Differential Equations II/B

Alex Kaltenbach

February 3, 2025

"The problem with linear theory is that it is not nonlinear"

— John A. Adam[1]

[1]Adam is a 1971 graduate, with first-class honours, from Queen Elizabeth College. He completed his Ph.D. in 1974 at University College London. His PhD thesis, with the title *A Theoretical Study of Magnetohydrodynamic Processes in Solar Active Regions*, was jointly supervised by astrochemist Gillian Peach and astrophysicist Carole Jordan.

Contents

0 Introduction **7**

1 Linear Functional Analysis **16**
1.1 Dual Spaces . . . 16
1.2 Hahn–Banach Theorem . . . 17
1.3 Bi-Dual Space and Reflexivity . . . 19
1.4 Notions of Convergence . . . 21

2 Monotone Operator Theory **28**
2.1 Notions of continuity and boundedness . . . 30
2.2 Notions of monotonicity . . . 33
2.3 Notion of coercivity . . . 40
2.4 Main theorem on monotone operators . . . 41
2.5 Applications of the main theorem on monotone operators . . . 50
2.5.1 p-Laplace equation . . . 50
2.5.2 p-Stokes equations . . . 55

3 Pseudo-Monotone Operator Theory **66**
3.1 Brief review of the condition (M) . . . 67
3.2 Pseudo-monotonicity . . . 70
3.3 Main theorem on pseudo-monotone operators . . . 75
3.4 Applications of the main theorem on pseudo-monotone operators . . . 77
3.4.1 p-Navier–Stokes equations . . . 77

4 Calculus of Variations **85**
4.1 Direct Method in the Calculus of Variations . . . 85

Bibliography **93**

0 Introduction

In linear functional analysis, the necessary technical instruments are developed to establish, for example, the weak solvability of linear elliptic partial differential equations.

As a typical motivating example serves the celebrated *Poisson*[1] *equation*: The Poisson problem can be used to model the deflection of a (linear) elastic membrane (made of an ideal material) fixed at the boundary of a domain under the influence of an external force. More precisely, the Poisson equation seeks for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, over which the membrane is spanned and to which (topological) boundary $\partial\Omega$ the membrane is fixed to, and a given external force $f\colon \Omega \to \mathbb{R}$ acting on the membrane (*e.g.*, gravity), for a *deflection field* $u\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega\,, \\ u &= 0 \qquad \text{on } \partial\Omega\,. \end{aligned} \tag{0.1}$$

In the system (0.1), the equation $(0.1)_1$ describes the deflection of the membrane influenced by the external force, while the equation $(0.1)_2$ describes that the membrane is fixed at the boundary of the domain (*cf.* Figure 1).

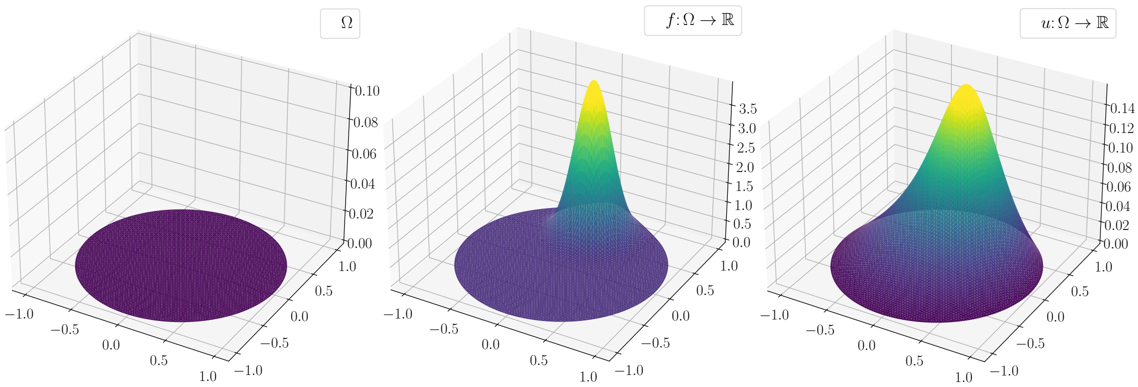

Figure 1: *left:* domain $\Omega := B_1^2(0) = \{x \in \mathbb{R}^2 \mid |x| < 1\}$; *middle:* external force $f\colon \Omega \to \mathbb{R}$, for every $x = (x_1, x_2)^\top \in \Omega$, defined by $f(x) := 4\exp\big(-\beta^2(x_1^2 + (x_2 - R_0)^2)\big)$, where $\beta = 4$ and $R_0 = 0.6$; *right:* deflection $u\colon \Omega \to \mathbb{R}$ approximated using element-wise affine finite elements.

[1]Simeon Denis Poisson, born 1781 in Pithiviers, died 1840 in Sceaux near Paris. After studying at the Ecole Polytechnique in Paris, Poisson taught there and later succeeded Fourier. Poisson was a mathematician, physicist and astronomer.

Another famous motivating example is given via the *Stokes*[2] *equations*: The Stokes equations can be used to model the laminar (*i.e.*, non-turbulent, *cf.* Figure 5) flow of a *Newtonian fluid* (*e.g.*, water) through a bounded domain under the influence of an external force. More precisely, the Stokes equations seek for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, occupied by the fluid and a given external force $\mathbf{f}\colon \Omega \to \mathbb{R}^d$ acting on the fluid, for a *velocity vector field* $\mathbf{v} := (v_1, \dots, v_d)^\top\colon \Omega \to \mathbb{R}^d$ and a *kinematic pressure* $\pi\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\Delta \mathbf{v} + \nabla \pi &= \mathbf{f} &&\text{in } \Omega\,,\\ \operatorname{div} \mathbf{v} &= 0 &&\text{in } \Omega\,,\\ \mathbf{v} &= \mathbf{0} &&\text{on } \partial\Omega\,. \end{aligned} \tag{0.2}$$

In the system (0.2), the equation $(0.2)_1$ describes the stress acting on the fluid influenced by the external force, the equation $(0.1)_2$ describes that the fluid is incompressible, while the equation $(0.1)_3$, typically called *no-slip boundary condition*, describes that the fluid is not moving at the boundary of the domain (*cf.* Figure 2).

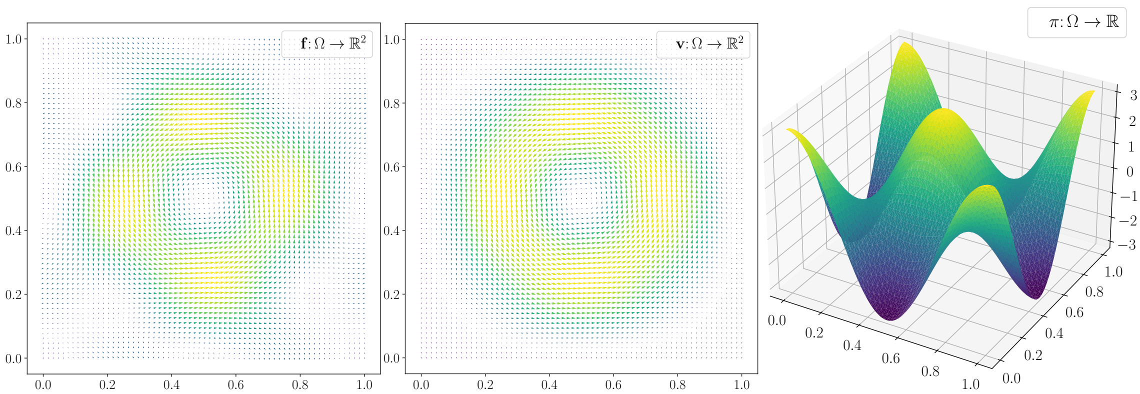

Figure 2: *left:* external force $\mathbf{f}\colon \Omega \to \mathbb{R}^2$, for every $x = (x_1, x_2)^\top \in \Omega := [0,1]^2$, defined by

$$\mathbf{f}(x) := \begin{pmatrix} -4\pi^2(2\cos(2\pi x_1) - 1)\sin(2\pi x_2) - 2\pi^2 \sin(2\pi x_1)\cos(2\pi x_2) \\ 4\pi^2(2\cos(2\pi x_2) - 1)\sin(2\pi x_1) - 2\pi^2 \sin(2\pi x_2)\cos(2\pi x_1) \end{pmatrix};$$

middle: velocity vector field $\mathbf{v}\colon \Omega \to \mathbb{R}^2$, for every $x = (x_1, x_2)^\top \in \Omega := [0,1]^2$, defined by

$$\mathbf{v}(x) := \begin{pmatrix} \sin(2\pi x_2)(1 - \cos(2\pi x_1)) \\ -\sin(2\pi x_1)(1 - \cos(2\pi x_2)) \end{pmatrix};$$

right: kinematic pressure $\pi\colon \Omega \to \mathbb{R}$, for every $x = (x_1, x_2)^\top \in \Omega := [0,1]^2$, defined by

$$\pi(x) := \pi \cos(2\pi x_1)\cos(2\pi x_2)\,.$$

[2]Sir George Gabriel Stokes, born 1819 in Skreen, died 1903 in Cambridge. Stokes was a professor of mathematics at Cambridge and his research interests included analysis, fluorescence and geodesy.

The existence of weak solutions to both the Poisson equation (*cf.* (0.1)) and the Stokes equations (*cf.* (0.2)) can be established resorting to the Lax–Milgram lemma[3].

Lemma 0.3 (Lax[4]–Milgram[5], 1954)
Let $(X, (\cdot,\cdot)_X)$ be a (real) Hilbert space. Moreover, let $A\colon X \to X^*$ be a linear operator with the following properties:

(i) Boundedness: *There exists a constant $\alpha > 0$ such that for every $x \in X$, it holds that*

$$\|Ax\|_{X^*} \leq \alpha \, \|x\|_X \,;$$

(ii) Strong positivity: *There exists a constant $\beta > 0$ such that for every $x \in X$, it holds that*

$$\langle Ax, x\rangle_X \geq \beta \, \|x\|_X^2 \,.$$

Then, $A\colon X \to X^*$ is an isomorphism, i.e., $A\colon X \to X^*$ is bijective and its inverse $A^{-1}\colon X^* \to X$ is continuos.

Proof. See [Emm04, Satz 3.4.6]. □

However, the majority of phenomena observable in nature cannot be described properly via linear mathematical models. Instead, we should focus more on non-linear mathematical models:

As a prototypical motivating example serves the *p-Laplace equation*: The p-Laplace equation can be used to model the deflection of a (non-linear) elastic membrane (made of a non-ideal material) fixed at the boundary of a domain under the influence of an external force. More precisely, the p-Laplace equation seeks for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, over which the membrane is spanned and to which (topological) boundary $\partial\Omega$ the membrane is fixed to, a given external force $f\colon \Omega \to \mathbb{R}$ acting on the membrane (*e.g.*, gravity), and a material parameter $p \in (1, +\infty)$ describing the non-linear elastic properties of the membrane, for a deflection field $u\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\mathrm{div}(|\nabla u|^{p-2}\nabla u) &= f && \text{in } \Omega\,, \\ u &= 0 && \text{on } \partial\Omega\,. \end{aligned} \tag{0.4}$$

In the system (0.4), again, the equation $(0.4)_1$ describes the deflection of the membrane influenced by the force, while the equation $(0.4)_2$ describes that the membrane is fixed at the boundary of the domain (*cf.* Figure 3).

[3] More precisely, for the Stokes equations (*cf.* (0.1)), the celebrated Lax–Milgram lemma (*cf.* Lemma 0.3) can only be employed to prove the existence of a velocity vector field solving the Stokes equations (*cf.* (0.1)) in a hydromechanical sense. Then, one can recover the pressure on the basis of the de Rham lemma.

[4] Peter David Lax, born in Hungary in 1926. Lax received his doctorate under Friedrichs and is a professor at the Courant Institute of Mathematical Sciences in New York. He works on the analysis and numerical analysis of partial differential equations and is particularly concerned with the basic equations of current mechanics. Lax was President of the American Mathematical Society from 1979 to 1980.

[5] Arthur Norton Milgram, born 1912, died 1961 in Minnesota. Milgram received his doctorate in 1937 in Pennsylvania and was a professor at Notre Dame, Princeton, Syracuse and Minnesota. He worked on questions of topology, algebra and functional analysis. Milgram is also known for publishing a classic on Galois theory based on lectures by Artin.

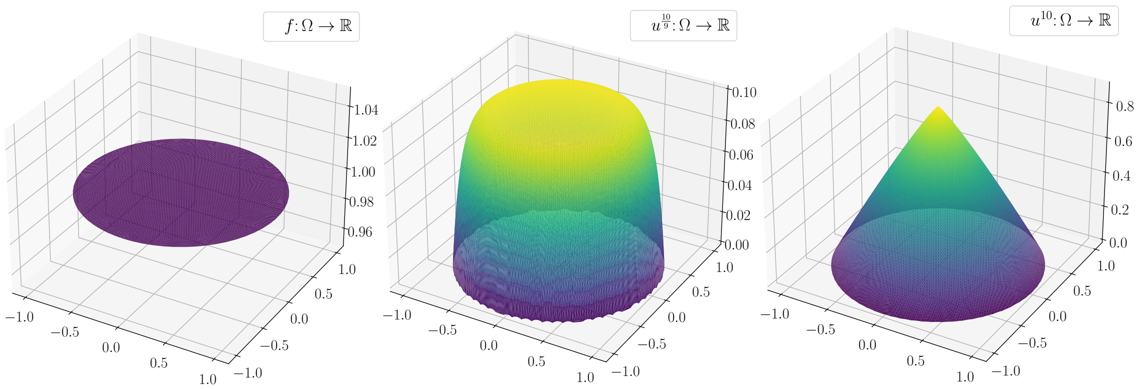

Figure 3: *left:* force $f\colon\Omega\to\mathbb{R}$, for every $x\in\Omega$, defined by $f(x)\coloneqq 4\exp\big(-(\beta^2)(x_1^2+(x_1-R_0)^2)\big)$, where $\Omega=B_1^2(0)$; *middle:* deflection $u^{\frac{10}{9}}\colon\Omega\to\mathbb{R}$, for every $x\in\Omega$, defined by $u^{\frac{10}{9}}(x)\coloneqq\frac{1}{10}(1-|\cdot|^{10})$; *right:* deflection $u^{10}\colon\Omega\to\mathbb{R}$, for every $x\in\Omega$, defined by $u^{10}(x)\coloneqq\frac{9}{10}(1-|\cdot|^{\frac{10}{9}})$. More generally, for every $p\in(1,+\infty)$, the deflection $u^p\colon\Omega\to\mathbb{R}$, for every $x\in\Omega$, defined by $u^p(x)\coloneqq\frac{1}{p'}(1-|\cdot|^{p'})$, where $p'\coloneqq\frac{p}{p-1}$, is a solution of the p-Laplace equation (0.4) with $\Omega=B_1^2(0)$ and $f\coloneqq 1$.

Another famous motivating example is given via the *p-Stokes equations*: The p-Stokes equations can be used to model the laminar (*i.e.*, non-turbulent) flow of a *non-Newtonian fluid* through a bounded domain under the influence of an external force. More precisely, the p-Stokes equations seek for a given bounded domain $\Omega\subseteq\mathbb{R}^d$, $d\geq 2$, occupied by the fluid and a given external force $\mathbf{f}\colon\Omega\to\mathbb{R}^d$ acting on the fluid, for a velocity vector field $\mathbf{v}\coloneqq(v_1,\dots,v_d)^\top\colon\Omega\to\mathbb{R}^d$ and a kinematic pressure $\pi\colon\Omega\to\mathbb{R}$ such that

$$\begin{aligned} -\operatorname{div}(\mathbf{S}(\mathbf{D}\mathbf{v}))+\nabla\pi&=\mathbf{f} &&\text{in }\Omega\,,\\ \operatorname{div}\mathbf{v}&=0 &&\text{in }\Omega\,,\\ \mathbf{v}&=\mathbf{0} &&\text{on }\partial\Omega\,. \end{aligned}\tag{0.5}$$

In the system (0.5), the *extra-stress tensor* $\mathbf{S}(\mathbf{D}\mathbf{v})\colon\Omega\to\mathbb{R}^{d\times d}_{\mathrm{sym}}$ depends on the *strain-rate tensor* $\mathbf{D}\mathbf{v}\coloneqq\frac{1}{2}(\nabla\mathbf{v}+\nabla\mathbf{v})^\top\colon\Omega\to\mathbb{R}^{d\times d}_{\mathrm{sym}}$, *i.e.*, the symmetric part of the velocity gradient $\nabla\mathbf{v}\coloneqq(\partial_j v_i)_{i,j\in\{1,\dots,d\}}\colon\Omega\to\mathbb{R}^{d\times d}$. In the case of a so-called *generalized Newtonian fluids*, which is a sub-class of the class of non-Newtonian fluids, the non-linearity $\mathbf{S}\colon\mathbb{R}^{d\times d}_{\mathrm{sym}}\to\mathbb{R}^{d\times d}_{\mathrm{sym}}$, describing extra-stress tensor, admits the form

$$\mathbf{S}(\mathbf{D}\mathbf{v})\coloneqq\nu(|\mathbf{D}\mathbf{v}|)\mathbf{D}\mathbf{v}\quad\text{ in }\Omega\,,\tag{0.6}$$

where the *(generalized) viscosity* $\nu(|\mathbf{D}\mathbf{v}|)\colon\Omega\to[0,+\infty)$ is a function from the *shear-rate* $|\mathbf{D}\mathbf{v}|\colon\Omega\to[0,+\infty)$, *i.e.*, the modulus of the strain-rate tensor. The (generalized) viscosity can be seen as a measure for the resistance of the fluid against deformation at a given shear-rate.

In this lecture, we will mainly be engaged with a sub-class of generalized Newtonian fluids, namely so-called *power-law fluids*, in which the (generalized) viscosity depends on the shear-rate $|\mathbf{D}\mathbf{v}|$ in a power function relation. More precisely, for a power-law fluid, the (generalized) viscosity admits the form

$$\nu(|\mathbf{D}\mathbf{v}|) \coloneqq \nu_0\, (\delta + |\mathbf{D}\mathbf{v}|)^{p-2} \quad \text{in } \Omega\,, \tag{0.7}$$

where $\nu_0 > 0$, $\delta \geq 0$, and $p \in (1, +\infty)$ is the so-called *power-law index*. Power-law fluids, depending on the value of the power-law index, may be categorized into three classes:

- ($p = 2$) The *Newtonian* case describes fluids (*e.g.*, water), where the viscosity does not change when the shear-rate increases;
- ($p > 2$) The *shear-thickening* (or dilatant) case describes fluids (*e.g.*, a mixture of cornstarch and water, sometimes called oobleck), where the viscosity increases when the shear-rate increases (*e.g*, the fluid behaves thicker);
- ($p < 2$) The *shear-thinning* (or pseudo-plastic) case describes fluids (*e.g.*, blood and paint), where the viscosity decreases when the shear-rate increases (*e.g*, the fluid behaves thinner).

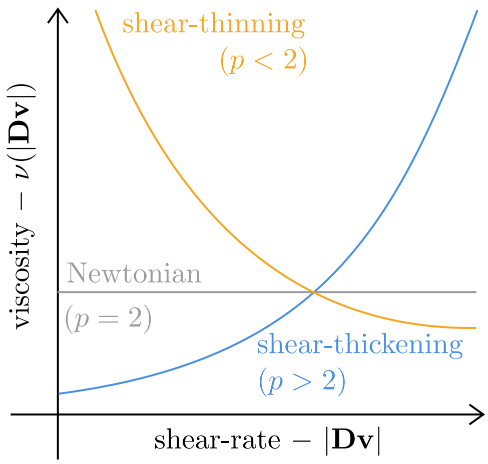

Figure 4: Classification of power-law fluids.

The existence of weak solutions to both the p-Laplace equation (*cf.* (0.4)) and the p-Stokes equations (*cf.* (0.5)), since both problems are non-linear, cannot any longer be established resorting to the Lax–Milgram lemma (*cf.* Lemma 0.3). The super-ordinate objective of this lecture is to generalize the methods (from linear functional analysis) developed for solving linear elliptic partial differential equations to methods (from non-linear functional analysis) for solving a large class of non-linear partial differential equations, including both the p-Laplace equation (*cf.* (0.4)) and the p-Stokes equations (*cf.* (0.5)). More precisely, the ultimate objective of this lecture is to establish the existence of weak solutions to the *p-Navier[6]–Stokes equations*: The p-Navier–Stokes equations can be used to model the laminar and turbulent flow of a *non-Newtonian fluid* through a bounded domain.

[6] Claude Louis Marie Henri Navier, born 1785 in Dijon, died 1836 in Paris. Navier was an engineer and professor at the Ecole des Ponts et Chaussees in Paris, which was founded in 1747 and still exists today. From 1831 he succeeded Cauchy as professor of analysis and mechanics at the Ecole Polytechnique, which was founded in 1794. He dealt with questions of elasticity theory, hydraulics and bridge construction.

More precisely, the p-Navier–Stokes equations seek for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, occupied by the fluid and a given external force $\mathbf{f}\colon \Omega \to \mathbb{R}^d$ acting on the fluid, for a velocity vector field $\mathbf{v} \coloneqq (v_1, \ldots, v_d)^\top \colon \Omega \to \mathbb{R}^d$ and a kinematic pressure $\pi\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\operatorname{div}(\mathbf{S}(\mathbf{D}\mathbf{v})) + \operatorname{div}(\mathbf{v} \otimes \mathbf{v}) + \nabla \pi &= \mathbf{f} && \text{in } \Omega\,, \\ \operatorname{div} \mathbf{v} &= 0 && \text{in } \Omega\,, \\ \mathbf{v} &= \mathbf{0} && \text{on } \partial\Omega\,. \end{aligned} \tag{0.8}$$

In the system (0.8), the extra-stress tensor $\mathbf{S}(\mathbf{D}\mathbf{v})\colon \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is defined via (0.7) and the *convective term* $\operatorname{div}(\mathbf{v} \otimes \mathbf{v}) = (\sum_{j=1}^d \partial_j(v_i v_j))_{i=1,\ldots,d}\colon \Omega \to \mathbb{R}^d$ models turbulences in the fluid (*cf.* Figure 5), which cannot be modelled by the p-Stokes equations (*cf.* (0.5)).

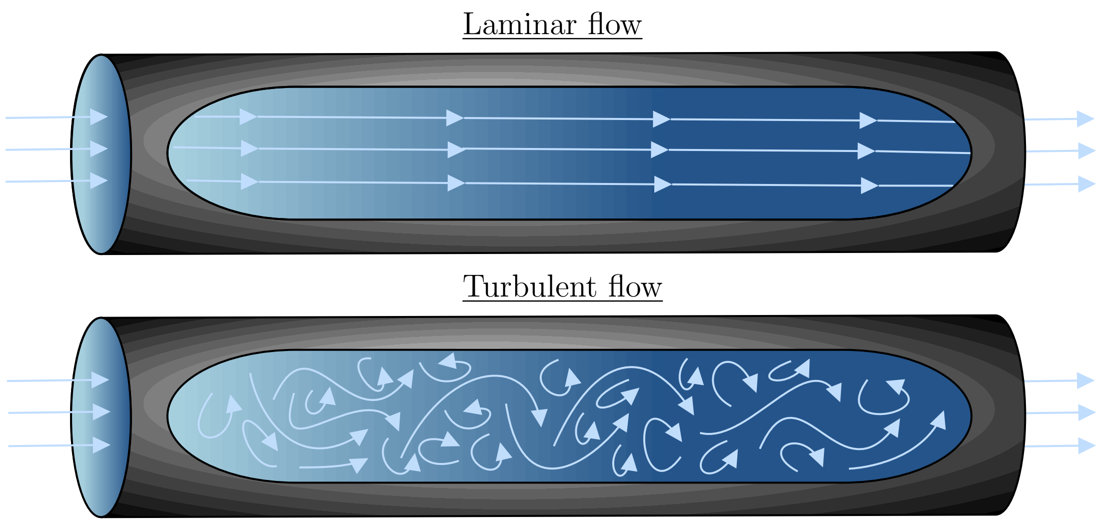

Figure 5: *top:* Laminar flow through a pipe; *bottom:* Turbulent flow through a pipe.

The ultimate objective of this lecture is to develop the mathematical framework that allows us to establish the weak solvability of the p-Navier–Stokes equations.

1 Linear Functional Analysis

1.1 Dual Spaces

In linear algebra, the *algebraic dual space* of a vector space X is defined by

$$X^{\#} := \left\{ x^*\colon X \to \mathbb{R} \;\middle|\; \begin{array}{r} \langle x^{\#}, \lambda x + \mu y\rangle_X = \lambda \langle x^{\#}, x\rangle_X + \mu \langle x^{\#}, y\rangle_X \\ \text{for all } x, y \in X\,,\ \ \lambda, \mu \in \mathbb{R} \end{array} \right\},$$

where the so-called *duality product* $\langle \cdot, \cdot \rangle_X \colon X^{\#} \times X \to \mathbb{R}$, for every $x^{\#} \in X^{\#}$ and $x \in X$, is defined by

$$\langle x^{\#}, x\rangle_X := x^{\#}(x)\,.$$

If X is a normed vector space, we can introduce the *topological dual space* X^*; a subspace of $X^{\#}$, which, then, forms a normed Banach space as well.

Definition 1.1 (dual space)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the topological dual space *of X is defined by*

$$X^* := \left\{ x^{\#} \in X^{\#} \mid \textit{there exists } M > 0 \textit{ such that } \langle x^{\#}, x\rangle_X \le M\, \|x\|_X \textit{ for all } x \in X \right\}.$$

In fact, the topological dual space of a normed vector space equipped with the so-called operator norm does not only form a normed vector space but even a Banach space, *i.e.*, a complete normed vector space (*cf.* Figure 1.1).

Proposition 1.2 (completeness of X^*)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the topological dual space X^* of X, equipped with the operator norm

$$\begin{aligned} \|x^*\|_{X^*} &= \sup_{x \in X \setminus \{0\}} \frac{|\langle x^*, x\rangle_X|}{\|x\|_X} \\ &= \sup_{x \in X \,:\, \|x\|_X \le 1} |\langle x^*, x\rangle_X| \\ &= \sup_{x \in X \,:\, \|x\|_X = 1} |\langle x^*, x\rangle_X|\,, \end{aligned}$$

forms a Banach space.

Proof. Completeness follows from completeness of $\mathbb{R}$ (Exercise). □

Apparently, in any case, we have that $X^* \subseteq X^\#$. Depending on the dimension of the pre-dual space X, the relation between algebraic dual space $X^\#$ and topological dual space X^* can be stated more explicitly.

Remark 1.3 (relation between $X^\#$ and X^*)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the following statements apply:

(i) If X is finite-dimensional, then

$$X^* = X^\# .$$

(ii) If X is a infinite-dimensional (real) normed vector space, then

$$X^\# \setminus X^* \neq \emptyset .$$

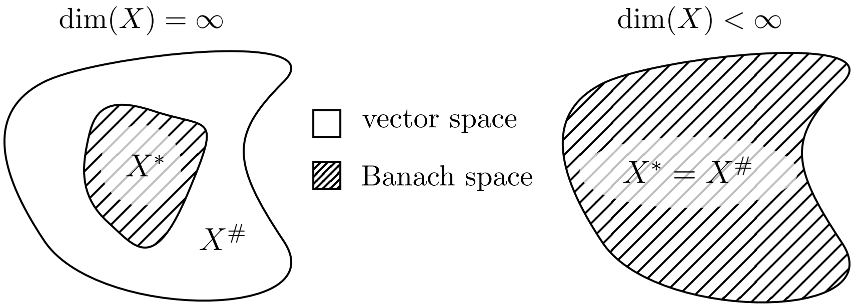

Figure 1.1: Relation between algebraic dual spaces $X^\#$ and topological dual spaces X^* depending on the dimension of the pre-dual space X.

