Contents

Part I Basic Course

1 Introduction ........................................... 3
   1.1 The brachistochrone problem ...................... 4
   1.2 The isoperimetric problem .......................... 6
   1.3 Electrostatics .......................................... 7
   1.4 Stationary states in quantum mechanics .......... 8
   1.5 Optimal saving and consumption .................. 9
   1.6 Sailing against the wind ......................... 10
   1.7 Hyperelasticity ......................................... 12
   1.8 Microstructure in crystals .......................... 14
   1.9 Phase transitions ...................................... 17
   1.10 Composite elastic materials ...................... 20

2 Convexity .................................................. 23
   2.1 The Direct Method ....................................... 24
   2.2 Functionals with convex integrands ............... 26
   2.3 Integrands with \( u \)-dependence .................. 31
   2.4 The Lavrentiev gap phenomenon ................... 33
   2.5 Integral side constraints ............................. 36
   2.6 The general theory of convex functions and duality 37
   Notes and historical remarks ............................ 42
   Problems .................................................. 43

3 Variations .................................................. 45
   3.1 The Euler–Lagrange equation ....................... 46
   3.2 Regularity of minimizers ............................. 53
   3.3 Lagrange multipliers ................................... 62
   3.4 Invariances and Noether’s theorem ............... 65
   3.5 Subdifferentials ........................................ 70
   Notes and historical remarks ............................ 73
4 Young Measures ................................................................. 77
  4.1 The Fundamental Theorem .............................................. 78
  4.2 Examples ........................................................................ 87
  4.3 Young measures and notions of convergence ...................... 89
  4.4 Gradient Young measures .............................................. 90
  4.5 Homogeneous gradient Young measures ......................... 93
  Notes and historical remarks .............................................. 95
  Problems ............................................................................ 97

5 Quasiconvexity .................................................................... 99
  5.1 Quasiconvexity ............................................................... 100
  5.2 Null-Lagrangians ............................................................ 107
  5.3 A Jensen-type inequality for gradient Young measures ......... 111
  5.4 Rigidity for gradients ...................................................... 112
  5.5 Lower semicontinuity ...................................................... 115
  5.6 Integrands with $u$-dependence ........................................ 119
  5.7 Regularity of minimizers ............................................... 121
  Notes and historical remarks .............................................. 122
  Problems ............................................................................ 123

6 Polyconvexity .................................................................... 127
  6.1 Polyconvexity ............................................................... 128
  6.2 Existence of minimizers ................................................. 129
  6.3 Global injectivity ........................................................... 137
  Notes and historical remarks .............................................. 139
  Problems ............................................................................ 140

7 Relaxation ........................................................................... 143
  7.1 Quasiconvex envelopes ..................................................... 144
  7.2 Relaxation of integral functionals ...................................... 148
  7.3 Generalized convexity notions and envelopes ..................... 152
  7.4 Young measure relaxation ............................................. 156
  7.5 Characterization of gradient Young measures ................... 160
  Notes and historical remarks .............................................. 169
  Problems ............................................................................ 169

Part II Advanced Topics

8 Rigidity .............................................................................. 173
  8.1 Two-gradient inclusions ................................................... 174
  8.2 Linear inclusions ............................................................ 176
  8.3 Relaxation and quasiconvex hulls of sets ......................... 180
  8.4 Multi-point inclusions ...................................................... 183
  8.5 The one-well inclusion .................................................... 194
### 1.3 \( \Gamma \)-Convergence

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1 Abstract ( \Gamma )-convergence</td>
<td>348</td>
</tr>
<tr>
<td>13.2 Sharp-interface limits</td>
<td>353</td>
</tr>
<tr>
<td>13.3 Higher-order sharp-interface limits</td>
<td>363</td>
</tr>
<tr>
<td>13.4 Periodic homogenization</td>
<td>365</td>
</tr>
<tr>
<td>13.5 Convex homogenization</td>
<td>374</td>
</tr>
<tr>
<td>13.6 Quadratic homogenization</td>
<td>376</td>
</tr>
</tbody>
</table>

### A Prerequisites

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1 Linear algebra</td>
<td>385</td>
</tr>
<tr>
<td>A.2 Functional analysis</td>
<td>388</td>
</tr>
<tr>
<td>A.3 Measure theory</td>
<td>389</td>
</tr>
<tr>
<td>A.4 Vector measures</td>
<td>394</td>
</tr>
<tr>
<td>A.5 Sobolev and other function spaces</td>
<td>395</td>
</tr>
<tr>
<td>A.6 Harmonic analysis</td>
<td>400</td>
</tr>
</tbody>
</table>

### References

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>403</td>
</tr>
</tbody>
</table>

### Index

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>415</td>
</tr>
</tbody>
</table>
Chapter 2
Convexity

In this chapter we start to develop the mathematical theory that will allow us to analyze the problems presented in the introduction, and many more. The basic minimization problem that we are considering is the following:

\[
\begin{cases}
\text{Minimize} & \mathcal{F}[u] := \int_{\Omega} f(x, u(x), \nabla u(x)) \, dx \\
\text{over all} & u \in W^{1,p}(\Omega; \mathbb{R}^m) \text{ with } u|_{\partial \Omega} = g.
\end{cases}
\]

Here, and throughout the text if not stated otherwise, we will make the standard assumption that \( \Omega \subset \mathbb{R}^d \) is a bounded Lipschitz domain, that is, \( \Omega \) is open, bounded, connected, and has a boundary that is the union of finitely many Lipschitz manifolds. The function \( f: \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R} \) is required to be measurable in the first and (jointly) continuous in the second and third arguments, which makes \( f \) a so-called Carathéodory integrand. Furthermore, in this chapter we (usually) let \( p \in (1, \infty) \) and for the prescribed boundary values \( g \) we assume

\[ g \in W^{1-1/p, p}(\partial \Omega; \mathbb{R}^m). \]

In this context recall that \( W^{1-1/p, p}(\partial \Omega; \mathbb{R}^m) \) is the space of traces of Sobolev maps in \( W^{1,p}(\Omega; \mathbb{R}^m) \), see Appendix A.5 for some background on Sobolev spaces.

Below, we will investigate the solvability of the above minimization problem (under additional technical assumptions). We first present the main ideas of the so-called Direct Method of the calculus of variations in an abstract setting, namely for (nonlinear) functionals on Banach spaces. Then we will begin our study of integral functionals, where we will in particular take a close look at the way in which convexity properties of \( f \) in its gradient (third) argument determine whether \( \mathcal{F} \) is lower semicontinuous. We also consider the question of which function space should be chosen for the candidate functions. Finally, we explain basic aspects of general convex analysis, in particular the Legendre–Fenchel duality.
2.1 The Direct Method

Fundamental to all of the existence theorems in this book is the conceptually simple, yet powerful, Direct Method of the calculus of variations. It is called “direct” since we prove the existence of solutions to minimization problems without the detour through a differential equation.

Let \( X \) be a complete metric space (e.g. a Banach space with the norm topology or a closed and convex subset of a reflexive Banach space with the weak topology). Let \( \mathcal{F} : X \to \mathbb{R} \cup \{+\infty\} \) be our objective functional that we require to satisfy the following two assumptions:

(H1) Coercivity: For all \( \Lambda \in \mathbb{R} \), the sublevel set
\[
\{ u \in X : \mathcal{F}[u] \leq \Lambda \}
\]
is sequentially precompact, that is, if \( \mathcal{F}[u_j] \leq \Lambda \) for a sequence \( (u_j) \subset X \) and some \( \Lambda \in \mathbb{R} \), then \( (u_j) \) has a converging subsequence in \( X \).

(H2) Lower semicontinuity: For all sequences \( (u_j) \subset X \) with \( u_j \to u \) in \( X \) it holds that
\[
\mathcal{F}[u] \leq \liminf_{j \to \infty} \mathcal{F}[u_j].
\]

Note that here and in all of the following we use the sequential notions of compactness and lower semicontinuity, which are better suited to our needs than the corresponding topological concepts. For more on this point see the notes section at the end of this chapter.

The Direct Method for the abstract problem

\[
\text{Minimize } \mathcal{F}[u] \text{ over all } u \in X
\]
is encapsulated in the following simple result.

**Theorem 2.1.** Assume that \( \mathcal{F} \) is both coercive and lower semicontinuous. Then, the abstract minimization problem (2.1) has at least one solution, that is, there exists a \( u_* \in X \) with \( \mathcal{F}[u_*] = \min \{ \mathcal{F}[u] : u \in X \} \).

