Nonlinear Functional Analysis

K. Cieliebak

Incomplete draft, 6 July 2018

Contents

Introduction

These are notes for a course on Nonlinear Functional Analysis given at University of Augsburg in Fall 2017.

Nonlinear Functional Analysis concerns the analysis of nonlinear maps between infinite dimensional spaces. So it relates to other subjects by the following table:

	Linear	Nonlinear
$\dim < \infty$	Linear Algebra	Analysis
$\dim = \infty$	Functional Analysis	Nonlinear Functional Analysis

A typical question concerns the structure of the zero set $f^{-1}(0)$ of a map $f : E \to F$ between vector spaces. If f is linear then $f^{-1}(0)$ is a linear subspace. For nonlinear f the zero set has in general no nice structure:

Problem 0.1. Show that every closed subset $A \subset \mathbb{R}^n$ arises as the zero set of a smooth (i.e. infinitely often differentiable) function $f : \mathbb{R}^n \to \mathbb{R}$.

If E, F are finite dimensional, say $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$, then a satisfactory answer is provided by the implicit function theorem: If $f : \mathbb{R}^n \to \mathbb{R}^m$ is smooth and 0 is a regular value of f (i.e. $Df(x)$ is surjective for all $x \in f^{-1}(0)$), then $f^{-1}(0) \subset \mathbb{R}^n$ is a smooth submanifold of dimension $n - m$.

In infinite dimensions the situation is more delicate and the validity of the implicit function theorem depends on the type of the spaces

E, F. For example, it holds in Hilbert and Banach spaces, but not in Fréchet spaces. Compensations for this failure of the implicit function theorem lie at the heart of many deep results in geometry and analysis and will be one recurring theme in this course.

To be continued.

Part I

Banach manifolds

Chapter 1

Banach spaces and their linear maps

In this chapter we recall – mostly without proof – some necessary background about Banach spaces and their linear maps. Most of the material in this chapter can be found in [19].

1.1 Metric spaces

Here we collect 3 basic results about metric spaces: the Banach fixed point theorem, the Baire category theorem, and the Arzela-Ascoli theorem.

Banach fixed point theorem. Recall that a metric space (X, d) is called *complete* if every Cauchy sequence converges to a limit in X.

Theorem 1.1 (Banach fixed point theorem)**.** *Let (X, d) be a complete metric space and $f : X \to X$ a map which is* contracting, *i.e., there exists a constant $\lambda < 1$ such that*

$$d\big(f(x), f(y)\big) \leq \lambda d(x, y) \quad \textit{for all } x, y \in X.$$

Then f has a unique fixed point $x_* \in X$ such that $f(x_*) = x_*$. Moreover, for any initial point $x_0 \in X$ the sequence $x_n := f^n(x_0)$ converges to x_*.

Problem 1.1. Prove the Banach fixed point theorem and show that it fails under the weaker hypothesis $d\big(f(x), f(y)\big) < d(x, y)$ for all $x, y \in X$.

Baire's theorem. A subset A of a topological space is called *nowhere dense* if its closure $\overline{A}$ has empty interior.

Theorem 1.2 (Baire). *Let (X, d) be a complete metric space and A_n, $n \in \mathbb{N}$ be closed nowhere dense subsets of X. Then $\bigcup_{n\in\mathbb{N}} A_n$ has empty interior.*

Notation. A subset of X is called *meagre* (or of *first category*) if it is contained in a countable union of nowhere dense closed sets. We will call the complement of a meagre set a *Baire set.* By definition a countable union of meagre sets is meagre and a countable intersection of Baire sets is a Baire set. Baire's theorem says that in a complete metric space a meagre set has empty interior, or equivalenly, a Baire set is dense.

Problem 1.2. Which of the following sets are Baire sets / meagre?
(a) $\mathbb{Q} \subset \mathbb{R}$;
(b) $\mathbb{R} \setminus \mathbb{Q} \subset \mathbb{R}$;
(c) the set of transcendental numbers in $\mathbb{R}$;
(d) the set of nowhere differentiable functions in the space of continuous functions $[0, 1] \to \mathbb{R}$ with the maximum norm.
Hint for (d): Consider for each $n, k \in \mathbb{N}$ the set

$$\begin{aligned} A_{n,k} := \{f \in C^0([0,1],\mathbb{R}) \mid \forall x \in [0,1]\, \exists y \in [0,1] \textit{ such that} \\ 0 < |x-y| < 1/k \textit{ and } |f(x) - f(y)| \geq n|x-y|\}. \end{aligned}$$

Show that the interior of $A_{n,k}$ is dense for all n, k (by adding small zigzag functions) and apply Baire's theorem.

Arzela-Ascoli theorem. Recall that a topological space is called *compact* if every open cover has a finite subcover. A subset of a topological space is called *relatively compact* if its closure is compact.

Let now X be a compact topological space and (Y, d) a metric space. Then the space

$$C(X, Y) := \{\text{continuous functions } f : X \to Y\}$$

becomes a metric space with the metric

$$d(f, g) := \max_{x \in X} d\big(f(x), f(y)\big).$$

If Y is complete then so is $C(X, Y)$.

Theorem 1.3 (Arzela-Ascoli)**.** *Let X be a compact topological space and (Y, d) a (not necessarily complete) metric space. Then a subset $\mathcal{F} \subset C(X, Y)$ is relatively compact if and only if it satisfies the following two conditions:*

(i) $\mathcal{F}$ is equicontinuous, *i.e., for every $x \in X$ and $\varepsilon > 0$ there exists an open neighbourhood $U(x) \subset X$ such that $d\big(f(x), f(y)\big) < \varepsilon$ for all $y \in U(x)$;*

(ii) for each $x \in X$ the set $\{f(x) \mid f \in \mathcal{F}\}$ is relatively compact in Y.

For $Y = \mathbb{R}^m$ property (ii) is sometimes called "uniform boundedness". The proof uses the following characterization of compactness for metric spaces.

Lemma 1.4. *A metric space (X, d) is compact if and only it is complete and* totally bounded *(i.e., for every $\varepsilon > 0$ the space X can be covered by finitely many open ε-balls).*

Proof of Theorem 1.3. We prove the "if" and leave the "only if" (which is simpler) as an exercise. In view of Lemma 1.4 it suffices to show that the closure $\overline{\mathcal{F}}$ is complete and totally bounded.

For completeness, let (f_n) be a Cauchy sequence in $\mathcal{F}$. Then for each $x \in X$, $f_n(x)$ is a Cauchy sequence in $\{f(x) \mid f \in \mathcal{F}\}$, which has a limit in Y by assumption (ii). So $f(x) := \lim_{n\to\infty} f_n(x)$ defines a map $f : X \to Y$. Now one shows that f is continuous and $f = \lim_{n\to\infty} f_n$.

For total boundedness, note that for each cover of $\mathcal{F}$ by open ε-balls the 2ε-balls cover $\overline{\mathcal{F}}$, so it suffices to show that $\mathcal{F}$ is totally bounded. Let $\varepsilon > 0$ be given.

The open sets $U(x)$, $x \in X$, from assumption (i) cover X. By compactness of X, finitely many of these sets $U(x_1), \dots, U(x_k)$ cover X. By assumption (ii), the set $\{f(x_i) \mid f \in \mathcal{F},\ 1 \le i \le k\}$ is covered by finitely many open ε-balls $B_\varepsilon(y_j)$, $j = 1, \dots, \ell$. For each map $\sigma : \{1, \dots, k\} \to \{1, \dots, \ell\}$ define

$$\mathcal{F}_\sigma := \{f \in \mathcal{F} \mid f(x_i) \in B_\varepsilon(y_{\sigma(i)}) \text{ for all } i = 1, \dots, k\} \subset \mathcal{F}.$$

By construction, $\bigcup_\sigma \mathcal{F}_\sigma = \mathcal{F}$. For $f, g \in \mathcal{F}_\sigma$ and each $x \in X$ we find an i with $x \in U(x_i)$ and estimate

$$\begin{aligned} d\big(f(x), g(x)\big) &\le d\big(f(x), f(x_i)\big) + d\big(f(x_i), y_{\sigma(i)}\big) \\ &\qquad + d\big(y_{\sigma(i)}, g(x_i)\big) + d\big(g(x_i), g(x)\big) \\ &< \varepsilon + \varepsilon + \varepsilon + \varepsilon = 4\varepsilon. \end{aligned}$$

Thus each $\mathcal{F}_\sigma$ has diameter $< 4\varepsilon$ and is therefore contained in an open 4ε-ball. Since ε was arbitrary, this proves total boundedness of $\mathcal{F}$. □

Example 1.5. In applications property (i) usually comes from uniform bounds on derivatives. For example, the set $\mathcal{F}$ of differentiable functions f with $|f'(x)| \le C$ for all x and $|f(0)| \le C$ is relatively compact in $C([0,1], \mathbb{R})$.

1.2 Banach and Hilbert spaces

Unless otherwise stated all vector spaces will be over $\mathbb{R}$. All definitions and results carry over to $\mathbb{C}$ with minor adjustments.

Definition 1.1. A *Banach space* $(E, \| \ \|)$ is a complete normed vector space. A *Hilbert space* $(E, \langle\ ,\ \rangle)$ is a vector space with a scalar product which is complete with respect to the induced norm $\|x\| = \langle x, x\rangle^{1/2}$.

Problem 1.3. Show: A norm $\| \ \|$ is induced by a scalar product if and only if it satisfies for all x, y the *parallelogram law*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Examples. Here are the most important examples of Banach spaces.

(1) $\mathbb{R}^n$ becomes a Banach space with each of the norms

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} & : 1 \le p < \infty, \\ \max_{1\le i\le n} |x_i| & : p = \infty. \end{cases}$$

For $p = 2$ the norm is induced by the scalar product $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ and we will write the corresponding Euclidean norm simply as $|x|$.

(2) For $1 \le p \le \infty$ the space of sequences

$$l^p := \{x = (x_i)_{i\in\mathbb{N}} \mid x_i \in \mathbb{R}, \|x\|_p < \infty\}$$

is a Banach space with the norm

$$\|x\|_p := \begin{cases} \left(\sum_{i\in\mathbb{N}} |x_i|^p\right)^{1/p} & : 1 \le p < \infty, \\ \max_{i\in\mathbb{N}} |x_i| & : p = \infty. \end{cases}$$

(3) For a measure space (Ω, μ) denote by $M(\Omega, \mathbb{R}^m)$ the vector space of measurable functions $\Omega \to \mathbb{R}^m$ modulo the subspace of functions which vanish outside a set of measure zero. Then for $1 \le p \le \infty$ the space

$$L^p(\Omega, \mathbb{R}^m) := \{f \in M(\Omega, \mathbb{R}^m) \mid \|f\|_{L^p} < \infty\}$$

is a Banach space with the norm

$$\|x\|_{L^p} := \begin{cases} \left(\int_\Omega |f|^p d\mu\right)^{1/p} & : 1 \le p < \infty, \\ \inf_{\mu(Z)=0} \sup_{\Omega \setminus Z} |f| & : p = \infty. \end{cases}$$

Note that (1) and (2) are special cases of this (with $m = 1$). For $p = 2$ this space is a Hilbert space with the scalar product

$$\langle f, g \rangle_{L^2} := \int_\Omega \langle f, g \rangle d\mu.$$

(4) The space $C(X, F)$ of continuous maps from a compact topological space X to a Banach space $(F, \| \|)$ becomes a Banach space with the norm

$$\|f\|_{\max} := \max_{x \in X} \|f(x)\|.$$

(5) Let $\Omega \subset \mathbb{R}^n$ be open. For each $k \in \mathbb{N}_0$ and $1 \le p < \infty$ the *Sobolev space*

$$\left(W^{k,p}(\Omega, \mathbb{R}^m), \| \|_{W^{k,p}}\right)$$

of functions $\Omega \to \mathbb{R}^m$ with k weak derivatives in L^p (see Section 1.6 below) is a Banach space. It is a Hilbert space for $p = 2$ and agrees with $L^p(\Omega, \mathbb{R}^m)$ for $k = 0$.

(6) Let $\Omega \subset \mathbb{R}^n$ be bounded open. For each $k \in \mathbb{N}_0$ and $0 \le \alpha \le 1$ the *Hölder space*

$$\left(C^{k,\alpha}(\overline{\Omega}, \mathbb{R}^m), \| \|_{C^{k,\alpha}}\right)$$

of functions $\overline{\Omega} \to \mathbb{R}^m$ with k derivatives in C^α (see Section 1.6 below) is a Banach space. This agrees with the space of C^k-functions for $\alpha = 0$ and with the space of Lipschitz continuous functions for $k = 0, \alpha = 1$.

By contrast, we will see later that the space $C^\infty([0,1],\mathbb{R})$ of smooth functions $[0,1] \to \mathbb{R}$ with the C^∞-topology is *not* a Banach space.

Notation. If the target is $\mathbb{R}$ we will sometimes write $L^p(\Omega)$ instead of $L^p(\Omega,\mathbb{R})$ etc.

Closed subspaces. Not every linear subspace F of a normed vector space E is closed. For example, the closure of the linear subspace $C^1([0,1],\mathbb{R}) \subset C^0([0,1],\mathbb{R})$ is all of $C^0([0,1],\mathbb{R})$. A subspace $F < E$ is closed if and only if

$$\|\xi\| := \inf_{x\in\xi} \|x\|$$

defines a norm on the quotient space E/F. In this case E/F is complete if E is.

One says that E is the *direct sum* of two subspaces F, G, and writes $E = F \oplus G$, if E is the algebraic direct sum of F and G (i.e., every $x \in E$ can be uniquely written as $x = y + z$ with $y \in F$ and $z \in G$) and F, G are closed. In this case G is called a *complement* of F. The existence of a complement is equivalent to splitting by a continuous linear map of the short exact sequence

$$0 \mapsto F \mapsto E \mapsto E/F \mapsto 0.$$

Problem 1.4. (a) Show that every finite dimensional subspace F of a normed vector space E is closed and has a complement. *Hint: Use the Hahn-Banach theorem 1.12 below.*
(b) Show that every finite codimensional closed subspace F of a normed vector space E has a complement.

(c) Find an example of a codimension 1 subspace of a Banach space which is not closed.
(d) Show that every closed subspace of a Hilbert space has a complement (namely its orthogonal complement).

Example 1.6. There exist closed subspaces of Banach spaces which have no complement. An example is the subspace of $L^1([0,1],\mathbb{C})$ consisting of those functions f whose Fourier coefficients $\widehat{f}(k) = \int_0^1 f(x)e^{-2\pi ikx}dx$ vanish for all $k<0$, see [42, Examples 5.19].

Separability. A topological space is called *separable* if it has a countable dense set.

Problem 1.5. (a) Prove that the following spaces are separable: $\mathbb{R}^n$, $C^{k,\alpha}(\overline{\Omega})$, ℓ^p, $L^p(\Omega)$, $W^{k,p}(\Omega)$ for $\Omega \subset \mathbb{R}^n$ open bounded and $1 \le p < \infty$.
(b) Prove that the spaces ℓ^∞ and $L^\infty(\Omega)$ for $\Omega \subset \mathbb{R}^n$ open bounded are not separable.

Bases. The notion of a basis in the sense of linear algebra is not very useful for infinite dimensional spaces. For example, one has

Problem 1.6. Prove: The vector space dimension of a Banach space is either finite or uncountably infinite. *Hint: Baire's theorem.*

Instead, the notion of a "basis" is modified by allowing infinite linear combinations as follows. Let E be an infinite dimensional Banach space. A sequence $e_1, e_2 \dots$ in E is called a *(Schauder) basis* of E if every $x \in E$ can be uniquely written as a convergent series $x = \sum_{k=1}^\infty x_k e_k$ with $x_k \in \mathbb{R}$. If E is a Hilbert space, then a basis (e_k) is called *orthonormal* (or a *Hilbert space basis*) if in addition $\langle e_k, e_\ell\rangle = \delta_{k\ell}$. Note that the existence of a basis implies that E is separable (finite linear combinations with rational coefficients are dense).

Every separable Hilbert space E has a basis, and every basis of

E can be turned into an orthonormal basis by the Gram-Schmidt procedure. Sending $x \in E$ to its coefficients (x_k) with respect to an orthonormal basis defines an isometry $E \cong \ell^2$. In particular, all separable Hilbert spaces are isometric to ℓ^2.

Not every separable Banach space has basis (this long-standing problem was only solved in 1973), but all the spaces in Problem 1.5(a) have bases.

Example 1.7. (a) The standard vectors e_k with 1 at the k-th position and 0 otherwise form a basis of ℓ^p for $1 \leq p < \infty$ (but not for $p = \infty$).
(b) The trigonometric polynomials $e_k(x) = e^{2\pi i k x}$ for $k \in \mathbb{Z}$ ordered as $\{0, 1, -1, 2, -2, \dots\}$ form a basis of $L^p([0,1], \mathbb{C})$ for all $1 \leq p < \infty$, but not for the space of periodic continuous functions $C^0(\mathbb{R}/\mathbb{Z}, \mathbb{C})$, see [3, Chapter 1].

Given a basis (e_k) for the Banach space E, define the linear maps $e_\ell^* : E \to \mathbb{R}$ sending $\sum_{k \in \mathbb{N}} x_k e_k$ to the coefficient x_ℓ, so each element of E is written as

$$x = \sum_{k \in \mathbb{N}} e_k^*(x) e_k.$$

It is a nontrivial fact that the maps e_k^* are continuous and thus represent elements of the dual space E^*, see [3, Chapter 1]. So we obtain continuous linear projections $P_n : E \to E$ sending $\sum_{k=1}^{\infty} x_k e_k$ to $\sum_{k=1}^{n} x_k e_k$. It is an immediate consequence of the Banach-Steinhaus Theorem 1.9 below that $\sup_n \|P_n\| < \infty$.

1.3 Continuous linear maps

In infinite dimensions, linear maps need not be continuous:

Problem 1.7. Let $E, F \neq \{0\}$ be normed vector spaces. Show:
(a) A linear map $A : E \to F$ is continuous if and only if there exists a constant C such that $\|Ax\| \leq C\|x\|$ for all $x \in E$.
(b) All linear maps $E \to F$ are continuous if and only if E is finite dimensional.

In view of the property in (a), continuous linear maps are also called *bounded linear maps.*[1] For normed spaces E, F the space

$$\mathcal{L}(E, F) := \{\text{continuous linear maps } A : E \to F\}$$

is again a normed vector space with the *operator norm*

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|.$$

If F is a Banach space then so is $\mathcal{L}(E, F)$. In particular, the *dual space*

$$E^* := \mathcal{L}(E, \mathbb{R})$$

of a normed space with the operator norm is a Banach space. Each $A \in \mathcal{L}(E, F)$ has an *adjoint operator* $A^* \in \mathcal{L}(F^*, E^*)$ defined by $(A^* f)(x) := f(Ax)$ for $f \in F^*$ and $x \in E$.

Problem 1.8. Show that the derivative $\frac{d}{dt}$ defines a continuous linear map $C^1(\mathbb{R}/\mathbb{Z}, \mathbb{R}) \to C^0(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ and determine its kernel and image.

A bijective continuous linear map $A : E \to F$ between normed vector spaces is called an *isomorphism.* If it also preserves the norms it is called an *isometry.* We write $E \cong F$ if they are isomorphic.

Problem 1.9. Let $1 \leq p, q \leq \infty$ be related by $1/p + 1/q = 1$. Prove: (a) $(\ell^p)^*$ is isometric to ℓ^q for $1 \leq p < \infty$.
(b) $(\ell^\infty)^*$ is *not* isomorphic to ℓ^1.

[1] Though of course the image of a nontrivial linear map is never bounded!

Remark 1.8. The space ℓ^p and ℓ^q are non-isomorphic for $p \neq q$, see [3, Corollary 2.1.6].

The following result is also called the **uniform boundedness principle**.

Theorem 1.9 (Banach-Steinhaus). *Let E be a Banach space and F a normed vector space. Let $\mathcal{A} \subset \mathcal{L}(E, F)$ be such that for all $x \in E$ there exists a constant C_x such that $\|Ax\| \leq C_x$ for all $A \in \mathcal{A}$. Then there exists a constant C such that $\|A\| \leq C$ for all $A \in \mathcal{A}$.*

Problem 1.10. Prove the Banach-Steinhaus theorem. *Hint: Apply Baire's theorem to conclude that one of the sets $E_n := \{x \in E \bigm| \|Ax\| \leq n \text{ for all } A \in \mathcal{A}\}$, $n \in \mathbb{N}$ has nonempty interior.*

Open mapping theorem. The following fundamental result will be proved later in a more general context.

Theorem 1.10 (Open mapping theorem). *Let $A : E \to F$ be a continuous linear map between Banach spaces. If A is surjective then it is open (i.e., images of open sets are open). In particular, if A is bijective then its inverse A^{-1} is also continuous.*

Corollary 1.11 (Closed graph theorem). *A linear map $A : E \to F$ between Banach spaces is continuous if and only if its graph $gr(A) = \{(x, Ax) \mid x \in E\}$ is closed in $E \oplus F$.*

Proof. The "only if" is easy: if $x_n \to x$ and $Ax_n \to y$, then continuity of A implies $Ax = y$ and thus $(x, y) \in gr(A)$.
For the "if", let $p_E : gr(A) \to E$ and $p_F : gr(A) \to F$ be the projections onto the two factors. Since p_E is continuous and bijective, it has a continuous inverse p_E^{-1} by the open mapping theorem, hence $A = p_F \circ p_E^{-1}$ is continuous. □

Problem 1.11. Find a surjective continuous map $f : [0, 1] \to [0, 1]$ which is not open.

Hahn-Banach theorem. The following important result allows the construction of lots of linear functionals.

Theorem 1.12 (Hahn-Banach theorem)**.** *Let F be a linear subspace of a normed vector space E. Then every $f \in F^*$ can be extended (non-uniquely) to an $e \in E^*$ with $\|e\| = \|f\|$.*

1.4 Compact and Fredholm operators

Compact operators. A linear map $K : E \to F$ is called *compact* if it maps bounded sets to relatively compact sets. This is equivalent to $K\big(B_1(0)\big)$ being relatively compact, and to the image (Kx_n) of each bounded sequence (x_n) having a convergent subsequence, and it implies continuity of K. Let

$$\mathcal{K}(E, F) := \{\text{compact linear maps } E \to F\}.$$

Proposition 1.13 (Properties of compact operators)**.**
Let E, F, G be Banach spaces. Then:
(a) $\mathcal{K}(E, F) \subset \mathcal{L}(E, F)$ is closed.
(b) $\mathcal{K}(F, G) \circ \mathcal{L}(E, F) \subset \mathcal{K}(E, G)$ and $\mathcal{L}(F, G) \circ \mathcal{K}(E, F) \subset \mathcal{K}(E, G)$.
(c) $K \in \mathcal{L}(E, F)$ is compact if and only if $K^* \in \mathcal{L}(F^*, E^*)$ is compact.
(d) If $K \in \mathcal{L}(E, F)$ is the limit in the operator norm of operators K_n of finite rank *(i.e., with finite dimensional image) then K is compact; the converse holds if F is a Hilbert space (but not in general if F is a Banach space).*

Proof of (d), see [13]. The first implication follows from (a) because finite rank operators are clearly compact. For the converse, let $K : E \to F$ be compact. Since the closure of $K\big(B_1(0)\big)$ is compact, hence separable, it follows that $G := \overline{\operatorname{im} K} \subset F$ is a separable closed subspace. If F is a Hilbert space, then G is a separable Hilbert space and therefore has a basis $e_1, e_2, \dots$ (this step fails if F is not Hilbert). Let $P_n : G \to \operatorname{span}\{e_1, \dots, e_n\}$ be the continuous linear map sending $\sum_{k=1}^{\infty} y_k e_k$ to $\sum_{k=1}^{n} y_k e_k$. Then $K_n := P_n K$ has finite rank for each $n \in \mathbb{N}$.

To show that $K_n \to K$ first note that $K_n x = P_n K x \to Kx$ for each x by definition of a basis. Now let $\varepsilon > 0$ be given. Since $K\big(\overline{B_1(0)}\big)$ is relatively compact, it is covered by finitely many open balls $B_\varepsilon(Kx_j)$, $j = 1, \dots, m$. Since $K_n x_j \to Kx_j$ for each j, there exists $N \in \mathbb{N}$ such that $\|K_n x_j - Kx_j\| < \varepsilon$ for all j and $n \geq N$. For $n \geq N$ and each $x \in \overline{B_1(0)}$ we find a j with $Kx \in B_\varepsilon(Kx_j)$ and obtain, using the uniform bound $\|P_n\| \leq C$,

$$\begin{aligned}\|Kx - K_n x\| &\leq \|Kx - Kx_j\| + \|Kx_j - K_n x_j\| + \|P_n(Kx_j - Kx)\| \\ &< \varepsilon + \varepsilon + C\varepsilon = (2 + C)\varepsilon.\end{aligned}$$

Since ε was arbitrary, this proves $K_n \to K$. □

Remark 1.14. For $E = F$, properties (a) and (b) in Proposition 1.13 say that the quotient space $\mathcal{L}(E, E)/\mathcal{K}(E, E)$ is a *Banach algebra*, i.e., an $\mathbb{R}$-algebra which is also a Banach space such that $\|xy\| \leq \|x\|\|y\|$ for all x, y.

Problem 1.12. Prove: The closed unit ball in a normed vector space E is compact if and only if E is finite dimensional. *Hint: If* $\dim E = \infty$ *construct a sequence* (e_n) *in* E *with* $\|e_n\| = 1$ *and* $\|e_m - e_n\| \geq 1$ *for all* $m \neq n$.

Fredholm operators. Consider the following properties for a continuous linear map $T : E \to F$ between Banach spaces:

(i) $\ker T$ is finite dimensional,

(ii) $\operatorname{im} T$ is closed,
(iii) $\operatorname{coker} T := F/\operatorname{im} T$ is finite dimensional.

The operator T is called

- *Fredholm* if it has properties (i), (ii) and (iii),
- *left-Fredholm* if it has properties (i) and (ii),
- *right-Fredholm* if it has properties (ii) and (iii), and
- *semi-Fredholm* if it has properties (ii) and ((i) or (iii)).

Problem 1.13. Prove: If $T \in \mathcal{L}(E, F)$ and $F/\operatorname{im} T$ is finite dimensional, then $\operatorname{im} T$ is closed. Hence condition (ii) above follows from (iii). *Hint: Complete T to a surjective operator $\widehat{T} : E \oplus \mathbb{R}^n \to F$ and apply the open mapping theorem.*

We denote the spaces of (left-/right-/semi-) Fredholm operators by

$$\mathcal{F}(E, F), \quad \mathcal{LF}(E, F), \quad \mathcal{RF}(E, F), \quad \mathcal{SF}(E, F),$$

so that

$$\begin{aligned} \mathcal{F}(E, F) &= \mathcal{LF}(E, F) \cap \mathcal{RF}(E, F), \\ \mathcal{SF}(E, F) &= \mathcal{LF}(E, F) \cup \mathcal{RF}(E, F). \end{aligned}$$

The *index* of a semi-Fredholm operator is

$$\operatorname{ind}(T) := \dim(\ker T) - \dim(\operatorname{coker} T) \in \mathbb{Z} \cup \{-\infty, +\infty\}.$$

This is well-defined because at most one of the two terms on the right hand side is infinite, and $\operatorname{ind}(T) \in \mathbb{Z}$ iff T is Fredholm. The main result about Fredholm operators is

Theorem 1.15. *Let E, F be Banach spaces. Then the sets $\mathcal{F}(E, F)$, $\mathcal{LF}(E, F)$, $\mathcal{RF}(E, F)$ and $\mathcal{SF}(E, F)$ are open in $\mathcal{L}(E, F)$ and the map $\operatorname{ind} : \mathcal{SF}(E, F) \to \mathbb{Z} \cup \{-\infty, +\infty\}$ is continuous (i.e., constant on connected components).*

Add proof.

The following criterion for the left-Fredholm property is very useful in applications to PDEs.

Proposition 1.16. *Let E, F, G be Banach spaces, $T \in \mathcal{L}(E, F)$ and $K \in \mathcal{K}(E, G)$ such that*

$$\|x\| \le C\|Tx\| + \|Kx\| \quad \text{for all } x \in E, \tag{1.1}$$

with some constant $C > 0$. Then T is left-Fredholm.

Proof. For property (i) note that $\|x\| \le \|Kx\|$ for all $x \in \ker T$. If (x_n) is a sequence in $\ker T$ with $\|x_n\| \le 1$, then compactness of K implies that a subsequence (Kx_{n_j}) is a Cauchy sequence, hence by the preceding inequality also (x_{n_j}), so (x_{n_j}) converges by completeness of E. This shows that the closed unit ball in $\ker T$ is compact, so $\ker T$ is finite dimensional by Problem 1.12.

For property (ii), after restricting T and K to a complement of $\ker T$ in E (which exists because $\dim \ker T < \infty$) we may assume that T is injective. Consider a sequence $x_n \in E$ such that $Tx_n \to y \in F$. Suppose that $\|x_n\|$ is unbounded, so after passing to a subsequence we may assume $\|x_n\| \to \infty$. Then by (1.1) the $x_n' := x_n/\|x_n\|$ satisfy

$$1 = \|x_n'\| \le C\|Tx_n'\| + \|Kx_n'\|$$

and $Tx_n' \to 0$. By compactness of K, after passing to a subsequence we may assume that (Kx_n') is a Cauchy sequence, hence by the displayed inequality so is (x_n'). It follows that $x_n' \to x'$ and $Tx' = \lim Tx_n' = 0$, thus $x' = 0$ by injectivity of T, contradicting $\|x_n'\| = 1$ for all n. Hence $\|x_n\|$ must be bounded. Then again, after passing to a subsequence we may assume that (Kx_n) is a Cauchy sequence, hence by (1.1) so is (x_n) (recall that (Tx_n) is a Cauchy sequence). Hence $x_n \to x \in E$ and $Tx_n \to Tx$, so $y = Tx \in \operatorname{im} T$. □

Corollary 1.17 (Riesz)**.** *Let $K : E \to E$ be a compact operator on a Banach space E. Then $T := \mathbb{1} - K$ is a Fredholm operator of index zero.*

Proof. Properties (i) and (ii) follow from Proposition 1.16 (with $C = 1$), and property (iii) follows from (i) applied to the compact operator K^* because $\operatorname{coker} K \cong \ker K^*$ (see the next problem). By Theorem 1.15 the index is constant on the path $T \mapsto \mathbb{1} - tK$ and therefore zero because $\operatorname{ind}(\mathbb{1}) = 0$. □

Problem 1.14. Let $T : E \to F$ be a continuous linear operator between Banach spaces with closed image $\operatorname{im} T$. Then there are canonical isomorphisms

$$\ker(T^*) \cong (\operatorname{coker} T)^*, \qquad \operatorname{coker}(T^*) \cong (\ker T)^*.$$

Invertibility modulo compact operators. The following lemma gives a useful reinterpretation of the (left-/right-/semi-) Fredholm property.

Lemma 1.18. *Let $T \in \mathcal{L}(E, F)$ for Banach spaces E, F. Then:*

(a) T is left Fredholm iff there exists $S_1 \in \mathcal{L}(F, E)$ such that $S_1T - \mathbb{1}$ is compact;

(b) T is right Fredholm iff there exists $S_2 \in \mathcal{L}(F, E)$ such that $TS_2 - \mathbb{1}$ is compact;

(c) T is Fredholm iff there exists $S \in \mathcal{L}(F, E)$ such that $ST - \mathbb{1}$ and $TS - \mathbb{1}$ are compact.

For $E = F$ this can be interpreted in terms of the Banach algebra $\mathcal{L}(E, E)/\mathcal{K}(E, E)$ from Remark 1.14: T is (left-/right-)Fredholm iff $[T]$ has a (left-/right-)inverse in $\mathcal{L}(E, E)/\mathcal{K}(E, E)$. This also explains the terminology “left/right-Fredholm”.

Proposition 1.19 (Properties of the Fredholm index). *Let E, F, G be Banach spaces and $T \in \mathcal{L}(E, F)$. Then:*

(a) If T is (semi-)Fredholm, then so is T^* and

$$\operatorname{ind}(T^*) = -\operatorname{ind}(T).$$

(b) If T is (semi-)Fredholm and $K : E \to F$ is compact, then $T + K$ is (semi-)Fredholm and

$$\operatorname{ind}(T + K) = \operatorname{ind}(T).$$

(c) If $S : F \to G$ and $T : E \to F$ are Fredholm, then so is $S \circ T : E \to F$ and

$$\operatorname{ind}(S \circ T) = \operatorname{ind}(S) + \operatorname{ind}(T).$$

Proof (see [19]. Part (a) follows directly by dualizing the criteria in Lemma 1.18 and using the isomorphisms in Problem 1.14. Note that if T is left-Fredholm then T^* is right-Fredholm and vice versa. Part (b) follows similarly from Lemma 1.18: If $S_1 T - \mathbb{1} = K_1$ is compact then $S_1(T + K) - \mathbb{1} = S_1 K + K_1$ is compact, and similary for the right-Fredholm case. For (c) consider the short exact sequences

$$0 \longrightarrow \ker T \longrightarrow \ker(S \circ T) \xrightarrow{T} \ker S \cap \ker T \longrightarrow 0$$

$$0 \longrightarrow \operatorname{im} S/\operatorname{im}(S \circ T) \longrightarrow G/\operatorname{im}(S \circ T) \longrightarrow G/\operatorname{im} S \longrightarrow 0$$

$$0 \longrightarrow \ker S/(\ker S \cap \operatorname{im} T) \longrightarrow F/\operatorname{im} T \xrightarrow{S} \operatorname{im} S/\operatorname{im}(S \circ T) \longrightarrow 0$$

$$0 \longrightarrow \ker S \cap \operatorname{im} T \longrightarrow \ker S \longrightarrow \ker S/(\ker S \cap \operatorname{im} T) \longrightarrow 0$$

This shows first that $\ker(S \circ T)$ and $G/\operatorname{im}(S \circ T)$ are finite dimensional, so $S \circ T$ is Fredholm by Problem 1.13. Next, note that the alternating sum of dimensions in each row is zero. Summing these up over the rows with alternating signs yields $\mathrm{ind}(S \circ T) = \mathrm{ind}(S) + \mathrm{ind}(T)$. $\square$

1.5 Topology of Fredholm operators

In this section we begin investigating the topology of the space of Fredholm operators. Everything is based on the following fundamental result, which stands in stark contrast to the finite dimensional situation.

Theorem 1.20 (Kuiper [30]). *For an infinite dimensional separable (real or complex) Hilbert space H the following spaces are contractible:*
(i) the unit sphere $S(H) = \{x \in H \mid \|x\| = 1\}$;
(ii) the general linear group $GL(H) = \{\text{isomorphisms } H \to H\}$ with the operator norm;
(iii) the orthogonal (or unitary in the complex case) group $O(H) = \{\text{isometries } H \to H\}$ with the operator norm.

Corollary 1.21. *For an infinite dimensional separable Hilbert space H the spaces $\mathcal{F}_k(H)$ of Fredholmo operators $H \to H$ of index k have the following properties:*
(a) $\mathcal{F}_k(H)$ and $\mathcal{F}_\ell(H)$ are homotopy equivalent for all $k, \ell \in \mathbb{Z}$.
(b) $\mathcal{F}_k(H)$ is path connected for all $k \in \mathbb{Z}$.

Proof. For (a) pick a Hilbert space basis $e_1, e_2, \dots$ and define for $n \in \mathbb{N}$ the *shift operator*

$$S_n : H \to H, \qquad e_i \mapsto e_{i+n}.$$

Its dual (with respect to the inner product on H) is given by

$$S_n^* : H \to H, \qquad e_i \mapsto \begin{cases} e_{i-n} & i > n, \\ 0 & i \le n. \end{cases}$$

They obviously satisfy $S_n \in \mathcal{F}_{-n}(H)$, $S_n^* \in \mathcal{F}_n(H)$ and $S_n^* \circ S_n = \mathbb{1}$. The other composition $S_n \circ S_n^*$ equals the orthogonal projection P onto $\mathrm{span}\{e_{n+1}, e_{n+2}, \dots\}$. Since $\mathbb{1} - P$ is of finite rank, hence compact, $P_t := P + t(\mathbb{1} - P)$ for $t \in [0,1]$ is a path in $\mathcal{F}_0(H)$ from P to $\mathbb{1}$. Now the maps

$$\begin{aligned} \Phi &: \mathcal{F}_k(H) \to \mathcal{F}_{k-n}(H), \qquad T \mapsto S_n \circ T, \\ \Psi &: \mathcal{F}_{k-n}(H) \to \mathcal{F}_k(H), \qquad T \mapsto S_n^* \circ T \end{aligned}$$

are continuous and $\Psi \circ \Phi = \mathrm{id}$. The map $\Phi \circ \Psi$ is given by composition with P, and is therefore homotopic to the identity via compositions with P_t. Thus Φ is a homotopy equivalence with homotopy inverse Ψ.

For (b), in view of (a) it suffices to consider the case $k = 0$. Let $T \in \mathcal{F}_0(H)$ and pick complements to $\ker T$ and $\mathrm{im}\, T$ such that

$$H = \ker T \oplus V = W \oplus \mathrm{im}\, T.$$

With respect to these decompositions T is represented by the matrix $\left(\begin{smallmatrix} 0 & 0 \\ 0 & B \end{smallmatrix}\right)$ with the isomorphism $B = T|_V : V \to \mathrm{im}\, T$. Since $\mathrm{ind}(T) = 0$, the finite dimensional spaces $\ker T$ and W have the same dimension. Pick any isomorphism $A : \ker T \to W$. Then the matrices $\left(\begin{smallmatrix} tA & 0 \\ 0 & B \end{smallmatrix}\right)$ for $t \in [0,1]$ define a path in $\mathcal{F}_0(H)$ from T to the identity $\mathbb{1} \in \mathcal{F}_0(H)$. □

We will see later that the spaces $\mathcal{F}_k(H)$ are not contractible but have fundamental group $\mathbb{Z}/2\mathbb{Z}$.

1.6 Sobolev and Hölder spaces

Here we collect some facts about Hölder and Sobolev spaces. For more details and proofs see e.g. [4, 39, 45].

Definitions. Let $\Omega \subset \mathbb{R}^n$ be open. For a function $f : \Omega \to \mathbb{R}$ and a multi-index $s = (s_1, \dots, s_n) \in \mathbb{N}_0^n$ we denote $|s| := s_1 + \dots + s_n$ and $D^s f := \frac{\partial^{|s|} f}{\partial x_1^{s_1} \cdots \partial x_n^{s_n}}$.

For $k \in \mathbb{N}_0$ and $1 \leq p < \infty$ the *Sobolev space*

$$W^{k,p}(\Omega)$$

is the space of L^1_{loc}-functions $f : \Omega \to \mathbb{R}$ with k weak derivatives in L^p. It is a Banach space equipped with the norm

$$\|f\|_{W^{k,p}} := \sum_{|s| \leq k} \|D^s f\|_{L^p}.$$

Now let $\Omega \subset \mathbb{R}^n$ be bounded open. In this case we will sometimes write $W^{k,p}(\overline{\Omega})$ for $W^{k,p}(\Omega)$. For $0 < \alpha \leq 1$ the *Hölder space*

$$C^{0,\alpha}(\overline{\Omega})$$

is the space of continuous functions $f : \overline{\Omega} \to \mathbb{R}$ such that

$$\mathrm{Höl}_\alpha(f) := \sup_{x \neq y \in \overline{\Omega}} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$$

is finite. It is a Banach space with the norm

$$\|f\|_{C^{0,\alpha}} := \|f\|_{\max} + \mathrm{Höl}_\alpha(f).$$

For $\alpha = 0$ we define $C^{0,0}(\overline{\Omega}) := C^0(\overline{\Omega})$ with the maximum norm. More generally, for $k \in \mathbb{N}_0$ and $0 \leq \alpha \leq 1$ the *Hölder space*

$$C^{k,\alpha}(\overline{\Omega})$$

is the space of functions $\overline{\Omega} \to \mathbb{R}$ with k derivatives in $C^{0,\alpha}$. It is a Banach space with the norm

$$\|f\|_{C^{k,\alpha}} := \sum_{|s|\le k} \|D^s f\|_{C^{0,\alpha}}.$$

The preceding definitions carry over verbatim to spaces of functions with values in $\mathbb{R}^m$ which we denote by $W^{k,p}(\Omega, \mathbb{R}^m)$ and $C^{k,\alpha}(\overline{\Omega}, \mathbb{R}^m)$, respectively. Moreover, we can replace the domain $\Omega \subset \mathbb{R}^n$ by an n-dimensional compact manifold N (possibly with boundary). For this, we pick finitely many coordinate charts $\phi_i : U_i \to V_i$ between open subsets $U_i \subset N$ and $V_i \subset \{x \in \mathbb{R}^n \mid x_n \ge 0\}$ such that $N = \bigcup_{i=1}^r U_i$, and a smooth partition of unity $\chi_i : N \to [0,1]$ subordinate to this covering (see Section 2.2). Then we define $W^{k,p}(N)$ to be the space of maps $f : N \to \mathbb{R}$ such that $(\chi_i f) \circ \phi_i^{-1} \in W^{m,p}(V_i)$ for all i.

Problem 1.15. For N and $W^{k,p}(N)$ as above show:
(a) The definition of the space $W^{k,p}(N)$ does not depend on the choice of coordinate covering and partition of unity.
(b) $W^{k,p}(N)$ becomes a Banach space with the norm

$$\|f\|_{W^{k,p}(N)} := \sum_{i=1}^{r} \|(\chi_i f) \circ \phi_i^{-1}\|_{W^{k,p}(V_i)}.$$

(c) Different choices of coordinate coverings and partitions of unity give equivalent norms on $W^{k,p}(N)$.

The space of Hölder functions $C^{k,\alpha}(N)$ is defined analogously. For a smooth vector bundle $E \to N$ we define the spaces $W^{k,p}(N, E)$ and $C^{k,\alpha}(N)$ of Sobolev resp. Hölder sections in E as above, using trivializations of E over the (sufficiently small) coordinate domains U_i to view $(\chi_i f) \circ \phi_i^{-1}$ as elements in $W^{k,p}(V_i, \mathbb{R}^m)$.

Embedding theorems. Next we collect the results concerning embeddings between Sobolev and Hölder spaces. The first one is an easy corollary of the Arzela-Ascoli theorem (see e.g. [4, Theorem 8.6]).

Corollary 1.22. *Let N be a compact manifold (possibly with boundary), $k, \ell \in \mathbb{N}_0$, and $\alpha, \beta \in [0,1]$. If $k + \alpha > \ell + \beta$, then the inclusion map $C^{k,\alpha}(N) \hookrightarrow C^{\ell,\beta}(N)$ is compact.*

The next theorem is usually refered to as the Sobolev embedding theorem, although assertion (b) is due to Rellich and Kondrachov. For the proof see e.g. [4, Theorems 8.9 and 8.13].

Theorem 1.23 (Sobolev embedding theorem)**.** *Let N be a compact n-dimensional manifold (possibly with boundary), $k, \ell \in \mathbb{N}_0$, $p, q \in [1, \infty)$, and $\alpha \in [0,1]$.*
(a) If $k \geq \ell$ and $k - n/p \geq \ell - n/q$, then $W^{k,p}(N) \subset W^{\ell,q}(N)$ and the inclusion map is continuous.
(b) If $k > \ell$ and $k - n/p > \ell - n/q$, then the inclusion map $W^{k,p}(N) \subset W^{\ell,q}(N)$ is compact.
(c) If $k - n/p \geq \ell + \alpha$ and $0 < \alpha < 1$, then $W^{k,p}(N) \subset C^{\ell,\alpha}(N)$ and the inclusion map is continuous.
(d) If $k - n/p > \ell + \alpha$, then the inclusion map $W^{k,p}(N) \subset C^{\ell,\alpha}(N)$ is compact.

Products. Products of Sobolev functions are in general not of the same Sobolev class. For example, the product of two L^2-functions on $[0,1]$ is in general in L^1 but not in L^2. By contrast, if a Sobolev space embeds into the continuous functions then it is closed under products (see [39, Corollary 9.7] and [9, Lemma 3.7]):

Theorem 1.24 (Products of Sobolev functions)**.** *Let N be a compact n-dimensional manifold (possibly with boundary), $k, \ell \in \mathbb{N}_0$ and $p, q \in [1, \infty)$ such that $k - n/p > 0$ and $k > \ell$. Then* *pointwise multiplication of functions turns $W^{k,p}(N)$ into a Banach algebra and $W^{\ell,q}(N)$ into a Banach module over $W^{k,p}(N)$. Moreover, multiplication $f \mapsto fg$ with a fixed $g \in W^{\ell,q}(N)$ defines a compact linear map $W^{k,p}(N) \to W^{\ell,q}(N)$.*

More explicitly, the first assertion of the theorem means that the product $(f, g) \mapsto fg$ defines continuous maps

$$W^{k,p}(N) \times W^{k,p}(N) \to W^{k,p}(N)$$

and

$$W^{k,p}(N) \times W^{\ell,q}(N) \to W^{\ell,q}(N)$$

satisfying with some constants C, C' the estimates

$$\|fg\|_{W^{k,p}} \le C\|f\|_{W^{k,p}}\|g\|_{W^{k,p}}, \qquad \|fg\|_{W^{\ell,q}} \le C'\|f\|_{W^{k,p}}\|g\|_{W^{\ell,q}}.$$

The corresponding statement for Hölder functions is simpler:

Problem 1.16 (Products of Hölder functions)*.* Let N be a compact manifold (possibly with boundary), $k, \ell \in \mathbb{N}_0$ and $\alpha, \beta \in [0, 1]$ such that $k+\alpha > \ell+\beta$. Then pointwise multiplication of functions turns $C^{k,\alpha}(N)$ into a Banach algebra and $C^{\ell,\beta}(N)$ into a Banach module over $C^{k,\alpha}(N)$. Moreover, multiplication $f \mapsto fg$ with a fixed $g \in C^{\ell,\beta}(N)$ defines a compact linear map $C^{k,\alpha}(N) \to C^{\ell,\beta}(N)$.

Chapter 2

Banach manifolds and smooth maps

2.1 Analysis on Banach spaces

Consider Banach spaces E, F and an open subset $U \subset E$. A map $f : U \to F$ is called *Fréchet differentiable* at $x \in U$ if there exists a continuous linear map $Df(x) : E \to F$ such that

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Df(x)h\|}{\|h\|} = 0,$$

where the limit is taken over all $h \in E$ such that $x + h \in U$. Then $Df(x) \in \mathcal{L}(E, F)$ is called the *Fréchet derivative* of f at x. It has the same properties as the total derivative of maps $\mathbb{R}^n \to \mathbb{R}^m$ (with the same proofs left as an exercise):

Proposition 2.1 (Properties of the Fréchet derivative)**.**
(a) Fréchet differentiability at x implies continuity at x.
(b) (Chain rule) If $f : E \supset U \to F$ is Fréchet differentiable at $x \in U$ and $g : F \supset V \to G$ is Fréchet differentiable at $f(x) \in V$, then $g \circ f$ is Fréchet differentiable at x and

$$D(g \circ f)(x) = Dg\big(f(x)\big) \circ Df(x).$$

Remark 2.2. A map $f : E \supset U \to F$ is called *Gâteaux differentiable* at $x \in U$ if for each $h \in E$ the directional derivative

$$Df(x)h = \lim_{t \to 0} \frac{f(x+th) - f(x)}{t}$$

exists and defines a continuous linear map in h. Fréchet differentiability obviously implies Gâteaux differentiability but not conversely: for example, the function $f(x,y) = x^4y/(x^6+y^3)$ on $\mathbb{R}^2$ is Gâteaux differentiable with $Df(0,0) = 0$ but not continuous at $(0,0)$. We will only use the Fréchet derivative and refer to it in the following simply as the *derivative*.

If $f : E \supset U \to F$ is differentiable at each $x \in U$, then its derivative defines a map $U \to \mathcal{L}(E,F)$ and we can define (if it exists) its second derivative $D^2f(x) \in \mathcal{L}\big(E, \mathcal{L}(E,F)\big)$. Inductively, we define the k-th derivative

$$D^k f(x) \in \mathcal{L}\Big(E, \big(E, \ldots \mathcal{L}(E,F)\ldots\big)\Big) =: \mathcal{L}^k(E,F).$$

Problem 2.1. Let E, F be Banach spaces, $U \subset E$ open and $f : U \to F$ continuously differentiable.
(a) Show: if $x \in U$ and $\xi \in E$ such that $x + t\xi \in U$ for all $t \in [0,1]$, then

$$f(x+\xi) - f(x) = \int_0^1 Df(x+t\xi)\cdot\xi\, dt.$$

(b) Formulate and prove Taylor's theorem if f is $(k+1)$ times continuously differentiable.

Implicit and inverse function theorems. The implicit and inverse function theorems carry over with essentially the same proofs to maps between Banach spaces.

Theorem 2.3 (Implicit function theorem)**.** *Let E, F, G be Banach spaces and $U \subset E$, $V \subset F$ open neighbourhoods of* 0. *Let $f : U \times V \to G$ be continuous with continuous Fréchet derivative $D_2 f : U \times V \to \mathcal{L}(F,G)$ with respect to the second variable. Suppose that $f(0,0) = 0$ and $D_2f(0,0)$ is an isomorphism. Then there exist open balls $B_\delta \subset U$ and $B_\varepsilon \subset V$ around* *the origin and a unique continuous function $g : B_\delta \to B_\varepsilon$ such that*

$$f(x, y) = 0 \iff y = g(x) \quad \textit{for } (x, y) \in B_\delta \times B_\varepsilon.$$

Moreover, if f is of class C^k for some $k \in \mathbb{N} \cup \{\infty\}$ then so is g.

Proof. After replacing f by $D_2 f(0,0)^{-1} \circ f$ (which has the same zero set and regularity as f) we may assume that $F = G$ and $D_2 f(0,0) = \mathbb{1}$. We pick $\delta, \varepsilon > 0$ such that $B_\delta \subset U$ and $B_\varepsilon \subset V$ and

(i) $\|\mathbb{1} - D_2 f(x, y)\| < 1/2$ for all $(x, y) \in B_\delta \times B_\varepsilon$,

(ii) $\|f(x, 0)\| < \varepsilon/2$ for all $x \in B_\delta$.

Here we first arrange (i) by continuity of $D_2 f$ and then shrink δ further to arrange (ii) by continuity of f. Now we proceed in two steps.

Step 1. We first show that for each $x \in B_\delta$ there exists a unique $g(x) \in B_\varepsilon$ with $f\big(x, g(x)\big) = 0$. For this, fix $x \in B_\delta$ and consider the map

$$T_x : \overline{B}_\varepsilon \to F, \qquad T_x y := y - f(x, y)$$

whose fixed points correspond to solutions y of $f(x, y) = 0$. For $y, z \in \overline{B}_\varepsilon$ we estimate

$$\begin{aligned} \|T_x y - T_x z\| &= \|y - z + f(x, z) - f(x, y)\| \\ &= \left\| \int_0^1 \Big(\mathbb{1} - D_2 f\big(x, (1-t)y + tz\big)\Big)(y - z)dt \right\| \\ &\leq \max\nolimits_{t \in [0,1]} \|\mathbb{1} - D_2 f\big(x, (1-t)y + tz\big)\| \, \|y - z\| \\ &\leq \frac{1}{2}\|y - z\|. \end{aligned} \tag{2.1}$$

Here the second equality follows from Problem 2.1, and the last inequality follows from (i) above because $(1-t)y + tz \in B_\delta$ for all

$t \in [0, 1]$. Setting $z := 0$ we obtain from (ii) above

$$\|T_x y\| \leq \|T_x y - T_x 0\| + \|T_x 0\| \leq \frac{1}{2}\|y\| + \|f(x, 0)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus T_x maps the complete metric space $\overline{B}_\varepsilon$ to its interior and is contracting, so by the Banach fixed point theorem it has a unique fixed point $g(x) \in B_\varepsilon$.

Step 2. Now we show that the map $g : B_\delta \to B_\varepsilon$ from Step 1 is continuous. For this, using $T_x g(x) = g(x)$, the definition of T_x and the contraction property (2.1) we estimate for $x, x' \in B_\delta$

$$\begin{aligned}\|g(x) - g(x')\| &= \|T_x g(x) - T_{x'} g(x')\| \\ &\leq \|T_x g(x) - T_{x'} g(x)\| + \|T_{x'} g(x) - T_{x'} g(x')\| \\ &\leq \|f\big(x', g(x)\big) - f\big(x, g(x)\big)\| + \frac{1}{2}\|g(x) - g(x')\|,\end{aligned}$$

hence $\|g(x) - g(x')\| \leq 2\|f\big(x', g(x)\big) - f\big(x, g(x)\big)\|$. Thus continuity of f in the first variable implies continuity of g.

Step 3. Finally, suppose that the derivative $D_1 f(0, 0)$ with respect to the first variable exists at the point $(0, 0)$. Then for $x \in B_\delta$ we can write

$$\begin{aligned}0 &= f\big(x, g(x)\big) - f(0, 0) \\ &= f\big(x, g(x)\big) - f(x, 0) + f(x, 0) - f(0, 0) \\ &= D_2 f(x, 0) g(x) + r_2(x) + D_1 f(0, 0) x + r_1(x)\end{aligned}$$

with functions r_1, r_2 satisfying

$$\lim_{x \to 0} \frac{r_1(x)}{\|x\|} = \lim_{x \to 0} \frac{r_2(x)}{\|g(x)\|} = 0.$$

By continuity of $D_2 f$, for x sufficiently small $D_2 f(x, 0)$ is invertible and $\|D_2 f(x, 0)^{-1}\| \leq C$ with a constant C independent of x. For such x we can solve the preceding equation for

$$g(x) = -D_2 f(x, 0)^{-1} \Big[D_1 f(0, 0) x + r_1(x) + r_2(x)\Big].$$

For sufficiently small x we deduce the estimate $\|g(x)\| \leq C_1\|x\| + \frac{1}{2}\|g(x)\|$ for a constant C_1, hence $\|g(x)\| \leq 2C_1\|x\|$. Now with $g(0) = 0$ we see that

$$\frac{\|g(x) - g(0) - D_2 f(0,0)^{-1} D_1 f(0,0)x\|}{\|x\|}$$
$$\leq \|D_2 f(x,0)^{-1} - D_2 f(0,0)^{-1}\| \|D_1 f(0,0)\| + C\Big[\frac{\|r_1(x)\|}{\|x\|} + \frac{\|r_2(x\|)}{\|x\|}\Big]$$

converges to 0 as $x \to 0$, so g is differentiable at 0 with derivative

$$Dg(0) = -D_2 f(0,0)^{-1} D_1 f(0,0).$$

If $D_1 f(x,y)$ exists for all $(x,y) \in U \times V$, then the same reasoning shows that g is differentiable at each $x \in B_\delta$ and

$$Dg(x) = -D_2 f\big(x, g(x)\big)^{-1} D_1 f\big(x, g(x)\big).$$

This formula shows that if f is of class C^k then so is g. □

Remark 2.4. In contrast to the finite dimensional case, in the preceding proof we could not apply the Banach fixed point theorem on the space $C(\overline{B}_\delta, \overline{B}_\varepsilon)$ of continuous maps $\overline{B}_\delta \to \overline{B}_\varepsilon$ because this is not a complete metric space if E is infinite dimensional.

For $k \in \mathbb{N} \cup \{\infty\}$ a C^k-map $f : U \to V$ between open subsets of Banach spaces is called a C^k-*diffeomorphism* if it is bijective and its inverse is also of class C^k. As a formal consequence of the implicit function theorem we obtain

Corollary 2.5 (Inverse function theorem)**.** *Let E, F be Banach spaces and $U \subset E$ an open neighbourhood of* 0. *Let $f : U \to F$ be of class C^k for some $k \in \mathbb{N} \cup \{\infty\}$. Suppose that $f(0) = 0$ and $Df(0) \in \mathcal{L}(E,F)$ is an isomorphism. Then there exists an open neighbourhood $V \subset U$ of the origin such that $f(V) \subset F$ is open and $f : V \to f(V)$ is a C^k-diffeomorphism.*

Proof. The function $\widehat{f} : F \times U \to F$, $\widehat{f}(x, y) := f(y) - x$ satisfies the hypotheses of the implicit function theorem. Thus there exist open balls $B_\delta \subset F$ and $B_\varepsilon \subset U$ around the origin and a C^k-function $g : B_\delta \to B_\varepsilon$ such that $f(y) = x \iff y = g(x)$ for $(x, y) \in B_\delta \times B_\varepsilon$. Then $V := g(B_\delta) = f^{-1}(B_\delta) \subset U$ and $f(V) = B_\delta$ are open neighbourhoods of the origin and $f : V \to f(V)$ is a C^k-diffeomorphism with inverse g. □

Problem 2.2. Show that, conversely, the implicit function theorem is also a formal consequence of the inverse function theorem.

Bump functions. A *bump function* (or *cutoff function*) on a Banach space E is a smooth (i.e. C^∞) function $\phi : E \to [0, 1]$ with $\phi(0) > 0$ and $\phi(x) = 0$ if $\|x\| \geq 1$. Such functions are an important tool in analysis, for example for the construction of partitions of unity (see the next section). Therefore, we collect here some results on their existence from [17].

Recall first that a bump fuction $\psi : \mathbb{R} \to [0, 1]$ is given by $\psi(x) = \chi(x + 1)\chi(-x - 1)$ for the smooth function

$$\chi(x) := \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \leq 0. \end{cases}$$

Let us call a norm $\| \ \|$ on E *smooth* if it is smooth as a function $E \to \mathbb{R}$ away from the origin. If E admits a smooth norm $\| \ \|$ which is equivalent to the given norm $\| \ \|_E$, then $\phi(x) := \psi(\|x\|)$ with ψ above defines a smooth bump function on E. Since the Euclidean norm on $\mathbb{R}^n$ and the norm $\|x\| = \langle x, x\rangle^{1/2}$ on a Hilbert space are smooth (exercise), it follows that each finite dimensional Banach space and each Hilbert space admits a bump function. In general, the situation is more subtle:

Problem 2.3. Let (Ω, μ) be a measure space with at least two disjoint subsets of positive measure. Then for $1 \leq p < \infty$ the

function $\| \|_p^p : L^p(\Omega, \mathbb{R}) \to \mathbb{R}$ is outside the origin of class
(i) C^∞ if p is an even integer,
(ii) $C^{p-1,1}$ if p is an odd integer,
(iii) $C^{[p],p-[p]}$ if p is not an integer.
Hint: Investigate first the regularity of the function $f(x) = |x|^p$ on $\mathbb{R}$ at the origin.

The following result gives criteria for the existence of bump functions. Recall that a Banach space E is called *reflexive* if the canonical embedding $\iota_E : E \hookrightarrow E^{**}$, $\iota_E(x)(f) := f(x)$ is an isomorphism.

Theorem 2.6 ([17])**.** *Consider the following properties for a Banach space E:*
(a) E (or equivalently E^*) is reflexive.
(b) E admits a smooth norm.
(c) E admits a smooth bump function.
(d) E^* is separable.
(e) E is separable.
Between these properties the following implications hold:

$$(a) \Longrightarrow (b) \Longrightarrow (c) \Longleftarrow (d) \Longrightarrow (e), \qquad (c)\&(e) \Longrightarrow (b)\&(d).$$

Corollary 2.7. *(a) The following spaces admit smooth bump functions: Hilbert spaces, $L^p(\Omega)$ and $W^{k,p}(\Omega)$ for $1 < p < \infty$, and the space $c_0 \subset \ell^\infty$ of sequences converging to zero.*
(b) The following spaces do not *admit smooth bump functions: $L^1(\Omega)$ and $W^{k,1}(\Omega)$, and $C^{k,\alpha}(\Omega)$.*

Proof. All spaces in (a) have separable dual spaces (the dual space of c_0 being ℓ^1), while the duals of the spaces in (b) are not separable: For example, in the dual space to $C([0, 1]$ the functionals $\delta_x(f) = f(x)$ for $x \in [0, 1]$ satisfy $\|\delta_x - \delta_y\| = 2$ for all $x \neq y$ because there exists a continuous function $f : [0,1] \to [-1,1]$ with $f(x) = 1$ and $f(y) = -1$. $\square$

2.2 Banach and Hilbert manifolds

The definition of a Banach manifold completely parallels the definition of a finite dimensional manifold, with $\mathbb{R}^n$ replaced by a Banach space.

Let X be a topological space and fix $k \in \mathbb{N}_0 \cup \{\infty\}$. A *Banach chart* (U, ϕ, E) on X is a homeomorphism $\phi : U \to \phi(U)$ from an open subset $U \subset X$ onto an open subset $\phi(U)$ of a Banach space E. Two charts (U_i, ϕ_i, E_i) and (U_j, ϕ_j, E_j) are called C^k-*compatible* if the transition map

$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \to \phi_i(U_i \cap U_j)$$

is a C^k-diffeomorphism. A C^k-*Banach atlas* $\mathcal{A}$ on X is a collection of pairwise C^k-compatible Banach charts $(U_i, \phi_i, E_i)_{i \in I}$ such that $\bigcup_{i \in I} U_i = X$. Two atlases $\mathcal{A}$, $\mathcal{B}$ are called *equivalent* if $\mathcal{A} \cup \mathcal{B}$ is an atlas, i.e., every chart of $\mathcal{A}$ is compatible with every chart of $\mathcal{B}$.

Remark 2.8. Equivalently, one could start with X being just a set. Then a chart (U, ϕ, E) is just a bijection $\phi : U \to \phi(U)$ from a subset $U \subset X$ onto an open subset $\phi(U)$ of a Banach space E, and we need to require in addition that for each pair of charts the set $\phi_i(U_i \cap U_j)$ is open in E_i. The notions of C^k-compatibility and a C^k-Banach atlas are unchanged. For a C^k-Banach atlas $\mathcal{A}$, the sets $\phi_i^{-1}(W)$ for all charts (U_i, ϕ_i, E_i) of $\mathcal{A}$ and all open subset $W \subset \phi_i(U_i)$ are then the basis of a topology on X, so that $\mathcal{A}$ becomes an atlas on the topological space X in the sense of the earlier definition.

Problem 2.4. Let $(X, \mathcal{A})$ be a topological space with a C^k-Banach

atlas, $k \in \mathbb{N}_0 \cup \{\infty\}$. Show:
(a) X is connected if and only if it is path connected.
(b) If X is connected and $k \geq 1$, then all the Banach spaces E_i appearing in the charts of $\mathcal{A}$ are isomorphic.

If all the Banach spaces E_i appearing in the charts of an atlas $\mathcal{A}$ are isomorphc to a fixed Banach space E, then we will drop the E_i from the notation of charts and say that $(X, \mathcal{A})$ is *modelled over* E. By the preceding problem, this is the case whenever X is connected.

Problem 2.5. Consider the space

$$X := \{0,1\} \times \mathbb{R}/(0,x) \sim (1,x) \text{ for all } x < 0$$

with the quotient topology. Show: the two charts $\phi_i : \{i\} \times \mathbb{R} \to \mathbb{R}$, $\phi_i(i,x) = x$ for $i = 0,1$ define a smooth atlas on X (modelled over $\mathbb{R}$), but X is not Hausdorff.

Pathological examples of this kind are ruled out by the following

Definition 2.1. A C^k-*Banach manifold* is a paracompact Hausdorff space X equipped with an equivalence class of C^k-Banach atlases. For $k = 0$ it is called a *topological Banach manifold* and for $k = \infty$ a *smooth Banach manifold*, or simply a *Banach manifold*. If X is modelled on a Hilbert space it is called a *Hilbert manifold*.

As is the finite dimensional case, a map $f : X \to Y$ between Banach manifolds is defined to be of class C^k if for all charts $\phi : X \supset U \to E$ and $\psi : Y \supset V \to F$ the composition $\psi \circ f \circ \phi^{-1}$ is of class C^k. It is called a C^k-*diffeomorphism* if it has an inverse of class C^k.

The Hausdorff property in the definition of a Banach manifold is crucial and underlies many constructions. Paracompactness (defined below) is less essential and sometimes omitted; it is satisfied in all examples of interest and ensures the existence of continuous partitions of unity that we will discuss next.

