Calculus of Variations

An Introduction with Applications in Spectral Theory

Erwin Brüning
School of Mathematical Sciences

January 2011

Contents

Chapter 1

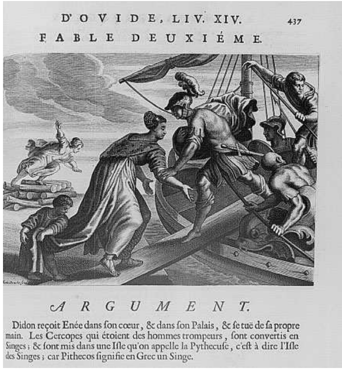
Introduction

Calculus of variations or 'variational calculus' is one of the oldest parts of mathematics and it has influenced the development of mathematics in many ways. The name for this theory goes back to L. Euler's publication *A method for finding curves enjoying certain maximum or minimum properties* in 1744. But the theory is much older, due to its intimate relation to problems humans encountered from early days of existence, namely to find 'optimal' solutions to practical problems, for instance to "minimize losses" and to "maximize benefits (profits)". In this wider sense of finding optimal solutions it is nearly 3000 years old (see next subsection).

1.1 History

Here we mention only the most important dates in the history of variational calculus. For a concise account we recommend the introduction to the book BB92 where one finds also further references. The following table presents some highlights:

≈ 800 B.C. **Legend of Queen Dido**; first recorded solution of an *isoperimetric problem* (i.e., the problem of finding the curve which encloses the largest area for a given length).



≈ 250 B.C. **ancient Greece**; several extreme value problems were solved, in particular it was shown that the shortest path joining two points is the straight line through the given points.

1662 **Fermat**: Principle of least time for the propagation of light.

July 1696 **birth of modern variational calculus**: *Johann Bernoulli* presented the *brachystochrone problem* in the Acta Eruditorum Lipsiae. The problem: A point mass glides without friction along a curve joining a point A with a lower point B. Which curve gives the shortest travel time if the point mass is moving only under the influence of gravity (accord-

July 1696 ing to Newton's law). Very soon many solutions were propoesed by well known mathematicians, among them by Jakob Bernoulli, Newton, Leibniz, and l'Hopital.

1684 **Leibniz** - first publication on differential calculus: *Nova methodus pro maximis et minimis itemque tangentibus*; **beginning of a mathematical theory of optimization**.

1744 **L. Euler**: first textbook entitled *A method for finding curves enjoying certain maximum or minimum properties*.

1753 **L. Euler**: *Dissertatio de principio minimae actionis*.

1786 **Legendre**: How to distinguish stationary points? Second necessary condition for an extremum, in terms of 'second variation.

1788 **Lagrange**: *Mécanique analytique*.

1837 **Jacobi**: *Theory of variational calculus and differential equations*; completed Legendre's study on the second variation.

$\approx$ 1850 **Dirichlet problem**: Dirichlet and Riemann took the existence of extreme value problems for granted, thus 'solving' the *Dirichlet problem*: Find a solution to the potential equation

$$\triangle u = 0 \quad \text{in G},$$

$G \subset \mathbb{R}^2$, with given values f on the smooth boundary ∂G of G, i.e., $u \restriction \partial G = f$, by arguing that the potential equation is just the

Euler-Lagrange equation of the *Dirichlet integral*

≈ 1850

$$D(u) = \int_G \left[(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2 \right] dxdy$$

which has a minimum and this minimum u solves the Euler-Lagrange equation.

1870 **K. Weierstraß**: Dirichlet, Gauß, Riemann ... are **wrong** by producing examples showing that in general

$$\inf \neq \min .$$

1901 **D. Hilbert**: Talk *On the Dirichlet Principle* at a DMV conference; proved correctness of this principle by establishing *weak lower semicontinuity* arguments.

1935 **Carathéodory**, the 'Isopérimaître incomparable', further extension of variational methods for partial differential equations in his book *Variational Calculus and First-order Partial Differential Equations*.

≈ 1965 **F. Browder**: systematic theory by solving (systems of) nonlinear partial differential equations by variational methods; monotonicity arguments, monotone operators.

1.2 Classical approach and direct methods

While the various solutions of the brachystochrone problem were based on ad hoc methods a systematic approach to solving variational problems started with Euler and Lagrange. They considered typically functionals of the form

$$f(u) = \int_a^b F(t, u(t), u'(t))dt \tag{1.1}$$

and tried to find a real-valued function u on the interval $[a, b]$ which minimizes this integral. (Recall that at Euler and Lagrange times no integration theory had been developed and thus they worked with an approximation which we would call an approximating Riemann sum). They found that the function u has to satisfy the well-known “Euler condition"

$$\frac{\partial F}{\partial u} - \frac{d}{dt}\frac{\partial F}{\partial u'} = 0.$$

In general, this is a nonlinear second order ordinary differential equation for u and certainly some regularity has to be assumed about the integrand F.

In the classical approach to this variational problem one tries to solve this Euler equation and uses additional arguments to show that solutions of this equation indeed give a minimum of the functional $f(u)$ by looking at the “second variation" (see later) as started by Legendre. In contrast, the direct methods of the calculus of variations use arguments of (infinite dimensional) nonlinear analysis, functional analysis and topology to find minima (maxima) of the functional directly and show that these extremal points are solutions of the corresponding Euler equation. In this way the direct methods of the calculus of variations have been established as a very powerful method to solve (systems of) nonlinear partial differential equations. In this case, for second order partial differential equations, one studies functionals of the form

$$f(u) = \int_\Omega F(x, u(x), Du(x))dx \tag{1.2}$$

where $\Omega \subset \mathbb{R}^n$ is an open nonempty subset and u is a real-valued function on Ω.

Chapter 2

Direct methods in the calculus of variations

The basic problem one wants to solve is the following.

> Given a nonempty set M and a real-valued function $f : M \longrightarrow \mathbb{R}$ one wants to know under which conditions on the pair (M, f) there is at least one $u \in M$ such that
>
> $$f(u) = \inf_{v \in M} f(v) \tag{2.1}$$
>
> holds.

If such a point $u \in M$ exists then the function f attains its minimal value at the point u and we say that u is a **minimizer for** f **on** M; and the minimization problem has been solved.

Clearly, when we know to solve minimization problems we also know to solve maximization problems for (M, f), i.e., problems in which we are looking for points $u \in M$ at which f attains its maximal possible value. Just note that the maximization problem for (M, f) is the minimization problem for $(M, -f)$.

In order to get some idea what is involved in solving a minimization problem it helps to recall an old well-known result of Weierstraß:

> A continuous function f attains its minimum and its maximum on a closed and bounded interval $[a, b]$.

We recommend that you have a careful look at the proof of this fundamental result. Then you will see that the above result 'contains' the result

> A lower semi-continuous function f attains its minimum on a closed and bounded interval $[a, b]$.

And a straight forward inspection of the proof of this result shows that one should expect the following statement to hold:

> A lower semi-continuous real-valued function f attains its minimum on a (sequentially) compact set M.

We recall the proof, assuming that the sequential characterization of lower semi-continuity is known.

In a first step one shows that f is bounded from below on M so that f has a finite infimum. Assume that f is not bounded from below. Then there is a sequence of points $x_n \in M$ such that $f(x_n) \leq -n$. Since M is sequentially compact there is a subsequence $(x_{n(j)})_{j\in\mathbb{N}}$ which converges to a point $y \in M$. Lower semi-continuity of f implies

$$f(y) \leq \liminf_{j\to\infty} f(x_{n(j)}).$$

By construction, $f(x_{n(j)}) \leq -n(j) \longrightarrow -\infty$ as $j \longrightarrow \infty$, a contradiction. Hence f has a finite infimum on M:

$$I = I(f, M) = \inf_{v\in M} f(v) > -\infty.$$

In the next step one constructs a 'minimizing sequence' which converges to a point in M. By definition of the infimum there is a sequence of points $x_n \in M$ such that

$$\lim_{n\longrightarrow\infty} f(x_n) = I.$$

Such a sequence is called a **minimization sequence**. (Note that in general a minimizing sequence will not converge in M). As above, by sequential compactness of M we can find a subsequence $(x_{n(j)})_{j\in\mathbb{N}}$ which converges to some point $y \in M$.

In the final step step one shows that the limit y of the convergent minimizing sequence $(x_{n(j)})_{j\in\mathbb{N}}$ is the minimizer we are looking for. Again lower semi-continuity of f implies

$$f(y) \leq \liminf_{j\to\infty} f(x_{n(j)}).$$

Observe $I \leq f(y)$, since $y \in M$. For a subsequence of a minimizing sequence one knows

$$\liminf_{j\to\infty} f(x_{n(j)}) = \lim_{j\to\infty} f(x_{n(j)}) = I,$$

and hence $I \leq f(y) \leq I$, i.e., $f(y) = I$, and y is a minimizer.

Remark 2.0.1 *In this proof compactness of the set M played a decisive role. It implied: (a) there is a finite infimum; (b) there are minimizing sequences which are bounded; (c) there are convergent minimizing sequences; (d) the limit of a convergent minimizing sequences belongs to the set M.*

Nevertheless, as we will learn later, one can do without compactness of the set if some additional restrictions are imposed on the function.

2.1 Outline of general strategy

As we had just seen compactness plays a decisive role in the proof of the minimization results given above. However this result is not useful in **infinite dimensional** minimization problems since in an infinite dimensional Banach space for instance there are nearly no compact sets as they occur in minimization problems (recall that in such a space compact sets have an empty interior). Therefore we need to develop a theory which does not rely on compactness of the set M (for the norm topology).

Here are the basic steps of a minimization theory.

1. Suppose M is a subset of the domain of the functional f and we want to find a minimum of f on M.

2. Through assumptions on f and/or M, assure that f has a finite infimum on M, i.e.,

$$\inf_{u \in M} f(u) = I(f, M) = I > -\infty. \tag{2.2}$$

Then there is a *minimizing sequence* $(u_n)_{n\in\mathbb{N}} \subset M$, i.e., a sequence in M such that

$$\lim_{n\to\infty} f(u_n) = I. \tag{2.3}$$

3. Suppose that we can find **one** minimizing sequence $(u_n)_{n\in\mathbb{N}} \subset M$ such that

$$u = \lim_{n\to\infty} u_n \in M, \tag{2.4}$$

$$f(u) \leq \liminf_{n\to\infty} f(u_n); \tag{2.5}$$

then the minimization problem is solved since then we have

$$I \leq f(u) \leq \liminf_{n\to\infty} f(u_n) = I$$

where the first inequality holds because of $u \in M$ and where the second identity holds because $(u_n)_{n\in\mathbb{N}}$ is a minimizing sequence. Obviously, for equation (2.4) a topology has to be specified on M.

4. Certainly, it is practically impossible to find one minimizing sequence with the two properties given above. Thus in explicit implementations of this strategy one works under conditions where the two properties hold for all convergent sequences, with respect to a suitable topology. If one looks at the proof of Weierstrass' theorem one expects to get a convergent minimizing sequence by taking a suitable subsequence of a given minimizing sequence. Recall: The coarser

the topology is, the easier it is for a sequence to have a convergent subsequence and to have a limit point, i.e., to have equation (2.4). On the other hand, the stronger the topology is the easier it is to satisfy inequality (2.5) which is a condition of lower semi-continuity.

5. The paradigmatic solution of this problem in infinite dimensional spaces is due to Hilbert who suggested using the weak topology, the main reason being that in a Hilbert space bounded sets are relatively sequentially compact for the weak topology while for the norm topology there are not too many compact sets of interest. Thus suppose that M is a weakly closed subset of a reflexive Banach space and that minimizing sequences are bounded (with respect to the norm). Then there is a weakly convergent subsequence whose weak limit belongs to M. Thus in order to conclude one verifies that inequality (2.5) holds for all weakly convergent sequences, i.e., that f is lower semi-continuous for the weak topology.

In the following sections the concepts and results which have been used above will be explained and some concrete existence results for extremal points will be formulated where the above strategy is implemented.

Suppose that with the direct methods of the calculus of variations we managed to show the existence of a local minimum of the functional f and that this functional is differentiable (in the sense of the classical methods). Then, if the local minimum occurs at an interior point u_0 of the domain of f, the Euler–Lagrange equation $f'(u_0) = 0$ holds and thus we have found a solution of this equation. If the functional f has the form (1.1) (or 1.2), then the equation $f'(u_0) = 0$ is a nonlinear ordinary (partial) differential equation and thus the direct methods become a powerful tool for solving nonlinear ordinary and partial differential equations. Some modern implementations of this strategy with many new results on nonlinear (partial) differential equations is described in good detail in the following books [Dac82, Dac89, BB92, JLJ98, Str00], in a variety of directions.

Note that a functional f can have other critical points than local extrema. These other critical points are typically not obtained by the direct methods as described above. However there are other, often topological methods of global analysis by which the existence of these other critical points can be established. We mention the minimax methods, index theory and mountain pass lemmas. These methods are developed and applied in [Zei85, BB92, Str00].

2.2 General existence results

From the Introduction we know that semi-continuity plays a fundamental role in direct methods in the calculus of variations. Accordingly we recall the definition and the basic characterization of lower semi-continuity. Upper semi-continuity of a function f is just lower semi-continuity of $-f$.

Definition 2.2.1 *Let M be a Hausdorff space. A function $f : M \to \mathbb{R} \cup \{+\infty\}$ is called* **lower semi-continuous at a point** *$x_0 \in M$ if, and only if, x_0 is an interior point of the set $\{x \in M : f(x) > f(x_0) - \epsilon\}$ for every $\epsilon > 0$. f is called* **lower semi-continuous on** *M if, and only if, f is lower semi-continuous at every point $x_0 \in M$.*

Lemma 2.2.2 *Let M be a Hausdorff space and $f : M \to \mathbb{R} \cup \{+\infty\}$ a function on M.*

a) If f is lower semi-continuous at $x_0 \in M$, then for every sequence $(x_n)_{n\in\mathbb{N}} \subset M$ converging to x_0, one has

$$f(x_0) \leq \liminf_{n\to\infty} f(x_n). \tag{2.6}$$

b) If M satisfies the first axiom of countability, i.e., if every point of M has a countable neighborhood basis, then the converse of a) holds.

Proof. For the simple proof we refer to the Exercises. □

In the Introduction we also learned that *compactness* plays a fundamental role too, more precisely, the direct methods use *sequential compactness* in a decisive way.

Definition 2.2.3 *Let M be a Hausdorff space. A subset $K \subset M$ is called* **sequentially compact** *if, and only if, every infinite sequence in K has a subsequence which converges in K.*

The following fundamental results proves the existence of a minimum. Replacing f by $-f$ it can easily be translated into a result on the existence of a maximum.

Theorem 2.2.4 (Existence of a minimizer) *Let $f : M \to \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous function on the Hausdorff space M. Suppose that there is a real number r such that*

a) $[f \leq r] = \{x \in M : f(x) \leq r\} \neq \emptyset$ and

b) $[f \leq r]$ is sequentially compact.

Then there is a minimizing point x_0 for f on M:

$$f(x_0) = \inf_{x \in M} f(x). \tag{2.7}$$

Proof. We begin by showing indirectly that f is lower bounded. If f is not bounded from below there is a sequence $(x_n)_{n \in \mathbb{N}}$ such that $f(x_n) < -n$ for all $n \in \mathbb{N}$. For sufficiently large n the elements of the sequence belong to the set $[f \leq r]$, hence there is a subsequence $y_j = x_{n(j)}$ which converges to a point $y \in M$. Since f is lower semi-continuous we know $f(y) \leq \liminf_{j \to \infty} f(y_j)$, a contradiction since $f(y_j) < -n(j) \to -\infty$. We conclude that f is bounded from below and thus has a finite infimum:

$$-\infty < I = I(f, M) = \inf_{x \in M} f(x) \leq r.$$

Therefore there is a minimizing sequence $(x_n)_{n \in \mathbb{N}}$ whose elements belong to $[f \leq r]$ for all sufficiently large n. Since $[f \leq r]$ is sequentially compact there is again a subsequence $y_j = x_{n(j)}$ which converges to a unique point $x_0 \in [f \leq r]$. Since f is lower semi-continuous we conclude

$$I \leq f(x_0) \leq \liminf_{j \to \infty} f(y_j) = \lim_{j \to \infty} f(y_j) = I.$$

□

Sometimes one can prove

Theorem 2.2.5 (Uniqueness of minimizer) *Suppose M is a convex set in a vector space E and $f : M \to \mathbb{R}$ is a strictly convex function on M. Then f has at most one minimizing point in M.*

Proof. Suppose there are two different minimizing points x_0 and y_0 in M. Since M is convex all points $x(t) = tx_0 + (1-t)y_0$, $0 < t < 1$, belong to M and therefore $f(x_0) = f(y_0) \leq f(x(t))$. Since f is strictly convex we know $f(x(t)) < tf(x_0) + (1-t)f(y_0) = f(x_0)$ and therefore the contradiction $f(x_0) < f(x_0)$. Thus there is at most one minimizing point. □

2.3 Minimization in Banach spaces

In interesting minimization problems we typically have at our disposal much more information about the set M and the function f than we have assumed in Theorem 2.2.4. If for instance one is interested in minimizing the functional (1.2) one would prefer to work in a suitable Banach space of functions, usually a Sobolev space. These function spaces and their properties are an essential input for applying them in the direct methods. A concise introduction to the most important of these function spaces can be found in [LL01].

Concerning the choice of a topology on Banach spaces which is suitable for the direct methods (compare our discussion in the Introduction) we begin by recalling the wellknown result of Riesz: The closed unit ball of a normed space is compact (for the norm topology) if, and only if, this space is finite dimensional. Thus, in infinite dimensional Banach spaces compact sets have an empty interior and therefore are not of much interest for must purposes of analysis, in particular not for the direct methods. Which other topology can be used? Recall that Weierstrass' result on the existence of extrema of continuous functions on closed and bounded sets uses in an essential way that in finite dimensional Euclidean spaces a set is compact if, and only if, it is closed and bounded. A topology with such a characterization of closed and bounded sets is known for infinite dimensional

Banach spaces too, the weak topology. Suppose E is a Banach space and E' is its topological dual space. Then the weak topology $\sigma = \sigma(E, E')$ on E is defined by the system $\{q_u(\cdot) : u \in E'\}$ of semi-norms q_u, $q_u(x) = |u(x)|$ for all $x \in E$. In most applications one can actually use reflexive Banach spaces and there the following important result is available.

Lemma 2.3.1 *In a reflexive Banach space E every bounded set (for the norm) is relatively compact for the weak topology $\sigma(E, E')$.*

A fairly detailed discussion about compact and weakly compact sets in Banach spaces, as they are relevant for the direct methods, is given in the Appendix of [BB92]. Prominent examples of reflexive Banach spaces are Hilbert spaces (see Chapter 18), the Lebesgue spaces L^p for $1 < p < \infty$, and the corresponding Sobolev spaces $W^{m,p}$, $m = 1, 2, \ldots$, $1 < p < \infty$.

Accordingly we decide to use mainly reflexive Banach spaces for the direct methods, whenever this is possible. Then, with the help of Lemma 2.3.1, we always get weakly convergent minimizing sequences whenever we can show that bounded minimizing sequences exist. Thus the problem of lower semi-continuity of the functional f for the weak topology remains. This is unfortunately not a simple problem. Suppose we consider a functional of the form (1.2) and, according to the growth restrictions on the integrand F, we decide to work in a Sobolev space $E = W^{1,p}(\Omega)$ or in a closed subspace of this space, $\Omega \subseteq \mathbb{R}^d$ open. Typically, the restrictions on F, which assure that f is well defined on E, imply that f is continuous (for the norm topology). However the question when such a functional is lower semi-continuous for the weak topology is quite involved, nevertheless a fairly comprehensive answer is known (see [Dac82]). Under certain technical assumptions on the integrand F the functional f is lower semi-continuous for the weak topology on $E = W^{1,p}(\Omega)$ if, and only if, for (almost) all $(x, u) \in \Omega \times \mathbb{R}^m$ the function $y \mapsto F(x, u, y)$ is convex (if $m = 1$), respectively quasi-convex (if $m > 1$).

Though in general continuity of a functional for the norm topology does not imply its continuity for the weak topology, there is a large and much used class of functionals where this implication holds. This is the class of convex functionals and for this reason *convex minimization* is relatively easy. We prepare the proof of this important result with a lemma.

Lemma 2.3.2 *Let E be a Banach space and M a weakly (sequentially) closed subset. A function $f : M \to \mathbb{R}$ is (sequentially) lower semi-continuous on M for the weak topology if, and only if, the sub-level sets $[f \leq r]$ are weakly (sequentially) closed for every $r \in \mathbb{R}$.*

Proof. We give the proof explicitly for the case of sequential convergence. For the general case one proceeds in the same way using nets.

Let f be weakly sequentially lower semi-continuous and for some $r \in \mathbb{R}$ let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $[f \leq r]$ which converges weakly to some point $x \in M$ (since M is weakly sequentially closed). By Lemma 2.2.2 we know $f(x) \leq \liminf_{n\to\infty} f(x_n)$ and therefore $f(x) \leq r$, i.e., $x \in [f \leq r]$. Therefore $[f \leq r]$ is closed.

Conversely assume that all the sub-level sets $[f \leq r]$, $r \in \mathbb{R}$, are weakly sequentially closed. Suppose f is not weakly sequentially lower semi-continuous on M. Then there is a weakly convergent sequence $(x_n)_{n\in\mathbb{N}} \subset M$ with limit $x \in M$ such that $\liminf_{n\to\infty} f(x_n) < f(x)$. Choose a real number r such that $\liminf_{n\to\infty} f(x_n) < r < f(x)$. Then there is a subsequence $y_j = x_{n(j)} \subset [f \leq r]$. This subsequence too converges weakly to x and, since $[f \leq r]$ is weakly sequentially closed, we know $x \in [f \leq r]$, a contradiction. We conclude that f is sequentially lower semi-continuous for the weak topology. □

Lemma 2.3.3 *Let E be a Banach space, M a convex closed subset and $f : M \to \mathbb{R}$ a continuous convex function. Then f is lower semi-continuous on M for the weak topology.*

Proof. Because f is continuous (for the norm topology) the sub-level sets $[f \leq r]$, $r \in \mathbb{R}$, are all closed. Since f is convex these sub-level sets are convex subsets of E ($x, y \in [f \leq r]$, $0 \leq t \leq 1 \Rightarrow f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \leq tr + (1-t)r = r$). As in Hilbert spaces one knows that a convex subset is closed if, and only if, it is weakly closed. We deduce that all the sub-level sets are weakly closed and conclude by Lemma 2.3.2. □

As a conclusion to this section we present a summary of our discussion in the form of two explicit results on the existence of a minimizer in reflexive Banach spaces.

Theorem 2.3.4 (Generalized Weierstrass theorem I) *A weakly sequentially lower semi-continuous function f attains its infimum on a* *bounded and weakly sequentially closed subset M of a real reflexive Banach space E, i.e., there is $x_0 \in M$ such that*

$$f(x_0) = \inf_{x \in M} f(x).$$

Proof. All the sub-level sets $[f \leq r]$, $r \in \mathbb{R}$, are bounded and therefore relatively weakly compact since we are in a reflexive Banach space (see Lemma 2.3.1). Now Lemma 2.3.2 implies that all hypotheses of Theorem 2.2.4 are satisfied. Thus we conclude by this theorem. □

In Theorem 2.3.4 one can replace the assumption that the set M is bounded by an assumption on the function f which implies that the sub-level sets of f are bounded. Then one obtains another generalized Weierstrass theorem.

Theorem 2.3.5 (Generalized Weierstrass theorem II) *Let E be a reflexive Banach space, $M \subset E$ a weakly (sequentially) closed subset, and $f : M \to \mathbb{R}$ a weakly (sequentially) lower semi-continuous function on M. If f is coercive , i.e., if $\|x\| \to \infty$ implies $f(x) \to +\infty$, then f has a finite minimum on M, i.e., there is a $x_0 \in M$ such that*

$$f(x_0) = \inf_{x \in M} f(x).$$

Proof. Since f is coercive the sub-level sets $[f \leq r]$ are not empty for sufficiently large r and are bounded. We conclude as in the previous result. □

For other variants of generalized Weierstrass theorems we refer to [Zei85]. Detailed results on the minimization of functionals of the form (1.2) can be found in [Dac89, JLJ98, Str00].

2.4 Minimization of special classes of functionals

For a self-adjoint compact operator A in the complex Hilbert space $\mathcal{H}$ consider the sesquilinear function $Q : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ defined by $Q(x,y) = \langle x, Ay \rangle + r\langle x, y \rangle$ for $r = \|A\| + c$ for some $c > 0$. This function has the following properties: $Q(x,x) \geq c \|x\|^2$ for all $x \in \mathcal{H}$ and for fixed $x \in \mathcal{H}$ the function $y \mapsto Q(x,y)$ is weakly continuous (since a compact operator maps weakly convergent sequences onto norm convergent ones). Then $f(x) =$

$Q(x, x)$ is a concrete example of a *quadratic functional* on $\mathcal{H}$ which has a unique minimum on closed balls B_r of $\mathcal{H}$. This minimization is actually a special case of the following result on the minimization of quadratic functionals on reflexive Banach spaces.

Theorem 2.4.1 (Minimization of quadratic forms) *Let E be a reflexive Banach space and Q a symmetric sesquilinear form on E having the following properties: There is a constant $c > 0$ such that $Q(x, x) \geq c \|x\|^2$ for all $x \in E$ and for fixed $x \in E$ the functional $y \mapsto Q(x, y)$ is weakly continuous on E. Then, for every $u \in E'$ and every $r > 0$, there is exactly one point $x_0 = x_0(u, r)$ which minimizes the functional*

$$f(x) = Q(x, x) - \operatorname{Re} u(x), \qquad x \in E$$

on the closed ball $B_r = \{x \in E : \|x\| \leq r\}$, i.e.,

$$f(x_0) = \inf_{x \in B_r} f(x).$$

Proof. Consider $x, y \in E$ and $0 < t < 1$, then a straightforward calculation gives

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y) - t(1-t)Q(x-y, x-y) < tf(x) + (1-t)f(y).$$

for all $x, y \in E$, $x \neq y$, since then $t(1-t)Q(x-y, x-y) > 0$, hence the functional f is strictly convex and thus has at most one minimizing point by Theorem 2.2.5.