1.2 Hahn–Banach Theorem

The Hahn–Banach theorem (*cf.* Theorem 1.4 (below)) is concerned with the extension of linear and continuous functionals defined on a linear subspace of the entire normed vector space – while maintaining the operator norm or related quantities; and, thus, is not only a mainstay of linear functional analysis, but also for non-linear functional analysis.

Theorem 1.4 (Hahn[1]–Banach[2], independent in 1926/1929)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space and Y a linear subspace (i.e., $(Y, \|\cdot\|_Y)$, where $\|y\|_Y \coloneqq \|y\|_X$ for all $y \in Y$, is also a (real) normed vector space). Then, for every $y^* \in Y^*$, there exists a norm-preserving extension $x^* \in X^*$ such that

$$\begin{aligned} \langle x^*, y\rangle_X &= \langle y^*, y\rangle_Y \qquad \text{for all } y \in Y\,,\\ \|x^*\|_{X^*} &= \|y^*\|_{Y^*}\,. \end{aligned}$$

Proof. See [Bré11, Thm. 1.1]. □

[1] Hans Hahn, born 1879 in Vienna, died 1934 ibid. After studying at the Vienna University of Technology, Hahn became a professor there in 1921. He dealt with functional analysis and calculus of variations.

[2] Stefan Banach, born 1892 in Krakow, died 1945 in Lviv. Banach was a mathematician in Lviv and founded the important Lviv school of functional analysis. See also Werner [Wer97, p. 41 ff.] and Saxe [Sax02] for some interesting remarks on Banach's life.

Even more famous and/or useful than the Hahn–Banach theorem (*cf.* Theorem 1.4), are its various corollaries, which will come in handy frequently in this lecture.

Corollary 1.5 (why we love the Hahn–Banach theorem…♡)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the following statements apply:

(i) For every $x \in X \setminus \{0\}$, there exists $x^* \in X^*$ such that

$$\langle x^*, x\rangle_X = \|x^*\|_{X^*}^2 = \|x\|_X^2 \,.$$

(ii) For every $x \in X$, it holds that

$$\begin{aligned}\|x\|_X &= \max_{x^* \in X^* \setminus \{0\}} \frac{|\langle x^*, x\rangle_X|}{\|x^*\|_{X^*}} \\ &= \max_{x^* \in X^* \,:\, \|x^*\|_{X^*} \le 1} |\langle x^*, x\rangle_X| \\ &= \max_{x^* \in X^* \,:\, \|x^*\|_{X^*} = 1} |\langle x^*, x\rangle_X| \,.\end{aligned}$$

(iii) Let $Y \subseteq X$ be a linear subspace with $\overline{Y}^{\|\cdot\|_X} \neq X$. Then, there exists $x^* \in X^*$ with $x^* \neq 0^*$ such that for every $y \in Y$, it holds that

$$\langle x^*, y\rangle_X = 0 \,.$$

In other words, if for every $x^* \in X^*$ from

$$\langle x^*, y\rangle_X = 0 \qquad \textit{for all } y \in Y \,,$$

it follows that $x^* = 0^*$, then $\overline{Y}^{\|\cdot\|_X} = X$.

(iv) If $x \in X$ is such that

$$\langle x^*, x\rangle_X = 0 \qquad \textit{for all } x^* \in X^* \,,$$

then $x = 0$ in X.

Proof. *ad (i).* Set $Y \coloneqq \mathbb{R}x$ and let $y^* \colon Y \to \mathbb{R}$, for every $\lambda \in \mathbb{R}$, be defined by

$$\langle y^*, \lambda x\rangle_Y \coloneqq \lambda \, \|x\|_X \,.$$

Then, we have that $y^* \in Y^*$, so that the Hahn–Banach theorem (*cf.* Theorem 1.4) yields a norm-preserving extension $x^* \in X^*$, *i.e.*, it holds that

$$\begin{aligned}\langle x^*, \lambda x\rangle_X &= \langle y^*, \lambda x\rangle_Y \qquad \text{for all } \lambda \in \mathbb{R}\,, \\ \|x^*\|_{X^*} &= \|y^*\|_{Y^*} \,.\end{aligned}$$

In particular, we have that

$$\begin{aligned}\langle x^*, \lambda x\rangle_X &= \langle y^*, 1\cdot x\rangle_Y = \|x\|_X\,,\\ \|x^*\|_{X^*} &= \|y^*\|_{Y^*} = \sup_{\lambda\in\mathbb{R}} \frac{\langle y^*, \lambda\, x\rangle_Y}{\|\lambda\, x\|_X} = 1\,.\end{aligned}$$

ad (iii). On the one hand, for every $x^* \in X^*$, we have that

$$|\langle x^*, x\rangle_X| \leq \|x^*\|_{X^*}\|x\|_X\,.$$

On the other hand, due to (i), there exists $x^* \in X^*$ with $\|x^*\|_{X^*} = 1$ and

$$\langle x^*, x\rangle_X = \|x\|_X\,.$$

ad (iv). See [Bré11, Cor. 1.8].
ad (v) Assume that $x \neq 0$. Then, due to (i), there exists $x^* \in X^*$ with

$$\langle x^*, x\rangle_X = \|x\|_X > 0\,,$$

which is a contradiction. Therefore, we have that $x = 0$ in X. □

1.3 Bi-Dual Space and Reflexivity

One decisive ingredient for being able to generalize results from linear functional analysis to the non-linear case is to have a generalization of the Bolzano–Weierstraß theorem. The latter in infinite-dimensional normed vector spaces cannot be expected to be true with respect to the norm topology. However, switching to a coarser topology (*i.e.*, the weak topology) a generalization of the Bolzano–Weierstraß theorem for infinite-dimensional normed vector spaces can be established if these normed vector spaces (with respect to a particular isometry, *i.e.*, the canonical isometry) are isometrically isomorphic to the topological dual space of their topological dual space.

Definition 1.6 (bi-dual space and canonical isometry)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the (topological) bi-dual space X^{**} *of X is defined as the (topological) dual space of the (topological) dual space X^*, i.e.,*

$$X^{**} \coloneqq (X^*)^*\,.$$

In addition, the canonical isometry $j_X\colon X \to X^{**}$, *for every $x \in X$ and $x^* \in X^*$, is defined by*

$$\langle j_X(x), x^*\rangle_{X^*} \coloneqq \langle x^*, x\rangle_X\,.$$

Proposition 1.7 (properties of the canonical isometry)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the canonical isometry $j_X \colon X \to X^{**}$ is linear and isometrical, i.e., for every $x \in X$, it holds that

$$\|j_X(x)\|_{X^{**}} = \|x\|_X \,,$$

and, thus, injective.

Proof. Linearity is evident. The isometry property is a consequence of Corollary 1.5(ii). □

Proposition 1.7 let us interpret X as a subspace of its topological bi-dual space X^{**}. If X coincides with X^{**} (up to the canonical isometry j_X), then we use the following notion.

Definition 1.8 (reflexivity)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, X is called reflexive *if $j_X(X) = X^{**}$.*

Proposition 1.9 (things to know about reflexivity)
Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the following statements apply:
(i) If X is reflexive, then X is a Banach space.
(ii) If X is reflexive, then X and X^{**} are isometrically isomorphic.
The opposite direction, in general, is not correct.
(iii) X is reflexive if and only if X^* is reflexive.
(iv) If X is reflexive and separable, then X^* is reflexive and separable. More precisely, since from the separability of X^* it always follows that X is separable, we have that X is reflexive and separable if and only if X^* is reflexive and separable.

Proof. *ad (i).* Follows from Proposition 1.2, that $X^{**} = (X^*)^*$ (*cf.* Definition 1.6), and that $j_X \colon X \to X^{**}$ is an isometric isomorphism.

ad (ii). See [Jam51], *i.e.*, if we define for every $x = (x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, the norm

$$\|x\|_X \coloneqq \sup_{(p_n)_{n\in\mathbb{N}}\in\mathbb{N}^{\mathbb{N}}\,:\,p_i,p_j\neq 0 \text{ for some } i,j\in\mathbb{N}} \left\{ (x_{p_{n+1}} - x_{p_1})^2 + \sum_{i=1}^{n} (x_{p_i} - x_{p_{i+1}})^2 \right\} \in [0, +\infty]$$

the space

$$X \coloneqq \big\{ x = (x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \|x\|_X < \infty \,,\; x_n \to 0 \,(n \to \infty) \big\} \,.$$

ad (iii). See [Bré11, Cor. 3.21].

ad (iv). See [Bré11, Cor. 3.27]. □

1.4 Notions of Convergence

In this subsection, we discuss three important notions of convergence that we will use frequently in the course of the lecture. In particular, two of these notions of convergence induce coarser topologies with respect to a generalization of the Bolzano–Weierstraß theorem for infinite-dimensional normed vector spaces holds true.

Definition 1.10 (strong and weak convergence)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is said to

(i) converge strongly *to $x \in X$, written*

$$x_n \to x \quad \text{in } X \quad (n \to \infty)\,,$$

if

$$\|x_n - x\|_X \to 0 \quad (n \to \infty)\,.$$

(ii) converge weakly *to $x \in X$, written*

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,,$$

if for every $x^* \in X^*$, it holds that

$$\langle x^*, x_n - x\rangle_X \to 0 \quad (n \to \infty)\,.$$

Let us collect some useful facts about strong and weak convergence in the following proposition.

Proposition 1.11 (some useful facts about strong and weak convergence)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, the following statements apply:

(i) Weak and strong limits are unique.

(ii) Strong convergence implies weak convergence, i.e., for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and $x \in X$, from

$$x_n \to x \quad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,.$$

(iii) Weakly convergent sequences are bounded, i.e., for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and $x \in X$, from

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$\sup_{n\in\mathbb{N}} \|x_n\|_X < \infty\,.$$

(iv) The norm function $\|\cdot\|_X \colon X \to [0,+\infty)$ *is weakly lower semi-continuous, i.e., for a sequence* $(x_n)_{n\in\mathbb{N}} \subseteq X$ *and* $x \in X$*, from*

$$x_n \rightharpoonup x \quad \text{in } X \quad (n\to\infty)\,,$$

it follows that

$$\|x\|_X \leq \liminf_{n\to\infty} \|x_n\|_X\,.$$

(v) For sequences $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ *and* $(x_n)_{n\in\mathbb{N}} \subseteq X$ *from*

$$\begin{aligned} x_n^* &\to x^* && \text{in } X^* && (n\to\infty)\,,\\ x_n &\rightharpoonup x && \text{in } X && (n\to\infty)\,, \end{aligned}$$

where $x^* \in X^*$ *and* $x \in X$*, it follows that*

$$\langle x_n^*, x_n\rangle_X \to \langle x^*, x\rangle_X \quad (n\to\infty)\,.$$

(vi) If X *is uniformly convex (and, thus, reflexive), i.e., for every* $\varepsilon > 0$*, there exists some* $\delta > 0$ *such that for every* $x, y \in X$ *from*

$$\|x\|_X \leq 1\,,\ \|y\|_X \leq 1\,,\ \|x-y\|_X \geq \varepsilon\,,$$

it follows that

$$\|x+y\|_X \leq 2\,(1-\delta)\,,$$

then for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ *and* $x \in X$*, from*

$$\begin{aligned} x_n &\rightharpoonup x && \text{in } X && (n\to\infty)\,,\\ \|x_n\|_X &\to \|x\|_X && \text{in } \mathbb{R} && (n\to\infty)\,, \end{aligned}$$

it follows that

$$x_n \to x \quad \text{in } X \quad (n\to\infty)\,.$$

(vii) If $\dim(X) < \infty$*, then for a sequence* $(x_n)_{n\in\mathbb{N}} \subseteq X$ *and* $x \in X$*, it holds that*

$$x_n \rightharpoonup x \quad \text{in } X \quad (n\to\infty)\,,$$

if and only if

$$x_n \to x \quad \text{in } X \quad (n\to\infty)\,.$$

Proof. See [Bré11, Prop. 3.5, Prop. 3.6]. □

Definition 1.12 (weak-$*$ convergence)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, a sequence $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ is said to converge weakly-$*$- *to $x^* \in X^*$, written*

$$x_n^* \overset{*}{\rightharpoonup} x^* \quad \text{in } X^* \quad (n \to \infty)\,,$$

if for every $x \in X$, it holds that

$$\langle x_n^* - x^*, x\rangle_X \to 0 \quad (n \to \infty)\,.$$

Proposition 1.13 (some useful facts about weak-$*$ convergence)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, the following statements apply:

(i) Weak-$*$ limits are unique.

(ii) Weak convergence implies weak-$*$ convergence, i.e., for a sequence $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ and $x^* \in X^*$, from

$$x_n^* \rightharpoonup x^* \quad \text{in } X^* \quad (n \to \infty)\,,$$

it follows that

$$x_n^* \overset{*}{\rightharpoonup} x^* \quad \text{in } X^* \quad (n \to \infty)\,.$$

(iii) Weakly-$*$ convergent sequences are bounded, i.e., for a sequence $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ and $x^* \in X^*$, from

$$x_n^* \overset{*}{\rightharpoonup} x^* \quad \text{in } X^* \quad (n \to \infty)\,,$$

it follows that

$$\sup_{n\in\mathbb{N}} \|x_n^*\|_{X^*} < \infty\,.$$

(iv) The norm function $\|\cdot\|_{X^*} \colon X^* \to [0, +\infty)$ is weakly-$*$ lower semi-continuous, i.e., for a sequence $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ and $x^* \in X^*$, from

$$x_n^* \overset{*}{\rightharpoonup} x^* \quad \text{in } X^* \quad (n \to \infty)\,,$$

it follows that

$$\|x^*\|_{X^*} \le \liminf_{n\to\infty} \|x_n^*\|_{X^*}\,.$$

(v) For sequences $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ and $(x_n)_{n\in\mathbb{N}} \subseteq X$ from

$$\begin{aligned} x_n^* &\overset{*}{\rightharpoonup} x^* && \text{in } X^* && (n \to \infty)\,,\\ x_n &\to x && \text{in } X && (n \to \infty)\,, \end{aligned}$$

where $x^* \in X^*$ and $x \in X$, it follows that

$$\langle x_n^*, x_n\rangle_X \to \langle x^*, x\rangle_X \quad (n \to \infty)\,.$$

(vi) If X is reflexive, then weak convergence and weak-$*$ convergence coincide.

Proof. See [Bré11, Prop. 3.13]. □

Using the notions of weak convergence and reflexivity, we arrive at the following generalization of the Bolzano–Weierstraß theorem for infinite-dimensional Banach spaces.

Theorem 1.14 (Eberlein[3]–Šmulian[4], 1940 ($\Rightarrow$), 1947 ($\Leftarrow$))

Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, X is reflexive if and only if each bounded sequence in X has a weakly convergent subsequence.

Proof. See [Bré11, Thm. 3.19] and references therein. $\square$

Using the notions of weak-$*$ convergence, we arrive at the following generalization of the Bolzano–Weierstraß theorem for topological dual spaces of infinite-dimensional Banach spaces.

Theorem 1.15 (Banach–Alaoglu[5], 1934, 1940)

Let $(X, \|\cdot\|_X)$ be a (real) normed vector space. Then, the closed unit ball in the topological dual space, i.e., $\overline{B}_1^{X^*}(0) := \{x^* \in X^* \mid \|x^*\|_{X^*} \le 1\}$, is weakly-$*$ compact.

Proof. See [Bré11, Thm. 3.16]. $\square$

Note that the Banach–Alaoglu theorem (*cf.* Theorem 1.15) does not imply that the closed unit ball in the (topological) dual space is weakly-$*$ sequentially compact, unless we assume that X is separable.

Corollary 1.16

Let $(X, \|\cdot\|_X)$ be a separable normed vector space. Then, each bounded sequence in X^* has a weakly-$*$ convergent subsequence.

Proof. See [Bré11, Cor. 3.30]. $\square$

In the course of the lecture, we will resort countless times to the following lemma.

Lemma 1.17 (Subsequence convergence principle (SCP))

Let $(X, \|\cdot\|_X)$ be a reflexive Banach space. Then, if for a bounded sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ each weakly convergent subsequence has the same limit $x \in X$, it follows that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,.$$

The assertion still holds true if we replace weak convergence with strong convergence.

Proof. See Exercise Sheet 1, Exercise 3. $\square$

[3] William Frederick Eberlein, born 1917 in Shawano, Wisconsin, died 1986 in Rochester, New York. Eberlein studied from 1936 to 1942 at the University of Wisconsin and at Harvard University, where he received in 1942 his doctorate for the thesis *Closure, Convexity, and Linearity in Banach Spaces* under the direction of Marshall Stone.

[4] Witold Lwowitsch Schmulian, born in 1914 in Kherson, died 1944 in Warsaw. Shmulian obtained his first degree in mathematics at Odessa State University in 1936 and continued his studies under Mark Grigoryevich Krein, who awakened his interest in functional analysis, especially geometry in Banach spaces. He published 20 papers between 1937 and 1941.

[5] Leonidas Alaoglu was born in Red Deer, Alberta to Greek parents. He received his Bachelor of Science in 1936, Master of Science in 1937, and doctorate in 1938 (at the age of 24), all from the University of Chicago. His dissertation, written under the direction of Lawrence M. Graves, was on *Weak topologies of normed linear spaces* and establishes Banach–Alaoglu's theorem.

2 Monotone Operator Theory

The ultimate objective of this part of the lecture is to prove a generalization of the inverse function theorem for strictly monotone functions for infinite-dimensional Banach spaces.

Theorem 2.1 (inverse function theorem for strictly monotone functions)
Let $A\colon \mathbb{R} \to \mathbb{R}$ be a function (or an operator[1]*) with the following properties:*

(a) <u>Continuity:</u> *For a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ and an element $x \in \mathbb{R}$ from*

$$x_n \to x \quad \text{in } \mathbb{R} \quad (n \to \infty)\,,$$

it follows that

$$Ax_n \to Ax \quad \text{in } \mathbb{R} \quad (n \to \infty)\,.$$

(b) <u>Coercivity:</u> *It holds that $Ax \to \pm\infty$ $(x \to \pm\infty)$.*

Then, the following statements apply:

(i) $A\colon \mathbb{R} \to \mathbb{R}$ is surjective;

(ii) If, in addition, $A\colon \mathbb{R} \to \mathbb{R}$ is <u>strictly monotone</u>*, i.e., for every $x, y \in \mathbb{R}$ from $x < y$, it follows that*

$$Ax < Ay\,,$$

then $A\colon \mathbb{R} \to \mathbb{R}$ is bijective and its inverse $A^{-1}\colon \mathbb{R} \to \mathbb{R}$ is continuous, i.e., $A\colon \mathbb{R} \to \mathbb{R}$ is a homeomorphism.

Proof. <u>*ad (i).*</u> Direct consequence of the intermediate value theorem.

<u>*ad (ii).*</u>

<u>*1. Bijectivity:*</u> Since $A\colon \mathbb{R} \to \mathbb{R}$ is strictly monotone, it is also injective and, therefore, together with (i) bijective. As a result, the inverse $A^{-1}\colon \mathbb{R} \to \mathbb{R}$ exists.

<u>*2. Continuity of inverse:*</u> Consider a sequence $(y_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ and an element $y \in \mathbb{R}$ such that

$$y_n \to y \quad \text{in } \mathbb{R} \quad (n \to \infty)\,.$$

Then, the sequence $(x_n)_{n\in\mathbb{N}} \coloneqq (A^{-1}y_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ must be bounded. In fact, if the sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ would be unbounded, then there would exist a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq \mathbb{R}$ such that

$$x_{n_k} \to \pm\infty \quad \text{in } \mathbb{R} \quad (n \to \infty)\,,$$

[1]Typically, in this lecture, we call mappings between infinite-dimensional spaces (*e.g.*, function spaces) *'operators'*, while mappings between finite-dimensional spaces are called *'functions'*.

which, by the coercivity of $A\colon \mathbb{R}\to\mathbb{R}$, in turn, would imply that

$$y_{n_k} = Ax_{n_k} \to \pm\infty \quad \text{in } \mathbb{R} \quad (k\to\infty)\,,$$

which would contradict to convergence of the sequence $(y_n)_{n\in\mathbb{N}}\subseteq\mathbb{R}$. As a consequence, the sequence $(x_n)_{n\in\mathbb{N}}\subseteq\mathbb{R}$ is bounded and the Bolzano–Weierstraß theorem yields a subsequence $(x_{n_k})_{k\in\mathbb{N}}\subseteq\mathbb{R}$ and an element $\widetilde{x}\in\mathbb{R}$ such that

$$x_{n_k}\to\widetilde{x} \quad \text{in } \mathbb{R} \quad (k\to\infty)\,.$$

From the continuity of $A\colon\mathbb{R}\to\mathbb{R}$, we infer that

$$Ax_{n_k}\to A\widetilde{x} \quad \text{in } \mathbb{R} \quad (k\to\infty)\,.$$

Then, by the uniquess of limits, we infer that $A\widetilde{x}=y$ in $\mathbb{R}$, *i.e.*, $\widetilde{x}=A^{-1}y$ in $\mathbb{R}$. Since this argumentation may be repeated for each converging subsequence of $(x_n)_{n\in\mathbb{N}}\subseteq\mathbb{R}$, $A^{-1}y\in\mathbb{R}$ is the sole accumulation point of the sequence $(x_n)_{n\in\mathbb{N}}\subseteq\mathbb{R}$. As a result, the subsequence convergence principle (*cf.* Lemma 1.17) yields that

$$A^{-1}y_n\to A^{-1}y \quad \text{in } \mathbb{R} \quad (n\to\infty)\,,$$

where we also exploit that in finite-dimensional space weak and strong convergence coincide (*cf.* Proposition 1.11(vii)). In other words, the inverse $A^{-1}\colon\mathbb{R}\to\mathbb{R}$ is continuous. □

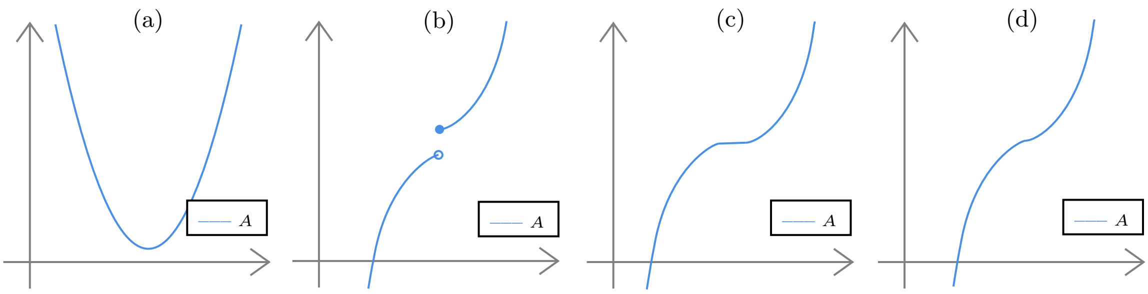

Figure 2.1: *(a):* $A\colon\mathbb{R}\to\mathbb{R}$ is continuous, but not coercive, and, thus, not surjective;
(b): $A\colon\mathbb{R}\to\mathbb{R}$ is coercive, but not continuous, and, thus, not surjective;
(c): $A\colon\mathbb{R}\to\mathbb{R}$ is continuous and coercive and, thus, surjective, but not strictly monotone, and, thus, not bijective;
(d): $A\colon\mathbb{R}\to\mathbb{R}$ is continuous, coercive, and strictly monotone, and, thus, bijective.

In order to prove an infinite-dimensional analogue of the inverse function theorem for strictly monotone functions (*cf.* Theorem 2.1), let us start with generalizing the properties of the function (or the operator) $A\colon\mathbb{R}\to\mathbb{R}$ assumed in Theorem 2.1, *i.e.*, let us generalize the notions continuity, coercivity, and (strict) monotonicity.

2.1 Notions of continuity and boundedness

As we will later rarely encounter strongly converging sequences, instead of working with the classical notion of continuity, we introduce generalized notions of continuity that are either tailored to weak convergence or turn out to harmonize with generalized concepts of monotonicity given only weak convergence.

Definition 2.2 (Notions of continuity and boundedness)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, an operator $A\colon X \to X^*$ is called

(i) continuous *if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from*

$$x_n \to x \qquad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$Ax_n \to Ax \quad \text{in } X^* \quad (n \to \infty)\,.$$

(ii) weakly continuous *if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from*

$$x_n \rightharpoonup x \qquad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$Ax_n \rightharpoonup Ax \quad \text{in } X^* \quad (n \to \infty)\,.$$

(iii) demi-continuous *if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from*

$$x_n \to x \qquad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$Ax_n \rightharpoonup Ax \quad \text{in } X^* \quad (n \to \infty)\,.$$

(iv) strongly continuous *if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from*

$$x_n \rightharpoonup x \qquad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$Ax_n \to Ax \quad \text{in } X^* \quad (n \to \infty)\,.$$

(v) hemi-continuous *if for every $x, y, z \in X$, the function*

$$(t \mapsto \langle A(x+ty), z\rangle_X)\colon [0,1] \to \mathbb{R}$$

is continuous.

(vi) radially continuous *if for every $x, y \in X$, the function*

$$(t \mapsto \langle A(x+ty), y\rangle_X)\colon [0,1] \to \mathbb{R}$$

is continuous.

(vii) locally bounded *if for every $x \in X$, there exist constants $\varepsilon = \varepsilon(x), M = M(x) > 0$ such that*

$$\|Ay\|_{X^*} \leq M \quad \text{for all } y \in \overline{B}_\varepsilon^X(x)\,,$$

where $\overline{B}_\varepsilon^X(x) \coloneqq \{y \in X \mid \|y - x\|_X \leq \varepsilon\}$.

Remark 2.3 (as a reminder...)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, an operator $A\colon X \to X^*$ is called

(i) bounded *if bounded sets in X are mapped into bounded sets in X^*.*
(ii) compact *if it is continuous and bounded sets in X are mapped into relatively compact sets in X^*.*

Lemma 2.4 (relations between notions of continuity and boundedness)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space and $A\colon X \to X^*$ an operator. Then, the following statements apply:

(i) The following implications apply:

$$\begin{tikzcd}[column sep=tiny]
& \textit{strongly continuous} \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
\textit{weakly continuous} \arrow[dr, Rightarrow] & & \textit{continuous} \arrow[dl, Rightarrow] \\
& \textit{demi-continuous} \arrow[d, Rightarrow] & \\
& \textit{hemi-continuous} \arrow[d, Rightarrow] & \\
& \textit{radially continuous} &
\end{tikzcd}$$

(ii) If $A\colon X \to X^*$ is strongly continuous and X reflexive, then $A\colon X \to X^*$ is compact.
(iii) If $A\colon X \to X^*$ is demi-continuous, then $A\colon X \to X^*$ is locally bounded.
(iv) If $A\colon X \to X^*$ is linear and compact, then $A\colon X \to X^*$ is strongly continuous.

