# Equilibrium Problems with Equilibrium Constraints in Banach Spaces: Stationarity Concepts, Applications and Algorithms

Vom Fachbereich Mathematik der Technischen Universität Darmstadt zur Erlangung des Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)
Genehmigte Dissertation von Jan Becker aus Aschaffenburg
Tag der Einreichung: 09.12.2020, Tag der Prüfung: 16.03.2021

Referentin: Prof. Dr. Alexandra Schwartz
Korreferent: Prof. Dr. Thomas M. Surowiec
Darmstadt – D17

TECHNISCHE UNIVERSITÄT DARMSTADT

Mathematics Department
Optimization
Graduate School of Computational Engineering

Equilibrium Problems with Equilibrium Constraints in Banach Spaces: Stationarity Concepts, Applications and Algorithms

Accepted doctoral thesis by Jan Becker

Referentin: Prof. Dr. Alexandra Schwartz
Korreferent: Prof. Dr. Thomas M. Surowiec

Date of submission: 09.12.2020
Date of thesis defense: 16.03.2021

Darmstadt – D17

Für Amélie.

# Acknowledgment

First and foremost, I would like to thank my supervisor Prof. Dr. Alexandra Schwartz for giving me the opportunity to work in her research group and the guidance during my Ph.d. studies. Although her research topic is located in finite dimensions, she always could point me in the right direction and kept me on track.
Next, I would like to thank Prof. Dr. Thomas M. Surowiec for acting as a referee of this thesis.
I am thankful to Prof. Dr. Stefan Ulbrich not only for being my co-supervisor at the Graduate School of Computational Engineering and being part in my examination committee, but also for the nice and funny Schafkopf events together with Anna, Paloma and Krissy. It was always a pleasure to revive a little bit of Bavarian tradition in Darmstadt.
Furthermore, I thank Prof. Dr. Jens Lang for being also part in my examination committee.
I am grateful to Dr. Adrian Sichau for his supporting help during my master's thesis and encouraging me to apply for a Ph.d. at Dolivostraße.
I am thankful to my friends Marco, Max, Philipp and Yannik, who chased all my errors and wrong formulations that I intentionally placed throughout this thesis. I am pretty sure that they discovered all of them. At this point, special thanks go to my friend and office colleague Daniel, who not only proofread parts of my work, but also had the time to answer all of my bureaucratic questions around the Ph.d..
Moreover, I would like to thank my colleagues and friends at the Research Group Optimization. It was always fun and I enjoyed the common time.
Without the support of my parents and my sister, this would not have been possible. Thank you for believing in me and giving me the best possible chances to succeed. Last but not least, I would like to thank Annika and my daughter Amélie. Thank you for being here, making my life so much easier and more beautiful and helping me when it was needed the most.
This work is supported by the DFG through DFG SPP 1962 Priority Programme "Non-Smooth and Complementarity-Based Distributed Parameter Systems: Simulation and Hierarchical Optimization" within project 21 and by the Graduate School of Computational Engineering at Technische Universität Darmstadt.

# Abstract

In mathematics, the field of non-cooperative game theory models the competition between several parties, which are called players. Therein, each player tries to reach an individual goal, which is described by an optimization problem. However and in contrast to classical nonlinear programming, there exists a dependency between the players, i.e. the choice of a suitable strategy influences the behavior and the reward of the player's opponents and vice versa. For this reason, a popular solution concept is given by Nash equilibria, which were introduced by John Forbes Nash in his Ph.D. thesis in 1950. In order to prove the existence of a Nash equilibrium, the convexity of the underlying optimization problem is a central requirement. However, this assumption does not hold in general.

This thesis is devoted to special equilibrium problems in Banach spaces, which can be described by equilibrium problems with equilibrium/complementarity constraints (EPEC/EPCC). Due to the structure of the underlying feasible set, those games are non-convex. Motivated by known results with respect to mathematical programs with complementarity constraints, we focus on weaker Nash equilibrium concepts, which can at least be seen as necessary conditions for a Nash equilibrium under suitable assumptions.

In the first part of this work, we concentrate on multi-leader multi-follower games, where the participating players are divided hierarchically into leaders and followers, which compete on their particular level with each other. Under suitable assumptions, the solution of the lower level is described by its necessary and sufficient first-order optimality system and can be written as an EPCC. In this context, we first analyze the latter problem in abstract Banach spaces and afterwards, consider the special case of a multi-leader single-follower game (MLFG), where the lower level is given by a quadratic problem in a Hilbert space. For the latter one, we show on the basis of two known penalization techniques that there exist sequences of auxiliary equilibrium problems, which approximate the corresponding EPCC. In the following application that extends known contributions on an optimal control framework of the obstacle problem, we use these auxiliary games and show that both generate sequences, which converge at least to an $\epsilon$-almost C-stationary Nash equilibrium of the original MLFG. The results are analyzed numerically on the basis of a Gauß-Seidel-type algorithm and are tested with respect to two examples.

The second part is motivated by the work [14] "A generalized Nash equilibrium approach

for optimal control problems of autonomous cars" by Axel Dreves and Matthias Gerdts, where a traffic scenario between several intelligent cars is modeled by a dynamic equilibrium problem. Due to the collision avoidance constraint, this game is non-convex. However, we show that it can be written as a generalized Nash equilibrium problem with mixed-integer variables (MINEP), which again is equivalent to an EPCC. In contrast to the first application, we now concentrate on problems in Lebesgue spaces. In the following, we compare known results from abstract Banach spaces and the corresponding ones in Lebesgue spaces. In particular, we show that for general MINEPs all weak Nash equilibrium concepts coincide. Based on these observations, we apply the results to the traffic scenario. In this context, we again use a penalization technique and deduce by the generated sequence of Nash equilibrium problems that we find a sequence of equilibrium points, which converge to an S-stationary Nash equilibrium of MINEP. We end up with a numerical analysis and test the results with two hypothetical traffic scenarios.

# Zusammenfassung

Als Teilgebiet der Mathematik modelliert die nicht-kooperative Spieltheorie den Wettbewerb zwischen mehreren Parteien, genannt Spieler. Diese verfolgen darin individuelle Ziele, welche durch die Minimierung einer Zielfunktion über einer zulässigen Menge dargestellt werden. Im Unterschied zu klassischen Optimierungsproblemen, existiert jedoch eine Abhängigkeit zwischen den Spielern, d.h. die Wahl einer eigenen Strategie hat direkte Auswirkungen auf das Verhalten sowie den Ertrag der Gegner und umgekehrt. Ein mögliches Lösungskonzept ist durch Nash-Gleichgewichte gegeben, welche durch den Namensgeber, John Nash, 1950 in seiner Dissertation untersucht worden sind.

Eine zentrale Voraussetzung für die Existenz von Nash-Gleichgewichten ist die Konvexität der einzelnen Optimierungsprobleme, welche jedoch in der Realität nicht immer gegeben ist. Die eingereichte Dissertation untersucht in diesem Zusammenhang spezielle Spiele in Banachräumen, welche durch Gleichgewichtsprobleme mit Gleichgewichtsbeschränkungen (engl. equilibrium problems with equilibrium constraints, EPEC) bzw. Gleichgewichtsprobleme mit Komplementaritätsrestriktionen (engl. equilibrium problems with complementarity constraints, EPCC) beschrieben werden können. Durch die namensgebenden Restriktionen an die zulässige Menge definieren diese nicht-konvexe Probleme in Funktionenräumen. Motiviert durch bekannte Resultate bezüglich mathematischer Programme mit Komplementaritätsnebenbedinungen, weichen wir deshalb auf schwächere Gleichgewichtskonzepte aus, welche unter gewissen Annahmen zumindest notwendige Bedingungen für ein Nash-Gleichgewicht liefern.

Der erste Teil der Arbeit beschäftigt sich mit sogenannten Multi-Leader Multi-Follower Spielen (engl. multi-leader multi-follower games, MLMFG), in denen die beteiligten Spieler hierarchisch in Leader und Follower klassifiziert werden und auf beiden Ebenen untereinader konkurrieren. Unter geeigneten Annahmen lässt sich darin die Lösung auf der unteren Stufe durch notwendige und hinreichende Optimalitätsbedingungen erster Ordnung charakterisieren und in ein Gleichgewichtsproblem mit Komplementaritätsrestriktionen überführen. Im weiteren Verlauf untersuchen wir zunächst das daraus resultierende EPCC in abstrakten Banachräumen und betrachten im Anschluss den Spezialfall eines Multi-Leader Single-Follower Spiels, in dem der Follower ein quadratisches Optimierungsproblem in einem Hilbertraum löst. Für dieses konstruieren wir, anhand zweier Penalisierungstechniken,

Folgen vereinfachter Gleichgewichtsprobleme und zeigen, dass diese das dazugehörige EPCC approximieren. Im daran anschließenden Anwendungsbeispiel, welches auf Arbeiten über die optimale Steuerung des Hindernisproblems in Sobolevräumen aufbaut, nutzen wir diese Hilfsprobleme als Grundlage, um Folgen von Nash-Gleichgewichten zu erzeugen. Für Letztere können wir in beiden Fällen nachweisen, dass sie unter geeigneten Annahmen mindestens gegen ein $\epsilon$-almost C-stationäres Nash-Gleichgewicht des ursprünglichen hierarchischen Spiels konvergieren. Die dabei erhaltenen Ergebnisse analysieren wir im Anschluss numerisch und nutzen Gauß-Seidel-artige Algorithmen um diese anhand ausgewählter Beispiele zu testen.

Der zweite Teil der Dissertation ist maßgeblich durch die Arbeit [14] "A generalized Nash equilibrium approach for optimal control problems of autonomous cars" von Axel Dreves und Mathias Gerdts motiviert, in dem ein Verkehrsszenario zwischen intelligenten Fahrzeugen durch ein dynamisches Gleichgewichtsproblem modelliert wird. Durch die auftretenden Kollisionsvermeidungs-Bedingungen, ist dieses ein nicht-konvexes Spiel, welches durch die spezielle Struktur der erwähnten Restriktion als ein gemischt-ganzzahliges Gleichgewichtsproblem (engl. generalized Nash equilibrium problem with mixed-integer variables, MINEP) beschrieben werden kann. Da die geforderte Ganzzahligkeit äquivalent zu einer Komplementaritätsbedingung ist, erhalten wir auch hier ein EPCC mit dem Unterschied, dass die zugrundeliegenden Komplementaritätsrestriktionen Teil eines Lebesgueraumes sind. Im Anschluss vergleichen wir zunächst die Resultate bezüglich abstrakter Banachräume mit den bereits bekannten Ergebnissen in Lebesgueräumen. Insbesondere zeigen wir, dass in diesem Szenario alle schwachen Gleichgewichtskonzepte identisch sind. Darauf aufbauend werden die Ergebnisse auf das eingangs beschriebene Verkehrszenario übertragen. In diesem Zusammenhang erzeugen wir analog zum ersten Teil durch einen Penalisierungsansatz eine Folge von Gleichgewichtsproblemen und weisen nach, dass die dadurch konstruierte Folge von Nash-Gleichgewichten unter zusätzlichen Voraussetzungen gegen ein S-stationäres Nash-Gleichgewicht von MINEP konvergiert. Die erzielten Resultate werden abschließend numerisch analysiert, implementiert und mittels fiktiver Verkehrssituationen getestet.

# Contents

# Abbreviations

| | |
|---|---|
| iff | if and only if |
| f.a.a. | for almost all |
| a.e. | almost everywhere |
| q.e. | quasi everywhere |
| w.r.t. | with respect to |
| i.e. | that is (id est) |
| e.g. | for example (exempli gratia) |
| w.l.o.g. | without loss of generality |
| (w)lsc | (weakly) lower semicontinuous |
| KRZCQ | Kurcyusz Robinson Zowe constraint qualification |
| FRCQ | Full Range constraint qualification |
| KKT | Karush-Kuhn-Tucker conditions |
| MPEC | mathematical problem with equilibrium constraints |
| MPCC | mathematical problem with complementarity constraints |
| RNLP | relaxed nonlinear program |
| TNLP | tightened nonlinear program |
| BOP | bilevel optimization problem |
| NEP | Nash equilibrium problem |
| GNEP | generalized Nash equilibrium problem |
| MLMFG | multi-leader multi-follower game |
| MLFG | multi-leader single-follower game |
| EPEC | equilibrium problem with equilibrium constraints |
| EPCC | equilibrium problem with complementarity constraints |
| MLOCP | multi-leader optimal control problem |
| MINEP | generalized Nash equilibrium problem with mixed-integer variables |
| NE | Nash equilibrium |
| NNE | normalized Nash equilibrium |
| VI | variational inequality |
| ODE | ordinary differential equation |
| PDE | partial differential equation |

# Notation

**Banach space theory:**

| | |
|---|---|
| $X$ | a Banach space |
| $L(X,Y)$ | normed space of linear operators between Banach spaces $X$ and $Y$ |
| $I_{XY}$ | identity operator from $X$ to $Y \subseteq X$ |
| $I_X$ | identity operator from $X$ to $X$ |
| $X^*$ | dual space of $X$ |
| $\|\cdot\|_X$ | norm of Banach space $X$ |
| $\langle\cdot,\cdot\rangle_X$ | dual pairing of Banach space $X$ |
| $(\cdot,\cdot)_H$ | scalar product in Hilbert space $H$ |
| $\mathbb{B}_\epsilon(x)$ | ball around $x$ with radius $\epsilon > 0$ |
| $\{x^k\} \subseteq X$ | sequence in $X$ |
| $x^k \to \bar{x}$ | strong convergence of sequence in $X$ |
| $x^k \rightharpoonup \bar{x}$ | weak convergence of sequence in $X$ |
| $\phi^k \rightharpoonup^* \bar{\phi}$ | weak-star convergence of sequence in $X^*$ |
| $X \hookrightarrow Y$ | continuous embedding between $X$ and $Y$ |
| $X \hookrightarrow\hookrightarrow Y$ | compact embedding between $X$ and $Y$ |
| $C \subseteq X$ | subset of $X$ |
| lin $C$ | linear hull of set $C$ |
| cl $C$ | closure of set $C$ |
| cone $C$ | conic hull of set $C$ |
| int $C$ | interior of set $C$ |
| $\partial C$ | boundary of set $C$ |
| $T^*$ | adjoint operator of operator $T$ |
| $T'(\bar{x})$ | (Fréchet) derivative of mapping $T$ at $\bar{x}$ |
| gph $F$ | graph of set-valued mapping $F$ |
| | |
| $\mathbb{N}$ | natural numbers |
| $\mathbb{N}_0$ | $\mathbb{N} \cup \{0\}$ |

| | |
|---|---|
| $\mathbb{N}_n$ | first $n$ natural numbers |
| $\mathbb{R}$ | real numbers |
| $\mathbb{R}^{n,+}$ | positive real numbers in $\mathbb{R}^n$ |
| $\mathbb{R}_0^{n,+}$ | nonnegative real numbers in $\mathbb{R}^n$ |
| $\Omega$ | domain in $\mathbb{R}^d$ with $d \in \mathbb{N}$ |
| $\bar{\Omega}$ | closure of $\Omega$ |
| $\omega$ | element in $\Omega \subseteq \mathbb{R}^d$ with $d > 1$ |
| $\tau$ | element in $\Omega \subseteq \mathbb{R}$ |
| $L^p(\Omega)$ | Lebesgue space of $p$-integrable functions $y : \Omega \to \mathbb{R}$ |
| $C_0^\infty(\Omega)$ | space of infinitely differentiable functions with compact support |
| $W^{k,p}(\Omega)$ | Sobolev space of functions $y : \Omega \to \mathbb{R}$, whose weak derivatives of order $k$ exist and are in $L^p(\Omega)$ |
| $W_0^{k,p}(\Omega)$ | closure of $C_0^\infty(\Omega)$ in $W^{k,p}(\Omega)$ |
| $H^k(\Omega)$ | $W^{k,2}(\Omega)$ |
| $H^{-1}(\Omega)$ | dual space of $H_0^1(\Omega)$ |
| $H^{-1}(\Omega)_+$ | dual cone of $H_0^1(\Omega)_0^+$ |
| $H^{-1}(\Omega)_-$ | polar cone of $H_0^1(\Omega)_0^+$ |
| $AC^{1,p}(\Omega)$ | space of absolutely continuous functions with weak derivatives in $L^p(\Omega)$ and $\Omega \subseteq \mathbb{R}$ |
| $C(\bar{\Omega})$ | space of continuous functions |
| $\text{cap}(\Omega_0)$ | capacity of $\Omega_0 \subseteq \Omega$ |
| $\mathcal{X}_{\Omega_0}$ | characteristic function w.r.t. subset $\Omega_0$ |
| $\mathcal{Q}_A$ | active set |
| $\mathcal{Q}_A^+$ | strongly/strictly active set |
| $\mathcal{Q}_I$ | inactive set |
| $\mathcal{Q}_A^0$ | biactive set |

**Variational analysis:**

| | |
|---|---|
| $\emptyset$ | empty set |
| $\mathcal{R}_C(\bar{x})$ | radial cone of $C$ at $\bar{x}$ |
| $\mathcal{T}_C(\bar{x})$ | Bouligand tangent cone of $C$ at $\bar{x}$ |
| $\mathcal{T}_C^c(\bar{x})$ | Clarke tangent cone of $C$ at $\bar{x}$ |
| $\hat{\mathcal{N}}_C(\bar{x})$ | Fréchet normal cone of $C$ at $\bar{x}$ |
| $\mathcal{N}_C(\bar{x})$ | limiting normal cone of $C$ at $\bar{x}$ |
| $\mathcal{N}_C^c(\bar{x})$ | Clarke normal cone of $C$ at $\bar{x}$ |

| | |
|---|---|
| $C^\circ$ | polar cone of $C$ |
| $C^+$ | dual cone of $C$ |
| $C^\perp$ | annihilator of $C$ |
| $\mathcal{K}_C(\bar{x}, \bar{\eta})$ | critical cone to $C$ w.r.t. $(\bar{x}, \bar{\eta})$ |
| $\leq_C$ | partial order induced by the closed, convex and pointed cone $C$ |
| $x_+$ | positive part of $x$ |
| $x_-$ | negative part of $x$ |

**Optimization and equilibrium problems:**

| | |
|---|---|
| $X_{\text{ad}}$ | feasible/admissible set of minimization/equilibrium problem |
| $\mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})$ | linearized tangent cone of $X_{\text{ad}}$ at $\bar{x}$ |
| $\Pi_{X_{\text{ad}}}$ | projection mapping onto $X_{\text{ad}}$ |
| $\Lambda(\bar{x})$ | set of Lagrange multipliers w.r.t. $\bar{x}$ |
| $\Lambda_0(\bar{x})$ | set of singular Lagrange multipliers w.r.t. $\bar{x}$ |
| $\mathcal{L}$ | Lagrangian of a minimization problem |
| | |
| $\Gamma$ | (generalized) Nash equilibrium problem |
| $\Gamma_{\text{MLFG}}$ | multi-leader-follower game |
| $\Gamma_{\text{MINEP}}$ | generalized Nash equilibrium problem with mixed-integer variables |
| $\Gamma_{\text{EPEC}}$ | equilibrium problem with equilibrium constraints |
| $\Gamma_{\text{EPCC}}$ | equilibrium problem with complementarity constraints |
| $\Gamma^{\text{lq}}_{\text{MLFG}}$ | multi-leader-follower game with linear-quadratic lower level |
| $\Gamma^{\text{lq}}_{\text{EPEC}}$ | corresponding EPEC to $\Gamma^{\text{lq}}_{\text{MLFG}}$ |
| $\Gamma^{\text{lq}}_{\text{EPCC}}$ | corresponding EPCC to $\Gamma^{\text{lq}}_{\text{MLFG}}$ |
| $\Gamma_{\text{AD}}$ | generalized Nash equilibrium problem of autonomous driving scenario |
| $\Gamma^{\text{AD}}_{\text{MINEP}}$ | generalized Nash equilibrium problem with mixed-integer variables of autonomous driving scenario |
| $N$ | number of players/leaders in GNEP/MLMFG |
| $M$ | number of followers in MLMFG |
| $f^\nu$ | objective functional of player/leader $\nu$ ($f^\nu \equiv f$ if $N = 1$) |
| $x^\nu, x^{-\nu}$ | strategy of player/leader $\nu$ and corresponding opponent vector |
| $\theta^i$ | objective functional of follower $i$ ($\theta^1 \equiv \theta$ if $M = 1$) |
| $y^i, y^{-i}$ | strategy of follower $i$ and corresponding |

| | |
|---|---|
| | opponent vector |
| $D_y h$ | partial derivative of mapping $h : X \times Y \to Z$ w.r.t. $y$ |
| $D^2_{yx} h$ | second order partial derivative of $h$ w.r.t. $y$ and $x$ |
| $\Psi : X \rightrightarrows Y$ | solution set of lower level equilibrium/minimization problem |
| $\dot{\theta}$ | time derivative of function $\theta(t, x)$ |
| $\nabla\theta$ | gradient of $\theta(t, x)$ |

# 1. Introduction

In a world, which is more and more connected, interactions between several parties appear more frequently. In general, these scenarios can be distinguished between a cooperative and a non-cooperative behavior, where in the former case all participants work together in order to achieve the maximum outcome. In contrast, the focus of the latter one lies on the maximum individual reward of each participant, regardless of his opponents. While the cooperative setting is modeled mathematically in most cases[1] by a multi-objective optimization framework, the non-cooperative version leads to a game theoretical approach. In mathematics, the field of *game theory* uses concepts of nonlinear programming to model competitions, which are called *games*. In this context, a game is comprised of a finite number of parties that we denote *players*. Therein, each player tries to optimize his[2] objective function with respect to his feasible *strategies*. Now, each optimization problem is solved simultaneously, where the major difference to classical nonlinear programming is the coupling of each player's problem with his opponents' strategies. So in particular, the outcome and the feasible strategy can be influenced by the choice of the others. Although there exist some contributions written in the 19th century, the beginning of game theory can be dated to 1928, when John von Neumann published his "Zur Theorie der Gesellschaftsspiele" [67], which focuses on the non-cooperative behavior in board games. However, it took more than 20 years until John Forbes Nash introduced a general solution concept [65] of *equilibrium problems*, which was later called a *Nash equilibrium* and verified its existence in [66] under suitable assumptions. In general, a Nash equilibrium point describes the state of a game, where no player has the incentive of an unilateral change of his decision. For this reason, a Nash equilibrium can be seen as a stable point between all participating parties.

Nowadays, the area of real-world applications is widespread. Beginning with economical examples discussed for instance in the contributions [24, 31, 37, 84], where the competition in a spot-market scenario for electricity or gas is considered, we also find models that focused on autonomous driving [14], airfoil shape optimization of an airplane [87] or the behavior of predators and preys in a closed ecosystem [72]. However, the existence of

[1] There also exist game theoretic approaches in order to analyze cooperative behavior.

[2] This formulation is only chosen for convenience and does not indicate the gender of the player.

Nash equilibria is not always satisfied, which can in general be traced back to the lack of convexity of the underlying problems.
This thesis is devoted to *equilibrium problems with equilibrium constraints* (EPEC) in function spaces that denote a special class of games. In particular, these problems are described by a special structure of each player's feasible set, which is typically non-convex. We refer to the monographs [9, 60, 84, 85] for contributions in finite dimensions and to [61, 63] for publications in function spaces. At this point, we emphasize [61], where the author analyzed EPECs in abstract function spaces with respect to variational analysis and a generalized derivative concept. In contrast, our contribution is based on the work by Gerd Wachsmuth and Patrick Mehlitz [56, 94], who considered *mathematical programs with complementarity constraints* (MPCC) in abstract Banach spaces. Due to the non-convexity, we focus on weaker Nash equilibrium concepts. Therefore and motivated by the corresponding contributions for MPCCs (see e.g. [56, 82]), we extend the definitions of weak (W-), Clarke (C-), Mordukhovich (M-) and strong (S-) stationary points to equilibrium problems in the particular Banach spaces and focus on special types of EPECs.
Now, the structure of this work is as follows. In Chapter 2, we introduce all necessary basics that are needed for the subsequent considerations. In this context, we start with a fundamental overview of Banach space theory in Section 2.1, where we focus in Subsection 2.1.1 on basic properties of abstract Banach spaces and consider the underlying topology. Afterwards, Subsection 2.1.2 is dedicated to special Banach spaces. After that we concentrate on basic concepts of variational analysis (Section 2.2) that are used for minimization/equilibrium problems, which are introduced in Section 2.3. Therein, we focus on standard minimization problems in arbitrary Banach spaces (Subsection 2.3.1) and recall necessary and sufficient first-order optimality conditions under suitable assumptions. Chapter 2 ends up with Subsection 2.3.2 that extends the concepts of Subsection 2.3.1 and introduces basic definitions of game theory. We characterize several equilibrium problems and list the corresponding equilibria. In particular, we give an overview of the current existence results and emphasize the importance of convexity in this context.
Chapter 3 is dedicated to equilibrium problems with equilibrium/complementarity constraints in arbitrary Banach spaces. Due to its non-convex structure, concepts derived in Subsection 2.3.2 are rarely applicable. For this reason, we first introduce the problem of interest in Section 3.1 and afterwards focus on mathematical programs with complementarity constraints in Section 3.2. Therein, we recall the basic concepts of strongly [95], Mordukhovich [19] and weakly stationary points [56] in abstract Banach spaces, which can be seen as weaker forms of the well-known KKT-conditions. These results are applied and extended to general EPCCs in Section 3.3, where we introduce W-, M- and S-stationary Nash equilibria and additionally consider two auxiliary problems. The latter ones are used as a starting point for the further analysis in regard of finding such Nash

equilibria.
After considering EPECs/EPCCs, we move forward and concentrate on multi-leader multi-follower games in abstract Banach spaces in Chapter 4. Therein, we introduce these hierarchical equilibrium problems in Section 4.1 and define the basic setting. By using the necessary and sufficient first-order optimality condition of the lower level, we derive the corresponding EPEC, define tailored strongly, Mordukhovich and weakly stationary Nash equilibria and end up with necessary conditions for leader-follower Nash equilibria. In Section 4.2 we concentrate on a special multi-leader single-follower game, where the lower level problem is given by a quadratic program with linear constraints in a Hilbert space. In the following, it is shown that the resulting EPCC can be approximated by two auxiliary equilibrium problems under suitable assumptions. This is done by the generalization of two penalization techniques, known from the analysis of MPCCs/MPECs in a special Sobolev space (see [29, 31]). For this reason, we consider in Subsection 4.2.1 the approach, motivated by [29], where the lower level is analyzed via its corresponding variational inequality. In contrast, Subsection 4.2.2 is inspired by [31]. In this context, we utilize the KKT-system of the lower level and only penalize the crucial part, i.e. the complementarity condition.
Chapter 5 can be interpreted as an application of the previous one. Therein, we extend the well-known problem of an optimal control framework of the obstacle problem in the Sobolev space $H_0^1(\Omega)$ (see e.g. [29, 48, 59, 86, 96]) towards a multi-leader optimal control setting. For this reason, Section 5.1 introduces and classifies this multi-leader single-follower problem in the context of previous equilibrium problems. Afterwards, Section 5.2 focuses on stationarity concepts in $H_0^1(\Omega)$ that are derived with basic ideas of capacity theory. Known results (see e.g. [7]) for MPCCs are used to close the gap between MLFGs in abstract Banach spaces and their counterparts in $H_0^1(\Omega)$. While Section 5.3 shows that the penalization techniques, introduced in Chapter 4, lead to so-called $\epsilon$-almost C-stationary Nash equilibria, Section 5.4 is devoted to the numerical analysis. In this context, we formulate algorithms for both approaches, which are based on a Gauß-Seidel method and test the results from previous sections with respect to selected examples.
In Chapter 6, we consider a further class of non-convex equilibrium problems that is motivated by the contribution [14] "*A generalized Nash equilibrium approach for optimal control problems of autonomous cars*" by Axel Dreves and Matthias Gerdts. Therein, the authors analyze a driving scenario on the basis of an equilibrium problem. We show that the latter game can be written as a special EPCC. For this reason, we recall the problem in Section 6.1 and consider the general case in abstract Lebesgue spaces in Section 6.2. In particular, we derive stationary Nash equilibria and show that in this scenario all concepts coincide. Afterwards, Section 6.3 applies the results from the previous section to the equilibrium problem derived for the traffic scenario. In this context, we show in Subsec-

tion 6.3.1 that the penalty technique that was used in Subsection 4.2.2 also generates a sequence converging towards an S-stationary Nash equilibrium in this setting. Finally, we end this chapter by the corresponding numerical analysis (see Subsection 6.3.2) and consider two hypothetical traffic scenarios.
We conclude this thesis in Chapter 7, where we briefly summarize the achieved results and discuss open problems that can be the topic of future research.
Detailed information of basic numerical methods that are used during this work and auxiliary results that are needed in Chapter 6 can be found in Appendix A and B, respectively.

# 2. Background Knowledge

In this chapter we give an overview of all basic tools and concepts that are needed for the main parts of this thesis. Beginning with Section 2.1, where we introduce the important fundamentals concerning Banach spaces in general (Subsection 2.1.1). This includes sequential and topological concepts in function spaces as well as mappings in order to describe the interaction between Banach spaces. Afterwards, we move on and consider three selected Banach space classes in Subsection 2.1.2. Section 2.2 is dedicated to basic concepts of variational analysis, where we focus on different types of tangent and normal cones that are immediately applied in Section 2.3 to derive first-order optimality conditions for nonlinear programs in Subsection 2.3.1 and generalized Nash equilibrium problems in Subsection 2.3.2, respectively.

## 2.1. Introduction to Banach Space Theory

### 2.1.1. Fundamentals of Functional Analysis

The following content summarizes basic concepts of functional analysis. For a more detailed description, we refer for instance to [2, 97].
Throughout this thesis and if not stated otherwise, let $X$ be a *Banach space*, i.e. a normed space which is complete[3] with respect to the norm $\|\cdot\|_X$. For the sake of simplicity and if it is clear from the context, we often write $\|\cdot\| \equiv \|\cdot\|_X$ and $X$ instead of the tuple $(X, \|\cdot\|)$. Furthermore, $\mathbb{B}_\epsilon(\bar{x}) := \{x \in X \mid \|x - \bar{x}\| \leq \epsilon\}$ denotes the closed ball around $\bar{x} \in X$ with radius $\epsilon > 0$.
If the norm is induced by a scalar product, i.e. there exists a symmetric, bilinear mapping, denoted by $(\cdot,\cdot)_X : X \times X \to \mathbb{R}$, such that

$$\|\cdot\|_X := \sqrt{(\cdot,\cdot)_X},$$

then $X$ is called a *Hilbert space*. Since we will mainly consider sequential properties of sets and mappings, we highlight the next definition.

[3]A normed space is called *complete*, if every Cauchy sequence in $X$ converges.

**Definition 2.1.1.** We say that a sequence $\{x^k\} \subseteq X$ *converges* (strongly) to $\bar{x}$ in $X$ and write $x^k \to \bar{x}$, if $\lim_{k\to\infty} x^k = \bar{x}$ or equivalently $\lim_{k\to\infty} \|x^k - \bar{x}\|_X = 0$ is satisfied.

For later issues, an arbitrary subset $C \subseteq X$ is called

i) *convex*, if for all $x, y \in C$ and $\lambda \in [0,1]$, it holds $\lambda x + (1-\lambda)y \in C$.

ii) *bounded*, if there exists $c_B > 0$ such that $\|x\| \leq c_B$ for all $x \in C$.

iii) *(sequentially) closed*, if for all sequences $\{x^k\} \subseteq C$ with $x^k \to \bar{x}$, it holds $\bar{x} \in C$.

iv) *(sequentially) compact*, if every sequence $\{x^k\} \subseteq C$ has a subsequence that converges to some point in $C$.

v) a *cone*, if it holds $\alpha x \in C$ for all $x \in C$ and $\alpha \geq 0$.

In addition, we use the common notation cl $C$, int $C$, lin $C$ and cone $C$ to describe the *closure*, the *interior*, the *linear hull* and the *conic hull* of $C$, respectively. The boundary is defined by $\partial C :=$ cl $C \setminus$ int $C$. At this point, recall that in contrast to finite dimensions, a bounded and closed set is not compact in function spaces.
After focusing on set properties, we now move on and introduce mappings.

**Definition 2.1.2.** Let $X$ and $Y$ be two arbitrary Banach spaces. Then $T : X \to Y$ is called a *mapping* between $X$ and $Y$.
We call $T$ an *operator*, if $T$ is linear and use the notation $Tx$ or $T[x]$ instead of $T(x)$ to indicate the evaluation of $T$ at $x$. If $Y = \mathbb{R}$, then T is called a *functional*.

**Definition 2.1.3.** $T$ is said to be *(sequentially) continuous* at $\bar{x} \in X$, if for all $\{x^k\} \subseteq X$ with $x^k \to \bar{x}$ it holds that $T(x^k) \to T(\bar{x})$. $T$ is called continuous, if $T$ is continuous at all $x \in X$.

For an operator $T$, it is furthermore known (see [97, Satz II.1.2]) that continuity is equivalent to boundedness, i.e. there exists $c > 0$ such that

$$\|Tx\|_Y \leq c\|x\|_X \quad \forall x \in X.$$

The continuity can further be used to describe the relation between Banach spaces $X$ and $Y$. We say that $X$ is *continuously embedded* in $Y$, denoted by $X \hookrightarrow Y$, if $X \subseteq Y$ and there exists $c_E > 0$ such that $\|x\|_Y \leq c_E\|x\|_X$. In particular, the identity

$$I_{XY} : X \to \ Y, I_{XY}[x] = x \in \ Y$$

defines a continuous operator. If $I_{XY}$ is additionally compact[4], then $X$ is *compactly embedded* in $Y$ and denoted by $X \hookrightarrow\hookrightarrow Y$.

[4]An operator $T \in L(X,Y)$ is called *compact*, if cl $\{Tx \mid x \in \mathbb{B}_1(0)\}$ is compact w.r.t. $\|\cdot\|_Y$.

**Lemma 2.1.4** ([1, Remark 6.4 (2)]). *Let $X$, $Y$ and $Z$ be arbitrary Banach spaces such that $X \hookrightarrow Y$ and $Y \hookrightarrow Z$ are satisfied. If one of these embeddings is compact, then the composite embedding is compact, i.e. $X \hookrightarrow\hookrightarrow Z$.*

The space of bounded operators from $X$ to $Y$, denoted by $L(X,Y)$ and equipped with the norm

$$\|T\|_{L(X,Y)} \coloneqq \sup_{\|x\|_X \leq 1} \|Tx\|_Y,$$

is a normed space itself and a Banach space, if $Y$ is a Banach space. For the special case $Y = \mathbb{R}$, $L(X,\mathbb{R})$ is called the *dual space* of $X$ and denoted by $X^*$. Clearly, $X^*$ is a Banach space by the observations above. Similarly, the *bidual space* $X^{**}$ is defined as the dual space of $X^*$, i.e. $X^{**} \coloneqq L(X^*,\mathbb{R})$.
For an arbitrary operator $T \in L(X,Y)$, the operator $T^* \in L(Y^*,X^*)$ that is defined by the identity

$$\langle T^*y^*, x\rangle_X = \langle y^*, Tx\rangle_Y,$$

where $x \in X$ and $y^* \in Y^*$ are arbitrary but fixed, is called the *adjoint (operator)* of $T$. In this context, $\langle\cdot,\cdot\rangle_X$ denotes the dual pairing (see below). If $T$ is a bijection, then the inverse operator $T^{-1}$ exists and is continuous (see [97, Korollar IV.3.4]) and hence, is bounded.
In order to characterize a Banach space in more detail, its dual and bidual spaces take prominent roles. For this reason, we first introduce the *dual pairing*

$$\langle\cdot,\cdot\rangle_X : X \times X \to \mathbb{R},\ \langle x^*, x\rangle_X \coloneqq x^*(x),$$

which is continuous in both entries by the Cauchy-Schwarz inequality and is defined as the evaluation $x^*(x)$ for arbitrary $(x,x^*) \in X \times X^*$. Moreover, we define the *canonical embedding*

$$\iota_X : X \to X^{**},\ \iota_X(x) \coloneqq \langle\cdot, x\rangle_X,$$

which is a bounded operator. We again omit the subscript $X$, if the underlying space is clear from context.
By the theorem of Hahn-Banach (see [97, Korollar III.1.7]), $\iota_X$ is injective and isometric, i.e. $\|\iota_X(x)\|_{X^{**}} = \|x\|_X$ for any $x \in X$. Hence, if $\iota_X$ is additionally surjective, then $\iota_X$ is an isometric isomorphism and $X$ can be identified with its bidual $X^{**}$, i.e. $X \cong X^{**}$. In this context, $X$ is called *reflexive*.
For the special case that $X$ is even a Hilbert space, it is known by the famous Riesz representation theorem (see [97, Theorem V.3.6]) that for every $x^* \in X^*$ there exists a unique $\tilde{x} \in X$ such that $x^*(x) = (\tilde{x}, x)_X$ for all $x \in X$. Hence, the dual space of $X$ can be identified by the Hilbert space itself.

After introducing the dual and bidual of a Banach space, we now concentrate on further convergence concepts.

**Definition 2.1.5.** Let $\{x^k\} \subseteq X$ and $\{x^{*,k}\} \subseteq X^*$ be arbitrary sequences. Then $\{x^k\}$ is called *weakly convergent* to some $\bar{x} \in X$ and denoted by $x^k \rightharpoonup \bar{x}$, if $x^*(x^k) \to x^*(\bar{x})$ for all $x^* \in X^*$. Moreover, $\{x^{*,k}\}$ is called *weakly* convergent* (weakly-star convergent) to some $\bar{x}^* \in X^*$ and denoted by $x^{*,k} \rightharpoonup^* \bar{x}^*$, if $x^{*,k}(x) \to \bar{x}^*(x)$ for all $x \in X$.

In this context, the classical convergence, introduced in Definition 2.1.1, is often called *strong* convergence. Moreover, it is known that for any sequence $\{x^k\} \subseteq X$ such that $x^k \rightharpoonup \bar{x}$ in $X$, it holds $x^k \to \bar{x}$ in $Y$, if $X \hookrightarrow\hookrightarrow Y$.
By the theorem of Banach-Steinhaus (see [97, Theorem IV.2.1 and Corollary IV.2.3]) weakly and weakly* convergent sequences in a normed space are bounded. If $X$ is a reflexive Banach space, then the reverse holds and weak and weak* convergence coincide in the dual space $X^*$. Finally, it is known (see [97, Korollar III.1.6]) that the weak and weak* limit points, if they exist, are unique. The following result is stated in [2, Remark 6.3] and shows the convergence behavior of the dual pairing.

**Lemma 2.1.6.** *Let $\{x^k\} \subseteq X$ and $\{x^{*,k}\} \subseteq X^*$ be two sequences and let $\bar{x} \in X$ and $\bar{x}^* \in X^*$. Furthermore, assume that one of the following two conditions is satisfied:*

*i) It holds $x^k \rightharpoonup \bar{x}$ in $X$ and $x^{k,*} \to \bar{x}^*$ in $X^*$.*

*ii) It holds $x^k \to \bar{x}$ in $X$ and $x^{*,k} \rightharpoonup^* \bar{x}^*$ in $X^*$.*

*Then we have $\langle x^{*,k}, x^k \rangle_X \to \langle \bar{x}^*, \bar{x} \rangle_X$.*

Based on Definition 2.1.5, we extend the continuity concept of mappings $T : X \to Y$.

**Definition 2.1.7.** A mapping $T : X \to Y$ is called

i) *weakly (sequentially) continuous* at $\bar{x}$, if for all $\{x^k\} \subseteq X$ with $x^k \rightharpoonup \bar{x}$ in $X$ it holds that $T(x^k) \rightharpoonup T(\bar{x})$ in $Y$.

ii) *completely (sequentially) continuous* at $\bar{x}$, if for all $\{x^k\} \subseteq X$ with $x^k \rightharpoonup \bar{x}$ in $X$ it holds that $T(x^k) \to T(\bar{x})$ in $Y$.

Clearly, $T$ is called (weakly, completely) continuous, if $T$ is (weakly, completely) continuous at all $x \in X$.

Since strongly convergent sequences are also weakly convergent, it is straightforward to prove that completely continuous mappings are also (weakly) continuous. Moreover, it is

known (see [75]) that a compact operator $T \in L(X,Y)$ is completely continuous, while the reverse direction is only satisfied if $X$ is additionally reflexive.
Recall that we defined a closed and compact set in the beginning with respect to the strong convergence. In particular, the latter property of a set is oftentimes too restrictive, as for instance it is known that the unit ball $\mathbb{B}_1(0)$ is compact if and only if the underlying normed space is finite dimensional (see e.g. [97, Satz I.2.8]). A possible solution is given by the *weak topology*[5]. Similar as above for the strong topology, this motivates the following definition.

**Definition 2.1.8** ([88])**.** A subset $C \subseteq X$ is called

i) *weakly sequentially closed*, if the limit of every weakly convergent sequence $\{x^k\} \subseteq C$ is contained in $C$.

ii) *weakly sequentially relatively compact*, if every sequence $\{x^k\} \subseteq C$ contains a weakly convergent subsequence.

iii) *weakly sequentially compact*, if it is weakly sequentially closed and weakly sequentially relatively compact.

It is known that any closed, convex subset of a normed space is weakly sequentially closed ([35, Theorem 1.16]) and every bounded subset of a reflexive Banach space is weakly sequentially relatively compact (see [88, Theorem 2.10]). Hence, a closed, convex and bounded subset $C$ of a reflexive Banach space is weakly sequentially compact and thus, any sequence in $C$ contains a subsequence, which is weakly convergent and whose weak limit point is contained in $C$ as well.
The gap between weak compactness, i.e. the compactness with respect to the weak topology, and weak sequential compactness is closed by the theorem of Eberlein and Šmulian (see [97, Theorem VIII.6.1]), which states the equivalence of both properties.
We end this subsection with the definition of several differentiability concepts.

**Definition 2.1.9.** A mapping $T : X \to Y$ between arbitrary Banach spaces $X$ and $Y$ is called

i) *directionally differentiable* at $\bar{x} \in X$ in direction $d_x \in X$, if the limit

$$T'(\bar{x}; d_x) := \lim_{t \to 0} \frac{T(\bar{x} + td_x) - T(\bar{x})}{t}$$

exists.

[5]The weak topology is the coarsest topology such that all $x^* \in X^*$ are continuous [83].

ii) *(Fréchet) differentiable* at $\bar{x} \in X$, if there exists a bounded operator $T'(\bar{x}) \in L(X, Y)$ such that

$$\lim_{\|h\|_X \to 0} \frac{T(\bar{x} + h) - T(\bar{x}) - T'(\bar{x})h}{\|h\|_X} = 0$$

is satisfied.

iii) *continuously (Fréchet) differentiable* at $\bar{x} \in X$, if $T$ is differentiable at $\bar{x}$, the mapping $T' \in L(X, L(X, Y))$ is well-defined in a neighborhood of $\bar{x}$ and continuous at $\bar{x}$.

$T$ is called (continuously, directionally) differentiable, if it is (continuously, directionally) differentiable at all $x \in X$.

### 2.1.2. Examples of Banach Spaces

In this section we consider three types of Banach spaces that are of particular interest. In order to present Lebesgue and Sobolev spaces, we follow the work by Robert A. Adams and John J.F. Fournier [1]. For this reason, we start by introducing basic concepts about measures. If not stated otherwise, we consider $\mathbb{R}^d$ with $d \in \mathbb{N}$.

**Definition 2.1.10** ([1, Definition 1.38])**.** Let $\Sigma$ be a $\sigma$-algebra[6] on $\mathbb{R}^d$. Then the function $\mu : \Sigma \to \mathbb{R} \cup \{+\infty\}$ is called a *measure* if the following properties are satisfied:

1. For all $\mathcal{A} \in \Sigma$, we have $\mu(\mathcal{A}) \geq 0$.
2. The empty set has measure zero, i.e. $\mu(\emptyset) = 0$.
3. $\mu$ is countably additive, i.e. for all $\mathcal{A}_j \in \Sigma$ with $\mathcal{A}_i \cap \mathcal{A}_k = \emptyset$ for $i \neq k$, it holds

$$\mu\left(\bigcup_{j=1}^{\infty} \mathcal{A}_j\right) = \sum_{j=1}^{\infty} \mu(\mathcal{A}_j).$$

**Theorem 2.1.11** ([1, Theorem 1.39])**.** *There exists a $\sigma$-algebra $\Sigma$ of subsets of $\mathbb{R}^d$ and a measure $\mu$ on $\Sigma$ having the following properties:*

- *Every open set of $\mathbb{R}^d$ belongs to $\Sigma$.*
- *If $\mathcal{A} \subseteq \mathcal{B}$, $\mathcal{B} \in \Sigma$ and $\mu(\mathcal{B}) = 0$, then $\mathcal{A} \in \Sigma$ and $\mu(\mathcal{A}) = 0$.*

[6] A collection $\Sigma$ of subsets of $\mathbb{R}^d$ such that
i) $\mathbb{R}^d \in \Sigma$;  ii) $\mathcal{A} \in \Sigma \Rightarrow \mathbb{R}^d \setminus \mathcal{A} \in \Sigma$;  iii) $A_j \in \Sigma$ $(j = 1, 2, \dots) \implies \bigcup_{j=1}^{\infty} \mathcal{A}_j \in \Sigma$.

- *If* $\mathcal{A} := \{x \in \mathbb{R}^d \mid a_j \leq x_j \leq b_j \ \forall j = 1, \ldots, d\}$ *with* $a, b \in \mathbb{R}^d$ *such that* $a_j \leq b_j$ *for all* $j = 1, \ldots, d$*, then it holds* $\mathcal{A} \in \Sigma$ *and* $\mu(\mathcal{A}) = \prod_{j=1}^{d} (b_j - a_j)$.

The elements of $\Sigma$ are called *(Lebesgue) measurable* subsets of $\mathbb{R}^d$ and $\mu$ the *(Lebesgue) measure* in $\mathbb{R}^d$, respectively. Additionally, we say that a condition $\mathcal{P}$ holds *almost everywhere* (a.e.) in $\mathcal{B} \subseteq \mathbb{R}^d$, if $\mathcal{P}$ holds everywhere in $\mathcal{B} \setminus \mathcal{A}$ with $\mathcal{A} \subseteq \mathcal{B}$ and $\mu(\mathcal{A}) = 0$.
Similar to sets, we call $f : \mathcal{A} \to \mathbb{R} \cup \{-\infty, +\infty\}$ *measurable*, if the sets $\mathcal{A}$ and

$$\big\{\omega \in \mathcal{A} \mid f(\omega) > \alpha\big\}$$

for all $\alpha \in \mathbb{R}$ are measurable. As a consequence, every continuous function defined on a measurable set is measurable itself.

**Definition 2.1.12.** Let $\Omega \subseteq \mathbb{R}^d$ be a measurable subset and $f : \Omega \to \mathbb{R}$ a measurable function. Then $f$ is called *(Lebesgue) integrable* on $\Omega$, if

$$\int_\Omega |f(\omega)| \ d\omega < \infty$$

and we denote by $L^1(\Omega) := \big\{f : \Omega \to \mathbb{R} \mid \int_\Omega |f(\omega)| \ d\omega < \infty\big\}$ the set of Lebesgue integrable functions.

In order to introduce more general Lebesgue spaces, let $\Omega \subseteq \mathbb{R}^d$ be a domain[7] and $p \in [1, \infty]$. Then, we define the norm

$$\|f\|_{L^p(\Omega)} := \begin{cases} \big(\int_\Omega |f(\omega)|^p \ d\omega\big)^{\frac{1}{p}} & \text{if } p \in [1, \infty), \\ \operatorname*{ess\,sup}_{\omega \in \Omega} |f(\omega)| := \inf\big\{c \geq 0 \mid |f(\omega)| \leq c \text{ a.e. in } \Omega\big\} & \text{if } p = \infty \end{cases}$$

and set

$$L^p(\Omega) := \big\{f : \Omega \to \mathbb{R} \mid \|f\|_{L^p(\Omega)} < \infty\big\}.$$

Notice that $(L^p(\Omega), \| \cdot \|_{L^p(\Omega)})$ does not define a normed space since $\|f\|_{L^p(\Omega)} = 0$ does not imply $f = 0$. However, by using the equivalence relation

$$f \sim g \quad :\Longleftrightarrow \quad f(\omega) = g(\omega) \text{ f.a.a. } \omega \in \Omega,$$

the space $L^p(\Omega) := L^p(\Omega)/\sim$ defines a Banach space of equivalence classes of almost everywhere identical functions with respect to the norm $\| \cdot \|_{L^p(\Omega)}$. We continue by describing the relation of different Lebesgue spaces.

[7] A set $\Omega \subseteq \mathbb{R}^d$ is called a *domain*, if it is an open and connected subset of $\mathbb{R}^d$.

**Theorem 2.1.13** ([1, Theorem 2.14]). *Suppose that $\mu(\Omega) < \infty$ and let $1 \leq p \leq q \leq \infty$. If $f \in L^q(\Omega)$, then we have $f \in L^p(\Omega)$ and*

$$\|f\|_{L^p(\Omega)} \leq \mu(\Omega)^{\frac{1}{p}-\frac{1}{q}} \|f\|_{L^q(\Omega)},$$

*i.e. it holds $L^q(\Omega) \hookrightarrow L^p(\Omega)$.*

For the further analysis of Lebesgue spaces, we focus on $p \in [1, \infty)$ and observe that the dual space of $L^p(\Omega)$ is isometrically isomorphic to $L^q(\Omega)$ (see [1, Theorem 2.44]), where $q \in (1, \infty]$ is known as the *conjugate coefficient* of $p$ and computed by $p^{-1} + q^{-1} = 1$ (with $p = 1 \implies q = \infty$). Thus, depending on the choice of $p$ this leads to additional structure of $L^p(\Omega)$:

- If $p \in (1, \infty)$, then $L^p(\Omega)$ is reflexive ([1, Theorem 2.46]).
- If $p = 2$, then $L^2(\Omega)$ is a Hilbert space with scalar product

$$(f, g)_{L^2(\Omega)} := \int_\Omega f(\omega) g(\omega) \; d\omega.$$

After considering Lebesgue spaces, we now move on and introduce a further class of function spaces that is especially needed to analyze partial differential equations (PDE). It is well known that they do not attend a classical solution in general, but a weaker one. For this reason, we define the multi-index $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^d$ of order $|\alpha| = \sum_{i=1}^d \alpha_i$. Furthermore, we denote the partial derivative of $f$ at $\omega$ of order $|\alpha|$ by

$$D^\alpha f(\omega) := \frac{\partial^{|\alpha|}}{\partial \omega_1^{\alpha_1} \cdots \partial \omega_n^{\alpha_n}} f(\omega)$$

and define the set of locally integrable functions

$$L^1_{\text{loc}}(\Omega) := \left\{ u \in L^1(\Omega) \;\middle|\; \int_\Omega u(\omega) \varphi(\omega) \; d\omega < \infty \quad \forall \varphi \in C_0^\infty(\Omega) \right\},$$

where $C_0^\infty(\Omega)$ denotes the space of all functions that are continuously differentiable infinitely many times and have compact support[8].

**Definition 2.1.14** ([35, Definition 1.12]). Let $f \in L^1_{\text{loc}}(\Omega)$. If there exists $v \in L^1_{\text{loc}}(\Omega)$ such that

$$\int_\Omega v(\omega) \varphi(\omega) \; d\omega = (-1)^{|\alpha|} \int_\Omega f(\omega) D^\alpha \varphi(\omega) \; d\omega$$

for all $\varphi \in C_0^\infty(\Omega)$, then $D^\alpha f := v$ is called the *weak partial derivative* of $f$ of order $|\alpha|$.

[8] The *support* of a functional $f : X \to \mathbb{R}$ is given by $\text{supp}(f) := \{x \in X \mid f(x) \neq 0\}$.

The weak derivative is uniquely determined and consistent with the classical derivative (see [35, Remark 1.4]).
For $k \in \mathbb{N}_0$ and $p \in [1, \infty]$, we then define the *Sobolev space*

$$W^{k,p}(\Omega) := \big\{ f \in L^p(\Omega) \,\big|\, f \text{ has a weak derivative } D^\alpha f \in L^p(\Omega) \quad \forall |\alpha| \leq k \big\}.$$

Equipped with the norm

$$\|f\|_{W^{k,p}(\Omega)} := \begin{cases} \big( \sum\limits_{|\alpha| \leq k} \|D^\alpha f\|^p_{L^p(\Omega)} \big)^{\frac{1}{p}} & \text{if } p \in [1, \infty), \\ \sum\limits_{|\alpha| \leq k} \|D^\alpha f\|_{L^\infty(\Omega)} & \text{if } p = \infty, \end{cases}$$

$W^{k,p}(\Omega)$ is a Banach space (in the sense of equivalence classes) and its properties are inherited from the corresponding Lebesgue space $L^p(\Omega)$. In particular, the Sobolev space $W^{k,p}(\Omega)$ is reflexive, if $p \in (1, \infty)$ and is a Hilbert space, if and only if $p = 2$. In the latter case, it is common to write $H^k(\Omega) \equiv W^{k,p}(\Omega)$ with scalar product

$$(f, g)_{H^k(\Omega)} := \sum_{|\alpha| \leq k} (D^\alpha f, D^\alpha g)_{L^2(\Omega)}.$$

Last but not least, we define the space

$$W_0^{k,p}(\Omega) := \overline{C_0^\infty(\Omega)}^{W^{k,p}},$$

i.e. the closure of $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,p}(\Omega)}$ for all $p \in [1, \infty]$. In this context, the space $H_0^1(\Omega)$ plays a crucial role in order to solve elliptic PDEs with Dirichlet boundary constraints (see Section 4.1).
For the analysis of differential and integral operators, we need similar to Lebesgue spaces a certain ordering, which is summarized in the so called Sobolev embedding theorem (see e.g. [35, Theorem 1.14]) below.

**Theorem 2.1.15** (Sobolev embedding theorem). *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz boundary*[9]*, $k \in \mathbb{N}$ and $p \in [1, \infty)$. Then, the following assertions hold:*

1. *For all $l \in \mathbb{N}_0$ and $\gamma \in (0, 1)$ such that*

$$k - \frac{d}{p} \geq l + \gamma \tag{2.1}$$

[9] See for instance [88].