Continuous partitions of unity. We begin with some definitions about a topological space X. A *cover* $\mathcal{U}$ of X is a collection $(U_i)_{i\in I}$ of subsets $U_i \subset X$ with $\bigcup_{i\in I} U_i = X$. It is called *open* if all the U_i are open, and *locally finite* if every point of X has a neighbourhood which intersects only finitely many of the U_i. A cover $(V_j)_{j\in J}$ is a *refinement* of $\mathcal{U}$ if for every $j \in J$ there is some $i \in I$ with $V_j \subset U_i$. The space X is called *paracompact* if every open cover has a locally finite open refinement.

A *continuous partition of unity* on X is a collection $(\chi_i)_{i\in I}$ of continuous functions $\chi_i : X \to [0, 1]$ such the their supports form a locally finite cover and $\sum_{i\in I} \chi_i(x) = 1$ for all $x \in X$ (due to local finiteness this is a finite sum for each x). The partition of unity is *subordinate to the open cover* $(U_i)_{i\in I}$ (with the same index set) if $\operatorname{supp} \chi_i \subset U_i$ for all $i \in I$.

The space X is called *metrizable* if its topology is induced by a metric; *locally metrizable* if every point has a metrizable open neighbourhood; and *normal* if any two disjoint closed subsets can be separated by open sets.

Proposition 2.9. *(a) A Hausdorff space admits continuous partitions of unity subordinate to every open cover if and only if it is paracompact.*
(b) A topological space is metrizable if and only if it is Hausdorff, paracompact and locally metrizable.
(c) Every paracompact Hausdorff space is normal.

Corollary 2.10. *(a) Every Banach manifold is metrizable and admits continuous partitions of unity subordinate to every open cover.*

(b) Every open subset of a Banach manifold is a Banach manifold.
(c) Every Banach space is a Banach manifold.

Proof. Part (a) follows from Proposition 2.9(a) and (b) because every Banach manifold is locally metrizable. Part (b) follows because metrizability is inherited by open subsets (whereas paracompactness is in general not), and (c) holds because every Banach space is metrizable. □

Problem 2.6. Show: (a) A topological space X is paracompact iff it admits a locally finite cover by paracompact open sets.
(b) A topological space X with a C^k-Banach atlas, $k \in \mathbb{N}_0 \cup \{\infty\}$, is paracompact iff it admits an equivalent locally finite atlas.

The following characterization of normal spaces is often useful.

Proposition 2.11. *For a topological space X the following are equivalent:*
(i) X is normal;
(ii) for any two disjoint closed sets A, B there exists a continuous function $f : X \to [0, 1]$ with $f = 0$ on A and $f = 1$ on B (Urysohn lemma);
(iii) any continuous function $A \to \mathbb{R}$ defined on a closed subset $A \subset X$ can be extended to a continuous function $X \to \mathbb{R}$ (Tietze extension theorem);
(iv) X admits continuous partitions of unity subordinate to every locally finite open cover.

The main step for deriving (iv) from the other properties is

Lemma 2.12. *For each locally finite open cover $(U_i)_{i\in I}$ of a normal space X there exists a locally finite open cover $(V_i)_{i\in I}$ with $\overline{V}_i \subset U_i$ for all $i \in I$.*

Indeed, given such covers $\mathcal{U}$ and $\mathcal{V}$, Urysohn's lemma provides continuous functions $\rho_i : X \to [0,1]$ with $\rho_i = 1$ on $\overline{V}_i$ and $\rho_i = 0$ outside U_i, and $\chi_i := \rho_i/(\sum_{j\in I}\rho_j)$ is a continuous partition of unity subordinate to $\mathcal{U}$. Conversely, Lemma 2.12 follows from (iv) by setting $V_i := \{\chi_i > 0\}$ for a partition of unity subordinate to $\mathcal{U}$.

Smooth partitions of unity. Now let X be a Banach manifold. The existence of *smooth partitions of unity* on X (defined as above with the χ_i being smooth) is closely related to the following property. We say that X has the *smooth approximation property* if for every continuous map $f : X \to [0,1]$ and every $\varepsilon > 0$ there exists a smooth map $\widetilde{f} : X \to [0,1]$ such that $\operatorname{supp}\widetilde{f} \subset \operatorname{supp} f$ and $\|\widetilde{f}(x) - f(x)\| < \varepsilon$ for all $x \in X$.

Proposition 2.13. *For a Banach manifold X modelled on the Banach space E the following are equivalent:*
(i) X admits smooth partitions of unity subordinate to every open cover;
(ii) X has the smooth approximation property;
(iii) the smooth approximation property holds for all continuous functions $f : E \to [0,1]$ with bounded support.

Proof. $(i) \implies (ii)$: For later use, instead of the interval $[0,1]$ consider more generally a convex subset C of a Banach space containing 0. Let a continuous map $f : X \to C$ and an $\varepsilon > 0$ be given. Let $U_0 := \{x \in X \mid \|f(x)\| < \varepsilon\}$, and for every $y \in Y := X \setminus U_0$ let $U_y := \{x \in X \mid \|f(x) - f(y)\| < \varepsilon\}$. The sets U_0 and $(U_y)_{y\in Y}$ form an open cover of X, so by assumption there exists a smooth partition of unity χ_0 and $(\chi_y)_{y\in Y}$ subordinate to this cover. The sum

$$\widetilde{f}(x) := \sum_{y \in Y} \chi_y(x) f(y)$$

is finite near each $x \in X$ and thus defines a smooth function $\widetilde{f} : X \to F$. Moreover, adding the term $\chi_0(x) \cdot 0$ and using convexity of C we see that $\widetilde{f}(x) \in C$. Since by construction $U_y \cap (X \setminus \operatorname{supp} f) = \emptyset$ for all $y \in Y$, we have $\widetilde{f} = 0$ on $X \setminus \operatorname{supp} f$ and thus $\operatorname{supp} \widetilde{f} \subset \operatorname{supp} f$. For $x \in Y$ we have $\chi_0(x) = 0$ and therefore

$$\begin{aligned} \|\widetilde{f}(x) - f(x)\| &= \|\sum_{y \in Y \cap U_x} \chi_y(x)\big(f(y) - f(x)\big)\| \\ &\le \sum_{y \in Y \cap U_x} \chi_y(x)\|f(y) - f(x)\| < \varepsilon. \end{aligned}$$

For $x \in U_0$ we have $\|f(x)\| < \varepsilon$ and therefore

$$\begin{aligned} \|\widetilde{f}(x) - f(x)\| &= \|\sum_{y \in Y \cap U_x} \chi_y(x)\big(f(y) - f(x)\big) - \chi_0(x) f(x)\| \\ &\le \sum_{y \in Y \cap U_x} \chi_y(x)\|f(y) - f(x)\| + \chi_0(x)\|f(x)\| < \varepsilon. \end{aligned}$$

(ii) $\Longrightarrow$ *(iii)* is easy: Given a continuous function $f : E \to [0,1]$ with support in a bounded set $V \subset E$, pick a coordinate chart $\phi : U \to V$ from an open subset $U \subset X$ onto V, extend the function $f \circ \phi : U \to \mathbb{R}$ by zero outside U, and apply the smooth approximation property on X.

(iii) $\Longrightarrow$ *(i)*: Let $\mathcal{U} = (U_i)_{i \in I}$ be a given open cover of X. Suppose that we can construct a smooth partition of unity $(\kappa_j)_{j \in J}$ subordinate to some open refinement $(V_j)_{j \in J}$ of $\mathcal{U}$. Let $\sigma : J \to I$ be a map such that $V_j \subset U_{\sigma(i)}$ for all $j \in J$. Then $\chi_i := \sum_{j \in \sigma^{-1}(i)} \kappa_j$ is a partition of unity subordinate to $\mathcal{U}$. So it suffices to construct a smooth partition of unity for some refinement of $\mathcal{U}$. Let

$\mathcal{V} = (V_j)_{j\in J}$ be a refinement of $\mathcal{U}$ consisting of domains of charts $\phi_j : V_j \to \phi_j(V_j)$ onto bounded open subsets of E, and let $\mathcal{W}$ be a locally finite refinement of $\mathcal{V}$ provided by paracompactness of X. After renaming $\mathcal{W}$ back to $\mathcal{U}$, we may therefore assume without loss of generality that $\mathcal{U}$ is locally finite and the U_i are domains of charts $\phi_i : U_i \to \phi_i(U_i)$ onto bounded open subsets of E.

By Lemma 2.12 there exists a locally finite open cover $(V_i)_{i\in I}$ with $\overline{V}_i \subset U_i$ for all $i \in I$. Urysohn's lemma provides continuous functions $\rho_i : X \to [0,1]$ with $\rho_i = 1$ on $\overline{V}_i$ and $\rho_i = 0$ outside U_i. Then $\rho_i \circ \phi_i^{-1}$ extends by zero outside $\phi_i(U_i)$ to a continuous function $f_i : E \to [0,1]$ with bounded support. Condition (iii) with $\varepsilon = 1$ yields a smooth function $\widetilde{f}_i : E \to [0,1]$ with $\operatorname{supp} \widetilde{f}_i \subset \operatorname{supp} f_i \subset \phi_i(U_i)$ and $\|\widetilde{f}_i(y) - f_i(y)\| < 1$ for all $y \in E$. Then $\widetilde{f}_i \circ \phi_i$ extends by zero outside U_i to a continuous function $\widetilde{\rho}_i : X \to [0,1]$ with support in U_i and $\|\widetilde{\rho}_i(x) - \rho_i(x)\| < 1$ for all $x \in X$, so $\widetilde{\rho}_i > 0$ on $\overline{V}_i$. By local finiteness of $\mathcal{U}$ it follows that $\sum_{j\in I} \widetilde{\rho}_j$ defines a positive smooth function on X and $\chi_i := \widetilde{\rho}_i / (\sum_{j\in I} \widetilde{\rho}_j)$ is a smooth partition of unity subordinate to $\mathcal{U}$. □

More generally, given a convex subset $C \subset F$ of a Banach space containing 0, we say that a Banach manifold X has the *smooth approximation property with values in* C if for every continuous map $f : X \to C$ and every $\varepsilon > 0$ there exists a smooth map $\widetilde{f} : X \to C$ such that $\operatorname{supp} \widetilde{f} \subset \operatorname{supp} f$ and $\|\widetilde{f}(x) - f(x)\| < \varepsilon$ for all $x \in X$.

Corollary 2.14. *For a Banach manifold X, the smooth approximation properties with values in the following sets C are equivalent:*
(i) $C = [0,1]$,
(ii) C any convex subset of a Banach space containing 0,

(iii) $C = \mathbb{R}$.

Proof. The implication $(i) \implies (ii)$ was shown in the proof of Proposition 2.13, and $(ii) \implies (iii)$ is trivial. For $(iii) \implies (i)$ let a continuous function $f : X \to [0,1]$ and $\varepsilon > 0$ be given. By (iii) there exists a smooth function $g : X \to \mathbb{R}$ with $\operatorname{supp} g \subset \operatorname{supp} f$ and $\|g(x) - f(x)\| < \varepsilon/2$ for all $x \in X$. Pick a smooth function $\tau : \mathbb{R} \to [0,1]$ with $\tau(y) = 0$ for $y \leq 0$, $\tau(y) = 1$ for $y \geq 0$, and $|\tau(y) - y| < \varepsilon/2$ for all $y \in (-\varepsilon/2, 1 + \varepsilon/2)$. Then the smooth function $\widetilde{f} := \tau \circ g : X \to [0,1]$ satisfies $\operatorname{supp} \widetilde{f} \subset \operatorname{supp} g \subset \operatorname{supp} f$, and since g takes values in $(-\varepsilon/2, 1+\varepsilon/2)$, for every $x \in X$ we get

$$|\widetilde{f}(x) - f(x)| \leq |\tau\big(g(x)\big) - g(x)| + |g(x) - f(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

Proposition 2.13 reduces the existence of smooth partitions of unity on a Banach manifold X modelled over E to the existence of smooth partitions of unity on the Banach space E itself. Clearly, a necesary condition for the existence of smooth partitions of unity on E is the existence of a smooth bump function. It is a major open problem in Banach space theory whether this is also sufficient. An affirmative answer is known for a restricted class of spaces:

Theorem 2.15. *If a Banach space E is separable or reflexive, then the existence of a smooth bump function on E implies the existence of smooth partitions of unity subordinate to every open cover of E.*

Remark 2.16. More generally, this holds if E is *weakly compatly generated* in the sense that there exits a weakly compact set $K \subset E$ with $\overline{\operatorname{span}(K)} = E$. It also holds for other regularities C^k, $k \in \mathbb{N}_0 \cup \{\infty\}$.

Combining Theorem 2.15 with Theorem 2.6 and Corollary 2.7 we obtain

Corollary 2.17. *If a Banach space E is either reflexive, or both E and E^* are separable, then E (and thus every Banach manifold modelled over X) admits smooth partitions of unity subordinate to every open cover. In particular, this holds if E is a Hilbert space, or a Sobolev space $W^{k,p}(\Omega)$ with $k \in \mathbb{N}_0$ and $1 < p < \infty$.*

Second countability and separability. A topological space is called *second countable* if it has a countable basis of the topology.

Problem 2.7. Show that second countability implies separability, and for a metric space the two notions are equivalent. Moreover, every subset of a second countable space is again second countable.

Hence for a Banach manifold second countability is equivalent to separability. Moreover, if a Banach manifold modelled over E is separable, then so is the Banach space E. We will encounter below some results that require Banach manifolds to be separable.

2.3 Banach manifolds of maps

The main examples of infinite dimensional Banach manifolds are spaces of maps between manifolds of some given Sovolev or Hölder class. In this section we sketch the construction of smooth atlases on such spaces. The main reference is Eliasson's article [16].

For this section we fix a compact n-dimensional manifold N (possibly with boundary) and an m-dimensional manifold M (without boundary but possibly noncompact).

The Banach manifold of continuous maps. We begin by constructing the structure of a smooth Banach manifold on the space $C(N, M)$ of continuous maps from N to M. For this, we pick a Riemannian metric $\langle\ ,\ \rangle$ on M. It induces a metric d_M on M by defining $d_M(p, p')$ as the infimum of all lengths of smooth paths from p to p'. We equip $C(N, M)$ with the metric $d(u, v) := \max_{q \in N} d_M\big(u(q), v(q)\big)$. The resulting topology does not depend on the Riemannian metric, and it is Hausdorff and paracompact because it is metrizable.

The Riemannian metric also induces for each $p \in M$ an exponential map $\exp_p : \mathcal{D}_p \to M$, which is defined on an open neighbourhood $\mathcal{D}_p$ of the origin in the tangent space T_pM and sends $x \in \mathcal{D}_p$ to $\gamma(1)$ for the geodesic γ with $\gamma(0) = p$ and $\dot{\gamma}(0) = x$. The exponential maps at all $p \in M$ fit together to a smooth map

$$\exp : \mathcal{D} \to M \times M, \qquad (p, x) \mapsto (p, \exp_p x)$$

defined on an open neighbourhood $\mathcal{D} \subset TM$ of the zero section.

Problem 2.8. Show that for $\mathcal{D}$ sufficiently small, the map $\exp : \mathcal{D} \to M \times M$ is a diffeomorphism onto an open neighbourhood $\exp(\mathcal{D})$ of the diagonal. Moreover, we can arrange $\mathcal{D}$ such that $\exp(\mathcal{D})$ is invariant under the diffeomorphism $\tau(p, q) = (q, p)$ of $M \times M$.

We fix a neighbourhood $\mathcal{D}$ as in the preceding problem and denote $\mathcal{D}_p = \mathcal{D} \cap T_pM$ for $p \in M$.

Now we construct a chart around a given map $u \in C(N, M)$. The pullback bundle $u^*TM \to N$ (with fibre $T_{u(q)}M$ over $q \in N$) is a continuous vector bundle with a fibrewise inner product induced by the Riemannian metric on M. So the space

$$E_u := C(N, u^*TM)$$

of continuous sections in u^*TM becomes a Banach space with the norm $\|\xi\| := \max_{q\in N}|\xi(q)|$. The set

$$u^*\mathcal{D} := \{(q,x) \mid \big(u(q),x\big) \in \mathcal{D}\} \subset u^*TM$$

is an open neighbourhood of the zero section, so

$$V_u := C(N,u^*\mathcal{D}) = \{\xi \mid \big(u(q),\xi(q)\big) \in \mathcal{D} \text{ for all } q \in N\} \subset E_u$$

is an open neighbourhood of the origin. The exponential map induces a continuous map

$$\exp_u : V_u \to C(N,M), \qquad (\exp_u \xi)(q) := \exp_{u(q)} \xi(q)$$

which is a homeomorphism onto its image

$$U_u := \{v \in C(N,M) \mid \big(u(q),v(q)\big) \in \exp(\mathcal{D}) \text{ for all } q \in N\}.$$

The invariance of $\exp(\mathcal{D})$ under swapping the factors in $M \times M$ implies that $v \in U_u$ is equivalent to $u \in U_v$.

Proposition 2.18. *(a) The charts $\exp_u^{-1} : U_u \to V_u$ for $u \in C(N,M)$ define a smooth Banach atlas on $C(N,M)$.*
(b) The atlases induced by different Riemannian metrics on M are equivalent, so the resulting Banach manifold structure is independent of the choice of Riemannian metric.
(c) The Banach manifold $C(N,M)$ is covered by the chart domains U_u centered at smooth maps $u \in C^\infty(N,M)$.

Proof. (a) For $u,v \in C(N,M)$ with $v \in U_u$ the transition map

$$\Psi = \exp_v^{-1} \circ \exp_u : \exp_u^{-1}(U_u \cap U_v) \to E_v$$

can be explicitly described as follows. Its domain is

$$\begin{aligned}&\exp_u^{-1}(U_u \cap U_v)\\ &= \{\xi \in C(N,u^*\mathcal{D}) \mid \big(v(q),\exp_{u(q)}\xi(q)\big) \in \exp(\mathcal{D}) \text{ for all } q \in N\}\\ &= C(N,\mathcal{O})\end{aligned}$$

with the open neighbourhood of the zero section

$$\mathcal{O} := \{(q,x) \in u^*\mathcal{D} \mid \big(v(q), \exp_{u(q)} x\big) \in \exp(\mathcal{D})\} \subset u^*TM. \tag{2.2}$$

The transition map

$$\Psi(\xi)(q) = \exp_{v(q)}^{-1} \circ \exp_{u(q)} \xi(q) = \psi\big(q, \xi(q)\big)$$

can be written as the composition

$$\Psi : C(N, \mathcal{O}) \to C(N, v^*TM), \qquad \xi \mapsto \psi \circ \xi$$

with the fibre preserving continuous map

$$\psi : \mathcal{O} \to v^*TM, \qquad (q,x) \mapsto \big(q, \psi(q,x)\big), \tag{2.3}$$
$$\psi(q,x) := \exp_{v(q)}^{-1} \circ \exp_{u(q)}(x).$$

The Fréchet derivative of Ψ applied to $\eta \in C(N, u^*TM)$ is

$$\Big(D\Psi(\xi) \cdot \eta\Big)(q) = D_2\psi\big(q, \xi(q)\big) \cdot \eta(q),$$

where $D_2\psi$ denoted the derivative of $\psi(q,x)$ with respect to the second variable x. Note that although $\psi(q,x)$ is only continuous in q (because u, v are only continuous), it is smooth in x for fixed q, so this derivative exists. By definition, this is the *fibre derivative* of the fibre preserving map $\psi : \mathcal{O} \to v^*TM$ So we have shown that the Fréchet derivative of Ψ exists and is given by $(D\Psi)(\xi) = (D_2\psi) \circ \xi$. Since the fibre derivative is again a fibre preserving map $D_2\psi : \mathcal{O} \to \mathrm{Hom}(u^*TM, v^*TM)$, we can iterate this process. Hence Ψ is smooth and its k-th Fréchet derivative is related to the k-th fibre derivative of ψ by

$$(D^k\Psi)(\xi) = (D_2^k\psi) \circ \xi.$$

This proves part (a). Part (b) is proved similarly.
For part (c), consider $v \in C(N, M)$ and the domain $U_v \subset C(N, M)$

defined above. By the Stone-Weierstrass theorem (applied in coordinate charts after multiplying with a partition of unity) the smooth maps are dense in $C(N, M)$, so there exists $u \in C^\infty(N, M) \cap U_v$. As observed above, $u \in U_v$ implies $v \in U_u$, so v is contained in the chart domain U_u centered at the smooth map u. □

Problem 2.9. (a) Prove part (b) of Proposition 2.18.
(b) Adapt the proof of Proposition 2.18 to show that for each $k \in \mathbb{N}_0$ the space $C^k(N, M)$ of C^k-maps from N to M is a Banach manifold.

Other Banach manifolds of maps. We wish to extend Proposition 2.18 to spaces of maps $\mathfrak{S}(N, M)$ from N to M of some other regularity, e.g. of a given Hölder or Sobolev class. Following [16], we fix N and consider a functor

$$\mathfrak{S} : VB(N) \to \mathcal{B}$$

from the category of smooth vector bundles over N to the category of *Banachable spaces*, i.e. Banach spaces up to equivalence of norms (or equivalently, topological vector spaces whose topology is induced by a complete norm). Thus $\mathfrak{S}$ assigns to each smooth vector bundle $E \to N$ a Banach space $\mathfrak{S}(E)$ and to each smooth bundle homomorphism $\phi : E \to F$ a continuous linear map $\mathfrak{S}(\phi) : \mathfrak{S}(E) \to \mathfrak{S}(F)$. We call $\mathfrak{S}$ a *section functor over* N if elements of $\mathfrak{S}(E)$ can be viewed as sections of E (not necessarily continuous and only defined up to a set of measure zero), $\mathfrak{S}(\phi)(\xi)$ is the composition $\phi \circ \xi$, and the map

$$\mathfrak{S} : C^\infty\big(Hom(E, F)\big) \to \mathcal{L}\big(\mathfrak{S}(E), \mathfrak{S}(F)\big), \qquad \phi \mapsto \mathfrak{S}(\phi)$$

is continuous linear.

Definition 2.2. A section functor $\mathfrak{S}$ over N is called a *manifold model* if in addition for all $E, F \in VB(N)$ the following holds:

(i) There exists a continuous linear inclusion $\mathfrak{S}(E) \hookrightarrow C(N,E)$.
(ii) There exists a continuous linear inclusion $\mathfrak{S}\big(\mathrm{Hom}(E,F)\big) \hookrightarrow \mathcal{L}\big(\mathfrak{S}(E),\mathfrak{S}(F)\big)$.
(iii) Let $\mathcal{O} \subset E$ be an open subset projecting onto N and $\psi : \mathcal{O} \to F$ be a smooth fibre preserving map. Then for each $\xi \in \mathfrak{S}(\mathcal{O}) := \{\xi \in \mathfrak{S}(E) \mid \xi(N) \in \mathcal{O}\}$ we have $\psi \circ \xi \in \mathfrak{S}(F)$ and the corresponding map

$$\mathfrak{S}(\psi) : \mathfrak{S}(\mathcal{O}) \to \mathfrak{S}(F), \qquad \xi \mapsto \psi \circ \xi$$

is continuous.

Properties (i)–(iii) are extracted from the proof of Proposition 2.18 where they were implicitly used. To see this, restrict to the case of smooth u and replace $C(N,\cdot)$ by $\mathfrak{S}$. Then property (i) is needed for the set $\mathfrak{S}(\mathcal{O})$ in (iii) to be well-defined and open, (iii) gives continuity of the transition map $\Psi = \mathfrak{S}(\psi)$, and (ii) is needed to equate $\mathfrak{S}(D_2\psi) \in \mathfrak{S}\big(\mathrm{Hom}(E,F)\big)$ to the Fréchet derivative $D\big(\mathfrak{S}(\psi)\big) \in \mathcal{L}\big(\mathfrak{S}(E),\mathfrak{S}(F)\big)$.
Conversely, every manifold model $\mathfrak{S}$ and manifold M give rise to a Banach manifold $\mathfrak{S}(N,M)$ as follows. For each smooth map $u : N \to M$ consider the inverse chart

$$\exp_u : C(N,u^*TM) \supset V_u \xrightarrow{\approx} U_u \subset C(N,M)$$

for $C(N,M)$ constructed above. Using property (i) in Definition 2.2 we set

$$\mathfrak{S}(V_u) := V_u \cap \mathfrak{S}(u^*TM), \qquad \mathfrak{S}(U_u) := \exp_u\big(\mathfrak{S}(V_u)\big)$$

and the bijection

$$\mathfrak{S}(\exp_u) := \exp_u|_{V_u} : \mathfrak{S}(V_u) \to \mathfrak{S}(U_u).$$

with inverse $\mathfrak{S}(\exp_u^{-1}) = \exp_u^{-1}|_{U_u} : \mathfrak{S}(U_u) \to \mathfrak{S}(V_u)$. Now we define the set

$$\mathfrak{S}(N,M) := \bigcup_{u\in C^\infty(N,M)} \mathfrak{S}(U_u) \subset C(N,M).$$

By definition, it is covered by the chart domains $\mathfrak{S}(U_u)$ for $u \in C^\infty(N,M)$. For $u,v \in C^\infty(N,M)$ with $\mathfrak{S}(U_u)\cap\mathfrak{S}(U_v) \neq \emptyset$ we have

$$\mathfrak{S}(\exp_u^{-1})\big(\mathfrak{S}(U_u)\cap\mathfrak{S}(U_v)\big) = \mathfrak{S}(u^*TM)\cap C(N,\mathcal{O}) = \mathfrak{S}(\mathcal{O}),$$

with the open neighbourhood of the zero section $\mathcal{O} \subset u^*TM$ defined in (2.2). Hence this set is open by property (i) in Definition 2.2, and according to Remark 2.8 the charts $\mathfrak{S}(\exp_u^{-1})$ induce a topology on $\mathfrak{S}(N,M)$. Again by property (i), the inclusion $\mathfrak{S}(N,M) \subset C(N,M)$ is continuous, so $\mathfrak{S}(N,M)$ is Hausdorff because $C(N,M)$ is. For u,v as above the transition map is the composition

$$\mathfrak{S}(\psi) : \mathfrak{S}(\mathcal{O}) \to \mathfrak{S}(v^*TM), \qquad \xi \mapsto \psi\circ\xi$$

with the smooth fibre preserving map $\psi : \mathcal{O} \to v^*TM$ from (2.3). By property (iii) in Definition 2.2 the map $\mathfrak{S}(\psi)$ is continuous, and by Problem 2.10 below it is smooth with k-th derivative $D^k\mathfrak{S}(\psi) = \mathfrak{S}(D_2^k\psi)$. So we have shown most of

Theorem 2.19 (Eliasson [16])**.** *Let $\mathfrak{S}$ be a manifold model over N and let M be a manifold. Then the charts $\mathfrak{S}(\exp_u^{-1}) : \mathfrak{S}(U_u) \to \mathfrak{S}(V_u)$ for $u \in C^\infty(N,M)$ define on $\mathfrak{S}(N,M)$ the structure of a smooth Banach manifold, which is independent of the choice of Riemannian metric. If the Banach space $\mathfrak{S}(N,\mathbb{R})$ is separable, then so is the Banach manifold $\mathfrak{S}(N,M)$ for each M.*

Proof. It remains to show metrizability and the last assertion. Use the Whitney embedding theorem to embed M as a closed submanifold into some $\mathbb{R}^\ell$. Then $\mathfrak{S}(N, M)$ is a closed subset of the Banach space $\mathfrak{S}(N, \mathbb{R}^\ell)$ with the induced subset topology (since the chart $\mathfrak{S}(\exp_u)$ is a homeomorphism with respect to the subspace topology on $\mathfrak{S}(N, M)$). Hence $\mathfrak{S}(N, M)$ is metrizable because $\mathfrak{S}(N, \mathbb{R}^\ell)$ is, and second countable (which by Problem 2.7 is equivalent to separability) whenever $\mathfrak{S}(N, \mathbb{R}^\ell)$ is. □

Problem 2.10 (Lemma 4.1 in [16]). Let $\mathfrak{S}$ be a manifold model over N and $\psi : \mathcal{O} \to F$ be as in property (iii) of Definition 2.2. Then $\mathfrak{S}(\psi) : \mathfrak{S}(\mathcal{O}) \to \mathfrak{S}(F)$ is smooth with k-th derivative $D^k\mathfrak{S}(\psi) = \mathfrak{S}(D_2^k\psi)$. *Hint: Use induction on k, and for $k = 1$ apply the functor $\mathfrak{S}$ to the linear Taylor expansion of ψ.*

Corollary 2.20. *Let N be a compact n-dimensional manifold (possibly with boundary) and M be an m-dimensional manifold (without boundary but possibly noncompact). Then the following spaces of maps $N \to M$ are separable smooth Banach manifolds:*
(a) the space $C^{k,\alpha}(N, M)$ of Hölder maps for $k \in \mathbb{N}_0$ and $\alpha \in [0, 1]$;
(b) the space $W^{k,p}(N, M)$ of Sobolev maps for $k \in \mathbb{N}$ and $1 \le p < \infty$ satisfying $k - n/p > 0$.

Sketch of proof. In view of Theorem 2.19 and Problem 1.5, it suffices to show that $C^{k,\alpha}$ and $W^{k,p}$ are manifold models over N. For the Hölder functors $C^{k,\alpha}$ this is Problem 2.11 below. For the Sobolev functors $W^{k,p}$, property (i) in Definition 2.2 follows from the Sobolev Embedding Theorem 1.23, property (ii) from the fact (Theorem 1.24) that $W^{k,p}(N, \mathbb{R})$ is a Banach algebra (both properties use the same condition $k - n/p > 0$), and property (iii) from the combination of both. See [39] for details.

□

Problem 2.11. Show that for all $k \in \mathbb{N}_0$ and $\alpha \in [0,1]$, the Hölder functor $C^{k,\alpha}$ is a manifold model over any compact manifold N.

2.4 Some geometry on Banach manifolds

Most geometric notions carry over from finite dimensional manifolds to Banach manifolds in a straightforward way: tangent spaces, submanifolds, transversality, Riemannian metrics, vector fields and flows, differential forms,.... The main difference to the finite dimensional case is that, as a general rule, one should assume that all linear maps are continuous and all linear subspaces are closed and have a complement. In this section we review the most important notions, for further discussion see [15].

Tangent bundle. For a smooth map $f : U \to V$ between open sets $U \subset E$ and $V \subset F$ of Banach spaces, its differential gives rise to a smooth map

$$Tf : TU := U \times E \to TV := V \times F, \qquad (x,\xi) \mapsto \big(f(x), Df(x)\xi\big).$$

Let now X be a Banach manifold modelled on E and pick a cover $X = \bigcup_{i\in I} U_i$ by domains of charts $\phi_i : U_i \xrightarrow{\cong} V_i \subset E$. We abbreviate

$$\phi_{ij} := \phi_i \circ \phi_j^{-1} : \phi_j(U_{ij}) \to \phi_i(U_{ij}), \qquad U_{ij} := U_i \cap U_j.$$

Then the charts induce a canonical diffeomorphism

$$X \cong \coprod_{i\in I} V_i \Big/ x \sim \phi_{ij}(x) \text{ for all } x \in \phi_j(U_{ij}).$$

The *tangent bundle* of X is

$$\begin{aligned} TX &:= \coprod_{i \in I} TV_i \Big/ y \sim T\phi_{ij}(y) \text{ for all } y \in T\big(\phi_j(U_{ij})\big) \\ &= \coprod_{i \in I} V_i \times E \Big/ (x, \xi) \sim \big(\phi_{ij}(x), D\phi_{ij}(x)\xi\big),\ (x, \xi) \in \phi_j(U_{ij}) \times E. \end{aligned}$$

The maps $(x, \xi) \mapsto x$ induce a canonical projection $\tau : TX \to X$. The fibre $T_x X := \tau^{-1}(x)$ is called the *tangent space* to X at x.

Next, consider a smooth map $f : X \to Y$ between Banach manifolds. Pick atlases $\phi_i : U_i \xrightarrow{\cong} \phi_i(U_i) \subset E$ for X and $\psi_k : W_k \xrightarrow{\cong} \psi_k(W_k) \subset F$ for Y. Then for all i, k we have smooth maps

$$f_{ki} := \psi_k \circ f \circ \phi_i^{-1} : \phi_i\big(U_i \cap f^{-1}(W_k)\big) \to \psi_k(W_k).$$

It is shown in the following problem that their tangent maps

$$Tf_{ki} : \phi_i\big(U_i \cap f^{-1}(W_k)\big) \times E \to \psi_k(W_k) \times F$$

induce a smooth map

$$Tf : TX \to TY$$

which restricts on each tangent space to a continuous linear map

$$T_x f : T_x X \to T_{f(x)} Y.$$

The map Tf is called the *tangent map* (or *differential*) of f.

Problem 2.12. Let X be a Banach manifold modelled on E. Show:
(a) The tangent bundle TX is a Banach manifold modelled on $E \oplus E$, which does not depend on the chosen atlas up to canonical diffeomorphism.
(b) The canonical projection $\tau : TX \to X$ is a smooth map.
(c) Each tangent space $T_x X$ carries the canonical structure of a Banachable space.

(d) For a smooth map $f : X \to Y$ between Banach manifolds, the tangent map $Tf : TX \to TY$ is well-defined and smooth, and it restricts on each tangent space to a continuous linear map $T_x f : T_x X \to T_{f(x)} Y$.

Submanifolds. Let E, F be Banach spaces. We call a closed linear subspace $V \subset E$ *complemented* if it has a complement W uch that $E = V \oplus W$. A *right inverse* of $T \in \mathcal{L}(E, F)$ is a map $R \in \mathcal{L}(F, E)$ such that $T \circ R = \mathbb{1}$. Similarly, A *left inverse* of $T \in \mathcal{L}(E, F)$ is a map $R \in \mathcal{L}(F, E)$ such that $L \circ T = \mathbb{1}$.

Problem 2.13. (a) $T \in \mathcal{L}(E, F)$ has a right inverse iff it is surjective and $\ker T$ is complemented.
(b) $T \in \mathcal{L}(E, F)$ has a left inverse iff it is injecive and $\operatorname{im} T$ is closed and complemented.
(c) A Fredholm operator $T \in \mathcal{F}(E, F)$ has a right (resp, left) inverse iff it is surjective (resp. injective).

Let now X be a Banach manifold modelled over E. A subset $A \subset X$ is called a *submanifold* if each $a \in A$ has an open neighbourhood $U \subset X$ and a chart $\phi : U \to \phi(U) \subset E$ such that $\phi(U \cap A) = \phi(U) \cap F$ for a *complemented* closed linear subspace $F \subset E$. We call A *closed* if it is closed as a subset of X.[1]

Problem 2.14. (a) Show that a submanifold of a Banach manifold is again a Banach manifold.
(b) Let $\mathfrak{S}$ be a manifold model over a compact manifold N (possibly with boundary). Let $M \subset P$ be a closed submanifold of a finite dimensional manifold P. Prove that $\mathfrak{S}(N, M)$ is a closed submanifold of the Banach manifold $\mathfrak{S}(N, P)$. *Hint: Construct a metric on P for which M is* totally geodesic, *i.e., each geodesic*

[1] In differential geometry a "closed manifold" often denotes a compact finite dimensional manifold without boundary; we will not use the term "closed" in this sense.

starting at a point of M in a direction tangent to M remains on M for all times.

Remark 2.21. By the (easy) *Whitney embedding theorem*, every n-dimensional manifold is diffeomorphic to a closed submanifold of $\mathbb{R}^{2n+1}$. It generalizes in infinite dimensions to *McAlpin's embedding theorem* [33] that every separable Hilbert manifold is diffeomorphic to a closed submanifold of some separable Hilbert space.

Regular and critical values. Let $f : X \to Y$ be a smooth map between Banach manifolds. A point $x \in X$ is called *regular* if the differential $T_x f : T_x X \to T_{f(x)} Y$ has a right inverse (or equivalently, $T_x f$ is surjective and $\ker T_x f$ is complemented). A *critical point* is one which is not regular. A *regular value* is $y \in Y$ such that all $x \in f^{-1}(y)$ are regular, and a *critical value* is one which is not regular.

More generally, let $B \subset Y$ be a submanifold. We say that f *is transverse to* B if for each $x \in f^{-1}(B)$ the composition of $T_x f : T_x X \to T_y Y$ (where $y = f(x) \in B$) with the projection $T_y Y \to T_y Y / T_y B$ has a right inverse. The implicit function theorem implies

Theorem 2.22. *Let $f : X \to Y$ be a smooth map between Banach manifolds which is transverse to a submanifold $B \subset Y$. Then $f^{-1}(B)$ is a submanifold of X.*

Proof. Consider $x \in X$ with $y = f(x) \in B$. Since B is a submanifold, using local charts we may assume that $X = E$, $Y = F_1 \oplus F_2$, $B = F_1 \oplus 0$, $x = 0$ and $y = (0,0)$. Denote by $f_2 : E \to F_1 \oplus F_2 \to F_2$ the second component of f. By hypothesis $Df_2(0)$ is surjective and $E_1 := \ker Df_2(0)$ has a complement E_2 such that $E = E_1 \oplus E_2$. Thus the partial derivative with respect to the second variable $D_2 f_2(0,0) : E_2 \to F_2$ is an isomorphism. By the implicit function theorem, the set $f^{-1}(B) = f_2^{-1}(0)$ is locally near $x = (0,0)$ the graph $\{x_2 = g(x_1)\}$ of a smooth function $g : E_1 \supset U_1 \to E_2$. By the inverse function theorem, the map $(x_1, x_2) \mapsto \big(x_1, x_2 - g(x_1)\big)$ is a local diffeomorphism sending the graph of g to $E_1 \oplus 0$, so the graph of g is a submanifold near $(0,0)$. □

Applying this to $B = \{y\}$ consisting of a single point we obtain

Corollary 2.23. *Let $f : X \to Y$ be a smooth map between Banach manifolds and $y \in Y$ be a regular value. Then $f^{-1}(y)$ is a closed submanifold of X.* □

2.5 Local normal forms of maps

The following lemma standardizes a smooth map near an arbitrary point. It will be particularly useful for the study of Fredholm maps in the next chapter.

Lemma 2.24 (Local normal form)**.** *Let $f : X \to Y$ be a smooth map between Banach manifolds and $x \in X$. Suppose that the kernel and image of $T_x f$ are closed and complemented. Then there exist charts $\phi : X \supset U \hookrightarrow E = P \oplus Q$ sending x to 0 and $\psi : Y \supset V \hookrightarrow F = P \oplus R$ sending $f(x)$ to 0 such that*

$$\psi \circ f \circ \phi^{-1}(p, q) = \big(p, g(p, q)\big)$$

with a smooth function g satisfying $g(0) = 0$ and $Dg(0) = 0$.

Proof. After composing with charts sending x and $f(x)$ to 0 it suffices to consider $f : U \to F$ with $f(0) = 0$ for Banach spaces E, F and an open neighbourhood $U \subset E$ of 0. By hypothesis there are splittings $E = P \oplus \ker Df(0)$ and $F = \operatorname{im} Df(0) \oplus R$ with respect to which $Df(0) = \left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$ for an isomorphism $A : P \stackrel{\cong}{\to} \operatorname{im} Df(0)$. After composing f with the linear map $\left(\begin{smallmatrix} A^{-1} & 0 \\ 0 & \mathbb{1} \end{smallmatrix}\right)$ we may assume that $F = P \oplus R$ and $Df(0) = \left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Set $Q := \ker Df(0)$, write $x = (p, q)$ and $f = (f_1, f_2)$, and define $\phi : U \to E$ by $\phi(p, q) := \big(f_1(x), q\big)$. Since $D\phi(0) = \mathbb{1}$, by the inverse function theorem $\phi : U_1 \to U_2 \subset U$ is a diffeomorphism between open neighbourhoods of 0 in E. Let $\widetilde{f} := f \circ \phi^{-1} : U_2 \to F$. Then $\widetilde{f}(x) = f(x') = (y_1, y_2)$ with $x = \phi(x') = \big(f_1(x'), q'\big)$ implies $y_1 = f_1(x') = p$. So $\widetilde{f}$ has the form $\widetilde{f}(p, q) = \big(p, g(x)\big)$ with a smooth function g, and it follows that $g(0) = 0$ and $Dg(0) = 0$. □

The next result provides conditions under which a smooth map near a point agrees, after composition with diffeomorphisms from both sides, with its linearization at this point.

Theorem 2.25 (Constant rank theorem). *Let E, F be Banach spaces, $U \subset E$ an open neighbourhood of 0, and $f : U \to F$ be a smooth map with $f(0) = 0$. Suppose that we have splittings $E = P \oplus \ker Df(0)$ and $F = \operatorname{im} Df(0) \oplus R$. Suppose further that $\operatorname{im} Df(x) = \operatorname{im} \beta(x)$ for all $x \in U$, with a continuous map $\beta : U \to \mathcal{L}(P, F)$ satisfying $\beta(0) = Df(0)|_P$. Then there exist diffeomorphisms $\phi : U \supset U_1 \to U_2$ and $\psi : V_1 \to V_2$ between open neighbourhoods of 0 such that*

$$\psi \circ f \circ \phi^{-1}(p, q) = (p, 0).$$

Proof. (cf. [46, Lecture 1]). Set $Q := \ker Df(0)$. After applying Lemma 2.24 we may assume that $P = \operatorname{im} Df(0)$ and $f(p, q) = \big(p, g(p, q)\big)$ with a smooth function $g : P \oplus Q \supset U_2 \to R$ satisfying $g(0) = 0$ and $Dg(0) = 0$. Then

$$Df(x) = \begin{pmatrix} \mathbb{1} & 0 \\ D_1 g(x) & D_2 g(x) \end{pmatrix}$$

By hypothesis, $\operatorname{im} Df(x) = \operatorname{im}\beta(x)$ for a continuous map $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} : U_2 \to \mathcal{L}(P, F)$ with $\beta(0) = \left(\begin{smallmatrix} \mathbb{1} \\ 0 \end{smallmatrix}\right)$. For each $w \in Q$ we then have

$$\begin{pmatrix} \mathbb{1} & 0 \\ D_1 g(x) & D_2 g(x) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ D_2 g(x) w \end{pmatrix} \in \operatorname{im}\beta(x),$$

so there must exist $v \in P$ such that

$$\begin{pmatrix} 0 \\ D_2 g(x) w \end{pmatrix} = \begin{pmatrix} \beta_1(x) v \\ \beta_2(x) v \end{pmatrix}.$$

Since $\beta_1(0) = \mathbb{1}$, after shrinking U_2 we may assume that $\beta_1(x)$ is invertible for all $x \in U_2$. So the previous equation implies $v = 0$, and therefore $D_2 g(x) w = 0$. This shows that $D_2 g(x) = 0$ for all $x \in U_2$. Choosing U_2 convex, we conclude that

$$g(p, q) - g(p, 0) = \int_0^1 D_2 g(p, tq) \cdot q \, dt = 0$$

for all $x = (p, q) \in U_2$, so $g(p, q)$ does not depend on q and f has the form $f(p, q) = \big(p, g(p)\big)$. This looks like a graph and can be flattened by the map $\psi(p, r) := \big(p, r - g(p)\big)$ defined on a neighbourhood of 0 in F. Since $D\psi(0) = \mathbb{1}$, by the inverse function theorem $\psi : V_1 \to V_2$ is a diffeomorphism between open neighbourhoods of 0 in F. Then $\psi \circ f : U_2 \to F$ (after possibly shrinking U_2) has the form

$$\psi \circ f(p, q) = \psi\big(p, g(p)\big) = \big(p, g(p) - g(p)\big) = (p, 0).$$

□

Problem 2.15. Show that for $E = R^m$ and $F = \mathbb{R}^n$ finite dimensional, the hypothesis on the continuous map β in the constant rank theorem is equivalent to the rank of $Df(x)$ being constant for $x \in U$. (This explains the name "constant rank theorem".)

Let us discuss some consequences of the constant rank theorem. A *foliation* of a Banach manifold X is a decomposition of X into a disjoint union of closed submanifolds (called the *leaves*) such that near each point of X there exists a chart $X \supset U \to P \oplus Q$ under which the leaves correspond to the linear subspaces $\{p\} \times Q$, $p \in P$. Then Theorem 2.25 can be restated as

Corollary 2.26. *Let $f : X \to Y$ be a smooth map between Banach manifolds. Suppose that for each $x \in X$ the kernel and image of $T_x f : T_x X \to T_{f(x)} Y$ are closed and complemented, and the images of the $T_x f$ form a continuous vector bundle over X. Then the preimages $f^{-1}(y)$, $y \in Y$, are the leaves of a smooth foliation of X. In particular, each $f^{-1}(y)$ is a closed submanifold of X.* □

A smooth map $f : X \to Y$ between Banach manifolds is called a *submersion* if $T_x f$ has a right inverse for all $x \in X$; an *immersion* if $T_x f$ has a right inverse for all $x \in X$; and an *embedding* if it is a diffemorphism onto a submanifold. In these cases the hypotheses of the constant rank theorem are trivially satisfied and we obtain

Corollary 2.27. *Let $f : X \to Y$ be a smooth map between Banach manifolds.*
(a) If f is a submersion, then the preimages $f^{-1}(y)$, $y \in Y$, are the leaves of a smooth foliation of X.
(b) If f is an immersion, then each $x \in X$ has an open neighbourhood $U \subset X$ such that $f|_U$ is an embedding. □

2.6 Retractions and Cartan's last theorem

A *retraction* of a Banach manifold X is a map $r : X \to X$ satisfying $r \circ r = r$. Then its image $\mathrm{im}(r)$ agrees with its *fixed* *point set* $\mathrm{Fix}(r) := \{x \in X \mid r(x) = x\}$. The following surprising result is known as "Cartan's last theorem" or "Cartan's retraction theorem".

Theorem 2.28 (H. Cartan [8])**.** *Let $r : X \to X$ be a smooth retraction of a Banach manifold X. Then its image $\mathrm{im}(r) = \mathrm{Fix}(r)$ is a closed submanifold of X.*

Proof. This "proof from the book" is Cartan's original one. Let $x \in \mathrm{Fix}(r)$. After conjugating with a chart sending x to 0, we can view r as a smooth retraction $r : U \to E$ defined on an open neighbourhood $U \subset E$ of the origin in a Banach space E. After replacing U by $U \cap r^{-1}(U)$ we may assume $r(U) \subset U$. Set $r_1 := Dr(0)$ and define $\alpha, \beta : U \to E$ by

$$\alpha := \beta + r_1 \circ r, \qquad \beta := (\mathbb{1} - r_1) \circ (\mathbb{1} - r).$$

Then $r \circ r = r$ and $r_1 \circ r_1 = r_1$ imply $\beta \circ r = 0$ and $r_1 \circ \beta = 0$ (using linearity of r_1), and therefore

$$\alpha \circ r = \beta \circ r + r_1 \circ r = r_1 \circ r.$$

Again using linearity of r_1, we similarly obtain

$$r_1 \circ \alpha = r_1 \circ \beta + r_1 \circ r = r_1 \circ r,$$

hence

$$\alpha \circ r = r_1 \circ \alpha : U \to E.$$

Since

$$D\alpha(0) = (\mathbb{1} - r_1) \circ (\mathbb{1} - r_1) + r_1 \circ r_1 = \mathbb{1} - 2r_1 + 2r_1 = \mathbb{1},$$

by the inverse function theorem α defines a diffeomorphism $\alpha : V \to W$ between open neighbourhoods of the origin with $V \subset U$.

After replacing V by $V \cap r^{-1}(V)$ we may assume that $r(V) \subset V$. Then $r_1(W) \subset W$ and

$$\alpha \circ r \circ \alpha^{-1} = r_1 : W \to W.$$

Hence α conjugates r to the *linear* retraction r_1. In particular, $\mathrm{Fix}(r) = \alpha^{-1}\big(\mathrm{Fix}(r_1)\big)$ is a submanifold because $\mathrm{Fix}(r_1)$ is the intersection of W with a linear subspace. □

Cartan's last theorem fails for sc-smooth retractions between sc-manifolds. This failure is the starting point for the theory of polyfolds which we will discuss later.

Linear projections. A continuous linear map $P : E \to E$ on a Banach space E is called a *projection* if $P \circ P = P$. For example, the differential of a retraction at a fixed point is a projection.

Problem 2.16. Let $P \in \mathcal{L}(E)$ be a projection. Then $\mathbb{1} - P$ is also a projection,

$$\ker P = \mathrm{im}\,(\mathbb{1} - P), \qquad \mathrm{im}\, P = \ker(\mathbb{1} - P),$$

and E decomposes as the direct sum

$$E = \ker P \oplus \mathrm{im}\, P = \mathrm{im}\,(\mathbb{1} - P) \oplus \ker(\mathbb{1} - P).$$

In particular, $\ker P$ and $\mathrm{im}\, P$ are closed and complemented.

Our goal is to describe how the kernel and image vary over the space of projections. We begin with a more general situation. A continuous linear map $A : E \to F$ between Banach spaces is called a *monomorphism* if it is injective and has closed image. By the open mapping theorem, this is equivalent to A being an isomorphism onto a closed subspace of F. We denote the space of all monomorphisms $E \to F$ by $\mathrm{Mon}(E, F)$.

Lemma 2.29. *Let E, F be Banach spaces and $T \in \mathcal{L}(E, F)$ such that $\ker T$ and $\operatorname{im} T$ are both closed and complemented. Then there exists an open neighbourhood U of T in $\mathcal{L}(E, F)$ and continuous maps*

$$\alpha : U \to \mathrm{Mon}(\ker T, E), \qquad \beta : U \to \mathrm{Mon}(\operatorname{im} T, F)$$

such that $\alpha(T)$ is the inclusion $\ker T \hookrightarrow E$, $\beta(T)$ is the inclusion $\operatorname{im} T \hookrightarrow F$, and for all $S \in U$ we have

$$\ker S \subset \operatorname{im} \alpha(S), \qquad \operatorname{im} \beta(S) \subset \operatorname{im} S.$$

Proof. By hypothesis we have splittings

$$E = E_1 \oplus E_2, \quad F = F_1 \oplus F_2, \quad E_2 = \ker T, \quad F_1 = \operatorname{im} T$$

with respect to which $T = \left(\begin{smallmatrix} A_0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ for an isomorphism $A_0 : E_1 \xrightarrow{\cong} F_1$. Each $S \in \mathcal{L}(E, F)$ has a canonical representation $S = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with respect to these splittings whose coefficients A, B, C, D depend continuously on S. Hence there exists an open neighbourhood $U \subset \mathcal{L}(E, F)$ of T such that for all $S \in U$ the coefficient A is an isomorphism. For such S we have $(x_1, x_2) \in \ker S$ iff $Ax_1 + Bx_2 = 0$ and $Cx_1 + Dx_2 = 0$, or equivalently $x_1 = -A^{-1}Bx_2$ and $(D - CA^{-1}B)x_2 = 0$. Thus $\ker S \subset \operatorname{im} \alpha(S)$ for the graphical monomorphism

$$\alpha(S) : \ker T = E_2 \hookrightarrow E = E_1 \oplus E_2, \qquad x_2 \mapsto (-A^{-1}Bx_2, x_2).$$

Similarly, $(y_1, y_2) \in \operatorname{im} S$ iff there exist (x_1, x_2) such that $Ax_1 + Bx_2 = y_1$ and $Cx_1 + Dx_2 = y_2$. For $x_2 = 0$ this becomes $x_1 = A^{-1}y_1$ and $y_2 = Cx_1 = CA^{-1}y_1$. Thus $\operatorname{im} \beta(S) \subset \operatorname{im} S$ for the graphical monomorphism

$$\beta(S) : \operatorname{im} T = F_1 \hookrightarrow F = F_1 \oplus F_2, \qquad y_1 \mapsto (y_1, CA^{-1}y_1).$$

Since the coefficients A, B, C, D depend continuously on S, so do the monomorphisms $\alpha(S)$ and $\beta(S)$. $\square$

Remark 2.30. (a) In the situation of Lemma 2.29 it follows that

$$\dim\ker S \le \dim\ker T, \qquad \dim\operatorname{coker} S \le \dim\operatorname{coker} T.$$

In particular, this shows that the set of Fredholm operators is open in $\mathcal{L}(E,F)$.
(b) The proof of Lemma 2.29 shows that the maps α, β are in fact smooth as maps $U \to \mathcal{L}(\ker T, E)$ resp. $U \to \mathcal{L}(\operatorname{im} T, F)$.

Now we specialize this to projections. We denote the space of projections $E \to E$ by $\mathcal{P}(E)$.

Corollary 2.31. *Let E be a Banach space and $P \in \mathcal{P}(E)$ a projection. Let $\alpha : U \to \operatorname{Mon}(\ker P, E)$ and $\beta : U \to \operatorname{Mon}(\operatorname{im} P, E)$ be the continuous maps from Lemma 2.29 (with $T = P$ and $F = E$) on an open neighbourhood U of P in $\mathcal{L}(E)$. Then, after possibly shrinking U, for all projections $Q \in U \cap \mathcal{P}(E)$ we have*

$$\ker Q = \operatorname{im}\alpha(Q), \qquad \operatorname{im} Q = \operatorname{im}\beta(Q).$$

Proof. In view of Problem 2.16, the projections P and $\mathbb{1}-P$ satisfy the hypotheses of Lemma 2.29. Let $\alpha' : U' \to \operatorname{Mon}\big(\ker(\mathbb{1}-P), E\big)$ and $\beta' : U' \to \operatorname{Mon}\big(\operatorname{im}(\mathbb{1}-P), E\big)$ be the corresponding maps associated to $\mathbb{1}-P$ defined on a neighbourhood U' of $\mathbb{1}-P$. Then $U'' := \{S \in U \mid \mathbb{1} - S \in U'\}$ is an open neighbourhood of P such that for each $Q \in U'' \cap \mathcal{P}(E)$ we obtain a map $\gamma(Q) \in \mathcal{L}(\operatorname{im} P)$ by the composition

$$\begin{aligned}\operatorname{im} P &\xrightarrow{\beta(Q)} \operatorname{im}\beta(Q) \subset \operatorname{im} Q = \ker(\mathbb{1}-Q)\\ &\hookrightarrow \operatorname{im}\alpha'(\mathbb{1}-Q) \xrightarrow{\alpha'(\mathbb{1}-Q)^{-1}} \ker(\mathbb{1}-P) = \operatorname{im} P.\end{aligned}$$

Since $\gamma(Q)$ depends continuously on Q and $\gamma(P) = \mathbb{1}$, there exists an open neighbourhood $\widetilde{U} \subset U''$ of P such that $\gamma(Q)$ is an isomorphism for all $Q \in \widetilde{U} \cap \mathcal{P}(E)$. For such Q the inclusions in the above chain must be equalities, so we have $\operatorname{im} Q = \operatorname{im} \beta(Q)$ and $\ker(\mathbb{1} - Q) = \operatorname{im} \alpha'(\mathbb{1} - Q)$. This proves the second equality in the corollary, and the first one follows by reversing the roles of Q and $\mathbb{1} - Q$. $\square$

In topological terms, Corollary 2.31 says that the kernels and images of projections form continuous Banach vector bundles

$$Ker \to \mathcal{P}(E) \quad \text{and} \quad Im \to \mathcal{P}(E).$$

Chapter 3

Fredholm maps and their degrees

3.1 Fredholm maps

Definition 3.1. Let X, Y be Banach manifolds and $k \in \mathbb{N} \cup \{\infty\}$. A C^k-map $f : X \to Y$ is called *Fredholm* if for all $x \in X$ its derivative $T_x f$ is a linear Fredholm operator. If X is connected, its *index* $\mathrm{ind}(f) := \mathrm{ind}(T_x f)$ is independent of $x \in X$. If X is disconnected we will assume that f has the same index on each component, so we can still unambiguously speak of the degree of f.

For Fredholm maps the constant rank theorem takes a particularly simple form:

Corollary 3.1 (Constant rank theorem for Fredholm maps)**.** *Let X, Y be Banach manifolds and $f : X \to Y$ be a smooth Fredholm map of index k. Suppose that* $\dim \ker Df(x) = s$ *for all $x \in X$. Then near each $x \in X$ there exist charts $\phi : X \supset U \hookrightarrow E = P \oplus \mathbb{R}^s$ sending x to 0 and $\psi : Y \supset V \hookrightarrow F = P \oplus \mathbb{R}^{s-k}$ sending $f(x)$ to 0 such that*

$$\psi \circ f \circ \phi^{-1}(p, q) = (p, 0)$$

Problem 3.1. Prove the constant rank theorem for Fredholm maps.

Proper and closed maps. A map between topological spaces is called *proper* if preimages of compact sets are compact, and *closed* if images of closed sets are closed.

Problem 3.2. Let $f : X \to Y$ be a continuous map between topological spaces. Show:
(a) If f is closed and preimages of points are compact, then f is proper.
(b) The converse holds if Y is a metric space.

Proposition 3.2. *Every Fredholm map $f : X \to Y$ is* locally proper, *i.e., each $x \in X$ has a closed neighbourhood N such that $f|_N : N \to Y$ is proper.*

Proof. By Lemma 2.24 (which is applicable because f is Fredholm), after composing with charts near $x \in X$ we may assume that $f : P \oplus Q \supset A \times B \to F = P \oplus R$ has the form $f(p, q) = \big(p, g(p, q)\big)$, where $A \subset P$ and $B \subset Q$ are closed balls around the origin. Since $Df(0)$ is Fredholm, its kernel Q is finite dimensional and thus B is compact. We claim that $f : A \times B \to F$ is proper. For this, it suffices to show that if (x_n) is a sequence in $A \times B$ and $f(x_n)$ converges to $y \in F$, then (x_n) has a convergent subsequence. Write $x_n = (p_n, q_n)$, so that $f(x_n) = \big(p_n, g(p_n, q_n)\big) \to y = (p, r)$ and thus $p_n \to p$. By compactness of B, the sequence (q_n) has a convergent subsequence and the proposition is proved. □

Problem 3.3. (a) Show: if a Fredholm map $f : X \to Y$ between Banach manifolds with $\dim X = \infty$ is closed, then it is proper.
(b) Find a counterexample if $\dim X < \infty$.
Hint for (a): Assume that there exists $y \in Y$ with $f^{-1}(y)$ noncompact and derive a contradiction by construcing a sequence (x_n) in X without a convergent subsequence such that

$f(x_n) \neq y$ for all n and $f(x_n) \to y$ as $n \to \infty$.

Compact maps. A map $k : X \to Y$ between topological spaces is called *compact* if its image is relatively compact (i.e. has compact closure).[1]

Problem 3.4. (a) Let X be a metric space and Y a normed vector space. Show: if $f : X \to Y$ is proper and $k : X \to Y$ is compact, then $f + k : X \to Y$ is proper.
(b) Let X, Y be Banach manifolds. Show: if $k : X \to Y$ is a compact C^1-map, then for each $x \in X$ the differential $T_x k : T_x X \to T_{k(x)} Y$ is a compact linear map. *Hint: For (b) show that the image of the unit ball under $T_x k$ is totally bounded.*

In particular, this problem implies the following nonlinear analog of Proposition 1.19(b).

Proposition 3.3. *Let X be a Banach manifold and Y a Banach space. If $f : X \to Y$ is a proper C^1-Fredholm map and $k : X \to Y$ is a compact C^1-map, then $f + k : X \to Y$ is a proper C^1-Fredholm map. In particular, if U is an open subset of a Banach space X and $k : U \to X$ a compact C^1-map, then $\mathbb{1} - k : U \to X$ is a proper C^1-Fredholm map.* □

3.2 The Sard-Smale theorem

Consider the following classical result of analysis.

Theorem 3.4 (Sard [44]). *Let M, N be manifolds of dimensions m, n with M second countable. Let $f : M \to N$ be a C^k-map with $k > \max\{m - n, 0\}$. Then the set of critical values of f has measure zero.*

[1] Note that a compact *linear* map $K : X \to Y$ is not compact in this sense, but rather its restriction to the unit ball is compact. The definition of compactness is not uniform in the literature, where it sometimes means that k maps bounded sets to precompact ones.

In order to generalize this to infinite dimensions, we first replace the notion of "measure zero":

Corollary 3.5 (Sard II). *Under the hypotheses of Theorem 3.4, the set of critical values of f is meagre and the set of regular values is a Baire set.*

Proof. Since M is second countable, we can cover it by countably many compact sets B_i, $i \in \mathbb{N}$. Each set C_i of critical values of $f|_{B_i}$ is compact (because the set of critical points in B_i is compact), hence closed, and it has empty interior because it has measure zero. Thus the set $\bigcup_{i\in\mathbb{N}} C_i$ of critical values of f is meagre and its complement is a Baire set. $\square$

In general, Corollary 3.5 fails in infinite dimensions:

Example 3.6 (R. Bonic, see [15]). Let $\alpha : \mathbb{R} \to [0,1]$ be a smooth function with $\alpha(t) = 0$ for $t \leq 0$ and $\alpha(t) = 1$ for $t \geq 1$. Define

$$f : E := L^\infty([0,1],\mathbb{R}) \to \mathbb{R}, \qquad f(x) := \int_0^1 \alpha\big(x(t)\big)dt.$$

This is a smooth map whose first derivative is given by

$$Df(x)v = \int_0^1 \alpha'\big(x(t)\big)v(t)dt.$$

For $s \in [0,1]$ let x_s be the characteristic function of $[0,s]$. Then $\alpha(0) = \alpha'(0) = \alpha'(1) = 0$ and $\alpha(1) = 1$ implies

$$\begin{aligned} f(x_s) &= \int_0^1 \alpha\big(x_s(t)\big)dt = \int_0^s \alpha(1)dt = s, \\ Df(x_s)v &= \int_0^1 \alpha'\big(x_s(t)\big)v(t)dt = 0 \end{aligned}$$

for all $v \in E$, hence $Df(x_s) = 0$. So the set of critical values of f contains the interval $[0,1] \subset \mathbb{R}$.

Similarly, Kupka [31] has constructed a smooth map $H \to \mathbb{R}$ on a separable Hilbert space H whose set of critical values contains the interval $[0,1]$. By contrast, Corollary 3.5 continues to hold for Fredholm maps:

What about maps $H \to H$? Smale claims this without proof.

Theorem 3.7 (Sard-Smale theorem [47])**.** *Let X, Y be Banach manifolds. Let $f : X \to Y$ be a C^k-Fredholm map with $k > \max\{\mathrm{ind}(f), 0\}$. Then the set of regular values of f is*
(a) open and dense if f is proper;
(b) a Baire set (in particular dense) if X is second countable.

Proof. (cf. [38, Lectures 18-19]). Denote the set of regular values of f by $\mathcal{R}(f)$. The proof has four steps.

Step 1. If f is proper, then $\mathcal{R}(f)$ is open. This holds because the set of critical points is closed, and the map f is closed by Problem 3.2, so the set of critical values of f is closed.

Step 2. Each $x \in X$ has a closed neighbourhood N such that the set $\mathcal{R}(f|_N) \subset Y$ of regular values of $f|_N$ is open and dense. To see this, let $x \in X$ be given. We pick a neighbourhood of x as in Lemma 2.24 on which, after composing with a chart, $f : N = A \times B \to F = P \oplus R$ has the form $f(p,q) = \big(p, g(p,q)\big)$, where $A \subset P$ and $B \subset Q$ are closed balls around the origin. Moreover, by Proposition 3.2 we may assume that $f|_N$ is proper, so $\mathcal{R}(f|_N)$ is open by Step 1 and it suffices to show that it is dense.
Now $Df(p,q) = \left(\begin{smallmatrix} \mathbb{1} & 0 \\ * & D_2g(p,q) \end{smallmatrix}\right)$ is surjective iff $D_2g(p,q)$ is surjective. So $y = (p,r)$ is a regular value of $f|_N$ iff $D_2g(p,q)$ is surjective for all $q \in Q$ such that $g(p,q) = r$, which is equivalent to r being a regular value of $g|_{\{p\}\times Q}$. But Q is finite dimensional (by the Fredholm property of f), so the set of regular values of $g|_{\{p\}\times Q}$ is dense in $\{p\} \times R$ by Sard's theorem. This shows that the set $\mathcal{R}(f|_N)$ intersects each slice $\{p\} \times R$ in a dense set, so it is dense

in $P \oplus R$.