Suppose a sequence $(x_n)_{n \in \mathbb{N}}$ in E converges weakly to $x_0 \in E$. Since $Q(x_n, x_n) = Q(x_0, x_0) + Q(x_0, x_n - x_0) + Q(x_n - x_0, x_0) + Q(x_n - x_0, x_n - x_0)$ and since Q is strictly positive it follows that

$$Q(x_n, x_n) \geq Q(x_0, x_0) + Q(x_n - x_0, x_0) + Q(x_0, x_n - x_0)$$

for all $n \in \mathbb{N}$. Since Q is symmetric and weakly continuous in the second argument the last two terms converge to 0 as $n \to \infty$ and this estimate implies

$$\liminf_{n \to \infty} Q(x_n, x_n) \geq Q(x_0, x_0).$$

Therefore the function $x \mapsto Q(x, x)$ is weakly lower semi-continuous, thus, for every $u \in E'$, $x \mapsto f(x) = Q(x, x) - \operatorname{Re} u(x)$ is weakly lower semi-continuous on E and we conclude by Theorem 2.3.4 (Observe that the closed balls B_r are weakly closed, as closed convex sets). □

Corollary 2.4.2 *Let A be a bounded symmetric operator in complex Hilbert space $\mathcal{H}$ which is strictly positive, i.e., there is a constant $c > 0$ such that $\langle x, Ax \rangle \geq c\langle x, x \rangle$ for all $x \in \mathcal{H}$. Then, for every $y \in \mathcal{H}$ the function $x \mapsto f(x) = \langle x, Ax \rangle - \operatorname{Re} \langle y, x \rangle$ has a unique minimizing point $x_0 = x_0(y, r)$ on every closed ball B_r, i.e., there is exactly one $x_0 \in B_r$ such that*

$$f(x_0) = \inf_{x \in B_r} f(x).$$

Proof. Using the introductory remark to this section one verifies easily that $Q(x,y) = \langle x, Ay\rangle$ satisfies the hypothesis of Theorem 2.4.1. □

2.5 Exercises

1. Prove Lemma 2.2.2.

2. Show without the use of Lemma 2.3.2 that the norm $\|\cdot\|$ on a Banach space E is weakly lower semi-continuous.

   Hints: Recall that $\|x_0\| = \sup_{u\in E',\, \|u\|'\leq 1} |u(x_0)|$ for $x_0 \in E$. If a sequence $(x_n)_{n\in\mathbb{N}}$ converges weakly to x_0, then for every $u \in E'$ one knows $u(x_0) = \lim_{n\to\infty} u(x_n)$.

3. Prove: The functional
$$f(u) = \int_0^1 (tu'(t))^2 dt,$$
   defined on all continuous functions on $[0,1]$ which have a weak derivative $u' \in L^2(0,1)$ and which satisfy $u(0) = 0$ and $u(1) = 1$, has 0 as infimum and there is no function in this class at which the infimum is attained.

4. On the space $E = C^1([-1,1],\mathbb{R})$ define the functional
$$f(u) = \int_{-1}^1 (tu'(t))^2 dt$$
   and show that it has no minimum under the boundary conditions $u(\pm 1) = \pm 1$.

   Hints: This variation of the previous problem is due to Weierstrass. Show first that on the class of functions u_ϵ, $\epsilon > 0$, defined by
$$u_\epsilon(x) = \frac{\arctan \frac{x}{\epsilon}}{\arctan \frac{1}{\epsilon}},$$
   the infimum of f is zero.

Chapter 3

Differential Calculus on Banach Spaces

and extremal Points of differentiable Functions

As is well known from calculus on finite dimensional Euclidean spaces, the behavior of a sufficiently smooth function f in a neighborhood of some point x_0 is determined by the first few derivatives $f^{(n)}(x_0)$, $n \leq m$, of f at this point, $m \in \mathbb{N}$ depending on f and the intended accuracy. For example, if f is a twice continuously differentiable real valued function on the open interval $\Omega \subset \mathbb{R}$ and $x_0 \in \Omega$, the Taylor expansion of order 2

$$f(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{1}{2!} f^{(2)}(x_0)(x - x_0)^2 + (x - x_0)^2 R_2(x, x_0) \quad (3.1)$$

with $\lim_{x \to x_0} R_2(x, x_0) = 0$ is available, and on the basis of this representation the values of $f^{(1)}(x_0)$ and $f^{(2)}(x_0)$ determine whether x_0 is a critical point of the function f, or a local minimum, or a local maximum, or an inflection point.

In variational problems too one has to determine whether a function f has critical points, local minima or maxima or inflection points, but in these problems the underlying spaces are typically infinite dimensional Banach spaces. Accordingly an expansion of the form (3.1) in this infinite dimensional case can be expected to be an important tool too. Obviously one needs differential calculus on Banach spaces to achieve this goal.

Recall that differentiability of a real valued function f on an open interval Ω at a point $x_0 \in \Omega$ is equivalent to the existence of a proper tangent to the graph of the function through the point $(x_0, f(x_0)) \in \mathbb{R}^2$. A proper tangent means that the difference between the values of the tangent and of the function f at a point $x \in \Omega$ is of higher order in $x - x_0$ than the linear term. Since the tangent has the equation $y(x) = f^{(1)}(x_0)(x - x_0) + f(x_0)$ this approximation means

$$f(x) - y(x) = f(x) - f^{(1)}(x_0)(x - x_0) - f(x_0) = o(x - x_0) \quad (3.2)$$

where o is some function on $\mathbb{R}$ with the properties $o(0) = o$ and $\lim_{h \to 0, h \neq 0} \frac{o(h)}{h}$. In the case of a real valued function of several variables the tangent plane takes the rôle of the tangent line. As we are going to show, this way to look at differentiability has a natural counterpart for functions defined on infinite dimensional Banach spaces.

3.1 The Fréchet derivative

Let E, F be two real Banach spaces with norms $\|\cdot\|_E$, respectively $\|\cdot\|_F$. As usual $\mathcal{L}(E, F)$ denotes the space of all continuous linear operators from E into F. In Functional Analysis one shows that the space $\mathcal{L}(E, F)$ is a real Banach space too. The symbol o denotes any function $E \to F$ which is of higher than linear order in its argument, i.e., any function satisfying

$$o(0) = 0, \qquad \lim_{h \to 0, h \in E \setminus \{0\}} \frac{\|o(h)\|_F}{\|h\|_E} = 0. \quad (3.3)$$

Definition 3.1.1 *Let $U \subset E$ be a nonempty open subset of the real Banach space E and $f : U \to F$ a function from U into the real Banach space F. f is called* **Fréchet differentiable at a point** $x_0 \in U$ *if, and only if, there is an $\ell \in \mathcal{L}(E, F)$ such that*

$$f(x) = f(x_0) + \ell(x - x_0) + o(x_0; x - x_0) \qquad \forall\, x \in U. \quad (3.4)$$

If f is differentiable at $x_0 \in U$ the continuous linear operator $\ell \in \mathcal{L}(E,F)$ is called the **derivative of** f **at** x_0 *and is denoted by*

$$f'(x_0) \equiv D_{x_0}f \equiv Df(x_0) \equiv \ell. \tag{3.5}$$

If f is differentiable at every point $x_0 \in U$, f is called **differentiable on** U *and the function $Df : U \to \mathcal{L}(E,F)$ which assigns to every point $x_0 \in U$ the derivative $Df(x_0)$ of f at x_0 is called the* **derivative of the function** f.

If the derivative $Df : U \to \mathcal{L}(E,F)$ is continuous, the function f is called **continuously differentiable on** U *or* **of class $\mathcal{C}^1$**, *also denoted by $f \in \mathcal{C}^1(U,F)$.*

This definition is indeed meaningful because of the following

Lemma 3.1.2 *Under the assumptions of Definition 3.1.1 there is at most one $\ell \in \mathcal{L}(E,F)$ satisfying equation (3.4).*

Proof. Suppose there are $\ell_1, \ell_2 \in \mathcal{L}(E,F)$ satisfying equation (3.4). Then, for all $h \in B_r$ where B_r denotes an open ball in E with center 0 and radius $r > 0$ such that $x_0 + B_r \subset U$, we have $f(x_0) + \ell_1(h) + o_1(x_0,h) = f(x_0 + h) = f(x_0) + \ell_2(h) + o_2(x_0,h)$ and hence the linear functional $\ell = \ell_2 - \ell_1$ satisfies $\ell(h) = o_1(x_0,h) - o_2(x_0,h)$ for all $h \in B_r$. A continuous linear operator can be of higher than linear order on an open ball only if it is the null operator (see Exercises). This proves $\ell = 0$ and thus uniqueness. □

Definition 3.1.1 is easy to apply. Suppose $f : U \to F$ is constant, i.e., for some $a \in F$ we have $f(x) = a$ for all $x \in U \subset E$. Then $f(x) = f(x_0)$ for all $x, x_0 \in U$ and with the choice of $\ell = 0 \in \mathcal{L}(E,F)$ condition (3.4) is satisfied. Thus f is continuously Fréchet differentiable on U with derivative zero.

As another simple example consider the case were E is some real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and $F = \mathbb{R}$. For a continuous linear operator $A : E \to E$ define a function $f : E \to \mathbb{R}$ by $f(x) = \langle x, Ax \rangle$ for all $x \in E$. For $x, h \in E$ we calculate $f(x+h) = f(x) + \langle A^*x + Ax, h \rangle + f(h)$. $h \mapsto \langle A^*x + Ax, h \rangle$ is certainly a continuous linear functional $E \to \mathbb{R}$ and $f(h) = o(h)$ is obviously of higher than linear order (actually second order) in h. Hence f is Fréchet differentiable on E with derivative $f'(x) \in \mathcal{L}(E, \mathbb{R})$ given by $f'(x)(h) = \langle A^*x + Ax, h\rangle$ for all $h \in E$.

In the Exercises the reader will be invited to show that the above definition of differentiability reproduces the wellknown definitions of differentiability for functions of finitely many variables.

The Fréchet derivative has all the properties which are well known for the derivative of functions of one real variable. Indeed the following results hold.

Proposition 3.1.3 *Let $U \subset E$ be an open nonempty subset of the Banach space E and F some other real Banach space.*

a) The Fréchet derivative D is a linear mapping $\mathcal{C}^1(U, F) \to \mathcal{C}(U, F)$, i.e., for all $f, g \in \mathcal{C}^1(U, F)$ and all $a, b \in \mathbb{R}$ one has

$$D(af + bg) = aDf + bDg.$$

b) The chain rule holds for the Fréchet derivative D: Let $V \subset F$ be an open set containing $f(U)$ and G a third real Banach space. Then for all $f \in \mathcal{C}^1(U, F)$ and all $g \in \mathcal{C}^1(V, G)$ we have $g \circ f \in \mathcal{C}^1(U, G)$ and for all $x \in U$

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

Proof. The proof of the first part is left as an exercise.

Since f is differentiable at $x \in U$ we know

$$f(x+h) - f(x) = f'(x)(h) + o_1(h) \qquad \forall\, h \in B_r, \quad x + B_r \subset U$$

and similarly, since g is differentiable at $y = f(x) \in V$,

$$g(y+k) - g(y) = g'(y)(k) + o_2(k) \qquad \forall\, k \in B_\rho, \quad y + B_\rho \subset V.$$

Since f is continuous one can find, for the radius $\rho > 0$ in the differentiability condition for g, a radius $r > 0$ such that $f(B_r) \subseteq B_\rho$ and such that the differentiability condition for f holds. Then, for all $h \in B_r$, the following chain of identities holds, taking the above differentiability conditions into account:

$$\begin{aligned} g \circ f(x+h) - g \circ f(x) &= g[f(x+h)] - g[f(x)] \\ &= g[f(x) + f'(x)(h) + o_1(h)] - g[f(x)] \\ &= g'(y)(f'(x)(h) + o_1(h)) + o_2(f'(x)(h) + o_1(h)) \\ &= g'(y)(f'(x)(h)) + o(h) \end{aligned}$$

where

$$o(h) = g'(y)(o_1(h)) + o_2(f'(x)(h) + o_1(h))$$

is indeed a higher order term as shown in the Exercises. Thus we conclude. □

Higher order derivatives can be defined in the same way. Suppose E, F are two real Banach spaces and $U \subset E$ is open and nonempty. Given a function $f \in \mathcal{C}^1(U, F)$ we know that $f' \in \mathcal{C}(U, \mathcal{L}(E, F))$, i.e., the derivative is a continuous function on U with values in the Banach space $\mathcal{L}(E, F)$. If this function f' is differentiable at $x_0 \in U$ (on U), the function f is called *twice differentiable* at $x_0 \in U$ (on U) and is denoted by

$$D^2 f(x_0) = f^{(2)}(x_0) = D^2_{x_0} f \equiv D(f')(x_0). \tag{3.6}$$

According to Definition 3.1.1 and equation (3.6) the second derivative of $f : U \to F$ is a continuous linear operator $E \to \mathcal{L}(E, F)$, i.e., an element of the space $\mathcal{L}(E, \mathcal{L}(E, F))$. There is a natural isomorphism of the space of continuous linear operators from E into the space of continuous linear operators from E into F and the space $\mathcal{B}(E \times E, F)$ of continuous bilinear operators from $E \times E$ into F,

$$\mathcal{L}(E, \mathcal{L}(E, F)) \cong \mathcal{B}(E \times E, F). \tag{3.7}$$

This natural isomorphism is defined and studied in the Exercises. Thus the second derivative $D^2 f(x_0)$ at a point $x_0 \in U$ is considered as a continuous bilinear map $E \times E \to F$. If the second derivative $D^2 f : U \to \mathcal{B}(E \times E, F)$ exists on U and is continuous, the function f is said to be *of class* $\mathcal{C}^2$ and we write $f \in \mathcal{C}^2(U, F)$.

The derivatives of higher order are defined in the same way. The derivative of order $n \geq 3$ is the derivative of the derivative of order $n - 1$, according to Definition 3.1.1:

$$D^n f(x_0) = D(D^{n-1} f)(x_0). \tag{3.8}$$

In order to describe $D^n f(x_0)$ conveniently we extend the isomorphism (3.7) to higher orders. Denote by $E^{\times n} = E \times \cdots \times E$

(n factors) and by $\mathcal{B}(E^{\times n}, F)$ the Banach space of all continuous n-linear operators $E^{\times n} \to F$. In the Exercises one shows for $n = 3, 4, \dots$

$$\mathcal{L}(E, \mathcal{B}(E^{\times n-1}, F)) \cong \mathcal{B}(E^{\times n}, F). \tag{3.9}$$

Under this isomorphism the third derivative at some point $x_0 \in U$ is then a continuous 3-linear map $E^{\times 3} \to F$, $D^3 f(x_0) \in \mathcal{B}(E^{\times 3}, F)$. Using the isomorphisms (3.9) the higher order derivatives are

$$D^n f(x_0) \in \mathcal{B}(E^{\times n}, F) \tag{3.10}$$

if they exist. If $D^n f : U \to \mathcal{B}(E^{\times n}, F)$ is continuous the function f is called *n-times continuously differentiable* or *of class $\mathcal{C}^n$*. Then we write $f \in \mathcal{C}^n(U, F)$.

As an illustration we calculate the second derivative of the function $f(x) = \langle x, Ax \rangle$ on a real Hilbert space E with inner product $\langle \cdot, \cdot \rangle$, A a bounded linear operator on E. The first Fréchet derivative has been calculated, $f'(x_0)(y) = \langle (A + A^*)x_0, y \rangle$ for all $y \in E$. In order to determine the second derivative we evaluate $f'(x_0 + h) - f'(x_0)$. For all $y \in E$ one finds through a simple calculation

$$(f'(x_0 + h) - f'(x_0))(y) = \langle (A + A^*)h, y \rangle.$$

Hence the second derivative of f exists and is given by the continuous bilinear form $(D^2 f)(x_0)(y_1, y_2) = \langle (A + A^*)y_1, y_2 \rangle$, $y_1, y_2 \in E$. We see in this example that the second derivative is actually a symmetric bilinear form. With some effort this can be shown for every twice differentiable function.

As we have mentioned, the first few derivatives of a differentiable function $f : U \to F$ at a point $x_0 \in U$ control the behavior of the function in a sufficiently small neighborhood of this point. The key to this connection is the Taylor expansion with remainder. In order to be able to prove this fundamental result in its strongest form we need the fundamental theorem of calculus for functions with values in a Banach space. And this in turn requires the knowledge of the Riemann integral for functions on the real line with values in a Banach space.

Suppose E is a real Banach space and $u : [a, b] \to E$ a continuous function on the bounded interval $[a, b]$. Roughly, a partition Z of the interval $[a, b]$ is an ordered family of points $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ and of some points $t'_j \in (t_{j-1}, t_j]$, $j = 1, \ldots, n$. For each partition we introduce the approximating sums

$$\Sigma(u, Z) = \sum_{j=1}^{n} u(t'_j)(t_j - t_{j-1}).$$

By forming the joint refinement of two partitions one shows, the following result: Given $\epsilon > 0$ there is $\delta > 0$ such that

$$\|\Sigma(u, Z) - \Sigma(u, Z')\|_E < \epsilon$$

for all partitions Z, Z' with $|Z'|, |Z| < \delta$,

$$|Z| = \max\left\{t_j - t_{j-1} : j = 1, \ldots, n\right\}.$$

This estimate implies that the approximating sums $\Sigma(u, Z)$ have a limit with respect to partitions Z with $|Z| \to 0$.

Theorem 3.1.4 *Suppose E is a real Banach space and $u : [a, b] \to E$ a continuous function. Then u has an integral over this finite interval, defined by the following limit in E:*

$$\int_a^b u(t)dt = \lim_{|Z| \to 0} \Sigma(u, Z). \tag{3.11}$$

This integral of functions with values in a Banach space has the standard properties, i.e., it is linear in the integrand, additive in the interval of integration, and is bounded by the maximum of the function multiplied by the length of the integration interval:

$$\left\|\int_a^b u(t)dt\right\|_E \leq (b - a) \max_{a \leq t \leq b} \|u(t)\|_E.$$

Proof. It is straightforward to verify that the approximating sums $\Sigma(u, Z)$ are linear in u and additive in the interval of integration. The basic rules of calculation for limits then prove the statements for the integral. For the estimate observe

$$\|\Sigma(u,Z)\|_E \leq \sum_{j=1}^{n} \left\|u(t_j')\right\|_E (t_j - t_{j-1}) \leq \sup_{a\leq t\leq b} \|u(t)\|_E \sum_{j=1}^{n} (t_j - t_{j-1})$$

which implies the above estimate for the approximating sums. Thus we conclude. □

Corollary 3.1.5 (Fundamental theorem of calculus) *Let E be a real Banach space, $[a, b]$ a finite interval and $u : [a, b] \to E$ a continuous function. For some $e \in E$ define a function $v : [a, b] \to E$ by*

$$v(t) = e + \int_a^t u(s)ds \qquad \forall\, s \in [a, b]. \tag{3.12}$$

Then v is continuously differentiable with derivative $v'(t) = u(t)$ and one thus has for all $a \leq c < d \leq b$,

$$v(d) - v(c) = \int_c^d v'(t)dt. \tag{3.13}$$

Proof. We prove differentiability of v at some interior point $t \in (a, b)$. At the end points of the interval the usual modifications apply. Suppose $\tau > 0$ such that $t + \tau \in [a, b]$. Then, by definition of v,

$$v(t+\tau) - v(t) = \int_a^{t+\tau} u(s)ds - \int_a^t u(s)ds = \int_t^{t+\tau} u(s)ds$$

since

$$\int_a^{t+\tau} u(s)ds = \int_a^t u(s)ds + \int_t^{t+\tau} u(s)ds.$$

The basic bound for integrals gives

$$\left\| \int_t^{t+\tau} [u(s) - u(t)]ds \right\|_E \leq \tau \sup_{t\leq s\leq t+\tau} \|u(s) - u(t)\|_E$$

and thus proves that this integral is of higher order in τ. We deduce $v(t + \tau) = v(t) + \tau u(t) + o(\tau)$ and conclude that v is differentiable at t with derivative $v'(t) = u(t)$. The rest of the proof is standard. □

Theorem 3.1.6 (Taylor expansion with remainder) *Suppose E, F are real Banach spaces, $U \subset E$ an open and nonempty subset, and $f \in \mathcal{C}^n(U, F)$. Given $x_0 \in U$ choose $r > 0$ such that $x_0 + B_r \subset U$ where B_r is the open ball in E with center 0 and radius r. Then for all $h \in B_r$ we have, using the abbreviation $(h)^k = (h, \ldots, h)$, k terms,*

$$f(x_0 + h) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(h)^k + R_n(x_0; h) \tag{3.14}$$

where the remainder R_n has the form

$$R_n(x_0;h) = \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}[f^{(n)}(x_0+th) - f^{(n)}(x_0)](h)^n dt \tag{3.15}$$

and thus is of order $o((h)^n)$, i.e.,

$$\lim_{h\to 0, h\in E\setminus\{0\}} \frac{\|R_n(x_0;h)\|_F}{\|h\|_E^n} = 0.$$

Proof . Basically the Taylor formula is obtained by applying the fundamental theorem of calculus repeatedly (n times) and transforming the multiple integral which is generated in this process by a change of the integration order into a one-dimensional integral.

However there is a simplification of the proof based on the following observation (see [YCB82]). Let v be a function on $[0,1]$ which is n times continuously differentiable, then

$$\frac{d}{dt}\sum_{k=0}^{n-1}\frac{(1-t)^k}{k!}v^{(k)}(t) = \frac{(1-t)^{n-1}}{(n-1)!}v^{(n)}(t) \qquad \forall\, t\in[0,1].$$

The proof of this identity follows simply by differentiation and grouping terms together appropriately.

Integrate this identity for the function $v(t) = f(x_0+th)$. Since $f \in \mathcal{C}^n(U,F)$ the application of the chain rule yields for $h \in B_r$,

$$v^{(k)}(t) = f^{(k)}(x_0+th)(h)^k$$

and thus the result of this integration is, using Equation 3.13,

$$f(x_0+h) = \sum_{k=0}^{n-1}\frac{1}{k!}f^{(k)}(x_0)(h)^k + R$$

with remainder

$$R = \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}f^{(n)}(x_0+th)(h)^n dt$$

which can be written as

$$R = \frac{1}{n!}f^{(n)}(x_0)(h)^n + \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}[f^{(n)}(x_0+th) - f^{(n)}(x_0)](h)^n dt.$$

The differentiability assumption for f implies that the function $h \mapsto f^{(n)}(x_0+th)$ from B_r into $\mathcal{B}(E^{\times n},F)$ is continuous, hence

$$\left\|[f^{(n)}(x_0+th) - f^{(n)}(x_0)]\right\|_{\mathcal{B}(E^{\times n},F)} \to 0$$

as $h \to 0$. Thus we conclude. □

3.2 Extrema of differentiable functions

Taylor's formula (3.14) says that a function $f : U \to F$ of class $\mathcal{C}^n$ is approximated at each point of a neighborhood of some point $x_0 \in U$ by a polynomial of degree n, and the error is of order $o((x-x_0)^n)$. We apply now this approximation for $n=2$ to characterize local extrema of a function of class $\mathcal{C}^2$ in terms of the first and second derivative of f. We begin with the necessary definitions.

Definition 3.2.1 *Let E be a real Banach space, $M \subseteq E$ a nonempty subset, and $f: M \to R$ a real valued function on M. A point $x_0 \in M$ is called a* **local minimum (maximum)** *of f on M if there is some $r>0$ such that*

$$f(x_0) \leq f(x), \quad (f(x_0) \geq f(x)) \qquad \forall\, x \in M \cap (x_0 + B_r).$$

A local minimum (maximum) is **strict** *if*

$$f(x_0) < f(x), \quad (f(x_0) > f(x)) \qquad \forall\, x \in M \cap (x_0 + B_r),\ x \neq x_0.$$

If $f(x_0) \leq f(x)$, $(f(x_0) \geq f(x))$ holds for all $x \in M$, we call x_0 a **global minimum (maximum)**.

Definition 3.2.2 *Suppose E, F are two real Banach spaces, $U \subset E$ an open nonempty subset, and $f: U \to F$ a function of class $\mathcal{C}^1$. A point $x_0 \in U$ is called a* **regular** *(***critical***)* **point** *of the function f if, and only if, the Fréchet derivative $Df(x_0)$ of f at x_0 is surjective (not surjective).*

Remark 3.2.3 *For the case $F = \mathbb{R}$ the Fréchet derivative $Df(x_0) = f'(x_0) \in \mathcal{L}(E, \mathbb{R})$ is not surjective, if and only if, $f'(x_0) = 0$; hence the notion of a critical point introduced above is nothing else than the generalization of the corresponding notion introduced in elementary calculus.*

For extremal points which are interior points of the domain M of the function f a fairly detailed description can be given. In this situation we can assume that the domain $M = U$ is an open set.

Theorem 3.2.4 (Necessary condition of Euler – Lagrange) *Suppose U is an open nonempty subset of the real Banach space E and*

$f \in \mathcal{C}^1(U, \mathbb{R})$. Then every extremal point (i.e., every local or global minimum and every local or global maximum) is a critical point of f.

Proof. Suppose that $x_0 \in U$ is a local minimum of f. Then there is an $r > 0$ such that $x_0 + B_r \subset U$ and $f(x_0) \leq f(x_0 + h)$ for all $h \in B_r$. Since $f \in \mathcal{C}^1(U, \mathbb{R})$ Taylor's formula applies, thus

$$f(x_0) \leq f(x_0 + h) = f(x_0) + f'(x_0)(h) + R_1(x_0, h) \qquad \forall\, h \in B_r$$

or

$$0 \leq f'(x_0)(h) + R_1(x_0, h) \qquad \forall\, h \in B_r.$$

Choose any $h \in B_r$, $h \neq 0$. Then all $th \in B_r$, $0 < t \leq 1$ and therefore $0 \leq f'(x_0)(th) + R_1(x_0, th)$. Since $\lim_{t \to 0} t^{-1} R_1(x_0, th) = 0$ we can divide this inequality by $t > 0$ and take the limit $t \to 0$. This gives $0 \leq f'(x_0)(h)$. This argument applies to any $h \in B_r$, thus in particular to $-h$ and therefore $0 \leq f'(x_0)(-h) = -f'(x_0)(h)$. We conclude that $0 = f'(x_0)(h)$ for all $h \in B_r$. The open nonempty ball B_r absorbs the points of E, i.e., every point $x \in E$ can be written as $x = \lambda h$ with some $h \in B_r$ and some $\lambda \in \mathbb{R}$. It follows that $0 = f'(x_0)(x)$ for all $x \in E$ and therefore $f'(x_0) = 0 \in \mathcal{L}(E, \mathbb{R}) = E'$.

If $x_0 \in U$ is a local maximum of f, then this point is a local minimum of $-f$ and we conclude as above. □

Theorem 3.2.5 (Nec. & suff. conditions for local extrema) *Suppose $U \subset E$ is a nonempty open subset of the real Banach space E and $f \in \mathcal{C}^2(U, \mathbb{R})$.*

a) If f has a local minimum at $x_0 \in U$, then the first Fréchet derivative of f vanishes at x_0, $f'(x_0) = 0$, and the second Fréchet derivative of f is nonnegative at x_0, $f^{(2)}(x_0)(h, h) \geq 0$ for all $h \in E$.

b) If conversely $f'(x_0) = 0$ and if the second Fréchet derivative of f is strictly positive at x_0, i.e., if

$$\inf \left\{ f^{(2)}(x_0)(h, h) : h \in E,\ \|h\|_E = 1 \right\} = c > 0,$$

then f has a local minimum at x_0.

Proof. Suppose $x_0 \in U$ is a local minimum of f. Then by Theorem 3.2.4 $f'(x_0) = 0$. Since $f \in \mathcal{C}^2(U, \mathbb{R})$ Taylor's formula implies

$$f(x_0) \leq f(x_0 + h) = f(x_0) + \frac{1}{2!} f^{(2)}(x_0)(h, h) + R_2(x_0, h) \qquad \forall\, h \in B_r \tag{3.16}$$

for some $r > 0$ such that $x_0 + B_r \subset U$. Choose any $h \in B_r$. Then for all $0 < t \leq 1$ we know $0 \leq \frac{1}{2!} f^{(2)}(x_0)(th, th) + R_2(x_0, th)$ or, after division by $t^2 > 0$.

$$0 \leq f^{(2)}(x_0)(h, h) + \frac{2}{t^2} R_2(x_0, th) \qquad \forall\, 0 < t \leq 1.$$

Since $R_2(x_0, th)$ is a higher order term we know $t^{-2} R_2(x_0, th) \to 0$ as $t \to 0$. This gives $0 \leq f^{(2)}(x_0)(h, h)$ for all $h \in B_r$ and since open balls are absorbing, $0 \leq f^{(2)}(x_0)(h, h)$ for all $h \in E$. This proves Part a).

