Nonlinear Analysis

Martin Heida

Extended Version
partially inspired by the lecture of Prof. Willi Jäger in 2006/07

February 19, 2020

Contents

Introduction

The lecture on nonlinear functional analysis has no canonic structure. It is thought to give the students at hand a "toolbox" of mathematical methods for the analysis of nonlinear equations in Banach spaces. These nonlinear problems typically arise from the field of mathematical modeling and often are given in the form of partial differential equations. We discuss stationary and dynamic problems.

In the part on stationary problems we discuss nonlinear equations

$$f(u,\lambda)=0\,, \tag{1}$$

where $f:\,\mathscr{X}\times(r,s)\to\mathscr{Y}$ for two Banach spaces $\mathscr{X}$ and $\mathscr{Y}$.

We will discuss this topic in three steps which related to a proceeding understanding of the functional f and hence form a natural hierarchy:

1. For given $\lambda=\lambda_0$ prove the existence of solutions to (1).
2. Suppose the existence of solutions is shown, i.e. $f(u_0,\lambda_0)=0$:
   (a) Are there further solutions $f(\tilde{u},\lambda_0)=0$ in a vicinity of u_0? What is the dimension of the space of solutions?
   (b) Are there solutions to
   $$f(u_v,\lambda_0)=v\,,$$
   and how are they related to u_0?
3. How does the space of solutions change if $\lambda\neq\lambda_0$?

In case $\mathscr{X}=\mathbb{R}^d$ and $\mathscr{Y}=\mathbb{R}^m$, these questions have partially been answered in the undergraduate courses Analysis I – III. To name these partial answers we remind some keywords:

1. Concerning the existence of solutions:
   (a) If $d=m=1$ we can use the *Intermediate value theorem.*
   (b) If $d=m=2$ we can use the *Residual theorem* from complex analysis.
   (c) For general $d=m$ we can us fixed point theorems.
2. Question number 2. can be answered by
   (a) The implicit function theorem
   (b) The inverse function theorem
3. Question number 3 is "new", in a sense that there is no equivalent in undergraduate courses.

In the second part of the course, we discuss time dependent problems.

$$\dot{u} = f(u, \lambda) \,.$$

We will see that periodic solutions may branch of from the stationary solutions, a phenomenon which is called Hopf bifurcation. We furthermore discuss equations of the more particular structure

$$\dot{u} = \partial\Psi^* \left(-D\mathcal{E}(u)\right) \,,$$

where Ψ^* is convex and $\mathcal{E}$ is a coercive functional. The later type of equations is called generalized gradient flow. Furthermore we discuss equations of monotone type

$$\dot{u} = g(u) \,,$$

where g is a monotone function.

Some of the results will not be proved in the lecture but at least partially provided in this script here. In order to emphasize the range of applications, we will provide some exercises including solutions.

Let us finally comment on literature. The roots of this lecture are probably in the beginning of the 80's. Including Section 3.1, the first two third of this notes follow the lecture of Prof. Willi Jäger which he gave in 2006/2007 at the University of Heidelberg, with some modifications according to my own taste. There is also a german lecture note by Prof. Ben Schweizer which is very close to this script. Accordingly I found that many ideas concerning the proofs in the part on fixed point theory, degreee and bifurcation theory are provided in the books [7, 14]. Concerning the part on the implicit function theorem, I also highlight [5]. A more recent book in german is [12]. The part on gradient flows is mostly inspired by the recent work [11] which is restricted to Hilbert spaces in comparison with the more abstract book [1]. Finally, the part on monotone operators and convex analysis is in large parts taken from [4, 8, 9, 3].

Berlin/München 2019/2020

Chapter 1

The degree

1.1 Definition and Properties of the Degree for continuous nonlinear functions

1.1.1 Introduction

We recall the following theorem from Analysis I.

Theorem. *Let $f : [-1,1] \to \mathbb{R}$ be continuous. If $f(-1) * f(1) < 0$, then there exists at least one solution $x_0 \in (-1,1)$ to the problem $f(x_0) = 0$.*

On the other hand, if $f(-1) * f(1) > 0$ we are not in the position to draw any conclusion unless we know something about f on the inner of $(-1,1)$. In order to generalize this theorem to higher dimension, it is useful to have a geometrical interpretation in mind. Throughout these lecture notes, we call $\mathbb{B}_1^d(0) := \left\{x \in \mathbb{R}^d \mid |x| < 1\right\}$ the unit ball and $\mathbb{S}^{d-1} = \partial\mathbb{B}^d$ the sphere in $\mathbb{R}^d$. Furthermore, we use the following two definitions.

Definition 1.1.1. A map $\varphi : \mathbb{S}^{d-1} \to \mathbb{R}^m \backslash \{0\}$ is called essential if every continuous extension to $\mathbb{B}_1^d(0)$ possesses a zero.

Definition 1.1.2. Let T_1, T_2 be topological spaces and $f, g \in C(T_1; T_2)$. We say that f and g are homotope, written $f \sim g$ if there exists $h \in C(T_1 \times [0,1]; T_2)$ such that $h(\cdot, 0) = f$ and $h(\cdot, 1) = g$. A function $g \in C(T_1; T_2)$ is null-homotope if $g \sim t$ for some constant $t \in T_2$.

Remark 1.1.3. Being homotope is an equivalence relation.

We can now reformulate the above theorem as follows:

Theorem. *Let $\varphi : \{-1,1\} \to \mathbb{R}\backslash\{0\}$. Then, φ is essential if and only if $\psi = \frac{\varphi}{|\varphi|}$ as a function $\{-1,1\} \to \{-1,1\}$ is not null-homotope.*

Exercise 1.1.4. Prove that this is equivalent to the intermediate value theorem.

We can formulate this last result more generally.

Theorem 1.1.5 (Intermediate value theorem)**.** *Let $\varphi : \mathbb{S}^{d-1} \to \mathbb{R}^m \backslash \{0\}$ be continuous with $\psi = \frac{\varphi}{|\varphi|}$. Then φ is essential if and only if $\psi : \mathbb{S}^{d-1} \to \mathbb{S}^{m-1}$ is not null-homotope.*

Proof. Step 1: If φ is not essential, then there exists a continuation $f:\mathbb{B}^d\to\mathbb{R}^m\backslash\{0\}$ and ψ is homotope with $\frac{f(0)}{|f(0)|}$ through $h(x,t)=\frac{f(tx)}{|f(tx)|}$.

Step 2: Now, assume that ψ is homotope with $y_0\in\mathbb{S}^{m-1}$ through a function $h\in C(\mathbb{S}^{d-1}\times[0,1];\mathbb{S}^{m-1})$ with $h(x,1)\equiv y_0\in\mathbb{S}^{m-1}$. We define

$$f(x)=\begin{cases}\left(1-|x|+|x|\left|\varphi\left(\frac{x}{|x|}\right)\right|\right)h\left(\frac{x}{|x|},1-|x|\right) & \text{if } x\neq 0\\ y_0 & \text{if } x=0\end{cases}$$

and prove that f is a continuous extension of φ and has no zeros. This implies that φ is not essential. First note that

$$f|_{\mathbb{S}^{d-1}}(x)=|\varphi|(x)\,h(x,0)=|\varphi|\frac{\varphi}{|\varphi|}=|\varphi|(x)\quad\text{since }|x|=1\,.$$

Furthermore, f is continuous in 0. To see this, let $x_k\to 0$ and note that $\frac{x_k}{|x_k|}\to\xi$ along a subsequence. Hence

$$h\left(\frac{x_k}{|x_k|},1-|x_k|\right)\to h(\xi,1)=y_0$$

and we find $f(x_k)\to y_0$. It remains to show that f has no zeros. We use $|h|$=1 and observe $f(x)=0$ iff

$$1=|x|\left(1-\left|\varphi\left(\frac{x}{|x|}\right)\right|\right)<1$$

as $\varphi\neq 0$. Hence φ is not essential. □

It is desirable to not only know about the existence of zeros but also their multiplicity. In this context, we are inspired by the residual theorem from complex analysis.

Theorem 1.1.6. *Let $f:\mathbb{B}^2_1(0)\to\mathbb{C}$ be analytical and $0\notin f(\mathbb{S}^1)$. Let $z_1,\dots z_k$ be an enumeration of the zeros of f and let $\upsilon_f(z_i)$ be the multiplicity of f in z_i, $i=1,\dots k$. Then*

$$\frac{1}{2\pi i}\oint\frac{f'(z)}{f(z)}\mathrm{d}z=\sum_i\upsilon_f(z_i)\,. \tag{1.1}$$

We make use of the representation $z=x+iy$ and $f(z)=f_r(z)+if_i(z)$ where the Cauchy-Riemannian differential equations imply $\partial_x f_r=\partial_y f_i$ and $\partial_y f_r=-\partial_x f_i$ and $f'(z)=\partial_x f_r(z)+i\partial_x f_i(z)$. Furthermore, we find that $z=\nu$ is the outer normal of $\mathbb{B}^2_1(0)$. Hence we have by use of $z=e^{i\varphi}$ that $\mathrm{d}z=iz\mathrm{d}\varphi$ and

$$\begin{aligned}\frac{1}{2\pi i}\oint\frac{f'(z)}{f(z)}\mathrm{d}z&=\frac{1}{2\pi}\int_0^{2\pi}\frac{f'(z)}{f(z)}z\mathrm{d}\varphi=\frac{1}{2\pi}\int_0^{2\pi}\frac{\overline{f}(z)f'(z)z}{|f(z)|^2}\mathrm{d}\varphi\\&=\frac{1}{2\pi}\int_0^{2\pi}\mathrm{Re}\frac{\overline{f}(z)f'(z)z}{|f(z)|^2}\mathrm{d}\varphi\\&=\frac{1}{2\pi}\int_0^{2\pi}\frac{\langle f\,,\,\mathrm{D}f\,\nu\rangle_{\mathbb{R}^2}}{|f|^2}\mathrm{d}\varphi\,,\end{aligned}\tag{1.2}$$

where in the last line we switched from $\mathbb{C}$ to $\mathbb{R}^2$ and hence $\mathrm{D}fz \rightsquigarrow \nu$ the normal vector of $\mathbb{S}^1$. The last formula suggests a straight generalization to $\mathbb{R}^d$. However, we need to be careful. First of all, we only dealt with analytic (complex differentiable) functions. Second, the Cauchy-Riemannian equations show that in two dimensions and for analytical functions $\mathrm{D}f$ is equivalent with $A_f = (A_{jk})_{j,k=1\dots2}$ where

$$A_{jk} = (-1)^{j+k} \det\left(\partial_l f^m\right) , \qquad l \neq k,\, m \neq j$$

is the cofactor matrix. The choice of A_f instead of $\mathrm{D}f$ will turn out to be the correct generalization.

The aim of this section is to generalize (1.2) to arbitrary finite dimensions. For the time being, we note that

$$\mathfrak{d}(f, \mathbb{B}_1(0), 0) := \frac{1}{2\pi i} \oint \frac{f'(z)}{f(z)} \mathrm{d}z$$

has the following properties

Axioms of the degree

(d1) (Whole number property) Let $\mathscr{X}$ be a Banach space, $G \subset \mathscr{X}$ be open and bounded, $f : G \to \mathscr{X}$ be continuous and $y_0 \notin f(\partial G)$. Then

$$\mathfrak{d}(f, G, y_0) \in \mathbb{Z}\,.$$

(d2) (Invertibility)

$$\mathfrak{d}(f, G, y_0) \neq 0 \qquad \Rightarrow \qquad \exists x \in G : \quad f(x) = y_0\,.$$

(d3) (Norming property)

$$\mathfrak{d}(\mathrm{id}, G, y_0) = \begin{cases} 1 & y_0 \in G \\ 0 & y_0 \notin G \end{cases} .$$

(d4) (Homotopy) Let $h : \overline{G} \times [0,1] \to \mathscr{X}$ be continuous, $y : [0,1] \to \mathscr{X}$ continuous and for every $t \in [0,1]$ let $y(t) \notin h(\partial G, t)$. Then

$$t \mapsto \mathfrak{d}\left(h(\cdot, t),\, G,\, y(t)\right) \quad \text{is constant.}$$

In case $\dim \mathscr{X} = \infty$ we require $h(x,t) = x + g(x,t)$, where g is compact on $G \times [0,1]$.

(d5) (Joining Property) If $G_1, G_2 \subset G$ are open, disjoint and $y_0 \notin f\left(\overline{G}\backslash(G_1 \cup G_2)\right)$ then

$$\mathfrak{d}(f, G, y_0) = \mathfrak{d}(f, G_1, y_0) + \mathfrak{d}(f, G_2, y_0)$$

Inspired by the above considerations, we aim to show in the following that for general dimensions the map

$$\mathfrak{d}\left(f, G, y_0\right) := \frac{1}{|\mathbb{S}^{d-1}|} \int_{\partial G} \frac{\langle f - y_0\,,\, A_f\,\nu\rangle}{|f - y_0|^d} \mathrm{d}\sigma\,, \tag{1.3}$$

where

$$A_{jk} = (-1)^{j+k} \det\left(\partial_l f^m\right)_{l\neq k,\, m\neq j}\,, \tag{1.4}$$

satisfies (d1)—(d5). We will follow a natural plan: We first demonstrate that the definition (1.3) makes sense in case of C^2-functions, then generalize to C^0-functions via approximation. Using the fact that compact operators can be approximated by finite dimensional operators, we will finally show that the degree mapping exists in arbitrary dimensions. We furthermore mention that

Exercise 1.1.7. In finite dimensions, the degree is uniquely defined through (d1), (d3)–(d5).

which we will prove later.

1.1.2 The Degree in Finite Dimensions

It turns out that in order to prove that $\mathfrak{d}$ is a suitable degree, it is favorable to first apply Gauss's theorem and transform the above expression to an integral over G rather than ∂G.

Proposition 1.1.8. *For every $f \in C^2(G;\mathbb{R}^d) \cap C^1(\overline{G};\mathbb{R}^d)$ and $y_0 \notin f(\partial G)$ it holds*

$$\mathfrak{d}(f, G, y_0) = \int_G \omega\left(|f(x) - y_0|\right) \mathcal{J}_f(x) \mathrm{d}x\,, \tag{1.5}$$

where $\omega \in C([0,\infty],\mathbb{R})$ with $\int_{\mathbb{R}^d} \omega(|x|)\mathrm{d}x = 1$ has support in $[0,\varepsilon)$, $0 < \varepsilon < \frac{1}{2}\mathrm{dist}\left(f(\partial G), y_0\right)$ and $\mathcal{J}_f$ is the Jacobi determinant of f.

Proof. We define

$$\Phi(y) := \varphi(|y|)\, y\,, \quad \text{where} \quad \varphi(r) := \frac{1}{r^d}\int_0^r \omega(t) t^{d-1}\mathrm{d}t$$

with the property $\varphi(r) = r^{-d}\left|\mathbb{S}^{d-1}\right|^{-1}$ for $r > \varepsilon$. Hence we observe

$$\begin{aligned}
\mathfrak{d}(f, G, y_0) &= \frac{1}{\left|\mathbb{S}^{d-1}\right|}\int_{\partial G} \frac{\langle f - y_0\,,\, A_f\,\nu\rangle}{|f - y_0|^d}\mathrm{d}\sigma \\
&= \int_{\partial G} \langle \Phi\left(f - y_0\right)\,,\, A_f\,\nu\rangle\, \mathrm{d}\sigma \\
&= \int_G \nabla \cdot \Big(\Phi\left(f - y_0\right)^T\, A_f\Big)\,.
\end{aligned}$$

One may easily verify that (see lemma below)

$$\sum_{k=1}^d \partial_k A_{jk} = 0$$

and hence

$$\begin{aligned}\nabla\cdot\left(\Phi(f-y_0)^T A_f\right) &= \sum_{j,k=1}^{d} A_{jk}\partial_k\left(\Phi_j\left(f-y_0\right)\right)+\sum_{j=1}^{d}\Phi_j\left(f-y_0\right)\sum_{k=1}^{d}\partial_k A_{jk}\\
&=\sum_{j,k=1}^{d}\partial_k\left(\Phi_j\left(f-y_0\right)\right)A_{jk}+0\\
&=\sum_{j,k,i=1}^{d}\partial_i\Phi_j\left(f-y_0\right)\partial_k f_i\,A_{jk}\\
&=\sum_{j,k=1}^{d}\partial_j\Phi_j\left(f-y_0\right)\partial_k f_j\,A_{jk}+\sum_{\substack{j,k,i=1\\ i\neq j}}^{d}\partial_i\Phi_j\left(f-y_0\right)\partial_k f_i\,A_{jk}\\
&=\left(\nabla\cdot\Phi\right)\left(f-y_0\right)\mathcal{J}_f+0\,.\end{aligned}$$

Here, we used that $\sum_k \partial_k f_j\,A_{jk}=\det\left(Df\right)$ and the second sum in particular becomes zero as

$$\sum_{k=1}^{d}\partial_k f_i\,A_{jk}=\mathcal{J}_{\tilde f}=0\,,$$

where $\tilde f$ stems from f by replacing f_j by f_i. It remains to observe that

$$\begin{aligned}\left(\nabla\cdot\Phi\right) &= \sum_k\partial_k\Phi_k=d\,\varphi\left(|y|\right)+r\,\partial_r\varphi\left(|y|\right)\\
&=d\,\varphi\left(|y|\right)-d\frac{1}{r}\,r\varphi\left(|y|\right)+r^{-d}\,\omega\left(r\right)r^{d-1}r\\
&=\omega(r)\,.\end{aligned}$$

□

Lemma (Supplemental material)**.** *Let $d\geq 2$ and $f\in C^2(G;\mathbb{R}^d)$ then*

$$\forall j\ :\qquad \sum_{k=1}^{d}\partial_k A_{jk}=0\,.$$

Proof. Define $a_j=(f_1,\dots f_{j-1},f_{j+1},\dots f_d)^T$ then

$$\begin{aligned}\partial_k A_{jk} &= (-1)^{j+k}\sum_{\substack{l=1\\ l\neq k}}^{d}\det\left(\partial_1 a_j,\dots\partial_{l-1}a_j,\partial_k\partial_l a_j,\partial_{l+1}a_j,\dots,\partial_d a_j\right)\\
&=(-1)^j\sum_{\substack{l=1\\ l\neq k}}^{d}\sigma_{kl}\det\left(\partial_k\partial_l a_j,\partial_{i\notin\{k,l\}}a_j\right)\,,\end{aligned}$$

where $\sigma_{kl} = (-1)^{k+l-1}$ if $k > l$ and $\sigma_{kl} = (-1)^{k+l-2}$ if $k < l$. Hence σ_{kl} is anti-symmetric, while

$$b_{kl} := \det\left(\partial_k \partial_l a_j, \partial_{i \notin \{k,l\}} a_j\right)$$

is symmetric. In particular, we obtain that

$$\sum_{k=1}^{d} \partial_k A_{jk} = \sum_{\substack{k,l=1 \\ k \neq l}}^{d} \sigma_{kl} b_{kl} = 0\,.$$

□

There exists a further equivalent expression for $\mathfrak{d}$. It can be derived for regular y_0 in the following sense.

Definition 1.1.9. Let $G \subset \mathbb{R}^d$ be open and $Y = \mathbb{R}^m$ and let $F \in C^1(G;Y)$.

1. $x_0 \in G$ is called *regular* if $\mathrm{D}F(x_0)$ has maximal rank.
2. Any non-regular $x_0 \in G$ is called *critical.*
3. A point $y_0 \in Y$ is called *critical value* of F if there exists a critical $x \in F^{-1}(y_0)$. Otherwise, it is called a *regular value.*

In regular values, the following representation formula holds.

Proposition 1.1.10. *Let y_0 be a regular value of $f \in C^2(G;\mathbb{R}^d) \cap C^1(\overline{G};\mathbb{R}^d)$ and $y_0 \notin f(\partial G)$. Then*

$$\mathfrak{d}\,(f, G, y_0) = \sum_{x \in f^{-1}(y_0)} \mathrm{sign}\mathcal{J}_f(x)\,. \tag{1.6}$$

Proof. First note that $f^{-1}(y_0) \subset G$ is bounded and does not accumulate at ∂G, hence is compact. Since y_0 is regular, $f^{-1}(y_0) = (x_k)_k$ is finite and $f|_{\mathbb{B}^d_\rho(x_k)} : \mathbb{B}^d_\rho(x_k) \to f\left(\mathbb{B}^d_\rho(x_k)\right)$ is a local isomorphism for sufficiently small ρ. In particular, $\mathrm{sign}\mathcal{J}_f$ is constant in $\mathbb{B}^d_\rho(x_k)$ and we can assume that the balls are disjoint.

We may decrease $\varepsilon = \frac{1}{2}\mathrm{dist}\,(f(\partial G), y_0)$ to ensure that also

$$\varepsilon \leq \frac{1}{2}\mathrm{dist}\left(\bigcup_k f\left(\partial \mathbb{B}^d_\rho(x_k)\right)\,,\, y_0\right)\,.$$

It then holds with (1.5) that

$$\begin{aligned}\mathfrak{d}\left(f,G,y_0\right) &= \int_G \omega\left(|f(x)-y_0|\right)\mathcal{J}_f(x)\mathrm{d}x\\ &= \sum_k \int_{\mathbb{B}^d_\rho(x_k)} \omega\left(|f(x)-y_0|\right)\mathcal{J}_f(x)\mathrm{d}x\\ &= \sum_k \mathrm{sign}\mathcal{J}_f(x_k)\int_{\mathbb{B}^d_\rho(x_k)} \omega\left(|f(x)-y_0|\right)\left|\mathcal{J}_f(x)\right|\mathrm{d}x\\ &= \sum_k \mathrm{sign}\mathcal{J}_f(x_k)\int_{f\left(\mathbb{B}^d_\rho(x_k)\right)} \omega\left(|y-y_0|\right)\mathrm{d}y\\ &= \sum_{x\in f^{-1}(y_0)} \mathrm{sign}\mathcal{J}_f(x)\,.\end{aligned}$$

Finally, in order to show that our above definition of the degree is reasonable, we need the following result. □

Lemma 1.1.11 (Sard's Lemma (simplified version))**.** *Let $G=[0,1]^d\subset\mathbb{R}^d$ and $F\in C^1(G;\mathbb{R}^d)$ with bounded derivatives. Then the set of critical values of F has Lebesgue-measure 0.*

Proof. We divide each edge of the cube into N pieces and by doing so also divide G into N^d cubes of equal size with edges of length N^{-1}. For two points x, x_0 in the same cube it holds by Taylor's formula and boundedness of the derivative

$$F(x)=F(x_0)+\mathrm{D}F(x_0)\left(x-x_0\right)+o\left(\frac{1}{N}\right)\,.$$

Furthermore, we have $No\left(\frac{1}{N}\right)\to 0$ uniformly in x and x_0 as $N\to\infty$.

Assuming there was a critical point x_0 of F in one of the cubes, denoted W. Then $\det\mathrm{D}F(x_0)=0$ and hence the values $F(x)$ lie in a $d-1$ dimensional hypermanifold which is comprised in a cuboid of volume $CN^{-d+1}o(N^{-1})$ around the hypersurface given by $F_{approx}(x)=F(x_0)+\mathrm{D}F(x_0)\left(x-x_0\right)$. Since there are at most N^d of these cubes, their total volume is bounded by

$$N^dCN^{-d+1}o(N^{-1})=CNo(N^{-1})\to 0\quad\text{as } N\to\infty\,.$$

As N is arbitrary, this gives the result. □

We can now proof the main results of this section.

Theorem 1.1.12. *Let $\mathfrak{d}$ be defined by (1.3). Then for every $f\in C^2(G;\mathbb{R}^d)\cap C^1(\overline{G};\mathbb{R}^d)$ the map $\mathfrak{d}$ satisfies (d1)–(d3), (d5). Furthermore, $\mathfrak{d}$ satisfies (d4) for $h\in C^2(\overline{G};\mathbb{R}^d)$.*

Proof. Property (d1) follows from the equivalent representation (1.6) in case y_0 is regular. If not, there exists a sequence $y_k\to y_0$ of regular points (by Lemma 1.1.11). From (1.5) we infer that $\mathfrak{d}(f,G,y_k)\to\mathfrak{d}(f,G,y_0)$.

If y_0 is regular, property (d2) follows from (1.6). Otherwise we can once more use the above sequence $y_k\to y_0$ of regular points and continuity of f: If $x_k\in f^{-1}(y_k)$, then x_k is compact and we find $x_k\to x_0$ along a subsequence where $f(x_k)=y_k\to y_0=f(x_0)$.

Property (d3) follows equivalently from the representation (1.5) or (1.6).

In order to show Property (d4) let

$$R = \{(t, h(x,t)) \,:\, x \in \partial G,\, t \in [0,1]\}$$
$$\text{and } Q = \{(t, y(t)) \,:\, t \in [0,1]\}$$

and let $\varepsilon < \frac{1}{2}\text{dist}\,(R,Q)$ in Proposition 1.1.8. But then the expression

$$\mathfrak{d}\,(h(\cdot,t), G, y(t)) = \int_G \omega\,(|h(x,t) - y(t)|)\,\mathcal{J}_{h(\cdot,t)}(x)\mathrm{d}x$$

depends continuously on t.

Finally, property (d5) follows immediately from (1.5). □

In view of (1.5), we finally provide the following result, stating that ω can be chosen out of a much more general class of functions.

Proposition 1.1.13. *For $f \in C^2(G;\mathbb{R}^d) \cap C^1(\overline{G};\mathbb{R}^d)$ and $y_0 \notin f(\partial G)$ let $0 < \varepsilon < \frac{1}{2d}\text{dist}\,(f(\partial G), y_0)$ and let*

$$\Omega_\varepsilon(f, y_0) := \left\{\omega \in C^1(\mathbb{R}^d;\mathbb{R}) \,:\, \int_{\mathbb{R}^d} \omega = 1,\, \operatorname{supp}\omega \subset \mathbb{B}^d_\varepsilon(0)\right\}\,.$$

Then for every $\omega_1, \omega_2 \in \Omega_\varepsilon(f, y_0)$ it holds

$$\int_G \omega_1\,(f(x) - y_0)\,\mathcal{J}_f(x)\mathrm{d}x = \int_G \omega_2\,(f(x) - y_0)\,\mathcal{J}_f(x)\mathrm{d}x\,. \tag{1.7}$$

The last proposition in particular implies that we can generalize the formula (1.5) to any $\omega \in \Omega_\varepsilon(f, y_0)$.

Proof. According to Lemma 1.1.14 below, there exists $w \in C^1_0((-\varepsilon,\varepsilon)^d;\mathbb{R}^d)$ such that $\nabla\cdot w = \omega_1 - \omega_2$. Defining $v_j = \sum_k w_k\,(f - y_0)\,A_{jk}$ we may follow the lines of the proof of Proposition 1.1.8 to obtain

$$\begin{aligned}\nabla\cdot v(x) &= \sum_j \sum_k (\nabla w_k) \circ (f(x) - y_0)\;\partial_j f(x) A_{jk}(x)\\ &= \mathcal{J}_f(x)\,(\nabla\cdot w)\circ(f - y_0) = \mathcal{J}_f\,(\omega_1 - \omega_2)\circ(f - y_0)\;.\end{aligned}$$

Since $w_k\,(f - y_0) = 0$ on ∂G we find $\int_G \nabla\cdot v = \int_{\partial G} 0 = 0$ and hence (1.7). □

Lemma 1.1.14. *Let $K^d_\varepsilon = (-\varepsilon,\varepsilon)^d$. For every $q \in C^1_0(K^d_\varepsilon;\mathbb{R})$ with $\int_{K^d_\varepsilon} q = 0$ there exists $w \in C^1_0(K^d_\varepsilon;\mathbb{R}^d)$ such that $\nabla\cdot w = q$.*

Proof. We prove the lemma via induction over d. For $d = 1$ set $w(x) = \int_{-\varepsilon}^x q(s)\,\mathrm{d}s$.

Assume the lemma was true in dimension d. We denote coordinates in $\mathbb{R}^{d+1}$ by $(y_1,\dots,y_d,t)$ and write

$$\tilde{q}(y) := \int_{-\varepsilon}^{\varepsilon} q(y,t)\mathrm{d}t\,.$$

Then $\tilde{q} \in C_0^1(K_\varepsilon^d;\mathbb{R})$: if for example $y_1 = \pm\varepsilon$, then $q(y,t) = 0$ for all t and hence $\tilde{q}(y) = 0$. Hence we find $\tilde{q} = \nabla \cdot \tilde{w}$ for some $\tilde{w} \in C_0^1(K_\varepsilon^d;\mathbb{R}^d)$.

Now, let $g \in C_0^1((-\varepsilon,\varepsilon);\mathbb{R})$ with $\int_{-\varepsilon}^{\varepsilon} g = 1$ and define

$$w_{d+1}(y,t) := \int_{-\varepsilon}^{t} \left(q(y,s) - g(s)\tilde{q}(y)\right) \mathrm{d}s\,.$$

Obviously, $w_{d+1} \in C_0^1(K_\varepsilon^d;\mathbb{R})$ since $w_{d+1}(y,-\varepsilon) = 0$ by definition and $w_{d+1}(y,\varepsilon) = \tilde{q}(y) - \tilde{q}(y)$. Furthermore, introducing $w_i(y,t) = g(t)\tilde{w}_i(y)$ for $i = 1,\dots d$ and we find

$$\nabla \cdot w(y,t) = g(t)\nabla \cdot \tilde{w}(y) + q(y,t) - g(t)\tilde{q}(y) = q(y,t)\,.$$

□

1.1.3 The Degree for Continuous Functions

Our initial aim was to define the degree for continuous functions, not only for differentiable functions. We will do this using the following

Theorem 1.1.15 (Rouché's Theorem)**.** *Let $f, g \in C^1(\overline{G};\mathbb{R}^d) \cap C^2(G,\mathbb{R}^d)$ and let $y_0 \in \mathbb{R}^d$. If*

$$\forall x \in \partial G: \qquad |f(x) - g(x)| < |f(x) - y_0|$$

then

$$\mathfrak{d}\left(f, G, y_0\right) = \mathfrak{d}\left(g, G, y_0\right)\,.$$

Proof. Let

$$h(x,t) = (1-t)\,f(x) + t\,g(x)\,.$$

This is a C^2-homotopy that can be used in (d4) since for every $x \in \partial G$ it holds

$$|h(x,t) - y_0| \geq |f(x) - y_0| - t\,|f(x) - g(x)| > 0\,.$$

□

Using the Rouché Theorem we will define the degree for continuous functions approximating them by smooth functions. This can be done using Tietzes theorem, which we will prove later:

Theorem. *Let X be a metric space and $A \subset X$ be closed. Further, let $\mathscr{Y}$ be a Banach space and $g: A \to \mathscr{Y}$ be continuous. Then there exists a continuous extension*

$$\overline{g}: X \to \operatorname{conv}\left(g\left(A\right)\right) \subset \mathscr{Y}\,, \qquad \overline{g}|_A = g\,.$$

Here $\operatorname{conv} B$ *is the convex hull of B in $\mathscr{Y}$.*

In particular, we obtain the following:

Lemma 1.1.16. *Let $G \subset \mathbb{R}^d$ be a bounded and open set and let $f \in C\left(\overline{G}\right)$. Then there exists a family $(f_k)_{k\in\mathbb{N}} \subset C^\infty\left(\mathbb{R}^d\right)$ such that $\|f_k - f\|_{C\left(\overline{G}\right)} \to 0$.*

Proof. First observe that $\operatorname{conv} f\left(\overline{G}\right)$ is bounded by compactness of $\overline{G}$ and continuity of f. The Tietze extension $\overline{f}$ of f hence is bounded and on every ball $\mathbb{B}_{R-1}(0)\supset\overline{G}$ the function $\overline{f}$ is uniformly continuous. Let η_k be a sequence of smooth mollifiers with the property that $\operatorname{supp}\eta_k=\mathbb{B}_{\frac{1}{k}}(0)$. Then $f_k(x):=\left(\overline{f}*\eta_k\right)(x)=\int\overline{f}(y)\eta_k\left(y-x\right)\mathrm{d}y$ satisfies $f_k\in C^\infty\left(\mathbb{R}^d\right)$ and by uniform continuity of $\overline{f}$ on $\mathbb{B}_R\left(0\right)$ we find for every $x\in\overline{G}\subset\mathbb{B}_{R-1}(0)$

$$\sup_{x\in\overline{G}}\left|f_k\left(x\right)-\overline{f}\left(x\right)\right|\leq\sup_{x\in\overline{G}}\int\left|\overline{f}(y)-\overline{f}\left(x\right)\right|\eta_k\left(y-x\right)\mathrm{d}y\leq\sup_{\substack{x\in\overline{G},y\in\mathbb{B}_R(0)\\|x-y|<\frac{1}{k}}}\left|\overline{f}(y)-\overline{f}\left(x\right)\right|\to0$$

as $k\to\infty$. □

In the following we write $\|f\|_{\infty,\partial G}:=\sup_{x\in\partial G}|f(x)|$. Let $g_1,g_2\in C^1(\overline{G};\mathbb{R}^d)\cap C^2(G,\mathbb{R}^d)$ both satisfying $\|f-g_i\|_{\infty,\partial G}<\frac{1}{4}\|f-y_0\|_{\infty,\partial G}$. In a first step, we obtain

$$\|g_1-g_2\|_{\infty,\partial G}\leq\|f-g_2\|_{\infty,\partial G}+\|g_1-f\|_{\infty,\partial G}\leq\frac{1}{2}\|f-y_0\|_{\infty,\partial G}\,.$$

Furthermore, we find

$$\|f-y_0\|_{\infty,\partial G}\leq\|f-g_1\|_{\infty,\partial G}+\|g_1-y_0\|_{\infty,\partial G}\leq\frac{1}{4}\|f-y_0\|_{\infty,\partial G}+\|g_1-y_0\|_{\infty,\partial G}$$

and hence $\|f-y_0\|_{\infty,\partial G}\leq\frac{4}{3}\|g_1-y_0\|_{\infty,\partial G}$. In total we find

$$\|g_1-g_2\|_{\infty,\partial G}<\|g_1-y_0\|_{\infty,\partial G}$$

and by Rouché's Theorem

$$\mathfrak{d}\left(g_1,G,y_0\right)=\mathfrak{d}\left(g_2,G,y_0\right)\,.$$

In particular, the following definition is well-posed.

Definition 1.1.17. Let $G\subset\mathbb{R}^d$ be bounded, open, $f\in C(\overline{G};\mathbb{R}^d)$ and $y_0\notin\partial G$. For $g\in C^1(\overline{G};\mathbb{R}^d)\cap C^2(G,\mathbb{R}^d)$ with $\|f-g\|_\infty<\frac{1}{4}\|f-y_0\|_\infty$ define

$$\mathfrak{d}\left(f,G,y_0\right):=\mathfrak{d}\left(g,G,y_0\right)\,.$$

Lemma 1.1.18. *The degree of Definition 1.1.17 satisfies (d1)–(d5).*

Proof. Properties (d1), (d3) and (d5) follow immediately from the definition and the corresponding properties for differentiable f.

Now, let $g_k\to f$ in $C(\overline{G};\mathbb{R}^d)$. Then $\mathfrak{d}\left(f,G,y_0\right)=\mathfrak{d}\left(g_k,G,y_0\right)\neq0$ and hence there exist x_k with $g_k(x_k)=y_0$ and $x_0\in\overline{G}$ with $x_k\to x_0$. Then $|f(x_0)-g_k(x_k)|\leq|f(x_0)-f(x_k)|+|f(x_k)-g_k(x_k)|\to0$ and hence $f(x_0)=\lim_k g_k(x_k)=y_0$.

Let $h\in C([0,1]\times\overline{G};\mathbb{R}^d)$ and $y\in C\left([0,1];\mathbb{R}^d\right)$ with $R=\{(t,h(t,x))\,:\,x\in\partial G,\,t\in[0,1]\}$ and $Q=\{(t,y(t))\,:\,t\in[0,1]\}$ and let $\varepsilon<\frac{1}{2}\operatorname{dist}\left(R,Q\right)$. Then there exists $\tilde{h}\in C^2([0,1]\times\overline{G};\mathbb{R}^d)$ with $\left\|\tilde{h}-h\right\|\leq\frac{1}{4}\varepsilon$ and hence

$$\mathfrak{d}(h(t,\cdot),G,y(t))=\mathfrak{d}(\tilde{h}(t,\cdot),G,y(t))$$

and the statement follows from Theorem 1.1.12. □

1.1.4 The Degree in Infinite Dimensions

In what follows we will consider Banach spaces $\mathscr{X}$ and functions $f = \mathrm{id} + g$, where $g : \mathscr{X} \to \mathscr{X}$ is a compact map. This has to be understood in the following sense.

Definition 1.1.19. Let $\mathscr{X}$, $\mathscr{Y}$ be Banach spaces and let $A \subset \mathscr{X}$. A function $g : A \to \mathscr{Y}$ is compact if it is continuous and $\overline{g(B)}$ is compact for all bounded $B \subset A$.

Since $\mathscr{Y}$ is complete it holds $\overline{A} \subset \mathscr{Y}$ is compact if and only if A is precompact. This means that for every $\varepsilon > 0$ there exists a finite covering of A by balls of radius ε.

If A in Definition 1.1.19 is bounded, it hence suffices to claim $g(A)$ to be precompact, since closed subsets of compact sets are compact.

In finite dimensions, a continuous map f is compact if and only if its range is bounded (due to the Heine-Borel Theorem). In infinite dimensions, a continuous function is compact if and only if it can be approximated by compact finite dimensional functions.

Notation 1.1.20. Denote $\mathbb{C}(A;\mathscr{Y})$ the family of compact mappings from A to $\mathscr{Y}$.

Lemma 1.1.21. *Let $A \subset \mathscr{X}$ be closed and bounded. Then the function $f : A \to \mathscr{Y}$ is compact if and only if there exists a sequence of functions $(f_n)_{n\in N} : A \to \mathscr{Y}$ such that $f_n \to f$ in $C(A;\mathscr{Y})$ and* $\operatorname{span}(f_n(A))$ *is finite dimensional and bounded.*

Proof. Let f be compact. Then $\overline{f(A)}$ is compact and for $\varepsilon = \frac{1}{n}$ there exists a finite covering

$$\overline{f(A)} \subset \bigcup_{k=1}^{K} \mathbb{B}_\varepsilon\left(f(x_k)\right) .$$

We chose a partition of unity $(\psi_k)_{k=1,\dots,K}$ with $\psi_k(y) = 0$ iff $y \notin \mathbb{B}_\varepsilon\left(f(x_k)\right)$. Defining

$$f_n(x) := \sum_{k=1}^{K} \psi_k\left(f(x)\right) f(x_k) \,,$$

every $\operatorname{span}(f_n(A))$ is finite dimensional and every f_n is continuous. Moreover, since $(\psi_k)_{k=1,\dots,K} \neq 0$ if and only if $f(x) \in \mathbb{B}_\varepsilon\left(f(x_k)\right)$ we find for every $x \in A$

$$\|f(x) - f_n(x)\| = \left\|\sum_{k=1}^{K} \psi_k\left(f(x)\right)\left(f(x_k) - f(x)\right)\right\| \leq \sum_{f(x)\in\mathbb{B}_\varepsilon(f(x_k))} \psi_k\left(f(x)\right)\|f(x_k) - f(x)\| = \varepsilon \,.$$

Vice versa, let $f_n \to f$ in $C(A;\mathscr{Y})$ be bounded with finite dimensional $\operatorname{span}(f_n(A))$. Given $\varepsilon > 0$ chose n such that $\|f_n - f\|_{C(A;\mathscr{Y})} < \frac{\varepsilon}{2}$. Furthermore, let $\left(\mathbb{B}_{\frac{\varepsilon}{2}}\left(f_n(x_k)\right)\right)_{k=1,\dots K}$ be a finite covering of $f_n(A)$ (which exists by the Heine-Borel Theorem). For every $x \in A$ there hence exists x_k such that $f_n(x) \in \mathbb{B}_{\frac{\varepsilon}{2}}\left(f_n(x_k)\right)$ and thus

$$\|f(x) - f_n(x_k)\|_{\mathscr{Y}} \leq \|f(x) - f_n(x)\|_{\mathscr{Y}} + \|f_n(x) - f_n(x_k)\|_{\mathscr{Y}} \leq \varepsilon \,,$$

which implies $f(x) \in \mathbb{B}_\varepsilon\left(f_n(x_k)\right)$. This implies precompactness of $f(A)$. □

The property that continuous functions map compact sets onto compact sets is of topological nature. In Banach spaces, we moreover have the following important property.

Lemma 1.1.22. *Let $G \subset \mathscr{X}$ be bounded, $g : \overline{G} \to \mathscr{X}$ compact. Then $f := \mathrm{id} + g$ is closed and proper (the preimages of compact sets are compact).*

Proof. Let $A \subset \overline{G}$ be closed and let $f(x_n)$ be a sequence in $f(A)$ such that $f(x_n) \to y$ converges. We have to show $y \in f(A)$. We first note that $f(x_n) = x_n + g(x_n)$ and since g is compact, we find w.l.o.g. $g(x_n) \to y_g$ implying convergence of $x_n = f(x_n) - g(x_n) \to x = y - y_g$ and closednes of A implies $x \in A$. Moreover, by continuity of f we find

$$f(x_n) \to f(x) = x + y_g = y \in f(A)\,.$$

Hence, $f(A)$ is closed.

If $B \subset f(A)$ is compact and $y_n = f(x_n) \to y$ is a convergent sequence in B, we find $g(x_n) \to y_g$ along a subsequence and hence $x_n = y_n - g(x_n) \to y - y_g =: x \in \overline{G}$ and continuity of g implies $y = x + g(x)$, i.e. $x \in f^{-1}(B)$. □

The definition of the degree in infinite dimensions is now based on the following result.