Step 3. To prove (a), assume that f is proper. Then $\mathcal{R}(f)$ is open by Step 1, so it remains to prove that it is dense. Let $y \in Y$ be given. Then $f^{-1}(y)$ is compact because f is proper, so by Step 2 it is covered by finitely open sets $U_1, \dots, U_r$ such that $\mathcal{R}(f|_{\overline{U}_i})$ is open and dense for all i. Thus $U := U_1 \cup \dots \cup U_r$ is an open neighbourhood of $f^{-1}(y)$ such that $\mathcal{R}(f|_{\overline{U}}) = \bigcap_{i=1}^r \mathcal{R}(f|_{\overline{U}_i})$ is open and dense. Since $f(X \setminus U)$ is closed (by closedness of f) and does not contain y, there exists an open neighbourhood V of y with $V \cap f(X \setminus U) = \emptyset$, so $f^{-1}(V) \subset U$. If follows that $V \cap \mathcal{R}(f) = V \cap \mathcal{R}(f|_{\overline{U}})$ is open and dense in V, so $\mathcal{R}(f)$ contains values arbitrary close to y.

Step 4. To prove (b), assume that X is second countable. Then by Step 2 we find a covering of X by countably many open sets U_i such that $\mathcal{R}(f|_{\overline{U}_i})$ is open and dense for all $i \in \mathbb{N}$, so $\mathcal{R}(f) = \bigcap_{i \in \mathbb{N}} \mathcal{R}(f|_{\overline{U}_i})$ is a Baire set. □

For smooth maps $f : X \to Z$ and $g : Y \to Z$ between Banach manifolds we say that f *is transverse to* g if for all $(x, y) \in X \times Y$ with $f(x) = g(y) = z$ we have $\operatorname{im} T_x f + \operatorname{im} T_y g = T_z Z$. This is equivalent to $f \times g : X \times Y \to Z \times Z$ being transverse to the diagonal in $Z \times Z$ in the sense of Section 2.4, and if g is an embedding it is equivalent to transversality of f to the image of g.

Move to Section 2.4.

Corollary 3.8 (Fredholm transversality)**.** *Let X, Y be Banach manifolds and N a second countable finite dimensional manifold. Let $f : X \to Y$ be a proper C^k-Fredholm map with $k >$ $\max\{\operatorname{ind}(f) + \dim N, 0\}$, and $g : N \hookrightarrow Y$ a C^r-embedding with $r \in \mathbb{N} \cup \{\infty\}$. Then there exists a C^r-embedding $\widetilde{g} : N \hookrightarrow Y$ arbitrarily C^r-close to g such that $\widetilde{g}$ is transverse to f. If g is already transverse to f over some closed subset $A \subset N$, then* *we can arrange $\widetilde{g} = g$ on A.*

Smale claims this without propernes How to prove this

Proof. By second countability of N we find open sets $U_i \subset \overline{U}_i \subset V_i \subset N$, $i \in \mathbb{N}$, such that $N = \bigcup_{i\in\mathbb{N}} U_i$ and each $\overline{V}_i$ is contained in a chart domain. Then we can inductively modify $\widetilde{g}$ to g on V_i, in each step keeping it fixed outside V_i and on the set $A \cup \overline{U}_1 \cup \cdots \cup \overline{U}_{i-1}$. Therefore, it suffices to consider the following semi-local situation: P, Q are Banach spaces with $\dim P = \dim N < \infty$; $U \subset \overline{U} \subset V \subset P$ are bounded open neighbourhoods of 0; $f : X \to \overline{V} \oplus Q$ is a proper C^k-Fredholm map and $g : P \hookrightarrow P \oplus Q$ is the embedding $g(p) = (p, 0)$. Moreover, we assume that g is transverse to f over a given compact subset $A \subset \overline{V}$. The proof of the following claim is left as a exercise (here we use properness of f).

Claim. *There exists a compact set $B \subset \overline{V}$ with $A \subset \mathrm{int}_{\overline{V}} B$ and a constant $\varepsilon > 0$ such that every smooth function $\widetilde{g} : P \to P \oplus Q$ with $\|\widetilde{g} - g\|_{C^1(B)} < \varepsilon$ is transverse to f over B.*

Fix a smooth cutoff function $\chi : P \to [0, 1]$ with support in $V \setminus A$ which equals 1 on $\overline{U} \setminus B$. Let $\pi : \overline{V} \oplus Q \to Q$ be the projection onto the second factor. Then the composition $\pi f : X \to Q$ is a proper C^k-Fredholm map of index $\mathrm{ind}(f) + \dim N$, hence by Theorem 3.7 its regular values are dense in Q. For a regular value q of πf define

$$\widetilde{g} : P \to P \oplus Q, \qquad \widetilde{g}(p) := \big(p, \chi(p)q\big).$$

Then on $\overline{U} \setminus B$ we have $\widetilde{g}(p) = (p, q)$, so for each $x \in X$ with $f(x) = (p, q)$ the differential $T_x(\pi f) = \pi \circ T_x f : T_x X \to Q$ is surjective because q is a regular value of πf. This shows that $\widetilde{g}$ is transverse to f over $\overline{U} \setminus B$. On the other hand $\|\widetilde{g} - g\|_{C^1(B)} = \|\chi\|_{C^1(B)} \|q\|$ can be made smaller than the constant ε in the claim by choosing q close to 0, so that $\widetilde{g}$ is transverse to f over B. Hence

$\widetilde{g}$ is transverse to f over $\overline{U}$ and the corollary is proved. □

Problem 3.5. Prove the claim in the proof of Corollary 3.8.

3.3 Caccioppoli–Smale degree

Let X, Y be Banach manifolds with Y connected. Let $f : X \to Y$ be a proper Fredholm map of index $k \geq 0$. For each regular value $y \in Y$ the preimage $f^{-1}(y)$ is a compact k-dimensional submanifold of X. For two regular values $y_0 \neq y_1$ pick an embedded path $g : [0,1] \hookrightarrow Y$ with $g(0) = y_0$ and $g(1) = y_1$ (which exists because Y is connected). By Corollary 3.8 we can choose g to be transverse to f. Then $f^{-1}(\operatorname{im} g)$ is a compact $(k+1)$-dimensional submanifold of X with boundary $\partial f^{-1}(\operatorname{im} g) = f^{-1}(y_1) \amalg f^{-1}(y_0)$. So $f^{-1}(y_0)$ and $f^{-1}(y_1)$ are "cobordant" in any of the following three meanings of the word:

Definition 3.2. (a) Two disjoint compact k-dimensional submanifolds $M_0, M_1 \subset X$ are called *embedded cobordant in* X if there exists a compact $(k+1)$-dimensional submanifold $W \subset X$ with boundary $\partial W = M_1 \amalg M_0$.
(b) Two (not necessarily disjoint) compact k-dimensional submanifolds $M_0, M_1 \subset X$ are called *cobordant in* X if there exists a compact $(k+1)$-dimensional submanifold $W \subset [0,1] \times X$ with boundary $\partial W = \{1\} \times M_1 \amalg \{0\} \times M_0$.
(c) Two compact k-dimensional abstract manifolds M_0, M_1 are called *(abstractly) cobordant* if there exists a compact $(k+1)$-dimensional manifold W with boundary $\partial W \cong M_1 \amalg M_0$.

Remark 3.9. In Definition 3.2 all cobordisms are *unoriented.* There are corresponding notions in the oriented category where $\partial W =$

$M_1 \amalg (-M_0)$ in (a), and similarly in (b) and (c). We will see in Section 3.6 that for general Fredholm maps there is no consistent way to assign orientations to the preimages $f^{-1}(y)$ of regular values, but such an assignment is possible for a restricted class of maps and gives rise to the Leray–Schauder degree.

Problem 3.6. Concerning the three notions of cobordism in Definition 3.2 show:
(i) We always have $(a) \Longrightarrow (b) \Longrightarrow (c)$, but $(c) \Longrightarrow (b)$ does not hold in general. What about $(b) \Longrightarrow (a)$?
(ii) If X is infinite dimensional, then $(b) \Longrightarrow (a)$ (assuming that M_0, M_1 are disjoint).
(iii) The notions (b) and (c) define equivalence relations.

Problem 3.7. Prove: the set Ω_k of cobordism classes of (unoriented) k-dimensional compact manifolds is an abelian group with disjoint union as addition, and the direct sum $\Omega := \bigoplus_{k=0}^{\infty} \Omega_k$ is a commutative ring with Cartesian product as multiplication. Ω is called the *unoriented cobordism ring*.

Definition 3.3 ([7, 47]). For Banach manifolds X, Y with Y connected and $k \geq 0$ let $\mathcal{PF}_k(X,Y)$ denote the set of proper C^{k+1}-Fredholm maps $X \to Y$ of index k. The *Caccioppoli–Smale degree*

$$\deg : \mathcal{PF}_k(X,Y) \to \Omega_k, \qquad \deg(f) := [f^{-1}(y)]$$

associates to f the abstract (unoriented) cobordism class $[f^{-1}(y)]$ for a regular value $y \in Y$. By the above discussion this does not depend on the choice of y. In particular, for $k = 0$, $\deg(f) \in \mathbb{Z}/2\mathbb{Z}$ is the number modulo 2 of preimages of a regular value.

Theorem 3.10 (Properties of the Caccioppoli–Smale degree I). *(Normalization)* $\deg(\mathrm{id} : X \to X) = 1$.

(Excision) For $f \in \mathcal{PF}_k(X, Y)$ and each nonempty connected open subset $V \subset Y$,

$$\deg(f) = \deg\big(f|_{f^{-1}(V)} : f^{-1}(V) \to V\big).$$

(Additivity) For $f_i \in \mathcal{PF}_k(X_i, Y)$,

$$\deg(f_1 \amalg f_2 : X_1 \amalg X_2 \to Y) = \deg(f_1) + \deg(f_2).$$

(Multiplicativity) For $f_i \in \mathcal{PF}_{k_i}(X_i, Y_i)$ of class $C^{k_1+k_2+1}$,

$$\deg(f_1 \times f_2 : X_1 \times X_2 \to Y_1 \times Y_2) = \deg(f_1) \cdot \deg(f_2).$$

(Boundary) For a Banach manifold X with boundary ∂X and a proper C^{k+2}-Fredholm map $f : X \to Y$ of index $k+1$,

$$\deg(f|_{\partial X} : \partial X \to Y) = 0.$$

(Homotopy) For a proper C^{k+2}-Fredholm map $f : [0,1] \times X \to Y$, $(t, x) \mapsto f_t(x)$ of index $k+1$,

$$\deg(f_0) = \deg(f_1).$$

(Diffeomorphism) For C^{k+1}-diffeomorphisms $\phi : X' \xrightarrow{\cong} X$ and $\psi : Y \xrightarrow{\cong} Y'$,

$$\deg(f) = \deg(\psi \circ f \circ \phi).$$

(Solution) If $\deg(f) \neq 0$, then $f \in \mathcal{PF}_k(X, Y)$ is surjective.

Proof. (Normalization), (Additivity), (Multiplicativity) and (Diffeomorphism) are clear.
(Excision) $U := f^{-1}(V)$ is open in X, thus a Banach manifold, and $f|_U : U \to V$ is proper. Now a regular value $y \in V$ of f is also regular for $f|_U$ and $f^{-1}(y) = (f|_U)^{-1}(y)$.
(Boundary) Let $y \in Y$ be a regular value of f and $f|_{\partial X}$. Then $W := f^{-1}(y) \subset X$ is a compact submanifold with boundary

$\partial W = W \cap \partial X = (f|_{\partial X})^{-1}(y)$.
(Homotopy) is a special case of (Boundary).
(Solution) If f is not surjective, then any $y \in Y \setminus \mathrm{im}\,(f)$ is regular and $\deg(f) = [\emptyset] = 0$. □

Applications for Fredholm maps of index zero. The first application is a formal consequence of the properties of the Caccioppoli–Smale degree.

Corollary 3.11 (No proper Fredholm retraction to boundary)**.** *Let X be a Banach manifold with boundary ∂X. Then there exists no proper index 0 Fredholm map $f : X \to X$ with $f(X) \subset \partial X$ and $f|_{\partial X} = \mathrm{id}$ (i.e. $f \circ f = f$).*

Proof. Suppose such a map f exists. Then we can view f as an index 1 Fredholm map $X \to \partial X$ and the (Boundary) property in Theorem 3.10 gives $\mathrm{ind}(f|_{\partial X} = \mathrm{id} : \partial X \to \partial X) = 0$, contradicting the (Normalization) property in Theorem 3.10. □

The next applications are taken from A. Tromba's article [50]. They are based on the following consequence of the constant rank theorem.

Theorem 3.12 (Tromba [50])**.** *Let X, Y be Banach manifolds and $f : X \to Y$ be an injective Fredholm operator. Then the* singular set $\Sigma(f) := \{x \in X \mid \ker(T_x f) \neq 0\}$ *is closed with empty interior.*

The proof uses the following simple lemma.

Lemma 3.13. *Let $T : X \to \mathcal{F}(E, F)$ be a continuous map from a nonempty topological space X to the space of linear Fredholm operators between Banach spaces E, F. Then there exists a nonempty open subset $U \subset X$ such that* $\dim\ker T(x)$ *is constant for $x \in U$.*

Proof. For each $s \in \mathbb{N}_0$ let $X_s := \{x \in X \mid \dim\ker T(x) \geq s\}$. Then $X = X_0 \supset X_1 \cdots$, so there exist a unique s_0 such that $X = X_{s_0} \neq X_{s_0+1}$. Pick $x_0 \in X_{s_0} \setminus X_{s_0+1}$, so that $\dim\ker T(x_0) = s_0$. By Remark 2.30 there exists an open neighbourhood U of x_0 such that for all $x \in U$ we have $\dim\ker T(x) \leq s_0$, hence $\dim\ker T(x) = s_0$ by definition of s_0. □

Proof of Theorem 3.12. By Remark 2.30 the set $X \setminus \Sigma(f) = \{x \mid \ker(T_x f) = 0\}$ is open, so $\Sigma(f)$ is closed. Arguing by contradiction, suppose $\Sigma(f)$ contains a nonempty open set V. After choosing V sufficiently small and composing with charts, we may assume that $f : E \supset V \to F$ for Banach spaces E, F with $\dim\ker Df(x) \geq 1$ for all $x \in V$. Applying Lemma 3.13 to the continuous map $Df : V \to \mathcal{F}(E, E)$, we find a nonempty open subset $U \subset V$ and some $s \in \mathbb{N}$ such that $\dim\ker Df(x) = s \geq 1$ for all $x \in U$. Then by Corollary 3.1 there exist charts ϕ, ψ near 0 such that $\psi \circ f \circ \phi^{-1}(p, q) = (p, 0)$ for $(p, q) \in P \oplus \mathbb{R}^s$, contradicting injectivity of f. □

Theorem 3.12 combined with Theorem 3.10 has a number of interesting corollaries. In all of them, X, Y are Banach manifolds and f is a C^1-Fredholm map of index 0.

Corollary 3.14. *Let $f : X \to Y$ be a proper Fredholm map of index* 0. *If Y is connected and f is injective, then f is surjective and* $\deg(f) = 1$.

Proof. By Theorem 3.12 there exists some $x \in X$ for which $T_x f$ is injective, hence also surjective because $\mathrm{ind}(f) = 0$. Then $y := f(x)$ is a regular value with $f^{-1}(y) = \{x\}$, and therefore $\deg(f) = 1$. Surjectivity of f now follows from the (Solution) property in Theorem 3.10. □

Corollary 3.15 (Invariance of domain for Fredholm maps)**.** *Let $f : X \to Y$ be a Fredholm map of index* 0*. If f is locally injective, then f is open.*

Proof. Let $W \subset X$ be open, $x \in W$ and $y := f(x)$. We need to find an open neighbourhood $V \subset f(W)$ of y. Since f is locally injective and Fredholm maps are locally proper (Proposition 3.2), x has a connected open neighbourhood $U \subset W$ such that $f|_{\overline{U}} : \overline{U} \to Y$ is proper and injective. It follows that $y \notin f(\partial U)$, and $f(\partial U) \subset Y$ is closed by properness and Problem 3.2. Let V be the connected component of $Y \setminus f(\partial U)$ containing y. Then V is an open neighbourhood of y, and $f(U) \subset V$ by connectedness of U. Moreover, $f(\partial U) \cap V = \emptyset$ implies $\overline{U} \cap f^{-1}(V) = U$, so $f|_U : U \to V$ is proper and injective, hence surjective by Corollary 3.14 and connectedness of V. Thus $V = f(U) \subset f(W)$ is the desired open neighbourhood of y. $\square$

Note that for linear Fredholm operators of index zero, injectivity implies surjectivity (and vice versa). For operators of the form $\mathbb{1} - K$ with K compact this statement is known as the *Fredholm alternative*. The following corollary can be seen as a nonlinear version of this.

Corollary 3.16 (Nonlinear Fredholm alternative)**.** *Let $f : X \to Y$ be a proper Fredholm map of index* 0*. If Y is connected and f is locally injective, then f is surjective and a finite covering map. In particular, if X is connected and Y is simply connected, then f is a homeomorphism.*

Proof. By Corollary 3.15 the map f is open, hence a local homeomorphism. In particular, its image is open, and it is also closed because f is proper and therefore closed. Since Y is connected, the image must be all of Y. Now Problem 3.8 below yields that f is a covering map, and the last assertion is a standard result in covering space theory. $\square$

Problem 3.8. Show: if a continuous map $f : X \to Y$ between metric spaces is surjective, proper and a local homeomorphism, then it is a finite covering map.

3.4 The Borsuk-Ulam theorem for Fredholm maps

In this section we prove a version of the Borsuk-Ulam theorem for Fredholm maps. The classical Borsuk-Ulam theorem and its applications will be discussed in the next section.

Caccioppoli–Smale degree for domains with boundary. There is a slight generalization of the Caccioppoli–Smale degree to the following setup for $k \in \mathbb{N}_0$:
- X, Y are Banach manifolds and $U \subset X$ is open;
- $f : \overline{U} \to Y$ is continuous and proper such that $f|_U$ is C^{k+1}-Fredholm of index k;
- $y \in Y \setminus f(\partial U)$.

Then we define

$$d(f, y) := [f^{-1}(y')] \in \Omega_k$$

for a regular value y' in the connected component of $Y \setminus f(\partial U)$ containing y. As before, this does not depend on the choice of y'. For $U = X$ and Y connected, $d(f, y)$ is independent of y and agrees with the earlier degree $\deg(f)$.

Theorem 3.17 (Properties of the Caccioppoli–Smale degree II)**.** *(Normalization) If M is a compact connected k-dimensional manifold, E a Banach space, and $f : M \times E \to E$ a C^{k+1}-map of the form $f(x, v) = \phi(x)v$ for a continuous map $\phi : M \to$*

$GL(E)$, then $d(f,0) = [M] \in \Omega_k$.
(Excision) For open subsets $V \subset U$ with $f^{-1}(y) \subset V$ and $W \subset Y$ with $f(\overline{V}) \subset W$,

$$d(f,y) = d(f|_{\overline{V}} : \overline{V} \to W, y).$$

(Additivity) If $U = U_1 \amalg U_2$ then

$$d(f,y) = d(f|_{\overline{U}_1}, y) + d(f|_{\overline{U}_2}, y).$$

(Homotopy) For a proper continuous map $f : [0,1] \times \overline{U} \to Y$, $(t,x) \mapsto f_t(x)$, such that $f|_{[0,1]\times U}$ is C^{k+2}-Fredholm of index $k+1$, and a continuous path $y : [0,1] \to Y$ such that $y(t) \notin f_t(\partial U)$ for all $t \in [0,1]$,

$$d(f_0, y_0) = d(f_1, y_1).$$

(Diffeomorphism) For C^{k+1}-diffeomorphisms $\phi : U' \xrightarrow{\cong} U$ extending continuously to $\overline{U'} \to \overline{U}$ and $\psi : Y \xrightarrow{\cong} Y'$,

$$d(f,y) = d\big(\psi \circ f \circ \phi, \psi(y)\big).$$

(Solution) If $d(f,y) \neq 0$, then there exists $x \in U$ with $f(x) = y$.
Moreover, $d(f,y)$ is uniquely determined by the first five properties on maps f of class C^{k+3}. C^{k+2} should suffic

Proof. The properties are verified as in the proof of Theorem 3.17 with the obvious modifications. For (Homotopy) note that the continuous path y can be approximated by a smooth path $\widetilde{y}$ with the same endpoints still satisfying $\widetilde{y}(t) \notin f_t(\partial U)$ for all t. Now we show in four steps that $d(f,y)$ is uniquely determined by the first four properties for f of class C^{k+2}.

Step 1. After connecting y by a path in $Y \setminus f(\partial U)$ to a regular value and applying (Homotopy), we may assume that y is regular.

Step 2. Write $f^{-1}(y) = M_1 \amalg \cdots \amalg M_r$ with compact connected k-manifolds M_i and pick disjoint open neighbourhoods $U_i \subset U$ of the M_i. Then by (Excision) and (Additivity) $d(f,y) = \sum_{i=1}^r d(f|_{U_i}, y)$, so it suffices to consider the case that $f^{-1}(y) = M$ is connected and U is an arbitrarily small neighbourhood of M.

Step 3. Since y is a regular value, for each $x \in M$ the differential $T_x f : T_x X \to T_y Y$ is surjective with kernel $T_x M$, so it induces an isomorphism $T_x X / T_x M \to T_y Y$. These isomorphisms fit together to a trivialization $Tf : \nu \to M \times T_y Y$ of the normal bundle $\nu \to M$ whose fibre over x is $\nu_x := T_x X / T_x M$. Hence by the tubular neighbourhood theorem (see [32]), a suitable neighbourhood U of M in X is diffeomorphic to $M \times V$ for an open neighbourhood V of the origin in $E := T_y Y$. After applying the (Diffeomorphism) property, we may therefore assume that $f : M \times V \to E$ for an open neighbourhood V of 0 in a Banach space E with regular value $y = 0$ and $f^{-1}(0) = M$.

State in earlier section.

Step 4. For $(x, v) \in M \times V$ we can write

$$f(x,v) = \phi(x)v + r(x,v)\|v\|, \qquad \phi(x) := D_2 f(x,0) \tag{3.1}$$

with continuous maps $\phi : M \to GL(E)$ and $r : M \times V \to E$ satisfying $r(x,0) = 0$. We define $F : [0,1] \times M \times V \to E$ by

$$F(t,x,v) := f_t(x,v) := f(x,v) + t\big(\phi(x)v - f(x,v)\big).$$

Since $D_2 f_t(x,0) = \phi(x) \in GL(E)$, the differential of F is Fredholm of index $k+1$ at points of $[0,1] \times M \times 0$, hence also on $[0,1] \times M \times U$ for some open neighbourhood $U \subset V$ of 0 because the Fredholm property is open. Since Fredholm operators are locally proper and M is compact, $\widetilde{F} := F|_{[0,1]\times M \times \overline{U}}$ is proper for U sufficiently small. Again by compactness of M and the open mapping theorem, there exists $\varepsilon > 0$ such that $\|\phi(x)v\| \geq \varepsilon \|v\|$

for all $(x, v) \in M \times E$, and for U sufficiently small we have $\|r(x, v)\| \leq \varepsilon/2$ for all $(x, v) \in M \times \overline{U}$. Using (3.1) we rewrite $f_t(x, v) = \phi(x)v + (1 - t)r(x, v)\|v\|$ and estimate for $(x, v) \in M \times \overline{U}$:

$$\begin{aligned}\|f_t(x, v)\| &\geq \|\phi(x)v\| - |1 - t|\,\|r(x, v)\|\,\|v\| \\ &\geq \varepsilon\|v\| - \frac{\varepsilon}{2}\|v\| = \frac{\varepsilon}{2}\|v\|.\end{aligned}$$

Thus $\widetilde{F}^{-1}(0) = [0, 1] \times M \times \{0\}$. Now we use the hypothesis that f is of class C^{k+3}, so $\widetilde{F}$ is of class C^{k+2}. This allows us to apply the (Homotopy) property to $\widetilde{F}$ and deduce

$$d(f, 0) = d(f_0|_{M\times U}, 0) = d(f_1|_{M\times U}, 0) = d(f_1, 0),$$

where the first and last equalities follow from (Excision). Since $f_1(x, v) = \phi(x)v$, (Normalization) yields $d(f_1, 0) = [M]$. □

Problem 3.9. Prove the following *composition formula* for the Caccioppoli–Smale degree: Let X, Y, Z be Banach manifolds and $U \subset X$ open. Let $f : \overline{U} \to Y$ and $g : Y \to Z$ be proper C^1-Fredholm maps of index 0. Then $g \circ f : \overline{U} \to Z$ is a proper C^1-Fredholm map of index zero and for each $z \in Z \setminus (g \circ f)(\partial U)$ we have

$$d(g \circ f, z) = \sum_i d(f, K_i)d(g|_{K_i}, z),$$

where K_i are the connected components of $Y \setminus f(\partial U)$, only finitely many terms in the sum are nonzero, and $d(f, K_i) := d(f, y_i)$ for any $y_i \in K_i$. In particular, if $U = X$ and Y, Z are connected, then

$$\deg(g \circ f) = \deg(g)\deg(f).$$

The Borsuk-Ulam theorem for Fredholm maps. The classical Borsuk-Ulam theorem has the following generalization to proper index 0 Fredholm maps.

Theorem 3.18 (Borsuk-Ulam theorem for Fredholm maps)**.** *Let $U \subset E$ be a bounded open subset of a Banach space with $-U = U$, and $f \in \mathcal{PF}_0(\overline{U}, E)$ with $0 \notin f(\partial U)$ and $f(-x) = -f(x)$ for all $x \in U$. Then $d(f, 0) \equiv 1 \mod 2$.*

The proof is based on the following very useful result.

Proposition 3.19 (Parametric transversality)**.** *Let X, P, Y be Banach manifolds, $f : X \times P \to Y$ a smooth map and $y \in Y$. Suppose that for all $(x, p) \in f^{-1}(y)$ the differential $T_{(x,p)}f$ is surjective and its restriction $T_{(x,p)}f|_{T_xX\times 0} : T_xX \to T_yY$ is Fredholm of index k. Then:*
(a) $M := f^{-1}(y)$ is a smooth submanifold;
(b) the projection $X \times P \to P$ onto the second factor restricts to an index k Fredholm map $\pi : M \to P$;
(c) $p \in P$ is a regular value of π if and only if y is a regular value of the map $f_p : X \to Y$, $f_p(x) := f(x, p)$.

We first consider the linear version of the proposition.

Problem 3.10. Let E, P, F be Banach spaces and $A : E \oplus P \to F$ a surjective continuous linear map. Let $A_1 := A|_{E\times 0} : E \to F$ and $\pi : \ker A \to P$ be the restriction of the projection $E \oplus P \to P$ onto the second factor. Then $\ker \pi = \ker A_1 \times 0$ and the map $A_2 := A|_{0\times P} : P \to F$ induces an algebraic linear isomorphism $P/\operatorname{im} \pi \stackrel{\cong}{\to} F/\operatorname{im} A_1$. In particular:
(a) if A_1 is Fredholm, then A has a right inverse;
(b) π is Fredholm of index k if and only if A_1 is;
(c) π is surjective if and only if A_1 is.

Proof of Proposition 3.19. For $(x, p) \in f^{-1}(y)$ the map $A := T_{(x,p)}f : T_xX \oplus T_pP \to T_yY$ and its restriction $A_1 := T_{(x,p)}f|_{T_xX\times 0} : T_xX \to T_yY$ satisfy the hypotheses of Problem 3.10. Now (a), (b),

(c) follow from the corresponding properties in Problem 3.10. First, A has a right inverse, hence y is a regular value and M is a submanifold by Corollary 2.23. Next, since $T_{(x,p)}\pi : T_{(x,p)}M = \ker A \to T_pP$ is the restriction of the projection $T_xX \oplus T_pP \to T_pP$ onto the second factor, it is Fredholm of index k because A_1 is. Finally, $T_{(x,p)}\pi$ is surjective iff $A_1 = T_xf_p$ is surjective. □

Proof of Theorem 3.18. We follow the proof in [18]. Since f is proper, and therefore closed, $f(\partial U)$ is closed. Hence $0 \notin f(\partial U)$ implies $\delta := \mathrm{dist}\big(f(\partial U), 0\big) > 0$. Since U is bounded, there exists a constant $M > 0$ such that $\|x\| \le M$ for all $x \in \partial U$. For $\varepsilon > 0$ let $\mathcal{K}_\varepsilon$ denote the space of compact linear operators $K : E \to E$ with $\|K\| < \varepsilon$, and for $K \in \mathcal{K}_\varepsilon$ define

$$f_K : \overline{U} \to E, \qquad f_K(x) := f(x) + Kx,$$

so $f_K \in \mathcal{PF}_0(\overline{U}, E)$. Moreover, for $x \in \partial U$ we have

$$\|f_K(x)\| \ge \|f(x)\| - \|K\|\,\|x\| \ge \delta - \varepsilon M > 0$$

if we choose $\varepsilon < \delta/M$, hence $0 \notin f_K(\partial U)$. Since the map $[0,1] \times \overline{U} \to E$, $(t,x) \mapsto f_{tK}(x)$ is proper and $0 \notin f_{tK}(\partial U)$ for all t, it follows that $d(f,0) = d(f_K,0)$. Note that $f_K(-x) = -f_K(x)$ because K is linear. This shows that we can perturb f by some $K \in \mathcal{K}_\varepsilon$ to compute the degree. Using this, we now conclude the proof in 3 steps.

Step 1. Consider $Df_K(0) = T + K$ with $T := Df(0)$. Since T is Fredholm of index 0, we have splittings $E = E_1 \oplus \ker T = \mathrm{im}\,T) \oplus E_2$ with $\dim E_2 = \dim\ker T$. Pick $K \in \mathcal{K}_\varepsilon$ such that $K|_{E_1} = 0$ and $K|_{\ker T} : \ker T \to E_2$ is an isomorphism, hence $T + K$ is an isomorphism. After renaming f_K back to f, we may therefore assume that $Df(0)$ is an isomorphism.

Step 2. Consider the map

$$F : U \times \mathcal{K}_\varepsilon \to E, \qquad (x, K) \mapsto f_K(x).$$

Its differential at (x, K),

$$DF(x, K) \cdot (v, L) = \big(Df(x) + K\big)v + Lx,$$

is surjective if $x \neq 0$ because then Lx takes arbitrary values in E as L varies in the space of compact linear operators $E \to E$. For $x = 0$ it is also surjective for ε sufficiently small because $Df(0)$ is an isomorphism. Since f_K is Fredholm of index 0 for each fixed K, the map F satsifies the hypotheses of Proposition 3.19 with $k = 0$. Hence $M := F^{-1}(0)$ is a submanifold and the map $\pi : M \to \mathcal{K}_\varepsilon$ induced by projection onto the second factor is Fredholm of index 0. According to Problem 3.11 below there exists a regular value $K \in \mathcal{K}_\varepsilon$ for π. Then by Proposition 3.19(c), 0 is a regular value of f_K, so after renaming f_K back to f we may assume that 0 is a regular value of f.

Step 3. If 0 is a regular value of f, properness and $f(-x) = -x$ imply that $f^{-1}(0) = \{0, x_1, -x_1, \dots, x_r, -x_r\}$, hence $d(f, 0) \equiv 1 \mod 2$. □

Problem 3.11. Find a regular value for the map $\pi : M \to \mathcal{K}_\varepsilon$ in Step 2 of the proof of Theorem 3.18 in two ways:

(a) If E is separable, show that the space of compact linear operators $E \to E$ is also separable and apply the Sard-Smale theorem.

(b) In the general case, use compactness of $\pi^{-1}(0)$ and the following local version of the Sard-Smale theorem (which follows from its proof): each point $m \in M$ has an open neighbourhood $W \subset M$ such that the set of regular values of $\pi|_W : W \to \mathcal{K}_\varepsilon$ is a Baire set.

3.5 Brouwer degree

In this and the next section we sketch, mostly without proof, two refinements of the Caccioppoli–Smale degree in more special situations: the Brouwer degree in finite dimensions, and the Leray–Schauder degree for maps of the form $\mathbb{1} - k$ with compact k.

The Brouwer degree. In the finite dimensional case the Fredholm hypothesis is automatically satisfied, and so is the properness assumption if we restrict to compact domains. Moreover, the differentiability assumptions are redundant because we can approximate continuous maps by smooth ones. Finally, on the class of oriented manifolds we can upgrade the degree to take values in the *oriented cobordism ring* $\Omega^+ = \bigoplus_{k=0}^{\infty} \Omega_k^+$ (consisting of compact oriented manifolds modulo oriented cobordisms) using the following orientation convention. If $f : M \to N$ is a smooth map between oriented manifolds and $K := f^{-1}(y)$ for a regular value $y \in N$, then at each $x \in K$ we have a short exact sequence

$$0 \longrightarrow T_x K \longrightarrow T_x M \xrightarrow{T_x f} T_y N \longrightarrow 0.$$

Now for each short exact sequence of finite dimensional vector spaces $0 \to A \to B \to C \to 0$, orientations of two of the spaces determine an orientation of the third one by requiring that oriented bases of C and A (in this order) give an oriented basis of $B \cong C \oplus A$. Thus orientations of $T_x M$ and $T_x N$ induce an orientation of $T_x K$.

In view of the preceding discussion, we consider the following setup:
- M, N are oriented manifolds of dimensions $m \geq n$, and $U \subset M$ is open and relatively compact;
- $f : \overline{U} \to N$ is continuous;
- $y \in N \setminus f(\partial U)$.

Then we uniformly approximate f by a continuous map $\widetilde{f} : \overline{U} \to N$ such that $y \in N \setminus \widetilde{f}(\partial U)$ and $\widetilde{f}|_U$ is smooth and define the *Brouwer degree*

$$d(f, y) := [\widetilde{f}^{-1}(\widetilde{y})] \in \Omega^+_{m-n}$$

for a regular value $\widetilde{y}$ in the connected component of $Y \setminus \widetilde{f}(\partial U)$ containing y. By the usual argument, this does not depend on the choices of $\widetilde{f}$ and $\widetilde{y}$. Ignoring orientations gives back the Caccioppoli–Smale degree (restricted to compact domains in finite dimension and extended to continuous maps).

Problem 3.12. Formulate and prove an analog of Theorem 3.17 for the Brouwer degree $d(f, y)$. Here the (Normalization) property reads $d(f, 0) = [M]$ if $\det \phi(x) > 0$ for some (and hence all) $x \in M$, and $d(f, 0) = [-M]$ otherwise, and the (Diffeomorphism) property holds for C^1-diffeomorphisms ϕ, ψ.

Problem 3.13. Let M, N be oriented manifolds of dimensions $m \geq n$ with N connected.
(a) Define a Brouwer degree $\deg(f) \in \Omega^+_{m-n}$ for each proper continuous map $f : M \to N$.
(b) Formulate and prove an analog of Theorem 3.10 for the Brouwer degree $\deg(f)$.

Applications. Let us now specialize to the case $m = n$ (i.e. of Fredholm index zero). The proof of Corollary 3.11 carries over verbatim to give

Corollary 3.20 (No continuous retraction to boundary)**.** *Let M be a compact finite dimensional manifold with boundary ∂M. Then there exists no continuous map $f : M \to M$ with $f(M) \subset \partial M$ and $f|_{\partial M} = \mathrm{id}$ (i.e. $f \circ f = f$).*

Corollary 3.21 (Brouwer fixed point theorem). *Every continuous map $k : \overline{B^n} \to \overline{B^n}$ on the closed unit ball $\overline{B^n} \subset \mathbb{R}^n$ has a fixed point.*

Proof. Set $B := \overline{B^n}$ and suppose $k(x) \neq x$ for all $x \in B$. Define the homotopy $f_t(x) := x - tk(x)$ from $f_0 = \mathbb{1}$ to $f_1 = \mathbb{1} - k$. If $f_t(x) = 0$ for $(t, x) \in [0, 1] \times \partial B$, then $1 = |x| = t|k(x)| \leq 1$ implies $t = 1$ and thus $x = tk(x) = k(x)$, contradicting the above assumption. Hence $f_t(x) \neq 0$ for $(t, x) \in [0, 1] \times \partial B$, and homotopy invariance and normalization of the Brouwer degree imply $d(f_1, 0) = d(f_0, 0) = 1$, so there exists $x \in B$ with $k(x) = x$, contradicting the above assumption. □

Problem 3.14. (a) Show that Corollary 3.21 follows formally from Corollary 3.20 for $M = \overline{B^n}$ and vice versa.
(b) Find a continuous map $f : B^n \to B^n$ on the *open* unit ball $B^n \subset \mathbb{R}^n$ without fixed points.

While the preceding two corollaries have only used the Brouwer degree mod 2, the following corollary needs the $\mathbb{Z}$-valued degree. Note that for $m = n$ the definition of the Brouwer degree becomes

$$d(f, y) = \sum_{x \in \widetilde{f}^{-1}(y)} \delta(x) \in \mathbb{Z},$$

where $\delta(x)$ equals $+1$ if $T_x f : T_x M \to T_y N$ is orientation preserving and -1 otherwise. For example, the antipodal map $x \mapsto -x$ on $\mathbb{R}^n$ and its restriction to the unit sphere $S^{n-1} \subset \mathbb{R}^n$ both have degree $(-1)^n$.

Corollary 3.22 (Hedgehog theorem). *Let v be a continuous vector field tangent to the unit sphere $S^{n-1} \subset \mathbb{R}^n$, i.e. a continuous map $v : S^{n-1} \to \mathbb{R}^n$ with $v(x) \perp x$ for all x. If n is odd, then there exists some $x \in S^{n-1}$ with $v(x) = 0$.*

Problem 3.15. Prove the hedgehog theorem in two ways:
(a) by showing that v without zeroes would give a homotopy of the identity map $S^{n-1} \to S^{n-1}$ to the antipodal map $x \mapsto -x$;
(b) by extending v to a continuous map $B \to B$ on the closed unit ball $B \subset \mathbb{R}^n$ and deforming it to $\mathbb{1} : B \to B$ and $-\mathbb{1} : B \to B$.

The Borsuk-Ulam theorem. All the corollaries in Section 3.3 continue to hold with "C^1-Fredholm map of index zero" replaced by "continuous map between n-dimensional manifolds". However, the proofs do not carry over directly to the continuous case because Tromba's Theorem 3.12 uses differentiability in an essential way. Instead, we will derive the corollaries from the classical

Theorem 3.23 (Borsuk-Ulam theorem)**.** *Let $U \subset \mathbb{R}^n$ be a bounded open subset with $-U = U$, and $f : \overline{U} \to \mathbb{R}^n$ be continuous with $0 \notin f(\partial U)$ and $f(-x) = -f(x)$ for all $x \in U$. Then $d(f, 0) \equiv 1 \mod 2$.*

Proof. Let $\delta := \min_{\partial U}|f| > 0$. Let $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a C^1-map with $\|\widetilde{f} - f\|_{C^0(\overline{U})} < \delta$ and define $g : \overline{U} \to \mathbb{R}^n$ by $g(x) := \frac{1}{2}\big(\widetilde{f}(x) - \widetilde{f}(-x)\big)$. For $x \in \overline{U}$ we have

$$\begin{aligned} |g(x) - f(x)| &= \left|\frac{1}{2}\big(\widetilde{f}(x) - \widetilde{f}(-x)\big) - \frac{1}{2}\big(f(x) - f(-x)\big)\right| \\ &\le \frac{1}{2}\Big(|\widetilde{f}(x) - f(x)| + |\widetilde{f}(-x) - f(-x)|\Big) < \delta. \end{aligned}$$

For $t \in [0, 1]$ let $f_t(x) := (1 - t)f(x) + tg(x)$. Then for $x \in \partial U$ we have

$$|f_t(x)| \ge |f(x)| - t|g(x) - f(x)| > \delta - t\delta \ge 0,$$

so homotopy invariance of the Brouwer degree implies $d(f, 0) = d(g, 0)$. Since g is of class C^1 and $g(-x) = -g(x)$, Theorem 3.18 yields $d(g, 0) \equiv 1 \mod 2$. □

In the literature the "Borsuk-Ulam theorem" usually refers to the following corollary. It says, for example, that on the surface of the earth there always exists a pair of antipodal points with exactly the same temperature and pressure.

Corollary 3.24 (Borsuk-Ulam antipodal theorem)**.** *For every continuous function $g : S^n \to \mathbb{R}^n$ on the unit sphere $S^n \subset \mathbb{R}^{n+1}$ there exists some $x \in S^n$ with $g(-x) = g(x)$.*

Proof. Consider $E := \mathbb{R}^{n+1}$, its proper subspace $F := \mathbb{R}^n \times \{0\}$, and the closed unit ball $B \subset E$. Suppose there exists a continuous function $g : \partial B \to F$ with $g(-x) \neq g(x)$ for all $x \in \partial B$. Extend g to a continuous function $\overline{g} : B \to F$ and define $f : B \to F \hookrightarrow E$ by $f(x) := \frac{1}{2}\big(\overline{g}(x) - \overline{g}(-x)\big)$. Then $f|_{\partial B} \neq 0$ and $f(-x) = -f(x)$, hence $d(f, 0) \neq 0$ by Theorem 3.23. Since $d(f, 0) = d(f, y)$ for all y with $|y| < \delta = \min_{\partial B} |f|$, it follows that the image $f(B)$ contains the open δ-ball around the origin, contradicting $f(B) \subset F$. □

The following problem answers the important question whether one can cut a ham-and-cheese sandwich into two pieces with equal amounts of ham, cheese and bread.

Problem 3.16. Prove the *sandwich theorem*: For n bounded measurable subsets $A_1, \dots, A_n \subset \mathbb{R}^n$ there exists a hyperplane which cuts each A_i into two pieces of equal volume. *Hint: For $x \in S^n$ define the half-space $H_x^+ := \{y \in \mathbb{R}^n \mid \sum_{i=1}^n x_i y_i \geq x_{n+1}\}$. Now apply the Borsuk-Ulam antipodal theorem to the function $f = (f_1, \dots, f_n) : S^n \to \mathbb{R}^n$ with $f_i(x) := \mathrm{vol}(A_i \cap H_x^+)$.*

Invariance of domain. Now we can prove the continuous analog of Corollary 3.15.

Theorem 3.25 (Invariance of domain, Brouwer 1912)**.**
Let $f : M \to N$ be a continuous map between n-dimensional manifolds. If f is locally injective, then f is open.

Proof. This comes down to the following local statement: If B_r denotes the open ball of radius r around the origin in $E := \mathbb{R}^n$ and $f : \overline{B}_r \to E$ is continuous and injective with $f(0) = 0$, then there exists $s > 0$ such that $B_s \subset f(B_r)$. To show this, consider the homotopy

$$f_t(x) := f\Big(\frac{x}{1+t}\Big) - f\Big(-\frac{tx}{1+t}\Big)$$

from $f_0 = f$ to the odd function f_1. If $f_t(x) = 0$ for $(t, x) \in [0, 1] \times \overline{B}_r$, then injectivity of f implies $\frac{x}{1+t} = -\frac{tx}{1+t}$ and thus $x = 0$. Hence $f_t(x) \neq 0$ for $(t, x) \in [0, 1] \times \partial B_r$, and homotopy invariance and Theorem 3.23 yield $d(f, 0) = d(f_1, 0) \neq 0$. Since $d(f, 0) = d(f, y)$ for all y with $|y| < s := \min_{\partial B_r} |f|$, it follows that $B_s \subset f(B_r)$. □

The proof of Corollary 3.16 carries over verbatim to give

Corollary 3.26 (Nonlinear Fredholm alternative in finite dimension)**.** *Let $f : M \to N$ be a proper continuous map between n-dimensional manifolds. If N is connected and f is locally injective, then f is surjective and a finite covering map. In particular, if M is connected and N is simply connected, then f is a homeomorphism.*

Example 3.27. If $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuous, locally injective and proper (meaning $|f(x)| \to \infty$ as $|x| \to \infty$), then f is a homeomorphism. Note that surjectivity of f may fail if we drop the assumption of local injectivity (take $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$) or properness (take $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto e^x$).

The following two important consequences of Theorem 3.25 assert that the dimension and the boundary of a topological manifold are well-defined (i.e., invariant under coordinate changes).

Corollary 3.28 (Invariance of dimension)**.** *If $f : U \to \mathbb{R}^m$ is a continuous injective map defined on an open subset $U \subset \mathbb{R}^n$, then $n \leq m$. In particular, if two open subsets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are homeomorphic, then $n = m$.*

Proof. If $n > m$, then the composition $\widetilde{f} : U \to \mathbb{R}^m \cong \mathbb{R}^m \times \{0\} \hookrightarrow \mathbb{R}^n$ is continuous injective, so by Theorem 3.25 $\widetilde{f}(U)$ is open in $\mathbb{R}^n$, contradicting $\widetilde{f}(U) \subset \mathbb{R}^m \times \{0\}$. $\square$

Corollary 3.29 (Invariance of boundary)**.** *If $f : U \to V$ is a homeomorphism between open subsets U, V of $H^n := \{x \in \mathbb{R}^n \mid x_n \geq 0\}$, then $f(U \cap \partial H^n) = V \cap \partial H^n$.*

Proof. Otherwise, after possibly replacing f by f^{-1}, there exists $x \in U \setminus \partial H^n$ with $f(x) \in \partial H^n$. Let $\widetilde{U} \subset U \setminus \partial H^n$ be an open neighbourhood of x and consider the continuous injective map $\widetilde{f} := f|_{\widetilde{U}} : \widetilde{U} \to \mathbb{R}^n$. Since $\widetilde{U}$ is open in $\mathbb{R}^n$, Theorem 3.25 implies that $\widetilde{f}(\widetilde{U})$ is open in $\mathbb{R}^n$, contradicting $\widetilde{f}(\widetilde{U}) \subset H^n$ and $\widetilde{f}(x) \in \partial H^n$. $\square$

3.6 Leray–Schauder degree

All the results on continuous maps of the previous section fail in infinite dimensions. For example:

- The map $f : \ell^2 \to \ell^2$, $x \mapsto (0, x_1, x_2, \dots)$ is continuous injective but not open, so invariance of domain (Theorem 3.25) fails.
- *Kakutani's map* $f : B \to B$, $x \mapsto (\sqrt{1 - \|x\|^2}, x_1, x_2, \dots)$ on the closed unit ball $B \subset \ell^2$ is continuous but has no fixed point, so Brouwer's fixed point theorem fails.
- Klee [29] has constructed a homeomorphism from ℓ^2 onto the closed half-space $\{x_1 \geq 0\} \subset \ell^2$, so invariance of the boundary fails.

• All infinite dimensional separable Banach spaces are homeomorphic (Kadec [28]), although e.g. ℓ^p is not isomorphic to ℓ^q for $p \neq q$ (see [3]), so there is no analog of invariance of dimension.

By contrast, we will see that many results carry over to continuous maps of the form $\mathbb{1} - k$ with k compact.

The Leray–Schauder degree. We follow the exposition in Deimling [14]. The starting point is the following topological result, which can be viewed as a nonlinear analog of Proposition 1.13(d).

Proposition 3.30 (Finite dimensional approximation)**.**
Let $k : X \to E$ be a compact map from a topological space X to a Banach space E. Then for every $\varepsilon > 0$ there exists a compact map $\widetilde{k} : X \to E$ whose image is contained in a finite dimensional subspace $F \subset E$ such that $\|\widetilde{k}(x) - k(x)\| < \varepsilon$ for all $x \in X$.

Proof. Since $k(X)$ is compact, it is covered by finitely many open ε-balls $B_\varepsilon(y_i)$, $1 \leq i \leq n$. Define a continuous partition of unity $\psi_i : k(X) \to [0,1]$ by $\psi_i := \rho_i / (\sum_j \rho_j)$ with $\rho_i(y) := \max\{0, \varepsilon - \|y - y_i\|\}$, so $\psi_i = 0$ outside $B_\varepsilon(y_i)$ and $\sum_i \psi_i = 1$ on $k(X)$. Now define $\widetilde{k} : X \to E$ by

$$\widetilde{k}(x) := \sum_{i=1}^n \psi_i\bigl(k(x)\bigr) y_i.$$

Its image is contained in $F := \operatorname{span}\{y_1, \dots, y_n\}$ and bounded, hence relatively compact, and for all $x \in X$

$$\begin{aligned} \|\widetilde{k}(x) - k(x)\| &= \|\sum_{i=1}^n \psi_i\bigl(k(x)\bigr)\bigl(y_i - k(x)\bigr)\| \\ &\leq \sum_{i=1}^n \psi_i\bigl(k(x)\bigr)\|y_i - k(x)\| < \varepsilon \end{aligned}$$

because $\|y_i - k(x)\| < \varepsilon$ whenever $\psi_i\big(k(x)\big) > 0$. □

Let us now fix a Banach space E. For each bounded open subset $U \subset E$ let $\mathcal{K}(\overline{U}, E)$ denote the space of compact continuous maps $k : \overline{U} \to E$. An *admissible triple* $(\mathbb{1} - k, U, y)$ consists of

- a bounded open subset $U \subset E$;
- a map $\mathbb{1} - k : \overline{U} \to E$ with $k \in \mathcal{K}(\overline{U}, E)$;
- a value $y \in E \setminus f(\partial U)$.

Theorem 3.31 (Leray–Schauder degree)**.**
There exists a unique function, called the Leray–Schauder degree, *which associates to each admissible triple a value $D(\mathbb{1} - k, U, y) \in \mathbb{Z}$ with the following four properties:*
(Normalization) $D(\mathbb{1}, U, y) = 1$ for $y \in U$.
(Excision) For an open subset $V \subset U$ with $(\mathbb{1} - k)^{-1}(y) \subset V$,

$$D(\mathbb{1} - k, U, y) = D(\mathbb{1} - k, V, y).$$

(Additivity) If $U = U_1 \amalg U_2$ then

$$D(\mathbb{1} - k, U, y) = D(\mathbb{1} - k, U_1, y) + D(\mathbb{1} - k, U_2, y).$$

(Homotopy) For a compact map $k : [0,1] \times \overline{U} \to E$, $(t, x) \mapsto k_t(x)$, and a continuous path $y : [0,1] \to E$ such that $y(t) \notin (\mathbb{1} - k_t)(\partial U)$ for all $t \in [0,1]$,

$$D(\mathbb{1} - k_0, U, y_0) = D(\mathbb{1} - k_1, U, y_1).$$

Moreover, the Leray–Schauder degree satisfies:
(Solution) If $d(\mathbb{1} - k, U, y) \neq 0$, then there exists $x \in U$ with $x - k(x) = y$.
(Brouwer) If there exists a finite dimensional linear subspace $E_1 \subset E$ with $y \in E_1$ and $k(\overline{U}) \subset E_1$, then

$$D(\mathbb{1} - k, U, y) = d\big((\mathbb{1} - k)|_{\overline{U} \cap E_1}, y\big)$$

where the right hand side is the Brouwer degree.

Proof. We first prove uniqueness. Let $(\mathbb{1}-k, U, y)$ be an admissible triple. By Problem 3.4 the map $\mathbb{1} - k : \overline{U} \to E$ is proper, hence closed, so $\varepsilon := \operatorname{dist}\big(y, (\mathbb{1}-k)(\partial U)\big) > 0$. By Proposition 3.30 there exists a compact continuous map $k_1 : \overline{U} \to E$ with $\|k_1 - k\|_{C(\overline{U})} < \varepsilon$ such that $k_1(\overline{U}) \subset E_1$ for a finite dimensional subspace $E_1 \subset E$. After increasing E_1 we may assume $y \in E_1$. Set $U_1 := U \cap E_1$.

Claim. $D(\mathbb{1} - k, U, y) = d\big((\mathbb{1} - k_1)|_{\overline{U}_1}, y\big)$, where the right hand side is the Brouwer degree.

To see this, note first that (Homotopy) applied to $k_t(x) = (1 - t)k(x) + tk_1(x)$ yields $D(\mathbb{1} - k, U, y) = D(\mathbb{1} - k_1, U, y)$. Let $P_i : E \to E_i$ be the projections with respect to a splitting $E = E_1 \oplus E_2$ (which exists because $\dim E_1 < \infty$). Pick any continuous extension $\widetilde{k}_1 : E_1 \to E_1$ of $k_1|_{\overline{U}_1}$ and consider the homotopy $h_t := (1 - t)k_1 + t\widetilde{k}_1 P_1 : \overline{U} \to E$. Since y and $h_t(x)$ are contained in E_1, the equation $x - h_t(x) = y$ for $x \in \overline{U}$ implies $x \in \overline{U}_1$, and therefore $h_t(x) = (1-t)k_1(x) + tk_1(x) = k_1(x)$ for all $t \in [0, 1]$. In particular, $y \notin h_t(\partial U)$ and (Homotopy) gives $D(\mathbb{1} - k_1, U, y) = D(\mathbb{1} - \widetilde{k}_1 P_1, U, y)$. Moreover, since $(\mathbb{1} - \widetilde{k}_1 P_1)^{-1}(y) \subset U_1$, by (Excision) we can replace U by any bounded open set containing U_1, for example $U_1 + B$ for the open unit ball $B \subset E_2$ around the origin, and obtain

$$D(\mathbb{1}-k, U, y) = D(\mathbb{1}-\widetilde{k}_1 P_1, U_1+B, y) = D(\mathbb{1}-k_1 P_1, U_1+B, y).$$

Now to every bounded open subset $V \subset E_1$, continuous map $f : \overline{V} \to E_1$ and $z \in E_1 \setminus f(\partial V)$ we associate the integer

$$d_1(f, z) := D\big(\mathbb{1} - (\mathbb{1} - f)P_1, V + B, z\big) = D(fP_1 + P_2, V + B, z).$$

The first 4 properties of the Leray–Schauder degree D imply that d_1 satisfies the first 4 properties of the Brouwer degree on bounded open subsets of E_1, so by uniqueness $d_1(f, z)$ agrees with the

Brouwer degree $d(f,z)$. Applying this to $f := (\mathbb{1} - k_1)|_{\overline{U}_1}$ and $z := y$ we obtain $D(\mathbb{1} - k_1 P_1, U_1 + B, y) = d\big((\mathbb{1} - k_1)|_{\overline{U}_1}, y\big)$. This proves the claim, and thus uniqueness as well as the (Brouwer) and (Solution) properties.

For existence, we *define* the Leray–Schauder degree via the Brouwer degree by $D(\mathbb{1} - k, U, y) := d\big((\mathbb{1} - k_1)|_{\overline{U} \cap E_1}, y\big)$ for any pair (k_1, E_1), where $k_1 : \overline{U} \to E$ is a compact continuous map such that $\|k_1 - k\|_{C(\overline{U})} < \varepsilon$ and $k_1(\overline{U}) \subset E_1$ for a finite dimensional subspace $E_1 \subset E$ with $y \in E_1$. To show that this is well-defined, consider another such pair (k_2, E_2). Set $E_0 := E_1 + E_2$ and $U_i := U \cap E_i$ for $i = 0, 1, 2$. Then

$$\begin{aligned} d\big((\mathbb{1} - k_1)|_{\overline{U}_1}, y\big) &= d\big((\mathbb{1} - k_1)|_{\overline{U}_0}, y\big) = d\big((\mathbb{1} - k_2)|_{\overline{U}_0}, y\big) \\ &= d\big((\mathbb{1} - k_2)|_{\overline{U}_2}, y\big), \end{aligned}$$

where the first and third equality follow from Problem 3.17 below, and the second equality follows from homotopy invariance of the Brouwer degree applied to the homotopy $f_t := \big(\mathbb{1} - (1-t)k_1 - tk_2\big)|_{\overline{U}_0}$. This concludes the proof of Theorem 3.31. □

Problem 3.17 (Dimension reduction for the Brouwer degree).
Let E be a finite dimensional vector space, $U \subset E$ a bounded open subset, and $k : \overline{U} \to E$ a continuous map. Prove: if $k(\overline{U}) \subset E_1$ for a linear subspace $E_1 \subset E$, then for each $y \in E_1 \setminus (\mathbb{1} - k)(\partial U)$ we have

$$d(\mathbb{1} - k, y) = d\big((\mathbb{1} - k)|_{\overline{U} \cap E_1}, y\big).$$

Problem 3.18. Let $(\mathbb{1} - k, U, y)$ be an admissible triple with $k \in C^1(U, E)$. Show:
(a) If y is a regular value of $\mathbb{1} - k$, then its Leray–Schauder degree is given by

$$D(\mathbb{1} - k, U, y) = \sum_{x \in (\mathbb{1} - k)^{-1}(y)} (-1)^{m(x)},$$

where $m(x)$ is the sum of the algebraic multiplicities of the real eigenvalues $\lambda > 1$ of $Dk(x)$.
(b) The Leray–Schauder degree $D(\mathbb{1} - k, U, y)$ modulo 2 agrees with the Caccioppoli–Smale degree.

Remark 3.32. Alternatively, one could define the Leray–Schauder degree of an admissible triple $(\mathbb{1} - k, U, y)$ as follows: approximate k by a C^1-map $\widetilde{k}$ and y by a regular value $\widetilde{y}$ and set $D(\mathbb{1} - k, U, y) := D(\mathbb{1} - \widetilde{k}, U, \widetilde{y})$, where the latter expression is defined by the formular in Problem 3.18. To find the C^1-approximation $\widetilde{k}$ we can use Proposition 3.30 to approximate k by a finite rank map $k_1 : \overline{U} \to F$, and then apply Corollary 2.14 to approximate k_1 by a C^1-map. However, this definition has the slight disadvantage that it requires a C^1-partition of unity on the Banach space E.

Applications. The proofs of the following four corollaries are straightforward adaptations of the proofs in the finite dimensional case.

Corollary 3.33 (Schauder fixed point theorem)**.** *Every compact map $k : B \to B$ on the closed unit ball $B \subset E$ in a Banach space has a fixed point.*

Corollary 3.34 (Borsuk-Ulam theorem for $\mathbb{1} - k$)**.**
Let $U \subset E$ be a bounded open subset of a Banach space with $-U = U$, and $k : \overline{U} \to E$ be compact with $0 \notin (\mathbb{1} - k)(\partial U)$ and $k(-x) = -k(x)$ for all $x \in U$. Then $D(\mathbb{1} - k, U, 0) \equiv 1 \mod 2$.

Corollary 3.35 (Borsuk-Ulam antipodal theorem for $\mathbb{1} - k$)**.**
Let E be a Banach space, $B \subset E$ the closed unit ball, and $F \subsetneqq E$ a proper linear subspace. Let $k : \partial B \to E$ be a compact continous map such that $(\mathbb{1} - k)(\partial B) \subset F$. Then there exists some $x \in \partial B$ such that $(\mathbb{1} - k)(x) = (\mathbb{1} - k)(-x)$.

Corollary 3.36 (Invariance of domain for $\mathbb{1} - k$)**.**
Let $U \subset E$ be an open subset of a Banach space and $k : \overline{U} \to E$ be a continuous map sending bounded sets to relative compact sets. If $\mathbb{1} - k$ is locally injective, then $\mathbb{1} - k$ is open.

Problem 3.19. Prove Corollaries 3.33, 3.34, 3.35 and 3.36.
Hint: For Corollary 3.34 approximate k by a map h with $h(-x) = -h(x)$ and finite dimensional image. For Corollary 3.35 extend k to a compact map $\bar{k} : B \to E$ by $\bar{k}(tx) := tk(x)$ for $x \in \partial B$ and $t \in [0, 1]$.

Problem 3.20. (a) Prove *Peano's existence theorem for ODEs:* Let $B^n(x_0, r) \subset \mathbb{R}^n$ be the open ball or radius r around $x_0 \in \mathbb{R}^n$ and $f : [a, b] \times \overline{B^n(x_0, r)} \to \mathbb{R}^n$ a continuous map. Then for every $t_0 \in (a, b)$ there exists $\delta > 0$ and a differentiable solution $x : [t_0 - \delta, t_0 + \delta] \to B^n(x_0, r)$ of the initial value problem

$$\dot{x}(t) = f\big(t, x(t)\big), \qquad x(t_0) = x_0.$$

Hint: For $I := [t_0 - \delta, t_0 + \delta]$ consider the closed ball $B := \{x \in E \mid \max_{t \in I} |x(t) - x_0| \le r\}$ in the Banach space $E := C(I, \mathbb{R}^n)$ and apply the Schauder fixed point theorem to the map

$$k : B \to E, \qquad k(x)(t) := x_0 + \int_{t_0}^{t} f\big(s, x(s)\big) ds.$$

(b) Find an example to which Peano's theorem applies but the solution is not unique.

3.7 An application in symplectic topology

In this subsection we discuss an application of the Caccioppoli–Smale degree to Hamiltonian dynamics. The presentation will be very informal; we refer to [35] for background on symplectic topology and to the original article [10] for detailed proofs.

Let M be a closed n-dimensional manifold. Let $\tau : T^*M \to M$ its cotangent bundle equipped with the canonical Liouville 1-form $\lambda = p\,dq$ and symplectic form $\omega = d\lambda = dp \wedge dq$. To a 1-periodic time-dependent smooth Hamiltonian $H : S^1 \times T^*M \to \mathbb{R}$ we associate the time-dependent Hamiltonian vector field X_{H_t} defined by $i_{X_{H_t}}\omega = -dH_t$, where $H_t(x) = H(t,x)$ and $S^1 = \mathbb{R}/\mathbb{Z}$. We pick a Riemannian metric on M and impose the following quadratic growth condition on the Hamiltonian for constants $\kappa, d_1, d_2 > 0$:

$$\langle p, \frac{\partial H}{\partial p}\rangle \geq \kappa |p|^2 - d_1 \quad \text{and} \quad |\frac{\partial^2 H}{\partial p^2}| + |\frac{\partial^2 H}{\partial p \partial q}| \leq d_2. \tag{3.2}$$

Theorem 3.37. *[10] Let $H : S^1 \times T^*M \to \mathbb{R}$ satisfy condition* (3.2). *If $\pi_1(M)$ is finite, then the Hamiltonian system $\dot{x} = X_{H_t}(x)$ possesses infinitely many contractible* 1*-periodic orbits.*

Here are some remarks on this result.

(1) Condition (3.2) can be replaced by a weaker and symplectically invariant one [12]. It is satisfied e.g. for classical Hamiltonians $H(t,q,p) = \frac{1}{2}|p - A(t,q)|^2 + V(t,q)$ with time-dependent vector potential A and scalar potential V.

(2) Theorem 3.37 fails in general if $\pi_1(M)$ is infinite:

Problem 3.21. Show: for $0 < \omega < 2\pi$ the planar pendulum with $M = S^1$ and autonomous Hamiltonian $H = \frac{1}{2}p^2 + \omega^2 \cos q$ has exactly two contractible 1-periodic orbits, the stable and unstable equilibria.

(3) Let $\pi_1(M)$ be arbitrary and denote by $\Lambda := H^{1,2}(S^1, M)$ the free loop space of M. For a connected component Λ_a of Λ denote by

$\mathcal{C}_a$ the set of 1-periodic orbits whose projection under $\tau : T^*M \to M$ belongs to Λ_a. Then we still get the estimate (see [10, 12])

$$\#\mathcal{C}_a \geq \mathrm{cuplength}_{\mathbb{Z}_2}(\Lambda_a) + 1 \geq 1,$$

where the $\mathbb{Z}_2$-*cuplength* of a topological space is the supremum of all $k \in \mathbb{N}_0$ such that there exist cohomology classes $\alpha_i \in H^{n_i}(X; \mathbb{Z}_2)$ with $n_i \geq 1$ and $\alpha_1 \cup \dots \cup \alpha_k \neq 0$.

Problem 3.22. (a) Show that the component $\Lambda_0 \subset \Lambda$ of contractible loops satisfies $\mathrm{cuplength}_{\mathbb{Z}_2}(\Lambda_0) \geq \mathrm{cuplength}_{\mathbb{Z}_2}(M)$.
(b) Show that each component of the loop space of the n-torus $M = T^n$ has cuplength n.
(c) The system on T^n given by n uncoupled pendula with frequencies $\omega_i < 2\pi$ has exactly 2^n contractible 1-periodic orbits. Try to find a system which realizes the cuplength bound of $n + 1$ contractible orbits.

