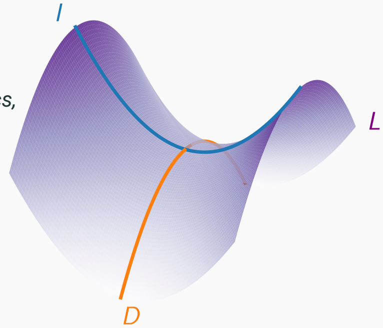
A priori and *a posteriori* error identities for convex minimization problems based on convex duality relations

Lecture 1

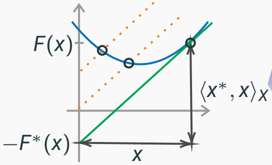
Alex Kaltenbach

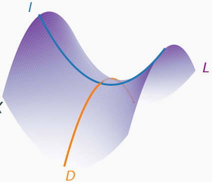
"Oberseminar" of the Department of Applied Mathematics,
University of Freiburg,
14th–20th August 2024



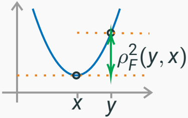
General objective

1. A posteriori error identities on the basis of <u>continuous</u> convex duality.



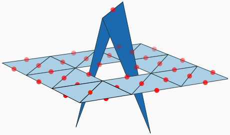


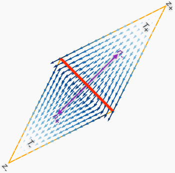
$$\rho^2_{\mathrm{tot}}(v,y) = \eta^2_{\mathrm{gap}}(v,y)$$



---

2. <u>Numerically practicable</u> a posteriori error identities on the basis of <u>discrete</u> convex duality.





$$z_h^{rt} = D_t\phi_h(\cdot, \nabla_h u_h^{cr}) + \frac{D_t\psi_h(\cdot, \Pi_h u_h^{cr})}{d}\left(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}\right)$$

A priori error analysis
and
a posteriori error analysis

Concepts of *a priori* and *a posteriori* error analysis

- **Continuous problem:** Seek $u \in X$ s.t.

$$I(u) = \inf_{v \in X} I(v)\,,$$

where $I\colon X \to \mathbb{R} \cup \{+\infty\}$.

- **Discrete problem:** Seek $u_h \in X_h$ ($\not\subseteq X$) s.t.

$$I_h(u_h) = \inf_{v_h \in X_h} I_h(v_h)\,,$$

where $I_h \colon X_h \to \mathbb{R} \cup \{+\infty\}$.

- ***A priori* error estimate:** (continuous into discrete)

$$\rho_{I_h}^2(u_h, \Pi_h^{X_h} u) \leq \eta_{I_h}^2(\Pi_h^{X_h} u) = \begin{pmatrix}\text{independent of } u_h \\ \text{dependent of } u\end{pmatrix},$$

where

- $\rho_{I_h}^2 \colon X_h \times X_h \to \mathbb{R}_{\geq 0}$ is a *discrete distance measure*;
- $\Pi_h^{X_h} \colon X \to X_h$ a *(quasi-)interpolation operator*;
- $\eta_{I_h}^2 \colon X_h \to \mathbb{R}_{\geq 0}$ an *a priori error estimator*.

Concepts of *a priori* and *a posteriori* error analysis

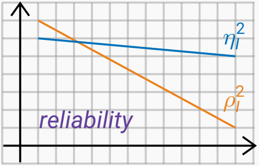
- ***A posteriori* error estimate:** (discrete into continuous)

$$\rho_I^2(u,\Pi_h^X u_h) \leq \eta_I^2(\Pi_h^X u_h) = \begin{pmatrix}\text{independent of } u\\ \text{dependent of } u_h\end{pmatrix},$$

  where
  - $\rho_I^2 : X \times X \to \mathbb{R}_{\geq 0}$ is a *distance measure*;
  - $\Pi_h^X : X_h \to X$ a *post-processing operator*;
  - $\eta_I^2 : X \to \mathbb{R}_{\geq 0}$ an *a posteriori error estimator*.

- **Reliability/Efficiency:** An *a posteriori* error estimator is called[1]
  - *reliable*, if

$$\rho_I^2(u,\Pi_h^X u_h) \lesssim \eta_I^2(\Pi_h^X u_h)\,;$$



[1] We write $A \lesssim B$ if $A \leq c\,B$ for a constant that does not depend on $h > 0$.

Concepts of *a priori* and *a posteriori* error analysis

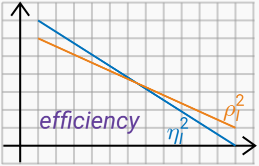
- ***A posteriori* error estimate:** (discrete into continuous)

$$\rho_I^2(u, \Pi_h^X u_h) \leq \eta_I^2(\Pi_h^X u_h) = \begin{pmatrix} \text{independent of } u \\ \text{dependent of } u_h \end{pmatrix},$$

  where

  - $\rho_I^2 : X \times X \to \mathbb{R}_{\geq 0}$ is a *distance measure*;
  - $\Pi_h^X : X_h \to X$ a *post-processing operator*;
  - $\eta_I^2 : X \to \mathbb{R}_{\geq 0}$ an *a posteriori error estimator*.

- **Reliability/Efficiency:** An *a posteriori* error estimator is called[2]
  - *reliable*, if
  $$\rho_I^2(u, \Pi_h^X u_h) \lesssim \eta_I^2(\Pi_h^X u_h)\,;$$
  - *efficient*, if
  $$\rho_I^2(u, \Pi_h^X u_h) \gtrsim \eta_I^2(\Pi_h^X u_h)\,.$$



[2] We write $A \lesssim B$ if $A \leq c\,B$ for a constant that does not depend on $h > 0$.

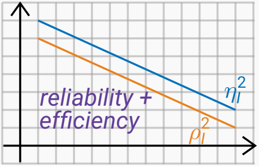
Concepts of *a priori* and *a posteriori* error analysis

- ***A posteriori* error estimate:** (discrete into continuous)

$$\rho_I^2(u,\Pi_h^X u_h) \leq \eta_I^2(\Pi_h^X u_h) = \begin{pmatrix}\text{independent of } u\\ \text{dependent of } u_h\end{pmatrix},$$

  where
  - $\rho_I^2\colon X\times X\to\mathbb{R}_{\geq 0}$ is a *distance measure*;
  - $\Pi_h^X\colon X_h\to X$ a *post-processing operator*;
  - $\eta_I^2\colon X\to\mathbb{R}_{\geq 0}$ an *a posteriori error estimator*.