*is satisfied, it holds*[10]

$$W^{k,p}(\Omega) \hookrightarrow C^{l,\gamma}(\bar{\Omega}) := \big\{ f \in C^l(\bar{\Omega}) \;\big|\; D^\alpha f \text{ is } \gamma\text{-Hölder continuous for } |\alpha| = l \big\},$$

*where $C^l(\bar{\Omega})$ denotes the function space of $l$-times continuously differentiable functions on $\bar{\Omega}$.*

2. *For all $q \geq 1$ and $l \in \mathbb{N}_0$ such that*

$$k - \frac{d}{p} \geq l - \frac{d}{q} \tag{2.2}$$

*is satisfied, it holds $W^{k,p}(\Omega) \hookrightarrow W^{l,q}(\Omega)$.*

Under stronger assumptions, we obtain the compact embedding by the Rellich-Kondrachov theorem.

**Corollary 2.1.16** (Rellich-Kondrachov theorem, [1, Theorem 6.3])**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz boundary. Then for all $l \in \mathbb{N}_0$ and $\gamma \in [0,1]$ such that* (2.1) *holds strictly, we have*

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow C^{l,\gamma}(\bar{\Omega}),$$

*i.e. $W^{k,p}(\Omega)$ is compactly embedded in $C^{l,\gamma}(\bar{\Omega})$. For all $l \in \mathbb{N}_0$ and $q \geq 1$ such that $l < k$ and* (2.2) *holds strictly, we obtain*

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow W^{l,q}(\Omega).$$

We end this section by looking at a special type of subset of the Sobolev space $W^{1,p}(\Omega)$, where $\Omega := (t_0, t_F) \subseteq \mathbb{R}$ is assumed to be an open interval with $t_0 < t_F$. The following definition can be found in [20, Section 2.3.2] and [49, Definition 3.1], respectively.

**Definition 2.1.17.** A function $f : \bar{\Omega} := [t_0, t_F] \to \mathbb{R}$ is called *absolutely continuous*, if for every $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that

$$\sum_{i=1}^{m} |b_i - a_i| < \delta(\epsilon) \implies \sum_{i=1}^{m} |f(b_i) - f(a_i)| < \epsilon$$

is satisfied, where $m \in \mathbb{N}$ and $(a_i, b_i) \subseteq \Omega$ are pairwise disjoint intervals for all $i = 1, \ldots, m$.

[10] A function $f : \bar{\Omega} \to \mathbb{R}$ is called *$\gamma$-Hölder continuous*, if there exists $c > 0$ such that

$$|f(\omega_1) - f(\omega_2)| \leq c\|\omega_1 - \omega_2\|^\gamma$$

for all $\omega_1, \omega_2 \in \bar{\Omega}$.

Now, let $y$ be an arbitrary absolutely continuous function. Then the fundamental theorem of calculus is valid (see [49, Theorem 3.20]) and it holds

$$y(\tau) = y(t_0) + \int_{t_0}^{\tau} \dot{y}(t)\; dt \tag{2.3}$$

for almost all $\tau \in \bar{\Omega}$. In particular, we see that $\dot{y} \in L^1(\Omega)$. The latter observation motivates the following definition.

**Definition 2.1.18.** Let $\Omega = (t_0, t_F)$. The function space

$$AC^{1,p}(\Omega) \coloneqq \left\{ y : \bar{\Omega} \to \mathbb{R} \;\middle|\; y(\tau) = y(t_0) + \int_{t_0}^{\tau} \dot{y}(t)\; dt \quad \text{f.a.a. } \tau \in \bar{\Omega},\; \dot{y} \in L^p(\Omega) \right\}$$

is called *space of absolutely continuous functions* on $\bar{\Omega}$ of order $p$.

Notice that Definition 2.1.18 and in particular (2.3) indicate that $AC^{1,p}(\Omega)$ can be identified with $\mathbb{R} \times L^p(\Omega)$, i.e. there exists a bijection $\iota : AC^{1,p}(\Omega) \to \mathbb{R} \times L^p(\Omega)$ such that $\iota[y] = (y(t_0), \dot{y})$. The following theorem, which is stated in [20, Section 2.3.3] without proof and different norms, further characterizes the space $AC^{1,p}(\Omega)$.

**Theorem 2.1.19.** *Let $p \in [1, \infty)$ and $\Omega = (t_0, t_F)$ be given. Then $AC^{1,p}(\Omega)$ equipped with the norm*

$$|y|_{AC^{1,p}(\Omega)} \coloneqq |y(t_0)| + \|\dot{y}\|_{L^p(\Omega)}$$

*is a Banach space. If $p = 2$, then $AC^{1,2}(\Omega)$ is a Hilbert space with inner product*

$$(y, v)_{AC^{1,2}(\Omega)} \coloneqq y(t_0)v(t_0) + \int_{t_0}^{t_F} \dot{y}(t)\dot{v}(t)\; dt.$$

*Proof.* One easily verifies that $|\cdot|_{AC^{1,p}(\Omega)}$ and $(\cdot,\cdot)_{AC^{1,2}(\Omega)}$ indeed define a norm and an inner product on $AC^{1,p}(\Omega)$ and $AC^{1,2}(\Omega)$, respectively. In order to show the completeness, let $1 \le p < \infty$ be arbitrary and consider a Cauchy sequence $\{y_k\}_{k\in\mathbb{N}} \subseteq AC^{1,p}(\Omega)$, i.e.

$$\forall \epsilon > 0\; \exists N(\epsilon) \in \mathbb{N} :\; \forall n, m \ge N(\epsilon) \quad |y_n - y_m|_{AC^{1,p}(\Omega)} \le \epsilon.$$

Hence, we get

$$\epsilon \ge |y_n - y_m|_{AC^{1,p}(\Omega)} = |y_n(t_0) - y_m(t_0)| + \|\dot{y}_n - \dot{y}_m\|_{L^p(\Omega)},$$

which implies that $\{y_k(t_0)\} \subseteq \mathbb{R}$ and $\{\dot{y}_k\} \subseteq L^p(\Omega)$ are Cauchy sequences in $\mathbb{R}$ and $L^p(\Omega)$, respectively. Since $\mathbb{R}$ and $L^p(\Omega)$ are Banach spaces, $(y_k(t_0), \dot{y}_k)$ converges to some $(\hat{y}, \gamma)$ in $\mathbb{R} \times L^p(\Omega)$. By Theorem 2.1.13, we get that $\gamma \in L^1(\Omega)$ and

$$y(\tau) \coloneqq \hat{y} + \int_{t_0}^{\tau} \gamma(t)\; dt$$

is absolutely continuous with $y(t_0) = \hat{y}$ and $\dot{y} = \gamma$. This implies that $y \in AC^{1,p}(\Omega)$ and hence, the space $AC^{1,p}(\Omega)$ is complete. The result for $AC^{1,2}(\Omega)$ follows similarly. □

Notice that $(\cdot,\cdot)_{AC^{1,2}(\Omega)}$ induces a norm, which differs from $|\cdot|_{AC^{1,2}(\Omega)}$. The following result will be useful for the further analysis in this work.

**Theorem 2.1.20.** *For any $\Omega = (t_0, t_F)$ with $t_F > t_0$, the space $AC^{1,2}(\Omega)$ equipped with the norm $|\cdot|_{AC^{1,2}(\Omega)}$ is compactly embedded in $L^2(\Omega)$, i.e. $AC^{1,2}(\Omega) \hookrightarrow\hookrightarrow L^2(\Omega)$.*

*Proof.* It was shown in [53, Theorem 2.10] that

$$AC^{1,2}(\Omega) \hookrightarrow\hookrightarrow C(\bar{\Omega})$$

holds. Moreover, we have $C(\bar{\Omega}) \subseteq L^2(\Omega)$ and for all $y \in C(\bar{\Omega})$ it follows

$$\|y\|_{L^2(\Omega)} = \left(\int_{t_0}^{t_F} |y(t)|\, dt\right)^{\frac{1}{2}} \leq \left(\int_{t_0}^{t_F} \|y\|^2_{C(\bar{\Omega})}\, dt\right)^{\frac{1}{2}} = (t_F - t_0)^{\frac{1}{2}} \|y\|_{C(\bar{\Omega})}$$

with $\|y\|_{C(\bar{\Omega})} \coloneqq \max_{\tau\in\bar{\Omega}} |y(\tau)|$. Hence, $C(\bar{\Omega}) \hookrightarrow L^2(\Omega)$ and the result follows by Lemma 2.1.4. □

## 2.2. Basics about Variational Analysis

This section is dedicated to basic concepts of variational analysis and serves as the theoretical background in order to describe optimization problems. It heavily relies on the books [3, 7, 62].

We start by defining the *Bouligand tangent cone* $\mathcal{T}_C(\bar{x})$ and the *Clarke tangent cone* $\mathcal{T}_C^c(\bar{x})$ to an arbitrary subset $C$ of a Banach space $X$ at $\bar{x} \in X$ by

$$\begin{aligned}
\mathcal{T}_C(\bar{x}) &\coloneqq \big\{d \in X \;\big|\; \exists\{(d_k, t_k)\} \subseteq X \times \mathbb{R} :\; d_k \to d, t_k \searrow 0 \text{ and } \bar{x} + t_k d_k \in C\ \forall k \in \mathbb{N}\big\},\\
\mathcal{T}_C^c(\bar{x}) &\coloneqq \big\{d \in X \;\big|\; \forall\{(x_k, t_k)\} \subseteq C \times \mathbb{R} \text{ with } x_k \to \bar{x} \text{ and } t_k \searrow 0 :\\
&\qquad\qquad \exists\{d_k\} \subseteq X :\; d_k \to d,\; x_k + t_k d_k \in C\ \forall k \in \mathbb{N}\big\},
\end{aligned}$$

if $\bar{x} \in C$ and assign the individual tangent cones to the empty set if $\bar{x} \notin C$.

By the definition above we see that $\mathcal{T}_C(\bar{x})$ is a closed but in general non-convex cone, while $\mathcal{T}_C^c(\bar{x})$ is a closed and convex cone (see [3, Prop. 4.1.6]). Furthermore, the inclusion

$$\mathcal{T}_C^c(\bar{x}) \subseteq \mathcal{T}_C(\bar{x})$$

is satisfied. If $C$ is additionally convex, both tangent cones coincide and satisfy (see [7, Prop. 2.55] and [3, Prop. 4.2.1]))

$$\mathcal{T}_C(\bar{x}) = \text{cl}\, \mathcal{R}_C(\bar{x}) = \overline{\bigcup_{\lambda>0} \lambda(C - \bar{x})} = \text{cl cone}(C - \{\bar{x}\}), \tag{2.4}$$

where $\mathcal{R}_C\,(\bar{x})$ denotes the *radial cone* or *cone of feasible directions*, which is given by

$$\mathcal{R}_C(\bar{x}) := \begin{cases} \left\{d \in X \;\middle|\; \exists \tilde{t} > 0 \text{ s.t. } \forall t \in (0, \tilde{t}] : \; \bar{x} + td \in C\right\} & \text{if } \bar{x} \in C, \\ \emptyset & \text{if } \bar{x} \notin C \end{cases}.$$

Moreover, it is known (see [3, 10]) that for the Cartesian product $C = C_1 \times \cdots \times C_n$ of sets $C_i \subseteq X_i$ with $x = (x_1, \ldots, x_n)$ and $x_i \in C_i$, we have

$$\mathcal{T}_C(x) \subseteq \prod_{i=1}^{n} \mathcal{T}_{C_i}(x_i)$$

with equality, if $C_1, \ldots, C_n$ are convex.

In contrast to tangent cones that describe local approximations of sets in the *primal* space $X$, it is also possible to consider concepts that use information from the dual space $X^*$. According to [62, Definition 1.1.], we define for $\epsilon \geq 0$ the set of $\epsilon$-normals to $C$ at $\bar{x}$ by

$$\hat{\mathcal{N}}_C^{\epsilon}(\bar{x}) := \begin{cases} \left\{\eta \in X^* \;\middle|\; \limsup\limits_{x \to \bar{x}, x \in C} \frac{\langle \eta, x - \bar{x}\rangle_X}{\|x - \bar{x}\|_X} \leq \epsilon\right\} & \text{if } \bar{x} \in C, \\ \emptyset & \text{if } \bar{x} \notin C \end{cases}.$$

Based on the latter set, we introduce the *Fréchet/regular normal cone* $\hat{\mathcal{N}}_C(\bar{x})$ and the *limiting/basic/Mordukhovich normal cone* $\mathcal{N}_C(\bar{x})$ by

$$\begin{aligned} \hat{\mathcal{N}}_C(\bar{x}) &:= \hat{\mathcal{N}}_C^{0}(\bar{x}), \\ \mathcal{N}_C(\bar{x}) &:= \Big\{\eta \in X^* \;\Big|\; \exists \{(\epsilon_k, x_k, \eta_k)\} \subseteq \mathbb{R}^+ \times C \times X^* : \\ &\qquad \epsilon_k \searrow 0,\; x_k \to \bar{x},\; \eta_k \rightharpoonup^* \eta,\; \eta_k \in \hat{\mathcal{N}}_C^{\epsilon_k}(x_k) \;\; \forall k \in \mathbb{N}\Big\}. \end{aligned}$$

In order to define the *Clarke normal cone*, we need the concept of polar cones. Given a nonempty set $C \subseteq X$, we introduce the *polar cone*, the *dual cone* and the *annihilator* of $C$ by

$$\begin{aligned} C^{\circ} &:= \left\{\eta \in X^* \;\middle|\; \langle \eta, x\rangle \leq 0 \quad \forall x \in C\right\}, \\ C^{+} &:= -C^{\circ} = \left\{\eta \in X^* \;\middle|\; \langle \eta, x\rangle \geq 0 \quad \forall x \in C\right\}, \\ C^{\perp} &:= C^{\circ} \cap C^{+} = \left\{\eta \in X^* \;\middle|\; \langle \eta, x\rangle = 0 \quad \forall x \in C\right\}. \end{aligned}$$

Notice that $C^\circ$ and $C^+$ define weakly*-closed, convex cones, while $C^\perp$ is a weakly*-closed subspace.

**Remark 2.2.1.** *To avoid confusion with the notation, the polar and dual cone, as well as the annihilator of a set $B$ in the dual space $X^*$ are described by subscripts, e.g.*

$$B_\circ := \{d \in X \mid \langle \eta, d\rangle \le 0 \quad \forall \eta \in B\}.$$

*Moreover, we emphasize that both notations coincide, if $X$ is reflexive.*

Now, the Clarke normal cone is defined as the polar cone of the Clarke tangent cone, i.e.

$$\mathcal{N}_C^c(\bar{x}) := \mathcal{T}_C^c(\bar{x})^\circ = \{\eta \in X^* \mid \langle \eta, d\rangle \le 0 \quad \forall d \in T_C^c(\bar{x})\}.$$

By definition, the Fréchet normal cone $\hat{\mathcal{N}}_C(\bar{x})$ and the Clarke normal cone $N_C^c(\bar{x})$ are closed and convex, while the basic normal cone $\mathcal{N}_C(\bar{x})$ is in general neither convex [62] nor closed [62, Example 1.7].
Moreover, note that if $X$ is reflexive, then we can choose $\epsilon_k = 0$ in the definition of the limiting normal cone (see [62, Theorem 2.35]).
Similarly to tangent cones, it holds

$$\hat{\mathcal{N}}_C(\bar{x}) \subseteq \mathcal{N}_C(\bar{x}) \subseteq \mathcal{N}_C^c(\bar{x}) \tag{2.5}$$

and all cones coincide with the normal cone of convex analysis ([62, Proposition 1.3]), if $C$ is convex. Thus, we have

$$\mathcal{N}_C(\bar{x}) = \{\eta \in X^* \mid \langle \eta, x - \bar{x}\rangle \le 0 \quad \forall x \in C\} = \big(C - \{\bar{x}\}\big)^\circ. \tag{2.6}$$

In particular, the identity $\mathcal{N}_C(\bar{x}) = T_C(\bar{x})^\circ$ is satisfied for nonempty, closed and convex sets $C$. Notice that we only have $\mathcal{N}_C(\bar{x}) \subseteq T_C(\bar{x})^\circ$, if $C$ is not convex (see [62, Theorem 1.10]).
Now assume that $C$ is a nonempty, closed and convex cone. Then we obtain by [7, Example 2.62] and (2.4) that the following identities hold:

$$\mathcal{R}_C(\bar{x}) = C + \mathrm{lin}\{\bar{x}\},\ \mathcal{T}_C(\bar{x}) = \mathrm{cl}(C + \mathrm{lin}\{\bar{x}\}),\ \mathcal{N}_C(\bar{x}) = C^\circ \cap \{\bar{x}\}^\perp. \tag{2.7}$$

The next definition further characterizes sets in $X$.

**Definition 2.2.2** ([7, Definition 3.51])**.** We say that a closed and convex set $C \subseteq X$ is *polyhedric* with respect to $(\bar{x}, \bar{\eta}) \in C \times \mathcal{N}_C(\bar{x})$, if

$$\mathrm{cl}(\mathcal{R}_C(\bar{x}) \cap \{\bar{\eta}\}_\perp) = \mathcal{T}_C(\bar{x}) \cap \{\bar{\eta}\}_\perp =: \mathcal{K}_C(\bar{x}, \bar{\eta}), \tag{2.8}$$

where $\mathcal{K}_C(x, \eta)$ is called *critical cone* to $C$ with respect to $(x, \eta)$. The set $C$ is polyhedric at $\bar{x} \in C$, if (2.8) is satisfied for all $\eta \in \mathcal{N}_C(\bar{x})$ and $C$ is polyhedric, if it is polyhedric at all $x \in C$.

An important class of polyhedric sets are polyhedral ones. Therefore, recall that a closed and convex set $C \subseteq X$ is called *polyhedral*, if there exist $m \in \mathbb{N}_0$, $\eta_1, \dots, \eta_m \in X^*$ and real constants $c_1, \dots, c_m \in \mathbb{R}$ such that

$$C = \{x \in X \mid \langle \eta_i, x \rangle \leq c_i \quad \forall i = 1, \dots, m\}, \tag{2.9}$$

i.e. $C$ is described via the intersection of finitely many half-spaces. In this case, it is known that $\mathcal{R}_C(x) = \mathcal{T}_C(x)$ for any $x \in C$ and hence, (2.8) is satisfied.

**Example 2.2.3** (Polyhedric sets in Lebesgue and Sobolev spaces [93]). *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain. Then the sets*

$$\{y \in L^p(\Omega) \mid y_l \leq y \leq y_u \text{ a.e. in } \Omega\} \text{ and } \{z \in W_0^{1,q}(\Omega) \mid z_l \leq z \leq z_u \text{ a.e. in } \Omega\}$$

*are polyhedric in $L^p(\Omega)$ for all $p \in [1, \infty], y_l, y_u \in L^p(\Omega)$ (with $y_l \leq y_u$ a.e. in $\Omega$ ) and in $W_0^{1,q}(\Omega)$ for all $q \in [1, \infty), z_l, z_u \in W_0^{1,q}(\Omega)$ (with $z_l \leq z_u$ a.e. in $\Omega$), respectively.*

The following lemma shows one of the benefits of polyhedric cones.

**Lemma 2.2.4** ([95, Lemma 4.3]). *Let $C$ be a closed, convex cone and let $(\bar{x}, \bar{\eta}) \in \operatorname{gph} \mathcal{N}_C$, where $\operatorname{gph} \mathcal{N}_C := \{(x, \eta) \in X \times X^* \mid \eta \in \mathcal{N}_C(x)\}$ denotes the graph of the normal cone mapping $\mathcal{N}_C : X \rightrightarrows X^*$. Then the following conditions are equivalent.*

*a) The cone $C$ is polyhedric w.r.t. $(\bar{x}, \bar{\eta})$.*

*b) The cone $C^\circ$ is polyhedric w.r.t. $(\bar{\eta}, \bar{x})$.*

*c) $\mathcal{K}_C(\bar{x}, \bar{\eta})^\circ = \mathcal{K}_{C^\circ}(\bar{\eta}, \bar{x})$.*

*d) $\mathcal{K}_{C^\circ}(\bar{\eta}, \bar{x})_\circ = \mathcal{K}_C(\bar{x}, \bar{\eta})$.*

Later in this section, we present further polyhedric sets. Before doing that, we define another structure. For this reason, let $C$ be a nonempty, closed and convex cone and consider the order relation '$\leq_C$' induced by $C$, which is defined by

$$x \leq_C y \qquad :\Longleftrightarrow \qquad y - x \in C. \tag{2.10}$$

Then it was shown in [42, Theorem 1.18] that '$\leq_C$' is a *partial ordering* of $X$, i.e. it satisfies

1) $x \leq_C x,$ (Reflexivity)
2) $x \leq_C y \ \wedge \ w \leq_C z \implies x + w \leq_C y + z,$
3) $x \leq_C y \ \wedge \ y \leq_C z \implies x \leq_C z,$ (Transitivity)
4) $x \leq_C y \implies \alpha x \leq_C \alpha y$

for all $w, x, y, z \in X$, $\alpha \geq 0$. Here, $C$ is called *ordering* or *positive cone* and $(X, \leq_C)$ the *partially ordered set*.
While condition $2)$ can be seen as an invariance with respect to addition, condition $4)$ describes the invariance with respect to scalar multiplication. If $C$ is additionally *pointed*, i.e. $C \cap (-C) = \{0\}$, then it holds

$$x \leq_C y \ \wedge \ y \leq_C x \quad \Longrightarrow \quad x = y. \qquad \text{(Antisymmetry)}$$

In the remaining part of this section, we assume that $C$ is pointed and call $s \in X$ the *supremum* or *least/smallest upper bound* of $x, y \in X$, if

$$s \in \mathcal{S}(x, y) := \big\{ z \in X \ \big| \ x \leq_C z \ \wedge \ y \leq_C z \big\}$$

and $s \leq_C w$ for all $w \in \mathcal{S}(x, y)$. Since $C$ is pointed and hence, the order relation is antisymmetric, there exists at most one supremum. In case of existence, we write $s := \max_C(x, y)$. In the next step, we define an important structure, which helps to describe polyhedric sets.

**Definition 2.2.5** ([7, Def. 3.56])**.** Let $C$ be a pointed ordering cone. We say that $C$ induces a *lattice structure* on $X$, if for any $x, y \in X$ the supremum $s$ exists and the operator $\max_C : X \times X \to X$ is continuous. The partially ordered set $(X, \leq_C)$ is called *vector lattice*.

**Remark 2.2.6.** *At this point, we mention that a vector lattice can be used to illustrate a convex feasible set of a minimization problem.*
*As we will see in Subsection 2.3.1, constraints are in general given by*

$$X_{\text{ad}} = \big\{ x \in X \ \big| \ g(x) \in C \big\},$$

*where $C \subseteq Y$ is a nonempty, closed and convex cone, which induces a partial ordering and $g : X \to Y$ denotes an arbitrary mapping. In order to guarantee convexity of $X_{\text{ad}}$, it is well-known (see e.g. [45, Lemma 2.4]) that $g$ has to be concave with respect to $C$ or equivalently convex with respect to $-C$, i.e.*

$$g(\gamma x_1 + (1-\gamma)x_2) - \gamma g(x_1) - (1-\gamma) g(x_2) \in C \tag{2.11}$$

*for all $x_1, x_2 \in X$ and $\gamma \in [0, 1]$. By using* (2.10)*, we see that* (2.11) *is equivalent to*

$$\gamma g(x_1) + (1-\gamma) g(x_2) \leq_C g(\gamma x_1 + (1-\gamma)x_2),$$

*which has the familiar notation of concave functions in $\mathbb{R}$ and hence, the naming is meaningful. Moreover, if $g$ is additionally Fréchet differentiable, it was shown in [42, Theorem 2.20] that*

$$g(x_1) + g'(x_1)(x_2 - x_1) \leq_C g(x_2) \tag{2.12}$$

*is satisfied for all $x_1, x_2 \in X$.*
*Additionally, we see that the supremum can be used to describe penalty functions, since we have the relation*

$$x \in X_{\text{ad}} \iff g(x) \in C \iff 0 \leq_C g(x) \iff \max_C(-g(x), 0) = 0.$$

Analogously to the supremum, it is possible to define the *infimum* or *greatest lower bound* $v$ of $(x, y) \in X$, i.e.

$$v \in \mathcal{V}(x, y) := \{z \in X \mid z \leq_C x \ \wedge \ z \leq_C y\}$$

and $w \leq_C v$ for all $w \in \mathcal{V}(x, y)$. If the infimum exists, we write $v := \min_C(x, y)$. Similar to the finite dimensional setting with the definition of maximum and minimum of two real numbers, we have for arbitrary $x, y \in X$ the identities

$$\begin{aligned} \max_C(x, y) &= x + y - \min_C(x, y), \\ \min_C(x, y) &= -\max_C(-x, -y), \\ x &= x_+ + x_- \end{aligned}$$

with *positive part* $x_+ := \max_C(x, 0)$ and *negative part* $x_- := \min_C(x, 0)$ of $x \in X$. Now, we formulate the theorem that closes the gap between a vector lattice and a polyhedric set.

**Theorem 2.2.7** ([7, Theorem 3.58])**.** *Let $C \subseteq X$ induce a lattice structure, i.e. $(X, \leq_C)$ is a vector lattice. Then $C$ is polyhedric.*

**Example 2.2.8.** *A direct consequence of Theorem 2.2.7 is the polyhedricity of the nonnegative cones (see [7, Corollary 6.46] and [53, Example 2.25])*

$$L^p(\Omega)_0^+ := \{y \in L^p(\Omega) \mid y \geq 0 \quad \text{a.e. in } \Omega\}$$

*with $p \in [1, \infty]$ and*

$$W_0^{1,q}(\Omega)_0^+ := \{y \in W_0^{1,q}(\Omega) \mid y \geq 0 \quad \text{a.e. in } \Omega\}$$

*with $q \in [2, \infty)$, respectively, which will be of greater importance in later sections.*

## 2.3. Basic Concepts about Nonlinear Programming and Game Theory

In this section we concentrate on basic tools in nonlinear programming (Subsection 2.3.1) and in game theory (Subsection 2.3.2). In particular, the former one serves as a further preparation for the latter, where we use the introduced concepts in order to derive optimality conditions for equilibrium points. In addition, we give a survey of known existing results of Nash equilibria and highlight the importance of convexity.
We refer to the books [5, 7, 41] for a more detailed introduction to nonlinear optimization and to [33, 34, 43] for contributions concerning equilibrium problems. We emphasize that although [43] is restricted to the finite dimensional setting, the basic ideas stay the same.

### 2.3.1. Nonlinear Programming in Banach Spaces

Throughout this section we consider the generic minimization problem

$$\begin{aligned} \min \quad & f(x) \\ \text{w.r.t.} \quad & x \in X, \\ \text{s.t.} \quad & x \in Q,\ g(x) \in C, \end{aligned} \tag{2.13}$$

with a functional $f : X \to \mathbb{R}$, a mapping $g : X \to Y$ from a Banach space $X$ into a Banach space $Y$ and nonempty, closed subsets $Q \subseteq X$ and $C \subseteq Y$. Moreover, we denote by

$$X_{\text{ad}} := \big\{x \in Q \;\big|\; g(x) \in C\big\} \tag{2.14}$$

the *feasible set* of (2.13) and say that $x$ is *feasible* for (2.13) if and only if $x \in X_{\text{ad}}$.

**Definition 2.3.1.** Let $\bar{x} \in X_{\text{ad}}$. Then $\bar{x}$ is called

- a *local solution* of (2.13), if there exists $\epsilon > 0$ such that

$$f(\bar{x}) \leq f(x) \qquad \forall x \in X_{\text{ad}} \cap \mathbb{B}_\epsilon(\bar{x}),$$

- a *global solution* of (2.13), if

$$f(\bar{x}) \leq f(x) \qquad \forall x \in X_{\text{ad}}.$$

In order to ensure the existence of a (local) solution, we first recall Weierstraß's existence theorem in $\mathbb{R}^d$, which states that each continuous function $f : \mathbb{R}^d \to \mathbb{R}$ attains a minimum on a nonempty, compact subset $B \subseteq \mathbb{R}^d$. However, it is known that this result is not applicable in function spaces under the same assumptions.
On the one hand, compactness with respect to the topology induced by the norm $\|\cdot\|_X$ is in general too strong and hence, we have to deal with the weak (sequential) topology. On the other hand, this leads directly to a strengthening of the continuity assumption on $f$.

**Definition 2.3.2.** Let $X$ be a Banach space. A functional $f : X \to \mathbb{R} \cup \{\infty\}$ is called

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

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y);$$

ii) *strictly convex*, if for all $x, y \in X$ and $\lambda \in [0,1]$, it holds

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y);$$

iii) *(weakly) lower semicontinuous* (w.l.s.c.) at $\bar{x}$, if for any sequence $\{x^k\} \subseteq X$ such that $x^k \to \bar{x}$ ($x^k \rightharpoonup \bar{x}$), it holds

$$f(\bar{x}) \leq \liminf_{k \to \infty} f(x^k);$$

iv) *coercive*, if it holds

$$\forall \{x^k\} \subseteq X \quad \text{s.t.} \quad \|x^k\| \to \infty \qquad \Longrightarrow \qquad f(x^k) \to \infty.$$

Clearly, any weakly lower semicontinuous functional is also lower semicontinuous. Then we can state the following existence result in Banach spaces.

**Theorem 2.3.3** ([7, Theorem 2.6]). *Let $f : X \to \mathbb{R}$ be l.s.c. and assume that $X_{\text{ad}}$ is nonempty and weakly compact. Then problem* (2.13) *has a solution.*

If $X_{\text{ad}}$ is only weakly sequentially closed (see Definition 2.1.8), we still can hope for existence under stronger assumptions on $f$. For this reason, observe that if $f$ is additionally coercive, we obtain for arbitrary $\tilde{x} \in X_{\text{ad}}$ that the set

$$\tilde{X}_{\text{ad}} := \left\{ x \in X_{\text{ad}} \;\middle|\; f(x) \leq f(\tilde{x}) \right\}$$

is bounded and hence, Theorem 2.3.3 is applicable.

**Corollary 2.3.4.** *Let $f$ be convex and continuous and assume that $X_{\text{ad}}$ is a nonempty, closed, convex and bounded subset of a reflexive Banach space $X$. Then problem* (2.13) *has a solution. Moreover, the solution set of* (2.13) *is convex. If $f$ is additionally strictly convex, then the solution set is a singleton.*

*Proof.* For the first two parts we refer to [41, Theorem 2.12 and Theorem 2.14].
Now, let $f$ be strictly convex and assume that there exist two solutions $x_1, x_2 \in X$ with $x_1 \neq x_2$. In this case, it holds

$$f\left(\frac{x_1+x_2}{2}\right) < \frac{f(x_1)+f(x_2)}{2} = f(x_1) < f\left(\frac{x_1+x_2}{2}\right),$$

which is a contradiction. Hence, we have $x_1 = x_2$. □

In order to describe and find solutions of (2.13) numerically, one often refers to first-order optimality conditions.

**Theorem 2.3.5** (see e.g. [35, Theorem 1.46])**.** *Let $\bar{x}$ be a local solution of* (2.13) *and assume that $f$ is continuously differentiable. Then $\bar{x}$ is Bouligand stationary (B-stationary), i.e. it holds*

$$\langle f'(\bar{x}), d\rangle_X \geq 0 \qquad \forall d \in \mathcal{T}_{X_{\text{ad}}}(\bar{x}). \tag{2.15}$$

*If $X_{\text{ad}}$ is convex, then we obtain*

$$\langle f'(\bar{x}), x - \bar{x}\rangle_X \geq 0 \qquad \forall x \in X_{\text{ad}} \tag{2.16}$$

*and if $f$ is additionally convex, then $\bar{x}$ is a global solution of* (2.13).

Notice that in case of convexity, condition (2.15) is by definition equivalent to

$$0 \in \ f'(\bar{x}) \ + \ \mathcal{N}_{X_{\text{ad}}}(\bar{x}).$$

A drawback of the notation of first-order optimality conditions in Theorem 2.3.5 is the fact that the optimality condition is quite cumbersome for further computations. Moreover, it does not use the whole algebraic structure of the feasible set $X_{\text{ad}}$.
For this reason, we assume that $g$ is continuously differentiable and define the *linearized tangent cone* $\mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})$ and the *linearized normal cone* $\mathcal{N}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})$ of the feasible set $X_{\text{ad}}$ at $\bar{x} \in X$ by

$$\begin{aligned}\mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x}) &:= \left\{d \in \mathcal{T}_Q(\bar{x}) \;\middle|\; g'(\bar{x})d \in \mathcal{T}_C(g(\bar{x}))\right\},\\ \mathcal{N}^{\text{lin}}_{X_{\text{ad}}}(\bar{x}) &:= \left\{g'(\bar{x})^*\mu + \xi \in X^* \;\middle|\; \mu \in \mathcal{T}_C(g(\bar{x}))^\circ,\ \xi \in \mathcal{T}_Q(\bar{x})^\circ\right\}.\end{aligned} \tag{2.17}$$

At this point, the goal is to derive optimality conditions that use $\mathcal{N}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})$ instead of $\mathcal{N}_{X_{\text{ad}}}(\bar{x})$. Therefore, observe that the relations $\mathcal{T}_{X_{\text{ad}}}(\bar{x}) \subseteq \mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})$ and $\mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})^{\circ} = \text{cl}^{*}\mathcal{N}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})$ (see e.g. [19, Proposition 3.2]) are valid. In order to guarantee equality of the first inclusion we need additional assumptions, so called *constraint qualifications*.

**Definition 2.3.6.** Let $\bar{x} \in X_{\text{ad}}$. Then $\bar{x}$ satisfies

- the *Guignard Constraint Qualification* (GCQ), if it holds
$$\mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})^{\circ} = \mathcal{N}^{\text{lin}}_{X_{\text{ad}}}(\bar{x});$$
- the *Abadie Constraint Qualification* (ACQ), if it holds
$$\begin{aligned}\mathcal{T}_{X_{\text{ad}}}(\bar{x}) &= \mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x}),\\ \mathcal{N}^{\text{lin}}_{\text{ad}}(\bar{x}) &= \text{cl}^{*}\, \mathcal{N}^{\text{lin}}_{\text{ad}}(\bar{x});\end{aligned}$$
- the *Kurcyusz Robinson Zowe Constraint Qualification* (KRZCQ), if it holds
$$g'(\bar{x})\,[\mathcal{R}_{Q}(\bar{x})] - \mathcal{R}_{C}(g(\bar{x})) = Y; \tag{2.18}$$

At this point, we refer to [19, Lemma 3.4] for a proof that

$$\text{KRZCQ} \quad \Longrightarrow \quad \text{ACQ} \quad \Longrightarrow \quad \text{GCQ}$$

is valid. For the further analysis, we require the following assumptions.

**Assumption 2.3.7.** *Let $f$ be convex and continuous differentiable and assume that the mapping $g$ is continuously differentiable and $-C$-convex (see Remark 2.2.6). Moreover, let $Q \subseteq X$ and $C \subseteq Y$ be nonempty, closed and convex in which $C$ is additionally a cone.*

In this context, we say that (2.13) is a *convex problem*, if Assumption 2.3.7 is satisfied. In particular, we refer to [45, Lemma 2.4] for a verification of the convexity. Furthermore, we introduce the *Lagrangian* of (2.13) by

$$\mathcal{L} : X \times Y^{*} \to \mathbb{R}, \quad \mathcal{L}(x, \mu) \coloneqq f(x) + \langle \mu, g(x) \rangle$$

and call a tuple $(\bar{x}, \bar{\mu}) \in X \times Y^{*}$ a *Karush-Kuhn-Tucker point* (KKT-point) of (2.13), if

$$-\mathcal{L}'(\bar{x}, \bar{\mu}) \in \mathcal{T}_{Q}(\bar{x})^{\circ}, \quad \bar{\mu} \in \mathcal{T}_{C}(g(\bar{x}))^{\circ}. \tag{2.19}$$

Note that if $(\bar{x}, \bar{\mu})$ is a KKT-point, then $\bar{x}$ is feasible, since we have by definition of the tangent cones that $\bar{x} \in Q$ and $g(\bar{x}) \in C$, respectively.

**Definition 2.3.8.** A point $\bar{x} \in X$ is called *stationary point* of (2.13), if there exists a *Lagrange multiplier* $\bar{\mu} \in Y^*$ such that $(\bar{x}, \bar{\mu})$ is a KKT-point of (2.13).

Now, we are able to formulate a more concrete first order condition that summarizes known results from [7] and [100].

**Theorem 2.3.9.** *Let $\bar{x}$ be a local solution of (2.13) and assume that KRZCQ is satisfied in $\bar{x}$. Then $\bar{x}$ is a stationary point of (2.13), i.e. there exists $\bar{\mu} \in Y^*$ such that (2.19) is satisfied. Moreover, the set of Lagrange multipliers*

$$\Lambda(\bar{x}) := \big\{(\xi, \mu) \in \mathcal{T}_Q(\bar{x})^\circ \times \mathcal{T}_C(g(\bar{x}))^\circ \;\big|\; f'(\bar{x}) + g'(\bar{x})^*\mu + \xi = 0\big\} \tag{2.20}$$

*is nonempty, convex, bounded and weakly*-compact.*
*In reverse, if (2.13) is convex and $\bar{x}$ is a stationary point of (2.13), then $\bar{x}$ is a global solution of (2.13).*

*Proof.* Since KRZCQ implies that ACQ and hence, GCQ it satisfied, the tangent cones coincide and the polar cone of the linearized tangent cone is equal to the linearized normal cone. Thus, it follows that

$$(2.15) \quad \Longleftrightarrow \; -f'(\bar{x}) \in \mathcal{T}^{\text{lin}}_{X_{\text{ad}}}(\bar{x})^\circ \quad \Longleftrightarrow \; -f'(\bar{x}) \in \mathcal{N}^{\text{lin}}_{X_{\text{ad}}}(\bar{x}),$$

which shows (2.19).
In order to show the properties of $\Lambda(\bar{x})$, we refer to [7, Theorem 3.6] or [100].
Now, let (2.13) be convex, let $(\bar{x}, \bar{\mu})$ solve (2.19) and assume that $\bar{x}$ is not a global solution of (2.13). Then there exists $\tilde{x} \in X_{\text{ad}}$ such that $f(\tilde{x}) < f(\bar{x})$. Since $\bar{x}$ is a stationary point, there exist multipliers $(\bar{\xi}, \bar{\mu}) \in \mathcal{N}_Q(\bar{x}) \times \mathcal{N}_C(g(\bar{x}))$ such that

$$0 = \langle f'(\bar{x}) + g'(\bar{x})^*\bar{\mu} + \bar{\xi}, \tilde{x} - \bar{x}\rangle_X = \langle f'(\bar{x}), \tilde{x} - \bar{x}\rangle_X + \langle \bar{\mu}, g'(\bar{x})(\tilde{x} - \bar{x})\rangle_X + \langle \bar{\xi}, \tilde{x} - \bar{x}\rangle_X.$$

Considering all three terms above separately, we observe that the last term is non-positive by the definition of $\bar{\xi}$. The same result can be deduced for the second term, as we have

$$\langle \bar{\mu}, g'(\bar{x})(\tilde{x} - \bar{x})\rangle_X \leq \langle \bar{\mu}, g(\bar{x}) + g'(\bar{x})(\tilde{x} - \bar{x}) - g(\tilde{x})\rangle_X \leq 0,$$

where we use the definition of $\bar{\mu}$, the feasibility of $\tilde{x}$ and that $g$ is $-C$-convex. In particular, the latter implies by (2.12) that $g(\bar{x}) + g'(\bar{x})(\tilde{x} - \bar{x}) - g(\tilde{x}) \in C$ is valid. Hence, we further estimate by the convexity of $f$ that

$$0 = \langle f'(\bar{x}) + g'(\bar{x})^*\bar{\mu} + \bar{\xi}, \tilde{x} - \bar{x}\rangle_X \leq \langle f'(\bar{x}), \tilde{x} - \bar{x}\rangle_X \leq f(\tilde{x}) - f(\bar{x}) < 0,$$

which is a contradiction and hence, $\bar{x}$ is a global solution of (2.13). $\square$

For later purposes the following theorem generalizes [94, Theorem 3.2] and gives a necessary condition that the Lagrange multiplier of (2.13) is uniquely determined.

**Theorem 2.3.10.** *Let $\bar{x}$ be a local solution of* (2.13)*, which satisfies both KRZCQ and*

$$\mathrm{cl}(g'(\bar{x})[(\mathcal{T}_Q(\bar{x})^\circ)_\perp] - (\mathcal{T}_C(\bar{x})^\circ)_\perp) = Y. \tag{2.21}$$

*Then there exists a unique Lagrange multiplier $\bar{\mu}$ such that $(\bar{x}, \bar{\mu})$ is a KKT point.*

*Proof.* By Theorem 2.3.9 the existence of a KKT-point $(\bar{x}, \bar{\mu})$ was already shown. With [53, Corollary 2.37], we obtain that (2.21) is equivalent to

$$\mathrm{cl}\left(\begin{pmatrix} g'(\bar{x}) \\ I \end{pmatrix}[X] - \begin{pmatrix} (\mathcal{T}_C(g(\bar{x})^\circ)_\perp \\ (\mathcal{T}_Q(\bar{x})^\circ)_\perp \end{pmatrix}\right) = \begin{bmatrix} Y \\ X \end{bmatrix}.$$

Defining $\tilde{g}(x) := (g(x), x)$ and $\tilde{C} := C \times Q$, the uniqueness follows by applying [94, Theorem 3.2]. □

Now, the subsequent corollary is motivated by [7, Section 2.3.4] in which the name of the presented constraint qualification was given in [19].

**Corollary 2.3.11.** *Let $\bar{x}$ be a local solution of* (2.13) *and assume that the Full Range Constraint Qualification (FRCQ)*

$$g'(\bar{x})\,[\mathcal{R}_Q(\bar{x})] = Y$$

*is satisfied. Then $\bar{x}$ is a stationary point of* (2.13)*. If $g'(\bar{x})$ is surjective and either $Q = X$ or $\bar{x} \in \mathrm{int}(Q)$ is valid, then the corresponding Lagrange multiplier is unique.*

*Proof.* In order to see the first claim, observe that FRCQ implies KRZCQ and the result follows by Theorem 2.3.9.
For the second assertion, we obtain by the additional assumptions on $Q$ that $\mathcal{R}_Q(\bar{x}) = X$ and hence, FRCQ implies (2.21) and the Lagrange multiplier is unique. □

**Remark 2.3.12.** *Condition* (2.21) *only ensures the existence of at most one Lagrange multiplier (see [95, Lemma 3.1 and Theorem 3.2] for the case $Q = X$). So in particular not the existence itself. Hence, KRZCQ or a weaker constraint qualification has to be satisfied as well. Moreover, notice that in finite dimensions KRZCQ coincides with the Mangasarian-Fromovitz constraint qualification (see [7, Section 2.3.4]).*

*If $C := \mathbb{R}^d_+$, then condition* (2.21) *is equivalent to the linear independent constraint qualification, which implies uniqueness and in this case even existence of a Lagrange multiplier. In order to see this, observe that we have*

$$\begin{aligned}(\mathcal{T}_C(\bar{x})^\circ)_\perp &= (\mathcal{T}_C(\bar{x})^\circ)_\circ \cap (-(\mathcal{T}_C(\bar{x})^\circ))_\circ = (\mathcal{T}_C(\bar{x})^\circ)_\circ \cap ((-\mathcal{T}_C(\bar{x}))^\circ)_\circ \\ &= \mathcal{T}_C(\bar{x}) \cap (-\mathcal{T}_C(\bar{x})) \\ &\subseteq \mathcal{T}_C(\bar{x})\end{aligned}$$

*for arbitrary nonempty, closed, not necessarily convex sets $C \subseteq X$ and $\bar{x} \in C$. Since we have $\mathcal{R}_C(\bar{x}) = \mathcal{T}_C(\bar{x})$ in finite dimension,* (2.21) *implies KRZCQ in $\mathbb{R}^d$.*

We end this section with an auxiliary result that is needed in Chapter 3.

**Lemma 2.3.13** ([7, Prop. 3.16])**.** *Assume that $\bar{x}$ satisfies KRZCQ. Then the set of singular Lagrange multipliers contains only the zero element, i.e.*

$$\Lambda_0(\bar{x}) := \big\{(\mu,\xi) \in \mathcal{T}_C(g(\bar{x}))^\circ \times \mathcal{T}_Q(\bar{x})^\circ \;\big|\; g'(\bar{x})^*\mu + \xi = 0\big\} = \big\{(0,0)\big\}. \tag{2.22}$$

### 2.3.2. Game Theory in Banach Spaces

After giving a short introduction about the basic concepts of nonlinear programming, we now take a step further and consider *generalized Nash games* or *generalized Nash equilibrium problems* (GNEPs). In this context, a GNEP $\Gamma := \{f^\nu, X^\nu_{\text{ad}}\}_{\nu=1}^N$ consists of $N \in \mathbb{N}$ *players*, where each player solves a parametric minimization problem of the form

$$\begin{aligned} \min \;\; & f^\nu(x^\nu, x^{-\nu}) \\ \text{w.r.t.} \;\; & x^\nu \in X_\nu, \\ \text{s.t.} \;\; & x^\nu \in X^\nu_{\text{ad}}(x^{-\nu}), \end{aligned} \tag{2.23}$$

where we assume that $X_\nu$ is a Banach space. In particular, player $\nu$'s opponents decisions, denoted by

$$x^{-\nu} := (x^1, \dots, x^{\nu-1}, x^{\nu+1}, \dots, x^N) \in X_{-\nu} := \prod_{j\neq\nu} X_j,$$

can influence his objective functional $f^\nu : X \to \mathbb{R}$ and his *strategy set*, which is given by the set-valued mapping $X^\nu_{\text{ad}} : X_{-\nu} \rightrightarrows X_\nu$.
For the sake of simplicity, we use the common game theoretic notations and summarize all player strategies by the vector $x := (x^1, \dots, x^N) \in X := X_1 \times \cdots \times X_N$ and write $x = (x^\nu, x^{-\nu}) \in X_\nu \times X_{-\nu}$ to emphasize player $\nu$'s strategy in $x$. We proceed by further characterizing the game $\Gamma$.

**Definition 2.3.14.** The generalized Nash equilibrium problem $\Gamma$ is called

i) *Nash equilibrium problem* (NEP), if for all $\nu = 1, \dots, N$ the strategy set is independent of player $\nu$'s opponent strategy vector $x^{-\nu}$, i.e. $X^\nu_{\text{ad}}(x^{-\nu}) \equiv X^\nu_{\text{ad}}$.

ii) *GNEP of joint/shared type*, if the strategy set is given by

$$X^\nu_{\text{ad}}(x^{-\nu}) = \left\{x^\nu \in X^\nu \;\middle|\; (x^\nu, x^{-\nu}) \in X_{\text{ad}}\right\}$$

with $X_{\text{ad}} \subseteq X$.

iii) *jointly convex GNEP*, if $\Gamma$ is a GNEP of joint type and each player's minimization problem is convex (see Assumption 2.3.7).

iv) *potential game*[16], if $\Gamma$ is a GNEP of joint type and there exists a continuous functional $p : X \to \mathbb{R}$ such that for all $\nu = 1, \dots, N$ and all $x^{-\nu} \in X_{-\nu}$ with $X^\nu_{\text{ad}}(x^{-\nu})$ nonempty, we obtain that for all $y^\nu, z^\nu \in X^\nu_{\text{ad}}(x^{-\nu})$

$$f^\nu(y^\nu, x^{-\nu}) - f^\nu(z^\nu, x^{-\nu}) > 0$$

implies

$$p(y^\nu, x^{-\nu}) - p(z^\nu, x^{-\nu}) \geq \sigma(f^\nu(y^\nu, x^{-\nu}) - f^\nu(z^\nu, x^{-\nu})),$$

where $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ is a forcing function[11].

The advantage of potential games is that the issue of solving an equilibrium problem is reduced to a single optimization problem. Its requirements are for instance satisfied, if each player's objective function is independent of his rivals strategies or is given with respect to a common term for all players, i.e.

$$f^\nu(x) \equiv f^\nu(x^\nu) \qquad \text{or} \qquad f^\nu(x) = f_1(x) + f_2^\nu(x^\nu).$$

In this context, $p$ is given by $p(x) = \sum_{\nu=1}^N f^\nu(x^\nu)$ and $p(x) = f_1(x) + \sum_{\nu=1}^N f_2^\nu(x^\nu)$, respectively.
Due to the classification in Definition 2.3.14, we can distinguish between several equilibrium concepts.

**Definition 2.3.15.** The point $\bar{x}$ is called

i) *Nash equilibrium* (NE) of GNEP $\Gamma$, if $\bar{x} \in X_{\text{ad}}(\bar{x})$ and for all $\nu = 1, \dots, N$ it holds

$$f^\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq f^\nu(x^\nu, \bar{x}^{-\nu}) \qquad \forall x^\nu \in X^\nu_{\text{ad}}(\bar{x}^{-\nu}); \tag{2.24}$$

[11] A function $\sigma : \mathbb{R}^+ \to \mathbb{R}^+$ is called *forcing function*, if $\lim\limits_{k\to\infty} \sigma(t^k) = 0$ implies that $\lim\limits_{k\to\infty} t^k = 0$.

ii) *normalized Nash equilibrium* (NNE), if $\Gamma$ is of shared type and it holds

$$\sum_{\nu=1}^{N} f^\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \sum_{\nu=1}^{N} f^\nu(x^\nu, \bar{x}^{-\nu}) \qquad \forall x \in X_{\text{ad}}; \tag{2.25}$$

iii) *stationary Nash equilibrium* of $\Gamma$, if for all $\nu = 1, \dots, N$ $f^\nu$ is continuously differentiable and it holds

$$-D_{x^\nu} f^\nu(\bar{x}) \in \mathcal{T}_{X^\nu_{\text{ad}}(\bar{x}^{-\nu})}(\bar{x}^\nu)^\circ. \tag{2.26}$$

**Remark 2.3.16.** *Both equilibrium concepts, NE and NNE, in Definition 2.3.15 are given in the global sense. However, a transfer on a local level in the sense of Definition 2.3.1 is straight forward.*

In an equilibrium point $\bar{x}$ player $\nu$'s strategy $\bar{x}^\nu$ is optimal with respect to his opponents' strategies $\bar{x}^{-\nu}$ and hence, no player has the incentive of an unilateral change of his strategy, if his opponents use their Nash equilibrium strategy. Consequently, if $\bar{x}$ is a NE of $\Gamma$, then

$$\bar{x}^\nu \in \operatorname{argmin} \left\{ f^\nu(x^\nu, \bar{x}^{-\nu}) \;\middle|\; x^\nu \in X^\nu_{\text{ad}}(\bar{x}^{-\nu}) \right\}$$

for all $\nu = 1, \dots, N$.

During the remainder of this section we focus on existence results of GNEs. For proofs of existence concerning NEs of classical Nash equilibirum problems, we refer to the monographs [43, 68]. However, notice that both references focus on the finite dimensional setting and hence, certain requirements, i.e. compactness and continuity assumptions have to be adapted in general Banach spaces. Coming back to general GNEPs, first assume that $\Gamma$ is a potential game. In this context, it follows that a solution of the minimization problem

$$\min\; p(x) \quad \text{s.t. } x \in X_{\text{ad}},$$

where $p$ denotes the potential function and $X_{\text{ad}}$ the shared feasible set, is a NE of $\Gamma$. Hence, the existence can be ensured by results presented in Subsection 2.3.1. However, notice that it is eminent that $p$ and $X_{\text{ad}}$ are convex (see Corollary 2.3.4).

Next, we consider $\Gamma$ as a GNEP of joint type. Due to its structure, it is known (see [26, Lemma 2.1]) that $x \in X_{\text{ad}}(x) := \prod_{\nu=1}^{N} X^\nu_{\text{ad}}(x^{-\nu})$ is equivalent to $x \in X_{\text{ad}}$ and $X_{\text{ad}}(x)$ is nonempty for all $x \in X_{\text{ad}}$. Hence, we see that we can deal with the set $X_{\text{ad}}$ instead of $X_{\text{ad}}(x)$ and the existence of a Nash equilibrium can be traced back to the existence of an NNE. In order to see this, assume that $\bar{x} \in X_{\text{ad}}$ satisfies (2.25). Thus, (2.25) is in particular valid for all $(x^\nu, \bar{x}^{-\nu}) \in X_{\text{ad}}$, where $\nu \in \{1, \dots, N\}$ is arbitrary but fixed, which

reduces to (2.24).
In this context, the so-called *Nikaido-Isoda-function* or *Ky-Fan-function*

$$\Psi : X \times X \to \mathbb{R},\ (x,y) \mapsto \Psi(x,y) := \sum_{\nu=1}^{N} f^\nu(x^\nu, x^{-\nu}) - f^\nu(y^\nu, x^{-\nu})$$

is widely used (see e.g. [34, 45]). In particular, observe that with Definition 2.3.15 it holds $\Psi(\bar{x}, y) \leq 0$ for all $y \in X_{\text{ad}}$, if $\bar{x}$ is a normalized Nash equilibrium and $\Psi(x,x) = 0$ for all $x \in X_{\text{ad}}$. Defining

$$V : X \to \mathbb{R},\ x \mapsto V(x) := \sup \left\{ \Psi(x,y) \;\middle|\; y \in X_{\text{ad}} \right\},$$

the latter implies that $V(x) \geq 0$ and hence, it follows that $\bar{x}$ is a normalized Nash equilibrium if and only if $\bar{x} \in X_{\text{ad}}$ and $V(\bar{x}) = 0$. The following result summarizes the discussion above and can also be found in [34, Section 3.1].

**Theorem 2.3.17.** *Let GNEP $\Gamma$ be jointly convex. Then the subsequent statements are equivalent:*

*a)* $\bar{x}$ *is a normalized Nash equilibrium of* $\Gamma$.

*b)* $\bar{x} \in X_{\text{ad}}$ *and* $V(\bar{x}) = 0$.

*c)* $\bar{x} \in \hat{y}(\bar{x})$ *with* $\hat{y}(x) := \operatorname{argmax}\{\Psi(x,y) \mid y \in X_{\text{ad}}\}$.