Lemma 1.1.23. *Let $G \subset \mathbb{R}^d$ be open and bounded and let $m < d$. Writing $E^m := \mathbb{R}^m \times \{0\}^{d-m} \subset \mathbb{R}^d$, let $\pi : E^m \to \mathbb{R}^m$ and $\pi^* : \mathbb{R}^m \to E^m$ be the identifications. For every $f = \mathrm{id} + g$ with $g \in \mathbb{C}(\overline{G}; E^m)$ let $y_0 \in E^m$ with $y_0 \notin f(\partial G)$. Then*

$$\mathfrak{d}_{\mathbb{R}^d}\left(f, G, y_0\right) = \mathfrak{d}_{\mathbb{R}^m}\left(\pi \circ f \circ \pi^*, \pi\left(G \cap E^m\right), \pi y_0\right)\,.$$

Proof. W.l.o.g. let $g \in C^1(\overline{G}; E^m)$ and let y_0 be a regular point. According to (1.6) it suffices to show $\mathrm{sign}\mathcal{J}_f(x) = \mathrm{sign}\mathcal{J}_{f^*}(\pi(x))$ for $f^* = \pi \circ f \circ \pi^*$ and for every $x \in f^{-1}(y_0)$. We also write $g^* = \pi \circ g \circ \pi^*$and obtain

$$\mathcal{J}_f = \det\begin{pmatrix} \mathrm{id}_{\mathbb{R}^m} + \mathrm{D}g^* & 0 \\ & \mathrm{id}_{\mathbb{R}^{d-m}} \end{pmatrix} = \det\left(\mathrm{id}_{\mathbb{R}^m} + \mathrm{D}g^*\right) = \mathcal{J}_{f^*}\,.$$

In particular, note that also the sign coincides. □

Based on the last lemma, we are able to generalize the concept of degree to infinite dimensions. Due to Lemma 1.1.22 $f(\partial G)$ is closed and hence for $y_0 \notin f(\partial G)$ we infer that $\varepsilon := \frac{1}{2}\mathrm{dist}\left(y_0, f(\partial G)\right) > 0$. According to Lemma 1.1.21 we can approximate g by a finite dimensional g_ε with $\|g - g_\varepsilon\|_{C(G;\mathscr{X})} < \frac{\varepsilon}{3}$ and $E_\varepsilon := \mathrm{span}\, g_\varepsilon(\mathscr{X})$. For given f, y_0 and $\varepsilon > 0$ let g_1, g_2 be two such functions with $E_i := \mathrm{span}\, g_i(\mathscr{X})$, $i = 1, 2$. Furthermore, let $F = \mathrm{span}\left(E_1 \cup E_2\right) \subset \mathscr{X}$. W.l.o.g. we assume $y_0 \in E_1 \cap E_2$ and Lemma 1.1.23 yields

$$\mathfrak{d}\left(f_i|_{E_i \cap G}, E_i \cap G, y_0\right) = \mathfrak{d}\left(f_i|_{F \cap G}, F \cap G, y_0\right)\,.$$

Furthermore, we find

$$\|f_1 - f_2\| < \frac{2\varepsilon}{3} = \frac{2}{3}\mathrm{dist}\left(y_0, f(\partial G)\right) \leq \mathrm{dist}\left(y_0, f_1(\partial G)\right)\,.$$

Now, Rouché's Theorem yields

$$\mathfrak{d}\left(f_1|_{F \cap G}, F \cap G, y_0\right) = \mathfrak{d}\left(f_2|_{F \cap G}, F \cap G, y_0\right)\,.$$

In particular, the expression

Definition 1.1.24.

$$\mathfrak{d}\,(f, G, y_0) := \mathfrak{d}\,(f_\varepsilon|_{E_\varepsilon\cap G}, E_\varepsilon \cap G, y_0) \tag{1.8}$$

is well defined and independent from g_ε as long as $\|g - g_\varepsilon\|_{C(G;\mathscr{X})} < \frac{\varepsilon}{3}$ for compact finite dimensional g_ε.

Theorem 1.1.25. *The degree from Definition 1.1.24 satisfies (d1)–(d5).*

Proof. The properties (d1), (d3) and (d5) (i.e. whole number, norming and joining properties) are satisfied by definition of $\mathfrak{d}$ and Rouché's Theorem.

In order to see (d2) let $\mathfrak{d}\,(f, G, y_0) \neq 0$. Let g_k be a sequence of finite dimensional functions $g_k : \overline{G} \to E_k$, $E_k \subset \mathscr{X}$ finite dimensional space, such that $g_k \to g$ in $C(\overline{G};\mathscr{X})$. Then for every k there exists $x_k \in G$ such that $y_0 = x_k + g_k(x_k)$. Since $g(x_k) \to y_g$ along a subsequence (compactness) and $g_k \to g$ uniformly, we find $\|g_k(x_k) - g(x_k)\|_{\mathscr{X}} \leq \|g_k - g\|_{C(G;\mathscr{X})} \to 0$. Hence

$$x_k = y_0 - g_k(x_k) \to y_0 - g_0 =: x_0$$

and by continuity $f(x_0) = y_0$.

In order to verify (d4), we note that we demand $h(\cdot, t) = \mathrm{id} + g(\cdot, t)$ where $g \in C(\overline{G} \times [0, 1];\mathscr{X})$ is compact. The function $K(t) := \mathrm{dist}\,(y(t), h(\partial G, t))$ is continuous and hence attains its minimum $K_0 > 0$ over $[0, 1]$. Due to compactness of g there exist a finite dimensional approximation $g_k \to g$ in $C(\overline{G} \times [0, 1]; E_k)$, $E_k \subset \mathscr{X}$ is a linear subspace and where we demand

$$\|g_k - g\|_{C(\overline{G}\times[0,1];\mathscr{X})} \leq \frac{1}{24} K_0\,.$$

Furthermore, the continuity of y implies compactness of $y([0, 1])$. We hence observe that $y(t)$ can be approximated by a finite dimensional piecewise affine curve y_k with $\|y - y_k\|_{C([0,T];\mathscr{X})} < \frac{1}{24} K_0$.

For fixed t we obtain by Lemma 1.1.23 that (uppon extending E_k to $\tilde{E}_k(t) := E_k \oplus \mathbb{R}y(t)$)

$$\mathfrak{d}\left(h_k(\cdot, t), G \cap \tilde{E}_k(t), y(t)\right) = \mathfrak{d}\left(h_k(\cdot, t), G \cap \tilde{E}_k, y_k(t)\right) \tag{1.9}$$

Now,

$$\mathfrak{d}\,(h_k(\cdot, t), G \cap E_k, y_k(t)) = const$$

and according to its definition in (1.8) and (1.9), the same holds for the degree

$$\mathfrak{d}\,(h(\cdot, t), G, y(t)) = const\,.$$

□

1.1.5 The Hopf Theorem

Given $\varphi : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ there exist various possible extensions $\overline{\varphi} : \mathbb{B}^d_1(0) \to \mathbb{B}^d_1(0)$. In the right coordinates, e.g. $\nu = e_d$, the expression $A_f e_d = \left((-1)^{j+d} \det\,(\partial_l f^m)_{l\neq d,\, m\neq j}\right)$ depends only on tangential derivatives of φ and hence the degree in (1.3) depends only on φ. Hence, one can define

$$\deg(\varphi) := \mathfrak{d}\left(\overline{\varphi}, \mathbb{B}^d_1(0), 0\right)\,.$$

From (d4) we know that two homotope functions φ_1 and φ_2 share the same degree $\deg(\varphi_1) = \deg(\varphi_2)$. The opposite holds true, too. Without any proof (which is lengthy and rather topological) we provide the following result:

Theorem 1.1.26 (Hopf)**.** *For any two continuous maps $\varphi_1, \varphi_2 : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ there holds*

$$\deg(\varphi_1) = \deg(\varphi_2) \qquad \Leftrightarrow \qquad \varphi_1 \sim \varphi_2 \,.$$

We refer to [6] for a prove of this theorem below. The particular benefit of Hopf's theorem is that one immediately obtains a characterization for a function $\varphi : \mathbb{S}^{d-1} \to \mathbb{R}^d \setminus \{0\}$ to be essential or not by moving to $\psi = \varphi / |\varphi|$.

Definition 1.1.27. Let $\mathscr{X}$ be a Banach space and $\mathbb{S}_{\mathscr{X}}$ the unit sphere in $\mathscr{X}$. Let

$$\varphi : \mathbb{S}_{\mathscr{X}} \to \mathscr{X} \setminus \{0\} \,, \qquad \varphi = \mathrm{id} + g \,, \quad g \in \mathbb{C}(\mathbb{S}_{\mathscr{X}}; \mathscr{X}) \,. \tag{1.10}$$

φ is called *essential* if for every $\overline{g} \in \mathbb{C}\left(\overline{\mathbb{B}_1^{\mathscr{X}}(0)};\, \mathscr{X}\right)$ with $\overline{g}|_{\mathbb{S}_{\mathscr{X}}} \equiv g$ the function $\overline{\varphi} := \mathrm{id} + \overline{g}$ has a zero in $\mathbb{B}_1^{\mathscr{X}}(0)$. If this will not cause confusion we sometimes identify φ with $\overline{\varphi}$.

In what follows, we discuss some properties of essential functions and provide an infinite dimensional version of Hopf's Theorem.

Lemma 1.1.28. *Let φ_1 be essential and let φ be of the form (1.10) such that $\varphi \sim \varphi_1$. Then φ is essential.*

Proof. We show that if φ is not essential then also φ_1 is not essential.

Let $\varphi = \mathrm{id} + g$ and $\varphi_1 = \mathrm{id} + g_1$. Let G be the homotopy from g to g_1 and $\overline{G}$ the extension of the homotopy to $\overline{\mathbb{B}_1^{\mathscr{X}}(0)}$ and let $h := \mathrm{id} + \overline{G}$. Let

$$A := \left\{ x \in \overline{\mathbb{B}_1^{\mathscr{X}}(0)} \mid \exists t \in [0,1] \,:\, h(x,t) = 0 \right\} \,.$$

The set A is closed by continuity of h and $A \cap \mathbb{S}_{\mathscr{X}} = \emptyset$. We chose $\tau : \overline{\mathbb{B}_1^{\mathscr{X}}(0)} \to [0,1]$ with $\tau|_{\mathbb{S}_{\mathscr{X}}} = 1$ and $\tau|_A = 0$. Define $h^*(x,t) = h(x, \tau(x)t)$ then $h^*(x,t) = h(x,t)$ if $x \in \mathbb{S}_{\mathscr{X}}$ and $h(x,0) = h^*(x,0)$ and

$$h^*(x,t) = 0 \quad \Rightarrow \quad x \in A \quad \Rightarrow \quad \tau(x) = 0 \,.$$

And hence

$$0 = h^*(x,t) = h^*(x, \tau(x)t) = h\left(x, 0\right) \,.$$

Since φ is not essential, we can chose $h(\cdot, 0)$ in a way that $h(x^*, 0) \neq 0$ for every $x^* \in \mathbb{B}_1^{\mathscr{X}}(0)$.

On the other hand, if φ_1 is essential, there is at least one element of A. But this is a contradiction. □

One particular difficulty in the context of infinite dimensions is the question how to move from general $\varphi : \mathbb{S}^{\mathscr{X}} \to \mathscr{X} \setminus \{0\}$ to functions $\mathbb{S}^{\mathscr{X}} \to \mathbb{S}^{\mathscr{X}}$. However, since in finite dimensions the two situations are equivalent in the sense that the latter case implies the first one, we restrict only to the first setting.

Lemma 1.1.29 (The Hopf theorem in infinite dimensions)**.** *φ of the form (1.10) is essential if and only if $\mathfrak{d}(\overline{\varphi},\mathbb{B}_1^{\mathscr{X}}(0),0) =: \deg(\varphi)\neq 0$ (which is independed from the particular extension).*

Proof. Let $\overline{\varphi}_1$ and $\overline{\varphi}_2$ be two extensions of φ. Then $h(x,t) = t\overline{\varphi}_1 + (1-t)\,\overline{\varphi}_2$ is a valid homotopy.

If $\mathfrak{d}(\overline{\varphi},\mathbb{B}_1^{\mathscr{X}}(0),0)\neq 0$ the function φ is essential by (d2).

If $\mathfrak{d}(\overline{\varphi},\mathbb{B}_1^{\mathscr{X}}(0),0) = 0$ let $\varepsilon > 0$ be small enough and g_1 be a finite dimensional approximation of g with values in a finite dimensional space $E_1\subset\mathscr{X}$. We can assume w.l.o.g. $g_1\in C^2(B)$.

By Definition of the degree it holds with $B := \mathbb{B}_1^{\mathscr{X}}(0)\cap E_1 = \mathbb{B}_1^{E_1}(0)$

$$\mathfrak{d}\left(\overline{\varphi},\mathbb{B}_1^{\mathscr{X}}(0),0\right) = \mathfrak{d}\left(\overline{\varphi}_1|_{E_1},B,0\right)$$

Since $n(x) := \|x+g_1(x)\|$ is strictly positive on ∂B, we can extend it to a strictly positive function on $\overline{B}$ and use the homotopy $h(x,t) = \frac{\overline{\varphi}_1(x)}{t+(1-t)n(x)}$ to a function $\psi(x)$ with $\psi:\mathbb{S}^{E_1}\to\mathbb{S}^{E_1}$. But now Hopf yields $\mathfrak{d}(\psi,B,0) = \mathfrak{d}(y,B,0)$ for any $y\in\mathbb{S}^{E_1}$ and hence ψ is homotope to some non-essential function. But then also $\overline{\varphi}_1$ is not essential by Lemma 1.1.28. Applying the Lemma again, we infer that $\overline{\varphi}$ is not essential. □

1.2 Applications of the Degree

1.2.1 Extension Theorems

Theorem 1.2.1 (Tietze)**.** *Let X be a metric space and $A \subset X$ be closed. Further, let $\mathscr{Y}$ be a Banach space and $g : A \to \mathscr{Y}$ be continuous. Then there exists a continuous extension*

$$\overline{g} : X \to \operatorname{conv}(g(A)) \subset \mathscr{Y}, \qquad \overline{g}|_A = g.$$

Here conv$B$ is the convex hull of B in $\mathscr{Y}$.

Remark 1.2.2.

1. It suffices to choose $\mathscr{Y}$ a topological space.
2. If g is compact it is possible to construct $\overline{g}$ as a compact function.

Proof. For every $x \in X\backslash A$ let $\delta(x) := \frac{1}{2}\operatorname{dist}(x, A) > 0$. We may choose a locally finite covering $\left(B_j = \mathbb{B}_{\delta(x_j)}(x_j)\right)_j$ of $X\backslash A$ and a corresponding partition of unity

$$\eta_j(x) := \frac{\operatorname{dist}(x, X\backslash B_j)}{\sum_i \operatorname{dist}(x, X\backslash B_i)}.$$

For every B_j there exists $a_j \in A$ such that $\operatorname{dist}(a_j, B_j) < 2\operatorname{dist}(A, B_j)$. Then, the function

$$\overline{g}(x) := \begin{cases} g(x) & x \in A \\ \sum_j \eta_j(x) g(a_j) & x \notin A \end{cases}$$

is continuous on $X\backslash\partial A$. It is evident that $\overline{g}(X) \subset \operatorname{conv}(g(A))$.

It remains to show continuity of $\overline{g}$ in ∂A. Let $x_0 \in \partial A$ and $x \in X\backslash A$. Then

$$\|g(x_0) - g(x)\|_{\mathscr{Y}} \leq \sup\left\{\|g(x_0) - g(a_j)\|_{\mathscr{Y}} : \text{ for } j \text{ s.t. } x \in B_j\right\}.$$

For j such that $x \in B_j$ we find $\operatorname{dist}(A, B_j) < \operatorname{dist}(x, x_0) =: \delta$ and hence by definition

$$\delta(x_j) \leq \frac{1}{2}\operatorname{dist}(x_j, A) \leq \frac{1}{2}\left(\operatorname{dist}(B_j, A) + \delta(x_j)\right)$$

implying

$$\delta(x_j) \leq \operatorname{dist}(B_j, A) \leq \delta.$$

Taking all together, we obtain

$$\begin{aligned} \operatorname{dist}(a_j x_0) &\leq \delta + \operatorname{dist}(a_j, x) \\ &\leq \delta + \operatorname{dist}(a_j, B_j) + \delta(x_j) \\ &\leq \delta + 2\operatorname{dist}(A, B_j) + \delta(x_j) \leq 4\delta. \end{aligned}$$

This in turn implies that $a_j \to x_0$ as $\delta \to 0$ and hence $\overline{g}(x) \to g(x_0)$. □

Lemma 1.2.3. *Let $W \subset \mathbb{R}^d$ be a cube and $K \subset W$ compact. Let $g : K \to \mathbb{R}^m\backslash\{0\}$ be continuous with $m > d$. Then g can be continuously extended to a function*

$$\overline{g} : W \to \mathbb{R}^m\backslash\{0\}.$$

Proof. Define $c = \inf_{x\in K} |g(x)|$ and let g_1 be an extension of g to $\mathbb{R}^d$ according to Tietze's Theorem. We choose $g_2 \in C^2(W;\mathbb{R}^m)$ such that $|g_2 - g_1| < \varepsilon$ for $\frac{c}{3} > \varepsilon > 0$. $g_2(W)$ is a hypermanifold in $\mathbb{R}^m$ and hence has Lebesgue measure zero. Moreover, there exists $y_0 \in \mathbb{B}_\varepsilon(0)$ such that $\operatorname{dist}(0, g_2(W) - y_0) > 0$. Without loss of generality, we assume $y_0 = 0$.

Now, introduce

$$\eta(t) := \begin{cases} \frac{3}{2c}t & t \le \frac{3c}{2} \\ 1 & t > \frac{2}{3}c \end{cases}$$

and set $g_3(x) := (g_2(x) - y_0)\,\eta\left(|g_2(x) - y_0|\right)^{-1} + y_0$. Then $|g_3| > \frac{1}{3}c$ with $g_3 = g_2$ on K and hence $|g_3 - g| = |g_2 - g| < \varepsilon$ on K.

We may now extend $\psi := g_3 - g$ as a function $K \to B^d_\varepsilon(0)$ to a function $\overline{\psi} : W \to B^d_\varepsilon(0)$. Hence

$$\left|g_3 - \overline{\psi}\right| \ge |g_3| - \left|\overline{\psi}\right| \ge \frac{c}{3} - \varepsilon\,,$$

and for ε small enough we discover that $g_3 - \overline{\psi}$ is the required extension of g. □

Definition 1.2.4. A set $G \subseteq \mathbb{R}^d$ is called symmetric if $G = -G$. A function $f : G \to \mathbb{R}^d$ is called odd if $f(-x) = -f(x)$ for every $x \in G$.

Lemma 1.2.5 (Odd Extension Lemma)**.** *Let $D \subset \mathbb{R}^d$ be open, bounded and symmetric such that $0 \notin \overline{D}$. Let $g : \partial D \to \mathbb{R}^m\backslash\{0\}$ be odd and continuous with $m > d$. Then there exists a continuous odd extension of g such that*

$$\overline{g} : \overline{D} \to \mathbb{R}^m\backslash\{0\}\,.$$

Proof. For $d = 1$ the statement is obvious.

Assume the statement was true for $d - 1 \ge 1$ and prove the statement for dimension d. To this aim, let $D_0 := D \cap \{x_d = 0\} \subset \mathbb{R}^{d-1}$. Then, $g : \partial D_0 \to \mathbb{R}^m$ can be extended to $\overline{D}_0$.

Defining $D_+ := D \cap \{x_d > 0\}$, the function g is now defined on ∂D_+ and hence can be extended to $\overline{g} : \overline{D}_+ \to \mathbb{R}^m\backslash\{0\}$ using Lemma 1.2.3. By $\overline{g}(-x) := -\overline{g}(x)$ if $x_d < 0$, the function can be extended to $\overline{D}$. □

1.2.2 Fixed Point Theorems

We recall the following result.

Theorem 1.2.6 (Banach fixed point theorem)**.** *Let (X, d) a complete metric space and $\varphi : X \to X$ a contraction, that is*

$$\exists \alpha \in (0,1) : \qquad \forall x, y \in X : d(\varphi(x), \varphi(y)) \le \alpha d(x,y)\,.$$

Then there exists a unique $x_0 \in X$ such that $x_0 = \varphi(x_0)$ and for every $y_0 \in X$ the sequence $y_k := \varphi(y_{k-1})$ converges to x_0.

In this section we seek for alternative fixed point theorems. The first one is the Schauder fixed point theorem

Theorem 1.2.7 (Schauder's Theorem)**.** *Let $\mathscr{X}$ be a Banach space, K be bounded, closed and convex and let $g : K \to K$ be continuous and compact. Then g has a fixed point. This also holds if K is homeomorphic to a closed convex set.*

Proof. Let $K = B = \mathbb{B}_1(0)$ and $f := \mathrm{id} - g$, $y_0 = 0$ and $G := \overline{B}$. Every fixed points of g is a zero of f. If f has a zero on ∂B, we are done. Otherwise it is sufficient to show

$$\mathfrak{d}\,(f, G, 0) \neq 0\,.$$

The function $h(x, t) = x - tg(x)$ is a valid homotopy since

$$\|h(x,t)\|_{\mathscr{X}} \geq 1 - t\,\|g(x)\|_{\mathscr{X}} \geq 1 - t > 0$$

for every $x \in \partial B$ and $t \neq 1$ and since f has no zeros on ∂B this also holds for $t = 1$. Because of (d4) and (d3) we find $\mathfrak{d}\,(f, G, 0) = \mathfrak{d}\,(\mathrm{id}, G, 0) = 1$. Hence f has a fixed point.

In the general case we can use Tietzes Theorem 1.2.1 to extend f to $\overline{g} : B \to K$ and Lemma 1.2.7 yields a fixed point of $\overline{g}$ which necessarily lies in K. In case K is homeomorphic to a convex set in B and Φ is the homeomorphism, consider $\Phi \circ g \circ \Phi^{-1}$ instead.

□

A special case is the Brouwer Theorem, which is an immediate consequence.

Theorem 1.2.8 (Brouwer's Theorem)**.** *Let $K \subset \mathbb{R}^d$ non-empty, compact and convex or homeomorphic to such a set. Let $g : K \to K$ be continuous. Then f has at least one fixed point.*

Let us now turn to a fundamental existence theorem for ordinary differential equations in Banach spaces. It will be based on the following generalization of the Arzela Ascoli theorem

Theorem 1.2.9 (Arzela-Ascoli theorem for Banach valued functions)**.** *Let $\mathscr{Y}$ be Banach spaces and let $A \subset C([0,T];\mathscr{Y})$. Then $\overline{A}$ is compact if and only if*

$$\sup_{f\in A} \|f\|_{C([0,T];\mathscr{Y})} < \infty\,, \qquad \forall t \in [0,T] \,:\, \{f(t) \,:\, f \in A\} \quad \textit{is precompact}$$
$$\sup_{f\in A} \|f(t) - f(s)\|_{\mathscr{Y}} \to 0 \quad \textit{as } |t-s| \to 0\,.$$

The proof is similar to the finite dimensional case, using precompactness of $\{f(t) \,:\, f \in A\}$ instead of Helly's theorem.

Theorem 1.2.10 (Peano's theorem)**.** *Let $f : [0,T] \times \mathscr{X} \to \mathscr{X}$ be compact. Then, for all initial data $u(0) = u_0$, there exists a solution small times to*

$$\dot{u}(t) = f(t, u(t))\,. \tag{1.11}$$

Proof. From small t_0 and $r > 0$ we find boundedness of f on $M := [0,t_0] \times \overline{\mathbb{B}_r^{\mathscr{X}}(u_0)}$, i.e. $\|f\|_{C(M)} \leq C$. Hence, $f(M)$ is compact and a solution of (1.11) satisfies

$$u(t) = u_0 + \int_0^t f(s, u(s))\mathrm{d}s\,. \tag{1.12}$$

We use this to define the operator

$$Su(t) := u_0 + \int_0^t f(s,u(s))\mathrm{d}s$$

and observe that $\|Su - u_0\|_{C([0,t_0];\mathscr{X})} \leq t_0 C$. In particular, defining

$$A := \left\{ u \in C([0,t_0];\mathscr{X}) \,:\, u(0) = u_0,\, \|u\|_{C([0,t_0];\mathscr{X})} \leq Ct_0 \right\} ,$$

we find $S:\, A \to A$. Provided S is compact, Schauder's theorem yields existence of a solution to (1.12).

Evidently, $\sup_{u\in A} \|Su\|_{C(M;\mathscr{X})} < \infty$ and for every $t \in [0,t_0]$ we find $\{Su(t) \,:\, u \in A\}$ is precompact. Furthermore, we find

$$\|Su(t_1) - Su(t_2)\| \leq C\,|t_1 - t_2| \to 0 \quad \text{as } |t_1 - t_2| \to 0\,.$$

The compactness of $S(A)$ and hence of S follows from Theorem 1.2.9. □

1.2.3 Borsuks Theorem and the Hedgehog Theorem

Theorem 1.2.11 (Hedgehog Theorem)**.** *Let d be odd let $0 \in G$ for the bounded open domain $G \subset \mathbb{R}^d$. Then, for every continuous $f:\, \partial G \to \mathbb{R}^d \backslash \{0\}$ there exists $x \in \partial G$ and $\lambda \in \mathbb{R}$ such that $f(x) = \lambda x$.*

Proof. Let $\overline{f}:\, \overline{G} \to \mathbb{R}^d$ be an extension by Tietze's Theorem. Since d is odd, we know that $\mathfrak{d}(-\mathrm{id}, \overline{G}, 0) = -1$.

Now, assuming $\mathfrak{d}\left(\overline{f}, \overline{G}, 0\right) \neq -1$ the function

$$h(x,t) := (1-t)\,\overline{f}(x) - t\,x$$

cannot be a valid homotopy in the sense of (d4). Hence there exists $x \in \partial G$ and $t \in [0,1]$ such that

$$0 = h(x,t) = (1-t)\,\overline{f}(x) - t\,x\,.$$

Since $t \in \{0,1\}$ is excluded, we find $\lambda \in \mathbb{R}\backslash\{0\}$ such that $f(x) = \lambda x$.

In case $\mathfrak{d}\left(\overline{f}, \overline{G}, 0\right) = -1$, we may perform the same calculation with id. □

Theorem 1.2.12 (Borsuks Theorem)**.** *Let $\mathscr{X}$ be a Banach space and let $G \subset \mathscr{X}$ be symmetric with $0 \in G$ and let $f:\, \overline{G} \to \mathscr{X}$ continuous with $0 \notin f(\partial G)$ and of the form $f = \mathrm{id} + g$ for some compact g. Furthermore, let $\frac{f(x)}{\|f(x)\|_{\mathscr{X}}} \neq \frac{f(-x)}{\|f(-x)\|_{\mathscr{X}}}$ for all $x \in \partial G$. Then $\mathfrak{d}(f, G, 0)$ is odd.*

A particular case of Borsuks Theorem is the case when $\mathscr{X} = \mathbb{R}^d$ and $f(x) = -f(-x)$, for all $x \in \partial G$, i.e. f is odd. We will see in the following proof that the general case always reduces to this particular case.

Proof. The function $\tilde{g}(x) := \frac{1}{2}(g(x) - g(-x))$ is odd. We show that

$$h(x,t) := x + (1-t)g(x) + t\tilde{g}(x)$$

is a valid homotopy between $\mathrm{id} + g$ and $\mathrm{id} + \tilde{g}$. This follows from the fact that $h(x,t) = 0$ is equivalent with

$$\begin{aligned} & \left(1 - \frac{t}{2}\right) g(x) + x = \frac{t}{2} g(-x) \\ \Leftrightarrow \quad & \left(1 - \frac{t}{2}\right)(g(x) + x) = \frac{t}{2}(g(-x) - x) \\ \Leftrightarrow \quad & \left(1 - \frac{t}{2}\right) f(x) = \frac{t}{2} f(-x) \,. \end{aligned}$$

Since there exists t with $\left(1 - \frac{t}{2}\right)\left(\frac{t}{2}\right)^{-1} = \|f(-x)\|_{\mathscr{X}} \|f(x)\|_{\mathscr{X}}^{-1}$ we conclude $x \notin \partial G$.

Thus, it suffices to consider odd compact functions g. Furthermore, by the definition of the infinite dimensional degree, we conclude that it is sufficient to consider finite dimensional image of g and hence, we can restrict to finite dimensional spaces $\mathscr{X} = \mathbb{R}^d$.

We chose $\varepsilon > 0$ such that $\overline{\mathbb{B}}_\varepsilon^d(0) \subset G$. Let $f_{\mathrm{id}}(x) = f(x)$ if $x \in \partial G$ and $f_{\mathrm{id}}(x) = x$ if $x \in \partial \mathbb{B}_\varepsilon^d$. Since both $f|_{\partial G}$ and $\mathrm{id}|_{\partial \mathbb{B}_\varepsilon^d}$ are odd, we can extend

$$f_{\mathrm{id}} : \partial\left(G \backslash \mathbb{B}_\varepsilon^d(0)\right) \cap \{x_d = 0\} \to \mathbb{R}^d \backslash \{0\}$$

by an application of the odd extension Lemma 1.2.5 to an odd function

$$f_{\mathrm{id}} : \overline{G \backslash \mathbb{B}_\varepsilon^d(0)} \cap \{x_d = 0\} \to \mathbb{R}^d \backslash \{0\} \,.$$

Now we apply Tietzes theorem and extend f_{id} to $\overline{G \backslash \mathbb{B}_\varepsilon^d(0)} \cap \{x_d \geq 0\}$ and using $f_{\mathrm{id}}(-x) := -f_{\mathrm{id}}(x)$ in case $x_d < 0$, we have thus extended f_{id} to an odd function

$$f_{\mathrm{id}} : \overline{G \backslash \mathbb{B}_\varepsilon^d(0)} \to \mathbb{R}^d \backslash \{0\} \,.$$

Since $\mathfrak{d}(g, \tilde{G}, 0)$ depends only on $g|_{\partial G}$ (at least for twice continuously differentiable functions), we obtain by (d3) and (d5) that

$$\begin{aligned} \mathfrak{d}(f, G, 0) = \mathfrak{d}(f_{\mathrm{id}}, G, 0) &= \mathfrak{d}\left(f_{\mathrm{id}}, \mathbb{B}_\varepsilon^d(0), 0\right) + \mathfrak{d}\left(f_{\mathrm{id}}, G \backslash \mathbb{B}_\varepsilon^d(0), 0\right) \\ &= 1 + \mathfrak{d}\left(f_{\mathrm{id}}, G \backslash \mathbb{B}_\varepsilon^d(0), 0\right) \\ &= 1 + 2\mathfrak{d}\left(f_{\mathrm{id}}, \left(G \backslash \mathbb{B}_\varepsilon^d(0)\right) \cap \{x_d \geq 0\}, 0\right) \,. \end{aligned}$$

This implies that $\mathfrak{d}(f, G, 0)$ is odd. □

The next result is the weather theorem. In a folkloristic way, it states that there are always two antipodean points on earth with same temperature and pressure.

Corollary 1.2.13 (The weather theorem)**.** *Let $G \subset \mathbb{R}^d$ be open, bounded and symmetric and let $0 \in G$. For $m < d$ let $f : G \to \mathbb{R}^m$ be continuous. Then there exists $x \in \partial G$ with $f(x) = f(-x)$.*

Proof. Since $m < d$ we may assume that f has values in $\mathbb{R}^d$ by setting $(f_i(x))_{i=m+1,\dots d} \equiv 0$. Assume now that $g(x) := f(x) - f(-x)$ had no zeros on ∂G. Extending g to $\overline{g} : \overline{G} \to \mathbb{R}^d$, the Borsuk Theorem yields $\mathfrak{d}(g, G, 0) \neq 0$. Due to the homotopy property, we furthermore conclude $\mathfrak{d}\,(g, G, \delta e_d) \neq 0$ for sufficiently small δ. By (d3) this is a contradiction with $g_d \equiv 0$. □

In the following, one may associate with A_1, A_2 and A_3 bread, cheese and ham. The following theorem states that a sandwich can always be cut into two parts such that all three ingredients are simultaneously divided into equal parts.

Corollary 1.2.14 (The sandwich theorem). *Let A_1, A_2 and A_3 be measurable subsets of $\mathbb{R}^3$ with finite Lebesgue measure. Then there exists a plane that cuts all three sets into two equal parts.*

Proof. Let $x \in \mathbb{S}^2$. We consider $E_x(t)$ the plane in tx that is orthogonal to x. Furthermore, we define

$$A_{k,-}(x,t) := \{y \in A_k \;:\; y \cdot x \leq t\}\,, \quad k = 1,2,3$$

the set of all elements in A_k that lie "below" $E_x(t)$. This can be seen from the $y \in A_{k,-}(x,t)$ if and only if $(y - tx)\cdot x \leq 0$. We furthermore denote $V_{k,-}(x,t) := |A_{k,-}(x,t)|$ and $V_{k,+}(x,t) := |A_k| - |A_{k,-}(x,t)|$.

Since $V_{k,\pm}(x,t)$ are monotone in t and sum up to $|A_k|$, there exist $-\infty < t_1(x) \leq t_2(x) < +\infty$ such that $E_x(t)$ cuts A_3 into equal parts if and only if $t \in [t_1(x), t_2(x)]$. We define $t_0(x) := \frac{1}{2}\left(t_1(x) + t_2(x)\right)$ (remark that $t_0(x) = -t_0(-x)$) and consider

$$V_{k,-}(x) := V_{k,-}(x, t_0(x)) = V_{k,+}(-x, -t_0(x)) = V_{k,+}(-x, t_0(-x))\,, \quad k = 1,2$$

the volume of all elements of A_k that lie "below" $E_x(t_0(x))$. $V_{k,-}(x)$ depends continuously on x and according to the weather theorem there exists $x_0 \in \mathbb{S}^2$ such that

$$\begin{aligned} V_{k,-}(x, t_0(x)) &= V_{k,-}(x) = V_{k,-}(-x) \\ &= V_{k,-}(-x, t_0(-x)) = V_{k,+}(x, -t_0(-x)) = V_{k,+}(x, t_0(x))\,. \end{aligned}$$

This yields that $E_{x_0}(t_0(x_0))$ is the desired plane. □

Theorem 1.2.15. *Let $\mathscr{X}$ be a Banach space, $G \subset \mathscr{X}$ open and $f : G \to \mathscr{X}$ continuous and locally injective. Then f is an open map.*

Proof. Assume without loss of generality $0 \in G$ and $f(0) = 0$. For arbitrary $r > 0$ we have to show the existence of $\delta > 0$ such that $\mathbb{B}_\delta(0) \subset f(\mathbb{B}_r(0))$. W.l.o.g. we can assume that f is injective on $\mathbb{B}_r(0)$. Hence

$$h(x,t) := f\left(\frac{1}{1+t}x\right) - f\left(\frac{-t}{1+t}x\right)$$

is a valid homotopy between f and the odd function $h(x,1) = f\left(\frac{1}{2}x\right) - f\left(-\frac{1}{2}x\right)$. Injectivity implies that

$$h(x,t) = 0 \qquad \Leftrightarrow \qquad \frac{1}{1+t}x = -\frac{t}{1+t}x \qquad \Leftrightarrow \qquad x = 0\,.$$

Since $0 \notin f(\partial\mathbb{B}_r(0))$ we obtain from (d4) and Borsuk's Theorem

$$\mathfrak{d}\,(f,G,0)=\mathfrak{d}\,(h(\cdot,1),G,0)\neq 0\,.$$

Hence continuity of $\mathfrak{d}$ implies that $f(x)=y$ has a solution in an open neighborhood of 0. □

Corollary 1.2.16. *Let $G\subset\mathbb{R}^d$ be symmetric open and bounded with $0\in G$. Let ∂G be covered by d closed sets $(A_k)_{k=1,\dots,d}$. Then one of the A_k contains two antipodes x and $-x$.*

Proof. Assume there exists $x\in\bigcap_k A_k$. Since $-x\in A_l$ for some l, we find $x,-x\in A_l$.

Assume the opposite, i.e. $\bigcap_k A_k=\emptyset$. Then we find

$$d_k(x):=\mathrm{dist}(x,A_k)\,,\qquad\text{and}\qquad d(x):=\sum_k d_k(x)>0\,.$$

Consider $f:\,\partial G\to\mathbb{R}^{d-1}$ defined through

$$f(x):=\left(\frac{d_1(x)}{d(x)},\dots,\frac{d_{d-1}(x)}{d(x)}\right)\,.$$

Due to the weather theorem there exists x such that $f(x)=f(-x)$ and $k\in\{1,\dots,d\}$ such that $x\in A_k$. If $k<d$ then $0=d_k(x)=f(-x)d(x)$ and hence $d_k(-x)=0$ which implies $-x\in A_k$. If $x\notin A_k$ for all $k<d$ then $d_k(x)=\frac{d(x)}{d(-x)}d_k(-x)\neq 0$ for all $k<d$ and hence $-x\notin A_k$. This implies $-x,x\in A_d$. □

Chapter 2

Calculus

2.1 Calculus in Infinite Dimensions

Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces. While the degree theory is focused on continuous functions and (after approximation arguments) does not rely on any further regularity of the functions, in this section we will deal with function with higher regularity than just continuity. In particular, we will introduce a concept of differentiability of functions $f : \mathscr{X} \to \mathscr{Y}$.

In what follows, we will write $\mathcal{L}(\mathscr{X}, \mathscr{Y})$ for the space of bounded linear functionals $\mathscr{X} \to \mathscr{Y}$. We recall from linear functional analysis that $\mathcal{L}(\mathscr{X}, \mathscr{Y})$ with norm

$$\|T\| := \|T\|_{\mathcal{L}(\mathscr{X},\mathscr{Y})} := \sup_{x\in\mathscr{X}\setminus\{0\}} \frac{\|Tx\|_{\mathscr{Y}}}{\|x\|_{\mathscr{X}}}$$

is a Banach space.

2.1.1 Derivatives

Definition 2.1.1. A function $f : \mathscr{X} \to \mathscr{Y}$ is (Frechet-) differentiable in $x_0 \in \mathscr{X}$ if there exists $A \in \mathcal{L}(\mathscr{X}, \mathscr{Y})$ such that for

$$\varphi(z) := f(x_0 + z) - f(x_0) - Az$$

it holds

$$\lim_{\|z\|_{\mathscr{X}}\to 0} \frac{\|\varphi(z)\|_{\mathscr{Y}}}{\|z\|_{\mathscr{X}}} = 0\,.$$

The Frechet differentiability is a straight forward generalization of the concept of differentiability in finite dimensional Analysis, while the Gateaux-differentiability is a straight forward generalization of directional derivatives. In the same way as in the finite dimensional case we have uniqueness of the derivative by the simple observation

$$\frac{\|A_1 z - A_2 z\|_{\mathscr{Y}}}{\|z\|_{\mathscr{X}}} \le \frac{\|\varphi_1(z)\|_{\mathscr{Y}}}{\|z\|_{\mathscr{X}}} + \frac{\|\varphi_2(z)\|_{\mathscr{Y}}}{\|z\|_{\mathscr{X}}} \to 0 \qquad \text{as } \|z\|_{\mathscr{X}} \to 0\,,$$

which implies $\|A_1 - A_2\|_{\mathcal{L}(\mathscr{X};\mathscr{Y})} = 0$. Since the derivative of f in $x \in \mathscr{X}$ is unique, we denote it by $\mathrm{D}f(x) := A$. The above concept can be generalized in a straight forward way to higher derivatives: if the function $\mathrm{D}f : \mathscr{X} \to \mathcal{L}(\mathscr{X};\mathscr{Y})$, $x \mapsto \mathrm{D}f(x)$ is differentiable we denote

$$\begin{aligned} \mathrm{D}^2 f :\ \mathscr{X} \to \mathcal{L}(\mathscr{X};\mathcal{L}(\mathscr{X};\mathscr{Y})) \equiv \mathcal{L}(\mathscr{X}\times\mathscr{X};\mathscr{Y}) , \\ \mathrm{D}^2 f(x) :\ (v,w) \mapsto \mathrm{D}^2 f(x)\langle v,w\rangle \ , \end{aligned}$$

and say $f \in C^2(G;\mathscr{Y})$. The above notation might be irritating in view of the second derivative in case $\mathscr{Y} = \mathbb{R}$, which was studied in Analysis II lectures:

$$\mathrm{D}^2 f(x) :\ (v,w) \mapsto \left\langle v, \mathrm{D}^2 f(x) w\right\rangle \ .$$

However, the our new notation is more easy to handle below.

The infinite dimensional derivative satisfies a chain rule in the sense of Analysis II. In particular, if $f : \mathscr{Y} \to \mathscr{Z}$, $g :\ \mathscr{X} \to \mathscr{Y}$ we obtain

$$\mathrm{D}\left(f \circ g\right)(x) = \mathrm{D}f\left(g(x)\right) \circ \mathrm{D}g(x) \,.$$

This can be shown by a similar calculation as in Analysis II. In particular, if

$$\begin{aligned} \varphi_f(z) &:= f(g(x_0)+z) - f(g(x_0)) - \mathrm{D}f(g(x_0))z \\ \varphi_g(y) &:= g(x_0+y) - g(x_0) - \mathrm{D}g(x_0)y \end{aligned}$$

then

$$\begin{aligned} f\left(g(x_0+y)\right) - f\left(g(x_0)\right) &= \mathrm{D}f\left(g(x_0)\right)\left(g(x_0+y)-g(x_0)\right) + \varphi_f\left(g(x_0+y)-g(x_0)\right) \\ &= \mathrm{D}f\left(g(x_0)\right)\left(\mathrm{D}g(x_0)y + \varphi_g(y)\right) + \varphi_f\left(g(x_0+y)-g(x_0)\right) \end{aligned}$$

and continuity of g and $\mathrm{D}f\left(g(x_0)\right)$ it follows that

$$\begin{aligned} &\lim_{y\to 0} \frac{1}{\|y\|_{\mathscr{X}}} \left\|\mathrm{D}f\left(g(x_0)\right)\varphi_g(y) + \varphi_f\left(g(x_0+y)-g(x_0)\right)\right\| \\ &\quad \le \lim_{y\to 0} \frac{\|\varphi_g(y)\|}{\|y\|_{\mathscr{X}}} \left\|\mathrm{D}f\left(g(x_0)\right)\right\| + \lim_{y\to 0} \frac{\|g(x_0+y)-g(x_0)\|}{\|y\|_{\mathscr{X}}} \frac{\left\|\varphi_f\left(g(x_0+y)-g(x_0)\right)\right\|}{\|g(x_0+y)-g(x_0)\|} \\ &\quad \to 0+0\,. \end{aligned}$$

The chain rule also opens the door to the definition of directional derivatives. In particular, if $g : \mathbb{R} \to \mathscr{X}$ is given by $t \mapsto x_0 + tx$. Then we obtain for every differentiable $f \in C^1(\mathscr{X};\mathscr{Y})$ $\mathrm{D}(f\circ g) = (\mathrm{D}f\circ g)\,\mathrm{D}g = (\mathrm{D}f\circ g)\,x$. In the particular case $\mathscr{Y} = \mathbb{R}$ we find $\mathrm{D}f(x_0) \in \mathscr{X}^*$ and hence

$$\mathrm{D}(f\circ g) = \langle \mathrm{D}f(x_0),\, x\rangle_{\mathscr{X}^*,\mathscr{X}} \ ,$$

where $\langle \mathrm{D}f(x_0),\, x\rangle_{\mathscr{X}^*,\mathscr{X}}$ is the dual pairing between $\mathrm{D}f(x_0)$ and x.