**Proof.** Let us assume that there exists at least one \( u \in X \) such that \( \mathcal{F}[u] < +\infty \); otherwise, any \( u \in X \) is a “solution” to the (degenerate) minimization problem.

To construct a minimizer we take a minimizing sequence \( (u_j) \subset X \) such that
\[
\lim_{j \to \infty} \mathcal{F}[u_j] \to \alpha := \inf \{ \mathcal{F}[u] : u \in X \} < +\infty.
\]

Then, there exists a \( \Lambda \in \mathbb{R} \) such that \( \mathcal{F}[u_j] \leq \Lambda \) for all \( j \in \mathbb{N} \). Hence, by the coercivity, we may select a subsequence, which we do not make explicit in our notation, such that
\[
u_j \to u_* \in X.
\]

By the lower semicontinuity we immediately conclude that
Thus, $\mathcal{F}[u] = \alpha$ and $u_*$ is the sought minimizer.

**Example 2.2.** Using the Direct Method, one can easily see that the lower semicontinuous function

$$h(t) := \begin{cases} 1 - t & \text{if } t < 0, \\ t & \text{if } t \geq 0, \end{cases}$$

has the minimizer $t = 0$.

Despite its nearly trivial proof, the Direct Method is very useful and flexible in applications. Indeed, it pushes the difficulty in proving the existence of a minimizer into establishing coercivity and lower semicontinuity. This, however, is a big advantage, since we have many tools at our disposal to establish these two hypotheses separately. In particular, for integral functionals, lower semicontinuity is tightly linked to convexity properties of the integrand, as we will see throughout this book.

At this point it is crucial to observe how coercivity and lower semicontinuity interact with the topology on $X$: If we choose a stronger topology, i.e., one for which there are fewer converging sequences, then it is easier for $\mathcal{F}$ to be lower semicontinuous, but harder for $\mathcal{F}$ to be coercive. The opposite holds if we choose a weaker topology. In the mathematical treatment of a problem from applications, we are most likely in a situation where $\mathcal{F}$ and the set $X$ are given. We then need to find a suitable topology in which we can establish both coercivity and lower semicontinuity. It is remarkable that the topology that turns out to be mathematically convenient is often also physically relevant.

In this book, $X$ will always be an infinite-dimensional Banach space (or a subset thereof) and we have a real choice between using the strong or weak convergence. Usually, it turns out that coercivity with respect to the strong convergence is false since strongly compact sets in infinite-dimensional spaces are very restricted, whereas coercivity with respect to the weak convergence is true under reasonable assumptions. On the other hand, while strong lower semicontinuity poses few challenges, lower semicontinuity with respect to weakly converging sequences is a more delicate matter and we will spend considerable time on this topic.

As a result of this discussion, we will almost always use the Direct Method in the following version:

**Theorem 2.3.** Let $X$ be a reflexive Banach space or a closed affine subset of a reflexive Banach space and let $\mathcal{F} : X \to \mathbb{R} \cup \{+\infty\}$. Assume the following:

(\textbf{WH1}) \textbf{Weak coercivity:} For all $\Lambda \in \mathbb{R}$ the sublevel set

$$\{ u \in X : \mathcal{F}[u] \leq \Lambda \}$$

is sequentially weakly precompact, that is, if $\mathcal{F}[u_j] \leq \Lambda$ for a sequence $(u_j) \subset X$ and some $\Lambda \in \mathbb{R}$, then $(u_j)$ has a weakly converging subsequence.
(WH2) **Weak lower semicontinuity:** For all sequences \((u_j) \subset X\) with \(u_j \rightharpoonup u\) in \(X\) (weak convergence) it holds that
\[
F[u] \leq \liminf_{j \to \infty} F[u_j].
\]
Then, the problem
Minimize \(F[u]\) over all \(u \in X\)
has at least one solution.

The proof of this theorem is analogous to the proof of Theorem 2.1, also taking into account the fact that all (strongly) closed affine subsets of a Banach space are weakly closed.

### 2.2 Functionals with convex integrands

As a first instance of the theory of integral functionals to be developed in this book, we now consider the minimization problem for
\[
F[u] := \int_{\Omega} f(x, \nabla u(x)) \, dx
\]
over all \(u \in W^{1,p}(\Omega; \mathbb{R}^m)\), where \(\Omega \subset \mathbb{R}^d\) is a bounded Lipschitz domain and \(p \in (1, \infty)\) will be chosen later (depending on growth properties of \(f\)). The reader is referred to Appendix A.5 for an overview of Sobolev spaces.

The following lemma shows that the integrand is measurable if \(f\) is a so-called Carathéodory integrand, which from now on we assume.

**Lemma 2.4.** Let \(f: \Omega \times \mathbb{R}^N \to \mathbb{R}\) be a Carathéodory integrand, that is,

1. \(x \mapsto f(x,A)\) is Lebesgue-measurable for every fixed \(A \in \mathbb{R}^N\);
2. \(A \mapsto f(x,A)\) is continuous for (Lebesgue-)almost every fixed \(x \in \Omega\).

Then, for any Borel-measurable map \(V: \Omega \to \mathbb{R}^N\) the composition \(x \mapsto f(x,V(x))\) is Lebesgue-measurable.

**Proof.** Assume first that \(V\) is a simple function,
\[
V = \sum_{k=1}^m v_k 1_{E_k},
\]
where the sets \(E_k \subset \Omega\) are Borel-measurable \((k \in \{1,\ldots,m\})\), \(\bigcup_{k=1}^m E_k = \Omega\), and \(v_k \in \mathbb{R}^N\). For \(t \in \mathbb{R}\) we have
\[
\{ x \in \Omega : f(x,V(x)) > t \} = \bigcup_{k=1}^m \{ x \in E_k : f(x,v_k) > t \},
\]
2.2 Functionals with convex integrands

which is a Lebesgue-measurable set by assumption. Hence, \( x \mapsto f(x, V(x)) \) is Lebesgue-measurable.

Turning to the general case, every Borel-measurable function \( V \) can be approximated by simple functions \( V_k \) with

\[
f(x, V(x)) \to f(x, V_k(x)) \quad \text{for all } x \in \Omega \quad \text{as } k \to \infty,
\]

see Lemma A.5. We conclude that the right-hand side is Lebesgue-measurable as the pointwise limit of Lebesgue-measurable functions. \( \square \)

It is possible that the (compound) integrand in \( \mathcal{F} \) is measurable, but that the integral is not well-defined. These pathological cases can, for example, be avoided if \( f \geq 0 \) or if one imposes the \( p \)-growth bound

\[
|f(x,A)| \leq M(1 + |A|^p), \quad (x,A) \in \Omega \times \mathbb{R}^{m \times d},
\]

for some \( M > 0 \), which implies the finiteness of \( \mathcal{F}[u] \) for all \( u \in W^{1,p}(\Omega; \mathbb{R}^m) \). In this chapter, however, this bound is not otherwise needed.

We next investigate the coercivity of \( \mathcal{F} \). If \( p \in (1, \infty) \), then the most basic assumption to guarantee coercivity, and the only one we consider here, is the \( p \)-coercivity bound

\[
\mu |A|^p \leq f(x,A), \quad (x,A) \in \Omega \times \mathbb{R}^{m \times d}, \tag{2.2}
\]

for some \( \mu > 0 \). This coercivity also determines the exponent \( p \) for the Sobolev space where we look for solutions. Note that in the literature sometimes the coercivity bound is given as the seemingly more general \( \mu |A|^p - C \leq f(x,A) \) for some \( \mu, C > 0 \). This, however, does not increase generality since we may pass from the integrand \( f(x,A) \) to the integrand \( \tilde{f}(x,A) := f(x,A) + C \), which now satisfies (2.2), without changing the minimization problem (recall that \( \Omega \) is assumed bounded throughout this book).

**Proposition 2.5.** If the Carathéodory integrand \( f : \Omega \times \mathbb{R}^{m \times d} \to [0, \infty) \) satisfies the \( p \)-coercivity bound (2.2) with \( p \in (1, \infty) \), then \( \mathcal{F} \) is weakly coercive on the space

\[
W^{1,p}_g(\Omega; \mathbb{R}^m) = \{ u \in W^{1,p}(\Omega; \mathbb{R}^m) : u|_{\partial \Omega} = g \},
\]

where \( g \in W^{1-1/p \cdot p}(\partial \Omega; \mathbb{R}^m) \).