*d)* $\bar{x}$ *is the solution of*

$$\min \sum_{\nu=1}^{N} f^\nu(y^\nu, \bar{x}^{-\nu}) \quad \text{w.r.t. } y \in X^\nu, \quad \text{s.t. } y \in X_{\text{ad}}. \tag{2.27}$$

In order to prove the existence of a (normalized) Nash equilibrium, Theorem 2.3.17 indicates several possible approaches.
In [34, Theorem 3.4], the authors used Kakutani's fixed point theorem to show that $\hat{y}(x)$ has a fixed point, while a different ansatz that is based on the Ky-Fan theorem (see [17, Theorem 1]) was used for instance in [40, Theorem 1.1] and [45, Theorem 2.3] to tackle point b). Again we want to emphasize the prominent role of convexity in all referenced results.
Finally, assume that $\Gamma$ is a classical GNEP presented in the beginning of this section without

further joint structure. Under sufficient smoothness assumptions on the objective functionals $f^\nu$ ($\nu = 1, \ldots, N$), first-order optimality conditions are derived via Theorem 2.3.5. Hence, let $\bar{x}$ be a Nash equilibrium of $\Gamma$, then $\bar{x}$ satisfies

$$\text{Find } x \in X_{\text{ad}}(x) \quad \text{s.t.} \quad \langle F(x), d\rangle_X \geq 0 \quad \forall d \in \prod_{\nu=1}^{N} \mathcal{T}_{X_{\text{ad}}^\nu(x^{-\nu})}(x^\nu) \tag{2.28}$$

with $F(x) := (D_{x^\nu} f^\nu(x))_{\nu=1}^N$. In particular, notice that $\bar{x}$ satisfying (2.28) is equivalent by Definition 2.3.15 to $\bar{x}$ being a stationary Nash equilibrium of $\Gamma$. If $X_{\text{ad}}^\nu(x^{-\nu})$ is convex-valued for all $\nu = 1, \ldots, N$, i.e. convex for all $x^{-\nu} \in X_{-\nu}$, then (2.28) can further be written as

$$\text{Find } x \in X_{\text{ad}}(x) \quad \text{s.t.} \quad \langle F(x), y - x\rangle \geq 0 \quad \forall y \in X_{\text{ad}}(x). \tag{2.29}$$

In reverse, a solution of (2.29) is a generalized Nash equilibrium, if $f^\nu$ is additional player-convex (see e.g. [83, Example 6.1]). Inequalities of the form (2.28) and (2.29) are both known as *quasi-variational inequalities* (QVI) and are of non-trivial type. The existence of a solution of (2.29) was shown for instance in [44, Theorem 4.4], which also requires convexity among other non-trivial assumptions.

# 3. Equilibrium Problems with Equilibrium Constraints in Banach Spaces

The following chapter focuses on equilibrium problems with equilibrium constraints (EPEC) and therein, on the special case of equilibrium problems with complementarity constraints (EPCC) in Banach spaces, which are characterized by a specific structure of the underlying feasible set that result in a non-convex program. As we have shown in Subsection 2.3.2, the lack of convexity is crucial, since existence results fail to hold in this case. Therefore, we have to deal with weaker Nash equilibrium concepts.

Up to now, EPCCs/EPECs were analyzed in the literature mainly in finite dimensional settings (see e.g. [9, 24, 38, 60, 84, 85]). Therein, the authors [9, 24, 85] used generalized tools of variational analysis in order to tackle the general form of equilibrium constraints and derive stationarity concepts. The only contribution, we are aware of in abstract Banach spaces is [61], which can be seen as an extension of [60] and focused on necessary optimality conditions on the basis of generalized differentiation theory. In contrast, our approach utilizes results in [53, 56, 95] on mathematical programs with complementarity constraints (MPCC) in general Banach spaces and follows similar to [84] in finite dimensions the idea to extend these results to equilibrium problems.

In regard to the numerical analysis, we derive sequences of Nash equilibrium problems that converge to the corresponding EPEC/EPCC. We emphasize that the idea was also used in [38], where the equilibrium constraints are given by a P-matrix linear complementarity system in finite spaces. Therein, the authors showed the convergence of the auxiliary problems towards the original EPCC. We extend this approach by considering the analogon in abstract Banach spaces by using a special penalization technique applied to the equilibrium constraints.

For this reason, this chapter is organized as follows. In Section 3.1, we introduce EPECs formally and classify them in the context of GNEPs. Subsequently, Section 3.2 is devoted to the underlying programs of each player, i.e. the MPCCs. Finally, those results are extended in Section 3.3 and we introduce two appropriate auxiliary NEPs in order to find stationary Nash equilibria of the original EPEC.

## 3.1. Problem Formulation

Using the same notation as introduced in Subsection 2.3.2, we consider throughout this section a special type of a generalized Nash equilibrium problem with $N$ players, denoted by $\Gamma_{\text{EPEC}}$, where player $\nu$ solves the following nonlinear program

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}) \\ \text{w.r.t.}\ & x^\nu \in X_\nu, \\ \text{s.t.}\ & x^\nu \in Q_\nu,\ g^\nu(x^\nu, x^{-\nu}) \in C_\nu \\ & 0 \in h^\nu(x^\nu, x^{-\nu}) + \mathcal{H}^\nu(x^\nu, x^{-\nu}). \end{aligned} \tag{3.1}$$

Here, the condition

$$0 \in h^\nu(x^\nu, x^{-\nu}) + \mathcal{H}^\nu(x^\nu, x^{-\nu}) \tag{3.2}$$

is known as *equilibrium constraint* and is in general comprised of a mapping $h^\nu : X \to Z_\nu^*$ and a set-valued mapping $\mathcal{H}^\nu : X \rightrightarrows Z_\nu^*$ between a Banach space $X$ and the dual space of a Banach space $Z_\nu$ for all $\nu = 1, \dots, N$. For this reason, we call $\Gamma_{\text{EPEC}}$ an *equilibrium problem with equilibrium constraints*. The representation (3.2) is quite general and we refer to [61, 63] for a detailed analysis of these problem classes. In our case, we concentrate on a more popular version. For this reason, we set $h^\nu(x) := -H^\nu(x)$ and $\mathcal{H}^\nu(x) := \mathcal{N}_{K_\nu}(G^\nu(x))$, where $K_\nu \subseteq Z_\nu$ is an arbitrary closed, convex set and observe that (3.2) is equivalent to

$$H^\nu(x) \in \mathcal{N}_{K_\nu}(G^\nu(x)). \tag{3.3}$$

The latter condition can be interpreted as the first-order optimality condition of a lower level mathematical program or stationary Nash equilibrium condition of a jointly convex GNEP, respectively, where $H^\nu$ denotes the corresponding derivative.
If $K_\nu$ is additionally a cone for all $\nu = 1, \dots, N$, then it follows by Section 2.2 that (3.3) is equivalent to

$$G^\nu(x) \in K_\nu, \quad H^\nu(x) \in K_\nu^\circ, \quad \langle H^\nu(x), G^\nu(x)\rangle_Z = 0, \tag{3.4}$$

that is known as *complementarity constraints*.

**Remark 3.1.1.** *Let $G^N : \mathbb{R}^n \to \mathbb{R}^m$ and $H^N : \mathbb{R}^n \to \mathbb{R}^m$ be arbitrary mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then a point $\bar{x} \in \mathbb{R}^n$ satisfies the complementarity conditions in finite dimensions, if for all $j = 1, \dots, m$ we have*

$$G_j^N(\bar{x}) \geq 0, \quad H_j^N(\bar{x}) \geq 0, \quad G_j^N(\bar{x}) \cdot H_j^N(\bar{x}) = 0.$$

*Recalling that any closed, convex cone induces the order relation '$\leq_K'$ (see Section 2.2), we obtain that $G(x) \geq_K 0$ and hence, the first constraint in* (3.4) *can be seen as a non-negativity. Since it holds $\langle H(x), z\rangle_Z \leq 0$ for all $z \in K$ by definition, the second condition in* (3.4) *can be interpreted as $H(x)$ being non-positive in the dual sense. Finally, the last condition ensures that the product is zero.*

As a result, $\Gamma_{\text{EPEC}}$ can be written as the subsequent *equilibrium problem with complementarity constraints*, denoted by $\Gamma_{\text{EPCC}}$, where player $\nu$ solves

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}) \\ \text{w.r.t. } & x^\nu \in X_\nu, \\ \text{s.t. } & x^\nu \in Q_\nu,\ g^\nu(x^\nu, x^{-\nu}) \in C_\nu, \\ & G^\nu(x^\nu, x^{-\nu}) \in K_\nu,\ H^\nu(x^\nu, x^{-\nu}) \in K_\nu^\circ, \\ & \langle H^\nu(x^\nu, x^{-\nu}), G^\nu(x^\nu, x^{-\nu})\rangle_{Z^\nu} = 0. \end{aligned} \tag{3.5}$$

We denote the feasible set of player $\nu$'s problem by

$$\begin{aligned} X_{\text{ad}}^\nu(x^{-\nu}) := \big\{x^\nu \in Q_\nu \mid g^\nu(x^\nu, x^{-\nu}) \in C_\nu,\ G^\nu(x^\nu, x^{-\nu}) \in K_\nu,\ H^\nu(x^\nu, x^{-\nu}) \in K_\nu^\circ, \\ \langle H^\nu(x^\nu, x^{-\nu}), G^\nu(x^\nu, x^{-\nu})\rangle_{Z^\nu} = 0\big\} \end{aligned}$$

and observe that $\Gamma_{\text{EPCC}}$ is

- a GNEP of joint type, if for all $\nu = 1, \dots, N$ $(g^\nu, G^\nu, H^\nu) \equiv (g, G, H)$ and $(C_\nu, K_\nu) \equiv (C, K)$;
- convex, if for all $\nu = 1, \dots, N$ $(g^\nu, G^\nu, H^\nu)$ is $(-C_\nu, -K_\nu, -K_\nu^\circ)$-convex and $\langle H^\nu(x), G^\nu(x)\rangle$ is linear.

In particular, the latter property is not satisfied in general, as this would require that one entry is constant while the other one is linear. As a consequence, we are in general not able to prove the existence of a generalized Nash equilibrium of $\Gamma_{\text{EPEC}}$ (see the discussion in Subsection 2.3.2).
Additionally, it is known (see [56, Lemma 3.1]) that any feasible point of $\Gamma_{\text{EPCC}}$ violates constraint qualifications of suitable strength and hence, necessary first-order optimality conditions are in general too strong. Both the missing convexity and the absent connection between a Nash equilibrium and the solution of a corresponding KKT system can be traced back to each player's MPCC. Thus, we proceed by analyzing these special nonlinear problems in the next step and transfer the results to the extended setting afterwards.

## 3.2. Introduction to Mathematical Programs with Complementarity Constraints

In this section we follow the work [94] on MPCCs in general (reflexive) Banach spaces and constitutive publications (see [19, 53, 56]) and consider the MPCC

$$\begin{aligned} \min\ & f(x) \\ \text{w.r.t.}\ & x \in X, \\ \text{s.t.}\ & x \in Q,\ g(x) \in C, \\ & G(x) \in K,\ H(x) \in K^\circ,\ \langle H(x), G(x)\rangle = 0, \end{aligned} \tag{3.6}$$

where we require that $f : X \to \mathbb{R}$, $g : X \to W$, $G : X \to Z$, $H : X \to Z^*$ and the sets $Q \subseteq X$, $C \subseteq W$ and $K \subseteq Z$ satisfy the following assumption.

**Assumption 3.2.1.** *Let $W$, $X$ and $Z$ be reflexive Banach spaces and assume that the mappings $f : X \to \mathbb{R}$, $g : X \to W$, $G : X \to Z$ and $H : X \to Z^*$ are continuously differentiable. Moreover, let $Q \subseteq X$ be non-empty, closed, convex and bounded, $C \subseteq W$ be non-empty, closed and convex and assume that $K \subseteq Z$ is a non-empty, closed and convex cone.*

Furthermore, we define the feasible set of (3.6) by

$$X_{\text{ad}}^{\text{MPCC}} := \big\{x \in Q \;\big|\; g(x) \in C,\ G(x) \in K,\ H(x) \in K^\circ,\ \langle H(x), G(x)\rangle_Z = 0\big\},$$

which is in general non-convex as already stated above. Nevertheless, we conclude by Theorem 2.3.5 that a local solution $\bar{x}$ of MPCC (3.6) satisfies $\bar{x} \in X_{\text{ad}}^{\text{MPCC}}$ and

$$\langle f'(\bar{x}), d\rangle \geq 0 \qquad \forall d \in \mathcal{T}_{X_{\text{ad}}^{\text{MPCC}}}(\bar{x}). \tag{3.7}$$

By similar argumentation as in Subsection 2.3.1, we are interested in a *KKT-like* first-order optimality condition. In this context, notice that a constraint qualification is needed. Unfortunately and as already indicated in Section 3.1, this is not the case for any feasible point of MPCC (3.6). Due to its importance, we recall the result that was given for instance in [56, Lemma 3.1] for the case $Q = X$.

**Lemma 3.2.2.** *The constraint qualification KRZCQ is violated at any feasible point $\bar{x} \in X_{\text{ad}}^{\text{MPCC}}$.*

*Proof.* Let $\bar{x} \in X_{\text{ad}}^{\text{MPCC}}$ be an arbitrary feasible point of MPCC (3.6) and assume that *KRZCQ* is satisfied, i.e. it holds

$$\tilde{G}'(\bar{x})[\mathcal{R}_Q(\bar{x})] - R_{\tilde{C}}(\tilde{G}(\bar{x})) = V$$

with

- $\tilde{G}(x) := (g(x), G(x), H(x), \langle H(x), G(x)\rangle)$,
- $\tilde{C} := C \times K \times K^\circ \times \{0\}$,
- $V := W \times Z \times Z^* \times \mathbb{R}$.

As stated in Lemma 2.3.13, KRZCQ implies that the set of singular Lagrange multipliers contains only the zero element, i.e. it holds

$$\big\{(\eta, \kappa) \in \mathcal{N}_{\tilde{C}}(\tilde{G}(\bar{x})) \times \mathcal{N}_Q(\bar{x}) \;\big|\; \tilde{G}'(\bar{x})^*[\eta] + \kappa = 0\big\} = \big\{(0,0)\big\}.$$

By definition, a singular Lagrange multiplier

$$(\kappa, \xi, \mu, \lambda, \rho) \in \mathcal{N}_Q(\bar{x}) \times \mathcal{N}_C(g(\bar{x})) \times K^\circ \times K \times \mathbb{R}$$

of MPCC (3.6) solves

$$\begin{aligned} g'(\bar{x})^*[\xi] + G'(\bar{x})^*[\mu] + H'(\bar{x})^*[\lambda] + \rho(G'(\bar{x})^*[H(\bar{x})] + H'(\bar{x})^*[G(\bar{x})]) + \kappa = 0, \\ \langle \mu, G(\bar{x})\rangle = \langle \lambda, H(\bar{x})\rangle = 0. \end{aligned} \tag{3.8}$$

Since the first condition is equivalent to

$$g'(\bar{x})^*[\xi] + G'(\bar{x})^*[\mu + \rho H(\bar{x})] + H'(\bar{x})^*[\lambda + \rho G(\bar{x})] + \kappa = 0,$$

we see that the vector $(\bar{\kappa}, \bar{\xi}, \bar{\mu}, \bar{\lambda}, \bar{\rho}) := (0, 0, H(\bar{x}), G(\bar{x}), -1)$ is a solution of (3.8) that is obviously not equal to zero. Hence, *KRZCQ* is not satisfied in $\bar{x}$. Since $\bar{x}$ was chosen arbitrarily, the claim is proven. □

As a consequence of Lemma 3.2.2, KKT-conditions as first-order optimality conditions are too strong. Furthermore, the reverse direction in Theorem 2.3.9 is also not valid by the non-convexity of $X_{\text{ad}}^{\text{MPCC}}$. Therefore, we have to deal with weaker stationarity concepts. Recall, that in finite dimensions weakly and strongly stationary points (W- and S-stationarity) are introduced (see [78]) as KKT-points of some auxiliary problems. In this context, the strongly active, the inactive and the biactive set are defined, respectively. Unfortunately, those are not available for MPCCs in general Banach spaces. To circumvent this issue, the author [95] defined two auxiliary problems. For this reason, let $\bar{x} \in X_{\text{ad}}^{\text{MPCC}}$ be feasible and consider the *relaxed nonlinear program* (RNLP)

$$\begin{aligned} \min\ & f(x) \\ \text{w.r.t.}\ & x \in X, \\ \text{s.t.}\ & x \in Q,\ g(x) \in C, \\ & G(x) \in \mathcal{N}_{K^\circ}(H(\bar{x})),\ H(x) \in \mathcal{N}_K(G(\bar{x})) \end{aligned} \tag{3.9}$$

and the *tightened nonlinear program* (TNLP)

$$
\begin{aligned}
\min\ & f(x) \\
\text{w.r.t.}\ & x \in X, \\
\text{s.t.}\ & x \in Q,\ g(x) \in C, \\
& G(x) \in \mathcal{N}_{K^\circ}(H(\bar{x})) \cap \mathcal{N}_K(G(\bar{x}))^\perp, \\
& H(x) \in \mathcal{N}_K(G(\bar{x})) \cap \mathcal{N}_{K^\circ}(H(\bar{x}))^\perp.
\end{aligned} \tag{3.10}
$$

Now, it is clear that $\bar{x}$ is feasible for both auxiliary problems. Moreover, notice that $X_{\text{ad}}^{\text{TNLP}} \subseteq X_{\text{ad}}^{\text{MPCC}}$ is also valid, where $X_{\text{ad}}^{\text{TNLP}}$ denotes the feasible set of TNLP (3.10). In order to see the latter inclusion, let $\tilde{x} \in X_{\text{ad}}^{\text{TNLP}}$ be arbitrary but fixed. Then we have by definition $\tilde{x} \in Q$, $g(\tilde{x}) \in C$ and

$$
\begin{aligned}
G(\tilde{x}) &\in \mathcal{N}_{K^\circ}(H(\bar{x})) \cap \mathcal{N}_K(G(\bar{x}))^\perp = K \cap \{H(\bar{x})\}^\perp \cap \mathcal{N}_K(G(\bar{x}))^\perp, \\
H(\tilde{x}) &\in \mathcal{N}_K(G(\bar{x})) \cap \mathcal{N}_{K^\circ}(H(\bar{x}))^\perp = K^\circ \cap \{G(\bar{x})\}^\perp \cap \mathcal{N}_{K^\circ}(H(\bar{x}))^\perp.
\end{aligned}
$$

So in particular, it holds that $(G(\tilde{x}), H(\tilde{x})) \in K \times K^\circ$. Since we additionally have $(G(\tilde{x}), H(\tilde{x})) \in \mathcal{N}_{K^\circ}(H(\bar{x})) \times \mathcal{N}_{K^\circ}(H(\bar{x}))^\perp$, the dual pairing formulation vanishes, which implies $\tilde{x} \in X_{\text{ad}}^{\text{MPCC}}$. Hence, if $\bar{x}$ is a local solution of MPCC (3.6), $\bar{x}$ is also a local solution of TNLP (3.10). However, it holds neither $X_{\text{ad}}^{\text{RNLP}} \subseteq X_{\text{ad}}^{\text{MPCC}}$ nor $X_{\text{ad}}^{\text{MPCC}} \subseteq X_{\text{ad}}^{\text{RNLP}}$ in general. Similar to the finite dimensional setting, W- and S-stationary points are the KKT-points of TNLP (3.10) and RNLP (3.9), respectively. In this context, it was proven in [56, Lemma 3.2] that it holds

$$
\begin{aligned}
\operatorname{cl}(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp &= \mathcal{N}_{\mathcal{N}_{K^\circ}(H(\bar{x})) \cap \mathcal{N}_K(G(\bar{x}))^\perp}(G(\bar{x})), \\
\operatorname{cl}(K - K \cap \{H(\bar{x})\}^\perp) \cap \{H(\bar{x})\}^\perp &= \mathcal{N}_{\mathcal{N}_{K^\circ}(H(\bar{x})) \cap \mathcal{N}_K(G(\bar{x}))^\perp}(H(\bar{x})),
\end{aligned}
$$

where the *Minkowski sum*[12] is used. Since the intersection operator is stronger than the Minkowski sum, we omit the brackets in the following and write for instance

$$
K - K \cap \{H(\bar{x})\}^\perp \equiv K - (K \cap \{H(\bar{x})\}^\perp).
$$

Analogously, one can show that

$$
\begin{aligned}
\mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x})) &= \mathcal{N}_{\mathcal{N}_{K^\circ}(H(\bar{x}))}(G(\bar{x})), \\
\mathcal{K}_K(G(\bar{x}), H(\bar{x})) &= \mathcal{N}_{\mathcal{N}_K(G(\bar{x}))}(H(\bar{x}))
\end{aligned}
$$

[12] For two arbitrary sets $\mathcal{A}, \mathcal{B} \subseteq X$, the Minkowski sum is defined by $\mathcal{A} + \mathcal{B} := \{a + b \mid a \in \mathcal{A},\ b \in \mathcal{B}\}$, where we set $\mathcal{A} - \mathcal{B} := \mathcal{A} + (-\mathcal{B})$.

are satisfied, which follows by the definition of the critical cone (see Definition 2.2.2) and that $\mathcal{N}_{K^\circ}(H(\bar{x}))$ is closed and convex, i.e.

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

The discussion above is summarized in the definition below.

**Definition 3.2.3** ([56, Definition 3.3])**.** A feasible point $\bar{x} \in X$ of MPCC (3.6) is called

- *weakly stationary* (W-stationary), if there exist multipliers

  $$(\kappa, \xi, \mu, \lambda) \in X^* \times W^* \times Z^* \times Z$$

  such that the system

  $$0 = f'(\bar{x}) + g'(\bar{x})^*\xi + G'(\bar{x})^*\mu + H'(\bar{x})^*\lambda + \kappa, \tag{3.11a}$$
  $$\kappa \in \mathcal{N}_Q(\bar{x}), \tag{3.11b}$$
  $$\xi \in \mathcal{N}_C(g(\bar{x})), \tag{3.11c}$$
  $$\mu \in \mathrm{cl}(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp, \tag{3.11d}$$
  $$\lambda \in \mathrm{cl}(K - K \cap \{H(\bar{x})\}^\perp) \cap \{H(\bar{x})\}^\perp \tag{3.11e}$$

  is satisfied.

- *strongly stationary* (S-stationary), if there exist multipliers

  $$(\kappa, \xi, \mu, \lambda) \in X^* \times W^* \times Z^* \times Z$$

  such that (3.11a)-(3.11c) hold and the system

  $$\mu \in \mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x})), \tag{3.12a}$$
  $$\lambda \in \mathcal{K}_K(G(\bar{x}), H(\bar{x})) \tag{3.12b}$$

  is satisfied.

At this point, we refer to [56, Lemma 3.4], which showed the expected implication

$$\text{S-stationarity} \quad \Longrightarrow \quad \text{W-stationarity.}$$

Moreover, notice that in contrast to MPCCs in $\mathbb{R}^n$ S-stationarity is in general strictly weaker than being a KKT-point (see [94, Lemma 4.2] and the corresponding discussion). However, it is possible to derive a certain type of first-order stationary condition under additional assumptions. Therefore, it was shown in [94, Lemma 4.5] that a strongly stationary point $\bar{x}$ is *linearized B-stationary*, i.e.

$$\langle f'(\bar{x}), d\rangle \geq 0 \qquad \forall d \in \mathcal{T}^{\text{lin}}_{X_{\text{ad}}^{\text{MPCC}}}(\bar{x}),$$

if $K$ is polyhedric, where $\mathcal{T}^{\text{lin}}_{X_{\text{ad}}^{\text{MPCC}}}(\bar{x})$ is the linearized tangent cone (see (2.17)) of $X_{\text{ad}}^{\text{MPCC}}$ with respect to $\bar{x}$. Hence and similar to local solutions, it is guaranteed that there do not exist any descent directions in a strongly stationary point under polyhedricity, as we have $\mathcal{T}_{X_{\text{ad}}^{\text{MPCC}}}(\bar{x}) \subseteq \mathcal{T}^{\text{lin}}_{X_{\text{ad}}^{\text{MPCC}}}(\bar{x})$.
Coming back to finite dimensional MPCCs once more, it is expected that there are more stationarity concepts than W- and S-stationarity. In order to define M-stationary (Mordukhovich stationary) points, observe that we have

$$G(x) \in K,\ H(x) \in K^{\circ},\ \langle H(x), G(x)\rangle = 0 \iff (G(x), H(x)) \in \text{gph}\, \mathcal{N}_K$$

due to the convexity of $K$ with $\text{gph}\, \mathcal{N}_K := \{(z, z^*) \in Z \times Z^* \mid z^* \in \mathcal{N}_K(z)\}$.

**Definition 3.2.4** ([19, Definition 4.1])**.** A feasible point $\bar{x} \in X$ of MPCC (3.6) is called *M-stationary*, if there exist multipliers $(\kappa, \xi, \mu, \lambda) \in X^* \times Y^* \times Z^* \times Z$ such that (3.11a)-(3.11c) hold and

$$(\mu, \lambda) \in \mathcal{N}_{\text{gph}\, \mathcal{N}_K}(G(\bar{x}), H(\bar{x})) \tag{3.13}$$

is satisfied.

Notice that (3.13) is given with respect to the limiting normal cone of the non-convex set $\text{gph}\, \mathcal{N}_K$, which reduces to the well-known concept of M-stationarity in $\mathbb{R}^n$ (see [18, Proposition 2.4]). Hence, the naming is appropriate.
Unfortunately and in contrast to W- and S-stationarity, the integration of M-stationarity into the chain of implication does not hold without further assumptions on the cone $K$ in general.
In [53, Lemma 3.9], it was shown that the identity

$$\hat{\mathcal{N}}_{\text{gph}\, \mathcal{N}_K}(G(x), H(x)) = \mathcal{K}_{K^{\circ}}(H(x), G(x)) \times \mathcal{K}_K(G(x), H(x))$$

is satisfied, if $K$ is polyhedric. Since the Fréchet normal cone is a subset of the limiting normal cone, we conclude by (2.5) that any S-stationary point is also M-stationary under polyhedricity.

In order to prove that an M-stationary point implies W-stationarity, we refer to [53, Lemma 3.10], which showed that

$$\begin{aligned}\mathcal{N}_{\text{gph } N_K}(G(x), H(x)) \subseteq \text{cl}(K^\circ - K^\circ \cap \{G(\bar{x})\}^\perp) \cap \{G(\bar{x})\}^\perp \times \\ \text{cl}(K - K \cap \{H(\bar{x})\}^\perp) \cap \{H(\bar{x})\}^\perp\end{aligned}$$

is valid, if $K$ and $K^\circ$ are additionally pointed, induce a vector lattice (see Section 2.2) and the following continuity properties

$$\begin{aligned}&\max_{K^\circ}(0,\cdot) : Z^* \to Z^* \text{ is continuous or} \\ &\max_K(0,\cdot) : Z \to Z \text{ is weakly continuous,}\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}&\max_{K^\circ}(0,\cdot) : Z^* \to Z^* \text{ is weakly continuous or} \\ &\max_K(0,\cdot) : Z \to Z \text{ is continuous,}\end{aligned} \tag{3.15}$$

are satisfied. Consequently, we end up with the following chain of inclusions

$$\text{S-stationarity} \quad \Longrightarrow \quad \text{M-stationarity} \quad \Longrightarrow \quad \text{W-stationarity,} \tag{3.16}$$

where we use again (2.5).

**Example 3.2.5.** *Let $X = H_0^1(\Omega)$ and consider the cone*

$$K = H_0^1(\Omega)_0^+ = \{y \in H_0^1(\Omega) \mid y \geq 0 \quad \text{a.e. in } \Omega\},$$

*which we introduced in Example 2.2.8. In particular, we know that $K$ is polyhedric and pointed. Moreover, we obtain by [7, Proposition 6.45] and [96, Lemma 4.1] that* $\max_K(0,\cdot)$ *is continuous and weakly continuous, respectively. Hence, the inclusion* (3.16) *is valid in* $H_0^1(\Omega)$.

In order to get necessary conditions for a local solution $\bar{x}$ of MPCC (3.6) to be an S- or W-stationary point, the derivation of the latter motivates to apply the first-order optimality conditions of nonlinear programming on the auxiliary problems. For this reason, the following theorem summarizes results from [56, 94] and extends them to the more general case $X = Q$.

**Theorem 3.2.6.** *Let $\bar{x}$ be a local solution of MPCC* (3.6). *If $\bar{x}$ satisfies*

$$\begin{pmatrix} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{pmatrix} [\mathcal{R}_Q(\bar{x})] - \begin{pmatrix} \mathcal{R}_C(g(\bar{x})) \\ \mathcal{R}_K(G(\bar{x})) \cap (-\mathcal{K}_K(G(\bar{x}), H(\bar{x}))) \\ \mathcal{R}_{K^\circ}(H(\bar{x})) \cap (-\mathcal{K}_{K^\circ}(H(\bar{x}), G(\bar{x}))) \end{pmatrix} = \begin{bmatrix} W \\ Z \\ Z^* \end{bmatrix}, \tag{3.17}$$

*then $\bar{x}$ is W-stationary.*
*If $\bar{x}$ satisfies at least one of the subsequent conditions, i.e.*

*i)* (3.17) *and*

$$\operatorname{cl}\left(\left(\begin{array}{c} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{array}\right)[\mathcal{N}_Q(\bar{x})_\perp] - \left(\begin{array}{c} \mathcal{N}_C(g(\bar{x}))_\perp \\ \mathcal{N}_{K\cap(-\mathcal{T}_K(G(\bar{x})))}(G(\bar{x}))^\perp \\ \mathcal{N}_{K^\circ\cap(-\mathcal{T}_{K^\circ}(H(\bar{x})))}(H(\bar{x}))^\perp \end{array}\right)\right) = \left[\begin{array}{c} W \\ Z \\ Z^* \end{array}\right], \tag{3.18}$$

*ii)*

$$\left(\begin{array}{c} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{array}\right)[\mathcal{R}_Q(\bar{x})] - \left(\begin{array}{c} \mathcal{R}_C(g(\bar{x})) \\ \mathcal{R}_K(G(\bar{x})) \cap \{H(\bar{x})\}^\perp \\ \mathcal{R}_{K^\circ}(H(\bar{x})) \cap \{G(\bar{x})\}^\perp \end{array}\right) = \left[\begin{array}{c} W \\ Z \\ Z^* \end{array}\right], \tag{3.19}$$

*iii)*

$$\left(\begin{array}{c} g'(\bar{x}) \\ G'(\bar{x}) \\ H'(\bar{x}) \end{array}\right)[\mathcal{R}_Q(\bar{x})] = \left[\begin{array}{c} W \\ Z \\ Z^* \end{array}\right], \tag{3.20}$$

*then $\bar{x}$ is S-stationary.*

*Proof.* In order to verify the first part, observe that (3.17) equals KRZCQ for TNLP (3.10), which was shown in [56, Theorem 3.6].
For a proof of i), we refer to [94, Theorem 4.6 and Prop. 4.7] for the case that $Q = X$. However, an extension to our more general setting is straightforward. In particular, notice that (3.18) coincides with (2.21) applied to TNLP (3.10).
Next, the condition (3.19) equals KRZCQ for RNLP (3.9). In fact, we have $G(\bar{x}) \in \{H(\bar{x})\}^\perp$ and hence, can apply [56, Lemma 2.2] to deduce that

$$\begin{aligned} \mathcal{R}_{\mathcal{N}_{K^\circ}(H(\bar{x}))}(G(\bar{x})) &= N_{K^\circ}(H(\bar{x})) + \operatorname{lin}(G(\bar{x})) = K \cap \{H(\bar{x})\}^\perp + \operatorname{lin}(G(\bar{x})) \\ &= (K + \operatorname{lin}(G(\bar{x}))) \cap \{H(\bar{x})\}^\perp = \mathcal{R}_K(G(\bar{x})) \cap \{H(\bar{x})\}^\perp. \end{aligned}$$

Similarly, one shows that $\mathcal{R}_{\mathcal{N}_K(G(\bar{x}))}(H(\bar{x})) = \mathcal{R}_{K^\circ}(H(\bar{x})) \cap \{G(\bar{x})\}^\perp$ is valid.
The last assertion denotes FRCQ for RNLP (3.9), which implies (3.19) and thus, the result follows by ii). ☐

**Remark 3.2.7.** *Theorem 3.2.6 focuses solely on CQs for W- and S-stationary points of MPCC* (3.6)*, since a more explicit formulation of M-stationary points is hardly at hand in abstract Banach spaces and therefore, rarely applicable. Nevertheless, we refer to [53, Proposition 3.6] for conditions that imply M-stationarity.*

Summarizing the key bottlenecks of MPCCs once more, we know that both implications of Theorem 2.3.9 fail to hold. Though, under additional assumptions it is at least possible

to show a suitable strength of S-stationary points. With respect to the numerical analysis, a major drawback is the lack of sequences that converge towards S-stationary points. Otherwise, W-stationarity is in general too weak and only qualifies for identifying potential critical points. Hence, a possible compromise lies in between.
We have already introduced M-stationarity, which is difficult to handle in general Banach spaces and requires set-valued analysis due to its implicit structure.
Another stationarity that is also known in finite dimensions is the concept of *Clarke stationarity* (C-stationarity), which is located between M- and W-stationarity. This concept will be of particular interest in Chapter 5. However, notice that an explicit characterization does not exist at the moment for abstract Banach spaces. In this context, see also the discussion in [53, Section 3.4].

## 3.3. Stationarity Concepts and Constraint Qualifications

Coming back to our equilibrium problem $\Gamma_{\text{EPCC}}$ that we introduced in the beginning of this chapter, the lack of existence results for Nash equilibria, the recent section and corresponding contributions on EPECs (see e.g. [9, 84]) motivate the following definition.

**Definition 3.3.1.** A point $\bar{x} \in X$ is called *W-, M-, S-stationary Nash equilibrium* of $\Gamma_{\text{EPCC}}$, if for all $\nu = 1, \dots, N$, $\bar{x}^\nu \in X_\nu$ is a W-, M-, S-stationary point of the parametric MPCC (3.5) with respect to $\bar{x}^{-\nu}$.

Let Assumption 3.2.1 be satisfied for player $\nu$'s MPCC (3.5) for all $\nu = 1, \dots, N$. Then we obtain with Definition 3.2.3 and Definition 3.2.4 the subsequent result.

**Theorem 3.3.2.** *The point $\bar{x} \in X$ is a*

*i) W-stationary Nash equilibrium of $\Gamma_{\text{EPCC}}$, if there exist multipliers*

$$(\kappa, \xi, \mu, \lambda) \in X^* \times W^* \times Z^* \times Z$$

*such that for all $\nu = 1, \dots, N$, it holds*

$$0 = D_{x^\nu} f^\nu(\bar{x}) + D_{x^\nu} g^\nu(\bar{x})^* \xi^\nu + D_{x^\nu} G^\nu(\bar{x})^* \mu^\nu + D_{x^\nu} H^\nu(\bar{x})^* \lambda^\nu + \kappa^\nu, \tag{3.21a}$$
$$\kappa^\nu \in \mathcal{N}_{Q_\nu}(\bar{x}^\nu), \tag{3.21b}$$
$$\xi^\nu \in \mathcal{N}_{C_\nu}(g^\nu(\bar{x})), \tag{3.21c}$$
$$\mu^\nu \in \text{cl}(K_\nu^\circ - K_\nu^\circ \cap \{G^\nu(\bar{x})\}^\perp) \cap \{G^\nu(\bar{x})\}^\perp, \tag{3.21d}$$
$$\lambda^\nu \in \text{cl}(K_\nu - K_\nu \cap \{H^\nu(\bar{x})\}^\perp) \cap \{H^\nu(\bar{x})\}^\perp; \tag{3.21e}$$

*ii) M-stationary Nash equilibrium of* $\Gamma_{\text{EPCC}}$*, if there exist multipliers*

$$(\kappa, \xi, \mu, \lambda) \in X^* \times W^* \times Z^* \times Z$$

*such that for all* $\nu = 1, \ldots, N$*, it holds* (3.21a)-(3.21c) *and*

$$(\mu^\nu, \lambda^\nu) \in \mathcal{N}_{\text{gph}\,\mathcal{N}_{K_\nu}}(G^\nu(\bar{x}), H^\nu(\bar{x})); \tag{3.22}$$

*iii) S-stationary Nash equilibrium of* $\Gamma_{\text{EPCC}}$*, if there exist multipliers*

$$(\kappa, \xi, \mu, \lambda) \in X^* \times W^* \times Z^* \times Z$$

*such that for all* $\nu = 1, \ldots, N$*, it holds* (3.21a)-(3.21c) *and*

$$\mu^\nu \in \mathcal{K}_{K_\nu^\circ}(H^\nu(\bar{x}), G^\nu(\bar{x})), \tag{3.23a}$$

$$\lambda^\nu \in \mathcal{K}_{K_\nu}(G^\nu(\bar{x}), H^\nu(\bar{x})). \tag{3.23b}$$

The strength of each type of Nash equilibrium is now inherited by the corresponding type of stationary point of the underlying MPCC. In particular, we obtain that an S-stationary Nash equilibrium is also a W-stationary Nash equilibrium. Furthermore, it holds

$$\text{S-stationary NE} \quad \Longrightarrow \quad \text{M-stationary NE} \quad \Longrightarrow \quad \text{W-stationary NE,}$$

if for all $\nu = 1, \ldots, N$, $K_\nu$ is polyhedric, the cones $K_\nu, K_\nu^\circ$ are pointed, induce a vector lattice and the corresponding supremum operators $\max_{K_\nu}(0, \cdot)$ and $\max_{K_\nu^\circ}$ satisfy (3.14) and (3.15), respectively.
Based on Definition 3.3.1, an extension of Theorem 3.2.6 can be derived straight forward for EPCCs.

**Theorem 3.3.3.** *Let* $\bar{x}$ *be a Nash equilibrium of* $\Gamma_{\text{EPCC}}$.

1. *If* $\bar{x}^\nu$ *satisfies CQ* (3.17) *for all* $\nu = 1, \ldots, N$*, then* $\bar{x}$ *is a W-stationary Nash equilibrium.*
2. *If at least one of the following conditions holds:* $\bar{x}^\nu$ *satisfies*
   - *CQ* (3.17) *and* (3.18)
   - *CQ* (3.19)
   - *CQ* (3.20)

   *for all* $\nu = 1, \ldots, N$*. Then* $\bar{x}$ *is an S-stationary Nash equilibrium.*

*Proof.* By Definition 3.3.1, the assertion follows by applying Theorem 3.2.6 to all players MPCCs. □

We end this section by giving a brief outlook. As we have discussed, we can not hope for existence of a Nash equilibrium of $\Gamma_{\text{EPEC}}$ in general or using the corresponding stationary Nash equilibrium. However, Theorem 3.3.3 indicates that weaker forms of stationary Nash equilibria can be seen as critical points under suitable regularity assumptions. Hence, it would be interesting to see whether we can find such points. A possible approach that we will focus on during the subsequent chapters is motivated by the consideration of a sequence of auxiliary equilibrium problems $\{\Gamma_k\}$ that are easier to handle and approximate the game $\Gamma_{\text{EPEC}}$. The latter property is meant in the sense that given a sequence $\{x^k\} \subseteq X$ of Nash equilibria of the auxiliary GNEPs, i.e. $x^k$ is a Nash equilibrium of $\Gamma_k$ for all $k \in \mathbb{N}$, then we find $\bar{x} \in X$ such that $x^k \to \bar{x}$ and $\bar{x}$ is a Nash equilibrium of $\Gamma_{\text{EPEC}}$. These auxiliary problems are typically constructed by penalization or regularization approaches (see e.g. [36, 38]). The structure of $\Gamma_k$ is then used to find a sequence of points that converges to a weaker stationary Nash equilibrium.

# 4. Multi-Leader Multi-Follower Games in Banach Spaces

Until now, we considered equilibrium problems in which all agents choose their strategy simultaneously. Thus, players that have a temporal advantage over their opponents can not be modeled. To overcome this issue, Chapter 4 is devoted to *hierarchical equilibrium problems*, where the players on the upper level are called *leaders* and on the lower level *followers*, respectively. In general, we distinguish between two types that can be seen in Figure 4.1. While Figure 4.1a depicts a *multi-leader multi-follower game* (MLMFG), we see in Figure 4.1b a *single-leader multi-follower game*. The latter one is known as *Stackelberg game* and named after Heinrich von Stackelberg, who first considered this type of problem in 1934 [92].

(a) Multi-leader-multi-follower game

(b) Stackelberg game

Figure 4.1.: Hierarchical equilibrium problems

Since Stackelberg games can also be formulated as MPECs or MPCCs, respectively, we mainly focus in this section on MLMFGs and within that on *multi-leader (single-)follower games*. We will denote the latter equilibrium problem by MLFG to emphasize the difference to MLMFGs.

Now, the basic idea of MLMFGs is that the leaders compete with each other in their equilibrium problem on the upper level, where they anticipate the followers' strategies. Afterwards, the followers themselves determine their equilibrium strategy by solving the lower level Nash game with respect to the leaders' solution.

Most previous works on MLMFGs and MLFGs focus on the finite dimensional case, see e.g.

[25, 36, 37, 39, 47, 50, 69]. In this context, we emphasize the contributions [37, 69] for a general introduction to MLMFGs in finite spaces. In order to prove the existence of an equilibrium, the authors considered special types of MLMFGs/MLFGs under very strict assumptions. In this context, [25] analyzed a quadratic multi-leader single-follower game, where the lower level has a unique solution and the leaders' objective function is linear dependent on the follower's strategy. Due to an implicit function approach, the authors were able to show the existence of a leader-follower equilibrium point. Furthermore, by using a nonlinear-complementarity (NCP) function they proved the convergence of a sequence to an S-stationary Nash equilibrium. In a similar way, [36] also obtained the convergence to the same type of equilibrium point. In [39], the authors considered a game of zero-sum type and were able to show the existence and uniqueness under strong convexity of the objective functions. A different approach was used in [47], where the authors focused on a potential game formulation and summarized the lower-level constraints to a joint constraint. In this context, they were able to prove the existence of an equilibrium of the original problem.
In abstract Banach spaces, one notable exception is the paper [61] that we already mentioned in Chapter 3 by Boris Mordukhovich, who considered abstract EPECs that can be seen as a superclass of MLMFGs. Now, our contribution is less restricted to a specific structure of the underlying problem and extends the basic concepts to the function space setting. The remaining chapter is organized as follows. In Section 4.1 we introduce basic definitions and assumptions, which are needed to analyze and describe MLMFGs in general. For this reason, we reformulate the hierarchical game to an EPEC/EPCC, where we can apply the results from Chapter 3 and in particular, focus on the relation between both games. Notice that the latter is no trivial task, since we obtain an additional parameter for the corresponding EPCC.
Section 4.2 considers a special MLFG, i.e. a multi-leader single-follower game, where the follower solves a quadratic program with linear constraints in Hilbert spaces. Therein, we consider two approaches. While the first one (see Subsection 4.2.1) reformulates the lower level with respect to a variational inequality, the second ansatz in Subsection 4.2.2 uses the KKT-formulation of the follower's problem. For both models it is shown that the resulting auxiliary NEPs are approximations of the corresponding EPEC under suitable assumptions.

## 4.1. Problem Formulation and Basic Definitions

If not stated otherwise, we denote by $N \in \mathbb{N}$ the number of leaders and by $M \in \mathbb{N}$ the amount of followers. Let $x^\nu \in X_\nu$ be the strategy of leader $\nu$ ($\nu = 1, \ldots, N$) and $y^i \in Y_i$ be the strategy of follower $i$ ($i = 1, \dots, M$), respectively. Moreover, we use the common game theoretic notations introduced in Subsection 2.3.2 and denote by $\Gamma_{\text{MLMFG}}$ the MLMFG, where leader $\nu$ solves the parametric *bilevel optimization problem* (BOP)

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}, y) \\ \text{w.r.t. }\ & x^\nu \in X_\nu, \\ \text{s.t. }\ & x^\nu \in Q_\nu,\ y \in \Psi(x^\nu, x^{-\nu}) \end{aligned} \tag{4.1}$$

with respect to $x^{-\nu} \in X_{-\nu}$ for all $\nu = 1, \dots, N$.
Here, we consider the Banach spaces $X := X_1 \times \cdots \times X_N$ and $Y := Y_1 \times \cdots \times Y_M$ and denote by $f^\nu : X \times Y \to \mathbb{R}$ each leader's cost functional. The set $Q_\nu \subseteq X_\nu$ is assumed to be nonempty, closed, convex and bounded for all $\nu = 1, \dots, N$, while the set-valued mapping $\Psi : X \rightrightarrows Y$ defines the set of Nash equilibria of the lower level parametric NEP $\Gamma_F(x) := \{\theta^i(\cdot, x), Y^i_{\text{ad}}(\cdot, x)\}_{i=1}^M$ with respect to $x$. In particular, the leaders' strategies can influence both the outcome and the feasible strategies of each follower.
For the sake of simplicity, we assume that the sets $Y^i_{\text{ad}}$ are independent of the opponent follower strategies $y^{-i}$ and all leaders share the same followers. Below, we consider a prominent example given in [69], which shows that even MLFGs, where each player faces an analytically simple problem, may be troublesome.

**Example 4.1.1** ([69, Example 4])**.** *Consider a two-leader single-follower game in which each leader chooses a strategy on the real line under the following data:*

- *Leader 1:* $f^1(x^1, x^2, y) := \frac{1}{2}x^1 + y$ *and* $Q_1 := [0, 1]$*;*
- *Leader 2:* $f^2(x^1, x^2, y) := -\frac{1}{2}x^2 - y$ *and* $Q_2 := [0, 1]$*;*
- $\Psi(x) := \operatorname{argmin}\{\frac{1}{2}y^2 + (x^1 + x^2 - 1)y \mid y \geq 0\}$.

*In this case,* $\Psi$ *denotes the set of solutions of a quadratic minimization problem with linear constraints. It is easily seen that the KKT-system is given by*

$$x^1 + x^2 - 1 + y - \lambda = 0, \quad y \geq 0, \quad \lambda \geq 0, \quad \lambda \cdot y = 0$$

*and hence, we obtain the following best response function*[13] *of the follower by*

$$y(x^1, x^2) = \max(0, 1 - x^1 - x^2).$$

[13] The optimal output with respect to the leaders' choices.

*Thus,* $\Psi(x) = \{y(x)\}$ *and we can write the hierarchical equilibrium problem as a single-level Nash game with cost functions*

$$\tilde{f}^1(x^1, x^2) := f^1(x^1, x^2, y(x^1, x^2)) = \max(\frac{1}{2}x^1, 1 - \frac{1}{2}x^1 - x^2),$$
$$\tilde{f}^2(x^1, x^2) := f^2(x^1, x^2, y(x^1, x^2)) = \min(-\frac{1}{2}x^2, x^1 - \frac{1}{2}x^2 - 1).$$

*Then the resulting best-response function for each leader can be derived by*

$$\varphi^1(x^2) = \{1 - x^2\} \quad \forall x^2 \in Q_2, \qquad \varphi_2(x^1) = \begin{cases} \{0\} & \text{if } x^1 \in [0, \frac{1}{2}), \\ \{0, 1\} & \text{if } x^1 = \frac{1}{2}, \\ \{1\} & \text{if } x^1 \in (\frac{1}{2}, 1]. \end{cases}$$

*In particular, we see that there are no points* $\hat{x}^1, \hat{x}^2 \in [0, 1]$ *such that* $\varphi^1(\hat{x}^2) = \varphi^2(\hat{x}^1)$ *and hence, the NEP* $\{\tilde{f}^\nu, Q_\nu\}_{\nu=1}^2$ *has no Nash equilibrium, which implies that the two-leader single-follower game has no NE either. Both graphs are displayed in Figure 4.2, which confirms the latter assertion.*

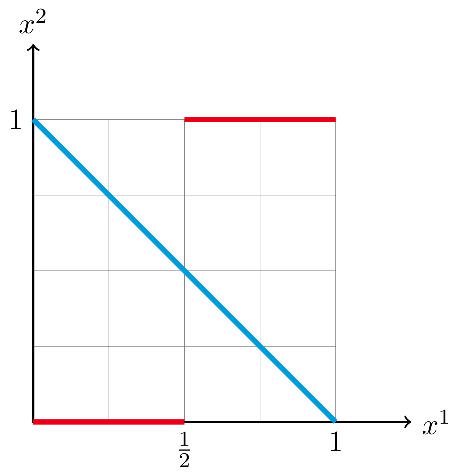



Figure 4.2.: Graphs of the best response function: $\varphi_1(x^2)$ (blue) and $\varphi_2(x^1)$ (red).

As a result of Example 4.1.1, we observe that although the leader problems are of linear type and the lower level is a quadratic problem with linear constraints, we still lack the existence of an equilibrium. The reason for that can be found in the reformulated single-level equilibrium problem, where we see that the reduced cost function of leader 2 is non-convex, which emphasizes once more the importance of convexity in game theory. In order to define a leader-follower Nash equilibrium, we have to distinguish two different cases depending on how each leader anticipates the followers' strategies, i.e. optimistically or pessimistically.

**Definition 4.1.2** ([37, Def. 2.6])**.** A strategy tuple $(\bar{x}, \bar{y}) \in X \times Y$ is called *optimistic leader-follower Nash equilibrium*, if $(\bar{x}, \bar{y}) \in Q \times \Psi(\bar{x})$ and for all $\nu = 1, \ldots, N$, $\bar{x}^\nu$ solves

$$\min_{x^\nu \in Q_\nu} \min_{y \in \Psi(x^\nu, \bar{x}^{-\nu})} f^\nu(x^\nu, \bar{x}^{-\nu}, y).$$

Analogously, $(\bar{x}, \bar{y})$ is called *pessimistic leader-follower Nash equilibrium*, if $(\bar{x}, \bar{y}) \in Q \times \Psi(\bar{x})$ and for all $\nu = 1, \ldots, N$, $\bar{x}^\nu$ solves

$$\min_{x^\nu \in Q_\nu} \max_{y \in \Psi(x^\nu, \bar{x}^{-\nu})} f^\nu(x^\nu, \bar{x}^{-\nu}, y).$$

Notice that unless the set-valued mapping $\Psi$ equals a singleton, i.e. the lower level Nash equilibrium problem admits a unique solution for all $x \in Q$, Definition 4.1.2 is not well-defined in general, since each leader can anticipate a different lower level Nash equilibrium. In order to circumvent these issues, it is quite popular (see e.g. [37, 39]) to require that the lower level problem possesses at most one solution for all $x \in X$. In this context, we assume for the subsequent analysis that $\Psi(x)$ is at most a singleton.

**Definition 4.1.3.** Let $\Psi(x) \equiv \{y(x)\}$. Then the tuple $(\bar{x}, \bar{y})$ is called a *leader-follower Nash equilibrium* (LF-Nash equilibrium), if $(\bar{x}, \bar{y}) \in Q \times Y_{\text{ad}}(\bar{x})$ and for all $\nu = 1, \ldots, N$, $\bar{x}^\nu$ solves

$$f^\nu(\bar{x}, \bar{y}) \leq f^\nu(x^\nu, \bar{x}^{-\nu}, y(x^\nu, \bar{x}^{-\nu})) \qquad \forall x^\nu \in Q_\nu \tag{4.2}$$

Analogously, $(\bar{x}, \bar{y})$ is called a *local LF-Nash equilibrium*, if (4.2) is replaced by

$$f^\nu(\bar{x}, \bar{y}) \leq f^\nu(x^\nu, \bar{x}^{-\nu}, y(x^\nu, \bar{x}^{-\nu})) \qquad \forall x^\nu \in Q_\nu \cap \mathbb{B}_\epsilon(\bar{x}^\nu).$$

Without loss of generality[14], we consider from now on the case $M = 1$ and write $\Gamma_{\text{MLFG}}$ instead of $\Gamma_{\text{MLMFG}}$ to indicate that the further focus is on a multi-leader single-follower

[14]The amount of followers does not change the complexity of (4.1), since it is still possible under suitable assumptions to replace $\Psi(x)$ by the corresponding first-order optimality conditions of the lower level.

game. Furthermore, $\Psi$ denotes the solution set of the following parametric minimization problem

$$\begin{aligned} \min \quad & \theta(x,y) \\ \text{w.r.t.} \quad & y \in Y, \\ \text{s.t.} \quad & y \in Y_{\text{ad}}(x) := \{ y \in Y \mid h(x,y) \in K \} \end{aligned} \tag{4.3}$$

with a functional $\theta : X \times Y \to \mathbb{R}$, a mapping $h : X \times Y \to Z$ into some Banach space $Z$ and a nonempty, closed and convex subset $K \subseteq Z$.
Now, the idea to tackle $\Gamma_{\text{MLFG}}$ is motivated by the analysis of each leader's bilevel optimization problem (4.1). In this context, two approaches are used regularly within the scope of MLFGs. The first and most intuitive one was already applied in Example 4.1.1, which uses the fact that we have an explicit solution of the lower level. Similar to the example, $\Gamma_{\text{MLFG}}$ reduces to an, in general non-smooth, single-level Nash equilibrium problem. For this reason, we refer to [25, 37, 39] for a detailed analysis in finite dimensions.
The second approach that we will apply in this thesis uses the first-order optimality system of (4.3) in order to generate an EPEC. Here, we refer to [36, 39, 50, 69] for further contributions.
The literature concerning bilevel programs is manifold and in particular, there are different ways to analyze these problems (e.g. by the optimal-value functional[15] of the lower level). At this point, we recommend [12] for a detailed and extensive overview of bilevel literature in finite dimensions and [19, 54, 56] for contributions in function spaces.
Through the remainder of Chapter 4 and if not stated otherwise, we require that the lower level satisfies the following assumptions.

