,h}(w, z, \varphi) = (h(w), g(z), T_z g \circ \varphi \circ (T_w h)^{-1})$. Fix $\sigma = (w, z, \varphi) \in \psi_{g,h}^{-1}\mathcal{R}$ and a dual pair (λ, v) on $T_w W$. We set $G = T_z g$ and $H = T_w h$. Both those maps are linear isomorphisms. We compute the slice corresponding to (σ, λ, v):

$$\begin{aligned}
\psi_{g,h}^{-1}\mathcal{R}(\sigma, \lambda, v) &= \mathrm{Conn}_{\varphi v}\left\{u \in T_w W \;\middle|\; (w, z, \varphi + (u - \varphi v) \otimes \lambda) \in \psi_{g,h}^{-1}\mathcal{R}\right\} \\
&= \mathrm{Conn}_{\varphi v}\left\{u \in T_w W \;\middle|\; (h(w), g(z), G \circ (\varphi + (u - \varphi v) \otimes \lambda) \circ H^{-1}) \in \mathcal{R}\right\} \\
&= G^{-1}\mathcal{R}(\psi_{g,h}\sigma, \lambda \circ H^{-1}, Hv).
\end{aligned}$$

Hence the slice $\psi_{g,h}^{-1}\mathcal{R}(\sigma, \lambda, v)$ is the image of a slice of $\mathcal{R}$ under a linear isomorphism, hence ample. □

Lemma 3.27. *The relation of immersions of M into N in positive codimension is open and ample.*

Proof. For every $\sigma = (x, y, \varphi)$ in the immersion relation $\mathcal{R}$, and for every dual pair (π, v), the slice $\mathcal{R}(\sigma, \pi, v)$ is the set of w which do not belong to the image of $\ker \pi$ under φ. Since $\dim M > \dim N$, this image has codimension at least 2 in $T_y N$, and Lemma 2.13 concludes. □

Theorem 3.28 (Gromov)**.** *For any manifolds M and N, any relation $\mathcal{R} \subset J^1(M, N)$ that is open and ample satisfies the full h-principle (relative, parametric and C^0-dense).*

We first explain how to get rid of parameters, using the relation $\mathcal{R}^P$ for families of solutions parametrized by P.

Lemma 3.29. *If $\mathcal{R}$ is ample then, for any parameter space P, $\mathcal{R}^P$ is also ample.*

Proof. We fix $\sigma = (x, y, \psi)$ in $\mathcal{R}^P$. For any $\lambda = (\lambda_X, \lambda_P) \in T_x^*X \times T_p^*P$ and $v = (v_X, v_P) \in T_x X \times T_p P$ such that $\lambda(v) = 1$, we need to prove that $\mathrm{Conv}\, \mathcal{R}^P_{\sigma,\lambda,v} = T_y Y$. Unfolding the definitions gives:

$$\mathcal{R}^P_{\sigma,\lambda,v} = \mathrm{Conn}_{\varphi(v)}\left\{w \in T_y Y \;;\; (x,\ y,\ \psi \circ \iota_{x,p} + (w - \psi(v)) \otimes \lambda_X) \in \mathcal{R}\right\}.$$

A degenerate but easy case is when $\lambda_X = 0$. Then the condition on w becomes $\psi \circ \iota_{x,p} \in \mathcal{R}$, which is true by definition of $\mathcal{R}^P$, so $\mathcal{R}^P_{\sigma,\lambda,v} = T_y Y$.

We now assume λ_X is not zero and choose $u \in T_x X$ such that $\lambda_X(u) = 1$. We then have $\mathcal{R}^P_{\sigma,\lambda,v} = \mathcal{R}_{\Psi\sigma,\lambda_X,u} + \psi(v) - \psi \circ \iota_{x,p}(u)$. Because $\mathcal{R}$ is ample and taking convex hull commutes with translation, we get that $\mathrm{Conv}\, \mathcal{R}^P_{\sigma,\lambda,v} = T_y Y$. □

Proof of Theorem 3.28. Lemmas 3.24 and 3.29 prove we can assume there are no parameters. So we start with a single formal solution F of $\mathcal{R}$, which is holonomic near some closed subset $A \subset X$.