**Proof.** We need to show that any sequence \( (u_j) \subset W^{1,p}_g(\Omega; \mathbb{R}^m) \) with

\[
\sup_{j \in \mathbb{N}} \mathcal{F}[u_j] < \infty
\]

is weakly precompact. From (2.2) we get

\[
\mu \cdot \sup_{j \in \mathbb{N}} \int_{\Omega} |\nabla u_j|^p \ dx \leq \sup_{j \in \mathbb{N}} \mathcal{F}[u_j] < \infty,
\]
whereby $\sup_j \|\nabla u_j\|_{L^p} < \infty$. Fix $u_0 \in W^{1,p}_0(\Omega; \mathbb{R}^m)$. Then, $u_j - u_0 \in W^{1,p}_0(\Omega; \mathbb{R}^m)$ and $\sup_j \|\nabla (u_j - u_0)\|_{L^p} < \infty$. From the Poincaré inequality, see Theorem A.26 (i), we therefore get

$$\sup_j \|u_j\|_{W^{1,p}} \leq \sup_j \|u_j - u_0\|_{W^{1,p}} + \|u_0\|_{W^{1,p}} < \infty.$$ 

This finishes the proof since bounded sets in separable and reflexive Banach spaces, like $W^{1,p}(\Omega; \mathbb{R}^m)$ for $p \in (1, \infty)$, are sequentially weakly precompact by Theorem A.2.

Having settled the question of weak coercivity, we can now investigate the weak lower semicontinuity. The following pivotal result (in the one-dimensional case) goes back to the work of Leonida Tonelli in the early 20th century; the generalization to higher dimensions is due to James Serrin.

**Theorem 2.6** (Tonelli 1920 & Serrin 1961 [242, 276]). Let $f : \Omega \times \mathbb{R}^{m \times d} \to [0, \infty)$ be a Carathéodory integrand such that $f(x, \cdot)$ is convex for almost every $x \in \Omega$.

Then, $\mathcal{F}$ is weakly lower semicontinuous on $W^{1,p}(\Omega; \mathbb{R}^m)$ for any $p \in (1, \infty)$.

**Proof.** 

**Step 1.** We first establish that $\mathcal{F}$ is strongly lower semicontinous, so let $u_j \rightharpoonup u$ in $W^{1,p}(\Omega; \mathbb{R}^m)$ and $\nabla u_j \to \nabla u$ almost everywhere, which holds after selecting a subsequence (not explicitly labeled), see Appendix A.3. By assumption we have that $f(x, \nabla u_j(x)) \geq 0$. Applying Fatou’s Lemma, we immediately conclude that

$$\mathcal{F}[u] = \int_\Omega f(x, \nabla u(x)) \, dx \leq \liminf_{j \to \infty} \int_\Omega f(x, \nabla u_j(x)) \, dx = \liminf_{j \to \infty} \mathcal{F}[u_j].$$

Since this holds for all subsequences, it also follows for our original sequence, see Problem □.

**Step 2.** To prove the claimed weak lower semicontinuity take $(u_j) \subset W^{1,p}(\Omega; \mathbb{R}^m)$ with $u_j \to u$ in $W^{1,p}$. We need to show that

$$\mathcal{F}[u] \leq \liminf_{j \to \infty} \mathcal{F}[u_j] =: \alpha. \quad (2.3)$$

Taking a subsequence (not explicitly labeled), we can in fact assume that $\mathcal{F}[u_j]$ converges to $\alpha$.

By the Mazur Lemma A.4 we may find convex combinations

$$v_j = \sum_{n=j}^{N(j)} \theta_n^{(j)} u_n,$$

where $\theta_n^{(j)} \in [0, 1]$ and $\sum_{n=j}^{N(j)} \theta_n^{(j)} = 1$,

such that $v_j \to u$ in $W^{1,p}$. As $f(x, \cdot)$ is convex for almost every $x$,
2.2 Functionals with convex integrands

\[ F[v_j] = \int_\Omega f \left( x, \sum_{n=J}^{N} \theta_n^{(j)} \nabla u_n(x) \right) \, dx \leq \sum_{n=J}^{N} \theta_n^{(j)} F[u_n]. \]

Since \( F[u_n] \to \alpha \) as \( n \to \infty \) and \( \sum_{n=J}^{N} \theta_n^{(j)} = 1 \), we arrive at

\[ \liminf_{j \to \infty} F[v_j] \leq \alpha. \]

On the other hand, from the first step and since \( v_j \to u \) strongly, we have 

\[ F[u] \leq \liminf_{j \to \infty} F[v_j]. \]

Thus, (2.3) follows and the proof is finished. \( \square \)

We can summarize our findings in the following existence theorem.

**Theorem 2.7.** Let \( f : \Omega \times \mathbb{R}^{m \times d} \to [0, \infty) \) be a Carathéodory integrand such that

(i) \( f \) satisfies the \( p \)-coercivity bound (2.2) with \( p \in (1, \infty) \);

(ii) \( f(x, \cdot) \) is convex for almost every \( x \in \Omega \).

Then, the associated functional \( F \) has a minimizer over \( W^{1, p}(\Omega; \mathbb{R}^m) \), where \( g \in W^{1-1/p, p}(\partial \Omega; \mathbb{R}^m) \).

**Proof.** This follows immediately from the Direct Method for the weak convergence, Theorem 2.3 with \( X = W^{1, p}(\Omega; \mathbb{R}^m) \) together with Proposition 2.5 and the Tonelli–Serrin Theorem 2.6. \( \square \)

**Example 2.8.** The Dirichlet functional (or Dirichlet integral) is 

\[ F[u] := \int_{\Omega} \frac{1}{2} |\nabla u(x)|^2 \, dx, \quad u \in W^{1, 2}(\Omega; \mathbb{R}^m). \]

We already encountered this integral functional when considering electrostatics in Section 1.3. It is easy to see that the Dirichlet functional satisfies all requirements of Theorem 2.7 and so there exists a minimizer for any prescribed boundary values \( g \in W^{1/2, 2}(\partial \Omega; \mathbb{R}^m) \).

We next show the following converse to the Tonelli–Serrin Theorem 2.6.

**Proposition 2.9.** Let \( F : W^{1, p}(\Omega; \mathbb{R}^m) \to \mathbb{R}, \ p \in [1, \infty), \) be an integral functional with continuous integrand \( f : \mathbb{R}^{m \times d} \to \mathbb{R} \) (not \( x \)-dependent). If \( F \) is weakly lower semicontinuous on \( W^{1, p}(\Omega; \mathbb{R}^m) \) and if either \( m = 1 \) or \( d = 1 \) (the scalar case and the one-dimensional case, respectively), then \( f \) is convex.

**Proof.** We only consider the case \( m = 1 \) and \( d \) arbitrary; the other case is proved in a similar manner. Assume that \( a, b \in \mathbb{R}^d \) with \( a \neq b \) and \( \theta \in (0, 1) \). Let \( v := \theta a + (1 - \theta)b, \ n := b - a, \) and set
The function
\[ \phi_0(x) := v \cdot x + \frac{1}{j} \phi_0(jx \cdot n - \lfloor jx \cdot n \rfloor), \quad x \in \Omega, \]
where \( \lfloor s \rfloor \) denotes the largest integer less than or equal to \( s \in \mathbb{R} \), and
\[ \phi_0(t) := \begin{cases} 
(1 - \theta)t & \text{if } t \in [0, \theta), \\
\theta t - \theta & \text{if } t \in [\theta, 1) .
\end{cases} \]
see Figure 2.1. We have that
\[ \nabla u_j(x) = \begin{cases} 
\theta a + (1 - \theta)b - (1 - \theta)(b - a) = a & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in [0, \theta), \\
\theta a + (1 - \theta)b + \theta(b - a) = b & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in [\theta, 1) .
\end{cases} \]
Hence, \( u_j \subset W^{1,\infty}(\Omega) \) and since the second term in the definition of \( u_j \) converges to zero uniformly, it holds that \( u_j \rightharpoonup v \cdot x \) in \( W^{1,p}(\Omega) \) (here and in the following, \( "v \cdot x" \) is a shorthand notation for the linear function \( x \mapsto v \cdot x \)). By the weak lower semi-continuity, we conclude that
\[ |\Omega| \cdot f(v) = \mathcal{F}[v \cdot x] \leq \liminf_{j \to \infty} \mathcal{F}[u_j] = |\Omega| \cdot (\theta f(a) + (1 - \theta)f(b)) . \]
This proves the claim. \( \square \)

In the vectorial case, i.e., \( m \neq 1 \) and \( d \neq 1 \), it turns out that convexity of the integrand (in the gradient variable) is far from being necessary for weak lower semicontinuity. In fact, there is indeed a weaker condition ensuring weak lower semicontinuity; we will explore this in Chapter 5.

