TECHNISCHE UNIVERSITÄT
BERGAKADEMIE FREIBERG
The University of Resources. Since 1765.

Contributions to complementarity and bilevel programming in Banach spaces

By the Faculty of Mathematics and Computer Sciences
of the Technische Universität Bergakademie Freiberg

approved

Thesis

to attain the academic degree of

Doctor rerum naturalium
(Dr. rer. nat.)

submitted by **M.Sc. Patrick Mehlitz**

born on the 21st December, 1988 in Spremberg, Germany

Assessor: Prof. Dr. Stephan Dempe, Freiberg (Germany)
Prof. Dr. Juan-Juan Ye, Victoria (Canada)
Prof. Dr. Matthias Gerdts, Munich (Germany)

Date of the award:Freiberg, 07th July, 2017

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von folgenden Personen erhalten:

- Professor Dr. rer. nat. habil. Stephan Dempe
- Dr. rer. nat. Gerd Wachsmuth

Weitere Personen waren an der Abfassung der vorliegenden Arbeit nicht beteiligt.

Die Hilfe eines Promotionsberaters habe ich nicht in Anspruch genommen. Weitere Personen haben von mir keine geldwerten Leistungen für Arbeiten erhalten, die nicht als solche kenntlich gemacht worden sind. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

M.Sc. Patrick Mehlitz

Declaration

I hereby declare that I completed this work without any improper help from a third party and without using any aids other than those cited. All ideas derived directly or indirectly from other sources are identified as such.

In the selection and use of materials and in the writing of the manuscript I received support from the following persons:

- Professor Dr. rer. nat. habil. Stephan Dempe
- Dr. rer. nat. Gerd Wachsmuth

Persons other than those above did not contribute to the writing of this thesis.

I did not seek the help of a professional doctorate-consultant. Only those persons identified as having done so received any financial payment from me for any work done for me. This thesis has not previously been published in the same or a similar form in Germany or abroad.

07th July, 2017 M.Sc. Patrick Mehlitz

Acknowledgement

Foremost, I would like to express my deepest gratitude to my supervisor Prof. Dr. Stephan Dempe for his indispensable mentorship and his continuous support throughout the last years. From the beginning of my studies at the Technische Universität Bergakademie Freiberg he guided me through the different facets of mathematical optimization. He introduced me to the topic of bilevel programming and motivated the study of this rather difficult problem class in the abstract setting of Banach spaces although we both only had narrow ideas how the final results may look like. Constantly, he pointed my attention to scientists working in this area and to relevant books and papers.

During the last two years, I enjoyed a pretty intense collaboration with Dr. Gerd Wachsmuth who introduced me to complementarity programming in Banach spaces. Many times, I benefited from his deep knowledge of functional analysis and optimization theory, and always he gave ear to my problems. Together we created two papers and many more ideas which hopefully result in some more publications in the future. My profound gratitude goes to him.

A special thanks I devote to Dr. Francisco Javier Benita Maldonado. He came to Freiberg in 2014 as a PhD student in order to work together with Prof. Dr. Stephan Dempe on the *natural gas cash-out problem* for one year. This induced our research on bilevel optimal control problems of ordinary differential equations with finite-dimensional lower level which already led to three publications and Francisco's PhD thesis.

Furthermore, I am honored by the willingness of Prof. Dr. Jane Ye and Prof. Dr. Matthias Gerdts to be the co-referees of this thesis. Especially, I thank Prof. Dr. Gerdts for enabling a short research stay at the Universität der Bundeswehr in October 2016.

Many colleagues and staff members of the Department of Mathematics and Computer Science of the Technische Universität Bergakademie Freiberg deserve my special thanks for creating a good working environment. Particularly, I want to mention Prof. Dr. Udo Hebisch who willingly accepted me as his scientific assistant and under whose guidance I gave several tutorials on mathematical algebra, number theory, as well as coding theory. This way, it was possible to clear up my mind in situations where my head was overflowing with mathematical problems related to my PhD project. Gladly, I benefited from his valuable experiences in teaching and bringing mathematics not only to the heads but also to the hearts of students. On the other hand, I want to thank Dr. Maria Pilecka for proofreading my thesis.

Last but not least, I would like to sincerely thank my family for their constant support during the last years. Especially, I owe my mother a thank-you for proofreading the Chapters 1 and 6 of this thesis. Finally, I thank you, Julia, for your emotional support and patience.

Contents

Abbreviations

General abbreviations

a.e.	almost everywhere	
e.g.	exempli gratia	
f.a.a.	for almost all	
i.e.	id est	
w.l.o.g.	without loss of generality	
w.r.t.	with respect to	
CEL	compactly epi-Lipschitz	20
ODE	ordinary differential equation	
PDE	partial differential equation	
SNC	sequentially normally compact	20

Constraint qualifications

KRZCQ	Kurcyusz Robinson Zowe constraint qualification	30, Remark 2.33
LICQ	linear independence constraint qualification	31, Remark 2.34
MFCQ	Mangasarian Fromovitz constraint qualification	30, Remark 2.33
MPCC-LICQ	tailored linear independence constraint qualification for standard complementarity problems	51, Remark 3.5
MPCC-MFCQ	tailored Mangasarian Fomovitz constraint qualification for standard complementarity problems	51, Remark 3.5
NNAMCQ	no nonzero abnormal multiplier constraint qualification	30, Remark 2.33
SDPMPCC-LICQ	tailored linear independence constraint qualification for semidefinite complementarity problems	104
SDPMPCC-MFCQ	tailored Mangasarian Fromovitz constraint qualification for semidefinite complementarity problems	104
SKRZC	strict Kurcyusz Robinson Zowe condition	31, Remark 2.34

Optimization problems

BOC	bilevel optimal control problem	107
BPP	bilevel programming problem	67
KKT	Karush-Kuhn-Tucker reformulation of the bilevel programming problem	82
MPCC	mathematical problem with complementarity constraints	47
OC	optimal control problem	117
OV	optimal value reformulation of the bilevel programming problem	68
RNLP	relaxed nonlinear problem	48
TNLP	tightened nonlinear problem	48

Notation

Banach spaces

$\mathcal{X}$	a Banach space
0	zero vector of a Banach space or the scalar zero
$\|\cdot\|_{\mathcal{X}}$	norm of the Banach space $\mathcal{X}$
$\|\cdot\|_p$	p-norm of the Banach space $\mathbb{R}^n$
$x \cdot y$	Euclidean inner product of two vectors $x, y \in \mathbb{R}^n$
$\mathcal{X}^\star$	topological dual of the Banach space $\mathcal{X}$
$\prod_{i=1}^n \mathcal{X}_i$	product space induced by the Banach spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$ equipped with the sum norm
$\langle \cdot, \cdot \rangle_{\mathcal{X}}$	dual pairing of the Banach space $\mathcal{X}$
$\mathbb{U}_{\mathcal{X}}$	open unit ball of the Banach space $\mathcal{X}$
$\mathbb{U}_{\mathcal{X}}^{\varepsilon}(\bar{x})$	open ball with radius ε around $\bar{x} \in \mathcal{X}$
$\mathbb{U}_{n,p}^{\varepsilon}(\bar{x})$	open ball with radius ε around $\bar{x} \in \mathbb{R}^n$ w.r.t. the p-norm
$\mathbb{B}_{\mathcal{X}}$	closed unit ball of the Banach space $\mathcal{X}$
$\mathbb{B}_{\mathcal{X}}^{\varepsilon}(\bar{x})$	closed ball with radius ε around $\bar{x} \in \mathcal{X}$
$\mathbb{B}_{n,p}^{\varepsilon}(\bar{x})$	closed ball with radius ε around $\bar{x} \in \mathbb{R}^n$ w.r.t. the p-norm
$\mathcal{X} \hookrightarrow \mathcal{Y}$	the Banach space $\mathcal{X}$ is continuously embedded into the Banach space $\mathcal{Y}$
$\{x_k\} \subseteq \mathcal{X}$	a sequence of vectors from the Banach space $\mathcal{X}$
$x_k \to \bar{x}$	a sequence $\{x_k\} \subseteq \mathcal{X}$ converging (strongly) to $\bar{x} \in \mathcal{X}$
$x_k \rightharpoonup \bar{x}$	a sequence $\{x_k\} \subseteq \mathcal{X}$ converging weakly to $\bar{x} \in \mathcal{X}$
$x_k^* \overset{\star}{\to} x^*$	a sequence $\{x_k^*\} \subseteq \mathcal{X}^\star$ converging weakly* to $x^* \in \mathcal{X}^\star$
$t_k \searrow 0$	a sequence $\{t_k\} \subseteq \mathbb{R}^+$ converging to 0
$t_k \downarrow 0$	a sequence $\{t_k\} \subseteq \mathbb{R}_0^+$ converging to 0

Operators

$\mathbb{L}[\mathcal{X}, \mathcal{Y}]$	Banach space of all bounded, linear operators mapping from a Banach space $\mathcal{X}$ to a Banach space $\mathcal{Y}$
$\mathtt{F}$	an element of $\mathbb{L}[\mathcal{X}, \mathcal{Y}]$
$\mathtt{F}^\star$	adjoint operator of $\mathtt{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$
$\mathtt{F}[A]$	image of $A \subseteq \mathcal{X}$ under $\mathtt{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$
$\mathtt{I}_{\mathcal{X}}$	identical operator of the Banach space $\mathcal{X}$
$\mathtt{0}$	appropriate zero operator

Sets and set operations

$\mathbb{N}$	natural numbers (without zero)
$\mathbb{N}_0$	natural numbers with zero
$\mathbb{Q}^+$	positive rational numbers
$\mathbb{R}$	real numbers
$\overline{\mathbb{R}}$	extended real line
$\mathbb{R}^+$	positive real numbers
$\mathbb{R}_0^+$	nonnegative real numbers
$\mathbb{R}^n$	set of all real vectors with n components
$\mathbb{R}^{n,+}$	set of all vectors from $\mathbb{R}^n$ with positive components
$\mathbb{R}_0^{n,+}$	set of all vectors from $\mathbb{R}^n$ with nonnegative components
Ξ_n	unit simplex in $\mathbb{R}^n$

$[a, b]$	a closed interval in $\mathbb{R}$
$(a, b]$	a half-open interval in $\mathbb{R}$
(a, b)	an open interval in $\mathbb{R}$
$A + A'$	Minkowski sum of sets $A, A' \subseteq \mathcal{X}$ in a Banach space $\mathcal{X}$
$A \setminus A'$	set of all elements from A which do not belong to A' where $A, A' \subseteq \mathcal{X}$
$X \times Y$	Cartesian product of sets X and Y
2^X	power set of X
sA	s-fold of a set $A \subseteq \mathcal{X}$
$\operatorname{cl} A$	closure of a set $A \subseteq \mathcal{X}$
$\operatorname{cl}^w A$	weak closure of a set $A \subseteq \mathcal{X}$
$\operatorname{cl}^w_{\mathsf{seq}} A$	weak sequential closure of a set $A \subseteq \mathcal{X}$
$\operatorname{cl}^\star B$	weak* closure of a set $B \subseteq \mathcal{X}^\star$
$\operatorname{int} A$	interior of a set $A \subseteq \mathcal{X}$
$\operatorname{rint} A$	relative interior of a set $A \subseteq \mathcal{X}$
$\operatorname{bd} A$	boundary of a set $A \subseteq \mathcal{X}$
$\operatorname{lin} A$	linear hull of a set $A \subseteq \mathcal{X}$, i.e. the smallest linear subspace of $\mathcal{X}$ comprising A
$\operatorname{conv} A$	convex hull of a set $A \subseteq \mathcal{X}$, i.e. the smallest convex set in $\mathcal{X}$ comprising A
$\overline{\operatorname{conv}} A$	closed, convex hull of a set $A \subseteq \mathcal{X}$, i.e. the smallest closed, convex set in $\mathcal{X}$ comprising A
$\operatorname{cone} A$	conic hull of a set $A \subseteq \mathcal{X}$, i.e. the smallest convex cone in $\mathcal{X}$ comprising A
$\overline{\operatorname{cone}} A$	closed, conic hull of a set $A \subseteq \mathcal{X}$, i.e. the smallest closed, convex cone in $\mathcal{X}$ comprising A
$\operatorname{proj}_A(x)$	projection of $x \in \mathcal{X}$ onto the nonempty, closed, convex set $A \subseteq \mathcal{X}$
A°	polar cone of $A \subseteq \mathcal{X}$
$A^\perp$	annihilator of $A \subseteq \mathcal{X}$
$B_\circ$	reverse polar cone of $B \subseteq \mathcal{X}^\star$
$B_\perp$	reverse annihilator of $B \subseteq \mathcal{X}^\star$
$\limsup^w_{k\to\infty} A_k$	weak sequential upper limit of a sequence of sets $\{A_k\} \subseteq 2^{\mathcal{X}}$
$\liminf_{k\to\infty} A_k$	(strong) sequential lower limit of a sequence of sets $\{A_k\} \subseteq 2^{\mathcal{X}}$

Partially ordered sets

(S, ϱ)	a partially ordered set, i.e. a set S equipped with a partial order $\varrho \subseteq S \times S$
$\leq_K$	partial order induced by the closed, convex, pointed cone $K \subseteq \mathcal{X}$
$\leq$	partial order induced by the cone $\mathbb{R}^{n,+}_0$, i.e. the common componentwise less-or-equal relation in $\mathbb{R}^n$
$\max_K\{x; y\}$	supremum of $x, y \in \mathcal{X}$ w.r.t. the partial order $\leq_K$
$\min_K\{x; y\}$	infimum of $x, y \in \mathcal{X}$ w.r.t. the partial order $\leq_K$

Measurability

$\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$	measure space induced by the σ-algebra Σ on Ω with measure $\mathfrak{m}$
$N(\mathfrak{M})$	sets of measure zero in the measure space $\mathfrak{M}$
$\mathcal{L}^0(\mathfrak{M}, \mathbb{R}^n)$	set of all measurable functions mapping from Ω to $\mathbb{R}^n$
$L^0(\mathfrak{M}, \mathbb{R}^n)$	set of all equivalence classes of measurable functions mapping from Ω to $\mathbb{R}^n$
$\Sigma_1 \otimes \Sigma_2$	smallest σ-algebra comprising the Cartesian product $\Sigma_1 \times \Sigma_2$ of the σ-algebras Σ_1 and Σ_2
$\mathfrak{B}(X)$	Borelean σ-algebra induced by the metric space X
$\mathfrak{B}^m$	Borelean σ-algebra induced by the Banach space $\mathbb{R}^m$
$\mathfrak{l}$	the Lebesgue measure
δ_x	Dirac measure of the singleton $\{x\}$
$\lvert\mathfrak{m}\rvert$	total variation of the measure $\mathfrak{m}$

χ_A	characteristic function of the set A
$\mathcal{Z}(A)$	system of all finite and disjoint partitions of $A \in \mathfrak{B}(X)$ w.r.t. the σ-algebra $\mathfrak{B}(X)$
$\mathcal{M}(X)$	Banach space of all signed, regular measures w.r.t. the measurable space $(X, \mathfrak{B}(X))$
$\mathbb{E}$, $\mathbb{K}$	decomposable sets
r_K	auxiliary function characterizing the convex hull of $K \subseteq \mathbb{R}^m$

Matrices

$S^{m\times n}$	set of all matrices with m rows and n columns whose entries come from $S \subseteq \mathbb{R}$
$\mathcal{S}_m$	set of all symmetric matrices from $\mathbb{R}^{m\times m}$
$\mathcal{O}_m$	set of all orthogonal matrices from $\mathbb{R}^{m\times m}$
$\mathbf{A}$	a matrix from $\mathbb{R}^{m\times n}$
$\mathbf{A}^\top$	transposed of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$
$\mathbf{A}^\dagger$	pseudo inverse matrix of $\mathbf{A} \in \mathbb{R}^{m\times n}$
$\mathbf{A}_{I,J}$	submatrix of $\mathbf{A} \in \mathbb{R}^{m\times n}$ whose rows are indexed by $I \subseteq \{1,\ldots,m\}$ and whose columns are indexed by $J \subseteq \{1,\ldots,n\}$
$\mathbf{A} \bullet \mathbf{A}'$	Hadamard product of two matrices $\mathbf{A}, \mathbf{A}' \in \mathbb{R}^{m\times n}$
$\mathbf{B}^{-1}$	inverse of a regular matrix $\mathbf{B} \in \mathbb{R}^{m\times m}$
$\operatorname{tr}\mathbf{B}$	trace of a matrix $\mathbf{B} \in \mathbb{R}^{m\times m}$
$\mathbf{B}^\mathbf{Q}$	the matrix product $\mathbf{Q}^\top\mathbf{B}\mathbf{Q}$ for $\mathbf{B} \in \mathbb{R}^{m\times m}$ and $\mathbf{Q} \in \mathcal{O}_m$
$\mathbf{B}^\mathbf{Q}_{IJ}$	the submatrix $(\mathbf{B}^\mathbf{Q})_{IJ}$ for $\mathbf{B} \in \mathbb{R}^{m\times m}$, $\mathbf{Q} \in \mathcal{O}_m$, and $I, J \subseteq \{1,\ldots,m\}$
α, β, γ	index sets corresponding to the positive, zero, and negative eigenvalues associated with an ordered eigenvalue decomposition of a matrix $\mathbf{B} \in \mathcal{S}_m$
$\mathbf{I}_m$	identity matrix from $\mathbb{R}^{m\times m}$
$\mathbf{O}$	zero matrix of appropriate dimensions
$\mathbf{E}$	all-ones matrix of appropriate dimensions
$\mathbf{e}^1, \ldots, \mathbf{e}^n$	the n unit vectors in $\mathbb{R}^n$

Function spaces

Ω	domain in $\mathbb{R}^d$
$\overline{\Omega}$	closure of $\Omega \subseteq \mathbb{R}^d$
D^α	the differential operator $D_1^{\alpha_1}\ldots D_d^{\alpha_d}$ with a multiindex $\alpha \in \mathbb{N}_0^d$ and $D_i := \frac{\partial}{\partial\omega_i}$, $i = 1,\ldots,d$, defined for functions mapping from $\Omega \subseteq \mathbb{R}^d$ to $\mathbb{R}$
Δ	the Laplacian operator
∇	the gradient operator
∇^2	the Hessian operator
∇_{ω_I}	the partial gradient operator w.r.t. the variables ω_i with $i \in I \subseteq \{1,\ldots,d\}$
$\nabla^2_{\omega_I\omega_J}$	the partial Hessian w.r.t. the variables ω_j with $j \in J \subseteq \{1,\ldots,d\}$ and the partial gradient ∇_{ω_I}
$L^p(\mathfrak{M}, \mathbb{R}^n)$	Lebesgue space of all p-integrable functions from $L^0(\mathfrak{M}, \mathbb{R}^n)$ with $1 \le p \le \infty$
$L^p(\mathfrak{M})$	Lebesgue space $L^p(\mathfrak{M}, \mathbb{R})$
$L^p(\Omega, \mathbb{R}^n)$	Lebesgue space of all p-integrable functions from $L^0(\mathfrak{M}, \mathbb{R}^n)$ with measure space $\mathfrak{M} = (\Omega, \mathfrak{B}(\Omega), \mathfrak{l})$
$L^p(\Omega)$	Lebesgue space $L^p(\mathfrak{M}, \mathbb{R})$ with measure space $\mathfrak{M} = (\Omega, \mathfrak{B}(\Omega), \mathfrak{l})$
p'	conjugate coefficient of p, i.e. the real number from the interval $(1, \infty]$ which satisfies $1/p + 1/p' = 1$ where $1 \le p < \infty$
$C^k(X)$	vector space of all k-times continuously differentiable functions mapping from the metric space $X \subseteq \mathbb{R}^d$ to $\mathbb{R}$
$C_0^k(X)$	vector space of all functions from $C^k(X)$ with support in X which is compact in $\mathbb{R}^d$
$C(X)$	vector space of all continuous functions from X to $\mathbb{R}$, i.e. $C^0(X)$
$C_0(X)$	vector space $C_0^0(X)$

$L^1_{\mathrm{loc}}(\Omega)$	vector space of all locally Lebesgue integrable functions mapping from Ω to $\mathbb{R}$
$W^{k,p}(\Omega)$	Sobolev space of all order k weakly differentiable functions mapping from Ω to $\mathbb{R}$ whose weak derivatives belong to $L^p(\Omega)$
$W^{k,p}_0(\Omega)$	closure of $C^\infty_0(\Omega)$ w.r.t. the Sobolev norm in $W^{k,p}(\Omega)$
$H^k(\Omega)$	Sobolev space $W^{k,2}(\Omega)$
$H^k_0(\Omega)$	Sobolev space $W^{k,2}_0(\Omega)$
$H^{-1}(\Omega)$	the dual of $H^1_0(\Omega)$
$AC(\Omega)$	vector space of all absolutely continuous functions mapping from $\overline{\Omega}$ to $\mathbb{R}$ where $\Omega := (0,T) \subseteq \mathbb{R}$ holds
$AC^{1,2}(\Omega,\mathbb{R}^n)$	Banach space of all functions with n components coming from $AC(\Omega)$ and possessing weak first-order derivatives in $L^2(\Omega)$

Cones

$\mathcal{R}_A(\bar{x})$	radial cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{T}_A(\bar{x})$	tangent cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{T}^\flat_A(\bar{x})$	inner tangent cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{T}^w_A(\bar{x})$	weak tangent cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{T}^c_A(\bar{x})$	Clarke tangent cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{K}_A(\bar{x},x^*)$	critical cone to a set $A \subseteq \mathcal{X}$ w.r.t. $\bar{x} \in A$ and $x^* \in \mathcal{T}_A(\bar{x})^\circ$
$\widehat{\mathcal{N}}^\sigma_A(\bar{x})$	set of all Fréchet σ-normals to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\widehat{\mathcal{N}}_A(\bar{x})$	Fréchet normal cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{N}_A(\bar{x})$	limiting normal cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{N}^s_A(\bar{x})$	strong limiting normal cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{N}^c_A(\bar{x})$	Clarke normal cone to a set $A \subseteq \mathcal{X}$ at $\bar{x} \in A$
$\mathcal{S}^+_m$	cone of all positive semidefinite matrices from $\mathcal{S}_m$
$\mathcal{S}^{++}_m$	cone of all positive definite matrices from $\mathcal{S}_m$
$\mathcal{S}^-_m$	cone of all negative semidefinite matrices from $\mathcal{S}_m$
$\mathcal{K}_m$	second-order cone in $\mathbb{R} \times \mathbb{R}^m$
$\mathcal{K}_\mathcal{H}$	$\mathcal{H}$-second-order cone in $\mathbb{R} \times \mathcal{H}$ where $\mathcal{H}$ is a Hilbert space
$C\left(\overline{\Omega}\right)^+_0$	cone of all nonnegative functions from $C\left(\overline{\Omega}\right)$
$L^p(\mathfrak{M})^+_0$	cone of all almost everywhere nonnegative functions from $L^p(\mathfrak{M})$
$L^p(\Omega)^+_0$	cone of all almost everywhere nonnegative functions from $L^p(\Omega)$
$W^{1,p}(\Omega)^+_0$	cone of all almost everywhere nonnegative functions from $W^{1,p}(\Omega)$
$H^1_0(\Omega)^+_0$	cone of all almost everywhere nonnegative functions from $H^1_0(\Omega)$
$AC^{1,2}(\Omega)^+_0$	cone of all almost everywhere nonnegative functions from $AC^{1,2}(\Omega,\mathbb{R})$

Functions and set-valued mappings

$F\colon \mathcal{X} \to \mathcal{Y}$	a function mapping from a Banach space $\mathcal{X}$ to a Banach space $\mathcal{Y}$
$\psi\colon \mathcal{X} \to \overline{\mathbb{R}}$	a functional mapping from a Banach space $\mathcal{X}$ to the extended real line
$\Psi\colon \mathcal{X} \rightrightarrows \mathcal{Y}$	a set-valued function mapping from a Banach space $\mathcal{X}$ to the power set $2^\mathcal{Y}$ of a Banach space $\mathcal{Y}$
$F'(\bar{x};\delta)$	directional derivative of F at $\bar{x}$ in direction δ
$F'(\bar{x})$	Fréchet derivative of F at $\bar{x}$
$F^{(2)}(\bar{x})$	second-order Fréchet derivative of F at $\bar{x}$
$F'_{x_i}(\bar{x})$	partial Fréchet derivative of F at $\bar{x}$ w.r.t. x_i
$F^{(2)}_{x_ix_j}(\bar{x})$	second-order partial Fréchet derivative of F at $\bar{x}$ w.r.t. x_i and x_j, i.e. the operator $(F'_{x_i})'_{x_j}(\bar{x})$
$\operatorname{epi}\psi$	epigraph of a functional ψ
$\psi^\circ(\bar{x};\delta)$	Clarke's generalized directional derivative of ψ at $\bar{x}$ in direction δ
$\partial^c\psi(\bar{x})$	Clarke subdifferential of ψ at $\bar{x}$
$\partial\psi(\bar{x})$	limiting subdifferential of ψ at $\bar{x}$

$\operatorname{gph}\Psi$	graph of the set-valued mapping Ψ
$\ker\Psi$	kernel of the set-valued mapping Ψ
$\Psi^{-1}\colon \mathcal{Y}\rightrightarrows\mathcal{X}$	inverse set-valued mapping of Ψ
$D^*_N\Psi(\bar{x},\bar{y})$	normal coderivative of Ψ at $(\bar{x},\bar{y})\in\operatorname{gph}\Psi$
$o\colon \mathbb{R}^+_0\to\mathcal{X}$	a function with the property $\lim_{t\searrow 0}\frac{o(t)}{t}=0$
$H\circ G$	composition of mappings $G\colon X\to Y$ and $H\colon Y\to Z$ where X,Y,Z are nonempty sets

Bilevel programming

X_{ad}	upper level feasible set
$\varphi\colon \mathcal{X}\to\overline{\mathbb{R}}$	lower level optimal value function
$\Psi\colon \mathcal{X}\rightrightarrows\mathcal{Y}$	lower level solution set mapping
$\Lambda(\bar{x},\bar{y})$	set of lower level regular Lagrange multipliers associated with $(\bar{x},\bar{y})\in\operatorname{gph}\Psi$

1. Introduction

The aim of this thesis is the derivation of necessary optimality conditions for bilevel programming problems in the rather general setting of Banach spaces. Therefore, we consider the model program

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ G(x) &\in C \\ y &\in \Psi(x) \end{aligned} \tag{1.1}$$

where $\Psi\colon \mathcal{X} \rightrightarrows \mathcal{Y}$ is a set-valued mapping between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ which assigns to every $x \in \mathcal{X}$ the (possibly empty) set of global optimal solutions of the parametric optimization problem

$$\begin{aligned} f(x,y) &\to \min_{y} \\ g(x,y) &\in K. \end{aligned} \tag{1.2}$$

Therein, $F, f\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, $G\colon \mathcal{X} \to \mathcal{W}$, and $g\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ are sufficiently smooth mappings, $\mathcal{W}$ as well as $\mathcal{Z}$ are Banach spaces, and $C \subseteq \mathcal{W}$ as well as $K \subseteq \mathcal{Z}$ are nonempty, closed, convex sets.
The original idea of bilevel programming dates back to Stackelberg who described a game between two players whose cost functions and sets of admissible strategies depend on the respective other player's decision in 1934, see [115]. In this thesis, we think of the following special decision order in (1.1): First, the upper level player or so-called leader, i.e. the decision maker in (1.1), chooses an instance of x which satisfies the constraint $G(x) \in C$. Next, the lower level player or so-called follower, i.e. the decision maker in (1.2), is capable of computing the set $\Psi(x)$ which is passed back to the upper level decision maker. This way the leader can determine the overall feasible set of (1.1) which allows him to solve the bilevel programming problem since he can evaluate his objective functional for every admissible point. Note that this procedure is related to the so-called optimistic approach to bilevel programming where the follower only passes back such $y \in \Psi(x)$ which minimize the leader's objective for fixed x over the set $\Psi(x)$, see [140, Proposition 5.3.1]. This reflects a cooperative behavior between the two players. The situation where the leader does not know how the follower will react on his choice of x is far more delicate. In order to minimize the damage caused by the follower's decision, the leader has to assume that (in the worst case) the follower passes back only those $y \in \Psi(x)$ which maximize his objective for fixed x over the set $\Psi(x)$. This idea on how to interpret a bilevel programming problem is called pessimistic approach and can be used as well in order to model a competitive behavior of leader and follower. It is also possible to interpret the classical bilevel optimization model where the leader minimizes his objective functional only w.r.t. x as an instance of set optimization where the objective $\mathcal{F}(x) := \bigcup_{y \in \Psi(x)} \{F(x,y)\}$ has to be minimized, see [101].

During the last decades, standard bilevel programming problems, i.e. problems of type (1.1) where all appearing Banach spaces are instances of $\mathbb{R}^n$, were studied extensively from a theoretical and numerical point of view, see the monographs [8, 26, 36, 113] and the references therein. Hierarchical decision structures appear in numerous applications from economics, logistics, engineering, or natural sciences.
Many models naturally comprise (ordinary or partial) differential equations (ODEs and PDEs for short) in at least one decision level, see [3, 4, 18, 58, 59, 74, 76, 77, 89]. A special instance of the bilevel programming problem in function spaces is the obstacle problem studied in [57, 64, 66, 67, 68, 73, 88, 96, 123, 125] and other papers by these authors where the solution of a variational inequality of the first kind has to be controlled. More precisely, for a bounded domain $\Omega \subseteq \mathbb{R}^d$ with sufficiently smooth boundary, a nonempty, closed, convex set $X_{\mathrm{ad}} \subseteq L^2(\Omega)$, a desired state $y_{\mathrm{d}} \in L^2(\Omega)$, and a regularization

parameter $\sigma > 0$, the obstacle problem in hierarchical form is given by

$$\begin{aligned} \tfrac{1}{2}\,\|y - y_d\|^2_{L^2(\Omega)} + \tfrac{\sigma}{2}\,\|x\|^2_{L^2(\Omega)} &\to \min_{x,y} \\ x &\in X_{\mathsf{ad}} \\ y &\in \Psi(x) \end{aligned} \tag{1.3}$$

where $\Psi\colon L^2(\Omega) \rightrightarrows H^1_0(\Omega)$ denotes the solution mapping of the parametric optimization problem

$$\begin{aligned} \tfrac{1}{2}\int_\Omega |\nabla y(\omega)|_2^2\,\mathrm{d}\omega - \int_\Omega x(\omega)y(\omega)\mathrm{d}\omega &\to \min_{y} \\ y(\omega) &\geq 0 \qquad \text{a.e. on } \Omega. \end{aligned} \tag{1.4}$$

Therein, $x \in L^2(\Omega)$ is the so-called control variable while $y \in H^1_0(\Omega)$ is referred to as state variable. Of course, using an appropriate discretization, one can easily transfer a so-called bilevel optimal control problem into a standard bilevel optimization problem which can be discussed with the aid of results from finite-dimensional programming (*first-discretize-then-optimize* approach), but from the current point of view it is unclear in many situations how the solutions of the discretized problem and the original bilevel optimal control problem are related. Thus, it seems reasonable to study (1.1) in the setting of Banach spaces which covers the situation where function spaces are under consideration. In this thesis, we aim for the derivation of necessary optimality conditions for (1.1) in the generalized setting. In terms of bilevel optimal control, this can be seen as the starting point of a *first-optimize-then-discretize* scheme for the numerical treatment of (1.1): First, the optimal solutions are characterized in the function space setting. Afterwards, the feasibility and optimality conditions can be discretized and the arising finite-dimensional and possibly nonlinear system can be treated numerically. For a detailed discussion of the *first-discretize-then-optimize* and the *first-optimize-then-discretize* approaches for single-level optimal control problems of PDEs, we refer the interested reader to [55] and [103].
Suppose that a standard bilevel programming problem is under consideration which comprises an unknown parameter p. Using the concept of *fuzziness*, this situation is discussed in [109]. Another possible way to handle the uncertainty is provided by the robust optimization approach, see e.g. [11, 12]. Given a so-called uncertainty set P, the appearing objective functions are replaced by their corresponding suprema over P w.r.t. the parameter p, whereas the constraints have to hold for all instances of $p \in P$. Depending on the structure of P, the *robustification* of the bilevel programming problem (1.1) is a hierarchical optimization problem comprising e.g. semidefinite cone constraints or second-order cone constraints, see [12]. Thus, it is reasonable to study bilevel programming problems in the vector space of symmetric matrices. Noting that the symmetric matrices form a Banach space and that the set of all positive semidefinite matrices is a closed, convex cone in this space, the general formulation (1.1) covers this situation as well.

In order to derive necessary optimality conditions for (1.1), we need to transfer the hierarchical model into a single-level surrogate problem. Let us briefly discuss the three approaches we use in this thesis for that purpose, see [98] as well.
First, we assume that the lower level solution set $\Psi(x)$ equals the singleton $\{\psi(x)\}$ for all $x \in \mathcal{X}$. Then it is possible to plug the function ψ into the upper level objective F, directly. Now, the resulting surrogate problem only has to be minimized w.r.t. the upper level variable x. For that purpose, it is necessary to study the (generalized) differentiability properties of the function ψ. In [25], the author investigates standard bilevel programming problems with a unique lower level solution. Under appropriate constraint qualifications the resulting lower level solution function is directionally differentiable, see e.g. [27, 104], and its directional derivative is continuous w.r.t. the direction which allows the formulation of necessary and sufficient optimality conditions. Reformulating the obstacle problem as the bilevel programming problem (1.3), it is well-known that its lower level problem (1.4) possesses a unique solution which is directionally differentiable and Lipschitz continuous w.r.t. the upper level variable, see [57, 75, 88]. This property is used in several papers to derive necessary optimality conditions for the underlying optimal control problem with and without control constraints, see [66, 67, 68, 123, 125]. In [84], we study a bilevel programming problem whose lower level is given by a formal parametric optimal control problem. We show that the lower level solution is unique and directionally differentiable under appropriate assumptions. Furthermore, it is shown that this directional derivative equals the solution of a variational inequality of the first kind.

Next, let us suppose that the lower level problem (1.2) is convex w.r.t. the follower's decision variable and sufficiently regular, whereas K is a closed, convex cone. Then the condition $y \in \Psi(x)$ in (1.1) can be equivalently replaced by some Karush-Kuhn-Tucker-type optimality condition (forthwith, we will use the abbreviation KKT for Karush-Kuhn-Tucker for brevity). This gives rise to the consideration of the surrogate problem

$$\begin{aligned} F(x,y) &\to \min_{x,y,\lambda} \\ G(x) &\in C \\ f_y'(x,y) + g_y'(x,y)^\star[\lambda] &= 0 \\ g(x,y) &\in K \\ \lambda &\in K^\circ \\ \langle \lambda, g(x,y) \rangle_{\mathcal{Z}} &= 0 \end{aligned} \tag{1.5}$$

where λ is the lower level Lagrange multiplier, $K^\circ \subseteq \mathcal{Z}^\star$ denotes the polar cone of K, and the mapping $\langle \cdot, \cdot \rangle_{\mathcal{Z}} : \mathcal{Z}^\star \times \mathcal{Z} \to \mathbb{R}$ represents the dual pairing in the Banach space $\mathcal{Z}$. The final three constraints in (1.5) form a so-called complementarity condition. In [28], it is shown that the surrogate problem (1.5) is not necessarily equivalent to (1.1) w.r.t. local optimal solutions anymore. Using this KKT approach, the authors derive necessary optimality conditions for standard bilevel optimization problems in [32, 33, 34, 140]. In [87], some of these results are generalized to the Banach space setting and applied to state necessary optimality conditions of Pontryagin-type for a bilevel optimal control problem with control constraints in the lower level problem. Note that under the postulated convexity assumptions it is also possible to replace the lower level problem (1.2) equivalently by a generalized equation or variational inequality. Necessary optimality conditions for mathematical problems with variational inequality constraints are widely studied in the finite-dimensional setting, see [23, 97, 116, 128, 133] and the references therein. Generalizations to the setting of Banach spaces can be found in [90] and several other publications of this author. We already mentioned a number of papers where optimal control problems comprising variational inequalities are under consideration.

Introducing the lower level optimal value by $\varphi(x) := \inf_y \{ f(x,y) \,|\, g(x,y) \in K \}$ for all parameters $x \in \mathcal{X}$, we can exploit the so-called optimal value function $\varphi \colon \mathcal{X} \to \overline{\mathbb{R}}$ of (1.2) in order to equivalently reformulate (1.1) by

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ G(x) &\in C \\ f(x,y) - \varphi(x) &\leq 0 \\ g(x,y) &\in K \end{aligned} \tag{1.6}$$

without any additional assumption. On the other hand, the function φ is likely to be nonsmooth or even discontinuous, see [7]. Anyway, under appropriate assumptions one can guarantee that φ is locally Lipschitz continuous or convex, and it is possible to compute its generalized derivative or subdifferential (in an appropriate sense) in terms of initial data, see [31, 84, 91, 93]. Thus, using some nonsmooth calculus, necessary optimality conditions of KKT-type for (1.1) via (1.6) seem to be within reach. However, it is shown for standard bilevel programming problems in [35] that constraint qualifications of reasonable strength fail to hold at all feasible points of (1.6) and similar difficulties are likely to appear in the generalized setting. In [14, 84], the authors show in different situations that Fritz-John-type necessary optimality conditions directly derived from (1.6) are degenerated in the sense that they hold at all feasible points of the bilevel programming problem (1.1). In order to overcome this difficulty, the authors propose a partial penalization approach w.r.t. the crucial constraint $f(x,y) - \varphi(x) \leq 0$ in [138]. Namely, they introduce the famous concept of *partial calmness* which is used to derive necessary optimality conditions for standard bilevel optimization problems in [29, 34, 92, 138], for bilevel programming problems with semidefinite lower level in [30], for bilevel optimal control problems of ODEs with finite-dimensional lower level in [13, 14, 74], for bilevel optimal control problems where leader and follower share a common dynamical system of ODEs at the lower level in [130, 131], and for bilevel optimal control problems of ODEs with control constrained optimal control problems at both decision levels in [84].

Thinking of a standard bilevel programming problem where K equals the nonpositive orthant, problem (1.5) is a special instance of a so-called mathematical program with complementarity constraints, MPCC for short. It is well-known from the literature, see e.g. [139, Proposition 1.1], that standard constraint qualifications like the Mangasarian Fromovitz constraint qualification, MFCQ for short, fail to hold at all feasible points of such an MPCC. Consequently, the KKT conditions turn out to be a too selective necessary optimality criterion for this problem class. Thus, the formulation of weaker stationarity notions and appropriate constraint qualifications became a hot topic during the last two decades, see [42, 43, 44, 45, 80, 111, 129, 132] and the references therein. Furthermore, the numerical approach to standard MPCCs is well-developed, see [4, 46, 70].
We mentioned earlier that the setting where K equals the positive semidefinite or second-order cone is reasonable as well. This justifies the study of a more general class of complementarity problems. In this thesis, we will consider the situation where the complementarity condition is given by

$$\begin{aligned} \tilde{G}(x) &\in K \\ \tilde{H}(x) &\in K^\circ \\ \langle \tilde{H}(x), \tilde{G}(x) \rangle_{\mathcal{Z}} &= 0 \end{aligned} \tag{1.7}$$

where $\tilde{G}\colon \mathcal{X} \to \mathcal{Z}$ and $\tilde{H}\colon \mathcal{X} \to \mathcal{Z}^\star$ are sufficiently smooth maps between Banach spaces and K is a closed, convex cone. In order to ensure the symmetry of the complementarity condition, we will assume that $\mathcal{Z}$ is reflexive. In [37, 119, 124, 127], the authors investigate the situation where K equals the positive semidefinite cone. Assuming that K is given by the second-order cone, such complementarity problems are discussed in [48, 79, 124, 135]. One may also think of optimal control problems with mixed control-state complementarity constraints which can be used to model switching effects, see [56]. Then the complementarity is induced by the cone of almost everywhere nonnegative functions in a Lebesgue space, see [86] where we discuss this setting. Finally, optimal control problems of ODEs with terminal complementarity constraints are the subject of discussion in [15]. A far more general consideration of optimization problems in Banach spaces which comprise constraints of the form (1.7) can be found in [87, 121, 124].

In Chapter 2 of this thesis, we present all the necessary fundamentals of mathematical programming in Banach spaces we need for our subsequent considerations. First, we briefly recall the basic terminology of functional analysis. Afterwards, we introduce all the Banach spaces which are considered here. Especially, we deal with certain function spaces in the sense of Lebesgue and Sobolev and list the required embedding theorems from [1]. Section 2.3.1 is dedicated to the study of different concepts of tangents and normals to a set. Here we also recall the famous notion of polyhedricity which dates back to [57, 88]. Furthermore, we study the property of sets in function spaces to be sequentially normally compact, see [90], which plays an important role for the variational calculus of Mordukhovich. In order to compare certain stationarity notions for general MPCCs, the algebraic concept of vector lattices is introduced in Section 2.3.2. In the subsequent section, we study some tools of generalized differentiation which are used in Section 2.3.4 to formulate necessary optimality conditions for different classes of single-level optimization problems in Banach spaces. Many optimal control problems are equipped with pointwise control constraints which are induced by a (in a certain sense) measurable set-valued mapping with not necessarily convex images. In order to state necessary optimality conditions for such problems, one needs to know more about the variational geometry of pointwise defined sets in Lebesgue spaces. Section 2.3.5 deals with this issue. We derive explicit formulae for several tangent and normal cones to these sets and obtain reasonable lower and upper bounds for the corresponding limiting normal cone. Here the argumentation parallels our considerations in [86].
Chapter 3 is dedicated to the investigation of general complementarity problems in Banach spaces. In Section 3.1, we derive the concepts of weak and Mordukhovich stationarity for abstract MPCCs and study corresponding constraint qualifications which ensure that a local minimizer of the problem of interest satisfies these stationarity conditions. Here we mainly follow our considerations in [47, 87]. Furthermore, we recall the concept of strong stationarity known from [121]. Finally, we study the relationship between weak, Mordukhovich, and strong stationarity. As we will see, the polyhedricity of the underlying complementarity cone is essential for our results which can be used to discuss MPCCs in Lebesgue spaces and Sobolev spaces. Afterwards, we apply our findings to the situations where the cone which induces the complementarity equals the set of all almost everywhere nonnegative functions in a Lebesgue space or is

polyhedral. Using our results on the variational geometry of pointwise defined sets in Lebesgue spaces, we will be able to show the surprising fact that weak and Mordukhovich stationarity coincide for MPCCs in Lebesgue spaces, see [86].
Necessary optimality conditions for the bilevel programming problem (1.1) are derived in Chapter 4. We start with the consideration of a bilevel model whose lower level problem represents a rather general parametric optimal control problem with a unique solution for all instances of the parameter. For the theoretical background, we investigate a special type of nonsmooth parametric equation in Section 4.1.1. We show that it possesses a unique solution for any choice of the parameter, and it is proven that the corresponding solution mapping is directionally differentiable under appropriate assumptions. We characterize the directional derivative as the solution of another nonsmooth equation which can be equivalently stated as a complementarity system. In Section 4.1.2, we apply these general results and our knowledge on abstract MPCCs from Chapter 3 to the bilevel programming problem of interest. In the case where the lower level control constraints are induced by a cone, it is possible to apply our findings from Chapter 3 directly to the bilevel programming problem. We state the resulting optimality conditions and compare them to the ones derived via the directional differentiability of the lower level solution mapping. An example from optimal control illustrates the theory. In Section 4.2, we focus our attention on the KKT reformulation (1.5) of the bilevel programming problem (1.1). First, the relationship of these two problems is discussed in detail. An example shows that the argumentation in [28] for local optimal solutions cannot be generalized to the abstract setting if $\mathcal{Z}$ is infinite-dimensional. Nevertheless, we present some conditions under which both problems are equivalent w.r.t. local optimal solutions. Partially following [87], we derive necessary optimality conditions for the bilevel programming problem (1.1) via the MPCC reformulation (1.5) in Section 4.2.2. Therefore, we exploit our results for general MPCCs from Chapter 3 again. We close the section on abstract bilevel programming in Banach spaces with the investigation of the optimal value transformation (1.6) in Section 4.3. First, we show that reasonable constraint qualifications from the theory of programming in Banach spaces fail to hold at all feasible points of this program. Afterwards, we restate the well-known concept of partial calmness, see [138], in the abstract setting in order to derive necessary optimality conditions for the bilevel programming problem (1.1). Therefore, we exploit some of our results on the Lipschitz continuity of the optimal value function in parametric optimization, see [31].
Using materials from Chapters 3 and 4, we derive necessary optimality conditions for three different classes of bilevel programming problems in Chapter 5. We start with the consideration of the problem class we already dealt with in Section 4.1 where the lower level control constraint is realized by demanding the control to come from the cone of positive semidefinite matrices in Section 5.1. After gathering some necessary results from variational analysis in the Banach space of symmetric matrices in Section 5.1.1, we apply our findings from Section 4.1 to the model problem in Section 5.1.2. Furthermore, we compare our results to the already existing ones on semidefinite complementarity programming [37, 124, 127] w.r.t. the presented optimality conditions and constraint qualifications. Our considerations deepen the results presented in [84]. In the subsequent Section 5.2, we consider a bilevel optimal control problem of ODEs with optimal control problems at both decision levels and lower level control constraints. After stating all the assumptions we need in order to apply our findings from Sections 3.2 and 4.2, we formulate the lower level necessary and sufficient optimality conditions of Pontryagin-type, see [102], and derive the KKT reformulation of the given bilevel programming problem in Section 5.2.1. As we will see in Section 3.2, Mordukhovich's stationarity concept is not suitable to deal with the situation at hand which is why we restrict ourselves to the derivation of the weak and strong stationarity conditions of the surrogate optimization problem in Section 5.2.2. Finally, in Section 5.2.3, we construct an applicable constraint qualification which ensures that all local optimal solutions of the bilevel optimal control problem are strongly stationary. Therefore, we use the concept of controllability of linear dynamical systems, see [9]. The content of this section is partially taken from our manuscript [87]. Section 5.3 is dedicated to the study of an abstract optimal control problem with control constraints and an implicitly given pointwise state constraint which is induced by a finite-dimensional lower level. More precisely, the lower level parameter equals the realization of the state function at a certain point of the underlying domain. The *natural gas cash-out problem*, see [74], represents a typical situation where such optimal control problems arise in practice. We start the section by providing an existence result for global optimal solutions of such a bilevel optimal control problem. The remaining part of the section is dedicated to the derivation of necessary optimality conditions. First, we state an abstract optimality criterion in Section 5.3.1. We finish our considerations applying our findings to the situation where the underlying optimal control problem is governed by a linear ODE or Poisson's equation which is a PDE in Sections 5.3.2 and 5.3.3, respectively. Thereby, we continue our research on this problem class we initiated in [13, 14, 15, 74].

2. Fundamentals of mathematical programming in Banach spaces

2.1. Preliminaries from functional analysis

Let $\mathcal{X}$ be a (real) Banach space with norm $\|\cdot\|_{\mathcal{X}}$ and zero vector 0. A second norm $\|\cdot\|_{\mathcal{X}}^*$ of $\mathcal{X}$ is said to be equivalent to $\|\cdot\|_{\mathcal{X}}$ if there exist positive real constants c_1 and c_2 satisfying $c_1\|x\|_{\mathcal{X}} \leq \|x\|_{\mathcal{X}}^* \leq c_2\|x\|_{\mathcal{X}}$ for any $x \in \mathcal{X}$. It is well-known that in a finite-dimensional Banach space, all norms are equivalent. Let us denote by $\mathbb{U}_{\mathcal{X}}$ and $\mathbb{B}_{\mathcal{X}}$ the open and closed unit ball of $\mathcal{X}$, respectively. Furthermore, for any $x \in \mathcal{X}$ and any positive scalar ε, we set $\mathbb{U}_{\mathcal{X}}^{\varepsilon}(x) := \{x\} + \varepsilon\mathbb{U}_{\mathcal{X}}$ and $\mathbb{B}_{\mathcal{X}}^{\varepsilon}(x) := \{x\} + \varepsilon\mathbb{B}_{\mathcal{X}}$. Therein, for arbitrary sets $A, B \subseteq \mathcal{X}$ and a scalar s, $A + B$ denotes the sum in Minkowski's sense, whereas we define $sA := \{sa \in \mathcal{X} \,|\, a \in A\}$. Additionally, we set $A - B := A + (-B)$. A set $C \subseteq \mathcal{X}$ is called a cone if for any $\alpha \geq 0$, $\alpha C \subseteq C$ holds true. Furthermore, we exploit $\operatorname{cl} A$, $\operatorname{int} A$, $\operatorname{rint} A$, $\operatorname{bd} A$, $\operatorname{lin} A$, $\operatorname{conv} A$, $\overline{\operatorname{conv}} A$, $\operatorname{cone} A$, and $\overline{\operatorname{cone}} A$ in order to denote the closure of A, the interior of A, the relative interior of A, the boundary of A, the smallest linear subspace of $\mathcal{X}$ comprising A, the convex hull of A, the closed convex hull of A, the smallest convex cone comprising A, and the smallest closed, convex cone comprising A, respectively. The set A is said to be locally closed (locally convex) at $\bar{x} \in A$ if there exists $\varepsilon > 0$ such that $A \cap \mathbb{B}_{\mathcal{X}}^{\varepsilon}(\bar{x})$ is closed (convex). The set A is said to be dense in $B \subseteq \mathcal{X}$ if $\operatorname{cl} A = B$ holds. The Banach space $\mathcal{X}$ is called separable if there exists a countable set A which is dense in $\mathcal{X}$.

Lemma 2.1. Let $\mathcal{X}$ be a Banach space, and let $A, B \subseteq \mathcal{X}$ be nonempty. Then

$$\operatorname{cl}(A + B) = \operatorname{cl}(\operatorname{cl} A + B) = \operatorname{cl}(\operatorname{cl} A + \operatorname{cl} B)$$

is valid.

Proof. The respective inclusions $\subseteq$ are obvious. Thus, we only need to show $\operatorname{cl}(\operatorname{cl} A + \operatorname{cl} B) \subseteq \operatorname{cl}(A + B)$. Therefore, observe that $\operatorname{cl} A + \operatorname{cl} B \subseteq \operatorname{cl}(A + B)$ holds. That is why we obtain

$$\operatorname{cl}(\operatorname{cl} A + \operatorname{cl} B) \subseteq \operatorname{cl}\big(\operatorname{cl}(A + B)\big) = \operatorname{cl}(A + B)$$

from the monotonicity of the operator cl. This completes the proof. □

Lemma 2.2. For a Banach space $\mathcal{X}$ and a nonempty set $S \subseteq \mathcal{X}$, $\operatorname{lin} S = \operatorname{conv} \bigcup_{\alpha\in\mathbb{R}} \alpha S$ holds true.

Proof. We start with the proof of the inclusion $\subseteq$. Observe that due to $S \subseteq \operatorname{conv} \bigcup_{\alpha\in\mathbb{R}} \alpha S$, it is sufficient to show that $L := \operatorname{conv} \bigcup_{\alpha\in\mathbb{R}} \alpha S$ is a linear space. Therefore, take $x, y \in L$ and $c^x, c^y \in \mathbb{R}$. Then there are integers $n, m \in \mathbb{N}$, vectors $s_1^x, \ldots, s_n^x, s_1^y, \ldots, s_m^y \in S$, scalars $\alpha_1^x, \ldots, \alpha_n^x, \alpha_1^y, \ldots, \alpha_m^y \in \mathbb{R}$, and nonnegative scalars $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m \in \mathbb{R}$ such that

$$x = \sum_{j=1}^{n} \lambda_j \alpha_j^x s_j^x, \quad y = \sum_{i=1}^{m} \mu_i \alpha_i^y s_i^y, \quad \sum_{j=1}^{n} \lambda_j = 1, \quad \sum_{i=1}^{m} \mu_i = 1.$$

That is why we obtain

$$c^x x + c^y y = \sum_{j=1}^{n} \lambda_j c^x \alpha_j^x s_j^x + \sum_{i=1}^{m} \mu_i c^y \alpha_i^y s_i^y = \sum_{j=1}^{n} \tfrac{\lambda_j}{2} 2c^x \alpha_j^x s_j^x + \sum_{i=1}^{m} \tfrac{\mu_i}{2} 2c^y \alpha_i^y s_i^y \in L$$

since $\frac{\lambda_1}{2}, \dots, \frac{\lambda_n}{2}, \frac{\mu_1}{2}, \dots, \frac{\mu_m}{2} \in \mathbb{R}$ are nonnegative scalars which satisfy

$$\sum_{j=1}^{n} \frac{\lambda_j}{2} + \sum_{i=1}^{m} \frac{\mu_i}{2} = 1.$$

This shows $\operatorname{lin} S \subseteq L$. The inclusion $L \subseteq \operatorname{lin} S$ follows from $\alpha S \subseteq \operatorname{lin} S$ for any $\alpha \in \mathbb{R}$ and the convexity of $\operatorname{lin} S$. □

Now, let $\mathcal{Y}$ be another Banach space. Then the Cartesian product $\mathcal{X} \times \mathcal{Y}$ is a Banach space as well if equipped, e.g., with the sum norm induced by $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$ which will be done throughout the thesis if not stated otherwise. Thus, for $n \in \mathbb{N}$ Banach spaces $\mathcal{X}_1, \dots, \mathcal{X}_n$, we can define the product space $\prod_{j=1}^{n} \mathcal{X}_j := \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ in a similar way by recursion. Its elements are addressed by n-tupels $x = (x_1, \dots, x_n)$ with $x_j \in \mathcal{X}_j$, $j = 1, \dots, n$, or column vectors denoted by

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

depending on which representation is more suitable in the corresponding situation. Especially, when dealing with linear operators between product spaces, the use of column vectors is more convenient. Similarly, for sets $X_j \subseteq \mathcal{X}_j$, $j = 1, \dots, n$, we exploit

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \quad \begin{pmatrix} \mathcal{X}_1 \\ \vdots \\ \mathcal{X}_n \end{pmatrix}$$

in order to represent $X_1 \times \cdots \times X_n$ and $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$, respectively. If $\mathcal{X} = \mathcal{X}_1 = \cdots = \mathcal{X}_n$ holds, we use the notation $\mathcal{X}^n := \prod_{j=1}^{n} \mathcal{X}$. Similarly, for $A \subseteq \mathcal{X}$, we denote by A^n its Cartesian product of order n.

A mapping $\Phi\colon \mathcal{X} \to \mathcal{Y}$ is called an isometry between $\mathcal{X}$ and $\mathcal{Y}$ if $\|\Phi(x)\|_{\mathcal{Y}} = \|x\|_{\mathcal{X}}$ holds for all $x \in \mathcal{X}$. If an isomorphism between $\mathcal{X}$ and $\mathcal{Y}$ exists which is an isometry, these spaces are called isometrically isomorphic, $\mathcal{X} \cong \mathcal{Y}$ for short.
The real vector space $\mathbb{L}[\mathcal{X}, \mathcal{Y}]$ of all continuous (or bounded, see [126, Satz II.1.2]) linear operators mapping from $\mathcal{X}$ to $\mathcal{Y}$ equipped with the norm

$$\forall \mathrm{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]\colon \quad \|\mathrm{F}\|_{\mathbb{L}[\mathcal{X},\mathcal{Y}]} := \sup_{x \in \mathbb{B}_{\mathcal{X}}} \|\mathrm{F}[x]\|_{\mathcal{Y}},$$

forms another Banach space, see [126, Satz II.1.4]. From this definition $\|\mathrm{F}[x]\|_{\mathcal{Y}} \leq \|\mathrm{F}\|_{\mathbb{L}[\mathcal{X},\mathcal{Y}]} \|x\|_{\mathcal{X}}$ for all $\mathrm{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$ and all $x \in \mathcal{X}$ is easily seen. We can introduce $\mathcal{X}^\star := \mathbb{L}[\mathcal{X}, \mathbb{R}]$, the (topological) dual space of $\mathcal{X}$. Let us define the so-called dual pairing $\langle \cdot, \cdot \rangle_{\mathcal{X}} : \mathcal{X}^\star \times \mathcal{X} \to \mathbb{R}$ of $\mathcal{X}$ by $\langle x^*, x \rangle_{\mathcal{X}} := x^*[x]$ for all $x^* \in \mathcal{X}^\star$ and $x \in \mathcal{X}$. Obviously, this mapping is bilinear and satisfies

$$\forall x^* \in \mathcal{X}^\star \, \forall x \in \mathcal{X}\colon \quad |\langle x^*, x \rangle_{\mathcal{X}}| \leq \|x^*\|_{\mathcal{X}^\star} \|x\|_{\mathcal{X}}.$$

The above relation is called Cauchy-Schwarz inequality. Note that the dual pairing is continuous w.r.t. both of its components.

Lemma 2.3. For any Banach spaces $\mathcal{X}_1, \dots, \mathcal{X}_n$, there is an isomorphism $\Phi\colon \left(\prod_{j=1}^{n} \mathcal{X}_j\right)^\star \to \prod_{j=1}^{n} \mathcal{X}_j^\star$ with the property

$$\forall x^* \in \left(\prod_{j=1}^{n} \mathcal{X}_j\right)^\star\colon \quad \|x^*\|_{(\prod_{j=1}^{n} \mathcal{X}_j)^\star} \leq \|\Phi(x^*)\|_{\prod_{j=1}^{n} \mathcal{X}_j^\star} \leq n \, \|x^*\|_{(\prod_{j=1}^{n} \mathcal{X}_j)^\star}. \tag{2.1}$$

Proof. In order to shorten the notation, we introduce the spaces $\mathcal{X} := \prod_{j=1}^{n} \mathcal{X}_j$, $\mathcal{P} := \left(\prod_{j=1}^{n} \mathcal{X}_j\right)^\star$, and $\mathcal{Q} := \prod_{j=1}^{n} \mathcal{X}_j^\star$. For any $x^* \in \mathcal{P}$, we define

$$\Phi(x^*) := \left(x_1^*, \dots, x_n^*\right)$$

where $x_i^* \in \mathcal{X}_i^\star$, $i = 1, \dots, n$ is defined by

$$\forall x_i \in \mathcal{X}_i\colon \quad x_i^*[x_i] := x^*[0, \dots, 0, x_i, 0, \dots, 0].$$

It is easy to see that this mapping is a linear bijection, i.e. an isomorphism between $\mathcal{P}$ and $\mathcal{Q}$. In order to show the validity of the presented inequalities, recall that $\mathcal{X}$ is equipped with the sum norm induced by $\|\cdot\|_{\mathcal{X}_1}, \dots, \|\cdot\|_{\mathcal{X}_n}$. That is why we obtain

$$\|x^*\|_{\mathcal{P}} = \sup_{(x_1,\dots,x_n)\in\mathbb{B}_{\mathcal{X}}} x^*[x_1, \dots, x_n] = \sup_{(x_1,\dots,x_n)\in\mathbb{B}_{\mathcal{X}}} \sum_{i=1}^n x_i^*[x_i] \leq \sum_{i=1}^n \sup_{x_i\in\mathbb{B}_{\mathcal{X}_i}} x_i^*[x_i] = \|\Phi(x^*)\|_{\mathcal{Q}}$$

which yields the first inequality. The second one follows from

$$\|\Phi(x^*)\|_{\mathcal{Q}} = \sum_{i=1}^n \sup_{x_i\in\mathbb{B}_{\mathcal{X}_i}} x_i^*[x_i] \leq \sup_{(x_1,\dots,x_n)\in n\mathbb{B}_{\mathcal{X}}} \sum_{i=1}^n x_i^*[x_i] = \sup_{(x_1,\dots,x_n)\in\mathbb{B}_{\mathcal{X}}} n\, x^*[x_1, \dots, x_n] = n\, \|x^*\|_{\mathcal{P}}\,.$$

This completes the proof. □

We may interpret Lemma 2.3 as follows: the Banach spaces $\left(\prod_{j=1}^n \mathcal{X}_j\right)^\star$ and $\prod_{j=1}^n \mathcal{X}_j^\star$ are isomorphic and (via an isomorphism) equipped with equivalent norms. Thus, it is reasonable to identify $\left(\prod_{j=1}^n \mathcal{X}_j\right)^\star$ and $\prod_{j=1}^n \mathcal{X}_j^\star$ with each other. This will be done throughout this thesis without mentioning it again.

The continuous linear mapping $\mathcal{X} \ni x \to \langle \cdot, x\rangle_{\mathcal{X}} \in \mathcal{X}^{\star\star}$, called canonical embedding of $\mathcal{X}$, is an injective isometry between $\mathcal{X}$ and $\mathcal{X}^{\star\star}$ by the theorem of Hahn-Banach, see [126, Korollar III.1.6]. We call $\mathcal{X}$ reflexive if the corresponding canonical embedding is surjective. Thus, for any reflexive Banach space $\mathcal{X}$, we obtain $\mathcal{X} \cong \mathcal{X}^{\star\star}$, and, hence, we may identify any reflexive Banach space $\mathcal{X}$ and its bidual space $\mathcal{X}^{\star\star}$ with each other. Note that any finite-dimensional Banach space is reflexive. We say that $\mathcal{X}$ is continuously embedded in the Banach space $\mathcal{Y}$, $\mathcal{X} \hookrightarrow \mathcal{Y}$ for short, if $\mathcal{X} \subseteq \mathcal{Y}$ holds and if a positive real constant C exists, such that

$$\forall x \in \mathcal{X}\colon \quad \|x\|_{\mathcal{Y}} \leq C\, \|x\|_{\mathcal{X}}$$

holds, i.e. if the identical mapping $\mathcal{X} \ni x \to x \in \mathcal{Y}$ is an element of $\mathbb{L}[\mathcal{X}, \mathcal{Y}]$. An operator $\mathtt{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$ is called compact if the closure of $\{\mathtt{F}[x] \,|\, x \in \mathbb{B}_{\mathcal{X}}\}$ w.r.t. the norm in $\mathcal{Y}$ is compact. We say that $\mathcal{X}$ is compactly embedded in $\mathcal{Y}$ if we have $\mathcal{X} \hookrightarrow \mathcal{Y}$ and if the identical mapping $\mathcal{X} \ni x \mapsto x \in \mathcal{Y}$ is a compact operator. Let $\mathcal{H}$ be a Hilbert space with inner product $(\cdot, \cdot)_{\mathcal{H}}$. Then by means of Riesz's representation theorem the mapping $\mathcal{H} \ni x \to (\cdot, x)_{\mathcal{H}} \in \mathcal{H}^\star$ is an isometric isomorphism between $\mathcal{H}$ and $\mathcal{H}^\star$. That is why it is possible to identify the inner product of $\mathcal{H}$ and its dual pairing with each other. This will be done throughout this thesis. On the other hand, it is also possible (but not always recommendable, e.g. when discussing certain function spaces, see [118, Section 2.13.2]) to identify $\mathcal{H}$ and $\mathcal{H}^\star$ with each other. We will point out whenever this property of Hilbert spaces is exploited.

For any $\mathtt{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$, $\mathtt{F}^\star \in \mathbb{L}[\mathcal{Y}^*, \mathcal{X}^*]$ defined by the relation $\langle \mathtt{F}^\star[y^*], x\rangle_{\mathcal{X}} = \langle y^*, \mathtt{F}[x]\rangle_{\mathcal{Y}}$ for any $x \in \mathcal{X}$ and $y^* \in \mathcal{Y}^*$ is called the adjoint operator of $\mathtt{F}$. From [126, Satz III.4.2] we know that the linear mapping $\mathbb{L}[\mathcal{X}, \mathcal{Y}] \ni \mathtt{F} \to \mathtt{F}^\star \in \mathbb{L}[\mathcal{Y}^*, \mathcal{X}^*]$ is an isometry. For a reflexive Banach space $\mathcal{X}$, an operator $\mathtt{G} \in \mathbb{L}[\mathcal{X}, \mathcal{X}^*]$ is called self-adjoint if $\mathtt{G} = \mathtt{G}^\star$ holds, and $\mathtt{G}$ is called monotone or positive, if

$$\forall x \in \mathcal{X}\colon \quad \langle \mathtt{G}[x], x\rangle_{\mathcal{X}} \geq 0$$

is satisfied. We call $\mathtt{G}$ elliptic if there is a constant $\gamma > 0$, such that

$$\forall x \in \mathcal{X}\colon \quad \langle \mathtt{G}[x], x\rangle_{\mathcal{X}} \geq \gamma\, \|x\|_{\mathcal{X}}^2$$

is valid. By $\mathtt{O} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$ and $\mathtt{I}_{\mathcal{X}} \in \mathbb{L}[\mathcal{X}, \mathcal{X}]$, we denote the zero operator defined by $\mathtt{O}[x] = 0$ for all $x \in \mathcal{X}$ and the identical operator of $\mathcal{X}$, respectively. For Banach spaces $\mathcal{Y}_1, \dots, \mathcal{Y}_m$ and linear operators $\mathtt{F}_{i,j} \in \mathbb{L}[\mathcal{X}_j, \mathcal{Y}_i]$, $i = 1, \dots, m$ and $j = 1, \dots, n$, we introduce the product operator

$$\forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \begin{pmatrix} \mathcal{X}_1 \\ \vdots \\ \mathcal{X}_n \end{pmatrix}\colon \quad \begin{bmatrix} \mathtt{F}_{1,1} & \dots & \mathtt{F}_{1,n} \\ \vdots & \ddots & \vdots \\ \mathtt{F}_{m,1} & \dots & \mathtt{F}_{m,n} \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^n \mathtt{F}_{1,j}[x_j] \\ \vdots \\ \sum_{j=1}^n \mathtt{F}_{m,j}[x_j] \end{pmatrix} \in \begin{pmatrix} \mathcal{Y}_1 \\ \vdots \\ \mathcal{Y}_m \end{pmatrix},$$

and a similar notation is used to represent images of sets under product operators. Note that for brevity, we leave the brackets $[\cdot]$ around the argument away whenever $n \geq 2$ holds. Clearly, the product operator is an element of $\mathbb{L}[\prod_{j=1}^{n} \mathcal{X}_j, \prod_{i=1}^{m} \mathcal{Y}_i,]$, and we obtain the relation

$$\begin{bmatrix} \mathrm{F}_{1,1} & \dots & \mathrm{F}_{1,n} \\ \vdots & \ddots & \vdots \\ \mathrm{F}_{m,1} & \dots & \mathrm{F}_{m,n} \end{bmatrix}^{\star} = \begin{bmatrix} \mathrm{F}_{1,1}^{\star} & \dots & \mathrm{F}_{m,1}^{\star} \\ \vdots & \ddots & \vdots \\ \mathrm{F}_{1,n}^{\star} & \dots & \mathrm{F}_{m,n}^{\star} \end{bmatrix} \tag{2.2}$$

from the above definitions.

A sequence $\{x_k\} \subseteq \mathcal{X}$ converges (strongly) to some point $\bar{x} \in \mathcal{X}$, $x_k \to \bar{x}$ for short, if the real sequence $\{\|x_k - \bar{x}\|_{\mathcal{X}}\}$ converges to zero. On the other hand, $\{x_k\}$ converges weakly to $\bar{x}$, expressed by $x_k \rightharpoonup \bar{x}$, if for any $x^* \in \mathcal{X}^\star$, the real sequence $\{\langle x^*, x_k\rangle_{\mathcal{X}}\}$ converges to $\langle x^*, \bar{x}\rangle_{\mathcal{X}}$. In case of existence, the weak limit point $\bar{x}$ is uniquely determined and satisfies $\bar{x} \in \overline{\mathrm{conv}}\{x^k \,|\, k \in \mathbb{N}\}$, see [126, Satz III.3.8]. Now, choose a sequence $\{x_k^*\} \subseteq \mathcal{X}^\star$. We call it weakly* convergent to $\bar{x}^* \in \mathcal{X}^\star$, $x_k^* \overset{\star}{\rightharpoonup} \bar{x}^*$ for short, if for any $x \in \mathcal{X}$, the real sequence $\{\langle x_k^*, x\rangle_{\mathcal{X}}\}$ converges to $\langle \bar{x}^*, x\rangle_{\mathcal{X}}$. Again, if a weak* limit point exists, it is unique. Using the theorem of Banach and Steinhaus, see [126, Satz IV.2.1], we obtain that weakly and weakly* convergent sequences are bounded. It is clear from the definition that

$$x_k^* \to \bar{x}^* \implies x_k^* \rightharpoonup \bar{x}^* \implies x_k^* \overset{\star}{\rightharpoonup} \bar{x}^*$$

holds, and whenever $\mathcal{X}$ is reflexive, weak and weak* convergence in $\mathcal{X}^\star$ are equivalent. Moreover, any bounded sequence of a reflexive Banach space contains at least a weakly convergent subsequence, see [126, Satz III.3.7]. In any finite-dimensional Banach space, the notions of strong, weak, and weak* convergence coincide.

Lemma 2.4. Let $\mathcal{X}$ be a Banach space. Choose sequences $\{x_k\} \subseteq \mathcal{X}$ and $\{x_k^*\} \subseteq \mathcal{X}^\star$ such that $x_k \to x$ and $x_k^* \overset{\star}{\rightharpoonup} x^*$ ($x_k \rightharpoonup x$ and $x_k^* \to x^*$) hold. Then we have $\langle x_k^*, x_k\rangle_{\mathcal{X}} \to \langle x^*, x\rangle_{\mathcal{X}}$.

Proof. Since $\{x_k^*\}$ is weakly* convergent, it is bounded. Thus, we have

$$\lim_{k\to\infty} |\, \langle x_k^*, x_k - x\rangle_{\mathcal{X}} \,| \leq \lim_{k\to\infty} \|x_k^*\|_{\mathcal{X}^\star} \, \|x_k - x\|_{\mathcal{X}} = 0.$$

This yields

$$\lim_{k\to\infty} \langle x_k^*, x_k\rangle_{\mathcal{X}} = \lim_{k\to\infty} \left(\langle x_k^*, x_k - x\rangle_{\mathcal{X}} + \langle x_k^*, x\rangle_{\mathcal{X}}\right) = \lim_{k\to\infty} \langle x_k^*, x\rangle_{\mathcal{X}} = \langle x^*, x\rangle_{\mathcal{X}}\,.$$

The proof is similar for the situation where $x_k \rightharpoonup x$ and $x_k^* \to x^*$ hold. □

For sets $U \subseteq \mathcal{X}$ and $V \subseteq \mathcal{X}^\star$, $\mathrm{cl}^w U$ and $\mathrm{cl}^\star V$ denote the corresponding closure w.r.t. weak and weak* topology (in $\mathcal{X}$ and $\mathcal{X}^\star$, respectively), whereas $\mathrm{cl}^w_{\mathrm{seq}} U$ is used to express the weak sequential closure of U, i.e. the set of all weak accumulation points of sequences in U. Note that we have $\mathrm{cl}^w_{\mathrm{seq}} U \subseteq \mathrm{cl}^w U$ in general, see [83, Section 2.5]. We call U weakly sequentially closed (weakly closed) whenever $U = \mathrm{cl}^w_{\mathrm{seq}} U$ ($U = \mathrm{cl}^w U$) holds. Any closed, convex set is weakly sequentially closed and, by means of the famous Hahn-Banach theorem, see [126, Satz VIII.2.12], weakly closed as well. Furthermore, U is said to be weakly sequentially compact if any sequence in U contains a weakly convergent subsequence whose weak limit belongs to U. Any compact set is weakly sequentially compact but not vice versa. Additionally, it is clear that any bounded, closed, convex subset of a reflexive Banach space is weakly sequentially compact. Finally, V is called weakly* closed if $V = \mathrm{cl}^\star V$ is satisfied.

A mapping $F\colon \mathcal{X} \to \mathcal{Y}$ between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ is called continuous at $\bar{x} \in \mathcal{X}$ if the condition

$$\forall \{x_k\} \subseteq \mathcal{X}\colon \quad x_k \to \bar{x} \implies F(x_k) \to F(\bar{x})$$

is satisfied. Similarly, we call F weakly-weakly continuous at $\bar{x}$ if

$$\forall \{x_k\} \subseteq \mathcal{X}\colon \quad x_k \rightharpoonup \bar{x} \implies F(x_k) \rightharpoonup F(\bar{x})$$

holds. The mapping F is said to be continuous (weakly-weakly continuous) if it is continuous (weakly-weakly continuous) at any point in $\mathcal{X}$. Let $\varphi\colon \mathcal{X} \to \overline{\mathbb{R}}$ be a functional of $\mathcal{X}$ where $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ denotes the extended real line. Then φ is called weakly lower semicontinuous at some point $\bar{x} \in \mathcal{X}$ with $|\varphi(\bar{x})| < \infty$ if the condition

$$\forall \{x_k\} \subseteq \mathcal{X}\colon \quad x_k \rightharpoonup \bar{x} \implies \varphi(\bar{x}) \leq \liminf_{k\to\infty} \varphi(x_k)$$

is satisfied. Furthermore, we call φ weakly lower semicontinuous if it possesses this property at any point from $\mathcal{X}$ where it is finite. Clearly, if a functional is weakly-weakly continuous, it is weakly lower semicontinuous. Obviously, any weakly lower semicontinuous functional is lower semicontinuous. On the other hand, there exist continuous functionals which are not weakly lower semicontinuous. Note that for any continuous convex functional φ, the level sets $\{x \in \mathcal{X} \,|\, \varphi(x) \leq \alpha\}$ are closed and convex and, thus, weakly closed for all $\alpha \in \mathbb{R}$. Particularly, it easily follows that φ is weakly lower semicontinuous in this case, see [71, Theorem 2.5]. The following result presents a version of the famous Weierstraß theorem applicable in infinite-dimensional Banach spaces.

Lemma 2.5. Let $\mathcal{X}$ be a real Banach space, let $\varphi\colon \mathcal{X} \to \mathbb{R}$ be a weakly lower semicontinuous functional, and let $M \subseteq \mathcal{X}$ be nonempty. Then φ attains a global minimum on M provided that one of the following assumptions holds:

1. M is weakly sequentially compact,
2. $\mathcal{X}$ is reflexive, φ is coercive, i.e. it satisfies

   $$\forall \{x_k\} \subseteq \mathcal{X}\colon \quad \|x_k\|_{\mathcal{X}} \to \infty \implies \varphi(x_k) \to \infty,$$

   and M is weakly sequentially closed.

Proof. The statement of the first assertion equals [71, Theorem 2.3]. For the proof of the second claim, choose $\tilde{x} \in M$ arbitrarily and consider $\tilde{M} := \{x \in M \,|\, \varphi(x) \leq \varphi(\tilde{x})\}$. Since φ is weakly lower semicontinuous while M is weakly sequentially closed, it is easy to check that $\tilde{M}$ is weakly sequentially closed as well. Moreover, $\tilde{M}$ is bounded due to the coercivity of φ and contains $\tilde{x}$. We exploit the reflexivity of $\mathcal{X}$ to see that $\tilde{M}$ is weakly sequentially compact. Clearly, $\operatorname{Argmin}\{\varphi(x) \,|\, x \in M\} = \operatorname{Argmin}\{\varphi(x) \,|\, x \in \tilde{M}\}$ holds by definition of $\tilde{M}$. Finally, $\operatorname{Argmin}\{\varphi(x) \,|\, x \in \tilde{M}\}$ is nonempty due to the first statement of this lemma. This completes the proof. □

Let us compare the above result with the classical Weierstraß theorem.

Remark 2.6. In the classical Weierstraß theorem, lower semicontinuity of the functional φ and compactness of the set M are demanded. However, compactness is a rather restrictive property in infinite-dimensional spaces (boundedness and closedness is not sufficient for compactness anymore). Thus, this assumption has to be weakened in order to obtain an acceptable version of the Weierstraß theorem. Here weak sequential compactness pays off. However, weakening the assumptions on M, stronger properties have to be postulated on the functional φ. In the above lemma, the weak lower semicontinuity of φ is used for that purpose. Thus, we may interpret the first statement of Lemma 2.5 as a generalized version of the Weierstraß theorem. In any finite-dimensional Banach space $\mathcal{X}$, this abstract result coincides with the classical version, obviously.

2.2. Examples of Banach spaces

In this section, we are going to give a brief overview of Banach spaces which will be used in this thesis. We start with some finite-dimensional spaces. Afterwards, certain function spaces together with some well-known embedding relations will be introduced.

2.2.1. Finite-dimensional Banach spaces

Let $n \in \mathbb{N}$ be arbitrarily chosen. By $\mathbb{R}^n$ we denote the set of all real column vectors with n components. Furthermore, $\mathbb{R}_0^{n,+}$ and $\mathbb{R}^{n,+}$ represent the set of all componentwise nonnegative real vectors with n components and the set of all componentwise positive real vectors with n components, respectively. Especially, we set $\mathbb{R}_0^+ := \mathbb{R}_0^{1,+}$ and $\mathbb{R}^+ := \mathbb{R}^{1,+}$. For $1 \leq p < \infty$, the norm $|\cdot|_p$ is defined by

$$\forall x \in \mathbb{R}^n: \quad |x|_p := \left(\sum\nolimits_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$$

where $|s|$ represents the absolute value of a scalar $s \in \mathbb{R}$. Furthermore, $|\cdot|_\infty$ is introduced by

$$\forall x \in \mathbb{R}^n: \quad |x|_\infty := \max\{|x_1|; \ldots; |x_n|\}.$$

As mentioned earlier, all the norms $|\cdot|_p$, $1 \leq p \leq \infty$, are equivalent. For any such p, $\mathbb{U}_{n,p}$ and $\mathbb{B}_{n,p}$ represent the open and closed unit ball of $\mathbb{R}^n$ w.r.t. the norm $|\cdot|_p$. Choosing $x \in \mathbb{R}^n$ and a positive real scalar ε arbitrarily, we set $\mathbb{U}_{n,p}^\varepsilon(x) := \{x\} + \varepsilon\mathbb{U}_{n,p}$ and $\mathbb{B}_{n,p}^\varepsilon(x) := \{x\} + \varepsilon\mathbb{B}_{n,p}$. For arbitrary vectors $x, y \in \mathbb{R}^n$, $x \cdot y$ denotes their Euclidean inner product. Clearly, this inner product induces the norm $|\cdot|_2$. We write $x \leq y$ ($x < y$) if $y - x \in \mathbb{R}_0^{n,+}$ ($y - x \in \mathbb{R}^{n,+}$) holds. For a sequence $\{t_k\} \subseteq \mathbb{R}$, $t_k \searrow 0$ is used to express that $\{t_k\} \subseteq \mathbb{R}^+$ and $t_k \to 0$ are valid. On the other hand, we write $t_k \downarrow 0$ in order to express $\{t_k\} \subseteq \mathbb{R}_0^+$ and $t_k \to 0$.

Remark 2.7. Let $\mathcal{X}_1, \ldots, \mathcal{X}_n$ be Banach spaces. As introduced before, for the norm of their corresponding product space, we have

$$\forall x = (x_1, \ldots, x_n) \in \prod\nolimits_{j=1}^n \mathcal{X}_j: \quad \|x\|_{\prod_{j=1}^n \mathcal{X}_j} = \left|\left(\|x_1\|_{\mathcal{X}_1}, \ldots, \|x_n\|_{\mathcal{X}_n}\right)\right|_1.$$

However, since all the norms $|\cdot|_p$ for $p \in [1, \infty]$ are equivalent, for any such p, an equivalent norm of the product space $\prod_{j=1}^n \mathcal{X}_j$ is given by

$$\forall x = (x_1, \ldots, x_n) \in \prod\nolimits_{j=1}^n \mathcal{X}_j: \quad \|x\|_{p,\prod_{j=1}^n \mathcal{X}_j} := \left|\left(\|x_1\|_{\mathcal{X}_1}, \ldots, \|x_n\|_{\mathcal{X}_n}\right)\right|_p.$$

The choice of such a norm in the product space does not change the statement of Lemma 2.3, one only needs to choose different constants in (2.1).

For $m \in \mathbb{N}$, let $\mathbb{R}^{m\times n}$ contain all real matrices with m rows and n columns. Especially, $\mathbb{R}^{m\times 1} = \mathbb{R}^m$ is obtained. We use $\mathbf{O} \in \mathbb{R}^{m\times n}$ and $\mathbf{E} \in \mathbb{R}^{m\times n}$ to represent the zero matrix and the all-ones matrix of appropriate dimensions, respectively. The symbol $\mathbf{I}_n$ is used to represent the identity matrix in $\mathbb{R}^{n\times n}$. Furthermore, for any matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{A}^\top$ denotes its transpose. Interpreting $\mathbf{A}$ as a linear operator between the product spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, we obtain $\mathbf{A}^\star = \mathbf{A}^\top$ by means of formula (2.2). Let arbitrary index sets $I \subseteq \{1, \ldots, m\}$ and $J \subseteq \{1, \ldots, n\}$ be given. Then $\mathbf{A}_{IJ} \in \mathbb{R}^{|I|\times|J|}$ denotes the submatrix of $\mathbf{A}$ which possesses the rows indexed by the elements of I and the columns indexed by the elements of J. For any quadratic and regular matrix $\mathbf{B} \in \mathbb{R}^{m\times m}$, $\mathbf{B}^{-1}$ expresses the inverse matrix of $\mathbf{B}$. In general, $\mathbf{A}^\dagger$ denotes the pseudo inverse matrix of $\mathbf{A}$. If $\mathbf{A}$ possesses full row rank m, $\mathbf{A}^\dagger = \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}$ is valid. We will exploit the Hadamard product in $\mathbb{R}^{m\times n}$ defined by

$$\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}: \quad \mathbf{A} \bullet \mathbf{B} := \begin{pmatrix} a_{1,1}b_{1,1} & a_{1,2}b_{1,2} & \ldots & a_{1,n}b_{1,n} \\ a_{2,1}b_{2,1} & a_{2,2}b_{2,2} & \ldots & a_{2,n}b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}b_{m,1} & a_{m,2}b_{m,2} & \ldots & a_{m,n}b_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

The set $\mathcal{S}_n$ shall comprise all symmetric matrices from $\mathbb{R}^{n\times n}$. We will make use of the Frobenius inner product on $\mathcal{S}_n$ defined as stated below for matrices $\mathbf{A} = (a_{i,j})_{i,j=1,\ldots,n}, \mathbf{B} = (b_{i,j})_{i,j=1,\ldots,n} \in \mathcal{S}_n$:

$$\langle \mathbf{A}, \mathbf{B} \rangle_{\mathcal{S}_n} := \sum_{i=1}^n \sum_{j=1}^n a_{i,j}b_{i,j}.$$

Forthwith, $\operatorname{tr}\mathbf{A}$ denotes the trace, i.e. the sum of the diagonal elements or eigenvalues, of $\mathbf{A}$. It is easily seen that $\langle \mathbf{A}, \mathbf{B} \rangle_{\mathcal{S}_n} = \operatorname{tr}(\mathbf{AB})$ is satisfied. Thus,

$$\|\mathbf{A}\|_{\mathcal{S}_n} := \sqrt{\operatorname{tr}(\mathbf{AA})}$$

induces a norm in $\mathcal{S}_n$. Consequently, $\mathcal{S}_n$ is a Banach space of dimension $\frac{1}{2}n(n+1)$.

2.2.2. Function spaces

Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete and σ-finite measure space (see [16] for a detailed introduction to the theory of measurable spaces and measure spaces), let $N(\mathfrak{M})$ be the system of all sets of measure zero in Σ, and fix some $n \in \mathbb{N}$. We assume $\mathfrak{m}(\Omega) > 0$ in order to exclude trivial situations. Let $\mathcal{L}^0(\mathfrak{M}, \mathbb{R}^n)$ be the set of measurable functions mapping from Ω to $\mathbb{R}^n$. We can define an equivalence relation on $\mathcal{L}^0(\mathfrak{M}, \mathbb{R}^n)$ as stated below:

$$\forall u, v \in \mathcal{L}^0(\mathfrak{M}, \mathbb{R}^n): \quad u \sim v \iff \exists N \in N(\mathfrak{M})\, \forall \omega \in \Omega \setminus N: \quad u(\omega) = v(\omega).$$

The corresponding factor set $\mathcal{L}^0(\mathfrak{M}, \mathbb{R}^n)/\sim$ is denoted by $L^0(\mathfrak{M}, \mathbb{R}^n)$ and its elements (equivalence classes) are expressed via $u\colon \Omega \to \mathbb{R}^n$ again, i.e. we identify an equivalence class with its representatives. For any $p \in [1, \infty]$, we denote by $L^p(\mathfrak{M}, \mathbb{R}^n)$ the Banach space of (equivalence classes of) functions from $L^0(\mathfrak{M}, \mathbb{R}^n)$ satisfying $\int_\Omega |u(\omega)|_2^p \,\mathrm{d}\mathfrak{m} < \infty$ for $p \in [1, \infty)$ or which are componentwise essentially bounded in the case $p = \infty$, whose norm is given as stated below:

$$\forall p \in [1, \infty)\, \forall u \in L^p(\mathfrak{M}, \mathbb{R}^n): \qquad \|u\|_{L^p(\mathfrak{M}, \mathbb{R}^n)} := \left(\int_\Omega |u(\omega)|_2^p \,\mathrm{d}\mathfrak{m} \right)^{\frac{1}{p}},$$

$$\forall u \in L^\infty(\mathfrak{M}, \mathbb{R}^n): \qquad \|u\|_{L^\infty(\mathfrak{M}, \mathbb{R}^n)} := \inf_{N \in N(\mathfrak{M})} \left(\sup_{\omega \in \Omega \setminus N} |u(\omega)|_2 \right).$$

In the case where $\Omega \subseteq \mathbb{R}^d$ is a domain (i.e. a nonempty, connected, open set) equipped with the Borelean σ-algebra induced by Ω and the corresponding Lebesgue measure $\mathfrak{l}$, we write $L^p(\Omega, \mathbb{R}^n)$ for any $p \in [1, \infty]$ and $\mathrm{d}\omega$ instead of $\mathrm{d}\mathfrak{m}$. On the other hand, in the case $n = 1$, we use $L^p(\mathfrak{M}) := L^p(\mathfrak{M}, \mathbb{R})$ for any $p \in [1, \infty]$. It is easily seen from Remark 2.7 that the spaces $L^2(\mathfrak{M}, \mathbb{R}^n)$ and $L^2(\mathfrak{M})^n$ are isomorphic and equipped with equivalent norms. Note that for any $p \in [1, \infty)$, the space $L^{p'}(\mathfrak{M}, \mathbb{R}^n)$ is isometrically isomorphic to $L^p(\mathfrak{M}, \mathbb{R}^n)^\star$ where the conjugate coefficient $p' \in (1, \infty]$ is characterized via $1/p + 1/p' = 1$ (for $p = 1$, we set $p' = \infty$). The corresponding dual pairing reads as

$$\forall p \in [1, \infty)\, \forall u \in L^p(\mathfrak{M}, \mathbb{R}^n)\, \forall v \in L^{p'}(\mathfrak{M}, \mathbb{R}^n): \quad \langle v, u \rangle_{L^p(\mathfrak{M}, \mathbb{R}^n)} := \int_\Omega u(\omega) \cdot v(\omega) \mathrm{d}\mathfrak{m}.$$

For any $p \in (1, \infty)$, $L^p(\mathfrak{M}, \mathbb{R}^n)$ is reflexive. The case $p = 2$ is of special interest because $L^2(\mathfrak{M}, \mathbb{R}^n)$ is a Hilbert space. For any set $A \in \Sigma$, $\chi_A\colon \Omega \to \mathbb{R}$ defined by

$$\forall \omega \in \Omega: \quad \chi_A(\omega) := \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

is called characteristic function of A. Obviously, we have $\chi_A \in L^\infty(\mathfrak{M})$ and $\|\chi_A\|_{L^p(\mathfrak{M})} = \mathfrak{m}(A)^{\frac{1}{p}}$ for any $p \in [1, \infty)$, i.e. $\chi_A \in L^p(\mathfrak{M})$ for all $p \in [1, \infty)$ if and only if $\mathfrak{m}(A) < \infty$.

Let $\Omega \subseteq \mathbb{R}^d$ be a domain. Any $\alpha \in \mathbb{N}_0^d$ is called a multiindex of order $|\alpha| := |\alpha|_1$. We write ω^α in order to express the monomial $\omega_1^{\alpha_1} \dots \omega_d^{\alpha_d}$. Let us introduce the differential operator $D^\alpha := D_1^{\alpha_1} \dots D_d^{\alpha_d}$ of order $|\alpha|$ where $D_i := \frac{\partial}{\partial \omega_i}$, $i = 1, \dots, d$, holds. We interpret $D^{(0,\dots,0)}$ to be the identity and list some other popular differential operators below:

$$\Delta := \sum_{i=1}^d D_i^2, \quad \nabla := \begin{pmatrix} D_1 \\ \vdots \\ D_d \end{pmatrix}, \quad \nabla^2 := \begin{pmatrix} | & & | \\ \nabla D_1 & \dots & \nabla D_d \\ | & & | \end{pmatrix}^\top.$$

Furthermore, we set

$$\nabla_{\omega_I} := \begin{pmatrix} \vdots \\ D_i \\ \vdots \end{pmatrix}_{i\in I}, \quad \nabla^2_{\omega_I\omega_J} := \begin{pmatrix} & | & \\ \dots & \nabla_{\omega_J} D_i & \dots \\ & | & \end{pmatrix}^\top_{i\in I}.$$

Therein, $I, J \subseteq \{1, \dots d\}$ are nonempty and $\omega_I := (\omega_i)_{i\in I}$ holds true. If $u\colon \Omega \to \mathbb{R}^n$ is a function such that all components $u_1, \dots, u_n\colon \Omega \to \mathbb{R}$ are differentiable, we define

$$\nabla u := \begin{pmatrix} | & & | \\ \nabla u_1 & \dots & \nabla u_n \\ | & & | \end{pmatrix}^\top, \qquad \nabla_{\omega_I} u := \begin{pmatrix} | & & | \\ \nabla_{\omega_I} u_1 & \dots & \nabla_{\omega_I} u_n \\ | & & | \end{pmatrix}^\top.$$

Let us assume that the domain Ω is bounded. We exploit the notation $\overline{\Omega} := \operatorname{cl}\Omega$ in order to stay close to standard literature. For any function $u\colon \Omega \to \mathbb{R}$, the set $\operatorname{supp} u := \operatorname{cl}\{\omega \in \Omega \,|\, u(\omega) \neq 0\}$ is called support of u. For any $k \in \mathbb{N}_0 \cup \{\infty\}$, we introduce $C^k(\Omega)$, the vector space of all k-times continuously differentiable real-valued functions on Ω and set $C(\Omega) := C^0(\Omega)$ to be the vector space of all continuous functions on Ω. Clearly, since Ω is not necessarily closed, the functions in $C^k(\Omega)$ do not need to be bounded. Additionally, we introduce $C^k_0(\Omega)$, the subspace of $C^k(\Omega)$ comprising all functions whose support is a subset of Ω and compact in $\mathbb{R}^d$. Clearly, any function from $C^k_0(\Omega)$ vanishes on $\operatorname{bd}\Omega$. Again, we stipulate $C_0(\Omega) := C^0_0(\Omega)$. Let us consider $C^k\left(\overline{\Omega}\right)$, the vector space of all k-times continuously differentiable functions on $\overline{\Omega}$, which becomes a Banach space when equipped with the norm

$$\forall u \in C^k\left(\overline{\Omega}\right): \quad \|u\|_{C^k\left(\overline{\Omega}\right)} := \max_{\alpha\in\mathbb{N}_0^d,\, |\alpha|\leq k} \left(\max_{\omega\in\overline{\Omega}} |D^\alpha u(\omega)| \right).$$

We define $C\left(\overline{\Omega}\right) := C^0\left(\overline{\Omega}\right)$, the Banach space of all functions continuous on $\overline{\Omega}$. Furthermore, we will exploit the vector space of all locally Lebesgue-integrable functions $L^1_{\mathsf{loc}}(\Omega)$ defined by

$$L^1_{\mathsf{loc}}(\Omega) := \left\{ u \in L^0(\Omega) \,\middle|\, \forall \phi \in C^\infty_0(\Omega)\colon \int_\Omega u(\omega)\phi(\omega)\mathrm{d}\omega < \infty \right\}.$$

Note that for any functions $u, \phi \in C^\infty_0(\Omega)$ and any multiindex $\alpha \in \mathbb{N}^d_0$, we obtain

$$\int_\Omega u(\omega) D^\alpha \phi(\omega)\mathrm{d}\omega = (-1)^{|\alpha|} \int_\Omega D^\alpha u(\omega)\phi(\omega)\mathrm{d}\omega$$

from integration by parts and the fact that the functions from $C^\infty_0(\Omega)$ vanish on $\operatorname{bd}\Omega$. This motivates the following definition of weak derivatives (also referred to as Sobolev derivatives), see [118].

Definition 2.1. Let $u \in L^1_{\mathsf{loc}}(\Omega)$ and a multiindex $\alpha \in \mathbb{N}^d_0$ be given. A function $v \in L^1_{\mathsf{loc}}(\Omega)$ which satisfies

$$\forall \phi \in C^\infty_0(\Omega)\colon \quad \int_\Omega u(\omega) D^\alpha \phi(\omega)\mathrm{d}\omega = (-1)^{|\alpha|} \int_\Omega v(\omega)\phi(\omega)\mathrm{d}\omega$$

is called weak derivative of order α of u and is denoted by $D^\alpha u$.

The concept of weak derivatives will become important when considering solutions of PDEs. It is well-known that several PDEs possess no classical (i.e. strong) solution but weak solutions in the Sobolev sense. From the definition of weak derivatives it is natural to call functions from $C^\infty_0(\Omega)$ test functions. Note that we will use the differential operators defined above in this weak sense as well.

Now, it is possible to introduce the so-called Sobolev spaces, see [1]. For any $p \in [1, \infty]$ and $k \in \mathbb{N}_0$, $W^{k,p}(\Omega)$ denotes the set of all functions $u \in L^p(\Omega)$ possessing weak derivatives $D^\alpha u \in L^p(\Omega)$ for any $\alpha \in \mathbb{N}^d_0$ such that $|\alpha| \leq k$. Defining the norm

$$\forall p \in [1,\infty)\, \forall u \in W^{k,p}(\Omega)\colon \qquad \|u\|_{W^{k,p}(\Omega)} := \left(\sum_{\alpha\in\mathbb{N}_0^d,\, |\alpha|\leq k} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}},$$

$$\forall u \in W^{k,\infty}(\Omega)\colon \qquad \|u\|_{W^{k,\infty}(\Omega)} := \max_{\alpha\in\mathbb{N}_0^d,\, |\alpha|\leq k} \|D^\alpha u\|_{L^\infty(\Omega)},$$

$W^{k,p}(\Omega)$ becomes a Banach space. If $p \in (1,\infty)$ holds, then $W^{k,p}(\Omega)$ is reflexive. Note that from [1, Theorem 3.17] $W^{k,p}(\Omega)$ is the completion of the set $\{u \in C^k(\Omega) \mid \|u\|_{W^{k,p}(\Omega)} < \infty\}$ w.r.t. the Sobolev norm $\|\cdot\|_{W^{k,p}(\Omega)}$. Furthermore, we will deal with the space $W_0^{k,p}(\Omega)$ which is the closure of $C_0^\infty(\Omega)$ w.r.t. the Sobolev norm $\|\cdot\|_{W^{k,p}(\Omega)}$. We obtain the trivial embeddings

$$W_0^{k,p}(\Omega) \hookrightarrow W^{k,p}(\Omega) \hookrightarrow L^p(\Omega).$$

For $p = 2$, we set $H^k(\Omega) := W^{k,2}(\Omega)$ and observe that for any $k \in \mathbb{N}_0$, this is a Hilbert space with inner product

$$\forall u, v \in H^k(\Omega): \quad (v,u)_{H^k(\Omega)} := \sum_{\alpha \in \mathbb{N}_0^d,\, |\alpha| \le k} \langle D^\alpha v, D^\alpha u \rangle_{L^2(\Omega)}.$$

Note that we do not identify $H^k(\Omega)$ with its dual (see [118] for some reasons) and, thus, did not use the notion of the dual pairing above. We introduce the Hilbert space $H_0^k(\Omega) := W_0^{k,2}(\Omega)$ for any $k \in \mathbb{N}_0$. The dual space of $H_0^1(\Omega)$ will be denoted by $H^{-1}(\Omega)$, and again, we abstain from identifying it with $H_0^1(\Omega)$. Using the definition of duality, we easily obtain the embeddings

$$H_0^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega),$$

i.e. the spaces $(H_0^1(\Omega), L^2(\Omega), H^{-1}(\Omega))$ form a so-called Gelfand triple, see [118].

In the following two theorems, we subsume the embeddings needed in this thesis. The results are parts of Sobolev's embedding theorem, see [1, Theorem 4.12], and the Rellich-Kondrachov embedding theorem, see [1, Theorem 6.3], respectively.

Theorem 2.8. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz continuous boundary for $d \ge 2$ and a bounded, open interval for $d = 1$. For any integer $k \in \mathbb{N}$ and any real numbers $p, q \in [1,\infty)$, the following assertions hold:

1. If $kp > d$, then
$$W^{k,p}(\Omega) \hookrightarrow C\left(\overline{\Omega}\right).$$
2. If $kp = d$ and $p \le q < \infty$, then
$$W^{k,p}(\Omega) \hookrightarrow L^q(\Omega).$$
3. If $kp < d$ and $p \le q < \frac{dp}{d-kp}$, then
$$W^{k,p}(\Omega) \hookrightarrow L^q(\Omega).$$

Theorem 2.9. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz continuous boundary for $d \ge 2$ and a bounded, open interval for $d = 1$. Then the embeddings from Theorem 2.8 are compact.

When dealing with ODEs defined on a real time interval $\Omega := (0,T)$, we will deal with the vector space $AC(\Omega)$ of absolutely continuous functions on $\overline{\Omega}$. Note that any function from $AC(\Omega)$ is differentiable almost everywhere and the corresponding derivative equals the (existing) weak derivative which is an element of $L^1(\Omega)$. For short, for any absolutely continuous function $u \in AC(\Omega)$, there is $v \in L^1(\Omega)$ such that $u(t) = u(0) + \int_0^t v(\tau)\mathrm{d}\tau$ holds for almost every $t \in \Omega$ and vice versa; and v is its weak derivative, see [49, Theorem 2.1.20]. Thus, any function $u \in AC(\Omega)$ may be identified with $(u(0), D^1 u) \in \mathbb{R} \times L^1(\Omega)$. In this thesis, we will work with the space $AC^{1,2}(\Omega,\mathbb{R}^n)$ which contains all functions mapping from Ω to $\mathbb{R}^n$ whose components are absolutely continuous with weak derivatives in $L^2(\Omega)$. This way, $AC^{1,2}(\Omega,\mathbb{R}^n)$ can be identified with $\mathbb{R}^n \times L^2(\Omega,\mathbb{R}^n)$ by means of the bijection

$$AC^{1,2}(\Omega,\mathbb{R}^n) \ni u \mapsto (u(0), \nabla u) \in \mathbb{R}^n \times L^2(\Omega,\mathbb{R}^n)$$

and this will be done throughout the thesis at several instances. Note that ∇u denotes the column vector of the weak derivatives corresponding to the components of u. A suitable norm in $AC^{1,2}(\Omega,\mathbb{R}^n)$ is given by

$$\forall u \in AC^{1,2}(\Omega,\mathbb{R}^n): \quad \|u\|_{AC^{1,2}(\Omega,\mathbb{R}^n)} := |u(0)|_2 + \|\nabla u\|_{L^2(\Omega,\mathbb{R}^n)}.$$

Note that $AC^{1,2}(\Omega,\mathbb{R}^n)$ is a Hilbert space whose dual will be identified with $AC^{1,2}(\Omega,\mathbb{R}^n)$ throughout the thesis. The dual pairing in $AC^{1,2}(\Omega,\mathbb{R}^n)$ is defined as stated below:

$$\forall u,v\in AC^{1,2}(\Omega,\mathbb{R}^n)\colon\quad \langle v,u\rangle_{AC^{1,2}(\Omega,\mathbb{R}^n)} := u(0)\cdot v(0)+\int_0^T \nabla u(t)\cdot\nabla v(t)\mathrm{d}t.$$

Observe that it is possible to interpret the above dual pairing as an inner product in $AC^{1,2}(\Omega,\mathbb{R}^n)$. However, this inner product induces a norm in $AC^{1,2}(\Omega,\mathbb{R}^n)$ which is different from $\|\cdot\|_{AC^{1,2}(\Omega,\mathbb{R}^n)}$.
The following embedding theorem will be important for our subsequent analysis.

Theorem 2.10. For any $T>0$, $\Omega:=(0,T)$, and $n\in\mathbb{N}$, we have

$$AC^{1,2}(\Omega,\mathbb{R}^n)\hookrightarrow C\left(\overline{\Omega}\right)^n$$

and this embedding is compact.