Definition 2.1.2. A function $f :\ \mathscr{X} \to \mathscr{Y}$ is called Gateaux-Differentiable if for all $g : \mathbb{R} \to \mathscr{X}$ given by $t \mapsto x_0 + tx$, the directional derivative $\mathrm{D}(f\circ g)$ exists.

In what follows, we will only deal with the concept of Frechet-differentiability. We close this section by the following

Lemma 2.1.3. *Let $\mathscr{X}$, $\mathscr{Y}$ be Banach spaces and let $f \in C(\mathscr{X};\mathscr{Y})$ be compact in an open ball $\mathbb{B}_\varepsilon^{\mathscr{X}}(x_0)$. Then $\mathrm{D}f(x_0) \in \mathcal{L}(\mathscr{X};\mathscr{Y})$ is compact.*

Proof. W.l.o.g. let $x_0 = 0$ and $f(x_0) = 0$. Writing $A := \mathrm{D}f(x_0)$, we have to show that every bounded sequence $(x_k)_{k\in\mathbb{N}}$ has a subsequence $x_{k'}$ such that $Ax_{k'}$ converges. The differentiability of f implies for

$$\psi(x) := \varphi(x)\,\|x\|^{-1} := \|x\|^{-1}\,(f(x) - Ax)$$

that $\psi(x) \to 0$ as $x \to 0$. Hence for every $n \in \mathbb{N}$ there exists $\delta(n)$ such that

$$\|f(\delta(n)x_k) - \delta(n)Ax_k\|_{\mathscr{Y}} \leq \frac{1}{n}\delta(n)\,.$$

By compactness of f we can extract a subsequence x_{k_n} such that $f(\delta(n)x_{k_n})$ converges and for k_n, l_n big enough we find

$$\begin{aligned}\frac{1}{n}\delta(n) &\geq \|f(\delta(n)x_{k_n}) - f(\delta(n)x_{l_n})\|_{\mathscr{Y}}\\ &\geq \|\delta(n)Ax_{k_n} - \delta(n)Ax_{l_n}\|_{\mathscr{Y}} - \frac{2}{n}\delta(n)\,,\end{aligned}$$

which implies

$$\|Ax_{k_n} - Ax_{l_n}\|_{\mathscr{Y}} \leq 3\frac{1}{n}\,.$$

using a Cantor argument, we obtain that $Ax_k \to y$ along a subsequence. □

A warning is in place here by the following

Lemma. *Let $\mathscr{X}$ be the set of Null-sequences with the absolute norm $\|\cdot\|_\infty$, i.e.*

$$\mathscr{X} := \left\{(x_k)_{k\in\mathbb{N}}\,:\,\lim_{k\to\infty} x_k = 0\right\}\,,\qquad \left\|(x_k)_{k\in\mathbb{N}}\right\|_\infty := \sup_{k\in\mathbb{N}} |x_k|\,,$$

and let $F:\mathscr{X}\to\mathscr{X}$, $x = (x_k)_{k\in\mathbb{N}} \mapsto (x_k^2)_{k\in\mathbb{N}}$. Then $F \in C^1(\mathscr{X};\mathscr{X})$ and for every $x \in \mathscr{X}$ the map $\mathrm{D}f(x)$ is compact, but F is not compact.

Proof. Let $e^i \in \mathscr{X}$ be the sequence with 1 at the i-th place and 0 else. Then $F(e^i) = e^i$ and $\|e^i - e^j\|_\infty = 1 - \delta_{ij}$. In particular, $\{F(e^i)\}_{i\in\mathbb{N}}$ does not have a convergent subsequence and hence is not compact.

However, it is differentiable:

$$\begin{aligned}F(x+y) &= \left([x_k + y_k]^2\right)_{k\in\mathbb{N}} = \left(x_k^2\right)_{k\in\mathbb{N}} + (2x_ky_k)_{k\in\mathbb{N}} + \left(y_k^2\right)_{k\in\mathbb{N}}\\ &= F(x) + (2x_ky_k)_{k\in\mathbb{N}} + F(y)\,.\end{aligned}$$

Now, observe that

$$\lim_{y\to 0}\frac{1}{\|y\|_\infty}\|F(y)\|_\infty = \lim_{y\to 0}\frac{\sup_k (y_k^2)}{\sup_k |y_k|} = \lim_{y\to 0}\sup_k |y_k| = 0\,.$$

Hence $\mathrm{D}f(x)y = (2x_ky_k)_{k\in\mathbb{N}}$. The operator $\mathrm{D}f(x)$ is compact because for

$$A_K y := (2a_ky_k)_{k\in\mathbb{N}}\,,\qquad a_k := \begin{cases} x_k & k \leq K\\ 0 & \text{else}\end{cases}$$

it holds

$$\|y\|_\infty^{-1}\,\|(A_K - \mathrm{D}f(x))\,y\| \leq \sup_{k>K}|x_k| \to 0 \qquad \text{as } K \to 0\,,$$

i.e. $\mathrm{D}f(x)$ can be approximated by finite dimensional A_K. □

2.1.2 The Euler-Lagrange equation

We consider a standard minimization problem: Let $u : [-1,1] \to \mathbb{R}$ be a continuously differentiable function satisfying $u(-1) = u(1) = 0$. More precisely, we consider

$$H_0^1(-1,1) := \left\{u \in H^1(-1,1) \,:\, u(-1) = u(1) = 0\right\} .$$

The length of the curve is given by

$$L(u) := \int_{-1}^{1} \sqrt{1 + |\partial_x u(s)|^2} \mathrm{d}s \,,$$

while the (signed) area between u and the s-axis is given by

$$F(u) = \int_{-1}^{1} u(s) \mathrm{d}s \,.$$

We want to maximize $L(u)$ under given $F \equiv F(u)$. Note that for any two functions $u_1, u_2 \in H_0^1(-1,1)$ with $F(u_1) = F(u_2)$ this implies $\int_{-1}^{1} u_1(s) - u_2(s)\mathrm{d}s = 0$. In particular, for given $u_0 \in H_0^1(-1,1)$ with $F(u_0) = F_0$ it suffices to maximize $L(u_0 + v)$ with respect to

$$v \in H^1_{(0),0} := \left\{ v \in H_0^1(-1,1) \,:\, \int_{-1}^{1} v(s)\mathrm{d}s = 0 \right\} .$$

In such a maximum $v_0 \in H^1_{(0),0}$ it holds

$$\forall v \in H^1_{(0),0} : \qquad \mathrm{D}L(u_0 + v_0)(v) = 0 \,. \tag{2.1}$$

Since

$$\mathrm{D}L(u_0 + v_0)(v) = \int_{-1}^{1} \frac{\partial_x u(s) \partial_x v(s)}{\sqrt{1 + |\partial_x u(s)|^2}} \mathrm{d}s \,, \qquad u = u_0 + v_0 \,,$$

(2.1) implies that $\partial_x \frac{\partial_x u(s)}{\sqrt{1+|\partial_x u(s)|^2}} = K_0$. Since $u(-1) = u(1) = 0$ we can furthermore assume $u(-1 + x) = u(1 - x)$ and integration over x yields

$$u(x) = \frac{1}{K_0} \sqrt{1 - K_0^2 x^2} \,,$$

and hence u describes a circle.

2.1.3 The Implicit function theorem

Theorem 2.1.4. *Let $\mathscr{X}$, $\mathscr{Y}$, $\mathscr{Z}$ be Banach spaces, let*

$$f : \mathscr{X} \times \mathscr{Y} \supset U \to \mathscr{Z}$$

be continuous in U and let $(x_0, y_0) \in U$ be such that $f(x_0, y_0) = 0$. Assume that $\mathrm{D}_{\mathscr{X}} f$ exists and is continuous a neighborhood of (x_0, y_0) and that $\mathrm{D}_{\mathscr{X}} f(x_0, y_0)$ is an isomorphism. Then

1. *There exists $r>0$ and a uniquely determined function $u:\mathbb{B}_r^{\mathscr{Y}}(y_0)\to\mathscr{X}$ such that*
$$u(y_0)=x_0 \quad \textit{and} \quad \forall y\in\mathbb{B}_r^{\mathscr{Y}}(y_0):\quad f\left(u(y),y\right)=0\,.$$
2. *For every $p\geq 1$ holds $f\in C^p(U;\mathscr{Z})$ implies $u\in C^p(\mathbb{B}_r^{\mathscr{Y}}(y_0);\mathscr{X})$ and*
$$\mathrm{D}_{\mathscr{Y}}u(y)=-\left[\mathrm{D}_{\mathscr{X}}f\left(u(y),y\right)\right]^{-1}\mathrm{D}_{\mathscr{Y}}f\left(u(y),y\right)\,.$$

In order to prove this theorem, we need the following Lemma.

Lemma 2.1.5. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces and let $A:\mathscr{X}\to\mathscr{Y}$ be a linear isomorphism. Then for every linear $B:\mathscr{X}\to\mathscr{Y}$ with $\|B\|\left\|A^{-1}\right\|\leq 1$ the operator $A+B$ is an isomorphism.*

Proof. Writing $\tilde{B}=A^{-1}B$ we may assume that $A=\mathrm{id}$. If $y_1+\tilde{B}y_1=y_2+\tilde{B}y_2$ we obtain the contradiction$\|y_1-y_2\|=\|By_1-By_2\|\leq\|B\|\,\|y_1-y_2\|<\|y_1-y_2\|$. Hence $\mathrm{id}+B$ is an isomorphism. The inverse operator theorem states that $(\mathrm{id}+B)^{-1}$ is continuous. □

Now we prove Theorem 2.1.4.

Proof. W.l.o.g. $x_0=0$ and $y_0=0$. For $A:=\mathrm{D}_{\mathscr{X}}f(0)$ we consider the residual estimate for
$$f(x,y)=0 \qquad\Leftrightarrow\qquad \begin{aligned}Ax&=-f(x,y)+Ax=:R(x,y)=\varphi_y(x)\\ x&=A^{-1}R(x,y)=:g(x,y)\,.\end{aligned}$$

With the aim to apply Banach's fixed point theorem, we show that $g(\cdot,y)$ is a contraction for $y\in B_r(0)$.

For $\varepsilon>0$ with $\varepsilon\left\|A^{-1}\right\|\leq\frac{1}{2}$ we calculate
$$\begin{aligned}R(x_1,y)-R(x_2,y)&=A(x_1-x_2)-\left(f(x_1,y)-f(x_2,y)\right)\\ &=A(x_1-x_2)-\int_0^1\mathrm{D}_{\mathscr{X}}f\left(tx_1-(1-t)x_2\right)\mathrm{d}t\,(x_1-x_2)\\ &=\int_0^1\left[\mathrm{D}_{\mathscr{X}}f(0,0)-\mathrm{D}_{\mathscr{X}}f\left(tx_1-(1-t)x_2\right)\mathrm{d}t\right](x_1-x_2)\end{aligned}$$

Since $\mathrm{D}_{\mathscr{X}}f$ is continuous, we find $r,\delta>0$ such that $\|A-\mathrm{D}_{\mathscr{X}}f\left(x,y\right)\|<\varepsilon$ for $\|x\|_{\mathscr{X}}<\delta$ and $\|y\|_{\mathscr{X}}<r$. Then there holds
$$\begin{aligned}\|g(x_1,y)-g(x_2,y)\|&\leq\left\|A^{-1}\right\|\|R(x_1,y)-R(x_2,y)\|\\ &\leq\frac{1}{2}\|x_1-x_2\|\,.\end{aligned}$$

If r is small enough there holds
$$\|g(0,y)\|\leq\frac{1}{2}\delta$$
and hence
$$\begin{aligned}\|g(x,y)\|&\leq\|g(x,y)-g(0,y)\|+\|g(0,y)\|\\ &\leq\frac{1}{2}\|x-0\|+\frac{1}{2}\delta<\delta\,.\end{aligned}$$

In particular, $g(\cdot,y):\mathbb{B}_\delta^{\mathscr{X}}\to\mathbb{B}_\delta^{\mathscr{X}}$ is a contraction for every $y\in\mathbb{B}_r^{\mathscr{Y}}(0)$. This yields unique existence of a fixed point $x=:u(y)$ of $g(\cdot,y)$.

Since $g(\cdot,y)$ is a contraction, we obtain

$$\begin{aligned}\|u(y_1)-u(y_2)\| &\leq \|g(u(y_1),y_1)-g(u(y_2),y_2)\|\\ &\leq \|g(u(y_1),y_1)-g(u(y_2),y_1)\|+\|g(u(y_2),y_1)-g(u(y_2),y_2)\|\\ &\leq \frac{1}{2}\|u(y_1)-u(y_2)\|+\|g(u(y_2),y_1)-g(u(y_2),y_2)\|\,.\end{aligned}$$

The first term on the right hand side can be adsorbed on the left hand side, while the second term on the right hand side converges to 0 for $y_1-y_2\to 0$ due to the continuity of g. Hence the first part of the theorem is proved.

In order to prove differentiability of u, let y and $y+\delta_y$ be in $\mathbb{B}_r^{\mathscr{Y}}(0)$ and let $\delta_u=u(y+\delta_y)-u(y)$. Since f is differentiable, we obtain

$$\left\|f\left(u(y+\delta_y),y+\delta_y\right)-f\left(u(y),y\right)-\mathrm{D}_{\mathscr{X}}f(u(y),y)\delta_u-\mathrm{D}_{\mathscr{Y}}f(u(y),y)\delta_y\right\|\leq\varphi_f(\delta_u,\delta_y)\,.$$

By the characterization of u, the first two terms in the norm vanish and we have

$$\|-\mathrm{D}_{\mathscr{X}}f(u(y),y)\delta_u-\mathrm{D}_{\mathscr{Y}}f(u(y),y)\delta_y\|\leq\varphi_f(\delta_u,\delta_y)\,.$$

By continuity of $\mathrm{D}_{\mathscr{X}}f(x,y)$ in $(0,0)$ and Lemma 2.1.5 it follows that $\mathrm{D}_{\mathscr{X}}f(u(y),y)$ is invertible for small r with bounded inverse. Hence, applying $\mathrm{D}_{\mathscr{X}}f(u(y),y)^{-1}$ to the above calculation yields

$$\left\|\delta_u+\mathrm{D}_{\mathscr{X}}f(u(y),y)^{-1}\mathrm{D}_{\mathscr{Y}}f(u(y),y)\delta_y\right\|\leq\left\|\mathrm{D}_{\mathscr{X}}f(u(y),y)^{-1}\right\|\varphi_f(\delta_u,\delta_y)\,.$$

We write $B:=-\mathrm{D}_{\mathscr{X}}f(u(y),y)^{-1}\mathrm{D}_{\mathscr{Y}}f(u(y),y)$ as well as $A_y:=\mathrm{D}_{\mathscr{X}}f(u(y),y)$ and find

$$\|\delta_u-B\delta_y\|\leq\left\|A^{-1}\right\|\frac{\varphi_f(\delta_u,\delta_y)}{\|\delta_u\|_{\mathscr{X}}+\|\delta_y\|_{\mathscr{Y}}}\left(\|\delta_u-B\delta_y\|_{\mathscr{X}}+(1+\|B\|)\|\delta_y\|_{\mathscr{Y}}\right)\,.$$

Due to differentiability of f there exists r small enough such that

$$\frac{\|A^{-1}\|\,\varphi_f(\delta_u,\delta_y)}{\|\delta_u\|_{\mathscr{X}}+\|\delta_y\|_{\mathscr{Y}}}\leq\frac{1}{2}\,,$$

and we obtain

$$\|u(y+\delta_y)-u(y)-B\delta_y\|\leq 2\left\|A^{-1}\right\|\frac{\varphi_f(\delta_u,\delta_y)}{\|u(y+\delta_y)-u(y)\|_{\mathscr{X}}+\|\delta_y\|_{\mathscr{Y}}}(1+\|B\|)\|\delta_y\|_{\mathscr{Y}}\,.$$

The right hand side has the property that

$$\lim_{\delta_y\to 0}2\left\|A^{-1}\right\|\frac{\varphi_f(\delta_u,\delta_y)}{\|u(y+\delta_y)-u(y)\|_{\mathscr{X}}+\|\delta_y\|_{\mathscr{Y}}}(1+\|B\|)=0$$

and hence $y\to u(y)$ is differentiable. Higher differentiability follows from the representation of $\mathrm{D}_{\mathscr{Y}}u$. □

Theorem 2.1.6 (Inverse function theorem)**.** *Let* $f : \mathscr{X} \supset U \to \mathscr{Y}$ *be* C^p*,* $p \geq 1$*,* $x_0 \in U$ *and* $f(x_0) = y_0$*. If* $\mathrm{D}_{\mathscr{X}} f(x_0)$ *is an isomorphism, then there exists* $r > 0$ *and a uniquely determined continuous function* $u : \mathbb{B}_r^{\mathscr{X}}(y_0) \to \mathscr{X}$ *with* $u(y_0) = x_0$ *and such that* $f(u(y)) = y$*. The inverse is of class* C^p*.*

Proof. Write $F(x, y) = f(x) - y$ and solve $F(x, y) = 0$. It holds $\mathrm{D}_{\mathscr{X}} F(x_0, y_0) = \mathrm{D}_{\mathscr{X}} f(x_0)$ and apply the implicit function theorem. □

Example (A counterexample)**.** Consider $\alpha \in (0, 1)$ and

$$f(x) := \begin{cases} \alpha x + x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases} .$$

One may verify that f is differentiable in $\mathbb{R}$ with $f'(0) \neq 0$ but f is not invertible in any environment of 0. Why is this not a contradiction to the inverse function theorem?

Theorem 2.1.7 (Closed complement)**.** *Let* $\mathscr{X}$ *be a Banach space and* $\mathscr{Y} \subset \mathscr{X}$ *a closed subspace and* $Z \subset \mathscr{X}$ *a subspace with* $\mathscr{X} = \mathscr{Y} + Z$*. Then the following are equivalent:*

1. *There exists a continuous projection* $P \in \mathcal{L}(\mathscr{X})$ *with* $R(P) = \mathscr{Y}$ *and* $\ker P = Z$*.*
2. Z *is closed.*

Based on Theorem 2.1.7 we make the following general statement.

Theorem 2.1.8 (Generalized inverse function theorem)**.** *Let* $f : \mathscr{X} \supset U \to \mathscr{Y}$ *be* C^1*,* $\tilde{x} \in U$ *and* $f(\tilde{x}) = \tilde{y}$*. If* $\mathrm{D}_{\mathscr{X}} f(\tilde{x})$ *is surjective, then there exists* $\delta, r > 0$ *such that for every* $y \in \mathbb{B}_r(\tilde{y})$ *there exist* $x \in \mathbb{B}_\delta(\tilde{x})$ *with* $f(x) = y$*.*

Proof. W.l.o.g. let $\tilde{y} = 0$ and $\tilde{x} = 0$. Write $A = \mathrm{D}_{\mathscr{X}} f(\tilde{x})$. The kernel $\mathscr{X}_0$ of A is closed, hence Theorem 2.1.7 yields the existence of projections $P : \mathscr{X} \to \mathscr{X}_0$ and $Q = 1 - P : \mathscr{X} \to \mathscr{X}_1 := Q\mathscr{X}$ such that $\mathscr{X}_1$ is closed, $A|_{\mathscr{X}_1} \to \mathscr{Y}$ is bijective and the splitting $x = x_0 + x_1$ with $x_0 = Px$ is well defined. By the implicit function theorem, there exists $g : \mathscr{X}_0 \to \mathscr{X}_1$ such that $g(0) = 0$ and

$$f(x_0, x_1) = 0 \qquad \Leftrightarrow \qquad x_1 = g(x_0) .$$

Writing $F(x_0, x_1) := f(x_0, g(x_0) + x_1)$ we see $F(x_0, 0) = 0$ in a neighborhood of 0. Furthermore, $F(0, \cdot)$ is C^1 with $D_{\mathscr{X}_1} F(0, 0) = A$, which is invertible. Hence we find a unique solution of $F(0, x_1) = y$ in a neighborhood of $(0, 0) \in \mathscr{X}_1 \times \mathscr{Y}$. □

2.1.4 Continuous dependence of ODE-solutions on Data

We consider the following problem

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} u(t) &= f(t, u(t), \lambda) , \\ u(\sigma) &= \xi , \end{aligned}$$

where $\xi \in \mathscr{X}$, $\sigma \in \mathbb{R}$, $\lambda \in \Lambda$, where $\mathscr{X}$ and Λ are Banach spaces.

Theorem 2.1.9. *Let $f : \mathbb{R} \times \mathscr{X} \times \Lambda \to \mathscr{X}$ be continuous and such that $\mathrm{D}_{\mathscr{X}} f$ exists and is continuous in a neighborhood of $(\sigma_0, \xi_0, \lambda_0)$. Then there exists $T > 0$ and a neighborhood U of $(\sigma_0, \xi_0, \lambda_0)$ such that for all $(\sigma, \xi, \lambda) \in U$ there exists a solution $u(\cdot; \sigma, \xi, \lambda)$ on $(\sigma - T, \sigma + T)$. If $f \in C^k$ then*

$$u \in C^k \left((\sigma - T, \sigma + T) \times U; \mathscr{X}\right) .$$

Proof. We consider the reparametrization $u(t) = \xi + U\left(\frac{t-\sigma}{T}\right)$ and observe that the equation for U reads

$$\frac{\mathrm{d}}{\mathrm{d}\tau} U(\tau) = T f\left(\sigma + T\tau, \xi + U(\tau), \lambda\right)$$

with $U(0) = 0$. We define

$$\begin{aligned}
\mathscr{A} &= \left\{U \in C^1([-1,1]; \mathscr{X}) \,:\, U(0) = 0\right\} ,\\
\mathscr{Y} &= \mathbb{R} \times \mathbb{R} \times \mathscr{X} \times \Lambda ,\\
\mathscr{B} &= C^0([-1,1]; \mathscr{X}) ,
\end{aligned}$$

and

$$\begin{aligned}
F : \qquad & \mathscr{A} \times \mathscr{Y} \to \mathscr{B} \\
(U, (T, \sigma, \xi, \lambda)) &\mapsto \frac{\mathrm{d}}{\mathrm{d}\tau} U - T f\left(\sigma + T\tau, \xi + U(\tau), \lambda\right)
\end{aligned} .$$

The theorem is proved if we can solve the equation $F\left(U, (T, \sigma, \xi, \lambda)\right) = 0$ for U. Hence, we calculate

$$\mathrm{D}_{\mathscr{A}} F\left(0, (0, \sigma_0, \xi_0, \lambda_0)\right) \,:\, \mathscr{A} \to \mathscr{B} , \quad U \mapsto \frac{\mathrm{d}}{\mathrm{d}\tau} U ,$$

where we used $T = 0$. Therefore, $\mathrm{D}_{\mathscr{A}} F\left(0, (0, \sigma_0, \xi_0, \lambda_0)\right)$ is an isomorphism (due to initial condition $U(0) = 0$). The implicit function theorem yields $U(T, \sigma, \xi, \lambda)$ and the differentiability properties. □

2.1.5 Global inverse

Theorem 2.1.10. *Let $f \in C^1(\mathscr{X}; \mathscr{Y})$ such that for every $x \in \mathscr{X}$: $(\mathrm{D}_{\mathscr{X}} f(x))^{-1} \in \mathcal{L}(\mathscr{Y}; \mathscr{X})$ and $\left\|(\mathrm{D}_{\mathscr{X}} f(x))^{-1}\right\| \leq \alpha \|x\| + \beta$. then f is a diffeomorphism, i.e. there exists a global inverse $g \in C^1(\mathscr{Y}; \mathscr{X})$ and $f \in C^k$ implies $g \in C^k$.*

Proof. Step 1: Let $x_1, x_2 \in \mathscr{X}$ with $f(x_1) = f(x_2) = 0$. Furthermore, let $\chi(t) := t x_1 + (1-t) x_2$ and $\gamma(t) := f(\chi(t))$. We observe that $\gamma(0) = \gamma(1) = 0$ and for continuous u with $u(0) = u(1) = 0$ the same holds for $f(\chi + u)$. Hence we consider the spaces of cycles in $\mathscr{X}$ and $\mathscr{Y}$:

$$\begin{aligned}
\mathscr{C}_{\mathscr{X}} &:= \left\{u \in C([0,1]; \mathscr{X}) \,:\, u(0) = u(1) = 0\right\} ,\\
\mathscr{C}_{\mathscr{Y}} &:= \left\{w \in C([0,1]; \mathscr{Y}) \,:\, w(0) = w(1) = 0\right\} ,
\end{aligned}$$

with the function

$$F : \mathscr{C}_{\mathscr{X}} \to \mathscr{C}_{\mathscr{Y}} , \quad F(u) := f(u + \chi) .$$

This function is differentiable with derivative

$$\forall v \in \mathscr{C}_{\mathscr{X}} \;:\quad \mathrm{D}_{\mathscr{C}_{\mathscr{X}}} F(u)\, v = \mathrm{D}_{\mathscr{X}} f(u + \chi)\, v .$$

Hence we infer that

$$\left\|(\mathrm{D}_{\mathscr{X}}F(u))^{-1}\right\| = \left\|(\mathrm{D}_{\mathscr{X}}f(u+\chi))^{-1}\right\| \leq \alpha\,\|u\| + \tilde{\beta}\,,$$

where $\tilde{\beta}$ depends only on α, β and χ. We then consider the homotopy $y(\tau) := \tau\gamma$ in $\mathscr{C}_{\mathscr{Y}}$ and **assume** there exists a continuous $u : \tau \to \mathscr{C}_{\mathscr{X}}$, $\tau \mapsto u(\tau)$ such that for every $\tau \in [0,1]$: $F(u(\tau)) = y(\tau)$. In $\tau = 0$ we then obtain

$$0 = F(u(0)) = f(u(0)+\chi)\,,$$

where $(u(0)+\chi)\,(0) = x_1$, $(u(0)+\chi)\,(1) = x_2$. Since the path $(u(0)+\chi)$ is continuous, we obtain that either $x_1 = x_2$ or that f is not locally invertible around $(x_1, 0)$ or $(x_2, 0)$ (which is a contradiction to the inverse function theorem).

Step 2: In what follows, we consider $\mathscr{X}$ and $\mathscr{Y}$ but according to Step 1, the result directly applies to $\mathscr{X} = \mathscr{C}_{\mathscr{X}}$ and $\mathscr{Y} = \mathscr{C}_{\mathscr{Y}}$. W.l.o.g. assume $f(0) = 0$. Given $y_0 \in \mathscr{Y}$ consider $y(t) := t\,y_0$ and the set

$$M := \left\{T \in (0,\infty) \,:\, \exists!\, x \in C^1([0,T];\mathscr{X}) \,:\, x(0) = 0,\, \forall t \in [0,T]\; f(x(t)) = y(t)\right\}\,.$$

We show that M is open, non-empty and closed, hence $M = (0,\infty)$. Together with Step 1, this proves the theorem.

First note that M is open and non-empty due to the inverse function theorem: For T small enough, the inverse of f exists on $\mathbb{B}^{\mathscr{Y}}_{T\|y_0\|}(0)$ and hence M is not empty. Given $T_0 \in M$ with $x(t)$ and $f(x(T_0)) = y(T_0)$, the local inverse of f in $T_0 y_0$ exists and hence we can prolongate x to $\tilde{T} > T_0$ in a unique way, i.e. M is open.

We now show that M is closed: It holds

$$x(t) = \int_0^t x'(s)\,\mathrm{d}s = \int_0^t \left(\mathrm{D}_{\mathscr{X}} f(x)\right)^{-1} y_0 \,\mathrm{d}s$$

$$\Rightarrow \quad \|x(t)\|_{\mathscr{X}} \leq \|y_0\|_{\mathscr{Y}} \int_0^t \left(\alpha \|x(s)\|_{\mathscr{X}} + \beta\right) \mathrm{d}s\,,$$

which implies that $x(t)$ remains bounded in finite time. Now let $(t_k)_{k\in\mathbb{N}} \subset M$ be a sequence with $t_k \to T$. If $C = c_0 \exp\left(\alpha \|y_0\|_{\mathscr{X}} 2T\right)$ we find $\|x'\|_{\mathscr{X}} \leq \alpha C + \beta$ on $(0,T)$ and hence $x(t_k)$ is a Cauchy sequence, $x(t_k) \to x^*$. Since f is continuous we obtain $f(x^*) = y_0 T$ and hence $T \in M$. □

2.1.6 Lagrange multipliers

We recall the standard minimization problem from Section 2.1.2: Let $u : [0,1] \to \mathbb{R}$ be a continuously differentiable function satisfying $u(0) = u(1) = 0$. The length of the curve is given by

$$L(u) := \int_0^1 \sqrt{1 + |\partial_x u(s)|^2}\mathrm{d}s\,,$$

while the (signed) area between u and the s-axis is given by

$$F(u) = \int_0^1 u(s)\mathrm{d}s\,.$$

A standard problem in geometric measure theory is: Minimize $L(u)$ for given $F(u)$. We know similar problems from Analysis II in finite dimensional spaces. There, the problem is solved using Lagrange multipliers. Our aim is to generalize this ansatz to the infinite dimensional setting.

Hilbert theory (skip in lecture)

Theorem 2.1.11 (Existence of Lagrange multipliers.). *Let $\mathscr{X}$, $\mathscr{Y}$ be Hilbert spaces and let $M \subset \mathscr{X}$ open. Let $\mathcal{E} : M \to \mathbb{R}$ and $\Phi : M \to \mathscr{Y}$ be continuously differentiable. Let $M_0 := \{x \in M \,:\, \Phi(x) = 0\}$ and let $m_0 \in M_0$ be an extremum of $\mathcal{E}$ in M_0. If $\mathrm{D}_{\mathscr{X}}\Phi(m_0) : \mathscr{X} \to \mathscr{Y}$ is surjective, then there exists $\lambda \in \mathscr{Y}$ such that*

$$\forall x \in \mathscr{X} \,:\qquad \mathrm{D}_{\mathscr{X}}\mathcal{E}(m_0)x = \left\langle \lambda\,,\, \mathrm{D}_{\mathscr{X}}\Phi(m_0)x \right\rangle\,. \tag{2.2}$$

The proof makes use of the following results from linear functional analysis.

1. Let $\mathscr{X}$ be a Hilbert space, then every $f \in \mathscr{X}'$ can be represented by $\lambda \in \mathscr{X}$. In particular, we obtain $f(x) = \langle \lambda, x\rangle$.

2. Let $\mathscr{X}_0 \subset \mathscr{X}$ be a closed subspace of the Hilbert space $\mathscr{X}$. Then the orthogonal space $\mathscr{X}_1 \perp \mathscr{X}_0$ is also closed and $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$.

3. The inverse functional theorem: If $\mathscr{X}$, $\mathscr{Y}$ are Banach spaces and $A \in \mathcal{L}(\mathscr{X};\mathscr{Y})$ is bijective, then $A^{-1} \in \mathcal{L}(\mathscr{Y};\mathscr{X})$.

Proof. We consider $\mathscr{X}_0 := \ker \mathrm{D}_{\mathscr{X}}\Phi(m_0) \subset \mathscr{X}$, which is a closed linear subspace of $\mathscr{X}$. Hence $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$ for the orthogonal closed spaces $\mathscr{X}_1 \perp \mathscr{X}_0$ and every x decomposes as $x = x_0 + x_1$. Since $A := \mathrm{D}_{\mathscr{X}}\Phi(m_0)|_{\mathscr{X}_1}$ is bijective, we find $A^{-1} \in \mathcal{L}(\mathscr{Y};\mathscr{X}_1)$.

Hence, the linear operator $f := \mathrm{D}_{\mathscr{X}}\mathcal{E}(m_0) \circ A^{-1} : \mathscr{Y} \to \mathbb{R}$ is continuous and by 1. can be represented as $\lambda \in \mathscr{Y}$: $f(y) = \langle \lambda, y \rangle$. In particular, setting $y = A^{-1}x_1$ we find

$$\forall x_1 \in \mathscr{X}_1 : \quad \mathrm{D}_{\mathscr{X}}\mathcal{E}(m_0)x_1 = \langle \lambda, Ax_1 \rangle \ .$$

For $x_0 \in \mathscr{X}_0$ it holds $\langle \lambda, \mathrm{D}_{\mathscr{X}}\Phi(m_0)x_0 \rangle = 0$ and in view of (2.2) it remains to show $\mathrm{D}_{\mathscr{X}}\mathcal{E}(m_0)x_0 = 0$. Indeed, for any path γ with $\gamma(0) = m_0$ and $\gamma'(0) = x_0$ we obtain by the maximum condition

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{E} \circ \gamma\right)|_{t=0} = \mathrm{D}_{\mathscr{X}}\mathcal{E}(m_0)x_0$$

and it only remains to prove the existence of such a path.

Let $\Psi(x_0, x_1) = \Phi(m_0 + x_0 + x_1)$ with $\mathrm{D}_{\mathscr{X}_1}\Psi(0) = \mathrm{D}_{\mathscr{X}}\Phi(m_0)|_{\mathscr{X}_1} = A$. Since A is invertible, we can resolve $x_1 = x_1(x_0)$ and $\gamma(t) = m_0 + tx_0 + x_1(tx_0)$. Then

$$\gamma'(t) = x_0 - A^{-1}\mathrm{D}_{\mathscr{X}}\Phi(m_0)|_{\mathscr{X}_0} = x_0 \,,$$

and the theorem is proved. □

The Theorem 2.1.11 is valid also in general reflexive Banach spaces. However, the general proof is much more involved than the one we provided above. Therefore, we will restrict ourselves to the following result.

Theorem 2.1.12. *Let $\mathscr{X}$ be a reflexive Banach space and $\mathscr{Y}$ be finite dimensional Banach space and let $M \subset \mathscr{X}$ open. Let $\mathcal{E} : M \to \mathbb{R}$ and $\Phi : M \to \mathscr{Y}$ be continuously differentiable. Let $M_0 := \{x \in M \,:\, \Phi(x) = 0\}$ and let $x_0 \in M_0$ be an extremum of $\mathcal{E}$ in M_0. If $\mathrm{D}_{\mathscr{X}}\Phi(x_0) : \mathscr{X} \to \mathscr{Y}$ is surjective, there exists $\lambda \in \mathscr{Y}$ such that*

$$\forall x \in \mathscr{X} : \qquad \mathrm{D}_{\mathscr{X}}\mathcal{E}(x_0)x = \langle \lambda \,,\, \mathrm{D}_{\mathscr{X}}\Phi(x_0)x \rangle \ .$$

In the following proof, we will again use 1. and 3. from above and replace 2. with the following Lemma. The most general case of the Lagrange-Multiplier theorem can be found in the book by Kunisch and Ito on *Lagrange Multiplier Approach to Variational Problems and Applications.*

Lemma 2.1.13. *Let $\mathscr{X}$ be a Banach space, let $A : \mathscr{X} \to \mathbb{R}^d$ be surjective and denote $\mathscr{X}_0 := \ker A \subset \mathscr{X}$. Then there exists a d -dimensional $\mathscr{X}_1 \subset \mathscr{X}$ such that $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$ and a continuous projection $P : \mathscr{X} \to \mathscr{X}_1$ such that* $\mathrm{id} - P : \mathscr{X} \to \mathscr{X}_0$.

Proof. Let $(e_i)_{i=1,\dots,d}$ be the orthogonal basis of $\mathbb{R}^d$ and chose $x_i \in A^{-1}(e_i)$. Then $\tilde{A}^{-1} : e_i \mapsto x_i$ is continuous and define $\mathscr{X}_1 := \operatorname{span}(x_i)_{i=1,\dots,d}$ the operator $P := \tilde{A}^{-1} \circ A : \mathscr{X} \to \mathscr{X}_1$ is continuous. Moreover $A(x - Px) = 0$, hence $x - Px \in \mathscr{X}_0$ and $x = (x - Px) + Px$. □

We will later generalize the last result in the context of Fredholm operators using the closed complement theorem. We are not in the position to prove Theorem 2.1.12.

Proof of Theorem 2.1.12. The proof follows the lines of Theorem 2.1.11. However, instead of the orthogonal decomposition, we make use of Lemma 2.1.13 for $\mathscr{X}_0 := \ker \mathrm{D}_{\mathscr{X}}\Phi(m_0) \subset \mathscr{X}$ and a corresponding space $\mathscr{X}_1$ isomorphic to $\mathscr{Y}$ (i.e. having the same (finite) dimension). Hence $\mathscr{X} = \mathscr{X}_0 + \mathscr{X}_1$ and we may proceed as in Theorem 2.1.11. □

General Banach theory

In general Banach space, we can provide a similar result as in finite deimensions. It reads as follows

Theorem 2.1.14. *Let $\mathscr{X}$, $\mathscr{Y}$ be Banach spaces and let $M \subset \mathscr{X}$ open. Let $\mathcal{E} : M \to \mathbb{R}$ and $\Phi : M \to \mathscr{Y}$ be continuously differentiable. Let $M_0 := \{x \in M \,:\, \Phi(x) = 0\}$ and let $m_0 \in M_0$ be an extremum of $\mathcal{E}$ in M_0. If $\mathrm{D}_{\mathscr{X}}\Phi(m_0) : \mathscr{X} \to \mathscr{Y}$ is surjective, then there exists $\lambda \in \mathscr{Y}^*$ such that*

$$\forall x \in \mathscr{X} : \qquad \mathrm{D}_{\mathscr{X}}\mathcal{E}(m_0)x = \langle \lambda \,,\, \mathrm{D}_{\mathscr{X}}\Phi(m_0)x \rangle \ .$$

The proof will rely on the the following result for adjoint operators. If $A : \mathscr{X} \to \mathscr{Y}$ is a linear operator, the adjoint $A^* : \mathscr{Y}^* \to \mathscr{X}^*$ is linear and satisfies

$$\forall x \in \mathscr{X}, y^* \in \mathscr{Y}^* : \qquad \langle Ax, y^* \rangle = \langle x, A^* y^* \rangle \ .$$

Theorem 2.1.15. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces and let $A \in \mathcal{L}(\mathscr{X};\mathscr{Y})$ such that $R(A)$ is closed. Then*

$$R(A^*) = (\ker A)^\perp \ .$$

Using this result we can prove Theorem 2.1.14.

Proof of Theorem 2.1.14. W.l.o.g. assume that m_0 is a minimum and consider the joint map $\tilde{\Phi} : M \to \mathbb{R} \times \mathscr{Y}$, $m \mapsto (\mathcal{E}(m), \Phi(m))$ with $\mathrm{D}\tilde{\Phi} = (\mathrm{D}\mathcal{E}, \mathrm{D}\Phi)$. If there was $x \in \mathscr{X}$ with $\langle \mathrm{D}\mathcal{E}(m_0), x\rangle \neq 0$ but $\langle \mathrm{D}\Phi(m_0), x\rangle = 0$, then $\mathrm{D}\tilde{\Phi}(m_0)$ would be surjective. Now, for every $\varepsilon > 0$ Theorem 2.1.8 would provide a solution m_ε for $\tilde{\Phi}(m) = (-\varepsilon, 0)$ with $m_\varepsilon \to m_0$ as $\varepsilon \to 0$, which is a contradiction to the minimizing property of m_0.