- **Reliability/Efficiency:** An *a posteriori* error estimator is called[3]
  - *reliable*, if
$$\rho_I^2(u,\Pi_h^X u_h) \lesssim \eta_I^2(\Pi_h^X u_h)\,;$$
  - *efficient*, if
$$\rho_I^2(u,\Pi_h^X u_h) \gtrsim \eta_I^2(\Pi_h^X u_h)\,.$$



- **Objective:** *A posteriori* error estimator $\eta_I^2\colon X\to\mathbb{R}_{\geq 0}$ that is *reliable* and *efficient*, *i.e.*,
$$\rho_I^2(u,\Pi_h^X u_h) \sim \eta_I^2(\Pi_h^X u_h)\,.$$

[3] We write $A\lesssim B$ if $A\leq c\,B$ for a constant that does not depend on $h>0$.

Concepts of *a priori* and *a posteriori* error analysis

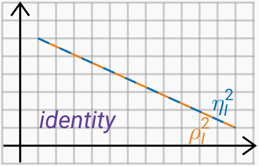
- ***A posteriori* error estimate:** (discrete into continuous)

$$\rho_I^2(u,\Pi_h^X u_h) \leq \eta_I^2(\Pi_h^X u_h) = \begin{pmatrix}\text{independent of } u\\ \text{dependent of } u_h\end{pmatrix},$$

where

- $\rho_I^2\colon X\times X\to\mathbb{R}_{\geq 0}$ is a *distance measure*;
- $\Pi_h^X\colon X_h\to X$ a *post-processing operator*;
- $\eta_I^2\colon X\to\mathbb{R}_{\geq 0}$ an *a posteriori error estimator*.

- **Reliability/Efficiency:** An *a posteriori* error estimator is called[4]
  - *reliable*, if
$$\rho_I^2(u,\Pi_h^X u_h) \lesssim \eta_I^2(\Pi_h^X u_h)\,;$$
  - *efficient*, if
$$\rho_I^2(u,\Pi_h^X u_h) \gtrsim \eta_I^2(\Pi_h^X u_h)\,.$$



- **Objective:** *A posteriori* error estimator $\eta_I^2\colon X\to\mathbb{R}_{\geq 0}$ that is *identical*, *i.e.*,

$$\rho_I^2(u,\Pi_h^X u_h) = \eta_I^2(\Pi_h^X u_h)\,.$$

[4] We write $A\lesssim B$ if $A\leq c\,B$ for a constant that does not depend on $h>0$.

Content of the course

- **Lecture 1: General convex duality theory**
- **Lecture 2: Convex duality theory for integral functionals**
- **Lecture 3: Convex duality theory for discrete integral functionals**
- **Lecture 4: Applications**

Content of the Lecture 1

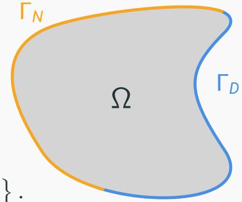
- **Lecture 1: General convex duality theory**
  - The Prager–Synge–Mikhlin identity.
  - Convex analysis.
  - Lagrange duality theory.
  - Fenchel duality theory.

The Prager–Synge–Mikhlin identity

Function spaces

- **Dirichlet/Neumann boundary:** Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded Lipschitz domain and

$$\Gamma_D, \Gamma_N \subseteq \partial\Omega \quad \text{s.t.} \quad \Gamma_D \dot{\cup} \Gamma_N = \partial\Omega\,.$$



- **Sobolev spaces:** For $p \in [1, +\infty)$, we define

$$W^{1,p}(\Omega) \coloneqq \{v \in L^p(\Omega) \mid \nabla v \in (L^p(\Omega))^d\}\,,$$
$$W^{1,p}_D(\Omega) \coloneqq \{v \in W^{1,p}(\Omega) \mid v = 0 \text{ a.e. on } \Gamma_D\}\,.$$

- H(**div**)**-spaces:** For $p \in (1, +\infty]$, we define

$$W^{p'}(\mathrm{div};\Omega) \coloneqq \{y \in (L^{p'}(\Omega))^d \mid \mathrm{div}\, y \in L^{p'}(\Omega)\}\,,$$
$$W^{p'}_N(\mathrm{div};\Omega) \coloneqq \{y \in W^{p'}(\mathrm{div};\Omega) \mid \langle y\cdot n, v\rangle_{W^{1-\frac{1}{p},p}(\partial\Omega)} = 0 \text{ for all } v \in W^{1,p}_D(\Omega)\}\,,$$

where $(y \mapsto y\cdot n)\colon W^{p'}(\mathrm{div};\Omega) \to (W^{1-\frac{1}{p},p}(\partial\Omega))^*$, for every $v \in W^{1,p}(\Omega)$ defined by

$$\langle y\cdot n, v\rangle_{W^{1-\frac{1}{p},p}(\partial\Omega)} = \int_\Omega y\cdot\nabla v\,\mathrm{d}x + \int_\Omega \mathrm{div}\, y\, v\,\mathrm{d}x\,,$$

is the *normal trace operator*.

- **Statement**: *A posteriori* error identity for both
  - ✓ weak formulation;
  - ✓ mixed formulation;

  based on geometric argument.


Figure: LEFT: William Prager (23 May 1903 – 17 March 1980); RIGHT: John Lighton Synge (23 March 1897 – 30 March 1995).