Proof. It is clear from the definition that $AC^{1,2}(\Omega,\mathbb{R}^n)\subseteq H^1(\Omega)^n$ is satisfied. Applying Theorem 2.8 with $d=1$, $m=1$, and $p=2$, we find $c>0$ such that

$$\forall v\in H^1(\Omega)\colon\quad \|v\|_{C\left(\overline{\Omega}\right)}\leq c\,\|v\|_{H^1(\Omega)}$$

holds along with $H^1(\Omega)\subseteq C\left(\overline{\Omega}\right)$. Thus, we have $AC^{1,2}(\Omega,\mathbb{R}^n)\subseteq C\left(\overline{\Omega}\right)^n$ and for any $u\in AC^{1,2}(\Omega,\mathbb{R}^n)$, we obtain

$$\begin{aligned}
\|u\|_{C\left(\overline{\Omega}\right)^n} &= \sum_{i=1}^n \|u_i\|_{C\left(\overline{\Omega}\right)} \leq c\sum_{i=1}^n \|u_i\|_{H^1(\Omega)} = c\sum_{i=1}^n \sqrt{\|u_i\|^2_{L^2(\Omega)}+\|D^1u_i\|^2_{L^2(\Omega)}}\\
&\leq c\sum_{i=1}^n \left(\|u_i\|_{L^2(\Omega)}+\left\|D^1u_i\right\|_{L^2(\Omega)}\right)\\
&= c\sum_{i=1}^n \left(\left\|u_i(0)+\int_0^{\cdot} D^1u_i(\tau)\mathrm{d}\tau\right\|_{L^2(\Omega)}+\left\|D^1u_i\right\|_{L^2(\Omega)}\right)\\
&\leq c\sum_{i=1}^n \left(\sqrt{T}|u_i(0)|+\left(\int_0^T\left(\int_0^t D^1u_i(\tau)\mathrm{d}\tau\right)^2\mathrm{d}t\right)^{\frac{1}{2}}+\left\|D^1u_i\right\|_{L^2(\Omega)}\right)\\
&\leq c\sum_{i=1}^n \left(\sqrt{T}|u_i(0)|+\left(\int_0^T t\left(\int_0^t D^1u_i(\tau)^2\mathrm{d}\tau\right)\mathrm{d}t\right)^{\frac{1}{2}}+\left\|D^1u_i\right\|_{L^2(\Omega)}\right)\\
&\leq c\sum_{i=1}^n \left(\sqrt{T}|u_i(0)|+\left(T\int_0^T\int_0^T D^1u_i(\tau)^2\mathrm{d}\tau\mathrm{d}t\right)^{\frac{1}{2}}+\left\|D^1u_i\right\|_{L^2(\Omega)}\right)\\
&\leq c\sum_{i=1}^n \left(\sqrt{T}|u_i(0)|+T\left\|D^1u_i\right\|_{L^2(\Omega)}+\left\|D^1u_i\right\|_{L^2(\Omega)}\right)\\
&= c\sqrt{T}\,|u(0)|_1+c(1+T)\,\|\nabla u\|_{L^2(\Omega)^n}\\
&\leq C_1c\sqrt{T}\,|u(0)|_2+C_2c(1+T)\,\|\nabla u\|_{L^2(\Omega,\mathbb{R}^n)}\\
&\leq \max\{C_1c\sqrt{T};C_2c(1+T)\}\,\|u\|_{AC^{1,2}(\Omega,\mathbb{R}^n)}
\end{aligned}$$

where $C_1,C_2>0$ are real constants characterizing the equivalent norms $|\cdot|_1$ and $|\cdot|_2$ as well as $\|\cdot\|_{L^2(\Omega)^n}$ and $\|\cdot\|_{L^2(\Omega,\mathbb{R}^n)}$, respectively. This shows $AC^{1,2}(\Omega,\mathbb{R}^n)\hookrightarrow C\left(\overline{\Omega}\right)^n$. On the other hand, the above argumentation may be reprised to see

$$\forall u\in H^1(\Omega)^n\colon\quad \|u\|_{H^1(\Omega)^n}\leq\max\{C_1\sqrt{T};C_2(1+T)\}\,\|u\|_{AC^{1,2}(\Omega,\mathbb{R}^n)}.$$

Thus, we have $AC^{1,2}(\Omega,\mathbb{R}^n) \hookrightarrow H^1(\Omega)^n$. Theorem 2.9 yields that $H^1(\Omega)^n \hookrightarrow C\left(\overline{\Omega}\right)^n$ is a compact embedding. Applying [1, Remark 6.4.2], we obtain that $AC^{1,2}(\Omega,\mathbb{R}^n) \hookrightarrow C\left(\overline{\Omega}\right)^n$ is a compact embedding as well. This completes the proof. □

Now, let $\Omega \subseteq \mathbb{R}^d$ be an arbitrary bounded domain again. Then $(\overline{\Omega}, \mathfrak{B}(\overline{\Omega}))$ is a measurable space where $\mathfrak{B}(\overline{\Omega})$ denotes the Borelean σ-algebra induced by $\overline{\Omega}$. On the other hand, $\overline{\Omega}$ becomes a metric space when equipped with the Euclidean distance. A mapping $\mu\colon \mathfrak{B}(\overline{\Omega}) \to \mathbb{R}$ which is σ-additive and satisfies $\mu(\varnothing) = 0$ is called a signed measure of $(\overline{\Omega}, \mathfrak{B}(\overline{\Omega}))$. For any signed measure μ, we define its variation $|\mu|\colon \mathfrak{B}(\overline{\Omega}) \to \mathbb{R}_0^+$ by

$$\forall A \in \mathfrak{B}(\overline{\Omega})\colon \quad |\mu|(A) := \sup_{\mathfrak{S}\in\mathcal{Z}(A)} \sum_{E\in\mathfrak{S}} |\mu(E)|$$

where $\mathcal{Z}(A)$ denotes the system of all finite and disjoint partitions of A in $\mathfrak{B}(\overline{\Omega})$. We call μ regular if its variation $|\mu|$ possesses only finite values on the compact subsets of $\overline{\Omega}$ and satisfies

$$\forall A \in \mathfrak{B}(\overline{\Omega})\colon \quad |\mu|(A) = \sup\{|\mu|(C) \,|\, C \subseteq A,\, C \text{ compact}\}.$$

Let $\mathcal{M}(\overline{\Omega})$ be the vector space of all signed and regular measures of $(\overline{\Omega}, \mathfrak{B}(\overline{\Omega}))$. Introducing

$$\forall \mu \in \mathcal{M}(\overline{\Omega})\colon \quad \|\mu\|_{\mathcal{M}(\overline{\Omega})} := |\mu|(\overline{\Omega}),$$

$\mathcal{M}(\overline{\Omega})$ becomes a nonreflexive Banach space. It is well-known that it is the dual space of $C(\overline{\Omega})$, see [126, Satz II.2.5].

2.3. Principles of variational analysis and optimization in Banach spaces

Set approximation and generalized differentiation are essential concepts for the consideration of optimization problems in Banach spaces. In this section, we are going to introduce the variational concepts needed in this thesis.

2.3.1. Polar, tangent, and normal cones

Let $\mathcal{X}$ be a Banach space and let $A \subseteq \mathcal{X}$ be a nonempty set. The polar cone and the annihilator of A are defined by

$$A^\circ := \{x^* \in \mathcal{X}^\star \,|\, \forall x \in A\colon\ \langle x^*, x\rangle_{\mathcal{X}} \le 0\} \quad \text{and} \quad A^\perp := \{x^* \in \mathcal{X}^\star \,|\, \forall x \in A\colon\ \langle x^*, x\rangle_{\mathcal{X}} = 0\},$$

respectively. Clearly, A° is a weakly* closed, convex cone, whereas $A^\perp$ is a weakly* closed subspace of $\mathcal{X}^\star$. It is easy to see that $A^\circ = (\operatorname{cl} A)^\circ$, $A^\circ = (\operatorname{conv} A)^\circ$, and $A^\circ = (\operatorname{cone} A)^\circ$ hold true. Furthermore, $A^\perp = A^\circ \cap (-A)^\circ$ is satisfied and, thus, $S^\perp = S^\circ$ holds true for any subspace $S \subseteq \mathcal{X}$. For any nonempty set $B \subseteq \mathcal{X}^\star$, we introduce the corresponding backward operations by

$$B_\circ := \{x \in \mathcal{X} \,|\, \forall x^* \in B\colon\ \langle x^*, x\rangle_{\mathcal{X}} \le 0\}, \quad B_\perp = \{x \in \mathcal{X} \,|\, \forall x^* \in B\colon\ \langle x^*, x\rangle_{\mathcal{X}} = 0\}.$$

These operations possess similar properties as the polarization and annihilation presented above. If $\mathcal{X}$ is reflexive, $\mathcal{X}^{\star\star} \cong \mathcal{X}$ holds true and, thus, it is consistent to identify $B^\circ = B_\circ$ as well as $B^\perp = B_\perp$. The following lemma is often called bipolar theorem, see [17, Proposition 2.40].

Lemma 2.11. Let $\mathcal{X}$ be a Banach space, and let $C \subseteq \mathcal{X}$ be a cone. Then $(C^\circ)_\circ = \overline{\operatorname{conv}}\, C$ holds.

For any subspace $S \subseteq \mathcal{X}$, Lemma 2.11 leads to $(S^\perp)_\perp = \operatorname{cl} S$. Especially, for a single point $\bar{x} \in \mathcal{X}$, $\operatorname{lin}\{\bar{x}\} = (\{\bar{x}\}^\perp)_\perp = (\{\bar{x}\}^\perp)_\circ$ is obtained. In the lemma below, we list some calculus rules for convex cones using the above operations.

Lemma 2.12. Let $\mathcal{X}$ be a Banach space, and let $C_1, C_2, C \subseteq \mathcal{X}$ as well as $K_1, K_2 \subseteq \mathcal{X}^\star$ be nonempty, closed, convex cones. Then we have

$$(C_1 + C_2)^\circ = C_1^\circ \cap C_2^\circ, \tag{2.3a}$$
$$(C_1 \cap C_2)^\circ = \mathrm{cl}^\star(C_1^\circ + C_2^\circ), \tag{2.3b}$$
$$(K_1 \cap K_2)_\circ = \mathrm{cl}\big((K_1)_\circ + (K_2)_\circ\big), \tag{2.3c}$$
$$(C^\circ)_\perp = C \cap (-C), \tag{2.3d}$$
$$(C^\perp)_\circ = \mathrm{cl}\,\mathrm{lin}\, C. \tag{2.3e}$$

Proof. For the proof of the statements (2.3a), (2.3b), and (2.3c), we refer to [17, formulae (2.31) and (2.32)]. Applying Lemma 2.11,

$$(C^\circ)_\perp = (C^\circ)_\circ \cap (-C^\circ)_\circ = (C^\circ)_\circ \cap ((-C)^\circ)_\circ = C \cap (-C)$$

is obtained which yields (2.3d). Finally, we exploit Lemma 2.11, (2.3c), and $\mathrm{lin}\, C = C - C$ to show

$$(C^\perp)_\circ = \big(C^\circ \cap (-C)^\circ\big)_\circ = \mathrm{cl}\big((C^\circ)_\circ - (C^\circ)_\circ\big) = \mathrm{cl}(C - C) = \mathrm{cl}\,\mathrm{lin}\, C$$

which yields (2.3e). □

Let $\mathcal{Y}$ be another Banach space and choose an operator $\mathtt{F} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$. For any set $A \subseteq \mathcal{X}$, we define the image of A under $\mathtt{F}$ by $\mathtt{F}[A] := \{\mathtt{F}[x] \in \mathcal{Y} \,|\, x \in A\}$. The set $\mathtt{F}[\mathcal{X}]$ is just called image of $\mathtt{F}$. Obviously, $\mathtt{F}[\mathcal{X}]$ is a linear subspace of $\mathcal{Y}$. Note that $\mathtt{F}[\mathcal{X}]$ does not need to be closed.
The following result is called generalized Farkas lemma and can be found in a more abstract form in [51, Theorem 1, Lemma 3]. In [47, Remark 2.1], its relation to the well-known Farkas lemma is illustrated. The third assertion is taken from [17, Proposition 2.201].

Lemma 2.13. Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, let $C \subseteq \mathcal{X}$ and $K \subseteq \mathcal{Y}$ be nonempty, closed, convex cones, and let $\mathtt{A} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$ be an arbitrary linear operator. Then we obtain

$$\{x \in C \,|\, \mathtt{A}[x] \in K\}^\circ = \mathrm{cl}^\star(C^\circ + \mathtt{A}^\star[K^\circ]).$$

If the condition $\mathtt{A}[C] - K = \mathcal{Y}$ holds, then $\mathrm{cl}^\star$ can be dropped in the above line.
On the other hand, if there are functionals $x_1^*, \ldots, x_n^* \in \mathcal{X}^\star$ such that C possesses the form

$$C = \{x \in \mathcal{X} \,|\, \forall j \in \{1, \ldots, n\}\colon\ \langle x_j^*, x\rangle_{\mathcal{X}} \leq 0\}$$

and $\mathtt{A}[\mathcal{X}]$ is closed, then

$$\{x \in C \,|\, \mathtt{A}[x] = 0\}^\circ = \mathrm{cone}\{x_1^*, \ldots, x_n^*\} + \mathtt{A}^\star[\mathcal{Y}^\star]$$

is satisfied.

Note that the first statement of the above lemma generalizes the calculus rule (2.3b).
The following example introduces the concept of ellipticity and is included since it presents a typical application of the above calculus rules for annihilators and polars. Note that the concept of elliptic operators is closely related to the concept of elliptic PDEs. The theoretical and numerical handling of optimal control problems which are governed by linear elliptic PDEs is well-developed, see [103, 118] and the references therein.

Example 2.14. Let $\mathcal{X}$ be a reflexive Banach space and let $\mathtt{A} \in \mathbb{L}[\mathcal{X}, \mathcal{X}^\star]$ be elliptic. Thus, there is some $\alpha > 0$ such that

$$\forall x \in \mathcal{X}\colon \quad \langle \mathtt{A}[x], x\rangle_{\mathcal{X}} \geq \alpha \|x\|_{\mathcal{X}}^2$$

holds. Using the inequality of Cauchy and Schwarz, this yields

$$\forall x \in \mathcal{X}\colon \quad \alpha \|x\|_{\mathcal{X}} \leq \|\mathtt{A}[x]\|_{\mathcal{X}^\star} \leq \|\mathtt{A}\|_{\mathbb{L}[\mathcal{X}, \mathcal{X}^\star]} \|x\|_{\mathcal{X}}\,. \tag{2.4}$$

It is easily seen from (2.4) that $\mathtt{A}$ is injective. On the other hand, $\mathtt{A}[\mathcal{X}]$ is closed. In order to show this, we choose a sequence $\{x_k\} \subseteq \mathcal{X}$ such that $\{\mathtt{A}[x_k]\} \subseteq \mathcal{X}^\star$ converges to some $x^* \in \mathcal{X}^\star$. Particularly, $\{\mathtt{A}[x_k]\}$ is bounded and due to (2.4), the same holds true for $\{x_k\}$. We exploit the reflexivity of $\mathcal{X}$ in order to extract a subsequence $\{x_{k_l}\}$ of $\{x_k\}$ converging weakly to $\bar{x} \in \mathcal{X}$. Then for any $x \in \mathcal{X}$,

$$\langle \mathtt{A}[\bar{x}], x\rangle_{\mathcal{X}} = \langle \mathtt{A}^\star[x], \bar{x}\rangle_{\mathcal{X}} = \lim_{l\to\infty} \langle \mathtt{A}^\star[x], x_{k_l}\rangle_{\mathcal{X}} = \lim_{l\to\infty} \langle \mathtt{A}[x_{k_l}], x\rangle_{\mathcal{X}} = \langle x^*, x\rangle_{\mathcal{X}}$$

is obtained and consequently, $x^* = \mathtt{A}[\bar{x}] \in \mathtt{A}[\mathcal{X}]$ is valid. This yields the closedness of $\mathtt{A}[\mathcal{X}]$. Now, using the definition of ellipticity, $\mathtt{A}[\mathcal{X}]^\perp = \{0\}$ is obtained. Thus, we finally apply Lemma 2.11 to see $\mathcal{X}^\star = \{0\}^\perp = \mathtt{A}[\mathcal{X}]^{\perp\perp} = \operatorname{cl} \mathtt{A}[\mathcal{X}] = \mathtt{A}[\mathcal{X}]$ which shows the surjectivity of $\mathtt{A}$. Hence, any elliptic operator is an isomorphism. ■

In this thesis, we will deal with several different concepts of tangent and normal cones which will be defined and discussed below. Therefore, choose a nonempty set $A \subseteq \mathcal{X}$ such that $\bar{x} \in \operatorname{cl} A$ is satisfied. The cone

$$\mathcal{R}_A(\bar{x}) := \{d \in \mathcal{X} \,|\, \exists t_0 > 0 \,\forall s \in (0, t_0]\colon\ \bar{x} + sd \in A\}$$

is called radial cone to A at $\bar{x}$. Furthermore, we introduce the tangent (or Bouligand) cone, the weak tangent cone, the inner (or adjacent) tangent cone, and the Clarke tangent cone to A at $\bar{x}$ as stated below, see [6, Section 4.1] and [90, Definition 1.8]:

$$\begin{aligned}
\mathcal{T}_A(\bar{x}) &:= \{d \in \mathcal{X} \,|\, \exists \{d_k\} \subseteq \mathcal{X}\, \exists \{t_k\} \subseteq \mathbb{R}\colon\ d_k \to d,\ t_k \searrow 0,\ \bar{x} + t_k d_k \in A\, \forall k \in \mathbb{N}\},\\
\mathcal{T}_A^w(\bar{x}) &:= \{d \in \mathcal{X} \,|\, \exists \{d_k\} \subseteq \mathcal{X}\, \exists \{t_k\} \subseteq \mathbb{R}\colon\ d_k \rightharpoonup d,\ t_k \searrow 0,\ \bar{x} + t_k d_k \in A\, \forall k \in \mathbb{N}\},\\
\mathcal{T}_A^\flat(\bar{x}) &:= \{d \in \mathcal{X} \,|\, \forall \{t_k\} \subseteq \mathbb{R}\,(t_k \searrow 0 \implies \exists \{d_k\} \subseteq \mathcal{X}\colon\ d_k \to d,\ \bar{x} + t_k d_k \in A\, \forall k \in \mathbb{N})\},\\
\mathcal{T}_A^c(\bar{x}) &:= \left\{ d \in \mathcal{X} \,\middle|\, \begin{array}{l} \forall \{x_k\} \subseteq A\, \forall \{t_k\} \subseteq \mathbb{R} \\ \qquad (x_k \to \bar{x},\ t_k \searrow 0 \implies \exists \{d_k\} \subseteq \mathcal{X}\colon\ d_k \to d,\ x_k + t_k d_k \in A\, \forall k \in \mathbb{N}) \end{array} \right\}.
\end{aligned}$$

Clearly, the cones $\mathcal{T}_A(\bar{x})$, $\mathcal{T}_A^\flat(\bar{x})$, and $\mathcal{T}_A^c(\bar{x})$ are closed by definition, see [6, Section 4] as well. Moreover, $\mathcal{T}_A^c(\bar{x})$ is convex, see [6, Proposition 4.1.6]. From the definitions, we obviously have the inclusions

$$\mathcal{T}_A^c(\bar{x}) \subseteq \mathcal{T}_A^\flat(\bar{x}) \subseteq \mathcal{T}_A(\bar{x}) \subseteq \mathcal{T}_A^w(\bar{x})$$

and

$$\mathcal{R}_A(\bar{x}) \subseteq \mathcal{T}_A^\flat(\bar{x}).$$

The set A is said to be derivable at $\bar{x}$ if $\mathcal{T}_A^\flat(\bar{x}) = \mathcal{T}_A(\bar{x})$ is satisfied. We call A derivable if it is derivable at all of its points. For convex sets A, we easily see $\mathcal{T}_A^w(\bar{x}) \subseteq \operatorname{cl}^w_{\text{seq}} \operatorname{cone}(A - \{\bar{x}\}) = \overline{\operatorname{cone}}(A - \{\bar{x}\})$. Thus, [6, Proposition 4.2.1] yields

$$\overline{\operatorname{cone}}(A - \{\bar{x}\}) = \operatorname{cl} \mathcal{R}_A(\bar{x}) = \mathcal{T}_A^c(\bar{x}) = \mathcal{T}_A^\flat(\bar{x}) = \mathcal{T}_A(\bar{x}) = \mathcal{T}_A^w(\bar{x})$$

in the convex case. Hence, any convex set is derivable. If $C \subseteq \mathcal{X}$ is a closed, convex cone containing $\bar{c}$, the formulae

$$\mathcal{R}_C(\bar{c}) = C + \operatorname{lin}\{\bar{c}\}, \quad \mathcal{T}_C(\bar{c}) = \operatorname{cl}(C + \operatorname{lin}\{\bar{c}\})$$

are easily obtained; $\mathcal{T}_C(\bar{c})^\circ = C^\circ \cap \{\bar{c}\}^\perp$ follows from (2.3a).

Lemma 2.15. Let $\mathcal{X}$ be a Banach space, let $D_1, \dots, D_k \subseteq \mathcal{X}$ be closed sets, and choose $\bar{x} \in D := \bigcup_{i=1}^k D_i$ arbitrarily. Define $I(\bar{x}) = \{i \in \{1, \dots, k\} \,|\, \bar{x} \in D_i\}$ and assume that for any $i \in I(\bar{x})$, D_i is derivable at $\bar{x}$. Then D is derivable at $\bar{x}$.

Proof. Applying some calculus rules for tangent and inner tangent cones, see [6, Tables 4.1 and 4.2], we obtain

$$\mathcal{T}_D(\bar{x}) = \bigcup_{i\in I(\bar{x})} \mathcal{T}_{D_i}(\bar{x}) = \bigcup_{i\in I(\bar{x})} \mathcal{T}_{D_i}^\flat(\bar{x}) \subseteq \mathcal{T}_D^\flat(\bar{x}) \subseteq \mathcal{T}_D(\bar{x}).$$

Thus, $\mathcal{T}_D(\bar{x}) = \mathcal{T}_D^\flat(\bar{x})$ is obtained and, hence, D is derivable at $\bar{x}$. □

From the above lemma the union of finitely many convex sets is derivable. Especially, the finite union of polyhedral sets is derivable. Recall that a set $A \subseteq \mathcal{X}$ is polyhedral if there exist functionals $x_1^*, \dots, x_n^* \in \mathcal{X}^\star$ and scalars $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, such that A possesses the representation

$$A = \{x \in \mathcal{X} \,|\, \forall i \in \{1, \dots, n\} \colon \ \langle x_i^*, x \rangle_{\mathcal{X}} \leq \alpha_i\}.$$

On the other hand, a closed, convex set A is said to be polyhedric w.r.t. $(\bar{x}, x^*) \in A \times \mathcal{T}_A(\bar{x})^\circ$ if

$$\mathcal{T}_A(\bar{x}) \cap \{x^*\}_\perp = \operatorname{cl}\left(\mathcal{R}_A(\bar{x}) \cap \{x^*\}_\perp\right)$$

is satisfied, and we call A polyhedric if it is polyhedric w.r.t. all points $(x, \eta) \in A \times \mathcal{T}_A(x)^\circ$. The closed cone

$$\mathcal{K}_A(\bar{x}, x^*) := \mathcal{T}_A(\bar{x}) \cap \{x^*\}_\perp$$

is called critical cone to A w.r.t. $(\bar{x}, x^*)$. Roughly speaking, a set is polyhedric if its boundary possesses no curvature. It is easy to see that the radial cone to a polyhedral set is always closed and, thus, equals the corresponding tangent cone. This shows that any polyhedral set is polyhedric. The notion of polyhedricity dates back to the seminal works [57] and [88] where it was used to characterize the directional differentiability of the projection operator onto closed, convex sets. Later on, polyhedricity turned out to be a useful property in infinite-dimensional programming, see [17, 68, 120, 121]. Generalizations of polyhedricity can be found in [17, Section 3.2.3] and [120]. The latter article presents an overview of polyhedric sets, calculus rules addressing set intersections involving polyhedric sets, and applications of polyhedricity in mathematical programming.
Let $\mathcal{X}$ be a reflexive Banach space, let $C \subseteq \mathcal{X}$ be a nonempty, closed, convex cone, and choose $\bar{c} \in C$ as well as $c^* \in C^\circ$ such that $\langle c^*, \bar{c} \rangle_{\mathcal{X}} = 0$ holds (i.e. $(\bar{c}, c^*) \in C \times \mathcal{T}_C(\bar{c})^\circ$). Then

$$\begin{aligned} C \text{ polyhedric w.r.t. } (\bar{c}, c^*) &\iff C^\circ \text{ polyhedric w.r.t. } (c^*, \bar{c}) \\ &\iff \mathcal{K}_C(\bar{c}, c^*)^\circ = \mathcal{K}_{C^\circ}(c^*, \bar{c}) \\ &\iff \mathcal{K}_{C^\circ}(c^*, \bar{c})^\circ = \mathcal{K}_C(\bar{c}, c^*) \end{aligned} \tag{2.5}$$

is obtained by straightforward calculations and Lemma 2.11, see [121, Lemma 5.2] as well. Clearly, C is always polyhedric w.r.t. $(0, 0)$.

Let us introduce some appropriate concepts of generalized normals which date back to Mordukhovich, see [90] for the historical details. Again, let $A \subseteq \mathcal{X}$ be a set with $\bar{x} \in \operatorname{cl} A$. Furthermore, we choose a scalar $\sigma \geq 0$ in order to define the set of σ-normals to A at $\bar{x}$ as stated below:

$$\widehat{\mathcal{N}}_A^\sigma(\bar{x}) := \left\{ \eta \in \mathcal{X}^\star \;\middle|\; \limsup_{x \to \bar{x},\, x \in A} \frac{\langle \eta, x - \bar{x} \rangle_{\mathcal{X}}}{\|x - \bar{x}\|_{\mathcal{X}}} \leq \sigma \right\}.$$

For $\sigma = 0$, $\widehat{\mathcal{N}}_A(\bar{x}) := \widehat{\mathcal{N}}_A^0(\bar{x})$ is called Fréchet (or regular) normal cone to A at $\bar{x}$. Furthermore, the cone

$$\mathcal{N}_A(\bar{x}) := \left\{ \eta \in \mathcal{X}^\star \;\middle|\; \exists \{\sigma_k\} \subseteq \mathbb{R}\, \exists \{x_k\} \subseteq A\, \exists \{\eta_k\} \subseteq \mathcal{X}^\star \colon \ \sigma_k \downarrow 0,\ x_k \to \bar{x},\ \eta_k \overset{\star}{\to} \eta,\ \eta_k \in \widehat{\mathcal{N}}_A^{\sigma_k}(x_k)\, \forall k \in \mathbb{N} \right\}$$

is called limiting (or basic, Mordukhovich) normal cone to A at $\bar{x}$, whereas

$$\mathcal{N}_A^s(\bar{x}) := \left\{ \eta \in \mathcal{X}^\star \;\middle|\; \exists \{\sigma_k\} \subseteq \mathbb{R}\, \exists \{x_k\} \subseteq A\, \exists \{\eta_k\} \subseteq \mathcal{X}^\star \colon \ \sigma_k \downarrow 0,\ x_k \to \bar{x},\ \eta_k \to \eta,\ \eta_k \in \widehat{\mathcal{N}}_A^{\sigma_k}(x_k)\, \forall k \in \mathbb{N} \right\}$$

defines the so-called strong limiting (or norm-limiting) normal cone to A at $\bar{x}$. From [90, Theorem 2.35] we can fix $\sigma_k \equiv 0$ in these definitions provided $\mathcal{X}$ is reflexive (or, to be more general, a so-called Asplund space, i.e. a Banach space whose separable subspaces possess separable duals) and A is closed in a neighborhood of $\bar{x}$. While the limiting normal cone enjoys full calculus, see [90, Sections 1.1 and 3.1.1], the strong limiting normal cone suffers from a lack of available calculus rules in infinite-dimensional spaces; in the finite-dimensional case, these cones obviously coincide. To the best of our knowledge, the strong limiting normal cone was introduced in [50] first. Finally, we define the Clarke normal cone to A at $\bar{x}$ by

$$\mathcal{N}_A^c(\bar{x}) := \mathcal{T}_A^c(\bar{x})^\circ.$$

Let us stipulate that all the introduced normal cones to A at some point $\tilde{x} \notin \operatorname{cl} A$ are empty. It is clear from the definitions that $\widehat{\mathcal{N}}_A(\bar{x})$ and $\mathcal{N}_A^c(\bar{x})$ are closed, convex cones. On the other hand, $\mathcal{N}_A(\bar{x})$ is neither convex nor closed in general, see [90, Example 1.7]. Note that $\mathcal{N}_A^s(\bar{x})$ is a closed cone which is not necessarily convex.

Lemma 2.16. For any set $A \subseteq \mathcal{X}$ and $\bar{x} \in \operatorname{cl} A$, the cone $\mathcal{N}_A^s(\bar{x})$ is closed.

Proof. Let $\{\eta_k\} \subseteq \mathcal{N}_A^s(\bar{x})$ be a sequence converging to $\eta \in \mathcal{X}^\star$. By definition, for any $k \in \mathbb{N}$, we find sequences $\{\sigma_{k,l}\} \subseteq \mathbb{R}$, $\{x_{k,l}\} \subseteq A$, and $\{\eta_{k,l}\} \subseteq \mathcal{X}^\star$ with $\sigma_{k,l} \downarrow 0$, $x_{k,l} \to \bar{x}$, and $\eta_{k,l} \to \eta_k$ as $l \to \infty$, and $\eta_{k,l} \in \widehat{\mathcal{N}}_A^{\sigma_{k,l}}(x_{k,l})$ for all $l \in \mathbb{N}$. Thus, for any $k \in \mathbb{N}$, we find $l_k \in \mathbb{N}$ such that the relations

$$\sigma_{k,l_k} \leq \tfrac{1}{k}, \qquad \|x_{k,l_k} - \bar{x}\|_{\mathcal{X}} \leq \tfrac{1}{k}, \qquad \|\eta_{k,l_k} - \eta_k\|_{\mathcal{X}^\star} \leq \tfrac{1}{k}$$

hold. Hence, we have

$$\|\eta_{k,l_k} - \eta\|_{\mathcal{X}^\star} \leq \|\eta_{k,l_k} - \eta_k\|_{\mathcal{X}^\star} + \|\eta_k - \eta\|_{\mathcal{X}^\star} \leq \tfrac{1}{k} + \|\eta_k - \eta\|_{\mathcal{X}^\star}$$

from the triangle inequality, and, consequently, the sequences $\{\sigma_{k,l_k}\}$, $\{x_{k,l_k}\}$, and $\{\eta_{k,l_k}\}$ satisfy $\sigma_{k,l_k} \downarrow 0$, $x_{k,l_k} \to \bar{x}$, and $\eta_{k,l_k} \to \eta$ as $k \to \infty$, and $\eta_{k,l_k} \in \widehat{\mathcal{N}}_A^{\sigma_{k,l_k}}(x_{k,l_k})$ for all $k \in \mathbb{N}$. Hence, $\eta \in \mathcal{N}_A^s(\bar{x})$ is valid. □

Applying [90, Proposition 2.45], the inclusions

$$\widehat{\mathcal{N}}_A(\bar{x}) \subseteq \mathcal{N}_A^s(\bar{x}) \subseteq \mathcal{N}_A(\bar{x}) \subseteq \mathcal{N}_A^c(\bar{x})$$

between the introduced normal cones are obtained. For any convex set A, all these normal cones coincide with the normal cone of convex analysis, see [24, Proposition 2.4.4] and [90, Proposition 1.3], i.e. we have

$$\widehat{\mathcal{N}}_A(\bar{x}) = \mathcal{N}_A^s(\bar{x}) = \mathcal{N}_A(\bar{x}) = \mathcal{N}_A^c(\bar{x}) = \{\eta \in \mathcal{X}^\star \,|\, \forall x \in A\colon\ \langle \eta, x - \bar{x}\rangle_{\mathcal{X}} \leq 0\} = (A - \{\bar{x}\})^\circ.$$

On the other hand, if $\mathcal{X}$ is reflexive, we obtain $\widehat{\mathcal{N}}_A(\bar{x}) = \mathcal{T}_A^w(\bar{x})^\circ$ and $\mathcal{N}_A^c(\bar{x}) = \overline{\operatorname{conv}}\,\mathcal{N}_A(\bar{x})$ for sets A which are locally closed at $\bar{x}$ from [90, Corollary 1.11 and Theorem 3.57], respectively.
Let $\mathcal{X}$ be an arbitrary Banach space again. The set A is called sequentially normally compact, SNC for short, at $\bar{x}$ if for any sequences $\{\sigma_k\} \subseteq \mathbb{R}$, $\{x_k\} \subseteq A$, $\{\eta_k\} \subseteq \mathcal{X}^\star$ satisfying $\sigma_k \downarrow 0$, $x_k \to \bar{x}$, $\eta_k \xrightarrow{\star} 0$, and $\eta_k \in \widehat{\mathcal{N}}_A^{\sigma_k}(x_k)$ for all $k \in \mathbb{N}$, $\eta_k \to 0$ is valid. It follows from [90, Theorem 1.21] that a singleton in $\mathcal{X}$ is SNC at its point if and only if $\mathcal{X}$ is finite-dimensional. Clearly, any subset of a finite-dimensional Banach space is SNC at all of its points. In the following lemma, we present a results which characterizes the SNC property of a closed, convex set in a reflexive Banach space.

Lemma 2.17. For a reflexive Banach space $\mathcal{X}$ and a nonempty, closed, convex set $A \subseteq \mathcal{X}$, the following statements are equivalent:

(i) A is everywhere SNC,

(ii) A is SNC at some point $\bar{x} \in A$,

(iii) $\operatorname{lin} A$ is closed, $\operatorname{rint} A$ is nonempty, and the dimension of the factor space $\mathcal{X}/\operatorname{lin} A$ is finite.

Proof. Recall that the set A is CEL (compactly epi-Lipschitzian) at a point $\bar{x} \in A$ if there are neighborhoods N of $\bar{x}$ and U of 0, some $\varepsilon > 0$, and a convex, compact set $C \subseteq \mathcal{X}$ which satisfy

$$\forall \alpha \in (0, \varepsilon)\colon \quad A \cap N + \alpha U \subseteq A + \alpha C,$$

see [19, Definition 2.1]. Since A is a closed subset of a reflexive Banach space, we see from [39, comments after Definition 2.1, Theorem 3.1] that A is SNC at $\bar{x}$ if and only if it is CEL there.
Exploiting the equivalence of the SNC and CEL property, the equivalence of (i) and (iii) follows from [19, Theorem 2.5, (i) and (vii)]. Moreover, (i) obviously implies (ii), i.e. we only need to show that (ii) implies (i). Assume that A is SNC, i.e. CEL, at $\bar{x} \in A$. Then we find neighborhoods N of $\bar{x}$ and U of 0, some $\varepsilon > 0$, and a convex, compact set $C \subseteq \mathcal{X}$ which satisfy $A \cap N + \frac{\varepsilon}{2} U \subseteq A + \frac{\varepsilon}{2} C$. This yields $\{\bar{x}\} + \frac{\varepsilon}{2} U \subseteq A + \frac{\varepsilon}{2} C$ and, thus, $\frac{\varepsilon}{2} U \subseteq A + \frac{\varepsilon}{2} C - \{\bar{x}\}$. Obviously, the set $C' := \frac{\varepsilon}{2} C - \{\bar{x}\}$ is convex as well as compact and satisfies $0 \in \operatorname{int}(A + C')$. Thus, by means of [19, Theorem 2.5, (i) and (ii)], A is CEL everywhere. The above arguments imply that A needs to be SNC everywhere. □

Observe that the above result is remarkable since, by definition, SNC seems to be a local property of a set.
For Banach spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$, let $A_j \subseteq \mathcal{X}_j$ be chosen such that $\bar{x}_j \in A_j$, $j = 1, \ldots, n$, holds. Then from [90, Proposition 1.2] we obtain the product formulae

$$\widehat{\mathcal{N}}_A(\bar{x}) = \prod_{j=1}^{n} \widehat{\mathcal{N}}_{A_j}(\bar{x}_j), \quad \mathcal{N}_A^s(\bar{x}) = \prod_{j=1}^{n} \mathcal{N}_{A_j}^s(\bar{x}_j), \quad \mathcal{N}_A(\bar{x}) = \prod_{j=1}^{n} \mathcal{N}_{A_j}(\bar{x}_j) \tag{2.6}$$

where $A := \prod_{j=1}^{n} A_j$ and $\bar{x} := (\bar{x}_1, \ldots, \bar{x}_n)$. For the Clarke normal cone, we only obtain the inclusion

$$\mathcal{N}_A^c(\bar{x}) \supseteq \prod_{j=1}^{n} \mathcal{N}_{A_j}^c(\bar{x}_j)$$

by straightforward calculations in general.
We will exploit the following calculus rule for the limiting normal cone to the intersection of sets, see [90, Corollary 3.5].

Lemma 2.18. Let $\mathcal{X}$ be a reflexive Banach space and let $A, A' \subseteq \mathcal{X}$ be sets which are locally closed at $\bar{x} \in A \cap A'$. Assume that one of the sets A or A' is SNC at $\bar{x}$ and that the condition

$$\mathcal{N}_A(\bar{x}) \cap \left(-\mathcal{N}_{A'}(\bar{x})\right) = \{0\}$$

is satisfied. Then we have

$$\mathcal{N}_{A\cap A'}(\bar{x}) \subseteq \mathcal{N}_A(\bar{x}) + \mathcal{N}_{A'}(\bar{x}).$$

If, additionally, A and A' are locally convex at $\bar{x}$, then equality holds.

By means of examples, see [90], it is easily seen that the SNC assumption in the above lemma cannot be omitted in general. On the other hand, the SNC property is very restrictive in several important function spaces. This has been already remarked, e.g., in [73] and [86, Lemma 4.8]. Here we state some results in order to point out the difficulties.

Lemma 2.19. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain and consider the following nonempty, closed, convex cones:

$$\begin{aligned} C\left(\overline{\Omega}\right)_0^+ &:= \left\{u \in C\left(\overline{\Omega}\right) \,\middle|\, u(\omega) \geq 0 \text{ for all } \omega \in \Omega\right\}, \\ L^p(\Omega)_0^+ &:= \left\{u \in L^p(\Omega) \,\middle|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\right\}, \\ W^{1,p}(\Omega)_0^+ &:= \left\{u \in W^{1,p}(\Omega) \,\middle|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\right\}. \end{aligned}$$

Then $C\left(\overline{\Omega}\right)_0^+$ and $L^\infty(\Omega)_0^+$ are SNC at all of their points. On the other hand, for any $p \in [1, \infty)$, the cone $L^p(\Omega)_0^+$ is nowhere SNC.
Let Ω possess a Lipschitz boundary and choose $p \in (1, \infty)$ arbitrarily. If $p \leq d$ is satisfied, then the cone $W^{1,p}(\Omega)_0^+$ is nowhere SNC. On the other hand, if $p > d$ holds, then $W^{1,p}(\Omega)_0^+$ is SNC everywhere.
For an interval $\Omega := (0, T) \subseteq \mathbb{R}$, the nonempty, closed, convex cone

$$AC^{1,2}(\Omega)_0^+ := \{u \in AC^{1,2}(\Omega, \mathbb{R}) \,|\, u(\omega) \geq 0 \text{ for all } \omega \in \Omega\}$$

is SNC at all of its points.

Proof. Since the cones $C\left(\overline{\Omega}\right)_0^+$ and $L^\infty(\Omega)_0^+$ possess a nonempty interior and satisfy $\operatorname{lin} C\left(\overline{\Omega}\right)_0^+ = C\left(\overline{\Omega}\right)$ and $\operatorname{lin} L^\infty(\Omega)_0^+ = L^\infty(\Omega)$, their property to be SNC everywhere follows from [90, Theorem 1.21].
Fix $p \in [1, \infty)$, an arbitrary function $\bar{u} \in L^p(\Omega)_0^+$, and a sequence $\{\Omega_k\} \subseteq 2^\Omega$ of measurable sets satisfying $\mathfrak{l}(\Omega_k) \searrow 0$. We define $u_k := \bar{u}(1 - \chi_{\Omega_k}) \in L^p(\Omega)_0^+$ for any $k \in \mathbb{N}$. Then $u_k \to \bar{u}$ in $L^p(\Omega)$ follows easily from Lemma A.1. Now, it is possible to introduce

$$\forall p \in (1, \infty)\, \forall \omega \in \Omega\colon \quad \eta_k(\omega) := -\mathfrak{l}(\Omega_k)^{-\frac{1}{p'}} \chi_{\Omega_k}(\omega)$$

or $\eta_k := -\chi_{\Omega_k} \in L^\infty(\Omega)$ in the case $p = 1$ for any $k \in \mathbb{N}$. Here p' is the conjugate coefficient of p. For all $p \in [1, \infty)$ and $k \in \mathbb{N}$, we easily obtain

$$\eta_k \in \left(-L^{p'}(\Omega)_0^+\right) \cap \{u_k\}^\perp = (L^p(\Omega)_0^+)^\circ \cap \{u_k\}^\perp = \mathcal{N}_{L^p(\Omega)_0^+}(u_k) = \widehat{\mathcal{N}}_{L^p(\Omega)_0^+}(u_k)$$

from [17, Example 2.64]. Choosing $p \in (1, \infty)$ and $v \in L^p(\Omega)$, we obviously have $v \in L^p(\Omega_k)$ for all $k \in \mathbb{N}$ and Hölder's inequality yields

$$\left|\langle \eta_k, v\rangle_{L^p(\Omega)}\right| = \mathfrak{l}(\Omega_k)^{-\frac{1}{p'}} \left|\int_{\Omega_k} \chi_{\Omega_k}(\omega) v(\omega) \mathrm{d}\omega\right| \leq \mathfrak{l}(\Omega_k)^{-\frac{1}{p'}} \|\chi_{\Omega_k}\|_{L^{p'}(\Omega_k)} \|v\|_{L^p(\Omega_k)} = \left(\int_{\Omega_k} |v(\omega)|^p \mathrm{d}\omega\right)^{\frac{1}{p}}.$$

Using Lemma A.1 again, the latter integral tends to zero, i.e. $\eta_k \xrightarrow{\star} 0$. On the other hand, $\|\eta_k\|_{L^{p'}(\Omega)} = 1$ holds true for any $k \in \mathbb{N}$. Now, let $p = 1$ and $v \in L^1(\Omega)$ be given. Similarly as above, the right hand side of

$$\left|\langle \eta_k, v\rangle_{L^1(\Omega)}\right| \leq \int_{\Omega_k} |v(\omega)| \mathrm{d}\omega$$

converges to zero as k tends to ∞, i.e. $\eta_k \xrightarrow{\star} 0$ follows. By construction $\|\eta_k\|_{L^\infty(\Omega)} = 1$ is satisfied for all $k \in \mathbb{N}$. Summing up all these considerations, we constructed a sequence $\{\eta_k\}$ such that $\eta_k \xrightarrow{\star} 0$ and $\eta_k \nrightarrow 0$ hold in parallel. Thus, $L^p(\Omega)_0^+$ cannot be SNC at $\bar{u}$.
For $p \in (1, \infty)$, the Banach space $W^{1,p}(\Omega)$ is reflexive. Choosing $u \in W^{1,p}(\Omega)$, the functions $\max\{0; u\}$ and $\max\{0; -u\}$ belong to $W^{1,p}(\Omega)$ as well, see [75, Theorem A.1.]. From $u = \max\{0; u\} - \max\{0; -u\}$ we deduce $\operatorname{lin} W^{1,p}(\Omega)_0^+ = W^{1,p}(\Omega)_0^+ - W^{1,p}(\Omega)_0^+ = W^{1,p}(\Omega)$. We distinguish between two cases.
Case I: Let $p \leq d$ hold. Following Lemma 2.17, we only need to show that $\operatorname{rint} W^{1,p}(\Omega)_0^+$ is empty in order to verify the lack of the SNC property. Clearly, since we have $\operatorname{lin} W^{1,p}(\Omega)_0^+ = W^{1,p}(\Omega)$, $\operatorname{rint} W^{1,p}(\Omega)_0^+$ equals $\operatorname{int} W^{1,p}(\Omega)_0^+$. Thus, it is sufficient to show $\operatorname{int} W^{1,p}(\Omega)_0^+ = \varnothing$. Therefore, we first verify the existence of functions in $W^{1,p}(\Omega) \setminus L^\infty(\Omega)$. We assume on the contrary that $W^{1,p}(\Omega) \subseteq L^\infty(\Omega)$ holds. Take a sequence $\{u_k\} \subseteq W^{1,p}(\Omega)$ converging to $\hat{u}$ in $W^{1,p}(\Omega)$ and assume that $\{u_k\}$ converges in $L^\infty(\Omega)$ to $\tilde{u}$ at the same time. Then we have

$$\|\hat{u} - \tilde{u}\|_{L^p(\Omega)} \leq \|\hat{u} - u_k\|_{L^p(\Omega)} + \|u_k - \tilde{u}\|_{L^p(\Omega)} \leq \|\hat{u} - u_k\|_{W^{1,p}(\Omega)} + \mathfrak{l}(\Omega)^{\frac{1}{p}} \|u_k - \tilde{u}\|_{L^\infty(\Omega)} \to 0$$

from the triangle inequality. Thus, $\hat{u} = \tilde{u}$ holds and, hence, we can deduce $W^{1,p}(\Omega) \hookrightarrow L^\infty(\Omega)$ from [126, Satz IV.4.5]. This contradicts Sobolev's embedding theorem since for $p < d$, only $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $p \leq q \leq \frac{dp}{d-p}$ holds and the upper bound on q is tight, or for $p = d$, only $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $p \leq q < \infty$ is satisfied, see Theorem 2.8. Thus, there exists $\bar{v} \in W^{1,p}(\Omega)_0^+ \setminus L^\infty(\Omega)$. Choose an arbitrary function $\bar{u} \in W^{1,p}(\Omega)_0^+$ and define u_k for any $k \in \mathbb{N}$ as stated below:

$$\forall \omega \in \Omega\colon \quad u_k(\omega) := \min\{k; \bar{u}(\omega)\} - \tfrac{1}{k}\bar{v}(\omega).$$

From Lemma A.4 we have $\{u_k\} \subseteq W^{1,p}(\Omega)$ and $u_k \to \bar{u}$ in $W^{1,p}(\Omega)$. On the other hand, the boundedness of $\omega \mapsto \min\{k; \bar{u}(\omega)\}$ and the unboundedness of $\bar{v}$ yield $u_k \notin W^{1,p}(\Omega)_0^+$. Thus, the interior of $W^{1,p}(\Omega)_0^+$ is empty and, hence, this cone does not possess the SNC property at its elements.
Case II: Suppose that $p > d$ is satisfied. Then we have $W^{1,p}(\Omega) \hookrightarrow C\left(\overline{\Omega}\right)$ by Theorem 2.8. Consequently, $\operatorname{int} W^{1,p}(\Omega)_0^+ \neq \varnothing$ is obtained from $\operatorname{int} C\left(\overline{\Omega}\right)_0^+ \neq \varnothing$. From above we have $\operatorname{lin} W^{1,p}(\Omega)_0^+ = W^{1,p}(\Omega)$. Combining these facts with Lemma 2.17, the cone $W^{1,p}(\Omega)_0^+$ is SNC at all of its points.
Finally, let $\Omega := (0, T) \subseteq \mathbb{R}$ be a real interval. Applying Theorem 2.10, it is not difficult to see the relations $\operatorname{lin} AC^{1,2}(\Omega)_0^+ = AC^{1,2}(\Omega, \mathbb{R})$ and $\operatorname{int} AC^{1,2}(\Omega)_0^+ \neq \varnothing$. Thus, the cone $AC^{1,2}(\Omega)_0^+$ is SNC at any of its points by means of Lemma 2.17. □

Adapting the proof of the above lemma, we obtain the following results for sets in function spaces defined via lower and upper bounds.

Corollary 2.20. Fix some $\varepsilon > 0$. For an arbitrary bounded domain $\Omega \subseteq \mathbb{R}^d$, the following sets are SNC at all of their points:

$$\left\{u \in C\left(\overline{\Omega}\right) \,\middle|\, a(\omega) \leq u(\omega) \leq b(\omega) \text{ for all } \omega \in \Omega\right\} \qquad a, b \in C\left(\overline{\Omega}\right),\ a(\omega) < b(\omega) \text{ for all } \omega \in \overline{\Omega},$$
$$\{u \in L^\infty(\Omega) \,|\, a(\omega) \leq u(\omega) \leq b(\omega) \text{ f.a.a. } \omega \in \Omega\} \qquad a, b \in L^\infty(\Omega),\ b(\omega) - a(\omega) \geq \varepsilon \text{ f.a.a. } \omega \in \Omega.$$

On the other hand, for any $p \in [1,\infty)$, the set

$$\{u \in L^p(\Omega) \,|\, a(\omega) \leq u(\omega) \leq b(\omega) \text{ f.a.a. } \omega \in \Omega\} \qquad a,b \in L^p(\Omega),\, a(\omega) \leq b(\omega) \text{ f.a.a. } \omega \in \Omega$$

is nowhere SNC.
For an interval $\Omega := (0,T) \subseteq \mathbb{R}$, the set

$$\{u \in AC^{1,2}(\Omega,\mathbb{R}) \,|\, a(\omega) \leq u(\omega) \leq b(\omega) \text{ for all } \omega \in \Omega\} \qquad a,b \in AC^{1,2}(\Omega,\mathbb{R}),\, a(\omega) < b(\omega) \text{ for all } \omega \in \overline{\Omega}$$

is SNC at all of its points.
Similar assertions hold true for sets defined via only upper or lower bounds and for the function spaces $C\left(\overline{\Omega}\right)^n$, $L^p(\Omega,\mathbb{R}^n)$, $p \in [1,\infty]$, and $AC^{1,2}(\Omega,\mathbb{R}^n)$.

Finally, we consider sets with upper and lower bounds in Sobolev spaces.

Lemma 2.21. Let $\Omega \subseteq \mathbb{R}^d$ be an arbitrary bounded domain with Lipschitz continuous boundary and fix $p \in (1,\infty)$. Furthermore, let $a,b \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ be functions which satisfy $b(\omega) - a(\omega) \geq \varepsilon$ almost everywhere on Ω where $\varepsilon > 0$ is a fixed constant. For $p \leq d$, the set

$$S := \{u \in W^{1,p}(\Omega) \,|\, a(\omega) \leq u(\omega) \leq b(\omega) \text{ f.a.a. } \omega \in \Omega\}$$

is nowhere SNC. On the other hand, if $p > d$ is satisfied, then S is SNC everywhere.

Proof. We first show $\lin S = W^{1,p}(\Omega) \cap L^\infty(\Omega)$. Therefore, we invoke Lemma 2.2 in order to see the relation $\lin S = \operatorname{conv} \bigcup_{\alpha\in\mathbb{R}} \alpha S$. Since we clearly have $\alpha S \subseteq W^{1,p}(\Omega) \cap L^\infty(\Omega)$ for any $\alpha \in \mathbb{R}$, the inclusion $\subseteq$ is obvious. Now, take any function $\bar{u} \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$. Then we find a real constant $C > 0$ such that $|\bar{u}(\omega)| \leq C$ holds almost everywhere on Ω. On the other hand, with the function $\bar{w} := a + \frac{1}{2}(b-a) \in S$ we obtain

$$S' := \{u \in W^{1,p}(\Omega) \,|\, -\tfrac{\varepsilon}{2} \leq u(\omega) \leq \tfrac{\varepsilon}{2} \text{ f.a.a. } \omega \in \Omega\} \subseteq S - \{\bar{w}\} \subseteq \lin S.$$

Moreover, $\bar{u} \in \frac{2C}{\varepsilon} S' \subseteq \lin S$ is satisfied. Hence, we have shown $W^{1,p}(\Omega) \cap L^\infty(\Omega) \subseteq \lin S$.
First, we assume $p \leq d$. Similarly as in the proof of Lemma 2.19, we can show the existence of a function $\bar{v} \in W^{1,p}(\Omega)_0^+ \setminus L^\infty(\Omega)$. On the other hand, the sequence $\{v_k\} \subseteq W^{1,p}(\Omega) \cap L^\infty(\Omega)$ defined by

$$\forall k \in \mathbb{N}\, \forall \omega \in \Omega\colon \quad v_k(\omega) := \min\{k; \bar{v}(\omega)\}$$

converges to $\bar{v}$ w.r.t. the $W^{1,p}(\Omega)$-norm, see Lemma A.4. Thus, $\lin S = W^{1,p}(\Omega) \cap L^\infty(\Omega)$ is not closed. Applying Lemma 2.17, S is nowhere SNC.
Finally, let us assume $p > d$. Then we have $W^{1,p}(\Omega) \hookrightarrow C\left(\overline{\Omega}\right)$ from Theorem 2.8 and, consequently, any function from $W^{1,p}(\Omega)$ needs to be bounded. This yields $\lin S = W^{1,p}(\Omega) \cap L^\infty(\Omega) = W^{1,p}(\Omega)$. On the other hand, we obtain $\bar{w} \in \operatorname{int} S = \operatorname{rint} S$ in this situation. Summing up these arguments, S is SNC everywhere by means of Lemma 2.17. □

2.3.2. Some facts on vector lattices

For some binary relation $\varrho \subseteq S \times S$ of a nonempty set S, we exploit the infix notation $x \varrho y$ in order to express $(x,y) \in \varrho$ for $x,y \in S$. Recall that ϱ is called

$$\begin{aligned}
&\text{reflexive} && \iff && \forall x \in S\colon\ x \varrho x,\\
&\text{antisymmetric} && \iff && \forall x,y \in S\colon\ x \varrho y \wedge y \varrho x \implies x = y,\\
&\text{transitive} && \iff && \forall x,y,z \in S\colon\ x \varrho y \wedge y \varrho z \implies x \varrho z.
\end{aligned}$$

A reflexive, antisymmetric, and transitive binary relation $\varrho \subseteq S \times S$ is called a partial order of S, the pair (S,ϱ) is said to be a partially ordered set.
Now, let (S,ϱ) be a partially ordered set and fix $x,y \in S$. Then $s \in S$ is called the supremum of x and y if $x \varrho s$ as well as $y \varrho s$ hold and the condition

$$\forall z \in S\colon \quad x \varrho z \wedge y \varrho z \implies s \varrho z$$

is satisfied, i.e. s is the smallest upper bound of $\{x, y\}$ in S w.r.t. ϱ. In case of existence, we denote the supremum s of x and y by $\max\{x; y\}$. If for all $x, y \in S$, $\max\{x; y\}$ exists, then (S, ϱ) is called an upper semilattice.
Let $\mathcal{X}$ be a Banach space and let $K \subseteq \mathcal{X}$ be a closed, convex cone which is pointed, i.e. which satisfies $K \cap (-K) = \{0\}$. We define a binary relation $\leq_K \subseteq \mathcal{X} \times \mathcal{X}$ as stated below:

$$\forall x, y \in \mathcal{X}: \quad x \leq_K y \iff y - x \in K.$$

It is easily seen, cf. [71, Theorem D.3], that $(\mathcal{X}, \leq_K)$ is a partially ordered set. Moreover, we obtain the following calculus rules for $\leq_K$ arising from the properties of the Banach space $\mathcal{X}$:

$$\begin{aligned} &\forall x, y, z \in \mathcal{X}: && x \leq_K y \implies x + z \leq_K y + z, \\ &\forall x, y \in \mathcal{X}\, \forall \alpha \geq 0: && x \leq_K y \implies \alpha x \leq_K \alpha y. \end{aligned}$$

If $(\mathcal{X}, \leq_K)$ is an upper semilattice, it is called a vector lattice.
Now, let $(\mathcal{X}, \leq_K)$ be a vector lattice. Then the above calculus rules imply

$$\begin{aligned} &\forall x, y, z \in \mathcal{X}: && \max_K\{x; y\} + z = \max_K\{x + z; y + z\}, \\ &\forall x, y \in \mathcal{X}\, \forall \alpha \geq 0: && \max_K\{\alpha x; \alpha y\} = \alpha \max_K\{x; y\}. \end{aligned}$$

Here we added the index K to the supremum operator in order to emphasize that it is induced by the cone K. Later on this notation will avoid confusion. We are going to exploit the natural definition $\min_K\{x; y\} := -\max_K\{-x; -y\}$ for any $x, y \in \mathcal{X}$. Note that

$$x = x + \max_K\{-x; 0\} - \max_K\{-x; 0\} = \max_K\{0; x\} - \max_K\{-x; 0\} = \max_K\{x; 0\} + \min_K\{x; 0\}$$

is satisfied for all $x \in \mathcal{X}$. Clearly, we can identify $\min_K\{x; y\}$ with the largest lower bound of $\{x, y\}$ in $\mathcal{X}$ w.r.t. $\leq_K$, i.e. with the infimum of the set $\{x, y\}$. Consequently, for arbitrary $x, y \in \mathcal{X}$, the infimum $\min_K\{x; y\}$ exists and, thus, $(\mathcal{X}, \leq_K)$ is a so-called lower semilattice as well. We easily check

$$\forall x, y \in \mathcal{X}: \quad x \leq_K y \iff \max_K\{x; y\} = y \iff \min_K\{x; y\} = x.$$

Thus, $(\mathcal{X}, \leq_K)$ is a lattice in classical sense, see [53]. This justifies the name *vector lattice*.

Example 2.22. In $\mathbb{R}^3$, consider the closed, convex, pointed, and polyhedral cones

$$\begin{aligned} K_1 &:= \{x \in \mathbb{R}^3 \,|\, (-1, 0, 0) \cdot x \leq 0,\ (0, -1, 0) \cdot x \leq 0,\ (0, 0, -1) \cdot x \leq 0\} = \mathbb{R}^{3,+}_0, \\ K_2 &:= \{x \in \mathbb{R}^3 \,|\, (-1, 1, 1) \cdot x \leq 0,\ (-1, 1, -1) \cdot x \leq 0,\ (-1, -1, -1) \cdot x \leq 0,\ (-1, -1, 1) \cdot x \leq 0\}. \end{aligned}$$

Clearly, the cone K_1 induces the common less-or-equal binary relation $\leq$ in $\mathbb{R}^3$ and $(\mathbb{R}^3, \leq)$ is obviously a vector lattice. On the other hand, K_2 does not induce a vector lattice in $\mathbb{R}^3$ since there does not exist a supremum of, e.g., the points $(0, 0, 0)$ and $(0, 1, 1)$. The set $\mathcal{U}$ of upper bounds of $\{(0, 0, 0), (0, 1, 1)\}$ takes the form

$$\mathcal{U} = \operatorname{conv}\{(1, 1, 0), (1, 0, 1)\} + K_2$$

but this set does not possess a smallest element w.r.t. $\leq_{K_2}$. ■

For a detailed introduction to the theory of vector lattices, we refer the interested reader to [110].
Later, we need some more calculus rules for supremum and infimum in connection with tangent cones.

Lemma 2.23. Let $(\mathcal{X}, \leq_K)$ be a vector lattice induced by the closed, convex, pointed cone $K \subseteq \mathcal{X}$ which satisfies the following condition:

$$\forall \{x_k\} \subseteq \mathcal{X}\, \forall \bar{x} \in \mathcal{X}: \quad x_k \to \bar{x} \implies \max_K\{x_k; 0\} \rightharpoonup \max_K\{\bar{x}; 0\}. \tag{2.7}$$

Furthermore, for $x \in K$, choose $d \in \mathcal{T}_K(x)$ and $r \in -\mathcal{T}_K(x)$. Then $\max_K\{d; 0\}, \min_K\{d; 0\} \in \mathcal{T}_K(x)$ and $\max_K\{r; 0\}, \min_K\{r; 0\} \in -\mathcal{T}_K(x)$ hold.

Proof. Under condition (2.7), the statements for d follow from [120, Lemma 4.12]. Now, let us prove the assertions on r. Observe that $\max_K\{-r; 0\} \in \mathcal{T}_K(x)$ holds. Hence, we have $-\max_K\{-r; 0\} \in -\mathcal{T}_K(x)$, i.e. $\min_K\{r; 0\} \in -\mathcal{T}_K(x)$. Analogously, we can show $\max_K\{r; 0\} \in -\mathcal{T}_K(x)$. □

As it was remarked in [120, Section 4.2], the property (2.7) is weaker than the property of $(\mathcal{X}, \leq_K)$ to be a Banach lattice, see also [110] for details. One may check [120, Lemma 4.8] for the proof of the following result presenting a condition which guarantees that (2.7) is valid.

Lemma 2.24. Assume that $\mathcal{X}$ is a reflexive Banach space and let $(\mathcal{X}, \leq_K)$ be a vector lattice induced by the closed, convex, pointed cone $K \subseteq \mathcal{X}$. If there is a constant $c > 0$ which satisfies

$$\forall x \in \mathcal{X}: \quad \|\max_K\{x; 0\}\|_{\mathcal{X}} \leq c \|x\|_{\mathcal{X}},$$

then condition (2.7) is valid.

Example 2.25. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete and σ-finite measure space. For $p \in [1, \infty]$, the cone

$$L^p(\mathfrak{M})^+_0 := \{u \in L^p(\mathfrak{M}) \,|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\}$$

induces the vector lattice $\left(L^p(\mathfrak{M}), \leq_{L^p(\mathfrak{M})^+_0}\right)$. Using the obvious inequality

$$\forall \alpha, \beta \in \mathbb{R}: \quad |\max\{\alpha; 0\} - \max\{\beta; 0\}| \leq |\alpha - \beta|,$$

we easily see that the mapping $L^p(\mathfrak{M}) \ni u \mapsto \max_{L^p(\mathfrak{M})^+_0}\{u; 0\} \in L^p(\mathfrak{M})$ is Lipschitz continuous with Lipschitz modulus $L = 1$. Thus, the assumption of Lemma 2.23 holds.
Now, let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain and consider

$$H^1_0(\Omega)^+_0 := \{u \in H^1_0(\Omega) \,|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\}.$$

The pair $\left(H^1_0(\Omega), \leq_{H^1_0(\Omega)^+_0}\right)$ is a vector lattice, see [17, Lemma 6.11 and Proposition 6.45], and the corresponding supremum operator $H^1_0(\Omega) \ni u \mapsto \max_{H^1_0(\Omega)^+_0}\{u; 0\} \in H^1_0(\Omega)$ is continuous, i.e. the property (2.7) holds. A similar argumentation is possible for $H^1(\Omega)$ equipped with the cone of all almost everywhere nonnegative functions which forms a vector lattice satisfying (2.7) as well, see [5, Theorem 5.8.2]. ■

2.3.3. Tools of generalized differentiation

Let $F\colon \mathcal{X} \to \mathcal{Y}$ be a mapping between Banach spaces $\mathcal{X}$ as well as $\mathcal{Y}$ and choose $\bar{x} \in \mathcal{X}$ arbitrarily. Then F is said to be directionally differentiable at $\bar{x}$ in direction $\delta \in \mathcal{X}$ if the limit

$$F'(\bar{x}; \delta) := \lim_{t \searrow 0} \frac{F(\bar{x} + t\delta) - F(\bar{x})}{t}$$

exists. In this case, $F'(\bar{x}; \delta)$ is called directional derivative of F at $\bar{x}$ in direction δ. If $F'(\bar{x}; \delta)$ exists for all δ, then F is called directionally differentiable at $\bar{x}$. Suppose that F is directionally differentiable at $\bar{x}$. If there is a function $o\colon \mathbb{R}^+_0 \to \mathcal{Y}$ which satisfies $\lim_{t \searrow 0} \frac{o(t)}{t} = 0$ and

$$\forall \delta \in \mathcal{X}: \quad F(\bar{x} + \delta) - F(\bar{x}) - F'(\bar{x}; \delta) = o(\|\delta\|_{\mathcal{X}}),$$

then F is called B-differentiable at $\bar{x}$. Note that whenever F is locally Lipschitz continuous at $\bar{x}$ while $\mathcal{X}$ is finite-dimensional, then F is directionally differentiable at $\bar{x}$ if and only if it is B-differentiable at this point, see [112, Proposition 3.5]. On the other hand, F is said to be Fréchet differentiable at $\bar{x}$ if there exists an operator $F'(\bar{x}) \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$ which satisfies

$$0 = \lim_{\|h\|_{\mathcal{X}} \searrow 0} \frac{F(\bar{x} + h) - F(\bar{x}) - F'(\bar{x})[h]}{\|h\|_{\mathcal{X}}}.$$

In case of existence, $F'(\bar{x})$ is called Fréchet derivative of F at $\bar{x}$. Obviously, if F is Fréchet differentiable at $\bar{x}$, it is continuous there. If the mapping $x \mapsto F'(x)$ is well-defined in a neighborhood of $\bar{x}$ and continuous at this point, then F is called continuously Fréchet differentiable at $\bar{x}$. Note that any mapping which is Fréchet differentiable at $\bar{x}$ is B-differentiable at $\bar{x}$ and for any direction $\delta \in \mathcal{X}$, $F'(\bar{x}; \delta) = F'(\bar{x})[\delta]$ holds true. The mapping F is called directionally differentiable (B-differentiable, Fréchet differentiable,

continuously Fréchet differentiable) if it possesses this property at all points $x \in \mathcal{X}$. If the mapping F is continuously Fréchet differentiable at $\bar{x}$, it is reasonable to check whether the mapping $F' \colon \mathcal{X} \to \mathbb{L}[\mathcal{X}, \mathcal{Y}]$ is Fréchet differentiable at $\bar{x}$. If the answer is positive, then we call F twice Fréchet differentiable at $\bar{x}$ and exploit the notation $F^{(2)}(\bar{x}) := (F')'(\bar{x}) \in \mathbb{L}[\mathcal{X}, \mathbb{L}[\mathcal{X}, \mathcal{Y}]]$. Note that we may interpret the operator $F^{(2)}(\bar{x})$ as a continuous, symmetric bilinear mapping from $\mathcal{X}^2$ to $\mathcal{Y}$, see [142, Section 10.5]. Recall that a bilinear mapping $\mathsf{b} \colon \mathcal{X}^2 \to \mathcal{Y}$ is continuous if and only if there is a constant $\alpha > 0$ which satisfies

$$\forall x, x' \in \mathcal{X} \colon \quad |\mathsf{b}[x, x']| \leq \alpha \, \|x\|_{\mathcal{X}} \, \|x'\|_{\mathcal{X}} \, .$$

For any $y^* \in \mathcal{Y}^\star$, we introduce $\left\langle y^*, F^{(2)}(\bar{x}) \right\rangle_{\mathcal{Y}} \in \mathbb{L}[\mathcal{X}, \mathcal{X}^\star]$ as stated below:

$$\forall x \in \mathcal{X} \colon \quad \left\langle y^*, F^{(2)}(\bar{x}) \right\rangle_{\mathcal{Y}} [x] := \left\langle y^*, F^{(2)}(\bar{x})[x, \cdot] \right\rangle_{\mathcal{Y}} = F^{(2)}(\bar{x})[x, \cdot]^\star[y^*].$$

If $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ holds for Banach spaces $\mathcal{X}_1$ and $\mathcal{X}_2$, then partial Fréchet differentiability can be defined considering $x_1 \mapsto F(x_1, \bar{x}_2)$ and $x_2 \mapsto F(\bar{x}_1, x_2)$ where $\bar{x} = (\bar{x}_1, \bar{x}_2)$. We write $F'_{x_i}(\bar{x}) \in \mathbb{L}[\mathcal{X}_i, \mathcal{Y}]$, $i = 1, 2$, in order to address the partial Fréchet derivatives if they exist. It is well-known that any Fréchet differentiable mapping is partially Fréchet differentiable w.r.t. all of its variables. On the other hand, if a mapping is continuously partially Fréchet differentiable w.r.t. all of its variables, it is continuously Fréchet differentiable, see [142, Section 10.4.2]. A similar notion will be used for partial directional derivatives. Partial second-order Fréchet derivatives can be introduced in an analogous way. The notation $F^{(2)}_{x_i x_j}(\bar{x}) := (F'_{x_i})'_{x_j}(\bar{x}) \in \mathbb{L}[\mathcal{X}_j, \mathbb{L}[\mathcal{X}_i, \mathcal{Y}]]$ is exploited, and $F^{(2)}_{x_i x_j}(\bar{x})$ will be interpreted as a continuous bilinear mapping from $\mathcal{X}_i \times \mathcal{X}_j$ to $\mathcal{Y}$. Calculus rules for the computation of directional and Fréchet derivatives are presented in [17] and [142].

Example 2.26. Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be Banach spaces and fix $\mathtt{A} \in \mathbb{L}[\mathcal{X}, \mathcal{Y}]$, $y_\mathsf{d} \in \mathcal{Y}$, and a symmetric, continuous bilinear mapping $\mathsf{b} \colon \mathcal{Y}^2 \to \mathcal{Z}$. Then straightforward calculations show that $F \colon \mathcal{X} \to \mathcal{Z}$ given by

$$\forall x \in \mathcal{X} \colon \quad F(x) := \tfrac{1}{2} \mathsf{b}[\mathtt{A}[x] - y_\mathsf{d}, \mathtt{A}[x] - y_\mathsf{d}]$$

is twice continuously Fréchet differentiable with the Fréchet derivatives

$$\forall h, h' \in \mathcal{X} \colon \quad F'(\bar{x})[h] = \mathsf{b}[\mathtt{A}[h], \mathtt{A}[\bar{x}] - y_\mathsf{d}], \quad F^{(2)}(\bar{x})[h, h'] := \mathsf{b}[\mathtt{A}[h], \mathtt{A}[h']]$$

at any $\bar{x} \in \mathcal{X}$. ■

Example 2.27. We consider a similar situation as in Example 2.26. Let $\mathcal{X}$ and $\mathcal{H}$ be a Banach space and a Hilbert space, respectively. Here we identify $\mathcal{H}$ and $\mathcal{H}^\star$ by means of Riesz's representation theorem. Fix $\mathtt{A} \in \mathbb{L}[\mathcal{X}, \mathcal{H}]$ and $y_\mathsf{d} \in \mathcal{H}$. Then the mapping $J \colon \mathcal{X} \to \mathbb{R}$ defined by

$$\forall x \in \mathcal{X} \colon \quad J(x) := \tfrac{1}{2} \left\langle \mathtt{A}[x] - y_\mathsf{d}, \mathtt{A}[x] - y_\mathsf{d} \right\rangle_{\mathcal{H}} = \tfrac{1}{2} \left\| \mathtt{A}[x] - y_\mathsf{d} \right\|^2_{\mathcal{H}}$$

is twice continuously Fréchet differentiable at any point $\bar{x} \in \mathcal{X}$ and the Fréchet derivatives take the form

$$\forall h, h' \in \mathcal{X} \colon \quad J'(\bar{x})[h] = \left\langle \mathtt{A}[h], \mathtt{A}[\bar{x}] - y_\mathsf{d} \right\rangle_{\mathcal{H}} , \quad J^{(2)}(\bar{x})[h, h'] = \left\langle \mathtt{A}[h], \mathtt{A}[h'] \right\rangle_{\mathcal{H}} .$$

Using the definition of adjoint operators, we find the reasonable representation

$$J'(\bar{x}) = \mathtt{A}^\star[\mathtt{A}[\bar{x}] - y_\mathsf{d}].$$

In optimal control, the so-called tracking-functional J is a typical objective. It is easily seen that J is a convex functional as well. Moreover, if $\mathtt{A}$ is elliptic (in the case $\mathcal{X} = \mathcal{H}$), then J is coercive since we have

$$\begin{aligned} J(x) &= \tfrac{1}{2} \left\| \mathtt{A}[x] \right\|^2_{\mathcal{H}} - \left\langle \mathtt{A}[x], y_\mathsf{d} \right\rangle_{\mathcal{H}} + \tfrac{1}{2} \left\| y_\mathsf{d} \right\|^2_{\mathcal{H}} \geq \tfrac{1}{2} \left\| \mathtt{A}[x] \right\|^2_{\mathcal{H}} - \left\| \mathtt{A}[x] \right\|_{\mathcal{H}} \left\| y_d \right\|_{\mathcal{H}} \\ &= \tfrac{1}{2} \left\| \mathtt{A}[x] \right\|_{\mathcal{H}} \left(\left\| \mathtt{A}[x] \right\|_{\mathcal{H}} - 2 \left\| y_\mathsf{d} \right\|_{\mathcal{H}} \right) \geq \tfrac{\alpha^2}{2} \left\| x \right\|_{\mathcal{X}} \left(\left\| x \right\|_{\mathcal{X}} - \tfrac{2}{\alpha} \left\| y_\mathsf{d} \right\|_{\mathcal{H}} \right) \end{aligned}$$

for all $x \in \mathcal{X}$ with $\|x\|_{\mathcal{X}} \geq \frac{2}{\alpha} \|y_\mathsf{d}\|_{\mathcal{H}}$ where $\alpha > 0$ denotes the constant from the characterization of ellipticity, see Example 2.14. Thus, for any nonempty, closed, convex set $M \subseteq \mathcal{H}$, the optimization problem

$$\begin{aligned} \tfrac{1}{2} \left\| y - y_\mathsf{d} \right\|^2_{\mathcal{H}} &\to \min \\ y &\in M \end{aligned}$$

possesses an optimal solution $\bar{y}$ by means of Lemma 2.5. Since the objective functional is strictly convex, this solution is unique. Clearly, $\bar{y}$ is the projection of y_d onto M. We exploit $\mathrm{proj}_M(y_\mathsf{d}) := \bar{y}$. ■

Example 2.28. Finally, let $\mathcal{X}$ be a reflexive Banach space and let $\mathcal{U}$ be an arbitrary Banach space. For a self-adjoint operator $\mathtt{A} \in \mathbb{L}[\mathcal{X}, \mathcal{X}^\star]$ and another arbitrary operator $\mathtt{B} \in \mathbb{L}[\mathcal{U}, \mathcal{X}^\star]$, we consider the functional $J\colon \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ given by

$$\forall x \in \mathcal{X}\, \forall u \in \mathcal{U}\colon \quad J(x,u) := \tfrac{1}{2} \langle \mathtt{A}[x], x \rangle_{\mathcal{X}} - \langle \mathtt{B}[u], x \rangle_{\mathcal{X}} .$$

At any point $(\bar{x}, \bar{u}) \in \mathcal{X} \times \mathcal{U}$, it is twice continuously Fréchet differentiable and satisfies

$$\begin{aligned} \forall h, h' \in \mathcal{X}\, \forall k \in \mathcal{U}\colon \qquad J'_x(\bar{x},\bar{u})[h] &= \langle \mathtt{A}[\bar{x}] - \mathtt{B}[\bar{u}], h \rangle_{\mathcal{X}} , & J'_u(\bar{x},\bar{u})[k] &= -\langle \mathtt{B}[k], \bar{x} \rangle_{\mathcal{X}} , \\ J^{(2)}_{xx}(\bar{x},\bar{u})[h,h'] &= \langle \mathtt{A}[h'], h \rangle_{\mathcal{X}} , & J^{(2)}_{xu}(\bar{x},\bar{u})[h,k] &= -\langle \mathtt{B}[k], h \rangle_{\mathcal{X}} . \end{aligned}$$

Exploiting the concept of adjoint operators, we obtain

$$J'_x(\bar{x},\bar{u}) = \mathtt{A}[\bar{x}] - \mathtt{B}[\bar{u}], \quad J'_u(\bar{x},\bar{u}) = -\mathtt{B}^\star[\bar{x}].$$

If $\mathtt{A}$ is elliptic, then $x \mapsto J(x,\bar{u})$ is coercive and strictly convex. If, additionally, the operator $\mathtt{B}^\star$ is compact, then J is weakly lower semicontinuous, since for any sequences $\{x_k\} \subseteq \mathcal{X}$ and $\{u_k\} \subseteq \mathcal{U}$ satisfying $x_k \rightharpoonup \bar{x}$ and $u_k \rightharpoonup \bar{u}$, we have $\mathtt{B}^\star[x_k] \to \mathtt{B}^\star[\bar{x}]$ and, by means of Lemma 2.4,

$$\lim_{k\to\infty} \langle \mathtt{B}[u_k], x_k \rangle_{\mathcal{X}} = \lim_{k\to\infty} \langle \mathtt{B}^\star[x_k], u_k \rangle_{\mathcal{U}} = \langle \mathtt{B}^\star[\bar{x}], \bar{u} \rangle_{\mathcal{U}} = \langle \mathtt{B}[\bar{u}], \bar{x} \rangle_{\mathcal{X}} .$$

Note that the lower level program of the obstacle problem, see (1.4), possesses an objective functional of the type presented in this example. ■

For a functional $\psi\colon \mathcal{X} \to \overline{\mathbb{R}}$, a point $\bar{x} \in \mathcal{X}$ where $\psi(\bar{x})$ is finite, and $\delta \in \mathcal{X}$, we define the Clarke generalized directional derivative of ψ at $\bar{x}$ in direction δ as stated below:

$$\psi^\circ(\bar{x};\delta) := \limsup_{x \to \bar{x},\, t \searrow 0} \frac{\psi(x+t\delta) - \psi(x)}{t}.$$

The set

$$\partial^c \psi(\bar{x}) := \{x^* \in \mathcal{X}^\star \,|\, \forall \delta \in \mathcal{X}\colon \ \langle x^*, \delta \rangle_{\mathcal{X}} \le \psi^\circ(\bar{x};\delta)\}$$

is referred to as Clarke subdifferential of ψ at $\bar{x}$. Supposing that ψ is locally Lipschitz continuous in a neighborhood of $\bar{x}$, the set $\partial^c\psi(\bar{x})$ is nonempty, closed, convex, and bounded, see [24, Proposition 2.1.2]. Let $\operatorname{epi}\psi := \{(x,\alpha) \in \mathcal{X} \times \mathbb{R} \,|\, \psi(x) \le \alpha\}$ denote the epigraph of ψ. Then

$$\partial^c \psi(\bar{x}) = \left\{ x^* \in \mathcal{X}^\star \,\middle|\, (x^*, -1) \in \mathcal{N}^c_{\operatorname{epi}\psi}(\bar{x}, \psi(\bar{x})) \right\} \tag{2.8}$$

is satisfied if ψ is locally Lipschitz continuous at $\bar{x}$. In [24], many other different calculus rules for Clarke subdifferentials are presented. Especially, the mapping $\psi \mapsto \partial^c\psi(\bar{x})$ is linear on the set of all functions which are locally Lipschitz at $\bar{x}$. On the other hand, it is well-known that due to the convexity of the Clarke subdifferential, this set is comparatively large. Thus, necessary optimality conditions derived via Clarke's tools turn out to be rather weak in many situations. Noting the representation in (2.8) and recalling the relationship of the Clarke normal cone to the limiting normal cone, it is reasonable to define the limiting (or basic, Mordukhovich) subdifferential of ψ at $\bar{x}$ via

$$\partial \psi(\bar{x}) := \{x^* \in \mathcal{X}^\star \,|\, (x^*, -1) \in \mathcal{N}_{\operatorname{epi}\psi}(\bar{x}, \psi(\bar{x}))\}$$

in order to try to overcome this shortcoming of the Clarke subdifferential. Clearly, for convex functionals, these subdifferential constructions coincide since the epigraph of a convex functional is a convex set. For a functional ψ which is continuously Fréchet differentiable at $\bar{x}$,

$$\partial \psi(\bar{x}) = \partial^c \psi(\bar{x}) = \{\psi'(\bar{x})\}$$

is obtained. In general, $\partial\psi(\bar{x})$ is a nonconvex set which is nonempty if ψ is locally Lipschitz at $\bar{x}$. Moreover, the mapping $\psi \mapsto \partial\psi(\bar{x})$ is positively homogeneous on the set of functions being locally Lipschitz at $\bar{x}$. In reflexive Banach spaces, we especially obtain $\partial^c\psi(\bar{x}) = \overline{\operatorname{conv}}\, \partial\psi(\bar{x})$ if ψ is locally Lipschitz at $\bar{x}$ which yields

$$\partial(-\psi)(\bar{x}) \subseteq \partial^c(-\psi)(\bar{x}) = -\partial^c\psi(\bar{x}) = -\,\overline{\operatorname{conv}}\, \partial\psi(\bar{x})$$

Other calculus rules for the limiting subdifferential can be found in [90].

Let $\Psi\colon \mathcal{X} \rightrightarrows \mathcal{Y}$ be a set-valued mapping, i.e. Ψ maps elements of $\mathcal{X}$ to subsets of $\mathcal{Y}$. We define its domain, kernel, and graph by $\operatorname{dom}\Psi := \{x \in \mathcal{X} \,|\, \Psi(x) \neq \varnothing\}$, $\ker\Psi := \{x \in \mathcal{X} \,|\, 0 \in \Psi(x)\}$, and $\operatorname{gph}\Psi := \{(x,y) \in \mathcal{X} \times \mathcal{Y} \,|\, y \in \Psi(x)\}$, respectively. For any set $C \subseteq \mathcal{Y}$, $\Psi^{-1}(C) := \{x \in \mathcal{X} \,|\, \Psi(x) \cap C \neq \varnothing\}$ represents the preimage of C under Ψ. The associated set-valued mapping $\Psi^{-1}\colon \mathcal{Y} \rightrightarrows \mathcal{X}$ is called the inverse of Ψ.
For any point $(\bar{x},\bar{y}) \in \operatorname{gph}\Psi$, the set-valued mapping $D_N^*\Psi(\bar{x},\bar{y})\colon \mathcal{Y}^\star \rightrightarrows \mathcal{X}^\star$ defined by

$$\forall y^* \in \mathcal{Y}^\star\colon \quad D_N^*\Psi(\bar{x},\bar{y})(y^*) := \{x^* \in \mathcal{X}^\star \,|\, (x^*,-y^*) \in \mathcal{N}_{\operatorname{gph}\Psi}(\bar{x},\bar{y})\}$$

is called normal coderivative of Ψ at $(\bar{x},\bar{y})$. If the continuously Fréchet differentiable mapping $F\colon \mathcal{X} \to \mathcal{Y}$ is interpreted as a set-valued mapping with singleton images, for any $\bar{x} \in \mathcal{X}$, the formula

$$\forall y^* \in \mathcal{Y}^\star\colon \quad D_N^*F(\bar{x},F(\bar{x}))(y^*) = \{F'(\bar{x})^\star[y^*]\}$$

is obtained, see [90, Theorem 1.38].
The set-valued mapping Ψ is called closed at $\bar{x} \in \operatorname{dom}\Psi$ if for any sequence $\{(x_k,y_k)\} \subseteq \operatorname{gph}\Psi$ satisfying $x_k \to \bar{x}$ and $y_k \to \bar{y} \in \mathcal{Y}$, we have $(\bar{x},\bar{y}) \in \operatorname{gph}\Psi$. Furthermore, Ψ is said to be locally bounded at $\bar{x}$ if there are $\delta > 0$ and a bounded set $B \subseteq \mathcal{Y}$ such that $\Psi(x) \subseteq B$ holds for all $x \in \mathbb{U}_{\mathcal{X}}^{\delta}(\bar{x})$. We say that Ψ is locally upper Lipschitzian at $\bar{x}$ if there exist constants $L > 0$ and $\delta > 0$, such that

$$\forall x \in \mathbb{U}_{\mathcal{X}}^{\delta}(\bar{x})\colon \quad \Psi(x) \subseteq \Psi(\bar{x}) + L\,\|x-\bar{x}\|_{\mathcal{X}}\,\mathbb{B}_{\mathcal{Y}}$$

is satisfied. Furthermore, Ψ is called calm at $(\bar{x},\bar{y})$ if there are $L > 0$, $\delta > 0$, and $\varepsilon > 0$ such that the condition

$$\forall x \in \mathbb{U}_{\mathcal{X}}^{\delta}(\bar{x})\colon \quad \Psi(x) \cap \mathbb{U}_{\mathcal{Y}}^{\varepsilon}(\bar{y}) \subseteq \Psi(\bar{x}) + L\,\|x-\bar{x}\|_{\mathcal{X}}\,\mathbb{B}_{\mathcal{Y}}$$

holds. Clearly, if Ψ is locally upper Lipschitz at $\bar{x}$, it is calm at all points $(\bar{x},y)$ where $y \in \Psi(\bar{x})$ is valid. The mapping Ψ is said to be inner semicontinuous at $(\bar{x},\bar{y})$ if for any sequence $\{x_k\} \subseteq \mathcal{X}$ which converges to $\bar{x}$, there exists a sequence $\{y_k\} \subseteq \mathcal{Y}$ converging to $\bar{y}$ which satisfies $y_k \in \Psi(x_k)$ for sufficiently large $k \in \mathbb{N}$. Finally, we call Ψ inner semicompact at $\bar{x}$ if for any sequence $\{x_k\} \subseteq \mathcal{X}$ converging to $\bar{x}$, there exists a sequence $\{y_k\} \subseteq \mathcal{Y}$ which possesses an accumulation point in $\Psi(\bar{x})$ and satisfies $y_k \in \Psi(x_k)$ for sufficiently large $k \in \mathbb{N}$.

2.3.4. Programming and constraint qualifications in Banach spaces

In general, a single-level optimization problem in Banach spaces is of the form

$$\begin{aligned} \psi(x) \;&\to\; \min \\ x \;&\in\; M \end{aligned} \tag{2.9}$$

where $\psi\colon \mathcal{X} \to \overline{\mathbb{R}}$ is called objective function, $\varnothing \neq M \subseteq \mathcal{X}$ is referred to as feasible set, and $\mathcal{X}$ is a Banach space. The following lemma comprises three different necessary optimality conditions for (2.9). The first two are taken from [90, Propositions 5.1 and 5.3], respectively. The proof of the third condition is standard and, thus, omitted.

Lemma 2.29. Let $\bar{x} \in M$ be a local optimal solution of (2.9) such that ψ is finite around $\bar{x}$. Then the following assertions hold:

1. if ψ is Fréchet differentiable at $\bar{x}$, then $-\psi'(\bar{x}) \in \widehat{\mathcal{N}}_M(\bar{x})$ is satisfied,
2. if ψ is locally Lipschitz continuous at $\bar{x}$ and $\mathcal{X}$ is reflexive, then $0 \in \partial\psi(\bar{x}) + \mathcal{N}_M(\bar{x})$ is satisfied,
3. if ψ is directionally differentiable at $\bar{x}$, then $\psi'(\bar{x};\delta) \geq 0$ holds for all $\delta \in \mathcal{R}_M(\bar{x})$.

Example 2.30. Let $M \subseteq \mathcal{H}$ be a nonempty, closed, convex subset of a real Hilbert space $\mathcal{H}$ which is identified with its dual by means of Riesz's representation theorem. Choose $y_\mathsf{d} \in \mathcal{H}$ arbitrarily. Then, combining Example 2.27 as well as Lemma 2.29 and keeping the convexity of the functional $y \mapsto \frac{1}{2}\|y - y_\mathsf{d}\|_{\mathcal{H}}$ in mind, we obtain the following well-known equivalences:

$$\bar{y} = \operatorname{proj}_M(y_\mathsf{d}) \iff y_\mathsf{d} - \bar{y} \in \mathcal{N}_M(\bar{y}) \iff \forall y \in M\colon\ \langle y - \bar{y}, y_\mathsf{d} - \bar{y}\rangle_{\mathcal{H}} \leq 0.$$

We will exploit these different representations of the projection onto a convex set in Section 4.1. ■

Often, the set M equals the preimage $F^{-1}(C)$ of some nonempty, closed set $C \subseteq \mathcal{Y}$ where $\mathcal{Y}$ is a Banach space and $F\colon \mathcal{X} \to \mathcal{Y}$ is a mapping equipped with certain (generalized) differentiability properties. Here we present some important lemmas which help to compute tangent and normal cones to preimages of sets under transformations. This will be necessary in order to state optimality conditions for (2.9) in terms of initial data.

Lemma 2.31. Let $F\colon \mathcal{X} \to \mathcal{Y}$ be a continuously Fréchet differentiable mapping between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, let $C \subseteq \mathcal{Y}$ be a nonempty, closed, convex set, and choose $\bar{x} \in M := F^{-1}(C)$ such that the constraint qualification

$$F'(\bar{x})[\mathcal{X}] - \mathcal{R}_C(F(\bar{x})) = \mathcal{Y} \tag{2.10}$$

is satisfied. Then

$$\mathcal{T}_M^c(\bar{x}) = \mathcal{T}_M^\flat(\bar{x}) = \mathcal{T}_M(\bar{x}) = \mathcal{T}_M^w(\bar{x}) = \{d \in \mathcal{X} \mid F'(\bar{x})[d] \in \mathcal{T}_C(F(\bar{x}))\}$$

is satisfied. If, additionally, $\mathcal{X}$ is reflexive, then we obtain

$$\widehat{\mathcal{N}}_M(\bar{x}) = \mathcal{N}_M^s(\bar{x}) = \mathcal{N}_M(\bar{x}) = \mathcal{N}_M^c(\bar{x}) = F'(\bar{x})^\star[\mathcal{N}_C(F(\bar{x}))].$$

Proof. We exploit [17, Corollary 2.91] in order to obtain

$$\mathcal{T}_M^c(\bar{x}) = \mathcal{T}_M^\flat(\bar{x}) = \mathcal{T}_M(\bar{x}) = \{d \in \mathcal{X} \mid F'(\bar{x})[d] \in \mathcal{T}_C(F(\bar{x}))\}$$

under the postulated constraint qualification. Thus, it is sufficient to show $\mathcal{T}_M^w(\bar{x}) \subseteq \mathcal{T}_M(\bar{x})$ to verify the first statement of the lemma since the inclusion $\mathcal{T}_M(\bar{x}) \subseteq \mathcal{T}_M^w(\bar{x})$ holds by definition of these cones. Hence, choose $d \in \mathcal{T}_M^w(\bar{x}) \setminus \{0\}$ arbitrarily (for $d = 0$, the statement is trivial). Then we find sequences $\{d_k\} \subseteq \mathcal{X}$ and $\{t_k\} \subseteq \mathbb{R}$ such that $d_k \rightharpoonup d$ as well as $t_k \searrow 0$ are satisfied and $F(\bar{x} + t_k d_k) \in C$ holds for any $k \in \mathbb{N}$. The convexity of C yields $\xi_k := \frac{1}{t_k}(F(\bar{x} + t_k d_k) - F(\bar{x})) \in \mathcal{R}_C(F(\bar{x}))$ for any $k \in \mathbb{N}$. Now, choose $y^* \in \mathcal{Y}^\star$ arbitrarily. Then we have

$$\begin{aligned}
\langle y^*, F'(\bar{x})[d]\rangle_{\mathcal{Y}} &= \lim_{k\to\infty} \langle y^*, F'(\bar{x})[d_k]\rangle_{\mathcal{Y}} \\
&= \lim_{k\to\infty} \left\langle y^*, \frac{F(\bar{x} + t_k d_k) - F(\bar{x}) - \big(F(\bar{x} + t_k d_k) - F(\bar{x}) - F'(\bar{x})[t_k d_k]\big)}{t_k} \right\rangle_{\mathcal{Y}} \\
&= \lim_{k\to\infty} \left(\langle y^*, \xi_k\rangle_{\mathcal{Y}} - \left\langle y^*, \|d_k\|_{\mathcal{X}} \frac{F(\bar{x} + t_k d_k) - F(\bar{x}) - F'(\bar{x})[t_k d_k]}{\|t_k d_k\|_{\mathcal{X}}} \right\rangle_{\mathcal{Y}} \right).
\end{aligned} \tag{2.11}$$

Due to its weak convergence, $\{d_k\}$ is bounded and, thus, $\{t_k d_k\}$ converges to zero. The definition of Fréchet differentiability yields

$$\begin{aligned}
0 &\leq \left| \lim_{k\to\infty} \left\langle y^*, \|d_k\|_{\mathcal{X}} \frac{F(\bar{x} + t_k d_k) - F(\bar{x}) - F'(\bar{x})[t_k d_k]}{\|t_k d_k\|_{\mathcal{X}}} \right\rangle_{\mathcal{Y}} \right| \\
&\leq \lim_{k\to\infty} \|y^*\|_{\mathcal{Y}^\star} \|d_k\|_{\mathcal{X}} \frac{\|F(\bar{x} + t_k d_k) - F(\bar{x}) - F'(\bar{x})[t_k d_k]\|_{\mathcal{Y}}}{\|t_k d_k\|_{\mathcal{X}}} = 0.
\end{aligned}$$

Combining this with (2.11),

$$\lim_{k\to\infty} \langle y^*, \xi_k\rangle_{\mathcal{Y}} = \langle y^*, F'(\bar{x})[d]\rangle_{\mathcal{Y}}$$

is obtained for all $y^* \in \mathcal{Y}^\star$, i.e. $\xi_k \rightharpoonup F'(\bar{x})[d]$. Now, we can deduce

$$F'(\bar{x})[d] \in \overline{\operatorname{conv}}\{\xi_k \,|\, k \in \mathbb{N}\} \subseteq \overline{\operatorname{conv}}\, \mathcal{R}_C(F(\bar{x})) = \operatorname{cl} \mathcal{R}_C(F(\bar{x})) = \mathcal{T}_C(F(\bar{x}))$$

which shows the first statement of the lemma.
Due to the reflexivity of $\mathcal{X}$ and the first part of the proof, we find

$$\widehat{\mathcal{N}}_M(\bar{x}) \subseteq \mathcal{N}_M^s(\bar{x}) \subseteq \mathcal{N}_M(\bar{x}) \subseteq \mathcal{N}_M^c(\bar{x}) = \mathcal{T}_M^c(\bar{x})^\circ = \mathcal{T}_M^w(\bar{x})^\circ = \widehat{\mathcal{N}}_M(\bar{x}),$$

i.e. all these cones coincide. The fact that these cones equal $F'(\bar{x})^\star[\mathcal{N}_C(F(\bar{x}))]$ follows from the second statement of Lemma 2.13 and the above representation of the tangent cones, observing that the postulated constraint qualification (2.10) implies

$$F'(\bar{x})[\mathcal{X}] - \mathcal{T}_C(F(\bar{x})) = \mathcal{Y}. \tag{2.12}$$

This completes the proof. □

Lemma 2.32. Let all assumptions apart from the reflexivity of $\mathcal{X}$ stated in Lemma 2.31 be satisfied. Additionally, suppose that $\bar{x} \in M$ is a local optimal solution of (2.9) such that ψ is Fréchet differentiable at this point. Then there is a so-called (regular) Lagrange multiplier $\lambda \in \mathcal{N}_C(F(\bar{x}))$ which satisfies the relation $0 = \psi'(\bar{x}) + F'(\bar{x})^\star[\lambda]$.

Proof. From Lemma 2.29 $-\psi'(\bar{x}) \in \widehat{\mathcal{N}}_M(\bar{x})$ is satisfied, and this yields $-\psi'(\bar{x}) \in \mathcal{N}_M^c(\bar{x})$. Due to the validity of the constraint qualification (2.10), we obtain

$$\mathcal{N}_M^c(\bar{x}) = \mathcal{T}_M^c(\bar{x})^\circ = \{d \in \mathcal{X} \,|\, F'(\bar{x})[d] \in \mathcal{T}_C(F(\bar{x}))\}^\circ = F'(\bar{x})^\star[\mathcal{T}_C(F(\bar{x}))^\circ]$$

from Lemmas 2.13 and 2.31. Finally, the convexity of C leads to $\mathcal{T}_C(F(\bar{x}))^\circ = \mathcal{N}_C(F(\bar{x}))$ which completes the proof. □

Remark 2.33. Again, let us postulate that $F\colon \mathcal{X} \to \mathcal{Y}$ is a continuously Fréchet differentiable mapping between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ while $C \subseteq \mathcal{Y}$ is a nonempty, closed, convex set and $\bar{x} \in F^{-1}(C)$ is fixed. The constraint qualification (2.10) was introduced by Robinson, see [105], in order to study stability properties of the solution set to nonlinear inequality systems in Banach spaces. Later, Kurcyusz and Zowe exploited the same condition to show the existence of Lagrange multipliers at local optimal solutions of differentiable programming problems in Banach spaces, see [143]. That is why we call (2.10) Kurcyusz Robinson Zowe constraint qualification, KRZCQ for brevity, throughout this thesis. Note that KRZCQ implies the condition (2.12) which plays an important role in the generalized Farkas lemma, see Lemma 2.13. Polarizing (2.12), we easily see that this condition implies

$$0 = F'(\bar{x})^\star[\lambda],\ \lambda \in \mathcal{N}_C(F(\bar{x})) \implies \lambda = 0 \tag{2.13}$$

which may be interpreted in the way that there do not exist nontrivial singular Lagrange multipliers at $\bar{x} \in F^{-1}(C)$ for optimization problems possessing the feasible set $F^{-1}(C)$. That is why it is often called NNAMCQ, no nonzero abnormal multiplier constraint qualification, in the literature. Exploiting Lemma 2.11, the latter condition is equivalent to

$$\operatorname{cl}\big(F'(\bar{x})[\mathcal{X}] - \mathcal{T}_C(F(\bar{x}))\big) = \mathcal{Y} \tag{2.14}$$

and, thus (see Lemma 2.1), to

$$\operatorname{cl}\big(F'(\bar{x})[\mathcal{X}] - \mathcal{R}_C(F(\bar{x}))\big) = \mathcal{Y}. \tag{2.15}$$

Consequently, we have

$$(2.10) \implies (2.12) \implies (2.13) \iff (2.14) \iff (2.15),$$

see [17, Proposition 2.97]. If $\mathcal{Y}$ is finite-dimensional, all these conditions are equivalent. On the other hand, if the set C possesses a nonempty interior (which is often too restrictive in the context of programming in Banach spaces), then all these conditions are equivalent and coincide with

$$\exists d \in \mathcal{X}\colon \quad F(\bar{x}) + F'(\bar{x})[d] \in \operatorname{int} C.$$

This constraint qualification may be seen as a generalized version of MFCQ. Clearly, for common finite-dimensional programming problems with inequality constraints, KRZCQ and MFCQ are equivalent. Note that the conditions (2.10), (2.12), (2.14), and (2.15) are postulated in the space $\mathcal{Y}$ which is why they are called primal constraint qualifications. On the other hand, (2.13) is a dual constraint qualification since it states a condition in $\mathcal{Y}^\star$. More information on KRZCQ and other representations of this constraint qualification in product spaces can be found in [17, Section 2.3.4] and [31].

Remark 2.34. The necessary optimality conditions presented in Lemma 2.32 are called the KKT conditions of (2.9) at $\bar{x}$ as long as C is a convex set. In the setting of programming in Banach spaces, these conditions were stated under validity of KRZCQ in [143] first. In the case where C is a nonempty, closed, convex cone, we obtain the more intuitive condition $\lambda \in C^\circ \cap \{F(\bar{x})\}^\perp$ for the Lagrange multiplier. Note that whenever $F'(\bar{x})$ is surjective, then KRZCQ holds at $\bar{x}$ and the Lagrange multiplier is unique. As it is shown in [121, Section 4], the uniqueness of the Lagrange multiplier also follows from the condition

$$\operatorname{cl}\left(F'(\bar{x})[\mathcal{X}] - \mathcal{N}_C(F(\bar{x}))_\perp\right) = \mathcal{Y} \tag{2.16}$$

which is obviously weaker than the surjectivity of $F'(\bar{x})$. However, since (2.16) does not imply KRZCQ or vice versa whenever $\mathcal{Y}$ is infinite-dimensional, it does not necessarily imply the existence of a Lagrange multiplier satisfying the KKT conditions. Clearly, from the inclusion

$$\mathcal{N}_C(F(\bar{x}))_\perp = \mathcal{T}_C(F(\bar{x})) \cap \left(-\mathcal{T}_C(F(\bar{x}))\right) \subseteq \mathcal{T}_C(F(\bar{x}))$$

condition (2.16) implies (2.14) which is equivalent to KRZCQ as long as $\mathcal{Y}$ is finite-dimensional, see Remark 2.33. Especially, in the case $\mathcal{Y} = \mathbb{R}^m$ and $C = -\mathbb{R}_0^{m,+}$, it is easily seen that (2.16) is equivalent to the linear independence constraint qualification, LICQ for short, which simply says that the vectors $\{F_i'(\bar{x}) \in \mathcal{X}^\star \,|\, i \in I(\bar{x})\}$ are linear independent. Here $F_1, \ldots, F_m \colon \mathcal{X} \to \mathbb{R}$ denote the component mappings of F and $I(\bar{x}) = \{i \in \{1, \ldots, m\} \,|\, F_i(\bar{x}) = 0\}$ is the set of active constraints.
Assume that at $\bar{x} \in M$, the KKT conditions hold with the multiplier $\lambda \in \mathcal{Y}^\star$. The so-called strict condition of Kurcyusz, Robinson, and Zowe, SKRZC for brevity, is said to hold at $(\bar{x}, \lambda)$ if

$$F'(\bar{x})[\mathcal{X}] - \mathcal{R}_{C(\bar{x},\lambda)}(F(\bar{x})) = \mathcal{Y} \tag{2.17}$$

is satisfied where $C(\bar{x}, \lambda) := \{y \in C \,|\, \langle \lambda, y - F(\bar{x})\rangle_{\mathcal{Y}} = 0\}$ is valid. Since this condition does already depend on a fixed Lagrange multiplier and, thus, on the objective ψ, we do not call it a constraint qualification for (2.9). However, it obviously implies KRZCQ and the uniqueness of the Lagrange multiplier, see [17, Proposition 4.47]. On the other hand, SKRZC is not related to the constraint qualification (2.16) in general. Note that whenever C is a cone, then SKRZC is equivalent to

$$F'(\bar{x})[\mathcal{X}] - \mathcal{R}_C(F(\bar{x})) \cap \{\lambda\}_\perp = \mathcal{Y}.$$

Especially, for common programming problems in $\mathbb{R}^n$ with inequality constraints, SKRZC equals SMFC, the strict Mangasarian Fromovitz condition, see [80].

Recall that under certain convexity assumptions, the KKT conditions of (2.9) provide a sufficient criterion for optimality.

Lemma 2.35. Let ψ be a convex functional, $F \colon \mathcal{X} \to \mathcal{Y}$ be a continuously Fréchet differentiable mapping between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, let $C \subseteq \mathcal{Y}$ be a nonempty, closed, convex cone, and assume that F is $-C$-convex, i.e. it satisfies

$$\forall x, x' \in \mathcal{X}\, \forall \alpha \in [0,1] \colon \quad F(\alpha x + (1-\alpha)x') - \alpha F(x) - (1-\alpha)F(x') \in C,$$

see [72, Definition 2.4]. Choose $\bar{x} \in M := F^{-1}(C)$ where ψ is Fréchet differentiable and assume that the KKT conditions hold at $\bar{x}$. Then $\bar{x}$ is a global optimal solution of the corresponding optimization problem (2.9).

Proof. Since the KKT conditions hold at $\bar{x}$ and C is a convex cone, we find a multiplier $\lambda \in C^\circ \cap \{F(\bar{x})\}^\perp$ such that $0 = \psi'(\bar{x}) + F'(\bar{x})^\star[\lambda]$ is satisfied. Suppose that there is some $x \in F^{-1}(C)$ such that $\psi(x) < \psi(\bar{x})$,

i.e. $\bar{x}$ is not globally optimal for (2.9). Since F is $-C$ convex, we obtain $F(\bar{x}) + F'(\bar{x})[x - \bar{x}] - F(x) \in C$, see [72, Theorem 2.20]. Then the convexity of ψ, $\langle \lambda, F(x) \rangle_{\mathcal{Y}} \leq 0$, and $\langle \lambda, F(\bar{x}) \rangle_{\mathcal{Y}} = 0$ yield

$$\begin{aligned} 0 &= \langle \psi'(\bar{x}) + F'(\bar{x})^\star[\lambda], x - \bar{x} \rangle_{\mathcal{X}} = \psi'(\bar{x})[x - \bar{x}] + \langle \lambda, F'(\bar{x})[x - \bar{x}] \rangle_{\mathcal{Y}} \\ &\leq \psi'(\bar{x})[x - \bar{x}] + \langle \lambda, F(\bar{x}) + F'(\bar{x})[x - \bar{x}] - F(x) \rangle_{\mathcal{Y}} \\ &\leq \psi'(\bar{x})[x - \bar{x}] \leq \psi(x) - \psi(\bar{x}) < 0 \end{aligned}$$

which is a contradiction. Hence, $\bar{x}$ is a global optimal solution of (2.9). □

In the case of a standard nonlinear program with inequality constraints, i.e. $\mathcal{Y} = \mathbb{R}^m$ and $C = -\mathbb{R}^{m,+}_0$, the mapping F is $\mathbb{R}^{m,+}_0$-convex if and only if its m component mappings $F_1, \ldots, F_m \colon \mathcal{X} \to \mathbb{R}$ are convex. The fact that the KKT conditions are sufficient for optimality in standard nonlinear convex programming is well-known, see [106, Section 28].
Later, it will be necessary to apply the primal and dual constraint qualifications introduced above to constraint systems in product spaces. In order to simplify these conditions, we will exploit some cancellation rules stated in the subsequent lemma.