Finally, we prove the following result concerning the uniqueness of the minimizer.

**Proposition 2.10.** Let \( \mathcal{F} : W^{1,p}(\Omega; \mathbb{R}^m) \to \mathbb{R} \), \( p \in [1, \infty) \), be an integral functional with Carathéodory integrand \( f : \Omega \times \mathbb{R}^{m \times d} \to \mathbb{R} \). If \( f \) is strictly convex, that is,
\[ f(x, \theta A + (1 - \theta)B) < \theta f(x, A) + (1 - \theta)f(x, B) \]
for every \( x \in \Omega \), \( \theta \in [0, 1] \), and \( A, B \in \mathbb{R}^{m \times d} \), then \( \mathcal{F} \) is lower semi-continuous and satisfies the direct method in weak convergence. Moreover, any minimizing sequence \( u_j \subset W^{1,p}(\Omega; \mathbb{R}^m) \) is a strongly convergent sequence in \( W^{1,p}(\Omega) \) and \( \mathcal{F}[u_j] \to \mathcal{F}[u] \) as \( j \to \infty \), where \( u \) is the unique minimizer of \( \mathcal{F} \) on \( \Omega \).
for all \( x \in \Omega, \ A, B \in \mathbb{R}^{m \times d} \) with \( A \neq B, \ \theta \in (0,1) \), then the minimizer \( u_* \in W^{1,p}_g(\Omega;\mathbb{R}^m) \) (\( g \in W^{-1/p,p}(\partial \Omega;\mathbb{R}^m) \)) of \( \mathcal{F} \), if it exists, is unique.

**Proof.** Assume there are two different minimizers \( u, v \in W^{1,p}_g(\Omega;\mathbb{R}^m) \) of \( \mathcal{F} \). Then set
\[
w := \frac{1}{2} u + \frac{1}{2} v \in W^{1,p}_g(\Omega;\mathbb{R}^m)
\]
and observe that
\[
\mathcal{F}[w] = \int_{\Omega} f\left(x, \frac{1}{2} \nabla u(x) + \frac{1}{2} \nabla v(x)\right) < \frac{1}{2} \mathcal{F}[u] + \frac{1}{2} \mathcal{F}[v] = \min_{W^{1,p}_g(\Omega;\mathbb{R}^m)} \mathcal{F},
\]
yielding an immediate contradiction. \( \square \)

2.3 Integrands with \( u \)-dependence

If we try to extend the results in the previous section to more general functionals
\[
\mathcal{F}[u] := \int_{\Omega} f(x, u(x), \nabla u(x)) \, dx,
\]
we discover that our proof strategy via the Mazur lemma runs into difficulties: We cannot “pull out” the convex combination inside
\[
\int_{\Omega} f\left(x, \sum_{n=1}^{N} \theta_n^{(j)} u_n(x), \sum_{n=1}^{N} \theta_n^{(j)} \nabla u_n(x)\right) \, dx
\]
any more. Nevertheless, a lower semicontinuity result analogous to the one for the \( u \)-independent case turns out to be true:

**Theorem 2.11.** Let \( f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to [0,\infty) \) be a Carathéodory integrand, which here means that

(i) \( x \mapsto f(x, v, A) \) is Lebesgue-measurable for every fixed \( (v, A) \in \mathbb{R}^m \times \mathbb{R}^{m \times d} \);

(ii) \( (v, A) \mapsto f(x, v, A) \) is continuous for (Lebesgue-)almost every fixed \( x \in \Omega \).

Assume also that
\[
f(x, v, \cdot) \text{ is convex for every } (x, v) \in \Omega \times \mathbb{R}^m.
\]

Then, for \( p \in (1,\infty) \), the functional
\[
\mathcal{F}[u] := \int_{\Omega} f(x, u(x), \nabla u(x)) \, dx, \quad u \in W^{1,p}(\Omega;\mathbb{R}^m),
\]
is weakly lower semicontinuous.
While it would be possible to give an elementary proof of this theorem here, we postpone the detailed study of integral functionals with \( u \)-dependent integrands until Section 5.6. There, using more advanced techniques, we will establish a much more general lower semicontinuity result, albeit under an additional \( p \)-growth assumption \( |f(x,v,A)| \leq M(1 + |v|^p + |A|^p) \). A proof of the above theorem without this growth assumption can be found in Section 3.2.6 of [76].

**Example 2.12.** In the prototypical problem of linearized elasticity from Section 1.7 we are tasked to solve

\[
\begin{align*}
\text{Minimize} \quad & \mathcal{F}[u] := \frac{1}{2} \int_{\Omega} 2\mu |\varepsilon u|^2 + \left( \kappa - \frac{2}{3} \mu \right) |\text{tr} \varepsilon u|^2 - b \cdot u \, dx \\
\text{over all} \quad & u \in W^{1,2}(\Omega;\mathbb{R}^3) \text{ with } u|_{\partial \Omega} = g,
\end{align*}
\]

where \( \mu, \kappa > 0 \), \( b \in L^2(\Omega;\mathbb{R}^3) \), and \( g \in W^{1/2,2}(\partial \Omega;\mathbb{R}^m) \). It is clear that \( \mathcal{F} \) has quadratic growth. We assume that \( \kappa - 2\mu/3 \geq 0 \) and \( g = 0 \) for simplicity. Then, we first show that

\[
\|\nabla u\|_{L^2} \leq \sqrt{2} \|\varepsilon u\|_{L^2} \tag{2.4}
\]

for all \( u \in W^{1,2}(\Omega;\mathbb{R}^3) \) with \( u|_{\partial \Omega} = 0 \). This can be seen as follows: An elementary computation shows that for \( \varphi \in C_c^\infty(\Omega;\mathbb{R}^3) \) it holds that

\[
2(\varepsilon \varphi : \varepsilon \varphi) - \nabla \varphi : \nabla \varphi = \text{div} \left[ (\nabla \varphi) \varphi - (\text{div} \varphi) \varphi \right] + (\text{div} \varphi)^2.
\]

Thus, by the divergence theorem,

\[
\begin{align*}
2\|\varepsilon \varphi\|_{L^2}^2 - \|\nabla \varphi\|_{L^2}^2 &= \int_{\Omega} \text{div} \left[ (\nabla \varphi) \varphi - (\text{div} \varphi) \varphi \right] \, dx + \int_{\Omega} (\text{div} \varphi)^2 \, dx \\
&= \int_{\Omega} (\text{div} \varphi)^2 \, dx \\
&\geq 0.
\end{align*}
\]

This is (2.4) for \( \varphi \). The general case follows from the density of \( C_c^\infty(\Omega;\mathbb{R}^3) \) in \( W^{1,2}_0(\Omega;\mathbb{R}^3) \). Then, using Young’s inequality and the Poincaré inequality (see Theorem A.26 (i), we denote the \( L^2 \)-Poincaré constant by \( C_\varphi > 0 \), we get for any \( \delta > 0 \),

\[
\begin{align*}
\mathcal{F}[u] &\geq \mu \|\varepsilon u\|_{L^2}^2 - \|b\|_{L^2} \|u\|_{L^2} \\
&\geq \mu \|\varepsilon u\|_{L^2}^2 - \frac{1}{2\delta} \|b\|_{L^2}^2 - \frac{\delta}{2} \|u\|_{L^2}^2 \\
&\geq \frac{\mu}{2} \|\nabla u\|_{L^2}^2 - \frac{1}{2\delta} \|b\|_{L^2}^2 - \frac{C_\varphi^2 \delta}{2} \|\nabla u\|_{L^2}^2.
\end{align*}
\]

Choosing \( \delta = \mu/(2C_\varphi^2) \), we obtain the coercivity estimate

\[
\mathcal{F}[u] \geq \frac{\mu}{4} \|\nabla u\|_{L^2}^2 - \frac{C_\varphi^2 \mu}{\mu} \|b\|_{L^2}^2.
\]
2.4 The Lavrentiev gap phenomenon

We have chosen the function space in which we look for the solution of a minimization problem from the scale of Sobolev spaces according to a coercivity assumption such as (2.2). However, at first sight, classically differentiable functions may appear to be more appealing. So the question arises whether the infimum value is actually the same when considering different function spaces. Formally, given two linear or affine spaces $X$ and $Y$ such that $X$ is dense in $Y$, and a functional $F: Y \to \mathbb{R}$, we ask whether

$$\inf_X F = \inf_Y F.$$ 

Note that even if the infima agree, it is a priori unlikely that this infimum is attained in both spaces unless we have additional regularity of a minimizer (which we will investigate in Section 3.2).