**Assumption 4.1.4.** *Let $X$, $Y$ and $Z$ be reflexive Banach spaces and assume that $\theta : X \times Y \to \mathbb{R}$ and $h : X \times Y \to Z$ are twice continuously differentiable. Furthermore, we require that $\theta(x,\cdot)$ is strictly convex and $h(x,\cdot)$ is $-K$-convex in $y$ for all $x \in X$. Finally, KRZCQ is satisfied, i.e.*

$$D_y h(x,y)[Y] - \mathcal{R}_K(h(x,y)) = Z$$

*for all $y \in Y_{\text{ad}}(x)$ and $x$ such that $Y_{\text{ad}}(x)$ is nonempty.*

Based on Theorem 2.3.9, we know that for all $x \in Q$ such that $\Psi(x)$ is nonempty, $y \in \Psi(x)$ is equivalent to the existence of a Lagrange multiplier $p \in Z^*$ such that

$$\begin{gathered} D_y\theta(x,y) + D_y h(x,y)^* p = 0, \\ h(x,y) \in K, \; p \in K^\circ, \; \langle p, h(x,y) \rangle_Z = 0 \end{gathered}$$

[15]The optimal value functional $\varphi$ of (4.3) is given by $\varphi(x) := \inf\{\theta(x,y) \mid h(x,y) \in K\}$, which motivates to replace $y \in \Psi(x)$ in (4.1) by $\theta(x,y) - \varphi(x) \leq 0, \; h(x,y) \in K$.

is satisfied. In particular, if a solution of the KKT-system exists, then it is unique. This motivates the following equilibrium problem, denoted by $\Gamma^{\text{MLFG}}_{\text{EPCC}}$ in which leader $\nu$ solves

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}, y) \\ \text{w.r.t. } & (x^\nu, y, p) \in X_\nu \times Y \times Z^*, \\ \text{s.t. } & x^\nu \in Q_\nu, \\ & D_y\theta(x,y) + D_y h(x,y)^* p = 0, \\ & h(x,y) \in K,\ p \in K^\circ,\ \langle p, h(x,y)\rangle_Z = 0. \end{aligned} \tag{4.4}$$

Then we observe that $\Gamma^{\text{MLFG}}_{\text{EPCC}}$ is an EPCC of joint type that contains an additional optimization parameter $p$, which has to be considered as well in the further analysis. However, notice that this inclusion can lead to some difficulties and not all relations between a global/local LF-Nash equilibrium of $\Gamma_{\text{MLFG}}$ and a global/local Nash equilibrium of $\Gamma^{\text{MLFG}}_{\text{EPCC}}$ are straight forward. Motivated by the corresponding contributions on bilevel problems and MPCCs (see [13, 53]), we recall the set of Lagrange multipliers of (4.3)

$$\Lambda(x,y) = \left\{p \in \mathcal{N}_K(h(x,y)) \;\middle|\; D_y\theta(x,y) + D_y h(x,y)^* p = 0\right\}$$

and start with the pleasant part.

**Theorem 4.1.5.** *Let $(\bar{x}, \bar{y}) \in X \times Y$ be a global/local LF-Nash equilibrium of $\Gamma_{\text{MLFG}}$, then $(\bar{x}, \bar{y}, p) \in X \times Y \times Z^*$ is a global/local Nash equilibrium of $\Gamma^{\text{MLFG}}_{\text{EPCC}}$ for any $p \in \Lambda(\bar{x}, \bar{y})$.*
*In reverse, let $(\hat{x}, \hat{y}, \hat{p}) \in X \times Y \times Z^*$ be a global Nash equilibrium of $\Gamma^{\text{MLFG}}_{\text{EPCC}}$, then $(\hat{x}, \hat{y})$ is a global LF-Nash equilibrium of $\Gamma_{\text{MLFG}}$.*

*Proof.* Let $(\bar{x}, \bar{y}) \in X \times Y$ be a global LF-Nash equilibrium of $\Gamma_{\text{MLFG}}$. Then, $(\bar{x}, \bar{y})$ is by definition a global solution of (4.1) with respect to $\bar{x}^{-\nu}$ for all $\nu = 1, \ldots, N$. By [53, Theorem 4.17], this is equivalent to the fact that for any $p \in \Lambda(\bar{x}, \bar{y})$, $(\bar{x}^\nu, \bar{y}, p)$ is a global solution of (4.4) with respect to $\bar{x}^{-\nu}$ for all $\nu = 1, \ldots, N$. Again by definition, we conclude that $(\bar{x}, \bar{y}, p)$ is a global Nash equilibrium of $\Gamma^{\text{MLFG}}_{\text{EPCC}}$ for any $p \in \Lambda(\bar{x}, \bar{y})$. In particular, this has also shown the reverse direction for global equilibria.
The assertion for local LF-Nash equilibria follows the same idea with the difference that we use [53, Theorem 4.18] for the transition to local solutions of the corresponding MPCC. □

In contrast to global equilibria, the reverse direction does not hold for the local counterparts in general. In particular, it has been proven in [13, Example 3.1] that a local solution of a finite dimensional MPCC is not automatically a local solution of the corresponding bilevel problem. In this context, it was shown in [13, Theorem 3.2] that if $(\hat{x}, \hat{y}, p)$ is a local solution of MPCC (4.4) for all $p \in \Lambda(\hat{x}, \hat{y})$, then $(\hat{x}, \hat{y})$ is a local solution of BOP (4.1). On the one hand, it is quite cumbersome to check that the assumptions are satisfied for all $p \in \Lambda(\hat{x}, \hat{y})$. On the other hand, the requirement that $Z$ is finite dimensional is crucial, as has been shown in [53, Example 4.20]. Hence and in order to circumvent these difficulties, we strengthen the assumptions in the infinite dimensional setting and obtain the following result.

**Corollary 4.1.6.** *Assume that the set of Lagrange multipliers $\Lambda(\bar{x}, \bar{y})$ is a singleton. Then the tuple $(\bar{x}, \bar{y}) \in X \times Y$ is a global/local LF-Nash equilibrium of $\Gamma_{\text{MLFG}}$ if and only if there exists $\bar{p} \in \Lambda(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, \bar{p}) \in X \times Y \times Z^*$ is a global/local Nash equilibrium of $\Gamma^{\text{MLFG}}_{\text{EPCC}}$.*

We refer to Subsection 2.3.1 for conditions, which imply the uniqueness of the Lagrange multiplier. After analyzing the consistency of the surrogate equilibrium problem $\Gamma^{\text{MLFG}}_{\text{EPCC}}$, we proceed with MLFG-tailored stationarity concepts. Motivated by the last results, we state the following definition.

**Definition 4.1.7.** A feasible point $(\bar{x}, \bar{y}) \in X \times Y$ of $\Gamma_{\text{MLFG}}$ is called *W-, M-, S-stationary Nash equilibrium*, if there exists a multiplier $\bar{p} \in \Lambda(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, \bar{p}) \in X \times Y \times Z^*$ is a W-, M-, S-stationary Nash equilibrium of $\Gamma^{\text{MLFG}}_{\text{EPCC}}$ in the sense of Definition 3.3.1.

As a consequence of Definition 4.1.7, explicit stationary conditions in the sense of Theorem 3.3.2 can be stated. Moreover, we derive necessary optimalty conditions for LF-Nash equilibria by using Theorem 3.3.3 and MLFG-tailored constraint qualifications. For a better clarity, we use the abbreviations $\bar{h} := h(\bar{x}, \bar{y})$ and $\bar{\theta} := \theta(\bar{x}, \bar{y})$.

**Theorem 4.1.8.** *Let $(\bar{x}, \bar{y})$ be a LF-Nash equilibrium of $\Gamma_{\text{MLFG}}$ and define the operator $P_\nu : X \times Y \times Z^* \to Y^* \times Z \times Z^*$ by*

$$P_\nu(x, y, p) := \begin{pmatrix} D^2_{yx^\nu}\theta(x, y) + \langle p, D^2_{yx^\nu} h(x, y)\rangle_Z & D^2_{yy}\theta(x, y) + \langle p, D^2_{yy} h(x, y)\rangle_Z & D_y h(x, y) \\ D_{x^\nu} h(x, y) & D_y h(x, y) & 0_{Z^*} \\ 0_X & 0_Y & I_{Z^*} \end{pmatrix}.$$

*a) If there exists a $p \in \Lambda(\bar{x}, \bar{y})$ such that*

$$P_\nu(\bar{x}, \bar{y}, p) \begin{bmatrix} \mathcal{R}_{Q_\nu}(\bar{x}^\nu) \\ Y \\ Z^* \end{bmatrix} - \begin{pmatrix} \{0\} \\ \mathcal{R}_K(\bar{h}) \cap (-\mathcal{K}_K(\bar{h}, p)) \\ \mathcal{R}_{K^\circ}(p) \cap (-\mathcal{K}_{K^\circ}(p, \bar{h})) \end{pmatrix} = \begin{bmatrix} Y^* \\ Z \\ Z^* \end{bmatrix} \tag{4.5}$$

*is satisfied for all $\nu = 1, \ldots, N$, then $(\bar{x}, \bar{y})$ is a W-stationary Nash equilibrium of $\Gamma_{\text{MLFG}}$.*

*b) If one of the following conditions holds:*

*i) The identity* (4.5) *holds and*

$$\mathrm{cl}\left( P_\nu(\bar{x}, \bar{y}, p) \begin{bmatrix} \mathcal{N}_{Q_\nu}(\bar{x}^\nu)_\perp \\ Y \\ Z^* \end{bmatrix} - \begin{pmatrix} \{0\} \\ \mathcal{N}_{K \cap (-\mathcal{T}_K(\bar{h}))}(\bar{h})^\perp \\ \mathcal{N}_{K^\circ \cap (-\mathcal{T}_{K^\circ}(p))}(p)^\perp \end{pmatrix} \right) = \begin{bmatrix} Y^* \\ Z \\ Z^* \end{bmatrix} \tag{4.6}$$

*is satisfied for all* $\nu = 1, \ldots, N$.

*ii) It holds*

$$P_\nu(\bar{x}, \bar{y}, p) \begin{bmatrix} \mathcal{R}_{Q_\nu}(\bar{x}^\nu) \\ Y \\ Z^* \end{bmatrix} - \begin{pmatrix} \{0\} \\ \mathcal{R}_K(\bar{h}) \cap \{p\}^\perp \\ \mathcal{R}_{K^\circ}(p) \cap \{\bar{h}\}^\perp \end{pmatrix} = \begin{bmatrix} Y^* \\ Z \\ Z^* \end{bmatrix}, \tag{4.7}$$

*for all* $\nu = 1, \ldots, N$.

*iii) The operator* $P_\nu(\bar{x}, \bar{y}, p)$ *is surjective and either* $Q_\nu = X^\nu$ *or* $\bar{x}^\nu \in \mathrm{int}(Q_\nu)$ *holds for all* $\nu = 1, \ldots, N$.

*Then* $(\bar{x}, \bar{y})$ *is an S-stationary Nash equilibrium of* $\Gamma_{\text{MLFG}}$.

*Proof.* As a special type of EPCC, the proof follows by Theorem 3.3.3 applied to $\Gamma^{\text{MLFG}}_{\text{EPCC}}$. □

Notice that the conditions (4.5)-(4.7) depend on the lower level Lagrange multiplier $p$. Therefore and unless the set $\Lambda(\bar{x}, \bar{y})$ is single-valued, these constraint qualifications need not to be satisfied for all $p \in \Lambda(\bar{x}, \bar{y})$ (see also [53, Example 4.30]). However, in the next section we focus on a special type of lower level problem where this scenario is more comfortable to handle.

## 4.2. Multi-Leader Single-Follower Games with Quadratic Lower Level in Hilbert Spaces

After considering MLFGs with a rather general lower level problem, we now focus on a special type of $\Gamma_{\text{MLFG}}$, where $\Psi$ is given as the solution mapping of the following quadratic minimization problem

$$\begin{aligned} \min \quad & \frac{1}{2}\langle Ay, y\rangle_Y - \langle Bx - b, y\rangle_Y \\ \text{w.r.t.} \quad & y \in Y, \\ \text{s.t.} \quad & y \in Y_{\text{ad}}(x). \end{aligned} \tag{4.8}$$

In order to distinguish between this model and the general equilibrium problem in the previous section, we use $\Gamma^{\mathrm{lq}}_{\mathrm{MLFG}}$ to emphasize the difference. Throughout the remainder of this section, the lower level satisfies the following standing assumptions.

**Assumption 4.2.1.** *We strengthen Assumption 4.1.4 with respect to the underlying spaces and assume that $Y$ is a Hilbert space. Moreover, the set-valued mapping $Y_{\mathrm{ad}}(x) \subseteq Y$ is given by*

$$Y_{\mathrm{ad}}(x) := \left\{ y \in Y \;\middle|\; h(x,y) := \hat{h}(x) + Hy + \psi \in K \right\}$$

*where $\hat{h} : X \to Z$ is continuously differentiable and $-K$-convex and $K \subseteq Z$ denotes a nonempty, closed and convex cone. Furthermore, let $H \in L(Y,Z)$ be coercive and $\psi \in Z$. The objective function is comprised of a symmetric and coercive operator $A \in L(Y,Y^*)$, $B \in L(X,Y^*)$ and $b \in Y^*$.*

At this point, we refer to Section 5.1, where we consider and discuss an MLFG in $H^1_0(\Omega)$. In particular, it is shown that Assumption 4.2.1 is reasonable. In this context, we recall that an operator $A$ in a Hilbert space $Y$ is called *symmetric* (or *self-adjoint*), if for arbitrary $y_1, y_2 \in Y$ it holds

$$(Ay_1, y_2)_Y = (y_1, Ay_2)_Y$$

and *coercive* (or *uniformly elliptic*), if there exists a constant $c > 0$ such that for all $y \in Y$ the inequality

$$(Ay, y)_Y \geq c\|y\|_Y^2$$

is satisfied. Due to Assumption 4.2.1, it is known (see Theorem 2.3.5 and the theorem of Lions and Stampacchia [46, Theorem 2.1]) that for all $x \in Q$ with $Y_{\mathrm{ad}}(x)$ nonempty, problem (4.8) admits a unique solution $y_x$, which is given as the solution of the following VI:

$$\text{Find } y \in Y_{\mathrm{ad}}(x) \text{ s.t. } \langle Ay - Bx - b, v - y\rangle_Y \geq 0 \quad \forall v \in Y_{\mathrm{ad}}(x). \tag{4.9}$$

Since $K$ is a cone and $H$ is coercive and hence, surjective, we can further apply Theorem 2.3.9 and conclude that (4.9) is equivalent to the existence of a unique Lagrange multiplier $p_x \in Z^*$ such that $(y_x, p_x)$ solves

$$\begin{aligned} Ay - Bx - b + H^*p &= 0, \\ h(x,y) \in K,\ p \in K^\circ,\ \langle p, h(x,y)\rangle_Z &= 0. \end{aligned} \tag{4.10}$$

Both first-order optimality conditions can be used in order to replace the condition $y \in \Psi(x)$ in leader $\nu$'s problem (4.1) for all $\nu = 1, \ldots, N$. The resulting auxiliary problems are now given on the one hand by an EPEC, denoted by $\Gamma^{\mathrm{lq}}_{\mathrm{EPEC}}$, where each leader solves

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}, y) \\ \text{w.r.t. }\ & (x^\nu, y) \in X_\nu \times Y, \\ \text{s.t. }\ & x^\nu \in Q_\nu,\ -Ay + Bx + b \in \mathcal{N}_{Y_{\mathrm{ad}}(x)}(y) \end{aligned} \tag{4.11}$$

and on the other hand as an EPCC, denoted by $\Gamma^{\mathrm{lq}}_{\mathrm{EPCC}}$, where each leader solves

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}, y) \\ \text{w.r.t. }\ & (x^\nu, y, p) \in X_\nu \times Y \times Z^*, \\ \text{s.t. }\ & x^\nu \in Q_\nu, \\ & Ay - B[x^\nu, x^{-\nu}] + H^* p = b, \\ & h(x^\nu, x^{-\nu}, y) \in K,\ p \in K^\circ,\ \langle p, h(x^\nu, x^{-\nu}, y)\rangle_Z = 0. \end{aligned} \tag{4.12}$$

With respect to bilevel optimization problems, both techniques are quite common in order to get rid of the bilevel structure. In the context of equilibrium problems, we will observe that the former one is more suitable as the resulting auxiliary equilibrium problems are easier to handle. In reverse, the second approach may lead to stronger types of stationary Nash equilibria. Furthermore, notice that the approach induced by the VI-formulation is stronger, since less requirements are needed. In particular, it does not rely on the conic property of $K$.
Motivated by Definition 4.1.7, the explicit forms of the different type of stationary Nash equilibria are given in the subsequent theorem.

**Theorem 4.2.2.** *Let $(\bar{x}, \bar{y}) \in X \times Y$ be a feasible point of $\Gamma^{\mathrm{lq}}_{\mathrm{MLFG}}$. Then $(\bar{x}, \bar{y})$ is a*

*i) W-stationary Nash equilibrium of $\Gamma_{\mathrm{MLFG}}$, if there exist multipliers $(p, \kappa, \xi, \mu) \in Z^* \times X^* \times Y^N \times (Z^*)^N$ such that*

$$\begin{aligned} D_{x^\nu} f^\nu(\bar{x}, \bar{y}) + \kappa^\nu + B_\nu^* \xi^\nu + D_{x^\nu} \hat{h}(\bar{x})^*[\mu^\nu] &= 0, \\ D_y f^\nu(\bar{x}, \bar{y}) + A\xi^\nu + H^* \mu^\nu &= 0, \\ Ay - Bx - b + H^* p &= 0, \\ \kappa^\nu &\in \mathcal{N}_{Q_\nu}(\bar{x}^\nu) \\ p &\in \mathcal{N}_K(h(\bar{x}, \bar{y})) \end{aligned} \tag{4.13}$$

*and*

$$\begin{aligned} \mu^\nu &\in \mathrm{cl}(K^\circ - K^\circ \cap \{h(\bar{x}, \bar{y})\}^\perp) \cap \{h(\bar{x}, \bar{y})\}^\perp, \\ -H\xi^\nu &\in \mathrm{cl}(K - K \cap \{p\}^\perp) \cap \{p\}^\perp \end{aligned}$$

*are satisfied for all $\nu = 1, \ldots, N$;*

*ii) M-stationary Nash equilibrium of* $\Gamma_{\text{MLFG}}$*, if there exist multipliers* $p \in \Lambda(\bar{x}, \bar{y})$ *and* $(\kappa, \xi, \mu) \in X^* \times Y \times Z^*$ *such that* (4.13) *is satisfied and it holds*

$$(\mu^\nu, -H\xi^\nu) \in \mathcal{N}_{\text{gph } N_K}(h(\bar{x}, \bar{y}), p);$$

*iii) S-stationary Nash equilibrium of* $\Gamma_{\text{MLFG}}$*, if there exist multipliers* $p \in \Lambda(\bar{x}, \bar{y})$ *and* $(\kappa, \xi, \mu) \in X^* \times Y \times Z^*$ *such that* (4.13) *is satisfied and it holds*

$$\begin{aligned} \mu^\nu &\in \mathcal{K}_{K^\circ}(p, h(\bar{x}, \bar{y})), \\ -H\xi^\nu &\in \mathcal{K}_K(h(\bar{x}, \bar{y}), p). \end{aligned}$$

*Proof.* According to Definition 4.1.7, the result follows by applying $\Gamma^{\text{lq}}_{\text{EPCC}}$ to Theorem 3.3.2. □

Similar to Section 4.1 for the general case, we are interested in the relation between $\Gamma^{\text{lq}}_{\text{MLFG}}$ and both surrogate problems. Starting with $\Gamma^{\text{lq}}_{\text{EPCC}}$, we emphasize the requirements in Corollary 4.1.6 are valid, which guarantees the equivalence between LF-Nash equilibria of $\Gamma^{\text{lq}}_{\text{MLFG}}$ and Nash equilibria of $\Gamma^{\text{lq}}_{\text{EPCC}}$.
In order to verify the equivalence of $\Gamma^{\text{lq}}_{\text{MLFG}}$ and $\Gamma^{\text{lq}}_{\text{EPEC}}$, observe that the latter one can be written as the following equilibrium problem, where leader $\nu$ solves

$$\begin{aligned} \min \;& F^\nu(x^\nu, x^{-\nu}) := f^\nu(x^\nu, x^{-\nu}, y(x^\nu, x^{-\nu})) \\ \text{w.r.t. } & x^\nu \in X_\nu, \\ \text{s.t. } & x^\nu \in Q_\nu. \end{aligned} \tag{4.14}$$

Here, we used an implicit function approach and inserted the solution mapping $y(x)$ of (4.9) into each leader's objective functional.

**Lemma 4.2.3.** *The tuple* $(\bar{x}, \bar{y})$ *is a global/local LF-Nash equilibrium of* $\Gamma^{\text{lq}}_{\text{MLFG}}$ *if and only if* $\bar{y} = y(\bar{x})$ *and* $\bar{x}$ *is a global/local Nash equilibrium of NEP* $\{F^\nu, Q_\nu\}_{\nu=1}^N$.

In finite dimensions the proof for global equilibria was given for instance in [38]. Since the transfer to infinite dimensions and local equilibria is straight forward, we omit the corresponding proof at this point.
During the remaining part of this section and already indicated at the end of Chapter 3, we are now interested in finding auxiliary equilibrium problems for both $\Gamma^{\text{lq}}_{\text{EPCC}}$ and $\Gamma^{\text{lq}}_{\text{EPEC}}$ that approximate the original problem. Motivated by several concepts for MPECs, we consider two approaches. While the first one is based on a penalization technique of the

variational inequality (4.9), the second approach concentrates on the critical part of the complementarity constraints.
Before we show both methods more detailed, we introduce an appropriate convergence concept for equilibrium problems. For this reason, we start by defining the *epi-convergence* of a sequence of functions $\{f^k\}_{k\in\mathbb{N}}$ and recall that the *epigraph* of an arbitrary functional $f : X \to \mathbb{R}$ is given by

$$\text{epi } f := \big\{(x,\alpha) \in X \times \mathbb{R} \;\big|\; f(x) \le \alpha\big\}.$$

**Definition 4.2.4.** A sequence of functions $f^k : X \to \mathbb{R}$ is said to *epi-converge* to $f$ if and only if the epigraph epi $f^k$ converges to epi $f$ in the sense of Kuratowski (see [11, Definition 4.10]).

In order to check the epi-convergence, one often uses an equivalent definition (see e.g. [74, Proposition 7.2]), which states that a sequence $f^k$ epi-converges to $f$, if the following two conditions are satisfied:

1. Whenever $\{x^k\} \subseteq X$ and $x^k \rightharpoonup \bar{x}$, then $f(\bar{x}) \le \liminf\limits_{k\to\infty} f^k(x^k)$;
2. For every $\bar{x} \in X$, there exists a sequence $\{x^k\} \subseteq X$ such that $x^k \to x$ and $f(\bar{x}) \ge \limsup\limits_{k\to\infty} f^k(x^k)$;

With respect to equilibrium problems, [22] introduced the concept of multi epi-convergence in finite dimensions, which was used in [38] to describe approximations of EPCCs. We extend this property to general Banach spaces and introduce this convergence in the sense of [38, Proposition 2.8] that establishes the relation of epi-convergence and multi epi-convergence.

**Definition 4.2.5.** A sequence of families $\{F^{k,\nu}\}$ with $F^{k,\nu} : X \to \mathbb{R}$ and $\nu = 1, \dots, N$ *multi epi-converges* to the family of functionals $\{F^\nu\}$ with $F^\nu : X \to \mathbb{R}$ on the set $X$ if and only if for all $\nu = 1, \dots, N$ and every sequence $\{x^{k,-\nu}\} \subseteq X_{-\nu}$ converging to some $\bar{x}^{-\nu} \in X_{-\nu}$, the sequence of functionals $\{\Psi^{k,\nu}\}$ with $\Psi^{k,\nu}(x^\nu) := F^{k,\nu}(x^\nu, x^{k,-\nu})$ epi-converges to $\{\Psi^\nu\}$ with $\Psi^\nu(x^\nu) := F^\nu(x^\nu, \bar{x}^{-\nu})$ on $X_\nu$.

In the next step, we define a suitable continuity concept for set-valued mappings. For this reason, we first need a certain type of set convergence (see [64]).

**Definition 4.2.6.** Let $\mathcal{S}$ and $\mathcal{S}_k$ with $k \in \mathbb{N}$ be arbitrary subsets of a Banach space $X$. Then the sequence $\{\mathcal{S}_k\}$ is called *Mosco-convergent* to $\mathcal{S}$, denoted by $\mathcal{S}_k \to^M \mathcal{S}$, if

- for every $\bar{x} \in \mathcal{S}$, there exists a sequence $\{x^k\}$ with $x^k \in \mathcal{S}_k$ for all $k \in \mathbb{N}$ such that $x^k \to \bar{x}$;
- whenever $x^k \in \mathcal{S}_k$ for all $k \in \mathbb{N}$ and $x^k \rightharpoonup \bar{x}$ for some $\bar{x} \in X$, then $\bar{x} \in \mathcal{S}$.

The latter definition can now be used to describe continuity of set-valued mappings.

**Definition 4.2.7.** Let $\Psi : X \rightrightarrows Y$ be a set-valued mapping between arbitrary Banach spaces $X$ and $Y$ and let $\bar{x} \in X$. Then $\Psi$ is called *weakly Mosco-continuous in* $\bar{x}$, if $x^k \rightharpoonup \bar{x}$ implies $\Psi(x^k) \to^M \Psi(\bar{x})$.
Clearly, $\Psi$ is called *weakly Mosco-continuous*, if $\Psi$ is weakly Mosco-continuous at all $x \in X$.

Then we can show the following approximation result for Nash equilibrium problems that can be seen as an extension of [22, Theorem 1] to general Banach spaces and set-valued mappings.

**Theorem 4.2.8.** *For all $\nu = 1, \ldots, N$, let $X_\nu$ be a Banach space and $X_{\text{ad}}^\nu : X_{-\nu} \rightrightarrows X_\nu$ a set-valued mapping. Assume that the sequence $\{F^{k,\nu}\}$ multi epi-converges to $\{F^\nu\}$ and that $X_{\text{ad}}^\nu$ is weakly Mosco-continuous for all $\nu = 1, \ldots, N$. Moreover, let $x^k$ be a global Nash equilibrium of GNEP $\{F^{k,\nu}, X_{\text{ad}}^\nu(x^{k,-\nu})\}_{\nu=1}^N$ for all $k \in \mathbb{N}$.*
*If $x^k$ converges strongly to $\bar{x}$, then $\bar{x}$ is a global Nash equilibrium of GNEP $\{F^\nu, X_{\text{ad}}^\nu(\bar{x}^{-\nu})\}_{\nu=1}^N$.*

*Proof.* By assumption, $X_{\text{ad}}^\nu(x^{k,-\nu})$ is Mosco-convergent to $X_{\text{ad}}^\nu(\bar{x}^{-\nu})$ and hence, $\bar{x}^\nu \in X_{\text{ad}}^\nu(\bar{x}^{-\nu})$ for all $\nu = 1, \ldots, N$. Moreover, we have

$$x^{k,\nu} \in \operatorname{argmin} \left\{ F^{k,\nu}(x^\nu, x^{k,-\nu}) \;\middle|\; x^\nu \in X_{\text{ad}}^\nu(x^{k,-\nu}) \right\}$$

for all $\nu = 1, \ldots, N$. By assumption and Definition 4.2.5, the sequence $\{\Psi^{k,\nu}\}$ epi-converges to $\Psi^\nu(x^\nu)$ for all $\nu = 1, \ldots, N$ and hence, we conclude by [11, Corollary 7.20] and the weak Mosco-continuity of $X_{\text{ad}}^\nu$ that

$$\bar{x}^\nu \in \operatorname{argmin} \left\{ F^\nu(x^\nu, \bar{x}^{-\nu}) \;\middle|\; x^\nu \in X_{\text{ad}}^\nu(\bar{x}^{-\nu}) \right\}.$$

Thus, $\bar{x}$ is a Nash equilibrium of GNEP $\{F^\nu, X_{\text{ad}}^\nu(\bar{x}^{-\nu})\}_{\nu=1}^N$. $\square$

Clearly, the concept of multi epi-convergence is quite technical. However, by using Definition 4.2.5 it is sufficient to check the epi-convergence of $\{\Psi^{k,\nu}\}$ for all $\nu = 1, \ldots, N$, which is guaranteed (see [11, Remark 4.9]), if the sequence of functions $\{\Psi^{k,\nu}\}$ is *continuously convergent*, i.e. for all $x^\nu \in X_{\text{ad}}^\nu$ and any sequence $x^{k,\nu}$ such that $x^{k,\nu} \to x^\nu$, we have

$$\Psi^{k,\nu}(x^{k,\nu}) \to \Psi^\nu(x^\nu).$$

In this context, we also refer to [51, Theorem 4.1], where the authors showed the approximation of a Nash equilibrium in arbitrary normed spaces under continuous convergence of $\{\Psi^{k,\nu}\}$. Moreover, notice that the strong convergence of $x^k$ is needed in order to apply [11, Corollary 7.20]. At this point, it is an open question whether it is enough to assume that $x^k$ is only weakly convergent to $\bar{x}$. However, for a specific type of problem, weak convergence is sufficient to ensure that the limit point is a Nash equilibrium.

**Theorem 4.2.9.** *For all $\nu = 1, \dots, N$ and $k \in \mathbb{N}$, let $F^{k,\nu}$,$F^\nu$ be such that it holds*

$$\begin{aligned} z^{k,-\nu} \rightharpoonup \bar{z}^{-\nu} \quad &\Longrightarrow \quad F^{k,\nu}(\cdot, z^{k,-\nu}) \to F^\nu(\cdot, \bar{z}^{-\nu}), \\ z^k \rightharpoonup \bar{z} \quad &\Longrightarrow \quad F^\nu(\bar{z}) \leq \liminf_{k\to\infty} F^{k,\nu}(z^k) \end{aligned}$$

*for arbitrary sequences $\{z^k\} \subseteq X$ with $z^k \rightharpoonup \bar{z}$. Let $\{x^k\} \subseteq X$ be a sequence such that $x^k$ is a Nash equilibrium of $\{F^{k,\nu}, X^\nu_{\text{ad}}\}$ for all $k \in \mathbb{N}$.*
*If there exists $\bar{x} \in X_{\text{ad}}$ such that $x^k \rightharpoonup \bar{x}$, then $\bar{x}$ is a Nash equilibrium of $\{F^\nu, X^\nu_{\text{ad}}\}$.*

*Proof.* Since $\bar{x}$ is feasible, it remains to verify the optimality. Therefore, assume not. Hence, there is $\nu \in \{1, \dots, N\}$ and $\hat{x}^\nu \in X^\nu_{\text{ad}}$ such that $F^\nu(\hat{x}^\nu, \bar{x}^{-\nu}) < F^\nu(\bar{x})$. With the additional properties of $F^{k,\nu}$, $F^\nu$ and that $x^k$ is a NE of $\{F^{k,\nu}, X^\nu_{\text{ad}}\}$, we deduce that

$$F^\nu(\hat{x}^\nu, \bar{x}^{-\nu}) < \liminf_{k\to\infty} F^{k,\nu}(x^{k,\nu}, x^{k,-\nu}) \leq \limsup_{k\to\infty} F^{k,\nu}(\hat{x}^\nu, x^{k,-\nu}) = F^\nu(\hat{x}^\nu, \bar{x}^{-\nu}),$$

which is a contradiction. Thus, $\bar{x}$ is a Nash equilibrium of $\{F^{k,\nu}, X^\nu_{\text{ad}}\}$. □

We refer to Example 4.2.18 for an objective function that satisfies the requirements above.
Now, the following subsections are devoted to auxiliary equilibrium problems of $\Gamma^{\text{lq}}_{\text{EPEC}}$ and $\Gamma^{\text{lq}}_{\text{EPCC}}$, respectively, where we start with the former one.

### 4.2.1. A Penalization Technique for Variational Inequalities

We consider $\Gamma^{\text{lq}}_{\text{EPEC}}$ and require that the lower level feasible set is given by

$$Y_{\text{ad}} := \left\{ y \in Y \;\middle|\; h(y) := Hy + \psi \in K \right\},$$

where $K \subseteq Z$ is a closed and convex set. Moreover, we require that for all $\nu = 1, \dots, N$, $X_\nu \equiv X$ is a Hilbert space and the relation

$$Y \hookrightarrow X \hookrightarrow Y^*$$

is satisfied for all $\nu = 1, \dots, N$.

**Example 4.2.10.** *An important representative are the spaces $L^2(\Omega)$ and $H_0^1(\Omega)$, since it holds*

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

*In particular, notice that the dual space of $H_0^1(\Omega)$ is not identified by the Hilbert space itself but $H^{-1}(\Omega)$.*

We recall that VI (4.9) is given with the assumptions above by

$$\text{Find } y \in Y_{\text{ad}} \text{ s.t. } \langle Ay - Bx - b, v - y\rangle \geq 0 \quad \forall v \in Y_{\text{ad}}. \tag{4.15}$$

and consider the unconstrained minimization problem

$$\min \; \frac{1}{2}\langle Ay, y\rangle_Y - \langle Bx, y\rangle_Y + \frac{1}{2\gamma}\|\pi_\epsilon(y)\|^2, \tag{4.16}$$

where $\gamma > 0$ is a penalty parameter and $\pi_\epsilon : Y \to Y^*$ denotes a mapping that satisfies the subsequent assumptions.

**Assumption 4.2.11.** *Let $\epsilon \geq 0$ and assume that $\pi_\epsilon : Y \to Y^*$*

- *is at least once continuously differentiable for all $\epsilon > 0$;*
- *denotes a penalty function with respect to $Y_{\text{ad}}$, i.e. $\pi_\epsilon(y) = 0 \iff y \in Y_{\text{ad}}$;*
- *is Lipschitz continuous with respect to $y$, i.e. there exists a constant $c_L > 0$ such that for all $y_1, y_2 \in Y$ it holds*

  $$\|\pi_\epsilon(y_1) - \pi_\epsilon(y_2)\|_{Y^*} \leq c_L\|y_1 - y_2\|_X$$

  *for all $\epsilon \geq 0$;*
- *is Lipschitz continuous with respect to $\epsilon$, i.e. there exists a constant $c_l > 0$ such that for all $\epsilon_1, \epsilon_2 > 0$ it holds*

  $$\|\pi_{\epsilon_1}(y) - \pi_{\epsilon_2}(y)\|_{Y^*} \leq c_l|\epsilon_1 - \epsilon_2|$$

  *for all $y \in Y$;*
- *is monotone with respect to $y$, i.e. for all $y_1, y_2 \in Y$, it holds*

  $$\langle \pi_\epsilon(y_1) - \pi_\epsilon(y_2), y_1 - y_2\rangle_Y \geq 0$$

  *for all $\epsilon \geq 0$.*

Notice that the Lipschitz-continuity with respect to $y$ is well-defined, since we assumed that $Y \hookrightarrow X$.
By Subsection 2.3.1, $\bar{y}$ is a solution of (4.16), if it satisfies

$$Ay + \frac{1}{\gamma}\pi_\epsilon(y) - Bx - b = 0, \tag{4.17}$$

which motivates the analysis of the latter equation.

**Example 4.2.12.** *Consider the setting with* $Y = Z = H_0^1(\Omega)$, $K = H_0^1(\Omega)_0^+$ *(see Example 2.2.8) and* $h(y) := y - \psi$*. Then it was shown in [29] that for all* $\epsilon > 0$*, the mappings* $\pi_\epsilon^1(y) := -\max_\epsilon^1(0, \psi - y)$ *and* $\pi_\epsilon^2(y) := -\max_\epsilon^2(0, \psi - y)$ *with*

$$\max_\epsilon^1(0,\omega) := \begin{cases} \omega - \frac{\epsilon}{2} & \text{if } \omega \geq \epsilon, \\ \frac{\omega^2}{2\epsilon} & \text{if } \omega \in (0,\epsilon), \\ 0 & \text{else} \end{cases} \quad \text{and } \max_\epsilon^2(0,\omega) := \begin{cases} \omega - \frac{\epsilon}{2} & \text{if } \omega \geq \epsilon, \\ \frac{\omega^3}{\epsilon^2} - \frac{\omega^4}{2\epsilon^3} & \text{if } \omega \in (0,\epsilon), \\ 0 & \text{else} \end{cases}$$

*satisfy Assumption 4.2.11. In particular, notice that both mappings converge to*

$$\pi(y) := -\max_K(0, -h(y))$$

*as* $\epsilon$ *tends to zero, which can be seen in Figure 4.3. At this point, we emphasize that* $\pi_\epsilon^1$


(a) $\epsilon = 0.3$ (b) $\epsilon = 0.1$

Figure 4.3.: Comparison of $\max_1^\epsilon(0,\omega)$ (blue) and $\max_2^\epsilon(0,\omega)$ (red) for different smoothing parameters $\epsilon$ with limit function $\max_K(0,\omega)$ (dashed).

*is continuously differentiable, while* $\pi_\epsilon^2$ *is even twice continuously differentiable. Further regularizations can be found for instance in [79, 86].*

Recall that one of the sub-goals of this section is to find an auxiliary equilibrium problem of $\Gamma_{\text{EPEC}}^{\text{lq}}$. For this reason, it is shown in the next step that (4.17) admits a unique solution for

all $x \in X$, denoted by $y^\gamma(x)$, and in regard of the consistency verify that $y^\gamma(x)$ converges to the unique solution of VI (4.15), if $\gamma$ tends to zero.
For this reason, we first introduce further properties of mappings between a Banach space and its dual.

**Definition 4.2.13.** A mapping $T : Y \to Y^*$ is called *strongly monotone*, if there exists $c > 0$ such that for all $y_1, y_2 \in Y$, it holds

$$\langle T(y_1) - T(y_2), y_1 - y_2 \rangle_Y \geq c \|y_1 - y_2\|_Y^2.$$

Furthermore, $T$ is called *hemicontinuous*, if the real-valued function

$$\varphi : [0,1] \to \mathbb{R},\ t \mapsto \varphi(t) := \langle T(y_1 + t y_2), y_3 \rangle_Y$$

is continuous on $[0,1]$ for all $y_1, y_2, y_3 \in Y$.

Clearly, any strongly monotone mapping is monotone as well. Now, we can show that equation (4.17) is uniquely solvable for all $x \in X$.

**Lemma 4.2.14.** *Consider the mapping $A_\gamma^\epsilon(y) := Ay + \frac{1}{\gamma}\pi_\epsilon(y)$ in which $\pi_\epsilon$ satisfies Assumption 4.2.11. Moreover, assume that $\langle \pi_\epsilon(y), y \rangle \geq 0$ for all $y \in Y$.*
*Then $A_\gamma^\epsilon$ is strongly monotone, hemicontinuous and coercive and* (4.17) *has a unique solution $y^\gamma(x)$ for all $x \in X^N$.*

*Proof.* We start by showing all properties separately.

a) Coercivity: For all $y \in Y$ we consider

$$\langle A_\gamma^\epsilon(y), y \rangle_Y = \langle Ay, y \rangle_Y + \frac{1}{\gamma} \langle \pi_\epsilon(y), y \rangle_Y.$$

By assumption, the second term is non-negative. Hence, the result follows by the coercivity of $A$.

b) Strong monotonicity: Let $y_1, y_2 \in Y$ be arbitrary. Then, we have

$$\langle A_\gamma^\epsilon(y_1) - A_\gamma^\epsilon(y_2), y_1 - y_2 \rangle_Y = \langle Ay_1 - Ay_2, y_1 - y_2 \rangle_Y + \frac{1}{\gamma} \langle \pi_\epsilon(y_1) - \pi_\epsilon(y_2), y_1 - y_2 \rangle_Y.$$

Similar as before, we observe that the latter term is non-negative by the monotonicity of $\pi_\epsilon$. Since coercivity and strong monotonicity are equivalent for linear operators, the result follows again by the coercivity of $A$, i.e.

$$\langle A_\gamma^\epsilon(y_1) - A_\gamma^\epsilon(y_2), y_1 - y_2 \rangle_Y \geq \langle A[y_1 - y_2], y_1 - y_2 \rangle_Y \geq c \|y_1 - y_2\|_Y^2.$$

c) Hemicontinuity: Fix $y_1, y_2, y_3 \in Y$. Then for all $t_1, t_2 \in [0,1]$ we have

$$\begin{aligned}|\varphi(t_1) - \varphi(t_2)| &= |\langle A^\epsilon_\gamma(y_1 + t_1 y_2) - A^\epsilon_\gamma(y_1 + t_2 y_2), y_3\rangle_Y| \\ &\leq |\langle (t_1 - t_2) A y_2, y_3\rangle_Y| + \frac{1}{\gamma}|\langle \pi_\epsilon(y_1 + t_1 y_2) - \pi_\epsilon(y_1 + t_2 y_2), y_3\rangle_Y|,\end{aligned}$$

where we used the linearity of $A$ and the triangle inequality. Due to the boundedness of $A$ and the Lipschitz continuity of $\pi$ with respect to $y$, we further obtain

$$|\varphi(t_1) - \varphi(t_2)| \leq ((\|A\|_{L(Y,Y^*)} + \frac{c_L}{\gamma})\|y_2\|\|y_3\|)|t_1 - t_2|,$$

which shows the continuity of $\varphi$ and hence, the hemicontinuity of $A^\epsilon_\gamma$.

Now, we can apply the main theorem on monotone operators (see [99, Theorem 26.A]), which shows that equation (4.17) has a unique solution. □

Observe that the additional assumption is satisfied if $\psi \in K$, since this implies that $0 \in Y_{\text{ad}}$ and we obtain $\langle \pi_\epsilon(y), y\rangle \geq 0$ for all $y \in Y$ by the monotonicity of $\pi_\epsilon$. Moreover, we emphasize at this point the requirement that $h(y)$ is independent of $x$, in order to use the main theorem on monotone operators [99, Theorem 26.A]. In particular, it is needed that the mapping $A^\epsilon_\gamma(y)$ only depends on $y$.

**Remark 4.2.15.** *Although we have introduced operators in the beginning as linear mappings, the mentioned theorem above is not restricted solely to linear mappings. The reason for the naming is that the terms operator and mapping are often used synonymously in literature.*

The subsequent lemma, which verifies the convergence of $y^\gamma(x)$ to the unique solution of VI (4.15) can be seen as a generalization of [29, Theorem 2.3] to abstract Hilbert spaces.

**Lemma 4.2.16.** *Let $\{\epsilon_k\} \subseteq \mathbb{R}^+$ be a bounded sequence and assume that $X \hookrightarrow\hookrightarrow Y^*$. Furthermore, let $\{x^k\} \subseteq X^N$ be a sequence such that $x^k \rightharpoonup \bar{x}$ for some $\bar{x} \in X^N$. Denote by $y^k$ the solution of*

$$Ay + \frac{1}{\gamma_k}\pi_{\epsilon_k}(y) - Bx^k - b = 0 \tag{4.18}$$

*with respect to $x^k$, where $\pi_{\epsilon_k}$ satisfies Assumption 4.2.11.*
*Then we have*

$$y^k \to y(\bar{x}) =: \bar{y} \quad \text{in } Y,$$

*if $\gamma_k \to 0$ as $k \to \infty$ and $\bar{y}$ is the solution of VI (4.15) w.r.t. $\bar{x}$.*

Motivated by [21], the proof basically follows the ideas in [29] for the special case in $H_0^1(\Omega)$.

*Proof.* Let $v \in Y_{\text{ad}}$ be arbitrary but fixed. Then we obtain by the monotonicity of $\pi_{\epsilon_k}$ that

$$0 \leq \langle \pi_{\epsilon_k}(y^k) - \pi_{\epsilon_k}(v), y^k - v\rangle_Y = \langle \pi_{\epsilon_k}(y^k), y^k - v\rangle_Y, \tag{4.19}$$

which together with the coercivity of $A$ implies that

$$\begin{aligned} c\|y^k\|_Y^2 &\leq \langle Ay^k, y^k\rangle_Y = \langle Bx^k + b, y^k - v\rangle_Y - \langle \pi_{\epsilon_k}(y^k), y^k - v\rangle_Y + \langle Ay^k, v\rangle_Y \\ &\leq \langle Bx^k + b, y^k - v\rangle_Y + \langle Ay^k, v\rangle_Y. \end{aligned}$$

Due to the boundedness of $A$ and $B$, the sequence $\{y^k\}$ is bounded itself and there exists $\bar{y} \in Y$ such that $y^k \rightharpoonup \bar{y}$ in $Y$ on a subsequence (we denote the same). Moreover, we deduce by (4.18) that

$$\lim_{k\to\infty} \pi_{\epsilon_k}(y^k) = \lim_{k\to\infty} \gamma_k(Bx^k + b - Ay^k) = 0, \tag{4.20}$$

if $\gamma_k$ tends to $0$ as $k \to \infty$, since the expression within the brackets is bounded.
By the boundedness of $\{\epsilon_k\}$, we again find $\bar{\epsilon} \in \mathbb{R}$ such that $\epsilon_k \to \bar{\epsilon}$ (on a subsequence) and obtain by the Lipschitz continuity of $\pi_{\epsilon_k}$, the triangle inequality and the strong convergence of $y^k$ in $X$ that

$$\begin{aligned} \|\pi_{\bar{\epsilon}}(\bar{y}) - \pi_{\epsilon_k}(y^k)\|_{Y^*} &\leq \|\pi_{\bar{\epsilon}}(\bar{y}) - \pi_{\bar{\epsilon}}(y^k)\|_{Y^*} + \|\pi_{\bar{\epsilon}}(y^k) - \pi_{\epsilon_k}(y^k)\|_{Y^*} \\ &\leq \max(c_L, c_l)(\|\bar{y} - y^k\|_X + |\bar{\epsilon} - \epsilon_k|) \\ &\to 0. \end{aligned}$$

Due to the uniqueness of weak limit points, the latter result and (4.20) imply that $\pi_{\bar{\epsilon}}(\bar{y}) = 0$ and hence, $\bar{y} \in Y_{\text{ad}}$.
In the next step, we show that $\bar{y}$ solves VI (4.15). For this reason, consider (4.18) again and observe that we obtain

$$\begin{aligned} \langle A\bar{y}, \bar{y}\rangle_Y &\leq \liminf_{k\to\infty} \langle Ay^k, y^k\rangle_Y = \liminf_{k\to\infty}(\langle Bx^k + b - \pi_{\epsilon_k}(y^k), y^k - v\rangle_Y + \langle Ay^k, v\rangle_Y) \\ &\leq \liminf_{k\to\infty}(\langle Bx^k + b, y^k - v\rangle_Y + \langle Ay^k, v\rangle_Y) \\ &= \langle B\bar{x} + b, \bar{y} - v\rangle_Y + \langle A\bar{y}, v\rangle_Y, \end{aligned}$$

where we used the coercivity of $A$, the strong convergence $x^k \to \bar{x}$ in $(Y^*)^N$ and (4.19). Since $v \in Y_{\text{ad}}$ was chosen arbitrarily, this implies that

$$\langle A\bar{y} - B\bar{x} - b, v - \bar{y}\rangle \geq 0 \qquad \forall v \in Y_{\text{ad}}$$

and hence, $\bar{y} = y(\bar{x})$ by the unique solvability of VI (4.15). The uniqueness of $\bar{y}$ implies that $y^k \rightharpoonup \bar{y}$ on the whole sequence. It remains to show the strong convergence of $y^k$ in $Y$. For this reason, we utilize again the coercivity of $A$, (4.18) and (4.19) in order to estimate

$$\begin{aligned} c\|y^k - \bar{y}\|_Y^2 &\leq \langle A[y^k - \bar{y}], y^k - \bar{y}\rangle_Y \\ &= \langle Bx^k + b, y^k - \bar{y}\rangle_Y - \langle \pi_{\epsilon_k}(y^k), y^k - \bar{y}\rangle_Y - \langle A\bar{y}, y^k - \bar{y}\rangle_Y \\ &\leq \langle Bx^k + b, y^k - \bar{y}\rangle_Y - \langle A\bar{y}, y^k - \bar{y}\rangle_Y. \end{aligned}$$

Thus, we obtain that

$$c\|y^k - \bar{y}\|_Y \leq \|Bx^k + b - A\bar{y}\|_{Y^*},$$

where the right hand side tends to zero by (4.18) and hence, $\|y^k - \bar{y}\| \to 0$ as $k \to \infty$. □

We denote by $y^k(x)$ the solution mapping of (4.18) and define the following equilibrium problem

$$\begin{aligned} \min\ & F^{k,\nu}(x^\nu, x^{-\nu}) := f^\nu(x^\nu, x^{-\nu}, y^k(x^\nu, x^{-\nu})) \\ \text{w.r.t.}\ & x^\nu \in X, \\ \text{s.t.}\ & x^\nu \in Q_\nu. \end{aligned}$$

As a consequence of Lemma 4.2.16, we are now able to prove the following consistency result. Therefore, we state an additional assumption.

**Theorem 4.2.17.** *Let $\{\epsilon_k\} \subseteq \mathbb{R}_+$ be a bounded sequence and $\{x^k\} \subseteq X^N$ be such that $x^k$ is a Nash equilibrium of $\{F^{k,\nu}, Q_\nu\}$ for all $k \in \mathbb{N}$.*
*Then there exists $(\bar{x}, \bar{y}) \in X^N \times Y$ such that $x^k \rightharpoonup \bar{x}$ in $X^N$ on a subsequence and $y^k \to \bar{y}$ in $Y$, if $\gamma_k \to 0$ as $k \to \infty$ and $(\bar{x}, \bar{y})$ is a feasible point of $\Gamma^{\mathrm{lq}}_{\mathrm{EPEC}}$.*
*If $F^{k,\nu}$ and $F^\nu$ satisfy the requirements in Theorem 4.2.9 for all $\nu = 1, \dots, N$, then $(\bar{x}, \bar{y})$ is a leader-follower Nash equilibrium of $\Gamma^{\mathrm{lq}}_{\mathrm{MLFG}}$.*

*Proof.* Since $Q$ is a closed, bounded and convex subset in a reflexive Banach space $X^N$, there exists $\bar{x} \in Q$ and $x^k \rightharpoonup \bar{x}$ in $X^N$ on a subsequence. Hence, we can apply the previous lemma and find $\bar{y} \in Y$ such that $y^k(x^k) \to y(\bar{x}) = \bar{y}$ in $Y$ and $\bar{y}$ is the unique solution of VI (4.15). This implies that $(\bar{x}, \bar{y})$ is feasible for $\Gamma^{\mathrm{lq}}_{\mathrm{EPEC}}$, which shows the first part.
The second part is a direct consequence of Theorem 4.2.9. □

The following example shows that the requirements given in Theorem 4.2.9 are not too strong.

**Example 4.2.18.** *Consider the tracking-type functional*

$$f^\nu(x^\nu, y) := \frac{1}{2}\|y - y_d^\nu\|_Y^2 + \frac{\alpha_\nu}{2}\|x^\nu\|_X^2$$

*and the corresponding reduced forms*

$$\begin{aligned} F^\nu(x^\nu, x^{-\nu}) &:= \frac{1}{2}\|y(x^\nu, x^{-\nu}) - y_d^\nu\|_Y^2 + \frac{\alpha_\nu}{2}\|x^\nu\|_X^2, \\ F^{k,\nu}(x^\nu, x^{-\nu}) &:= \frac{1}{2}\|y^k(x^\nu, x^{-\nu}) - y_d^\nu\|_Y^2 + \frac{\alpha_\nu}{2}\|x^\nu\|_X^2 \end{aligned}$$

*where $y(x)$ and $y^k(x)$ denote the solution mappings of VI* (4.15) *and the monotone equation* (4.18)*, respectively. For an arbitrary sequence $\{x^k\}$ with $x^k \rightharpoonup \bar{x}$, we obtain on the one hand that*

$$F^\nu(\bar{x}) \leq \liminf_{k\to\infty} F^{k,\nu}(x^k),$$

*since $f^\nu$ is continuous with respect to $y$ and weakly lower semicontinous with respect to $x^\nu$. On the other hand it follows for arbitrary $\hat{x}^\nu \in X$ by by Lemma 4.2.16 that*

$$\begin{aligned} F^{k,\nu}(\hat{x}^\nu, x^{k,-\nu}) &= \frac{1}{2}\|y^k(\hat{x}^\nu, x^{k,-\nu}) - y_d^\nu\|_Y^2 + \frac{\alpha_\nu}{2}\|\hat{x}^\nu\|_X^2 \\ &\to \frac{1}{2}\|y(\hat{x}^\nu, \bar{x}^{-\nu}) - y_d^\nu\|_Y^2 + \frac{\alpha_\nu}{2}\|\hat{x}^\nu\|_X^2 = F^\nu(\hat{x}^\nu, \bar{x}^{-\nu}) \end{aligned}$$

*and hence, the requirements of Theorem 4.2.9 are satisfied.*

### 4.2.2. A Regularization Technique for Complementarity Constraints

In contrast to Subsection 4.2.1, we now focus on $\Gamma_{\text{EPCC}}^{\text{lq}}$ and introduce the auxiliary equilibrium problem $\Gamma_2^k$, where leader $\nu$ solves

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}, y) - \frac{1}{\gamma_k}\langle p, h(x^\nu, x^{-\nu}, y)\rangle_Z \\ \text{w.r.t. }\ & (x^\nu, y, p) \in X_\nu \times Y \times Z^*, \\ \text{s.t. }\ & x^\nu \in Q_\nu,\ h(x^\nu, x^{-\nu}, y) \in K,\ p \in K^\circ, \\ & Ay - B[x^\nu, x^{-\nu}] + H^* p = b \end{aligned} \tag{4.21}$$

for all $\nu = 1, \dots, N$. Moreover, we require that Assumption 4.2.1 is satisfied. In particular, we consider again $Y_{\text{ad}}(x)$ with the cone $K$. Additionally, we assume that $f^\nu$ is bounded from below for all $\nu = 1, \dots, N$. Similar to the previous section, our goal is to show that

$\Gamma_2^k$ approximates $\Gamma_{\text{EPCC}}^{\text{lq}}$ in the sense of Theorem 4.2.8. For this reason, we first prove an auxiliary result that can be seen as a generalization of [31, Lemma 2.2] to our abstract scenario. In particular, the authors [31] considered a tracking-type setting in $H_0^1(\Omega)$ and $L^2(\Omega)$ with the lower level feasible set $\{y \in H_0^1(\Omega) \mid y \geq 0\}$. For this reason, notice that additional assumptions are required such that our counterpart, i.e.

$$Y_{\text{ad}}(x) = \left\{y \in Y \,\middle|\, h(x,y) = \hat{h}(x) + Hy + \psi \in K\right\},$$

is applicable.