Hence, we conclude

$$\forall x \in \mathscr{X} : \qquad \langle \mathrm{D}\Phi(m_0), x\rangle = 0 \quad \Rightarrow \quad \langle \mathrm{D}\mathcal{E}(m_0), x\rangle = 0 \,.$$

In particular, since $R(\mathrm{D}\Phi(m_0))$ is closed, Theorem 2.1.15 yields with $A := \mathrm{D}\Phi(m_0)$

$$\mathrm{D}\mathcal{E}(m_0) \in R(A^*) \ ,$$

and there exists $y^* \in \mathscr{Y}^*$ such that

$$\mathrm{D}\mathcal{E}(m_0) = -A^* y^* \,.$$

Interpreting $\lambda = -y^*$, we infer the statement. □

Examples

Example 2.1.16. With the above insights, we might take once more a look at the Euler problem. Maximizing $L(u)$ with respect to $F(u) \equiv F_0$ yields a Lagrange multiplier $\lambda \in H_0^1(-1,1)$ such that

$$\int_1^1 \frac{\partial_x u(s)\partial_x v(s)}{\sqrt{1+|\partial_x u(s)|^2}} \mathrm{d}s = \lambda \int_{-1}^1 v(s)\mathrm{d}s \,.$$

This implies that $\partial_x \frac{\partial_x u(s)}{\sqrt{1+|\partial_x u(s)|^2}} = \lambda$. Note in particular, that $\lambda = K_0$ is given by the curvature(!). The multiplier λ can be recovered from the above representation of u given by

$$u(x) = \frac{1}{\lambda}\sqrt{1-\lambda^2 x^2} \quad \text{and} \quad \int_{-1}^1 u(s)\mathrm{d}s = F \,.$$

Example 2.1.17. We will now consider a different example. More precisely, we consider $\mathscr{X} = W_0^{1,p}(\Omega)$ for some bounded open domain $\Omega \subset \mathbb{R}^d$. The functional $\mathcal{E}(u) := \int_\Omega \frac{1}{p}|\nabla u|^p$ is complemented by the condition $\Phi(u) = 0$, where $\Phi(u) = 1 - \frac{1}{r}\int_\Omega |u|^r$. By Theorem 2.1.12 we infer that every extremum of $\mathcal{E}$ under the constraint $\Phi(u) = 0$ satisfies

$$\forall v \in \mathscr{X} : \qquad \mathrm{D}_{\mathscr{X}}\mathcal{E}(u)v = \lambda \mathrm{D}_{\mathscr{X}}\Phi(u)v \,.$$

This implies

$$-\int_\Omega \nabla \cdot \left(|\nabla u|^{p-2}\nabla u\right) v = \lambda \int |u|^{r-2} u\, v$$

or

$$-\nabla \cdot \left(|\nabla u|^{p-2}\nabla u\right) = \lambda |u|^{r-2} u \,.$$

2.2 Bifurcation Theory

2.2.1 Fredholm operators and the Index

Let $\mathscr{X}$, $\mathscr{Y}$ be Banach spaces. For $L \in \mathcal{L}(\mathscr{X};\mathscr{Y})$ we denote by $\ker L := \{x \in \mathscr{X} \,:\, L(x) = 0\}$ the kernel and by $R(L) := \{y \in \mathscr{Y} \,:\, \exists x \in \mathscr{X} : L(x) = y\}$ the range of L.

Definition 2.2.1. An operator $L \in \mathcal{L}(\mathscr{X};\mathscr{Y})$ is called Fredholm if

$$\dim\ker(L) < \infty \quad \text{and} \quad \operatorname{codim} R(L) < \infty.$$

The Fredholm condition implies that $\mathscr{Y} = \mathscr{Y}_0 \oplus R(L)$ with finite dimensional $\mathscr{Y}_0$. The integer number $\dim\ker(L) - \operatorname{codim} R(L)$ is called Index of L. Below, we will frequently use the decomposition $\mathscr{X} = \mathscr{X}_0 + \mathscr{X}_1$, where $\mathscr{X}_0 = \ker(L)$ is finite dimensional and $\mathscr{Y} = \mathscr{Y}_0 + \mathscr{Y}_1$ where $\mathscr{Y}_1 = R(L)$ and $\mathscr{Y}_0$ is finite dimensional.

This splitting is well defined, as the Theorem 2.1.7 on the closed complement shows. Based on this theorem, we make the following observations:

1. Let $\mathscr{X}$ be a Banach space and $\mathscr{X}_0$ a finite dimensional subspace. Then there exists a continuous projection onto $\mathscr{X}_0$ in particular, the complement of $\mathscr{X}_0$ is closed.
   Proof: let $\mathscr{X}_0 = \operatorname{span}(e_1, \dots, e_n)$ and $\lambda_k : \mathscr{X}_0 \to \mathbb{R}$ the linear functionals with $\lambda_k(e_j) = \delta_{kj}$ respectively. Using the Hahn-Banach theorem we may extend λ_k to $\mathscr{X}$ and

   $$Px := \sum_{k=1}^{n} \lambda_k(x) e_k$$

   is a continuous linear projection $\mathscr{X} \to \mathscr{X}_0$. In particular, $\mathscr{X}_1$ can be assumed to be closed.

2. If $\dim\ker(L) < \infty$ and $\operatorname{codim} R(L) < \infty$ then $R(L)$ is closed.
   Proof: Let $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$ with $\mathscr{X}_0 = \ker L$ and $\mathscr{X}_1$ closed. The operator $\overline{L} : \mathscr{X}_1 \times \mathscr{Y}_0 \to \mathscr{Y}$, $(x_1, y_0) \to Lx_1 + y_0$ is bijective, linear continuous and hence has an inverse $B := \overline{L}^{-1}$. Therefore:

   $$R(L) = L(\mathscr{X}_1) = \overline{L}(\mathscr{X}_1 \times \{0\}) = B^{-1}(\mathscr{X}_1 \times \{0\})$$

   is closed.

3. Since $\mathscr{X}_1$ and $\mathscr{X}_0$ are closed, and since $R(L)$ is closed by assumption, the inverse operator theorem yields that $L|_{\mathscr{X}_1} \to R(L)$ is continuously invertible.

Fredholm operators are closely related to compact operators.

Theorem 2.2.2. *If $K \in \mathcal{L}(\mathscr{X};\mathscr{X})$ is compact, then $L := \mathrm{id} + K$ is Fredholm with index* 0.

Proof. We prove the Theorem in 5 Steps.

Step 1: $\dim\ker(L) < \infty$. This follows from the fact that $\mathbb{B}_1^{\mathscr{X}}(0) \cap \ker(L) = \left\{x \in \mathbb{B}_1^{\mathscr{X}}(0) \,:\, x = Kx\right\}$ is a precompact ball and hence $\ker(L)$ is finite dimensional.

Step 2: $R(L)$ is closed: This follows from Lemma 1.1.22.

Step 3: $\ker(L) = \{0\}$ implies $R(L) = \mathscr{X}$: This follows from $\mathfrak{d}(L, B_1^{\mathscr{X}}(0), 0) = 1$ and the fact that $\mathfrak{d}$ is continuous under valid homotopies.

Step 4: $\mathrm{codim} R(L) \leq \dim\ker(L)$: Let $n = \dim\ker(L)$ and $x_1, \dots, x_n$ a basis of $\ker(L)$. Assuming the statement was wrong, there exist $n+1$ linear independent vectors $y_1, \dots y_{n+1} \in \mathscr{X} \backslash R(L)$. Furthermore, there exist n continuous linear functionals $x_1', \dots, x_n' \in \mathscr{X}'$ such that $x_l'(x_k) = \delta_{lk}$ and hence

$$\tilde{K}x := Kx + \sum_{i=1}^{n} x_i'(x) y_i$$

is compact. Moreover, $\ker\left(\mathrm{id} + \tilde{K}\right) = \{0\}$ and hence $R\left(\mathrm{id} + \tilde{K}\right) = \mathscr{X}$. This is a contradiction to $y_{n+1} \notin R\left(\mathrm{id} + \tilde{K}\right)$.

Step 5: $\mathrm{codim} R(L) \geq \dim\ker(L)$: We proceed similar to Step 4, but now $m = \mathrm{codim} R(L) < \ker(L) = n$ implies existence of $y_1, \dots y_m \in \mathscr{X} \backslash R(L)$ such that

$$\tilde{K}x := Kx + \sum_{i=1}^{m} x_i'(x) y_i \,.$$

Now $R(\tilde{L}) = \mathscr{X}$ and it remains to show for such operators that $\ker(\tilde{L}) = \{0\}$. Hence, w.l.o.g. assume $R(L) = \mathscr{X}$.

From the above considerations we find $\mathscr{X} = \mathscr{X}_1 + \ker(L)$ and $L : \mathscr{X}_1 \to \mathscr{X}$ is invertible with inverse $\tilde{L}$. Hence, $\ker \tilde{L} = \{0\}$ and

$$\mathrm{id} - \tilde{L} = \tilde{L}\,(L - \mathrm{id}) = \tilde{L}K$$

is compact. But then from Step 3. we obtain $R(\tilde{L}) = \mathscr{X}$ and hence $\ker(L) = \{0\}$. □

Lemma 2.2.3. *Let $L : \mathscr{X} \to \mathscr{Y}$ be Fredholm with index 0. Then there exists a Banach isomorphism $I : \mathscr{Y} \to \mathscr{X}$ such that $I \circ L = \mathrm{id} + K$ for some $K \in \mathbb{C}(\mathscr{X})$.*

Proof. We once more use $\overline{L} : \mathscr{X}_1 \times \mathscr{Y}_0 \to \mathscr{Y}$, $(x_1, y_0) \to Lx_1 + y_0$ with inverse $B = \overline{L}^{-1}$. $B \circ L : \mathscr{X} \to \mathscr{X}_1 \times \mathscr{Y}_0$ is then the identity on $\mathscr{X}_1$. Since $\mathscr{X}_0$ and $\mathscr{Y}_0$ have the same dimension they are isomorphic by an isomorphism B_0. More precisely, we define $C : \mathscr{X}_1 \times \mathscr{Y}_0 \to \mathscr{X}_1 \times \mathscr{X}_0$ through $(x_1, y_0) \mapsto (x_1, B_0 y_0)$ and hence $I = C \circ B : \mathscr{Y} \to \mathscr{X}$ is an isomorphism with $I \circ L|_{\mathscr{X}_1} = \mathrm{id}$ and hence $I \circ L = \mathrm{id} + K$ for some $K \in \mathbb{C}(\mathscr{X})$. □

Lemma 2.2.4. *Let $L : \mathscr{X} \to \mathscr{Y}$ be continuous Fredholm and $K : \mathscr{X} \to \mathscr{Y}$ be continuous compact. Then $L + K$ is Fredholm with the same index as L.*

Proof. We may extend $\mathscr{X}$ or $\mathscr{Y}$ by a finite space and hence assume w.l.o.g. hat the index of L is 0. According to Lemma 2.2.3 there exist and isomorphism $I : \mathscr{Y} \to \mathscr{X}$ such that $\tilde{L} := I \circ L = \mathrm{id} + K_L$ for some compact operator $K_L : \mathscr{X} \to \mathscr{X}$. Moreover, $\tilde{K} := I \circ K$ is also continuous and compact and so is $K_L + \tilde{K}$. By Theorem 2.2.2 $\mathrm{id} + K_L + \tilde{K}$ hast index 0 and hence $L + K = I^{-1} \circ \left(\mathrm{id} + K_L + \tilde{K}\right)$ has index 0. □

For later use, we finally note the following important consequence.

Theorem 2.2.5 (Fredholm alternative)**.** *Let $\lambda \neq 0$ a real value and $K \in \mathcal{L}(\mathscr{X};\mathscr{X})$ compact. Then either $(\lambda - K)\,x = 0$ has a non-trivial solution or $\lambda - K$ is invertible.*

Proof. Consider $L = \lambda - K$. According to Theorem 2.2.2 L is Fredholm with index 0. Hence, if L is not invertible, this implies $\dim\ker(L) \neq 0$. □

2.2.2 Ljapunov-Schmidt Reduction

Let $f : \mathscr{X} \times \Lambda \to \mathscr{Y}$ be a continuous mapping for Banach spaces Λ, $\mathscr{X}$ and $\mathscr{Y}$ and consider the equation

$$f(x,\lambda) = 0\,. \tag{2.3}$$

Under the assumption that $f(0,0) = 0$, we could resolve $x = x(\lambda)$ in case $\mathrm{D}_{\mathscr{X}} f(0,0)$ was invertible (using the implicit function theorem). However, this is often not the case.

In the following, we will assume that $\mathrm{D}_{\mathscr{X}} f(0,0)$ is "close to invertible" in the sense that $\mathrm{D}_{\mathscr{X}} f(0,0)$ is assumed to be Fredholm. Like in the last section, we can split $\mathscr{X} = \mathscr{X}_0 + \mathscr{X}_1$ and $\mathscr{Y} = \mathscr{Y}_0 + \mathscr{Y}_1$ with projections

$$P : \mathscr{X} \to \mathscr{X}_0\,, \qquad Q : \mathscr{Y} \to \mathscr{Y}_0\,.$$

Then (2.3) is equivalent with

$$\begin{aligned} Q\,f(x,\lambda) &= 0\,,\\ (1-Q)\,f(x,\lambda) &= 0\,. \end{aligned}$$

Using $x = x_0 + x_1$, $(x_0, x_1) \in \mathscr{X}_0 \times \mathscr{X}_1$, we write this equation as

$$\begin{aligned} Q\,f(x_0 + x_1,\lambda) &= 0\,,\\ (1-Q)\,f(x_0 + x_1,\lambda) &= 0\,. \end{aligned}$$

In this formulation, we can solve the second equation for x_1 because

$$\mathrm{D}_{\mathscr{X}_1}\left[(1-Q)\,f(x_0 + x_1,\lambda)\right] = (1-Q)\,\mathrm{D}_{\mathscr{X}} f(0,0) \,:\, \mathscr{X}_1 \to \mathscr{Y}_1$$

is invertible. Hence we reduced (2.3) to

$$Q\,f(x_0 + x_1(x_0,\lambda),\lambda) = 0\,.$$

This equation is called "reduced equation" or "bifurcation equation". We formulate this in the following more general Lemma:

Theorem 2.2.6. *Let $\mathscr{X}$ and Λ be Banach spaces and $f \in C(\mathscr{X} \times \Lambda)$ with $\mathrm{D}_{\mathscr{X}} f \in C(\mathscr{X} \times \Lambda; \mathcal{L}(\mathscr{X};\mathscr{Y}))$. Let $L = \mathrm{D}_{\mathscr{X}} f(0,0)$ such that $R(L)$ and the projection $P : \mathscr{X} \to \mathscr{X}$ onto $\ker L$ and $\mathrm{id} - Q : \mathscr{Y} \to \mathscr{Y}$ onto $R(L)$ are continuous. Then there exist neighborhoods U of 0 in $P\mathscr{X}$ and V of 0 in $(\mathrm{id} - P)\mathscr{X}$ and W of 0 in Λ such that solving $f(x,\lambda) = 0$ can be reduced around $(0,0)$ to the localized equations*

$$x = Px + v(Px,\lambda) \qquad Qf(Px + v(Px,\lambda)),\lambda) = 0\,,$$

where $v : \ker L \times \Lambda \supset U \times W \to V$ is given by the implicit function theorem.

Proof. Let $u = Px$ and $v = (1 - P)x$. Then

$$f(x,\lambda) = 0 \qquad \Leftrightarrow \qquad \begin{cases} F(u,v,\lambda) = (\mathrm{id} - Q)\, f(x,\lambda) = 0 \\ \Phi(u,v,\lambda) = Q f(x,\lambda) = 0 \end{cases}$$

It holds $F(0,0,0) = 0$ and

$$\mathrm{D}_V F(0,0,0) = (\mathrm{id} - Q)\, \mathrm{D}_{\mathscr{X}} f(0,0) : \, (\mathrm{id} - P)\mathscr{X} \to Q\mathscr{Y}$$

is bijective. Hence the inverse of $\mathrm{D}_V F(0,0,0)$ is continuous and the implicit function theorem yields $v : U \times W \to V$ such that $F(u, v(u,\lambda), \lambda) = 0$ with

$$\begin{aligned} \mathrm{D}_U v(0,0) &= -\mathrm{D}_V F(0,0,0)^{-1} \mathrm{D}_U F(0,0,0) \\ &= -\left((\mathrm{id} - Q)\, \mathrm{D}_{\mathscr{X}} f(0,0)\right)^{-1} (\mathrm{id} - Q)\, \mathrm{D}_{\mathscr{X}} f(0,0) P \,. \end{aligned}$$

□

2.2.3 The idea behind bifurcation theory

For given $\lambda \in \mathbb{R}$, equations like $x = \lambda$ or $x^3 = \lambda$ always have a unique solution in $\mathbb{R}$. In contrast with that, the equation $x^2 - \lambda = 0$ may have no solutions (if $\lambda < 0$), one solution (if $\lambda = 0$) or two solutions (if $\lambda > 0$). Similar phenomena arise in the equations

$$x(x^2 - \lambda) = 0 \qquad\qquad x(x^2 - \lambda)(x^4 - \lambda) = 0$$
$$x(x - \lambda) = 0 \qquad\qquad x(x^2 - \lambda)(x^2 - 2\lambda) = 0$$

Example 2.2.7. Let $\mathscr{X} = \mathbb{R}^2$ and $\Lambda = \mathbb{R}$ as well as

$$f(x_1, x_2, \lambda) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \lambda \begin{pmatrix} -2x_1 \\ x_2 \end{pmatrix} + \lambda \begin{pmatrix} x_1^3 \\ 0 \end{pmatrix} .$$

Then $(0,0,\lambda)$ is the trivial solution. Furthermore, dividing by x_1 resp. x_2 we obtain that $f(x_1, x_2, \lambda) = 0$ if one of the following two conditions is fulfilled:

$$\begin{aligned} 0 = 1 - \lambda(-2 - x_1^2) \quad &\text{and} \quad x_2 = 0 \\ \text{or} \qquad x_1 = 0 \quad &\text{and} \quad 1 - \lambda = 0 \,. \end{aligned}$$

The first condition is equivalent with $x_2 = 0$ and

$$x_1 = \sqrt{-\frac{1}{\lambda}\,(1 + 2\lambda)} \,.$$

From the last formula one sees that additional solutions emerge for $\lambda > -\frac{1}{2}$. The derivative of f at $(0,0,\lambda)$ is given through

$$\mathrm{D} f(0,0,\lambda) = \begin{pmatrix} 1 + 2\lambda & 0 \\ 0 & 1 - \lambda \end{pmatrix}$$

and degenerates precisely in $\lambda = -\frac{1}{2}$ and $\lambda = 1$.

The latter observation holds more generally:

Theorem 2.2.8. *Let $\mathscr{X}$, $\mathscr{Y}$ and Λ be Banach spaces, $f : \mathscr{X} \times \Lambda \to \mathscr{Y}$ be C^1 and $f(x_0, \lambda_0) = 0$. If*

$$\mathrm{D}_{\mathscr{X}} f(x_0, \lambda_0) : \mathscr{X} \to \mathscr{Y} \qquad \textit{is and isomorphism,}$$

then $f(x, \lambda) = 0$ can be uniquely solved by $x = g(\lambda)$ in a neighborhood of (x_0, λ_0).

In contrast with the last theorem, we introduce the following definition.

Definition 2.2.9. Let $f : \mathscr{X} \times \Lambda \to \mathscr{X}$ be continuous with a trivial branch, i.e. for every $\lambda \in \Lambda$ holds $f(0, \lambda) = 0$. Then we say that $(0, \lambda_0)$ is a point bifurcation if every neighborhood of $(0, \lambda_0)$ contains a nontrivial solution $f(x, \lambda) = 0$, $x \neq 0$.

In particular, in the the context of Theorem 2.2.8 the point (x_0, λ_0) is not a point of bifurcation. However, degeneracy of $\mathrm{D}_{\mathscr{X}} f(x_0, \lambda_0)$ is necessary but not sufficient. Below we will find sufficient conditions for bifurcation.

Remark 2.2.10. (i) The problems bifurcation arises despite the smoothness of f. They are independent from applying any smooth transformations like rotations.

(ii) The equation $f(x, \lambda) = 0$ can be solved locally almost everywhere except some "lower dimensional manifold", often this consists of isolated points.

2.2.4 Bifurcation in a simple eigenvalue

For $f : \mathscr{X} \times \mathscr{Y} \to \mathscr{Z}$ we find $\mathrm{D}_{\mathscr{X}} f : \mathscr{X} \times \mathscr{Y} \to \mathcal{L}(\mathscr{X}; \mathscr{Z})$ with $(x, y) \mapsto \mathrm{D}_{\mathscr{X}} f(x, y)$ and provided f is regular enough, we might calculate $\mathrm{D}_{\mathscr{Y}} \mathrm{D}_{\mathscr{X}} f : \mathscr{X} \times \mathscr{Y} \to \mathcal{L}(\mathscr{Y}; \mathcal{L}(\mathscr{X}; \mathscr{Z}))$, $(x, y) \mapsto \mathrm{D}_{\mathscr{Y}} \mathrm{D}_{\mathscr{X}} f(x, y)$. This is particularly possible if $f \in C^2(\mathscr{X} \times \mathscr{Y}; \mathscr{Z})$.

Theorem 2.2.11 (Bifurcation in a simple eigenvalue)**.** *Let $\Lambda = \mathbb{R}$ and $\mathscr{X}$,$\mathscr{Y}$ be Banach spaces. Let $f \in C^2(\mathscr{X} \times \Lambda; \mathscr{Y})$ where we assume*

1. *Trivial branch:* $\forall \lambda \in \Lambda : f(0, \lambda) = 0$.
2. *Fredholm property:* $L(\lambda) := \mathrm{D}_{\mathscr{X}} f(0, \lambda)$ *is Fredholm with index* 0.
3. *Simple eigenvalue:* $\mathscr{X}_0 := \ker L(0) = \mathbb{R}\tilde{x}_0$
4. *Transversality:* $\mathrm{D}_\Lambda \mathrm{D}_{\mathscr{X}} f(0, 0)\tilde{x}_0 \notin R(L(0))$.

Then for $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$ and neighborhoods $(0, 0, 0) \in U \times V \times W \subset \mathscr{X}_0 \times \mathscr{X}_1 \times \Lambda$ there exist unique continuous $\tilde{v} : U \to V$, $\tilde{\lambda} : U \to W$, $\tilde{v}(0) = 0$, $\tilde{\lambda}(0) = 0$ such that

$$f(x_0 + x_1, \lambda) = 0 \quad \Leftrightarrow \quad \begin{cases} x = x_0 + x_1 = 0 & \textit{or} \\ (x_1, \lambda) = \Big(\tilde{v}(x_0), \, \tilde{\lambda}_0(x_0)\Big) \end{cases} .$$

Furthermore, $f \in C^k$ implies $\tilde{v}, \tilde{\lambda} \in C^{k-1}$.

Example 2.2.12. To explain the name of the theorem, consider $f(x,\lambda) = (\mu_0 + \lambda)\,x - Tx$, where $\mu_0 \in \mathbb{R}\backslash\{0\}$ and T is nonlinear differentiable and compact with $T(0) = 0$. We consider $L := \mathrm{D}_{\mathscr{X}}T(0)$. We check the under what additional assumptions the conditions 1.–4. of Theorem 2.2.11 are satisfied.

1. This is always satisfied.
2. requires that $\mathrm{D}_{\mathscr{X}}f(0,\lambda) = (\mu_0 + \lambda)\,x - Lx$ is Fredholm, which is always true in a neighborhood of $\lambda = 0$.
3. Requires that $\mathscr{X}_0 := \ker(\mu_0 x - Lx) = \mathbb{R}\tilde{x}_0$, i.e. μ_0 is a simple eigenvalue of L (i.e. geometric multiplicity 1).
4. Since $\mathrm{D}_\Lambda L(\lambda) = \mathrm{id}$, this implies that $\tilde{x}_0 \notin R(\mu_0 - L)$. This implies that μ_0 has algebraic multiplicity 1.

Proof. We will first reduce the problem to the 1-dimensional case and then solve this particular problem.

Step 1: We will use a Ljapunov-Schmidt reduction: Let $\mathscr{Y} = \mathscr{Y}_0 \oplus \mathscr{Y}_1$ where $\mathscr{Y}_1 = R(L(0))$ with a continuous projection $Q : \mathscr{Y} \to \mathscr{Y}_0$ with $\ker Q = \mathscr{Y}_1$. Then we can locally solve $v : U \times W \to \mathscr{X}_1$

$$f(x,\lambda) = 0 \qquad \Leftrightarrow \qquad Qf(x_0 + v(x_0,\lambda),\lambda) = 0\,,$$

where

$$\mathrm{D}_{\mathscr{X}_0,\Lambda}v = -\left((1-Q)\,\mathrm{D}_{\mathscr{X}}f(0,0)|_{\mathscr{X}_1}\right)^{-1}(1-Q)\,\mathrm{D}_{\mathscr{X}_0,\Lambda}f(0,0)$$

Now we set $\Phi(r,\lambda) := Qf(r\tilde{x}_0 + v(r\tilde{x}_0,\lambda),\lambda)$ with the evident identification of $\mathbb{R}\tilde{x}_0$ with $\mathbb{R}$. Then the problem is reduced to a 1-dimensional problem and it remains to show that Φ satisfies 1.–4..

1. $Qf(0 + v(0,\lambda),\lambda) = 0$ since $v(x_0,\lambda)$ is the unique solution of $Qf(x_0 + v(x_0,\lambda),\lambda) = 0$.
2. $\mathrm{D}_r\Phi(0,\lambda) : \mathbb{R} \to \mathbb{R}$ – as a linear map between finite dimensional spaces – is Fredholm.
3. We calculate

$$\mathrm{D}_r\Phi(0,0) = Q\mathrm{D}_{\mathscr{X}}f(0,0)\,(\mathrm{id} + \mathrm{D}_{\mathscr{X}_0,\Lambda}v(0,0))\,\tilde{x}_0 = 0$$

since $\mathscr{X}_0 := \ker L(0) = \mathbb{R}\tilde{x}_0$ and hence $\mathrm{D}_{\mathscr{X}_0}f(0,0)\tilde{x}_0 = 0$.

4.

$$\begin{aligned}\mathrm{D}_\Lambda\mathrm{D}_r\Phi(0,0) &= Q\mathrm{D}_\Lambda\left((\mathrm{D}_{\mathscr{X}}f\,(r\tilde{x}_0 + v(r\tilde{x}_0,\lambda),\lambda))\,(\tilde{x}_0 + \mathrm{D}_{\mathscr{X}_0}v(r\tilde{x}_0,\lambda)\,\tilde{x}_0)\right)\\ &= Q\mathrm{D}_{\mathscr{X}}\mathrm{D}_{\mathscr{X}}f(0,0)\,\langle\tilde{x}_0 + \mathrm{D}_{\mathscr{X}_0}v(r\tilde{x}_0,\lambda)\,\tilde{x}_0,\ \mathrm{D}_\Lambda v(0,0)\rangle\\ &\quad + Q\partial_\lambda\mathrm{D}_{\mathscr{X}}f(0,0)\,(\tilde{x}_0 + \mathrm{D}_{\mathscr{X}_0}v(0,0)\,\tilde{x}_0)\\ &\quad + Q\mathrm{D}_{\mathscr{X}}f(0,0)\,\partial_\lambda\mathrm{D}_{\mathscr{X}_0}v(0,0)\,\tilde{x}_0\,.\end{aligned}$$

The last term equals zero because $Q\mathrm{D}_xf(0,0) = 0$. Since $v(0,\lambda) = 0$, we find $\mathrm{D}_\Lambda v(0,0) = 0$ and hence the first term vanishes, too. Since $\mathrm{D}_{\mathscr{X}_0}f(0,0)\tilde{x}_0 = 0$, the second term reduces to $Q\partial_\lambda\mathrm{D}_{\mathscr{X}}f(0,0)\tilde{x}_0$. However, note that $\partial_\lambda\mathrm{D}_{\mathscr{X}}f(0,0)\tilde{x}_0 \notin R(L(0))$ and hence $Q\partial_\lambda\mathrm{D}_{\mathscr{X}}f(0,0)\tilde{x}_0 \neq 0$.

Hence, if $\Phi = 0$ can be solved by $\tilde\lambda(x_0)$, the theorem follows with $x_1 = v(x_0, \tilde\lambda(x_0))$.

Step 2: We now assume $\mathscr{X} = \mathscr{Y} = \mathbb{R}$ and consider

$$\psi(x,\lambda) := \begin{cases} \frac{1}{x} f(x,\lambda) & \text{for } x \neq 0\,, \\ \mathrm{D}_x f(0,\lambda) & \text{for } x = 0\,. \end{cases}$$

In case $x \neq 0$, we find

$$f(x,\lambda) = 0 \qquad \Leftrightarrow \qquad \psi(x,\lambda) = 0\,.$$

Hence, in order to show existence of non-trivial solutions it remains to solve the equation for ψ. We find

$$\partial_\lambda \psi(x,\lambda) = \begin{cases} \frac{1}{x} \partial_\lambda f(x,\lambda) & \text{for } x \neq 0\,, \\ \mathrm{D}_x \partial_\lambda f(0,\lambda) & \text{for } x = 0\,. \end{cases}$$

and since $\partial_\lambda f(0,\lambda) = 0$, we find continuity of $\partial_\lambda \psi(x,\lambda)$. Furthermore since f is twice differentiable, we deduce that $\partial_x \psi$ exists from $f(x,\lambda) = 0 + x\partial_x f(0,\lambda) + x^2 \partial_x^2 f(0,\lambda) + o(x^2)$. Hence $\psi \in C^1$. Since $\partial_\lambda \psi(0,0) = \mathrm{D}_x \partial_\lambda f(0,0) \neq 0$, we can apply the implicit function theorem to solve $\psi(x, \tilde\lambda(x)) = 0$ in a neighborhood of $(0,0)$, where $\tilde\lambda \in C^1$. □

Example 2.2.13. Consider

$$f(x,\lambda) := \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} x + g(x) = 0\,,$$

where $g \in C^2(\mathbb{R}^3;\mathbb{R}^3)$, $g(0) = 0$, $\mathrm{D}g(0) = 0$. The above theorem yields existence of a non-trivial branch that bifurcates at $(x,\lambda) = (0,0)$. We check the prerequisites:

1. and 2. are satisfied since $f(0) = 0$ and $\mathscr{X} = \mathbb{R}^3$ is finite dimensional. 3. holds since

$$\ker \mathrm{D}_x f(0,0) = \ker \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{R} e_1\,.$$

Finally, 4. holds because

$$\begin{aligned} \mathrm{D}_\lambda \mathrm{D}_x f(0,0)\, e_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} e_1 = e_1 \\ \notin R \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{R} e_2 + \mathbb{R} e_3\,. \end{aligned}$$

Example 2.2.14. We consider $f \in C^2(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}^d)$ with trivial branch $f(0,\lambda) = 0$. We assume that $L(\lambda) = \mathrm{D}_x f(0,\lambda)$ has the eigenvalues

$$\mu_1(\lambda) \leq \mu_2(\lambda) \leq \dots < \mu_k(\lambda) < \dots \leq \mu_d(\lambda)$$

which depend in a differentiable way on λ. Furthermore, let $\mu_k(\lambda_0) = 0$ for some critical λ_0. we assume that $\mu_k(\lambda_0)$ is singular and $\partial_\lambda \mu_k(\lambda_0) \neq 0$. Then f has a non-trivial branch that bifurcates in $(0, \lambda_0)$.

In particular, 1., 2. and 3. are evidently satisfied. In order to verify 4. let $x_0(\lambda)$ denote the eigenvector of $\mu_k(\lambda)$ note that $x_0(\lambda_0)$ lies orthogonal to $R(L(\lambda_0))$ since $\mu_k(\lambda_0) = 0$. We differentiate the eigenvalue equation

$$\mu_k(\lambda)x_0(\lambda) = L(\lambda)x_0(\lambda)$$

and obtain

$$\partial_\lambda \mu_k(\lambda)x_0(\lambda) + \mu_k(\lambda)\partial_\lambda x_0(\lambda) = \partial_\lambda L(\lambda)x_0(\lambda) + L(\lambda)\partial_\lambda x_0(\lambda)\,.$$

By assumption, $\partial_\lambda \mu_k(\lambda_0) =: \alpha \neq 0$, $\mu_k(\lambda_0) = 0$ and $L(\lambda_0)\partial_\lambda x_0(\lambda_0) \in R(L(\lambda_0))$. Because the eigenspace of $\mu_k(\lambda_0)$ is 1-dimensional, we find

$$\partial_\lambda L(\lambda_0)x_0(\lambda_0) \notin R(L(\lambda_0))\,.$$

2.2.5 The Index of a solution and bifurcation

Let $\mathscr{X}$ be a real Banach space and $G \subset \mathscr{X}$. For a locally invertible function $f \in C^1(G; \mathscr{X})$ and a regular point x_0 we can define the index

$$\mathrm{Index}(f, x_0) := \mathfrak{d}(f, \mathbb{B}_\varepsilon^{\mathscr{X}}(x_0), 0)\,.$$

For a finite dimensional space $\mathscr{X} = \mathbb{R}^d$, we obtain from the representation (1.6) that

$$\begin{aligned}\mathrm{Index}(f, x_0) &= \mathfrak{d}(f, \mathbb{B}_\varepsilon^{\mathscr{X}}(x_0), 0) = \operatorname{sign}\det \mathcal{J}_f(x_0)\\ &= \operatorname{sign}\det \mathrm{D}f(x_0) = \mathfrak{d}(\mathrm{D}f(x_0)(\cdot - x_0), \mathbb{B}_\varepsilon^{\mathscr{X}}(x_0), 0)\\ &= (-1)^\beta\,,\end{aligned}$$

where β is amount of negative eigenvalues of $\mathrm{D}f(x_0)$ (with multiplicity!). In particular, the index is independent from ε for ε small enough. Note that Eigenvalues $\lambda \in \mathbb{C}\backslash\mathbb{R}$ do not enter this formula since λ is an eigenvalue of $\mathrm{D}f(x_0)$ iff $\overline{\lambda}$ is an eigenvalue and $\lambda\overline{\lambda} > 0$.

In this context, we recall the definition of the multiplicity of an eigenvalue of an operator K:

$$n_\lambda(K) := \dim\left(\bigcup_{p=1}^{\infty} \ker(\lambda - K)^p\right)\,,$$

where for some p_0 $\ker(\lambda - K)^p = \ker(\lambda - K)^{p_0}$ for all $p > p_0$ and the kernel of $\lambda - K$ remains unchanged. The space $\bigcup_{p=1}^{n_\lambda} \ker(\lambda - K)^p$ is the generalized eigenspace. Furthermore, we use the following result, which is part of the spectral theorem for compact operators.

Lemma 2.2.15 (Spectral Lemma). *Let $\mathscr{X}$ be a Banach space, $K \in \mathcal{L}(\mathscr{X})$ be compact. Then for every $\alpha > 0$ $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$, where $\mathscr{X}_0$ is finite dimensional and spanned by all generalized eigenspaces with respect to $|\lambda| \geq \alpha$ and $\mathscr{X}_0$ and $\mathscr{X}_1$ are invariant under K, i.e. $K\mathscr{X}_0 \subset \mathscr{X}_0$ and $K\mathscr{X}_1 \subset \mathscr{X}_1$.*

Theorem 2.2.16 (Leray-Schauder)**.** *Let $\mathscr{X}$ be a real Banach space, $G \subset \mathscr{X}$, $f \in C^1(G;\mathscr{X})$ and $\tilde{x}$ a zero of f. Furthermore, assume*

1. *$\tilde{x}$ is regular, i.e. $\mathrm{D}_{\mathscr{X}} f(\tilde{x})$ is invertible*
2. *$\mathrm{D}_{\mathscr{X}} f(\tilde{x})$ is Fredholm with $A := \mathrm{D}_{\mathscr{X}} f(\tilde{x}) = \mathrm{id} + K$ where $K \in \mathcal{L}(\mathscr{X})$ is compact.*

Then

$$\mathrm{Index}(f,\tilde{x}) := \mathfrak{d}(f, \mathbb{B}_\varepsilon^{\mathscr{X}}(\tilde{x}), 0) = (-1)^\beta\,,$$

where

$$\beta = \sum_{\lambda\in\mathbb{R},\,\lambda<-1} n_\lambda(K)\,.$$

Proof. W.l.o.g. we assume $\tilde{x} = 0$. We consider the homotopy $h(x,t) = (1-t)f(x) + tAx$. This homotopy is valid if $h(x,t) = 0$ implies $x \notin \partial\mathbb{B}_\varepsilon^{\mathscr{X}}(0)$. We verify through a short calculation:

$$\begin{aligned} h(x,t) = 0 \quad &\Leftrightarrow \quad Ax = \quad (1-t)\left(f(x) - A(x)\right) \\ &\Leftrightarrow \quad\; x = A^{-1}(1-t)\left(f(x) - A(x)\right)\,. \end{aligned}$$

Since f is differentiable with invertible $A = \mathrm{D}_{\mathscr{X}} f(0)$ we obtain $\|A^{-1}\left(f(x) - A(x)\right)\| \leq \frac{\varepsilon}{2}$ provided $x \in \mathbb{B}_\varepsilon^{\mathscr{X}}(0)$. In particular, we obtain

$$\mathfrak{d}(f, \mathbb{B}_\varepsilon^{\mathscr{X}}(0), 0) = \mathfrak{d}(A, \mathbb{B}_\varepsilon^{\mathscr{X}}(0), 0)\,.$$

Using Lemma 2.2.15 we obtain $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$ with projections P_0 and P_1 where $\mathscr{X}_0$ is finite dimensional and generated by the generalized eigenspaces for $|\lambda| \geq \frac{1}{2}$ and $\mathscr{X}_1$ is invariant under K. Let $\tilde{K}(t) := K \circ P_0 + (1-t)K \circ P_1$ be the homotopy between K and $K \circ P_0$. We claim that $\mathrm{id} + \tilde{K}(t)$ is a valid homotopy between $\mathrm{id} + K$ and $\mathrm{id} + K_0$. Indeed, we find due to the invariance of $\mathscr{X}_0$ and $\mathscr{X}_1$:

$$x + \tilde{K}(t)x = 0 \quad \Leftrightarrow \quad x_0 = -Kx_0 \quad \text{and} \quad x_1 = (t-1)Kx_1\,.$$

Since $\mathrm{id} - K = \mathrm{D}_{\mathscr{X}} f(0)$ is invertible we obtain $x_0 = 0$. Since K has no eigenvalues on $\mathscr{X}_1$ with $|\lambda| > \frac{1}{2}$, and we obtain $x_1 = 0$. Hence we obtain from the finite dimensional formula

$$\begin{aligned} \mathfrak{d}(f, \mathbb{B}_\varepsilon^{\mathscr{X}}(0), 0) &= \mathfrak{d}(\mathrm{id} + K_0, \mathbb{B}_\varepsilon^{\mathscr{X}}(0), 0) \\ &= (-1)^\beta \end{aligned}$$

with

$$\beta = \sum_{\lambda\in\mathbb{R},\,\lambda<-1} n_\lambda(K_0) = \sum_{\lambda\in\mathbb{R},\,\lambda<-1} n_\lambda(K)\,.$$

□

The concept of topological index of a solution can be used to provide a sufficient condition for bifurcation.

Theorem 2.2.17. *Let $\mathscr{X}$ be a finite dimensional Banach space and $\Lambda = \mathbb{R}$, $f : \mathscr{X} \times \Lambda \to \mathscr{X}$ continuously differentiable with $f(0, \lambda) = 0$ for every λ and $\mathrm{D}_{\mathscr{X}} f(0, \lambda)$ invertible for $\lambda \neq \lambda_0$. If*

$$\begin{aligned} &\mathrm{Index}\,(f(\cdot, \lambda), 0) = \sigma \in \{-1, 1\} \qquad \forall \lambda > \lambda_0\,, \\ &\mathrm{Index}\,(f(\cdot, \lambda), 0) = -\sigma \qquad \forall \lambda < \lambda_0\,, \end{aligned}$$

then λ_0 is a point of bifurcation, i.e. there exists $\varepsilon_0 > 0$ such that for every $\varepsilon \in (0, \varepsilon_0)$ there exists $x(\varepsilon), \lambda(\varepsilon)$ s.t.

$$f(x(\varepsilon), \lambda(\varepsilon)) = 0\,, \qquad \|x(\varepsilon)\| = \varepsilon\,.$$

Moreover, it is possible to choose $\lambda(\varepsilon) \to \lambda_0$ for $\varepsilon \to 0$.

Hence, in a neighborhood of $(0, \lambda_0)$ there exists for every λ a non-trivial solution additionally to the trivial solution.

Proof. We chose a sequence $\mu_k \searrow 0$ and consider $\lambda_{\pm k} = \lambda_0 \pm \mu_k$. For every k there exists $\varepsilon_k > 0$ s.t.

$$\forall \varepsilon < \varepsilon_k : \qquad \mathrm{Index}(f(\cdot, \lambda_{\pm k})) = \mathfrak{d}\left(f(\cdot, \lambda_{\pm k}), \mathbb{B}^d_\varepsilon(0), 0\right) = \pm\sigma\,.$$

In particular, we find

$$\mathfrak{d}\left(f(\cdot, \lambda_0 + \mu_k), \mathbb{B}^d_\varepsilon(0), 0\right) \neq \mathfrak{d}\left(f(\cdot, \lambda_0 - \mu_k), \mathbb{B}^d_\varepsilon(0), 0\right)\,.$$

However, we can chose $\varepsilon_k \searrow 0$ strictly monotone decreasing and for $\varepsilon \in (\varepsilon_{k+1}, \varepsilon_k]$ we can chose $\lambda_\pm := \lambda_{\pm k}$. Then $f(x, \lambda)$ is a homotopy from $f(x, \lambda_-)$ to $f(x, \lambda_+)$ which changes the Index, hence the degree. Therefore, the homotopy is not valid and for some $\lambda \in (\lambda_-, \lambda_+)$ there exists $x \in \partial\mathbb{B}^d_\varepsilon(0)$ with $f(x, \lambda) = 0$. This yields the claim. □

2.2.6 Krasnoselskii's Theorem

The Krasnoselskii Theorem yields another sufficient condition for the existence of bifurcation. Let $\mathscr{X}$ be a Banach space and $f : \mathscr{X} \times \mathbb{R} \to \mathscr{X}$ be of the form

$$f(x, \lambda) = x - (\mu_0 + \lambda)\, Tx + g(x, \lambda)\,,$$

where we impose the following conditions:

1. $\mu_0 \neq 0$,
2. $T \in \mathcal{L}(\mathscr{X})$ is compact
3. $g \in \mathbb{C}(\mathscr{X} \times \mathbb{R}; \mathscr{X})$,
4. $g(0, \lambda) = 0$ for every λ and $g(x, \lambda) = o\,(\|x\|)$ uniformly in $\lambda \in (-\varepsilon, \varepsilon)$.

The last condition implies that $\mathrm{D}_{\mathscr{X}} f = \mathrm{id} - (\mu_0 + \lambda)\, T$. In particular, we infer from Theorem 2.2.8 that $\mathrm{id} - \mu_0 T$ needs to be non-invertible (non-surjective) such that $(0, 0)$ can be a bifurcation point of f. The Krasnoselskii Theorem tells us that this assumption is already sufficient.