Conversely assume that $f'(x_0) = 0$ and that $f^{(2)}(x_0)$ is strictly positive. Choose $r > 0$ such that $x_0 + B_r \subset U$. The second order Taylor expansion gives

$$f(x_0 + h) - f(x_0) = \frac{1}{2!} f^{(2)}(x_0)(h, h) + R_2(x_0, h) \qquad \forall h \in B_r,$$

and thus for all $h \in E$ with $\|h\|_E = 1$ and all $0 < s < r$,

$$\begin{aligned} f(x_0 + sh) - f(x_0) &= \frac{1}{2!} f^{(2)}(x_0)(sh, sh) + R_2(x_0, sh) \\ &= s^2 [\frac{1}{2!} f^{(2)}(x_0)(h, h) + s^{-2} R_2(x_0, sh)]. \end{aligned}$$

Since $R_2(x_0, sh)$ is a higher order term there is an $s_0 \in (0, r)$ such that $|s^{-2} R_2(x_0, sh)| < c/2$ for all $0 < s \leq s_0$, and since $\frac{1}{2!} f^{(2)}(x_0)(h, h) \geq c/2$ for all $h \in E$, $\|h\|_E = 1$, we get $[\frac{1}{2!} f^{(2)}(x_0)(h, h) + s^{-2} R_2(x_0, sh)] \geq 0$ for all $0 < s < s_0$ and all $h \in E$, $\|h\|_E = 1$. It follows that $f(x_0 + h) - f(x_0) \geq 0$ for all $h \in B_{s_0}$ and therefore the function f has a local minimum at x_0. □

As we mentioned before a function f has a local maximum at some point x_0 if, and only if, the function $-f$ has a local minimum at this point. Therefore Theorem 3.2.5 easily implies necessary and sufficient conditions for a local maximum.

3.3 Convexity and monotonicity

We begin with the discussion of an interesting connection between convexity of a functional and monotonicity of its first Fréchet derivative which has far-reaching implications for optimization problems. For differentiable real valued functions of one real variable these results are well known.

The following theorem states this connection in detail and provides the relevant definitions.

Theorem 3.3.1 (Convexity – Monotonicity) *Let U be a convex open subset of the real Banach space E and $f \in \mathcal{C}^1(U, \mathbb{R})$. Then the following statements are equivalent:*

a) f is **convex**, *i.e., for all $x, y \in U$ and all $0 \leq t \leq 1$ one has*

$$f(tx + (1 - t)y) \leq t f(x) + (1 - t) f(y); \tag{3.17}$$

b) The Fréchet derivative $f' : E \to E'$ is **monotone**, *i.e., for all $x, y \in U$ one has*

$$\langle f'(x) - f'(y), x - y \rangle \geq 0 \tag{3.18}$$

where $\langle \cdot, \cdot \rangle$ denotes the canonical bilinear form on $E' \times E$.

Proof. If f is convex inequality (3.17) implies, for $x, y \in U$ and $0 < t \leq 1$,

$$f(y + t(x-y)) - f(y) \leq tf(x) + (1-t)f(y) - f(y) = t(f(x) - f(y)).$$

If we divide this inequality by $t > 0$ and then take the limit $t \to 0$ the result is

$$\langle f'(y), x - y \rangle \leq f(x) - f(y).$$

If we exchange the rôles of x and y in this argument we obtain

$$\langle f'(x), y - x \rangle \leq f(y) - f(x).$$

Now add the two inequalities to get

$$\langle f'(y), x - y \rangle + \langle f'(x), y - x \rangle \leq 0,$$

thus condition (3.18) follows and therefore f' is monotone.

Suppose conversely that the Fréchet derivative $f' : E \to E'$ is monotone. For $x, y \in U$ and $0 \leq t \leq 1$ consider the function $p : [0,1] \to \mathbb{R}$ defined by $p(t) = f(tx + (1-t)y) - tf(x) - (1-t)f(y)$. This function is differentiable with derivative $p'(t) = \langle f'(x(t)), x - y \rangle - f(x) + f(y)$, $x(t) = tx + (1-t)y$, and satisfies $p(0) = 0 = p(1)$. The convexity condition is equivalent to the condition $p(t) \leq 0$ for all $t \in [0,1]$. We prove this condition indirectly. Thus we assume that there is some point in $(0,1)$ at which p is positive. Then there is some point $t_0 \in (0,1)$ at which p attains its positive maximum. For $t \in (0,1)$ calculate

$$\begin{aligned}(t - t_0)(p'(t) - p'(t_0)) &= (t - t_0)\langle f'(x(t)) - f'(x(t_0)), x - y \rangle \\ &= \langle f'(x(t)) - f'(x(t_0)), x(t) - x(t_0) \rangle.\end{aligned}$$

Since f' is monotone it follows that $(t - t_0)(p'(t) - p'(t_0) \geq 0$. Since p attains its maximum at t_0, $p'(t_0) = 0$, and thus $(t - t_0)p'(t) \geq 0$, hence $p'(t) \geq 0$ for all $t_0 < t \leq 1$, a contradiction. We conclude $p(t) \leq 0$ and thus condition (3.17). □

Corollary 3.3.2 *Let U be a nonempty convex open subset of the real Banach space E and $f \in \mathcal{C}^1(U, \mathbb{R})$. If f is convex, then every critical point of f is actually a minimizing point , i.e., a point at which f has a local minimum.*

Proof. If $x_0 \in U$ is a critical point, there is an $r > 0$ such that $x_0 + B_r \subset U$. Then for every $h \in B_r$ the points $x(t) = x_0 + th$, $0 \leq t \leq 1$, belong to $x_0 + B_r$. Since f is differentiable we find

$$f(x_0 + h) - f(x_0) = \int_0^1 \frac{d}{dt} f(x(t))dt = \int_0^1 \langle f'(x(t)), h \rangle dt.$$

Since $x(t) - x_0 = th$ the last integral can be written as:

$$= \lim_{\epsilon \downarrow 0} \int_\epsilon^1 \langle f'(x(t)) - f'(x_0), x(t) - x_0 \rangle \frac{dt}{t}.$$

Theorem 3.3.1 implies that the integrand of this integral is non-negative, hence $f(x_0+h) - f(x_0) \geq 0$ for all $h \in B_r$ and f has a local minimum at the critical point x_0. $\square$

Corollary 3.3.3 *Let U be a nonempty convex open subset of the real Banach space E and $f \in \mathcal{C}^1(U, \mathbb{R})$. If f is convex, then f is weakly lower semi-continuous.*

Proof. Suppose that a sequence $(x_n)_{n\in\mathbb{N}} \subset U$ converges weakly to $x_0 \in U$. Again differentiability of f implies

$$f(x_n) - f(x_0) = \int_0^1 \frac{d}{dt} f(x_0 + t(x_n - x_0))dt = \int_0^1 \langle f'(x_0 + t(x_n - x_0)), x_n - x_0\rangle dt$$
$$= \int_0^1 \langle f'(x_0 + t(x_n - x_0)) - f'(x_0), x_n - x_0\rangle dt + \langle f'(x_0), x_n - x_0\rangle.$$

As in the proof of the previous corollary, monotonicity of f' implies that the integral is not negative, hence

$$f(x_n) - f(x_0) \geq \langle f'(x_0), x_n - x_0\rangle.$$

As $n \to \infty$ the righthand side of this estimate converges to 0 and thus $\liminf_{n\to\infty} f(x_n) - f(x_0) \geq 0$ or $\liminf_{n\to\infty} f(x_n) \geq f(x_0)$. This shows that f is weakly lower semi-continuous at x_0. Since $x_0 \in U$ was arbitrary, we conclude. $\square$

Corollary 3.3.4 *Let U be a nonempty convex open subset of the real Banach space E and $f \in \mathcal{C}^2(U, \mathbb{R})$. Then f is convex if, and only if, $f^{(2)}(x_0)$ is non-negative for all $x_0 \in U$, i.e., $f^{(2)}(x_0)(h,h) \geq 0$ for all $h \in E$.*

Proof. By Theorem 3.3.1 we know that f is convex if, and only if, its Fréchet derivative f' is monotone. Suppose f' is monotone and $x_0 \in U$. Then there is an $r > 0$ such that $x_0 + B_r \subset U$ and $\langle f'(x_0+h) - f'(x_0), h\rangle \geq 0$ for all $h \in B_r$. Since $f \in \mathcal{C}^2(U, \mathbb{R})$, Taylor's Theorem implies that

$$\langle f'(x_0+h) - f'(x_0), h\rangle = f^{(2)}(x_0)(h,h) + R_2(x_0,h),$$

hence

$$0 \leq f^{(2)}(x_0)(h,h) + R_2(x_0,h) \qquad \forall\, h \in B_r.$$

Since $R_2(x_0,h) = o((h)^2)$ we deduce, as in the proof of Theorem 3.2.5, that $0 \leq f^{(2)}(x_0)(h,h)$ for all $h \in E$. Thus $f^{(2)}$ is nonnegative at $x_0 \in U$.

Conversely assume that $f^{(2)}$ is nonnegative on U. For $x, y \in U$ we know

$$\langle f'(x) - f'(y), x - y\rangle = \int_0^1 f^{(2)}(y + t(x-y))(x-y, x-y)dt.$$

By assumption the integrand is nonnegative, and it follows that f' is monotone. $\square$

3.4 Gâteaux derivatives and variations

For functions $f : \mathbb{R}^n \to \mathbb{R}$ one has the concepts of the total differential and that of partial derivatives. The Fréchet derivative has been introduced as the generalization of the total differential to the case of infinite dimensional Banach spaces. Now we introduce the Gâteaux derivatives as the counterpart of the partial derivatives.

Definition 3.4.1 *Let E, F be two real Banach spaces, $U \subseteq E$ a nonempty open subset, and $f : U \to F$ a mapping from U into F. The* **Gâteaux differential of f at a point** *$x_0 \in U$ is a mapping $\delta f(x_0, \cdot) : E \to F$ such that, for all $h \in E$,*

$$\lim_{t\to 0, t\neq 0} \frac{1}{t}(f(x_0 + th) - f(x_0)) = \delta f(x_0, h). \tag{3.19}$$

$\delta f(x_0, h)$ is called the **Gâteaux differential of f at the point** *x_0* **in the direction** *$h \in E$. If the Gâteaux differential of f at x_0 is a continuous linear map $E \to F$, one writes*

$$\delta f(x_0, h) = \delta_{x_0} f(h)$$

and calls $\delta_{x_0} f$ **the Gâteaux derivative of f at the point** *x_0.*

Basic properties of the Gâteaux differential, respectively derivative, are collected in the following

Lemma 3.4.2 *Let E, F be two real Banach spaces, $U \subseteq E$ a nonempty open subset, and $f : U \to F$ a mapping from U into F.*

a) If the Gâteaux differential of f exists at a point $x_0 \in U$, it is a homogeneous map $E \to F$, i.e., $\delta f(x_0, \lambda h) = \lambda \delta f(x_0, h)$ for all $\lambda \in \mathbb{R}$ and all $h \in E$;

b) If the Gâteaux derivatives exist at a point $x \in U$, they are linear in f, i.e., for $f, g : U \to F$ and $\alpha, \beta \in \mathbb{R}$ one has $\delta_x(\alpha f + \beta g) = \alpha \delta_x f + \beta \delta_x g$;

c) *If f is Gâteaux differentiable at a point $x \in U$, then f is continuous at x in every direction $h \in E$;*

d) *Suppose G is a third real Banach space, $V \subseteq F$ a nonempty open subset such that $f(U) \subseteq V$ and $g : V \to G$ a mapping from V into G. If f has a Gâteaux derivative at $x \in U$ and g has a Gâteaux derivative at $y = f(x)$, then $g \circ f : U \to G$ has a Gâteaux derivative at $x \in U$ and the chain rule*

$$\delta_x(g \circ f) = \delta_y g \circ \delta_x f$$

holds.

Proof: Parts a) and b) follow easily from the basic rules of calculation for limits. Part c) is obvious from the definitions. The proof of the chain rule is similar but easier than the proof of this rule for the Fréchet derivative and thus we leave it as an exercise. □

The following result establishes the important connection between Fréchet and Gâteaux derivatives, as a counterpart of the connection between total differential and partial derivatives for functions of finitely many variables.

Lemma 3.4.3 *Let E, F be two real Banach spaces, $U \subseteq E$ a nonempty open subset, and $f : U \to F$ a mapping from U into F.*

a) *If f is Fréchet differentiable at a point $x \in U$, then f is Gâteaux differentiable at x and both derivatives are equal: $\delta_x f = D_x f$.*

b) *Suppose that f is Gâteaux differentiable at all points in a neighborhood V of the point $x_0 \in U$ and that $x \mapsto \delta_x f \in \mathcal{L}(E, F)$ is continuous on V. Then f is Fréchet differentiable at x_0 and $\delta_{x_0} f = D_{x_0} f$.*

Proof. If f is Fréchet differentiable at $x \in U$ we know, for all $h \in E$, $f(x + th) = f(x) + (D_x f)(th) + o(th)$, hence

$$\lim_{t\to 0} \frac{1}{t}(f(x + th) - f(x)) = (D_x f)(h) + \lim_{t\to 0} \frac{o(th)}{t} = (D_x f)(h),$$

and Part a) follows.

If f is Gâteaux differentiable in the neighborhood V of $x_0 \in U$, there is an $r > 0$ such that f is Gâteaux differentiable at all points $x_0 + h$, $h \in B_r$. Given $h \in B_r$ it follows that $g(t) = f(x_0 + th)$ is differentiable at all points $t \in [0,1]$ and $g'(t) = (\delta_{x_0+th} f)(h)$. This implies

$$g(1) - g(0) = \int_0^1 g'(t)dt = \int_0^1 (\delta_{x_0+th} f)(h)dt$$

and thus

$$f(x_0 + h) - f(x_0) - (\delta_{x_0} f)(h) = g(1) - g(0) - (\delta_{x_0} f)(h) = \int_0^1 [(\delta_{x_0+th} f)(h) - (\delta_{x_0} f)(h)]dt.$$

The integral can be estimated in norm by

$$\sup_{0 \le t \le 1} \left\| (\delta_{x_0+th} f) - (\delta_{x_0} f) \right\|_{\mathcal{L}(E,F)} \|h\|_E$$

and therefore

$$\|f(x_0 + h) - f(x_0) - (\delta_{x_0} f)(h)\|_F \le \sup_{0 \le t \le 1} \left\| (\delta_{x_0+th} f) - (\delta_{x_0} f) \right\|_{\mathcal{L}(E,F)} \|h\|_E .$$

Continuity of $(\delta_x f)$ in $x \in x_0 + B_r$ implies $f(x_0 + h) - f(x_0) - (\delta_{x_0} f)(h) = o(h)$ and thus f is Fréchet differentiable at x_0 and $(D_{x_0} f)(h) = (\delta_{x_0} f)(h)$ for all $h \in B_r$ and therefore for all $h \in E$. □

Lemma 3.4.3 can be very useful in finding the Fréchet derivative of functions. We give a simple example. On the Banach space $E = L^p(\mathbb{R}^n)$, $1 < p < 2$, consider the functional

$$f(u) = \int_{\mathbb{R}^n} |u(x)|^p dx, \qquad \forall\, u \in E.$$

To prove directly that f is continuously Fréchet differentiable on E is not so simple. If however Lemma 3.4.3 is used the proof becomes a straightforward calculation. We only need to verify the hypotheses of this lemma. In the Exercises the reader is asked to show that there are constants $0 < c < C < \infty$ such that

$$c|s|^p \le |1 + s|^p - 1 - ps \le C|s|^p \qquad \forall\, s \in \mathbb{R}.$$

Insert $s = t\frac{h(x)}{u(x)}$, for all points $x \in \mathbb{R}^n$ with $u(x) \neq 0$ and multiply with $|u(x)|^p$. The result is

$$c|th(x)|^p \le |u(x) + th(x)|^p - |u(x)|^p - pth(x)|u(x)|^{p-1}\mathrm{sgn}(u(x)) \le C|th(x)|^p.$$

Integration of this inequality gives

$$c|t|^p f(h) \leq f(u+th) - f(u) - pt \int_{\mathbb{R}^n} h(x)v(x)dx \leq C|t|^p f(h)$$

where

$$v(x) = |u(x)|^{p-1}\mathrm{sgn}(u(x)).$$

Note that $v \in L^q(\mathbb{R}^n)$, $\frac{1}{q} + \frac{1}{p} = 1$ and that $L^q(\mathbb{R}^n)$ is (isomorphic to) the topological dual of $E = L^p(\mathbb{R}^n)$. This estimate allows us to determine easily the Gâteaux derivative of f:

$$\delta f(u,h) = \lim_{t \to 0} \frac{1}{t}[f(u+th) - f(u)] = p \int_{\mathbb{R}^n} v(x)h(x)dx.$$

Hölder's inequality implies that the absolute value of this integral is bounded by $\|v\|_q \ \|h\|_p$, hence $h \mapsto \delta f(u,h)$ is a continuous linear functional on E and

$$\|\delta_u f\|_{\mathcal{L}(E,\mathbb{R})} = \|v\|_q = \|u\|_p^{p/q}.$$

Therefore $u \mapsto \delta_u f$ is a continuous map from $E \to \mathcal{L}(E,\mathbb{R})$ and Lemma 3.4.3 implies that f is Fréchet differentiable with derivative $D_u f(h) = \delta_u f(h)$.

Suppose that M is a nonempty subset of the real Banach space E which is not open, for instance M has a nonempty interior and part of the boundary of M belongs to M. Suppose furthermore that a function $f : M \to \mathbb{R}$ attains a local minimum at the boundary point x_0. Then we cannot investigate the behavior of f in terms of the first few Fréchet or Gâteaux derivatives of f at the point x_0 as we did previously since this required that a whole neighborhood of x_0 is contained in M. In such situations the *variations of the function* in suitable directions are a convenient tool to study the local behavior of f.

Assume that $h \in E$ and that there is some $r = r_h > 0$ such that $x(t) = x_0 + th \in M$ for all $0 \leq t < r$. Then we can study the function $f_h(t) = f(x(t))$ on the interval $[0, r)$. Certainly, if f has a local minimum at x_0, then $f_h(0) \leq f_h(t)$ for all $t \in [0, r)$ (if necessary we can decrease the value of r) and this gives restrictions on the first few derivatives of f_h, if they exist. These derivatives are then called the variations of f.

Definition 3.4.4 *Let $M \subset E$ be a nonempty subset of the real Banach space E and $x_0 \in M$. For $h \in E$ suppose that there is an $r > 0$ such that $x_0 + th \in M$ for all $0 \leq t < r$. Then the n***th variation** *of f in the direction h is defined as*

$$\triangle^n f(x_0, h) = \frac{d^n}{dt^n} f(x_0 + th)|_{t=0} \qquad n = 1, 2, \ldots \tag{3.20}$$

if these derivatives exist.

In favorable situations obviously the first variation is just the Gâteaux derivative:

Lemma 3.4.5 *Suppose that M is a nonempty subset of the real Banach space E, x_0 an interior point of M, and f a real valued function on M. Then the Gâteaux derivative $\delta_{x_0} f$ of f at x_0 exists if, and only if, the first variation $\triangle f(x_0, h)$ exists for all $h \in E$ and $h \mapsto \triangle f(x_0, h)$ is a continuous linear functional on E.*

In this case one has $\triangle f(x_0, h) = \delta_{x_0} f$.

Proof. A straightforward inspection of the respective definitions easily proves this lemma. □

3.5 Exercises

1. Complete the proof of Lemma 3.1.2.
2. Let E and F be two real normed spaces and $A : E \to F$ a continuous linear operator such that $Ax = o(x)$ for all $x \in E$, $\|x\| < 1$. Prove: $A = 0$.

3. For a function $f : U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, assume that it is differentiable at a point $x_0 \in U$. Use Definition 3.1.1 to determine the Fréchet derivative $f'(x_0)$ of f at x_0 and relate it to the Jabobi matrix $\frac{\partial f}{\partial x}(x_0)$ of f at x_0.

4. Prove Part a) of Proposition 3.1.3.

5. Prove that $o(h) = g'(y)(o_1(h)) + o_2(f'(x)(h) + o_1(h))$ is a higher order term, under the assumptions of Proposition 3.1.3, Part b).

6. Let $I = [a, b]$ be some finite closed interval. Equip the space $E = \mathcal{C}^1(I, \mathbb{R})$ of all continuously differentiable real valued functions (one-sided derivatives at the end points of the interval) with the norm

$$\|u\|_{I,1} = \sup \left\{ |u^{j)}(t)| : t \in I,\ j = 0, 1 \right\}.$$

Under this norm $E = \mathcal{C}^1(I, \mathbb{R})$ is a Banach space. For a given continuously differentiable function $F : I \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, define a function $f : E \to \mathbb{R}$ by

$$f(u) = \int_a^b F(t, u(t), u'(t))dt$$

and show that f is Fréchet differentiable on E. Show in particular

$$f'(u)(v) = \int_a^b [F_{,u}(t, u(t), u'(t)) - \frac{d}{dt} F_{,u'}(t, u(t), u'(t))] v(t) dt \\ + F_{,u'}(t, u(t), u'(t)) v(t)|_a^b.$$

for all $v \in E$. $F_{,u}$ denotes the derivative of F with respect to the second argument and similarly, $F_{,u'}$ denotes the partial derivative with respect to the third argument.

Now consider $M = \{u \in E : u(a) = c,\ u(b) = d\}$ for some given values $c, d \in \mathbb{R}$ and show that the derivative of the restriction of f to M is

$$f'(u)(v) = \int_a^b [F_{,u}(t, u(t), u'(t)) - \frac{d}{dt} F_{,u'}(t, u(t), u'(t))] v(t) dt \tag{3.21}$$

for all $v \in E$, $v(a) = 0 = v(b)$. Deduce the *Euler – Lagrange equation*

$$F_{,u}(t, u(t), u'(t)) - \frac{d}{dt} F_{,u'}(t, u(t), u'(t)) = 0 \tag{3.22}$$

Hints: Use the Taylor expansion with remainder for F.

7. Suppose that E, F are two real Banach spaces. Prove the existence of the natural isomorphism $\mathcal{L}(E, \mathcal{L}(E, F)) \cong \mathcal{B}(E \times E, F)$.

   Hints: For $h \in \mathcal{L}(E, \mathcal{L}(E, F))$ define $\hat{h} \in \mathcal{B}(E \times E, F)$ by $\hat{h}(e_1, e_2) = h(e_1)(e_2)$ for all $e_1.e_2 \in E$ and for $b \in \mathcal{B}(E \times E, F)$ define $\check{b} \in \mathcal{L}(E, \mathcal{L}(E, F))$ by $\check{b}(e_1)(e_2) = b(e_1, e_2)$ and then show that these mappings are inverse to each other. Write down the definition of the norms of the spaces $\mathcal{L}(E, \mathcal{L}(E, F))$ and $\mathcal{B}(E \times E, F)$ explicitly and show that the mappings $h \mapsto \hat{h}$ and $b \mapsto \check{b}$ both have a norm ≤ 1.

8. Prove the existence of the natural isomorphism (3.9) for $n = 2, 3, 4, \ldots$.

9. Prove the chain rule for the Gâteaux derivative.

10. Complete the proof of Part d) of Lemma 3.4.3.

11. Let V be a function $\mathbb{R}^3 \to \mathbb{R}$ of class $\mathcal{C}^1$. Find the Euler – Lagrange equation (3.22) explicitly for the functional $I(u) = \int_a^b (\frac{m}{2}(u'(t))^2 - V(u(t)))dt$ on differentiable functions $u : [a, b] \to \mathbb{R}^3$, $u(a) = x$, $u(b) = y$ for given points $x, y \in \mathbb{R}^3$.

12. Consider the function $g(s) = \frac{|1+s|^p - 1 - ps}{|s|^p}$, $s \in \mathbb{R} \setminus \{0\}$, and show $g(s) \to 1$ as $|s| \to \infty$, $g(s) \to 0$ as $s \to 0$. Conclude

that there are constants $0 < c < C < \infty$ such that $c|s|^p \leq g(s)|s|^p \leq C|s|^p$ for all $s \in \mathbb{R}$.

Chapter 4

Constrained Minimization Problems

(Method of Lagrange Multipliers)

In the calculus of variations we have often to do with the following problem: Given a real valued function f on a nonempty open subset U of a real Banach space E, find the minimum (maximum) of f on all those points x in U which satisfy a certain restriction or constraint. A very important example of such a constraint is that the points have to belong to a level surface of some function g, i.e., have to satisfy $g(x) = c$ where the constant c distinguishes the various level surfaces of the function g. In elementary situations, and typically also in Lagrangian mechanics, one introduces a so-called *Lagrange multiplier* λ as a new variable and proceeds to minimize the function $f(\cdot) + \lambda(g(\cdot) - c)$ on the set U. In simple problems (typically finite dimensional) this strategy is successful. The problem is to prove the existence of a Lagrange multiplier.

As numerous successful applications have shown the following setting is an appropriate framework for such constrained minimization problems:

> Let E, F be two real Banach spaces, $U \subseteq E$ an open nonempty subset, $g : U \to F$ a mapping of class $\mathcal{C}^1$, $f : U \to \mathbb{R}$ a function of class $\mathcal{C}^1$, and y_0 some point in F. The *optimization problem for the function f under the*

constraint $g(x) = y_0$ is the problem of finding extremal points of the function $f_{|M} : M \to \mathbb{R}$ where $M = [g = y_0]$ is the level surface of g through the point y_0.

In this chapter we present a comprehensive solution for the infinite dimensional case, mainly based on ideas of Ljusternik [Lju34]. A first section explains in a simple setting the geometrical interpretation of the existence of a Lagrange multiplier. As an important preparation for the main results the existence of tangent spaces to level surfaces of $\mathcal{C}^1$-functions is shown in substantial generality. Finally the existence of a Lagrange multiplier is proven and some simple applications are discussed.

In the following chapter, after the necessary preparations, we will use the results on the existence of a Lagrange multiplier to solve eigenvalue problems, for linear and nonlinear partial differential operators.

4.1 Geometrical interpretation of constrained minimization

In order to develop some intuition about constrained minimization problems and the rôle of the Lagrange multiplier we consider such a problem first on a space of dimension two and discuss heuristically in geometrical terms how to obtain the solution. Let $U \subset \mathbb{R}^2$ be a nonempty open subset. Our goal is to determine the minimum of a continuous function $f : U \to \mathbb{R}$ under the constraint $g(x) = c$ where the constraint function $g : U \to \mathbb{R}$ is continuous. This means: Find $x_0 \in U$ satisfying $g(x_0) = c$ and $f(x_0) \leq f(x)$ for all $x \in U$ such that $g(x) = c$. In this generality the problem does not have a solution. If however both f and g are continuously differentiable on U, then the level surfaces of both functions have well-defined tangents, and then we expect a solution to exist, because of the following heuristic considerations.