- **Weak formulation**: Seek $u \in W^{1,2}_D(\Omega)$ s.t. for every $v \in W^{1,2}_D(\Omega)$, it holds that

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega f\, v \,\mathrm{d}x\,. \qquad (f \in L^2(\Omega))$$

- **Mixed formulation**: Seek $(z,u)^\top \in W^2_N(\mathrm{div};\Omega) \times L^2(\Omega)$ s.t. for every $(y,v)^\top \in W^2_N(\mathrm{div};\Omega) \times L^2(\Omega)$, it holds that

$$\begin{aligned} \int_\Omega z \cdot y \,\mathrm{d}x + \int_\Omega u \,\mathrm{div}\, y \,\mathrm{d}x &= 0 && (\Leftrightarrow \quad z = \nabla u)\,, \\ \int_\Omega \mathrm{div}\, z\, v \,\mathrm{d}x &= -\int_\Omega f\, v\,\mathrm{d}x && (\Leftrightarrow \quad \mathrm{div}\, z = -f)\,. \end{aligned} \qquad (f \in L^2(\Omega))$$

Prager–Synge identity

Theorem: (Prager–Synge identity, *cf.* [3, Prager & Synge, '47])

For every $v \in W^{1,2}_D(\Omega)$ and $y \in W^2_N(\mathrm{div} = -f; \Omega)$, it holds that

$$\underbrace{\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x}_{\text{(primal error)}} + \underbrace{\frac{1}{2}\int_\Omega |y - z|^2 \,\mathrm{d}x}_{\text{(dual error)}} = \underbrace{\frac{1}{2}\int_\Omega |\nabla v - y|^2 \,\mathrm{d}x}_{\text{(primal-dual gap)}} .$$

- **Proof (based on geometric argument).**
  - **Orthogonality relation:** Due to

$$\int_\Omega (y - z)\cdot(\nabla v - \nabla u)\,\mathrm{d}x = \int_\Omega (\underbrace{\mathrm{div}\, z}_{=-f} - \underbrace{\mathrm{div}\, y}_{=-f})\cdot(v - u)\,\mathrm{d}x = 0\,,$$

we have that

$$y - z \perp \nabla v - \nabla u \quad \text{in } (L^2(\Omega))^d\,.$$

$\Rightarrow$ Using the *Pythagoras theorem*, we conclude that

$$\int_\Omega |\nabla v - \nabla u|^2\,\mathrm{d}x + \int_\Omega |y - z|^2\,\mathrm{d}x = \int_\Omega |\nabla v \underbrace{- \nabla u + z}_{=0} - y|^2\,\mathrm{d}x$$

$$= \int_\Omega |\nabla v - y|^2\,\mathrm{d}x\,.$$

■

Mikhlin estimate

- **Statement**: *A posteriori* error estimate for only
  - ✓ primal problem;
  - ✗ dual problem;

  based on convex duality argument.

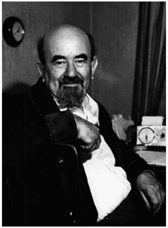


Figure: Solomon Grigorjewitsch Mikhlin (23 April 1908 – 29 August 1990).

- **Primal problem**: Seek $u \in W^{1,2}_D(\Omega)$ minimal for $I\colon W^{1,2}_D(\Omega) \to \mathbb{R}$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$I(v) \coloneqq \frac{1}{2}\int_\Omega |\nabla v|^2\,\mathrm{d}x - \int_\Omega f\,v\,\mathrm{d}x\,. \qquad (f \in L^2(\Omega))$$

- **Dual problem**: Seek $z \in W^2_N(\mathrm{div} = -f;\Omega)$ maximal for $D\colon W^2_N(\mathrm{div} = -f;\Omega) \to \mathbb{R}$, for every $y \in W^2_N(\mathrm{div} = -f;\Omega)$ defined by

$$D(y) \coloneqq -\frac{1}{2}\int_\Omega |y|^2\,\mathrm{d}x\,. \qquad (f \in L^2(\Omega))$$

Mikhlin estimate

Theorem: (Mikhlin estimate, *cf.* [2, Mikhlin, '64])

For every $v \in W^{1,2}_D(\Omega)$ and $y \in W^2_N(\mathrm{div} = -f; \Omega)$, it holds that

$$\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x \leq \frac{1}{2}\int_\Omega |\nabla v - y|^2 \,\mathrm{d}x\,.$$

- **Proof (based on convex duality argument).**
  - **Ingredient 1: (strong convexity)**

$$\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x \leq I(v) - I(u)\,.$$

  - **Ingredient 2: (weak duality)**

$$I(u) \geq D(y)\,.$$

$\Rightarrow$ Using the *integration-by-parts formula*, we conclude that

$$\begin{aligned}\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x &\leq I(v) - D(y)\\ &= \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x - \int_\Omega \overbrace{f}^{= -\mathrm{div}\, y} v \,\mathrm{d}x + \frac{1}{2}\int_\Omega |y|^2 \,\mathrm{d}x\\ &= \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x - \int_\Omega \nabla v \cdot y \,\mathrm{d}x + \frac{1}{2}\int_\Omega |y|^2 \,\mathrm{d}x\,.\end{aligned}$$

∎

Theorem: (Prager–Synge–Mikhlin identity)

For every $v \in W^{1,2}_D(\Omega)$ and $y \in W^2_N(\mathrm{div} = -f; \Omega)$, it holds that

$$\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x + \frac{1}{2}\int_\Omega |y - z|^2 \,\mathrm{d}x = \frac{1}{2}\int_\Omega |\nabla v - y|^2 \,\mathrm{d}x \,.$$

- **Proof (based on convex duality argument).**
  - **Ingredient 1: (optimal strong convexity)**

$$\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x = I(v) - I(u)\,,$$

$$\frac{1}{2}\int_\Omega |y - z|^2 \,\mathrm{d}x = D(z) - D(y)\,.$$

  - **Ingredient 2: (strong duality)**

$$I(u) = D(z)\,.$$

⇒ Using the *integration-by-parts formula*, we conclude that

$$\frac{1}{2}\int_\Omega |\nabla v - \nabla u|^2 \,\mathrm{d}x + \frac{1}{2}\int_\Omega |y - z|^2 \,\mathrm{d}x = I(v) - \underbrace{[I(u) - D(z)]}_{=0} - D(y)$$

$$= I(v) - D(y)\,.$$

∎

Convex analysis

Convex, proper, and lower semi-continuous functionals

Definition: (convex functions)

Let X be a Banach space. Then, a functional $F\colon X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is called

(i) *convex*, if for every $x, y \in X$ and $\lambda \in [0, 1]$, it holds that

$$F(\lambda x + (1-\lambda) y) \leq \lambda F(x) + (1-\lambda) F(y)\,,$$

whenever the right-hand side is well-defined, *i.e.*, $\infty - \infty$ does not occur.

(ii) *proper*, if $F(x) > -\infty$ for all $x \in X$ and $\operatorname{dom} F \neq \emptyset$, where

$$\operatorname{dom} F \coloneqq \{x \in X \mid F(x) < \infty\}\,,$$

denotes the *effective domain* of F.