Theorem 2.2.18 (Krasnoselskii)**.** *Let 1.-4. hold and let* $\frac{1}{\mu_0}$ *be an eigenvalue of* T *with odd multiplicity. Then* $(0,0)$ *is a point of bifurcation for* f*.*

Proof. Assume $(0,0)$ was not a point of bifurcation. Then there are only trivial solutions in $\mathbb{B}_\varepsilon(0)\times(-\varepsilon,\varepsilon)$ and the index

$$\mathfrak{d}\left(f(\cdot,\lambda),\mathbb{B}_\varepsilon(0),0\right)\in\mathbb{Z}$$

is well defined and due to homotopy invariance, it is independent from $\lambda\in(-\varepsilon,\varepsilon)$. This will lead to a contradiction.

By the spectral lemma 2.2.15, we can assume that $(\mu_0+\lambda)^{-1}$ is not an eigenvalue of T. Hence, the Leray-Schauder Theorem 2.2.16 and $\mathrm{D}_{\mathscr{X}}f(0,\lambda)=\mathrm{id}-(\mu_0+\lambda)\,T$ we find that

$$\mathfrak{d}\left(f(\cdot,\lambda),\mathbb{B}_\varepsilon(0),0\right)=(-1)^{\beta(\lambda)}\,,\qquad\beta(\lambda)=\sum_{\sigma>1}n_\sigma\left((\mu_0+\lambda)\,T\right)\,.$$

Now observe that

$$\begin{aligned}&(\mu_0+\lambda)\,Tx=\sigma x\,,\qquad\sigma>1\\ \Leftrightarrow\quad &TX=\overline{\sigma}\,,\qquad\overline{\sigma}>\frac{1}{\mu_0+\lambda}\,.\end{aligned}$$

Hence

$$\beta(\lambda)=\sum_{\overline{\sigma}>(\mu_0+\lambda)^{-1}}n_{\overline{\sigma}}(T)\,.$$

Since μ_0^{-1} is an isolated eigenvalue, we obtain for ε small enough

$$\beta(\lambda_+)-\beta(\lambda_-)=n_{\mu_0}(T)\in2\mathbb{Z}+1\,.$$

Hence, the homotopy invariance yields

$$\begin{aligned}(-1)^{\beta(\lambda_+)}&=\mathfrak{d}\left(f(\cdot,\lambda_+),\mathbb{B}_\varepsilon(0),0\right)\\ &=\mathfrak{d}\left(f(\cdot,\lambda_-),\mathbb{B}_\varepsilon(0),0\right)=(-1)^{\beta(\lambda_-)}=-\left(-1\right)^{\beta(\lambda_+)}\,.\end{aligned}$$

□

Example 2.2.19. We recall Example 2.2.7:

$$f(x_1,x_2,\lambda)=\begin{pmatrix}x_1\\x_2\end{pmatrix}-\left(-\frac{1}{2}+\lambda\right)\begin{pmatrix}-2x_1\\x_2\end{pmatrix}+\lambda\begin{pmatrix}x_1^3\\0\end{pmatrix}\,,$$

which we reformulate as

$$\begin{aligned}&f(x_1,x_2,\lambda)=\begin{pmatrix}x_1\\x_2\end{pmatrix}-(\mu_0+\lambda)\begin{pmatrix}-2x_1\\x_2\end{pmatrix}+\lambda\begin{pmatrix}x_1^3\\0\end{pmatrix}\,,\\ \text{or}\quad&f(x_1,x_2,\tilde\lambda)=\begin{pmatrix}x_1\\x_2\end{pmatrix}-\left(\tilde\mu_0+\tilde\lambda\right)\begin{pmatrix}-2x_1\\x_2\end{pmatrix}+\left(\tilde\lambda+\frac{3}{2}\right)\begin{pmatrix}x_1^3\\0\end{pmatrix}\,,\end{aligned}$$

where $\mu_0=\frac{-1}{2}$ or $\tilde\mu_0=1$ respectively and the derivative of f at $(0,0,\lambda)$ given through

$$\mathrm{D}f(0,0,\lambda)=\begin{pmatrix}1+2\,(\mu_0+\lambda)&0\\0&1-(\mu_0+\lambda)\end{pmatrix}$$

which degenerates precisely in $\mu_0=-\frac{1}{2}$ and $\mu_0=1$ with $n_{\mu_0}=1$.

However, if $n_{\mu_0} = 2$ the statement of Krasnoselskii might indeed be wrong.

Example 2.2.20. As an example, we consider

$$f(x_1, x_2, \lambda) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - (1+\lambda) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} -x_2^3 \\ x_1^3 \end{pmatrix} .$$

Then $T = \mathrm{id}$ and $\mu_0 = 1$ is an eigenvalue with multiplicity 2. Assuming $x = (x_1, x_2)$ is a solution to $f(x, \lambda) = 0$ then

$$0 = f(x, \lambda) \cdot (-x_2, x_1) = x_1^4 + x_2^4$$

and hence $x = 0$. In particular f has only trivial solutions.

One unsolved issue is the fact that we want to solve general equations of the form

$$f(x, \lambda) = 0 , \qquad f : \mathscr{X} \times \mathbb{R} \to \mathscr{Y} ,$$

where $\mathscr{X}$, $\mathscr{Y}$ are general Banach spaces. We then find the following result.

Theorem 2.2.21. *Let $f : \mathscr{X} \times \mathbb{R} \to \mathscr{Y}$ be differentiable in $\mathscr{X}$ such that $\mathrm{D}_{\mathscr{X}} f(0,0)$ is Fredholm with index 0 and let $I : \mathscr{Y} \to \mathscr{X}$ be the isomorphism of Lemma 2.2.3 and such that*

$$I\, f(x, \lambda) = x - (\mu_0 + \lambda)\, Tx + g(x, \lambda)$$

satisfies the conditions of Theorem 2.2.18. Then $(0,0)$ is a point of bifurcation for f.

2.2.7 Global Bifurcation

We will now look at the question how the non-trivial branches of $f(x, \lambda) = 0$ behave once they leave the neighborhood of the bifurcation point.

Let $\mathscr{X}$ be a Banach space, $\hat{\Omega} \subset \mathscr{X} \times \mathbb{R}$ an open neighborhood of $(0, \lambda_0)$, $T : \mathscr{X} \to \mathscr{X}$ linear and compact, as well as $g \in \mathbb{C}(\mathrm{cl}(\hat{\Omega}); \mathscr{X})$ with $g(x, \lambda) = o(\|x\|)$ uniformly in λ. We consider

$$f(x, \lambda) = x - \lambda T x + g(x, \lambda) ,$$

where $\frac{1}{\lambda_0}$ is an eigenvalue of T with odd algebraic multiplicity.

We study the set of non-trivial solutions

$$M := \left\{ (x, \lambda) \in \hat{\Omega} \, : \; f(x, \lambda) = 0 \, , \; x \neq 0 \right\}$$

and the branch that originates in λ_0:

$$C(\lambda_0) = \text{connected component of } \overline{M} \text{ that contains } (0, \lambda_0) ,$$

where $A \subset \overline{M}$ is connected iff there do not exists disjoint open sets U_1, U_2 with $A \subset U_1 \cup U_2$.

Lemma 2.2.22. *$\overline{M}$ equipped with the topology*

$$d\left((x_1, \lambda_1) , (x_2, \lambda_2) \right) := \|x_1 - x_2\| + |\lambda_1 - \lambda_2|$$

is a locally compact metric space.

Proof. Let $\overline{f} : \mathscr{X} \times \mathbb{R} \to \mathscr{X} \times \mathbb{R}$, $\overline{f}(x,\lambda) = (f(x,\lambda),\lambda)$. By Lemma 1.1.22 we observe that $\mathbb{B}_\lambda^{\mathscr{X}\times\mathbb{R}}(0) \cap f^{-1}(\{0\} \times [-\lambda,\lambda])$ are precompact. □

Then $C(\lambda_0)$ is closed: For $(x_k,\lambda_k)_{k\in\mathbb{N}} \subset C(\lambda_0)$ with $x_k \to x$. If $x \notin C(\lambda_0)$ then $A := \{x\} \cup C(\lambda_0)$ is not connected and there exist disjoint open neighborhoods $U_1, U_2 \subset M$, $x \in U_1$ with $A \subset U_1 \cup U_2$ and hence $C \subset \tilde{U}_1 \cup U_2$, for $\tilde{U}_1 = U_1 \backslash \{x\}$. Furthermore, $U_1 \cap C(\lambda_0) = \emptyset$ since $C(\lambda_0) \subset U_1 \cup U_2$ is connected and $U_1 \cap U_2 = \emptyset$. On the other hand, $U_1 \supset B_\varepsilon(x) \cap \overline{M}$ for suitable $\varepsilon > 0$ and we find that x has positive distance from $(x_k,\lambda_k)_{k\in\mathbb{N}}$, a contradiction.

The main result of this section is the following.

Theorem 2.2.23 (Rabinowitz)**.** *$C(\lambda_0)$ tends to $\partial\hat{\Omega}$ or returns to the trivial branch* $(\{0\} \times \mathbb{R}) \setminus \{(0,\lambda_0)\}$.

In case $\hat{\Omega} = \mathscr{X} \times \mathbb{R}$ the first option implies that $C(\lambda_0)$ tends to ∞.

Proof. Assume the theorem was wrong. Then there exists a bounded subset $\Omega \subset \hat{\Omega}$ such that $C(\lambda_0) \cap \partial\Omega = \emptyset$ and $C(\lambda_0) \cap \{0\} \times \mathbb{R} = \{(0,\lambda_0)\}$. From the above considerations, we infer that $C(\lambda_0) \subset \overline{M}$ is locally compact and closed and hence $C(\lambda_0) \cap \overline{\Omega}$ is compact. Since $C(\lambda_0) \cap \partial\Omega = \emptyset$ this implies that $C(\lambda_0)$ has a positive distance from $\partial\Omega$ and there exists an open environment Ω_0 of $C(\lambda_0)$ that separates it from $\partial\Omega$. This will lead to a contradiction.

Without loss of generality, we might assume that there exists small δ with

$$\overline{\Omega}_0 \cap (\{0\} \times \mathbb{R}) = \{0\} \times [\lambda_0 - \delta, \lambda_0 + \delta] \,, \tag{2.4}$$

$$\mathbb{B}_\delta^{\mathscr{X}}(0) \times (\lambda_0 - \delta, \lambda_0 + \delta) \subset \Omega_0 \,. \tag{2.5}$$

Using the notation

$$\Omega_0(\lambda) := \{(x,\lambda) \in \Omega_0\} \,, \qquad \mathfrak{d}(\lambda) := \mathfrak{d}\left(f(\cdot,\lambda), \Omega_0(\lambda), 0\right) \,,$$

let us first observe that $\mathfrak{d}(\lambda) = \mathit{const}$. Indeed, this follows from the fact that $C(\lambda_0)$ has a positive distance from $\partial\hat{\Omega}$(because of (2.5)) and hence for every $\tilde{\lambda} \in (\lambda_0 - \delta, \lambda_0 + \delta)$ we can replace $\Omega_0(\lambda)$ by $\Omega_0(\tilde{\lambda})$ in a neighborhood (using (d5)) and then apply the homotopy invariance (d4).

We choose λ_1, λ_2 close to λ_0 with $\lambda_1 < \lambda_0 < \lambda_2$. Because of (d5)

$$\begin{aligned}\mathfrak{d}\left(f(\cdot,\lambda_i), \Omega_0(\lambda_i), 0\right) = \mathfrak{d}\left(f(\cdot,\lambda_i), \Omega_0(\lambda_i) \backslash \mathbb{B}_\rho(0), 0\right) \\ + \mathfrak{d}\left(f(\cdot,\lambda_i), \mathbb{B}_\rho(0), 0\right) \,.\end{aligned}$$

For λ_1 small enough, resp. λ_2 large enough, the first term on the right hand side becomes 0 (note that $C(\lambda)$ is compact). The second term is the index of $f(\cdot,\lambda_i)$ in 0. Since $\frac{1}{\lambda_0}$ is an odd eigenvalue, the index jumps in λ_0 and this is in contradiction with the above observation that the left hand side is constant. □

Example 2.2.24 (Example for global bifurcation)**.** On the interval $(0,\pi)$ we consider

$$\mathscr{X} = H_0^2(0,\pi) = \left\{u \in H^2(0,\pi) \,:\, u(0) = u(\pi) = 0\right\}$$

and

$$\partial_x^2 u + \lambda u - \Phi(\cdot, u(\cdot), \partial_x u(\cdot)) = 0 \,.$$

Here, $\Phi : (0,\pi)\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is nonlinear and satisfies for all (s,x,p): $\Phi(x,s,p)\leq C\left(|s|^2+|p|^2\right)$. The Laplace operator $\partial_x^2 : \mathscr{X}\to L^2(0,\pi)$ is invertible and writing $T := \left(\partial_x^2\right)^{-1}$ (with zero Dirichlet-conditions) we infer

$$f(u,\lambda) := u+\lambda Tu - N(u) = 0\,.$$

This equation satisfies the conditions of Theorem 2.2.11 and we find bifurcation in all eigenvalues $\lambda_k = -k^2$, $k\in\mathbb{N}$. Let C_k be the corresponding connected components. Since we have only countably many branches, we can globally track the behavior of the bifurcation branches.

Theorem 2.2.25. *For every $k\in\mathbb{N}$ the set $C^k := C(\lambda_k)$ is unbounded in $\mathscr{X}\times\mathbb{R}$ and consists of functions u with $k-1$ simple zeros in $(0,\pi)$.*

Proof. Let S^k be the set of functions $u\in\mathscr{X}$ with $k-1$ simple zeros in $(0,\pi)$ and $\partial_x u(0)\neq 0$, $\partial_x u(\pi)\neq 0$. The two branches bifurcating in $(0,\lambda_k)$ can be distinguished by $\partial_x u(0)>0$ and $\partial_x u(0)<0$.

Then S^k is open in $\mathscr{X}$ and $S^k\cap S^m=\emptyset$ for $k\neq m$. For the branch C_k we know that locally $u(s)=su_k+o(s)$, where $u_k=\sin kx$ is the eigenfunction to $\lambda_k=-k^2$. Hence locally $(u,\lambda)\in C^k$ satisfies $u\in S^k$.

Since C^k is connected, and S^n, S^m are mutually disjoint, we infer $C_k\not\subset\bigcup_n S^n$. Therefore, assuming existence of $(u,\lambda)\in C^k$ such that $u\notin S^k$ and assuming we are on the branch $\partial_x u(0)>0$, the connectedness of C_k implies existence of $(u,\lambda)\in C^k$ with a double zero, i.e. $\partial_x u(s)=u(s)=0$ for some $s\in(0,\pi)$. Since u satisfies a second order quasilinear ODE we find $u\equiv 0$ and hence $(u,\lambda)=(0,\lambda)$. This implies $(u,\lambda)=(0,\lambda_m)$, $m\neq k$, a contradiction. Hence we can conclude $(u,\lambda)\in C^k \Rightarrow u\in S^k$ and additionally that C^k does not return to the trivial branch, i.e. bifurcates to ∞. □

Chapter 3

ODE in infinite dimensions

3.1 Stability and Bifurcation

3.1.1 Linear ODE and Stability of Stationary Points

We consider the following ODE in $\mathbb{R}^d$:

$$\dot{f}(t) = A\, f(t)\,, \qquad f(0) = f_0\,, \tag{3.1}$$

where $A \in \mathbb{R}^{d\times d}$ is a diagonal matrix. It is easy to see that

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{pmatrix} \qquad \Rightarrow \qquad A^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_d^k \end{pmatrix}.$$

from the 1-dimensional case we infer that a solution to (3.1) is given by

$$f(t) = \exp(At)\, f_0 := \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k\, f_0\,.$$

The last formula also holds for general matrices $A \in \mathbb{R}^{d\times d}$. Furthermore, we can observe the following for Hilbert spaces $\mathscr{Y}$ and $A : \mathscr{Y} \to \mathscr{Y}$:

$$\begin{aligned} A\, \exp(At)\, f_0 : &= A \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k\, f_0 \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^{k+1}\, f_0 = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{k=0}^{\infty} \frac{1}{(k+1)!} t^{k+1} A^{k+1}\, f_0 \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \exp(At)\, f_0\,, \end{aligned}$$

and hence $f(t) = \exp(At)\, f_0$ is a solution of (3.1) in general Hilbert spaces.

What do we gain from this insight?

In the finite dimensional case, let $\lambda_i < 0$ for all $i = 1, \ldots, d$. We then infer that $\|f(t)\| \to 0$ decreases exponentially with rate $\min|\lambda_i|$. On the other hand, if one of the eigenvalues of A

is positive with eigenvector e_i and $f_0 \cdot e_i \neq 0$, this implies that $\|f(t)\| \to \infty$ exponentially. We intend to use this insight for the discussion of nonlinear ODE.

More precisely, we study

$$\dot{u} = f(t, u(t)) \,, \qquad u(0) = u_0 \,,$$

where $f \in C^2([0,T] \times \mathbb{R}^d)$ for simplicity. Using a Taylor expansion for f in $u \approx \tilde{u}$ and assuming $f(\cdot, \tilde{u}) = const$, the above ODE takes the form

$$\dot{u} = f(u(t)) = f(\tilde{u}) + \mathrm{D}f(\tilde{u})\,(u - \tilde{u}) + o\,(u - \tilde{u}) \;.$$

If $\mathrm{D}f(\tilde{u})$ is non-degenerate, and u_0 is close enough to $\tilde{u}$, the behavior of the right hand side is strongly dominated by the second term $\mathrm{D}f(\tilde{u})\,(u-\tilde{u})$ since $\frac{o(u-\tilde{u})}{|u-\tilde{u}|} \to 0$.

Hence, if all eigenvalues of $L = \mathrm{D}f(\tilde{u})$ are negative and $|u_0|$ is small enough, we find that $u(t) \to \tilde{u}$ as $t \to \infty$. On the other hand, if L has a positive eigenvalue $\lambda_1 > 0$ with eigenvector x_1 and $u_0 \cdot x_1 \neq 0$ we find $|u(t) - \tilde{u}|$ monotone increasing at least for short times.

What else can happen?

Consider

$$u(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \qquad \dot{u}(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} u(t) \,.$$

We call this a periodic orbit. Note that we have proven the existence of periodic orbits for periodic right hand sides in Proposition **??**. Note that the later calculation becomes much more complicated for general rotation matrices.

Based on the above observations, we define the following:

Definition 3.1.1. A stationary solution x_0 of $f(x_0) = 0$ is called

- stable if for some $\alpha > 0$ all spectral values λ of $\mathrm{D}_{\mathscr{X}} f(x_0)$ satisfy $\lambda < -\alpha$.
- unstable if there exists a positive eivenvalue.

As an outlook to the next section, note that equations of the form (3.1) can be interpreted as a a so called gradient flow: For a symmetric $A \in \mathbb{R}^{d \times d}$ define $E(u) := \frac{1}{2}\langle u,\, Au \rangle$. Then (3.1) is equivalent with

$$\dot{u} = \nabla E(u) \,.$$

This can be brought in a more general setting. For a lower semicontinuous functional $\mathcal{E} : \mathscr{X} \to \mathbb{R}$, where $\mathscr{X}$ is a reflexive Banach space with dual $\mathscr{X}^*$ and a convex functional $\Psi^* : \mathscr{X}^* \to \mathbb{R}$ consider

$$\dot{u} = \mathrm{D}_{\mathscr{X}} \Psi\left(-\mathrm{D}\mathcal{E}(u)\right) \,.$$

This is the doubly nonlinear gradient equation, which we will study in Section 3.2.

3.1.2 Stability Criteria

In this section, we analyze under what conditions stationary solutions of

$$\dot{u} = f(u)$$

are stable.

Let us recall Theorem 2.2.11. We are particularly interested in the case when the bifurcating eigenvalue is simple. Since we have a bifurcating branch in $(0,0)$, we particularly infer that $\mathrm{D}_{\mathscr{X}} f(0,0)$ has a simple eigenvalue 0. Provided all eigenvalues on the trivial branch are negative for $\lambda < 0$, the trivial solution is stable and we are interested in stability of the three remaining branches, that is we are interested in the stability of the bifurcating branches $(u(s), \lambda(s))$.

Definition 3.1.2. Let $T_0, K : \mathscr{X} \to \mathscr{Y}$ be linear continuous, then μ_0 if called K-simple eigenvalue of T_0 if

1. $T_0 - \mu_0 K$ is a Fredholm operator with index 0,
2. $\ker(T_0 - \mu_0 K) = \mathbb{R}\tilde{x}_0$,
3. and $K\tilde{x}_0 \notin R(T_0 - \mu_0 K)$.

If $\mathscr{X} = \mathscr{Y}$, $K = \mathrm{id}$ and T_0 is compact, we call μ_0 a simple eigenvalue of T_0.

Lemma 3.1.3 (Continuation of simple eigenvalue)**.** *Let μ_0 be a K-simple eigenvalue of T_0 with eigenvector $\tilde{x}_0$. Then there exists a continuous extension*

$$\mu : \mathbb{B}_\delta^{\mathcal{L}(\mathscr{X};\mathscr{Y})}(T_0) \to \mathbb{R}$$

with $\mu(T_0) = \mu_0$ and $\mu(T)$ is K-simple eigenvalue of T This extension μ is unique. Furthermore it holds

$$\ker(T - \mu(T)K) = \mathbb{R}x(T)\,, \qquad x(T) = \tilde{x}_0 + x_1(T) \in \mathscr{X} = \mathbb{R}\tilde{x}_0 \oplus \mathscr{X}_1\,.$$

μ and x_1 are analytic in T.

Proof. W.l.o.g. let $\mu_0 = 0$. In view of the statement of the Lemma, we want to solve

$$(T - \mu(T)K)\, x(T) = 0\,.$$

Hence we define in an environment of $(0, 0, T_0) \in \mathbb{R} \times \mathscr{X}_1 \times \mathcal{L}(\mathscr{X};\mathscr{Y})$ the function

$$F(r, x_1, T) := (T - rK)\,(\tilde{x}_0 + x_1) \in \mathscr{Y}\,,$$

where T is now a parameter of F.

We make use of $\mu_0 = 0$ which implies that $\tilde{x}_0$ is eigenvector of T_0 with eigenvalue 0 and hence $\mathscr{Y} = \mathscr{Y}_0 + \mathscr{Y}_1 = \mathscr{Y}_0 + R(T_0)$, $\dim \mathscr{Y}_0 = 1$ and with corresponding projections P_0 and $P_1 = \mathrm{id} - P_0$. Then

$$\begin{aligned}\mathrm{D}_{r,\mathscr{X}_1} F(0, 0, T_0) : \ (r, x_1) &\mapsto -rK\tilde{x}_0 + T_0 x_1 \\ &= (-rP_0 K\tilde{x}_0,\ P_1(-rK\tilde{x}_0) + P_1 T_0 x_1)\ .\end{aligned}$$

Since $P_0(K\tilde{x}_0) \neq 0$ and $P_1T_0 : \mathscr{X}_1 \to \mathscr{Y}_1$ is invertible, the functional $\mathrm{D}_{r,\mathscr{X}_1}F(0,0,T_0)$ is invertible. Hence the implicit function theorem yields existence of $\mu(T)$ and $x_1(T)$ close to T_0.

It remains to show that these solutions a) are globally unique and b) satisfy properties 1.–3. of a K-simple eigenvalue.

a) The ansatz $x = \tilde{x}_0 + x_1$ has a very particular structure, especially because of the assumption that x_1 is small. For a different value of x_1, particularly with larger norm, there might exist a second solution. Hence, in what follows, we assume that $x = \beta\tilde{x}_0 + x_1$ is a normalized solution, i.e. $\|x\| = 1$, of

$$\begin{aligned} &(T - rK)\,(\beta\tilde{x}_0 + x_1) = 0 \\ \Leftrightarrow \qquad & T_0x_1 - r\beta K\tilde{x}_0 = (T - T_0)\,(\beta\tilde{x}_0 + x_1) + rKx_1\,. \end{aligned} \tag{3.2}$$

The terms on the left hand side have the structure $T_0x_1 \in R(T_0)$ and $r\beta K\tilde{x}_0 \notin R(T_0)$. Therefore, there holds an estimate

$$\|x_1\| + |r\beta| \leq C\left(\,\|T_0 - T\|\,(|\beta| + \|x_1\|) + |r|\,\|x_1\|\,\right)\,.$$

This follows from a contradiction argument: otherwise there exist $x_{1,n}$, β_n, T_n and r_n such that

$$1 = \|x_{1,n}\| + |r_n\beta_n| \geq n\left(\,\|T_0 - T_n\|\,(|\beta_n| + \|x_{1,n}\|) + |r_n|\,\|x_{1,n}\|\,\right)\,,$$

but by (3.2) this would imply in the limit $n \to \infty$

$$T_0x_{1,n} - r_n\beta_nK\tilde{x}_0 \to 0\,.$$

Since $T_0x_{1,n} \in R(T_0)$ and $r_n\beta_nK\tilde{x}_0 \notin R(T_0)$ this would imply $x_{1,n} \to 0$ and $r_n\beta_n \to 0$, a contradiction.

Assuming that $\|T_0 - T\|$ and r are small, we can adsorb the terms depending on x_1 on the left hand side and obtain

$$\|x_1\| + |r\beta| \leq C\,\|T_0 - T\|\,|\beta|\,.$$

Since β is bounded (i.e. $\|x\| = 1$), we obtain that $\|x_1\|$ and r are small. Smallness of $\|x_1\|$ yields smallness of $(1 - \beta)$ (again $\|x\| = 1$). But for small x_1 and $\beta \approx 1$ the implicit function theorem already provided uniqueness of solutions.

b) Property 2. follows from the uniqueness of $x_1(T)$. In particular, if $\ker(T - rK) = \mathbb{R}\tilde{x}_0 \oplus \mathbb{R}\hat{x}_0$, then

$$(T - rK)\,(\tilde{x}_0 + x_1) = (T - rK)\,(\tilde{x}_0 + x_1 + \mathbb{R}\hat{x}_0)\,.$$

Property 3. follows from the fact that

$$\hat{T} :\, (x_1, r) \mapsto (T - \mu(T)K)\,x_1 + rK\,(\tilde{x}_0 + x_1(T)) \in \mathscr{Y}$$

is invertible at T_0 and hence also for small (x_1, r). From this invertibility, we obtain

$$rK\,(\tilde{x}_0 + x_1(T)) \notin R\,(T - \mu(T)K)\,.$$

Property 1. follows from the following Lemma 3.1.4. □

Lemma 3.1.4. *Let $T_0 \in \mathcal{L}(\mathscr{X};\mathscr{Y})$ be Fredholm. For all T in a neighborhood of T_0 holds that T is also a Fredholm operator and*

$$\dim\ker T \le \dim\ker T_0\,,\qquad \operatorname{codim} R(T) \le \operatorname{codim} R(T_0)\,.$$

If the dimensions on the right hand side are 1, then T is Fredholm with index 0.

Proof. For $T_0 : \mathscr{X} \to \mathscr{Y}$ let $\mathscr{X} = \mathscr{X}_0 \oplus \mathscr{X}_1$, $\mathscr{Y} = \mathscr{Y}_0 \oplus \mathscr{Y}_1$ such that $T_0|_{\mathscr{X}_1} \to \mathscr{Y}_1$ is invertible. Hence $\tilde{T}_0 : \mathscr{X}_1 \times \mathscr{Y}_0 \to \mathscr{Y}$, $(x_1, y_0) \mapsto T_0 x_1 + y_0$ is invertible and for operators T such that $\|T - T_0\|$ small also $\tilde{T} : (x_1, y_0) \mapsto T x_1 + y_0$ is invertible, implying $\tilde{T}$ is Fredholm with index 0. The estimates on the dimensions follow from the the fact that we restricted T to $\mathscr{X}_1$.

In order to show the last statement, we show $R(T) = \mathscr{Y}$ implies $\ker T = \{0\}$. In case $T(\mathscr{X}_1) = \mathscr{Y}$, the map $\tilde{T}$ would not be injective. Hence $T(\mathscr{X}_1) \ne \mathscr{Y}$, but then also $Tx_0 \notin T(\mathscr{X}_1)$ since $R(T) = \mathscr{Y}$. This implies $\ker T = \{0\}$. □

In the following we consider the situation that $\mathscr{X} \subset \mathscr{Y}$ and $K = j$ is the inclusion. Furthermore assume that the situation of Theorem 2.2.11 is given. In particular, we have two branches $(0, \lambda)$ and $(x(s), \lambda(s))$, where $s \in (-\varepsilon, \varepsilon)$ is sufficiently small and $x(s) = s\tilde{x}_0 + x_1(s)$. We are interested in the eigenvalues of $\mathrm{D}_{\mathscr{X}} f$ on the branches. According to Lemma 3.1.3 there exist $\tilde{u}$, $\tilde{\mu}$, $\bar{u}$, $\bar{\mu}$ such that

$$\mathrm{D}_{\mathscr{X}} f(0,\lambda)\bar{u}(\lambda) = \bar{\mu}(\lambda)\bar{u}(\lambda)\,,$$
$$\mathrm{D}_{\mathscr{X}} f(x(s),\lambda(s))\tilde{u}(s) = \tilde{\mu}(s)\tilde{u}(s)\,.$$

We will now determine the signs of $\bar{\mu}$ and $\tilde{\mu}$.

Theorem 3.1.5 (Crandall-Rabinowitz)**.** *In the above situation (i.e. under the prerequisites of Theorem 2.2.11) it holds $\bar{\mu}'(0) \ne 0$, $s\lambda(s)' \to 0$ and $\tilde{\mu}(s) \to 0$ for $s \to 0$. Moreover for every sequence $s \to 0$ with $\tilde{\mu}(s) \ne 0$ it holds*

$$\frac{s\lambda'(s)\bar{\mu}'(0)}{\tilde{\mu}(s)} \to -1 \qquad \text{as } s \to 0\,.$$

Proof. Trivial branch: We differentiate the eigenvalue equality

$$\mathrm{D}_{\mathscr{X}} f(0,\lambda)\bar{u}(\lambda) = \bar{\mu}(\lambda)\bar{u}(\lambda)$$

after λ and obtain

$$\partial_\lambda \mathrm{D}_{\mathscr{X}} f(0,\lambda)\bar{u}(\lambda) + \mathrm{D}_{\mathscr{X}} f(0,\lambda)\partial_\lambda\bar{u}(\lambda) = \partial_\lambda\bar{\mu}(\lambda)\bar{u}(\lambda) + \bar{\mu}(\lambda)\partial_\lambda\bar{u}(\lambda)\,.$$

Because $\partial_\lambda \mathrm{D}_{\mathscr{X}} f(0,0)\bar{u}(0) = \partial_\lambda L(0)\bar{u}(0) \notin R(L(0))$ and $\mathrm{D}_{\mathscr{X}} f(0,0)\partial_\lambda\bar{u}(0) \in R(L(0))$, the left hand side is not 0 in $\lambda = 0$. On the right hand side, we obtain $\bar{\mu}(0) = 0$ and hence $\partial_\lambda\bar{\mu}(0) \ne 0$.

For later purpose, consider $y^* : \mathscr{Y} \to \mathbb{R}$ continuous, linear with $\ker y^* = R(L(0))$. Then

$$\left\langle y^*,\, \partial_\lambda \mathrm{D}_{\mathscr{X}} f(0,0)\bar{u}(0)\right\rangle = \partial_\lambda\bar{\mu}(0)\left\langle y^*,\, \bar{u}(0)\right\rangle\,. \tag{3.3}$$

Non-trivial branch: We use the eigenvalue equation

$$\mathrm{D}_{\mathscr{X}} f(x(s),\lambda(s))\tilde{u}(s) = \tilde{\mu}(s)\tilde{u}(s)\,.$$

and the solution equation $f(x(s), \lambda(s)) = 0$, which we differentiate after s:

$$\mathrm{D}_{\mathscr{X}} f(x(s), \lambda(s))\partial_s x(s) + \mathrm{D}_\Lambda f(x(s), \lambda(s))\partial_s \lambda(s) = 0\,.$$

We take the difference of the above equations and obtain

$$\mathrm{D}_{\mathscr{X}} f(x(s), \lambda(s))\left(\partial_s x(s) - \tilde{u}(s)\right) + \mathrm{D}_\Lambda f(x(s), \lambda(s))\partial_s \lambda(s) = \tilde{\mu}(s)\tilde{u}(s)\,.$$

In what follows, we develop $\mathrm{D}_{\mathscr{X}} f$, $\mathrm{D}_\Lambda f$ and $\tilde{u}$ as "Taylor-series" in $s = 0$ and obtain (noting that $o(1)$ means "small in s"):

$$\begin{aligned} \mathrm{D}_{\mathscr{X}} f(0,0)\left(\partial_s x(s) - \tilde{u}(s)\right) + o(1)\left(\partial_s x(s) - \tilde{u}(s)\right) \\ + \mathrm{D}_{\mathscr{X}}\mathrm{D}_\Lambda f(0,0)\left\langle \partial_s x(s), s\right\rangle \partial_s\lambda(s) + \mathrm{D}_\Lambda\mathrm{D}_\Lambda f(0,0)\left\langle \partial_s \lambda(s), s\right\rangle \partial_s\lambda(s) \\ + o(s)\partial_s\lambda(s) + \tilde{\mu}(s)\tilde{u}(0) + \tilde{\mu}(s)o(1) = 0\,. \end{aligned}$$

We use $x(s) = s\tilde{x}_0 + x_1(s)$ from Theorem 2.2.11 and $\tilde{u}(s) = \tilde{x}_0 + u_1(s)$ from Lemma 3.1.3. This yields $\partial_s x(s) - \tilde{u}(s) \in \mathscr{Y}_1 = R(L(0))$. Since $L(0) : \mathscr{X}_1 \to \mathscr{Y}_1$ is invertible, the above equation yields

$$\|\partial_s x(s) - \tilde{u}(s)\| \le C\left(|s\partial_s\lambda(s)| + |\tilde{\mu}(s)|\right)\,.$$

We us y^* from (3.3) and find because of $\langle y^*, \mathrm{D}_{\mathscr{X}} f(0,0)\dots\rangle = 0$ and $\mathrm{D}_\Lambda\mathrm{D}_\Lambda f(0,0) = 0$ and $\ker y^* = R(L(0))$ that

$$\left\langle y^*, \mathrm{D}_{\mathscr{X}}\mathrm{D}_\Lambda f(0,0)\left\langle \tilde{x}_0 + \partial_s x_1(s)\right\rangle\right\rangle s\partial_s\lambda(s) + \tilde{\mu}(s)\left\langle y^*, \tilde{x}_0\right\rangle = o(1)\left(|s\partial_s\lambda(s)| + |\tilde{\mu}(s)|\right)$$

Using (3.3) we obtain

$$s\partial_s\lambda(s)\partial_\lambda\bar{\mu}(0) + \tilde{\mu}(s) = o(1)\left(|s\partial_s\lambda(s)| + |\tilde{\mu}(s)|\right)$$

which implies the statement. □

Example 3.1.6. Let $G \subset \mathbb{R}^d$ be a bounded domain with smooth boundary and consider the PDE

$$\begin{aligned} f(u,\lambda) = \Delta u + (\mu_0 + \lambda)u + u^2 &= 0 && \text{in } G\,, \\ u &= 0 && \text{on } \partial G\,. \end{aligned}$$

As shown in the theory of PDE, the eigenvalues of $-\Delta$ are positive with $0 < \mu_0 < \mu_1 \le \dots$ with $\mu_k \to \infty$. Here, μ_0 is a simple eigenvalue with a positive eigenfunction u_0. In particular, the inverse operator with Dirichlet boundary conditions $K := (-\Delta)^{-1}$ is compact (As an Operator on $\mathscr{X} = C(\overline{G})$ and on $\mathscr{X} = L^2(G)$) with eigenvalues $\frac{1}{\mu_0} > \frac{1}{\mu_1} \ge \dots \ge 0$, smoot eigenfunctions $u_0, u_1, \dots$ and we may instead consider the problem

$$\tilde{f}(u,\lambda) = -u + (\mu_0 + \lambda)Ku + Ku^2 = 0\,.$$

From Theorem 2.2.11 we find that the trivial branch of f bifurcates in $\lambda = 0$ since $\mathrm{D}_{\mathscr{X}}\tilde{f}(0,0) = -id + \mu_0 K$ degenerates with a simple eigenvector $u_0 \in C^\infty$ and

$$\mathrm{D}_\lambda\mathrm{D}_{\mathscr{X}} f(0,0)u_0 = Ku_0 \notin R(\mathrm{id} - \mu_0 K)\,.$$

Hence there exists a non-trivial branch bifurcating in $(0,0)$ of the form $(u(s),\lambda(s))$, where $u(s)=su_0+u_1(s)$.

We will now analyze the bifurcation diagram in terms of $(s,\lambda(s))$. We differentiate the equation $\tilde f(u(s),\lambda(s))=0$ w.r.t. s and obtain

$$(\mathrm{id}-\mu_0 K)\left(u_0+\partial_s u_1(s)\right)=\lambda(s)K\partial_s u(s)+\partial_s\lambda(s)Ku(s)+2Ku(s)\partial_s u(s)\,.$$

Differentiating once more and evaluating in $s=0$ (note that $u(0)=0$) we obtain

$$(\mathrm{id}-\mu_0 K)\,\partial_s^2 u_1(0)=2\partial_s\lambda(0)K\partial_s u(0)+2K\left|\partial_s u(0)\right|^2\,.$$

Note that $\partial_s^2 u_1(0)\in\mathscr{X}_1$ and hence the left hand side is orthogonal (in $L^2(G)$) to u_0. We use $\partial_s u(0)=u_0$ and self-adjointness of K: Multiplying the last equation with u_0 and integrating over G, we hence obtain (using non-negativity of the first eigenfunction, $u_0>0$)

$$0\geq 2\partial_s\lambda(0)\left\|u_0\right\|_{L^2}^2+\int_G 2\left|u_0\right|^3\,.$$

This particularly implies $\lambda'(0)<0$. Hence, $\lambda(s)>0$ iff $s<0$ and $\lambda(s)<0$ iff $s>0$.

We can conclude the following: For $\lambda<0$, the eigenvalues of $\mathrm{D}_{\mathscr{X}}\tilde f$ are all negative and for $\lambda>0$ $\mathrm{D}_{\mathscr{X}}\tilde f$ has a positive eigenvalue. Hence the trivial branch is stable for $\lambda<0$ and unstable for $\lambda>0$. Since $\partial_\lambda\overline{\mu}=1>0$, we obtain from Crandall-Rabinowitz that

$$\mathrm{sign}(s\lambda'(s))=-\mathrm{sign}(s),\ \mathrm{sign}(\partial_\lambda\overline{\mu}(0))=1\quad\Rightarrow\quad\mathrm{sign}(\tilde\mu(s))=\mathrm{sign}(s)\,,$$

and the non-trivial branch is stable for $s<0$ and unstable for $s>0$.

Example 3.1.7. With the same G consider

$$\begin{aligned} f(u,\lambda)=\Delta u-(\mu_0+\lambda)u+u^3&=0 &&\text{in } G\,,\\ u&=0 &&\text{on } \partial G\,. \end{aligned}$$

Again we find that the trivial branch of f bifurcates in $\lambda=0$ with $(u(s),\lambda(s))$, where $u(s)=su_0+u_1(s)$.

We differentiate the equation $f(u(s),\lambda(s))=0$ w.r.t. s and obtain

$$(\Delta-\mu_0)\left(u_0+\partial_s u_1(s)\right)=\lambda(s)\partial_s u(s)+\partial_s\lambda(s)u(s)+3u^2(s)\partial_s u(s)\,,$$

i.e. $\partial_s u_1(0)=0$. Differentiating once more and evaluating in s we obtain

$$(\Delta-\mu_0)\,\partial_s^2 u_1(s)=\lambda(s)\partial_s^2 u(s)+2\partial_s\lambda(s)\partial_s u(s)+\partial_s^2\lambda(s)u(s)+3u^2(s)\partial_s^2 u(s)+6u(s)\left|\partial_s u(s)\right|^2\,.$$

We take the L^2-product with u_0 and obtain

$$0=2\partial_s\lambda(s)\left\|u_0\right\|_{L^2}^2+\partial_s^2\lambda(s)s\left\|u_0\right\|_{L^2}^2+6\int_G s\left|u_0\right|^2\left|\partial_s u(s)\right|^2+O(s^2)\,.$$

Using $\partial_s u_1(0)=0$ we obtain yields $\partial_s\lambda(0)=0$. Hence a Taylor expansion in $\partial_s\lambda(0)$ yields

$$2\partial_s^2\lambda(0)s\left\|u_0\right\|_{L^2}^2+\partial_s^2\lambda(s)s\left\|u_0\right\|_{L^2}^2+6\int_G s\left|u_0\right|^2\left|\partial_s u(s)\right|^2+O(s^2)$$

and dividing by $s = 0$ and $s \to 0$ we obtain

$$\partial_s^2 \lambda(0) = -\frac{2}{\|u_0\|^2} \int_G |u_0|^4 \,.$$

In particular, $\partial_s^2\lambda(0) < 0$ and the bifurcation diagram yields a bifurcation "to the left". This behavior is called "subcritical". The trivial branch has been analyzed above and we obtain from Crandall-Rabinowitz that the non-trivial branch is stable.

3.1.3 Hopf Bifurcation

In what follows, let $\mathscr{X}$ be - as usual - a Banach space and let $T(x,t,\lambda) : \mathscr{X} \to \mathscr{X}$ be a family of "dynamical systems". In the following, we basically rely on the property $T(x,t,\lambda) \circ T(x,s,\lambda) = T(x,s+t,\lambda)$. The map $T(\cdot,t,\lambda)$ is called Poincaré map if for every $u_0 \in \mathscr{X}$ the function $T(u_0,t,\lambda)$ solves the evolution equation

$$\dot{u} = f(u(t),\lambda)\,, \qquad u(0) = u_0\,.$$

Lemma 3.1.8. *Let $T : \mathscr{X} \times \mathbb{R} \times \mathbb{R} \to \mathscr{X}$ be C^2 and let*

$$L(t,\lambda) : \mathscr{X} \to \mathscr{X}\,, \qquad x \mapsto \mathrm{D}_{\mathscr{X}} T(0,t,\lambda)\,\langle x\rangle \,.$$

We make the following assumptions

1. $\mathrm{id} - L(t_0,\lambda_0)$ *is Fredholm Operator with index* 0 *and has a* 2*-dimensional kernel* $\mathscr{X}_0 = \mathrm{span}(u,v)$*. In particular we find a projection* P *onto the co-space of* $\mathscr{Y}_1 := R(\mathrm{id} - L(t_0,\lambda_0))$ *with* $\ker P = \mathscr{Y}_1$.