**Lemma 4.2.19.** *Let $\{x^k\} \subseteq X$ be a bounded sequence and $\{(y^k, p^k)\} \subseteq Y \times Z^*$ be such that*

$$Ay^k + H^* p^k - Bx^k = b, \tag{4.22a}$$

$$h(x^k, y^k) \in K, \; p^k \in K^\circ \tag{4.22b}$$

*and (for some $\hat{c} > 0$)*

$$|\langle p^k, Hy^k \rangle_Y| \leq \hat{c} \tag{4.23}$$

*are satisfied for all $k \in \mathbb{N}$. Furthermore, let the mapping $\hat{h}$ be weakly sequentially continuous. Then there exists $(\bar{x}, \bar{y}, \bar{p}) \in X \times Y \times Z^*$ such that*

$$(x^k, y^k, p^k) \rightharpoonup (\bar{x}, \bar{y}, \bar{p}) \text{ in } X \times Y \times Z^*$$

*on a subsequence, which satisfies* (4.22). *If $\hat{h}$ and $B$ are completely continuous, then it holds*

$$\liminf_{k\to\infty} \langle p^k, h(x^k, y^k) \rangle_Z \leq \langle \bar{p}, h(\bar{x}, \bar{y}) \rangle_Z \leq 0 \tag{4.24}$$

*and if*

$$\langle p^k, h(x^k, y^k) \rangle_Z \to 0 \quad \text{as } k \to \infty$$

*is additionally valid, then $(\bar{x}, \bar{y}, \bar{p})$ fulfills the complementarity condition, i.e. it holds $\langle \bar{p}, h(\bar{x}, \bar{y}) \rangle_Z = 0$.*

*Proof.* Let $\{(x^k, y^k, p^k)\}$ be a sequence such that the assumptions hold. Due to (4.22) and the coercivity of $A$, there exists a constant $c > 0$ such that

$$c\|y^k\|_Y^2 \leq \langle Ay^k, y^k \rangle_Y = \langle Bx^k + b, y^k \rangle_Y - \langle H^* p^k, y^k \rangle_Y.$$

Hence, (4.23) and the boundedness of $B$ and $\{x^k\}$ imply that $\{y^k\}$ is bounded in $Y$. As a consequence, we directly obtain that

$$\|p^k\|_{Z^*} = \|(H^*)^{-1}(Bx^k + b - Ay^k)\|_{Z^*},$$

and thus, $p^k$ is bounded in $Z^*$. Here, we used the coercivity of $H$, which implies the existence of the bounded operator $(H^*)^{-1}$. Thus, there exists $(\bar{x}, \bar{y}, \bar{p}) \in X \times Y \times Z^*$ such that

$$(x^k, y^k, p^k) \rightharpoonup (\bar{x}, \bar{y}, \bar{p}) \text{ in } X \times Y \times Z^*$$

on a subsequence that we denote the same. Since $K$ and $K^\circ$ are closed and convex, they are weakly sequentially closed (see Subsection 2.1.1). Together with the weak sequential continuity of $h$ w.r.t $x$ and the linearity w.r.t $y$, we further conclude that

$$h(x^k, y^k) = \hat{h}(x^k) + Hy^k + \psi \rightharpoonup \hat{h}(\bar{x}) + H\bar{y} + \psi = h(\bar{x}, \bar{y})$$

and hence, it holds $h(\bar{x}, \bar{y}) \in K$ and $\bar{p} \in K^\circ$. Moreover, we obtain that

$$0 = Ay^k + H^*p^k - Bx^k - b \rightharpoonup A\bar{y} + H^*\bar{p} - B\bar{x} - b \quad \text{in } Y^*, \tag{4.25}$$

which implies $A\bar{y} + H^*\bar{p} = B\bar{x} + b$ and consequently, the first part of the proof is shown. In order to verify the second assertion, we compute

$$\begin{aligned}
&\liminf_{k\to\infty} \langle p^k, h(x^k, y^k)\rangle_Z = \\
&= \liminf_{k\to\infty} \langle (H^{-1})^*(Bx^k + b - Ay^k), \hat{h}(x^k) + Hy^k + \psi\rangle_Z \\
&= \liminf_{k\to\infty} (\langle Bx^k + b, y^k\rangle_Y - \langle Ay^k, y^k\rangle_Y + \langle (H^{-1})^*(Bx^k + b - Ay^k), \hat{h}(x^k) + \psi\rangle_Z) \\
&\leq \langle B\bar{x} + b, \bar{y}\rangle_Y - \langle A\bar{y}, \bar{y}\rangle_Y + \langle (H^{-1})^*(B\bar{y} + b - A\bar{y}), \hat{h}(\bar{x}) + \psi\rangle_Z \\
&= \langle B\bar{x} + b - A\bar{y}, (H^{-1}H)\bar{y}\rangle_Y + \langle (H^{-1})^*(B\bar{y} + b - A\bar{y}), \hat{h}(\bar{x}) + \psi\rangle_Z \\
&= \langle (H^{-1})^*(B\bar{x} + b - A\bar{y}), \hat{h}(\bar{x}) + H\bar{y} + \psi\rangle_Z \\
&= \langle \bar{p}, h(\bar{x}, \bar{y})\rangle_Z \leq 0,
\end{aligned}$$

where we used the identity $(H^*)^{-1} = (H^{-1})^*$, the coercivity of $A$ and the complete continuity of $\hat{h}$ and $B$, respectively.
The last part follows directly with (4.24). □

Notice that the price of generalizing [31, Theorem 2.2] goes along with strong requirements on the involved mappings and operators. In particular, the complete continuity is not a standard assumption. However, notice (and already mentioned in Subsection 2.1.1) that the identity $I_{YX}$ between compactly embedded Banach spaces $X$ and $Y$ denotes a compact operator and thus, is completely continuous.
Next, we define the feasible set of player $\nu$'s auxiliary problem by

$$\tilde{X}^\nu_{\text{ad}}(x^{-\nu}) = \left\{(x^\nu, y, p) \in Q_\nu \times Y \times K^\circ \;\middle|\; Ay - B[x^\nu, x^{-\nu}] + H^*p = b,\ h(x^\nu, x^{-\nu}, y) \in K\right\}$$

and observe that each leader $\nu$'s feasible set in $\Gamma^{\mathrm{lq}}_{\mathrm{EPCC}}$ is given by

$$X^{\nu}_{\mathrm{ad}}(x^{-\nu}) = \big\{(x^{\nu}, y, p) \in \tilde{X}^{\nu}_{\mathrm{ad}}(x^{-\nu}) \;\big|\; \langle p, h(x^{\nu}, x^{-\nu}, y)\rangle_Z = 0\big\}.$$

The subsequent result verifies that the limit point of a sequence of Nash equilibria is a feasible point of $\Gamma^{\mathrm{lq}}_{\mathrm{EPCC}}$ under suitable assumptions.

**Theorem 4.2.20.** *Let $\{(x^k, y^k, p^k)\} \subseteq X \times Y \times Z^*$ be a sequence, where $(x^k, y^k, p^k)$ is a Nash equilibrium of $\Gamma^k_2$, $|\langle p^k, Hy^k\rangle_Y|$ is bounded and $X^{\nu}_{\mathrm{ad}}(x^{k,-\nu})$ is nonempty for all $\nu = 1, \dots, N$ and $k \in \mathbb{N}$. Furthermore, assume that $B$ and $\hat{h}$ are completely continuous and for all $\nu = 1, \dots, N$, $f^{\nu}$ is bounded in $x^{-\nu}$.*
*Then there exists $(\bar{x}, \bar{y}, \bar{p}) \in X \times Y \times Z^*$ such that*

$$x^k \rightharpoonup \bar{x} \quad \text{in } X,$$

*on a subsequence (denoted the same),*

$$(y^k, p^k) \to (\bar{y}, \bar{p}) \quad \text{in } Y \times Z^*$$

*on the whole sequence and $(\bar{x}, \bar{y}, \bar{p})$ is a feasible point of $\Gamma^{\mathrm{lq}}_{\mathrm{EPCC}}$.*

*Proof.* Since $Q_{\nu}$ is a nonempty, closed, convex and bounded subset of a reflexive Banach space $X_{\nu}$ for all $\nu = 1, \dots, N$, it follows by observations in Subsection 2.1.1 that $Q_{\nu}$ is weakly sequentially compact and hence, there exists $\bar{x} \in Q$ such that $x^k \rightharpoonup \bar{x}$. By assumption, $(x^k, y^k, p^k)$ satisfies the system (4.22). Thus, we can apply the first part of Lemma 4.2.19 and conclude the existence of $(\bar{y}, \bar{p}) \in Y \times Z^*$ such that $y^k \rightharpoonup \bar{y}$ in $Y$ and $p^k \rightharpoonup \bar{p}$ in $Z^*$ on a subsequence (denoted the same), which fulfill (4.22).
Since $X^{\nu}_{\mathrm{ad}}(x^{k,-\nu})$ is nonempty for all $\nu = 1, \dots, N$, we find $(\hat{x}^{\nu}, \hat{y}, \hat{p}) \in X^{\nu}_{\mathrm{ad}}(x^{k,-\nu})$ such that it holds

$$\begin{aligned} f^{\nu}(\hat{x}^{\nu}, x^{k,-\nu}, \hat{y}) &= f^{\nu}(\hat{x}^{\nu}, x^{k,-\nu}, \hat{y}) - \frac{1}{\gamma_k}\langle \hat{p}, h(\hat{x}^{\nu}, x^{k,-\nu}, \hat{y})\rangle_Z \\ &\geq f^{\nu}(x^k, y^k) - \frac{1}{\gamma_k}\langle p^k, h(x^k, y^k)\rangle_Z \\ &\geq c_{\nu} - \frac{1}{\gamma_k}\langle p^k, h(x^k, y^k)\rangle_Z \geq c_{\nu}. \end{aligned}$$

where we used the optimality of $(x^k, y^k, p^k)$ and that $f^{\nu}$ is bounded from below (with constant $c_{\nu} \in \mathbb{R}$). Hence, we deduce by the boundedness of $f^{\nu}$ in $x^{-\nu}$ that $\gamma_k^{-1}\langle p^k, h(x^k, y^k)\rangle_Z$ is uniformly bounded, which implies

$$\langle p^k, h(x^k, y^k)\rangle_Z \to 0,$$

as $k \to \infty$. Thus, we can apply the remaining parts of Lemma 4.2.19 and conclude that $(\bar{x}, \bar{y}, \bar{p})$ is feasible for $\Gamma^{\text{lq}}_{\text{EPCC}}$.
In order to show the strong convergence, observe that

$$\begin{aligned} c\|y^k - \bar{y}\|_Y^2 &\le \langle A[y^k - \bar{y}], y^k - \bar{y}\rangle_Y \\ &= \langle B[x^k - \bar{x}], y^k - \bar{y}\rangle_Y - \langle H^*[p^k - \bar{p}], y^k - \bar{y}\rangle_Y \end{aligned}$$

holds by (4.25) and the coercivity of $A$. Since $B$ is completely continuous, the first term tends to zero by Lemma 2.1.6. For the second term, we obtain by the complete continuity of $\hat{h}$, the weak convergence of $(y^k, p^k)$ and $\langle p^k, h(x^k, y^k)\rangle \to 0$ (which implies by Lemma 4.2.19 that $\langle \bar{p}, h(\bar{x}, \bar{y})\rangle_Z = 0$) that

$$\begin{aligned} \langle H^*[p^k - \bar{p}], y^k - \bar{y}\rangle_Y &= \langle p^k - \bar{p}, H[y^k - \bar{y}]\rangle_Z \\ &= \langle p^k, h(x^k, y^k) - h(\bar{x}, \bar{y})\rangle_Z - \langle p^k, \hat{h}(x^k) - \hat{h}(\bar{x})\rangle_Z \\ &\quad - \langle \bar{p}, h(x^k, y^k) - h(\bar{x}, \bar{y})\rangle_Z + \langle \bar{p}, \hat{h}(x^k) - \hat{h}(\bar{x})\rangle_Z \\ &\to 0. \end{aligned}$$

Combining both observations yield the strong convergence of $y^k$ to $\bar{y}$.
Similarly, we estimate

$$\begin{aligned} \|p^k - \bar{p}\|_{Z^*} &= \|(H^*)^{-1}(B[x^k - \bar{x}] - A[y^k - \bar{y}])\|_{Z^*} \\ &\le c_H(\|B[x^k - \bar{x}]\|_{Y^*} + c_A\|y^k - \bar{y}\|_Y), \end{aligned}$$

where the complete continuity of $B$ and the boundedness of $A$ and $(H^*)^{-1}$ (with constants $c_A$ and $c_H$) are used. As a consequence, $\{p^k\}$ converges strongly to $\bar{p}$ in $Z^*$. ☐

In contrast to the penalization technique presented in Subsection 4.2.1, the approach we use here results in a feasible set for each leader that is not independent of his opponents' strategies. In order to apply Theorem 4.2.8, we need to assure that $x^{k,-\nu} \rightharpoonup \bar{x}^{-\nu}$ implies $\tilde{X}^{\nu}_{\text{ad}}(x^{k,-\nu}) \to^M X^{\nu}_{\text{ad}}(\bar{x}^{-\nu})$ for all $\nu = 1, \dots, N$. In particular, the existence of a strongly convergent sequence in $\tilde{X}^{\nu}_{\text{ad}}(x^{k,-\nu})$ to an arbitrary element in $X^{\nu}_{\text{ad}}(\bar{x}^{-\nu})$ (see Definition 4.2.6) is no trivial task and needs in general the validity of a regularity condition (see e.g. [33]). In [33, Theorem 2.5], the authors required the existence of point that is located in the interior of the feasible set of the original equilibrium problem. Due to the structure of $\Gamma^{\text{lq}}_{\text{EPCC}}$, this assumption does not fit in our scenario. As consequence, it is not obvious whether a limit point that is generated in Theorem 4.2.20 is also a LF Nash equilibrium of $\Gamma^{\text{lq}}_{\text{MLFG}}$. The theorem below gives a sufficient condition such that we obtain a LF Nash equilibrium.

**Theorem 4.2.21.** *Let* $(x^k, y^k, p^k) \in X \times Y \times Z^*$ *be a Nash equilibrium of* $\Gamma_2^k$ *and assume that it is also a feasible point of* $\Gamma_{\text{EPCC}}^{\text{lq}}$. *Then* $(x^k, y^k, p^k)$ *is a LF Nash equilibrium of* $\Gamma_{\text{MLFG}}^{\text{lq}}$.

*Proof.* Assume that $(x^k, y^k, p^k)$ is no LF Nash equilibrium. Then there exists $\nu \in \{1, \ldots, N\}$ and $(\hat{x}^\nu, \hat{y}, \hat{p}) \in X_{\text{ad}}^\nu(x^{k,-\nu})$ such that

$$\begin{aligned} f^\nu(\hat{x}^\nu, x^{k,-\nu}, \hat{y}) &< f^\nu(x^k, y^k) \\ &= f^\nu(x^k, y^k) - \frac{1}{\gamma_k}\langle p^k, h(x^k, y^k)\rangle \\ &\leq f^\nu(\hat{x}^\nu, x^{k,-\nu}, \hat{y}), \end{aligned}$$

where the latter inequality follows by the inclusion $X_{\text{ad}}^\nu(x^{k,-\nu}) \subseteq \tilde{X}_{\text{ad}}^\nu(x^{k,-\nu})$ and that $(x^k, y^k, p^k)$ is a NE of $\Gamma_2^k$. Hence, we have a contradiction and $(x^k, y^k)$ is a LF Nash equilibrium of $\Gamma_{\text{MLFG}}^{\text{lq}}$. □

The assumption that there is $\tilde{k} \in \mathbb{N}$ such that $(x^{\tilde{k}}, y^{\tilde{k}}, p^{\tilde{k}})$ is a feasible point of $\Gamma_{\text{EPEC}}^{\text{lq}}$ is not too strict as the $l1$-regularization that we used here is an exact penalization method. However, notice that the requirement of being a Nash equilibrium of $\Gamma_k^2$ is the main task that can not be proven in general (see Subsection 2.3.2). Despite these difficulties, the latter regularization approach is still worth the effort to analyze as it may also lead to stronger stationary Nash equilibria (see Section 5.3).

# 5. A Multi-Leader Optimal Control Framework of the Obstacle Problem

After introducing general MLMFGs and a special MLFG with a quadratic lower level program in Chapter 4, our next goal is the analysis of an hierarchical game in a particular Hilbert space setting. Motivated by a high number of publications (see e.g. [29, 30, 31, 32, 48, 59, 79, 95, 96]) in recent years, we extend the well-known optimal control of the obstacle problem in $H_0^1(\Omega)$ to equilibrium problems.
Starting with the early work by Yvon [98], the latter problem class gained increasing interest in the past decades (see e.g. [4, 52, 58, 59]), whose results are still used today. For instance, in [59] the authors derived and analyzed necessary conditions of local solutions in the absence of control constraints. In the last decade, contributions by Hintermüller and several co-authors focused on penalty techniques to derive C-stationary points with [29, 31, 32] and without [30] control constraints. While in [30] the authors used a relaxation scheme in the sense of a Scholtes-regularization [80] to generate an auxiliary problem, a $l1$-penalty scheme was applied in [31]. Moreover, the works [29, 79] approximated the whole variational inequality with the difference that the latter contribution considered a stronger concept of C-stationarity. Notice that in contrast to [29, 30, 31, 32, 48, 79], the other publications [32, 59, 95, 96] assumed less regular data, which influenced the corresponding stationarity concepts. In this context, [32, 95] extended the work in [59] towards an optimal control problem with control constraints, while in [96] the same author focused on a regularization scheme that lead to an M-stationary point of the original problem.
In this chapter we extend these ideas to equilibrium problems and consider a multi-leader optimal control problem (MLOCP) with pointwise control constraints. For this reason, Section 5.1 is dedicated to the general problem formulation and to the basic classification into the setting of MLFGs. In Section 5.2, weaker forms of stationary Nash equilibria in $H_0^1(\Omega)$ are considered, where we additionally present a short overview of basic definitions of capacity theory. Therein, these concepts can be seen as an extension of the stationarity conditions for MPCCs in $H_0^1(\Omega)$ that can be found in [96]. In Section 5.3 we apply the methods introduced in Subsection 4.2.1 and Subsection 4.2.2, respectively, to verify the

convergence to an $\epsilon$-almost C-stationary Nash equilibrium. In this context, the latter equilibrium concept denotes a generalization of the corresponding stationarity condition, which was stated in [30]. Finally, Section 5.4 is devoted to the numerical analysis.

## 5.1. Problem Formulation

Throughout this section, we consider a multi-leader optimal control problem, denoted by $\Gamma_{\text{MLOCP}}$, where each leader $\nu = 1, \dots, N$ solves a parametric bilevel optimal control problem of tracking type, i.e.

$$\begin{aligned} \min \;\; & f^\nu(x^\nu, y) := \frac{1}{2}\|y - y_d^\nu\|_{L^2(\Omega)}^2 + \frac{\alpha_\nu}{2}\|x^\nu\|_{L^2(\Omega)}^2 \\ \text{w.r.t.} \;\; & (x^\nu, y) \times L^2(\Omega) \times H_0^1(\Omega), \\ \text{s.t.} \;\; & a_\nu \le x^\nu \le b_\nu, \\ & y \in \Psi(x^\nu, x^{-\nu}). \end{aligned} \tag{5.1}$$

Here, $\Psi(x)$ denotes the solution set of the following quadratic program

$$\begin{aligned} \min \;\; & \frac{1}{2}\langle Ay, y\rangle_{H_0^1(\Omega)} - \langle Bx + b, y\rangle_{H_0^1(\Omega)} \\ \text{w.r.t.} \;\; & y \in H_0^1(\Omega), \\ \text{s.t.} \;\; & y \ge 0. \end{aligned} \tag{5.2}$$

Moreover, we assume that $\Omega \subseteq \mathbb{R}^d$ with $d \in \{1, 2, 3\}$ is an open and bounded subset and $\partial\Omega$ is sufficiently regular, i.e. a Lipschitz-boundary (see e.g. [88]).
In this context, the follower's strategy $y \in H_0^1(\Omega)$ is typically called the *state* of the underlying system, while the leader strategies are known as *controls*.
In order to establish the relationship to general MLFGs considered in Section 4.2, observe that we have the following data:

- $X_\nu = L^2(\Omega)$, $Y = Z = H_0^1(\Omega)$ and $Y^* = H^{-1}(\Omega)$;
- $Q_\nu = \{x^\nu \in L^2(\Omega) \mid a_\nu \le x^\nu \le b_\nu \text{ a.e. in } \Omega\}$ with $a_\nu, b_\nu \in L^2(\Omega)$ and $a_\nu \le b_\nu$ a.e. in $\Omega$;
- $Y_{\text{ad}} = \{y \in H_0^1(\Omega) \mid y \ge 0\}$, i.e. we have $h : H_0^1(\Omega) \to H_0^1(\Omega),\ h(y) = y$, $H \equiv I_{H_0^1(\Omega)}$ and $K = H_0^1(\Omega)_0^+$;
- $y_d^\nu \in L^2(\Omega)$, $\alpha_\nu > 0$ and $b \in H^{-1}(\Omega)$.

Notice that this setting also allows to consider state constraints $y \geq \psi$ for arbitrary $\psi \in H^1(\Omega)$ that satisfy $\psi \leq 0$ on $\partial\Omega$, since the latter inequality is equivalent to $\tilde{y} \geq 0$ with $\tilde{y} := y - \psi$ and hence, we obtain a quadratic program that is similar to (5.2). Moreover, we emphasize that $f^\nu$ and $Q_\nu$ satisfy the assumptions stated in Section 4.2.
Recall that the operator $A \in L(H_0^1(\Omega), H^{-1}(\Omega))$ is required to be symmetric and coercive and $B \in L(L^2(\Omega)^N, H^{-1}(\Omega))$ in order to be consistent with the general setting in Section 4.2 (see Assumption 4.2.1). Therefore, we show on the basis of a practical application that these assumptions are not too strict. For this reason, consider the following partial differential equation (PDE) with Dirichlet boundary condition

$$\begin{aligned} \mathcal{A}y &= f \qquad \text{in } \Omega, \\ y &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{5.3}$$

Here, we have $f \in H^{-1}(\Omega)$ and $\mathcal{A}$ defines a second-order differential operator, which is given f.a.a. $\omega \in \Omega$ by

$$\mathcal{A}y(\omega) := -\sum_{i,j=1}^{d} D_{\omega_i}\left(a_{ij}(\omega)D_{\omega_j}y(\omega)\right).$$

Therein, the coefficient functions $a_{ij}$

- are elements of $L^\infty(\Omega)$;
- are symmetric, i.e. $a_{ij}(\omega) = a_{ji}(\omega)$ for all $i, j \in \{1, \dots, d\}$ and f.a.a. $\omega \in \Omega$;
- satisfy the condition of uniform ellipticity, i.e.

  $$\sum_{i,j=1}^{d} a_{ij}(\omega)\xi_i\xi_j \geq \gamma_0\|\xi\|^2$$

  for some $\gamma_0 > 0$ and all $\xi \in \mathbb{R}^d$.

Then, it follows that $\mathcal{A}$ is symmetric, bounded and coercive and the PDE (5.3) is called *elliptic* in this context (see [15, 88]).

**Example 5.1.1.** *If*

$$a_{ij}(\omega) \equiv \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{else,} \end{cases}$$

*then* $\mathcal{A}y$ *reduces to* $-\sum_{j=1}^{n} D^2_{\omega_j\omega_j}y =: -\Delta$*, where the operator* $\Delta$ *is known as Laplace operator. The resulting PDE is called Poisson's equation, which can be used to describe natural phenomenons in physics, for instance Newtonian gravity, electrostatics and fluid- or thermodynamics (see e.g. [15]).*

Notice that the right hand side of (5.3) is allowed to be discontinuous in this setting. Hence, we can not expect classical solutions, i.e. $y \in C^2(\Omega) \cap C(\bar{\Omega})$ and thus, are restricted to so-called weak solutions. A function $y \in H_0^1(\Omega)$ is said to be a *weak solution* of (5.3) if

$$a(y, \varphi) = F[\varphi] \tag{5.4}$$

is satisfied for all $\varphi \in H_0^1(\Omega)$ with

$$\begin{aligned} &a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R},\ a(y, \varphi) := \int_\Omega \sum_{i,j=1}^n a_{ij} D_{\omega_i} y D_{\omega_j} \varphi \, d\omega, \\ &F : H_0^1(\Omega) \to \mathbb{R},\ F[\varphi] := \langle f, \varphi \rangle_{H_0^1(\Omega)}. \end{aligned}$$

We refer to [35, 88] for a more detailed introduction of elliptic PDEs.
Since (5.4) has to be satisfied for all $\varphi \in H_0^1(\Omega)$, we obtain the equivalent condition

$$Ay = \tilde{b} \qquad \text{in } H^{-1}(\Omega), \tag{5.5}$$

where $A \in L(H_0^1(\Omega), H^{-1}(\Omega))$ and $\tilde{b} \in H^{-1}(\Omega)$ are defined by $\langle Ay, \cdot \rangle_{H_0^1(\Omega)} := a(y, \cdot)$ and $\tilde{b} := \langle f, \cdot \rangle_{H_0^1(\Omega)}$, respectively. In particular, it follows that $A$ is symmetric and coercive. Now, equation (5.5) denotes the KKT-condition of the unconstrained problem

$$\min \ \frac{1}{2} \langle Ay, y \rangle_Y - \langle \tilde{b}, y \rangle_Y$$

and hence, the lower level (5.2) can be interpreted as finding the weak solution in $Y_{\text{ad}}$ of the elliptic PDE (5.3) with respect to $\tilde{b} := Bu + b$.
The next section is dedicated to stationary Nash equilibrium concepts of $\Gamma_{\text{MLOCP}}$.

## 5.2. Stationary Nash Equilibria

Motivated by the previous sections, stationary concepts of $\Gamma_{\text{MLOCP}}$ are similarly derived by considering the corresponding EPCC, so in particular the application of Theorem 4.2.2 to our concrete setting. For this reason, we recall that leader $\nu$ solves

$$\begin{aligned} \min \ \ & \frac{1}{2} \|y - y_d^\nu\|_{L^2(\Omega)}^2 + \frac{\alpha_\nu}{2} \|x^\nu\|_{L^2(\Omega)}^2 \\ \text{w.r.t.} \ \ & (x^\nu, y, p) \in L^2(\Omega) \times H_0^1(\Omega) \times H^{-1}(\Omega), \\ \text{s.t.} \ \ & x^\nu \in Q_\nu, \\ & Ay - B[x^\nu, x^{-\nu}] + p = b, \\ & y \in K,\ p \in K^\circ,\ \langle p, y \rangle_Y = 0, \end{aligned} \tag{5.6}$$

where $K$ is given by $H_0^1(\Omega)_0^+$ (see Example 2.2.8). We immediately see that the corresponding dual cone can be written by

$$K^\circ = \left\{y^* \in H^{-1}(\Omega) \;\middle|\; \langle y^*, y\rangle_{H^{-1}(\Omega)} \leq 0 \quad \forall y \in H_0^1(\Omega)_0^+\right\} =: H^{-1}(\Omega)_-.$$

In order to derive stationarity conditions, we are interested in the explicit representation of the tangent and the normal cone with respect to $H_0^1(\Omega)_0^+$. For this reason, basic concepts about *capacity theory* are required, where we refer to [7] for a more detailed introduction.

**Definition 5.2.1** ([7, Definition 6.47])**.** Let $\Omega_0 \subseteq \Omega$ be an arbitrary subset and $\gamma \in \mathbb{R}$.

1. The *capacity* of $\Omega_0$ is defined by

$$\text{cap}(\Omega_0) := \inf\left\{\|\nabla y\|^2_{L^2(\Omega)} \;\middle|\; y \in H_0^1(\Omega) \text{ and } y \geq 1 \text{ a.e. in a neighborhood of } \Omega_0\right\}.$$

2. A function $y : \Omega \to \mathbb{R} \cup \{\pm\infty\}$ is called *quasi-continuous*, if for all $\epsilon > 0$ there exists an open set $G_\epsilon \subseteq \Omega$ with cap$(G_\epsilon) < \epsilon$ such that $y$ is continuous on $\Omega \setminus G_\epsilon$.

3. A set $O \subseteq \Omega$ is called *quasi-open*, if for all $\epsilon > 0$ there exists an open set $G_\epsilon \subseteq \Omega$ such that cap$(G_\epsilon) < \epsilon$ and $O \cup G_\epsilon$ is open. Similarly to the classical topology, the set $Q \subseteq \Omega$ is *quasi-closed*, if the complement $\Omega \setminus Q$ is quasi-open.

**Remark 5.2.2.** *In this context, the second requirement in* cap$(\Omega)$ *is meant in the sense that there exists a sequence* $\{y^k\} \subseteq H_0^1(\Omega)$ *with* $y^k \to y$ *in* $H_0^1(\Omega)$ *such that* $y^k \geq 1$ *almost everywhere in a neighborhood of* $\Omega_0$*.*

Since the infimum over an empty set is defined by $\infty$, we conclude that cap$(\Omega) = \infty$. Moreover, it is known (see [23, Theorem 3.2]) that cap$(\emptyset) = 0$ and cap$(\Omega_1) \leq$ cap$(\Omega_2)$ for arbitrary sets $\Omega_1, \Omega_2 \subseteq \Omega$ with $\Omega_1 \subseteq \Omega_2$.
Similar to the 'almost everywhere'-notation in Lebesgue spaces, we say that a property $\mathcal{P}$ holds *quasi everywhere* (q.e.) on $\Omega_0 \subseteq \Omega$, if $\mathcal{P}$ is only violated on a set with zero capacity, i.e.

$$\text{cap}(\{x \in \Omega_0 \;|\; \mathcal{P} \text{ does not hold at } x\}) = 0.$$

**Example 5.2.3.** *Let* $y : \Omega \to \mathbb{R}$ *be a quasi-continuous function. Then the set*

$$\{y > 0\} := \{x \in \Omega \;|\; y(x) > 0\}$$

*is quasi-open and it holds*

$$y > 0 \quad \text{q.e. on } \Omega \qquad \Longleftrightarrow \qquad \text{cap}(\{y \leq 0\} \cap \Omega) = 0.$$

Furthermore, it is known (see [7, Lemma 6.51]) that every $y \in H_0^1(\Omega)$ possesses a quasi-continuous representative $\tilde{y}$, which is uniquely determined up to sets of zero capacity. As a consequence and similar to equivalence classes in Lebesgue spaces, we always refer to $\tilde{y}$ when speaking about a function $y \in H_0^1(\Omega)$.
The next result shows the relation between 'almost everywhere' and 'quasi everywhere'.

**Lemma 5.2.4** ([23, Lemma 3.4]). *Let $y : \Omega \to \mathbb{R}$ be quasi-continuous and $\Omega_0 \subseteq \Omega$ quasi-open. Then we have*

$$y \geq 0 \quad \text{q.e. on } \Omega_0 \qquad \Longleftrightarrow \qquad y \geq 0 \quad \text{a.e. on } \Omega_0.$$

Now, a major advantage of capacity theory is that we are able to describe functionals $\mu \in K^+ = H^{-1}(\Omega)_+$ as measures. Therefore, we quote [73, Lemma 2.5], which can be seen as a summary of results from [7, Section 6] and [23, Section 3].

**Lemma 5.2.5** ([73, Lemma 2.5]). *Let $\mu \in H^{-1}(\Omega)_+$ be given. Then, we have the following:*

1. *The functional $\mu$ can be identified with a regular Borel measure*[16] *on $\Omega$ that is finite on compact sets and that possesses the following property: For every Borel set*[17] *$\mathcal{B} \subseteq \Omega$ with* cap$(\mathcal{B}) = 0$*, it holds $\mu(\mathcal{B}) = 0$.*
2. *Every function $y \in H_0^1(\Omega)$ is $\mu$-integrable and it holds*
$$\langle \mu, y \rangle_{H_0^1(\Omega)} = \int_\Omega y \, d\mu.$$
3. *There exists a quasi-closed set* f-supp$(\mu) \subseteq \Omega$ *with the property that for all $y \in K$, it holds*
$$\langle \mu, y \rangle_{H_0^1(\Omega)} = 0 \qquad \Longleftrightarrow \qquad y = 0 \quad \text{q.e. on f-supp}(\mu).$$
*The set* f-supp$(\mu)$ *is uniquely defined up to a set of zero capacity and is called the fine support of $\mu$.*

Notice that Lemma 5.2.5 considers functionals $\mu$ in $H^{-1}(\Omega)_+$. However, if $\mu$ is in $H^{-1}(\Omega)_-$, as in our case, we set f-supp$(\mu)$ := f-supp$(-\mu)$. For more information we refer to [95, Appendix A].
After this short introduction to capacity theory, we are now able to give an explicit representation of the tangent, normal and the critical cone of $K$, respectively.

[16] A measure defined on a $\sigma$-algebra (see Section 2.1.2.) of Borel sets.

[17] A set is called *Borel set*, if it is generated by a countable number of operations (e.g. intersection, union, complement).

**Lemma 5.2.6.** *Let $\bar{z} \in K = H_0^1(\Omega)_0^+$ be given. Then we have*

$$\mathcal{T}_K(\bar{z}) = \big\{z \in H_0^1(\Omega) \;\big|\; z \geq 0 \text{ q.e. on } \{\bar{z} = 0\}\big\},$$
$$\mathcal{N}_K(\bar{z}) = \big\{z^* \in H^{-1}(\Omega)_- \;\big|\; \bar{z} = 0 \text{ q.e. on f-supp}(z^*)\big\}.$$

*Moreover, let $z^* \in \mathcal{N}_K(\bar{z})$. Then the critical cone is given by*

$$\mathcal{K}_K(\bar{z}, z^*) = \big\{z \in H_0^1(\Omega) \;\big|\; z \geq 0 \text{ q.e. on } \{\bar{z} = 0\} \text{ and } z = 0 \text{ q.e. on f-supp}(z^*)\big\}.$$

*Proof.* In order to show the identity of the tangent cone, we refer to the proof of [7, Theorem 6.57].
Since $K$ is convex, the identity $N_K(\bar{z}) = K^\circ \cap \{\bar{z}\}^\perp$ holds, which implies that

$$N_K(\bar{z}) = \big\{z^* \in H^{-1}(\Omega)_- \;\big|\; \langle z^*, \bar{z}\rangle_{H_0^1(\Omega)} = 0\big\}$$

and the result follows by Lemma 5.2.5 (3).
Finally, we obtain the representation of the critical cone via

$$\begin{aligned}
\mathcal{K}_K(\bar{z}, z^*) &= T_K(\bar{z}) \cap \{z^*\}^\perp \\
&= \big\{z \in H_0^1(\Omega) \;\big|\; z \geq 0 \text{ q.e. on } \{\bar{z} = 0\}\big\} \cap \big\{z \in H_0^1(\Omega) \;\big|\; \langle z^*, z\rangle_{H_0^1(\Omega)} = 0\big\} \\
&= \big\{z \in H_0^1(\Omega) \;\big|\; z \geq 0 \text{ q.e. on } \{\bar{z} = 0\} \text{ and } z = 0 \text{ q.e. on f-supp}(z^*)\big\}.
\end{aligned}$$

□

Now, we are able to formulate the stationarity concepts in $H_0^1(\Omega)$. Similar to the finite dimensional case but in contrast to the general setting in Chapter 3, we introduce the *active, strongly active, inactive* and *biactive set* of the complementarity conditions

$$y \in H_0^1(\Omega)_0^+, \quad p \in H^{-1}(\Omega)_-, \quad \langle p, y\rangle_{H_0^1(\Omega)} = 0$$

by

$$\mathcal{Q}_A \coloneqq \{y = 0\}, \quad \mathcal{Q}_A^+ \coloneqq \text{f-supp}(p), \quad \mathcal{Q}_I \coloneqq \{y > 0\}, \quad \mathcal{Q}_A^0 \coloneqq \mathcal{Q}_A \setminus \mathcal{Q}_A^+.$$

We emphasize that all defined sets above are understood in the quasi-everywhere sense.

**Remark 5.2.7.** *As a comparison, we recall the situation in finite dimensions. Here, the index sets of complementarity conditions in $\mathbb{R}^m$ of the form $0 \leq y_i \perp p_i \leq 0$ for all $i = 1, \dots, m$ are given by*

$$\mathcal{Q}_A = \big\{i \in \{1, \dots, m\} \;\big|\; y_i = 0\big\}, \quad \mathcal{Q}_A^+ = \big\{i \in \mathcal{Q}_A \;\big|\; p_i < 0\big\},$$

$$\mathcal{Q}_I = \big\{i \in \{1, \dots, m\} \;\big|\; y_i > 0\big\}, \quad \mathcal{Q}_A^0 = \big\{i \in \mathcal{Q}_A \;\big|\; p_i = 0\big\}.$$

The subsequent result can be used to transfer W- and S-stationary Nash equilibria in abstract Banach spaces to their counterparts in $H_0^1(\Omega)$.

**Lemma 5.2.8.** *Let $K = H_0^1(\Omega)_0^+$. Then we have*

$$\begin{aligned} \mathrm{cl}(K - K \cap \{p\}^\perp) \cap \{p\}^\perp &= \big\{v \in H_0^1(\Omega) \;\big|\; v = 0 \text{ q.e. on } \mathcal{Q}_A^+\big\}, \\ \mathrm{cl}(K^\circ - K^\circ \cap \{y\}^\perp) \cap \{y\}^\perp &= \big\{v \in H_0^1(\Omega) \;\big|\; v = 0 \text{ q.e. on } \mathcal{Q}_A\big\}^\circ \end{aligned} \tag{5.7}$$

*and*

$$\begin{aligned} \mathcal{K}_K(y,p) &= \big\{v \in H_0^1(\Omega) \;\big|\; v \geq 0 \text{ q.e. on } \mathcal{Q}_A \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+\big\}, \\ \mathcal{K}_{K^\circ}(p,y) &= \big\{v \in H_0^1(\Omega) \;\big|\; v \geq 0 \text{ q.e. on } \mathcal{Q}_A \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+\big\}^\circ. \end{aligned} \tag{5.8}$$

*Proof.* For a proof of (5.7), we refer to [23, Theorem 5.1].
For the second part, recall that the representation of $\mathcal{K}_K(y,p)$ was shown in Lemma 5.2.6. Since $H_0^1(\Omega)_0^+$ is polyhedric (see Example 2.2.8), we can apply Lemma 2.2.4 and observe that $\mathcal{K}_{K^\circ}(p,y) = \mathcal{K}_K(y,p)^\circ$, which implies the assertion. Thus, (5.8) is satisfied. ☐

The following definition is motivated by the previous lemma, Theorem 4.2.2 and [96].

**Definition 5.2.9.** A feasible point $(\bar{x}, \bar{y}) \in L^2(\Omega)^N \times H_0^1(\Omega)$ of $\Gamma_{\text{MLOCP}}$ is a

- *W-stationary Nash equilibrium*, if there exist multipliers

$$(p, \xi, \mu) \in H^{-1}(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$$

  such that for all $\nu = 1, \dots, N$ it holds

$$\bar{x}^\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu + (\alpha_\nu^{-1} B_\nu^* \xi^\nu - b_\nu)_+ - (a_\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu)_+ = 0, \tag{5.9a}$$

$$A\xi^\nu - \bar{y} + y_d^\nu - \mu^\nu = 0, \tag{5.9b}$$

$$A\bar{y} + p - B\bar{x} - b = 0, \tag{5.9c}$$

$$p \leq 0, \; \bar{y} \geq 0, \; \langle p, \bar{y} \rangle = 0 \tag{5.9d}$$

  and

$$\mu^\nu \in \big\{v \in H_0^1(\Omega) \;\big|\; v = 0 \text{ q.e. on } \mathcal{Q}_A\big\}^\circ, \tag{5.10a}$$

$$-\xi^\nu \in \big\{v \in H_0^1(\Omega) \;\big|\; v = 0 \text{ q.e. on } \mathcal{Q}_A^+\big\}; \tag{5.10b}$$

- *C-stationary Nash equilibrium*, if there exist multipliers

$$(p, \xi, \mu) \in H^{-1}(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$$

  such that for all $\nu = 1, \dots, N$ (5.9), (5.10) hold as well as

$$\langle \mu^\nu, \xi^\nu \rangle_{H_0^1(\Omega)} \leq 0; \tag{5.11}$$

- *M-stationary Nash equilibrium*, if there exist multipliers

$$(p,\xi,\mu) \in H^{-1}(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$$

  and a disjoint decomposition of the biactive set $\mathcal{Q}_A^0 = \hat{\mathcal{Q}}_I \cup \hat{\mathcal{Q}}_A^0 \cup \hat{\mathcal{Q}}_A^+$ such that for all $\nu = 1,\dots,N$ (5.9) holds as well as

$$\mu^\nu \in \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \hat{\mathcal{Q}}_A^0 \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+ \cup \hat{\mathcal{Q}}_A^+\}^\circ, \quad \text{(5.12a)}$$

$$-\xi^\nu \in \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \hat{\mathcal{Q}}_A^0 \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+ \cup \hat{\mathcal{Q}}_A^+\}; \quad \text{(5.12b)}$$

- *S-stationary Nash equilibrium*, if there exist multipliers

$$(p,\xi,\mu) \in H^{-1}(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$$

  such that for all $\nu = 1,\dots,N$ (5.9) holds as well as

$$\mu^\nu \in \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \mathcal{Q}_A \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+\}^\circ, \quad \text{(5.13a)}$$

$$-\xi^\nu \in \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \mathcal{Q}_A \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+\}. \quad \text{(5.13b)}$$

It is clear that W- and S-stationary Nash equilibria are a direct result of Lemma 5.2.8 and Theorem 4.2.2 applied to $\Gamma_{\text{MLOCP}}$. In particular, observe that due to the latter theorem, there exist multipliers $(\kappa,\xi) \in L^2(\Omega)^N \times H_0^1(\Omega)^N$ such that for all $\nu = 1,\dots,N$, $\bar{x}^\nu$ satisfies

$$\alpha_\nu \bar{x}^\nu + \kappa^\nu - B_\nu^* \xi^\nu = 0, \quad \kappa^\nu \in \mathcal{N}_{Q_\nu}(\bar{x}^\nu).$$

Hence, $\bar{x}^\nu$ solves

$$(\alpha_\nu \bar{x}^\nu - B_\nu^* \xi^\nu, x^\nu - \bar{x}^\nu) \geq 0$$

for all $x^\nu \in Q_\nu$. Moreover, it is known (see e.g. [35, Lemma 1.11]) that the latter is equivalent to

$$\bar{x}^\nu - \Pi_{Q_\nu}(\bar{x}^\nu - \gamma^\nu(\alpha_\nu \bar{x}^\nu - B_\nu^* \xi^\nu)) = 0$$

for all $\gamma^\nu > 0$, where $\Pi_{Q_\nu}$ denotes the projection onto $Q_\nu$. By choosing $\gamma^\nu := \alpha_\nu^{-1}$ and using the structure of $Q_\nu$, we conclude that (5.9a) is satisfied. Below, we comment on some aspects concerning C- and M-stationary Nash equilibria.

1. In contrast to abstract Banach spaces, we have introduced an additional stationarity concept, named after *Clarke*. Here, the derivation is motivated by the theory in finite spaces (see e.g. [82, Definition 5.13]) and uses the variant given for instance in [29, 86]. Further (stronger) forms were considered in [79].

2. Similar to C-stationarity, we use the M-stationarity concept motivated by the finite dimensional case and introduced in [96]. For this reason, recall (see e.g. [82, Definition 5.13]) that a point $x \in \mathbb{R}^n$ is M-stationary, if it is weakly stationary and the multipliers $\mu, \lambda \in \mathbb{R}^m$ that are associated to the complementarity conditions (see Remark 5.2.7) satisfy either $\mu_j, \lambda_j \geq 0$ or $\mu_j \cdot \lambda_j = 0$ for all $j \in \mathcal{Q}_A^0$. Now, this condition can also be written as

$$\begin{aligned} \mu_j &= 0 && \forall j \in Q_I \cup \hat{Q}_I, \\ \lambda_j &= 0 && \forall j \in Q_A^+ \cup \hat{Q}_A^+, \\ \mu_j, \lambda_j &\geq 0 && \forall j \in \hat{Q}_A^0, \end{aligned}$$

   which is more suitable for the infinite dimensional approach. Therein, $\hat{Q}_I$, $\hat{Q}_A^+$ and $\hat{Q}_A^0$ describe a disjoint decomposition of $Q_A^0$, i.e. $Q_A^0 = \hat{Q}_I \cup \hat{Q}_A^+ \cup \hat{Q}_A^0$.
   Moreover, observe that the concept of M-stationarity introduced above and its counterpart in abstract Banach spaces are not equivalent in general. In particular, it was shown in [23, Theorem 5.3 and Theorem 5.4] that both concepts coincide, if $\Omega \subseteq \mathbb{R}$, i.e. $d = 1$.

We have already mentioned the amount of contributions referring to the optimal control of the obstacle problem in the beginning of Chapter 5. For this reason, we recall that the majority (see e.g. [29, 30, 31, 48, 79, 86]) considered more regular data, i.e. $Bx + b \in L^2(\Omega)$ instead of $Bx + b \in H^{-1}(\Omega)$. As a consequence of this lifting, we get that $y$ and $p$ are more regular themselves (see [46]), i.e. $y \in H^2(\Omega) \cap H_0^1(\Omega)$ and $p \in L^2(\Omega)$. As a consequence, it was shown in [96, Lemma 4.6] that (5.10a) implies

$$\langle \mu^\nu, \varphi \rangle_{H_0^1(\Omega)} = 0 \tag{5.14}$$

for all $\varphi \in \{v \in H_0^1(\Omega) \mid v = 0 \text{ a.e. in } \mathcal{Q}_A\}$ and (5.10b) is equivalent to

$$\xi^\nu = 0 \text{ a.e. in } \tilde{\mathcal{Q}}_A^+, \tag{5.15}$$

with $\tilde{\mathcal{Q}}_A^+ \coloneqq \{p < 0\}$. Notice that the latter condition is defined up to sets of Lebesgue measure zero, i.e. in the almost-everywhere sense. Furthermore, observe that the concept of W- and C-stationarity coincide with the corresponding definitions given in the references above. However, we do not obtain a similar relation for an S-stationary Nash equilibrium. For this reason, observe that we have

$$\begin{aligned} -\xi^\nu \in \mathcal{M}_S \coloneqq &\big\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \mathcal{Q}_A \text{ and } v = 0 \text{ q.e. on } \mathcal{Q}_A^+\big\} \\ = &\big\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \mathcal{Q}_A\big\} \cap \big\{v \in H_0^1(\Omega) \mid v = 0 \text{ q.e. on } \mathcal{Q}_A^+\big\}. \end{aligned}$$

By the discussion above, the second set within the intersection is equivalent to (5.15). Since

$$v \geq 0 \text{ q.e. on } \mathcal{Q}_A \quad \Longleftrightarrow \quad \operatorname{cap}\left(\{v<0\} \cap \mathcal{Q}_A\right)=0$$

and the measure of $\{v<0\} \cap \mathcal{Q}_A$, denoted by $\operatorname{meas}(\{v<0\} \cap \mathcal{Q}_A)$, satisfies

$$\operatorname{meas}(\{v<0\} \cap \mathcal{Q}_A) \leq c \operatorname{cap}\left(\{v<0\} \cap \mathcal{Q}_A\right)$$

for positive constants $c$ (see e.g. [23, Lemma 3.2b)]), we obtain $\operatorname{meas}\left(\{v<0\} \cap \mathcal{Q}_A\right)=0$. Hence, we only get the inclusion

$$\left\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \mathcal{Q}_A\right\} \subseteq \left\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e. on } \mathcal{Q}_A\right\},$$

since $\{v<0\} \cap \mathcal{Q}_A$ is quasi-closed, and it holds

$$\mathcal{M}_S \subseteq \left\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e. in } \mathcal{Q}_A \text{ and } v=0 \text{ a.e. in } \tilde{\mathcal{Q}}_A^{+}\right\}. \tag{5.16}$$

According to [30, Definition 4.1.], a feasible point is S-stationary if and only if it holds (5.9a)-(5.9d) and

$$\mu^\nu \in \left\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e. in } \mathcal{Q}_A \text{ and } v=0 \text{ a.e. in } \tilde{\mathcal{Q}}_A^{+}\right\}^{\circ}, \tag{5.17a}$$

$$-\xi^\nu \in \left\{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e. in } \mathcal{Q}_A \text{ and } v=0 \text{ a.e. in } \tilde{\mathcal{Q}}_A^{+}\right\} \tag{5.17b}$$

are satisfied. However, with the inclusion (5.16) and the relation $\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{B}^\circ \subseteq \mathcal{A}^\circ$ for arbitrary sets $\mathcal{A}, \mathcal{B} \in H_0^1(\Omega)$, we observe on the one hand that condition (5.13a) is stronger than (5.17a) while on the other hand (5.13b) is weaker than (5.17b). See also [95] for further discussions.

The upcoming section applies the penalization and regularization techniques that were introduced and analyzed in Chapter 4 to the corresponding EPECs of $\Gamma_{\text{MLOCP}}$.

## 5.3. Convergence to a Stationary Nash Equilibrium

After introducing stationarity concepts in $H_0^1(\Omega)$ for $\Gamma_{\text{MLOCP}}$, we now move on to analyze the corresponding EPEC/EPCC derived in Section 4.2. In this context, we recall both auxiliary equilibrium problems. By using the variational inequality formulation as the first-order optimality condition of (5.2), we end up with the following Nash equilibrium problem, denoted by $\Gamma_{\text{MLOCP}}^{1,k}$, where leader $\nu$ solves

$$\begin{aligned}
\min\ & \frac{1}{2}\|y-y_d^\nu\|_{L^2(\Omega)}^2+\frac{\alpha_\nu}{2}\|x^\nu\|_{L^2(\Omega)}^2 \\
\text{w.r.t. }\ & (x^\nu, y) \in L^2(\Omega) \times H_0^1(\Omega), \\
\text{s.t. }\ & x^\nu \in Q_\nu,\ Ay-\frac{1}{\gamma_k}\max{}_K^{\epsilon_k}(0,-y)-B[x^\nu, x^{-\nu}]=b.
\end{aligned} \tag{5.18}$$

In particular, we use one of the penalty functions that was presented in Example 4.2.12. By applying the corresponding KKT-system as first-order optimality condition of (5.2), we consider the subsequent Nash equilibrium problem, denoted by $\Gamma^{2,k}_{\text{MLOCP}}$, where leader $\nu$ solves

$$\begin{aligned} \min\ & \frac{1}{2}\|y - y_d^\nu\|^2_{L^2(\Omega)} + \frac{\alpha_\nu}{2}\|x^\nu\|^2_{L^2(\Omega)} - \frac{1}{\gamma_k}\langle p, y\rangle_{H_0^1(\Omega)}, \\ \text{w.r.t.}\ & (x^\nu, y, p) \in L^2(\Omega) \times H_0^1(\Omega) \times L^2(\Omega) \\ \text{s.t.}\ & x^\nu \in Q_\nu,\ y \geq 0,\ p \leq 0, \\ & Ay + p - B[x^\nu, x^{-\nu}] = b. \end{aligned} \tag{5.19}$$

Notice that we assume more regular data, i.e. $Bx + b \in L^2(\Omega)$. Furthermore, we write $\max^{\epsilon_k}$ instead of $\max^{\epsilon_k}_K$, since it is clear from context.
A comparison with Section 4.2 reveals that $\Gamma^{1,k}_{\text{MLOCP}}$ satisfies all the assumptions, which are postulated in Subsection 4.2.1 (see also Example 4.2.18). For the second problem, observe that we need the boundedness of $|\langle p^k, y^k\rangle|$ in order to apply Lemma 4.2.19 in the proof of Theorem 4.2.20. However, notice that the latter requirement is satisfied, since we obtain during the corresponding proof that $\langle p^k, y^k\rangle \to 0$, if $k \to \infty$.
Although, the verification of the approximation of $\Gamma^{2,k}_{\text{MLOCP}}$ to $\Gamma_{\text{MLOCP}}$ is not as comfortable as for $\Gamma^{1,k}_{\text{MLOCP}}$, we can still use the former auxiliary problem to show the convergence towards an certain type of stationary Nash equilibrium. Therefore, we proceed by looking for suitable sequences and start with the auxiliary equilibrium problem $\Gamma^{2,k}_{\text{MLOCP}}$.

**Theorem 5.3.1.** *Let $(x^k, y^k, p^k) \in L^2(\Omega)^N \times H_0^1(\Omega) \times L^2(\Omega)$ be a Nash equilibrium of $\Gamma^{2,k}_{\text{MLOCP}}$. Then $(x^k, y^k, p^k)$ is a stationary Nash equilibrium, i.e. there exist multipliers*

$$(\xi^k, \tilde{\mu}^k, \tilde{\lambda}^k) \in H_0^1(\Omega)^N \times H^{-1}(\Omega)^N \times L^2(\Omega)^N$$

*such that the system*

$$x^{k,\nu} - \alpha_\nu^{-1} B_\nu^* \xi^{k,\nu} + (\alpha_\nu^{-1} B_\nu^* \xi^{k,\nu} - b_\nu)_+ - (a_\nu - \alpha_\nu^{-1} B_\nu^* \xi^{k,\nu})_+ = 0, \tag{5.20a}$$

$$A\xi^{k,\nu} + \tilde{\mu}^{k,\nu} - \frac{1}{\gamma_k} p^k - y_d^\nu + y^k = 0, \tag{5.20b}$$

$$Ay^k + p^k - Bx^k - b = 0, \tag{5.20c}$$

$$y^k \geq 0,\ \tilde{\mu}^{k,\nu} \leq 0,\ \langle \tilde{\mu}^{k,\nu}, y^k\rangle_{H_0^1(\Omega)} = 0, \tag{5.20d}$$

$$p^k \leq 0,\ \tilde{\lambda}^{k,\nu} \geq 0,\ (p^k, \tilde{\lambda}^{k,\nu})_{L^2(\Omega)} = 0, \tag{5.20e}$$

$$\tilde{\lambda}^{k,\nu} - \frac{1}{\gamma_k} y^k + \xi^{k,\nu} = 0, \tag{5.20f}$$

*is satisfied for all* $\nu = 1, \dots, N$.