Proof. *ad (i).* Direct consequence of the respective definitions (*cf.* Definition 2.2).

ad (ii).
1. Continuity: Due to (i), $A\colon X \to X^*$ is continuous.
2. Relative compact images of bounded sets: Let $M \subseteq X$ be a bounded set. We aim to show that $A(M)$ is relatively compact in X^*. To this end, consider an arbitrary sequence $(y_n)_{n\in\mathbb{N}} \subseteq A(M)$. Then, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq M$ such that

$$y_n = Ax_n \quad \text{in } X^* \quad \text{for all } n \in \mathbb{N}\,.$$

Due to the boundedness of $M \subseteq X$, the sequence $(x_n)_{n\in\mathbb{N}} \subseteq M$ is bounded as well. Hence, by the Eberlein–Šmuljian theorem (*cf.* Theorem 1.14), where we exploit that X is reflexive, there exists a subsequence $(x_{n_k})_{n\in\mathbb{N}} \subseteq M$ and an element $x \in X$ such that

$$x_{n_k} \rightharpoonup x \quad \text{in } X \quad (k \to \infty)\,.$$

Since $A\colon X \to X^*$ is strongly continuous, we infer that

$$y_{n_k} = Ax_{n_k} \to Ax \quad \text{in } X \quad (k \to \infty)\,.$$

In other words, every sequence in $A(M)$ has a strongly convergent subsequence in X^*, *i.e.*, $A(M)$ is relatively compact in X^*. In summary, we have shown that $A\colon X \to X^*$ is compact.

ad (iii). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be sequence and $x \in X$ an element such that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,.$$

Since $A\colon X \to X^*$ is compact, by Remark 2.3(ii), it is also continuous. Since linear and continuous operators are weakly continuous (*cf.* Exercise Sheet 2, Exercise 1), we deduce that

$$Ax_n \rightharpoonup Ax \quad \text{in } X^* \quad (n \to \infty)\,.$$

Due to Proposition 1.11(iii), the sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is bounded. Since $A\colon X \to X^*$ is compact, there exists a subsequence $(Ax_{n_k})_{k\in\mathbb{N}} \subseteq X^*$ and an element $x^* \in X^*$ such that

$$Ax_{n_k} \to x^* \quad \text{in } X^* \quad (k \to \infty)\,.$$

Since strong convergence implies weak convergence (*cf.* Proposition 1.11(ii)) and since weak limits are unique (*cf.* Proposition 1.11(i)), we find that $x^* = Ax$ in X^*. Since this argument may be repeated for any subsequence of $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$, $Ax \in X^*$ is the sole weak accumulation point of the sequence $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$. Thus, the subsequence convergence principle (*cf.* Lemma 1.17), which also applies for strong convergence, yields that

$$Ax_n \to Ax \quad \text{in } X^* \quad (n \to \infty)\,.$$

In other words, we have shown that $A\colon X \to X^*$ is strongly continuous.

ad (iv). Suppose that $A\colon X \to X^*$ is not locally bounded. Then, there exists an element $x \in X$ and a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ such that

$$\begin{aligned} x_n \to x \quad &\text{in } X \qquad &(n \to \infty)\,,\\ \|Ax_n\|_{X^*} \to +\infty & &(n \to \infty)\,, \end{aligned}$$

i.e., the sequence $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$ is unbounded. Since $A\colon X \to X^*$ is demi-continuous from the first convergence, we obtain

$$Ax_n \rightharpoonup Ax \quad \text{in } X^* \quad (n \to \infty)\,,$$

which contradicts the unboundedness of the sequence $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$. In other words, we have shown that $A\colon X \to X^*$ is locally bounded. $\square$

2.2 Notions of monotonicity

It is typically less obvious how to generalize strict monotonicity to infinite-dimensional Banachs spaces, as it is a concept that seems to rely on fact that $(\mathbb{R}, <)$ is a strictly order set. For a general infinite-dimensional Banach space (or even for $\mathbb{R}^d$, $d \in \mathbb{N}$) such a strict order is, in general, not available. A remedy is, then, provided by the observation that strict monotonicity can equivalently be expressed without directly using the strictly ordering of $\mathbb{R}$. More precisely, it is an exercise (*cf.* Homework 2, Problem 1) to verify that a function (or an operator) $A\colon \mathbb{R} \to \mathbb{R}$ is strictly monotone if and only if for every $x, y \in \mathbb{R}$ with $x \neq y$, it holds that

$$(Ax - Ay) \cdot (x - y) > 0\,.$$

We observe that using products we can avoid using the strict ordering of $\mathbb{R}$. And this observation helps us to generalize strict monotonicity to infinite-dimensional Banach spaces, since in Subsection 1.1 we recalled that on each Banach space X there exists a product, not with the Banach space X itself, but with its topological dual space X^*, *i.e.*, the duality product $\langle \cdot, \cdot \rangle_X \colon X^* \times X \to \mathbb{R}$. Therefore, if we restrict to operators that map a Banach space X into its topological dual space X^*, *i.e.*, operators of type $A\colon X \to X^*$, then we are in the position to generalize strict monotonicity to infinite-dimensional Banach spaces.

Definition 2.5 (Notions of monotonicity)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, an operator $A\colon X \to X^*$ is called

(i) monotone *if for every $x, y \in X$, it holds that*

$$\langle Ax - Ay, x - y \rangle_X \geq 0\,.$$

(ii) strictly monotone *if for every $x, y \in X$ with $x \neq y$, it holds that*

$$\langle Ax - Ay, x - x \rangle_X > 0\,.$$

(iii) d-monotone *if there exists a strictly non-decreasing function $\alpha\colon [0, +\infty) \to \mathbb{R}$ such that for every $x, y \in X$, it holds that*

$$\langle Ax - Ay, x - y \rangle_X \geq \big(\alpha(\|x\|_X) - \alpha(\|y\|_X)\big)(\|x\|_X - \|y\|_X)\,.$$

(iv) uniformly monotone *if there exists a strictly non-decreasing function $\rho\colon [0, +\infty) \to \mathbb{R}$ with $\rho(0) = 0$ such that for every $x, y \in X$, it holds that*

$$\langle Ax - Ay, x - y \rangle_X \geq \rho(\|x - y\|_X)\,.$$

(v) strongly monotone *if there exists a constant $m > 0$ such that for every $x, y \in X$, it holds that*

$$\langle Ax - Ay, x - y \rangle_X \geq m\, \|x - y\|_X^2\,.$$

Lemma 2.6 (relations between notions of monotonicity)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, the following implications apply:

$$
\begin{tikzcd}
 & \textit{strongly monotone} \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
\textit{d-monotone} \arrow[dr, Rightarrow, "X \textit{ strictly convex}"] & & \textit{uniformly monotone} \arrow[dl, Rightarrow] \\
 & \textit{strictly monotone} \arrow[d, Rightarrow] & \\
 & \textit{monotone} &
\end{tikzcd}
$$

Proof. For the implication '*d-monotone (& X strictly convex)* $\Rightarrow$ *strictly monotone*, see Exercise Sheet 2, Exercise 4. The remaining implications are direct consequences of the definitions (*cf.* Definition 2.5). $\square$

As prototypical examples for operators meeting the notions of monotonicity in Definition 2.5, in this lecture, serve operators with so-called *p-structure*.

Example 2.7 (an operator with p-structure)
For $p \in (1, +\infty)$, let the operator $A_p \colon \mathbb{R} \to \mathbb{R}$, for every $x \in \mathbb{R}$, be defined by

$$A_p x := |x|^{p-2} x \,.$$

Then, the following statements apply (cf. Figure 2.2):

(i) $A_p \colon \mathbb{R} \to \mathbb{R}$ *is continuous (since* $|A_p x| = |x|^{p-1} \to 0$ $(x \to 0)$ *as* $p > 1$*);*

(ii) $A_p \colon \mathbb{R} \to \mathbb{R}$ *is bounded (follows from (i) as* $\mathbb{R}$ *finite-dimensional);*

(iii) $A_p \colon \mathbb{R} \to \mathbb{R}$ *is coercive (since* $|A_p x| = |x|^{p-1} \to +\infty$, $\operatorname{sign}(A_p x) = \operatorname{sign}(x)$*, and, thus,* $A_p x \to \pm\infty$ $(x \to \pm\infty)$*);*

(iv) $A_p \colon \mathbb{R} \to \mathbb{R}$ *is d-monotone and, thus, stricly monotone (Exercise Sheet 2, Exercise 3);*

(v) $A_p \colon \mathbb{R} \to \mathbb{R}$ *is bijective and its inverse* $A_p^{-1} \colon \mathbb{R} \to \mathbb{R}$ *is continuous, bounded, coercive, d-monotone, and strictly monotone. More precisely, we have that (Exercise Sheet 2, Exercise 3)*

$$A_p^{-1} = A_{p'} \,.$$

(vi) If $p \in [2, +\infty)$*, then* $A_p \colon \mathbb{R} \to \mathbb{R}$ *is uniformly continuous; more precisely, there exists a constant* $\mu > 0$ *such that for every* $x, y \in \mathbb{R}$*, it holds that (Homework 3, Problem 3)*

$$\big(A_p x - A_p y\big)(x - y) \geq \mu\, |x - y|^p \,.$$

In particular, in the case $p = 2$, $A_p \colon \mathbb{R} \to \mathbb{R}$ *is strongly monotone;*

(vii) If $p \in [2, +\infty)$*, then* $A_p \colon \mathbb{R} \to \mathbb{R}$ *is locally Lipschitz continuous; more precisely, for every* $x, y \in \mathbb{R}$*, it holds that (Exercise)*

$$|A_p x - A_p y| \leq (p-2)(\max\{|x|, |y|\})^{p-1} |x - y| \,.$$

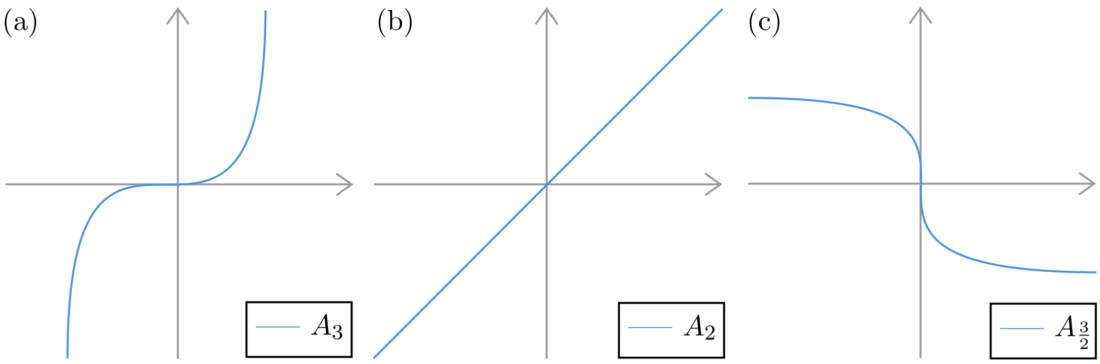

Figure 2.2: (a) $A_3\colon \mathbb{R}\to\mathbb{R}$, defined by $A_3x:=|x|^2x$ for all $x\in\mathbb{R}$, which satisfies $A_3^{-1}=A_{\frac{3}{2}}$;
(b) $A_2\colon \mathbb{R}\to\mathbb{R}$, defined by $A_2x:=x$ for all $x\in\mathbb{R}$, which satisfies $A_2^{-1}=A_2$;
(c) $A_{\frac{3}{2}}\colon \mathbb{R}\to\mathbb{R}$, defined by $A_{\frac{3}{2}}x:=|x|^{-\frac{1}{2}}x$ for all $x\in\mathbb{R}$, which satisfies $A_{\frac{3}{2}}^{-1}=A_3$.

One cornerstone of monotone operator theory constitutes Minty's Trick, which shows that for establishing demi-continuity of a monotone operator is suffices to establish radial continuity, which typically is straightforward to verify.

Lemma 2.8 (Minty's Trick, 1962)
Let $(X,\|\cdot\|_X)$ be a (real) Banach space. Moreover, let $A\colon X\to X^*$ be a monotone and radially continuous operator. Then, the following statements apply:

(i) The operator $A\colon X\to X^*$ is maximal monotone*, i.e., for a couple $(x,x^*)^\top\in X\times X^*$ from*

$$\langle x^*-Ay,x-y\rangle_X\geq 0\qquad\text{for all } y\in X\,,$$

it follows that $Ax=x^*$ in X^*;

(ii) The operator $A\colon X\to X^*$ is of type (M)[2]*, i.e., for a sequence $(x_n)_{n\in\mathbb{N}}\subseteq X$ and element $x\in X$, $x^*\in X^*$ from*

$$\begin{aligned} x_n &\rightharpoonup x &&\text{in } X &&(n\to\infty)\,,\\ Ax_n &\rightharpoonup x^* &&\text{in } X^* &&(n\to\infty)\,,\\ \limsup_{n\to\infty}\,&\langle Ax_n,x_n\rangle_X\leq\langle x^*,x\rangle_X\,, \end{aligned}$$

it follows that $Ax=x^*$ in X^*.

[2]As far as I know, *'(M)'* stands for *'Minty'*.

Proof. ad (i). If we replace $y \in X$ by $x \pm t\,y \in X$, where $y \in X$ and $t \in (0,1]$ are arbitrary, for every $y \in X$, we find that

$$\pm t\langle x^* - A(x + t\,y), y\rangle_X \geq 0\,.$$

Next, dividing by $t > 0$, for every $y \in X$ and $t \in (0,1]$, we infer that

$$\pm\langle x^* - A(x + t\,y), y\rangle_X \geq 0\,.$$

Inasmuch as $A\colon X \to X^*$ is radially continuous, by passing for $t \to 0$, for every $y \in X$, we infer that

$$\pm\langle x^* - Ax, y\rangle_X \geq 0\,,$$

i.e., for every $y \in X$, we have that

$$\langle x^* - Ax, y\rangle_X = 0\,,$$

which implies that $Ax = x^*$ in X^*.

ad (ii). Since $A\colon X \to X^*$ is monotone, for every $y \in X$ and $n \in \mathbb{N}$, we have that

$$\langle Ax_n - Ay, x_n - y\rangle_X \geq 0\,,$$

which, equivalently, can be reformulated as

$$\langle Ax_n, x_n\rangle_X \geq \langle Ax_n, y\rangle_X + \langle Ay, x_n - y\rangle_X\,.$$

Hence, taking the limit superior with respect to $n \to \infty$ on both sides, for every $y \in X$, we arrive at

$$\langle x^*, x\rangle_X \geq \langle x^*, y\rangle_X + \langle Ay, x - y\rangle_X\,,$$

which, equivalently, can be reformulated as

$$\langle x^* - Ay, x - y\rangle_X \geq 0\,.$$

Since $A\colon X \to X^*$, due to (i), is maximal monotone, we conclude that $Ax = x^*$ in X^*. $\square$

Minty's Trick (*cf.* Lemma 2.8) is next crucial for proving the third statement in the following lemma, which collects useful properties of monotone operators.

Lemma 2.9

Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, the following statements apply:

(i) If $A\colon X \to X^*$ is monotone, then $A\colon X \to X^*$ is locally bounded;

(ii) If $A\colon X \to X^*$ is linear and monotone, then $A\colon X \to X^*$ is continuous/bounded;

(iii) If $A\colon X \to X^*$ is monotone and radially continuous and X reflexive, then $A\colon X \to X^*$ is demi-continuous.

In order to prove Lemma 2.9(i), we need to resort to the Banach–Steinhaus theorem.

Theorem 2.10 (Banach–Steinhaus[3])

Let $(X, \|\cdot\|_X)$ be a (real) Banach space and $(Y, \|\cdot\|_Y)$ a (real) normed vector space. Moreover, let $A_n\colon X \to Y$, $n \in \mathbb{N}$, be a point-wise bounded sequence of linear and continuous operators, i.e., for every $x \in X$, it holds that

$$\sup_{n\in\mathbb{N}} \|A_n x\|_Y < \infty\,.$$

Then, the sequence of operators is uniformly bounded, i.e.,

$$\sup_{n\in\mathbb{N}} \|A_n\|_{L(X;Y)} < \infty\,.$$

Proof. See [Bré11, Thm. 2.2]. □

Proof (of Lemma 2.9). ad (i). Suppose that $A\colon X \to X^*$ is not locally bounded. Then, there exists and element $x \in X$ and a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ such that

$$\begin{aligned} x_n &\to x \quad \text{in } X & &(n\to\infty)\,,\\ \|Ax_n\|_{X^*} &\to +\infty & &(n\to\infty)\,. \end{aligned}$$

Next, let the sequence $(a_n)_{n\in\mathbb{N}} \subseteq [0,1]$, for every $n \in \mathbb{N}$, be defined by

$$a_n \coloneqq \frac{1}{1 + \|Ax_n\|_{X^*}\|x_n - x\|_X}\,.$$

[3] Hugo Steinhaus, born 1887 in Jaslo, died 1972 in Wrodaw. Steinhaus studied in Gottingen under Klein and Carathédory, among others, and completed his doctorate under Hilbert. He belonged to the famous Lviv school around Banach and dealt with questions of functional analysis and probability theory, among other things.

Then, for every $y \in X$ and $n \in \mathbb{N}$, it holds that

$$\begin{aligned} 0 &\leq \langle Ax_n - Ay, x_n - y\rangle_X \\ &= \langle Ax_n - Ay, (x_n - x) + (x - y)\rangle_X \,, \end{aligned}$$

and, thus,

$$\begin{aligned} a_n\langle Ax_n, y - x\rangle_X &\leq a_n\big\{\langle Ax_n, x_n - x\rangle_X + \langle Ay, x_n - y\rangle_X\big\} \\ &= a_n\big\{\|Ax_n\|_{X^*}\|x_n - x\|_X + \|Ay\|_{X^*}(\|x_n\|_X + \|y\|_X)\big\} \\ &\leq 1 + c(y, x)\,, \end{aligned}$$

where

$$c(y, x) \coloneqq \|Ay\|_{X^*}\Big(\sup_{n\in\mathbb{N}}\|x_n\|_X + \|y\|_X\Big) > 0\,.$$

If we replace $y \in X$ by $x \pm y \in X$ for arbitrary $y \in X$ above, then, for every $y \in X$, we find that

$$\sup_{n\in\mathbb{N}}\big\{a_n|\langle Ax_n, y\rangle_X|\big\} \leq 1 + \max\{c(x + y, x), c(x - y, x)\}\,,$$

so that, by the Banach–Steinhaus theorem (*cf.* Theorem 2.10), we arrive that

$$\sup_{n\in\mathbb{N}}\|a_n Ax_n\|_{X^*} \leq c(x)\,,$$

where $c(x) > 0$ is independent of $n \in \mathbb{N}$. Next, we choose $n_0 \in \mathbb{N}$ so large that for every $n \in \mathbb{N}$ with $n \geq n_0$, it holds that

$$\|x_n - x\|_X < \frac{1}{2c(x)}\,.$$

Then, for every $n \in \mathbb{N}$ with $n \geq n_0$, we conclude that

$$\begin{aligned} \|Ax_n\|_{X^*} &\leq \frac{c(x)}{a_n} \\ &= c(x)(1 + \|Ax_n\|_{X^*}\|x_n - x\|_X) \\ &\leq c(x) + \frac{\|Ax_n\|_{X^*}}{2}\,, \end{aligned}$$

and, thus,

$$\sup_{n\in\mathbb{N}\,:\,n\geq n_0}\|Ax_n\|_{X^*} \leq 2c(x)\,,$$

which contradicts the unboundedness of the sequence $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$. In other words, we have shown that $A\colon X \to X^*$ is locally bounded.

ad (ii). Since $A\colon X\to X^*$ is monotone, due to (i), $A\colon X\to X^*$ is locally bounded. Therefore, there exists constants $M,\varepsilon>0$ such that for every $y\in \overline{B}_\varepsilon^X(0)$, it holds that

$$\|Ay\|_{X^*}\leq M\,,$$

which, by the linearity of $A\colon X\to X^*$, for every $y\in X$, implies that

$$\|Ay\|_{X^*}\leq \frac{M}{\varepsilon}\|y\|_X\,,$$

since for every $y\in X\setminus\{0\}$, we have that

$$\left\|A\left(\varepsilon\frac{y}{\|y\|_X}\right)\right\|_X\leq M\,.$$

In other words, we have shown that $A\colon X\to X^*$ is bounded and, thus, continuous.

ad (iii). Let $(x_n)_{n\in\mathbb{N}}\subseteq X$ be a sequence and $x\in X$ an element such that

$$x_n\to x\quad \text{in } X\quad (n\to\infty)\,.$$

Since $A\colon X\to X^*$ is locally bounded, the sequence of images $(Ax_n)_{n\in\mathbb{N}}\subseteq X^*$ is bounded as well. Hence, since X^* is reflexive, by the Eberlein–Šmulian theorem (*cf.* Theorem 1.14), there exists a subsequence $(Ax_{n_k})_{k\in\mathbb{N}}\subseteq X^*$ and an element $x^*\in X^*$ such that

$$Ax_{n_k}\rightharpoonup x^*\quad \text{in } X^*\quad (k\to\infty)\,.$$

Due to Proposition 1.13(ii)&(v), we have that

$$\langle Ax_{n_k},x_{n_k}\rangle_X\to\langle x^*,x\rangle_X\quad (n\to\infty)\,,$$

which, apparently, implies that

$$\lim_{k\to\infty}\langle Ax_{n_k},x_{n_k}\rangle_X=\langle x^*,x\rangle_X\,.$$

In summary, by Minty's Trick (*cf.* Lemma 2.8), we obtain $Ax=x^*$ in X^*. In other words, we have that

$$Ax_{n_k}\rightharpoonup Ax\quad \text{in } X^*\quad (k\to\infty)\,.$$

As this argumentation remains valid for each subsequence of $(Ax_n)_{n\in\mathbb{N}}\subseteq X^*$, $Ax\in X^*$ is the sole weak accumulation point of the sequence $(Ax_n)_{n\in\mathbb{N}}\subseteq X^*$. Hence, by the subsequence convergence principle (*cf.* Lemma 1.17), we conclude that

$$Ax_n\rightharpoonup Ax\quad \text{in } X^*\quad (n\to\infty)\,.$$

In other words, we have shown that $A\colon X\to X^*$ is demi-continuous. □

2.3 Notion of coercivity

In the same manner we generalized monotonicity to infinite-dimensional Banach spaces, we can generalize coercivity to infinite-dimensional Banach spaces. Again, it is less obvious how to generalize coercivity, which is also a concept that seems to rely on fact that $(\mathbb{R}, <)$ is a strictly order set. For a general infinite-dimensional Banach space (or even for $\mathbb{R}^d$, $d \in \mathbb{N}$) such a strict order is, in general, not available. A remedy is, then, provided by the observation that coercivity can equivalently expressed without directly using the strictly ordering of $\mathbb{R}$. More precisely, it is an exercise (*cf.* Exercise Sheet 2, Micro Exercise 5) to verify that a function (or an operator) $A\colon \mathbb{R} \to \mathbb{R}$ is coercive if and only if

$$\frac{(Ax) \cdot x}{|x|} \to +\infty \quad (|x| \to +\infty)\,.$$

Definition 2.11 (coercivity)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, an operator $A\colon X \to X^*$ is called coercive *if there exists a function $\gamma\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfying $\gamma(s) \to +\infty$ $(s \to +\infty)$ such that for every $x \in X$, it holds that*

$$\langle Ax, x\rangle_X \geq \gamma(\|x\|_X)\|x\|_X\,,$$

or equivalently

$$\frac{\langle Ax, x\rangle_X}{\|x\|_X} \to +\infty \quad (\|x\|_X \to +\infty)\,.$$

A first class of operators that are coercive is given through strongly monotone operators.

Lemma 2.12
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Moreover, let $A\colon X \to X^*$ be a strongly monotone operator. Then, $A\colon X \to X^*$ is coercive.

Proof. Since $A\colon X \to X^*$ is strongly monotone, there exists a constant $m > 0$ such that for every $x, y \in X$, it holds that

$$\langle Ax - Ay, x - y\rangle_X \geq m\,\|x - y\|_X^2\,.$$

In particular, for every $x \in X$, we have that

$$\langle Ax - A0, x\rangle_X \geq m\,\|x\|_X^2\,.$$

Therefore, for every $x \in X$, we infer that

$$\left.\begin{aligned}\frac{\langle Ax, x\rangle_X}{\|x\|_X} &\geq m\,\|x\|_X + \frac{\langle A0, x\rangle_X}{\|x\|_X}\\ &\geq m\,\|x\|_X - \|A0\|_{X^*}\end{aligned}\right\} \to +\infty \quad (\|x\|_X \to +\infty)\,.$$

In other words, we have shown that the operator $A\colon X \to X^*$ is coercive. $\square$

2.4 Main theorem on monotone operators

After having generalized the notions of continuity, coercivity, and strict monotonicity from the inverse function theorem for striclty monotone operators (*cf.* Theorem 2.1) to infinite-dimensional Banach spaces, we have everything at our disposal to, eventually, establish a generalization of the inverse function theorem for striclty monotone operators to infinite-dimensional Banach spaces (*cf.* Theorem 2.1).

Theorem 2.13 (Browder–Minty, 1963)
Let $(X, \|\cdot\|_X)$ be a (real) separable, reflexive Banach space. Moreover, let $A\colon X \to X^*$ be a monotone, radially continuous, and coercive operator. Then, the following statements apply:

(i) The operator $A\colon X \to X^*$ is surjective and for every $x^* \in X^*$, the set of solutions (or the pre-image)

$$\mathbb{L}(x^*) \coloneqq \big\{x \in X \mid Ax = x^*\big\}\big(= A^{-1}x^*\big)\,,$$

is convex, closed, and, thus, weakly closed.

(ii) If, in addition, $A\colon X \to X^*$ is strictly monotone, then $A\colon X \to X^*$ is bijective and its inverse $j_X \circ A^{-1}\colon X^* \to X^{**}$ is strictly monotone, bounded, and demi-continuous;

(iii) If, in addition, $A\colon X \to X^*$ is strongly monotone, then $j_X \circ A^{-1}\colon X^* \to X^{**}$[4] *Lipschitz continuous;*

(iv) If, in addition, $A\colon X \to X^*$ is strictly monotone and Lipschitz continuous, then $j_X \circ A^{-1}\colon X^* \to X^{**}$ strongly monotone.

Remark 2.14 (a few comments) *(i) Theorem 2.13 is typically called the* Browder–Minty theorem *or the* main theorem on monotone operators*;*

(ii) Theorem 2.13 remains true if we drop the assumption that X is separable. Then, however, one cannot approximate X with a increasing sequence of finite-dimensional subspaces. Instead, one then performs a Galerkin approximation over all finite-dimensional subspaces. As this approximation is completely unordered, the passage to the limit is more technically demanding. The idea is generate for each finite-dimensional subspace an increasing sequence of finite-dimensional subspaces and, then, to perform the passage to the limit for each finite-dimensional subspace (cf. [Růž04, Folg. 3.51]).

We postpone the proof of Theorem 2.13(i) and start with proving Theorem 2.13(ii)–(iv), for which we will initially assume that Theorem 2.13(i) is true.

Proof (of Theorem 2.13(ii)–(iv)). *ad (ii).*

1. Existence of $A^{-1}\colon X^* \to X$: According to (i), the operator $A\colon X \to X^*$ is surjective. In addition, the operator $A\colon X \to X^*$ is injective since for $x, y \in X$ from $Ax = Ay$ in X^*, it follows that $x = y$ (otherwise, if $x \neq y$, by the strict monotonicity of $A\colon X \to X^*$, it would follow that $\langle Ax - Ay, x - y\rangle_X > 0$, in contrast to $Ax = Ay$ in X^*). In summary, we have shown that $A\colon X \to X^*$ is bijective and, thus, that the inverse $A^{-1}\colon X^* \to X$ exists.