2. *The following map is invertible:*

$$\mathrm{D}_{(t,\lambda)} P\left(L(t_0,\lambda_0)\right)\langle u\rangle : \mathbb{R}\times\mathbb{R} \to \mathscr{Y}_0\,.$$

Then there exists a family of fixed points (the dynamical system has a family of periodic solutions): For small s we find starting points $x(s) = su + x_1(s)$ with periods $t(s) = t_0 + o(1)$ and $\lambda(s) = \lambda_0 + o(1)$ such that

$$T\left(x(s),t(s),\lambda(s)\right) = x(s)\,.$$

Proof. We use a Lyapunov-Schmidt reduction with $\Lambda = \mathbb{R}^2$ for the equation $x - T(x,t,\lambda) = 0$:

$$\begin{aligned} x &= x_0 + x_1(x_0,t,\lambda)\,,\\ \Phi(x_0,t,\lambda) &= Px_0 - P\,T\left(x_0 + x_1(x_0,t,\lambda),t,\lambda\right) = 0\,. \end{aligned}$$

Note that $x_1(0,t,\lambda) = 0$ and $\mathrm{D}_{\mathscr{X}_0} x_1(0,t_0,\lambda_0) = 0$. Using the particular ansatz $x_0 = su$ the bifurcation equation takes the equivalent form (compare to the proof of Theorem 2.2.11)

$$\Psi(s,t,\lambda) = \begin{cases} \frac{1}{s}\Phi(su,t,\lambda) & \text{for } s \neq 0\,,\\ \mathrm{D}_{\mathscr{X}}\Phi(0,t,\lambda) & \text{for } s = 0\,. \end{cases}$$

The non-trivial zeros of Φ correspond to the zeros of Ψ and Ψ is of class C^1.

A given non-trivial zero is

$$\Psi\left(0,t_0,\lambda_0\right) = P\,u - P\,L(t_0,\lambda_0)\left\langle u\right\rangle = 0\,.$$

We solve $(t,\lambda)=(t(s),\lambda(s))$ for small s using the implicit function theorem applied to

$$\mathrm{D}_{(t,\lambda)}\Psi(0,t_0,\lambda_0) = \mathrm{D}_{(t,\lambda)}\left(P\,u - P\,L(t,\lambda)\left\langle u\right\rangle\right)\Big|_{(t_0,\lambda_0)} = \mathrm{D}_{(t,\lambda)}P\,L(t_0,\lambda_0)\left\langle u\right\rangle\,,$$

which is invertible by assumption. □

Theorem 3.1.9. *Let $f\in C^2(\mathbb{R}^d\times\mathbb{R};\mathbb{R}^d)$ such that $f(0,\lambda)=0$ for every $\lambda\in(-\delta,\delta)$. We assume*

1. *The linear map $A(\lambda):=\mathrm{D}_{\mathscr{X}}f(0,\lambda)$ has for $\lambda=\lambda_0$ the eigenfalues $\pm i\beta_0$, $\beta_0\in\mathbb{R}_+$. (imaginary eigenvalues)*
2. *All multiples $ki\beta_0$ are not eigenvalues of $A(\lambda_0)$ (Resonance)*
3. *The C^1-continuation $\alpha(\lambda)+i\beta(\lambda)$ of the eigenvalue $i\beta_0$ satisfies $\partial_\lambda\alpha(\lambda_0)\neq 0$.*

Then the equation $\dot{x}=f(x,\lambda)$ has a family of non-trivial periodic solutions of the form $x(t,s)$ to parameters $\lambda(s)=\lambda_0+o(1)$ with period $t(s)=\frac{2\pi}{\beta_0}+o(1)$ with inital values $x(0,s)=su+o(s)$ for small s.

Proof. We verify that Lemma 3.1.8 can be applied for $t_0=\frac{2\pi}{\beta_0}$ and T the dynamical system for f. We start with the following claim:

If $T(\cdot,t,\lambda)$ is the Poincaré mapping corresponding to $x\mapsto f(x,\lambda)$, then $L(t,\lambda)\left\langle\cdot\right\rangle$ is the Poincaré mapping corresponding to $x\mapsto Ax$. This follows from differentiating $\partial_t T(x_0,t,\lambda)=f(T(x_0,t,\lambda),\lambda)$ w.r.t. direction $\overline{x}$:

$$\begin{aligned}\partial_t L(t,\lambda) &= \mathrm{D}_{\mathscr{X}}\partial_t T(0,t,\lambda)\left\langle\overline{x}\right\rangle\\ &= \mathrm{D}_{\mathscr{X}}f(T(0,t,\lambda),\lambda)\mathrm{D}_{\mathscr{X}}T(0,t,\lambda)\left\langle\overline{x}\right\rangle = A\,L(t,\lambda)\,.\end{aligned}$$

We consider the evolution in $\mathbb{C}^d$ and split it in the generalized eigenspaces to A. The evolution of an eigenvector v with eigenvalue μ has the form

$$v(t)=e^{\mu t}v\,,\qquad \dot{v}=A\,v\,.$$

Hence the kernel of $\mathrm{id}-L(t_0,\lambda_0)$ is spanned by the eigenvectors of A with $e^{\mu t_0}=1$, i.e. $\mu=ik\beta_0$. Assumptions 1. and 2. hence imply that the kernel is 2-dimensional.

We chose two eigenvectors u, v such that $Au=i\beta_0u$, $Av=-i\beta_0v$. For the corresponding real basis vectors we chose $w_1=\mathrm{Re}u$ and $w_2=\mathrm{Im}u$. These vectors satisfy

$$Aw_1=\mathrm{Re}\left(Au\right)=\mathrm{Re}\left(i\beta_0u\right)=-\beta_0\mathrm{Im}u=\beta_0w_2$$

and similar $Aw_2=\beta_0w_1$. The projection P is the projection onto $\mathrm{span}\left(w_1,w_2\right)$.

We extend u and v to eigenvectors $u(\lambda)$, $v(\lambda)$ with $P(u(\lambda)) = u$, $P(v(\lambda)) = v$ and eigenvalues $\alpha(\lambda) \pm i\beta(\lambda)$. For these eigenvectors we calculate

$$\mathrm{D}_{(t,\lambda)} P\, L\,(t_0, \lambda_0) : \mathbb{R} \times \mathbb{R} \to \mathscr{X}_0\,.$$

It holds

$$L(t, \lambda) : \; u(\lambda) \mapsto \exp\left(\left(\alpha(\lambda) \pm i\beta(\lambda)\right) t\right) u(\lambda)\,,$$

and similar to PL. We differentiate in t and λ and obtain

$$\begin{aligned} \partial_t PL\,(t_0, \lambda_0)\,\langle u \rangle &= i\beta_0 u\,, \\ \partial_\lambda PL\,(t_0, \lambda_0)\,\langle u \rangle &= \left(\alpha'(\lambda_0) t_0 + i\beta'(\lambda_0)\right) u\,. \end{aligned}$$

Using real vectors w_1 and w_2 this reads

$$\begin{aligned} \partial_t PL\,(t_0, \lambda_0)\,\langle w_1 \rangle &= -\beta_0 w_2\,, \\ \partial_\lambda PL\,(t_0, \lambda_0)\,\langle w_1 \rangle &= \alpha'(\lambda_0) t_0 w_1 - \beta'(\lambda_0) w_2\,. \end{aligned}$$

Since $\alpha'(\lambda_0) \neq 0$, the map $\mathrm{D}_{(t,\lambda)} P\, L\,(t_0, \lambda_0)\,\langle w_1 \rangle$ is invertible and we can apply Lemma 3.1.8. □

Example 3.1.10. Consider the damped oscillator

$$y'' + (y')^3 - \lambda y' + y = 0\,.$$

Writing $u = (y, y')$, we arrive at

$$u' = \begin{pmatrix} 0 & 1 \\ -1 & \lambda \end{pmatrix} u - \begin{pmatrix} 0 \\ u_2^3 \end{pmatrix}\,.$$

The eigenvalues are given by $\mu_0 \in \mathbb{C}$ with

$$-\mu\,(-\mu + \lambda) + 1 = 0 \quad \Rightarrow \quad \mu_{1,2} = \frac{1}{2}\left(\lambda \pm \sqrt{\lambda^2 - 4}\right)\,.$$

In case $\lambda = 0$ we find conjugate complex eigenvalues with $\partial_\lambda \mathrm{Re}\mu_{1,2} = \frac{1}{2} \neq 0$. Hence we have Hopf-bifurcation in $\lambda = 0$ and there are non-trivial solutions with small amplitude close to 0.

Example 3.1.11 (Van der Pol oscillator)**.** Consider

$$y'' + f(y)y' + g(y) = 0\,, \qquad \Phi(y) := \int_0^y f(s)\mathrm{d}s\,.$$

we set $u_1 = y$ and $u_2 := y' + \Phi(y)$. Then

$$u' = \begin{pmatrix} u_2 - \Phi(u_1) \\ -g(u_1) \end{pmatrix}\,.$$

Here we consider $g(y) = y$ and $f(y) = \lambda(y^2 - 1)$ and $\Phi(y) = \lambda\left(\frac{1}{3}\lambda^3 - y\right)$:

$$u' = \begin{pmatrix} -\lambda \frac{1}{3} u_1^3 + \lambda u_1 + u_2 \\ -u_1 \end{pmatrix}\,.$$

Like in last example, linearization around $\lambda = 0$ and $u = 0$ yields

$$u' = \begin{pmatrix} \lambda & 1 \\ -1 & 0 \end{pmatrix} u\,.$$

Example 3.1.12 (Hamilton System)**.** We consider the Hamilton system:

$$\partial_t x = -\partial_y H(x,y)\,, \qquad \partial_t y = \partial_x H(x,y)\,, \qquad x,y\in\mathbb{R}^d\,,$$

where for simplicity we consider

$$H(x,y) = \sum_{j=1}^{d} \frac{\nu_j}{2}\left(x_j^2+y_j^2\right) + o\left(|x|^2+|y|^2\right)\,,$$

where $\nu_j \neq k\nu_1$, $k\in\mathbb{Z}$. Then consider the pertubation

$$\begin{aligned}\partial_t x &= \lambda\partial_x H(x,y) - \partial_y H(x,y)\\ \partial_t y &= \lambda\partial_y H(x,y) + \partial_x H(x,y)\end{aligned}$$

From

$$\frac{\mathrm{d}}{\mathrm{d}t}H = \nabla_u H\cdot\frac{\mathrm{d}}{\mathrm{d}t}u = \lambda\left|\nabla_u H\right|^2$$

we conclude that the system is energy conserving iff $\lambda=0$ and for $\lambda\neq0$ the only periodic solution is the trivial one. With $N := \operatorname{diag}(\nu_1,\dots\nu_d)$ the linearization yields in 0

$$\begin{aligned}\partial_t x &= \lambda N x - N y\\ \partial_t y &= \lambda N y + N x\end{aligned}$$

With the $2d\times2d$ matrices $\tilde N = \begin{pmatrix} N & 0\\ 0 & N\end{pmatrix}$ and $\tilde M = \begin{pmatrix} 0 & -N\\ N & 0\end{pmatrix}$ and $\zeta=(x,y)$ we infer

$$\dot\zeta = \lambda\tilde N\zeta + \tilde M\zeta\,.$$

For a suitable choice of coordinates, there are solutions

$$\begin{pmatrix} x_1(s)\\ y_1(s)\end{pmatrix} = \begin{pmatrix}\cos\left(|\nu_1(t)|\right)+o(s)\\ \sin\left(|\nu_1(t)|\right)+o(s)\end{pmatrix}$$

and other coordinates are of order $o(s)$.

3.2 Gradient Flows

3.2.1 Convex Analysis and Monotone Operators

We recall some well known results from convex analysis on a reflexive Banach space $\mathscr{B}$ with dual $\mathscr{B}^*$ and dual pairing $\langle b, b^* \rangle_{\mathscr{B},\mathscr{B}^*}$. A good outline of the standard theory of convex functions and functionals can be found the book by Ekeland and Témam [2] or by Rockafellar [9].

A function $\varphi : \mathscr{B} \to \mathbb{R} \cup \{+\infty\}$ is convex if

$$\forall b_1, b_2 \in \mathscr{B}\,,\ \lambda \in (0,1)\ : \qquad \varphi\left(b_1 + \lambda\left(b_2 - b_1\right)\right) \leq \left(1-\lambda\right)\varphi(b_1) + \lambda\varphi(b_2)\,, \tag{3.4}$$

which in equivalent to the convexity of its epigraph

$$\mathrm{epi}\varphi := \left\{(b,r) \in \mathscr{B}\times\mathbb{R}\ :\ \varphi(b) \leq r\right\}\,.$$

A function $\varphi : \mathscr{B} \to \mathbb{R} \cup \{+\infty\}$ is strictly convex if in (3.4) the strict inequality holds.

Lemma 3.2.1. *A convex function φ is lower semicontinuous (l.s.c.) if and only if* epiφ *is closed.*

Proof. $\Rightarrow$: This is evident by definition of epiφ.

$\Leftarrow$: Since $(b_n, \varphi(b_n)) \in \mathrm{epi}\varphi$ and $b_n \to b$, $\varphi(b_n) \to \phi$ with $(b, \phi) \in \mathrm{epi}\varphi$, it holds $\phi \geq \varphi(b)$. □

Let us observe that subsets of φ are convex. This follows from the fact that for $\varphi(b_1), \varphi(b_2) \leq r$ also $\varphi\left(b_1 + \lambda\left(b_2 - b_1\right)\right)$. This implies that

Lemma 3.2.2. *if a convex φ is l.s.c., then φ remains l.s.c. if $\mathscr{B}$ is equipped with its weak topology.*

Proof. Let $b_n \rightharpoonup b$ such that w.l.o.g. $\varphi(b_n) \to r := \liminf_n \varphi(b_n)$. Then $C_n := \{a \in \mathscr{B}\ :\ \varphi(a) \leq \varphi(b_n)\}$ are closed, convex with $(b_j)_{j\geq n} \subset C_n$ and hence $b \in C_n$ on noting that the weakly closed sets are the closed convex sets. This implies $b \in \bigcap_n C_n$ and hence $\varphi(b) \leq \varphi(b_n)$ for every n. □

Lemma 3.2.3. *A convex function $\varphi : \mathscr{B} \to \mathbb{R}$ is continuous in b if and only if it is bounded in a neighborhood of b.*

Proof. If φ is continuous, it is bounded in a neighborhood b.

If φ is convex and bounded in a neighborhood of b, we reduce to the case $b = 0$ and $\varphi(0) = 0$ by translation. Let V be a neighborhood of 0 such that $\varphi(v) \leq a < +\infty$ for all $v \in V$. Defining $W := V \cap -V$ (a symmetric neighborhood of 0), let us take $\varepsilon \in (0,1)$.Vor every $v \in \varepsilon W$ we have by convexity:

$$\begin{aligned}
\forall \frac{v}{\varepsilon} \in W\ : \qquad & \varphi(v) \leq (1-\varepsilon)\varphi(0) + \varepsilon\varphi\left(\frac{v}{\varepsilon}\right) \leq \varepsilon a\,, \\
\forall -\frac{v}{\varepsilon} \in W\ : \qquad & \varphi(v) \geq (1+\varepsilon)\varphi(0) - \varepsilon\varphi\left(-\frac{v}{\varepsilon}\right) \geq -\varepsilon a\,.
\end{aligned}$$

This implies $|F| \leq \varepsilon a$ in εW. □

The domain of a convex function φ is the subset of $\mathscr{B}$ on which φ is finite, i.e.

$$\mathrm{dom}(\varphi) := \{b \in \mathscr{B} : \varphi(b) < +\infty\} \ .$$

For a convex function φ the Legendre-Fenchel conjugate φ^* is defined through

$$\varphi^* : \mathscr{B}^* \to \mathbb{R} \cup \{+\infty\}, \quad b^* \mapsto \sup_{b \in \mathscr{B}} \{\langle b^*, b\rangle - \varphi(b)\}.$$

Theorem 3.2.4. *For every l.s.c. convex function $\varphi : \mathscr{B} \to (-\infty, +\infty]$ it holds $\varphi = \varphi^{**}$.*

Proof. For every $b^* \in \mathscr{B}^*$ and every $b \in \mathscr{B}$ we have by definition of φ^*:

$$\varphi^*(b^*) \geq \langle b, b^*\rangle - \varphi(b) \quad \Leftrightarrow \quad \varphi(b) \geq \langle b, b^*\rangle - \varphi^*(b^*)$$

and $\varphi \geq \varphi^{**}$. Moreover, an affine function $f_{b^*,\alpha} : b \mapsto \langle b, b^*\rangle - \alpha$ is everywhere below φ iff $\varphi(b) \geq \langle b, b^*\rangle - \alpha$. In that sense, $\alpha = \varphi^*(b^*)$ is the smallest value of α such that the affine function $f_{b^*,\alpha}$ lies below $\varphi(b)$. If we can show

$$\varphi(b) = \sup \{f(b) \ : \ f \text{ is affine function below } \varphi\}$$

the claim follows from

$$\begin{aligned} \varphi^{**}(b) &= \sup_{b^*} \left(\langle b, b^*\rangle - \varphi^*(b^*)\right) \\ &\geq \sup_{b^*} \left(f_{b^*,\alpha}(b)\right) \\ &= \sup \{f(b) \ : \ f \text{ is affine function below } \varphi\} \ . \end{aligned}$$

□

Theorem 3.2.5. *For every l.s.c. convex function $\varphi : \mathscr{B} \to (-\infty, +\infty]$ it holds*

$$\varphi(b) = \sup \{f(b) \ : \ f \ \textit{is affine function below}\ \varphi\} \ .$$

Proof. If $\varphi \equiv +\infty$, this is obvious. On the other hand, assume $\overline{b} \in \mathscr{B}$ and $\overline{a} \in \mathbb{R}$ such that $\overline{a} < \varphi(\overline{b})$. We will show that there exists an affine function f such that $f(\overline{b}) > \overline{a}$ and $f \leq \varphi$. This will imply the theorem.

Since epiφ is convex, we can separate it from $(\overline{b}, \overline{a})$ by a hyperplane

$$\mathcal{H} := \{(b, a) \in \mathscr{B} \times \mathbb{R} \ : \ l(b) + \alpha a = \beta\} \ ,$$

where $l : \mathscr{B} \to \mathbb{R}$ is linear.

Now, if $\varphi(\overline{b}) < +\infty$, we find

$$l(b) + \alpha\varphi(b) > \beta > l(\overline{b}) + \alpha\overline{a}. \tag{3.5}$$

and for $b = \overline{b}$ that

$$\alpha\left(\varphi(\overline{b}) - \overline{a}\right) > 0$$

and hence $\alpha > 0$. This in turn implies that $f(b) := \frac{\beta}{\alpha} - \frac{l(b)}{\alpha}$ (note that the kernel is $\mathcal{H}$) is the solution.

In case $\varphi(\bar{b}) = +\infty$, either $\alpha \neq 0$ and we are in the above case (where $\varphi(\bar{b})$ replaced by $\overline{\varphi} > \overline{a}$) or $\alpha = 0$. If $\alpha = 0$ (where (3.5) makes no sense), we still find $\beta - l(\bar{b}) > 0$ and $\beta - l(b) < 0$ for $b \in \mathrm{dom}\varphi$. Hence, for every affine function $\gamma - m(b) < \varphi(b)$ and $c > 0$ it holds

$$\gamma - m(b) + c\,(\beta - l(b)) < \varphi(b)$$

and for c large enough we obtain $\gamma - m(\bar{b}) + c\left(\beta - l(\bar{b})\right) \geq \overline{a}$. □

The subdifferential $\partial\varphi : \mathrm{dom}(\varphi) \to \mathcal{P}(\mathscr{B}^*)$ (for $\mathcal{P}(\mathscr{B}^*)$ being the power set of $\mathscr{B}^*$) of a convex function is the set

$$\partial\varphi(b) = \left\{ b^* \in \mathscr{B}^* \text{ such that } \varphi(\bar{b}) \geq \varphi(b) + \left\langle (\bar{b} - b), b^* \right\rangle_{\mathscr{B},\mathscr{B}^*} \quad \forall\, \bar{b} \in \mathscr{B} \right\}. \tag{3.6}$$

We note that $\partial\varphi(b) = \emptyset$ is possible, i.e. $\mathrm{dom}(\varphi) \neq \mathrm{dom}(\partial\varphi)$ in general.

Related to the notion of subdifferentials are the monotone operators. A multivalued operator $f : \mathrm{dom}(f) \subset \mathscr{B} \to \mathcal{P}(\mathscr{B}^*)$ is called *monotone* if

$$\langle b_1 - b_2, b_1^* - b_2^* \rangle \geq 0, \quad \forall\, b_i \in \mathrm{dom}(f), \quad \forall\, b_i^* \in f(b_i),\ (i = 1, 2).$$

In what follows, we frequently use the following usefull properties of convex functionals[**?**].

Lemma 3.2.6. *For every*

$$\varphi : \mathscr{B} \to \mathbb{R} \cup \{+\infty\},\ \textit{convex and lower-semicontinuous, with } \varphi \not\equiv +\infty. \tag{3.7}$$

it holds

(i) φ^* *is convex, lower-semicontinuous, and* $dom(\varphi^*) \neq \emptyset$*,*

(ii) $\partial\varphi,\ \partial\varphi^*$ *are monotone operators,*

(iii) $\varphi(b) + \varphi^*(b^*) \geq \langle b, b^* \rangle, \qquad \forall\ (b, b^*) \in \mathscr{B} \times \mathscr{B}^*.$

(iv) $b \in dom(\varphi)$ *and* $b^* \in \partial\varphi(b) \ \Leftrightarrow\ b^* \in dom(\varphi^*)$ *and* $b \in \partial\varphi^*(b^*)$*.*

(v) $b^* \in dom(\varphi^*)$ *and* $b \in \partial\varphi^*(b^*) \quad \Leftrightarrow\ \varphi(b) + \varphi^*(b^*) = \langle b, b^* \rangle$*.*

We refer to (v) as Fenchel's equality and to (iii) as Fenchel's *in*equality.

Proof. (i)

$$\begin{aligned}\varphi^*\left(\lambda b_1^* + (1-\lambda) b_2^*\right) &= \sup_{b \in \mathscr{B}} \left\{ \langle (\lambda b_1^* + (1-\lambda) b_2^*)\,,\, b \rangle - (\lambda + (1-\lambda))\, \varphi(b) \right\} \\ &\leq \lambda \sup_{b \in \mathscr{B}} \left\{ \langle b_1^*, b \rangle - \varphi(b) \right\} + (1-\lambda) \sup_{b \in \mathscr{B}} \left\{ \langle b_2^*,\, b \rangle - \varphi(b) \right\}\,.\end{aligned}$$

In order to verify the l.s.c. property, note that $\mathrm{epi}\varphi^* = \bigcap_{b \in \mathscr{B}} \mathrm{epi}\left(\langle b, \cdot \rangle - \varphi(b)\right)$ and hence is the intersection of closed convex sets. Therefore, $\mathrm{epi}\varphi^*$ is closed and convex and hence φ^* is l.s.c.

By Theorem 3.2.5 there exists a continuous affine function $\langle \cdot, b^* \rangle + \alpha \leq \varphi$ and hence

$$\varphi^*(b^*) \leq \sup_b \{\langle b, b^* \rangle - \langle b, b^* \rangle - \alpha\} = -\alpha\,.$$

(ii) follows from the definition of the subdifferential
(iii) follows from definition of φ^*.
(iv) and (v) If $b \in dom(\varphi)$ and $b^* \in \partial\varphi(b)$ we have for every $\bar{b} \in \mathscr{B}$:

$$\langle b, b^* \rangle - \varphi(b) \geq \left\langle \bar{b}, b^* \right\rangle - \varphi(\bar{b})$$

and hence $\varphi^*(b^*) = \langle b, b^* \rangle - \varphi(b)$, i.e. (v). Furthermore, the above inequality implies for $\tilde{b}^* \in \mathscr{B}$

$$\varphi(\tilde{b}^*) \geq \left\langle b, \tilde{b}^* \right\rangle - \varphi(b) \geq \left\langle \bar{b}, b^* \right\rangle - \varphi(\bar{b}) + \left\langle b, \tilde{b}^* - b^* \right\rangle = \varphi^*(b^*) + \left\langle b, \tilde{b}^* - b^* \right\rangle\,,$$

which implies the claim. the oposite direction follows similarly using $\varphi^{**} = \varphi$ □

3.2.2 Integral Convex Functionals

In this section, we will study functionals of the form

$$I_f(u) := \int_M f\,(m, u(m)))\mathrm{d}\mu(m)\,,$$

where $f : M \times \mathbb{R}^D \to \mathbb{R}$ is measurable on $M \times \mathbb{R}^D$ and convex, lower semi-continuous in the second variable. It is by no means clear the the above functionals are well defined on spaces of integrable functions, though it is clear that *if* they are well defined, they are convex functionals. Functionals of the above type have been studied in detail by Rockafellar in [8] and we will summarize his main findings below.

Definition 3.2.7. Let $(M, \mathscr{F}, \mu)$ be a measure space and let $f : M \times \mathbb{R}^D \to \mathbb{R}$. We call f a convex integrand if f is proper (i.e. $f \not\equiv +\infty$) and for every $m \in M$ the function $x \mapsto f(m, x)$ is convex. We call f a *normal* convex integrand if

1. f is lower semi-continuous in $\mathbb{R}^D$ (i.e. f is a convex integrand)
2. There exists a countable family $(u_i)_{i\in\mathbb{N}}$ of measurable functions from M to $\mathbb{R}^D$ such that
   (a) for every $i \in \mathbb{N}$ the function $m \mapsto f(m, u_i(m))$ is measurable
   (b) for every m:

$$\{u_i(m)\ :\ i \in \mathbb{N}\} \cap \mathrm{dom} f(m, \cdot) \quad \text{is dense in } \mathrm{dom} f(m, \cdot)\,. \tag{3.8}$$

The class of normal convex integrands is not empty, as the following lemma shows.

Lemma 3.2.8. *Let f be a convex integrand such that for every $x \in \mathbb{R}^D$ the function $m \mapsto f(m,x)$ is measurable and such that for every m the function $x \mapsto f(m,x)$ is lower semi-continuous and has interior points in its effective domain* $\mathrm{dom} f(m) := \{x : f(m,x) < \infty\}$. *Then f is a normal convex integrand.*

Proof. The first point of Definition 3.2.7 is satisfied. Now let $(r_i)_{i\in\mathbb{N}} \subset \mathbb{R}^D$ be a countable dense subset and let $u_i(m) \equiv r_i$. Then, $m \mapsto f(m, u_i(m))$ is measurable by hypothesis. Furthermore (3.8) is satisfied by density of $(r_i)_{i\in\mathbb{N}} \subset \mathbb{R}^D$. □

Corollary 3.2.9. *Let f be a convex integrand having only finite values such that for every $x \in \mathbb{R}^D$ the map $m \mapsto f(m,x)$ is measurable. Then f is a normal convex integrand.*

Proof. The function $f(m,\cdot)$ is continuous since $\mathrm{dom} f(m,\cdot) = \mathbb{R}^D$. □

Lemma 3.2.10. *Let f be a normal convex interand and consider $I_f : L^p(M) \to \mathbb{R}$. For every $u \in L^p(M)$ it holds $\varphi \in \partial I_f(u)$ if and only if $\varphi \in L^q(M)$, $\frac{1}{p} + \frac{1}{q} = 1$, and $\varphi(m) \in \partial f(m, u(m))$ for μ-almost every $m \in M$.*

Proof. Let $\varphi \in \partial I_f(u)$. Then

$$\forall v \in L^p(M) : \qquad I_f(v) \geq I_f(u) + \int_M (v-u)\cdot\varphi\,.$$

For every measurable $\tilde{M} \subset M$ let $\tilde{v}(m) := u(m)$ on $M\backslash\tilde{M}$ and $\tilde{v}(m) = v(m)$ on $\tilde{M}$. Hence, we infer from the last inequality

$$\forall \tilde{M} \subset M \qquad \int_{\tilde{M}} \left(f(m,u(m)) - f(m,v(m)) + (v(m)-u(m))\cdot\varphi\right) \leq 0$$

and hence $\varphi(m) \in \partial f(m,u(m))$ for μ-almost every $m \in M$.

Conversely, let $\varphi \in L^q(M)$, $\frac{1}{p} + \frac{1}{q} = 1$, and $\varphi(m) \in \partial f(m,u(m))$ for μ-almost every $m \in M$. Since we have

$$\text{for almost every } m \in M : \qquad f(m,u(m)) - f(m,v(m)) + (v(m)-u(m))\cdot\varphi\,,$$

it follows that $\varphi \in \partial I_f(u)$. □

Lemma 3.2.11. *If f is a normal convex integrand, then f^* is a normal convex integrand, too. Furthermore, $f(m,u(m))$ is measurable for every measurable function $u : M \to \mathbb{R}^D$.*

Proof. By definition, $x \mapsto f(m,x)$ is convex l.s.c. (Lemma 3.2.6) for a.e. $m \in M$. Since for every x the function $m \mapsto f(m,x)$ is measurable, we find by definition of the Fenchel conjugate

$$f^*(m,u^*(m)) = \sup_x \langle u^*(m), x\rangle - f(m,x)$$

is the pointwise supremum of measurable functions, hence measurable. Furthermore, $f^*(m,\cdot)$ is by definition convex l.s.c.. Hence it remains to verify 2.(b) of Definition 3.2.7 for f^*. Note at this point that by $f^{**} = f$ and the above observations we obtain measurability of $f(m,u(m))$.

The verification of 2.(b) of Definition 3.2.7 can be found in Rockafellar [8]. □

We are now able to state the well-definitness of the involved integrals:

Theorem 3.2.12. *Let $1 < p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ and let f be a normal convex integrand. Suppose there exists at least one $u^* \in L^q(M)$ such that $f^*(m, u^*(m))$ is integrable. Then*

$$I_f : \; L^p(M) \to \mathbb{R}\,, \qquad I_f(u) := \int_M f(m, u(m)) \mathrm{d}\mu(m)\,,$$

is a well-defined convex functional on $L^p(M)$ with values in $(-\infty, +\infty]$.

Proof. Lemma 3.2.11 ensures the measurability of $m \mapsto f(m, u(m))$. Furthermore, for $u^* \in L^q(M)$ from the hypothesis, it follows for a.e. $m \in M$ that

$$f(m, u(m)) \geq u(m) \cdot u^*(m) - f^*(m, u^*(m)) > -\infty$$

and hence $I_f(u) > -\infty$ and measurability implies that I_f is well defined. The convexity of I_f follows immediately from the convexity of f through

$$\begin{aligned} I_f\left(u_1 + \lambda\left(u_2 - u_1\right)\right) &= \int_M f\left(\cdot\,,\, u_1 + \lambda\left(u_2 - u_1\right)\right) \\ &\leq \int_M \left(1 - \lambda\right) f(\cdot\,, u_1) + \lambda f(\cdot\,, u_2) \\ &= \left(1 - \lambda\right) I_f(\cdot\,, u_1) + \lambda I_f(\cdot\,, u_2)\,. \end{aligned}$$

□

In what follows, we will show that I_f and I_{f^*} are conjugate.

Theorem 3.2.13. *Let $1 < p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ and let f be a normal convex integrand such that $m \mapsto f(m, u(m))$ is integrable in m for at least one $u \in L^p(M)$ and $f^*(m, u^*(m))$ is summable for at least one $u^* \in L^q(M)$. Then I_f and I_{f^*} are proper convex functionals conjugate to each other.*

Proof. The general proof is to lengthy for our purpose, hence we restrict to the case $f(m, u) = f(u)$ is continuously differentiable. In particular, the l.s.c. of f implies that $u \mapsto \int_M f(u)$ is strongly *l.s.c.* as well as weakly *l.s.c.*. Hence

$$\left(I_f\right)^*\left(u^*\right) = \sup_{u \in \mathscr{B}} \int_M u\, u^* - \int_M f(u)\,.$$

Let u_n be a minimizing sequence of the right hand side. W.l.o.g. we can assume $u_n \rightharpoonup u$ and hence $\int_M f(u) \leq \liminf_n \int_M f(u_n)$ which implies

$$\left(I_f\right)^*\left(u^*\right) \leq \int_M u\, u^* - \int_M f(u) \leq \int_M f^*(u^*)\,.$$

Let $\beta < I_{f^*}(u^*)$ and let $\alpha(m)$ be a function with $\alpha(m) \leq f^*(u^*(m))$ and $\int_M \alpha > \beta$. Then

$$-\alpha(m) \geq \inf_x \left(f(x) - x \cdot u^*(m)\right)\,.$$

Since ∂f is continuous an monotone, we can provide a measurable function $x(m)$ which minimizes the above expression. In particular, we infer

$$\alpha(m) \leq f(x(m)) - x(m) \cdot u^*(m)$$

and hence $I_{f^*}(u^*) \geq \left(I_f\right)^*(u^*)$. □

3.2.3 Existence Theory for a Class of Gradient Flows

The homogenization of convex functionals opens the door to the homogenization of a hughe class of dynamic problems (note that the homogenization of convex functionals was a purely stationary problem). A deep introduction into the field of gradient flows is given in [1]. We consider the following equation in a reflexive Banach space $\mathscr{B}$:

$$\dot{u} \in \partial\Psi^* \left(-\mathrm{D}\mathscr{E}(t,u)\right) . \tag{3.9}$$

where we impose the following assumptions:

$$\mathscr{E}(t,u) = \mathcal{E}(u) - \langle f(t), u\rangle \ ,$$

where $f \in W^{1,q}(0,T;\mathscr{B})$. In particular, we solve the equation

$$\dot{u} \in \partial\Psi^* \left(-\mathrm{D}\mathcal{E}(t,u) + f\right) . \tag{3.10}$$

The notion "gradient flow" comes from the following observation: Consider $\mathcal{E} : \mathbb{R} \to \mathbb{R}$ a lower semi-continuous functional with $\mathcal{E}(u) \to \infty$ as $|u| \to \infty$ and the equation

$$\dot{u} = -\nabla\mathcal{E}(u) .$$

This is equivalent with the system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u(t)) = -\left|\nabla\mathcal{E}(u(t))\right|\left|\dot{u}(t)\right| , \qquad \left|\nabla\mathcal{E}(u(t))\right| = \left|\dot{u}(t)\right| ,$$

or according to Lemma 3.2.6

$$\Psi^*\left(-\mathrm{D}\mathcal{E}(u)\right) + \Psi\left(\dot{u}\right) = -\left\langle \mathrm{D}\mathcal{E}(u)\,,\,\dot{u}\right\rangle ,$$

with $\Psi^*(u) = \frac{1}{2}u^2$. In the late 90's it was observed that this structure can be made use of in much more general context.

Example 3.2.14 (Diffusion equation)**.** Consider $\mathcal{E}(u) = \int_\Omega \frac{1}{2}\left|\nabla u\right|^2 - \int_\Omega fu$ with $\Psi^*(\xi) = \int_\Omega \frac{1}{2}\xi^2$. Then $\mathrm{D}\mathcal{E}(u) = -\Delta u - f$ and the above formalism leads to

$$\dot{u} = \Delta u + f .$$

On the other hand, consider $\mathcal{E}(u) = \int_\Omega \frac{1}{2}\left|u\right|^2 - \int_\Omega u\left(-\Delta\right)^{-1} f$ with $\Psi^*(\xi) = \int_\Omega \frac{1}{2}\left|\nabla\xi\right|^2$. This leads to the same result.

Both approaches are not physical. The driving physical quantity is the entropy $\mathcal{E}(u) = \int_\Omega u\ln u$. The corresponding Ψ^* is $\Psi^*_u(\xi) = \int_\Omega \frac{1}{2}u\left|\nabla\xi\right|^2$. This is the Jordan-Kinderlehrer-Otto formalism of the Diffusion equation in the Wasserstein metric space

$$\dot{u} \in \partial\Psi^*_u\left(-\mathrm{D}\mathcal{E}(u)\right) .$$

Assumption 3.2.15. *The functional* $\mathcal{E} : [0,T]\times\mathscr{B} \to \mathbb{R}$ *and* $\Psi : \mathscr{B}^* \to \mathbb{R}$ *satisfy*

1. *$\mathcal{E}$ is a strongly l.s.c. function with compact sublevels, that is*

$$F \subset \mathscr{B} \ \sup_{u\in F} \mathcal{E}(u) < \infty \qquad \Rightarrow \qquad F \text{ is precompact,}$$

   and such that $u_n \to u$, $\mathrm{D}\mathcal{E}(u_n) \rightharpoonup \xi$ and $\sup_n \mathcal{E}(u_n) < \infty$ implies

$$\mathcal{E}(u_n) \to \mathcal{E}(u)\,, \qquad \xi = \mathrm{D}\mathcal{E}(u)\,. \tag{3.11}$$

2. *Ψ is l.s.c., p-homogeneous (i.e. $\Psi(\alpha b) = |\alpha|^p\Psi(b)$) and convex, $1 < p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$.*

The particular property of p-homogeneous functions which we will use here is the following:

Lemma 3.2.16. *If Ψ is l.s.c., p-homogeneous (i.e. $\Psi(\alpha b) = |\alpha|^p\Psi(b)$) and convex, $1 < p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ then Ψ^* is q-homogeneous.*

Proof.

$$\begin{aligned}\Psi^*(\alpha b^*) &= \sup_{b\in\mathscr{B}} \langle b, \alpha b^*\rangle - \Psi(b) = \sup_{b\in\mathscr{B}} \left\langle \frac{|\alpha|^q}{\alpha} b, \alpha b^* \right\rangle - \Psi\left(\frac{|\alpha|^q}{\alpha} b\right)\\ &= \sup_{b\in\mathscr{B}} |\alpha|^q \langle b, b^*\rangle - |\alpha|^{(q-1)p}\,\Psi(b) = |\alpha|^q\,\Psi^*(b^*)\ .\end{aligned}$$

□

From Lemma 3.2.6

$$\Psi^*\left(-\mathrm{D}\mathscr{E}(t,u)\right) + \Psi(\dot{u}) \geq -\left\langle \mathrm{D}\mathscr{E}(t,u)\,,\,\dot{u}\right\rangle\ .$$

we infer that (3.9) holds if and only if

$$\Psi^*\left(-\mathrm{D}\mathscr{E}(t,u)\right) + \Psi(\dot{u}) = -\left\langle \mathrm{D}\mathscr{E}(t,u)\,,\,\dot{u}\right\rangle\ .$$

Integrating the last equality over time yields

$$\mathcal{E}(T,u(T)) + \int_0^T \left(\Psi^*\left(-\mathrm{D}\mathcal{E}(s,u(s))\right) + \Psi\left(\dot{u}(s)\right)\right)\mathrm{d}s = \mathcal{E}(0,u_0) + \int_0^T \partial_t\mathcal{E}(s,u(s))\mathrm{d}s\,,$$

which is equivalent to

$$\mathcal{E}(T,u(T)) + \int_0^T \left(\Psi^*\left(-\mathrm{D}\mathcal{E}(\cdot,u) + f\right) + \Psi(\dot{u})\right) = \mathcal{E}(0,u_0) + \int_0^T \left\langle \dot{f}, u\right\rangle\ .$$

where $u(0) = u_0$. A very general existence theory is outlined in [10] and for the case of quadratic Ψ^*, the most general theory is outlined in [11]. Here, we will consider the following simplified setting.

By the above considerations, it is trivial to see that

$$\mathcal{E}(T,u(T)) + \int_0^T \left(\Psi^*\left(-\mathrm{D}\mathcal{E}(\cdot,u) + f\right) + \Psi(\dot{u})\right) \geq \mathcal{E}(0,u_0) + \int_0^T \left\langle \dot{f}, u\right\rangle$$

always holds and the tough part is to prove the inverse inequality.

Theorem 3.2.17. *Under the above assumptions and the additional assumption that there exists $c>0$ with $c\left\|\xi^*\right\|_{\mathscr{B}^*}^q \leq \Psi^*(\xi^*)$ and $c\left\|\xi\right\|_{\mathscr{B}}^p \leq \Psi(\xi)$ for all $(\xi,\xi^*)\in\mathscr{B}\times\mathscr{B}^*$ the gradient flow inequality*

$$\mathscr{E}(T,u(T))+\int_0^T\left(\Psi^*\left(-\mathrm{D}\mathcal{E}(\cdot,u)+f\right)+\Psi\left(\dot{u}\right)\right)\leq\mathscr{E}(0,u_0)+\int_0^T\left\langle\dot{f},u\right\rangle \tag{3.12}$$

has at least one solution $u\in L^\infty(0,T;\mathscr{B})$, $\dot{u}\in L^p(0,T;\mathscr{B})$ and $\mathrm{D}\mathcal{E}(\cdot,u)\in L^q(0,T;\mathscr{B}^*)$ with $u(0)=u_0$.

Remark 3.2.18. a) Note that the oposite inequality in (3.12) holds for all sufficiently regular functions u due to Lemma 3.2.6.

b) Theorem 3.2.17 also implies that $u(t)\in\operatorname{dom}D\mathscr{E}(t,\cdot)$ for almost every t.