For $X = C^\infty$ and $Y = W^{1,p}$ the equality of infima turns out to be true under suitable growth conditions:

**Theorem 2.13.** Let $f: \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$ be a Carathéodory integrand with $p$-growth, i.e.,

$$|f(x,v,A)| \leq M(1 + |v|^p + |A|^p), \quad (x,v,A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d},$$

for some $M > 0$, $p \in [1, \infty)$. Then, the functional

$$\mathcal{F}[u] := \int_\Omega f(x,u(x),\nabla u(x)) \, dx, \quad u \in W^{1,p}(\Omega;\mathbb{R}^m),$$

is strongly continuous. Consequently,

$$\inf_{W^{1,p}(\Omega;\mathbb{R}^m)} \mathcal{F} = \inf_{C^\infty(\Omega;\mathbb{R}^m)} \mathcal{F}.$$

The same equality of infima also holds with fixed boundary values.

**Proof.** Let $u_j \to u$ in $W^{1,p}(\Omega;\mathbb{R}^m)$ and additionally assume that $u_j \to u$, $\nabla u_j \to \nabla u$ almost everywhere (which holds after selecting a subsequence). Then, from the $p$-growth assumption we get

Hence, applying the Poincaré inequality again, $F[u]$ controls $\|u\|_{W^{1,2}}$ and our functional is weakly coercive. Moreover, it is clear that the integrand is convex in the $\varepsilon u$-argument. Hence, Theorem 2.11 yields the existence of a solution $u_\varepsilon \in W^{1,2}(\Omega;\mathbb{R}^3)$ to our minimization problem of linearized elasticity. In fact, one could also argue using the Tonelli–Serrin Theorem 2.6 and the elementary fact that the lower-order term $\int_\Omega b(x) \cdot u(x) \, dx$ is weakly continuous on $W^{1,2}$. More on the topic of linearized elasticity can be found in Sections 6.2 and 6.3 of [64].
\[ F[u_j] = \int_{\Omega} f(x, u_j, \nabla u_j) \, dx \leq \int_{\Omega} M(1 + |u_j|^p + |\nabla u_j|^p) \, dx \]

and via Pratt’s Theorem A.10 we infer that

\[ F[u_j] \to F[u]. \]

Since this holds for a subsequence of any subsequence of the original sequence \((u_j)\), we have established the continuity of \(F\) with respect to the strong convergence in \(W^{1,p}(\Omega; \mathbb{R}^m)\).

The assertion about the equality of infima now follows readily since \(C^\infty(\Omega; \mathbb{R}^m)\) is dense in \(W^{1,p}(\Omega; \mathbb{R}^m)\). The equality of the infima under an additional boundary value constraint follows from the continuity of the trace operator under the \(W^{1,p}\)-convergence, see Theorem A.24, and the fact that any map in \(W^{1,p}(\Omega; \mathbb{R}^m)\) can be approximated with smooth functions with the same boundary values, see Theorem A.29.

If we dispense with the \(p\)-growth assumption, however, the infimum over different spaces may indeed be different – this is called the Lavrentiev gap phenomenon, discovered in 1926 by Mikhail Lavrentiev. Here, we give an example between the spaces \(W^{1,1}(0,1)\) and \(W^{1,\infty}(0,1)\) (with boundary conditions):

**Example 2.14 (Manià 1934 [178]).** Consider the minimization problem

\[
\begin{align*}
\text{Minimize} & \quad \mathcal{F}[u] := \int_0^1 (u(t)^3 - t)^2 \dot{u}(t)^6 \, dt \\
\text{subject to} & \quad u(0) = 0, u(1) = 1
\end{align*}
\]

for \(u\) from either \(W^{1,1}(0,1)\) or \(W^{1,\infty}(0,1)\). We claim that

\[ \inf_{W^{1,1}(0,1)} \mathcal{F} < \inf_{W^{1,\infty}(0,1)} \mathcal{F}, \]

where here and in the following these infima are to be taken only over functions \(u\) with boundary values \(u(0) = 0, u(1) = 1\).

Clearly, \(\mathcal{F} \geq 0\), and for \(u_\ast(t) := t^{1/3} \in (W^{1,1} \setminus W^{1,\infty})(0,1)\) we have \(\mathcal{F}[u_\ast] = 0\). Thus,

\[ \inf_{W^{1,1}(0,1)} \mathcal{F} = 0. \]

On the other hand, every \(u \in W^{1,\infty}(0,1)\) is Lipschitz continuous. Thus, also using \(u(0) = 0, u(1) = 1\), there exists a \(\tau \in (0,1)\) with

\[ u(t) \leq h(t) := \frac{t^{1/3}}{2} \quad \text{for all } t \in [0, \tau] \quad \text{and} \quad u(\tau) = h(\tau). \]

Then, \(u(t)^3 - t \leq h(t)^3 - t\) for \(t \in [0, \tau]\) and, since both of these terms are negative,
2.4 The Lavrentiev gap phenomenon

\[(u(t)^3 - t)^2 \geq (h(t)^3 - t)^2 = \frac{7^2}{8^2} t^2 \quad \text{for all } t \in [0, \tau].\]

We then estimate

\[F[u] \geq \int_0^\tau (u(t)^3 - t)^2 \dot{u}(t)^6 \, dt \geq \frac{7^2}{8^2} \int_0^\tau t^2 \dot{u}(t)^6 \, dt.\]

Further, by Hölder’s inequality,

\[\int_0^\tau \dot{u}(t) \, dt = \int_0^\tau t^{-1/3} \cdot t^{1/3} \dot{u}(t) \, dt \leq \left( \int_0^\tau t^{-2/5} \, dt \right)^{5/6} \cdot \left( \int_0^\tau t^2 \dot{u}(t)^6 \, dt \right)^{1/6} = \frac{5^{5/6}}{3^{5/6}} \cdot \frac{1}{3^{1/2}} \left( \int_0^\tau t^2 \dot{u}(t)^6 \, dt \right)^{1/6}.\]

Since also

\[\int_0^\tau \dot{u}(t) \, dt = u(\tau) - u(0) = h(\tau) = \frac{\tau^{1/3}}{2},\]

we arrive at

\[F[u] \geq \frac{7^2 \cdot 3^5}{8^2 \cdot 5^{26} \tau} > \frac{7^2 \cdot 3^5}{8^2 \cdot 5^{26}} > 0.\]

Thus,

\[\inf_{W^{1,\infty}(0,1)} F > \inf_{W^{1,1}(0,1)} F,\]

and \(F\) can be seen to exhibit the Lavrentiev gap phenomenon.

In a more recent example, Ball & Mizel \[33\] showed that the problem

\[
\begin{aligned}
\text{Minimize} \quad & F[u] := \int_{-1}^{1} (t^4 - u(t)^6)^2 |\dot{u}(t)|^{2m} + \varepsilon \dot{u}(t)^2 \, dt \\
\text{subject to} \quad & u(-1) = \alpha, \ u(1) = \beta
\end{aligned}
\]

also exhibits the Lavrentiev gap phenomenon between the spaces \(W^{1,2}\) and \(W^{1,\infty}\) if \(m \in \mathbb{N}\) satisfies \(m > 13\), \(\varepsilon > 0\) is sufficiently small, and \(-1 \leq \alpha < 0 < \beta \leq 1\). This example is significant because the Ball–Mizel functional is \textit{coercive} on \(W^{1,2}(-1,1)\) thanks to the second term of the integrand.