*Proof.* Let $(x^k, y^k, p^k)$ be a Nash equilibrium of $\Gamma^{2,k}_{\text{MLOCP}}$ and observe that the equation (see [31, Proposition 2.5])

$$D_{z^\nu} G(z^{k,\nu}, x^{k,-\nu}) \begin{bmatrix} \mathcal{R}_{X^\nu_{\text{ad}}}(x^{k,\nu}) \\ H^1_0(\Omega) \\ L^2(\Omega) \end{bmatrix} - \mathcal{R}_{\tilde{K}}(G(z^{k,\nu}, x^{k,-\nu})) = W \tag{5.21}$$

holds for all $\nu = 1, \dots, N$, where we set

- $z^{k,\nu} := (x^{k,\nu}, y^k, p^k)$,
- $V := L^2(\Omega) \times H^1_0(\Omega) \times L^2(\Omega)$, $W := H^{-1}(\Omega) \times H^1_0(\Omega) \times L^2(\Omega)$,
- $G : V \times L^2(\Omega)^{N-1} \to W,\ G(x, y, p) := (Ay + p - Bx - b, y, -p)$,
- $\tilde{K} := \{0\} \times H^1_0(\Omega)^+ \times L^2(\Omega)^+ \subseteq W$.

Hence, KRZCQ is satisfied for every leader's problem and the result follows by applying Theorem 2.3.9 for all $\nu = 1, \dots, N$. □

In particular, the proof showed that any Nash equilibrium of $\Gamma^{2,k}_{\text{MLOCP}}$ satisfies KRZCQ for each player's problem. Similar, but without the requirement that further constraint qualifications have to be valid, the analogous result for $\Gamma^{1,k}_{\text{MLOCP}}$ is straightforward.

**Theorem 5.3.2.** *Let* $(x^k, y^k) \in L^2(\Omega)^N \times H^1_0(\Omega)$ *be a Nash equilibrium of* $\Gamma^{1,k}_{\text{MLOCP}}$. *Then it is a stationary Nash equilibrium, i.e. there exists* $\xi^k \in H^1_0(\Omega)^N$ *such that the system*

$$x^{k,\nu} - \alpha_\nu^{-1} B_\nu^* \xi^{k,\nu} + (\alpha_\nu^{-1} B_\nu^* \xi^{k,\nu} - b_\nu)_+ - (a_\nu - \alpha_\nu^{-1} B_\nu^* \xi^{k,\nu})_+ = 0, \tag{5.22a}$$

$$A\xi^{k,\nu} + \frac{1}{\gamma_k} D_y \max^{\epsilon_k}(0, -y^k)\xi^{k,\nu} - y^\nu_d + y^k = 0, \tag{5.22b}$$

$$Ay^k - \frac{1}{\gamma_k} \max^{\epsilon_k}(0, -y^k) - Bx^k - b = 0 \tag{5.22c}$$

*is satisfied for all* $\nu = 1, \dots, N$.

During the remainder of this section, we proceed by showing that we find sequences of stationary Nash equilibria of $\Gamma^{2,k}_{\text{MLOCP}}$ and $\Gamma^{1,k}_{\text{MLOCP}}$, respectively, which converge to a stationary-type Nash equilibrium. Similar to the single-player scenario, i.e. $N = 1$ (see e.g. [29] and [31]), and without further assumptions, we need an additional stationarity concept that is weaker than C-stationary Nash equilibria. Therefore, the following definition is motivated by [30, Definition 4.2] for $\epsilon$-almost C-stationary points of MPCCs, where we recall that the strongly active set is given by $\tilde{Q}_A^+ = \{p < 0\}$.

**Definition 5.3.3.** A point $(\bar{x}, \bar{y}) \in L^2(\Omega)^N \times H_0^1(\Omega)$ is called *$\epsilon$-almost C-stationary Nash equilibrium* of $\Gamma_{\text{MLOCP}}$, if there exist multipliers $(p, \xi, \mu) \in L^2(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$ such that for all $\nu = 1, \dots, N$ the system

$$\bar{x}^\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu + (\alpha_\nu^{-1} B_\nu^* \xi^\nu - b_\nu)_+ - (a_\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu)_+ = 0, \tag{5.23a}$$
$$A\xi^\nu - \mu^\nu - y_d^\nu + \bar{y} = 0, \tag{5.23b}$$
$$A\bar{y} + p - B\bar{x} - b = 0, \tag{5.23c}$$
$$p \le 0,\ \bar{y} \ge 0,\ \langle p, \bar{y} \rangle_{H_0^1(\Omega)} = 0, \tag{5.23d}$$
$$\xi^\nu = 0 \quad \text{a.e. in } \tilde{Q}_A^+, \tag{5.23e}$$
$$\langle \mu^\nu, \xi^\nu \rangle_{H_0^1(\Omega)} \le 0, \tag{5.23f}$$
$$\langle \mu^\nu, \bar{y} \rangle_{H_0^1(\Omega)} = 0 \tag{5.23g}$$

is satisfied and further for every $\epsilon > 0$, there exists $\mathcal{U}_\epsilon \subseteq \mathcal{Q}_I$ with measure $\operatorname{meas}(\mathcal{Q}_I \backslash \mathcal{U}_\epsilon) \le \epsilon$ such that it holds

$$\langle \mu^\nu, \varphi \rangle_{H_0^1(\Omega)} = 0 \quad \forall \varphi \in \left\{ v \in H_0^1(\Omega) \,\middle|\, v = 0 \quad \text{a.e. in } \Omega \setminus \mathcal{U}_\epsilon \right\}. \tag{5.24}$$

Comparing Definition 5.3.3 and Definition 5.2.9 reveals that a C-stationary Nash equilibrium differs only in condition (5.14), i.e. it holds

$$\langle \mu^\nu, \varphi \rangle_{H_0^1(\Omega)} = 0 \quad \forall \varphi \in \left\{ v \in H_0^1(\Omega) \,\middle|\, v = 0 \text{ a.e. in } \{\bar{y} = 0\} \right\}.$$

However, observe that the condition above implies (5.23g) and (5.24), respectively. Hence, if $(\bar{x}, \bar{y})$ is a C-stationary Nash equilibrium, it is an $\epsilon$-almost C-stationary Nash equilibrium. The next result that can be seen as the extension to equilibrium problems of the results presented in [31, Section 2.2] shows the convergence of a sequence generated via $\Gamma_{\text{MLOCP}}^{2,k}$ at least to an $\epsilon$-almost C-stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}$.

**Theorem 5.3.4.** *Let $\{\gamma_k\} \subseteq \mathbb{R}^+$ and $\{(x^k, y^k, p^k)\} \subseteq L^2(\Omega)^N \times H_0^1(\Omega) \times L^2(\Omega)$ be two sequences such that $\gamma_k \to 0$ as $k \to \infty$ is satisfied and $(x^k, y^k, p^k)$ is a stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}^{2,k}$. Additionally, assume that there is $c > 0$ such that $\langle p^k, y^k \rangle_{H_0^1(\Omega)} \le c$ holds for all $k \in \mathbb{N}$. Then there exist limit points*

$$(\bar{x}, \bar{y}, \bar{p}, \bar{\xi}, \bar{\mu}) \in L^2(\Omega)^N \times H_0^1(\Omega) \times H^{-1}(\Omega) \times L^2(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$$

*such that*

$$\begin{aligned}(x^{k,\nu}, y^k, p^k) &\to (\bar{x}^\nu, \bar{y}, \bar{p}) &&\in L^2(\Omega) \times H_0^1(\Omega) \times H^{-1}(\Omega),\\ (\xi^{k,\nu}, \tilde{\mu}^{k,\nu} + \tfrac{1}{\gamma_k} p^k =: \mu^{k,\nu}) &\rightharpoonup (\bar{\xi}^\nu, \bar{\mu}^\nu) &&\in H_0^1(\Omega) \times H^{-1}(\Omega)\end{aligned}$$

*for all $\nu = 1, \dots, N$ on a subsequence and $(\bar{x}, \bar{y})$ is an $\epsilon$-almost C-stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}$.*
*If $Q_I$ is a Lipschitz domain and the conditions*

$$\langle \tilde{\mu}^{k,\nu}, \bar{y} \rangle_{H_0^1(\Omega)} \to 0, \qquad \langle p^k, \bar{y} \rangle_{H_0^1(\Omega)} \to 0 \tag{5.25}$$

*are satisfied for all $\nu = 1, \dots, N$, then $(\bar{x}, \bar{y})$ is a C-stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}$.*

*Proof.* Let $\nu \in \{1, \dots, N\}$ be arbitrary but fixed and define the multiplier

$$\lambda^{k,\nu} := \tilde{\lambda}^{k,\nu} - \frac{1}{\gamma_k} y^k.$$

Due to the boundedness of $\langle p^k, y^k \rangle$ and $Q_\nu$, it was shown in [31, Proposition 2.7] that there exists a constant $c > 0$, independent of $\gamma_k$, such that

$$\|\lambda^{k,\nu}\|_{H_0^1(\Omega)} \le c, \tag{5.26a}$$
$$\|\mu^{k,\nu}\|_{H^{-1}(\Omega)} \le c, \tag{5.26b}$$
$$\langle \mu^{k,\nu}, \lambda^{k,\nu} \rangle_{H_0^1(\Omega)} \ge 0, \tag{5.26c}$$
$$0 \le -\frac{1}{\gamma_k^2} \langle p^k, y^k \rangle_{H_0^1(\Omega)} \le c. \tag{5.26d}$$

In particular, we find $(\bar{\mu}^\nu, \bar{\lambda}^\nu) \in H_0^1(\Omega) \times H^{-1}(\Omega)$ such that $(\mu^{k,\nu}, \lambda^{k,\nu}) \rightharpoonup (\bar{\mu}^\nu, \bar{\lambda}^\nu)$ in $H_0^1(\Omega) \times H^{-1}(\Omega)$. Moreover, it follows by the weak sequential compactness of $Q$ that there exists $\bar{x} \in Q$ such that $x^{k,\nu} \rightharpoonup \bar{x}^\nu$ in $L^2(\Omega)$ and by the compact embedding of $L^2(\Omega)$ in $H^{-1}(\Omega)$ that $x^{k,\nu} \to \bar{x}^\nu$ in $H^{-1}(\Omega)$. Since (5.26d) implies that $\langle p^k, y^k \rangle \to 0$, we can now apply Lemma 4.2.19 and obtain

$$(y^k, p^k) \to (\bar{y}, \bar{p}) \text{ in } H_0^1(\Omega) \times H^{-1}(\Omega).$$

Furthermore, $(\bar{x}, \bar{y}, \bar{p})$ satisfies the conditions (5.23c) and (5.23d).
Next, we estimate with the coercivity of $A$ that

$$c\|\xi^{k,\nu}\|_{H_0^1(\Omega)}^2 \le \langle A\xi^{k,\nu}, \xi^{k,\nu} \rangle_{H_0^1(\Omega)} = \langle \mu^{k,\nu}, \xi^{k,\nu} \rangle_{H_0^1(\Omega)} - \langle y^k - y_d^\nu, \xi^{k,\nu} \rangle_{H_0^1(\Omega)}$$

and deduce by the boundedness of $\{\mu^{k,\nu}\}$ (see (5.26b)) and $\{y^k\}$ that $\{\xi^{k,\nu}\}$ is bounded in $H_0^1(\Omega)$. Hence, we find $\bar{\xi}^\nu \in H_0^1(\Omega)$ such that $\xi^{k,\nu} \rightharpoonup \bar{\xi}^\nu$ in $H_0^1(\Omega)$ and $\xi^{k,\nu} \to \bar{\xi}^\nu$ in $L^2(\Omega)$ on a subsequence we denote the same. By the weak continuity of $A$, it follows that

$$0 = A\xi^{k,\nu} - \mu^{k,\nu} - y_d^\nu + y^k \rightharpoonup A\bar{\xi}^\nu - \bar{\mu}^\nu - y_d^\nu + \bar{y}$$

and thus, the adjoint equation (5.23b) holds by the uniqueness of the weak limit. Similarly, we obtain that

$$0 = \lambda^{k,\nu} + \xi^{k,\nu} \rightharpoonup \bar{\lambda}^\nu + \bar{\xi}^\nu \quad \text{ in } H_0^1(\Omega),$$

which implies $\bar{\lambda}^\nu = -\bar{\xi}^\nu$. Together with (5.26c), we end up with

$$-\langle \bar{\mu}^\nu, \bar{\xi}^\nu \rangle_{H_0^1(\Omega)} = \langle \bar{\mu}^\nu, \bar{\lambda}^\nu \rangle_{H_0^1(\Omega)} \geq \liminf_{k \in \mathbb{N}} \langle \mu^{k,\nu}, \lambda^{k,\nu} \rangle_{H_0^1(\Omega)} \geq 0$$

and consequently, $\bar{\mu}^\nu$ and $\bar{\xi}^\nu$ satisfy (5.23f).
In order to show (5.23a), we use the equivalent representation discussed below of Definition 5.2.9 and consider

$$x^{k,\nu} = \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \xi^{k,\nu}),$$

with projection operator $\Pi_{Q_\nu}$. With the non-expansiveness of the projection (see e.g. [35, Lemma 1.10]) and the boundedness of $B_\nu$ (with constant $c_\nu > 0$), we get

$$\begin{aligned}\|x^{k,\nu} - \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \bar{\xi}^\nu)\|_{L^2(\Omega)} &= \|\Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \xi^{k,\nu}) - \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \bar{\xi}^\nu)\|_{L^2(\Omega)} \\ &\leq \|\alpha_\nu^{-1} B_\nu^* [\xi^{k,\nu} - \bar{\xi}^\nu]\|_{L^2(\Omega)} \leq \frac{c_\nu}{\alpha_\nu} \|\xi^{k,\nu} - \bar{\xi}^\nu\|_{L^2(\Omega)} \to 0\end{aligned}$$

as $k \to \infty$. As a result, it holds $x^{k,\nu} \to \bar{x}^\nu = \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \bar{\xi}^\nu)$ in $L^2(\Omega)$ and (5.23a) is satisfied.
Next, observe that

$$\langle p^k, \xi^{k,\nu} \rangle_{H_0^1(\Omega)} = -\langle p^k, \lambda^{k,\nu} \rangle_{H_0^1(\Omega)} = \frac{1}{\gamma_k} \langle p^k, y^k \rangle_{H_0^1(\Omega)} - \langle p^k, \tilde{\lambda}^{k,\nu} \rangle_{H_0^1(\Omega)} = \frac{1}{\gamma_k} \langle p^k, y^k \rangle_{H_0^1(\Omega)}.$$

Using (5.26d), we get that $\langle \bar{p}, \bar{\xi}^\nu \rangle = 0$, which implies that $\bar{\xi}^\nu = 0$ if $\bar{p} > 0$ and consequently, the limit points fulfill (5.23e).
It remains to check the validity of (5.23g) and (5.24). The former one can be deduced by the strong convergence of $y^k$ and (5.26d), i.e. it holds

$$\langle \bar{\mu}^\nu, \bar{y} \rangle_{H_0^1(\Omega)} = \lim_{k \to \infty} (\langle \tilde{\mu}^{k,\nu}, y^k - \psi \rangle_{H_0^1(\Omega)} + \frac{1}{\gamma_k} \langle p^k, y^k \rangle_{H_0^1(\Omega)}) = 0,$$

which shows (5.23g).
Finally, we verify that condition (5.24) is valid as well. Similar to the proof of [30, Theorem 4.6], it is sufficient by Egorov's theorem (see e.g. [2]) to prove that $\mu^{k,\nu} \to 0$ pointwise almost everywhere in $\mathcal{Q}_I$.
For this reason, let $\tilde{\omega} \in \mathcal{Q}_I$ be arbitrary but fixed. Then there exists $\tilde{k} \geq 0$ such that $y^k(\tilde{\omega}) > \bar{y}(\tilde{\omega}) > 0$ for all $k \geq \tilde{k}$. Since $\{\mu^{k,\nu}\}$ is bounded and $\gamma_k \to 0$ as $k \to \infty$, it follows by the definition of $\mu^{k,\nu}$ that $p^k(\tilde{\omega}) \to 0$. Moreover, the complementarity condition in (5.20d) implies that

$$\tilde{\mu}^{k,\nu}(\tilde{\omega})y^k(\tilde{\omega}) = 0.$$

Thus, we obtain $\tilde{\mu}^{k,\nu}(\tilde{\omega}) = 0$ and $\tilde{\mu}^{k,\nu} \to 0$ pointwise almost everywhere in $\mathcal{Q}_I$. Consequently, $\mu^{k,\nu}$ converges pointwise to zero. Now, by Egorov's theorem there exists $\mathcal{U}_\epsilon \subseteq \Omega$ for every $\epsilon > 0$ such that $\text{meas}(\mathcal{Q}_I \setminus \mathcal{U}_\epsilon) \leq \epsilon$ and $\mu^{k,\nu}$ converges uniformly in $\mathcal{U}_\epsilon$ for all $\nu = 1, \dots, N$. Now, let $\varphi \in \{\phi \in H_0^1(\Omega) \mid \phi = 0 \text{ a.e. in } \Omega \setminus \mathcal{U}_\epsilon\}$ be arbitrary but fixed. Then, we compute for all $\nu = 1, \dots, N$

$$\begin{aligned}
\langle \bar{\mu}^\nu, \varphi \rangle_{H_0^1(\Omega)} &= \lim_{k\to\infty} \langle \mu^{k,\nu}, \varphi \rangle_{H_0^1(\Omega)} = \lim_{k\to\infty} \int_\Omega \mu^{k,\nu}(\omega)\varphi(\omega)\, d\omega \\
&= \lim_{k\to\infty} \int_{\Omega\setminus\mathcal{U}_\epsilon} \mu^{k,\nu}(\omega)\varphi(\omega)\, d\omega + \lim_{k\to\infty} \int_{\mathcal{U}_\epsilon} \mu^{k,\nu}(\omega)\varphi(\omega)\, d\omega \\
&\to 0,
\end{aligned}$$

which yields condition (5.24). Since $\nu$ was chosen arbitrarily, this shows the claim and $(\bar{x}, \bar{y})$ is an $\epsilon$-almost C-stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}$.
If $\mathcal{Q}_I$ is additional a Lipschitz domain and (5.25) is satisfied for all $\nu = 1, \dots, N$, then the assertion follows by [30, Lemma 4.9]. □

At this point notice that the authors [30] used a different penalty scheme that is based on a Scholtes regularization (see [80]), i.e. $\langle p, y \rangle \leq \alpha$ with some regularization parameter $\alpha > 0$. However, we still can apply [30, Lemma 4.9] to our setting in order to obtain the stronger result. Moreover, we emphasize that although we have more regular multipliers, i.e. $p^k \in L^2(\Omega)$, we only get the strong convergence in $H^{-1}(\Omega)$.
The following result emphasizes the potential of the chosen regularization approach.

**Theorem 5.3.5.** *Let $(x^k, y^k, p^k)$ be a stationary Nash equilibria of $\Gamma^{2,k}_{\text{MLOCP}}$ and assume that the tuple is also a feasible point of $\Gamma^{\text{lq}}_{\text{EPCC}}$. Then, $(x^k, y^k, p^k)$ is an S-stationary Nash equilibrium of $\Gamma^{\text{lq}}_{\text{MLFG}}$.*

*Proof.* The result was shown in [31, Theorem 2.12] for an MPCC. By Definition 4.1.7 the result is also valid for $\Gamma^{\text{lq}}_{\text{MLFG}}$. □

At first sight, the assumptions above seem very strict. However, notice that we use an $l^1$-penalty scheme, which acts exact under appropriate conditions. In this context, there exists $\tilde{k} \in \mathbb{N}$ with $\gamma_{\tilde{k}} > 0$ and $\langle p^{\tilde{k}}, y^{\tilde{k}} \rangle = 0$. Hence, it is enough to consider a finite penalty parameter in order to obtain a solution of $\Gamma_{\text{MLFG}}^{\text{lq}}$.
In a similar way, we show below that a sequence of stationary Nash equilibria of $\Gamma_{\text{MLOCP}}^{1,k}$ converges to an $\epsilon$-almost C-stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}$.

**Theorem 5.3.6.** *Let* $\{\epsilon_k\} \subseteq \mathbb{R}^+$ *be a bounded sequence and* $\{(x^k, y^k)\} \subseteq L^2(\Omega)^N \times H_0^1(\Omega)$ *be such that* $(x^k, y^k)$ *is a stationary Nash equilibria of* $\Gamma_{\text{MLOCP}}^{1,k}$ *for all* $k \in \mathbb{N}$. *Then there exist limit points*

$$(\bar{x}, \bar{y}, \bar{p}, \bar{\xi}, \bar{\mu}) \in L^2(\Omega)^N \times H_0^1(\Omega) \times L^2(\Omega) \times H_0^1(\Omega)^N \times H^{-1}(\Omega)^N$$

*such that*

$$\begin{array}{lll} (x^{k,\nu}, y^k, p^k) & \to (\bar{x}^\nu, \bar{y}, \bar{p}) & \in L^2(\Omega) \times H_0^1(\Omega) \times H^{-1}(\Omega), \\ (\xi^{k,\nu}, \mu^{k,\nu}) & \rightharpoonup (\bar{\xi}^\nu, \bar{\mu}^\nu) & \in H_0^1(\Omega) \times H^{-1}(\Omega) \end{array}$$

*on a subsequence, if* $\gamma_k \to 0$ *as* $k \to \infty$ *for all* $\nu = 1, \ldots, N$ *with*

$$p^k := -\frac{1}{\gamma_k} \max{}^{\epsilon_k}(0, -y^k), \qquad \mu^{k,\nu} := -\frac{1}{\gamma_k} D_y \max{}^{\epsilon_k}(0, -y^k) \xi^{k,\nu}$$

*and* $(\bar{x}, \bar{y})$ *is an* $\epsilon$*-almost C-stationary Nash equilibrium of* $\Gamma_{\text{MLOCP}}$.

*Proof.* The claim was shown in [29, Theorem 3.4] for the case $N = 1$, which can directly be transferred to equilibrium problems. □

All in all and restating the discussion in the beginning of this section, we see that both approaches lead to the same result, i.e. both converge at least strongly to an $\epsilon$-almost C-stationary Nash equilibrium of $\Gamma_{\text{MLOCP}}$. Under additional assumptions, stronger results can be expected for the approach induced by the $l1$-regularization.

## 5.4. Numerical Analysis and Examples

In the next steps, we focus on the numerical computation of an $\epsilon$-almost C-stationary Nash equilibrium. Beginning with the stationary conditions of $\Gamma_{\text{MLOCP}}^{1,k}$ (see Theorem 5.3.2), we observe that (5.22) is equivalent to

$$\mathcal{G}_\nu^k(x^{-\nu}, y^k, \xi^{k,\nu}) = 0 \tag{5.27}$$

with

$$\mathcal{G}_\nu^k(x^{-\nu}, y, \xi^\nu) := \begin{pmatrix} Ay - \frac{1}{\gamma_k}\max^{\epsilon_k}(0,-y) - B_\nu \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \xi^\nu) - B_{-\nu} x^{-\nu} - b \\ A\xi^\nu + \frac{1}{\gamma_k} D_y \max^{\epsilon_k}(0,-y)\xi^\nu + y - y_d^\nu \end{pmatrix},$$

which leads directly to Algorithm 1.

**Algorithm 1** Gauss-Seidel method

**Input:** $N \in \mathbb{N}$, $\gamma_0 > 0$, $(x^0, y^0, \xi^0) \in L^2(\Omega)^N \times H_0^1(\Omega) \times H_0^1(\Omega)^N$.

Set $k = 0$.
**while** $(x^k, y^k)$ is not an $\epsilon$-almost C-stationary Nash equilibrium **do**
  **for all** $\nu = 1, \dots, N$ **do**
    Compute $(y^{k+1}, \xi^{k+1,\nu})$ such that

$$\mathcal{G}_\nu^k(x^{k,-\nu}, y^{k+1}, \xi^{k+1,\nu}) = 0$$

    w.r.t. $x^{k,-\nu} = (x^{k+1,1}, \dots, x^{k+1,\nu-1}, x^{k,\nu}, \dots, x^{k,N})$.
    Set $x^{k+1,\nu} = \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \xi^{k+1,\nu})$.
  **end for**
  Choose $\gamma_{k+1} < \gamma_k$ and set $k = k+1$.
**end while**
**return** $\epsilon$-almost C-stationary Nash equilibrium $(\bar{x}, \bar{y})$

Algorithm 1 is based on a Gauss-Seidel iteration that is typically known for solving linear systems of equations. In game theory, this approach appears quite natural, where we solve each leader's stationarity system separately by using the already computed controls of his predecessors, i.e.

$$x^{k,-\nu} = (x^{k+1,1}, \dots, x^{k+1,\nu-1}, x^{k,\nu}, \dots, x^{k,N})$$

for leader $\nu$ in the $k$th iteration.
Obviously, the main task is the computation of a solution of (5.27). Since the projection operator is not differentiable but only Lipschitz continuous, we apply the semismooth Newton method (see Appendix A) and recall the following result.

**Theorem 5.4.1** ([28, Proposition 4.1]). *Let $p, q \in \mathbb{R}$ be such that $1 \le p < q \le \infty$ is satisfied. Then the mapping* $\max(0,\cdot) : L^q(\Omega) \to L^p(\Omega)$ *is Newton differentiable on* $L^q(\Omega)$ *and*

$$G_\delta(y)(x) := \begin{cases} 1 & \text{if } y(x) > 0, \\ 0 & \text{if } y(x) < 0, \\ \delta & \text{if } y(x) = 0, \end{cases}$$

*with* $\delta \in \mathbb{R}$ *arbitrary but fixed, is a Newton derivative of* $\max(0, \cdot)$.

As a consequence of Theorem 5.4.1, a Newton derivative of $G_\nu^k$ with respect to $(y, \xi^\nu)$ is given by

$$DG_\nu^k(y,\xi^\nu) := \begin{pmatrix} A + \frac{1}{\gamma_k} D_y \max^{\epsilon_k}(0,-y) & -\frac{1}{\alpha_\nu}\mathcal{X}_{\mathcal{I}_{Q_\nu}} B_\nu B_\nu^* \\ I - \frac{1}{\gamma_k} D_{yy}^2 \max^{\epsilon_k}(0,-y)\xi^\nu & A + \frac{1}{\gamma_k} D_y \max^{\epsilon_k}(0,-y) \end{pmatrix},$$

with $\mathcal{I}_{Q_\nu} := \{\omega \in \Omega \mid a_\nu(\omega) \le \alpha_\nu^{-1}(B_\nu^*\xi^\nu)(\omega) \le b_\nu(\omega)\}$. Here, we used $G_0(y)$ as Newton derivative of $\max(0, \cdot)$ and that $x^\nu = \Pi_{Q_\nu}(\alpha_\nu^{-1} B_\nu^* \xi^\nu)$ is equivalent to

$$x^\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu + (\alpha_\nu^{-1} B_\nu^* \xi^\nu - b_\nu)_+ - (a_\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu)_+ = 0.$$

**Remark 5.4.2.** *At this point, notice that we do not consider* $\max(0, \cdot)$ *as a function from* $L^2(\Omega)$ *to* $L^2(\Omega)$*, which would contradict Theorem 5.4.1, but as a mapping from* $L^6(\Omega)$ *to* $L^2(\Omega)$. *In this context, the latter relation is a result of Theorem 2.1.15, which guarantees that* $H_0^1(\Omega) \hookrightarrow L^6(\Omega)$ *for* $\Omega \subseteq \mathbb{R}^d$ *with* $d = 1, 2, 3$.

The computation of (5.27) is summarized in Algorithm 2. In particular, it is known that this method converges superlinearly in a neighborhood around the solution under suitable assumptions on $DG_\nu^k$ (see Theorem A.0.5).
Algorithm 2 terminates, if the sum of the relative errors of each leader's system

$$\sum_{\nu=1}^{N} \frac{\|G_\nu^k(x^{-\nu}, y^k, \xi^{k,\nu}) - G_\nu^k(x^{-\nu}, y^{k-1}, \xi^{k-1,\nu})\|}{\|G_\nu^k(x^{-\nu}, y^k, \xi^{k,\nu}) - G_\nu^k(x^{-\nu}, y^0, \xi^{0,\nu})\|},$$

drops below a certain threshold.
The outer loop, i.e. Algorithm 1, stops, if the relative error of the residual

$$\left| \frac{\mathrm{res}^k - \mathrm{res}^{k-1}}{\mathrm{res}^k - \mathrm{res}^0} \right|, \tag{5.28}$$

satisfies a given termination criterion. In this context, the residual is given by

$$\begin{aligned} \mathrm{res}^k := & \|Ay^k - p^k - Bx^k - b\|_{H_0^1(\Omega)} + \|p^k + \max(0, -(p^k + y^k))\|_{L^2(\Omega)} \\ & + \sum_{\nu=1}^{N} (\|A\xi^{k,\nu} - \mu^{k,\nu} - y_d^\nu + y^k\|_{H_0^1(\Omega)} + |\max(0, \langle \mu^{k,\nu}, \xi^{k,\nu}\rangle_{H_0^1(\Omega)})| \\ & + \|\xi^{k,\nu}\|_{L^2(\tilde{\mathcal{Q}}_A^+)} + |\langle \mu^{k,\nu}, y^k \rangle_{H_0^1(\Omega)}|). \end{aligned}$$

**Algorithm 2** Semismooth Newton's method for leader $\nu$

**Input:** $(y^0, \xi^{0,\nu}) \in H_0^1(\Omega) \times H_0^1(\Omega)$, $x^{-\nu} \in L^2(\Omega)^{N-1}$.

1: Set $k := 0$.

2: **while** $(y^k, \xi^{k,\nu})$ does not solve (5.27) **do**

3: Solve

$$DG_\nu^k(y^k, \xi^{k,\nu}) \begin{pmatrix} d_y \\ d_\xi \end{pmatrix} = -G_\nu^k(x^{-\nu}, y^k, \xi^{k,\nu}).$$

4: Set $y^{k+1} := y^k + d_y$, $\xi^{k+1,\nu} = \xi^{k,\nu} + d_\xi$ and $k = k+1$.

5: **end while**

6: **return** Solution $(y^k, \xi^{k,\nu})$ of the nonsmooth equation (5.27).

A similar idea can be used to deal with the second approach. However, we consider a different ansatz that does not compute the solutions of auxiliary problems but the C-stationary conditions directly. For this reason, recall that the latter stationarity system is given by

$$x^\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu + (\alpha_\nu^{-1} B_\nu^* \xi^\nu - b_\nu)_+ - (a_\nu - \alpha_\nu^{-1} B_\nu^* \xi^\nu)_+ = 0, \tag{5.29a}$$

$$Ay + p - Bx - b = 0, \tag{5.29b}$$

$$p + \max(0, -(p+y)) = 0, \tag{5.29c}$$

$$A\xi^\nu - \mu^\nu + y - y_d^\nu = 0, \tag{5.29d}$$

$$\langle \mu^\nu, y \rangle = 0,\ (p, \xi^\nu) = 0,\ \langle \mu^\nu, \xi^\nu \rangle \le 0, \tag{5.29e}$$

for all $\nu = 1, \dots, N$. Motivated by [86], we use an active set strategy and start by solving

$$\mathcal{H}(x, y, p) := \begin{pmatrix} Ay + p - Bx - b \\ p + \max(0, -(p+y)) \end{pmatrix} = 0 \tag{5.30}$$

with respect to $(y, p)$, where $x$ is assumed to be fixed. Then (5.30) is equivalent to the coupled system (5.29b) and (5.29c). Notice that $\mathcal{H}(x, y, p)$ is again a semismooth mapping. In contrast to the former approach, we can not apply the semismooth Newton's method directly, since we have $\max(0, \cdot) : L^2(\Omega) \to L^2(\Omega)$, which is not Newton differentiable (see Theorem 5.4.1). However, this is not an issue on a discretized level.
We continue by recalling the strongly active, biactive and inactive sets that are given by

$$\begin{aligned} \tilde{\mathcal{Q}}_A^+ &= \{\omega \in \Omega \mid y(\omega) = 0 \wedge p(\omega) < 0\}, \\ \mathcal{Q}_A^0 &= \{\omega \in \Omega \mid y(\omega) = 0 \wedge p(\omega) = 0\}, \\ \mathcal{Q}_I &= \{\omega \in \Omega \mid y(\omega) > 0\}. \end{aligned}$$

Then observe that on the one hand (5.29d) can be written as

$$\begin{array}{rcll} \mu^\nu & = & -y_\mathrm{d}^\nu & \text{a.e. in } \tilde{\mathcal{Q}}_A^+, \\ A\xi^\nu - \mu^\nu & = & y_\mathrm{d}^\nu & \text{a.e. in } \mathcal{Q}_A^0, \\ A\xi^\nu & = & y_\mathrm{d}^\nu - y & \text{a.e. in } \mathcal{Q}_I \end{array}$$

and on the other hand (5.29e) reduces to

$$\begin{array}{lll} \xi^\nu & = 0 & & \text{a.e. in } \tilde{\mathcal{Q}}_A^+, \\ \langle \mu^\nu, \xi^\nu \rangle & \leq 0 & & \text{a.e. in } \mathcal{Q}_A^0, \\ \langle \mu^\nu, v \rangle & = 0 & \forall v \in H_0^1(\Omega) & \text{a.e. in } \mathcal{Q}_I \end{array}$$

for all $\nu = 1, \dots, N$. The derivations above are summarized in Algorithm 3. Similar to Algorithm 1 we stop, if (5.28) drops below a given threshold.

At this point, we can not give any convergence results for both methods. For contributions that apply a Gauss-Seidel method, we refer to [33, 36, 43]. In this context, it was shown in [43, Satz 6.1] that this approach finds a Nash equilibrium of a convex game, if the whole sequence generated by the algorithm converges towards that point. In [36], the authors analyzed a MLFG in finite dimensions and used the same method to find S-stationary Nash equilibria of the original problem. Again under the assumption that the whole sequence converges. In our setting, we have only convergence on a subsequence in general.
Similarly, it is not guaranteed that the strongly active set in Algorithm 3 is convergent in any sence of set-convergence, which is due to the lower regularity of $p$. In particular, the lower level Lagrange multiplier need not to be continuous. Nevertheless, we observed a good behavior for our tested examples below.
Throughout the remainder of this section, we assume four leaders, i.e. $N = 4$ and consider the two-dimensional unit square $\Omega = [0, 1]^2$ as our underlying domain. The lower level data is given by

- $A = -\Delta$ (see Example 5.1.1),
- $B_\nu = \mathcal{X}_{\Omega_\nu}$, where $\Omega_\nu \subseteq \Omega$ and $\mathcal{X}_{\Omega_\nu}$ denotes the corresponding characteristic function for all $\nu = 1, \dots, 4$.

The domain $\Omega$ and the Laplacian are discretized via a uniform grid with $n^2$ elements and a standard finite difference 5-point stencil, respectively. Moreover, we define the mesh size by $h := n^{-1}$ and use the starting points $(u^0, y^0, \xi^0) = (0, 0, 0)$ for Algorithm 1 and $(u^0, y^0, p^0) = (0, 0, 0)$ for Algorithm 3. Both methods were tested with mesh sizes up to $n = 2^9$.
In case of Algorithm 1 we set $\gamma_0 = 10^{-1}$ with update scheme $\gamma_{k+1} = 0.1\gamma_k$ and smoothing

**Algorithm 3** Active-set strategy

**Input:** $N \in \mathbb{N}$ and $(x^0, y^0, p^0) \in L^2(\Omega)^N \times H_0^1(\Omega) \times L^2(\Omega)$.

1: Set $k = 0$.
2: **while** $(x^k, y^k)$ is not an C-stationary Nash equilibrium **do**
3: Compute $(y^{k+1}, p^{k+1})$ such that

$$\mathcal{H}(x^k, y^{k+1}, p^{k+1}) = 0.$$

4: Set

$$\begin{aligned}
\tilde{\mathcal{Q}}^+_{A,k+1} &:= \{\omega \in \Omega \mid y^{k+1}(\omega) = 0 \wedge p^{k+1}(\omega) < 0\}, \\
\mathcal{Q}^0_{A,k+1} &:= \{\omega \in \Omega \mid y^{k+1}(\omega) = 0 \wedge p^{k+1}(\omega) = 0\}, \\
\mathcal{Q}_{I,k+1} &:= \{\omega \in \Omega \mid y^{k+1}(\omega) > 0\}.
\end{aligned}$$

5: **for all** $\nu = 1, \ldots, N$ **do**
6: Compute $(\xi^{k+1,\nu}, \mu^{k+1,\nu})$ by solving

$$\begin{aligned}
\xi^{k+1,\nu} &= 0 \quad \text{a.e. in } \tilde{\mathcal{Q}}^+_{A,k+1}, \\
\mu^{k+1,\nu} + y_\mathrm{d}^\nu &= 0 \quad \text{a.e. in } \tilde{\mathcal{Q}}^+_{A,k+1}, \\
\\
A\xi^{k+1,\nu} - \mu^{k+1,\nu} - y_\mathrm{d}^\nu &= 0 \quad \text{a.e. in } \mathcal{Q}^0_{A,k+1}, \\
\max(0, \langle \mu^{k+1,\nu}, \xi^{k+1,\nu}) &= 0 \quad \text{a.e. in } \mathcal{Q}^0_{A,k+1}, \\
\\
\mu^{k+1,\nu} &= 0 \quad \text{a.e. in } \mathcal{Q}_{I,k+1}, \\
\xi^{k+1,\nu} - A^{-1}(y_\mathrm{d}^\nu - y^{k+1}) &= 0 \quad \text{a.e. in } \mathcal{Q}_{I,k+1},
\end{aligned}$$

7: Compute $x^{k+1,\nu}$ by solving (5.29a).
8: **end for**
9: Set $k = k + 1$.
10: **end while**
11: **return** C-stationary Nash equilibrium $(\bar{x}, \bar{y})$.

function $\epsilon(\gamma) := \gamma^2$. In both cases the outer while-loop is terminated, if the stopping criteria drops below the threshold $\mathrm{tol}_\mathrm{out} = 10^{-6}$. The semismooth Newton's methods stop, if the residual is below $5 \cdot 10^{-4}$.
While Application 5.4.3 is a modification of an equilibrium problem with linear state

equation (see [33, Example 4.1]) to an equilibrium problem with variational inequality constraints, Application 5.4.4 can be seen as an extension of an MPCC motivated by [30, Example 6.1] to an EPCC.

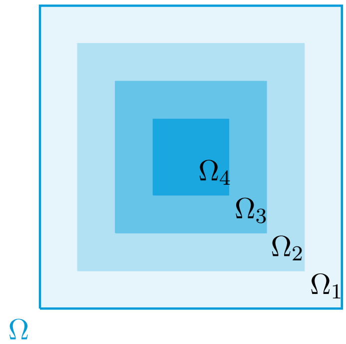



Figure 5.1.: Control domains of each leader.

Both examples are applied to control domains $\Omega_\nu$ as depicted in Figure 5.1.

**Application 5.4.3.** *For all $\nu = 1, \dots, N$, leader $\nu$'s desired state, depicted in Figure 5.2, is given by*

$$y_{\mathrm{d}}^{\nu}(\omega_1, \omega_2) := 10^3 \max(0, 1 - 4\max(|\omega_1 - q_1^{\nu}|, |\omega_2 - q_2^{\nu}|))$$

*with $q_1 = (0.25, 0.75, 0.25, 0.75)$ and $q_2 = (0.25, 0.25, 0.75, 0.75)$.*

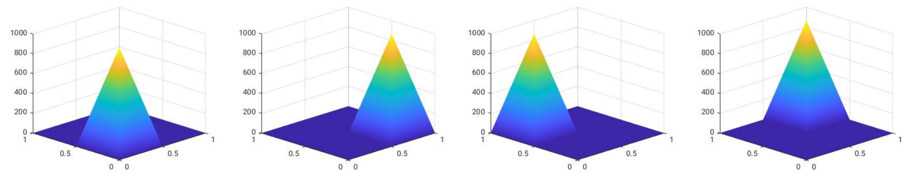


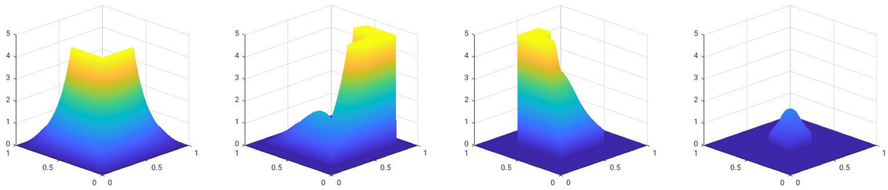
Figure 5.2.: Desired states of each leader.

*The obstacle $\psi$ and an external force are defined by*

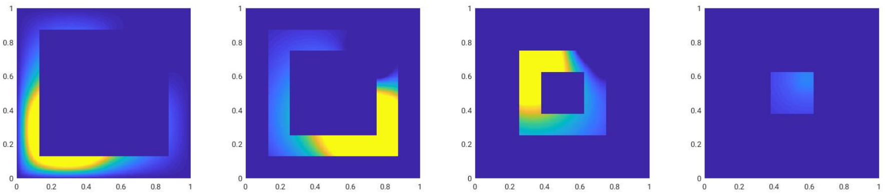
- $\psi(\omega_1, \omega_2) := \cos(2\sqrt{(\omega_1^2 - 0.5)^2 + (\omega_2^2 - 0.5)^2}) - 0.7,$
- $b := \mathcal{X}_{\mathcal{B}}(\Delta\psi + 5)$ *with* $\mathcal{B} := [0.35, 0.65]^2$.

*In particular, we consider a scenario with an obstacle not equal to zero. Hence, a substitution is needed in advance to be consistent with our model. Moreover, we have uniform and constant box constraints, i.e. $[a_\nu, b_\nu] = [0, 5]$ and $\alpha_\nu = 1$ for all $\nu = 1, \dots, 4$.*

The results of Application 5.4.3 are now shown below. Figure 5.3 depicts each leader's control $\bar{x}^\nu$ of the computed $\epsilon$-almost C-stationary Nash equilibrium.



(a) Three-dimensional view



(b) Top view

Figure 5.3.: Computed controls $\bar{x}$ of Application 5.4.3.

In contrast, Figure 5.4 shows the corresponding state $\bar{y}$ and lower-level Lagrange multiplier $\bar{p}$. In particular, observe that the maximum value of $\bar{y}$ is attained approximately at $0.4$. Comparing this result with the desired states that each leader aims to reach, we see that no one gains an advantage and the state stays on a lower level above the obstacle in the middle of the underlying domain $\Omega$. Moreover, notice that the biactive set is empty. In this context, all weaker stationary Nash equilibria coincide.

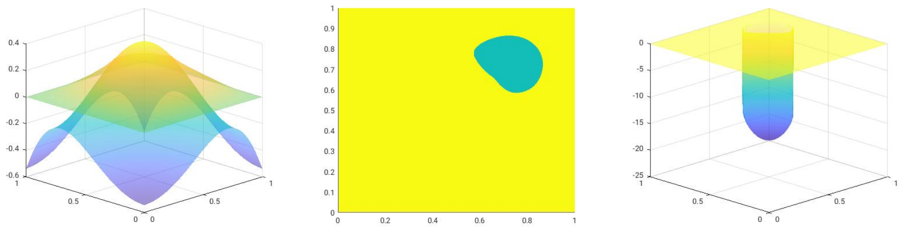


Figure 5.4.: Computed state $\bar{y}$ with obstacle (left), inactive set in yellow and strongly active set in green (middle), Lagrange multiplier $\bar{p}$ (right).

In Figure 5.5, we see the error estimation of both methods. In this context, notice

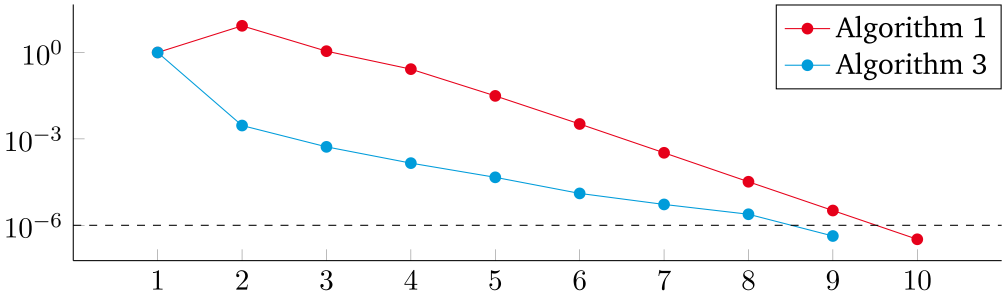
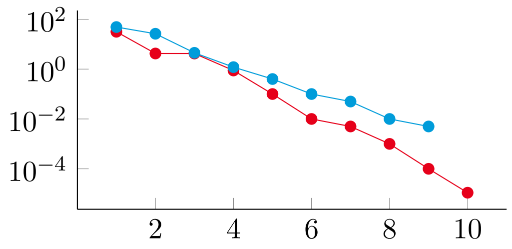



Figure 5.5.: Relative error of residuals.

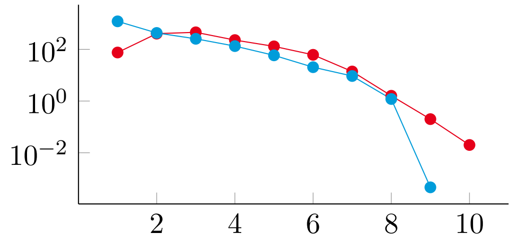
that both algorithms need nearly the same number of outer iterations until the given termination criterion is satisfied. However, we observed that Algorithm 3 is much faster than its counterpart. The latter is not surprising as Algorithm 1 solves $N$ semismooth Newton's methods, while the active-set strategy in general only solves a nonsmooth equation once per iteration.

Finally, Figure 5.6 visualizes the convergence rate of the important parameters $(x, y, p)$ that once again confirms the promising behavior of both methods with respect to the chosen example.

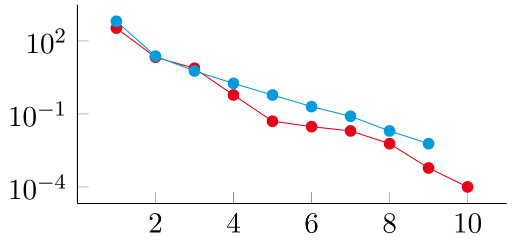
In contrast to Application 5.4.3, the subsequent example constructs an equilibrium problem with nonempty biactive set $\mathcal{Q}_A^0$.



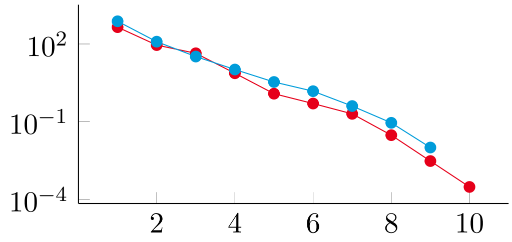
(a) $\|y^{k+1} - y^k\|$



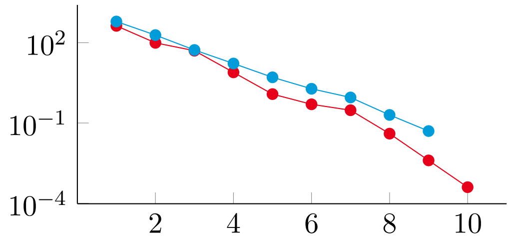
(b) $\|p^{k+1} - p^k\|$



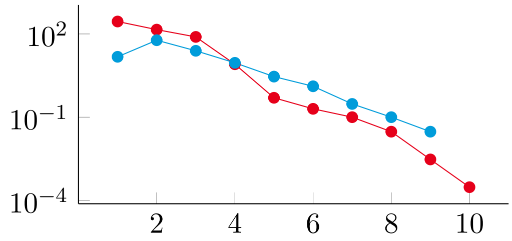
(c) $\|x^{k+1,1} - x^{k,1}\|$



(d) $\|x^{k+1,2} - x^{k,2}\|$



(e) $\|x^{k+1,3} - x^{k,3}\|$



(f) $\|x^{k+1,4} - x^{k,4}\|$

Figure 5.6.: Convergence rate of state, lower level Lagrange multiplier and the individual controls for Application 5.4.3 with respect to Algorithm 1 and Algorithm 3.

**Application 5.4.4.** *In order to obtain a nonempty biactive set, we define*

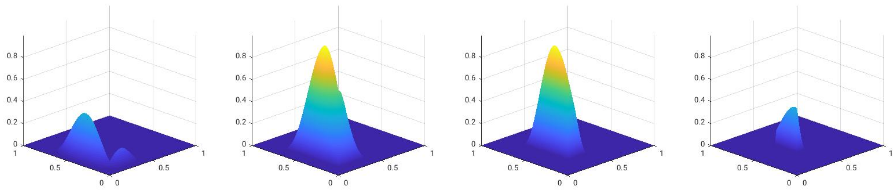
- $\bar{y}(\omega_1, \omega_2) := \mathcal{X}_{\mathcal{A}}(\omega_1, \omega_2) z_1(\omega_1) z_2(\omega_2)$ *with*
  - *a)* $\mathcal{A} := (0, 0.5) \times (0, 0.8)$,
  - *b)* $z_1(\omega_1) := -4096\omega_1^6 + 6144\omega_1^5 - 3072\omega_1^4 + 512\omega_1^3$,
  - *c)* $z_2(\omega_2) := -244.140625\omega_2^6 + 585.9375\omega_2^5 - 468.9375\omega_2^4 + 125\omega_2^3$,
- $\bar{x}^\nu(\omega_1, \omega_2) := \mathcal{X}_{\Omega_\nu} \bar{y}(\omega_1, \omega_2)$ *for all* $\nu = 1, \dots, 4$,
- $\bar{p}(\omega_1, \omega_2) := -2\max(0, -|\omega_1 - 0.8| - |(\omega_2 - 0.2)\omega_1 - 0.3| + 0.35)$,

*and set*

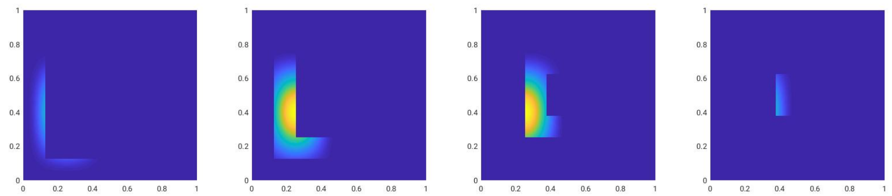
$$b = A\bar{y} - B\bar{x} + \bar{p}, \qquad y_d^\nu = \bar{y} - \bar{p} + A\bar{u}^\nu$$

*for all* $\nu = 1\dots, N$. *Moreover, we set* $\psi \equiv 0$ *and* $\alpha_\nu = 1$ *for all* $\nu = 1, \dots, N$. *In particular, notice that we do not have any control constraints.*

Using the same tolerances as in Application 5.4.3, we end up with the corresponding results depicted in Figure 5.7 and Figure 5.8.



(a) Three-dimensional view



(b) Top view

Figure 5.7.: Computed controls $\bar{x}$ of Application 5.4.4.

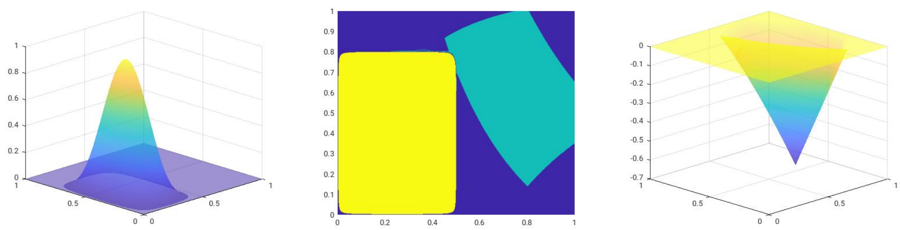


Figure 5.8.: Computed state $\bar{y}$ with obstacle (left), inactive set in yellow, strongly active set in green and biactive set in blue (middle), Lagrange multiplier $\bar{p}$ (right).

As can be seen in Figure 5.9, we again observe that Algorithm 3 applied to Application 5.4.4 is not only faster than Algorithm 1 but also needs less total iterations to pass the desired threshold. Nevertheless, we emphasize that both methods find satisfying solutions that is confirmed in Figure 5.10.

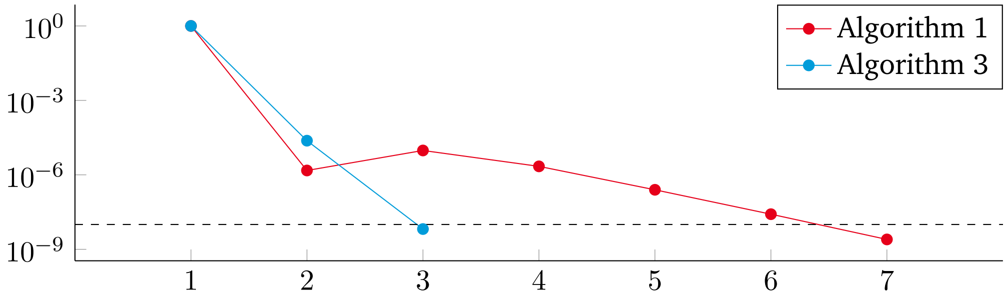
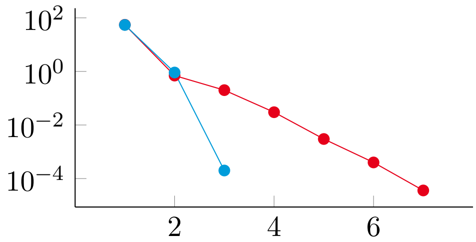
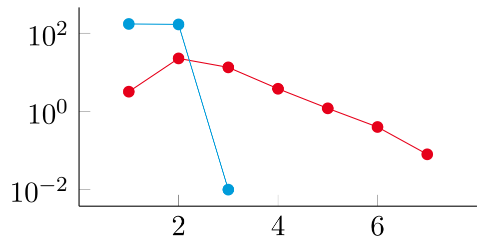


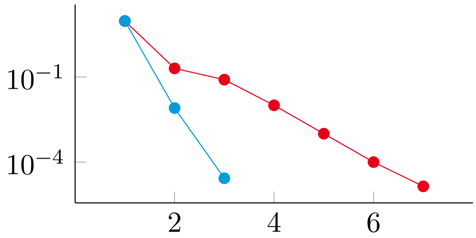

Figure 5.9.: Relative error of residuals.