[4] Recall that $j_X : X \to X^{**}$ is the canonical isometry (*cf.* Definition 1.6), which for a reflexive Banach space X is an isometric isomorphism (*cf.* Definition 1.8).

2. Strict monotonicity: For every $x^*, y^* \in X^*$ with $x^* \neq y^*$, due to the bijectivity of $A\colon X \to X^*$, there exist $x, y \in X$ with $x \neq y$ such that $Ax = x^*$ in X^* and $Ay = y^*$ in X^*. Then, by the definition of $j_X : X \to X^{**}$ (*cf.* Definition 1.6) and the strict monotonicity of $A\colon X \to X^*$, we obtain

$$\begin{aligned}\langle (j_X \circ A^{-1})x^* - (j_X \circ A^{-1})y^*, x^* - y^* \rangle_{X^*} &= \langle x^* - y^*, A^{-1}x^* - A^{-1}y^* \rangle_X \\ &= \langle Ax - Ay, x - y \rangle_X > 0 \,.\end{aligned}$$

In other words, the operator $j_X \circ A^{-1} : X^* \to X^{**}$ is strictly monotone.

3. Boundedness: Let $M \subseteq X^*$ be a bounded set. Suppose that $(j_X \circ A^{-1})(M) \subseteq X^{**}$ is unbounded. Then, there exists a sequence $(x_n^*)_{n\in\mathbb{N}} \subseteq M$ such that

$$\|(j_X \circ A^{-1})x_n^*\|_{X^{**}} = \|A^{-1}x_n^*\|_X \to \infty \quad (n \to \infty)\,,$$

where we used the isometry property of $j_X : X \to X^{**}$ (*cf.* Proposition 1.7), so that, by the coercivity of $A\colon X \to X^*$, we infer that

$$\frac{\langle A(A^{-1}x_n^*), A^{-1}x_n^* \rangle_X}{\|A^{-1}x_n^*\|_X} \to +\infty \quad (n \to \infty)\,.$$

On the other hand, since $M \subseteq X^*$ is bounded, for every $n \in \mathbb{N}$, we have that

$$\frac{\langle A(A^{-1}x_n^*), A^{-1}x_n^* \rangle_X}{\|A^{-1}x_n^*\|_X} = \frac{\langle x_n^*, A^{-1}x_n^* \rangle_X}{\|A^{-1}x_n^*\|_X} \leq \sup_{n\in\mathbb{N}} \|x_n^*\|_{X^*} < \infty\,,$$

i.e., a contradiction. Therefore, the set $(j_X \circ A^{-1})(M) \subseteq X^{**}$ is bounded. In other words, the operator $j_X \circ A^{-1} : X^* \to X^{**}$ is bounded.

4. Demi-continuity: Let $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ be a sequence and $x^* \in X^*$ an element such that

$$x_n^* \to x^* \quad \text{in } X^* \quad (n \to \infty)\,.$$

Due to the boundedness of $j_X \circ A^{-1} : X^* \to X$ and the isometry property of $j_X : X \to X^{**}$ (*cf.* Proposition 1.7), the sequence of pre-images $(x_n)_{n\in\mathbb{N}} := (A^{-1}x_n^*) \subseteq X$ is bounded as well. Since X is reflexive, by the Eberlein–Šmulian theorem (*cf.* Theorem 1.14), there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq X$ and an element $x \in X$ such that

$$A^{-1}x_{n_k}^* = x_{n_k} \rightharpoonup x \quad \text{in } X \quad (k \to \infty)\,.$$

In addition, due to the monotonicity of $A\colon X \to X^*$, for every $y \in Y$ and $k \in \mathbb{N}$, we have that

$$\langle x_{n_k}^* - Ay, x_{n_k} - y \rangle_X \geq 0\,.$$

Then, by passing for $k \to \infty$, using Proposition 1.13(ii)&(v) in doing so, for every $y \in Y$, we arrive at

$$\langle x^* - Ay, x - y \rangle_X \geq 0\,.$$

As a result, Minty's Trick (*i.e.*, the maximal monotonicity of $A\colon X \to X^*$, *cf.* Lemma 2.8(i)) yields that $Ax = x^*$ in X^*, *i.e.*, $x = A^{-1}x^*$ in X. Therefore, we have that

$$A^{-1}x_{n_k}^* = x_{n_k} \rightharpoonup x = A^{-1}x^* \quad \text{in } X \quad (k \to \infty)\,.$$

As the entire argumentation applies to weakly convergent subsequence of $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$, each weakly convergent subsequence of $(x_n)_{n\in\mathbb{N}} = (A^{-1}x_n^*)_{n\in\mathbb{N}} \subseteq X$ converges weakly to $x = A^{-1}x^* \in X$. In consequence, the subsequence convergence principle (*cf.* Lemma 1.17) yields that

$$A^{-1}x_n^* \rightharpoonup A^{-1}x^* \quad \text{in } X \quad (n \to \infty)\,.$$

In other words, the operator is demi-continuous $j_X \circ A^{-1} \colon X^* \to X^{**}$ is demi-continuous.

ad (iii),(iv). See Exercise Sheet 4, Exercise 2. $\square$

Let us next prove Theorem 2.13(i). To this end, we first prove an analogue of Theorem 2.13(i) for Euclidean spaces.

Theorem 2.15 (Browder–Minty theorem for Euclidean spaces)
Let $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$, be a continuous vector field. Then, the following statements apply:

(i) If there exists a constant $r > 0$ such that for every $y \in \partial B_r^d(0) \coloneqq \{z \in \mathbb{R}^d \mid |z| = r\}$, it holds that

$$\mathbf{f}(y) \cdot y \geq 0\,,$$

then there exists $x \in \overline{B}_r^d(0) \coloneqq \{z \in \mathbb{R}^d \mid |z| \leq r\}$ such that $\mathbf{f}(x) = 0$.

(ii) If $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$ is coercive, i.e.,

$$\frac{\mathbf{f}(x)\cdot x}{|x|} \to +\infty \quad (|x| \to +\infty)\,,$$

then $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$ is surjective.

(iii) If $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$ is coercive and strictly monotone, i.e., for every $x, y \in \mathbb{R}^d$ with $x \neq y$, it holds that

$$\big(\mathbf{f}(x) - \mathbf{f}(y)\big) \cdot (x - y) > 0\,,$$

then $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$ is bijective and its inverse $\mathbf{f}^{-1}\colon \mathbb{R}^d \to \mathbb{R}^d$ is strictly monotone, bounded, and continuous.

Remark 2.16 *(i) In the case $d = 1$, the condition*

$$\mathbf{f}(y)\cdot y \geq 0 \quad \text{for all } y \in \partial B_r^1(0) = \{-r, r\}\,,$$

is equivalent to

$$\operatorname{sign}(\mathbf{f}(-r)) \neq \operatorname{sign}(\mathbf{f}(r))\,,$$

i.e., in the case $d = 1$, the above condition simply indicates a change of the sign, which for continuous functions, by the intermediate value theorem, implies that the function admits a root, i.e., there exists $x \in [-r, r]$ such that $\mathbf{f}(x) = 0$. For this reason, Theorem 2.15(i) is a canonical generalization of the intermediate value theorem for continuous vector fields $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$;

(ii) Theorem 2.15(ii) is a canonical generalization of Theorem 2.1(i) for continuous vector fields $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$;

(iii) Theorem 2.15(iii) is a canonical generalization of Theorem 2.1(ii) for continuous vector fields $\mathbf{f}\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$.

Key ingredient in the proof of Theorem 2.15(i) is the Brouwer fixed point theorem. In fact, it is possible to show that Theorem 2.15(i) is equivalent to the Brouwer fixed point theorem (Exercise Sheet 3, Exercise 1).

Theorem 2.17 (Brouwer[5], 1912)
Let $d\in\mathbb{N}$ *and* $r>0$*. Moreover, let* $\mathbf{g}\colon\overline{B}_r^d(0)\to\overline{B}_r^d(0)$ *be continuous. Then,* $\mathbf{g}\colon\overline{B}_r^d(0)\to\overline{B}_r^d(0)$ *admits a fixed point, i.e., there exists* $x\in\overline{B}_r^d(0)$ *such that*

$$\mathbf{g}(x)=x\,.$$

In other words, any continuous self-mapping on a closed ball in $\mathbb{R}^d$ *admits a fixed point.*

Proof. See [Růž04, Satz 2.14]. □

Proof (of Theorem 2.15). ad (i). Suppose that $\mathbf{f}(y)\neq 0$ for all $y\in\overline{B}_r^d(0)$. Then, the vector field $\mathbf{g}\colon\mathbb{R}^d\to\mathbb{R}^d$, for every $y\in\mathbb{R}^d$, defined by

$$\mathbf{g}(y):=-r\frac{\mathbf{f}(y)}{|\mathbf{f}(y)|}\,,$$

is well-defined, continuous, and satisfies $\mathbf{g}(\overline{B}_r^d(0))\subseteq\partial B_r^d(0)$. Therefore, by the Brouwer fixed point theorem (*cf.* Theorem 2.17), there exists $x\in\overline{B}_r^d(0)$ such that

$$\mathbf{g}(x)=x\,,$$

which, apparently, implies that $x\in\partial B_r^d(0)$ as

$$|x|=r\frac{|\mathbf{f}(y)|}{|\mathbf{f}(y)|}=r\,.$$

As a consequence, we find that

$$0\leq\mathbf{f}(x)\cdot x=-\frac{|\mathbf{f}(x)|}{r}\mathbf{g}(x)\cdot x=-|x|^2\frac{|\mathbf{f}(x)|}{r}=-r|\mathbf{f}(x)|<0\,,$$

which is a contradiction. In summary, we have shown that there exists $x\in\overline{B}_r^d(0)$ such that $\mathbf{f}(x)=0$.

ad (ii)/(iii). Exercise. □

Eventually, we have everything at out disposal to prove Theorem 2.13(i).

[5]Luitzen Egbertus Jan Brouwer, born 1881 in Overschie (now part of Rotterdam), died 1966 in Blaricum. Brouwer studied in Amsterdam under Korteweg and later succeeded him. He was primarily concerned with topology and the logical and epistemological foundations of mathematics. Alongside Weyl, Brouwer was the leading representative of intuitionism and denied the application of the theorem of the excluded third, provided that infinite sets came into play. Later, Kolmogorov established a calculus of constructive problem-solving in accordance with the rules of intuitionistic logic and Borel also spoke out in favour of exclusively constructive definitions.

Proof (of Theorem 2.13(i)). *1. Surjectivity:* Let $x^* \in X^*$ be fixed, but arbitrary.

Objective: We want to show that there exists some $x \in X$ such that for every $y \in X$, it holds that

$$\langle Ax, y\rangle_X = \langle x^*, y\rangle_X\,,$$

i.e., $Ax = x^*$ in X^*.

To this end, we resort to the *Galerkin method*, *i.e.*, we approximate the Banach space X through a (strictly) increasing sequence of finite-dimensional Banach spaces $(X_n)_{n\in\mathbb{N}}$ (to which Theorem 2.15 can be applied). More precisely, due to the separability of the Banach space X, there exists a countable dense subset $\{b_n \mid n \in \mathbb{N}\}$. Without loss of generality, we may assume that the elements in $\{b_n \mid n \in \mathbb{N}\}$ are linear independent. Then, we define the sequence of finite-dimensional spaces $(X_n)_{n\in\mathbb{N}}$, for every $n \in \mathbb{N}$, by

$$X_n \coloneqq \operatorname{span}(\{b_1, \ldots, b_n\})\,.$$

Then, the following statements apply:

- $\dim(X_n) = n$ and $X_n \subseteq X_{n+1} \subseteq X$ for all $n \in \mathbb{N}$;
- $\bigcup_{n\in\mathbb{N}} X_n$ is dense in X.

Given (strictly) increasing sequence of finite-dimensional spaces $(X_n)_{n\in\mathbb{N}}$, we introduce the a sequence of finite-dimensional *Galerkin systems*: *i.e.*, for every $n \in \mathbb{N}$, we seek for a *Galerkin solution* $x_n \in X_n$ such that for every $y_n \in X_n$, it holds that

$$\langle A_n x_n, y_n\rangle_{X_n} = \langle x_n^*, y_n\rangle_{X_n}\,, \tag{GS}$$

i.e., $(\mathrm{id}_{X_n})^* A x_n = (\mathrm{id}_{X_n})^* x^*$ in X_n^*, where $(\mathrm{id}_{X_n})^* \colon X^* \to X_n^*$ is the adjoint operator of the identity mapping $\mathrm{id}_{X_n} \colon X_n \to X$.

By means of the sequence of Galerkin systems (GS), we prove the existence of a solution $x \in X$ of $Ax = x^*$ in X^* in three main steps:

1. *Well-posedness of Galerkin systems (GS):* We establish the existence of Galerkin solutions of the sequence Galerkin systems (GS);
2. *Stability of Galerkin systems (GS):* We derive *a priori* bounds for the sequence of Galerkin solutions of the sequence Galerkin systems (GS);
3. *(Weak) convergence of Galerkin systems (GS):* We establish the (weak) convergence of the sequence of Galerkin solutions of the sequence Galerkin systems (GS) to a solution $x \in X$ of $Ax = x^*$ in X^*.

1. Well-posedness of the Galerkin systems (GS): We aim to resort to the Browder–Minty theorem for Euclidean spaces (*cf.* Theorem 2.15(i)). To this end, we need to translate the sequence of Galerkin systems (GS) to the Euclidean setting. Since $\dim(X_n) = \dim(\mathbb{R}^n)$ for all $n \in \mathbb{N}$, the finite-dimensional space X_n for every $n \in \mathbb{N}$ is isomorphic to the Euclidean space $\mathbb{R}^n$. More precisely, a canonical isomorphism is given via the basis transformation mapping $\boldsymbol{\Psi}_n \colon \mathbb{R}^n \to X_n$, for every $\boldsymbol{\beta}_n \coloneqq (\beta_n^1, \ldots, \beta_n^n)^\top \in \mathbb{R}^n$ defined by

$$\boldsymbol{\Psi}_n(\boldsymbol{\beta}_n) \coloneqq \sum_{i=1}^n \beta_n^i\, b_i \quad \text{in } X_n\,.$$

With the help of the basis transformation mappings $\mathbf{\Psi}_n\colon \mathbb{R}^n \to X_n$, $n \in \mathbb{N}$, the sequence of the Galerkin systems (GS) can equivalently re-written as follows: for every $n \in \mathbb{N}$, seek for $\boldsymbol{\alpha}_n \in \mathbb{R}^n$ such that

$$\mathbf{f}_n(\boldsymbol{\alpha}_n) = \mathbf{0} \quad \text{in } \mathbb{R}^n ,$$

where the mapping $\mathbf{f}_n\colon \mathbb{R}^n \to \mathbb{R}^n$, for every $\boldsymbol{\beta}_n \in \mathbb{R}^n$, is defined by

$$\mathbf{f}_n(\boldsymbol{\beta}_n) \coloneqq (\langle A_n \mathbf{\Psi}_n(\boldsymbol{\beta}_n) - x^*, b_i\rangle_X)_{i=1,\dots,n} \quad \text{in } \mathbb{R}^n .$$

Therefore, it is left to show that $\mathbf{f}_n\colon \mathbb{R}^n \to \mathbb{R}^n$, for every $n \in \mathbb{N}$, meets the assumptions of the Browder–Minty theorem for Euclidean spaces (*cf.* Theorem 2.15(i)):

1. Continuity: Let $(\boldsymbol{\alpha}_n^k)_{k\in\mathbb{N}} \subseteq \mathbb{R}^n$ be a sequence and $\boldsymbol{\alpha}_n \in \mathbb{R}^n$ an element such that

$$\boldsymbol{\alpha}_n^k \to \boldsymbol{\alpha}_n \quad \text{in } \mathbb{R}^n \quad (k \to \infty) .$$

Since the basis transformation mapping $\mathbf{\Psi}_n\colon \mathbb{R}^n \to X_n$ is continuous, we infer that

$$\mathbf{\Psi}_n(\boldsymbol{\alpha}_n^k) \to \mathbf{\Psi}_n(\boldsymbol{\alpha}_n) \quad \text{in } X_n \quad (k \to \infty) .$$

Then, since $X_n \hookrightarrow X$ and $A\colon X \to X^*$ is demi-continuous (*cf.* Lemma 2.9(iii)), we conclude that

$$A\mathbf{\Psi}_n(\boldsymbol{\alpha}_n^k) \rightharpoonup A\mathbf{\Psi}_n(\boldsymbol{\alpha}_n) \quad \text{in } X^* \quad (k \to \infty) ,$$

and, consequently, for every $i = 1, \dots, n$,

$$\langle A\mathbf{\Psi}_n(\boldsymbol{\alpha}_n^k) - x^*, b_i\rangle_X \to \langle A\mathbf{\Psi}_n(\boldsymbol{\alpha}_n) - x^*, b_i\rangle_X \quad (k \to \infty) .$$

In other words, we have just proved that

$$\mathbf{f}_n(\boldsymbol{\alpha}_n^k) \to \mathbf{f}_n(\boldsymbol{\alpha}_n) \quad \text{in } \mathbb{R}^n \quad (k \to \infty) ,$$

i.e., $\mathbf{f}_n\colon \mathbb{R}^n \to \mathbb{R}^n$ is continuous.

2. Generalized change of sign: For every $\boldsymbol{\beta}_n \in \mathbb{R}^n$, we have that

$$\begin{aligned} \mathbf{f}_n(\boldsymbol{\beta}_n) \cdot \boldsymbol{\beta}_n &= \langle A\mathbf{\Psi}_n(\boldsymbol{\beta}_n) - x^*, \mathbf{\Psi}_n(\boldsymbol{\beta}_n)\rangle_X \\ &\geq (\gamma(\|\mathbf{\Psi}_n(\boldsymbol{\beta}_n)\|_X) - \|x^*\|_{X^*})\|\mathbf{\Psi}_n(\boldsymbol{\beta}_n)\|_X , \end{aligned}$$

where $\gamma\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfies $\gamma(s) \to +\infty$ $(s \to +\infty)$ and exists due to the coercivity of $A\colon X \to X^*$ (*cf.* Definition 2.11). Then, there exists a constant $r > 0$ such that for every $s \geq r$, it holds that

$$\gamma(s) \geq \|x^*\|_{X^*} .$$

Since $\|\boldsymbol{\Psi}_n(\cdot)\|_X$ and $|\cdot|$ are equivalent norms on $\mathbb{R}^n$, there exists a constant $c_n > 0$, depending on $n \in \mathbb{N}$, such that for every $\boldsymbol{\beta}_n \in \mathbb{R}^n$, it holds that

$$c_n^{-1}\,|\boldsymbol{\beta}_n| \leq \|\boldsymbol{\Psi}_n(\boldsymbol{\beta}_n)\|_X \leq c_n\,|\boldsymbol{\beta}_n|\,.$$

Therefore, for every $\boldsymbol{\beta}_n \in \mathbb{R}^n$ with $|\boldsymbol{\beta}_n| = c_n r$, *i.e.*, $\|\boldsymbol{\Psi}_n(\boldsymbol{\beta}_n)\|_X \geq c_n^{-1}|\boldsymbol{\beta}_n| = r$, we have that

$$\mathbf{f}_n(\boldsymbol{\beta}_n)\cdot\boldsymbol{\beta}_n \geq 0\,.$$

In summary, the assumptions of the Browder–Minty theorem for Euclidean spaces (*cf.* Theorem 2.15(i)) are met and the latter yields some $\boldsymbol{\alpha}_n \in \mathbb{R}^n$ such that

$$\mathbf{f}_n(\boldsymbol{\alpha}_n) = 0 \quad \text{in } \mathbb{R}^n\,,$$

and, setting $x_n \coloneqq \boldsymbol{\Psi}_n(\boldsymbol{\alpha}_n) \in X_n$, we just proved the existence of a Galerkin solution of the Galerkin system (GS).

2. Stability of the Galerkin systems (GS):

2.1. Boundedness of sequence of solutions: Suppose that the sequence of Galerkin solutions $(x_n)_{n\in\mathbb{N}} \subseteq X$ of the sequence of Galerkin schemes (GS) is not bounded, *i.e.*, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq X$ such that

$$\|x_{n_k}\|_X \to +\infty \quad (k \to \infty)\,.$$

Then, by the coercivity of $A\colon X \to X^*$, it follows that

$$\frac{\langle Ax_{n_k}, x_{n_k}\rangle_X}{\|x_{n_k}\|_X} \to +\infty \quad (k \to \infty)\,.$$

On the other hand, due to $\langle Ax_n, y_n\rangle_X = \langle x^*, y_n\rangle_X$ for all $y_n \in X_n$ (*cf.* (GS)), choosing $y_n = x_n \in X_n$ in (GS), for every $n \in \mathbb{N}$, we have that

$$\frac{\langle Ax_n, x_n\rangle_X}{\|x_n\|_X} = \frac{\langle x^*, x_n\rangle_X}{\|x_n\|_X} \leq \|x^*\|_{X^*}\,,$$

i.e., a contradiction. In summary, we have shown that the sequence of Galerkin solutions $(x_n)_{n\in\mathbb{N}} \subseteq X$ of the sequence of Galerkin schemes (GS) is bounded.

2.2. Boundedness of images of sequence of solutions: Since $A\colon X \to X^*$ is locally bounded (*cf.* Lemma 2.9(i)), there exist constants $M, \varepsilon > 0$ such that for every $y \in \overline{B}_\varepsilon^X(0)$, it holds that

$$\|Ay\|_{X^*} \leq M\,.$$

Therefore, by the monotonicity of $A\colon X\to X^*$, for every $n\in\mathbb{N}$ and $y\in\overline{B}_{\varepsilon}^{X}(0)$, we have that

$$\begin{aligned}\langle Ax_n,y\rangle_X&\leq\langle Ax_n,x_n\rangle_X-\langle Ay,x_n-y\rangle_X\\&=\langle x^*,x_n\rangle_X-\langle Ay,x_n-y\rangle_X\\&\leq(\|x^*\|_{X^*}+M)\left(\sup_{n\in\mathbb{N}}\|x_n\|_X+\varepsilon\right).\end{aligned}$$

Eventually, using a scaled version of the operator norm, we find that

$$\begin{aligned}\|Ax_n\|_{X^*}&=\sup_{y\in\overline{B}_1^X(0)}\langle Ax_n,y\rangle_X\\&=\sup_{y\in\overline{B}_{\varepsilon}^X(0)}\frac{1}{\varepsilon}\langle Ax_n,y\rangle_X\\&\leq\frac{1}{\varepsilon}(\|x^*\|_{X^*}+M)\left(\sup_{n\in\mathbb{N}}\|x_n\|_X+\varepsilon\right),\end{aligned}$$

i.e., the sequence of images of Galerkin solutions $(Ax_n)_{n\in\mathbb{N}}\subseteq X^*$ is bounded.

3. (Weak) convergence of Galerkin scheme: Due to Step 2, the sequence of Galerkin solutions $(x_n)_{n\in\mathbb{N}}\subseteq X$ of the sequence Galerkin schemes (GS) as well as the corresponding sequence of images $(Ax_n)_{n\in\mathbb{N}}\subseteq X^*$ are bounded. As a consequence, by the reflexivity of X, the Eberlein–Šmulian theorem (*cf.* Theorem 1.14) yields subsequences $(x_{n_k})_{k\in\mathbb{N}}\subseteq X$, $(Ax_{n_k})_{k\in\mathbb{N}}\subseteq X^*$ as well as elements $x\in X$, $\xi^*\in X^*$ such that

$$\begin{aligned}x_{n_k}&\rightharpoonup x&&\text{in } X&&(k\to\infty),\\Ax_{n_k}&\rightharpoonup \xi^*&&\text{in } X^*&&(k\to\infty).\end{aligned}$$

First, we show that $\xi^*=x^*$ in X^*. To this end, let $y\in X$ be fixed, but arbitrary. Then, there exists a sequence $y_n\in X_n$, $n\in\mathbb{N}$, such that

$$y_n\to y\quad\text{in } X\quad(n\to\infty).$$

Hence, by Proposition 1.13(ii)&(v), we deduce that

$$\begin{aligned}\langle\xi^*,y\rangle_X&=\lim_{k\to\infty}\langle Ax_{n_k},y_{n_k}\rangle_X\\&=\lim_{k\to\infty}\langle x^*,y_{n_k}\rangle_X\\&=\langle x^*,y\rangle_X.\end{aligned}$$

Since $y\in X$ was chosen arbitrary, we infer that $\xi^*=x^*$ in X^*. As a result, we obtain

$$\begin{aligned}\limsup_{k\to\infty}\langle Ax_{n_k},x_{n_k}\rangle_X&=\lim_{k\to\infty}\langle x^*,x_{n_k}\rangle_X\\&=\langle x^*,x\rangle_X,\end{aligned}$$

so that, by Minty's Trick (*i.e.*, $A\colon X\to X^*$ is of type (M), *cf.* Lemma 2.8(ii)), we conclude that $Ax=x^*$ in X^*.

2. Properties of set of solutions: Let $x^*\in X^*$ be fixed, but arbitrary.
2.1 Non-emptyness of $\mathbb{L}(x^*)$*:* Follows from the surjectivity of $A\colon X\to X^*$.
2.2 Boundedness of $\mathbb{L}(x^*)$*:* Suppose that $\mathbb{L}(x^*)$ is unbounded, *i.e.*, there exists a sequence $(x_n)_{n\in\mathbb{N}}\subseteq \mathbb{L}(x^*)$ such that

$$\|x_n\|_X\to+\infty\quad (k\to\infty)\,.$$

Then, by the coercivity of $A\colon X\to X^*$, it follows that

$$\frac{\langle Ax_n,x_n\rangle_X}{\|x_n\|_X}\to+\infty\quad (n\to\infty)\,.$$

On the other hand, due to $\langle Ax_n,y_n\rangle_X=\langle x^*,y_n\rangle_X$ for all $y_n\in X_n$ (*cf.* (GS)), choosing $y_n=x_n\in X_n$ in (GS), for every $n\in\mathbb{N}$, we have that

$$\frac{\langle Ax_n,x_n\rangle_X}{\|x_n\|_X}=\frac{\langle x^*,x_n\rangle_X}{\|x_n\|_X}\le\|x^*\|_{X^*}\,,$$

i.e., a contradiction. In summary, we have shown that $\mathbb{L}(x^*)$ is bounded.
2.3 Convexity of $\mathbb{L}(x^*)$*:* Let $x,y\in\mathbb{L}(x^*)$ and $\lambda\in[0,1]$. Then, due to the monotonicity of $A\colon X\to X^*$ and $Ax=Ay=x^*$ in X^*, for every $z\in X$, we have that

$$\begin{aligned}\langle x^*-Az,(\lambda\,x+(1-\lambda)\,y)-z\rangle_X&=\lambda\,\langle x^*-Az,x-z\rangle_X\\&\quad+(1-\lambda)\,\langle x^*-Az,y-z\rangle_X\\&=\lambda\,\langle Ax-Az,x-z\rangle_X\\&\quad+(1-\lambda)\,\langle Ay-Az,y-z\rangle_X\\&\ge 0\,.\end{aligned}$$

In consequence, since $A\colon X\to X^*$ is maximal monotone (*cf.* Lemma 2.8), we conclude that $A(\lambda\,x+(1-\lambda)\,y)=x^*$ in X^*, *i.e.*, $\lambda\,x+(1-\lambda)\,y\in\mathbb{L}(x^*)$. In other words, we have shown that $\mathbb{L}(x^*)$ is convex.
2.4 Closedness of $\mathbb{L}(x^*)$*:* Let $(x_n)_{n\in\mathbb{N}}\subseteq\mathbb{L}(x^*)$ be a sequence and $x\in X$ an element such that

$$x_n\to x\quad\text{ in }X\quad (n\to\infty)\,.$$

Since $A\colon X\to X^*$ is demi-continuous (*cf.* Lemma 2.9(iii)), we infer that

$$Ax_n\to Ax\quad\text{ in }X^*\quad (n\to\infty)\,.$$

Then, due to $Ax_n=x^*$ in X^* for all $n\in\mathbb{N}$, we find that $Ax=x^*$ in X^*, *i.e.*, $x\in\mathbb{L}(x^*)$. In other words, we have shown that $\mathbb{L}(x^*)$ is closed.
2.5 Weak closedness of $\mathbb{L}(x^*)$*:* Follows from the closedness and convexity of $\mathbb{L}(x^*)$, since, by Mazur's theorem (*cf.* [Bré11, Cor. 3.8]), a convex set is weakly closed if and only if it is (strongly) closed. □

2.5 Applications of the main theorem on monotone operators

In this section, we address two famous applications of the main theorem on monotone operators (*cf.* Theorem 2.13).