(iii) *lower semi-continuous or l.s.c.*, if for every $(x_n)_{n\in\mathbb{N}} \subseteq X$ and $x \in X$, it holds that

$$x_n \to x \quad \text{in } X \quad (n \to \infty) \qquad \Rightarrow \qquad F(x) \leq \liminf_{n\to\infty} F(x_n)\,.$$

(iv) We define

$$\Gamma(X) \coloneqq \{F\colon X \to \overline{\mathbb{R}} \mid F \text{ convex, l.s.c., and } F > -\infty\} \,\dot{\cup}\, \{-\infty\}\,,$$
$$\Gamma_0(X) \coloneqq \{F \in \Gamma(X) \mid F \text{ proper}\}\,.$$

Fenchel conjugate

Definition: (Fenchel conjugate)

Let X be a Banach space and $F\colon X \to \overline{\mathbb{R}}$ a functional.

Then, the *Fenchel conjugate* $F^*: X^* \to \overline{\mathbb{R}}$, for every $x^* \in X^*$, is defined by

$$F^*(x^*) \coloneqq \sup_{x\in X} \left\{\langle x^*, x\rangle_X - F(x)\right\} .$$

- **Remarks:**
  - **Properties:** For every $x^* \in X^*$, it holds that

    $$F^*(x^*) = \sup_{x\in \operatorname{dom} F} \left\{\langle x^*, x\rangle_X - F(x)\right\},$$

    i.e., $F^* \in \Gamma(X^*)$;
  - **Fenchel–Young inequality:** For every $x \in X$ and $x^* \in X^*$, it holds that

    $$\langle x^*, x\rangle_X \leq F^*(x^*) + F(x) ,$$
    $$\text{with} \quad `{=}\text{'} \quad \Leftrightarrow \quad x^* \in \partial F(x) ,$$

    whenever $\infty - \infty$ does not occur.

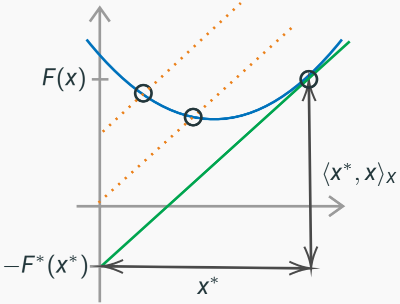


Figure: Geometric interpretation of Fenchel conjugate in 1D.

Sub-differential

Definition: (sub-differential)

Let X be a Banach space. Then, $F\colon X\to\overline{\mathbb{R}}$ is called *sub-differentiable* at $x\in X$ if $F(x)\neq\pm\infty$ and there exists $x^*\in X^*$, a so-called *sub-gradient*, s.t.

$$F(y)\geq F(x)+\langle x^*,y-x\rangle_X \quad \text{for all } y\in X\,. \qquad (*)$$

Then, the *sub-differential* $\partial F\colon X\to 2^{X^*}$, for every $x\in X$, is defined by

$$\partial F(x)\coloneqq\begin{cases}\emptyset & \text{if } F(x)=\pm\infty\,,\\ \{x^*\in X^*\mid (*) \text{ applies}\} & \text{else}\;.\end{cases}$$

- **Remarks:**
  - **Consistency:** If $F\colon X\to\mathbb{R}$ is convex and Gâteaux-differentiable at $x\in X$, then
    $$\partial F(x)=\{DF(x)\}\,.$$
  - **Optimality relation:** If $F\colon X\to\overline{\mathbb{R}}$ is proper, then
    $$F(x)=\inf_{y\in X}F(y)\quad\Leftrightarrow\quad 0\in\partial F(x)\,.$$

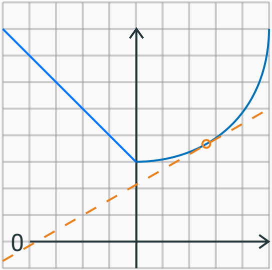


Figure: Geometric interpretation in 1D.

Sub-differential

Definition: (sub-differential)

Let X be a Banach space. Then, $F\colon X\to\overline{\mathbb{R}}$ is called *sub-differentiable* at $x\in X$ if $F(x)\neq\pm\infty$ and there exists $x^*\in X^*$, a so-called *sub-gradient*, s.t.

$$F(y)\geq F(x)+\langle x^*,y-x\rangle_X\quad\text{for all } y\in X\,. \qquad (*)$$

Then, the *sub-differential* $\partial F\colon X\to 2^{X^*}$, for every $x\in X$, is defined by

$$\partial F(x)\coloneqq\begin{cases}\emptyset & \text{if } F(x)=\pm\infty\,,\\ \{x^*\in X^*\mid (*)\text{ applies}\} & \text{else}\,.\end{cases}$$

- **Remarks:**
  - **Consistency:** If $F\colon X\to\mathbb{R}$ is convex and Gâteaux-differentiable at $x\in X$, then

$$\partial F(x)=\{DF(x)\}\,.$$

  - **Optimality relation:** If $F\colon X\to\overline{\mathbb{R}}$ is proper, then

$$F(x)=\inf_{y\in X}F(y)\quad\Leftrightarrow\quad 0\in\partial F(x)\,.$$

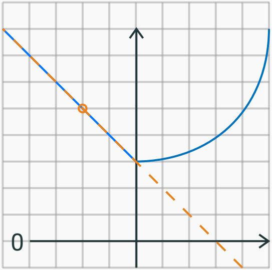


Figure: Geometric interpretation in 1D.

Definition: (sub-differential)

Let X be a Banach space. Then, $F\colon X \to \overline{\mathbb{R}}$ is called *sub-differentiable* at $x \in X$ if $F(x) \neq \pm\infty$ and there exists $x^* \in X^*$, a so-called *sub-gradient*, s.t.

$$F(y) \geq F(x) + \langle x^*, y - x\rangle_X \quad \text{for all } y \in X\,. \qquad (*)$$

Then, the *sub-differential* $\partial F\colon X \to 2^{X^*}$, for every $x \in X$, is defined by

$$\partial F(x) \coloneqq \begin{cases} \emptyset & \text{if } F(x) = \pm\infty\,, \\ \{x^* \in X^* \mid (*) \text{ applies}\} & \text{else}\,. \end{cases}$$

Remarks:

- **Consistency:** If $F\colon X \to \mathbb{R}$ is convex and Gâteaux-differentiable at $x \in X$, then

$$\partial F(x) = \{DF(x)\}\,.$$

- **Optimality relation:** If $F\colon X \to \overline{\mathbb{R}}$ is proper, then

$$F(x) = \inf_{y \in X} F(y) \quad \Leftrightarrow \quad 0 \in \partial F(x)\,.$$

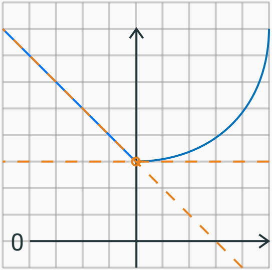


Figure: Geometric interpretation in 1D.