We apply Lemma 3.7 to get some localisation data $(\varphi : \iota \to \mathcal{P}X, \psi : \iota' \to \mathcal{P}Y, j)$ for $\mathrm{bs}\, F : X \to Y$. Lemma 3.9 then provides a continuous function $\varepsilon : X \to \mathbb{R}_{>0}$ such that every function g with $d(\mathrm{bs}\, F, g) < \varepsilon$ sends each $\varphi_i(E)$ into $\psi_{j(i)}(F)$. Since ι is finite or countable we can assume it is a convenient indexing set (see Definition 3.4). We denote by $\pi : \mathbb{N} \to \iota$ and $\sigma : \iota \to \mathbb{N}$ the corresponding structure maps.

We will construct by induction on n a sequence of homotopies of sections $F_n : [0,1] \times X \to J^1(X, Y)$ such that, for all n,

- $F_{n,t}(x)$ coincides with $F(x)$ for all t if $n = 0$ or x is close to A ;
- for all t, $d(\operatorname{bs} F, \operatorname{bs} F_{n,t}) < \varepsilon$;
- if $\pi(n+1) = \pi(n)$ then $F_{n+1,t}$ coincides with $F_{n,t}$ for all t ;
- $F_{n+1,1}$ is holonomic on $\bigcup_{i<\pi(n)} \varphi_i(\bar{B}_E)$;
- each $F_{n+1,t}$ coincides with $F_{n,t}$ outside $\varphi_{\pi(n)}(E)$.

The induction construction starts with setting $F_{0,t} = F$ for all t, which has the required properties (the first two conditions are clear and the other ones don't say anything about F_0). Now assume F_n has been constructed. If $\pi(n+1) = \pi(n)$ then we set $F_{n+1} = F_n$. Otherwise we have $\pi(n+1) > \pi(n)$.

For any t, since $d(\operatorname{bs} F, \operatorname{bs} F_{n,t}) < \varepsilon$ by induction hypothesis, $\operatorname{bs} F_{n,t}$ sends $\varphi_i(E)$ into $\psi_{j(i)}(F)$.

Definition 3.20 then turns F_n into a homotopy of sections $\mathcal{F}$ of $J^1(E, F)$. According to Lemma 3.21, each $\mathcal{F}_t$ is a formal solution of the relation $\mathcal{R}_i$ induced by $\mathcal{R}$ in $J^1(E, F)$ via $\varphi_{\pi(n)}$ and $\psi_{j(\pi(n))}$, $\mathcal{F}$ is relative to $\varphi_{\pi(n)}^{-1}(A)$ and $\mathcal{F}_1$ is holonomic near $\varphi_{\pi(n)}^{-1}(A \cup \bigcup_{i<\pi(n)} \varphi_i(\bar{B}_E))$.

The new homotopy F_{n+1} will be constructed by updating F_n using some homotopy $\mathcal{F}'$ of sections of $J^1(E, F)$. In order to ensure $d(\operatorname{bs} F, \operatorname{bs} F_{n+1}) < \varepsilon$, it suffices to ensure that, for each x and t, either $\mathcal{F}_{n+1,t}(x) = \mathcal{F}_{n,t'}(x)$ for some t' or $d(\operatorname{bs} F_{n,1}(x), \operatorname{bs} F_{n+1,t}(x)) < \varepsilon(x) - d(\operatorname{bs} F(x), \operatorname{bs} F_{n,1}(x))$. The latter will hold as soon as, for all e and t, $\| \operatorname{bs} \mathcal{F}_1(e) - \operatorname{bs} \mathcal{F}'_t(e) \| < \eta$ for some positive η given by Lemma 3.3 (applied to $P = \mathbb{R}$, M and N). So Lemma 2.16 gives a suitable $\mathcal{F}'$.