Lemma 2.36. For Banach spaces $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_1$, and $\mathcal{Y}_2$, linear operators $\mathtt{F} \in \mathbb{L}[\mathcal{X}_1, \mathcal{Y}_1]$, $\mathtt{G} \in \mathbb{L}[\mathcal{X}_2, \mathcal{Y}_1]$, $\mathtt{U}, \mathtt{U}' \in \mathbb{L}[\mathcal{X}_1, \mathcal{Y}_2]$, and $\mathtt{V}, \mathtt{V}' \in \mathbb{L}[\mathcal{X}_2, \mathcal{Y}_2]$, as well as nonempty sets $X_1 \subseteq \mathcal{X}_1$, $S_i \subseteq \mathcal{Y}_i$, $i = 1, 2$, and $T \subseteq \mathcal{X}_2$, we consider the following conditions:

$$\begin{bmatrix} \mathtt{F} & \mathtt{G} \\ \mathtt{U} & \mathtt{V} \\ \mathtt{U}' & \mathtt{V}' \\ \mathtt{0} & \mathtt{I}_{\mathcal{X}_2} \end{bmatrix} \begin{pmatrix} X_1 \\ \mathcal{X}_2 \end{pmatrix} - \begin{pmatrix} S_1 \\ \{0\} \\ S_2 \\ T \end{pmatrix} = \begin{pmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \\ \mathcal{Y}_2 \\ \mathcal{X}_2 \end{pmatrix}, \tag{2.18a}$$

$$\begin{bmatrix} \mathtt{F} \\ \mathtt{U} \\ \mathtt{U}' - \mathtt{U} \end{bmatrix} [X_1] - \begin{bmatrix} \mathtt{I}_{\mathcal{Y}_1} & \mathtt{0} & -\mathtt{G} \\ \mathtt{0} & \mathtt{0} & -\mathtt{V} \\ \mathtt{0} & \mathtt{I}_{\mathcal{Y}_2} & \mathtt{V} - \mathtt{V}' \end{bmatrix} \begin{pmatrix} S_1 \\ S_2 \\ T \end{pmatrix} = \begin{pmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \\ \mathcal{Y}_2 \end{pmatrix}, \tag{2.18b}$$

$$\operatorname{cl}\left(\begin{bmatrix} \mathtt{F} & \mathtt{G} \\ \mathtt{U} & \mathtt{V} \\ \mathtt{U}' & \mathtt{V}' \\ \mathtt{0} & \mathtt{I}_{\mathcal{X}_2} \end{bmatrix} \begin{pmatrix} X_1 \\ \mathcal{X}_2 \end{pmatrix} - \begin{pmatrix} S_1 \\ \{0\} \\ S_2 \\ T \end{pmatrix} \right) = \begin{pmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \\ \mathcal{Y}_2 \\ \mathcal{X}_2 \end{pmatrix}, \tag{2.18c}$$

$$\operatorname{cl}\left(\begin{bmatrix} \mathtt{F} \\ \mathtt{U} \\ \mathtt{U}' - \mathtt{U} \end{bmatrix} [X_1] - \begin{bmatrix} \mathtt{I}_{\mathcal{Y}_1} & \mathtt{0} & -\mathtt{G} \\ \mathtt{0} & \mathtt{0} & -\mathtt{V} \\ \mathtt{0} & \mathtt{I}_{\mathcal{Y}_2} & \mathtt{V} - \mathtt{V}' \end{bmatrix} \begin{pmatrix} S_1 \\ S_2 \\ T \end{pmatrix} \right) = \begin{pmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \\ \mathcal{Y}_2 \end{pmatrix}. \tag{2.18d}$$

Then (2.18a) and (2.18b) are equivalent, whereas (2.18c) and (2.18d) are equivalent as well.

Proof. We only provide a proof for the equivalence of (2.18c) and (2.18d). The validation of the lemma's first statement follows from a similar (but easier) argumentation.
Let (2.18c) hold and choose $y_1 \in \mathcal{Y}_1$ and $y_2, y_2' \in \mathcal{Y}_2$ arbitrarily. Fix some $\varepsilon > 0$. Since 0 belongs to $\mathcal{X}_2$, we find $x_1^\varepsilon \in X_1$, $x_2^\varepsilon \in \mathcal{X}_2$, $s_i^\varepsilon \in S_i$, $i = 1, 2$, and $t^\varepsilon \in T$ such that

$$\|\mathtt{F}[x_1^\varepsilon] + \mathtt{G}[x_2^\varepsilon] - s_1^\varepsilon - y_1\|_{\mathcal{Y}_1} \leq \varepsilon, \tag{2.19a}$$

$$\|\mathtt{U}[x_1^\varepsilon] + \mathtt{V}[x_2^\varepsilon] - y_2\|_{\mathcal{Y}_2} \leq \varepsilon, \tag{2.19b}$$

$$\|\mathtt{U}'[x_1^\varepsilon] + \mathtt{V}'[x_2^\varepsilon] - s_2^\varepsilon - (y_2' + y_2)\|_{\mathcal{Y}_2} \leq \varepsilon, \tag{2.19c}$$

$$\|x_2^\varepsilon - t^\varepsilon\|_{\mathcal{X}_2} \leq \varepsilon. \tag{2.19d}$$

Using the triangle inequality, we find

$$\begin{aligned} \|\mathtt{F}[x_1^\varepsilon] - s_1^\varepsilon + \mathtt{G}[t^\varepsilon] - y_1\|_{\mathcal{Y}_1} &\leq \|\mathtt{F}[x_1^\varepsilon] + \mathtt{G}[x_2^\varepsilon] - s_1^\varepsilon - y_1\|_{\mathcal{Y}_1} + \|\mathtt{G}[t^\varepsilon - x_2^\varepsilon]\|_{\mathcal{Y}_1} \\ &\leq \|\mathtt{F}[x_1^\varepsilon] + \mathtt{G}[x_2^\varepsilon] - s_1^\varepsilon - y_1\|_{\mathcal{Y}_1} + \|\mathtt{G}\|_{\mathbb{L}[\mathcal{X}_2, \mathcal{Y}_1]} \|t^\varepsilon - x_2^\varepsilon\|_{\mathcal{X}_2} \\ &\leq \varepsilon \left(1 + \|\mathtt{G}\|_{\mathbb{L}[\mathcal{X}_2, \mathcal{Y}_1]}\right) \end{aligned}$$

and, similarly,

$$\|\mathtt{U}[x_1^\varepsilon] + \mathtt{V}[t^\varepsilon] - y_2\|_{\mathcal{Y}_2} \leq \varepsilon \left(1 + \|\mathtt{V}\|_{\mathbb{L}[\mathcal{X}_2,\mathcal{Y}_2]}\right)$$

is derived. The same trick yields

$$\begin{aligned}
&\|(\mathtt{U}' - \mathtt{U})[x_1^\varepsilon] - s_2^\varepsilon + (\mathtt{V}' - \mathtt{V})[t^\varepsilon] - y_2'\|_{\mathcal{Y}_2} \\
&\quad \leq \|\mathtt{U}'[x_1^\varepsilon] + \mathtt{V}'[t^\varepsilon] - s_2^\varepsilon - (y_2' + y_2)\|_{\mathcal{Y}_2} + \|y_2 - \mathtt{U}[x_1^\varepsilon] - \mathtt{V}[t^\varepsilon]\|_{\mathcal{Y}_2} \\
&\quad \leq \varepsilon \left(1 + \|\mathtt{V}'\|_{\mathbb{L}[\mathcal{X}_2,\mathcal{Y}_2]}\right) + \varepsilon \left(1 + \|\mathtt{V}\|_{\mathbb{L}[\mathcal{X}_2,\mathcal{Y}_2]}\right) \\
&\quad = \varepsilon \left(2 + \|\mathtt{V}'\|_{\mathbb{L}[\mathcal{X}_2,\mathcal{Y}_2]} + \|\mathtt{V}\|_{\mathbb{L}[\mathcal{X}_2,\mathcal{Y}_2]}\right).
\end{aligned}$$

Taking the limit $\varepsilon \searrow 0$, we see that (2.18d) is valid.
Now, assume that (2.18d) holds and choose $\tilde{y}_1 \in \mathcal{Y}$, $\tilde{y}_2, \tilde{y}_2' \in \mathcal{Y}_2$, and $\tilde{x} \in \mathcal{X}_2$ arbitrarily. Furthermore, fix $\varepsilon > 0$. Then there are $\tilde{x}_1^\varepsilon \in X_1$, $\tilde{s}_i^\varepsilon \in S_i$, $i = 1, 2$, and $\tilde{t}^\varepsilon \in T$ such that the estimates

$$\begin{aligned}
\left\|\mathtt{F}[\tilde{x}_1^\varepsilon] - \tilde{s}_1^\varepsilon + \mathtt{G}[\tilde{t}^\varepsilon] - (\tilde{y}_1 - \mathtt{G}[\tilde{x}])\right\|_{\mathcal{Y}_1} &\leq \varepsilon, \\
\left\|\mathtt{U}[\tilde{x}_1^\varepsilon] + \mathtt{V}[\tilde{t}^\varepsilon] - (\tilde{y}_2 - \mathtt{V}[\tilde{x}])\right\|_{\mathcal{Y}_2} &\leq \varepsilon, \\
\left\|(\mathtt{U}' - \mathtt{U})[\tilde{x}_1^\varepsilon] - \tilde{s}_2^\varepsilon + (\mathtt{V}' - \mathtt{V})[\tilde{t}^\varepsilon] - (\tilde{y}_2' - \tilde{y}_2 - (\mathtt{V}' - \mathtt{V})[\tilde{x}])\right\|_{\mathcal{Y}_2} &\leq \varepsilon
\end{aligned}$$

are valid. Define $\tilde{x}_2^\varepsilon := \tilde{x} + \tilde{t}^\varepsilon$. Then we have

$$\begin{aligned}
\|\mathtt{F}[\tilde{x}_1^\varepsilon] + \mathtt{G}[\tilde{x}_2^\varepsilon] - \tilde{s}_1^\varepsilon - \tilde{y}_1\|_{\mathcal{Y}_1} &\leq \varepsilon, \\
\|\mathtt{U}[\tilde{x}_1^\varepsilon] + \mathtt{V}[\tilde{x}_2^\varepsilon] - \tilde{y}_2\|_{\mathcal{Y}_2} &\leq \varepsilon, \\
\left\|\tilde{x}_2^\varepsilon - \tilde{t}^\varepsilon - \tilde{x}\right\|_{\mathcal{X}_2} &= 0.
\end{aligned}$$

Moreover, the triangle inequality yields

$$\begin{aligned}
&\|\mathtt{U}'[\tilde{x}_1^\varepsilon] + \mathtt{V}'[\tilde{x}_2^\varepsilon] - \tilde{s}_2^\varepsilon - \tilde{y}_2'\|_{\mathcal{Y}_2} \\
&\quad \leq \left\|(\mathtt{U}' - \mathtt{U})[\tilde{x}_1^\varepsilon] - \tilde{s}_2^\varepsilon + (\mathtt{V}' - \mathtt{V})[\tilde{t}^\varepsilon] - (\tilde{y}_2' - \tilde{y}_2 - (\mathtt{V}' - \mathtt{V})[\tilde{x}])\right\|_{\mathcal{Y}_2} + \left\|\mathtt{U}[\tilde{x}_1^\varepsilon] + \mathtt{V}[\tilde{t}^\varepsilon] - (\tilde{y}_2 - \mathtt{V}[\tilde{x}])\right\|_{\mathcal{Y}_2} \\
&\quad \leq 2\varepsilon.
\end{aligned}$$

Thus, taking the limit $\varepsilon \searrow 0$ shows that (2.18c) is satisfied. □

Corollary 2.37. For Banach spaces $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{Y}$, linear operators $\mathtt{F} \in \mathbb{L}[\mathcal{X}_1, \mathcal{Y}]$ and $\mathtt{G} \in \mathbb{L}[\mathcal{X}_2, \mathcal{Y}]$, as well as nonempty sets $X_1 \subseteq \mathcal{X}_1$, $S \subseteq \mathcal{Y}$, and $T \subseteq \mathcal{X}_2$, we consider the following conditions:

$$\begin{bmatrix} \mathtt{F} & \mathtt{G} \\ \mathtt{O} & \mathtt{I}_{\mathcal{X}_2} \end{bmatrix} \begin{pmatrix} X_1 \\ \mathcal{X}_2 \end{pmatrix} - \begin{pmatrix} S \\ T \end{pmatrix} = \begin{pmatrix} \mathcal{Y} \\ \mathcal{X}_2 \end{pmatrix}, \tag{2.20a}$$

$$\mathtt{F}[X_1] + \mathtt{G}[T] - S = \mathcal{Y}, \tag{2.20b}$$

$$\operatorname{cl}\left(\begin{bmatrix} \mathtt{F} & \mathtt{G} \\ \mathtt{O} & \mathtt{I}_{\mathcal{X}_2} \end{bmatrix} \begin{pmatrix} X_1 \\ \mathcal{X}_2 \end{pmatrix} - \begin{pmatrix} S \\ T \end{pmatrix}\right) = \begin{pmatrix} \mathcal{Y} \\ \mathcal{X}_2 \end{pmatrix}, \tag{2.20c}$$

$$\operatorname{cl}\left(\mathtt{F}[X_1] + \mathtt{G}[T] - S\right) = \mathcal{Y}. \tag{2.20d}$$

Then (2.20a) and (2.20b) are equivalent, whereas (2.20c) and (2.20d) are equivalent as well.

The results in Lemma 2.31 heavily rely on the convexity of C. The following lemma which combines [90, Corollary 1.15, Theorem 1.17, Theorem 3.8] shows some calculus rules for normals to inverse images of not necessarily convex sets. Related results are presented in [6, Section 4] and [108, Section 6] as well.

Lemma 2.38. Let $F\colon \mathcal{X} \to \mathcal{Y}$ be a continuously Fréchet differentiable mapping between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, let $C \subseteq \mathcal{Y}$ be a nonempty, closed set, and choose $\bar{x} \in M := F^{-1}(C)$. If $F'(\bar{x})$ is a surjective operator, then

$$\widehat{\mathcal{N}}_M(\bar{x}) = F'(\bar{x})^\star\left[\widehat{\mathcal{N}}_C(F(\bar{x}))\right], \quad \mathcal{N}_M(\bar{x}) = F'(\bar{x})^\star[\mathcal{N}_C(F(\bar{x}))]$$

hold. On the other hand, if $\mathcal{X}$ and $\mathcal{Y}$ are reflexive, C is SNC at $F(\bar{x})$, and the constraint qualification (2.13) is satisfied, then the following upper approximation for the limiting normal cone is valid:

$$\mathcal{N}_M(\bar{x}) \subseteq F'(\bar{x})^\star[\mathcal{N}_C(F(\bar{x}))].$$

Combining Lemmas 2.29 and 2.38, we obtain necessary optimality conditions again: Supposing that $\bar{x} \in \mathcal{X}$ is a local optimal solution of (2.9) where ψ is Fréchet differentiable and the feasible set M is given as presented in the above lemma, we have

$$\exists \lambda \in \widehat{\mathcal{N}}_C(F(\bar{x})): \quad 0 = \psi'(\bar{x}) + F'(\bar{x})^\star[\lambda]$$

provided $F'(\bar{x})$ is surjective. Often, this assumption is too restrictive. If it is weakened to (2.13) in the setting of reflexive Banach spaces, then the weaker necessary optimality condition

$$\exists \lambda \in \mathcal{N}_C(F(\bar{x})): \quad 0 = \psi'(\bar{x}) + F'(\bar{x})^\star[\lambda]$$

holds provided C is SNC at $F(\bar{x})$. Clearly, both of these conditions equal the KKT conditions for a convex set C.

An important instance of (2.9) are problems with finitely many equality and inequality constraints.

Lemma 2.39. Let $\varphi_1, \ldots, \varphi_l \colon \mathcal{X} \to \mathbb{R}$ be locally Lipschitz continuous functionals of a reflexive Banach space $\mathcal{X}$ and let $S \subseteq \mathcal{X}$ be nonempty and closed. For $k \in \{0, \ldots, l\}$, we consider the set

$$M := \left\{ x \in S \;\middle|\; \begin{array}{ll} \varphi_i(x) \leq 0 & \text{for } i = 1, \ldots, k \\ \varphi_i(x) = 0 & \text{for } i = k+1, \ldots, l \end{array} \right\}$$

at some arbitrary point $\bar{x} \in M$. Define $I(\bar{x}) := \{i \in \{1, \ldots, k\} \mid \varphi_i(\bar{x}) = 0\}$ and suppose that the constraint qualification

$$\left. \begin{array}{l} 0 \in \displaystyle\sum_{i \in I(\bar{x})} \theta_i \partial \varphi_i(\bar{x}) + \sum_{i=k+1}^{l} \theta_i \partial^c \varphi_i(\bar{x}) + \mathcal{N}_S(\bar{x}), \\ \forall i \in I(\bar{x}): \; \theta_i \geq 0 \end{array} \right\} \implies \forall i \in I(\bar{x}) \cup \{k+1, \ldots, l\}: \; \theta_i = 0. \tag{2.21}$$

is satisfied. Then we have the following upper estimate for the limiting normal cone to M at $\bar{x}$:

$$\begin{aligned} \mathcal{N}_M(\bar{x}) \subseteq & \left\{ \sum_{i \in I(\bar{x})} \theta_i x_i^* \;\middle|\; \forall i \in I(\bar{x}): \; \theta_i \geq 0, \, x_i^* \in \partial \varphi_i(\bar{x}) \right\} \\ & + \left\{ \sum_{i=k+1}^{l} \theta_i x_i^* \;\middle|\; \forall i \in \{k+1, \ldots, l\}: \; \theta_i \in \mathbb{R}, \, x_i^* \in \partial^c \varphi_i(\bar{x}) \right\} + \mathcal{N}_S(\bar{x}). \end{aligned}$$

Proof. Let $F \colon \mathcal{X} \to \mathbb{R}^l$ be the mapping possessing the components $\varphi_1, \ldots, \varphi_l$. Clearly, F is locally Lipschitz continuous. Moreover, let us introduce the convex set $C := (-\mathbb{R}_0^{k,+}) \times \{0\} \subseteq \mathbb{R}^l$ and the preimage $\widetilde{M} := F^{-1}(C)$. Observe that we have $M = \widetilde{M} \cap S$ by construction.
We apply the scalarization property of the limiting subdifferential, see [90, Theorem 3.28], and the sum rule for locally Lipschitzian functionals, see [90, Theorem 3.36] in order to obtain

$$D_N^* F(\bar{x}, F(\bar{x}))(\vartheta) = \partial \left(\sum_{i=1}^{l} \vartheta_i \varphi_i \right)(\bar{x}) \subseteq \sum_{i=1}^{l} \partial(\vartheta_i \varphi_i)(\bar{x})$$

for any $\vartheta \in \mathbb{R}^l$. Choosing $\theta \in \mathcal{N}_C(F(\bar{x}))$, we easily see $\theta_i = 0$ for all $i \in \{1, \ldots, k\} \setminus I(\bar{x})$ and $\theta_i \geq 0$ for all $i \in I(\bar{x})$ which yields

$$D_N^* F(\bar{x}, F(\bar{x}))(\theta) \subseteq \sum_{i \in I(\bar{x})} \theta_i \partial \varphi_i(\bar{x}) + \sum_{i=k+1}^{l} \partial^c(\theta_i \varphi_i)(\bar{x}) = \sum_{i \in I(\bar{x})} \theta_i \partial \varphi_i(\bar{x}) + \sum_{i=k+1}^{l} \theta_i \partial^c \varphi_i(\bar{x}).$$

Thus, the postulated constraint qualification implies

$$\mathcal{N}_C(F(\bar{x})) \cap \ker D_N^* F(\bar{x}, F(\bar{x})) = \{0\}.$$

Observing that C is SNC at $F(\bar{x})$ as a subset of a finite-dimensional Banach space, $\widetilde{M}$ is SNC at $\bar{x}$ by [90, Corollary 1.69, Theorem 3.84] and satisfies

$$\mathcal{N}_{\widetilde{M}}(\bar{x}) \subseteq \left\{ \sum_{i\in I(\bar{x})} \theta_i x_i^* \,\middle|\, \forall i \in I(\bar{x})\colon\ \theta_i \geq 0,\ x_i^* \in \partial\varphi_i(\bar{x}) \right\} + \left\{ \sum_{i=k+1}^{l} \theta_i x_i^* \,\middle|\, \forall i \in \{k+1,\dots,l\}\colon\ \theta_i \in \mathbb{R},\ x_i^* \in \partial^c\varphi_i(\bar{x}) \right\},$$

see [90, Theorem 3.8]. Thus, the postulated constraint qualification implies $\mathcal{N}_{\widetilde{M}}(\bar{x}) \cap \big(-\mathcal{N}_S(\bar{x})\big) = \{0\}$. Recalling that $\widetilde{M}$ is SNC at $\bar{x}$, the statement of this lemma is a direct consequence of Lemma 2.18. □

Finally, we present necessary optimality conditions for problem (2.9) whose feasible set is given as presented in Lemma 2.39.

Lemma 2.40. Let $\mathcal{X}$ be a reflexive Banach space, let the feasible set M of (2.9) be given as described in Lemma 2.39, assume that $\bar{x} \in M$ is a local optimal solution of (2.9), and let ψ be locally Lipschitz continuous at $\bar{x}$. Then there exist scalars $\theta_0 \geq 0$, $\theta_i \geq 0$ for $i \in I(\bar{x})$, and $\theta_i \in \mathbb{R}$ for $i = k+1,\dots,l$ not all equal to zero at the same time such that

$$0 \in \theta_0\partial\psi(\bar{x}) + \sum_{i\in I(\bar{x})} \theta_i\partial\varphi_i(\bar{x}) + \sum_{i=k+1}^{l} \theta_i\partial^c\varphi(\bar{x}) + \mathcal{N}_S(\bar{x})$$

is satisfied. If the constraint qualification (2.21) is satisfied, then we can choose $\theta_0 = 1$.

Proof. The necessity of the Fritz-John-type optimality condition was validated in [90, Theorem 5.21(iii)]. If (2.21) is satisfied, then $\theta_0 = 0$ would imply that all all the other scalar multipliers are zero as well. This, however, would be a contradiction to the fact that not all these scalars vanish at the same time. Thus, we have $\theta_0 > 0$ in this case and a scalarization yields the claim. □

2.3.5. Variational geometry of decomposable sets in Lebesgue spaces

Frequently, in optimal control, the set of feasible controls is given by the pointwise defined set

$$U_{\mathsf{ad}} = \{u \in L^p(\Omega,\mathbb{R}^m) \,|\, u(\omega) \in U(\omega) \text{ f.a.a. } \omega \in \Omega\}$$

where $p \in (1,\infty)$ holds, $\Omega \subseteq \mathbb{R}^d$ is a bounded domain, and the set-valued mapping $U\colon \Omega \rightrightarrows \mathbb{R}^m$ has nonempty and closed images. In order to apply the necessary optimality conditions from Section 2.3.4 to optimization problems whose feasible sets comprise control constraints of this type, it is necessary to compute different normal cones to U_{ad}. This has been done for convex-valued mappings U in [17, Section 6.3.3]. Here we want to substantially generalize these considerations, see [86].

Let (Ω,Σ) be a measurable space and let $\mathcal{X}$ be a Banach space. For an arbitrary set-valued mapping $\Psi\colon \Omega \rightrightarrows \mathcal{X}$, we introduce its domain, its graph, and its preimages as it was done at the end of Section 2.3.3 for set-valued mappings between Banach spaces. We call Ψ measurable if for any open set $O \subseteq \mathcal{X}$, the preimage $\Psi^{-1}(O)$ is measurable. If Ψ is a closed-valued mapping and $\operatorname{dom}\Psi = \Omega$ holds, then it is measurable if and only if there is a sequence of measurable functions $\psi_k\colon \Omega \to \mathcal{X}$, $k \in \mathbb{N}$, such that $\Psi(\omega) = \operatorname{cl}\{\psi_k(\omega) \,|\, k \in \mathbb{N}\}$ is satisfied everywhere on Ω, see [100, Theorem 6.3.19]. Clearly, Ψ is measurable if and only if $\operatorname{cl}\Psi\colon \Omega \rightrightarrows \mathcal{X}$ defined by $(\operatorname{cl}\Psi)(\omega) := \operatorname{cl}\Psi(\omega)$ for any $\omega \in \Omega$ is measurable. Thus, measurability of set-valued mappings is a concept which is not suited for set-valued mappings with nonclosed images since it somehow disregards information on nonclosedness. We call Ψ graph-measurable if $\operatorname{gph}\Psi$ is measurable w.r.t. the measurable space $(\Omega\times\mathcal{X}, \Sigma\otimes\mathfrak{B}(\mathcal{X}))$ where $\mathfrak{B}(\mathcal{X})$ denotes the Borelean σ-algebra induced by $\mathcal{X}$, i.e. the smallest σ-algebra which contains all open sets of $\mathcal{X}$, and

$\Sigma \otimes \mathfrak{B}(\mathcal{X})$ represents the smallest σ-algebra containing the Cartesian product $\Sigma \times \mathfrak{B}(\mathcal{X})$. Note that we use $\mathfrak{B}^m := \mathfrak{B}(\mathbb{R}^m)$. Any measurable mapping with closed images is graph-measurable by means of [100, Proposition 6.2.10]. For any Banach space $\mathcal{Y}$, a mapping $\psi\colon \Omega \times \mathcal{X} \to \mathcal{Y}$ is called a Carathéodory function if for any $\omega \in \Omega$, the mapping $\psi(\omega, \cdot)$ is continuous, whereas $\psi(\cdot, x)$ is measurable for any fixed $x \in \mathcal{X}$. Any Carathéodory function is $\Sigma \otimes \mathfrak{B}(\mathcal{X})$-measurable, see [100, Theorem 6.2.6].

In the upcoming considerations, let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete, σ-finite, and nonatomic measure space satisfying $\mathfrak{m}(\Omega) > 0$. Here nonatomic simply means that for any $M \in \Sigma$ with $\mathfrak{m}(M) > 0$, we find $\widetilde{M} \in \Sigma$ such that $\mathfrak{m}(M) > \mathfrak{m}(\widetilde{M}) > 0$ holds.
Next, we state two lemmas we need in order to compute the variational objects of interest related to pointwise defined sets in Lebesgue spaces.

Lemma 2.41. Let $\mathcal{X}$ be a Banach space, let $\Psi\colon \Omega \rightrightarrows \mathcal{X}$ be a measurable set-valued mapping with compact images, and let $\psi\colon \Omega \times \mathcal{X} \to \mathbb{R}$ be a Carathéodory function. Then the set $M \subseteq \Omega$ defined below is measurable:

$$M := \{\omega \in \Omega \,|\, \forall x \in \Psi(\omega)\colon \psi(\omega, x) \leq 0\}.$$

Proof. Let us define a set-valued mapping $\Phi\colon \Omega \rightrightarrows \mathbb{R}$ by $\Phi(\omega) := \{\psi(\omega, x) \,|\, x \in \Psi(\omega)\}$ for any $\omega \in \Omega$. Then $M = \{\omega \in \Omega \,|\, \Phi(\omega) \subseteq (-\infty, 0]\} = \Omega \setminus \Phi^{-1}((0, \infty))$ holds. Since for any $\omega \in \Omega$, $\psi(\omega, \cdot)$ is continuous while $\Psi(\omega)$ is compact, $\Phi(\omega)$ is closed. Thus, by means of [6, Theorem 8.2.8], Φ is measurable. Consequently, the preimage $\Phi^{-1}((0, \infty))$ is measurable as well which shows the claim. □

Lemma 2.42. Let $\{\psi_k\}$ be a sequence of measurable functions mapping from Ω to $\mathbb{R}$. Assume that for almost every $\omega \in \Omega$, there is a number $N(\omega) \in \mathbb{N}$ such that $\psi_k(\omega) \geq 0$ holds true for all $k \geq N(\omega)$. Then the function $\psi\colon \Omega \to \mathbb{R}$ defined below is measurable:

$$\forall \omega \in \Omega\colon \quad \psi(\omega) := \min\{n \in \mathbb{N} \,|\, \forall k \geq n\colon\ \psi_k(\omega) \geq 0\}.$$

Proof. Let us set $\psi_0 \equiv -1$. Then we have

$$\psi^{-1}(\{n\}) = \psi_{n-1}^{-1}((-\infty, 0)) \cap \bigcap_{i \in \mathbb{N},\, i \geq n} \psi_i^{-1}([0, \infty))$$

for any $n \in \mathbb{N}$, i.e. the preimage of $\{n\}$ under ψ is the countable intersection of measurable sets and, thus, measuable. Since $\mathbb{N}$ is countable, this yields the measurability of ψ already. □

Let $\Xi_n \subseteq \mathbb{R}^n$ denote the unit simplex defined by

$$\Xi_n := \left\{\kappa \in \mathbb{R}^n \,\middle|\, \textstyle\sum_{i=1}^n \kappa^i = 1,\ \kappa^1, \ldots, \kappa^n \geq 0\right\}.$$

For a nonempty, closed set $K \subseteq \mathbb{R}^m$ and a point $\mathrm{u} \in \operatorname{conv} K$, we introduce the notation

$$r_K(\mathrm{u}) := \min\left\{ \max_{i=1,\ldots,m+1} \left|\mathrm{u}^i\right|_2 \,\middle|\, \exists \mathrm{u}^1, \ldots, \mathrm{u}^{m+1} \in K\ \exists \kappa \in \Xi_{m+1}\colon\ \textstyle\sum_{i=1}^{m+1} \kappa^i \mathrm{u}^i = \mathrm{u} \right\}.$$

Note that Carathéodory's theorem yields $r_K(\mathrm{u}) < \infty$. Moreover, the coercivity of the maximum norm implies that the minimum is attained. By definition, $r_K(\mathrm{u})$ is the smallest radius $r \geq 0$ such that the relation $\mathrm{u} \in \operatorname{conv}\big(K \cap \mathbb{B}^r_{m,2}(0)\big)$ holds. Thus, we obtain $r_K(\mathrm{u}) = |\mathrm{u}|_2$ for $\mathrm{u} \in K$. Exploiting Carathéodory's theorem once more, we easily see

$$r_K(\mathrm{u}) = \min\left\{ \max_{i=1,\ldots,l} \left|\mathrm{u}^i\right|_2 \,\middle|\, \exists l \in \mathbb{N}\ \exists \mathrm{u}^1, \ldots, \mathrm{u}^l \in K\ \exists \kappa \in \Xi_l\colon\ \textstyle\sum_{i=1}^{l} \kappa^i \mathrm{u}^i = \mathrm{u} \right\}. \tag{2.22}$$

For the sake of completeness, let us set $r_K(\tilde{\mathrm{u}}) = \infty$ for all $\tilde{\mathrm{u}} \notin \operatorname{conv} K$.
Now, let $K\colon \Omega \rightrightarrows \mathbb{R}^m$ be a measurable set-valued mapping with closed images and let $u\colon \Omega \to \mathbb{R}^m$ be a measurable function satisfying $u(\omega) \in K(\omega)$ for almost every $\omega \in \Omega$. Then [6, Theorem 8.2.11] yields that the marginal function $\omega \mapsto r_{K(\omega)}(u(\omega))$ is measurable as well.
Let us recall the notion of decomposable sets which can be retraced to [107] where a similar concept was introduced.

Definition 2.2. A nonempty set $\mathbb{K} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ is said to be decomposable if for any $A \in \Sigma$ and $u_1, u_2 \in \mathbb{K}$, the relation $\chi_A u_1 + (1 - \chi_A) u_2 \in \mathbb{K}$ is satisfied.

In [65], the authors present different properties and calculus rules for decomposable sets. A convenient overview of the available theory on decomposable sets can be found in the recent monograph [100]. Clearly, if we find a measurable set-valued mapping $K \colon \Omega \rightrightarrows \mathbb{R}^m$ with $\operatorname{dom} K = \Omega$ and closed images such that

$$\mathbb{K} := \{u \in L^p(\mathfrak{M}, \mathbb{R}^m) \,|\, u(\omega) \in K(\omega) \text{ f.a.a. } \omega \in \Omega\} \tag{2.23}$$

is satisfied, then $\mathbb{K}$ is closed and decomposable. The converse holds true as well by means of [100, Theorem 6.4.6]. The following lemma presents some essential calculus rules for decomposable sets.

Lemma 2.43. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete, σ-finite, and nonatomic measure space. Furthermore, let $\mathbb{K} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ be a nonempty, closed, and decomposable set with its associated set-valued mapping $K \colon \Omega \rightrightarrows \mathbb{R}^m$ possessing nonempty, closed images.

(a) We have

$$\operatorname{cl}^w \mathbb{K} = \overline{\operatorname{conv}}\, \mathbb{K} = \{u \in L^p(\mathfrak{M}, \mathbb{R}^m) \,|\, u(\omega) \in \overline{\operatorname{conv}}\, K(\omega) \text{ f.a.a. } \omega \in \Omega\}$$

where $p' \in (1, \infty)$ satisfying $1/p + 1/p' = 1$ is the conjugate coefficient of p. Since the set-valued mapping $\omega \mapsto \overline{\operatorname{conv}}\, K(\omega)$ is measurable, the closed set $\operatorname{cl}^w \mathbb{K}$ is decomposable.
If $L^q(\mathfrak{M})$ is separable for all $q \in [1, \infty)$, then we additionally have

$$\operatorname{conv} \mathbb{K} \subseteq \left\{u \in L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, r_{K(\cdot)}(u(\cdot)) \in L^p(\mathfrak{M})\right\} \subseteq \operatorname{cl}^w_{\mathsf{seq}} \mathbb{K}.$$

(b) Assume that $0 \in \mathbb{K}$ holds true. Then we have

$$\mathbb{K}^\circ = \left\{\eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \in K(\omega)^\circ \text{ f.a.a. } \omega \in \Omega\right\}$$

and since $\omega \mapsto K(\omega)^\circ$ is measurable, the closed set $\mathbb{K}^\circ$ is decomposable.

Proof. Let us start with the proof of (a). The first calculus rule follows combining [100, Propositions 6.4.14, 6.4.19]. Here the fact that $(\Omega, \Sigma, \mathfrak{m})$ is nonatomic is of essential importance. The measurability of the mapping $\omega \mapsto \overline{\operatorname{conv}}\, K(\omega)$ is shown in [6, Theorem 8.2.2].
Next, we prove the presented inclusions giving a lower and upper estimate of the set

$$S := \left\{u \in L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, r_{K(\cdot)}(u(\cdot)) \in L^p(\mathfrak{M})\right\}.$$

First, we show $\mathbb{K} \subseteq S$. Therefore, choose $\bar{u} \in \mathbb{K}$ arbitrarily. Then we have $r_{K(\omega)}(\bar{u}(\omega)) = |\bar{u}(\omega)|_2$. By definition of $L^p(\mathfrak{M}, \mathbb{R}^m)$, we have $|\bar{u}(\cdot)|_2 \in L^p(\mathfrak{M})$, i.e. $\bar{u} \in S$ is valid. Next, let us show that S is convex. Thus, we choose $u, \tilde{u} \in S$ and $\alpha \in [0, 1]$ arbitrarily. Since $u(\omega), \tilde{u}(\omega) \in \operatorname{conv} K(\omega)$ holds for almost all $\omega \in \Omega$, we obtain $\alpha u(\omega) + (1 - \alpha)\tilde{u}(\omega) \in \operatorname{conv} K(\omega)$ for almost every $\omega \in \Omega$. Moreover, $\alpha u + (1 - \alpha)\tilde{u}$ is measurable. Thus, $\omega \mapsto r_{K(\omega)}(\alpha u(\omega) + (1 - \alpha)\tilde{u}(\omega))$ is well-defined and measurable. By definition, for almost all $\omega \in \Omega$, we find $u^1(\omega), \ldots, u^{m+1}(\omega), \tilde{u}^1(\omega), \ldots, \tilde{u}^{m+1}(\omega) \in K(\omega)$ and $\kappa(\omega), \tilde{\kappa}(\omega) \in \Xi_{m+1}$ which satisfy

$$\begin{aligned} \textstyle\sum_{i=1}^{m+1} \kappa^i(\omega)\, u^i(\omega) &= u(\omega), \qquad & r_{K(\omega)}(u(\omega)) &= \max_{i=1,\ldots,m+1} \left|u^i(\omega)\right|_2, \\ \textstyle\sum_{i=1}^{m+1} \tilde{\kappa}^i(\omega)\, \tilde{u}^i(\omega) &= \tilde{u}(\omega), \qquad & r_{K(\omega)}(\tilde{u}(\omega)) &= \max_{i=1,\ldots,m+1} \left|\tilde{u}^i(\omega)\right|_2. \end{aligned}$$

Thus, we have $(\alpha\kappa(\omega), (1 - \alpha)\tilde{\kappa}(\omega)) \in \Xi_{2m+2}$ and

$$\textstyle\sum_{i=1}^{m+1} \alpha\, \kappa^i(\omega)\, u^i(\omega) + \sum_{i=1}^{m+1} (1 - \alpha)\, \tilde{\kappa}^i(\omega)\, \tilde{u}^i(\omega) = \alpha\, u(\omega) + (1 - \alpha)\, \tilde{u}(\omega).$$

Consequently, exploiting the representation (2.22), we obtain

$$\begin{aligned} 0 &\leq r_{K(\omega)}(\alpha\, u(\omega) + (1 - \alpha)\, \tilde{u}(\omega)) \\ &\leq \max\left\{\max_{i=1,\ldots,m+1} \left|u^i(\omega)\right|_2; \max_{i=1,\ldots,m+1} \left|\tilde{u}^i(\omega)\right|_2\right\} = \max\left\{r_{K(\omega)}(u(\omega)); r_{K(\omega)}(\tilde{u}(\omega))\right\} \end{aligned}$$

for almost all $\omega \in \Omega$. Hence, $r_{K(\cdot)}(u(\cdot)), r_{K(\cdot)}(\tilde{u}(\cdot)) \in L^p(\mathfrak{M})$ leads to $r_{K(\cdot)}(\alpha u(\cdot) + (1-\alpha)\tilde{u}(\cdot)) \in L^p(\mathfrak{M})$ which yields the relation $\alpha u + (1-\alpha)\tilde{u} \in S$. This shows the convexity of S. Since S is a convex superset of $\mathbb{K}$, we obtain $\operatorname{conv}\mathbb{K} \subseteq S$, i.e. the first inclusion we wanted to show.
For the proof of the upper estimate of S, we first assume that $\mathfrak{m}$ is a finite measure, i.e. that $\mathfrak{m}(\Omega) < \infty$ holds. Choose $u \in S$ arbitrarily. Then we have $u(\omega) \in \operatorname{conv} K(\omega)$ for almost every $\omega \in \Omega$. Let us define a set-valued mapping $\Upsilon_u \colon \Omega \rightrightarrows \mathbb{R}^{m(m+1)} \times \mathbb{R}^{m+1}$ as stated below:

$$\forall \omega \in \Omega\colon \quad \Upsilon_u(\omega) := \left\{ (\mathfrak{u}^1, \dots, \mathfrak{u}^{m+1}, \kappa) \;\middle|\; \begin{array}{l} \mathfrak{u}^1, \dots, \mathfrak{u}^{m+1} \in K(\omega) \cap \mathbb{B}_{m,2}^{r_{K(\omega)}(u(\omega))}(0), \\ \kappa \in \Xi_{m+1},\ \sum_{i=1}^{m+1} \kappa^i \mathfrak{u}^i = u(\omega) \end{array} \right\}.$$

By definition of r, the images of Υ_u are nonempty. It is easy to see from [6, Theorem 8.2.9] that Υ_u is measurable. Applying the measurable selection theorem, see [6, Theorem 8.1.3], yields the existence of measurable functions $\kappa \colon \Omega \to \Xi_{m+1}$ and $u^1, \dots, u^{m+1} \colon \Omega \to \mathbb{R}^m$ with

$$u^1(\omega), \dots, u^{m+1}(\omega) \in K(\omega) \cap \mathbb{B}_{m,2}^{r_{K(\omega)}(u(\omega))}(0), \qquad \sum_{i=1}^{m+1} \kappa^i(\omega)\, u^i(\omega) = u(\omega)$$

for almost every $\omega \in \Omega$. Recalling $r_{K(\cdot)}(u(\cdot)) \in L^p(\mathfrak{M})$, $u^1, \dots, u^{m+1} \in L^p(\mathfrak{M}, \mathbb{R}^m)$ is obtained.
Next, we define the set-valued mapping $E \colon \Omega \rightrightarrows \mathbb{R}^{m+1}$ via $E(\omega) := \{\mathbf{e}^i \in \mathbb{R}^{m+1} \,|\, i = 1, \dots, m+1\}$ for any $\omega \in \Omega$ where $\mathbf{e}^1, \dots, \mathbf{e}^{m+1} \in \mathbb{R}^{m+1}$ denote the $m+1$ unit vectors in $\mathbb{R}^{m+1}$. By $\mathbb{E} \subseteq L^p(\mathfrak{M}, \mathbb{R}^{m+1})$ we denote the closed, decomposable set associated with E. Note that $\mathbb{E}$ is nonempty since $\mathfrak{m}(\Omega) < \infty$ is assumed. Using the first formula of this lemma, we have

$$\operatorname{cl}^w \mathbb{E} = \overline{\operatorname{conv}}\, \mathbb{E} = \left\{ \xi \in L^p(\mathfrak{M}, \mathbb{R}^{m+1}) \,\middle|\, \xi(\omega) \in \Xi_{m+1} \text{ f.a.a. } \omega \in \Omega \right\}$$

and, thus, $\kappa \in \operatorname{cl}^w \mathbb{E}$ is obtained. Noting that $\operatorname{cl}^w \mathbb{E}$ is bounded (it lies within an $\sqrt{m+1}\,\mathfrak{m}(\Omega)^{1/p}$-ball around zero) while $L^{p'}(\mathfrak{M})$ is separable, we have $\kappa \in \operatorname{cl}^w_{\mathsf{seq}} \mathbb{E}$, see [83, Corollary 2.6.20]. Thus, we find a sequence $\{w_k\} \subseteq \mathbb{E}$ with $w_k \rightharpoonup \kappa$. Note that $\{w_k\}$ is even bounded in $L^\infty(\mathfrak{M}, \mathbb{R}^{m+1})$. By means of Lemma A.3 we already have $w_k \overset{\star}{\to} \kappa$ in $L^\infty(\mathfrak{M}, \mathbb{R}^{m+1})$. This yields

$$\textstyle\sum_{i=1}^{m+1} (w_k)_i u^i \rightharpoonup \sum_{i=1}^{m+1} \kappa^i u^i = u$$

in $L^p(\mathfrak{M}, \mathbb{R}^m)$. Moreover, $\{w_k\} \subseteq \mathbb{E}$ yields $\sum_{i=1}^{m+1} (w_k)_i u^i \in \mathbb{K}$ for all $k \in \mathbb{N}$. Thus, we have $u \in \operatorname{cl}^w_{\mathsf{seq}} \mathbb{K}$. This shows the claim for $\mathfrak{m}(\Omega) < \infty$. In the case where $\mathfrak{M}$ is an infinite measure space, we can prove this inclusion by working on a countable partition $\{\Omega_n\} \subseteq \Sigma$ of Ω with $0 < \mathfrak{m}(\Omega_n) < \infty$ for all $n \in \mathbb{N}$.
Finally, we show the statement (b). The inclusion $\supseteq$ follows in a straightforward way from the definition of the polar cone. For the proof of the inclusion $\subseteq$, we choose $\bar{\eta} \in \mathbb{K}^\circ$ arbitrarily and assume that $\bar{\eta}$ does not belong to the set on the right. Then we find $\varepsilon > 0$ and a set $\Omega' \in \Sigma$ of positive, finite measure such that

$$\forall \omega \in \Omega'\, \exists u(\omega) \in K(\omega)\colon \quad \bar{\eta}(\omega) \cdot u(\omega) \geq \varepsilon$$

is valid. We define a set-valued mapping $\Phi \colon \Omega \rightrightarrows \mathbb{R}^m$ by

$$\forall \omega \in \Omega\colon \quad \Phi(\omega) := \begin{cases} \{\mathfrak{u} \in \mathbb{R}^m \,|\, \mathfrak{u} \in K(\omega),\, \bar{\eta}(\omega) \cdot \mathfrak{u} \geq \varepsilon\} & \text{if } \omega \in \Omega', \\ \{0\} & \text{if } \omega \in \Omega \setminus \Omega'. \end{cases}$$

Noting that $K \colon \Omega \rightrightarrows \mathbb{R}^m$ is closed-valued and measurable while $(\omega, \mathfrak{u}) \mapsto \bar{\eta}(\omega) \cdot \mathfrak{u}$ is a Carathéodory function, Φ is measurable by [6, Theorem 8.2.9] and, thus, possesses a measurable selection $\bar{u}$, i.e. $\bar{u}$ is a measurable function from $L^0(\mathfrak{M}, \mathbb{R}^m)$ and satisfies $\bar{u}(\omega) \in \Phi(\omega)$ almost everywhere on Ω, see [6, Theorem 8.1.3]. For any $k \in \mathbb{N}$, we define a measurable set $A_k \subseteq \Omega$ as stated below:

$$A_k := \{\omega \in \Omega' \,|\, k \geq |\bar{u}(\omega)|_2\}.$$

Obviously, we have $A_k \subseteq A_{k+1}$ for all $k \in \mathbb{N}$ and $\bigcup_{k \in \mathbb{N}} A_k = \Omega'$. Let us set $\Omega'_n := \bigcup_{k=1}^n A_k$ which is measurable again for any $n \in \mathbb{N}$. Now, choose $N \in \mathbb{N}$ so large such that $\mathfrak{m}(\Omega'_N) \geq \frac{1}{2}\mathfrak{m}(\Omega')$ is valid. Let us set $\hat{u} := \bar{u}\chi_{\Omega'_N}$. By construction, $\hat{u}$ is essentially bounded on Ω'_N and vanishes on $\Omega \setminus \Omega'_N$. Thus, we have

$\hat{u} \in L^p(\mathfrak{M}, \mathbb{R}^m)$. Moreover, $\hat{u} \in \mathbb{K}$ is obtained from the definition of $\bar{u}$ and $0 \in K(\omega)$ for almost every $\omega \in \Omega$. This leads to

$$\langle \bar{\eta}, \hat{u} \rangle_{L^p(\mathfrak{M}, \mathbb{R}^m)} = \int_{\Omega'_N} \bar{\eta}(\omega) \cdot \bar{u}(\omega) \mathrm{d}\mathfrak{m} \geq \varepsilon\, \mathfrak{m}(\Omega'_N) \geq \frac{\varepsilon}{2}\, \mathfrak{m}(\Omega') > 0$$

which contradicts $\bar{\eta} \in \mathbb{K}^\circ$. This shows the inclusion $\subseteq$. The measurability of $\omega \mapsto K(\omega)^\circ$ follows from [108, Exercise 14.12]. This completes the proof. □

Applying the above lemma, we can characterize the weak closure of a decomposable set equivalently by taking the closure of its weak sequential closure.

Proposition 2.44. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete, σ-finite, and nonatomic measure space such that $L^q(\mathfrak{M})$ is separable for all $q \in [1, \infty)$. Furthermore, let $\mathbb{K} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ be a nonempty, closed, and decomposable set. Then we have

$$\mathrm{cl}\, \mathrm{cl}^w_{\mathsf{seq}}\, \mathbb{K} = \mathrm{cl}^w\, \mathbb{K}.$$

Proof. We invoke Lemma 2.43 in order to see

$$\mathrm{conv}\, \mathbb{K} \subseteq \mathrm{cl}^w_{\mathsf{seq}}\, \mathbb{K} \subseteq \mathrm{cl}^w\, \mathbb{K}.$$

Taking the closure yields

$$\overline{\mathrm{conv}}\, \mathbb{K} \subseteq \mathrm{cl}\, \mathrm{cl}^w_{\mathsf{seq}}\, \mathbb{K} \subseteq \mathrm{cl}\, \mathrm{cl}^w\, \mathbb{K} = \mathrm{cl}^w\, \mathbb{K}.$$

Since we have $\overline{\mathrm{conv}}\, \mathbb{K} = \mathrm{cl}^w\, \mathbb{K}$ from Lemma 2.43, we deduce that these inclusions are in fact equalities. This yields the claim. □

As a result, we obtain an interesting property of decomposable sets.

Corollary 2.45. Let the assumptions of Proposition 2.44 hold. Then $\mathbb{K}$ is weakly sequentially closed if and only if it is weakly closed.

Proof. Clearly, if $\mathbb{K}$ is weakly closed, then we have $\mathbb{K} \subseteq \mathrm{cl}^w_{\mathsf{seq}}\, \mathbb{K} \subseteq \mathrm{cl}^w\, \mathbb{K} = \mathbb{K}$, i.e. $\mathbb{K} = \mathrm{cl}^w_{\mathsf{seq}}\, \mathbb{K}$, and $\mathbb{K}$ is weakly sequentially closed as well.
On the other hand, if $\mathbb{K}$ is weakly sequentially closed, we obtain $\mathbb{K} = \mathrm{cl}\, \mathbb{K} = \mathrm{cl}\, \mathrm{cl}^w_{\mathsf{seq}}\, \mathbb{K} = \mathrm{cl}^w\, \mathbb{K}$ from the closedness of $\mathbb{K}$ and Proposition 2.44. This shows that $\mathbb{K}$ is weakly closed. □

The property of closed, decomposable sets described in the above corollary is remarkable and does not hold for general closed sets in infinite-dimensional Banach spaces. For some Hilbert space $\mathcal{H}$ with orthonormal basis $\{e_k \,|\, k \in \mathbb{N}\} \subseteq \mathcal{H}$, one could consider the closed set $H := \{\sqrt{k}\, e_k \,|\, k \in \mathbb{N}\}$ which is closed and weakly sequentially closed since the sequence $\{\sqrt{k} e_k\}$ does not converge weakly. However, we have $0 \in \mathrm{cl}^w\, H$ and, thus, H cannot be weakly closed.
Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete and σ-finite measure space. It is well-known from [6, Corollary 8.5.2] that for any nonempty, closed, decomposable set $\mathbb{K} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ whose associated set-valued mapping $K \colon \Omega \rightrightarrows \mathbb{R}^m$ possesses derivable images and any $\bar{u} \in \mathbb{K}$, we have

$$\mathcal{T}_{\mathbb{K}}(\bar{u}) = \left\{ d \in L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, d(\omega) \in \mathcal{T}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}. \tag{2.24}$$

Moreover, the mapping $\omega \mapsto \mathcal{T}_{K(\omega)}(\bar{u}(\omega))$ is measurable, see [108, Theorem 14.26], and closed-valued. Thus, $\mathcal{T}_{\mathbb{K}}(\bar{u})$ is decomposable. We will use the above formula and Lemma 2.43 in order to compute some other variational objects of interest. Therefore, let us fix the following assumptions throughout the whole section.

Assumption 2.1. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete, σ-finite, as well as nonatomic measure space and let $K \colon \Omega \rightrightarrows \mathbb{R}^m$ be a measurable set-valued mapping with closed and derivable images. For $p \in (1, \infty)$, we consider the corresponding decomposable set $\mathbb{K} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ defined in (2.23) and fix an arbitrary point $\bar{u} \in \mathbb{K}$. Furthermore, $p' \in (1, \infty)$ satisfying $1/p + 1/p' = 1$ denotes the conjugate coefficient of p. Finally, we assume that $L^q(\mathfrak{M})$ is separable for all $q \in [1, \infty)$.

Choosing Ω to be a domain in $\mathbb{R}^d$ equipped with the Borelean σ-algebra induced by Ω and Lebesgue's measure, the corresponding measure space $\mathfrak{M}$ is σ-finite and nonatomic. Moreover, $L^q(\Omega)$ is separable for all $q \in [1,\infty)$, see [1, Theorem 2.21]. Thus, the formal completion of $(\Omega, \mathfrak{B}(\Omega), \mathfrak{l})$, see [16, Section 1.5], is a complete, σ-finite, and nonatomic measure space such that $L^q(\Omega)$ is separable for all $q \in [1,\infty)$. Consequently, in the classical setting of optimal control, the assumptions on the underlying measure space and the Lebesgue spaces of interest are naturally valid.
As a direct consequence of Lemma 2.43 and the formula for the tangent cone in (2.24), we obtain

$$\mathcal{K}_{\mathbb{K}}(\bar{u},\bar{\eta}) = \left\{ d \in L^p(\mathfrak{M},\mathbb{R}^m) \,\middle|\, d(\omega) \in \mathcal{K}_{K(\omega)}(\bar{u}(\omega),\bar{\eta}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}$$

for any $\bar{\eta} \in \mathcal{T}_{\mathbb{K}}(\bar{u})^\circ$, where the latter set possesses a pointwise characterization as well, see Lemma 2.43. Let us check how the weak tangent cone to a decomposable set behaves.

Proposition 2.46. We have

$$\mathcal{T}_{\mathbb{K}}(\bar{u}) \subseteq \mathcal{T}^w_{\mathbb{K}}(\bar{u}) \subseteq \overline{\operatorname{conv}}\, \mathcal{T}_{\mathbb{K}}(\bar{u}).$$

Proof. The first inclusion follows naturally from the definitions of tangent cone and weak tangent cone. For the proof of the second inclusion, choose $d \in \mathcal{T}^w_{\mathbb{K}}(\bar{u})$ arbitrarily and assume on the contrary that $d \notin \overline{\operatorname{conv}}\, \mathcal{T}_{\mathbb{K}}(\bar{u})$ is satisfied. Applying Lemma 2.11, we find $\eta \in \mathcal{T}_{\mathbb{K}}(\bar{u})^\circ$ which satisfies $\langle \eta, d \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} > 0$. From the definition of the dual pairing there must exist a set $E_1 \in \Sigma$ with positive measure and a scalar $\alpha > 0$ such that $\eta(\omega) \cdot d(\omega) \geq \alpha$ holds true for almost all $\omega \in E_1$. On the other hand, there are sequences $\{d_k\} \subseteq L^p(\mathfrak{M},\mathbb{R}^m)$ and $\{t_k\} \subseteq \mathbb{R}$ satisfying $d_k \rightharpoonup d$ and $t_k \searrow 0$ such that $u_k := \bar{u} + t_k d_k \in \mathbb{K}$ holds for all $k \in \mathbb{N}$. Thus, we have $u_k \to \bar{u}$ in $L^p(\mathfrak{M},\mathbb{R}^m)$ and, by choosing a subsequence which we denote by $\{u_k\}$ again, we find $E_2 \in \Sigma$ satisfying $E_2 \subseteq E_1$, $\mathfrak{m}(E_2) \geq \frac{2}{3}\mathfrak{m}(E_1)$, and $u_k \to \bar{u}$ in $L^\infty(\mathfrak{M}|_{E_2},\mathbb{R}^m)$, see Lemma A.2. Let us define $M > 0$ and $\varepsilon > 0$ by

$$M := \|\eta\|_{L^{p'}(\mathfrak{M},\mathbb{R}^m)} \sup \left\{ \|d_k\|_{L^p(\mathfrak{M},\mathbb{R}^m)} \,\middle|\, k \in \mathbb{N} \right\}, \qquad \varepsilon := \frac{\alpha}{4}\frac{\mathfrak{m}(E_1)}{M}.$$

Note that M is well-defined since $\{d_k\}$ is bounded. By means of Lemma 2.41, for any $\gamma > 0$, the set

$$E_\gamma := E_2 \cap \left\{ \omega \in \Omega \,\middle|\, \forall \mathfrak{u} \in K(\omega) \cap \mathbb{B}^\gamma_{m,2}(\bar{u}(\omega)) \colon\ \eta(\omega) \cdot (\mathfrak{u} - \bar{u}(\omega)) \leq \varepsilon |\eta(\omega)|_2 |\mathfrak{u} - \bar{u}(\omega)|_2 \right\}$$

is measurable. Since $\eta(\omega) \in \mathcal{T}_{K(\omega)}(\bar{u}(\omega))^\circ$ holds almost everywhere on Ω, see Lemma 2.43, we find a sufficiently small $\bar{\gamma}$ such that $\mathfrak{m}(E_{\bar{\gamma}}) \geq \frac{1}{2}\mathfrak{m}(E_1)$ is satisfied. From $\|u_k - \bar{u}\|_{L^\infty(\mathfrak{M}|_{E_2},\mathbb{R}^m)} \to 0$, for sufficiently large $k \in \mathbb{N}$, we have $u_k(\omega) \in K(\omega) \cap \mathbb{B}^{\bar{\gamma}}_{m,2}(\bar{u}(\omega))$ almost everywhere on $E_{\bar{\gamma}}$. On the other hand, the mappings $\omega \to |\eta(\omega)|_2$ and $\omega \mapsto |u_k(\omega) - \bar{u}(\omega)|_2$ come from $L^{p'}(\mathfrak{M})$ and $L^p(\mathfrak{M})$, respectively. Thus, we can exploit $d_k \rightharpoonup d$ and Hölder's inequality to obtain

$$\begin{aligned}
\frac{2\alpha}{3}\mathfrak{m}(E_{\bar{\gamma}}) &= \frac{2}{3}\int_{E_{\bar{\gamma}}} \alpha \mathrm{d}\mathfrak{m} \leq \int_{E_{\bar{\gamma}}} \eta(\omega) \cdot d_k(\omega) \mathrm{d}\mathfrak{m} = \frac{1}{t_k}\int_{E_{\bar{\gamma}}} \eta(\omega) \cdot (u_k(\omega) - \bar{u}(\omega)) \mathrm{d}\mathfrak{m} \\
&\leq \frac{\varepsilon}{t_k}\int_{E_{\bar{\gamma}}} |\eta(\omega)|_2 |u_k(\omega) - \bar{u}(\omega)|_2 \mathrm{d}\mathfrak{m} \leq \frac{\varepsilon}{t_k} \|\eta\|_{L^{p'}(\mathfrak{M},\mathbb{R}^m)} \|u_k - \bar{u}\|_{L^p(\mathfrak{M},\mathbb{R}^m)} \\
&= \varepsilon \|\eta\|_{L^{p'}(\mathfrak{M},\mathbb{R}^m)} \|d_k\|_{L^p(\mathfrak{M},\mathbb{R}^m)} \leq \varepsilon M = \frac{\alpha}{4}\mathfrak{m}(E_1) \leq \frac{\alpha}{2}\mathfrak{m}(E_{\bar{\gamma}})
\end{aligned}$$

for sufficiently large k. This, however, is a contradiction. □

Since the Lebesgue space $L^p(\mathfrak{M},\mathbb{R}^m)$ is reflexive, we immediately obtain the following formula for the Fréchet normal cone from Lemma 2.43 and the above result for the weak tangent cone.

Corollary 2.47. We have

$$\widehat{\mathcal{N}}_{\mathbb{K}}(\bar{u}) = \left\{ \eta \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \eta(\omega) \in \widehat{\mathcal{N}}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}.$$

Especially, $\widehat{\mathcal{N}}_{\mathbb{K}}(\bar{u})$ is a nonempty, closed, decomposable set.

Proof. Using Proposition 2.46 as well as $A^\circ = (\overline{\operatorname{conv}} A)^\circ$ for arbitrary sets in Banach spaces, we obtain

$$\mathcal{T}_{\mathbb{K}}(\bar{u})^\circ = \big(\overline{\operatorname{conv}}\, \mathcal{T}_{\mathbb{K}}(\bar{u})\big)^\circ \subseteq \mathcal{T}^w_{\mathbb{K}}(\bar{u})^\circ \subseteq \mathcal{T}_{\mathbb{K}}(\bar{u})^\circ.$$

Consequently, Lemma 2.43 and the reflexivity of $L^p(\mathfrak{M}, \mathbb{R}^m)$ yield

$$\begin{aligned}\widehat{\mathcal{N}}_{\mathbb{K}}(\bar{u}) = \mathcal{T}^w_{\mathbb{K}}(\bar{u})^\circ = \mathcal{T}_{\mathbb{K}}(\bar{u})^\circ &= \Big\{\eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\Big|\, \eta(\omega) \in \mathcal{T}_{K(\omega)}(\bar{u}(\omega))^\circ \text{ f.a.a. } \omega \in \Omega\Big\} \\ &= \Big\{\eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\Big|\, \eta(\omega) \in \widehat{\mathcal{N}}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega\Big\}.\end{aligned}$$

Since we have $\widehat{\mathcal{N}}_{\mathbb{K}}(\bar{u}) = \mathcal{T}_{\mathbb{K}}(\bar{u})^\circ$ from above while $\mathcal{T}_{\mathbb{K}}(\bar{u})$ is nonempty, closed, and decomposable, the same holds true for $\widehat{\mathcal{N}}_{\mathbb{K}}(\bar{u})$ by means of Lemma 2.43. This completes the proof. □

Trying to compute the limiting normal cone by definition using the above formula for the Fréchet normal cone, it would be necessary to apply a so-called measurable selection theorem, see [6, Theorem 8.1.3] or [100, Theorem 6.3.17], to the set-valued mapping $\omega \mapsto \operatorname{gph} \widehat{\mathcal{N}}_{K(\omega)}$. However, since the images of this mapping are not closed in general, these results cannot be exploited. On the other hand, we can use the graph-measurability of $\omega \mapsto \operatorname{gph} \widehat{\mathcal{N}}_{K(\omega)}$ provided in the upcoming lemma.

Lemma 2.48. The mapping $\omega \mapsto \operatorname{gph} \widehat{\mathcal{N}}_{K(\omega)}$ is graph-measurable, i.e. the set

$$\operatorname{gph} \operatorname{gph} \widehat{\mathcal{N}}_{K(\cdot)} := \Big\{(\omega, \mathrm{u}, \mathrm{v}) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^m \,\Big|\, \mathrm{u} \in K(\omega),\, \mathrm{v} \in \widehat{\mathcal{N}}_{K(\omega)}(\mathrm{u})\Big\}$$

is measurable w.r.t. the σ-algebra $\Sigma \otimes \mathfrak{B}^m \otimes \mathfrak{B}^m$.

Proof. First, we show that for an arbitrary closed set $C \subseteq \mathbb{R}^m$, the set $\operatorname{gph} \widehat{\mathcal{N}}_C$ can be represented as a countable intersection of countably many unions of sets. Thus, observe that from the definition of the Fréchet normal cone

$$\begin{aligned}\operatorname{gph} \widehat{\mathcal{N}}_C &= \Big\{(\mathrm{u}, \mathrm{v}) \in \mathbb{R}^m \times \mathbb{R}^m \,\Big|\, \mathrm{u} \in C,\, \mathrm{v} \in \widehat{\mathcal{N}}_C(\mathrm{u})\Big\} \\ &= \{(\mathrm{u}, \mathrm{v}) \in C \times \mathbb{R}^m \,|\, \forall \{c_k\} \subseteq C\colon\ c_k \to \mathrm{u} \implies \mathrm{v} \cdot (c_k - \mathrm{u}) \leq o(|c_k - \mathrm{u}|_2)\} \\ &= \Big\{(\mathrm{u}, \mathrm{v}) \in C \times \mathbb{R}^m \,\Big|\, \forall \varepsilon \in \mathbb{Q}^+ \exists \gamma \in \mathbb{Q}^+ \forall c \in \mathbb{U}^\gamma_{m,2}(\mathrm{u}) \cap C\colon\ \mathrm{v} \cdot (c - \mathrm{u}) \leq \varepsilon |c - \mathrm{u}|_2\Big\} \\ &= \bigcap_{\varepsilon \in \mathbb{Q}^+} \bigcup_{\gamma \in \mathbb{Q}^+} \Big\{(\mathrm{u}, \mathrm{v}) \in C \times \mathbb{R}^m \,\Big|\, \forall c \in \mathbb{U}^\gamma_{m,2}(\mathrm{u}) \cap C\colon\ \mathrm{v} \cdot (c - \mathrm{u}) \leq \varepsilon |c - \mathrm{u}|_2\Big\}\end{aligned}$$

is satisfied. Therein, $\mathbb{Q}^+$ denotes the set of all positive rational numbers. Let $\{c_k\} \subseteq C$ be a sequence which is dense in C. Then $\{c_k\} \cap \mathbb{U}^\gamma_{m,2}(\mathrm{u})$ is dense in $C \cap \mathbb{B}^\gamma_{m,2}(\mathrm{u})$. Thus, we obtain

$$\begin{aligned}&\Big\{(\mathrm{u}, \mathrm{v}) \in C \times \mathbb{R}^m \,\Big|\, \forall c \in \mathbb{U}^\gamma_{m,2}(\mathrm{u}) \cap C\colon\ \mathrm{v} \cdot (c - \mathrm{u}) \leq \varepsilon |c - \mathrm{u}|_2\Big\} \\ &= \bigcap_{k \in \mathbb{N}} \Big(\big[(C \setminus \mathbb{U}^\gamma_{m,2}(c_k)) \times \mathbb{R}^m\big] \cup \{(\mathrm{u}, \mathrm{v}) \in C \times \mathbb{R}^m \,|\, \mathrm{v} \cdot (c_k - \mathrm{u}) \leq \varepsilon |c_k - \mathrm{u}|_2\}\Big) \\ &= \bigcap_{k \in \mathbb{N}} \Big((C \times \mathbb{R}^m) \cap \Big[\big[(\mathbb{R}^m \setminus \mathbb{U}^\gamma_{m,2}(c_k)) \times \mathbb{R}^m\big] \cup \{(\mathrm{u}, \mathrm{v}) \in \mathbb{R}^m \times \mathbb{R}^m \,|\, \mathrm{v} \cdot (c_k - \mathrm{u}) \leq \varepsilon |c_k - \mathrm{u}|_2\}\Big]\Big).\end{aligned}$$

Now, let es turn our attention to $\operatorname{gph} \operatorname{gph} \widehat{\mathcal{N}}_{K(\cdot)}$. Since K is closed-valued and measurable, we find a sequence of measurable functions $\psi_k\colon \Omega \to \mathbb{R}^m$, $k \in \mathbb{N}$, such that $\{\psi_k(\omega) \,|\, k \in \mathbb{N}\}$ is dense in $K(\omega)$ for all $\omega \in \Omega$. Let us introduce

$$\begin{aligned}S_1 &:= \operatorname{gph} K \times \mathbb{R}^m, \\ S_2(\gamma, k) &:= \{(\omega, \mathrm{u}, \mathrm{v}) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^m \,|\, |\mathrm{u} - \psi_k(\omega)|_2 \geq \gamma\}, \\ S_3(\varepsilon, k) &:= \{(\omega, \mathrm{u}, \mathrm{v}) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^m \,|\, \mathrm{v} \cdot (\psi_k(\omega) - \mathrm{u}) \leq \varepsilon |\psi_k(\omega) - \mathrm{u}|_2\}.\end{aligned}$$

Then from the above considerations we obtain

$$\operatorname{gph} \operatorname{gph} \widehat{\mathcal{N}}_{K(\cdot)} = \bigcap_{\varepsilon \in \mathbb{Q}^+} \bigcup_{\gamma \in \mathbb{Q}^+} \bigcap_{k \in \mathbb{N}} \Big(S_1 \cap \big[S_2(\gamma, k) \cup S_3(\varepsilon, k)\big]\Big). \tag{2.25}$$

Since K is a measurable set-valued mapping with closed images, it is graph-measurable and, hence, S_1 is a measurable set. For arbitrary $\gamma, \varepsilon \in \mathbb{Q}^+$ and $k \in \mathbb{N}$, define functions $\varphi_{\gamma,k}, \vartheta_{\varepsilon,k} \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ by

$$\begin{aligned} \forall \omega \in \Omega\, \forall \mathrm{u}, \mathrm{v} \in \mathbb{R}^m \colon \quad & \varphi_{\gamma,k}(\omega, \mathrm{u}, \mathrm{v}) := \gamma - |\mathrm{u} - \psi_k(\omega)|_2, \\ & \vartheta_{\varepsilon,k}(\omega, \mathrm{u}, \mathrm{v}) := \mathrm{v} \cdot (\psi_k(\omega) - \mathrm{u}) - \varepsilon |\psi_k(\omega) - \mathrm{u}|_2. \end{aligned}$$

Then $\varphi_{\gamma,k}(\cdot, \mathrm{u}, \mathrm{v})$ and $\vartheta_{\varepsilon,k}(\cdot, \mathrm{u}, \mathrm{v})$ are measurable, whereas $\varphi_{\gamma,k}(\omega, \cdot, \cdot)$ and $\vartheta_{\varepsilon,k}(\omega, \cdot, \cdot)$ are continuous. Thus, $\varphi_{\gamma,k}$ and $\vartheta_{\varepsilon,k}$ are Carathéodory functions and, hence, measurable w.r.t. $\Sigma \otimes \mathfrak{B}^m \otimes \mathfrak{B}^m$. Consequently, the sets

$$S_2(\gamma, k) = \varphi_{\gamma,k}^{-1}((-\infty, 0]), \quad S_3(\varepsilon, k) = \vartheta_{\varepsilon,k}^{-1}((-\infty, 0])$$

are measurable w.r.t. $\Sigma \otimes \mathfrak{B}^m \otimes \mathfrak{B}^m$. Now, the claim follows from the representation (2.25). □

Using Lemma 2.48 and the obvious fact $\operatorname{gph} \mathcal{N}_{\mathbb{K}}^s = \operatorname{cl} \operatorname{gph} \widehat{\mathcal{N}}_{\mathbb{K}}$, we obtain an explicit representation of the strong limiting normal cone to decomposable sets.

Proposition 2.49. We have

$$\mathcal{N}_{\mathbb{K}}^s(\bar{u}) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \in \mathcal{N}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}.$$

Especially, $\mathcal{N}_{\mathbb{K}}^s(\bar{u})$ is a nonempty, closed, decomposable set.

Proof. Exploiting the graph-measurability of $\omega \mapsto \operatorname{gph} \widehat{\mathcal{N}}_{K(\omega)}$ which was obtained in Lemma 2.48, it is possible to invoke [100, Proposition 6.4.20] in order to see

$$\begin{aligned} \operatorname{gph} \mathcal{N}_{\mathbb{K}}^s &= \operatorname{cl} \operatorname{gph} \widehat{\mathcal{N}}_{\mathbb{K}} \\ &= \operatorname{cl} \left\{ (u, \eta) \in L^p(\mathfrak{M}, \mathbb{R}^m) \times L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, (u(\omega), \eta(\omega)) \in \operatorname{gph} \widehat{\mathcal{N}}_{K(\omega)} \text{ f.a.a. } \omega \in \Omega \right\} \\ &= \left\{ (u, \eta) \in L^p(\mathfrak{M}, \mathbb{R}^m) \times L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, (u(\omega), \eta(\omega)) \in \operatorname{cl} \operatorname{gph} \widehat{\mathcal{N}}_{K(\omega)} \text{ f.a.a. } \omega \in \Omega \right\} \\ &= \left\{ (u, \eta) \in L^p(\mathfrak{M}, \mathbb{R}^m) \times L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, (u(\omega), \eta(\omega)) \in \operatorname{gph} \mathcal{N}_{K(\omega)} \text{ f.a.a. } \omega \in \Omega \right\} \end{aligned}$$

Therein, the second equality follows from Corollary 2.47, whereas the last one is a simple consequence of the fact that in finite-dimensional spaces, the strong limiting normal cone coincides with the limiting normal cone. Now, the pointwise representation of the strong limiting normal cone follows from applying the definition of the graph of a set-valued mapping. Since the mapping $\omega \mapsto \mathcal{N}_{K(\omega)}(\bar{u}(\omega))$ is measurable, see [108, Theorem 14.26], and closed-valued, see Lemma 2.16, $\mathcal{N}_{\mathbb{K}}^s(\bar{u})$ is closed and decomposable. □

The above proposition is useful to prove the upcoming result. It presents upper and lower estimates for the limiting normal cone to a decomposable set which satisfies Assumption 2.1.

Proposition 2.50. We have

$$\operatorname{cl}_{\mathsf{seq}}^w \mathcal{N}_{\mathbb{K}}^s(\bar{u}) \subseteq \mathcal{N}_{\mathbb{K}}(\bar{u}) \subseteq \overline{\operatorname{conv}}\, \mathcal{N}_{\mathbb{K}}^s(\bar{u}).$$

Proof. Let us show the first inclusion. Therefore, we choose $\eta \in \operatorname{cl}_{\mathsf{seq}}^w \mathcal{N}_{\mathbb{K}}^s(\bar{u})$ arbitrarily. Thus, there is a sequence $\{\eta_k\} \subseteq \mathcal{N}_{\mathbb{K}}^s(\bar{u})$ which satisfies $\eta_k \rightharpoonup \eta$. By definition of the strong limiting normal cone and the reflexivity of $L^{p'}(\mathfrak{M}, \mathbb{R}^m)$, for every $k \in \mathbb{N}$, we find sequences $\{u_{k,l}\} \subseteq \mathbb{K}$ and $\{\eta_{k,l}\} \subseteq L^{p'}(\mathfrak{M}, \mathbb{R}^m)$ with $u_{k,l} \to \bar{u}$ and $\eta_{k,l} \to \eta_k$ as $l \to \infty$, and $\eta_{k,l} \in \widehat{\mathcal{N}}_{\mathbb{K}}(u_{k,l})$ for all $l \in \mathbb{N}$. Consequently, for any $k \in \mathbb{N}$, we find $l_k \in \mathbb{N}$ which satisfies

$$\|u_{k,l_k} - \bar{u}\|_{L^p(\mathfrak{M}, \mathbb{R}^m)} \leq \tfrac{1}{k}, \qquad \|\eta_{k,l_k} - \eta_k\|_{L^{p'}(\mathfrak{M}, \mathbb{R}^m)} \leq \tfrac{1}{k}.$$

For arbitrary $v \in L^p(\mathfrak{M}, \mathbb{R}^m)$, we have

$$\begin{aligned} \left| \langle \eta_{k,l_k} - \eta, v \rangle_{L^p(\mathfrak{M}, \mathbb{R}^m)} \right| &\leq \left| \langle \eta_{k,l_k} - \eta_k, v \rangle_{L^p(\mathfrak{M}, \mathbb{R}^m)} \right| + \left| \langle \eta_k - \eta, v \rangle_{L^p(\mathfrak{M}, \mathbb{R}^m)} \right| \\ &\leq \|\eta_{k,l_k} - \eta_k\|_{L^{p'}(\mathfrak{M}, \mathbb{R}^m)} \|v\|_{L^p(\mathfrak{M}, \mathbb{R}^m)} + \left| \langle \eta_k - \eta, v \rangle_{L^p(\mathfrak{M}, \mathbb{R}^m)} \right| \\ &\leq \tfrac{1}{k} \|v\|_{L^p(\mathfrak{M}, \mathbb{R}^m)} + \left| \langle \eta_k - \eta, v \rangle_{L^p(\mathfrak{M}, \mathbb{R}^m)} \right| \xrightarrow{k \to \infty} 0, \end{aligned}$$

i.e. $\eta_{k,l_k} \rightharpoonup \eta$ as $k \to \infty$. From $u_{k,l_k} \to \bar{u}$ and $\eta_{k,l_k} \in \widehat{\mathcal{N}}_{\mathbb{K}}(u_{k,l_k})$ for all $k \in \mathbb{N}$, $\eta \in \mathcal{N}_{\mathbb{K}}(\bar{u})$ is obtained. For the proof of the second inclusion, we assume on the contrary that $\mathcal{N}_{\mathbb{K}}(\bar{u}) \nsubseteq \overline{\operatorname{conv}} \mathcal{N}^s_{\mathbb{K}}(\bar{u})$ is satisfied. Then we find $\eta \in \mathcal{N}_{\mathbb{K}}(\bar{u})$ and, by Lemma 2.11, $d \in \mathcal{N}^s_{\mathbb{K}}(\bar{u})^\circ$ such that $\langle \eta, d \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} > 0$ is satisfied. Thus, there are a set $E_1 \in \Sigma$ of positive measure and a scalar $\alpha > 0$ which satisfy $\eta(\omega) \cdot d(\omega) \geq \alpha$ for almost every $\omega \in E_1$. On the other hand, there are sequences $\{u_k\} \subseteq \mathbb{K}$ and $\{\eta_k\} \subseteq L^{p'}(\mathfrak{M}, \mathbb{R}^m)$ satisfying $\eta_k \in \widehat{\mathcal{N}}_{\mathbb{K}}(u_k)$ for all $k \in \mathbb{N}$ and $u_k \to \bar{u}$ as well as $\eta_k \rightharpoonup \eta$. Invoking Lemma A.2, there are a subsequence of $\{u_k\}$, which we denote by $\{u_k\}$ again, and $E_2 \in \Sigma$ with $\mathfrak{m}(E_2) \geq \frac{2}{3}\mathfrak{m}(E_1)$ and $E_2 \subseteq E_1$ such that $u_k \to \bar{u}$ holds true in $L^\infty(\mathfrak{M}|_{E_2}, \mathbb{R}^m)$. We introduce $M > 0$ and $\varepsilon > 0$ as stated below:

$$M := \|d\|_{L^p(\mathfrak{M},\mathbb{R}^m)} \sup\left\{ \|\eta_k\|_{L^{p'}(\mathfrak{M},\mathbb{R}^m)} \,\middle|\, k \in \mathbb{N} \right\}, \qquad \varepsilon := \frac{\alpha}{4}\frac{\mathfrak{m}(E_1)}{M}.$$

For almost every fixed $\omega \in E_2$, we obtain $d(\omega) \in \mathcal{N}_{K(\omega)}(\bar{u}(\omega))^\circ$ from Lemma 2.43 as well as Proposition 2.49 and $\eta_k(\omega) \in \widehat{\mathcal{N}}_{K(\omega)}(u_k(\omega))$ for any $k \in \mathbb{N}$ from Corollary 2.47. Next, we show that for almost every $\omega \in E_2$, there is a number $N(\omega) \in \mathbb{N}$ such that

$$\forall k \in \mathbb{N}\colon \quad k \geq N(\omega) \implies \eta_k(\omega) \cdot d(\omega) \leq \varepsilon |\eta_k(\omega)|_2 |d(\omega)|_2$$

is satisfied. If this is not the case, we have $d(\omega) \neq 0$ and, hence, there needs to be a subsequence $\{\eta_{k_l}(\omega)\}$ of $\{\eta_k(\omega)\}$ with

$$\forall l \in \mathbb{N}\colon \quad \eta_{k_l}(\omega) \cdot d(\omega) > \varepsilon |\eta_{k_l}(\omega)|_2 |d(\omega)|_2.$$

Since $\eta_{k_l}(\omega)$ cannot vanish and $u_{k_l}(\omega) \to \bar{u}(\omega)$ holds true almost everywhere on E_2, we obtain that the bounded sequence $\{\eta_{k_l}(\omega)/|\eta_{k_l}(\omega)|_2\}$ converges w.l.o.g. to $\mathrm{v} \in \mathcal{N}_{K(\omega)}(\bar{u}(\omega))$. We deduce

$$0 \geq \mathrm{v} \cdot d(\omega) = \lim_{l \to \infty} \frac{\eta_{k_l}(\omega)}{|\eta_{k_l}(\omega)|_2} \cdot d(\omega) \geq \varepsilon |d(\omega)|_2$$

which contradicts $d(\omega) \neq 0$. Hence, a number $N(\omega)$ with the desired properties exists for almost every $\omega \in E_2$. Thus, the function $P\colon \Omega \to \mathbb{N}_0$ given by

$$\forall \omega \in \Omega\colon \quad P(\omega) := \begin{cases} \min\{n \in \mathbb{N} \,|\, \forall k \in \mathbb{N}\colon\ k \geq n \implies \eta_k(\omega) \cdot d(\omega) \leq \varepsilon |\eta_k(\omega)|_2 |d(\omega)|_2\} & \text{if } \omega \in E_2, \\ 0 & \text{if } \omega \in \Omega \setminus E_2 \end{cases}$$

is well-defined and $E_N := \{\omega \in E_2 \,|\, N \geq P(\omega)\}$ is measurable for any $N \in \mathbb{N}$ by Lemma 2.42. For some $\bar{N} \in \mathbb{N}$ sufficiently large, $\mathfrak{m}(E_{\bar{N}}) \geq \frac{1}{2}\mathfrak{m}(E_1)$ holds. Thus, for sufficiently large $k \in \mathbb{N}$, we obtain

$$\begin{aligned} \frac{2\alpha}{3}\mathfrak{m}(E_{\bar{N}}) \leq \frac{2}{3}\int_{E_{\bar{N}}} \alpha \mathrm{d}\mathfrak{m} \leq \int_{E_{\bar{N}}} \eta_k(\omega) \cdot d(\omega) \mathrm{d}\mathfrak{m} \leq \varepsilon \int_{E_{\bar{N}}} |\eta_k(\omega)|_2 |d(\omega)|_2 \mathrm{d}\mathfrak{m} \\ \leq \varepsilon \|d\|_{L^p(\mathfrak{M},\mathbb{R}^m)} \|\eta_k\|_{L^{p'}(\mathfrak{M},\mathbb{R}^m)} \leq \varepsilon M = \frac{\alpha}{4}\mathfrak{m}(E_1) \leq \frac{\alpha}{2}\mathfrak{m}(E_{\bar{N}}) \end{aligned}$$

which is a contradiction. This completes the proof. □

Now, we are able to state the main result of this section.

Proposition 2.51. We have

$$\operatorname{conv} \mathcal{N}^s_{\mathbb{K}}(\bar{u}) \subseteq \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, r_{\mathcal{N}_{K(\cdot)}(\bar{u}(\cdot))}(\eta(\cdot)) \in L^{p'}(\mathfrak{M}) \right\} \subseteq \mathcal{N}_{\mathbb{K}}(\bar{u}) \subseteq \mathcal{N}^c_{\mathbb{K}}(\bar{u}).$$

Moreover, we obtain

$$\operatorname{cl} \mathcal{N}_{\mathbb{K}}(\bar{u}) = \mathcal{N}^c_{\mathbb{K}}(\bar{u}) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \in \mathcal{N}^c_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}.$$

Especially, $\mathcal{N}^c_{\mathbb{K}}(\bar{u})$ is a nonempty, closed, decomposable set.

Proof. The first assertion follows from Lemma 2.43 and Propositions 2.49 as well as 2.50.
We use Proposition 2.50 once more to find

$$\operatorname{conv}\mathcal{N}^s_{\mathbb{K}}(\bar{u}) \subseteq \mathcal{N}_{\mathbb{K}}(\bar{u}) \subseteq \overline{\operatorname{conv}}\,\mathcal{N}^s_{\mathbb{K}}(\bar{u}).$$

This yields $\operatorname{cl}\mathcal{N}_{\mathbb{K}}(\bar{u}) = \overline{\operatorname{conv}}\,\mathcal{N}^s_{\mathbb{K}}(\bar{u}) = \overline{\operatorname{conv}}\,\mathcal{N}_{\mathbb{K}}(\bar{u}) = \mathcal{N}^c_{\mathbb{K}}(\bar{u})$. Applying Lemma 2.43 and Proposition 2.49 again, the pointwise characterization of the Clarke normal cone follows.
Since $\mathcal{N}^s_{\mathbb{K}}(\bar{u})$ is nonempty, closed, and decomposable, the same holds true for $\overline{\operatorname{conv}}\,\mathcal{N}^s_{\mathbb{K}}(\bar{u}) = \mathcal{N}^c_{\mathbb{K}}(\bar{u})$, see Lemma 2.43. This completes the proof. □

Recalling the convexity of Clarke's tangent cone and exploiting Lemma 2.43, we obtain the following corollary from Proposition 2.51.

Corollary 2.52. We have

$$\mathcal{T}^c_{\mathbb{K}}(\bar{u}) = \left\{ d \in L^p(\mathfrak{M},\mathbb{R}^m) \,\middle|\, d(\omega) \in \mathcal{T}^c_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}.$$

Especially, $\mathcal{T}^c_{\mathbb{K}}(\bar{u})$ is a nonempty, closed, decomposable set.

Proposition 2.51 shows that the limiting normal cone to any closed, decomposable set in a Lebesgue space contains the convex hull of the strong limiting normal cone and is dense in the corresponding Clarke normal cone. In particular, if it is closed, it equals Clarke's normal cone. Note that the nonatomicity of the underlying measure space $\mathfrak{M}$ is essential for this result. The above proposition is related to Lyapunov's convexity theorem, see [6, Theorem 8.7.3], which says that the set $\{\int_S u(\omega)\mathrm{d}\mathfrak{m} \,|\, S \in \Sigma\}$ is convex and compact for any $u \in L^1(\mathfrak{M},\mathbb{R}^m)$. This theorem implies the convexity of certain integral functions whenever the measure space of interest is nonatomic, see [6, Theorem 8.6.4] as well as [94, Theorem 2.8, Proposition 2.10] and the references therein. Actually, it was already shown in [95] that the image of a decomposable set under componentwise integration is convex which can be interpreted as an extension of Lyapunov's convexity theorem.
Note that whenever the set $\operatorname{conv}\mathcal{N}^s_{\mathbb{K}}(\bar{u})$ is closed, then we already have $\mathcal{N}_{\mathbb{K}}(\bar{u}) = \mathcal{N}^c_{\mathbb{K}}(\bar{u})$ from Propositions 2.50 and 2.51. On the other hand, we easily see from the definition of the function r and Proposition 2.51 that

$$\begin{aligned}
&\left\{ \eta \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, r_{\mathcal{N}_{K(\cdot)}(\bar{u}(\cdot))}(\eta(\cdot)) \in L^{p'}(\mathfrak{M}) \right\} \\
&\quad\subseteq \left\{ \eta \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \eta(\omega) \in \operatorname{conv}\mathcal{N}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\} \\
&\quad\subseteq \left\{ \eta \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \eta(\omega) \in \overline{\operatorname{conv}}\,\mathcal{N}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\} = \mathcal{N}^c_{\mathbb{K}}(\bar{u})
\end{aligned}$$

holds true. Thus, Proposition 2.51 does not provide a precise characterization of the limiting normal cone whenever $\operatorname{conv}\mathcal{N}_{K(\omega)}(\bar{u}(\omega))$ is not closed almost everywhere on Ω since in the latter case, the lower bounds on $\mathcal{N}_{\mathbb{K}}(\bar{u})$ are strict subsets of $\mathcal{N}^c_{\mathbb{K}}(\bar{u})$.
Proposition 2.51 suggests that Mordukhovich's (weak) limiting constructions are not appropriate for the discussion of optimal control problems with nonconvex control constraints. In this situation, the consideration of strong limiting normals seems to be reasonable by means of Proposition 2.49. However, strong limiting normals do not enjoy a good calculus in infinite-dimensional spaces.
The following examples are included to visualize the above results. While the first example presents a situation where the limiting normal cone to a decompoable set is computable from Proposition 2.51, the second one shows that this does not need to be the case in general.

Example 2.53. Let us consider a bounded domain $\Omega \subseteq \mathbb{R}^d$, the nonempty, closed, decomposable set

$$\mathbb{K} := \left\{ u \in L^2(\Omega,\mathbb{R}^2) \,\middle|\, u_1(\omega)u_2(\omega) = 0 \text{ f.a.a. } \omega \in \Omega \right\},$$

and its element 0. Introducing

$$K := \left\{ \mathtt{u} \in \mathbb{R}^2 \,\middle|\, \mathtt{u}_1\mathtt{u}_2 = 0 \right\},$$

we obtain the equivalent representation

$$\mathbb{K} = \left\{ u \in L^2(\Omega,\mathbb{R}^2) \,\middle|\, u(\omega) \in K \text{ f.a.a. } \omega \in \Omega \right\}.$$

Clearly, the closed set K is the union of two closed, convex sets and, consequently, derivable by means of Lemma 2.15. We apply Corollary 2.47 as well as Proposition 2.49 to obtain

$$\widehat{\mathcal{N}}_{\mathbb{K}}(0) = \{0\}, \quad \mathcal{N}^s_{\mathbb{K}}(0) = \mathbb{K}.$$

Obviously, we have $\operatorname{conv} \mathcal{N}^s_{\mathbb{K}}(0) = \operatorname{conv} \mathbb{K} = L^2(\Omega, \mathbb{R}^2)$. Thus

$$\mathcal{N}_{\mathbb{K}}(0) = \mathcal{N}^c_{\mathbb{K}}(0) = L^2(\Omega, \mathbb{R}^2)$$

follows from Proposition 2.51. ■

Example 2.54. We define the nonempty, closed set $K \subseteq \mathbb{R}^3$ by

$$K := \left\{ \mathtt{u} \in \mathbb{R}^3 \,\middle|\, \mathtt{u}_1 \geq \sqrt{\mathtt{u}_2^2 + \mathtt{u}_3^2} \right\} \cup \left\{ \mathtt{u} \in \mathbb{R}^3 \,\middle|\, \mathtt{u}_1 - \mathtt{u}_3 = 0 \right\}.$$

Note that K is the union of two convex sets and, thus, derivable, see Lemma 2.15. Let us consider the associated decomposable set

$$\mathbb{K} := \left\{ u \in L^2(\Omega, \mathbb{R}^3) \,\middle|\, u(\omega) \in K \text{ f.a.a. } \omega \in \Omega \right\}$$

where $\Omega \subseteq \mathbb{R}^d$ is a bounded domain. It is easy to see the relations

$$\begin{aligned}
\widehat{\mathcal{N}}_K(0) &= \operatorname{cone}\{(-1,0,1)\}, \\
\mathcal{N}_K(0) &= \left\{ \mathtt{v} \in \mathbb{R}^3 \,\middle|\, -\mathtt{v}_1 = \sqrt{\mathtt{v}_2^2 + \mathtt{v}_3^2} \right\} \cup \operatorname{lin}\{(-1,0,1)\}, \\
\operatorname{conv} \mathcal{N}_K(0) &= \left\{ \mathtt{v} \in \mathbb{R}^3 \,\middle|\, \mathtt{v}_1 + \mathtt{v}_3 < 0 \right\} \cup \operatorname{lin}\{(-1,0,1)\}, \\
\mathcal{N}^c_K(0) = \overline{\operatorname{conv}}\, \mathcal{N}_K(0) &= \left\{ \mathtt{v} \in \mathbb{R}^3 \,\middle|\, \mathtt{v}_1 + \mathtt{v}_3 \leq 0 \right\}.
\end{aligned}$$

Since $\operatorname{conv} \mathcal{N}_K(0)$ is not closed, the lower bounds for the limiting normal cone $\mathcal{N}_{\mathbb{K}}(0)$ provided in Proposition 2.51 are strict subsets of

$$\mathcal{N}^c_{\mathbb{K}}(0) = \left\{ \eta \in L^2(\Omega, \mathbb{R}^3) \,\middle|\, \eta_1(\omega) + \eta_3(\omega) \leq 0 \text{ f.a.a. } \omega \in \Omega \right\}$$

and thus, Proposition 2.51 cannot be used to obtain a precise formula for $\mathcal{N}_{\mathbb{K}}(0)$. ■

As a direct consequence of Proposition 2.51 we obtain the following well-known result for sets in Lebesgue spaces defined by box constraints. Note that we only present the scalar situation $m = 1$ here but a similar result holds for vector functions as well.

Corollary 2.55. Let $a, b \in L^p(\mathfrak{M})$ be given such that $a(\omega) < b(\omega)$ holds true for almost every $\omega \in \Omega$ and consider the set

$$\mathbb{K} := \left\{ u \in L^p(\mathfrak{M}) \,\middle|\, a(\omega) \leq u(\omega) \leq b(\omega) \text{ f.a.a. } \omega \in \Omega \right\}.$$

Fix some $\bar{u} \in \mathbb{K}$. Then we have

$$\begin{aligned}
\mathcal{N}_{\mathbb{K}}(\bar{u}) = &\left\{ \eta_a \in L^{p'}(\mathfrak{M}) \,\middle|\, \begin{array}{ll} \eta_a(\omega) \leq 0 & \text{if } a(\omega) = \bar{u}(\omega) \\ \eta_a(\omega) = 0 & \text{if } a(\omega) < \bar{u}(\omega) \end{array} \text{ f.a.a. } \omega \in \Omega \right\} \\
&+ \left\{ \eta_b \in L^{p'}(\mathfrak{M}) \,\middle|\, \begin{array}{ll} \eta_b(\omega) \geq 0 & \text{if } \bar{u}(\omega) = b(\omega) \\ \eta_b(\omega) = 0 & \text{if } \bar{u}(\omega) < b(\omega) \end{array} \text{ f.a.a. } \omega \in \Omega \right\}.
\end{aligned}$$

Proof. Clearly, we obtain

$$\mathcal{N}_{\mathbb{K}}(\bar{u}) = \left\{ \eta \in L^{p'}(\mathfrak{M}) \,\middle|\, \begin{array}{ll} \eta(\omega) \leq 0 & \text{if } a(\omega) = \bar{u}(\omega) \\ \eta(\omega) = 0 & \text{if } a(\omega) < \bar{u}(\omega) < b(\omega) \text{ f.a.a. } \omega \in \Omega \\ \eta(\omega) \geq 0 & \text{if } \bar{u}(\omega) = b(\omega) \end{array} \right\}$$

from Proposition 2.51 since for convex sets, the limiting normal cone and the Clarke normal cone always coincide. Thus, for any $\eta \in \mathcal{N}_{\mathbb{K}}(\bar{u})$, we can define $\eta_a := \min_{L^{p'}(\mathfrak{M})^+_0}\{\eta; 0\}$ and $\eta_b := \max_{L^{p'}(\mathfrak{M})^+_0}\{\eta; 0\}$ to see that the decomposition $\eta = \eta_a + \eta_b$ with $\eta_a, \eta_b \in L^{p'}(\mathfrak{M})$ is possible. Thus, the limiting normal cone can be characterized as described above. □

We want to close this section with a short look at decomposable sets which are defined by pointwise inverse images. Therefore, let $G\colon \Omega \times \mathbb{R}^m \to \mathbb{R}^n$ be a Carathéodory function continuously differentiable w.r.t. its second component such that the corresponding partial derivative $\nabla_u G\colon \Omega \times \mathbb{R}^m \to \mathbb{R}^{n\times m}$ is a Carathéodory function again. Moreover, let $\Upsilon\colon \Omega \rightrightarrows \mathbb{R}^n$ be a measurable set-valued mapping with nonempty, closed, convex images and define

$$\forall \omega \in \Omega\colon \quad K(\omega) := \{\mathtt{u} \in \mathbb{R}^m \,|\, G(\omega, \mathtt{u}) \in \Upsilon(\omega)\}.$$

By [6, Theorem 8.2.9], the set-valued mapping $K\colon \Omega \rightrightarrows \mathbb{R}^m$ is measurable and closed-valued. Assuming that for almost every $\omega \in \Omega$, the constraint qualification

$$\forall \mathtt{u} \in K(\omega)\colon \quad \nabla_u G(\omega, \mathtt{u})^\top \mathtt{l} = 0,\, \mathtt{l} \in \mathcal{N}_{\Upsilon(\omega)}(G(\omega, \mathtt{u})) \implies \mathtt{l} = 0$$

is satisfied, the images of K are derivable as well, see Lemma 2.31 and Remark 2.33. Thus, the corresponding decomposable set $\mathbb{K}$ satisfies Assumption 2.1. For the derivation of necessary optimality conditions for optimal control problems with control constraints of the form $u \in \mathbb{K}$, we need an explicit formula for the limiting normal cone to the set $\mathbb{K}$.

Proposition 2.56. Fix an arbitrary point $\bar{u} \in \mathbb{K}$. In the setting described above, we have

$$\mathcal{N}_{\mathbb{K}}(\bar{u}) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \exists \lambda \in L^0(\mathfrak{M}, \mathbb{R}^n)\colon \begin{array}{l} \eta(\omega) = \nabla_u G(\omega, \bar{u}(\omega))^\top \lambda(\omega), \\ \lambda(\omega) \in \mathcal{N}_{\Upsilon(\omega)}(G(\omega, \bar{u}(\omega))) \end{array} \text{ f.a.a. } \omega \in \Omega \right\}.$$

Proof. We apply Lemma 2.31, Remark 2.33, as well as Proposition 2.49 to obtain

$$\begin{aligned} \mathcal{N}_{\mathbb{K}}(\bar{u}) \supseteq \mathcal{N}^s_{\mathbb{K}}(\bar{u}) &= \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \in \mathcal{N}_{K(\omega)}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega \right\} \\ &= \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \exists \lambda(\omega) \in \mathbb{R}^n\colon \begin{array}{l} \eta(\omega) = \nabla_u G(\omega, \bar{u}(\omega))^\top \lambda(\omega) \\ \lambda(\omega) \in \mathcal{N}_{\Upsilon(\omega)}(G(\omega, \bar{u}(\omega))) \end{array} \text{ f.a.a. } \omega \in \Omega \right\}. \end{aligned} \tag{2.26}$$

That means the inclusion $\supseteq$ is already verified.
For some $\eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m)$, we introduce the set-valued mapping $\Theta_\eta\colon \Omega \rightrightarrows \mathbb{R}^n$ by

$$\forall \omega \in \Omega \quad \Theta_\eta(\omega) := \left\{ \mathtt{l} \in \mathcal{N}_{\Upsilon(\omega)}(G(\omega, \bar{u}(\omega))) \,\middle|\, \nabla_u G(\omega, \bar{u}(\omega))^\top \mathtt{l} - \eta(\omega) = 0 \right\}.$$

Since G is a Carathéodory function, $\omega \mapsto G(\omega, \bar{u}(\omega))$ is measurable. Now, [108, Theorem 14.26] yields that the set-valued mapping $\omega \mapsto \mathcal{N}_{\Upsilon(\omega)}(G(\omega, \bar{u}(\omega)))$ is measurable as well. On the other hand, define a function $q_\eta\colon \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ by

$$\forall \omega \in \Omega\, \forall \mathtt{l} \in \mathbb{R}^n\colon \quad q_\eta(\omega, \mathtt{l}) := \nabla_u G(\omega, \bar{u}(\omega))^\top \mathtt{l} - \eta(\omega).$$

This mapping is a Carathéodory function as well since $\omega \mapsto \nabla_u G(\omega, \bar{u}(\omega))$ and η are measurable. Invoking [6, Theorem 8.2.9], Θ_η is a measurable set-valued mapping since it possesses the representation

$$\Theta_\eta(\omega) = \left\{ \mathtt{l} \in \mathcal{N}_{\Upsilon(\omega)}(G(\omega, \bar{u}(\omega))) \,\middle|\, q_\eta(\omega, \mathtt{l}) = 0 \right\}$$

for almost all $\omega \in \Omega$. Moreover, the images of Θ_η are obviously closed almost everywhere on Ω.
Choose $\bar{\eta} \in \mathcal{N}_{\mathbb{K}}(\bar{u})$ arbitrarily. Recalling Proposition 2.51 as well as Lemma 2.31 and Remark 2.33, we easily see

$$\mathcal{N}_{\mathbb{K}}(\bar{u}) \subseteq \mathcal{N}^c_{\mathbb{K}}(\bar{u}) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \Theta_\eta(\omega) \neq \varnothing \text{ f.a.a. } \omega \in \Omega \right\}.$$

Thus, the measurable set-valued mapping $\Theta_{\bar{\eta}}$ possesses nonempty and closed images almost everywhere on Ω. Thus, by means of [6, Theorem 8.1.3], we find a measurable selection of $\Theta_{\bar{\eta}}$, i.e. a function $\lambda \in L^0(\mathfrak{M}, \mathbb{R}^n)$ which satisfies $\lambda(\omega) \in \Theta_{\bar{\eta}}(\omega)$ for almost every $\omega \in \Omega$. Recalling the definition of $\Theta_{\bar{\eta}}$, this shows the inclusion $\subseteq$ and completes the proof. □

In order to obtain a higher degree of regularity for the appearing multiplier λ, i.e. $\lambda \in L^{q'}(\mathfrak{M}, \mathbb{R}^n)$ where q' is the conjugate coefficient of some $q \in [1, p)$, one has to postulate certain growth conditions on the so-called Nemytskii operator induced by G, i.e. the mapping $u \mapsto G(\cdot, u(\cdot))$, see [52, Theorem 7, Remark 4].

3. Mathematical problems with complementarity constraints

In this chapter, we take a closer look at mathematical problems with complementarity constraints which appear frequently in many different applications. The model program of our interest possesses the following form:

$$\begin{aligned} \psi(x) &\to \min \\ g(x) &\in C \\ G(x) &\in K \\ H(x) &\in K^\circ \\ \langle H(x), G(x)\rangle_{\mathcal{Z}} &= 0. \end{aligned} \tag{MPCC}$$

Forthwith, the feasible set of (MPCC) is denoted by M. Below, we list our standing assumptions on this problem.

Assumption 3.1. Let $\psi\colon \mathcal{X} \to \mathbb{R}$ be Fréchet differentiable and let $g\colon \mathcal{X} \to \mathcal{Y}$, $G\colon \mathcal{X} \to \mathcal{Z}$, as well as $H\colon \mathcal{X} \to \mathcal{Z}^\star$ be continuously Fréchet differentiable. Here $\mathcal{X}$ and $\mathcal{Y}$ are arbitrary Banach spaces while $\mathcal{Z}$ is a reflexive Banach space. Finally, $C \subseteq \mathcal{Y}$ is assumed to be a nonempty, closed, and convex set, whereas $K \subseteq \mathcal{Z}$ is a nonempty, closed, convex cone.

Note that the reflexivity of $\mathcal{Z}$ guarantees that the complementarity constraint in (MPCC) is symmetric w.r.t. G and H. The standard finite-dimensional situation, i.e. where $\mathcal{X} = \mathbb{R}^n$, $\mathcal{Y} = \mathbb{R}^q$, and $\mathcal{Z} = \mathbb{R}^m$ hold and the complementarity in (MPCC) is induced by the cone $K = \mathbb{R}^{m,+}_0$, is well-studied. It is well-known that common constraint qualifications of reasonable strength like MFCQ fail to hold at any feasible point of the problem, and, consequently, the Fritz-John conditions hold at all feasible points. Thus, in the past, huge effort was put in constructing suitable regularity and stationarity conditions, see e.g. [42, 43, 44, 45, 80, 111, 129, 132], as well as numerical methods, see e.g. [46, 70] and the contained references, in order to handle these standard complementarity problems.
Recently, complementarity constraints induced by the cone of positive semidefinite matrices

$$\mathcal{S}^+_m := \{\mathbf{A} \in \mathcal{S}_m \,|\, \forall x \in \mathbb{R}^m\colon\ x^\top \mathbf{A} x \geq 0\},$$

see [37, 48, 124, 127], and the second-order cone

$$\mathcal{K}_m := \{(t, x) \in \mathbb{R} \times \mathbb{R}^m \,|\, t \geq |x|_2\},$$

see [79, 119, 135], were studied. In [124], the author investigates complementarity constraints induced by the more general $\mathcal{H}$-second-order cone

$$\mathcal{K}_{\mathcal{H}} := \{(t, x) \in \mathbb{R} \times \mathcal{H} \,|\, t \geq \|x\|_{\mathcal{H}}\}$$

where $\mathcal{H}$ is an arbitrary Hilbert space. Whenever $\mathcal{Z}$ is infinite-dimensional, only a manageable amount of publications is at hand which mainly focus on optimal control of variational inequalities as they appear for instance when the obstacle problem is considered, see [64, 66, 67, 68, 73, 88, 96, 125]. The situation where the complementarity constraint in governed by the cone of nonnegative functions in a Lebesgue space was recently considered in [86]. On the other hand, in [56], the authors derive necessary optimality conditions for optimal control problems with mixed control state complementarity constraints using pointwise error bound conditions and a measurabe selection theorem. Thus, their approach naturally leads

to weak results since the Lagrange multipliers corresponding to the complementarity constraint are, in general, only measurable functions and no L^p-regularity can be shown. To the best of our knowledge, [121] is the first contribution which introduces a general concept of strong stationarity which fits (MPCC). In [124], the author considers the situation where the cone K is nonpolyhedric in more detail. A generalized concept of weak stationarity is presented in the recent paper [87]. Finally, a first step towards a generalization of Mordukhovich's stationarity concept for (MPCC) has been done in [47].