Sub-differential

Definition: (sub-differential)

Let X be a Banach space. Then, $F\colon X\to\overline{\mathbb{R}}$ is called *sub-differentiable* at $x\in X$ if $F(x)\neq\pm\infty$ and there exists $x^*\in X^*$, a so-called *sub-gradient*, s.t.

$$F(y)\geq F(x)+\langle x^*,y-x\rangle_X \quad \text{for all } y\in X\,. \qquad (*)$$

Then, the *sub-differential* $\partial F\colon X\to 2^{X^*}$, for every $x\in X$, is defined by

$$\partial F(x):=\begin{cases}\emptyset & \text{if } F(x)=\pm\infty\,,\\ \{x^*\in X^*\mid (*)\text{ applies}\} & \text{else}\ .\end{cases}$$

Remarks:

- **Consistency:** If $F\colon X\to\mathbb{R}$ is convex and Gâteaux-differentiable at $x\in X$, then

$$\partial F(x)=\{DF(x)\}\,.$$

- **Optimality relation:** If $F\colon X\to\overline{\mathbb{R}}$ is proper, then

$$F(x)=\inf_{y\in X}F(y) \quad\Leftrightarrow\quad 0\in\partial F(x)\,.$$

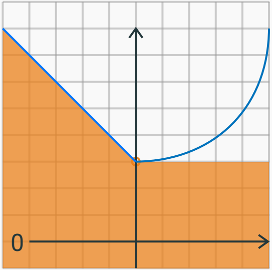


Figure: Geometric interpretation in 1D.

Definition: (bi-conjugate)

Let X be a Banach space and $F\colon X \to \overline{\mathbb{R}}$ a functional.

Then, the *Fenchel bi-conjugate* $F^{**}\colon X \to \overline{\mathbb{R}}$, for every $x \in X$, is defined by

$$F^{**}(x) \coloneqq \sup_{x^* \in X^*} \left\{ \langle x^*, x \rangle_{X^*} - F^*(x^*) \right\}.$$

- **Remarks:**
  - **Properties:** For every $x \in X$, it holds that

$$\begin{aligned} F^{**}(x) &= \sup_{x^* \in \operatorname{dom} F^*} \left\{ \langle x^*, x \rangle_X - F^*(x^*) \right\} \\ &\leq F(x)\,, \end{aligned}$$

    i.e., $F^{**} \in \Gamma(X)$;
  - **Fenchel–Moreau Theorem:** If $F \in \Gamma(X)$, then for every $x \in X$, we have that

$$F(x) = F^{**}(x) = \sup_{x^* \in X^*} \left\{ \langle x^*, x \rangle_{X^*} - F^*(x^*) \right\}.$$

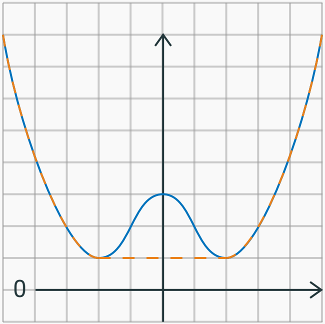


Figure: Double-well potential F (blue) and bi-conjugate F^{**} (orange).

Lagrange duality theory

Definition: (saddle point)

Let X, Y be Banach spaces and $L\colon X\times Y^*\to\overline{\mathbb{R}}$ a *Lagrange functional*.

A tuple $(u,z^*)^\top\in X\times Y^*$ is called *saddle point* of $L\colon X\times Y^*\to\overline{\mathbb{R}}$ if

$$\inf_{x\in X}\sup_{y^*\in Y^*} L(x,y^*) = L(u,z^*) = \sup_{y^*\in Y^*}\inf_{x\in X} L(x,y^*)\,.$$

- **Neumann's minimax theorem:**

(L.1) For every $y^*\in Y^*$, it holds that

$$L(\cdot,y^*)\in\Gamma_0(X)\,;$$

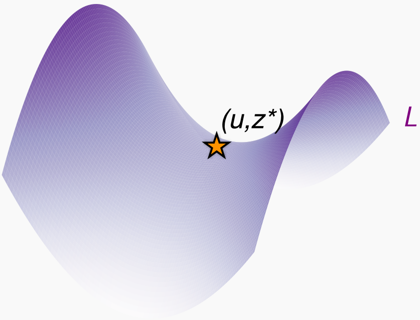
(L.2) For every $x\in X$, it holds that

$$-L(x,\cdot)\in\Gamma_0(Y^*)\,;$$

(L.3) There exists $y_0^*\in Y^*$ (or $x_0\in X$) s.t.

$$L(x,y_0^*)\to\infty\qquad(\|x\|_X\to\infty)\,,$$

$$(\text{or}\quad -L(x_0,y^*)\to\infty\qquad(\|y\|_{Y^*}\to\infty))\,.$$



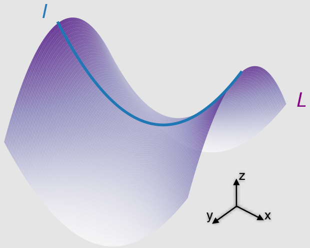
Definition: (primal problem & dual problem)

Let X, Y be Banach spaces and $L\colon X \times Y^* \to \overline{\mathbb{R}}$ a *Lagrange functional*.

