Nash-Moser and the h-principle

Utrecht University

Master's Thesis
Mathematical Sciences

Supervisor:
Dr. Álvaro del Pino Gómez

Author:
Ket Bottelier

Second reader:
Prof. Dr. Marius Crainic

November 2024

Nash-Moser and the h-principle

Abstract

The h-principle is a fruitful and active area of research with applications in many areas of differential geometry. Its main goal is to prove that the space of various geometric structures is (weakly) homotopy equivalent to more manageable spaces. There are various ways to obtain such results; from these we will introduce Eliashberg and Mishachev's famous holonomic ($\mathcal{R}$)-approximation theorem.

Nash-Moser theorems are a powerful class of theorems generalizing the inverse function theorem to spaces more difficult than Banach spaces, with wide applications in PDEs and differential geometry. They originate from Nash's seminal paper on isometric Riemannian embeddings. The method of proof was later adapted by Moser into a more general theorem which he subsequently used to solve various PDEs. Many other versions have since arisen; in this thesis we treat versions by Gromov and Hamilton.

Motivated by the problem of finding an h-principle for isotropic immersions into almost symplectic manifolds, we show how the Nash-Moser theorem may be used hand in hand with holonomic approximation to obtain h-principles for a certain kind of differential relation.

Acknowledgements

First and foremost, I would like to thank dr. Álvaro del Pino Gómez for being a great supervisor. Accomodating and patient throughout the process of writing this thesis, I never felt unwelcome to ask questions. The talks we had were always very thought-provoking and supplied me with plenty of new material to explore, as well as having a pleasant and relaxed atmosphere.

I would also like to thank prof. dr. Marius Crainic for taking the time and effort to be the second reader of this thesis.

Of course I thank my parents, who have supported me throughout the process and were always there for me. I am also very grateful to Joris, for lightening the mood and providing perspective when needed.

Contents

Chapter 1

Introduction

1.1 Isotropic subspaces and immersions

Symplectic geometry is an important field in mathematics with many applications in other fields, notably physics, from which it originates. Recall that a (real) symplectic vector space is a real vector space V equipped with a nondegenerate antisymmetric bilinear map $\Omega : V \times V \to \mathbb{R}$, called a symplectic bilinear map. This could be considered a counterpart to the definition of an inner product, which is of course required to be symmetric. The antisymmetry allows for subspaces $W \subseteq V$ on which Ω vanishes (meaning $\Omega(w, w') = 0$ for all $w, w' \in W$) and such subspaces are called *isotropic*. Isotropic subspaces of maximal dimension are called *Lagrangian*.

Passing to differential geometry, we call a manifold M equipped with a smooth choice of symplectic map $\omega_p : T_pM \times T_pM \to \mathbb{R}$ for each $p \in M$ an *almost symplectic manifold*. Stated differently: it is a manifold equipped with a nondegenerate two-form. A *symplectic manifold* (M, ω) is a manifold M equipped with a *closed* nondegenerate two-form ω. In either case, an immersion $f : N \to M$ is then *isotropic* if $\operatorname{Im} T_pf \subseteq T_{f(p)}M$ is isotropic for all $p \in N$, and Lagrangian immersions are defined similarly. It is now natural to attempt to extend the understanding of isotropic subspaces and investigate isotropic embeddings and immersions.

In the symplectic case, the closedness of the form ω leads to various results describing the symplectic manifold locally. The Darboux theorem states that we can find coordinates x^i, y^i around any p such that $\omega = \sum_i dx^i \wedge dy^i$, which amounts to identifying a neighborhood of p with an open subset of the standard symplectic vector space $\mathbb{R}^{2n}$. From this one can obtain many small isotropic embeddings. Similarly, if we are already given a Lagrangian submanifold L, Weinstein's tubular neighborhood theorem identifies a neighborhood of L with a neighborhood of the zero section of the cotangent bundle T^*L equipped with the the canonical two-form. From these, one can again find many nearby Lagrangian and isotropic submanifolds, for example by using generating functions.

The "global" question whether for some manifold N there exists an immersion into a given symplectic manifold (M, ω) requires a bit more setup. One way to approach it is by means of the h-principle.

1.2 The h-principle

The statement that $f : N \to (M, \omega)$ is isotropic can be written succinctly as $f^*\omega = 0$. Note that this defines a PDE. The theory of the h-principle deals with the reduction of PDE's to homotopy-theoretic problems. To give a quick idea of what this technique entails, consider a PDE for some function $f : \mathbb{R}^n \to \mathbb{R}$ given by:

$$\Phi(x, f(x), \partial^\alpha f \mid 1 \leq |\alpha| \leq n) = 0, \tag{1.1}$$

where $\Phi : \mathbb{R}^q \to \mathbb{R}$ (with $q = |\{1 \leq |\alpha| \leq n\}| + n + 1$) is some smooth function and α is used to denote multi-indices. A necessary requirement for this equation to be solvable is that there exists some function $F : \mathbb{R}^n \to \mathbb{R}^{q-n}$ such that $\Phi(x, F(x)) = 0$. Such an F is called a *formal solution*. Indeed, if Equation (1.1) had an *actual* solution f, then we could take:

$$F(x) = (f(x), \partial^\alpha f(x) \mid 1 \leq |\alpha| \leq n).$$

Such a formal solution, which corresponds to an actual solution, is called *holonomic*. We can extend the notion of formal solutions to statements that are defined by *inequalities* instead of equalities involving partial derivatives, for example replacing the equal sign in Equation (1.1) with $>$ or $\neq$. This is is then an example of a *partial differential relation* (PDR). A PDR is said to satisfy an *h-principle* if there is some homotopy-theoretic statement relating the spaces of formal and actual solutions. One variant of the h-principle, perhaps the most basic, states that any formal solution is homotopic to an actual solution. A stronger variant, the *full h-principle*, is satisfied if the spaces of formal and actual solutions are weakly homotopy equivalent.

The notion of a formal solution of a PDR can be generalized to maps between manifolds and, more generally, sections of fiber bundles. It turns out that many interesting and natural PDRs in differential geometry satisfy some form of h-principle. An early example of this was the study of regular homotopies: these are homotopies f_t such that for all $t \in [0,1]$ the map f_t is an immersion. Note that the property of being an immersion can be defined as a PDR, essentially generalizing the statement "$f' \neq 0$" for $f : \mathbb{R} \to \mathbb{R}^n$. One early result by Whitney ([Whi37]) classifies regular homotopies of maps $S^1 \to \mathbb{R}^2$. Smale showed that any two immersions of the sphere into $\mathbb{R}^3$ are regularly homotopic ([Sma59a]) (which yields the famous and surprising corollary that the sphere can be turned "inside out" through immersions) and classified immersions $S^n \to \mathbb{R}^m$ with $m \geq n+1$ ([Sma59b]). Hirsch extended his work to arbitrary manifolds, showing in particular that the existence of an immersion $N \to M$ is equivalent to a fiberwise injective bundle map $TN \to TM$ (a "formal immersion"), and that regular homotopies correspond to homotopies of such bundle maps ([Hir59]). Phillips provided similar results for submersions ([Phi67]).

The terms PDR and h-principle were introduced by Gromov ([Gro71]), who previously generalized the work of Smale, Hirsch and Phillips in his thesis ([Gro69]). His landmark book [Gro86] compiles a large amount of his work on the h-principle as well as results by others that can be framed in an h-principle setting. The book also outlines various methods for proving h-principles. One of these, the *method of continuous sheaves*, was later adapted by Eliashberg and Mishachev into the versatile *holonomic approximation theorems*. The "standard" holonomic approximation theorem, first published in [EM01], deals with the approximation of formal maps $F : \mathbb{R}^n \to \mathbb{R}^{q-n}$ (with q as in Equation (1.1)) by holonomic maps $x \mapsto (f(x), \partial^\alpha f(x))$. Another variant, *holonomic $\mathcal{R}$-approximation* ([CEM24], p. 283), states that under certain conditions the approximation can be taken to remain in a given PDR. This is done through a procedure that can intuitively be described as "gluing" solutions on small opens to obtain a solution near some submanifold. In order to do this, the PDR is required to satisfy the properties of *local integrability* and *microflexibility*. The former of these amounts to the existence of the small solutions, whereas the latter makes the gluing possible. Holonomic approximation will be our main h-principle-theoretic tool in this thesis.

Areas in which the h-principle has been influential include foliation theory, due to contributions by Thurston who related (possibly non-integrable) distributions to integrable distributions ([Thu74], [Thu76]) and, of particular interest to us, symplectic geometry (see [Eli15] for a survey).

1.3 Isotropic immersions

Eliashberg and Mishachev devote a large part of their celebrated book [EM02] to applications of the h-principle to symplectic geometry. Here they show that the h-principle is satisfied for subcritical (i.e. non-Lagrangian) isotropic embeddings, as well as istropic immersions (including Lagrangians, subject to some cohomological condition). As for virtually all proofs in symplectic geometry, these results rely heavily on the closedness of the symplectic form.

The symplectic form being closed has various implications. On the one hand, it means that there is a lack of local invariants, as seen in the Darboux theorem: locally, all symplectic forms "are the same". On the other hand symplectic forms possess a "large amount" of symplectomorphisms, as seen in the Moser theorems ([Sil01], pp. 46-49, or [CEM24], p. 230). These properties of are used by the authors of [EM02] to show that the conditions for the application of holonomic-$\mathcal{R}$-approximation (or variants thereof) are satisfied. As mentioned before, for example, the Darboux theorem implies that one can always find small isotropics around a given point, which shows that we have local integrability.

In the almost symplectic case, the existence of local invariants means that the property of being a symplectomorphism is much more restrictive: since $d\omega$ is nonzero an automorphism $\phi : (M,\omega) \to (M,\omega)$ must not only satisfy $\phi^*\omega = \omega$, but also $\phi^* d\omega = d\omega$. Without requirements on $d\omega$, there might for instance be $p, q \in M$ such that $(d\omega)_p \neq 0$ and $(d\omega)_q = 0$, meaning there is no symplectomorphism ϕ such that $\phi(p) = q$. More generally, $d\omega$ might have different rank at various points [1].

On p.325 of [CEM24], the authors pose the following open problems:

Problem 1.3.1. *Is there any form of the h-principle for:*

1. *Lagrangian and isotropic immersions into an almost symplectic manifold?*
2. *isosymplectic immersions between almost symplectic manifolds?*
3. *coisotropic and isometric coisotropic immersions into almost symplectic manifolds?*

In this thesis we attempt to solve the first problem of the above. As discussed, the lack of almost-symplectomorphisms leads us to consider methods not found in [CEM24].

1.4 The Nash-Moser theorem

As it happens, a differential equation similar to the one defining isotropics:

$$f^*\omega = 0$$

was considered by Gromov in [Gro96], pp. 249-260 (see also [Pan16]). Instead of the two-form ω the authors were interested in a (set of) one-forms defining a tangent distribution, with f being a horizontal immersion [2]. The main theorem on which the construction of horizontal immersions hinges is Gromov's Nash-Moser theorem ([Gro86], p. 117), although this is not applied directly. Instead it is used to prove microflexibility, which ultimately allows the gluing of local solutions to produce global solutions.

Nash-Moser theorems can be thought of as versions of the submersion theorem. Recall that this theorem says that any submersion $f : \mathbb{R}^m \to \mathbb{R}^n$ (with $m \geq n$) can locally be written as the projection $(x_1, \dots, x_m) \mapsto (x_1, \dots, x_n)$. One implication of this is that the map f is open and we have a right inverse $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, 0, \dots, 0)$. Note that the only requirement is that the map f be a submersion at a *single point* $p \in \mathbb{R}^m$. A similar result also holds for Banach spaces which, in the same way as the finite-dimensional case, derives from the inverse function theorem. For other kinds of functional spaces however, it is not sufficient that a map between them has a surjective derivative at a single point. This is the case, for example, for *Fréchet spaces*, see [Ham82] pp. 121-129.

One problem where this arises is the isometric embedding problem treated by Nash in [Nas56], which is the origin of the Nash-Moser theorem. In this paper, Nash shows that for any n-dimensional Riemannian manifold (M, g) there are $m \in \mathbb{N}$ [3] and an isometric embedding $f : M \to \mathbb{R}^m$, meaning that:

$$f^* g_{\text{std}} = g, \tag{1.2}$$

with g_{std} being the standard metric on $\mathbb{R}^m$. The following approach is taken: first, show that for small enough error $h^* g_{\text{std}} - g$, one can perturb h to find f that satisfies the isometry condition above. Then show that one can indeed construct some approximating embedding $h : M \to \mathbb{R}^m$ such that the error is arbitrarily small.

The perturbation process has been adapted and generalized by many, particulary Moser [Mos61], into what are now known as Nash-Moser theorems. They rely on the fact that the linearization of Equation (1.2) can be solved in a neighborhood (with respect to an appropriate functional space) of f. An abstract description of the setting is as follows: suppose we have spaces of sections $\Gamma(E)$, $\Gamma(E')$ of bundles E, E' over M and a *differential operator* $D : \Gamma(E) \to \Gamma(E')$. With the right setup, we can treat the spaces of sections $\Gamma(E)$ and $\Gamma(E')$ as manifolds. Consequently, we can define their tangent bundles $T\Gamma(E)$ and $T\Gamma(E')$ as well as a tangent map $TD : T\Gamma(E) \to T\Gamma(E')$. Then, informally, Nash-Moser says the following:

Theorem 1.4.1 (Sketch of a Nash-Moser theorem)**.**
Suppose that we have a right inverse to the map:

$$T_f D : T_f \Gamma(E) \to T_{D(f)} \Gamma(E')$$

[1] By the rank of a k-from θ, we mean the rank of the interior multiplication $X \mapsto \iota_X \theta$.

[2] For $H \subseteq TM$ a tangent distribution, a map $f : N \to TM$ is said to be horizontal if $\operatorname{Im} T_p f \subseteq H_{f(p)}$

[3] The requirement on m is shown to be $m \geq n(3n + 11)/2$ for compact manifolds and $m \geq n(n + 1)(3n + 11)/2$ for noncompact manifolds.

for f in some neighborhood U of $f_0 \in \Gamma(E)$. Then we have a local right inverse to the map:

$$D|_U : U \subseteq \Gamma(E) \to V := D(U) \subseteq \Gamma(E'),$$

that is: a map $R : V \to U$ such that $D(R(g)) = g$ for all $g \in V$.

In the context of the isometric immersion problem, take $\Gamma(E) = C^\infty(M, \mathbb{R}^m)$ and $\Gamma(E')$ the space of symmetric covariant two-tensors and note that $D : f \mapsto f^* g_{\mathrm{std}}$ is a differential operator. If we find a neighborhood V and map h such that $D(h) \in V$ and $g \in V$, then $f = R(g)$ is isometric.

Gromov states his version of Theorem 1.4.1 ([Gro86], pp.117-118) specifically for spaces of sections (as above), whereas Moser ([Mos61]) and Hamilton ([Ham82], pp. 171-172) replace these with suitable abstract functional spaces. Both approaches have their advantages. Gromov's specificity gives it great power for applications to section spaces. Hamilton's theorem on the other hand can be stated quite succinctly, and appears very similar to the finite-dimensional case, once the required background is established. However, this does come at the price of abstracting away certain properties, notably *locality* (the property that an operator only depends on the local behavior of a section), which is necessary for Gromov's argument to obtain microflexibility.

Besides stating Gromov's Nash-Moser theorem we develop the theory needed to state Hamilton's. In particular we rigorously define manifold structures on section spaces over manifolds, which hopefully allows the reader to transplant some intuition of finite-dimensional manifolds into this field. It should also serve to illustrate some constructions that are valid for non-compact manifolds more concretely.

1.5 Overview

1.5.1 Part I

This thesis consists of two parts. In the first part we develop the abstract theory of Fréchet spaces and manifolds following [Ham82]. In addition to this, we provide a detailed construction of the appropriate structures on spaces of sections. This will occur in the final section of each chapter. We conclude this part of the thesis with Hamilton's Nash-Moser theorem, stated abstractly for *Fréchet spaces*.

We start off by defining these spaces and establishing some basic calculus results in Chapter 2. As an example of these, we consider the space of sections of a vector bundle.

With the theory of Chapter 2 we define Fréchet manifolds in Chapter 3 and treat some basic theory regarding these such as the definition of tangent bundles and tangent maps. Building on the example of sections of vector bundles, we show that sections of fiber bundles form a Fréchet manifold. We investigate its tangent bundle and some particular maps between section spaces.

In preparation of stating Hamilton's Nash-Moser theorem we treat his notion of *tameness* in Chapter 4, which in the subsequent section Chapter 5 we state and prove. The proof of Hamilton's theorem serves to give the reader an idea of the use of *smoothing operators* which are a common element in the proof of all Nash-Moser theorems.

1.5.2 Part II

The second part is mostly dedicated to introducing the h-principle, holonomic approximation and the necessary preliminaries. We lay the foundation of the theory by introducing jet spaces in Chapter 6, where we also introduce another means of topologizing section spaces over non-compact manifolds. We show that in some cases, these topologies coincide with the manifold topology defined before.

The notion of jet space is fundamental to the general definition of PDRs (also simply called *differential relations*), which are treated in Chapter 7, where we also define the h-principle.

Continuing this line in Chapter 8, we look at holonomic ($\mathcal{R}$-)approximation, Eliashberg and Mishachev's versatile h-principle tool, introducing the requisite notions of *local integrability* and *microflexibility* along the way.

The general theory of (differential) operators on spaces of sections reaches its conclusion in Chapter 9. Here we use jets to define differential operators in full generality. The theory of Part I makes a reappearance, as we show that differential operators defined in this way fit in the framework of tame Fréchet spaces. We subsequently state

Gromov's Nash-Moser theorem, and use it to show that we may apply holonomic approximation to differential equations. Under the right conditions, this will also allow us to state an h-principle.

Finally, in Chapter 10 we use the developed theory to consider the isometric immersion problem and provide an outlook for further research.

Part I

Fréchet spaces and global analysis

Chapter 2

Calculus in Fréchet Spaces

In this chapter we define Fréchet spaces and develop the theory to do basic calculus on them. Fréchet spaces are a kind of *topological vector space* that is more general than the well-known Banach spaces. Many constructions that work for Banach spaces, notably the derivative of maps between spaces and the definition of the integral, also work for Fréchet space. With these definitions we will be able to define a notion of smoothness. Just as in the finite-dimensional case this allows us to define smoothly compatible charts and therefore smooth (Fréchet) manifolds, which we will do in Chapter 3.

The main example of a Fréchet space to keep in mind is the space of smooth functions $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$. After developing the general theory we will treat this space, as well as its geometric generalizations, in Section 2.5.

The development of the introductory material and basic definitions in this chapter follows [CB17], whereas the construction of the calculus follows Hamilton's paper [Ham82].

2.1 Fréchet spaces

In this section we provide an overview of the basics of Fréchet spaces. As with all the vector spaces in this thesis, we will assume it to be over $\mathbb{R}$. First of we recall the most general notion of a topologized vector space whose topology "interacts nicely" with the vector space structure.

Definition 2.1.1. *A* ***topological vector space*** *(TVS) F is a vector space F together with a topology $\mathcal{T}$ on F such that scalar multiplication:*

$$(\lambda, f) \mapsto \lambda f, \quad \mathbb{R} \times F \to F$$

and addition:

$$(f, g) \mapsto f + g, \quad F \times F \to F$$

are continuous.

The topology on Banach spaces is induced by a norm, although such a norm is not unique. Fréchet spaces, being more general, are instead induced by families of *seminorms*, which are similar but are not required to be nondegenerate. For completeness we state the definition:

Definition 2.1.2. *A* ***seminorm*** *on a vector space F is a function $\|\ldots\| : F \to [0, \infty)$ such that:*

1. *$\|f\| \geq 0$ for all $f \in F$;*
2. *$\|f + g\| \leq \|f\| + \|g\|$;*
3. *$\|cf\| = |c| \, \|f\|$ for all scalars c and $f \in F$.*

Note that a seminorm is a norm whenever $\|f\| = 0$ if and only if $f = 0$. A collection $P = \{\|\ldots\|_i\}$ of seminorms on F induces a TVS structure on F in the following way: for seminorms $\|\ldots\|_{i_k}$ with $1 \leq k \leq n$, define:

$$B^\epsilon_{i_1,\ldots,i_n}(0) = \left\{ f \in F \mid \|f\|_{i_k} < \epsilon \quad \forall 1 \leq k \leq n \right\}.$$

Then we define $\mathcal{B}_P(0) = \{B^{\epsilon}_{i_1,\ldots,i_n}\}$ to be the basis around zero of the topology induced by $\{\|\ldots\|_i\}$. Since in a TVS translation by a vector is a homeomorphism by requirement, the topology is determined by this basis of neighborhoods at 0.

Definition 2.1.3. *Let F be a TVS. For $f \in F$, we call a set of the kind $B^{\epsilon}_{i_1,\ldots,i_n}(f)$ as defined above a **seminorm ball**.*

Definition 2.1.4. *We call a TVS whose topology is induced by a family of seminorms a **locally convex vector space** or **LCVS** for short.*

Note that a sequence $\{f_j\}$ in an LCVS converges to $\{f\}$ if and only if $\|f_j - f\| \to 0$ for any seminorm $\|\ldots\|$.

Remark 2.1.5. An equivalent definition for locally convex vector spaces is a TVS such that zero (hence any point, since it is a TVS) has a basis of convex neighborhoods, hence the name "locally convex". From these convex neighborhoods, one can define seminorms inducing an equivalent basis as described above. Conversely: for a collection of seminorms P, each seminorm ball $B^{\epsilon}_{i_1,\ldots,i_n}$ of $\mathcal{B}_P(0)$ is convex. △

The collection of seminorms defining the topology is not unique: suppose we have a collection P of seminorms on F and a subcollection $P_0 \subseteq P$. If P_0 has the property that for any $\|\ldots\| \in P$ we have a $\|\ldots\|_0 \in P_0$ with $\|\ldots\|_0 \geq \|\ldots\|$, then P_0 also defines a neighborhood basis for the topology induced by P, and the topologies are therefore equivalent.

Suppose we have a linear map $A : F \to G$ where F and G are LCVSs whose topologies are induced by collections of seminorms P and Q respectively. Since A is linear and F, G are TVS's, continuity of A is determined by continuity around zero, meaning that for any ball $B' \in \mathcal{B}_Q(0)$ we must have $A^{-1}(B') \in \mathcal{B}_P(0)$. A more useful form of writing this is: for every $\|\ldots\|_q \in Q$ there are $\|\ldots\|_k \in P$, with $1 \leq k \leq n$, and a constant C_q (depending on q) such that:

$$\|A(f)\|_q \leq C_q \max_{1\leq k\leq n} \{\|f\|_k\}$$

for every $f \in F$.

Fréchet spaces are then a special type of LCVS space:

Definition 2.1.6. *A **Fréchet space** is a complete Hausdorff LCVS whose topology is induced by a countable collection of seminorms.*

Remark 2.1.7. Some of the defining properties of a Fréchet space F can be stated more explicitly:

- Hausdorffness can be expressed in terms of seminorms as: for every nonzero $f \in F$, there is a seminorm $\|\ldots\|$ such that $\|f\| \neq 0$.
- The property that the collection of seminorms can be taken to be countable is equivalent to F being metrizable. For given a collection of seminorms $\{\|\ldots\|_n\}$, we can define $d : F \times F \to \mathbb{R}$ by:

  $$d(f,g) = \sum_{n\in\mathbb{N}} 2^{-n} \frac{\|f-g\|_n}{1 + \|f-g\|_n},$$

  which can be shown to satisfy the properties of a metric. Conversely, if the topology is metrizable the collections of seminorms can be taken to be countable. In particular, continuity of maps between Fréchet spaces is determined by convergence of sequences rather than nets.
- Completeness means that any Cauchy sequence converges. A Cauchy sequence $\{f_k\}$ in a LCVS can be characterised by the property that for any $\epsilon > 0$ and seminorm $\|\ldots\|$ there is an M such that for all $n, m > M$ we have: $\|f_n - f_m\| < \epsilon$. Note that this is equivalent to F being complete with respect to the metric defined above.

△

The following construction of a Fréchet space given another Fréchet space F will be used in the upcoming section:

Definition 2.1.8. *Let $a < b \in \mathbb{R}$ and F a Fréchet space. Then we define $\mathcal{C}([a,b], F)$ to be the space of continuous functions $f : [a,b] \to F$, where we take the seminorms to be:*

$$\|f\|'_n = \sup_t \|f(t)\|_n,$$

with $\{\|\ldots\|_n\}$ the seminorms of F.

Showing that $\mathcal{C}([a,b],F)$ is a Fréchet space is done in much the same way as showing that $\mathcal{C}([a,b],\mathbb{R})$ is a Banach space.

2.2 Integrals and derivatives of curves

We preface this section by reminding the reader of the Hahn-Banach theorem, stated below, which holds for general vector spaces and therefore in particular for Fréchet spaces. It will be used in proofs throughout this section. Recall that a *sublinear functional* is a function $p : V \to \mathbb{R}$ (where V is a vector space) such that:

- $p(v+w) \leq p(v) + p(w)$ for all $v, w \in V$;
- $p(\lambda v) = \lambda p(v)$ for all $v \in V$ and $\lambda \geq 0$.

Theorem 2.2.1 (Hahn-Banach)**.** *Let F be vector space and $G \subseteq F$ a subspace. Let $l : G \to \mathbb{R}$ be a linear functional on G and $p : F \to \mathbb{R}_{\geq 0}$ a sublinear functional such that $l(g) \leq p(g)$ for all $g \in G$. Then there is a linear functional $L : F \to \mathbb{R}$ extending l such that $L(f) \leq p(f)$ for all $f \in F$.*

Its proof can be found in [Con89], pp. 77-82, or almost any other functional analysis textbook. This result will often be used in the following way:

Corollary 2.2.2. *Let F be a Fréchet space and $f \in F$ an element of F. Then $f = 0$ if and only $l(f) = 0$ for any continuous linear functional $l : F \to \mathbb{R}$.*

Proof. If $f = 0$ then clearly $l(f) = 0$ for any linear functional. For the converse, take any nonzero $f \in F$. On the subspace $\mathbb{R}f \subseteq F$, we can define the functional $l_f : \lambda f \mapsto \lambda$. Since F is Hausdorff, there is a seminorm $\|\ldots\|_f$ such that $\|f\|_f \neq 0$, which is a norm on the subspace $\mathbb{R}f$. Then l_f is continuous on $\mathbb{R}f$ with this norm, which means $|l_f(f)| \leq C \|f\|_f$ for some constant $C > 0$. Since a seminorm is a sublinear functional, we can extend l_f to a continuous $l : F \to \mathbb{R}$ and by definition $l(f) = 1 \neq 0$. ☐

In Theorem 2.2.5, we will show that there is an appropriate notion of an integral on Fréchet spaces. In preparation for this result we define the following spaces and state a lemma regarding them.

Definition 2.2.3. *We call a function $f : [a,b] \to F$ **linear** if $f(t) = f_1 + f_2 t$ with $f_1, f_2 \in F$ constant and we call it **piecewise linear** if there is a partition $a = a_0 < a_1, \ldots, a_k = b$ such that f is linear on each subinterval $[a_i, a_{i+1}]$.*

We denote by $\mathcal{L}([a,b],F)$ the space of piecewise linear curves $[a,b] \to F$.

Lemma 2.2.4. *The space $\mathcal{L}([a,b],F)$ of piecewise linear curves is dense in $\mathcal{C}([a,b],F)$ (see Definition 2.1.8).*

Proof. We have to show that any ball $B^{\epsilon}_{j_1,\ldots,j_l}(f)$ around $f \in \mathcal{C}([a,b],F)$ has nonempty intersection with $\mathcal{L}([a,b],F)$. Let therefore $\epsilon > 0$ and seminorms $\|\ldots\|_{j_i}$ be given. Since F is locally convex, we can cover $f([a,b])$ by balls of the form $B^{\epsilon}_{j_1,\ldots,j_l}(f(t))$, with $t \in [a,b]$. Since $f([a,b])$ is compact, we take finitely many of such balls B_i, from which we obtain a finite open cover $\{f^{-1}(B_i)\}$ of $[a,b]$, which we may take (by passing to connected components and taking a minmal cover) to be intervals (c_i, d_i) such that $(c_i, d_i) \cap (c_{i+2}, d_{i+2}) = \emptyset$. Pick points a_i such that $c_{i+1} < a_i < d_i$ (with the first and last equaling a and b), and define $g(t) = f(a_i) + t(f(a_{i+1}) - f(a_i))$ on each subinterval $[a_i, a_{i+1}]$. This is well-defined and g is a piecewise linear function such that $\sup_t \|(g-f)(t)\|_{j_i} < 2\epsilon$ for every seminorm $\|\ldots\|_{j_i}$ defining the ball $B^{2\epsilon}_{j_1,\ldots,j_l}(f(t)) \subseteq F$, hence $g \in B^{2\epsilon}_{j_1,\ldots,j_l}(f) \subseteq \mathcal{C}([a,b],F)$. ☐

With this lemma we are ready to show the following. Note that the construction of the integral in Fréchet space mimics the definition of the Riemann integral.

Theorem 2.2.5. *For $a, b \in \mathbb{R}$ there is a unique element $\int_a^b f(t)dt \in F$ such that:*

1. *For every continuous linear functional $l : F \to \mathbb{R}$ we have:*

$$l\left(\int_a^b f(t)dt\right) = \int_a^b l(f(t))dt;$$

2. *For every continuous seminorm* $\|\ldots\|$ *we have:*

$$\left\| \int_a^b f(t)dt \right\| \leq \int_a^b \|f(t)\| \, dt;$$

3. *For* $c \in \mathbb{R}$ *we have:*

$$\int_a^b f(t)dt + \int_b^c f(t)dt = \int_a^c f(t)dt;$$

4. *The assignment* $f \mapsto \int_a^b f(t)dt$ *is linear in* f.

Proof. We define the element $\int_a^b g(t)dt$ for a $g : [a, b] \to F$ piecewise linear on $[a_i, a_{i+1}]$ by:

$$\int_a^b g(t)dt = \sum_i \tfrac{1}{2}(a_{i+1} - a_i)(g_{i+1}(a_{i+1}) - g_i(a_i)).$$

This is a continuous functional $\mathcal{L}([a, b], F) \to F$ since:

$$\left\| \int_a^b g(t)dt \right\|_n \leq \tfrac{1}{2}(b - a) \, \|g\|_n' \, ,$$

and by density (Lemma 2.2.4) it extends to a continuous functional $\mathcal{C}([a, b], F) \to F$. This shows the existence of $\int_a^b f(t)dt$. For any continuous linear functional $l : F \to \mathbb{R}$ the functionals $\mathcal{C}([a, b], F) \to \mathbb{R}$ given by:

$$f \mapsto l\left(\int_a^b f(t)dt \right)$$

$$f \mapsto \int_a^b l(f(t))dt$$

are continuous and equal on $\mathcal{L}([a, b], F)$. Since this space is dense they are equal on $\mathcal{C}([a, b], F)$, which shows (1). Property (2) follows from density and the fact that it holds for piecewise linear functions, and so do (3) and (4). Uniqueness follows from (1) and Corollary 2.2.2. □

The following theorem shows how the integral on Fréchet space behaves under a reparametrization of the interval $[a, b]$. Unsurprisingly it behaves very much like the ordinary integral on $\mathbb{R}$.

Theorem 2.2.6. *Let* $\gamma : [a, b] \to [\gamma(a), \gamma(b)]$ *be a* C^1 *monotone function, then:*

$$\int_{\gamma(a)}^{\gamma(b)} f(u)du = \int_a^b f(\gamma(t))\gamma'(t)dt.$$

Proof. For any continuous functional $l : F \to \mathbb{R}$ we see that:

$$l\left(\int_{\gamma(a)}^{\gamma(b)} f(u)du \right) = \int_{\gamma(a)}^{\gamma(b)} (l \circ f)(u)du = \int_a^b (l \circ f)(\gamma(t))\gamma'(t)dt = l\left(\int_a^b f(\gamma(t))\gamma'(t)dt \right),$$

by applying the change of variables for $(l \circ f) : \mathbb{R} \to \mathbb{R}$. Equality follows from Corollary 2.2.2. □

Lemma 2.2.7. *For* X *an arbitrary topological space,* $[a, b] \subseteq \mathbb{R}$ *a compact interval and* $f : X \times [a, b] \to F$ *a continuous map into a Fréchet space, the map* $\tilde{f} : X \to \mathcal{C}([a, b], F)$ *defined by:*

$$x \mapsto f(x, -)$$

is continuous.

Proof. Let $\{\|\ldots\|_1, \ldots, \|\ldots\|_n\}$ be a finite set of seminorms of F. For $u \in F$ and $g \in \mathcal{C}([a,b],F)$ define:

$$B^\epsilon(u) = \{v \in F \mid \|v-u\|_i < \epsilon \text{ for all } 1 \le i \le n\}$$
$$D^\epsilon(g) = \{h \in \mathcal{C}([a,b],F) \mid \|h(t)-g(t)\|_i < \epsilon \text{ for all } 1 \le i \le n \text{ and } t \in [a,b]\}.$$

Fix $x \in X$. Then for any $t \in [a,b]$ there are opens U and I such that $(x,t) \in U \times I \subseteq f^{-1}(B^\epsilon(f(x,t)))$. Since $\{x\} \times [a,b]$ is compact, we can extract a finite cover $U_j \times I_j$ from this cover. Note that for any $y \in \cap_j U_j$ we have that $f(y,t) \in B^\epsilon(f(x,t))$. This means that $\cap_j U_j \subseteq \tilde{f}^{-1}(D^\epsilon)$, hence $\tilde{f}$ is continuous. ☐

Theorem 2.2.8. *For X an arbitrary topological space, and $f : X \times [a,b] \to F$ a continuous map into a Fréchet space, the map $g : X \to F$ defined by:*

$$g : x \mapsto \int_a^b f(x,t)dt$$

is also continuous.

Proof. By the lemma above and continuity of the integral on $\mathcal{C}([a,b],F)$ we see that the composition:

$$g : x \mapsto f(x,-) \mapsto \int_a^b f(x,t)dt$$

is continuous. ☐

Definition 2.2.9. *For $f \in \mathcal{C}([a,b],F)$ we define:*

$$f'(t) = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$$

if the limit exists. If the limit exists for all $t \in [a,b]$ and $t \mapsto f'(t)$ is continuous then we say that f is C^1.

We will prove analogues of the fundamental theorem of calculus for Fréchet spaces:

Theorem 2.2.10. *Let $f : [a,b] \to F$ be a C^1 curve, then:*

$$f(b) - f(a) = \int_a^b f'(t)dt.$$

Proof. Note that for any continuous linear functional l we have that: $(l \circ f)' = l \circ f'$. Then by the fundamental theorem of calculus for functions $\mathbb{R} \to \mathbb{R}$, we have:

$$l(f(a) - f(b)) = (l \circ f)(b) - (l \circ f)(a) = \int_a^b (l \circ f)'(t)dt = l\left(\int_a^b f'(t)dt\right),$$

from which the result once again follows by Corollary 2.2.2. ☐

The converse also holds:

Theorem 2.2.11. *Let $f \in \mathcal{C}([a,b],F)$ and define:*

$$g(\theta) = \int_a^\theta f(t)dt.$$

Then g is a C^1 curve and $g'(t) = f(t)$.

Proof. The following change of variables formula:

$$\int_t^{t+h} f(\theta)d\theta = h\int_0^1 f(t+hu)du$$

holds for curves in a Fréchet space F by Theorem 2.2.6. Then we see that $(g(t+h)-g(t))/h = \int_0^1 f(t+hu)du$. Since $(h,u) \mapsto f(t+hu)$ is continuous for fixed $t \in \mathbb{R}$, so is the map $h \mapsto \int_0^1 f(t+hu)du$ by Theorem 2.2.8, with $0 \mapsto f(t)$. Therefore we have $\lim_{h\to 0}(g(t+h)-g(t))/h = f(t)$. Since this exists for every t for which f (and therefore g) is defined, this also shows that g is a C^1 curve. ☐

The following corollaries are immediate from the above:

Corollary 2.2.12. *For $f, g \in \mathcal{C}([a,b], F)$:*

- *if $f(a) = g(a)$, the curves f, g are both C^1 and $f'(t) = g'(t)$ on $[a,b] \subseteq \mathbb{R}$, then $f(b) = g(b)$;*
- *if f is C^1 on $[a,b] \subseteq \mathbb{R}$ and $\|f'(t)\| \leq K$ for some seminorm $\|\ldots\|$ and constant K, then:*

$$\|f(b) - f(a)\| \leq K(b-a).$$

2.3 Derivatives of maps between Fréchet spaces

In the following, we will let $P : U \subseteq F \to G$ be a map between an open subset U of a Fréchet space F into another Fréchet space G.

Definition 2.3.1. *The **derivative** of $P : U \subseteq F \to G$ at a point $f \in U$ in the direction of $h \in F$ is defined by:*

$$DP(f)h = \lim_{t \to 0} \frac{P(f+th) - P(f)}{t}.$$

*We call P **directionally differentiable at** f if this limit exists for any $h \in F$. We call P a C^1 (or **continuously differentiable**) map if P is differentiable at any $f \in U$ and the map $(f,h) \mapsto DP(f)h$ is continuous.*

The rest of this section is devoted to proving some results, analogous to those in the finite dimensional case, for the derivative in Fréchet space and in particular with its interaction with the integral.

Lemma 2.3.2. *Let $P : U \subseteq F \to G$ a C^1 map and define the curve $p : t \mapsto P(f+th)$ for fixed $f \in U, h \in F$. Then p is a C^1 curve and:*

$$p'(t) = P(f+th)' = DP(f+th)h.$$

Proof. This follows by writing out the definitions of the derivative for both curves and maps. Note that $p'(t)$ is continuous because of the continuity of $(f,h) \mapsto DP(f)h$. □

Theorem 2.3.3. *Let $P : U \subseteq F \to G$ a C^1 map and $f \in U$, $h \in F$ such that the line from f to $f+h$ lies in U. Then we have:*

$$P(f+h) - P(f) = \int_0^1 DP(f+th)h dt$$

Proof. We consider the path $p : t \to P(f+th)$ as in Lemma 2.3.2. By Theorem 2.2.10 we have:

$$P(f+h) - P(f) = \int_0^1 P(f+th)'dt = \int_0^1 DP(f+th)h\, dt.$$

□

Note that the requirement that the line between f and $f+h$ lies in U is always satisfied for sufficiently small h: take a locally convex neighborhood V around f and pick $h \in V - f$. The preceding result, in combination with Theorem 2.2.8, will be a crucial tool in proofs in this section.