We note that the Lavrentiev gap phenomenon is a major obstacle for the numerical approximation of minimization problems. For instance, standard (piecewise affine) finite element approximations are in \(W^{1,\infty}\) and hence in the presence of the Lavrentiev gap phenomenon (between \(W^{1,p}\) and \(W^{1,\infty}\)) we cannot approximate the true solution with such finite elements. Thus, one is forced to work with non-conforming elements and other advanced schemes. This issue does not only affect “academic” examples such as the ones above, but is also of great concern in applied problems, such as nonlinear elasticity theory.
2.5 Integral side constraints

In some minimization problems the class of candidate functions is restricted to include one or more integral side constraints. To establish the existence of a minimizer in these cases, we first need to extend the Direct Method to this scenario.

**Theorem 2.15.** Let $X$ be a Banach space or a closed affine subset of a Banach space and let $\mathcal{F}, \mathcal{H} : X \to \mathbb{R} \cup \{+\infty\}$. Assume the following:

(\text{WH1}) **Weak coercivity of $\mathcal{F}$**: For all $L \in \mathbb{R}$ the sublevel set \( \{ u \in X : \mathcal{F}[u] \leq L \} \) is sequentially weakly precompact, that is, if $\mathcal{F}[u_j] \leq L$ for a sequence $(u_j) \subset X$ and some $L \in \mathbb{R}$, then $(u_j)$ has a weakly converging subsequence.

(\text{WH2}) **Weak lower semicontinuity of $\mathcal{F}$**: For all sequences $(u_j) \subset X$ with $u_j \rightharpoonup u$ in $X$ it holds that
\[
\mathcal{F}[u] \leq \liminf_{j \to \infty} \mathcal{F}[u_j].
\]

(\text{WH3}) **Weak continuity of $\mathcal{H}$**: For all sequences $(u_j) \subset X$ with $u_j \rightharpoonup u$ in $X$ it holds that
\[
\mathcal{H}[u_j] \to \mathcal{H}[u].
\]

Assume also that there exists at least one $u_0 \in X$ with $\mathcal{H}[u_0] = 0$. Then, the minimization problem

Minimize $\mathcal{F}[u]$ over all $u \in X$ with $\mathcal{H}[u] = 0$

has a solution.

**Proof.** The proof is almost exactly the same as the one for the standard Direct Method in Theorem 2.3. The only difference is that we need to select the $u_j$ for a minimizing sequence with $\mathcal{H}[u_j] = 0$. Then, by (WH3), this property also holds for any weak limit $u$ of a subsequence of the $u_j$’s, which then is the sought minimizer. \(\square\)

A large class of side constraints can be treated using the following simple result.

**Lemma 2.16.** Let $h : \Omega \times \mathbb{R}^m \to \mathbb{R}$ be a Carathéodory integrand and let $p \in [1, \infty)$ such that there exists an $M > 0$ with
\[
|h(x,v)| \leq M(1 + |v|^q), \quad (x,v) \in \Omega \times \mathbb{R}^m,
\]

for some $q \in [1, dp/(d - p))$ if $p \leq d$, or no growth condition if $p > d$. Then, the functional $\mathcal{H} : W^{1,p}(\Omega; \mathbb{R}^m) \to \mathbb{R}$ defined through
\[
\mathcal{H}[u] := \int_{\Omega} h(x,u(x)) \, dx, \quad u \in W^{1,p}(\Omega; \mathbb{R}^m),
\]

is weakly continuous.
Proof. We only prove the lemma in the case $p \leq d$. The proof for $p > d$ is analogous, but easier.

Let $u_j \rightharpoonup u$ in $W^{1,p}(\Omega; \mathbb{R}^m)$, whereby after selecting a subsequence and employing the Rellich–Kondrachov Theorem A.28 and Lemma A.8, $u_j \rightarrow u$ in $L^q$ and almost everywhere. By assumption we have

$$\pm h(x,v) + M(1 + |v|^q) \geq 0.$$ 

Thus, applying Fatou’s lemma separately to these two integrands, we get

$$\liminf_{j \rightarrow \infty} \left( \pm \mathcal{H}[u_j] + \int_{\Omega} M(1 + |u_j|^q) \, dx \right) \geq \pm \mathcal{H}[u] + \int_{\Omega} M(1 + |u|^q) \, dx.$$ 

Since $\|u_j\|_{L^q} \rightarrow \|u\|_{L^q}$, we can combine these two assertions to get $\mathcal{H}[u_j] \rightarrow \mathcal{H}[u]$. This holds for a subsequence of any subsequence of $(u_j)$, hence it also holds for our original sequence. $\square$

Combining this lemma with Theorems 2.7 and 2.15 and also the Rellich–Kondrachov Theorem A.28, we immediately get the following existence result.

**Theorem 2.17.** Let $f : \Omega \times \mathbb{R}^{m \times d} \rightarrow [0, \infty)$ and $h : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}$ be Carathéodory integrands such that

(i) $f$ satisfies the $p$-coercivity bound (2.2), where $p \in (1, \infty)$;
(ii) $f(x,\cdot)$ is convex for all $x \in \Omega$;
(iii) $h$ satisfies the $q$-growth condition (2.5) for some $q \in [1,dp/(d-p))$ if $p \leq d$, or no growth condition if $p > d$.

Then, there exists a minimizer $u \in W^{1,p}_g(\Omega; \mathbb{R}^m)$, where $g \in W^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$, of the functional

$$\mathcal{F}[u] := \int_{\Omega} f(x,u(x),\nabla u(x)) \, dx, \quad u \in W^{1,p}_g(\Omega; \mathbb{R}^m),$$

under the side constraint

$$\mathcal{H}[u] := \int_{\Omega} h(x,u(x)) \, dx = 0.$$

2.6 The general theory of convex functions and duality

We finish this chapter by briefly considering the general theory of convex functions.

In all of the following let $X$ be a (real) reflexive Banach space (finite or infinite-dimensional) with dual space $X^*$, see Appendix A.2. We denote by $\langle x,x^* \rangle = x^*(x)$ the duality product between $x \in X$ and $x^* \in X^*$. For a set $A \subset X$ we write $\text{co} A$, $\text{cl} A$ for its convex hull and closed convex hull, respectively. These hulls are defined
to be the smallest (closed) convex set containing $A$, or, equivalently, the intersection of all (closed) convex sets containing $A$. For $A \subset X$ we furthermore define the characteristic function $\chi_A : X \to \mathbb{R} \cup \{+\infty\}$ as

$$\chi_A(x) := \frac{1}{\mathbb{1}_A(x)} - 1 = \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{if } x \notin A. \end{cases}$$

Let $F : X \to \mathbb{R} \cup \{+\infty\}$. The function $F$ is called proper if it is not identically $+\infty$. We define the effective domain $\text{dom} F \subset X$ and the epigraph $\text{epi} F \subset X \times \mathbb{R}$ of $F$ as follows:

$$\text{dom} F := \{ x \in X : F(x) < +\infty \},$$

$$\text{epi} F := \{ (x, \alpha) \in X \times \mathbb{R} : \alpha \geq F(x) \}.$$

It can be shown (see Problems [14.4, 14.5]) that $F$ is convex if and only if $\text{epi} F$ is convex (as a set), and that $f$ is (sequentially) lower semicontinuous if and only if $\text{epi} F$ is (sequentially) closed; this holds with respect to both the strong and the weak convergence.

**Lemma 2.18.** If $\dim X < \infty$, then every convex function $F : X \to \mathbb{R} \cup \{+\infty\}$ is locally bounded on the interior of its effective domain.

**Proof.** If $x \in X$ is in the interior of the effective domain of $F$, then $x$ lies in the convex hull $\text{co} \{x_1, \ldots, x_{n+1}\}$ of $n + 1$ affinely independent points $x_k$ (i.e., $\sum \alpha_k x_k = 0$ for some $\alpha_k \in \mathbb{R}$ with $\sum \alpha_k = 0$ implies $\alpha_1 = \alpha_2 = \cdots = \alpha_{n+1} = 0$) with $F(x_k) < +\infty$, where $n = \dim X$. Thus, there exists an open ball around $x$ inside $\text{co} \{x_1, \ldots, x_{n+1}\}$ on which $F$ is bounded by $\sup \{F(x_1), \ldots, F(x_{n+1})\}$. \qed

**Lemma 2.19.** Let $\mathcal{A}$ be a non-empty family of continuous affine functions $a(x) = \langle x, x^* \rangle + \alpha$ for some $x^* \in X^*$, $\alpha \in \mathbb{R}$. Then, $F : X \to \mathbb{R} \cup \{+\infty\}$ defined through

$$F(x) := \sup_{a \in \mathcal{A}} a(x)$$

is convex and lower semicontinuous. Conversely, every convex and lower semicontinuous function can be written in this form.