2.5.1 p-Laplace equation

As a first application of the main theorem on monotone operators (*cf.* Theorem 2.13), we obtain the well-posedness of a weak formulation of the p-Laplace equation, which models the deflection of a (non-linear) elastic membrane (made of a non-ideal material) fixed at the boundary of a domain under the influence of an external force. More precisely, the p-Laplace equation, for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, over which the membrane is spanned and to which (topological) boundary $\partial\Omega$ the membrane is fixed to, and a given external force $f\colon \Omega \to \mathbb{R}$ acting on the membrane, seeks for a deflection field $u\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\operatorname{div}(|\nabla u|^{p-2}\nabla u) &= f \qquad && \text{in } \Omega\,, \\ u &= 0 \qquad && \text{on } \partial\Omega\,, \end{aligned} \tag{2.18}$$

where $p \in (1, +\infty)$ is a material parameter that characterizes the non-linear elastic behavior.

Definition 2.19 (weak formulation of the p-Laplace equation)
Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded domain, $p \in (1, +\infty)$, and $f^* \in (W_0^{1,p}(\Omega))^*$. Then, a function $u \in W_0^{1,p}(\Omega)$ is called weak solution of the p-Laplace equation *if for every $v \in W_0^{1,p}(\Omega)$, it holds that*

$$\int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \,\mathrm{d}x = \langle f^*, v\rangle_{W_0^{1,p}(\Omega)}\,.$$

Remark 2.20 (properties of $W^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$)
Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded domain and $p \in (1, +\infty)$. Then, the spaces $(W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}(\Omega)})$ and $(W_0^{1,p}(\Omega), \|\cdot\|_{W_0^{1,p}(\Omega)})$, where

$$\begin{aligned} \|\cdot\|_{W^{1,p}(\Omega)} &:= (\|\cdot\|_{L^p(\Omega)}^p + \|\nabla\cdot\|_{(L^p(\Omega))^d})^{\frac{1}{p}}\,, \\ \|\cdot\|_{W_0^{1,p}(\Omega)} &:= (\|\cdot\|_{L^p(\Omega)}^p + \|\nabla\cdot\|_{(L^p(\Omega))^d})^{\frac{1}{p}}\,, \end{aligned}$$

form separable and uniformly convex (and, thus, reflexive) Banach spaces (cf. [Ada75, Thm. 3.3 & Thm. 3.6] or Exercise Sheet 4, Exercise 3).

Remark 2.21 (characterizations of $(W^{1,p}(\Omega))^*$ and $(W_0^{1,p}(\Omega))^*$)
Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded domain and $p \in (1, +\infty)$. Then, we have that (cf. [Ada75, Thm. 3.9 & Thm. 3.12] or Exercise Sheet 4, Exercise 4):

$$\begin{aligned} (W^{1,p}(\Omega))^* &= \left\{\left(v \mapsto \int_\Omega f\, v \,\mathrm{d}x + \int_\Omega \mathbf{F}\cdot\nabla v\,\mathrm{d}x\right) \,\middle|\, f \in L^{p'}(\Omega)\,,\ \mathbf{F} \in (L^{p'}(\Omega))^d\right\}\,, \\ (W_0^{1,p}(\Omega))^* &= \left\{\left(v \mapsto \int_\Omega \mathbf{F}\cdot\nabla v\,\mathrm{d}x\right) \,\middle|\, \mathbf{F} \in (L^{p'}(\Omega))^d\right\}\,, \end{aligned}$$

where $p' := \frac{p}{p-1}$. With a slight abuse of notation, the above expresses that

$$\begin{aligned} (W^{1,p}(\Omega))^* &= L^{p'}(\Omega) + \operatorname{div}((L^{p'}(\Omega))^d)\,, \\ (W_0^{1,p}(\Omega))^* &= \operatorname{div}((L^{p'}(\Omega))^d)\,. \end{aligned}$$

Theorem 2.22 (well-posedness of the weak formulation of the p-Laplace equation) *Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded domain and $p \in (1, +\infty)$. Then, the operator $A\colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$, for every $u, v \in W_0^{1,p}(\Omega)$ defined by*

$$\langle Au, v\rangle_{W_0^{1,p}(\Omega)} := \int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \,\mathrm{d}x\,,$$

is well-defined, bounded, continuous, strictly monotone, coercive, bijective, and its inverse $A^{-1}\colon (W_0^{1,p}(\Omega))^* \to W_0^{1,p}(\Omega)$[6] is strictly monotone, bounded, and demi-continuous.

In other words, for every $f^* \in (W_0^{1,p}(\Omega))^*$, there exists a unique $u \in W_0^{1,p}(\Omega)$ solution of the weak formulation of the p-Laplace equation (cf. Definition 2.19), which depends demi-continuous on the data, i.e., if $(f_n^*)_{n\in\mathbb{N}} \subseteq (W_0^{1,p}(\Omega))^*$ is a sequence such that

$$f_n^* \to f^* \quad \text{in } (W_0^{1,p}(\Omega))^* \quad (n \to \infty)\,,$$

for the corresponding sequence of solutions $(u_n)_{n\in\mathbb{N}} := (A^{-1}f_n^*)_{n\in\mathbb{N}} \subseteq W_0^{1,p}(\Omega)$ of the weak formulation of the p-Laplace equation (cf. Definition 2.19), it follows that

$$u_n \rightharpoonup u \quad \text{in } W_0^{1,p}(\Omega) \quad (n \to \infty)\,.$$

Remark 2.23 (Well-posedness in the sense of Hadamard) *(i) A mathematical problem is called well-posed in the sense of Hadamard if the following three conditions are met:*

(a) For given data, the mathematical problem has at least *one solution;*
(b) For given data, the mathematical problem has at most *one solution;*
(c) The solution of the mathematical problems depends continuously *on the data.*

(ii) Theorem 2.22 shows that the weak formulation of the p-Laplace equation (cf. Definition 2.22) is well-posed in the sense of Hadamard.

Proof (of Theorem 2.22). *1. Well-definedness:* Let $u \in W_0^{1,p}(\Omega)$ be fixed, but arbitrary. Then, for every $v \in W_0^{1,p}(\Omega)$, by the Hölder inequality and due to $p'(p-1) = p$, we have that

$$\begin{aligned}
\left|\int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \,\mathrm{d}x\right| &\le \int_\Omega |\nabla u|^{p-1}|\nabla v| \,\mathrm{d}x \\
&\le \bigg(\int_\Omega |\nabla u|^{(p-1)p'} \,\mathrm{d}x\bigg)^{\frac{1}{p'}} \bigg(\int_\Omega |\nabla u|^{p} \,\mathrm{d}x\bigg)^{\frac{1}{p}} \\
&= \bigg(\int_\Omega |\nabla u|^{p} \,\mathrm{d}x\bigg)^{\frac{p-1}{p}} \bigg(\int_\Omega |\nabla u|^{p} \,\mathrm{d}x\bigg)^{\frac{1}{p}} \\
&= \|\nabla u\|_{L^p(\Omega)}^{p-1}\|\nabla v\|_{L^p(\Omega)} \\
&\le \|u\|_{W_0^{1,p}(\Omega)}^{p-1}\|v\|_{W_0^{1,p}(\Omega)}\,,
\end{aligned}$$

which implies that

$$\bigg(v \mapsto \int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \,\mathrm{d}x\bigg) \in (W_0^{1,p}(\Omega))^*\,.$$

In other words, the operator $A\colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ is well-defined.

[6] In this case, we omit the canonical mapping $j_{W_0^{1,p}(\Omega)}\colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^{**}$ in favor of readability.

2. Boundedness: Due to the above inequality and the definition of the operator norm (*cf.* Proposition 1.2), for every $u \in W_0^{1,p}(\Omega)$, we have that

$$\begin{aligned}\|Au\|_{(W_0^{1,p}(\Omega))^*} &= \sup_{\|v\|_{W_0^{1,p}(\Omega)}\leq 1} |\langle Au, v\rangle_{W_0^{1,p}(\Omega)}| \\ &\leq \|u\|_{W_0^{1,p}(\Omega)}^{p-1}\,.\end{aligned}$$

In other words, the operator $A\colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ is bounded.

3. Continuity: In order to prove the continuity of $A\colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ we need the Vitali convergence theorem:

Theorem 2.24 (Vitali)
Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded measurable set and let $p \in [1, +\infty)$. Then, for a sequence $(f_n)_{n\in\mathbb{N}} \subseteq L^p(\Omega)$ and a function $f \in L^p(\Omega)$, the following statements are equivalent:
(i) $f_n \to f$ *in* $L^p(\Omega)$ $(n \to \infty)$*;*
(ii) The following two conditions are satisfied:
  (a) $f_n \to f$ *in measure in* Ω $(n \to \infty)$*;*
  (b) The sequence $(f_n)_{n\in\mathbb{N}} \subseteq L^p(\Omega)$ is p-uniformly integrable, *i.e., for every $\varepsilon > 0$, there exists $\delta > 0$ such that for a measurable set $\omega \subseteq \Omega$, from $\lambda^d(\omega) \leq \delta$, it follows that*

$$\sup_{n\in\mathbb{N}} \int_\omega |f_n|^p \,\mathrm{d}x \leq \varepsilon\,.$$

Proof. See [Els09, Kap. VI., §5, Satz 5.6, S. 262]. □

Let $(u_n)_{n\in\mathbb{N}} \subseteq W_0^{1,p}(\Omega)$ be a sequence and $u \in W_0^{1,p}(\Omega)$ a function such that

$$u_n \to u \quad \text{in } W_0^{1,p}(\Omega) \quad (n \to \infty)\,.$$

Then, it holds that:
(a) By the inverse direction of the dominated convergence theorem, there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}} \subseteq W_0^{1,p}(\Omega)$ such that

$$\nabla u_{n_k} \to \nabla u \quad \text{a.e. in } \Omega \quad (k \to \infty)\,.$$

Since the mapping $(t \mapsto |t|^{p-2}t)\colon \mathbb{R}^d \to \mathbb{R}^d$ is continuous (*cf.* Exercise Sheet 2, Exercise 3), we infer that

$$|\nabla u_{n_k}|^{p-2}\nabla u_{n_k} \to |\nabla u|^{p-2}\nabla u \quad \text{a.e. in } \Omega \quad (k \to \infty)\,.$$

(b) By the Vitali convergence theorem (*cf.* Theorem 2.24), the sequence $(\nabla u_n)_{n\in\mathbb{N}} \subseteq (L^p(\Omega))^d$ is p-uniformly integrable, *i.e.*, for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every (Lebesgue) measurable set $\omega \subseteq \Omega$ with $\lambda^d(\omega) \leq \delta$, it follows that

$$\sup_{n\in\mathbb{N}} \int_\omega |\nabla u_n|^p \,\mathrm{d}x \leq \varepsilon\,.$$

Then, the sequence $(|\nabla u_n|^{p-2}\nabla u_n)_{n\in\mathbb{N}} \subseteq (L^{p'}(\Omega))^d$ is p'-uniformly integrable since for every (Lebesgue) measurable set $\omega \subseteq \Omega$, due to $(p-1)p' = p$, it holds that

$$\int_\Omega \big||\nabla u_n|^{p-2}\nabla u_n\big|^{p'}\,\mathrm{d}x = \int_\Omega |\nabla u_n|^p\,\mathrm{d}x\,.$$

In summary, the Vitali convergence theorem (*cf.* Theorem 2.24) yields that

$$|\nabla u_{n_k}|^{p-2}\nabla u_{n_k} \to |\nabla u|^{p-2}\nabla u \quad \text{in } (L^{p'}(\Omega))^d \quad (k\to\infty)\,.$$

The subsequence convergence principle (*cf.* Lemma 1.17) further yields that

$$|\nabla u_n|^{p-2}\nabla u_n \to |\nabla u|^{p-2}\nabla u \quad \text{in } (L^{p'}(\Omega))^d \quad (n\to\infty)\,.$$

Eventually, by the definition of the operator norm (*cf.* Proposition 1.2) and the Hölder inequality, we conclude that

$$\begin{aligned}\|Au_n - Au\|_{(W_0^{1,p}(\Omega))^*} &= \sup_{\|v\|_{W_0^{1,p}(\Omega)}\le 1} |\langle Au_n - Au, v\rangle_{W_0^{1,p}(\Omega)}| \\ &\le \||\nabla u_n|^{p-2}\nabla u_n - |\nabla u|^{p-2}\nabla u\|_{L^{p'}(\Omega)} \to 0 \quad (n\to\infty)\,.\end{aligned}$$

In other words, the operator $A\colon W_0^{1,p}(\Omega)\to (W_0^{1,p}(\Omega))^*$ is continuous.

<u>*4. Strict monotonicity:*</u>

<u>*4.1 Monotonicity:*</u> For every $u, v \in W_0^{1,p}(\Omega)$, it holds that

$$\begin{aligned}\langle Au - Av, u - v\rangle_{W_0^{1,p}(\Omega)} &= \int_\Omega \underbrace{\big(|\nabla u|^{p-2}\nabla u - |\nabla v|^{p-2}\nabla v\big)\cdot(\nabla u - \nabla v)}_{\ge 0}\,\mathrm{d}x \\ &\ge 0\,,\end{aligned}$$

where we used that $(t \to |t|^{p-2}t)\colon \mathbb{R}^d \to \mathbb{R}^d$ is monotone (*cf.* Exercise Sheet 2, Exercise 3).

<u>*4.2 Strict Monotonicity:*</u> For $u, v \in W_0^{1,p}(\Omega)$, from

$$\begin{aligned}0 &= \langle Au - Av, u - v\rangle_{W_0^{1,p}(\Omega)} \\ &= \int_\Omega \big(|\nabla u|^{p-2}\nabla u - |\nabla v|^{p-2}\nabla v\big)\cdot(\nabla u - \nabla v)\,\mathrm{d}x\,,\end{aligned}$$

it follows that

$$\big(|\nabla u|^{p-2}\nabla u - |\nabla v|^{p-2}\nabla v\big)\cdot(\nabla u - \nabla v) = 0 \quad \text{a.e. in } \Omega\,.$$

Due to the d-monotonicity and, thus, strict monotonicity of $(t \to |t|^{p-2}t)\colon \mathbb{R}^d \to \mathbb{R}^d$ (*cf.* Exercise Sheet 2, Exercise 3), we find that $\nabla u = \nabla v$ a.e. Ω, which, by the Poincaré inequality, *i.e.*, for every $w \in W_0^{1,p}(\Omega)$, we have that

$$\|w\|_{L^p(\Omega)} \le c_{\mathrm{P}}\,\|\nabla w\|_{L^p(\Omega)}\,, \tag{PI}$$

where $c_{\mathrm{P}} > 0$ denotes the so-called Poincaré constant, implies that

$$0 \leq \|u - v\|_{L^p(\Omega)} \leq c_{\mathrm{P}} \, \|\nabla(u - v)\|_{L^p(\Omega)} = 0 \,,$$

that $u = v$ a.e. in Ω. In other words, the operator $A \colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ is strictly monotone.

5. *Coercivity:* For every $u \in W_0^{1,p}(\Omega)$, we have that

$$\begin{aligned}
\langle Au, u \rangle_{W_0^{1,p}(\Omega)} &= \int_\Omega |\nabla u|^{p-2} \nabla u \cdot \nabla u \, \mathrm{d}x \\
&= \int_\Omega |\nabla u|^p \, \mathrm{d}x \\
&= \|\nabla u\|_{L^p(\Omega)}^p \,.
\end{aligned}$$

By means of the Poincaré inequality (PI), for every $u \in W_0^{1,p}(\Omega)$, we have that

$$\begin{aligned}
\|u\|_{W_0^{1,p}(\Omega)}^p &= \|u\|_{L^p(\Omega)}^p + \|\nabla u\|_{L^p(\Omega)}^p \\
&\leq (c_{\mathrm{P}}^p + 1) \|\nabla u\|_{L^p(\Omega)}^p \,.
\end{aligned}$$

Hence, for every $u \in W_0^{1,p}(\Omega)$, we find that

$$\langle Au, u \rangle_{W_0^{1,p}(\Omega)} \geq \frac{1}{1 + c_{\mathrm{P}}^p} \|u\|_{W_0^{1,p}(\Omega)}^p \,,$$

Therefore, due to $p > 1$, we conclude that

$$\frac{\langle Au, u \rangle_{W_0^{1,p}(\Omega)}}{\|u\|_{W_0^{1,p}(\Omega)}} \geq \frac{1}{1 + c_{\mathrm{P}}^p} \|u\|_{W_0^{1,p}(\Omega)}^{p-1} \to +\infty \quad (\|u\|_{W_0^{1,p}(\Omega)} \to +\infty) \,.$$

In other words, the operator $A \colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ is coercive.

In summary, we have shown that the operator $A \colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ is well-defined, bounded, continuous, strictly monotone, and coercive. Therefore, since the space $W_0^{1,p}(\Omega)$ is separable and reflexive (*cf.* Remark 2.20), the main theorem on monotone operators (*cf.* Theorem 2.13) yields that $A \colon W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ is bijective and that its inverse $A^{-1} \colon (W_0^{1,p}(\Omega))^* \to W_0^{1,p}(\Omega)$ is strictly monotone, bounded, and demi-continuous. $\square$

2.5.2 p-Stokes equations

As a second application of the main theorem on monotone operators (*cf.* Theorem 2.13), we obtain the well-posedness of the weak formulation of the p-Stokes equations, which model the laminar (*i.e.*, non-turbulent) flow of a *non-Newtonian fluid* through a bounded domain under the influence of an external force. More precisely, the p-Stokes equations, for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, occupied by the fluid and a given external force $\mathbf{f}\colon \Omega \to \mathbb{R}^d$ acting on the fluid, seek for a velocity vector field $\mathbf{v} \coloneqq (v_1, \dots, v_d)^\top \colon \Omega \to \mathbb{R}^d$ and a kinematic pressure $\pi\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\operatorname{div}(\mathbf{S}(\mathbf{D}\mathbf{v})) + \nabla\pi &= \mathbf{f} && \text{in } \Omega\,, \\ \operatorname{div} \mathbf{v} &= 0 && \text{in } \Omega\,, \\ \mathbf{v} &= \mathbf{0} && \text{on } \partial\Omega\,. \end{aligned} \tag{2.25}$$

In this subsection, we restrict to the case of a power-law fluid (*cf.* Introduction 0), *i.e.*, $\mathbf{S}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$, for every $\mathbf{A} \in \mathbb{R}^{d\times d}$, is defined by

$$\mathbf{S}(\mathbf{A}) \coloneqq \nu_0\,(\delta + |\mathbf{A}|)^{p-2}\mathbf{A} \quad \text{in } \mathbb{R}^{d\times d}\,,$$

where $\delta \geq 0$, $\nu_0 > 0$, and $p \in (1, +\infty)$ the power-law index.

Definition 2.26 (weak formulation of the p-Stokes equations)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain, $p \in (1, +\infty)$, and $\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^*$. Then, a couple $(\mathbf{v}, \pi)^\top \in (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$, where

$$L_0^{p'}(\Omega) \coloneqq \left\{ \eta \in L^{p'}(\Omega) \;\middle|\; \int_\Omega \eta \,\mathrm{d}x = 0 \right\}\,,$$

is called weak solution of the p-Stokes equations *if for every $(\boldsymbol{\varphi}, \eta)^\top \in (W_0^{1,p}(\Omega))^d \times L^{p'}(\Omega)$, it holds that*[7]

$$\begin{aligned} \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi} \,\mathrm{d}x - \int_\Omega \pi \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x &= \langle \mathbf{f}^*, \boldsymbol{\varphi} \rangle_{(W_0^{1,p}(\Omega))^d}\,, \\ \int_\Omega \eta \operatorname{div} \mathbf{v} \,\mathrm{d}x &= 0\,. \end{aligned}$$

Remark 2.27 (zero mean condition)
The zero mean condition imposed on the pressure (i.e., $\int_\Omega \pi \,\mathrm{d}x = 0$) is needed to enforce its uniqueness. Without this condition, for every constant $c \in \mathbb{R}$, the function $\pi + c \in L^{p'}(\Omega)$ is a solution as well since, by the Gauss theorem, for every $\boldsymbol{\varphi} \in (W_0^{1,p}(\Omega))^d$, it holds that

$$\begin{aligned} \int_\Omega (\pi + c) \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x &= \int_\Omega \pi \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x + c \int_\Omega \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x \\ &= \int_\Omega \pi \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x + c \cancel{\int_{\partial\Omega} \boldsymbol{\varphi} \cdot \boldsymbol{\nu} \,\mathrm{d}s}\,, \end{aligned}$$

where $\mathrm{d}s$ is the surface measure on the (topological) boundary $\partial\Omega$ and $\boldsymbol{\nu}\colon \partial\Omega \to \mathbb{R}^d$ is the outwards unit vector field, which exists since $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, is a bounded Lipschitz domain.

[7] Here, for two matrices $\mathbf{A} = (A_{ij})_{i,j\in\{1,\dots,d\}}, \mathbf{B} = (B_{ij})_{i,j\in\{1,\dots,d\}} \in \mathbb{R}^{d\times d}$, the *Frobenius inner product* is defined by $\mathbf{A} : \mathbf{B} \coloneqq \sum_{i,j=1,\dots,d} A_{ij}B_{ij}$.

One major difficulty of the mathematical treatment of the p-Stokes equations (2.25) (compared to the p-Laplace equation (2.18)) consists in its non-linear saddle point structure, *i.e.*, it cannot directly be re-written as an operator equation with a coercive operator (*cf.* Theorem 2.13). A remedy is then provides by initially passing to the hydro-mechanical formulation of the p-Stokes equations (2.25), which initially neglects the pressure via restricting the energy space for the velocity vector field to incompressible (*i.e.*, divergence-free) vector fields. The hydro-mechanical formulation of the p-Stokes equations (2.25), in turn, can be re-written as an operator equation with a coercive operator to which the main theorem on pseudo-monotone operators (*cf.* Theorem 2.13) can be applied to.

Lemma 2.28 (hydro-mechanical formulation of the p-Stokes equations)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain, $p \in (1, +\infty)$, and $\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^*$. Then, for a vector field $\mathbf{v} \in (W_0^{1,p}(\Omega))^d$, the following two statements are equivalent:

(i) There exists a function $\pi \in L_0^{p'}(\Omega)$ such that $(\mathbf{v}, \pi)^\top \in (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$ is a weak solution of the p-Stokes equations (cf. Definition 2.26);

(ii) The vector field $\mathbf{v} \in (W_0^{1,p}(\Omega))^d$ satisfies $\mathbf{v} \in V_p$, where

$$V_p \coloneqq \left\{\boldsymbol{\varphi} \in (W_0^{1,p}(\Omega))^d \mid \operatorname{div} \boldsymbol{\varphi} = 0 \text{ in } \Omega\right\},$$

and is a weak solution of the hydro-mechanical p-Stokes equations, i.e., for every $\boldsymbol{\varphi} \in V_p$, it holds that

$$\int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi} \,\mathrm{d}x = \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} .$$

Remark 2.29 (properties of V_p)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in (1, +\infty)$. Then, the space $(V_p, \|\cdot\|_{V_p})$, where

$$\|\cdot\|_{V_p} \coloneqq \left(\|\cdot\|_{L^p(\Omega)}^p + \|\nabla(\cdot)\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}},$$

is a closed subspace of $(W_0^{1,p}(\Omega))^d$ and, thus, separable and uniformly convex (and, thus, reflexive).

The stated equivalence between the weak formulation (*cf.* Definition 2.26) and the hydro-mechanical formulation (*cf.* Lemma 2.28(ii)) of the p-Stokes equations in Lemma 2.28 is a consequence of the de Rham lemma, which also allows us to establish the existence of the pressure solving weak formulation of the p-Stokes equations (*cf.* Definition 2.26) after establishing the existence of a velocity vector field solving hydro-mechanical formulation of the p-Stokes equations (*cf.* Lemma 2.28(ii)).

Lemma 2.30 (de Rham)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in (1,+\infty)$. Then, for every $\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^*$, from

$$\langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} = 0 \quad \text{for all } \boldsymbol{\varphi} \in V_p\,,$$

it follows the existence of $\pi \in L_0^{p'}(\Omega)$ such that

$$\langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} = -\int_\Omega \pi \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x \quad \text{for all } \boldsymbol{\varphi} \in (W_0^{1,p}(\Omega))^d\,.$$

Proof. See [BF13, Thm. IV.2.3.], for the special case $p = 2$. □

Remark 2.31 (alternative interpretation of the de Rham lemma (*cf.* Lemma 2.30))
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in (1,+\infty)$. Then, the de Rham lemma (cf. Lemma 2.30) states that the distributional gradient $\nabla\colon L^{p'}(\Omega) \to ((W_0^{1,p}(\Omega))^d)^*$, *for every $\eta \in L^{p'}(\Omega)$ defined by*

$$\langle \nabla\eta, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} := -\int_\Omega \eta \operatorname{div} \boldsymbol{\varphi} \,\mathrm{d}x \quad \text{for all } \boldsymbol{\varphi} \in (W_0^{1,p}(\Omega))^d\,, \tag{2.32}$$

yields an isomorphism between $L_0^{p'}(\Omega)$ and the annihilator

$$(V_p)^\circ := \big\{\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^* \mid \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} = 0 \text{ for all } \boldsymbol{\varphi} \in V_p\big\}\,,$$

i.e., we have that

$$L_0^{p'}(\Omega) \overset{\nabla}{\cong} (V_p)^\circ\,.$$

Proof (of Lemma 2.28). ad "⇒". Let $\pi \in L_0^{p'}(\Omega)$ be a function such that for every $(\boldsymbol{\varphi},\eta)^\top \in (W_0^{1,p}(\Omega))^d \times L^{p'}(\Omega)$, it holds that

$$\begin{aligned} &\text{(I)} && \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x - \int_\Omega \pi \operatorname{div}\boldsymbol{\varphi}\,\mathrm{d}x = \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d}\,,\\ &\text{(II)} && \int_\Omega \eta \operatorname{div}\mathbf{v}\,\mathrm{d}x = 0\,. \end{aligned}$$

Then, from (II), it follows that $\mathbf{v} \in V_p$, while from (I), for every $\boldsymbol{\varphi} \in V_p$, it follows that

$$\int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x = \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d}\,.$$

In other words, the vector field $\mathbf{v} \in V_p$ is a solution of the hydro-mechanical formulation of the p-Stokes equations (*cf.* Lemma 2.28(ii)).

ad "$\Leftarrow$". Let $\mathbf{v} \in V_p$ is a solution of the hydro-mechanical formulation of the p-Stokes equations (*cf.* Lemma 2.28(ii)), *i.e.*, for every $\boldsymbol{\varphi} \in V_p$, it follows that

$$\int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x = \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d}\,.$$

Then, we have that

$$\Big(\boldsymbol{\varphi} \mapsto \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} - \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\Big) \in (V_p)^\circ\,.$$

Therefore, the de Rham lemma (*cf.* Lemma 2.30) yields a function $\pi \in L_0^{p'}(\Omega)$ such that for every $\boldsymbol{\varphi} \in (W_0^{1,p}(\Omega))^d$, it holds that

$$-\int_\Omega \pi \operatorname{div}\boldsymbol{\varphi}\,\mathrm{d}x = \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d} - \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\,.$$

In other words, the couple $(\mathbf{v}, \pi)^\top \in (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$ is a solution of the weak formulation of the p-Stokes equations (*cf.* Definition 2.26). $\square$

Motivated by Lemma 2.28, in order to establish the well-posedness of the weak formulation of the p-Stokes equations (*cf.* Definition 2.26), we first establish the well-posedness of the hydro-mechanical formulation of the p-Stokes equations (*cf.* Lemma 2.28(ii)), which will be a consequence of the main theorem on monotone operators (*cf.* Theorem 2.13).