Theorem 2.3.4. *Let $P : U \subseteq F \to G$ be a C^1 map. Then for any $f \in U$ the map $h \mapsto DP(f)h$ is linear.*

Proof. First, we remark that preservation of scalar multiplication follows from a standard property of limits. Take $\lambda \neq 0$ (for $\lambda = 0$ the result is straightforward), then:

$$DP(f)\lambda h = \lim_{t \to 0} \frac{P(f+\lambda th) - P(f)}{t} = \lim_{s \to 0} \frac{\lambda(P(f+sh) - P(f)}{s} = \lambda DP(f)h$$

by taking $s = \lambda t$.
It remains to show that $DP(f)$ preserves addition. We first rewrite:

$$P(f + t(h_1+h_2)) - P(f) = (P(f+th_1+th_2) - P(f+th_1)) + (P(f+th_1) - P(f)).$$

Now we note that, since a Fréchet space is locally convex, we can always take t small enough to ensure that the path from f to $f + th_1$ and from $f + th_1$ to $f + th_1 + th_2$ is contained in U. Applying Theorem 2.3.3 and dividing by t, we get:

$$\begin{aligned}\frac{P(f + t(h_1 + h_2)) - P(f)}{t} &= \frac{1}{t}\left(\int_0^1 DP(f + th_1 + uth_2)th_2 du + \int_0^1 DP(f + uth_1)th_1)\right) \\ &= \int_0^1 DP(f + th_1 + uth_2)h_2 du + \int_0^1 DP(f + uth_1)h_1 du.\end{aligned}$$

By Theorem 2.2.8 and the C^1 property of P, we may take the limit $t \to 0$ "inside the integral", and we see that: $DP(f)(h_1 + h_2) = DP(f)h_1 + DP(f)h_2$. □

Another standard result for derivatives is the chain rule, for which we use the following lemma:

Lemma 2.3.5. *Let $P : U \subseteq F \to G$ with U convex. Then P is continuously differentiable if and only if there exists a continuous map $L : U \times U \times F$ that is linear in the last variable, denoted by: $(f_0, f_1, h) \mapsto L(f_0, f_1)h$, such that:*

$$P(f_1) - P(f_0) = L(f_0, f_1)(f_1 - f_0).$$

In this case $DP(f)h = L(f, f)h$.

Remark 2.3.6. Note that the convexity requirement isn't very strong since differentiability is a local property and by local convexity we can always restrict P to a convex open. △

Proof. First assume that P is C^1, and define:

$$L(f_0, f_1)h = \int_0^1 DP(f_0 + t(f_1 - f_0))h\, dt$$

(note the use of convexity of U here). By Theorem 2.3.3, we see that $L(f_0, f_1)(f_1 - f_0) = P(f_1) - P(f_0)$ and by Theorem 2.2.8 we see that L is continuous. Since DP is linear, so is L, and therefore L satisfies all the requirements in the statement of the result.

Conversely, suppose we have such an L. Then we see that:

$$\frac{P(f + th) - P(f)}{t} = L(f, f + th)h,$$

whenever $t \neq 0$. Since the right hand side is well defined and continuous even for $t = 0$, we see by taking the limit that $DP(f)h = L(f, f)h$, showing that P is C^1. □

Theorem 2.3.7 (Chain rule)**.** *If $P : U \subseteq F \to V \subseteq G$ and $Q : V \subseteq G \to H$ are C^1 then so is $Q \circ P : U \subseteq F \to H$ and:*

$$D(Q \circ P)(f)h = DQ(P(f))DP(f)h.$$

Proof. Note that we may take U and V smaller and convex if necessary. Then we take functions L and M as in Lemma 2.3.5, such that:

$$P(f_1) - P(f_0) = L(f_0, f_1)(f_1 - f_0), \quad Q(g_1) - Q(g_0) = M(g_0, g_1)(g_1 - g_0).$$

Defining: $N(f_0, f_1)h = M(P(f_0), P(f_1))L(f_0, f_1)h$ we see that:

$$N(f_0, f_1)(f_1 - f_0) = M(P(f_0), P(f_1))(P(f_1) - P(f_0)) = (Q \circ P)(f_1) - (Q \circ P)(f_0).$$

Furthermore N is linear in h and continuous (as composition of continuous maps). Therefore $(Q \circ P)$ is C^1 by Lemma 2.3.5 and its derivative equals:

$$N(f, f)h = M(P(f), P(f))L(f, f)h = DQ(P(f))DP(f)h.$$

□

2.4 Multivariate maps and partial derivatives

In addition to maps from a single Fréchet space into another, we will come across maps taking arguments in multiple Fréchet spaces. We have already seen one example: the derivative $DP : U \times F \to G$ of a C^1 map $P : U \subseteq F \to G$. We would like to treat these variables separately when taking derivatives.

Definition 2.4.1. *Let:*

$$P : U_1 \times \ldots \times U_n \subseteq F_1 \times \ldots \times F_n \to G$$

be a map taking arguments in multiple (opens in) Fréchet spaces. We define the ***i-th partial derivative*** $D_i P(f_1, \ldots, f_n)h$ *for* $h \in F_i$ *as:*

$$\begin{aligned} D_i P(f_1, \ldots, f_n)h &= D\left[f_i \mapsto P(f_1, \ldots, f_n)\right] h \\ &= \lim_{t \to 0} \frac{P(f_1, \ldots, f_{i-1}, f_i + th, f_{i+1}, \ldots, f_n) - P(f_1, \ldots, f_n)}{t}. \end{aligned}$$

Remark 2.4.2. If the map is denoted by $(f, g) \mapsto P(f, g)$, then we might also denote $D_1 P(f, g)$ by $D_f P(f, g)$ (and similar for $D_2 P(f, g)$). △

In the remainder of this section, we shall restrict our proofs to maps $P : U_1 \times U_2 \subseteq F_1 \times F_2 \to G$ of two variables, since the general proofs are conceptually the same but notationally more cumbersome.

The following result is proven in almost exactly the same way as Lemma 2.3.5:

Lemma 2.4.3. *The partial derivative $D_1 P(f, g)$ exists if and only if there is a map $L : U_1 \times U_1 \times U_2 \times F_1$ linear in the last argument such that: $P(f_1, g) - P(f_0, g) = L(f_0, f_1, g)(f_1 - f_0)$. In this case $D_1 P(f, g)h = L(f, f, g)h$.*

Theorem 2.4.4. *The partial derivatives $D_1 P$ and $D_2 P$ exist and are continuous if and only if P is C^1 (where we consider $P : U_1 \times U_2 \subseteq F_1 \times F_2 \to G$ as a map from a single Fréchet space into G).*

Proof. If the derivative:

$$DP(f, g)(h, k) = \lim_{t \to 0} \frac{P(f + th, g + tk) - P(f, g)}{t}$$

exists, then by taking $h = 0$ or $k = 0$ we obtain $D_1 P$ and $D_2 P$ respectively. Conversely, if the partial derivatives exist we find L_1 and L_2 as in Lemma 2.4.3 (after restricting to convex opens) such that:

$$P(f_1, g) - P(f_0, g) = L_1(f_0, f_1, g)(f_1 - f_0), \quad P(f, g_1) - P(f, g_0) = L_2(g_0, g_1, f)(g_1 - g_0).$$

Define:

$$L(f_0, g_0, f_1, g_1)(h, k) = L_1(f_0, f_1, g_1)h + L_2(g_0, g_1, f_0)k,$$

then:

$$L(f_0, g_0, f_1, g_1)(f_1 - f_0, g_1 - g_0) = P(f_1, g_1) - P(f_0, g_1) + P(f_0, g_1) - P(f_0, g_0) = P(f_1, g_1) - P(f_0, g_0).$$

Furthermore, L is linear in the last argument and continuous. Then by Lemma 2.3.5, P is C^1 with derivative:

$$L(f, f, g, g)(h, k) = L_1(f, f, g)h + L_2(g, g, f)k = D_1 P(f, g)h + D_2 P(f, g)k.$$

□

Remark 2.4.5. For a map L that is linear in multiple arguments, we separate these arguments into linear and (not necessarily) linear by a vertical delimiter. For example, if we have a map:

$$L : U_1 \times U_2 \times F_3 \times F_4 \to G$$

that is linear in the last two arguments, we denote it by:

$$(f_1, f_2, f_3, f_4) \mapsto L(f_1, f_2 \mid f_3, f_4).$$

When we wish to consider $L(f_1, f_2 \mid \ldots, \ldots)$ as a (multi)linear map in its own right, we will denote it by:

$$L_{f_1, f_2} = L(f_1, f_2 \mid \ldots, \ldots)$$

△

Note that the derivative of a linear function $L : F \to G$ is simply L itself:

$$DL(f)h = \lim_{t\to 0} \frac{L(f+th) - L(f)}{t} = \lim_{t\to 0} \frac{tL(h)}{t} = L(h).$$

Since differentiating with respect to the linear variables "gives no new information", we adopt the convention to disregard these when taking the total derivative. For example: the derivative of a family of linear maps $(f, h) \mapsto L(f)h$ is taken to be:

$$DL(f \mid h, k) = \lim_{t\to 0} \frac{L(f+tk)h - L(f)h}{t}$$

The following follows directly from linearity of derivatives:

Theorem 2.4.6. *If a map $L : U \subseteq F \times G \to H$ given by $(f, g) \mapsto L(f)g$ is C^1 then $DL : U \times G \times F \to H$ given by $(f, g, k) \mapsto DL(f \mid h, k)$ is linear in both h and k.*

Using this, we can inductively define notions of higher-order differentiability:

Definition 2.4.7. *We say that $P : U \subseteq F \to G$ is C^n if P is C^{n-1} and the limit:*

$$\lim_{t\to 0} \frac{D^{n-1}P(f + th_n \mid h_1, \dots h_{n-1}) - D^{n-1}P(f \mid h_1, \dots h_{n-1})}{t} =: D^n P(f \mid h_1, \dots, h_n)$$

exists and the map $(f, h_1, \dots h_n) \mapsto D^n P(f|h_1, \dots h_n)$ is continuous.

For example, P is C^2 if $(f, h) \mapsto DP(f)h$ is differentiable. It turns out that $D^2P(f \mid h, k)$ is not only bilinear but also symmetric. We first prove the following:

Theorem 2.4.8. *If $P : U \subseteq F \to G$ is C^2 then we have:*

$$D^2P(f \mid h, k) = \lim_{t,u\to 0} \frac{P(f + th + uk) - P(f + th) - P(f + uk) + P(f)}{tu}.$$

Proof. By Theorem 2.3.3 we have:

$$\begin{aligned}
\frac{P(f+th) - P(f)}{t} &= \int_0^1 DP(f + \theta th)h \, d\theta \\
\frac{P(f+uk+th) - P(f+uk)}{t} &= \int_0^1 DP(f + uk + \theta th)h \, d\theta \\
\frac{DP(f + \theta th + uk)h - DP(f + \theta th)h}{u} &= \int_0^1 D^2P(f + \theta th + \eta uk \mid h, k) d\eta.
\end{aligned}$$

Subtracting the first two equalities, dividing by u and integrating, we obtain:

$$\frac{P(f + th + uk) - P(f + th) - P(f + uk) + P(f)}{tu} = \int_{\theta=0}^1 \int_{\eta=0}^1 P(f + \theta th + \eta uk \mid h, k) d\eta d\theta.$$

As in earlier proofs in this section, we take $u, t \to 0$ and use Theorem 2.2.8 ☐

Corollary 2.4.9. *If P is C^2 then the second derivative is symmetric, i.e.:*

$$D^2P(f \mid h, k) = D^2P(f \mid k, h).$$

Proof. This follows from Theorem 2.4.8 of the statement in the result above. ☐

This symmetry is not only present in the second derivative:

Theorem 2.4.10. *If P is C^n then, $D^nP(f \mid h_1, \dots, h_n)$ is completely symmetric, that is:*

$$D^nP(f \mid h_1, \dots, h_n) = D^nP(f \mid h_{\sigma(1)}, \dots, h_{\sigma(n)}),$$

for any permutation σ of $\{1, \dots, n\}$.

Proof. Assume $n \geq 2$ (since for $n = 2$ the result is shown), and assume that the result is true for n. We proceed by induction. For a C^{n+1} map we know that $DP(f)\{h_1, \ldots, h_n, h_{n+1}\}$ is symmetric in $h_1, \ldots h_n$, and by Corollary 2.4.9 it is symmetric in h_n and h_{n+1}. Since the permutations of $\{n, n+1\} \subseteq \{1, \ldots, n+1\}$ and $\{1, \ldots, n\} \subseteq \{1, \ldots, n+1\}$ generate the permutations of $\{1, n+1\}$ (as shown in a group theory course), the result follows. □

For certain situations, it is more useful to use the following "version" of the derivative:

Definition 2.4.11. *For a C^1 map $P : U \subseteq F \to G$ we define the **tangent map** as:*

$$TP : (f, h) \mapsto (P(f), DP(f)h), \quad U \times F \to G \times G$$

For higher orders, define $T^n P$ as $T(T^{n-1}P)$.

Note that $T^n P$ is defined and continuous if and only if P is C^n. Furthermore, we see that:

$$(TQ \circ TP)(f, h) = ((Q \circ P)(f), DQ(P(f))DP(f)h) = T(Q \circ P)(f, h),$$

by the chain rule. By induction, we get the following result:

Theorem 2.4.12. *For P and Q of class C^n, their composition is also C^n and:*

$$T^n(Q \circ P) = (T^n Q) \circ (T^n P).$$

Definition 2.4.13. *We say that a map $P : U \subseteq F \to G$ is **smooth** if P is C^n for all $n \in \mathbb{N}$. We call P a **diffeomorphism** if $P : U \to P(U)$ is smooth, bijective and has smooth inverse.*

Finally, we have the following version of Taylor's theorem that we will use later.

Theorem 2.4.14. *If $P : U \subseteq F \to G$ is C^2 and U convex then:*

$$P(f + h) = P(f) + DP(f)h + \int_0^1 (1 - t)D^2P(f + th \mid h, h)dt.$$

Proof. Some rewriting gives:

$$DP(f + th)h = (1 - t)D^2P(f + h \mid h, h) - ((1 - t)DP(f + th)h)',$$

and integrating this over $[0, 1]$ gives us:

$$P(f + h) - P(f) = \int_0^1 DP(f + th)hdt = \int_0^1 (1 - t)D^2P(f + th \mid h, h)dt + DP(f)h.$$

□

2.5 Vector bundle sections and other examples

In this section we will cease to develop the abstract theory on Fréchet spaces and look to apply it to concrete examples instead. We examine the spaces of smooth functions $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^m)$, which turn out to be Fréchet spaces, in depth. We then look at an important kind of map between these spaces, which arise naturally from maps on (subsets of) the underlying Euclidean space. These results for the space of smooth functions will then be carried over, in the classic differential geometry way, to sections of vector bundle sections. In Chapter 3, specifically Section 3.3, we will see that the space of sections of fiber bundles is a Fréchet manifold that is "locally modeled" on such spaces (just as manifolds are modeled on $\mathbb{R}^n$).

We start with the following intuitive example of a Fréchet space:

Example 2.5.1. Let $\mathbb{R}^{\mathbb{N}}$ be the space of functions $f : \mathbb{N} \to \mathbb{R}$ with seminorms:

$$\|f\|_k = \sup_{1 \leq i \leq k} |f(i)|.$$

for $k \in \mathbb{N}$. To see that this is indeed a Fréchet space, note that the collection of seminorms is obviously countable and that for any $f \neq g \in \mathbb{R}^{\mathbb{N}}$ we can take k such that $\|f - g\|_k > 0$. Furthermore, since the seminorm $\|\ldots\|_k$ is a complete norm on the subspace:

$$\mathbb{R}^k \cong \{f \mid f(i) = 0 \text{ for } i > k\} \subseteq \mathbb{R}^\infty,$$

the space is complete as a Fréchet space. △

2.5.1 The space of smooth functions

For the upcoming results we make heavy use of the results and notation introduced in Appendix A. Just as in this appendix, we will denote:

$$D_x^k f(h_1, \dots, h_k) = D^k f(x \mid h_1, \dots, h_k)$$

for the derivative of a map $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ (as defined in Definition 2.4.7) and $x \in U$.

Note that $D_x^k f \in S^k(\mathbb{R}^n, \mathbb{R}^m)$ (see Proposition A.0.5), and the latter space inherits an inner product $\langle \dots, \dots \rangle$ and consequently a norm $|\dots|$ from Euclidean space (see Remark A.0.4). The following kinds of Fréchet space are the local case of the kind that will be of main importance to us.

Definition 2.5.2. *For a compact set $K \subseteq \mathbb{R}^n$, we denote $\mathcal{C}^\infty(K, \mathbb{R}^m)$ for the space of smooth functions $K \to \mathbb{R}^m$, with the topology endowed by the following seminorms:*

$$\|f\|_k = \sup_{x \in K} \sum_{i=1}^{k} \left| D_x^i f \right|.$$

For an open $U \subseteq \mathbb{R}^n$, we again denote $\mathcal{C}^\infty(U, \mathbb{R}^m)$ for the space of smooth functions. The topology is taken to be the one induced by:

$$\|f\|_{K_i,k} = \sup_{x \in K_i} \sum_{i=1}^{k} \left| D_x^i f \right|,$$

where $\{K_i \mid i \in \mathbb{N}\}$ is a compact exhaustion of U.

As usual, the space of seminorms inducing the topology is not unique. For example, the same topology on $\mathcal{C}^\infty(K, \mathbb{R}^m)$ is induced by the seminorms:

$$\|f\|'_k = \sup_{x \in K} \sup_{1 \le i \le k} \left| D_x^i f \right|,$$

since we have $\|f\|'_k \le \|f\|_k \le k \|f\|'_k$, which implies that convergence with respect to $\|\dots\|_k$ is equivalent to convergence with respect to $\|\dots\|'_k$. We now show that the spaces defined above are in fact Fréchet spaces, starting with the compact case.

Proposition 2.5.3. *For $K \subseteq \mathbb{R}^n$ compact, the space $\mathcal{C}^\infty(K, \mathbb{R}^m)$ is a Fréchet space.*

Proof. The collection of seminorms is countable and since they are all norms the Hausdorff property (as in Remark 2.1.7) is satisfied. Furthermore, a classic result in analysis (see, for example, [Rud+64], Theorem 7.17) says that a uniformly convergent sequence of functions $f_j \to f$ with uniformly convergent derivatives Df_j satisfies $\lim_{j \to \infty} Df_j = Df$. Repeating this inductively in the order of the derivative shows that $\mathcal{C}^\infty(K)$ is complete. ☐

Corollary 2.5.4. *The space $\mathcal{C}^\infty(U, \mathbb{R}^m)$ of smooth functions on an open subset $U \subseteq \mathbb{R}^n$ is a Fréchet space.*

Proof. The collection of seminorms $\|\dots\|_{K_i,k}$ defined in Definition 2.5.2 remains coutable. For the Hausdorff property, remark that if $f(x) \neq 0$ for some $x \in U$, then $x \in K_i$ for some compact of the exhaustion and $\|f\|_{K_i,0} \neq 0$. Completeness follows from the uniform convergence on compact subsets.. ☐

We turn to give an example of a smooth map between the spaces defined above. For this discussion we will need to define the following kind of open subsets:

Definition 2.5.5. *Let $V \subseteq K \times \mathbb{R}^n$ be an open subset, then define:*

$$\Gamma(V) := \{f \in \mathcal{C}^\infty(K, \mathbb{R}^m) \mid \mathrm{Graph}(f) \subseteq V\}.$$

For $V \subseteq U \times \mathbb{R}^n$ and $K \subseteq U$ compact, we define:

$$\Gamma(V|_K) := \{f \in \mathcal{C}^\infty(U, \mathbb{R}^m) \mid \mathrm{Graph}(f|_K) \subseteq V\}.$$

Observe that $\Gamma(V)$ and $\Gamma(V|_K)$ are indeed open in their respective function spaces. Now let $V \subseteq K \times \mathbb{R}^m$ open and let $\phi : V \to \mathbb{R}^m$ be a smooth map. Then define the following map $\Gamma(V) \subseteq \mathcal{C}^\infty(K, \mathbb{R}^m) \to \mathcal{C}^\infty(K, \mathbb{R}^m)$:

$$(\phi_*)(f)(x) = \phi(x, f(x)). \tag{2.1}$$

Proposition 2.5.6. *For $\phi : V \subseteq K \times \mathbb{R}^n \to \mathbb{R}^m$ smooth, the map $\phi_* : \Gamma(V) \to \mathcal{C}^\infty(K, \mathbb{R}^m)$ defined in Equation (2.1) is smooth.*

Proof. The derivative of ϕ_* is given by:

$$D_f(\phi_*)(h)(x) = \lim_{t \to 0} \frac{\phi(x, f(x) + th(x)) - \phi(x, f(x))}{t} = (D_2\phi)(x, f(x) \mid h(x))$$

The convergence of this limit can be shown to be uniform over K. Now note that for $\psi : V \times \mathbb{R}^m \to \mathbb{R}^m$ given by $\psi(x, y, z) = (D_2\phi)(x, y \mid z)$ we have:

$$D_f(\phi_*)(h) = \psi_*,$$

hence the derivative of ϕ_* is again induced by a smooth map of the same kind as ϕ, only now defined on an open subset of $K \times \mathbb{R}^{2m}$. By induction, ϕ is infinitely many times differentiable. ☐

By the same reasoning we get the following:

Corollary 2.5.7. *Let $U \subseteq \mathbb{R}^n$ be open, $K \subseteq U$ compact and let $\phi : V \subseteq U \times \mathbb{R}^n \to \mathbb{R}^m$ be a smooth map defined on the open subset V. Then the map $\phi_* : \Gamma(V|_K) \to \mathcal{C}^\infty(U, \mathbb{R}^m)$ defined by Equation (2.1) is smooth.*

2.5.2 Sections of a vector bundle

The space of smooth functions serves as the "local case" for the space of sections of a vector bundle. Indeed, smooth maps $K \to \mathbb{R}^m$ can be considered to be sections of the bundle $K \times \mathbb{R}^m \twoheadrightarrow K$. These spaces will also serve to examine and prove results for the "global case". First we introduce the requisite notation and recall some basic definitions. The first of these is the space for which we will define the Fréchet structure:

Definition 2.5.8. *Given a vector bundle $E \xrightarrow{\pi} M$ over M, we define the* ***space of sections*** *$\Gamma(M, E)$ to be the set of maps $s : M \to E$ such that $\pi \circ s = Id_M$. When the base manifold is understood, we simply write $\Gamma(E)$.*

Given two vector bundles E and E' over the same base manifold M, we will make use of maps $E \to E'$ "preserving basepoints" in M. We define these more generally for open subsets of E.

Definition 2.5.9. *Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M$ be vector bundles over M and $U \subseteq E$ be open. We call a map $\phi : E \to E'$ such that $\pi' \circ \phi = \pi$ a* ***(vector) bundle map****.*

In order to generalize Definition 2.5.2 to spaces of sections, we need a notion of "size" for the vector bundle. This brings us to the final basic definition we recall:

Definition 2.5.10. *Let $E \twoheadrightarrow M$ be a vector bundle over M. Then a* ***(Riemannian) fiber metric*** *is a smooth section g of the bundle $E^* \otimes E^* \twoheadrightarrow M$ such that:*

$$g_p : E_p \otimes E_p \to \mathbb{R}$$

is symmetric and positive-definite.

That is: g is a set of inner products on E_p that varies smoothly with respect to the basepoint $p \in M$. Alternatively, a fiber metric could have been defined as a symmetric, positive-definite $\mathcal{C}^\infty(M)$-bilinear map $\Gamma(E) \times \Gamma(E) \to \mathcal{C}^\infty(M)$.

Now suppose we have a fiber metric on $E \twoheadrightarrow M$ as well as a connection ∇^E on E and a connection ∇^{TM} on TM. Recall that for a section $s \in \Gamma(E)$ the map $\Gamma(TM) \to \Gamma(E)$ given by:

$$\nabla^1 s : X \mapsto \nabla^E_X s$$

is $\mathcal{C}^\infty(M)$ linear and therefore may be seen as a section of $\mathrm{Hom}(TM, E) \cong T^*M \otimes E$. The connections ∇^E and ∇^{TM} induce a connection $\nabla^{T^*M \otimes E}$ on this bundle, therefore applying this connection to $\nabla^1 s$ gives a $\mathcal{C}^\infty(M)$-linear map $\Gamma(TM) \to \Gamma(T^*M, E)$. Such a map is equivalent to a section:

$$\mathrm{Hom}(TM, T^*M \otimes E) \cong T^*M \otimes T^*M \otimes E,$$

which we will call $\nabla^2 s$. As a map $\Gamma(TM) \times \Gamma(TM) \to \Gamma(E)$ this is given by:

$$(\nabla^2 s)(X, Y) = \nabla_X^{T^*M \otimes E} \left(\nabla_Y^E s \right).$$

Repeating this process inductively we get the following notion of higher order derivatives:

Definition 2.5.11. *Let $E \twoheadrightarrow M$ be a vector bundle and let ∇^E and ∇^{TM} be connections on E and TM respectively. Then for $s \in \Gamma(E)$ and $k \geq 1$ we define the* ***k-th covariant derivative*** *$(\nabla^k s) : \Gamma(TM)^k \to \Gamma(E)$ inductively as:*

$$(\nabla^k s) = \nabla^{(T^*M)^{\otimes k} \otimes E}(\nabla^{k-1} s),$$

where we take $\nabla^0 s = s$.

Recall that a fiber metric g on E induces a metric on the dual bundle E^* as well as the bundle $(E^*)^{\otimes k}$. We will denote the metric induced by g on these derived vector bundles by g as well. Now we have all the ingredients to define the following seminorms on $\Gamma(E)$:

Definition 2.5.12. *Let $E \twoheadrightarrow M$ be a vector bundle with M compact. Let g be a fiber metric on E and assume that we have connections on E and TM. Let ∇^k (for $k \in \mathbb{N}$) be the higher order derivatives as defined in Definition 2.5.11. Then we define the norms $\|\ldots\|_k : \Gamma(E) \to \mathbb{R}$ by:*

$$\|s\|_k = \sup_{p \in M} \sum_{i=0}^{k} \left| (\nabla^i s)(p) \right|_g,$$

where $|s|_g = \sqrt{g(s,s)}$.

Before we proceed, we assert the following:

Lemma 2.5.13. *The topology on $\Gamma(E)$ induced by the seminorms defined in Definition 2.5.12 does not depend on the choice of metric g and connection ∇.*

Proof. Let g' and g be two fiberwise metrics on E. Then we can choose a (small) constant $C_p > 0$ such that $g_p - C_p g'_p$ is positive definite on E_p. Such a choice can be made continuously in p. By compactness of M we can take $C = \min_{p \in M} C_p > 0$. Then with notation as above we have $C|\ldots|_{g'} \leq |\ldots|_g$.

Similarly, the topology does not depend on the choice of connections. To see this, take a set of sections $e_1, \ldots, e_q$ spanning E. Then $|\nabla^k s|_g$ is completely determined by any decomposition $s = \sum_i \lambda^i e_i$ and $\sqrt{g(\nabla^k e_i, \nabla^k e_j)}$. Now let $(\nabla')^k$ be a different set of higher order derivatives. Since $\sqrt{g(\nabla^k e_i, \nabla^k e_j)}$ and $\sqrt{g((\nabla)'^k e_i, (\nabla')^k e_j)}$ are continuous functions over a compact space, we can estimate one by the other. □

We now show that these seminorms induce a Fréchet space structure.

Proposition 2.5.14. *For a vector bundle $E \twoheadrightarrow M$ with M compact, the space $\Gamma(E)$ with seminorms as in Definition 2.5.12 is a Fréchet space.*

Proof. As in Definition 2.5.2, we have countably many norms (not just seminorms!), so it remains to check completeness. Take $(s_j)_j$ a Cauchy sequence of sections with respect to $\|\ldots\|_k$. Now pass to a compact neighborhood over which E is trivializable to reduce to a sequence of smooth maps $f_j : K \subseteq \mathbb{R}^m \to \mathbb{R}^n$, (where m is the dimension of M and n the dimension of the fibers of E). By the discussion above, the topology does not depend on the choice of connection so f_j is Cauchy with respect to the standard seminorms as in Definition 2.5.2 (which, as one can check, are simply the seminorms defined in Definition 2.5.12 for the standard connections and metric on Euclidean space). This gives us local convergence: over trivializable compacts $K \subseteq M$, we have $s_j|_K \to s|_K$. For global convergence we cover M by finitely many trivializible compact neighborhoods and take the maximum index j. □

Remark 2.5.15. The proof above relates to another approach to define seminorms on $\Gamma(E)$ by passing to a finite open cover of M over which E is trivializable. For this approach see [CB17]. △

The results shown so far all hold for compact base space M, but we can extend the definitions and results to noncompact base spaces M just as we did in the local case (see Definition 2.5.2, Corollary 2.5.4).

Definition 2.5.16. *Let $E \twoheadrightarrow M$ be a vector bundle, g be a fiber metric on E and assume that we have connections on E and TM. Let ∇^k (for $k \in \mathbb{N}$) be the higher order derivatives as defined in Definition 2.5.11. For $\{K_i \mid i \in \mathbb{N}\}$ a compact exhaustion of M, we define the norms $\|\ldots\|_{K_i,k} : \Gamma(E) \to \mathbb{R}$ by:*

$$\|s\|_{K_i,k} = \sup_{p \in K_i} \sum_{i=0}^{k} \left|(\nabla^i s)(p)\right|_g ,$$

where $|s|_g = \sqrt{g(s,s)}$.

Since we have:

$$\|s\|_{K_i,k} = \|s|_{K_i}\|_k ,$$

with the right hand side defined by Definition 2.5.12, we obtain the following results from Lemma 2.5.13 and Proposition 2.5.14 by considering any compact subset $K \subseteq M$.

Proposition 2.5.17. *The seminorms defined in Definition 2.5.16 define a Fréchet structure that does not depend on the choice of metric and connections.*

Next we consider a particular kind of map on spaces of sections, induced by smooth maps on vector bundles preserving the bundle structure. These will be generalizations of the maps defined in Equation (2.1). As in the local case, we first define the open subsets on which this kind of maps is defined. This definition is the global analogue of Definition 2.5.5.

Definition 2.5.18. *Let $E \twoheadrightarrow M$ be a vector bundle. For $V \subseteq E$ open and $K \subseteq M$ compact, we define:*

$$\Gamma(V|_K) := \{s \in \Gamma(E) \mid \mathrm{Im}(s|_K) \subseteq V\}.$$

If M is compact we simply write $\Gamma(V|_M) = \Gamma(V)$.

Note that the sets above are indeed open with respect to the Fréchet structure, both in the case of noncompact M and compact M (where we may take $K = M$). Continuing to mimic the development of the previous subsection, we extend the definition of maps as in Equation (2.1) to operators induced by bundle maps (Definition 2.5.9).

Definition 2.5.19. *Let E, E' be vector bundles over M and $V \subseteq E$ open. For $\phi : V \to E'$ a bundle map, we call a map of the form:*

$$\phi_*(s)(p) = \phi(p, s(p)),$$

*a **(vector) bundle operator**.*

The example of a smooth map $\mathcal{C}^\infty(K, \mathbb{R}^m) \to \mathcal{C}^\infty(K, \mathbb{R}^m)$ defined in Equation (2.1) is a bundle operator for the trivial bundle $K \times \mathbb{R}^m \twoheadrightarrow K$. Unsurprisingly, vector bundle operators are smooth in general. In fact, we use Proposition 2.5.6 to show this.

Proposition 2.5.20. *Any vector bundle operator $\phi_* : \Gamma(V) \to \Gamma(E')$ is smooth.*

Proof. Over a trivializible compact neighborhood K a bundle operator ϕ_* has the same form as the map defined in Equation (2.1). On this neighborhood K, therefore, $(\phi|_K)_*$ is differentiable. This implies that the following limit:

$$\lim_{t \to 0} \frac{(\phi|_K)_*(s|_K + th|_K) - (\phi|_K)_*(s|_K)}{t}$$

(where $s \in \Gamma(V)$, $h \in \Gamma(E)$) converges uniformly over *any* compact neighborhood $K \subseteq M$, hence the operator ϕ_* is "globally" differentiable.

The differential is once again a map of the same form:

$$D\phi_* : (s,h) \in \Gamma(U) \times \Gamma(E) \cong \Gamma(U \oplus E) \mapsto D_s\phi_* h \in \Gamma(E'),$$

and therefore differentiable. By induction, we see that ϕ_* is smooth. □

Chapter 3

Fréchet manifolds

Now that we have some tools to do calculus and in particular a notion of smoothness in Fréchet space, we can define Fréchet manifolds. Once we have defined these, we define some basic differential-geometric notions such as smooth maps, the tangent space and tangent maps. All of this is mostly the same as in the finite dimensional case. As in Chapter 2, we end the chapter by focusing on the main example of a Fréchet manifold: the space of sections of a fiber bundle. This will build on the results and definitions shown in Section 2.5.

3.1 Definitions

The following is very reminiscent of the definition for "standard" manifolds, and as such can be safely skipped by the reader who has seen it (too) many times. We start by defining charts and the notion of smooth compatibility, which gives rise to the notion of a smooth atlas.

Definition 3.1.1. *For F a Fréchet space, M a topological space and $U \subseteq M$ an open subset, we call an injective continuous map $\varphi : U \to F$ a **chart**.*

*We call two two charts $\varphi_1 : U_1 \subseteq M \to F_1$ and $\varphi_2 : U_2 \subseteq M \to F_2$ **smoothly compatible** if either $U_1 \cap U_2 = \emptyset$ or:*

$$\varphi_2 \circ {\varphi_1}^{-1} : \varphi_1(U_1 \cap U_2) \subseteq F_1 \to \varphi_2(U_1 \cap U_2) \subseteq F_2$$

is a diffeomorphism.

Definition 3.1.2. *For a topological space M, we call a collection of charts $\{\varphi_i : U_i \subseteq M \to F_i\}$ and corresponding domains a **Fréchet atlas** if every chart is smoothly compatible with any other and the U_i cover M.*

Definition 3.1.3. *A **Fréchet manifold** is a Hausdorff topological space with a (maximal) Fréchet atlas.*

Just as for finite-dimensional manifolds, we have a notion of submanifolds, which generalize the notion of a closed subspace of a Fréchet space.

Definition 3.1.4. *Let M be a Fréchet manifold. A **Fréchet submanifold** N of M is a subset $N \subseteq M$ such that for each point $f \in N$ there is a chart $\varphi : U \ni f \to F \times G$ (with F and G Fréchet spaces) such that:*

$$g \in N \cap U \textbf{ if and only if } \varphi(g) \in F \times 0.$$

*Such a chart is called a **submanifold chart**.*

The notion of smooth maps is also defined in the same way as usual.

Definition 3.1.5. *Let M, N be Fréchet manifolds and $\phi : M \to N$ a map. Then we say that ϕ is **smooth** if for any (hence every) pair of charts $\chi : \widetilde{U} \to U \subseteq F$ and $\psi : \widetilde{V} \to V \subseteq G$ of M and N respectively (such that $\phi(\widetilde{U}) \subseteq \widetilde{V}$) the map:*

$$\psi \circ \phi \circ \chi^{-1} : U \to V$$

is smooth as a map between Fréchet spaces.

Finally, we iterate the definition of a vector bundle.

Definition 3.1.6. *Let E, M be Fréchet manifolds with $\pi : E \to M$ a surjective smooth map. Denote $E|_{\widetilde{U}} = \pi^{-1}(\widetilde{U})$ for $\widetilde{U} \subseteq M$ open. Then we say that E is a* ***(Fréchet) vector bundle over*** *M if for any chart $\chi : \widetilde{U} \to U \subseteq F$ of M there is a chart $\chi_E : E|_{\widetilde{U}} \to U \times G$ (where F and G are Fréchet spaces) such that:*

$$\begin{array}{ccc} E|_{\tilde{U}} & \xrightarrow{\quad \chi_E \quad} & U \times G \subseteq F \times G \\ \big\downarrow{\scriptstyle \pi} & & \big\downarrow{\scriptstyle \pi_1} \\ \tilde{U} & \xrightarrow{\quad \chi \quad} & U \subseteq F \end{array}$$

commutes. Here π_1 is the projection onto the first factor.

Remark 3.1.7. The definition above has an obvious extension to define general *fiber* bundles over a Fréchet manifold by taking G an arbitrary Fréchet manifold. However we will not make use of Fréchet fiber bundles anywhere. △

3.2 The Tangent Bundle

In order to complete our manifold machinery we also need a notion of tangent spaces, which we define in this section. In the finite dimensional case, the two main constructions for an abstract tangent vector are to define them as derivations or equivalence classes of curves. In the infinite dimensional case these are not equivalent, as shown in Chapter VI of [KM97]. We will take the "kinematic" approach, defining tangent vectors as classes of curves. This construction will once again appear very similar to those familiar with the finite-dimensional case.

In order to define the tangent bundle we must start by defining a tangent space at a point. Let M be a Fréchet manifold and take $p \in M$. Define $\mathcal{C}_p$ to be the following space:

$$\mathcal{C}_p = \{\gamma : I \to M \mid 0 \in I \subseteq \mathbb{R} \text{ an open interval}, \gamma \text{ smooth such that } \gamma(0) = p\}.$$

Definition 3.2.1. *We define the equivalance relation $\sim$ on $\mathcal{C}_p$ by:*

$$\gamma_0 \sim \gamma_1 \iff (\chi \circ \gamma_0)'(0) = (\chi \circ g)'(0) \textit{ for any chart } \chi \textit{ around } p.$$

We will show that $\mathcal{C}_p/\sim$ can be endowed with a linear structure. First note that, since the transition maps of a manifold are diffeomorphisms between (opens in) Fréchet spaces, it suffices that $(\chi \circ \gamma_0)'(0) = (\chi \circ \gamma_1)'(0)$ for any one chart χ. Therefore fix such a chart $\chi : \tilde{U} \to U \subseteq F$ with F a Fréchet space and define:

$$\begin{aligned} \Psi_\chi : \quad & \mathcal{C}_p/\sim \to F \\ & [\gamma] \mapsto (\chi \circ \gamma)'(0), \end{aligned}$$

where $[\gamma]$ denotes the equivalence class of γ. By definition of the equivalence relation, this is a bijection. Its inverse is:

$$\Psi_\chi^{-1} : \ v \in F \mapsto \left[t \mapsto \chi^{-1}(\chi(p) + tv)\right] \in \mathcal{C}_p/\sim.$$

By identifying $\mathcal{C}_p/\sim$ and F through this bijection, we can define both a vector space and a Fréchet space structure on $\mathcal{C}_p/\sim$. The topology can be defined either by defining sets of the form $\Psi^{-1}(U) \subseteq \mathcal{C}_p/\sim$ (with $U \subseteq F$ open) to be open or, perhaps more concretely, by defining the seminorms:

$$\|\Psi(\dots)\|_k$$

on $\mathcal{C}_p/\sim$, with $\|\dots\|_k$ a set of seminorms on F inducing its topology.