- The *primal problem* is given via the minimization of $I\colon X \to \overline{\mathbb{R}}$, for every $x \in X$ defined by

$$I(x) \coloneqq \sup_{y^* \in Y^*} L(x, y^*)\,.$$

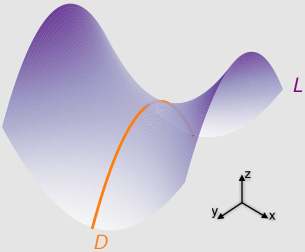
A minimizer $u \in X$ is called *primal solution*;



- The *dual problem* is given via the maximization of $D\colon Y^* \to \overline{\mathbb{R}}$, for every $y^* \in Y^*$ defined by

$$D(y^*) \coloneqq \inf_{x \in X} L(x, y^*)\,.$$

A maximizer $z^* \in Y^*$ is called *dual solution*.



Theorem: (weak duality & strong duality)

Let X, Y be Banach spaces and $L\colon X \times Y^* \to \overline{\mathbb{R}}$ a *Lagrange functional*.

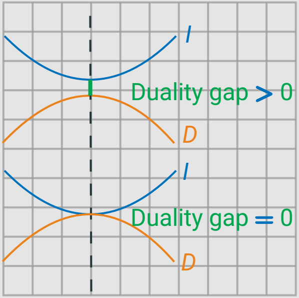
(i) A *weak duality relation* applies, *i.e.*,

$$\inf_{x\in X} I(x) \geq \sup_{y^*\in Y^*} D(y^*)\,;$$

(ii) A *strong duality relation* applies, *i.e.*,

$$I(u) = D(z^*)\,,$$

if and only if $(u,z^*)^\top \in X \times Y^*$ is a saddle point of $L\colon X \times Y^* \to \overline{\mathbb{R}}$.



♦ **Proof.**

ad (i).

$$\begin{aligned}\inf_{x\in X} I(x) &= \inf_{x\in X}\sup_{y^*\in Y^*} L(x,y^*)\\ &\geq \sup_{y^*\in Y^*}\inf_{x\in X} L(x,y^*)\\ &= \sup_{y^*\in Y^*} D(y^*)\,.\end{aligned}$$

$$\left(\sup_{y^*\in Y^*} L(x,y^*) \geq \sup_{y^*\in Y^*}\inf_{x\in X} L(x,y^*)\right)$$

Theorem: (weak duality & strong duality)

Let X, Y be Banach spaces and $L : X \times Y^* \to \overline{\mathbb{R}}$ a *Lagrange functional*.

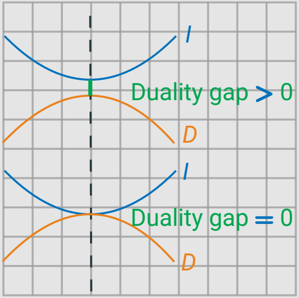
(i) A *weak duality relation* applies, *i.e.*,

$$\inf_{x\in X} I(x) \geq \sup_{y^*\in Y^*} D(y^*)\,;$$

(ii) A *strong duality relation* applies, *i.e.*,

$$I(u) = D(z^*)\,,$$

if and only if $(u, z^*)^\top \in X \times Y^*$ is a saddle point of $L : X \times Y^* \to \overline{\mathbb{R}}$.



Proof.

ad (ii). Due to **(i)**, it holds that

$$I(u) = D(z^*) \quad\Leftrightarrow\quad \sup_{y^*\in Y^*} L(u, y^*) = \inf_{x\in X} L(x, z^*)$$

$$\Leftrightarrow\quad \inf_{x\in X}\sup_{y^*\in Y^*} L(x, y^*) = L(u, z^*) = \sup_{y^*\in Y^*}\inf_{x\in X} L(x, y)\,. \quad ■$$

Fenchel duality theory

(Fenchel) primal problem

Definition: ((Fenchel) primal problem)

Let X, Y be Banach spaces, $F \in \Gamma_0(X)$, $G \in \Gamma_0(Y)$, and $\Lambda \in L(X;Y)$.

The *(Fenchel) primal problem* is given via the minimization of $I\colon X \to \mathbb{R} \cup \{+\infty\}$, for every $x \in X$ defined by

$$I(x) \coloneqq F(x) + G(\Lambda x)\,.$$

Example: Poisson problem

- $F \in \Gamma_0(W_D^{1,2}(\Omega))$, for every $v \in W_D^{1,2}(\Omega)$, defined by
$$F(v) \coloneqq -\int_\Omega f\, v \,\mathrm{d}x\,.$$
- $G \in \Gamma_0((L^2(\Omega))^d)$, for every $y \in (L^2(\Omega))^d$, defined by
$$G(y) \coloneqq \frac{1}{2}\int_\Omega |y|^2 \,\mathrm{d}x\,.$$
- $\Lambda \in L(W_D^{1,2}(\Omega);(L^2(\Omega))^d)$, for every $v \in W_D^{1,2}(\Omega)$, defined by
$$\Lambda v \coloneqq \nabla v\,.$$

Figure: Moritz Werner Fenchel (3 May 1905 – 24 January 1988).

Proposition: ((Fenchel) Lagrange functional)

Let X, Y be Banach spaces, $F \in \Gamma_0(X)$, $G \in \Gamma_0(Y)$, and $\Lambda \in L(X;Y)$.

Then, the *Lagrange functional* $L\colon X \times Y^* \to \overline{\mathbb{R}}$, for every $(x,y)^\top \in X \times Y^*$ defined by

$$L(x,y^*) \coloneqq F(x) + \langle y^*, \Lambda x\rangle_Y - G^*(y^*)\,,$$

induces the (Fenchel) primal problem, *i.e.*, for every $x \in X$, it holds that

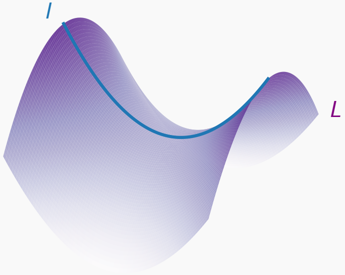
$$I(x) = \sup_{y^* \in Y^*} L(x,y^*)\,.$$

Proof. For every $y \in Y$, we have that

$$\begin{aligned} G(y) &= G^{**}(y) \\ &\coloneqq \sup_{y^* \in Y^*} \left\{\langle y^*, y\rangle_Y - G^*(y^*)\right\}, \end{aligned}$$

so that for every $x \in X$, it follows that

$$\begin{aligned} I(x) &\coloneqq F(x) + G(\Lambda x) \\ &= F(x) + \sup_{y^* \in Y^*} \left\{\langle y^*, \Lambda x\rangle_Y - G^*(y^*)\right\} \\ &= \sup_{y^* \in Y^*} \left\{F(x) + \langle y^*, \Lambda x\rangle_Y - G^*(y^*)\right\}. \blacksquare \end{aligned}$$



(Fenchel) dual problem

Proposition: ((Fenchel) dual problem)

Let X, Y be Banach spaces, $F \in \Gamma_0(X)$, $G \in \Gamma_0(Y)$, and $\Lambda \in L(X;Y)$.