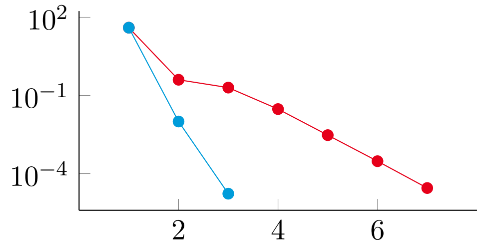

(a) $\|y^{k+1} - y^k\|$



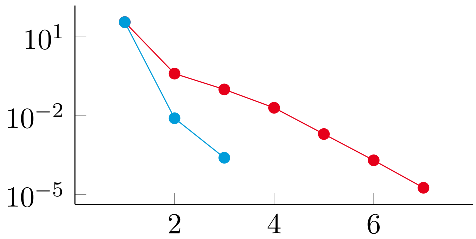

(b) $\|p^{k+1} - p^k\|$



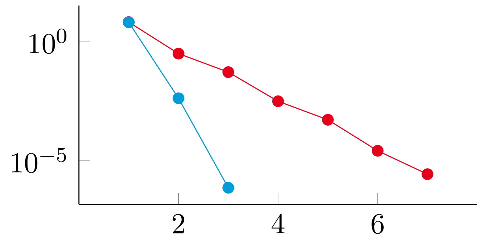

(c) $\|x^{k+1,1} - x^{k,1}\|$




(d) $\|x^{k+1,2} - x^{k,2}\|$




(e) $\|x^{k+1,3} - x^{k,3}\|$




(f) $\|x^{k+1,4} - x^{k,4}\|$

Figure 5.10.: Convergence rate of state, lower level Lagrange multiplier and the individual controls for Application 5.4.4 with respect to Algorithm 1 and Algorithm 3.

# 6. Autonomous Driving: A Generalized Nash Equilibrium Problem in Lebesgue Spaces

After considering a multi-leader optimal control framework of the obstacle problem in Chapter 5 that was based on a multi-leader single-follower game, we move on and focus on another class of non-convex equilibrium problems.
Motivated by the contribution [14] "A generalized Nash equilibrium approach for optimal control problems of autonomous cars", we consider an equilibrium problem, which models the competition of several vehicles in a simplified traffic scenario. Although the underlying feasible set is non-convex, the authors were able to show the existence of a normalized Nash equilibrium.
Our contribution focuses on the corresponding EPCC. For this reason, Section 6.1 starts with a general problem description. Moreover, we highlight the non-convex part of the feasible set and use the so-called Big-M formulation, which transforms the non-convex constraint into a system of inequalities with mixed-integer variables. In this context, the latter denotes that each player's strategy set is comprised of both continuous and integer variables. Notice that this approach is not new and we refer for instance to [76, 77] for contributions with respect to mixed-integer equilibrium problems in finite spaces and to [6] for a general survey of mixed-integer nonlinear optimization. Section 6.2 is devoted to an abstract generalized Nash equilibrium problem with mixed-integer variables (MINEP) in Lebesgue spaces. We verify its relationship to EPCCs and derive stationary Nash equilibria in the sense of Definition 3.3.1. In particular, we show that all weaker forms of stationary Nash equilibria are equivalent, which is therefore denoted by S-stationary Nash equilibrium. The derived results are then applied in Section 6.3 to the optimal control problem introduced in the beginning of this chapter. In this context, we use the same penalization technique that we considered in Subsection 4.2.2 for MLFGs in order to show the convergence of a sequence of stationary Nash equilibria to an S-stationary Nash equilibrium (Subsection 6.3.1). Afterwards, the results are analyzed under a numerical perspective and visualized on the basis of two selected traffic scenarios.

## 6.1. Problem Formulation

Motivated by the contribution on autonomous driving in [14], we consider a scenario with $N \in \mathbb{N}$ vehicles that are able to communicate with each other. In this context, let $\Omega := (0, T)$ (with $T > 0$) be a given time interval and denote by $\gamma^\nu : \mathbb{R}_0^+ \to \mathbb{R}^2$ the preassigned path of vehicle $\nu$ for all $\nu = 1, \ldots, N$. Then player $\nu$'s goal is to reach his destination or at least drive as far as possible in the given time, which can be seen exemplarily in Figure 6.1 for $N = 3$.

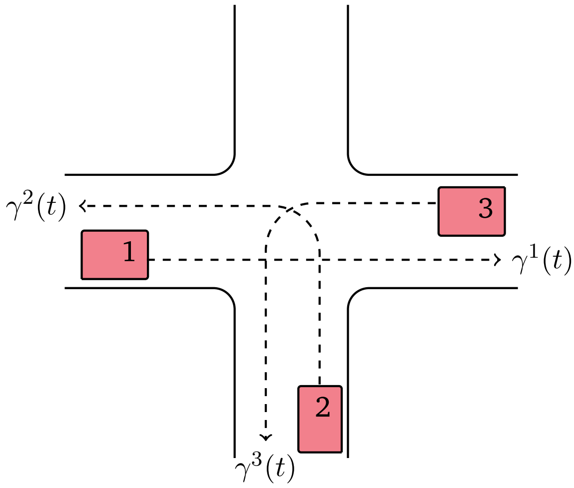



Figure 6.1.: Traffic scenario with $N = 3$ vehicles.

In order to describe the dynamics of this system, we use the following initial value problem

$$\dot{y}^\nu(\tau) = \begin{pmatrix} y_2^\nu(\tau) \\ x^\nu(\tau) \end{pmatrix} \text{ for almost all } \tau \in \Omega, \tag{6.1a}$$

$$y^\nu(0) = \begin{pmatrix} y_0^\nu \\ v_0^\nu \end{pmatrix}. \tag{6.1b}$$

Here, player $\nu$'s state is given by $y^\nu = (y_1^\nu, y_2^\nu) \in Y := AC^{1,2}(\Omega)^2$, where $y_1^\nu$ represents the position on the path $\gamma^\nu$ and $y_2^\nu$ denotes the velocity. The latter variable is controlled by the acceleration $x^\nu \in X := L^2(\Omega)$ and the parameters $y_0^\nu$ and $v_0^\nu$ in (6.1b) characterize the initial condition, i.e. the start position and the velocity in the beginning, where the latter is assumed to be strictly positive. Hence and in contrast to Chapter 5, the state of

our system is given by an *ordinary differential equation* (ODE) instead of a PDE.
Now, this scenario can be modeled by an equilibrium problem, denoted by $\Gamma_{\text{AD}}$, where player $\nu$ solves

$$\begin{aligned} \min \quad & -y_1^\nu(T) + \frac{\alpha_\nu}{2}\|x^\nu\|_{L^2(\Omega)} \\ \text{w.r.t.} \quad & (x^\nu, y^\nu) \in L^2(\Omega) \times AC^{1,2}(\Omega)^2, \\ \text{s.t.} \quad & x^\nu \in Q_\nu, \\ & \dot{y}^\nu = \begin{pmatrix} y_2^\nu \\ x^\nu \end{pmatrix}, \ y^\nu(0) = \begin{pmatrix} y_0^\nu \\ v_0^\nu \end{pmatrix}, \\ & G(y_1^\nu, y_1^\mu) \geq 0 \qquad \forall \mu \in \{1, \ldots, N\} \setminus \{\nu\}. \end{aligned} \tag{6.2}$$

In this context, the set $Q_\nu \subseteq L^2(\Omega)$, which is assumed to be nonempty, closed, convex and bounded, describes additional control constraints for all vehicles $\nu = 1, \ldots, N$. The coupled inequality constraint $G$ is known as *collision avoidance* that is in general given by a non-convex mapping. Motivated by [14], we define

$$G(y^\nu, y^\mu) := \max\{|y_1^\nu(\tau) - \hat{y}_{\nu\mu}|, |y_1^\mu(\tau) - \hat{y}_{\mu\nu}|\} - \delta \tag{6.3}$$

for all $\tau \in \Omega$ with $\delta > 0$. Moreover, $(\hat{y}_{\nu\mu}, \hat{y}_{\mu\nu}) \in \mathbb{R}^2$ is defined by

$$(\hat{y}_{\nu\mu}, \hat{y}_{\mu\nu}) := \operatorname*{argmin}_{(t_\nu, t_\mu) \in [0,T]^2} \left\{\frac{1}{2}\|\gamma^\nu(t_\nu) - \gamma^\mu(t_\mu)\|_2^2\right\},$$

where $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^2$.

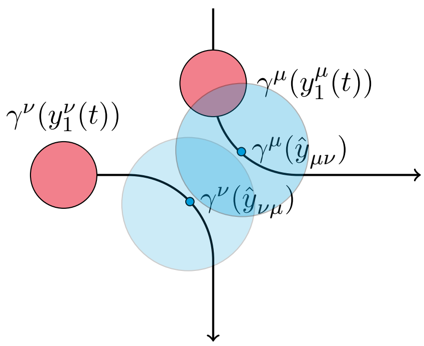



Figure 6.2.: Collision avoidance between two vehicles.

Geometrically, the points $(\hat{y}_{\nu\mu}, \hat{y}_{\mu\nu})$ describe the position on two disjoint paths $\gamma^\nu$ and $\gamma^\mu$, respectively, such that their distance to each other is minimal. In particular, the collision avoidance is not restricted to common intersection points (see Figure 6.1), but ensures that at most one vehicle is within the potential conflict area (see Figure 6.2). In this case, the parameter $\delta$ can be interpreted as the radius of the critical areas of both paths.

**Remark 6.1.1.** *As it may happen that two paths have multiple conflict points/zones, it is usually required to consider the collision avoidance constraint* (6.3) *for all of them separately. To keep the notation compact, we focus only on one mutually agreed conflict point* $(\hat{y}_{\nu\mu}, \hat{y}_{\mu\nu})$ *for each pair of players, which is w.l.o.g. the conflict area that is the closest to player* $\nu$*'s starting position (see Figure 6.3). However, it is possible to extend the model by regarding multiple collision avoidance constraints* $G^j$ *with* $j = 1, \ldots, n_C$*, where* $n_C \in \mathbb{N}$ *denotes the maximum number of conflict zones between two different vehicles.*

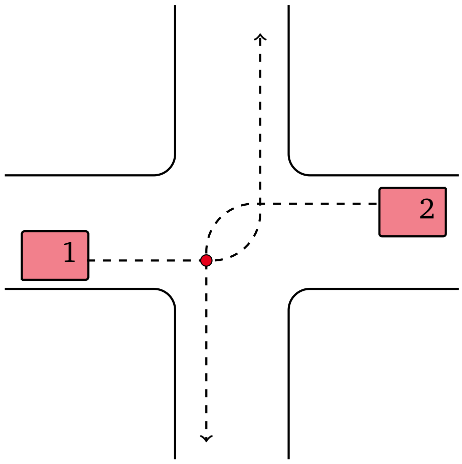



Figure 6.3.: Traffic scenario with two players and intersection point.

Since all cost functionals only depend on the individual player variables, $\Gamma_{\text{AD}}$ is a potential game. This property was used in [14] to show the existence of a Nash equilibrium. However, less was done with regard to optimality conditions. As a consequence, we first transfer the feasible set of $\Gamma_{\text{AD}}$ into a set that fits the structure of $\Gamma_{\text{EPCC}}$.
In this context, we need an intermediate step and use a well-known idea, the so-called *Big-M formulation* (see for instance [71],[81] and Appendix B.1 for further details). We fix two players $\nu, \mu \in \{1, \ldots, N\}$ with w.l.o.g. $\nu < \mu$. Then (6.3) can be replaced by the

following linear inequality

$$\tilde{H}\begin{bmatrix} y_1^\nu - \hat{y}_{\nu\mu} \\ y_1^\mu - \hat{y}_{\mu\nu} \end{bmatrix} + \tilde{\mathcal{M}}\begin{bmatrix} z_{\mu,1}^\nu \\ z_{\mu,2}^\nu \end{bmatrix} - \tilde{b} \geq 0, \tag{6.4a}$$

$$z_\mu^\nu \in (\{0,1\} \cap L^2(\Omega))^2 \tag{6.4b}$$

with

$$\tilde{H} := (h_1\ h_2) := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix},\ \tilde{\mathcal{M}} := \begin{pmatrix} M & M \\ M & -M \\ -M & M \\ -M & -M \end{pmatrix} \text{ and } \tilde{b} := \begin{pmatrix} \delta \\ \delta - M \\ \delta - M \\ \delta - 2M \end{pmatrix}.$$

Here, $M >> 0$ is assumed to be sufficiently large with $M > \delta$ and condition (6.4b) shows that $z_\mu^\nu$ is a binary variable and hence, an integer for almost all $\tau \in \Omega$. In particular, the feasible set of $z_\mu^\nu$ is given by

$$\left\{z \in L^2(\Omega) \;\middle|\; z(\tau) = 1 \wedge z(\tau) = 0 \text{ f.a.a. } \tau \in \Omega\right\},$$

which includes for instance the Dirichlet function $1_{\mathbb{Q}}$, general characteristic functions $\mathcal{X}_{\mathcal{A}}$ with arbitrary sets $\mathcal{A} \subseteq \Omega$ and the trivial functions $z \equiv 1$ and $z \equiv 0$. By defining the vector-valued function

$$z^\nu := (z_1^\nu, \ldots, z_{\nu-1}^\nu, z_{\nu+1}^\nu, \ldots, z_N^\nu)^T \in (\{0,1\} \cap L^2(\Omega))^{2(N-1)}$$

of all binary variables of player $\nu$ and the vectors

$$\begin{aligned} \hat{y}^\nu &:= (\hat{y}_{\nu 1}, \ldots, \hat{y}_{\nu\nu-1}, \hat{y}_{\nu\nu+1}, \ldots, \hat{y}_{\nu N})^T \in L^2(\Omega)^{N-1}, \\ \hat{y}^{-\nu} &:= (\hat{y}_{1\nu}, \ldots, \hat{y}_{\nu-1\nu}, \hat{y}_{\nu+1\nu}, \ldots, \hat{y}_{N\nu})^T \in L^2(\Omega)^{N-1}, \end{aligned}$$

which summarize the critical points of player $\nu$ with respect to his opponents and the critical points of player $\nu$'s opponents with respect to the player himself, respectively, we can write player $\nu$'s $N-1$ constraints of type (6.4a) as the following single inequality

$$Hy^\nu + H_{\text{opp}}y^{-\nu} + \mathcal{M}z^\nu - (b + \hat{H}_1\hat{y}^\nu + \hat{H}_2\hat{y}^{-\nu}) \geq 0. \tag{6.5}$$

Here, the operators

$$\begin{aligned} H &\in L(AC^{1,2}(\Omega)^2, L^2(\Omega)^{4(N-1)}), \\ H_{\text{opp}} &\in L(AC^{1,2}(\Omega)^{2(N-1)}, L^2(\Omega)^{4(N-1)}), \\ \hat{H}_\nu, \hat{H}_{-\nu} &\in L(L^2(\Omega)^{N-1}, L^2(\Omega)^{4(N-1)}, \\ \mathcal{M} &\in L(L^2(\Omega)^{2(N-1)}, L^2(\Omega)^{4(N-1)}) \end{aligned}$$

and $b \in L^2(\Omega)^{4(N-1)}$ are obtained by concatenation of (6.4a) (see Example 6.1.2).

**Example 6.1.2.** *We consider a scenario with four players and focus on player* 3. *Then* (6.4) *is given by*

$$\begin{pmatrix} h_1 & 0 \\ h_1 & 0 \\ h_1 & 0 \end{pmatrix} y^3 + \begin{pmatrix} h_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_2 & 0 \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ y^4 \end{pmatrix} + \begin{pmatrix} \tilde{M} & 0 & 0 \\ 0 & \tilde{M} & 0 \\ 0 & 0 & \tilde{M} \end{pmatrix} \begin{pmatrix} z_1^3 \\ z_2^3 \\ z_4^3 \end{pmatrix}$$
$$- \begin{pmatrix} \tilde{b} \\ \tilde{b} \\ \tilde{b} \end{pmatrix} - \begin{pmatrix} h_1 & 0 & 0 \\ 0 & h_1 & 0 \\ 0 & 0 & h_1 \end{pmatrix} \begin{pmatrix} \hat{y}_1^3 \\ \hat{y}_2^3 \\ \hat{y}_4^3 \end{pmatrix} - \begin{pmatrix} h_2 & 0 & 0 \\ 0 & h_2 & 0 \\ 0 & 0 & h_2 \end{pmatrix} \begin{pmatrix} \hat{y}_3^1 \\ \hat{y}_3^2 \\ \hat{y}_3^4 \end{pmatrix} \geq 0.$$

Summarizing the discussion above, $\Gamma_{\text{AD}}$ can equivalently be written as the following MINEP, denoted by $\Gamma_{\text{MINEP}}^{\text{AD}}$, where player $\nu$ solves

$$\begin{aligned} \min \quad & -y_1^\nu(T) + \frac{\alpha_\nu}{2}\|x^\nu\|_{L^2(\Omega)}, \\ \text{w.r.t.} \quad & (x^\nu, y^\nu, z^\nu) \in L^2(\Omega) \times AC^{1,2}(\Omega)^2 \times L^2(\Omega)^{2(N-1)}, \\ \text{s.t.} \quad & x^\nu \in Q_\nu, \ z^\nu \in \{0,1\}^{2(N-1)} \\ & \dot{y}^\nu = \begin{pmatrix} y_2^\nu \\ x^\nu \end{pmatrix}, \ y^\nu(0) = \begin{pmatrix} y_0^\nu \\ v_0^\nu \end{pmatrix}, \\ & Hy^\nu + H_{\text{opp}} y^{-\nu} + \mathcal{M} z^\nu - c^\nu \geq 0 \end{aligned} \tag{6.6}$$

with $c^\nu := b + \hat{H}_1 \hat{y}^\nu + \hat{H}_2 \hat{y}^{-\nu}$.

**Remark 6.1.3.** *Due to this reformulation, the number of optimizing variables increases by* $2(N-1)$. *However, notice that the latter number can be further reduced by only considering opponent vehicles that can pass the path of player* $\nu$ *in a distance smaller than* $2\delta$ *instead of all vehicles* $\mu \neq \nu$.

The equilibrium problem $\Gamma_{\text{MINEP}}^{\text{AD}}$ is considered again in Section 6.3. For now, we focus on general MINEPs in Lebesgue spaces in the subsequent section.

## 6.2. Generalized Nash Equilibrium Problems with Mixed-Integer Variables in Lebesgue Spaces

In order to define stationary Nash equilibria of $\Gamma_{\text{MINEP}}^{\text{AD}}$, we now focus on general GNEPs with mixed-integer variables. For this reason, we consider the equilibrium problem, denoted by $\Gamma_{\text{MINEP}}$, where player $\nu$ solves

$$\begin{aligned} \min\ & f^\nu(x^\nu, x^{-\nu}, z^\nu, z^{-\nu}) \\ \text{w.r.t.}\ & (x^\nu, z^\nu) \in X_\nu \times Z_\nu, \\ \text{s.t.}\ & x^\nu \in Q_\nu,\ z^\nu \in \{0,1\}^{n_\nu}, \\ & g^\nu(x^\nu, x^{-\nu}, z^\nu, z^{-\nu}) \in C_\nu \end{aligned} \tag{6.7}$$

for all $\nu = 1, \dots, N$. Here, we set $Z_\nu := L^p(\Omega)^{n_\nu}$ with $2 \leq p < \infty$, where $\Omega \subseteq \mathbb{R}^d$ is an arbitrary measurable subset. Notice that in this context $Z_\nu$ describes a reflexive Banach space and $\Omega$ is not restricted on the real line anymore. Moreover, we use the compact notation $Q = Q_1 \times \cdots \times Q_N$, $W = W_1 \times \cdots \times W_N$, $X = X_1 \times \cdots \times X_N$ and $Z = Z_1 \times \cdots \times Z_N$ in order to summarize the involved sets and spaces, respectively. The remaining data are given by the mappings $f^\nu : X \times Z \to \mathbb{R}$, $g^\nu : X \times Z \to W_\nu$ and the sets $Q_\nu \subseteq X_\nu$, $C_\nu \subseteq W_\nu$ for arbitrary Banach spaces $W_\nu$ and $X_\nu$ for all $\nu = 1, \dots, N$ under the following assumptions.

**Assumption 6.2.1.** *For all $\nu = 1, \dots, N$, we assume that $f^\nu$ and $g^\nu$ are continuously differentiable w.r.t $(x^\nu, z^\nu)$. Moreover, $Q_\nu \subseteq X_\nu$ and $C_\nu \subseteq W_\nu$ are nonempty, closed and convex.*

Now, notice that for all $\nu = 1, \dots, N$ and $i = 1, \dots, n_\nu$, we have

$$z_i^\nu \in \{0,1\} \cap L^p(\Omega) \quad \iff \quad z_i^\nu \in L^p(\Omega) \text{ solves } 0 \leq z_i^\nu \perp 1 - z_i^\nu \geq 0,$$

where the complementarity constraints have to be understood in the 'almost everywhere'-sense. In particular, the right hand side is well-defined, since the dual space of $L^p(\Omega)$ is given by $L^q(\Omega)$ with $q = \frac{p}{p-1}$, which implies $q \leq 2$. Then Theorem 2.1.13 guarantees that $1 - z_i^\nu \in L^q(\Omega)$ is valid. Thus, $\Gamma_{\text{MINEP}}$ can equivalently be written as the following EPCC, denoted by $\Gamma^{\text{MINEP}}_{\text{EPCC}}$, where player $\nu$ solves

$$\begin{aligned} \min\ & f^\nu\left(x^\nu, x^{-\nu}, z^\nu, z^{-\nu}\right) \\ \text{w.r.t.}\ & (x^\nu, z^\nu) \in X_\nu \times Z_\nu, \\ \text{s.t.}\ & x^\nu \in Q_\nu, \\ & g^\nu(x^\nu, x^{-\nu}, z^\nu, z^{-\nu}) \in C_\nu, \\ & z^\nu \in K_\nu,\ z^\nu - 1 \in K_\nu^\circ,\ \langle z^\nu - 1, z^\nu \rangle_{L^p(\Omega)} = 0 \end{aligned} \tag{6.8}$$

with

$$K_\nu := \left\{ v \in L^p(\Omega)^{n_\nu} \;\middle|\; \forall i \in \mathcal{N}_\nu := \{1, \dots, n_\nu\} :\ v_i(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega \right\}. \tag{6.9}$$

In this context, it was shown for instance in [55, Lemma 3.4] that the corresponding dual cone $K_\nu^\circ$ is given by

$$K_\nu^\circ = \big\{\eta \in L^q(\Omega)^{n_\nu} \;\big|\; \forall i \in \mathcal{N}_\nu : \; \eta_i(\omega) \leq 0 \text{ f.a.a. } \omega \in \Omega\big\}.$$

As a consequence, the normal cone can be computed by

$$\begin{aligned} \mathcal{N}_{K_\nu}(\bar{z}^\nu) &= K_\nu^\circ \cap \{\bar{z}^\nu\}^\perp \\ &= \big\{\eta \in L^q(\Omega)^{n_\nu} \;\big|\; \forall i \in \mathcal{N}_\nu : \; \eta_i(\omega) \leq 0 \wedge \eta_i(\omega)\bar{z}_i^\nu(\omega) = 0 \text{ f.a.a. } \omega \in \Omega\big\}, \end{aligned} \tag{6.10}$$

where we used that $K_\nu$ is a convex cone. In a similar way as for EPCCs in $H_0^1(\Omega)$, we proceed by defining the corresponding stationarity concepts in $L^p(\Omega)$.
For the subsequent analysis we denote by $q$ the conjugate coefficient of $p$ and consider the generic complementarity constraints in $L^p(\Omega)$, i.e.

$$G(z) \in K, \; H(z) \in K^\circ, \; \langle H(z), G(z)\rangle_{L^p(\Omega)} = 0$$

with

- $G : L^p(\Omega)^n \to L^p(\Omega)^n$,
- $H : L^p(\Omega)^n \to L^q(\Omega)^n$,
- $K = \big\{v \in L^p(\Omega)^n \;\big|\; \forall i \in \mathcal{N} := \{1, \ldots, n\} : \; v_i(\omega) \geq 0 \text{ f.a.a. } \omega \in \Omega\big\}$,
- $K^\circ = \big\{\eta \in L^q(\Omega)^n \;\big|\; \forall i \in \mathcal{N} : \; \eta_i(\omega) \leq 0 \text{ f.a.a. } \omega \in \Omega\big\}$.

Moreover, we define for all $i \in \mathcal{N}$ the inactive, strongly active and biactive set by

$$\begin{aligned} \mathcal{I}_{+0}(z, i) &:= \big\{\omega \in \Omega \;\big|\; G(z)_i(\omega) > 0 \wedge H(z)_i(\omega) = 0\big\}, \\ \mathcal{I}_{0-}(z, i) &:= \big\{\omega \in \Omega \;\big|\; G(z)_i(\omega) = 0 \wedge H(z)_i(\omega) < 0\big\}, \\ \mathcal{I}_{00}(z, i) &:= \big\{\omega \in \Omega \;\big|\; G(z)_i(\omega) = 0 \wedge H(z)_i(\omega) = 0\big\}. \end{aligned}$$

Now, the following lemma can be seen as the analogon of Lemma 5.2.8 in $L^p(\Omega)^n$, which was also verified in the proof of [53, Theorem 3.14].

**Lemma 6.2.2.** *Let $K$ be given as* (6.9) *and set $(\bar{G}, \bar{H}) := (G(\bar{z}), H(\bar{z}))$. Then we have*

$$\begin{aligned} \mathrm{cl}(K - K \cap \{\bar{H}\}^\perp) \cap \{\bar{H}\}^\perp &= \big\{\lambda \in L^p(\Omega)^n \;\big|\; \forall i \in \mathcal{N} : \lambda_i(\omega) = 0 \text{ f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z}, i)\big\}, \\ \mathrm{cl}(K^\circ - K^\circ \cap \{\bar{G}\}^\perp) \cap \{\bar{G}\}^\perp &= \big\{\mu \in L^q(\Omega)^n \;\big|\; \forall i \in \mathcal{N} : \mu_i(\omega) = 0 \text{ f.a.a. } \omega \in \mathcal{I}_{+0}(\bar{z}, i)\big\} \end{aligned}$$

*and*

$$\mathcal{K}_K(\bar{G},\bar{H}) = \left\{ \lambda \in L^p(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \begin{array}{ll} \lambda_i(\omega) = 0 & \text{f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z},i), \\ \lambda_i(\omega) \geq 0 & \text{f.a.a. } \omega \in \mathcal{I}_{00}(\bar{z},i) \end{array} \right\},$$

$$\mathcal{K}_{K^\circ}(\bar{H},\bar{G}) = \left\{ \mu \in L^q(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \begin{array}{ll} \mu_i(\omega) = 0 & \text{f.a.a. } \omega \in \mathcal{I}_{+0}(\bar{z},i), \\ \mu_i(\omega) \leq 0 & \text{f.a.a. } \omega \in \mathcal{I}_{00}(\bar{z},i) \end{array} \right\}.$$

*Proof.* Similar as in (6.10), we compute

$$\begin{aligned} K \cap \{\bar{H}\}^\perp &= \left\{ \lambda \in L^p(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \; \lambda_i(\omega) \geq 0 \wedge \lambda_i(\omega)\bar{H}_i(\omega) = 0 \;\text{ f.a.a. } \omega \in \Omega \right\} \\ &= \left\{ \lambda \in L^p(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \begin{array}{ll} \lambda_i(\omega) \geq 0 & \text{f.a.a. } \omega \in \mathcal{I}_{+0}(\bar{z},i) \cup \mathcal{I}_{00}(\bar{z},i) \\ \lambda_i(\omega) = 0 & \text{f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z},i) \end{array} \right\} \end{aligned}$$

and hence,

$$K - K \cap \{\bar{H}\}^\perp = \left\{ \lambda \in L^p(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \; \lambda_i(\omega) \geq 0 \text{ f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z},i) \right\}.$$

The latter set is closed and thus, the assertion follows. The second identity is verified in the same way.
In order to prove the representation of the critical cones, we first focus on $\mathcal{K}_K(\bar{G},\bar{H})$. As the critical cone is defined by $\mathcal{T}_K(\bar{G}) \cap \{\bar{H}\}^\perp$, we look for a representation of the tangent cone in $L^p(\Omega)^n$. In this context, it was shown in [7, Example 2.64] that

$$\mathcal{T}_K(\bar{G}) = \left\{ \lambda \in L^p(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \; \lambda_i(\omega) \geq 0 \text{ f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z},i) \cup \mathcal{I}_{00}(\bar{z},i) \right\}$$

and we conclude that

$$\mathcal{K}_K(\bar{G},\bar{H}) = \left\{ \lambda \in L^p(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \begin{array}{ll} \lambda_i(\omega) = 0 & \text{f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z},i), \\ \lambda_i(\omega) \geq 0 & \text{f.a.a. } \omega \in \mathcal{I}_{00}(\bar{z},i) \end{array} \right\}.$$

For the representation of $\mathcal{K}_{K^\circ}(\bar{H},\bar{G})$, we refer to Example 2.2.8, which states the polyhedricity of the cone $K$. By Lemma 2.2.4, this implies $\mathcal{K}_{K^\circ}\left(\bar{H},\bar{G}\right) = \mathcal{K}_K(\bar{G},\bar{H})^\circ$ and thus, it remains to show that

$$\mathcal{K}_K(\bar{G},\bar{H})^\circ = \left\{ \mu \in L^q(\Omega)^n \;\middle|\; \forall i \in \mathcal{N}: \begin{array}{ll} \mu_i(\omega) = 0 & \text{f.a.a. } \omega \in \mathcal{I}_{+0}(\bar{z},i), \\ \mu_i(\omega) \leq 0 & \text{f.a.a. } \omega \in \mathcal{I}_{00}(\bar{z},i) \end{array} \right\} =: \mathcal{P}.$$

Therefore, let $\mu \in \mathcal{P}$ and observe that for arbitrary $\lambda \in \mathcal{K}_K(\bar{G}, \bar{H})$ we obtain for all $i \in \mathcal{N}$ that

$$\begin{aligned} \mu_i(\omega)\lambda_i(\omega) &= 0 && \text{f.a.a. } \omega \in \mathcal{I}_{+0}(\bar{z}, i), \\ \mu_i(\omega)\lambda_i(\omega) &= 0 && \text{f.a.a. } \omega \in \mathcal{I}_{0-}(\bar{z}, i), \\ \mu_i(\omega)\lambda_i(\omega) &\leq 0 && \text{f.a.a. } \omega \in \mathcal{I}_{00}(\bar{z}, i) \end{aligned}$$

and hence, $\mu \in \mathcal{K}_K(\bar{G}, \bar{H})^\circ$. In reverse, let $\mu \in \mathcal{K}_K(\bar{G}, \bar{H})^\circ$ and choose $\lambda \in \mathcal{K}_K(\bar{G}, \bar{H})$ arbitrarily but fixed. Then it holds $\mu_i(\omega)\lambda_i(\omega) \leq 0$ for all $i \in \mathcal{N}$ and f.a.a. $\omega \in \Omega$, which directly follows by the representation of $\mathcal{K}_K(\bar{G}, \bar{H})$ and thus, $\mu \in \mathcal{P}$ is valid. $\square$

Lemma 6.2.2 directly shows that an S-stationary Nash equilibrium is a W-stationary Nash equilibrium. since it holds

$$\mathcal{K}_K(\bar{G}, \bar{H}) \times \mathcal{K}_{K^\circ}(\bar{H}, \bar{G}) \subseteq \text{cl}(K - K \cap \{\bar{H}\}^\perp) \cap \{\bar{H}\}^\perp \times \text{cl}(K^\circ - K^\circ \cap \{\bar{G}\}^\perp) \cap \{\bar{G}\}^\perp.$$

Coming back to the the special type of complementarity constraints in $\Gamma_{\text{EPCC}}^{\text{MINEP}}$, Lemma 6.2.2 yields the following result.

**Corollary 6.2.3.** *For all $\nu = 1, \dots, N$, define*

$$G^\nu(z^\nu) := z^\nu \quad \text{and} \quad H^\nu(z^\nu) := z^\nu - 1$$

*and assume that there exists $\bar{z} \in Z$ such that*

$$\bar{G}^\nu \in K_\nu, \quad \bar{H}^\nu \in K_\nu^\circ, \quad \langle \bar{H}^\nu, \bar{G}^\nu \rangle_{Z^\nu} = 0$$

*hold for all $\nu = 1, \dots, N$ with $\bar{G}^\nu := G^\nu(\bar{z}^\nu)$ and $\bar{H}^\nu := H^\nu(\bar{z}^\nu)$. Then we have*

$$\begin{aligned} \mathcal{K}_{K_\nu}(\bar{G}^\nu, \bar{H}^\nu) &= \text{cl}(K_\nu - K_\nu \cap \{\bar{H}^\nu\}^\perp) \cap \{\bar{H}_\nu\}^\perp, \\ \mathcal{K}_{K_\nu^\circ}(\bar{H}^\nu, \bar{G}^\nu) &= \text{cl}(K_\nu^\circ - K_\nu^\circ \cap \{\bar{G}^\nu\}^\perp) \cap \{\bar{G}^\nu\}^\perp. \end{aligned}$$

*Proof.* Let $\nu \in \{1, \dots, N\}$ and $i \in \mathcal{N}_\nu$ be arbitrary but fixed. Thus, it holds

$$\begin{aligned} \mathcal{I}_{00}^\nu(\bar{z}, i) &= \big\{\omega \in \Omega \;\big|\; \bar{G}_i^\nu(\omega) = 0 \wedge \bar{H}_i^\nu(\omega) = 0\big\} \\ &= \big\{\omega \in \Omega \;\big|\; \bar{z}_i^\nu(\omega) = 0 \wedge \bar{z}_i^\nu(\omega) = 1\big\}. \end{aligned}$$

Hence, $\mathcal{I}_{00}^\nu(\bar{z}, i)$ is empty and the assertion follows by Lemma 6.2.2. $\square$

A direct consequence of Corollary 6.2.3 is that the concept of W-stationary and S-stationary Nash equilibria coincide for $\Gamma_{\text{EPCC}}^{\text{MINEP}}$.
The discussion above is summarized in the subsequent definition, which follows by the general concepts in Section 3.3. Moreover, notice that we use the term 'S-stationary Nash equilibrium' although all concepts coincide.

**Definition 6.2.4.** A feasible point $(\bar{x}, \bar{z}) \in X \times Z$ is called *S-stationary Nash equilibrium* of $\Gamma_{\text{MINEP}}$, if there exist multipliers

$$(\xi, \mu, \lambda) \in W^* \times Z^* \times Z$$

such that for all $\nu = 1, \ldots, N$, it holds

$$(D_{x^\nu} f^\nu(\bar{x}, \bar{z}) + D_{x^\nu} g^\nu(\bar{x}, \bar{z})\xi^\nu, x^\nu - \bar{x}^\nu) \geq 0 \qquad \forall x^\nu \in Q_\nu, \tag{6.11a}$$

$$D_{z^\nu} f^\nu(\bar{x}, \bar{z}) + D_{z^\nu} g^\nu(\bar{x}, \bar{z})\xi^\nu + \mu^\nu + \lambda^\nu = 0, \tag{6.11b}$$

$$\xi^\nu \in \mathcal{N}_{C_\nu}(g^\nu(\bar{x}, \bar{z})), \tag{6.11c}$$

$$\forall i \in \mathcal{N}_\nu : \ \lambda_i^\nu(\omega) = 0 \quad \text{f.a.a. } \omega \in \mathcal{I}_{0-}^\nu(\bar{z}, i), \tag{6.11d}$$

$$\forall i \in \mathcal{N}_\nu : \ \mu_i^\nu(\omega) = 0 \quad \text{f.a.a. } \omega \in \mathcal{I}_{+0}^\nu(\bar{z}, i), \tag{6.11e}$$

where the strongly active and the inactive set are given by

$$\mathcal{I}_{0-}^\nu(\bar{z}, i) = \{\omega \in \Omega \mid \bar{z}_i^\nu(\omega) = 0\} \quad \text{and} \quad \mathcal{I}_{+0}^\nu(\bar{z}, i) = \{\omega \in \Omega \mid \bar{z}_i^\nu(\omega) = 1\}.$$

Similar to MLFGs, we derive MINEP-tailored necessary conditions for a Nash equilibrium to be an S-stationary Nash equilibrium in the sense of Theorem 3.3.3. However, observe that for an arbitrary feasible tuple $(\bar{x}, \bar{z}) \in X \times Z$, FRCQ (see (3.20)) is not applicable. In order to see this, assume that FRCQ is satisfied in $(\bar{x}, \bar{z})$, i.e. for all $\nu = 1, \ldots, N$ it holds

$$\begin{pmatrix} D_{x^\nu} g(\bar{x}, \bar{z}) & D_{z^\nu} g(\bar{x}, \bar{z}) \\ 0_{X^\nu} & I_{L^p(\Omega)^{n_\nu}} \\ 0_{X^\nu} & -I_{L^p(\Omega)^{n_\nu} L^q(\Omega)^{n_\nu}} \end{pmatrix} \begin{bmatrix} \mathcal{R}_{Q_\nu}(\bar{x}^\nu) \\ L^p(\Omega)^{n_\nu} \end{bmatrix} = \begin{bmatrix} W_\nu \\ Lp^{n_\nu} \\ L^q(\Omega)^{n_\nu} \end{bmatrix}. \tag{6.12}$$

Then the second and third condition are given by

$$\begin{pmatrix} I_{L^p(\Omega)^{n_\nu}} \\ -I_{L^p(\Omega)^{n_\nu} L^q(\Omega)^{n_\nu}} \end{pmatrix} L^p(\Omega)^{n_\nu} = \begin{pmatrix} L^p(\Omega)^{n_\nu} \\ L^q(\Omega)^{n_\nu} \end{pmatrix}.$$

Hence, for arbitrary elements $(w_1, w_2) \in L^p(\Omega)^{n_\nu} \times L^q(\Omega)^{n_\nu}$ we find $v \in L^p(\Omega)^{n_\nu}$ such that

$$v = w_1 = -w_2$$

is satisfied. It is obvious that this system is in general not satisfied. As a consequence, we have to deal with weaker CQs in this context (see Theorem 3.3.3) in order to state necessary optimality conditions.

## 6.3. Theoretical and Numerical Analysis

In the following sections, we focus again on the equilibrium problem introduced in Section 6.1 and show that we find a sequence that converges to an S-stationary Nash equilibrium (Subsection 6.3.1). The results are further analyzed numerically and visualized by means of two hypothetical traffic scenarios in Subsection 6.3.2.

### 6.3.1. Convergence to an S-stationary Nash equilibrium

By the observations in the previous section, we know that $\Gamma^{\text{AD}}_{\text{MINEP}}$ can be written as an EPCC. In the following, we focus on a slightly different equilibrium problem, denoted by $\Gamma^{\text{AD}}_{\text{EPCC}}$, where player $\nu$ solves

$$\begin{aligned}
\min \quad & -y_1^\nu(T) + \frac{\alpha_\nu}{2}\|x^\nu\|^2_{L^2(\Omega)} + \frac{\beta_\nu}{2}\|z^\nu\|^2_{H^1(\Omega)} \\
\text{w.r.t.} \quad & (x^\nu, y^\nu, z^\nu) \in L^2(\Omega) \times AC^{1,2}(\Omega)^2 \times H^1(\Omega)^{2(N-1)}, \\
\text{s.t.} \quad & x^\nu \in Q_\nu, \\
& \dot{y}^\nu = \begin{pmatrix} y_2^\nu \\ x^\nu \end{pmatrix}, \; y^\nu(0) = \begin{pmatrix} y_0^\nu \\ v_0^\nu \end{pmatrix}, \\
& Hy^\nu + H_{\text{opp}}y^{-\nu} + \mathcal{M}z^\nu - c^\nu \geq 0, \\
& z^\nu \geq 0, \; z^\nu - 1 \leq 0, \; (z^\nu, z^\nu - 1) = 0.
\end{aligned} \tag{6.13}$$

Here, we strengthened the setting by looking for solutions $z^\nu$ in $H^1(\Omega)^{2(N-1)}$ instead of $L^2(\Omega)^{2(N-1)}$ and added a cost term with $\beta_\nu > 0$ and $\|z^\nu\|^2_{H^1(\Omega)} = \|z^\nu\|^2_{L^2(\Omega)} + \|Dz^\nu\|^2_{L^2(\Omega)}$ with differential operator $D$. As a brief outlook for the subsequent analysis, this assumption on $z$ is needed in order to obtain the strong convergence of $z$ in $L^2(\Omega)$. Nevertheless, notice that this assumption seems naturally as we do not wish that $z^\nu$ has arbitrarily many jumps, which happens for instance if $z^\nu$ is the Dirichlet function.
Furthermore, notice that the underlying domain $\Omega$ is given again by $\Omega = (0, T)$ with $T > 0$. Before we adopt Definition 6.2.4, we show two auxiliary results that are needed in the proof of Theorem 6.3.3 below. While the first lemma reduces the solution of two equalities to the computation of one single ODE with end point condition, the second lemma can be seen as the integration by parts applied to two special functions.

**Lemma 6.3.1.** *Let $f : \mathbb{R} \times AC^{1,2}(\Omega) \to L^2(\Omega)$ and $x_F \in \mathbb{R}$ be given. If $\theta \in AC^{1,2}(\Omega)$ solves*

$$\dot{\theta} = f(\tau, \theta), \quad \theta(T) = x_F,$$

*then* $\Theta(\tau) := \theta(0) + \int_0^\tau \theta(t)\, dt$ *satisfies the following two equalities:*

$$\Theta(0) - x_F + \int_0^T f(t, \theta(t))\, dt = 0, \qquad \dot{\Theta}(\tau) - x_F + \int_\tau^T f(t, \theta(t))\, dt = 0. \tag{6.14}$$

*Proof.* On the one hand we have

$$\Theta(0) - x_F + \int_0^T f(t, \theta(t))\, dt = \theta(0) - x_F + \int_0^T \dot{\theta}(t)\, dt = \theta(0) - x_F + \theta(T) - \theta(0) = 0$$

and on the other hand

$$\dot{\Theta}(\tau) - x_F + \int_\tau^T f(t, \theta(t))\, dt = \theta(\tau) - x_F + \int_\tau^T \dot{\theta}(t)\, dt = \theta(\tau) - x_F + \theta(T) - \theta(\tau) = 0$$

and hence, (6.14) is satisfied. □

**Lemma 6.3.2.** *Let* $u$ *and* $v$ *be two arbitrary integrable functions on* $\Omega = (0, T)$ *with* $T > 0$*. Then it holds*

$$\int_0^T \left( v(t) \int_0^t u(\tau)\, d\tau \right) dt = \int_0^T \left( u(t) \int_t^T v(\tau)\, d\tau \right) dt.$$

*Proof.* We apply integration by parts and obtain

$$\begin{aligned} \int_0^T \left( v(t) \int_0^t u(\tau)\, d\tau \right) dt &= \int_0^t v(\tau)\, d\tau \int_0^t u(\tau)\, d\tau \Big|_{t=0}^{T} - \int_0^T \left( u(t) \int_0^t v(\tau)\, d\tau \right) dt \\ &= \int_0^T \left( u(t) \int_0^T v(\tau)\, d\tau \right) dt - \int_0^T \left( u(t) \int_0^t v(\tau)\, d\tau \right) dt \\ &= \int_0^T \left( u(t) \int_t^T v(\tau)\, d\tau \right) dt, \end{aligned}$$

which shows the result. □

Now, we can state the S-stationary Nash equilibrium conditions for $\Gamma^{\mathrm{AD}}_{\mathrm{EPCC}}$.

**Theorem 6.3.3.** *A feasible point* $(\bar{x}, \bar{y}, \bar{z}) \in L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times L^2(\Omega)^{2N(N-1)}$ *of* $\Gamma^{\mathrm{AD}}_{\mathrm{EPCC}}$ *is an S-stationary Nash equilibrium, if there exist multipliers*

$$(\rho, \theta, \mu) \in AC^{1,2}(\Omega)^{2N} \times L^2(\Omega)^{4N(N-1)} \times L^2(\Omega)^{2N(N-1)}$$

*such that* $(\bar{x}, \bar{y}, \bar{z}, \rho, \theta, \mu)$ *solves the system*

$$(\alpha_\nu x^\nu - \rho_2^\nu, w^\nu - x^\nu) \geq 0 \quad \forall w^\nu \in Q_\nu, \tag{6.15a}$$

$$\begin{aligned} \dot{y}_1^\nu - y_2^\nu &= 0, & y_1^\nu(0) &= y_0^\nu, \\ \dot{y}_2^\nu - x^\nu &= 0, & y_2^\nu(0) &= v_0^\nu, \end{aligned} \tag{6.15b}$$

$$\begin{aligned} \dot{\rho}_1^\nu + H_1^* \theta^\nu &= 0, & \rho_1^\nu(T) &= 1, \\ \dot{\rho}_2^\nu + \rho_1^\nu &= 0, & \rho_2^\nu(T) &= 0, \end{aligned} \tag{6.15c}$$

$$0 \leq \theta^\nu \perp Hy^\nu + H_{\text{opp}} y^{-\nu} + \mathcal{M} z^\nu - c^\nu \geq 0, \tag{6.15d}$$

*for all* $\nu = 1, \ldots, N$. *Moreover, for all* $i = 1, \ldots, 2(N-1)$ *it holds*

$$\begin{aligned} \mu_i^\nu(\tau) &= 0 \quad \text{f.a.a. } \tau \in \mathcal{I}_{+0}^\nu(z, i), \\ \mu_i^\nu(\tau) - (\mathcal{M}^* \theta^\nu)_i(\tau) + \beta_\nu (D^* D z^\nu)_i(\tau) &= 0 \quad \text{f.a.a. } \tau \in \mathcal{I}_{0-}^\nu(z, i). \end{aligned} \tag{6.16}$$

*Proof.* Let $\nu \in \{1, \ldots, N\}$ be arbitrary but fixed and set

$$v^\nu := (x^\nu, y_1^\nu, y_2^\nu) \in V := L^2(\Omega) \times AC^{1,2}(\Omega) \times AC^{1,2}(\Omega).$$

Moreover, we define the mapping

$$g^\nu(v^\nu, y^{-\nu}, z^\nu) := \begin{pmatrix} v_2^\nu - y_0^\nu - \int_0^\cdot v_3^\nu(t)\, dt \\ v_3^\nu - v_0^\nu - \int_0^\cdot v_1^\nu(t)\, dt \\ H[v_2^\nu, v_3^\nu] + H_{\text{opp}} y^{-\nu} + \mathcal{M} z^\nu - c^\nu \end{pmatrix},$$

which is comprised of player $\nu$'s constraints (see (6.13)). Thus, player $\nu$ solves

$$\begin{aligned} \min \quad & -v_2^\nu(T) + \frac{\alpha_\nu}{2} \|v_1^\nu\|_{L^2(\Omega)}^2 + \frac{\beta_\nu}{2} \|z^\nu\|_{H^1(\Omega)}^2 \\ \text{w.r.t.} \quad & (v^\nu, z^\nu) \in V \times L^2(\Omega)^{2(N-1)}, \\ \text{s.t.} \quad & v^\nu \in \tilde{Q}_\nu, \ g^\nu(v^\nu, y^{-\nu}, z^\nu) \in \tilde{C}_\nu \end{aligned}$$

with

$$\begin{aligned} &\tilde{Q}_\nu := Q_\nu \times AC^{1,2}(\Omega) \times AC^{1,2}(\Omega), \\ &\tilde{C}_\nu := \{0\} \times \{0\} \times L^2(\Omega)_0^{4(N-1),+} \subseteq W := AC^{1,2}(\Omega) \times AC^{1,2}(\Omega) \times L^2(\Omega)^{4(N-1)}. \end{aligned}$$

By Definition 3.3.1, $(\bar{v}, \bar{z})$ is an S-stationary Nash equilibrium of $\Gamma_{\text{EPCC}}^{\text{AD}}$ iff there exist multipliers $(\theta, \mu, \lambda) \in W^N \times L^2(\Omega)^{2N(N-1)} \times L^2(\Omega)^{2N(N-1)}$ such that for all $\nu = 1, \ldots, N$

$(\bar{v}^\nu, \bar{z}^\nu)$ satisfies

$$
\begin{aligned}
(\alpha_\nu \bar{v}_1^\nu + D_{v_1^\nu} g^\nu(\bar{v}^\nu, \bar{y}^{-\nu}, \bar{z}^\nu)^* \theta^\nu, v^\nu - \bar{v}_1^\nu)_{L^2(\Omega)} &\geq 0 \qquad \forall v^\nu \in Q_\nu, \\
D_{v_2^\nu} g^\nu(\bar{v}^\nu, \bar{y}^{-\nu}, \bar{z}^\nu)^* \theta^\nu - 1 &= 0, \\
D_{v_3^\nu} g^\nu(\bar{v}^\nu, \bar{y}^{-\nu}, \bar{z}^\nu)^* \theta^\nu &= 0, \\
D_{z^\nu} g^\nu(\bar{v}^\nu, \bar{y}^{-\nu}, \bar{z}^\nu)^* \theta^\nu + \mu^\nu + \lambda^\nu + \beta_\nu(z^\nu + D^* D z^\nu) &= 0, \\
\theta^\nu &\in \mathcal{N}_{\tilde{C}_\nu}(g^\nu(\bar{v}^\nu, \bar{y}^{-\nu}, \bar{z}^\nu))
\end{aligned}
\tag{6.17}
$$

and for all $\nu = 1, \dots, N$ and all $i = 1, \dots, 2(N-1)$ it holds

$$
\begin{aligned}
\mu_i^\nu(\tau) &= 0 \quad \text{f.a.a. } \tau \in \mathcal{I}_{+0}^\nu(\bar{z}, i), \\
\lambda_i^\nu(\tau) &= 0 \quad \text{f.a.a. } \tau \in \mathcal{I}_{0-}^\nu(\bar{z}, i).
\end{aligned}
\tag{6.18}
$$

Since the gradient of a functional highly depends on the chosen scalar product, we are looking for a more explicit representation. Thus, we recall (see Theorem 2.1.19) that the scalar product in $AC^{1,2}(\Omega)$ is given by

$$
(u, v)_{AC^{1,2}(\Omega)} = u(0)v(0) + \int_0^T \dot{u}(t)\dot{v}(t)\, dt.
$$

Now, let $h := (h_x, h_{y_1}, h_{y_2}, h_z) \in L^2(\Omega) \times AC^{1,2}(\Omega) \times AC^{1,2}(\Omega) \times L^2(\Omega)^{2(N-1)}$ be arbitrary but fixed and set $\theta^\nu = (\theta_1^\nu, \theta_2^\nu, \theta_3^\nu) \in W$. In order to compute the adjoint of the partial derivatives of $g^\nu$ with respect to $(u^\nu, y_1^\nu, y_2^\nu, z^\nu)$, we calculate

$$
\begin{aligned}
&(\theta^\nu, D_{(v^\nu, z^\nu)} g^\nu(v^\nu, y^{-\nu}, z^\nu) h)_V = \\
&\quad = \left(\theta^\nu, \left(h_{y_1} - \int_0^\cdot h_{y_2}(t)\, dt,\ h_{y_2} - \int_0^\cdot h_x(t)\, dt,\ H_1 h_{y_1} + \mathcal{M} h_z\right)\right)_V \\
&\quad = \left(\theta_1^\nu, h_{y_1} - \int_0^\cdot h_{y_2}(t)\, dt\right)_{AC^{1,2}(\Omega)} + \left(\theta_2^\nu, h_{y_2} - \int_0^\cdot h_u(t)\, dt\right)_{AC^{1,2}(\Omega)} \\
&\qquad + (\theta_3^\nu, H_1 h_{y_1} + \mathcal{M} h_z)_{L^2(\Omega)},
\end{aligned}
$$

where $H_1$ denotes the first column of the matrix $H$. Since the second one is zero by definition (see Example 6.1.2), the derivative of the collision avoidance constraint (6.5) with respect to $y_2^\nu$ vanishes. We continue by applying the definitions of the scalar products and obtain

$$
\begin{aligned}
&(\theta^\nu, D_{(v^\nu,z^\nu)} g^\nu(x^\nu, y^{-\nu}, z^\nu)h)_V = \\
&= \theta_1^\nu(0) h_{y_1}(0) + \theta_2^\nu(0) h_{y_2}(0) + \int_0^T \dot\theta_1^\nu(t)(\dot h_{y_1}(t) - h_{y_2}(t)) + \dot\theta_2^\nu(t)(\dot h_{y_2}(t) - h_x(t))\; dt \\
&\qquad + \int_0^T \theta_3^\nu(t)(H_1 h_{y_1}(t) + \mathcal{M} h_z(t))\; dt.
\end{aligned}
$$

Since we have $h_{y_1}, h_{y_2} \in AC^{1,2}(\Omega)$, the identity $h_{y_j}(\tau) = h_{y_j}(0) + \int_{t_0}^{\tau} \dot h_{y_j}(t)\, dt$ for $j \in \{1,2\}$ is valid and by Lemma 6.3.2, we further conclude that

$$
\begin{aligned}
&(\theta^\nu, D_{(v^\nu,z^\nu)} g^\nu(x^\nu, y^{-\nu}, z^\nu)h)_V = \\
&= \theta_1^\nu(0) h_{y_1}(0) + \int_0^T \dot\theta_1^\nu(t) \left( \dot h_{y_1}(t) - h_{y_2}(0) - \int_0^t \dot h_{y_2}(\tau)\; d\tau \right) dt + \theta_2^\nu(0) h_{y_2}(0) \\
&\; + \int_0^T \dot\theta_2^\nu(t)(\dot h_{y_2}(t) - h_x(t)) dt + \int_0^T \theta_3^\nu(t) \left( H_1 \left( h_{y_1}(0) + \int_0^t \dot h_{y_1}(\tau)\; d\tau \right) + \mathcal{M} h_z(t) \right) dt \\
&= h_{y_1}(0) \left( \theta_1^\nu(0) + \int_0^T H_1^* \theta_3^\nu(t)\; dt \right) + \int_0^T \left( \dot h_{y_1}(t) \dot\theta_1^\nu(t) + H_1^* \theta_3^\nu(t) \int_0^t \dot h_{y_1}(\tau)\; d\tau \right) dt \\
&\quad + h_{y_2}(t_0) \left( \theta_2^\nu(0) - \int_0^{t_F} \dot\theta_1^\nu(t)\; dt \right) + \int_0^T \left( \dot h_{y_2}(t) \dot\theta_2^\nu(t) - \dot\theta_1^\nu(t) \int_0^t \dot h_{y_2}(\tau)\; d\tau \right) dt \\
&\quad + \int_0^T \mathcal{M}^* \theta_3^\nu(t) h_z(t) - \dot\theta_2^\nu(t) h_x(t)\; dt \\
&= h_{y_1}(0) \left( \theta_1^\nu(0) + \int_0^T H_1^* \theta_3^\nu(t)\; dt \right) + \int_0^T \dot h_{y_1}(t) \left( \dot\theta_1^\nu(t) + \int_t^T H_1^* \theta_3^\nu(\tau)\; d\tau \right) dt \\
&\qquad + h_{y_2}(0) \left( \theta_2^\nu(0) - \int_0^T \dot\theta_1^\nu(t)\; dt \right) + \int_0^T \dot h_{y_2}(t) \left( \dot\theta_2^\nu(t) - \int_t^T \dot\theta^\nu(\tau)\; d\tau \right) dt \\
&\qquad + \int_0^T \mathcal{M}^* \theta_3^\nu(t) h_z(t)\; dt - \int_0^T \dot\theta_2^\nu(t) h_x(t)\; dt.
\end{aligned}
$$

Hence, the adjoints are explicitly given by

$$
\begin{aligned}
&D_{v_1^\nu} g^\nu(v^\nu, y^{-\nu}, z^\nu)^* \theta^\nu = -\dot{\theta}_2^\nu, \\
&D_{v_2^\nu} g^\nu(v^\nu, y^{-\nu}, z^\nu)^* \theta^\nu = \left(\theta_1^\nu(0) + \int_0^T H_1^* \theta_3^\nu(t)\, dt,\ \dot{\theta}_1^\nu + \int_\tau^T H_1^* \theta_3^\nu(t)\, dt\right), \\
&D_{v_3^\nu} g^\nu(v^\nu, y^{-\nu}, z^\nu)^* \theta^\nu = \left(\theta_2^\nu(0) - \int_0^T \dot{\theta}_1^\nu(t)\, dt,\ \dot{\theta}_2^\nu - \int_\tau^T \dot{\theta}_1^\nu(t)\, dt\right), \\
&D_{z^\nu} g^\nu(x^\nu, y^{-\nu}, z^\nu)^* \theta^\nu = \mathcal{M}^* \theta_3^\nu.
\end{aligned}
$$

Note that the representation in $\mathbb{R} \times L^2(\Omega)$ for the adjoints in $AC^{1,2}(\Omega)$ is used. Consequently, (6.17) and (6.18) are equivalent to

$$(\alpha_\nu \bar{x}^\nu - \dot{\theta}_2^\nu, x^\nu - \bar{x}^\nu)_{L^2(\Omega)} \geq 0 \qquad \forall x^\nu \in Q_\nu, \tag{6.19a}$$

$$\theta_2^\nu(0) - \int_0^T \dot{\theta}_1^\nu(t)\, dt = 0, \tag{6.19b}$$

$$\dot{\theta}_2^\nu(\tau) - \int_\tau^T \dot{\theta}_1^\nu(t)\, dt = 0 \quad \text{f.a.a. } \tau \in \Omega, \tag{6.19c}$$

$$\theta_1^\nu(0) + \int_0^T H_1^* \theta_3^\nu(t)\, dt - 1 = 0, \tag{6.19d}$$

$$\dot{\theta}_1^\nu(\tau) + \int_\tau^T H_1^* \theta_3^\nu(t)\, dt - 1 = 0 \quad \text{f.a.a. } \tau \in \Omega, \tag{6.19e}$$

$$0 \geq \theta_3^\nu \perp H\bar{y}^\nu + H_{\text{opp}}\bar{y}^{-\nu} + \mathcal{M}\bar{z}^\nu - c^\nu \geq 0 \tag{6.19f}$$

and for all $i = 1, \ldots, 2(N-1)$

$$
\begin{aligned}
\mu_i^\nu(\tau) &= 0 \quad \text{f.a.a. } \tau \in \mathcal{I}_{+0}^\nu(\bar{z}, i), \\
\mu_i^\nu(\tau) + (\mathcal{M}^* \theta_3^\nu)_i(\tau) + \beta_\nu (D^* D z^\nu)_i(\tau) &= 0 \quad \text{f.a.a. } \tau \in \mathcal{I}_{0-}^\nu(\bar{z}, i).
\end{aligned}
$$

With Lemma 6.3.1 we see that, if $\rho_1^\nu$ solves

$$\dot{\rho}_1^\nu = H_1^* \theta_3^\nu, \quad \rho_1^\nu(T) = 1,$$

then $\theta_1^\nu(\tau) := \rho_1^\nu(0) + \int_0^\tau \rho_1^\nu(t)\, dt$ satisfies (6.19d) and (6.19e). Similarly, we conclude that, if $\rho_2^\nu$ solves the ODE

$$\dot{\rho}_2^\nu = -\rho_1^\nu, \quad \rho_2^\nu(T) = 0,$$

then $\theta_2^\nu(\tau) := \rho_2^\nu(0) + \int_0^\tau \rho_2^\nu(t)\, dt$ satisfies (6.19b) and (6.19c). With the substitution $\theta^\nu := -\theta_3^\nu$, we end up with (6.15) and (6.16). Since $\nu$ was chosen arbitrarily, this proves the claim. □

In order to obtain a sequence that converges to an S-stationary Nash equilibrium of $\Gamma^{\text{AD}}_{\text{EPCC}}$, we use the Nemytskij operator $\Phi : H^1(\Omega)^{2(N-1)} \to L^2(\Omega)^{2(N-1)}$ defined for all $u \in H^1(\Omega)^{2(N-1)}$ and $j = 1, \dots, 2(N-1)$ by

$$\Phi_j(u)(\omega) := \phi(u_j(\omega), 1 - u_j(\omega)) \qquad \forall \omega \in \Omega,$$

where $\phi : \mathbb{R}^2 \to \mathbb{R}$ denotes an NCP-function, i.e.

$$\forall a, b \in \mathbb{R} : \quad \phi(a,b) = 0 \iff 0 \leq a \perp b \geq 0.$$

As a direct consequence, we see that

$$\Phi(z^\nu) = 0 \iff z^\nu \in Z_{\text{ad}} := \left\{ v \in [0,1]^{2(N-1)} \;\middle|\; v(\omega)(1 - v(\omega)) = 0 \quad \text{f.a.a. } \omega \in \Omega \right\}$$

is valid.