Introduce the level surface

$$[g = c] = \{x \in U : g(x) = c\}$$

and similarly the family of level surfaces $[f = d]$, $d \in \mathbb{R}$, for the function f. If a level surface $[f = d]$ does not intersect the level surface $[g = c]$, then no point on this level surface of f satisfies the constraint and is thus not relevant for our problem. If for a certain value of d the level surfaces $[f = d]$ and $[g = c]$ intersect in exactly one point (at some finite angle), then for all values d' close to d the level surfaces $[g = c]$ and $[f = d']$ also intersect at exactly one point, and thus d is not the minimum of f under the constraint $g(x) = c$. Next consider a value of d for which the level surfaces $[g = c]$ and $[f = d]$ intersect in at least two distinct points (at finite angles). Again for all values d' sufficiently close to d the level surfaces $[f = d']$ and $[g = c]$ intersect in at least two distinct points and therefore d is not the minimum of f under the given constraint. Finally consider a value d_0 for which the level surfaces $[g = c]$ and $[f = d_0]$ 'touch' in exactly one point x_0, i.e., $[g = c] \cap [f = d_0] = \{x_0\}$ and the tangents to both level surfaces at this point coincide. In this situation small changes of the value of d lead to an intersection which is either empty or consists of at least two points, hence these values $d' \neq d_0$ do not produce a minimum under the constraint $g(x) = c$. We conclude that d_0 is the minimum value of f under the given constraint and that x_0 is the minimizing point. The following figure shows in a two dimensional problem three of the cases discussed above. Given the level surface $[g = c]$ of the constraint function g, three different level surfaces of the function f are considered.

Recall that the level surfaces $[g = c]$ and $[f = d]$ are level surfaces of smooth functions over an open set $U \subset \mathbb{R}^2$. Assume (or prove under appropriate assumptions with the help of the implicit function theorem) that in a neighborhood of the point

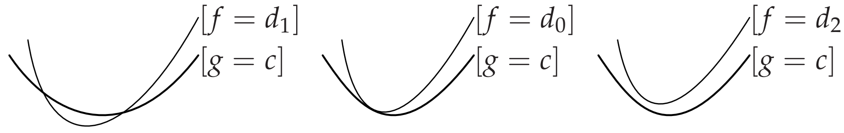


Figure 4.1: Level surface $[g = c]$ and $[f = d_i]$, $i = 0, 1, 2$; $d_1 < d_0 < d_2$; $i = 1$ two points of intersection, $i = 0$ touching level surfaces; $i = 2$ no intersection.

$x_0 = (x_1^0, x_2^0)$ these level surfaces have the explicit representation $x_2 = y(x_1)$, respectively $x_2 = \xi(x_1)$. Under these assumptions it is shown in the Exercises that the tangent to these touching level surfaces coincide if, and only if,

$$(Df)(x_0) = \lambda(Dg)(x_0) \tag{4.1}$$

for some $\lambda \in \mathbb{R} \cong \mathcal{L}(\mathbb{R}, \mathbb{R})$.

4.2 Tangent spaces of level surfaces

In our setting a constraint minimization problem is a problem of analysis on level surfaces of $\mathcal{C}^1$ mappings. It requires that we can do differential calculus on these surfaces which in turn relies on the condition that these level surfaces are differential manifolds. The following approach does not assume this but works under the hypothesis that one has, at the points of interest on these level surfaces, the essential element of a differential manifold, namely a proper tangent space.

Recall that in infinite dimensional Banach spaces E a closed subspace K does not always have a topological complement, i.e., a closed subspace L such that E is the direct sum of these two subspaces (see for instance [RR73]). Thus in our fundamental result on the existence of a proper tangent space this property is assumed but later we will show when and how it holds.

Theorem 4.2.1 (Existence of a tangent space) *Let E, F be real Banach spaces, $U \subseteq E$ a nonempty open subset, and $g : U \to F$ a mapping of class $\mathcal{C}^1$. Suppose that x_0 is a point of the level surface $[g = y_0]$*

of the mapping g. If x_0 is a regular point of g at which the null-space $N(g'(x_0))$ of the derivative of g has a topological complement in E, then the set

$$\begin{aligned} T_{x_0}[g = y_0] &= \{x \in E : \exists u \in N(g'(x_0)),\ x = x_0 + u\} \\ &= x_0 + N(g'(x_0)) \end{aligned} \tag{4.2}$$

is a proper tangent space of the level surface $[g = y_0]$ at the point x_0, i.e., there is a homeomorphism χ of a neighborhood U' of x_0 in $T_{x_0}[g = y_0]$ onto a neighborhood V of x_0 in $[g = y_0]$ with the following properties:

a) $\chi(x_0 + u) = x_0 + u + \varphi(u)$ for all $x_0 + u \in U'$;

b) φ is continuous and of higher than linear order in u, $\varphi(u) = o(h)$.

Proof. Since x_0 is a regular point of g, the derivative $g'(x_0)$ is a surjective continuous linear mapping from E onto F. By assumption the null-space $K = N(g'(x_0))$ of the mapping has a topological complement L in E so that the Banach space E is the direct sum of these two closed subspaces, $E = K + L$. It follows (see [RR73]) that there are continuous linear mappings p and q of E onto K and L, respectively, which have the following properties: $K = \operatorname{ran} p = N(q)$, $L = N(p) = \operatorname{ran} q$, $p^2 = p$, $q^2 = q$, $p + q = id$.

Since U is open there is $r > 0$ such that the open ball B_r in E with center 0 and radius r satisfies $x_0 + B_r + B_r \subset U$. Now define a mapping $\psi : K \cap B_r \times L \cap B_r \to F$ by

$$\psi(u, v) = g(x_0 + u + v) \qquad \forall\, u \in K \cap B_r, \quad \forall\, v \in L \cap B_r. \tag{4.3}$$

By the choice of the radius r this map is well defined. The chain rule implies that it has the following properties: $\psi(0, 0) = g(x_0) = y_0$, ψ is continuously differentiable and

$$\psi_{,u}(0, 0) = g'(x_0)_{|K} = 0 \in \mathcal{L}(K, F), \quad \psi_{,v}(0, 0 = g'(x_0)_{|L} \in \mathcal{L}(L, F).$$

On the complement L of its null-space the surjective mapping $g'(x_0) : E \to F$ is bijective, thus $\psi_{,v}(0, 0)$ is a bijective continuous linear mapping of the Banach space L onto the Banach space F. The inverse mapping theorem (see Appendix 34.5) implies that the inverse $\psi_{,v}(0, 0)^{-1} : F \to L$ is a continuous linear operator too. Thus all hypotheses of the implicit function theorem (see, for example, [Die69]) are satisfied for the problem

$$\psi(u, v) = y_0.$$

This theorem implies that there is $0 < \delta < r$ and a unique function $\varphi : K \cap B_\delta \to L$ which is continuously differentiable such that

$$y_0 = \psi(u, \varphi(u)) \quad \forall\, u \in K \cap B_\delta \quad \text{and} \quad \varphi(0) = 0.$$

Since in general $\varphi'(0) = -\psi_{,v}(0, 0)^{-1}\, \psi_{,u}(0, 0)$ we have here $\varphi'(0) = 0$ and thus $\varphi(u) = o(u)$.

Define a mapping $\chi : x_0 + K \cap B_\delta \to M$ by $\chi(x_0 + u) = x_0 + u + \varphi(u)$. Clearly χ is continuous. By construction, $y_0 = \psi(u, \varphi(u)) = g(x_0 + u + \varphi(u))$, hence χ maps into $M = [g = y_0]$. By construction, u and $\varphi(u)$ belong to complementary subspaces of E, therefore χ is injective and thus invertible on

$$V = \{x_0 + u + \varphi(u) : u \in K \cap B_\delta\} \subset M.$$

Its inverse is $\chi^{-1}(x_0 + u + \varphi(u)) = x_0 + u$. Since $\operatorname{ran} p = K$ and $N(p) = L$ the inverse can be represented as

$$\chi^{-1}(x_0 + u + \varphi(u)) = x_0 + p(u + \varphi(u))$$

and this shows that χ^{-1} is continuous too. Therefore χ is a homeomorphism from $U' = x_0 + K \cap B_\delta$ onto $V \subset M$. This concludes the proof. $\square$

Apart from the natural assumption about the regularity of the point x_0 this theorem uses the technical assumption that the nullspace $K = N(g'(x_0))$ of $g'(x_0) \in \mathcal{L}(E, F)$ has a topological complement in E. We show now that this assumption is quite adequate for the general setting by proving that it is automatically satisfied for three large and frequent classes of special cases.

Proposition 4.2.2 *Let E, F be real Banach spaces and $A : E \to F$ a surjective continuous linear operator. The nullspace $K = N(A)$ has a topological complement in E , in the following three cases:*

a) E is a Hilbert space;

b) F is a finite dimensional Banach space;

c) $N(A)$ is finite dimensional, for instance $A : E \to F$ is a Fredholm operator (i.e., an operator with finite dimensional null-space and closed range of finite codimension).

Proof. If K is a closed subspace of the Hilbert space E, the projection theorem guarantees existence of the topological complement $L = K^{\perp}$ and thus proves Part a).

If F is a finite dimensional Banach space, there exist linearly independent vectors $e_1, \dots, e_m \in E$ such that $\{f_1 = Ae_1, \dots, f_m = Ae_m\}$ is a basis of F. The vectors $e_1, \dots, e_m$ generate a linear subspace V of E of dimension m and it follows that A now is represented by $Ax = \sum_{j=1}^{m} a_j(x) f_j$ with continuous linear functionals $a_j : E \to \mathbb{R}$. Define $px = \sum_{j=1}^{m} a_j(x) e_j$ and $qx = x - px$. One proves easily that $p^2 = p$, $q^2 = q$, $p + q = id$, $V = pE$ and that both maps are continuous. Thus $V = pE$ is the topological complement of $N(A) = qE$. This proves b).

Suppose $\{e_1, \dots, e_m\}$ is a basis of $N(A)$. There are continuous linear functionals a_j on E such that $a_i(e_j) = \delta_{ij}$ for $i, j = 1, \dots, m$. (Use the Hahn–Banach theorem). As above define $px = \sum_{j=1}^{m} a_j(x) e_j$ and $qx = x - px$ for all $x \in E$. Now we conclude as in Part b). (See the Exercises) □

Corollary 4.2.3 *Suppose that E, F are real Banach spaces, $U \subset E$ a nonempty open set and $g : U \to F$ a map of class C^1. In each of the three cases mentioned in Proposition 4.2.2 for $A = g'(x_0)$ the tangent space of the level surface $[g = y_0]$ at every regular point $x_0 \in [g = y_0]$ of g is given by equation (4.2).*

Proof. Proposition 4.2.2 ensures the hypotheses of Theorem 4.2.1. □

4.3 Existence of Lagrange multipliers

The results on the existence of the tangent spaces of level surfaces allow us to translate the heuristic considerations on the existence of a Lagrange multiplier into precise statements. The result which we present now is primarily useful for the explicit calculation of the extremal points once their existence has been established, say as a consequence of the direct methods discussed earlier.

Theorem 4.3.1 (Existence of Lagrange multipliers) *Let E, F be real Banach spaces, $U \subset E$ open and nonempty, $g : U \to F$ and $f : U \to \mathbb{R}$ of class $\mathcal{C}^1$. Suppose that f has a local extremum at the point $x_0 \in U$ subject to the constraint $g(x) = y_0 = g(x_0)$. If x_0 is a regular point of the map g and if the null-space $K = N(g'(x_0))$ of $g'(x_0)$ has a topological complement L in E, then there exists a continuous linear functional $\ell : F \to \mathbb{R}$ such that x_0 is a critical point of the function $F = f - \ell \circ g : U \to \mathbb{R}$, that is*

$$f'(x_0) = \ell \circ g'(x_0). \tag{4.4}$$

Proof. The restriction H of $g'(x_0)$ to the topological complement L of its kernel K is a continuous injective linear map from the Banach space L onto the Banach space F since x_0 is a regular point of g. The inverse mapping theorem (see Appendix) implies that H has an inverse H^{-1} which is a continuous linear operator $F \to L$.

According to Theorem 4.2.1 the level surface $[g = y_0]$ has a proper tangent space at x_0. Thus the points x of this level surface, in a neighborhood V of x_0, are given by $x = x_0 + u + \varphi(u)$, $u \in K \cap B_\delta$ where $\delta > 0$ is chosen as in the proof of Theorem 4.2.1. Suppose that f has a local minimum at x_0 (otherwise consider $-f$). Then there is an $r \in (0, \delta)$ such that $f(x_0) \leq f(x_0 + u + \varphi(u))$ for all $u \in K \cap B_r$, hence by Taylor's theorem

$$0 \leq f'(x_0)(u) + f'(x_0)(\varphi(u)) + o(u + \varphi(u)) \quad \forall\, u \in K \cap B_r.$$

Since we know that $\varphi(u) = o(u)$, this implies $f'(x_0)(u) = 0$ for all $u \in K \cap B_r$. But $u \in K \cap B_r$ is absorbing in K, therefore $f'(x_0)(u) = 0$ for all $u \in K$, i.e.,

$$K = N(g'(x_0)) \subseteq N(f'(x_0)). \tag{4.5}$$

By assumption, E is the direct sum of the closed subspaces K, L, $E = K + L$. Denote the canonical projections onto K and L by p respectively q. If $x_1, x_2 \in E$ satisfy $q(x_1) = q(x_2)$, then $x_1 - x_2 \in K$ and thus equation (4.5) implies $f'(x_0)(x_1) = f'(x_0)(x_2)$. Therefore a continuous linear functional $\hat{f}'(x_0) : L \to \mathbb{R}$ is well defined by $\hat{f}'(x_0)(qx) = f'(x_0)(x)$ for all $x \in E$. This functional is used to define

$$\ell = \hat{f}'(x_0) \circ H^{-1} : F \to \mathbb{R}$$

as a continuous linear functional on the Banach space F which satisfies equation (4.4), since for every $x \in E$

$$\ell \circ g'(x_0)(x) = \ell \circ g'(x_0)(qx) = \ell \circ H(qx) = \hat{f}'(x_0)(qx) = f'(x_0)(x).$$

We conclude that x_0 is a critical point of the function $F = f - \ell \circ g$, by using the chain rule. □

To illustrate some of the strengths of this theorem we consider a simple example. Suppose E is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and A a bounded self-adjoint operator on E. The problem is to minimize the function $f(x) = \langle x, Ax \rangle$ under the constraint $g(x) = \langle x, x \rangle = 1$. Obviously both functions are of class $\mathcal{C}^1$. Their derivatives are given by $f'(x)(u) = 2\langle Ax, u \rangle$, respectively by $g'(x)(u) = 2\langle x, u \rangle$ for all $u \in E$. It follows that all points of the level surface $[g = 1]$ are regular points of g. Corollary 4.2.3 implies that Theorem 4.3.1 can be used to infer the existence of a Lagrange multiplier $\lambda \in \mathbb{R}$ if x_0 is a minimizing point of f under the constraint $g(x) = 1$: $f'(x_0) = \lambda g'(x_0)$ or $Ax_0 = \lambda x_0$, i.e., the Lagrange multiplier λ is an eigenvalue of the operator A and x_0 is the corresponding normalized eigenvector. This simple example suggests a strategy to determine eigenvalues of operators. Later we will explain this powerful strategy in some detail, not only for linear operators.

In the case of finite dimensional Banach spaces we know that the technical assumptions of Theorem 4.3.1 are naturally satisfied. In this theorem assume that $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$. Every continuously linear functional ℓ on $\mathbb{R}^m$ is characterized uniquely by some m-tuple $(\lambda_1, \ldots, \lambda_m)$ of real numbers. Explicitly Theorem 4.3.1 takes now the form

Corollary 4.3.2 *Suppose that $U \subset \mathbb{R}^n$ is open and nonempty, and consider two mappings $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}^m$ of class $\mathcal{C}^1$. Furthermore assume that the function f attains a local extremum at a regular point $x_0 \in U$ of the mapping g (i.e., the Jacobi matrix $g'(x_0)$ has maximal rank m) under the constraint $g(x) = y_0 \in \mathbb{R}^m$. Then there exist real numbers $\lambda_1, \ldots, \lambda_m$ such that*

$$\frac{\partial f}{\partial x_i}(x_0) = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(x_0), \qquad i = 1, \ldots, n. \tag{4.6}$$

Note that equation (4.6) of Corollary 4.3.2 and the equation $g(x_0) = y_0 \in \mathbb{R}^m$ give us exactly $n + m$ equations to determine the $n + m$ unknowns $(\lambda, x_0) \in \mathbb{R}^m \times U$.

Theorem 4.3.1 can also be used to derive necessary and sufficient conditions for extremal points under constraints. For more details we have to refer to chapter 4 of the book [BB92].

4.3.1 Comments on Dido's problem

According to the brief discussion in the introduction to Part C Dido's original problem is a paradigmatic example of constrained minimization. Though intuitively the solution is clear (a circle where the radius is determined by the given length) a rigorous proof is not very simple even with the help of the abstract results which we have developed in this section. Naturally Dido's problem and its solution have been discussed much in the history of the calculus of variations (see [Gol80]). Weierstrass solved this problem in his lectures in 1872 and 1879. There is also an elegant geometrical solution based on symmetry considerations due to Steiner.

In the Exercises we invite the reader to find the solution by two different methods. The first method suggests parametrizing the curve we are looking for by its arc length and using Parseval's relation in the Hilbert space $\mathcal{H} = L^2([0, 2\pi])$. This means that we assume that this curve is given in parametric form by a parametrization $(x(t), y(t)) \in \mathbb{R}^2$, $0 \leq t \leq 2\pi$ where x, y are differentiable functions satisfying $\dot{x}(t)^2 + \dot{y}(t)^2 = 1$ for all $t \in [0, 2\pi]$. With this normalization and parametrization the total length of the curve is $L = \int_0^{2\pi} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt = 2\pi$ and the area enclosed by this curve is

$$A = \int_0^{2\pi} x(t)\dot{y}(t)dt.$$

Proposition 4.3.3 *For all parametrizations of the form described above one has $A \leq \pi$. Furthermore, $A = \pi$ if, and only if, the curve is a* *circle of radius* 1.

Proof. See the Exercises. □

The second approach uses the Lagrange multiplier method as explained above. Suppose that the curve is to have the total length $2L_0$. Choose a parameter a such that $2a < L_0$. In a suitable coordinate system the curve we are looking for is given as $y = u(x)$, $-a \le x \le a$, and $u(x) \ge 0$, $u(\pm a) = 0$ with a function u of class $\mathcal{C}^1$. Its length is $\int_{-a}^{a} \sqrt{1+u'(x)^2}dx = L(u)$ and the area enclosed by the x-axis and this curve is $A(u) = \int_{-a}^{a} u(x)dx$. The problem then is to determine u such that $A(u)$ is maximal under the constraint $L(u) = L_0$.

Proposition 4.3.4 *For the constrained minimization problem for $A(u)$ under the constraint $L(u) = L_0$ there is a Lagrange multiplier λ satisfying $\frac{s}{\sqrt{1+s^2}} = \frac{a}{\lambda}$ for some $s \in \mathbb{R}$ and a solution $u(x) = \lambda[\sqrt{1-(\frac{x}{\lambda})^2} - \sqrt{1-(\frac{a}{\lambda})^2}]$, $-a \le x \le a$. One has $L_0 = 2\lambda\theta(a)$ with $\theta(a) =$* arcsin $\frac{a}{\lambda} \in [0, \frac{\pi}{2}]$. *For this curve the area is*

$$A(u) = \lambda^2\theta(a) - a\sqrt{\lambda^2 - a^2}.$$

Proof. See the Exercises. □

Since $L_0 = 2\lambda\theta(a)$ the Lagrange multiplier λ is a function of a and hence one can consider $A(u)$ as a function of a. Now it is not difficult to determine a so that the enclosed area $A(u)$ is maximal. For $a = \lambda = \frac{L_0}{\pi}$ this area is maximal and is given by $A(u) = a^2\pi/2$. This is the area enclosed by a half-circle of radius $a = \frac{L_0}{\pi}$.

Remark 4.3.5 *There is an interesting variation of Dido's problem which has found important applications in modern probability theory (see [LT91]) and which we mention briefly. Let $A \subset \mathbb{R}^n$ be a bounded domain with a sufficiently smooth boundary and for $t > 0$ consider the set*

$$A_t = \{x \in \mathbb{R}^n \backslash A : \|x - y\| \le t, \ \forall\, y \in A\}.$$

Now minimize the volume $|A_t|$ of the set A_t under the constraint that the volume $|A|$ of A is fixed. The answer is known: This minimum is attained when A is a ball in $\mathbb{R}^n$. This is of particular interest in the case of very high dimensions $n \to \infty$ since then it is known that practically the volume of $A_t \cup A$ is equal to the volume of A_t. For the proof of this result we refer to the book [BZ88] and the article [Oss78].

4.4 Exercises

1. Let $U \subset \mathbb{R}^2$ be open and nonempty. Suppose $f, g \in \mathcal{C}^1(U, \mathbb{R})$ have level surfaces $[g = c]$ and $[f = d]$ which touch in a point $x_0 \in U$ in which the functions f, g have nonvanishing derivatives with respect to the second argument. Prove Equation 4.1.

2. Prove in detail: A finite dimensional subspace V of a Banach space E has a topological complement.

3. Prove Corollary 4.3.2.

4. Prove Proposition 4.3.3.

   Hints: Use the Fourier expansion for x, y:

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kt + b_k \sin kt)$$
$$y(t) = \frac{\alpha_0}{2} + \sum_{k=1}^{\infty}(\alpha_k \cos kt + \beta_k \sin kt).$$

   Calculate $\dot{x}(t), \dot{y}(t)$ and calculate $\int_0^{2\pi}[\dot{x}(t)^2 + \dot{y}(t)^2]dt$ as

$$\langle \dot{x}(t), \dot{x}(t)\rangle_2 + \langle \dot{y}(t), \dot{y}(t)\rangle_2$$

   using $\langle \cos kt, \sin jt\rangle_2 = 0$ and $\langle \cos kt, \cos kt\rangle_2 = \langle \sin kt, \sin kt\rangle_2 = \pi$. Similarly one can calculate $A = \langle x, \dot{y}\rangle_2 = \pi \sum_{k=1}^{\infty} k(a_k\beta_k -$

$b_k\alpha_k)$. This gives

$$2\pi - 2A = \pi \sum_{k=1}^{\infty}(k^2 - k)[a_k^2 + b_k^2 + \alpha_k^2 + \beta_k^2] +$$

$$+\pi \sum_{k=1}^{\infty} k[(a_k - \beta_k)^2 + (\alpha_k + b_k)^2].$$

Now it is straightforward to conclude.

5. Prove Proposition 4.3.4.

   Hints: 1. Calculate the Fréchet derivative of the constraint functional $L(u)$ and show that all points of a level surface $[L = L_0]$ are regular points of the mapping L, for $2a < L_0$. 2. Prove that $|u(x)| \leq L(u)$ for all $x \in [-a, a]$ and hence $A(u) \leq 2aL(u) = 2aL_0$. 3. Prove that $A(u)$ is (upper semi-) continuous for the weak topology on $E = H_0^1(-a, a)$. 4. Conclude that a maximizing element $u \in E$ and a Lagrange multiplier λ exist. 5. Solve the differential equation $A'(u) = \lambda L'(u)$ under the boundary condition $u(-a) = u(a) = 0$. 6. Calculate $L(u)$ for this solution and equate the result to L_0. 7. Calculate the area $A(u)$ for this solution.

Chapter 5

Surjectivity of monotone coercive operators

In the theory of linear operators on Hilbert spaces there is a famous result by Lax-Milgram which says that strictly positive continuous sesquilinear forms are given by continuous surjective linear operators. This results reads

Theorem 5.0.1 (Lemma of Lax-Milgram) *Let a be a continuous sesquilinear form on a (complex) Hilbert space $\mathcal{H}$ which is positive in the sense that for some $c > 0$ one has*

$$a(u,u) \geq c\,\|u\|^2 \quad \textit{for all} \quad u \in \mathcal{H}.$$

Then there is exactly one bounded linear operator A on $\mathcal{H}$ such that

$$a(u,v) = \langle Au, v\rangle \quad \textit{for all} \quad u,v \in \mathcal{H}.$$

The operator A is bijective and has a continuous inverse A^{-1}.

5.1 The result of Browder and Minty

Around 1963 this result has been extended considerably to nonlinear operators by F. E. Browder and independently by G. Minty and is known as the surjectivity of coercive monotone operators in real Banach spaces. We discuss here only the case of separable Banach spaces

Definition 5.1.1 *Let E be a real Banach space and E' its topological dual. A (not necessarily linear) map $T : E \longrightarrow E'$ is called* **monotone** *if, and only if, for all $u, v \in E$ one has*

$$\langle T(u) - T(v), u - v \rangle \geq 0. \tag{5.1}$$

Here $\langle \cdot, \cdot \rangle$ denotes the duality between E' and E.

Definition 5.1.2 *Let E be a real Banach space and E' its topological dual. A (not necessarily linear) map $T : E \longrightarrow E'$ is called* **coercive** *if, and only if, for all $u \in E$ one has*

$$\langle T(u), u \rangle \geq C(\|u\|) \, \|u\| \tag{5.2}$$

for some function $C : [0, \infty) \longrightarrow \mathbb{R}$ with the property $C(s) \longrightarrow \infty$ as $s \longrightarrow \infty$.

With these definitions the surjectivity of continuous coercive monotone maps follows.

Theorem 5.1.3 (Browder-Minty) *Let E be a separable reflexive Banach space with topological dual E' and let $T : E \longrightarrow E'$ be a continuous, coercive, monotone mapping. Then T is surjective, i.e.,*

$$T(E) = E'.$$

For fixed $f \in E'$ the solution set $T^{-1}(f)$ is a bounded, convex, closed subset of E.

Remark: The proof of this important result relies mainly on three facts:

a) Continuous coercive operators on finite dimensional Banach spaces are coercive (this follows from Brouwer's fixed point theorem as we will see);

b) a convenient characterization of the solution set $T^{-1}(f)$ for given $f \in E'$;

c) as a result of monotonicity a generalized Galerkin approximation converges.

Accordingly we prepare the proof of the Theorem of Browder-Minty by the following lemmas.

Lemma 5.1.4 *Let $T : F \longrightarrow F'$ be a continuous coercive mapping on the finite-dimensional Banach space F. Then T is surjective: $T(F) = F'$.*

Lemma 5.1.5 *Let E be a Banach space and $T : E \longrightarrow E'$ be a monotone mapping.*

a) If T is continuous, then for given $f \in E'$, every solution u of $T(u) = f$ is characterized by the inequality

$$\langle T(v) - f, v - u\rangle \geq 0 \quad \textit{for all} \quad v \in E. \tag{5.3}$$

The set of solutions $T^{-1}(f) = \{u \in E : T(u) = f\}$ is closed and convex.

b) If T is coercive, then the solution set $T^{-1}(f)$ is bounded for every $f \in E'$:

$$T^{-1}(f) \subseteq \{u \in E : \|u\| \leq \hat{C}(\|f\|')\},$$

where $\hat{C}(s) = \sup\{r \in [0, \infty) : C(r) \leq s\}$.