Now that the inductive construction is completed, we apply Lemma 3.10 to learn our sequence F_n is locally ultimately constant, hence it converges pointwise to a smooth homotopy relative to A and ending at a holonomic section of $\mathcal{R}$. $\square$

Theorem 3.30 (Smale 1958)**.** *There is a homotopy of immersions of $\mathbb{S}^2$ into $\mathbb{R}^3$ from the inclusion map to the antipodal map $a : q \mapsto -q$.*

Proof. We denote by ι the inclusion of $\mathbb{S}^2$ into $\mathbb{R}^3$. We set $j_t = (1-t)\iota + ta$. This is a homotopy from ι to a (but not an immersion for $t = 1/2$). Using the canonical trivialization of the tangent bundle of $\mathbb{R}^3$, we can set, for $(q, v) \in T\mathbb{S}^2$, $G_t(q, v) = \operatorname{Rot}_{Oq}^{\pi t}(v)$, the rotation around axis Oq with angle πt. The family $\sigma : t \mapsto (j_t, G_t)$ is a homotopy of formal immersions relating $j^1\iota$ to $j^1 a$. It is homotopic by reparametrization to a homotopy of formal immersions relating $j^1\iota$ to $j^1 a$ which are holonomic for t near the 0 and 1.

The above theorem ensures this family is homotopic, relative to $t = 0$ and $t = 1$, to a family of holonomic formal immersions, ie a family $t \mapsto j^1 f_t$ with $f_0 = \iota$, $f_1 = a$, and each f_t is an immersion. $\square$

Appendix A

Local sphere eversion

The local theory of Chapter 2 is already enough to deduce Smale's sphere eversion theorem, although it is less natural than going through the general results of Chapter 3. The goal of this appendix is to explain how to do so. In this section E denote a finite dimensional real vector space equipped with an inner product. Later we will assume it is 3-dimensional. We denote by $\mathbb{S}$ the unit sphere in E.

We want to study immersions of $\mathbb{S}$ into E, but we want to work only with functions defined on the whole E. An immersion of $\mathbb{S}$ into E is a smooth map $f : E \to E$ such that for every x in $\mathbb{S}$, $Df(x)$ is injective on $T_x\mathbb{S} = x^\perp$. At face value this may sound slightly stronger than the definition of an abstract immersion from manifold theory. But one can easily prove that any abstract immersion extends to an immersion in the elementary sense. In any case, using a definition of immersion that is too strong would only make Smale's theorem stronger. We introduce two slightly artificial notions. We denote by B the open ball with radius 9/10 around the origin in E and set:

$$\mathcal{R} := \{(x, y, \varphi) \in J^1(E, E) \mid x \notin B \Rightarrow \varphi|_{x^\perp} \text{ is injective}\}.$$

Of course solutions of this relation restrict to immersions of $\mathbb{S}$.

Lemma A.1. *The relation $\mathcal{R}$ above is open.*

Proof. The main task is to fix $x_0 \notin B$ and $\varphi_0 \in L(E, E)$ which is injective on $x_0^\perp$ and prove that, for every x close to x_0 and φ close to φ_0, φ is injective on $x^\perp$. This is a typical situation where geometric intuition makes it feel like there is nothing to prove.

One difficulty is that the subspace $x^\perp$ moves with x. We reduce to a fixed subspace by considering the restriction to $x_0^\perp$ of the orthogonal projection onto $x^\perp$. One can check this is an isomorphism as long as x is not perpendicular to x_0. More precisely, we consider $f : J^1(E, E) \to \mathbb{R} \times L(x_0^\perp, E)$ which sends (x, y, φ) to $(\langle x_0, x\rangle, \varphi \circ \mathrm{pr}_{x^\perp} \circ j_0)$ where j_0 is the inclusion of $x_0^\perp$ into E. The set U of injective linear maps is open in $L(x_0^\perp, E)$ and the map f is continuous hence the preimage of $\{0\}^c \times U$ is open. This is good enough for us because injectivity of $\varphi \circ \mathrm{pr}_{x^\perp} \circ j_0$ implies injectivity of φ on the image of $\mathrm{pr}_{x^\perp} \circ j_0$ which is $x^\perp$ whenever $\langle x_0, x\rangle \neq 0$. □