We introduce a modified Moreau-Yosida approximation defining

$$\begin{cases}\mathcal{E}^*(\tau,\mathfrak{f},\tilde{u};u):=\tau\Psi\left(\dfrac{u-\tilde{u}}{\tau}\right)+\mathcal{E}(u)-\langle u,\mathfrak{f}\rangle\ ,\\ \qquad\mathcal{E}_\sigma(\mathfrak{f},\tilde{u}):=\inf\limits_{u\in\mathscr{B}}\mathcal{E}^*(\sigma,\mathfrak{f},\tilde{u};u) \qquad\qquad \sigma>0\end{cases}$$

and note that there exists at least one minimizer in case $\tilde{u}\in\operatorname{dom}\mathcal{E}$. In particular, the following set is not empty

$$J_\sigma(\mathfrak{f},\tilde{u}):=\operatorname{argmin}_{u\in\mathscr{B}}\mathcal{E}^*(\sigma,\mathfrak{f},\tilde{u};u)\,. \tag{3.13}$$

We note that $u(t)\in J_\sigma(\mathfrak{f},\tilde{u})$ satisfies

$$\partial\Psi\left(\frac{u-\tilde{u}}{\tau}\right)+\mathrm{D}\mathcal{E}(u)-\mathfrak{f}=0 \tag{3.14}$$

Lemma 3.2.19. *Under the above assumptions, the map $\sigma\mapsto\mathcal{E}_\sigma(\tilde{u};\mathfrak{f})$ is locally Lipschitz for every $\tilde{u}\in\operatorname{dom}(\mathcal{E})$, and*

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathcal{E}_\sigma\left(\tilde{u};\mathfrak{f}\right)=-(p-1)\Psi\left(\frac{u_\sigma-\tilde{u}}{\sigma}\right)\qquad\forall\sigma\in(0,\tau_*)\backslash\mathcal{N}_{\tilde{u}}\,. \tag{3.15}$$

In particular, for all σ_0 and $u_{\sigma_0}\in J_{\sigma_0}\left(\mathfrak{f},\tilde{u}\right)$ we have

$$\int_0^{\sigma_0}\Psi\left(\frac{u_{\sigma_0}-\tilde{u}}{\sigma_0}\right)+\int_0^{\sigma_0}\Psi^*\left(-\mathrm{D}\mathcal{E}(u_\sigma)+\mathfrak{f}\right)d\sigma\leq\mathcal{E}(\tilde{u})-\mathcal{E}(u_{\sigma_0})-\left\langle\tilde{u}-u_{\sigma_0},\mathfrak{f}\right\rangle\,. \tag{3.16}$$

Proof. We first observe that $\mathcal{E}_\sigma(\tilde{u};\mathfrak{f})\leq\mathcal{E}(\tilde{u})-\langle\tilde{u},\mathfrak{f}\rangle$. Hence, $\bigcup_{\sigma\in\mathbb{R}}J_\sigma(\mathfrak{f},\tilde{u})$ is bounded and also $\sup_\sigma\sigma\Psi\left(\frac{u_\sigma-\tilde{u}}{\sigma}\right)<\infty$. First we have

$$\begin{aligned}\mathcal{E}_{\sigma_1}\left(\tilde{u};\mathfrak{f}\right)-\mathcal{E}_{\sigma_2}\left(\tilde{u};\mathfrak{f}\right)&\leq\mathcal{E}(u_{\sigma_2})-\left\langle u_{\sigma_2},f\right\rangle+\sigma_1^{1-p}\Psi\left(u_{\sigma_2}-\tilde{u}\right)-\mathcal{E}(u_{\sigma_2})+\left\langle u_{\sigma_2},f\right\rangle-\sigma_2^{1-p}\Psi\left(u_{\sigma_2}-\tilde{u}\right)\\&\leq\frac{\left(\sigma_2^{p-1}-\sigma_1^{p-1}\right)\Psi\left(u_{\sigma_2}-\tilde{u}\right)}{\sigma_2^{p-1}\sigma_1^{p-1}}\end{aligned}$$

for every $\sigma_1, \sigma_2 \in (0,1)$ and $u_{\sigma_i} \in J_{\sigma_i}(\tilde{u}, \mathfrak{f})$. Interchanging the sign of the inequality and σ_1 and σ_2 we obtain

$$\mathcal{E}_{\sigma_1}(\tilde{u};\mathfrak{f}) - \mathcal{E}_{\sigma_2}(\tilde{u};\mathfrak{f}) \geq \frac{\left(\sigma_2^{p-1} - \sigma_1^{p-1}\right)\Psi\left(u_{\sigma_1} - \tilde{u}\right)}{\sigma_2^{p-1}\sigma_1^{p-1}}\,.$$

we devide both relations by $\sigma_2 - \sigma_1$ and obtain

$$\frac{\left(\sigma_2^{p-1} - \sigma_1^{p-1}\right)\Psi\left(u_{\sigma_2} - \tilde{u}\right)}{\sigma_2^{p-1}\sigma_1^{p-1}\left(\sigma_2 - \sigma_1\right)} \geq \frac{\mathcal{E}_{\sigma_1}(\tilde{u};\mathfrak{f}) - \mathcal{E}_{\sigma_2}(\tilde{u};\mathfrak{f})}{(\sigma_2 - \sigma_1)} \geq \frac{\left(\sigma_2^{p-1} - \sigma_1^{p-1}\right)\Psi\left(u_{\sigma_1} - \tilde{u}\right)}{\left(\sigma_2 - \sigma_1\right)\sigma_2^{p-1}\sigma_1^{p-1}}\,.$$

The last inequality implies that $\mathcal{E}_{\sigma_1}(\tilde{u};\mathfrak{f})$ is Lipschitz continuous because

$$\lim_{\sigma_1\to\sigma_2} \frac{\left(\sigma_2^{p-1} - \sigma_1^{p-1}\right)\Psi\left(u_{\sigma_2} - \tilde{u}\right)}{(\sigma_2 - \sigma_1)\sigma_2^{p-1}\sigma_1^{p-1}} = \frac{(p-1)\sigma_2^{p-2}\Psi\left(u_{\sigma_2} - \tilde{u}\right)}{\sigma_2^{2p-2}} = (p-1)\Psi\left(\frac{u_{\sigma_2} - \tilde{u}}{\sigma_2}\right)$$

and hence the left hand side is bounded. It remains to show that for almost every σ_2 it holds

$$\sigma_1 \to \sigma_2 \quad \Rightarrow \quad \frac{\left(\sigma_2^{p-1} - \sigma_1^{p-1}\right)\Psi\left(u_{\sigma_1} - \tilde{u}\right)}{\left(\sigma_2 - \sigma_1\right)\sigma_2^{p-1}\sigma_1^{p-1}} \to (p-1)\Psi\left(\frac{u_{\sigma_2} - \tilde{u}}{\sigma_2}\right)$$

Since $\sigma \to \Psi(u_\sigma - \tilde{u})$ is bounded, it is sufficient to show that $\sigma \to \Psi(u_\sigma - \tilde{u})$ is also monotone, which implies it is almost everywhere continuous. But monotonicity follows from the definition of u_σ:

$$\begin{aligned}
\sigma_1\Psi\left(\frac{u_{\sigma_1} - \tilde{u}}{\sigma_1}\right) + \mathcal{E}(u_{\sigma_1}) - \langle u_{\sigma_1}, \mathfrak{f}\rangle &\leq \sigma_1^{1-p}\Psi\left(u_{\sigma_2} - \tilde{u}\right) + \mathcal{E}(u_{\sigma_2}) - \langle u_{\sigma_2}, \mathfrak{f}\rangle \\
&\leq \left(\sigma_1^{1-p} - \sigma_2^{1-p}\right)\Psi\left(u_{\sigma_2} - \tilde{u}\right) + \sigma_2^1\Psi\left(\frac{u_{\sigma_2} - \tilde{u}}{\sigma_2}\right) + \mathcal{E}(u_{\sigma_2}) - \langle u_{\sigma_2}, \mathfrak{f}\rangle \\
&\leq \left(\sigma_1^{1-p} - \sigma_2^{1-p}\right)\Psi\left(u_{\sigma_2} - \tilde{u}\right) + \sigma_2^1\Psi\left(\frac{u_{\sigma_1} - \tilde{u}}{\sigma_2}\right) + \mathcal{E}(u_{\sigma_1}) - \langle u_{\sigma_1}, \mathfrak{f}\rangle
\end{aligned}$$

and hence

$$\left(\sigma_1^{1-p} - \sigma_2^{1-p}\right)\Psi\left(u_{\sigma_1} - \tilde{u}\right) \leq \left(\sigma_1^{1-p} - \sigma_2^{1-p}\right)\Psi\left(u_{\sigma_2} - \tilde{u}\right)\,,$$

which implies $\Psi(u_{\sigma_1} - \tilde{u}) \leq \Psi(u_{\sigma_2} - \tilde{u})$ in case $\sigma_2 > \sigma_1$.

In order to proceed note that

$$\partial_\alpha\left(\Psi^*(\alpha u)\right) = \langle u,\, \partial\Psi^*(\alpha u)\rangle = \partial_\alpha\left(\alpha^q\Psi^*(u)\right) = q\alpha^{q-1}\Psi^*(u)\,,$$

and for $\alpha = 1$ we conclude $\langle u,\, \partial\Psi^*(u)\rangle = p\Psi^*(u)$. Furthermore, we infer from (3.14) and Lemma 3.2.6 that

$$(p-1)\Psi\left(\frac{u_\sigma - \tilde{u}}{\sigma}\right) = (p-1)\left\langle -\mathrm{D}\mathcal{E}(u_\sigma) + \mathfrak{f},\, \frac{u_\sigma - \tilde{u}}{\sigma}\right\rangle - (p-1)\Psi^*\left(-\mathrm{D}\mathcal{E}(u_\sigma) + \mathfrak{f}\right)$$

and becase $\frac{u_\sigma - \tilde{u}}{\sigma} \in \partial\Psi^*\left(-\mathrm{D}\mathcal{E}(u_\sigma) + \mathfrak{f}\right)$ we find

$$(p-1)\Psi\left(\frac{u_\sigma - \tilde{u}}{\sigma}\right) = (p-1)\left(q-1\right)\Psi^*\left(-\mathrm{D}\mathcal{E}(u_\sigma) + \mathfrak{f}\right) = \Psi^*\left(-\mathrm{D}\mathcal{E}(u_\sigma) + \mathfrak{f}\right)\,.$$

Integrating (3.15) over σ we deduce

$$\int_0^{\sigma_0} \Psi\left(\frac{u_{\sigma_0}-\tilde{u}}{\sigma_0}\right) + \int_0^{\sigma_0} \Psi^*\left(-\mathrm{D}\mathcal{E}(u_\sigma)+\mathfrak{f}\right) d\sigma = \lim_{\sigma\to 0}\mathcal{E}_\sigma(\tilde{u}) - \mathcal{E}(u_{\sigma_0}) + \langle u_{\sigma_0},\mathfrak{f}\rangle_* \,.$$

Finally, note that $\mathcal{E}(u_\sigma) - \langle u_\sigma,\mathfrak{f}\rangle \leq \mathcal{E}_\sigma(\tilde{u}) \leq \mathcal{E}(\tilde{u}) - \langle \tilde{u},\mathfrak{f}\rangle$ which implies the claim. □

For fixed $0 < T < \infty$ and time step $0 < \tau < T$, there corresponds a partition of $(0,T)$ as

$$t_0 := 0 < t_1 < \cdots < t_j < \cdots < \cdots < t_{N-1} < T \leq t_N, \qquad t_j := j\tau, \quad N \in \mathbb{N}\,.$$

We set

$$\overline{F}_\tau(t) := F_\tau^j = \frac{1}{\tau}\int_{t_{j-1}}^{t_j} f(s)ds \quad \text{for } t\in(t_{j-i},t_j], \quad j = 1,\dots,N$$

and note that

$$\begin{aligned}\tau \left\|\overline{F}_\tau(t)\right\|_{\mathscr{B}^*}^q &\leq \|f\|^q_{L^q(t_{j-1},t_j;\mathscr{B}^*)} \qquad &&\forall t\in(t_{j-1},t_j)\\ \left\|\overline{F}_\tau\right\|^q_{L^q(t_m,t_n;\mathscr{B}^*)} &\leq \|f\|^q_{L^q(t_m,t_n;\mathscr{B}^*)} \qquad &&\forall 1\leq m<n\leq N\\ \overline{F}_\tau \to f \quad &\text{as } \tau\to 0 \text{ strongly in } L^q(0,T;\mathscr{B}^*)\,.\end{aligned}$$

Let $u_\tau^0 := u_0$ for all τ. For $j = 1,\dots,N$ we let

$$u_\tau^j \in J_\tau(F_\tau^j, u_\tau^{j-1})\,. \tag{3.17}$$

We denote the piecewise linear interpolant u_τ and define the constant interpolant $\overline{u}_\tau$ by$\overline{u}_\tau(t) := u_\tau^j$, $t\in(t_{j-1},t_j]$, as well as the De Giorgi variational interpolant $\tilde{u}_\tau$ through $\tilde{u}_\tau(0) = u_\tau^0$ and

$$\tilde{u}_\tau(t) \in J_\sigma(F_\tau^j, u_\tau^{j-1}) \qquad \text{for } t = t_{j-1}+\sigma \in (t_{j-1},t_j]\,. \tag{3.18}$$

We note that $\tilde{u}_\tau(t)$ satisfies

$$\partial\Psi\left(\frac{\tilde{u}_\tau(t)-u_\tau^{j-1}}{t-t_\tau^{j-1}}\right) + \mathrm{D}\mathcal{E}(\tilde{u}_\tau(t)) - F_\tau^j = 0 \tag{3.19}$$

The tool which fill finally allow us to pass to the limit is the following.

Theorem 3.2.20 (Simon's Theorem [13])**.** *Let $\mathscr{B}$ be a Banach space and let $F\subset L^p(0,T;\mathscr{B})$. F is relatively compact in $L^p(0,T;\mathscr{B})$ for $1\leq p<\infty$, or in $C(0,T;\mathscr{B})$ for $p=\infty$ if and only if*

$$\forall 0<t_1<t_2<T\ :\quad \left\{\int_{t_1}^{t_2} f(t)\mathrm{d}t\ :\ f\in F\right\} \text{ is relatively compact in } \mathscr{B}\,, \tag{3.20}$$

$$\sup_{f\in F}\|f(\cdot+h)-f\|_{L^p(0,T;\mathscr{B})} \to 0 \quad \text{as } h\to 0\,. \tag{3.21}$$

The proof of Simon's Theorem is based on the Ascoli characterization of compact sets in $C([0,T];\mathscr{B})$.

Lemma 3.2.21 (Ascoli compactness criterium)**.** *Let $\mathscr{B}$ be a Banach space and let $F \subset C([0,T];\mathscr{B})$. F is relatively compact if and only if it is pointwise compact and uniformly equicontinuous that is*

$$\forall t \in [0,T] : \quad F(t) := \{f(t) \,:\, f \in F\} \text{ is relatively compact in } \mathscr{B}\,, \tag{3.22}$$

$$\forall \varepsilon > 0 \exists \eta > 0 : \qquad |t_1 - t_2| < \eta \quad \Rightarrow \quad \forall f \in F \,:\, \|f(t_1) - f(t_2)\|_{\mathscr{B}} \leq \varepsilon\,. \tag{3.23}$$

The proof of this lemma is very similar to the finite dimensional case and one only needs to replace the Heine-Borel property with (3.22).

Proof of Theorem 3.2.20. Assume that F is relatively compact. Then $f \mapsto \int_{t_1}^{t_2} f(t)\mathrm{d}t$ is continuous and hence (3.20).

Since F is compact and $C([0,T];\mathscr{B})$ is dense in $L^p(0,T;\mathscr{B})$, for every $\varepsilon > 0$ exist finitely many $(f_i)_{i=1\dots n} \subset C([0,T];\mathscr{B})$ such that

$$F \subset \bigcup_i \mathbb{B}_{\frac{\varepsilon}{3}}(f_i)\,.$$

Furthermore, there exists $h_0 > 0$ such that for all $h < h_0$ and all i it holds

$$\|f_i(\cdot + h) - f_i(\cdot)\|_{L^p(0,T-h;\mathscr{B})} < \frac{\varepsilon}{3}\,.$$

From this we conclude (3.21) using the classical ansatz

$$f(\cdot + h) - f(\cdot) = f(\cdot + h) - f_i(\cdot + h) + f_i(\cdot + h) - f_i(\cdot) + f_i(\cdot) - f(\cdot)\,.$$

Now assume that Fsatisfies (3.20)–(3.21). We only treat the case $1 \leq p < \infty$, the other case follows analogously in $C([0,T];\mathscr{B})$. We prove compactness of F in three steps.

Step 1: For every $f \in F$ and $a > 0$ we define the right mean function $M_a f(t) := \frac{1}{a}\int_t^{t+a} f(s)\mathrm{d}s$. Then $M_a f \in C([0,T-a];\mathscr{B})$ and for every $t_1, t_2 \in [0,T-a]$ one has

$$\|M_a f(t_1) - M_a f(t_2)\|_{\mathscr{B}} = \left\| \frac{1}{a}\int_{t_1}^{t_1+a} (f(\cdot - t_1 + t_2) - f(\cdot)) \right\|_{\mathscr{B}} \leq \frac{1}{a}\|f(\cdot - t_1 + t_2) - f(\cdot)\|_{L^1(0,T;\mathscr{B})}\,.$$

Then (3.21) implies that $M_a F$ is uniformly equiconinuous. Furthermore, (3.20) applied to $t_1 = t$, $t_2 = t + a$ implies that $M_a F(t)$ is relatively compact in $\mathscr{B}$. From the Ascoli-Lemma, we obtain that $M_a F$ is relatively compact in $C([0,T-a];\mathscr{B})$. By that, $M_a F$ is relatively compact in $L^p(0,T;\mathscr{B})$.

Step 2: It holds

$$M_a f - f = \frac{1}{a}\int_0^a (f(\cdot + s) - f(\cdot))\,\mathrm{d}s$$

and hence

$$\|M_a f - f\|_{L^p(0,T-a;\mathscr{B})} \leq \sup_{0 \leq h \leq a} \|f(\cdot + h) - f\|_{L^p(0,T-a;\mathscr{B})}\,.$$

From (3.21) we conclude that for every $\varepsilon > 0$ there exists a such that $h < a$ implies

$$\forall f \in F : \qquad \|f(\cdot + h) - f\|_{L^p(0,T;\mathscr{B})} < \frac{\varepsilon}{3}\,.$$

Since M_aF is relatively compact in $L^p(0,T-a;\mathscr{B})$, there exist finitely many balls $\mathbb{B}_{\frac{\varepsilon}{3}}(f_i) \subset L^p(0,T-a;\mathscr{B})$ such that $M_aF \leq \bigcup_i \mathbb{B}_{\frac{\varepsilon}{3}}(f_i)$ and with the last two inequalities this implies $F \subset \bigcup_i \mathbb{B}_\varepsilon(f_i)$ is compact in $L^p(0,T-a;\mathscr{B})$.

Step 3: By change of direction of time, we infer F is compact in $L^p(a,T;\mathscr{B})$. Together, this implies compactnes of F. □

Finally, we will need Helly's theorem on convergence of monotone functions.

Theorem 3.2.22 (Helly's selection principle)**.** *Let $g_n : [0,T] \to \mathbb{R}$ be a sequence of non-increasing functions, such that $\sup_n \|g_n\|_\infty < \infty$. Then there exists a non-increasing function $g : [0,T] \to \mathbb{R}$ such that along a subsequence $g_n(t) \to g(t)$ for a.e. $t \in [0,T]$.*

Proof. (Short version) Let $R = (r_k)_{k\in\mathbb{N}}$ be an enumeration of $\mathbb{Q} \cap [0,T]$. By the standard Cantor argument, we find a subsequence such that $g_n(r_k) \to g_k$ for every $k \in \mathbb{N}$ and we define $g(r_k) := g_k$. Since $r_k \mapsto g(r_k)$ is non-increasing (monotonicity is preserved in the limit), the limit $g(t+) := \lim_{r_k \searrow t} g(r_k)$ and $g(t-) := \lim_{r_k \nearrow t} g(r_k)$ are well defined and non-increasing. This implies that $t \mapsto g(t-)$ is continuous except for a countable subset of $[0,T]$.

In such continuity points, let $\varepsilon > 0$ be given and let $r_k < t < r_l$ such that $|g(r_k) - g(r_l)| < \frac{\varepsilon}{2}$ and let $n \in \mathbb{N}$ such that $|g_n(r_k) - g(r_k)| + |g_n(r_l) - g(r_l)| < \frac{\varepsilon}{2}$. Then $|g_n(t) - g(t)| < \varepsilon$. □

We are now in the position to prove Theorem 3.2.17.

Proof of Theorem 3.2.17. We add up (3.16) from $t = 0$ to $t = t_J$ and obtain for all $j > 0$ that

$$\mathcal{E}(u_\tau^J) - \left\langle F_\tau^J, u_\tau^J \right\rangle + \sum_{j=1}^J \tau\Psi\left(\frac{u_\tau^j - u_\tau^{j-1}}{\tau}\right) + \int_0^{t_J} \Psi^*\left(-\mathrm{D}\mathcal{E}(\tilde{u}_\tau(t)) + F_\tau(t)\right) dt$$
$$= \mathcal{E}(u_0) - \langle f(0), u_0 \rangle - \sum_{j=1}^J \left\langle F_\tau^j - F_\tau^{j-1}, u_\tau^j \right\rangle . \quad (3.24)$$

Hence, since $f \in W^{1,q}(0,T;\mathscr{B}^*)$, we obtain from the last equation and the estimates on Ψ^* resp. Ψ that

$$\|\dot{u}_\tau\|_{L^p((0,t^J);\mathscr{B})} + \|u_\tau\|_{L^\infty(0,t^J;\mathscr{B})} + \|\mathrm{D}\mathcal{E}(\tilde{u}_\tau(t))\|_{L^q((0,t^J);\mathscr{B}^*)} \leq C\,,$$

and with help of Simon's Theorem we obtain the existence of

$$u \in W^{1,p}(0,T;\mathscr{B}) \cap L^\infty(0,T;\mathrm{dom}\mathcal{E})$$

such that as $\tau \to 0$ we have for a subsequence $u_\tau \rightharpoonup^* u$ weakly$*$ in the respective space and such that

$$u_\tau \to u \quad \text{strongly in } L^p((0,T);\mathscr{B})\,.$$

and for almost every $t \in [0,T]$

$$u_\tau(t) \to u(t) \quad \text{strongly in } \mathscr{B}\,. \quad (3.25)$$

On the other hand, we first note that

$$\left\|u_\tau(t)-u_\tau^{j-1}\right\|_{\mathscr{B}}\leq\int_{t_{j-1}}^{t_j}\|\dot u_\tau(s)\|_{\mathscr{B}}\,\mathrm{d}s\leq\tau^{1-\frac1p}\|\dot u_\tau\|_{L^p((0,t^J);\mathscr{B})}\,,$$

while on the other hand (3.24) yields

$$\Psi\left(\tilde u_\tau(t)-u_\tau^{j-1}\right)\leq\tau^{q-1}\left(\mathcal{E}(u_0)-\langle f(0),u_0\rangle-\sum_{j=1}^J\left\langle F_\tau^j-F_\tau^{j-1},u_\tau^j\right\rangle+\left\langle F_\tau^J,u_\tau^J\right\rangle\right)$$

Hence we find for almost every $t\in[0,T]$ that

$$\tilde u_\tau(t)\to u(t)\text{ and }\overline{u}_\tau(t)\to u(t)\quad\text{strongly in }\mathscr{B}\text{ as }\tau\to0\,.\tag{3.26}$$

We define the sequence of functions e_τ and ζ_τ through

$$e_\tau(t_j):=\mathcal{E}(u_\tau^j)\,,\qquad\zeta_\tau(t_j):=\sum_{i=1}^j\left\langle u_\tau^{i-1}-u_\tau^i,F_i\right\rangle_*\,.$$

extending these functions to $[0,T]$ by $e_\tau(t)=e_\tau(t_{j-1})$ for $t\in[t_{j-1},t_j)$ and similary for ζ_τ. By (3.24), the functions $\varphi_\tau:=e_\tau+\zeta_\tau$ are nonincreasing lower semi-continuous and by Helly's theorem, there exists a limit function $\varphi(t)$, such that for almost every $t\in[0,T]$ it holds $\varphi_\tau(t)\to\varphi(t)$ and this convergence holds particularly for $t=0$. Furthermore, observe that

$$\begin{aligned}\zeta_\tau(t_J)&=\left\langle u_\tau^J,F_J\right\rangle_*-\left\langle u_\tau^0,F_1\right\rangle_*-\sum_{j=1}^{J-1}\left\langle u_\tau^j,F_{j+1}-F_j\right\rangle_*\\&\to\langle u(t),f(t)\rangle_*-\langle u_0,f(0)\rangle_*-\int_0^t\left\langle u(s),\dot f(s)\right\rangle_*\mathrm{d}s\,.\end{aligned}$$

Hence, by lower semi-continuity of $\mathcal{E}$ and the pointwise convergence (3.25)–(3.26) that

$$\mathcal{E}(u(t))\leq\liminf_{\tau\to0}\mathcal{E}(\overline{u}_\tau(t))=\lim_{\tau\to0}\left(\varphi_\tau(t)-\zeta_\tau(t)\right)$$

exists. We finally obtain that (3.12) holds. □

The Cahn-Hilliard equation

We study study the quasilinear Cahn-Hilliard equation on a domain Ω

$$\partial_t u+\Delta\left(\Delta u-s_0'(u)\right)=0\,,$$

with Neumann boundary conditions $\partial_\nu u=0$ and $\partial_\nu\left(\Delta u-s_0'(u)\right)=0$ and rewrite

$$\partial_t u=-\Delta\left[-\left(-\Delta u+s_0'(u)\right)\right]=-\Delta\left[-D_{L^2}\mathcal{E}(u)\right]\,,$$

where

$$\mathcal{E}(u):=\begin{cases}\int_\Omega\left(s_0(u)+\frac12|\nabla u|^2\right) & \text{for }u\in H^1(\Omega)\,,\\+\infty & \text{otherwise.}\end{cases}\tag{3.27}$$

With our experience from above, we discover immediately

$$\partial_t u = \partial\Psi^*\left(-\mathrm{D}\mathcal{E}(u)\right) ,$$

where $\Psi^*(\xi) = \int_{\mathbb{R}^d} \frac{1}{2}\left|\nabla\xi\right|^2$, which suggests that $\mathscr{B} = H^{-1}_{(0)}(\mathbf{Q})$, where $H^1_{(0)}(\mathbf{Q}) = H^1(\mathbf{Q}) \cap \left\{\int u = 0\right\}$.

A typical choice for s_0 is given by $s_0(u) = u^4 - u^2$ which easily implies that $\mathrm{D}_{L^2}\mathcal{E}(u) = s_0'(u) - \Delta u$ is well defined on $\mathrm{dom}\mathrm{D}\mathcal{E} = H^2(\Omega)$. The Cahn-Hilliard equation is a so called phase-separation model, i.e. in a system of two immiscible fluids (air and water, oil and water, ...) $u = 1$ indicates one phase (say water) and $u = -1$ indicates air. All intermediate values of $u \in (-1, 1)$ indicate the so-called mushy region. In this context, it is more reasonable to restrict the domain of $\mathcal{E}$ to $u \in [-1, 1]$, and a general result from Abels and Wilke justifies this approach:

Lemma 3.2.23. *[**?**] Let $s_1 \equiv 0$ and $s_0 : [a, b] \to \mathbb{R}$ be a continuous and convex function that is twice continuously differentiable in (a, b) and satisfies $\lim_{x\to a} s_0'(x) = -\infty$, $\lim_{x\to b} s_0'(x) = +\infty$. Moreover, we set $s_0' = +\infty$ for $x \notin (a, b)$ and let $\mathcal{E}$ be defined as in (3.27). Then, for the L^2-subdifferential it holds*

$$D\left(\frac{\delta^0\mathcal{E}}{\delta u}\right) = \left\{c \in H^2(\Omega) \cap L^2_{(0)}(\Omega) \,:\, s_0'(c) \in L^2(\Omega),\, s_0''(c)\left|\nabla c\right|^2 \in L^1(\Omega),\, \partial_n c\Big|_{\partial\Omega} = 0\right\} \tag{3.28}$$

and

$$\mathrm{D}_{L^2}\mathcal{E}(\tilde{u}) = -\Delta\tilde{u} + s_0'(\tilde{u})\,. \tag{3.29}$$

The norm on $H^1_{(0)}(\mathbf{Q})$ is given by $\|u\|^2_{H^1_{(0)}} := \int_{\mathbf{Q}}\left|\nabla u\right|^2$ and the Riesz-isomorphism with $H^{-1}_{(0)}(\mathbf{Q})$ is given by $f \mapsto \left(-\Delta_N\right)^{-1} f$, where $\left(-\Delta_N\right)$ is the Laplace-operator with Neumann BC. This follows from the equation

$$\forall\varphi \in H^1_{(0)}(\mathbf{Q}) \,: \qquad \langle f, \varphi\rangle_{H^{-1}_{(0)}, H^1_{(0)}} = \langle u, \varphi\rangle_{H^1_{(0)}} = \int_{\mathbf{Q}} \nabla u \cdot \nabla\varphi\,.$$

In particular, the norm on $H^{-1}_{(0)}(\mathbf{Q})$ is given by $\|f\|^2_{H^{-1}_{(0)}} := \int_{\mathbf{Q}}\left|\nabla\left(-\Delta_N\right)^{-1} f\right|^2$. With all this knowledge, we might set $\Psi^*(\xi) = \frac{1}{2}\|\xi\|^2_{H^{-1}_{(0)}}$, provided we can show that $\mathrm{D}_{H^{-1}_{(0)}}\mathcal{E}(u) = -\Delta\left(-\Delta u + s_0'(u)\right)$.

First note that the domain of $\mathcal{E}$ does not change if we consider it as a functional on $H^{-1}_{(0)}(\mathbf{Q})$ instead of $L^2(\mathbf{Q})$. On the other hand, we already know that $\mathrm{D}_{L^2}\mathcal{E}(u) = -\Delta u + s_0'(u)$. Hence, for $u, u' \in \mathrm{dom}\mathcal{E}$ we find

$$\begin{aligned}\mathcal{E}(u') - \mathcal{E}(u) &= \langle \mathrm{D}_{L^2}\mathcal{E}(u),\, u' - u\rangle + o\left(\|u' - u\|_{L^2}\right)\\ &= \left\langle \nabla\mathrm{D}_{L^2}\mathcal{E}(u),\, \nabla\left(-\Delta_N\right)^{-1} u' - u\right\rangle + Co\left(\|u' - u\|_{H^{-1}_{(0)}}\right) ,\end{aligned}$$

where we used that $\|u' - u\|_{H^{-1}_{(0)}} \leq C\,\|u' - u\|_{L^2}$. From the last equality, we infer $\mathrm{D}_{H^{-1}_{(0)}}\mathcal{E}(u) = -\Delta\mathrm{D}_{L^2}\mathcal{E}(u)$. Since $\mathscr{B} = H^{-1}_{(0)}$ is a Hilbert space, we find $\mathscr{B} = \mathscr{B}^*$ and $\Psi = \Psi^*$.

The p-Laplace equation

We consider

$$\mathcal{E}(u) := \int_{\mathbf{Q}} \frac{1}{p} |\nabla u|^p + \int_{\mathbf{Q}} e(u) , \qquad \Psi^*(\xi) = \int_{\mathbf{Q}} \frac{1}{r} |\xi|^r .$$

With $\frac{1}{r} + \frac{1}{s} = 1$ we find $\Psi(\xi) = \frac{1}{s} \int_{\mathbf{Q}} |\xi|^s$ and the underlying PDE reads

$$\dot{u} = |\xi|^{r-2} \xi , \qquad \xi = -\nabla \cdot \left(|\nabla u|^{p-2} \nabla u \right) + e'(u) + f(t)$$

and the existence theorem yields at least one $u \in W^{1,r}(0,T; W_0^{1,p}(\mathbf{Q}))$ that solves the above problem.

3.3 Monotone Operators

3.3.1 Monotone Operators

We introduce the fundamentals of the theory of monotone operators. In this chapter, we follow [4].

Definition 3.3.1. An operator $A:\,\mathscr{B}\to\mathscr{B}^*$ is called

- radially continuous, if for every $u,v\in\mathscr{B}$ the function $s\to\langle A(u+sv),\,v\rangle$ is continuous on $[0,1]$
- hemi continuous, if for every $u,v,w\in\mathscr{B}$ the function $s\to\langle A(u+sv),\,w\rangle$ is continuous on $[0,1]$
- demi continuous, if $u_n\to u$ in $\mathscr{B}$ implies $Au_n\rightharpoonup^* Au$ weakly$*$ in $\mathscr{B}^*$
- Lipschitz continuous, if there exists $L>0$ such that for every $u,v\in\mathscr{B}$ it holds $\|Au-Av\|_{\mathscr{B}^*}\leq L\,\|u-v\|_{\mathscr{B}}$
- monotone, if for every $u,v\in\mathscr{B}$ it holds $\langle Au-Av,\,u-v\rangle\geq 0$
- strictly monotone, if for every $u,v\in\mathscr{B}$, $u\neq v$, it holds $\langle Au-Av,\,u-v\rangle>0$
- d-monotone, if for every $u,v\in\mathscr{B}$ it holds

  $$\langle Au-Av,\,u-v\rangle\geq\left(\alpha\left(\|u\|_{\mathscr{B}}\right)-\alpha\left(\|v\|_{\mathscr{B}}\right)\right)\left(\|u\|_{\mathscr{B}}-\|v\|_{\mathscr{B}}\right)$$

  for a monoton increasing function α
- uniformly monotone if there exists a positive monotone increasing function $\rho:\,[0,\infty)\to\mathbb{R}$ such that $u,v\in\mathscr{B}$ it holds $\langle Au-Av,\,u-v\rangle\geq\rho\left(\|u-v\|_{\mathscr{B}}\right)$
- strongly monotone if there exists a constant m such that $u,v\in\mathscr{B}$ it holds $\langle Au-Av,\,u-v\rangle\geq m\,\|u-v\|_{\mathscr{B}}^2$

Remark. The above concepts can be directly generalized to $A:\,\mathrm{dom}A\subset\mathscr{B}\to\mathscr{B}^*$.

Definition 3.3.2. An Operator $A:\,\mathscr{B}\to\mathscr{B}^*$ is called coercive if there exists a function $\gamma:\,[0,\infty)\to\mathbb{R}$ with $\lim_{s\to\infty}\gamma(s)=+\infty$ such that

$$\langle Au,u\rangle\geq\gamma\left(\|u\|\right)\|u\|\ .$$

Remark 3.3.3. a) every uniformly monotone operator $A:\mathscr{B}\to\mathscr{B}^*$ is coercive w.r.t. $\gamma(s)=(s-1)\rho(1)-\|A0\|_{\mathscr{B}^*}$. This can be verified by the following calculation: Defining $v:=u\,\|u\|^{-1}$ and n the biggest integer number small than $\|u\|$, it holds

$$\begin{aligned}\langle Au,u\rangle&\geq\|u\|\,\langle Au,v\rangle=\|u\|\left(\langle Au-A(nv),v\rangle+\langle A(nv)-A0,v\rangle+\langle A0,v\rangle\right)\\&\geq\|u\|\left(\langle A(nv)-A0,v\rangle-\|A0\|_{\mathscr{B}^*}\right)\\&\geq\|u\|\left(\sum_{i=1}^n\langle A(iv)-A\left((i-1)v\right),v\rangle-\|A0\|_{\mathscr{B}^*}\right)\\&\geq\|u\|\left(n\rho(1)-\|A0\|_{\mathscr{B}^*}\right)\geq\|u\|\left(\left(\|u\|-1\right)\rho(1)-\|A0\|_{\mathscr{B}^*}\right)\ .\end{aligned}$$

b) let A be d-monotone w.r.t. a function α and $\lim_{s\to\infty}\alpha(s)=\infty$. Then A is coercive w.r.t. $\gamma(s)=\alpha(s)-\alpha(0)$.

Definition 3.3.4. An operator $A:\mathscr{B}\to\mathscr{B}^*$ is Gateaux-differentiable in $b_0\in\mathscr{B}$ if there exists $A'\in(\mathscr{B}\to\mathcal{L}(\mathscr{B},\mathscr{B}^*))$ such that for every $u,v,w\in\mathscr{B}$ holds

$$\lim_{h\to 0}\frac{1}{h}\left\langle A(u+hw)-Au\,,\,v\right\rangle=\left\langle A'(u)w\,,\,v\right\rangle\,.$$

Lemma 3.3.5. *a) An operator $A:\mathscr{B}\to\mathscr{B}^*$ is monotone if and only if for every $u,v\in\mathscr{B}$ the function*

$$t\mapsto\varphi_{u,v}(t):=\left\langle A(u+tv),v\right\rangle$$

is monotone increasing on $[0,1]$.

b) If A is Gateaux-differentiable and $h\mapsto\left\langle A'(u+hv)v\,,\,v\right\rangle$ is continuous on $[0,1]$ then A is monotone if and only if $\left\langle A'(u)v\,,\,v\right\rangle\geq 0$.

Proof. a) "$\Rightarrow$": for $t_1\leq t_2$ it holds

$$\begin{aligned}\varphi_{u,v}(t_2)-\varphi_{u,v}(t_1)&=\left\langle A(u+t_2v),v\right\rangle-\left\langle A(u+t_1v),v\right\rangle\\&=\frac{1}{t_2-t_1}\left\langle A(u+t_2v)-A(u+t_1v)\,,\,u+t_2v-(u+t_1v)\right\rangle\\&\geq 0\,.\end{aligned}$$

a) "$\Leftarrow$": Let $v=w-u$ then

$$\left\langle Au-Aw\,,\,u-w\right\rangle=\varphi_{u,v}(1)-\varphi_{u,v}(0)\geq 0\,.$$

b) "$\Rightarrow$": For $0<s<1$ the intermediate value theorem yields $0<s_0<s$ s.t.

$$0\leq\left\langle A(u+sv)-Au,\,sv\right\rangle=\int_0^s\left\langle A'(u+tv)v,\,sv\right\rangle\mathrm{d}t=s^2\left\langle A'(u+s_0v)v,\,v\right\rangle\,.$$

Deviding by s^2 and performing $s\to 0$ yields $\left\langle A'(u)v\,,\,v\right\rangle\geq 0$.

b) "$\Leftarrow$": We find

$$\left\langle Au-Av\,,\,u-v\right\rangle=\int_0^1\left\langle A'(v+t(u-v))(u-v)\,,\,u-v\right\rangle\mathrm{d}t\geq 0\,.$$

□

Lemma 3.3.6. *Every monotone operator $A:\mathscr{B}\to\mathscr{B}^*$ is locally bounded.*

Proof. We assume A is not locally bounded. Then there exists a sequence $(u_n)_{n\in\mathbb{N}}$ with $u_n\to u$ in $\mathscr{B}$ and $\|Au_n\|_{\mathscr{B}^*}\to\infty$. We define $\alpha_n:=1+\|Au_n\|_{\mathscr{B}^*}\|u_n-u\|$. For arbitrary $v\in\mathscr{B}$ we find by monotonicity of A:

$$\begin{aligned}\frac{1}{\alpha_n}\left\langle Au_n\,,\,v\right\rangle&=\frac{1}{\alpha_n}\left(\left\langle Au_n\,,\,u_n-u\right\rangle+\left\langle A(u+v)\,,\,v+u-u_n\right\rangle+\left\langle Au_n-A(u+v)\,,\,v-u_n+u\right\rangle\right)\\&\leq\frac{1}{\alpha_n}\left(\left\langle Au_n\,,\,u_n-u\right\rangle+\left\langle A(u+v)\,,\,v+u-u_n\right\rangle\right)\\&\leq 1+\frac{1}{\alpha_n}\|A(u+v)\|_{\mathscr{B}^*}\left(\|v\|+\|u-u_n\|\right)\leq M_1\end{aligned}$$

with M_1 beeing an independent constant. A similar estimate holds for $-v$. Hence

$$\limsup_{n\to\infty}\left|\frac{1}{\alpha_n}\langle Au_n\,,\,v\rangle\right|<\infty\qquad\forall v\in\mathscr{B}\,.$$

The Banach-Steinhaus theorem yields $\frac{1}{\alpha_n}\langle Au_n\,,\,v\rangle\leq M$, i.e.

$$\|Au_n\|_{\mathscr{B}^*}\leq M\alpha_n=M\left(1+\|Au_n\|_{\mathscr{B}^*}\|u_n-u\|\right)\,.$$

We chose n_0 s.t. for $n>n_0$ $M\|u-u_n\|\leq\frac{1}{2}$. Then the last inequality implies

$$\|Au_n\|_{\mathscr{B}^*}\leq 2M\,.$$

This is a contradiction to the assumption $\|Au_n\|_{\mathscr{B}^*}\to\infty$. □

Corollary 3.3.7. *Every linear monotone operator A is continuous.*

Proof. Let $u_n\to u$ in $\mathscr{B}$ and define $v_n:=\frac{u_n-u}{\|u_n-u\|^{\frac{1}{2}}}$ if $u_n\neq u$ and $v_n=0$ else. Then $v_n\to 0$ and hence $\|Av_n\|_{\mathscr{B}^*}\leq M$. This implies by linearity

$$\|Au_n-Au\|_{\mathscr{B}^*}\leq M\|u_n-u\|^{\frac{1}{2}}\to 0\,.$$

□

Corollary 3.3.8. *Let A be a monotone operator and $K\subset\mathscr{B}$ a set with two constants M_1,M_2 such that*

$$\sup_{u\in K}\|u\|\leq M_1\,,\qquad\sup_{u\in K}\langle Au\,,\,u\rangle\leq M_2\,,$$

then there exists a constant M such that

$$\sup_{u\in K}\|Au\|\leq M\,.$$

Proof. By Lemma 3.3.6 A is locally bounded around 0. Hence for some $\varepsilon>0$ we obtain by monotonicity

$$\begin{aligned}\|Au\|_{\mathscr{B}^*}&=\sup_{\|y\|=\varepsilon}\frac{1}{\varepsilon}\langle Au\,,\,y\rangle\leq\sup_{\|y\|=\varepsilon}\frac{1}{\varepsilon}\left(\langle Au\,,\,u\rangle+\langle Ay\,,\,y\rangle-\langle Ay\,,\,u\rangle\right)\\&\leq\frac{1}{\varepsilon}\left(M_2+\sup_{\|y\|=\varepsilon}\|Ay\|_{\mathscr{B}^*}(\varepsilon+M_1)\right)\,.\end{aligned}$$

□

Lemma 3.3.9. *Let $A:\mathscr{B}\to\mathscr{B}^*$ be a monotone operator. Then the following statements are equivalent:*

a) A is radially continuous

b) If f is such that for every $v\in\mathscr{B}$ $\langle f-Av,u-v\rangle\geq 0$ then $Au=f$

c) If $u_n \rightharpoonup u$ *in* $\mathscr{B}$ *and* $Au_n \rightharpoonup f$ *in* $\mathscr{B}^*$ *and*

$$\limsup_{n\to\infty} \langle Au_n,\, u_n\rangle \le \langle f,\, u\rangle$$

then $Au = f$*.*

d) A *is demi continuous*

e) If K *is dense in* $\mathscr{B}$ *and* f *is such that for every* $v \in K$ $\langle f - Av, u - v\rangle \ge 0$ *then* $Au = f$

Proof. a)⇒b): Let $v \in \mathscr{B}$ and set $v_t := u + tv$. It holds $0 \le t\,\langle f - Av_t,\, v\rangle$ and, deviding by t, $0 \le \langle f - Av_t,\, v\rangle$. For $t \to 0$ it follows by radial continuity of A the relation $0 \le \langle f - Au,\, v\rangle$. Since v was arbitrary, it follows $f - Au = 0$.

b)⇒c): Let $u_n \rightharpoonup u$ and $Au_n \rightharpoonup f$ and $\limsup_{n\to\infty} \langle Au_n, u_n\rangle \le \langle f,\, u\rangle$. Then one obtains for arbitrary $v \in \mathscr{B}$

$$\begin{aligned}\langle f - Av\,,\, u - v\rangle &= \langle f, u\rangle - \langle f, v\rangle - \langle Av,\, u - v\rangle\\ &\ge \limsup_{n\to\infty} \left(\langle Au_n,\, u_n\rangle - \langle f, v\rangle - \langle Av,\, u - v\rangle\right)\\ &= \limsup_{n\to\infty} \left(\langle Au_n,\, u_n\rangle - \langle Au_n, v\rangle - \langle Av,\, u_n - v\rangle\right)\\ &= \limsup_{n\to\infty} \langle Au_n - Av,\, u_n - v\rangle \ge 0\,.\end{aligned}$$

Due to b) this implies c).

c)⇒d): Let $(u_n)_{n\in\mathbb{N}}$ be a sequence with $u_n \to u$ in $\mathscr{B}$. By local boundedness of A (Lemma 3.3.6) the sequence $\|Au_n\|_{\mathscr{B}^*}$ is bounded. W.l.o.g. assume $Au_n \rightharpoonup f$ in $\mathscr{B}^*$. Then

$$\lim_{n\to\infty} \langle Au_n\,,\, u_n\rangle = \langle f\,,\, u\rangle$$

and thus $Au = f$ and $Au_n \rightharpoonup Au$. Hence the whole sequence converges weakly.

d)⇒e): A is demi continuous and hence radially condinuous. because of a)⇒b) it suffices to show that

$$(\forall v \in K:\quad \langle f - Av\,,\, u - v\rangle \ge 0) \quad\Rightarrow\quad (\forall v \in \mathscr{B}:\quad \langle f - Av\,,\, u - v\rangle \ge 0)\;.$$

By density of K in $\mathscr{B}$ there exists $(v_n)_n \subset K$ with $v_n \to v$. Demi continuity implies

$$\langle f - Av\,,\, u - v\rangle = \lim_{n\to\infty} \langle f - Av_n\,,\, u - v_n\rangle\;.$$

e)⇒a): For the special case $K = \mathscr{B}$ e) coincides with *b*). But b) implies demi continuity and hence radial symmetry of A. This concludes the proof. □

Corollary 3.3.10. *Let* $A : \mathscr{B} \to \mathscr{B}^*$ *be a radialy continuous monotone operator. Then for every* $f \in \mathscr{B}^*$ *the set* K *of solutions to* $Au = f$ *is convex weakly closed.*

Proof. Let $u_1, u_2 \in K$ and let $u_t = tu_1 + (1-t)u_2$, $t \in (0,1)$ then vor every $v \in \mathscr{B}$

$$\langle f - Av,\, u_t - v\rangle = t\,\langle Au_1 - Av,\, u_1 - v\rangle + (1-t)\,\langle Au_2 - Av,\, u_2 - v\rangle \ge 0$$

and by Lemma 3.3.9 $Au_t = f$, i.e. K is convex.