(4) Similar methods can be used to prove cuplength estimates for Lagrangian intersections [20].

The remainder of this section is devoted to the proof of Theorem 3.37.

The space of Floer cylinders. We now return to the setup of Theorem 3.37. One-periodic orbits of $\dot{x} = X_{H_t}(x)$ are exactly the critical points of the Hamiltonian action

$$\mathcal{A}_H : C^\infty(S^1, T^*M), \qquad \mathcal{A}_H(x) := \int_0^1 \big(x^*\lambda - H(t, x)dt\big).$$

We identify $TM \cong T^*M$ using the fixed a Riemannian metric on M. Let J be the ω-compatible almost complex structure with matrix $\left(\begin{smallmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{smallmatrix}\right)$ with respect to the splitting $TTM = T^vTM \oplus T^hTM$ into vertical and horizontal subspaces defined by the Levi-Civita connection. Let $\langle\ ,\ \rangle = \omega(\cdot, J\cdot)$ be the induced Riemannian metric on TM. The gradient of $\mathcal{A}_H$ with respect to the corresponding L^2-metric equals

$$\nabla_{L^2}\mathcal{A}_H(x) = -J(x)\big(\dot{x} - X_{H_t}(x)\big) = -J(x)\dot{x} + \nabla H_t(x),$$

so L^2-gradient trajectories are smooth maps $u : \mathbb{R} \times S^1 \to TM$ satisfying the Floer equation

$$\bar{\partial}_H(u) := \partial_s u + J(u)\partial_t u + \nabla H_t(u) = 0. \tag{3.3}$$

Their *energy* is given by

$$E(u) := \int_{\mathbb{R}\times S^1} |\partial_s u|^2 ds\, dt = \lim_{s\to\infty} \mathcal{A}_H\big(u(s)\big) - \lim_{s\to-\infty} \mathcal{A}_H\big(u(s)\big),$$

where we write $u(s) := u(s,\cdot) : S^1 \to TM$ and the limits are well-defined (though possibly infinite) because

$$\frac{d}{ds}\mathcal{A}_H\big(u(s)\big) = \int_0^1 |\partial_s u(s,t)|^2 dt \geq 0.$$

The space of finite energy gradient trajectories

$$X := \{u \in C^\infty(\mathbb{R} \times S^1, TM) \mid (3.3),\ E(u) < \infty\}$$

becomes a metric space equipped with the C^∞_{loc}-topology. It carries a continuous flow $\mathbb{R} \times X \to X$ given by negative time shift

$$(\sigma \cdot u)(s,t) := u(s-\sigma, t),$$

and a continuous function

$$\Phi : X \to \mathbb{R}, \qquad \Phi(u) := \mathcal{A}_H\big(u(0)\big).$$

Lemma 3.38. *The pair (X, Φ) has the following properties:*
(a) Rest points of the flow on X are precisely the critical points of $\mathcal{A}_H$, whose set we denote by $\mathrm{Crit}(\mathcal{A}_H)$.
(b) For each $u \in X$ and each sequence $s_k \to \pm\infty$, some subsequence of $u(s_k)$ converges to some $x_\pm \in$ $\mathrm{Crit}(\mathcal{A}_H)$.
(c) Φ is a strict Lyapunov function for the flow on X.
(d) $\Phi : X \to \mathbb{R}$ is bounded from below.

Proof. Part (b) follows from the Palais-Smale condition for $\mathcal{A}_H$, strictness in (c) uses unique continuation, and (d) follows from (c) and the fact that $\mathcal{A}_H$ is bounded from below on $\mathrm{Crit}(\mathcal{A}_H)$. For details see [10]. □

Consider now the continuous evaluation map

$$\pi : X \to \Lambda = W^{1,2}(S^1, M), \qquad u \mapsto \tau \circ u(0)$$

and the function

$$\Phi_+ : X \to \mathbb{R}, \qquad \Phi_+(u) := \lim_{s\to\infty} \mathcal{A}_H\big(u(s)\big).$$

By Lemma 3.38(b) and (d) the function Φ_+ takes values in $\mathrm{Crit}(\mathcal{A}_H)$ and is bounded from below (since $\Phi \le \Phi_+$). It is not continuous but *lower semicontinuous*, i.e., $\Phi_+(u) \le \liminf_{k\to\infty} \Phi_+(u_k)$ whenever $u_k \to u$ in X.

In the following proposition $\check{H}^*$ denotes Alexander-Spanier cohomology [48], which on a metric space is equivalent to Cech cohomology. Its advantage over singular cohomology is its *tautness* property $\check{H}^*(A) = \varinjlim \check{H}^*(U)$ for every subset $A \subset X$ of a metric space X, where the direct limit is taken over all open subsets $U \subset X$ containing A, see [49].

Proposition 3.39. *(A) The function $\Phi_+ : X \to \mathbb{R}$ is proper, so the set $\Phi_+^b := \{u \in X \mid \Phi_+(u) \le b\}$ is compact for each $b \in \mathbb{R}$.*
(B) The induced map $\pi^*\check{H}^*(\Lambda; \mathbb{Z}_2) \to \check{H}(X; \mathbb{Z}_2)$ is injective. More precisely, for each $0 \ne \alpha \in \check{H}^*(\Lambda; \mathbb{Z}_2)$ there exists $b_\alpha \in \mathbb{R}$ such that $0 \ne (\pi|_{\Phi_+^{b_\alpha}})^*\alpha \in \check{H}^*(\Phi_+^{b_\alpha}; \mathbb{Z}_2)$.

Proof of Theorem 3.37 assuming Proposition 3.39. Suppose that $\pi_1(M)$ is finite. After replacing M by its (compact) universal cover, we may assume that $\pi_1(M) = 0$. For $s \in \mathbb{R}$ define

$$\Phi_s : X \to \mathbb{R}, \qquad \Phi_s(u) := \mathcal{A}_H\big(u(s)\big).$$

For each $0 \neq \alpha \in \check{H}^*(\Lambda; \mathbb{Z}_2)$ choose b_α as in Proposition 3.39(B) and define

$$\begin{aligned} \mathcal{F}_\alpha &:= \{A \subset \Phi_+^{b_\alpha} \text{ compact} \mid (\pi_A)^*\alpha \neq 0\}, \\ c_\alpha^s &:= \inf_{A \in \mathcal{F}_\alpha} \sup_A \Phi_s \in \mathbb{R}. \end{aligned}$$

Problem 3.23. For $0 \neq \alpha \in \check{H}^*(\Lambda; \mathbb{Z}_2)$ and $s \in \mathbb{R}$ show:
(i) $\mathcal{F}_\alpha$ is nonempty and invariant under the flow on X;
(ii) $-\infty < c_\alpha^s = c_\alpha^0 < \infty$;
(iii) there exists a rest point $u \in X$ of the flow with $\Phi_s(u) = c_\alpha^s$.

Since the rest point u in (iii) is a 1-periodic orbit x with $\mathcal{A}_H(x) = c_\alpha$, the theorem follows from the following

Claim. The values c_α^0 for $0 \neq \alpha \in \check{H}^*(\Lambda; \mathbb{Z}_2)$ are not bounded from above.

Arguing by contradiction, suppose the values $c_\alpha^s = c_\alpha^0$ are $\leq c - 1$ for some $c \in \mathbb{R}$. By definition of c_α^s, the compact set $K_{\alpha,s}^c := \Phi_+^{b_\alpha} \cap \Phi_s^c$ satisfies $(\pi|_{K_{\alpha,s}^c})^*\alpha \neq 0$. By weak continuity of Alexander-Spanier cohomology [48, Theorem 6.6.6] we have $\check{H}^*(\bigcap_{s \geq 0} K_{\alpha,s}^c; \mathbb{Z}_2) \cong \varinjlim_{s \to \infty} \check{H}^*(K_{\alpha,s}^c; \mathbb{Z}_2)$, so it follows that $(\pi|_{\bigcap_{s \geq 0} K_{\alpha,s}^c})^*\alpha \neq 0$. But $\bigcap_{s \geq 0} K_{\alpha,s}^c \subset \Phi_+^c$, hence $(\pi|_{\Phi_+^c})^*\alpha \neq 0$, and therefore $i_K^*\alpha \neq 0$ for the inclusion $i_K : K := \pi(\Phi_+^c) \hookrightarrow \Lambda$. For $\varepsilon > 0$ smaller than the injectivity radius and $k \in \mathbb{N}$ let

$$P_k M := \{(p_1, \ldots, p_k) \in M^k \mid \sum_{i=1^k} d(p_{i-1}, p_i)^2 \leq \varepsilon^2\},$$

where $p_0 := p_k$, and denote by $g_k : P_k M \hookrightarrow \Lambda$ the inclusion as piecewise geodesics. Since $K \subset \Lambda$ is compact, there exists a $k \in \mathbb{N}$

and a continuous map $r : K \to P_k M$ with $g_k \circ r = i_K$ (see [6]). Then the preceding discussion implies that $g_k^* : \check{H}^*(\Lambda; \mathbb{Z}_2) \to \check{H}^*(P_k M; \mathbb{Z}_2)$ is injective, so $\check{H}^*(\Lambda; \mathbb{Z}_2)$ is finite dimensional because $P_k M$ is a compact finite dimensional manifold with boundary. But this contradicts Sullivan's theorem that for a simply connected closed manifold M infinitely many Betti numbers of its loops space Λ are nonzero. This proves the claim and thus Theorem 3.37. □

Approximation by finite Floer cylinders. It remains to prove Proposition 3.39. We will not say anything on part (A), which involves analysis of Floer cylinders using hypothesis (H) on the Hamiltonian H, but rather explain the proof of part (B). The idea is to approximate the space X of Floer cylinders by suitable spaces X_R of finite Floer cylinders and let their length go to infinity. Fix $R > 0$ and consider the finite cylinder $Z_R := [-R, R] \times S^1$. For each $q \in C^\infty(S^1, M)$ consider the Hilbert manifold

$$\mathcal{E}_q := \{u \in W^{2,2}(Z_R, TM) \mid u(-R,t) \in M,\ u(R,t) \in T_{q(t)}M\ \forall t\}.$$

The perturbed Cauchy-Riemann operator $\bar\partial_H$ defines a section in the Hilbert space bundle $\mathcal{F} \to \mathcal{E}_q$ with fibres $\mathcal{F}_u = W^{1,2}(Z_R, u^*TTM)$. Since by Kuiper's theorem the bundle $\mathcal{F}$ is isomorphic to the trivial bundle $\mathcal{E}_q \times F$ with $F = W^{1,2}(Z_R, \mathbb{C}^n)$, we can view $\bar\partial_H$ as a map $\mathcal{E}_q \to F$.

Lemma 3.40. *The map $\bar\partial_H : \mathcal{E}_q \to F$ is a smooth Fredholm map of index zero.*

Let $\mathcal{J} : \Lambda \to C^\infty(S^1, M)$ be a smoothing operator and consider the Hilbert manifold

$$\mathcal{E} := \{(u,q) \in W^{2,2}(Z_R, TM) \times \Lambda \mid u \in \mathcal{E}_{\mathcal{J}q}\}.$$

By the preceding lemma the map

$$f : \mathcal{E} \to F \times \Lambda, \qquad (u, q) \mapsto (\bar\partial_H u, q)$$

is a Fredholm map of index zero.

Proposition 3.41. *For each compact subset $K \subset \Lambda$ there exists an open neighbourhood U of $f^{-1}(\{0\} \times K)$ such that $f|_{\overline{U}}$ is proper with Caccioppoli–Smale degree $d\big(f|_{\overline{U}}, (0, q)\big) = 1$ for all $q \in K$.*

Sketch of proof. Using condition (3.2) and elliptic regularity one shows that $f^{-1}(\{0\} \times K)$ is compact. Since Fredholm operators are nocally proper, there exists an open neighbourhood U of $f^{-1}(\{0\} \times K)$ such that $f|_{\overline{U}}$ is proper.

To compute the degree of $f|_{\overline{U}}$, we homotope f to the map $f_0(u, q) := (\bar\partial u, q)$, where $\bar\partial u = \partial_s u + J(u)\partial_t u$ is the Floer operator without Hamiltonian term. After enlarging K, we may assume that K is path connected and contains a constant loop $q_0 \in M$. Consider $u \in f_0^{-1}(0, q_0)$. Its energy is

$$\int_{Z_R} |\partial_s u|^2 ds\, dt = \int_0^1 u(R)^*\lambda - \int_0^1 u(-R)^*\lambda = 0$$

because $u(R, t) \in T_{q_0}M$ and $u(-R, t) \in M$, so both integrals of $\lambda = p\, dq$ vanish. Thus $\partial_s u \equiv 0$, and $\bar\partial u = 0$ together with the boundary conditions implies that u is the constant map $u_0 = (q_0, 0) \in T^*M$. By Problem 3.24 below the differential $T_{u_0}f$ is an isomorphism, so $(0, q_0)$ is a regular value of f_0 with $f^{-1}(0, q_0) = \{u_0\}$ and homotopy invariance implies

$$d\big(f|_{\overline{U}}, (0, q)\big) = d\big(f_0|_{\overline{U}}, (0, q_0)\big) = 1.$$

□

Problem 3.24. For $R > 0$ consider the Hilbert space

$$E := \{\xi \in W^{2,2}(Z_R, \mathbb{C}^n) \mid \xi(-R,t) \in \mathbb{R}^n,\ \xi(R,t) \in i\mathbb{R}^n\ \forall t\}.$$

Show that the linear operator

$$\bar\partial : E \to F = W^{1,2}(Z_R, \mathbb{C}^n), \qquad \bar\partial u = \partial_s u + i\bar\partial_t u$$

is an isomorphism. *Hint: Use Fourier series and the adjoint operator.*

Consider now the space of finite Floer cylinders

$$X_R := \{u \in C^\infty(Z_R, TM) \mid (3.3),\ u(-R,t) \in M\ \forall t\}$$

with the continuous map

$$\pi_R : X_R \to \Lambda, \qquad u \mapsto \tau \circ u(R).$$

Recall the smoothing operator $\mathcal{J}$ and the inclusion $g_m : P_m M \hookrightarrow \Lambda$ with image $K_m := g_m(P_m M)$.

Proposition 3.42. *For each $0 \neq \alpha \in \check{H}^*(\Lambda; \mathbb{Z}_2)$ there exists $m \in \mathbb{N}$ such that $\big(\pi_R|_{\pi_R^{-1}(\mathcal{J}K_m)}\big)^* \alpha \neq 0$ for all $R > 0$.*

Proof. Pick $m \in \mathbb{N}$ such that $g_m^* \alpha \neq 0$ and any $R > 0$. By Proposition 3.41 there exists an open neighbourhood U of $f^{-1}(\{0\} \times K_m)$ such that $f|_{\overline{U}}$ is proper with $d\big(f|_{\overline{U}}, (0,q)\big) = 1$ for all $q \in K_m$. Pick a C^1-approximation $h : K_m \hookrightarrow F \times \Lambda$ of the inclusion map $q \mapsto (0,q)$ which is transverse to $f|_{\overline{U}}$. Then $B := h(K_m)$ and $A := f|_{\overline{U}}^{-1}(B)$ are compact manifolds-with-boundary of the same dimension and $f|_A : A \to B$ is a C^1-map of degree 1. By Problem 3.25 below $(f|_A)^* : \check{H}^*(B; \mathbb{Z}_2) \to \check{H}^*(A; \mathbb{Z}_2)$ is injective, so $0 \neq (f|_A)^* \pi_\Lambda^* \alpha \in \check{H}^*(A; \mathbb{Z}_2)$, where $\pi_\Lambda : F \times \Lambda \to \Lambda$ denotes the projection. Now $A \subset U$ implies $(f|_U)^* \pi_\Lambda^* \alpha \neq 0$. Since this holds for every sufficiently small neighbourhood U of

$f^{-1}(\{0\} \times K_m)$, tautness of Alexander-Spanier cohomology yields $(f|_{f^{-1}(\{0\}\times K_m)})^* \pi_\Lambda^* \alpha \neq 0$. From the commuting diagram

$$\begin{tikzcd} f^{-1}(\{0\} \times K_m) \arrow[r, "p_1"] \arrow[d, "f"] & \pi_R^{-1}(\mathcal{J} K_m) \arrow[d, "\pi_R"] \\ F \times \Lambda \arrow[r, "\mathcal{J} \circ \pi_\Lambda"] & \Lambda, \end{tikzcd}$$

where $p_1(u, q) = u$ and $\mathcal{J}$ is homotopic to the identity, we finally conclude $\left(\pi_R|_{\pi_R^{-1}(\mathcal{J} K_m)}\right)^* \alpha \neq 0$. $\square$

Problem 3.25. Let A, B be compact n-dimensional manifolds, possibly with boundary. Show: if $f : A \to B$ is a C^1-map with $f(\partial A) \subset \partial B$ and Brouwer degree $d(f, b) = 1$ for all $b \in B \setminus \partial B$, then $f^* : H^*(B; \mathbb{Z}_2) \to H^*(A; \mathbb{Z}_2)$ is injective. *Hint: Show that f_* maps the fundamental class $[A] \in H_n(A, \partial A; \mathbb{Z}_2)$ to $[B]$ and use Poincaré-Lefschetz duality.*

Proof of Proposition 3.39(B). Let $0 \neq \alpha \in \check{H}^*(\Lambda; \mathbb{Z}_2)$ be given. Pick $m \in \mathbb{N}$ as in Proposition 3.42 so that $(\pi_R|_{X_{R,m}})^* \alpha \neq 0$ for all $R > 0$, where we have set $X_{R,m} := \pi_R^{-1}(\mathcal{J} K_m)$. We pick a smooth function $\phi_R : \mathbb{R} \to [-R, R]$ with $\phi(s) = s$ for $|s| \leq R - 1$ and define a continuous map

$$P : X_{R,m} \to C^\infty(\mathbb{R} \times S^1, TM), \qquad Pu(s,t) := u\big(\phi_R(s), t\big).$$

Now we invoke two nontrivial analytical results. First, there exists $b_\alpha \in \mathbb{R}$ such that $\mathcal{A}_H\big(u(R)\big) \leq b_\alpha$ for all $u \in X_{R,m}$ and all $R > 0$. Second, for every open neighbourhood U of $\Phi_+^{b_\alpha}$ in $C^\infty(\mathbb{R} \times S^1, TM)$ there exists some (large) $R > 0$ with $P(X_{R,m}) \subset U$. For such R we get a commuting diagram

$$\begin{tikzcd} X_{R,m} \arrow[r, "P"] \arrow[d, "\pi"] & U \arrow[d, "\pi"] \\ \Lambda \arrow[r, "\mathrm{id}"] & \Lambda, \end{tikzcd}$$

where $\pi u(t) = \tau \circ u(0,t)$. Since $\pi \sim \pi_R : X_{R,m} \to \Lambda$ and $(\pi_R|_{X_{R,m}})^*\alpha \neq 0$, it follows that $(\pi|_U)^*\alpha \neq 0$ for each open neighbourhood U of $\Phi_+^{b_\alpha}$, hence $(\pi|_{\Phi_+^{b_\alpha}})^*\alpha \neq 0$ by tautness of Alexander-Spanier cohomology. This concludes the proof of Proposition 3.39(B). $\square$

3.8 Framed degree

Discuss results in [1].

3.9 Elliptic PDEs

Part II

Polyfolds

Chapter 4

Fréchet manifolds

4.1 Fréchet spaces and their linear maps

Unless otherwise stated all vector spaces will be over $\mathbb{R}$. All definitions and results carry over to $\mathbb{C}$ with minor adjustments.

Definition 4.1. A *topological vector space* is a vector space E with a topology such that points are closed, and scalar multiplication $\mathbb{R} \times E \to E$ and addition $E \times E \to E$ are continuous maps.

Problem 4.1. Let E be a topological vector space. Show:
(a) The closure of a linear subspace is again a linear subspace.
(b) If $U \subset E$ is open then so are $x + U$ and tU for all $x \in E$ and $t \in \mathbb{R} \setminus \{0\}$.
(c) For each open neighbourhood $W \subset E$ of 0 there exists an open neighbourhood U of 0 which is symmetric (i.e. $U = -U$) and satisfies $U + U \subset W$.
(d) For $A \subset E$ closed and $K \subset V$ compact with $A \cap K = \emptyset$ there exists an open neighbourhood $V \subset E$ of 0 such that $(A + V) \cap (K + V) = \emptyset$; in particular, E is Hausdorff.

Definition 4.2. A *metric vector space* is a vector space E with a metric d turning it into a topological vector space and such that $d(x + z, y + z) = d(x, y)$ for all $x, y, z \in E$.

A complete metric vector space is also called an “F-space”, but we will not use this term. Most metric vector spaces arise from families of seminorms:

Definition 4.3. A *seminorm* on a vector space E is a function $\| \, \| : E \to \mathbb{R}_{\geq 0}$ satisfying $\|tx\| = |t| \, \|x\|$ and $\|x+y\| \leq \|x\|+\|y\|$ for all $x, y \in E$ and $t \in \mathbb{R}$. It is a *norm* if in addition $\|x\| = 0$ implies $x = 0$.

Consider a countable family of seminorms $\| \, \|_k$, $k \in \mathbb{N}$, on a vector space E. For $x \in E$, $K \in \mathbb{N}$ and $\varepsilon > 0$ consider the sets

$$U(x, K, \varepsilon) := \{y \in E \mid \|y - x\|_k < \varepsilon \text{ for all } k = 1, \ldots, K\}.$$

Problem 4.2. (a) The sets $U(x, K, \varepsilon)$ are the basis of a topology which makes E a topological vector space.
(b) Assume that $\|x\|_k = 0$ for all $k \in \mathbb{N}$ implies $x = 0$. Then

$$d(x, y) := \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\|x - y\|_k}{1 + \|x - y\|_k} \tag{4.1}$$

defines a metric on E which induces the topology in (a) and makes E a metric vector space.
(c) In the situation of (b), a sequence (x_n) in E converges to x with respect to the metric d iff it converges to x with respect to each seminorm (i.e. $\|x_n - x\|_k \to 0$ for each k), and it is a Cauchy sequence for d iff it is a Cauchy sequence for each seminorm.

Definition 4.4. A *Fréchet space* is a vector space E with a countable family of seminorms $\| \, \|_k$, $k \in \mathbb{N}$, such that $\|x\|_k = 0$ for all $k \in \mathbb{N}$ implies $x = 0$, and such that E is complete with respect to the metric (4.1).

Example 4.1. Let $C(X, F)$ be the space of continuous maps from a Hausdorff space X to a normed vector space $(F, \| \, \|)$.

(a) If X is compact, then $C(X, F)$ becomes a normed vector space with the norm $\|u\| := \max_{x\in X}\|u(x)\|$.
(b) If X is noncompact but has an exhaustion $\bigcup_{k\in\mathbb{N}} X_k = X$ by compact sets $X_1 \subset X_2 \subset \dots$, then $C(X, F)$ becomes a Fréchet space with the seminorms $\|u\|_k := \max_{x\in X_k}\|u(x)\|$.

Example 4.2. Let $C^\infty(N)$ be the space of smooth functions $u : N \to \mathbb{R}$ on a manifold N (possibly with boundary).
(a) If N is compact, then $C^\infty(N)$ becomes a Fréchet space with the seminorms (actually, norms) $\|u\|_k := \|u\|_{C^k(N)}$.
(b) If N is noncompact but has an exhaustion $\bigcup_{k\in\mathbb{N}} N_k = N$ by compact sets $N_1 \subset N_2 \subset \dots$, then $C^\infty(N)$ becomes a Fréchet space with the seminorms $\|u\|_k := \|u\|_{C^k(N_k)}$.

A subset $A \subset E$ of a topological vector space E is called *bounded* if for each open neighbourhood $V \subset E$ of 0 there exists $s > 0$ such that $A \subset tV$ for all $t > s$. In a Fréchet space this is equivalent to each seminorm $\| \ \|_k$ being bounded on A; note that it is *not* equivalent to A having bounded diameter with respect to the metric (4.1).

Problem 4.3. For the space $C^\infty(N)$ in Example 4.2 prove:
(a) A nonempty bounded subset $A \subset C^\infty(N)$ is never open.
(b) $C^\infty(N)$ has the *Heine-Borel property*, i.e., every bounded closed subset is compact.

Linear maps. Let E, F be topological vector spaces. A linear map $T : E \to F$ is called *bounded* if it maps bounded sets to bounded sets.

Problem 4.4. Let $T : E \to F$ be a linear map between topological vector spaces. Prove:
(a) T is continuous iff it is continuous at 0, i.e., $T^{-1}(V)$ is open for each open neighbourhood $V \subset F$ of 0.

(b) If T is continuous then it is bounded, and the converse holds if E is Fréchet.

Let $\mathcal{L}(E, F)$ denote the space of all continuous linear maps $T : E \to F$. For a bounded subset $A \subset E$ and an open neighbourhood $V \subset F$ of 0 let

$$\mathcal{U}(A, V) := \{T \in \mathcal{L}(E, F) \mid T(A) \subset V\}.$$

Translates of these sets are the basis of a topology on $\mathcal{L}(E, F)$ which we call the *bounded-open topology.*[1] If E, F are normed vector spaces, then this topology agrees with the one induced by the operator norm.

Problem 4.5. For toplogical vector spaces E, F consider $\mathcal{L}(E, F)$ with the bounded-open topology. Show:
(a) $\mathcal{L}(E, F)$ is a topological vector space.
(b) If E, F are Fréchet spaces, then the topology on $\mathcal{L}(E, F)$ is generated by an uncountable family of seminorms, but in general $\mathcal{L}(E, F)$ has no countable neighbourhood basis of 0 (hence it is not Fréchet).

The notions of isomorphisms, compact operators, Fredholm operators and projections carry over verbatim to the Fréchet setting. The open mapping theorem, closed graph theorem and the Banach-Steinhaus theorem continue to hold for continuous linear maps between Fréchet spaces, see [42]. By contrast, many results concerning families of linear operators fail in the Fréchet setting. The following example (to which we will return later) shows that, in contrast to Banach spaces, the following may happen for Fréchet spaces E, F:

- The space of isomorphisms $E \to F$ need not be open in $\mathcal{L}(E, F)$.

[1] It is also called the "topology of bounded convergence" or the "topology of uniform convergence on bounded sets".

• The space of linear Fredholm operators need not be open in $\mathcal{L}(E, F)$.
• There may exist continuous families $\mathbb{R} \to \mathcal{P}(E)$ of linear projections in which the dimensions of the kernel and cokernels jump (by finite or infinite amounts).

Example 4.3. Consider $X = \mathbb{R}$ with the compact subsets $X_k := [-k, k]$ for $k \in \mathbb{R}_+$. As in Example 4.1(b), the space $E := C(\mathbb{R}, \mathbb{R})$ becomes a Fréchet space with the seminorms $\|u\|_k = \|u\|_{C(X_k)}$. Fix some $\beta \in E$ with $\operatorname{supp} \beta \subset [-1, 1]$ and $\beta(0) = 1$. For $k \in \mathbb{R}_+$ define the closed subspaces

$$E_k := \{u \in E \mid u|_{X_k} \equiv 0\}, \qquad E_k^\perp := \{u \in E \mid u|_{\mathbb{R}\setminus X_k} \equiv 0\}.$$

We have a topological splitting

$$E = E_k \oplus E_k^\perp \oplus L_k,$$

where $L_k \subset E$ is the subspace spanned by the two functions $\beta(\cdot - k)$ and $\beta(\cdot + k)$. Let $P_k : E \to E$ be the projection onto E_k along $E_k^\perp \oplus L_k$. For $k \geq 1$ it is given explicitly by

$$P_k u(x) = \begin{cases} u(x) - u(k)\beta(x-k) - u(-k)\beta(x+k) & |x| > k, \\ 0 & |x| \leq k. \end{cases}$$

We claim that $P_k \to 0$ as $k \to \infty$ in $\mathcal{L}(E)$ with respect to the bounded-open topology. To see this, consider a basic open neighbourhood $\mathcal{U}(V, A)$ of 0 as above with a bounded set $A = \{u \in E \mid \|u\|_k \leq c_k \text{ for all } k \in \mathbb{N}\}$ and an open set $V = \{u \in E \mid \|u\|_k < \varepsilon_k \text{ for } k = 1, \ldots, N\}$; then $P_k(A) \subset P_k(E) \subset V$ and thus $P_k \in \mathcal{U}(A, V)$ for all $k \geq N$.
Note that $\ker(\mathbb{1} - P_k) = \operatorname{coker}(\mathbb{1} - P) = \operatorname{im} P_k = E_k$ has infinite dimension and codimension for all $k > 0$. In particular, this shows that no open neighbourhood of $\mathbb{1}$ in $\mathcal{L}(E)$ consists entirely

of Fredholm maps.

Pick a continuous function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\phi(t) \to \infty$ as $t \to 0$. Then

$$p_t := \begin{cases} P_{\phi(t)} & t > 0, \\ 0 & t \leq 0 \end{cases}$$

defines a continuous family $\mathbb{R} \to \mathcal{P}(E)$ of projections whose kernels and images do not form continuous vector bundles over $\mathbb{R}$: $\ker p_t$ equals E for $t \leq 0$ and has infinite codimension for $t > 0$, while $\operatorname{im} p_t$ equals $\{0\}$ for $t \leq 0$ and has infinite dimension for $t > 0$.

Another continuous family $\mathbb{R} \to \mathcal{P}(E)$ of projections π_t is given by

$$\pi_t u := \begin{cases} \langle u, \beta_t \rangle_{L^2(\mathbb{R})} \beta_t & t > 0, \\ 0 & t \leq 0 \end{cases}, \qquad \beta_t(x) := \beta\bigl(x + \phi(t)\bigr).$$

Here $\ker \pi_t$ equals E for $t \leq 0$ and has codimension one for $t > 0$, while $\operatorname{im} \pi_t$ equals $\{0\}$ for $t \leq 0$ and has dimension one for $t > 0$.

4.2 Analysis on Fréchet spaces

Consider topological vector spaces E, F, an open subset $U \subset E$, and a map $f : U \to F$. If $0 \in U$ and $f(0) = 0$, then we say that f has *derivative zero* at 0 if for each open neighbourhood $W \subset F$ of 0 there exists an open neighbourhood $V \subset E$ of 0 and a function $o : (-1, 1) \to \mathbb{R}$ such that

$$\lim_{t \to 0} \frac{o(t)}{t} = 0 \quad \text{and} \quad f(tV) \subset o(t)W \text{ for all } t \in (-1, 1).$$

In general, f is called *(Fréchet) differentiable* at $x \in U$ if there exists a continuous linear map $Df(x) : E \to F$ such that the map $h \mapsto f(x + h) - f(x) - Df(x)h$ has derivative zero at 0. Then

$Df(x) \in \mathcal{L}(E, F)$ is called the *(Fréchet) derivative* of f at x. Since the space $\mathcal{L}(E, F)$ is again a topological vector space, we can define higher order derivatives as usual.

The basic properties of the Fréchet derivative continue to hold on topological vector spaces: the Fréchet derivative satisfies the chain rule, and Fréchet differentiability implies continuity and Gâteaux differentiability. By contrast, the following two problems show that the implicit function theorem and Cartan's last theorem may fail in the Fréchet setting.

Problem 4.6. Let $E = C(\mathbb{R}, \mathbb{R})$ be the Fréchet space as in Example 4.3.
(a) Show that $f(u) := e^u$ defines a C^1-map $f : E \to E$ and compute its derivative $Df(u)$.
(b) Show that $Df(u)$ is invertible for each $u \in E$ but the image of f is not open (so the inverse function theorem fails on Fréchet spaces).

Problem 4.7. (a) Show: if in the situation of Example 4.3 the functions β and ϕ are smooth, then the map

$$r : \mathbb{R} \times E \to \mathbb{R} \times E, \qquad r(t, u) := (t, \pi_t u)$$

is a smooth retraction whose image is not a submanifold (so Cartan's last theorem fails on Fréchet spaces).
(b) Show that the same is true if $E = C(\mathbb{R}, \mathbb{R})$ is replaced by $C^\infty(\mathbb{R}, \mathbb{R})$ with its Fréchet structure as in Example 4.2(b).

Chapter 5

Sc-manifolds

The basic references for this chapter are [24, 26].

5.1 Sc-Banach spaces and their linear maps

Definition 5.1. An *sc-structure* (or scale structure) on a Banach space is a decreasing sequence $(E_m)_{m\geq 0}$ of Banach spaces

$$E = E_0 \supset E_1 \supset E_2 \supset \dots$$

such that
(i) each inclusion $E_{m+1} \hookrightarrow E_m$ is compact, and
(ii) the intersection $E_\infty := \bigcap_{i\geq 0} E_i$ is dense in each E_m.

Here each E_m is equipped with its own Banach norm $\| \, \|_m$ which is *not* the restriction of the norm of E. An *sc-Banach space* is a Banach space with an sc-structure. Points in E_m are called *points of regularity* m and points in E_∞ are called *smooth points*.

Problem 5.1. Let E_∞ be the set of smooth points in an sc-Banach space E. Show that the norms $\| \, \|_m$ make E_∞ a Fréchet space having the Heine-Borel property.

A finite dimensional Banach space E has the unique sc-structure $E_0 = E_1 = \dots$. The direct sum of two sc-Banach spaces E, F

is again naturally an sc-Banach space with $(E \oplus F)_m = E_m \oplus F_m$. The following is the basic infinite dimensional example for all applications.

Example 5.1 (Weighted Sobolev spaces). Let N be a manifold N (possibly with boundary), $\xi \to N$ a vector bundle, and $p > 1$.
(a) If N is compact, then the space $E = L^p(N, \xi)$ of L^p-sections in ξ becomes an sc-Banach space with $E_m = W^{m,p}(N, \xi)$. Here property (i) follows from the Sobolev embedding theorem. For $p = 2$ it is an sc-Hilbert space and we write $H^m(N, \xi) := W^{m,2}(N, \xi)$.
(b) Suppose now that N is noncompact with *cylindrical end*, i.e., there exists a compact subset $K \subset N$ such that $N \setminus K \cong \mathbb{R}_+ \times P$ for a compact manifold P. Suppose in addition that the restriction of ξ to the cylindrical end is the pullback of a bundle over P under the projection $\mathbb{R}_+ \times P \to P$. Then the space $E = L^p(N, \xi)$ becomes an sc-Banach space with $E_m = W^{m,p,\delta_m}(N, \xi)$ for any strictly increasing sequence $0 = \delta_0 < \delta_1 < \dots$. Here the *weighted Sobolev space* $W^{m,p,\delta_m}(N, \xi)$ consists of sections having weak derivatives up to order m which belong to L^p after being multiplied by a continuous function that equals $e^{\delta_m s}$ in the coordinate $s \in \mathbb{R}_+$ on the cylindrical end $\mathbb{R}_+ \times P$. Again, for $p = 2$ this is an sc-Hilbert space and we write $H^{m,\delta_m}(N, \xi) := W^{m,2,\delta_m}(N, \xi)$.
Typical examples are $N = S^1$ for (a) and $N = \mathbb{R}$ for (b) with a trivial bundle ξ.

Each subset A of an sc-Banach space E inherits a filtration $A_m := A \cap E_m$. Note that if $A \subset E$ is open (resp. closed), then $A_m \subset E_m$ is open (resp. closed) for each $m \geq 0$. For $k \geq 0$ we denote by A^k the set A_k with the induced filtration

$$A^k_m := A_{k+m}, \qquad m \geq 0.$$

In particular, E^k is again an sc-Banach space.

Subspaces and quotients. A linear subspace F of an sc-Banach space E is called an *sc-subspace* if it is closed and $F_m = F \cap E_m$ defines an sc-structure on F, i.e., $F_\infty = F \cap E_\infty$ is dense in each F_m. An *sc-complement* of an sc-subspace F is an sc-subspace G such that $E_m = F_m \oplus G_m$ for each $m \geq 0$. Then $E = F \oplus G$ is called an *sc-splitting*.

Proposition 5.2. *Let E be an sc-Banach space. Then:*
(a) A finite dimensional subspace $F \subset E$ is an sc-subspace if and only if $F \subset E_\infty$.
(b) Each finite dimensional sc-subspace $F \subset E$ has an sc-complement.
(c) For an sc-subspace $F \subset E$, the quotient E/F is an sc-Banach space with $(E/F)_m = E_m/F_m$.

Proof. (a) If F is an sc-subspace, then $F_\infty = F \cap E_\infty$ is dense in F; by finite dimensionality this implies $F_\infty = F$, so $F \subset E_\infty$. Conversely, if $F \subset E_\infty$, then $F_\infty = F \cap E_\infty = F$, so by finite dimensionality F is an sc-subspace.
(b) Pick a basis $e_1, \dots, e_k$ for F. By the Hahn-Banach theorem, its dual basis extends to continuous linear functionals $\lambda_1, \dots, \lambda_k \in E^*$. Then $P(x) := \sum_{i=1}^k \lambda_i(x) e_i$ defines a continuous projection $P : E \to E$ onto F. Thus the closed subspace $G := \ker P = \mathrm{im}\,(\mathbb{1} - P)$ yields a topological splitting $E = F \oplus G$. We claim that G is an sc-complement of F. For this, we need to show that $G_\infty = G \cap E_\infty$ is dense in each $G_m = G \cap E_m$. For this, let $z \in G_m$ be given. There exists a sequence (x_n) in E_∞ converging in E_m to z. Since the λ_i restrict to continuous functionals $E_m \to \mathbb{R}$ and $F \subset E_m$, the projection P restricts to a continuous map $E_m \to E_m$. Therefore, the sequence $(\mathbb{1} - P)(x_n) \in G_m$ converges in G_m to $(\mathbb{1} - P)(z) = z$.
(c) is left as an easy exercise. $\square$

Problem 5.2. Prove Proposition 5.2(c).

Linear operators. A linear map $T : E \to F$ between sc-Banach spaces is called an *sc-operator* if $T(E_m) \subset F_m$ and $T : E_m \to F_m$ is continuous for all $m \geq 0$.

Problem 5.3. Show that every sc-operator $T : E \to F$ induces a continuous linear map $T : E_\infty \to F_\infty$ with respect to their Fréchet topologies, but not every continuous linear map $E_\infty \to F_\infty$ is induced by an sc-operator $E \to F$.

In general the inverse of a bijective sc-operator $T : E \to F$ need not be an sc-operator. For example, take an infinite dimensional sc-Banach space F and equip $E := F$ with the filtration $E_0 := F_0$ and $E_m := F_{m+1}$ for all $m \geq 1$; then $\mathbb{1} : E \to F$ is a bijective sc-operator whose inverse is not an sc-operator. By the open mapping theorem applied on each level, the inverse of T is an sc-operator if and only if $T(E_m) = F_m$ for each $m \geq 0$. In this case T is called an *sc-isomorphism*.

An sc-operator $S : E \to F$ is called an *sc⁺-operator* if $S(E_m) \subset F_{m+1}$ for all $m \geq 0$ and $S : E \to F^1$ is an sc-operator. Since the inclusion $F_{m+1} \hookrightarrow F_m$ is compact, the induced operator $S : E_m \to F_m$ is compact for each m, so sc$^+$-operators are the analogues of compact operators in the sc-category.

Definition 5.2. An sc-operator $T : E \to F$ is called *sc-Fredholm* if there exist sc-splittings $E = K \oplus X$ and $F = C \oplus Y$ such that $K = \ker T$ and C are finite dimensional and $T : X \to Y = \operatorname{im} T$ is an sc-isomorphism. Its *index* is then

$$\operatorname{ind}(T) := \dim K - \dim C = \dim \ker T - \dim \operatorname{coker} T.$$

It follows from Proposition 5.2 that $K \subset E_\infty$ and $C \subset F_\infty$, and for each $m \geq 0$ we have topological splittings $E_m = K \oplus X_m$ and $F_m = C \oplus Y_m$ with $T(X_m) = Y_m$.

Proposition 5.3 (Sc-Fredholm regularity and stability).
(a) An sc-operator $T : E \to F$ is sc-Fredholm if and only if
(i) $T : E_m \to F_m$ is Fredholm for all $m \geq 0$, and
(ii) T is regularizing, *i.e., $e \in E$ and $Te \in F_m$ implies $e \in E_m$.*
(b) Sc-Fredholm operators are stable under compact perturbations, *i.e., for an sc-Fredhom operator $T : E \to F$ and an* sc$^+$*-operator $S : E \to F$ the sum $T + S$ is again sc-Fredholm.*

Proof. (a) Assume first that T is sc-Fredholm. Then condition (i) clearly holds. For (ii), let $E = K \oplus X$ and $F = C \oplus Y$ be the splittings in Definition 5.2 and consider $e \in E$ with $Te \in F_m$. Since $Te \in Y \cap F_m = Y_m$ and $T : X_m \to Y_m$ is an isomorphism, there exists $x \in X_m$ with $Tx = Te$. Then $k := e - x \in \ker T = K \subset E_\infty$, hence $e = x + k \in E_m$.

Now assume that T satisfies conditions (i) and (ii). Since $T : E_m \to F_m$ is Fredholm for each $m \geq 0$, its kernel K_m is finite dimensional. Since T is regularizing, $K_0 = K_1 = \cdots = K_\infty \subset E_\infty$. Hence, by Proposition 5.2, $K := K_0$ is a finite dimensional sc-subspace and it has an sc-complement X such that $E = K \oplus X$.

Since $T : E \to F$ is Fredholm, its image $Y = \mathrm{im}\,(T)$ is closed in F. We claim that $Y_\infty = Y \cap F_\infty$ is dense in each $Y_m = Y \cap F_m$. To see this, let $y \in Y_m$ be given and write it as $y = Te$ with $e \in E$. Since T is regularizing, it follows that $e \in E_m$. There exists a sequence (e_n) in E_∞ converging in E_m to e, and $Te_n \in Y_\infty$ converges in Y_m to $Te = y$. This proves the claim, which implies that $Y \subset F$ is an sc-subspace.

Since T is Fredholm, its cokernel F/Y is finite dimensional. Let $p : F \to F/Y$ be the projection. Since $F_\infty \subset F$ is dense, its image $p(F_\infty)$ is dense in F/Y. So we find $f_1, \ldots, f_k \in F_\infty$ representing a basis of F/Y. Then $C := \mathrm{span}\{f_1, \ldots, f_k\} \subset F_\infty$ is an sc-subspace of F such that $F = C \oplus Y$ is a topological splitting. Since $F_m = (C \oplus Y) \cap F_m = C \oplus Y_m$ for each m, this is an sc-splitting.

It remains to show that $T : X \to Y$ is an sc-isomorphism, i.e., $T(X_m) = Y_m$ for all $m \geq 0$. To see this, let $y \in Y_m$ be given. Since $T : X \to Y$ is surjective, $Tx = y$ for some $x \in X$. Since T is regularizing, it follows that $x \in X \cap E_m = X_m$ and part (a) is proved.

(b) It suffices to show that $T + S$ satisfies conditions (i) and (ii) in part (a). Since $T : E_m \to F_m$ is Fredholm and $S : E_m \to F_m$ is compact, $T + S : E_m \to F_m$ is Fredholm. To see that $T + S$ is regularizing, let $e \in E$ be given with $(T + S)e = f \in F_m$. Since S is $\mathrm{sc}^+$and T is regularizing by part (a), $Te = f - Se \in F_1$ implies $e \in E_1$, so $Te = f - Se \in F_2$, and continuing inductively we see that $e \in E_m$. □

Example 5.4 (Linear elliptic differential operators). Let $\xi, \eta \to N$ be vector bundles over a manifold N as in Example 5.1, and denote the corresponding sc-Banach spaces of sections by $E = L^p(N, \xi)$ and $F = L^p(N, \eta)$. Suppose first that N is closed. Then each linear partial differential operator $T : E_\infty \to F_\infty$ of order d induces an sc-operator $T : E^d \to F$. *If T is elliptic, then $T : E^d \to F$ is sc-Fredholm.* This follows from Proposition 5.3(a) because $T : E_{m+d} \to F_m$ is Fredholm by a classical result on elliptic operators, and $T : E^d \to F$ is regularizing by elliptic regularity. Moreover, each linear partial differential operator $S : E_\infty \to F_\infty$ of order $\leq d - 1$ induces an sc^+-operator $S : E^d \to F$.

The same holds true if N has boundary and E is replaced by a space with suitable boundary conditions, or if N has cylindrical ends and we use weighted Sobolev spaces as in Example 5.1(b).

Typical examples are the Laplace-Beltrami operator $\Delta = dd^* +$

$d^*d : \Omega^k(N) \to \Omega^k(N)$ with respect to some Riemannian metric on N (with Dirichlet or Neumann boundary conditions), or the Cauchy-Riemann operator $\bar{\partial} : C^\infty(\Sigma, \xi) \to \Omega^{0,1}(\Sigma, \xi)$ on a complex vector bundle $\xi \to \Sigma$ over a Riemann surface Σ (with Lagrangian boundary conditions).

5.2 Calculus on sc-Banach spaces

In this section we introduce the notion of sc-differentiability and prove its basic properties. We begin with sc-continuity.

Definition 5.3. A subset $C \subset E$ of an sc-Banach space is a *partial quadrant* if there exists an sc-isomorphism $T : E \to \mathbb{R}^n \oplus W$ (for some n and W) such that $T(C) = [0, \infty)^n \times W$.
An *sc-triple* (U, C, E) consists of an sc-Banach space E, a partial quadrant $C \subset E$ and a relatively open subset $U \subset C$.
Given sc-triples (U, C, E) and (V, D, F), a map $f : U \to V$ is called $\mathbf{sc}^0$ or *sc-continuous* if $f(U_m) \subset V_m$ and $f : U_m \to V_m$ is continuous for all $m \geq 0$.

The partial quadrant C is included in this definition in order to accomodate boundaries and corners. At first reading one may simply think of the case $C = E$. Next we define sc-differentiability.

Definition 5.4. The *tangent* of an sc-triple (U, C, E) is the sc-triple $T(U, C, E) = (TU, TC, TE)$ defined by

$$TU := U^1 \oplus E, \quad TC := C^1 \oplus E, \quad TE := E^1 \oplus E.$$

Definition 5.5. Let (U, C, E) and (V, D, F) be sc-triples. An sc^0-map $f : U \to V$ is called $\mathbf{sc}^1$ if
(i) For each $x \in U_1$ there exists a continuous linear operator

$Df(x) : E_0 \to F_0$ such that for $h \in E_1$ with $x + h \in U_1$,

$$\lim_{|h|_1 \to 0} \frac{|f(x+h) - f(x) - Df(x)h|_0}{|h|_1} = 0.$$

(ii) The *tangent map*

$$Tf : TU \to TV, \qquad (x,h) \mapsto \big(f(x), Df(x)h\big)$$

is sc^0.

The crucial point in this definition is the shift in regularity, which makes it differ from the usual Fréchet differentiability. Here are some comments on the definition.

(1) Condition (i) says that $f : U_1 \to F_0$ is differentiable in the usual sense and its derivative $df(x) : E_1 \to F_0$ extends to a continuous linear map $Df(x) : E_0 \to F_0$. Denoting the difference quotient by $\Delta(h)$, vanishing of the limit means that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that $\Delta(h) < \varepsilon$ for all h with $|h|_1 > \delta$. Arguing by contradiction, it suffices to show that each sequence (h_n) with $|h_n|_1 \to 0$ has a subsequence (h_{n_k}) for which $\Delta(h_{n_k}) \to 0$.

(2) Condition (ii) for fixed x implies that the derivative at $x \in U_{m+1}$ induces continuous linear maps $Df(x) : E_k \to F_k$ for all $0 \le k \le m$. In particular, $Df(x) : E \to F$ is an sc-operator for $x \in U_\infty$ (but not in general for non-smooth points).

(3) Condition (ii) provides a weak version of continuity for the derivative. In general, the map $Df : U_1 \to \mathcal{L}(E_0, F_0)$ is *not* continuous with respect to the operator topology. However, the following problem shows that it is continuous with respect to the *compact-open topology* on $\mathcal{L}(E_0, F_0)$, and the map $df : U_1 \to \mathcal{L}(E_1, F_0)$ is continuous with respect to the operator topology.

Problem 5.4. Let X be a topological space and D, E, F be topological vector spaces. The *compact-open topology* on $\mathcal{L}(E, F)$ is

defined analogously to the bounded-open topology, with bounded sets replaced by compact ones. Show:
(a) If $f : X \times E \to F$ is continuous and linear in the second variable, then the induced map $X \to \mathcal{L}(E, F)$, $x \mapsto f(x, \cdot)$ is continuous with respect to the compact-open topology.
(b) If $S : D \to E$ is linear and compact, then the induced map $\mathcal{L}(E, F) \to \mathcal{L}(D, F)$, $T \mapsto T \circ S$ is continuous with respect to the compact-open topology on $\mathcal{L}(E, F)$ and the bounded-open topology on $\mathcal{L}(D, F)$.
(c) For f and S as in (a) and (b), the induced map $X \to \mathcal{L}(D, F)$, $x \mapsto f(x, S\,\cdot)$ is continuous with respect to the bounded-open topology.

Since the tangent map $Tf : TU \to TV$ of an sc^1-map $f : U \to V$ is sc^0, we can iterate Definition 5.5 and define maps of class $\mathbf{sc}^k$ and $\mathbf{sc}^\infty$ in the obvious way.

Remark 5.5. Let us say that an sc^0 map $f : U \to V$ is $\mathbf{ssc}^1$ (or *strong* sc^1 if $f : U_m \to V_m$ is of class C^1 for each $m \geq 0$, and similarly for $\mathbf{ssc}^k$ and $\mathbf{ssc}^\infty$. This stronger notion of differentiability just leads to the usual calculus on Banach spaces on each level.

Basic properties of sc-differentiability. The first result characterizes sc-differentiability in terms of classical differentiability on each level.

Proposition 5.6 (Characterization of sc^1 in terms of C^1)**.**
Let (U, C, E) and (V, D, F) be sc-triples. An sc^0-map $f : U \to V$ is sc^1 if and only if it satisfies the following two conditions:
(a) For each $m \geq 0$ the induced map $f : U_{m+1} \to F_m$ is of class C^1; in particular, its derivative

$$df : U_{m+1} \to \mathcal{L}(E_{m+1}, F_m)$$

is continuous (with respect to the operator topology).
(b) For each $m \geq 0$ and $x \in U_{m+1}$ the derivative $df(x) : E_{m+1} \to F_m$ extends to a continuous linear operator $Df(x) : E_m \to F_m$, and the map

$$U_{m+1} \oplus E_m \to F_m, \qquad (x,h) \mapsto Df(x)h$$

is continuous.

Proof. Suppose first that f satisfies (a) and (b). Then (a) and (b) for $m = 0$ imply condition (i) in Definition 5.5, and (b) for all $m \geq 0$ implies condition (ii).
Conversely, suppose that f is sc^1. Then $f : U_1 \to F_0$ is differentiable, and for each $x \in U_1$ the differential $df(x) \in \mathcal{L}(E_1, F_0)$ extends to $Df(x) \in \mathcal{L}(E_0, F_0)$. By Problem 5.4 the map $df : U_1 \to \mathcal{L}(E_1, F_0)$ is continuous, so $f : U_1 \to F_0$ is of class C^1. By condition (ii) in Definition 5.5, $(x,h) \mapsto Df(x)h$ induces a continuous map $U_{m+1} \oplus E_m \to F_m$ for each $m \geq 0$, so (b) holds. By Problem 5.4 this implies continuity of the map $Df : U_{m+1} \to \mathcal{L}(E_{m+1}, F_m)$ with respect to the operator topology. So it only remains to show that for each $m \geq 0$ the map $f : U_{m+1} \to F_m$ is of class C^1 with derivative $df = Df$.
To see this, we estimate for $x \in U_{m+1}$ and $0 \neq h \in E_{m+1}$:

$$\begin{aligned}
&\frac{1}{|h|_{m+1}}\Big|f(x+h) - f(x) - Df(x)h\Big| \\
&= \frac{1}{|h|_{m+1}}\Big|\int_0^1 \Big[Df(x+th)h - Df(x)h\Big]dt\Big|_m \\
&\leq \int_0^1 \Big|Df(x+th)\frac{h}{|h|_{m+1}} - Df(x)\frac{h}{|h|_{m+1}}\Big|_m dt \\
&\leq \int_0^1 \|Df(x+th) - Df(x)\|_{\mathcal{L}(E_{m+1},F_m)}dt \\
&\longrightarrow 0 \text{ as } |h|_{m+1} \to 0.
\end{aligned}$$

Here the first equality holds because $f : U_1 \to F_0$ is C^1, the first inequality because $t \mapsto Df(x+th)h - Df(x)h$ defines a continuous map $[0,1] \to F_m$, the second inequality by definition of the operator norm on $\mathcal{L}(E_{m+1}, F_m)$, and the convergence to 0 because the map $Df : U_{m+1} \to \mathcal{L}(E_{m+1}, F_m)$ is continuous with respect to the operator topology. □

Corollary 5.7 (Lifting indices)**.** *If $f : U \to V$ is an* sc^k*-map, then the induced map $f : U^1 \to V^1$ is also* sc^k*.*

Proof. We prove this by induction over k. The case $k = 1$ follows immediately from Proposition 5.6 because conditions (a) and (b) are preserved under replacing U, F by U^1, F^1. For the induction step from k to $k+1$, let $f : U \to V$ be an sc^{k+1}-map. Then the tangent map $Tf : TU \to TV$ is sc^k, so by induction hypothesis $Tf : (TU)^1 = T(U^1) \to (TV)^1 = T(V^1)$ is sc^k, which means that $f : U^1 \to V^1$ is sc^{k+1}. □

Problem 5.5. Prove that for each sc^1-map $f : E \supset U \to F$ the induced map $f : E_\infty \supset U_\infty \to F_\infty$ between Fréchet spaces is of class C^1. Upgrade to C^∞.

The next result gives necessary and sufficient conditions for the sc^k property in terms of the C^k property on each level.

Proposition 5.8 (Conditions for sc^k in terms of C^k)**.**
Let (U, C, E) and (V, D, F) be sc-triples and $f : U \to V$.
(a) If $f : U \to V$ is sc^k*, then the induced map $f : U_{m+\ell} \to V_m$ is of class C^ℓ for each $m \geq 0$ and $0 \leq \ell \leq k$. In particular, $f : U_{m+k} \to V_m$ is of class C^k.*
(b) If the induced map $f : U_{m+\ell} \to V_m$ is of class $C^{\ell+1}$ for each $m \geq 0$ and $0 \leq \ell \leq k$, then $f : U \to V$ is sc^{k+1}*.*

Proof. In part (a) it suffices to prove the last assertion because every sc^k-map is also sc^ℓ for each $0 \le \ell \le k$. For $k = 0$ the assertion is obvious, and for $k = 1$ it is condition (a) in Proposition 5.6.
Next consider part (b) for $k = 0$, so $f : U_m \to V_m$ is of class C^1 for all $m \ge 0$. For $x \in U_0$ let $Df(x) \in \mathcal{L}(E_0, F_0)$ be its derivative. Restricting to $x \in U_1$, $h \in E_1$ and replacing $|h|_0$ by $|h_1| \ge C|h_0|$, we see that $Df(x)$ satisfies condition (i) in Definition 5.5. Condition (ii) follows by restricting the continuous map $U_m \oplus E_m \to F_m$, $(x, h) \mapsto Df(x)h$ to $U_{m+1} \subset U_m$.
This proves part (a) for $k = 0, 1$ and part (b) for $k = 0$. The general cases follow by a somewhat tedious induction on k which is omitted, see [24]. □

Chain rule. The following chain rule is rather surprising because the definition of sc-differentiability involves loss of one regularity level, so one might expect a loss of two levels for a composition. The proof shows that this loss can be avoided using compactness of the embedding $E_1 \hookrightarrow E_0$.

Theorem 5.9 (Chain rule)**.** *Let (U, C, E), (V, D, F) and (W, Q, G) be sc-triples. If $f : U \to V$ and $g : V \to W$ are* sc^1*-maps, then $g \circ f : U \to W$ is also* sc^1 *and*

$$T(g \circ f) = Tg \circ Tf.$$

Proof. Fix $x \in U_1$. Since V_1 is an open neighbourhood of $f(x)$ in the cone $D_1 \subset F_1$, there exists $\delta > 0$ such that

$$\phi(t, h) := tf(x + h) + (1 - t)f(x) \in V_1$$

for all $t \in [0, 1]$ and $h \in \widetilde{E}_1 := \{h \in E_1 \mid x + h \in U_1,\ |h|_1 < \delta\}$.

Since $g : U_1 \to G_0$ is of class C^1, we can write

$$\begin{aligned} g\big(f(x+h)\big)-g\big(f(x)\big) &- Dg\big(f(x)\big)Df(x)h \\ &= \int_0^1 Dg\big(\phi(t,h)\big)\Big[f(x+h)-f(x)-Df(x)h\Big] \\ &+ \int_0^1 \Big[Dg\big(\phi(t,h)\big)-Dg\big(f(x)\big)\Big]Df(x)h\,dt. \end{aligned}$$

Dividing by $|h|_1$, the first integral becomes

$$\int_0^1 Dg\big(\phi(t,h)\big)\alpha(h), \qquad \alpha(h) := \frac{f(x+h)-f(x)-Df(x)h}{|h|_1}. \tag{5.1}$$

Since f is sc^1, we have $\alpha(h) \to 0$ in F_0 as $|h|_1 \to 0$. Note that $\phi : [0,1] \times \widetilde{E}_1 \to V_1$ is continuous and $\phi(t,h) \to f(x)$ uniformly in t as $|h|_1 \to 0$. By continuity of the map $V_1 \times F_0 \to G_0$, $(y,k) \mapsto Dg(y)k$, it follows that $Dg\big(\phi(t,h)\big)\alpha(h) \to 0$ uniformly in t as $|h|_1 \to 0$, so the integral (5.1) vanishes as $|h|_1 \to 0$.

The second integral divided by $|h|_1$ becomes

$$\int_0^1 \Big[Dg\big(\phi(t,h)\big)-Dg\big(f(x)\big)\Big]\frac{Df(x)h}{|h|_1}dt. \tag{5.2}$$

Since the inclusion $E_1 \hookrightarrow E_0$ is compact and $Df(x) \in \mathcal{L}(E_0, F_0)$, the set of all $Df(x)h/|h|_1$ with $0 \neq h \in \widetilde{E}_1$ has compact closure in F_0. Since $Dg : V_1 \to \mathcal{L}(F_0, G_0)$ is continuous with respect to the compact-open topology, it follows as above that as $|h|_1 \to 0$ the integrand in (5.2) converges to 0 uniformly in t, so the integral (5.2) vanishes as $|h|_1 \to 0$.

So we have shown that $g \circ f$ satisfies condition (i) in Definition 5.5 with the continuous linear operator $Dg\big(f(x)\big)Df(x) : E_0 \to G_0$. Hence its tangent map is given by $T(g \circ f) = Tg \circ Tf : TU \to TW$, which is sc^0 because Tf and Tg are sc^0, so $g \circ f$ is sc^1. $\square$

Boundary recognition. Consider a partial quadrant C in an sc-Banach space E. Choose a linear sc-isomorphism $T : E \to \mathbb{R}^n \oplus W$ with $T(C) = [0,\infty)^n \oplus W$. For $x \in C$ write $T(x) = (a_1, \dots, a_n, w) \in [0,\infty)^n \oplus W$ and define its *degeneracy index*

$$d_C(x) := \#\{i \in \{1,\dots,n\} \mid a_i = 0\} \in \mathbb{N} \cup \{0\}.$$

Thus $x \in C$ is an interior point if $d_C(x) = 0$, a codimension one boundary point if $d_C(x) = 1$, and a corner point if $d_C(x) \geq 2$.

Problem 5.6. Show that the degeneracy index d_C does not depend on the choice of the linear sc-isomorphism $T : E \to \mathbb{R}^n \oplus W$.

The following result, whose proof we omit (see [24]), shows that the degeneracy index is also invariant under sc^1-*diffeomorphisms*, i.e., sc^1-maps with an sc^1-inverse.

Theorem 5.10. *Given sc-triples* (U, C, E), (V, D, F) *and an* sc^1*-diffeomorphism* $f : U \to V$, *for every* $x \in U$ *the degeneracy index satisfies*

$$d_C(x) = d_D\big(f(x)\big).$$

Reparametrization action. For $S^1 = \mathbb{R}/\mathbb{Z}$ and $p > 1$ consider $E := L^p(S^1, \mathbb{R}^N)$ with is usual sc-structure and the *reparametrization action*

$$\rho : \mathbb{R} \oplus E \to E, \quad \rho(s,x) := s * x := x(\cdot + s).$$

Note that ρ is linear in ξ but not in s. For smooth x its derivative is given by $D\rho(s,x) \cdot (\sigma, \xi) = D_1\rho(s,x) \cdot \sigma + D_2\rho(s,x) \cdot \xi$ with

$$D_1\rho(s,x) \cdot \sigma = (s * \dot{x})\sigma, \qquad D_2\rho(s,x) \cdot \xi = s * \xi. \tag{5.3}$$

Since this involves a derivative of x, the map $\rho : \mathbb{R} \oplus E_m \to E_m$ cannot be differentiable for any m. By contrast, we have the following result which was one of the main motivations for the notion of sc-differentiability.

Theorem 5.11. *The reparametrization action $\rho : \mathbb{R} \oplus E \to E$ between sc-Banach spaces is* sc^∞*.*

Proof. We will only prove that ρ is sc^1, for the higher derivatives see [24]. The proof has 3 steps.
Step 1. ρ is sc^0, i.e., $\rho : \mathbb{R} \oplus E_m \to E_m$ is continuous for all $m \geq 0$.
To see this, let $(s, x) \in \mathbb{R} \oplus E_m$ and $\varepsilon > 0$ be given. We pick $\widetilde{x} \in E_\infty$ with $|\widetilde{x} - x|_m < \varepsilon$. Now for $(t, y) \in \mathbb{R} \oplus E_m$ we write

$$\begin{aligned} |t * y - s * x|_m &\leq |t * y - t * \widetilde{x}|_m + |t * \widetilde{x} - s * \widetilde{x}|_m + |s * \widetilde{x} - s * x|_m \\ &=: I + II + III \end{aligned}$$

and estimate the three terms by

$$\begin{aligned} I &= |y - \widetilde{x}|_m \leq |y - x|_m + |x - \widetilde{x}|_m < |y - x|_m + \varepsilon, \\ II &\leq \sum_{k=0}^{m} \max\nolimits_{u \in [0,1]} |D^k \widetilde{x}(u + t) - D^k \widetilde{x}(u + s)| \leq C|t - s|, \\ III &= |\widetilde{x} - x|_m < \varepsilon, \end{aligned}$$

where $C := \|\widetilde{x}\|_{C^{m+1}}$. Together we obtain

$$|t * y - s * x|_m < 2\varepsilon + |y - x|_m + C|t - s| < 4\varepsilon$$

for $|y - x|_m < \varepsilon$ and $|t - s| < \varepsilon/C$, which proves continuity at (s, x).
Step 2. ρ satisfies condition (i) in Definition 5.5 with $D\rho(s, x)$ given by (5.3).
For this, let $s, \sigma \in \mathbb{R}$ and $x, \xi \in E_1$ be given and estimate

$$\begin{aligned} &|\rho(s + \sigma, x + \xi) - \rho(s, x) - D\rho(s, x) \cdot (\sigma, \xi)|_0 \\ &= |(s + \sigma) * (x + \xi) - s * x - (s * \dot{x})\sigma - s * \xi|_0 \\ &\leq |(s + \sigma) * x - s * x - (s * \dot{x})\sigma|_0 + |(s + \sigma) * \xi - s * \xi|_0. \end{aligned}$$

After replacing x, ξ by $s * x, s * \xi$ we may assume without loss of generality that $s = 0$. To estimate the first term, we pick a sequence of C^1-functions x_k converging in E_1 to x. Then for each $t \in [0, 1]$ we can write

$$(\sigma * x_k - x_k - \dot{x}_k \sigma)(t) = \int_0^1 \big(\dot{x}_k(t + \tau\sigma) - \dot{x}_k(t)\big) d\tau \cdot \sigma.$$

As $k \to \infty$ both sides converge in E_0 to the equality in E_0

$$\sigma * x - x - \dot{x}\sigma = \int_0^1 \big((\tau\sigma) * \dot{x} - \dot{x}\big) d\tau \cdot \sigma.$$

Dividing by σ we obtain

$$\frac{|\sigma * x - x - \dot{x}\sigma|_0}{|\sigma|} \leq \int_0^1 \left|(\tau\sigma) * \dot{x} - \dot{x}\right|_0 d\tau.$$

Due to continuity of the map $\mathbb{R} \to E_0$, $\sigma \mapsto \sigma * \dot{x}$ (by Step 1) the integrand converges to 0 uniformly in τ as $\sigma \to 0$, so the integral converges to 0. For the second term we apply the same argument to ξ instead of x to estimate

$$\frac{|\sigma * \xi - \xi|_0}{|\sigma|} \leq \int_0^1 |(\tau\sigma) * \dot{\xi}|_0 d\tau = \int_0^1 |\dot{\xi}|_0 d\tau = |\dot{\xi}|_0 \leq |\xi|_1,$$

which converges to 0 as $|\xi|_1 \to 0$. Together this proves

$$\lim_{|\sigma| + |\xi|_1 \to 0} \frac{|\rho(s + \sigma, x + \xi) - \rho(s, x) - D\rho(s, x) \cdot (\sigma, \xi)|_0}{|\sigma| + |\xi|_1} = 0.$$

Step 3. The map $\mathbb{R} \oplus E_{m+1} \oplus \mathbb{R} \oplus E_m \to E_m$, $(s, x, \sigma, \xi) \mapsto D\rho(s, x) \cdot (\sigma, \xi)$ is continuous for all $m \geq 0$.