Theorem 2.33 (well-posedness of the hydro-mechanical form. of the p-Stokes equations) *Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in (1, +\infty)$. Then, the operator $\widehat{S}\colon V_p \to (V_p)^*$, for every $\mathbf{v}, \boldsymbol{\varphi} \in V_p$ defined by*

$$\langle \widehat{S}\mathbf{v}, \boldsymbol{\varphi}\rangle_{V_p} := \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\,,$$

is well-defined, bounded, strictly monotone, continuous, coercive, bijective, and its inverse $\widehat{S}^{-1}\colon (V_p)^* \to V_p$ is strictly monotone, bounded, and demi-continuous.

In other words, for every $\mathbf{f}^* \in (V_p)^*$, there exists a unique solution hydro-mechanical formulation of the p-Stokes equations (cf. Lemma 2.28(ii)) $\mathbf{v} \in V_p$, which depends demi-continuous on the data, i.e., if $(\mathbf{f}_n^*)_{n\in\mathbb{N}} \subseteq (V_p)^*$ is a sequence such that

$$\mathbf{f}_n^* \to \mathbf{f}^* \quad \text{in } (V_p)^* \quad (n \to \infty)\,,$$

for the corresponding sequence of solutions $(\mathbf{v}_n)_{n\in\mathbb{N}} := (\widehat{S}^{-1}\mathbf{f}_n^*)_{n\in\mathbb{N}} \subseteq V_p$ of the hydro-mechanical formulation of the p-Stokes equations (cf. Definition 2.26), it follows that

$$\mathbf{v}_n \rightharpoonup \mathbf{v} \quad \text{in } V_p \quad (n \to \infty)\,.$$

Proof. <u>*1. Well-posedness:*</u> Let $\mathbf{v} \in V_p$ be fixed, but arbitrary. Then, for every $\boldsymbol{\varphi} \in V_p$, by the Hölder inequality, $(\delta + |\mathbf{D}\mathbf{v}|)^{p-2}|\mathbf{D}\mathbf{v}| \leq (\delta + |\mathbf{D}\mathbf{v}|)^{p-1}$ a.e. in Ω, and $p'(p-1) = p$, we have that

$$
\begin{aligned}
\bigg|\int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\bigg| &\leq \nu_0 \int_\Omega (\delta + |\mathbf{D}\mathbf{v}|)^{p-2}|\mathbf{D}\mathbf{v}||\mathbf{D}\boldsymbol{\varphi}|\,\mathrm{d}x \\
&\leq \nu_0 \int_\Omega (\delta + |\mathbf{D}\mathbf{v}|)^{p-1}|\mathbf{D}\boldsymbol{\varphi}|\,\mathrm{d}x \\
&\leq \nu_0 \left(\int_\Omega (\delta + |\mathbf{D}\mathbf{v}|)^{(p-1)p'}\,\mathrm{d}x\right)^{\frac{1}{p'}} \left(\int_\Omega |\mathbf{D}\boldsymbol{\varphi}|^p\,\mathrm{d}x\right)^{\frac{1}{p}} \\
&\leq \nu_0 \left(\int_\Omega (\delta + |\mathbf{D}\mathbf{v}|)^{p}\,\mathrm{d}x\right)^{\frac{p-1}{p}} \left(\int_\Omega |\mathbf{D}\boldsymbol{\varphi}|^p\,\mathrm{d}x\right)^{\frac{1}{p}} \\
&= \nu_0 \,\|\delta + |\mathbf{D}\mathbf{v}|\|_{L^p(\Omega)}^{p-1}\|\mathbf{D}\boldsymbol{\varphi}\|_{L^p(\Omega)} \\
&\leq \nu_0 \,(\delta\,(\lambda^d(\Omega))^{\frac{1}{p}} + \|\mathbf{v}\|_{V_p})^{p-1}\|\boldsymbol{\varphi}\|_{V_p}\,,
\end{aligned}
$$

which implies that

$$
\left(\boldsymbol{\varphi} \mapsto \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\right) \in (V_p)^*\,.
$$

In other words, the operator $\widehat{S}\colon V_p \to (V_p)^*$ is well-defined.

<u>*2. Boundedness:*</u> Due to the above inequality and the definition of the operator norm (*cf.* Proposition 1.2), for every $\mathbf{v} \in V_p$, we have that

$$
\begin{aligned}
\|\widehat{S}\mathbf{v}\|_{(V_p)^*} &:= \sup_{\|\boldsymbol{\varphi}\|_{V_p} \leq 1} |\langle \widehat{S}\mathbf{v}, \boldsymbol{\varphi}\rangle_{V_p}| \\
&\leq \nu_0\,(\delta\,(\lambda^d(\Omega))^{\frac{1}{p}} + \|\mathbf{v}\|_{V_p})^{p-1}\,.
\end{aligned}
$$

In other words, the operator $\widehat{S}\colon V_p \to (V_p)^*$ is bounded.

<u>*3. Continuity:*</u> Let $(\mathbf{v}_n)_{n\in\mathbb{N}} \subseteq V_p$ be a sequence and $\mathbf{v} \in V_p$ a vector field such that

$$
\mathbf{v}_n \to \mathbf{v} \quad \text{in } V_p \quad (n \to \infty)\,.
$$

Then, it holds that:

(a) By the inverse direction of the dominated convergence theorem, there exists a subsequence $(\mathbf{v}_{n_k})_{k\in\mathbb{N}} \subseteq V_p$ such that

$$
\mathbf{D}\mathbf{v}_{n_k} \to \mathbf{D}\mathbf{v} \quad \text{a.e. in } \Omega \quad (k \to \infty)\,.
$$

Since the mapping $\mathbf{S}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is continuous, we infer that

$$
\mathbf{S}(\mathbf{D}\mathbf{v}_{n_k}) \to \mathbf{S}(\mathbf{D}\mathbf{v}) \quad \text{a.e. in } \Omega \quad (k \to \infty)\,.
$$

(b) By the Vitali convergence theorem (*cf.* Theorem 2.24), the sequence $(\mathbf{D}\mathbf{v}_n)_{n\in\mathbb{N}} \subseteq (L^p(\Omega))^{d\times d}$ is p-uniformly integrable. Then, the sequence $(\mathbf{S}(\mathbf{D}\mathbf{v}_n))_{n\in\mathbb{N}} \subseteq (L^{p'}(\Omega))^{d\times d}$ is p'-uniformly integrable since for every (Lebesgue) measurable set $\omega \subseteq \Omega$, due to $(p-1)p' = p$, it holds that

$$\int_\omega |\mathbf{S}(\mathbf{D}\mathbf{v}_n)|^{p'}\,\mathrm{d}x \le \int_\omega (\delta + |\mathbf{D}\mathbf{v}_n|)^{(p-2)p'}\,\mathrm{d}x = \int_\omega (\delta + |\mathbf{D}\mathbf{v}_n|)^{p}\,\mathrm{d}x\,.$$

In summary, the Vitali convergence theorem (*cf.* Theorem 2.24) yields that

$$\mathbf{S}(\mathbf{D}\mathbf{v}_{n_k}) \to \mathbf{S}(\mathbf{D}\mathbf{v}) \quad \text{in } (L^{p'}(\Omega))^{d\times d} \quad (k\to\infty)\,.$$

The subsequence convergence principle (*cf.* Lemma 1.17) further yields that

$$\mathbf{S}(\mathbf{D}\mathbf{v}_{n}) \to \mathbf{S}(\mathbf{D}\mathbf{v}) \quad \text{in } (L^{p'}(\Omega))^{d\times d} \quad (n\to\infty)\,.$$

Eventually, by the definition of the operator norm (*cf.* Proposition 1.2) and the Hölder inequality, we conclude that

$$\begin{aligned}\|\widehat{S}\mathbf{v}_n - \widehat{S}\mathbf{v}\|_{(V_p)^*} &= \sup_{\|\boldsymbol{\varphi}\|_{V_p}\le 1} |\langle \widehat{S}\mathbf{v}_n - \widehat{S}\mathbf{v}, \boldsymbol{\varphi}\rangle_{V_p}| \\ &\le \|\mathbf{S}(\mathbf{D}\mathbf{v}_n) - \mathbf{S}(\mathbf{D}\mathbf{v})\|_{L^{p'}(\Omega)} \to 0 \quad (n\to\infty)\,.\end{aligned}$$

In other words, the operator $\widehat{S}\colon V_p \to (V_p)^*$ is continuous.

4. Strict monotonicity:
4.1 Monotonicity: For every $\mathbf{v}, \boldsymbol{\varphi} \in V_p$, it holds that

$$\begin{aligned}\langle \widehat{S}\mathbf{v} - \widehat{S}\boldsymbol{\varphi}, \mathbf{v} - \boldsymbol{\varphi}\rangle_{V_p} &= \int_\Omega \underbrace{\big(\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\boldsymbol{\varphi})\big)\cdot(\mathbf{D}\mathbf{v} - \mathbf{D}\boldsymbol{\varphi})}_{\ge 0}\,\mathrm{d}x \\ &\ge 0\,,\end{aligned}$$

where we used that $\mathbf{S}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is monotone.
4.2 Strict monotonicity: For $\mathbf{v}, \boldsymbol{\varphi} \in V_p$, from

$$\begin{aligned}0 &= \langle \widehat{S}\mathbf{v} - \widehat{S}\boldsymbol{\varphi}, \mathbf{v} - \boldsymbol{\varphi}\rangle_{V_p} \\ &= \int_\Omega \big(\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\boldsymbol{\varphi})\big)\cdot(\mathbf{D}\mathbf{v} - \mathbf{D}\boldsymbol{\varphi})\,\mathrm{d}x\,,\end{aligned}$$

it follows that

$$\big(\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\boldsymbol{\varphi})\big)\cdot(\mathbf{D}\mathbf{v} - \mathbf{D}\boldsymbol{\varphi}) = 0 \quad \text{a.e. in } \Omega\,.$$

Due to the strict monotonicity of $\mathbf{S}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$, we find that $\mathbf{D}\mathbf{v} = \mathbf{D}\boldsymbol{\varphi}$ a.e. Ω, which, by the Korn inequality, *i.e.*, for every $\mathbf{w} \in (W^{1,p}_0(\Omega))^d$, we have that

$$\|\nabla\mathbf{w}\|_{L^p(\Omega)} \le c_{\mathrm{K}}\,\|\mathbf{D}\mathbf{w}\|_{L^p(\Omega)}\,, \tag{KI}$$

where $c_{\mathrm{K}} > 0$ denotes the so-called Korn constant, and the Poincaré inequality (PI), implies that

$$\begin{aligned}0 &\leq \|\mathbf{v}-\boldsymbol{\varphi}\|_{L^p(\Omega)}\\ &\leq c_{\mathrm{P}}\,\|\nabla(\mathbf{v}-\boldsymbol{\varphi})\|_{L^p(\Omega)}\\ &\leq c_{\mathrm{P}}c_{\mathrm{K}}\,\|\mathbf{D}(\mathbf{v}-\boldsymbol{\varphi})\|_{L^p(\Omega)} = 0\,,\end{aligned}$$

that $\mathbf{u}=\boldsymbol{\varphi}$ a.e. in Ω. In other words, the operator $S\colon V_p\to (V_p)^*$ is strictly monotone.

5. *Coercivity:* For every $\mathbf{v}\in V_p$, using that $(\delta+|\mathbf{D}\mathbf{v}|)^{p-2}|\mathbf{D}\mathbf{v}|^2\geq \frac{1}{2}|\mathbf{D}\mathbf{v}|^p-\delta^p$ a.e. in Ω, we have that

$$\begin{aligned}\langle\widehat{S}\mathbf{v},\mathbf{v}\rangle_{V_p} &= \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}):\mathbf{D}\mathbf{v}\,\mathrm{d}x\\ &= \nu_0\int_\Omega (\delta+|\mathbf{D}\mathbf{v}|)^{p-2}|\mathbf{D}\mathbf{v}|^2\,\mathrm{d}x\\ &\geq \nu_0\int_\Omega \left\{\frac{1}{2}|\mathbf{D}\mathbf{v}|^p-\delta^p\right\}\,\mathrm{d}x\\ &= \frac{\nu_0}{2}\|\mathbf{D}\mathbf{v}\|^p_{L^p(\Omega)}-\nu_0\delta^p\lambda^d(\Omega)\,.\end{aligned}$$

Using the Korn inequality (KI) and the Poincaré inequality (PI), for every $\mathbf{v}\in V_p$, we deduce that

$$\langle\widehat{S}\mathbf{v},\mathbf{v}\rangle_{V_p}\geq \frac{\nu_0}{2}\frac{1}{c_{\mathrm{K}}^p+(c_{\mathrm{P}}c_{\mathrm{K}})^p}\|\mathbf{v}\|^p_{V_p}-\nu_0\delta^p\lambda^d(\Omega)\,,$$

Therefore, due to $p>1$, we conclude that

$$\frac{\langle\widehat{S}\mathbf{v},\mathbf{v}\rangle_{V_p}}{\|\mathbf{v}\|_{V_p}}\geq \frac{\nu_0}{2}\frac{1}{c_{\mathrm{K}}^p+(c_{\mathrm{P}}c_{\mathrm{K}})^p}\|\mathbf{v}\|^{p-1}_{V_p}-\frac{\nu_0\delta^p\lambda^d(\Omega)}{\|\mathbf{v}\|_{V_p}}\to+\infty\quad(\|\mathbf{v}\|_{V_p}\to+\infty)\,.$$

In other words, the operator $\widehat{S}\colon V_p\to (V_p)^*$ is coercive.

In summary, we have shown that the operator $\widehat{S}\colon V_p\to(V_p)^*$ is well-defined, bounded, continuous, strictly monotone, and coercive. Therefore, since the space V_p is separable and reflexive (*cf.* Remark 2.29), the main theorem on monotone operators (*cf.* Theorem 2.13) yields that $\widehat{S}\colon V_p\to(V_p)^*$ is bijective and that its inverse $\widehat{S}^{-1}\colon (V_p)^*\to V_p$ is strictly monotone, bounded, and demi-continuous. $\square$

Up to this point, we merely proved the well-posedness of the hydro-mechanical formulation of the p-Stokes equations (*cf.* Lemma 2.28(ii)). By means of Lemma 2.28(ii) and the de Rham lemma (*cf.* Lemma 2.30), we can finally prove the well-posedness of the weak formulation of the p-Stokes equations (*cf.* Definition 2.26).

Corollary 2.34 (well-posedness of the weak formulation of the p-Stokes equations)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in (1, +\infty)$. Then, for every $\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^*$, there exists a unique solution weak formulation of the p-Stokes equations (cf. Definition 2.26) $(\mathbf{v}, \pi)^\top \in (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$, which depends demi-continuous on the data, i.e., if $(\mathbf{f}_n^*)_{n \in \mathbb{N}} \subseteq ((W_0^{1,p}(\Omega))^d)^*$ is a sequence such that

$$\mathbf{f}_n^* \to \mathbf{f}^* \quad \text{in } ((W_0^{1,p}(\Omega))^d)^* \quad (n \to \infty)\,,$$

for the corresponding sequence of solutions $((\mathbf{v}_n, \pi_n)^\top)_{n \in \mathbb{N}} \subseteq (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$ of the weak formulation of the p-Stokes equations (cf. Definition 2.26), it follows that

$$(\mathbf{v}_n, \pi_n)^\top \rightharpoonup (\mathbf{v}, \pi)^\top \quad \text{in } (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega) \quad (n \to \infty)\,.$$

Proof. See Homework 5, Problem 3. $\square$

3 Pseudo-Monotone Operator Theory

We cannot apply the main theorem on monotone operators (*cf.* Theorem 2.13) to the p-Navier–Stokes equations since the convective term $\operatorname{div}(\mathbf{v}\otimes\mathbf{v})$ will not lead to a monotone operator (*cf.* Homework 5, Problem 1). Thus, we need an existence result for infinite-dimensional Banach spaces, which also applies to operators with non-monotone parts. To this end, we recall that the inverse function theorem for strictly monotone functions (*cf.* Theorem 2.1), contained a statement that did not assume strict monotonicity of the operator. More precisely, this statement was the following version of the intermediate value theorem.

Theorem 3.1 (intermediate value theorem)
Let $A\colon \mathbb{R}\to\mathbb{R}$ be a function (or an operator) with the following properties:

(a) <u>Continuity:</u> *For a sequence $(x_n)_{n\in\mathbb{N}}\subseteq\mathbb{R}$ and an element $x\in\mathbb{R}$ from*

$$x_n\to x \quad \text{in } \mathbb{R} \quad (n\to\infty)\,,$$

it follows that

$$Ax_n\to Ax \quad \text{in } \mathbb{R} \quad (n\to\infty)\,.$$

(b) <u>Coercivity:</u> *It holds that $Ax\to\pm\infty$ $(x\to\pm\infty)$.*

Then, then the $A\colon\mathbb{R}\to\mathbb{R}$ operator is surjective.

Proof. See Theorem 2.1(i). □

We observe that in the intermediate value theorem (*cf.* Theorem 3.1), if we drop the strict monotonicity property of the operator $A\colon\mathbb{R}\to\mathbb{R}$ imposed in the inverse function theorem for strictly monotone functions (*cf.* Theorem 2.1), then we loose the bijectivity of $A\colon\mathbb{R}\to\mathbb{R}$ and, thus, the existence and continuity of the inverse $A^{-1}\colon\mathbb{R}\to\mathbb{R}$. However, we still obtain the surjectivity of of $A\colon\mathbb{R}\to\mathbb{R}$, which is at least enough to meet the first requirement of well-posedness in the sense of Hadamard (*cf.* Remark 2.23(i)). Therefore, the ultimate objective of this section of the lecture is to prove a generalization of the intermediate value theorem (*cf.* Theorem 3.1) for infinite-dimensional Banach spaces. In fact, we already generalized Theorem 3.1 to the Euclidean case in Theorem 2.15(i). Similar as for the main theorem on monotone operators (*cf.* Theorem 2.13), we cannot work with the classical notion of continuity and first need to identify a suitable substitute. In fact, we already encountered this substitute in Minty's Trick (*cf.* Lemma 2.8), which showed that every radially continuous and monotone operator has this property. More precisely, we need to replace continuity with the condition (M).

3.1 Brief review of the condition (M)

We recall that, by Minty's Trick (*cf.* Lemma 2.8(ii)), radial continuity together with monotonicity implies that an operator is of type (M). In addition, we recall that being of type (M) was a crucial condition for an operator to be able to perform the passage to the limit with the Galerkin scheme (GS) in the proof of the main theorem of monotone operators (*cf.* Theorem 2.13). Therefore, the natural question that arises is whether we can prove at least the surjectivity statement in the main theorem of monotone operators (*cf.* Theorem 2.13(i)) only imposing that the operator is of type (M) instead of radially continuous and monotone. The answer to this question is partly positive and will be given in the forthcoming sections. Before we do so, let us first start with a brief reminder on operators of type (M) recalling the definition and the most important related results.

Definition 3.2

Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, an operator $A\colon X \to X^*$ is of type (M) *or satisfies the* condition (M) *if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from*

$$\begin{aligned} x_n &\rightharpoonup x && \text{in } X && (n\to\infty)\,, \\ Ax_n &\rightharpoonup x^* && \text{in } X^* && (n\to\infty)\,, \\ \limsup_{n\to\infty}\, &\langle Ax_n, x_n\rangle_X \le \langle x^*, x\rangle_X\,, \end{aligned}$$

it follows that $Ax = x^*$ in X^*.

The condition (M) together with locally boundedness is a consistent generalization of continuity from finite-dimensional to infinite-dimensional Banach spaces.

Remark 3.3 (consistency)

If $\dim X < \infty$, then an operator $A\colon X \to X^*$ is continuous if and only if of type (M) and locally bounded (cf. Homework 5, Problem 2).

The following lemma collects important examples for and properties of operators of type (M).

Lemma 3.4 (important examples for and properties of operators of type (M))

Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, the following statements apply:

(i) If $A\colon X \to X^*$ is monotone and radially continuous, then $A\colon X \to X^*$ is of type (M);
(ii) If $A\colon X \to X^*$ is weakly continuous, then $A\colon X \to X^*$ is of type (M);
(iii) If $A\colon X \to X^*$ is strongly continuous, then $A\colon X \to X^*$ is of type (M);
(iv) If $A\colon X \to X^*$ is of type (M) and $B\colon X \to X^*$ is strongly continuous, then the sum $A + B\colon X \to X^*$ is of type (M).
(v) If $A\colon X \to X^*$ is of type (M) and $B\colon X \to X^*$ is monotone and weakly continuous, then the sum $A + B\colon X \to X^*$ is of type (M).
(vi) If $A\colon X \to X^*$ is of type (M) and locally bounded and if, in addition, X is reflexive, then $A\colon X \to X^*$ is demi-continuous.

Proof. *ad (i).* See Minty's Trick (*cf.* Lemma 2.8(ii)).

ad (ii). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a sequence and $x \in X$ an element such that

$$\begin{aligned} x_n &\rightharpoonup x &&\text{in } X &&(n\to\infty)\,,\\ Ax_n &\rightharpoonup x^* &&\text{in } X^* &&(n\to\infty)\,,\\ &\limsup_{n\to\infty} \langle Ax_n, x_n\rangle_X \le \langle x^*, x\rangle_X\,. \end{aligned}$$

Since $A\colon X\to X^*$ is weakly continuous, we infer that

$$Ax_n \rightharpoonup Ax \quad \text{in } X^* \quad (n\to\infty)\,,$$

which, by the uniqueness of weak limits, implies that $Ax = x^*$ in X^*. In other words, the operator $A\colon X\to X^*$ is of type (M).

ad (iii). Follows from (iv), in the case $A = 0$.

ad (iv). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a sequence and $x \in X$ an element such that

$$\begin{aligned} x_n &\rightharpoonup x &&\text{in } X &&(n\to\infty)\,,\\ (A+B)x_n &\rightharpoonup x^* &&\text{in } X^* &&(n\to\infty)\,,\\ &\limsup_{n\to\infty} \langle (A+B)x_n, x_n\rangle_X \le \langle x^*, x\rangle_X\,. \end{aligned}$$

Since $B\colon X\to X^*$ is strongly continuous, we infer that

$$Bx_n \to Bx \quad \text{in } X^* \quad (n\to\infty)\,,$$

which implies that

$$Ax_n \rightharpoonup x^* - Bx \quad \text{in } X^* \quad (n\to\infty)\,,$$

and

$$\begin{aligned} \limsup_{n\to\infty} \langle Ax_n, x_n\rangle_X &= \limsup_{n\to\infty} \langle (A+B)x_n, x_n\rangle_X + \lim_{n\to\infty} \langle -Bx_n, x_n\rangle_X \\ &\le \langle x^* - Bx, x\rangle_X\,. \end{aligned}$$

Since $A\colon X\to X^*$ is of type (M), we conclude that $Ax = x^* - Bx$ in X^*, *i.e.*, $(A+B)x = x^*$ in X^*. In other words, the operator $A\colon X\to X^*$ is of type (M).

ad (v). See Homework 6, Problem 2.

ad (vi). Follows analogously to Minty's Trick (*cf.* Lemma 2.8(ii)) (*cf.* Homework 6, Problem 3). □

The following famous examples reveals the greatest weakness of the condition (M). The condition (M) is not stable under summation.

Example 3.5 (type (M) + type (M) $\not\Leftarrow$ type (M), *cf.* [Bré67])
Let $(X, (\cdot,\cdot)_X)$ be a Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{N}} \subseteq X$, i.e., $(e_n, e_m)_X := \delta_{nm}$ for all $n, m \in \mathbb{N}$. Then, the following statements apply:

(i) The operator $A\colon X \to X^*$, for every $x \in X$, defined by

$$Ax := -R_X x \quad \text{in } X^*\,,$$

where $R_X\colon X \to X^*$ denotes the Riesz isomorphism, is weakly continuous and, thus, of type (M) (cf. Lemma 3.4(ii)).

(ii) The operator $B\colon X \to X^*$, for every $x \in X$, defined by

$$\begin{aligned} Bx &:= R_X\Big(\operatorname*{arg\,min}_{y\in \overline{B}_1^X(0)} \|x-y\|_X^2\Big)\,,\\ &\overset{\text{Exercise Sheet 6, Exercise 3}}{\Leftrightarrow} \quad (R_X^{-1}Bx - x, R_X^{-1}Bx - y)_X \le 0 \quad \text{for all } y \in \overline{B}_1^X(0)\,, \end{aligned}$$

is monotone and a contraction (i.e., Lipschitz continuous with Lipschitz constant less or equal to 1), since, for every $x, y \in X$, we have that

$$\begin{aligned} \langle Bx - By, x-y\rangle_X &= (R_X^{-1}Bx - R_X^{-1}By, x-y)_X \\ &= (R_X^{-1}Bx - R_X^{-1}By, x - R_X^{-1}Bx)_X \\ &\quad + \|R_X^{-1}Bx - R_X^{-1}By\|_X^2 \\ &\quad + (R_X^{-1}Bx - R_X^{-1}By, R_X^{-1}By - y)_X \\ &\ge \|R_X^{-1}Bx - R_X^{-1}By\|_X^2 = \|Bx - By\|_{X^*}^2\,, \end{aligned}$$

and, thus, of type (M) (cf. Lemma 3.4(i)).

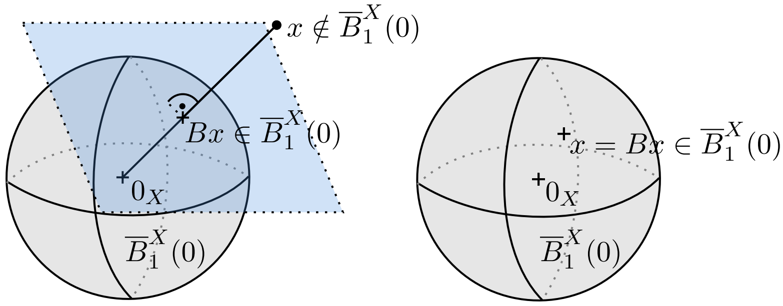

Figure 3.1: Sketch of orthogonal projection operator $B\colon X \to X^*$.