3.1. Stationarity concepts for MPCCs

For the purpose of completeness, let us verify that KRZCQ is violated at any feasible point of (MPCC).

Lemma 3.1. The constraint qualification KRZCQ fails to hold at any feasible point of (MPCC).

Proof. Let $\bar{x} \in \mathcal{X}$ be a feasible point of (MPCC). We show that there is a nonvanishing singular Lagrange multiplier of (MPCC) at $\bar{x}$. By means of Remark 2.33 this yields the violation of KRZCQ.
A nonvanishing quadruple $(\lambda, \mu, \nu, \kappa) \in \mathcal{N}_C(g(\bar{x})) \times K^\circ \times K \times \mathbb{R}$ is a singular Lagrange multiplier of (MPCC) at $\bar{x}$ if and only if it satisfies

$$0 = g'(\bar{x})^\star[\lambda] + G'(\bar{x})^\star[\mu] + H'(\bar{x})^\star[\nu] + \kappa G'(\bar{x})^\star[H(\bar{x})] + \kappa H'(\bar{x})^\star[G(\bar{x})],\ \mu \in \{G(\bar{x})\}^\perp,\ \nu \in \{H(\bar{x})\}^\perp.$$

Rearranging leads to

$$0 = g'(\bar{x})^\star[\lambda] + G'(\bar{x})^\star[\mu + \kappa H(\bar{x})] + H'(\bar{x})^\star[\nu + \kappa G(\bar{x})],\ \mu \in \{G(\bar{x})\}^\perp,\ \nu \in \{H(\bar{x})\}^\perp,$$

and, thus, respecting the feasibility of $\bar{x}$, $(0, H(\bar{x}), G(\bar{x}), -1)$ is a nonvanishing singular Lagrange multiplier of (MPCC) at $\bar{x}$. □

Since KRZCQ is violated at any feasible point of (MPCC), we cannot expect the KKT conditions to be satisfied at the local optimal solutions of this problem. That is why we aim for weaker necessary optimality conditions and corresponding constraint qualifications. To this end, we first follow [87] to derive the notions of weak and strong stationarity via appropriate surrogate problems.
For fixed $\bar{x} \in M$, we define the relaxed nonlinear problem by

$$\begin{aligned} \psi(x) &\to \min \\ g(x) &\in C \\ G(x) &\in K \cap \{H(\bar{x})\}^\perp \\ H(x) &\in K^\circ \cap \{G(\bar{x})\}^\perp. \end{aligned} \tag{RNLP}$$

Obviously, $\bar{x}$ is a feasible point of (RNLP) but there may exist points in M which are not feasible for this surrogate problem. On the other hand, a feasible point of (RNLP) does not necessarily satisfy the complementarity condition in (MPCC). In order to overcome these difficulties, we introduce the tightened nonlinear problem

$$\begin{aligned} \psi(x) &\to \min \\ g(x) &\in C \\ G(x) &\in K \cap \{H(\bar{x})\}^\perp \cap (K^\circ \cap \{G(\bar{x})\}^\perp)^\perp \\ H(x) &\in K^\circ \cap \{G(\bar{x})\}^\perp \cap (K \cap \{H(\bar{x})\}^\perp)^\perp. \end{aligned} \tag{TNLP}$$

Again, $\bar{x}$ is a feasible point of (TNLP) but the feasible set of (TNLP) is a subset of M. Thus, if $\bar{x}$ is a local optimal solution of (MPCC), it is a local optimal solution of (TNLP) as well and the latter problem may satisfy standard constraint qualifications like KRZCQ, i.e. $\bar{x}$ may satisfy the KKT conditions of (TNLP) under suitable assumptions. In order to find an explicit representation of the KKT conditions of (TNLP), we need to study the geometry of the cones $P(\bar{x})$ and $Q(\bar{x})$ defined below:

$$P(\bar{x}) := K \cap \{H(\bar{x})\}^\perp \cap (K^\circ \cap \{G(\bar{x})\}^\perp)^\perp, \qquad Q(\bar{x}) := K^\circ \cap \{G(\bar{x})\}^\perp \cap (K \cap \{H(\bar{x})\}^\perp)^\perp.$$

Lemma 3.2. For any $\bar{x} \in M$, we have

$$\mathcal{N}_{P(\bar{x})}(G(\bar{x})) = \mathrm{cl}\big(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp\big) \cap \{G(\bar{x})\}^\perp, \tag{3.1a}$$
$$\mathcal{N}_{Q(\bar{x})}(H(\bar{x})) = \mathrm{cl}\big(K - K \cap \{H(\bar{x})\}^\perp\big) \cap \{H(\bar{x})\}^\perp, \tag{3.1b}$$
$$\mathcal{N}_{K\cap\{H(\bar{x})\}^\perp}(G(\bar{x})) = \mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x})), \tag{3.1c}$$
$$\mathcal{N}_{K^\circ\cap\{G(\bar{x})\}^\perp}(H(\bar{x})) = \mathcal{K}_{K}(G(\bar{x}), H(\bar{x})). \tag{3.1d}$$

Proof. Let us show (3.1a) first. The relation (3.1b) follows from similar arguments. Observe that from $\bar{x} \in M$ we obtain $H(\bar{x}) \in K^\circ \cap \{G(\bar{x})\}^\perp$. Exploiting the monotonicity of the operator lin, we derive $\mathrm{lin}\{H(\bar{x})\} \subseteq \mathrm{lin}(K^\circ \cap \{G(\bar{x})\}^\perp)$. This leads to

$$\mathrm{lin}\{H(\bar{x})\} + \mathrm{lin}(K^\circ \cap \{G(\bar{x})\}^\perp) = \mathrm{lin}(K^\circ \cap \{G(\bar{x})\}^\perp).$$

Since K° is a convex cone, we have $K^\circ + K^\circ \cap \{G(\bar{x})\}^\perp = K^\circ$. Putting these observations together with the Lemmas 2.1 and 2.12, we obtain

$$\begin{aligned} P(\bar{x})^\circ &= \mathrm{cl}\big(K^\circ + \mathrm{lin}\{H(\bar{x})\} + \mathrm{cl}\,\mathrm{lin}(K^\circ \cap \{G(\bar{x})\}^\perp)\big) = \mathrm{cl}\big(K^\circ + \mathrm{lin}\{H(\bar{x})\} + \mathrm{lin}(K^\circ \cap \{G(\bar{x})\}^\perp)\big) \\ &= \mathrm{cl}\big(K^\circ + \mathrm{lin}(K^\circ \cap \{G(\bar{x})\}^\perp)\big) = \mathrm{cl}\big(K^\circ + K^\circ \cap \{G(\bar{x})\}^\perp - K^\circ \cap \{G(\bar{x})\}^\perp\big) \\ &= \mathrm{cl}\big(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp\big). \end{aligned}$$

The fact that $P(\bar{x})$ is a closed, convex cone implies $\mathcal{N}_{P(\bar{x})}(G(\bar{x})) = P(\bar{x})^\circ \cap \{G(\bar{x})\}^\perp$ which yields the claim.
Next, we show (3.1c). Exploiting that $K \cap \{H(\bar{x})\}^\perp$ is a closed, convex cone, we easily see

$$\begin{aligned} \mathcal{N}_{K\cap\{H(\bar{x})\}^\perp}(G(\bar{x})) &= (K \cap \{H(\bar{x})\}^\perp)^\circ \cap \{G(\bar{x})\}^\perp \\ &= \mathrm{cl}(K^\circ + \mathrm{lin}\{H(\bar{x})\}) \cap \{G(\bar{x})\}^\perp = \mathcal{T}_{K^\circ}(H(\bar{x})) \cap \{G(\bar{x})\}^\perp = \mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x})). \end{aligned}$$

Property (3.1d) follows analogously. □

Using Lemma 3.2, it is possible to introduce the weak and strong stationarity conditions of (MPCC) as the KKT conditions of (TNLP) and (RNLP), respectively.

Definition 3.1. Let $\bar{x} \in M$ be arbitrarily chosen.

1. The point $\bar{x}$ is called weakly stationary (W-stationary for short) for (MPCC) provided there exist multipliers $\lambda \in \mathcal{Y}^\star$, $\mu \in \mathcal{Z}^\star$, and $\nu \in \mathcal{Z}$ which solve the system

$$0 = \psi'(\bar{x}) + g'(\bar{x})^\star[\lambda] + G'(\bar{x})^\star[\mu] + H'(\bar{x})^\star[\nu], \tag{3.2a}$$
$$\lambda \in \mathcal{N}_C(g(\bar{x})), \tag{3.2b}$$
$$\mu \in \mathrm{cl}\big(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp\big) \cap \{G(\bar{x})\}^\perp, \tag{3.2c}$$
$$\nu \in \mathrm{cl}\big(K - K \cap \{H(\bar{x})\}^\perp\big) \cap \{H(\bar{x})\}^\perp. \tag{3.2d}$$

2. The point $\bar{x}$ is called strongly stationary (S-stationary for short) for (MPCC) provided there exist multipliers $\lambda \in \mathcal{Y}^\star$, $\mu \in \mathcal{Z}^\star$, and $\nu \in \mathcal{Z}$ which satisfy (3.2a), (3.2b), and the conditions

$$\begin{aligned} &\mu \in \mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x})), \\ &\nu \in \mathcal{K}_{K}(G(\bar{x}), H(\bar{x})). \end{aligned} \tag{3.3}$$

Corollary 3.3. If $\bar{x} \in M$ is an S-stationary point of (MPCC), it is W-stationary as well.

Proof. For the proof, it is sufficient to show the inclusions

$$\mathcal{T}_{K^\circ}(H(\bar{x})) \subseteq \mathrm{cl}\big(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp\big), \qquad \mathcal{T}_{K}(G(\bar{x})) \subseteq \mathrm{cl}\big(K - K \cap \{H(\bar{x})\}^\perp\big).$$

In order to verify the first one, observe that $H(\bar{x}) \in K^\circ \cap \{G(\bar{x})\}^\perp$ follows due to the feasibility of $\bar{x}$ for (MPCC). Thus, $K^\circ + \operatorname{lin}\{H(\bar{x})\} \subseteq K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp$ is obtained. The fact that K° is a closed, convex cone leads to

$$\mathcal{T}_{K^\circ}(H(\bar{x})) = \operatorname{cl}(K^\circ + \operatorname{lin}\{H(\bar{x})\}) \subseteq \operatorname{cl}\big(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp\big).$$

Similarly, we show the second inclusion. □

In the following proposition, we state constraint qualifications which imply that local optimal solutions of (MPCC) satisfy the W- and S-stationarity conditions from Definition 3.1.

Proposition 3.4. Let $\bar{x} \in M$ be a local optimal solution of (MPCC).

1. Assume that the constraint qualification

$$\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} [\mathcal{X}] - \begin{pmatrix} \mathcal{R}_C(g(\bar{x})) \\ \mathcal{R}_K(G(\bar{x})) \cap \big(-\mathcal{K}_K(G(\bar{x}), H(\bar{x}))\big) \\ \mathcal{R}_{K^\circ}(H(\bar{x})) \cap \big(-\mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x}))\big) \end{pmatrix} = \begin{pmatrix} \mathcal{Y} \\ \mathcal{Z} \\ \mathcal{Z}^\star \end{pmatrix} \tag{3.4}$$

   is satisfied. Then $\bar{x}$ is W-stationary.

2. Assume that the constraint qualifications (3.4) and

$$\operatorname{cl}\left(\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} [\mathcal{X}] - \begin{pmatrix} \mathcal{N}_C(g(\bar{x}))_\perp \\ \mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x}))^{\circ\perp} \\ \mathcal{T}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(H(\bar{x})))}(H(\bar{x}))^{\circ\perp} \end{pmatrix}\right) = \begin{pmatrix} \mathcal{Y} \\ \mathcal{Z} \\ \mathcal{Z}^\star \end{pmatrix} \tag{3.5}$$

   are satisfied. Then $\bar{x}$ is S-stationary.

3. Assume that the constraint qualification

$$\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} \text{ is surjective} \tag{3.6}$$

   is satisfied. Then $\bar{x}$ is S-stationary.

Proof. For the proof of the first assertion, it is sufficient to show that (3.4) equals KRZCQ for (TNLP). Hence, we only need to verify

$$\begin{aligned} \mathcal{R}_{P(\bar{x})}(G(\bar{x})) &= \mathcal{R}_K(G(\bar{x})) \cap \big(-\mathcal{K}_K(G(\bar{x}), H(\bar{x}))\big), \\ \mathcal{R}_{Q(\bar{x})}(H(\bar{x})) &= \mathcal{R}_{K^\circ}(H(\bar{x})) \cap \big(-\mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x}))\big). \end{aligned}$$

Since $G(\bar{x})$ is an element of the linear space $\{H(\bar{x})\}^\perp \cap (K^\circ \cap \{G(\bar{x})\}^\perp)^\perp$, we obtain

$$\begin{aligned} \mathcal{R}_{P(\bar{x})}(G(\bar{x})) &= K \cap \{H(\bar{x})\}^\perp \cap (K^\circ \cap \{G(\bar{x})\}^\perp)^\perp + \operatorname{lin}\{G(\bar{x})\} \\ &= \big(K + \operatorname{lin}\{G(\bar{x})\}\big) \cap \{H(\bar{x})\}^\perp \cap (K^\circ \cap \{G(\bar{x})\}^\perp)^\perp \\ &= \mathcal{R}_K(G(\bar{x})) \cap \{H(\bar{x})\}^\perp \cap \mathcal{T}_K(G(\bar{x}))^{\circ\perp} \\ &= \mathcal{R}_K(G(\bar{x})) \cap \{H(\bar{x})\}^\perp \cap \mathcal{T}_K(G(\bar{x})) \cap \big(-\mathcal{T}_K(G(\bar{x}))\big) \\ &= \mathcal{R}_K(G(\bar{x})) \cap \big(-\mathcal{T}_K(G(\bar{x}))\big) \cap \{H(\bar{x})\}^\perp \\ &= \mathcal{R}_K(G(\bar{x})) \cap \big(-\mathcal{K}_K(G(\bar{x}), H(\bar{x}))\big) \end{aligned}$$

from Lemma 2.12. The formula for $\mathcal{R}_{Q(\bar{x})}(H(\bar{x}))$ follows similarly.
The proof of the second assertion follows from [121, Theorem 5.2, Proposition 5.1] provided we can show that the constraint qualification (2.16) for (TNLP) equals (3.5). Hence, it is sufficient to derive the formulae

$$\mathcal{N}_{P(\bar{x})}(G(\bar{x}))^\perp = \mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x}))^{\circ\perp}, \qquad \mathcal{N}_{Q(\bar{x})}(H(\bar{x}))^\perp = \mathcal{T}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(H(\bar{x})))}(H(\bar{x}))^{\circ\perp}.$$

We exploit Lemmas 2.11, 2.12, and 3.2 in order to see

$$\begin{aligned}
\mathcal{N}_{P(\bar{x})}(G(\bar{x}))^{\perp} &= \Big(\mathrm{cl}\big(K^{\circ} - K^{\circ} \cap \{G(\bar{x})\}^{\perp}\big) \cap \{G(\bar{x})\}^{\perp}\Big)^{\perp} \\
&= \Big(\mathrm{cl}\big(K^{\circ} - K^{\circ} \cap \{G(\bar{x})\}^{\perp}\big) \cap \{G(\bar{x})\}^{\perp}\Big)^{\circ} \cap \Big(\mathrm{cl}\big(K^{\circ} \cap \{G(\bar{x})\}^{\perp} - K^{\circ}\big) \cap \{G(\bar{x})\}^{\perp}\Big)^{\circ} \\
&= \mathrm{cl}\Big(\big(K^{\circ} - \mathcal{T}_K(G(\bar{x}))^{\circ}\big)^{\circ} + \mathrm{lin}\{G(\bar{x})\}\Big) \cap \mathrm{cl}\Big(\big(\mathcal{T}_K(G(\bar{x}))^{\circ} - K^{\circ}\big)^{\circ} + \mathrm{lin}\{G(\bar{x})\}\Big) \\
&= \mathrm{cl}\Big(K \cap \big(-\mathcal{T}_K(G(\bar{x}))\big) + \mathrm{lin}\{G(\bar{x})\}\Big) \cap \mathrm{cl}\Big(-\big(K \cap \big(-\mathcal{T}_K(G(\bar{x}))\big)\big) - \mathrm{lin}\{G(\bar{x})\}\Big) \\
&= \mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x})) \cap \big(-\mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x}))\big) \\
&= \mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x}))^{\circ\perp}.
\end{aligned}$$

By similar argumentation, we obtain the formula for $\mathcal{N}_{Q(\bar{x})}(H(\bar{x}))^{\perp}$.
Finally, the third assertion follows from the second one observing that the surjectivity of the given operator implies the constraint qualifications (3.4) and (3.5). This completes the proof. □

Remark 3.5. Whenever $\mathcal{Y}$ and $\mathcal{Z}$ are finite-dimensional, (3.5) implies (3.4), see Remark 2.34. Following Remarks 2.33 and 2.34, in the situation $\mathcal{Y} = \mathbb{R}^q$, $\mathcal{Z} = \mathbb{R}^m$, and $K = \mathbb{R}^{m,+}_0$, the constraint qualifications (3.4) and (3.5) equal MFCQ and LICQ for (TNLP) at $\bar{x}$, respectively. These qualification conditions are called MPCC-MFCQ and MPCC-LICQ, see [44, Definition 2.1].

Let $\bar{x} \in M$ be arbitrarily chosen. It is shown in [121, Lemma 5.1] that $\bar{x}$ satisfies the classical KKT conditions of (MPCC) if and only if there are multipliers $\lambda \in \mathcal{Y}^{\star}$, $\mu \in \mathcal{Z}^{\star}$, and $\nu \in \mathcal{Z}$ which satisfy (3.2a), (3.2b), and

$$\begin{aligned}
&\mu \in \mathcal{R}_{K^{\circ}}(H(\bar{x})) \cap \{G(\bar{x})\}^{\perp}, \\
&\nu \in \mathcal{R}_{K}(G(\bar{x})) \cap \{H(\bar{x})\}^{\perp}.
\end{aligned} \tag{3.7}$$

Thus, the KKT conditions are even stronger than the S-stationarity conditions. In the recent paper [135], the authors illustrate this phenomenon by means of MPCCs whose complementarity condition is induced by the second-order cone and rename the corresponding KKT conditions as K-stationarity conditions in order to depict that they provide a possible necessary optimality condition for (MPCC) as well. However, in this thesis, we are not going to use the term K-stationarity. Clearly, if K is a polyhedral cone, then $\mathcal{R}_K(G(\bar{x}))$ and $\mathcal{R}_{K^{\circ}}(H(\bar{x}))$ are already closed and, thus, in this case, the KKT conditions equal the S-stationarity conditions. Especially, this holds true for MPCCs with $\mathcal{Z} = \mathbb{R}^m$ and $K = \mathbb{R}^{m,+}_0$. However, an example from semidefinite complementarity programming presented in [121, Section 6.2] reveals that even in the presence of the constraint qualification (3.6), a local optimal solution of (MPCC) does not need to satisfy the KKT conditions in general.
Let $\bar{x} \in M$ be an S-stationary point of (MPCC) such that K is polyhedric w.r.t. $(G(\bar{x}), H(\bar{x}))$. It was shown in [121, Theorem 5.1] that in this case, there do not exist first-order decent directions of (MPCC), i.e. we have $\psi'(\bar{x})[\delta] \geq 0$ for all $\delta \in \mathcal{T}_M(\bar{x})$ (actually, this holds even when $\mathcal{T}_M(\bar{x})$ is replaced by its larger linearization cone). Thus, the S-stationarity conditions possess a reasonable strength whenever K is polyhedric. In the absence of polyhedricity, one can use a linearization approach and exploit the fact that any cone is polyhedric w.r.t. the origin in order to obtain S-stationarity conditions of appropriate strength provided $\mathcal{Z}$ is a Hilbert space such that the projection onto K is directionally differentiable in the sense of Haraux, see [57]. This procedure is presented in [124] and applied to the cases where K is the positive semidefinite cone $\mathcal{S}^{+}_m$ and the $\mathcal{H}$-second-order cone $\mathcal{K}_{\mathcal{H}}$.
For the study of common finite-dimensional MPCCs, there exist several other stationarity concepts stronger than W- but weaker than S-stationarity, see e.g. [129]. Here we want to generalize the stationarity concept of Mordukhovich to (MPCC). Therefore, we introduce the normal cone mapping $\mathcal{N}_K \colon \mathcal{Z} \rightrightarrows \mathcal{Z}^{\star}$ induced by the cone K which maps any $z \in K$ to the limiting normal cone to K at z and any $\tilde{z} \notin K$ to the empty set. Since K is a convex cone, we obtain

$$\mathrm{gph}\, \mathcal{N}_K = \{(z, z^*) \in K \times K^{\circ} \mid \langle z^*, z \rangle_{\mathcal{Z}} = 0\}.$$

Thus, (MPCC) is equivalent to

$$\begin{aligned} \psi(x) &\to \min \\ g(x) &\in C \\ (G(x), H(x)) &\in \operatorname{gph} \mathcal{N}_K. \end{aligned}$$

This justifies the following definition.

Definition 3.2. A feasible point $\bar{x} \in M$ of (MPCC) is called Mordukhovich stationary (M-stationary for short) for (MPCC) provided there exist multipliers $\lambda \in \mathcal{Y}^\star$, $\mu \in \mathcal{Z}^\star$, and $\nu \in \mathcal{Z}$ which satisfy (3.2a), (3.2b), and

$$(\mu, \nu) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(G(\bar{x}), H(\bar{x})). \tag{3.8}$$

We obtain the following necessary optimality conditions of M-stationarity-type.

Proposition 3.6. Let $\bar{x} \in M$ be a local optimal solution of (MPCC) where the constraint qualification

$$\left.\begin{aligned} &0 = g'(\bar{x})^\star[\lambda] + G'(\bar{x})^\star[\mu] + H'(\bar{x})^\star[\nu], \\ &\lambda \in \mathcal{N}_C(g(\bar{x})), \\ &(\mu, \nu) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(G(\bar{x}), H(\bar{x})) \end{aligned}\right\} \implies \lambda = 0,\ \mu = 0,\ \nu = 0 \tag{3.9}$$

is satisfied and ψ is continuously Fréchet differentiable, and assume that $\mathcal{X}$ and $\mathcal{Y}$ are reflexive. Then any of the conditions stated below is sufficient for $\bar{x}$ to be M-stationary for (MPCC):

1. $C \times \operatorname{gph} \mathcal{N}_K$ is SNC at $(g(\bar{x}), G(\bar{x}), H(\bar{x}))$,
2. C is SNC at $g(\bar{x})$ and $\mathcal{Z}$ is finite-dimensional,
3. C is SNC at $g(\bar{x})$ and the operator
$$\begin{bmatrix} G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix}$$
is surjective,
4. $g'(\bar{x})$ is surjective and $\operatorname{gph} \mathcal{N}_K$ is SNC at $(G(\bar{x}), H(\bar{x}))$.

Proof. Under the first condition, the assertion follows combining Lemmas 2.29 and 2.38. Note that the second condition implies the first one. For the proof of the assertion under the third condition, we introduce

$$M_1 := g^{-1}(C), \qquad M_2 := \{x \in \mathcal{X} \mid (G(x), H(x)) \in \operatorname{gph} \mathcal{N}_K\}.$$

Obviously, we have $M = M_1 \cap M_2$. Exploiting the given assumptions, the constraint qualification (3.9) implies

$$\mathcal{N}_{M_1}(\bar{x}) \subseteq g'(\bar{x})^\star[\mathcal{N}_C(g(\bar{x}))], \qquad \mathcal{N}_{M_2}(\bar{x}) = \{G'(\bar{x})^\star[\mu] + H'(\bar{x})^\star[\nu] \in \mathcal{X}^\star \mid (\mu, \nu) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(G(\bar{x}), H(\bar{x}))\},$$

see Lemma 2.38. Moreover, M_1 is SNC at $\bar{x}$, see [90, Theorem 3.84]. Exploiting the constraint qualification (3.9) once more, from Lemma 2.18 we obtain $\mathcal{N}_M(\bar{x}) \subseteq \mathcal{N}_{M_1}(\bar{x}) + \mathcal{N}_{M_2}(\bar{x})$. Thus, the assertion follows from Lemma 2.29. The proof of the lemma's assertion under the fourth set of conditions is analogous. □

Without a specific setting, the definition of M-stationarity and the above necessary optimality conditions are rarely applicable since they comprise the implicitly given limiting normal cone to the graph of the normal cone mapping induced by K which is usually difficult to compute. A general ready-to-use formula for this variational object does not exist. However, if K equals $\mathbb{R}_0^{m,+}$, $\mathcal{S}_m^+$, or $\mathcal{K}_m$, such formulae are at hand, see [45], [37], or [134], respectively. In all these cases, $\mathcal{Z}$ is finite-dimensional, and if $\mathcal{Y}$ is finite-dimensional as well, then $C \times \operatorname{gph} \mathcal{N}_K$ is SNC everywhere and Proposition 3.6 applies for the derivation

of necessary optimality conditions.
From the theory of standard MPCCs we expect the relations

$$\text{S-stationarity} \implies \text{M-stationarity} \implies \text{W-stationarity} \tag{3.10}$$

between the three stationarity concepts introduced above, see [129]. In the following, we are going to study whether these relations hold in the situation, where K is polyhedric. As remarked earlier, MPCCs with polyhedric complementarity cone are of special importance. We need the following lemma in order to analyze this situation. These results were already presented in [124] for the case where $\mathcal{Z}$ is a Hilbert space. Here we provide a slight generalization to the situation where $\mathcal{Z}$ is only reflexive.

Lemma 3.7. Let $(z, z^*) \in \operatorname{gph} \mathcal{N}_K$ be chosen such that K is polyhedric w.r.t. (z, z^*). Then the following calculus rules hold:

$$\begin{aligned} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) &= \{(d, d^*) \in \mathcal{K}_K(z, z^*) \times \mathcal{K}_{K^\circ}(z^*, z) \mid \langle d^*, d \rangle_{\mathcal{Z}} = 0\}, \\ \operatorname{conv} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) &= \mathcal{K}_K(z, z^*) \times \mathcal{K}_{K^\circ}(z^*, z). \end{aligned}$$

Proof. From [78, Theorem 3.1] we obtain

$$\mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) = \{(d, d^*) \in \mathcal{K}_K(z, z^*) \times \mathcal{K}_K(z, z^*)^\circ \mid \langle d^*, d \rangle_{\mathcal{Z}} = 0\},$$

and $\mathcal{K}_K(z, z^*)^\circ = \mathcal{K}_{K^\circ}(z^*, z)$ holds due to the polyhedricity of K w.r.t. (z, z^*). For the proof of the second formula, we first remark that the inclusion $\subseteq$ is clear from the first part of the proof. For the validation of the converse inclusion, choose $d \in \mathcal{K}_K(z, z^*)$ and $d^* \in \mathcal{K}_{K^\circ}(z^*, z)$ arbitrarily. Then, obviously, $(2d, 0), (0, 2d^*) \in \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)$ is satisfied due to the first formula. Thus, we obtain the relation $(d, d^*) = \frac{1}{2}(2d, 0) + \frac{1}{2}(0, 2d^*) \in \operatorname{conv} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)$ which completes the proof. □

Below, we present another result we need for the subsequent proofs.

Lemma 3.8. Let $(z, z^*) \in \operatorname{gph} \mathcal{N}_K$ be arbitrarily chosen. Then the following inclusions hold:

$$K \cap \big(-\mathcal{T}_K(z)\big) \subseteq K \cap \{z^*\}^\perp, \quad K^\circ \cap \big(-\mathcal{T}_{K^\circ}(z^*)\big) \subseteq K^\circ \cap \{z\}^\perp.$$

Proof. We only show the first inclusion since the validation of the second one can be done in a similar way.
Choose $d \in K \cap \big(-\mathcal{T}_K(z)\big)$. Then we find sequences $\{z_k\} \subseteq K$ and $\{t_k\} \subseteq \mathbb{R}$ such that $z_k \to z$, $t_k \searrow 0$, and $\frac{1}{t_k}(z_k - z) \to -d \in -K$ hold true. From $\langle z^*, z \rangle_{\mathcal{Z}} = 0$ we obtain

$$0 \leq \left\langle z^*, -\frac{z_k}{t_k} \right\rangle_{\mathcal{Z}} = \left\langle z^*, -\frac{z_k - z}{t_k} \right\rangle_{\mathcal{Z}},$$

and taking the limit $k \to \infty$ yields $0 \leq \langle z^*, d \rangle_{\mathcal{Z}}$. On the other hand, from $d \in K$ we have $\langle z^*, d \rangle_{\mathcal{Z}} \leq 0$. Hence, $d \in \{z^*\}^\perp$ follows and the proof is completed. □

Now, we can start to consider the relations between the three introduced stationarity notions.

Lemma 3.9. Let $(z, z^*) \in \operatorname{gph} \mathcal{N}_K$ be chosen such that K is polyhedric w.r.t. (z, z^*). Then we have

$$\widehat{\mathcal{N}}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) = \mathcal{K}_{K^\circ}(z^*, z) \times \mathcal{K}_K(z, z^*).$$

Proof. First, let us show the inclusions

$$\mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) \subseteq \mathcal{T}^w_{\operatorname{gph} \mathcal{N}_K}(z, z^*) \subseteq \operatorname{conv} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*). \tag{3.11}$$

The first one is clear from the definition of these tangent cones. Choose $(d, d^*) \in \mathcal{T}^w_{\operatorname{gph} \mathcal{N}_K}(z, z^*)$ arbitrarily. Then we find sequences $\{(z_k, z_k^*)\} \subseteq \operatorname{gph} \mathcal{N}_K$ and $\{t_k\} \subseteq \mathbb{R}$ satisfying $z_k \to z$, $z_k^* \to z^*$, $t_k \searrow 0$, $\frac{1}{t_k}(z_k - z) \rightharpoonup d$, and $\frac{1}{t_k}(z_k^* - z^*) \rightharpoonup d^*$. Since we have $\frac{1}{t_k}(z_k - z) \in \mathcal{R}_K(z)$ and $\frac{1}{t_k}(z_k^* - z^*) \in \mathcal{R}_{K^\circ}(z^*)$

for any $k \in \mathbb{N}$, we deduce $d \in \mathcal{T}_K(z)$ and $d^* \in \mathcal{T}_{K^\circ}(z^*)$ from the closedness and convexity of these cones. For any $w \in K \cap \{z^*\}^\perp$,

$$\left\langle \frac{z_k^* - z^*}{t_k}, w - z_k \right\rangle_{\mathcal{Z}} \leq \left\langle -\frac{z^*}{t_k}, w - z_k \right\rangle_{\mathcal{Z}} = \left\langle -\frac{z^*}{t_k}, z - z_k \right\rangle_{\mathcal{Z}} = \left\langle z^*, \frac{z_k - z}{t_k} \right\rangle_{\mathcal{Z}} \leq 0$$

holds, and taking the limit $k \to \infty$ yields

$$\langle d^*, w - z \rangle_{\mathcal{Z}} \leq \langle z^*, d \rangle_{\mathcal{Z}} \leq 0,$$

see Lemma 2.4. Using $w := z$, we see $d \in \{z^*\}^\perp$ and, hence, $d \in \mathcal{K}_K(z, z^*)$. On the other hand, testing with $w = 0$ and $w = 2z$ yields $d^* \in \{z\}^\perp$ which leads to $d^* \in \mathcal{K}_{K^\circ}(z^*, z)$. Now, we apply Lemma 3.7 and obtain $(d, d^*) \in \operatorname{conv} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)$.
Let us polarize (3.11). Then we obtain

$$\mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)^\circ \supseteq \widehat{\mathcal{N}}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) \supseteq \big(\operatorname{conv} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)\big)^\circ = \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)^\circ$$

since $\mathcal{Z} \times \mathcal{Z}^\star$ is reflexive. Consequently,

$$\widehat{\mathcal{N}}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) = \big(\operatorname{conv} \mathcal{T}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)\big)^\circ = \mathcal{K}_K(z, z^*)^\circ \times \mathcal{K}_{K^\circ}(z^*, z)^\circ = \mathcal{K}_{K^\circ}(z^*, z) \times \mathcal{K}_K(z, z^*)$$

follows from Lemma 3.7 and the polyhedricity of K w.r.t. (z, z^*). □

In the case where the cone K is even polyhedral, the above assertion was already provided in [60, Proposition 3.2]. Obviously, Lemma 3.9 shows that any S-stationary point of (MPCC) is also M-stationary provided the cone K is polyhedric w.r.t. the reference point.
Below, we study whether any M-stationary point of (MPCC) is also W-stationary in the situation where K is a polyhedric cone such that at least one of the pairs $(\mathcal{Z}, \leq_K)$ and $(\mathcal{Z}^\star, \leq_{K^\circ})$ is a vector lattice.

Lemma 3.10. Let $(z, z^*) \in \operatorname{gph} \mathcal{N}_K$ be fixed. Assume that K is a polyhedric cone and let the following two conditions hold:

(i) One of the following properties is valid:

(ia) K° is pointed, $(\mathcal{Z}^\star, \leq_{K^\circ})$ is a vector lattice, and the mapping $\mathcal{Z}^\star \ni \tilde{z}^* \mapsto \max_{K^\circ}\{\tilde{z}^*; 0\} \in \mathcal{Z}^\star$ is continuous,

(ib) K is pointed, $(\mathcal{Z}, \leq_K)$ is a vector lattice, and the mapping $\mathcal{Z} \ni \tilde{z} \mapsto \max_K\{\tilde{z}; 0\} \in \mathcal{Z}$ is weakly-weakly continuous.

(ii) One of the following properties is valid:

(iia) K is pointed, $(\mathcal{Z}, \leq_K)$ is a vector lattice, and the mapping $\mathcal{Z} \ni \tilde{z} \mapsto \max_K\{\tilde{z}; 0\} \in \mathcal{Z}$ is continuous,

(iib) K° is pointed, $(\mathcal{Z}^\star, \leq_{K^\circ})$ is a vector lattice, and the mapping $\mathcal{Z}^\star \ni \tilde{z}^\star \mapsto \max_{K^\circ}\{\tilde{z}^*; 0\} \in \mathcal{Z}^\star$ is weakly-weakly continuous.

Then we have

$$\mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(z, z^*) \subseteq \Big[\operatorname{cl}(K^\circ - K^\circ \cap \{z\}^\perp) \cap \{z\}^\perp\Big] \times \Big[\operatorname{cl}(K - K \cap \{z^*\}^\perp) \cap \{z^*\}^\perp\Big].$$

Proof. Let $(\eta^*, \eta) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(z, z^*)$ be arbitrarily chosen. Then we find sequences $\{(z_k, z_k^*)\} \subseteq \operatorname{gph} \mathcal{N}_K$ and $\{(\eta_k^*, \eta_k)\} \subseteq \mathcal{Z}^\star \times \mathcal{Z}$ which satisfy $z_k \to z$, $z_k^* \to z^*$, $\eta_k^* \rightharpoonup \eta^*$, $\eta_k \rightharpoonup \eta$, and $(\eta_k^*, \eta_k) \in \widehat{\mathcal{N}}_{\operatorname{gph} \mathcal{N}_K}(z_k, z_k^*)$ for all $k \in \mathbb{N}$. Exploiting Lemma 3.9, the latter condition is equivalent to $\eta_k^* \in \mathcal{K}_{K^\circ}(z_k^*, z_k)$ as well as $\eta_k \in \mathcal{K}_K(z_k, z_k^*)$ for all $k \in \mathbb{N}$. Hence, we obtain $\eta_k^* \in \{z_k\}^\perp$ and $\eta_k \in \{z_k^*\}^\perp$ for any $k \in \mathbb{N}$. Applying Lemma 2.4, $\eta^* \in \{z\}^\perp$ and $\eta \in \{z^*\}^\perp$ follow. Thus, we only need to show

$$\eta^* \in \operatorname{cl}\big(K^\circ - K^\circ \cap \{z\}^\perp\big), \quad \eta \in \operatorname{cl}(K - K \cap \{z^*\}^\perp)$$

in order to complete the proof.
Here we only prove that condition (i) implies $\eta \in \mathrm{cl}(K - K \cap \{z^*\}^\perp)$. Similarly, one can proceed in order to see that condition (ii) implies $\eta^* \in \mathrm{cl}\big(K^\circ - K^\circ \cap \{z\}^\perp\big)$.
First, let us assume that the condition (ia) holds. Here we need some appropriate limiting operators for sequences of sets, cf. [6, 108]. Thus, for a sequence $\{U_k\} \subseteq 2^{\mathcal{U}}$ in the Banach spaces $\mathcal{U}$, we set

$$\begin{aligned}\limsup_{k\to\infty}{}^w U_k &:= \{u \in \mathcal{U} \,|\, \exists \{U_{k_l}\} \subseteq \{U_k\}\, \forall l \in \mathbb{N}\, \exists u_{k_l} \in U_{k_l} \colon\ u_{k_l} \rightharpoonup u\},\\ \liminf_{k\to\infty} U_k &:= \{u \in \mathcal{U} \,|\, \exists k_0 \in \mathbb{N}\, \forall k \geq k_0\, \exists u_k \in U_k \colon\ u_k \to u\}.\end{aligned}$$

These sets are called the weak sequential upper and (strong) sequential lower limit of $\{U_k\}$, respectively, see [6, Section 1.1]. First, note that we have

$$\eta \in \limsup_{k\to\infty}{}^w \mathcal{T}_K(z_k).$$

For later use, we need to verify the following formula:

$$K^\circ \cap \big(-\mathcal{T}_{K^\circ}(z^*)\big) \subseteq \liminf_{k\to\infty} K^\circ \cap \big(-\mathcal{T}_{K^\circ}(z_k^*)\big). \tag{3.12}$$

Take $\xi^* \in K^\circ \cap \big(-\mathcal{T}_{K^\circ}(z^*)\big)$. Due to [6, Definition 4.1.7, Theorem 4.2.2], we find $\{\xi_k^*\} \subseteq \mathcal{Z}^\star$ satisfying $\xi_k^* \in -\mathcal{T}_{K^\circ}(z_k^*)$ for all $k \in \mathbb{N}$ and $\xi_k^* \to \xi^*$. Taking the supremum w.r.t. K°, we obtain the relation $\max_{K^\circ}\{\xi_k^*; 0\} \in -\mathcal{T}_{K^\circ}(z_k^*)$ for all $k \in \mathbb{N}$, see Lemma 2.23. By definition of the supremum operator, $\max_{K^\circ}\{\xi_k^*; 0\} \in K^\circ$ is satisfied for any $k \in \mathbb{N}$ as well. From the continuity assumption on the supremum operator we obtain $\max_{K^\circ}\{\xi_k^*; 0\} \to \max_{K^\circ}\{\xi^*; 0\}$. Since $\xi^* \in K^\circ$ is satisfied, $\max_{K^\circ}\{\xi^*; 0\} = \xi^*$ holds. This shows the formula in (3.12).
Using the bipolar theorem, see Lemma 2.11, [6, Theorem 1.1.8], Lemma 3.8, (3.12), and Lemma 2.12, we derive the following chain of inclusions:

$$\begin{aligned}\limsup_{k\to\infty}{}^w \mathcal{T}_K(z_k) &= \limsup_{k\to\infty}{}^w \mathcal{T}_K(z_k)^{\circ\circ} \subseteq \overline{\mathrm{conv}} \limsup_{k\to\infty}{}^w \mathcal{T}_K(z_k)^{\circ\circ}\\ &= \left(\liminf_{k\to\infty} \mathcal{T}_K(z_k)^\circ\right)^\circ = \left(\liminf_{k\to\infty} K^\circ \cap \{z_k\}^\perp\right)^\circ\\ &\subseteq \left(\liminf_{k\to\infty} K^\circ \cap \big(-\mathcal{T}_{K^\circ}(z_k^*)\big)\right)^\circ \subseteq \big(K^\circ \cap \big(-\mathcal{T}_{K^\circ}(z^*)\big)\big)^\circ\\ &= \mathrm{cl}\big(K - K \cap \{z^*\}^\perp\big).\end{aligned}$$

This shows $\eta \in \mathrm{cl}\big(K - K \cap \{z^*\}^\perp\big)$ under condition (ia).
Now, we assume that condition (ib) is satisfied. Recall that for all $k \in \mathbb{N}$, we have $\eta_k \in \mathcal{T}_K(z_k)$. Furthermore, $\eta_k \rightharpoonup \eta$ holds. For any $k \in \mathbb{N}$, we decompose η_k into $\eta_k^+ := \max_K\{\eta_k; 0\}$ and $\eta_k^- := \min_K\{\eta_k; 0\}$. Due to the assumption, the sequences $\{\eta_k^+\}$ and $\{\eta_k^-\}$ converge weakly to $\max_K\{\eta; 0\}$ and $\min_K\{\eta; 0\}$, respectively. Especially, $\eta_k^+ + \eta_k^- = \eta_k \rightharpoonup \eta = \max_K\{\eta; 0\} + \min_K\{\eta; 0\}$ holds. Due to the definition of supremum and infimum, we obtain $\eta_k^+ \in K$ and $\eta_k^- \in -K$ for all $k \in \mathbb{N}$. Furthermore, for all $k \in \mathbb{N}$, $\eta_k^- \in \mathcal{T}_K(z_k)$ follows from Lemma 2.23. We invoke Lemma 3.8 in order to obtain $\eta_k^- \in (-K) \cap \{z_k^*\}^\perp$. Since $\{\eta_k^-\}$ converges weakly, whereas $\{z_k^*\}$ converges strongly, we obtain $\min_K\{\eta; 0\} \in (-K) \cap \{z^*\}^\perp$ from Lemma 2.4. The relation $\max_K\{\eta; 0\} \in K$ follows from the convexity and closedness of K. Combining the above observations, $\eta \in K - K \cap \{z^*\}^\perp \subseteq \mathrm{cl}\big(K - K \cap \{z^*\}^\perp\big)$ is obtained.
Summing up these results, condition (i) implies $\eta \in \mathrm{cl}\big(K - K \cap \{z^*\}^\perp\big)$. □

Combining Lemmas 3.9 and 3.10 with the definitions of S-, M-, and W-stationarity, we obtain the following result.

Proposition 3.11. Consider (MPCC) where the cone K is polyhedric and choose an arbitrary feasible point $\bar{x} \in M$. Then the implication

$$\text{S-stationarity} \implies \text{M-stationarity}$$

holds for $\bar{x}$. If, additionally, the conditions (i) as well as (ii) from Lemma 3.10 are satisfied, then the implications in (3.10) are valid for $\bar{x}$.

Let us discuss the above result by means of the following two examples.

Example 3.12. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete as well as σ-finite measure space and fix $p \in (1, \infty)$ as well as the corresponding conjugate coefficient $p' \in (1, \infty)$. We consider the reflexive Banach space $\mathcal{Z} = L^p(\mathfrak{M})$ and the closed, convex, pointed cone

$$L^p(\mathfrak{M})_0^+ = \{u \in L^p(\mathfrak{M}) \,|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\}.$$

As we remarked in Example 2.25, $L^p(\mathfrak{M})_0^+$ induces a vector lattice in $\mathcal{Z}$ and the corresponding supremum operator $\max_{L^p(\mathfrak{M})_0^+}$ is continuous. On the other hand,

$$\left(L^p(\mathfrak{M})_0^+\right)^\circ = \left\{\eta \in L^{p'}(\mathfrak{M}) \,\middle|\, \eta(\omega) \leq 0 \text{ f.a.a. } \omega \in \Omega\right\}$$

is easily seen. Repeating the same reasoning as used to discuss $L^p(\mathfrak{M})_0^+$, we obtain that its polar cone induces a vector lattice in $\mathcal{Z}^\star = L^{p'}(\mathfrak{M})$ whose corresponding supremum operator is continuous. Thus, the conditions (i) and (ii) of Lemma 3.10 hold since (ia) and (iia) are valid. Note that $L^p(\mathfrak{M})_0^+$ is polyhedric by [17, Theorem 3.58]. Thus, the relations in (3.10) are valid. A detailed discussion about MPCCs whose complementarity cone equals $L^p(\mathfrak{M})_0^+$ is presented in Section 3.2. ■

Example 3.13. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain. We take a closer look at the reflexive Banach space $H_0^1(\Omega)$ and consider the closed, convex, pointed cone

$$H_0^1(\Omega)_0^+ = \left\{u \in H_0^1(\Omega) \,\middle|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\right\}.$$

This cone is polyhedric by means of [17, Corollary 6.46] and induces a vector lattice in $H_0^1(\Omega)$. The corresponding supremum operator is continuous, see Example 2.25, and weakly-weakly continuous, see [125, Lemma 4.1]. Consequently, the conditions (i) and (ii) of Lemma 3.10 are valid since (ib) and (iia) are satisfied and the relations in (3.10) hold. Explicit representations of the W- and S-stationarity conditions of an MPCC whose complementarity cone is given by $H_0^1(\Omega)_0^+$ can be derived from the results in [123]. In the case $d = 1$, the M-stationarity conditions are discussed in [73]. In [125], some M-stationarity-type conditions for the obstacle problem are derived via a regularization approach for domains with dimension $d \geq 2$. Whether these conditions equal the M-stationarity conditions from Definition 3.2 has to be clarified in the future.
Note that $\left(H^{-1}(\Omega), \leq_{(H_0^1(\Omega)_0^+)^\circ}\right)$ is no vector lattice: in [122, Appendix B], the author shows that the negative part of a measure $\mu \in H^{-1}(\Omega) \cap \mathcal{M}(\Omega)$, i.e. the possible maximum of μ and 0 w.r.t. $(H_0^1(\Omega)_0^+)^\circ$, does not necessarily need to be an element of $H^{-1}(\Omega)$ again. Thus, the conditions (ia) and (iib) from Lemma 3.10 cannot be satisfied. ■

In the following, we want to consider two important types of MPCCs with polyhedric complementarity cone in more detail: mathematical problems with complementarity constraints in Lebesgue spaces and optimization problems with polyhedral complementarity constraints.

3.2. Complementarity programming in Lebesgue spaces

Here we study the problem (MPCC) where $\mathcal{Z} := L^p(\mathfrak{M}, \mathbb{R}^m)$ holds and the closed, convex, pointed cone K is given as stated below:

$$K := \{u \in L^p(\mathfrak{M}, \mathbb{R}^m) \,|\, u(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\}. \tag{3.13}$$

Therein, $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ is a complete, σ-finite, as well as nonatomic measure space such that $L^q(\mathfrak{M})$ is separable for all $q \in [1, \infty)$. We fix $p \in (1, \infty)$ and denote by $p' \in (1, \infty)$ the corresponding conjugate coefficient. Moreover, we set $I := \{1, \ldots, m\}$. Clearly, K is a decomposable set which satisfies Assumption 2.1. That is why for any $\bar{u} \in K$, we obtain

$$\begin{aligned}\mathcal{N}_K(\bar{u}) &= \left\{\eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \in \mathcal{N}_{\mathbb{R}_0^{m,+}}(\bar{u}(\omega)) \text{ f.a.a. } \omega \in \Omega\right\} \\ &= \left\{\eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \leq 0,\ \eta(\omega) \cdot \bar{u}(\omega) = 0 \text{ f.a.a. } \omega \in \Omega\right\}\end{aligned}$$

from Propositions 2.49 and 2.51. Moreover, K is polyhedric, see [17, Theorem 3.58]. We introduce a closed set $\Xi \subseteq \mathbb{R}^m \times \mathbb{R}^m$ by

$$\Xi := \{(a,b) \in \mathbb{R}^m \times \mathbb{R}^m \,|\, a \geq 0,\, b \leq 0,\, a \cdot b = 0\}.$$

Clearly, Ξ can be represented as the finite union of convex sets and, thus, is derivable, see Lemma 2.15. Taking these observations together,

$$\operatorname{gph} \mathcal{N}_K = \left\{(u,\eta) \in L^p(\mathfrak{M},\mathbb{R}^m) \times L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, (u(\omega),\eta(\omega)) \in \Xi \text{ f.a.a. } \omega \in \Omega\right\}$$

holds true. Moreover, the decomposable set $\operatorname{gph} \mathcal{N}_K$ satisfies Assumption 2.1. In the following theorem, we present explicit representations of the introduced stationarity conditions of (MPCC) where K is given as in (3.13). We already know from Proposition 3.11 that the implications in (3.10) hold, see Example 3.12.

Theorem 3.14. Let $\bar{x} \in M$ be a feasible point of (MPCC) where the cone K is given as in (3.13). Then the following assertions hold:

1. The point $\bar{x}$ is W-stationary if and only if there are $\lambda \in \mathcal{Y}^\star$ as well as functions $\mu \in L^{p'}(\mathfrak{M},\mathbb{R}^m)$ and $\nu \in L^p(\mathfrak{M},\mathbb{R}^m)$ which satisfy the conditions (3.2a), (3.2b), and
$$\begin{aligned} &\forall i \in I\colon \quad \mu_i(\omega) = 0 \quad \text{f.a.a. } \omega \in I^{+0}(\bar{x},i), \\ &\forall i \in I\colon \quad \nu_i(\omega) = 0 \quad \text{f.a.a. } \omega \in I^{0-}(\bar{x},i). \end{aligned} \tag{3.14}$$

2. The point $\bar{x}$ is M-stationary if and only if it is W-stationary.

3. The point $\bar{x}$ is S-stationary if and only if there are $\lambda \in \mathcal{Y}^\star$ as well as functions $\mu \in L^{p'}(\mathfrak{M},\mathbb{R}^m)$ and $\nu \in L^p(\mathfrak{M},\mathbb{R}^m)$ which satisfy the conditions (3.2a), (3.2b), (3.14), and
$$\forall i \in I\colon \quad \mu_i(\omega) \leq 0,\, \nu_i(\omega) \geq 0 \quad \text{f.a.a. } \omega \in I^{00}(\bar{x},i). \tag{3.15}$$

Therein, for any $i \in I$, the measurable sets $I^{+0}(\bar{x},i)$, $I^{0-}(\bar{x},i)$, and $I^{00}(\bar{x},i)$ are defined as stated below:

$$\begin{aligned} I^{+0}(\bar{x},i) &:= \{\omega \in \Omega \,|\, G(\bar{x})_i(\omega) > 0,\, H(\bar{x})_i(\omega) = 0\}, \\ I^{0-}(\bar{x},i) &:= \{\omega \in \Omega \,|\, G(\bar{x})_i(\omega) = 0,\, H(\bar{x})_i(\omega) < 0\}, \\ I^{00}(\bar{x},i) &:= \{\omega \in \Omega \,|\, G(\bar{x})_i(\omega) = 0,\, H(\bar{x})_i(\omega) = 0\}. \end{aligned}$$

Proof. For brevity, we set $z := G(\bar{x})$ and $z^* := H(\bar{x})$.
Let us start with the proof of the first assertion. We need to show

$$\begin{aligned} \operatorname{cl}(K^\circ - K^\circ \cap \{z\}^\perp) \cap \{z\}^\perp &= \left\{\mu \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \forall i \in I\colon\ \mu_i(\omega) = 0 \text{ f.a.a. } \omega \in I^{+0}(\bar{x},i)\right\}, \\ \operatorname{cl}(K - K \cap \{z^*\}^\perp) \cap \{z^*\}^\perp &= \left\{\nu \in L^{p}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \forall i \in I\colon\ \nu_i(\omega) = 0 \text{ f.a.a. } \omega \in I^{0-}(\bar{x},i)\right\}. \end{aligned}$$

Here we verify only the first equation, the second one follows in a similar way. Since K is a decomposable set which satisfies Assumption 2.1, we obtain

$$\begin{aligned} K^\circ \cap \{z\}^\perp = \mathcal{N}_K(z) &= \left\{\mu \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \mu(\omega) \leq 0,\, \mu(\omega) \cdot z(\omega) = 0 \text{ f.a.a. } \omega \in \Omega\right\} \\ &= \left\{\mu \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \forall i \in I\colon \begin{array}{ll} \mu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x},i) \\ \mu_i(\omega) \leq 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x},i) \cup I^{00}(\bar{x},i) \end{array}\right\}. \end{aligned}$$

Since we have

$$K^\circ = \left\{\eta \in L^{p'}(\mathfrak{M},\mathbb{R}^m) \,\middle|\, \eta(\omega) \leq 0 \text{ f.a.a. } \omega \in \Omega\right\}$$

from Lemma 2.43, we easily obtain

$$K^\circ - K^\circ \cap \{z\}^\perp = \left\{ \mu \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \forall i \in I\colon\ \mu_i(\omega) \leq 0 \text{ f.a.a. } \omega \in I^{+0}(\bar{x}, i) \right\}.$$

Due to the closedness of this set, the first assertion follows in a straightforward way.
Let us derive the second assertion. Therefore, we apply Proposition 2.51 to obtain

$$\begin{aligned} &\mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(z, z^*) \subseteq \mathcal{N}^c_{\operatorname{gph}\mathcal{N}_K}(z, z^*) \\ &\quad = \left\{ (\mu, \nu) \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \times L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, (\mu(\omega), \nu(\omega)) \in \mathcal{N}^c_\Xi(z(\omega), z^*(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}. \end{aligned} \tag{3.16}$$

Let us introduce $\widetilde{\Xi} \subseteq \mathbb{R}^m \times \mathbb{R}^m$ by

$$\widetilde{\Xi} := \{(a, b) \in \mathbb{R}^m \times \mathbb{R}^m \,|\, a \geq 0,\ b \geq 0,\ a \cdot b = 0\}.$$

Then we clearly have

$$\Xi = \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & -\mathbf{I}_m \end{bmatrix} [\widetilde{\Xi}]$$

which yields

$$\mathcal{N}_\Xi(\bar{a}, \bar{b}) = \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & -\mathbf{I}_m \end{bmatrix} \left[\mathcal{N}_{\widetilde{\Xi}}(\bar{a}, -\bar{b})\right] \tag{3.17}$$

for any $(\bar{a}, \bar{b}) \in \Xi$, see Lemma 2.38. Using e.g. [45, Proposition 2.4], we obtain

$$\mathcal{N}_{\widetilde{\Xi}}(\bar{a}, -\bar{b}) = \left\{ (\eta, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m \,\middle|\, \begin{array}{ll} \eta_i = 0 & \text{if } i \in J^{+0}(\bar{a}, \bar{b}) \\ \zeta_i = 0 & \text{if } i \in J^{0-}(\bar{a}, \bar{b}) \\ (\eta_i < 0 \wedge \zeta_i < 0) \vee \eta_i \zeta_i = 0 & \text{if } i \in J^{00}(\bar{a}, \bar{b}) \end{array} \right\} \tag{3.18}$$

where the index sets $J^{+0}(\bar{a}, \bar{b})$, $J^{0-}(\bar{a}, \bar{b})$, and $J^{00}(\bar{a}, \bar{b})$ are defined as stated below:

$$\begin{aligned} J^{+0}(\bar{a}, \bar{b}) &:= \{i \in I \,|\, \bar{a}_i > 0,\ \bar{b}_i = 0\}, \\ J^{0-}(\bar{a}, \bar{b}) &:= \{i \in I \,|\, \bar{a}_i = 0,\ \bar{b}_i < 0\}, \\ J^{00}(\bar{a}, \bar{b}) &:= \{i \in I \,|\, \bar{a}_i = 0,\ \bar{b}_i = 0\}. \end{aligned}$$

Now, we easily see

$$\begin{aligned} \mathcal{N}^c_\Xi(\bar{a}, \bar{b}) = \overline{\operatorname{conv}}\, \mathcal{N}_\Xi(\bar{a}, \bar{b}) &= \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & -\mathbf{I}_m \end{bmatrix} \left[\overline{\operatorname{conv}}\, \mathcal{N}_{\widetilde{\Xi}}(\bar{a}, -\bar{b})\right] \\ &= \left\{ (\eta, \zeta) \in \mathbb{R}^m \times \mathbb{R}^m \,\middle|\, \begin{array}{ll} \eta_i = 0 & \text{if } i \in J^{+0}(\bar{a}, \bar{b}) \\ \zeta_i = 0 & \text{if } i \in J^{0-}(\bar{a}, \bar{b}) \end{array} \right\}. \end{aligned}$$

Hence, (3.16) yields that any pair $(\mu, \nu) \in \mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(z, z^*)$ satisfies the conditions in (3.14). Note that this result is also a consequence of Example 3.12.
Now, choose $(\tilde{\mu}, \tilde{\nu}) \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \times L^p(\mathfrak{M}, \mathbb{R}^m)$ which satisfy (3.14). Combining (3.17), (3.18), and Proposition 2.49, we obtain

$$\begin{aligned} &\mathcal{N}^s_{\operatorname{gph}\mathcal{N}_K}(z, z^*) \\ &\quad = \left\{ (\mu, \nu) \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \times L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \begin{array}{lll} & \mu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \\ \forall i \in I\colon & \nu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \\ & (\mu_i(\omega), \nu_i(\omega)) \in \Theta & \text{f.a.a. } \omega \in I^{00}(\bar{x}, i) \end{array} \right\} \end{aligned}$$

where $\Theta \subseteq \mathbb{R}^2$ is the set defined below:

$$\Theta := \{(\alpha, \beta) \in \mathbb{R}^2 \,|\, (\alpha < 0 \wedge \beta > 0) \vee \alpha\beta = 0\}.$$

Let us introduce functions $\mu^1, \mu^2 \in L^{p'}(\mathfrak{M}, \mathbb{R}^m)$ and $\nu^1, \nu^2 \in L^p(\mathfrak{M}, \mathbb{R}^m)$ by $\mu^1 := 2\tilde{\mu}$, $\mu^2 := 0$, $\nu^1 := 0$, and $\nu^2 := 2\tilde{\nu}$. Then it is easy to see from the above representation of the strong limiting normal cone that $(\mu^1, \nu^1), (\mu^2, \nu^2) \in \mathcal{N}^s_{\operatorname{gph}\mathcal{N}_K}(z, z^*)$ holds. Consequently,

$$(\tilde{\mu}, \tilde{\nu}) = \tfrac{1}{2}(\mu^1, \nu^1) + \tfrac{1}{2}(\mu^2, \nu^2) \in \operatorname{conv} \mathcal{N}^s_{\operatorname{gph}\mathcal{N}_K}(z, z^*) \subseteq \mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(z, z^*)$$

is obtained from Proposition 2.51. Summing up the above arguments, we have

$$\mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(z, z^*) = \left\{ (\mu, \nu) \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \times L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \forall i \in I\colon \begin{array}{ll} \mu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \\ \nu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \end{array} \right\}$$

which yields the second assertion of the theorem.
The proof of the third assertion reduces to the validation of

$$\begin{aligned} \mathcal{K}_{K^\circ}(z^*, z) &= \left\{ \mu \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \forall i \in I\colon \begin{array}{ll} \mu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \\ \mu_i(\omega) \leq 0 & \text{f.a.a. } \omega \in I^{00}(\bar{x}, i) \end{array} \right\}, \\ \mathcal{K}_K(z, z^*) &= \left\{ \nu \in L^p(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \forall i \in I\colon \begin{array}{ll} \nu_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \\ \nu_i(\omega) \geq 0 & \text{f.a.a. } \omega \in I^{00}(\bar{x}, i) \end{array} \right\}. \end{aligned} \tag{3.19}$$

We only show the first formula since the proof of the second one is analogous. Applying the results from Section 2.3.5, we easily obtain

$$\mathcal{K}_{K^\circ}(z^*, z) = \left\{ \mu \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \,\middle|\, \eta(\omega) \in \mathcal{K}_{-\mathbb{R}^{m,+}_0}(z^*(\omega), z(\omega)) \text{ f.a.a. } \omega \in \Omega \right\}.$$

Using the above notation, it is not difficult to see

$$\begin{aligned} \mathcal{K}_{-\mathbb{R}^{m,+}_0}(\bar{b}, \bar{a}) = \mathcal{T}_{-\mathbb{R}^{m,+}_0}(\bar{b}) \cap \{\bar{a}\}^\perp &= \{\eta \in \mathbb{R}^m \mid \eta_i \leq 0 \text{ if } i \in J^{+0}(\bar{a}, \bar{b}) \cup J^{00}(\bar{a}, \bar{b})\} \cap \{\bar{a}\}^\perp \\ &= \left\{ \eta \in \mathbb{R}^m \,\middle|\, \begin{array}{ll} \eta_i = 0 & \text{if } i \in J^{+0}(\bar{a}, \bar{b}) \\ \eta_i \leq 0 & \text{if } i \in J^{00}(\bar{a}, \bar{b}) \end{array} \right\}. \end{aligned}$$

Thus, the desired result follows once more from pointwise evaluation. □

The above theorem shows that the implications in (3.10) can be strengthened to

$$\text{S-stationarity} \implies \text{M-stationarity} \iff \text{W-stationarity}$$

in the case where the complementarity cone is given as presented in (3.13). Thus, M-stationarity is uncomfortably weak in this case. This result should be alarming in the sense that not all results from the theory of finite-dimensional complementarity programming can be generalized one-to-one to more abstract cases. It is a question of future research whether similar problems appear if the complementarity is induced in Sobolev spaces or other reflexive function spaces.
Note that our stationarity concepts are essentially different from those ones derived in [56]. Especially, our multiplier functions μ and ν possess a much higher degree of regularity. Furthermore, we do not need two sets of multipliers in order to define M- and S-stationarity, see [56, Definition 3.1].

Lemma 3.15. Let $\bar{x} \in M$ be a feasible point for (MPCC) where the complementarity cone K is given as presented in (3.13). Then $\operatorname{gph}\mathcal{N}_K$ is not SNC at $(G(\bar{x}), H(\bar{x}))$.

Proof. First, suppose that $I^{00}(\bar{x}, 1)$ is a set of positive measure. Since $\mathfrak{M}$ is nonatomic, we find a sequence $\{\Omega_k\} \subseteq \Sigma$ of measurable subsets of $I^{00}(\bar{x}, 1)$ such that $\mathfrak{m}(\Omega_k) \searrow 0$ holds true, see [16, Corollary 1.12.10]. For any $k \in \mathbb{N}$, we define

$$\forall \omega \in \Omega\colon \quad \mu_{k,1}(\omega) := -\mathfrak{m}(\Omega_k)^{-\frac{1}{p'}} \chi_{\Omega_k}(\omega), \quad \nu_{k,1}(\omega) := 0$$

and

$$\forall i \in \{2, \ldots, m\}\, \forall \omega \in \Omega\colon \quad \mu_{k,i}(\omega) = \nu_{k,i}(\omega) = 0.$$

From (3.19) and Lemma 3.9 we have $(\mu_k, \nu_k) \in \widehat{\mathcal{N}}_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}), H(\bar{x}))$ for all $k \in \mathbb{N}$. Choose an arbitrary function $u \in L^p(\mathfrak{M}, \mathbb{R}^m)$ and observe that $u_1 \in L^p(\mathfrak{M}|_{\Omega_k})$ holds for all $k \in \mathbb{N}$. We apply Hölder's inequality in order to obtain

$$\begin{aligned}\left|\langle \mu_{k,1}, u_1\rangle_{L^p(\mathfrak{M})}\right| &= \mathfrak{m}(\Omega_k)^{-\frac{1}{p'}} \left|\int_{\Omega_k} u_1(\omega)\mathrm{d}\mathfrak{m}\right| \\ &\leq \mathfrak{m}(\Omega_k)^{-\frac{1}{p'}} \|\chi_{\Omega_k}\|_{L^{p'}(\mathfrak{M}|_{\Omega_k})} \|u_1\|_{L^p(\mathfrak{M}|_{\Omega_k})} = \left(\int_{\Omega_k} |u_1(\omega)|^p \mathrm{d}\mathfrak{m}\right)^{\frac{1}{p}}.\end{aligned}$$

Lemma A.1 shows that the latter integral tends to zero. Thus, $\mu_k \rightharpoonup 0$ is satisfied. However, from $\|\mu_k\|_{L^{p'}(\mathfrak{M},\mathbb{R}^m)} = 1$ for any $k \in \mathbb{N}$, $\{\mu_k\}$ does not converge strongly to 0. This leads to $(\mu_k, \nu_k) \rightharpoonup (0,0)$ and $(\mu_k, \nu_k) \nrightarrow (0,0)$. Hence, $\operatorname{gph}\mathcal{N}_K$ cannot be SNC at $(G(\bar{x}), H(\bar{x}))$ in this case.
Now, suppose that $I^{00}(\bar{x}, 1)$ is of measure zero (w.l.o.g. we assume $I^{00}(\bar{x}, 1) = \varnothing$). Then one of the sets $I^{+0}(\bar{x}, 1)$ or $I^{0-}(\bar{x}, 1)$ possesses positive measure. Let us assume $\mathfrak{m}(I^{+0}(\bar{x}, 1)) > 0$. Again, we find a sequence $\{\Omega_k\} \subseteq \Sigma$ of measurable subsetes of $I^{+0}(\bar{x}, 1)$ satisfying $\mathfrak{m}(\Omega_k) \searrow 0$. For any $k \in \mathbb{N}$, we define

$$\forall \omega \in \Omega\colon \quad u_{k,1}(\omega) := \begin{cases}(1 - \chi_{\Omega_k}(\omega))G(\bar{x})_1(\omega) & \text{if } \omega \in I^{+0}(\bar{x}, 1)\\ 0 & \text{if } \omega \in I^{0-}(\bar{x}, 1)\end{cases}$$

as well as

$$\forall i \in \{2, \ldots, m\}\, \forall \omega \in \Omega\colon \quad u_{k,i}(\omega) := G(\bar{x})_i(\omega).$$

From Lemma A.1 we obtain $u_k \to G(\bar{x})$ in $L^p(\mathfrak{M}, \mathbb{R}^m)$. Let us set

$$\forall \omega \in \Omega\colon \quad \mu_{k,1}(\omega) := -\mathfrak{m}(\Omega_k)^{-\frac{1}{p'}} \chi_{\Omega_k}(\omega), \quad \nu_{k,1}(\omega) := 0$$

and

$$\forall i \in \{2, \ldots, m\}\, \forall \omega \in \Omega\colon \quad \mu_{k,i}(\omega) = \nu_{k,i}(\omega) := 0$$

for any $k \in \mathbb{N}$. Then from (3.19) and Lemma 3.9 we have $(\mu_k, \nu_k) \in \widehat{\mathcal{N}}_{\operatorname{gph}\mathcal{N}_K}(u_k, H(\bar{x}))$ for any $k \in \mathbb{N}$. Similar as above, we show $\mu_k \rightharpoonup 0$ and $\mu_k \nrightarrow 0$ which implies that $\operatorname{gph}\mathcal{N}_K$ is not SNC at $(G(\bar{x}), H(\bar{x}))$ in this case as well. Finally, an analogous argumentation shows the lack of the SNC property whenever $\mathfrak{m}(I^{0-}(\bar{x}, 1)) > 0$ holds. This completes the proof. □

Remark 3.16. Due to Theorem 3.14, one may use Proposition 3.6 in order formulate constraint qualifications which ensure a local optimal solution of (MPCC) where K is given as in (3.13) to be W-stationary. However, due to Lemma 3.15, the set C needs to possess the SNC property in this case. In view of Lemma 2.19 and Corollary 2.20, and Lemma 2.21, this property fails for many reasonable function spaces and, thus, the consideration of optimal control problems with complementarity constraints on the control via Mordukhovich's approach does not seem to be advisable. In the absence of the constraints $g(x) \in C$, the only applicable constraint qualification from Proposition 3.6 already implies S-stationarity of local solutions, see Proposition 3.4. Thus, we focus on Proposition 3.4 for necessary optimality conditions and constraint qualifications.

The following result is a direct consequence of Proposition 3.4.

Proposition 3.17. Let $\bar{x} \in M$ be a local solution of (MPCC) where K is given as in (3.13). We set

$$S := \left\{ (u, \eta) \in L^p(\mathfrak{M}, \mathbb{R}^m) \times L^{p'}(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \begin{array}{ll} \exists \alpha \geq 0\, \exists \beta \geq 0\, \forall i \in I\colon & \\ \quad u_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \cup I^{00}(\bar{x}, i)\\ \quad u_i(\omega) + \alpha G(\bar{x})_i(\omega) \geq 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i)\\ \quad \eta_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \cup I^{00}(\bar{x}, i)\\ \quad \eta_i(\omega) + \beta H(\bar{x})_i(\omega) \leq 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \end{array} \right\}$$

and

$$T := \left\{ (u, \eta) \in L^p(\mathfrak{M}, \mathbb{R}^m) \times L^{p'}(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \forall i \in I\colon \begin{array}{ll} u_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \cup I^{00}(\bar{x}, i)\\ \eta_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \end{array} \right\}.$$

Suppose that the constraint qualification

$$\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} [\mathcal{X}] - \begin{pmatrix} \mathcal{R}_C(g(\bar{x})) \\ S \end{pmatrix} = \begin{pmatrix} \mathcal{Y} \\ L^p(\mathfrak{M}, \mathbb{R}^m) \\ L^{p'}(\mathfrak{M}, \mathbb{R}^m) \end{pmatrix}$$

is satisfied. Then $\bar{x}$ is W-stationary (and, thus, M-stationary). If, additionally, the constraint qualification

$$\operatorname{cl}\left(\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} [\mathcal{X}] - \begin{pmatrix} \mathcal{N}_C(g(\bar{x}))_\perp \\ T \end{pmatrix} \right) = \begin{pmatrix} \mathcal{Y} \\ L^p(\mathfrak{M}, \mathbb{R}^m) \\ L^{p'}(\mathfrak{M}, \mathbb{R}^m) \end{pmatrix}$$

holds, then $\bar{x}$ is S-stationary.

Proof. Due to Proposition 3.4, we only need to show

$$S = \Big(\mathcal{R}_K(G(\bar{x})) \cap \big(-\mathcal{K}_K(G(\bar{x}), H(\bar{x})) \big) \Big) \times \Big(\mathcal{R}_{K^\circ}(H(\bar{x})) \cap \big(-\mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x})) \big) \Big)$$

and

$$T = \Big(\mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x}))^{\circ\perp} \Big) \times \Big(\mathcal{T}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(H(\bar{x})))}(H(\bar{x}))^{\circ\perp} \Big).$$

The formula for S follows from

$$\mathcal{R}_K(G(\bar{x})) = \left\{ u \in L^p(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \begin{array}{ll} \exists \alpha \geq 0 \, \forall i \in I\colon & \\ \quad u_i(\omega) \geq 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \\ \quad u_i(\omega) + \alpha G(\bar{x})_i(\omega) \geq 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \end{array} \right\},$$

$$\mathcal{R}_{K^\circ}(H(\bar{x})) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \begin{array}{ll} \exists \beta \geq 0 \, \forall i \in I\colon & \\ \quad \eta_i(\omega) \leq 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \\ \quad \eta_i(\omega) + \beta H(\bar{x})_i(\omega) \leq 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \end{array} \right\},$$

and (3.19). On the other hand, the relations

$$K \cap \big(-\mathcal{T}_K(G(\bar{x})) \big) = \left\{ u \in L^p(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \forall i \in I\colon \begin{array}{ll} u_i(\omega) \geq 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \\ u_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \end{array} \right\},$$

$$K^\circ \cap \big(-\mathcal{T}_{K^\circ}(H(\bar{x})) \big) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \forall i \in I\colon \begin{array}{ll} \eta_i(\omega) \leq 0 & \text{f.a.a. } \omega \in I^{0-}(\bar{x}, i) \\ \eta_i(\omega) = 0 & \text{f.a.a. } \omega \in I^{+0}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \end{array} \right\}$$

lead to

$$\mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x})) = \left\{ u \in L^p(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \forall i \in I\colon \; u_i(\omega) = 0 \text{ f.a.a. } \omega \in I^{0-}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \right\},$$

$$\mathcal{T}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(H(\bar{x})))}(H(\bar{x})) = \left\{ \eta \in L^{p'}(\mathfrak{M}, \mathbb{R}^m) \;\middle|\; \forall i \in I\colon \; \eta_i(\omega) = 0 \text{ f.a.a. } \omega \in I^{+0}(\bar{x}, i) \cup I^{00}(\bar{x}, i) \right\}.$$

This shows the formula for T since the above tangent cones are already closed subspaces of $L^p(\mathfrak{M}, \mathbb{R}^m)$ and $L^{p'}(\mathfrak{M}, \mathbb{R}^m)$, respectively. □

3.3. Complementarity programming with polyhedral cones

Let $\{z_1^*, \ldots, z_m^*\} \subseteq \mathcal{Z}^\star$ be a set of linear independent functionals where the reflexive Banach space $\mathcal{Z}$ is arbitrarily chosen. Here we consider the polyhedral cone

$$K := \{ z \in \mathcal{Z} \,|\, \forall i \in \{1, \ldots, m\}\colon \; \langle z_i^*, z \rangle_{\mathcal{Z}} \leq 0 \} \tag{3.20}$$

inducing the complementarity in (MPCC). Note that we cannot apply Proposition 3.11 in order to derive the relations (3.10) directly since from the above assumptions, it is not clear whether or not K induces a vector lattice in $\mathcal{Z}$. Consider, e.g., $\mathcal{Z} = \mathbb{R}^3$ and the polyhedral cone

$$K := \{z \in \mathbb{R}^3 \,|\, (1,-1,0)\cdot z \leq 0,\ (-1,-1,0)\cdot z \leq 0\}$$

whose generating elements $z_1^* = (1,-1,0)$ and $z_2^* = (-1,-1,0)$ are linearly independent. Then K is not pointed and, consequently, the corresponding binary relation $\leq_K$ is not antisymmetric. Hence, K does not induce a lattice structure in $\mathbb{R}^3$. On the other hand, the cone

$$K^\circ = \{z^* \in \mathbb{R}^3 \,|\, (-1,1,0)\cdot z^* \leq 0,\ (1,1,0)\cdot z^* \leq 0,\ (0,0,1)\cdot z^* \leq 0,\ (0,0,-1)\cdot z^* \leq 0\}$$

does not induce a lattice structure in $\mathbb{R}^3$ as well since, e.g., the set $\{(0,0,0),(0,0,1)\}$ does not possess an upper bound w.r.t. $\leq_{K^\circ}$. Thus, we cannot apply Lemma 3.10 to the situation at hand.
Following Lemma 2.13, we obtain

$$K^\circ = \operatorname{cone}\{z_1^*,\ldots,z_m^*\}.$$

Fix a feasible point $\bar{x} \in M$ of (MPCC) where K is given as in (3.20). Then we find unique scalars $\alpha_1,\ldots,\alpha_m \geq 0$ such that $H(\bar{x}) = \sum_{i=1}^m \alpha_i z_i^*$ is satisfied. Thus, it makes sense to define index sets $I(\bar{x})$ and $J(\bar{x})$ as stated below:

$$I(\bar{x}) := \{i \in \{1,\ldots,m\} \,|\, \langle z_i^*, G(\bar{x})\rangle_{\mathcal{Z}} = 0\}, \qquad J(\bar{x}) := \{i \in I(\bar{x}) \,|\, \alpha_i > 0\}.$$

Stipulating $\operatorname{lin}\varnothing = \operatorname{cone}\varnothing = \{0\}$, it is easily seen that

$$\begin{aligned}\mathcal{R}_K(G(\bar{x})) &= \{d \in \mathcal{Z} \,|\, \forall i \in I(\bar{x}):\ \langle z_i^*, d\rangle_{\mathcal{Z}} \leq 0\} = \mathcal{T}_K(G(\bar{x})),\\ \mathcal{R}_{K^\circ}(H(\bar{x})) &= \operatorname{lin}\{z_i^* \,|\, i \in J(\bar{x})\} + \operatorname{cone}\{z_i^* \,|\, i \in \{1,\ldots,m\}\setminus J(\bar{x})\} = \mathcal{T}_{K^\circ}(H(\bar{x}))\end{aligned}$$

hold true since the radial cones are both closed, see Lemma 2.13. Especially, K is polyhedric w.r.t. $(G(\bar{x}),H(\bar{x}))$. Moreover, we easily obtain

$$\begin{aligned}\mathcal{K}_K(G(\bar{x}),H(\bar{x})) &= \left\{ d \in \mathcal{Z} \;\middle|\; \begin{array}{ll} \forall i \in I(\bar{x})\setminus J(\bar{x}): & \langle z_i^*, d\rangle_{\mathcal{Z}} \leq 0 \\ \forall i \in J(\bar{x}): & \langle z_i^*, d\rangle_{\mathcal{Z}} = 0 \end{array} \right\},\\ \mathcal{K}_{K^\circ}(H(\bar{x}),G(\bar{x})) &= \operatorname{lin}\{z_i^* \,|\, i \in J(\bar{x})\} + \operatorname{cone}\{z_i^* \,|\, i \in I(\bar{x})\setminus J(\bar{x})\}.\end{aligned} \tag{3.21}$$

For $J(\bar{x}) \subseteq P \subseteq Q \subseteq I(\bar{x})$, we define

$$\begin{aligned}C_{Q,P} &:= \operatorname{lin}\{z_i^* \,|\, i \in P\} + \operatorname{cone}\{z_i^* \,|\, i \in Q\setminus P\},\\ D_{Q,P} &:= \left\{ d \in \mathcal{Z} \;\middle|\; \begin{array}{ll} \forall i \in Q\setminus P: & \langle z_i^*, d\rangle_{\mathcal{Z}} \leq 0 \\ \forall i \in P: & \langle z_i^*, d\rangle_{\mathcal{Z}} = 0 \end{array} \right\}.\end{aligned}$$

Note that we have $C_{I(\bar{x}),J(\bar{x})} = \mathcal{K}_{K^\circ}(H(\bar{x}),G(\bar{x}))$ and $D_{I(\bar{x}),J(\bar{x})} = \mathcal{K}_K(G(\bar{x}),H(\bar{x}))$ from (3.21). We obtain the following result which was partially presented in [60] and [62] .

Lemma 3.18. Let $\bar{x} \in M$ be feasible for (MPCC) where K is given as in (3.20). Exploiting the above notations, we obtain

$$\begin{aligned}\widehat{\mathcal{N}}_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x})) &= C_{I(\bar{x}),J(\bar{x})} \times D_{I(\bar{x}),J(\bar{x})},\\ \mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x})) &= \bigcup_{J(\bar{x})\subseteq P\subseteq Q\subseteq I(\bar{x})} C_{Q,P} \times D_{Q,P},\\ \mathcal{N}^c_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x})) &= C_{I(\bar{x}),I(\bar{x})} \times D_{J(\bar{x}),J(\bar{x})}.\end{aligned}$$

Proof. The formula for the Fréchet normal cone follows from Lemma 3.9. The second assertion precisely equals [60, Theorem 4.2]. Thus, we only need to prove the formula for Clarke's normal cone.
Note that for any index sets $J(\bar{x}) \subseteq P \subseteq Q \subseteq I(\bar{x})$, we have $C_{Q,P} \subseteq C_{I(\bar{x}),I(\bar{x})}$ and $D_{Q,P} \subseteq D_{J(\bar{x}),J(\bar{x})}$. Since $\mathcal{Z}$ is reflexive, we obtain

$$\begin{aligned}\mathcal{N}^c_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x})) &= \overline{\operatorname{conv}}\,\mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x}))\\ &\subseteq \overline{\operatorname{conv}}\big(C_{I(\bar{x}),I(\bar{x})} \times D_{J(\bar{x}),J(\bar{x})}\big) = C_{I(\bar{x}),I(\bar{x})} \times D_{J(\bar{x}),J(\bar{x})}\end{aligned}$$

from the second formula of this lemma. Thus, the inclusion $\subseteq$ holds. For the proof of the converse inclusion, choose $(\mu,\nu) \in C_{I(\bar{x}),I(\bar{x})} \times D_{J(\bar{x}),J(\bar{x})}$ arbitrarily. Then $(2\mu,0) \in C_{I(\bar{x}),I(\bar{x})} \times D_{I(\bar{x}),I(\bar{x})}$ and $(0,2\nu) \in C_{J(\bar{x}),J(\bar{x})} \times D_{J(\bar{x}),J(\bar{x})}$ are obvious. Hence, we can conclude

$$\begin{aligned}(\mu,\nu) &\in \operatorname{conv}\Big(\big(C_{I(\bar{x}),I(\bar{x})} \times D_{I(\bar{x}),I(\bar{x})}\big) \cup \big(C_{J(\bar{x}),J(\bar{x})} \times D_{J(\bar{x}),J(\bar{x})}\big)\Big)\\ &\subseteq \overline{\operatorname{conv}}\,\mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x})) = \mathcal{N}^c_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}),H(\bar{x}))\end{aligned}$$

from the second formula of this lemma and the reflexivity of $\mathcal{Z}$. This completes the proof. □

Another important observation we report in the subsequent lemma.

Lemma 3.19. Let $\bar{x} \in M$ be a feasible point of (MPCC) where K is given as in (3.20). Exploiting the above notations, we obtain

$$\begin{aligned}\operatorname{cl}(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp &= C_{I(\bar{x}),I(\bar{x})},\\ \operatorname{cl}(K - K \cap \{H(\bar{x})\}^\perp) \cap \{H(\bar{x})\}^\perp &= D_{J(\bar{x}),J(\bar{x})}.\end{aligned}$$

Proof. We start with the proof of the first equation. Observe that $K^\circ \cap \{G(\bar{x})\}^\perp = \operatorname{cone}\{z_i^* \,|\, i \in I(\bar{x})\}$ holds. This yields

$$K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp = \operatorname{lin}\{z_i^* \,|\, i \in I(\bar{x})\} + \operatorname{cone}\{z_i^* \,|\, i \in \{1,\dots,m\} \setminus I(\bar{x})\}$$

and the latter set is closed due to Lemma 2.13. Thus, we obtain

$$\operatorname{cl}(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp = (K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp = \operatorname{lin}\{z_i^* \,|\, i \in I(\bar{x})\} = C_{I(\bar{x}),I(\bar{x})},$$

i.e. the first assertion of the lemma is valid.
In order to prove the second statement, we show both inclusions separately. First, observe that

$$K \cap \{H(\bar{x})\}^\perp = \left\{ d \in \mathcal{Z} \;\middle|\; \begin{array}{ll} \forall i \in \{1,\dots,m\} \setminus J(\bar{x})\colon & \langle z_i^*, d\rangle_{\mathcal{Z}} \le 0 \\ \forall i \in J(\bar{x})\colon & \langle z_i^*, d\rangle_{\mathcal{Z}} = 0 \end{array} \right\}$$

is satisfied. This yields

$$K - K \cap \{H(\bar{x})\}^\perp \subseteq \{d \in \mathcal{Z} \,|\, \forall i \in J(\bar{x})\colon\ \langle z_i^*, d\rangle_{\mathcal{Z}} \le 0\}$$

and, since the set on the right is closed,

$$\begin{aligned}\operatorname{cl}(K - K \cap \{H(\bar{x})\}^\perp) \cap \{H(\bar{x})\}^\perp &\subseteq \{d \in \mathcal{Z} \,|\, \forall i \in J(\bar{x})\colon\ \langle z_i^*, d\rangle_{\mathcal{Z}} \le 0\} \cap \{H(\bar{x})\}^\perp\\ &= \{z_i^* \,|\, i \in J(\bar{x})\}^\perp = D_{J(\bar{x}),J(\bar{x})}\end{aligned}$$

where we put $\varnothing^\perp = \mathcal{Z}$ if necessary. This shows the inclusion $\subseteq$. On the other hand, the set $\{z_i^* \,|\, i \in J(\bar{x})\}^\perp$ is a subset of $\{H(\bar{x})\}^\perp$. Furthermore, the above representation of $K \cap \{H(\bar{x})\}^\perp$ enables us to deduce

$$\begin{aligned}D_{J(\bar{x}),J(\bar{x})} = \{z_i^* \,|\, i \in J(\bar{x})\}^\perp &\subseteq K \cap \{H(\bar{x})\}^\perp - K \cap \{H(\bar{x})\}^\perp\\ &\subseteq K - K \cap \{H(\bar{x})\}^\perp \subseteq \operatorname{cl}(K - K \cap \{H(\bar{x})\}^\perp).\end{aligned}$$

Thus, the inclusion $\supseteq$ holds as well and the proof is completed. □

Combining Lemmas 3.18 and 3.19, we are able to state the W-, M-, and S-stationarity conditions of MPCCs whose complementarity cone is given as stated in (3.20).

Theorem 3.20. Let $\bar{x} \in M$ be a feasible point of (MPCC) where the cone K is given as in (3.20). Then the following assertions hold:

1. The point $\bar{x}$ is W-stationary if and only if there are $\lambda \in \mathcal{Y}^\star$, $\mu_1, \ldots, \mu_m \in \mathbb{R}$, and $\nu \in \mathcal{Z}$ which satisfy the conditions (3.2b) and

$$\begin{aligned} &0 = \psi'(\bar{x}) + g'(\bar{x})^\star[\lambda] + \textstyle\sum_{i=1}^m \mu_i G'(\bar{x})^\star[z_i^*] + H'(\bar{x})^\star[\nu], \\ &\forall i \in \{1, \ldots, m\} \setminus I(\bar{x}) \colon\ \mu_i = 0, \\ &\forall i \in J(\bar{x}) \colon\ \langle z_i^*, \nu \rangle_{\mathcal{Z}} = 0. \end{aligned} \tag{3.22}$$

2. The point $\bar{x}$ is M-stationary if and only if there are $\lambda \in \mathcal{Y}^\star$, $\mu_1, \ldots, \mu_m \in \mathbb{R}$, and $\nu \in \mathcal{Z}$ which satisfy the conditions (3.2b), (3.22), and

$$\forall i \in I(\bar{x}) \setminus J(\bar{x}) \colon\ (\mu_i > 0 \wedge \langle z_i^*, \nu \rangle_{\mathcal{Z}} < 0) \vee \langle \mu_i z_i^*, \nu \rangle_{\mathcal{Z}} = 0. \tag{3.23}$$

3. The point $\bar{x}$ is S-stationary if and only if there are $\lambda \in \mathcal{Y}^\star$, $\mu_1, \ldots, \mu_m \in \mathbb{R}$, and $\nu \in \mathcal{Z}$ which satisfy the conditions (3.2b), (3.22), and

$$\forall i \in I(\bar{x}) \setminus J(\bar{x}) \colon\ \mu_i \geq 0,\ \langle z_i^*, \nu \rangle_{\mathcal{Z}} \leq 0. \tag{3.24}$$

Proof. We only need to comment on the M-stationarity conditions since the other representations are clear from Lemmas 3.18 and 3.19. Therefore, choose $(\mu, \nu) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(G(\bar{x}), H(\bar{x}))$. By Lemma 3.18 there exist index sets $J(\bar{x}) \subseteq P \subseteq Q \subseteq I(\bar{x})$ such that $\mu \in C_{Q,P}$ and $\nu \in D_{Q,P}$ hold. Hence, there are $\mu_1, \ldots, \mu_m \in \mathbb{R}$ satisfying $\mu_i = 0$ for all $i \in \{1, \ldots, m\} \setminus Q$, $\mu_i \geq 0$ for all $i \in Q \setminus P$, and $\mu = \sum_{i=1}^m \mu_i z_i^*$. Especially, $\mu_i = 0$ holds for all $i \in \{1, \ldots, m\} \setminus I(\bar{x})$. On the other hand, ν satisfies $\langle z_i^*, \nu \rangle_{\mathcal{Z}} \leq 0$ for all $i \in Q \setminus P$ and $\langle z_i^*, \nu \rangle_{\mathcal{Z}} = 0$ for all $i \in P$. Particularly, $\langle z_i^*, \nu \rangle_{\mathcal{Z}} = 0$ for $i \in J(\bar{x})$ follows. Observe that $I(\bar{x}) \setminus J(\bar{x}) = (P \setminus J(\bar{x})) \cup (Q \setminus P) \cup (I(\bar{x}) \setminus Q)$ holds.
For any $i \in P \setminus J(\bar{x})$, we have $\langle z_i^*, \nu \rangle_{\mathcal{Z}} = 0$ and, thus, $\langle \mu_i z_i^*, \nu \rangle_{\mathcal{Z}} = 0$. Similarly, for $i \in I(\bar{x}) \setminus Q$, we obtain $\mu_i = 0$ which leads to $\langle \mu_i z_i^*, \nu \rangle_{\mathcal{Z}} = 0$. Finally, choose $i \in Q \setminus P$. Then we have $\mu_i \geq 0$ and $\langle z_i^*, \nu \rangle_{\mathcal{Z}} \leq 0$. If $\mu_i = 0$ or $\langle z_i^*, \nu \rangle_{\mathcal{Z}} = 0$ holds true, then $\langle \mu_i z_i^*, \nu \rangle_{\mathcal{Z}} = 0$ follows again. Otherwise, we have $\mu_i > 0$ and $\langle z_i^*, \nu \rangle_{\mathcal{Z}} < 0$. This completes the proof. □

Due to the above theorem, we have the strict relations (3.10) between the introduced stationarity notions although Proposition 3.11 is not generally applicable here.
Now, necessary optimality conditions of the corresponding MPCC may be derived from Propositions 3.4 and 3.6. Note that for any feasible point $\bar{x} \in M$ of (MPCC) where K is given as in (3.20), the constraint qualification (3.4) takes the form

$$\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} [\mathcal{X}] - \begin{pmatrix} \mathcal{R}_C(g(\bar{x})) \\ \{z_i^* \mid i \in I(\bar{x})\}^\perp \\ \operatorname{lin}\{z_i^* \mid i \in J(\bar{x})\} \end{pmatrix} = \begin{pmatrix} \mathcal{Y} \\ \mathcal{Z} \\ \mathcal{Z}^\star \end{pmatrix}, \tag{3.25}$$

whereas the constraint qualification (3.5) becomes

$$\operatorname{cl} \left(\begin{bmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{bmatrix} [\mathcal{X}] - \begin{pmatrix} \mathcal{N}_C(g(\bar{x}))_\perp \\ \{z_i^* \mid i \in I(\bar{x})\}^\perp \\ \operatorname{lin}\{z_i^* \mid i \in J(\bar{x})\} \end{pmatrix} \right) = \begin{pmatrix} \mathcal{Y} \\ \mathcal{Z} \\ \mathcal{Z}^\star \end{pmatrix}.$$

Assume that the constraint qualification (3.25) holds. Polarizing this equation, we obtain

$$\left.\begin{aligned} &0 = g'(\bar{x})^\star[\lambda] + \textstyle\sum_{i=1}^m \mu_i G'(\bar{x})^\star[z_i^*] + H'(\bar{x})^\star[\nu], \\ &\lambda \in \mathcal{N}_C(g(\bar{x})), \\ &\forall i \in \{1, \ldots, m\} \setminus I(\bar{x}) \colon\ \mu_i = 0, \\ &\forall i \in J(\bar{x}) \colon\ \langle z_i^*, \nu \rangle_{\mathcal{Z}} = 0 \end{aligned}\right\} \implies \lambda = 0,\ \mu_1 = \cdots = \mu_m = 0,\ \nu = 0$$

which is stronger than the constraint qualification (3.9), see Lemma 3.18 for the characterization of the limiting normal cone to the complementarity set. Thus, under some additional SNC assumptions, (3.25) is already sufficient for M-stationarity of local optimal solutions of MPCCs whose complementarity cone K is polyhedral, see Proposition 3.6 as well.

Corollary 3.21. Let $\bar{x} \in M$ be a local optimal solution of (MPCC) where ψ is continuously Fréchet differentiable. Let the cone K be given as in (3.20) and assume that $\mathcal{X}$ as well as $\mathcal{Y}$ are reflexive. Suppose that (3.25) is satisfied and that one of the additional SNC conditions from Proposition 3.6 holds. Then $\bar{x}$ is M-stationary.

It is easily seen that the general setting in this section covers standard MPCCs with $\mathcal{X} = \mathbb{R}^n$, $\mathcal{Y} = \mathbb{R}^q$, $\mathcal{Z} = \mathbb{R}^m$, and $K = \mathbb{R}^{m,+}_0$. Indeed, if $\mathbf{e}^1, \ldots, \mathbf{e}^m \in \mathbb{R}^m$ denote the m unit vectors of $\mathbb{R}^m$, then we have

$$\mathbb{R}^{m,+}_0 = \{z \in \mathbb{R}^m \,|\, \forall i \in \{1, \ldots, m\} \colon\ (-\mathbf{e}^i) \cdot z \leq 0\}.$$

One may check that the stationarity notions characterized in Theorem 3.20 equal the well-known ones, see [129]. As we mentioned in Remark 3.5, (3.25) equals MPCC-MFCQ in this situation which already implies M-stationarity of local optimal solutions by means of [42]. We obtained the same result in more general form in Corollary 3.21.
In [10], the authors discuss a conic linear problem governed by the nonpolyhedric second-order cone $\mathcal{K}_m$. They used a polyhedral approximation of $\mathcal{K}_m$ in order to simplify the original problem. It was shown that this approximation is reasonably good under mild assumptions. Transferring this idea to (MPCC) with complementarity constraints governed by $\mathcal{K}_m$, one could think of approximating the nonpolyhedric complementarity problem by a polyhedral one and applying the results obtained above to study the surrogate problem. How this can be done explicitly and how the two problems behave is, however, beyond the scope of this thesis and left as a topic of future research.

3.4. Additional remarks on complementarity programming

We want to close this chapter on complementarity programming with some additional remarks.
Firstly, consider (MPCC) where K is given as in (3.13) or (3.20) and choose a feasible point $\bar{x} \in M$ of it. Observe that from the proof of Theorem 3.14 and Lemmas 3.18 as well as 3.19, we see that the set of multipliers (μ, ν) satisfying (3.2c) and (3.2d) equals $\mathcal{N}^c_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}), H(\bar{x}))$. Thus, the question arises whether we have

$$\mathcal{N}^c_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}), H(\bar{x})) = \Big(\operatorname{cl}(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp\Big) \times \Big(\operatorname{cl}(K - K \cap \{H(\bar{x})\}^\perp) \cap \{H(\bar{x})\}^\perp\Big)$$

in general or at least in the situation where K is polyhedric w.r.t. $(G(\bar{x}), H(\bar{x}))$. Currently, there is no proof available for this result since we still suffer from a lack of knowledge on a general representation of the limiting normal cone $\mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}), H(\bar{x}))$. On the other hand, it might be possible to derive the formula

$$\mathcal{T}^c_{\operatorname{gph}\mathcal{N}_K}(G(\bar{x}), H(\bar{x})) = \mathcal{T}_{K \cap (-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x})) \times \mathcal{T}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(H(\bar{x})))}(H(\bar{x})),$$

which is precisely the corresponding dual statement (see proof of Proposition 3.4), directly.
Secondly, it is a question of future research whether M-stationarity of a feasible point of (MPCC) implies its W-stationarity in general provided the complementarity cone K is polyhedric (or even without the polyhedricity assumption). In the proof of Lemma 3.10, the condition that K induces a vector lattice in $\mathcal{Z}$ or that K° induces a vector lattice in $\mathcal{Z}^\star$ is indispensable. On the other hand, we saw in the context of MPCCs whose complementarity cone is polyhedral that this property is not necessary in some cases.
We will see later, see Remark 5.11, that in the setting $\mathcal{Z} = \mathcal{S}_p$ and $K = \mathcal{S}^+_p$, the implications (3.10) do not hold since M- and S-stationarity (in the sense of Definitions 3.2 and 3.1, respectively) do not imply each other. Furthermore, this example shows that our generalized notions of stationarity do not always coincide with the well-known stationarity concepts from the literature, see [37]. Thus, in the absence of polyhedricity, the situation is far more complicated and has to be analyzed carefully for the different complementarity cones.
Since we know from Example 3.13 that the common relationship between W-, M-, and S-stationarity, see (3.10), is valid for $\mathcal{Z} = H^1_0(\Omega)$ and the corresponding cone of almost everywhere nonnegative functions $H^1_0(\Omega)^+_0$, it would be interesting to compute the limiting normal cone to the complementarity set $\operatorname{gph}\mathcal{N}_{H^1_0(\Omega)^+_0}$ in order to find an explicit representation of the M-stationarity conditions stated in Definition 3.2. However, this is a challenging task which requires a deep knowledge of Sobolev spaces and capacity

theory, see [17, Section 6.4.3], which is clearly beyond the scope of this thesis but a promising object of future research.
Finally, we want to mention that other stationarity concepts for generalized MPCCs exist in literature apart from W-, M-, and S-stationarity e.g. Clarke's stationarity concept, see [37, 66, 79, 129]. A possible approach on how to derive the so-called C-stationarity conditions is described below. Assume that $\mathcal{Z}$ is a Hilbert space which is identified with its dual by means of Riesz's representation theorem. Then we have

$$(G(x), H(x)) \in \operatorname{gph} \mathcal{N}_K \iff H(x) \in \mathcal{N}_K(G(x)) \iff G(x) = \operatorname{proj}_K(G(x) + H(x))$$

from Example 2.30. Thus, we can restate (MPCC) equivalently by

$$\begin{aligned} \psi(x) &\to \min \\ g(x) &\in C \\ G(x) - \operatorname{proj}_K(G(x) + H(x)) &= 0 \end{aligned}$$

which is an optimization problem with a single nonsmooth constraint. Applying Mordukhovich's tools of generalized differentiation, it is possible to derive the M-stationarity conditions of (MPCC) under appropriate constraint qualifications via this problem. On the other hand, since the projection operator $\operatorname{proj}_K(\cdot)$ possesses certain Lipschitz properties, it is reasonable to apply the concept of Clarke's generalized Jacobian, see [24, Section 2.6], in order to exploit some abstract differential information to derive first order optimality conditions of (MPCC). This has been done for $K = \mathcal{S}^+_m$ in [37] and for $K = \mathcal{K}_m$ in [79]. The procedure is possible for standard MPCCs as well. This way, the C-stationarity conditions can be deduced. However, the definition of Clarke's generalized Jacobian heavily relies on Rademacher's theorem which only applies in our setting if $\mathcal{X}$ and $\mathcal{Z}$ are finite-dimensional. Thus, this approach is limited to the finite-dimensional situation which we do not presume here.

4. Bilevel programming in Banach spaces

In this section, we take a closer look at the general bilevel programming model given by

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ G(x) &\in C \\ y &\in \Psi(x) \end{aligned} \tag{BPP}$$

where $\Psi\colon \mathcal{X} \rightrightarrows \mathcal{Y}$ is the solution set mapping of the parametric optimization problem

$$\begin{aligned} f(x,y) &\to \min_{y} \\ g(x,y) &\in K. \end{aligned} \tag{4.1}$$

The following general assumptions on (BPP) shall hold.

Assumption 4.1. The mapping $F\colon \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$ is Fréchet differentiable while the mappings $f\colon \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$, $G\colon \mathcal{X} \to \mathcal{W}$, and $g\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ are continuously Fréchet differentiable. Therein, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{W}$, and $\mathcal{Z}$ are Banach spaces. Furthermore, the sets $C \subseteq \mathcal{W}$ and $K \subseteq \mathcal{Z}$ are nonempty, closed, and convex.

Let us briefly explain the decision order in (BPP): First, $x \in X_{\mathsf{ad}} := \{x \in \mathcal{X} \,|\, G(x) \in C\}$ has to be chosen (so-called upper level decision). Afterwards, the set $\Psi(x)$ is computed (lower level decision) and the overall objective functional F can be evaluated for the pairs (x,y) where $y \in \Psi(x)$ holds. Thus, in contrast to the original bilevel optimization model, see [115], where F is only minimized w.r.t. x, our model (BPP) is well-defined even if the lower level solution is not unique for some x satisfying the upper level constraints. In [26, Theorem 5.2], one can find a criterion which ensures the existence of a global optimal solution of (BPP) in the finite-dimensional situation. The proof of this result heavily relies on the classical Weierstraß theorem. However, we cannot transfer this proof to the general situation since the set $\operatorname{gph}\Psi$ which is part of the constraints of (BPP) is generally nonconvex and, thus, the generalized Weierstraß theorem, see Lemma 2.5, is rarely applicable. Actually, we have to discuss the existence of global optimal solutions of (BPP) for the specific instances of this problem separately exploiting given structures. If the lower level solution is unique, then one may utilize the properties of the corresponding solution operator to show the existence of global optimal solutions. This procedure is possible for the consideration of e.g. the obstacle problem, see [125]. Postulating quite restrictive assumptions, the author verifies the existence of global optimal solutions of a bilevel model possessing optimal control problems of ODEs at upper and lower level in [20].
In order to derive necessary optimality conditions for the bilevel model (BPP), it is a common idea to transfer it to a single-level surrogate problem. Here we are going to distinguish three different approaches, see [98]:

- If the lower level problem (4.1) possesses a unique solution $\psi(x)$ for all $x \in X_{\mathsf{ad}}$, then the original bilevel programming problem can be replaced by the equivalent model

$$\begin{aligned} F(x,\psi(x)) &\to \min_{x} \\ G(x) &\in C. \end{aligned}$$

  Now, the derivation of necessary optimality conditions heavily relies on the properties of the mapping $X_{\mathsf{ad}} \ni x \mapsto \psi(x) \in \mathcal{Y}$ and, thus, further discussion depends on the problem's structure. This approach is used to tackle finite-dimensional bilevel programming problems in e.g. [25] and [35].

In [85], we consider an abstract setting in Banach spaces where the lower level is given by an abstract parametric optimal control problem. Furthermore, all results which address the obstacle problem (1.3), see e.g. [67, 68, 125], can be settled here. In these papers, the authors exploit uniqueness and stability results for the solution mapping of a certain given variational inequality.