We treat $D_1\rho$ and $D_2\rho$ reparately. Identifying $y \in E_m$ with the linear map $\mathbb{R} \to E_m$, $\sigma \mapsto y\sigma$ we see that

$$D_1\rho : \mathbb{R} \oplus E_{m+1} \to E_m \cong \mathcal{L}(\mathbb{R}, E_m), \qquad (s, x) \mapsto s * \dot{x}$$

is the composition of the continuous maps $\mathbb{R} \oplus E_{m+1} \to \mathbb{R} \oplus E_m$, $(s, x) \mapsto (s, \dot{x})$ and $\rho : \mathbb{R} \oplus E_m \to E_m$ (by Step 1), hence continuous. In particular, the map $\mathbb{R} \oplus E_{m+1} \oplus \mathbb{R} \to E_m$, $(s, x, \sigma) \mapsto D_1\rho(s, x) \cdot \sigma = (s * \dot{x})\sigma$ is continuous. Again by Step 1, the map $\mathbb{R} \oplus E_{m+1} \oplus E_m \to E_m$, $(s, x, \xi) \mapsto D_2\rho(s, x) \cdot \xi = s * \xi$ is continuous (note that this map in actually independent of x).
Steps 1-3 together show that ρ is sc^1. □

Remark 5.12. The preceding proof shows a fundamental difference between the partial derivatives $D_1\rho$ and $D_2\rho$. The first one, $D_1\rho(s, x) = s * \dot{x}$, involves a loss of derivative but defines a contionuous map $\mathbb{R} \oplus E_{m+1} \to \mathcal{L}(\mathbb{R}, E_m)$ with respect to the operator norm. The second one, $D_2\rho(s, x) \cdot \xi = s * \xi$, involves no loss of derivative but (dropping the irrelevant variable x) the map $D_2\rho : \mathbb{R} \to \mathcal{L}(E_m, E_m)$, $D_2\rho(s) \cdot \xi = s * \xi$ is *not* continuous with respect to the operator norm. Indeed, for each $s \in (0, 1/2)$ there exists some $x_s \in E_m$ with $|x_s|_m = 1$ and $\operatorname{supp} x_s \subset [0, s]$, so that

$$\|D_2\rho(s)-D_2\rho(0)\|_{\mathcal{L}(E_m,E_m)} = \sup_{|x|_m=1} |s*x-x|_m \geq |s*x_s-x_s|_m \geq 2.$$

Definition 5.5 (loss of a derivative in (i), pointwise continuity in (ii)) is designed precisely to make the reparametrization map sc^1.

More generally, the group G_∞ of smooth orientation preserving diffeomorphisms $g : S^1 \to S^1$ acts on E (from the right) by reparametrization

$$\rho : G_\infty \times E \to E, \qquad \rho(g, x) := x \circ g.$$

We can realize G_∞ as the smooth points in an sc-manifold (actually an sc-Lie group, see below for the definitions) G as follows. Let $F := W^{2,p}(S^1, \mathbb{R})$ with its canonical sc-structure. By the Sobolev embedding theorem we have a continuous embedding

$F \hookrightarrow C^1(S^1, \mathbb{R})$, so

$V := \{\gamma \in F \mid g(t) := t+\gamma(t)$ defines a C^1-diffeomorphism $S^1 \to S^1\}$

is an open subset of F. Now the induced sc-structure on $V \subset F$ defines an sc-structure on

$$G := \{g : S^1 \to S^1 \mid g(t) = t + \gamma(t) \text{ for } \gamma \in V\}.$$

Problem 5.7. For G and E as above prove:

To be checked. It everything sc$^\infty$?

(a) G is a group, and the maps $G \times G \to G$, $(g,h) \mapsto g \circ h$ and $G \to G$, $g \mapsto g^{-1}$ are sc^0.

(b) The reparametrization $\rho(g,x) := x \circ g$ defines an sc^0-map $G \times E \to E$.

(c) The map $\rho : G \times E \to E$ is sc^1 with derivative

$$D\rho(g,x) \cdot (\gamma, \xi) = (\dot{x} \circ g)\gamma + \xi \circ g.$$

Remark 5.13 (holomorphic spheres mod Aut(S^2)).

5.3 Sc-manifolds

Now it is clear how to define sc-manifolds. We will restrict to the smooth (sc$^\infty$) case, the sc^{k}case for finite k being entirely analogous. Let X be a topological space. An *sc-chart* $\big(V, \phi, (U, C, E)\big)$ on X is a homeomorphism $\phi : V \to U$ where $V \subset X$ is an open subset and (U, C, E) is an sc-triple. Two sc-charts $\big(V, \phi, (U, C, E)\big)$ and $\big(V', \phi', (U', C', E')\big)$ are called *sc-smoothly compatible* if the transition map

$$\phi' \circ \phi^{-1} : \phi(V \cap V') \to \phi'(V \cap V')$$

and its inverse are sc-smooth. An *sc-smooth atlas* on X is a collection of pairwise sc-smoothly compatible sc-charts $\big(V, \phi, (U, C, E)\big)$ whose domains V cover X. Two atlases are called *equivalent* if their union is again an sc-smooth atlas.

Definition 5.6. An *sc-manifold* is a paracompact Hausdorff space X equipped with an equivalence class of sc-smooth atlases.

A map $f : X \to Y$ between sc-manifolds is called sc-smooth if for all sc-charts $\big(V, \phi, (U, C, E)\big)$ for X and $\big(V', \phi', (U', C', E')\big)$ for Y the composition $\phi' \circ f \circ \phi^{-1}$ is sc-smooth. It is called an *sc-diffeomorphism* if it has an sc-smooth inverse.

Let X be an sc-manifold. We define the *degeneracy index* of $x \in X$ by $d_X(x) := d_C\big(\phi(x)\big)$ for an sc-chart $\big(V, \phi, (U, C, E)\big)$ around x. By Theorem 5.10 this definition does not depend on the sc-chart. We call x an *interior point* if $d_X(x) = 0$, a *boundary point* if $d_X(x) \geq 1$, and a *corner point* if $d_X(x) \geq 2$.

We say that a point $x \in X$ is on *level m* if $\phi(x) \in U_m$ for some (and thus every) sc-chart $\big(V, \phi, (U, C, E)\big)$ around x. Let $X_m \subset X$ be the set of all points on level m, so we have a filtration

$$X = X_0 \supset X_1 \supset \cdots \supset X_\infty := \bigcap_{m \geq 0} X_m.$$

Since the transition maps induce continuous maps on each level, each X_m inherits the structure of a topological (but generally not C^1) Banach manifold. By Problem 5.5 (upgraded from C^1 to C^∞), X_∞ is a smooth Fréchet manifold (with boundary and corners).

Example 5.14. Let N be a compact n-dimensional manifold (possibly with boundary) and M be a manifold (without boundary but possibly noncompact). Let $k \in \mathbb{N}$ and $1 \leq p < \infty$ satisfy $k - n/p > 0$. By Corollary 2.20, the space $X := W^{k,p}(N, M)$ of Sobolev maps $N \to M$ is a separable smooth Banach manifold. The charts, defined via the exponential map centered as smooth maps $N \to M$, restrict to Banach manifold charts on $X_m := W^{k+m,p}(N, M)$ for each $m \geq 0$. So they give X the structure of a *strong sc-manifold* (i.e., the transition maps are ssc^∞).

An analogous discussion applies to the space $C^{k,\alpha}(N, M)$ of Hölder maps with $k \in \mathbb{N}_0$ and $\alpha \in [0, 1]$.

Example 5.15. For $p > 1$ consider the loop space $X := W^{2,p}(S^1, M)$ with its strong sc-manifold structure as in the previous example. As explained in the previous section, the reparametrization action $S^1 \times X \to X$, $(s, x) \mapsto x(\cdot + s)$ is not classically differentiable but it is sc-smooth. Note that the action is free on the open subset $X^* \subset X$ of loops x that are *somewhere injective*, i.e., there exists $t \in S^1$ such that $\dot{x}(t) \neq 0$ and $x^{-1}(x(t)) = \{t\}$. A suitable implicit function theorem (to be proved later) implies that the quotient X^*/S^1 inherits the structure of an sc-manifold. On the other hand, X^*/S^1 does not appear to be a Banach manifold in a natural way.

Give reference.

What does this mean?

Remark 5.16. For an sc-manifold X, let $\psi : U \to U'$ be a transition map between two charts into sc-triples (U, C, E) and (U', C', E'). Then its differential at a smooth point $e \in U_\infty$ (which always exists) yields a linear sc-isomorphism $D\phi(e) : E \to E'$. Hence for a connected sc-manifold all the sc-Banach spaces corresponding to charts are sc-isomorphic.

Remark 5.17. Let X be an sc-manifold with its filtration $X = X_0 \supset X_1 \supset \cdots \supset X_\infty$. Then the topological Banach manifolds X_m and the Fréchet manifold X_∞ each carry their own topology and one may wonder whether they are all homotopy equivalent. For the sc-manifolds in Example 5.14 this is true due to the existence of *smoothing operators*, i.e., continuous maps $\mathcal{J}_\varepsilon : X \to X_\infty$ for $\varepsilon \in (0, 1]$ and $\mathcal{J}_0 = \mathbb{1} : X \to X$ such that $\mathcal{J} : [0, 1] \times X_m \to X_m$ is continuous for all m, see [9].

We will not develop the theory of sc-manifolds any further at this point. It will arise as a special case of the theory of M-polyfolds to which we turn now.

Chapter 6

M-polyfolds

The basic references for this chapter are [24, 26].

6.1 Sc-smooth retracts

Recall that an sc-triple (U, C, E) consists of an sc-Banach space E, a partial quadrant $C \subset E$ and a relatively open subset $U \subset C$. An sc-smooth map $r : U \to U$ satisfying $r \circ r = r$ is called an *sc-smooth retraction.*

Definition 6.1. An *sc-smooth retract* (O, C, E) consists of a partial quadrant $C \subset E$ in an sc-Banach space and the image $O = r(U)$ of an sc-smooth retraction $r : U \to U$ of a relatively open subset $U \subset C$.

Recall that by Cartan's last theorem the image of a smooth retraction on a Banach space is a Banach submanifold. By contrast, we will see below that an sc-smooth retract is in general not a submanifold but can consist of pieces of jumping dimension, a phenomenon that we already encountered on Fréchet spaces. Nevertheless, we will show now how to do analysis on such spaces.

Lemma 6.1. *Let (O, C, E) be an sc-smooth retract. If $r : U \to U$ and $s : V \to V$ are two sc-smooth retractions of open* *subsets $U, V \subset C$ with $r(U) = s(V) = O$, then their tangent maps satisfy*

$$Tr(TU) = Ts(TV).$$

Proof. Recall that the tangent map is defined by $Tr : U_1 \oplus E \to U_1 \oplus E$, $(x, h) \mapsto \big(r(x), Dr(x)h\big)$. Now first observe that for a retraction r with $r(U) = O$ we have

$$Tr(TU) = \{(x, h) \in O_1 \oplus E \mid Dr(x)h = h\}. \tag{6.1}$$

Indeed, $(x, h) \in Tr(TU)$ implies that $(x, h) = \big(r(y), Dr(y)k\big)$ for some $(y, k) \in U_1 \oplus E$, so $r \circ r = r$ implies $Dr(x)h = Dr(x)Dr(y)k = Dr(y)k = h$, and the converse inclusion is obvious. Next note that $r(y) \in O = s(V)$ and $s \circ s = s$ implies $s\big(r(y)\big) = r(y)$ for all $y \in U$, hence

$$s \circ r = r \quad \text{and} \quad r \circ s = s. \tag{6.2}$$

Using this and (6.1), for $(x, h) \in Tr(TU)$ we deduce $Ds(x)h = Ds(x)Dr(x)h = Dr(x)h = h$, so $(x, h) \in Ts(TV)$. This proves $Tr(TU) \subset Ts(TV)$, and the converse inclusion is proved analogously. □

Definition 6.2. The *tangent bundle* of an sc-smooth retract (O, C, E) is the triple

$$T(O, C, E) := (TO, TC, TE),$$

where $TO := Tr(TU)$ for an sc-smooth retraction $r : U \to U$ with $r(U) = O$.

By Lemma 6.1 this definition does not depend on the retraction r. Since $Tr : TU \to TU$ is an sc-smooth retraction, the tangent bundle $T(O, C, E)$ is again an sc-smooth retract. Note that

$$TO = \{(x, h) \in TU \mid Tr(x, h) = (x, h)\} \tag{6.3}$$

and we have an sc-smooth projection

$$p : TO \to O^1, \qquad (x,h) \mapsto x.$$

We define the *tangent space* of $O = r(U)$ at $x \in O_1$ by

$$T_xO := p^{-1}(x) = \{h \in E \mid Dr(h) = h\}.$$

It inherits from E the structure of a Banach space, and of an sc-Banach space if $x \in O_\infty$.

Proposition 6.2. *Let (O, C, E) and (O', C', E') be sc-smooth retracts and $f : O \to O'$ be a map. Let $r : U \to U$ and $s : V \to V$ be two sc-smooth retractions of open subsets $U, V \subset C$ with $r(U) = s(V) = O$. If $f \circ r : U \to E'$ is sc-smooth, then so is $f \circ s : V \to E'$ and the tangent maps satisfy*

$$T(f \circ r)|_{TO} = T(f \circ s)|_{TO} : TO \to TO'.$$

Proof. Let $f \circ r$ be sc$^\infty$. Since s is sc$^\infty$and $r \circ s = s$ by equation (6.2), it follows that $f \circ s = (f \circ r) \circ s$ is sc$^\infty$. Next consider $(x, h) \in TO$, so $Ts(x, h) = (x, h)$. Then the chain rule yields

$$T(f\circ r)(x,h) = T(f\circ r)Ts(x,h) = T(f\circ r\circ s)(x,h) = T(f\circ s)(x,h).$$

Finally let $r' : U' \to U'$ be an sc-smooth retraction with $r'(U') = O'$. Then $f(O) \subset O'$ implies $r' \circ f = f$, hence $r' \circ (f \circ r) = f \circ r$ and again the chain rule yields

$$Tr'T(f \circ r)(x,h) = T(r' \circ f \circ r)(x,h) = T(f \circ r)(x,h),$$

so $T(f \circ r)(x,h) \in TO'$ for each $(x,h) \in TO$. □

Definition 6.3. Let (O, C, E) and (O', C', E') be sc-smooth retracts and $r : U \to U$ be an sc-smoooth retraction with $r(U) = O$. A map $f : O \to O'$ is called *sc-smooth* if $f \circ r : U \to E'$ is sc-smooth, and in this case its *tangent map* is defined by

$$Tf := T(f \circ r)|_{TO} : TO \to TO'.$$

By Proposition 6.2 this definition does not depend on the retraction r. Note that for $(x,h)\in TO$ we have

$$Tf(x,h)=T(f\circ r)(x,h)=\big(f\circ r(x),D(f\circ r)(x)h\big)=\big(f(x),Df(x)h\big)$$

if we define

$$Df(x):=D(f\circ r)(x):T_xO\to T_{f(x)}O'.$$

Theorem 6.3 (Chain rule)**.** *Let (O,C,E), (O',C',E') and (O'',C'',E'') be sc-smooth retracts. If $f:O\to O'$ and $g:O'\to O''$ are sc-smooth, then $g\circ f:O\to O''$ is also sc-smooth and*

$$T(g\circ f)=Tg\circ Tf.$$

Proof. Let $r:U\to U$ and $r':U'\to U'$ be sc-smooth retractions onto O and O', respectively. Then $g\circ f=g\circ r'\circ f$, so sc-smoothness of $f\circ r$ and $g\circ r'$ implies sc-smoothness of $g\circ f\circ r=(g\circ r')\circ(f\circ r)$, hence of $g\circ f$. Moreover, $Tf=T(f\circ r)|_{TO}$ and $Tg=T(g\circ r')|_{TO'}$ imply via the chain rule

$$\begin{aligned}Tg\circ Tf&=T(g\circ r')T(f\circ r)|_{TO}=T(g\circ r'\circ f\circ r)|_{TO}\\&=T(g\circ f\circ r)|_{TO}=T(g\circ f).\end{aligned}$$

□

Problem 6.1. Let (O,C,E) be an sc-smooth retract. Prove:
(a) If $O'\subset O$ is open, then (O',C,E) is an sc-smooth retract.
(b) If $V\subset O$ is open and $s:V\to V$ is sc-smooth (which makes sense in view of (a)) with $s\circ s=s$, then $\big(s(V),C,E\big)$ is an sc-smooth retract.

6.2 M-polyfolds

M-polyfolds[1] are spaces that are locally modelled on sc-smooth retracts. Their definition is completely analogous to sc-manifolds, with sc-triples replaces by sc-smooth retracts. Thus an *M-polyfold chart* $\big(V, \phi, (O, C, E)\big)$ on a topological space X is a homeomorphism $\phi : V \to O$ where $V \subset X$ is an open subset and (O, C, E) is an sc-smooth retract. Two M-polyfold charts $\big(V, \phi, (O, C, E)\big)$ and $\big(V', \phi', (O', C', E')\big)$ are called *sc-smoothly compatible* if the transition map

$$\phi' \circ \phi^{-1} : \phi(V \cap V') \to \phi'(V \cap V')$$

and its inverse are sc-smooth. An *M-polyfold atlas* on X is a collection of pairwise sc-smoothly compatible M-polyfold charts $\big(V, \phi, (O, C, E)\big)$ whose domains V cover X. Two atlases are called *equivalent* if their union is again an sc-smooth atlas.

Definition 6.4. An *M-polyfold* is a paracompact Hausdorff space X equipped with an equivalence class of M-polyfold atlases.

Sc-smooth maps and *sc-diffeomorphisms* between M-polyfolds are defined in the obvious way. Every sc-manifold is an M-polyfold, and in view of Problem 6.1(a) an open subset of an M-polyfold is again an M-polyfold. An M-polyfold X carries a filtration

$$X = X_0 \supset X_1 \supset \cdots \supset X_\infty := \bigcap_{m \geq 0} X_m.$$

by points on level m. In general the X_m are *not* topological Banach manifolds and X_∞ is *not* a topological Fréchet manifold. However, each X_m inherits the structure of an M-polyfold which we denote by X^m.

[1] The "M" in "M-polyfold" stands for "manifold", the name "polyfold" being reserved to an orbifold version of these.

Tangent bundle. Let X be an M-polyfold.

Definition 6.5. The *tangent bundle* TX is the set of equivalence classes of tuples $\big(x, V, \phi, (O, C, E), h\big)$ consisting of a point $x \in X_1$, a chart $\big(V, \phi, (O, C, E)\big)$ around x and a tangent vector $h \in T_{\phi(x)}O$, where $\big(x, V, \phi, (O, C, E), h\big) \sim \big(x', V', \phi', (O', C', E'), h'\big)$ iff

$$x = x' \quad \text{and} \quad D(\phi' \circ \phi^{-1})\big(\phi(x)\big)h = h'.$$

It comes with a projection

$$p : TX \to X^1, \qquad [x, V, \phi, (O, C, E), h] \mapsto x$$

and we call $T_xX := p^{-1}(x)$ the *tangent space* at $x \in X_1$.

Let us discuss the structure of TX. To simplify notation we will just write (x, ϕ, h) for $\big(x, V, \phi, (O, C, E), h\big)$.
(1) Each tangent space T_xX is naturally a vector space with

$$\lambda[x, \phi, h] + \mu[x, \phi, k] := [x, \phi, \lambda h + \mu k], \qquad \lambda, \mu \in \mathbb{R}.$$

(2) For a fixed chart $\big(V, \phi, (O, C, E)\big)$ the map

$$T\phi : TV := p^{-1}(V) \to TO, \qquad [x, \phi, h] \mapsto \big(\phi(x), h\big)$$

is a bijection which restricts at each $x \in V_1$ to a linear isomorphism

$$T_x\phi : T_xX = T_xV \xrightarrow{\cong} T_{\phi(x)}O.$$

Thus T_xX inherits from $T_{\phi(x)}O$ the structure of a Banach space, and of an sc-Banach space if $x \in X_\infty$.
(3) The set TX carries a natural topology:

Problem 6.2. For an M-polyfold X show:
(a) The sets $(T\phi)^{-1}(W) \subset X$ for charts $\big(V, \phi, (O, C, E)\big)$ and open sets $W \subset TO$ are the basis of a topology on TX.
(b) The projection $p : TX \to X^1$ is continuous and open.
(c) TX is Hausdorff and metrizable, hence paracompact.

(4) By definition of the topology, for each chart $\big(V, \phi, (O, C, E)\big)$ the map $T\phi : TV \to TO$ is a homeomorphism onto an sc-smooth retract. For two charts ϕ, ϕ' the transition map is given at $a = \phi(x) \in O$ by

$$\begin{aligned} T\phi' \circ (T\phi)^{-1} : &\big(\phi(x), h\big) \mapsto [x, \phi, h] = [x, \phi', D(\phi' \circ \phi^{-1})(a)h] \\ &\mapsto \big(\phi'(x), D(\phi' \circ \phi^{-1})(a)h\big), \end{aligned}$$

thus $T\phi' \circ (T\phi)^{-1} = T(\phi' \circ \phi^{-1})$. This is sc-smooth because $\phi' \circ \phi^{-1}$ is, so we have shown (the last assertion being obvious)

Proposition 6.4. *The tangent bundle TX of an M-polyfold X is again an M-polyfold and the projection $p : TX \to X^1$ is sc-smooth.* □

Problem 6.3. Define the notion of an *M-polyfold bundle*, the analogue of a smooth vector bundle in the category of M-polyfolds, and show that $p : TX \to X^1$ is an M-polyfold bundle.

An sc-smooth map $f : X \to Y$ between M-polyfolds induces an sc-smooth map

$$Tf : TX \to TY, \qquad [x, \phi, h] \mapsto [f(x), \psi, D(\psi \circ f \circ \phi^{-1})\big(\phi(x)\big)h],$$

where ϕ and ψ are charts around x and $f(x)$, respectively. The restriction

$$T_x f : T_x X \to T_{f(x)} Y$$

is linear, and an sc-operator if $x \in X_\infty$.

Sub-M-polyfolds. Let X be an M-polyfold.

Definition 6.6. A subset $A \subset X$ is called a *sub-M-polyfold* if for each $a \in A$ there exists an open neighbourhood V and an sc-smooth retraction $r : V \to V$ such that $r(V) = A \cap V$.

Proposition 6.5. *Each sub-M-polyfold $A \subset X$ inherits the structure of an M-polyfold. Moreover:*
(a) The inclusion $A \hookrightarrow X$ is an sc-smooth homeomorphism onto its image.
(b) For $r : V \to V$ as in Definition 6.6 the induced map $r : V \to A$ is sc-smooth and $T_a r(T_a X) = T_a A$ for each $a \in A \cap V$.
(c) For $a \in A_\infty$ the tangent space $T_a A \subset T_a X$ is a complemented sc-subspace.

Proof. Consider $a \in A$ and $r : V \to V$ as in Definition 6.6. Let $\big(W, \phi, (O, C, E)\big)$ be a chart around a. After replacing V and W by $V \cap W \cap r^{-1}(W)$ we may assume that $V = W$, so we have the commuting diagram

$$\begin{tikzcd} V \arrow[r, "\phi", "\cong"'] \arrow[d, "r"] & O \arrow[d, "s"] \\ V \arrow[r, "\phi", "\cong"'] & O \end{tikzcd}$$

with horizontal homeomorphisms and the sc-smooth retraction

$$s := \phi \circ r \circ \phi^{-1} : O \to O.$$

From $r(V) = A \cap V$ it follows that $\phi(A \cap V) = \phi\big(r(V)\big) = s\big(\phi(V)\big) = s(O)$, so ϕ restricts to a homeomorphism

$$\psi := \phi|_{A \cap V} : A \cap V \xrightarrow{\cong} s(O).$$

Since $s(O)$ is an sc-smooth retract by Problem 6.1(b), the tuple $\big(A \cap V, \psi, (s(O), C, E)\big)$ is a chart on A around a. For two such charts we get the diagram

$$\begin{tikzcd} O \arrow[d, "s"] & V \cap V' \arrow[l, "\phi"'] \arrow[r, "\phi'"] \arrow[d, "r'"] & O \arrow[d, "s'"] \\ s(O) & A \cap V \cap V' \arrow[l, "\psi"'] \arrow[r, "\psi'"] & O' \end{tikzcd}$$

Since $s \circ \phi(V \cap V') = \psi(A \cap V \cap V') \subset s(O)$ is open, by Problem 6.1(a) we find an sc-smooth retraction $\rho : U \to U$ of an open subset $U \subset C$ with $\rho(U) = \psi(A \cap V \cap V')$. Since $\phi' \circ \phi^{-1}$ and ρ are sc-smooth, so is the transition map $\psi' \circ \psi^{-1} \circ \rho = \phi' \circ \phi^{-1} \circ \rho$. This shows that the above charts define an M-polyfold atlas on A. Assertions (a) and (b) now follow easily and are left as exercises. Assertion (c) follows from (b) which implies that $T_a r : T_a X \to T_a A$ at a smooth point a is a surjective sc-operator, so it induces the sc-splitting $T_a X = T_a A \oplus \ker(T_a r)$. □

Problem 6.4. Prove assertions (a) and (b) in Proposition 6.5.

Boundary and corners. For M-polyfolds the recognition of boundary and corner points is more subtle than for sc-manifolds, even in finite dimensions:

Example 6.6. For $a \in [0, \infty)$ consider the orthogonal projection

$$r_a : \mathbb{R}^2 \to \mathbb{R}^2, \qquad (x, y) \mapsto \frac{x + ay}{1 + a^2}(1, a)$$

onto the line $L_a = \{(x, ax) \mid a \in \mathbb{R}\}$. It restricts to a smooth retraction $r_a : C \to C$ of the cone $C = [0, \infty)^2$ with $r_a(C) = L_a \cap C$. Thus $\phi_a(x) := (x, ax)$ defines a global chart $\phi_a : X \to r_a(C)$ on the M-polyfold $X = [0, \infty)$. Another global chart is given by the identity $\phi' = \mathrm{id} : X \to C' = [0, \infty)$. The transition map $\psi := \phi_a \circ (\phi')^{-1} : [0, \infty) \to \mathbb{R}^2$, $x \mapsto (x, ax)$ is smooth and so is its inverse $\psi^{-1} : r_a(C) \to \mathbb{R}$ because $\psi^{-1} \circ r_a : C \to \mathbb{R}$, $(x, y) \mapsto \frac{x+ay}{1+a^2}$ is smooth. Thus the charts ϕ_a, ϕ' are compatible. However, the degeneracy indices of a point $x \in X$ in these charts depend on the charts:

$$d_C\big(\phi_a(x)\big) = \begin{cases} 2 & x = 0, \\ 1 & x \neq 0, a = 0 \\ 0 & x \neq 0, a > 0 \end{cases}, \qquad d_{C'}\big(\phi'(x)\big) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0. \end{cases}$$

In the preceding example, the "correct" degeneracy indices are clearly given by the chart ϕ'. This motivates the following

Definition 6.7. The *degeneracy index* $d_X(x)$ of a point x in an M-polyfold X is the minimum of $d_C\big(\phi(x)\big)$ over all charts around x.

With this definition, we again call x an *interior point* if $d_X(x) = 0$, a *boundary point* if $d_X(x) \geq 1$, and a *corner point* if $d_X(x) \geq 2$. In Example 6.6 the retraction $r_a : C \to C$ does not preserve the degeneracy index and its image $r_a(C)$ is not transverse to the stratification of the quadrant C by degeneracy indices. These phenomena are ruled out by the following

Definition 6.8. Let (U, C, E) be an sc-triple. An sc-smooth retraction $r : U \to U$ is called *tame* if
(i) $d_C\big(r(x)\big) = d_C(x)$ for all $x \in U$, and
(ii) for each smooth point $x \in O = r(U)$ there exists an sc-subspace $A \subset E$ satisfying $E = T_xO \oplus A$ and $A \subset E_x$, where E_x is the tangent space to the stratum of C through x.

Now an sc-smooth retract (O, C, E) is called *tame* if O is the image of a tame sc-smooth retraction $U \to U$, and an M-polyfold is called *tame* if it possesses an equivalent M-polyfold atlas modeled on tame sc-smooth retracts. It is proved in [26] that for a tame M-polyfold X the degeneracy index is independent of the chart, i.e., for each chart $\big(V, \phi, (O, C, E)\big)$ we have

$$d_X(x) = d_C\big(\phi(x)\big).$$

We refer to [26] for further discussion of the boundary and corner structure on tame M-polyfolds.

Problem 6.5. Let $r : C \to C$ be a smooth retraction of $C := [0, \infty)^n$. Show:

(a) r satisfies condition (ii) in Definition 6.8 iff $r(C)$ is transverse to each stratum of C.
(b) If r is tame, then $r(C)$ is an open neighbourhood of 0 in C.
Hint: Show $r(0) = 0$ *and use part (a).*

6.3 Examples of M-polyfolds

In Example 4.3 and Problem 4.7 we constructed a smooth family of projections on the Fréchet space $C(\mathbb{R}, \mathbb{R})$ whose images have jumping dimension, showing that Cartan's last theorem fails on Fréchet spaces. Now we carry over this example to an sc-Banach space.

Example 6.7. Consider the sc-Hilbert space $E = L^2(\mathbb{R}) = L^2(\mathbb{R}, \mathbb{R})$ with its filtration $E_m = H^{m,\delta_m}(\mathbb{R})$ for an increasing sequence $0 = \delta_0 < \delta_1 < \dots$. Fix a smooth function $\beta : \mathbb{R} \to [0, \infty)$ with $\operatorname{supp} \beta \subset [-1, 1]$ and $\int_{\mathbb{R}} \beta(s)^2 ds = 1$. Let

$$\phi : \mathbb{R}_+ \to \mathbb{R}_+, \qquad \phi(t) := e^{1/t}.$$

and for $t > 0$ define

$$\beta_t : \mathbb{R} \to \mathbb{R}, \qquad \beta_t(s) := \beta\big(s + \phi(t)\big).$$

As in Example 4.3 this gives rise to a family of projections

$$\pi_t : E \to E, \qquad \pi_t u := \begin{cases} \langle u, \beta_t \rangle_{L^2(\mathbb{R})} \beta_t & t > 0, \\ 0 & t \leq 0 \end{cases}$$

whose image $\operatorname{im} \pi_t$ equals $\{0\}$ for $t \leq 0$ and has dimension one for $t > 0$. We will show below that the retraction

$$r : \mathbb{R} \oplus E \to \mathbb{R} \oplus E, \qquad (t, u) \mapsto (t, \pi_t u)$$

is sc-smooth. Thus $O := r(\mathbb{R} \oplus E)$ is an sc-smooth retract homeomorphic to the set $(-\infty, 0] \times \{0\} \cup (0, \infty) \times \mathbb{R}$, which is not a topological manifold. Note that $O = O_\infty$ consists entirely of smooth points (because $\operatorname{im} \pi_t \subset E_\infty$ for all t).

Proposition 6.8. *The retraction $r : \mathbb{R} \oplus E \to \mathbb{R} \oplus E$ in Example 6.7 is sc-smooth.*

Proof. Clearly we only need to consider the second component $(t, u) \mapsto \pi_t u$. More generally, let us consider a map

$$\Phi : \mathbb{R} \oplus E \to E, \qquad \Phi(t, u) := \begin{cases} \langle u, f_t \rangle_{L^2} g_t & t > 0, \\ 0 & t \leq 0 \end{cases}$$

where $f_t = f(t, \cdot)$ and $g_t = g(t, \cdot)$ for $f, g \in \mathcal{F}$. Here $\mathcal{F}$ is the class of smooth functions $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, $f(t, s) = f_t(s)$, such that

$$\operatorname{supp} f_t \subset I_t := [-\phi(t) - A, \phi(t) + A]$$

and for all $k, \ell \in \mathbb{N}_0$ there exist constants $C_{k,\ell}$ such that

$$\max\nolimits_{s \in I_t} |\partial_t^k \partial_s^\ell f(t, s)| \leq C_{k,\ell} \frac{\phi(t)^k}{t^{2k}}. \tag{6.4}$$

Recall that $\phi(t) = e^{1/t}$ and note that the function $(t, s) \mapsto \beta_t(s) = \beta\big(s + \phi(t)\big)$ belongs to the class $\mathcal{F}$.

In the following C will denote a generic constant and $t > 0$ will be so small that $I_t \subset (-\infty, 0)$. For $f \in \mathcal{F}$ condition (6.4) with $k = \ell = 0$ yields $\max_{s \in I_t} |f(t, s)| \leq C$, so for $u \in E_m$ and $t > 0$ we obtain

$$\begin{aligned} \langle u, f_t \rangle_{L^2}^2 &= \int_{I_t} |u(s)|^2 f_t(s)^2 e^{2\delta_m s} e^{-2\delta_m s} ds \\ &\leq C e^{-2\delta_m \phi(t)} \int_{I_t} |u(s)|^2 e^{-2\delta_m s} ds. \end{aligned} \tag{6.5}$$

Now we prove sc-smoothness of Φ. This is clear for $t \neq 0$, so we only have to check it at $t = 0$.

Step 1. Φ is sc^0, i.e., $\Phi : \mathbb{R} \oplus E_m \to E_m$ is continuous at $(0, u_0)$ for each $u_0 \in E_m$.

For this, let u_0 and $\varepsilon > 0$ be given. For $t > 0$ and $h \in E_m$ we write $u := u_0 + h$. Since $\Phi(0.u_0) = 0$, we estimate using (6.5) and $\max_{s \in I_t} |\partial_s^\ell g(t,s)| \leq C_\ell$:

$$\begin{aligned}
|\Phi(t,u) - \Phi(0,u_0)|_m^2 &= \langle u, f_t \rangle_{L^2}^2 \sum_{\ell \leq m} \int_{I_t} |\partial_s^\ell g(t,s)|^2 e^{-2\delta_m s} ds \\
&\leq C e^{-2\delta_m \phi(t)} \int_{I_t} |u(s)|^2 e^{-2\delta_m s} ds \int_{I_t} e^{-2\delta_m s} ds \\
&\leq C \int_{I_t} |u(s)|^2 e^{-2\delta_m s} ds \\
&\leq C \int_{I_t} |u_0(s)|^2 e^{-2\delta_m s} ds + C|h|_m^2 \\
&< \varepsilon^2
\end{aligned}$$

for $|h|_m$ and t sufficiently small, since $\int_{I_t} |u_0(s)|^2 e^{-2\delta_m s} ds \to 0$ as $t \to 0$ (because I_t moves to $-\infty$).

Step 2. The map $\Phi : \mathbb{R} \oplus E_{m+1} \to E_m$ is Fréchet differentiable at $(0, u_0)$ with $D\Phi(0, u_0) = 0$ (for $m = 0$ this is condition (i) in Definition 5.5).

For this, consider $t > 0$, $h \in E_{m+1}$ and write $u := u_0 + h \in E_{m+1}$. Then from Step 1 with we obtain with $\delta' := \delta_{m+1} - \delta_m > 0$:

$$\begin{aligned}
|\Phi(t,u) - \Phi(0,u_0)|_m^2 &\leq C \int_{I_t} |u(s)|^2 e^{-2\delta_m s} ds \\
&\leq C e^{-2\delta' \phi(t)} \int_{I_t} |u(s)|^2 e^{-2\delta_{m+1} s} ds \\
&\leq C e^{-2\delta' \phi(t)} |u|_{m+1}^2,
\end{aligned}$$

hence

$$\frac{|\Phi(t, u_0 + h) - \Phi(0, u_0)|_m}{|t| + |h|_{m+1}} \leq C \frac{e^{-\delta' \phi(t)}}{|t|} |u_0 + h|_{m+1}$$

converges to 0 as $|t| + |h|_{m+1} \to 0$.

Step 3. Now we prove that Φ is sc^{k+1} for each $k \geq 0$ by induction, assuming each such map is sc^k.

The sc-derivative of $\Phi : \mathbb{R} \oplus E_1 \to E_0$ at (t, u) with $t > 0$ is given on $(\tau, h) \in \mathbb{R} \oplus E$ by

$$D\Phi(t,u)(\tau,h) = \langle h, f_t \rangle_{L^2} g_t + \langle u, \partial_t f_t \rangle_{L^2} g_t \tau + \langle u, f_t \rangle_{L^2} \partial_t g_t \tau. \tag{6.6}$$

For $t \leq 0$ we have $D\Phi(t, u) = 0$. The first term on the right hand side defines by induction hypothesis an sc^k-map $\mathbb{R} \oplus E \to E$, $(t, h) \mapsto \langle h, f_t \rangle_{L^2} g_t$. For the second term we rewrite

$$\langle u, \partial_t f_t \rangle_{L^2} = \langle e^{-\delta_1 s} u, \widetilde{f}_t \rangle_{L^2}$$

where the function $\widetilde{f}(t, s) := e^{\delta_1 s} \partial_t f(t, s)$ belongs to the class $\mathcal{F}$. Let us introduce the sc-Hilbert space $\widetilde{E} := L^2(\mathbb{R})$ with the filtration $\widetilde{E}_m := H^{m,\delta_{m+1}-\delta_1}(\mathbb{R})$. Identifying $\mathcal{L}(\mathbb{R}, E)$ with E the second term is then the composition of the linear sc-operator

$$\mathbb{R} \oplus E^1 \to \mathbb{R} \oplus \widetilde{E}, \qquad (t, u) \mapsto (t, e^{-\delta_1 s} u)$$

with the map (set to 0 for $t \leq 0$)

$$\mathbb{R} \oplus \widetilde{E} \to E, \qquad (t, v) \mapsto \langle v, \widetilde{f}_t \rangle_{L^2} g_t,$$

which is sc^k by induction hypothesis (after replacing E by $\widetilde{E}$). The third term is not exactly of the same type, so we cannot apply the induction hypothesis. Continuity as a map $\mathbb{R} \oplus E_{m+1} \to E_m$ follows from the estimates in Step 1 for $t > 0$ and $u = u_0 + h \in$

E_{m+1}:

$$\begin{aligned}|\langle u, f_t\rangle_{L^2}\partial_t g_t|_m^2 &\le C\|\partial_t g_t\|_{C^m(I_t)} e^{-2(\delta_{m+1}-\delta_m)\phi(t)} \int_{I_t} |u(s)|^2 e^{-2\delta_{m+1}s} ds \\ &\le C \int_{I_t} |u(s)|^2 e^{-2\delta_{m+1}s} ds \\ &\le C \int_{I_t} |u_0(s)|^2 e^{-2\delta_{m+1}s} ds + C|h|_{m+1}^2 \\ &< \varepsilon^2\end{aligned}$$

for $|h|_{m+1}$ and t sufficiently small. This shows that $T\Phi$ is sc^0, so Φ is sc^1, and a modified induction scheme (see [24]) also shows that it is sc^∞. □

Remark 6.9. One may wonder if there is a notion of dimension for an M-polyfold X. The dimension of the sc-Banach space E corresponding to local charts is clearly not the right one even in finite dimensions. One could define the "dimension" of X at $x \in X$ as the dimension of the tangent space $T_x X$. Then Example 6.7 and some variants of it show that this "dimension" can jump even within a single chart, by arbitrary finite or infinite amounts.

6.4 Gluing and antigluing

Chapter 7

Fredholm theory on M-polyfolds

The basic references for this chapter are [24, 26]. We will only consider the case of M-polyfolds without boundary and corners, occasionally mentioning the modifications in the general case in remarks.

7.1 Strong bundles over M-polyfolds

7.2 Germs and fillings

Definition 7.1. A *filling* of an sc-smooth section germ $f : (\mathcal{O}, 0) \to K$ is an sc-smooth section germ $g : (E, 0) \to E \triangleleft F$ with the following properties.
(i) $g(x) = f(x)$ for each $x \in \mathcal{O}$ close to 0.
(ii) If $x \in E$ close to 0 satisfies $g(x) = \rho(r(x))g(x)$, then $x \in \mathcal{O}$.
(iii) The linearization of the map $x \mapsto \big(\mathbb{1} - \rho(r(x))\big)g(x)$ at 0 restricts to a continuous linear isomorphism $\ker Dr(0) \xrightarrow{\cong} \ker \rho(0)$.

Proposition 7.1.

Problem 7.1. If f is an sc-smooth germ at each point, then it is sc-smooth.

7.3 Contraction germs

Recall the proof of the implicit function theorem for Banach spaces V, W, Z (adjusted to the notation in [26]). We are given a continuous germ $f : \big(V \oplus W, (0,0)\big) \to (Z, 0)$ with continuous Fréchet derivative $D_2 f : V \times W \to \mathcal{L}(W, Z)$ with respect to the second variable such that $D_2 f(0,0)$ is an isomorphism. We want to find a continuous germ $\delta : (V, 0) \to (W, 0)$ such that

$$f(a, w) = 0 \Longleftrightarrow w = \delta(a) \quad \text{for } (a, w) \text{ near } (0,0),$$

and if f is of class C^k then so is δ.
After replacing f by $D_2 f(0,0)^{-1} \circ f$ we may assume that $W = Z$ and $D_2 f(0,0) = \mathbb{1}$, so we can write

$$f(a, w) = w - B(a, w)$$

with $B(0,0) = 0$ and $D_2 B(0,0) = 0$. By continuity of $D_2 f$ (hence $D_2 B$), for each $\varepsilon > 0$ there exists a product $U' = V' \times W' \subset V \oplus W$ of open balls around the origin (depending on ε) such that

$$|D_2 B(a, w)| \leq \varepsilon \quad \text{for all } (a, w) \in U'.$$

By integration, this implies the contraction property

$$|B(a, w) - B(a, w')| \leq \varepsilon |w - w'| \quad \text{for all } (a, w), (a, w') \in U'.$$

Using this with continuity of f (hence B) and $B(0,0) = 0$, by shrinking V' for fixed W' we can arrange $B(a, w) \in W'$ for $(a, w) \in V' \times \overline{W}'$. Choosing $\varepsilon < 1$, we see that for each $a \in V'$ the map $\overline{W}' \to W'$, $w \mapsto B(a, w)$ is contracting, so it has a unique fixed point $\delta(a) \in W'$. Continuity of f in the first variable implies that δ is continuous, and a similar argument yields higher regularity.

Suppose now that V, W, Z are sc-Banach spaces and f is an sc^k-germ. The reduction to the case $W = Z$ and $D_2 f(0,0) = \mathbb{1}$ works as before, so we can again write $f(a,w) = w - B(a,w)$ with $B(0,0) = 0$ and $D_2 B(0,0) = 0$. But in the next step we encounter two difficulties: the sc-derivative $D_2 B : V_1 \oplus W_1 \to \mathcal{L}(W,W)$ loses one level of regularity and it is not continuous with respect to the operator norm on $\mathcal{L}(W,W)$. It appears that we cannot proceed without further hypotheses. The solution in [26] is to skip this step and require the contraction property right away:

Definition 7.2 (sc^0-contraction germ)**.** Let $E = \mathbb{R}^n \oplus W$ be a split sc-Banach space. An sc^0-germ $f : (E,0) \to (W,0)$ is an **sc^0-contraction germ** if it has the form

$$f(a,w) = w - B(a,w),$$

and for each $m \in \mathbb{N}_0$, $\varepsilon > 0$ there exists an open neighbourhood $U_m \subset E_m$ of the origin (depending on m, ε) such that

$$|B(a,w) - B(a,w')|_m \leq \varepsilon |w - w'|_m \quad \text{for all } (a,w), (a,w') \in U_m.$$

Remark 7.2. (a) In the general case E is replaced by a partial quadrant $C = \big([0,\infty)^k \oplus \mathbb{R}^{n-k}\big) \oplus W$.
(b) The first factor in the splitting $E = \mathbb{R}^n \oplus W$ is chosen finite dimensional because this is the situation that arises for Fredholm sections. This restriction does not appear to be indispensable, but it simplifies the theory because no loss of derivatives occurs in the first factor.

With this definition, the proof of the implicit function theorem can be adapted to the sc-setting and yields (see [26, Theorem 3.9])

Theorem 7.3 (Implicit function theorem for contraction germs)**.** *Let* $f : (E,0) \to (W,0)$ *be an* sc^0*-contraction germ. Then there* *exists a unique* sc^0*-germ* $\delta : (\mathbb{R}^n, 0) \to (W, 0)$ *such that*

$$f(a,w) = 0 \Longleftrightarrow w = \delta(a) \quad \textit{for } (a,w) \textit{ near } (0,0).$$

Moreover, if f *is of class* sc^k *then so is* δ*.*

Problem 7.2. Prove the sc^0-assertion in Theorem 7.3. (The assertion for higher sc^k is more difficult [26]).

We call δ in Theorem 7.3 the *solution germ*. In view of Proposition 5.8, sc-smoothness of δ is equivalent to the existence of a nested sequence $V_0 \supset V_1 \supset \cdots$ of open neighbourhoods of the origin in $\mathbb{R}^n$ such that for each m we have

$$\delta(V_m) \subset W_m \text{ and } \delta : V_m \to W_m \text{ is of class } C^m.$$

The next lemma describes the linear geometry of sc^0-contraction germs and their dependence on the sc-splitting.

Lemma 7.4. *Let* $f : (E, 0) \to (W, 0)$ *be an* sc^0*-contraction germ of class* sc^1*. Then:*
(a) $D_2 f(0,0) = \mathbb{1}$*, and* $D\delta(0) = -D_1 f(0,0)$ *for each* sc^1*-solution germ* δ *of* $f\big(a, \delta(a)\big) = 0$*.*
(b) For $K := \ker Df(0,0)$*, the* sc^1*-germ* $\widehat{f} : (K \oplus W, 0) \to (W, 0)$ *corresponding to* f *under the sc-splitting* $E = K \oplus W$ *is an* sc^0*-contraction germ satisfying* $D_1 \widehat{f}(0,0) = 0$*.*

Proof. (a) Let $f(a,w) = w - B(a,w)$ as in Definition 7.2. For each $\varepsilon > 0$ we have $|B(0.w) - B(0,0)|_0 \le \varepsilon |w|_0$ for $|w|_0$ sufficiently small. Thus $\lim_{|w|_0 \to 0} \frac{|B(0,w)-B(0,0)|_0}{|w|_0} = 0$, which shows that $D_2 B(0,0) = 0$ and therefore $D_2 f(0,0) = \mathbb{1}$. The second assertion follows by taking the derivative of $f\big(a, \delta(a)\big) = 0$, which gives $0 = D_1 f(0,0) + D\delta(0)$.
(b) Since $Df(0,0)(a,w) = w - D_1 B(0,0)a$, its kernel equals $K = \{(a, La) \mid a \in \mathbb{R}^n\}$ with $L := D_1 B(0,0)$. It follows that, with

the projection $\pi : \mathbb{R}^n \oplus W \to \mathbb{R}^n$, the linear coordinate change between the splittings $\mathbb{R}^n \oplus W = K \oplus W$ sends (a, w) to $(k, v) = \big((a, La), w - La\big)$, and its inverse sends (k, v) to $(a, v + La)$ with $a = \pi k$. Thus with respect to the splitting $K \oplus W$ the map f takes the form

$$\widehat{f}(k, v) = f(a, v + La) = v + La - B(a, v + La), \qquad a = \pi k.$$

This has the form $\widehat{f}(k, v) = v - \widehat{B}(k, v)$ with

$$\widehat{B}(k, v) = B(a, v + La) - La, \qquad a = \pi k.$$

Since for each $m \geq 0$ with $a = \pi k$ we have

$$|\widehat{B}(k, v) - \widehat{B}(k, v')|_m = |B(a, v + La) - B(a, v' + La)|_m \leq \varepsilon |v - v'|_m$$

for (k, v) and (k, v') sufficiently close to $(0, 0)$, the germ $\widehat{f}$ is an sc^0-contraction germ. It satisfies $D_1\widehat{f}(0, 0) = 0$ because $K = \ker Df(0, 0)$. □

Corollary 7.5. *If f in Theorem 7.3 is of class at least* sc^1*, then after changing the sc-splitting to $E = K \oplus W$ with $K =$* $\ker Df(0, 0)$ *we may assume that the solution germ $\delta : (K, 0) \to (W, 0)$ satisfies in addition $D\delta(0) = 0$.* □

7.4 Basic germs

The following definition will be the basic building block for the definition of an sc-Fredholm operator.

Definition 7.3 (basic germ)**.** Let $E = \mathbb{R}^n \oplus W$ and $F = \mathbb{R}^N \oplus W$ be split sc-Banach spaces and $P : \mathbb{R}^N \oplus W \to W$ be the projection. A **basic germ** is an sc-smooth germ $f : (E, 0) \to (F, 0)$ such that $P \circ f : (E, 0) \to (W, 0)$ is an sc^0-contraction germ.

Remark 7.6. Note that we require sc-smoothness in the definition. In the general case E is again replaced by the partial quadrant $C = \big([0,\infty)^k \oplus \mathbb{R}^{n-k}\big) \oplus W$.

The Germ Implicit Function Theorem 7.3 extends from sc^0-contraction germs to basic germs (see [26, Theorem 3.12])

Theorem 7.7 (Implicit function theorem for basic germs)**.** *Let $f : (E,0) \to (F,0)$ be a basic germ such that $Df(0)$ is surjective. Then there exists an sc-splitting $E = K \oplus Y$ with $K = \ker Df(0)$ and a unique sc-smooth germ $\delta : (K,0) \to (Y,0)$ such that*

$$f(k,y) = 0 \Longleftrightarrow y = \delta(k) \quad \textit{for } (k,y) \textit{ near } (0,0).$$

Moreover, δ satisfies $D\delta(0) = 0$.

Proof. By definition of a basic germ, $f_1 := P \circ f : (E,0) \to (W,0)$ is an sc^0-contraction germ. Thus by Corollary 7.5, for the sc-splitting $E = K_1 \oplus W$ with $K_1 = \ker Df_1(0)$ there exists a unique sc-smooth solution germ $\delta_1 : (K_1,0) \to (W,0)$ such that $D\delta_1(0) = 0$ and

$$f_1(a,w) = 0 \Longleftrightarrow w = \delta_1(a) \quad \text{for } (a,w) \in K_1 \oplus W \text{ near } (0,0).$$

Set $f_0 := (\mathbb{1} - P) \circ f$ and consider the sc-smooth germ

$$g : (K_1,0) \to (\mathbb{R}^N,0), \qquad g(a) := f_0\big(a, \delta_1(a)\big).$$

The condition $D\delta_1(0) = 0$ implies $Dg(0) = Df_0(0)|_{K_1}$. By hypothesis $Df(0) = \big(Df_0(0), Df_1(0)\big) : E \to \mathbb{R}^N \oplus W$ is surjective. Thus for each $b \in \mathbb{R}^N$ there exists $e \in E$ with $Df(0) = (b,0)$, which shows that $Dg(0) = Df_0(0)|_{K_1}$ is surjective. Since K_1 is finite dimensional, there exists an open neighbourhood $U_1 \subset K_1$ of 0 on which g is of class C^1. By the proof of the finite dimensional implicit function theorem (see the discussion at the beginning of Section 7.3) this implies that g is an sc^0-contraction germ. Again by Corollary 7.5, for any splitting $K_1 = K \oplus X$ with $K = \ker Dg(0) = \ker Df(0)$ there exists a unique sc-smooth solution germ $\delta_0 : (K, 0) \to (X, 0)$ such that $D\delta_0(0) = 0$ and

$$g(k, x) = 0 \iff x = \delta_0(k) \quad \text{for } (k, x) \in K \oplus X \text{ near } (0, 0).$$

Altogether we have an sc-splitting $E = K \oplus X \oplus W$ such that for (k, x, w) near $(0, 0, 0)$,

$$\begin{aligned} f(k, x, w) = 0 &\iff w = \delta_1(k, x) \text{ and } x = \delta_0(k) \\ &\iff (x, w) = \big(\delta_0(k), \delta_1(k, \delta_0(k)\big) =: \delta(k). \end{aligned}$$

Setting $Y := X \oplus W$, this shows that $\delta : (K, 0) \to (Y, 0)$ is the unique sc-smooth solution germ for f. Since $D\delta(0) = 0$, this concludes the proof of Theorem 7.7. □

The following result corresponds to [26, Theorem 3.10].

Proposition 7.8 (Stability of basic germs under sc$^+$-perturbations)**.** *Let $f : (E, 0) \to (F, 0)$ be a basic germ and $s : (E, 0) \to (F^1, 0)$ be an sc-smooth germ. Then there exists a linear sc-isomorphism $\Psi : F \to F$ such that $\big(\Psi \circ (f + s)\big)^1 : (E^1, 0) \to (F^1, 0)$ (with levels of regularity raised by 1) is a basic germ.*

Remark 7.9. Using the terminology of Section 7.1 one should view f as the germ of a section of the strong bundle $E \triangleleft F \to E$, s as the germ of an sc$^+$-section in $E \triangleleft F \to E$, and Ψ as the germ of a strong bundle isomorphism $E \triangleleft F \to E \triangleleft F$ covering the identity $E \to E$.

Proof. Let $P : \mathbb{R}^N \oplus W \to W$ be the projection. Then $A := P \circ D_2 s(0) : W \to W$ is a linear sc$^+$-operator, so $\mathbb{1} + A$ is an sc-Fredholm operator of index 0. It follows that we have splittings

$$W = K \oplus X = C \oplus R$$

with $K = \ker(\mathbb{1}+A)$ and C finite dimensional and $R = \operatorname{im}(\mathbb{1}+A)$. By definition of an sc$^+$-section, for each $m \geq 1$ we find an open neighbourhood $U_m \subset E_m$ of 0 such that $s : U_m \to F_m$ is of class C^1. It follows that

$$P \circ s(a,w) = Aw + S(a,w),$$

where $S(a,w) = P \circ s(a,w) - Aw$ satisfies $D_2 S(0) = 0$ and defines a C^1-map $S : U_m \to F_m$ for each $m \geq 1$. Hence

$$P \circ (f+s)(a,w) = w - B(a,w) + Aw + S(a,w) = (\mathbb{1}+A)w - \overline{B}(a,w),$$

where $\overline{B}(a,w) = B(a,w) - S(a,w)$ satisfies the contraction estimate in Definition 7.2 for each $m \geq 1$. Introduce the projections

$$P_1 : W = K \oplus X \to X, \qquad P_2 : W = C \oplus R \to R$$

and the sc-smooth germ $\phi : (E,0) \to (R,0)$,

$$\phi(a,w) := P_2 \circ P \circ (f+s)(a,w) = (\mathbb{1}+A)P_1 w - P_2 \overline{B}(a,w),$$

where we have used $(\mathbb{1}+A)P_1 = \mathbb{1}+A = P_2(\mathbb{1}+A)$. Applying the inverse of the linear sc-isomorphism $L := (\mathbb{1}+A)|_X : X \to R$ yields

$$L^{-1}\phi(a,w) = P_1 w - L^{-1} P_2 \overline{B}(a,w).$$

Writing $w = (\mathbb{1} - P_1)w + P_1 w \in K \oplus X$, we see that $\big(L^{-1}\phi\big)^1 : (E^1,0) \to (X^1,0)$ is an sc^0-contraction germ with respect to the splitting $E = (\mathbb{R}^n \oplus K) \oplus X$. Here we need the regularity by 1 because $L^{-1} P_2 \overline{B}(a,w)$ satisfies the contraction estimate in Definition 7.2 only for $m \geq 1$.

Pick a linear isomorphism $\tau : C \to K$ (which exists because $\dim K = \dim C < \infty$) and define the linear sc-isomorphism

$$\Psi = \mathbb{1} \times \tau \times L^{-1} : F = \mathbb{R}^N \oplus C \oplus R \to F = \mathbb{R}^N \oplus K \oplus R.$$

Then with the projections $P_1P : F \to X$ and $P_2P : F \to R$ we have

$$P_1P\Psi \circ (f+s) = L^{-1}P_2P \circ (f+s) = L^{-1}\phi,$$

so $\big(\Psi \circ (f+s)\big)^1 : (E^1,0) \to (F^1,0)$ is a basic germ. $\square$

The following result corresponds to [26, Theorem 3.11].

Proposition 7.10 (Local compactness of basic germs)**.**
Let $f : (E,0) \to (F,0)$ be a basic germ. Then there exists a nested sequence $U(0) \supset U(1) \supset U(2)\cdots$ of open neighbourhoods of 0 in E_0 (!) such that for each $m \geq 0$ the E_0-closure $\mathrm{cl}_{E_0}\{x \in U(m) \mid f(x) = 0\}$ is a compact subset of E_m.

Proof. Since the zero set of f is a closed subset of the zero set of its projection $P \circ f$, we can without loss of generality replace f by the contraction germ $P \circ f$. Then by the Germ Implicit Function Theorem 7.3 there exists an sc-splitting $E = K \oplus Y$ with $K = \ker Df(0)$ and a unique sc-smooth germ $\delta : (K,0) \to (Y,0)$ such that

$$f(k,y) = 0 \Longleftrightarrow y = \delta(k) \quad \text{for } (k,y) \text{ near } (0,0).$$

Sc-continuity of the germ δ means that there exist open balls $B^K(\tau_m) \subset K$ and $B^{Y_m}(\sigma_m) \subset Y_m$ of radii τ_m, σ_m around the origin such that $\delta : \overline{B}^K(\tau_m) \to \overline{B}^{Y_m}(\sigma_m)$ is continuous for each $m \geq 0$. Moreover, we can choose the radii such that the open sets $U(m) := B^K(\tau_m) \times B^{Y_0}(\sigma_m) \subset E_0$ satisfy $U(0) \supset U(1) \supset U(2)\cdots$ and $f(k,y) = 0 \Longleftrightarrow y = \delta(k)$ for $(k,y) \in \overline{U}(0)$. We claim that

$$\begin{aligned} \mathrm{cl}_{E_0}\{x \in U(m) \mid f(x) = 0\} &= \{\big(k,\delta(k)\big) \mid k \in \overline{B}^K(\tau_m)\} \\ &\subset \overline{B}^K(\tau_m) \times \overline{B}^{Y_m}(\sigma_m) \subset E_m. \end{aligned}$$

Indeed, for $(k,y) \in \mathrm{cl}_{E_0} U(m)$ with $f(k,y) = 0$ there exists a sequence $(k_n, y_n) \in U(m)$ with $f(k_n, y_n) = 0$ and $(k_n, y_n) \to (k,y)$ in E_0. From $k_n \in \overline{B}^K(\tau_m)$ it follows that $y_n = \delta(k_n) \in \overline{B}^{Y_m}(\sigma_m)$ and $k_n \to k$ in $\overline{B}^K(\tau_m)$, hence $y = \lim \delta(k_n) = \delta(k) \in \overline{B}^{Y_m}(\sigma_m)$ by continuity of $\delta : \overline{B}^K(\tau_m) \to \overline{B}^{Y_m}(\sigma_m)$. This proves the claim. Compactness of $\mathrm{cl}_{E_0}\{x \in U(m) \mid f(x) = 0\}$ in E_m now follows from compactness of $\overline{B}^K(\tau_m)$ (because K is finite dimensional) and continuity of $\delta : \overline{B}^K(\tau_m) \to \overline{B}^{Y_m}(\sigma_m)$. □

Corollary 7.11. *Let $f : (E,0) \to (F,0)$ be a basic germ and let $(x_k) \subset E_0$ be a sequence with $f(x_k) = 0$ and $x_k \to 0$ in E_0. Then for each $m \geq 0$ we have $x_k \in E_m$ for large k and $x_k \to 0$ in E_m.* □

The following corollary says that if $Df(0)$ is surjective for a basic germ f, then $Df(x)$ is surjective for x near 0 in E_0. Note that this is non-obvious because $x \mapsto Df(x)$ is not continuous with respect to the operator norm.

Corollary 7.12 (Local regularity of basic germs)**.**
Let $f : (E = \mathbb{R}^n \oplus W, 0) \to (F = \mathbb{R}^N \oplus W, 0)$ be a basic germ such that $Df(0)$ is surjective. Then there exists an open neighbourhood $U \subset E_0$ of 0 such that if $x \in U$ satisfies $f(x) = 0$, then $x \in E_1$ and $Df(x) : E_0 \to F_0$ is a surjective Fredholm operator of index $n - N$.

Proof. Let $P : \mathbb{R}^N \oplus W \to W$ be the projection and write $f = (f_0, f_1)$ with $f_1 = P \circ f$ and $f_0 = (\mathbb{1} - P) \circ f$. By hypothesis $f_1(a,w) = w - B(a,w)$ is an sc^0-contraction germ. According to Problem 7.3, there exists an open neighbourhood $U \subset E_0$ of 0 such that $|D_2 B(a,w)|_{\mathcal{L}(W_0,W_0)} < 1/2$ for all $(a,w) \in U \cap E_1$, so $D_2 f_1(a,w) = \mathbb{1} - D_2 B(a,w) : W_0 \to W_0$ is invertible. Since the linear operators $D_1 f_1(a,w) : \mathbb{R}^n \to W_0$ and

$Df_0(a,w): E_0 \to \mathbb{R}^N$ are finite dimensional, hence compact, this shows that $Df(a,w)$ is a Fredholm operator of index $n-N$ for each $(a,w) \in U \cap E_1$. By Proposition 7.10, we can choose U such that $x \in E_1$ for each $x \in U$ with $f(x)=0$.

It remains to show surjectivity of $Df(x)$ for $x \in E_0$ sufficiently close to 0 with $f(x)=0$. Using the splitting $E = K \oplus Y$ with $K = \ker Df(0)$ and the fact that $Df(x)|_Y : Y \to F$ is a Fredholm operator of index 0, this is equivalent to injectivity of $Df(x)|_Y$. Arguing by contradiction, suppose there exist sequences $(x_k) \in E_0$ with $f(x_k)=0$ (hence $x_k \in E_1$) and $x_k \to 0$ in E_0, and $(\alpha_k,\zeta_k) \in (\mathbb{R}^n \oplus W) \cap Y$ with $|(\alpha_k,\zeta_k)|_0 = 1$ and $Df(x_k)(\alpha_k,\zeta_k) = 0$. Applying the projection P to the last condition yields

$$\big((\mathbb{1} - D_2B(x_k)\big)\zeta_k = D_1B(x_k)\alpha_k. \tag{7.1}$$

Since the sequence $(\alpha_k) \subset \mathbb{R}^n$ is bounded, we may assume that $\alpha_k \to \alpha \in \mathbb{R}^n$. By Corollary 7.11 we have $x_k \to 0$ in E_1, so it follows that $z_k := D_1B(x_k)\alpha_k \to D_1B(0)\alpha =: z$ in W_0. In view of (7.1), invertibility of $\mathbb{1} - D_2B(x)$, and $D_2B(0)=0$ this implies $\zeta_k = \big((\mathbb{1} - D_2B(x_k)\big)^{-1} z_k \to z$ in W_0. Thus (α_k,ζ_k) converges in E_0 to $(\alpha,z) \in Y$ and we conclude $Df(0)(\alpha,z) = \lim_{k\to\infty} Df(x_k)(\alpha_k,\zeta_k) = 0$, contradicting injectivity of $Df(0)|_Y$. □

Problem 7.3. Let $f: (E = \mathbb{R}^n \oplus W, 0) \to (W,0)$, $f(a,w) = w - B(a,w)$ be an sc-smooth sc^0-contraction germ. Then for every $\varepsilon > 0, m \in \mathbb{N}_0$ there exists an open neighbourhood $U_m \subset E_m$ of 0 (depending on ε) such that $|D_2B(a,w)|_{\mathcal{L}(W_m,W_m)} < \varepsilon$ for all $(a,w) \in U_m \cap E_{m+1}$. Move up.