5.2 The proofs

5.2.1 The proof of the Browder-Minty theorem

a) In order to show that a given continuous, monotone, and coercive functional $T_0 : E \longrightarrow E'$ is surjective, we proceed as follows: Given $f \in E'$ consider the functional $T(\cdot) = T_0(\cdot) - f$ and show that $0 \in E'$ belongs to the range of T, i.e., there is $u_0 \in E$ such that $T(u) = 0$. Then obviously $T_0(u_0) = f$.

It is straightforward to see that with T_0 also the functional T is continuous, monotone, and coercive. Thus in order to prove the theorem we proceed to show that $0 \in T(E)$.

b) Definition of a suitable Galerkin approximation: Since E is separable, there is an increasing sequence of finite-dimensional subspaces E_n whose union is dense in E. Let $\Phi_n : E_n \longrightarrow E$ denote the identical embedding and let $\Phi'_n : E' \longrightarrow E'_n$ denote adjoint projection. Then for given $T : E \longrightarrow E'$ introduce the sequence $T_n : E_n \longrightarrow E'_n$ of auxiliary mappings by

$$T_n = \Phi'_n \circ T \circ \Phi_n. \tag{5.4}$$

Clearly, all the mappings T_n are continuous. A straightforward calculation shows that they are monotone: For arbitrary $u, v \in E_n$ we have

$$\begin{aligned}\langle T_n(u) - T_n(v), u - v\rangle &= \langle \Phi'_n(T \circ \Phi_n(u) - T \circ \Phi_n(v)), u - v\rangle \\ &= \langle T(\Phi_n(u)) - T(\Phi_n(v)), \Phi_n(u) - \Phi_n(v)\rangle \geq 0.\end{aligned}$$

Coerciveness of T implies uniform coerciveness of the mappings T_n: For all $u \in E_n$ the following estimate holds:

$$\begin{aligned}\langle T_n(u), u\rangle &= \langle T(\Phi_n(u)), \Phi_n(u)\rangle \\ &\geq C(\|\Phi_n(u)\|)\, \|\Phi_n(u)\| = C(\|u\|)\, \|u\|\,.\end{aligned}$$

c) Proof of surjectivity of the auxiliary mappings: For every $n \in \mathbb{N}$ $T_n : E_n \longrightarrow E'_n$ is a continuous coercive mapping on the finite-dimensional Banach space E_n. Hence Lemma 5.1.4 applies and proves that T_n is surjective, i.e., there is $u_n \in E_n$ such that $T_n(u_n) = 0$. Part b) of Lemma 5.1.5 gives the uniform estimate

$$\|u_n\| \leq M \equiv \hat{C}(0).$$

d) Passage to the limit: Since E is a reflexive Banach space the bounded sequence of the solutions u_n has a weakly convergent subsequence which we denote in the same way. Thus there is $u_0 \in E$ such that $w - \lim_{n\to\infty} u_n = u_0$. We claim $T(u_0) = 0$.

To prove this take any $v \in E_m$, $m \in \mathbb{N}$, arbitrary but fixed. By construction, for all $n \geq m$ we know $E_m \subseteq E_n$, and thus, because of monotonicity of T_n on E_n,

$$0 \leq \langle T_n(v) - T_n(u_n), v - u_n \rangle = \langle T_n(v), v - u_n \rangle = \langle T(v), v - u_n \rangle.$$

Taking the limit gives $0 \leq \langle T(v), v - u_0 \rangle$, and since v was arbitrary, we conclude

$$0 \leq \langle T(v), v - u_0 \rangle \quad \text{for all} \quad v \in \cup_{m \in \mathbb{N}} E_m.$$

Since T is continuous and $\cup_{m \in \mathbb{N}} E_m$ is dense in E, we can extend this estimate by continuity to all of E and thus get

$$0 \leq \langle T(v), v - u_0 \rangle \quad \text{for all} \quad v \in E.$$

Now we apply Part a) of Lemma 5.1.5 to conclude $T(u_0) = 0$, i.e., $0 \in T(E)$ and therefore $T(E) = E'$.

5.2.2 Proof of Lemma 5.1.4

a) As discussed above it suffices to show $0 \in T(F)$. Note furthermore that the hypotheses of this lemma remain true if we pass to an equivalent Banach space (the function C of the coercivity hypothesis might change but not its properties). In particular we can assume that F is a finite-dimensional Hilbert space so that F and its dual F' can be identified.

b) For sufficiently large $R > 0$ we know $C(R) > 0$. For such a number R we thus have, for $s(u) = u - T(u)$,

$$\langle s(u), u \rangle < \|u\|^2 \quad \text{for all} \quad u \in F,\ \|u\| = R.$$

The *radial retraction* r of the space F to the closed ball $B_R = \{u \in F : \|u\| \leq r\}$ is defined by

$$r(v) = \begin{cases} v & \text{for v} \in \mathrm{B_R}, \\ \frac{R}{\|v\|} v & \text{for v} \notin \mathrm{B_R}. \end{cases}$$

It follows that $r(B_R^c) \subseteq \partial B_R$ and that $f(u) = r \circ s(u)$ is a continuous mapping from B_R into B_R. By Brouwer's fixed point theorem, f has a fixed point u_0 in B_R.

If $\|u_0\| < R$ then the definition of the retraction map r and the fixed point property $f(u_0) = u_0$ show that $f(u_0) = s(u_0) = u_0$ since then it follows $s(u_0) \in B_R$. The last identity $s(u_0) = u_0$ implies $T(u_0) = 0$ and thus proves the lemma for this case.

The case $\|u_0\| = R$ can be excluded: If

$$\|u_0\| = R$$

then $u_0 = f(u_0)$ implies that $s(u_0) \in B_R^c \cup \partial B_R$ and thus $\rho = \|s(u_0)\| \geq R$. The definition of the retraction then says $f(u_0) = r(s(u_0)) = \frac{R}{\rho} s(u_0)$ and therefore, by choice of R,

$$R^2 = \|u_0\|^2 = \langle f(u_0), u_0 \rangle = \frac{R}{\rho} \langle s(u_0), u_0 \rangle < \frac{R}{\rho} \|u_0\|^2 < R^2,$$

a contradiction.

5.2.3 Proof of Lemma 5.1.5

a) Suppose $T(u) = f$ holds; then by monotonicity of T we get for all $v \in E$

$$\langle T(v) - f, v - u \rangle = \langle T(v) - T(u), v - u \rangle \geq 0,$$

hence condition (5.3) holds. Conversely assume that (5.3) holds. Apply it for $v = u + tz$, $z \in E$ arbitrary. It follows

$$0 \leq \langle T(u + tz) - f, tz \rangle.$$

Divide by $t > 0$ to get $0 \leq \langle T(u + tz) - f, z \rangle$ and take the limit $t \longrightarrow 0$. Continuity of T implies $0 \leq \langle T(u) - f, z \rangle$. Since $z \in E$ is arbitrary, we conclude $T(u) - f = 0$.

By condition (5.3) the solution set has the representation

$$T^{-1}(f) = \cap_{v \in E} \{ u \in E : \langle T(v) - f, v - u \rangle \geq 0 \}.$$

This represents $T^{-1}(f)$ as an intersection of closed half-spaces and therefore $T^{-1}(f)$ is closed and convex.

b) Suppose $u \in T^{-1}(f)$; coerciveness of T implies

$$\|u\| \, C(\|u\|) \leq \langle T(u), u \rangle = \langle f, u \rangle \leq \|f\|' \, \|u\| \, ;$$

hence $C(\|u\|) \leq \|f\|'$ or $\|u\| \leq \hat{C}(\|f\|')$.

5.3 An important variation of the Browder-Minty result

Note that in the proof of the Browder-Minty Theorem monotonicity of T was used only in the last step of the proof, to conclude that $u_0 = w - \lim_{n \longrightarrow \infty} u_n$ implies $T(u_0) = \lim_{n \longrightarrow \infty} T(u_n)$. Accordingly one expects that every other property of T which allows to reach this conclusion also implies surjectivity of coercive continuous maps T. We present such a result which is based on the **Smale condition**.

Definition 5.3.1 *Let E be a Banach space with topological dual E' and T a mapping from E into E'. T is said to satisfy the* **Smale condition** *if, and only if, for every sequence $(u_n)_{n \in \mathbb{N}} \subset E$ the two conditions*

$$u = w - \lim_{n \longrightarrow \infty} u_n \quad \text{and} \quad \lim_{n \longrightarrow \infty} \langle \mathrm{T(u_n) - T(u), u_n - u} \rangle = 0$$

imply

$$u = \lim_{n \longrightarrow \infty} u_n .$$

Remark: If a mapping T satisfies the Smale condition then weak convergence of a sequence implies strong convergence under the hypothesis $\lim_{n \longrightarrow \infty} \langle T(u_n) - T(u), u_n - u \rangle = 0$. Accordingly it is a kind of compactness condition. This additional hypothesis can be viewed as a generalized monotonicity condition.

Theorem 5.3.2 *Let E be a reflexive separable Banach space and $T : E \longrightarrow E'$ a continuous coercive mapping which is bounded (i.e., T maps bounded sets into bounded sets) and which satisfies the Smale condition. Then T is surjective: $T(E) = E'$*

Proof. Similar to the proof of the previous result. For details see [BB92].

Chapter 6

Sobolev spaces

6.1 Motivation

As we explained in the Introduction all major developments in the calculus of variations were driven by concrete problems, mainly in Physics. In these applications the underlying Banach space is a suitable function space, depending on the context as we are going to see explicitly later. Major parts of the existence theory of solutions of nonlinear partial differential use variational methods (some are treated in Chapter 7). Here the function spaces which are used are often the so-called **Sobolev spaces** and the successful application of variational methods rests on various types of **embeddings** for these spaces. Accordingly we present here the classical aspects of the theory of Sobolev spaces as they are used in later applications. We assume that the reader is familiar with the basics aspects of the theory of Lebesgue spaces.

6.2 Basic definitions

Let $\Omega \subset \mathbb{R}^n$ be a nonempty open subset, and for $k = 0, 1, 2, \ldots$ and $1 \le p \le \infty$ introduce the vector space

$$\mathcal{C}^{k,p}(\Omega) = \left\{ u \in \mathcal{C}^k(\Omega) : D^\alpha u \in L^p(\Omega), |\alpha| \le k \right\}.$$

Here $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an n-tuple of number $\alpha_i = 0, 1, 2, \ldots$ and $|\alpha| = \sum_{i=1} \alpha_i$, and $D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}}$. On this vector space define a norm for $1 \leq p < \infty$ by

$$\|u\|_{k,p} = \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_p^p \right)^{1/p} . \tag{6.1}$$

and for $p = \infty$ by

$$\|f\|_{k,\infty} = \sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^\infty(\Omega)} . \tag{6.2}$$

The *Sobolev space* $W^{k,p}(\Omega)$ is by definition the completion of $\mathcal{C}^{k,p}(\Omega)$ with respect to this norm. These Banach spaces are naturally embedded into each other according to

$$W^{k,p}(\Omega) \subset W^{k-1,p}(\Omega) \cdots \subset W^{0,p}(\Omega) = L^p(\Omega).$$

Since the Lebesgue spaces $L^p(\Omega)$ are separable for $1 \leq p < \infty$ one can show that these Sobolev spaces are separable too. For $1 < p < \infty$ the spaces $L^p(\Omega)$ are reflexive, and it follows that for $1 < p < \infty$ the Sobolev spaces $W^{k,p}(\Omega)$ are separable reflexive Banach spaces.

There is another equivalent definition of the Sobolev spaces in term of weak (or distributional) derivatives due to Meyers and Serrin (1964):

$$W^{k,p}(\Omega) = \{ f \in L^p(\Omega) : D^\alpha f \in L^p(\Omega) \text{ (weakly) for all } |\alpha| \leq k \} . \tag{6.3}$$

Here $D^\alpha f$ stands for the weak derivative of f, i.e. for all $\phi \in \mathcal{C}_c^\infty(\Omega)$ one has in the sense of Schwartz distributions on Ω

$$\langle D^\alpha f, \phi \rangle = (-1)^{|\alpha|} \int f(x) D^\alpha \phi(x) dx .$$

Theorem 6.2.1 *Equipped with (6.1) respectively (6.2) the set $W^{k,p}(\Omega)$ is a Banach space. In the case $p = 2$ the space $W^{k,2}(\Omega) = H^k(\Omega)$ is actually a Hilbert space with the inner product*

$$\langle f, g\rangle_{H^k(\Omega)} = \sum_{|\alpha|\leq k} \int_\Omega \overline{D^\alpha f} \cdot D^\alpha g dx. \tag{6.4}$$

The spaces $W^{k,p}(\Omega)$ are called **Sobolev spaces of order (k,p)**.

Proof. Since the space $L^p(\Omega)$ is a vector space, the set $W^{k,p}(\Omega)$ is a vector space too, as a subspace of $L^p(\Omega)$. The norm properties of $\|\cdot\|_{L^p(\Omega)}$ easily imply that $\|\cdot\|_{W^{k,p}(\Omega)}$ is also a norm.

The **local** Sobolev spaces $W^{k,p}_{\mathrm{loc}}(\Omega)$ are obtained when in the above construction the Lebesgue space $L^p(\Omega)$ is replaced by the local Lebesgue space $L^p_{\mathrm{loc}}(\Omega)$. Elements in a Sobolev space can be approximated by smooth functions, i.e., these spaces allow **mollification**. In details one has the following result.

Theorem 6.2.2 *Let Ω be an open subset of $\mathbb{R}^n$, $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $1 \leq p < \infty$. Then the following holds:*

a) For $u \in W^{k,p}_{\mathrm{loc}}(\Omega)$ there exists a sequence $u_m \in \mathcal{C}^\infty_c(\Omega)$ of $\mathcal{C}^\infty$ functions on Ω which have a compact support such that $u_m \to u$ in $W^{k,p}_{\mathrm{loc}}(\Omega)$.

b) $\mathcal{C}^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.

c) $\mathcal{C}^\infty_c(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$.

Proof. Here we have to refer to the literature, for instance [Ada75] or the PDE lecture notes by B. Driver (see his home page at www.math.ucsd.edu).

Naturally, the space $\mathcal{C}^\infty_c(\Omega)$ is contained in $W^{k,p}(\Omega)$ for all $k = 0, 1, 2, \ldots$ and all $1 \leq p < \infty$. The closure of this space in $W^{k,p}(\Omega)$ is denoted $W^{k,p}_0(\Omega)$. In general $W^{k,p}_0(\Omega)$ is a proper subspace of $W^{k,p}(\Omega)$. For $\Omega = \mathbb{R}^n$ however equality holds.

The fact that $W^{k,p}_0(\Omega)$ is, in general, a proper subspace of $W^{k,p}(\Omega)$ plays a decisive role in the formulation of boundary value problems. Roughly one can say the following: If the boundary $\Gamma =$

$\partial\Omega$ is sufficiently smooth, then elements $u \in W^{k,p}(\Omega)$ together with their normal derivatives of order $\leq k-1$ can be restricted to Γ. And elements in $W_0^{k,p}(\Omega)$ can then be characterized by the fact that this restriction vanishes. (There is a fairly technical theory involved here). A concrete example of a result of this type is the following theorem.

Theorem 6.2.3 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open subset whose boundary $\Gamma = \partial\Omega$ is piecewise $\mathcal{C}^1$. Then the following holds:*

(a) every $u \in H^1(\Omega)$ has a restriction $\gamma_0 u = u|\Gamma$ to the boundary;

(b) $H_0^1(\Omega) = \ker\gamma_0 = \{u \in H^1(\Omega) : \gamma_0(u) = 0\}$.

Obviously, the Sobolev space $W^{k,p}(\Omega)$ embeds naturally into the Lebesgue space $L^p(\Omega)$. Depending on the value of the exponent p in relation to the dimension n of the underlying space $\mathbb{R}^n$ it embeds also into various other functions spaces, expressing various degrees of smoothness of elements in $W^{k,p}(\Omega)$. The following few sections present a number of (classical) estimates for elements in $W^{k,p}(\Omega)$ which then allow to prove the main **Sobolev embeddings**.

A simple example indicates what can be expected. Take $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\psi(x) = 1$ for all $|x| \leq 1$ and define $f(x) = |x|^q\psi(x)$ for $x \in \mathbb{R}^n$, for some $q \in \mathbb{R}$. Then $\nabla f \in L^p(\mathbb{R}^n)$ requires $n + (q-1)p \geq 0$, or

$$q \geq 1 - \frac{n}{p}.$$

Therefore, if $1 \leq p < n$ then $q < 0$ is allowed and thus f can have a singularity (at $x = 0$). If however $p \geq n$, then only exponents $q \geq 0$ are allowed, and then f is continuous at $x = 0$. The following estimates give a much more accurate picture. These estimates imply first that we get continuous embeddings and at a later stage we will show that for exponents $1 \leq p < n$ these embeddings are actually compact.

We start with the case $n < p \leq +\infty$.

6.3 Morrey's inequality

Denote the unit sphere in $\mathbb{R}^n$ by S and introduce for a measurable set $\Gamma \subset S$ with $\sigma(S) > 0$ ($\sigma(S)$ denotes the surface measure of S) the sets

$$\Gamma_{x,r} = \{x + t\omega : \omega \in \Gamma, 0 \leq t \leq r\}, \quad x \in \mathbb{R}^n, \quad r > 0.$$

Note that for measurable functions f one has

$$\int_{\Gamma_{x,r}} f(y)dy = \int_0^r dt t^{n-1} \int_\Gamma f(x + t\omega)d\sigma(\omega). \tag{6.5}$$

Choosing $f = 1$ we find for the Lebesgue measure of $\Gamma_{x,r}$:

$$|\Gamma_{x,r}| = r^n \sigma(\Gamma)/n. \tag{6.6}$$

Lemma 6.3.1 *If S, x, r are as above and $u \in C^1(\overline{\Gamma_{x,r}})$ then*

$$\int_{\Gamma_{x,r}} |u(y) - u(x)|dy \leq \frac{r^n}{n} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x - y|^{n-1}} dy. \tag{6.7}$$

Proof. For $y = x + t\omega$, $0 \leq t \leq r$ and $\omega \in \Gamma$ one has

$$u(x + t\omega) - u(x) = \int_0^t \omega \cdot \nabla u(x + s\omega)ds,$$

thus integration over Γ yields

$$\begin{aligned}
\int_\Gamma |u(x + t\omega) - u(x)|d\sigma(\omega) &\leq \int_0^t \int_\Gamma |\nabla u(x + s\omega)|d\sigma(\omega)ds \\
&= \int_0^t s^{n-1} \int_\Gamma \frac{|\nabla u(x + s\omega)|}{|x + s\omega - x|^{n-1}} d\sigma(\omega)ds \\
&= \int_{\Gamma_{x,t}} \frac{|\nabla u(y)|}{|y - x|^{n-1}} dy \leq \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|y - x|^{n-1}} dy
\end{aligned}$$

If we multiply this inequality with t^{n-1} and integrate from 0 to r and observe Equation (6.5) we get (6.7).

Corollary 6.3.2 *For any $n < p \leq +\infty$, any $0 < r < \infty$, any $x \in \mathbb{R}^n$ and any Borel measurable subset $\Gamma \subset S$ such that $\sigma(\Gamma) > 0$, one has, for all $u \in C^1(\overline{\Gamma}_{x,r})$*

$$|u(x)| \leq C(\sigma(\Gamma), r, n, p) \|u\|_{W^{1,p}(\Gamma_{x,r})} \tag{6.8}$$

with

$$C(\sigma(\Gamma),r,n,p) = \frac{r^{1-n/p}}{\sigma(\Gamma)^{1/p}} \max\left\{\frac{n^{-1/p}}{r}, \left(\frac{p-1}{p-n}\right)^{1-1/p}\right\}.$$

Proof. Clearly, $|u(x)| \leq |u(y)| + |u(x)-u(y)|$, for any $y \in \Gamma_{x,r}$; integration over $\Gamma_{x,r}$ and application of (6.7) gives

$$|\Gamma_{x,r}||u(x)| = \int_{\Gamma_{x,r}} |u(x)|dy \leq \int_{\Gamma_{x,r}} |u(y)|dy + \frac{r^n}{n}\int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{n-1}}dy.$$

Now apply Hölder's inequality to get

$$\leq \|u\|_{L^p(\Gamma_{x,r})}\|1\|_{L^q(\Gamma_{x,r})} + \frac{r^n}{n}\|\nabla u\|_{L^p(\Gamma_{x,r})}\left\|\frac{1}{|x-\cdot|^{n-1}}\right\|_{L^p(\Gamma_{x,r})} \tag{6.9}$$

where q is the Hölder conjugate exponent of p, i.e., $q = \frac{p}{p-1}$. Calculate

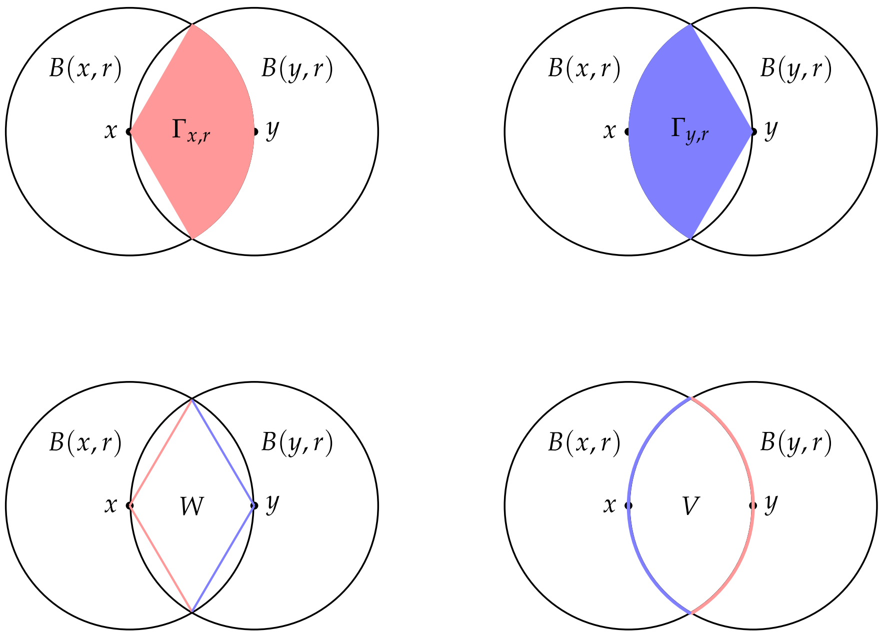
$$\left\|\frac{1}{|\cdot|^{n-1}}\right\|_{L^q(\Gamma_{0,r})} = r^{1-n/p}\left(\sigma(\Gamma)\frac{p-1}{p-n}\right)^{\frac{p-1}{p}} \tag{6.10}$$

and insert the result into (6.9). A rearrangement and a simple estimate finally gives (6.8).

Corollary 6.3.3 *Consider $n \in \mathbb{N}$ and $p \in (n,+\infty]$. There are constants $A = A_n$ and $B = B_n$ such that for any $u \in C^1(\mathbb{R}^n)$ and any $x,y \in \mathbb{R}^n$ one has ($r = |x-y|$, $B(x,r)$ is the ball with center x and radius r)*

$$|u(y)-u(x)| \leq 2BA^{1/p}\left(\frac{p-1}{p-n}\right)^{\frac{p-1}{p}} \|\nabla u\|_{L^p(B(x,r)\cap B(y,r))}|x-y|^{1-\frac{n}{p}}. \tag{6.11}$$

Proof. Certainly, the intersection $V = B(x,r)\cap B(y,r)$ of the two balls is not empty. Introduce the following subsets of the unit sphere in $\mathbb{R}^n$: $\Gamma = \frac{1}{r}(\partial B(x,r)) \cap B(y,r)$ and $\Lambda = \frac{1}{r}(\partial B(y,r))\cap B(x,r) = -\Gamma$ so that we have $x + r\Gamma = (\partial B(x,r))\cap B(y,r)$ and $y + r\Lambda = (\partial B(y,r))\cap B(x,r)$. It is instructive to draw a picture of the sets introduced above and of the related sets $\Gamma_{x,r}$ and $\Lambda_{y,r}$.



$W = \Gamma_{x,r} \cap \Gamma_{y,r}$ $\qquad$ $r = |x-y|$ $\qquad$ $V = B(x,r) \cap B(y,r)$

Since $\Gamma_{x,r} = r\Gamma_{x,1}$ and $\Lambda_{y,r} = r\Lambda_{y,1}$ we find that

$$B_n = \frac{|\Gamma_{x,r} \cap \Lambda_{y,r}|}{|\Gamma_{x,r}|} = \frac{|\Gamma_{x,1} \cap \Lambda_{y,1}|}{|\Gamma_{x,1}|}$$

is a number between 0 and 1 which only depends on the dimension n. It follows $|\Gamma_{x,r}| = |\Lambda_{y,r}| = B_n|W|$, $W = \Gamma_{x,r} \cap \Lambda_{y,r}$.

Now we estimate, using Lemma 6.3.1 and Hölder's inequality

$$\begin{aligned}
|u(x)-u(y)||W| &\le \int_W |u(x)-u(z)|dz + \int_W |u(z)-u(y)|dz \le \int_{\Gamma_{x,r}} |u(x)-u(z)|dz + \int_{\Lambda_{y,r}} |u(z)-u(y)|dz \\
&\le \frac{r^n}{n}\int_{\Gamma_{x,r}} \frac{|\nabla u(z)|}{|x-y|^{n-1}}dz + \frac{r^n}{n}\int_{\Lambda_{y,r}} \frac{|\nabla u(z)|}{|z-y|^{n-1}}dz \\
&\le \frac{r^n}{n}\left(\|\nabla u\|_{L^p(\Gamma_{x,r})} \left\| \frac{1}{|x-\cdot|^{n-1}} \right\|_{L^q(\Gamma_{x,r})} + \|\nabla u\|_{L^p(\Lambda_{y,r})} \left\| \frac{1}{|y-\cdot|^{n-1}} \right\|_{L^q(\Lambda_{y,r})} \right) \\
&\le 2\frac{r^n}{n} \|\nabla u\|_{L^p(V)} \left\| \frac{1}{|\cdot|^{n-1}} \right\|_{L^q(\Gamma_{0,r})}.
\end{aligned}$$

Taking (6.10) and (6.6) into account and recalling $r = |x-y|$, estimate (6.11) follows with $A = \sigma(\Gamma)^{-1}$.

Theorem 6.3.4 (Morrey's inequality) *Suppose $n < p \le +\infty$ and $u \in W^{1,p}(\mathbb{R}^n)$. Then there is a unique version u^* of u (i.e., $u^* = u$ almost everywhere) which is Hölder[1] continuous of exponent $1 - \frac{n}{p}$, i.e., $u^* \in C^{0,1-\frac{n}{p}}(\mathbb{R}^n)$ and satisfies*

$$\|u^*\|_{C^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \le C\,\|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{6.12}$$

where $C = C(n,p)$ is a universal constant. In addition the estimates in (6.7), (6.8) and (6.11) hold when u is replaced by u^*.