Definition 3.2.2. *The* ***tangent space*** *T_pM of M at p is defined as $\mathcal{C}_p/\sim$ with a Fréchet structure and topology defined as above.*

Remark 3.2.3. Note that since the transition maps and their derivatives are isomorphisms between Fréchet spaces, both the vector space structure and the topology on T_pM are well-defined, i.e. independent of the choice of chart χ. △

Next we show that we can define a Fréchet manifold structure on the tangent space TM which, set-theoretically, equals the disjoint union $\sqcup_{p\in M} T_pM$. We denote $TM|_{\widetilde{U}} = \sqcup_{p\in\widetilde{U}} T_pM$ for the restriction of this bundle to a subset $\widetilde{U} \subseteq M$. Then for a chart $\chi : \widetilde{U} \to U \subseteq F$ of M, we define the following chart on TM:

$$\begin{aligned} T\chi : \quad & TM|_{\widetilde{U}} \to (U \subseteq F) \times F \\ & (p, [f_p]) \mapsto (\chi(p), (\chi \circ f_p)'(0)) \end{aligned} \tag{3.1}$$

Definition 3.2.4. *For M a Fréchet manifold with atlas of charts $\{\chi_i\}$, we define the* ***tangent bundle*** *TM of M to be the set:*

$$\bigsqcup_{p\in M} T_pM,$$

with smooth structure induced by the atlas of charts $\{T\chi_i\}$ as defined in Equation (3.1)

We check that this indeed defines a smooth manifold. Take a chart as in Equation (3.1). Then for a different chart $\varphi : \widetilde{V} \to V \subseteq G$ with G a Fréchet space, the second component of the transition map $T\varphi \circ (T\chi)^{-1}$ component is given by:

$$v \mapsto \left[t \mapsto \chi^{-1}(\chi(p) + tv)\right] \mapsto \left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \chi^{-1})(\chi(p) + tv) = D_{\chi(p)}(\varphi \circ \chi^{-1})v,$$

hence the total transition map is simply $T(\varphi \circ \chi^{-1})$ (as defined in Definition 2.4.11). This is again smooth and therefore gives TM a smooth structure.

Note that TM has the structure of a vector bundle as in Definition 3.1.6. Take $\pi : TM \to M$ to be the map sending the fiber T_pM to p. By Equation (3.1) we have the following:

$$\begin{array}{ccc} TM|_{\widetilde{U}} & \xrightarrow{\quad T\chi \quad} & (U \subseteq F) \times F \\ {\scriptstyle \pi}\big\downarrow & & \big\downarrow{\scriptstyle \pi_1} \\ \widetilde{U} & \xrightarrow{\quad \chi \quad} & U \subseteq F \end{array} \quad ,$$

where the horizontal arrows are charts. Note that this also shows smoothness of the projection π.

Finally, we generalize the tangent map of Definition 2.4.11:

Definition 3.2.5. *Let M, N be Fréchet manifolds and $\phi : M \mapsto N$ a smooth map. Then we define the* ***tangent map*** *$T\phi : TM \mapsto TN$ by:*

$$T\phi : (p, [\gamma]) \mapsto (\phi(p), [\phi \circ \gamma]).$$

Observe that for given charts $\varphi : \widetilde{U} \to U$ and $\psi : \widetilde{V} \to V$ of M and N respectively (such that $\phi(\widetilde{U}) \subseteq \widetilde{V}$) we have:

$$T(\psi \circ \phi \circ \varphi^{-1}) = T\psi \circ T\phi \circ T\varphi^{-1}$$

where on the left hand side we have the tangent map as in Definition 2.4.11 and on the right hand side the charts as in Definition 3.2.4. This shows that $T\phi$ is a smooth map.

3.3 Fiber bundle sections

As in the previous chapter, we end this chapter by putting the abstract concepts introduced so far in a more concrete setting. Since we showed in Section 2.5 that the space of sections over vector bundle is a Fréchet space, it is perhaps unsurprising that the space of sections of a more general *fiber* bundle is a Fréchet *manifold*, although this only holds over compact base spaces. Before defining the Fréchet structure on this set we go through some basic definitions for completeness.

3.3.1 Prerequisites

Definition 3.3.1. *Let $E \xrightarrow{\pi} M$ be two manifolds with π a smooth surjective submersion. We call E a **fiber bundle** with fiber F (which is some manifold) if for any $p \in M$ there is an open $U \ni p$ and a **trivialization** $\eta : E|_U \to U \times F$ such that the following diagram:*

$$\begin{array}{ccc} E|_U & \xrightarrow{\eta} & U \times F \\ & {\scriptstyle\pi}\searrow \quad \swarrow{\scriptstyle\pi_1} & \\ & U & \end{array}$$

commutes.

*Given two fiber bundles E, E' over M and $V \subseteq E$ open, we call a smooth map $\phi : V \subseteq E \to E'$ such that $\pi_{E'} \circ \phi = \pi_E$ a **(fiber) bundle map**.*

The notation we use for sections of a fiber bundle is the same as the one used for sections of a vector bundle.

Definition 3.3.2. *Let $E \xrightarrow{\pi} M$ be a fiber bundle over M. We define the **space of sections** $\Gamma(M, E)$ to be the set of maps $s : M \to E$ such that $\pi \circ s = Id_M$. When the base manifold is understood, we also simply denote this with $\Gamma(E)$.*

The construction of the Fréchet manifold structure on $\Gamma(E)$ will rely in a crucial way on the following definition:

Definition 3.3.3. *Given a fiber bundle $E \xrightarrow{\pi} M$, we define the **vertical bundle** to be:*

$$\mathrm{Vert}(E) = \mathrm{Ker}\, T\pi \subseteq TE.$$

For a point $e \in E$ we denote the fiber of this bundle by $\mathrm{Vert}_e(E) = \mathrm{Ker}\, T_e\pi = T_eE \cap \mathrm{Vert}(E)$. Similarly, for an (open) subset $V \subseteq E$, we define $\mathrm{Vert}(V) = \cup_{e \in V} \mathrm{Vert}_e(E)$.

Note that $\mathrm{Vert}(E)$ is a vector bundle over E, since by the diagram of Definition 3.3.1 $T\pi$ is of constant rank. The final geometric construction we will need in the upcoming is the following:

Definition 3.3.4. *Let $E \xrightarrow{\pi} M$ be a fiber bundle. Then we define a **(fiber) bundle metric** on E to be a fiber metric of $\mathrm{Vert}(E) \to E$ as defined in Definition 2.5.10.*

Remark 3.3.5. Note that for some bundle metric g of a fiber bundle E and $\varphi : E|_U \to U \times F$ a trivialization of E, we get a family of Riemannian metrics g_p on F that is smoothly parametrized by $p \in U$. This shows that this notion defined above truly gives us a "metric on each fiber". In particular, geodesics of the connection on $\mathrm{Vert}(E)$ induced by g are contained in fibers of E.

Also note that if E is a vector bundle, the notion above corresponds with Definition 2.5.10 if g is constant on each fiber. △

3.3.2 The Fréchet manifold structure of $\Gamma(E)$

Now we are ready to start constructing the Fréchet manifold structure on section spaces. The first order of business is defining charts on $\Gamma(E)$, which in turn requires that we choose some "model space" in which these take values. We will take the model space around some section $s_0 : M \to E$ to be the pullback bundle $s_0^* \mathrm{Vert}(E)$. Let $M \subseteq s_0^* \mathrm{Vert}(E)$ denote the zero section. Then we can map a neighborhood of M diffeomorphically onto an open in E containing the image of s_0 as follows. Choose a fiberwise metric with respect to the bundle $\mathrm{Vert}(E) \to E$ and let exp be the exponential map associated with this metric. Then the map $s_0^* \mathrm{Vert}(E) \to E$ defined by:

$$\exp_{s_0}(p, v) := \exp(s_0(p), v) \tag{3.2}$$

is a local diffeomorphism around the zero section $M \subseteq s_0^* \mathrm{Vert}(E)$. Additionally it is bijective close enough to M, hence the restriction of the above map to suitable opens $V \subseteq s_0^* \mathrm{Vert}(E)$ and $\widetilde{V} \subseteq E$:

$$\exp_{s_0}|_V : V \supseteq M \to \widetilde{V} \supseteq \mathrm{Im}(s_0)$$

is a diffeomorphism. We write $\varphi = (\exp_{s_0}|_V)^{-1} : \widetilde{V} \to V$

Now consider the subset of (smooth) sections with images in V, denoted by $\Gamma(V)$. Since φ above is a diffeomorphism, the map:

$$\varphi_* : \Gamma(\widetilde{V}) \subseteq \Gamma(E) \to \Gamma(V) \subseteq \Gamma(s_0^* \operatorname{Vert}(E)), \quad s \mapsto \varphi \circ s \tag{3.3}$$

is a bijection. As remarked below Definition 2.5.18, this set is open with respect to the toplogy defined by Definition 2.5.12 whenever the base manifold M is *compact*. This brings us to the following convention and definition.

Remark 3.3.6. We will assume throughout the remainder of this section that **all fiber bundles** $E \twoheadrightarrow M$ **are over compact base manifolds** M. △

Definition 3.3.7. *Let $E \twoheadrightarrow M$ be a fiber bundle. Then we define the smooth structure on $\Gamma(E)$ to be the one defined by charts:*

$$\varphi_* : \Gamma(\widetilde{V}) \to \Gamma(V)$$

with φ_* as in Equation (3.3)*.*

We check that these charts indeed define a smooth atlas:

Proposition 3.3.8. *Let $E \twoheadrightarrow M$ be a fiber bundle over M. Let $s_0, s_1 \in G$, and $(\varphi_i)_* : \Gamma(\widetilde{V_i}) \to \Gamma(V_i)$ $(i \in \{0,1\})$ charts as in Equation* (3.3) *around s_0 and s_1 respectively. Suppose that $\Gamma(\widetilde{V_0}) \cap \Gamma(\widetilde{V_1}) = \Gamma(\widetilde{V_0 \cap V_1}) \neq \emptyset$. Then the map:*

$$(\varphi_1)_* \circ (\varphi_0)_*^{-1} : (\varphi_0)_*(\widetilde{V_0 \cap V_1}) \to (\varphi_1)_*(\widetilde{V}_0 \cap \widetilde{V}_1)$$

is smooth.

Proof. Note that this map is a vector bundle operator (Definition 2.5.19) between open subsets of the vector bundles $s_i^* \operatorname{Vert}(E)$. It is induced by the diffeomorphism:

$$\varphi_1 \circ \varphi_0^{-1} : \varphi_0(V_0 \cap V_1) \subseteq s_0^* \operatorname{Vert}(E) \to \varphi_1(V_0 \cap V_1) \subseteq s_1^* \operatorname{Vert}(E)$$

By Proposition 2.5.20 it is a smooth map between Fréchet spaces. □

Now that we have a well-defined smooth structure on $\Gamma(E)$, we examine a particular kind of smooth map between (opens in) $\Gamma(E)$ and $\Gamma(E')$. Just as fiberwise maps between vector bundles induce smooth maps between the corresponding Fréchet spaces, fiberwise maps of fiber bundles induce smooth maps between the corresponding Fréchet manifolds.

Definition 3.3.9. *Let $E \xrightarrow{\pi_E} M$ and $E' \xrightarrow{\pi_{E'}} M$ be fiber bundles over M and $V \subseteq E$ open. A map $\phi_* : \Gamma(V) \subseteq \Gamma(E) \to \Gamma(E')$ induced by a bundle map ϕ:*

$$\phi_*(s)(p) = \phi(p, s(p)),$$

is called a ***(fiber) bundle operator****.*

Remark 3.3.10. Note that if the fiber F of a fiber bundle E is a vector space, then E is a vector bundle. In this case it should be clear that the notions of a fiber bundle operator and vector bundle operator (or map) coincide. Because of this, we will usually just use the terms “bundle operator” and “bundle map”. △

Proposition 3.3.11. *Fiber bundle operators are smooth.*

Proof. Note that with respect to the local charts as in Definition 3.3.7 a fiber bundle operator ϕ_* is a (nonlinear) vector bundle operator, which is smooth by Proposition 2.5.20. In fact, the charts of Definition 3.3.7 are themselves bundle operators and compositions of bundle operators. □

3.3.3 The tangent bundle of $\Gamma(E)$

It is now a natural question to ask what the tangent bundle of $\Gamma(E)$ looks like. It turns out that this is again a space of sections, namely $\Gamma(\operatorname{Vert}(E))$.

Let $t \in I \mapsto s_t \in \Gamma(E)$ be a smooth family of sections in $\Gamma(E)$ over some open interval containing 0. Note that two such "curves" s_t and s'_t are equivalent under the relation defined in Definition 3.2.1 and for charts as in Definition 3.3.7 if and only if:

$$\left.\frac{d}{dt}\right|_{t=0} s_t(p) = \left.\frac{d}{dt}\right|_{t=0} s'_t(p)$$

for all $p \in M$. Note that $\left.\frac{d}{dt}\right|_{t=0} s_t(p) \in \operatorname{Vert}_{s_0(p)}(E)$. Therefore the derivative at $t = 0$ defines a lift $\widetilde{s_0}$ of s_0 in the following diagram:

$$\begin{array}{ccc} & & \operatorname{Vert}(E) \\ & \overset{\widetilde{s_0}}{\nearrow} & \downarrow \\ M & \xrightarrow[s_0]{} & E \end{array} \quad . \tag{3.4}$$

In other terms $\widetilde{s_0}$ is a section of $s_0^* \operatorname{Vert}(E)$. Conversely, it is clear that given any $\widetilde{s_0} \in \Gamma(s_0^* \operatorname{Vert}(E))$ we can take a family s_t such that $\left.\frac{d}{dt}\right|_{t=0} s_t = \widetilde{s_0}$. For example, we may take a chart φ_* mapping into $s_0^* \operatorname{Vert}(E)$ as in Definition 3.3.7 and take $t \mapsto (\varphi_*)^{-1}(t\widetilde{s_0})$. We therefore have the following:

Proposition 3.3.12. *Let $E \twoheadrightarrow M$ be a fiber bundle and $\Gamma(E)$ its space of sections. Then, with the terminology as described above, the map $T\Gamma(E) \to \Gamma(\operatorname{Vert}(E))$ defined by:*

$$\Phi : (s_0, [s_t]) \mapsto \widetilde{s_0} = \left.\frac{d}{dt}\right|_{t=0} s_t$$

(where $[s_t]$ denotes the equivalence class as in Definition 3.2.1) is a diffeomorphism.

Proof. The discussion above shows that the map is a bijection. We will have to check smoothness of Φ and its inverse. In this proof we assume all bundles to be over M and consequently Vert denotes the vertical bundle with respect to the projection to M.

Take $s_0 \in \Gamma(E)$ some section. Then the local model space around s_0 is, as seen before, the space $E_0 = s_0^* \operatorname{Vert}(E)$. The charts into this model space are induced by maps φ defined by:

$$\eta := \varphi^{-1} : (p, v) \mapsto \exp(s_0(p), v),$$

(As seen in Definition 3.3.7). The map η is a diffeomorphism between some sufficiently small open V around the zero section M and some open $\widetilde{V} \supseteq \operatorname{Im}(s_0)$. The (inverse of the) chart on $\Gamma(E)$ induced by η is then given by:

$$\eta_*(s)(p) = \eta(p, s(p)) = \exp(s_0(p), s(p)), \quad \eta_* : \Gamma(V) \mapsto \Gamma(\widetilde{V})$$

To show the smoothness of Φ, we will compute $\Phi \circ T\varphi_*^{-1} = \Phi \circ T\eta_*$.

The tangent map $T\eta_*$ is given by:

$$(s, \sigma) \mapsto (s, [t \mapsto \eta_*(s + t\sigma)]),$$

Therefore $\Phi \circ T\eta_*$ is given by:

$$(s, \sigma) \mapsto \left.\frac{d}{dt}\right|_{t=0} \eta_*(s + t\sigma)$$

We will compute this pointwise. For fixed $p \in M$, write $\eta_p : T_{s_0(p)}E \to \widetilde{V}$. Then for $v_0 \in T_{s_0(p)}E$ such that $(p, v_0) \in V$ and some $v \in T_{v_0}(E_0)$ we have:

$$\left.\frac{d}{dt}\right|_{t=0} (\eta_p(v_0 + tv)) = T_{v_0}\eta_p v$$

Observe that since E_0 is a vector bundle, there is a canonical isomorphism $\operatorname{Vert}_{v_0} E_0 \cong (E_0)_p = \operatorname{Vert}_{s_0(p)} E$. Since the curve $t \mapsto \eta_p(v_0 + tv)$ is tangent to the fiber at p, we see that $T_{v_0}\eta_p : T_{v_0}E_0 \to T_{\eta_p(v_0)}E$ restricts to a map:

$$T_{v_0}\eta_p : \operatorname{Vert}_{s_0(p)} E \cong \operatorname{Vert}_{v_0} E_0 \to \operatorname{Vert}_{\eta_p(v_0)} E.$$

This map is a linear isomorphism since η is a local diffeomorphism on V.

Applying this to the induced map of sections gives:

$$(\Phi \circ T\eta_*)(s,\sigma)(p) = \frac{d}{dt}\bigg|_{t=0} \eta_*(s+t\sigma)(p) = \frac{d}{dt}\bigg|_{t=0} \eta_p(s(p)+t\sigma(p)) = T_{s(p)}\eta_p\sigma_p$$

This shows that $\Phi \circ T\eta_* : \Gamma(V) \times \Gamma(E_0) = \Gamma(V \oplus E_0) \to \Gamma(\mathrm{Vert}(E)|_{\widetilde{V}})$ is an invertible bundle operator induced by the bundle map:

$$(p, v_0, v) \mapsto T_{v_0}\eta_p v.$$

By Proposition 3.3.11 bundle operators are smooth hence Φ is smooth. Since it is locally a diffeomorphism[1] and bijective, it is a diffeomorphism. □

Remark 3.3.13. By the canonical identification above, we will refer to $\Gamma(\mathrm{Vert}(E))$ as the tangent bundle of $\Gamma(E)$. Observe that for E a vector bundle, we have that $\mathrm{Vert}(E) = E \oplus E$, hence $T\Gamma(E) = \Gamma(E \oplus E) = \Gamma(E) \times \Gamma(E)$. This resembles the finite-dimensional case, where $T\mathbb{R}^n = \mathbb{R}^{2n}$.

At a given section $s_0 \in \Gamma(E)$, the tangent space at s_0 is the subset of sections in $\mathrm{Vert}(E)$ that lift s_0, as in diagram Equation (3.4). This is precisely the set $\Gamma(s_0^* \mathrm{Vert}(E))$, therefore we will also say that $T_{s_0}\Gamma(E) = \Gamma(s_0^* \mathrm{Vert}(E))$. △

Now we ask ourselves what the tangent map of a bundle operator is. As one might expect, it involves the tangent map of the underlying bundle map.

Proposition 3.3.14. *Let E, E' be fiber bundles over M, $V \subseteq E$ open and $\phi : V \to E'$ a bundle map. Then we have:*

$$T(\phi_*) = (T\phi|_{\mathrm{Vert}})_* : T\Gamma(V) \subseteq T\Gamma(E) \to T\Gamma(E'),$$

where $T\phi|_{\mathrm{Vert}}$ denotes the restriction of $T\phi$ to $\mathrm{Vert}(E) \subseteq TE$.

Proof. First note that since ϕ is a bundle map, the image of $T\phi|_{\mathrm{Vert}}$ lies in $\mathrm{Vert}(E')$, so the map on the right hand side of the equality is well-defined.

Let $\widetilde{s_0} \in \Gamma(\mathrm{Vert}(V))$ such that $\pi_E^{\mathrm{Vert}(E)}(\widetilde{s_0}) = s_0$. Under the identification in Proposition 3.3.12, this corresponds to $(s_0, [s_t]) \in T\Gamma(V)$, where s_t is is some family of sections such that $\frac{d}{dt}\big|_{t=0} s_t = \widetilde{s_0}$. The tangent map of ϕ_* sends this to:

$$(s_0, [s_t]) \mapsto (\phi_*(s_0), [\phi_*(s_t)])$$

We compute the derivative at zero of s_t pointwise. For fixed $p \in M$ write $\phi_p(x) = \phi(p, x)$, where $x \in E_p$. Then:

$$\frac{d}{dt}\bigg|_{t=0} (\phi_*(s_t)(p)) = \frac{d}{dt}\bigg|_{t=0} \phi_p(s_t(p)) = T_{s_0(p)}(\phi_p)\widetilde{s_0}(p),$$

which shows the result. □

[1]Note that for Fréchet spaces, being a local diffeomorphism is *not* equivalent to the tangent map being invertible at any point.

Chapter 4

Tame Fréchet spaces and maps

In this chapter we build on the abstractions constructed in Chapter 2 and Chapter 3 and add a new construction: tameness. This concept and related results will be used in Chapter 5 for the statement of Hamilton's Nash-Moser theorem. The contents of this chapter, which can indeed be seen as a preperation for Chapter 5, are based on [Ham82]. We have attempted to provide more detail than the original work.

Recall that a map $P : F \to G$ between Fréchet spaces (or more generally, locally convex vector spaces) is continuous if and only if:

$$\|P(f)\|' \leq \|f\|$$

for seminorms $\|\ldots\|$ and $\|\ldots\|'$ of F and G respectively. Now in many Fréchet spaces, there is a natural "ordering" $\|\ldots\|_k$ of seminorms (see, for example, Example 2.5.1 and Definition 2.5.2). This will be examined more carefully in Section 4.1. Tameness, in short, then states that the estimate $\|P(f)\|'_j \leq \|f\|_i$ should be satisfied by indices i and j that "do not lie too far apart". This behavior is seen, for example, by a linear differential operator:

$$f \mapsto \sum_{0 \leq |\alpha| \leq r} c_\alpha \partial^\alpha f$$

on $\mathcal{C}^\infty(U, \mathbb{R}^m)$ (with Fréchet structure as in Definition 2.5.2).

After defining the aforementioned "orderings" on Fréchet spaces and tameness, we will prove some technical estimates for tame maps that will be used in Chapter 5. Also in preparation for Chapter 5, we will define *smoothing operators*, a common theme in all Nash-Moser theorems. We end the chapter by showing that the familiar space of vector bundle sections is tame (if the base manifold is compact), and that vector bundle operators are tame.

4.1 Graded Fréchet spaces

In this section we define the "ordering" mentioned in the introduction of the chapter. As mentioned there, the main example to keep in mind are the ones defined in Example 2.5.1 and Definition 2.5.2, although we will shortly provide others.

Definition 4.1.1. *A* ***grading*** *of a Fréchet space F is a countable collection $\{\|\ldots\|_n \mid n \in \mathbb{N}\}$ of seminorms on F its topology such that $\|\ldots\|_{n+1} \geq \|\ldots\|_n$ for all $n \in \mathbb{N}$. A* ***graded (Fréchet) space*** *is a Fréchet space with a choice of grading.*

Many properties of a graded space will be preserved with a grading that is not too different from the original:

Definition 4.1.2. *We say two gradings $\{\|\ldots\|_n\}$ and $\{\|\ldots\|'_n\}$ are* ***(tamely) equivalent*** *if there are r and b such that:*

$$\|f\|_n \leq C_n \|f\|'_{n+r} \text{ and } \|f\|'_n \leq C'_n \|f\|_{n+r} \text{ for all } n \geq b.$$

The constants r and b are called the ***degree*** *and* ***base*** *of the equivalence.*

The following example of a graded Fréchet space will be important for the definition of tameness.

Definition 4.1.3. *Let $(B, \|\ldots\|)$ be a Banach space, define $\Sigma(B)$ as the space of sequences $(f_k)_k$ in B such that:*

$$\|(f_k)_k\|_n = \sum_{k\in\mathbb{N}} e^{nk} \|f_k\| < \infty$$

for all $n \in \mathbb{N}$, with seminorms given by $\|\ldots\|_n$.

It is straightforward to check that with these seminorms, $\Sigma(B)$ is a graded space. However, there are other collections of graded seminorms that have "the same structure".

Lemma 4.1.4. *Given a Banach space B, the following graded collections of seminorms (where n ranges over $\mathbb{N}$) on $\Sigma(B)$ are equivalent to the grading in definition 4.1.3:*

$$\|(f_k)_k\|_{n,\infty} = \sup_k e^{nk} \|f_k\|$$

$$\|(f_k)_k\|_{n,p} = \left(\sum_{k\in\mathbb{N}} e^{pnk} \|f_k\|^p\right)^{1/p} \quad \textit{for } p \geq 1.$$

Proof. We will show that we have the following inequalities:

$$\|(f_k)_k\|_{n,\infty} \leq \|(f_k)_k\|_{n,p} \leq \|(f_k)_k\|_n \leq C \|(f_k)_k\|_{n+r,\infty}.$$

The first inequality is straightforward. For the second inequality, recall that for the sequence spaces l^p and l^q with $1 \leq q \leq p$ we have that $l^p \subseteq l^q$ and $\|\ldots\|_{l^p} \leq \|\ldots\|_{l^q}$ (this is a standard result in functional analysis). Take $q = 1$ to obtain the result. To see the last inequality, note that $e^{nk} \|f_k\| \leq \|(f_k)_k\|_{n+r,\infty} e^{-rk}$. By summing over k we see that we may take $C = \sum_k e^{-rk}$. □

The following more concrete example will be used in Section 4.4 to show that the space of smooth sections of a vector bundle (over a compact manifold) is tame.

Example 4.1.5. Let (X, μ) be a measure space and $w : X \to \mathbb{R}_{\geq 0}$ a positive weight function. We define $L_1(X, \mu, w)$ to be the space of measurable functions $f : X \to \mathbb{R}$ such that:

$$\|f\|_n = \int_X e^{nw} |f| d\mu < \infty$$

for any $n \in \mathbb{N}$. This is a Fréchet space and has a grading since $\|f\|_n \leq \|f\|_m$ for $n \leq m$. △

Definition 4.1.6. *A **tame linear map** $L : F \to G$ between two graded Fréchet spaces is a linear map for which there are r and b such that:*

$$\|Lf\|_n \leq C \|f\|_{n+r} \textit{ for all } n \geq b,$$

*and for any $f \in F$. Once more the constants r and b are called the **degree** and **base** respectively. We say L is a **tame isomorphism** if L is a linear isomorphism and both L and L^{-1} are tame.*

Note that tameness automatically implies continuity and that the composition of two tame linear maps is again a tame linear map. Note also that two gradings are tamely equivalent if and only if the identity map (from a space with one grading into the same space with a different grading) is a tame isomorphism. With the definition of tame linear maps we can finally give the definition of the objects of the category we are interested in: tame Fréchet spaces.

Definition 4.1.7. *We say that a graded space F is a **tame direct summand** of another graded space G if there are tame linear maps:*

$$F \xrightarrow{L} G \xrightarrow{R} F$$

such that $R \circ L = Id_F$.

Lemma 4.1.8. *If F is a tame direct summand of G then G is tamely isomorphic to $F \oplus H$, for some subspace $H \subseteq G$. Here H is given the graded collection of seminorms induced by $G \supseteq H$ and $F \oplus H$ has the graded collection of seminorms $\|\ldots\|_n = \|\ldots\|_n^F + \|\ldots\|_n^H$.*

Proof. Suppose that we have tame linear maps L, R such that the composition $F \xrightarrow{L} G \xrightarrow{R} F$ equals the identity on F. Let r_L and r_R denote the degrees of L and R respectively, which we assume to be non-negative (we will ignore the base, since we can always just take the maximum). Let $H = \operatorname{Ker} R$. Then the map $N : F \oplus H \to G$ given by $(f, h) \mapsto L(f) + h$ has inverse:

$$N^{-1} : g \mapsto (Rg, g - LRg).$$

We see that N is tame since:

$$\|L(f) + h\|_n \leq C \left(\|f\|_{n+r_L} + \|h\|_n \right) \leq C \left(\|f\|_{n+r_L} + \|h\|_{n+r_L} \right).$$

Its inverse is tame as well:

$$\|g - LRg\|_n + \|Rg\|_n \leq C(\|g\|_{n+r_L+r_R}).$$

□

Definition 4.1.9. *A* ***tame (Fréchet) space*** *F is a graded Fréchet space that is the direct summand of $\Sigma(B)$ for some Banach space B.*

Lemma 4.1.10. *Let F, G and H be graded spaces. If G is a tame direct summand of H, and F is a tame direct summand G, then F is a tame direct summand of H. In particular, a tame direct summand of a tame space is also a tame space.*

Proof. Note that we have tame linear maps such that the composition:

$$F \to G \to H \to G \to F$$

is the identity. If G is tame, then we can take $H = \Sigma(B)$ for some Banach space B and we see that F is also tame. □

The following result will later on be used to show that the space of sections of a vector bundle is tame. It also serves to illustrate the use of Lemma 4.1.10.

Proposition 4.1.11. *The space $L_1(X, \mu, w)$ defined in Example 4.1.5 is tame.*

Proof. Define $X_k = \{k \leq w < k+1\}$ and let χ_k the characteristic function on X_k and define:

$$L : f \mapsto (\chi_k f)_k, \quad L_1(X, \mu, w) \to \Sigma(L^1(X, \mu))$$
$$R : (f_k)_k \mapsto \sum_k \chi_k f_k, \quad \Sigma(L^1(X, \mu)) \to L_1(X, \mu, w).$$

First we show that $\|Lf\|_n \leq \|f\|_n$. Since on X_k we have $k \leq w$, we see:

$$\sum_k e^{nk} \|\chi_k f\|_{L^1} = \sum_k e^{nk} \int_X \chi_k |f| d\mu = \int_X \sum_k e^{nk} \chi_k |f| d\mu \leq \int_X e^{nw} |f| d\mu,$$

(where we may interchange the summation and integral by monotone convergence), proving the estimate. For R we have the estimate:

$$\int_X e^{nw} \sum_k |\chi_k f_k| d\mu \leq \int_X \sum_k e^{n(k+1)} \chi_k f_k d\mu \leq \sum_k e^{n(k+1)} \int_X \chi_k f_k d\mu \leq e^n \sum_k e^{nk} \|f_k\|_{L^1},$$

so $\|R((f_k)_k)\|_n \leq e^n \|(f_k)_k\|_n$, which means that R is tame. Since the composition $R \circ L$ is the identity on $L_1(X, \mu, w)$, this shows that it is a tame space. □

4.2 Nonlinear maps and some estimates

So far we have only looked at *linear* tame maps. The class of non-linear maps we are interested in are also required to "interact nicely" with the tame structure of tame spaces, in the following way:

Definition 4.2.1. *Let F and G be graded spaces and $P : U \subseteq F \to G$ a map. We say P* ***satisfies a tame estimate*** *if there are r, b (the* ***degree*** *and* ***base****)*

such that:

$$\|P(f)\|_n \leq C_n \left(1 + \|f\|_{n+r}\right)$$

for all $n \geq b$ and for all $f \in U$. We say that a map $P : U \subseteq F \to G$ is ***tame*** *if U is open, P is continuous and the domain U can be covered by opens on which P satisfies a tame estimate.*

Note that the constant may depend on n, which we assume from now on.

Lemma 4.2.2. *A map $L : F \to G$ is tame linear if and only if it is a tame map that is linear.*

Proof. It is straightforward to show that a tame linear map is tame. Suppose that $L : F \to G$ is tame and linear, then for some $b, r \in \mathbb{N}$ and on some neighborhood U of zero it satisfies an inequality of the form:

$$\|L(f)\|_n \leq C(1 + \|f\|_{n+r}),$$

for all $n \geq b$ and $f \in U$. Since U is a neighborhood, there are an $a \in \mathbb{N}$ and $\delta > 0$ such that $\{\|f\|_a < \delta\} \subseteq U$. Then for any nonzero $f \in F$, $\frac{1}{2} f / \|f\|_a \in U$ and therefore:

$$\frac{\delta}{2\|f\|_a} \|L(f)\|_n \leq C \left(1 + \|f\|_{n+r} \frac{\delta}{2\|f\|_a}\right)$$

Then for all $n \geq a - r$ we have:

$$\|L(f)\|_n \leq C \left(\frac{2\|f\|_a}{\delta} + \|f\|_{n+r}\right) \leq C \|f\|_{n+r},$$

so L satisfies a tame linear estimate of degree r and base $\max b, a - r$. □

Theorem 4.2.3. *The composition of tame maps is tame.*

Proof. Suppose we have tame maps:

$$U \subseteq F \xrightarrow{P} V \subseteq G \xrightarrow{Q} H,$$

such that $P(U) \subseteq V$. Let $V_0 \subseteq V$ be a neighborhood on which Q satisfies a tame estimate of base c and degrees s, we may restrict $P^{-1}(V_0)$ to a neighborhoud U_0 on which P satisfies a tame estimate of base b and degree r. Then we have:

$$\|Q(P(f))\|_n \leq C(1 + \|P(f)\|_{n+s}) \leq C(1 + C(1 + \|f\|_{n+r+s})) \leq C(1 + \|f\|_{n+r+s}),$$

where the first estimate holds for $n \geq c$ and the second for $n \geq b - s$. It follows that $Q \circ P$ satisfies a tame estimate of degree $r + s$ and base $\max\{c, b - s\}$. □

We will also need a notion of tameness for multivariate maps:

Definition 4.2.4. *We say a multivariate map $(f, g) \mapsto P(f, g)$ satisfies a tame estimate of base b, degree r in f and degree s in g on some neighborhood U if:*

$$\|P(f, g)\|_n \leq C(1 + \|f\|_{n+r} + \|g\|_{n+s})$$

for all $n \geq b$.

The kind of maps to which one may apply the Nash-Moser theorem are the following:

Definition 4.2.5. *A* ***smooth tame map*** *$P : U \subseteq F \to G$ between graded spaces is a map that is smooth and whose derivatives $D^k P$ are tame (in the multivariate sense above) for all k.*

For (families of) linear maps, we can take tame estimates to be of a more convenient form:

Lemma 4.2.6. *Suppose $L : (U \subseteq F) \times G \to H$ is linear in the last argument and satisfies the tame estimate:*

$$\|L(f \mid g)\|_n \leq C(1 + \|f\|_{n+r} + \|g\|_{n+s})$$

for $n \geq b'$ on some neigborhood $(f_0, 0) \in V \subseteq U \times G$. Then for some (possibly higher) base b, L satisfies the estimate:

$$\|L(f \mid g)\|_n \leq C(\|f\|_{n+r} \|g\|_{b+s} + \|g\|_{n+s})$$

on some $\|\ldots\|_{b+r}$ neighborhood of f_0 and for all $g \in G$.

Proof. The neighborhood V contains some product neighborhood:

$$W_U \times W_G = \{\|f - f_0\|_k < \delta\} \times \{\|g\|_l < \epsilon\},$$

where we can choose the integers $k = b + r$ and $l = b + s$ by increasing them if needed. Given any $g \in G$, we see that $\tilde{g} := \frac{1}{2}\epsilon g / \|g\|_{b+s}$ is contained W_G. Applying the estimate gives for $f \in W_U$ gives:

$$\|L(f)\tilde{g}\| \leq C(1 + \|f\|_{n+r} + \|\tilde{g}\|_{n+s}).$$

Multiplying both sides by $2\|g\|_{b+s}/\epsilon$ gives:

$$\|L(f)g\| \leq C\left(\frac{2}{\epsilon}\left(\|g\|_{b+s} + \|f\|_{n+r}\right) + \|h\|_{n+s}\right) \leq C(\|h\|_{n+s} + \|f\|_{n+r}\|h\|_{b+s}),$$

for $b \leq n$, where in the last line we enlarge our constant C if necessary and use $\|h\|_{b+s} \leq \|h\|_{n+s}$. □

By a similar proof we obtain the result for a family of bilinear maps:

Lemma 4.2.7. *Suppose $L : (U \subseteq F) \times G \times H \to K$ is linear in the last two arguments and satisfies the tame estimate:*

$$\|L(f \mid g, h)\|_n \leq C(1 + \|f\|_{n+r} + \|g\|_{n+s} + \|h\|_{n+t})$$

for $n \geq b'$ on some neigborhood $(f_0, 0, 0) \in V \subseteq U \times G \times H$. Then for some possibly higher base b, L satisfies the estimate:

$$\|L(f \mid g, h)\|_n \leq C\left(\|f\|_{n+r}\|g\|_{b+s}\|h\|_{b+t} + \|g\|_{n+s}\|h\|_{b+t} + \|g\|_{b+s}\|h\|_{n+t}\right)$$

on some $\|\ldots\|_{b+r}$ neighborhood of f_0 and for all $g \in G$.

4.3 The smoothing operator and interpolation

One of the main tools in the proof of the Nash-Moser theorem is the so-called "smoothing operator" S_t (defined below). Intuitively we think of elements $f = (f_k)_k \in \Sigma(B)$ as a representation of functions such that the k-order entry of f represents k order regularity in some sense[1]. The smoothing operator will set entries f_k with $k > t$ to zero, allowing us to consider a space where we only have to worry about some fixed regularity smaller than t. However, as $t \to \infty$, the operator S_t converges to the identity.

Definition 4.3.1. *Let $\psi : \mathbb{R} \to [0, 1]$ be a non-decreasing bump function such that $\psi(x) = 0$ for $x \leq 0$ and $\psi(x) = 1$ for $x \geq 1$ [2]. We define the **smoothing operator** $S : \Sigma(B) \times \mathbb{R} \to \Sigma(B)$, with B some Banach space, by:*

$$S_t(f) := S(f, t) = (\psi(t - k) f_k)_k$$

The following result makes the statement that S_t "converges to the identity" precise.

Lemma 4.3.2. *Let $f : \mathbb{R} \to \Sigma(B)$ such that $f(t) \to f_\infty$ as $t \to \infty$. Then $S_t f(t) \to f_\infty$ as $t \to \infty$.*

[1] See the proof of the tameness of $C^\infty(M)$ in Theorem 4.4.2, where a function g is sent to some frequency space, which in turn is sent to $\Sigma(L^1(\mathbb{R}))$

[2] Besides these properties, the precise definition of ψ does not matter.

Proof. We have to show that for any n, $\|S_t f(t) - f_\infty\|_n \to 0$ as $t \to \infty$. Let $\epsilon > 0$ be given and take T such that for $t > T$ we have $\|f(t) - f_\infty\|_n < \frac{1}{2}\epsilon$. For fixed n, there is a K such that:

$$\sum_{k>K} e^{nk} \|(f_\infty)_k\| < \tfrac{1}{2}\epsilon,$$

Then for $t > \max\{K+1, T\}$ we have:

$$\|S_t f(t) - f_\infty\|_n = \sum_{k\leq K} e^{nk} \|f(t)_k - (f_\infty)_k\| + \sum_{k>K} e^{nk} \|\psi(t-k)f(t)_k - (f_\infty)_k\| < \epsilon,$$

since $\psi(t-k) \in [0,1]$ for all t, k. □

Applying the smoothing operator to some element f allows us to estimate it by some lower order seminorm.