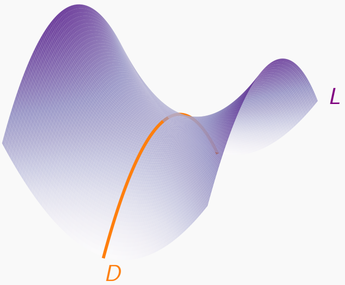
Then, the *(Fenchel) dual problem* is given via the maximization of $D\colon Y^*\to\mathbb{R}\cup\{-\infty\}$, for every $y^* \in Y^*$ defined by

$$D(y^*) \coloneqq -F^*(-\Lambda^* y^*) - G^*(y^*)\,.$$

- **Proof.** For every $y^* \in Y^*$, it holds that

$$\begin{aligned} D(y^*) &\coloneqq \inf_{x\in X} L(x,y^*) \\ &= \inf_{x\in X}\{F(x) + \langle y^*, \Lambda x\rangle_Y - G^*(y^*)\} \\ &= \inf_{x\in X}\{F(x) + \langle y^*, \Lambda x\rangle_Y\} - G^*(y^*) \\ &= \inf_{x\in X}\{F(x) + \langle \Lambda^* y^*, x\rangle_X\} - G^*(y^*) \\ &= -\sup_{x\in X}\{\langle -\Lambda^* y^*, x\rangle_X - F(x)\} - G^*(y^*) \\ &= -F^*(-\Lambda^* y^*) - G^*(y^*)\,. \end{aligned}$$

■



Lemma: (strong duality relation ⇔ convex optimality relations)

Let X, Y be Banach spaces, $F \in \Gamma_0(X)$, $G \in \Gamma_0(Y)$, and $\Lambda \in L(X;Y)$.

For $u \in X$ and $z^* \in Y^*$, the following statements are equivalent:

(i) A *strong duality relation* applies, *i.e.*, it holds that

$$I(u) = D(z^*)\,;$$

(ii) *Convex optimality relations* apply, *i.e.*, it holds that

$$G^*(z^*) - \langle z^*, \Lambda u\rangle_Y + G(\Lambda u) = 0\,,$$
$$F^*(-\Lambda^* z^*) - \langle -\Lambda^* z^*, u\rangle_X + F(u) = 0\,.$$

Proof. By the *Fenchel–Young inequality* and $\langle z^*, \Lambda u\rangle_Y = \langle \Lambda^* z^*, u\rangle_X$, it holds that

$$\text{(i)} \quad \Leftrightarrow \quad 0 = I(u) - D(z^*)\,.$$

$$\Leftrightarrow \quad 0 = \underbrace{\{G^*(z^*) - \langle z^*, \Lambda u\rangle_Y + G(\Lambda u)\}}_{\geq 0} + \underbrace{\{F^*(-\Lambda^* z^*) - \langle -\Lambda^* z^*, u\rangle_X + F(u)\}}_{\geq 0}\,.$$

$$\Leftrightarrow \quad \text{(ii)}\,.$$

■

Fenchel's duality theorem

Theorem: (Fenchel's duality theorem, *cf.* [1, Ekeland & Temam, '99])

Let X, Y be Banach spaces, $F \in \Gamma_0(X)$, $G \in \Gamma_0(Y)$, and $\Lambda \in L(X;Y)$.

If there exists $x_0 \in \mathrm{dom}(F)$ with $y_0 := \Lambda x_0 \in \mathrm{dom}(G)$ s.t.

$$y \to y_0 \quad \text{in } Y \qquad \Rightarrow \qquad G(y) \to G(y_0)\,,$$

then there exists a dual solution $z^* \in Y^*$ and

$$\inf_{v \in X} I(v) = D(z^*)\,.$$

Corollary: (*cf.* [1, Ekeland & Temam, '99])

Let the assumptions of Fenchel's duality theorem I be satisfied.

If X is reflexive and $I\colon X \to \mathbb{R} \cup \{+\infty\}$ is *(weakly) coercive*, *i.e.*,

$$\|x\|_X \to +\infty \qquad \Rightarrow \qquad I(x) \to +\infty\,,$$

then there exists a primal solution $u \in X$, a dual solution $z^* \in Y^*$, and

$$I(u) = D(z^*)\,.$$

Thank You for today!

Examples of Fenchel conjugates

Example: (p-Dirichlet density)

For $p \in (1, +\infty)$, let $F\colon X = \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, for every $x \in \mathbb{R}^d$, be defined by

$$F(x) \coloneqq \frac{1}{p}|x|^p\,.$$

Then, $F^*\colon X^* \cong \mathbb{R}^d \to \mathbb{R}$, for every $x^* \in \mathbb{R}^d$, is given via

$$F^*(x^*) = \frac{1}{p'}|x^*|^{p'}\,. \qquad (p' = \tfrac{p}{p-1})$$

Proof. For every $x^* \in \mathbb{R}^d$, we have that

$$F^*(x^*) \coloneqq \sup_{x \in \mathbb{R}^d}\left\{x^* \cdot x - \frac{1}{p}|x|^p\right\}.$$

The supremum is attained at $x_0 \in X$ if

$$0 \overset{!}{=} \frac{\mathrm{d}}{\mathrm{d}x}\left\{x^* \cdot x - \frac{1}{p}|x|^p\right\}\Big|_{x=x_0} = x^* - |x_0|^{p-2}x_0\,.$$

$$\Leftrightarrow \quad x_0 = |x^*|^{p'-2}x^*\,.$$

Inserting $x_0 \in X$, we obtain

$$F^*(x^*) = x^* \cdot x_0 - \frac{1}{p}|x_0|^p = \frac{1}{p'}|x^*|^{p'}\,. \qquad \blacksquare$$

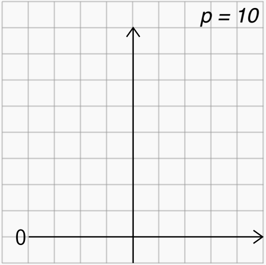


Figure: F (blue) and F^* (orange).