Proof. At first consider the case $n < p < \infty$. For $u \in \mathcal{C}_c^1(\mathbb{R}^n)$ Corollaries 6.3.2 and 6.3.3 imply ($BC(\mathbb{R}^n)$ denotes the spaces of bounded continuous functions on $\mathbb{R}^n$)

$$\|u\|_{BC(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(\mathbb{R}^n)} \quad \text{and} \quad \frac{|u(y)-u(x)|}{|y-x|^{1-\frac{n}{p}}} \leq C\,\|\nabla u\|_{L^p(\mathbb{R}^n)}\,.$$

This implies

$$\|u\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(\mathbb{R}^n)}\,. \tag{6.13}$$

If $u \in W^{1,p}(\mathbb{R}^n)$ is given, there is a sequence of functions $u_j \in \mathcal{C}_c^1(\mathbb{R}^n)$ such that $u_j \to u$ in $W^{1,p}(\mathbb{R}^n)$. Estimate (6.13) implies that this sequence is also a Cauchy sequence in $\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)$ and thus converges to a unique element u^* in this space. Clearly Estimate (6.12) holds for this limit element u^* and $u^* = u$ almost everywhere.

The case $p = \infty$ and $u \in W^{1,p}(\mathbb{R}^n)$ can be proven via a similar approximation argument.

Corollary 6.3.5 *[Morrey's inequality] Let Ω be an open bounded subset of $\mathbb{R}^n$ with smooth boundary ($\mathcal{C}^1$) and $n < p \leq \infty$. Then for every $u \in W^{1,p}(\Omega)$ there exists a unique version u^* in $\mathcal{C}^{0,1-\frac{n}{p}}(\Omega)$ satisfying*

$$\|u^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\Omega)} \leq C\,\|u\|_{W^{1,p}(\Omega)}\,. \tag{6.14}$$

with a universal constant $C = C(n, p, \Omega)$.

Proof. Under the assumptions of the corollary there exists a continuous extension operator $J : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$. Then, given $u \in W^{1,p}(\Omega)$, Theorem 6.3.4 implies that there is a continuous version $U^* \in \mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)$ of Ju which satisfies (6.12). Now define $u^* = U^*|_\Omega$. It follows

$$\|u^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\Omega)} \leq \|U^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \leq C\,\|Ju\|_{W^{1,p}(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(\Omega)}\,.$$

6.4 Gagliardo-Nirenberg-Sobolev inequality

This very important inequality is of the form

$$\|u\|_{L^q} \leq C\,\|\nabla u\|_{L^p}\,, \quad u \in \mathcal{C}_c^1(\mathbb{R}^n) \tag{6.15}$$

for a suitable exponent q for a given exponent p, $1 \leq p \leq n$. This exponent is easily determined through the scale covariance of the quantities in this inequality. For $\lambda > 0$ introduce u_λ by setting $u_\lambda(x) = u(\lambda x)$. A simple calculation shows $\|u_\lambda\|_{L^q} = \lambda^{-n/q}\,\|u\|_{L^q}$ and $\|\nabla u_\lambda\|_{L^p} = \lambda^{1-n/p}\,\|\nabla u\|_{L^p}$. Thus inserting u_λ into (6.15 gives

$$\lambda^{-n/q}\,\|u\|_{L^q} \leq C\lambda^{1-n/p}\,\|\nabla u\|_{L^p}$$

for all $\lambda > 0$. This is possible for all $u \in \mathcal{C}_c^1(\mathbb{R}^n)$ only if

$$1 - n/p + n/q = 0, \quad \text{i.e.,} \quad \frac{1}{\mathrm{p}} = \frac{1}{\mathrm{n}} + \frac{1}{\mathrm{q}}. \tag{6.16}$$

It is a standard notation to denote the exponent q which solves (6.16) by p^*, i.e.,

$$p^* = \frac{np}{n-p}$$

with the understanding that $p^* = \infty$ if $p = n$.

Since the case $1 < p < n$ can easily be reduced to the case $p = 1$, we have to prove this inequality for $p = 1$, i.e., $p^* = 1^* = \frac{n}{n-1}$.

Theorem 6.4.1 *For all $u \in W^{1,1}(\mathbb{R}^n)$ one has*

$$\|u\|_{1^*} = \|u\|_{\frac{n}{n-1}} \le \prod_{i=1}^n \left(\int_{\mathbb{R}^n} |\partial_i u(x)| dx \right)^{\frac{1}{n}} \le n^{-\frac{1}{2}} \|\nabla u\|_1 \tag{6.17}$$

Proof. According to Theorem 6.2.2 every element $u \in W^{1,1}(\mathbb{R}^n)$ is the limit of a sequence of elements $u_j \in \mathcal{C}_c^1(\mathbb{R}^n)$. Hence is suffices to prove this inequality for $u \in \mathcal{C}_c^1(\mathbb{R}^n)$, and this is done by induction on the dimension n.

We suggest that the reader proves the GNS inequality for $n = 1$ and $n = 2$. Here we present first the case $n = 3$ before we come to the general case.

Suppose that $u \in \mathcal{C}_c^1(\mathbb{R}^3)$ is given. Observe that now $1^* = 3/2$. Introduce the notation $x^1 = (y_1, x_2, x_3)$, $x^2 = (x_1, y_2, x_3)$, and $x^3 = (x_1, x_2, y_3)$. The fundamental theorem of calculus implies for $i = 1, 2, 3$

$$|u(x)| \le \int_{-\infty}^{x_i} |\partial_i u(x^i)| dy_i \le \int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i,$$

hence multiplication of these three inequalities gives

$$|u(x)|^{\frac{3}{2}} \le \prod_{i=1}^3 \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{2}}.$$

Now integrate this inequality with respect to x_1 and note that the first factor on the right does not depend on x_1:

$$\int_{\mathbb{R}} |u(x)|^{\frac{3}{2}} dx_1 \le \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| dy_1 \right)^{\frac{1}{2}} \int_{\mathbb{R}} \prod_{i=2}^3 \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{2}} dx_1$$

Apply Hölder's inequality (for $p = q = 2$) to the second integral, this gives the estimate

$$\le \left(\int_{-\infty}^{\infty} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{2}} \prod_{i=2}^3 \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{2}}.$$

Next we integrate this inequality with respect to x_2 and apply again Hölder's inequality to get

$$\begin{aligned} \int_{\mathbb{R}^2} |u(x)|^{\frac{3}{2}} dx_1 dx_2 &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x) dx_1 dx_2 \right)^{\frac{1}{2}} \int_{\mathbb{R}} \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| dx_1 \right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} |\partial_3 u(x^3)| dx_1 dy_3 \right)^{\frac{1}{2}} dx_2 \\ &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x) dx_1 dx_2 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x^1)| dx_1 dx_2 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\partial_3 u(x)| dx_1 dx_2 dx_3 \right)^{\frac{1}{2}} \end{aligned}$$

A final integration with respect to x_3 and applying Hölder's inequality as above implies

$$\int_{\mathbb{R}^3} |u(x)|^{\frac{3}{2}} dx_1dx_2dx_3 \le \prod_{i=1}^{3} \left(\int_{\mathbb{R}^3} |\partial_i u(x)| dx_1dx_2dx_3 \right)^{\frac{1}{2}} \le \left(\int_{\mathbb{R}^3} |\nabla u(x)| dx_1dx_2dx_3 \right)^{\frac{1}{2}}$$

which is the GNS inequality for $n = 3$.

The general case uses the same strategy. Naturally some more steps are necessary. Now we have $1^* = \frac{n}{n-1}$. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ introduce the variables $x^i = (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)$. The fundamental theorem of calculus implies for $i = 1, \dots, n$

$$|u(x)| \le \int_{\mathbb{R}} |\partial_i u(x^i)| dy_i$$

and thus

$$|u(x)|^{\frac{n}{n-1}} \le \prod_{i=1}^{n} \left(\int_{\mathbb{R}} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{n-1}}. \tag{6.18}$$

Recall Hölder's inequality for the product of $n-1$ functions in the form

$$\left\| \prod_{i=2}^{n} f_i \right\|_1 \le \prod_{i=2}^{n} \|f_i\|_{n-1} \tag{6.19}$$

and integrate (6.18) with respect to x_1 to get

$$\begin{aligned}
\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 &\le \left(\int_{\mathbb{R}} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{n-1}} \int_{\mathbb{R}} \prod_{i=2}^{n} \left(\int_{\mathbb{R}} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&\le \left(\int_{\mathbb{R}} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{n-1}} \prod_{i=2}^{n} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{n-1}} \\
&\le \left(\int_{\mathbb{R}} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{n-1}} \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2 \right)^{\frac{1}{n-1}} \prod_{i=3}^{n} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{n-1}}
\end{aligned}$$

Now integrate this inequality with respect to x_2 and apply (6.19) again. This implies

$$\begin{aligned}
\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1dx_2 &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1dx_2 \right)^{\frac{1}{n-1}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{n-1}} \prod_{i=3}^{n} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{n-1}} dx_2 \\
&\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1dx_2 \right)^{\frac{1}{n-1}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x)| dx_1dx_2 \right)^{\frac{1}{n-1}} \prod_{i=3}^{n} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| dx_1dx_2 dy_i \right)^{\frac{1}{n-1}}.
\end{aligned}$$

Obviously one can repeat these steps successively for $x_3, \dots, x_n$ and one proves by induction that for $k \in \{1, \dots, n\}$ we get the estimate

$$\begin{aligned}
\int_{\mathbb{R}^k} |u(x)|^{\frac{n}{n-1}} dx_1dx_2 \cdots dx_k \le& \prod_{i=1}^{k} \left(\int_{\mathbb{R}^k} |\partial_i u(x)| dx_1dx_2 \cdots dx_k \right)^{\frac{n}{n-1}} \\
&\times \prod_{i=k+1}^{n} \left(\int_{\mathbb{R}^{k+1}} |\partial_i u(x)| dx_1dx_2 \cdots dx_k dy_i \right)^{\frac{n}{n-1}}
\end{aligned}$$

where naturally for $k = n$ the second product does not occur. Thus for $k = n$ one has

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}} dx_1dx_2 \cdots dx_n &\le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\partial_i u(x)| dx_1dx_2 \cdots dx_n \right)^{\frac{n}{n-1}} \\
&\le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\nabla u(x)| dx \right)^{\frac{n}{n-1}}.
\end{aligned} \tag{6.20}$$

In order to improve this estimate recall Young's inequality in the elementary form $\prod_{i=1}^{n} A_i \le \frac{1}{n} \sum_{i=1}^{n} A_i^n$, where $A_i \ge 0$. Thus we get

$$\|u\|_{\frac{n}{n-1}} \le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\partial_i u(x)| dx \right)^{\frac{1}{n}} \le \frac{1}{n} \sum_{i=1}^{n} \int_{\mathbb{R}^n} |\partial_i u(x)| dx$$

and by Hölder's inequality one knows $\sum_{i=1}^{n} |\partial_i u(x)| \le \sqrt{n}\, |\nabla u(x)|$, hence $\|u\|_{\frac{n}{n-1}} \le \frac{1}{\sqrt{n}} \|\nabla u\|_1$.

Remark 6.4.2 *The starting point of our estimates was the identity $u(x) = \int_{-\infty}^{x_i} \partial_i u(x^i) dy_i$ and the resulting estimate*

$$|u(x)| \leq \int_{\mathbb{R}} |\partial_i u(x^i)| dy_i, \quad i = 1, \ldots, n.$$

If we write

$$u(x) = \frac{1}{2}\left(\int_{-\infty}^{x_i} \partial_i u(x^i) dy_i - \int_{x_i}^{\infty} \partial_i u(x^i) dy_i\right),$$

we can improve this estimate to

$$|u(x)| \leq \frac{1}{2}\int_{\mathbb{R}} |\partial_i u(x^i)| dy_i, \quad i = 1, \ldots, n.$$

Next we look at the case $1 < p < n$. As we will see it can easily be reduced to the case $p = 1$.

Theorem 6.4.3 *If $1 \leq p < n$ then, for all $u \in W^{1,p}(\mathbb{R}^n)$, with $p^* = \frac{np}{n-p}$,*

$$\|u\|_{p^*} \leq \frac{1}{\sqrt{n}} \frac{p(n-1)}{n-p} \|\nabla u\|_p . \tag{6.21}$$

Proof. Since elements in $W^{1,p}(\mathbb{R}^n)$ can be approximated by elements in $\mathcal{C}_c^1(\mathbb{R}^n)$ it suffices to prove Estimate (6.21) for $u \in \mathcal{C}_c^1(\mathbb{R}^n)$. For such a function u consider the function $v = |u|^s \in \mathcal{C}_c^1(\mathbb{R}^n)$ for an exponent $s > 1$ to be determined later. We have $\nabla v = s|u|^{s-1}\mathrm{sgn(u)}\nabla \mathrm{u}$ and thus by applying (6.17) to v we get

$$\| |u|^s \|_{1^*} \leq \frac{1}{\sqrt{n}} \|\nabla |u|^s\|_1 = \frac{s}{\sqrt{n}} \left\| |u|^{s-1} \nabla u \right\|_1 \leq \frac{s}{\sqrt{n}} \left\| |u|^{s-1} \right\|_q \|\nabla u\|_p \tag{6.22}$$

where q is the Hölder conjugate exponent of p. Note that this estimate can be written as

$$\|u\|_{s1^*}^s \leq \frac{s}{\sqrt{n}} \|u\|_{(s-1)q}^{s-1} \|\nabla u\|_p .$$

Now choose s such that $s1^* = (s-1)q$. This gives $s = \frac{q}{q-1^*} = \frac{p^*}{1^*}$ and accordingly the last estimate can be written as

$$\|u\|_{p^*}^s \leq \frac{s}{\sqrt{n}} \|u\|_{p^*}^{s-1} \|\nabla u\|_p .$$

Inserting the value $s = \frac{p(n-1}{n-p}$ of s now yields (6.21).

Corollary 6.4.4 *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set with $\mathcal{C}^1$-boundary. Then for all $p \in [1, n)$ and $1 \leq q \leq p^*$ there is a constant $C = C(\Omega, p, q)$ such that for all $u \in W^{1,p}(\Omega)$*

$$\|u\|_q \leq C \|u\|_{1,p} .$$

Proof. Under the given conditions on Ω one can show that every $u \in W^{1,p}(\Omega)$ has an extension to $Ju \in W^{1,p}(\mathbb{R}^n)$ (i.e., $Ju|\Omega = u$ and $J : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ is continuous). Then for $u \in \mathcal{C}^1(\bar{\Omega}) \cap W^{1,p}(\Omega)$

$$\|u\|_{L^{p^*}(\Omega)} \leq C\,\|Ju\|_{L^{p^*}(\mathbb{R}^n)} \leq C\,\|\nabla(Ju)\|_{L^p(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(\Omega)}\,. \tag{6.23}$$

Since $\mathcal{C}^1(\Omega)$ is dense in $W^{1,p}(\Omega)$, this estimate holds for all $u \in W^{1,p}(\Omega)$. If now $1 \leq q < p^*$ a simple application of Hölder's inequality gives

$$\|u\|_{L^q(\Omega)} \leq \|u\|_{L^{p^*}(\Omega)}\,\|1\|_{L^s(\Omega)} = \|u\|_{L^{p^*}(\Omega)}\,|\Omega|^{1/s} \leq C|\Omega|^{1/s}\,\|u\|_{W^{1,p}(\Omega)}$$

where $\frac{1}{s} + \frac{1}{p^*} = \frac{1}{q}$.

6.4.1 Continuous Embeddings of Sobolev spaces

In this short review of the classical theory of Sobolev spaces we can only discuss the main embeddings results. In the literature one finds many additional cases.

For convenience of notation let us introduce, for a given number $r \geq 0$,

$$r_+ = \begin{cases} r & \text{if r} \notin \mathbb{N}_0 \\ r+\delta & \text{if r} \in \mathbb{N}_0 \end{cases}$$

where $\delta > 0$ is some arbitrary small number.

Lemma 6.4.5 *For $i \in \mathbb{N}$ and $p \geq n$ and $i > n/p$ (i.e., $i \geq 1$ if $p > n$ and $i \geq 2$ if $p = n$) one has*

$$W^{i,p}(\Omega) \hookrightarrow \mathcal{C}^{i-(n/p)_+}(\Omega)$$

and there is a constant $C > 0$ such that for all $u \in W^{i,p}(\Omega)$

$$\|u\|_{\mathcal{C}^{i-(n/p)_+}(\Omega)} \leq C\,\|u\|_{i,p} \tag{6.24}$$

Proof. For a number $r = k + \alpha$ with $k \in \mathbb{N}_0$ and $0 \leq \alpha < 1$ we write $\mathcal{C}^r(\Omega)$ for $\mathcal{C}^{k,\alpha}$. As earlier it suffices to prove (6.24) for $u \in \mathcal{C}^\infty(\Omega)$. For such u and $p > n$ and $|\alpha| \leq i-1$ apply Morrey's inequality to get

$$\|D^\alpha u\|_{\mathcal{C}^{0,1-n/p}(\Omega)} \leq C\,\|D^\alpha u\|_{i,p}$$

and therefore with $\mathcal{C}^{i-n/p}(\Omega) \equiv \mathcal{C}^{i-1,1-n/p}(\Omega)$, we get (6.24).

If $p = n$ (and thus $i \geq 2$) choose $q \in (1,n)$ close to n so that $i > n/q$ and $q^* = \frac{qn}{n-q} > n$. Then, by the first part of Theorem (6.4.6) and what we have just shown

$$W^{i,n}(\Omega) \hookrightarrow W^{i,q}(\Omega) \hookrightarrow W^{i-1,q^*}(\Omega) \hookrightarrow \mathcal{C}^{i-2,1-n/q^*}(\Omega).$$

As $q \uparrow n$ implies $n/q^* \downarrow 0$, we conclude $W^{i,n}(\Omega) \hookrightarrow \mathcal{C}^{i-2,\alpha}(\Omega)$ for any $\alpha \in (0,1)$ which is written as

$$W^{i,n}(\Omega) \hookrightarrow \mathcal{C}^{i-(n/n)_+}(\Omega).$$

Theorem 6.4.6 (Sobolev Embedding Theorems) *Assume that $\Omega = \mathbb{R}^n$ or that Ω is a bounded open subset of $\mathbb{R}^n$ with a $\mathcal{C}^1$-boundary; furthermore assume that $1 \le p < \infty$ and $k, m \in \mathbb{N}$ with $m \le k$. Then one has:*

(1) If $p < n/m$, then $W^{k,p}(\Omega) \hookrightarrow W^{k-m,q}(\Omega)$ for $q = \frac{np}{n-pm}$ or $\frac{1}{q} = \frac{1}{p} - \frac{m}{n} > 0$, and there is a constant $C > 0$ such that

$$\|u\|_{k-m,q} \le C \|u\|_{k,p} \quad \text{for all } u \in W^{k,p}(\Omega). \tag{6.25}$$

(2) If $p > n/k$, then $W^{k,p}(\Omega) \hookrightarrow \mathcal{C}^{k-(n/p)_+}(\Omega)$ and there is a constant $C > 0$ such that

$$\|u\|_{\mathcal{C}^{k-(n/p)_+}(\Omega)} \le C \|u\|_{k,p} \quad \text{for all } u \in W^{k,p}(\Omega). \tag{6.26}$$

Proof. Suppose $p < n/m$ and $u \in W^{k,p}(\Omega)$; then $D^\alpha u \in W^{1,p}(\Omega)$ for all $|\alpha| \le k-1$. Corollary 6.4.4 implies $D^\alpha u \in L^{p^*}(\Omega)$ for all $|\alpha| \le k-1$ and therefore $W^{k,p}(\Omega) \hookrightarrow W^{k-1,p^*}(\Omega)$ and there is a constant $C_1 > 0$ such that

$$\|u\|_{k-1,p_1} \le C_1 \|u\|_{k,p} \tag{6.27}$$

for all $u \in W^{k,p}(\Omega)$, with $p_1 = p^*$. Next define p_j, $j \ge 2$, inductively by $p_j = p^*_{j-1}$. Thus $\frac{1}{p_j} = \frac{1}{p_{j-1}} - \frac{1}{n}$ and since $p < n/m$ we have $\frac{1}{p_m} = \frac{1}{p} - \frac{m}{n} > 0$. Therefore we can (6.27 repeatedly and find that the following inclusion maps are all bounded:

$$W^{k,p}(\Omega) \hookrightarrow W^{k-1,p_1}(\Omega) \hookrightarrow W^{k-2,p_2}(\Omega) \cdots \hookrightarrow W^{k-m,p_m}(\Omega)$$

and part (1) follows.

In order to prove part (2) consider $p > n/k$. For $p \ge n$ the statement follows from Lemma 6.4.5. Now consider the case $n > p > n/k$ and choose the largest m such that $1 \le m < k$ and $n/m > p$. Define $q \ge n$ by $q = \frac{np}{n-mp}$ (i.e., $\frac{1}{q} = \frac{1}{p} - \frac{m}{n} > 0$). Then, by what we have established above, the following inclusion maps are all bounded:

$$W^{k,p}(\Omega) \hookrightarrow W^{k-m,q}(\Omega) \hookrightarrow \mathcal{C}^{k-m-(n/q)_+}(\Omega) = \mathcal{C}^{k-m-(\frac{n}{p}-m)_+}(\Omega) = \mathcal{C}^{k-(n/p)_+}(\Omega)$$

which is the estimate of Part (2).

In the case $p = 2$ and $\Omega = \mathbb{R}^n$ one has the Fourier transform $\mathcal{F}$ available as a unitary operator on $L^2(\mathbb{R}^n)$. This allows to give a convenient characterization of the Sobolev space $H^k(\mathbb{R}^n) = W^{k,2}(\mathbb{R}^n)$ and to prove a useful embedding result.

Recall that for $u \in H^k(\mathbb{R}^n)$ one has $\mathcal{F}(D^\alpha u)(p) = i^{|\alpha|} p^\alpha \mathcal{F}(u)(p)$. Hence we can characterize this space as

$$\begin{aligned} H^k(\mathbb{R}^n) &= \{u \in L^2(\mathbb{R}^n) : p^\alpha \mathcal{F}(u) \in L^2(\mathbb{R}^n), |\alpha| \le k\} \\ &= \left\{u \in L^2(\mathbb{R}^n) : (1+p^2)^{k/2} \mathcal{F}(u) \in L^2(\mathbb{R}^n)\right\}. \end{aligned}$$

This definition can be extended to arbitrary $s \in \mathbb{R}$ and thus we can introduce the spaces

$$H^s(\mathbb{R}^n) = \left\{ u \in L^2(\mathbb{R}^n) : (1+p^2)^{s/2}\mathcal{F}(u) \in L^2(\mathbb{R}^n) \right\}.$$

As we are going to show this space can be continuously embedded into the space

$$\mathcal{C}_b^k(\mathbb{R}^n) = \left\{ f \in \mathcal{C}^k(\mathbb{R}^n) : \|f\|_{k,\infty} = \sup_{|\alpha|\le k} \sup_{x\in\mathbb{R}^n} |D^\alpha f(x)| < \infty \right\}.$$

Theorem 6.4.7 *For $k \in \mathbb{N}$ and $s > k + n/2$ the Sobolev space $H^s(\mathbb{R}^n)$ is continuously embedded into the space $\mathcal{C}_b^k(\mathbb{R}^n)$ and one has for all $u \in H^s(\mathbb{R}^n)$*

$$\|u\|_{k,\infty} \le C \|u\|_{s,2}, \quad \lim_{|x|\to\infty} |D^\alpha u(x)| = 0, \ |\alpha| \le k.$$

Proof. Recall that the Lemma of Riemann-Lebesgue says that the Fourier transform of an $L^1(\mathbb{R}^n)$ function is continuous and vanishes at infinity. For $|\alpha| \le k$ and $s > k + n/2$ one knows

$$\int_{\mathbb{R}^n} \frac{|p^{2\alpha}|}{(1+p^2)^s} dp = C_\alpha^2 < \infty.$$

Thus, for $u \in H^s(\mathbb{R}^n)$ we can estimate

$$\int_{\mathbb{R}^n} |p^\alpha (\mathcal{F}u)(p)| dp \le C_\alpha \left(\int_{\mathbb{R}^n} (1+p^2)^s |\mathcal{F}u(p)|^2 dp \right)^{1/2} = C_\alpha \|u\|_{s,2}$$

and therefore for all $x \in \mathbb{R}^n$

$$|D^\alpha u(x)| = |\int_{\mathbb{R}^n} e^{\mathrm{ipx}} p^\alpha (\mathcal{F}u)(p) dp| \le C_\alpha \|u\|_{s,2}.$$

It follows $\|u\|_{k,\infty} \le \|u\|_{s,2}$. By applying the Lemma of Riemann-Lebesgue we conclude.

6.4.2 Compact Embeddings of Sobolev spaces

Here we show that some of the continuous embeddings established above are actually compact, that is they map bounded subsets into precompact sets. There are various ways to prove these compactness results. We present a proof which is based on the characterization of compact subsets $M \subset L^q(\mathbb{R}^n)$, due to Kolmogorov and Riesz [Kol31, Rie33].

Theorem 6.4.8 (Kolmogorov-Riesz compactness criterion) *Suppose $1 \le q < \infty$. Then a subset $M \subset L^q(\mathbb{R}^n)$ is precompact if, and only if M satisfies the following threes conditions:*

(a) M is bounded, i.e.,

$$\exists_{C<\infty} \forall_{f\in M} \ \|f\| \le C;$$

(b)

$$\forall_{\epsilon>0} \exists_{R<\infty} \forall_{f\in M} \left\|\pi_R^{\perp} f\right\|_q < \epsilon;$$

(c)

$$\forall_{\epsilon>0} \exists_{r>0} \forall_{f\in M} \forall_{\substack{y\in\mathbb{R}^n \\ |y|<r}} \ \left\|\tau_y(f) - f\right\|_q < \epsilon.$$

Here the following notation is used: $\pi_R^{\perp}$ is the operator of multiplication with the characteristic function of the set $\{x \in \mathbb{R}^n : |x| > R\}$ and τ_y denotes the operator of translation by $y \in \mathbb{R}^n$, i.e., $\tau_y(f)(x) = f(x+y)$.

Remark 6.4.9 *If $\Omega \subset \mathbb{R}^n$ is an open bounded subset we can consider $L^q(\Omega)$ as a subset of $L^q(\mathbb{R}^n)$ by extending all elements $f \in L^q(\Omega)$ by 0 to all of $\mathbb{R}^n$. Then the above characterization provides also a characterization of precompact subset $M \subset L^q(\Omega)$ where naturally condition (b) is satisfied always and where in condition (c) we have to use these extensions.*

There are several versions of compact embedding results depending on the assumptions on the domain $\Omega \subset \mathbb{R}^n$ which are used. The following version is already quite comprehensive though there are several newer results of this type.

Theorem 6.4.10 (Rellich-Kondrachov compactness theorem) *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Assume that the boundary of Ω is sufficiently smooth and that $1 \le p < \infty$ and $k = 1, 2, \ldots$. Then the following holds:*

(a) The following embeddings are compact:

(i) $kp < n$: $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$, $1 \le q < p^* = \frac{np}{n-kp}$;

(ii) $kp = n$: $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$, $1 \le q < \infty$;

(iii) $kp > n$: $W^{k,p}(\Omega) \hookrightarrow \mathcal{C}_b^0(\Omega)$.

(b) For the subspaces $W_0^{k,p}(\Omega)$ the embeddings (i) - (iii) are compact for arbitrary open sets Ω.