Lemma 4.3.3. *For all $n, q \leq n$ in $\mathbb{N}$ we have:*

$$\|S_t f\|_n \leq C e^{qt} \|f\|_{n-q}$$
$$\|(1-S_t)f\|_n \leq C e^{-qt} \|f\|_{n+q}.$$

Proof. For the first estimate, we compute:

$$\|S_t f\|_n \leq \sum_{k\leq t} e^{nk} \|f_k\| = \sum_{k\leq t} e^{qk} e^{(n-q)k} \|f_k\| \leq e^{qt} \sum_{k\leq t} e^{(n-q)k} \|f_k\| \leq e^{qt} \|f\|_{n-q}.$$

We omit the computation for the second estimate, which is analogous. □

Proposition 4.3.4 (Interpolation estimate)**.** *For $l \leq m \leq n \in \mathbb{N}$ we have the estimate:*

$$\|f\|_m^{n-l} \leq C \|f\|_l^{n-m} \|f\|_n^{m-l}$$

Proof. By Lemma 4.3.3 we have for all t:

$$\|f\|_m \leq \|S_t f\|_m + \|(1-S_t)f\|_m \leq C\left(e^{(m-l)t} \|f\|_l + e^{(m-n)t} \|f\|_n\right).$$

We solve for t in the equation $e^{(m-l)t} \|f\|_l = e^{(m-n)t} \|f\|_n$ to obtain:

$$t = \frac{1}{(n-l)} \log \frac{\|f\|_n}{\|f\|_l}$$

plugging it back in we get:

$$e^{(m-l)t} \|f\|_l = e^{(m-n)t} \|f\|_n = \|f\|_n^{\frac{m-l}{n-l}} \|f\|_l^{\frac{n-m}{n-l}},$$

which shows:

$$\|f\|_m \leq C \|f\|_n^{\frac{m-l}{n-l}} \|f\|_l^{\frac{n-m}{n-l}}.$$

□

The following particular form of the estimate will be used regularly in the proof ahead, allowing us to "trade orders of regularity" in expressions where two seminorms are multiplied.

Corollary 4.3.5. *Let $n, k, a \geq 0$ be integers with $k \leq n$ and $f \in \Sigma(B)$ for some Banach space B. Then:*

$$\|f\|_n \|f\|_{k+a} \leq C \|f\|_{n+a} \|f\|_k$$

Proof. Using the interpolation estimate we obtain:

$$\|f\|_n^{n+a-k} \leq C \|f\|_{n+a}^{n-k} \|f\|_k^a,$$
$$\|f\|_{k+a}^{n+a-k} \leq C \|f\|_{n+a}^a \|f\|_k^{n-k}$$

Multiplying these and taking the $(n+a-k)$-th root gives the result. □

As a final note on smoothing operators, we note that they can be extended (in a sense) to general tame spaces, by combining S_t with the maps L and R seen in Section 4.1. This does, however, come at the cost of adding a nonzero degree to the smoothing operator.

Lemma 4.3.6. *Let F be a tame Fréchet space. Then then there are numbers $b, r \in \mathbb{N}$ and a family of operators $S_{F,t} : F \to F$ such that:*

$$\|S_{F,t}f\|_n \leq Ce^{qt} \|f\|_{n+r-q},$$
$$\|(1 - S_{F,t})f\|_n \leq Ce^{-qt} \|f\|_{n+r+q}$$

for all $n \geq b$.

Proof. Since F is tame, there are tame linear maps $F \xrightarrow{L} \Sigma(B) \xrightarrow{R} F$ for some Banach space B, such that $R \circ L = \mathrm{Id}_F$. Denote with r_L and r_R the degrees of L and R respectively and set $r = r_L + r_R$. Then define

$$S_{F,t} = R \circ S_t \circ L,$$

with S_t defined by Definition 4.3.1. Then the result follows by Lemma 4.3.3 and choosing the appropriate base b. ☐

Corollary 4.3.7. *Let F be a tame Fréchet space. Then there exist integers b, r such that for $m \geq b$ and all $l \leq m \leq n$:*

$$\|f\|_m^{n-l} \leq C \|f\|_{n+r}^{m-l} \|f\|_{l+r}^{n-m}.$$

Additionally, for $k + a \geq b$ and $k \leq m$ we have:

$$\|f\|_m \|f\|_{k+a} \leq C \|f\|_{m+a+r} \|f\|_{k+r}.$$

Proof. This follows from repeating the proofs of Proposition 4.3.4 and Corollary 4.3.5, using Lemma 4.3.6 instead of Lemma 4.3.3. ☐

4.4 Tameness of vector bundles and bundle operators

In this section we prove that the space of real-valued functions on a compact manifold M is tame. This will also imply that the space of smooth sections of a vector bundle is tame. Then we show that (nonlinear) vector bundle operators (see Definition 2.5.19) are tame.

Observe that the seminorms defined in Definition 2.5.12 have a natural grading. In order to show tameness of the space of sections, we start by asserting that fiberwise linear maps induce tame linear operators.

Proposition 4.4.1. *Let $E, E' \twoheadrightarrow M$ be vector bundles over a compact manifold M and $\phi \in \Gamma(\mathrm{Hom}(E, E'))$ a fiberwise linear map. Then $\phi_* : \Gamma(E) \to \Gamma(E')$ defined by:*

$$(\phi_* s)(p) = \phi \circ s(p)$$

is a tame linear map with base and degree zero.

Proof. We argue in the local case. Let $f : U \to \mathbb{R}^n$ be a smooth map and $A : U \to \mathbb{R}^{m\times n}$ a smooth map of matrices. Then note that:

$$|\partial^\alpha A(p)f(p)| \leq C \sum_{\beta\leq\alpha} |(\partial^\beta A)(p)||(\partial^{\alpha-\beta} f)(p)|,$$

with α and β multi-indices. On any compact set $K \subseteq U$ we have that $\partial^\beta A$ is bounded. Hence for $\|\ldots\|_{K,k}$ as defined in Definition 2.5.2 we see that:

$$\|Af\|_{K,k} \leq C \|f\|_{K,k}.$$

This holds in particular if we take U compact in the first place. By a partition of unity argument over trivializible neighborhoods, the same holds for smooth manifolds.

Since the proof of Proposition 2.5.20 shows that the derivative of any bundle operator is again a bundle operator, we see that all derivatives of ϕ_* are tame as well. ☐

Now we use the Section 4.1 and the Lemma 4.1.10 to show the first major result of this section:

Theorem 4.4.2. *If M is a compact manifold then $C^\infty(M)$ is tame.*

Proof. We will prove that $C^\infty(M)$ is a tame direct summand of the space $L_1(\mathbb{R}^d, d\xi, w)$ with Lebesgue measure $d\xi$ and weight function $w = \log(1 + |\xi|)$. Then by Lemma Lemma 4.1.10 and Section 4.1 it follows that $C^\infty(M)$ is tame.

First of all we embed M into $\mathbb{R}^d$ for some d (possible if $d \geq 2 \dim M + 1$ by the Whitney embedding theorem, see [GG74], p.62) and since M is compact the embedding will lie in some large ball B. We will write $C_0^\infty(\mathbb{R}^d)$ for the Fréchet space of smooth functions whose derivatives all tend to zero at infinity with graded norms:

$$\|f\|_n = \sup_{|\alpha| \leq n} \sup_M |\partial^\alpha f(x)| \,.$$

Let $C_0^\infty(B)$ be the subspace of these functions that vanish outside the ball B. Using a tubular neighborhood and bump function supported in the neighborhood equalling one on M, we can define an extension operator $\epsilon : C^\infty(M) \to C_0^\infty(B)$ such that $\|\epsilon f\|_n \leq C_n \|f\|_n$ (where the constant C_n depends only on the bump function and its derivatives of order $\leq n$). Additionally, we have tame inclusion and restriction maps i, ρ respectively:

$$C^\infty(M) \xrightarrow{\epsilon} C_0^\infty(B) \xrightarrow{i} C_0^\infty(\mathbb{R}^d) \xrightarrow{\rho} C^\infty(M).$$

We can factor i through the Fourier transform by:

$$\begin{array}{ccccc} & & L_1(\mathbb{R}^d, d\xi, \log(1+|\xi|)) & & \\ & \overset{\mathcal{F}}{\nearrow} & & \overset{\mathcal{F}^{-1}}{\searrow} & \\ C_0^\infty(B) & & \xrightarrow[i]{\hspace{6em}} & & C_0^\infty(\mathbb{R}^d) \end{array} .$$

$\mathcal{F}$ is well-defined since functions in $C_0^\infty(B)$ are Schwarz functions. For $f \in L_1(\mathbb{R}^d, d\xi, \log(1 + |\xi|))$, we have that $(\xi \mapsto \xi^\alpha f(\xi)) \in L^1(\mathbb{R}^d)$ for any multi-index α. The operation $\mathcal{F}^{-1}$ maps $\xi^\alpha f(\xi)$ to some constant (of norm one) times $\partial^\alpha \mathcal{F}(f)$, which by the Riemann-Lebesgue lemma tends to zero at infinity. Therefore also $\mathcal{F}^{-1}$ with domain and codomain as in the diagram above is well-defined.

Now denote $\|\ldots\|_\infty$ for the supremum over $\mathbb{R}^d$ and $\|\ldots\|'_n$ for the seminorms of $L_1(\mathbb{R}^d, d\xi, w)$. We see that:

$$|\xi^\alpha \mathcal{F}(f)(\xi)| = |(\mathcal{F}(\partial^\alpha f))(\xi)| \leq \int_B |\partial^\alpha f| \, dx \leq C \|\partial^\alpha f\|_\infty \leq C \|f\|_{n+r} \,,$$

where the last inequality follows by taking the maximum over $|\alpha| \leq n + r$. Summing over the various cross-terms in the expansion of $(1 + |\xi|)^{n+r}$, (and after some straightforward estimations) we get:

$$\left\|(1 + |\xi|)^{n+r} \mathcal{F}(f)\right\|_\infty \leq C \|f\|_{n+r} \,, \tag{4.1}$$

for some larger C, from which it follows that:

$$\|\mathcal{F}(f)\|'_n = \int_{\mathbb{R}^d} (1 + |\xi|)^n |\mathcal{F}(f)| d\xi \leq \left(\int_{\mathbb{R}^d} \frac{1}{(1 + |\xi|)^r} d\xi \right) C \|f\|_{n+r} \,,$$

where the integral is finite when $r > d$. For $\mathcal{F}^{-1}(f)$ we observe that:

$$\sup_{\mathbb{R}^d} |\partial^\alpha \mathcal{F}^{-1}(f)| \leq \int_{\mathbb{R}^d} |\xi^\alpha||f(\xi)| d\xi \leq C \int_{\mathbb{R}^d} (1 + |\xi|)^n |f(\xi)| d\xi \tag{4.2}$$

from which it follows that $\left\|\mathcal{F}^{-1}(f)\right\|_n \leq C \|f\|'_n$.

From the factorisations in equations 4.1 and 4.2, we see that $C^\infty(M)$ is a tame direct summand of $L_1(\mathbb{R}^d, d\xi, \log(1+|\xi|))$. ☐

A well-known result in differential geometry states that for any vector bundle E over M there is another vector bundle E' (also over M) such that $E \oplus E' \cong M \times \mathbb{R}^k$ for some $k \in \mathbb{N}$. This yields the following:

Corollary 4.4.3. *Let E be a vector bundle over a compact manifold M, then $\Gamma(E)$ is tame.*

Proof. First remark that tameness of $\mathcal{C}^\infty(M)^k = \mathcal{C}^\infty(M, \mathbb{R}^k)$ follows immediately from Theorem 4.4.2. As remarked above there are fiberwise linear maps:

$$\phi_L : E \hookrightarrow E \oplus 0 \subseteq E \times E' \cong M \times \mathbb{R}^k$$
$$\phi_R : M \times \mathbb{R}^k \cong E \oplus E' \to E$$

such that $\phi_R \circ \phi_L = \mathrm{Id}_E$. These induce vector bundle operators L, R such that $\Gamma(E) \xrightarrow{L} \mathcal{C}^\infty(M)^k \xrightarrow{R} \Gamma(E)$, which by Proposition 4.4.1 and Lemma 4.1.10 shows the result. □

Finally we argue that more generally, nonlinear bundle operators as defined in Definition 2.5.19 are also tame (as defined in Definition 4.2.1).

Proposition 4.4.4. *Let $E, E' \twoheadrightarrow M$ be vector bundles over M, $V \subseteq E$ open and $\phi : V \to E$ a (not necessarily linear) bundle map. Then the bundle operator $\phi_* : \Gamma(V) \to \Gamma(E')$ defined in Definition 2.5.19 is a smooth tame map.*

Sketch of proof. We treat the local one-dimensional case as an illustration, the general case being computationally very involved. Let I be a compact interval in $\mathbb{R}$, $V \subseteq I \times \mathbb{R}$ open and $\phi : V \to \mathbb{R}$ a smooth function. Then we can write:

$$\left(\frac{d}{dx}\right)^n \phi(x, f(x)) = \sum_{(0 \le k \le n)} \sum_{(j + i_1 + \ldots + i_k = n)} \partial_1^j \partial_2^k \phi(x, f(x)) \frac{d^{i_1} f}{dx} \cdots \frac{d^{i_k} f}{dx},$$

where ∂_1 and ∂_2 denote the derivatives of ϕ with respect to the first and second coordinates respectively.

Now consider a neigborhood in $\Gamma(V)$ of the form:

$$U = \{f \in \Gamma(V) \mid \|f - f_0\|_{\max b, r} < \delta\},$$

where we take norms as in Definition 2.5.2. The boundedness of any f in this neighborhood ensures that for $j, k \le n$ the first term in the above is bounded:

$$|\partial_1^j \partial_2^k \phi(x, f(x))| < C_n = C,$$

with the constant depending on n. Hence we obtain:

$$\|\phi_* f\|_n \le C \sum_{1 \le k \le n} \sum_{i_1 + \ldots + i_k \le n} \|f\|_{i_1} \cdots \|f\|_{i_k}$$

Since I is compact, we may apply Theorem 4.4.2 to obtain that $\mathcal{C}^\infty(I)$ is tame. Now we can apply Corollary 4.3.7 to see that for any $f \in U$:

$$\|f\|_i^n \le C \|f\|_{n+r}^i \|f\|_r^{n-i} \le C \|f\|_{n+r}^i$$

for any $i \ge b$. For $i \le b$, we may directly estimate $\|f\|_i \le \|f\|_b$. So we see that:

$$\|\phi_* f\|_n \le C \sum \|f\|_{n+r}^\theta,$$

with $\theta \le 1$. Then either $\|f\|_{n+r}^\theta \le 1$ or $\|f\|_{n+r}^\theta \le \|f\|_{n+r}$ so $\|\phi_* f\|_n \le C(1 + \|f\|_{n+r})$, and we see that ϕ_* satisfies a tame estimate as in Definition 4.2.1.

For passing to manifolds one may use partitions of unity, noting that multiplication by a bump function does not change the base or degree by Proposition 4.4.1. Finally, the proof of Proposition 2.5.20 shows that the derivative of a bundle operator is once again a bundle operator, from which it follows by induction that all derivatives of ϕ_* are smooth and tame. □

Chapter 5

The Nash-Moser theorem

In this technical chapter we will prove Hamilton's Nash-Moser theorem, published in [Ham82], in detail. Note that by the final result of Chapter 4, this applies result applies to vector bundle operators.

5.1 Statement of the theorem

The statement of the main theorem is the following:

Theorem 5.1.1 (Hamilton's Nash-Moser Theorem). *Let F and G be tame spaces and $P : (U \subseteq F) \to G$ a smooth tame map such that the derivative:*

$$DP : U \times F \to G$$

and its family of linear inverses:

$$(DP)^{-1} : U \times G \to F$$

are smooth and tame. Then P is locally invertible and its inverse is smooth and tame.

Remark 5.1.2. Note the distinction between $D(P^{-1})$, the derivatives of the inverse of P, and $(DP)^{-1}$, by which we mean the family of inverses to DP. △

5.2 Reducing to $F = \Sigma(B)$ and normalising tame estimates

We first reduce Theorem 5.1.1 to the setting of the standard space $\Sigma(B)$ with B a Banach space, defined in Definition 4.1.3.

Lemma 5.2.1. *Without loss of generality, we may take $F = G = \Sigma(B)$ for some Banach space B.*

Proof. Since F is a tame space, it is a tame direct summand of $\Sigma(B)$ for some Banach space B. Then by Lemma 4.1.8, we have that $F \times H$ is tamely isomorphic to $\Sigma(B)$ for some tame space H. Then note that $P : U \to G$ satisfies the conditions of the Nash-Moser theorem if and only if $\tilde{P} : U \times H \to U \times H$ given by $(f, h) \mapsto (P(f), h)$ does. □

We know that the maps P, DP, D^2P and $(DP)^{-1}$ all satisfy tame estimates on U (by restricting U if necessary). In order to keep notation manageable, we show that we can take the degrees to be either 0, r or $2r$ for some $r > 0$ and the base to be zero. To this end we introduce the following operator on $\Sigma(B)$ and $\Sigma(C)$:

Definition 5.2.2. *On a space $\Sigma(B)$ with B some Banach space, we define the* ***shift operator*** *$\nabla^p : \Sigma(B) \to \Sigma(B)$ by:*

$$\nabla^p : f = (f_k)_k \mapsto (e^{pk} f_k)_k,$$

Where $p \in \mathbb{Z}$.

The motivation for this operator is that it satisfies $\|\nabla^p f\|_n = \|f\|_{n+p}$ for $p \geq -n$ and $\|\nabla^p f\|_n \leq \|f\|_0$ for $p \leq -n$, so in general:

$$\|\nabla^p f\|_n \leq \|f\|_{n+p}\,,$$

where we quietly take $n+p$ to be 0 whenever it is negative in order not to complicate notation. In particular, the operator ∇^p is a tame isomorphism.

Suppose we have a tame map Q satisfying the tame estimates:

$$\|Q(f)\|_n \leq C(1 + \|f\|_{n+s})$$

for $n \geq b$ on some neighborhood of zero that we may assume to be of the form $\{\|f\|_a \leq \delta\}$ for some a and δ. Observe that $\tilde{Q} = \nabla^q \circ Q \circ \nabla^{-p}$ satisfies:

$$\left\|\tilde{Q}(f)\right\|_n = \left\|(\nabla^q \circ Q \circ \nabla^{-p})(f)\right\|_{n+q} \leq C(1 + \|f\|_{n+q-p+s})$$

for $n+q \geq b$ and $\|f\|_{a-p} \leq \delta$. Choosing p and q large, we see that our new map $\tilde{Q}$ satisfies a tame estimate of base zero on a neighborhood of the form $\{\|f\|_0 \leq \delta\}$. We may furthermore choose the difference $p - q$ such that the degree of the tame estimate is zero, since we may always choose p and q larger.

We apply this to the maps we want to consider: P, DP, D^2P and $(DP)^{-1}$. By choosing p and q large enough and setting $\tilde{P} = \nabla^q \circ P \circ \nabla^{-p}$, we can ensure that $\tilde{P}$, $D\tilde{P}$, $D^2\tilde{P}$ and $D\tilde{P}^{-1}$ satisfy tames estimates of base zero on some neighborhoord $\{\|f\|_0 \leq \delta\}$. By precomposing with the map $f \mapsto \frac{1}{2}\delta f$ we can take the neighborhood to be of the form $\{\|f\|_0 \leq 1\}$. Note that this doesn't affect the (left or right) invertibility of the derivative.

Additionally we can choose the difference $p - q$ such that (at least) one of the degrees of the maps is zero. We choose $D\tilde{P}^{-1}$ for this. Then we may take the other degrees to be r or $2r$ by increasing if necessary (because of the graded Fréchet structure).

Since by translation we may always assume that for $P : U \to \Sigma(B)$ we have $0 \in U$ and $P(0) = 0$, this argument shows that we can take estimates of the kind:

Lemma 5.2.3 (Normalized estimates)**.** *We may assume that the maps P, DP, D^2P and $(DP)^{-1}$ satisfy the estimates:*

$$\begin{aligned}
\|P(f)\|_n &\leq C\,\|f\|_{n+2r}\\
\|DP(f \mid h\|)_n &\leq C(\|f\|_{n+2r}\,\|h\|_r + \|h\|_{n+r})\\
\left\|DP^{-1}(f \mid k)\right\|_n &\leq C(\|f\|_{n+2r}\,\|k\|_0 + \|k\|_n)\\
\left\|D^2P(f \mid h_1, h_2)\right\|_n &\leq C(\|f\|_{n+2r}\,\|h_1\|_r\,\|h_2\|_r + \|h_1\|_{n+r}\,\|h_2\|_r + \|h_1\|_r\,\|h_2\|_{n+r})
\end{aligned}$$

for all $n \geq 0$ and $\|f\|_{2r} \leq 1$.

Proof. The last three inequalities are consequences of the argument above and Lemmas 4.2.6 and 4.2.7. For the first, use $P(0) = 0$ to write:

$$\|P(f)\|_n = \left\|\int_0^1 DP(tf)f\,dt\right\|_n \leq C(\|f\|_{n+r} + \|f\|_{n+2r}\,\|f\|_r),$$

using the second inequality in the statement above. Now $\|f\|_{n+r} \leq \|f\|_{n+2r}$ and on the neighborhood $\{\|f\|_{2r} \leq 1\}$, we have $\|f\|_r \leq 1$ so the above is estimated by $C\,\|f\|_{n+2r}$ as desired. □

5.3 Local injectivity

In this section we only require that DP^{-1} is a linear tame left inverse of DP, still satisfying the estimates in Lemma 5.2.3. We then obtain the following:

Proposition 5.3.1. *There exists a $1 \geq \delta > 0$ such that for $\|f_1\|_{2r}, \|f_2\|_{2r} < \delta$ we have:*

$$\|f_2 - f_1\|_0 \leq C\,\|P(f_2) - P(f_1)\|_0$$

This implies that P is injective on $\{\|f\|_{2r} \leq \delta\}$.

Proof. We require $\delta \leq 1$ in advance. Using Taylor's theorem (Theorem 2.4.14), write:

$$DP(f_1 \mid f_2 - f_1) = P(f_2) - P(f_1) - \int_0^1 (1-t)D^2P(f_t \mid f_2 - f_1, f_2 - f_1)dt,$$

where $f_t = tf_2 + (1-t)f_1$. Applying the inverse $(DP)^{-1}$ and denoting the left-hand side with K, we obtain:

$$\|f_2 - f_1\|_0 = \left\|(DP)^{-1}(f_1 \mid K)\right\|_0 \leq C(\|K\|_0 + \|f_1\|_{2r}\|K\|_0) \leq C\|K\|_0,$$

by Lemma 5.2.3 and our assumption $\|f_1\|_{2r} < \delta \leq 1$. Continuing our estimation, we find:

$$\begin{aligned}\|f_2 - f_1\|_0 &\leq C\left\|P(f_2) - P(f_1) - \int_0^1 (1-t)D^2P(f_t \mid f_2 - f_1, f_2 - f_1)dt\right\|_0 \\ &\leq C\|P(f_2) - P(f_1)\|_0 + \int_0^1 C(\|f_t\| + 2)\|f_2 - f_1\|_r^2 dt \leq C\|P(f_2) - P(f_1)\|_0 + C\|f_2 - f_1\|_r^2\end{aligned}$$

using the estimate for D^2P in Lemma 5.2.3. By interpolation, we have:

$$\|f_2 - f_1\|_r^2 \leq C\|f_2 - f_1\|_{2r}\|f_2 - f_1\|_0 \leq 2C\delta\|f_2 - f_1\|_{2r},$$

so the final result becomes:

$$\|f_2 - f_1\|_0 \leq C\|P(f_2) - P(f_1)\|_0 + 2C\delta\|f_2 - f_1\|_0.$$

Choosing δ sufficiently small yields the required result. Note that since the since the seminorm $\|\ldots\|_0$ is a norm, we have that $P(f_2) = P(f_1)$ implies $f_2 = f_1$, showing injectivity. ☐

5.4 Local surjectivity

Our approach will be motivated by the following: suppose $P : B \to B$ were a map on a Banach space B instead, again with $P(0) = 0$. For some g near zero, we wish to find an f_∞ such that $P(f_\infty) = g$. Let $\alpha > 0$ be some arbitrary constant. Then for $f : [0, \infty) \to B$ satisfying:

$$\frac{d}{dt}P(f(t)) = DP(f(t))f'(t) = \alpha(g_0 - P(f(t))),$$

the path $t \to P(f(t))$ is a straight line from zero to g_0. If $f(t)$ were to converge to some f_∞ as $t \to \infty$, and therefore $f'(t)$ would go to zero, then we see from the above that $g = P(f_\infty)$. Such an f as above is a solution of the ODE:

$$f'(t) = \alpha(DP)^{-1}(f(t))(g - P(f(t))), \tag{5.1}$$

which in Banach spaces always has solution on small enough intervals. On Banach spaces we can also restrict to a neighborhood such that the operator norm of $(DP)^{-1}(u)$ is bounded for $u \in U$. Note that for the "error" at time t, given by $E(t) = g - P(f(t))$, we have $E'(t) = -\alpha E(t)$, showing that it decays exponentially. Together with the bound on the operator $(DP)^{-1}$ (such that $\alpha\left\|(DP)^{-1}\right\| < 1$), this allows us to show convergence of $f(t)$.

In Fréchet spaces, the ODE above may not always a have a solution (see [Ham82], p. 129). It is also not clear how to obtain a bound on $(DP)^{-1}$: one may have an estimate such as in Lemma 5.2.3, but the constant $C = C_n$ may increase with n and therefore $\alpha C_n < 1$ may not hold for all n. It is, however, possible to *modify* Equation (5.1) such that local solutions exist. This result relies on Appendix B.3, where it is shown that ODEs in Fréchet space that factor through Banach space have local solutions.

To obtain local surjectivity we will show that the following ODE:

$$f'(t) = \alpha(DP)^{-1}(S_t f_t \mid S_t(g - P(f_t))), \tag{5.2}$$

with $\alpha > 0$ some constant, has a solution such that $g - P(f(t)) \to 0$ as $t \to \infty$.

In order to show the existence of a solution to the ODE, we introduce the following two functions:

$$Q:\begin{cases}(f,t) & \mapsto (S_t(f), S_t(g-P(f)))\\ \Sigma(B)\times\mathbb{R} & \to \Sigma(B)\times\Sigma(C)\end{cases}$$
$$R:\begin{cases}(g,h) & \mapsto \alpha(DP)^{-1}(g\mid h)\\ \Sigma(B)\times\Sigma(C) & \to \Sigma(B)\end{cases},$$

such that the ODE in Equation (5.2) becomes $f'(t)=R(Q(f))$. Note that when we restrict Q to $\Sigma(B)\times[0,T]$ for any (finite) time T, its image lies in the space:

$$\Sigma_T(B)=\{f\in\Sigma(B)\mid f_k=0 \text{ for all } k>T\}$$

which (with the subspace topology induced from $\Sigma(B)$) is a Banach space. Therefore, by Theorem B.3.3 we see that Equation (5.2) has a solution for $t\in[0,\epsilon)$, where $\epsilon>0$ is some constant.

As for bounding f by the input g, we will seek to prove the following result:

Theorem 5.4.1. *Let $f:[0,T)\to\Sigma(B)$ be a solution to Equation (5.2). Then there is $1\geq\delta>0$ such that for $g\in\Sigma(B)$ with $\|g\|_{2r}\leq\delta$ we have:*

$$\int_0^T \|f'(\tau)\|_n\,d\tau \leq C\,\|g\|_n\,, \tag{5.3}$$

with constant C independent of T and for all n greater than some constant [1].

This will give us what we need by the following proposition.

Proposition 5.4.2. *Suppose f is a solution of the ODE in Equation (5.2) on $[0,T_0)$ and satisfies the inequality in Equation (5.3) above for some C independent of T and f. Then f may be extended to a solution on $[0,\infty)$. Furthermore, $f(t)\to f_\infty$ for some f_∞ as $t\to\infty$ and:*

$$\lim_{t\to\infty}P(f(t))=P(f_\infty)=g$$

Proof. Suppose $[0,T_0)$ is the maximal interval on which f is a solution and can be defined. Since the constant in Equation (5.3) is independent of T, we see the estimate still holds for $T=T_0$ by taking the limit. Then we see that:

$$\lim_{t_1,t_2\to T_0}\|f(t_1)-f(t_2)\|\leq\lim_{t_1,t_2\to T_0}\left|\int_{t_1}^{t_2}\|f'(\tau)\|\,d\tau\right|\to 0$$

This shows that $f(t)$ converges to some f_{T_0} as $t\to T$: remark that $(f(T_0-1/k))_{k\geq 1}$ is a Cauchy sequence. Now we may apply the Banach ODE theorem starting at f_{T_0} to extend our solution, hence T_0 cannot be maximal.

It follows that the solution f can be extended to $[0,\infty)$. In particular, the bound of equation Equation (5.3) holds for any $T\in[0,\infty)$ with the constants C and δ independent of T. This implies that $\int_0^\infty\|f'(t)\|_n\,dt\leq C\,\|g\|_n$. Therefore $f'(t)\to 0$ as $t\to\infty$ and by a similar argument to the above $f(t)\to f_\infty$ for some f_∞. Taking $t\to\infty$ in our ODE yields:

$$0=\lim_{t\to\infty}f'(t)=\lim_{t\to\infty}\alpha(DP)^{-1}(S_tf_t\mid S_t(g-P(f_t)))=(DP)^{-1}(f_\infty\mid g-P(f_\infty)),$$

where we used Lemma 4.3.2. By inverting this expression we get $g-P(f_\infty)=0$, as required. □

Remark 5.4.3. From now on in this section, $f:[0,T)\to\Sigma(B)$ **will always denote the (hypothetical) solution of ODE Equation (5.2)**. △

[1] This constant will turn out to be $2r$.

5.4.1 Estimation of ODEs

We rewrite Equation (5.2) in a perhaps more well-known form:

Lemma 5.4.4. *Let $f : \mathbb{R} \to \Sigma(B)$ satisfy Equation* (5.2) *and define:*

$$E(t) = g - P(f(t))$$
$$l(t) = \Big(DP(S(t)f(t)) - DP(f(t))\Big)f'(t).$$

Then $E, l : \mathbb{R} \to \Sigma(B)$ satisfy:

$$E'(t) + \alpha S(t)E(t) = l(t), \tag{5.4}$$

We will refer to $E(t)$ as the ***error term****.*

Proof. The proof is a straightforward calculation. □

In this subsection we will find estimates for general solutions to the ODE in Equation (5.4). These will be used in the following section to finally obtain the estimate in Equation (5.3).

We remind the reader of Definition 4.3.1, where the smoothing operator $S(t)$ was defined using a non-decreasing bump function $\psi : \mathbb{R} \to [0, 1]$ such that $\psi\big((-\infty, 0]\big) = \{0\}$ and $\psi\big([1, \infty)\big) = \{1\}$. Given this bump function, the smoothing operator was set to:

$$\big(S(t)\,(f)\,\big)_k = \psi(t-k)f_k$$

for $f = (f_k)_k \in \Sigma(B)$. Therefore the equation above amounts to a set of equations $E_k'(t) + \alpha\psi(t-k)E_k(t) = l_k(t)$ for $k \in \mathbb{N}$. This linear ODE has solution:

$$E(t)_k = e^{-\alpha\int_0^t \psi(\tau-k)d\tau}E(0)_k + \int_0^t e^{-\alpha\int_\theta^t \psi(\tau-k)d\tau}l_k(\theta)d\theta,$$

as shown in a standard differential equations course.

We will only consider $t \geq 0$ and $\alpha > 0$. In this case the expression above is well-defined as an element of $\Sigma(B)$, for any $t \in \mathbb{R}_{\geq 0}$. We will write:

$$A(\theta, t)_k = e^{-\alpha\int_\theta^t \psi(\tau-k)}d\tau,$$

such that $E(t)_k = A(0,t)_k E(0)_k + \int_0^t A(\theta,t)_k l(\theta)_k d\theta$. This coefficient satisfies the following:

Lemma 5.4.5. *For $0 < q < \alpha$ we have:*

$$\int_0^\infty e^{qt}A(\theta,t)dt \leq C_q(e^{qk} + e^{q\theta}),$$

where the constant C_q does not depend on k and θ.

Proof. If $\theta \geq k$ then $\int_\theta^t \psi(\tau-k)d\tau + 1 \geq t - \theta$, so we can estimate:

$$e^{qt}A(\theta,t)_k \leq Ce^{qt+\alpha(\theta-t)} = Ce^{\alpha\theta}e^{(q-a)t}.$$

This estimate shows that we may integrate the right hand side from θ to ∞, which gives:

$$\int_\theta^\infty e^{qt}A(\theta,t)_k dt \leq Ce^{\alpha\theta}\int_\theta^\infty e^{(q-\alpha)t}dt = Ce^{\alpha\theta}Ce^{(q-\alpha)\theta} = Ce^{q\theta}$$

In case $\theta \leq k$ we split the integral. In this case, we have $\int_\theta^t \psi(\tau-k)d\tau + 1 \geq t-k$ and therefore $A(\theta,t)_k \leq Ce^{\alpha(k-t)}$. We may take the following integral and estimate it:

$$\int_k^\infty e^{qt}A(\theta,t)_k dt \leq Ce^{qk}.$$

Finally, use $A(\theta,t) \leq 1$ to obtain:

$$\int_\theta^k e^{qt}A(\theta,t)_k \leq \int_0^k e^{qt}dt \leq qe^{qk} = C_q e^{qk}.$$

Adding these estimates gives the result, where the the constant C_q depends only on q. □

Theorem 5.4.6. *If $E : \mathbb{R} \to \Sigma(B)$ satisfies Equation (5.4), then for all $p \geq 0$, $0 < q < \alpha$ and $T \geq 0$ we have:*

$$\int_0^T e^{qt} \left\|E(t)\right\|_p dt \leq C \left\|E(0)\right\|_{p+q} + C \int_0^T e^{qt} \left\|l(t)\right\|_p + \left\|l(t)\right\|_{p+q} dt.$$

where the constant C is independent of T.

Proof. We use the known expression for the solution of Equation (5.4) to obtain:

$$e^{qt} E(t)_k = e^{qt} A(0,t)_k E(0)_k + e^{qt} \int_0^t A(\theta,t)_k l(\theta)_k d\theta. \tag{5.5}$$

By Lemma 5.4.5, we estimate the first term by:

$$\int_0^T e^{qt} A(0,t)_k \left\|E(0)_k\right\| dt \leq C(e^{qk}+1) \left\|E(0)_k\right\| = Ce^{qk} \left\|E(0)_k\right\|.$$

The "superexponential" convergence of the series $\sum_k A(0,t)_k$ and $\sum_k E(0)_k$ allows summation to obtain the following estimate for any $p \geq 0$:

$$\int_0^T e^{qt} \left\|A(0,t)E(0)\right\|_p dt = \sum_{k=0}^{\infty} e^{pk} \int_0^T e^{qt} A(0,t)_k \left\|E(0)_k\right\| dt \leq C \sum_{k=0}^{\infty} e^{pk} e^{qk} E(0)_k = C \left\|E(0)\right\|_{p+q}.$$

We estimate the second term in Equation (5.5) in much the same way, first reparametrizing the domain of integration:

$$\begin{aligned}
\int_{t=0}^T e^{qt} \int_{\theta=0}^t \left\|A(\theta,t)l(\theta))\right\|_p d\theta dt &= \int_{\theta=0}^T \int_{t=\theta}^T e^{qt} \left\|A(\theta,t)l(\theta))\right\|_p dt d\theta \\
&\leq \int_{\theta=0}^T \sum_k e^{pk} C(e^{qk} + e^{q\theta}) \left\|l(\theta)_k\right\| d\theta \\
&\leq C \int_{\theta=0}^T e^{q\theta} \left\|l(\theta)\right\|_p + \left\|l(\theta)\right\|_{p+q} d\theta.
\end{aligned}$$

□

5.4.2 A priori estimates

In the following we will assume that f satisfies Equation (5.2) and that $\left\|f(t)\right\|_{2r} \leq 1$ for t in some interval of the kind $[0, T)$. By this "a priori" estimate, we will show that for g small enough the solution f of Equation (5.2) will remain in the set $\{\left\|\ldots\right\|_{2r} \leq 1\}$ even for $t > T$ (whenever it is defined there for such t).

The main idea is to use theorem Theorem 5.4.6 of the previous subsection to show that f cannot escape a certain domain. In order to do this, we need to estimate $l(t)$.

We start off by estimating the derivative of f by the error term:

Lemma 5.4.7. *For f satistying the ODE (Equation (5.2)) we have:*

$$\left\|f'(t)\right\|_{n+q} \leq Ce^{qt} \left(\left\|E(t)\right\|_n + \left\|f(T)\right\|_{n+2r} \left\|E(t)\right\|_0\right)$$

for all $n, q \geq 0$.