7.5 Sc-Fredholm sections

Let $P : Y \to X$ be a strong bundle over an M-polyfold X.

Definition 7.4. An sc-smooth section germ $f : (X, x) \to Y$ near a smooth point $x \in X_\infty$ is called an **sc-Fredholm germ** if there exists a strong bundle chart $\Phi : \Phi^{-1}(V) \to K$ covering $\phi : (V, x) \to (\mathcal{O}, 0)$ such that $\widetilde{f} = \Phi_* f = \Phi \circ f \circ \phi^{-1} : (\mathcal{O}, 0) \to K$ has a filling $g : (U, 0) \to U \triangleleft F$, and there exists an sc$^+$-section germ $s : (U, 0) \to U \triangleleft F$ with $s(0) = g(0)$ such that $g - s$ is conjugated to a basic germ h.

The definition is depicted in the following diagram where $t = \Psi_* s$:

$$\begin{tikzcd}
Y \supset P^{-1}(V) \arrow[r, "\Phi", "\cong"'] & K & U \triangleleft F \arrow[l, "R"'] \arrow[r, "\Psi", "\cong"'] & U' \triangleleft F' \\
X \supset V \arrow[u, "f"] \arrow[r, "\phi", "\cong"'] & \mathcal{O} \arrow[u, "\widetilde{f}"] & U \arrow[l, "r"'] \arrow[u, "g"] \arrow[r, "\psi", "\cong"'] & U'. \arrow[u, "h+t"']
\end{tikzcd} \tag{7.2}$$

Definition 7.5. An sc-smooth section $f : X \to Y$ is called an **sc-Fredholm section** if
(i) f is regularizing, i.e., if $x \in X_m$ and $f(x) \in Y_{m,m+1}$ then $x \in X_{m+1}$;
(ii) the germ of f at each smooth point $x \in X_\infty$ is an sc-Fredholm germ.

The following result corresponds to [26, Theorem 3.13].

Theorem 7.13 (Implicit function theorem for sc-Fredholm sections)**.** *Let f be an sc-Fredholm section of a strong bundle $Y \to X$. Let $x \in X_\infty$ such that $f(x) = 0$ and $T_x^v f : T_x X \to Y_x$ is surjective. Then there exists an open neighbourhood $V \subset X$ of x such that $M := \{y \in V \mid f(y) = 0\}$ is a sub-M-polyfold of X, and the induced M-polyfold structure on M is equivalent to the structure of a finite dimensional smooth manifold.*

Remark 7.14. Here as alaways in this chapter we assume that X has no boundary and corners. In the general case we have to require that the strong bundle $Y \to X$ is tame and the sc-Fredholm germ of f at x is in good position, and the conclusion is that M is a smooth manifold with boundary and corners.

Proof. Consider the diagram (7.2) for the sc-Fredholm germ $f : (X, x) \to Y$. Since $f'(x) : T_x X \to Y_x$ is surjective, so is $T\widetilde{f}(0) : T_0\mathcal{O} \to K_0$. By Proposition 7.1 this implies that $Dg(0) : E \to F$ is surjective, hence so is $Dg'(0) : E' \to F'$ for $g' := h + t$. Recall that h is a basic germ and t is an sc$^+$-germ. Thus by Proposition 7.8, after raising regularity levels by 1 and composing Ψ with a linear sc-isomorphism we may assume that g' is a basic germ. By the Germ Implicit Function Theorem 7.7 there exists an sc-splitting $E' = N' \oplus Y'$ with $N' = \ker Dg'(0)$ and a unique sc-smooth germ $\delta : (N', 0) \to (Y', 0)$ such that $D\delta(0) = 0$ and

$$g'(v, y) = 0 \Longleftrightarrow y = \delta(v) \quad \text{for } (v, y) \in U' = V' \times W'.$$

Here $V' \subset N'$ and $W' \subset Y'$ are open neighbourhoods of the origin such that $\delta : V'_0 \to W'_0$ is continuous. In view of the diagram (7.2) and again Proposition 7.1, with $V := (\psi \circ \phi)^{-1}(U')$ this implies that

$$\begin{aligned} &\psi \circ \phi\{y \in V \mid f(y) = 0\} = \psi\{u \in \mathcal{O} \mid \widetilde{f}(u) = 0\} \\ &= \psi\{u \in U \mid g(u) = 0\} = \{u' \in U' \mid g'(u') = 0\} \\ &= \{v + \delta(v) \mid v \in V'\}. \end{aligned}$$

Next note that $M := \{y \in V \mid f(y) = 0\}$ is contained in X_∞ because f is regularizing. Moreover, by hypothesis f is an sc-Fredholm germ near each $y \in M$. Set $n := \dim\ker f'(x) = \dim\ker Dg'(0)$. By Corollary 7.12, for U' sufficiently small we may assume that $Dg'(u')$ is surjective with $\dim\ker Dg'(y) = n$

for each $u \in U'$ with $g'(u') = 0$, hence $f'(y) : T_yX \to Y_y$ is surjective with $\dim \ker f'(y)$ for each $y \in M$. Thus we can repeat the preceding discussion near each $y \in M$ to write M as the graph of an sc-smooth germ δ near y. This shows that the set $M \subset X$ has the *n-dimensional tangent germ property* in the sense of Definition 7.6 below, and Proposition 7.15 below yields that M is a sub-M-polyfold and an n-dimensional manifold. □

Definition 7.6. Let X be an M-polyfold and $n \in \mathbb{N}_0$. We say that a subset $M \subset X$ has the *n-dimensional tangent germ property* if $M \subset X_\infty$ and near each $x \in M$ there exists an M-polyfold chart $\phi : U \to \mathcal{O}$ with $\phi(x) = 0$ from an open neighbourhood $U \subset X$ of x onto an sc-smooth retract $\mathcal{O}$ in an sc-Banach space E, an sc-splitting $E = N \oplus Y$ with $\dim N = n$, and a continuous map $\delta : V \to Y$ defined on an open neighbourhood $V \subset N$ of 0 such that
(i) $\phi\{M \cap U) = \{v + \delta(v) \mid v \in V\} \subset N \oplus Y$, and
(ii) $\delta : (N, 0) \to (Y, 0)$ is an sc-smooth germ with $D\delta(0) = 0$.

Problem 7.4. Show that the n-dimensional tangent germ property is diffeomorphism invariant: if $\psi : X \to X'$ is an sc-diffeomorphism of M-polyfolds and $M \subset X$ has the n-dimensional tangent germ property, then so does $\psi(M) \subset X'$.

The following result corresponds to [26, Theorem 3.6].

Proposition 7.15. *Let X be an M-polyfold and $M \subset X$ be a subset with the n-dimensional tangent germ property. Then M is a sub-M-polyfold of X, and the induced M-polyfold structure on M is equivalent to the structure of an n-dimensional smooth manifold.*

Proof. Consider a point $x \in M$. Let $\phi : X \supset U \to \mathcal{O} \subset E = N \oplus Y$, $\phi(x) = 0$, and $\delta : N \supset V \to Y$ be as in Definition 7.6

such that $\phi\{M \cap U) = \{v + \delta(v) \mid v \in V\} \subset N \oplus Y$. We choose V so small that $\delta : V \to Y_0$ is of class C^1. Let $\pi : N \oplus Y \to N$ be the projection. We will show that

$$\pi \circ \phi|_{M \cap U} : M \cap U \to V$$

is an M-polyfold chart for M. Note that this is the restriction of the sc-smooth map $\pi \circ \phi : U \to N$, it is continuous and bijective, and its inverse is the map

$$\gamma : V \to M \cap U, \qquad v \mapsto \phi^{-1}\big(v + \delta(v)\big).$$

Sc-smoothness of γ is equivalent to

Claim 1: The map $\delta : V \to Y$ (not just its germ at 0!) is sc-smooth.

To see this, consider a point $v_0 \in V$ and set $x_0 := \phi^{-1}\big(v_0 + \delta(v_0)\big) \in M \cap U$. Let $\psi : X \supset U' \to \mathcal{O}' \subset E' = N' \oplus Y'$, $\psi(x_0) = 0$, and $\tau : N' \supset V' \to Y'$ be as in Definition 7.6 such that $\psi\{M \cap U') = \{w + \tau(w) \mid w \in V'\} \subset N' \oplus Y'$. To compute the transition map between the two charts for M arising from ϕ and ψ, consider $v \in V$ and $w \in V'$ that correspond to the same point $x \in M$ via $\phi(x) = v + \delta(v)$ and $\psi(x) = w + \tau(w)$. This is equivalent to

$$v = \pi \circ \phi \circ \psi^{-1}\big(w + \tau(w)\big) := f(w),$$

where the germ $f : (N', 0) \to (N, v_0)$ is (classically) of class C^∞ at 0. Since $D\tau(0) = 0$, the differential of f at 0 is

$$Df(0) = \pi \circ D(\phi \circ \psi^{-1})(0) : N' \to N.$$

Now $D(\phi \circ \psi^{-1})(0) : N' \oplus Y' \to N \oplus Y$ is an isomorphism and maps the linear subspace $T_0\{w + \tau(w) \mid w \in V'\} = N'$ onto the subspace $T_{v_0+\delta(v_0)}\{v + \delta(v) \mid v \in V\} = \{v + D\delta(v_0)v \mid v \in N\}$. This shows that $Df(0) : N' \to N$ is an isomorphism. By the classical inverse function theorem, there exists an inverse germ $(N, v_0) \to (N', 0)$, $v \mapsto w(v)$ which is C^∞ at v_0. Thus the germ

$$v \mapsto \phi \circ \psi^{-1}\big(w(v) + \tau(w(v))\big) = v + \delta(v)$$

is C^∞ at v_0 and the claim follows.

Claim 2: $M \subset X$ is a sub-M-polyfold.

To see this, consider near $x \in M$ an M-polyfold chart $\phi : X \supset U \to \mathcal{O} \subset E = N \oplus Y$, $\phi(x) = 0$, and $\delta : N \supset V \to Y$ as above such that $\phi\{M \cap U) = \{v + \delta(v) \mid v \in V\} \subset N \oplus Y$. Here $\delta : V \to Y$ is an sc-smooth map by Claim 1. We have $\mathcal{O} = r(W)$ for an sc-smooth retraction $r : W \to W$ of an open subset $W \subset E$. Define the open subset $\Sigma := V \oplus Y \subset N \oplus Y$. Then r maps $r^{-1}(\Sigma \cap \mathcal{O})$ to itself, so $\Sigma \cap \mathcal{O}$ is an sc-smooth retract. Define the sc-smooth map

$$t : \Sigma \cap \mathcal{O} \to \Sigma \cap \mathcal{O}, \qquad v + y \mapsto v + \delta(v).$$

Then $t \circ t(v + y) = t\big(v + \delta(v)\big) = v + \delta(v) = t(v + y)$, so t is an sc-smooth retraction onto the set $t(\Sigma \cap \mathcal{O}) = \{v + \delta(v) \mid v \in V\} = \phi(M \cap U)$. It follows that $s := \phi^{-1} \circ t \circ \phi : U \to U$ is an sc-smooth retraction onto the set $s(U) = \phi^{-1} \circ t(\Sigma \cap \mathcal{O}) = M \cap U$ and Claim 2 is proved.

Claim 3: The induced M-polyfold structure on M is equivalent to the structure of an n-dimensional smooth manifold.

To see this, consider near $x \in M$ an M-polyfold chart $\phi : X \supset U \to \mathcal{O} \subset E = N \oplus Y$ and an sc-smooth retraction $s = \phi^{-1} \circ t \circ \phi : U \to U$ with $s(U) = M \cap U$ as in the proof of Claim 2. The induced M-polyfold chart for M is given by the restriction

(retaining the earlier notation)

$$\begin{aligned}
&\phi|_{M\cap U} : M \cap U \to \phi(M \cap U) = \{v + \delta(v) \mid v \in V\},\\
&x \mapsto \phi(x) = \phi\big(s(x)\big) = t\big(\phi(x)\big) = \pi \circ \phi(x) + \delta\big(\pi \circ \phi(x)\big),
\end{aligned}$$

where $\pi : N \oplus Y \to N$ is the projection. Since π induces a diffeomorphism $\{v + \delta(v) \mid v \in V\} \to V$, composition with π yields an equivalent M-polyfold chart $\pi \circ \phi|_{M\cap U} : M \cap U \to V$ onto the open subset $V \subset N \cong \mathbb{R}^n$. The collection of these charts defines on M the structure of an n-dimensional smooth manifold and Proposition 7.15 is proved. □

7.6 The sc-Fredholm section in Morse homology

In this section we show that the differential operator describing (broken) gradient flow lines of a Morse function is an sc-Fredholm section. The discussion follows [51].

In general, one considers an n-dimensional Riemannian manifold (M, g) and a proper Morse function $H : W \to \mathbb{R}$ such the pair (H, g) is Morse-Smale. For critical points p^-, p^+ one wants to construct

- an M-polyfold X of paths from p^- to p^+, modelled as maps from $\mathbb{R}$ to M of suitable weighted Sobolev class, possibly broken at other critical points;
- a strong bundle $Y \to X$ whose fibre over $u \in X$ is modeled over sections in the bundle u^*TM;
- an sc-Fredholm section f in the bundle $Y \to X$ sending $u : \mathbb{R} \to M$ to the section $\partial_s u + \nabla H(u)$, so that the zero set of f corresponds to (broken) negative gradient lines.

We will restrict to the simplest possible setup:

- $M = \mathbb{R}^n$ with the Euclidean metric;
- $H : \mathbb{R}^n \to \mathbb{R}$ has a nondegenerate critical point at the origin;
- all paths start and end at the origin;
- we consider only a neighbourhod of the once broken gradient line $(u_1, u_2) = (0, 0)$ consisting of two constant maps at the origin.

Although geometrically uninteresting, this situation already exhibits all the essential analytical features. At the end of the section we will briefly describe the necessary adjustments for the general case.

The M-polyfold chart. The M-polyfold chart near $(u_1, u_2) = (0, 0)$ is given by a small modification of the construction in Section 6.4, replacing the domains $\mathbb{R}_\pm$ by $\mathbb{R}$ and gluing in a more symmetric way.

For weights $0 < \delta_0 < \delta_1 < \delta_2 < \cdots < \delta_A$ (with δ_A to be chosen later) consider the sc-Banach space

Do we need $\delta_0 > 0$ and higher level of regularity?

$$E := H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n) \oplus H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n)$$

with scale structure

$$E_m := H^{m+1,\delta_m}(\mathbb{R}, \mathbb{R}^n) \oplus H^{m+1,\delta_m}(\mathbb{R}, \mathbb{R}^n), \qquad m \geq 0.$$

As a set, the M-polyfold is

$$X := \{0\} \times E \cup \bigcup_{a \in (0,1/2)} \{a\} \times H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n)$$

with the obvious scale structure, where $a \in [0, 1/2)$ is the *gluing parameter*. We pick the exponential *gluing profile*

$$\varphi : (0, 1] \xrightarrow{\cong} [0, \infty), \qquad \varphi(a) := e^{1/a} - e$$

to associate to each gluing parameter $a \in (0,1)$ its *gluing length*

$$R = \varphi(a).$$

We choose a smooth cutoff function $\beta : \mathbb{R} \to [0,1]$ satisfying

$$\beta(s) = 1 \text{ for } s \leq -1, \quad \beta(-s)+\beta(s) \equiv 1, \quad \beta'(s) < 0 \text{ for } s \in (-1,1).$$

For $(u_1, u_2) \in E$ and $a > 0$ define the *gluing and antigluing* maps

$$\begin{aligned} \oplus_a(u_1,u_2)(s) &:= \beta(s)u_1(s+R) + \big((1-\beta(s)\big)u_2(s-R), \\ \ominus_a(u_1,u_2)(s) &:= \big(\beta(s)-1\big)u_1(s+R) + \beta(s)u_2(s-R). \end{aligned}$$

With the notation $\tau_{\pm R}u(s) := u(s \pm R)$ this can be written more concisely as the *total gluing map*

$$\begin{gathered} \boxdot_a = (\oplus_a, \ominus_a) : E \to G^a = H^{1,\delta_0}(\mathbb{R},\mathbb{R}^n) \oplus H^{1,\delta_0}(\mathbb{R},\mathbb{R}^n), \\ \boxdot_a(u_1,u_2) = \begin{pmatrix} \beta & 1-\beta \\ \beta - 1 & \beta \end{pmatrix} \begin{pmatrix} \tau_R u_1 \\ \tau_{-R} u_2 \end{pmatrix}. \end{gathered}$$

For $a = 0$ we set $G^0 := E \oplus \{0\}$ and $\boxdot_0 := (\mathbb{1}_E, 0)$. Then we have

Lemma 7.16. *For each $a \in [0,1/2)$ the total gluing map $\boxdot_a : E \xrightarrow{\cong} G^a$ is a linear sc-isomorphism whose inverse for $a > 0$ is*

$$\boxdot_a^{-1} = \begin{pmatrix} \tau_{-R} & 0 \\ 0 & \tau_R \end{pmatrix} \frac{1}{\gamma} \begin{pmatrix} \beta & \beta - 1 \\ 1-\beta & \beta \end{pmatrix}, \quad \gamma = \beta^2 + (1-\beta)^2 > 0. \tag{7.3}$$

It follows that for each $a \in [0,1/2)$ we have

$$E = \ker \oplus_a \oplus \ker \ominus_a,$$

and $\oplus_a$ restricts to an isomorphism

$$\oplus_a|_{\ker \ominus_a} : \ker \ominus_a \xrightarrow{\cong} H^{1,\delta_0}(\mathbb{R},\mathbb{R}^n)$$

Let $\pi_a : E \to E$ be the projection onto $\ker \ominus_a$ along $\ker \oplus_a$. Thus $\pi_0 = \mathbb{1}_E$ and for $a > 0$ it is given by

$$\pi_a = \boxdot_a^{-1} \begin{pmatrix} \oplus_a \\ 0 \end{pmatrix} = \begin{pmatrix} \tau_{-R} & 0 \\ 0 & \tau_R \end{pmatrix} \frac{1}{\gamma} \begin{pmatrix} \beta & \beta - 1 \\ 1 - \beta & \beta \end{pmatrix} \begin{pmatrix} \oplus_a \\ 0 \end{pmatrix} .$$

Proposition 7.17. *The map*

$$r : [0, \varepsilon) \oplus E \to [0, \varepsilon) \oplus E, \qquad (a, u_1, u_2) \mapsto \big(a, \pi_a(u_1, u_2)\big)$$

is an sc-smooth retraction onto its image

$$\mathcal{O} = \{0\} \times E \cup \bigcup_{a \in (0,1/2)} \{a\} \times \ker \ominus_a .$$

We now define the M-polyfold structure on X with the global chart

$$\phi^{-1} : \mathcal{O} \xrightarrow{\cong} X$$

given by $\mathbb{1}_E$ for $a = 0$ and by $\oplus_a : \ker \ominus_a \xrightarrow{\cong} H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n)$ for $a > 0$.

The strong bundle. The strong bundle $Y \to X$ is constructed in exactly the same way, just starting with one level of regularity less. Thus we define the sc-Banach space

$$F := H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n) \oplus H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n)$$

with scale structure

$$F_m := H^{m,\delta_m}(\mathbb{R}, \mathbb{R}^n) \oplus H^{m,\delta_m}(\mathbb{R}, \mathbb{R}^n), \qquad m \geq 0.$$

The same formulas as before define a total gluing map (which we decorate with a $'$ to distinuish it from the previous one)

$$\boxdot_a' = (\oplus_a', \ominus_a') : F \to H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n) \oplus H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n),$$

extended by $\boxdot'_0 := (\mathbb{1}_F, 0)$ for $a = 0$. As before, let $\pi'_a : F \to F$ be the projection onto $\ker \ominus'_a$ along $\ker \oplus'_a$. We obtain a local strong bundle

$$\big([0,1/2) \oplus E\big) \triangleleft F \to [0,1/2) \oplus E$$

and an sc-smooth strong bundle retraction

$$\begin{gathered} R : \big([0,\varepsilon) \oplus E\big) \triangleleft F \to \big([0,\varepsilon) \oplus E\big) \triangleleft F, \\ (a, u_1, u_2, v_1, v_2) \mapsto \big(a, \pi_a(u_1, u_2), \pi'_a(v_1, v_2)\big) \end{gathered}$$

onto its image

$$\mathcal{K} = \{0\} \times E \triangleleft F \cup \bigcup_{a \in (0,1/2)} \{a\} \times \ker \ominus_a \triangleleft \ker \ominus'_a.$$

We define as a set

$$Y := \{0\} \times E \triangleleft F \cup \bigcup_{a \in (0,\varepsilon)} \{a\} \times H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n) \triangleleft H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n)$$

with its obvious projection $P : Y \to X$ and the global strong bundle chart

$$\Phi^{-1} : \mathcal{K} \xrightarrow{\cong} Y$$

covering ϕ^{-1} given by $\mathbb{1}_E \times \mathbb{1}_F$ for $a = 0$ and by $\oplus_a \times \oplus'_a$: $\ker \ominus_a \triangleleft \ker \ominus'_a \xrightarrow{\cong} H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n) \triangleleft H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n)$ for $a > 0$.

The section. For a map $u : \mathbb{R} \to \mathbb{R}^n$ we set

$$\partial_H(u) := \partial_s u + \nabla H(u).$$

This defines a section $f : X \to Y$ whose principal part is

$$\begin{aligned} f(0, u_1, u_2) &:= \big(\partial_H(u_1), \partial_H(u_2)\big) \in F \quad \text{on } \{0\} \times E, \\ f(a, u) &:= \partial_H(u) \in H^{0,\delta_0}(\mathbb{R}, \mathbb{R}^n) \quad \text{on } \{a\} \times H^{1,\delta_0}(\mathbb{R}, \mathbb{R}^n). \end{aligned}$$

The principal part of its local representative

$$\widetilde{f} := \Phi \circ f \circ \phi^{-1} : \mathcal{O} \to \mathcal{K}$$

is given for $a=0$ by $\widetilde{f}|_{\{0\}\times E} = f|_{\{0\}\times E}$, and for $a>0$ via (7.3) by

$$\begin{aligned}\widetilde{f}(a,u_1,u_2) &= (\boxdot'_a)^{-1}\begin{pmatrix}\partial_H \oplus_a (u_1,u_2)\\ 0\end{pmatrix}\\ &= \begin{pmatrix}\tau_{-R} & 0\\ 0 & \tau_R\end{pmatrix}\frac{1}{\gamma}\begin{pmatrix}\beta\partial_H \oplus_a (u_1,u_2)\\ (1-\beta)\partial_H \oplus_a (u_1,u_2)\end{pmatrix}.\end{aligned}$$

Since $\oplus_a \circ \pi_a = \oplus_a$, the principal part of the composition $\widetilde{f}\circ r : [0,1/2)\oplus E \to [0,1/2)\oplus F$ is given by the same formula and we read off that it is sc-smooth, hence so is the section f. The goal of this section is to prove

Theorem 7.18. *$f: X \to Y$ is an sc-Fredholm section.*

The filling. We first need to find a filling g for $\widetilde{f}$ fitting into the diagram

$$\begin{array}{ccccc} Y & \xrightarrow[\cong]{\Phi} & K & \xleftarrow{R} & [0,1/2)\oplus E \triangleleft F\\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle \widetilde{f}} & & \uparrow{\scriptstyle g}\\ X & \xrightarrow[\cong]{\phi} & \mathcal{O} & \xleftarrow{r} & [0,1/2)\oplus E. \end{array}$$

The idea is to fill the zero second component in the definition of $\widetilde{f}$ using the antigluing $\ominus_a(u_1,u_@)$ and a linear isomorphism $\ker \oplus_a \xrightarrow{\cong} \ker \oplus'_a$. A suitable isomorphism is given by

$$\partial_A := \partial_s + A, \qquad A := D^2_0 H,$$

where $D^2_p H$ denotes the Hessian of H at $p\in\mathbb{R}^n$, in view of *Problem* 7.5. For each nondegenerate symmetric $n\times n$-matrix A the first order linear differential operator $\partial_A = \partial_s + A$ defines an isomorphism

$$\partial_A : H^{m+1,w}(\mathbb{R},\mathbb{R}^n) \xrightarrow{\cong} H^{m,w}(\mathbb{R},\mathbb{R}^n)$$

for each $m\in\mathbb{N}_0$ and each weight $w\in\mathbb{R}$ satisfying $|w|<\delta_A$, where δ_A denotes the minimal distance of an eigenvalue of A to $0\in\mathbb{R}$.

We define the principal part $g : [0,1/2) \oplus E \to F$ of the section g by

$$g(a, u_1, u_2) := (\boxdot'_a)^{-1} \begin{pmatrix} \partial_H \oplus_a (u_1, u_2) \\ \partial_A \ominus_a (u_1, u_2) \end{pmatrix} .$$

In the following we will often abbreviate $u = (u_1, u_2)$.

Lemma 7.19. *The section g is a filling of $\widetilde{f}$.*

Proof. We need to check properties (i)–(iii) in Definition 7.2.

(i) $g(x) = \widetilde{f}(x)$ for each $x = (a, u) \in \mathcal{O}$:
For $a = 0$ we have $g(0, u) = (\partial_H u_1, \partial_H u_2) = \widetilde{f}(0, u)$. If $a > 0$ then $(a, u) \in \mathcal{O}$ implies $u \in \ker \ominus_a$ and thus

$$g(a, u) = (\boxdot'_a)^{-1} \begin{pmatrix} \partial_H \oplus_a (u_1, u_2) \\ 0 \end{pmatrix} = \widetilde{f}(a, u)$$

(ii) If $x = (a, u) \in [0, 1/2) \oplus E$ satisfies $g(x) = \rho(r(x))g(x)$, then $x \in \mathcal{O}$, where $R(x, h) = \big(r(x), \rho(x)h\big)$:
Here $\rho(a, u) = \pi'_a : F \to F$ is independent of u and we obtain the equivalences

$$\begin{aligned} & g(a, u) = \rho(r(a, u))g(a, u) = \pi'_a g(a, u) \\ & \iff g(a, u) \in \ker \ominus'_a \iff \partial_A \ominus_a (u_1, u_2) = 0 \\ & \iff \ominus_a(u_1, u_2) = 0 \iff (u_1, u_2) \in \ker \ominus_a \\ & \iff (a, u) \in \mathcal{O}. \end{aligned}$$

(iii) The linearization of the map $x \mapsto \big(\mathbb{1} - \rho(r(x))\big)g(x)$ at 0 restricts to a continuous linear isomorphism $\ker Dr(0) \xrightarrow{\cong} \ker \rho(0)$:
This is automatic because $Dr(0) = \mathbb{1}_{\mathbb{R}\oplus E}$ and $\rho(0) = \mathbb{1}_F$ imply $\ker Dr(0) = \{0\} = \ker \rho(0)$. □

Conjugation to a basic germ. It remains to show that, after possibly raising regularity levels by one, g is conjugated to a basic germ. For this observe that for each fixed $a \in [0, 1/2)$ and each $m \geq 0$ the map $E_m \mapsto F_m$, $u \mapsto g(a, u)$ is *classically smooth*. By contrast, the dependence on the gluing parameter a involves reparametrizations and is therefore only sc-smooth. In such a situation, a criterion due to Katrin Wehrheim expresses conjugation to a basic germ entirely in terms of the partial derivative $D_E g : [0, 1/2) \oplus E \to \mathcal{L}(E, F)$ with respect to the second variable u. The criterion is stated and proved in the next subsection. Here we will apply it to prove the following proposition which concludes the proof of Theorem 7.18.

Proposition 7.20. *After raising regularity levels by one, the filling $g : [0, 1/2) \oplus E \to F$ above is conjugated to a basic germ.*

Proof. We will check the conditions for Wehrheim's criterion in Definition 7.7. Condition (i) holds by ellitpic regularity for the operators ∂_H and ∂_A. For conditions (ii) and (iii) we compute the partial derivative $D_E g : [0, 1/2) \oplus E \to \mathcal{L}(E, F)$. At $a = 0$ and $u = (u_1, u_2) \in E$ it is given on $\xi = (\xi_1, \xi_2) \in E$ by

$$D_E g(0, u)\xi = \begin{pmatrix} D\partial_H(u_1)\xi_1 \\ D\partial_H(u_2)\xi_2 \end{pmatrix} .$$

For $a > 0$ we abbreviate

$$u_a := \oplus_a(u_1, u_2), \qquad u_a^- := \tau_{-R} u_a.$$

Then

$$D_E g(a, u)\xi = (\boxdot_a')^{-1} \begin{pmatrix} K^+\xi \\ K^-\xi \end{pmatrix} = \begin{pmatrix} \tau_{-R}\Big(\frac{\beta}{\gamma}K^+\xi + \frac{\beta-1}{\gamma}K^-\xi\Big) \\ \tau_R\Big(\frac{1-\beta}{\gamma}K^+\xi + \frac{\beta}{\gamma}K^-\xi\Big) \end{pmatrix}$$

with the linear operators

$$\begin{aligned}
K^+\xi &= D(\partial_H \circ \oplus_a)(u)\xi \\
&= D\partial_H(u_a) \oplus_a (\xi_1, \xi_2) \\
&= (\partial_s + D^2_{u_a})\big(\beta\tau_R\xi_1 + (1-\beta)\tau_{-R}\xi_2\big) \\
&= \beta'(\tau_R\xi_1 - \tau_{-R}\xi_2) + \beta(\partial_s + D^2_{u_a})\tau_R\xi_1 + (1-\beta)(\partial_s + D^2_{u_a})\tau_{-R}\xi_2 \\
&= \beta'(\tau_R\xi_1 - \tau_{-R}\xi_2) + \beta(\partial_s + A)\tau_R\xi_1 + \beta(D^2_{u_a} - A)\tau_R\xi_1 \\
&\quad + (1-\beta)(\partial_s + D^2_{u_a})\tau_{-R}\xi_2,
\end{aligned}$$

$$\begin{aligned}
K^-\xi &= D(\partial_A \circ \ominus_a)(u)\xi \\
&= \partial_A \ominus_a (\xi_1, \xi_2) \\
&= (\partial_s + A)\big((\beta-1)\tau_R\xi_1 + \beta\tau_{-R}\xi_2\big) \\
&= \beta'(\tau_R\xi_1 + \tau_{-R}\xi_2) + (\beta-1)(\partial_s + A)\tau_R\xi_1 + \beta(\partial_s + A)\tau_{-R}\xi_2.
\end{aligned}$$

Plugging this into the formula above for $D_E g(a,u)\xi$ we obtain after some manipulations

$$D_E g(a,u)\xi = \begin{pmatrix} (\partial_s + A)\xi_1 + E_1(a,u)\xi \\ (\partial_s + A)\xi_2 + E_2(a,u)\xi \end{pmatrix}$$

with the zero order operator

$$\begin{aligned}
E_1(a,u)\xi = B_s^{[R-1,R+1]}&\tau_{-R}\big(2\beta - 1)\tau_R\xi_1 - \tau_{-R}\xi_2\big) \\
&+ B^{(-\infty,R+1]}(D^2_{u_a^-}H - A)\xi_1 \\
&+ B^{[R-1,R+1]}(D^2_{u_a^-}H - A)\tau_{-2R}\xi_2
\end{aligned} \tag{7.4}$$

and a similar formula for $E_2(a,u)\xi$, where we have introduced the functions (with superscript denoting their support)

$$B_s^{[R-1,R+1]} := \tau_{-R}\left(\frac{\beta'}{\gamma}\right), \quad B_s^{(-\infty,R+1]} := \tau_{-R}\left(\frac{\beta^2}{\gamma}\right),$$
$$B^{[R-1,R+1]} := \tau_{-R}\left(\frac{\beta(1-\beta)}{\gamma}\right).$$

Now we are ready to verify conditions (ii) and (iii) in Definition 7.7.

(iii) By Problem 7.5, the linear operator

$$D_E g(0,0)\xi = \begin{pmatrix} \partial_A \xi_1 \\ \partial_A \xi_2 \end{pmatrix}$$

defines a sc-isomorphism $D_E g(0,0) : E \to F$, hence in particular an sc-Fredholm operator of index 0. Next note that for $u = 0$ only the first term remains in the expression (7.4) for $E_1(a,u)$ and this term has compact support in $[R-1, R+1]$, and similarly for $E_2(a,u)$. Thus by the Sobolev–Rellich embedding theorem, $D_E g(a,0) = D_E g(0,0) + B(a) : E_0 \to F_0$ for a compact operator $B(a)$, so it is Fredholm of index 0.

(ii) The computation above shows that for each a the map $g(a,\cdot)$ is differentiable with derivative $D_E g(a,u)$ at $u \in E$. For part (a) we need to uniformly estimate for $u, v, \xi \in E_m$ the difference

$$|D_E g(a,u)\xi - D_E(a,v)\xi|_{F_m} = \sum_{i=1}^{2} |E_i(a,u)\xi - E_i(a,v)\xi|_{H^{m,\delta_m}}.$$

Let us cosider for example the second term in (7.4), which looks most dangerous because it is not compactly supported. We estimate its contribution by

$$\begin{aligned} &|B^{(-\infty,R+1]}(D^2_{u_a^-}H - D^2_{v_a^-}H)\xi_1|_{H^{m,\delta_m}} \\ &\le |B^{(-\infty,R+1]}|_{C^m} |D^2_{u_a^-}H - D^2_{v_a^-}H|_{C^m} |\xi_1|_{H^{m,\delta_m}}, \end{aligned}$$

where $|B^{(-\infty,R+1]}|_{C^m} = |\tau_{-R}(\beta^2/\gamma)|_{C^m} \le C$ and

$$|D^2_{u_a^-}H - D^2_{v_a^-}H|_{C^m} \le C|u_a^- - v_a^-|_{C^m} \le C|u-v|_{C^m} \le C|u-v|_{H^{m+1,\delta_m}}$$

for a generic constant independent of a, u, v, ξ and the last inequality follows from the Sobolev embedding theorem. Similar estimates for the other terms yield altogether an estimate

$$|D_E g(a,u)\xi - D_E(a,v)\xi|_{F_m} \le C|u-v|_{E_m}|\xi_{E_m}$$

which proves the desired uniform continuity and thus part (a).

For part (b) consider sequences $a^\nu \in (0,1/2)$ and $(\xi^\nu) \subset B_1^{E_m}$ with $a^\nu \to 0$ and $|D_E g(a^\nu,0)\xi^\nu|_{F_m} \to 0$. We will show that (even without passing to a subsequence) $|D_E g(0,0)\xi^\nu|_{F_m} \to 0$, or equivalently $|D_E g(a^\nu,0)\xi^\nu - D_E g(0,0)\xi^\nu|_{F_m} \to 0$. Recall from (7.4) that $D_E g(a,0) = D_E g(0,0) + B(a)$ for a linear operator $B(a)$ whose first component is given by

$$B_1(a)\xi = B_s^{[R-1,R+1]}\tau_{-R}\big(2\beta-1)\tau_R\xi_1 - \tau_{-R}\xi_2\big)$$

and whose second component $B_2(a)$ is similar. The strategy now is to estimate $B(a)\xi$ in terms of $D_E g(a,0)\xi$. For this, we first use formula (7.3) for $\boxdot_a^{-1}$ to express (ξ_1,ξ_2) is terms of $\oplus_a(\xi)$ and $\ominus_a(\xi)$ as

$$\begin{aligned}\xi_1 &= \tau_{-R}\Big(\frac{\beta}{\gamma}\oplus_a(\xi) + \frac{\beta-1}{\gamma}\ominus_a(\xi)\Big),\\ \xi_2 &= \tau_R\Big(\frac{1-\beta}{\gamma}\oplus_a(\xi) + \frac{\beta}{\gamma}\ominus_a(\xi)\Big).\end{aligned}$$

Inserting this we obtain after a short computation

$$(2\beta-1)\tau_R\xi_1 - \tau_{-R}\xi_2 = \frac{2\beta^2-1}{\gamma}\oplus_a(\xi) + \frac{2\beta^2-4\beta+1}{\gamma}\ominus_a(\xi),$$

which combined with $B_s^{[R-1,R+1]} = \tau_{-R}(\beta'/\gamma)$ gives

$$B_1(a)\xi = \tau_{-R}\Big(\frac{\beta'(2\beta^2-1)}{\gamma^2}\oplus_a(\xi) + \frac{\beta'(2\beta^2-4\beta+1)}{\gamma^2}\ominus_a(\xi)\Big)$$

and a similar expression for $B_2(a)\xi$. So with $w=\delta_m$ and a generic constant C independent of a,ξ we can estimate

$$\begin{aligned}|B_1(a)\xi|_{H^{m,w}} &\le C\Big(|\tau_{-R}\oplus_a(\xi)|_{H^{m,w}([R-1,R+1])} + |\tau_{-R}\ominus_a(\xi)|_{H^{m,w}([R-1,R+1])}\Big)\\ &\le Ce^{wR}\Big(|\oplus_a(\xi)|_{H^m([-1,1])} + |\ominus_a(\xi)|_{H^m([-1,1])}\Big),\end{aligned} \tag{7.5}$$

and similarly for $B_2(a)\xi$ with the same right hand side. Now comes the crucial argument. Using $u_a = 0$ for $u = (0,0)$ we invert the equation

$$\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} = D_E g(a,0)\xi = (\boxdot_a')^{-1}\begin{pmatrix}\partial_A \oplus_a(\xi)\\ \partial_A \ominus_a(\xi)\end{pmatrix}$$

to obtain

$$\begin{pmatrix}\beta\tau_R\phi_1 + (1-\beta)\tau_{-R}\phi_2\\ (\beta-1)\phi_1 + \beta\tau_{-R}\phi_2\end{pmatrix} = \begin{pmatrix}\partial_A \oplus_a(\xi)\\ \partial_A \ominus_a(\xi)\end{pmatrix}.$$

So invertibility of ∂_A from Problem 7.5 allows us to estimate the second tern in (7.5) by

$$\begin{aligned}
C^{-1}|\ominus_a(\xi)|_{H^m([-1,1])} &\le C^{-1}|\ominus_a(\xi)|_{H^{m+1,w}(\mathbb{R})}\\
&\le |\partial_A \ominus_a(\xi)|_{H^{m,w}(\mathbb{R})}\\
&= |(\beta-1)\tau_R\phi_1 + \beta\tau_{-R}\phi_2|_{H^{m,w}(\mathbb{R})}\\
&\le |\beta|_{C^m}\Big(|\tau_R\phi_1|_{H^{m,w}([-1,\infty))} + |\tau_{-R}\phi_2|_{H^{m,w}((-\infty,1])}\Big)\\
&\le Ce^{-wR}\Big(|\phi_1|_{H^{m,w}([R-1,\infty))} + |\phi_2|_{H^{m,w}((-\infty,1-R])}\Big)\\
&\le Ce^{-wR}\Big(|\phi_1|_{H^{m,w}(\mathbb{R})} + |\phi_2|_{H^{m,w}(\mathbb{R})}\Big)\\
&\le Ce^{-wR}|D_E g(a,0)\xi|_{H^{m,w}}.
\end{aligned}$$

The same argument applied to the first tern in (7.5) leads to a term e^{wR} which does not suffice for the proof. Instead, we pick a smooth family of cutoff functions $\psi_T : \mathbb{R} \to [0,1]$, $T \ge 2$, supported in $[-T,T]$ with $\psi_T|_{[-T+1,T-1]} \equiv 1$ and with uniformly bounded derivatives. Applying Problem 7.5 with a *negative* weight

$w' \in (-\Delta_A, -w)$, we estimate for $T \leq R/2$:

$$\begin{aligned}
&C^{-1}|\oplus_a(\xi)|_{H^m([-1,1])} \leq C^{-1}|\psi_T \oplus_a(\xi)|_{H^{m+1,w'}(\mathbb{R})} \\
&\leq |\partial_A\big(\psi_T \oplus_a(\xi)\big)|_{H^{m,w'}(\mathbb{R})} \\
&\leq |\psi_T' \oplus_a(\xi)|_{H^{m,w'}} + |\psi_T\big(\beta\tau_R\phi_1 + (1-\beta)\tau_{-R}\phi_2\big)|_{H^{m,w}(\mathbb{R})} \\
&\leq |\psi_T|_{C^{m+1}}|\beta|_{C^m}\Big(|\tau_R\xi_1|_{H^{m,w'}(\operatorname{supp}\psi_T')} + |\tau_{-R}\xi_2|_{H^{m,w'}(\operatorname{supp}\psi_T')} \\
&\quad + |\tau_R\phi_1|_{H^{m,w'}([-T,T])} + |\tau_{-R}\phi_2|_{H^{m,w'}([-T,T])}\Big) \\
&\leq Ce^{-wR}\Big(e^{(w'+w)T}\sum_{i=1}^{2}|\xi_i|_{H^{m,w}} + e^{wT}\sum_{i=1}^{2}|\phi_i|_{H^{m,w}}\Big) \\
&= Ce^{-wR}\Big(e^{(w'+w)T}|\xi|_{H^{m,w}} + e^{wT}|D_Eg(a,0)\xi|_{H^{m,w}}\Big).
\end{aligned}$$

Here the last inequality follows from the estimates

$$\begin{aligned}
|\tau_{\pm R}\xi|_{H^{m,w'}(\operatorname{supp}\psi_T')} &\leq Ce^{w'T}|\tau_{\pm R}\xi|_{H^m([-T,T])} \\
&= Ce^{w'T}|\xi|_{H^m([\pm R-T,\pm R+T])} \\
&\leq Ce^{w'T}e^{-w(R-T)}|\xi|_{H^{m,w}(\mathbb{R})}, \\
|\tau_{\pm R}\phi|_{H^{m,w'}([-T,T])} &\leq |\tau_{\pm R}\phi|_{H^m([-T,T])} \\
&= |\phi|_{H^m([\pm R-T,\pm R+T])} \\
&\leq Ce^{-w(R-T)}|\phi|_{H^{m,w}(\mathbb{R})}.
\end{aligned}$$

Inserting the preceding estimates into (7.5) and using $|\xi^\nu|_{H^{m,w}} \leq 1$ we obtain

$$|B(a)\xi^\nu|_{H^{m,w}} \leq C\Big(e^{(w'+w)T^\nu} + e^{wT^\nu}|D_Eg(a^\nu,0)\xi^\nu|_{H^{m,w}}\Big)$$

for any sequence of numbers $T^\nu \in [2, R^\nu/2]$. Since $|D_Eg(a^\nu,0)\xi^\nu|_{H^{m,w}} \to 0$ as $\nu \to \infty$, we can choose the T^ν such that $T^\nu \to \infty$ and the second term still converges to zero. Then the first term converges to zero because $w' + w < 0$ and we conclude $|B(a)\xi^\nu|_{H^{m,w}} \to 0$.

In view of the discussion at the beginning of the proof of (ii)(b) this establishes condition (ii)(b) and thus concludes the proof of Proposition 7.20. □

To do. **Adjustments in the general case.**

7.7 Wehrheim's criterion for conjugation to a basic germ

We denote by B_r^E the open ball of radius r in a Banach space E.

Definition 7.7 (Wehrheim [51])**.** Let E, F be sc-Banach spaces. An sc-smooth germ $g : (E, 0) \to F$ is an *sc-Fredholm germ with respect to an sc-splitting* $E = \mathbb{R}^d \oplus E'$ if for each $m \in \mathbb{N}_0$ there exists an $r_m > 0$ with the following properties.
(i) (regularity). g is regularizing as a germ, i.e., $e \in B_{r_m}^{E_m}$ and $g(e) \in F_{m+1}$ implies $e \in E_{m+1}$.
(ii) (continuity). For each $a \in B_{r_m}^{\mathbb{R}^d}$ the map $g(a, \cdot) : B_{r_m}^{E'_m} \to F_m$ is differentiable and the partial derivative $D_{E'}g(a, u)\xi := \frac{d}{dt}|_{t=0}g(a, u+t\xi)$ satisfies:
(a) For each $a \in B_{r_m}^{\mathbb{R}^d}$ the map

$$B_{r_m}^{E'_m} \to \mathcal{L}(E'_m, F_m), \qquad u \mapsto D_{E'}g(a, u)$$

is continuous, uniformly for (a, u) in a neighbourhood of $(0, 0)$. That is, for each $\varepsilon > 0$ there exists a δ with $0 < \delta < r_m$ such that

$$|D_{E'}g(a, u) - D_{E'}g(a, v)|_{\mathcal{L}(E'_m, F_m)} < \varepsilon \quad \text{for all } a \in B_\delta^{\mathbb{R}^d},\ u, v \in B_\delta^{E'_m}.$$

(b) Given sequences $(a^\nu) \subset \mathbb{R}^d$ and $(\xi^\nu) \subset B_1^{E'_m}$ with $a^\nu \to 0$ and $|D_{E'}g(a^\nu, 0)\xi^\nu|_{F_m} \to 0$, there exists a subsequence of (ξ^ν) for which $|D_{E'}g(0, 0)\xi^\nu|_{F_m} \to 0$.
(iii) (linear Fredholm). The linear operator $D_{E'}g(0, 0)$:

$E' \to F$ is sc-Fredholm, and for each $a \in B^{\mathbb{R}^d}_{r_0}$ the linear operator $D_{E'}g(a,0) : E'_0 \to F_0$ is Fredholm with the same index as $D_{E'}g(0,0)$.

Remark 7.21. The additional condition under (iii) in [51] that $D_{E'}g(a,0)$ is weakly regularizing in the sense that $\ker D_{E'}g(a,0) \subset E'_1$ is not needed for the following proposition.

Proposition 7.22 (Wehrheim's criterion [51])**.** *Let $g : (E,0) \to F$ be an sc-Fredholm germ with respect to an sc-splitting $E = \mathbb{R}^d \oplus E'$. Then $g^1 : (E^1,0) \to F^1$ (with regularity raised by 1) is an sc-Fredholm germ.*

Proof. Since $D_{E'}g(0,0) : E' \to F$ is an sc-Fredholm operator, there exist sc-splittings

$$E' = K \oplus W, \qquad F = \operatorname{im} D_{E'}g(0,0) \oplus C$$

with $K = \ker D_{E'}g(0,0)$ and C finite dimensional such that $D_{E'}g(0,0)|_W : W \xrightarrow{\cong} \operatorname{im} D_{E'}g(0,0)$ is an sc-isomorphism. We denote by $\pi_C : F \to C$ and $\pi_C^\perp := \mathbb{1} - \pi_C : F \to \operatorname{im} D_{E'}g(0,0)$ the projections and set for each $a \in \mathbb{R}^d$

$$L(a) := \pi_C^\perp \circ D_{E'}g(a,0)|_W : W \to \operatorname{im} D_{E'}g(0,0).$$

Claim 1. For each $m \geq 1$ there exist $r_m, C_m > 0$ such that

$$|w|_{W_m} \leq C_m |L(a)w|_{F_m} \qquad \text{for all } a \in B^{\mathbb{R}^d}_{r_m},\ w \in W_m. \tag{7.6}$$

To see this, note first that (7.6) holds for $a = 0$. Now we argue by contradiction and suppose that for some $m \geq 1$ there exist sequences $(a^\nu) \subset \mathbb{R}^d$ and $(w^\nu) \subset W_m$ with $a^\nu \to 0$ and $|w^\nu|_{W_m} = 1$ such that $|L(a^\nu)w^\nu|_{F_m} \to 0$.
Suppose first that $|\pi_C D_{E'}g(a^\nu,0)w^\nu|_{F_m}$ is uniformly bounded. Since $\dim C < \infty$, after passing to a subsequence we may assume that

$\pi_C D_{E'} g(a^\nu, 0)w^\nu \to c \in C$. Since $(w^\nu) \subset W_m$ is bounded and the embedding $W_m \hookrightarrow W_{m-1}$ is compact, we may also assume that w^ν converges in W_{m-1} to some $w \in W_{m-1}$. Since g is sc^1 at $(0,0)$, it follows that $D_{E'} g(a^\nu, 0)w^\nu \to D_{E'} g(0,0)w$ in F_{m-1}. Now continuity of π_C implies $\pi_C D_{E'} g(0,0)w = c$, hence $c = 0$ because $D_{E'} g(0,0)w \in \operatorname{im} D_{E'} g(0,0)$. It follows that $|D_{E'} g(a^\nu, 0)w^\nu|_{F_m}|to0$, hence by condition (ii)(b) there exists a subsequence of (w^ν) for which $|D_{E'} g(0,0)w^\nu|_{F_m} \to 0$, thus $|w^\nu|_{F_m} \to 0$ by (7.6) for $a = 0$, contradicting the hypothesis $|w^\nu| = 1$.
Next suppose that for some subsequence we have $t^\nu := |\pi_C D_{E'} g(a^\nu, 0)w^\nu|_{F_m}$ ∞. Then $\overline{w}^\nu := w^\nu / t^\nu \to 0$ in W_m and $|\pi_C D_{E'} g(a^\nu, 0)\overline{w}^\nu|_{F_m} = 1$ for all ν. Since $\dim C < \infty$, after passing to a subsequence we may assume that $\pi_C D_{E'} g(a^\nu, 0)\overline{w}^\nu \to c \in C$ with $|c|_{F_m} = 1$. Since g is sc^1 at $(0,0)$, it follows that $D_{E'} g(a^\nu, 0)\overline{w}^\nu \to D_{E'} g(0,0) \cdot 0 = 0$ in F_{m-1}. Now continuity of π_C implies $c = \pi_C D_{E'} g(0,0) \cdot 0 = 0$, contradicting $|c|_{F_m} = 1$, and the claim is proved.

We raise regularity by one and define the splitting and open subset

$$U^1 := (B^{\mathbb{R}^d}_{r_1} \oplus K) \oplus W^1 \subset (\mathbb{R}^d \oplus K) \oplus W^1 = E^1.$$

The map

$$\Phi : U^1 \triangleleft F^1 \to U^1 \triangleleft (C \oplus W^1), \quad (e, f) \mapsto \Big(e, \pi_C f, L(a)^{-1} \pi_C^\perp f\Big)$$

where $e = (a, k, w)$ defines a strong bundle isomorphism over $\mathbb{1} : U^1 \to U^1$ because the right hand side depends only on $a \in \mathbb{R}^d$. Note that here we need to raise regularity by one in order to apply Claim 1. Since the constant section $g(0,0)$ is sc^+, the following claim concludes the proof of Proposition 7.22.

Claim 1. $\Phi \circ \big(g - g(0,0)\big) : U^1 \to C \oplus W^1$ defines a basic germ at 0.

To see this, we write its projection onto W as

$$\pi_W \circ \Phi \circ \big(g - g(0,0)\big)(a,k,w) = w - B(a,k,w)$$

with

$$B(a,k,w) = w - L(a)^{-1}\pi_C^\perp\big(g(a,k+w) - g(0,0)\big).$$

We will show that this is an sc^0-contraction germ. Applying $L(a) = \pi_C^\perp \circ D_{E'}g(a,0)|_W$ to both sides we obtain

$$-L(a)B(a,k,w) = \pi_C^\perp\big(g(a,k+w) - g(0,0) - D_{E'}g(a,0)w\big).$$

Using this together with Claim 1 we estimate for each $m \geq 1$, $a \in B^{\mathbb{R}^d}_{r_m}$, $k \in B^K_{r_m/2}$ and $w, w' \in B^{E'}_{r_2/2}$:

$$\begin{aligned}
&|B(a,k,w) - B(a,k,w')|_{W_m} \\
&\leq C_m|L(a)B(a,k,w) - L(a)B(a,k,w')|_{F_m} \\
&= C_m|\pi_C^\perp\big(g(a,k+w) - g(a,k+w') - D_{E'}g(a,0)(w-w')\big)|_{F_m} \\
&\leq C_m|g(a,k+w) - g(a,k+w') - D_{E'}g(a,0)(w-w')|_{F_m} \\
&= C_m\Big|\int_0^1 \big[D_{E'}g\big(a,k+tw+(1-t)w'\big) - D_{E'}g(a,0)\big](w-w')dt\Big|_{F_m},
\end{aligned}$$

where the last equality follows from continuous differentiability of the map $t \mapsto g\big(a,k+tw+(1-t)w'\big)$. By condition (ii)(a) in Definition 7.7, for each $\varepsilon > 0$ there exists $\delta > 0$ such that for $|k|, |w|, |w'| < \delta$ the norm of the integral in the last line is estimated by $\frac{\varepsilon}{C_m}|w-w'|_{W_m}$, and therefore $|B(a,k,w)-B(a,k,w')|_{W_m} \leq \varepsilon|w-w'|_{W_m}$. This is the required property for an sc^0-contraction germ in Definition 7.2, so Claim 2 and Proposition 7.22 are proved.

□

Chapter 8

Perturbation theory on M-polyfolds

The goal of this chapter is to prove the following result: Given an sc-Fredholm section f in a strong bundle $P : Y \to X$ with compact solution set $f^{-1}(0)$, then there exists an sc$^+$-section s such that $(f+s)^{-1}(0)$ is compact and transversely cut out. For this, we will first construct a space Γ of sc$^+$-sections such that $(f+s)^{-1}(0)$ is compact for all $s \in \Gamma$. In the second step we will find a dense subset $\Gamma^{\mathrm{reg}} \subset \Gamma$ such that $f+s$ is transverse to the zero section for all $s \in \Gamma^{\mathrm{reg}}$.

The reference for this chapter is [26, Chapter 5]. Throughout this chapter, $P : Y \to X$ denotes a strong bundle over am M-polyfold X.

8.1 Auxiliary norms

Recall the double filtration $Y_{m,k}$ of base regularity m and fibre regularity $k \leq m+1$, so an sc$^+$-section s of P gives in particular a continuous section of the topological Banach bundle $P : Y_{0,1} \to X_0$. The following definition introduces a suitable class of norms on this bundle to measure "smallness" of sc$^+$-sections.

Definition 8.1. An *auxiliary norm* on $P : Y \to X$ is a contin-

uous map $N : Y_{0,1} \to \mathbb{R}_{\geq 0}$ such that
(i) the restriction of N to each fibre is a complete norm;
(ii) if (y_k) is a sequence in $Y_{0,1}$ such that $P(y_k) \to x$ in X_0 and $N(y_k) \to 0$, then $y_k \to 0_x$ in $Y_{0,1}$.

Lemma 8.1. *Let $N : Y_{0,1} \to \mathbb{R}_{\geq 0}$ be a continuous map whose restriction to each fibre is a complete norm. Then the following are equivalent:*
(a) N is an auxiliary norm.
(b) For each $x \in X$ there exists an open neighbourhood $V \subset X$ of x, a strong bundle isomorphism

$$\begin{tikzcd} Y \supset P^{-1}(V) \arrow[r, "\Phi", "\cong"'] \arrow[d, "P"] & K \arrow[d] \\ X \supset V \arrow[r, "\phi", "\cong"'] & \mathcal{O} \end{tikzcd} \tag{8.1}$$

with $\phi(x) = 0$, and a constant $c > 0$ such that

$$\frac{1}{c} N(y) \leq |h|_1 \leq c N(y) \tag{8.2}$$

for all $y \in P^{-1}(V)$, where $\Phi(y) = (p, h)$.

Proof. Clearly (b) implies (a). For the converse implication, pick a strong bundle isomorphism as in (8.1) with $\phi(x) = 0$. and view N as a function on $K_{0,1} \subset E_0 \oplus F_1$. Then the first inequality in (8.2) follows from continuity of N at $(0, 0)$: there exists a neighbourhood $U \subset \mathcal{O}$ of 0 and a $\delta > 0$ such that $p \in U$ and $|h|_1 \leq \delta$ implies $N(p, h) \leq 1$, so for $p \in U$ and each $h \neq 0$ we have $\delta N(p, h) = |h|_1 N(p, \frac{\delta h}{|h|_1}) \leq |h|_1$. For the second inequality in (8.2) we argue by contradiction: if the inequality does not hold, then there exists a sequence $(p_k, h_k) \in K$ with $p_k \to 0$ and $|h_k|_1 = 1$ such that $N(p_k, h_k) \to 0$; since N is an auxiliary norm, this implies that $h_k \to 0$ in F_1, contradicting $|h_k|_1 = 1$ for all k. $\square$

Corollary 8.2. *On each strong bundle $P : Y \to X$ there exists an auxiliary norm N. Moreover, for any two auxiliary norms N_0, N_1 there exists a continuous positive function $\rho : X \to \mathbb{R}_{>0}$ such that for all $y \in P^{-1}(x) \cap Y_{0,1}$,*

$$\frac{1}{\rho(x)} N_0(y) \leq N_1(y) \leq \rho(x) N_0(y).$$

Proof. The norm $|\ |_1$ pulls back under the strong bundle isomorphism Φ in Lemma 8.1 to an auxiliary norm on $P^{-1}(V)$, and adding these local auxiliary norms with a continuous partition of unity yields an auxiliary norm on Y. Next consider two auxiliary norms N_0, N_1 on P. Lemma 8.1 provides for each $x \in X$ an open neighbourhood $U_x \subset X$ and a constant $c_x > 0$ such that $\frac{1}{c_x} N_0 \leq N_1 \leq c_x N_0$ on $P^{-1}(U_x)$. Hence there exist a locally finite overing $\{U_i\}_{i \in I}$ of X and constants $c_i > 0$ such that $\frac{1}{c_i} N_0 \leq N_1 \leq c_i N_0$ on $P^{-1}(U_i)$. Let $\{\rho_i\}_{i \in I}$ be a continuous partition of unity subordinate to this covering and set $\rho := \sum_{i \in I} c_i \rho_i : X \to \mathbb{R}_{>0}$. Then for each $y \in Y_{0,1}$ with $P(y) = x$ we obtain $N_1(y) = \sum_i \phi_i(x) N_1(y) \leq \sum_i \phi_i(x) c_i N_0(y) = f(x) N_0(y)$, and similarly $N_0(y) \leq f(x) N_1(y)$. □

8.2 Compact perturbations

Definition 8.2. Let f be an sc-smooth section of the strong bundle $P : Y \to X$. We say that f has *compact solution set* if its solution set

$$S_f := f^{-1}(0)$$

is compact in X_0. We say that f is *proper* if for each auxiliary norm N on P there exists an open neighbourhood U of $f^{-1}(0)$

such that the closure

$$\mathrm{cl}_{X_0}\{x \in U \mid N\big(f(x)\big) \leq 1\}$$

is compact in X_0.

Here we have extended the auxiliary norm to a function $N : Y \to [0, \infty]$ by setting $N := \infty$ on $Y \setminus Y_{0,1}$. We emphasize that this definition involves compactness only on the level X_0.

Proposition 8.3 (Properness of sc-Fredholm sections)**.** *An sc-Fredholm section f has compact solution set if and only if it is proper.*

Proof. (1) Since the solution set S_f is a closed subset of $\mathrm{cl}_{X_0}\{x \in U \mid N\big(f(x)\big) \leq 1\}$, properness of f implies compactness of S_f. Conversely, suppose that f has compact solution set and let N be an auxiliary norm. By compactness of S_f, it suffices to find for each $x_0 \in S_f$ an open neighbourhood U such that $\mathrm{cl}_{X_0}\{x \in U \mid N\big(f(x)\big) \leq 1\}$ is compact. Since f is sc-Fredholm, near x_0 we have a local trivialization with $\phi(x_0) = 0$ and a commuting diagram

$$\begin{tikzcd}
Y \supset P^{-1}(V) \arrow[r, "\Phi", "\cong"'] & K & \widetilde{U} \lhd \widetilde{F} \arrow[l, "R"'] \arrow[r, "\Psi", "\cong"'] & U \lhd F \\
X \supset V \arrow[u, "f"] \arrow[r, "\phi", "\cong"'] & \mathcal{O} \arrow[u, "\widetilde{f}"] & \widetilde{U} \arrow[l, "r"'] \arrow[u, "\widetilde{g}"] \arrow[r, "\psi", "\cong"'] & U, \arrow[u, "g=h+t"']
\end{tikzcd}$$

where $\widetilde{g}$ is a filling of $\widetilde{f}$, $h : (E, 0) \to (F, 0)$ is a basic germ, and $t : (E, 0) \to (F^1, 0)$ is an sc$^+$-germ. Since $\widetilde{g} = \widetilde{f}$ on $\mathcal{O}$, the set $\mathrm{cl}_{E_0}\{x \in \mathcal{O} \mid N\big(\widetilde{f}(x)\big) \leq 1\}$ is a closed subset of $\mathrm{cl}_{E_0}\{x \in \widetilde{U} \mid N\big(\widetilde{g}(x)\big) \leq 1\}$, so it suffices to show compactness of $\mathrm{cl}_{E_0}\{x \in U \mid N\big(g(x)\big) \leq 1\}$. We identify g with its principal part $g : (E, 0) \to F$. By Lemma 8.1 there exists a constant $c > 0$ such that $N(x, v) \le 1$ implies $|v|_1 \le c$ for all $(x, v) \in U_0 \oplus F_1$, so it suffices to show compactness of $\mathrm{cl}_{E_0}\{x \in U \mid |g(x)|_1 \le c\}$.

(2) By definition of a basic germ there exist sc-splittings $E = \mathbb{R}^n \oplus W$ and $F = \mathbb{R}^N \oplus W$ with projection $P : F \to W$ such that $P \circ h(a, w) = w - B(a, w)$ is an sc^0-contraction germ. By definition of a contraction germ and continuity of t and h there exists $\sigma > 0$ such that for $|a|_0, |w|_0, |w'|_0 \le \sigma$ we have

$$|B(a, w) - B(a, w')|_0 \le \frac{1}{4}|w - w'|,$$
$$|t(a, w)|_1 \le c, \qquad |(\mathbb{1} - P)h(a, w)|_1 \le c.$$

By continuity of B there exists $\tau > 0$ such that

$$|B(a, 0)|_0 \le \frac{\sigma}{8} \quad \text{for } |a|_0 \le \tau.$$

Denoting by B_r^V the closed ball in V of radius r around the origin, we consider the map

$$F : B_\tau^{\mathbb{R}^n} \times B_{\sigma/2}^W \times B_\sigma^W \to W, \qquad F(a, z, w) := B(a, w) + z.$$

It satisfies the estimates

$$\begin{aligned} |F(a, z, w)|_0 &\le |B(a, w) - B(a, 0)|_0 + |B(a, 0)|_0 + |z|_0 \\ &\le \frac{\sigma}{4} + \frac{\sigma}{8} + \frac{\sigma}{2} < \sigma, \\ |F(a, z, w) - F(a, z, w')|_0 &= |B(a, w) - B(a, w')|_0 \le \frac{1}{4}|w - w'|_0. \end{aligned}$$

Thus $F(a, z, \cdot)$ is a parametrized contraction of B_σ^W, uniform in $(a, z) \in B_\tau^{\mathbb{R}^n} \times B_{\sigma/2}^W$, and the parametrized Banach fixed point theorem (as in the proof of the implicit function theorem) yields a unique continuous map $\delta : B_\tau^{\mathbb{R}^n} \times B_{\sigma/2}^W \to B_\sigma^W$ such that

$$F(a, z, w) = w \iff w = \delta(a, z). \tag{8.3}$$

(3) We claim that the open neighbourhood

$$U := \{(a,w) \mid |a|_0 < \tau,\ |w|_0 < \sigma/4\} \subset E$$

of $(0,0)$ has the desired compactness property at the end of (1). As preparation, consider $(a,w) \in U$ and write

$$\begin{aligned} y := g(a,w) &= h(a,w) + t(a,w) \\ &= w - B(a,w) + (\mathbb{1} - P) \circ h(a,w) + t(a,w). \end{aligned}$$

Suppose that $y \in F_1$ and $|y|_1 \leq c$. Then $z := w - B(a,w) \in W_1$ and the estimates in (2) yield

$$\begin{aligned} |z|_1 &\leq |y|_1 + |(\mathbb{1} - P) \circ h(a,w)|_1 + |t(a,w)|_1 \\ &\leq c + c + c = 3c, \\ |z|_0 &\leq |w|_0 + |B(a,w) - B(a,0)|_0 + |B(a,0)|_0 \\ &\leq \frac{\sigma}{4} + \frac{\sigma}{16} + \frac{\sigma}{8} < \frac{\sigma}{2}. \end{aligned}$$

Hence $(a,z,w) \in B_\tau^{\mathbb{R}^n} \times B_{\sigma/2}^W \times B_\sigma^W$ and $w = \delta(a,z)$ for the continuous map δ from (2).