Proof. Because of time constraints we discuss here only the proof of embedding (i) of part (a) for $k = 1$. According to Corollary 6.4.4 the inclusion mapping $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ is continuous for $1 \le q \le p^*$. We have to show that every bounded subset $M \subset W^{k,p}(\Omega)$ is precompact in $L^q(\Omega)$ for $1 \le q < p^*$. This is done by the Kolmogorov-Riesz compactness criterion. By Remark 6.4.9 only conditions (a) and (c) have to be verified for M considered as a subset of $L^q(\Omega)$. Since we know that this inclusion map is continuous, it follows that M is bounded in $L^q(\Omega)$ too and thus Condition (a) of the Kolmogorov-Riesz criterion is verified and we are left with verifying Condition (c).

Observe that for $1 \le q < p^*$ Hölder's inequality implies

$$\|u\|_q \le \|u\|_1^{\alpha} \, \|u\|_{p^*}^{1-\alpha}, \quad \alpha = \frac{1}{q}\frac{p^* - q}{p^* - 1} \in (0,1).$$

Now let $M \subset W^{1,p}(\Omega)$ be bounded; then this set is bounded in $L^{p^*}(\Omega)$ and hence there is a constant $C < \infty$ such that for all $u \in M$ we have

$$\|u\|_q \le C\, \|u\|_1^{\alpha}$$

and it follows

$$\left\|\tau_y u - u\right\|_q \le 2C \left\|\tau_y u - u\right\|_1^{\alpha}, \quad \forall\, u \in M. \tag{6.28}$$

Therefore it suffices to verify condition (c) of Theorem 6.4.8 for the norm $\|\cdot\|_1$. For $i = 1, 2, \ldots$ introduce the sets

$$\Omega_i = \{x \in \Omega : d(x, \partial\Omega) > 2/i\},$$

where $d(x, \partial\Omega)$ denotes the distance of the point x from the boundary $\partial\Omega$ of Ω. Another application of Hölder's inequality gives, for all $u \in M$,

$$\begin{aligned}\int_{\Omega\setminus\Omega_i} |u(x)| dx &\le \left(\int_{\Omega\setminus\Omega_i} |u(x)|^{p^*} dx\right)^{1/p^*} \left(\int_{\Omega\setminus\Omega_i} dx\right)^{1-\frac{1}{p^*}} \\ &\le \|u\|_{p^*} |\Omega\setminus\Omega_i|^{1-\frac{1}{p^*}} \le C_M |\Omega\setminus\Omega_i|^{1-\frac{1}{p^*}}\end{aligned}$$

where C_M is a bound for M in $L^{p^*}(\Omega)$. Given $\epsilon > 0$ we can therefore find $i_0 = i_0(\epsilon)$ such that

$$\int_{\Omega\setminus\Omega_i} |u(x)| dx < \epsilon/4$$

holds for all $u \in M$. Extend $u \in M$ outside Ω by 0 to get

$$\hat{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & \text{otherwise.}\end{cases}$$

For a fixed $i \ge i_0$ and $y \in \mathbb{R}^n$, $|y| < 1/i$, we estimate

$$\begin{aligned}\left\|\tau_y u - u\right\|_1 &= \int_{\Omega_i} |u(x+y) - u(x)| dx + \int_{\Omega\setminus\Omega_i} |\hat{u}(x+y) - \hat{u}(x)| dx \\ &\le \int_{\Omega_i} |u(x+y) - u(x)| dx + \epsilon/2\end{aligned}$$

And the integral is estimated as follows:

$$= \int_{\Omega_i} \Big| \int_0^1 \frac{d}{dt} u(x+ty)dt \Big| dx = \int_{\Omega_i} \Big| \int_0^1 y \cdot \nabla u(x+ty) dt \Big| dx \le |y| \int_{\Omega_{2i}} |\nabla u(x)| dx$$
$$\le |y| |\Omega_{2i}|^{\frac{1}{p'}} \|\nabla u\|_{L^p(\Omega_{2i})} \le |y| |\Omega_{2i}|^{\frac{1}{p'}} C \le |y| |\Omega|^{\frac{1}{p'}} C$$

It follows that there is $r_0 > 0$ such that $\|\tau_y u - u\|_1 < \epsilon$ for all $|y| < r_0$. By estimate (6.28) we conclude that Condition (c) of Theorem 6.4.8 holds and therefore by this theorem $M \subset W^{1,p}(\Omega)$ is precompact in $L^q(\Omega)$.

Remark 6.4.11 *The general case of $W^{k,p}(\Omega)$ with $k > 1$ follows from the following observation which can be proven similarly.*

> *For $m \ge 1$ and $\frac{1}{q} > \frac{1}{p} - \frac{m}{n} > 0$ the inclusion of $W^{k,p}(\Omega)$ into $W^{k-m,q}(\Omega)$ is compact.*

In the proof of compactness of the Sobolev embeddings the fact that the underlying set $\Omega \subset \mathbb{R}^n$ is bounded entered in an important way. We now mention a result for the pre-compactness in $L^q(\mathbb{R}^n)$ of certain bounded sets $M \subset W^{1,p}(\mathbb{R}^n)$.

Theorem 6.4.12 *Let $M \subset W^{1,p}(\mathbb{R}^n)$ be bounded and suppose that*

$$\forall_{\epsilon>0} \, \exists_{R<\infty} \, \forall_{u\in M} \; \left\|\pi_R^{\perp} u\right\|_{1,p} < \epsilon \tag{6.29}$$

holds. Then M is pre-compact in $L^q(\mathbb{R}^n)$ provided

$$\frac{n}{p} - 1 < \frac{n}{q} \le \frac{n}{p}.$$

Chapter 7

Some Applications

7.1 Spectral Theory for compact Operators

This chapter serves to illustrate how the variational methods can be used in a variety of mathematical problems. Naturally, in this introduction on we only consider those which are typical and technically not too involved. We begin with two classical results. Then we proceed to solve partial differential equations by variational methods. Here we distinguish between linear and nonlinear problems and between those which are defined over bounded and unbounded domains $\Omega \subset \mathbb{R}^n$.

Theorem 7.1.1 (Spectral Theorem for compact self-adjoint Ops.) *Let $\mathcal{H}$ be a separable Hilbert space and $A \neq 0$ a compact self-adjoint operator on $\mathcal{H}$. Then there are an orthonormal system $\{e_j : j \in \mathbb{N}\}$ of eigenvectors and a system of real eigenvalues $\{\lambda_j : j \in \mathbb{N}\}$, i.e., $Ae_j = \lambda_j e_j$, with the following properties.*

a) The eigenvalues are arranged in descending order of magnitude, i.e., $|\lambda_1| \geq \cdots |\lambda_j| \geq |\lambda_{j+1}| \geq \cdots$.

b) Either finitely many of the eigenvalues are nonzero, or all eigenvalues are nonzero, and $\lim_{j\to\infty} \lambda_j = 0$, i.e., 0 is the only cluster point of the eigenvalues.

c) The multiplicity of each nonzero eigenvalues is finite:

$$\dim \ker(A - \lambda_j I) < \infty \quad \textit{for all} \quad j \in \mathbb{N}.$$

d) The orthonormal system of eigenvectors $\{e_j : j \in \mathbb{N}\}$ is complete if, and only if, A is injective.

Proof. Naturally we discuss here only the 'variational part' of the proof. In a first step we collect some basic facts about operators on Hilbert spaces. Recall that a compact operator in a separable Hilbert space $\mathcal{H}$ is characterized by the property that, for every weakly convergent sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$, the sequence of images $(Ax_n)_{n\in\mathbb{N}}$ converges strongly. This fact implies immediately that the functions

$$x \to \|Ax\| \quad \text{and} \quad \mathrm{x} \to \langle \mathrm{x}, \mathrm{Ax} \rangle$$

are weakly continuous on $\mathcal{H}$.

Recall furthermore that the norm of an operator A can be calculated as follows:

$$\|A\| = \sup_{u\in B_1(\mathcal{H})} \|Au\| = \sup_{u\in S_1(\mathcal{H})} \|Au\| = \|A\|_{B(\mathcal{H})},$$

$$B_1(\mathcal{H}) = \{x \in \mathcal{H} : \|x\| \le 1\}, \qquad S_1(\mathcal{H}) = \{x \in \mathcal{H} : \|x\| = 1\}.$$

In the second step we determine the largest (in absolute value) eigenvalue of A. Since the closed unit ball $B_1(\mathcal{H})$ of a Hilbert space is weakly (sequentially) compact and the functions $u \to \|Au\|$ and $u \to |\langle u, Au\rangle|$ are weakly continuous, Theorem 2.3.4 can be applied. Hence there is $e_1 \in B_1(\mathcal{H})$ such that

$$\|Ae_1\| = \sup_{u\in B_1(\mathcal{H})} \|Au\| = \|A\|.$$

$A \neq 0$ implies $e_1 \neq 0$, and so $\|e_1\| = 1$, since if $0 < \|e_1\| < 1$ then

$$\|Ae_1\| < \frac{\|Ae_1\|}{\|e_1\|} = \|A\hat{e}_1\|,$$

a contradiction to $\hat{e}_1 = \frac{1}{\|e_1\|}e_1 \in S_1(\mathcal{H})$.

In the third step it is shown that the maximizing element $e_1 \in S_1(\mathcal{H})$ is an eigenvalue of A. To this end we note that the function $x \to Q(x) = \langle x, Ax\rangle$ is differentiable on $\mathcal{H}$ with the Fréchet derivative $Q'(x)(h) = 2\langle Ax, h\rangle$ for all $h \in \mathcal{H}$ (here we use self-adjointness $A^* = A$ of A) and, as we had mentioned above,

$$\sup_{x\in S_1(\mathcal{H})} Q(x) = \pm\|A\|.$$

The constraint $\phi(x) = \langle x, x\rangle - 1 = 0$ is defined by a mapping ϕ which is regular at all points of the level surface $\phi^{-1}(0) = S_1(\mathcal{H})$, since $\phi'(x)(h) = 2\langle x, h\rangle$ for all $h \in \mathcal{H}$.

Since we know that $e_1 \in \phi^{-1}(0)$ is an extremal point of the functional Q, Theorem 4.3.1 (existence of a Lagrange multiplier) assures that there is $\lambda_1 \in \mathbb{R}$ such that

$$Q'(e_1) = \lambda_1\phi'(e_1).$$

It follows $Ae_1 = \lambda_1 e_1$ and $|\lambda_1| = \|A\|$.

In a fourth step a reduction of the problem to a 'smaller' space $\mathcal{H}_1 = \{e_1\}^\perp \subset \mathcal{H}$ is done. Since A is self-adjoint, it maps the space $\mathcal{H}_1$ into itself. Thus the restriction $A_1 = A \restriction \mathcal{H}_1$ of A to the subspace $\mathcal{H}_1$ is well defined. It also is self-adjoint and compact. Its norm is not larger than that of A: $\|A_1\|_{B(\mathcal{H}_1)} \le \|A\|_{B(\mathcal{H})}$. If $A_1 \neq 0$, then our considerations from above apply to the compact self-adjoint operator A_1 in the reduced Hilbert space $\mathcal{H}_1$. Hence there are a normalized eigenvector $e_2 \in \mathcal{H}_1$ of A_1 for an eigenvalue λ_2, such that

$$|\lambda_2| = \|A_1\|_{B(\mathcal{H}_1)} \le \|A\|_{B(\mathcal{H})} = |\lambda_1| \quad \text{and} \quad \langle \mathrm{e}_2, \mathrm{e}_1\rangle = 0$$

hold. This reduction can be iterated and produces a sequence of subspaces $\mathcal{H}_j = \{e_{j-1}\}^\perp \subset \mathcal{H}_{j-1}$ and operators $A_j = A \restriction \mathcal{H}_j$. Obviously this iteration stops for j when $A_j = 0$; otherwise it continues. The rest of the proof does not use anymore variational arguments is omitted.

The above result establishes the existence and some of the properties of eigenvalues of a compact self-adjoint operator. The question arises whether there is some method to calculate these eigenvalues. And indeed such a method has been worked out long ago and is known as the *classical minimax principle of Courant–Weyl–Fischer–Poincaré–Rayleigh–Ritz*.

Theorem 7.1.2 (Minimax Principle) *Let $\mathcal{H}$ be a real separable Hilbert space and $A \geq 0$ a self-adjoint operator on $\mathcal{H}$ with spectrum $\sigma(A) = \{\lambda_m : m \in \mathbb{N}\}$ ordered according to size, $\lambda_m \leq \lambda_{m+1}$. For $m = 1, 2, \ldots$ denote by $\mathcal{E}_m$ the family of all m-dimensional subspaces E_m of $\mathcal{H}$. Then the eigenvalue λ_m can be calculated as*

$$\lambda_m = \min_{E_m \in \mathcal{E}_m} \max_{v \in E_m} \frac{\langle v, Av \rangle}{\langle v, v \rangle}. \tag{7.1}$$

Proof. The proof is obtained by determining the lower bound for the values of the Rayleigh quotient $R(v) = \frac{\langle v, Av \rangle}{\langle v, v \rangle}$. In order to do this we expand every $v \in \mathcal{H}$ in terms of eigenvectors e_j of A. This gives $v = \sum_{i=1}^{\infty} a_i e_i$ and $\langle v, v \rangle = \sum_{i=1}^{\infty} a_i^2$. In this form the Rayleigh quotient reads

$$R(v) = \frac{\sum_{i=1}^{\infty} \lambda_i a_i^2}{\sum_{i=1}^{\infty} a_i^2}.$$

Denote by V_m the linear subspace generated by the first m eigenvectors of A. It follows that

$$\max_{v \in V_m} R(v) = \max_{(a_1, \ldots, a_m) \in \mathbb{R}^m} \frac{\sum_{i=1}^{m} \lambda_i a_i^2}{\sum_{i=1}^{m} a_i^2} = \lambda_m = R(e_m),$$

and thus we are left with showing $\max_{v \in E_m} R(v) \geq \lambda_m$ for every other subspace $E_m \in \mathcal{E}_m$. Let $E_m \neq V_m$ be such a subspace; then $E_m \cap V_m^{\perp} \neq \{0\}$ and therefore

$$\max_{v \in E_m} R(v) \geq \max_{v \in E_m \cap V_m^{\perp}} R(v).$$

Every $v \in E_m \cap V_m^{\perp}$ is of the form $v = \sum_{i \geq m+1} a_i e_i$ and for such vectors we have

$$R(v) = \frac{\sum_{i \geq m+1} \lambda_i a_i^2}{\sum_{i \geq m+1} a_i^2} \geq \lambda_{m+1} \geq \lambda_m.$$

This then completes the proof. □

Theorem 7.1.2 implies for the smallest eigenvalue of the operator A the simple formula

$$\lambda_1 = \min_{v \in E,\, v \neq 0} \frac{\langle v, Av \rangle}{\langle v, v \rangle}. \tag{7.2}$$

7.2 Some linear boundary and eigenvalue problems

7.2.1 The Dirichlet–Laplace operator

The goal of this section is to illustrate the application of the general strategy and the results developed thus far. This is done by solving several relatively simple linear boundary and eigenvalue problems. The typical example is the Laplace operator with Dirichlet boundary conditions on a bounded domain Ω. Naturally, for these concrete problems we have to use concrete function spaces, and we need to know a number of basic facts about them. In this brief introduction we have to refer the reader to the literature for the proof of these facts. We recommend the books [LL01, JLJ98, BB92].

For a bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary $\partial\Omega$ consider the real Hilbert space $L^2(\Omega)$ with inner product $\langle \cdot, \cdot \rangle_2$. Then define a space $H^1(\Omega)$ as

$$H^1(\Omega) = \left\{ u \in L^2(\Omega) : \partial_j u \in L^2(\Omega),\ j = 1, \ldots, n \right\}. \qquad (7.3)$$

Here naturally the partial derivatives $\partial_j u$ are understood in the weak (distributional) sense. One shows that $H^1(\Omega)$ is a Hilbert space with the inner product

$$\langle u, v \rangle = \langle u, v \rangle_2 + \langle Du, Dv \rangle_2 \qquad \forall\, u, v \in H^1(\Omega) \qquad (7.4)$$

where $Du = (\partial_1 u, \ldots, \partial_n u)$ and where in the second term the natural inner product of $L^2(\Omega)^{\times n}$ is used. This space is the *Sobolev* space $W^{1,2}(\Omega)$ as introduced in Chapter 6. Next define a subspace $H_0^1(\Omega)$ of this space as the closure of the space $\mathcal{D}(\Omega)$ of $\mathcal{C}^\infty$-functions on Ω with compact support:

$$H_0^1(\Omega) = \text{closure of } \mathcal{D}(\Omega) \text{ in } H^1(\Omega). \qquad (7.5)$$

Intuitively, $H_0^1(\Omega)$ is the subspace of those $u \in H^1(\Omega)$ whose restriction to the boundary $\partial\Omega$ vanishes, $u_{|\partial\Omega} = 0$.

The Sobolev space $H^1(\Omega)$ is by definition contained in the Hilbert space $L^2(\Omega)$, however for us of much greater importance is the fact that the following embeddings for $2 \leq n$,

$$H^1(\Omega) \hookrightarrow L^q(\Omega), \quad 1 \leq q < 2^* = \frac{2n}{n-2}, \quad 2 < n \tag{7.6}$$

and

$$H^1(\Omega) \hookrightarrow L^q(\Omega), \quad 1 \leq q < \infty, \quad 2 = n \tag{7.7}$$

are compact (see Theorem 6.4.10). This means that every weakly convergent sequence in $H^1(\Omega)$ converges strongly in $L^q(\Omega)$. In addition we are going to use the important *Sobolev inequality*

$$\|u\|_q \leq S\,\|Du\|_p = S(\sum_{j=1}^{n} \|\partial_j u\|_p^p)^{1/p} \qquad \forall\, u \in H^1(\Omega), \tag{7.8}$$

where S is the Sobolev constant depending on q, n and where q is in the range indicated in (7.6) respectively (7.7) (compare Corollary 6.4.4) .

Now we are in the position to show that the famous Dirichlet problem has a solution.

Theorem 7.2.1 (Dirichlet problem) *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary and $v_0 \in H^1(\Omega)$ some given element. Then the Dirichlet integral*

$$f(v) = \int_\Omega |Dv(x)|^2 dx = \int_\Omega \sum_{j=1}^{n} |\partial_j v(x)|^2 dx \tag{7.9}$$

is minimized on $M = v_0 + H_0^1(\Omega)$ by an element $v \in M$ satisfying

$$\triangle v = 0 \quad \text{in } \Omega \qquad \text{and} \qquad v_{|\partial\Omega} = v_{0|\partial\Omega}. \tag{7.10}$$

Proof. Observe that $f(u) = Q(u,u)$ with the quadratic functional $Q(u,v) = \langle Du, Dv\rangle_2$. This quadratic form satisfies, because of inequality (7.8), the estimate

$$c\,\|u\|^2 \leq Q(u,u) \leq \|u\|^2$$

for some $c > 0$. It follows that Q is a strictly positive continuous quadratic form on $H^1(\Omega)$ and thus f is a strictly convex continuous function on this space (see the proof of Theorem 2.4.1). We conclude, by Lemma 2.3.2 or Theorem 2.4.1, that f is weakly lower semi-continuous on $H^1(\Omega)$.

As a Hilbert space, $H_0^1(\Omega)$ is weakly complete and thus the set $M = v_0 + H_0^1(\Omega)$ is weakly closed. Therefore Theorem 2.3.5 applies and we conclude that there is a minimizing element v for the functional f on M.

Since the minimizing element $v = v_0 + u \in M$ satisfies $f(v) = f(v_0 + u) \leq f(v_0 + w)$ for all $w \in H_0^1(\Omega)$ we deduce as earlier that $f'(v)(w) = 0$ for all $w \in H_0^1(\Omega)$ and thus

$$0 = f'(v)(w) = \int_\Omega Dv(x) \cdot Dw(x) dx \qquad \forall\, w \in \mathcal{D}(\Omega).$$

Recalling the definition of differentiation in the sense of distributions, this means $-\triangle v = 0$ in the sense of $\mathcal{D}'(\Omega)$. Now the Lemma of Weyl (see [JLJ98, BB92]) implies that $-\triangle v = 0$ also holds in the classical sense, i.e., as an identity for functions of class $\mathcal{C}^2$.

Because for $u \in H_0^1(\Omega)$ one has $u_{|\partial\Omega} = 0$ the minimizer v satisfies the boundary condition too. Thus we conclude. □

As a simple application of the theory of constrained minimization we solve the eigenvalue problem for the Laplace operator on an open bounded domain Ω with Dirichlet boundary conditions, i.e., the problem is to find a number λ and a function $u \neq 0$ satisfying

$$-\triangle u = \lambda u \quad \text{in} \quad \Omega, \qquad u_{|\partial\Omega} = 0. \tag{7.11}$$

The strategy is simple. On the Hilbert space $H_0^1(\Omega)$ we minimize the functional $f(u) = \frac{1}{2}\langle Du, Du\rangle_2$ under the constraint $g(u) = \frac{1}{2}$ for the constraint functional $g(u) = \frac{1}{2}\langle u, u\rangle_2$. The derivative of g is easily calculated; it is $g'(u)(v) = \langle u, v\rangle_2$ for all $v \in H_0^1(\Omega)$ and thus the level surface $[g = \frac{1}{2}]$ consists only of regular points of the mapping g.

Since we know that f is weakly lower semi-continuous and coercive on $H_0^1(\Omega)$ we can prove the existence of a minimizer for the functional f on $[g = \frac{1}{2}]$ by verifying that $[g = \frac{1}{2}]$ is weakly closed and then to apply Theorem 2.3.5.

Suppose a sequence $(u_j)_{j\in\mathbb{N}}$ converges to u weakly in $H_0^1(\Omega)$. According to the Sobolev embedding (7.6) the space $H_0^1(\Omega)$ is compactly embedded into the space $L^2(\Omega)$ and thus this sequence converges strongly in $L^2(\Omega)$ to u. It follows that $g(u_j) \to g(u)$ as $j \to \infty$, i.e., g is weakly continuous on $H_0^1(\Omega)$ and its level surfaces are weakly closed.

Theorem 2.3.5 implies the existence of a minimizer of f under the constraint $g(u) = 1/2$. Using Corollary 4.2.3 and Theorem 4.3.1 we deduce that there is a Lagrange multiplier $\lambda \in \mathbb{R}$

for this constrained minimization problem, i.e., a real number λ satisfying $f'(u) = \lambda g'(u)$. In detail this identity reads

$$\int_\Omega Du(x) \cdot Dv(x) dx = \lambda \int_\Omega u(x)v(x) dx \qquad \forall\, v \in H_0^1(\Omega),$$

and in particular for all $v \in \mathcal{D}(\Omega)$, thus $-\triangle u = \lambda u$ in $\mathcal{D}'(\Omega)$; and by elliptic regularity theory (see for instance Section 9.3 of [BB92]) we conclude that this identity holds in the classical sense. Since the solution u belongs to the space $H_0^1(\Omega)$ it satisfies the boundary condition $u_{|\partial\Omega} = 0$. This proves

Theorem 7.2.2 (Dirichlet Laplacian) *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary $\partial\Omega$. Then the eigenvalue problem for the Laplace operator with Dirichlet boundary conditions (7.11) has a solution.*

The above argument which proved the existence of the lowest eigenvalue λ_1 of the Dirichlet–Laplace operator can be repeated on the orthogonal complement of the eigenfunction u_1 of the first eigenvalue and thus gives an eigenvalue $\lambda_2 \geq \lambda_1$ (some additional arguments show $\lambda_2 > \lambda_1$). In this way one proves actually the existence of an infinite sequence of eigenvalues for the Dirichlet–Laplace operator. By involving some refined methods of the theory of Hilbert space operators it can be shown that these eigenvalues are of the order $\lambda_k \approx$ constant $(\frac{k}{|\Omega|})^{\frac{2}{n}}$ (see for instance [LL01]).

7.2.2 Linear elliptic differential operators

Next we consider more generally the following class of second order linear partial differential operators A defined on sufficiently smooth functions u by

$$Au = A_0 u - \sum_{j=1}^{n} \partial_j \Big(\sum_{i=1}^{n} a_{ji} \partial_i u\Big). \tag{7.12}$$

The matrix a of coefficient functions $a_{ji} = a_{ij} \in L^\infty(\Omega)$ satisfies for almost all $x \in \Omega$ and all $\xi \in \mathbb{R}^n$,

$$m \sum_{j=1}^{n} \xi_j^2 \leq \sum_{i,j=1}^{n} \xi_j a_{ji}(x) \xi_i \leq M \sum_{j=1}^{n} \xi_j^2 \tag{7.13}$$

for some constants $0 < m < M$. A_0 is a bounded symmetric operator in $L^2(\Omega)$ which is bounded from below, $\langle u, A_0 u\rangle_2 \geq -r \|u\|_2^2$ for some positive number r satisfying $0 \leq r < \frac{m}{c^2}$. Here m is the constant in condition (7.13) and c is the smallest constant for which $\|u\|_2 \leq c \|Du\|_2$ holds for all $u \in H_0^1(\Omega)$.

As we are going to show, under these assumptions, the arguments used for the study of the Dirichlet problem and the eigenvalue problem for the Dirichlet–Laplace operator still apply. The associated quadratic form

$$Q(u,v) = \langle u, A_0 v\rangle_2 + \sum_{i,j=1}^{n} \langle \partial_j v, a_{ji} \partial_i u\rangle_2 \qquad \forall\, u, v \in H_0^1(\Omega)$$

is strictly positive since the ellipticity condition (7.13) and the lower bound for A_0 imply

$$\begin{aligned} Q(u,u) &= \langle u, A_0 u\rangle_2 + \int_\Omega \sum_{i,j=1}^{n} \partial_j v(x) a_{ji}(x) \partial_i u(x) dx \\ &\geq -r \|u\|_2^2 + \int_\Omega m \sum_{j=1}^{n} (\partial_j u(x))^2 dx = -r \|u\|_2^2 + m \|Du\|_2^2 \\ &\geq (-rc^2 + m) \|Du\|_2^2 = c_0 \|Du\|_2^2, \quad c_0 = -rc^2 + m > 0. \end{aligned}$$

As earlier we deduce that the functional $f(u) = Q(u,u)$ is coercive and weakly lower semi-continuous on $H^1(\Omega)$. Hence Theorem 2.3.5 allows us to minimize f on $M = v_0 + H_0^1(\Omega)$ and thus to solve the boundary value problem for a given $v_0 \in H^1(\Omega)$ or on the level surface $[g = \frac{1}{2}]$ for the constraint function $g(u) = \frac{1}{2}\langle u, u\rangle_2$ on $H_0^1(\Omega)$. The conclusion is that the *linear elliptic partial differential operator* (7.12) with Dirichlet boundary conditions has an increasing sequence of eigenvalues, as it is the case for the Laplace operator.

7.3 Some nonlinear boundary and eigenvalues problems

In order to be able to minimize functionals of the general form (1.2) we first have to find a suitable domain of definition and then to have enough information about it. We begin with the description of several important aspects from the theory of Lebesgue spaces. A good reference for this are paragraphs 18–20 of [Vai64].