Proof. We have $f(t) = (DP)^{-1}(S_t(f(t)) \mid S_t E(t))$, which by our normalised estimates (Lemma 5.2.3) and Lemma 4.3.3 means:

$$\left\|f'(t)\right\|_{n+q} \leq C(\left\|S_t(f(t))\right\|_{n+q+2r} \left\|S_t(E(t))\right\|_0 + \left\|S_t E(t)\right\|_{n+q}) \leq e^{qt} C(\left\|f(t)\right\|_{n+2r} \left\|E(t)\right\|_0 + \left\|E(t)\right\|_n).$$

□

The following special case is immediate from the above and will be used often:

Corollary 5.4.8. *For $\|f(t)\|_{2r} \leq 1$ and $n \in \mathbb{N}$*

$$\|f'(t)\|_n \leq Ce^{nt} \|E(t)\|_0 .$$

Now we estimate $l(t)$ in the alternate form of our ODE (Equation (5.4)):

Lemma 5.4.9. *For $n \in \mathbb{N}$:*

$$\|l(t)\|_n \leq C \|f(t)\|_{n+2r} \|E(t)\|_0 .$$

Proof. For fixed t, we write:

$$l(t) = \Big(DP(S_t(f(t))) - DP(f(t)\Big)f'(t)) = \int_0^1 D^2P(\tau S_t(f(t)) + (1-\tau)f(t) \mid f'(t), (S_t - 1)f(t))d\tau$$

Using the normalized estimates of Lemma 5.2.3 we obtain:

$$\begin{aligned} \|l(t)\|_n \leq C\big(& \|f'(t)\|_r \|(1-S_t)f(t)\|_{n+r} + \|f'(t)\|_{n+r} \|(1-S_t)f(t)\|_r + \\ & (\|S_t(f(t))\|_{n+2r} + \|f(t)\|_{n+2r}) \|f'(t)\|_r \|(1-S_t)f(t)\|_r \big). \end{aligned}$$

We estimate the first two terms by the previous corollary and Lemma 4.3.3:

$$\begin{aligned} \|f'(t)\|_r \|(1-S_t)f(t)\|_{n+r} &\leq Ce^r \|E(t)\|_0 e^{-r} \|f(t)\|_{n+2r} \\ \|f'(t)\|_{n+r} \|(1-S_t)f(t)\|_r &\leq Ce^{n+r} \|E(t)\|_0 e^{-(n+r)} \|f(t)\|_{n+2r} . \end{aligned}$$

For the final term we note that $\|S_t(f(t))\|_{n+2r} \leq \|f(t)\|_{n+2r}$ and (similarly to the above) we can take:

$$\|f'(t)\|_r \|(1-S_t)f(t)\|_r \leq C \|E(t)\|_0 \|f(t)\|_{2r} \leq C \|E(t)\|_0 ,$$

using $\|f(t)\|_{2r} \leq 1$. □

Combining Corollary 5.4.8 and Lemma 5.4.9, we get the following estimation of $l(t)$ entirely in terms of the error:

Corollary 5.4.10. *For all $n \in N$ we have:*

$$\|l(t)\|_n \leq C \|E(t)\|_0 \int_0^t e^{(n+2r)\tau} \|E(\tau)\|_0 d\tau.$$

Proof. This follows from:

$$\begin{aligned} \|l(t)\|_n &\leq C \|E(t)\|_0 \|f(t)\|_{n+2r} \\ &\leq C \|E(t)\|_0 \int_0^t \|f'(\tau)\|_{n+2r} d\tau \end{aligned}$$

and Corollary 5.4.8. □

Proposition 5.4.11 (Barrier estimate)**.** *Denoting*

$$K_T := \int_0^T e^{2rt} \|E(t)\|_0 dt,$$

we have:

$$K_T \leq C(\|g\| + K_T^2).$$

This implies:

$$K_T \leq 2C \|g\|_{2r} \text{ or } K_T > \frac{1}{2C}.$$

Proof. Our estimate for solutions of first order ODEs (Theorem 5.4.6) gives us:

$$\int_0^T e^{2rt} \|E(t)\|_0 \, dt \leq C \|g\|_{2r} + C \int_0^T e^{2rt} \|l(t)\|_0 + \|l(t)\|_{2r} \, dt.$$

By the corollary above, we obtain:

$$\begin{aligned}
\int_0^T e^{2rt} \|l(t)\|_0 \, dt &\leq C \int_0^T e^{2rt} \|E(t)\|_0 \int_0^t e^{2r\tau} \|E(\tau)\|_0 \, d\tau dt \\
&\leq C(K_T)^2, \\
\int_0^T \|l(t)\|_{4r} \, dt &\leq C \int_0^T \|E(t)\|_0 \int_0^t e^{4r\tau} \|E(\tau)\|_0 \, d\tau dt \\
&\leq C \int_0^T \|E(t)\|_0 \, e^{2rt} \int_0^t e^{2r\tau} \|E(\tau)\|_0 \, d\tau dt \\
&\leq C(K_T)^2
\end{aligned}$$

For the final statement, note that $K_T(1 - CK_T) \leq C \|g\|_{2r}$. If $K_T \leq 1/(2C)$ then:

$$\tfrac{1}{2} K_T \leq K_T(1 - CK_T) \leq C \|g\|_{2r} .$$

□

Corollary 5.4.12. *If $\|g\|_{2r} \leq \delta$ for δ sufficiently small we have:*

$$K_T = \int_0^T e^{2rt} \|E(t)\|_0 \, dt \leq C \|g\|_{2r} .$$

Proof. We may choose $\|g\|_{2r}$ small enough such that $2C \|g\|_{2r} < 1/(2C)$. Now K_T is continuous in T and $K_0 = 0$, which implies that K_T cannot cross the region $[C \|g\|_r , 1/C)$ (where we absorb the factor 2 in the constant C). □

Corollary 5.4.13. *For δ as above and $t \leq T$, we have:*

$$\|f(t)\|_{2r} \leq \int_0^t \|f'(\tau)\|_{2r} \, d\tau \leq C \|g\|_{2r}$$

Proof. This follows from $\|f'(t)\|_{2r} \leq Ce^{2rt} \|E(t)\|_0$ and the corollary above. □

This shows that we may take g small in order for the solution f of the ODE to remain in $\{\|\ldots\|_{2r} \leq 1\}$ for all time. What remains is to show that:

$$\int_0^T \|f'(\tau)\|_n \, d\tau \leq C \|g\|_n$$

for seminorms of order $n \geq 2r$. This will be done in theorem Theorem 5.4.15 below, which uses the following lemma in its proof for the induction step:

Lemma 5.4.14. *For δ sufficiently small and $\|g\|_{2r} \leq \delta$ we have:*

$$\int_0^T e^{(2r+1)t} \|E(t)\|_0 \, dt \leq C \|g\|_{2r+1}$$

Proof. Corollaries 5.4.10 and 5.4.12, with δ chosen appropriately, give us:

$$\begin{aligned}
\|l(t)\|_0 &\leq C \|E(t)\|_0 \int_0^t e^{2r\tau} \|E(\tau)\|_0 \, d\tau \leq C \|E(t)\|_0 \|g\|_{2r} \\
\|l(t)\|_{2r+1} &\leq C \|E(t)\|_0 \int_0^t e^{(4r+1)\tau} \|E(\tau)\|_0 \, d\tau \leq Ce^{(2r+1)t} \|E(t)\|_0 \|g\|_{2r}
\end{aligned}$$

The estimate for first-order ODEs estimate then becomes:

$$\int_0^T e^{(2r+1)t}\left\|E(t)\right\|_0 dt \le C\left\|g\right\|_{2r+1} + C\int_0^T e^{(2r+1)t}\left\|l(t)\right\|_0 + \left\|l(t)\right\|_{2r+1} dt$$
$$\le C\left\|g\right\|_{2r+1} + C\left\|g\right\|_{2r}\int_0^T e^{(2r+1)t}\left\|E(\tau)\right\|_0 dt$$

Now choose δ smaller if necessary such that $C\delta < 1$. Then $C\left\|g\right\|_{2r} < 1$ and we can subtract the second term of the right hand side above. □

Finally, the following theorem will immediately imply Proposition 5.4.2, concluding the proof of local surjectivity.

Theorem 5.4.15. *There exists δ such that if $\left\|g\right\|_{2r} \le \delta$ then for all $n \ge 2r$ and $q \ge 0$ we have:*

$$\int_0^T \left\|f'(t)\right\|_{n+q} dt \le Ce^{qT}\left\|g\right\|_n$$

Proof. We proceed by induction and write $n = p + 2r$ for $n \ge 2r$. For $p = 0$, Corollary 5.4.8 and Lemma 5.4.14 above give us the base case:

$$\int_0^T \left\|f'(t)\right\|_{1+q+2r} dt \le \int_0^T Ce^{(1+q+2r)t}\left\|E(t)\right\|_0 dt \le Ce^{qT}\int_0^T Ce^{(1+2r)t}\left\|E(t)\right\|_0 \le Ce^{qT}\left\|g\right\|_{2r+1}.$$

Now we assume that for $p \ge 1$ and any $q \ge 0$ we have:

$$\int_0^T \left\|f'(t)\right\|_{q+p+2r} dt \le Ce^{qT}\left\|g\right\|_{p+2r}.$$

We begin our estimation with the use of Lemma 5.4.7:

$$\int_0^T \left\|f'(t)\right\|_{q+p+2r+1} dt \le C\int_0^T e^{(q+2r+1)t}\left\|E(t)\right\|_p dt + C\int_0^T e^{(q+2r+1)t}\left\|f(t)\right\|_{p+2r}\left\|E(t)\right\|_0 dt. \tag{5.6}$$

We will estimate the two terms on the right hand side seperately. For the first term, Theorem 5.4.6 gives us:

$$\int_0^T e^{(2r+1)t}\left\|E(t)\right\|_p dt \le \left\|g\right\|_{p+2r+1} + \int_0^T e^{(2r+1)t}\left\|l(t)\right\|_p + \left\|l(t)\right\|_{p+2r+1} dt.$$

Now our induction hypothesis implies in particular that:

$$\left\|f(t)\right\|_{p+2r} \le C\left\|g\right\|_{p+2r} \text{ and } \left\|f(t)\right\|_{p+4r+1} \le Ce^{(2r+1)t}\left\|g\right\|_{p+2r},$$

which combined with Lemma 5.4.9 gives:

$$e^{(2r+1)t}\left\|l(t)\right\|_p + \left\|l(t)\right\|_{p+2r+1} \le Ce^{(2r+1)t}\left\|E(t)\right\|_0\left\|g\right\|_{p+2r}.$$

We integrate the right hand side and make use of Lemma 5.4.14 to obtain:

$$\left\|g\right\|_{p+2r}\int_0^T e^{(2r+1)t}\left\|E(t)\right\|_0 dt \le C\left\|g\right\|_{p+2r}\left\|g\right\|_{2r+1}.$$

This shows that we can estimate the first term in Equation (5.6) by:

$$\int_0^T e^{(q+2r+1)t}\left\|E(t)\right\|_p dt \le Ce^{qT}\left(\left\|g\right\|_{p+2r+1} + \left\|g\right\|_{p+2r}\left\|g\right\|_{2r+1}\right).$$

For the second term, note that $\left\|f_t\right\|_{p+2r} \le C\left\|g\right\|_{p+2r}$ by hypothesis and

$$\int_0^T e^{(2r+1)t}\left\|E(t)\right\|_0 dt \le C\left\|g\right\|_{2r+1}$$

by Lemma 5.4.14. Hence we have:

$$\int_0^T \left\|f'(t)\right\|_{q+p+2r+1} dt \le Ce^{qT}\left(\left\|g\right\|_{p+2r+1} + \left\|g\right\|_{p+2r}\left\|g\right\|_{2r+1}\right).$$

By interpolation (specifically Corollary 4.3.5), we can take:

$$\left\|g\right\|_{p+2r}\left\|g\right\|_{2r+1} \le C\left\|g\right\|_{p+2r+1}\left\|g\right\|_{2r},$$

and for $\left\|g\right\|_{2r} \le \delta \le 1$ we obtain the result. □

5.4.3 The local inverse and regularity

The previous section shows that for our map $P : (U \subseteq \Sigma(B)) \to \Sigma(B)$ there is a neighborhood V around zero such that for any $g \in V$ there is an f such that $P(f) = g$. Before we showed that we may restrict P to a neighborhood such that it is injective. This means that for some neighborhoods $U, V \ni 0$, the map $P : U \to V$ is bijective. For $g \in V$, we have the estimate

$$\left\| P^{-1}(g) \right\|_n \leq C \left\| g \right\|_n ,$$

showing tameness (and therefore continuity) of P^{-1}.

It remains to show that the inverse is *smooth* and tame (see Definition 4.2.5). We will need the following two results.

Lemma 5.4.16. *There exists a $1 \geq \delta > 0$ such that for $\|f_1\|_{2r}$, $\|f_0\|_{2r} < \delta$ we have:*

$$\|f_2 - f_1\|_n \leq C \|P(f_2) - P(f_1)\|_n + C(\|f_2\|_{n+2r} + \|f_1\|_{n+2r}) \|P(f_2) - P(f_1)\|_0$$

Proof. As in the proof of proposition Proposition 5.3.1, we write:

$$DP(f_1 \mid f_2 - f_1) = P(f_2) - P(f_1) - \int_0^1 (1-t) D^2 P(f_t \mid f_2 - f_1, f_2 - f_1) dt = K,$$

where $f_t = t f_2 - (1-t) f_1$. Applying the inverse $(DP)^{-1}$ and using the normalized estimates in Lemma 5.2.3, we obtain:

$$\left\| (DP)^{-1}(f_1 \mid K) \right\|_n \leq C(\|f_1\|_{n+2r} \|K\|_0 + \|K\|_n).$$

We proceed the estimation term by term. For $\|K\|_n$, we use Lemma 5.2.3 once again to obtain:

$$\begin{aligned} \left\| D^2 P(f_t \mid f_2 - f_1, f_2 - f_1) \right\|_n &\leq C \|f_2 - f_1\|_{n+r} \|f_2 - f_1\|_r + C \|f_t\|_{n+2r} \|f_2 - f_1\|_r^2 \\ &\leq \|f_2 - f_1\|_{n+2r} \|f_2 - f_1\|_0 + C(\|f_2\|_{n+2r} + \|f_1\|_{n+2r}) \|f_2 - f_1\|_{2r} \|f_2 - f_1\|_0 , \end{aligned}$$

where the last step follows from interpolation (Corollary 4.3.5). Now $\|f_2 - f_1\|_{2r} \leq 2$ and for δ as in Proposition 5.3.1 we have $\|f_2 - f_1\|_0 \leq C \|P(f_2) - P(f_1)\|_0$, which gives us:

$$\|K\|_n \leq C \|P(f_2) - P(f_1)\|_0 + C(\|f_2\|_{n+2r} + \|f_1\|_{n+2r}) \|P(f_2) + P(f_1)\|_0 .$$

For the second term, we note that from the above we have $\|f_2 - f_1\|_r^2 \leq C \|P(f_2) - P(f_1)\|_0$. As seen in the proof of Proposition 5.3.1 we have:

$$\|K\|_0 \leq C \|P(f_2) - P(f_1)\|_0 + C \|f_2 - f_1\|_r^2 ,$$

hence:

$$\|f_1\|_{n+2r} \|K\|_0 \leq C(\|f_2\|_{n+2r} + \|f_1\|_{n+2r}) \|P(f_2) - P(f_1)\|_0 .$$

Combining the estimates for the individual terms gives the result. □

Corollary 5.4.17. *Let δ as above and $n \geq m \geq 0$ integers. Then for f_1, f_2 with $\|f_1\|_{2r}, \|f_2\|_{2r} \leq \delta$ we have:*

$$\|f_2 - f_1\|_n \|f_2 - f_1\|_m \leq R_{n,m}(f_2, f_1) \|P(f_2) - P(f_1)\|_n^2 ,$$

with $R_{n,m} : \Sigma(B) \times \Sigma(B) \to [0, \infty)$ some continuous function such that:

$$\lim_{f_1, f_2 \to 0} R_{n,m}(f_1, f_2) = 0.$$

Proof. Estimate $\|f_2 - f_1\|_n$ and $\|f_2 - f_1\|_m$ using the Lemma 5.4.16 and multiply the expressions. Then use $\|\ldots\|_n \geq \|\ldots\|_m$. □

Proposition 5.4.18. *Let δ as in Proposition 5.3.1 and Theorem 5.4.1 (by taking the minimum δ), such that:*

$$P : U = \{\|\ldots\|_{2r} \leq \delta\} \to P(U) = V$$

is bijective. Then the inverse $P^{-1} : V \to U$ is a smooth tame map.

Proof. We claim that:

$$D(P^{-1})(g \mid k) = (DP)^{-1}(P^{-1}(g) \mid k),$$

where we recall that by $(DP)^{-1}(f \mid \ldots)$ we mean the inverse of the linear map $DP(f \mid \ldots)$. Since $(DP)^{-1}$ and P^{-1} are tame, this shows the derivative is tame. Furthermore, the derivative is written as a composition of the C^1 map P^{-1} and the smooth map $(DP)^{-1}$, hence the expression above is differentiable. Repeating this process inductively shows that P^{-1} is a smooth tame map.

To show the claim, write $J = \int_0^1 (1-t) D^2 P(f_t \mid f_2 - f_1, f_2 - f_1) dt$. From the proof of lemma Lemma 5.4.16 above we see:

$$\left\| f_2 - f_1 - (DP)^{-1}(f_1 \mid P(f_2) - P(f_1)) \right\|_n = \left\| (DP)^{-1}(f_1 \mid J) \right\|_n \leq C \left\| f_1 \right\|_{n+2r} \left\| J \right\|_0 + C \left\| J \right\|_n .$$

The same proof also shows:

$$\begin{aligned} \|J\|_n &\leq C \|f_2 - f_1\|_{n+r} \|f_2 - f_1\|_r + C(\|f_2\|_{n+2r} + \|f_1\|_{n+2r}) \|f_2 - f_1\|_r^2 \\ &\leq C R_{n+r,r} \|P(f_2) - P(f_1)\|_{n+r}^2 + C(\|f_2\|_{n+2r} + \|f_1\|_{n+2r}) R_{r,r} \|P(f_2) - P(f_1)\|_{n+r}^2 , \end{aligned}$$

with R as in Corollary 5.4.17 above. This shows the estimate:

$$\left\| f_2 - f_1 - (DP)^{-1}(f_1 \mid P(f_2) - P(f_1)) \right\|_n \leq Z(f_1, f_2) \left\| P(f_1) - P(f_2) \right\|_m^2 ,$$

for some $m \geq n + r$ and $Z : \Sigma(B)^2 \to [0, \infty)$ satisfying the same properties as R.

Now given $g \in V$, take t small enough such that $g + tk \in V$. Take $f_2 = P^{-1}(g + tk)$ and $f_1 = P^{-1}(g)$ in the estimate above to obtain:

$$\left\| P^{-1}(g + tk) - P^{-1}(g) - (DP)^{-1}(P^{-1}(g + tk) \mid tk) \right\| \leq Z(P^{-1}(g + tk), P^{-1}(g)) \left\| tk \right\|_m^2 .$$

Dividing this expression by $|t|$ and taking $t \to 0$ shows the claim. □

Part II

Differential relations

Chapter 6

Jet Spaces

In this section we introduce the so called *jet bundles* associated with a fiber bundle $E \twoheadrightarrow M$. This is a space that, intuitively, consists of the Taylor polynomials of all possible sections $M \to E$. We refer to the appendix on Taylor polynomials for much of the terminology used in this section.

Once we have defined this "space of Taylor polynomials", we will use it to define a topology on the space of sections $\Gamma(E)$ that is stronger than the one defined by Definition 2.5.16 (and, deriving from this, Definition 3.3.7) for noncompact base manifolds M. The notions and results introduced in this chapter will be foundational for the upcoming chapters.

6.1 Jet bundles

We proceed as usual: first defining jet space "locally" over Euclidean space and then globally (over manifolds) as spaces that locally reduce to the Euclidean case.

6.1.1 Jets over Euclidean space

Before we treat jet spaces, we introduce the following notation, which proves to be very convenient in this context:

Definition 6.1.1. *Let M be a topological space (in our case this will almost always be a manifold) and $K \subseteq M$ a closed subset. Then we denote $Op(K)$ for an arbitrarily small open $U \supseteq K$.*

We will mostly use this notion for $\mathrm{Op}(p)$ with $p \in M$ a a point. The "smallness" refers to the fact that usually (in the context of differential geometry) shrinking an open allows one to assume more about it. The classical example is that for a "small enough" open $U \supseteq p$ in a manifold, it is diffeomorphic to a subspace of $\mathbb{R}^n$.

Definition 6.1.2. *Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ be open subsets and $f, g : Op(p) \subseteq U \to V$ smooth maps defined around p. We say f and g are r-**tangent** at $p \in U$ if:*

$$D_p^k f = D_p^k g$$

for all $0 \leq k \leq r$.

In particular, 0-tangency simply means $f(p) = D_p^0 f = D_p^0 g = g(p)$. Note that r-tangency at a point p is an equivalence relation, which we denote by $f \sim_{p,r} g$. We write $[f]_{p,r}$ for the r-tangency equivalence class of f at p.

Definition 6.1.3. *The r-order **jet space** $J^r(U, V)$ (with $r \in \mathbb{N}$) for opens $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ is defined as:*

$$J^r(U, V) = \{[f]_{p,r} \mid f : Op(p) \to V \text{ smooth }, p \in U\}.$$

When discussing jet space of order r, we will often simply denote $[f]_p$ for $[f]_{p,r} \in J^r(U, V)$.

This abstract space of equivalence classes can in fact be given a linear structure by the following:

Proposition 6.1.4. *There is a bijection:*

$$\Phi : J^r(U,V) \xrightarrow{\sim} U \times V \times S^r_{n,m},$$

given by:

$$[f]_{p,r} \mapsto (p, f(p), D^1_p f, D^2_p f, \dots D^r_p f)$$

Its inverse is given by:

$$\Phi^{-1} : (p,q,A) \mapsto [(P_{p,q,A})]_{p,r},$$

where $P_{p,q,A}$ is the polynomial associated to (p,q,A) (see Definition A.0.10).

Proof. Observe that by Corollary A.0.9 we have that:

$$\Phi([P_{p,q,A}]_{p,r}) = (p,q,A).$$

For the other direction, note that the polynomial associated to $p, f(p)$ and $(D^k_p f \mid 1 \leq k \leq r)$ is simply the Taylor polynomial of f. By Proposition A.0.12 we see that $P_{p,f(p),D^k_p f}$ is r-tangent to f. □

From now on we will assume $J^r(U,V)$ to be endowed with a vector space structure as above.

Proposition 6.1.5. *Let $U_0, U_1 \subseteq \mathbb{R}^n$ and $V_0, V_1 \subseteq \mathbb{R}^m$ be opens. Let $\alpha : U_0 \to U_1$ be a diffeomorphism and let $\beta : V_0 \to V_1$ be a smooth map. Then the following maps are well-defined and smooth:*

- $\beta_* : J^r(U_0,V_0) \to J^r(U_0,V_1)$ *given by* $[s]_p \mapsto [\beta \circ s]_p$,
- $\alpha^* : J^r(U_0,V_0) \to J^r(U_1,V_0)$ *given by* $[s]_p \mapsto [s \circ \alpha^{-1}]_{\alpha(p)}$,
- $\beta_* \circ \alpha^* = \alpha^* \circ \beta_* : J^r(U_0,V_0) \to J^r(U_1,V_1)$ *given by* $[s]_p \mapsto [\beta \circ s \circ \alpha^{-1}]_{\alpha(p)}$.

Proof. It suffices to prove the latter since the former two are special cases with $\beta = \mathrm{Id}_V$ and $\alpha = \mathrm{Id}_U$ respectively. We have to check smoothness with respect to the linear structure. With respect to the maps $\Phi_i : J^r(U_i,V_i) \to U_i \times V_i \times S^r_{n,m}$ $(i \in \{0,1\})$ as in Proposition 6.1.4 that define this linear structure, the map takes the form:

$$\Phi_1 \circ \beta_* \circ \alpha^* \circ {\Phi_0}^{-1} : (p,q,A_1,\dots,A_r) \mapsto \left(\alpha(p), \beta(q), D^1_{\alpha(p)}(\beta \circ P_{p,q,A} \circ \alpha^{-1}), \dots, D^r_{\alpha(p)}(\beta \circ P_{p,q,A} \circ \alpha^{-1})\right).$$

Now by the smoothness of α and β the right hand side depends smoothly on the input variables, which one can verify by doing an arduous computation.

□

We will use the following result to show that the jet space of a fiber bundle is a manifold:

Corollary 6.1.6. *Let $U_0, U_1 \subseteq \mathbb{R}^n$ and $V_0, V_1 \subseteq \mathbb{R}^m$ be open subsets and let $\alpha : U_0 \to U_1$ and $\beta : V_0 \to V_1$ be diffeomorphisms. Write $\Psi_{\alpha,\beta} = \alpha^* \circ \beta_* = \beta_* \circ \alpha^*$ and let Φ_i as in the proof of Proposition 6.1.5. Then:*

$$\Phi_1 \circ \Psi_{\alpha,\beta} \circ {\Phi_0}^{-1} : U_0 \times V_0 \times S^r_{n,m} \to U_1 \times V_1 \times S^r_{n,m}$$

is a diffeomorphism.

Proof. Proposition 6.1.5 shows smoothness and the given map has a two-sided inverse:

$$\Phi_0 \circ \Psi_{\alpha^{-1},\beta^{-1}} \circ {\Phi_1}^{-1}$$

which is smooth by the same result. □

In addition to these natural maps defined on jet space, there is also the following map that "lifts" a map to jet space:

Definition 6.1.7. *Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a smooth map. We define its r-**jet extension** to be the map $U \to J^r(U,V)$ defined by:*

$$j^r f : p \mapsto [f]_{p,r}.$$

Note that with respect to the linear structure of Proposition 6.1.4, the jet extension of f becomes:

$$j^r f : p \mapsto (p, f(p), D^1_p f, \dots, D^r_p f),$$

therefore it is smooth.

6.1.2 Jets over manifolds

We now turn to the definition of the jet bundle over a fiber bundle $E \twoheadrightarrow M$.

Definition 6.1.8. *Let $E \twoheadrightarrow M$ be a fiber bundle over M. We say that two local sections $s_0, s_1 : \widetilde{U} \to E|_{\widetilde{U}}$ over some open neighborhood $\widetilde{U} \subseteq M$ are r-**tangent** if they are r-tangent with respect to any local trivialization.*

More precisely, given a trivialisation $\tau : E|_{\widetilde{U}} \to \widetilde{U} \times F$ (with F the fiber of E) and charts:

- $\varphi : \widetilde{U} \subseteq M \to U \subseteq \mathbb{R}^n$
- $\psi : \widetilde{V} \subseteq F \to V \subseteq \mathbb{R}^m$,

we say that (local) sections $s_0, s_1 : \widetilde{U} \to E|_{\widetilde{U}}$ are r-tangent at p if $(\psi \circ \pi_F \circ \tau \circ s_0 \circ \varphi^{-1})$ and $(\psi \circ \pi_F \circ \tau \circ s_1 \circ \varphi^{-1})$ are r-tangent at $\varphi(p)$ according to Definition 6.1.2

Just as in the local case, r-tangency is an equivalence relation. Observe that r-tangency for one set of charts and trivialization imply r-tangency for any other. To see this, let s_0 and s_1 be sections such that $(\pi_F \circ \tau \circ s_i)(p) \in V \subseteq F$ and take charts and trivialization φ, ψ, τ as above. Then (after restricting $\widetilde{U}$ if necessary) the local representation of the sections s_i ($i \in \{0, 1\}$) given by:

$$\overline{s}_i = (\psi \circ \pi_F \circ \tau \circ s_i \circ \varphi^{-1}) : U \to V.$$

These are simply smooth functions between opens in Euclidean space. For a different choice of charts and trivialization, we would have local representations $\overline{s_i'} : U' \to V'$ on some different open subsets of Euclidean space. By the properties of charts and local trivializations, these representations are related by $\overline{s'}_i = \beta \circ \overline{s}_i \circ \alpha^{-1}$ for some diffeomorphisms α and β. on sufficiently small opens. Since $[\overline{s'}]_{\overline{p},r} \mapsto [\beta \circ \overline{s} \circ \alpha^{-1}]_{\alpha(\overline{p}),r}$ (with $\overline{p} = \varphi(p)$) is a bijection, $\overline{s_0'}$ and $\overline{s_1'}$ are r-tangent at $\alpha(\overline{p})$ if and only of $\overline{s_0}$ and $\overline{s_1}$ are r-tangent at $\overline{p}$.

Definition 6.1.9. *Let $E \twoheadrightarrow M$ a fiber bundle over M. We define the **order** r **jet space of** E to be:*

$$J^r(E) = \{[s]_{p,r} \mid s : Op(p) \to E \text{ a section }, p \in M\},$$

where $[s]_{p,r}$ is the class of local sections r-tangent to s at p. When the order of the equivalence class is clear, we will also simply denote it by $[s]_p \in J^r(E)$. When the fiber bundle $M \times N \twoheadrightarrow M$ is trivial we will denote the jet space by $J^r(M, N) = J^r(M \times N)$.

Note that we have a canonical surjective map $\pi^r_M : J^r(E) \twoheadrightarrow M$ given by $[s]_p \mapsto p$. Although so far $J^r(E)$ is just a set of equivalence classes, we can give it a manifold and fiber bundle structure.

Proposition 6.1.10. *The jet space $J^r(E)$ has a manifold structure such that $J^r(E) \xrightarrow{\pi^r_M} M$ is a fiber bundle with fiber $\mathbb{R}^m \times S^r_{n,m}$.*

Proof. For an open $\widetilde{U} \subseteq M$, we will write $J^r(E)|_{\widetilde{U}} = \{[s]_p \mid p \in \widetilde{U}\} \subseteq J^r(E)$. Given a set of smooth charts and trivialization:

- $\tau : E|_{\widetilde{U}} \to \widetilde{U} \times F$,
- $\varphi : \widetilde{U} \subseteq M \to U \subseteq \mathbb{R}^n$,
- $\psi : \widetilde{V} \subseteq F \to V \subseteq \mathbb{R}^m$,

as in Definition 6.1.8. As in the discussion above, we will denote $\overline{s} = (\psi \circ \pi_F \circ \tau \circ s \circ \varphi^{-1})$ for the representation of a local section $s : \widetilde{U} \to E|_{\widetilde{U}}$ with respect to these charts.

We define the chart $\Xi = \Xi_{\tau,\varphi,\psi}$ on $W = \{[s]_p \mid p \in \widetilde{U}, (\pi_F \circ \tau \circ s)(p) \in \widetilde{V}\}$:

$$\Xi : [s]_p \mapsto \left(\varphi(p), \overline{s}(p), D^1_{\varphi(p)}\overline{s}, \ldots, D^r_{\varphi(p)}\overline{s}\right) \in U \times V \times S^r_{n,m}$$

Given another trivialization $\tau' : \widetilde{U}' \to \widetilde{U}' \times F$ and other charts $\varphi' : \widetilde{U}' \to U', \psi' : \widetilde{V}' \to V'$, we denote $\overline{s'}$ for the representation of s. As above, $\overline{s}$ and $\overline{s'}$ are related by diffeomorphisms. Then for the chart Ξ' defined similarly to Ξ, the map:

$$\Xi' \circ \Xi^{-1} : U \times V \times S^r_{n,m} \to U' \times V' \times S^r_{n,m}$$

has the same form as in the statement of Corollary 6.1.6 and is therefore a diffeomorphism. This shows that these Ξ induce a manifold structure.

Furthermore, note that Ξ also induces a local trivialization with fiber $S^r_{n,m}$. Locally, the map $\pi^r_M : J^r(X) \to M$ is given by the projection:

$$U \times V \times S^r_{n,m} \twoheadrightarrow U,$$

showing that π^r_M is a submersion. Finally, we can always choose the chart ψ (by restricting $\widetilde{V}$ if necessary) such that $V = \mathbb{R}^m$ showing that we have fibers $\mathbb{R}^m \times S^r_{n,m}$ (indeed, we can also take any fiber diffeomorphic to $\mathbb{R}^m$ such as the open disk $\overset{\circ}{\mathbb{D}^m} \subseteq \mathbb{R}^m$). □

When two sections s_0, s_1 are r-tangent at p, then it is clear that they are also l-tangent at p for any $0 \leq l \leq r$. This gives us a projection:

Proposition 6.1.11. *For $E \twoheadrightarrow M$ a fiber bundle, let $J^l(E)$ and $J^r(E)$ be the jet bundles of orders $0 \leq l \leq r$. Then the map $\pi^r_l : J^r(E) \to J^l(E)$ defined by:*

$$[s]_{p,r} \mapsto [s]_{p,l}$$

is a smooth, surjective submersion.

Proof. The comment above the statement shows that the map is well-defined. Surjectivity is immediate since on the right hand of the expression above we can obtain any $[s]_{p,l}$. For smoothness and submersivity, note that with respect to trivializations as in the proof of Proposition 6.1.10, the projection π^r_l simply becomes the projection of (subsets of) Euclidean spaces:

$$U \times V \times S^r_{n,m} \mapsto U \times V \times S^l_{n,m}.$$

□

Definition 6.1.12. *Given a section $s : M \to E$ of a fiber bundle $E \twoheadrightarrow M$, we define its r**-jet extension** $M \to J^r(E)$ by*

$$(j^r s)(p) = [s]_p.$$

Once again, since locally this is just the jet extension in Euclidean space, the map $j^r s$ is smooth.

We also have a global anologue of Proposition 6.1.5:

Proposition 6.1.13. *Let E and E' be fiber bundles over M, $\phi : E \to E'$ a bundle map. Then there is an induced smooth map $J^r\phi : J^r(E) \to J^r(E')$ defined by:*

$$J^r\phi : j^r s(p) \mapsto j^r \phi_* s(p),$$

where s is a local section of E.

Proof. We have to check the above is well-defined is smooth. For well-definedness, note that if two sections s, s' are r-tangent at p, then the pushforwards $\phi_* s$ and $\phi_* s'$ by ϕ are also r-tangent at p. Therefore the r-jet extensions at p are equal. Note also that (by definition) any point $x \in J^r(E)$ is of the form $x = j^r s(p)$ for some local section s and $p = \pi_M(x)$.

Smoothness follows since locally (i.e. with respect to the charts as in the proof of Proposition 6.1.10) this map is of the form Proposition 6.1.5. □

In fact, we have the following:

Corollary 6.1.14. *Let $FB(M)$ denote the category of fiber bundles over M and bundle maps between them. Then for any $r \in \mathbb{N}$, J^r defines a functor $FB(M) \to FB(M)$.*

Proof. For E a fiber bundle, Proposition 6.1.10 states that $J^r(E)$ is again a fiber bundle over M. Proposition 6.1.13 defines the action of J^r on arrows and it is straightforward to check that it preserves composition, that is — for bundle maps $\phi_0 : E_0 \to E_1$ and $\phi_1 : E_1 \to E_2$ we have:

$$J^r(\phi_1 \circ \phi_0) = J^r\phi_1 \circ J^r\phi_0.$$

□

The "other direction" Proposition 6.1.5 can be generalized to spaces of sections as follows:

Proposition 6.1.15. *Let $E \twoheadrightarrow M$ be a fiber bundle and $\psi : N \to M$ be a diffeomorphism. Then there is a smooth map $\psi^* : J^r(M,E) \to J^r(N,E)$ defined by:*

$$\psi^* : [s]_p \to [s \circ \psi]_{\psi^{-1}(p)}$$

Proof. Well-definedness follows as in the proof of Proposition 6.1.13 and smoothness once again follows from Proposition 6.1.5. □

We have seen that for some fiber bundle $E \twoheadrightarrow M$, the r-jet space $J^r(E) \twoheadrightarrow M$ is again a fiber bundle. In what follows, we will be interested in those sections $s : M \to J^r(E)$ that arise as an r-jet extension from some section $s_0 : M \to E$.

Definition 6.1.16. *We call a section $s : M \to J^r(E)$ **holonomic** if $s = j^r(\pi_0^r \circ s)$, (with $\pi_0^r : J^r(E) \twoheadrightarrow J^0(E) \cong E$ the jet space projection). We denote the subset of holonomic sections with $\mathrm{Hol}(J^r E)$.*

Note that there is a bijection $J^r : \Gamma(E) \xrightarrow{\sim} \mathrm{Hol}(J^r E)$.

6.2 The Whitney Topologies

In order to treat operators on the space of sections, we need to topologize this space. Hamilton's Nash-Moser theorem (Theorem 5.1.1) is stated for tame Fréchet spaces which, as noted before, restricts us to compact base manifolds. In order to state Gromov's Nash-Moser theorem, we will introduce a different topology over arbitrary base manifolds. This is done using jet spaces. In short, we define the opens to be those sets of sections whose jet extensions lie in some open in the fiber bundle. We show that over compact base spaces this topology coincides with the one defined in Definition 3.3.7.

Definition 6.2.1. *Let E be a fiber bundle over M. Then we define the **strong Whitney topologies** in the following way:*

- *For an open $U \subseteq J^k(E)$, we will denote the set of sections s such that $\mathrm{Im}\, j^k s \subseteq U$ with $\Gamma_{j^k}(U)$;*
- *For $k \in \mathbb{N}_0$ we define the **strong (or fine) C^k Whitney topology** on $\Gamma(E,M)$ to be the topology generated by:*
$$\mathcal{B}_k = \left\{ \Gamma_{j^k}(U) \mid U \subseteq J^k(E) \text{ open } \right\};$$
- *We define the **strong (or fine) C^∞ Whitney topology** to be the topology generated by $\mathcal{B}_\infty = \cup_{k\in\mathbb{N}} \mathcal{B}_k$.*

*In a very similar manner, we define the **weak Whitney topologies**:*

- *For an open subset $U \subseteq J^k(E)$ and a compact subset $K \subseteq M$, we will denote the set of sections s such that $\mathrm{Im}(j^k s|_K) \subseteq U$ with $\Gamma_{j^k}(U|_K)$;*
- *For $k \in \mathbb{N}_0$ we define the **weak C^k Whitney topology** on $\Gamma(E,M)$ to be the topology generated by:*
$$\mathcal{B}_k^{wk} = \left\{ \Gamma_{j^k}(U|_K) \mid U \subseteq J^k(E) \text{ open }, K \subseteq M \text{ compact } \right\};$$
- *We define the **weak C^∞ Whitney topology** to be the topology generated by $\mathcal{B}_\infty^{wk} = \cup_{k\in\mathbb{N}} \mathcal{B}_k^{wk}$.*

Remark 6.2.2. Observe that the strong and weak Whitney topologies coincide when M is compact. For noncompact M, convergence with respect to the strong Whitney topology is quite restrictive, see [GG74], pp. 43-44.

When not stated explicitly **we will always assume that $\Gamma(E)$ is endowed with the smooth (C^∞) topology** (whether weak or strong). △

We briefly check that the definition above is sound, i.e.: it indeed defines a topology basis.

Lemma 6.2.3. *The sets $\mathcal{B}_k$ and $\mathcal{B}_k^{wk}$ defined above are topology bases for any $k \in \mathbb{N} \cup \{\infty\}$.*

Proof. We argue only for the strong topology since the argument for the weak topology is very similar.

Since the sets $\Gamma_{j^k}(U)$ satisfy $\Gamma_{j^k}(U) \cap \Gamma_{j^k}(V) = \Gamma_{j^k}(U \cap V)$, the set $\mathcal{B}_k$ is a topology basis. Furthermore, $\mathcal{B}_\infty$ is a basis as well. To see this, take $\Gamma_{j^k}(U), \Gamma_{j^l}(V) \in \mathcal{B}_\infty$ and assume $k \geq l$. Then with the projection π^k_l defined in Proposition 6.1.11 we see that $\Gamma_{j^l}(V) = \Gamma_{j^k}((\pi^k_l)^{-1}(V))$ hence:

$$\Gamma_{j^l}(V) \cap \Gamma_{j^k}(U) = \Gamma_{j^k}((\pi^k_l)^{-1}(V) \cap U) \in \mathcal{B}_k \subseteq \mathcal{B}_\infty.$$

□

From now on, when referring to the spaces of section $\Gamma(E)$ of some fiber bundle E, we assume it to be endowed with the strong C^∞ Whitney topology. For compact base manifolds, Proposition 6.2.10 will yield that this coincides with these previously defined topologies (see Definition 3.3.7 and Definition 2.5.12).

Having just defined these topologies, it is natural to consider some continuous mappings on the space $\Gamma(E)$. This result is essentially a consequence of Proposition 6.1.13.

Proposition 6.2.4. *Let E and E' be fiber bundles over some manifold M and let $\phi : E \to E'$ be a bundle map. Then the pushforward $\phi_* : \Gamma(E) \to \Gamma(E')$ (defined in Definition 3.3.9) is continuous with respect to both the weak and strong Whitney topologies.*

Proof. Let $\Gamma_{j^k}(U) \subseteq \Gamma(E')$ be a basic open neighborhood for some $k \in \mathbb{N}$ and $U \subseteq J^k(E')$. Then we see that:

$$\phi_*{}^{-1}\left(\Gamma_{j^k}(U)\right) = \left\{s \in \Gamma(E) \mid \operatorname{Im}(j^k\phi_* s) \subseteq U\right\} = \Gamma_{j^k}\left((J^k\phi)^{-1}(U)\right) \subseteq \Gamma(E)$$

is again a basic open neighborhood. A similar argument holds for the weak Whitney topology. □

Similarly to Corollary 6.1.14, we have the following functorial property:

Corollary 6.2.5. *Let $FB(M)$ denote the category of fiber bundles over M and bundle maps between them. Then the assignment $E \mapsto \Gamma(X)$ (endowed with either the weak or strong Whitney topology) and:*

$$(\phi : E \to E') \mapsto (\phi_* : \Gamma(E) \to \Gamma(E'))$$

defines a functor $FB(M) \to Top$ (the category of topological spaces).