Consider now a sequence (a_k, w_k) in U with $|g(a_k,w_k)|_1 \leq c$. By the preceding discussion, $z_k := w_k - B(a_k,w_k) \in W_1$ and $|z_k|_1 \leq 3c$. By finite dimensionality of $\mathbb{R}^n$ and compactness of the embedding $W_1 \hookrightarrow W_0$, there exists a subsequence such that $a_k \to a$ in $\mathbb{R}^n$ and $z_k \to z$ in W_0. By the preceding discussion and continuity of δ we have $w_k = \delta(a_k, z_k) \to \delta(a,z) =: w$ in W_0, and therefore $(a_k, w_k) \to (a,w)$ in E_0. This concludes the proof of Proposition 8.3. $\square$

8.3 Transversality

In this section we consider

- a tame M-polyfold X which admits smooth sc$^\infty$-bump functions (e.g. one modeled over sc-Hilbert spaces);
- a strong bundle $P : Y \to X$;
- an auxiliary norm N on P.

Lemma 8.4 (Existence of sc$^+$-sections)**.** *For each $x \in X_\infty$, $e \in P^{-1}(x) \cap Y_\infty$, $\varepsilon > 0$, and open neighbourhood $U \subset X$ of x there exists an* sc$^+$*-section $s : X \to Y$ satisfying*

$$s(x) = e, \quad \mathrm{supp}\,(s) \subset U, \quad N\big(s(y)\big) < N(e) + \varepsilon \textit{ for all } y \in X.$$

Proof. After passing to a local trivialization, it suffices to consider a local strong bundle $K = R(U \triangleleft F) \to \mathcal{O} = r(U)$ with $e = (0, e') \in K_\infty$ and $x = 0 \in \mathcal{O}_\infty$. Since R is a strong bundle retraction, the map

$$t : \mathcal{O} \to K, \qquad y \mapsto R(y, e')$$

defines an sc$^+$-section with $t(0) = R(0, e') = (0, e')$. Since $N : K \to \mathbb{R}$ is continuous, there exists $\delta > 0$ such that $N\big(t(y)\big) < N\big(t(0)\big) + \varepsilon$ for all y in $\mathcal{O}' := \{y \in \mathcal{O} \mid |y|_0 < \delta\}$. Since X admits sc$^\infty$-bump functions, there exists an sc$^\infty$-function $\beta : \mathcal{O} \to [0, 1]$ with $\beta(0) = 1$ and $\mathrm{supp}\,(\beta) \subset \mathcal{O}'$. Then $s(y) := \beta(y)t(y)$ is an sc$^+$-section with the desired properties. $\square$

We denote by $\Gamma^+(P)$ the set of sc$^+$-sections $s : X \to Y$. Given an open subset $U \subset X$ we define the space of *allowable* sc$^+$*-sections*

$$\Gamma_U := \big\{s \in \Gamma^+(P) \mid \mathrm{supp}\,(s) \subset U,\ N\big(s(x)\big) \le 1 \text{ for all } x \in X\big\}.$$

We equip Γ_U with the (non-complete) metric

$$\rho(s, s') := \sup_{x \in X} N\big(s(x) - s'(x)\big).$$

Theorem 8.5 (Perturbation theorem)**.** *Let X be a tame M-polyfold X which admits smooth* sc^∞*-bump function, $P : Y \to X$ a strong bundle, and N an auxiliary norm on P. Let f be an sc-Fredholm section on P with compact solution set $S_f = f^{-1}(0)$. Then:*
(a) There exists an open neighbourhood $U \subset X$ of S_f such that for all $s \in \Gamma_U$ the solution set $S_{f+s} = \{x \in X \mid f(x)+s(x) = 0\}$ is compact in X.
(b) There exists a dense subset $\Gamma_U^{\mathrm{reg}} \subset \Gamma_U$ such that for each $s \in \Gamma_U^{\mathrm{reg}}$ the section $f + s$ is transverse to the zero section (i.e., the linearization $T_x^v(f + s) : T_xX \to Y_x$ is surjective for all $x \in S_{f+s}$). In particular, $S_{f+s} \subset X$ is a sub-M-polyfold whose induced M-polyfold structure is equivalent to the structure of a compact smooth manifold with corners.

Proof. As usually, we will treat only the case where X has no boundary and refer to [26] for the general case. Part (a) is an immediate consequence of Proposition 8.3 and the definition of properness.

For part (b) note first that $S_f \subset X_\infty$ because f is regularizing. For each $x \in S_f$ the linearization $T_x^v f$ is an sc-Fredholm operator, so its image in Y_x has finite codimension. By Lemma 8.4 we thus find sc^+-sections $s_1^x, \dots s_{k^x}^x \in \Gamma_U$ (whose number k^x depends on x) which together with the image of $T_x^v f$ span Y_x. Hence

$$F^x(\lambda, y) := f(y) + \sum_{j=1}^{k^x} \lambda_j s_j^x(y)$$

defines an sc-Fredholm section $F^x : \mathbb{R}^{k^x} \oplus X \to Y$ whose linearization at $(0, x)$ is surjective. By local regularity of sc-Fredholm germs (see Corollary 7.12 and the proof of Theorem 7.13) there exists an open neighbourhood $U^x \subset X$ of x such that $T^v_{(0,y)}F^x$ is surjective for all $y \in U^x \cap S_f$. By compactness of S_f, the open cover $\{U^x\}_{x\in S_f}$ of S_f has a finite subcover $\{U^{x_i}\}_{1\le i\le p}$. We number the sections $s_j^{x_i}$ for $1 \le i \le p$ and $1 \le j \le k^{x_i}$ as $t_1, \dots, t_m$. Then

$$F(\lambda, x) := f(x) + \sum_{j=1}^{m} \lambda_j t_j(x)$$

defines an sc-Fredholm section $F : \mathbb{R}^m \oplus X \to Y$ whose linearization at $(0, x)$ is surjective for all $x \in S_f$.

Now let $\varepsilon > 0$ be given. By the Implicit Function Theorem 7.13 there exists an open neighbourhood

$$V \subset \{\lambda \in \mathbb{R}^m \mid \sum_{j=1}^{m} |\lambda_j| < \varepsilon\} \oplus X$$

of $\{0\} \times S_f$ such that $T^v_{\lambda,x}F$ is surjective for all $(\lambda, x) \in V$ with $F(\lambda, x) = 0$. In particular, $S_F := \{(\lambda, x) \in V \mid F(\lambda, x) = 0\}$ is a smooth finite dimensional manifold. By Sard's theorem, the smooth map $S_F \to \mathbb{R}^m$, $(\lambda, x) \mapsto \lambda$ has a regular value λ^*. It follows that $s(x) := \sum_{j=1}^{m} \lambda^*_j t_j(x)$ defines an sc$^+$-section with support in U and $N\big(s(x)\big) \le \sum_{j=1}^{m} |\lambda_j| N\big(t_j(x)\big) \le \varepsilon$ for all $x \in X$ such that $f + s$ is transverse to the zero section. This shows that the set Γ_U^{reg} of $s \in \Gamma_U$ such that $f + s$ is transverse to the zero section contains elements arbitrarily close to 0.

To show that Γ_U^{reg} is dense in Γ_U, let $s_0 \in \Gamma_U$ and $\varepsilon > 0$ be given. Then $s_1 := (1-\varepsilon)s_0 \in \Gamma_U$ with $N(s_1) \le 1 - \varepsilon$. Applying the preceding argument to the sc-Fredholm section $f + s_1$, we find an $s_2 \in \Gamma_U$ such that $N(s_2) \le \varepsilon$ and $f + s_1 + s_2$ is transverse to the zero section. Since $N(s_1 + s_2) \le (1-\varepsilon) + \varepsilon = 1$, this means that $s := s_1 + s_2 \in \Gamma_U^{\mathrm{reg}}$. Since $N(s - s_0) = N(s_2 - \varepsilon s_0) \le 2\varepsilon$, this shows that $s_0 \in \Gamma_U$ can be arbitrarily well approximated by elements $s \in \Gamma_U^{\mathrm{reg}}$ and the theorem is proved. □

8.4 Morse homology

In this section we explain how the perturbation theory on M-polyfolds can be used to define Morse homology. We consider the following setup.

- M is an n-dimensional manifold (without boundary but not necessarily closed);
- $H : M \to \mathbb{R}$ is a Morse function which is proper and bounded from below;
- g is a Riemannian metric on M.

Unbroken paths. For a critical point p of H let $\delta(p)$ be the minimal distance of an eigenvalue of its Hessian $\mathrm{Hess}_p H$ to 0. For each p we pick a sequence of positive weights $0 < \delta_0(p) < \delta_1(p) < \cdots < \delta(p)$. For critical points p^-, p^+ we pick charts $\psi^\pm : M \supset U^\pm \to \mathbb{R}^n$ with $\psi^\pm(p^\pm) = 0$. For $m \in \mathbb{N}_0$ we define

$$\widetilde{\mathcal{E}}_m(p^-, p^+)$$

to be the space of H^{m+2}_{loc}-maps $u : \mathbb{R} \to M$ for which there exists s_0 such that $u(s) \in U^\pm$ whenever $\pm s \geq s_0$ and

$$\psi^+ \circ u|_{[s_0,\infty)} \in H^{m+2,\delta_m(p^+)}([s_0,\infty), \mathbb{R}^n),$$
$$\psi^- \circ u|_{(-\infty,-s_0]} \in H^{m+2,\delta_m(p^-)}((-\infty,-s_0], \mathbb{R}^n).$$

Problem 8.1. For critical points p^-, p^+ of H show:
(a) The definition of $\widetilde{\mathcal{E}}_m(p^-, p^+)$ does not depend on the charts $\psi^\pm$ and yields a smooth sc-manifold

$$\widetilde{\mathcal{E}}(p^-, p^+) := \widetilde{\mathcal{E}}_0(p^-, p^+) \supset \widetilde{\mathcal{E}}_1(p^-, p^+) \supset \cdots$$

(b) There exists a canonical strong sc-bundle $\widetilde{\mathcal{F}}(p^-,p^+) \to \widetilde{\mathcal{E}}(p^-,p^+)$ whose fibre over u on level m is the space $H^{m+1,\delta_m(p^\pm)}(\mathbb{R}, u^*TM)$ with weights $\delta_m(p^\pm)$ near $\pm\infty$, defined with respect to trivializations of u^*TM induced by charts $\psi^\pm$ near $p^\pm$.
(c) The assignment

$$\widetilde{\partial}_H(u) := \partial_s u + \nabla H(u)$$

defines an sc-Fredholm section in the bundle $\widetilde{\mathcal{F}}(p^-,p^+) \to \widetilde{\mathcal{E}}(p^-,p^+)$.
(d) The $\mathbb{R}$-translation $u \mapsto u(\cdot + t)$ defines smooth $\mathbb{R}$-actions on $\widetilde{\mathcal{E}}(p^-,p^+)$ and $\widetilde{\mathcal{F}}(p^-,p^+)$ for which $\widetilde{\partial}_H$ is equivariant. Moreover, for $p^- \neq p^+$ this action is free and the quotients

$$\mathcal{F}(p^-,p^+) := \widetilde{\mathcal{F}}(p^-,p^+)/\mathbb{R} \to \mathcal{E}(p^-,p^+) := \widetilde{\mathcal{E}}(p^-,p^+)/\mathbb{R}$$

define a smooth sc-bundle to which $\widetilde{\partial}_H$ descends as an sc-Fredholm section ∂_H. *Hint: Construct local slices for the* $\mathbb{R}$*-action as in the proof of Theorem 8.7.*

The Fredholm index of ∂_H is computed in the following two exercises.

Problem 8.2. For $u \in \widetilde{\mathcal{E}}(p^-,p^+)$ with $\widetilde{\partial}_H(u) = 0$ show:
(a) The map $u : \mathbb{R} \to M$ is smooth and converges exponentially to $p^\pm$ as $s \to \pm\infty$ with rates $\delta^\pm = \delta(p^\pm)$.
(b) In a trivialization of u^*TM induced by charts $\phi^\pm$ near $p^\pm$, the linearization of $\widetilde{\partial}_H$ at u is given on level m by a linear operator

$$L_A = \partial_s - A(s) : H^{m+2,w^\pm}(\mathbb{R},\mathbb{R}^n) \to H^{m+1,w^\pm}(\mathbb{R},\mathbb{R}^n)$$

with weights $w^\pm = \delta_m(p^\pm)$, where $A : \mathbb{R} \to \mathbb{R}^{n\times n}$ is a smooth path of matrices converging to the asymptotic operators

$$A^\pm = -\mathrm{Hess}_{p^\pm} H \tag{8.4}$$

(note the minus sign) in the sense that $e^{\pm\delta^\pm s}\big(A(s) - A^\pm\big)$ remains bounded as $s \to \pm\infty$. Let us call such a path A *admissible.*

Problem 8.3. Let L_A be a linear operator as in Problem 8.2 with nondegenerate symmetric asymptotic operators $A^\pm$. Show:
(a) For all $m \in \mathbb{N}_0$ and weights $w^\pm$ satisfying $|w^\pm| < \delta^\pm$ the operator $L_A : H^{m+1,w^\pm}(\mathbb{R}, \mathbb{R}^n) \to H^{m,w^\pm}(\mathbb{R}, \mathbb{R}^n)$ is Fredholm with index not depending on $m, w^\pm$. *Hint: See [40, Theorem 2.1] and [34, Theorem C.2.3].*
(b) The Fredholm index of L_A is invariant under homotopies of admissible paths $A : \mathbb{R} \to \mathbb{R}^{n\times n}$, where the asymptotic operators are allowed to vary in the space of nondegenerate symmetric matrices.
(c) For two admissible paths $A_i : \mathbb{R} \to \mathbb{R}^{n_i\times n_i}$ the path $A_1 \oplus A_2 : \mathbb{R} \to \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)}$ is admissible and

$$\mathrm{ind}(L_{A_1\oplus A_2}) = \mathrm{ind}(L_{A_1}) + \mathrm{ind}(L_{A_2}).$$

(d) The Fredholm index of L_A equals the difference of the Morse indices (= number of negative eigenvalues) of the asymptotic operators,

$$\mathrm{ind}(L_A) = \mathrm{Morse}(A^+) - \mathrm{Morse}(A^-).$$

Hint: Use (b) and (c) to reduce this to the case $n = 1$.

In view of (8.4), the preceding problem implies

Corollary 8.6. *The index of the sc-Fredholm section $\widetilde{\partial}_H : \widetilde{\mathcal{E}}(p^-, p^+) \to \widetilde{\mathcal{F}}(p^-, p^+)$ equals*

$$\mathrm{ind}(\widetilde{\partial}_H) = \mathrm{Morse}(p^-) - \mathrm{Morse}(p^+).$$

Thus the index of the induced sc-Fredholm section $\partial_H : \mathcal{E}(p^-, p^+) \to \mathcal{F}(p^-, p^+)$ on the quotient equals

$$\mathrm{ind}(\partial_H) = \mathrm{Morse}(p^-) - \mathrm{Morse}(p^+) - 1.$$

Broken paths. For critical points $p_0, \dots, p_k$, $k \geq 1$ with

$$H(p_0) > H(p_1) > \cdots > H(p_k)$$

we define the space of broken paths from p_0 to p_k as

$$\mathcal{E}(p_0, p_1, \dots, p_k) := \mathcal{E}(p_0, p_1) \times \cdots \times \mathcal{E}(p_{k-1}, p_k).$$

This is an sc-manifold over which we have the sc-bundle

$$\mathcal{F}(p_0, p_1, \dots, p_k) := \mathcal{F}(p_0, p_1) \times \cdots \times \mathcal{F}(p_{k-1}, p_k)$$

with the sc-Fredholm section

$$\bar{\partial}_H \times \cdots \times \bar{\partial}_H : \mathcal{E}(p_0, p_1, \dots, p_k) \to \mathcal{F}(p_0, p_1, \dots, p_k). \tag{8.5}$$

For critical points p^-, p^+ with $H(p^-) > H(p^+)$ we define the set

$$X(p^-, p^+) := \mathcal{E}(p^-, p^+) \amalg \coprod_{k \geq 1} \coprod_{p_1, \dots, p_k} \mathcal{E}(p^-, p_1, \dots, p_k, p^+),$$

where the disjoint union is taken over all critical points $p_1, \dots, p_k$ with $H(p^-) > H(p_1) > \cdots > H(p_k) > H(p^+)$, and similarly

$$Y(p^-, p^+) := \mathcal{F}(p^-, p^+) \amalg \coprod_{k \geq 1} \coprod_{p_1, \dots, p_k} \mathcal{F}(p^-, p_1, \dots, p_k, p^+).$$

Theorem 8.7 (M-polyfold Fredholm setup for Morse homology)**.**
(a) The set $X(p^-, p^+)$ carries the natural structure of a tame M-polyfold whose interior is $X^{(0)}(p^-, p^+) = \mathcal{E}(p^-, p^+)$ and whose set of points of degeneracy index $k \geq 1$ is

$$X^{(k)}(p^-, p^+) = \coprod_{p_1, \dots, p_k} \mathcal{E}(p^-, p_1, \dots, p_k, p^+),$$

and such that the induced M-polyfold structure on $X^{(k)}(p^-, p^+)$ agrees with the sc-manifold structure defined above.
(b) The set $Y(p^-, p^+)$ carries the natural structure of a strong bundle over $X(p^-, p^+)$ which agrees on each stratum with the strong sc-bundle $Y^{(k)}(p^-, p^+) \to X^{(k)}(p^-, p^+)$ defined above.

(c) The products (8.5) *of* ∂_H *on the strata* $X^{(k)}(p^-, p^+)$ *induce an sc-Fredholm section* $f : X(p^-, p^+) \to Y(p^-, p^+)$ *of index*

$$\mathrm{ind}(f) = \mathrm{Morse}(p^-) - \mathrm{Morse}(p^+) - 1$$

and with compact solution set $S_f = f^{-1}(0)$*.*

Sketch of proof. (a) We will treat the case of once broken paths; the case of multiply broken paths differs only in notation. Consider critical points $p^\pm, p$ with $H(p^-) > H(p) > H(p^+)$ and paths $u^- \in \widetilde{\mathcal{E}}(p^-, p)$, $u^+ \in \widetilde{\mathcal{E}}(p, p^+)$. After reparametrization we may assume that $u^\pm$ have nonvanishing derivative at $s = 0$. Pick embedded $(n-1)$-balls $\Sigma^\pm$ transverse to $u^\pm$ at $u^\pm(0)$. Pick also an open embedding $\psi : N \hookrightarrow \mathbb{R}^n$ of a neighbourhood N of the union of the images of u^- and u^- such that $\psi(N)$ is starshaped with respect to $\psi(p) = 0$. Since $\widetilde{\mathcal{E}}(p^-, p^+) \subset C^1(\mathbb{R}, W)$ by the Sobolev embedding theorem, the condition that (smooth or once broken) paths are contained in N and transverse to Σ^- and Σ^+ defines a subset

$$\mathcal{U} \subset X(p^-, p^+)$$

containing the broken path $[u^-, u^+]$, whose intersections with the strata $\mathcal{E}(p^-, p^+)$ and $\mathcal{E}(p^-, p) \times \mathcal{E}(p, p^+)$ are open. In particular, the two factors of $\mathcal{U} \cap \big(\mathcal{E}(p^-, p) \times \mathcal{E}(p, p^+)\big)$ define open neighbourhoods of $[u^\pm]$ in $\mathcal{E}(p^-, p)$ resp. $\mathcal{E}(p, p^+)$, which we can represent by open subsets of the local slices to the $\mathbb{R}$-actions given by

$$\mathcal{V}^\pm \subset \{v^\pm \in \widetilde{\mathcal{E}}(p^-, p) \text{ resp. } \widetilde{\mathcal{E}}(p, p^+) \mid v^\pm(0) \in \Sigma^\pm\}.$$

Via the embedding $\psi : N \hookrightarrow \mathbb{R}^n$ we can view elements in $\mathcal{V}^\pm$ as maps $v^\pm : \mathbb{R} \to \mathbb{R}^n$ and write them uniquely as

$$v^\pm = u^\pm + \xi^\pm, \qquad \xi^\pm \in H^{2,\delta_0}(\mathbb{R}, \mathbb{R}^n).$$

Hence we are in the situation of Section 7.6 with the sc-Banach space

$$E := H^{2,\delta_0}(\mathbb{R}, \mathbb{R}^n) \oplus H^{2,\delta_0}(\mathbb{R}, \mathbb{R}^n)$$

(except that the regularity level has been raised by 1). Let

$$r : [0,\varepsilon) \oplus E \to [0,\varepsilon) \oplus E, \qquad (a, u_1, u_2) \mapsto \big(a, \pi_a(u_1, u_2)\big)$$

be the sc-smooth retraction from Proposition 7.17 onto its image

$$\mathcal{O} = \{0\} \times E \cup \bigcup_{a\in(0,1/2)} \{a\} \times \ker \ominus_a.$$

Recall that $\pi_a : E \to E$ is the projection onto $\ker \ominus_a$ along $\ker \oplus_a$, where the gluing and antigluing maps are defined by

$$\begin{aligned}
\oplus_a(u_1, u_2)(s) &:= \beta(s)u_1(s + R) + \big((1 - \beta(s)\big)u_2(s - R),\\
\ominus_a(u_1, u_2)(s) &:= \big(\beta(s) - 1\big)u_1(s + R) + \beta(s)u_2(s - R).
\end{aligned}$$

with $R = \varphi(a) = e^{1/a} - e$. This shows that for $a > 0$ and $v^\pm = u^\pm + \xi^\pm \in \mathcal{V}^\pm$ the glued map $v := \oplus_a(v^-, v^+) = \oplus_a(u^-, u^+) + \oplus_a(\xi^-, \xi^+)$ satisfies $v(\pm R) \in \Sigma^\pm$. Let us write

$$\begin{aligned}
\mathcal{V} &:= \mathcal{V}^- \oplus \mathcal{V}^+ = (u^-, u^+) + \mathcal{W},\\
\mathcal{W} &:= \{(\xi^-, \xi^+) \in E \mid u^\pm + \xi^\pm \in \mathcal{V}^\pm\}.
\end{aligned}$$

Since the gluing map defines an sc-isomorphism $\oplus_a : \ker \ominus_a \xrightarrow{\cong} H^{2,\delta_0}(\mathbb{R}, \mathbb{R}^n)$, by the preceding discussion it induces an sc-diffeomorphism

$$\mathcal{W} \cap \ker \ominus_a \xrightarrow{\cong} \mathcal{V}_a, \qquad \xi \mapsto \oplus_a(u + \xi)$$

onto an open neighbourhood

$$\mathcal{V}_a \subset \{v \in \widetilde{\mathcal{E}}(p^-, p^+) \mid v(\pm R) \in \Sigma^\pm\}$$

of $u_a := \oplus_a(u^-, u^+)$. Conversely, for every $v \in \widetilde{\mathcal{E}}(p^-, p^+)$ representing a class $[v] \in \mathcal{U}$ there exist unique times $s^\pm$ such that

$v(s^\pm) \in \Sigma^\pm$. Setting $R := (s^+ - s^-)/2$ we can uniquely reparametrize v such that $v(\pm R) \in \Sigma^\pm$, so that $v \in \mathcal{V}_a$ with $a = \varphi^{-1}(R)$. This shows that the map given by $(\xi^-, \xi^+) \mapsto (u^- + \xi^-, u^+ + \xi^+)$ for $a = 0$ and by $(\xi^-, \xi^+) \mapsto \oplus_a(u^- + \xi^-, u^+ + \xi^+)$ for $a > 0$ defines a bijection

$$\begin{aligned}\phi^{-1} : \big([0,1/2) \oplus \mathcal{W}\big) \cap \mathcal{O} &= \{0\} \times \mathcal{W} \cup \bigcup_{a \in (0,1/2)} \{a\} \times (\mathcal{W} \cap \ker \ominus_a) \\ &\xrightarrow{\cong} \{0\} \times \mathcal{V} \cup \bigcup_{a \in (0,1/2)} \{a\} \times \mathcal{V}_a = \mathcal{U}.\end{aligned}$$

Its inverse ϕ is the desired M-polyfold chart for $X(p^-, p^+)$ near $[u^-, u^+]$ and part (a) is proved. Part (b) is proved analogously.

In part (c) the Fredholm property follows from Theorem 7.18 by passing to the quotient, and the Fredholm index follows from Corollary 8.6. It remains to show compactness of the solution set. For this, consider a sequence $(u_k) \subset X(p^-, p^+)$ with $f(u_k) = 0$. After passing to a subsequence and by treating the pieces of a broken trajectory individually, we may assume without loss of generality that $(u_k) \subset \mathcal{E}(p^-, p^+)$. Since H is proper, a standard argument in Morse homology yields a subsequence which converges "geometrically" to a broken trajectory $u = (u^{(1)}, \dots, u^{(\ell)}) \in \mathcal{E}(p^-, p_1, \dots, p_{\ell-1}, p^+)$ in the following sense: there exist sequences of times $t_k^{(i)}$ such that $u_k(\cdot - t_k^{(i)}) \to u^{(i)}$ in C^∞_{loc} for each $i = 1, \dots, \ell$. We will prove that $u_k \to u$ in $X(p^-, p^+)$.

To see this, we restrict ourselves again to the case $\ell = 2$ so that $u = [u^-, u^+] \in \mathcal{E}(p^-, p) \times \mathcal{E}(p, p^+)$. The preceding geometric convergence implies that $u_k \in \mathcal{U}$ for large k, where $\mathcal{U} \subset X(p^-, p^+)$ is the neighbourhood of $[u^-, u^+]$ defined in the proof of part (a). So we can apply the coordinate chart ϕ from part (a) and obtain using (7.3)

$$\phi(u_k) = \boxdot_{a_k}^{-1}\begin{pmatrix} u_k \\ 0 \end{pmatrix} = \begin{pmatrix} \tau_{-R_k} & 0 \\ 0 & \tau_{R_k} \end{pmatrix} \frac{1}{\gamma} \begin{pmatrix} \beta & \beta - 1 \\ 1-\beta & \beta \end{pmatrix} \begin{pmatrix} u_k \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} \tau_{-R_k}\left(\frac{\beta}{\gamma} u_k\right) \\ \tau_{R_k}\left(\frac{1-\beta}{\gamma} u_k\right) \end{pmatrix} =: \begin{pmatrix} u_k^- \\ u_k^+ \end{pmatrix} \in E.$$

The preceding geometric convergence yields $u_k^\pm \to u^\pm$ in C^∞_{loc}, and we need to show $u_k^\pm \to u^\pm$ in $H^{2,\delta_0}(\mathbb{R}, \mathbb{R}^n)$. Note that in general C^∞_{loc}-convergence does not imply convergence in H^{2,δ_0}, since the former topology is local while the latter is global. What saves us is the fact that the trajectory u_k stays near the critical point p for very long times as k becomes large. For simplicity, let us assume that the metric is standard and

$$H(x,y) = \frac{1}{2}(|x|^2 - |y|^2)$$

on an open ball B around $p = 0$ in $\mathbb{R}^n = \mathbb{R}^{n_1} \oplus \mathbb{R}^{n_2}$. Then the negative gradient flow equation for $v = (x, y) : (a, b) \to B \subset \mathbb{R}^n$ is

$$\partial_s x = -x, \quad \partial_s y = y \tag{8.6}$$

and has the solutions

$$x(s) = x^0 e^{-s}, \quad y(s) = y^0 e^s.$$

We choose the transverse $(n-1)$-balls $\Sigma^\pm$ to be contained in B, so that $u_k(\pm R_k) \in \Sigma_k$ implies $u_k([-R_k, R_k]) \subset B$. Let us write

$$u_k^- = \beta_k v_k \quad \text{with} \quad \beta_k := \tau_{-R_k}\left(\frac{\beta}{\gamma}\right), \quad v_k := \tau_{-R_k} u_k = (x_k, y_k).$$

Then the components of v_k have the form

$$x_k(s) = x_k^0 e^{-s}, \quad y_k(s) = y_k^0 e^s \qquad \text{for } s \in [0, 2R_k],$$

and $v_k([0,2R_k]) \in B$ yields a constant c such that

$$|y_k^0| \le ce^{-2R_k}. \tag{8.7}$$

The trajectory u^- satisfies $u^-(0) \in \Sigma^-$ and thus has the form $u^-(s) = (x(s),0) = (x^0 e^{-s},0)$ for $s \in [0,\infty)$. The geometric converge above implies that the initial values converge,

$$x_k^0 \to x_0, \qquad y_k^0 \to 0.$$

We claim that u_k^- converges to u^- in $H^{m,w}([0,\infty),\mathbb{R}^n)$ for each $m \in \mathbb{N}_0$ and weight $w \in [0,1)$. To see this, we use $|\beta_k(s)| \le 2$ to estimate the weighted L^2-norm of the first components:

$$\begin{aligned}
&\int_0^\infty e^{2ws}|\beta_k(s)x_k(s)-x(s)|^2 ds = \int_0^\infty e^{2(w-1)s}|\beta_k(s)x_k^0 - x^0|^2 ds \\
&\le \int_0^{R_k-1} e^{2(w-1)s}|x_k^0-x^0|^2 ds + \int_{R_k-1}^{R_k+1} e^{2(w-1)s}(2|x_k^0|+|x^0|)^2 ds \\
&\le \frac{1}{2(1-w)}|x_k^0-x^0| + 2(2|x_k^0|+|x^0|)^2 e^{2(w-1)(R_k-1)}.
\end{aligned}$$

The first term in the last line converges to zero because $x_k^0 \to x^0$, and the second term converges to zero because $w<1$ and $R_k \to \infty$. The weighted second components can be estimated using $s \le R_k+1$ and (8.7) by

$$\begin{aligned}
e^{ws}|\beta_k(s)y_k(s)-0| &= e^{(w+1)s}|\beta_k(s)y_k^0| \\
&\le 2ce^{(w+1)(R_k+1)-2R_k} = c_1 e^{(w-1)R_k}
\end{aligned}$$

with $c_1 = 2ce^{w+1}$. So its L^2-norm

$$\int_0^\infty e^{2ws}|\beta_k(s)y_k(s)-0|^2 ds \le c_1^2(R_k+1)e^{2(w-1)R_k}$$

converges to zero because $w<1$ and $R_k \to \infty$. This proves the claim for $m=0$, and in view of the ODE (8.6) it follows for each

$m \geq 0$.

Since u^- and the u_k^- converge exponentially to p^- as $s \to -\infty$, convergence of the initial conditions implies $H^{m,w}$-convergence on $(-\infty, 0]$ for each $w < \delta(p^-)$. This shows that $u_k^- \to u^-$ in $H^{2,\delta_0}(\mathbb{R}, \mathbb{R}^n)$, and an analogous argument for u_k^+ concludes the proof of Theorem 8.7. $\square$

Abstract perturbations. By construction, the sc-Fredholm section f makes the following diagram commute for each $k \geq 1$:

$$\begin{tikzcd}
Y^{(k)}(p^-, p^+) \arrow[r, equal] & \coprod_{p_1,\ldots,p_k} \mathcal{F}(p^-, p_1, \ldots, p_k, p^+) \\
X^{(k)}(p^-, p^+) \arrow[r, equal] \arrow[u, "f|_{X^{(k)}(p^-,p^+)}"] & \coprod_{p_1,\ldots,p_k} \mathcal{E}(p^-, p_1, \ldots, p_k, p^+). \arrow[u, "f\times\cdots\times f"']
\end{tikzcd}$$

Note that, by induction, this diagram commutes for all $k \geq 1$ if and only if it commutes for $k = 1$:

$$\begin{tikzcd}
Y^{(1)}(p^-, p^+) \arrow[r, equal] & \coprod_p \mathcal{F}(p^-, p) \times \mathcal{F}(p, p^+) \\
X^{(1)}(p^-, p^+) \arrow[r, equal] \arrow[u, "f|_{X^{(1)}(p^-,p^+)}"] & \coprod_p \mathcal{E}(p^-, p) \times \mathcal{E}(p, p^+). \arrow[u, "f\times f"']
\end{tikzcd}$$

Denoting by $\partial X(p^-, p^+) \subset X(p^-, p^+)$ the set of points of degeneracy index ≥ 1, this is equivalent to commutativity of the diagram

$$\begin{tikzcd}
\partial Y(p^-, p^+) \arrow[r, equal] & \coprod_p Y(p^-, p) \times Y(p, p^+) \\
\partial X(p^-, p^+) \arrow[r, equal] \arrow[u, "f|_{\partial X(p^-,p^+)}"] & \coprod_p X(p^-, p) \times X(p, p^+). \arrow[u, "f\times f"']
\end{tikzcd} \tag{8.8}$$

Theorem 8.8 (Abstract perturbations for Morse homology)**.** *In the setup of Theorem 8.7, there exists an* sc$^+$*-section s of the strong bundle*

$$Y := \coprod_{H(p^-)>H(p^+)} Y(p^-, p^+) \to X := \coprod_{H(p^-)>H(p^+)} X(p^-, p^+)$$

such that the sc-Fredholm section $f+s : X \to Y$ has the following properties:
(i) For all p^-, p^+ the solution set $\{u \in X(p^-,p^+) \mid f(u)+s(u) = 0\}$ is compact.
(ii) $f+s$ is transverse to the zero section.
(iii) For all p^-, p^+ the diagram (8.8) *commutes with f replaced by $f+s$.*

Sketch of proof. Theorem 8.5 provided a section s satisfying properties (i) and (ii), but not (iii). To ensure property (iii) we construct s inductively as follows. Since the Morse function H is proper and bounded from below, its critical values can be ordered as $c_0 < c_1 < c_2 < \cdots$, where the sequence c_k is either finite or converges to ∞ as $k \to \infty$. For each $k \geq 1$ define

$$X[k] := \coprod_{i-j=k} \coprod_{H(p^-)=c_i,\, H(p^+)=c_j} X(p^-,p^+) \subset X.$$

Denote by $\partial X \subset X$ the set of points of degeneracy index ≥ 1 and set $\partial X[k] := X[k] \cap \partial X$. Then $\partial X[1] = \varnothing$ and we use Theorem 8.5 to find a perturbation $s[1] : X[1] \to Y$ of $f|_{X[1]}$ satisfying properties (i) and (ii) on $X[1]$. The product $s[1] \times s[1]$ then defines a section on the boundary $\partial X[2] = X[1] \times X[1]$. To proceed we need two more results:

- an *extension theorem* [26, Theorem 5.9] providing an extension of $s[1] \times s[1]$ from $\partial X[2]$ to an sc$^+$-section $t[2] : X[2] \to Y$ such that the solution set of $f|_{X[2]} + t[2]$ intersects each $X(p^-,p^+)$ is a compact set;

- and a *refined perturbation theorem* [26, Theorem 5.7] providing a perturbation $s[2]$ of $t[2]$ such that $s[2] = t[2]$ on $\partial X[2]$ and $f|_{X[2]} + s[2]$ satisfies properties (i) and (ii) on $X[2]$.

By construction, the diagram (8.8) commutes with f replaced by $f|_{X[2]} + s[2]$ over the set $X[2]$. Now the products $s[1] \times s[2]$ and $s[2] \times s[1]$ define a section on the boundary $\partial X[2] = X[1] \times X[2] \amalg X[2] \times X[1]$, and we proceed as before to extend it to $s[3]$ on $X[3]$ such that properties (i)–(iii) hold over $X[3]$. Continuing inductively we find the desired section s. $\square$

Theorem 8.8 together with the Implicit Function Theorem 7.13 implies

Corollary 8.9. *For the* sc^+*-section s from Theorem 8.8, each solution set*

$$\mathcal{M}(p^-, p^+) := \{u \in X(p^-, p^+) \mid f(u) + s(u) = 0\}$$

is a smooth compact manifold with corners of dimension

$$\dim \mathcal{M}(p^-, p^+) = \mathrm{Morse}(p^-) - \mathrm{Morse}(p^+) - 1.$$

Moreover, the diagram (8.8) *induces for the boundaries $\partial\mathcal{M}(p^-, p^+) = \mathcal{M}(p^-, p^+) \cap \partial X$ canonical diffeomorphisms*

$$\partial\mathcal{M}(p^-, p^+) \cong \coprod_p \mathcal{M}(p^-, p) \times \mathcal{M}(p, p^+). \tag{8.9}$$

Definition of Morse homology. Now we use Corollary 8.9 to define Morse homology with coefficients in $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$. For an exhausting Morse function $H : M \to \mathbb{R}$ let $MC_k(H)$ be the $\mathbb{Z}_2$-vector space with basis the critical points of index k. For a metric g and an abstract perturbation s as in Theorem 8.8 we define a linear map $\partial_k : MC_k(H) \to MC_{k-1}(H)$ on basis elements by

$$\partial_k p^- := \sum_{\mathrm{Morse}(p^+)=k-1} \langle \partial_k p^-, p^+ \rangle p^+, \qquad \langle \partial_k p^-, p^+ \rangle := \#\mathcal{M}(p^-, p^+).$$

Here $\#\mathcal{M}(p^-, p^+)$ denotes the number of points in $\mathcal{M}(p^-, p^+)$ modulo two, which is well-defined because $\mathcal{M}(p^-, p^+)$ is a compact manifold of dimension $\mathrm{Morse}(p^-) - \mathrm{Morse}(p^+) - 1 = 0$. We claim that

$$\partial_{k-1} \circ \partial_k = 0.$$

To see this, for $\mathrm{Morse}(p^-) = k$ and $\mathrm{Morse}(p^+) = k-2$ we use (8.9) to write

$$\begin{aligned}
\langle \partial_{k-1} \circ \partial_k p^-, p^+ \rangle &= \sum_{\mathrm{Morse}(p)=k-1} \langle \partial_k p^-, p \rangle \langle \partial_{k-1} p, p^+ \rangle \\
&= \sum_{\mathrm{Morse}(p)=k-1} \#\mathcal{M}(p^-, p) \#\mathcal{M}(p, p^+) \\
&= \#\partial\mathcal{M}(p^-, p^+).
\end{aligned}$$

Since $\mathcal{M}(p^-, p^+)$ is a compact 1-dimensional manifold with boundary, the number of its boundary points is even and the claim follows. Hence we can define the *Morse homology groups*

$$MH_k(H, g, s) := \ker \partial_k / \mathrm{im}\, \partial_{k+1}.$$

Properties of Morse homology. The following problems are devoted to the basic properties of Morse homology. Some of them are difficult and you may wish to consider [34] or [5].

Problem 8.4. (a) Prove that $MH_k(H, g, s)$ is independent of the metric g and the abstract perturbation s; we will denote it by $MH_k(H)$.
(b) For a regular value b of H let $MC_k^b(H) \subset MC_k(H)$ be the subspace generated by the critical points p with $H(p) < b$. Show: this defines a subcomplex whose homology we denote by $MH_k^b(H)$; for $a < b$ the obvious inclusion induces a homomorphism $MH_k^a(H) \to MH_k^b(H)$; and $MH_k(H) = \lim_{b \to \infty} MH_k^b(H)$.
(c) Prove: $MH_k(H)$ is independent of the exhausting Morse function H. *Hint: Use part (b) to reduce this to the case that M is a compact manifold with boundary, and $H : M \to \mathbb{R}$ is a*

Morse function without critical points on ∂M and attaining its maximal value on ∂M.

Problem 8.5. Prove: if the pair (H, g) is *Morse–Smale* (i.e., all stable and unstable manifolds of critical points intersect transversely), then the sc-Fredholm section f in Theorem 8.7 is transverse to the zero section, so Corollary 8.9 holds with $s = 0$.

In view of Problem 8.4(c), $MH_k(H)$ does not depend on the exhausting Morse function $H : M \to \mathbb{R}$ and is therefore an invariant of the manifold M. The following result identifies this invariant as the singular homology with $\mathbb{Z}_2$-coefficients.

Theorem 8.10. *For each exhausting Morse function $H : M \to \mathbb{R}$ we have*

$$MH_k(H) \cong H_k(M; \mathbb{Z}_2).$$

There are various proofs of this theorem in the literature, see [34, 5]. One approach would be using Problem 8.4(b) to reduce it to the case of a compact manifold with boundary, and choosing a Morse function H whose critical points are the barycenters of the faces of a smooth triangulation of M.

Problem 8.6. Use appropriate Morse functions to compute $H_*(M; \mathbb{Z}_2)$ for the following manifolds:
(a) the sphere S^n;
(b) the torus T^n;
(c) a closed orientable surface Σ_g of genus g;
(d) complex projective space $\mathbb{C}P^n$;
(e) real projective space $\mathbb{R}P^n$.

Problem 8.7. The *Morse cohomology* $MH^k(H)$ is defined as the homology of the dual chain complex $\mathrm{Hom}_{\mathbb{Z}_2}\big(CM_k(H), \mathbb{Z}_2\big)$. Show:
(a) $MH^k(H) \cong H^k(M; \mathbb{Z}_2)$.
(b) For M a closed n-manifold there exists a natural isomorphism

$MH^k(H) \cong MH_{n-k}(-H)$ which together with (a) yields Poincaré duality with $\mathbb{Z}_2$-coefficients.

8.5 Hamiltonian Floer homology in the aspherical case

In this section we explain how the perturbation theory on M-polyfolds can be used to define Floer homology in the aspherical case. See [34] for background on Floer homology. We consider the following setup:

- (M, ω) is a $2n$-dimensional symplectic manifold which is closed and *symplectically spherical*, i.e. $[\omega]$ and $c_1(TM)$ vanish on $\pi_2(M)$;
- $H : S^1 \times M \to \mathbb{R}$ is a time dependent Hamiltonian function, where $S^1 = \mathbb{R}/\mathbb{Z}$, such that all 1-periodic orbits of its Hamiltonian system are nondegenerate;
- J is a compatible almost complex structure on M.

Here the Hamiltonian system is the ODE

$$\dot{x}(t) = X_{H_t}\big(x(t)\big),$$

where the Hamiltonian vector field X_{H_t} of $H_t = H(t, \cdot)$ is defined by $dH_t + \omega(X_{H_t}, \cdot) = 0$. For each contractible loop $x : S^1 \to M$ we define its *Hamiltonian action*

$$\mathcal{A}_H(x) := \int_D \overline{x}^*\omega - \int_0^1 H\big(t, x(t)\big)dt,$$

where $\overline{x} : D \to M$ is a smooth map from the closed unit disk $D \subset \mathbb{C}$ with $\overline{x}|_{\partial D} = x$. This is independent of the choice of $\overline{x}$ because $[\omega]|_{\pi_2(M)} = 0$. Critical points of $\mathcal{A}_H$ are contractible

1-periodic orbits, and positive L^2-gradient trajectories of $\mathcal{A}_H$ are solutions $u : \mathbb{R} \to S^1 \to M$ of the *Floer equation*

$$\bar{\partial}_H(u) := \partial_s u + J(u)\partial_t u + \nabla H(t,u) = 0. \tag{8.10}$$

Now we proceed as in the previous section, replacing the domain $\mathbb{R}$ by $\mathbb{R} \times S^1$ and the negative gradient flow equation by the Floer equation, so we will only indicate the necessary modifications.

Unbroken cylinders. The Hessian of $\mathcal{A}_H$ at a 1-periodic orbit x defines a symmetric operator $S(x)$ on sections of the pullback bundle $x^*TM \to S^1$, and we define $\delta(x) > 0$ to be the minimal distance of an eigenvalue of $S(x)$ to 0. We pick positive weights $0 < \delta_0(x) < \delta_1(x) < \cdots < \delta(x)$. For contractible 1-periodic orbits x^-, x^+ with $\mathcal{A}_H(x^-) < \mathcal{A}_H(x^+)$ and $m \in \mathbb{N}_0$ we define $\widetilde{\mathcal{E}}_m(x^-, x^+)$ to be the space of H^{m+3}_{loc}-maps $u : \mathbb{R} \times S^1 \to M$ with exponential convergence to $x^\pm$ with rates $\delta_m(x^\pm)$ as $s \to \pm\infty$. Note the increased level of regularity $m+3$, which ensures that $\widetilde{\mathcal{E}}_0(x^-, x^+) \subset C^1(\mathbb{R} \times S^1, M)$. These spaces define an sc-manifold $\widetilde{\mathcal{E}}(x^-, x^+)$ with a free $\mathbb{R}$-action $u \mapsto u(\cdot + s_0, \cdot)$, hence a quotient sc-manifold $\mathcal{E}(x^-, x^+)$. The Floer operator in (8.10) descends to an sc-Fredholm section in a strong sc-bundle $\mathcal{F}(x^-, x^+) \to \mathcal{E}(x^-, x^+)$ which we still denote by $\bar{\partial}_H$. Its index is computed in [40] to be

$$\mathrm{ind}(\bar{\partial}_H) = \mathrm{CZ}(x^+) - \mathrm{CZ}(x^-) - 1,$$

where CZ denotes the *Conley-Zehnder index*. Note that compared to Corollary 8.6 the roles of x^+ and x^- are interchanged, which reflects the fact the the Floer equation describes the *positive* gradient flow of the Hamiltonian action functional. This is compatible with Corollary 8.6 because the Morse and Conley-Zehnder index of a critical point p of H are related by $\mathrm{CZ}(p) = n - \mathrm{Morse}(p)$.

Broken cylinders. As in the previous section we define for $\mathcal{A}_H(x_0) < \cdots < \mathcal{A}_H(x_k)$ sc-manifolds of broken cylinders

$$\mathcal{E}(x_0, x_1, \dots, x_k) := \mathcal{E}(x_0, x_1) \times \cdots \times \mathcal{E}(x_{k-1}, x_k)$$

and sets

$$X(x^-, x^+) := \mathcal{E}(x^-, x^+) \amalg \coprod_{k \geq 1} \coprod_{x_1, \dots, x_k} \mathcal{E}(x^-, x_1, \dots, x_k, x^+)$$

with corresponding bundles $\mathcal{F}(x_0, x_1, \dots, x_k) \to \mathcal{E}(x_0, x_1, \dots, x_k)$ and $Y(x^-, x^+) \to X(x^-, x^+)$. Then Theorem 8.7 has the following analogue.

Theorem 8.11 (M-polyfold Fredholm setup for Floer homology)**.**
(a) The set $X(x^-, x^+)$ carries the natural structure of a tame M-polyfold whose interior is $X^{(0)}(x^-, x^+) = \mathcal{E}(x^-, x^+)$ and whose set of points of degeneracy index $k \geq 1$ is

$$X^{(k)}(x^-, x^+) = \coprod_{x_1, \dots, x_k} \mathcal{E}(x^-, x_1, \dots, x_k, x^+),$$

and such that the induced M-polyfold structure on $X^{(k)}(x^-, x^+)$ agrees with the sc-manifold structure.
(b) The set $Y(x^-, x^+)$ carries the natural structure of a strong bundle over $X(x^-, x^+)$ which agrees on each stratum with the strong sc-bundle $Y^{(k)}(x^-, x^+) \to X^{(k)}(x^-, x^+)$.
(c) The products of $\bar\partial_H$ on the strata $X^{(k)}(x^-, x^+)$ induce an sc-Fredholm section $f : X(x^-, x^+) \to Y(x^-, x^+)$ of index

$$\mathrm{ind}(f) = \mathrm{CZ}(x^+) - \mathrm{CZ}(x^-) - 1$$

and with compact solution set $S_f = f^{-1}(0)$.

The proof is completely analogous to that of Theorem 8.7, the only difference being that for compactness of the solution set the

ODE arguments in Morse theory need to be replaced by elliptic regularity arguments in Floer theory, see [34]. As in the previous section we get commuting diagrams

$$\begin{tikzcd} \partial Y(x^-,x^+) \arrow[r, equal] & \coprod_p Y(x^-,x) \times Y(x,x^+) \\ \partial X(x^-,x^+) \arrow[r, equal] \arrow[u, "f|_{\partial X(x^-,x^+)}"] & \coprod_p X(x^-,x) \times X(x,x^+), \arrow[u, "f\times f"'] \end{tikzcd} \tag{8.11}$$

and the proof of Theorem 8.8 carries over to give

Theorem 8.12 (Abstract perturbations for Floer homology)**.** *In the setup of Theorem 8.11, there exists an* sc^+*-section* s *of the strong bundle*

$$Y := \coprod_{\mathcal{A}_H(x^-)<\mathcal{A}_H(x^+)} Y(x^-,x^+) \to X := \coprod_{\mathcal{A}_H(x^-)<\mathcal{A}_H(x^+)} X(p^-,p^+)$$

such that the sc-Fredholm section $f+s : X \to Y$ *has the following properties:*
(i) For all x^-,x^+ *the solution set* $\{u \in X(x^-,x^+) \mid f(u)+s(u)=0\}$ *is compact.*
(ii) $f+s$ *is transverse to the zero section.*
(iii) For all x^-,x^+ *the diagram* (8.11) *commutes with* f *replaced by* $f+s$*.*

Together with the Implicit Function Theorem 7.13 this implies

Corollary 8.13. *For the* sc^+*-section* s *from Theorem 8.12, each solution set*

$$\mathcal{M}(x^-,x^+) := \{u \in X(x^-,x^+) \mid f(u)+s(u)=0\}$$

is a smooth compact manifold with corners of dimension

$$\dim \mathcal{M}(x^-,x^+) = \mathrm{CZ}(x^+) - \mathrm{CZ}(x^-) - 1.$$

Moreover, the diagram (8.11) *induces for the boundaries* $\partial\mathcal{M}(x^-,x^+) = \mathcal{M}(x^-,x^+)\cap\partial X$ *canonical diffeomorphisms*

$$\partial\mathcal{M}(x^-,x^+) \cong \coprod_x \mathcal{M}(x^-,x)\times\mathcal{M}(x,x^+). \tag{8.12}$$

Definition of Floer homology. As in the previous section, we now use Corollary 8.13 to define Floer homology with coefficients in $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$. Let $FC_k(H)$ be the $\mathbb{Z}_2$-vector space with basis the contractible 1-periodic orbits of Conley-Zehnder index k. For a compatible almost complex structure J and an abstract perturbation s as in Theorem 8.12 we define a linear map $\partial_k : FC_k(H) \to FC_{k-1}(H)$ on basis elements by

$$\partial_k x^+ := \sum_{\mathrm{CZ}(x^-)=k-1} \langle\partial_k x^+, x^-\rangle x^-, \qquad \langle\partial_k x^+, x^-\rangle := \#\mathcal{M}(x^-,x^+).$$

Note again the reversed roles of x^+ and x^-, which reflects the fact that we use positive gradient trajectories of the Hamiltonian action to define its homology, so the boundary operator maps x^+ to x^-. The same argument as before yields $\partial_{k-1}\circ\partial_k = 0$ and we define the *Floer homology groups*

$$FH_k(H,J,s) := \ker\partial_k/\operatorname{im}\partial_{k+1}.$$

As in Problem 8.4 it follows that $FH_k(H,J,s) = FH_k(H)$ is independent of the almost complex structure J, the abstract perturbation s, and the Hamiltonian function H, so it is an invariant of the symplectic manifold (M,ω). In fact, this invariant is nothing new:

Theorem 8.14 (Floer, Salamon–Zehnder [43])**.** *For each closed symplectically apherical manifold* (M,ω) *and exhausting Hamiltonian function* $H : M \to \mathbb{R}$ *with nondegenerate contractible* 1*-periodic orbits we have*

$$FH_k(H) \cong MH_k(H) \cong H_k(M;\mathbb{Z}_2).$$

The main step in the proof consists in showing that for a sufficiently C^2-small Morse function $H : M \to \mathbb{R}$ all Floer cylinders connecting two critical points p^-, p^+ of index difference 1 are independent of the S^1-coordinate and thus negative Morse gradient trajectories, so the Floer complex for H equals the Morse complex.

We will outline in Section 9.6 below an alternative proof of this fact. Will we?

Extensions:
, possibly with cylindrical end $\big([1,\infty)\times N, d(r\alpha)\big)$ modeled over a closed contact manifold (N,α); which on the cylindrical end is a linear function of r whose slope is not the action of a closed Reeb orbit on N; which is cylindrical on the cylindrical end.

Chapter 9

Groupoids and orbifolds

Many geometric elliptic PDEs arise in the following general setup of a "$\mathcal{G}$-moduli problem":

- $\mathcal{E} \to \mathcal{B}$ is a vector bundle with a Lie group $\mathcal{G}$ acting on the total space and the base in a compatible way;
- $\mathcal{F} : \mathcal{B} \to \mathcal{E}$ is a $\mathcal{G}$-equivariant section which is "Fredholm modulo $\mathcal{G}$" and such that the moduli space $\mathcal{M} = \mathcal{F}^{-1}(0)/\mathcal{G}$ is compact.

Here we are deliberately vague about the kinds of spaces: $\mathcal{E} \to \mathcal{B}$ might be a finite dimensional vector bundle, a Hilbert space bundle, or a strong bundle over an M-polyfold, and $\mathcal{G}$ might be finite or infinite dimensional. For example, $\mathcal{G} = \mathbb{R}$ in Morse theory, $\mathcal{G} = PSL(2, \mathbb{C})$ in genus zero Gromov-Witten theory, and $\mathcal{G}$ is the infinite dimensional gauge group in Donaldson or Seiberg–Witten theory. The property "Fredholm modulo $\mathcal{G}$" means on the linear level that

$$0 \longrightarrow \mathrm{Lie}\,\mathcal{G} \xrightarrow{L_x} T_x\mathcal{B} \xrightarrow{T_x^v\mathcal{F}} \mathcal{E}_x \longrightarrow 0$$

is a Fredholm complex, where L_x denotes the infinitesimal action. In this setup, $\mathcal{M}$ should be a finite dimensional compact manifold, provided that $\mathcal{F}$ is transverse to the zero section and $\mathcal{G}$ acts freely on $\mathcal{F}^{-1}(0)$. Now several cases can arise.

(1) $\mathcal{G}$ acts freely on $\mathcal{F}^{-1}(0)$. Then after shrinking $\mathcal{B}$ we may assume that $\mathcal{G}$ acts freely on $\mathcal{B}$ and $\mathcal{E}$, so $\mathcal{E}/\mathcal{G} \to \mathcal{B}/\mathcal{G}$ is a smooth vector bundle and we can perturb the induced section $\overline{\mathcal{F}} : \mathcal{B}/\mathcal{G} \to \mathcal{E}/\mathcal{G}$ to achieve transversality. This case arises e.g. in Morse homology and Hamiltonian Floer homology in the aspherical case and was treated in the previous chapter.
(2) $\mathcal{G}$ acts on $\mathcal{F}^{-1}(0)$ with finite isotropy groups. This case arises e.g. in Gromov-Witten theory and general Hamiltonian Floer homology (due to multiply covered holomorphic curves) and will be treated in this chapter.
(3) $\mathcal{G}$ has infinite isotropy groups on $\mathcal{F}^{-1}(0)$. This case arises e.g. in Donaldson and Seiberg–Witten theory (due to reducible connections which have isotropy group S^1) and lies outside the scope of this lecture.

In case (2) we can distinguish two subcases:
(a) The section $\mathcal{F}$ can be perturbed to make it transverse to the zero section (keeping it $\mathcal{G}$-equivariant). Then $\mathcal{M} = \mathcal{F}^{-1}(0)/\mathcal{G}$ is a finite dimensional *orbifold* as in the following problem (where $\mathcal{F}^{-1}(0) = S^3$).
(b) As the following problem shows, it may not be possible to make $\mathcal{F}$ transverse to the zero section, keeping it $\mathcal{G}$-equivariant. We will see that in this case we can still achieve transversality using *multivalued perturbations*. The resulting moduli spaces will in general not be orbifolds but *weighted branched manifolds*.

Problem 9.1. Consider the trivial bundle $E = T^2 \times \mathbb{R}^2 \to B = T^2 = \mathbb{R}^2/\mathbb{Z}^2$ with the generator of $G = \mathbb{Z}_2$ acting on the total space by $(x_1, x_2, y_1, y_2) \mapsto (x_1, -x_2, -y_1, -y_2)$. Prove: every $\mathbb{Z}_2$-equivariant section $F : B \to E$ must vanish on $S^1 \times \{0\}$ and is therefore not transverse to the zero section.

The references for this chapter are [2, 36, 11, 41, 22, 23, 27, 25, 26].

9.1 Orbifolds

This section follows [2].

Group actions. We begin with some basic notions on group actions.

Definition 9.1. Let $G \times X \to X$ be the action of a group G on a set X. For $x \in X$ we define

- the *orbit* $G \cdot x := \{gx \mid g \in G\} \subset X$ of x;
- the *isotropy subgroup* $G_x := \{g \in G \mid gx = x\} \subset G$ of x.

We denote by X/G the *orbit (or quotient) space.* The action is called

- *effective* if $gx = x$ for all $x \in X$ implies $g = 1$;
- *free* if $G_x = \{1\}$ for all $x \in X$;
- *locally free* if G_x is finite for all $x \in X$.
- *proper* if the map $G \times X \to X \times X$, $(g, x) \mapsto (x, gx)$ is proper (provided that G and X carry topologies).

Note that properness implies that all isotropy groups G_x are compact. If G is a Lie group and $X = M$ a manifold we will always assume that the map $G \times M \to M$ is smooth.

Problem 9.2. Prove: if $G \times M \to M$ is a proper free action of a Lie group G on a manifold M, then the quotient space M/G inherits the structure of a manifold.

Problem 9.3. Prove the *local slice theorem*: if $G \times M \to M$ is a proper action of a Lie group G on a manifold M, then through each $x \in M$ there exist a *local slice* (U, ϕ) where

- $U \subset \mathbb{R}^n$ is an open disk around 0 with an orthogonal linear action $G_x \times U \to U$;
- $\phi : U \hookrightarrow M$ is a G_x-equivariant embedding with $\phi(0) = x$ such that the induced map

$$G \times_{G_x} U \to X. \qquad [g, y] \mapsto g\phi(y)$$

is a G-equivariant diffeomorphism onto a G-invariant neighbourhood of $G \cdot x$. Here $G \times_{G_x} U = (G \times U)/\sim$ with $(g, x) \sim (gh^{-1}, hx)$ for $h \in G_x$.

Hint: Use compactness of G_x to find a G_x-invariant Riemannian metric on M, show that $\exp_x : T_xM \to M$ *is G_x-equivariant, and take for U a small disk in the orthogonal complement to $T_x(G \cdot x)$ in T_xM.*

Orbifolds. Orbifolds arise as quotient spaces M/G of proper *locally finite* actions of a Lie group G on a manifold M. By the local slice theorem, M/G is locally homeomorphic to the quotient U/G_x of an open disk $U \subset \mathbb{R}^n$ by an orthogonal linear action of the *finite* group G_x. This motivates the following

Definition 9.2. Let X be a topological space and $n \in \mathbb{N}_0$.
- An n-dimensional *orbifold chart* (U, G, ϕ) consists of a connected open subset $U \subset \mathbb{R}^n$, a finite group G acting smoothly and *effectively* on U, and a G-invariant map $\phi : U \to X$ inducing a homeomorphism from U/G onto an open subset $\widetilde{U} \subset X$.
- An *embedding* $\lambda : (U, G, \phi) \hookrightarrow (V, H, \psi)$ of orbifold charts is a smooth open embedding $\lambda : U \hookrightarrow V$ with $\psi \circ \lambda = \phi$; see Problem 9.4.
- An *orbifold atlas* on X is a family $\mathcal{U}$ of orbifold charts covering X such that for any two charts (U, G, ϕ), (V, H, ψ) and $x \in \widetilde{U} \cap \widetilde{V}$ there exists a chart (W, K, ϑ) with $x \in \widetilde{W} \subset \widetilde{U} \cap \widetilde{V}$ and embeddings $(W, K, \vartheta) \hookrightarrow (U, G, \phi)$ and $(W, K, \vartheta) \hookrightarrow (V, H, \psi)$.

• An orbifold atlas $\mathcal{U}$ is a *refinement* of $\mathcal{V}$ if every chart of $\mathcal{U}$ has an embedding into a chart of $\mathcal{V}$. Two orbifold atlases are *equivalent* if they have a common refinement.
• An *n-dimensional orbifold* is a paracompact Hausdorff space X equipped with an equivalence class of n-dimensional orbifold atlases.

Problem 9.4. (a) Show that every orbifold has an atlas consisting of charts (U, G, ϕ) for which the group G acts *linearly and orthogonally* on $U \subset \mathbb{R}^n$.
(b) Let $\lambda : (U, G, \phi) \hookrightarrow (V, H, \psi)$ be an embedding of orbifold charts. Show that there exist a unique injective group homomorphism $\widehat{\lambda} : G \hookrightarrow H$ which makes $\lambda : U \to V$ G-equivariant in the sense that $\lambda(gx) = \widehat{\lambda}(g)\lambda(x)$ for all $g \in G$ and $x \in U$.
(c) Define the *isotropy group* at a point x in an orbifold X as $G_x := G_y$ for any orbifold chart (U, G, ϕ) and $y \in U$ with $\phi(y) = x$. Show that G_x is independent of the choices up to group isomorphism. We call x *singular* if $G_x \neq \{1\}$.

Problem 9.5. Prove that the quotient M/G of a proper locally finite effective action of a Lie group G on a manifold M inherits the structure of an orbifold.

Problem 9.6. Prove the converse to Problem 9.5: Every n-orbifold X is diffeomorphic to the quotient of a proper locally finite effective action of the compact Lie group $O(n)$ on a manifold M. *Hint: For an orbifold chart (U, G, ϕ) as in Problem 9.4(a) let $FU \to U$ be the bundle of orthonormal frames whose fibre over $x \in U$ are the orthogonal maps $\mathbb{R}^n \to T_xU$. It carries a right action of $O(n)$ and a free left action of G, so FU/G is a manifold. These manifolds glue to a manifold FX with a right $O(n)$-action whose quotient is X.*

Problem 9.7. Consider the action of $S^1 = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$ on

$S^3 = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z|^2 + |z_2|^2 = 1\}$ by $\lambda(z_1, z_2) = (\lambda^p z_1, \lambda^q z_2)$ with relatively prime integers $p, q \geq 2$. Show that the quotient S^3/S^1 is an orbifold homeomorphic to the 2-sphere and compute the isotropy groups at its two singular points.

Problem 9.8 (Moduli space of elliptic curves).
(a) Prove that the action of $PSL(2, \mathbb{R})$ on the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$ by Möbius transformations is proper and locally free, so the quotient $\mathcal{M}_{1,0} := \mathbb{H}/PSL(2, \mathbb{R})$ is a 2-dimensional orbifold.
(b) Show that $\mathcal{M}_{1,0}$ is homeomorphic to $\mathbb{C}$ with two singular points and compute their local groups.

9.2 Lie groupoids

Moerdijk and Pronk [37] introduced an alternative description of orbifolds in terms of groupoids. This point of view has the advantage that functoriality properties become more transparent, and it also lends itself to later generalizations to polyfolds. We follow the exposition in [36].

Definition 9.3. A *groupoid* X is a small category such that every morphism has an inverse. Thus it is specified by its object and morphism sets X, $\mathbf{X}$ together with its five structure maps:

- source and target maps $s, t : \mathbf{X} \to X$;
- composition $m : \mathbf{X}\,{}_s\!\times_t \mathbf{X} = \{(f, g) \in \mathbf{X} \times \mathbf{X} \mid s(f) = t(g)\} \to \mathbf{X}$, $(f, g) \mapsto f \circ g$;
- unit $u : X \to \mathbf{X}$, $x \mapsto \mathbb{1}_x$;
- inverse $i : \mathbf{X} \to \mathbf{X}$, $g \mapsto g^{-1}$.