- If the lower level problem (4.1) is convex w.r.t. y, it seems to be reasonable to replace the condition $y \in \Psi(x)$ in (BPP) by the lower level necessary and sufficient optimality conditions. This can be done by means of a variational inequality, see e.g. [97, 128, 133] for the finite-dimensional situation and [64, 66, 73] where the authors study optimal control problems of PDEs with variational inequality constraints. It is also possible (in the presence of a constraint qualification) to use the KKT conditions of (4.1) to replace the lower level problem. However, introducing the corresponding Lagrange multiplier as a new decision variable, the resulting surrogate problem does not need to be equivalent to the original bilevel programming problem anymore, see [28] for the finite-dimensional case. We will show that similar or even harder difficulties may arise in a more general setting. The so-called KKT approach is used to derive necessary optimality conditions for (BPP) in [32, 33, 140] for the finite-dimensional situation and in [87] for the setting in Banach spaces with applications to a bilevel optimal control problem of ODEs.
- Let us introduce the so-called optimal value function $\varphi\colon \mathcal{X} \to \overline{\mathbb{R}}$ of (OV) by

$$\forall x \in \mathcal{X}\colon \quad \varphi(x) := \inf_y \{ f(x,y) \,|\, g(x,y) \in K \}. \tag{4.2}$$

  Since we clearly have $y \in \Psi(x)$ if and only if $f(x,y) \leq \varphi(x)$ and $g(x,y) \in K$ hold, the problem

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ G(x) &\in C \\ f(x,y) - \varphi(x) &\leq 0 \\ g(x,y) &\in K \end{aligned} \tag{OV}$$

  is fully equivalent to (BPP), see [34, Theorem 3.1] for the finite-dimensional case as well. However, this problem is still a challenging one since it contains the implicitly known function φ which is likely to be nonsmooth or even discontinuous. From [35, Theorem 3.1] and other contributions we see that constraint qualifications of reasonable strength applicable to nonsmooth programs fail to be satisfied at the feasible points of (OV). Especially, Fritz-John-type optimality conditions may hold at all feasible points of (OV), see [84, Section 3.2]. Using partial penalization w.r.t. the constraint $f(x,y) - \varphi(x) \leq 0$, Ye and Zhu initiated the study of optimality conditions and constraint qualifications for (OV) with their seminal work [138]. This paper inspired the theory in [29, 34, 92, 137] for finite-dimensional bilevel programming and the results in [13, 14, 74, 84, 130, 131] where bilevel optimal control problems of ODEs are studied. These results rely on a careful study of the subdifferentiability properties of the function φ, see e.g. [31, 91, 93]. In this thesis, we will exploit local Lipschitz properties of the optimal value function.

One may check the monographs [8, 26, 36] for further information and references on several other aspects of bilevel programming.

4.1. On a special class of bilevel programming problems with unique lower level solution

Here we are going to study the bilevel programming model (BPP) where the corresponding lower level problem is given as stated below:

$$\begin{aligned} \tfrac{1}{2} \left\| \mathtt{C}[y] - \mathtt{P}[x] \right\|_{\mathcal{M}}^2 + \tfrac{\sigma}{2} \left\| u - \mathtt{Q}[x] \right\|_{\mathcal{U}}^2 &\to \min_{y,u} \\ \mathtt{A}[y] - \mathtt{B}[u] - h(x) &= 0 \\ u &\in U_{\mathsf{ad}}. \end{aligned} \tag{4.3}$$

One may interpret (4.3) as a parametric optimal control problem of PDEs with control constraints governed by an elliptic (differential) operator $\mathtt{A}$. Here y and u represent the state and control function, respectively. The parameters can play the role of desired state and control. In this case, the overall bilevel programming problem may be seen as a parameter identification problem where the desired state is unknown and shall be reconstructed by measurements in some observation space $\mathcal{M}$.
In addition to the standing assumptions listed in Assumption 4.1, the data of (4.3) shall satisfy the following requirements.

Assumption 4.2. The Banach space $\mathcal{Y} = \mathcal{Y}_s \times \mathcal{U}$ is the product of a reflexive Banach space $\mathcal{Y}_s$ and a Hilbert space $\mathcal{U}$. Furthermore, $\mathcal{M}$ is a Hilbert space as well. We identify $\mathcal{U}$ and $\mathcal{M}$ with their corresponding dual spaces $\mathcal{U}^\star$ and $\mathcal{M}^\star$ by means of Riesz's representation theorem, respectively. Moreover, we assume that the norm in $\mathcal{U}$ and $\mathcal{M}$ is induced by the respective inner product. The mapping $h\colon \mathcal{X} \to \mathcal{Y}_s^\star$ is Lipschitz continuous and Fréchet differentiable. The set $U_{\mathsf{ad}} \subseteq \mathcal{U}$ is nonempty, closed, and convex. The linear operators $\mathtt{A} \in \mathbb{L}[\mathcal{Y}_s, \mathcal{Y}_s^\star]$, $\mathtt{B} \in \mathbb{L}[\mathcal{U}, \mathcal{Y}_s^\star]$, $\mathtt{C} \in \mathbb{L}[\mathcal{Y}_s, \mathcal{M}]$, $\mathtt{P} \in \mathbb{L}[\mathcal{X}, \mathcal{M}]$, and $\mathtt{Q} \in \mathbb{L}[\mathcal{X}, \mathcal{U}]$ are fixed. Moreover, $\mathtt{A}$ is an isomorphism. Finally, $\sigma > 0$ is a fixed constant.

Observe that the so-called state equation $\mathtt{A}[y] - \mathtt{B}[u] - h(x) = 0$ is equivalent to $y = (\mathtt{A}^{-1} \circ \mathtt{B})[u] + (\mathtt{A}^{-1} \circ h)(x)$ since $\mathtt{A}$ is a bijection. We introduce the control-to-observation operator $\mathtt{S} := \mathtt{C} \circ \mathtt{A}^{-1} \circ \mathtt{B} \in \mathbb{L}[\mathcal{U}, \mathcal{M}]$ in order to transfer (4.3) into the so-called reduced problem

$$\begin{aligned} \tfrac{1}{2} \left\| \mathtt{S}[u] - \big(\mathtt{P} - (\mathtt{C} \circ \mathtt{A}^{-1} \circ h)\big)(x) \right\|_{\mathcal{M}}^2 + \tfrac{\sigma}{2} \left\| u - \mathtt{Q}[x] \right\|_{\mathcal{U}}^2 \;&\to \min_u \\ u \;&\in U_{\mathsf{ad}}. \end{aligned} \tag{4.4}$$

Keeping in mind the state equation, the problems (4.3) and (4.4) are equivalent.

Proposition 4.1. Let $\bar{x} \in \mathcal{X}$ be arbitrarily chosen. Then for fixed parameter $x = \bar{x}$, problem (4.3) possesses a unique solution $(\bar{y}, \bar{u}) \in \mathcal{Y}_s \times \mathcal{U}$. Additionally, $(\bar{y}, \bar{u})$ is the unique solution of the following system:

$$\bar{y} = (\mathtt{A}^{-1} \circ \mathtt{B})[\bar{u}] + (\mathtt{A}^{-1} \circ h)(\bar{x}), \tag{4.5a}$$

$$\bar{u} = \operatorname{proj}_{U_{\mathsf{ad}}}\Big(\big(\tfrac{1}{\sigma}(\mathtt{S}^\star \circ \mathtt{P}) + \mathtt{Q} - \tfrac{1}{\sigma}(\mathtt{S}^\star \circ \mathtt{C} \circ \mathtt{A}^{-1} \circ h) \big)(\bar{x}) - \tfrac{1}{\sigma}(\mathtt{S}^\star \circ \mathtt{S})[\bar{u}] \Big). \tag{4.5b}$$

Proof. Similar as in Example 2.27 we see that for fixed $x = \bar{x}$, the objective functional of the reduced problem (4.4) is coercive and strictly convex. Thus, (4.4) possesses the unique solution $\bar{u}$ which is necessarily characterized by

$$-\big(\mathtt{S}^\star \circ \mathtt{S} + \sigma \mathtt{I}_{\mathcal{U}}\big)[\bar{u}] + \big((\mathtt{S}^\star \circ \mathtt{P}) + \sigma \mathtt{Q} - (\mathtt{S}^\star \circ \mathtt{C} \circ \mathtt{A}^{-1} \circ h)\big)(\bar{x}) \in \widehat{\mathcal{N}}_{U_{\mathsf{ad}}}(\bar{u}),$$

see Lemmas 2.5 and 2.29. Due to the inherent convexity of the reduced problem (4.4), this condition is also sufficient for the optimality of $\bar{u}$. Noting that $\widehat{\mathcal{N}}_{U_{\mathsf{ad}}}(\bar{u})$ is a cone, by means of Example 2.30 the above generalized equation is equivalent to condition (4.5b). The corresponding uniquely determined optimal state $\bar{y}$ is computed via the modified state equation (4.5a). This completes the proof. □

The above result justifies the definition of single-valued mappings $\psi_y\colon \mathcal{X} \to \mathcal{Y}_s$ and $\psi_u\colon \mathcal{X} \to \mathcal{U}$ such that ψ_u maps any parameter $\bar{x} \in \mathcal{X}$ to the solution $\bar{u} = \psi_u(\bar{x})$ of the nonsmooth equation (4.5b), whereas ψ_y equals $(\mathtt{A}^{-1} \circ \mathtt{B} \circ \psi_u) + (\mathtt{A}^{-1} \circ h)$, i.e. $\bar{y} = \psi_y(\bar{x})$ is valid. Thus, the bilevel programming problem (BPP) is equivalent to

$$\begin{aligned} \tilde{F}(x) := F(x, \psi_y(x), \psi_u(x)) \;&\to \min_x \\ G(x) \;&\in C. \end{aligned} \tag{4.6}$$

Clearly, in the presence of control constraints, the function ψ_u is expected to be nonsmooth and, thus, the same is true for ψ_y. However, in order to derive necessary optimality criteria for (BPP), we will state conditions which ensure that ψ_u and ψ_y are at least directionally differentiable. Moreover, we will present formulae which characterize the corresponding directional derivatives. Applying Lemma 2.29 as well as

an appropriate chain rule to (4.6) and using the theory of MPCCs presented in Chapter 3, we will finally obtain necessary optimality conditions.

For brevity, we introduce a Lipschitz continuous and Fréchet differentiable function $\eta\colon \mathcal{X} \to \mathcal{U}$ and a bounded, linear operator $\mathtt{E} \in \mathbb{L}[\mathcal{U},\mathcal{U}]$ by

$$\eta := \tfrac{1}{\sigma}(\mathtt{S}^\star \circ \mathtt{P}) + \mathtt{Q} - \tfrac{1}{\sigma}(\mathtt{S}^\star \circ \mathtt{C} \circ \mathtt{A}^{-1} \circ h), \qquad \mathtt{E} := \tfrac{1}{\sigma}(\mathtt{S}^\star \circ \mathtt{S}).$$

Note that $\mathtt{E}$ is a monotone operator since we have

$$\langle \mathtt{E}[u], u\rangle_{\mathcal{U}} = \tfrac{1}{\sigma}\langle (\mathtt{S}^\star \circ \mathtt{S})[u], u\rangle_{\mathcal{U}} = \tfrac{1}{\sigma}\langle \mathtt{S}[u], \mathtt{S}[u]\rangle_{\mathcal{M}} = \tfrac{1}{\sigma}\|\mathtt{S}[u]\|^2_{\mathcal{M}} \geq 0$$

for all $u \in \mathcal{U}$.

4.1.1. Nonsmooth equations governed by monotone operators

In this section, we are going to provide the theory which is necessary in order to see that the solution mapping of the nonsmooth equation (4.5b) is directionally differentiable under some additional assumptions. Furthermore, we implicitly characterize its directional derivative as the solution of another nonsmooth equation. Therefore, the parameter-dependent abstract nonsmooth equation

$$h = \mathrm{proj}_H\big(\upsilon(b) - \mathtt{U}[h]\big) \tag{4.7}$$

will be studied under the following standing assumptions.

Assumption 4.3. The mapping $\upsilon\colon \mathcal{B} \to \mathcal{H}$ is Fréchet differentiable and Lipschitz continuous with Lipschitz modulus $l > 0$. Therein, $\mathcal{B}$ is a Banach space, whereas $\mathcal{H}$ is a Hilbert space. We identify $\mathcal{H}$ and its dual $\mathcal{H}^\star$ by means of Riesz's representation theorem. Furthermore, we suppose that the norm in $\mathcal{H}$ is induced by its inner product. The bounded, linear operator $\mathtt{U} \in \mathbb{L}[\mathcal{H},\mathcal{H}]$ is monotone. Finally, $H \subseteq \mathcal{H}$ is a nonempty, closed, convex set.

First, we show that for any $b \in \mathcal{B}$, (4.7) possesses a unique solution which depends in a Lipschitz continuous way on the choice of b.

Lemma 4.2. For any $b \in \mathcal{B}$, the nonsmooth equation (4.7) possesses a unique solution $h_b \in \mathcal{H}$. Furthermore, the mapping $\psi\colon \mathcal{B} \to \mathcal{H}$ which maps $b \in \mathcal{B}$ to the unique solution $h_b \in \mathcal{H}$ of (4.7) is Lipschitz continuous with Lipschitz modulus l.

Proof. Fix an arbitrary parameter $b \in \mathcal{B}$. According to Example 2.30, $h_b \in \mathcal{H}$ is a solution of (4.7) if and only if it satisfies

$$\forall h \in \mathcal{H}\colon \quad \langle (\mathtt{I}_{\mathcal{H}} + \mathtt{U})[h_b], h - h_b\rangle_{\mathcal{H}} \geq \langle \upsilon(b), h - h_b\rangle_{\mathcal{H}}\,. \tag{4.8}$$

Thus, let us show that the variational problem (4.8) possesses a unique solution h_b. Therefore, observe that the linear operator $\mathtt{I}_{\mathcal{H}} + \mathtt{U}$ is elliptic. Indeed, we have

$$\forall h \in \mathcal{H}\colon \quad \langle (\mathtt{I}_{\mathcal{H}} + \mathtt{U})[h], h\rangle_{\mathcal{H}} = \|h\|^2_{\mathcal{H}} + \langle \mathtt{U}[h], h\rangle_{\mathcal{H}} \geq \|h\|^2_{\mathcal{H}}$$

from the monotonicity of $\mathtt{U}$. Consequently, (4.8) possesses a unique solution by means of [75, Theorem 2.1]. Thus, the solution mapping ψ of (4.7) is single-valued.
Now, choose $b, b' \in \mathcal{B}$ arbitrarily and fix the corresponding solutions $h_b, h_{b'} \in \mathcal{H}$ of (4.7), i.e. we have $h_b = \psi(b)$ and $h_{b'} = \psi(b')$. From (4.8) we obtain the inequalities

$$\langle (\mathtt{I}_{\mathcal{H}} + \mathtt{U})[h_b], h_{b'} - h_b\rangle_{\mathcal{H}} \geq \langle \upsilon(b), h_{b'} - h_b\rangle_{\mathcal{H}}\,, \quad \langle (\mathtt{I}_{\mathcal{H}} + \mathtt{U})[h_{b'}], h_b - h_{b'}\rangle_{\mathcal{H}} \geq \langle \upsilon(b'), h_b - h_{b'}\rangle_{\mathcal{H}}\,.$$

Summing them up and exploiting the bilinearity of the dual pairing yields

$$\langle (\mathtt{I}_{\mathcal{H}} + \mathtt{U})[h_b - h_{b'}], h_{b'} - h_b\rangle_{\mathcal{H}} \geq \langle \upsilon(b) - \upsilon(b'), h_{b'} - h_b\rangle_{\mathcal{H}}$$

or, equivalently,

$$\langle h_{b'} - h_b + \mathtt{U}[h_{b'} - h_b], h_{b'} - h_b\rangle_{\mathcal{H}} \leq \langle \upsilon(b') - \upsilon(b), h_{b'} - h_b\rangle_{\mathcal{H}}\,.$$

We exploit the monotonicity of $\mathtt{U}$ and the Lipschitz continuity of υ to see

$$\|h_{b'} - h_b\|^2_{\mathcal{H}} \leq \langle h_{b'} - h_b + \mathtt{U}[h_{b'} - h_b], h_{b'} - h_b\rangle_{\mathcal{H}} \leq \langle \upsilon(b') - \upsilon(b), h_{b'} - h_b\rangle_{\mathcal{H}} \leq l\,\|b' - b\|_{\mathcal{B}}\,\|h_{b'} - h_b\|_{\mathcal{H}}\,.$$

Thus, by definition of ψ, we have

$$\forall b, b' \in \mathcal{B}\colon \quad \|\psi(b') - \psi(b)\|_{\mathcal{H}} \leq l\,\|b' - b\|_{\mathcal{B}}\,,$$

i.e. the solution mapping ψ of (4.7) is Lipschitz continuous with Lipschitz modulus l. □

Clearly, if H does not equal the whole space $\mathcal{H}$, we cannot expect the Lipschitz continuous solution mapping ψ of (4.7) to be Fréchet differentiable. However, it seems to be reasonable that ψ possesses similar properties as the projection operator proj_H. For a deeper analysis of ψ, we will exploit the following result by Haraux which says that the projection operator proj_H is directionally differentiable if some additional properties hold, e.g. if H is polyhedric, see [57, Theorems 1 and 2].

Lemma 4.3. Choose $h, \bar{h} \in \mathcal{H}$ such that $\bar{h} = \mathrm{proj}_H(h)$ holds. Assume the existence of a self-adjoint operator $\mathtt{L} \in \mathbb{L}[\mathcal{H}, \mathcal{H}]$ which possesses the following two properties:

$$\mathtt{L} \circ \mathrm{proj}_{\mathcal{K}_H(\bar{h}, h-\bar{h})} = \mathrm{proj}_{\mathcal{K}_H(\bar{h}, h-\bar{h})} \circ \mathtt{L}, \tag{4.9a}$$

$$\forall \delta \in \mathcal{K}_H(\bar{h}, h - \bar{h})\colon \ \lim_{t \searrow 0} \frac{\mathrm{proj}_H(h + t\delta) - \mathrm{proj}_H(h)}{t} = \mathtt{L}^2[\delta]. \tag{4.9b}$$

Then proj_H is directionally differentiable at h and the following formula holds:

$$\forall \delta \in \mathcal{H}\colon \quad \mathrm{proj}'_H(h; \delta) = \left(\mathtt{L}^2 \circ \mathrm{proj}_{\mathcal{K}_H(\bar{h}, h-\bar{h})}\right)(\delta). \tag{4.10}$$

Furthermore, if H is polyhedric w.r.t. $(\bar{h}, h - \bar{h})$, then the conditions (4.9) hold true with $\mathtt{L} := \mathtt{I}_{\mathcal{H}}$, i.e. proj_H is directionally differentiable and the corresponding directional derivative satisfies (4.10) with $\mathtt{L} := \mathtt{I}_{\mathcal{H}}$.

Note that (4.9b) already implies the directional differentiability of proj_H at h in all directions coming from the critical cone $\mathcal{K}_H(\bar{h}, h - \bar{h})$.
Below, we present a criterion which can be applied in order to check whether the condition (4.9a) is satisfied.

Lemma 4.4. Choose $h, \bar{h} \in \mathcal{H}$ such that $\bar{h} = \mathrm{proj}_H(h)$ holds. Assume the existence of a self-adjoint automorphism $\mathtt{L} \in \mathbb{L}[\mathcal{H}, \mathcal{H}]$ satisfying $\mathtt{L}\left[\mathcal{K}_H(\bar{h}, h - \bar{h})\right] = \mathcal{K}_H(\bar{h}, h - \bar{h})$ and

$$\forall \delta \in \mathcal{K}_H(\bar{h}, h - \bar{h})\, \forall \delta^* \in \mathcal{K}_H(\bar{h}, h - \bar{h})^\circ\colon \quad \langle \delta^*, \delta\rangle_{\mathcal{H}} = 0 \iff \langle \mathtt{L}[\delta^*], \mathtt{L}[\delta]\rangle_{\mathcal{H}} = 0.$$

Then (4.9a) is valid.

Proof. First, we show that $\mathtt{L}\left[\mathcal{K}_H(\bar{h}, h - \bar{h})^\circ\right] = \mathcal{K}_H(\bar{h}, h - \bar{h})^\circ$ holds. We exploit the assumptions of the lemma in order to see the following equivalences:

$$\begin{aligned}
\delta^* \in \mathtt{L}\left[\mathcal{K}_H(\bar{h}, h - \bar{h})^\circ\right] &\iff \exists \theta^* \in \mathcal{K}_H(\bar{h}, h - \bar{h})^\circ\colon\ \delta^* = \mathtt{L}[\theta^*] \\
&\iff \exists \theta^* \in \mathcal{H}\, \forall \delta \in \mathcal{K}_H(\bar{h}, h - \bar{h})\colon\ \langle \theta^*, \delta\rangle_{\mathcal{H}} \leq 0 \ \wedge\ \delta^* = \mathtt{L}[\theta^*] \\
&\iff \exists \theta^* \in \mathcal{H}\, \forall \delta \in \mathcal{K}_H(\bar{h}, h - \bar{h})\colon\ \langle \theta^*, \mathtt{L}[\delta]\rangle_{\mathcal{H}} \leq 0 \ \wedge\ \delta^* = \mathtt{L}[\theta^*] \\
&\iff \exists \theta^* \in \mathcal{H}\, \forall \delta \in \mathcal{K}_H(\bar{h}, h - \bar{h})\colon\ \langle \mathtt{L}[\theta^*], \delta\rangle_{\mathcal{H}} \leq 0 \ \wedge\ \delta^* = \mathtt{L}[\theta^*] \\
&\iff \forall \delta \in \mathcal{K}_H(\bar{h}, h - \bar{h})\colon\ \langle \delta^*, \delta\rangle_{\mathcal{H}} \leq 0 \\
&\iff \delta^* \in \mathcal{K}_H(\bar{h}, h - \bar{h})^\circ.
\end{aligned}$$

Let us choose $\delta, \bar{\delta} \in \mathcal{H}$ arbitrarily. Due to the above result, the assumptions of the lemma, Example 2.30, and the fact that $\mathcal{K}_H(\bar{h}, h-\bar{h})$ is a cone, we obtain:

$$\begin{aligned}
\bar{\delta} &= \Big(\mathtt{L} \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)(\delta) \\
&\iff \bar{\delta} = \mathtt{L}[\theta] \wedge \theta = \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}(\delta) \\
&\iff \bar{\delta} = \mathtt{L}[\theta] \wedge \theta \in \mathcal{K}_H(\bar{h}, h-\bar{h}) \wedge \delta - \theta \in \mathcal{K}_H(\bar{h}, h-\bar{h})^\circ \wedge \langle \theta, \delta - \theta \rangle_{\mathcal{H}} = 0 \\
&\iff \mathtt{L}^{-1}[\bar{\delta}] \in \mathcal{K}_H(\bar{h}, h-\bar{h}) \wedge \delta - \mathtt{L}^{-1}[\bar{\delta}] \in \mathcal{K}_H(\bar{h}, h-\bar{h})^\circ \wedge \big\langle \mathtt{L}^{-1}[\bar{\delta}], \delta - \mathtt{L}^{-1}[\bar{\delta}] \big\rangle_{\mathcal{H}} = 0 \\
&\iff \bar{\delta} \in \mathcal{K}_H(\bar{h}, h-\bar{h}), \mathtt{L}[\delta] - \bar{\delta} \in \mathcal{K}_H(\bar{h}, h-\bar{h})^\circ \wedge \big\langle \bar{\delta}, \mathtt{L}[\delta] - \bar{\delta} \big\rangle_{\mathcal{H}} = 0 \\
&\iff \bar{\delta} = \Big(\operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})} \circ \mathtt{L}\Big)(\delta).
\end{aligned}$$

This completes the proof. □

In order to show our main result of this section, we need the following observation. Its proof is straightforward and, hence, omitted.

Lemma 4.5. Let $\mathcal{X}$ be a Banach space. For mappings $\mathtt{f}, \mathtt{g} \colon \mathcal{X} \to \mathcal{X}$, we consider the three nonlinear systems

$$\left.\begin{aligned} x &= \mathtt{f}(y) \\ y &= \mathtt{g}(x) \end{aligned}\right\} \text{(I)} \qquad \left.\begin{aligned} x &= \mathtt{f}(y) \\ y &= (\mathtt{g} \circ \mathtt{f})(y) \end{aligned}\right\} \text{(II)} \qquad \left.\begin{aligned} x &= (\mathtt{f} \circ \mathtt{g})(x) \\ y &= \mathtt{g}(x) \end{aligned}\right\} \text{(III)}.$$

Then $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{X}$ solves (I) if and only if it solves (II) if and only if it solves (III). Especially, (I) possesses a unique solution if and only if $\mathtt{g} \circ \mathtt{f}$ possesses a unique fixpoint if and only if $\mathtt{f} \circ \mathtt{g}$ possesses a unique fixpoint.

Now, we are able to prove that under the assumptions of Lemma 4.3, the solution mapping ψ of (4.7) is directionally differentiable. A similar result validated by related proof techniques is presented in [54, Theorem 4.3].

Proposition 4.6. Let $b \in \mathcal{B}$ be arbitrarily chosen and set $\bar{h} := \psi(b)$ and $h := \upsilon(b) - \mathtt{U}[\bar{h}]$. Assume the existence of a self-adjoint operator $\mathtt{L} \in \mathbb{L}[\mathcal{H}, \mathcal{H}]$ possessing the properties (4.9) and let

$$\forall h' \in \mathcal{H} \colon \quad \operatorname{proj}_H(h') - \operatorname{proj}_H(h) - \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)(h'-h) = o(\|h'-h\|_{\mathcal{H}})$$

for some function $o \colon \mathbb{R}_0^+ \to \mathcal{H}$ with $\lim_{t \searrow 0} \frac{o(t)}{t} = 0$ be satisfied, i.e. proj_H is supposed to be B-differentiable at h. Then ψ is directionally differentiable at b, and for any direction $\delta_b \in \mathcal{B}$, the corresponding directional derivative $\psi'(b; \delta_b)$ is the unique solution of the following nonsmooth equation:

$$\delta_h = \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\big(\upsilon'(b)[\delta_b] - \mathtt{U}[\delta_h]\big). \tag{4.11}$$

Furthermore, the mapping $\delta_b \mapsto \psi'(b; \delta_b)$ is Lipschitz continuous.

Proof. Fix $b \in \mathcal{B}$ and an arbitrary direction $\delta_b \in \mathcal{B}$. According to (4.9a) and Lemma 4.5, (4.11) possesses a unique solution if and only if the following system possesses a unique solution:

$$\delta_h = \mathtt{L}[\theta_h], \tag{4.12a}$$

$$\theta_h = \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big(\big(\mathtt{L} \circ \upsilon'(b)\big)[\delta_b] - \big(\mathtt{L} \circ \mathtt{U}\big)[\delta_h]\Big). \tag{4.12b}$$

We apply Lemma 4.5 once more in order to see that $(\bar{\delta}_h, \bar{\theta}_h)$ solves (4.12) if and only if it solves

$$\delta_h = \mathtt{L}[\theta_h], \tag{4.13a}$$

$$\theta_h = \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big(\big(\mathtt{L} \circ \upsilon'(b)\big)[\delta_b] - \big(\mathtt{L} \circ \mathtt{U} \circ \mathtt{L}\big)[\theta_h]\Big). \tag{4.13b}$$

Clearly, $\mathtt{L} \circ \upsilon'(b)$ is a bounded linear operator and, thus, Fréchet differentiable and Lipschitz continuous. Furthermore, it is easy to check that $\mathtt{L} \circ \mathtt{U} \circ \mathtt{L}$ is monotone since $\mathtt{U}$ is monotone and $\mathtt{L}$ is self-adjoint. Thus, by means of Lemma 4.2, (4.13b) possesses a unique solution $\bar{\theta}_h$ and the mapping $\delta_b \mapsto \bar{\theta}_h$ is Lipschitz continuous. Retracing the above arguments, (4.11) possesses a unique solution $\bar{\delta}_h$ and the mapping $\delta_b \mapsto \bar{\delta}_h$ is Lipschitz continuous.
Now, we are going to show $\psi'(b;\delta_b) = \bar{\delta}_h$. In order to exclude trivial situations, we assume $\delta_b \neq 0$ since for $\delta_b = 0$, the assertion of the lemma is obviously satisfied. Fix some $t \geq 0$. Then Lemma 4.3 and the B-differentiability of proj_H at h yield

$$\begin{aligned}\psi(b+t\delta_b) - \psi(b) &= \operatorname{proj}_H\big(\upsilon(b+t\delta_b) - \mathtt{U}[\psi(b+t\delta_b)]\big) - \operatorname{proj}_H\big(\upsilon(b) - \mathtt{U}[\psi(b)]\big) \\ &= \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\big(\upsilon(b+t\delta_b) - \upsilon(b) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\big) + r_1(t)\end{aligned} \tag{4.14}$$

where

$$r_1(t) := o(\|\upsilon(b+t\delta_b) - \upsilon(b) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\|_{\mathcal{H}})$$

satisfies

$$\|\upsilon(b+t\delta_b) - \upsilon(b) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\|_{\mathcal{H}} \to 0 \implies \frac{\|r_1(t)\|_{\mathcal{H}}}{\|\upsilon(b+t\delta_b) - \upsilon(b) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\|_{\mathcal{H}}} \to 0.$$

The Lipschitz continuity of υ, $\mathtt{U}$, and ψ, see Lemma 4.2, guarantee the existence of a constant $\alpha > 0$ such that

$$\|\upsilon(b+t\delta_b) - \upsilon(b) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\|_{\mathcal{H}} \leq \alpha t$$

is satisfied. Thus, we obtain

$$t \to 0 \implies \frac{\|r_1(t)\|_{\mathcal{H}}}{t} \to 0. \tag{4.15}$$

Rearranging (4.14) and exploiting the Fréchet differentiability of υ leads to

$$\begin{aligned}\psi(b+t\delta_b) - \psi(b) - r_1(t) &= \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\big(\upsilon(b+t\delta_b) - \upsilon(b) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\big) \\ &= \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\big(t\upsilon'(b)[\delta_b] + o_1(t) - \mathtt{U}[\psi(b+t\delta_b) - \psi(b)]\big) \\ &= \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\big(t\upsilon'(b)[\delta_b] - \mathtt{U}[\psi(b+t\delta_b) - \psi(b) - r_1(t)] + r_2(t)\big)\end{aligned}$$

where $r_2(t) := o_1(t) - \mathtt{U}[r_1(t)]$ shall hold and $o_1 \colon \mathbb{R}_0^+ \to \mathcal{H}$ satisfies $\lim_{t \searrow 0} \frac{o_1(t)}{t} = 0$. Since $\mathcal{K}_H(\bar{h}, h-\bar{h})$ is a cone, we obtain

$$\frac{\psi(b+t\delta_b) - \psi(b) - r_1(t)}{t} = \Big(\mathtt{L}^2 \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\left(\upsilon'(b)[\delta_b] - \mathtt{U}\left[\frac{\psi(b+t\delta_b) - \psi(b) - r_1(t)}{t}\right] + \frac{r_2(t)}{t}\right)$$

for positive t. From (4.15)

$$t \to 0 \implies \frac{\|r_2(t)\|_{\mathcal{H}}}{t} \to 0 \tag{4.16}$$

follows easily. We introduce a function $\xi \colon \mathbb{R}^+ \to \mathcal{H}$ by

$$\forall t \in \mathbb{R}^+ \colon \quad \xi(t) := \frac{\psi(b+t\delta_b) - \psi(b) - r_1(t)}{t}$$

and exploit (4.9a) to obtain

$$\xi(t) = \Big(\mathtt{L} \circ \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\Big)\left(\big(\mathtt{L} \circ \upsilon'(b)\big)[\delta_b] - \big(\mathtt{L} \circ \mathtt{U}\big)[\xi(t)] + \mathtt{L}\left[\tfrac{r_2(t)}{t}\right]\right). \tag{4.17}$$

We interpret the above term as a nonsmooth equation in $\xi(t)$ which is parameterized by $\frac{r_2(t)}{t}$. Applying Lemma 4.5, we can split this equation equivalently into two:

$$\begin{aligned}\xi(t) &= \mathtt{L}[\eta(t)], \\ \eta(t) &= \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\left(\big(\mathtt{L} \circ \upsilon'(b)\big)[\delta_b] - \big(\mathtt{L} \circ \mathtt{U}\big)[\xi(t)] + \mathtt{L}\left[\tfrac{r_2(t)}{t}\right]\right).\end{aligned}$$

Again, by Lemma 4.5, this system possesses the same solutions as

$$\xi(t) = \mathtt{L}[\eta(t)], \tag{4.18a}$$

$$\eta(t) = \operatorname{proj}_{\mathcal{K}_H(\bar{h},h-\bar{h})}\left(\big(\mathtt{L}\circ\upsilon'(b)\big)[\delta_b] - \big(\mathtt{L}\circ\mathtt{U}\circ\mathtt{L}\big)[\eta(t)] + \mathtt{L}\left[\tfrac{r_2(t)}{t}\right]\right). \tag{4.18b}$$

Note that the mapping $\frac{r_2(t)}{t} \mapsto \big(\mathtt{L}\circ\upsilon'(b)\big)[\delta_b] + \mathtt{L}\left[\frac{r_2(t)}{t}\right]$ is Fréchet differentiable and Lipschitz continuous, whereas the operator $\mathtt{L}\circ\mathtt{U}\circ\mathtt{L}$ is monotone. Invoking Lemma 4.2, (4.18b) possesses a unique solution $\bar{\eta}(t)_{r_2(t)/t}$ for any choice of $\frac{r_2(t)}{t} \in \mathcal{H}$ and the mapping $\frac{r_2(t)}{t} \mapsto \bar{\eta}(t)_{r_2(t)/t}$ is Lipschitz continuous. The same holds true for the overall solution mapping $\frac{r_2(t)}{t} \mapsto (\bar{\xi}(t)_{r_2(t)/t}, \bar{\eta}(t)_{r_2(t)/t})$ of the system (4.18). Particularly, $\frac{r_2(t)}{t} \mapsto \bar{\xi}(t)_{r_2(t)/t}$ is the single-valued and Lipschitz continuous solution mapping of (4.17). Now, observe that $\bar{\xi}(t)_0 = \bar{\delta}_h$ holds. On the other hand, we have $\bar{\xi}(t)_{r_2(t)/t} = \frac{1}{t}(\psi(b+t\delta_b) - \psi(b) - r_1(t))$ by construction. This leads to

$$\left\|\frac{\psi(b+t\delta_b)-\psi(b)-r_1(t)}{t} - \bar{\delta}_h\right\|_{\mathcal{H}} = \left\|\bar{\xi}(t)_{r_2(t)/t} - \bar{\xi}(t)_0\right\|_{\mathcal{H}} \leq \beta\frac{\|r_2(t)\|_{\mathcal{H}}}{t}$$

where $\beta > 0$ is a fixed constant. Combining the last inequality with (4.15) and (4.16) yields

$$\left\|\frac{\psi(b+t\delta_b)-\psi(b)}{t} - \bar{\delta}_h\right\|_{\mathcal{H}} \leq \frac{\|r_1(t)\|_{\mathcal{H}} + \beta\,\|r_2(t)\|_{\mathcal{H}}}{t} \to 0$$

for $t \searrow 0$. This shows $\psi'(b;\delta_b) = \bar{\delta}_h$ which completes the proof. □

Remark 4.7. The assumption on proj_H to be B-differentiable is essential for the proof of the above proposition. Note that for a finite-dimensional space $\mathcal{H}$, proj_H is directionally differentiable if and only if it is B-differentiable at a certain reference point since it is a Lipschitz continuous function, see [112, Proposition 3.5].
On the other hand, B-differentiability of the projection is not inherent if the underlying space is infinite-dimensional. In [54, Remark 4], the authors show that the projection onto $L^2(\Omega)^+_0$, where $\Omega \subseteq \mathbb{R}^d$ is a bounded domain, is not necessarily B-differentiable from $L^2(\Omega)$ to $L^2(\Omega)$ although it is directionally differentiable since $L^2(\Omega)^+_0$ is polyhedric, see Lemma 4.3. Nevertheless, under more restrictive assumptions on υ and $\mathtt{U}$, the results of Proposition 4.6 stay true even if the space $\mathcal{H} = L^2(\Omega)$ is considered, see [54, Theorem 4.3] for the details.

4.1.2. Necessary optimality conditions

Here we exploit the results obtained in Chapter 3 and Section 4.1.1 in order to find necessary optimality conditions for the bilevel programming problem (BPP) with lower level (4.3). Let us start with the following consequence of Proposition 4.6 and the chain rule for the composition of directionally differentiable mappings, see [112, Proposition 3.6].

Lemma 4.8. Fix an arbitrary parameter $\bar{x} \in \mathcal{X}$ and set $\bar{y} := \psi_y(\bar{x})$, $\bar{u} := \psi_u(\bar{x})$, as well as $\bar{w} := \eta(\bar{x}) - \mathtt{E}[\bar{u}]$. Suppose that there exists a self-adjoint operator $\mathtt{L} \in \mathbb{L}[\mathcal{U},\mathcal{U}]$ which satisfies the following two conditions:

$$\mathtt{L}\circ\operatorname{proj}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})} = \operatorname{proj}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}\circ\mathtt{L}, \tag{4.19a}$$

$$\forall \delta \in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})\colon\ \lim_{t\searrow 0}\frac{\operatorname{proj}_{U_{\mathrm{ad}}}(\bar{w}+t\delta) - \operatorname{proj}_{U_{\mathrm{ad}}}(\bar{w})}{t} = \mathtt{L}^2[\delta]. \tag{4.19b}$$

Moreover, let some function $o\colon \mathbb{R}^+_0 \to \mathcal{U}$ with $\lim_{t\searrow 0}\frac{o(t)}{t} = 0$ exist which satisfies

$$\forall w \in \mathcal{U}\colon \quad \operatorname{proj}_{U_{\mathrm{ad}}}(w) - \operatorname{proj}_{U_{\mathrm{ad}}}(\bar{w}) - \left(\mathtt{L}^2\circ\operatorname{proj}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}\right)(w-\bar{w}) = o(\|w-\bar{w}\|_{\mathcal{U}}), \tag{4.20}$$

i.e. $\operatorname{proj}_{U_{\mathrm{ad}}}$ is B-differentiable at $\bar{w}$. Then ψ_y and ψ_u are directionally differentiable at $\bar{x}$. For all $\delta_x \in \mathcal{X}$, the directional derivative $\psi'_u(\bar{x};\delta_x)$ is the unique solution of the nonsmooth equation

$$\delta_u = \left(\mathtt{L}^2\circ\operatorname{proj}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}\right)\big(\eta'(\bar{x})[\delta_x] - \mathtt{E}[\delta_u]\big),$$

whereas $\psi_y'(\bar{x};\delta_x)$ can be computed as stated below:

$$\psi_y'(\bar{x};\delta_x) = \big(\mathtt{A}^{-1}\circ\mathtt{B}\big)[\psi_u'(\bar{x};\delta_x)] + \big(\mathtt{A}^{-1}\circ h'(\bar{x})\big)[\delta_x].$$

Furthermore, the mappings $\delta_x \mapsto \psi_y'(\bar{x};\delta_x)$ and $\delta_x \mapsto \psi_u'(\bar{x};\delta_x)$ are Lipschitz continuous.

Now, we are able to show the following result.

Proposition 4.9. Let $(\bar{x},\bar{y},\bar{u}) \in \mathcal{X}\times\mathcal{Y}_s\times\mathcal{U}$ be a local optimal solution of (BPP) with lower level (4.3) where F is continuously Fréchet differentiable. Let the constraint qualification

$$G'(\bar{x})[\mathcal{X}] - \mathcal{R}_C(G(\bar{x})) = \mathcal{W} \tag{4.21}$$

be satisfied. Set $\bar{w} := \eta(\bar{x}) - \mathtt{E}[\bar{u}]$. Suppose that there exists a self-adjoint operator $\mathtt{L} \in \mathbb{L}[\mathcal{U},\mathcal{U}]$ which satisfies the conditions (4.19) and (4.20). Then $(\bar{\delta}_x,\bar{\delta}_u,\bar{\delta}_\pi) := (0,0,0)$ is a global optimal solution of the following MPCC:

$$\begin{aligned}
\big(F_x'(\bar{x},\bar{y},\bar{u}) + F_y'(\bar{x},\bar{y},\bar{u})\circ\mathtt{A}^{-1}\circ h'(\bar{x})\big)[\delta_x] & \\
+\big(F_y'(\bar{x},\bar{y},\bar{u})\circ\mathtt{A}^{-1}\circ\mathtt{B} + F_u'(\bar{x},\bar{y},\bar{u})\big)[\delta_u] &\to \min_{\delta_x,\delta_u,\delta_\pi} \\
G'(\bar{x})[\delta_x] &\in \mathcal{T}_C(G(\bar{x})) \\
\delta_u - \mathtt{L}^2[\delta_\pi] &= 0 \\
\delta_\pi &\in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u}) \\
\eta'(\bar{x})[\delta_x] - \mathtt{E}[\delta_u] - \delta_\pi &\in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})^\circ \\
\langle \eta'(\bar{x})[\delta_x] - \mathtt{E}[\delta_u] - \delta_\pi, \delta_\pi\rangle_{\mathcal{U}} &= 0.
\end{aligned} \tag{4.22}$$

Proof. Due to Proposition 4.1, $\bar{x}$ is a local optimal solution of (4.6). By means of Lemma 4.8 ψ_y and ψ_u are directionally differentiable at $\bar{x}$. Furthermore, these functions are Lipschitz continuous, see Lemma 4.2. Recall that $\tilde{F}\colon \mathcal{X}\to\mathbb{R}$ denotes the objective functional of (4.6). Since F is continuously Fréchet differentiable at $(\bar{x},\bar{y},\bar{u})$, it is locally Lipschitz continuous there. Thus, $\tilde{F}$ is locally Lipschitz continuous at $\bar{x}$. Invoking the chain rule [112, Proposition 3.6], $\tilde{F}$ is directionally differentiable at $\bar{x}$ as well. Using Lemma 4.8 once more, we obtain

$$\begin{aligned}
\tilde{F}'(\bar{x};\delta_x) &= F_x'(\bar{x},\psi_y(\bar{x}),\psi_u(\bar{x}))[\delta_x] + F_y'(\bar{x},\psi_y(\bar{x}),\psi_u(\bar{x}))[\psi_y'(\bar{x};\delta_x)] + F_u'(\bar{x},\psi_y(\bar{x}),\psi_u(\bar{x}))[\psi_u'(\bar{x};\delta_x)] \\
&= F_x'(\bar{x},\bar{y},\bar{u})[\delta_x] + F_y'(\bar{x},\bar{y},\bar{u})\Big[\big(\mathtt{A}^{-1}\circ\mathtt{B}\big)[\psi_u'(\bar{x};\delta_x)] + \big(\mathtt{A}^{-1}\circ h'(\bar{x})\big)[\delta_x]\Big] + F_u'(\bar{x},\bar{y},\bar{u})[\psi_u'(\bar{x};\delta_x)] \\
&= \big(F_x'(\bar{x},\bar{y},\bar{u}) + F_y'(\bar{x},\bar{y},\bar{u})\circ\mathtt{A}^{-1}\circ h'(\bar{x})\big)[\delta_x] + \big(F_y'(\bar{x},\bar{y},\bar{u})\circ\mathtt{A}^{-1}\circ\mathtt{B} + F_u'(\bar{x},\bar{y},\bar{u})\big)[\psi_u'(\bar{x};\delta_x)]
\end{aligned}$$

for arbitrary directions $\delta_x \in \mathcal{X}$, and the mapping $\delta_x \mapsto \tilde{F}'(\bar{x};\delta_x)$ is Lipschitz continuous.
Let us define the perturbation mapping $\Delta\colon \mathcal{W}\rightrightarrows\mathcal{X}$ as stated below:

$$\forall w\in\mathcal{W}\colon \quad \Delta(w) := \{x\in\mathcal{X} \mid G(x)+w\in C\}.$$

The postulated constraint qualification (4.21) implies that Δ is calm at $(0,\bar{x})$, see [17, Theorem 2.87 and Remark 2.88] and [61, Section 1]. Hence, we can apply [90, Lemma 5.47] in order to see that there exists a constant $c>0$ such that $(\bar{x},G(\bar{x}))$ is a local optimal solution of the penalized problem

$$\begin{aligned}
\tilde{F}(x) + c\,\|G(x)-w\|_{\mathcal{W}} &\to \min_{x,w} \\
w &\in C.
\end{aligned}$$

Observe that $(x,w)\mapsto c\,\|G(x)-w\|_{\mathcal{W}}$ is directionally differentiable at $(\bar{x},G(\bar{x}))$ with directional derivative $(\delta_x,\delta_w)\mapsto c\,\|G'(\bar{x})[\delta_x]-\delta_w\|_{\mathcal{W}}$, see [112, Proposition 3.6], and this mapping is Lipschitz continuous. We apply Lemma 2.29 in order to find the necessary optimality condition

$$\forall\delta_x\in\mathcal{X}\,\forall\delta_w\in\mathcal{R}_C(G(\bar{x}))\colon \quad \tilde{F}'(\bar{x};\delta_x) + c\,\|G'(\bar{x})[\delta_x]-\delta_w\|_{\mathcal{W}} \geq 0.$$

Next, we exploit the convexity of C and the continuity of the appearing directional derivatives to obtain the stronger condition

$$\forall \delta_x \in \mathcal{X}\, \forall \delta_w \in \mathcal{T}_C(G(\bar{x})): \quad \tilde{F}'(\bar{x};\delta_x) + c\,\|G'(\bar{x})[\delta_x] - \delta_w\|_{\mathcal{W}} \geq 0.$$

This implies

$$\forall \delta_x \in \mathcal{X}: \quad G'(\bar{x})[\delta_x] \in \mathcal{T}_C(G(\bar{x})) \implies \tilde{F}(\bar{x};\delta_x) \geq 0.$$

Consequently, $\bar{\delta}_x = 0$ solves the surrogate problem

$$\begin{aligned} \tilde{F}'(\bar{x};\delta_x) &\to \min_{\delta_x} \\ G'(\bar{x})[\delta_x] &\in \mathcal{T}_C(G(\bar{x})). \end{aligned}$$

Let us introduce the additional variable $\delta_u := \psi_u'(\bar{x};\delta_x)$ in order to handle the directional derivative of $\tilde{F}$ properly. Furthermore, we set

$$\delta_\pi := \mathrm{proj}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}\big(\eta'(\bar{x})[\delta_x] - \mathtt{E}[\delta_u]\big).$$

Since $\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})$ is a closed, convex cone, we can invoke Example 2.30 and Lemma 4.8 in order to see that the relation $\delta_u = \psi_u'(\bar{x};\delta_x)$ is equivalent to the complementarity system

$$\begin{aligned} \delta_u - \mathtt{L}^2[\delta_\pi] &= 0 \\ \delta_\pi &\in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u}) \\ \eta'(\bar{x})[\delta_x] - \mathtt{E}[\delta_u] - \delta_\pi &\in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})^\circ \\ \langle \eta'(\bar{x})[\delta_x] - \mathtt{E}[\delta_u] - \delta_\pi, \delta_\pi\rangle_{\mathcal{U}} &= 0. \end{aligned}$$

For $\delta_x = 0$, its unique solution is given by $\bar{\delta}_u = \bar{\delta}_\pi = 0$. Combining this observation with the above formula for the directional derivative of $\tilde{F}$, the proposition's assertion is proven. □

Let us exploit the theory of MPCCs presented in Chapter 3 in order to formulate necessary optimality conditions of KKT-type for the bilevel programming problem under consideration.

Theorem 4.10. Let $(\bar{x},\bar{y},\bar{u}) \in \mathcal{X}\times\mathcal{Y}_s\times\mathcal{U}$ be a local optimal solution of (BPP) with lower level (4.3) and let all the assumptions of Proposition 4.9 hold. Then the following statements are valid:

1. Assume that the constraint qualification

$$\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}} \\ \eta'(\bar{x}) & -\mathtt{E}-\mathtt{I}_{\mathcal{U}} \end{bmatrix}\begin{pmatrix}\mathcal{X}\\ \mathcal{U}\end{pmatrix} - \begin{bmatrix} \mathtt{I}_{\mathcal{W}} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{L}^2 & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}}-\mathtt{L}^2 & \mathtt{I}_{\mathcal{U}} \end{bmatrix}\begin{pmatrix}\mathcal{T}_C(G(\bar{x}))\\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})^{\circ\perp}\\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})^{\perp}\end{pmatrix} = \begin{pmatrix}\mathcal{W}\\ \mathcal{U}\\ \mathcal{U}\end{pmatrix} \tag{4.23}$$

holds. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mu,\nu \in \mathcal{U}$, and $p \in \mathcal{Y}_s$ which solve the following system:

$$0 = F_x'(\bar{x},\bar{y},\bar{u}) + h'(\bar{x})^\star[p] + G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\nu], \tag{4.24a}$$
$$0 = F_u'(\bar{x},\bar{y},\bar{u}) + \mathtt{B}^\star[p] + \mu - (\mathtt{E}+\mathtt{I}_{\mathcal{U}})[\nu], \tag{4.24b}$$
$$0 = \mathtt{A}^\star[p] - F_y'(\bar{x},\bar{y},\bar{u}), \tag{4.24c}$$
$$\rho \in \mathcal{N}_C(G(\bar{x})), \tag{4.24d}$$
$$\nu + \mathtt{L}^2[\mu-\nu] \in \mathrm{cl}\big(\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})^\circ - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})^\circ\big), \tag{4.24e}$$
$$\nu \in \mathrm{cl}\big(\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u}) - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})\big). \tag{4.24f}$$

2. Let $\mathcal{X}$ and $\mathcal{W}$ be reflexive. Assume that the constraint qualification

$$\left.\begin{aligned} &0 = G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\nu], \\ &0 = \mu - (\mathtt{E}+\mathtt{I}_{\mathcal{U}})[\nu], \\ &\rho \in \mathcal{N}_C(G(\bar{x})), \\ &(\nu + \mathtt{L}^2[\mu-\nu], \nu) \in \mathcal{N}_{\mathrm{gph}\,\mathcal{N}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}}(0,0) \end{aligned}\right\} \implies \rho = 0,\ \mu = 0,\ \nu = 0 \tag{4.25}$$

holds, whereas one of the following conditions is valid:

a) $\mathcal{T}_C(G(\bar{x}))$ is SNC at 0 and $\mathcal{U}$ is finite-dimensional,

b) the set $\mathcal{T}_C(G(\bar{x})) \times \operatorname{gph} \mathcal{N}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}$ is SNC at $(0,0,0)$.

Then there are multipliers $\rho \in \mathcal{W}^\star$, $\mu, \nu \in \mathcal{U}$, and $p \in \mathcal{Y}_s$ which satisfy the conditions (4.24a) - (4.24d) and

$$\big(\nu + \mathtt{L}^2[\mu - \nu], \nu\big) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}}(0,0). \tag{4.26}$$

3. Assume that the constraint qualifications (4.23) and

$$\operatorname{cl}\left(\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{U} \end{pmatrix} - \begin{bmatrix} \mathtt{I}_{\mathcal{W}} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{L}^2 & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}} - \mathtt{L}^2 & \mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{N}_C(G(\bar{x}))_\perp \\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^{\circ\perp} \\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^{\perp} \end{pmatrix}\right) = \begin{pmatrix} \mathcal{W} \\ \mathcal{U} \\ \mathcal{U} \end{pmatrix} \tag{4.27}$$

are valid. Then there are multipliers $\rho \in \mathcal{W}^\star$, $\mu, \nu \in \mathcal{U}$, and $p \in \mathcal{Y}_s$ which satisfy the conditions (4.24a) - (4.24d) and

$$\nu + \mathtt{L}^2[\mu - \nu] \in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^\circ, \tag{4.28a}$$
$$\nu \in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u}). \tag{4.28b}$$

Proof. First, we note that by means of Proposition 4.9 $(\bar{\delta}_x, \bar{\delta}_u, \bar{\delta}_\pi) = (0,0,0)$ is a global optimal solution of the MPCC (4.22).
For the proof of the first assertion, we invoke Lemma 2.36 in order to see that (4.23) is equivalent to

$$\begin{bmatrix} G'(\bar{x}) & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}} & -\mathtt{L}^2 \\ \mathtt{0} & \mathtt{0} & \mathtt{I}_{\mathcal{U}} \\ \eta'(\bar{x}) & -\mathtt{E} & -\mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{U} \\ \mathcal{U} \end{pmatrix} - \begin{pmatrix} \mathcal{T}_C(G(\bar{x})) \\ \{0\} \\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^{\circ\perp} \\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^{\perp} \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{U} \\ \mathcal{U} \\ \mathcal{U} \end{pmatrix}.$$

Thus, we can apply the first statement of Proposition 3.4 in order to find multipliers $\rho \in \mathcal{W}^\star$ and $\kappa, \vartheta, \nu \in \mathcal{U}$ which satisfy (4.24d) and

$$\begin{aligned} 0 &= F_x'(\bar{x}, \bar{y}, \bar{u}) + F_y'(\bar{x}, \bar{y}, \bar{u}) \circ \mathtt{A}^{-1} \circ h'(\bar{x}) + G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\nu], \\ 0 &= F_y'(\bar{x}, \bar{y}, \bar{u}) \circ \mathtt{A}^{-1} \circ \mathtt{B} + F_u'(\bar{x}, \bar{y}, \bar{u}) + \kappa - \mathtt{E}[\nu], \\ 0 &= -\mathtt{L}^2[\kappa] + \vartheta - \nu, \\ \vartheta &\in \operatorname{cl}\big(\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^\circ - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})^\circ\big), \\ \nu &\in \operatorname{cl}\big(\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u}) - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{w} - \bar{u})\big). \end{aligned} \tag{4.29}$$

We define $\mu := \kappa + \nu$ in order to obtain the conditions (4.24e) and (4.24f). Finally, we introduce an adjoint variable $p \in \mathcal{Y}_s$ as presented in (4.24c) to state the first two conditions in (4.29) as (4.24a) and (4.24b), equivalently.
The proof of the second assertion is similar to the argumentation above. Under condition 2.a), the set $\mathcal{T}_C(G(\bar{x})) \times \{0\} \times \operatorname{gph} \mathcal{N}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}$ is SNC at $(0,0,0,0)$ and combining Lemmas 2.29 and 2.38 yields the claim if the constraint qualification

$$\left.\begin{aligned} &0 = G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\nu], \\ &0 = \kappa - \mathtt{E}[\nu], \\ &0 = -\mathtt{L}^2[\kappa] + \vartheta - \nu, \\ &\rho \in \mathcal{N}_C(G(\bar{x})),\ \kappa \in \mathcal{U}, \\ &(\vartheta, \nu) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{K}_{U_{\mathrm{ad}}}(\bar{u},\bar{w}-\bar{u})}}(0,0) \end{aligned}\right\} \Longrightarrow \rho = 0,\ \kappa = 0,\ \vartheta = 0,\ \nu = 0 \tag{4.30}$$

is satisfied. However, (4.30) is equivalent to (4.25):
If $(\rho, \mu, \nu) \in \mathcal{W}^\star \times \mathcal{U} \times \mathcal{U}$ satisfy the premise of (4.25), then $(\rho, \kappa, \vartheta, \nu)$ with $\kappa := \mu - \nu$ and $\vartheta := \nu + \mathtt{L}^2[\mu - \nu]$ satisfy the premise of (4.30). Thus, if the latter constraint qualification holds, we obtain $\rho = 0$, $\kappa = 0$,

$\vartheta = 0$, and $\nu = 0$ which leads to $\mu = \kappa + \nu = 0$, i.e. (4.25) holds. On the other hand, let (4.25) hold and assume that $(\rho', \kappa', \vartheta', \nu') \in \mathcal{W}^\star \times \mathcal{U} \times \mathcal{U} \times \mathcal{U}$ satisfy the premise of (4.30). Setting $\mu' := \kappa' + \nu'$, the triplet (ρ', μ', ν') satisfies the premise of (4.25) and, thus, we have $\rho' = 0$, $\mu' = 0$, and $\nu' = 0$. This leads to $\kappa' = \mu' - \nu' = 0$ and $\vartheta' = \nu' + \mathtt{L}^2[\kappa'] = 0$, i.e. (4.30) is valid.
Postulating 2.b) and observing that the mapping $(\delta_u, \delta_\pi) \mapsto \delta_u - \mathtt{L}^2[\delta_\pi]$ is surjective, we can adapt the proof of Proposition 3.6 and the above argumentation in order to verify the second assertion.
The proof of the theorem's third statement is analogous to the validation of its first assertion using the second statement of Proposition 3.4. □

Observe that the constraint qualifications (4.21) and (4.23) are both implied by the condition

$$\begin{bmatrix} G'(\bar{x}) & 0 \\ 0 & \mathtt{I}_{\mathcal{U}} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{U} \end{pmatrix} - \begin{bmatrix} \mathtt{I}_{\mathcal{W}} & 0 & 0 \\ 0 & \mathtt{L}^2 & 0 \\ 0 & \mathtt{I}_{\mathcal{U}} - \mathtt{L}^2 & \mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{R}_C(G(\bar{x})) \\ \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{w} - \bar{u})^{\circ\perp} \\ \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{w} - \bar{u})^{\perp} \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{U} \\ \mathcal{U} \end{pmatrix}. \tag{4.31}$$

On the other hand, (4.21) and (4.23) together do not need to imply (4.31), see Example 4.15.
Below, we comment on Theorem 4.10.

Remark 4.11. We can interpret the optimality conditions presented in the first, second, and third statement of Theorem 4.10 as W-, M-, and S-stationarity-type conditions, respectively, since they were derived via the surrogate MPCC (4.22). Since zero is a global minimizer of (4.22) for any local optimal solution of (BPP) with lower level problem (4.3) under the assumptions of the theorem, whereas any closed, convex cone is polyhedric w.r.t. $(0, 0)$, the optimality conditions of S-stationary-type (i.e. the optimality conditions presented in the theorem's third statement) possess reasonable strength, see Chapter 3. Moreover, these S-stationarity-type conditions are more restrictive than the M-stationarity-type optimality conditions (i.e. the optimality conditions presented in the theorem's second statement), see Proposition 3.11.
If U_{ad} is polyhedric w.r.t. $(\bar{u}, \bar{w} - \bar{u})$, then the operator $\mathtt{L}$ can be chosen to be $\mathtt{I}_{\mathcal{U}}$ in (4.24e), (4.26), and (4.28a), see Lemma 4.3. Especially, the optimality system of S-stationarity-type reduces to the classical system of S-stationarity.

Remark 4.12. From the proof of Theorem 4.10 the system of W-stationarity-type (4.24) possesses a solution provided the system (4.24d), (4.29) possesses a solution. The converse of this statement is also true: Let $\rho \in \mathcal{W}^\star$, $\mu, \nu \in \mathcal{U}$ and $p \in \mathcal{Y}_s$ solve (4.24) and set $\kappa := \mu - \nu$ as well as $\vartheta := \nu + \mathtt{L}^2[\kappa]$. Respecting the definition of p in (4.24c), the multipliers ρ, κ, ϑ, and ν solve (4.24d), (4.29). Carrying out the proofs of the theorem's second and third assertion in a detailed way leads to optimality systems similar to (4.29) as well, and these systems are equivalent to the optimality conditions of M- and S-stationarity-type, respectively.

Note that there exist different ways of how to derive necessary optimality conditions for the bilevel programming problem of interest. Recalling the proof of Proposition 4.1, we can replace the condition $(y, u) \in \Psi(x)$ equivalently by

$$\begin{aligned} \mathtt{A}[y] - \mathtt{B}[u] - h(x) &= 0 \\ \eta(x) - \big(\mathtt{E} + \mathtt{I}_{\mathcal{U}}\big)[u] &\in \mathcal{N}_{U_{\mathsf{ad}}}(u). \end{aligned}$$

Thus, the original bilevel programming problem is equivalent to an optimization problem whose feasible set comprises a generalized equation. Problems of this kind were considered in [90, 97, 128, 133] and many other publications. Here we do not want to discuss this approach in more detail.

Let us take a closer look at the situation where U_{ad} is a cone. Then we can rewrite the original bilevel programming problem equivalently as the following MPCC:

$$\begin{aligned} F(x, y, u) &\to \min_{x,y,u} \\ G(x) &\in C \\ \mathtt{A}[y] - \mathtt{B}[u] - h(x) &= 0 \\ u &\in U_{\mathsf{ad}} \\ \eta(x) - \big(\mathtt{E} + \mathtt{I}_{\mathcal{U}}\big)[u] &\in U_{\mathsf{ad}}^\circ \\ \big\langle \eta(x) - \big(\mathtt{E} + \mathtt{I}_{\mathcal{U}}\big)[u], u \big\rangle_{\mathcal{U}} &= 0. \end{aligned} \tag{4.32}$$

Using the results of Propositions 3.4 and 3.6, it is possible to obtain necessary optimality conditions for this problem under appropriate constraint qualifications. Observe that the following results are essentially different from those obtained in Theorem 4.10. Later, we will visualize these differences in the situation where U_{ad} equals the nonpolyhedric cone $\mathcal{S}_p^+$ in Section 5.1.

Proposition 4.13. Let $(\bar{x}, \bar{y}, \bar{u}) \in \mathcal{X} \times \mathcal{Y}_s \times \mathcal{U}$ be a local optimal solution of (BPP) with lower level (4.3) where U_{ad} is a cone. Furthermore, set $\bar{w} := \eta(\bar{x}) - \mathtt{E}[\bar{u}]$. Then the following statements are valid:

1. Assume that the constraint qualification

$$\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{U} \end{pmatrix} - \begin{pmatrix} \mathcal{R}_C(G(\bar{x})) \\ \mathcal{R}_{U_{\mathsf{ad}}}(\bar{u}) \cap \big(-\mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{w} - \bar{u})\big) \\ \mathcal{R}_{U_{\mathsf{ad}}^\circ}(\bar{w} - \bar{u}) \cap \big(-\mathcal{K}_{U_{\mathsf{ad}}^\circ}(\bar{w} - \bar{u}, \bar{u})\big) \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{U} \\ \mathcal{U} \end{pmatrix} \tag{4.33}$$

holds. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mu, \nu \in \mathcal{U}$, and $p \in \mathcal{Y}_s$ which satisfy the conditions (4.24a) - (4.24d) and

$$\begin{aligned} \mu &\in \mathrm{cl}\big(U_{\mathsf{ad}}^\circ - U_{\mathsf{ad}}^\circ \cap \{\bar{u}\}^\perp\big) \cap \{\bar{u}\}^\perp, \\ \nu &\in \mathrm{cl}\big(U_{\mathsf{ad}} - U_{\mathsf{ad}} \cap \{\bar{w} - \bar{u}\}^\perp\big) \cap \{\bar{w} - \bar{u}\}^\perp. \end{aligned}$$

2. Let $\mathcal{X}$ and $\mathcal{W}$ be reflexive, and let F be continuously Fréchet differentiable at $(\bar{x}, \bar{y}, \bar{u})$. Assume that the constraint qualification

$$\left.\begin{aligned} &0 = G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\nu], \\ &0 = \mu - (\mathtt{E} + \mathtt{I}_{\mathcal{U}})[\nu], \\ &\rho \in \mathcal{N}_C(G(\bar{x})), \\ &(\mu, \nu) \in \mathcal{N}_{\mathrm{gph}\,\mathcal{N}_{U_{\mathsf{ad}}}}(\bar{u}, \bar{w} - \bar{u}) \end{aligned}\right\} \implies \rho = 0,\ \mu = 0,\ \nu = 0$$

holds, whereas one of the following conditions is valid:

   a) C is SNC at $G(\bar{x})$ and $\mathcal{U}$ is finite-dimensional,

   b) the set $C \times \mathrm{gph}\,\mathcal{N}_{U_{\mathsf{ad}}}$ is SNC at $(G(\bar{x}), \bar{u}, \bar{w} - \bar{u})$.

   Then there are multipliers $\rho \in \mathcal{W}^\star$, $\mu, \nu \in \mathcal{U}$, and $p \in \mathcal{Y}_s$ which satisfy (4.24a) - (4.24d) and

$$(\mu, \nu) \in \mathcal{N}_{\mathrm{gph}\,\mathcal{N}_{U_{\mathsf{ad}}}}(\bar{u}, \bar{w} - \bar{u}).$$

3. Assume that the constraint qualifications (4.33) and

$$\mathrm{cl}\left(\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{U}} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{U}} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{U} \end{pmatrix} - \begin{pmatrix} \mathcal{N}_C(G(\bar{x}))_\perp \\ \mathcal{T}_{U_{\mathsf{ad}} \cap (-\mathcal{T}_{U_{\mathsf{ad}}}(\bar{u}))}(\bar{u})^{\circ\perp} \\ \mathcal{T}_{U_{\mathsf{ad}}^\circ \cap (-\mathcal{T}_{U_{\mathsf{ad}}^\circ}(\bar{w} - \bar{u}))}(\bar{w} - \bar{u})^{\circ\perp} \end{pmatrix} \right) = \begin{pmatrix} \mathcal{W} \\ \mathcal{U} \\ \mathcal{U} \end{pmatrix} \tag{4.34}$$

are satisfied. Then there are multipliers $\rho \in \mathcal{W}^\star$, $\mu, \nu \in \mathcal{U}$, and $p \in \mathcal{Y}_s$ which satisfy (4.24a) - (4.24d) and

$$\begin{aligned} \mu &\in \mathcal{K}_{U_{\mathsf{ad}}^\circ}(\bar{w} - \bar{u}, \bar{u}), \\ \nu &\in \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{w} - \bar{u}). \end{aligned}$$

Remark 4.14. Even in the situation where U_{ad} is a polyhedric cone, there might be a difference between the results in Theorem 4.10 and Proposition 4.13 although we can put $\mathtt{L} := \mathtt{I}_{\mathcal{U}}$. In order to see this, we can consider $\mathcal{U} := L^2(\mathfrak{M})$ and $U_{\mathsf{ad}} := L^2(\mathfrak{M})_0^+$ for some complete, σ-finite, and nonatomic measure space $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ such that $L^q(\mathfrak{M})$ is separable for all $q \in [1, \infty)$. Then we have from (3.19) and Proposition 3.17 that the constraint qualification (4.33) is stronger than (4.31), whereas (4.27) is equivalent to (4.34). Observe that the corresponding systems of W- and S-stationarity (i.e. the first and third optimality system

in Theorem 4.10 and Proposition 4.13) are equivalent, see Theorem 3.14 and (3.19). Clearly, the M-stationarity-type optimality conditions are not applicable due to the absence of the SNC property for the complementarity set in Lebesgue spaces, see Lemma 3.15. Thus, in this situation, the results in Theorem 4.10 are superior to those in Proposition 4.13. In Section 5.1, we will analyze the situation where $\mathcal{U} = \mathcal{S}_p$ and $U_{\mathsf{ad}} := \mathcal{S}_p^+$ are under consideration.

We want to close this section with an illustrative example.

Example 4.15. We choose $\Omega := (0,1) \subseteq \mathbb{R}$, $\mathcal{X} = \mathcal{U} = L^2(\Omega)$, $\mathcal{Y}_s := \{y \in AC^{1,2}(\Omega, \mathbb{R}) \,|\, y(0) = 0\}$,

$$C := \{x \in L^2(\Omega) \,|\, x(\omega) \in [-1,1] \text{ f.a.a. } \omega \in \Omega\},$$

as well as $U_{\mathsf{ad}} = L^2(\Omega)_0^+$ and consider the bilevel programming problem

$$\begin{aligned} \tfrac{1}{2}\int_0^1 (y(\omega)-1)^2 \mathrm{d}\omega \;&\to \min_{x,y,u} \\ x \;&\in C \\ (y,u) \;&\in \Psi(x) \end{aligned} \tag{4.35}$$

where $\Psi\colon L^2(\Omega) \rightrightarrows \mathcal{Y}_s \times L^2(\Omega)$ denotes the solution mapping of

$$\begin{aligned} \tfrac{1}{2}\int_0^1 (u(\omega)-x(\omega))^2 \mathrm{d}\omega &\to \min_{y,u} \\ y(\omega) - \int_0^\omega u(\tau)\mathrm{d}\tau - \int_0^\omega x(\tau)\mathrm{d}\tau \;&= 0 \qquad\qquad \text{a.e. on } \Omega \\ u \;&\in L^2(\Omega)_0^+. \end{aligned} \tag{4.36}$$

Clearly, the lower level dynamics can be expressed equivalently by $\nabla y = u + x$ almost everywhere on Ω and $y(0) = 0$. We identify $\mathcal{Y}_s^\star$ with its dual by means of Riesz's representation theorem. Thus, we obtain $\sigma = 1$, $\mathtt{C} = \mathtt{0}$, $\mathtt{P} = \mathtt{0}$, $\mathtt{Q} = \mathtt{I}_{L^2(\Omega)}$, $\mathtt{A} = \mathtt{I}_{\mathcal{Y}_s}$, and

$$\forall u, x \in L^2(\Omega)\, \forall \omega \in \Omega\colon \quad \mathtt{B}[u](\omega) = \int_0^\omega u(\tau)\mathrm{d}\tau, \quad h(x)(\omega) = \int_0^\omega x(\tau)\mathrm{d}\tau = \mathtt{B}[x](\omega).$$

This yields $\eta = \mathtt{I}_{L^2(\Omega)}$ and $\mathtt{E} = \mathtt{0}$.

One can easily check that a global optimal solution of (4.35) is given by $(\bar{x}, \bar{y}, \bar{u}) \in L^2(\Omega) \times \mathcal{Y}_s \times L^2(\Omega)$ defined below:

$$\forall \omega \in \Omega\colon \quad \bar{x}(\omega) = \bar{u}(\omega) = \begin{cases} 1 & \text{if } \omega \in (0, \frac{1}{2}), \\ 0 & \text{if } \omega \in [\frac{1}{2}, 1), \end{cases} \qquad \bar{y}(\omega) = \begin{cases} 2\omega & \text{if } \omega \in (0, \frac{1}{2}), \\ 1 & \text{if } \omega \in [\frac{1}{2}, 1). \end{cases}$$

As said in Remark 4.7, the projection onto $L^2(\Omega)_0^+$ is not B-differentiable from $L^2(\Omega)$ to $L^2(\Omega)$ so we cannot apply Theorem 4.10 directly. On the other hand, in order to prove Proposition 4.9 and, thus, Theorem 4.10, we only need the directional differentiability of the solution mapping to the nonsmooth equation (4.5b) which reduces to $\bar{u} = \mathrm{proj}_{U_{\mathsf{ad}}}(\bar{x})$ in our setting. Thus, its solution mapping equals $\mathrm{proj}_{U_{\mathsf{ad}}}$. Clearly, since $U_{\mathsf{ad}} = L^2(\Omega)_0^+$ is polyhedric, the directional differentiability of the projection is inherent from Lemma 4.3. Especially, we have $\mathtt{L} = \mathtt{I}_{L^2(\Omega)}$. Thus, we can apply Theorem 4.10.

Clearly, the constraint qualification (4.21) holds. Using Lemma 2.12 and the results from Section 2.3.5, we obtain

$$\begin{aligned} \mathcal{N}_C(\bar{x}) &= \left\{ v \in L^2(\Omega) \;\middle|\; \begin{array}{l} v(\omega) \geq 0 \text{ f.a.a. } \omega \in (0, \frac{1}{2}) \\ v(\omega) = 0 \text{ f.a.a. } \omega \in [\frac{1}{2}, 1) \end{array} \right\}, \\ \mathcal{N}_C(\bar{x})_\perp &= \left\{ w \in L^2(\Omega) \,|\, w(\omega) = 0 \text{ f.a.a. } \omega \in (0, \tfrac{1}{2}) \right\}, \\ \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{x} - \bar{u}) &= \left\{ w \in L^2(\Omega) \,|\, w(\omega) \geq 0 \text{ f.a.a. } \omega \in [\tfrac{1}{2}, 1) \right\}, \\ \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{x} - \bar{u})^\circ &= \left\{ v \in L^2(\Omega) \;\middle|\; \begin{array}{l} v(\omega) = 0 \text{ f.a.a. } \omega \in (0, \frac{1}{2}) \\ v(\omega) \leq 0 \text{ f.a.a. } \omega \in [\frac{1}{2}, 1) \end{array} \right\}, \\ \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{x} - \bar{u})^\perp &= \{0\}, \\ \mathcal{K}_{U_{\mathsf{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\circ\perp} &= \left\{ w \in L^2(\Omega) \,\middle|\, w(\omega) = 0 \text{ f.a.a. } \omega \in [\tfrac{1}{2}, 1) \right\}. \end{aligned}$$

From above, we easily see

$$\mathcal{N}_C(\bar{x})_\perp - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\circ\perp} - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\perp} = L^2(\Omega).$$

Using Corollary 2.37 twice, this is equivalent to

$$\begin{bmatrix} \mathtt{I}_{L^2(\Omega)} & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{L^2(\Omega)} \\ \mathtt{I}_{L^2(\Omega)} & -\mathtt{I}_{L^2(\Omega)} \end{bmatrix} \begin{pmatrix} L^2(\Omega) \\ L^2(\Omega) \end{pmatrix} - \begin{pmatrix} \mathcal{N}_C(\bar{x})_\perp \\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\circ\perp} \\ \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\perp} \end{pmatrix} = \begin{pmatrix} L^2(\Omega) \\ L^2(\Omega) \\ L^2(\Omega) \end{pmatrix}$$

and the latter condition implies (4.23) and (4.27) for (4.35). By means of Example 2.27 the objective functional of (4.35) is continuously Fréchet differentiable. Consequently, we know that the S-stationarity-type optimality conditions from Theorem 4.10 hold at $(\bar{x}, \bar{y}, \bar{u})$. Note that since we have

$$\mathcal{R}_C(\bar{x}) = \operatorname{cone}\left\{ u \in L^2(\Omega) \;\middle|\; \begin{array}{l} u(\omega) \in [-2, 0] \text{ f.a.a. } \omega \in (0, \frac{1}{2}) \\ u(\omega) \in [-1, 1] \text{ f.a.a. } \omega \in [\frac{1}{2}, 1) \end{array} \right\},$$

all functions from $\mathcal{R}_C(\bar{x})$ need to be essentially bounded, i.e. they come from $L^\infty(\Omega)$. Thus, we obtain

$$\mathcal{R}_C(\bar{x}) - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\circ\perp} - \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^{\perp} \neq L^2(\Omega)$$

which shows (apply Corollary 2.37 again) that (4.31) is violated for our problem of interest. Following Remark 4.14, the constraint qualification (4.33) is violated as well. Especially, we cannot apply the W- and S-stationarity conditions from Proposition 4.13. Since $\operatorname{gph}\mathcal{N}_{U_{\mathrm{ad}}}$ fails to be SNC everywhere, see Lemma 3.15, the corresponding M-stationarity conditions from Proposition 4.13 cannot be used as well.
In order to evaluate the optimality conditions, we interpret $F_y'(\bar{x}, \bar{y}, \bar{u})$ as a function in $AC^{1,2}(\Omega, \mathbb{R})$. Using Lemma A.5, we obtain

$$\forall \omega \in \Omega\colon \quad F_y'(\bar{x}, \bar{y}, \bar{u})(\omega) = \int_0^\omega \int_s^1 (\bar{y}(\tau) - 1) \mathrm{d}\tau \mathrm{d}s.$$

Applying the definition of the adjoint operator yields $\mathtt{B}^\star[y] = \nabla y$ for all functions $y \in \mathcal{Y}_s$. Defining a function $\psi \in L^2(\Omega)$ by $\psi := \mathtt{B}^\star[F_y'(\bar{x}, \bar{y}, \bar{u})] = h'(\bar{x})^\star[F_y'(\bar{x}, \bar{y}, \bar{u})]$, the S-stationarity conditions from the third assertion of Theorem 4.10 reduce to

$$\begin{aligned} &0 = \psi + \rho + \nu, \\ &0 = \psi + \mu - \nu, \\ &\rho \in \mathcal{N}_C(\bar{x}), \\ &\mu \in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u})^\circ, \\ &\nu \in \mathcal{K}_{U_{\mathrm{ad}}}(\bar{u}, \bar{x} - \bar{u}). \end{aligned} \tag{4.37}$$

Due to

$$\forall \omega \in \Omega\colon \quad \psi(\omega) = \int_\omega^1 (\bar{y}(s) - 1) \mathrm{d}s = \begin{cases} \omega - \omega^2 - \frac{1}{4} & \text{if } \omega \in (0, \frac{1}{2}), \\ 0 & \text{if } \omega \in [\frac{1}{2}, 1), \end{cases}$$

the system (4.37) possesses the solution $\rho := -2\psi$, $\mu = 0$, and $\nu = \psi$. Thus, the S-stationarity-type conditions are valid. ■

4.2. The KKT reformulation of the bilevel programming problem

In this section, we want to discuss the replacement of the lower level problem by necessary and sufficient optimality conditions comprising multipliers. Let us first postulate our standing assumptions which shall hold throughout the whole section.

Assumption 4.4. Let Assumption 4.1 hold. Furthermore, f and g are twice continuously Fréchet differentiable, whereas K is a nonempty, closed, convex cone. For any $x \in X_{\mathsf{ad}}$, $f(x,\cdot)$ is convex, whereas $g(x,\cdot)$ is $-K$-convex. Finally, for any $x \in X_{\mathsf{ad}}$ and any $y \in \mathcal{Y}$ satisfying $g(x,y) \in K$, the constraint qualification

$$g'_y(x,y)[\mathcal{Y}] - \mathcal{R}_K(g(x,y)) = \mathcal{Z}$$

shall hold.

These assumptions guarantee that for any $x \in X_{\mathsf{ad}}$, the condition $y \in \Psi(x)$ is equivalent to

$$\exists \lambda \in K^\circ \cap \{g(x,y)\}^\perp : \quad f'_y(x,y) + g'_y(x,y)^\star[\lambda] = 0,\ g(x,y) \in K,$$

see Lemmas 2.32 and 2.35. Thus, it is reasonable to study the surrogate problem

$$\begin{aligned} F(x,y) &\to \min_{x,y,\lambda} \\ G(x) &\in C \\ f'_y(x,y) + g'_y(x,y)^\star[\lambda] &= 0 \\ g(x,y) &\in K \\ \lambda &\in K^\circ \\ \langle \lambda, g(x,y)\rangle_{\mathcal{Z}} &= 0 \end{aligned} \tag{KKT}$$

which is an MPCC, see Chapter 3. We call (KKT) the KKT reformulation of (BPP). Similar as presented in Lemma 3.1 we easily see that KRZCQ fails to be satisfied at the feasible points of (KKT). That is why we need to invoke the theory developed in Chapter 3 in order to state applicable necessary optimality conditions and constraint qualifications. Furthermore, in view of [28], it is necessary to clarify the relationship between the two optimization models (BPP) and (KKT).

4.2.1. On the relationship between original and surrogate problem

In this section, we want to compare the models (BPP) and (KKT) w.r.t. their global and local optimal solutions. As mentioned earlier, by means of [28] we expect some delicate results here when local optimal solutions are under consideration. The core of our analysis relies on the properties of the Lagrange multiplier mapping $\Lambda \colon \mathcal{X} \times \mathcal{Y} \rightrightarrows \mathcal{Z}^\star$ of (4.1) defined below:

$$\forall x \in \mathcal{X}\, \forall y \in \mathcal{Y}: \quad \Lambda(x,y) := \{\lambda \in K^\circ \cap \{g(x,y)\}^\perp \mid f'_y(x,y) + g'_y(x,y)^\star[\lambda] = 0\}.$$

Note that due to Assumption 4.4 as well as Lemmas 2.32 and 2.35, we have $y \in \Psi(x)$ for $x \in X_{\mathsf{ad}}$ with $g(x,y) \in K$ if and only if $\Lambda(x,y) \neq \varnothing$ holds. In the following lemma, we subsume some important properties of Λ. Recall that Assumption 4.4 holds. Especially, KRZCQ holds at any lower level feasible point if the corresponding parameter is feasible for the upper level problem.

Lemma 4.16. Let $(\bar{x},\bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be feasible for (4.1) and assume that $\Lambda(\bar{x},\bar{y})$ is nonempty. Then the following assertions hold:

1. For any sequences $\{(x_k,y_k)\} \subseteq \mathcal{X} \times \mathcal{Y}$ and $\{\lambda_k\} \subseteq \mathcal{Z}^\star$ satisfying $x_k \to \bar{x}$, $y_k \to \bar{y}$, $\lambda_k \overset{\star}{\to} \bar{\lambda}$, and $\lambda_k \in \Lambda(x_k,y_k)$ for all $k \in \mathbb{N}$, we have $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$.

2. Λ is locally bounded at $(\bar{x},\bar{y})$.

3. If the set-valued mapping $\Upsilon \colon \mathcal{Y}^\star \times \mathcal{Z} \rightrightarrows \mathcal{Z}^\star$ defined by

$$\forall y^* \in \mathcal{Y}^\star\, \forall z \in \mathcal{Z}: \quad \Upsilon(y^*,z) := \{\lambda \in K^\circ \cap \{z\}^\perp \mid y^* + g'_y(\bar{x},\bar{y})^\star[\lambda] = 0\} \tag{4.38}$$

is locally upper Lipschitzian at $(f'_y(\bar{x},\bar{y}), g(\bar{x},\bar{y}))$, then Λ is locally upper Lipschitzian at $(\bar{x},\bar{y})$.

Proof. Let us start with the proof of the first assertion. Therefore, choose sequences $\{(x_k, y_k)\} \subseteq \mathcal{X} \times \mathcal{Y}$ and $\{\lambda_k\} \subseteq \mathcal{Z}^\star$ which satisfy $x_k \to \bar{x}$, $y_k \to \bar{y}$, $\lambda_k \overset{\star}{\to} \bar{\lambda}$, and $\lambda_k \in \Lambda(x_k, y_k)$ for all $k \in \mathbb{N}$. Then we have

$$f'_y(x_k, y_k) + g'_y(x_k, y_k)^\star[\lambda_k] = 0, \quad \langle \lambda_k, g(x_k, y_k) \rangle_{\mathcal{Z}} = 0, \quad \lambda_k \in K^\circ$$

for any $k \in \mathbb{N}$. Since K° is weakly* closed, we have $\bar{\lambda} \in K^\circ$ as well. The continuity of g and Lemma 2.4 lead to $\langle \bar{\lambda}, g(\bar{x}, \bar{y}) \rangle_{\mathcal{Z}} = 0$. Since f and g are continuously Fréchet differentiable, we obtain the convergencies $f'_y(x_k, y_k) \to f'_y(\bar{x}, \bar{y})$ and $g'_y(x_k, y_k) \to g'_y(\bar{x}, \bar{y})$ in $\mathcal{Y}^\star$ and $\mathbb{L}[\mathcal{Y}, \mathcal{Z}]$, respectively. We apply Lemma 2.4 once more in order to see

$$\lim_{k\to\infty} \langle g'_y(x_k, y_k)^\star[\lambda_k], y \rangle_{\mathcal{Y}} = \lim_{k\to\infty} \langle \lambda_k, g'_y(x_k, y_k)[y] \rangle_{\mathcal{Z}} = \langle \bar{\lambda}, g'_y(\bar{x}, \bar{y})[y] \rangle_{\mathcal{Z}} = \langle g'_y(\bar{x}, \bar{y})^\star[\bar{\lambda}], y \rangle_{\mathcal{Y}}$$

for any $y \in \mathcal{Y}$, i.e. $g'_y(x_k, y_k)^\star[\lambda_k] \overset{\star}{\to} g'_y(\bar{x}, \bar{y})^\star[\bar{\lambda}]$ holds. Hence, for any $y \in \mathcal{Y}$, we have

$$\langle f'_y(\bar{x}, \bar{y}) + g'_y(\bar{x}, \bar{y})^\star[\bar{\lambda}], y \rangle_{\mathcal{Y}} = \lim_{k\to\infty} \langle f'_y(x_k, y_k) + g'_y(x_k, y_k)^\star[\lambda_k], y \rangle_{\mathcal{Y}} = 0$$

which yields $f'_y(\bar{x}, \bar{y}) + g'_y(\bar{x}, \bar{y})^\star[\bar{\lambda}] = 0$. Summarizing the above calculations, we arrive at $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$. The fact that Λ is locally bounded at $(\bar{x}, \bar{y})$ follows from [17, Proposition 4.43]. The final statement of the lemma is a consequence of [17, Lemma 4.44]. □

Note that the property of Λ which we discussed in the first statement of the above lemma is stronger than its closedness. Clearly, whenever $\mathcal{Z}$ is finite-dimensional, then both properties are equivalent.
Now, we are prepared to start our analysis of the relationship between the two problems (BPP) and (KKT). For global optimal solutions, the situation is calm and parallels [28, Theorems 2.1 and 2.3]. The proof of the subsequent result is straightforward and, hence, omitted.

Theorem 4.17. If $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ is a global optimal solution of (BPP), then $(\bar{x}, \bar{y}, \lambda)$ is a global optimal solution of (KKT) for any $\lambda \in \Lambda(\bar{x}, \bar{y})$. On the other hand, if $(\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}^\star$ is a global optimal solution of (KKT), then $(\tilde{x}, \tilde{y})$ is a global optimal solution of (BPP).

Now, we take a look at the relationship of (BPP) and (KKT) w.r.t. local optimal solutions. Note that local optimality is considered w.r.t. all appearing variables of (KKT) in this thesis. In [137], the authors use a different notion of local optimality where the norm of the lower level Lagrange multiplier λ is not taken into account.
Let us start with the following observation. Its proof is, again, standard and, thus, omitted.

Theorem 4.18. If $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ is a local optimal solution of (BPP), then $(\bar{x}, \bar{y}, \lambda)$ is a local optimal solution of (KKT) for any $\lambda \in \Lambda(\bar{x}, \bar{y})$.

Clearly, the most interesting question is whether a local optimal solution of (KKT) corresponds to a local optimal solution of the original bilevel programming problem since we want to solve the surrogate problem (KKT) instead of dealing with the hierarchical optimization problem (BPP). In the situation where $\mathcal{Z}$ is finite-dimensional, we obtain the following result which parallels [28, Theorem 3.2].

Theorem 4.19. Suppose that $\mathcal{Z}$ is finite-dimensional and let $(\bar{x}, \bar{y}, \lambda) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}^\star$ be a local optimal solution of (KKT) for every $\lambda \in \Lambda(\bar{x}, \bar{y})$. Then $(\bar{x}, \bar{y})$ is a local optimal solution of (BPP).

Proof. Suppose that $(\bar{x}, \bar{y})$ is no local optimal solution of (BPP). Then there is a sequence $\{(x_k, y_k)\}$ of points from $\operatorname{gph} \Psi \cap (X_{\mathsf{ad}} \times \mathcal{Y})$ converging to $(\bar{x}, \bar{y})$ which satisfy $F(x_k, y_k) < F(\bar{x}, \bar{y})$ for all $k \in \mathbb{N}$. Since KRZCQ holds for the lower level problem at (x_k, y_k), we find $\lambda_k \in \Lambda(x_k, y_k)$ for all $k \in \mathbb{N}$. Applying Lemma 4.16, $\{\lambda_k\}$ is a bounded sequence in a finite-dimensional Banach space and, thus, contains a convergent subsequence with limit $\bar{\lambda} \in \mathcal{Z}^\star$. Due to the closedness of Λ at $(\bar{x}, \bar{y})$, see Lemma 4.16, we obtain $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$. Thus, $(\bar{x}, \bar{y}, \bar{\lambda})$ is no local optimal solution of (KKT) which contradicts the assumptions we made. □

Let $(\bar{x}, \bar{y})$ be a feasible point of (BPP). It is presented in [28, Example 3.1] in the case of finite-dimensional bilevel programming that there may exist only one element $\tilde{\lambda}$ in the set of regular Lagrange multipliers $\Lambda(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, \lambda)$ is a local optimal solution of (KKT) for all $\lambda \in \Lambda(\bar{x}, \bar{y}) \setminus \{\tilde{\lambda}\}$ whereas $(\bar{x}, \bar{y}, \tilde{\lambda})$ is not, and $(\bar{x}, \bar{y})$ is no local optimal solution of (BPP). Moreover, $\tilde{\lambda}$ may be a point in the relative interior of $\Lambda(\bar{x}, \bar{y})$.
The proof of Theorem 4.19 heavily relies on the fact that the bounded sequence of Lagrange multipliers contains a convergent subsequence which is natural when $\mathcal{Z}$ is finite-dimensional. However, this argumentation is not possible anymore if we drop the assumption on $\mathcal{Z}$ to be a Banach space of finite dimension. In the following example, we show that the situation is even worse in more general cases.

Example 4.20. For $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$ and

$$\begin{aligned} \mathcal{Z} &= C_p([0, 2\pi]) := \{u \in C([0, 2\pi]) \,|\, u(0) = u(2\pi)\}, \\ K &:= \{u \in C_p([0, 2\pi]) \,|\, u(t) \leq 0 \text{ for all } t \in [0, 2\pi]\}, \end{aligned}$$

we consider the parametric optimization problem

$$\begin{aligned} x \cdot y \;&\to \min_y \\ g(y) \;&\in K \end{aligned} \tag{4.39}$$

where $g\colon \mathbb{R}^2 \to C_p([0, 2\pi])$ is given by

$$\forall y \in \mathbb{R}^2 \, \forall t \in [0, 2\pi]\colon \quad g(y)(t) := y_1 \cos(t) + y_2 \sin(t) - 1.$$

Note that $g(y) \in K$ is equivalent to $|y|_2 \leq 1$: Assume $g(y) \in K$ and $|y|_2 > 1$. Then there is a unique $\hat{t} \in [0, 2\pi)$ such that $y/|y|_2 = (\cos(\hat{t}), \sin(\hat{t}))$ holds. This yields

$$g(y)(\hat{t}) = y_1 \cos(\hat{t}) + y_2 \sin(\hat{t}) - 1 = \frac{1}{|y|_2}(y_1^2 + y_2^2) - 1 = \frac{|y|_2^2}{|y|_2} - 1 = |y|_2 - 1 > 0$$

which is a contradiction. On the other hand, $|y|_2 \leq 1$ yields

$$g(y)(t) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \cdot \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} - 1 \leq |y|_2 \sqrt{\cos^2(t) + \sin^2(t)} - 1 = |y|_2 - 1 \leq 0$$

for any $t \in [0, 2\pi]$, i.e. $g(y) \in K$.
Thus, for any $x \in \mathbb{R}^2 \setminus \{0\}$, the unique optimal solution of (4.39) is given by $\bar{y}(x) := -x/|x|_2$, and there is a unique $t(x) \in [0, 2\pi)$ which satisfies $\bar{y}(x) = (\cos(t(x)), \sin(t(x)))$. The affine mapping g is continuous and, hence, continuously Fréchet differentiable with Fréchet derivative $\mathtt{G} \in \mathbb{L}[\mathbb{R}^2, C_p([0, 2\pi])]$ given below:

$$\forall d \in \mathbb{R}^2 \, \forall t \in [0, 2\pi]\colon \quad \mathtt{G}[d](t) = d_1 \cos(t) + d_2 \sin(t).$$

Note that we have

$$g(\bar{y}(x)) + g'(\bar{y}(x))[-\bar{y}(x)] = g(\bar{y}(x)) - \mathtt{G}[\bar{y}(x)] \equiv -1 \in \operatorname{int} K$$

which shows that KRZCQ holds for (4.39) at the optimal solution, see Remark 2.33. It is reasonable to interpret $C_p([0, 2\pi])^\star = \mathcal{M}([0, 2\pi))$ where the latter vector space contains all signed and regular measures of $([0, 2\pi), \mathfrak{B}([0, 2\pi)))$ equipped with the common variation norm of measure spaces. That is why we have

$$\forall \mu \in \mathcal{M}([0, 2\pi))\colon \quad \mathcal{G}^\star[\mu] = \left(\int_{[0,2\pi)} \cos(t) \mathrm{d}\mu(t), \int_{[0,2\pi)} \sin(t) \mathrm{d}\mu(t) \right)$$

as well as

$$K^\circ = \{\mu \in \mathcal{M}([0, 2\pi)) \,|\, \forall A \in \mathfrak{B}([0, 2\pi))\colon \; \mu(A) \geq 0\},$$

and the KKT conditions of (4.39) at $(x, \bar{y}(x))$ take the form

$$x_1 + \int_{[0,2\pi)} \cos(t) \mathrm{d}\lambda(t) = 0, \quad x_2 + \int_{[0,2\pi)} \sin(t) \mathrm{d}\lambda(t) = 0, \quad \int_{[0,2\pi)} g(\bar{y}(x))(t) \mathrm{d}\lambda(t) = 0, \quad \lambda \in K^\circ.$$

The unique solution of this system is given by $\lambda(x) := |x|_2 \, \delta_{t(x)}$ where $\delta_{t(x)}$ denotes the Dirac measure of the singleton $\{t(x)\}$.
For an arbitrary continuously differentiable function $F\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, we consider the bilevel programming problem

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ |x|_2 - 1 &= 0 \\ y &\in \Psi(x) \end{aligned} \tag{4.40}$$

where $\Psi\colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ denotes the solution set mapping of (4.39). This problem satisfies Assumption 4.4. Due to the above considerations, its KKT reformulation is given by

$$\begin{aligned} F(x,y) &\to \min_{x,y,\lambda} \\ |x|_2 - 1 &= 0 \\ \frac{x}{|x|_2} + y &= 0 \\ \lambda - |x|_2 \, \delta_{t(x)} &= 0. \end{aligned} \tag{4.41}$$

Choosing two different feasible points (x^i, y^i, λ^i), $i = 1, 2$, of (4.41), we have $t(x^1) \neq t(x^2)$ and, thus,

$$\left\|\lambda^1 - \lambda^2\right\|_{\mathcal{M}([0,2\pi))} = \left\|\delta_{t(x^1)} - \delta_{t(x^2)}\right\|_{\mathcal{M}([0,2\pi))} = 2.$$

Consequently, any feasible point of the KKT reformulation (4.41) is a local optimal solution of this problem since it is isolated. However, for reasonable objective functions F, not every feasible point will be locally optimal for (4.40). Note that this example satisfies all the assumptions of Theorem 4.19 apart from the fact that $\mathcal{Z}$ is infinite-dimensional. Choosing a sequence $\{(x_k, y_k)\} \subseteq (X_{\mathsf{ad}} \times \mathbb{R}^2) \cap \operatorname{gph} \Psi$ converging to some point $(\bar{x}, \bar{y})$ (with $\bar{x} \neq (1, 0)$), we easily see $t(x_k) \to t(\bar{x})$. Thus, for any $u \in C_p([0, 2\pi])$, we obtain

$$\begin{aligned} \lim_{k\to\infty} \langle \lambda(x_k), u \rangle_{C_p([0,2\pi])} &= \lim_{k\to\infty} \int_{[0,2\pi)} u(t) \mathrm{d}\delta_{t(x_k)}(t) = \lim_{k\to\infty} u(t(x_k)) \\ &= u(t(\bar{x})) = \int_{[0,2\pi)} u(t) \mathrm{d}\delta_{t(\bar{x})}(t) = \langle \lambda(\bar{x}), u \rangle_{C_p([0,2\pi])}, \end{aligned}$$

i.e. the bounded sequence of corresponding Lagrange multipliers $\{\lambda(x_k)\}$ converges weakly* to a multiplier $\lambda(\bar{x}) \in \Lambda(\bar{x}, \bar{y})$. However, the convergence is not strong whenever the sequence $\{x_k\}$ does not become stationary. This shows exemplary why the proof of Theorem 4.19 does not apply to the situation where $\mathcal{Z}$ is infinite-dimensional. ■

In order to avoid the difficulties depicted in the above example, we need to formulate stronger assumptions than in Theorem 4.19.

Theorem 4.21. Let K be a polyhedral cone and let $(\bar{x}, \bar{y}, \bar{\lambda}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}^\star$ be a local optimal solution of (KKT) such that $\Lambda(\bar{x}, \bar{y})$ equals the singleton $\{\bar{\lambda}\}$. Then $(\bar{x}, \bar{y})$ is a local optimal solution of (BPP).

Proof. Since K is polyhedral, it follows from [17, Theorem 2.208, Example 2.209] that the set-valued mapping Υ defined in (4.38) is locally upper Lipschitzian at $(f'_y(\bar{x}, \bar{y}), g(\bar{x}, \bar{y}))$. Thus, Λ is locally upper Lipschitz continuous at $(\bar{x}, \bar{y})$, see Lemma 4.16. Taking into account $\Lambda(\bar{x}, \bar{y}) = \{\bar{\lambda}\}$, there are constants $L > 0$ and $\varepsilon > 0$ such that

$$\forall (x,y) \in \mathbb{U}^\varepsilon_{\mathcal{X}\times\mathcal{Y}} \, \forall \lambda \in \Lambda(x,y)\colon \quad \left\|\lambda - \bar{\lambda}\right\|_{\mathcal{Z}^\star} \leq L(\|x - \bar{x}\|_{\mathcal{X}} + \|y - \bar{y}\|_{\mathcal{Y}}) \tag{4.42}$$

holds. Suppose that $(\bar{x}, \bar{y})$ is no local optimal solution of (BPP). Then there is a sequence $\{(x_k, y_k)\}$ in $(X_{\mathsf{ad}} \times \mathcal{Y}) \cap \operatorname{gph} \Psi$ converging to $(\bar{x}, \bar{y})$ which satisfies $F(x_k, y_k) < F(\bar{x}, \bar{y})$ for all $k \in \mathbb{N}$. On the other hand, we find a sequence $\{\lambda_k\} \subseteq \mathcal{Z}^\star$ such that $\lambda_k \in \Lambda(x_k, y_k)$ holds for all $k \in \mathbb{N}$, i.e. (x_k, y_k, λ_k) is feasible for (KKT) for all $k \in \mathbb{N}$. From (4.42) we derive $\left\|\lambda_k - \bar{\lambda}\right\|_{\mathcal{Z}} \leq L(\|x_k - \bar{x}\|_{\mathcal{X}} + \|y_k - \bar{y}\|_{\mathcal{Y}})$ for sufficiently large $k \in \mathbb{N}$. Hence, we have $\lambda_k \to \bar{\lambda}$. This contradicts the local optimality of $(\bar{x}, \bar{y}, \bar{\lambda})$ for (KKT). □

Theorem 4.22. Let $(\bar{x}, \bar{y}, \bar{\lambda}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}^\star$ be a local optimal solution of (KKT) where the condition

$$g'_y(\bar{x}, \bar{y})[\mathcal{Y}] - \mathcal{R}_K(g(\bar{x}, \bar{y})) \cap \{\bar{\lambda}\}_\perp = \mathcal{Z}$$

is satisfied. Then $(\bar{x}, \bar{y})$ is a local optimal solution of (BPP).

Proof. Obviously, the postulated condition equals SKRZC for the lower level problem (4.1) for fixed $x = \bar{x}$ and $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$. Recalling Remark 2.34, we have $\Lambda(\bar{x}, \bar{y}) = \{\bar{\lambda}\}$. On the other hand, [17, Proposition 4.47] shows that the set-valued mapping Υ defined in (4.38) is locally upper Lipschitzian at $(\bar{x}, \bar{y})$. The remaining part of the argumentation parallels the proof of Theorem 4.21. □

Recalling Example 4.20 where any feasible point of the KKT reformulation was already a local optimal solution, we obtain that the cone K defined therein is nonpolyhedral and that SKRZC does not hold at the feasible points of the corresponding lower level problem (4.39).
Combining Theorems 4.17, 4.18, and 4.22 yields the following corollary.

Corollary 4.23. Suppose that for any point $(x, y) \in X_{\text{ad}} \times \mathcal{Y}$ which satisfies $g(x, y) \in K$, the operator $g'_y(x, y)$ is surjective. Then the mapping Λ is at most singleton-valued on $X_{\text{ad}} \times \mathcal{Y}$. Furthermore, a point $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ is a global (local) optimal solution of (BPP) if and only if there is $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, \bar{\lambda})$ is a global (local) optimal solution of (KKT).

4.2.2. Necessary optimality conditions

Here we are going to apply the results obtained in Chapter 3 to the surrogate problem (KKT) in order to derive necessary optimality conditions for (BPP). Therefore, we assume that $\mathcal{Z}$ is a reflexive Banach space in order to ensure the symmetry of the complementarity condition. Let us define the following stationarity notions for the bilevel programming problem.

Definition 4.1. A feasible point $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ of the bilevel programming problem (BPP) is called W-stationary (M-stationary, S-stationary) for (BPP) if there is some $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, \bar{\lambda})$ is a W-stationary (M-stationary, S-stationary) point of (KKT) in the sense of Definition 3.1 (Definition 3.2, Definition 3.1).

We need to mention that our definition of the various stationarity notions differs from the definitions postulated in [33] where the authors demand that $(\bar{x}, \bar{y}, \lambda)$ satisfies the corresponding stationarity conditions of (KKT) for all $\lambda \in \Lambda(\bar{x}, \bar{y})$. On the one hand, this stronger notion seems to respect the results in Theorems 4.17 and 4.19. On the other hand, we may run into some trouble w.r.t. appropriate constraint qualifications, see the forthcoming Remark 4.29 and Example 4.30. Additionally, as we revealed in the last section, the local equivalence of (BPP) and (KKT) may be generally guaranteed only in the case where the lower level multiplier is unique a priori whenever $\mathcal{Z}$ is infinite-dimensional, see Theorems 4.21 and 4.22 as well as Corollary 4.23, and in this case, our notions of stationarity do not differ from the ones in [33]. Hence, we rely on the weaker stationarity notions from Definition 4.1.
In the following proposition, we state equivalent notions of the W-, M-, and S-stationarity conditions of the bilevel programming problem which easily follow from Definition 4.1.

Proposition 4.24. Let $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a feasible point of the bilevel programming problem (BPP).

1. The point $(\bar{x}, \bar{y})$ is W-stationary for (BPP) if and only if there are $\lambda \in \Lambda(\bar{x}, \bar{y})$, $\rho \in \mathcal{W}^\star$, $\kappa \in \mathcal{Y}^{\star\star}$, and

$\mu \in \mathcal{Z}^\star$ which satisfy the following conditions:

$$0 = F'_x(\bar{x},\bar{y}) + G'(\bar{x})^\star[\rho] + f^{(2)}_{yx}(\bar{x},\bar{y})^\star[\kappa] + \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y}) \right\rangle^\star_{\mathcal{Z}}[\kappa] + g'_x(\bar{x},\bar{y})^\star[\mu], \tag{4.43a}$$

$$0 = F'_y(\bar{x},\bar{y}) + f^{(2)}_{yy}(\bar{x},\bar{y})^\star[\kappa] + \left\langle \lambda, g^{(2)}_{yy}(\bar{x},\bar{y}) \right\rangle^\star_{\mathcal{Z}}[\kappa] + g'_y(\bar{x},\bar{y})^\star[\mu], \tag{4.43b}$$

$$\rho \in \mathcal{N}_C(G(\bar{x})), \tag{4.43c}$$

$$\mu \in \mathrm{cl}(K^\circ - K^\circ \cap \{g(\bar{x},\bar{y})\}^\perp) \cap \{g(\bar{x},\bar{y})\}^\perp, \tag{4.43d}$$

$$-g'_y(\bar{x},\bar{y})^{\star\star}[\kappa] \in \mathrm{cl}(K - K \cap \{\lambda\}^\perp) \cap \{\lambda\}^\perp. \tag{4.43e}$$

2. The point $(\bar{x},\bar{y})$ is M-stationary for (BPP) if and only if there are $\lambda \in \Lambda(\bar{x},\bar{y})$, $\rho \in \mathcal{W}^\star$, $\kappa \in \mathcal{Y}^{\star\star}$, and $\mu \in \mathcal{Z}^\star$ which satisfy the conditions (4.43a) - (4.43c) and

$$(\mu, -g'_y(\bar{x},\bar{y})^{\star\star}[\kappa]) \in \mathcal{N}_{\mathrm{gph}\,\mathcal{N}_K}(g(\bar{x},\bar{y}),\lambda). \tag{4.44}$$

3. The point $(\bar{x},\bar{y})$ is S-stationary for (BPP) if and only if there are $\lambda \in \Lambda(\bar{x},\bar{y})$, $\rho \in \mathcal{W}^\star$, $\kappa \in \mathcal{Y}^{\star\star}$, and $\mu \in \mathcal{Z}^\star$ which satisfy the conditions (4.43a) - (4.43c) and

$$\begin{aligned} &\mu \in \mathcal{K}_{K^\circ}(\lambda, g(\bar{x},\bar{y})), \\ &-g'_y(\bar{x},\bar{y})^{\star\star}[\kappa] \in \mathcal{K}_K(g(\bar{x},\bar{y}),\lambda). \end{aligned} \tag{4.45}$$

In the following remark, we comment on the above notation.