Let u_n be a sequence in K with $u_n \rightharpoonup u$ weakly in $\mathscr{B}$. For arbitrary $v \in \mathscr{B}$ we find

$$\langle f - Av,\, u - v\rangle = \lim_{n\to\infty} \langle f - Av,\, u_n - v\rangle = \lim_{n\to\infty} \langle Au_n - Av,\, u_n - v\rangle \ge 0$$

and by Lemma 3.3.9 $Au_t = f$, i.e. K is weakly closed. □

3.3.2 Stationary Problems with Monotone Operators

Lemma 3.3.11. *Let $B:\mathbb{R}^d\to\mathbb{R}^d$ be a continuous map satisfying*

$$\langle Ba\,,\,a\rangle\geq 0\qquad\text{for all } |a|=r \text{ for some } R>0.$$

Then there exists $a\in\overline{\mathbb{B}_R(0)}$ with $Ba=0$.

Proof. Assume that $Ba\neq 0$ for every $a\in\overline{\mathbb{B}_R(0)}$. Then $a\mapsto -R\frac{Ba}{|Ba|}$ is continuous and by Brouwers Fixedpoint theorem there exists $a\in\overline{\mathbb{B}_R(0)}$ with $a=-R\frac{Ba}{|Ba|}$. Then $|a|=R$ and $\langle Ba\,,\,a\rangle=-R|Ba|^{-1}\langle Ba\,,\,Ba\rangle<0$, a contradiction. □

Theorem 3.3.12 (Browder-Minty)**.** *Let $A:\mathscr{B}\to\mathscr{B}^*$ be monotone, radialy continuous and coercive. Then for every $f\in\mathscr{B}^*$ the set K of solutions to*

$$Au=f$$

is convex, weakly closed and not empty.

Proof. Because of Corollary 3.3.10 it suffices to show that there exists at least one solution.

Let $(h_d)_{n\in\mathbb{N}}$ be a complete system of linearly independent elements and let $\mathscr{B}_d:=\operatorname{span}(h_1,\dots h_d)$. The map

$$C:\quad\mathbb{R}^d\to\mathscr{B}_d\,,\qquad(a_1,\dots a_d)\to\sum_{i=1}^d a_ih_i$$

is continuous and invertible and $|a|_C:=\|Ca\|_{\mathscr{B}}$ defines an equivalent norm on $\mathbb{R}^d$. Next we define

$$B:\,\mathbb{R}^d\to\mathbb{R}^d\,,\qquad a\mapsto\left(\langle ACa-f,\,h_i\rangle\right)_{i=1\dots d}\,,$$

and since A is radialy continuous, it is also demi-continuous and hence B is continuous. The coercivity of A implies for sufficiently large R_1:

$$\left(\frac{\langle Au_d,\,u_d\rangle}{\|u_d\|}-\|f\|_{\mathscr{B}^*}\right)\|u_d\|\geq 0\quad\text{for }\ \|u_d\|>R_1\,.$$

Thus for $|a|=R=R_1c$ it holds

$$\langle Ba,\,a\rangle=\sum_{i=1}^d b_ia_i=\langle Au_d,\,u_d\rangle-\langle f,\,u_d\rangle\geq\left(\frac{\langle Au_d,\,u_d\rangle}{\|u_d\|}-\|f\|_{\mathscr{B}^*}\right)\|u_d\|\geq 0\,.$$

By Lemma 3.3.11 there exists $a\in\mathbb{R}^d$ such that $Ba=0$, i.e. $u_d=Ca$ satisfies

$$\forall i=1\dots d:\quad\langle Au_d,\,h_i\rangle=\langle f,\,h_i\rangle\ .$$

The estimate

$$\frac{\langle Au_d,\,u_d\rangle}{\|u_d\|}\leq\|f\|_{\mathscr{B}^*}$$

and coercivity of A imply $\|u_d\| \leq M_1$ and $\langle Au_d,\, u_d\rangle \leq M_2$ for every $d \in \mathbb{N}$. By Corollary 3.3.8 it holds $\|Au_d\|_{\mathscr{B}^*} \leq M$ and further it holds

$$\lim_{d\to\infty} \langle Au_d,\, h\rangle = \langle f,\, h\rangle \qquad \forall h \in \bigcup_d \mathscr{B}_d\,.$$

Therefore $Au_d \rightharpoonup f$ in $\mathscr{B}^*$. Now let $u_d \rightharpoonup u$ along a subsequence. We find

$$\lim_{d\to\infty} \langle Au_d,\, u_d\rangle = \lim_{d\to\infty} \langle f,\, u_d\rangle = \langle f,\, u\rangle\;,$$

and hence by Lemma 3.3.9c) $Au = f$. This provides the statement. □

Theorem 3.3.13. *Let $A : \mathscr{B} \to \mathscr{B}^*$ be radially continuous, strictly monotone and coercive. Then $A^{-1} : \mathscr{B}^* \to \mathscr{B}$ exists, is monotone and bounded. If A is strongly monotone, then A^{-1} is Lipschitz continuous.*

Proof. We provide the proof in 4 steps.

Step 1: By Theorem 3.3.12 it suffices to prove that $Au = f$ has only one solution. Let u, v be solutions then $\langle Au - Av,\, u - v\rangle = 0$ and the strict monotonicity provides $u = v$.

Step 2: For $f, g \in \mathscr{B}^*$ and $u := A^{-1}f$, $v := A^{-1}g$ we find by strict monotonicity of A

$$\left\langle f - g,\, A^{-1}f - A^{-1}g\right\rangle = \langle Au - Av,\, u - v\rangle > 0\,.$$

Step 3: Let $\|f\|_{\mathscr{B}^*} < M$ and let $Au = f$. Then

$$\langle Au,\, u\rangle \geq \|u\|\,\gamma\left(\|u\|\right)\,, \quad \text{i.e. } \gamma\left(\|u\|\right) \leq \|f\|_{\mathscr{B}^*}\,.$$

Because $\gamma(x) \to \infty$ for $x \to \infty$ this implies $\|u\| = \|A^{-1}f\| \leq K$ for a constant K depending on f.

Step 4: For $f, g \in \mathscr{B}^*$ and $u := A^{-1}f$, $v := A^{-1}g$ we find

$$\|f - g\|_{\mathscr{B}^*}\,\|u - v\|_{\mathscr{B}} \geq \langle Au - Av,\, u - v\rangle \geq m\,\|u - v\|_{\mathscr{B}}^2 = m\,\|u - v\|_{\mathscr{B}}\left\|A^{-1}f - A^{-1}g\right\|_{\mathscr{B}}\,.$$

□

3.3.3 Evolution Equations with Continuous Nonlinear Operators

We study the following ODE on $S = [0, T]$

$$\dot{u}(t) + G(t)u(t) = f(t)\,,$$

where $G(t)$ is a family of nonlinear operators $\mathscr{B} \to \mathscr{B}$ and $f : S \to \mathscr{B}$ is sufficiently regular. In Theorem 1.2.10 we studied this problem in the context of compact operators G. Here we will study a different setting. In particular, we will consider the following setting

1. For every $b \in \mathscr{B}$ the function $t \mapsto G(t)b$ is continuous.
2. The family $G(t) : \mathscr{B} \to \mathscr{B}$ is uniformly Lipschitz continuous. I.e. there exists $L \in \mathbb{R}$ such that for all $a, b \in \mathscr{B}$ and all $t \in S$

$$\|G(t)a - G(t)b\|_{\mathscr{B}} \leq L\,\|a - b\|_{\mathscr{B}}\,.$$

Lemma 3.3.14. *Let 1.-2. hold then for every $u \in C(S;\mathscr{B})$ it holds $Gu \in C(S;\mathscr{B})$.*

Proof. Let $t_n \to t$ in S as $n \to \infty$. Then it holds

$$\|G(t_n)u(t_n) - G(t)u(t)\|_{\mathscr{B}} \leq L\,\|u(t_n) - u(t)\|_{\mathscr{B}} + \|G(t_n)u(t) - G(t)u(t)\|_{\mathscr{B}}$$

Then the Lemma follows from convergence $u(t_n) \to u(t)$ and 1. □

Lemma 3.3.15. *The norms*

$$\|u\|_{C,k} := \sup_{t\in S}\left\{e^{kt}\,\|u(t)\|_{\mathscr{B}}\right\}$$

are equivalent with the standard norm on $C(S;\mathscr{B})$.

Proof. Evidently, $\|\cdot\|_{C,k}$ are norms and we find

$$\|u\|_{C,k} \leq \|u\|_{C(S;\mathscr{B})} \leq e^{kT}\,\|u\|_{C,k} \,.$$

□

Lemma 3.3.16 (Gronwall)**.** *Let $f : S \to \mathbb{R}$ continuous and $g : S \to \mathbb{R}$ non-decreasing . If for some $c > 0$*

$$\forall t \in S : \quad f(t) \leq g(t) + c\int_0^t f(s)\mathrm{d}s\,, \tag{3.30}$$

then

$$\forall t \in S : \qquad f(t) \leq e^{ct}g(t)\,.$$

In particular, for $g \equiv 0$ and $f \geq 0$ we infer $f \equiv 0$.

Proof. From (3.30) we infer

$$f(t) \leq g(t) + c\int_0^t \left(g(s) + c\int_0^s f(r)\mathrm{d}r\right)\mathrm{d}s\,,$$

and hence by induction

$$f(t) \leq g(t)\sum_{k=0}^{n}\frac{(ct)^k}{k!} + R_{n+1}(t)\,,$$

where

$$R_{n+1}(t) = c^{n+1}\int_0^t\int_0^{s^1}\cdots\int_0^{s_n} f(s_{n+s})\mathrm{d}s_{n+1}\cdots\mathrm{d}s_1\,.$$

With $M = \sup_{t\in S} f(t)$ we find

$$\forall t \in S : \qquad |R_{n+1}(t)| \leq \frac{M(cT)^{n+1}}{(n+1)!}\,.$$

Since $R_{n+1}(t) \to 0$ uniformly on S as $n \to \infty$, the claim follows. □

Theorem 3.3.17. *For every $f \in C(S;\mathscr{B})$ and $u_0 \in \mathscr{B}$ the Cauchy-problem*

$$\dot{u}(t) + G(t)u(t) = f(t)\,, \qquad u(0) = u_0 \tag{3.31}$$

has a unique solution $u \in C(S;\mathscr{B})$. The mapping $\mathscr{B} \times C(S;\mathscr{B}) \to C^1(S;\mathscr{B})$, $(u_0, f) \to u$ is continuous.

Proof. We use the Banach fixed point theorem and define for $t \in S$:

$$(\mathcal{U}u)(t) := u_0 - \int_0^t (G(s)u(s) - f(s))\,\mathrm{d}s\,.$$

Evidently, we can diffentiate the last expression and hence $\mathcal{U} : C(S;\mathscr{B}) \to C^1(S;\mathscr{B})$ is continuous. We show that $\mathcal{U}$ is a contraction with respect to one of the above introduced norsm $\|\cdot\|_{C,k}$ for suitable k. By definition of $\mathcal{U}$, we obtain for $u, v \in C(S;\mathscr{B})$

$$\begin{aligned}
\|\mathcal{U}u(t) - \mathcal{U}v(t)\|_{\mathscr{B}} &\leq \int_0^t \|G(s)u(s) - G(s)v(s)\|_{\mathscr{B}}\, e^{-ks}e^{ks}\mathrm{d}s \\
&\leq L\int_0^t \|u(s) - v(s)\|_{\mathscr{B}}\, e^{-ks}e^{ks}\mathrm{d}s \\
&\leq L\,\|u - v\|_{C,k}\,\frac{e^{kt} - 1}{k}\,.
\end{aligned}$$

Multiplying both sides with e^{-kt} and taking the supremum over S, we obtain

$$\|\mathcal{U}(u - v)\|_{C,k} \leq \frac{L}{k}\left(1 - e^{-kT}\right)\|u - v\|_{C,k}\,.$$

Choosing $k > L$, the Banach fixed point theorem yields exists of a unique element $u \in C^1(S;\mathscr{B})$ such that $u = \mathcal{U}u$ and differentiating the last equation in t, we obtain (3.31).

In order to prove the continuity, let $f_1, f_2 \in C(S;\mathscr{B})$ and $u_{0,1}, u_{0,2} \in \mathscr{B}$. Then the corresponding solutions $u_1, u_2 \in C^1(S;\mathscr{B})$ of (3.31) satisfy

$$u_i(t) := u_{0,i} - \int_0^t (G(s)u_i(s) - f_i(s))\,\mathrm{d}s\,.$$

Taking the difference of these two equations we find

$$\|u_1(t) - u_2(t)\|_{\mathscr{B}} \leq \|u_{0,1} - u_{0,2}\|_{\mathscr{B}} + T\,\|f_1 - f_2\|_{C(S;\mathscr{B})} + L\int_0^t \|u_1(s) - u_2(s)\|_{\mathscr{B}}\,\mathrm{d}s\,.$$

With help of the Gronwall Lemma, we obtain

$$\|u_1 - u_2\|_{C(S;\mathscr{B})} \leq K\left(\|u_{0,1} - u_{0,2}\|_{\mathscr{B}} + \|f_1 - f_2\|_{C(S;\mathscr{B})}\right)$$

From the original ODE, we furthermore obtain by the last estimate

$$\|\dot{u}_1 - \dot{u}_2\|_{C(S;\mathscr{B})} \leq K\left(\|u_{0,1} - u_{0,2}\|_{\mathscr{B}} + \|f_1 - f_2\|_{C(S;\mathscr{B})}\right)\,.$$

Together, this implies the claimes continuity $(u_0, f) \to u$. □

Definition 3.3.18. Let $\mathscr{B}_1$, $\mathscr{B}_2$ be Banch spaces. A map $G:\,\mathrm{dom}G\to\{S\to\mathscr{B}_2\}$, $\mathrm{dom}G\subset S$, is called Volterra operator if for almost all $t\in S$ it holds

$$(u(s)=v(s)\text{ for a.e. } s\in[0,t])\qquad\Rightarrow\qquad (Gu(s)=Gv(s)\text{ for a.e. } s\in[0,t])\,.$$

In what follows, we consider $\mathscr{B}:=\mathscr{B}_1=\mathscr{B}_2$ and make the following additional assumption:

$$\exists L>0:\quad \|Gu-Gv\|_{C(S;\mathscr{B})}\leq L\,\|u-v\|_{C(S;\mathscr{B})}\,. \tag{3.32}$$

Lemma 3.3.19. *If a Volterra-Operator G satisfies condition (3.32) then for all $u,v\in C(S;\mathscr{B})$ and all $t\in S$:*

$$\|Gu-Gv\|_{C([0,t];\mathscr{B})}\leq L\sup_{s\in[0,t]}\|u(s)-v(s)\|_{\mathscr{B}}\,.$$

Proof. Let $t\in S$; We define

$$u_t(s):=\begin{cases}u(s) & \text{for } 0\leq s\leq t\\ u(t) & \text{for } t<s<T\end{cases}$$

and v_t correspondingly. Evidently, $u_t,v_t\in C(S;\mathscr{B})$ and

$$\|Gu-Gv\|_{C([0,t];\mathscr{B})}\leq\|Gu_t-Gv_t\|_{C(S;\mathscr{B})}\leq L\sup_{s\in S}\|u_t(s)-v_t(s)\|_{\mathscr{B}}\leq L\sup_{s\in[0,t]}\|u(s)-v(s)\|_{\mathscr{B}}\,.$$

□

Example 3.3.20. Let $h\in C(S)$ and $0\leq h(t)\leq t$ on S. Furthermore, let $\{G(t)\}$ be a family of operators satisfying 1. and 2. Then $Gu(t):=G(t)u(h(t))$ is a Volterra-Operator.

Theorem 3.3.21. *Let the Volterra-Operator G satisfy condition (3.32). For every $f\in C(S;\mathscr{B})$ and $u_0\in\mathscr{B}$ the Cauchy-problem*

$$\dot u(t)+G(t)u(t)=f(t)\,,\qquad u(0)=u_0$$

has a unique solution $u\in C(S;\mathscr{B})$. The mapping $\mathscr{B}\times C(S;\mathscr{B})\to C^1(S;\mathscr{B})$, $(u_0,f)\to u$ is continuous.

Proof. Like in the proof of Theorem 3.3.17 we define for $t\in S$:

$$(\mathcal{U}u)\,(t):=u_0-\int_0^t\left((Gu)\,(s)-f(s)\right)\mathrm{d}s\,. \tag{3.33}$$

Since $f\in C(S;\mathscr{B})$ and $G:\,C(S;\mathscr{B})\to C(S;\mathscr{B})$ it holds $\mathcal{U}:\,C(S;\mathscr{B})\to C^1(S;\mathscr{B})$. From (3.33) it follows

$$\begin{aligned}\|\mathcal{U}u(t)-\mathcal{U}v(t)\|_{\mathscr{B}}&\leq\int_0^t\|(Gu)\,(s)-(Gv)\,(s)\|_{\mathscr{B}}\,\mathrm{d}s\\&\leq\int_0^t\|Gu-Gv\|_{C([0,s];\mathscr{B})}\,\mathrm{d}s\leq L\int_0^t\|u-v\|_{C([0,s];\mathscr{B})}\,\mathrm{d}s\\&\leq L\int_0^t\sup_{0\leq r\leq s}\|u(r)-v(r)\|_{\mathscr{B}}\,e^{-ks}e^{ks}\mathrm{d}s\\&\leq L\int_0^t\sup_{0\leq r\leq s}\|u(r)-v(r)\|_{\mathscr{B}}\,e^{-kr}e^{ks}\mathrm{d}s\end{aligned}$$

Using the $\|\cdot\|_{C,k}$-norms as above, we end up with

$$\|\mathcal{U}u-\mathcal{U}v\|_{C,k}\leq\frac{L}{k}\left(1-e^{-kT}\right)\|u-v\|_{C,k}\,.$$

From here, we can conclude as in the proof of Theorem 3.3.17. □

3.3.4 Pseudoparabolic Evolution Equations

We study the two equations

$$\begin{aligned}A(t)\dot{u}(t)+(Bu)(t)&=f(t)\,,\\ \dot{\overline{Au}}(t)+(Bu)(t)&=f(t)\,,\end{aligned}$$

where A will satisfy the following assumptions.

The family $A=A(t)$ constits of nonlinear mappings such that

$$\forall t\in[0,T]:\qquad A(t):\mathscr{B}\to\mathscr{B}^*\text{is radially continuous}\tag{3.34}$$
$$\forall x\in\mathscr{B}:\quad (t\mapsto A(t)x)\in C(S;\mathscr{B}^*)\tag{3.35}$$
$$\forall t\in[0,T]:\quad \forall x,y\in\mathscr{B}\,\langle A(t)x-A(t)y,\,x-y\rangle\geq m\|x-y\|_{\mathscr{B}}^p\tag{3.36}$$
$$\exists L>0:\quad \|Bu-Bv\|_{C(S;\mathscr{B}^*)}\leq L\|u-v\|_{C(S;\mathscr{B})}\tag{3.37}$$

Lemma 3.3.22. *Assuming (3.34)–(3.36) the operators $A^{-1}(t):\mathscr{B}^*\to\mathscr{B}$ exist and for every $t\in S$*

$$\left\|A^{-1}(t)x^*-A^{-1}(t)y^*\right\|_{\mathscr{B}}\leq m^{-1}\|x-y\|_{\mathscr{B}^*}^{\frac{1}{p-1}}\,.\tag{3.38}$$

Furthermore, for every $x^*\in\mathscr{B}^*$ the function $t\mapsto A^{-1}(t)x^*$ is continuous.

Proof. The existence of the inverse and estimate (3.38) follows from Theorem 3.3.13.

The second part follows from

$$\begin{aligned}\left\|A^{-1}(t)x^*-A^{-1}(s)x^*\right\|_{\mathscr{B}}&=\left\|A^{-1}(s)A(s)A^{-1}(t)x^*-A^{-1}(s)x^*\right\|_{\mathscr{B}}\\&\leq\frac{1}{m}\left\|A(s)A^{-1}(t)x^*-x^*\right\|_{\mathscr{B}^*}^{\frac{1}{p-1}}=\frac{1}{m}\left\|A(s)A^{-1}(t)x^*-A(t)A^{-1}(t)x^*\right\|_{\mathscr{B}^*}^{\frac{1}{p-1}}\to0\,.\end{aligned}$$

□

Theorem 3.3.23. *Let (3.34)–(3.37) be satisfied. For every $f\in C(S;\mathscr{B}^*)$ and $u_0\in\mathscr{B}$ the Cauchy-problem*

$$A(t)\,\dot{u}(t)+(Bu)(t)=f(t)\,,\qquad u(0)=u_0\,.$$

has a unique solution $u\in C(S;\mathscr{B})$. The mapping $\mathscr{B}\times C(S;\mathscr{B})\to C^1(S;\mathscr{B})$, $(u_0,f)\to u$ is continuous.

Proof. Instead of the original problem, we consider

$$\dot{u}(t) + (Gu)(t) = f(t), \qquad u(0) = u_0$$

where

$$\forall u \in C(S;\mathscr{B}),\, t \in S: \qquad (Gu)(t) := -A^{-1}(t)\left(-(Bu)(t) + f(t)\right).$$

Hence, it suffices to verify the conditions of Theorem 3.3.21. Since B is Volterra operator, so is G. Defining $v_0 := f(t_0) - (Bu)(t_0)$ we obtain

$$\begin{aligned}\lim_{t\to t_0} \|(Gu)(t) - (Gu)(t_0)\| \leq \frac{1}{m} &\lim_{t\to t_0} \left(\|(Bu)(t) - (Bu)(t_0)\|_{\mathscr{B}^*} + \|f(t) - f(t_0)\|_{\mathscr{B}^*}\right)^{\frac{1}{p-1}} \\ &+ \lim_{t\to t_0} \left\|A^{-1}(t)v_0 - A^{-1}(t_0)v_0\right\|\end{aligned}$$

which implies $G: C(s;\mathscr{B}) \to C(S;\mathscr{B})$.

For $u, v \in C(S;\mathscr{B})$ we obtain using Lemma 3.3.22

$$\begin{aligned}\|(Gu)(t) - (Gv)(t)\| &\leq \frac{1}{m}\|(Bu)(t) - (Bv)(t)\|_{\mathscr{B}^*} \\ &\leq \frac{1}{m}\|Bu - Bv\|_{C(S;\mathscr{B}^*)} \leq \frac{L}{m}\|u - v\|_{C(S;\mathscr{B})}\end{aligned}$$

and hence

$$\|Gu - Gv\|_{C(S;\mathscr{B})} \leq \frac{L}{m}\|u - v\|_{C(S;\mathscr{B})}.$$

□

3.3.5 Fitzpatrick-Theory and Very Generalized Gradient Flows

In this last chapter of the lecture, we will draw a link from the theory of monotone operators back to the theory of gradient flows. The following link between convex funtionals and maximal monotone graphs has been developed by Fitzpatrick in a seminal paper [3]. The knowledge on the relation to gradient flows is not widely spread, though.

Throughout this section, $\mathscr{B}$ is assumed to be locally convex and Hausdorff, implying $\mathscr{B}$ is *reflexive*.

Definition 3.3.24. The graph of a monotone operator $A: \mathscr{B} \to \mathscr{B}^*$ is defined through

$$G(A) := \{(b, b^*): \quad b^* \in Ab\}.$$

A is called maximal monotone, if there is no other monotone operator $\tilde{A}$ such that $G(A) \subset G(\tilde{A})$ is properly contained.

As we have seen in Lemma 3.2.6, the subdifferential of a convex l.s.c. function is monotone. In the following, we will study convex functions which depend on both $\mathscr{B}$ and $\mathscr{B}^*$. Let $\psi: \mathscr{B} \times \mathscr{B}^* \to \mathbb{R}$ be a convex function and denote $\psi_b: \mathscr{B}^* \to \mathbb{R}$, $b^* \mapsto \psi(b, b^*)$ and $\psi^{b^*}: \mathscr{B} \to \mathbb{R}$, $b \mapsto \psi(b, b^*)$. Recalling the definition of sub-derivative in (3.6) we denote $\partial_{\mathscr{B}}\psi(b, b^*) = \partial\psi^{b^*}(b)$ and $\partial_{\mathscr{B}^*}\psi(b, b^*) = \partial\psi_b(b^*)$.

Definition 3.3.25. For every convex function $\psi : \mathscr{B} \times \mathscr{B}^* \to \mathbb{R}$ we define (and observe)

$$\mathrm{T}_\psi b := \{b^* \in \mathscr{B}^* \,:\, (b^*, b) \in \partial\psi(b, b^*)\}$$

and because

$$\langle b^* - a^*,\, b - a\rangle = \frac{1}{2}\left\langle (b^*, b) - (a^*, a)\,,\; (b, b^*) - (a, a^*)\right\rangle \geq 0\,,$$

we find T_ψ is monotone for convex ψ.

Example 3.3.26. Fitzpatrick choses as an example some convex and l.s.c. $p : \mathscr{B} \to \mathbb{R}$ and defines the function

$$\psi(b, b^*) := p(b) + p^*(b^*) = \sup_{a \in \mathscr{B}} \left(p(b) - p(a) + \langle b^*, a\rangle\right)$$

with the corresponding $\mathrm{T}_\psi = \partial p$: In particular we find from Lemma 3.2.6

$$\langle b, b^*\rangle \leq \psi(b, b^*) \quad \Leftrightarrow \quad b^* \in \partial p(b) \quad \Leftrightarrow \quad b \in \partial p^*(b^*) \quad \Leftrightarrow \quad (b^*, b) \in \partial\psi(b, b^*) \quad \Leftrightarrow \quad b^* \in \mathrm{T}_\psi b\,.$$

We hope to establish the latter relation in more generality. To this aim, we first observe the following.

Lemma 3.3.27. *Let $\psi : \mathscr{B} \times \mathscr{B}^* \to \mathbb{R}$ be convex with $\psi(b, b^*) \geq \langle b, b^*\rangle$for all (b, b^*) in a neighborhood U of (b_0, b_0^*). If $\psi(b_0, b_0^*) = \langle b_0, b_0^*\rangle$ then then $b_0^* \in \mathrm{T}_\psi b_0$.*

Proof. Let $(a, a^*) \in \mathscr{B} \times \mathscr{B}^*$ and $s > 0$ such that $(b_0 + sa, b_0^* + sa^*) \in U$. Then

$$\begin{aligned} f\left(b_0 + a, b_0^* + a^*\right) - f\left(b_0, b_0^*\right) &\geq \frac{1}{s}\left(f\left(b_0 + sa, b_0^* + sa^*\right) - f\left(b_0, b_0^*\right)\right) \\ &\geq \langle b_0 + sa, b_0^* + sa^*\rangle - \langle b_0, b_0^*\rangle \\ &= \langle b_0, a^*\rangle + \langle a, b_0^*\rangle + s\,\langle a, a^*\rangle\;. \end{aligned}$$

Letting $s \to 0$ we have

$$f\left(b_0 + a, b_0^* + a^*\right) - f\left(b_0, b_0^*\right) \geq \langle b_0, a^*\rangle + \langle a, b_0^*\rangle\;,$$

which is $(b_0^*, b_0) \in \partial\psi(b_0, b_0^*)$ and hence $b_0^* \in \mathrm{T}_\psi(b_0)$. □

Lemma 3.3.28. *Let $\psi : \mathscr{B} \times \mathscr{B}^* \to \mathbb{R}$ be convex and $b \in \partial\psi_b(b^*)$ such that*

$$\sup_{v^* \in \mathscr{B}^*} \left(\langle b, v^*\rangle - \psi(b, v^*)\right) = 0\,. \tag{3.39}$$

Then $\psi(b, b^*) = \langle b, b^*\rangle$.

Proof. Since $b \in \partial\psi_b(b^*)$ it holds for every $v^* \in \mathscr{B}^*$ that $\langle v^*, b\rangle \leq \psi_b(b^* + v^*) - \psi_b(b^*)$ and

$$\langle v^* + b^*, b\rangle - \psi_b(b^* + v^*) \leq \langle b^*, b\rangle - \psi_b(b^*)\,.$$

Taking the supremum over v^* it holds $0 \leq \langle b^*, b\rangle - \psi_b(b^*)$. However, putting $v^* = b^*$ in (3.39) it holds $\langle b^*, b\rangle - \psi_b(b^*) \leq 0$ and hence $\langle b^*, b\rangle = \psi_b(b^*)$. □

Theorem 3.3.29. *Let*

$$\forall b \in \mathscr{B} : \quad \sup_{b^* \in \mathscr{B}^*} \left(\langle b, b^* \rangle - \psi(b, b^*) \right) = 0 \,.$$

Then $b \in \partial\psi_b\,(b^*)$ if and only if $b^* \in \mathrm{T}_\psi b$.

Proof. Let $b \in \partial\psi_b\,(b^*)$. Then Lemma 3.3.28 implies $\psi(b, b^*) = \langle b, b^* \rangle$ and Lemma 3.3.27 implies $b^* \in \mathrm{T}_\psi b$.

Vice versa, let $b^* \in \mathrm{T}_\psi b$. Then $b \in \partial\psi_b\,(b^*)$ follows by definition. □

Lemma 3.3.30. *For every maximal monotone operator* T *let*

$$L_{\mathrm{T}}(b, b^*) := \sup_{(a,a^*) \in G(\mathrm{T})} \left(\langle b^*, a \rangle + \langle a^*,\, b - a \rangle \right) \,.$$

Then L_{T} is convex and l.s.c..

Proof. Convexity follows from

$$\begin{aligned} L_{\mathrm{T}}(sb + (1-s)\tilde{b}, sb^* + (1-s)\tilde{b}^*) &:= \sup_{(a,a^*) \in G(\mathrm{T})} \left(\left\langle sb^* + (1-s)\tilde{b}^*, a \right\rangle + \left\langle a^*,\, sb + (1-s)\tilde{b} - a \right\rangle \right) \\ &\leq s \sup_{(a,a^*) \in G(\mathrm{T})} \left(\langle b^*, a \rangle + \langle a^*,\, b - a \rangle \right) + (1-s) \sup_{(a,a^*) \in G(\mathrm{T})} \left(\left\langle \tilde{b}^*, a \right\rangle + \left\langle a^*,\, \tilde{b} - a \right\rangle \right) \,. \end{aligned}$$

Let $(b_n, b_n^*, c_n)_{n\in\mathbb{N}} \subset \operatorname{epi} L_{\mathrm{T}}$ a sequence with $(b_n, b_n^*, c_n) \to (b, b^*, c)$. Then for all $(a, a^*) \in G(\mathrm{T})$ it holds

$$c_n \geq \langle b_n^*, a \rangle + \langle a^*,\, b_n - a \rangle$$

and hence

$$c \geq \langle b^*, a \rangle + \langle a^*,\, b - a \rangle \,.$$

This implies lower semicontinuity by Lemma 3.2.1. □

Lemma 3.3.31. *If* T *is a monotone operator on $\mathscr{B}$ and $(a, a^*) \in G(\mathrm{T})$ and for some $(b, b^*) \in \mathscr{B} \times \mathscr{B}^*$ we have*

$$L_{\mathrm{T}}\,(b, b^*) = \langle a, b^* \rangle + \langle b - a,\, a^* \rangle \,,$$

then $(a^*, a) \in \partial L_{\mathrm{T}}\,(b, b^*)$. Furthermore, if $(b, b^*) \in G(\mathrm{T})$ then $L_{\mathrm{T}}\,(b, b^*) = \langle b, b^* \rangle$ and $(b^*, b) \in \partial L_{\mathrm{T}}\,(b, b^*)$.

Proof. For each $u \in \mathscr{B}$ and $u^* \in \mathscr{B}^*$ we have

$$\begin{aligned} L_{\mathrm{T}}\,(b + u, b^* + u^*) - L_{\mathrm{T}}(b, b^*) &= \\ &= \sup_{(y,y^*) \in G(\mathrm{T})} \left(\langle y, b^* + u^* \rangle + \langle b + u, y^* \rangle - \langle y, y^* \rangle \right) - L_{\mathrm{T}}\,(b, b^*) \\ &\geq \langle a, b^* + u^* \rangle + \langle b + u, a^* \rangle - \langle a, a^* \rangle - \langle a, b^* \rangle - \langle b - a,\, a^* \rangle \\ &= \langle u, a^* \rangle + \langle a, u^* \rangle \,, \end{aligned}$$

i.e. $(a^*, a) \in \partial L_{\mathrm{T}}\,(b, b^*)$.

If $(b, b^*) \in G(\mathrm{T})$ then monotonicity implies $\langle b, b^* \rangle \geq \langle a, b^* \rangle + \langle b - a, a^* \rangle$ implying $L_{\mathrm{T}}(b, b^*) \leq \langle b, b^* \rangle$. On the other hand by definition

$$L_{\mathrm{T}}(b, b^*) \geq \langle b, b^* \rangle + \langle b - b, b^* \rangle = \langle b, b^* \rangle \,,$$

implying the statement by Lemma 3.3.27. □

We will find the following result useful.

Lemma 3.3.32. *If* T *is a monotone operator on* $\mathscr{B}$ *then* T *is maximal monotone if and only if* $L_{\mathrm{T}}(b,b^*) > \langle b,b^*\rangle$ *whenever* $b\in\mathscr{B}$ *and* $b^*\in\mathscr{B}^*\backslash\mathrm{T}(b)$.

Proof. Let $L_{\mathrm{T}}(b,b^*)\leq\langle b,b^*\rangle$ such that for all $(a,a^*)\in G(\mathrm{T})$ it holds

$$\langle a,b^*\rangle+\langle b-a,b^*\rangle\leq\langle b,b^*\rangle\quad\Rightarrow\quad\langle a-b,a^*-b^*\rangle\geq 0\,.$$

Since T is maximal monotone, this implies $b^*\in\mathrm{T}b$.

If T is not maximal monotone, then there exists $b\in\mathscr{B}$ and $b^*\in\mathscr{B}^*\backslash\mathrm{T}b$ such that for all $(a,a^*)\in G(\mathrm{T})$ it holds $\langle a-b,a^*-b^*\rangle\geq 0$. But then $L_{\mathrm{T}}(b,b^*)\leq\langle b,b^*\rangle$. □

We are now able to prove our main result of this section.

Theorem 3.3.33. *Let* A *be maximal monotone on* $\mathscr{B}$*. Then* $L_A(b,b^*)\geq\langle b,b^*\rangle$ *for every* $(b,b^*)\in\mathscr{B}\times\mathscr{B}^*$ *and* $L_A(b,b^*)=\langle b,b^*\rangle$ *if and only if* $(b,b^*)\in G(A)$.

Proof. This follows from Lemmas 3.3.32 and 3.3.27. □

We make use of Theorem 3.3.33 as follows. Assume we study the equation

$$\dot u = g\left(-\mathrm{D}\mathcal{E}(u)+f\right)\,,$$

similar to Section 3.2.3, but with $\partial\Psi$ replaced by a maximal monotone operator g. In particular, we find $(\dot u,\,-\mathrm{D}\mathcal{E}(u)+f)\in G(g)$ and hence we might proceed as in Section to find that solutions to satisfy

$$\mathcal{E}(T,u(T))+\int_0^T L_g\left(\dot u,\,-\mathrm{D}\mathcal{E}(\cdot,u)+f\right)\leq\mathcal{E}(0,u_0)+\int_0^T\left\langle\dot f,u\right\rangle\,,$$

while for all curves $u\in C^1(0,T;\mathscr{B})$ we find

$$\mathcal{E}(T,u(T))+\int_0^T L_g\left(\dot u,\,-\mathrm{D}\mathcal{E}(\cdot,u)+f\right)\geq\mathcal{E}(0,u_0)+\int_0^T\left\langle\dot f,u\right\rangle\,.$$

Bibliography

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[2] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*, volume 28. Siam, 1999.

[3] Simon Fitzpatrick et al. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization*, pages 59–65. Centre for Mathematics and its Applications, Mathematical Sciences Institute, 1988.

[4] Herbert Gajewski, Konrad Gröger, and Klaus Zacharias. Nichtlineare operatorgleichungen und operatordifferentialgleichungen. *Mathematische Nachrichten*, 67(22):iv–iv, 1975.

[5] David G Luenberger. *Optimization by vector space methods*. John Wiley & Sons, 1997.

[6] John Milnor and David W Weaver. *Topology from the differentiable viewpoint*. Princeton university press, 1997.

[7] L. Nirenberg. *Topics in Nonlinear Functional Analysis*. Courant Institute, 1974.

[8] Ralph Rockafellar. Integrals which are convex functionals. *Pacific journal of mathematics*, 24(3):525–539, 1968.

[9] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.

[10] Riccarda Rossi, Alexander Mielke, and Giuseppe Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze-Serie IV*, 7(1):97, 2008.

[11] Riccarda Rossi and Giuseppe Savaré. Gradient flows of non convex functionals in hilbert spaces and applications. *ESAIM: Control, Optimisation and Calculus of Variations*, 12(3):564–614, 2006.

[12] Michael Ruzicka. *Nichtlineare Funktionalanalysis: Eine Einführung*. Springer-Verlag, 2006.

[13] Jacques Simon. Compact sets in the spacel p (o, t; b). *Annali di Matematica pura ed applicata*, 146(1):65–96, 1986.

[14] E. Zeidler. *Nonlinear Functional Analysis I, Fixed Point Theorems*. Springer, 1986.