(iii) The operator $A + B\colon X \to X^*$ is not of type (M), since for the sequence $(x_n)_{n\in\mathbb{N}} := (e_1 + e_n)_{n\in\mathbb{N}} \subseteq X$, it holds that

$$\begin{aligned} x_n &\rightharpoonup e_1 && \text{in } X && (n\to\infty)\,,\\ (A+B)x_n &\rightharpoonup \Big(\frac{1}{\sqrt{2}} - 1\Big) R_X e_1 && \text{in } X^* && (n\to\infty)\,,\\ \limsup_{n\to\infty} \langle (A+B)x_n, x_n\rangle_X &< \Big\langle \Big(\frac{1}{\sqrt{2}} - 1\Big) R_X e_1, e_1\Big\rangle_X\,, \end{aligned}$$

where we used that $Bx_n = \frac{1}{\sqrt{2}} R_X x_n$ in X^* and, thus, $\langle (A+B)x_n, x_n\rangle_X = 2(\frac{1}{\sqrt{2}} - 1)$ for all $n \in \mathbb{N}$, but $(A+B)e_1 = -R_X e_1 + R_X e_1 = 0 \neq (\frac{1}{\sqrt{2}} - 1) R_X e_1$ in X^*.

3.2 Pseudo-monotonicity

A class of operators, which is stable under summation and (strictly) 'intermediate' between radially continuous, monotone operators and operaters of type (M), is given through the class of pseudo-monotone operators.

Definition 3.6 (pseudo-monotonicity)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, an operator $A\colon X \to X^*$ is pseudo-monotone *if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from*

$$\begin{aligned} &x_n \rightharpoonup x \quad \text{in } X \quad (n\to\infty)\,,\\ &\limsup_{n\to\infty} \langle Ax_n, x_n - x\rangle_X \leq 0\,, \end{aligned}$$

for every $y \in X$, it follows that

$$\langle Ax, x-y\rangle_X \leq \liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X\,.$$

Remark 3.7 *(i) If $\dim X < \infty$, then an operator $A\colon X \to X^*$ is continuous if and only if pseudo-monotone and locally bounded (Homework 6, Problem 1). Thus, in the finite-dimensional case, pseudo-monotonicity and the condition (M) coincide to continuity;*

(ii) Pseudo-monotonicity is actually a notion of continuity for the treatment of variational inequalities (cf. [Růž04, Satz 3.54(i)]): more precisely, if $C \subseteq X$ is a non-empty, closed, and convex subset of a reflexive Banach space X, then, if the operator $A\colon X \to X^*$ is bounded, pseudo-monotone, and coercive, for every $x^* \in X^*$, there exists $x \in C$ such that for every $y \in C$, it holds that

$$\langle Ax, x-y\rangle_X \geq \langle x^*, x-y\rangle_X\,.$$

(iii) Note that, if $A\colon X \to X^*$ is pseudo-monotone, then for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$ from

$$\begin{aligned} &x_n \rightharpoonup x \quad \text{in } X \quad (n\to\infty)\,,\\ &\limsup_{n\to\infty} \langle Ax_n, x_n - x\rangle_X \leq 0\,, \end{aligned}$$

it follows that

$$\begin{aligned} 0 = \langle Ax, x-x\rangle_X &\leq \liminf_{n\to\infty} \langle Ax_n, x_n - x\rangle_X\\ &\leq \liminf_{n\to\infty} \langle Ax_n, x_n - x\rangle_X \leq 0\,, \end{aligned}$$

i.e.,

$$\lim_{n\to\infty} \langle Ax_n, x_n - x\rangle_X = 0\,,$$

which, for every $y \in X$, implies that

$$\begin{aligned} \langle Ax, x-y\rangle_X &\leq \liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X\\ &= \liminf_{n\to\infty} \langle Ax_n, x-y\rangle_X + \cancel{\lim_{n\to\infty} \langle Ax_n, x_n - x\rangle_X}\\ &= \liminf_{n\to\infty} \langle Ax_n, x-y\rangle_X\,. \end{aligned}$$

As a result, if there exists a functional $x^* \in X^*$ such that

$$Ax_n \rightharpoonup x^* \quad \text{in } X^* \quad (n \to \infty)\,, \tag{3.8}$$

for every $y \in X$, we conclude that

$$\langle Ax, x - y\rangle_X \leq \langle x^*, x - y\rangle_X\,,$$

i.e., $Ax = x^*$ in X^*. This indicates that pseudo-monotonicity is a notion of continuity.

The following lemma collects important examples for and properties of pseudo-monotone operators.

Lemma 3.9 (important examples for and properties of pseudo-monotone operators) *Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, the following statements apply:*

- *(i) If $A\colon X \to X^*$ is monotone and radially continuous, then $A\colon X \to X^*$ is pseudo-monotone;*
- *(ii) If $A\colon X \to X^*$ is strongly continuous, then $A\colon X \to X^*$ is pseudo-monotone;*
- *(iii) If $A\colon X \to X^*$ is pseudo-monotone, then $A\colon X \to X^*$ is of type (M);*
- *(iv) If $A, B\colon X \to X^*$ are pseudo-monotone, then the sum $A + B\colon X \to X^*$ is pseudo-monotone;*
- *(v) If $A\colon X \to X^*$ is pseudo-monotone and locally bounded and if, in addition, X is reflexive, then $A\colon X \to X^*$ is demi-continuous.*

Proof. *ad (i).* Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a sequence and $x \in X$ an element such that

$$\begin{aligned} x_n &\rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,,\\ \limsup_{n\to\infty}\, &\langle Ax_n, x_n - x\rangle_X \leq 0\,. \end{aligned}$$

From Remark 3.7, it follows that

$$\lim_{n\to\infty} \langle Ax_n, x_n - x\rangle_X = 0\,.$$

Next, let $y \in X$ and $t \in (0, 1]$ be fixed, but arbitrary, and set $y_t \coloneqq x + t(y - x) \in X$. Then, by the monotonicity of $A\colon X \to X^*$, for every $n \in \mathbb{N}$, we have that

$$\langle Ay_t - Ax_n, y_t - x_n\rangle_X \geq 0\,,$$

which, by the definition of $y_t \in X$, for every $n \in \mathbb{N}$, is equivalent to

$$\begin{aligned} t\langle Ax_n, x_n - y\rangle_X \geq -(1-t)\langle Ax_n, x_n - x\rangle_X + (1-t)\langle Ay_t, x_n - x\rangle_X\\ + t\langle Ay_t, x_n - y\rangle_X\,. \end{aligned}$$

Therefore, if we divide both sides by $t > 0$ and take the limit inferior with respect to $n \to \infty$ on both sides, for every $y \in X$ and $t \in (0, 1]$, we arrive at

$$\liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X \geq \langle Ay_t, x - y\rangle_X\,.$$

Then, by the radial continuity of $A\colon X \to X^*$, if we pass for $t \searrow 0$, for every $y \in X$, we conclude that

$$\liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X \geq \langle Ax, x - y\rangle_X\,.$$

In other words, the operator $A\colon X \to X^*$ is pseudo-monotone.

ad (ii). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a sequence and $x \in X$ an element such that

$$\begin{aligned} x_n &\rightharpoonup x \quad \text{in } X \quad (n\to\infty)\,,\\ \limsup_{n\to\infty}\,&\langle Ax_n, x_n - x\rangle_X \le 0\,. \end{aligned}$$

Since $A\colon X \to X^*$ is strongly continuous, we infer that

$$Ax_n \to Ax \quad \text{in } X^* \quad (n\to\infty)\,.$$

which, by Proposition 1.13(ii)&(v), for every $y \in X$, implies that

$$\langle Ax, x-y\rangle_X = \lim_{n\to\infty} \langle Ax_n, x_n - y\rangle_X\,.$$

In other words, the operator $A\colon X \to X^*$ is pseudo-monotone.

ad (iii). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a sequence and $x \in X$ an element such that

$$\begin{aligned} x_n &\rightharpoonup x \qquad \text{in } X \qquad (n\to\infty)\,,\\ Ax_n &\rightharpoonup x^* \qquad \text{in } X^* \qquad (n\to\infty)\,,\\ \limsup_{n\to\infty}\,&\langle Ax_n, x_n\rangle_X \le \langle x^*, x\rangle_X\,. \end{aligned}$$

Then, we have that

$$\begin{aligned} \limsup_{n\to\infty} \langle Ax_n, x_n - x\rangle_X &= \limsup_{n\to\infty} \langle Ax_n, x_n\rangle_X + \lim_{n\to\infty} -\langle Ax_n, x\rangle_X\\ &\le \langle x^*, x\rangle_X - \langle x^*, x\rangle_X\\ &= 0\,. \end{aligned}$$

Therefore, following the argumentation of Remark 3.7, we conclude that $Ax = x^*$ in X^*. In other words, the operator $A\colon X \to X^*$ is of type (M).

More precisely, since $A\colon X \to X^*$ is pseudo-monotone, for every $y \in X$, it follows that

$$\begin{aligned} \langle Ax, x-y\rangle_X &\le \liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X\\ &\le \limsup_{n\to\infty} \langle Ax_n, x_n - y\rangle_X\\ &\le \limsup_{n\to\infty} \langle Ax_n, x_n\rangle_X + \lim_{n\to\infty} \langle Ax_n, -y\rangle_X\\ &\le \langle x^*, x\rangle_X - \langle x^*, y\rangle_X\\ &\le \langle x^*, x-y\rangle_X\,. \end{aligned}$$

In particular, replacing $y \in X$ by $x \pm y \in X$, where $y \in X$ is arbitrary, for every $y \in X$, we conclude that

$$\pm\langle Ax, y\rangle_X \le \pm\langle x^*, y\rangle_X\,,$$

i.e., $Ax = x^*$ in X^*. In other words, the operator $A\colon X \to X^*$ is of type (M).

ad (iv). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a sequence and $x \in X$ an element such that

$$\begin{aligned} x_n \rightharpoonup x \quad &\text{in } X \quad (n \to \infty)\,,\\ \limsup_{n\to\infty} \langle (A+B)x_n, x_n - x\rangle_X &\leq 0\,. \end{aligned}$$

For every $n \in \mathbb{N}$, we define

$$\begin{aligned} a_n &\coloneqq \langle Ax_n, x_n - x\rangle_X\,,\\ b_n &\coloneqq \langle Bx_n, x_n - x\rangle_X\,. \end{aligned}$$

Then, we have that

$$\begin{aligned} \limsup_{n\to\infty} a_n &= \limsup_{n\to\infty} \langle Ax_n, x_n - x\rangle_X \leq 0\,,\\ \limsup_{n\to\infty} b_n &= \limsup_{n\to\infty} \langle Bx_n, x_n - x\rangle_X \leq 0\,. \end{aligned}$$

In fact, suppose the contrary, *e.g.*, that $\limsup_{n\to\infty} a_n = a > 0$. Then, there exists a subsequence $(a_{n_k})_{k\in\mathbb{N}}$ and an element $a \in \mathbb{R}$ such that

$$a_{n_k} \to a \quad \text{in } \mathbb{R} \quad (k \to \infty)\,,$$

and, consequently,

$$\limsup_{k\to\infty} b_{n_k} = \limsup_{k\to\infty} \big\{a_{n_k} + b_{n_k}\big\} - \lim_{k\to\infty} a_{n_k} \leq -a < 0\,,$$

i.e., a contradiction, since, then, the pseudo-monotonicity of $B : X \to X^*$, $n \in \mathbb{N}$, choosing $y = x$, implies that $0 \leq \liminf_{k\to\infty} b_{n_k} < 0$. Therefore, the pseudo-monotonicity of $A, B\colon X \to X^*$, for every $y \in X$, implies that

$$\begin{aligned} \langle Ax, x - y\rangle_X &\leq \liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X\,,\\ \langle Bx, x - y\rangle_X &\leq \liminf_{n\to\infty} \langle Bx_n, x_n - y\rangle_X\,. \end{aligned}$$

For every $y \in X$, summing these inequalities yields that

$$\begin{aligned} \langle (A+B)x, x - y\rangle_X &\leq \liminf_{n\to\infty} \langle Ax_n, x_n - y\rangle_X + \liminf_{n\to\infty} \langle Bx_n, x_n - y\rangle_X\\ &\leq \liminf_{n\to\infty} \langle (A+B)x_n, x_n - y\rangle_X\,. \end{aligned}$$

In other words, the operator $A + B\colon X \to X^*$ is pseudo-monotone.

ad (v). Follows, due to (iii), from Lemma 3.4(v). □

According to Lemma 3.9, the class of pseudo-monotone operators is an 'intermediate' class between radially continuous and monotone operators and operators of type (M). The following examples clarify that class of pseudo-monotone operators is even a 'strict intermediate' class.

The first example shows that the class of pseudo-monotone operators is a strict subclass of the class of operators of type (M).

Example 3.10 (type (M) $\not\Rightarrow$ pseudo-monotone)
Let $(X,(\cdot,\cdot)_X)$ be a Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{N}}\subseteq X$. Moreover, let the operator $A\colon X\to X^*$ be defined as in Example 3.5. Then, $A\colon X\to X^*$ is of type (M) (cf. Lemma 3.4(ii)), but for the sequence $(x_n)_{n\in\mathbb{N}}\coloneqq(e_n)_{n\in\mathbb{N}}\subseteq X$, it holds that

$$\begin{aligned} x_n &\rightharpoonup 0 \quad \text{in } X \quad (n\to\infty)\,,\\ \lim_{n\to\infty}\langle Ax_n, x_n-0\rangle_X &= 0\,, \end{aligned}$$

where we used that $\langle Ax_n,x_n\rangle_X=-1$ and $\langle Ax_n,0\rangle_X=0$ for all $n\in\mathbb{N}$, but

$$\langle A0,0-0\rangle_X = 0 > -1 = \lim_{n\to\infty}\langle Ax_n, x_n-e_1\rangle_X\,.$$

In other words, $A\colon X\to X^*$ is not pseudo-monotone.

The second example shows that the class of radially continuous and monotone operators is a strict subclass of the class of pseudo-monotone operators.

Example 3.11 (pseudo-monotone $\not\Rightarrow$ radially continuous and monotone)
Let $\Omega\subseteq\mathbb{R}^d$, $d\geq 2$, be a bounded domain and $p\in[\frac{3d}{d+2},+\infty)$. Then, the weak convective term $C\colon V_p\to(V_p)^*$, for every $\mathbf{v},\boldsymbol{\varphi}\in V_p$ defined by

$$\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p} \coloneqq -\int_\Omega \mathbf{v}\otimes\mathbf{v}:\mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\,,$$

is well-defined, bounded, and pseudo-monotone (cf. Lemma 3.18), but not monotone, since for every $\mathbf{v},\boldsymbol{\varphi}\in V_p$, due to $\langle C\mathbf{v},\mathbf{v}\rangle_{V_p}=\langle C\boldsymbol{\varphi},\boldsymbol{\varphi}\rangle_{V_p}=0$ (cf. Lemma 3.18), it holds that

$$\langle C\mathbf{v}-C\boldsymbol{\varphi},\mathbf{v}-\boldsymbol{\varphi}\rangle_{V_p} = -\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p} - \langle C\boldsymbol{\varphi},\mathbf{v}\rangle_{V_p}\,.$$

If $C\colon V_p\to(V_p)^*$ would be monotone, then for every $\mathbf{v},\boldsymbol{\varphi}\in V_p$ and $\rho\in\mathbb{R}$, replacing above $\mathbf{v}$ by $\rho\mathbf{v}$, we would have that

$$0\leq -\rho^2\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p} - \rho\langle C\boldsymbol{\varphi},\mathbf{v}\rangle_{V_p}\,.$$

Then, for fixed $\mathbf{v},\boldsymbol{\varphi}\in V_p$ such that $\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p}>0$, there exists a sufficiently large $\rho\in\mathbb{R}$ such that

$$-\rho^2\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p} - \rho\langle C\boldsymbol{\varphi},\mathbf{v}\rangle_{V_p} < 0\,,$$

which is a contradiction. Therefore, $C\colon V_p\to(V_p)^*$ is not monotone. Note that $\mathbf{v},\boldsymbol{\varphi}\in V_p$ with $\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p}>0$ exist, since for $\mathbf{v}\in V_p\setminus\{0\}$, we have that $C\mathbf{v}\in(V_p)^*\setminus\{0\}$, so that there exists some $\boldsymbol{\varphi}\in V_p$ such that $\langle C\mathbf{v},\boldsymbol{\varphi}\rangle_{V_p}>0$.

3.3 Main theorem on pseudo-monotone operators

After having generalized the notions of continuity from the intermediate value theorem (*cf.* Theorem 3.1) to infinite-dimensional Banach spaces as well as non-monotone operators, we have everything at our disposal to, finally, establish a generalization of the intermediate value theorem (*cf.* Theorem 3.1) to infinite-dimensional Banach spaces.

Theorem 3.12 (Brézis, 1963)
Let $(X, \|\cdot\|_X)$ be a (real) separable, reflexive Banach space. Moreover, let $A\colon X \to X^*$ be bounded, pseudo-monotone, and coercive. Then, the operator $A\colon X \to X^*$ is surjective.

Remark 3.13 (a few comments) *(i) Theorem 3.12 is typically called the* Brézis theorem *or the* main theorem on pseudo-monotone operators*;*
(ii) Theorem 3.12 remains true if we drop the assumption that X is separable;
(iii) Theorem 3.12 remains true if we replace pseudo-monotonicity with the condition (M);
(iv) Theorem 3.12 remains true if we drop the assumption that $A\colon X \to X^*$ is bounded.

Proof (of Theorem 3.12(i)). Let $x^* \in X^*$ be fixed, but arbitrary.

Objective: We want to show that there exists $x \in X$ such that for every $y \in X$, it holds that

$$\langle Ax, y\rangle_X = \langle x^*, y\rangle_X \,,$$

i.e., $Ax = x^*$ in X^*.

To this end, we resort to the *Galerkin method*, *i.e.*, we approximate the Banach space X through a (strictly) increasing sequence of finite-dimensional Banach spaces $(X_n)_{n\in\mathbb{N}}$ (to which Theorem 2.15 can be applied). More precisely, due to the separability of the Banach space X, there exists a countable dense subset $\{b_n \mid n \in \mathbb{N}\}$. Without loss of generality, we may assume that the elements in $\{b_n \mid n \in \mathbb{N}\}$ are linear independent. Then, we define the sequence of finite-dimensional spaces $(X_n)_{n\in\mathbb{N}}$, for every $n \in \mathbb{N}$, by

$$X_n \coloneqq \operatorname{span}(\{b_1, \ldots, b_n\}) \,.$$

Then, the following statements apply:

- $\dim(X_n) = n$ and $X_n \subsetneq X_{n+1} \subseteq X$ for all $n \in \mathbb{N}$;
- $\bigcup_{n\in\mathbb{N}} X_n$ is dense in X.

Given (strictly) increasing sequence of finite-dimensional spaces $(X_n)_{n\in\mathbb{N}}$, we introduce the a sequence of finite-dimensional *Galerkin systems*: *i.e.*, for every $n \in \mathbb{N}$, we seek for a *Galerkin solution* $x_n \in X_n$ such that for every $y_n \in X_n$, it holds that

$$\langle A_n x_n, y_n\rangle_{X_n} = \langle x_n^*, y_n\rangle_{X_n} \,, \tag{GS}$$

i.e., $(\mathrm{id}_{X_n})^* A x_n = (\mathrm{id}_{X_n})^* x^*$ in X_n^*.

By means of the sequence of Galerkin systems (GS), we prove the existence of a solution $x \in X$ of $Ax = x^*$ in X^* in three main steps:

1. *Well-posedness of Galerkin systems* (GS)*:* We establish the existence of Galerkin solutions of the sequence Galerkin systems (GS);
2. *Stability of Galerkin systems* (GS)*:* We derive *a priori* bounds for the sequence of Galerkin solutions of the sequence Galerkin systems (GS);
3. *(Weak) convergence of Galerkin systems* (GS)*:* We establish the (weak) convergence of the sequence of Galerkin solutions of the sequence Galerkin systems (GS) to a solution $x \in X$ of $Ax = x^*$ in X^*.

1. Well-posedness of the Galerkin systems (GS)*:* We can proceed analogously to the respective step in the main theorem on monotone operators (*cf.* Theorem 2.13), since, in this step, it is only needed that $A\colon X \to X^*$ is coercive and demi-continuous, which follows from Lemma 3.9(v) (*cf.* Lemma 3.4(vi), for the case that pseudo-monotonicity is replaced by the condition (M)).

2. Stability of the Galerkin systems (GS)*:*

2.1. Boundedness of sequence of solutions: We can proceed analogously to the respective step in the main theorem on monotone operators (*cf.* Theorem 2.13), since, in this step, it is only needed that $A\colon X \to X^*$ is coercive.

2.2. Boundedness of images of sequence of solutions: Since $A\colon X \to X^*$ is bounded, the sequence of images of Galerkin solutions $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$ is bounded.

3. (Weak) convergence of Galerkin scheme: Due to Step 2, the sequence of Galerkin solutions $(x_n)_{n\in\mathbb{N}} \subseteq X$ of the sequence Galerkin schemes (GS) as well as the corresponding image sequence $(Ax_n)_{n\in\mathbb{N}} \subseteq X^*$ are bounded. As a consequence, by the reflexivity of X, the Eberlein–Šmulian theorem (*cf.* Theorem 1.14) yields subsequences $(x_{n_k})_{k\in\mathbb{N}} \subseteq X$, $(Ax_{n_k})_{k\in\mathbb{N}} \subseteq X^*$ as well as elements $x \in X$, $\xi^* \in X^*$ such that

$$\begin{aligned} x_{n_k} &\rightharpoonup x &\quad& \text{in } X &\quad& (k \to \infty)\,,\\ Ax_{n_k} &\rightharpoonup \xi^* && \text{in } X^* && (k \to \infty)\,. \end{aligned}$$

First, we show that $\xi^* = x^*$ in X^*. To this end, let $y \in X$ be fixed, but arbitrary. Then, there exists a sequence $y_n \in X_n$, $n \in \mathbb{N}$, such that

$$y_n \to y \quad \text{in } X \quad (n \to \infty)\,.$$

Hence, by Proposition 1.13(ii),(v), we deduce that

$$\begin{aligned} \langle \xi^*, y\rangle_X &= \lim_{k\to\infty} \langle Ax_{n_k}, y_{n_k}\rangle_X \\ &= \lim_{k\to\infty} \langle x^*, y_{n_k}\rangle_X \\ &= \langle x^*, y\rangle_X\,. \end{aligned}$$

Since $y \in X$ was chosen arbitrary, we infer that $\xi^* = x^*$ in X^*. As a result, we obtain

$$\begin{aligned} \limsup_{k\to\infty} \langle Ax_{n_k}, x_{n_k}\rangle_X &= \lim_{k\to\infty} \langle x^*, x_{n_k}\rangle_X \\ &= \langle x^*, x\rangle_X\,, \end{aligned}$$

since $A\colon X \to X^*$ is of type (M) (*cf.* Lemma 3.9(iii)), we conclude that $Ax = x^*$ in X^*. $\square$

3.4 Applications of the main theorem on pseudo-monotone operators

In this section, we address the most famous application of the main theorem on pseudo-monotone operators (*cf.* Theorem 3.12).

3.4.1 p-Navier–Stokes equations

As an application of the main theorem on pseudo-monotone operators (*cf.* Theorem 3.12), we obtain the solvability of the weak formulation of the p-Navier–Stokes equations, which model the laminar and turbulent flow of a *non-Newtonian fluid* through a bounded domain under the influence of an external force. More precisely, the p-Navier–Stokes equations, for a given bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, occupied by the fluid and a given external force $\mathbf{f}\colon \Omega \to \mathbb{R}^d$ acting on the fluid, seek for a velocity vector field $\mathbf{v} := (v_1, \dots, v_d)^\top \colon \Omega \to \mathbb{R}^d$ and a kinematic pressure $\pi\colon \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\operatorname{div}(\mathbf{S}(\mathbf{D}\mathbf{v})) + \operatorname{div}(\mathbf{v} \otimes \mathbf{v}) + \nabla\pi &= \mathbf{f} && \text{in } \Omega\,, \\ \operatorname{div}\mathbf{v} &= 0 && \text{in } \Omega\,, \\ \mathbf{v} &= \mathbf{0} && \text{on } \partial\Omega\,. \end{aligned} \tag{3.14}$$

In this subsection, we restrict to the case of a power-law fluid (*cf.* Introduction 0), *i.e.*, $\mathbf{S}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$, for every $\mathbf{A} \in \mathbb{R}^{d\times d}$, is defined by

$$\mathbf{S}(\mathbf{A}) := \nu_0\,(\delta + |\mathbf{A}|)^{p-2}\mathbf{A} \quad \text{in } \mathbb{R}^{d\times d}\,,$$

where $\nu_0 > 0$, $\delta \geq 0$, and $p \in (1, +\infty)$ the power-law index.

Definition 3.15 (weak formulation of the p-Navier–Stokes equations)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain, $p \in [\frac{3d}{d+2}, +\infty)$, and $\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^*$. Then, a couple $(\mathbf{v}, \pi)^\top \in (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$ is called weak solution of the p-Navier–Stokes equations *if for every $(\boldsymbol{\varphi}, \eta)^\top \in (W_0^{1,p}(\Omega))^d \times L^{p'}(\Omega)$, it holds that*

$$\begin{aligned} \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x - \int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x - \int_\Omega \pi \operatorname{div}\boldsymbol{\varphi}\,\mathrm{d}x &= \langle \mathbf{f}^*, \boldsymbol{\varphi}\rangle_{(W_0^{1,p}(\Omega))^d}\,, \\ \int_\Omega \eta \operatorname{div}\mathbf{v}\,\mathrm{d}x &= 0\,. \end{aligned}$$

Analogously to the weak formulation of the p-Stokes equations (*cf.* Definition 2.26), the weak formulation of the p-Navier–Stokes equations (*cf.* Definition 3.15) cannot directly be re-written as an operator equation with a coercive operator (*cf.* Theorem 3.12). Therefore, again, we first pass to the hydro-mechanical formulation of the p-Navier–Stokes equations (3.14), which can be re-written as an operator equation with a coercive operator to which the main theorem on pseudo-monotone operators (*cf.* Theorem 3.12) can be applied to.

Lemma 3.16 (hydro-mechanical formulation of the p-Navier–Stokes equations)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain, $p \in [\frac{3d}{d+2}, +\infty)$, and $\mathbf{f}^* \in ((W_0^{1,p}(\Omega))^d)^*$. Then, for a vector field $\mathbf{v} \in (W_0^{1,p}(\Omega))^d$, the following two statements are equivalent:

(i) There exists a function $\pi \in L_0^{p'}(\Omega)$ such that $(\mathbf{v}, \pi)^\top \in (W_0^{1,p}(\Omega))^d \times L_0^{p'}(\Omega)$ is a weak solution of the p-Navier–Stokes equations (cf. Definition 3.15);

(ii) The vector field $\mathbf{v} \in (W_0^{1,p}(\Omega))^d$ satisfies $\mathbf{v} \in V_p$ and is a weak solution of the hydro-mechanical p-Navier–Stokes equations, i.e., for every $\boldsymbol{\varphi} \in V_p$, it holds that

$$\int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi} \, \mathrm{d}x - \int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\boldsymbol{\varphi} \, \mathrm{d}x = \langle \mathbf{f}^*, \boldsymbol{\varphi} \rangle_{(W_0^{1,p}(\Omega))^d} \,.$$

Proof. The proof follows analogously to the proof of Lemma 2.28 up to minor adjustments. □

Motivated by Lemma 3.16, in order to establish the solvability of the weak formulation of the p-Navier–Stokes equations (*cf.* Definition 2.26), we first establish the solvability of the hydro-mechanical formulation of the p-Navier–Stokes equations (*cf.* Lemma 3.16(ii)), which will follow from the main theorem on pseudo-monotone operators (*cf.* Theorem 3.12).