With some basic modification, the proof of Proposition 6.2.4 and Proposition 6.1.15 give the following:

Proposition 6.2.6. *Let $E \to M$ be a fiber bundle, N a manifold and $\psi : N \to M$ a diffeomorphism. Then the pullback $\psi^* : \Gamma(M, E) \to \Gamma(N, E)$ is continuous.*

The following result gives a more concrete basis of open neighborhoods around a given section that will be useful later on.

Definition 6.2.7. *For $E \twoheadrightarrow M$ a fiber bundle, $s_0 \in \Gamma(E)$ and $\epsilon : M \to \mathbb{R}_{>0}$ continuous. Suppose we have a fiberwise metric on $J^r(E)$ (see Definition 3.3.4) and let d denote the induced distance function. Then we denote:*

$$B_\epsilon(s_0) = \{s \in \Gamma(E) \mid d(j^r s(p), j^r s_0(p)) < \epsilon(p) \text{ for all } p \in M\}.$$

From now on we will always quietly assume that we have a metric on any manifold. The particular choice of metric will not matter for the results discussed.

Proposition 6.2.8. *For any $s_0 \in \Gamma(E)$ the set:*

$$\{B_\epsilon(s_0) \mid \epsilon \in C^\infty(M, \mathbb{R}_{>0})\}$$

is a basis of neighborhoods around s_0 in the strong Whitney topology.

Proof. We must show that for any basic Whitney-open neighborhood $\Gamma_{j^r}(U) \ni s_0$ (with $U \subseteq J^r(E)$ open) there is an ϵ such that $B_\epsilon(s_0) \subseteq \Gamma_{j^r}(U)$. By definition of the Whitney topology, this amounts to finding an open of the form:

$$\{x \in J^r(E) \mid d(j^r s_0(\pi^r_M(x)), x) < \epsilon(\pi^r_M(x))\} \subseteq U,$$

where $\pi^r_M : J^r(E) \twoheadrightarrow M$ is the usual projection. It is not hard to see that it is always possible to find an ϵ satisfying the above, for example by drawing a picture. □

6.2.1 Equivalence with seminorm topology

In this section we prove that the strong Whitney topology over a compact base manifold is equivalent to the topology defined in Definition 3.3.7, wich in turn derives from the seminorms defined in Definition 2.5.12. Our first step is to reduce to the case where $E \twoheadrightarrow M$ is a vector bundle.

Lemma 6.2.9. *Let E be a fiber bundle over a compact manifold M and $s \in \Gamma(E)$ a section. Then there is a vector bundle $E' \twoheadrightarrow M$ and homeomorphic open neighborhoods $W \subseteq \Gamma(E)$ and $W' \subseteq \Gamma(E')$ for the strong Whitney topology, with $s \in W$ and $0 \in W'$ (with 0 denoting the zero-section of $\Gamma(E)$).*

Proof. For the vector bundle $s^* \operatorname{Vert}(E)$, the map defined in Equation (3.3) is an invertible bundle map on a suitable open W. By Corollary 6.2.5 its pushforward is a homeomorphism. □

For the vector bundle case, we can in fact show the following, stronger result:

Proposition 6.2.10. *Let $E \twoheadrightarrow M$ be a vector bundle. Then the weak Whitney topology $\mathcal{T}_W$ and the seminorm topology $\mathcal{T}_S$ (induced by Definition 2.5.16) on the space of sections $\Gamma(E)$ are equal.*

Proof. First we note that for $K \subseteq \mathbb{R}^n$ compact and $E = K \times \mathbb{R}^m \to K$ the trivial bundle, the result holds. To see this, note that in this case $\Gamma(E)$ equals $\mathcal{C}^\infty(K, \mathbb{R}^m)$ with $\mathcal{T}_S$ induced by the seminorms seen in Definition 2.5.2. Let $U \subseteq J^r(E) = K \times \mathbb{R}^m \times S^r_{n,m}$ be any open and s_0 a section such that $\operatorname{Im} s_0 \subseteq U$. Then by the tube lemma, there is a number $\epsilon > 0$ such that:

$$B_\epsilon(s_0) := \{s \in \Gamma(E) \mid \|s(p) - s_0(p)\|_r < \epsilon \text{ for all } p \in K\} \subseteq \Gamma_{j^r}(U).$$

Conversely, remark that $B_\epsilon(s_0)$ is a basic open neighborhood as defined in Definition 6.2.1.

For the global case, let $\widetilde{K} \subseteq M$ be a compact subset. Note that the restriction $\Gamma_W(E) \to \Gamma_W(E|_{\widetilde{K}})$ is continuous with respect to the Whitney topologies. To see this, write $\iota : E|_{\widetilde{K}} \to E$ for the inclusion. Then the restriction is given by the pullback $\iota : s \mapsto s \circ \iota$. Note that ι is a closed map since $E|_{\widetilde{K}} \subseteq E$ is closed. Then on the basis of open sets defined in Definition 6.2.1, we have:

$$(\iota^*)^{-1}\Gamma_{j^r}(U) = \Gamma_{j^r}\left(\iota(U^c)^c\right).$$

Now assume (by taking $\widetilde{K}$ small enough) that there is a trivialization $\tau : E|_{\widetilde{K}} \to \widetilde{K} \times \mathbb{R}^m$ and a chart $\varphi : \widetilde{K} \to K \subseteq \mathbb{R}^n$. Then by Proposition 6.2.4 and Proposition 6.2.6 the map $\Gamma(\widetilde{K}, E|_{\widetilde{K}}) \to \Gamma(K, \mathbb{R}^m)$ given by:

$$s \mapsto \tau \circ s \circ \varphi^{-1}$$

is continuous for the Whitney topologies. It follows that the composition:

$$\begin{aligned}\Phi_K : \Gamma(M, E) &\to \Gamma(K, \mathbb{R}^m),\\ s &\mapsto s|_{\widetilde{K}} \mapsto \tau \circ s|_{\widetilde{K}} \circ \varphi^{-1}\end{aligned}$$

is also continuous.

The same map is continuous with respect to the seminorm topologies $\mathcal{T}_S$; first remark that the restriction to $\widetilde{K}$ obviously satisfies $\left\|s|_{\widetilde{K}}\right\|_r \leq \|s\|_r$ and the pushforward $s \mapsto \tau \circ s$ is continuous since it is a vector bundle operator. Finally, note that the pullback $s \mapsto s \circ \varphi^{-1}$ is continuous as well, as can be seen from definition Definition 2.5.16 and the chain rule.

For the vector bundle $E \twoheadrightarrow M$ it is possible to cover M by countably many compact sets $\widetilde{K}_i$ as above, meaning we have:

- vector bundle trivializations $\tau_i : E|_{\widetilde{K}_i} \to \widetilde{K_i} \times \mathbb{R}^m$;
- charts $\varphi_i : \widetilde{K}_i \to K_i \subseteq \mathbb{R}^n$.

Then we have a continuous map (with respect to either topology)

$$\Phi : \Gamma(E) \to \prod_i \Gamma(K, \mathbb{R}^m), \quad s \mapsto (\tau_i \circ s|_{\widetilde{K}_i} \circ {\varphi_i}^{-1})_i,$$

taking the product topology in the codomain of Φ. This map is clearly injective and one can show that it is an embedding with respect to both $\mathcal{T}_W$ and $\mathcal{T}_S$ (in the latter case we may even argue with convergent sequences). Therefore the topologies coincide. $\square$

Since the weak and strong Whitney topologies coincide over compact base manifolds, we obtain the following from Lemma 6.2.9 and Proposition 6.2.10.

Corollary 6.2.11. *Let $E \twoheadrightarrow M$ be a fiber bundle with M compact. Then the strong Whitney topology defined in Definition 6.2.1 and the topology defined by Definition 2.5.12 and Definition 3.3.7 coincide.*

Remark 6.2.12. The proof of Proposition 6.2.10 indicates a different way to define the Whitney topologies on a space of sections without the use of jet bundles. Indeed, we may define it to be induced by the injection Φ above by requiring that Φ is an embedding. For this approach we refer to [Hir12], p. 35.

From now on, we assume that for M compact and $E \twoheadrightarrow M$ a fiber bundle, the space $\Gamma(E)$ with the Whitney topology is also endowed with the smooth structure defined by Definition 3.3.7. $\triangle$

Chapter 7

Differential relations and the h-principle

We introduce one of the main tools in order to tackle the isotropic embedding problem: differential relations. These can be thought of as generalizations of differential equations. Their advantage, however, is that there is an obvious notion of "formal solutions", which will play an important role in the next section.

7.1 Differential relations

Definition 7.1.1. *Let $E \to M$ be a fiber bundle. A* ***differential relation of order r*** *is a subset $\mathcal{R} \subseteq J^r(E)$. We denote $\Gamma(\mathcal{R}) = \operatorname{Sec}(\mathcal{R})$ for the set of sections $s : M \to J^r(E)$ such that $\operatorname{Im} s \subseteq \mathcal{R}$.*

A differential relation is said to be ***open*** *or* ***closed*** *if $\mathcal{R} \subseteq J^r(E)$ is open or closed respectively.*

Following [EM02], we denote $\operatorname{Sec}\mathcal{R}$ for the set of sections with image in $\mathcal{R}$ and $\operatorname{Hol}\mathcal{R} \subseteq \operatorname{Sec}\mathcal{R}$ for the subset of holonomic sections. The set $\operatorname{Sec}\mathcal{R}$ is also called the set of ***formal solutions*** *whereas $\operatorname{Hol}\mathcal{R}$ is called the set of solutions.*

Note that for an open differential relation $\mathcal{R}$ the set $\operatorname{Hol}\mathcal{R} \subseteq \Gamma(E)$ of holonomic sections is open in the Whitney topology. The following example shows how this generalizes differential equations:

Example 7.1.2. Let $q = |\{\alpha \mid |\alpha| \leq r\}|$ be the number of n-indices of degree less than k. Let Φ be some smooth function $\mathbb{R}^{q+1} \to \mathbb{R}^k$. Then we know that:

$$\Phi(p, \partial_\alpha f(p) \mid 0 \leq |\alpha| \leq r) = 0, \quad \forall p \in \mathbb{R}^n$$

is called *differential equation* on the space of smooth functions. A function $f \in C^\infty(\mathbb{R}^n, \mathbb{R})$ is of course the same as a section of the trivial bundle $\mathbb{R}^n \times \mathbb{R} \twoheadrightarrow \mathbb{R}^n$, and its partial derivatives are coordinates of the r-jet extension $j^r f : \mathbb{R}^n \to J^r(\mathbb{R}^n, \mathbb{R}) = \mathbb{R} \times \mathbb{R}^{q+1}$. Now observe that f satisfies the equation above if and only if the image of $j^r f$ takes values in the zero set $\{\Phi = 0\} \subseteq J^r(\mathbb{R}^n, \mathbb{R})$. The differential equation therefore defines a closed differential relation and the *holonomic* sections of the differential relation satisfy the differential equation.

Similarly, we can define relations by taking (strict) inequalities. Taking $k = 1$ in the above example, the differential inequality $\Phi(p, \partial_\alpha f(p) \mid 0 \leq |\alpha| \leq r) > 0$ defines the open relation $\{\Phi > 0\} \subseteq J^r(\mathbb{R}^n, \mathbb{R})$. △

Not every differential relation admits holonomic sections:

Example 7.1.3. On $J^1(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}^3$, define the relation $\mathcal{R}_\varepsilon = \{(x, y, z) \mid |z - 1| < \varepsilon, |y| < \varepsilon\}$. Then one can check that no differentiable function $\mathbb{R} \to \mathbb{R}$ satisifes this relation for small ε. △

Quite often, the differential relation is not specified by some equation on the set $J^r(E)$ but rather by some requirement on the (holonomic) solutions. This means any $[\sigma]_p \in J^r(E)$ belongs to the differential relation if and only if there is a holonomic representative $\sigma : \operatorname{Op}(p) \to E$.

Example 7.1.4.

- The relation of functions $\mathbb{R} \to \mathbb{R}$ such that $f'(x) \neq 0$ for all $x \in \mathbb{R}$ corresponds the relation $\{(x, y, z) \mid z \neq 0\} \subseteq \mathbb{R}^3 \cong J^1(\mathbb{R}, \mathbb{R})$.

- The relation $\mathbb{R}_{\text{imm}} \subseteq J^1(N,M)$ is defined as the differential relation of immersions $N \to M$. This means that it contains exactly those jets $[s]_p \in J^1(N,M)$ representing an injective map $T_pN \to T_{\sigma(p)}M$.
- The relation $\mathbb{R}_{\text{subm}} \subseteq J^1(N,M)$ is defined in a similar way.
- The relation $\mathbb{R}_{\text{iso}} \subseteq J^1(N,M)$ of isotropic immersions into an (almost) symplectic manifold (M,ω) contains exactly those $x \in J^1(N,M)$ such that the linear map $T_pN \to T_qM$ (where $p = \pi_N(x)$ and is the target of x) corresponding to x has isotropic image with respect to ω_q.

△

7.2 Diffeomorphism-invariant relations

We take a look at differential relations that are invariant under transformations of the base manifold M. We will see in Theorem 8.1.2 that we have an h-principle for such relations on open manifolds.

To motivate the upcoming definitions we give the following example:

Example 7.2.1. Consider the relation $\mathcal{R}_{\text{imm}} \subseteq J^1(N,M)$ of immersions defined in example Example 7.1.4. Note that for any immersion $f : N \to M$ and diffeomorphism $\phi : N \to N$, the map $f \circ \phi^{-1}$ is still an immersion. The same holds for the differential relation $\mathcal{R}_{\text{subm}}$ of submersions.

Similarly, let $f : N \to (M,\omega)$ be an isotropic immersion into a (almost) symplectic manifold, then so is $f \circ \phi^{-1}$. △

In these examples, the transformation $\phi : N \to N$ has a natural lift to the bundle $N \times M \twoheadrightarrow N$ given by $\phi \times \mathrm{Id}_M : (p,q) \mapsto (\phi(p), q)$. Another example where such a natural lift is available is the tangent bundle $TN \twoheadrightarrow N$, where it is given by the tangent map $T\phi$.

Definition 7.2.2. *Let $E \overset{\pi}{\twoheadrightarrow} M$ be a fiber bundle. We denote $\mathrm{Diff}_M E$ for the group of fiber preserving diffeomorphisms of E — that is, diffeomorphisms $\phi_E : E \to E$ for which there exists a diffeomorphism $\phi_M : M \to M$ such that:*

$$\begin{array}{ccc} E & \xrightarrow{\phi_E} & E \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ M & \xrightarrow{\phi_M} & M \end{array}$$

commutes.

We call a fiber bundle E together with some homomorphism $\mathrm{Diff}_M \to \mathrm{Diff}_M E$ of lifts as above a ***natural bundle****. This homomorphism will be denoted by $\phi \mapsto \phi_E$.*

We will mainly be concerned with the case where $E = N \times M \twoheadrightarrow N$ is trivial and the lift is given by $\phi \mapsto \phi_E = \phi \times \mathrm{Id}_M$.

Lemma 7.2.3. *Let $E \twoheadrightarrow M$ be a natural bundle, Then J^rE is also a natural bundle*

Proof. Given $\phi : M \to M$ a diffeomorphism, write $\phi_E : E \to E$ for the natural lift of the natural bundle E. Then define $\phi_E^r \in \mathrm{Diff}_M J^rE$ by:

$$\phi_E^r : [s]_p \mapsto [\phi_E \circ s \circ \phi^{-1}]_{\phi(p)}. \tag{7.1}$$

It is straightforward to check that the assignment $\phi \mapsto \phi_E^r$ is a group homomorphism. □

The main definition of this section is the following:

Definition 7.2.4. *Let $E \twoheadrightarrow M$ be a natural bundle and let $\mathcal{R} \subseteq J^rE$ be a differential relation. Then we say that $\mathcal{R}$ is $\mathrm{Diff}\,M$* ***invariant*** *(or simply Diff* ***invariant****) if for any diffeomorphism $\phi : M \to M$ the lift $\phi_E^r : J^rE \to J^rE$ defined in Equation (7.1) leaves $\mathcal{R}$ invariant, that is:*

$$\phi_E^r(\mathcal{R}) \subseteq \mathcal{R}.$$

Equivalently, $\mathcal{R} \subseteq J^rE$ is a natural subbundle.

Example 7.2.5. The differential relations in Example 7.2.1 are all defined for trivial bundles $N \times M \twoheadrightarrow M$, for which the natural lift is given by $\phi \mapsto \phi \times \mathrm{Id}_M$. Then we see that the lift to jet spaces of some diffeomorphism $\phi : N \to N$ defined in Equation (7.1) is simply the pullback of the inverse $[s]_p \mapsto [s \circ \phi_p^{-1}]_{\phi(p)}$. Hence the statement of Example 7.2.1 amounts to stating that these relations are Diff_M invariant. △

Remark 7.2.6. Note that by definition of ϕ_E^r (Equation (7.1)), the action of ϕ_E^r preserves holonomic sections. △

7.3 The h-principle

As seen before, some differential relations do not have any holonomic sections. It turns out however, that for many important and interesting relations the existence of formal solutions implies existence of actual solutions. In fact for many such relations there are homotopies from formal solutions to holonomic solutions. This an example of an h-principle (from *homotopy principle*), a homotopic theoretic statement relating the spaces $\mathrm{Sec}\,\mathcal{R}$ and $\mathrm{Hom}\,\mathcal{R}$

Definition 7.3.1. *A differential relation is said to satisfy the* ***h-principle*** *if any section in* $\mathrm{Sec}\,\mathcal{R}$ *is homotopic to a section in* $\mathrm{Hol}\,\mathcal{R}$.

There are many variations of the h-principle. We mention in particular the following, which is a rather strong variation:

Definition 7.3.2. *A differential relation is said to satisfy the* ***full h-principle*** *if the inclusion* $\mathrm{Hol}\,\mathcal{R} \hookrightarrow \mathrm{Sec}\,\mathcal{R}$ *is a weak homotopy equivalence.*

Remark 7.3.3. The fact that $\mathrm{Hol}\,\mathcal{R} \hookrightarrow \mathrm{Sec}\,\mathcal{R}$ is a weak homotopy equivalence can be characterized in the following way. Suppose we have any continuous family s_x of sections in $\mathrm{Sec}\,\mathcal{R}$ for x in the closed k-dimensional disk $\mathbb{D}^k$, (meaning that the map $x \mapsto s_x$ defines a continuous map $\mathbb{D}^k \to \mathrm{Sec}\,\mathcal{R} \subseteq \Gamma(E)$ with respect to the weak Whitney topology) such that $s_x \in \mathrm{Hol}\,\mathcal{R}$ for $x \in \partial\mathbb{D}^k$. Then there is a homotopy s_x^τ for $\tau \in [0,1]$ such that:

- $s_x^0 = s_x$ for all $x \in \mathbb{D}^k$,
- $s_x^1 \in \mathrm{Hol}\,\mathcal{R}$ for all $x \in \mathbb{D}^k$ and
- $s_x^\tau = s_x$ for all $x \in \partial\mathbb{D}^k$ and $\tau \in [0,1]$.

This can be shown using the long exact sequence of the higher homotopy groups, see [Sag], pp. 86-89. △

Chapter 8

Holonomic Approximation Theorems

In this section we take a look at a powerful tool for showing the existence of h principles: holonomic approximation. Aftewards we examine the more general holonomic $\mathcal{R}$-approximation theorem. These results are due to Eliashberg and Mishachev and feature in their book ([CEM24], Ch.3 and [CEM24], Ch.21).

First we state the "standard" holonomic approximation, which as the name implies deals with the approximation of formal sections $M \to J^r E$ by holonomic ones. This version has applications to open relations $\mathcal{R}$, since any section close enough to a section of $\mathcal{R}$ also lies in $\mathcal{R}$.

For closed relations $\mathcal{R} \subseteq J^r E$, we will need a different tool, which is provided by "$\mathcal{R}$-approximation". This theorem essentially entails that, under certain conditions, we can apply the procedure of holonomic approximation to a formal section $F : M \to \mathbb{R}$ while retaining a section into $\mathcal{R}$ along the way. The conditions required, *local integrability* and *microflexibility*, will be introduced in Section 8.2 before the statement of the theorem.

Remark 8.0.1. Throughout this chapter we quietly assume that all manifolds and fiber bundles are endowed with metrics. On fiber bundles, we assume these to be fiberwise. We denote distance function derived from any of the metrics with d. △

8.1 Holonomic Approximation

The problem holonomic approximation aims to solve is the following. As seen before, some differential relations $\mathcal{R}$ admit no holonomic solutions. In fact, it may happen that any holonomic section is "far away" from any formal solution (or indeed the relation $\mathcal{R}$ itself). This is even the case for formal solutions near some (compact) subset, as seen in the following example (that extends Example 7.1.3):

Example 8.1.1. On $J^1(\mathbb{R}^n, \mathbb{R})$, consider the differential relation of maps $f : \mathbb{R}^n \to \mathbb{R}$ such that $|\partial_n f - 1| < \varepsilon$ and $|f| < \varepsilon$. For $k < n$, let $F : \mathrm{Op}([0,1]^k) \subseteq \mathbb{R}^n \to J^2(\mathbb{R}^n \mathbb{R})$ be a formal solution. Then for ε small one cannot approximate F by an actual solution — that is: there exists a $\delta_0 > 0$ such that there is no actual solution satisfying $|F - j^1 f| < \delta_0$.

It follows that for any holonomic section $j^1 f : \mathrm{Op}([0,1]^k) \to J^1(\mathbb{R}^n, \mathbb{R})$ and formal solution F on the same domain, there is a point p in the domain where $|F - j^1 f(p)| \geq \delta_0$. △

Surprisingly, it is possible to approximate any formal solutions on a neighborhood of a *deformation* of a subset such as the above. For the above example, this means that for any $\delta > 0$ we can find a diffeotopy $h^\tau : \mathbb{R}^n \to \mathbb{R}^n$ and a holonomic section $\widetilde{F} : \mathrm{Op}(h^1([0,1]^k)) \to J^2(\mathbb{R}^n, \mathbb{R})$ such that:

$$|\widetilde{F} - F| < \delta.$$

The statement of the theorem for manifolds is as follows:

Theorem 8.1.2 (Holonomic Approximation)**.**
Let $E \to M$ be a fiber bundle, $P \subseteq M$ a submanifold of positive codimension and $F : Op(P) \to J^r(E)$ a section. We assume that we are given metrics on all manifolds. Then for any $\epsilon, \delta > 0$, there is:

- *a δ-small diffeotopy $h^\tau : M \to M$ ($\tau \in [0,1]$) and*
- *a holonomic section $\widetilde{F} : Op(h^1(P)) \to J^r(E)$ with $\operatorname{Dom} \widetilde{F} \subseteq \operatorname{Dom} F$*

such that:

- *$d(\widetilde{F}, F) < \epsilon$ on $\operatorname{Dom} \widetilde{F}$*

Remark 8.1.3. For the proof Theorem 8.1.2 we refer once again to [CEM24], Ch.3, where the authors show that the theorem holds more generally for polyhedra $P \subseteq M$. △

If M is noncompact we take ϵ and δ to be positive functions on M.

Suppose we are interested in a differential relation $\mathcal{R}$ that is open. Then for any $s_0 \in \operatorname{Sec} \mathcal{R}$ we can find an ϵ such that (with respect to some fiberwise metric) we have:

$$B_\epsilon(s_0) \subseteq \operatorname{Sec}(\mathcal{R}) \subseteq \Gamma(J^r(E))$$

(compare Proposition 6.2.8). Theorem 8.1.2 now tells us that any holonomic section close enough to s_0 in jet space will satisfy the differential relation near some deformed submanifold. In the following section we show that under certain constraints this allows us to show the h-principle for open relations.

One important variation of Theorem 8.1.2 is the following. Note the similarity with Remark 7.3.3, which hints that this result may be used to prove the full h-principle.

Theorem 8.1.4 (Parametric Holonomic Approximation)**.**
Let $E \to M$ be a fiber bundle, $P \subseteq M$ a submanifold of positive codimension and $F_x : Op(P) \to J^r(E)$ a continuous family of sections $x \in I^k$ such that F_x is holonomic for $x \in Op(\partial I^k)$.

Then for any $\epsilon, \delta > 0$ we have:

- *δ-small diffeotopies $h_x^\tau : M \to M$ (for $\tau \in [0,1]$ and $x \in I^k$) with $h_x^\tau = Id_M$ for $x \in Op(\partial I^k)$,*
- *holonomic sections $\widetilde{F}_x : Op\left(h^1(P)\right) \to J^r(E)$ with $\operatorname{Dom} \widetilde{F} \subseteq \operatorname{Dom} F$ and $\widetilde{F}_x = F_x$ for $x \in Op(\partial I^k)$*

such that:

- *$|\widetilde{F}_x - F_x| < \epsilon$ on $\operatorname{Dom} \widetilde{F}_x$*

The following theorem was first proved by Gromov in [Gro69] although the proof below, using Holonomic Approximation, is from [EM02]. For its statement we give the following definition and result (without proof).

Definition 8.1.5. *A manifold M is said to be* ***open*** *if none of its path components are closed (that is: compact without boundary).*

Theorem 8.1.6. *For any open manifold M, there is a polyhedron P of positive codimension such that for any neighborhood U of P there is an isotopy $g_t : M \to M$ ($t \in [0,1]$) with $g_0 = Id_M$ and $g_1(M) \subseteq U$.*

A sketch of the proof is found in [EM02], p. 36.

Theorem 8.1.7. *Let $E \twoheadrightarrow M$ be a natural fiber bundle over an open manifold and $\mathcal{R} \subseteq J^rE$ an open* $\operatorname{Diff} M$*-invariant differential relation. Then the full h-principle holds for $\mathcal{R}$.*

Sketch of proof. By Remark 7.3.3, we have to prove that for any family of sections $F_x : M \to E$ with $x \in I^k$, such that F_x is holonomic for x near the boundeary of I^k, we have to find a homotopy F_x^t relative to ∂I^k.

Take a polyhedron $P \subseteq M$ of positive codimension provided by the theorem above. Let ϵ be a positive function on M such that $B_\epsilon(s_x) \subseteq \operatorname{Sec} \mathcal{R}$ for all $x \in I^k$. For these values, take diffeotopies h_x^τ and holonomic sections $\widetilde{F}_x$ as in Theorem 8.1.4 and write $h_x := h_x^1$. Note that the $\widetilde{F}_x$ are holonomic sections of $\mathcal{R}$.

By compactness, there exists an open neighborhood U of P such that $h_x(P) \subseteq h_x(U) \subseteq \operatorname{Dom} \widetilde{F}_x$ for all $x \in I^k$. Let $g^t : M \to M$ be the isotopy into U as in Theorem 8.1.6. Then for any $x \in I^k$ the isotopy $g_x^t : h_x \circ g^t \circ {h_x}^{-1}$ satisfies $g_x^0 = \mathrm{Id}_M$ and $g_x^1(M) \subseteq \operatorname{Dom} \widetilde{F}_x$. By $\operatorname{Diff} M$ invariance of $\mathcal{R}$,

$$F_x^t = (g_x^1)_*^{-1}(\widetilde{F}_x)$$

gives the required homotopy. □

8.2 $\mathcal{R}$-Approximation

For the problem of isotropic embeddings the relation under consideration is not open and therefore we cannot apply Theorem 8.1.7. Instead we will use another theorem in [EM02] which relies on the so-called *microflexibility* and *local integrability* (defined below) of the differential relation. The latter property serves as a base case: we can find a continuous family of holonomic sections around points in some region. Then microflexibility allows us to create holonomic sections around (deformed) line segments. Continuing inductively, we get families of holonomic sections on small planes and cubes, until we end up with a holonomic section of the entire region under consideration. In Section 9.4 we show how to prove microflexibility and local integrability for certain differential relations.

8.2.1 Local integrability

As described above, local integrability is the property that we can extend a continuous choice of points in some differential relation $\mathcal{R}$ into a continuous family of local holonomic sections.

Definition 8.2.1. *Let $E \twoheadrightarrow M$ a fiber bundle and $\mathcal{R} \subseteq J^r E$ a differential relation.*

*We say that $\mathcal{R}$ is **locally integrable** if for any $j \in \mathcal{R}$ and $p = \pi_M(j)$ there is a holonomic section $F : Op(p) \to \mathcal{R}$ such that $F(p) = j$.*

*The relation $\mathcal{R}$ is said to be **parametrically locally integrable** if for any smooth $h : [0,1]^k \to M$ and smooth $g : [0,1]^k \to \mathcal{R}$ lifting h, there is a smooth family of holonomic sections $F_x : Op(h(x)) \to \mathcal{R}$ for $x \in [0,1]^k$ such that:*

$$F_x(h(x)) = g(x).$$

The smoothness requirement means that for $\operatorname{Graph} h = \{(x, h(x))\} \subseteq [0,1] \times M$, the map $F : Op(\operatorname{Graph} h) \to \mathbb{R}$ given by $(x,p) \mapsto F_x(p)$ is smooth.

*Finally, we say $\mathcal{R}$ is **relatively parametrically locally integrable** if given a smooth map $h : [0,1]^k \to M$ and lift $g : [0,1]^k \to \mathcal{R}$ as before and a smooth family of holonomic extensions $F_x : Op(\pi_M(h(x))$ for $x \in Op(\partial I^k)$, we can find a smooth family F_x for $x \in [0,1]^k$ as before agreeing with the given F_x on $Op(\partial I^k)$.*

Example 8.2.2. Open differential relations satisfy all forms of local integrability. To see this, take some map $h : [0,1]^k \to M$ and lift $g : [0,1]^k \to \mathcal{R}$ as in the definition above. For any $x_0 \in [0,1]^k$, take open subsets $U_0 \ni h(x_0)$ in M and $V_0 \ni q_0 := \pi_1^r g(x_0)$ in E over which $J^r E$ is trivializable. Then for points (x,q) near $(h(x_0), q_0)$ we can simply take the polynomial $P(x,q)$ associated to $A = g(q)$ (as in Definition A.0.10), which depends smoothly on (x,q). Note that by openness of the differential relation $\mathcal{R}$ the polynomial satisfies $\operatorname{Im} P(x,q) \subseteq \mathcal{R}$ when restricted to a sufficiently small open.

To obtain the smooth family $F_x : \mathrm{Op}(h(x)) \to \mathcal{R}$ on the entire image of h, one uses a partition of unity argument. More precisely, on overlaps of trivializable neighborhoods, pass to the trivial jet space of the form $U \times V \times \mathbb{R}^l$ and use the partition of unity to interpolate between the various definitions.

Similarly, if we are already given a holonomic family near $\mathrm{Op}(\partial I^k)$ we can use bump functions on trivializible neighborhoods to interpolate between the given and constructed holonomic sections, resulting in a family. △

Remark 8.2.3. From now on we will refer to "relatively parametrically locally integrable" simply by "locally integrable". △

8.2.2 Microflexibility

Flexibility intuitively means that we can extend some homotopy of sections over a small set to a homotopy over a large set. *Microflexibility* means we can extend only the initial stage of the homotopy.

Definition 8.2.4. *Let $E \twoheadrightarrow M$ be a fiber bundle and $\mathcal{R} \subseteq J^r E$ a differential relation. We say that $\mathcal{R}$ is **flexible** if for any compact subsets $K \subseteq K' \subseteq M$, holonomic section $G : Op(K') \to \mathcal{R}$ and homotopy of holonomic sections $F_t : Op(K) \to \mathcal{R}$ ($t \in [0,1]$) such that $F_0 = G|_{Op(k)}$, there is a homotopy of holonomic sections $G_t : Op(K') \to \mathcal{R}$ such that:*

- *$G_0 = G$ and*
- *$(G_t)|_{Op(K)} = F_t$ for all $t \in [0,1]$.*

*We say that $\mathcal{R}$ is **microflexible** if such a homotopy G_t can be found for $t \in [0, \varepsilon]$ with $0 < \varepsilon \leq 1$.*

Example 8.2.5. Open relations $\mathcal{R} \subseteq J^r E$ are microflexible. To see this, let $B \subseteq A \subseteq M$ with B compact. Given a holonomic section $G_0 = j^r g_0 : \mathrm{Op}(A) \to \mathcal{R}$ and a homotopy of holonomic sections $F_t = j^r f_t : \mathrm{Op}(B) \to \mathcal{R}$ such that $F_0 = G_0|_{\mathrm{Op}(B)}$. Then we can extend the homotopy F_t to $\mathrm{Op}(A)$ for small t in the following way. Choose an open $U \supseteq B$ such that F_t is defined on U for all t. Then pick a bump function φ such that $\varphi = 1$ on B and $\varphi = 0$ on $M \setminus U$. Then define:

$$g_t(p) = f_{\varphi(p)t}(p)$$

and set $G_t(p) = (j^r g_t)(p)$. By the properties of φ we see that we have:

$$G_t(p) = F_t(p) \text{ for } p \in B \text{ and}$$
$$G_t(p) = G_0(p) \text{ for } p \in M \setminus U,$$

for all $t \in [0, 1]$. However, the bump function in the definition of g_t causes its r-jet extension to differ from $j^r f_t$ by the chain rule, so G_t does not satisfy the holonomic relation for all $t \in [0, 1]$. However, since the bump function is fixed and the relation is open, we will have $\operatorname{Im} G_t \subseteq \mathcal{R}$ for $t \in [0, \varepsilon]$ with ε small. △

Definition 8.2.4 is quite general and is essentially the same as the definition of a *microflexible sheaf* in [Gro86] (see p.40 and p.76). For the purpose of stating Holonomic $\mathcal{R}$-approximation we will use a more specialized variation of microflexbility, which is only required for pairs of subspaces of the following kind:

Definition 8.2.6. *Write $K^n = [-1, 1]^n$ for $n \in \mathbb{N}$. For $k < m$ we say that a pair of subsets (A, B) with $B \subseteq A$ of an m-dimensional manifold M is a θ_k**-pair** if there is a diffeomorphism $\phi : A \to [-1, 1]^m$ such that $\phi(B) = [-1, 1]^k \subseteq [-1, 1]^m$.*

Then the notion of microflexibility we need is the following. Note that a major difference is the requirement that the homotopy be constant near some piece of the boundary, and that the extension also be constant on the boundary of the larger set:

Definition 8.2.7. *We say that a differential relation $\mathcal{R} \subseteq J^r E$ is θ_k**-microflexible** if for any small enough open $U \subseteq M$, and:*

- *θ_k pair (A, B) inside U,*
- *holonomic section $G : Op(A) \to \mathcal{R}$ and*
- *homotopy of holonomic sections $F^t : Op(B) \to \mathcal{R}$ with $F^0 = G|_{Op(B)}$ that is constant on $Op(\partial B)$ (meaning $F^t(p) = F^0(p)$ for $p \in Op(B)$)*

there is an extension $G^t : Op(A) \to J^r E$ of holonomic solutions up to some $\varepsilon \leq 1$ that is constant on $Op(\partial A)$.

*A differential relation $\mathcal{R} \subseteq J^r E$ is said to be **parametrically** θ_k**-microflexibile** if for smooth families of θ_k-pairs (A_x, B_x), holonomic sections G_x and homotopies F^t_x with $x \in [0, 1]^m$ as above there is some ε and a smooth family G^t_x of homotopies up to ε as above.*

The way we showed microflexibility for open relations in Example 8.2.5 also shows that an open relation is θ_k-microflexible for any $k < \dim M$. The same argument can also be modified to show parametric θ_k-microflexibility.

Remark 8.2.8. Since from now on we will usually require parametric θ_k-microflexibility, we will simply refer to this as θ_k-microflexibility. △

8.2.3 Statement of $\mathcal{R}$-approximation

Now that we have the requisite notions of microflexibility and local integrability we are ready to state the main theorem of this section.

Theorem 8.2.9 (Holonomic $\mathcal{R}$-Approximation)**.**
Let $E \to M$ be a fiber bundle, $\mathcal{R} \subseteq J^r E$ a locally integrable and microflexible differential relation, $P \subseteq M$ a polyhedron of positive codimension and $F : Op(P) \to J^r(E)$ a section. Then for any $\epsilon, \delta > 0$, there is:

- *a δ-small diffeotopy $h^\tau : M \to M$ $(\tau \in [0, 1])$ and*
- *a holonomic section $\widetilde{F} : Op\left(h^1(P)\right) \to \mathcal{R}$ with $\operatorname{Dom} \widetilde{F} \subseteq \operatorname{Dom} F$*

Remark 8.2.10. The parametric variant of $\mathcal{R}$ approximation also holds. △

From the parametric $\mathcal{R}$ approximation theorem, we get an anologue of Theorem 8.1.7 in much the same way:

Theorem 8.2.11. *Let $E \twoheadrightarrow M$ be a natural fiber bundle over an open manifold and $\mathcal{R} \subseteq J^r E$ a locally integrable, microflexible and* Diff *M-invariant differential relation. Then the full h-principle holds for $\mathcal{R}$.*

Chapter 9

Differential operators

Differential operators are a well-known concept from calculus. In this section, we define them in a very general way for manifolds using jet spaces. We show that they are smooth tame maps (when the base manifold as compact).

Our main goal is to apply a Nash-Moser theorem to differential operators in order to show that the differential relation obtained from a differential equation (seen in Example 7.1.2 but defined more generally in Section 9.3) is locally integrable and microflexible. This will allow us to apply holonomic $\mathcal{R}$-approximation to this relation.

In order to do this, we will need define the differential, often called *linearization* in the context of differential operators. We show that the linearization of a differential operator is once again a differential operator.

After having done this, we state Gromov's Nash-Moser theorem without proof. This theorem turns out to to be more suitable to the application at hand than Theorem 5.1.1. We then rigorously define the relation that we will be looking to apply $\mathcal{R}$-approximation to, and using Gromov's theorem we show that this relation is indeed locally integrable and microflexible. As an immediate consequence, we get an a h-principle statement similar to Theorem 8.2.11.

9.1 Definition and linearization of differential operators

The following general definition of differential operators defines them, in short, as operators factoring through the jet extension j^r and a bundle map.