Let $\Omega \subset \mathbb{R}^n$ be a nonempty open set and $h : \Omega \times \mathbb{R} \to \mathbb{R}$ a function such that $h(\cdot, y)$ is measurable on Ω for every $y \in \mathbb{R}$ and $y \mapsto h(x, y)$ is continuous for almost every $x \in \Omega$. Such functions are often called *Carathéodory functions*. If now $u : \Omega \to \mathbb{R}$ is (Lebesgue) measurable, define $\hat{h}(u) : \Omega \to \mathbb{R}$ by $\hat{h}(u)(x) = h(x, u(x))$ for almost every $x \in \Omega$. Then $\hat{h}(u)$ is measurable too. For our purposes it is enough to consider $\hat{h}$ on Lebesgue integrable functions $u \in L^p(\Omega)$ and we need that the image $\hat{h}(u)$ is Lebesgue integrable too, for instance $\hat{h}(u) \in L^q(\Omega)$ for some exponents $1 \leq p, q$. Therefore the following lemma will be useful.

Lemma 7.3.1 *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set and $h : \Omega \times \mathbb{R} \to \mathbb{R}$ a Carathéodory function. Then $\hat{h}$ maps $L^p(\Omega)$ into $L^q(\Omega)$ if, and only if, there are $0 \leq a \in L^q(\Omega)$ and $b \geq 0$ such that for almost all $x \in \Omega$ and all $y \in \mathbb{R}$,*

$$|h(x, y)| \leq a(x) + b|y|^{p/q}. \tag{7.14}$$

If this condition holds the map $\hat{h} : L^p(\Omega) \to L^q(\Omega)$ is continuous.

This result extends naturally to Carathéodory functions $h : \Omega \times \mathbb{R}^{n+1} \to \mathbb{R}$. For $u_j \in L^{p_j}(\Omega)$, $j = 0, 1, \ldots, n$ define

$$\hat{h}(u_0, \ldots, u_n)(x) = h(x, u_0(x), \ldots, u_n(x))$$

for almost every $x \in \Omega$. Then $\hat{h} : L^{p_0}(\Omega) \times \cdots \times L^{p_n}(\Omega) \to L^q(\Omega)$ if, and only if, there are $0 \leq a \in L^q(\Omega)$ and $b \geq 0$ such that

$$|h(x, y_0, \ldots, y_n)| \leq a(x) + b \sum_{j=0}^{n} |y_j|^{p_j/q}. \tag{7.15}$$

And $\hat{h}$ is continuous if this condition holds.

As a last preparation define, for every $u \in W^{1,p}(\Omega)$, the functions $y(u) = (y_0(u), y_1(u), \ldots, y_n(u))$ where $y_0(u) = u$ and $y_j(u) = \partial_j u$ for $j = 1, \ldots, n$. By definition of the Sobolev space $W^{1,p}(\Omega)$ we know that

$$y : W^{1,p}(\Omega) \to L^p(\Omega) \times \cdots \times L^p(\Omega) = L^p(\Omega)^{\times(n+1)}$$

is a continuous linear map.

Now suppose that the integrand in formula (1.2) is a Carathéodory function and satisfies the bound

$$|F(x, y)| \leq a(x) + b \sum_{j=0}^{n} |y_j|^p, \tag{7.16}$$

for all $y \in \mathbb{R}^{n+1}$ and almost all $x \in \Omega$, for some $0 \leq a \in L^1(\Omega)$ and some constant $b \geq 0$. Then, as a composition of continuous mappings, $\hat{F} \circ y$ is a well-defined continuous mapping $W^{1,p}(\Omega) \to L^1(\Omega)$. We conclude that under the growth restriction (7.16) the Sobolev space $W^{1,p}(\Omega)$ is a suitable domain for the functional

$$f(u) = \int_\Omega F(x, u(x), Du(x))dx. \tag{7.17}$$

For $1 < p < \infty$ the Sobolev spaces $W^{1,p}(\Omega)$ are known to be separable reflexive Banach spaces, and thus well suited for the direct methods ([LL01]).

Proposition 7.3.2 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $F : \Omega \times \mathbb{R}^{n+1} \to \mathbb{R}$ a Carathéodory function.*

a) *If F satisfies the growth restriction (7.16), then a functional f : $W^{1,p}(\Omega) \to \mathbb{R}$ is well defined by (7.17). It is polynomially bounded according to*

$$|f(u)| \leq \|a\|_1 + b\,\|u\|_p^p + b\,\|Du\|_p^p \qquad \forall u \in W^{1,p}(\Omega). \tag{7.18}$$

b) *If F satisfies a lower bound of the form*

$$F(x,y) \geq -\alpha(x) - \beta|y_0|^r + c|\underline{y}|^p \tag{7.19}$$

for all $y = (y_0, \underline{y}) \in \mathbb{R}^{n+1}$ and almost all $x \in \Omega$, for some $0 \leq \alpha \in L^1(\Omega)$, $\beta \geq 0$, $c > 0$ and $0 \leq r < p$, then the functional f is coercive.

c) *If $y \mapsto F(x,y)$ is convex for almost all $x \in \Omega$, then f is lower semi-continuous for the weak topology on $W^{1,p}(\Omega)$.*

Proof. To complete the proof of Part a) we note that the assumed bound for F implies that $|F \circ y(u)(x)| \leq a(x) + b\sum_{j=0}^{n} |y_j(u)(x)|^p$ and thus by integration the polynomial bound follows.

Integration of the lower bound $F(x,u(x),Du(x)) \geq -\alpha(x) - \beta|u(x)|^r + c|Du(x)|^p$ for almost all $x \in \Omega$ gives $f(u) \geq -\|\alpha\|_1 - \beta\,\|u\|_r^r + c\,\|Du\|_p^p$. By inequality (7.8), $\|u\|_r^r \leq S^r\,\|Du\|_p^r$, hence $f(u) \to \infty$ as $\|Du\|_p \to \infty$ since $r < p$ and $c > 0$.

For any $u, v \in W^{1,p}(\Omega)$ and $0 \leq t \leq 1$ we have $\hat{F}(y(tu + (1-t)v)) = \hat{F}(ty(u) + (1-t)y(v)) \leq t\hat{F}(y(u)) + (1-t)\hat{F}(y(v))$ since F is assumed to be convex with respect to y. Hence integration over Ω gives $f(tu + (1-t)v) \leq tf(u) + (1-t)f(v)$. This shows that f is a convex functional. According to Part a), f is continuous on $W^{1,p}(\Omega)$, therefore Lemma 2.3.2 implies that f is weakly lower semi-continuous on $W^{1,p}(\Omega)$. □

Let us remark that the results presented in Part c) of Proposition 7.3.2 are not optimal (see for instance [Dac82, JLJ98, Str00]). But certainly the result given above has the advantage of a very simple proof. The above result uses stronger assumptions insofar as convexity with respect to u and Du is used whereas in fact convexity with respect to Du is sufficient.

Suppose we are given a functional f of the form (7.17) for which parts a) and c) of Proposition 7.3.2 apply. Then, by Theorem 2.3.4 we can minimize f on any bounded weakly closed subset $M \subset W^{1,p}(\Omega)$. If in addition f is coercive, i.e., if Part b) of Proposition 7.3.2 applies too, then we can minimize f on any weakly closed subset $M \subset W^{1,p}(\Omega)$.

In order to relate these minimizing points to solutions of nonlinear partial differential operators we need differentiability of the functional f. For this we will not consider the most general case but make assumptions which are typical and allow a simple proof.

Let us assume that the integrand F of the functional f is of class $\mathcal{C}^1$ and that all derivatives $F_j = \frac{\partial F}{\partial y_j}$ are again Carathéodory functions. Assume furthermore that there are functions $0 \leq a_j \in L^{p'}(\Omega)$ and constants $b_j > 0$ such that for all $y \in \mathbb{R}^{n+1}$ and almost all $x \in \Omega$,

$$|F_j(x,y)| \leq a_j(x) + b_j \sum_{j=0}^{n} |y_j|^{p-1}, \qquad j = 0,1,\ldots,n \tag{7.20}$$

where p' denotes the Hölder conjugate exponent, $\frac{1}{p} + \frac{1}{p'} = 1$. Since $(p-1)p' = p$ we get for all $u \in W^{1,p}(\Omega)$ the simple identity $\|y_j(u)\|_{p'}^{p'} = \|y_j(u)\|_p^p$ and it follows that $\hat{F}_j(y(u)) \in L^{p'}(\Omega)$ for all $u \in W^{1,p}(\Omega)$ and $j = 0,1,\ldots,n$. This implies the estimates, for all $u,v \in W^{1,p}(\Omega)$,

$$\left\|\hat{F}_j(y(u))y_j(v)\right\|_1 \leq \left\|\hat{F}_j(y(u))\right\|_{p'} \left\|y_j(v)\right\|_p, \qquad j = 0,1,\ldots,n$$

and thus

$$v \mapsto \int_\Omega \sum_{j=0}^{n} F_j(x, y(u)(x)) y_j(v)(x) dx \tag{7.21}$$

is a continuous linear functional on $W^{1,p}(\Omega)$, for every $u \in W^{1,p}(\Omega)$. Now it is straightforward (see Exercises) to calculate the derivative of the functional f, by using Taylor's Theorem. The result is the functional

$$f'(u)(v) = \int_\Omega \sum_{j=0}^{n} F_j(x, y(u)(x)) y_j(v)(x) dx \quad \forall\, u,v \in W^{1,p}(\Omega). \tag{7.22}$$

As further preparation for the solution of nonlinear eigenvalue problems we specify the relevant properties of the class of constraint functionals

$$g(u) = \int_\Omega G(x, u(x))dx, \qquad u \in W^{1,p}(\Omega) \tag{7.23}$$

which we are going to use. Here G is a Carathéodory function which has a derivative $G_0 = \frac{\partial G}{\partial u}$ which itself is a Carathéodory function. Since we are working on the space $W^{1,p}(\Omega)$ we assume the following growth restrictions. There are functions $0 \leq \alpha \in L^1(\Omega)$ and $0 \leq \alpha_0 \in L^{p'}(\Omega)$ and constants $0 \leq \beta, \beta_0$ such that for all $u \in \mathbb{R}$ and almost all $x \in \Omega$,

$$|G(x,u)| \leq \alpha(x) + \beta|u|^q, \qquad |G_0(x,u)| \leq \alpha_0(x) + \beta_0|u|^{q-1} \tag{7.24}$$

with an exponent q satisfying $2 \leq q < p^*$. Because of Sobolev's inequality (7.8) the functional g is well defined and continuous on $W^{1,p}(\Omega)$ and its absolute values are bounded by $|g(u)| \leq \|\alpha\|_1 + \beta \|u\|_q^q$.

Since $2 \leq q < p^*$ there is an exponent $1 \leq r < p^*$ such that $(q-1)r' < p^*$ (in the Exercises the reader is asked to show that any choice of r with $\frac{p^*}{p^*+1-q} < r < p^*$ satisfies this requirement). Then Hölder's inequality implies $\| \, |u|^{q-1}v\|_1 \leq \| \, |u|^{q-1}\|_{r'} \, \|v\|_r$. Therefore the bound for G_0 shows that for every $u \in W^{1,p}(\Omega)$ the functional $v \mapsto \int_\Omega G_0(x, u(x))v(x)dx$ is well defined and continuous on $W^{1,p}(\Omega)$. Now it is straightforward to show that the functional g is Fréchet differentiable on $W^{1,p}(\Omega)$ with derivative

$$g'(u)(v) = \int_\Omega G_0(x, u(x))v(x)dx \qquad \forall\, u, v \in W^{1,p}(\Omega). \tag{7.25}$$

Finally we assume that g has a level surface $[g = c]$ with the property that $g'(u) \neq 0$ for all $u \in [g = c]$.

A simple example of a function G for which all the assumptions formulated above are easily verified is $G(x,u) = au^2$ for some constant $a > 0$. Then all level surfaces $[g = c]$, $c > 0$, only contain regular points of g.

The nonlinear eigenvalue problems which can be solved by the strategy indicated above are those of divergence type , i.e., those which are of the form (7.26) below.

Theorem 7.3.3 (Nonlinear eigenvalue problem) *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary $\partial\Omega$ and $F : \Omega \times \mathbb{R}^{n+1} \to \mathbb{R}$ a Carathéodory function which satisfies all the hypotheses of Proposition 7.3.2 and in addition the growth restrictions (7.20) for its derivatives F_j. Furthermore let $G : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function with derivative G_0 which satisfies the growth conditions (7.24). Finally assume that the constraint functional g defined by G has a level surface $[g = c]$ which consists of regular points of g. Then the nonlinear eigenvalue problem*

$$F_0(x, u(x), Du(x)) - \sum_{j=1}^{n} \partial_j F_j(x, u(x), Du(x)) = \lambda G_0(x, u(x)) \tag{7.26}$$

with Dirichlet boundary conditions has a nontrivial solution $u \in W_0^{1,p}(\Omega)$.

Proof. Because of the Dirichlet boundary conditions we consider the functionals f and g on the closed subspace

$$E = W_0^{1,p}(\Omega) = \quad \text{closure of } \mathcal{D}(\Omega) \text{ in } W^{1,p}(\Omega). \tag{7.27}$$

Proposition 7.3.2 implies that f is a coercive continuous and weakly lower semi-continuous functional on E. The derivative of f is given by the restriction of the identity (7.22) to E.

Similarly, the functional g is defined and continuous on E and its derivative is given by the restriction of the identity (7.25) to E. Furthermore the bound (7.24) implies that g is defined and thus continuous on $L^q(\Omega)$.

Now consider a level surface $[g = c]$ consisting of regular points of g. Suppose $(u_n)_{n\in\mathbb{N}}$ is a weakly convergent sequence in E, with limit u. Because of the compact embedding of E into $L^q(\Omega)$ this sequence converges strongly in $L^q(\Omega)$. Since g is continuous on $L^q(\Omega)$ we conclude that $(g(u_n))_{n\in\mathbb{N}}$ converges to $g(u)$, thus g is weakly continuous on E. Therefore all level surface of g are weakly closed.

Theorem 2.3.5 implies that the functional f has a minimizing element $u \in [g = c]$ on the level surface $[g = c]$. By assumption, u is a regular point of g, hence Theorem 4.3.1 on the existence of a Lagrange multiplier applies and assures the existence of a number $\lambda \in \mathbb{R}$ such that

$$f'(u) = \lambda g'(u). \tag{7.28}$$

In detail this equations reads: $f'(u)(v) = \lambda g'(u)(v)$ for all $v \in E$ and thus for all v in the dense subspace $\mathcal{D}(\Omega)$ of $E = W_0^{1,p}(\Omega)$.

For $v \in \mathcal{D}(\Omega)$ we calculate

$$\begin{aligned} f'(u)(v) &= \int_\Omega F_0(x, u(x), Du(x))v(x)dx + \int_\Omega \sum_{j=1}^n F_j(x, u(x), Du(x))\partial_j v(x)dx \\ &= \int_\Omega F_0(x, u(x), Du(x))v(x)dx + \int_\Omega \sum_{j=1}^n \partial_j [F_j(x, u(x), Du(x))v(x)]dx \\ &- \int_\Omega \sum_{j=1}^n (\partial_j F_j(x, u(x), Du(x)))v(x)dx \\ &= \int_\Omega [F_0(x, u(x), Du(x)) - \sum_{j=1}^n (\partial_j F_j(x, u(x), Du(x))]v(x)dx \end{aligned}$$

since the second integral vanishes because of the Gauss divergence theorem and $v \in \mathcal{D}(\Omega)$. Hence equation (7.28) implies

$$\int_\Omega [F_0(x, u(x), Du(x)) - \sum_{j=1}^n (\partial_j F_j(x, u(x), Du(x)) - \lambda G_0(x, u(x)]v(x)dx = 0$$

for all $v \in \mathcal{D}(\Omega)$. We conclude that u solves the eigenvalue equation (7.26). □

Remark 7.3.4 1. *A very important assumption in the problems we solved in this section was that the domain $\Omega \subset \mathbb{R}^n$ on which we studied differential operators is bounded so that compact Sobolev embeddings can be used. Certainly, this strategy breaks down if Ω is not bounded. Nevertheless there are many important problems on unbounded domains Ω and one has to modify the strategy presented above. In the last twenty years considerable progress has been made in solving these global problems. The interested reader is referred to the books [BB92, LL01] and in particular to the book [Str00] for a comprehensive presentation of the new strategies used for the global problems.*

2. *As is well known, a differentiable function can have other critical points than minima or maxima for which we have developed a method to prove their existence and in favorable situations to calculate them.For these other critical points of functionals (saddle points or mountain passes) a number of other, mainly topological methods have been shown to be quite effective in proving their existence, such as index theories, mountain pass lemmas, perturbation theory). Modern books which treat these topics are [Str00, JLJ98] where one also finds many references to original articles.*

3. *The well-known mountain pass lemma of Ambrosetti and Rabinowitz is a beautiful example of results in variational calculus where elementary intuitive considerations have lead to a powerful analytical tool for finding critical points of functionals f on infinite dimensional Banach spaces E.*

   To explain this lemma in intuitive terms consider the case of a function f on $E = \mathbb{R}^2$ which has only positive values. We can imagine that f gives the height of the surface of the earth over a certain reference plane. Imagine further a town T_0 which is surrounded by a mountain chain. Then, in order to get to another town T_1 beyond this mountain chain, we have to cross the mountain chain at some point S. Certainly we want to climb as little as possible, i.e., at a point S with minimal height $f(S)$. Such a point is a mountain pass of minimal height which is a saddle point of the function f. All other mountain passes M have a height $f(M) \geq f(S)$. Furthermore we know $f(T_0) < f(S)$ and $f(T_1) < f(S)$. In order to get from town T_0 to town T_1 we go along a continuous path γ which has to wind through the mountain chain, $\gamma(0) = T_0$ and $\gamma(1) = T_1$. As described above we know $\sup_{0 \leq t \leq 1} f(\gamma(t)) \geq f(S)$ and for one path γ_0 we know $\sup_{0 \leq t \leq 1} f(\gamma_0(t)) = f(S)$. Thus, if we denote by Γ the set of all continuous paths γ from T_0 to T_1 we get

   $$f(S) = \inf_{\gamma \in \Gamma} \sup_{0 \leq t \leq 1} f(\gamma(t)),$$

   i.e., the saddle point S of f is determined by a 'minimax' principle.

4. *If $u \in E$ is a critical point of a differentiable functional f of the form (7.17) on a Banach space E, then this means that u satisfies $f'(u)(v) = 0$ for all $v \in E$. This means that u is a weak solution of the (nonlinear) differential equation $f'(u) = 0$. But in most cases we are actually interested in a strong solution of this equation, i.e., a solution which satisfies the equation $f'(u) = 0$ at least point-wise almost everywhere. For a classical solution this*

equation should be satisfied in the sense of functions of class $\mathcal{C}^2$. For the linear problems which we have discussed in some detail we have used the special form of the differential operator to argue that for these problems a weak solution is automatically a classical solution. The underlying theory is the theory of elliptic regularity. The basic results of this theory are presented in the books [BB92, JLJ98].

7.4 Generalized Dirichlet Forms and variational BVPs

Suppose we want to solve a quasi-linear differential equation in divergence form, i.e., an equation of the form

$$A(u)(x) \equiv A_0(x, u(x), Du(x)) - \sum_{j=1}^{n} \partial_j A_j(x, u(x), Du(x)) = f(x) \tag{7.29}$$

using variational methods over an open subset $G \subset \mathbb{R}^n$ under suitable growth restrictions on the coefficient functions A_j and the inhomogeneous term f. According to these growth restrictions we determine an exponent p and work in the Sobolev space $E \equiv W^{1,p}(G)$. Using the map $y : W^{1,p}(G) \longrightarrow L^p(G)^{n+1}$ defined by $y(u) = (u, \partial_1 u, \dots, \partial_n u)$ we introduce the **generalized Dirichlet form** on $W^{1,p}(G)$ by

$$a(u,v) = \langle A_0(\cdot, y(u)), v\rangle_2 + \sum_{j=1}^{n} \langle A_j(\cdot, y(u)), \partial_j v\rangle_2 \tag{7.30}$$

with the abbreviation

$$\langle f, g\rangle_2 = \int_G f(x) g(x) d^n x,$$

assuming real valued functions.

Under suitable growth restrictions on the coefficient functions one can establish an estimate of the form

$$|a(u,v)| \leq h(\|u\|_{1,p})\, \|v\|_{1,p} \quad \text{for all} \quad u, v \in E \tag{7.31}$$

with the norm

$$\|u\|_{1,p} = \left(\sum_{|\alpha| \leq 1} \int_G |D^\alpha u(x)|^p d^n x \right)^{1/p}$$

on E, for some bounded function $h : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$. If (7.31) holds, then for fixed $u \in E$, $v \longrightarrow a(u,v)$ is a continuous linear form on E which thus can be written as

$$\langle T(u), v \rangle = a(u,v) \quad \text{for all} \quad u, v \in E, \tag{7.32}$$

where T is a (nonlinear) operator $E \longrightarrow E'$ and where $\langle \cdot, \cdot \rangle$ indicates the canonical duality between E' and E.

Boundary conditions are introduced in an abstract way by using a closed subspace V of $E = W^{1,p}(G)$ satisfying

$$W_0^{1,p}(G) \subseteq V \subseteq W^{1,p}(G) \tag{7.33}$$

and the form (7.30) is restricted to V, equipped with its natural relative topology. Then the operator T maps from V into V'.

Definition 7.4.1 (Variational boundary value problem) *Given a generalized Dirichlet form a under boundary conditions expressed by a Banach space V satisfying (7.33) and an element $f \in V'$, determine all $u \in V$ which satisfy*

$$a(u,v) = \langle f, v \rangle \quad \textit{for all} \quad v \in V. \tag{7.34}$$

Next we formulate a set $(H_1), (H_2), (H_3)$ of hypotheses on the coefficient functions A_j under which the variational boundary value problem can be solved.

(H_1) *Growth restrictions*. The coefficient functions A_j are Carathéodory functions on $G \times \mathbb{R}^{n+1}$ having polynomials bounds in the variables y with exponents $\leq p - 1$ (for details see page 180 of [BB92]).

(H_2) *Monotonicity*. For almost all $x \in G$ and all $y, y' \in \mathbb{R}^{n+1}$ we have

$$\sum_{j=0}^{n} [A_j(x,y) - A_j(x,y')][y_j - y_j'] \geq 0. \tag{7.35}$$

(H_3) *Coerciveness*. There are a constant $\alpha > 0$ and there are non-negative functions $g_j \in L^{p/(p-r_j)}(G)$, $0 \leq r_j < p$, such that for almost all $x \in G$ and all $y \in \mathbb{R}^{n+1}$ we have

$$\sum_{j=0}^{n} A_j(x,y) y_j \geq \alpha \sum_{j=0}^{n} |y_j|^p - \sum_{j=0}^{n} g_j(x) |y_j|^{r_j}. \tag{7.36}$$

Theorem 7.4.2 (Solution of variational bvp's) *Let $G \subset \mathbb{R}^n$ be a nonempty open subset and V a closed subspace satisfying condition (7.33). Suppose that the coefficient functions A_j in (7.30) satisfy the hypotheses (H_i), $i = 1,2,3$. Then the variational boundary value problem (7.34) has a solution for every $f \in V'$. For fixed $f \in V'$ the family of all solutions is a closed convex and bounded subset of V.*

Proof. For details we refer to pages 182ff of [BB92]. Here we just give some comments. This result follows from the Browder-Minty Theorem and the proof consists in showing that the hypotheses (H_i), $i = 1,2,3$ imply the hypotheses of the Browder-Minty Theorem. The growth hypothesis (H_1) implies that $T : V \longrightarrow V'$ is a bounded continuous mapping. Hypothesis (H_2) implies that T is monotone while hypothesis (H_3) guarantees coerciveness of T.

7.5 Some global problems

The problems we discussed thus far relied in an essential way on at least one of the following assumptions:

1. The domain $\Omega \subset \mathbb{R}^d$ over which the problem is studied, is bounded.

2. The exponents which are used in the bounds for the nonlinear terms are strictly smaller than the *critical Sobolev exponent* p^* defined by

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{d} \quad \text{or} \quad \mathrm{p}^* = \frac{\mathrm{dp}}{\mathrm{d} - \mathrm{p}} \tag{7.37}$$

when one looks for solutions in the Sobolev space $W^{1,p}(\Omega)$, $1 \leq p < d$.

Recall that the reason for this restriction was that in the proofs we wanted to be able to use compact Sobolev embeddings. In this section we are discussing problems where the underlying domain Ω is not assumed to be bounded. Naturally this requires some new tools some of which will be presented here.

Depending on time which is available more details will be discussed.

7.6 Problems with critical growth

7.7 Problems not treated here

Bibliography

[Ada75] R.A. Adams. *Sobolev spaces*. Academic Press, New York, 1975.

[BB92] Ph. Blanchard and E. Brüning. *Variational Methods in Mathematical Physics. A unified approach.* Texts and Monographs in Physics. Springer-Verlag, Berlin Heidelberg New York London Paris Tokyo Hong Kong Barcelona Budapest, 1992.

[BZ88] Y. D. Burago and V. A. Zalgaller. *Geometric inequalities*, volume 285 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*. Springer-Verlag, Berlin, 1988.

[Dac82] B. Dacorogna. *Weak continuity and weak lower semicontinuity of non-linear functionals*, volume 922 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1982.

[Dac89] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied mathematical sciences*. Springer-Verlag, Berlin, 1989.

[Die69] J. A. Dieudonné. *Foundations of modern analysis*. Academic Press, New York, 1969.

[Gol80] H. H. Goldstine. *A history of the calculus of variations from the 17th through the 19th century*, volume 5 of *Studies in the history of mathematics and physical sciences*. Springer-Verlag, New York, 1980.

[JLJ98] J. Jost and X. Li-Jost. *Calculus of variations*, volume 64 of *Cambridge Studies in advanced mathematics*. Cambridge University Press, 1998.

[Kol31] A. N. Kolmorov. über kompaktheit der funktionenmengen bei der konvergenz im mittel. *Nachr. Ges. Wiss. Göttingen*, 9:60–63, 1931.

[Lju34] L. A. Ljusternik. On conditional extrema of functions. *Mat. Sbornik*, 41(3), 1934.

[LL01] E. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. AMS, 2nd edition, 2001.

[LT91] M. Ledoux and M. Talagrand. *Probability in Banach spaces : isoperimetry and processes*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete ; Folge 3*. Springer-Verlag, Berlin, 1991.

[Oss78] R. Osserman. The isoperimetric inequality. *Bull. Amer. Math. Soc.*, 84:1182–1238, 1978.

[Rie33] M. Riesz. Sur les ensembles compacts de fonctions sommable. *Acta Szeged Sect. Math.*, 6:136–142, 1933.

[RR73] A. P. Robertson and W. J. Robertson. *Topological vector spaces*. Cambridge University Press, Cambridge, 1973.

[Str00] M. Struwe. *Variational methods : applications to nonlinear partial differential equations and Hamiltonian systems*, volume 34 of *Ergebnisse der Mathematik und ihrer Grenzgebiete, Folge 3*. Springer-Verlag, Berlin, 3rd edition, 2000.

[Vai64] M. M. Vainberg. *Variational methods for the study of nonlinear operators*. Holden Day, London, 1964.

[YCB82] C. Dewitt-Morette with M. Dillard-Bleick Y. Choquet-Bruhat. *Analysis, manifolds and physics*. North Holland Publ. Comp., 1982.

[Zei85] E. Zeidler. *Variational methods and optimization*, volume 3 of *Nonlinear functional analysis and its applications*. Springer-Verlag, New York, 1985.