**Proof.** The convexity of $F$ is clear since all the affine functions $a \in \mathcal{A}$ are in particular convex. For the lower semicontinuity we just need to realize that the pointwise supremum of continuous functions is always lower semicontinuous. Indeed, for a sequence $x_j \to x$ in $X$ we have for all $\tilde{a} \in \mathcal{A}$ that

$$\tilde{a}(x) = \lim_{j \to \infty} \tilde{a}(x_j) \leq \liminf_{j \to \infty} \sup_{a \in \mathcal{A}} a(x_j) = \liminf_{j \to \infty} F(x_j).$$

Taking the supremum over all $\tilde{a} \in \mathcal{A}$, the lower semicontinuity follows.

For the converse, we may assume that $F$ is proper; otherwise the result is trivial. Let $x \in X$ with $F(x) < +\infty$. The epigraph $\text{epi} F$ of $F$ is closed and convex by
Fig. 2.2 The convex conjugate

assumption. Hence, by the Hahn–Banach Separation Theorem A.1, for every \( x \in X \) and every \( \beta < F(x) \) we can find an affine function \( a_{x, \beta} : X \to \mathbb{R} \) whose graph separates the point \((x, \beta)\) from \( \text{epi} F \). In particular, \( \beta < a_{x, \beta}(x) < F(x) \) and \( a_{x, \beta} \) lies everywhere below the graph of \( F \). Letting \( \beta \uparrow F(x) \), we arrive at

\[
F(x) = \sup \left\{ a_{x, \beta}(x) : (x, \beta) \in X \times \mathbb{R} \text{ with } \beta < F(x) \right\}.
\]

A similar argument also applies if \( F(x) = +\infty \). Collecting all these \( a_{x, \beta} \) for \((x, \beta) \in X \times \mathbb{R} \) into the set \( \mathcal{A} \), the conclusion follows. \( \Box \)

**Proposition 2.20.** Every proper convex function is continuous on the interior of its effective domain.

We will prove this in more generality later, see Lemma 5.6 in conjunction with Lemma A.22.

One important object in the general theory of convex functions is the (**convex**) conjugate, or **Legendre–Fenchel transform**, \( F^* : X^* \to \mathbb{R} \cup \{ +\infty \} \) of a proper function \( F : X \to \mathbb{R} \cup \{ +\infty \} \) (not necessarily convex), which is defined as follows:

\[
F^*(x^*) := \sup_{x \in X} \left[ \langle x, x^* \rangle - F(x) \right], \quad x^* \in X^*.
\]

Of course, we may restrict to \( x \in \text{dom} F \) in the supremum. The intuition here is that for a given \( x^* \) we may consider all affine hyperplanes with normal \( x^* \) (recall that all hyperplane normals are elements of \( X^* \)) that lie below \( \text{epi} F \). Then, \( -F^*(x^*) \) is the supremum of the heights at which these hyperplanes intersect the (vertical) \( (\mathbb{R} \cup \{ +\infty \}) \)-axis, see Figure A.12. Indeed, let \( \alpha \in \mathbb{R} \) be such that \( F(x) \geq \langle x, x^* \rangle - \alpha \) for all \( x \in X \). Then, \( \alpha \geq \langle x, x^* \rangle - F(x) \) for all \( x \in X \), so the highest supporting hyperplane with normal \( x^* \) is \( x \mapsto \langle x, x^* \rangle - F^*(x^*) \), which intersects the vertical axis in \( -F^*(x^*) \).

The following **Fenchel inequality** is immediate from the definition:

\[
\langle x, x^* \rangle \leq F(x) + F^*(x^*), \quad \text{for all } x \in X, x^* \in X^*.
\] (2.6)

We next collect some properties of the conjugate function:
Proposition 2.21. Let $F, G : X \to \mathbb{R} \cup \{+\infty\}$ be proper and $F^*, G^* : X^* \to \mathbb{R} \cup \{+\infty\}$ be their conjugates.

(i) $F^*$ is convex and lower semicontinuous.
(ii) $F^*(0) = -\inf F$.  
(iii) If $F \leq G$, then $G^* \leq F^*$.  
(iv) If for $\lambda > 0$ we denote by $F_\lambda$ the scaled function $F_\lambda(x) := F(\lambda x)$, then $F_\lambda^*(x^*) = F^*(x^*/\lambda)$.  
(v) $(\lambda F)^*(x^*) = \lambda F^*(x^*/\lambda)$ for all $\lambda > 0$.  
(vi) If for $a \in X$ we denote by $F_a$ the translated function $F_a(x) := F(x - a)$, then $F_a^*(x^*) = F^*(x^*) + \langle a, x^* \rangle$.  
(vii) If for a $g \in X$ we denote by $F_g$ the translated function $F_g(x) := F(x - g)$, then $F_g^*(x^*) = F^*(x^*) + \langle g, x^* \rangle$.

Proof. The first assertion follows from Lemma 2.19, all the others are straightforward calculations, see Problem 2.8. \(\square\)

We now consider a few canonical examples of convex functions.

Example 2.22 (Support function). Let $\chi_A$ be the characteristic function of $A \subset X$. Then, for the conjugate function we get

$$\sigma_A(x^*) := \chi_A^*(x^*) = \sup_{x \in A} \langle x, x^* \rangle, \quad x^* \in X^*,$$

which is called the support function of $A$. It is always convex, lower semicontinuous, and positively 1-homogeneous, i.e., $\sigma_A(\alpha x^*) = \alpha \sigma_A(x^*)$ for all $x^* \in X^*$ and $\alpha \geq 0$, see Problem 2.8.

Example 2.23. Let $p, q \in (1, \infty)$ with $1/p + 1/q = 1$, that is, $p, q$ are conjugate exponents. Then,

$$\phi(t) := \frac{1}{p} |t|^p \quad \text{and} \quad \phi^*(t) := \frac{1}{q} |t|^q, \quad t \in \mathbb{R},$$

are conjugate. From the Fenchel inequality (2.6) we recover the Young inequality

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \text{for all } x, y \geq 0.$$

Example 2.24. For the absolute value function $\phi(t) := |t|$ we get

$$\phi^*(t) = \chi_{[-1,1]}(t) = \begin{cases} 0 & \text{if } |t| \leq 1, \\ +\infty & \text{if } |t| > 1, \end{cases} \quad t \in \mathbb{R}. \quad (2.7)$$

Example 2.25. The conjugate of the exponential function is

$$\exp^*(t) = \begin{cases} +\infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ t \ln t - t & \text{if } t > 0, \end{cases} \quad t \in \mathbb{R}.$$
2.6 The general theory of convex functions and duality

In this case, (2.6.3) gives the inequality

\[ xy \leq \exp(x) + y \ln y - y \quad \text{for all } x, y > 0. \]

**Example 2.26.** Let \( \phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\} \) be proper, convex, and lower semicontinuous, and let \( \|\cdot\|, \|\cdot\|_\ast \) be the norms on \( X \) and on \( X^\ast \), respectively. Then the functions

\[ G(x) := \phi(||x||) \quad \text{and} \quad G^\ast(x^\ast) := \phi^\ast(||x^\ast||), \quad x \in X, x^\ast \in X^\ast, \]

are conjugate. In particular, \( ||\cdot||_p \) and \( ||\cdot||_q \) for \( 1/p + 1/q = 1 \) are conjugate. The verification of these statements is the task of Problem 2.10.

**Example 2.27.** Let \( X = \mathbb{R}^n \) and let \( S \in \mathbb{R}^{n \times n} \) be a symmetric, positive definite matrix. Then,

\[ F(x) := \frac{1}{2} x^T S x \quad \text{and} \quad F^\ast(y) := \frac{1}{2} y^T S^{-1} y, \quad x, y \in \mathbb{R}^n, \]

are conjugate.

Iterating the construction of the conjugate, we denote by \( F^{**} : X \to \mathbb{R} \cup \{+\infty\} \) the **biconjugate** of \( F \), that is, the function

\[ F^{**}(x) := \sup_{x^\ast \in X^\ast} \left[ (x, x^\ast) - F^\ast(x^\ast) \right], \quad x \in X. \]

**Proposition 2.28.** The biconjugate \( F^{**} \) is the convex, lower semicontinuous envelope of \( F \), that is, the greatest convex, lower semicontinuous function below \( F \). Moreover, \( F^{**} = F^\ast \).