Theorem 3.17 (solvability of the hydro-mech. form. of the p-Navier–Stokes equations)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in [\frac{3d}{d+2}, +\infty)$. Then, the operator $A\colon V_p \to (V_p)^*$, for every $\mathbf{v}, \boldsymbol{\varphi} \in V_p$ defined by

$$\langle A\mathbf{v}, \boldsymbol{\varphi} \rangle_{V_p} := \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi} \, \mathrm{d}x - \int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\boldsymbol{\varphi} \, \mathrm{d}x \,,$$

is well-defined, bounded, pseudo-monotone, coercive, and surjective.

Since we have already thoroughly examined the weak extra-stress tensor $\widehat{S}\colon V_p \to (V_p)^*$ for its properties (*cf.* Theorem 2.33), in order to prove the solvability of the hydro-mechanical formulation of the p-Navier–Stokes equations (*cf.* Theorem 3.17), it is sufficient to examine the weak convective term $C\colon V_p \to (V_p)^*$ (*cf.* Example 3.11) for its properties.

Lemma 3.18 (properties of the weak convective term)
Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in [\frac{3d}{d+2}, +\infty)$. Then, the weak convective term *$C\colon V_p \to (V_p)^*$, for every $\mathbf{v}, \boldsymbol{\varphi} \in V_p$ defined by*

$$\langle C\mathbf{v}, \boldsymbol{\varphi} \rangle_{V_p} := -\int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\boldsymbol{\varphi} \, \mathrm{d}x \,,$$

is well-defined, bounded, pseudo-monotone, and has the canceling property, i.e., for every $\mathbf{v} \in V_p$, we have that

$$\langle C\mathbf{v}, \mathbf{v} \rangle_{V_p} = 0 \,. \tag{3.19}$$

In addition, if $p > \frac{3d}{d+2}$, then $C\colon V_p \to (V_p)^*$ is strongly continuous.

Proof. *1. Well-definedness:* To begin with, we observe that $p^* \geq 2p'$ if and only if $p \geq \frac{3d}{d+2}$, where $p^* \in [1, +\infty]$ is the Sobolev embedding exponent, *i.e.*, $p^* := \frac{dp}{d-p}$ if $p \in [1, d)$ and $p^* \in [1, +\infty)$ arbitrarily large, but finite, if $p \geq d$. Therefore, for every $\mathbf{v} \in V_p$, by the Hölder inequality and Sobolev embedding theorem, we have that

$$\|\mathbf{v}\|_{L^{2p'}(\Omega)} \leq c\,\|\mathbf{v}\|_{L^{p^*}(\Omega)} \leq c\,\|\mathbf{v}\|_{V_p}\,.$$

As a consequence, for every $\mathbf{v}, \boldsymbol{\varphi} \in V_p$, using that $|\mathbf{v} \otimes \mathbf{v}| \leq |\mathbf{v}|^2$ and $|\mathbf{D}\boldsymbol{\varphi}| \leq |\nabla\boldsymbol{\varphi}|$ a.e. in Ω, the Hölder inequality, and the above inequality, we obtain

$$\begin{aligned}
\left| -\int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x \right| &\leq \int_\Omega |\mathbf{v}|^2 |\nabla\boldsymbol{\varphi}|\,\mathrm{d}x \\
&\leq \bigg(\int_\Omega |\mathbf{v}|^{2p'}\,\mathrm{d}x|\bigg)^{\frac{2}{2p'}} \bigg(\int_\Omega |\nabla\boldsymbol{\varphi}|^p\,\mathrm{d}x\bigg)^{\frac{1}{p}} \\
&= \|\mathbf{v}\|_{L^{2p'}(\Omega)}^2 \|\boldsymbol{\varphi}\|_{V_p} \\
&\leq c\,\|\mathbf{v}\|_{V_p}^2 \|\boldsymbol{\varphi}\|_{V_p}\,,
\end{aligned}$$

which implies that

$$\bigg(\boldsymbol{\varphi} \mapsto -\int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\bigg) \in (V_p)^*\,.$$

In other words, the operator $C\colon V_p \to (V_p)^*$ is well-defined.

2. Boundedness: Follows from the above inequality and the definition of the operator norm (*cf.* Proposition 1.2).

3. Canceling property: For every $\mathbf{v} = (v_1, \ldots, v_d)^\top \in V_p$, by integration-by-parts and $\operatorname{div}\mathbf{v} = 0$ a.e. in Ω, we find that

$$\begin{aligned}
\langle C\mathbf{v}, \mathbf{v}\rangle_{V_p} &= -\int_\Omega \mathbf{v} \otimes \mathbf{v} : \mathbf{D}\mathbf{v}\,\mathrm{d}x \\
&= -\int_\Omega \mathbf{v} \otimes \mathbf{v} : \nabla\mathbf{v}\,\mathrm{d}x \\
&= \sum_{i,j=1}^d -\int_\Omega v_i v_j \partial_j v_i\,\mathrm{d}x \\
&= \sum_{i,j=1}^d \int_\Omega \partial_j(v_i v_j) v_i\,\mathrm{d}x \\
&= \sum_{i,j=1}^d \int_\Omega \{(\partial_j v_i) v_i v_j + (\partial_j v_j)|v_i|^2\}\,\mathrm{d}x \\
&= \int_\Omega \{\mathbf{v} \otimes \mathbf{v} : \mathbf{D}\mathbf{v} + \cancel{\operatorname{div}\mathbf{v}}|\mathbf{v}|^2\}\,\mathrm{d}x \\
&= -\langle C\mathbf{v}, \mathbf{v}\rangle_{V_p}\,,
\end{aligned}$$

i.e., $\langle C\mathbf{v}, \mathbf{v}\rangle_{V_p} = 0$.

4. Pseudo-monotonicity: Let $(\mathbf{v}_n)_{n\in\mathbb{N}} \subseteq V_p$ be a sequence and $\mathbf{v} \in V_p$ be an element such that

$$\begin{aligned} \mathbf{v}_n &\rightharpoonup \mathbf{v} \quad \text{in } V_p \quad (n \to \infty)\,,\\ \limsup_{n\to\infty} \langle C\mathbf{v}_n, \mathbf{v}_n - \mathbf{v}\rangle_{V_p} &\leq 0\,. \end{aligned}$$

Then, since the embedding $V_p \hookrightarrow (L^{2p'}(\Omega))^d$ is linear and continuous and, thus, weakly continuous and since, by the Rellich theorem, $V_p \hookrightarrow\hookrightarrow (L^1(\Omega))^d$, we infer that

$$\begin{aligned} \mathbf{v}_n &\rightharpoonup \mathbf{v} &&\text{in } (L^{2p'}(\Omega))^d &&(n \to \infty)\,,\\ \mathbf{v}_n &\rightharpoonup \mathbf{v} &&\text{in } (L^1(\Omega))^d &&(n \to \infty)\,. \end{aligned}$$

which together implies that (*cf.* Exercise Sheet 7, Exercise 1)

$$\mathbf{v}_n \otimes \mathbf{v}_n \rightharpoonup \mathbf{v} \otimes \mathbf{v} \quad \text{in } (L^{p'}(\Omega))^{d\times d} \quad (n \to \infty)\,.$$

In other words, for every $\boldsymbol{\varphi} \in V_p$, we have that

$$\langle C\mathbf{v}_n, \boldsymbol{\varphi}\rangle_{V_p} \to \langle C\mathbf{v}, \boldsymbol{\varphi}\rangle_{V_p} \quad (n \to \infty)\,.$$

On the other hand, by the calling property (3.19), for every $n \in \mathbb{N}$, we have that

$$\langle C\mathbf{v}_n, \mathbf{v}_n\rangle_{V_p} = 0\,.$$

In summary, for every $\boldsymbol{\varphi} \in V_p$, we arrive at

$$\begin{aligned} \langle C\mathbf{v}, \mathbf{v} - \boldsymbol{\varphi}\rangle_{V_p} &= -\langle C\mathbf{v}, \boldsymbol{\varphi}\rangle_{V_p}\\ &= \lim_{n\to\infty} -\langle C\mathbf{v}_n, \boldsymbol{\varphi}\rangle_{V_p}\\ &= \lim_{n\to\infty} \langle C\mathbf{v}_n, \mathbf{v}_n - \boldsymbol{\varphi}\rangle_{V_p}\,. \end{aligned}$$

In other words, we have shown that $C\colon V_p \to (V_p)^*$ is pseudo-monotone.

5. Strong continuity: See Exercise Sheet 7, Exercise 2. □

Proof (of Theorem 3.17). Let es first summarize what we already know:

(a) According to Theorem 2.33, the weak extra-stress tensor $\widehat{S}\colon V_p \to (V_p)^*$, for every $\mathbf{v}, \boldsymbol{\varphi} \in V_p$ defined by

$$\langle \widehat{S}\mathbf{v}, \boldsymbol{\varphi}\rangle_{V_p} := \int_\Omega \mathbf{S}(\mathbf{D}\mathbf{v}) : \mathbf{D}\boldsymbol{\varphi}\,\mathrm{d}x\,,$$

is well-defined, bounded, continuous, strictly monotone, and coercive. By Lemma 3.9(i), the weak extra-stress tensor $\widehat{S}\colon V_p \to (V_p)^*$ is also pseudo-monotone.

(b) According to Lemma 3.18, the weak convective term $C\colon V_p \to (V_p)^*$ is well-defined, bounded, and pseudo-monotone.

Therefore, since $A \coloneqq \widehat{S} + C\colon V_p \to (V_p)^*$ and since boundedness and pseudo-monotonicity are stable with respect to summation (*cf.* Lemma 3.9(iv)), the operator $A\colon V_p \to (V_p)^*$ is bounded and pseudo-monotone as well. It is only left to verify coercivity. To this end, by the canceling property (3.19), for every $\mathbf{v} \in V_p$, we observe that

$$\langle A\mathbf{v}, \mathbf{v}\rangle_{V_p} = \langle \widehat{S}\mathbf{v}, \mathbf{v}\rangle_{V_p}\,,$$

so that, from the coercivity of the weak extra-stress tensor $\widehat{S}\colon V_p \to (V_p)^*$, we infer that

$$\frac{\langle A\mathbf{v}, \mathbf{v}\rangle_{V_p}}{\|\mathbf{v}\|_{V_p}} = \frac{\langle \widehat{S}\mathbf{v}, \mathbf{v}\rangle_{V_p}}{\|\mathbf{v}\|_{V_p}} \to +\infty \quad (\|\mathbf{v}\|_{V_p} \to +\infty)\,.$$

In order words, the operator $A\colon V_p \to (V_p)^*$ is coercive.

In summary, we have shown that the operator $A\colon V_p \to (V_p)^*$ is well-defined, bounded, pseudo-monotone, and coercive. Therefore, since the space V_p is separable and reflexive (*cf.* Remark 2.29), the main theorem on pseudo-monotone operators (*cf.* Theorem 2.13) yields that $A\colon V_p \to (V_p)^*$ is surjective. $\square$

Up to this point, we merely proved the solvability of the hydro-mechanical formulation of the p-Navier–Stokes equations (*cf.* Lemma 3.16(ii)). By means of Lemma 3.16(ii), we can finally prove the solvability of the weak formulation of the p-Navier–Stokes equations (*cf.* Definition 3.15).

Corollary 3.20 (solvability of the weak formulation of the p-Navier–Stokes equations) *Let $\Omega \subseteq \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and $p \in [\frac{3d}{d+2}, +\infty)$. Then, for every $\mathbf{f}^* \in ((W^{1,p}_0(\Omega))^d)^*$, there exists a solution $(\mathbf{v}, \pi)^\top \in (W^{1,p}_0(\Omega))^d \times L^{p'}_0(\Omega)$ of weak formulation of the p-Navier–Stokes equations (cf. Definition 3.15).*

Proof. Direct consequence of Theorem 3.17 together with Lemma 3.16. $\square$

4 Calculus of Variations

4.1 Direct Method in the Calculus of Variations

Definition 4.1 ((weak) lower semi-continuity)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, a functional $F\colon X \to \mathrm{R} \cup \{+\infty\}$ is called

(i) lower semi-continuous *(or* l.s.c.*), if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$, from*

$$x_n \to x \quad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$F(x) \leq \liminf_{n\to\infty} F(x_n)\,;$$

(ii) weakly lower semi-continuous *(or* w.l.s.c.*), if for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and an element $x \in X$, from*

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,,$$

it follows that

$$F(x) \leq \liminf_{n\to\infty} F(x_n)\,.$$

Remark 4.2 *(i) In Definition 4.1, the case $F(x) \leq \liminf_{n\to\infty} F(x_n) = +\infty$ is allowed;*
(ii) The condition $F(x) \leq \liminf_{n\to\infty} F(x_n)$ means that $F(x) \in \mathbb{R}\cup\{+\infty\}$ is lower bound for each accumulation point of $(F(x_n))_{n\in\mathbb{N}} \subseteq \mathbb{R} \cup \{+\infty\}$.

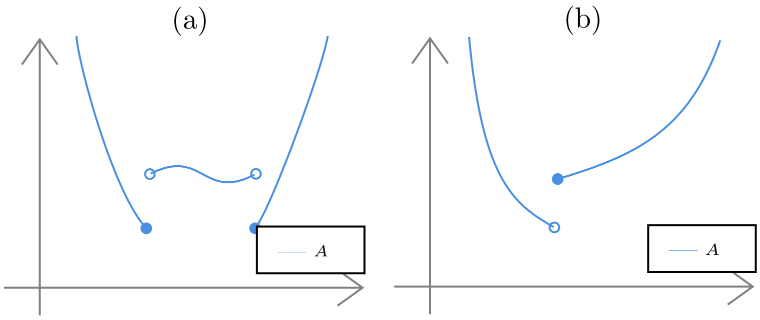

Figure 4.1: *(a)* a lower semi-continuous function $A\colon \mathbb{R} \to \mathbb{R}$;
(b) a not lower semi-continuous function $A\colon \mathbb{R} \to \mathbb{R}$.

Lemma 4.3 (relation between l.s.c. and w.l.s.c.)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. For a functional $F\colon X \to \mathbb{R} \cup \{+\infty\}$, the following statements apply:

(i) If $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c., then $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is l.s.c.;
(ii) If $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and convex, i.e., for every $x, y \in X$ and $\lambda \in [0,1]$, it holds that

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y)\,,$$

then $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c..

Proof. *ad (i).* Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ and $x \in X$ be such that

$$x_n \to x \quad \text{in } X \quad (n \to \infty)\,.$$

Then, due to Proposition 1.11(ii), we have that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,,$$

so that, since $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c., we find that

$$F(x) \leq \liminf_{n\to\infty} F(x_n)\,.$$

In other words, the functional $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is l.s.c..

ad (ii). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ and $x \in X$ be such that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,.$$

Then, we choose a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq X$ such that

$$F(x_{n_k}) \to \liminf_{n\to\infty} F(x_n) \quad (k \to \infty)\,.$$

Moreover, by Mazur's lemma, there exist $\lambda_k^n \in [0,1]$, $k = n, \dots, m_n$, $m_n \in \mathbb{N}$, with $\sum_{k=n}^{m_n} \lambda_k^n = 1$ and

$$y_n \coloneqq \sum_{k=n}^{m_n} \lambda_k^n x_{n_k} \to x \quad \text{in } X \quad (n \to \infty)\,.$$

Since $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and convex, we find that

$$\begin{aligned}
F(x) &\leq \liminf_{n\to\infty} F(y_n) \\
&\leq \liminf_{n\to\infty} \left\{ \sum_{k=n}^{m_n} \lambda_k^n F(x_{n_k}) \right\} \\
&\leq \liminf_{n\to\infty} \left\{ \max_{k=n,\dots,m_n} F(x_{n_k}) \right\} \left\{ \sum_{k=n}^{m_n} \lambda_k^n \right\} \\
&= \lim_{k\to\infty} F(x_{n_k}) \\
&= \liminf_{n\to\infty} F(x_n)\,.
\end{aligned}$$

In other words, the functional $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c.. $\square$

Example 4.4 (Indicator functional I)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, for $\mathcal{A} \subseteq X$, the indicator functional *$I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$, for every $x \in X$, is defined by*

$$I_{\mathcal{A}}(x) := \begin{cases} 0 & \text{if } x \in \mathcal{A}\,, \\ +\infty & \text{else}\,. \end{cases}$$

Then, the following equivalences apply:
(a) $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c. if and only if $\mathcal{A}$ is weakly closed;
(b) $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. if and only if $\mathcal{A}$ is (strongly) closed;
(c) $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if $\mathcal{A}$ is convex.

Proof (of (a)–(c)). *ad (a).*
ad '⇒'. Assume that $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c. and let $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ be such that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,.$$

Then, since $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c. and $I_{\mathcal{A}}(x_n) = 0$ for all $n \in \mathbb{N}$, we infer that

$$0 \le I_{\mathcal{A}}(x) \le \liminf_{n\to\infty} I_{\mathcal{A}}(x_n) = 0\,,$$

i.e., $I_{\mathcal{A}}(x) = 0$, which is equivalent to $x \in \mathcal{A}$. In other words, the set $\mathcal{A}$ is weakly closed.
ad '⇐'. Assume that $\mathcal{A}$ is weakly closed and let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be such that

$$x_n \rightharpoonup x \quad \text{in } X \quad (n \to \infty)\,.$$

Then, we distinguish two cases:
Case $x \in \mathcal{A}$. In this case, due to $I_{\mathcal{A}}(x_n) \ge 0$ for all $n \in \mathbb{N}$, we have that

$$I_{\mathcal{A}}(x) = 0 \le \liminf_{n\to\infty} I_{\mathcal{A}}(x_n)\,.$$

Case $x \notin \mathcal{A}$. In this case, since $\mathcal{A}^c$ is weakly open (as $\mathcal{A}$ is weakly closed), there exists some $n_0 \in \mathrm{N}$ such that for every $n \in \mathbb{N}$ with $n \ge n_0$, we have that $x_n \in \mathcal{A}^c$ and, thus, $I_{\mathcal{A}}(x_n) = +\infty$. Therefore, we have that

$$I_{\mathcal{A}}(x) \le +\infty = \liminf_{n\to\infty} I_{\mathcal{A}}(x_n)\,.$$

In other words, the indicator functional $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c..
ad (b). The claimed equivalence follows analogously to the equivalence (a) up to minor adjustments.
ad (c).
ad '⇒'. Assume that $I_{\mathcal{A}} \colon X \to \mathbb{R} \cup \{+\infty\}$ is convex and let $x, y \in \mathcal{A}$ and $\lambda \in [0, 1]$. Then, we have that

$$0 \le I_{\mathcal{A}}(\lambda x + (1-\lambda)y) \le \lambda I_{\mathcal{A}}(x) + (1-\lambda) I_{\mathcal{A}}(y) = 0\,,$$

i.e., $I_{\mathcal{A}}(\lambda x + (1-\lambda)y) = 0$, which is equivalent to $\lambda x + (1-\lambda)y \in \mathcal{A}$. In other words, the set $\mathcal{A}$ is convex.

ad '$\Leftarrow$'. Assume that $\mathcal{A}$ is convex and let $x, y \in X$ and $\lambda \in [0,1]$. Then, we distinguish two cases:

Case $x, y \in \mathcal{A}$. In this case, due to the convexity of $\mathcal{A}$, we have that $\lambda x + (1-\lambda)y \in \mathcal{A}$ and, thus,

$$I_{\mathcal{A}}(\lambda x + (1-\lambda)y) = 0 = \lambda I_{\mathcal{A}}(x) + (1-\lambda) I_{\mathcal{A}}(y)\,.$$

Case $x \notin \mathcal{A}$ or $y \notin \mathcal{A}$. In this case, we have that $I_{\mathcal{A}}(x) = +\infty$ or $I_{\mathcal{A}}(y) = +\infty$ and, thus,

$$I_{\mathcal{A}}(\lambda x + (1-\lambda)y) \leq +\infty = \lambda I_{\mathcal{A}}(x) + (1-\lambda) I_{\mathcal{A}}(y)\,.$$

In other words, the set $\mathcal{A}$ is convex. $\square$

Definition 4.5 (properness and effective domain)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, a functional $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is called proper *if the* effective domain, *i.e.,*

$$\mathrm{dom}(F) \coloneqq \{x \in X \mid F(x) < +\infty\}\,,$$

is non-empty.

Example 4.6 (indicator functional)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space and $\mathcal{A} \subseteq X$. Then, the indicator functional $I_{\mathcal{A}}\colon X \to \mathbb{R} \cup \{+\infty\}$ is proper if and only if $\mathcal{A} \neq \emptyset$.

Definition 4.7 (weak coercivity)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space. Then, a functional $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is called weakly coercive *if*

$$F(x) \to +\infty \quad (\|x\|_X \to +\infty)\,.$$

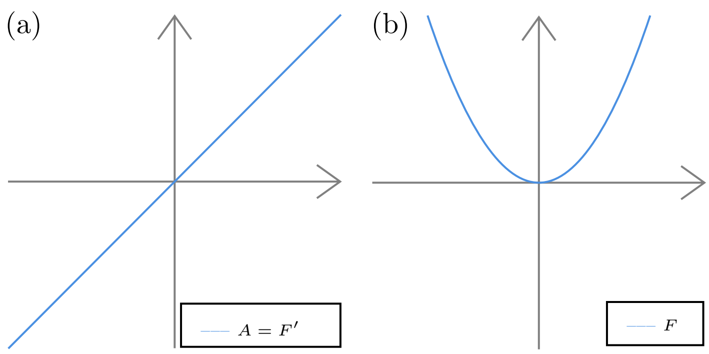

Figure 4.2: *(a)* a coercive function $A = F'\colon \mathbb{R} \to \mathbb{R}$;
(b) a weakly coercive function $F\colon \mathbb{R} \to \mathbb{R}$.

Example 4.8 (indicator functional)
Let $(X, \|\cdot\|_X)$ be a (real) Banach space and $\mathcal{A} \subseteq X$. Then, the indicator functional $I_{\mathcal{A}}\colon X \to \mathbb{R} \cup \{+\infty\}$ is weakly coercive if and only if $\mathcal{A}$ is bounded.

Proof. ad '$\Leftarrow$'. Assume that $\mathcal{A}$ is bounded and let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be such that

$$\|x_n\|_X \to +\infty \quad (n \to \infty)\,.$$

Then, there exists some $n_0 \in \mathbb{N}$ such that for every $n \in \mathbf{N}$ with $n \geq n_0$, we have that $x_n \in \mathcal{A}$ and, thus, $I_{\mathcal{A}}(x_n) = +\infty$, which implies that

$$I_{\mathcal{A}}(x_n) \to +\infty \quad (n \to \infty)\,.$$

In other words, the indicator functional $I_{\mathcal{A}}\colon X \to \mathbb{R} \cup \{+\infty\}$ is weakly coercive.

ad '$\Rightarrow$'. Assume that $I_{\mathcal{A}}\colon X \to \mathbb{R} \cup \{+\infty\}$ is weakly coercive. Moreover, suppose that $\mathcal{A}$ is unbounded. Then, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ be such that

$$\|x_n\|_X \to +\infty \quad (n \to \infty)\,.$$

Since $I_{\mathcal{A}}\colon X \to \mathbb{R} \cup \{+\infty\}$ is weakly coercive, we infer that

$$I_{\mathcal{A}}(x_n) \to +\infty \quad (n \to \infty)\,,$$

which is a contradiction since $I_{\mathcal{A}}(x_n) = +\infty$ for all $n \in \mathbb{N}$. □

Theorem 4.9 (direct method in the calculus of variations)
Let $(X, \|\cdot\|_X)$ be a (real) reflexive Banach space. Moreover, let $F\colon X \to \mathbb{R} \cup \{+\infty\}$ be a proper, w.l.s.c., and weakly lower semi-continuous functional. Then, there exists $x \in \operatorname{dom}(F)$ such that

$$F(x) = \inf_{y\in X} F(y)\,.$$

Proof. Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a minimizing sequence, *i.e.*, it holds that

$$F(x_n) \to \inf_{y\in X} F(y) \quad (n \to \infty)\,.$$

Note that, up to this point, we only know that $\inf_{y\in X} F(y) \in [-\infty, +\infty)$ (which we will improve later). Since $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is weakly coercive, we deduce that the sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is bounded. Therefore, since X is reflexive, by the Eberlein–Šmulian theorem, we find a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq X$ and a weak limit $x \in X$ such that

$$x_{n_k} \rightharpoonup x \quad \text{in } X \quad (k \to \infty)\,.$$

Then, since $F\colon X \to \mathbb{R} \cup \{+\infty\}$ is w.l.s.c., we find that

$$\begin{aligned} F(x) &\leq \liminf_{k\to\infty} F(x_{n_k}) \\ &= \lim_{k\to\infty} F(x_{n_k}) \\ &\inf_{y\in X} F(y)\,, \end{aligned}$$

which implies that $\inf_{y\in X} F(y) \in \mathbb{R}$ and the assertion. □

Bibliography

[Ada75] R.A. Adams, *Sobolev spaces*, Pure and Applied Mathematics, vol. 65, Academic Press, New York-London, 1975.

[BF13] Franck Boyer and Pierre Fabrie, *Mathematical tools for the study of the incompressible navier-stokes equations and related models*, Applied mathematical sciences, vol. 183, Springer, New York; Heidelberg; London [u.a.], 2013 (eng).

[Bré67] Haïm Brézis, *Inéquations d'évolution abstraites*, C. R. Acad. Sci., Paris, Sér. A **264** (1967), 732–735 (French).

[Bré11] Haïm Brézis, *Functional analysis, sobolev spaces and partial differential equations*, Universitext, Mathematics, Springer, New York; Heidelberg [u.a.], 2011 (eng).

[Els09] Jürgen Elstrodt, *Maß- und Integrationstheorie*, 6. ed., Springer-Lehrbuch, Grundwissen Mathematik, Springer, Berlin; Heidelberg [u.a.], 2009 (ger).

[Emm04] Etienne Emmrich, *Gewöhnliche und Operator-Differentialgleichungen. Eine integrierte Einführung in Randwertprobleme und Evolutionsgleichungen für Studierende*, Wiesbaden: Vieweg, 2004 (German).

[Jam51] Robert C. James, *A non-reflexive banach space isometric with its second conjugate space*, Proceedings of the National Academy of Sciences of the United States of America **37** (1951), no. 3, 174–177.

[Růž04] Michael Růžička, *Nichtlineare Funktionalanalysis: eine Einführung*, Springer, Berlin; Heidelberg [u.a.], 2004 (ger).

[Sax02] Karen Saxe, *Beginning functional analysis*, Undergraduate Texts Math., New York, NY: Springer, 2002 (English).

[Wer97] Dirk Werner, *Funktionalanalysis*, 2., überarb. Aufl. ed., Springer-Lehrb., Berlin: Springer, 1997 (German).