Definition 9.1.1. *Let E, E' be fiber bundles over M. A map $D : \Gamma(E) \to \Gamma(E')$ is called a **differential operator** if it factors as:*

$$\begin{array}{ccc} \Gamma(E) & \xrightarrow{s \mapsto J^r s} & \Gamma(J^r(E)) \\ & \searrow^{D} & \downarrow{\scriptstyle \sigma(D)_*} \\ & & \Gamma(E') \end{array}$$

*where $\sigma(D)_*$ is the bundle operator induced by some bundle map $\sigma(D) : J^r(E) \to E'$ called the **symbol** of the operator D.*

Now we turn to show the claim from the introduction of this chapter that the linearization of a differential operator is again a differential operator. We wish to compute $T_{s_0}D$ for some section $s_0 \in \Gamma(E)$. Therefore we take a tangent vector $\widetilde{s_0} \in \Gamma(\mathrm{Vert}(E))$ such that $s_0 = \pi \circ \widetilde{s_0}$, and compute the tangent map of a differential operator D by taking a family of curves s_t such that $\frac{d}{dt}\big|_{t=0} s_t = \widetilde{s_0}$ (through the identification of Proposition 3.3.12). Then at a point $p \in M$ it is given by:

$$(T_{s_0} D\widetilde{s_0})(p) = \frac{d}{dt}\bigg|_{t=0} D(s_t)(p) = \frac{d}{dt}\bigg|_{t=0} \sigma(D)_* J^r s_t(p) = T_{J^r s_0(p)} \sigma(D) \; \frac{d}{dt}\bigg|_{t=0} J^r s_t(p) \tag{9.1}$$

In this expression, $\frac{d}{dt}\big|_{t=0} J^r s_t(p)$ is some tangent vector in $\mathrm{Vert}_{J^r s_0(p)} J^r(E)$. However, for TD to be a differential operator, it must factor through a jet space. We need the following:

Lemma 9.1.2. *There is a canonical isomorphism* $\mathrm{Vert}(J^r(E)) \to J^r(\mathrm{Vert}(E))$ *determined by:*

$$(J^r s_0(p), [t \mapsto J^r s_t(p)]) \mapsto J^r \widetilde{s_0}(p),$$

where s_t *is a smooth family of sections,* $[t \mapsto J^r s_t(p)]$ *is a tangent vector equivalence class and:*

$$\widetilde{s_0} = \left(\frac{d}{dt}\bigg|_{t=0} s_t \right) \in \Gamma(\mathrm{Vert}(E)).$$

Proof. First we show that the expression above in fact determines a map on $\mathrm{Vert}(J^r(E))$. A tangent vector at some point $J^r s(p) \in J^r(E)$ is the derivative of some curve $\gamma : I \to J^r(E)$ such that $\gamma(0) = J^r s(p)$. We claim that any *vertical* tangent vector in jet space is a derivative of a curve of the form:

$$t \mapsto J^r s_t,$$

that is, there is an "underlying" curve of actual (local) sections.

To show this we pass to the "local" jet space $U \times V \times S^r_{n,m}$ by charts as in the proof of Proposition 6.1.10, where $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are opens. This is the jet space of the "bundle" $U \times V \twoheadrightarrow U$, and $\mathrm{Vert}(U \times V) = U \times TV = U \times V \times \mathbb{R}^m$. Therefore we see that:

$$\begin{aligned} J^r \,\mathrm{Vert}(U \times V) &\cong U \times TV \times S^r_{n,2m} \cong U \times V \times \mathbb{R}^m \times S^r_{n,2m} \\ \mathrm{Vert}\, J^r(U \times V) &= U \times T(V \times S^r_{n,m}) \cong U \times V \times \mathbb{R}^m \times S^r_{n,2m} \end{aligned}$$

In the bundle $J^r(U \times V) = U \times V \times S^r_{n,m}$ a vertical curve over a point $p \in U$ is simply a smooth function $(q, A) : I \to V \times S^r_{n,m}$, with $0 \in I \subseteq \mathbb{R}$ an open interval. The equivalence class $[t \mapsto (q(t), A(t))]$ is determined by:

$$(q'(0), A'(0)) = \frac{d}{dt}\bigg|_{t=0} (q(t), A(t)) \in \mathbb{R}^m \times S^r_{n,m}.$$

Now given such a curve, we can take the associated polynomial (see Definition A.0.10) of this curve at fixed p, and define this as the family of sections s_t:

$$s_t(x) = P_{p,q(t),A(t)}(x) = q(t) + \sum_{i=1}^{r} A_i(t)(x - p, \dots, x - p), \quad x \in U.$$

This has the property that $J^r s_t(p) = (p, q(t), A(t))$, which shows the claim.

Continuing the construction above, we see that locally the map is given by:

$$(p, q(0), A(0), q'(0), A'(0)) \mapsto (s_0(p), [t \mapsto J^r s_t(p)]) \mapsto (s_0(p), \frac{d}{dt}\bigg|_{t=0} J^r s_t(p)) = (p, q(0), A(0), q'(0), A'(0)),$$

where we have taken identifications $J^r \,\mathrm{Vert}(U \times W) = \mathrm{Vert}\, J^r(U \times W) = U \times V \times \mathbb{R}^m \times S^r_{n,2m}$. Hence with respect to all these identifications and trivializations, the map is simply the identity. This shows that it is smooth and well-defined. ∎

Proposition 9.1.3. *Let* $D : \Gamma(E) \to \Gamma(E')$ *be differential operator of order* r*. Then its linearization* TD : $\Gamma(\mathrm{Vert}(E)) \to \Gamma(\mathrm{Vert}(E'))$ *is also a differential operator of order* r*.*

Proof. The linearization of D is defined by Equation (9.1). With the identification as in Lemma 9.1.2 we see that for $D = \sigma(D)_* \circ J^r$ we have:

$$\begin{array}{ccc} \Gamma(\mathrm{Vert}(E)) & \xrightarrow{\;J^r\;} & \Gamma(J^r(\mathrm{Vert}(E))) \cong \Gamma(\mathrm{Vert}(J^r(E))) \\ & \searrow^{TD} & \downarrow T\sigma|_{\mathrm{Vert}(E)} \\ & & \Gamma(\mathrm{Vert}(E')) \end{array}.$$

This shows that the linearization of a differential operator is a again a differential operator. ∎

By repeating the result above inductively, we obtain the following:

Corollary 9.1.4. *Let E, E' be fiber bundles over a compact manifold M and $D : \Gamma(E) \to \Gamma(E')$ a differential operator. Then D is a smooth map between Fréchet manifolds.*

Proof. Note that the limit defining the derivative in Equation (9.1) converges uniformly for p in a compact set K. For compact M, therefore, Equation (9.1) defines the derivative in the Fréchet manifold sense. □

For *vector bundles* over compact M we show that the differential operator is not only smooth, but tame as well. Showing this first requires checking that the jet extension j^r is tame.

Lemma 9.1.5. *Let $E \twoheadrightarrow M$ be a vector bundle with M compact. Then for $r \in \mathbb{N}$ the jet extension:*

$$j^r : \Gamma(E) \to \Gamma(J^r E)$$

is a tame linear map.

Proof. For E a vector bundle, the jet bundle $J^r E$ is a vector bundle as well. From the local case (Definition 2.5.2) one can see that for sections $s : M \to E$ the estimate $\|j^r s\|_k \leq C \|s\|_{k+r}$ holds, with the constant C depending on the choice of seminorms (see Definition 2.5.12 and Lemma 2.5.13). □

Proposition 9.1.6. *Let E, E' be vector bundles over M compact, $U \subseteq E$ open and $D : \Gamma(U) \to \Gamma(E')$ a differential operator of order r. Then D is a tame smooth map.*

Proof. First we recall that $\Gamma(E)$ and $\Gamma(E')$ are indeed tame spaces by Corollary 4.4.3. By Lemma 9.1.5 above, the jet extension j^r is a tame linear map. Furthermore, by Proposition 4.4.4, the bundle operator $\sigma(D)_*$ induced by the symbol of D is tame. Now by Section 4.2 the composition:

$$D = \sigma(D)_* \circ j^r$$

is tame. Since D is smooth by Corollary 9.1.4 and any derivative of D is again a differential operator by Proposition 9.1.3 and therefore tame, D is a smooth tame map. □

Remark 9.1.7. Observe that this means that Hamilton's Nash-Moser theorem (Theorem 5.1.1) applies to differential operators D between vector bundles over some compact manifold M In particular, we can apply the result of Theorem 5.1.1 if we can invert the linearization TD some differential operator R, which is similar to the statement of Theorem 9.2.3 below. △

9.2 Infinitesimal inversions and Gromov's Theorem

In this section we provide the preliminaries for and statement of Gromov's Nash-Moser theorem. Throughout this section, all sections spaces are endowed with the *strong* Whitney topology.

First off, note that when E' is a vector bundle and $s \in \Gamma(E')$ any section, we have that $s^* \operatorname{Vert}(E') \cong E'$. Hence for a differential operator $D : \Gamma(E) \to \Gamma(E')$ of order r, the linearisation at any $s \in \Gamma(E)$ becomes a map:

$$T_s D : \Gamma(s^* \operatorname{Vert}(E)) = T_s\Gamma(E) \to T_{D(s)}\Gamma(E') = \Gamma(D(s)^* \operatorname{Vert}(E')) = \Gamma(E').$$

For the statement of Gromov's theorem we need the following:

Definition 9.2.1. *Let $E \twoheadrightarrow M$ be a fiber bundle, $E' \twoheadrightarrow M$ a vector bundle and $D : \Gamma(E) \to \Gamma(E')$ a differential operator of order r. Suppose that there is some $d \geq r$ and open subset $A \subseteq J^d(E)$ such that for any section $s \in \mathcal{A} := \Gamma_{j^d}(A)$ (see Definition 6.2.1 for this notation) we have a differential operator $R_s : \Gamma(E') \to T_s\Gamma(E)$ of order l satisfying:*

- *The global operator $R : \mathcal{A} \times \Gamma(E') \to T\Gamma(E)$ given by $(s, u) \mapsto R_s(u)$ is a differential operator of order l,*
- *At each $s \in \mathcal{A}$, R_s is a right inverse of $T_s D$, that is:*

$$T_s D(R_s(u)) = u \in \Gamma(E').$$

Then we say that D is ***infinitesimally invertible over*** *$\mathcal{A}$ with defect d and order l.*

The statement of Gromov's theorem includes not only spaces of sections that are smooth, but also spaces of C^k sections. In order to deal with this generality we introduce the following notation:

Definition 9.2.2. *Given a fiber bundle $E \twoheadrightarrow M$, we will write $\Gamma^k(M, E)$ for the space of C^k sections. When the base manifold is assumed to be understood and $k = \infty$, we will simply write $\Gamma(E) = \Gamma^\infty(E)$.*

For some subset $\mathcal{A} \subseteq \Gamma^l(E)$ and $k \geq l$, we write:

$$\mathcal{A}^k := \mathcal{A} \cap \Gamma^k(E).$$

Similarly, for E' another fiber bundle, $\mathcal{B} \subseteq \Gamma^l(E) \times \Gamma^b(E')$ and $k \geq l, a \geq b$ we write:

$$\mathcal{B}^{k,a} = \mathcal{B} \cap \left(\Gamma^k(E) \times \Gamma^a(E')\right).$$

Note that for $r \leq k$ and, the r-jet extension gives a map:

$$J^r : \Gamma^k(E) \to \Gamma^{k-r}(J^r(E)), \quad s \mapsto j^r s$$

This map is continuous with respect to the C^k and C^{k-r} Whitney topologies (see [GG74], p. 46). Since a differential operator $D : \Gamma(E) \to \Gamma(E')$ of order r is obtained by composing J^r with a smooth bundle map $\sigma(D)$, we see that $D : \Gamma^k(E) \to \Gamma^{k-r}(E')$ is continuous.

Theorem 9.2.3 (Gromov's Nash-Moser Theorem)**.** *Let $E \twoheadrightarrow M$ be fiber bundle and $E' \twoheadrightarrow M$ a vector bundle, and g be some Riemannian metric on M. Let $D : \Gamma(E) \to \Gamma(E')$ be a differential operator that is infinitesimally invertible over A with defect d and order l. Let $\sigma > \bar{l} = \max\{d, 2r + l\}$. Then there exist Whitney $C^{\sigma+l}$ opens $\mathcal{B}_s \subseteq \Gamma(E')$ and maps:*

$$D_s^{-1} : \mathcal{B}_s \to \mathcal{A} \subseteq \Gamma(E),$$

such that:

1. *$D_s^{-1}(0) = s$ (where $0 \in \Gamma(E')$ denotes the zero section);*
2. *$D(D_s^{-1}(g)) = D(s) + g$;*
3. *The set:*

$$\mathcal{B} = \cup_{s \in \mathcal{A}} \{s\} \times \mathcal{B}_s \subseteq \mathcal{A} \times \Gamma(E')$$

   is open,

4. *Let η_1, σ_1 satisfy either $\eta_1 = \sigma_1 > \sigma$ or $\eta_1 > \sigma_1 > \sigma$. Then the global operator:*

$$D^{-1} : \mathcal{B}^{\eta_1 + r + l, \sigma + l} \to \mathcal{A}^\sigma, \quad (s, u) \mapsto D_s^{-1}(u)$$

   is continuous.

5. *The operator is local: denote $B_p(1)$ for the metric ball of radius 1 centered at p. Then for $(s, u), (s', u') \in \Gamma(A) \times \Gamma(E')$ such that $(s, u)|_{B_p(1)} = (s', u')|_{B_p(1)}$ we have $D_s^{-1}(u)(p) = D_{s'}^{-1}(u')(p)$.*

Remark 9.2.4. Observe the similarity between Theorem 5.1.1 and Theorem 9.2.3. A substantial difference, however, is the locality property of the theorem above. This property will be crucial in the application to holonomic approximation, which we carry out in Section 9.4. △

The theorem below states that infinitesimally invertible operators have local (in appropriate section space) inverses.

Corollary 9.2.5 (Implicit Function Theorem)**.** *For every section $s \in \mathcal{A}$ there is a (strong) Whitney open neighborhood of $\mathcal{B} \ni s$ in $\sec E'$ such that for every smooth $g \in \mathcal{B}$ the equation:*

$$\mathcal{D}(s') = \mathcal{D}(s) + g$$

has a solution in s'.

Example 9.2.6. Consider a manifold N and another manifold M equipped with a two-form ω. Then for the bundles $N \times M \to N$ and $\Lambda^2 T^*N \to N$ over N we have an operator $\mathcal{D} : \Gamma(N \times M) \to \Omega^2(N)$ given by:

$$f \mapsto f^*\omega.$$

Given some $f \in \Gamma(N \times M)$, a point in the tangent space of $\Gamma(N \times M)$ at f is given by a section $\tilde{f}$ of $f^*\left(\Lambda^2 TM\right)$, that is, a map:

$$\begin{array}{ccc} & & \Lambda^2 T^*M \\ & \overset{\tilde{f}}{\nearrow} & \downarrow \\ N & \xrightarrow{\;f\;} & M \end{array} .$$

Given a family f_t of sections such that $f_0 = f$ and $d/dt|_{t=0} f_t = \tilde{f}$, the derivative of $\mathcal{D}$ is given by equation Equation (9.1). Using Cartan's magic formula we see that the linearisation $L_f : \Gamma\left(f^*\Lambda^2 T^*M\right) \to \Omega^2(N)$ is given by:

$$L_f : \tilde{f} \mapsto \Big(d(\tilde{f} \cdot \omega) + \tilde{f} \cdot (d\omega)\Big) \circ df_0$$

△

9.3 The differential relation of a differential equation

In Chapter 7 (in particular Example 7.1.2) we have seen that differential relations generalize differential equations. In particular, a differential equation defines a differential relation. A common way to phrase differential equations is of course in terms in differential operators, which we have defined in a general setting on manifolds above. In this section, therefore, we investigate the differential relation defined by the equation $D(s) = \eta$.

For a differential operator of order r, the differential relation should contain those elements of $[s]_{p_0} \in J^r(E)$ that can be represented by a local section $s : \mathrm{Op}(p_0) \to E$ that satisfies the differential equation. Thus we must have $D(s)(p_0) = \eta(p_0)$, which by definition of a differential operator means:

$$\sigma(D)(j^r s(p_0)) = \eta(p_0).$$

This equation defines the following closed differential relation in $J^r(E)$:

$$\mathcal{R}(D, \eta) = \{x \in J^r E \mid \sigma(D)(x) = \eta(p) \text{ where } p = \pi_M(x) \in M\}, \tag{9.2}$$

Note that holonomic solutions $j^r s : M \to E$ of $\mathcal{R}(D, \eta)$ satisfy the differential equation it is derived from since:

$$D(s)(p) = \sigma(D)(j^r s(p)) = \eta(p)$$

In contrast, for some $[s]_{p_0} \in \mathcal{R}(D, \eta)$ we do not automatically have $D(s)(p) = \eta(p)$ for all $p \in \mathrm{Op}(p_0)$. However, the requirement that $D(s)(p_0) = \eta(p_0)$ does imply, by Taylor's theorem (Theorem A.0.13), that:

$$d(D(s)(p), \eta(p)) \leq C d(p, p_0),$$

where the distance functions d on both sides derives from some metric on their respective bundles and C is some constant. In order to obtain a relation whose representative germs are better approximations, we do the following. Recall that by Corollary 6.1.14 J^a defines a functor over the category of fiber bundles over M, for any $a \in \mathbb{N}$. In particular, given the bundle map $\sigma(D) : J^r(E) \to E'$ we can pass to:

$$J^a \sigma(D) : J^a(J^r(E)) \to J^a(E').$$

Since there is a canonical embedding $J^{r+a}(E) \subseteq J^a(J^r(E))$, we can extend the differential relation $\mathcal{R}(D, \eta)$ defined in Equation (9.2) in the following way:

Definition 9.3.1. *Let $D : \Gamma(E) \to \Gamma(E')$ be a differential operator of order r and let $a \in \mathbb{N}$. Then we denote:*

$$\sigma^a(D) := J^a\, \sigma(D)|_{J^{r+a}(E)}\,.$$

And define the relation $\mathcal{R}^a(D, \eta) \subseteq J^{r+a}(E)$ by:

$$\mathcal{R}^a(D, \eta) = \left\{x \in J^{r+a} E \mid \sigma^a(D)(x) = (j^a \eta)(p) \text{ where } p = \pi_M(x) \in M\right\}.$$

Now for some germ $s : \mathrm{Op}(p_0) \to E$ representing an element $[s]_{p_0} \in \mathcal{R}^a(D,\eta)$, observe that the condition $\sigma^a(D)(j^{r+a}s(p_0)) = j^a\eta(p_0)$ implies that the Taylor polynomials of $D(s)$ and η agree up to order a. By Theorem A.0.13 this implies:

$$d(D(s)(p),\eta(p)) \leq Cd(p,p_0)^{a+1},$$

again for certain (arbitrary) choices of metric (compare [Pan16], pp. 14-15).

When considering differential operators that are infinitesimally invertible over some relation $A \subseteq J^k(E)$, we will need to consider the intersection of jets that satisfy both $\mathcal{R}^a(D,\eta)$ and A. In order to do this, we simply lift A to a higher order jet space.

Definition 9.3.2. *Let $D : \Gamma(E) \to \Gamma(E')$ be a differential operator of order r and $A \subseteq J^kE$ a differential relation. Write $A^{k+a} = (\pi_k^{k+a})^{-1}(A) \subseteq J^{k+a}E$; then for $a \geq k-r$ we define:*

$$\mathcal{R}^a(D,A,\eta) = \mathcal{R}^a(D,\eta) \cap A^{a+r} \subseteq J^{a+r}(E)$$

Once again observe that $\mathrm{Sol}(\mathcal{R}^a(D,A,\eta))$ corresponds to the intersection of the set of solutions to $D(s) = \eta$ and $\mathrm{Sol}(A)$.

9.4 Local integrability and microflexibility of operator relations

In this section we show the main result of the thesis: using Theorem 9.2.3, we obtain that the relation $\mathcal{R}^a(D,A,\eta)$ is microflexible and locally integrable for sufficiently high a. Therefore we can apply holonomic $\mathcal{R}$-approximation and obtain the h-principle in case the relation is Diff-invariant and the base manifold is open.

We begin by showing microflexibility.

Theorem 9.4.1. *Let $E \twoheadrightarrow M$ and $E' \twoheadrightarrow M$ be a fiber and vector bundle over M respectively and $D : \Gamma(E) \to \Gamma(E')$ infinitesimally invertible over some open relation A with defect d and order l. Then for $a > \max(d+l, 2r+2l)$, the differential relation $\mathcal{R}^a(D,A,\eta)$ is θ_k-microflexible for any $0 \leq k < \dim M = n$.*

Proof. We write $\mathcal{R} = \mathcal{R}^a(D,A,\eta)$ to simplify notation. Take some θ_k pair with (A,B) inside M, with B diffeomorphic to $[-1,1]^k$. Suppose we are given a holonomic solution $\widetilde{F} : \mathrm{Op}(A) \to \mathcal{R}$ and a deformation $F_t : \mathrm{Op}(B) \to \mathcal{R}$ such that $F_0 = \widetilde{F}\big|_{\mathrm{Op}(B)}$ that is constant on $\mathrm{Op}(\partial B)$ (meaning $F_t = F_0$ on $\mathrm{Op}(\partial B)$).

Fix a small open neighborhood V of ∂B such F_t is defined on V for all t and is constant on V. Set $B' = B \setminus V$ and take a small open neighborhood of $U \supseteq B'$ such that $U \cap \mathrm{Op}(\partial B) = \emptyset$ and F_t is defined on U for all $t \in [0,1]$. Then find precompact opens $B' \subseteq U_0 \subseteq U_1 \subseteq U$ such that $\overline{U_0} \subseteq U_1$ and $\overline{U_1} \subseteq U$. Now choose a bump function $\varphi : M \to [0,1]$ supported in U_1 such that $\varphi|_{U_0} \equiv 1$. Then we define:

$$\widetilde{F}_t(p) = \begin{cases} F_{t\varphi(p)}(p) & \text{for } p \in U \\ \widetilde{F}(p) & \text{for } p \in \mathrm{Op}(A) \setminus U \end{cases}.$$

Note that $\widetilde{F}_0 = \widetilde{F}$ and $\widetilde{F}_t(p) = F_t(p)$ for all $p \in B$.

Now we apply Theorem 9.2.3, where we take $M = \mathrm{Op}(A)$ and restrict the bundles and D appropriately (in order to avoid having to extend every local section over $\mathrm{Op}(A)$) and a metric such that $d(B, M \setminus U_0) > 1$ and $d(U_1, M \setminus U) > 1$. Note that for small t — belonging to some interval $[0,\varepsilon]$ — we have that:

$$\eta - D(\widetilde{F}_t) \in \mathcal{B}_{\widetilde{F}_t}$$

by openness of the set on the right hand side, $D(\widetilde{F}_0) = \eta$ and continuity of D. Therefore we may apply $D^{-1}_{\widetilde{F}_t}$ to obtain the following section on M:

$$G_t = D^{-1}_{\widetilde{F}_t}\left(\eta - D(\widetilde{F}_t)\right)$$

The requirement on the metric ensures that:

- any metric ball of radius one centered in B is contained in U_0

- any metric ball of radius one centered in $M \setminus U$ is contained in $M \setminus U_1$.

Since the bump function φ is constant on U_0 and $M \setminus U_1$, the locality (property 5. of Theorem 9.2.3) implies that $G_t|_B = F_t$ and $G_t|_{M\setminus U} = \widetilde{F}_0\Big|_{M\setminus U}$. In particular G_t is constant on $\mathrm{Op}(\partial A)$ since U avoids ∂A. Finally, applying the differential operator D gives us:

$$D(G_t) = D\left(D^{-1}_{\widetilde{F}_t}\left(\eta - D(\widetilde{F}_t)\right)\right) = D(\widetilde{F}_t) + \eta - D(\widetilde{F}_t) = \eta,$$

and openness of the relation A ensures that $G_t \in \mathrm{Sol}(A)$ for small t.

The argument above is straightforward to adapt to the parametric case, i.e. when given families of θ_k pairs (A_x, B_x), sections $\widetilde{F}_x : \mathrm{Op}(A_x) \to \mathcal{R}$ and homotopies $F_{x,t} : \mathrm{Op}(B_x) \to \mathcal{R}$ parametrized by x in some compact set P. ☐

Additionally, Theorem 9.2.3 implies local integrability.

Theorem 9.4.2. *Let $E \twoheadrightarrow M$ and $E' \twoheadrightarrow M$ be a fiber and vector bundle over M respectively and $D : \Gamma(E) \to \Gamma(E')$ infinitesimally invertible over some open relation A with defect d and order l. Then for $a > \max(d+l, 2r+2l)$, the differential relation $\mathcal{R}^a(D, A, \eta)$ is locally integrable.*

Proof. Suppose we are given maps $h : [0,1]^k \to M$ and $g : [0,1]^k$ as in Definition 8.2.1. First off, we can construct a smooth family of sections $F_x : \mathrm{Op}(h(x)) \to J^r(E)$ as in Example 8.2.2. Given some metric on M, we can cover $\mathrm{Graph}\, h \subseteq [0,1]^k \times M$ by square opens of the form $U_{x_0} \times V_{x_0}$, where V_{x_0} is some metric ball and $U_x \ni x$ a cube with sufficiently small sides δ.

This gives us continuous parametrized families of sections $F_x : V_{x_0} \to J^r E$ for $x \in U_{x_0}$. Since $F_x \in \mathcal{R}^a(D, A, \eta)$ we have that:

$$e_x = (\eta - D(F_x))(h(x)) \in \Gamma(E')$$

satisfies $j^a e = 0$. By choosing an appropriate bump function we can obtain a smooth family sections $e'_x \in \Gamma(E')$ such that:

- $e_x \equiv e'_x$ on $\mathrm{Op}(h(x))$,
- $j^a e'_x$ is arbitrariliy small.

Now we apply Theorem 9.2.3 and choose e' such that it lies in the Whitney C^a neighborhood $\mathcal{B}_{F_x} \in \Gamma(E')$ on which we have a right inverse $D^{-1}_{F_x}$. Then we see that:

$$G_x = D^{-1}_{F_x}(e'_x)$$

satisifies:

$$D(G_x) = D\left(D^{-1}_{F_x}(e'_x)\right) = D(F_x) + e'_x.$$

By the definition of e'_x, we see that $D(G_x) = \eta$ on $\mathrm{Op}(h(x))$. Therefore we have obtained a family of local solutions G_x such that $G_x(h(x)) = F_x(h(x))$.

To extend this "patchwork" of local families of solutions to $\mathrm{Op}(\mathrm{Graph}\, h)$, we do the following. On overlaps $W = U' \times V' \cap U \times V \neq \emptyset$, we may interpolate between the families G'_x and G_x on $U' \times V'$ and $U \times V$ respectively. One can do this, for example, by letting these overlaps (or indeed the sets $U \times V$ themselves) be small enough to be trivializable, and in this trivialization one may interpolate linearly.

Let $\widetilde{G}^t_x$ be such an interpolation on the overlap W such the value $t = 1$ corresponds to G_x and $t = 0$ to G'_x. Set $\widetilde{G}_x = \widetilde{G}^{\varphi(x)}_x$ for some bump function supported φ in $U' \times V'$ such that $\varphi \equiv 1$ on a large open inside $U' \times V$" (specifically on $\mathrm{Op}(\partial(U \times V)) \cap U' \times V$"). Applying the procudure using Theorem 9.2.3 exactly as above on W, we obtain a family of solutions that coincides with G_x and G'_x near the edges of W.

This argument works readily in case $[0,1]^k = [0,1]$ is the interval; for the general case the result follows from a (highly technical) combinatorial inductive argument, which we will not state here. ☐

Theorems 9.4.1 and 9.4.2 immediately imply the following application of Theorem 8.2.11:

Theorem 9.4.3. *Let $E \twoheadrightarrow M$ and $E' \twoheadrightarrow M$ be a fiber bundle and vector bundle over an open manifold M respectively Let $D : \Gamma(E) \to \Gamma(E')$ be a differential operator of order r that is infinitesimally invertible over A with defect d and order l. For*

$$a > \max(d + l, 2r + 2l),$$

write $\mathcal{R}^a(D, A, \eta)$ for the relation defined in Definition 9.3.2. If this relation is diffeomorphism invariant, then it satisfies the h-principle.

Remark 9.4.4. Observe that if $\mathcal{R}^0(D, \eta)$ and A are diffeomorphism invariant, then so is $\mathcal{R}^a(D, A, \eta)$. △

Chapter 10

Outlook

In this section we apply the results and theories developed so far to make some inroads into solving the problem of almost symplectic isotropic immersions. The approach is based on [Gro96], section 4.2 and [Pan16] sections 3 - 6.

10.1 Linearization of the pullback operator

The setting of our problem is as follows. Given an almost symplectic manifold (M,ω) and some manifold N, a function $f : N \to M$ is equivalent to a section of the fiber bundle $N \times M \twoheadrightarrow N$. Then we consider the following differential operator:

$$D : \Gamma(N \times M) \to \Gamma(\wedge^2 T^*N) = \Omega^2(N), \quad f \mapsto f^*\omega.$$

We wish to find an h-principle for maps such that:

$$D(f) = f^*\omega = 0. \tag{10.1}$$

In order to do this, we apply Theorem 9.4.3. First we remark that the equation $f^*\omega = 0$ is diffeomorphism invariant: for any diffeomorphism $\phi : N' \to N$, we see that $(f \circ \phi)^* = \phi^* f^* \omega = 0$ (see also Example 7.2.5). For the application of Theorem 9.4.3 we also need to show that we can infinitesimally invert the linearization TD of D over some differential relation. To compute the linearization of D we use Cartan's formula:

$$T_{f_0}D(X) = \frac{d}{dt}\bigg|_{t=0} f_t^*\omega = (f_0)^* \left(\iota_X d\omega + d\iota_X \omega\right)$$

where $X : N \to TM$ is the map such that $X = \frac{d}{dt}\big|_{t=0} f_t$. Therefore we have to solve for X given some two-form $\eta \in \Omega^2(N)$ and map $f : N \to M$ in the following equation:

$$f^* \left(\iota_X d\omega + d\iota_X \omega\right). \tag{10.2}$$

A general strategy for solving the linearized equation is by making assumptions or finding conditions on f and X such that one of the terms in Equation (10.2) vanishes. Note that at a point $p \in N$, the pullback by the immersion f can be seen as the restriction to the subspace $V_p := \operatorname{Im} T_p f \subseteq T_{f(p)}M$. Denoting V_p^ω for the symplectic orthogonal of V_p, we may for example assume that $X_p \in V_p^\omega$ and so Equation (10.2) becomes:

$$(T_f DX)(p) = \left(\iota_{X_p} d\omega_{f(p)} + d\iota_{X_p}\omega_{f(p)}\right)\big|_{V_p} = \iota_{X_p} d\omega_{f(p)}\big|_{V_p} = \eta_p$$

Under favorable conditions we expect the above to be algebraically solvable at a single point p and that we can choose such algebraic inverses smoothly with respect to p. This means we have an infinitesimal inverse of order zero.

One might also try to make the other term in Equation (10.2) vanish if (generically) $\operatorname{Ker} d\omega_{f(p)} \neq \{0\}$ and $X_p \in \operatorname{Ker} d\omega_{f(p)}$. Then we get the equation:

$$d\iota_{X_p}\omega_p\big|_{V_p} = \eta_p \tag{10.3}$$

However, we remark that in this case we take a derivative of the "vector field" $X : N \to TM$ that we are solving for. As such we expect that solving for this X involves finding a primitive of some kind so in general the assignment $R_f : \eta \mapsto X$ is not a differential operator. For instance, we do not expect it to be local, meaning that $\eta|_U = \eta'|_U$ on any open U implies that $R_f(\eta)|_U = R_f(\eta')|_U$. This means we cannot apply Theorem 9.4.3.

We will therefore continue our approach by finding conditions on f on which we can solve:

$$f^*(\iota_X d\omega) = \eta. \tag{10.4}$$

10.2 Regular subspaces

We start off by examining the linear algebra related to our problem. First we require that we can make the first term of Equation (10.2) vanish, so we require that there exists $X_p \in (V_p)^\omega \setminus V_p$. This is, however, always the case as long as $\dim V_p < 2\dim M$, which we assume from now on. In order to solve Equation (10.4) pointwise, we must have that:

$$X_p \mapsto \left(\iota_{X_p} d\omega_{f(p)}\right)\big|_{V_p}$$

is surjective for $X_p \in (V_p)^\omega$. To simplify notation, we consider an $2n$-dimensional symplectic vector space (W, ω), supposed to represent (T_pM, ω_p), and a k-dimensional subspace $V \subseteq W$.

Definition 10.2.1. *Let (W, ω) be a symplectic vector space and $\beta \in \Lambda^3 W^*$ a three-form. We say that a subspace $V \subseteq W$ is (ω, β)-**regular** if the map $G = G_{\omega,\beta,V} : V^\omega \to \Lambda^2 V^*$ defined by:*

$$w \mapsto (\iota_w \beta)|_V.$$

*is surjective. When the pair (ω, β) is clear from the discussion, we will simply refer to such V as "regular". We also say that the pair (ω, β) is k-**admissible** if there exists a regular subspace of dimension k for the pair, dropping the dimension to speak of "admissible" whenever it is clear from the context.*

Note that in order for G to be surjective, we see that we must have $\dim(V^\omega) = 2n - k \geq \binom{k}{2} = \dim(\Lambda^2 V)$. We state the following without proof:

Proposition 10.2.2. *For a fixed k-regular pair (ω, β), the set of k-dimensional regular spaces is an open subset of the Grassmannian $\mathrm{Gr}_k(W)$.*

Intuitively, this can be justified by stating that a small perturbation of V will not change the surjectivity of $G_{\omega,\beta,V}$.

Now we return to the manifold setting. First up we will define the geometric analogues of the notions defined in Definition 10.2.1:

Definition 10.2.3. *We say $\omega \in \Omega^2(M)$ is k-**admissible** if $(\omega_q, d\omega_q)$ is k-admissible for every $q \in M$. An immersion $f : N \to (M, \omega)$ is said to be ω-**regular** if $\mathrm{Im}\, T_p f$ is $(\omega_{f(p)}, d\omega_{f(p)})$-regular at every $p \in N$.*

Since the set of regular isotropic subspaces in a vector space is open, we see that:

Proposition 10.2.4. *Let N be a manifold and (M, ω) a $\dim N$-admissible almost symplectic manifold. Then the differential relation $\mathcal{R}_{reg}$ of regular maps $f : N \to M$ is open.*

The following is adapted from [Pan16], p. 16. It shows that pointwise regularity of $f : N \to (M, \omega)$ suffices to obtain a smooth right inverse $\eta \mapsto X$ of $X \mapsto f^*\iota_X d\omega = \eta$.

Proposition 10.2.5. *Let (M, ω) be an almost symplectic manifold and $f : N \to (M, \omega)$ an ω-regular map. Then there exists a smooth right inverse to $X \mapsto f^*\iota_X d\omega$.*

Proof. From Proposition 10.2.2 we see that the relation $\mathcal{R}_{\mathrm{reg}} \subseteq J^1(N, M)$ of regular immersions is an open subset. For every $(p, q, A : T_pN \to T_qM) \in \mathcal{R}_{\mathrm{reg}}$, the map:

$$G_{p,q,A} := G_{\omega_q,\beta_q,\mathrm{Im}\, A} : (\mathrm{Im}\, A)^{\omega_p} \to \Lambda^2(\mathrm{Im}\, A)^*$$

(as defined in Definition 10.2.1) is surjective and therefore admits a right inverse $R_{p,q,A}$. The space of right inverses is affine since for any map L such that $G_{p,q,A} \circ L = 0$, $R_{p,q,A} + L$ is again a right inverse. Varying over points (p, q, A) in $\mathcal{R}_{\mathrm{reg}}$ we get a smooth bundle $\mathrm{RightInv}(N, M, \omega)$ consisting of right inverses to $G_{p,q,A}$. Since the fibers are affine and therefore contractible, this admits a smooth section $S : \mathcal{R}_{\mathrm{reg}} \to \mathrm{RighInv}(N, M, \omega)$. Then for $f \in \mathrm{Sol}\, \mathcal{R}_{\mathrm{reg}}$ a right inverse to $X \mapsto f^*\iota_X d\omega$ is given by $S \circ j^1 f$. $\square$

With Proposition 10.2.5 in mind, we concentrate on the problem of pointwise algebraic invertibility of the differential operator. First we consider surfaces and afterwards we extend this to higher dimensional manifolds N, by finding a condition on the dimensions that implies existence of a $\dim N$-admissible almost symplectic form ω on M (or at least *almost everywhere* admissible).

10.3 Isotropic surfaces

We are interested in the case where $\dim V = 2$ and $\dim W = 2 \cdot l$ for $l > 2$. In the following discussion, we denote ω and β for a nondegenerate two-form and arbitrary three-form on W respectively. Note that $\mathrm{Span}\{x, y\}$ is regular if it satisfies:

$$\begin{cases} \omega(w, x) &= 0 \\ \omega(w, y) &= 0, \end{cases}$$

and $\beta(w, x, y) \neq 0$. This is because for $\dim V = 2$, the map $w \mapsto \iota_w \beta$ being surjective amounts to stating that $\iota_w \beta \neq 0$, which in turn means that $\beta(w, x, y) \neq 0$ for some w and all independent $x, y \in V$. Denote $Z_{\omega,\beta}$ for the pairs of triples (w, x, y) satisfying the equations above.

We claim that 2-admissibility is an open condition on pairs of nondegenerate two-forms and arbitrary three-forms. To see this, first observe that the condition on β is open: if there are w, x, y such that $\beta(w, x, y) \neq 0$, then for β_1 sufficiently close to β we still have $\beta_1(w, x, y) \neq 0$.

For the condition on ω, remark that for ω_1 near ω we can find a symplectic linear isomorphism $L : (W, \omega) \to (W, \omega_1)$ with $|L - 1|$ small. Since (ω, β) is admissible we may pick $w, x, y \in Z_{\omega,\beta}$. Observe that:

$$\begin{cases} \omega_1(Lw, Lx) &= \omega(w, x) = 0 \\ \omega_1(Lw, Ly) &= \omega(w, y) = 0 \end{cases}$$

Furthermore, for a small L we maintain $\beta(Lw, Lx, Ly) \neq 0$. Indeed, this is the requirement that we take for ω_1 to be near ω. We therefore see that:

$$(Lw, Lx, Ly) \in Z_{\omega_1,\beta}.$$

This discussion shows the first part of the following:

Proposition 10.3.1. *The set of 2-admissible pairs (ω, β) for $\dim W = 2 \cdot l, l > 2$ is Zariski open. In fact it is dense, since it is nonempty.*

Proof. Openness was shown above. For any $w, x, y \in W$, we can construct a pair (ω, β) such that $(w, x, y) \in Z_1 \setminus Z_2$. For example, we can take a basis $\{a_1, \dots, a_n, b_1, \dots, b_n\}$ such that $a_1 = w, a_2 = x, a_3 = y$ and define:

$$\omega = \sum_i a_i \wedge b_i$$

and $\beta = w \wedge x \wedge y$. Hence the set of 2-admissible pairs (ω, β) is nonempty. Density follows since nonempty Zariski opens are dense. □

In the discussion so far we remark that β is supposed to represent $d\omega$, which depends solely on $j^1\omega$. By the genericity of (ω, β) being admissible, we have the following:

Proposition 10.3.2. *The relation $\mathcal{R}_{2-adm} \subseteq J^1(\Lambda^2 T^* M)$ of $\omega \in \Omega^2(M)$ such that $(\omega, d\omega)$ is 2-admissible is open and dense for $\dim M \geq 6$ even. Therefore a generic ω on M is 2-admissible on a dense subset of M.*

This result says that in general we expect to be ω to be 2-admissible on a $2 \cdot l$-dimensional (with $l > 2$) manifold M, or at least at almost every point of M. If (perhaps after removing these points), we obtain a 2-admissible almost symplectic manifold, we obtain the following:

Proposition 10.3.3. *Let (M, ω) be an almost symplectic manifold of dimension greater that four with 2-admissible form ω. Then the operator $D : f \mapsto f^*\omega$ on $C^\infty(M, N)$ is infinitesimally invertible over $\mathcal{R}_{reg}$ with defect 1 and order 0. Therefore the relation:*

$$\mathcal{R}^a(D, \mathcal{R}_{reg}, 0)$$

satisfies the parametric h-principle for all $a \geq 3$.