Its *orbit space* is the set $|X|$ of equivalence classes, where $x \sim y$ iff there exists $g \in \mathbf{X}$ with $s(g) = x$ and $t(g) = y$. The morphisms from $x \in X$ to itself form the *isotropy group* $G_x = \mathrm{Mor}_X(x, x)$.

Definition 9.4. A *Lie groupoid* is a groupoid whose object and morphism spaces are manifolds such that all structure maps are smooth. Moreover, the source and target maps are required to be submersions, so that $\mathbf{X}\,{}_s\!\times_t\mathbf{X}$ is again a manifold and smoothness of the composition map makes sense. A Lie groupoid is called

- *étale* if the source and target maps are local diffeomorphisms;
- *proper* if the map $s \times t : \mathbf{X} \to X \times X$ is proper.

We will abbreviate "étale proper Lie groupoid" as *ep Lie groupoid*.

Problem 9.9. For a Lie groupoid X prove:
(a) All points in the orbit of $x \in X$ have isomorphic isotropy groups.
(b) If the source map s is a submersion (resp. a local diffeomorphism), then so is the target map t.
(c) X is proper if and only if every point $x \in X$ has an open neighbourhood $V \subset X$ such that the map $t : s^{-1}(\overline{V}) \to X$ is proper. (This more complicated formulation of properness will be the appropriate one for the generalization to polyfolds.)

Example 9.1 (Examples of Lie groupoids).
(a) Each Lie group G can be viewed as a Lie groupoid with $X = \mathrm{pt}$ and $\mathbf{X} = G$. It is proper iff G is compact, and étale iff G is discrete.
(b) Each manifold M can be viewed as an ep Lie groupoid with $\mathbf{X} = X = M$ and $s = t = \mathbb{1}_M$. More interestingly, consider an atlas $\mathcal{U}$ for $M = \bigcup_{i\in I} V_i$ consisting of coordinate charts $\phi_i : \mathbb{R}^n \supset U_i \xrightarrow{\cong} V_i \subset M$. Set $U := \coprod_{i\in I} U_i$ with the local diffeomorphism

$\phi : U \to M$ given by ϕ_i on U_i. Then

$$\mathbf{M}_{\mathcal{U}} := \{(x, y) \in U \times U \mid \phi(x) = \phi(y)\} \xrightarrow{s,t} M_{\mathcal{U}} := U$$

with $s(x, y) = x$ and $t(x, y) = y$ is an ep Lie groupoid whose orbit space is canonically diffeomorphic to M. It is called the *refinement* of M with respect to $\mathcal{U}$ and denoted by $M_{\mathcal{U}}$.

(c) Combining examples (a) and (b), consider an effective action $G \times M \to M$ of a Lie group G on a manifold M. Then

$$G \times M \xrightarrow{s,t} M$$

with $s(g, x) = x$ and $t(g, x) = gx$ is a Lie groupoid with orbit space M/G. It is called the *translation* (or *action*) *groupoid* and denoted by $G \ltimes M$. It is proper iff the action is proper, and étale iff G is discrete. Note that the isotropy group of this groupoid at $x \in M$ equals G_x, so the action groupoid has finite isotropy groups iff the action is locally free.

Suppose now that the action is proper. Pick a collection $\mathcal{U}$ of local slices $\phi_i : \mathbb{R}^n \supset U_i \hookrightarrow M$, $i \in I$ such that every orbit meets some $\phi_i(U_i)$. Set $U := \coprod_{i \in I} U_i$ with the local embedding $\phi : U \to M$ given by ϕ_i on U_i. Then

$$(\mathbf{G} \ltimes \mathbf{M})_{\mathcal{U}} := \{(x, g, y) \in U \times G \times U \mid g\phi(x) = \phi(y)\} \xrightarrow{s,t} U$$

with $s(x, g, y) = x$ and $t(x, g, y) = y$ is a proper Lie groupoid whose orbit space is canonically homeomorphic to M/G. It is called the *refinement* of $G \ltimes M$ with respect to $\mathcal{U}$ and denoted by $(G \ltimes M)_{\mathcal{U}}$. It is étale iff the action is locally free. Note that there is a canonical functor from this groupoid to the action groupoid given on objects by $\phi : U \to M$ and on morphisms by

$$(\mathbf{G} \ltimes \mathbf{M})_{\mathcal{U}} \to G \times M, \qquad (x, g, y) \mapsto \big(g, \phi(x)\big).$$

(d) Let M be a connected manifold. The *fundamental groupoid* $\Pi(M)$ has object space M, with morphisms from x to y the homotopy classes of continuous paths from x to y. This is a Lie groupoid with isotropy group $\Pi(M)_x = \pi_1(X, x)$ the fundamental group at $x \in M$. Note that $s \times t : \mathrm{Mor}\Pi(M) \to M \times M$ is a covering with discrete fibre $\pi_1(M)$, so $\Pi(M)$ is proper iff $\pi_1(M)$ is finite, and étale iff $M = \{\mathrm{pt}\}$.

Problem 9.10. Verify all the assertions in Example 9.1.

Smooth functors and equivalences. We begin with some notions from category theory.

Definition 9.5. Let X, Y be categories and $F, G : X \to Y$ functors. A *natural transformation* $\alpha : F \to G$ associates to each object $x \in X$ a morphism $\alpha(x) \in \mathrm{Mor}_Y\big(F(x), G(x)\big)$ such that for each $\phi \in \mathrm{Mor}_X(x, x')$ the following diagram commutes:

$$\begin{tikzcd} F(x) \arrow[r, "\alpha(x)"] \arrow[d, "F(\phi)"'] & G(x) \arrow[d, "G(\phi)"] \\ F(x') \arrow[r, "\alpha(x')"] & G(x') \end{tikzcd}$$

It is called a *natural equivalence* if each $\alpha(x)$ is an isomorphism. A functor $F : X \to Y$ is called

- *faithful* (resp. *full*) if for all objects $x, x' \in X$ the induced map $\mathrm{Mor}_X(x, x') \to \mathrm{Mor}_Y\big(F(x), F(y)\big)$ is injective (resp. surjective);
- *essentially surjective* if for every object $y \in Y$ there exists an object $x \in X$ and an isomorphism $F(x) \to y$;
- an *equivalence* if there exists a functor $G : Y \to X$ and natural equivalences $FG \to \mathbb{1}_Y$ and $GF \to \mathbb{1}_X$.

Problem 9.11. For categories X, Y show:
(a) A functor $F : X \to Y$ is an equivalence if and only if it is faithful, full and essentially surjective.
(b) An equivalence $F : X \to Y$ induces a bijection $|F| : |X| \to |Y|$ between the orbit sets.

ets? Terminology?

Definition 9.6. Let X, Y be ep Lie groupoids. A functor $F : X \to Y$ is called *smooth* if the corresponding maps on objects and morphisms are smooth. It is called an *equivalence* if it is an equivalence of categories and a local diffeomorphism on objects.

Problem 9.12. Show that a smooth functor $F : X \to Y$ between ep groupoids is an equivalence iff

(i) $F : \mathrm{Mor}_X(x, x') \to \mathrm{Mor}_Y\big(F(x), F(x')\big)$ is a bijection for all $x, x' \in X$;

(ii) $|F| : |X| \to |Y|$ is a homeomorphism;

(iii) $F : X \to Y$ is a local diffeomorphism.

The following two problems show that an equivalence between ep Lie groupoids need not have an inverse (so the term "equivalence" is rather misleading, but this is the terminology used in [36] and [26]).

Problem 9.13. Let G be a finite group acting freely on the manifold M, so that M/G is again a manifold. Show:
(a) The canonical functor $F : G \ltimes M \to M/G$ is an equivalence.
(b) If M is connected and $G \neq \{1\}$, then there exists no equivalence $M/G \to G \ltimes M$.

Problem 9.14. Let X be an ep groupoid and $\mathcal{U}$ be an open covering of the object manifold X. Define the *refinement* $X_{\mathcal{U}}$ of X with respect to $\mathcal{U}$ by its object and morphism spaces

$$X_{\mathcal{U}} := \coprod_{U\in\mathcal{U}} U = \{(x,U) \mid x \in U, U \in \mathcal{U}\},$$
$$\mathbf{X}_{\mathcal{U}} := \{(U,\phi,V) \in \mathcal{U} \times \mathbf{X} \times \mathcal{U} \mid s(\phi) \in U, t(\phi) \in V\}$$

with $s(U,\phi,V) := \big(s(\phi),U\big)$ and $t(U,\phi,V) := \big(t(\phi),V\big)$. Prove:
(a) $X_{\mathcal{U}}$ is an ep Lie groupoid and the canonical map $F : X_{\mathcal{U}} \to X$, $(x,U) \mapsto x$ is an equivalence.
(b) Upon replacing the $U_i \subset \mathbb{R}^n$ by their images in M, the groupoids $M_{\mathcal{U}}$ and (for a locally free action) $(G \ltimes M)_{\mathcal{U}}$ in Example 9.1(b) and (c) are special cases of this construction.
(c) If M is connected and not covered by a single set from $\mathcal{U}$, then there exists no equivalence $M \to M_{\mathcal{U}}$.

Orbifolds via ep Lie groupoids. Now we can give an equivalent definition of an orbifold in terms of ep Lie groupoids.

Definition 9.7. Let Z be a topological space.
- An *orbifold structure* on Z is a pair (X,α) consisting of an ep Lie groupoid X and a homeomorphism $\alpha : |X| \to Z$.
- Two orbifold structures (X,α) and (Y,β) are called *equivalent* if there exists an ep groupoid A and equivalences

$$X \xleftarrow{F} A \xrightarrow{G} Y$$

such that $\alpha \circ |F| = \beta \circ |G| : |A| \to Z$.
- An *orbifold* is a paracompact Hausdorff space Z equipped with an equivalence class of orbifold structures.

The slightly complicated formulation of equivalence of orbifolds structures is necessary because equivalences of ep groupoids need not have inverses. It corresponds to the notion of a common refinement of open covers.

Theorem 9.2. *Definitions 9.2 and 9.7 of an orbifold are equivalent.*

Proof. □

9.3 Moduli spaces of Riemann surfaces

In this and the following section we outline the construction of orbifold structures on moduli spaces of Riemann surfaces and their Deligne–Mumfod compactifications. We mostly follow [21], see also [41] for a similar approach. In ths section we consider Riemann surfaces without nodes and marked points.

Definition 9.8. In these notes, a *Riemann surface* (S, j) means a closed connected oriented surface S of genus g equipped with an almost complex structure j (which is automatically integrable). An *isomorphism* $\phi : (S, j) \to (S', j')$ of Riemann surfaces is a diffeomorphism $\phi : S \to S'$ such that $T\phi \circ j = j' \circ T\phi$. The *Riemann moduli space* $\mathcal{M}_g$ is the set of isomorphism classes of Riemann surfaces of genus g.

For $g = 0$ the space $\mathcal{M}_0$ consists of one point $[S^2, j_{\rm st}]$ with automorphism group $\mathrm{Aut}(S^2, j_{\rm st}) = PSL(2, \mathbb{C})$. For $g = 1$ the space $\mathcal{M}_1 \cong \mathbb{H}/PSL(2, \mathbb{Z})$ agrees with the moduli space of smooth elliptic curves, and the automorphism group of each elliptic curve C contains C acting on itself by translations. So both these spaces are not orbifolds due to presence of infinite automorphism groups. By contrast we have

Theorem 9.3. *For $g \geq 2$, the Riemann moduli space $\mathcal{M}_g$ is a complex orbifold of complex dimension $3g - 3$.*

In the following we will outline the proof of this result. Since all closed connected oriented surfaces of genus g are diffeomorphic, we fix a smooth model surface S of genus $g \geq 2$. Let $\mathcal{J}$ be the space of almost complex structures on S defining the orientation of S, which is a Fréchet manifold with the C^∞-topology. The group $\mathcal{G} = \mathrm{Diff}_+(S)$ of orientation preserving diffeomorphisms of S is a Fréchet Lie group with the C^∞-topology. It acts smoothly on $\mathcal{J}$ by conjugation $\phi_* j := T\phi \circ j \circ (T\phi)^{-1}$ and the quotient is the Riemann moduli space

$$\mathcal{M}_g = \mathcal{J}/\mathcal{G}.$$

We wish to show that this action is locally free and proper. As a first step we consider the linearized action at $j \in \mathcal{J}$,

$$L_j : \mathrm{Lie}\,\mathcal{G} \to T_j\mathcal{J}.$$

Here $\mathrm{Lie}\,\mathcal{G} = \Omega^0(TS)$ is the space of vector fields on S, i.e., smooth sections in the bundle $TS \to S$. Linearization of the equation $j(z)^2 = -1$ shows that $T_j\mathcal{J} = \Omega^{0,1}(TS)$ is the space of smooth sections η in the bundle $\mathrm{Hom}(TS, TS)$ which are $\mathbb{C}$-antilinear in the sense that $\eta(z)j(z) + j(z)\eta(z) = 0$ for all $z \in S$. The linearized action is defined by $L_j\xi = \frac{d}{dt}|_{t=0}\phi_{t*}j$, where ϕ_t is the flow of $\xi \in \Omega^0(TS)$. In a local holomorphic coordinate z on $U \subset \mathbb{C}$ we have $j = i$ and

$$(\phi_{t*}i)\big(\phi_t(z)\big) = D\phi_t(z) \circ i \circ D\phi_t(z)^{-1},$$

so the derivative at $t = 0$ yields

$$L_j\xi(z) = D\xi(z) \circ i - i \circ D\xi(z) = -2i\bar{\partial}\xi(z)$$

with the Cauchy-Riemann operator

$$\bar{\partial}\xi := \frac{1}{2}(D\xi + i \circ D\xi \circ i).$$

Hence the linearized action is (up to multiplication by $-2i$) given by

$$\bar{\partial} : \Omega^0(TS) \to \Omega^{0,1}(TS). \tag{9.1}$$

To proceed further we need some background.

Holomorphic line bundles. Let $E \to S$ be a holomorphic line bundle of *degree* $d = \langle c_1(E), [S] \rangle$ over a genus g Riemann surface S. Denote by $\Omega^0(E)$ the space of smooth sections in the bundle $E \to S$ and by $\Omega^{0,1}(E)$ the space of smooth sections in the bundle $\mathrm{Hom}^{0,1}(TS, E)$ of $\mathbb{C}$-antilinear bundle maps $TS \to E$. The *Cauchy–Riemann operator*

$$\bar{\partial} : \Omega^0(E) \to \Omega^{0,1}(E)$$

is defined in local trivializations by

$$\bar{\partial}\xi = \frac{\partial \xi}{\partial \overline{z}} d\overline{z} = \frac{1}{2}(T\xi + j \circ T\xi \circ j).$$

It is a $\mathbb{C}$-linear first order elliptic differential operator whose kernel and cokernel are finite dimensional $\mathbb{C}$-vector spaces

$$H^0(E) := \ker \bar{\partial}, \qquad H^1(E) := \operatorname{coker} \bar{\partial}.$$

Theorem 9.4. *Let $E \to S$ be a holomorphic line bundle of degree $d = \langle c_1(E), [S] \rangle$ over a genus g Riemann surface S. Then the following results hold:*

(Riemann–Roch theorem)

$$\mathrm{ind}_{\mathbb{C}}(\bar{\partial}) = \dim_{\mathbb{C}} H^0(E) = \dim_{\mathbb{C}} H^1(E) = 1 - g + d.$$

(Serre duality)

$$H^1(E) \cong H^0(T^*S \otimes E^*)^*.$$

(Kodaira vanishing theorem)

$$d < 0 \Longrightarrow H^0(E) = \{0\}.$$

We apply this to the linearized action (9.1), i.e., to the holomorphic line bundle $E = TS$. Its degree is $d = \chi(S) = 2 - 2g < 0$ due to our assumption $g \geq 2$. Hence the Kodaira vanishing theorem yields $\ker \bar{\partial} = \{0\}$, so the action is *locally free*. From Riemann–Roch we get

$$\dim_{\mathbb{C}} H^1(TS) = \mathrm{ind}_{\mathbb{C}}(\bar{\partial}) = 1 - g + (2 - 2g) = 3 - 3g.$$

The "tangent space" to $\mathcal{M}_g$ at $[j]$ is $H^1(TS)$, which by Serre duality is isomorphic to the space $H^0(T^*S \otimes T^*S)$ of *holomorphic quadratic differentials*. Its complex dimension is

$$\dim_{\mathbb{C}} H^1(TS) = -\mathrm{ind}_{\mathbb{C}}(\bar{\partial}) = 3g - 3.$$

Next we consider properness of the action. It will follow from

Theorem 9.5 (Uniformization theorem)**.** *Let (S, j) be a (closed) Riemann surface of genus $g \geq 2$. Then S carries a unique hyperbolic metric h (i.e., h has constant curvature -1) compatible with j in the sense that $h(v, jv) = 0$ for all $v \in T_pS$. Moreover, h depends continuously on j with respect to the C^∞-topologies.*

Remark 9.6. (a) The uniformization theorem implies that each Riemann surface S of genus $g \geq 2$ is isomorphic to $\mathbb{H}/\Gamma$ for some discrete subgroup $\Gamma < PSL(2, \mathbb{R})$.
(b) The uniformization theorem also has versions for genus 0 and 1. For $g = 0$ there exists a unique compatible metric of constant curvature $+1$, so (S, j) is isomorphic to (S^2, j_{st}). For $g = 1$ there exists a compatible metric of constant curvature 0, unique up to scaling by a positive constant, so (S, j) is isomorphic to $(\mathbb{C}/\Lambda, i)$ for some lattice $\Lambda \subset \mathbb{C}$.

To verify properness, consider sequences $(j_k) \subset \mathcal{J}$ and $(\phi_k) \subset \mathcal{G}$ such that $j_k \to j$ and $\phi_{k*} j_k \to j'$. Then the hyperbolic metrics h_k, h, h' associated to j_k, j, j' by the uniformization theorem satisfy $h_k \to h$ and $\phi_{k*}h_k \to h'$. It follows that the derivatives (also higher ones) of ϕ_k with respect to the metrics h and h' are uniformly bounded as $k \to \infty$, so by the Arzela-Ascoli theorem a subsequence of ϕ_k converges to an (h, h')-isometry ϕ and properness is proved. In summary we have shown:

Proposition 9.7. *The action of* $\mathcal{G} = \mathrm{Diff}_+(S)$ *on the space* $\mathcal{J}$ *of orientation preserving almost complex structures is locally free and proper. The linearized action at* $j \in \mathcal{J}$ *is equivalent to*

$$\bar\partial : \Omega^0(TS) \to \Omega^{0,1}(TS),$$

which is an injective, complex linear, elliptic differential operator of index $\mathrm{ind}_{\mathbb{C}}(\bar\partial) = 3 - 3g$. □

If $\mathcal{J}$ were a Banach manifold and $\mathcal{G}$ a Banach Lie group, then it would follow via the implicit function that $\mathcal{M}_g = \mathcal{J}/\mathcal{G}$ is a complex orbifold of complex dimension $3g - g$. This argument fails in the Fréchet setting due to the lack of the implicit function theorem, and taking Sobolev completions of $\mathcal{J}$ and $\mathcal{G}$ doesn't work either because then the action is not differentiable. The following remark sketches a solution to this problem in the Fréchet setting; a solution in the Banach setting is described in the remainder of this section.

Remark 9.8 (Teichmüller space). The study of $\mathcal{M}_g$ can be simplified by first considering the normal subgroup $\mathcal{G}_0 \subset \mathcal{G}$ consisting of diffeomorphisms isotopic to the identity, which acts *freely* on $\mathcal{J}$. (To see this, consider an automorphism ϕ of (S, j) which is smoothly isotopic to the identity; then ϕ preserves all closed geodesics in the hyperbolic metric h associated to j, hence a geodesic pair-of-pants decomposition, and it follows that $\phi = \mathbb{1}$ because the only automorphism of a pair-of-pants isotopic to the identity is the identity.) So it suffices to show that *Teichmüller space* $\mathcal{T}_g := \mathcal{J}/\mathcal{G}_0$ carries the structure of a complex *manifold* of dimension $3g-3$, because then the remaining action of the discrete *mapping class group* $\Gamma_g := \mathcal{G}/\mathcal{G}_0$ on $\mathcal{T}_g$ is proper and locally free and thus induces a complex orbifold structure on $\mathcal{M}_g = \mathcal{T}_g/\Gamma_g$. There are several classical constructions of complex manifold structures on $\mathcal{T}_g$. For example, the *Bers embedding* realizes $\mathcal{T}_g$ as an open subset of $\mathbb{C}^{3g-3}$. Earle and Eels show that $\mathcal{T}_g$ is a smooth manifold diffeomorphic to $\mathbb{R}^{6g-6}$ by constructing a smooth *global slice* $\mathbb{R}^{6g-6} \hookrightarrow \mathcal{J}$ to the action of $\mathcal{G}_0$.

Local slices (good families). Our first goal is to construct local slices for the action $\mathcal{G} \times \mathcal{J} \to \mathcal{J}$ through $j \in \mathcal{J}$. Recall that the candidate for the tangent space to $\mathcal{M}_g$ at $[j]$ is

$$E := H^1(TS, j) = \Omega^{0,1}(TS, j)/\mathrm{im}\,\bar{\partial}_j$$

(where we include j in the notation because it will vary in the sequel) modulo the $\mathbb{C}$-linear action of the finite automorphism group

$$G := \mathrm{Aut}(S, j).$$

So a local slice should be a G-equivariant smooth embedding

$$E \supset V \hookrightarrow \mathcal{J}, \qquad v \mapsto j(v), \quad 0 \mapsto j$$

of a G-invariant open neighbourhood $V \subset E$ of 0 which is transverse to the $\mathcal{G}$-orbits. To formulate the last condition, consider the differential

$$Dj(v) : E \to T_{j(v)}\mathcal{J} = \Omega^{0,1}\big(TS, j(v)\big)$$

and its composition with the projection onto H^1, the *Kodaira differential*

$$[Dj(v)] : E \to \Omega^{0,1}\big(TS, j(v)\big) \to H^1\big(TS, j(v)\big).$$

Definition 9.9. A *deformation of* j is a smooth map $j : V \to \mathcal{J}$, $v \mapsto j(v)$ from an open neighbourhood $V \subset E := H^1(TS, j)$ of 0 with $j(0) = j$. It is called

- *effective* if the Kodaira differential $[Dj(v)]$ is an $\mathbb{R}$-linear isomorphism for each $v \in V$;
- *complex* if $[Dj(v)]$ is $\mathbb{C}$-linear for each $v \in V$;
- *symmetric* if V is invariant under $G = \mathrm{Aut}(S, j)$ and $g_*(j(v)) = j(g_* v)$ for each $g \in G$ and $v \in V$, so g induces an isomorphism $g : (S, j) \to \big(S, j(g_* v)\big)$.

An effective symmetric deformation is called a *good family*.

Note that effectiveness corresponds to transversality to the $\mathcal{G}$-orbits and symmetry to G-equivariance, so a good family is nothing but a *local slice* for the action $\mathcal{G} \times \mathcal{J} \to \mathcal{J}$.

The construction of a good family is based on the following linear algebra exercise.

Problem 9.15. Let (V, ω, J) be a Hermitian vector space. Show that the *generalized Cayley transform*

$$A \mapsto J(\mathbb{1} + A)(\mathbb{1} - A)^{-1}$$

defines a diffeomorphism from the space of $\mathbb{C}$-antilinear maps $A : V \to V$ with $\|A\| < 1$ onto the space of almost complex structures K tamed by ω (i.e. such that $\omega(v, Jv) > 0$ for all $v \neq 0$) with inverse

$$K \mapsto (K + J)^{-1}(K - J).$$

Now we return to our closed oriented surface S of genus $g \geq 2$. Recall that $\mathcal{J}$ denotes the space of almost complex structures on S defining the orientation of S, which is equivalent to being tamed by a positive area form. Fix $j \in \mathcal{J}$ with $G = \mathrm{Aut}(S, j)$ and denote by $|v| := \max_{z\in S}|v(z)|$ the maximum norm of $v \in \Omega^{0,1}(TS, j)$. Applying the preceding exercise pointwise at all $z \in S$ we thus obtain a diffeomorphism

$$j : \{v \in \Omega^{0,1}(TS, j) \mid |v| < 1\} \to \mathcal{J}, \qquad j(v) := j(\mathbb{1}{+}v)(\mathbb{1}{-}v)^{-1}.$$

Pick a G-invariant complex subspace $E \subset \Omega^{0,1}(TS, j)$ such that the projection $\Omega^{0,1}(TS, j) \to H^1(TS, j)$ restricts to an isomorphism $E \xrightarrow{\cong} H^1(TS, j)$ (for example the orthogonal complement to $\mathrm{im}\,\bar\partial_j$ with respect to a G-invariant L^2-inner product).

Proposition 9.9. *For $\varepsilon > 0$ sufficiently small, the map*

$$j : V := \{v \in E \mid |v| < \varepsilon\} \to \mathcal{J}, \qquad j(v) := j(\mathbb{1} + v)(\mathbb{1} - v)^{-1}$$

is a complex good family with $j(0) = j$.

Proof. The derivative $Dj(v) : E \to H^1\big(TS, j(v)\big)$ of j at $v \in V$ is given by

$$Dj(v)w = jw(\mathbb{1}{-}v)^{-1}{+}j(\mathbb{1}{+}v)(\mathbb{1}{-}v)^{-1}w(\mathbb{1}{-}v)^{-1} = \big(j{+}j(v)\big)w(\mathbb{1}{-}v)^{-1}.$$

It is complex linear because

$$Dj(v)jw = \big(j{+}j(v)\big)jw(\mathbb{1}{-}v)^{-1} = j(v)\big(j{+}j(v)\big)w(\mathbb{1}{-}v)^{-1} = j(v)Dj(v)w,$$

and since $j(0) = 0$ it is an isomorphism for ε sufficiently small, so the deformation is complex and effective. Symmetry follows from $j \circ Tg = Tg \circ j$ via

$$\begin{aligned} j(g_*v) &= j(\mathbb{1} + Tg \circ v \circ Tg^{-1})(\mathbb{1} - Tg \circ v \circ Tg^{-1})^{-1} \\ &= j \circ Tg \circ (\mathbb{1} + v) \circ Tg^{-1} \circ Tg \circ (\mathbb{1} - v) \circ Tg^{-1} \\ &= Tg \circ j(v) \circ Tg^{-1} = g_*j(v). \end{aligned}$$

□

Orbifold charts (uniformizers).

Proposition 9.10. *Let $E \supset V_0 \to \mathcal{J}$, $v \mapsto j(v)$ be a good family with $j(0) = j$. Then there exists a G-invariant open neighbourhood $V \subset \overline{V} \subset V_0$ of 0 such that:*
(a) The set $\mathcal{U}$ of isomorphism classes $[S, j(v)]$ for $v \in V$ is open in $\mathcal{M}_g$.
(b) The projection $p : V \to \mathcal{U}$, $v \mapsto [S, j(v)]$ induces a homeomorphism $V/G \to \mathcal{U}$.
(c) If for $v, w \in V$ there is an isomorphism $\phi : \big(S, j(v)\big) \to \big(S, j(w)\big)$, then $\phi \in G$ and $\phi_* v = w$.

Conditions (a) and (b) say that (V, G, p) is an *orbifold chart* (a *uniformizer* in the terminology of [21]) for $\mathcal{M}_g$ around $[j]$. Note that (b) implies that in the situaton of (c) we must have $w = g_* v$ for some $g \in G$, while (c) makes the stronger assertion that each such ϕ belongs to G and satisfies $\phi_* v = w$. This stronger property will be needed for compatibility between these orbifolds charts.

The proof of Proposition 9.10 comes down to an implicit function theorem which we now derive. For part (a), given $k \in \mathcal{J}$ sufficiently close to j we need to find $v \in V$ and an isomorphism $\phi : \big(S, j(v)\big) \stackrel{\cong}{\to} (S, k)$, i.e. such that $T\phi \circ j(v) = k \circ T\phi$. This condition is equivalent to vanishing of the section

$$F(k, v, \phi) := \frac{1}{2}\Big(T\phi + k \circ T\phi \circ j(v)\Big).$$

in the bundle $\mathcal{F} \to \mathcal{J} \times V_0 \times \mathcal{G}$ whose fibre over (k, v, ϕ) is the space of sections in the bundle $\mathrm{Hom}^{0,1}\big((TS, j(v)), (TS, k)\big)$. So we want to solve $F(k, v, \phi) = 0$ for (v, ϕ) as a function of k. For this we compute the derivative of F with respect to the last two variables on $w \in E$ and $\xi \in T_{\mathrm{id}}\mathcal{G} = \Omega^0(TS)$,

$$D_{23}F(j, 0, \mathrm{id})(w, \xi) = \frac{1}{2} j \circ Dj(0)w + \frac{1}{2}\Big(T\xi + j \circ T\xi \circ j\Big),$$

thus

$$D_{23}F(j,0,\mathrm{id}) = \Big(\frac{1}{2}j \circ Dj(0), \bar\partial_j\Big) : E \oplus \Omega^0(TS) \to \Omega^{0,1}(TS,j).$$

By effectiveness of the deformation $v \mapsto j(v)$ this is an isomorphism. In order to apply the implicit function theorem, we take a Sobolev completion $\mathcal{G}^s$ making it a Banach manifold. By the implicit function theorem on Banach manifolds (applied pointwise for $k \in \mathcal{J}$) there exist open neighbourhoods $\mathcal{K} \subset \mathcal{J}$ of j, $V \subset V_0$ of 0 and $\mathcal{H}^s \subset \mathcal{G}^s$ of id and a unique map $\Phi : \mathcal{K} \to V \times \mathcal{H}^s$ such that

$$F(k,v,\phi) = 0 \Longleftrightarrow (v,\phi) = \Phi(k) \quad \text{for } (k,v,\phi) \in \mathcal{K} \times V \times \mathcal{H}^s.$$

Applying the implicit function theorem for finite dimensional smooth families $W \mapsto \mathcal{K}$, we see that Φ is *smooth* in the sense that it depends smoothly on $w \in W$ for each such family. Moreover, $(k,v,\phi) = 0$ implies that ϕ is smooth, so we may replace $\mathcal{H}^s$ by its smooth part $\mathcal{H} \subset \mathcal{G}$ and drop the superscript s. We may choose the neighbourhoods such that V is G-invariant and the translated sets $\mathcal{H}g$ for $g \in G$ are pairwise disjoint. For $g \in G$ the symmetry property $j(g_*v) = Tg \circ j(v) \circ Tg^{-1}$ implies

$$\begin{aligned} F(k, g_*v, \phi \circ g^{-1}) &= \frac{1}{2}\Big(T\phi \circ Tg^{-1} + k \circ T\phi \circ Tg^{-1} \circ j(g_*v)\Big) \\ &= F(k,v,\phi) \circ Tg^{-1}. \end{aligned}$$

We define $\Phi_g : \mathcal{K} \to V \times \mathcal{H}g$ by

$$\Phi_g(k) := (g_*^{-1}v, \phi \circ g) \quad \text{for } \Phi(k) = (v,\phi).$$

Then $F\big(k, \Phi_g(k)\big) = 0$, and $\Phi(j) = (0,\mathrm{id})$ implies $\Phi_g(j) = (0,g)$, so we have shown the first two assertions in

Lemma 9.11 ([21], Theorem 3.15)**.** *Let $E \supset V_0 \to \mathcal{J}$, $v \mapsto j(v)$ be a good family with $j(0) = j$. Then there exist open*

neighbourhoods $\mathcal{K} \subset \mathcal{J}$ *of* j, $V \subset V_0$ *of* 0 *and* $\mathcal{H} \subset \mathcal{G}$ *of* id *with the following properties:*
(a) V *is* G*-invariant and the sets* $\mathcal{H}g$ *for* $g \in G$ *are pairwise disjoint.*
(b) For each $g \in G$ *there exists a unique map* $\Phi_g : \mathcal{K} \to V \times \mathcal{H}g$ *such that* $\Phi_g(j) = (0, g)$ *and*

$$F(k, v, \phi) = 0 \Longleftrightarrow (v, \phi) = \Phi_g(k) \quad \textit{for } (k, v, \phi) \in \mathcal{K} \times V \times \mathcal{H}g.$$

The map Φ_g *is smooth in the sense that it depends smoothly on* $w \in W$ *for each finite dimensional smooth family* $W \to \mathcal{K}$.
(c) For every solution $(k, v, \phi) \in \mathcal{K} \times V \times \mathcal{G}$ *of* $F(k, v, \phi) = 0$ *there exists a unique* $g \in G$ *such that* $\phi \in \mathcal{H}g$ *and* $(v, \phi) = \Phi_g(k)$.

Proof. It only remains to prove part (c). Let $\mathcal{K}, V, \mathcal{H}$ be neighbourhoods satisfying parts (a) and (b). Arguing by contradiction, suppose there exist sequences $k_n \to j$ and $v_n \to 0$ and isomorphisms $\phi_n : \big(S, j(v_n)\big) \to (S, k_n)$ staying outside the neighbourhood $\amalg_{g \in G} \mathcal{H}g$ of G. By properness of the action $\mathcal{G} \times \mathcal{J} \to \mathcal{J}$ (Proposition 9.7), a subsequence of ϕ_n converges to an isomorphism $\phi \in \mathrm{Aut}(S, j) = G$, contradicting the assumption. Hence there exist open neighbourhoods $\mathcal{K}' \subset \mathcal{K}$ and $V' \subset V$ such that for every solution $(k, v, \phi) \in \mathcal{K}' \times V' \times \mathcal{G}$ of $F(k, v, \phi) = 0$ there exists a unique $g \in G$ with $\phi \in \mathcal{H}g$. Part (b) yields $(v, \phi) = \Phi_g(k)$. Pick a neighbourhood $\mathcal{H}' \subset \mathcal{H}$ such that $\mathcal{K}', V', \mathcal{H}'$ satisfy again parts (a) and (b). Then $\Phi_g(k) \in \mathcal{H}'$, so part (c) holds for the new neighbourhoods $\mathcal{K}', V', \mathcal{H}'$. □

Proof of Proposition 9.10. Let $\Phi_g : \mathcal{K} \to V \times \mathcal{H}g$ be as in Lemma 9.11. Then $\mathcal{U}$ is the image of $\mathcal{K}$ under the projection $\mathcal{J} \to \mathcal{J}/\mathcal{G}$, hence open, so part (a) holds. For part (c), note first that uniqueness of Φ_g (or its construction) implies

$$\Phi_{\mathrm{id}}\big(j(v)\big) = (v, \mathrm{id}) \quad \text{and} \quad \Phi_g\big(j(v)\big) = (g_*^{-1}v, g).$$

Suppose now that for $v, w \in V$ there exists an isomorphism $\phi : \big(S, j(v)\big) \to \big(S, j(w)\big)$, so $F\big(j(w), v, \phi\big) = 0$. By Lemma 9.11(c) there exists a unique $g \in G$ such that $\phi \in \mathcal{H}g$ and $(v, \phi) = \Phi_g\big(j(w)\big) = (g_*^{-1}w, g)$, hence $\phi = g \in G$ and $\phi_* v = w$. This proves (c) as well as bijectivity of the map $p : V/G \to \mathcal{U}$ in (b). One readily verifies that $[S, k] \mapsto [v]$ where $\Phi_{\mathrm{id}}(k) = (v, \phi)$ defines a continuous inverse to p and Proposition 9.10 is proved. □

The orbifold structure on $\mathcal{M}_g$. Now we will use the uniformizers in Proposition 9.10 to define an orbifold structure on $\mathcal{M}_g$ in terms of an ep groupoid. For this, consider first two good families

$$j : H^1(TS, j) \supset V \to \mathcal{J}, \qquad k : H^1(TS, k) \supset W \to \mathcal{J}$$

as in Proposition 9.10 with $j(0) = j$ and $k(0) = k$. Define

$$\mathbf{T} := \big\{(v, \phi, w) \mid \phi : \big(S, j(v)\big) \xrightarrow{\cong} \big(S, j(v)\big)\big\} \subset V \times \mathcal{G} \times W$$

with its induced topology and the obvious projections $s : \mathbf{T} \to V$, $t : \mathbf{T} \to W$.

Corollary 9.12. *(a) The maps s, t are local homeomorphisms and the map $s \times t : \mathbf{T} \to V \times W$ is proper.*
(b) The transition maps $s \circ t^{-1}$, $t \circ s^{-1}$ (where they are defined) are smooth, and holomorphic if the two families are complex.

Proof. For part (a), consider a point $(v_0, \phi_0, w_0) \in \mathbf{T}$. Since $\phi_0^* w_0 = v_0$, we find an open neighbourhood $W_0 \subset W$ of w_0 such that $\phi_0^* k(w)$ is contained in the neighbourhood $\mathcal{K} \subset \mathcal{J}$ of j from Lemma 9.11 for all $w \in W_0$. Then the composition

$$w \mapsto \phi_0^* k(w) \overset{\Phi_{\mathrm{id}}}{\mapsto} (v_w, \psi_w)$$

with the map $\Phi_{1\!\!1}$ from Lemma 9.11 yields a smooth map $W_0 \to V \times \mathcal{K}$ with $\psi_w : \big(S, j(v_w)\big) \xrightarrow{\cong} \big(S, \phi_0^* k(w)\big)$. Then

$$\phi_w := \phi_0 \circ \psi_w : \big(S, j(v_w)\big) \xrightarrow{\cong} \big(S, k(w)\big),$$

so $(v_w, \phi_w, w) \in \mathbf{T}$ and by Lemma 9.11 the smooth map

$$W_0 \to \mathbf{T}, \qquad w \mapsto (v_w, \phi_w, w)$$

defines a local inverse to the target map t near (v_0, ϕ_0, w_0). This shows that t is a local homeomorphism, and similarly for s. Properness of $s \times t$ follows immediately from properness of the action $\mathcal{G} \times \mathcal{J} \to \mathcal{J}$.

For part (b), note that the local transition map $s \circ t^{-1}(v) = w_v$ is smooth by part (a), and it is uniquely defined by the condition $j(v_w) = \phi_w^* k(w)$. Taking derivatives we find $Dj(v_w) \circ D(s \circ t^{-1})(w) = \phi_0^* Dk(w)$, so passing to quotients the following diagram commutes:

$$\begin{tikzcd}
F \arrow[r, "{[Dk(w)]}"] \arrow[d, "D(s\circ t^{-1})(w)"'] & H^1\big(TS, k(w)\big) \arrow[d, "\phi_w^*"] \\
E \arrow[r, "{[Dj(v_w)]}"] & H^1\big(TS, j(v_w)\big).
\end{tikzcd}$$

Note that all the maps in this diagram are isomorphisms and ϕ_w^* is $\mathbb{C}$-linear. If the families $j(v)$, $k(w)$ are complex, then their Kodaira differentials are $\mathbb{C}$-linear, hence so is $D(s \circ t^{-1}(w)$ and the transition map is holomorphic. $\square$

Now we pick a countable collection of complex good families $E_i \supset V_i \to \mathcal{J}$, $i \in \mathbb{N}$, whose images cover $\mathcal{M}_g$. For each $i, j \in \mathbb{N}$ define $\mathbf{T}_{i,j} \to V_i \times V_j$ as above. By Corollary 9.12 the groupoid

$$\mathbf{X} := \coprod_{i,j \in \mathbb{N}} \mathbf{T}_{i,j} \xrightarrow{s,t} X := \coprod_{i \in \mathbb{N}} V_i$$

is étale and proper and defines the desired orbifold structure on $\mathcal{M}_g$. This concludes the proof of Theorem 9.3.

9.4 Deligne–Mumford spaces

In this section we will indicate how to extend the results of the previous section to marked nodal Riemann surfaces.

Definition 9.10. A *(marked) nodal Riemann surface* $\alpha = (S, j, M, D)$ consists of

- a possibly disconnected closed Riemann surface (S, j);
- a finite ordered set $M \subset S$ of marked points;
- a finite unordered set D of unordered pairs $\{x, y\} \in S \times S$

such that all points in $|D| \cup M$ are distinct, where $|D|$ denotes the union of all points from pairs in D, so $\#|D| = 2\#D$. We assume that the associated nodal surface $\overline{S}$, obtained by identifying $x \sim y$ for each pair $\{x, y\} \in D$, is connected. Moreover, we assume the *stability condition*

$$2g_C + \#M_C \geq 3$$

for each component C of S, where g_C is the genus of C and M_C denotes the set of *special* (i.e., marked or nodal) points on C. The *arithmetic genus* of α is

$$g_\alpha = 1 + \sum_C (g_C - 1) + \#D.$$

An *isomorphism* $\phi : (S, j, M, D) \to (S', j', M', D')$ of nodal Riemann surfaces is an isomorphism $\phi : (S, j) \to (S', j')$ sending M to M' (preserving the ordering) and D to D'. For $2g + m \geq 3$, the *Deligne–Mumford space* $\overline{\mathcal{M}}_{g,m}$ is the set of isomorphism classes of nodal Riemann surfaces of arithmetic genus g with m marked points.

Note that the stability condition reads $\chi(C \setminus M_C) = 2 - 2g_C - \#M_C < 0$, which is equivalent to the automorphism group of C (preserving M_C as a set) being finite for each component C. The arithmetic genus is the genus of the smooth surface obtained by resolving the double points in $\overline{S}$, as seen from the Euler characteristic computation (since each connected sum decreases the Euler characteristic by 2)

$$2 - 2g_\alpha = \sum_C (2 - 2g_C) - 2\#D.$$

The Deligne–Mumford space $\overline{\mathcal{M}}_{g,m}$ is compact with respect to the *Gromov topology*, and the moduli space of *smooth* (i.e., without nodes) marked Riemann surfaces $\mathcal{M}_{g,m} \subset \overline{\mathcal{M}}_{g,m}$ is open and dense. Note that $\mathcal{M}_{g,0} = \mathcal{M}_g$ for $g \geq 2$.

The main result of this section is the following extension of Theorem 9.3.

Theorem 9.13. *For $2g + m \geq 3$, the Deligne–Mumford space $\overline{\mathcal{M}}_{g,m}$ is a complex orbifold of complex dimension $3g - 3 + m$.*

Problem 9.16. Show: (a) $\overline{\mathcal{M}}_{0,3} = \{\mathrm{pt}\}$ and $\overline{\mathcal{M}}_{0,4} \cong \mathbb{C}P^1$;
Check. (b) $\overline{\mathcal{M}}_{1,1} \cong (\mathbb{H}/PSL(2,\mathbb{Z})) \cup \{\infty\}$ equals $\mathbb{C}P^1$ with two orbifold singularities;
(c) $\overline{\mathcal{M}}_{0,5} \cong \mathbb{C}P^2 \# 4\overline{\mathbb{C}P}^2$ is the blow-up of $\mathbb{C}P^2$ at 4 points. *Hint: Consider the space Q of quadrics through 4 generic points $P_1, \ldots, P_4 \in \mathbb{C}P^2$. Then $Q \cong \mathbb{C}P^1$ consists of smooth quadrics and 3 pairs of lines, so $\mathrm{Bl}_{P_1,\ldots,P_4}\mathbb{C}P^2 \to Q$ is a Lefschetz fibration with generic fibre $\mathbb{C}P^1$ and 3 nodal fibres, which realizes the map $\pi : \overline{\mathcal{M}}_{0,5} \to \overline{\mathcal{M}}_{0,4} \cong \mathbb{C}P^1$ forgetting the last marked point.*

Orbifold structure with fixed nodal type. As preparation for the proof of Theorem 9.13, let us first fix a *nodal type* $\tau = (S, M, D)$ with arithmetic genus g and $m = \#M$ marked points. Let $\mathcal{G}_\tau := \mathrm{Diff}_+(S, M, D)$ be the group of orientation preserving diffeomorphisms of S fixing M (with its ordering) and D. It acts on the space $\mathcal{J}$ of positively oriented almost complex structures on S and its quotient $\mathcal{M}_\tau := \mathcal{J}/\mathcal{G}_\tau$ is the *moduli space of nodal Riemann surfaces of type* τ. The linearized action at $\alpha = (S, j, M, D)$ is given by

$$\bar{\partial}_j : \Omega^0(TS, M \cup |D|) \to \Omega^{0,1}(TS, j),$$

where $\Omega^0(TS, M \cup |D|)$ is the space of smooth vector fields on S vanishing at $M \cup |D|$. It follows again from the Kodaira vanishing theorem that $\ker \bar{\partial}_j = \{0\}$, and from the Riemann–Roch theorem (applied on each component C) that $H^1(\alpha) := \Omega^{0,1}(TS, j)/\mathrm{im}\, \bar{\partial}_j$ is a complex vector space of dimension

$$\begin{aligned}
\dim_{\mathbb{C}} H^1(\alpha) &= \sum_C (3g_C - 3 + \#M_C) \\
&= 3\sum_C (g_C - 1) + \#M + 2\#D \\
&= 3(g_\alpha - 1 - \#D) + m + 2\#D \\
&= 3g - 3 + m - \#D.
\end{aligned}$$

As before, we find a G-invariant complex subspace $E \subset \Omega^{0,1}(TS, j)$ such that the projection $\Omega^{0,1}(TS, j) \to H^1(\alpha)$ restricts to an isomorphism $E \xrightarrow{\cong} H^1(\alpha)$, and (with a *complex good family* defined as before) the proof of Proposition 9.9 carries over to give

Proposition 9.14. *For $\varepsilon > 0$ sufficiently small, the map*

$$j : V := \{v \in E \mid |v| < \varepsilon\} \to \mathcal{J}, \qquad j(v) := j(\mathbb{1} + v)(\mathbb{1} - v)^{-1}$$

is a complex good family on (S, M, D) with $j(0) = j$. $\square$

As in the previous section, the complex good families give rise to an orbifold structure and we obtain

Theorem 9.15. *The moduli space $\mathcal{M}_\tau$ of nodal Riemann surfaces of type τ is a complex orbifold of complex dimension $3g-3+m-\#D$.* $\square$

Small disk structures. Consider a nodal surface $\alpha = (S, j, M, D)$ with $G := \mathrm{Aut}(\alpha)$ the finite group of automorphisms of (S, j) preserving M (with its ordering) and D.

Definition 9.11. A *small disk structure* on α is a collection $\{D_x\}_{x\in|D|}$ of pairwise disjoint embedded disks $D_x \subset S\backslash M$, one for each nodal point x, such that their union $\amalg_{x\in|D|}D_x$ is G-invariant.

For the proof of the following result see [21].

Proposition 9.16 (small disk structures and good families)**.**
(a) For each open neighbourhood $U \subset S$ of $|D|$ there exists a small disk structure on α with $\amalg_{x\in|D|}D_x \subset U$.
(b) For each small disk structure $\{D_x\}_{x\in|D|}$ on α there exists a complex good family $V \ni v \mapsto j(v)$ with $j(0) = j$ such that $j(v) = j$ on $\amalg_{x\in|D|}D_x$ for all $v \in V$.

Let us fix a small disk structure and biholomorphic maps

$$\overline{h}_x : (\mathbb{D}, 0) \xrightarrow{\cong} (D_x, x), \qquad x \in |D|,$$

where $\mathbb{D} \subset \mathbb{C}$ denotes the closed unit disk. For an unordered pair $\{x, y\} \in D$ we pick an ordering (x, y) and introduce positive/negative holomorphic polar coordinates (where $S^1 := \mathbb{R}/\mathbb{Z}$)

$$h_x(s,t) := \overline{h}_x(e^{-2\pi(s+it)}), \qquad (s,t) \in [0,\infty) \times S^1,$$
$$h_y(s',t') := \overline{h}_y(e^{2\pi(s'+it')}), \qquad (s',t') \in (-\infty,0] \times S^1.$$

By G-invariance of a small disk structure, and automorphism $g \in G$ induces biholomorphisms

$$g : D_x \xrightarrow{\cong} D_{g(x)}, \qquad g : D_y \xrightarrow{\cong} D_{g(y)}.$$

Since the only automorphisms of $\mathbb{D}$ fixing the origin are rotations, it follows that the representations of these maps in holomorphic polar coordinates take the form

$$\begin{aligned} g^+(s,t) &:= h_{g(x)}^{-1} \circ g \circ h_x(s,t) = (s, t + \vartheta^+), \\ g^-(s',t') &:= h_{g(y)}^{-1} \circ g \circ h_y(s',t') = (s', t' + \vartheta^-) \end{aligned} \tag{9.2}$$

for $\vartheta^\pm = \vartheta^\pm_{(x,y)} \in S^1$.

Gluing at nodes. Now we introduce the space of *gluing parameters*

$$N := \bigoplus_{\{x,y\} \in D} \mathbb{C},$$

which we write as tuples $a = \{a_{\{x,y\}}\}$ of complex numbers associated to the nodes. Automorphisms $g \in G$ act on N by

$$g_* \{a_{\{x,y\}}\} := \{e^{-2\pi i (\vartheta^+_{(x,y)} - \vartheta^-_{(x,y)})} a_{\{x,y\}}\}.$$

Note that this action preserves the subset

$$B := \bigoplus_{\{x,y\} \in D} \mathbb{D} \subset N.$$

Problem 9.17. Show that interchanging the roles of x, y at a node does not change the action of G on N.

We fix a *gluing profile*

$$\varphi : (0,1] \xrightarrow{\cong} [0, \infty)$$

and associate to each gluing parameter $a_{\{x,y\}} \in \mathbb{D}$ a set $Z^{\{x,y\}}_{a_{\{x,y\}}}$ as follows. For $a_{\{x,y\}} = 0$ we set $Z^{\{x,y\}}_0 := D_x \amalg D_y$. For $a_{\{x,y\}} \neq 0$

we write

$$a_{\{x,y\}} = re^{-2\pi i\vartheta}, \qquad R = \varphi(r)$$

and define

$$Z^{\{x,y\}}_{a_{\{x,y\}}} := [0,R] \times S^1 \cong [-R,0] \times S^1,$$

where $(s,t) \in [0,R] \times S^1$ is identified with $(s',t') \in [-R,0] \times S^1$ via the relation

$$s - s' = R, \qquad t - t' = \vartheta.$$

Note that the chosen holomorphic polar coordinates yield canonical holomorphic embeddings

$$h_x : Z^{\{x,y\}}_{a_{\{x,y\}}} \hookrightarrow D_x, \qquad h_y : Z^{\{x,y\}}_{a_{\{x,y\}}} \hookrightarrow D_y.$$

Cutting out $h_x\big((R,\infty) \times S^1\big)$ from D_x and $h_y\big((-\infty,-R) \times S^1\big)$ from D_y and gluing the remaining cylinders via $Z^{\{x,y\}}_{a_{\{x,y\}}}$ at each node, we thus obtain for each $a \in B$ a *(partially) glued nodal surface* (S_a, j_a, M_a, D_a). Here the nodes $\{x,y\}$ with $a_{\{x,y\}} \neq 0$ get replaced by smooth cyinders, while the nodes with $a_{\{x,y\}} = 0$ remain nodes and form the new set of nodes D_a. The new set of marked points is just $M_a = M$, which lies outside $\amalg_{x\in|D|} D_x$ and is thus not affected by the gluing construction.

Problem 9.18. Show that interchanging the roles of x, y at a node leads to a glued surface which is canonically isomorphic to (S_a, j_a, M_a, D_a).

Consider now a good family $V \ni v \mapsto j(v)$ as in Proposition 9.16. Since $j(v) = j$ on $\amalg_{x\in|D|} D_x$, it does not interact with the preceding gluing operation and we obtain a deformation of nodal Riemann surfaces

$$B \times V \ni (a,v) \mapsto \big(S_a, j(a,v), M_a, D_a\big).$$

This deformation is again symmetric:

Lemma 9.17. *Each* $g \in G = \operatorname{Aut}(\alpha)$ *induces for all* $(a, v) \in B \times V$ *canonical isomorphisms*

$$g_a : \big(S_a, j(a,v), M_a, D_a\big) \xrightarrow{\cong} \big(S_{g(a)}, j(g_*a, g_*v), M_{g(a)}, D_{g(a)}\big).$$

Proof. Since the deformation $v \mapsto j(v)$ is symmetric, g induces for each $v \in V$ an isomorphism

$$g : \big(S, j(v), M, D\big) \xrightarrow{\cong} \big(S, j(g_*v), M, D\big).$$

Since $\amalg_{x\in|D|} D_x$ is G-invariant, g restricts for each $x \in |D|$ to a biholomorphism $g : D_x \xrightarrow{\cong} D_{g(x)}$. For $a \in B$ we modify g to the desired biholomorphism $g_a : S_a \to S_{g(a)}$ as follows. We set $g_a := g : S \setminus \amalg_{x\in|D|} D_x \to S \setminus \amalg_{x\in|D|} D_x$. For a node $\{x, y\}$ with $a_{\{x,y\}} = 0$ we set $g_a := g : D_x \amalg D_y \to D_{g(x)} \amalg D_{g(y)}$. If $a_{\{x,y\}} \neq 0$ we define $g_a : Z^{\{x,y\}}_{a_{\{x,y\}}} \to Z^{\{g(x),g(y)\}}_{(g_*a)_{\{g(x),g(y)\}}}$ by

$$\begin{gathered}[0, R] \times S^1 \ni (s, t) \mapsto (s, t + \vartheta^+) \in [0, R] \times S^1 \text{ resp.}\\ [-R, 0] \times S^1 \ni (s', t') \mapsto (s', t' + \vartheta^-) \in [-R, 0] \times S^1.\end{gathered}$$

This matches the representations (9.2) of g in holomorphic polar coordinates and thus fits together holomorphically with $g|_{S\setminus(D_x\amalg D_y)}$. To see that it is well-defined, recall that the positive/negative polar coordinates are related by $s - s' = R$ and $t - t' = \vartheta$, where $a_{\{x,y\}} = re^{-2\pi i\vartheta}$ and $R = \varphi(r)$. Then their images are related by $s - s' = R$ and $(t + \vartheta^+) - (t' + \vartheta^-) = \vartheta + \vartheta^+ - \vartheta^-$, which is exactly the identification in the cylinder $Z^{\{g(x),g(y)\}}_{(g_*a)_{\{g(x),g(y)\}}}$ because

$$g_*a = e^{-2\pi i(\vartheta^+ - \vartheta^-)} a = re^{-2\pi i(\vartheta + \vartheta^+ - \vartheta^-)}.$$

$\square$

As in the previous section one uses the deformations $(a, v) \mapsto j(a, v)$, with an arbitrary gluing profile φ, to construct a smooth orbifold structure on $\overline{\mathcal{M}}_{g,m}$. Note, however, that this is more difficult in the presence of nodes due to the varying domains S_a; see [21].

Finally, it is shown in [21] that the *logarithmic gluing profile*

$$\varphi(r) = -\frac{1}{2\pi}\log(r)$$

equips $\overline{\mathcal{M}}_{g,m}$ with the structure of a *complex* orbifold. The emergence of this gluuing profile can be understood from the following *local model for a node.* Consider the holomorphic family of curves

$$C_a := \{(z,w) \in \mathbb{C}^2 \mid zw = a,\ \max\{|z|,|w|\} \le 1\}$$

parametrized by $a \in \mathbb{C}$ with $|a| < 1$. Then C_a is a cylinder for $a \neq 0$, while C_0 is the union of two unit disks meeting at the origin, i.e. a node. For $a \neq 0$ we can solve for $w = a/z$, so $|w| = |a|/|z| \le 1$ and thus

$$C_a \cong \{z \in \mathbb{C} \mid |a| \le |z| \le 1\} \cong \left[0, -\frac{1}{2\pi}\log|a|\right] \times \mathbb{R}/\mathbb{Z}$$

using holomorphic polar coordinates $z = e^{-2\pi(s+it)}$. We see that the modulus of the cyinder C_a is given by the logarithmic gluing profile $R = \varphi(|a|) = -\frac{1}{2\pi}\log|a|$.

9.5 Gromov–Witten invariants

9.6 G-moduli problems

Chapter 10

Polyfolds with operations

Topics: operads

Applications: symplectic field theory

Bibliography

[1] A. Abbondandolo and T. O. Rot. On the homotopy classification of proper Fredholm maps into a Hilbert space. *arXiv:1610.07039v2*, 2017.

[2] A. Adem, J. Leida, and Y. Ruan. *Orbifolds and stringy topology*, volume 171 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2007.

[3] F. Albiac and N. J. Kalton. *Topics in Banach space theory*, volume 233 of *Graduate Texts in Mathematics*. Springer, New York, 2006.

[4] H. W. Alt. *Linear functional analysis*. Universitext. Springer-Verlag London, Ltd., London, 2016. An application-oriented introduction, Translated from the German edition by Robert Nürnberg.

[5] M. Audin and M. Damian. *Morse theory and Floer homology*. Universitext. Springer, London; EDP Sciences, Les Ulis, 2014. Translated from the 2010 French original by Reinie Erné.

[6] R. Bott. Lectures on Morse theory, old and new. *Bull. Amer. Math. Soc. (N.S.)*, 7(2):331–358, 1982.

[7] R. Caccioppoli. Sulle corrispondenze funzionali inverse diramate: teoria generale e applicazioni ad alcune equazioni fun-

zionali nonlineari e al problema di Plateau, I, II. *Rend. Accad. Naz. Lincei,*, 24:258–263, 416–421, 1936.

[8] H. Cartan. Sur les rétractions d'une variété. *C. R. Acad. Sci. Paris Sér. I Math.*, 303(14):715, 1986.

[9] K. Cieliebak. *Pseudo-holomorphe Kurven und periodische Orbits auf Cotangential-Bündeln.* Diplomarbeit, Ruhruniversität Bochum. 1992.

[10] K. Cieliebak. Pseudo-holomorphic curves and periodic orbits on cotangent bundles. *J. Math. Pures Appl. (9)*, 73(3):251–278, 1994.

[11] K. Cieliebak, I. Mundet i Riera, and D. A. Salamon. Equivariant moduli problems, branched manifolds, and the Euler class. *Topology*, 42(3):641–700, 2003.

[12] K. Cieliebak and E. Séré. Pseudoholomorphic curves and the shadowing lemma. *Duke Math. J.*, 99(1):41–73, 1999.

[13] J. B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1985.

[14] K. Deimling. *Nonlinear Functional Analysis*. Springer-Verlag, Berlin, 1985.

[15] J. Eells, Jr. A setting for global analysis. *Bull. Amer. Math. Soc.*, 72:751–807, 1966.

[16] H. I. Eliasson. Geometry of manifolds of maps. *J. Differential Geometry*, 1:169–194, 1967.

[17] R. Fry and S. McManus. Smooth bump functions and the geometry of Banach spaces: a brief survey. *Expo. Math.*, 20(2):143–183, 2002.

[18] K. Geba and A. Granas. A proof of the Borsuk antipodal theorem for Fredholm maps. *J. Math. Anal. Appl.*, 96(1):196–202, 1983.

[19] F. Hirzebruch and W. Scharlau. *Einführung in die Funktionalanalysis.* Bibliographisches Institut, Mannheim-Vienna-Zurich, 1971. B. I.-Hochschultaschenbücher, No. 296*.

[20] H. Hofer. Lusternik-Schnirelman-theory for Lagrangian intersections. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 5(5):465–499, 1988.

[21] H. Hofer, K. Wysocki, and E. Zehnder. Deligne–mumford–type spaces with a view towards symplectic field theory. *private communication.*

[22] H. Hofer, K. Wysocki, and E. Zehnder. A general Fredholm theory. III. Fredholm functors and polyfolds. *Geom. Topol.*, 13(4):2279–2387, 2009.

[23] H. Hofer, K. Wysocki, and E. Zehnder. Integration theory on the zero sets of polyfold Fredholm sections. *Math. Ann.*, 346(1):139–198, 2010.

[24] H. Hofer, K. Wysocki, and E. Zehnder. sc-smoothness, retractions and new models for smooth spaces. *Discrete Contin. Dyn. Syst.*, 28(2):665–788, 2010.

[25] H. Hofer, K. Wysocki, and E. Zehnder. Applications of polyfold theory I: The polyfolds of Gromov-Witten theory. *Mem. Amer. Math. Soc.*, 248(1179):v+218, 2017.

[26] H. Hofer, K. Wysocki, and E. Zehnder. Polyfold and Fredholm Theory. *arXiv:1707.08941*, 2017.

[27] H. H. W. Hofer. Polyfolds and Fredholm theory. In *Lectures on geometry*, Clay Lect. Notes, pages 87–158. Oxford Univ. Press, Oxford, 2017.

[28] M. I. Kadec. A proof of the topological equivalence of all separable infinite-dimensional Banach spaces. *Funkcional. Anal. i Priložen.*, 1:61–70, 1967.

[29] V. L. Klee, Jr. Convex bodies and periodic homeomorphisms in Hilbert space. *Trans. Amer. Math. Soc.*, 74:10–43, 1953.

[30] N. H. Kuiper. The homotopy type of the unitary group of Hilbert space. *Topology*, 3:19–30, 1965.

[31] I. Kupka. Counterexample to the Morse-Sard theorem in the case of infinite-dimensional manifolds. *Proc. Amer. Math. Soc.*, 16:954–957, 1965.

[32] S. Lang. *Fundamentals of differential geometry*, volume 191 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999.

[33] J. H. McAlpin. *Infinite dimensional Manifolds and Morse Theory*. ProQuest LLC, Ann Arbor, MI, 1965. Thesis (Ph.D.)–Columbia University.

[34] D. McDuff and D. Salamon. *J-holomorphic curves and symplectic topology*, volume 52 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[35] D. McDuff and D. Salamon. *Introduction to symplectic topology*. Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, third edition, 2017.

[36] I. Moerdijk. Orbifolds as groupoids: an introduction. In *Orbifolds in mathematics and physics (Madison, WI, 2001)*, volume 310 of *Contemp. Math.*, pages 205–222. Amer. Math. Soc., Providence, RI, 2002.

[37] I. Moerdijk and D. A. Pronk. Orbifolds, sheaves and groupoids. *K-Theory*, 12(1):3–21, 1997.

[38] T. Mrowka. *Geometry of Manifolds*. Lecture Notes. https://ocw.mit.edu/courses/mathematics/18-965-geometry-of-manifolds-fall-2004/lecture-notes/, 2004.

[39] R. S. Palais. *Foundations of global non-linear analysis*. W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[40] J. Robbin and D. Salamon. The spectral flow and the Maslov index. *Bull. London Math. Soc.*, 27(1):1–33, 1995.

[41] J. W. Robbin and D. A. Salamon. A construction of the Deligne-Mumford orbifold. *J. Eur. Math. Soc. (JEMS)*, 8(4):611–699, 2006.

[42] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[43] D. Salamon and E. Zehnder. Morse theory for periodic solutions of Hamiltonian systems and the Maslov index. *Comm. Pure Appl. Math.*, 45(10):1303–1360, 1992.

[44] A. Sard. The measure of the critical values of differentiable maps. *Bull. Amer. Math. Soc.*, 48:883–890, 1942.

[45] B. Schmidt. *Partielle Differentialgleichungen*. Vorlesungsskript. https://www.math.uni-

augsburg.de/prof/ana/arbeitsgruppe/schmidt/skripten/pdg1.pdf, 2012.

[46] P. Seidel. *Geometry of Manifolds*. Lecture Notes. https://math.mit.edu/ ebelmont/965-notes.pdf, 2013.

[47] S. Smale. An infinite dimensional version of Sard's theorem. *Amer. J. Math.*, 87:861–866, 1965.

[48] E. H. Spanier. *Algebraic topology*. McGraw-Hill Book Co., New York-Toronto, Ont.-London, 1966.

[49] E. H. Spanier. Tautness for Alexander-Spanier cohomology. *Pacific J. Math.*, 75(2):561–563, 1978.

[50] A. J. Tromba. Some theorems on Fredholm maps. *Proc. Amer. Math. Soc.*, 34:578–585, 1972.

[51] K. Wehrheim. Fredholm notions in scale calculus and hamiltonian floer theory. *arXiv:1209.4040*, 2012.