Lemma A.2. *The relation $\mathcal{R}$ above is ample.*

Proof. The core fact here is that if one fixes vector spaces F and F', a dual pair (π, v) on F and an injective linear map $\varphi : F \to F'$ then the updated map $\Upsilon_p(\varphi, w)$ is injective if and only if w is not in $\varphi(\ker \pi)$. First we assume $\Upsilon_p(\varphi, \varphi(u))$ is injective for some u in $\varphi(\ker \pi)$ and derive a contradiction. We have $\Upsilon_p(\varphi, \varphi(u))\, v = \varphi(u)$ by the general definition of updating and also $\Upsilon_p(\varphi, \varphi(u))\, u = \varphi(u)$ since u is in $\ker \pi$. Hence injectivity of φ ensure $u = v$, which is absurd since $\pi(u) = 0$ and $\pi(v) = 1$. Conversely assume w is not in $\varphi(\ker \pi)$ and let us prove $\Upsilon_p(\varphi, w)$ is injective. Assume x is in the kernel of $\Upsilon_p(\varphi, w)$. Decompose $x = u + tv$ with $u \in \ker \pi$ and t a real number. We have $\Upsilon_p(\varphi, w)(x) = \varphi(u) + tw$. Hence our assumption on x implies t vanishes otherwise we would have $w = -t^{-1}\varphi(u)$ contradicting that w isn't in $\varphi(\ker \pi)$. This vanishing and the assumption on x then imply $\varphi(u) = 0$. Since φ is injective we conclude that $u = 0$ and finally $x = 0$.

We now turn to $\mathcal{R}$. It suffices to prove that for every $\sigma = (x, y, \varphi) \in \mathcal{R}$ and every dual pair $p = (\pi, v)$ on E, the slice $\mathcal{R}(\sigma, p)$ is ample. If x is in B then $\mathcal{R}(\sigma, p)$ is the whole E which is obviously ample. So we assume x is not in B. Since σ is in $\mathcal{R}$, φ is injective on $x^\perp$. The slice is the set of w such that $\Upsilon_p(\varphi, w)$ is injective on $x^\perp$. Assume first $\ker \pi = x^\perp$. Then $\Upsilon_p(\varphi, w)$ coincides with φ on $x^\perp$ hence the slice is the whole E. Assume now that $\ker \pi \neq x^\perp$. The slice is not very easy to picture in this case. But one should remember that, up to affine isomorphism, the slice depends only on $\ker \pi$. More specifically, if we keep π but change v then the slice is simply translated in E. Here we replace v by the projection on $x^\perp$ of the vector dual to π rescaled to keep the property $\pi(v) = 1$. What has been gained is that we now have $v \in x^\perp$ and $x^\perp = (x^\perp \cap \ker \pi) \oplus \mathbb{R}v$. Since φ is injective on $x^\perp$, $\varphi(x^\perp \cap \ker \pi)$ is a hyperplane in $x^\perp$ and $\Upsilon_p(\varphi, w)$ is injective on $x^\perp$ if and only if w is in the complement of $\varphi(x^\perp \cap \ker \pi)$ according to the core fact above. Since it is an hyperplane in $x^\perp$, it has codimension at least 2 in E hence its complement is ample. □

Theorem A.3 (Smale 1958)**.** *There is a homotopy of immersion of $\mathbb{S}^2$ into $\mathbb{R}^3$ from the inclusion map to the antipodal map $a \colon q \mapsto -q$.*

Proof. We denote by ι the inclusion of $\mathbb{S}^2$ into $\mathbb{R}^3$. We set $j_t = (1-t)\iota + ta$. This is a homotopy from ι to a (but not an immersion for $t = 1/2$). Using the canonical trivialization of the tangent bundle of $\mathbb{R}^3$, we can set, for $(q, v) \in T\mathbb{S}^2$, $G_t(q, v) = \mathrm{Rot}_{Oq}^{\pi t}(v)$, the rotation around axis Oq with angle πt. The family $\sigma : t \mapsto (j_t, G_t)$ is a homotopy of formal immersions relating $j^1\iota$ to $j^1 a$. Those formal solutions are holonomic when t is zero or one, so we can reparametrize the family to make such it is holonomic when t is close to zero or one. Then we can extend it to a homotopy of formal solutions of $\mathcal{R}$ using a suitable cut-off ensuring smoothness near the orign. The relation $\mathcal{R}$ is ample according to Lemma A.2 and then Lemma 3.23 ensures its 1-parameter version $\mathcal{R}^{\mathbb{R}}$ is also ample. The relation $\mathcal{R}$ is open according to Lemma A.1 hence $\mathcal{R}^{\mathbb{R}}$ is also ample. So we can use Lemma 2.15 to deform our family of formal solutions into a holonomic one. □