Example: (inequality constraint)

Let $F\colon X=\mathbb{R}\to\mathbb{R}$, for every $x\in\mathbb{R}$, be defined by

$$F(x) \coloneqq I_+(x) \coloneqq \begin{cases} 0 & \text{if } x\geq 0\,, \\ +\infty & \text{else}\,. \end{cases}$$

Then, $F^*\colon X^*\cong\mathbb{R}\to\mathbb{R}$, for every $x^*\in\mathbb{R}$, is given via

$$F^*(x^*) = I_-(x^*) \coloneqq \begin{cases} 0 & \text{if } x^*\leq 0\,, \\ +\infty & \text{else}\,. \end{cases}$$

Proof. For every $x^*\in\mathbb{R}$, we have that

$$\begin{aligned} F^*(x^*) &\coloneqq \sup_{x\in\mathbb{R}}\{x^*x - I_+(x)\} \\ &= \sup_{x\in\mathbb{R}_{\geq 0}} x^*x \\ &= \begin{cases} 0 & \text{if } x^*\leq 0\,, \\ +\infty & \text{else}\,. \end{cases} \\ &\eqqcolon I_-(x^*)\,. \end{aligned}$$

■

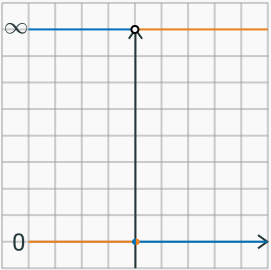


Figure: F (blue) and F^* (orange).

Example: (linear functional)

For $m \in \mathbb{R}$, let $F\colon X = \mathbb{R} \to \mathbb{R}$, for every $x \in \mathbb{R}$, be defined by

$$F(x) \coloneqq m\,x\,.$$

Then, $F^*\colon X^* \cong \mathbb{R} \to \mathbb{R}$, for every $x^* \in \mathbb{R}$, is given via

$$F^*(x^*) \coloneqq I_{\{m\}}(x^*) \coloneqq \begin{cases} 0 & \text{if } x^* = m\,, \\ +\infty & \text{else}\ . \end{cases}$$

Proof. For every $x^* \in \mathbb{R}$, we have that

$$\begin{aligned} F^*(x^*) &\coloneqq \sup_{x\in\mathbb{R}} \{x^*\,x - m\,x\} \\ &= \sup_{x\in\mathbb{R}} (x^* - m)\,x \\ &= \begin{cases} 0 & \text{if } x^* = m\,, \\ +\infty & \text{else}\ . \end{cases} \\ &\eqqcolon I_{\{m\}}(x^*)\,. \end{aligned}$$

■

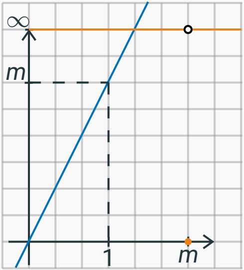


Figure: F (blue) and F^* (orange).

Literature recommendations

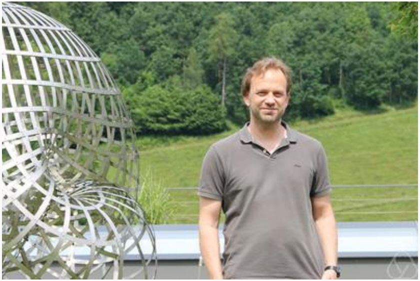


Figure: Sören Bartels

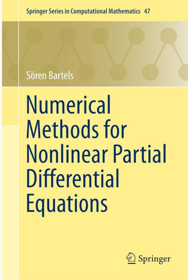


Figure: his book

Literature recommendations:

- S. Bartels, *Numerical methods for nonlinear partial differential equations*, **Springer Series in Computational Mathematics** 47, Springer, Cham, 2015. doi: 10.1007/978-3-319-13797-1.
- S. Bartels and M. Milicevic, *Primal-dual gap estimators for a posteriori error analysis of nonsmooth minimization problems*, **ESAIM Math. Model. Numer. Anal.** 54 (2020), 1635–1660.doi: 10.1051/m2an/2019074.
- S. Bartels and Z. Wang, *Orthogonality relations of Crouzeix-Raviart and Raviart-Thomas finite element spaces*, **Numer. Math.** 148 no. 1 (2021), 127–139. doi: 10.1007/s00211-021-01199-3.
- S. Bartels, *Nonconforming discretizations of convex minimization problems and precise relations to mixed methods*, **Comput. Math. Appl.** 93 (2021), 214–229. doi: 10.1016/j.camwa.2021.04.014.

Literature recommendations

Figure: Sergey I. Repin

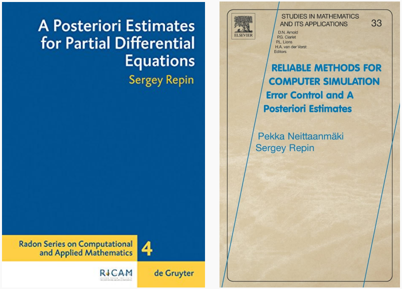


Figure: his books

Literature recommendations:

- S. I. Repin, *A posteriori error estimation for variational problems with uniformly convex functionals*, **Math. Comp.** 69 no. 230 (2000), 481–500. doi: 10.1090/S0025-5718-99-01190-4.
- P. Neittaanmäki and S. Repin, *Reliable methods for computer simulation*, **Studies in Mathematics and its Applications** 33, Elsevier, Amsterdam, 2004.
- S. Repin, *A posteriori estimates for partial differential equations*, **Radon Series on Computational and Applied Mathematics** 4, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
- S. Repin and J. Valdman, *Error identities for variational problems with obstacles*, **ZAMM Z. Angew. Math. Mech.** 98 no. 4 (2018), 635–658. doi: 10.1002/zamm.201700105.

References i

I. Ekeland and R. Témam.
Convex analysis and variational problems, volume 28 of Classics in Applied Mathematics.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999.
Translated from the French.

S. G. Mikhlin.
Variational methods in mathematical physics.
Pergamon, Oxford, 1964.

W. Prager and J. L. Synge.
Approximations in elasticity based on the concept of function space.
Quart. Appl. Math., 5:241–269, 1947.