**Proof.** For the moment denote the convex lower semicontinuous envelope of \( F \) by \( F_{\text{clsc}} \).

\[ F_{\text{clsc}}(x) := \sup \left\{ H(x) : H \leq F \text{ convex, lower semicontinuous} \right\}, \quad x \in X. \]

Also define

\[ G(x) := \sup \left\{ a(x) : a \leq F \text{ affine} \right\}, \quad x \in X. \]

Since \( G \leq F \) is convex and lower semicontinuous by Lemma 2.19, \( G \leq F_{\text{clsc}} \). On the other hand, for every convex and lower semicontinuous \( H \) from the definition of \( F_{\text{clsc}} \), we have \( H(x) = \sup_{b \in \mathcal{A}} b(x) \) for a collection of affine functions \( b \leq H \), again by the said lemma. However, \( b \leq F \) for all \( b \in \mathcal{A} \) and thus \( b \) is included in the collection in the definition of \( G \). Hence, \( H \leq G \), whereby \( F_{\text{clsc}} \leq G \). In conclusion, \( F_{\text{clsc}} = G \).

Every affine \( a \leq F \) has the form \( a(x) = (x, x^\ast) - \alpha \) for some \( x^\ast \in X^\ast \) and \( \alpha \in \mathbb{R} \). We can restrict ourselves to such \( a \) with \( \alpha \) minimal while still preserving the property \( a \leq F \). We see first that \( a \leq F \) if and only if \( \alpha \geq (y, x^\ast) - F(y) \) for all \( y \in X \). According to the definition of the conjugate function, this condition is nothing else than
Thus, $\alpha$ is minimal when $\alpha = F^*(x^*)$ and we get

$$F_{\text{clsc}}(x) = G(x) = \sup_{x^* \in X^*} \left[ \langle x, x^* \rangle - F^*(x^*) \right] = F^{**}(x), \quad x \in X.$$  

For the second assertion it suffices to observe that $F^*$ is convex and lower semi-continuous by Proposition 2.21(i) and to apply the first assertion. \hfill \Box

As a particular consequence of the preceding result, we see that conjugation facilitates a bijection between the proper, convex, and lower semicontinuous functions on $X$ and those on $X^*$, which is self-inverse in the sense above.

**Corollary 2.29.** $\text{epi} F^{**} = \overline{\text{coepi}} F$.

**Proof.** The process of taking the convex lower semicontinuous envelope of $F$ amounts to finding the closed convex hull of the epigraph. \hfill \Box

**Example 2.30.** For the characteristic function $\chi_A$ of $A \subset X$ we get

$$\chi^{**} = \sigma^*_A = \chi_{\overline{\text{co}A}}.$$  

In particular, $A$ and $\overline{\text{co}A}$ have the same support function.

**Notes and historical remarks**

The basic ideas concerning the Direct Method as well as lower semicontinuity and its connection to convexity are due to Leonida Tonelli and were established in a series of articles in the early 20th century [275–277]. In the 1960s James Serrin generalized the results to higher dimensions [242].

Most of the material in this chapter is very classical and can be found in a variety of books on the calculus of variations, we refer in particular to [76, 77, 137]. We note that a very general lower semicontinuity theorem for convex integrands can be found in Theorem 3.23 of [76].

All of our abstract results on the Direct Method are formulated using sequences and not using general topology tools like nets. This is justified since the weak topology on a separable, reflexive Banach space and the weak*-topology on a dual space with a separable predual are metrizable on norm-bounded sets. Thus, if the functionals under investigation satisfy suitable coerciveness assumptions, one can work with sequences. The only case where one has to be careful is when one uses the weak topology on a non-reflexive Banach space with a non-separable dual space because then the weak topology might not be metrizable. For instance, in the sequence space $l^1$ (with non-separable dual space $l^{1*}$), weak convergence of sequences is equivalent to strong convergence, but the weak and strong topologies still differ.
Problems

2.1. Let $\mathcal{F} : X \to \mathbb{R}$, where $X$ is a complete metric space. Show that if every subsequence of the sequence $(u_j) \subset X$ with $u_j \to u$ in $X$ has a further subsequence $(u_{j(k)})_k$ such that $\mathcal{F}[u] \leq \liminf_{k \to \infty} \mathcal{F}[u_{j(k)}]$, then also $\mathcal{F}[u] \leq \liminf_{j \to \infty} \mathcal{F}[u_j]$.

2.2. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Define $V := \left\{ u \in W^{1,2}(\Omega) : \int_{\Omega} u(x) \, dx = 0 \right\}$.

Assume furthermore that $f : \Omega \times \mathbb{R}^d \to \mathbb{R}$ is continuously differentiable with

$$
\mu |A|^2 \leq f(x,A) \quad \text{for some } \mu > 0 \text{ and all } (x,A) \in \Omega \times \mathbb{R}^d,
$$

$$
|D_A f(x,A)| \leq M (1 + |A|^2) \quad \text{for some } M > 0 \text{ and all } (x,A) \in \Omega \times \mathbb{R}^d,
$$

and that $A \mapsto f(x,A)$ is convex for all $x \in \Omega$. Finally, let $g \in L^2(\Omega)$. Consider the following minimization problem:
\[
\begin{aligned}
\begin{cases}
\text{Minimize} & \mathcal{F}[u] := \int_{\Omega} f(x, \nabla u(x)) - g(x)u(x) \, dx \\
\text{over all} & u \in V.
\end{cases}
\end{aligned}
\]

(i) Show that \( \mathcal{F} \) is coercive on \( V \), that is, there exists a \( \mu > 0 \) such that
\[
\mathcal{F}[u] \geq \mu \|u\|_{W^{1,2}}^2 - \mu^{-1} \quad \text{for all} \quad u \in V.
\]

(ii) Show that \( \mathcal{F} \) is also weakly lower semicontinuous on \( V \) (weak convergence in \( W^{1,2} \)) and hence there exists a minimizer \( u_* \in V \) of \( \mathcal{F} \) (minimized over \( V \)).

This problem is continued in Problem 3.9 in the next chapter.

2.3. Show that the function \( f : \mathbb{R}^2 \to \mathbb{R} \) given by \( f(x, y) = xy \) is \textbf{separately convex}, that is, \( x \mapsto f(x, y) \) is convex for fixed \( y \in \mathbb{R} \) and \( y \mapsto f(x, y) \) is convex for fixed \( x \in \mathbb{R} \), but \( f \) is not convex.

2.4. Let \( f : \mathbb{R}^d \to [0, \infty) \) be twice continuously differentiable and assume that there are constants \( \mu, M > 0 \) with
\[
\mu |b|^2 \leq D^2 f(a)[b, b] \leq M |b|^2 \quad \text{for all} \quad a, b \in \mathbb{R}^d,
\]
where
\[
D^2 f(a)[b, b] := \left. \frac{d^2}{dt^2} f(a + tb) \right|_{t=0} \quad \text{for all} \quad a, b \in \mathbb{R}^d.
\]
Show that \( f \) is convex and that \(|f(v)| \leq C(1 + |v|^2)\) for some \( C > 0 \) and all \( v \in \mathbb{R}^d \).

2.5. Let \( f : \mathbb{R}^d \to \mathbb{R} \) be convex and fix \( x_0 \in \mathbb{R}^d \). Set
\[
M := \max_{i=1, \ldots, d} \left( |f(x_0 + e_i) - f(x_0)|, |f(x_0 - e_i) - f(x_0)| \right).
\]
Prove that if \( y \in \mathbb{R}^d \) satisfies \(|y|_1 := |y_1| + \cdots + |y_d| \leq 1\), then \( f(x_0 + y) - f(x_0) \leq M \).

2.6. Show that \( F : X \to \mathbb{R} \cup \{+\infty\} \) is convex if and only if \( \text{epi} \, F \) is convex (as a set).

2.7. Show that \( F : X \to \mathbb{R} \cup \{+\infty\} \) is (sequentially) lower semicontinuous if and only if \( \text{epi} \, F \) is (sequentially) closed.

2.8. Prove the statements of Proposition 2.4.4.1.

2.9. Verify the statements in Example 2.4.2 about the support function.

2.10. Prove the assertion in Example 2.4.6.
References

References

References


30 References


References