Proof. Follows from the preceding discussion and Theorem 9.4.3. □

10.4 Isotropics of higher dimensions

We extend the previous discussion to $\dim N \geq k$.

Proposition 10.4.1. *Let W be a $2n$-dimensional vector space. Then the set of k-admissible pairs on W is open.*

Proof. If the set of admissible pairs is empty, then it is open. Therefore we assume it is nonempty and choose (ω, β) k-admissible, with V a k-dimensional regular subspace. We argue that a pair (ω_1, β_1) close to (ω, β) is k-admissible.

We first consider β_1 close to β. By definition of the regularity of V, the map:

$$G_{\omega,\beta,V} : V^\omega \to \Lambda^2 V^*, \qquad w \mapsto (\iota_w \beta)|_V$$

has full rank. A linear map having full rank is an open condition on the space of linear maps, so any linear map $L : V^\omega \to \bigwedge^2 V^*$ close to $G_{\omega,\beta,V}$ has full rank. In particular this holds for $G_{\omega,\beta_1,V}$ with β_1 close to β.

Now we consider some $\omega_1 \neq \omega$ and let L be a linear isomorphism such that $L^*\omega_1 = \omega$. Write $L|_V : V \to LV$ for the restriction to V. We claim that we have the following:

$$G_{\omega_1,(L^{-1})^*\beta,LV} = (L|_V^*)^{-1} \circ G_{\omega,\beta,V} \circ L^{-1}.$$

With this result, it follows that for ω_1 close enough to ω (meaning that $|L - \mathrm{Id}|$ is small) the map $G_{\omega_1,\beta,LV}$ is surjective. To see this, remark that we have previously shown that for small L the map $G_{\omega,L^*\beta,V}$ is surjective. By the equation above,

$$G_{\omega_1,\beta,LV} = G_{\omega_1,(L^{-1})^*L^*\beta,LV}$$

is obtained by composing this map with isomorphisms and is therefore surjective as well.

We proceed to show the claim. First we observe the following:

$$V^\omega = L^{-1}\left(L(V)^{\omega_1}\right).$$

Additionally, we will use the following in the upcoming computation. Writing $j_V : V \hookrightarrow W$ for any inclusion, we have $j_V = L^{-1} \circ j_{LV} \circ (L|_V)$. Since restriction of forms is defined by the pullback of the inclusion, we get for any (two-)form η that:

$$\eta|_V = (L|_V)^* \left[(L^{-1})^*\eta\right]\Big|_{LV}. \tag{10.5}$$

Now let $w \in (LV)_1^\omega$. Then $L^{-1}w \in V^\omega$ so we may take:

$$G_{\omega,\beta,V}(L^{-1}w) = [\iota_{L^{-1}w}\beta]|_V.$$

By Equation (10.5) we see:

$$G_{\omega,\beta,V}(L^{-1}w) = (L|_V)^* \left(\left(L^{-1}\right)^* [\iota_{L^{-1}w}\beta]\right)\Big|_{LV}.$$

Note that $(L^{-1})^*(\iota_{L^{-1}w}\beta) = \iota_w((L^{-1})^*\beta)$. Substituting this in the above gives us:

$$G_{\omega,\beta,V}(L^{-1}w) = (L|_V)^* \left(\iota_w(L^{-1})^*\beta\right)\Big|_{LV} = \left((L|_V)^* G_{\omega_1,(L^{-1})^*\beta,LV}\right)(w).$$

From this it follows that:

$$G_{\omega,\beta,V} \circ L^{-1} = (L|_V)^* \circ G_{\omega_1,(L^{-1})^*,LV},$$

as required. □

We have shown that the algebraic set of admissible pairs is open, meaning it is a Zariski open set. If we are able to show that it is nonempty it follows that it is dense. The following results obtains a condition on the dimension k such that the set of k-admissible pairs is nonempty.

Proposition 10.4.2. *For W be a $2n$-dimensional subspace, the space of k-admissible pairs (ω, β) is nonempty, and therefore dense, whenever:*

$$n \geq \tfrac{1}{2}\binom{k}{2} + k.$$

Proof. We take any k-dimensional subspace $V \subseteq W$ and construct (ω, β) such that V is regular. By taking a suitable linear isomorphism we take $W = \mathbb{R}^{2n}$ with coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ and:

$$V = \mathbb{R}^k \times \{0\} = \{(x_1, \ldots, x_k, 0 \ldots, 0)\}.$$

For ω we take the standard symplectic form: $\omega = \sum_i dx^i \wedge dy^i$. Then we see:

$$V^\omega = \{(x_1, \ldots, x_n, 0, \ldots, y_{k+1}, \ldots y_n)\}.$$

Now if $2(n-k) \geq \binom{k}{2}$, it is possible to uniquely match to any pair $1 \leq i < j \leq k$ one of the coordinates $x_{k+1}, \ldots, x_n$ or $y_{k+1}, \ldots, y_n$. Denote $w_{i,j}$ for the coordinate vector matched to such a pair (i, j). Then we see that:

$$\beta = \sum_{1 \leq i < j \leq k} dw_{i,j} \wedge dx^i \wedge dx^j$$

makes the subspace V regular. Indeed, since $\bigwedge^2 V^*$ is spanned by $dx^i \wedge dx^j$, it suffices to observe that there is some coordinate vector $w_{i,j} \in V^\omega$ for any (i, j) such that $\iota_{w_{i,j}} \beta = dx^i \wedge dx^j$. □

10.5 Further research

The preceding sections provide the first step to proving an h-principle for almost symplectic istropic immersions. Here provide some possible directions for generalizing or strengthening the results above.

Above, we have taken the approach to reduce Equation (10.2) to Equation (10.4). We mentioned in Section 10.1 that we expect to encounter obstacles when reducing Equation (10.2) to Equation (10.3), since the inverse would require some kind of "primitive" of the dependent variable X. Although this would mean the assignment $\eta \mapsto X$ is not a differential operator, the inverse of the linearization might very well still be *tame*, which puts forward the question whether Theorem 5.1.1 (Hamilton's Nash-Moser Theorem) could be applied in this situation.

Another difficulty to overcome is that for generic almost symplectic ω, the infinitesimal inversion we describe is still only possible for a dense set of points. It is therefore natural to ask what conditions may ensure that generic ω are regular everywhere. Alternatively, there could be a more detailed investigation of the singularities, perhaps providing a method of resolving these in a suitable way.

Finally, we note the possibility of extending the results so far, applicable only to *open* manifolds, to *closed manifolds*. A clear candidate method would be using a form of the *microextension trick*, discussed in Chapter 9 of [CEM24].

Appendix A

Taylor polynomials

In this appendix we develop some useful notation and machinery for Taylor polynomials of maps $U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, where U is an open subset. The convention that U is open will be used throughout this chapter. For the reader's convenience, we repeat the definition of higher-order derivatives (see also Definition 2.4.7):

Definition A.0.1. *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be smooth, then the* ***first order (total) derivative*** *of f is defined as:*

$$(Df)(p, h) = \lim_{t \to 0} \frac{f(p + th) - f(p)}{t}.$$

Inductively, we define the ***k-th order derivative*** *$D^k f : U \times (\mathbb{R}^n)^k \to \mathbb{R}^m$ for $k \geq 2$ as:*

$$(Df)(p, h_1, \ldots, h_k) = \lim_{t \to 0} \frac{Df(p + th_k, h_1, \ldots, h_{k-1}) - Df(p, h_1, \ldots, h_{k-1})}{t}.$$

Remark A.0.2. We will denote the higher-order derivative of a function $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ (as defined in above and in Definition 2.4.7) with:

$$D^k_p f(h_1, \ldots, h_k) := (D^k f)(p, h_1, \ldots, h_k).$$

We will also use the convention that $D^0_x f = f(x)$. △

As shown in Chapter 2, the k-th order derivative is symmetric and multilinear. On finite-dimensional spaces, we denote the space of symmetric multilinear maps as follows.

Definition A.0.3. *We define $S^k(\mathbb{R}^n, \mathbb{R}^m)$ for the space of symmetric k-linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. We will also denote:*

$$S^r_{n,m} = \bigoplus_{k=1}^{r} S_k = S^1 \times \ldots \times S^r$$

Remark A.0.4. We remark that $S^k(\mathbb{R}^n, \mathbb{R}^m)$ inherits an inner product from Euclidean space in the following way. For $A, B \in S^k(\mathbb{R}^n, \mathbb{R}^m)$, we define:

$$\langle A, B \rangle_{k,n,m} = \sum_{1 \leq i_1, \ldots, i_k \leq n} \langle A(e_{i_1}, \ldots, e_{i_k}), B(e_{i_1}, \ldots, e_{i_k}) \rangle_m,$$

where $\langle \ldots, \ldots \rangle_m$ is the standard inner product on $\mathbb{R}^m$. We will do without the indices and denote all these inner products simply with $\langle \ldots, \ldots \rangle$. Given such an inner product, we will denote the induced norm by:

$$|A| = \sqrt{\langle A, A \rangle}.$$

△

We repeat the result of Theorem 2.4.10:

Proposition A.0.5. *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be smooth. Then:*

$$D_p^k f \in S^k(\mathbb{R}^n, \mathbb{R}^m)$$

for all $p \in U$ and $k \in \mathbb{N}$.

Definition A.0.6. *We denote $\mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m)$ for the space of polynomial maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, and $\mathrm{Pol}^k(\mathbb{R}^n, \mathbb{R}^m)$ for the subspace consisting of homogeneous order k maps.*

Note that the spaces above are all (finite-dimensional) vector spaces. We state the following well-known result without proof:

Proposition A.0.7. *There is a canonical isomorphism between $S^k(\mathbb{R}^n, \mathbb{R}^m)$ and $\mathrm{Pol}^k(\mathbb{R}^n, \mathbb{R}^m)$ that maps $A \in S^k(\mathbb{R}^n, \mathbb{R}^m)$ to:*

$$\left(\mathring{A} : x \mapsto A(x, \ldots, x)\right) \in \mathrm{Pol}^k(\mathbb{R}^n, \mathbb{R}^m)$$

Lemma A.0.8. *Let $A \in S^k(\mathbb{R}^n, \mathbb{R}^m)$. Then for $0 \leq l \leq k$ and $\mathring{A}$ as above we have:*

$$D_x^l \mathring{A}(h_1, \ldots, h_l) = \frac{k!}{(k-l)!} A(x, \ldots, x, h_1, \ldots, h_l)$$

(That is: we add $k - l$ copies of x in the first entries.) In particular, we have:

$$D_x^k \mathring{A}(h_1, \ldots, h_k) = k! A(h_1, \ldots, h_k).$$

Proof. The derivative of $\mathring{A}$ is given by:

$$\begin{aligned} D_x \mathring{A}(h) &= \lim_{t \to 0} \frac{A(x + th, \ldots, x + th) - A(x, \ldots, x)}{t} \\ &= \lim_{t \to 0} \frac{kA(x + th, x, \ldots x) + O(t^2)}{t} \\ &= kA(h, x, \ldots, x) = kA(x, \ldots, x, h) \end{aligned}$$

Continuing this argument inductively proves the result. □

The result above gives us a means to explicitly specify a function that has given derivatives up to a certain order:

Corollary A.0.9. *Let $p \in \mathbb{R}^n$, $q \in \mathbb{R}^m$ and $A_k \in S^k(\mathbb{R}^n, \mathbb{R}^m)$ for $1 \leq k \leq r$. Then the function $f : \mathbb{R}^n \to \mathbb{R}^m$ defined by:*

$$f(x) = q + \sum_{k=1}^{r} \frac{1}{k!} \mathring{A}_k(x - p) = q + \sum_{k=1}^{r} \frac{1}{k!} A_k(x - p, \ldots, x - p)$$

satisfies:

$$D_p^k f(h_1, \ldots, h_k) = A_k(h_1, \ldots, h_k)$$

for all $0 \leq k \leq r$.

Proof. This follows from Lemma A.0.8 and the observation that for $l < k$ we have:

$$D_p^l \mathring{A}_k(h_1, \ldots, h_l) = \frac{k!}{(k-l)!} A(p - p, \ldots, p - p, h_1, \ldots h_l) = 0.$$

On the other hand, for $l > k$, we see that:

$$D_x^l \mathring{A}_k = D_x^{l-k}(k! A_k(h_1, \ldots h_k)) = 0,$$

since the expression within the derivative on the right hand side does not depend on x. □

Definition A.0.10. *For $p \in \mathbb{R}^n$, $q \in \mathbb{R}^m$ and $A = (A_1, \ldots, A_r) \in S_{n,m}^r$ we call the polynomial function $P_{p,q,A} : \mathbb{R}^n \to \mathbb{R}^m$ defined by:*

$$P_{p,q,A}(x) = q + \sum_{k=1}^{r} \frac{1}{k!} A_k(x - p, \ldots, x - p)$$

*the **polynomial associated to** p, q and A.*

It is now natural to introduce the following generalization of the Taylor polynomial (which we will simply refer to as Taylor polynomial):

Definition A.0.11. *Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a smooth function. For $r \in \mathbb{N}$ and $p \in U$ we define the function $P^r_{f,p} : \mathbb{R}^n \to \mathbb{R}^m$ to be the polynomial associated to $p, f(p)$ and $(D^k_p f \mid 1 \leq k \leq r)$. Explicitly:*

$$P^f_{p,r}(x) = \sum_{k=0}^{r} \frac{1}{k!} D^k_x f(x-p, \ldots, x-p).$$

*We will call $P^f_{p,r}$ the **r-order Taylor polynomial at p (of f)**.*

This definition for the Taylor polynomial indeed has the desirable property that the derivatives of f and $P^f_{p,r}$ coincide up to order r at p.

Proposition A.0.12. *Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a smooth function. Then for any $p \in U$ and $0 \leq k \leq r \in \mathbb{N}$ we have:*

$$D^k_p \left(P^f_{p,r}\right)(h_1, \ldots, h_k) = D^k_p f(h_1, \ldots, h_k)$$

for any $h_1, \ldots, h_k \in \mathbb{R}^n$.

Proof. Follows immediately from Corollary A.0.9 □

The Taylor polynomial of a function also approximates the function in the sense of the folowing theorem, which is a adapted from the one found in [Lee12].

Theorem A.0.13 (Taylor's Theorem ([Lee12], Theorem C.15))**.** *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a C^{l+1} smooth function and $p \in U$ fixed. Then for any compact convex subset $p \in W \subseteq \mathbb{R}^n$ and $k \leq l$ we have:*

$$f(x) = P^f_{p,k}(x) + R_k(x)$$

for all $x \in W$, where the remainder term $R_k(x) : W \to \mathbb{R}^m$ satisfies:

$$|R_k(x)| \leq C|x-a|^{k+1}.$$

Appendix B

The Inverse Function Theorem and ODEs factoring through Banach space

In this appendix we show the "regular" inverse function theorem for maps between Banach spaces, following [Ham82]. We also show that a certain class of ODEs in Fréchet space has unique solutions, which is used in the proof of the Nash-Moser theorem. For both of these, we will use some preliminary general results about differentiable maps between Fréchet spaces (hence in particular Banach spaces).

B.1 Estimates of differentiable maps

Lemma B.1.1. *Let $L : (U \subseteq F) \times G \to H$ be a continuous family of linear maps. Then for any point $f_0 \in U$ and seminorm $\|\ldots\|_H$ there is a neighborhood $f_0 \in V \subseteq U$, a seminorm $\|\ldots\|_G$ of G and a constant $C > 0$ such that:*

$$\|L(f \mid g)\|_H \leq C \|g\|_G$$

for all $f \in V$ and $g \in G$.

Proof. Since L is continuous and $L(f_0 \mid 0) = 0$, there is a neighborhood W with $(f_0, 0) \in W \subseteq U \times G$ such that $\|L(f \mid g)\|_H < 1$. By definition of the product topology, there is some neighborhood V containing f_0 and some seminorm $\|\ldots\|_G$ and $\epsilon > 0$ such that:

$$V \times \{\|g\|_G < \epsilon\} \subseteq W$$

For any nonzero $g \in G$ and $f \in V$ $(f, \frac{1}{2}\epsilon g / \|g\|_H)$ is contained in the left-hand side above so we have:

$$\|L(f \mid g)\|_H \leq \frac{1}{\epsilon} \|g\|_G ,$$

which shows the result. □

By a similar proof, we obtain:

Lemma B.1.2. *Let $L : (U \subseteq F) \times G \times H \to K$ a family of bilinear maps. Then for any $f_0 \in U$ and seminorm $\|\ldots\|_K$ there is a neighborhood $f_0 \in V \subseteq U$ and seminorms $\|\ldots\|_G$, $\|\ldots\|_H$ such that:*

$$\|L(f \mid g, h)\|_K \leq C \|g\|_G \|h\|_H ,$$

for all $f \in V$, $g \in G$ and $h \in H$.

Proposition B.1.3. *Let $P : (U \subseteq F) \to G$ be a C^1 map. Then for any $f_0 \in U$ and seminorm $\|\ldots\|_G$ of G there is an $\epsilon > 0$ and a seminorm $\|\ldots\|_F$ of F such that for any f_1, f_2 with $\|f_1 - f_0\|_F < \epsilon$ and $\|f_2 - f_0\|_F < \epsilon$ we have:*

$$\|P(f_2) - P(f_1)\| \leq C \|f_2 - f_1\| .$$

Proof. By Lemma B.1.1 there is a neighborhood V of f_0 such that we have:

$$\|DP(f \mid k)\|_G \leq C \|k\|_F$$

for some seminorm $\|\ldots\|_F$ and some constant C. By taking the maximum of two seminorms if necessary, there is an $\epsilon > 0$ such that V contains $\{\|f - f_0\|_F < \epsilon\}$. Replace V by this open set. Now by Theorem 2.3.3 we can write $P(f_2) - P(f_1)$ as:

$$P(f_2) - P(f_1) = \int_0^1 DP(f_t \mid f_2 - f_1)dt,$$

where $f_t = tf_2 - (1-t)f_1$. Since V is convex, $f_t \in V$ whenever $f_2, f_1 \in V$ and therefore we obtain:

$$\|P(f_2) - P(f_1)\|_G \leq \int_0^1 \|DP(f_t \mid f_2 - f_1)\|_G \, dt \leq C \|f_2 - f_1\|_F$$

for all $f_2, f_1 \in V$, as desired. □

Proposition B.1.4. *Let $P : (0 \in U \subseteq F) \to G$ be a C^2 map such that $P(0) = 0$ and $DP(0 \mid f) = 0$ for all $f \in F$. For any seminorm $\|\ldots\|_G$ of G there is a seminorm $\|\ldots\|_F$ of F and $\epsilon > 0$ such that whenever $\|f_2\|_F < \epsilon$ and $\|f_1\|_F < \epsilon$ we have:*

$$\|P(f_2) - P(f_1)\|_G \leq C \left(\|f_1\|_F + \|f_2\|_F\right) \|f_2 - f_1\|_F .$$

Proof. By Theorem 2.3.3 we can write:

$$P(f_2) - P(f_1) = \int_0^1 DP(f_t \mid f_2 - f_1)dt$$

$$DP(f_t \mid f_2 - f_1) = DP(f_t \mid f_2 - f_1) - DP(0 \mid f_2 - f_1) = \int_0^1 D^2P(sf_t \mid f_2 - f_1, f_t)ds,$$

where $f_t = tf_2 + (1-t)f_1$. Now we use the estimate of Lemma B.1.2, where as in the proof of Proposition B.1.3 we can take the neighborhood V to be of the form $V = \{\|f\| < \epsilon\}$ for some $\epsilon > 0$. Note that by convexity of V, $sf_t \in V$ for all $s, t \in [0, 1]$. Combining the expressions above with the estimate we obtain:

$$\begin{aligned}\|P(f_2) - P(f_1)\|_G &\leq \int_0^1 \int_0^1 \left\|D^2P(sf_t \mid f_2 - f_1, f_t)\right\|_G dtds \\ &\leq C \int_0^1 \int_0^1 \|f_t\|_F \|f_2 - f_1\|_F \, dtds \leq C \left(\|f_2\|_F + \|f_1\|_F\right) \|f_2 - f_1\|_F .\end{aligned}$$

□

B.2 The inverse function theorem for Banach spaces

We will use the estimates in the previous section to show the Banach inverse function theorem. First, we recall the following standard definition and result:

Definition B.2.1. *Let (X, d) be a metric space. A map $f : X \to X$ that satisfies:*

$$d(f(x_0), f(x_1)) \leq \rho d(x_0, x_1)$$

for some $\rho < 1$ is called a ***contraction****.*

Theorem B.2.2 (Fixed point theorem)**.** *Let (X, d) be a complete metric space and $f : X \to X$ a contraction. Then there is a unique point $p \in X$ such that $f(p) = p$.*

We will now consider a parametric version of this result.

Proposition B.2.3. *Let X, Y metric spaces with X complete and $P : X \times Y \to X$ such that:*

$$d(P(x_0, y), P(x_1, y)) < \rho d(x_0, x_1)$$

for some $\rho < 1$. Then for every y, there is a unique $x \in X$ such that $P(x, y) = x$ and the assignment $y \to x$ is continuous.

Proof. We write $P(x, y) = P_y(x)$, then $P_y : X \to X$ is a contraction for each y. Then by the fixed point theorem, there is a unique fixed point of P_y, showing the first part of the result. Denote the fixed point of P_y with $Q(y)$ and the k-times successive application of P_y with $Q_k(y) = (P_y)(x_0)$. Note that each Q_k is continuous in y. We claim that locally it converges uniformly to $Q(y)$, which makes the latter continuous as well. To see this, note observe that by the contraction property:

$$d(Q_{k+1}(y), Q_k(y)) \leq \rho^n = d(Q_1(y), Q_0(y)) = d(P_y(x_0), x_0)$$

From this, we obtain the following expression:

$$d(Q_{k+l}(y), Q_k(y)) = \sum_{i=0}^{l-1} d(Q_{k+i}(y), Q_{k+i+1}(y)) \leq d(Q_1(y), x_0) \sum_{i=0}^{l-1} \rho^{k+i} \leq d(Q_1(y), x_0) \frac{\rho^k}{1-\rho}.$$

Taking the limit $l \to \infty$ yields $d(Q(y), Q_k(y)) \leq \frac{\rho^k}{1-\rho} d(Q_1(y), x_0)$. Now fix y_0 and let $K = d(Q_1(y_0), x_0)$. By continuity of Q_1 we can define a neighborhood $V = \{y \mid d(y_0, y) < \delta\}$ such that $d(Q_1(y), Q_1(y_0)) \leq 1$ for $y \in V$. The for any $y \in V$ we have:

$$d(Q(y), Q_k(y)) \leq d(Q_1(y), x_0) \frac{\rho^k}{1-\rho} \leq (K+1) \frac{\rho^k}{1-\rho},$$

which indeed implies uniform convergence on V. □

Theorem B.2.4 (Inverse function theorem)**.** *Let $U \subseteq F$ and $V \subseteq G$ be neighborhoods in Banach spaces and $F : U \to V$ a C^1 map. Suppose that for $f_0 \in U$ the derivative $DP(f_0) : F \to G$ is invertible, then there are neighborhoods $f_0 \in U_0 \subseteq U$ and $P(f_0) \in V_0 \subseteq V$ such that:*

$$P : U_0 \to V_0$$

has a continuous inverse.

Proof. Since translations are isomorphisms, we may assume $f_0 = 0$ and $P(f_0) = 0$. Identifying F and G by the isomorpishm $DP(0)$, we also assume $DP(0) = \mathrm{Id}$ Define $Q(f) = P(f) - f$ such that $Q(0) = 0$ and $DQ(0) = 0$. Then we may apply Proposition B.1.4 to obtain an $\epsilon > 0$ such that for $\|f_1\|, \|f_2\| < \epsilon$ we have:

$$\|Q(f_1) - Q(f_2)\| \leq C(\|f_1\| + \|f_2\|) \|f_2 - f_1\|.$$

In particular, we have $\|Q(f)\| < C \|f\|^2$ for $\|f\| < \epsilon$. The strategy is now to define a "family of contractions" as in Proposition B.2.3 and use that result show continuity of the local inverse. We define:

$$R(f, g) = f - P(f) + g.$$

Note that $R(f, g) = f$ is equivalent to $P(f) = g$. We have the following estimates for R:

$$\begin{aligned} \|R(f1, g) - R(f_2, g)\| = \|Q(f_1) - Q(f_2)\| &\leq C(\|f_1\| + \|f_2\|) \|f_2 - f_1\| < 2C\epsilon \|f_2 - f_1\| \\ \|R(f, g)\| &\leq C \|f\|^2 + \|g\|. \end{aligned}$$

Choose $\epsilon < 1/2C$ (such that we also have $C\epsilon^2 \leq \frac{1}{2}\epsilon$) and let $\delta = \frac{1}{4}\epsilon$. Then for the following sets:

$$U_0 = \{\|f\| \leq \epsilon\}, \quad V_0 = \{\|g\| \leq \delta\},$$

we see that R maps $U_0 \times V_0$ into U_0 and is a contraction. By Proposition B.2.3, we see that for any $g \in V_0$ there is a unique $Q(g) \in U_0$ such that $PQ(g) = g$ and that $g \mapsto Q(g)$ is continuous. The uniqueness of $Q(g)$ implies that Q is also a left inverse of P. □

B.3 ODEs in Fréchet spaces

In general, ODEs in Fréchet space need not have (unique) solutions. In this section we will prove that for a special class of ODEs, namely those that can be factored through a Banach space, it *is* possible to find unique solutions.

Definition B.3.1. *Let $P : U \subseteq F \to F$ be a map defined on an open in Fréchet space. We say P is a **Banach map** if there is an open V in a Banach space B and maps $R : U \to B$ and $Q : V \subseteq B \to F$ such that the following diagram commutes:*

$$\begin{array}{ccc} U \subseteq F & \xrightarrow{\quad P \quad} & F \\ & {\scriptstyle R}\searrow \qquad \nearrow{\scriptstyle Q} & \\ & V \subseteq B & \end{array} .$$

Proposition B.3.2. *Let $P : U \subseteq F \to F$ be a C^1 Banach map. Then for any $f_0 \in U$ there is an ϵ and $f : [0, \epsilon) \to U$ such that:*

$$\begin{cases} f'(t) & = P(f(t)) \\ f(0) & = f_0 \end{cases},$$

with the solution f depending continuously on f_0.

Proof. Given f_0, we may replace P by $P - f_0$, since the constant vanishes upon taking the time derivative.

The idea of the proof is to define a parametric family of contractions and invoke Proposition B.2.3. Let $\|\ldots\|_B$ be the norm of the Banach space B. By Proposition B.1.3 there is a seminorm $\|\ldots\|_F$ and a δ_R such that for $\|f_1\|_F, \|f_2\|_F \leq \delta_R$:

$$\|R(f_2) - R(f_1)\|_B \leq C \|f_2 - f_1\|_F, \tag{B.1}$$

with $\|\ldots\|_B$ the unique (up to equivalence) norm on B. By Proposition B.1.3 again we see that there is a δ_Q such that for $\|g_1\|_B, \|g_2\|_B \leq \delta_Q$ we have:

$$\|Q(g_2) - Q(g_1)\|_F \leq C \|g_2 - g_1\|_B. \tag{B.2}$$

In particular, choosing $\delta_Q \leq 1$ we have the bound $\|Q(g)_F\| \leq C$ for g such that $\|g\|_B \leq \delta_Q$.

Now we define the following two spaces:

$$\begin{aligned} Y_\gamma &= \{f_0 \mid \|f_0\|_F \leq \gamma\} \\ X_\epsilon &= \left\{ g : [0, \epsilon] \to B \,\middle|\, \sup_{0 \leq t \leq \epsilon} \|g(t)\|_B \leq \delta_Q \right\} \subseteq \mathcal{C}([0, \epsilon], B), \end{aligned}$$

and we denote $\|\ldots\|'_B$ for the norm on $\mathcal{C}[0, \epsilon], B)$ (as in Definition 2.1.8). We claim that for appropriate values of $\gamma, \epsilon, \delta_Q$ and δ_R the map $M : X_\epsilon \times Y_\gamma \to X_\epsilon$ given by:

$$M(g, f_0)(t) = R\left(f_0 + \int_0^t Q(g(\tau))d\tau\right)$$

is well-defined and a contraction given any $f_0 \in Y_\gamma$.

To prove the claim, first observe that for $g \in X_\epsilon$ and $0 \leq t \leq \epsilon$

$$\left\| f_0 + \int_0^t Q(g(\tau))d\tau \right\|_F \leq \|f_0\|_F + \epsilon \sup_{0 \leq \tau \leq \epsilon} \|Q(g(\tau)\|_F \leq \gamma + \epsilon C.$$

By choosing γ and ϵ small, we can ensure that $\left\| f_0 + \int_0^t Q(g(\tau))d\tau \right\|_F \leq \delta_R$ and therefore by Equation (B.1):

$$\|M(g, f_0)(t)\|'_B = \left\| R\left(f_0 + \int_0^t Q(g(\tau))d\tau\right) \right\|_B \leq C\delta_R.$$

Then choose δ_R small enough such that $C\delta_R \leq \delta_Q$, which implies that the image of M lies in X_ϵ, making the map well-defined [1].

[1]To be precise: first choose $\delta_R \leq \delta_Q / C$, with both small enough to satisfy Equations (B.1) and (B.2), then choose ϵ, γ small compared to δ_R.

Now for any $f_0 \in Y_\gamma$ and $g_1, g_2 \in X_\epsilon$, note that:

$$\|M(g_2, f_0)(t) - M(g_1, f_0)(t)\|_B = \left\| R(\int_0^t Q(g_2(t)) - Q(g_1(t))) \right\|_B \leq \epsilon C \|g_2(t) - g_1(t)\|_B ,$$

by Equations (B.1) and (B.2). By choosing ϵ smaller if necessary, M is a contraction.

It follows by Proposition B.2.3 that the map $\tilde{S} : f_0 \to g$ assigning to each f_0 the unique fixed point g of $M(\dots, f_0) : X_\epsilon \to X_\epsilon$ is continuous. Furthermore, by Lemma 2.2.7 the mapping:

$$T : g \in X_\epsilon \mapsto \left(f : t \mapsto f_0 + \int_0^t Q(g(\tau)) d\tau \right) \in \mathcal{C}([0, \epsilon], F)$$

is continuous. Then $S = T \circ \tilde{S}$ is continuous. Now for g such that $M(f_0, g) = g$ and $f = Tg$ we see that $R(f(t)) = g(t)$ hence $P(f(t)) = (Q \circ R)(f(t)) = Q(g(t))$, so:

$$f(t) = f_0 + \int_0^t P(f(\tau)) d\tau,$$

showing that f satisfies the ODE. □

The proposition above can be strengthened as follows: if P is smooth, then so is the solution and the dependence on the "starting value" f_0:

Theorem B.3.3. *Let $P : U \subseteq F \to F$ be a C^1 Banach map. Then for any $f_0 \in U$ there is an ϵ and $f : [0, \epsilon) \to U$ such that:*

$$\begin{cases} f'(t) &= P(f(t)) \\ f(0) &= f_0 \end{cases}, \tag{B.3}$$

with the solution f depending smoothly on f_0.

Proof. We will use Lemma 2.3.5 to show that that the dependance on f_0 is differentiable.

By smoothness of P and Lemma 2.3.5, there is an L such that:

$$P(g) - P(f) = L(f, g)(g - f);$$
$$DP(f)h = L(f, f)h,$$

for $f, g \in U$ and $h \in F$. Since $P = Q \circ R$, we have $DP(f)h = DQ(R(f))DR(f)h$, with $DR : U \times F \to B$, so DP is a Banach map. The map L is given by:

$$L(f, g)h = \int_0^1 DP(f + t(g - f))(g - f) dt,$$

and is therefore also a Banach map. This means that by Proposition B.3.2 the following ODE:

$$\begin{pmatrix} f'(t) \\ g'(t) \\ h'(t) \end{pmatrix} = \begin{pmatrix} P(f(t)) \\ P(g(t)) \\ L(f(t), g(t))h(t) \end{pmatrix} \tag{B.4}$$

has a solution on $[0, \epsilon)$ given any starting values f_0, g_0 and h_0. Furthermore, there is a continuous map:

$$S : \begin{cases} U \times U \times F & \to \mathcal{C}([0, \epsilon], F)^3 \\ (f_0, g_0, h_0) & \mapsto (f, g, h) \end{cases}$$

assigning the solutions to the starting values. Since P and therefore L are smooth, Equation (B.4) shows that the solutions f, g and h are smooth.

Note that both the first and second component of S simply give the solution to $f'(t) = P(f(t))$ given any starting value. Write $M = S_1 = S_2$, such that $f = M(f_0)$ and $g = M(g_0)$. Note also that the third component of S is linear in h_0, we will denote it by $h = J(f_0, g_0)h_0$.

We claim that M and J satisfy the requirements for Lemma 2.3.5. First, M and J are both smooth. Second, compute:

$$\frac{d}{dt}\left(g(t) - f(t)\right) = P(g(t)) - P(f(t)) = L(f, g)(g - f),$$

so $(f, g, g - f)$ solves Equation (B.4) for starting values $(f_0, g_0, g_0 - f_0)$. Put differently:

$$M(g_0) - M(f_0) = J(f_0, g_0)(g_0 - f_0),$$

and now by Lemma 2.3.5 we see that M is C^1.

We have shown that for any ODE $f'(t) = P(f(t))$ with P a Banach map, the map M that assigns solutions to starting values is C^1. Furthermore, by taking $g = f$ the above shows that the tangent map TM equals the map assigning solutions to the ODE:

$$\begin{pmatrix} f'(t) \\ h'(t) \end{pmatrix} = \begin{pmatrix} P(f(t)) \\ DP(f(t))h(t) \end{pmatrix} = TP(f(t), h(t)).$$

Since TP is again a Banach map, TM is C^1 and therefore M is C^2. Repeating this process inductively shows that M is smooth.

□

References

[Whi37] Hassler Whitney. "On regular closed curves in the plane". In: *Compositio Mathematica* 4 (1937), pp. 276–284.

[Nas56] John Nash. "The imbedding problem for Riemannian manifolds". In: *Annals of mathematics* 63.1 (1956), pp. 20–63.

[Hir59] Morris W Hirsch. "Immersions of manifolds". In: *Transactions of the American Mathematical Society* 93.2 (1959), pp. 242–276.

[Sma59a] Stephen Smale. "A classification of immersions of the two-sphere". In: *Transactions of the American Mathematical Society* 90.2 (1959), pp. 281–290.

[Sma59b] Stephen Smale. "The classification of immersions of spheres in Euclidean spaces". In: *Annals of mathematics* 69.2 (1959), pp. 327–344.

[Mos61] Jurgen Moser. "A new technique for the construction of solutions of nonlinear differential equations". In: *Proceedings of the National Academy of Sciences* 47.11 (1961), pp. 1824–1831.

[Rud+64] Walter Rudin et al. *Principles of mathematical analysis*. Vol. 3. McGraw-hill New York, 1964.

[Phi67] Anthony Phillips. "Submersions of open manifolds". In: *Topology* 6.2 (1967), pp. 171–206.

[Gro69] Mikhail Leonidovich Gromov. "Stable mappings of foliations into manifolds". In: *Mathematics of the USSR-Izvestiya* 3.4 (1969), p. 671.

[Gro71] Mikhael Gromov. "A topological technique for the construction of solutions of differential equations and inequalities". In: *Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 2*. Gauthier-Villars Éditeur, Paris, 1971, pp. 221–225.

[GG74] M. Golubitsky and V. Guillemin. *Stable Mappings and Their Singularities*. Graduate texts in mathematics. Springer, 1974. ISBN: 9787506200448. URL: `https://books.google.nl/books?id=1RfgQgAACAAJ`.

[Thu74] William Thurston. "The theory of foliations of codimension greater than one". In: *Commentarii Mathematici Helvetici* 49 (1974), pp. 214–231.

[Thu76] William P Thurston. "Existence of codimension-one foliations". In: *Annals of Mathematics* 104.2 (1976), pp. 249–268.

[Ham82] Richard S Hamilton. "The inverse function theorem of Nash and Moser". In: *Bulletin of the American Mathematical Society* 7.1 (1982), pp. 65–222.

[Gro86] Mikhael Gromov. *Partial Differential Relations*. Vol. 9. Springer Science & Business Media, 1986.

[Con89] John B. Conway. *A Course in Functional Analysis*. Second Edition. 1989.

[Gro96] Mikhael Gromov. "Carnot-Carathéodory spaces seen from within". In: (1996), pp. 79–323.

[KM97] Andreas Kriegl and Peter W Michor. *The convenient setting of global analysis*. Vol. 53. American Mathematical Soc., 1997.

[EM01] Yakov M Eliashberg and Nikolai M Mishachev. "Holonomic approximation and Gromov's h-principle". In: *arXiv preprint math/0101196* (2001).

[Sil01] Ana Cannas da Silva. *Lectures on symplectic geometry*. Vol. 1764. Springer, 2001.

[EM02] Yakov Eliashberg and Nikolai M Mishachev. *Introduction to the h-principle*. 48. American Mathematical Soc., 2002.

[Hir12] M.W. Hirsch. *Differential Topology*. Graduate Texts in Mathematics. Springer New York, 2012. ISBN: 9781468494495. URL: https://books.google.nl/books?id=emTmBwAAQBAJ.

[Lee12] John M. Lee. *Introduction to Smooth Manifolds*. Second Edition. 2012.

[Eli15] Yakov Eliashberg. "Recent advances in symplectic flexibility". In: *Bulletin of the American Mathematical Society* 52.1 (2015), pp. 1–26. URL: https://www.ams.org/bull/2015-52-01/S0273-0979-2014-01470-3/S0273-0979-2014-01470-3.pdf.

[Pan16] Pierre Pansu. "Submanifolds and differential forms on Carnot manifolds, after M. Gromov and M. Rumin". In: *arXiv preprint arXiv:1604.06333* (2016).

[CB17] Marius Crainic and Erik van der Ban. *Analysis on Manifolds*. 2017. URL: https://www.ime.usp.br/~toscano/disc/2021/CrainicAS.pdf.

[CEM24] Kai Cieliebak, Y Eliashberg, and N Mishachev. *Introduction to the h-Principle*. Vol. 239. American Mathematical Society, 2024.

[Sag] Steffen Sagave. *Lecture notes for the master course Algebraic Topology*. URL: https://www.math.ru.nl/~sagave/teaching/at-lecturenotes-2020.pdf.