Remark 4.25. For fixed $(\bar{x},\bar{y}) \in \mathcal{X}\times\mathcal{Y}$, we interpret $f^{(2)}_{yx}(\bar{x},\bar{y}) \in \mathbb{L}[\mathcal{X},\mathcal{Y}^\star]$ and $f^{(2)}_{yy}(\bar{x},\bar{y}) \in \mathbb{L}[\mathcal{Y},\mathcal{Y}^\star]$. Now, fix some $\lambda \in \mathcal{Z}^\star$. Then we have

$$\forall \delta_x \in \mathcal{X}\colon \quad \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y}) \right\rangle_{\mathcal{Z}}[\delta_x] = \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y})[\cdot,\delta_x] \right\rangle_{\mathcal{Z}} = g^{(2)}_{yx}(\bar{x},\bar{y})[\cdot,\delta_x]^\star[\lambda] \in \mathcal{Y}^\star$$

by definition and, thus, for any $\kappa \in \mathcal{Y}^{\star\star}$ and $\delta_x \in \mathcal{X}$, we obtain

$$\left\langle \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y}) \right\rangle^\star_{\mathcal{Z}}[\kappa], \delta_x \right\rangle_{\mathcal{X}} = \left\langle \kappa, g^{(2)}_{yx}(\bar{x},\bar{y})[\cdot,\delta_x]^\star[\lambda] \right\rangle_{\mathcal{Y}^\star} = \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y})[\cdot,\delta_x]^{\star\star}[\kappa] \right\rangle_{\mathcal{Z}}.$$

Analogously, we interpret the operator $\left\langle \lambda, g^{(2)}_{yy}(\bar{x},\bar{y}) \right\rangle^\star_{\mathcal{Z}} \in \mathbb{L}[\mathcal{Y}^{\star\star},\mathcal{Y}^\star]$.
Let $\mathcal{Y}$ be reflexive. Then we deduce $\mathcal{Y} \cong \mathcal{Y}^{\star\star}$ and

$$\forall \delta_x \in \mathcal{X}\colon \quad f^{(2)}_{yx}(\bar{x},\bar{y})^\star[\kappa][\delta_x] = f^{(2)}_{yx}(\bar{x},\bar{y})[\kappa,\delta_x]$$

as well as

$$\forall \delta_x \in \mathcal{X}\colon \quad \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y}) \right\rangle^\star_{\mathcal{Z}}[\kappa][\delta_x] = \left\langle \lambda, g^{(2)}_{yx}(\bar{x},\bar{y})[\kappa,\delta_x] \right\rangle_{\mathcal{Z}}$$

are obtained for any $\kappa \in \mathcal{Y}$ since $g^{(2)}_{yx}(\bar{x},\bar{y})[\cdot,\delta_x]^{\star\star} = g^{(2)}_{yx}(\bar{x},\bar{y})[\cdot,\delta_x]$ holds for all $\delta_x \in \mathcal{X}$. Especially, we have

$$\forall \delta_y \in \mathcal{Y}\colon \quad f^{(2)}_{yy}(\bar{x},\bar{y})^\star[\kappa][\delta_y] = f^{(2)}_{yy}(\bar{x},\bar{y})[\kappa,\delta_y] = f^{(2)}_{yy}(\bar{x},\bar{y})[\kappa][\delta_y]$$

as well as

$$\begin{aligned} \forall \delta_y \in \mathcal{Y}\colon \quad \left\langle \lambda, g^{(2)}_{yy}(\bar{x},\bar{y}) \right\rangle^\star_{\mathcal{Z}}[\kappa][\delta_y] &= \left\langle \lambda, g^{(2)}_{yy}(\bar{x},\bar{y})[\kappa,\delta_y] \right\rangle_{\mathcal{Z}} \\ &= \left\langle \lambda, g^{(2)}_{yy}(\bar{x},\bar{y})[\delta_y,\kappa] \right\rangle_{\mathcal{Z}} = \left\langle \lambda, g^{(2)}_{yy}(\bar{x},\bar{y}) \right\rangle_{\mathcal{Z}}[\kappa][\delta_y]. \end{aligned}$$

Furthermore, we can replace $g'_y(\bar{x},\bar{y})^{\star\star}$ in (4.43e), (4.44), and (4.45) by $g'_y(\bar{x},\bar{y})$.

First, we formulate constraint qualifications which ensure that local optimal solutions of (BPP) are W- or S-stationary. The following result is the counterpart of Proposition 3.4 for (KKT).

Theorem 4.26. Let $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a local solution of (BPP).

1. Assume that there exists $\lambda \in \Lambda(\bar{x}, \bar{y})$ such that the constraint qualification

$$\mathtt{Q}(\bar{x}, \bar{y}, \lambda) \begin{pmatrix} \mathcal{X} \\ \mathcal{Y} \end{pmatrix} - \mathtt{P}(\bar{x}, \bar{y}) \begin{pmatrix} \mathcal{R}_C(G(\bar{x})) \\ \mathcal{R}_{K^\circ}(\lambda) \cap \left(-\mathcal{K}_{K^\circ}(\lambda, g(\bar{x}, \bar{y}))\right) \\ \mathcal{R}_K(g(\bar{x}, \bar{y})) \cap \left(-\mathcal{K}_K(g(\bar{x}, \bar{y}), \lambda)\right) \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{Y}^\star \\ \mathcal{Z} \end{pmatrix} \tag{4.46}$$

is satisfied where $\mathtt{Q}(\bar{x}, \bar{y}, \lambda) \in \mathbb{L}[\mathcal{X} \times \mathcal{Y}, \mathcal{W} \times \mathcal{Y}^\star \times \mathcal{Z}]$ and $\mathtt{P}(\bar{x}, \bar{y}) \in \mathbb{L}[\mathcal{W} \times \mathcal{Z}^\star \times \mathcal{Z}, \mathcal{W} \times \mathcal{Y}^\star \times \mathcal{Z}]$ are defined as stated below:

$$\mathtt{Q}(\bar{x}, \bar{y}, \lambda) := \begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ f^{(2)}_{yx}(\bar{x}, \bar{y}) + \left\langle \lambda, g^{(2)}_{yx}(\bar{x}, \bar{y}) \right\rangle_{\mathcal{Z}} & f^{(2)}_{yy}(\bar{x}, \bar{y}) + \left\langle \lambda, g^{(2)}_{yy}(\bar{x}, \bar{y}) \right\rangle_{\mathcal{Z}} \\ g'_x(\bar{x}, \bar{y}) & g'_y(\bar{x}, \bar{y}) \end{bmatrix},$$

$$\mathtt{P}(\bar{x}, \bar{y}) := \begin{bmatrix} \mathtt{I}_{\mathcal{W}} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & -g'_y(\bar{x}, \bar{y})^\star & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{I}_{\mathcal{Z}} \end{bmatrix}.$$

   Then $(\bar{x}, \bar{y})$ is W-stationary for (BPP).

2. Assume that there exists $\lambda \in \Lambda(\bar{x}, \bar{y})$ such that the constraint qualifications (4.46) and

$$\operatorname{cl}\left(\mathtt{Q}(\bar{x}, \bar{y}, \lambda) \begin{pmatrix} \mathcal{X} \\ \mathcal{Y} \end{pmatrix} - \mathtt{P}(\bar{x}, \bar{y}) \begin{pmatrix} \mathcal{N}_C(G(\bar{x}))_\perp \\ \mathcal{T}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(\lambda))}(\lambda)^{\circ\perp} \\ \mathcal{T}_{K \cap (-\mathcal{T}_K(g(\bar{x}, \bar{y})))}(g(\bar{x}, \bar{y}))^{\circ\perp} \end{pmatrix}\right) = \begin{pmatrix} \mathcal{W} \\ \mathcal{Y}^\star \\ \mathcal{Z} \end{pmatrix} \tag{4.47}$$

   are satisfied. Then $(\bar{x}, \bar{y})$ is S-stationary for (BPP).

3. Assume that there exists $\lambda \in \Lambda(\bar{x}, \bar{y})$ such that $\mathtt{Q}(\bar{x}, \bar{y}, \lambda)$ is surjective. Then $(\bar{x}, \bar{y})$ is S-stationary for (BPP).

Proof. The proof of the first two statements follows from Proposition 3.4 and the cancellation rule for constraint qualifications in product spaces, see Corollary 2.37. The final assertion obviously follows from the first two. □

Similar as stated above, we can adapt Proposition 3.6 in order to find constraint qualifications ensuring that local optimal solutions of (BPP) are M-stationary. Be aware that the singleton $\{0\}$ in $\mathcal{Y}^\star$, which appears on the right hand side of the constraints in (KKT), is SNC if and only if $\mathcal{Y}$ is finite-dimensional. This has to be taken into account during the formulation of appropriate constraint qualifications. In the subsequent theorem, we only present two possible versions of the qualification condition. One may check the proof of Proposition 3.6 in order to see how other constraint qualifications can be constructed.

Theorem 4.27. Let $\mathcal{W}$, $\mathcal{X}$, as well as $\mathcal{Y}$ be reflexive Banach spaces and let $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a local optimal solution of (BPP) where F is continuously Fréchet differentiable. Assume that there exists a multiplier $\lambda \in \Lambda(\bar{x}, \bar{y})$ such that the constraint qualification

$$\left.\begin{aligned} &0 = \mathtt{Q}(\bar{x}, \bar{y}, \lambda)^\star \begin{pmatrix} \rho \\ \kappa \\ \mu \end{pmatrix}, \\ &\rho \in \mathcal{N}_C(G(\bar{x})), \\ &(\mu, -g'_y(\bar{x}, \bar{y})[\kappa]) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_K}(g(\bar{x}, \bar{y}), \lambda) \end{aligned}\right\} \implies \rho = 0,\ \kappa = 0,\ \mu = 0 \tag{4.48}$$

is satisfied and $C \times \operatorname{gph} \mathcal{N}_K$ is SNC at $(G(\bar{x}), g(\bar{x}, \bar{y}), \lambda)$. Then any of the conditions stated below is sufficient for $(\bar{x}, \bar{y})$ to be M-stationary for (BPP):

(a) $\mathcal{Y}$ is finite-dimensional,

(b) the operator

$$\left[f_{yx}^{(2)}(\bar{x},\bar{y}) + \left\langle \lambda, g_{yx}^{(2)}(\bar{x},\bar{y})\right\rangle_{\mathcal{Z}} \quad f_{yy}^{(2)}(\bar{x},\bar{y}) + \left\langle \lambda, g_{yy}^{(2)}(\bar{x},\bar{y})\right\rangle_{\mathcal{Z}} \quad g_y'(\bar{x},\bar{y})^\star\right] \in \mathbb{L}[\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}^\star,\mathcal{Y}^\star]$$

is surjective.

Proof. We remark that the constraint qualification (4.48) is equivalent to

$$\left.\begin{aligned} &0 = \mathtt{Q}(\bar{x},\bar{y},\lambda)^\star \begin{pmatrix}\rho\\ \kappa\\ \mu\end{pmatrix},\\ &\rho \in \mathcal{N}_C(G(\bar{x})),\\ &0 = g_y'(\bar{x},\bar{y})[\kappa] + \nu,\\ &(\mu,\nu) \in \mathcal{N}_{\mathrm{gph}\,\mathcal{N}_K}(g(\bar{x},\bar{y}),\lambda) \end{aligned}\right\} \Longrightarrow \rho = 0,\ \kappa = 0,\ \mu = 0,\ \nu = 0.$$

Thus, under the first postulated condition of the theorem, the assertion follows combining Lemmas 2.29 and 2.38. If the second condition is valid, we can show the assertion similarly as we did in the proof of Proposition 3.6. □

Furthermore, we obtain the following result from Corollary 3.21.

Theorem 4.28. Let $\mathcal{W}$, $\mathcal{X}$, as well as $\mathcal{Y}$ be reflexive Banach spaces and let $(\bar{x},\bar{y}) \in \mathcal{X}\times\mathcal{Y}$ be a local optimal solution of (BPP) where F is continuously Fréchet differentiable. Furthermore, let K be polyhedral. Assume that there exists $\lambda \in \Lambda(\bar{x},\bar{y})$ such that the constraint qualification (4.46) is satisfied and suppose that $C \times \mathrm{gph}\,\mathcal{N}_K$ is SNC at $(G(\bar{x}), g(\bar{x},\bar{y}),\lambda)$. Then $(\bar{x},\bar{y})$ is M-stationary for (BPP) provided one of the conditions (a) and (b) from Theorem 4.27 is valid.

Obviously, the constraint qualifications in the Theorems 4.26 and 4.27 may depend not only on the feasible point of (BPP) but also on the choice of the corresponding lower level Lagrange multiplier. This may cause some trouble when checking whether these constraint qualifications hold or not.

Remark 4.29. Let $(\bar{x},\bar{y}) \in \mathcal{X}\times\mathcal{Y}$ be a local optimal solution of (BPP) where $\Lambda(\bar{x},\bar{y})$ is not a singleton. Then it may happen that the constraint qualifications (4.46), (4.47), and (4.48) hold for some but not all multipliers from $\Lambda(\bar{x},\bar{y})$. This means that the choice of the multipliers in the Theorems 4.26 and 4.27 is of essential importance.

Example 4.30. For $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathbb{R}^2$ and $K = -\mathbb{R}^{2,+}_0$, we consider the bilevel programming problem

$$\begin{aligned} \tfrac{1}{2}|x|_2^2 + \tfrac{1}{2}|y|_2^2 &\to \min_{x,y}\\ x &\in \Psi(x) \end{aligned}$$

where $\Psi\colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ represents the solution mapping of the parametric optimization problem

$$\begin{aligned} \tfrac{1}{2}(y_2-1)^2 &\to \min_{y}\\ y_1^2 + y_2 - x_1 &\le 0\\ y_2 - x_2 &\le 0. \end{aligned}$$

The lower level problem is convex w.r.t. y and its feasible set possesses interior points for any choice of $x \in \mathbb{R}^2$ which yields that the postulated constraint qualification in Assumption 4.4 holds everywhere. Obviously, the unique global optimal solution of the presented bilevel programming problem is given by $(\bar{x},\bar{y}) := (0,0)$. One can easily check that the corresponding set of lower level Lagrange multipliers is given by $\Lambda(\bar{x},\bar{y}) = \mathrm{conv}\{(1,0),(0,1)\}$. For any $\lambda \in \Lambda(\bar{x},\bar{y})$, we obtain

$$\mathtt{Q}(\bar{x},\bar{y},\lambda) = \begin{pmatrix} 0 & 0 & 2\lambda_1 & 0\\ 0 & 0 & 0 & 1\\ -1 & 0 & 0 & 1\\ 0 & -1 & 0 & 1 \end{pmatrix}, \quad \mathtt{P}(\bar{x},\bar{y}) = \begin{pmatrix} 0 & 0 & 0 & 0\\ -1 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{pmatrix}$$

which shows that the constraint qualifications (4.46) and (4.47) fail to hold for $\tilde{\lambda} := (0,1)$ and are valid for all the other multipliers from $\Lambda(\bar{x},\bar{y})\setminus\{\tilde{\lambda}\}$. Especially, $(\bar{x},\bar{y})$ is S-stationary for the bilevel programming problem. For any $\lambda \in \Lambda(\bar{x},\bar{y})$, the corresponding constraint qualification (4.48) reduces to

$$\left.\begin{aligned} 0 &= -\mu_1,\\ 0 &= -\mu_2,\\ 0 &= 2\lambda_1\kappa_1,\\ 0 &= \kappa_2 + \mu_1 + \mu_2,\\ \left(\mu, \begin{pmatrix} -\kappa_2 \\ -\kappa_2 \end{pmatrix}\right) &\in \mathcal{N}_{\operatorname{gph}\mathcal{N}_K}(0,\lambda) \end{aligned}\right\} \implies \kappa = 0,\ \mu = 0.$$

Clearly, this condition is satisfied for any choice of $\lambda \in \Lambda(\bar{x},\bar{y}) \setminus \{\tilde{\lambda}\}$ and violated for $\tilde{\lambda}$. ■

If the lower level constraints are of special affine type, i.e. if there are a linear operator $\mathtt{B} \in \mathbb{L}[\mathcal{Y},\mathcal{Z}]$ and a continuously Fréchet differentiable function $h\colon \mathcal{X} \to \mathcal{Z}$ such that

$$\forall x \in \mathcal{X}\, \forall y \in \mathcal{Y}\colon \quad g(x,y) := h(x) + \mathtt{B}[y]$$

is satisfied, the situation is more comfortable. Here the operator $\mathtt{Q}$ defined in Theorem 4.26 does not depend on the lower level Lagrange multiplier but only on the feasible point of the bilevel programming problem. Thus, the surjectivity of $\mathtt{Q}$ is a handy constraint qualification implying local minima of (BPP) to be S-stationary, see Theorem 4.26. In the following example, we study a situation where $\mathtt{Q}$ does not even depend on the choice of the feasible point of (BPP).

Example 4.31. Let $\mathcal{Y}$ be reflexive. Suppose that there are linear operators $\mathtt{A} \in \mathbb{L}[\mathcal{X},\mathcal{Z}]$, $\mathtt{B} \in \mathbb{L}[\mathcal{Y},\mathcal{Z}]$, $\mathtt{C} \in \mathbb{L}[\mathcal{X},\mathcal{W}]$, as well as $\mathtt{S} \in \mathbb{L}[\mathcal{X},\mathcal{Y}^\star]$, a self-adjoint and elliptic operator $\mathtt{R} \in \mathbb{L}[\mathcal{Y},\mathcal{Y}^\star]$, and vectors $c \in \mathcal{W}$ as well as $d \in \mathcal{Z}$ such that the mappings G, f, and g take the following form:

$$\forall x \in \mathcal{X}\, \forall y \in \mathcal{Y}\colon \quad G(x) := \mathtt{C}[x] - c, \qquad f(x,y) := \tfrac{1}{2}\left\langle \mathtt{R}[y], y\right\rangle_{\mathcal{Y}} + \left\langle \mathtt{S}[x], y\right\rangle_{\mathcal{Y}}, \qquad g(x,y) := \mathtt{A}[x] + \mathtt{B}[y] - d.$$

Furthermore, let the constraint qualification

$$\mathtt{B}[\mathcal{Y}] - \mathcal{R}_K(\mathtt{A}[x] + \mathtt{B}[y] - d) = \mathcal{Z}$$

be satisfied at any point $(x,y) \in \mathcal{X}\times\mathcal{Y}$ where $\mathtt{A}[x] + \mathtt{B}[y] - d \in K$ is valid. Then we easiliy see from Example 2.28 that Assumption 4.4 is satisfied. If the operator

$$\mathtt{Q} := \begin{bmatrix} \mathtt{C} & \mathtt{O} \\ \mathtt{S} & \mathtt{R} \\ \mathtt{A} & \mathtt{B} \end{bmatrix} \in \mathbb{L}[\mathcal{X}\times\mathcal{Y}, \mathcal{W}\times\mathcal{Y}^\star\times\mathcal{Z}]$$

is surjective, any local optimal solution $(\bar{x},\bar{y}) \in \mathcal{X}\times\mathcal{Y}$ of the corresponding bilevel programming problem is S-stationary, i.e. there are multipliers $\lambda \in \Lambda(\bar{x},\bar{y})$, $\rho \in \mathcal{W}^\star$, $\kappa \in \mathcal{Y}$, and $\mu \in \mathcal{Z}^\star$ which solve the system

$$\begin{aligned} 0 &= F_x'(\bar{x},\bar{y}) + \mathtt{C}^\star[\rho] + \mathtt{S}^\star[\kappa] + \mathtt{A}^\star[\mu],\\ 0 &= F_y'(\bar{x},\bar{y}) + \mathtt{R}[\kappa] + \mathtt{B}^\star[\mu],\\ \rho &\in \mathcal{N}_C(\mathtt{C}[\bar{x}] - c),\\ \mu &\in \mathcal{K}_{K^\circ}(\lambda, \mathtt{A}[\bar{x}] + \mathtt{B}[\bar{y}] - d),\\ -\mathtt{B}[\kappa] &\in \mathcal{K}_K(\mathtt{A}[\bar{x}] + \mathtt{B}[\bar{y}] - d, \lambda). \end{aligned}$$

In the absence of upper level constraints, one can fix $\mathtt{A} = \mathtt{O}$ and assume the surjectivity of $\mathtt{B}$ and $\mathtt{S}$ in order to ensure the surjectivity of $\mathtt{Q}$. This setting is called ample parameterization, see [38].
Note that the (appropriately) discretized obstacle problem (1.3) possesses an ample-parameterized lower level problem and, thus, its local optimal solutions are S-stationary in the absence of control constraints. Let us take a short look at the infinite-dimensional setting. Then we have $\mathcal{X} = L^2(\Omega)$ as well as $\mathcal{Y} = H_0^1(\Omega)$ for some bounded domain $\Omega \subseteq \mathbb{R}^d$, and $\mathtt{S}$ is given by the natural embedding from $L^2(\Omega)$ into $H^{-1}(\Omega)$. Since $L^2(\Omega)$ is dense in $H^{-1}(\Omega)$, $\mathtt{S}$ possesses a dense range but is not surjective, i.e. the lower level program of the (infinite-dimensional) obstacle problem cannot be ample-parameterized. However, its local optimal solutions are always S-stationary in the absence of control constraints. This classical result is known from [88, Section 4]. ■

4.3. The optimal value reformulation of the bilevel programming problem

In this section, we consider the so-called optimal value reformulation (OV) of the general bilevel programming model (BPP) in more detail. As we pointed out earlier, from

$$\forall x \in \mathcal{X}: \quad \Psi(x) = \{y \in \mathcal{Y} \,|\, g(x,y) \in K,\ f(x,y) \leq \varphi(x)\}$$

it is easily seen that (BPP) and (OV) are equivalent optimization problems. Therein, φ denotes the optimal value function of (4.1) defined in (4.2). Our main issue is to illustrate the difficulties arising from this equivalent reformulation and to depict some approaches to overcome these problems. Finally, we state KKT-type necessary optimality conditions for (BPP).
It is well-known from parametric optimization, see e.g. [7, 90, 91, 92, 93], that the function φ does not need to be smooth. Thus, (OV) is a nonsmooth optimization problem in general. Secondly, the problem (OV) is likely to be nonconvex due to the following observation: If f is fully convex and g is $-K$-convex, then φ is convex (one can easily adapt the proof of [40, Proposition 2.1]). Thus, (OV) possesses a constraint function given by the difference of two convex functions which is nonconvex in general. Another disadvantage of this surrogate problem is its inherent lack of regularity. If φ is continuously Fréchet differentiable at the reference point, then KRZCQ is violated. In the case where φ is at least locally Lipschitz continuous at the point of interest, the constraint qualification (2.21) is likely to fail as well. Thus, the standard constraint qualifications implying that local minimizers satisfy KKT-type optimality conditions do not hold.

Lemma 4.32. Let $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a feasible point of (OV). Then the following assertions are true.

1. Suppose that φ is continuously Fréchet differentiable at $\bar{x}$. Then KRZCQ for (OV) is violated at $(\bar{x}, \bar{y})$.

2. Let $\mathcal{X}$ and $\mathcal{Y}$ be reflexive. Assume that φ is locally Lipschitz continuous at $\bar{x}$ and let the constraint qualification
$$\begin{bmatrix} G'(\bar{x}) & 0 \\ g'_x(\bar{x},\bar{y}) & g'_y(\bar{x},\bar{y}) \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{Y} \end{pmatrix} - \begin{pmatrix} \mathcal{R}_C(G(\bar{x})) \\ \mathcal{R}_K(g(\bar{x},\bar{y})) \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{Z} \end{pmatrix} \tag{4.49}$$
be satisfied. Then for $M := \{(x,y) \in \mathcal{X} \times \mathcal{Y} \,|\, G(x) \in C,\ g(x,y) \in K\}$, the constraint qualification
$$\left.\begin{array}{l} 0 \in \theta \partial(f-\varphi)(\bar{x},\bar{y}) + \mathcal{N}_M(\bar{x},\bar{y}), \\ \theta \geq 0 \end{array}\right\} \Longrightarrow \theta = 0$$
from Lemma 2.39 is violated.

Proof. By definition of φ, for any point $(x,y) \in \mathcal{X} \times \mathcal{Y}$ satisfying $g(x,y) \in K$, we have $f(x,y) - \varphi(x) \geq 0$. Thus, since $(\bar{x}, \bar{y})$ is feasible for (OV), $(\bar{x}, \bar{y})$ is a global optimal solution of

$$\begin{array}{rl} f(x,y) - \varphi(x) & \to \min\limits_{x,y} \\ g(x,y) & \in K. \end{array} \tag{4.50}$$

Let φ be continuously Fréchet differentiable at $\bar{x}$. Suppose that KRZCQ holds at $(\bar{x}, \bar{y})$ for (OV). Then the constraint qualifications
$$\begin{bmatrix} (f-\varphi)'(\bar{x},\bar{y}) \\ g'(\bar{x},\bar{y}) \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{Y} \end{pmatrix} - \begin{pmatrix} -\mathbb{R}^+_0 \\ \mathcal{R}_K(g(\bar{x},\bar{y})) \end{pmatrix} = \begin{pmatrix} \mathbb{R} \\ \mathcal{Z} \end{pmatrix} \tag{4.51}$$
and
$$g'(\bar{x},\bar{y})[\mathcal{X} \times \mathcal{Y}] - \mathcal{R}_K(g(\bar{x},\bar{y})) = \mathcal{Z} \tag{4.52}$$
are valid as well. The global optimality of $(\bar{x}, \bar{y})$ for (4.50) and (4.52) imply the existence of $\lambda \in \mathcal{Z}^\star$ which satisfies
$$(f-\varphi)'(\bar{x},\bar{y}) + g'(\bar{x},\bar{y})^\star[\lambda] = 0,\ \lambda \in \mathcal{N}_K(g(\bar{x},\bar{y})),$$

see Lemma 2.32. Therefore, the constraint qualification

$$\left.\begin{aligned} 0 &= \theta(f-\varphi)'(\bar{x},\bar{y}) + g'(\bar{x},\bar{y})^\star[\lambda], \\ \theta &\geq 0,\ \lambda \in \mathcal{N}_K(g(\bar{x},\bar{y})) \end{aligned}\right\} \implies \theta = 0,\ \lambda = 0$$

is violated. Following Remark 2.33, (4.51) is violated as well. This is a contradiction.
Now, assume that φ is locally Lipschitz continuous at $\bar{x}$ and that the constraint qualification (4.49) is valid. Then we have

$$\mathcal{N}_M(\bar{x},\bar{y}) = G'(\bar{x})^\star\big[\mathcal{N}_C(G(\bar{x}))\big] \times \{0\} + g'(\bar{x},\bar{y})^\star\big[\mathcal{N}_K(g(\bar{x},\bar{y}))\big]$$

from Lemma 2.31. Since (4.49) implies (4.52), we have

$$0 \in \partial(f-\varphi)(\bar{x},\bar{y}) + g'(\bar{x},\bar{y})^\star\big[\mathcal{N}_K(g(\bar{x},\bar{y}))\big]$$

from Lemmas 2.29 as well as 2.31 and the fact that $(\bar{x},\bar{y})$ solves (4.50). Since we obtain the inclusion $g'(\bar{x},\bar{y})^\star\big[\mathcal{N}_K(g(\bar{x},\bar{y}))\big] \subseteq \mathcal{N}_M(\bar{x},\bar{y})$ from above, the second statement of the lemma is true as well. □

In order to overcome the inherent lack of regularity when facing (OV), we use a penalization approach introduced in [138]. Therefore, we take a look at

$$\begin{aligned} F(x,y) + \kappa(f(x,y)-\varphi(x)) &\to \min_{x,y} \\ G(x) &\in C \\ g(x,y) &\in K \end{aligned} \tag{OV$_\kappa$}$$

where $\kappa > 0$ is the penalization parameter. The validity of the so-called partial calmness condition at a local minimum of (OV), see [138, Definition 3.1] for the finite-dimensional case and the forthcoming Definition 4.2 for the general case, guarantees that for some finite κ, the point of interest is a local optimal solution of the partially penalized problem (OV$_\kappa$) as well under a not too restrictive additional assumption. Since (OV$_\kappa$) is a program which is likely to satisfy standard constraint qualifications, KKT-type necessary optimality conditions come within reach provided φ possesses certain (generalized) differentiability properties. The procedure described above was used to derive necessary optimality conditions for common finite-dimensional bilevel programming problems in [29, 34, 63, 92, 137, 138], for semidefinite bilevel programming problems in [30], and for bilevel optimal control problems in [13, 14, 74, 84, 130, 131]. Below, we show how this theory generalizes to a very abstract setting which covers all the aforementioned types of bilevel programming problems.

Definition 4.2. Let $(\bar{x},\bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a local optimal solution of (OV). This program is called partially calm at $(\bar{x},\bar{y})$ if there are a neighborhood U of $(\bar{x},\bar{y},0)$ and a constant $\eta > 0$, such that the following implication is valid:

$$\forall (x,y,r) \in U\colon \quad G(x) \in C,\ g(x,y) \in K,\ f(x,y)-\varphi(x) \leq r \implies F(x,y) - F(\bar{x},\bar{y}) + \eta r \geq 0.$$

As mentioned earlier, we have the following result which parallels [138, Proposition 3.3] and [14, Lemma 3.3]. However, since there is no proof provided in [138] and the result in [14] addresses a very special optimal control problem, we decided to present a proof for the sake of completeness.

Proposition 4.33. Let $(\bar{x},\bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a local optimal solution of (OV) where F is continuously Fréchet differentiable. Then (OV) is partially calm at $(\bar{x},\bar{y})$ if and only if there is some $\bar{\kappa} > 0$ such that $(\bar{x},\bar{y})$ is a local optimal solution of (OV$_\kappa$) for any $\kappa \geq \bar{\kappa}$.

Proof. We set $M := \{(x,y) \in \mathcal{X} \times \mathcal{Y} \,|\, G(x) \in C,\ g(x,y) \in K\}$. Furthermore, some $\varepsilon > 0$ can be chosen such that F is Lipschitz continuous on $\mathbb{U}^\varepsilon_{\mathcal{X}\times\mathcal{Y}}(\bar{x},\bar{y})$ with Lipschitz modulus $L > 0$ since F is continuously Fréchet differentiable at $(\bar{x},\bar{y})$. Supposing that (OV) is partially calm at $(\bar{x},\bar{y})$, we find $\delta > 0$ and $\eta > 0$ such that

$$\forall (x,y,r) \in (M \times \mathbb{R}) \cap \mathbb{U}^\delta_{\mathcal{X}\times\mathcal{Y}\times\mathbb{R}}(\bar{x},\bar{y},0)\colon \quad f(x,y)-\varphi(x) \leq r \implies F(x,y) - F(\bar{x},\bar{y}) + \eta r \geq 0$$

is valid. Define $\epsilon := \min\{\varepsilon; \frac{\delta}{2}\}$, choose $(x,y) \in M \cap \mathbb{U}^{\epsilon}_{\mathcal{X}\times\mathcal{Y}}(\bar{x},\bar{y})$ arbitrarily, and set $r := f(x,y) - \varphi(x)$. Note that $f(\bar{x},\bar{y}) - \varphi(\bar{x}) = 0$ and $r \geq 0$ hold true.
If (x,y,r) is an element of $\mathbb{U}^{\delta}_{\mathcal{X}\times\mathcal{Y}\times\mathbb{R}}(\bar{x},\bar{y},0)$, then the partial calmness yields

$$F(\bar{x},\bar{y}) + \eta(f(\bar{x},\bar{y}) - \varphi(\bar{x})) = F(\bar{x},\bar{y}) \leq F(x,y) + \eta r = F(x,y) + \eta(f(x,y) - \varphi(x)).$$

On the other hand, if (x,y,r) does not belong to $\mathbb{U}^{\delta}_{\mathcal{X}\times\mathcal{Y}\times\mathbb{R}}(\bar{x},\bar{y},0)$, then we have $r \geq \frac{\delta}{2}$ from our choice $(x,y) \in \mathbb{U}^{\epsilon}_{\mathcal{X}\times\mathcal{Y}}(\bar{x},\bar{y})$. The Lipschitz continuity of F around $(\bar{x},\bar{y})$ leads to

$$F(\bar{x},\bar{y}) + L(f(\bar{x},\bar{y}) - \varphi(\bar{x})) = F(\bar{x},\bar{y}) \leq F(x,y) + L\tfrac{\delta}{2} \leq F(x,y) + Lr = F(x,y) + L(f(x,y) - \varphi(x)).$$

Setting $\bar{\kappa} := \max\{\eta; L\}$ shows that $(\bar{x},\bar{y})$ is a local solution of (OV$_\kappa$) for any $\kappa \geq \bar{\kappa}$.
For the converse direction of the proof, we assume that there is $\bar{\kappa} > 0$ such that $(\bar{x},\bar{y})$ solves (OV$_\kappa$) locally for any $\kappa \geq \bar{\kappa}$. That is why we find a constant $\gamma > 0$ such that

$$\forall (x,y) \in M \cap \mathbb{U}^{\gamma}_{\mathcal{X}\times\mathcal{Y}}\colon \quad F(x,y) + \bar{\kappa}(f(x,y) - \varphi(x)) \geq F(\bar{x},\bar{y})$$

holds true. Consequently, if $(x,y,r) \in (M \times \mathbb{R}) \cap \mathbb{U}^{\gamma}_{\mathcal{X}\times\mathcal{Y}\times\mathbb{R}}(\bar{x},\bar{y},0)$ satisfies $f(x,y) - \varphi(x) \leq r$, then

$$F(x,y) - F(\bar{x},\bar{y}) + \bar{\kappa} r \geq F(x,y) - F(\bar{x},\bar{y}) + \bar{\kappa}(f(x,y) - \varphi(x)) \geq 0$$

is valid, i.e. (OV) is partially calm at $(\bar{x},\bar{y})$. □

If the bilevel programming problem (BPP) possesses a minimax structure, i.e. if the upper level objective function F equals $-f$, then the partial calmness property holds at all local optimal solutions of (OV), see [84, Remark 3], [131], or [138, Section 4.1]. One may check [14, 34, 63, 138] for other criteria which ensure that the partial calmness condition holds at a given local minimum of (OV). Later, we will make use of the presence of a so-called uniformly weak sharp minimum of the lower level problem (4.1), see [138, Section 5].

Definition 4.3. The lower level problem (4.1) possesses a uniformly weak sharp minimum if there exists a constant $\gamma > 0$ which satisfies

$$\forall (x,y) \in \mathcal{X} \times \mathcal{Y}\colon \quad g(x,y) \in K \implies f(x,y) - \varphi(x) \geq \gamma \cdot \min_{y' \in \Psi(x)} \|y - y'\|_{\mathcal{Y}}\,. \tag{4.53}$$

Particularly, the minimum on the right needs to be attained if the solution set $\Psi(x)$ is nonempty.

One can check [136, 138] for criteria which ensure that the lower level problem possesses a uniformly weak sharp minimum. In the upcoming example, we characterize a certain class of parametric programs where this property is inherent.

Example 4.34. Let $\mathcal{Z}'$ be an arbitrary Banach space and set $\mathcal{Z} = \mathcal{Z}' \times \mathbb{R}^p$. For operators $\mathtt{A} \in \mathbb{L}[\mathcal{X},\mathcal{Z}']$ and $\mathtt{B} \in \mathbb{L}[\mathcal{Y},\mathcal{Z}']$, functionals $a_1^*,\ldots,a_p^* \in \mathcal{X}^\star$ and $b_1^*,\ldots,b_p^*,c^* \in \mathcal{Y}^\star$, scalars $\beta_1,\ldots,\beta_p \in \mathbb{R}$, as well as $\xi \in \mathcal{Z}'$, we consider

$$\forall x \in \mathcal{X}\, \forall y \in \mathcal{Y}\colon \quad f(x,y) := \langle c^*, y\rangle_{\mathcal{Y}}\,, \quad g(x,y) := \begin{bmatrix} \mathtt{A} \\ a_1^* \\ \vdots \\ a_p^* \end{bmatrix} [x] + \begin{bmatrix} \mathtt{B} \\ b_1^* \\ \vdots \\ b_p^* \end{bmatrix} [y] - \begin{pmatrix} \xi \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}.$$

We set $K := \{0\} \times \left(-\mathbb{R}^{p,+}_0\right)$. Finally, we assume that $\mathcal{Y}$ is reflexive and that $\mathtt{B}$ possesses a closed range. Fix some $(\tilde{x},\tilde{y}) \in \mathcal{X} \times \mathcal{Y}$ which satisfy $g(\tilde{x},\tilde{y}) \in K$ and assume that $\Psi(\tilde{x})$ is nonempty. Since the latter set is convex and closed (due to the linearity of f and g), there exists $\bar{y} \in \operatorname{Argmin}_{y' \in \Psi(\tilde{x})} \|y' - \tilde{y}\|_{\mathcal{Y}}$, see Lemma 2.5. Obviously, we have

$$\Psi(\tilde{x}) = \{y \in \mathcal{Y} \,|\, \langle c^*, y - \bar{y}\rangle_{\mathcal{Y}} = 0,\, g(\tilde{x},y) \in K\}.$$

Invoking Hoffmann's lemma, see [17, Theorem 2.200], and $g(\tilde{x},\tilde{y}) \in K$, there is a scalar $\varrho > 0$ depending only on $\mathtt{B}$ and $b_1^*,\dots,b_p^*,c^*$ such that

$$\begin{aligned}\min_{y'\in\Psi(\tilde{x})} \|y'-\tilde{y}\|_{\mathcal{Y}} &= \|\bar{y}-\tilde{y}\|_{\mathcal{Y}}\\ &\leq \varrho\left(\|\mathtt{A}[\tilde{x}]+\mathtt{B}[\tilde{y}]-\xi\|_{\mathcal{Z}'}+\sum_{i=1}^{p}\max\{\langle a_i^*,\tilde{x}\rangle_{\mathcal{X}}+\langle b_i^*,\tilde{y}\rangle_{\mathcal{Y}}-\beta_i;0\}+\big|\langle c^*,\tilde{y}-\bar{y}\rangle_{\mathcal{Y}}\big|\right)\\ &= \varrho\big|\langle c^*,\tilde{y}-\bar{y}\rangle_{\mathcal{Y}}\big| = \varrho\big(\langle c^*,\tilde{y}\rangle_{\mathcal{Y}}-\langle c^*,\bar{y}\rangle_{\mathcal{Y}}\big) = \varrho\big(f(\tilde{x},\tilde{y})-\varphi(\tilde{x})\big)\end{aligned}$$

is valid. Since ϱ is independent of the choice of $(\tilde{x},\tilde{y})$, the considered parametric optimization problem possesses a uniformly weak sharp minimum. ■

In the following proposition, which is related to [138, Proposition 5.1], we show that the presence of a uniformly weak sharp minimum for (4.1) implies that (OV) is partially calm at all local optimal solutions where the objective function is continuously Fréchet differentiable.

Proposition 4.35. Let $(\bar{x},\bar{y}) \in \mathcal{X}\times\mathcal{Y}$ be a local optimal solution of (BPP) where F is continuously Fréchet differentiable. Furthermore, assume that (4.1) possesses a uniformly weak sharp minimum. Then (OV) is partially calm at $(\bar{x},\bar{y})$.

Proof. Since $(\bar{x},\bar{y})$ is a local optimal solution of (BPP) where F is continuously Fréchet differentiable, we find a constant $\varepsilon > 0$ such that F is Lipschitz continuous on $\mathbb{U}^{\varepsilon}_{\mathcal{X}\times\mathcal{Y}}(\bar{x},\bar{y})$ with Lipschitz modulus $L > 0$ and satisfies $F(x,y) \geq F(\bar{x},\bar{y})$ for all $(x,y) \in \mathbb{U}^{\varepsilon}_{\mathcal{X}\times\mathcal{Y}}(\bar{x},\bar{y})$ which are feasible for (BPP). Furthermore, there is a constant $\gamma > 0$ which satisfies (4.53).
We set $\delta := \min\{\frac{\varepsilon}{4};\frac{\varepsilon\gamma}{4}\}$ and choose $(x,y,r) \in \mathbb{U}^{\delta}_{\mathcal{X}\times\mathcal{Y}\times\mathbb{R}}(\bar{x},\bar{y},0)$ which satisfies $G(x) \in C$, $g(x,y) \in K$, and $f(x,y)-\varphi(x) \leq r$. Due to the definition of the uniformly weak sharp minimum, we find a point $y(x) \in \operatorname{Argmin}_{y'\in\Psi(x)} \|y-y'\|_{\mathcal{Y}}$. This yields

$$\|y(x)-\bar{y}\|_{\mathcal{Y}} \leq \|y(x)-y\|_{\mathcal{Y}}+\|y-\bar{y}\|_{\mathcal{Y}} < \tfrac{1}{\gamma}\big(f(x,y)-\varphi(x)\big)+\tfrac{\varepsilon}{4} \leq \tfrac{r}{\gamma}+\tfrac{\varepsilon}{4} < \tfrac{\varepsilon}{2}.$$

Clearly, $(x,y(x)) \in \mathbb{U}^{\varepsilon}_{\mathcal{X}\times\mathcal{Y}}(\bar{x},\bar{y})$ is feasible for (BPP) and, thus, satisfies $F(x,y(x)) \geq F(\bar{x},\bar{y})$. Finally, this leads to

$$F(x,y)-F(\bar{x},\bar{y}) \geq F(x,y)-F(x,y(x)) \geq -L\,\|y-y(x)\|_{\mathcal{Y}} \geq -\tfrac{L}{\gamma}\big(f(x,y)-\varphi(x)\big) \geq -\tfrac{L}{\gamma}r,$$

i.e. (OV) is partially calm at $(\bar{x},\bar{y})$. □

As we mentioned earlier, we need to ensure that the function φ possesses certain generalized differentiability properties in order to derive KKT-type necessary optimality conditions for the problem (OV_κ). Here we just focus on the local Lipschitz continuity of φ which ensures the nonemptyness of its limiting and, thus, Clarke subdifferential. The following result is taken from [31, Lemma 3.1, Theorem 3.1].

Lemma 4.36. Let $(\bar{x},\bar{y}) \in \operatorname{gph}\Psi$ be a point where Ψ is inner semicontinuous. Assume that the Banach spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are reflexive. Moreover, let K be SNC at $g(\bar{x},\bar{y})$ and let the constraint qualification

$$g'_y(\bar{x},\bar{y})[\mathcal{Y}] - \mathcal{R}_K(g(\bar{x},\bar{y})) = \mathcal{Z} \tag{4.54}$$

be valid. Then φ is locally Lipschitz continuous at $\bar{x}$ and the following formula holds true:

$$\partial^c\varphi(\bar{x}) \subseteq \{f'_x(\bar{x},\bar{y})+g'_x(\bar{x},\bar{y})^\star[\lambda] \,|\, \lambda\in\Lambda(\bar{x},\bar{y})\}.$$

Therein, $\Lambda(\bar{x},\bar{y})$ denotes the set of lower level Lagrange multipliers at $(\bar{x},\bar{y})$ given by

$$\Lambda(\bar{x},\bar{y}) := \big\{\lambda\in\mathcal{N}_K(g(\bar{x},\bar{y})) \,\big|\, 0 = f'_y(\bar{x},\bar{y})+g'_y(\bar{x},\bar{y})^\star[\lambda]\big\}.$$

Note that the inner semicontinuity of the solution set mapping Ψ needed in the above lemma is a very restrictive assumption. However, we exploited it here to get rid of compactness assumptions (as they are postulated in e.g. [26]) which are rarely satisfied in the infinite-dimensional setting. In [29, Remark 3.2], the authors present an overview of conditions ensuring the inner semicontinuity of Ψ at a given point. Recalling Proposition 4.1 and Lemma 4.2, the solution set mapping of the parametric optimization problem (4.3) is singleton-valued as well as Lipschitz continuous and, thus, inner semicontinuous. Clearly, the corresponding optimal value function is locally Lipschitz continuous.
It is worth to mention that the results of Lemma 4.36 stay valid if the inner semicontinuity of Ψ at the point $(\bar{x}, \bar{y}) \in \operatorname{gph} \Psi$ is weakend to φ-inner-semicontinuity and φ is lower semicontinuous at $\bar{x}$, see [91, Section 5] for the definition of φ-inner-semicontinuity and [91, Theorem 5.2] as well as [93, Theorem 7] for the validation of this result. However, the usual way to guarantee the lower semicontinuity of an optimal value function is to postulate certain compactness assumptions on the underlying data, see [7, Theorems 4.2.1. and 4.2.2.], which, as we mentioned earlier, is often too restrictive when infinite-dimensional parametric optimization problems are under consideration.
Finally, we would like to emphasize that results similar to Lemma 4.36 can be given in the case where Ψ is only inner semicompact (φ-inner-semicompact) at the reference point $\bar{x}$, see [91, 93] for the details.
We combine Proposition 4.33 and Lemma 4.36 in order to obtain the following necessary optimality conditions of KKT-type. Although the technique of their validation is the same as used in [29, 30, 34] to derive similar results, we decided to present the proof here in order to show the reader how all the above preliminaries come together.

Theorem 4.37. Let $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a local optimal solution of (BPP) where F in continuously Fréchet differentiable, Ψ is inner semicontinuous, and the constraint qualifications (4.21) as well as (4.54) are valid. Furthermore, let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be reflexive, whereas K is SNC at $g(\bar{x}, \bar{y})$. Finally, assume that (OV) is partially calm at $(\bar{x}, \bar{y})$. Then there exist multipliers $\rho \in \mathcal{W}^\star$ and $\lambda, \bar{\lambda} \in \mathcal{Z}^\star$ as well as a scalar $\kappa > 0$ which satisfy the following conditions:

$$0 = F'_x(\bar{x}, \bar{y}) + G'(\bar{x})^\star[\rho] + g'_x(\bar{x}, \bar{y})^\star[\lambda - \kappa\bar{\lambda}], \tag{4.55a}$$
$$0 = F'_y(\bar{x}, \bar{y}) + \kappa f'_y(\bar{x}, \bar{y}) + g'_y(\bar{x}, \bar{y})^\star[\lambda], \tag{4.55b}$$
$$0 = f'_y(\bar{x}, \bar{y}) + g'_y(\bar{x}, \bar{y})^\star[\bar{\lambda}], \tag{4.55c}$$
$$\rho \in \mathcal{N}_C(G(\bar{x})), \tag{4.55d}$$
$$\lambda \in \mathcal{N}_K(g(\bar{x}, \bar{y})), \tag{4.55e}$$
$$\bar{\lambda} \in \mathcal{N}_K(g(\bar{x}, \bar{y})). \tag{4.55f}$$

Proof. Due to Proposition 4.33, we find some $\kappa > 0$ such that $(\bar{x}, \bar{y})$ is a local optimal solution of (OV_κ). The theorem's assumptions guarantee that the objective function of the latter program is locally Lipschitz continuous at $(\bar{x}, \bar{y})$, see Lemma 4.36. We set $M := \{(x, y) \in \mathcal{X} \times \mathcal{Y} \,|\, G(x) \in C,\, g(x, y) \in K\}$ and obtain

$$0 \in \partial^c\big(F + \kappa(f - \varphi)\big)(\bar{x}, \bar{y}) + \mathcal{N}_M(\bar{x}, \bar{y})$$

from Lemma 2.29. Invoking the sum rule for Clarke's subdifferential, see [24, Corollary 2 in Section 2.3], we have

$$\partial^c\big(F + \kappa(f - \varphi)\big)(\bar{x}, \bar{y}) = \big\{F'(\bar{x}, \bar{y}) + \kappa f'(\bar{x}, \bar{y})\big\} - \kappa\partial^c\varphi(\bar{x}) \times \{0\}.$$

Clearly, the validity of the constraint qualifications (4.21) as well as (4.54) implies that the constraint qualification (4.49) holds as well. Hence, we can apply Lemma 2.31 in order to obtain

$$\mathcal{N}_M(\bar{x}, \bar{y}) = \big\{\big(G'(\bar{x})^\star[\rho] + g'_x(\bar{x}, \bar{y})^\star[\lambda], g'_y(\bar{x}, \bar{y})^\star[\lambda]\big) \in \mathcal{X}^\star \times \mathcal{Y}^\star \,\big|\, \rho \in \mathcal{N}_C(G(\bar{x})),\, \lambda \in \mathcal{N}_K(g(\bar{x}, \bar{y}))\big\}.$$

Consequently, we find $x^* \in \partial^c\varphi(\bar{x})$, $\rho \in \mathcal{W}^\star$, and $\lambda \in \mathcal{Z}^\star$ which satisfy

$$0 = F'_x(\bar{x}, \bar{y}) + \kappa(f'_x(\bar{x}, \bar{y}) - x^*) + G'(\bar{x})^\star[\rho] + g'_x(\bar{x}, \bar{y})^\star[\lambda], \tag{4.56}$$

(4.55b), (4.55d), and (4.55e). Finally, Lemma 4.36 yields the existence of $\bar{\lambda} \in \mathcal{Z}^\star$ satisfying (4.55c), (4.55f), and

$$x^* = f'_x(\bar{x}, \bar{y}) + g'_x(\bar{x}, \bar{y})^\star[\bar{\lambda}].$$

Putting this representation of x^* into (4.56) leads to (4.55a) and, consequently, the proof is completed. □

5. Selected applications of bilevel programming

In this chapter, we are going to apply the theory developed in the above sections in order to state necessary optimality conditions and constraint qualifications for three different types of bilevel programming problems. First, we discuss a special hierarchical semidefinite optimization problem whose lower level is governed by a certain operator equation. Therefore, we exploit the results obtained in Section 4.1. Afterwards, bilevel optimal control problems of ODEs with lower level control constraints are considered in more detail. Here our findings from Chapter 3 and Section 4.2 are useful. Finally, we study an optimal control problem with an implicit pointwise state constraint determined by a finite-dimensional optimization problem. Necessary optimality conditions and constraint qualifications for this problem are derived via the optimal value reformulation of the hierarchical optimization problem.

5.1. A special class of hierarchical semidefinite programming problems

Here we want to illustrate the theory from Section 4.1 for the situation where $\mathcal{U}$ equals the Hilbert space $\mathcal{S}_p$ and U_{ad} is given by the positive semidefinite cone in $\mathcal{S}_p^+$. Note that in this case, the surrogate MPCCs (4.22) and (4.32) will be semidefinite complementarity problems. Thus, it will be possible to compare our results to the achievements in [37, 124, 127]. Bilevel programming problems with finite-dimensional decision variables and a semidefinite lower level problem were recently considered in [30].
Let us motivate the setting of this section by means of the following example.

Example 5.1. Let $C \subseteq \mathcal{S}_p$ be a closed, convex set of real symmetric matrices, let $\mathbf{Y}_{\mathrm{d}} \in \mathcal{S}_p$ be a given matrix, and let $\lambda_0 \in \mathbb{R}$ be a fixed real number. We consider the problem of finding a matrix $\mathbf{Y}$ *not too far away from* C, whose eigenvalues are at least as large as λ_0, and whose distance to $\mathbf{Y}_{\mathrm{d}}$ is minimal. In order to emphasize that $\mathbf{Y}$ does not necessarily need to be an element of C while the eigenvalue condition has to be satisfied in any case, a possible formulation of this problem can be stated as follows:

$$\begin{aligned} \tfrac{1}{2}\left\|\mathbf{Y}-\mathbf{Y}_{\mathrm{d}}\right\|_{\mathcal{S}_p}^2 &\to \min_{\mathbf{X},\mathbf{Y},\mathbf{U}} \\ \mathbf{X} &\in C \\ (\mathbf{Y},\mathbf{U}) &\in \operatorname*{Argmin}_{\mathbf{Y},\mathbf{U}} \left\{ \tfrac{1}{2}\left\|\mathbf{Y}-\mathbf{X}\right\|_{\mathcal{S}_p}^2 + \tfrac{\sigma}{2}\left\|\mathbf{U}\right\|_{\mathcal{S}_p}^2 \;\middle|\; \begin{array}{l} \mathbf{Y}-\mathbf{U}-\lambda_0\mathbf{I}_p = \mathbf{O} \\ \mathbf{U} \in \mathcal{S}_p^+ \end{array} \right\}. \end{aligned}$$

Therein, the fixed parameter $\sigma > 0$ controls the preference between the goals *stay close to the set* C (σ small) and *stay close to the matrix* $\lambda_0\mathbf{I}_p$ (σ large). ■

We start the paragraph with a short introduction to variational analysis in the Hilbert space $\mathcal{S}_p$. Afterwards, we apply our findings to state necessary optimality conditions for the bilevel program of interest and compare the results to the ones in literature. Especially, we will show that there are essential differences between the results in Theorem 4.10 and Proposition 4.13.

5.1.1. Variational analysis in $\mathcal{S}_p$

Let $\mathcal{O}_p$ denote the set of all real orthogonal matrices from $\mathbb{R}^{p\times p}$. For an arbitrary matrix $\mathbf{M} \in \mathcal{S}_p$, there exist a matrix $\mathbf{P} \in \mathcal{O}_p$ and a diagonal matrix $\mathbf{\Lambda} \in \mathbb{R}^{p\times p}$ whose diagonal entries are ordered nonincreasingly

such that $\mathbf{M} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^\top$ holds true. This representation is called an ordered eigenvalue decomposition of $\mathbf{M}$. For fixed index sets $I, J \subseteq \{1, \ldots, p\}$ and an orthogonal matrix $\mathbf{Q} \in \mathcal{O}_p$, we define $\mathbf{M}^{\mathbf{Q}} := \mathbf{Q}^\top \mathbf{M}\mathbf{Q}$ and $\mathbf{M}^{\mathbf{Q}}_{IJ} := (\mathbf{M}^{\mathbf{Q}})_{IJ}$.
Recall that $\mathcal{S}_p^+$ denotes the cone of all positive semidefinite matrices in $\mathcal{S}_p$. Furthermore, we use $\mathcal{S}_p^{++}$ to denote the (nonclosed) convex cone of all positive definite matrices in $\mathcal{S}_p$. Clearly, we have the relation $\mathcal{S}_p = \mathcal{S}_p^{++} - \mathcal{S}_p^{++}$. In our subsequent analysis, we need to characterize the positive definiteness of block matrices. Therefore, we will exploit the following well-known result taken from [141, Theorem 1.12].

Lemma 5.2. Let $n, m \in \mathbb{N}$ satisfy $n + m = p$ and let $\mathbf{R} \in \mathcal{S}_n$, $\mathbf{S} \in \mathbb{R}^{n\times m}$, as well as $\mathbf{T} \in \mathcal{S}_m$ be fixed. Then the block matrix

$$\mathbf{M} := \begin{bmatrix} \mathbf{R} & \mathbf{S} \\ \mathbf{S}^\top & \mathbf{T} \end{bmatrix} \in \mathcal{S}_p$$

is positive definite if and only if $\mathbf{T} \in \mathcal{S}_m$ and $\mathbf{R} - \mathbf{S}\mathbf{T}^{-1}\mathbf{S}^\top \in \mathcal{S}_n$ are both positive definite.

Fix some matrix $\mathbf{A} \in \mathcal{S}_p$ and let $\bar{\mathbf{A}}$ be the projection of $\mathbf{A}$ onto $\mathcal{S}_p^+$. Furthermore, let $\mathbf{P}\mathbf{\Lambda}\mathbf{P}^\top$ be an ordered eigenvalue decomposition of $\mathbf{A}$. Then, by means of [69, Section 4.2.3], we obtain

$$\bar{\mathbf{A}} = \mathbf{P}\max\{\mathbf{\Lambda}; \mathbf{O}\}\mathbf{P}^\top \in \mathcal{S}_p^+, \qquad \mathbf{A} - \bar{\mathbf{A}} = \mathbf{P}\min\{\mathbf{\Lambda}; \mathbf{O}\}\mathbf{P}^\top \in \mathcal{S}_p^-$$

where minimum and maximum are interpreted in entrywise fashion and $\mathcal{S}_p^-$ denotes the closed, convex cone of all negative semidefinite and symmetric matrices from $\mathbb{R}^{p\times p}$. Clearly, $\mathcal{S}_p^+$ and $\mathcal{S}_p^-$ are polar to each other. Let α, β, and γ denote the index sets corresponding to the positive, zero, and negative eigenvalues of $\mathbf{\Lambda}$.

Lemma 5.3. Using the above notation, we have

$$\begin{aligned} \mathrm{cl}\big(\mathcal{S}_p^+ - \mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp\big) \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp &= \big\{\mathbf{W} \in \mathcal{S}_p \,\big|\, \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}\big\}, \\ \mathrm{cl}\big(\mathcal{S}_p^- - \mathcal{S}_p^- \cap \{\bar{\mathbf{A}}\}^\perp\big) \cap \{\bar{\mathbf{A}}\}^\perp &= \big\{\mathbf{V} \in \mathcal{S}_p \,\big|\, \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O}\big\}. \end{aligned}$$

Proof. We only show the first assertion since the proof of the second one is analogous.
For an arbitrary matrix $\mathbf{W} \in \mathcal{S}_p$, we obtain

$$\begin{aligned} \mathbf{W} \in \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp &\iff \mathrm{tr}\,\big((\mathbf{A} - \bar{\mathbf{A}})\mathbf{W}\big) = 0 \\ &\iff \mathrm{tr}\,\big((\mathbf{A} - \bar{\mathbf{A}})^{\mathbf{P}}\mathbf{W}^{\mathbf{P}}\big) = 0 \\ &\iff \mathrm{tr}\,\big(\min\{\mathbf{\Lambda}; \mathbf{O}\}\mathbf{W}^{\mathbf{P}}\big) = 0 \iff \mathrm{tr}\,\big(\mathbf{\Lambda}_{\gamma\gamma}\mathbf{W}^{\mathbf{P}}_{\gamma\gamma}\big) = 0. \end{aligned} \tag{5.1}$$

Thus, if $\mathbf{W}$ comes from $\mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp$, then we have $\mathbf{W}^{\mathbf{P}}_{\gamma\gamma} \in \mathcal{S}^+_{|\gamma|}$ and, due to the above arguments, $\mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}$. We conclude

$$\mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp = \Big\{\mathbf{W} \in \mathcal{S}_p \,\big|\, \mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\alpha\cup\beta} \in \mathcal{S}^+_{|\alpha\cup\beta|},\, \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\, \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\, \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}\Big\}. \tag{5.2}$$

Choose $\mathbf{W} \in \mathrm{cl}\big(\mathcal{S}_p^+ - \mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp\big) \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp$ arbitrarily. Then there are sequences $\{\mathbf{S}_k\} \subseteq \mathcal{S}_p^+$ and $\{\mathbf{T}_k\} \subseteq \mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp$ such that $\mathbf{S}_k - \mathbf{T}_k \to \mathbf{W}$ holds true. From (5.2) we deduce the relation $(\mathbf{S}_k - \mathbf{T}_k)^{\mathbf{P}}_{\gamma\gamma} = (\mathbf{S}_k)^{\mathbf{P}}_{\gamma\gamma} \in \mathcal{S}^+_{|\gamma|}$ for all $k \in \mathbb{N}$. Taking the limit $k \to \infty$ yields $\mathbf{W}^{\mathbf{P}}_{\gamma\gamma} \in \mathcal{S}^+_{|\gamma|}$. Since we have $\mathbf{W} \in \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp$, $\mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}$ can be derived from (5.1). This shows the inclusion $\subseteq$.
For the proof of the other inclusion, we pick a matrix $\mathbf{W} \in \mathcal{S}_p$ such that $\mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}$ holds. Due to (5.1), we only need to verify $\mathbf{W} \in \mathrm{cl}\big(\mathcal{S}_p^+ - \mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp\big)$. From $\mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\alpha\cup\beta} \in \mathcal{S}_{|\alpha\cup\beta|}$ we find two matrices $\mathbf{X}, \mathbf{Y} \in \mathcal{S}^{++}_{|\alpha\cup\beta|}$ which satisfy $\mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\alpha\cup\beta} = \mathbf{X} - \mathbf{Y}$. For $k \in \mathbb{N}$, we define matrices $\mathbf{S}_k, \mathbf{T}_k \in \mathcal{S}_p$ as stated below:

$$\mathbf{S}_k := \mathbf{P}\begin{bmatrix} k\mathbf{X} + k^2\mathbf{I}_{|\alpha\cup\beta|} & \mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\gamma} \\ \mathbf{W}^{\mathbf{P}}_{\gamma,\alpha\cup\beta} & \frac{1}{k}\mathbf{I}_{|\gamma|} \end{bmatrix}\mathbf{P}^\top, \qquad \mathbf{T}_k := \mathbf{P}\begin{bmatrix} \mathbf{Y} + (k-1)\mathbf{X} + k^2\mathbf{I}_{|\alpha\cup\beta|} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}\mathbf{P}^\top.$$

Obviously, $\mathbf{T}_k^{\mathbf{P}}$ is the sum of the three positive semidefinite matrices

$$\begin{bmatrix} \mathbf{Y} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \qquad (k-1)\begin{bmatrix} \mathbf{X} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \qquad k^2\begin{bmatrix} \mathbf{I}_{|\alpha\cup\beta|} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

and, thus, positive semidefinite. Taking a look at (5.2) yields $\mathbf{T}_k \in \mathcal{S}_p^+ \cap \{\mathbf{A} - \bar{\mathbf{A}}\}^\perp$. Next, we show $\mathbf{S}_k \in \mathcal{S}_p^{++}$ for sufficiently large k. Therefore, we use Lemma 5.2. First, observe that the matrix $\frac{1}{k}\mathbf{I}_{|\gamma|}$ is positive definite. Clearly, the eigenvalues of

$$\left(\mathbf{X} + k\mathbf{I}_{|\alpha\cup\gamma|}\right) - \mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\gamma}\mathbf{W}^{\mathbf{P}}_{\gamma,\alpha\cup\beta}.$$

strictly increase for $k \to \infty$. Thus, this matrix is positive definite for sufficiently large $k \in \mathbb{N}$. Consequently, the matrix

$$\left(k\mathbf{X} + k^2\mathbf{I}_{|\alpha\cup\beta|}\right) - \mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\gamma}\left(\tfrac{1}{k}\mathbf{I}_{|\gamma|}\right)^{-1}\mathbf{W}^{\mathbf{P}}_{\gamma,\alpha\cup\beta}$$

is positive definite for sufficiently large $k \in \mathbb{N}$. By means of Lemma 5.2 $\mathbf{S}_k^{\mathbf{P}} \in \mathcal{S}_p^{++}$ is valid for large enough k, and since $\mathbf{P}$ is orthogonal, the same holds true for $\mathbf{S}_k$. Combining these observations,

$$\mathbf{P}\begin{bmatrix}\mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\alpha\cup\beta} & \mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\gamma}\\ \mathbf{W}^{\mathbf{P}}_{\gamma,\alpha\cup\beta} & \frac{1}{k}\mathbf{I}_{|\gamma|}\end{bmatrix}\mathbf{P}^\top = \mathbf{P}\begin{bmatrix}\mathbf{X}-\mathbf{Y} & \mathbf{W}^{\mathbf{P}}_{\alpha\cup\beta,\gamma}\\ \mathbf{W}^{\mathbf{P}}_{\gamma,\alpha\cup\beta} & \frac{1}{k}\mathbf{I}_{|\gamma|}\end{bmatrix}\mathbf{P}^\top = \mathbf{S}_k - \mathbf{T}_k \in \mathcal{S}_p^+ - \mathcal{S}_p^+ \cap \{\mathbf{A}-\bar{\mathbf{A}}\}^\perp$$

is obtained for sufficiently large $k \in \mathbb{N}$, and taking the limit $k \to \infty$ yields $\mathbf{W} \in \operatorname{cl}\left(\mathcal{S}_p^+ - \mathcal{S}_p^+ \cap \{\mathbf{A}-\bar{\mathbf{A}}\}^\perp\right)$. This completes the proof. □

In [17, Section 5.3.1], the authors provide the following formula for the radial cone to $\mathcal{S}_p^+$ at $\bar{\mathbf{A}}$:

$$\mathcal{R}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}) = \left\{\mathbf{W} \in \mathcal{S}_p \,\middle|\, \mathbf{W}^{\mathbf{P}}_{\beta\cup\gamma,\beta\cup\gamma} \in \mathcal{S}^+_{|\beta\cup\gamma|},\, \exists \mathbf{X} \in \mathbb{R}^{|\alpha|\times|\beta\cup\gamma|}\colon\, \mathbf{W}^{\mathbf{P}}_{\alpha,\beta\cup\gamma} = \mathbf{X}\mathbf{W}^{\mathbf{P}}_{\beta\cup\gamma,\beta\cup\gamma}\right\}.$$

Moreover, in [69, Section 5], one can find explicit formulae for the tangent and normal cone to the cone of positive semidefinite matrices:

$$\begin{aligned}\mathcal{T}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}) &= \left\{\mathbf{W} \in \mathcal{S}_p \,\middle|\, \mathbf{W}^{\mathbf{P}}_{\beta\cup\gamma,\beta\cup\gamma} \in \mathcal{S}^+_{|\beta\cup\gamma|}\right\},\\ \mathcal{N}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}) &= \left\{\mathbf{V} \in \mathcal{S}_p \,\middle|\, \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\beta\cup\gamma,\beta\cup\gamma} \in \mathcal{S}^-_{|\beta\cup\gamma|}\right\}.\end{aligned}$$

Since we have $\bar{\mathbf{A}} = \operatorname{proj}_{\mathcal{S}_p^+}(\mathbf{A})$ and $\left\langle \bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}}\right\rangle_{\mathcal{S}_p} = 0$, it is reasonable to consider the critical cone $\mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})$. An explicit formula for this cone is stated in [99] and presented below:

$$\mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}}) = \left\{\mathbf{W} \in \mathcal{S}_p \,\middle|\, \mathbf{W}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^+_{|\beta|},\, \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\, \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}\right\}.$$

Lemma 5.4. Using the above notations, we have

$$\mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})^\circ = \left\{\mathbf{V} \in \mathcal{S}_p \,\middle|\, \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^-_{|\beta|}\right\}.$$

Especially, $\mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})^\circ \neq \mathcal{K}_{\mathcal{S}_p^-}(\mathbf{A}-\bar{\mathbf{A}}, \bar{\mathbf{A}})$ is valid provided $\alpha \neq \varnothing$ and $\gamma \neq \varnothing$ hold, i.e. $\mathcal{S}_p^+$ is not polyhedric w.r.t. $(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})$ in general.

Proof. Let us introduce

$$D := \left\{\mathbf{V} \in \mathcal{S}_p \,\middle|\, \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^-_{|\beta|}\right\}.$$

Choosing $\mathbf{V} \in D$, we easily see $\mathbf{V} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})^\circ$, i.e. $D \subseteq \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})^\circ$.
For the proof of the converse inclusion, we choose $\mathbf{V} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})^\circ$ arbitrarily. For any matrix $\mathbf{W} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})$, we obtain

$$\begin{aligned}0 \geq \langle \mathbf{V}, \mathbf{W}\rangle_{\mathcal{S}_p} &= \operatorname{tr}(\mathbf{V}\mathbf{W}) = \operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}\mathbf{W}^{\mathbf{P}}\right)\\ &= \operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}_{\alpha\alpha}\mathbf{W}^{\mathbf{P}}_{\alpha\alpha}\right) + 2\operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}_{\alpha\beta}\mathbf{W}^{\mathbf{P}}_{\beta\alpha}\right) + 2\operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}_{\alpha\gamma}\mathbf{W}^{\mathbf{P}}_{\gamma\alpha}\right) + \operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}_{\beta\beta}\mathbf{W}^{\mathbf{P}}_{\beta\beta}\right).\end{aligned}$$

Since there is no information on the blocks $\mathbf{W}^{\mathbf{P}}_{\alpha\alpha}$, $\mathbf{W}^{\mathbf{P}}_{\beta\alpha}$, and $\mathbf{W}^{\mathbf{P}}_{\gamma\alpha}$, we deduce $\mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O}$, $\mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O}$, and $\mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}$. Moreover, the block $\mathbf{V}^{\mathbf{P}}_{\beta\beta}$ needs to satisfy $0 \geq \operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}_{\beta\beta}\mathbf{W}^{\mathbf{P}}_{\beta\beta}\right) = \left\langle \mathbf{V}^{\mathbf{P}}_{\beta\beta}, \mathbf{W}^{\mathbf{P}}_{\beta\beta}\right\rangle_{\mathcal{S}_{|\beta|}}$ for any

$\mathbf{W}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^+_{|\beta|}$. Thus, we obtain $\mathbf{V}^{\mathbf{P}}_{\beta\beta} \in (\mathcal{S}^+_{|\beta|})^\circ = \mathcal{S}^-_{|\beta|}$ and, hence, $\mathbf{V} \in D$. This shows $D = \mathcal{K}_{\mathcal{S}^+_p}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})^\circ$. The last statement of the lemma follows from

$$\mathcal{K}_{\mathcal{S}^-_p}(\mathbf{A}-\bar{\mathbf{A}}, \bar{\mathbf{A}}) = \left\{ \mathbf{V} \in \mathcal{S}_p \;\middle|\; \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^-_{|\beta|} \right\}$$

and the characterization of polyhedric cones in (2.5). □

Let us introduce a matrix $\boldsymbol{\Xi} \in \mathcal{S}_p$ as stated below:

$$\forall i,j \in \{1,\ldots,p\}\colon \quad \xi_{i,j} := \begin{cases} \frac{\max\{\lambda_{i,i};0\}-\max\{\lambda_{j,j};0\}}{\lambda_{i,i}-\lambda_{j,j}} & \text{if } (i,j) \in (\alpha\times\gamma)\cup(\gamma\times\alpha), \\ 1 & \text{otherwise.} \end{cases} \tag{5.3}$$

We obtain from [99, Proposition 9] that the metric projection onto $\mathcal{S}^+_p$ is directionally differentiable at $\mathbf{A}$ and satisfies

$$\forall \boldsymbol{\Delta} \in \mathcal{S}_p\colon \quad \mathrm{proj}'_{\mathcal{S}^+_p}(\mathbf{A};\boldsymbol{\Delta}) = \mathbf{P} \begin{bmatrix} \boldsymbol{\Delta}^{\mathbf{P}}_{\alpha\alpha} & \boldsymbol{\Delta}^{\mathbf{P}}_{\alpha\beta} & \boldsymbol{\Xi}_{\alpha\gamma} \bullet \boldsymbol{\Delta}^{\mathbf{P}}_{\alpha\gamma} \\ \boldsymbol{\Delta}^{\mathbf{P}}_{\beta\alpha} & \mathrm{proj}_{\mathcal{S}^+_{|\beta|}}\left(\boldsymbol{\Delta}^{\mathbf{P}}_{\beta\beta}\right) & \mathbf{O} \\ \boldsymbol{\Xi}_{\gamma\alpha} \bullet \boldsymbol{\Delta}^{\mathbf{P}}_{\gamma\alpha} & \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{P}^\top.$$

Since the projection is Lipschitz continuous and $\mathcal{S}_p$ possesses the finite dimension $\frac{1}{2}p(p+1)$, $\mathrm{proj}_{\mathcal{S}^+_p}$ is already B-differentiable. On the other hand, [99, Proposition 9] yields

$$\forall \boldsymbol{\Delta} \in \mathcal{S}_p\colon \quad \mathrm{proj}'_{\mathcal{S}^+_p}(\mathbf{A};\boldsymbol{\Delta}) = \left(\mathrm{proj}_{\mathcal{K}_{\mathcal{S}^+_p}(\bar{\mathbf{A}},\mathbf{A}-\bar{\mathbf{A}})} \circ \mathrm{L}^2\right)(\boldsymbol{\Delta}) \tag{5.4}$$

where $\mathrm{L} \in \mathbb{L}[\mathcal{S}_p, \mathcal{S}_p]$ is the linear operator defined below:

$$\forall \boldsymbol{\Delta} \in \mathcal{S}_p\colon \quad \mathrm{L}[\boldsymbol{\Delta}] := \mathbf{P}\left(\sqrt{\boldsymbol{\Xi}} \bullet \boldsymbol{\Delta}^{\mathbf{P}}\right)\mathbf{P}^\top. \tag{5.5}$$

Therein, the entries of the matrix $\sqrt{\boldsymbol{\Xi}}$ are given by the square roots of the entries of $\boldsymbol{\Xi}$. In the lemma below, we study the properties of the operator L in more detail.

Lemma 5.5. Using the above notations, the operator L from (5.5) is a self-adjoint automorphism which satisfies

$$\mathrm{L} \circ \mathrm{proj}_{\mathcal{K}_{\mathcal{S}^+_p}(\bar{\mathbf{A}},\mathbf{A}-\bar{\mathbf{A}})} = \mathrm{proj}_{\mathcal{K}_{\mathcal{S}^+_p}(\bar{\mathbf{A}},\mathbf{A}-\bar{\mathbf{A}})} \circ \mathrm{L}, \tag{5.6a}$$

$$\forall \boldsymbol{\Delta} \in \mathcal{K}_{\mathcal{S}^+_p}(\bar{\mathbf{A}}, \mathbf{A}-\bar{\mathbf{A}})\colon \quad \mathrm{proj}'_{\mathcal{S}^+_p}(\mathbf{A};\boldsymbol{\Delta}) = \mathrm{L}^2[\boldsymbol{\Delta}]. \tag{5.6b}$$

Proof. The fact that L is an isomorphism is easily seen. Its inverse is given by

$$\forall \boldsymbol{\Delta} \in \mathcal{S}_p\colon \quad \mathrm{L}^{-1}[\boldsymbol{\Delta}] = \mathbf{P}\left(\frac{1}{\sqrt{\boldsymbol{\Xi}}} \bullet \boldsymbol{\Delta}^{\mathbf{P}}\right)\mathbf{P}^\top$$

where the matrix $\frac{1}{\sqrt{\boldsymbol{\Xi}}}$ contains entrywise the reciprocal entries of $\sqrt{\boldsymbol{\Xi}}$.
Next, we show that L is self-adjoint. Therefore, choose $\boldsymbol{\Delta}, \boldsymbol{\Theta} \in \mathcal{S}_p$ arbitrarily and observe

$$\begin{aligned} \langle \boldsymbol{\Theta}, \mathrm{L}[\boldsymbol{\Delta}] \rangle_{\mathcal{S}_p} &= \mathrm{tr}\left(\boldsymbol{\Theta}\,\mathrm{L}[\boldsymbol{\Delta}]\right) = \mathrm{tr}\left(\mathbf{P}\boldsymbol{\Theta}^{\mathbf{P}}\mathbf{P}^\top\mathbf{P}\left(\sqrt{\boldsymbol{\Xi}} \bullet \boldsymbol{\Delta}^{\mathbf{P}}\right)\mathbf{P}^\top\right) \\ &= \mathrm{tr}\left(\boldsymbol{\Theta}^{\mathbf{P}}\left(\sqrt{\boldsymbol{\Xi}} \bullet \boldsymbol{\Delta}^{\mathbf{P}}\right)\right) = \mathrm{tr}\left(\left(\sqrt{\boldsymbol{\Xi}} \bullet \boldsymbol{\Theta}^{\mathbf{P}}\right)\boldsymbol{\Delta}^{\mathbf{P}}\right) \\ &= \mathrm{tr}\left(\mathbf{P}\left(\sqrt{\boldsymbol{\Xi}} \bullet \boldsymbol{\Theta}^{\mathbf{P}}\right)\mathbf{P}^\top\mathbf{P}\boldsymbol{\Delta}^{\mathbf{P}}\mathbf{P}^\top\right) = \mathrm{tr}\left(\mathrm{L}[\boldsymbol{\Theta}]\,\boldsymbol{\Delta}\right) = \langle \mathrm{L}[\boldsymbol{\Theta}], \boldsymbol{\Delta} \rangle_{\mathcal{S}_p}. \end{aligned}$$

This shows that L is self-adjoint.
Now, we validate property (5.6a). Since L is a self-adjoint automorphism, we can invoke Lemma 4.4 for

this issue. For brevity, we set $\mathcal{K} := \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A} - \bar{\mathbf{A}})$. First, we need to show $\mathtt{L}[\mathcal{K}] = \mathcal{K}$. Therefore, choose $\mathbf{V} \in \mathtt{L}[\mathcal{K}]$ arbitrarily. Then there exists a matrix $\mathbf{W} \in \mathcal{S}_p$ satisfying $\mathbf{W}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^+_{|\beta|}$ and

$$\mathbf{V} = \mathbf{P} \begin{bmatrix} \mathbf{W}^{\mathbf{P}}_{\alpha\alpha} & \mathbf{W}^{\mathbf{P}}_{\alpha\beta} & \sqrt{\boldsymbol{\Xi}}_{\alpha\gamma} \bullet \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} \\ \mathbf{W}^{\mathbf{P}}_{\beta\alpha} & \mathbf{W}^{\mathbf{P}}_{\beta\beta} & \mathbf{O} \\ \sqrt{\boldsymbol{\Xi}}_{\gamma\alpha} \bullet \mathbf{W}^{\mathbf{P}}_{\gamma\alpha} & \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{P}^\top.$$

Consequently, we have $\mathbf{V}^{\mathbf{P}}_{\beta\beta} = \mathbf{W}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^+_{|\beta|}$, $\mathbf{V}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O}$, and $\mathbf{V}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}$, i.e. $\mathbf{V} \in \mathcal{K}$. This shows the inclusion $\mathtt{L}[\mathcal{K}] \subseteq \mathcal{K}$. Similarly, we can show $\mathtt{L}^{-1}[\mathcal{K}] \subseteq \mathcal{K}$, and applying $\mathtt{L}$ to this relation yields the other inclusion $\mathcal{K} \subseteq \mathtt{L}[\mathcal{K}]$. Next, we need to verify

$$\forall \mathbf{W} \in \mathcal{K}\, \forall \mathbf{V} \in \mathcal{K}^\circ\colon \quad \langle \mathbf{V}, \mathbf{W} \rangle_{\mathcal{S}_p} = 0 \iff \langle \mathtt{L}[\mathbf{V}], \mathtt{L}[\mathbf{W}] \rangle_{\mathcal{S}_p} = 0.$$

However, choosing $\mathbf{W} \in \mathcal{K}$ as well as $\mathbf{V} \in \mathcal{K}^\circ$ arbitrarily and respecting Lemma 5.4, we easily see

$$\begin{aligned} \langle \mathtt{L}[\mathbf{V}], \mathtt{L}[\mathbf{W}] \rangle_{\mathcal{S}_p} = 0 &\iff \operatorname{tr}\left(\mathbf{P} \left(\sqrt{\boldsymbol{\Xi}} \bullet \mathbf{V}^{\mathbf{P}} \right) \mathbf{P}^\top \mathbf{P} \left(\sqrt{\boldsymbol{\Xi}} \bullet \mathbf{W}^{\mathbf{P}} \right) \mathbf{P}^\top \right) = 0 \\ &\iff \operatorname{tr}\left(\left(\sqrt{\boldsymbol{\Xi}} \bullet \mathbf{V}^{\mathbf{P}} \right) \left(\sqrt{\boldsymbol{\Xi}} \bullet \mathbf{W}^{\mathbf{P}} \right) \right) \\ &\iff \operatorname{tr}\left(\mathbf{V}^{\mathbf{P}}_{\beta\beta} \mathbf{W}^{\mathbf{P}}_{\beta\beta} \right) = 0 \\ &\iff \operatorname{tr}\left(\mathbf{V}^{\mathbf{P}} \mathbf{W}^{\mathbf{P}} \right) = 0 \\ &\iff \operatorname{tr}(\mathbf{V}\mathbf{W}) = 0 \\ &\iff \langle \mathbf{V}, \mathbf{W} \rangle_{\mathcal{S}_p} = 0. \end{aligned}$$

By means of Lemma 4.4, (5.6a) is valid.
Property (5.6b) simply follows combining (5.4) and $\mathtt{L}[\mathcal{K}] = \mathcal{K}$ which was shown earlier. This completes the proof. □

Finally, we want to take a closer look at the complementarity set $\mathcal{C} := \operatorname{gph} \mathcal{N}_{\mathcal{K}}$ which is induced by the closed, convex cone $\mathcal{K} := \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{A}}, \mathbf{A} - \bar{\mathbf{A}})$. From Lemma 5.4 we obtain

$$\mathcal{C} = \left\{ (\mathbf{W}, \mathbf{V}) \in \mathcal{S}_p \times \mathcal{S}_p \;\middle|\; \begin{array}{l} \mathbf{W}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^+_{|\beta|},\, \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\, \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \\ \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\, \mathbf{V}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^-_{|\beta|}, \\ \left\langle \mathbf{V}^{\mathbf{P}}_{\beta\beta}, \mathbf{W}^{\mathbf{P}}_{\beta\beta} \right\rangle_{\mathcal{S}_{|\beta|}} = 0 \end{array} \right\}.$$

In the following lemma, we compute the limiting normal cone to $\mathcal{C}$ at $(\mathbf{O}, \mathbf{O})$.

Lemma 5.6. Using the above notations, we have

$$\mathcal{N}_{\mathcal{C}}(\mathbf{O}, \mathbf{O}) = \left\{ (\mathbf{M}, \mathbf{N}) \in \mathcal{S}_p \times \mathcal{S}_p \;\middle|\; \begin{array}{l} \mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\, \mathbf{M}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\, \mathbf{M}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}, \\ \mathbf{N}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\, \mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \\ (\mathbf{M}^{\mathbf{P}}_{\beta\beta}, \mathbf{N}^{\mathbf{P}}_{\beta\beta}) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{S}^+_{|\beta|}}}(\mathbf{O}, \mathbf{O}) \end{array} \right\}$$

where

$$\mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{S}^+_{|\beta|}}}(\mathbf{O}, \mathbf{O}) = \bigcup_{\substack{\{\beta_+, \beta_0, \beta_-\} \in \mathcal{P}(\beta) \\ \boldsymbol{\Sigma} \in [0,1]^{|\beta_+| \times |\beta_-|} \\ \mathbf{Q} \in \mathcal{O}_{|\beta|}}} \left\{ (\mathbf{V}, \mathbf{W}) \in \mathcal{S}_{|\beta|} \times \mathcal{S}_{|\beta|} \;\middle|\; \begin{array}{l} \mathbf{V}^{\mathbf{Q}}_{\beta_+\beta_+} = \mathbf{O},\, \mathbf{V}^{\mathbf{Q}}_{\beta_+\beta_0} = \mathbf{O}, \\ \mathbf{Q}^\top_{\beta\beta_0} \mathbf{V} \mathbf{Q}_{\beta\beta_0} \in \mathcal{S}^-_{|\beta_0|}, \\ \mathbf{W}^{\mathbf{Q}}_{\beta_0\beta_-} = \mathbf{O},\, \mathbf{W}^{\mathbf{Q}}_{\beta_-\beta_-} = \mathbf{O}, \\ \mathbf{Q}^\top_{\beta\beta_0} \mathbf{W} \mathbf{Q}_{\beta\beta_0} \in \mathcal{S}^+_{|\beta_0|}, \\ \boldsymbol{\Sigma} \bullet \mathbf{V}^{\mathbf{Q}}_{\beta_+\beta_-} + (\mathbf{E} - \boldsymbol{\Sigma}) \bullet \mathbf{W}^{\mathbf{Q}}_{\beta_+\beta_-} = \mathbf{O} \end{array} \right\}$$

holds. Here $\mathcal{P}(\beta)$ denotes the set of all partitions of β.

Proof. It is sufficient to show the equation involving $\mathcal{N}_{\mathcal{C}}(\mathbf{O},\mathbf{O})$ since the formula for the limiting normal cone to $\operatorname{gph}\mathcal{N}_{\mathcal{S}^+_{|\beta|}}$ at $(\mathbf{O},\mathbf{O})$ can be found in [37, Proposition 3.3].
We introduce an isomorphism $\mathtt{F}\in\mathbb{L}[\mathcal{S}_p,\mathcal{S}_p]$ by $\mathtt{F}[\mathbf{\Delta}]:=\mathbf{\Delta}^{\mathbf{P}}$ for all $\mathbf{\Delta}\in\mathcal{S}_p$. Moreover, let us define a set $S\subseteq\mathcal{S}_p\times\mathcal{S}_p$ as stated below:

$$S:=\left\{(\mathbf{X},\mathbf{Y})\in\mathcal{S}_p\times\mathcal{S}_p\;\middle|\;\begin{array}{l}\mathbf{X}_{\beta\gamma}=\mathbf{O},\,\mathbf{X}_{\gamma\gamma}=\mathbf{O},\\ \mathbf{Y}_{\alpha\alpha}=\mathbf{O},\,\mathbf{Y}_{\alpha\beta}=\mathbf{O},\,\mathbf{Y}_{\alpha\gamma}=\mathbf{O},\\ (\mathbf{X}_{\beta\beta},\mathbf{Y}_{\beta\beta})\in\operatorname{gph}\mathcal{N}_{\mathcal{S}^+_{|\beta|}}\end{array}\right\}.$$

Then we easily see $\mathcal{C}=\{(\mathbf{W},\mathbf{V})\in\mathcal{S}_p\times\mathcal{S}_p\,|\,(\mathtt{F}[\mathbf{W}],\mathtt{F}[\mathbf{V}])\in S\}$. Exploiting the surjectivity of $\mathtt{F}$,

$$\mathcal{N}_{\mathcal{C}}(\mathbf{O},\mathbf{O})=\left\{\left(\mathtt{F}^\star\big[\widetilde{\mathbf{M}}\big],\mathtt{F}^\star\big[\widetilde{\mathbf{N}}\big]\right)\in\mathcal{S}_p\times\mathcal{S}_p\;\middle|\;\left(\widetilde{\mathbf{M}},\widetilde{\mathbf{N}}\right)\in\mathcal{N}_S(\mathbf{O},\mathbf{O})\right\}$$

is obtained from Lemma 2.38. We apply the product rule for limiting normals (2.6) in order to see

$$\mathcal{N}_S(\mathbf{O},\mathbf{O})=\left\{\left(\widetilde{\mathbf{M}},\widetilde{\mathbf{N}}\right)\in\mathcal{S}_p\times\mathcal{S}_p\;\middle|\;\begin{array}{l}\widetilde{\mathbf{M}}_{\alpha\alpha}=\mathbf{O},\,\widetilde{\mathbf{M}}_{\alpha\beta}=\mathbf{O},\,\widetilde{\mathbf{M}}_{\alpha\gamma}=\mathbf{O},\\ \widetilde{\mathbf{N}}_{\beta\gamma}=\mathbf{O},\,\widetilde{\mathbf{N}}_{\gamma\gamma}=\mathbf{O},\\ \left(\widetilde{\mathbf{M}}_{\beta\beta},\widetilde{\mathbf{N}}_{\beta\beta}\right)\in\mathcal{N}_{\operatorname{gph}\mathcal{N}_{\mathcal{S}^+_{|\beta|}}}(\mathbf{O},\mathbf{O})\end{array}\right\}.$$

Thus, the formula for the limiting normal cone to $\mathcal{C}$ follows from

$$\forall\mathbf{\Theta}\in\mathcal{S}_p\colon\quad\mathtt{F}^\star[\mathbf{\Theta}]=\mathbf{P}\mathbf{\Theta}\mathbf{P}^\top$$

which is easily obtained from the definition of the adjoint operator. This completes the proof. □

5.1.2. Necessary optimality conditions and constraint qualifications

We consider the bilevel programming problem

$$\begin{aligned}F(x,y,\mathbf{U})\;&\to\;\min_{x,y,\mathbf{U}}\\ G(x)\;&\in\;C\\ (y,\mathbf{U})\;&\in\;\Psi(x)\end{aligned}\tag{5.7}$$

where $\Psi\colon\mathcal{X}\rightrightarrows\mathcal{Y}_s\times\mathcal{S}_p$ denotes the solution mapping of the parametric optimization problem

$$\begin{aligned}\tfrac{1}{2}\left\|\mathtt{C}[y]-\mathtt{P}[x]\right\|^2_{\mathcal{M}}+\tfrac{\sigma}{2}\left\|\mathbf{U}-\mathtt{Q}[x]\right\|^2_{\mathcal{S}_p}\;&\to\;\min_{y,\mathbf{U}}\\ \mathtt{A}[y]-\mathtt{B}[\mathbf{U}]-h(x)\;&=\;0\\ \mathbf{U}\;&\in\;\mathcal{S}^+_p.\end{aligned}\tag{5.8}$$

Here Assumption 4.2 shall hold with $\mathcal{U}=\mathcal{S}_p$ and $U_{\mathsf{ad}}=\mathcal{S}^+_p$. From Proposition 4.1 we already know that for any $x\in\mathcal{X}$, the lower level problem possesses a unique solution which can be characterized by the projection operator onto the positive semidefinite cone. As we mentioned in Section 5.1.1, the mapping $\operatorname{proj}_{\mathcal{S}^+_p}$ is B-differentiable everywhere and the corresponding directional derivative can be characterized in the sense of Haraux's lemma, see (5.4) and Lemma 5.5. Thus, all assumptions of Lemma 4.8 hold. We define the Lipschitz continuous and Fréchet differentiable function $\eta\colon\mathcal{X}\to\mathcal{S}_p$ and the bounded, linear operator $\mathtt{E}\in\mathbb{L}[\mathcal{S}_p,\mathcal{S}_p]$ as we did at the beginning of Section 4.1. Invoking Proposition 4.9, we obtain the following intermediate result.

Proposition 5.7. Let $(\bar{x},\bar{y},\bar{\mathbf{U}})\in\mathcal{X}\times\mathcal{Y}_s\times\mathcal{S}_p$ be a local optimal solution of (5.7) with lower level (5.8) where the function F is continuously Fréchet differentiable. Let the constraint qualification (4.21) be satisfied. Set $\bar{\mathbf{W}}:=\eta(\bar{x})-\mathtt{E}[\bar{\mathbf{U}}]$, let $\mathbf{P}\mathbf{\Lambda}\mathbf{P}^\top$ be an ordered eigenvalue decomposition of $\bar{\mathbf{W}}$, and let $\mathtt{L}\in\mathbb{L}[\mathcal{S}_p,\mathcal{S}_p]$

be the operator defined in (5.5) where $\boldsymbol{\Xi} \in \mathcal{S}_p$ is given in (5.3).
Then $(\bar{\delta}_x, \bar{\boldsymbol{\Delta}}_{\mathbf{U}}, \bar{\boldsymbol{\Delta}}_{\boldsymbol{\Pi}}) := (0, \mathbf{O}, \mathbf{O})$ is a global optimal solution of the following MPCC:

$$\begin{aligned}
&\big(F_x'(\bar{x},\bar{y},\bar{\mathbf{U}}) + F_y'(\bar{x},\bar{y},\bar{\mathbf{U}}) \circ \mathtt{A}^{-1} \circ h'(\bar{x})\big)[\delta_x] \\
&+\big(F_y'(\bar{x},\bar{y},\bar{\mathbf{U}}) \circ \mathtt{A}^{-1} \circ \mathtt{B} + F_{\mathbf{U}}'(\bar{x},\bar{y},\bar{\mathbf{U}})\big)[\boldsymbol{\Delta}_{\mathbf{U}}] \;\to\; \min_{\delta_x, \boldsymbol{\Delta}_{\mathbf{U}}, \boldsymbol{\Delta}_{\boldsymbol{\Pi}}} \\
G'(\bar{x})[\delta_x] &\in \mathcal{T}_C(G(\bar{x})) \\
\boldsymbol{\Delta}_{\mathbf{U}} - \mathtt{L}^2[\boldsymbol{\Delta}_{\boldsymbol{\Pi}}] &= 0 \\
\boldsymbol{\Delta}_{\boldsymbol{\Pi}} &\in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \\
\eta'(\bar{x})[\delta_x] - \mathtt{E}[\boldsymbol{\Delta}_{\mathbf{U}}] - \boldsymbol{\Delta}_{\boldsymbol{\Pi}} &\in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^\circ \\
\langle \eta'(\bar{x})[\delta_x] - \mathtt{E}[\boldsymbol{\Delta}_{\mathbf{U}}] - \boldsymbol{\Delta}_{\boldsymbol{\Pi}}, \boldsymbol{\Delta}_{\boldsymbol{\Pi}} \rangle_{\mathcal{S}_p} &= 0.
\end{aligned} \tag{5.9}$$

Before we apply our optimality conditions from Theorem 4.10 to (5.7), we state the following supplementary result which helps us to characterize the constraint qualifications (4.23) and (4.27) in the semidefinite case.

Lemma 5.8. Let $(\bar{x}, \bar{y}, \bar{\mathbf{U}}) \in \mathcal{X} \times \mathcal{Y}_s \times \mathcal{S}_p$ be a feasible point of (5.7). Define $\bar{\mathbf{W}} := \eta(\bar{x}) - \mathtt{E}[\bar{\mathbf{U}}]$, let $\mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^\top$ be an ordered eigenvalue decomposition of $\bar{\mathbf{W}}$, let α, β, and γ denote the index sets corresponding to the positive, zero, and negative eigenvalues of $\boldsymbol{\Lambda}$, and let $\mathtt{L} \in \mathbb{L}[\mathcal{S}_p, \mathcal{S}_p]$ be the operator defined in (5.5) where $\boldsymbol{\Xi} \in \mathcal{S}_p$ is given in (5.3).
We define a set $W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \subseteq \mathcal{S}_p \times \mathcal{S}_p$ as stated below:

$$W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) := \left\{ (\mathbf{W}, \mathbf{V}) \in \mathcal{S}_p \times \mathcal{S}_p \;\middle|\; \begin{array}{l} \mathbf{W}^{\mathbf{P}}_{\beta\beta} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \\ \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\beta\beta} = \mathbf{O}, \\ \boldsymbol{\Xi}_{\alpha\gamma} \bullet \big(\mathbf{W}^{\mathbf{P}}_{\alpha\gamma} + \mathbf{V}^{\mathbf{P}}_{\alpha\gamma}\big) - \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O} \end{array} \right\}.$$

Then we have

$$W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) = \begin{bmatrix} \mathtt{L}^2 & \mathtt{O} \\ \mathtt{I}_{\mathcal{S}_p} - \mathtt{L}^2 & \mathtt{I}_{\mathcal{S}_p} \end{bmatrix} \begin{pmatrix} \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^{\circ\perp} \\ \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^{\perp} \end{pmatrix}.$$

Proof. We show both inclusions separately. Choosing $(\mathbf{W}, \mathbf{V}) \in W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})$, we define

$$\widetilde{\mathbf{W}}^{\mathbf{P}} := \frac{1}{\boldsymbol{\Xi}} \bullet \mathbf{W}^{\mathbf{P}} = \begin{bmatrix} \mathbf{W}^{\mathbf{P}}_{\alpha\alpha} & \mathbf{W}^{\mathbf{P}}_{\alpha\beta} & \big(\frac{1}{\boldsymbol{\Xi}}\big)_{\alpha\gamma} \bullet \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} \\ \mathbf{W}^{\mathbf{P}}_{\beta\alpha} & \mathbf{O} & \mathbf{O} \\ \big(\frac{1}{\boldsymbol{\Xi}}\big)_{\gamma\alpha} \bullet \mathbf{W}^{\mathbf{P}}_{\gamma\alpha} & \mathbf{O} & \mathbf{O} \end{bmatrix}, \qquad \widetilde{\mathbf{V}}^{\mathbf{P}} := \begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{V}^{\mathbf{P}}_{\beta\gamma} \\ \mathbf{O} & \mathbf{V}^{\mathbf{P}}_{\gamma\beta} & \mathbf{V}^{\mathbf{P}}_{\gamma\gamma} \end{bmatrix}.$$

Therein, $\frac{1}{\boldsymbol{\Xi}} \in \mathcal{S}_p$ denotes the matrix which contains entrywise the reciprocal entries of $\boldsymbol{\Xi}$. From Lemmas 2.12 and 5.4 we obtain $\widetilde{\mathbf{W}} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^{\circ\perp}$ and $\widetilde{\mathbf{V}} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^{\perp}$. The definition of the set $W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})$ yields $\mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \big(\big(\frac{1}{\boldsymbol{\Xi}}\big)_{\alpha\gamma} - \mathbf{E}\big) \bullet \mathbf{W}^{\mathbf{P}}_{\alpha\gamma}$. Thus, we derive

$$(\mathbf{E} - \boldsymbol{\Xi}) \bullet \widetilde{\mathbf{W}}^{\mathbf{P}} + \widetilde{\mathbf{V}}^{\mathbf{P}} = \left(\frac{1}{\boldsymbol{\Xi}} - \mathbf{E}\right) \bullet \mathbf{W}^{\mathbf{P}} + \widetilde{\mathbf{V}}^{\mathbf{P}} = \mathbf{V}^{\mathbf{P}}$$

and this leads to

$$\mathbf{W} = \mathtt{L}^2[\widetilde{\mathbf{W}}], \quad \mathbf{V} = (\mathtt{I}_{\mathcal{S}_p} - \mathtt{L}^2)[\widetilde{\mathbf{W}}] + \widetilde{\mathbf{V}}.$$

This shows the inclusion $\subseteq$.
For the proof of the converse inclusion, choose $\widetilde{\mathbf{W}} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^{\circ\perp}$ and $\widetilde{\mathbf{V}} \in \mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^{\perp}$ arbitrarily. This yields $\widetilde{\mathbf{W}}^{\mathbf{P}}_{\beta\beta} = \mathbf{O}$, $\widetilde{\mathbf{W}}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O}$, $\widetilde{\mathbf{W}}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}$, $\widetilde{\mathbf{V}}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O}$, $\widetilde{\mathbf{V}}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O}$, $\widetilde{\mathbf{V}}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}$, and $\widetilde{\mathbf{V}}^{\mathbf{P}}_{\beta\beta} = \mathbf{O}$. Thus, we obtain

$$\mathtt{L}^2[\widetilde{\mathbf{W}}] = \mathbf{P} \begin{bmatrix} \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\alpha} & \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\beta} & \boldsymbol{\Xi}_{\alpha\gamma} \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\gamma} \\ \widetilde{\mathbf{W}}^{\mathbf{P}}_{\beta\alpha} & \mathbf{O} & \mathbf{O} \\ \boldsymbol{\Xi}_{\gamma\alpha} \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\gamma\alpha} & \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{P}^\top$$

and

$$(\mathtt{I}_{\mathcal{S}_p} - \mathtt{L}^2)[\widetilde{\mathbf{W}}] + \widetilde{\mathbf{V}} = \mathbf{P} \begin{bmatrix} \mathbf{O} & \mathbf{O} & (\mathbf{E} - \Xi_{\alpha\gamma}) \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\gamma} \\ \mathbf{O} & \mathbf{O} & \widetilde{\mathbf{V}}^{\mathbf{P}}_{\beta\gamma} \\ (\mathbf{E} - \Xi_{\gamma\alpha}) \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\gamma\alpha} & \widetilde{\mathbf{V}}^{\mathbf{P}}_{\gamma\beta} & \widetilde{\mathbf{V}}^{\mathbf{P}}_{\gamma\gamma} \end{bmatrix} \mathbf{P}^\top.$$

We easily see

$$\Xi_{\alpha\gamma} \bullet \Big[\big[\Xi_{\alpha\gamma} \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\gamma}\big] + \big[(\mathbf{E} - \Xi_{\alpha\gamma}) \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\gamma}\big] \Big] - \big[\Xi_{\alpha\gamma} \bullet \widetilde{\mathbf{W}}^{\mathbf{P}}_{\alpha\gamma}\big] = \mathbf{O}.$$

This shows $\big(\mathtt{L}^2[\widetilde{\mathbf{W}}], (\mathtt{I}_{\mathcal{S}_p} - \mathtt{L}^2)[\widetilde{\mathbf{W}}] + \widetilde{\mathbf{V}}\big) \in W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})$ and, thus, completes the proof. □

Now, we are in position to restate the necessary optimality conditions from Theorem 4.10 in terms of problem (5.7).

Proposition 5.9. Let $(\bar{x}, \bar{y}, \bar{\mathbf{U}}) \in \mathcal{X} \times \mathcal{Y}_s \times \mathcal{S}_p$ be a local optimal solution of (5.7) where F is continuously Fréchet differentiable. Suppose that the constraint qualification (4.21) holds. Set $\bar{\mathbf{W}} := \eta(\bar{x}) - \mathtt{E}[\bar{\mathbf{U}}]$. Let $\mathbf{P}\mathbf{\Lambda}\mathbf{P}^\top$ be an ordered eigenvalue decomposition of $\bar{\mathbf{W}}$, let α, β, and γ denote the index sets corresponding to the positive, zero, and negative eigenvalues of $\mathbf{\Lambda}$, and let Ξ be the matrix defined in (5.3). Then the following statements hold:

1. Assume that the constraint qualification

$$\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{S}_p} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{S}_p} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{S}_p \end{pmatrix} - \begin{pmatrix} \mathcal{T}_C(G(\bar{x})) \\ W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{S}_p \\ \mathcal{S}_p \end{pmatrix} \tag{5.10}$$

holds. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$, and $p \in \mathcal{Y}_s$ which satisfy (4.24d) as well as

$$0 = F'_x(\bar{x}, \bar{y}, \bar{\mathbf{U}}) + h'(\bar{x})^\star[p] + G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\mathbf{N}], \tag{5.11a}$$
$$\mathbf{O} = F'_{\mathbf{U}}(\bar{x}, \bar{y}, \bar{\mathbf{U}}) + \mathtt{B}^\star[p] + \mathbf{M} - (\mathtt{E} + \mathtt{I}_{\mathcal{S}_p})[\mathbf{N}], \tag{5.11b}$$
$$0 = \mathtt{A}^\star[p] - F'_y(\bar{x}, \bar{y}, \bar{\mathbf{U}}), \tag{5.11c}$$
$$\mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{M}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O}, \tag{5.11d}$$
$$\mathbf{N}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \tag{5.11e}$$
$$\Xi_{\alpha\gamma} \bullet (\mathbf{M}^{\mathbf{P}}_{\alpha\gamma} - \mathbf{N}^{\mathbf{P}}_{\alpha\gamma}) + \mathbf{N}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}. \tag{5.11f}$$

2. Let $\mathcal{X}$ and $\mathcal{W}$ be reflexive. Suppose that the constraint qualification

$$\left.\begin{array}{l} 0 = G'(\bar{x})^\star[\rho] + \eta'(\bar{x})^\star[\mathbf{N}], \\ \mathbf{O} = \mathbf{M} - (\mathtt{E} + \mathtt{I}_{\mathcal{S}_p})[\mathbf{N}], \\ \rho \in \mathcal{N}_C(G(\bar{x})), \\ \mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{M}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O}, \\ \mathbf{N}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \\ \Xi_{\alpha\gamma} \bullet (\mathbf{M}^{\mathbf{P}}_{\alpha\gamma} - \mathbf{N}^{\mathbf{P}}_{\alpha\gamma}) + \mathbf{N}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}, \\ (\mathbf{M}^{\mathbf{P}}_{\beta\beta}, \mathbf{N}^{\mathbf{P}}_{\beta\beta}) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{S}^+_{|\beta|}}}(\mathbf{O}, \mathbf{O}) \end{array}\right\} \implies \rho = 0,\ \mathbf{M} = \mathbf{O},\ \mathbf{N} = \mathbf{O} \tag{5.12}$$

is satisfied while $\mathcal{T}_C(G(\bar{x}))$ is SNC at 0. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$, and $p \in \mathcal{Y}_s$ which satisfy (4.24d), (5.11), and

$$(\mathbf{M}^{\mathbf{P}}_{\beta\beta}, \mathbf{N}^{\mathbf{P}}_{\beta\beta}) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{S}^+_{|\beta|}}}(\mathbf{O}, \mathbf{O}). \tag{5.13}$$

3. Assume that the constraint qualifications (5.10) and

$$\operatorname{cl}\left(\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{S}_p} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{S}_p} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{S}_p \end{pmatrix} - \begin{pmatrix} \mathcal{N}_C(G(\bar{x}))_\perp \\ W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \end{pmatrix} \right) = \begin{pmatrix} \mathcal{W} \\ \mathcal{S}_p \\ \mathcal{S}_p \end{pmatrix} \tag{5.14}$$

hold. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$, and $p \in \mathcal{Y}_s$ which satisfy (4.24d), (5.11), and

$$(\mathbf{M}^{\mathbf{P}}_{\beta\beta}, \mathbf{N}^{\mathbf{P}}_{\beta\beta}) \in \mathcal{S}^-_{|\beta|} \times \mathcal{S}^+_{|\beta|}. \tag{5.15}$$

Proof. Applying Lemmas 5.4, 5.6, and 5.8, the results directly follow from Theorem 4.10. □

Obviously, the constraint qualifications (4.21) and (5.10) are both implied by

$$\begin{bmatrix} G'(\bar{x}) & 0 \\ 0 & \mathtt{I}_{\mathcal{S}_p} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{S}_p} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{S}_p \end{pmatrix} - \begin{pmatrix} \mathcal{R}_C(G(\bar{x})) \\ W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{S}_p \\ \mathcal{S}_p \end{pmatrix} \tag{5.16}$$

which was introduced in [124, Definition 5.5] under the name SDPMPCC-MFCQ. In [124, Definition 5.7], the author refers to the constraint qualification (5.14) as SDPMPCC-LICQ. Observe that SDPMPCC-LICQ does not need to imply SDPMPCC-MFCQ as long as $\mathcal{W}$ is infinite-dimensional. Using [124, Lemma 5.4], we obtain

$$W(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}})^\circ = \left\{ (\mathbf{M}, \mathbf{N}) \in \mathcal{S}_p \times \mathcal{S}_p \;\middle|\; \begin{array}{l} \mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{M}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O}, \\ \mathbf{N}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \\ \mathbf{\Xi}_{\alpha\gamma} \bullet (\mathbf{M}^{\mathbf{P}}_{\alpha\gamma} - \mathbf{N}^{\mathbf{P}}_{\alpha\gamma}) + \mathbf{N}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O} \end{array} \right\}.$$

Thus, if (5.10) holds, then by polarization and Remark 2.33 we easily see that (5.12) is satisfied as well. Especially, SDPMPCC-MFCQ implies (5.12).
If we interpret the first, second, and third set of optimality conditions provided in Proposition 5.9 as some W-, M-, and S-stationarity-type conditions, respectively, then we easily see from Lemma 5.6 that the S-stationarity-type conditions are stronger than the M-stationarity-type conditions, whereas the M-stationarity-type conditions are stronger than the W-stationarity-type conditions. Due to the fact that these optimality conditions were derived via the surrogate MPCC (5.9) whose complementarity cone is trivially polyhedric w.r.t. the point of interest (i.e. the zero vector), the S-stationarity-type conditions possess reasonable strength. From above we obtain that a local minimizer of (5.7) where SDPMPCC-MFCQ is valid, $\mathcal{X}$ and $\mathcal{W}$ are reflexive, and $\mathcal{T}_C(G(\bar{x}))$ is SNC at 0 satisfies the M-stationarity-type conditions. This observation is related to [127, Theorem 3.5] where a similar result was obtained for general semidefinite MPCCs in finite-dimensional spaces.
Since $\mathcal{S}^+_p$ is a closed, convex cone, (5.7) is fully equivalent to the MPCC

$$\begin{aligned} F(x, y, \mathbf{U}) &\to \min_{x,y,\mathbf{U}} \\ G(x) &\in C \\ \mathtt{A}[y] - \mathtt{B}[\mathbf{U}] - h(x) &= 0 \\ \mathbf{U} &\in \mathcal{S}^+_p \\ \eta(x) - (\mathtt{E} + \mathtt{I}_{\mathcal{S}_p})[\mathbf{U}] &\in \mathcal{S}^-_p \\ \big\langle \eta(x) - (\mathtt{E} + \mathtt{I}_{\mathcal{S}_p})[\mathbf{U}], \mathbf{U} \big\rangle_{\mathcal{S}_p} &= 0. \end{aligned} \tag{5.17}$$

It is easy to see that the W-, M-, and S-stationarity-type necessary optimality conditions from Proposition 5.9 coincide with the W-, M-, and S-stationarity conditions as they are introduced in [127, Definition 3.3] for the above MPCC (5.17). This underlines the strength of the derived conditions. In [124, Theorem 5.8], the author shows that SDPMPCC-MFCQ and SDPMPCC-LICQ together imply that a local optimal solution of a general semidefinite complementarity problem satisfies the S-stationarity-type conditions. Here we obtained an analogous result for our special problem class (5.7) in Proposition 5.9.
Let us check how our notions of W-, M-, and S-stationarity from Definitions 3.1 and 3.2 can be used to derive necessary optimality conditions for (5.7) via the equivalent surrogate problem (5.17). Therefore, we just recall Proposition 4.13.

Proposition 5.10. Let $(\bar{x}, \bar{y}, \bar{\mathbf{U}}) \in \mathcal{X} \times \mathcal{Y}_s \times \mathcal{S}_p$ be a local optimal solution of (5.7). Set $\bar{\mathbf{W}} := \eta(\bar{x}) - \mathtt{E}[\bar{\mathbf{U}}]$, let $\mathbf{P}\mathbf{\Lambda}\mathbf{P}^\top$ be an ordered eigenvalue decomposition of $\bar{\mathbf{W}}$, and let α, β, and γ denote the index sets

corresponding to the positive, zero, and negative eigenvalues of $\mathbf{\Lambda}$. We define $T(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \subseteq \mathcal{S}_p \times \mathcal{S}_p$ as stated below:

$$T(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) := \left\{ (\mathbf{W}, \mathbf{V}) \in \mathcal{S}_p \times \mathcal{S}_p \;\middle|\; \begin{array}{l} \mathbf{W}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\beta\beta} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \\ \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\beta\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O} \end{array} \right\}.$$

Then the following statements hold.

1. Assume that the constraint qualification

$$\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{S}_p} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{S}_p} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{S}_p \end{pmatrix} - \begin{pmatrix} \mathcal{R}_C(G(\bar{x})) \\ T(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \end{pmatrix} = \begin{pmatrix} \mathcal{W} \\ \mathcal{S}_p \\ \mathcal{S}_p \end{pmatrix} \tag{5.18}$$

   is valid. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$, and $p \in \mathcal{Y}_s$ which satisfy (4.24d), (5.11a) - (5.11c), and

$$\mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}.$$

2. Let $\mathcal{X}$ and $\mathcal{W}$ be reflexive, and let F be continuously Fréchet differentiable at $(\bar{x}, \bar{y}, \bar{\mathbf{U}})$. Assume that the constraint qualification (5.12) is satisfied while C is SNC at $G(\bar{x})$. Then there exist multipliers $\rho \in \mathcal{W}^\star$, $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$, and $p \in \mathcal{Y}_s$ which satisfy (4.24d), (5.11), and (5.13).

3. Suppose that the constraint qualifications (5.18) and

$$\mathrm{cl}\left(\begin{bmatrix} G'(\bar{x}) & \mathtt{0} \\ \mathtt{0} & \mathtt{I}_{\mathcal{S}_p} \\ \eta'(\bar{x}) & -\mathtt{E} - \mathtt{I}_{\mathcal{S}_p} \end{bmatrix} \begin{pmatrix} \mathcal{X} \\ \mathcal{S}_p \end{pmatrix} - \begin{pmatrix} \mathcal{N}_C(G(\bar{x}))_\perp \\ T(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \end{pmatrix} \right) = \begin{pmatrix} \mathcal{W} \\ \mathcal{S}_p \\ \mathcal{S}_p \end{pmatrix} \tag{5.19}$$

   are satisfied. Then there are multipliers $\rho \in \mathcal{W}^\star$, $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$, and $p \in \mathcal{Y}_s$ which satisfy (4.24d), (5.11a) - (5.11e), and (5.15).

Proof. The proof mainly follows applying Proposition 4.13 to (5.17).
Let us start with the validation of the first statement. Due to Lemma 5.3, it is sufficient to show

$$T(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) = \Big(\mathcal{R}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}) \cap \big(-\mathcal{K}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) \big) \Big) \times \Big(\mathcal{R}_{\mathcal{S}_p^-}(\bar{\mathbf{W}} - \bar{\mathbf{U}}) \cap \big(-\mathcal{K}_{\mathcal{S}_p^-}(\bar{\mathbf{W}} - \bar{\mathbf{U}}, \bar{\mathbf{U}}) \big) \Big).$$

This, however, is a simple consequence of the formulae for radial and critical cone to the semidefinite cone provided in Section 5.1.1.
The proof of the second statement follows easily from [37, Theorem 3.1] where an explicit formula for the limiting normal cone to $\mathrm{gph}\, \mathcal{N}_{\mathcal{S}_p^+}$ is presented.
Recalling the formula for the critical cone to $\mathcal{S}_p^+$, we only need to verify

$$T(\bar{\mathbf{U}}, \bar{\mathbf{W}} - \bar{\mathbf{U}}) = \left(\mathcal{T}_{\mathcal{S}_p^+ \cap (-\mathcal{T}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}))}(\bar{\mathbf{U}})^{\circ\perp} \right) \times \left(\mathcal{T}_{\mathcal{S}_p^- \cap (-\mathcal{T}_{\mathcal{S}_p^-}(\bar{\mathbf{W}} - \bar{\mathbf{U}}))}(\bar{\mathbf{W}} - \bar{\mathbf{U}})^{\circ\perp} \right).$$

This formula easily follows from

$$\mathcal{S}_p^+ \cap \big(-\mathcal{T}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}) \big) = \left\{ \mathbf{W} \in \mathcal{S}^p \;\middle|\; \begin{array}{l} \mathbf{W}^{\mathbf{P}}_{\alpha\alpha} \in \mathcal{S}^+_{|\alpha|},\ \mathbf{W}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}, \\ \mathbf{W}^{\mathbf{P}}_{\beta\beta} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O} \end{array} \right\},$$

$$\mathcal{S}_p^- \cap \big(-\mathcal{T}_{\mathcal{S}_p^-}(\bar{\mathbf{W}} - \bar{\mathbf{U}}) \big) = \left\{ \mathbf{V} \in \mathcal{S}^p \;\middle|\; \begin{array}{l} \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}, \\ \mathbf{V}^{\mathbf{P}}_{\beta\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\gamma\gamma} \in \mathcal{S}^-_{|\gamma|} \end{array} \right\},$$

and, thus,

$$\mathcal{T}_{\mathcal{S}_p^+ \cap (-\mathcal{T}_{\mathcal{S}_p^+}(\bar{\mathbf{U}}))}(\bar{\mathbf{U}}) = \left\{ \mathbf{W} \in \mathcal{S}^p \;\middle|\; \begin{array}{l} \mathbf{W}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\beta\beta} = \mathbf{O}, \\ \mathbf{W}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{W}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O} \end{array} \right\},$$

$$\mathcal{T}_{\mathcal{S}_p^- \cap (-\mathcal{T}_{\mathcal{S}_p^-}(\bar{\mathbf{W}} - \bar{\mathbf{U}}))}(\bar{\mathbf{W}} - \bar{\mathbf{U}}) = \left\{ \mathbf{V} \in \mathcal{S}^p \;\middle|\; \begin{array}{l} \mathbf{V}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}, \\ \mathbf{V}^{\mathbf{P}}_{\beta\beta} = \mathbf{O},\ \mathbf{V}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O} \end{array} \right\}.$$

Now, we can apply Lemma 2.12 in order to complete the proof. □

Here we see that the linearization approach we used in Section 4.1.2 leads to much better results in terms of the W- and S-stationarity conditions than the direct consideration of the equivalent MPCC (5.17): It is obvious that the constraint qualification (5.18) is much stronger than SDPMPCC-MFCQ while (5.19) is much stronger than SDPMPCC-LICQ. On the other hand, the W- and S-stationarity conditions provided in Proposition 5.10 are much weaker than the W- and S-stationarity-type conditions we derived in Proposition 5.9. Interestingly, the corresponding M-stationarity-type results are nearly the same; they only differ in the SNC assumption and the additional constraint qualification (4.21) needed in Proposition 5.9.
Following Lemma 5.3, Theorem [37, Theorem 3.1], and the formula for the critical cone to $\mathcal{S}_p$ stated earlier, we obtain how the W-, M-, and S-stationarity conditions according to Definitions 3.1 and 3.2 for generalized MPCCs whose cone inducing the complementarity constraint equals $\mathcal{S}_p^+$ look like.

Remark 5.11. Consider the mathematical program (MPCC) under Assumption 3.1 with $\mathcal{Z} := \mathcal{S}_p$ and $K := \mathcal{S}_p^+$. Let the point $\bar{x} \in \mathcal{X}$ be feasible for this problem. Furthermore, let $\mathbf{P}\mathbf{\Lambda}\mathbf{P}^\top$ be an ordered eigenvalue decomposition of $G(\bar{x}) + H(\bar{x})$ and let α, β, and γ denote the index sets of positive, zero, and negative eigenvalues of $\mathbf{\Lambda}$.

1. The point $\bar{x}$ is W-stationary for the corresponding problem (MPCC) in the sense of Definition 3.1 if and only if there are multipliers $\lambda \in \mathcal{Y}^\star$ and $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$ which solve the system

$$0 = \psi'(\bar{x}) + g'(\bar{x})^\star[\lambda] + G'(\bar{x})^\star[\mathbf{M}] + H'(\bar{x})^\star[\mathbf{N}], \tag{5.20a}$$
$$\lambda \in \mathcal{N}_C(G(\bar{x})), \tag{5.20b}$$
$$\mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O}, \tag{5.20c}$$
$$\mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}. \tag{5.20d}$$

2. The point $\bar{x}$ is M-stationary for the corresponding problem (MPCC) in the sense of Definition 3.2 if and only if there are multipliers $\lambda \in \mathcal{Y}^\star$ and $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$ which satisfy (5.20a), (5.20b), and

$$\mathbf{M}^{\mathbf{P}}_{\alpha\alpha} = \mathbf{O},\ \mathbf{M}^{\mathbf{P}}_{\alpha\beta} = \mathbf{O}, \tag{5.21a}$$
$$\mathbf{N}^{\mathbf{P}}_{\beta\gamma} = \mathbf{O},\ \mathbf{N}^{\mathbf{P}}_{\gamma\gamma} = \mathbf{O}, \tag{5.21b}$$
$$\mathbf{\Xi}_{\alpha\gamma} \bullet (\mathbf{M}^{\mathbf{P}}_{\alpha\gamma} - \mathbf{N}^{\mathbf{P}}_{\alpha\gamma}) + \mathbf{N}^{\mathbf{P}}_{\alpha\gamma} = \mathbf{O}, \tag{5.21c}$$
$$(\mathbf{M}^{\mathbf{P}}_{\beta\beta}, \mathbf{N}^{\mathbf{P}}_{\beta\beta}) \in \mathcal{N}_{\operatorname{gph} \mathcal{N}_{\mathcal{S}^+_{|\beta|}}}(\mathbf{O}, \mathbf{O}) \tag{5.21d}$$

   where $\mathbf{\Xi}$ is defined in (5.3).

3. The point $\bar{x}$ is S-stationary for the corresponding problem (MPCC) in the sense of Definition 3.1 if and only if there are multipliers $\lambda \in \mathcal{Y}^\star$ and $\mathbf{M}, \mathbf{N} \in \mathcal{S}_p$ which satisfy (5.20a), (5.20b), (5.21a), (5.21b), and

$$\mathbf{M}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^-_{|\beta|},$$
$$\mathbf{N}^{\mathbf{P}}_{\beta\beta} \in \mathcal{S}^+_{|\beta|}.$$

Note that our notions of W- and S-stationarity are substantially weaker than the ones used in [37], [124], and [127]. For the S-stationarity conditions, this phenomenon was depicted in [121] already. One reason for that might be the fact that $\mathcal{S}_p^+$ is not polyhedric. This points out that the generalized stationarity notions for W- and S-stationarity from Chapter 3 for nonpolyhedric complementarity cones may turn out to be too weak to yield good necessary optimality conditions. In this section, by means of a linearization approach, we overcame this drawback. A related idea was exploited in [121] and [124].
Using the notions of stationarity for semidefinite complementarity programs from Remark 5.11, we easily see that any M-stationary point is W-stationary as well. On the other hand, there is no implication between the S- and M-stationarity conditions. Thus, the implications (3.10) do not hold for semidefinite MPCCs. This shows once more that the generalized concepts of stationarity for (MPCC) and their relation to each other need to be discussed carefully for different choices for the complementarity cone.

5.2. Bilevel optimal control of linear ODEs with lower level control constraints

For some instance $T > 0$, we set $\Omega := (0, T)$. In this section, we want to use the KKT approach to derive necessary optimality conditions for the bilevel optimal control problem

$$\begin{aligned} F_0(x(T), y(T)) + \int_0^T F_1(t, x(t), y(t), u(t), v(t))\mathrm{d}t &\to \min_{x,u,y,v} \\ \nabla x(t) - \mathbf{C}_x x(t) - \mathbf{C}_u u(t) &= 0 \qquad \text{a.e. on } \Omega \qquad \text{(BOC)} \\ x(0) - x_0 &= 0 \\ (y, v) &\in \Psi(x, u) \end{aligned}$$

where $\Psi\colon AC^{1,2}(\Omega, \mathbb{R}^n) \times L^2(\Omega, \mathbb{R}^k) \rightrightarrows AC^{1,2}(\Omega, \mathbb{R}^m) \times L^2(\Omega, \mathbb{R}^l)$ denotes the solution set mapping of the parametric optimal control problem stated below:

$$\begin{aligned} \tfrac{1}{2} y(T) \cdot \big(\mathbf{R} y(T)\big) + \tfrac{1}{2} \int_0^T \Big[y(t) \cdot \big[\mathbf{R}_y y(t) + 2\mathbf{P} x(t)\big] + v(t) \cdot \big(\mathbf{R}_v v(t)\big)\Big] \mathrm{d}t &\to \min_{y,v} \\ \nabla y(t) - \mathbf{A}_x x(t) - \mathbf{B}_y y(t) - \mathbf{A}_u u(t) - \mathbf{B}_v v(t) &= 0 \qquad \text{a.e. on } \Omega \qquad (5.22) \\ y(0) - y_0 &= 0 \\ \mathbf{D}_u u(t) + \mathbf{D}_v v(t) - \mathbf{d}(t) &\geq 0 \qquad \text{a.e. on } \Omega. \end{aligned}$$

Many mathematical models arising from practical applications possess a hierarchical structure where at least one decision level is a dynamic program of ODEs, see e.g. [3, 4, 41, 58, 59, 74, 76, 77, 89] for examples and numerical solution approaches. However, there exist surprisingly few theoretical results and optimality conditions for such problems which are not based on discretization, see [14, 18, 130] and other publications by these authors. Here we want to study the bilevel optimal control problem (BOC) in order to reduce the size of this gap.
Below, we list our standing assumptions on the model problem.

Assumption 5.1. The function $F_0\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable, whereas the function $F_1\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^l \to \mathbb{R}$ is (Lebesgue-)measurable w.r.t. its first argument and continuously differentiable w.r.t. its last four arguments. Furthermore, there are scalars $C, C' > 0$ such that the following estimates hold for all $t \in (0, T)$, $\mathtt{x} \in \mathbb{R}^n$, $\mathtt{y} \in \mathbb{R}^m$, $\mathtt{u} \in \mathbb{R}^k$, and $\mathtt{v} \in \mathbb{R}^l$:

$$\begin{aligned} |F_1(t, \mathtt{x}, \mathtt{y}, \mathtt{u}, \mathtt{v})| &\leq C\big(1 + |\mathtt{x}|_2^2 + |\mathtt{y}|_2^2 + |\mathtt{u}|_2^2 + |\mathtt{v}|_2^2\big), \\ \big|\nabla_{(x,y,u,v)} F_1(t, \mathtt{x}, \mathtt{y}, \mathtt{u}, \mathtt{v})\big|_2 &\leq C'\big(1 + |\mathtt{x}|_2 + |\mathtt{y}|_2 + |\mathtt{u}|_2 + |\mathtt{v}|_2\big). \end{aligned}$$

The matrices $\mathbf{A}_x \in \mathbb{R}^{m\times n}$, $\mathbf{A}_u \in \mathbb{R}^{m\times k}$, $\mathbf{B}_y \in \mathbb{R}^{m\times m}$, $\mathbf{B}_v \in \mathbb{R}^{m\times l}$, $\mathbf{C}_x \in \mathbb{R}^{n\times n}$, $\mathbf{C}_u \in \mathbb{R}^{n\times k}$, $\mathbf{D}_u \in \mathbb{R}^{q\times k}$, $\mathbf{D}_v \in \mathbb{R}^{q\times l}$, $\mathbf{P} \in \mathbb{R}^{m\times n}$, $\mathbf{R}, \mathbf{R}_y \in \mathbb{R}^{m\times m}$, and $\mathbf{R}_v \in \mathbb{R}^{l\times l}$ are fixed. Furthermore, we assume that $\mathbf{R}$, $\mathbf{R}_y$, as well as $\mathbf{R}_v$ are symmetric and positive semidefinite, whereas $\mathbf{D}_v$ possesses full row rank q. Additionally, the function $\mathbf{d} \in L^2(\Omega, \mathbb{R}^q)$ and the initial states $x_0 \in \mathbb{R}^n$ as well as $y_0 \in \mathbb{R}^m$ are fixed. The decision spaces are fixed to $AC^{1,2}(\Omega, \mathbb{R}^n)$ and $AC^{1,2}(\Omega, \mathbb{R}^m)$ for the state functions x and y, as well as to $L^2(\Omega, \mathbb{R}^k)$ and $L^2(\Omega, \mathbb{R}^l)$ for the control functions u and v, respectively.

We want to mention that the upcoming results stay valid in the case where all the autonomous matrices are replaced by time-dependent ones under not too hard assumptions. However, for the purpose of simplicity, we restict ourselves to the investigation of the autonomous case. Furthermore, one can modify the lower level objective function such that it contains desired targets as long as its quadratic structure is preserved.
Let us transfer (BOC) into a bilevel progamming problem discussed in Section 4.2 which satisfies Assumption 4.4. We introduce the spaces $\mathcal{X} := AC^{1,2}(\Omega, \mathbb{R}^n) \times L^2(\Omega, \mathbb{R}^k)$, $\mathcal{Y} := AC^{1,2}(\Omega, \mathbb{R}^m) \times L^2(\Omega, \mathbb{R}^l)$, $\mathcal{W} := AC^{1,2}(\Omega, \mathbb{R}^n)$, as well as $\mathcal{Z} := AC^{1,2}(\Omega, \mathbb{R}^m) \times L^2(\Omega, \mathbb{R}^q)$ and define the mappings $F\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$,

$G\colon \mathcal{X} \to \mathcal{W}$, $f\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, and $g\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ as stated below for all $(x,u) \in \mathcal{X}$ and $(y,v) \in \mathcal{Y}$:

$$\begin{aligned}
F(x,u,y,v) &:= F_0(x(T),y(T)) + \int_0^T F_1(t,x(t),y(t),u(t),v(t))\mathrm{d}t,\\
G(x,u) &:= x(\cdot) - x_0 - \int_0^\cdot \Big[\mathbf{C}_x x(\tau) + \mathbf{C}_u u(\tau)\Big]\mathrm{d}\tau,\\
f(x,u,y,v) &:= \tfrac{1}{2} y(T)\cdot\big(\mathbf{R}y(T)\big) + \tfrac{1}{2}\int_0^T \Big[y(t)\cdot\big[\mathbf{R}_y y(t) + 2\mathbf{P}x(t)\big] + v(t)\cdot\big(\mathbf{R}_v v(t)\big)\Big]\mathrm{d}t,\\
g(x,u,y,v) &:= \left(y(\cdot) - y_0 - \int_0^\cdot \Big[\mathbf{A}_x x(\tau) + \mathbf{B}_y y(\tau) + \mathbf{A}_u u(\tau) + \mathbf{B}_v v(\tau)\Big]\mathrm{d}\tau,\ \mathbf{D}_u u(\cdot) + \mathbf{D}_v v(\cdot) - \mathbf{d}(\cdot)\right).
\end{aligned}$$

We introduce closed, convex cones $C \subseteq \mathcal{W}$ and $K \subseteq \mathcal{Z}$ by $C := \{0\}$ and $K := \{0\} \times L^2(\Omega,\mathbb{R}^q)^+_0$, respectively.
The function F is Fréchet differentiable by means of [117, Lemma 3.1(b)]. Due to its quadratic structure, f is twice continuously Fréchet differentiable, see Examples 2.27 and 2.28 which provide a strategy for the validation of this result. Reprising the argumentation which was used to prove Lemma A.6, we easily see that G and g are continuous affine operators and, thus, arbitrarily often continuously Fréchet differentiable. Due to the positive semidefiniteness of the matrices $\mathbf{R}$, $\mathbf{R}_y$, and $\mathbf{R}_v$, the mapping $f(x,u,\cdot,\cdot)$ is convex for fixed $(x,u) \in \mathcal{X}$. Since g is affine, it is also $-K$-convex.
In the upcoming lemma, we show that g possesses a surjective partial Fréchet derivative w.r.t. the lower level decision variables which implies that KRZCQ holds at all lower level feasible points.

Lemma 5.12. For any point $(\bar{x},\bar{u},\bar{y},\bar{v}) \in \mathcal{X} \times \mathcal{Y}$, we have

$$\forall (h_y,h_v) \in \mathcal{Y}\colon \quad g'_{(y,v)}(\bar{x},\bar{u},\bar{y},\bar{v})[h_y,h_v] = \left(h_y(\cdot) - \int_0^T \Big[\mathbf{B}_y h_y(\tau) + \mathbf{B}_v h_v(\tau)\Big]\mathrm{d}\tau,\ \mathbf{D}_v h_v(\cdot)\right)$$

and this operator is surjective.

Proof. The representation of the Fréchet derivative follows from the fact that g is an affine and continuous mapping.
For the proof of the surjectivity of $\mathtt{G} := g'_{(y,v)}(\bar{x},\bar{u},\bar{y},\bar{v})$, we choose $(z_1,z_2) \in \mathcal{Z}$ arbitrarily and consider the equation $\mathtt{G}[h_y,h_v] = (z_1,z_2)$ for variables $(h_y,h_v) \in \mathcal{Y}$. Due to Assumption 5.1, $\mathbf{D}_v$ possesses full row rank q. Thus, we can set $\bar{h}_v(\cdot) := \mathbf{D}_v^\top\big(\mathbf{D}_v\mathbf{D}_v^\top\big)^{-1} z_2(\cdot) = \mathbf{D}_v^\dagger z_2(\cdot)$. Next, consider

$$h_y(\cdot) - \int_0^\cdot \mathbf{B}_y h_y(\tau)\mathrm{d}\tau = z_1(\cdot) + \int_0^\cdot \mathbf{B}_v\mathbf{D}_v^\dagger z_2(\tau)\mathrm{d}\tau.$$

Since the right hand side of this equation belongs to $AC^{1,2}(\Omega,\mathbb{R}^m)$, a solution $\bar{h}_y \in AC^{1,2}(\Omega,\mathbb{R}^m)$ of it can be explicitly constructed using the fundamental matrix function $\Phi\colon \Omega \to \mathbb{R}^{m\times m}$ given as the solution of the matrix differential equation

$$\nabla\Phi(t) = \mathbf{B}_y\Phi(t) \quad \text{f.a.a. } t \in \Omega, \qquad\qquad \Phi(0) = \mathbf{I}_m,$$

see [71, proof of Theorem 5.19]. Thus, we have $\mathtt{G}[\bar{h}_y,\bar{h}_v] = (z_1,z_2)$ which shows the surjectivity of $\mathtt{G}$. □

Combining the above observations and keeping Assumption 4.4 in mind, it is reasonable to discuss the model problem (BOC) using the KKT approach from Section 4.2. Note that by means of Corollary 4.23 and Lemma 5.12 we already know that the bilevel optimal control problem (BOC) and its KKT reformulation are equivalent w.r.t. local and global optimal solutions.

5.2.1. The lower level KKT conditions and the KKT reformulation of the original problem

In order to study the KKT reformulation of (BOC), we have to derive the KKT conditions of the lower level problem (5.22) first. Therefore, we fix a lower level feasible point $\mathfrak{p} := (\bar{x}, \bar{u}, \bar{y}, \bar{v}) \in \mathcal{X} \times \mathcal{Y}$ and observe that

$$f'_{(y,v)}(\mathfrak{p})[h_y, h_v] = h_y(T) \cdot \big(\mathbf{R}\bar{y}(T)\big) + \int_0^T h_y(t) \cdot \big[\mathbf{R}_y\bar{y}(t) + \mathbf{P}\bar{x}(t)\big]\mathrm{d}t + \int_0^T h_v(t) \cdot \big(\mathbf{R}_v\bar{v}(t)\big)\mathrm{d}t$$

holds true for arbitrary $(h_y, h_v) \in \mathcal{Y}$, see Example 2.27. We apply Lemma A.5 in order to obtain the alternative representation

$$f'_{(y,v)}(\mathfrak{p}) = \left(\left(\mathbf{R}\bar{y}(T) + \int_0^T \big[\mathbf{R}_y\bar{y}(\tau) + \mathbf{P}\bar{x}(\tau)\big]\mathrm{d}\tau,\, \mathbf{R}\bar{y}(T) + \int_\cdot^T \big[\mathbf{R}_y\bar{y}(\tau) + \mathbf{P}\bar{x}(\tau)\big]\mathrm{d}\tau\right),\, \mathbf{R}_v\bar{v}(\cdot)\right).$$

Combining Lemmas 5.12 and A.6, we have

$$g'_{(y,v)}(\mathfrak{p})^\star[\theta_y, \lambda] = \left(\left(\theta_y(0) - \int_0^T \mathbf{B}_y^\top\nabla\theta_y(\tau)\mathrm{d}\tau,\, \nabla\theta_y(\cdot) - \int_\cdot^T \mathbf{B}_y^\top\nabla\theta_y(\tau)\mathrm{d}\tau\right),\, \mathbf{D}_v^\top\lambda(\cdot) - \mathbf{B}_v^\top\nabla\theta_y(\cdot)\right)$$

for any $(\theta_y, \lambda) \in \mathcal{Z}^\star \cong \mathcal{Z}$. Consequently, the lower level KKT conditions reduce to the existence of functions $\theta_y \in AC^{1,2}(\Omega, \mathbb{R}^m)$ and $\lambda \in L^2(\Omega, \mathbb{R}^q)$ which satisfy

$$0 = \mathbf{R}\bar{y}(T) + \theta_y(0) + \int_0^T \big[\mathbf{R}_y\bar{y}(\tau) + \mathbf{P}\bar{x}(\tau) - \mathbf{B}_y^\top\nabla\theta_y(\tau)\big]\mathrm{d}\tau, \tag{5.23a}$$

$$0 = \mathbf{R}\bar{y}(T) + \nabla\theta_y(t) + \int_t^T \big[\mathbf{R}_y\bar{y}(\tau) + \mathbf{P}\bar{x}(\tau) - \mathbf{B}_y^\top\nabla\theta_y(\tau)\big]\mathrm{d}\tau \quad \text{f.a.a. } t \in \Omega, \tag{5.23b}$$

$$0 = \mathbf{R}_v\bar{v}(t) + \mathbf{D}_v^\top\lambda(t) - \mathbf{B}_v^\top\nabla\theta_y(t) \quad \text{f.a.a. } t \in \Omega, \tag{5.23c}$$

$$0 \geq \lambda(t) \quad \text{f.a.a. } t \in \Omega, \tag{5.23d}$$

$$0 = \lambda(t) \cdot \big(\mathbf{D}_u\bar{u}(t) + \mathbf{D}_v\bar{v}(t) - \mathbf{d}(t)\big) \quad \text{f.a.a. } t \in \Omega, \tag{5.23e}$$

see Section 3.2 for the derivation of the expression of the complementarity conditions (5.23d), (5.23e). Equation (5.23b) yields that $p := \nabla\theta_y$ solves the boundary value problem

$$0 = \nabla p(t) - \mathbf{R}_y\bar{y}(t) - \mathbf{P}\bar{x}(t) + \mathbf{B}_y^\top p(t) \quad \text{f.a.a. } t \in \Omega, \tag{5.24a}$$

$$0 = p(T) + \mathbf{R}\bar{y}(T) \tag{5.24b}$$

and, thus, is already an element of $AC^{1,2}(\Omega, \mathbb{R}^m)$. On the other hand, if $p \in AC^{1,2}(\Omega, \mathbb{R}^m)$ solves (5.24), then defining $\theta_y(t) := p(0) + \int_0^t p(\tau)\mathrm{d}\tau$ for all $t \in \Omega$ produces a function which satisfies the conditions (5.23a) and (5.23b).
Consequently, the KKT conditions (5.23) are equivalent to the existence of functions $p \in AC^{1,2}(\Omega, \mathbb{R}^m)$ and $\lambda \in L^2(\Omega, \mathbb{R}^q)$ which satisfy (5.23d), (5.23e), (5.24), and

$$0 = \mathbf{R}_v\bar{v}(t) + \mathbf{D}_v^\top\lambda(t) - \mathbf{B}_v^\top p(t) \quad \text{f.a.a. } t \in \Omega. \tag{5.25}$$

These optimality conditions for (5.22) are of so-called Pontryagin-type, see [102], since they comprise an adjoint equation (5.24a) on the so-called adjoint state p, transversality conditions (i.e. boundary conditions) (5.24b) on p, and Pontryagin's (linearized) Maximum Principle (5.25) where the appearing Lagrange multiplier λ is characterized in (5.23d), (5.23e) .
The above considerations show that the KKT reformulation of the bilevel programming problem (BOC) is

equivalent to the optimal control problem

$$\begin{aligned}
F_0(x(T),y(T)) + \int_0^T F_1(t,x(t),y(t),u(t),v(t))\mathrm{d}t &\to \min_{x,u,y,v,p,\lambda,\xi} \\
\nabla x(t) - \mathbf{C}_x x(t) - \mathbf{C}_u u(t) &= 0 && \text{a.e. on } \Omega \\
\nabla y(t) - \mathbf{A}_x x(t) - \mathbf{B}_y y(t) - \mathbf{A}_u u(t) - \mathbf{B}_v v(t) &= 0 && \text{a.e. on } \Omega \\
\nabla p(t) - \mathbf{R}_y y(t) - \mathbf{P} x(t) + \mathbf{B}_y^\top p(t) &= 0 && \text{a.e. on } \Omega \\
x(0) - x_0 &= 0 \\
y(0) - y_0 &= 0 \\
p(0) - \xi &= 0 \\
p(T) + \mathbf{R} y(T) &= 0 \\
\mathbf{R}_v v(t) + \mathbf{D}_v^\top \lambda(t) - \mathbf{B}_v^\top p(t) &= 0 && \text{a.e. on } \Omega \\
\mathbf{D}_u u(t) + \mathbf{D}_v v(t) - \mathbf{d}(t) &\geq 0 && \text{a.e. on } \Omega \\
\lambda(t) &\leq 0 && \text{a.e. on } \Omega \\
\lambda(t) \cdot \big(\mathbf{D}_u u(t) + \mathbf{D}_v v(t) - \mathbf{d}(t)\big) &= 0 && \text{a.e. on } \Omega
\end{aligned} \tag{5.26}$$

which comprises a complementarity constraint in the Lebesgue space $L^2(\Omega,\mathbb{R}^q)$. Note that we introduced a dummy variable $\xi \in \mathbb{R}^m$. This will be beneficial when deriving applicable constraint qualifications to handle (5.26). Recall that the state functions in this problem are x, y, and p, whereas the control functions are given by u, v, and λ. From Theorem 4.18 we know that if $(\bar{x},\bar{u},\bar{y},\bar{v}) \in \mathcal{X} \times \mathcal{Y}$ is a local optimal solution of (BOC), then there are $p \in AC^{1,2}(\Omega,\mathbb{R}^m)$ and $\lambda \in L^2(\Omega,\mathbb{R}^q)$ such that $(\bar{x},\bar{u},\bar{y},\bar{v},p,\lambda,p(0))$ is a local optimal solution of (5.26).

In order to write the above program in a compact way, we introduce matrices $\hat{\mathbf{M}} \in \mathbb{R}^{(n+2m)\times(n+2m)}$, $\hat{\mathbf{N}} \in \mathbb{R}^{(n+2m)\times(k+l+q)}$, $\hat{\mathbf{K}} \in \mathbb{R}^{(n+2m)\times m}$, $\hat{\mathbf{P}} \in \mathbb{R}^{l\times(n+2m)}$, $\hat{\mathbf{Q}} \in \mathbb{R}^{l\times(k+l+q)}$, and $\hat{\mathbf{R}} \in \mathbb{R}^{m\times(n+2m)}$ as stated below:

$$\hat{\mathbf{M}} := \begin{bmatrix} \mathbf{C}_x & \mathbf{O} & \mathbf{O} \\ \mathbf{A}_x & \mathbf{B}_y & \mathbf{O} \\ \mathbf{P} & \mathbf{R}_y & -\mathbf{B}_y^\top \end{bmatrix}, \qquad \hat{\mathbf{N}} := \begin{bmatrix} \mathbf{C}_u & \mathbf{O} & \mathbf{O} \\ \mathbf{A}_u & \mathbf{B}_v & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{bmatrix}, \qquad \hat{\mathbf{K}} := \begin{bmatrix} \mathbf{O} \\ \mathbf{O} \\ \mathbf{I}_m \end{bmatrix},$$
$$\hat{\mathbf{P}} := \begin{bmatrix} \mathbf{O} & \mathbf{O} & -\mathbf{B}_v^\top \end{bmatrix}, \qquad \hat{\mathbf{Q}} := \begin{bmatrix} \mathbf{O} & \mathbf{R}_v & \mathbf{D}_v^\top \end{bmatrix}, \qquad \hat{\mathbf{R}} := \begin{bmatrix} \mathbf{O} & \mathbf{R} & \mathbf{I}_m \end{bmatrix}.$$

Using these matrices, we can write (5.26) equivalently as

$$\begin{aligned}
F_0(x(T),y(T)) + \int_0^T F_1(t,x(t),y(t),u(t),v(t))\mathrm{d}t &\to \min_{x,y,p,u,v,\lambda,\xi} \\
\begin{pmatrix} x(t) \\ y(t) \\ p(t) \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ 0 \end{pmatrix} - \hat{\mathbf{K}}\xi - \int_0^t \left[\hat{\mathbf{M}} \begin{pmatrix} x(\tau) \\ y(\tau) \\ p(\tau) \end{pmatrix} + \hat{\mathbf{N}} \begin{pmatrix} u(\tau) \\ v(\tau) \\ \lambda(\tau) \end{pmatrix} \right] \mathrm{d}\tau &= 0 && \text{a.e. on } \Omega \\
\hat{\mathbf{P}} \begin{pmatrix} x(t) \\ y(t) \\ p(t) \end{pmatrix} + \hat{\mathbf{Q}} \begin{pmatrix} u(t) \\ v(t) \\ \lambda(t) \end{pmatrix} &= 0 && \text{a.e. on } \Omega \\
\hat{\mathbf{R}} \begin{pmatrix} x(T) \\ y(T) \\ p(T) \end{pmatrix} &= 0 \\
\mathbf{D}_u u(t) + \mathbf{D}_v v(t) - \mathbf{d}(t) &\geq 0 && \text{a.e. on } \Omega \\
\lambda(t) &\leq 0 && \text{a.e. on } \Omega \\
\lambda(t) \cdot \big(\mathbf{D}_u u(t) + \mathbf{D}_v v(t) - \mathbf{d}(t)\big) &= 0 && \text{a.e. on } \Omega.
\end{aligned}$$

5.2.2. The W- and S-stationarity conditions of the bilevel optimal control problem

In this section, we want to derive explicit representations of the W- and S-stationarity conditions which correspond to the bilevel programming problem (BOC). Recalling Definition 4.1, we only need to state

the W- and S-stationarity conditions of the MPCC (5.26) for that purpose. Note that the complementarity constraint of this program is induced by the nonnegative cone in $L^2(\Omega, \mathbb{R}^q)$ and, thus, we already know that the consideration of the M-stationarity concept is not reasonable here since it coincides with W-stationarity and the corresponding constraint qualifications are not applicable, see Section 3.2 for the details.
In order to stay close to the notation used in Chapter 3, we define the Banach spaces

$$\begin{aligned}\tilde{\mathcal{X}} &:= AC^{1,2}(\Omega, \mathbb{R}^{n+2m}) \times L^2(\Omega, \mathbb{R}^{k+l+q}) \times \mathbb{R}^m,\\ \tilde{\mathcal{Y}} &:= AC^{1,2}(\Omega, \mathbb{R}^{n+2m}) \times L^2(\Omega, \mathbb{R}^{l}) \times \mathbb{R}^m,\\ \tilde{\mathcal{Z}} &:= L^2(\Omega, \mathbb{R}^{q}).\end{aligned}$$

We introduce the notation $\mathfrak{x} := (x, y, p, u, v, \lambda, \xi) \in \tilde{\mathcal{X}}$ as well as functions $\psi\colon \tilde{\mathcal{X}} \to \mathbb{R}$, $\tilde{g}\colon \tilde{\mathcal{X}} \to \tilde{\mathcal{Y}}$, $\tilde{G}\colon \tilde{\mathcal{X}} \to \tilde{\mathcal{Z}}$, and $\tilde{H}\colon \tilde{\mathcal{X}} \to \tilde{\mathcal{Z}}^\star$ as stated below for arbitrary $\mathfrak{x} \in \tilde{\mathcal{X}}$:

$$\begin{aligned}\psi(\mathfrak{x}) &:= F(x, u, y, v),\\ \tilde{g}(\mathfrak{x}) &:= \left(\begin{pmatrix} x(\cdot)\\ y(\cdot)\\ p(\cdot)\end{pmatrix} - \begin{pmatrix} x_0\\ y_0\\ 0\end{pmatrix} - \hat{\mathbf{K}}\xi - \int_0^\cdot \left[\hat{\mathbf{M}} \begin{pmatrix} x(\tau)\\ y(\tau)\\ p(\tau)\end{pmatrix} + \hat{\mathbf{N}} \begin{pmatrix} u(\tau)\\ v(\tau)\\ \lambda(\tau)\end{pmatrix} \right] \mathrm{d}\tau, \right.\\ &\qquad \left. \hat{\mathbf{P}} \begin{pmatrix} x(\cdot)\\ y(\cdot)\\ p(\cdot)\end{pmatrix} + \hat{\mathbf{Q}} \begin{pmatrix} u(\cdot)\\ v(\cdot)\\ \lambda(\cdot)\end{pmatrix}, \hat{\mathbf{R}} \begin{pmatrix} x(T)\\ y(T)\\ p(T)\end{pmatrix} \right),\\ \tilde{G}(\mathfrak{x}) &:= \mathbf{D}_u u(\cdot) + \mathbf{D}_v v(\cdot) - \mathbf{d}(\cdot),\\ \tilde{H}(\mathfrak{x}) &:= \lambda(\cdot).\end{aligned}$$

Finally, we fix the closed, convex cones $\tilde{C} := \{0\}$ and $\tilde{K} := L^2(\Omega, \mathbb{R}^q)^+_0$ in order to see that the KKT reformulation (5.26) of the bilevel programming problem (BOC) is equivalent to

$$\begin{aligned}\psi(\mathfrak{x}) &\to \min\\ \tilde{g}(\mathfrak{x}) &\in \tilde{C}\\ \tilde{G}(\mathfrak{x}) &\in \tilde{K}\\ \tilde{H}(\mathfrak{x}) &\in \tilde{K}^\circ\\ \left\langle \tilde{H}(\mathfrak{x}), \tilde{G}(\mathfrak{x}) \right\rangle_{\tilde{\mathcal{Z}}} &= 0.\end{aligned}$$

From Section 5.2.1 we already know that ψ is Fréchet differentiable. Lemma A.6 can be used to show that $\tilde{g}$, $\tilde{G}$, and $\tilde{H}$ are continuous, affine operators and, thus, continuously Fréchet differentiable.
Below, we characterize the Fréchet derivatives of the functions ψ, $\tilde{g}$, $\tilde{G}$, and $\tilde{H}$ at some point $\mathfrak{x} \in \tilde{\mathcal{X}}$. Therefore, we fix some direction $\mathfrak{d} := (d_x, d_y, d_p, d_u, d_v, d_\lambda, d_\xi) \in \tilde{\mathcal{X}}$. Using [117, Lemma 3.1], we obtain

$$\begin{aligned}\psi'(\mathfrak{x}) = \Bigg(&\left(\nabla_x F_0(x(T), y(T)) + \int_0^T \nabla_x F_1(t, x(t), y(t), u(t), v(t)) \mathrm{d}t, \right.\\ &\qquad \left. \nabla_x F_0(x(T), y(T)) + \int_\cdot^T \nabla_x F_1(t, x(t), y(t), u(t), v(t)) \mathrm{d}t \right),\\ &\left(\nabla_y F_0(x(T), y(T)) + \int_0^T \nabla_y F_1(t, x(t), y(t), u(t), v(t)) \mathrm{d}t, \right.\\ &\qquad \left. \nabla_y F_0(x(T), y(T)) + \int_\cdot^T \nabla_y F_1(t, x(t), y(t), u(t), v(t)) \mathrm{d}t \right),\\ &(0,0),\ \nabla_u F_1(\cdot, x(\cdot), y(\cdot), u(\cdot), v(\cdot)),\ \nabla_v F_1(\cdot, x(\cdot), y(\cdot), u(\cdot), v(\cdot)),\ 0,\ 0 \Bigg).\end{aligned}$$

The affine structure of $\tilde{g}$, $\tilde{G}$, and $\tilde{H}$ makes it easy to see

$$\begin{aligned}\tilde{g}'(\mathfrak{x})[\mathfrak{d}] = \Bigg(&\begin{pmatrix} d_x(\cdot) \\ d_y(\cdot) \\ d_p(\cdot) \end{pmatrix} - \hat{\mathbf{K}} d_\xi - \int_0^\cdot \left[\hat{\mathbf{M}} \begin{pmatrix} d_x(\tau) \\ d_y(\tau) \\ d_p(\tau) \end{pmatrix} + \hat{\mathbf{N}} \begin{pmatrix} d_u(\tau) \\ d_v(\tau) \\ d_\lambda(\tau) \end{pmatrix} \right] \mathrm{d}\tau, \\ &\hat{\mathbf{P}} \begin{pmatrix} d_x(\cdot) \\ d_y(\cdot) \\ d_p(\cdot) \end{pmatrix} + \hat{\mathbf{Q}} \begin{pmatrix} d_u(\cdot) \\ d_v(\cdot) \\ d_\lambda(\cdot) \end{pmatrix}, \hat{\mathbf{R}} \begin{pmatrix} d_x(T) \\ d_y(T) \\ d_p(T) \end{pmatrix} \Bigg), \\ \tilde{G}'(\mathfrak{x})[\mathfrak{d}] &:= \mathbf{D}_u d_u(\cdot) + \mathbf{D}_v d_v(\cdot), \\ \tilde{H}'(\mathfrak{x})[\mathfrak{d}] &:= d_\lambda(\cdot).\end{aligned}$$

Now, we characterize the corresponding adjoint operators. First, we recall $\tilde{\mathcal{Y}}^\star \cong \tilde{\mathcal{Y}}$ and choose a vector $\mathfrak{w} := (w_x, w_y, w_p, \upsilon, s) \in \tilde{\mathcal{Y}}$ arbitrarily. Then Lemma A.6 yields

$$\begin{aligned}\tilde{g}'(\mathfrak{x})^\star[\mathfrak{w}] = \Bigg(&\Bigg(\begin{pmatrix} w_x(0) \\ w_y(0) \\ w_p(0) \end{pmatrix} + \hat{\mathbf{R}}^\top s + \int_0^T \left[\hat{\mathbf{P}}^\top \upsilon(\tau) - \hat{\mathbf{M}}^\top \begin{pmatrix} \nabla w_x(\tau) \\ \nabla w_y(\tau) \\ \nabla w_p(\tau) \end{pmatrix} \right] \mathrm{d}\tau, \\ &\begin{pmatrix} \nabla w_x(\cdot) \\ \nabla w_y(\cdot) \\ \nabla w_p(\cdot) \end{pmatrix} + \hat{\mathbf{R}}^\top s + \int_\cdot^T \left[\hat{\mathbf{P}}^\top \upsilon(\tau) - \hat{\mathbf{M}}^\top \begin{pmatrix} \nabla w_x(\tau) \\ \nabla w_y(\tau) \\ \nabla w_p(\tau) \end{pmatrix} \right] \mathrm{d}\tau \Bigg), \\ &\hat{\mathbf{Q}}^\top \upsilon(\cdot) - \hat{\mathbf{N}}^\top \begin{pmatrix} \nabla w_x(\cdot) \\ \nabla w_y(\cdot) \\ \nabla w_p(\cdot) \end{pmatrix}, -\hat{\mathbf{K}}^\top \begin{pmatrix} w_x(0) \\ w_y(0) \\ w_p(0) \end{pmatrix} \Bigg) \\ = \Bigg(&\Bigg(w_x(0) - \int_0^T \left[\mathbf{C}_x^\top \nabla w_x(\tau) + \mathbf{A}_x^\top \nabla w_y(\tau) + \mathbf{P}^\top \nabla w_p(\tau) \right] \mathrm{d}\tau, \\ &\nabla w_x(\cdot) - \int_\cdot^T \left[\mathbf{C}_x^\top \nabla w_x(\tau) + \mathbf{A}_x^\top \nabla w_y(\tau) + \mathbf{P}^\top \nabla w_p(\tau) \right] \mathrm{d}\tau \Bigg), \\ &\Bigg(w_y(0) + \mathbf{R} s - \int_0^T \left[\mathbf{B}_y^\top \nabla w_y(\tau) + \mathbf{R}_y \nabla w_p(\tau) \right] \mathrm{d}\tau, \\ &\nabla w_y(\cdot) + \mathbf{R} s - \int_\cdot^T \left[\mathbf{B}_y^\top \nabla w_y(\tau) + \mathbf{R}_y \nabla w_p(\tau) \right] \mathrm{d}\tau \Bigg), \\ &\Bigg(w_p(0) + s - \int_0^T \left[\mathbf{B}_v \upsilon(\tau) - \mathbf{B}_y \nabla w_p(\tau) \right] \mathrm{d}\tau, \\ &\nabla w_p(\cdot) + s - \int_\cdot^T \left[\mathbf{B}_v \upsilon(\tau) - \mathbf{B}_y \nabla w_p(\tau) \right] \mathrm{d}\tau \Bigg), \\ &- \mathbf{C}_u^\top \nabla w_x(\cdot) - \mathbf{A}_u^\top \nabla w_y(\cdot), \mathbf{R}_v \upsilon(\cdot) - \mathbf{B}_v^\top \nabla w_y(\cdot), \mathbf{D}_v \upsilon(\cdot), -w_p(0) \Bigg).\end{aligned}$$

For $z \in L^2(\Omega, \mathbb{R}^q)$, we obtain

$$\begin{aligned}\tilde{G}'(\mathfrak{x})^\star[z] &= \left(0,\, 0,\, 0,\, \mathbf{D}_u^\top z(\cdot),\, \mathbf{D}_v^\top z(\cdot),\, 0,\, 0\right), \\ \tilde{H}'(\mathfrak{x})^\star[z] &= (0,\, 0,\, 0,\, 0,\, 0,\, z,\, 0)\,.\end{aligned}$$

We are well-prepared to state the W- and S-stationarity conditions for the bilevel optimal control problem of interest. Let $\bar{\mathfrak{p}} = (\bar{x}, \bar{u}, \bar{y}, \bar{v}) \in \mathcal{X} \times \mathcal{Y}$ be a fixed feasible point of the bilevel optimal control problem (BOC). Then we find $\bar{p} \in AC^{1,2}(\Omega, \mathbb{R}^m)$ and $\bar{\lambda} \in L^2(\Omega, \mathbb{R}^q)$ such that $\bar{\mathfrak{x}} = (\bar{x}, \bar{y}, \bar{p}, \bar{u}, \bar{v}, \bar{\lambda}, \bar{p}(0)) \in \tilde{\mathcal{X}}$ is a local optimal solution of (5.26). Similar as in Section 3.2, for all $i \in Q := \{1, \ldots, q\}$, we define measurable sets $I^{+0}(\bar{\mathfrak{x}}, i)$, $I^{0-}(\bar{\mathfrak{x}}, i)$, and $I^{00}(\bar{\mathfrak{x}}, i)$ as stated below:

$$\begin{aligned}I^{+0}(\bar{\mathfrak{x}}, i) &:= \{t \in \Omega \,|\, \tilde{G}(\bar{\mathfrak{x}})_i(t) > 0,\, \tilde{H}(\bar{\mathfrak{x}})_i(t) = 0\}, \\ I^{0-}(\bar{\mathfrak{x}}, i) &:= \{t \in \Omega \,|\, \tilde{G}(\bar{\mathfrak{x}})_i(t) = 0,\, \tilde{H}(\bar{\mathfrak{x}})_i(t) < 0\}, \\ I^{00}(\bar{\mathfrak{x}}, i) &:= \{t \in \Omega \,|\, \tilde{G}(\bar{\mathfrak{x}})_i(t) = 0,\, \tilde{H}(\bar{\mathfrak{x}})_i(t) = 0\}.\end{aligned} \tag{5.27}$$

Then we know from Theorem 3.14 that $\bar{\mathfrak{x}}$ is W-stationary for (5.26) if and only if there are $\mathfrak{w} \in \tilde{\mathcal{Y}}^\star$ and $\mu, \nu \in L^2(\Omega, \mathbb{R}^q)$ which satisfy the following set of conditions:

$$
\begin{aligned}
0 &= \nabla_x F_0(\bar{x}(T), \bar{y}(T)) + w_x(0) \\
&\quad + \int_0^T \big[\nabla_x F_1(\tau, \bar{x}(\tau), \bar{y}(\tau), \bar{u}(\tau), \bar{v}(\tau)) - \mathbf{C}_x^\top \nabla w_x(\tau) - \mathbf{A}_x^\top \nabla w_y(\tau) - \mathbf{P}^\top \nabla w_p(\tau)\big] \mathrm{d}\tau, && \text{(5.28a)} \\
0 &= \nabla_x F_0(\bar{x}(T), \bar{y}(T)) + \nabla w_x(\cdot) \\
&\quad + \int_\cdot^T \big[\nabla_x F_1(\tau, \bar{x}(\tau), \bar{y}(\tau), \bar{u}(\tau), \bar{v}(\tau)) - \mathbf{C}_x^\top \nabla w_x(\tau) - \mathbf{A}_x^\top \nabla w_y(\tau) - \mathbf{P}^\top \nabla w_p(\tau)\big] \mathrm{d}\tau, && \text{(5.28b)} \\
0 &= \nabla_y F_0(\bar{x}(T), \bar{y}(T)) + w_y(0) + \mathbf{R}s \\
&\quad + \int_0^T \big[\nabla_y F_1(\tau, \bar{x}(\tau), \bar{y}(\tau), \bar{u}(\tau), \bar{v}(\tau)) - \mathbf{B}_y^\top \nabla w_y(\tau) - \mathbf{R}_y \nabla w_p(\tau)\big] \mathrm{d}\tau, && \text{(5.28c)} \\
0 &= \nabla_y F_0(\bar{x}(T), \bar{y}(T)) + \nabla w_y(\cdot) + \mathbf{R}s \\
&\quad + \int_\cdot^T \big[\nabla_y F_1(\tau, \bar{x}(\tau), \bar{y}(\tau), \bar{u}(\tau), \bar{v}(\tau)) - \mathbf{B}_y^\top \nabla w_y(\tau) - \mathbf{R}_y \nabla w_p(\tau)\big] \mathrm{d}\tau, && \text{(5.28d)} \\
0 &= w_p(0) + s + \int_0^T \big[\mathbf{B}_y \nabla w_p(\tau) - \mathbf{B}_v \upsilon(\tau)\big] \mathrm{d}\tau, && \text{(5.28e)} \\
0 &= \nabla w_p(\cdot) + s + \int_\cdot^T \big[\mathbf{B}_y \nabla w_p(\tau) - \mathbf{B}_v \upsilon(\tau)\big] \mathrm{d}\tau, && \text{(5.28f)} \\
0 &= \nabla_u F_1(\cdot, \bar{x}(\cdot), \bar{y}(\cdot), \bar{u}(\cdot), \bar{v}(\cdot)) - \mathbf{C}_u^\top \nabla w_x(\cdot) - \mathbf{A}_u^\top \nabla w_y(\cdot) + \mathbf{D}_u^\top \mu(\cdot), && \text{(5.28g)} \\
0 &= \nabla_v F_1(\cdot, \bar{x}(\cdot), \bar{y}(\cdot), \bar{u}(\cdot), \bar{v}(\cdot)) + \mathbf{R}_v \upsilon(\cdot) - \mathbf{B}_v^\top \nabla w_y(\cdot) + \mathbf{D}_v^\top \mu(\cdot), && \text{(5.28h)} \\
0 &= \mathbf{D}_v \upsilon(\cdot) + \nu(\cdot), && \text{(5.28i)} \\
0 &= -w_p(0), && \text{(5.28j)} \\
\forall i \in Q\colon &\quad \begin{array}{ll} \mu_i(t) = 0 & \text{f.a.a. } t \in I^{+0}(\bar{\mathfrak{x}}, i), \\ \nu_i(t) = 0 & \text{f.a.a. } t \in I^{0-}(\bar{\mathfrak{x}}, i). \end{array} && \text{(5.28k)}
\end{aligned}
$$

For the S-stationarity conditions, the additional condition

$$\forall i \in Q\colon \quad \mu_i(t) \leq 0,\, \nu_i(t) \geq 0 \quad \text{f.a.a. } t \in \Omega$$

needs to be satisfied as well.
Similar as in Section 5.2.1 we can show that introducing $\phi_x \in AC^{1,2}(\Omega, \mathbb{R}^n)$, $\phi_y \in AC^{1,2}(\Omega, \mathbb{R}^m)$, and $\phi_p \in AC^{1,2}(\Omega, \mathbb{R}^m)$ by means of $\phi_x := \nabla w_x$, $\phi_y := \nabla w_y$, and $\phi_p := \nabla w_p$, the conditions (5.28a) - (5.28f) and (5.28j) are equivalent to

$$
\begin{aligned}
0 &= \nabla \phi_x(t) + \mathbf{C}_x^\top \phi_x(t) + \mathbf{A}_x^\top \phi_y(t) + \mathbf{P}^\top \phi_p(t) - \nabla_x F_1(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{v}(t)) && \text{f.a.a. } t \in \Omega, \\
0 &= \nabla \phi_y(t) + \mathbf{B}_y^\top \phi_y(t) + \mathbf{R}_y \phi_p(t) - \nabla_y F_1(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{v}(t)) && \text{f.a.a. } t \in \Omega, \\
0 &= \nabla \phi_p(t) - \mathbf{B}_y \phi_p(t) + \mathbf{B}_v \upsilon(t) && \text{f.a.a. } t \in \Omega, \\
0 &= \phi_p(0), \\
0 &= \phi_x(T) + \nabla_x F_0(\bar{x}(T), \bar{y}(T)), \\
0 &= \phi_y(T) - \mathbf{R} \phi_p(T) + \nabla_y F_0(\bar{x}(T), \bar{y}(T)).
\end{aligned}
$$

Adding the lower level KKT system to this set of conditions, we obtain the following result.

Proposition 5.13. Let $\bar{\mathfrak{p}} := (\bar{x}, \bar{u}, \bar{y}, \bar{v}) \in \mathcal{X} \times \mathcal{Y}$ be a feasible point of the bilevel optimal control problem (BOC). The point $\bar{\mathfrak{p}}$ is a W-stationary point of the bilevel programming problem if and only if there exist functions $\phi_x \in AC^{1,2}(\Omega, \mathbb{R}^n)$, $\bar{p}, \phi_y, \phi_p \in AC^{1,2}(\Omega, \mathbb{R}^m)$, $\upsilon \in L^2(\Omega, \mathbb{R}^l)$, and $\bar{\lambda}, \mu \in L^2(\Omega, \mathbb{R}^q)$ which satisfy

the following set of conditions:

$$\begin{aligned}
0 &= \nabla\bar{p}(t) + \mathbf{B}_y^\top \bar{p}(t) - \mathbf{P}\bar{x}(t) - \mathbf{R}_y\bar{y}(t) && \text{f.a.a. } t \in \Omega,\\
0 &= \nabla\phi_x(t) + \mathbf{C}_x^\top \phi_x(t) + \mathbf{A}_x^\top \phi_y(t) + \mathbf{P}^\top \phi_p(t) - \nabla_x F_1(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{v}(t)) && \text{f.a.a. } t \in \Omega,\\
0 &= \nabla\phi_y(t) + \mathbf{B}_y^\top \phi_y(t) + \mathbf{R}_y\phi_p(t) - \nabla_y F_1(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{v}(t)) && \text{f.a.a. } t \in \Omega,\\
0 &= \nabla\phi_p(t) - \mathbf{B}_y\phi_p(t) + \mathbf{B}_v\upsilon(t) && \text{f.a.a. } t \in \Omega,
\end{aligned} \tag{5.29a}$$

$$\begin{aligned}
0 &= \phi_p(0),\\
0 &= \mathbf{R}\bar{y}(T) + \bar{p}(T),\\
0 &= \phi_x(T) + \nabla_x F_0(\bar{x}(T), \bar{y}(T)),\\
0 &= \phi_y(T) - \mathbf{R}\phi_p(T) + \nabla_y F_0(\bar{x}(T), \bar{y}(T)),
\end{aligned} \tag{5.29b}$$

$$\begin{aligned}
0 &= \mathbf{R}_v\bar{v}(t) + \mathbf{D}_v^\top \bar{\lambda}(t) - \mathbf{B}_v^\top \bar{p}(t) && \text{f.a.a. } t \in \Omega,\\
0 &= \nabla_u F_1(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{v}(t)) - \mathbf{C}_u^\top \phi_x(t) - \mathbf{A}_u^\top \phi_y(t) + \mathbf{D}_u^\top \mu(t) && \text{f.a.a. } t \in \Omega,\\
0 &= \nabla_v F_1(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{v}(t)) + \mathbf{R}_v\upsilon(t) - \mathbf{B}_v^\top \phi_y(t) + \mathbf{D}_v^\top \mu(t) && \text{f.a.a. } t \in \Omega,
\end{aligned} \tag{5.29c}$$

$$\begin{aligned}
0 &\leq \mathbf{D}_u\bar{u}(t) + \mathbf{D}_v\bar{v}(t) - \mathbf{d}(t) && \text{f.a.a. } t \in \Omega,\\
0 &\geq \bar{\lambda}(t) && \text{f.a.a. } t \in \Omega,\\
0 &= \bar{\lambda}(t) \cdot \big(\mathbf{D}_u\bar{u}(t) + \mathbf{D}_v\bar{v}(t) - \mathbf{d}(t)\big) && \text{f.a.a. } t \in \Omega,
\end{aligned} \tag{5.29d}$$

$$\begin{aligned}
0 &= \mu_i(t) && \text{f.a.a. } t \in I^{+0}(\bar{\mathfrak{x}}, i),\ i \in Q,\\
0 &= \big(\mathbf{D}_v\upsilon(t)\big)_i && \text{f.a.a. } t \in I^{0-}(\bar{\mathfrak{x}}, i),\ i \in Q.
\end{aligned} \tag{5.29e}$$

Furthermore, $\bar{\mathfrak{p}}$ is an S-stationary point of the bilevel programming problem if and only if there exist functions $\phi_x \in AC^{1,2}(\Omega, \mathbb{R}^n)$, $\bar{p}, \phi_y, \phi_p \in AC^{1,2}(\Omega, \mathbb{R}^m)$, $\upsilon \in L^2(\Omega, \mathbb{R}^l)$, and $\bar{\lambda}, \mu \in L^2(\Omega, \mathbb{R}^q)$ which satisfy (5.29) and

$$\begin{aligned}
0 &\geq \mu_i(t) && \text{f.a.a. } t \in I^{00}(\bar{\mathfrak{x}}, i),\ i \in Q,\\
0 &\geq \big(\mathbf{D}_v\upsilon(t)\big)_i && \text{f.a.a. } t \in I^{00}(\bar{\mathfrak{x}}, i),\ i \in Q.
\end{aligned}$$

Therein, we set $\bar{\mathfrak{x}} := (\bar{x}, \bar{y}, \bar{p}, \bar{u}, \bar{v}, \bar{\lambda}, \bar{p}(0))$, and the measurable sets $I^{+0}(\bar{\mathfrak{x}}, i)$, $I^{0-}(\bar{\mathfrak{x}}, i)$, as well as $I^{00}(\bar{\mathfrak{x}}, i)$ are defined in (5.27) for all $i \in Q$.

Remark 5.14. Using the terminology of optimal control, the dynamical system (5.29a) is called the adjoint system, its boundary conditions (5.29b) are called the transversality conditions, and one refers to the algebraic equations (5.29c) as Pontryagin's (linearized) Maximum Principle, see [102].

5.2.3. A constraint qualification implying S-stationarity of local optimal solutions

We want to close our considerations addressing (BOC) by constructing a constraint qualification which implies that the local optimal solutions of this problem are always S-stationary points. Therefore, we exploit the notation introduced in the previous sections.
Let $\bar{\mathfrak{p}} := (\bar{x}, \bar{u}, \bar{y}, \bar{v}) \in \mathcal{X} \times \mathcal{Y}$ be a local optimal solution of (BOC) and let $\bar{\mathfrak{x}} := (\bar{x}, \bar{y}, \bar{p}, \bar{u}, \bar{v}, \bar{\lambda}, \bar{p}(0)) \in \tilde{\mathcal{X}}$ denote a feasible point of the KKT reformulation (5.26) associated to $\bar{\mathfrak{p}}$. Due to Theorem 4.18, $\bar{\mathfrak{x}}$ is a local optimal solution of (5.26) as well. By definition $\bar{\mathfrak{p}}$ is an S-stationary point of (BOC) provided $\bar{\mathfrak{x}}$ is an S-stationary point of (5.26). Thus, following Proposition 3.4, the surjectivity of the continuous linear operator

$$\mathtt{H}(\bar{\mathfrak{x}}) := \begin{bmatrix} \tilde{g}'(\bar{\mathfrak{x}}) \\ \tilde{G}'(\bar{\mathfrak{x}}) \\ \tilde{H}'(\bar{\mathfrak{x}}) \end{bmatrix} \in \mathbb{L}[\tilde{\mathcal{X}}, \tilde{\mathcal{Y}} \times \tilde{\mathcal{Z}} \times \tilde{\mathcal{Z}}^\star]$$

implies the S-stationarity of $\bar{\mathfrak{p}}$ for (BOC). Obviously, $\mathtt{H} = \mathtt{H}(\bar{\mathfrak{x}})$ is independent of $\bar{\mathfrak{x}}$. Noting that $\tilde{H}'(\bar{\mathfrak{x}})$ equals the projection onto the λ-component and setting

$$\underline{\mathcal{X}} := AC^{1,2}(\Omega,\mathbb{R}^n) \times AC^{1,2}(\Omega,\mathbb{R}^m) \times AC^{1,2}(\Omega,\mathbb{R}^m) \times L^2(\Omega,\mathbb{R}^k) \times L^2(\Omega,\mathbb{R}^l) \times \mathbb{R}^m,$$
$$\underline{\mathcal{Y}} := AC^{1,2}(\Omega,\mathbb{R}^n) \times AC^{1,2}(\Omega,\mathbb{R}^m) \times AC^{1,2}(\Omega,\mathbb{R}^m) \times L^2(\Omega,\mathbb{R}^l) \times L^2(\Omega,\mathbb{R}^q) \times \mathbb{R}^m,$$

the surjectivity of $\mathtt{H}$ is equivalent to the surjectivity of the operator $\tilde{\mathtt{H}} \in \mathbb{L}[\underline{\mathcal{X}},\underline{\mathcal{Y}}]$ defined below for any $\mathfrak{h} = (h_x, h_y, h_p, h_u, h_v, h_\xi) \in \underline{\mathcal{X}}$:

$$\tilde{\mathtt{H}}[\mathfrak{h}] := \left(\begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} - \hat{\mathbf{K}} h_\xi - \int_0^\cdot \left[\hat{\mathbf{M}} \begin{pmatrix} h_x(\tau) \\ h_y(\tau) \\ h_p(\tau) \end{pmatrix} + \widetilde{\mathbf{N}} \begin{pmatrix} h_u(\tau) \\ h_v(\tau) \end{pmatrix} \right] \mathrm{d}\tau, \right.$$
$$\left. \widetilde{\mathbf{P}} \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{Q}} \begin{pmatrix} h_u(\cdot) \\ h_v(\cdot) \end{pmatrix}, \hat{\mathbf{R}} \begin{pmatrix} h_x(T) \\ h_y(T) \\ h_p(T) \end{pmatrix} \right).$$

Therein, the matrices $\widetilde{\mathbf{N}} \in \mathbb{R}^{(n+2m)\times(k+l)}$, $\widetilde{\mathbf{P}} \in \mathbb{R}^{(l+q)\times(n+2m)}$, and $\widetilde{\mathbf{Q}} \in \mathbb{R}^{(l+q)\times(k+l)}$ are given by

$$\widetilde{\mathbf{N}} := \begin{bmatrix} \mathbf{C}_u & \mathbf{O} \\ \mathbf{A}_u & \mathbf{B}_v \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \qquad \widetilde{\mathbf{P}} := \begin{bmatrix} \mathbf{O} & \mathbf{O} & -\mathbf{B}_v^\top \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{bmatrix}, \qquad \widetilde{\mathbf{Q}} := \begin{bmatrix} \mathbf{O} & \mathbf{R}_v \\ \mathbf{D}_u & \mathbf{D}_v \end{bmatrix}.$$

Now, choose an arbitrary vector $\mathfrak{y} = (w_x, w_y, w_p, \upsilon, z, s) \in \underline{\mathcal{Y}}$ and consider the linear equation $\tilde{\mathtt{H}}[\mathfrak{h}] = \mathfrak{y}$. Assume that $\widetilde{\mathbf{Q}}$ possesses full row rank $l+q$ while $q < k$ is satisfied. Let $\mathbf{Y} \in \mathbb{R}^{(k+l)\times(k-q)}$ be a matrix whose columns form a basis of the null space of $\widetilde{\mathbf{Q}}$, i.e. a matrix with full column rank $k-q$ which satisfies $\widetilde{\mathbf{Q}}\mathbf{Y} = \mathbf{O}$. Considering

$$\widetilde{\mathbf{P}} \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{Q}} \begin{pmatrix} h_u(\cdot) \\ h_v(\cdot) \end{pmatrix} = \begin{pmatrix} \upsilon(\cdot) \\ z(\cdot) \end{pmatrix},$$

an explicit solution is given by

$$\begin{pmatrix} h_u(\cdot) \\ h_v(\cdot) \end{pmatrix} = \mathbf{Y}\vartheta(\cdot) + \widetilde{\mathbf{Q}}^\dagger \left(\begin{pmatrix} \upsilon(\cdot) \\ z(\cdot) \end{pmatrix} - \widetilde{\mathbf{P}} \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} \right)$$

for any function $\vartheta \in L^2(\Omega,\mathbb{R}^{k-q})$. Putting this solution into the variational part of the definition of $\tilde{\mathtt{H}}$, we only need to find a solution of the ODE

$$\begin{pmatrix} \nabla h_x(\cdot) \\ \nabla h_y(\cdot) \\ \nabla h_p(\cdot) \end{pmatrix} - \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}} \right) \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} - \widetilde{\mathbf{N}}\mathbf{Y}\vartheta(\cdot) = \begin{pmatrix} \nabla w_x(\cdot) \\ \nabla w_y(\cdot) \\ \nabla w_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger \begin{pmatrix} \upsilon(\cdot) \\ z(\cdot) \end{pmatrix} \tag{5.30}$$

which satisfies the boundary conditions

$$\begin{pmatrix} h_x(0) \\ h_y(0) \\ h_p(0) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ h_\xi \end{pmatrix} = \begin{pmatrix} w_x(0) \\ w_y(0) \\ w_p(0) \end{pmatrix}, \quad \hat{\mathbf{R}} \begin{pmatrix} h_x(T) \\ h_y(T) \\ h_p(T) \end{pmatrix} = s \tag{5.31}$$

(due to the presence of the variable h_ξ, the initial value of h_p is actually free). First, we find a solution $(\tilde{h}_x, \tilde{h}_y, \tilde{h}_p)$ of the linear Volterra equation of the second kind

$$\begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} - \int_0^\cdot \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}} \right) \begin{pmatrix} h_x(\tau) \\ h_y(\tau) \\ h_p(\tau) \end{pmatrix} \mathrm{d}\tau = \begin{pmatrix} w_x(\cdot) \\ w_y(\cdot) \\ w_p(\cdot) \end{pmatrix} + \int_0^\cdot \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger \begin{pmatrix} \upsilon(\tau) \\ z(\tau) \end{pmatrix} \mathrm{d}\tau,$$

see the proof of Lemma 5.12 and the reference therein. Next, we consider the homogeneous system

$$\begin{pmatrix} \nabla h_x(\cdot) \\ \nabla h_y(\cdot) \\ \nabla h_p(\cdot) \end{pmatrix} = \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}} \right) \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{N}}\mathbf{Y}\vartheta(\cdot)$$

equipped with the boundary conditions

$$\begin{pmatrix} h_x(0) \\ h_y(0) \\ h_p(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ h_\xi \end{pmatrix}, \quad \hat{\mathbf{R}} \begin{pmatrix} h_x(T) \\ h_y(T) \\ h_p(T) \end{pmatrix} = s - \hat{\mathbf{R}} \begin{pmatrix} \tilde{h}_x(T) \\ \tilde{h}_y(T) \\ \tilde{h}_p(T) \end{pmatrix}.$$

Assuming that it possesses a solution $(h'_x, h'_y, h'_p, \vartheta', h'_\xi)$, we find that $(\tilde{h}_x + h'_x, \tilde{h}_y + h'_y, \tilde{h}_p + h'_p, \vartheta', h'_\xi)$ provides a solution of (5.30), (5.31). This, however, means that $\tilde{\mathtt{H}}$ is surjective. Thus, we need to demand that the homogeneous ODE system

$$\begin{pmatrix} \nabla h_x(\cdot) \\ \nabla h_y(\cdot) \\ \nabla h_p(\cdot) \end{pmatrix} = \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}}\right) \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{N}}\mathbf{Y}\vartheta(\cdot), \quad \begin{pmatrix} h_x(0) \\ h_y(0) \\ h_p(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ h_\xi \end{pmatrix}, \quad \hat{\mathbf{R}} \begin{pmatrix} h_x(T) \\ h_y(T) \\ h_p(T) \end{pmatrix} = r \tag{5.32}$$

possesses a solution for any $r \in \mathbb{R}^m$ in order to obtain the surjectivity of $\tilde{\mathtt{H}}$ and, consequently, of $\mathtt{H}$.
Note that the matrix $\hat{\mathbf{R}}$ possesses full row rank by definition while h_ξ is a variable. Thus, we have the surjectivity of $\mathtt{H}$ if

$$\begin{pmatrix} \nabla h_x(\cdot) \\ \nabla h_y(\cdot) \\ \nabla h_p(\cdot) \end{pmatrix} = \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}}\right) \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{N}}\mathbf{Y}\vartheta(\cdot), \quad \begin{pmatrix} h_x(0) \\ h_y(0) \\ h_p(0) \end{pmatrix} = 0, \quad \begin{pmatrix} h_x(T) \\ h_y(T) \\ h_p(T) \end{pmatrix} = r'$$

possesses a solution for any $r' \in \mathbb{R}^{n+2m}$. A sufficient condition for this property is the controllability of the linear dynamical system

$$\begin{pmatrix} \nabla h_x(\cdot) \\ \nabla h_y(\cdot) \\ \nabla h_p(\cdot) \end{pmatrix} = \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}}\right) \begin{pmatrix} h_x(\cdot) \\ h_y(\cdot) \\ h_p(\cdot) \end{pmatrix} + \widetilde{\mathbf{N}}\mathbf{Y}\vartheta(\cdot), \tag{5.33}$$

see [9] for a detailed introduction to the theory of linear dynamical systems, their properties, and their behavior. By means of the famous Kalman theorem, see [9, Theorem 4.1], the system (5.33) is controllable if and only if its controllability matrix

$$\left[\widetilde{\mathbf{N}}\mathbf{Y} \quad \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}}\right)\widetilde{\mathbf{N}}\mathbf{Y} \quad \dots \quad \left(\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^\dagger\widetilde{\mathbf{P}}\right)^{n+2m-1}\widetilde{\mathbf{N}}\mathbf{Y}\right] \in \mathbb{R}^{(n+2m)\times(n+2m)(k-q)} \tag{5.34}$$

possesses full row rank $n + 2m$.
Summarizing the above considerations, we obtain our final result of this section.

Theorem 5.15. Let $\tilde{\mathbf{Q}} \in \mathbb{R}^{(l+q)\times(k+l)}$ defined above possess full row rank $l + q$ where $q < k$ is satisfied. Moreover, let $\mathbf{Y} \in \mathbb{R}^{(k+l)\times(k-q)}$ be a matrix whose columns form a basis of the null space of $\tilde{\mathbf{Q}}$.
Suppose that one of the following conditions is valid:

1. For any $r \in \mathbb{R}^m$, the system (5.32) possesses a solution.

2. The controllability matrix (5.34) possesses full row rank $n + 2m$.

Then any local optimal solution $(\bar{x}, \bar{u}, \bar{y}, \bar{v}) \in \mathcal{X} \times \mathcal{Y}$ of (BOC) satisfies the S-stationarity conditions of Proposition 5.13.

We close this section with the following numerical example which illustrates that the controllability of (5.33) is a reasonable assumption.

Example 5.16. Consider the bilevel optimal control problem (BOC) with $n = m = 1$, $k = 2$, $l = 1$, and $q = 1$ as well as

$$\begin{aligned} &\mathbf{A}_x = 0, && \mathbf{A}_u = \begin{pmatrix} 1 & 1 \end{pmatrix}, && \mathbf{B}_y = 1, && \mathbf{B}_v = 1, && \mathbf{C}_x = 0, && \mathbf{C}_u = \begin{pmatrix} 1 & 0 \end{pmatrix}, \\ &\mathbf{D}_u = \begin{pmatrix} 1 & 0 \end{pmatrix}, && \mathbf{D}_v = 1, && \mathbf{P} = 0, && \mathbf{R} = 0, && \mathbf{R}_y = 1, && \mathbf{R}_v = 1. \end{aligned}$$

Then we have

$$\hat{\mathbf{M}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad \widetilde{\mathbf{N}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \widetilde{\mathbf{P}} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \widetilde{\mathbf{Q}} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

which yields

$$\widetilde{\mathbf{Q}}^{\dagger} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

We choose

$$\mathbf{Y} := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

to obtain

$$\hat{\mathbf{M}} - \widetilde{\mathbf{N}}\widetilde{\mathbf{Q}}^{\dagger}\widetilde{\mathbf{P}} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad \widetilde{\mathbf{N}}\mathbf{Y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, the controllability matrix (5.34) of the corresponding linear dynamical system (5.33) equals

$$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Clearly, this matrix possesses full row rank 3. Consequently, all local optimal solutions of any corresponding bilevel optimal control problem are S-stationary by means of Theorem 5.15. ■

5.3. Optimal control problems with an implicit pointwise state constraint

For some bounded domain $\Omega \subseteq \mathbb{R}^d$ with $d \in \{1, 2, 3\}$, we study the abstract optimal control problem

$$\begin{aligned} F_0(x_u(\bar{\omega}), y) + \tfrac{\sigma_x}{2} \left\| x_u - x_{\mathsf{d}} \right\|^2_{L^2(\Omega, \mathbb{R}^n)} + \tfrac{\sigma_u}{2} \left\| u - u_{\mathsf{d}} \right\|^2_{L^2(\Omega, \mathbb{R}^k)} &\to \min_{u,y} \\ u &\in U_{\mathsf{ad}} \\ y &\in \Psi(u) \end{aligned} \tag{OC}$$

where $\Psi \colon L^2(\Omega, \mathbb{R}^k) \rightrightarrows \mathbb{R}^m$ denotes the solution set mapping of the parametric optimization problem

$$\begin{aligned} f(x_u(\bar{\omega}), y) &\to \min_{y} \\ g(x_u(\bar{\omega}), y) &\leq 0. \end{aligned} \tag{5.35}$$

Therein, for any control $u \in L^2(\Omega, \mathbb{R}^k)$, the state $x_u := \mathtt{S}[u] \in \mathcal{F}(\Omega, \mathbb{R}^n)$ denotes the unique (weak) solution of a given linear (ordinary or partial) differential equation whose solution operator is denoted by $\mathtt{S} \in \mathbb{L}[L^2(\Omega, \mathbb{R}^k), \mathcal{F}(\Omega, \mathbb{R}^n)]$ where $\mathcal{F}(\Omega, \mathbb{R}^n)$ is a certain function space of vector-valued functions with n components. Below, we list our standing assumptions on (OC).

Assumption 5.2. We fix $\bar{\omega} \in \Omega$ as well as $n = k = 1$ for $d \in \{2, 3\}$ and $\bar{\omega} := T$ as well as $\Omega = (0, T)$ for $d = 1$. The functions $F_0, f \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $g \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l$ are continuously differentiable, $x_{\mathsf{d}} \in L^2(\Omega, \mathbb{R}^n)$ and $u_{\mathsf{d}} \in L^2(\Omega, \mathbb{R}^k)$ are the fixed desired state and control, respectively, $U_{\mathsf{ad}} \subseteq L^2(\Omega, \mathbb{R}^k)$ is a nonempty, closed, convex set, and $\sigma_x, \sigma_u \geq 0$ are fixed weights. Finally, we assume that $\mathcal{F}(\Omega, \mathbb{R}^n)$ satisfies $\mathcal{F}(\Omega, \mathbb{R}^n) \hookrightarrow C\left(\overline{\Omega}\right)^n$ and that this embedding is compact.

Let us discuss our assumption on $\mathcal{F}(\Omega,\mathbb{R}^n)$: From the compactness of $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow C\left(\overline{\Omega}\right)^n$ and the obvious embedding $C\left(\overline{\Omega}\right) \hookrightarrow L^2(\Omega)$ we already get that $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow L^2(\Omega,\mathbb{R}^n)$ is compact, see [1, Remark 6.4.2]. Especially, the objective function of (OC) is well-defined. Note that the continuity of the state function x_u ensures that the lower level problem (5.35) is meaningful. In the case $d = 1$ where $\Omega = (0,T)$ is some bounded, open interval, we think of $\mathcal{F}(\Omega,\mathbb{R}^n) := AC^{1,2}(\Omega,\mathbb{R}^n)$, see Theorem 2.10. For $d \in \{2,3\}$, a possible choice is given by $\mathcal{F}(\Omega,\mathbb{R}) := H^2(\Omega)$ where Ω possesses a Lipschitz continuous boundary, see Theorem 2.9.
One may think of (OC) as a control constrained optimal control problem where some penalty cost depending on the state function's value at $\bar{\omega}$ are added to the objective functional, and the penality cost is calculated by means of the program (5.35). A typical example for such an optimal control problem of ODEs is given by the so-called natural gas cash-out problem, see [13, 14, 74] and the references therein. For an example in the context of PDE control, one can think of heating a potato in an oven where the potato's core temperature is used to compute a certain quality measure considered in the problem's objective. It is not difficult to generalize the upcoming theory to the situation where the lower level problem depends in a parametric way on the state function's value at finitely many points from Ω as long as $d \geq 2$ holds.
At the beginning of Chapter 4 we mentioned that it can be very difficult to verify the existence of global optimal solutions for bilevel programming problems where the decision spaces are infinite-dimensional. However, due to Assumption 5.2, we have the following existence result for (OC).

Theorem 5.17. In addition to Assumption 5.2, let the set U_{ad} of admissible controls be bounded, assume that the set $\{(\mathtt{x},y) \in \mathbb{R}^n \times \mathbb{R}^m \,|\, g(\mathtt{x},y) \leq 0\}$ is nonempty and compact, and that $\tilde{\varphi}\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by $\tilde{\varphi}(\mathtt{x}) := \inf_y\{f(\mathtt{x},y) \,|\, g(\mathtt{x},y) \leq 0\}$ for any $\mathtt{x} \in \mathbb{R}^n$ is upper semicontinuous. Finally, assume that (OC) possesses at least one feasible point. Then (OC) has a global optimal solution.

Proof. From Section 4.3 we know that (OC) is equivalent to

$$\begin{aligned} F_0(x_u(\bar{\omega}),y) + \tfrac{\sigma_x}{2}\left\|x_u - x_{\mathsf{d}}\right\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\left\|u - u_{\mathsf{d}}\right\|^2_{L^2(\Omega,\mathbb{R}^k)} &\to \min_{u,y} \\ u &\in U_{\mathsf{ad}} \\ f(x_u(\bar{\omega}),y) - \tilde{\varphi}(x_u(\bar{\omega})) &\leq 0 \\ g(x_u(\bar{\omega}),y) &\leq 0. \end{aligned} \tag{5.36}$$

Let $\{(u_k,y_k)\} \subseteq L^2(\Omega,\mathbb{R}^k) \times \mathbb{R}^m$ be a minimizing sequence of (5.36), i.e. a sequence of points feasible for (5.36) with

$$\lim_{k\to\infty}\left(F_0(x_{u_k}(\bar{\omega}),y_k) + \tfrac{\sigma_x}{2}\left\|x_{u_k} - x_{\mathsf{d}}\right\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\left\|u_k - u_{\mathsf{d}}\right\|^2_{L^2(\Omega,\mathbb{R}^k)}\right) = \alpha$$

where $\alpha \in \mathbb{R} \cup \{-\infty\}$ denotes the infimal objective value of (5.36). Due to the assumptions of the theorem, U_{ad} is weakly sequentially compact. Thus, $\{u_k\}$ possesses a weakly convergent subsequence whose weak limit $\bar{u}$ belongs to U_{ad}. W.l.o.g. we assume $u_k \rightharpoonup \bar{u}$. Similary, since $\{y_k\}$ is bounded, it possesses an accumulation point $\bar{y} \in \mathbb{R}^m$ and w.l.o.g. we can assume $y_k \to \bar{y}$.
From $u_k \rightharpoonup \bar{u}$ in $L^2(\Omega,\mathbb{R}^k)$ we deduce $x_{u_k} \rightharpoonup x_{\bar{u}}$ in $\mathcal{F}(\Omega,\mathbb{R}^n)$. Due to the compactness of the embeddings $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow C\left(\overline{\Omega}\right)^n$ and $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow L^2(\Omega,\mathbb{R}^n)$, we obtain $x_{u_k} \to x_{\bar{u}}$ in $C\left(\overline{\Omega}\right)^n$ and $L^2(\Omega,\mathbb{R}^n)$. Especially, $x_{u_k}(\bar{\omega}) \to x_{\bar{u}}(\bar{\omega})$ holds true due to the definition of the norm in $C\left(\overline{\Omega}\right)^n$. That is why the continuity of f and g as well as the upper semicontinuity of $\tilde{\varphi}$ lead to

$$\begin{aligned} f(x_{\bar{u}}(\bar{\omega}),\bar{y}) - \tilde{\varphi}(x_{\bar{u}}(\bar{\omega})) &\leq \liminf_{k\to\infty}\left[f(x_{u_k}(\bar{\omega}),y_k) - \tilde{\varphi}(x_{u_k}(\bar{\omega}))\right] \leq 0, \\ g(x_{\bar{u}}(\bar{\omega}),\bar{y}) &= \lim_{k\to\infty} g(x_{u_k}(\bar{\omega}),y_k) \leq 0, \end{aligned}$$

i.e. $(\bar{u},\bar{y})$ is feasible for (5.36). Exploiting the continuity of F_0, $x_{u_k} \to x_{\bar{u}}$ in $L^2(\Omega,\mathbb{R}^n)$, and the weak lower

semicontinuity of $L^2(\Omega,\mathbb{R}^k) \ni u \mapsto \frac{\sigma_u}{2}\|u-u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)} \in \mathbb{R}$, we derive

$$\begin{aligned}
&F_0(x_{\bar u}(\bar\omega),\bar y) + \tfrac{\sigma_x}{2}\|x_{\bar u}-x_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\|\bar u-u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)}\\
&\quad= \lim_{k\to\infty} F_0(x_{u_k}(\bar\omega),y_k) + \lim_{k\to\infty}\tfrac{\sigma_x}{2}\|x_{u_k}-x_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\|\bar u-u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)}\\
&\quad\le \lim_{k\to\infty} F_0(x_{u_k}(\bar\omega),y_k) + \lim_{k\to\infty}\tfrac{\sigma_x}{2}\|x_{u_k}-x_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)} + \liminf_{k\to\infty}\tfrac{\sigma_u}{2}\|u_k-u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)}\\
&\quad= \liminf_{k\to\infty}\left[F_0(x_{u_k}(\bar\omega),y_k) + \tfrac{\sigma_x}{2}\|x_{u_k}-x_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\|u_k-u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)}\right]\\
&\quad= \alpha.
\end{aligned}$$

The feasibility of $(\bar u,\bar y)$ for (5.36) yields that it is a global optimal solution of this problem and, consequently, the proof is completed. □

Note that the above proof is possible without exploiting the compactness of $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow L^2(\Omega,\mathbb{R}^n)$: one only needs to use the weak lower semicontinuity of $L^2(\Omega,\mathbb{R}^n) \ni v \mapsto \frac{\sigma_x}{2}\|v-x_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)}$ and the fact that for real sequences $\{a_k\}$ and $\{b_k\}$, $\liminf_{k\to\infty} a_k + \liminf_{k\to\infty} b_k \le \liminf_{k\to\infty}[a_k+b_k]$ holds true.
Due to the equivalence of our model problem (OC) and its optimal value reformulation (5.36), it is reasonable to name (OC) an optimal control problem with implicit pointwise state constraint: The constraints of (5.36) which result from the reformulation of the lower level problem restrict the choice of the state function's value at $\bar\omega$. However, due to the presence of the (in general) unknown function $\tilde\varphi$, we need to call this pointwise state constraint *implicit*.
Note that the upper semicontinuity of the function $\tilde\varphi$ demanded in Theorem 5.17 is a reasonable assumption which holds for example if the lower level problem (5.35) is sufficiently regular. More precisely, $\tilde\varphi$ is upper semicontinuous if MFCQ is valid at all lower level feasible points, see [26, proof of Theorem 4.3]. Other criteria for the lower semicontinuity of optimal value functions can be found in [7, Section 4].
The above result justifies the search for conditions which characterize the optimal solutions of (OC). This will be done in the subsequent parts of this chapter. We will use the optimal value reformulation (5.36) for that purpose. However, following the arguments in [15], it seems to be possible to apply the KKT approach as well provided the lower level problem (5.35) possesses certain convexity and regularity properties, see Section 4.2.

5.3.1. The abstract case

In this section, we first want to study the differentiability properties of the mappings appearing in the problem (OC).
Recall that $\mathsf{S} \in \mathbb{L}[L^2(\Omega,\mathbb{R}^k),\mathcal{F}(\Omega,\mathbb{R}^n)]$ denotes the linear solution operator of a given differential equation. Let $\mathtt{E} \in \mathbb{L}[\mathcal{F}(\Omega,\mathbb{R}^n),L^2(\Omega,\mathbb{R}^n)]$ represent the compact embedding $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow L^2(\Omega,\mathbb{R}^n)$ and define $\bar{\mathsf{S}} := \mathtt{E}\circ\mathsf{S} \in \mathbb{L}[L^2(\Omega,\mathbb{R}^k),L^2(\Omega,\mathbb{R}^n)]$. Furthermore, let $\mathtt{E}_{\bar\omega} \in \mathbb{L}[\mathcal{F}(\Omega,\mathbb{R}^n),\mathbb{R}^n]$ be the pointwise evaluation operator definded by $\mathtt{E}_{\bar\omega}[x] := x(\bar\omega)$ for all $x \in \mathcal{F}(\Omega,\mathbb{R}^n)$. Since we have $\mathcal{F}(\Omega,\mathbb{R}^n) \hookrightarrow C\left(\overline{\Omega}\right)^n$, $\mathtt{E}_{\bar\omega}$ is a continuous, linear operator. Finally, we put $\mathsf{S}_{\bar\omega} := \mathtt{E}_{\bar\omega}\circ\mathsf{S} \in \mathbb{L}[L^2(\Omega,\mathbb{R}^k),\mathbb{R}^n]$.
We define mappings $\bar F\colon L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^m \to \mathbb{R}$, $\bar f\colon L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^m \to \mathbb{R}$, $\bar g\colon L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^m \to \mathbb{R}^l$, and $\bar\varphi\colon L^2(\Omega,\mathbb{R}^k) \to \overline{\mathbb{R}}$ formally by

$$\begin{aligned}
\bar F(u,y) &:= F_0(\mathsf{S}_{\bar\omega}[u],y) + \tfrac{\sigma_x}{2}\|\bar{\mathsf{S}}[u]-x_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\|u-u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)},\\
\bar f(u,y) &:= f(\mathsf{S}_{\bar\omega}[u],y),\\
\bar g(u,y) &:= g(\mathsf{S}_{\bar\omega}[u],y),\\
\bar\varphi(u) &:= \tilde\varphi(\mathsf{S}_{\bar\omega}[u])
\end{aligned}$$

for all $u \in L^2(\Omega,\mathbb{R}^k)$ and $y \in \mathbb{R}^m$. Note that the function $\tilde\varphi$ is defined in Theorem 5.17.
We will exploit the following auxiliary result.

Lemma 5.18. Let $\theta\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a given function and define $\Theta\colon L^2(\Omega,\mathbb{R}^k) \to \overline{\mathbb{R}}$ by $\Theta := \theta\circ\mathsf{S}_{\bar\omega}$. Furthermore, fix $\bar u \in L^2(\Omega,\mathbb{R}^k)$.

Suppose that θ is locally Lipschitz continuous at $\mathtt{S}_{\bar{\omega}}[\bar{u}]$. Then Θ is locally Lipschitz continuous at $\bar{u}$. Additionally, if $v \in \partial^c \Theta(\bar{u})$ holds true, then there exists $a \in \partial^c \theta(\mathtt{S}_{\bar{\omega}}[\bar{u}])$ which satisfies

$$\forall d \in L^2(\Omega, \mathbb{R}^k): \quad \langle v, d \rangle_{L^2(\Omega,\mathbb{R}^k)} = a \cdot \mathtt{S}_{\bar{\omega}}[d] = \int_\Omega a \cdot \mathtt{S}[d](\omega) \mathrm{d}\delta_{\bar{\omega}}(\omega)$$

where $\delta_{\bar{\omega}}$ denotes the Dirac measure of the singleton $\{\bar{\omega}\}$.
If θ is continuously differentiable at $\mathtt{S}_{\bar{\omega}}[\bar{u}]$, then Θ is continuously Fréchet differentiable at $\bar{u}$ and we have

$$\forall d \in L^2(\Omega, \mathbb{R}^k): \quad \Theta'(\bar{u})[d] = \nabla\theta(\mathtt{S}_{\bar{\omega}}[\bar{u}]) \cdot \mathtt{S}_{\bar{\omega}}[d] = \int_\Omega \nabla\theta(\mathtt{S}_{\bar{\omega}}[\bar{u}]) \cdot \mathtt{S}[d](\omega) \mathrm{d}\delta_{\bar{\omega}}(\omega).$$

Proof. Let θ be locally Lipschitz continuous at $\mathtt{S}_{\bar{\omega}}[\bar{u}]$. Since $\mathtt{S}_{\bar{\omega}}$ is a continuous, linear operator, it is Lipschitz continuous. Consequently, Θ is the composition of the mappings θ and $\mathtt{S}_{\bar{\omega}}$ which are locally Lipschitz continuous at $\mathtt{S}_{\bar{\omega}}[\bar{u}]$ and $\bar{u}$, respectively. Thus, Θ is locally Lipschitz continuous at $\bar{u}$. We apply Clarke's chain rule, see [24, Theorem 2.3.10], in order to obtain $\partial^c \Theta(\bar{u}) \subseteq \{a \circ \mathtt{S}_{\bar{\omega}} \,|\, a \in \partial^c \theta(\mathtt{S}_{\bar{\omega}}[\bar{u}])\}$ which shows the first claim.
In the case where θ is continuously differentiable at $\mathtt{S}_{\bar{\omega}}[\bar{u}]$, Θ is continuously Fréchet differentiable at $\bar{u}$ due to the chain rule for Fréchet differentiable mappings (clearly, $\mathtt{S}_{\bar{\omega}}$ is continuously Fréchet differentiable since it is a continuous, linear mapping), see [118, Satz 2.20]. The latter result yields $\Theta'(\bar{u}) = \theta'(\mathtt{S}_{\bar{\omega}}[\bar{u}]) \circ \mathtt{S}_{\bar{\omega}}$ which shows the claim. □

Using the above lemma and Example 2.27, we obtain the following corollaries.

Corollary 5.19. The mapping $\bar{F}$ is continuously Fréchet differentiable. For $(\bar{u}, \bar{y}) \in L^2(\Omega, \mathbb{R}^k) \times \mathbb{R}^m$, we have

$$\begin{aligned}\forall d \in L^2(\Omega, \mathbb{R}^k): \quad \bar{F}'_u(\bar{u}, \bar{y})[d] = &\int_\Omega \nabla_x F_0(\mathtt{S}_{\bar{\omega}}[\bar{u}], \bar{y}) \cdot \mathtt{S}[d](\omega) \mathrm{d}\delta_{\bar{\omega}}(\omega) + \sigma_x \int_\Omega \bar{\mathtt{S}}^\star \big[\bar{\mathtt{S}}[\bar{u}] - x_{\mathtt{d}}\big](\omega) \cdot d(\omega) \mathrm{d}\omega \\ &+ \sigma_u \int_\Omega \big[\bar{u}(\omega) - u_{\mathtt{d}}(\omega)\big] \cdot d(\omega) \mathrm{d}\omega\end{aligned}$$

and $\bar{F}'_y(\bar{u}, \bar{y}) = \nabla_y F_0(\mathtt{S}_{\bar{\omega}}[\bar{u}], \bar{y})$.

Corollary 5.20. The mappings $\bar{f}$ and $\bar{g}$ are continuously Fréchet differentiable. We have

$$\begin{aligned}\forall d \in L^2(\Omega, \mathbb{R}^k): \quad \bar{f}'_u(\bar{u}, \bar{y})[d] &= \int_\Omega \nabla_x f(\mathtt{S}_{\bar{\omega}}[\bar{u}], \bar{y}) \cdot \mathtt{S}[d](\omega) \mathrm{d}\delta_{\bar{\omega}}(\omega), \\ \bar{g}'_u(\bar{u}, \bar{y})[d] &= \int_\Omega \nabla_x g(\mathtt{S}_{\bar{\omega}}[\bar{u}], \bar{y})\, \mathtt{S}[d](\omega) \mathrm{d}\delta_{\bar{\omega}}(\omega),\end{aligned}$$

$\bar{f}'_y(\bar{u}, \bar{y}) = \nabla_y f(\mathtt{S}_{\bar{\omega}}[\bar{u}], \bar{y})$, and $\bar{g}'_y(\bar{u}, \bar{y}) = \nabla_y g(\mathtt{S}_{\bar{\omega}}[\bar{u}], \bar{y})$ for every $(\bar{u}, \bar{y}) \in L^2(\Omega, \mathbb{R}^k) \times \mathbb{R}^m$.

Now, we can start to derive necessary optimality conditions for (OC) using the penalization approach from Section 4.3. For that purpose, we exploit the following lemma which provides a criterion for the partial calmness property to hold at the local optimal solutions of (5.36).

Lemma 5.21. Suppose that the parametric optimization problem

$$\begin{aligned} f(\mathtt{x}, y) &\to \min_y \\ g(\mathtt{x}, y) &\leq 0 \end{aligned} \tag{5.37}$$

where the parameter $\mathtt{x}$ comes from $\mathbb{R}^n$ possesses a uniformly weak sharp minimum. Then (5.36) is partially calm at its local optimal solutions.

Proof. First, observe that the presence of a uniformly weak sharp minimum for (5.37) implies that

$$\begin{aligned} \bar{f}(u,y) &\to \min_y \\ \bar{g}(u,y) &\leq 0 \end{aligned}$$

possesses a uniformly weak sharp minimum as well. Noting that the function $\bar{F}$ is continuously Fréchet differentiable, see Corollary 5.19, the assertion follows from Proposition 4.35. □

Let us denote by $\tilde{\Psi}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the solution set mapping of the parametric optimization problem (5.37). Then we obtain the following abstract necessary optimality conditions from Theorem 4.37.

Theorem 5.22. Let $(\bar{u},\bar{y}) \in L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^m$ be a local optimal solution of (OC) and set $\bar{x} := \bar{\mathsf{S}}[\bar{u}]$. Assume that $\tilde{\Psi}$ is inner semicontinuous at $(\bar{x}(\bar{\omega}),\bar{y})$ and that the constraint qualification

$$\forall \lambda \in \mathbb{R}^l\colon \quad 0 = \nabla_y g(\bar{x}(\bar{\omega}),\bar{y})^\top \lambda,\ \lambda \geq 0,\ 0 = \lambda \cdot g(\bar{x}(\bar{\omega}),\bar{y}) \implies \lambda = 0$$

is valid. Finally, suppose that (5.37) possesses a uniformly weak sharp minimum. Then there are multipliers $p \in L^2(\Omega,\mathbb{R}^k)$, $\xi \in L^2(\Omega,\mathbb{R}^k)$, $\kappa > 0$, and $\lambda,\bar{\lambda} \in \mathbb{R}^l$ which satisfy the following conditions:

$$\begin{aligned} 0 = &\int_\Omega \big[\nabla_x F_0(\bar{x}(\bar{\omega}),\bar{y}) + \nabla_x g(\bar{x}(\bar{\omega}),\bar{y})^\top[\lambda - \kappa\bar{\lambda}]\big]\cdot \mathsf{S}[d](\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) \\ &+ \int_\Omega \big[\sigma_x p(\omega) + \sigma_u(\bar{u}(\omega) - u_{\mathsf{d}}(\omega)) + \xi(\omega)\big]\cdot d(\omega)\mathrm{d}\omega \quad \text{for all } d \in L^2(\Omega,\mathbb{R}^k), && \text{(5.38a)} \\ 0 = &\bar{\mathsf{S}}^\star[\bar{x} - x_{\mathsf{d}}] - p, && \text{(5.38b)} \\ 0 = &\nabla_y F_0(\bar{x}(\bar{\omega}),\bar{y}) + \kappa\nabla_y f(\bar{x}(\bar{\omega}),\bar{y}) + \nabla_y g(\bar{x}(\bar{\omega}),\bar{y})^\top\lambda, && \text{(5.38c)} \\ 0 = &\nabla_y f(\bar{x}(\bar{\omega}),\bar{y}) + \nabla_y g(\bar{x}(\bar{\omega}),\bar{y})^\top\bar{\lambda}, && \text{(5.38d)} \\ &\xi \in \mathcal{N}_{U_{\mathsf{ad}}}(\bar{u}), && \text{(5.38e)} \\ &\lambda \geq 0,\ 0 = \lambda\cdot g(\bar{x}(\bar{\omega}),\bar{y}), && \text{(5.38f)} \\ &\bar{\lambda} \geq 0,\ 0 = \bar{\lambda}\cdot g(\bar{x}(\bar{\omega}),\bar{y}). && \text{(5.38g)} \end{aligned}$$

Proof. The proof parallels the validation of Theorem 4.37. First, we apply Lemma 5.21 and Proposition 4.33 in order to find $\kappa > 0$ such that $(\bar{u},\bar{y})$ solves

$$\begin{aligned} \bar{F}(u,y) + \kappa\big(\bar{f}(u,y) - \bar{\varphi}(u)\big) &\to \min_{u,y} \\ u &\in U_{\mathsf{ad}} \\ \bar{g}(u,y) &\leq 0 \end{aligned}$$

locally. Invoking Lemma 4.36 and keeping Remark 2.33 in mind, the function $\tilde{\varphi}$ is locally Lipschitz continuous at $\bar{x}(\bar{\omega})$ and satisfies

$$\partial^c\tilde{\varphi}(\bar{x}(\bar{\omega})) \subseteq \{\nabla_x f(\bar{x}(\bar{\omega}),\bar{y}) + \nabla_x g(\bar{x}(\bar{\omega}),\bar{y})^\top\bar{\lambda} \,|\, \bar{\lambda} \text{ satisfies (5.38d) and (5.38g)}\}.$$

Recalling $\bar{\varphi} = \tilde{\varphi}\circ\mathsf{S}_{\bar{\omega}}$, $\bar{\varphi}$ is locally Lipschitz continuous at $\bar{u}$ and a formula for the corresponding Clarke subdifferential can be easily derived from Lemma 5.18. Moreover, all the other appearing mappings are continuously Fréchet differentiable at $(\bar{u},\bar{y})$ due to Corollaries 5.19 and 5.20. Finally, introducing $p \in L^2(\Omega,\mathbb{R}^k)$ via (5.38b) allows us to formulate the whole optimality system (4.55) in the form (5.38). □

Below, we characterize a setting where all assumptions of the above theorem are valid, see [29, Remark 3.2(c)], Example 4.34, and [17, Propositions 2.104, 2.106].

Remark 5.23. Let $(\bar{u},\bar{y}) \in L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^m$ be a local optimal solution of (OC) and set $\bar{x} := \bar{\mathsf{S}}[\bar{u}]$. Assume that there are matrices $\mathbf{c} \in \mathbb{R}^l$, $\mathbf{d} \in \mathbb{R}^m$, $\mathbf{C} \in \mathbb{R}^{l\times n}$, and $\mathbf{D} \in \mathbb{R}^{l\times m}$ such that

$$\forall s \in \mathbb{R}^n\,\forall y \in \mathbb{R}^m\colon \quad f(s,y) := \mathbf{d}\cdot y, \quad g(s,y) := \mathbf{C}s + \mathbf{D}y - \mathbf{c}$$

is valid and assume that there is some $\tilde{y} \in \mathbb{R}^m$ with $g(\bar{x}(\bar{\omega}),\tilde{y}) < 0$. Then all the assumptions of Theorem 5.22 hold and, thus, the corresponding necessary optimality conditions are valid at $(\bar{u},\bar{y})$.

5.3.2. Linear ODE constrained optimal control

We choose $\Omega := (0,T)$ and set $\mathcal{F}(\Omega,\mathbb{R}^n) := AC^{1,2}(\Omega,\mathbb{R}^n)$ as well as $\bar{\omega} := T$. As we already mentioned earlier, this choice for $\mathcal{F}(\Omega,\mathbb{R}^n)$ is reasonable due to Theorem 2.10.
For matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times k}$, let $\mathtt{S} \in \mathbb{L}[L^2(\Omega,\mathbb{R}^k), AC^{1,2}(\Omega,\mathbb{R}^n)]$ be the solution operator of the following linear system of ODEs:

$$\nabla x(t) = \mathbf{A}x(t) + \mathbf{B}u(t) \quad \text{f.a.a. } t \in \Omega, \qquad x(0) = 0. \tag{5.39}$$

We introduce the so-called fundamental matrix function $\Phi\colon \Omega \to \mathbb{R}^{n\times n}$ as the (uniquely determined) solution of the matrix differential equation

$$\nabla\Phi(t) = \mathbf{A}\Phi(t) \quad \text{f.a.a. } t \in \Omega, \qquad \Phi(0) = \mathbf{I}_n.$$

Note that $\Phi(t)$ is regular for all $t \in \Omega$ and, thus, we can define $\Phi^{-1}(t) := \Phi(t)^{-1}$ for all $t \in \Omega$. For detailed information on the matrix function Φ and its properties, we refer the interested reader to [2, Sections 17, 18]. Exploiting $\Phi(t)\Phi^{-1}(t) = \mathbf{I}_n$, we obtain $\nabla\Phi(t)\Phi^{-1}(t) + \Phi(t)\nabla\Phi^{-1}(t) = \mathbf{O}$ from the product rule and, consequently,

$$\nabla\Phi^{-1}(t) = -\Phi^{-1}(t)\nabla\Phi(t)\Phi^{-1}(t) = -\Phi^{-1}(t)\mathbf{A}\Phi(t)\Phi^{-1}(t) = -\Phi^{-1}(t)\mathbf{A} \tag{5.40}$$

for almost all $t \in \Omega$.
Due to [2, equation (18.14)], we have the following explicit representation of the solution operator $\mathtt{S}$:

$$\forall u \in L^2(\Omega,\mathbb{R}^k)\colon \quad \mathtt{S}[u] := \Phi(\cdot)\int_0^{\cdot} \Phi^{-1}(\tau)\mathbf{B}u(\tau)\mathrm{d}\tau. \tag{5.41}$$

Remark 5.24. Note that the assumption $x(0) = 0$ in (5.39) is not restrictive. If $x(0) = x_0$ is demanded for some $x_0 \in \mathbb{R}^n$, then the corresponding affine solution operator $\tilde{\mathtt{S}}\colon L^2(\Omega,\mathbb{R}^k) \to AC^{1,2}(\Omega,\mathbb{R}^n)$ is given by

$$\forall u \in L^2(\Omega,\mathbb{R}^k)\colon \quad \tilde{\mathtt{S}}(u) := \Phi(\cdot)\left(x_0 + \int_0^{\cdot} \Phi^{-1}(\tau)\mathbf{B}u(\tau)\mathrm{d}\tau\right) = \Phi(\cdot)x_0 + \mathtt{S}[u],$$

see [2, Section 18]. We introduce $\tilde{x}_\mathsf{d}(\cdot) := x_\mathsf{d}(\cdot) - \Phi(\cdot)x_0 \in L^2(\Omega,\mathbb{R}^n)$ and consider the new objective function

$$\tilde{F}(u,y) := F_0(\mathtt{S}_T[u] + \Phi(T)x_0, y) + \tfrac{\sigma_x}{2}\,\|\bar{\mathtt{S}}[u] - \tilde{x}_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^n)} + \tfrac{\sigma_u}{2}\,\|u - u_\mathsf{d}\|^2_{L^2(\Omega,\mathbb{R}^k)}$$

defined for all $u \in L^2(\Omega,\mathbb{R}^k)$ and all $y \in \mathbb{R}^m$. Here $\bar{\mathtt{S}}$ and $\mathtt{S}_T$ denote the operators associated to (5.39). Similarly, we can modify the definitions of $\bar{f}$, $\bar{g}$, and $\bar{\varphi}$. Hence, we transferred the possibly nonvanishing initial condition into a vanishing one.

For fixed $a \in \mathbb{R}^n$, we can define an operator $\mathtt{S}_{T,a} \in \mathbb{L}[L^2(\Omega,\mathbb{R}^k),\mathbb{R}]$ by

$$\forall d \in L^2(\Omega,\mathbb{R}^k)\colon \quad \mathtt{S}_{T,a}[d] := a \cdot \mathtt{S}_T[d].$$

For the evaluation of the necessary optimality conditions postulated in Theorem 5.22, we need to find an appropriate representation of $\mathtt{S}_{T,a}$ in the space $L^2(\Omega,\mathbb{R}^k) \cong \mathbb{L}[L^2(\Omega,\mathbb{R}^k),\mathbb{R}] = L^2(\Omega,\mathbb{R}^k)^\star$, and we have to calculate the adjoint operator of $\bar{\mathtt{S}}$. This is summarized in the lemmas below.

Lemma 5.25. For some $a \in \mathbb{R}^n$, let the operator $\mathtt{S}_{T,a}$ be given as defined above. Then we have

$$\forall d \in L^2(\Omega,\mathbb{R}^k)\colon \quad \mathtt{S}_{T,a}[d] = \int_0^T \left(\mathbf{B}^\top\zeta(t)\right)\cdot d(t)\mathrm{d}t$$

where $\zeta \in AC^{1,2}(\Omega,\mathbb{R}^n)$ is the unique solution of the linear ODE

$$\nabla\zeta(t) = -\mathbf{A}^\top\zeta(t) \quad \text{f.a.a. } t \in \Omega, \qquad \zeta(T) = a.$$

Proof. For some $d \in L^2(\Omega, \mathbb{R}^k)$, we use (5.41) to obtain

$$\begin{aligned}\mathtt{S}_{T,a}[d] &= a \cdot \mathtt{S}_T[d] = a \cdot \mathtt{E}_T\big[\mathtt{S}[d]\big] = a \cdot \Phi(T) \int_0^T \Phi^{-1}(\tau)\mathbf{B}d(\tau)\mathrm{d}\tau \\ &= \int_0^T a \cdot \Phi(T)\Phi^{-1}(\tau)\mathbf{B}d(\tau)\mathrm{d}\tau = \int_0^T \big(\mathbf{B}^\top \Phi^{-1}(\tau)^\top \Phi(T)^\top a\big) \cdot d(\tau)\mathrm{d}\tau.\end{aligned}$$

We set $\zeta(t) := \Phi^{-1}(t)^\top \Phi(T)^\top a$ for all $t \in \Omega$. Clearly, we have $\zeta(T) = a$, and

$$\nabla\zeta(t) = (\nabla\Phi^{-1}(t))^\top \Phi(T)^\top a = -\mathbf{A}^\top \Phi^{-1}(t)^\top \Phi(T)^\top a = -\mathbf{A}^\top \zeta(t)$$

follows from (5.40) for all $t \in \Omega$. This completes the proof. □

Lemma 5.26. We have

$$\forall v \in L^2(\Omega, \mathbb{R}^n): \quad \bar{\mathtt{S}}^\star[v] = \mathbf{B}^\top \psi$$

where $\psi \in AC^{1,2}(\Omega, \mathbb{R}^n)$ is the unique solution of the linear ODE

$$\nabla\psi(t) = -\mathbf{A}^\top \psi(t) - v(t) \quad \text{f.a.a. } t \in \Omega, \qquad \psi(T) = 0.$$

Proof. For $u \in L^2(\Omega, \mathbb{R}^k)$ and $v \in L^2(\Omega, \mathbb{R}^n)$, we use integration by parts and (5.41) in order to obtain

$$\begin{aligned}\big\langle \bar{\mathtt{S}}^\star[v], u\big\rangle_{L^2(\Omega,\mathbb{R}^k)} &= \big\langle v, \bar{\mathtt{S}}[u]\big\rangle_{L^2(\Omega,\mathbb{R}^n)} = \int_0^T v(t) \cdot \left(\Phi(t) \int_0^t \Phi^{-1}(\tau)\mathbf{B}u(\tau)\mathrm{d}\tau\right)\mathrm{d}t \\ &= \left(\int_0^T \Phi(t)^\top v(t)\mathrm{d}t\right) \cdot \left(\int_0^T \Phi^{-1}(\tau)\mathbf{B}u(\tau)\mathrm{d}\tau\right) \\ &\qquad - \int_0^T \left(\int_0^t \Phi(\tau)^\top v(\tau)\mathrm{d}\tau\right) \cdot \big(\Phi^{-1}(t)\mathbf{B}u(t)\big)\mathrm{d}t \\ &= \int_0^T \left(\int_t^T \Phi(\tau)^\top v(\tau)\mathrm{d}\tau\right) \cdot \big(\Phi^{-1}(t)\mathbf{B}u(t)\big)\mathrm{d}t \\ &= \int_0^T \left[\mathbf{B}^\top \Phi^{-1}(t)^\top \left(\int_t^T \Phi(\tau)^\top v(\tau)\mathrm{d}\tau\right)\right] \cdot u(t)\mathrm{d}t.\end{aligned}$$

For all $t \in \Omega$, we set

$$\psi(t) := \Phi^{-1}(t)^\top \left(\int_t^T \Phi(\tau)^\top v(\tau)\mathrm{d}\tau\right).$$

Obviously, $\psi(T) = 0$ holds by definition. Furthermore, we exploit (5.40) and the product rule of differentiation to see

$$\begin{aligned}\nabla\psi(t) &= (\nabla\Phi^{-1}(t))^\top \left(\int_t^T \Phi(\tau)^\top v(\tau)\mathrm{d}\tau\right) + \Phi^{-1}(t)^\top \big(-\Phi(t)^\top v(t)\big) \\ &= -\mathbf{A}^\top \Phi^{-1}(t)^\top \left(\int_t^T \Phi(\tau)^\top v(\tau)\mathrm{d}\tau\right) - v(t) = -\mathbf{A}^\top \psi(t) - v(t)\end{aligned}$$

for any $t \in \Omega$. This shows the claim. □

Combining the two above lemmas with Theorem 5.22, we arrive at the following explicit necessary optimality conditions.

Theorem 5.27. Let $(\bar{u}, \bar{y}) \in L^2(\Omega, \mathbb{R}^k) \times \mathbb{R}^m$ be a local optimal solution of (OC) where $\mathtt{S}$ is the solution operator of the linear ODE (5.39) and set $\bar{x} := \bar{\mathtt{S}}[\bar{u}]$. Furthermore, let all the assumptions of Theorem

5.22 be satisfied. Then there exist functions $q \in AC^{1,2}(\Omega, \mathbb{R}^n)$ and $\xi \in L^2(\Omega, \mathbb{R}^k)$, a scalar $\kappa > 0$, and vectors $\lambda, \bar{\lambda} \in \mathbb{R}^l$ which satisfy (5.38c) - (5.38g) for $\bar{\omega} := T$ and

$$0 = \nabla q(t) + \mathbf{A}^\top q(t) + \sigma_x(\bar{x}(t) - x_\mathsf{d}(t)) \qquad \text{f.a.a. } t \in \Omega, \tag{5.42a}$$

$$0 = q(T) - \nabla_x F_0(\bar{x}(T), \bar{y}) - \nabla_x g(\bar{x}(T), \bar{y})^\top [\lambda - \kappa\bar{\lambda}], \tag{5.42b}$$

$$0 = \mathbf{B}^\top q(t) + \sigma_u(\bar{u}(t) - u_\mathsf{d}(t)) + \xi(t) \qquad \text{f.a.a. } t \in \Omega. \tag{5.42c}$$

Proof. Let $\psi \in AC^{1,2}(\Omega, \mathbb{R}^n)$ be the unique solution of the linear ODE

$$\nabla\psi(t) = -\mathbf{A}^\top \psi(t) - (\bar{x}(t) - x_\mathsf{d}(t)) \quad \text{f.a.a. } t \in \Omega, \qquad \psi(T) = 0.$$

Then we have $\bar{\mathsf{S}}^\star[\bar{x} - x_\mathsf{d}] = \mathbf{B}^\top\psi$ from Lemma 5.26. Applying Theorem 5.22, we find $\xi \in L^2(\Omega, \mathbb{R}^k)$, $\kappa > 0$, and vectors $\lambda, \bar{\lambda} \in \mathbb{R}^l$ which satisfy (5.38c) - (5.38g) for $\bar{\omega} := T$ and

$$\begin{aligned} 0 = &\int_0^T \left[\nabla_x F_0(\bar{x}(T), \bar{y}) + \nabla_x g(\bar{x}(T), \bar{y})^\top [\lambda - \kappa\bar{\lambda}]\right] \cdot \mathsf{S}[d](t)\mathrm{d}\delta_T(t) \\ &+ \int_0^T \left[\sigma_x \mathbf{B}^\top \psi(t) + \sigma_u(\bar{u}(t) - u_\mathsf{d}(t)) + \xi(t)\right] \cdot d(t)\mathrm{d}t \qquad \text{for all } d \in L^2(\Omega, \mathbb{R}^k). \end{aligned}$$

We set $\bar{a} := \nabla_x F_0(\bar{x}(T), \bar{y}) + \nabla_x g(\bar{x}(T), \bar{y})^\top [\lambda - \kappa\bar{\lambda}]$ in order to see that the latter is equivalent to

$$0 = \mathsf{S}_{T,\bar{a}}[d] + \int_0^T \left[\sigma_x \mathbf{B}^\top \psi(t) + \sigma_u(\bar{u}(t) - u_\mathsf{d}(t)) + \xi(t)\right] \cdot d(t)\mathrm{d}t \qquad \text{for all } d \in L^2(\Omega, \mathbb{R}^k).$$

Defining $\zeta \in AC^{1,2}(\Omega, \mathbb{R}^n)$ to be the unique solution of the linear ODE

$$\nabla\zeta(t) = -\mathbf{A}^\top \zeta(t) \quad \text{f.a.a. } t \in \Omega, \qquad \zeta(T) = \bar{a},$$

we can transfer the above equation to

$$0 = \int_0^T \left[\mathbf{B}^\top(\zeta(t) + \sigma_x\psi(t)) + \sigma_u(\bar{u}(t) - u_\mathsf{d}(t)) + \xi(t)\right] \cdot d(t)\mathrm{d}t \qquad \text{for all } d \in L^2(\Omega, \mathbb{R}^k)$$

which is equivalent to

$$0 = \mathbf{B}^\top(\zeta(t) + \sigma_x\psi(t)) + \sigma_u(\bar{u}(t) - u_\mathsf{d}(t)) + \xi(t) \quad \text{f.a.a. } t \in \Omega,$$

see Lemma 5.25. Now, we only need to define $q \in AC^{1,2}(\Omega, \mathbb{R}^n)$ by means of $q(t) := \zeta(t) + \sigma_x\psi(t)$ for all $t \in \Omega$. A simple calculation reveals that this function satisfies (5.42a) and (5.42b). On the other hand, the above considerations suggest that (5.42c) is valid as well. This completes the proof. □

Remark 5.28. The necessary optimality conditions provided by Theorem 5.27 comprise the classical elements of optimality criteria known from optimal control: the adjoint equation (5.42a) which characterizes the adjoint state q, transversality conditions (5.42b), and Pontryagin's (linearized) Maximum Principle (5.42c). However, there also appear other types of conditions which reflect the bilevel structure of our model problem (OC).
Optimality conditions of related type can be found in [13, 14, 74] where the authors consider optimal control problems of ODEs with implicit pointwise state constraints in different settings.

5.3.3. Optimal control of Poisson's equation

In this section, for a bounded domain $\Omega \subseteq \mathbb{R}^d$, $d \in \{2, 3\}$, with C^2-boundary $\mathrm{bd}\,\Omega$ and $n = k = 1$, we consider Poisson's equation given by

$$\begin{aligned} -\Delta x(\omega) &= \beta(\omega)u(\omega) \qquad \text{f.a.a. } \omega \in \Omega, \\ x(\omega) &= 0 \qquad\qquad\quad \text{f.a.a. } \omega \in \mathrm{bd}\,\Omega. \end{aligned} \tag{5.43}$$

Therein, $\beta \in L^\infty(\Omega)$ is a fixed function which governs where and how much the control function u influences the dynamics. Especially, for some measurable set $\Omega' \subseteq \Omega$, the choice $\beta := \chi_{\Omega'}$ is possible. Note that (5.43) can be seen as a simple model which describes the (stationary) heating process of the domain Ω where u is the source of the heating energy, see [118].
Since the control $u \in L^2(\Omega)$ might be a function possessing jumps, it is reasonable that (5.43) does not need to possess a classical solution in $C^2\left(\overline{\Omega}\right)$. That is why we consider the variational formulation of the dynamics given by

$$\forall \phi \in C_0^\infty(\Omega): \quad -\int_\Omega \Delta x(\omega)\phi(\omega)\mathrm{d}\omega = \int_\Omega \beta(\omega)u(\omega)\phi(\omega)\mathrm{d}\omega.$$

Exploiting integration by parts on the left hand side of this equation and inserting the boundary condition $x|_{\mathrm{bd}\,\Omega} \equiv 0$, we derive the so-called weak formulation of the PDE (5.43):

$$\forall \phi \in C_0^\infty(\Omega): \quad \int_\Omega \nabla x(\omega) \cdot \nabla \phi(\omega)\mathrm{d}\omega = \int_\Omega \beta(\omega)u(\omega)\phi(\omega)\mathrm{d}\omega. \tag{5.44}$$

We have the following classical result.

Proposition 5.29. For any $u \in L^2(\Omega)$, there is a uniquely determined function $x_u \in H^2(\Omega) \cap H_0^1(\Omega)$ which satisfies (5.44) and

$$\|x_u\|_{H^2(\Omega)} \leq \gamma \|u\|_{L^2(\Omega)}.$$

Therein, the constant $\gamma > 0$ does not depend on the choice of u.

Proof. Due to [17, Lemma 6.14], (5.44) possesses a unique solution in $H^2(\Omega) \cap H_0^1(\Omega)$ for any control from $L^2(\Omega)$. The same result yields the existence of a constant $c > 0$ which satisfies

$$\forall u \in L^2(\Omega): \quad \|x_u\|_{H^2(\Omega)} \leq c \|\beta u\|_{L^2(\Omega)}.$$

Since we have $\|\beta u\|_{L^2(\Omega)} \leq \|\beta\|_{L^\infty(\Omega)} \|u\|_{L^2(\Omega)}$, the claim follows choosing $\gamma := c\|\beta\|_{L^\infty(\Omega)}$. □

We define $\mathcal{F}(\Omega, \mathbb{R}) := H^2(\Omega) \cap H_0^1(\Omega)$ and equip this space with the $H^2(\Omega)$-norm. Using Theorems 2.8 and 2.9, we obtain $H^2(\Omega) \cap H_0^1(\Omega) \hookrightarrow C\left(\overline{\Omega}\right)$ and that this embedding is compact due to our choice of the domain's dimension. Finally, we denote by $\mathtt{S} \in \mathbb{L}[L^2(\Omega), H^2(\Omega) \cap H_0^1(\Omega)]$ the solution operator of the weak PDE (5.44). Clearly, $\mathtt{S}$ is well-defined by Proposition 5.29.
Let us consider the optimal control problem (OC) where $\mathtt{S}$ is given as defined above. We first comment on the homogeneous boundary condition in (5.43).

Remark 5.30. Suppose that there is a measurable function $b\colon \mathrm{bd}\,\Omega \to \mathbb{R}$ such that the PDE (5.43) with homogeneous boundary condition is replaced by

$$\begin{aligned} -\Delta x(\omega) &= \beta(\omega)u(\omega) && \text{f.a.a. } \omega \in \Omega, \\ x(\omega) &= b(\omega) && \text{f.a.a. } \omega \in \mathrm{bd}\,\Omega. \end{aligned} \tag{5.45}$$

Clearly, if b is a discontinuous function, then the weak solution of (5.45) (if it exists) cannot be in $C\left(\overline{\Omega}\right)$ and, thus, not in $H^2(\Omega)$. However, under certain regularity assumptions on the function b, we can transfer the corresponding optimal control problem into a problem of type (OC) where $\mathtt{S}$ is the weak solution operator of the homogeneous Poisson equation (5.43) again, see [5, Theorem 6.1.3] for details. Therefore, one only has to introduce an appropriate desired state $\tilde{x}_\mathrm{d}$ which replaces x_d. The procedure is similar to the one described in Remark 5.24 for ODEs.

As we saw earlier, Assumption 5.2 is valid. In order to formulate the necessary optimality conditions from Theorem 5.22 in terms of this paragraph's setting, we need to find an explicit representation of the adjoint operator of $\bar{\mathtt{S}}$ and a reasonable way to deal with the integral $\int_\Omega \mathtt{S}[d](\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega)$ for $d \in L^2(\Omega)$ which appears in (5.38a).
The operator $\bar{\mathtt{S}}^\star$ is characterized in the following well-known lemma, see [118, Lemma 2.24] for a proof.

Lemma 5.31. We have

$$\forall v \in L^2(\Omega)\colon \quad \bar{\mathsf{S}}^\star[v] := \beta\psi$$

where $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ is the unique solution of

$$\forall \phi \in C_0^\infty(\Omega)\colon \quad \int_\Omega \nabla\psi(\omega) \cdot \nabla\phi(\omega)\mathrm{d}\omega = \int_\Omega v(\omega)\phi(\omega)\mathrm{d}\omega.$$

The latter is the weak formulation of the PDE

$$\begin{aligned} -\Delta\psi(\omega) &= v(\omega) && \text{f.a.a. } \omega \in \Omega, \\ \psi(\omega) &= 0 && \text{f.a.a. } \omega \in \mathrm{bd}\,\Omega. \end{aligned}$$

Next, we characterize the integral mentioned earlier.

Lemma 5.32. We have

$$\forall d \in L^2(\Omega)\colon \quad \int_\Omega \mathsf{S}[d](\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) = \int_\Omega \beta(\omega)\zeta(\omega)d(\omega)\mathrm{d}\omega$$

where $\zeta \in W_0^{1,1}(\Omega)$ is the unique solution of

$$\forall \phi \in C_0^\infty(\Omega)\colon \quad -\int_\Omega \zeta(\omega)\Delta\phi(\omega)\mathrm{d}\omega = \int_\Omega \phi(\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) = \phi(\bar{\omega}). \tag{5.46}$$

Proof. First, by means of [17, Lemma 6.38] there is a uniquely determined function $\zeta \in W_0^{1,1}(\Omega)$ satisfying (5.46). Fix some $d \in L^2(\Omega)$ and set $x^d := \mathsf{S}[d]$. Then we have

$$\forall \phi \in C_0^\infty(\Omega)\colon \quad \int_\Omega \nabla x^d(\omega) \cdot \nabla\phi(\omega)\mathrm{d}\omega = \int_\Omega \beta(\omega)d(\omega)\phi(\omega)\mathrm{d}\omega.$$

Recalling that $x^d \in H^2(\Omega)$ holds due to Proposition 5.29, we find

$$\forall \phi \in C_0^\infty(\Omega)\colon \quad -\int_\Omega \Delta x^d(\omega)\phi(\omega)\mathrm{d}\omega = \int_\Omega \beta(\omega)d(\omega)\phi(\omega)\mathrm{d}\omega$$

from the definition of the weak derivative. By definition, $C_0^\infty(\Omega)$ is dense in $W_0^{1,1}(\Omega)$. This yields

$$-\int_\Omega \Delta x^d(\omega)\zeta(\omega)\mathrm{d}\omega = \int_\Omega \beta(\omega)d(\omega)\zeta(\omega)\mathrm{d}\omega.$$

On the other hand, since we have $x^d \in H^2(\Omega) \cap H_0^1(\Omega)$ and $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$,

$$\int_\Omega x^d(\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) = -\int_\Omega \Delta x^d(\omega)\zeta(\omega)\mathrm{d}\omega$$

follows from (5.46). Taking these observations together, we have

$$\int_\Omega \mathsf{S}[d](\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) = \int_\Omega x^d(\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) = -\int_\Omega \Delta x^d(\omega)\zeta(\omega)\mathrm{d}\omega = \int_\Omega \beta(\omega)d(\omega)\zeta(\omega)\mathrm{d}\omega$$

which completes the proof. □

Combining Lemmas 5.31 and 5.32 with Theorem 5.22, we derive the following explicit optimality conditions for (OC) in terms of Poisson's equation.

Theorem 5.33. Let $(\bar{u}, \bar{y}) \in L^2(\Omega) \times \mathbb{R}^m$ be a local optimal solution of (OC) where S is the solution operator of the weak PDE (5.44) and set $\bar{x} := \bar{\mathsf{S}}[\bar{u}]$. Furthermore, let all the assumptions of Theorem 5.22 be valid. Then there exist functions $q \in W_0^{1,1}(\Omega)$ as well as $\xi \in L^2(\Omega)$, a scalar $\kappa > 0$, and vectors $\lambda, \bar{\lambda} \in \mathbb{R}^l$ which satisfy (5.38c) - (5.38g) and

$$\begin{aligned} 0 = {} & \int_\Omega q(\omega)\Delta\phi(\omega)\mathrm{d}\omega + \sigma_x \int_\Omega \big(\bar{x}(\omega) - x_\mathsf{d}(\omega)\big)\phi(\omega)\mathrm{d}\omega \\ & + \big(\nabla_x F_0(\bar{x}(\bar{\omega}), \bar{y}) + \nabla_x g(\bar{x}(\bar{\omega}), \bar{y})^\top[\lambda - \kappa\bar{\lambda}]\big)\phi(\bar{\omega}) \quad \text{for all } \phi \in C_0^\infty(\Omega), \end{aligned} \tag{5.47a}$$

$$0 = \beta(\omega)q(\omega) + \sigma_u(\bar{u}(\omega) - u_\mathsf{d}(\omega)) + \xi(\omega) \quad \text{f.a.a. } \omega \in \Omega. \tag{5.47b}$$

Proof. Let $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ be the unique solution of

$$\forall \phi \in C_0^\infty(\Omega): \quad \int_\Omega \nabla\psi(\omega) \cdot \nabla\phi(\omega)\mathrm{d}\omega = \int_\Omega \big(\bar{x}(\omega) - x_\mathrm{d}(\omega)\big)\phi(\omega)\mathrm{d}\omega.$$

Due to Lemma 5.31, we have $\bar{\mathsf{S}}^\star[\bar{x} - x_\mathrm{d}] = \beta\psi$. Applying Theorem 5.22, we find $\xi \in L^2(\Omega)$, $\kappa > 0$, and $\lambda, \bar{\lambda} \in \mathbb{R}^l$ which satisfy (5.38c) - (5.38g) and

$$\begin{aligned} 0 = &\int_\Omega \big[\nabla_x F_0(\bar{x}(\bar{\omega}), \bar{y}) + \nabla_x g(\bar{x}(\bar{\omega}), \bar{y})^\top [\lambda - \kappa\bar{\lambda}]\big]\mathsf{S}[d](\omega)\mathrm{d}\delta_{\bar{\omega}}(\omega) \\ &+ \int_\Omega \big[\sigma_x\beta(\omega)\psi(\omega) + \sigma_u(\bar{u}(\omega) - u_\mathrm{d}(\omega)) + \xi(\omega)\big]d(\omega)\mathrm{d}\omega \quad \text{for all } d \in L^2(\Omega). \end{aligned}$$

We set $\bar{a} := \nabla_x F_0(\bar{x}(\bar{\omega}), \bar{y}) + \nabla_x g(\bar{x}(\bar{\omega}), \bar{y})^\top [\lambda - \kappa\bar{\lambda}]$ and use Lemma 5.32 in order to see that this variational equation is equivalent to

$$0 = \int_\Omega \big[\beta(\omega)\big(\bar{a}\zeta(\omega) + \sigma_x\psi(\omega)\big) + \sigma_u(\bar{u}(\omega) - u_\mathrm{d}(\omega)) + \xi(\omega)\big]d(\omega)\mathrm{d}\omega \quad \text{for all } d \in L^2(\Omega)$$

where $\zeta \in W_0^{1,1}(\Omega)$ is the function defined via (5.46). Introducing $q(\omega) := \bar{a}\,\zeta(\omega) + \sigma_x\psi(\omega)$ for all $\omega \in \Omega$, (5.47b) is valid. Clearly, we have $H^2(\Omega) \cap H_0^1(\Omega) \subseteq W_0^{1,1}(\Omega)$ and, thus, $q \in W_0^{1,1}(\Omega)$. Finally, for arbitrary $\phi \in C_0^\infty(\Omega)$, we check

$$\begin{aligned} -\int_\Omega q(\omega)\Delta\phi(\omega)\mathrm{d}\omega &= -\bar{a}\int_\Omega \zeta(\omega)\Delta\phi(\omega)\mathrm{d}\omega - \sigma_x\int_\Omega \psi(\omega)\Delta\phi(\omega)\mathrm{d}\omega \\ &= \bar{a}\,\phi(\bar{\omega}) + \sigma_x\int_\Omega \nabla\psi(\omega)\cdot\nabla\phi(\omega)\mathrm{d}\omega = \bar{a}\,\phi(\bar{\omega}) + \sigma_x\int_\Omega \big(\bar{x}(\omega) - x_\mathrm{d}(\omega)\big)\phi(\omega)\mathrm{d}\omega \end{aligned}$$

which shows (5.47a). This completes the proof. □

Remark 5.34. From the proof of Theorem 5.33 we easily see that the adjoint function $q \in W_0^{1,1}(\Omega)$ can be decomposed into the regular part $\sigma_x\psi$ coming from $H^2(\Omega) \cap H_0^1(\Omega)$ and a less regular part $\bar{a}\,\zeta$ which belongs only to $W_0^{1,1}(\Omega)$. However, due to the appearence of state constraints in (OC) and, thus, in (5.36), this phenomenon had to be expected. It is documented in other papers where optimal control problems of PDEs with finitely many pointwise state constraints are considered, see [21, 22].
Note that the adjoint state in Theorem 5.27 possesses the same degree of regularity as the state function. This is not surprising since the pointwise state constraint actually is nothing else but an additional terminal condition and it is well-known from the theory of ODE control that this type of constraints only influences the resulting transversality conditions.

Remark 5.35. Let us consider the weak formulation of the slightly more general linear PDE

$$\begin{aligned} -\Delta x(\omega) + \alpha x(\omega) &= \beta(\omega)u(\omega) && \text{f.a.a. } \omega \in \Omega, \\ x(\omega) &= 0 && \text{f.a.a. } \omega \in \mathrm{bd}\,\Omega. \end{aligned}$$

for some constant $\alpha \geq 0$. Its solution operator still maps from $L^2(\Omega)$ to $H^2(\Omega) \cap H_0^1(\Omega)$, see [17, Proposition 6.15]. The derivation of the necessary optimality conditions from Theorem 5.22 for the corresponding problem (OC) parallels our above argumentation.

6. Conclusions and outlook

In this thesis, we derived new results which address variational analysis in function spaces, MPCCs in Banach spaces, and bilevel programming problems in Banach spaces. We used our theoretical findings to state applicable necessary optimality conditions for three different classes of bilevel programming problems.

Chapter 2 was dedicated to the gathering of preliminary results from functional analysis and optimization theory we needed in order to deal with MPCCs and bilevel programming problems in Banach spaces. Since these problems generally suffer from an inherent lack of convexity and/or smoothness, we decided to study tools of variational analysis introduced by Boris Mordukhovich. Especially, we took a closer look at the SNC property of pointwise defined sets in different function spaces in Section 2.3.1. Amongst others, it has been shown that reasonable sets which are often used in optimal control to define control or state constraints in the reflexive function spaces $L^p(\Omega)$ and $W^{1,p}(\Omega)$ with $1 < p < \infty$ are nowhere SNC. In view of [90] and other publications by Mordukhovich where the SNC property is demanded somehow carefree in most of the results, this observation is quite alarming.
One of the main issues of this thesis is the consideration of MPCCs whose complementarity constraints are given in a Lebesgue space. This abstract model covers optimal control problems with mixed control-state complementarity constraints which were recently studied in [56]. The complementarity constraint can be reformulated as an abstract constraint comprising a pointwise defined set in a Lebesgue space induced by a measurable set-valued mapping with nonconvex images. For the derivation of necessary optimality conditions for the underlying optimization problem, it is essential to be familiar with the variational geometry of this set. This motivated our study of the broad class of so-called decomposable sets in Section 2.3.5. As a supplementary result, we obtained that a nonempty, closed, decomposable set is weakly sequentially closed if and only if it is weakly closed which is a remarkable property. We derived an explicit formula for the weak closure of a decomposable set. In the future, we aim for a formula which characterizes the associated weak sequential closure. Under mild assumptions, we derived explicit formulae for the associated Bouligand and Clarke tangent cone as well as the Fréchet, strong limiting, and Clarke normal cone. Furthermore, we were able to show that the limiting normal cone to a decomposable set is a superset of the weak sequential closure of the associated strong limiting normal cone and, additionally, always dense in the associated Clarke normal cone, see Propositions 2.50 and 2.51. Although these results are strong enough to obtain an explicit characterization of the limiting normal cone to the complementarity set described above, we did not obtain explicit formulae for the limiting normal cone and the weak tangent cone to general decomposable sets. This is a nearby topic of future research.

In Chapter 3, we studied general MPCCs in Banach spaces. Since reasonable constraint qualifications like the regularity condition of Kurcyusz, Robinson, and Zowe fail to be satisfied at any feasible point of such a problem, the KKT conditions turn out to be a too restrictive necessary criterion for local optimality. Thus, one is in need of weaker necessary optimality conditions and constraint qualifications in order to deal with MPCCs. Gerd Wachsmuth introduced and studied a reasonable concept of strong stationarity in [121, 124]. We proceeded this research by introducing generalized concepts of weak and Mordukhovich stationarity, and we investigated the relationship between these three stationarity notions. It turned out that strong stationarity always implies weak stationarity. Furthermore, we were able to show that any strongly stationary point is also Mordukhovich stationary if the underlying cone which induces the complementarity condition is polyhedric. Using the theory of vector lattices and the polyhedricity of the complementarity cone, we were in position to formulate conditions which ensure that a Mordukhovich stationary point is weakly stationary as well. Furthermore, we presented constraint qualifications which imply that a local minimizer of an MPCC satisfies the aforementioned stationarity conditions. Subsequently, we applied the

obtained results to MPCCs whose complementarity cone equals the cone of nonnegative functions in a reflexive Lebesgue space or is polyhedral. An important consequence of Section 2.3.5 turned out to be the equivalence of Mordukhovich and weak stationarity for MPCCs in Lebesgue spaces. Furthermore, we depicted that the constraint qualifications arising from Mordukhovich's theory of variational analysis are not applicable to these MPCCs since their complementarity set is nowhere SNC. In the future, it needs to be investigated whether some pointwise counterpart of the finite-dimensional concept of Mordukhovich stationarity can be derived as an applicable necessary optimality condition for MPCCs in Lebesgue spaces. In view of Proposition 2.49, this requires some knowledge on the calculus of strong limiting normals which is not available yet. An important task for our future research seems to be a more general clarification of the relationship between the introduced stationarity notions under less restrictive assumptions. Furthermore, it is an open question whether other notions of stationarity which are well-known from standard complementarity programming can be generalized to the setting of Banach spaces. Finally, if MPCCs are considered whose complementarity constraint is induced by the nonnegative cone in $H_0^1(\Omega)$, then we know from Example 3.13 that the common relations between the concepts of strong, Mordukhovich, and weak stationarity hold. However, it is completely unclear what the limiting normal cone to the corresponding complementarity set looks like. This has to be investigated in the future.

We proceeded by considering a general bilevel optimization problem in Banach spaces in Chapter 4. The three main approaches of transformation used to convert the hierarchical optimization model into a single-level program (unique lower level solution, KKT reformulation, and optimal value reformulation), see [98], were applied to derive necessary optimality conditions.
First, we investigated a bilevel programming problem whose lower level is fully convex and governed by a so-called state equation equipped with control constraints. It has been shown that under certain assumptions, the lower level solution is unique, directionally differentiable, and that the directional derivative can be computed as the solution of a nonsmooth equation or, equivalently, a complementarity model. Afterwards, we used the theory of MPCCs from Chapter 3 in order to derive necessary optimality conditions for the corresponding bilevel model and discussed the case where the lower level control constraint set is a cone in more detail. Since our lower level of interest reflects a parametric optimal control problem with control constraints, we should transfer our results to the function space setting. However, due to Remark 4.7, this is not possible without additional assumptions on the underlying data. The technical details need to be discussed in the future.
Under certain convexity and regularity assumptions on a more general lower level problem, it is possible to replace it by its KKT conditions which we did in Section 4.2. However, in light of [28], this should not be done too light-hearted. Thus, we studied the relationship between the original bilevel optimization problem and its KKT reformulation in more detail. Both problems are (in a certain sense) equivalent w.r.t. global optimal solutions. By means of Example 4.20 we have shown that we cannot generalize the considerations of [28] to the infinite-dimensional situation if local optimal solutions are investigated. Nevertheless, we were able to show the local equivalence of the problems under more restrictive assumptions than in the finite-dimensional setting. However, these assumptions always imply the uniqueness of the lower level Lagrange multiplier which is quite restrictive. In the future, it needs to be clarified whether the local equivalence of the models can be preserved under less restrictive assumptions which allow the lower level Lagrange multiplier set to be no singleton. We continued our considerations by formulating necessary optimality conditions for the bilevel optimization model via its KKT reformulation. Therefore, we used the results of Chapter 3 again. In Example 4.30, we presented that the necessary constraint qualifications may depend on the choice of the lower level Lagrange multiplier.
Finally, we exploited the lower level optimal value function in order to find an equivalent single-level surrogate problem of the bilevel programming model. For the derivation of necessary optimality conditions via this approach, we used a concept of partial penalization. Our argumentation mainly generalized the one in [29] and [138] to the Banach space setting. In order to follow this approach in the future, we need to find some more results on the generalized differentiability of marginal functions to parametric optimal control problems. Furthermore, it has to be examined whether there exist classes of bilevel programming problems in function spaces where the so-called partial calmness property is inherent or at least easy to check.

We concluded this thesis by applying the results of Chapters 3 and 4 to special classes of bilevel programming problems in Chapter 5.

First, we studied a hierarchical model comprising a semidefinite lower level problem whose solution is unique. After we had carried out some variational analysis in the space $\mathcal{S}_p$, the findings of Section 4.1 turned out to be applicable. Thus, we obtained necessary optimality conditions for the corresponding bilevel programming problem. Since it was possible to state the latter equivalently as a semidefinite MPCC, we compared the obtained results to the existing literature on semidefinite complementarity programming. Next, we used our results from Chapter 3 and Section 4.2 in order to derive necessary optimality conditions for a bilevel optimal control model with optimal control problems of ODEs at both levels and lower level control constraints. We stated the corresponding weak and strong stationarity conditions and were able to construct a constraint qualification implying all local optimal solutions of the bilevel programming problem to be strongly stationary. This regularity condition is easy to check since it reduces to the controllability of a linear system of ODEs.
Finally, we considered a nonspecified optimal control problem with control constraints and an implicit pointwise state constraint arising from a finite-dimensional parametric optimization problem whose parameter equals a certain realization of the state function at a fixed point of the underlying domain. We were able to show the existence of a global solution to that problem under mild assumptions. Following our results of Section 4.3, we stated abstract optimality conditions for the general model. We specified these conditions in terms of linear ODEs and Poisson's equation in order to show that the theory is applicable to optimal control problems of ODEs and PDEs. This way, we continued the consideration of this problem class we already studied in [13, 14, 15, 74].

A. Supplementary results

Here we provide some results supporting our argumentation in the main part of the thesis. All the subsequent lemmas address analytical problems in function spaces and possess mainly technical but standard proofs.

First, we present a simple consequence of the dominated convergence theorem which can be found in [16, Theorem 2.8.1] (general form) or [114, Theorem 5.2.2] (tailored to L^p-spaces with $p \in [1, \infty)$).

Lemma A.1. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete and σ-finite measure space, let $\{\Omega_k\} \subseteq \Sigma$ be a sequence satisfying $\mathfrak{m}(\Omega_k) \downarrow 0$, and let $p \in [1, \infty)$ as well as $u \in L^p(\mathfrak{M})$ be arbitrarily chosen. Then $\chi_{\Omega_k} u \to 0$ and $(1 - \chi_{\Omega_k})u \to u$ hold true w.r.t. the convergence in $L^p(\mathfrak{M})$.

Proof. By definition of the characteristic function and $\mathfrak{m}(\Omega_k) \downarrow 0$, the sequences $\{\chi_{\Omega_k} u\}$ and $\{(1 - \chi_{\Omega_k})u\}$ converge pointwise almost everywhere on Ω to 0 and u, respectively. Furthermore, these sequences are both majorized by $|u| \in L^p(\mathfrak{M})$, i.e. we have

$$\forall \omega \in \Omega\colon \quad |\chi_{\Omega_k}(\omega) u(\omega)| \leq |u(\omega)|, \quad |(1 - \chi_{\Omega_k}(\omega))u(\omega)| \leq |u(\omega)|.$$

Thus, the lemma's assertion follows from the dominated convergence theorem in Lebesgue spaces, see [114, Theorem 5.2.2]. □

Furthermore, we need the following two technical convergence results in Lebesgue spaces. Their proofs follow from standard arguments but, however, are included for the reader's convenience.

Lemma A.2. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete and σ-finite measure space, let $p \in (1, \infty)$ and $m \in \mathbb{N}$ be fixed, and choose $\{u_k\} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ such that $u_k \to \bar{u}$ in $L^p(\mathfrak{M}, \mathbb{R}^m)$ holds for some $\bar{u} \in L^p(\mathfrak{M}, \mathbb{R}^m)$. Then for any $\varepsilon > 0$, there is a set $E \in \Sigma$ with $\mathfrak{m}(E) \leq \varepsilon$ and a subsequence $\{u_{k_l}\}$ of $\{u_k\}$ such that $u_{k_l} \to \bar{u}$ holds true in $L^\infty(\mathfrak{M}|_{\Omega \setminus E}, \mathbb{R}^m)$.

Proof. Due to the postulated convergence in $L^p(\mathfrak{M}, \mathbb{R}^m)$, we can choose $\{u_{k_l}\}$ with the following property:

$$\forall l \in \mathbb{N} \forall h \geq l\colon \quad \|u_{k_h} - \bar{u}\|_{L^p(\mathfrak{M}, \mathbb{R}^m)} \leq 2^{-2l}.$$

For all $l \in \mathbb{N}$, let us define sets $\Omega_l \in \Sigma$ by

$$\Omega_l := \{\omega \in \Omega \,|\, |u_{k_l}(\omega) - \bar{u}(\omega)|_2 \geq 2^{-l}\}.$$

Then we have

$$2^{-lp} \mathfrak{m}(\Omega_l) = \int_{\Omega_l} 2^{-lp} \mathrm{d}\mathfrak{m} \leq \int_{\Omega_l} |u_{k_l}(\omega) - \bar{u}(\omega)|_2^p \mathrm{d}\mathfrak{m} \leq \|u_{k_l} - \bar{u}\|^p_{L^p(\mathfrak{M}, \mathbb{R}^m)} \leq 2^{-2lp}$$

and, thus, $\mathfrak{m}(\Omega_l) \leq 2^{-lp}$. Let us introduce $E_j := \bigcup_{l=j}^\infty \Omega_l$ for all $j \in \mathbb{N}$. Then we have

$$\mathfrak{m}(E_j) \leq \sum_{l=j}^\infty 2^{-lp} < \sum_{l=j}^\infty 2^{-l} = 2^{1-j}$$

for any $j \in \mathbb{N}$. Choose $j_0 \in \mathbb{N}$ such that $\varepsilon \geq 2^{1-j_0}$ holds and set $E := E_{j_0}$. Then, for any $l \geq j_0$ and any $\omega \in \Omega \setminus E$, we obtain $|u_{k_l}(\omega) - \bar{u}(\omega)|_2 \leq 2^{-l}$ and, thus, we have

$$\sup_{\omega \in \Omega \setminus E} |u_{k_l}(\omega) - \bar{u}(\omega)|_2 \leq 2^{-l}$$

which shows $\|u_{k_l} - \bar{u}\|_{L^\infty(\mathfrak{M}|_{\Omega \setminus E}, \mathbb{R}^m)} \to 0$ as $l \to \infty$. □

Lemma A.3. Let $\mathfrak{M} = (\Omega, \Sigma, \mathfrak{m})$ be a complete and finite measure space, let $L^1(\mathfrak{M})$ be separable, let $p \in (1, \infty)$ as well as $m \in \mathbb{N}$ be fixed, and choose $\{u_k\} \subseteq L^p(\mathfrak{M}, \mathbb{R}^m)$ such that $u_k \rightharpoonup \bar{u}$ in $L^p(\mathfrak{M}, \mathbb{R}^m)$ holds for some $\bar{u} \in L^p(\mathfrak{M}, \mathbb{R}^m)$. If $\{u_k\}$ is bounded in $L^\infty(\mathfrak{M}, \mathbb{R}^m)$, then we have $u_k \overset{\star}{\rightharpoonup} \bar{u}$ in $L^\infty(\mathfrak{M}, \mathbb{R}^m)$.

Proof. By $p' \in (1, \infty)$ we denote the conjugate coefficient of p. First, we want to show $\bar{u} \in L^\infty(\mathfrak{M}, \mathbb{R}^m)$. Therefore, observe that the boundedness of $\{u_k\}$ in $L^\infty(\mathfrak{M}, \mathbb{R}^m)$ and the separability of $L^1(\mathfrak{M}, \mathbb{R}^m)$ imply that $\{u_k\}$ possesses a weakly* convergent subsequence $\{u_{k_l}\}$ with weak* limit $\tilde{u} \in L^\infty(\mathfrak{M}, \mathbb{R}^m)$, see [5, Corollary 2.4.2]. From $L^{p'}(\mathfrak{M}, \mathbb{R}^m) \subseteq L^1(\mathfrak{M}, \mathbb{R}^m)$ we deduce for any $v \in L^{p'}(\mathfrak{M}, \mathbb{R}^m)$:

$$\lim_{k\to\infty} \langle v, u_k \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} = \lim_{l\to\infty} \langle v, u_{k_l} \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} = \lim_{l\to\infty} \langle u_{k_l}, v \rangle_{L^1(\mathfrak{M},\mathbb{R}^m)} = \langle \tilde{u}, v \rangle_{L^1(\mathfrak{M},\mathbb{R}^m)} = \langle v, \tilde{u} \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)}.$$

Thus, we have $u_k \rightharpoonup \tilde{u}$ in $L^p(\mathfrak{M}, \mathbb{R}^m)$ and due to the uniqueness of the weak limit, $\bar{u} = \tilde{u} \in L^\infty(\mathfrak{M}, \mathbb{R}^m)$ is obtained.
Now, we start to verify $u_k \overset{\star}{\rightharpoonup} \bar{u}$ in $L^\infty(\mathfrak{M}, \mathbb{R}^m)$. Let $w \in L^1(\mathfrak{M}, \mathbb{R}^m)$ be given. Since $L^{p'}(\mathfrak{M}, \mathbb{R}^m)$ is dense in $L^1(\mathfrak{M}, \mathbb{R}^m)$, for any $l \in \mathbb{N}$, we find $w_l \in L^{p'}(\mathfrak{M}, \mathbb{R}^m)$ satisfying $\|w - w_l\|_{L^1(\mathfrak{M},\mathbb{R}^m)} \leq \frac{1}{l}$. This leads to

$$\begin{aligned} \left| \langle u_k - \bar{u}, w \rangle_{L^1(\mathfrak{M},\mathbb{R}^m)} \right| &= \left| \langle w_l, u_k - \bar{u} \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} + \langle u_k - \bar{u}, w - w_l \rangle_{L^1(\mathfrak{M},\mathbb{R}^m)} \right| \\ &\leq \left| \langle w_l, u_k - \bar{u} \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} \right| + \|u_k - \bar{u}\|_{L^\infty(\mathfrak{M},\mathbb{R}^m)} \|w - w_l\|_{L^1(\mathfrak{M},\mathbb{R}^m)}. \end{aligned}$$

Noting that $\{u_k - \bar{u}\}$ is bounded in $L^\infty(\mathfrak{M}, \mathbb{R}^m)$, we find a constant $\gamma > 0$ such that $\|u_k - \bar{u}\|_{L^\infty(\mathfrak{M},\mathbb{R}^m)} \leq \gamma$ is valid for all $k \in \mathbb{N}$. On the other hand, we have $u_k \rightharpoonup \bar{u}$ in $L^p(\mathfrak{M}, \mathbb{R}^m)$ which leads to

$$\limsup_{k\to\infty} \left| \langle u_k - \bar{u}, w \rangle_{L^1(\mathfrak{M},\mathbb{R}^m)} \right| \leq \limsup_{k\to\infty} \left| \langle w_l, u_k - \bar{u} \rangle_{L^p(\mathfrak{M},\mathbb{R}^m)} \right| + \frac{\gamma}{l} = \frac{\gamma}{l}$$

for all $l \in \mathbb{N}$. Taking the limit $l \to \infty$, we infer

$$\langle u_k - \bar{u}, w \rangle_{L^1(\mathfrak{M},\mathbb{R}^m)} \to 0.$$

Since $w \in L^1(\mathfrak{M}, \mathbb{R}^m)$ was chosen arbitrarily, we have $u_k \overset{\star}{\rightharpoonup} \bar{u}$ in $L^\infty(\mathfrak{M}, \mathbb{R}^m)$ which completes the proof. □

The next lemma provides a truncation result we need in order to prove Lemmas 2.19 and 2.21. The proof is similar to the validation of [5, Theorem 5.8.2] or [122, Theorem A.2].

Lemma A.4. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz boundary and let $p \in (1, \infty)$ be fixed. For $k \in \mathbb{N}$, we define the truncation $T_k \colon \mathbb{R} \to \mathbb{R}$ by

$$\forall x \in \mathbb{R} \colon \quad T_k(x) := \min\{k; x\}.$$

Then the associated Nemytskii operator, i.e. the mapping $u \mapsto T_k \circ u$, denoted by T_k as well, maps $W^{1,p}(\Omega)$ to $W^{1,p}(\Omega)$ and we have

$$\forall u \in W^{1,p}(\Omega)\, \forall i \in \{1, \dots, d\} \colon \quad D_i(T_k u)(\omega) = \begin{cases} D_i u(\omega) & \text{if } u(\omega) \leq k, \\ 0 & \text{if } u(\omega) > k \end{cases}$$

for almost every $\omega \in \Omega$. Moreover, for any $u \in W^{1,p}(\Omega)$, we have $T_k u \to u$ in $W^{1,p}(\Omega)$ as $k \to \infty$.

Proof. We invoke [82, Theorem 2.1] in order to see that $T_k u \in W^{1,p}(\Omega)$ holds for any $u \in W^{1,p}(\Omega)$ and $k \in \mathbb{N}$. Note that T_k is not differentiable, i.e. classical chain rules as presented in [81] are not applicable. Fix some $k \in \mathbb{N}$. For $\sigma > 0$, we define a differentiable approximation $T_k^\sigma \colon \mathbb{R} \to \mathbb{R}$ of T_k by

$$\forall x \in \mathbb{R} \colon \quad T_k^\sigma(x) := \begin{cases} \frac{\sigma}{2} + x & \text{if } x < k - \sigma, \\ k - \frac{1}{2\sigma}(k - x)^2 & \text{if } k - \sigma \leq x < k, \\ k & \text{if } x \geq k. \end{cases}$$

It is easy to see

$$\forall x \in \mathbb{R}\colon \quad \tfrac{\partial}{\partial x}(T_k^\sigma)(x) = \begin{cases} 1 & \text{if } x < k-\sigma, \\ \frac{1}{\sigma}(k-x) & \text{if } k-\sigma \le x < k, \\ 0 & \text{if } x \ge k. \end{cases}$$

We apply [82, Theorem 2.1] once more in order to obtain $T_k^\sigma u \in W^{1,p}(\Omega)$ for any fixed $u \in W^{1,p}(\Omega)$. Due to the differentiability of T_k^σ, we exploit the chain rule, see [81, Theorem 2.1], in order to obtain

$$\forall i \in \{1,\dots,d\}\colon \quad D_i(T_k^\sigma u)(\omega) = \tfrac{\partial}{\partial x}(T_k^\sigma)(u(\omega))D_i u(\omega)$$

almost everywhere on Ω. Thus, a promising candidate for $D_i(T_k u)$, $i \in \{1,\dots,d\}$, is given by

$$\forall \omega \in \Omega\colon \quad v_i(\omega) := \begin{cases} D_i u(\omega) & \text{if } u(\omega) \le k, \\ 0 & \text{if } u(\omega) > k. \end{cases}$$

We obtain

$$|T_k^\sigma(u(\omega)) - T_k(u(\omega))| \le \tfrac{3\sigma}{2}$$

almost everywhere on Ω. Consequently, $T_k^\sigma u$ converges a.e. on Ω pointwise to $T_k u$ as σ falls to zero. Since the above estimate provides an integrable upper bound, $T_k^\sigma u$ converges to $T_k u$ in $L^p(\Omega)$ as σ falls to zero by the dominated convergence theorem. For fixed $i \in \{1,\dots,d\}$, we have the pointwise convergence of $D_i(T_k^\sigma u)$ to v_i a.e. on Ω as σ tends to zero. Additionally, taking a look at the above results,

$$|D_i(T_k^\sigma u)(\omega) - v_i(\omega)| \le 2|D_i u(\omega)|$$

follows almost everywhere on Ω, i.e. by the dominated convergence theorem, $D_i(T_k^\sigma u)$ converges to v_i in $L^p(\Omega)$ as σ falls to zero. Using Hölder's inequality, for any function $\phi \in C_0^\infty(\Omega)$,

$$\int_\Omega v_i(\omega)\phi(\omega)\mathrm{d}\omega \xleftarrow{\sigma \searrow 0} \int_\Omega D_i(T_k^\sigma u)(\omega)\phi(\omega)\mathrm{d}\omega = -\int_\Omega (T_k^\sigma u)(\omega) D_i\phi(\omega)\mathrm{d}\omega \xrightarrow{\sigma \searrow 0} -\int_\Omega (T_k u)(\omega) D_i\phi(\omega)\mathrm{d}\omega$$

is valid and, thus, $v_i = D_i(T_k u)$ holds true.
Finally, we want to show that $T_k u \to u$ in $W^{1,p}(\Omega)$ holds true as $k \to \infty$. Therefore, we fix $u \in W^{1,p}(\Omega)$. Let us define $\Omega_k := \{\omega \in \Omega \,|\, u(\omega) > k\}$ for any $k \in \mathbb{N}$. Clearly, all these sets are measurable and satisfy $\mathfrak{l}(\Omega_k) \downarrow 0$ (otherwise, we would not have $u \in L^p(\Omega)$). We have $u(\omega) - (T_k u)(\omega) = (u(\omega) - k)\chi_{\Omega_k}(\omega)$ for any $k \in \mathbb{N}$ and $\omega \in \Omega$. Thus, $T_k u$ converges pointwise to u almost everywhere on Ω. Moreover, we have

$$|u(\omega) - (T_k u)(\omega)| = |(u(\omega) - k)\chi_{\Omega_k}(\omega)| \le |u(\omega)|$$

almost everywhere on Ω and, thus, the dominated convergence theorem yields $T_k u \to u$ in $L^p(\Omega)$ as $k \to \infty$. Similarly, we have $D_i u - D_i(T_k u) = D_i u\, \chi_{\Omega_k}$ which, by means of the dominated convergence theorem, shows $D_i(T_k u) \to D_i u$ in $L^p(\Omega)$ for all $i \in \{1,\dots,d\}$ as $k \to \infty$. Hence, $T_k u \to u$ in $W^{1,p}(\Omega)$ as $k \to \infty$ is satisfied and the proof is completed. □

For the discussion of the final two results, we need a nonempty, bounded interval $\Omega := (0,T) \subseteq \mathbb{R}$ and a positive natural number $n \in \mathbb{N}$. Identifying the Hilbert space $AC^{1,2}(\Omega,\mathbb{R}^n)$ with its dual by means of Riesz's representation theorem, it will be necessary to identify elements of $AC^{1,2}(\Omega,\mathbb{R}^n)^\star$ with a vector function in $AC^{1,2}(\Omega,\mathbb{R}^n)$. Therefore, we included the following lemma which is related to [117, Lemma 3.1(b)].

Lemma A.5. Let $v \in L^2(\Omega,\mathbb{R}^n)$ be fixed and let $a_v \in \mathbb{R}^n$ be a vector. We consider the dual vector $v^* \in AC^{1,2}(\Omega,\mathbb{R}^n)^\star$ given by

$$\forall u \in AC^{1,2}(\Omega,\mathbb{R}^n)\colon \quad v^*[u] := a_v \cdot u(T) + \int_0^T v(t)\cdot u(t)\mathrm{d}t.$$

Then v^* can be identified with a function in $AC^{1,2}(\Omega,\mathbb{R}^n)$ defined below:

$$(v^*(0), \nabla v^*) = \left(a_v + \int_0^T v(t)\mathrm{d}t,\, a_v + \int_\cdot^T v(t)\mathrm{d}t\right).$$

Proof. For the proof, we use integration by parts and the definition of the dual pairing in $AC^{1,2}(\Omega,\mathbb{R}^n)$ to obtain

$$\begin{aligned}
\langle v^*, u\rangle_{AC^{1,2}(\Omega,\mathbb{R}^n)} &= v^*[u] = a_v \cdot u(T) + \int_0^T v(t)\cdot u(t)\mathrm{d}t \\
&= a_v \cdot u(T) + \left(\int_0^T v(t)\mathrm{d}t\right)\cdot u(T) - \int_0^T \left(\int_0^t v(s)\mathrm{d}s\right)\cdot \nabla u(t)\mathrm{d}t \\
&= \left(a_v + \int_0^T v(t)\mathrm{d}t\right)\cdot\left(u(0) + \int_0^T \nabla u(t)\mathrm{d}t\right) - \int_0^T \left(\int_0^t v(s)\mathrm{d}s\right)\cdot\nabla u(t)\mathrm{d}t \\
&= \left(a_v + \int_0^T v(t)\mathrm{d}t\right)\cdot u(0) + \int_0^T a_v\cdot\nabla u(t)\mathrm{d}t + \int_0^T\left(\int_t^T v(s)\mathrm{d}s\right)\cdot\nabla u(t)\mathrm{d}t \\
&= \underbrace{\left(a_v + \int_0^T v(t)\mathrm{d}t\right)}_{=v^*(0)}\cdot u(0) + \int_0^T \underbrace{\left(a_v + \int_t^T v(s)\mathrm{d}s\right)}_{=\nabla v^*(t)}\cdot\nabla u(t)\mathrm{d}t.
\end{aligned}$$

This shows the claim. □

Finally, we show how the adjoint of a certain operator which describes linear constraints of an optimal control problem with ODE constraints can be computed. Again, we deal with the space $AC^{1,2}(\Omega,\mathbb{R}^n)$. A related result can be found in [87, Appendix 1].

Lemma A.6. For natural numbers $n,m,k,l\in\mathbb{N}$ and real matrices $\mathbf{M}\in\mathbb{R}^{n\times n}$, $\mathbf{N}\in\mathbb{R}^{n\times m}$, $\mathbf{P}\in\mathbb{R}^{k\times n}$, $\mathbf{Q}\in\mathbb{R}^{k\times m}$, and $\mathbf{R}\in\mathbb{R}^{l\times n}$, we define $\mathtt{D}\in\mathbb{L}[AC^{1,2}(\Omega,\mathbb{R}^n)\times L^2(\Omega,\mathbb{R}^m), AC^{1,2}(\Omega,\mathbb{R}^n)\times L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^l]$ as stated below for all $z\in AC^{1,2}(\Omega,\mathbb{R}^n)$ and $u\in L^2(\Omega,\mathbb{R}^m)$:

$$\mathtt{D}[z,u] := \left(z(\cdot) - \int_0^{\cdot}\left[\mathbf{M}z(t)+\mathbf{N}u(t)\right]\mathrm{d}t,\ \mathbf{P}z(\cdot)+\mathbf{Q}u(\cdot),\ \mathbf{R}z(T)\right).$$

Its adjoint is given by

$$\begin{aligned}
\mathtt{D}^\star[w,v,r] = \Bigg(\Bigg(&w(0) + \mathbf{R}^\top r + \int_0^T\left[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\right]\mathrm{d}t, \\
&\nabla w(\cdot) + \mathbf{R}^\top r + \int_{\cdot}^T\left[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\right]\mathrm{d}t\Bigg),\ \mathbf{Q}^\top v(\cdot) - \mathbf{N}^\top\nabla w(\cdot)\Bigg)
\end{aligned}$$

for arbitrary $w\in AC^{1,2}(\Omega,\mathbb{R}^n)$, $v\in L^2(\Omega,\mathbb{R}^k)$, and $r\in\mathbb{R}^l$.

Proof. We set $\mathcal{X} := AC^{1,2}(\Omega,\mathbb{R}^n)\times L^2(\Omega,\mathbb{R}^m)$ as well as $\mathcal{Y} := AC^{1,2}(\Omega,\mathbb{R}^n)\times L^2(\Omega,\mathbb{R}^k)\times\mathbb{R}^l$ and choose $(z,u)\in\mathcal{X}$ and $(w,v,r)\in\mathcal{Y}$ arbitrarily. First, we show the continuity of $\mathtt{D}$. Therefore, we observe

$$\begin{aligned}
\|z\|_{L^2(\Omega,\mathbb{R}^n)} &= \left\|z(0) + \int_0^{\cdot}\nabla z(t)\mathrm{d}t\right\|_{L^2(\Omega,\mathbb{R}^n)} \le \sqrt{T}\,|z(0)|_2 + \left\|\int_0^{\cdot}\nabla z(t)\mathrm{d}t\right\|_{L^2(\Omega,\mathbb{R}^n)} \\
&= \sqrt{T}\,|z(0)|_2 + \left(\int_0^T\left|\int_0^t\nabla z(\tau)\mathrm{d}\tau\right|_2^2\mathrm{d}t\right)^{\frac12} \le \sqrt{T}\,|z(0)|_2 + \left(T\int_0^T\int_0^T|\nabla z(\tau)|_2^2\,\mathrm{d}\tau\mathrm{d}t\right)^{\frac12} \\
&\le \sqrt{T}\,|z(0)|_2 + T\,\|\nabla z\|_{L^2(\Omega,\mathbb{R}^n)} \le \max\{\sqrt{T};T\}\,\|z\|_{AC^{1,2}(\Omega,\mathbb{R}^n)}.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
|z(T)|_2 &= \left|z(0) + \int_0^T\nabla z(t)\mathrm{d}t\right|_2 \le |z(0)|_2 + \left|\int_0^T\nabla z(t)\mathrm{d}t\right|_2 \\
&\le |z(0)|_2 + \left(T\int_0^T|\nabla z(t)|_2^2\,\mathrm{d}t\right)^{\frac12} = |z(0)|_2 + \sqrt{T}\,\|\nabla z\|_{L^2(\Omega,\mathbb{R}^n)} \le \max\{1;\sqrt{T}\}\,\|z\|_{AC^{1,2}(\Omega,\mathbb{R}^n)}.
\end{aligned}$$

For the derivation of these estimates, we used Hölder's inequality componentwise. Now, we find scalars $\mu > 0$ only depending on $\mathbf{M}$ as well as $\mathbf{P}$, $\nu > 0$ only depending on $\mathbf{N}$ as well as $\mathbf{Q}$, and $\rho > 0$ only depending on $\mathbf{R}$ such that

$$\begin{aligned}\|\mathtt{D}[z,u]\|_{\mathcal{Y}} &= |z(0)|_2 + \|\nabla z(\cdot) - \mathbf{M}z(\cdot) - \mathbf{N}u(\cdot)\|_{L^2(\Omega,\mathbb{R}^n)} + \|\mathbf{P}z(\cdot) + \mathbf{Q}u(\cdot)\|_{L^2(\Omega,\mathbb{R}^k)} + |\mathbf{R}z(T)|_2 \\ &\leq \|z\|_{AC^{1,2}(\Omega,\mathbb{R}^n)} + \mu\,\|z\|_{L^2(\Omega,\mathbb{R}^n)} + \nu\,\|u\|_{L^2(\Omega,\mathbb{R}^m)} + \rho\,|z(T)|_2 \\ &\leq \Big(1 + \mu\max\{\sqrt{T};T\} + \rho\max\{1;\sqrt{T}\}\Big)\,\|z\|_{AC^{1,2}(\Omega,\mathbb{R}^n)} + \nu\,\|u\|_{L^2(\Omega,\mathbb{R}^m)} \\ &\leq \max\Big\{1 + \mu\max\{\sqrt{T};T\} + \rho\max\{1;\sqrt{T}\};\nu\Big\}\,\|(z,u)\|_{\mathcal{X}}\end{aligned}$$

holds true, i.e. $\mathtt{D}$ is continuous.
We use integration by parts and the definition of the dual pairing in the appearing function spaces to come up with

$$\begin{aligned}&\langle \mathtt{D}^\star[w,v,r],(z,u)\rangle_{\mathcal{X}} = \langle (w,v,r),\mathtt{D}[z,u]\rangle_{\mathcal{Y}} \\ &\quad= z(0)\cdot w(0) + \int_0^T \left[\nabla z(t) - \mathbf{M}\left(z(0) + \int_0^t \nabla z(\tau)\mathrm{d}\tau\right) - \mathbf{N}u(t)\right]\cdot \nabla w(t)\mathrm{d}t \\ &\qquad + \int_0^T \left[\mathbf{P}\left(z(0) + \int_0^t \nabla z(\tau)\mathrm{d}\tau\right) + \mathbf{Q}u(t)\right]\cdot v(t)\mathrm{d}t + \left[\mathbf{R}\left(z(0) + \int_0^T \nabla z(\tau)\mathrm{d}\tau\right)\right]\cdot r \\ &\quad= z(0)\cdot\left(w(0) + \mathbf{R}^\top r + \int_0^T \big[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\big]\mathrm{d}t\right) + \int_0^T u(t)\cdot\big[\mathbf{Q}^\top v(t) - \mathbf{N}^\top\nabla w(t)\big]\mathrm{d}t \\ &\qquad + \int_0^T \left[\nabla z(t)\cdot\big[\nabla w(t) + \mathbf{R}^\top r\big] + \left(\int_0^t \nabla z(\tau)\mathrm{d}\tau\right)\cdot\big[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\big]\right]\mathrm{d}t \\ &\quad= z(0)\cdot\left(w(0) + \mathbf{R}^\top r + \int_0^T \big[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\big]\mathrm{d}t\right) + \int_0^T u(t)\cdot\big[\mathbf{Q}^\top v(t) - \mathbf{N}^\top\nabla w(t)\big]\mathrm{d}t \\ &\qquad + \int_0^T \nabla z(t)\cdot\big[\nabla w(t) + \mathbf{R}^\top r\big]\mathrm{d}t + \left(\int_0^T \nabla z(t)\mathrm{d}t\right)\cdot\left(\int_0^T \big[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\big]\mathrm{d}t\right) \\ &\qquad - \int_0^T \nabla z(t)\cdot\left(\int_0^t \big[\mathbf{P}^\top v(\tau) - \mathbf{M}^\top\nabla w(\tau)\big]\mathrm{d}\tau\right)\mathrm{d}t \\ &\quad= z(0)\cdot\left(w(0) + \mathbf{R}^\top r + \int_0^T \big[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\big]\mathrm{d}t\right) + \int_0^T u(t)\cdot\big[\mathbf{Q}^\top v(t) - \mathbf{N}^\top\nabla w(t)\big]\mathrm{d}t \\ &\qquad + \int_0^T \nabla z(t)\cdot\big[\nabla w(t) + \mathbf{R}^\top r\big]\mathrm{d}t + \int_0^T \nabla z(t)\cdot\left(\int_t^T \big[\mathbf{P}^\top v(\tau) - \mathbf{M}^\top\nabla w(\tau)\big]\mathrm{d}\tau\right)\mathrm{d}t \\ &\quad= z(0)\cdot\left(w(0) + \mathbf{R}^\top r + \int_0^T \big[\mathbf{P}^\top v(t) - \mathbf{M}^\top\nabla w(t)\big]\mathrm{d}t\right) \\ &\qquad + \int_0^T \nabla z(t)\cdot\left(\nabla w(t) + \mathbf{R}^\top r + \int_t^T \big[\mathbf{P}^\top v(\tau) - \mathbf{M}^\top\nabla w(\tau)\big]\mathrm{d}\tau\right)\mathrm{d}t \\ &\qquad + \int_0^T u(t)\cdot\big[\mathbf{Q}^\top v(t) - \mathbf{N}^\top\nabla w(t)\big]\mathrm{d}t.\end{aligned}$$

This yields the presented formula for the adjoint operator $\mathtt{D}^\star$. □

Bibliography

[1] R. A. Adams, J. J. F. Fournier, *Sobolev spaces*, Elsevier Science, Oxford, 2003.

[2] R. P. Agarwal, D. O'Regan, *An Introduction to Ordinary Differential Equations*, Springer, New York, 2008.

[3] S. Albrecht, M. Leibold, M. Ulbrich, *A bilevel optimization approach to obtain optimal cost functions for human arm movements*, Numerical Algebra, Control and Optimization 2 (1) (2012) 105–127, DOI: 10.3934/naco.2012.2.105.

[4] S. Albrecht, M. Ulbrich, *Mathematical programs with complementarity constraints in the context of inverse optimal control for locomotion*, Optimization Methods and Software (2016) 1–29, DOI: 10.1080/10556788.2016.1225212.

[5] H. Attouch, G. Buttazzo, G. Michaille, *Variational analysis in Sobolev and BV spaces*, volume 6, MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[6] J.-P. Aubin, H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, MA, 2009.

[7] B. Bank, J. Guddat, D. Klatte, B. Kummer, K. Tammer, *Non-Linear Parametric Optimization*, Akademie-Verlag, Berlin, 1982.

[8] J. F. Bard, *Practical Bilevel Optimization: Algorithms and Applications*, Kluwer Academic, Dordrecht, 1998.

[9] S. Barnett, R. G. Cameron, *Introduction to Mathematical Control Theory*, Oxford University Press, Oxford, 1990.

[10] A. Ben-Tal, A. Nemirovski, *On Polyhedral Approximations of the Second-Order Cone*, Mathematics of Operations Research 26 (2) (2001) 193–205, DOI: 10.1287/moor.26.2.193.10561.

[11] A. Ben-Tal, A. Nemirovski, *Robust Convex Optimization*, Mathematics of Operations Research 23 (4) (1998) 769–805, DOI: 10.1287/moor.23.4.769.

[12] A. Ben-Tal, A. Nemirovski, *Robust Optimization*, Princeton University Press, Princeton, 2002.

[13] F. Benita, S. Dempe, P. Mehlitz, *Bilevel Optimal Control Problems with Pure State Constraints and Finite-dimensional Lower Level*, SIAM Journal on Optimization 26 (1) (2016) 564–588, DOI: 10.1137/141000889.

[14] F. Benita, P. Mehlitz, *Bilevel Optimal Control With Final-State-Dependent Finite-Dimensional Lower Level*, SIAM Journal on Optimization 26 (1) (2016) 718–752, DOI: 10.1137/15M1015984.

[15] F. Benita, P. Mehlitz, *Optimal Control Problems with Terminal Complementarity Constraints*, Preprint TU Bergakademie Freiberg (2016) 1–22, URL: http://tu-freiberg.de/fakult1/forschung/preprints.

[16] V. I. Bogachev, *Measure Theory*, Springer, Berlin, 2007.

[17] J. F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[18] H. Bonnel, J. Morgan, *Optimality Conditions for Semivectorial Bilevel Convex Optimal Control Problems*, in: *Computational and Analytical Mathematics: In Honor of Jonathan Borwein's 60th Birthday*, edited by D. H. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. D. Vanderwerff, H. Wolkowicz, Springer, New York, 2013 45–78, DOI: 10.1007/978-1-4614-7621-4_4.

[19] J. Borwein, Y. Lucet, B. S. Mordukhovich, *Compactly epi-Lipschitzian Convex Sets and Functions in Normed Spaces*, Journal of Convex Analysis 7 (2) (2000) 375–393.

[20] D. A. Carlson, *Existence of Optimal Controls for a Bi-Level Optimal Control Problem*, in: *Advances in Dynamic Games: Theory, Applications, and Numerical Methods*, edited by V. Křivan, G. Zaccour, Springer, Cham, 2013 71–84, DOI: 10.1007/978-3-319-02690-9_4.

[21] E. Casas, *Necessary and sufficient optimality conditions for elliptic control problems with finitely many pointwise state constraints*, ESAIM 14 (3) (2008) 575–589, DOI: 10.1051/cocv:2007063.

[22] E. Casas, M. Mateos, *Second Order Optimality Conditions for Semilinear Elliptic Control Problems with Finitely Many State Constraints*, SIAM Journal on Control and Optimization 40 (5) (2002) 1431–1454, DOI: 10.1137/S0363012900382011.

[23] M. Červinka, *Hierarchical Structures in Equilibrium Problems*, PhD thesis, Charles University Prague, 2008.

[24] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[25] S. Dempe, *A necessary and a sufficient optimality condition for bilevel programming problems*, Optimization 25 (4) (1992) 341–354, DOI: 10.1080/02331939208843831.

[26] S. Dempe, *Foundations of Bilevel Programming*, Kluwer, Dordrecht, 2002.

[27] S. Dempe, *On the directional derivative of the optimal solution mapping without linear independence constraint qualification*, Optimization 20 (4) (1989) 401–414, DOI: 10.1080/02331938908843460.

[28] S. Dempe, J. Dutta, *Is bilevel programming a special case of a mathematical program with complementarity constraints?*, Mathematical Programming 131 (1) (2012) 37–48, DOI: 10.1007/s10107-010-0342-1.

[29] S. Dempe, J. Dutta, B. S. Mordukhovich, *New necessary optimality conditions in optimistic bilevel programming*, Optimization 56 (5-6) (2007) 577–604, DOI: 10.1080/02331930701617551.

[30] S. Dempe, F. Mefo Kue, P. Mehlitz, *Optimality Conditions for Special Semidefinite Bilevel Optimization Problems*, Preprint TU Bergakademie Freiberg (2016) 1–25, URL: http://tu-freiberg.de/fakult1/forschung/preprints.

[31] S. Dempe, P. Mehlitz, *Lipschitz continuity of the optimal value function in parametric optimization*, Journal of Global Optimization 61 (2) (2015) 363–377, DOI: 10.1007/s10898-014-0169-z.

[32] S. Dempe, A. B. Zemkoho, *KKT Reformulation and Necessary Conditions for Optimality in Nonsmooth Bilevel Optimization*, SIAM Journal on Optimization 24 (4) (2014) 1639–1669, DOI: 10.1137/130917715.

[33] S. Dempe, A. B. Zemkoho, *On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem*, Nonlinear Analysis: Theory, Methods & Applications 75 (3) (2012) 1202 –1218, DOI: 10.1016/j.na.2011.05.097.

[34] S. Dempe, A. B. Zemkoho, *The bilevel programming problem: reformulations, constraint qualifications and optimality conditions*, Mathematical Programming 138 (1) (2013) 447–473, DOI: 10.1007/s10107-011-0508-5.

[35] S. Dempe, A. B. Zemkoho, *The Generalized Mangasarian-Fromovitz Constraint Qualification and Optimality Conditions for Bilevel Programs*, Journal of Optimization Theory and Applications 148 (1) (2011) 46–68, DOI: 10.1007/s10957-010-9744-8.

[36] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, N. Kalashnykova, *Bilevel Programming Problems - Theory, Algorithms and Applications to Energy Networks*, Springer, Berlin, 2015.

[37] C. Ding, D. Sun, J. J. Ye, *First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints*, Mathematical Programming 147 (1-2) (2014) 539–579, DOI: 10.1007/s10107-013-0735-z.

[38] A. L. Dontchev, R. T. Rockafellar, *Ample Parameterization of Variational Inclusions*, SIAM Journal on Optimization 12 (1) (2001) 170–187, DOI: 10.1137/S1052623400371016.

[39] M. Fabian, B. S. Mordukhovich, *Sequential normal compactness versus topological normal compactness in variational analysis*, Nonlinear Analysis: Theory, Methods & Applications 54 (6) (2003) 1057 –1067, DOI: 10.1016/S0362-546X(03)00126-3.

[40] A. V. Fiacco, J. Kyparisis, *Convexity and concavity properties of the optimal value function in parametric nonlinear programming*, Journal of Optimization Theory and Applications 48 (1) (1986) 95–126, DOI: 10.1007/BF00938592.

[41] F. Fisch, J. Lenz, F. Holzapfel, G. Sachs, *On the Solution of Bilevel Optimal Control Problems to Increase the Fairness in Air Races*, Journal of Guidance, Control, and Dynamics (4) (2012) 1292–1298, DOI: 10.2514/1.54407.

[42] M. L. Flegel, C. Kanzow, *A direct proof for M-stationarity under MPEC-GCQ for mathematical programs with equilibrium constraints*, in: *Optimization with Multivalued Mappings*, edited by S. Dempe, V. Kalashnikov, volume 2, Springer Optimization and Its Applications, Springer, New York, 2006 111–122, DOI: 10.1007/0-387-34221-4_6.

[43] M. L. Flegel, C. Kanzow, *A Fritz John Approach to First Order Optimality Conditions for Mathematical Programs with Equilibrium Constraints*, Optimization 52 (3) (2003) 277–286, DOI: 10.1080/0233193031000120020.

[44] M. L. Flegel, C. Kanzow, *Abadie-Type Constraint Qualification for Mathematical Programs with Equilibrium Constraints*, Journal of Optimization Theory and Applications 124 (3) (2005) 595–614, DOI: 10.1007/s10957-004-1176-x.

[45] M. L. Flegel, C. Kanzow, *On M-stationary points for mathematical programs with equilibrium constraints*, Journal of Mathematical Analysis and Applications 310 (1) (2005) 286 –302, DOI: 10.1016/j.jmaa.2005.02.011.

[46] R. Fletcher, S. Leyffer, D. Ralph, S. Scholtes, *Local Convergence of SQP Methods for Mathematical Programs with Equilibrium Constraints*, SIAM Journal on Optimization 17 (1) (2006) 259–286, DOI: 10.1137/S1052623402407382.

[47] S. Franke, P. Mehlitz, M. Pilecka, *Optimality conditions for the simple convex bilevel programming problem in Banach spaces*, Preprint TU Bergakademie Freiberg (2016) 1–26, URL: http://tu-freiberg.de/fakult1/forschung/preprints.

[48] M. Fukushima, Z.-Q. Luo, P. Tseng, *Smoothing Functions for Second-Order-Cone Complementarity Problems*, SIAM Journal on Optimization 12 (2) (2002) 436–460, DOI: 10.1137/S1052623400380365.

[49] M. Gerdts, *Optimal Control of ODEs and DAEs*, de Gruyter, Berlin, 2012.

[50] W. Geremew, B. S. Mordukhovich, N. M. Nam, *Coderivative calculus and metric regularity for constraint and variational systems*, Nonlinear Analysis: Theory, Methods & Applications 70 (1) (2009) 529 –552, DOI: 10.1016/j.na.2007.12.025.

[51] B. M. Glover, *A generalized Farkas lemma with applications to quasidifferentiable programming*, Zeitschrift für Operations Research 26 (1) (1982) 125–141, DOI: doi:10.1007/BF01917106.

[52] H. Goldberg, W. Kampowsky, F. Tröltzsch, *On Nemytskij Operators in L_p-Spaces of Abstract Functions*, Math. Nachr. 155 (1992) 127–140, DOI: 10.1002/mana.19921550110.

[53] G. Grätzer, *General Lattice Theory*, Akademie-Verlag, Berlin, 1978.

[54] R. Griesse, T. Grund, D. Wachsmuth, *Update strategies for perturbed nonsmooth equations*, Optimization Methods and Software 23 (3) (2008) 321–343, DOI: 10.1080/10556780701523551.

[55] M. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, 2002.

[56] L. Guo, J. J. Ye, *Necessary Optimality Conditions for Optimal Control Problems with Equilibrium Constraints*, SIAM Journal on Control and Optimization 54 (5) (2016) 2710–2733, DOI: 10.1137/15M1013493.

[57] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities*, J. Math. Soc. Japan 29 (4) (1977) 615–631, DOI: 10.2969/jmsj/02940615.

[58] K. Hatz, *Efficient Numerical Methods for Hierarchical Dynamic Optimization with Application to Cerebral Palsy Gait Modeling*, PhD thesis, University of Heidelberg, Germany, 2014.

[59] K. Hatz, J. P. Schlöder, H. G. Bock, *Estimating Parameters in Optimal Control Problems*, SIAM Journal on Scientific Computing 34 (3) (2012) A1707–A1728, DOI: 10.1137/110823390.

[60] R. Henrion, B. S. Mordukhovich, N. M. Nam, *Second-Order Analysis of Polyhedral Systems in Finite and Infinite Dimensions with Applications to Robust Stability of Variational Inequalities*, SIAM Journal on Optimization 20 (5) (2010) 2199–2227, DOI: 10.1137/090766413.

[61] R. Henrion, J. V. Outrata, *Calmness of constraint systems with applications*, Mathematical Programming 104 (2) (2005) 437–464, DOI: 10.1007/s10107-005-0623-2.

[62] R. Henrion, W. Römisch, *On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling*, Applications of Mathematics 52 (6) (2007) 473–494, DOI: 10.1007/s10492-007-0028-z.

[63] R. Henrion, T. M. Surowiec, *On calmness conditions in convex bilevel programming*, Applicable Analysis 90 (6) (2011) 951–970, DOI: 10.1080/00036811.2010.495339.

[64] R. Herzog, C. Meyer, G. Wachsmuth, *B- and Strong Stationarity for Optimal Control of Static Plasticity with Hardening*, SIAM Journal on Optimization 23 (1) (2013) 321–352, DOI: 10.1137/110821147.

[65] F. Hiai, H. Umegaki, *Integrals, conditional expectations, and martingales of multivalued functions*, Journal of Multivariate Analysis 7 (1) (1977) 149 –182, DOI: 10.1016/0047-259X(77)90037-9.

[66] M. Hintermüller, I. Kopacka, *Mathematical Programs with Complementarity Constraints in Function Space: C- and Strong Stationarity and a Path-Following Algorithm*, SIAM Journal on Optimization 20 (2) (2009) 868–902, DOI: 10.1137/080720681.

[67] M. Hintermüller, B. S. Mordukhovich, T. M. Surowiec, *Several approaches for the derivation of stationarity conditions for elliptic MPECs with upper-level control constraints*, Mathematical Programming 146 (1) (2014) 555–582, DOI: 10.1007/s10107-013-0704-6.

[68] M. Hintermüller, T. M. Surowiec, *First-Order Optimality Conditions for Elliptic Mathematical Programs with Equilibrium Constraints via Variational Analysis*, SIAM Journal on Optimization 21 (4) (2011) 1561–1593, DOI: 10.1137/100802396.

[69] J.-B. Hiriart-Urruty, J. Malick, *A Fresh Variational-Analysis Look at the Positive Semidefinite Matrices World*, Journal of Optimization Theory and Applications 153 (3) (2012) 551–577, DOI: 10.1007/s10957-011-9980-6.

[70] T. Hoheisel, C. Kanzow, A. Schwartz, *Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints*, Mathematical Programming 137 (1) (2013) 257–288, DOI: 10.1007/s10107-011-0488-5.

[71] J. Jahn, *Introduction to the Theory of Nonlinear Optimization*, Springer, Berlin, 1996.

[72] J. Jahn, *Vector Optimization*, Springer, Berlin, 2004.

[73] J. Jarušek, J. V. Outrata, *On sharp necessary optimality conditions in control of contact problems with strings*, Nonlinear Analysis: Theory, Methods & Applications 67 (4) (2007) 1117 –1128, DOI: 10.1016/j.na.2006.05.021.

[74] V. Kalashnikov, F. Benita, P. Mehlitz, *The natural gas cash-out problem: A bilevel optimal control approach*, Math. Probl. Eng. (2015) 1–17, DOI: 10.1155/2015/286083.

[75] D. Kinderlehrer, G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, 1980.

[76] M. Knauer, C. Büskens, *Hybrid Solution Methods for Bilevel Optimal Control Problems with Time Dependent Coupling*, in: *Recent Advances in Optimization and its Applications in Engineering: The 14th Belgian-French-German Conference on Optimization*, edited by M. Diehl, F. Glineur, E. Jarlebring, W. Michiels, Springer, Berlin, 2010 237–246, DOI: 10.1007/978-3-642-12598-0_20.

[77] M. Knauer, C. Büskens, P. Lasch, *Real-Time Solution of Bi-Level Optimal Control Problems*, Proceedings in Applied Mathematics and Mechanics 5 (2005) 749–750, DOI: 10.1002/pamm.200510349.

[78] A. B. Levy, *Sensitivity of Solutions to Variational Inequalities on Banach Spaces*, SIAM Journal on Control and Optimization 38 (1) (1999) 50–60, DOI: 10.1137/S036301299833985X.

[79] Y.-C. Liang, X.-D. Zhu, G.-H. Lin, *Necessary Optimality Conditions for Mathematical Programs with Second-Order Cone Complementarity Constraints*, Set-Valued and Variational Analysis 22 (1) (2014) 59–78, DOI: 10.1007/s11228-013-0250-7.

[80] Z.-Q. Luo, J.-S. Pang, D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[81] M. Marcus, V. J. Mizel, *Absolute continuity on tracks and mappings of Sobolev spaces*, Archive for Rational Mechanics and Analysis 45 (1972) 294–320, DOI: 10.1007/BF00251378.

[82] M. Marcus, V. J. Mizel, *Nemitsky Operators on Sobolev Spaces*, Archive for Rational Mechanics and Analysis 51 (1973) 347–370, DOI: 10.1007/BF00263040.

[83] R. E. Megginson, *An Introduction to Banach Space Theory*, Graduate Texts in Mathematics, Springer, New York, 1998.

[84] P. Mehlitz, *Bilevel programming problems with simple convex lower level*, Optimization 65 (6) (2016) 1203–1227, DOI: 10.1080/02331934.2015.1122006.

[85] P. Mehlitz, *Necessary optimality conditions for a special class of bilevel programming problems with unique lower level solution*, Optimization (2017) 1–30, DOI: 10.1080/02331934.2017.1349123.

[86] P. Mehlitz, G. Wachsmuth, *On the Limiting Normal Cone to Pointwise Defined Sets in Lebesgue Spaces*, Set-Valued and Variational Analysis (2016), DOI: 10.1007/s11228-016-0393-4.

[87] P. Mehlitz, G. Wachsmuth, *Weak and strong stationarity in generalized bilevel programming and bilevel optimal control*, Optimization 65 (5) (2016) 907–935, DOI: 10.1080/02331934.2015.1122007.

[88] F. Mignot, *Contrôle dans les inéquations variationelles elliptiques*, Journal of Functional Analysis 22 (2) (1976) 130 –185, DOI: 10.1016/0022-1236(76)90017-3.

[89] K. Mombaur, A. Truong, J.-P. Laumond, *From human to humanoid locomotion—an inverse optimal control approach*, Autonomous Robots 28 (3) (2010) 369–383, DOI: 10.1007/s10514-009-9170-7.

[90] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, Springer, Berlin, 2006.

[91] B. S. Mordukhovich, N. M. Nam, *Variational Stability and Marginal Functions via Generalized Differentiation*, Mathematics of Operations Research 30 (4) (2005) 800–816, DOI: 10.1287/moor.1050.0147.

[92] B. S. Mordukhovich, N. M. Nam, H. M. Phan, *Variational Analysis of Marginal Functions with Applications to Bilevel Programming*, Journal of Optimization Theory and Applications 152 (3) (2012) 557–586, DOI: 10.1007/s10957-011-9940-1.

[93] B. S. Mordukhovich, N. M. Nam, N. D. Yen, *Subgradients of marginal functions in parametric mathematical programming*, Mathematical Programming 116 (1) (2009) 369–396, DOI: 10.1007/s10107-007-0120-x.

[94] B. S. Mordukhovich, N. Sagara, *Subdifferentials of nonconvex integral functionals in Banach spaces with applications to stochastic dynamic programming*, to appear in J. Convex Anal. (2016), URL: http://arxiv.org/abs/1508.02239.

[95] C. Olech, *The Lyapunov Theorem: Its extensions and applications*, in: *Methods of Nonconvex Analysis*, edited by A. Cellina, volume 1446, Lecture Notes in Mathematics, Springer, Berlin, 1990 84–103, DOI: 10.1007/BFb0084932.

[96] J. Outrata, J. Jarušek, J. Stará, *On Optimality Conditions in Control of Elliptic Variational Inequalities*, Set-Valued and Variational Analysis 19 (1) (2011) 23–42, DOI: 10.1007/s11228-010-0158-4.

[97] J. V. Outrata, *On Optimization Problems with Variational Inequality Constraints*, SIAM Journal on Optimization 4 (2) (1994) 340–357, DOI: 10.1137/0804019.

[98] J. V. Outrata, *On the numerical solution of a class of Stackelberg problems*, Zeitschrift für Operations Research 34 (4) (1990) 255–277, DOI: 10.1007/BF01416737.

[99] J.-S. Pang, D. Sun, J. Sun, *Semismooth Homeomorphisms and Strong Stability of Semidefinite and Lorentz Complementarity Problems*, Math. Oper. Res. 28 (1) (2003) 39–63, DOI: 10.1287/moor.28.1.39.14258.

[100] N. S. Papageorgiou, S. T. Kyritsi-Yiallourou, *Handbook of applied analysis*, volume 19, Advances in Mechanics and Mathematics, Springer, New York, 2009, DOI: 10.1007/b120946.

[101] M. Pilecka, *Set-valued optimization and its application to bilevel optimization*, PhD thesis, Technische Universität Bergakademie Freiberg, 2016.

[102] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Interscience, New York, 1962.

[103] U. Prüfert, *Solving optimal PDE control problems. Optimality conditions, algorithms and model reduction*, habilitation, Technische Universität Bergakademie Freiberg, 2016.

[104] D. Ralph, S. Dempe, *Directional derivatives of the solution of a parametric nonlinear program*, Mathematical Programming 70 (1) (1995) 159–172, DOI: 10.1007/BF01585934.

[105] S. M. Robinson, *Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems*, SIAM Journal on Numerical Analysis 13 (4) (1976) 497–513, DOI: 10.1137/0713043.

[106] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[107] R. T. Rockafellar, *Integrals which are convex functionals*. Pacific J. Math. 24 (3) (1968) 525–539.

[108] R. T. Rockafellar, R. J.-B. Wets, *Variational Analysis*, volume 317, Grundlehren der mathematischen Wissenschaften, Springer, Berlin, 1998.

[109] A. Ruziyeva, *Fuzzy Bilevel Programming*, PhD thesis, Technische Universität Bergakademie Freiberg, 2013.

[110] H. H. Schaefer, *Banach lattices and positive operators*, Die Grundlehren der mathematischen Wissenschaft, Band 215, Springer, Berlin, 1974.

[111] S. Scheel, S. Scholtes, *Mathematical Programs with Complementarity Constraints: Stationarity, Optimality, and Sensitivity*, Mathematics of Operations Research 25 (1) (2000) 1–22, DOI: 10.1287/moor.25.1.1.15213.

[112] A. Shapiro, *On concepts of directional differentiability*, Journal of Optimization Theory and Applications 66 (3) (1990) 477–487, DOI: 10.1007/BF00940933.

[113] K. Shimizu, Y. Ishizuka, J. F. Bard, *Nondifferentiable and two-level mathematical programming*, Kluwer Academic, Dordrecht, 1997.

[114] M. Simonnet, *Measures and Probabilities*, Springer, New York, 1996.

[115] H. v. Stackelberg, *Marktform und Gleichgewicht*, Springer, Berlin, 1934.

[116] T. M. Surowiec, *Explicit Stationarity Conditions and Solution Characterization for Equilibrium Problems with Equilibrium Constraints*, PhD thesis, Humboldt-Universität zu Berlin, 2009.

[117] N. T. Toan, B. T. Kien, *Subgradients of the Value Function to a Parametric Optimal Control Problem*, Set-Valued and Variational Analysis 18 (2) (2010) 183–203, DOI: 10.1007/s11228-009-0125-0.

[118] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen*, Vieweg, Wiesbaden, 2009.

[119] P. Tseng, *Merit functions for semi-definite complementarity problems*, Mathematical Programming 83 (1) (1998) 159–185, DOI: 10.1007/BF02680556.

[120] G. Wachsmuth, *A guided tour of polyhedric sets*, Preprint TU Chemnitz (2016) 1–39, URL: https://www.tu-chemnitz.de/mathematik/part_dgl/publications/Wachsmuth__A_guided_tour_of_polyhedric_sets.pdf.

[121] G. Wachsmuth, *Mathematical Programs with Complementarity Constraints in Banach Spaces*, Journal on Optimization Theory and Applications 166 (2015) 480–507, DOI: 10.1007/s10957-014-0695-3.

[122] G. Wachsmuth, *Pointwise Constraints in Vector-Valued Sobolev Spaces*, Applied Mathematics & Optimization (2016) 1–28, DOI: 10.1007/s00245-016-9381-1.

[123] G. Wachsmuth, *Strong Stationarity for Optimal Control of the Obstacle Problem with Control Constraints*, SIAM Journal on Optimization 24 (4) (2014) 1914–1932, DOI: 10.1137/130925827.

[124] G. Wachsmuth, *Strong stationarity for optimization problems with complementarity constraints in absence of polyhedricity*, Set-Valued and Variational Analysis 25 (1) (2016) 133–175, DOI: 10.1007/s11228-016-0370-y.

[125] G. Wachsmuth, *Towards M-Stationarity for Optimal Control of the Obstacle Problem with Control Constraints*, SIAM Journal on Control and Optimization 54 (2) (2016) 964–986, DOI: 10.1137/140980582.

[126] D. Werner, *Funktionalanalysis*, Springer, Berlin, 1995.

[127] J. Wu, L. Zhang, Y. Zhang, *Mathematical Programs with Semidefinite Cone Complementarity Constraints: Constraint Qualifications and Optimality Conditions*, Set-Valued and Variational Analysis (22) (2014) 155–187, DOI: 10.1007/s11228-013-0242-7.

[128] J. J. Ye, *Constraint Qualifications and Necessary Optimality Conditions for Optimization Problems with Variational Inequality Constraints*, SIAM Journal on Optimization 10 (4) (2000) 943–962, DOI: 10.1137/S105262349834847X.

[129] J. J. Ye, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, Math. Anal. Appl. 307 (2005) 350–369, DOI: 10.1016/j.jmaa.2004.10.032.

[130] J. J. Ye, *Necessary Conditions for Bilevel Dynamic Optimization Problems*, SIAM Journal on Control and Optimization 33 (4) (1995) 1208–1223, DOI: 10.1137/S0363012993249717.

[131] J. J. Ye, *Optimal Strategies For Bilevel Dynamic Problems*, SIAM Journal on Control and Optimization 35 (2) (1997) 512–531, DOI: 10.1137/S0363012993256150.

[132] J. J. Ye, *Optimality Conditions for Optimization Problems with Complementarity Constraints*, SIAM Journal on Optimization 9 (2) (1999) 374–387, DOI: 10.1137/S1052623497321882.

[133] J. J. Ye, X. Y. Ye, *Necessary Optimality Conditions for Optimization Problems with Variational Inequality Constraints*, Mathematics of Operations Research 22 (4) (1997) 977–997, DOI: 10.1287/moor.22.4.977.

[134] J. J. Ye, J. Zhou, *Exact formulas for the proximal/regular/limiting normal cone of the second-order cone complementarity set*, Mathematical Programming 162 (1) (2017) 33–50, DOI: 10.1007/s10107-016-1027-1.

[135] J. J. Ye, J. Zhou, *First-Order Optimality Conditions for Mathematical Programs with Second-Order Cone Complementarity Constraints*, SIAM Journal on Optimization 26 (4) (2016) 2820–2846, DOI: 10.1137/16M1055554.

[136] J. J. Ye, D. L. Zhu, *A note on optimality conditions for bilevel programming problems*, Optimization 39 (4) (1997) 361–366, DOI: 10.1080/02331939708844290.

[137] J. J. Ye, D. L. Zhu, *New Necessary Optimality Conditions for Bilevel Programs by Combining the MPEC and Value Function Approaches*, SIAM Journal on Optimization 20 (4) (2010) 1885–1905, DOI: 10.1137/080725088.

[138] J. J. Ye, D. L. Zhu, *Optimality conditions for bilevel programming problems*, Optimization 33 (1) (1995) 9–27, DOI: 10.1080/02331939508844060.

[139] J. J. Ye, D. L. Zhu, Q. J. Zhu, *Exact Penalization and Necessary Optimality Conditions for Generalized Bilevel Programming Problems*, SIAM Journal on Optimization 7 (2) (1997) 481–507, DOI: 10.1137/S1052623493257344.

[140] A. Zemkoho, *Bilevel Programming: Reformulations, Regularity, and Stationarity*, PhD thesis, Technische Universität Bergakademie Freiberg, 2012.

[141] F. Zhang, *The Schur Complement and its Applications*, volume 4, Numerical Methods and Algorithms, Springer, New York, 2005.

[142] V. A. Zorich, *Analysis II*, Springer, Berlin, 2007.

[143] J. Zowe, S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Applied Mathematics and Optimization 5 (1) (1979) 49–62, DOI: 10.1007/BF01442543.