**Example 6.3.4.** *Prominent examples of NCP-functions are for instance the minimum function*

$$\phi_{\text{min}}(a,b) := \min(a,b) = a - (a-b)_+$$

*or the Fischer-Burmeister function*

$$\phi_{\text{FB}}(a,b) := \sqrt{a^2 + b^2} - a - b.$$

For this reason, let $\{\gamma_k\} \subseteq \mathbb{R}^+$ be a sequence such that $\gamma_k \to 0$ as $k \to \infty$ and consider the following auxiliary Nash equilibrium problem, denoted by $\Gamma^k_{\text{AD}}$, where player $\nu$ solves

$$\begin{aligned}
\min \quad & F^{k,\nu}(x^\nu, y^\nu, z^\nu) := -y_1^\nu(T) + \frac{\alpha_\nu}{2}\|x^\nu\|^2 + \frac{\beta_\nu}{2}\|z^\nu\|^2 + \frac{1}{2\gamma_k}\|\Phi(z^\nu)\|^2 \\
\text{w.r.t.} \quad & (x^\nu, y^\nu, z^\nu) \in L^2(\Omega) \times AC^{1,2}(\Omega)^2 \times H^1(\Omega)^{2(N-1)}, \\
\text{s.t.} \quad & x^\nu \in Q_\nu, \\
& \dot{y}^\nu = \begin{pmatrix} y_2^\nu \\ x^\nu \end{pmatrix}, \; y^\nu(0) = \begin{pmatrix} y_0^\nu \\ v_0^\nu \end{pmatrix}, \\
& H y^\nu + H_{\text{opp}} y^{-\nu} + \mathcal{M} z^\nu - c^\nu \geq 0.
\end{aligned} \tag{6.20}$$

The following result shows that any sequence of feasible points of $\Gamma^k_{\text{AD}}$ weakly converges to a feasible point of $\Gamma^{\text{AD}}_{\text{EPCC}}$.

**Theorem 6.3.5.** *Let $\{(x^k, y^k, z^k)\} \subseteq L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}$ be a sequence such that $(x^k, y^k, z^k)$ is a feasible point of $\Gamma^k_{\text{AD}}$ for all $k \in \mathbb{N}$. Moreover, assume that the*

*sequence* $\{F^{k,\nu}(x^{k,\nu}, y^{k,\nu}, z^{k,\nu})\}$ *is bounded from above for all* $\nu = 1, \ldots, N$.
*Then there exists* $(\bar{x}, \bar{y}, \bar{z}) \in L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}$ *such that*

$$
\begin{aligned}
(x^{k,\nu}, y^{k,\nu}, z^{k,\nu}) &\rightharpoonup (\bar{x}^\nu, \bar{y}^\nu, \bar{z}^\nu) \qquad && \text{in } L^2(\Omega) \times AC^{1,2}(\Omega)^2 \times H^1(\Omega)^{2(N-1)} \\
z^{k,\nu} &\to \bar{z}^\nu && \text{in } L^2(\Omega)^{2(N-1)}
\end{aligned}
$$

*for all* $\nu = 1, \ldots, N$ *and* $(\bar{x}, \bar{y}, \bar{z})$ *is a feasible point of* $\Gamma^{\text{AD}}_{\text{EPCC}}$.

*Proof.* We split the proof into two parts, i.e. the verification of the boundedness of $\{(x^k, y^k, z^k)\}$ and the feasibility of the weak limit point $(\bar{x}, \bar{y}, \bar{z})$.

- Boundedness of $\{(x^k, y^k, z^k)\}$:
  Since $Q$ is weakly sequentially compact by assumption, we obtain the boundedness of $\{x^k\}$ in $L^2(\Omega)^N$, which implies

  $$|y_2^{k,\nu}|^2_{AC^{1,2}(\Omega)} = |y_2^{k,\nu}(0)| + \|\dot{y}_2^{k,\nu}\|_{L^2(\Omega)} = |v_0^\nu| + \|x^{k,\nu}\|_{L^2(\Omega)} \leq c_2^\nu$$

  for some $c_2^\nu > 0$ and all $\nu = 1, \ldots, N$. The latter estimation directly leads to

  $$|y_1^{k,\nu}|^2_{AC^{1,2}(\Omega)} = |y_1^{k,\nu}(0)| + \|y_2^{k,\nu}\|_{L^2(\Omega)} \leq |y_0^\nu| + |y_2^{k,\nu}|_{AC^{1,2}(\Omega)} \leq |y_0^\nu| + c_2^\nu \leq c_1^\nu$$

  for some $c_1^\nu > 0$, where we used the embedding $AC^{1,2} \hookrightarrow L^2(\Omega)$ (see Theorem 2.1.20) for the second estimation.
  Next, observe that it holds

  $$C_1^\nu \leq -y_1^{k,\nu}(T) \leq F^{k,\nu}(x^{k,\nu}, y^{k,\nu}, z^{k,\nu}) \leq C_2^\nu$$

  for some constant $C_1^\nu, C_2^\nu \in \mathbb{R}$ and all $\nu = 1, \ldots, N$. Since the norm is bounded and coercive, the boundedness of $\{z^{k,\nu}\}$ in $H^1(\Omega)$ follows.
  Hence, there exists $(\bar{x}, \bar{y}, \bar{z})$ such that

  $$(x^k, y^k, z^k) \rightharpoonup (\bar{x}, \bar{y}, \bar{z}) \in L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}$$

  on a subsequence, which we denote the same.

- Feasibility of $(\bar{x}, \bar{y}, \bar{z})$:
  Since the feasible set of $\Gamma^k_{\text{AD}}$ is independent of $\gamma_k$ and comprised of (affine) linear constraints, which preserve the weak continuity and weakly sequentially closed sets that contain the weak limits, we conclude that $(\bar{u}, \bar{y}, \bar{z})$ is a feasible point of $\Gamma^k_{\text{AD}}$ for all $k \in \mathbb{N}$.

Thus, it remains to show the complementarity constraint. Therefore, observe that it holds
$$0 \leq \frac{1}{2}\|\Phi(z^{k,\nu})\| \to 0$$
as $\gamma_k \to 0$ by the boundedness of $\{F^{k,\nu}(x^{k,\nu}, y^{k,\nu}, z^{k,\nu})\}$. Since $z^k \to \bar{z}$ in $L^2(\Omega)$ by the compact embedding of $H^1(\Omega)$ in $L^2(\Omega)$, we obtain by the observations above that $\Phi(\bar{z}^\nu) = 0$ and hence, by definition of $\Phi$ that $(\bar{x}, \bar{y}, \bar{z})$ is a feasible point of $\Gamma^{\text{AD}}_{\text{EPCC}}$.

□

At this point, we emphasize once more that the strong convergence of $z^{k,\nu}$ in $L^2(\Omega)$ was required in order to get the feasibility of $(\bar{x}, \bar{y}, \bar{z})$. We continue by finding an appropriate sequence that converges to an S-stationary Nash equilibrium.

**Theorem 6.3.6.** *Let $(x^k, y^k, z^k) \in L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}$ be a Nash equilibrium of $\Gamma^k_{\text{AD}}$ and assume that for all $(w_1, w_2, w_3) \in AC^{1,2}(\Omega)^N \times AC^{1,2}(\Omega)^N \times L^2(\Omega)^{4N(N-1)}$, there exist $(\eta_z, \eta_\theta) \in \mathbb{R}^{2N,+}_0$ and $v \in L^2(\Omega)^{4(N-1),+}_0$ such that it holds*

$$\begin{aligned} \eta_z^\nu \mathcal{M} z^\nu(\tau) - v^\nu(\tau) = w_3^\nu(\tau) - H(w_1^\nu(\tau) + \int_0^\tau w_2^\nu(t)\, dt) + \eta_z^\nu \mathcal{M} z^{k,\nu}(\tau) \\ + \eta_\theta^\nu (H_1 y^{k,\nu}(\tau) + H_{\text{opp}} y^{k,-\nu}(\tau) + \mathcal{M} z^{k,\nu}(\tau) - c^\nu) \end{aligned} \tag{6.21}$$

*f.a.a. $\tau \in \Omega$ and for all $\nu = 1, \dots, N$.*
*Then $(x^k, y^k, z^k)$ is a stationary Nash equilibrium of $\Gamma^k_{\text{AD}}$, i.e. there exist multipliers*

$$(\rho^k, \theta^k) \in AC^{1,2}(\Omega)^{2N} \times L^2(\Omega)^{4N(N-1)}$$

*such that for all $\nu = 1, \dots, N$ it holds*

$$(\alpha_\nu x^{k,\nu} - \rho_2^{k,\nu}, x^\nu - x^{k,\nu}) \geq 0 \quad \forall x^\nu \in Q_\nu, \tag{6.22a}$$

$$\begin{aligned} \dot{y}_1^{k,\nu} - y_2^{k,\nu} = 0, \qquad y_1^{k,\nu}(0) = y_0^\nu, \\ \dot{y}_2^{k,\nu} - x^{k,\nu} = 0, \qquad y_2^{k,\nu}(0) = v_0^\nu, \end{aligned} \tag{6.22b}$$

$$\begin{aligned} \dot{\rho}_1^{k,\nu} + H_1^* \theta^{k,\nu} = 0, \qquad \rho_1^{k,\nu}(T) = 1, \\ \dot{\rho}_2^{k,\nu} + \rho_1^{k,\nu} = 0, \qquad \rho_2^{k,\nu}(T) = 0, \end{aligned} \tag{6.22c}$$

$$\frac{1}{\gamma_k} \Phi'(z^{k,\nu}) \Phi(z^{k,\nu}) + \beta_\nu (z^{k,\nu} + D^* D z^{k,\nu}) - \mathcal{M}^* \theta^{k,\nu} = 0 \tag{6.22d}$$

*and*

$$0 \leq \theta^{k,\nu} \perp H y^{k,\nu} + H_{\text{opp}} y^{k,-\nu} + \mathcal{M} z^{k,\nu} - c^\nu \geq 0. \tag{6.23}$$

*Proof.* Let $(x^k, y^k, z^k)$ be a Nash equilibrium of $\Gamma^k_{\text{AD}}$ and define for all $\nu = 1, \ldots, N$ the mapping

$$g^\nu(x^\nu, y, z^\nu) := \begin{pmatrix} y_1^\nu(\cdot) - y_0^\nu - \int_0^\cdot y_2^\nu(t)\, dt \\ y_2^\nu(\cdot) - v_0^\nu - \int_0^\cdot x^\nu(t)\, dt \\ Hy^\nu + H_{\text{opp}} y^{-\nu} + \mathcal{M} z^\nu - c^\nu \end{pmatrix}.$$

Now, it is not difficult to show that (6.21) equals KRZCQ (see (2.18)) for all $\nu = 1, \ldots, N$. For this reason, let $(w_1, w_2, w_3) \in AC^{1,2}(\Omega)^N \times AC^{1,2}(\Omega)^N \times L^2(\Omega)^{4N(N-1)}$ be arbitrary but fixed and observe that KRZCQ is given by the conditions

$$\begin{aligned} h_{y_1}(\tau) - \int_0^\tau h_{y_2}(t)\, dt &= w_1(\tau), \\ h_{y_2}(\tau) - \eta_x^\nu \int_0^\tau h_x(t) - x^{k,\nu}(t)\, dt &= w_2(\tau) \end{aligned}$$

and

$$\begin{aligned} &H_1 h_{y_1}(\tau) + \eta_z^\nu \mathcal{M}(h_z(\tau) - z^{k,\nu}(\tau)) - v^\nu(\tau) \\ &\qquad + \eta_\theta^\nu (H_1 y^{k,\nu}(\tau) + H_{\text{opp}} y^{k,-\nu}(\tau) + \mathcal{M} z^{k,\nu}(\tau) - c^\nu) = w_3(\tau) \end{aligned}$$

f.a.a. $\tau \in \Omega$, where we additionally have the following requirements:

- $\eta_x, \eta_z, \eta_\theta \in \mathbb{R}^{N,+}$;
- $(h_x, h_{y_1}, h_{y_2}, h_z) \in Q \times AC^{1,2}(\Omega) \times AC^{1,2}(\Omega) \times H^1(\Omega)^{2(N-1)}$ and $v \in L^2(\Omega)_0^{4N(N-1),+}$;

Now, set $\eta_x \equiv 0$ and observe that (6.21) is equivalent to KRZCQ. Hence, there exists multipliers $(\rho^k, \theta^k, \tilde{\mu}^k, \tilde{\lambda}^k)$ such that (6.22) and (6.23) are satisfied for all $\nu = 1, \ldots, N$. □

In this context, $\Phi'(z^\nu)$ denotes the derivative of $\Phi(z^\nu)$. Observe that this mapping is continuous since the biactive set $\mathcal{I}_{00}^\nu$ is empty (see proof of Corollary 6.2.3). Throughout the remainder of this work and if not stated otherwise, we set $\phi(\cdot,\cdot) := \phi_{\text{FB}}(\cdot,\cdot)$ (see Example 6.3.4). In particular, the derivative is given by

$$\Phi'(z^\nu) = \frac{2z^\nu - 1}{\sqrt{(z^\nu)^2 + (1 - z^\nu)^2}}.$$

Now, we can show the following convergence result.

**Theorem 6.3.7.** *Let $\{(x^k, y^k, z^k)\} \subseteq L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}$ be a sequence, where $(x^k, y^k, z^k)$ is a stationary Nash equilibrium of $\Gamma^k_{\text{AD}}$ for all $k \in \mathbb{N}$ and $\frac{1}{\gamma_k}\|\Phi'(z^{k,\nu})\Phi(z^{k,\nu})\|$ be bounded for all $\nu = 1, \dots, N$.*
*Then there exists a tuple $(\bar{x}, \bar{y}, \bar{z}, \bar{\rho}, \bar{\theta}, \bar{\mu})$ such that for all $\nu = 1, \dots, N$ it holds*

$$\begin{aligned} (x^{k,\nu}, y^{k,\nu}, z^{k,\nu}) &\to (\bar{x}^\nu, \bar{y}^\nu, \bar{z}^\nu) \qquad && \text{in } L^2(\Omega) \times AC^{1,2}(\Omega)^2 \times L^2(\Omega)^{2(N-1)}, \\ (\rho^{k,\nu}, \theta^{k,\nu}) &\rightharpoonup (\bar{\rho}^\nu, \bar{\theta}^\nu) && \text{in } AC^{1,2}(\Omega)^2 \times L^2(\Omega)^{4(N-1)} \end{aligned}$$

*on a subsequence and $(\bar{x}, \bar{y}, \bar{z})$ is an S-stationary Nash equilibrium of $\Gamma^{\text{AD}}_{\text{EPCC}}$.*

Before we give a proof, we first consider an auxiliary result.

**Lemma 6.3.8.** *Let $\{(\theta^k, y^k, z^k)\} \subseteq L^2(\Omega)^{4(N-1)} \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}$ be a sequence such that* (6.23) *is valid for all $k \in \mathbb{N}$ and $\nu = 1, \dots, N$. Then $\{\theta^k\}$ is bounded in $L^2(\Omega)$.*

*Proof.* Let $\nu = 1, \dots, N$ be arbitrary but fixed and assume that $\{\theta^{k,\nu}\}$ is not bounded. Since (6.23) is satisfied for all $k \in \mathbb{N}$, this implies

$$\lim_{k\to\infty} Hy^{k,\nu} + H_{\text{opp}}y^{k,-\nu} + \mathcal{M}z^{k,\nu} - c^\nu = 0.$$

By the continuity of the latter term, we find a point $(y^{\tilde{k},\nu}, y^{\tilde{k},-\nu}, z^{\tilde{k},\nu})$ such that

$$(y^{\tilde{k},\nu}, y^{\tilde{k},-\nu}, z^{\tilde{k},\nu}) = \lim_{k\to\infty} (y^{k,\nu}, y^{k,-\nu}, z^{k,\nu})$$

and

$$Hy^{\tilde{k},\nu} + H_{\text{opp}}y^{\tilde{k},-\nu} + \mathcal{M}z^{\tilde{k},\nu} - c^\nu = 0. \tag{6.24}$$

W.l.o.g. we continue by focusing on the setting for two players. Then we deduce by (6.24) that $(y_1^{\tilde{k},\nu}, y_2^{\tilde{k},\nu}, z_1^{\tilde{k},\nu}, z_2^{\tilde{k},\nu})$ with $\nu = 1, 2$ (see (6.4a)) is a solution of the linear system

$$\begin{pmatrix} 1 & 0 & M & M \\ 0 & 1 & M & -M \\ -1 & 0 & -M & M \\ 0 & -1 & -M & -M \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} c_1^\nu + \hat{y}_{12} \\ c_2^\nu + \hat{y}_{21} \\ c_3^\nu - \hat{y}_{12} \\ c_4^\nu - \hat{y}_{21} \end{pmatrix},$$

which is equivalent to

$$\begin{pmatrix} 1 & 0 & M & 0 \\ 0 & 1 & M & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\beta_1 - \beta_3) \\ \beta_2 + \frac{1}{2}(\beta_1 + \beta_3) \\ \frac{1}{2}(\beta_1 + \beta_3) \\ \beta_1 + \beta_2 + \beta_3 + \beta_4 \end{pmatrix},$$

with $\beta_1^\nu := c_1^\nu + \hat{y}_{12}$, $\beta_2^\nu := c_2^\nu + \hat{y}_{21}$, $\beta_3^\nu := c_3^\nu - \hat{y}_{12}$ and $\beta_4^\nu := c_4^\nu - \hat{y}_{21}$.

Now, the last entry on the right hand side is equivalent (see (6.6)) to $4(\delta - M)$, which is by the assumptions on $\delta$ and $M$ in Section 6.1 nonzero. Hence, the linear system admits no solution and thus, $\{\theta^{k,\nu}\}$ is bounded. □

Now, we give the proof of Theorem 6.3.7.

*Proof of Theorem 6.3.7.* Similar as in the proof of Theorem 6.3.5, we start by showing that the sequence $\{(x^k, y^k, z^k, \rho^k, \theta^k)\}$ is bounded and continue with the verification of the S-stationary conditions (6.15) and (6.16) of the weak limit point.

- Boundedness of $\{(x^k, y^k, z^k, \rho^k, \theta^k)\}$:
  Since we have $\{x^k\} \subseteq Q$, we obtain the boundedness of $\{x^k\}$ in $L^2(\Omega)$ that directly implies the boundedness of $\{y^k\}$ in $AC^{1,2}(\Omega)$ (see the proof of Theorem 6.3.5).
  Lemma 6.3.8 showed the boundedness of $\{\theta^k\}$, which can be used to derive the result for $\{\rho^k\}$ in the same way we verified the boundedness of $y^k$. In this context, observe that the adjoint equation (6.22c) can be written as an initial value problem by using a substitution rule.
  Hence, it remains to prove the boundedness of $\{z^k\}$. Therefore, we have

$$\begin{aligned}
\beta_\nu \|z^{k,\nu}\|^2_{H^1(\Omega)} &= \beta_\nu(\|z^{k,\nu}\|^2_{L^2(\Omega)} + (D^*Dz^{k,\nu}, z^{k,\nu})_{L^2(\Omega)}) \\
&= |(M^*\theta^{k,\nu} - \frac{1}{\gamma_k}\Phi'(z^{k,\nu})\Phi(z^{k,\nu}), z^{k,\nu})_{L^2(\Omega)}| \\
&\leq (\|M^*\|\|\theta^{k,\nu}\| + \frac{1}{\gamma_k}\|\Phi'(z^{k,\nu})\Phi(z^{k,\nu})\|)\|z^{k,\nu}\|_{L^2(\Omega)}.
\end{aligned}$$

  Since $\|z^{k,\nu}\|_{L^2(\Omega)} \leq \|z^{k,\nu}\|_{H^1(\Omega)}$ holds by definition, we obtain by assumption and the boundedness of $\{\theta^{k,\nu}\}$ that $\{z^{k,\nu}\}$ is bounded in $H^1(\Omega)$.
  Summarizing, we find a tuple $(\bar{x}, \bar{y}, \bar{z}, \bar{\rho}, \bar{\theta})$ such that

$$\begin{aligned}
(x^k, y^k, z^k) &\rightharpoonup (\bar{x}, \bar{y}, \bar{z}) \text{ in } L^2(\Omega)^N \times AC^{1,2}(\Omega)^{2N} \times H^1(\Omega)^{2N(N-1)}, \\
(\rho^k, \theta^k) &\rightharpoonup (\bar{\rho}, \bar{\theta}) \text{ in } AC^{1,2}(\Omega)^{2N} \times L^2(\Omega)^{4N(N-1)}.
\end{aligned}$$

- Validity of S-stationarity conditions (6.15) and (6.16):
  Due to the equivalence of (6.22a) and

$$x^{k,\nu} = \Pi_{Q_\nu}(\alpha_\nu^{-1}\rho_2^{k,\nu})$$

(see e.g. [35, Lemma 1.11]), we estimate

$$\begin{aligned}\|x^{k,\nu} - \Pi_{Q_\nu}(\alpha_\nu^{-1}\bar{\rho}_2^\nu)\|_{L^2(\Omega)} &= \|\Pi_{Q_\nu}(\alpha_\nu^{-1}\rho_2^{k,\nu}) - \Pi_{Q_\nu}(\alpha_\nu^{-1}\bar{\rho}_2^\nu)\|_{L^2(\Omega)} \\ &\leq \alpha_\nu^{-1}\|\rho_2^{k,\nu} - \bar{\rho}_2^\nu\|_{L^2(\Omega)}.\end{aligned}$$

Since $\rho_2^{k,\nu} \to \bar{\rho}_2^\nu$ in $L^2(\Omega)$ by the compact embedding $AC^{1,2}(\Omega) \hookrightarrow\hookrightarrow L^2(\Omega)$, we obtain that

$$x^{k,\nu} \to \Pi_{Q_\nu}(\alpha_\nu^{-1}\bar{\rho}_2^\nu) \text{ in } L^2(\Omega).$$

Due to the uniqueness of the limit point, we conclude that $\bar{x}^\nu = \Pi_{Q_\nu}(\alpha_\nu^{-1}\bar{\rho}_2^\nu)$, $x^{k,\nu}$ converges strongly to $\bar{x}^\nu$ in $L^2(\Omega)$ and $\bar{x}^\nu$ satisfies (6.15a).
Since (6.22b) and (6.22c) are comprised of linear and continuous mappings, which are in particular weakly continuous, the limit points $\bar{y}^\nu$ and $\bar{\rho}^\nu$ satisfy (6.15b) and (6.15c). Moreover, it follows by the strong convergence of $x^{k,\nu}$ in $L^2(\Omega)$ that

$$\begin{aligned}|y_2^{k,\nu} - \bar{y}_2^\nu|_{AC^{1,2}(\Omega)} &= |y_2^{k,\nu}(0) - \bar{y}_2^\nu(0)| + \|x^{k,\nu} - \bar{x}^\nu\|_{L^2(\Omega)} \\ &= |v_0^\nu - v_0^\nu| + \|x^{k,\nu} - \bar{x}^\nu\|_{L^2(\Omega)} \to 0.\end{aligned}$$

Hence, it holds $y_2^{k,\nu} \to \bar{y}_2^\nu$ in $AC^{1,2}(\Omega)$. Similarly, one can show that $y_1^{k,\nu} \to \bar{y}_1^\nu$ in $AC^{1,2}(\Omega)$ is valid. By the same argument as for (6.22b) and (6.22c), we see that the limit point satisfies (6.23), where the orthogonality follows by the strong convergence of $y^{k,\nu}$ and $z^{k,\nu}$ in $L^2(\Omega)^2$ and $L^2(\Omega)^{2(N-1)}$, respectively. Thus, (6.15d) is valid.
In order to show (6.16), first define

$$\begin{aligned}\mu^{k,\nu} &:= \frac{z^{k,\nu} - 1}{\gamma_k\sqrt{(z^{k,\nu})^2 + (1 - z^{k,\nu})^2}}\Phi(z^{k,\nu}) \\ \lambda^{k,\nu} &:= \frac{z^{k,\nu}}{\gamma_k\sqrt{(z^{k,\nu})^2 + (1 - z^{k,\nu})^2}}\Phi(z^{k,\nu})\end{aligned}$$

and observe that (6.22d) is equivalent to

$$\mu^{k,\nu} + \lambda^{k,\nu} - M^*\theta^{k,\nu} + \beta_\nu(z^{k,\nu} + D^*Dz^{k,\nu}) = 0$$

for all $\nu = 1, \dots, N$. Hence, the assertion (6.16) follows by the strong convergence of $z^{k,\nu}$ in $L^2(\Omega)$ and the discussion above.

□

In the following section we concentrate on numerical solution methods for computing an S-stationary Nash equilibrium of $\Gamma_{\text{EPCC}}^{\text{AD}}$.

### 6.3.2. Numerical Analysis and Examples

In order to solve (6.22) and (6.23) numerically, we assume that for all $\nu = 1, \dots, N$ the set $Q_\nu$ is given by

$$Q_\nu := \{x^\nu \in L^2(\Omega) \mid a_\nu \leq x^\nu \leq b_\nu\}$$

with $a_\nu, b_\nu \in L^2(\Omega)$ such that $a_\nu(\tau) \leq b_\nu(\tau)$ f.a.a. $\tau \in \Omega = (0, T)$ and define

$$G^{k,\nu}(y, z, \theta) := \begin{pmatrix} \frac{1}{\gamma_k}\Phi'(z^\nu)\Phi(z^\nu) + \beta_\nu(z^\nu + D^*Dz^\nu) - \mathcal{M}^*\theta^\nu \\ \Phi_{\text{min}}(\theta^\nu, Hy^\nu + H_{\text{opp}}y^{-\nu} + \mathcal{M}z^\nu - c^\nu) \end{pmatrix},$$

where $\Phi_{\text{min}}$ denotes the NCP-function induced by $\phi_{\text{min}}$.

**Algorithm 4** Numerical computation of an S-stationary Nash equilibrium of $\Gamma^{\text{AD}}_{\text{EPCC}}$.

**Input:** $N \in \mathbb{N}$, $\gamma_0 > 0$, tol and $(x^0, y^0, z^0, \rho^0, \theta^0, \tilde{\lambda}^0, \mu^0)$.
Set $k = 0$.
**while** $(x^k, y^k, z^k)$ is no S-stationary Nash equilibrium **do**
    Compute $y^{k+1,\nu}$ as the solution of (6.22b) w.r.t. $x^{k,\nu}$ for all $\nu = 1, \dots, N$.
    **for all** $\nu = 1, \dots, N$ **do**
        Find $(z^{k+1,\nu}, \theta^{k+1,\nu})$ s.t. $G^{k,\nu}(y^{k+1}, z^{k+1,\nu}, \theta^{k+1,\nu}) = 0$.
        Compute $\rho^{k+1,\nu}$ as the solution of (6.22c) w.r.t. $\theta^{k+1,\nu}$.
        Set $x^{k+1,\nu} = \Pi_{Q_\nu}(\alpha_\nu^{-1}\rho_2^{k+1,\nu})$.
    **end for**
    Choose $\gamma_{k+1} < \gamma_k$ and set $k = k + 1$.
**end while**
**return** S-stationary Nash equilibrium $(\bar{x}, \bar{y}, \bar{z})$.

In order to compute an S-stationary Nash equilibrium, we consider Algorithm 4. Under the hypothesis that the assumptions in Theorem 6.3.7 are valid, we know that a sequence generated by Algorithm 4 converges towards an S-stationary Nash equilibrium of $\Gamma^{\text{AD}}_{\text{EPCC}}$. Obviously, the most crucial step is the computation of

$$G^{k,\nu}(y^{k+1}, z^{k+1,\nu}, \theta^{k+1,\nu}) = 0, \tag{6.25}$$

since $G^{k,\nu}$ is in general not smooth but semismooth. However, notice that $G^{k,\nu}$ is not Newton differentiable (see Theorem 5.4.1) since $\Phi_{\text{min}}(\cdot, Hy^\nu + H_{\text{opp}}y^{-\nu} + \mathcal{M}z^\nu - c^\nu)$ maps from $L^2(\Omega)$ to $L^2(\Omega)$. On a discrete level, this is not an issue. Therefore, set

$$0 = t_1 < \cdots < t_j := \frac{j}{m} < \cdots < t_m = T$$

for some $m \in \mathbb{N}$, i.e. we consider an equidistant discretization of $\Omega$. Furthermore, we introduce for all $\nu = 1, \ldots, N$ the vectors $\mathbf{x}^\nu, \mathbf{y}_1^\nu, \mathbf{y}_2^\nu, \boldsymbol{\rho}_1^\nu, \boldsymbol{\rho}_2^\nu \in \mathbb{R}^m$ and $\mathbf{y}^\nu, \boldsymbol{\rho}^\nu \in \mathbb{R}^{2m}$, which are exemplarily given by

$$\mathbf{y}_{1,j}^\nu := y_1^\nu(t_j) \quad \text{and} \quad \begin{pmatrix} \mathbf{y}_{2j-1}^\nu \\ \mathbf{y}_{2j}^\nu \end{pmatrix} := \begin{pmatrix} y_1^\nu(t_j) \\ y_2^\nu(t_j) \end{pmatrix}.$$

and the matrices $\mathbf{z}^\nu, \tilde{\boldsymbol{\lambda}}^\nu \in \mathbb{R}^{2(N-1)\times m}$, $\boldsymbol{\theta}^\nu \in \mathbb{R}^{4(N-1)\times m}$ in a similar way with components

$$\mathbf{z}_{ij}^\nu := z_i^\nu(t_j), \quad \boldsymbol{\theta}_{ij}^\nu := \theta_i^\nu(t_j), \quad \tilde{\boldsymbol{\lambda}}_{ij}^\nu := \tilde{\lambda}_i^\nu(t_j).$$

Then we define

$$G_h^{k,\nu} : \mathbb{R}^{2m} \times \cdots \times \mathbb{R}^{2m} \times \mathbb{R}^{2(N-1)\times m} \times \mathbb{R}^{4(N-1)\times m} \to \mathbb{R}^{6(N-1)\times m}$$

by

$$G_h^{k,\nu}(\mathbf{y}, \mathbf{z}^\nu, \boldsymbol{\theta}^\nu) := \begin{pmatrix} \frac{1}{\gamma_k}\phi'(\mathbf{z}^\nu)\phi(\mathbf{z}^\nu) + \beta_\nu(\mathbf{z}^\nu + D_h^T D_h \mathbf{z}^\nu) - \mathcal{M}^*\boldsymbol{\theta}^\nu \\ \phi_{\min}(\boldsymbol{\theta}^\nu, H\mathbf{y}^\nu + H_{\text{opp}}\mathbf{y}^{-\nu} + \mathcal{M}\mathbf{z}^\nu - \mathbf{c}^\nu) \end{pmatrix},$$

where $\phi(X_1)$ and $\phi_{\min}(X_2, X_3)$ have to be understood componentwise for matrices $X_1 \in \mathbb{R}^{2(N-1)\times m}$ and $X_2, X_3 \in \mathbb{R}^{4(N-1)\times m}$. Moreover, $D_h$ denotes the discrete difference operator. Hence, the discretization of (6.25) is given by

$$G_h^{k,\nu}(\mathbf{y}, \mathbf{z}^\nu, \boldsymbol{\theta}^\nu) = 0. \tag{6.26}$$

In order to discretize the state and adjoint equation, we use the classical Euler method[18] (see e.g. [8] for more details). Due to the structure of both equations, i.e. the initial and the end point condition, we apply the explicit Euler for (6.22b) and the implicit Euler for (6.22c). Therefore, we set $h := m^{-1}$ and observe that the former approach leads for all $j = 1, \ldots, m-1$ to

$$\frac{1}{h}\begin{pmatrix} \mathbf{y}_{1,j+1}^\nu - \mathbf{y}_{1,j}^\nu \\ \mathbf{y}_{2,j+1}^\nu - \mathbf{y}_{2,j}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{y}_{2,j}^\nu \\ \mathbf{x}_j^\nu \end{pmatrix}, \tag{6.27a}$$

$$\begin{pmatrix} \mathbf{y}_{1,1}^\nu \\ \mathbf{y}_{2,1}^\nu \end{pmatrix} = \begin{pmatrix} y_0^\nu \\ v_0^\nu \end{pmatrix}. \tag{6.27b}$$

[18] For an arbitrary continuous mapping $f : \bar{\Omega} \times H^1(\Omega) \to L^2(\Omega), (\tau, y) \to f(\tau, y)$ the explicit Euler method of the ODE $\dot{y} = f(\tau, y)$ with initial value $y(t_0) = y_0$ is given recursively by $y_{k+1} = y_k + h_k f(t_k, y_k)$ and $y_1 = y_0$ with $h_k = t_{k+1} - t_k$ and $y_k = y(t_k)$. Similarly, the implicit Euler method is defined by $y_{k+1} = y_k + h_k f(t_{k+1}, y_{k+1})$.

By rearranging (6.27a), we obtain

$$\underbrace{\begin{pmatrix} -1 & -h & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} \mathbf{y}^{\nu}_{1,j} \\ \mathbf{y}^{\nu}_{2,j} \\ \mathbf{y}^{\nu}_{1,j+1} \\ \mathbf{y}^{\nu}_{2,j+1} \end{pmatrix} - \underbrace{\begin{pmatrix} 0 \\ h \end{pmatrix}}_{=:b} \mathbf{x}^{\nu}_j = 0,$$

which can be extended to

$$A_y \mathbf{y}^{\nu} - B_x \mathbf{x}^{\nu} - \mathbf{y}^{\nu}_0 = 0. \tag{6.28}$$

In this context, $A_y \in \mathbb{R}^{2m\times 2m}$ and $B_x \in \mathbb{R}^{2m\times m}$ arise via blockwise concatenation of $A$ and $b$, respectively, under consideration of (6.27b) that is contained in $\mathbf{y}^{\nu}_0 \in \mathbb{R}^{2m}$.
Similar results can be expected for the adjoint equation (6.22b). Due to its structure, we use the implicit Euler method, which leads for all $j = 1, \dots, m-1$ to the approximation

$$\frac{1}{h} \begin{pmatrix} \boldsymbol{\rho}^{\nu}_{1,j+1} - \boldsymbol{\rho}^{\nu}_{1,j} \\ \boldsymbol{\rho}^{\nu}_{2,j+1} - \boldsymbol{\rho}^{\nu}_{2,j} \end{pmatrix} = \begin{pmatrix} -H_1^T \boldsymbol{\theta}^{\nu}_{j+1} \\ -\boldsymbol{\rho}^{\nu}_{1,j+1} \end{pmatrix}, \tag{6.29a}$$

$$\begin{pmatrix} \boldsymbol{\rho}^{\nu}_{1,m} \\ \boldsymbol{\rho}^{\nu}_{2,m} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{6.29b}$$

In particular, notice that we choose the implicit approach, since we have to solve the equation backwards. Now, (6.29a) can be written as

$$\begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & h & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\rho}^{\nu}_{1,j} \\ \boldsymbol{\rho}^{\nu}_{2,j} \\ \boldsymbol{\rho}^{\nu}_{1,j+1} \\ \boldsymbol{\rho}^{\nu}_{2,j+1} \end{pmatrix} + \begin{pmatrix} h \\ 0 \end{pmatrix} H_1^T \boldsymbol{\theta}^{\nu}_{j+1} = 0,$$

and hence, we derive the system

$$A_\rho \boldsymbol{\rho}^{\nu} + \underbrace{B_\theta \begin{pmatrix} H_1^T \boldsymbol{\theta}^{\nu}_1 \\ \vdots \\ H_1^T \boldsymbol{\theta}^{\nu}_M \end{pmatrix}}_{=:D_\theta[\boldsymbol{\theta}^{\nu}]} + \boldsymbol{\rho}_m = 0 \tag{6.30}$$

by analogous steps as for the state equation. In this case, we have block matrices $A_\rho \in \mathbb{R}^{2m\times 2m}$, $B_\theta \in \mathbb{R}^{2m\times m}$ that again follow by a suitable concatenation and a vector $\boldsymbol{\rho}_m \in \mathbb{R}^{2m}$ that contains (6.29b).
In order to solve (6.26) numerically, it is needed to either transfer $G^{k,\nu}_h$ into a pure vector

valued function in advance or to consider the condition for each column $j = 1, \dots, m$ separately. During the remainder of this section, we focus on the latter approach and fix $j \in \{1, \dots, m\}$. Then the Newton derivative of $G_h^{k,\nu}$ with respect to the jth column of $(\mathbf{z}^\nu, \boldsymbol{\theta}^\nu)$, i.e. $(\mathbf{z}^\nu_{\cdot j}, \boldsymbol{\theta}^\nu_{\cdot j})$, is given by

$$\begin{aligned} &DG_h^{k,\nu}(\mathbf{y}_{\cdot j}, \mathbf{z}^\nu_{\cdot j}, \boldsymbol{\theta}^\nu_{\cdot j}) \\ &\quad = \begin{pmatrix} \frac{1}{\gamma_k}(\phi''(\mathbf{z}^\nu_{\cdot j})\phi(\mathbf{z}^\nu_{\cdot j}) + \phi'(\mathbf{z}^\nu_{\cdot j})^2) + \beta_\nu(I + D_h^T D_h) & -M^T \\ \mathcal{X}^\nu(\mathbf{y}_{\cdot j}, \mathbf{z}^\nu_{\cdot j}, \boldsymbol{\theta}^\nu_{\cdot j})M & I + \mathcal{X}^\nu(\mathbf{y}_{\cdot j}, \mathbf{z}^\nu_{\cdot j}, \boldsymbol{\theta}^\nu_{\cdot j}) \end{pmatrix}, \end{aligned}$$

where $I$ denotes the identity matrix in the corresponding space and $\mathcal{X}^\nu \in \mathbb{R}^{4(N-1)\times 4(N-1)}$ is the diagonal matrix defined by

$$\mathcal{X}^\nu_{ii}(\mathbf{y}, \mathbf{z}^\nu, \boldsymbol{\theta}^\nu) = \begin{cases} 1 & \text{if } (\boldsymbol{\theta}^\nu_{\cdot j} - H\mathbf{y}^\nu_{\cdot j} - H_{\text{opp}}\mathbf{y}^{-\nu}_{\cdot j} - \mathcal{M}\mathbf{z}^\nu_{\cdot j} + \mathbf{c}^\nu_{\cdot j})_i > 0, \\ 0 & \text{else.} \end{cases}$$

Again, we recall that the semismooth Newton method (see Algorithm 7 in Appendix A) is well-defined and converges locally superlinear, if $Dg^{k,\nu}$ is nonsingular in a neighborhood around the solution and its inverse is locally bounded.
In our context, the outer while-loop is terminated, if the condition

$$\max_{\nu=1,\dots,N} \text{res}^{k,\nu} \leq \text{tol}$$

is satisfied, where player $\nu$'s residual is defined by

$$\begin{aligned} \text{res}^{k,\nu} :=& \|A_y\mathbf{y}^{k,\nu} - B_x\mathbf{x}^{k,\nu} - \mathbf{y}_0^\nu\| + \|\boldsymbol{\theta}^{k,\nu} - (\boldsymbol{\theta}^{k,\nu} - H\mathbf{y}^{k,\nu} - H_{\text{opp}}\mathbf{y}^{k,-\nu} + \mathcal{M}\mathbf{z}^{k,\nu} + \mathbf{c}^\nu)_+\| \\ &+ \left( - \sum_{i=1}^{2(N-1)} \sum_{j \in \hat{\mathcal{I}}^\nu_{+0}(\mathbf{z}^{k,\nu}, i)} \boldsymbol{\mu}^{k,\nu}_{ij} \right)_+ \\ &+ \left( \sum_{i=1}^{2(N-1)} \sum_{j \in \hat{\mathcal{I}}^\nu_{0-}(\mathbf{z}^{k,\nu}, i)} \boldsymbol{\mu}^{k,\nu}_{ij} - (\mathcal{M}^T\boldsymbol{\theta}^{k,\nu})_{ij} + \beta_\nu(D_h^T[D_h\mathbf{z}^{k,\nu}])_{ij} \right)_+ \end{aligned}$$

with the index sets

$$\hat{\mathcal{I}}^\nu_{+0}(\mathbf{z}^\nu, i) := \{j = 1, \dots, m \mid \mathbf{z}^\nu_{ij} = 1\}, \quad \hat{\mathcal{I}}^\nu_{0-}(\mathbf{z}^\nu, i) := \{j = 1, \dots, m \mid \mathbf{z}^\nu_{ij} = 0\}$$

and

$$\boldsymbol{\lambda}^{k,\nu} = \tilde{\boldsymbol{\lambda}}^{k,\nu} - \gamma_k^{-1}\mathbf{z}^{k,\nu}, \quad \boldsymbol{\mu}^{k,\nu} = \boldsymbol{\lambda}^{k,\nu} - \mathcal{M}^T\boldsymbol{\theta}^{k,\nu} + \beta_\nu D_h^T[D_h\mathbf{z}^{k,\nu}].$$

We end this section with two traffic scenarios in which we consider the time interval $\Omega := (0, 1)$ and $m = 100$ time steps.

**Application 6.3.9.** *The following hypothetical example considers a scenario with four vehicles in roundabout traffic, i.e. each vehicle intends to leave the setting in the same direction as it has entered (see Figure 6.4).*

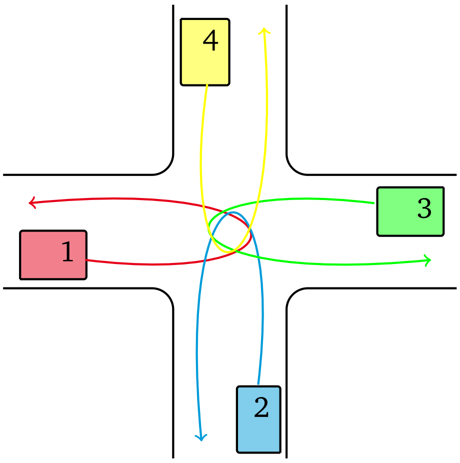



Figure 6.4.: Traffic scenario with four vehicles.

*It is obvious that there exists a great conflict potential in the center. For this reason, we assume for all $\nu = 1, \dots, N$ uniform data (see Table 6.1), the update scheme*

$$\gamma_{k+1} = \frac{1}{2}\gamma_k, \qquad \gamma_1 := 5 \cdot 10^{-1}$$

*and the initial values*

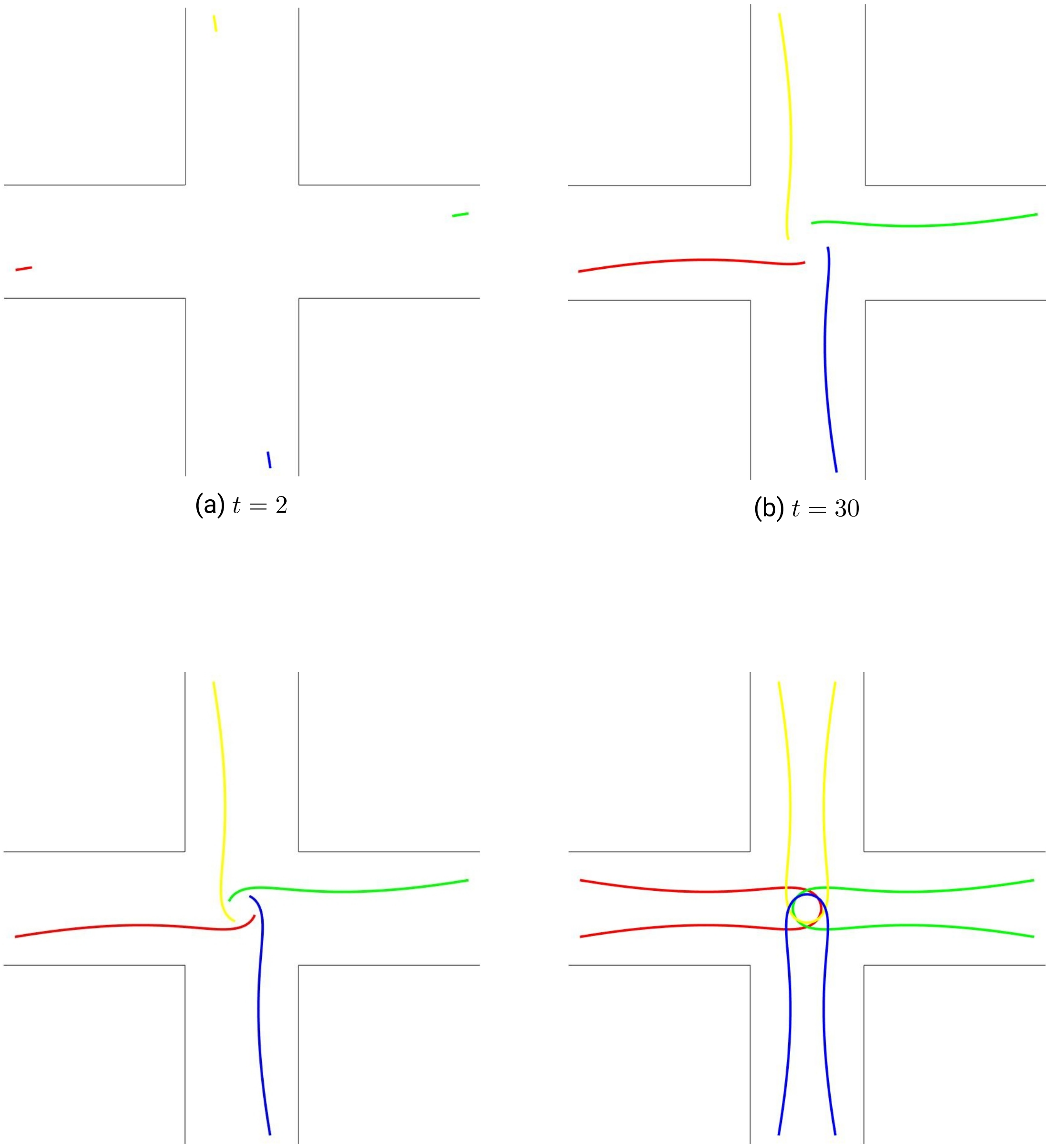
$$(\mathbf{y}^{0,\nu}, \boldsymbol{\rho}^{0,\nu}, \mathbf{z}^{0,\nu}, \boldsymbol{\theta}^{0,\nu}, \tilde{\boldsymbol{\lambda}}^{0,\nu}) = (\mathbf{y}_0^{\nu}, \boldsymbol{\rho}_m, \mathbb{1}, \mathbb{1}, \mathbb{1}),$$

*where $\mathbb{1}$ denotes the matrix with all ones in the particular matrix space. Moreover, we*

| Name | Parameter | Value |
|---|---|---|
| Weighting parameter | $\alpha$ | $(1, 1, 1, 1)$ |
| Lower bound of the control | $\mathbf{a}$ | $(0, 0, 0, 0)$ |
| Upper bound of the control | $\mathbf{b}$ | $(1, 1, 1, 1)$ |
| Initial state position | $y_0$ | $(0, 0, 0, 0)$ |
| Initial state velocity | $v_0$ | $(1, 1, 1, 1)$ |

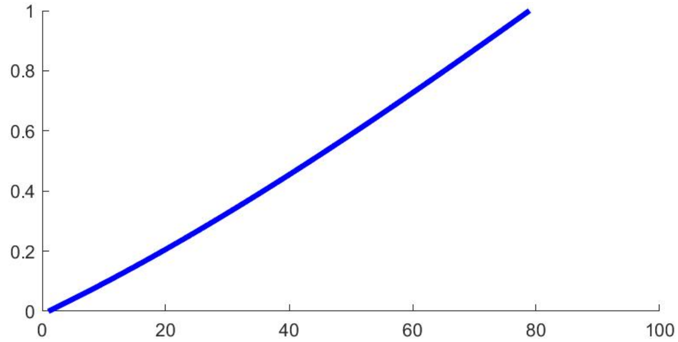
Table 6.1.: Uniform data for Application 6.3.9.

*set* $(M, \delta) := (50, 4h)$ *and stop the algorithm, if all residuals drop below the threshold* $\mathrm{tol} = 2 \cdot 10^{-5}$. *In particular, the uniform data guarantees the same conditions for all vehicles. For the simulations, we use B-spline curves in order to represent each path. In this context, we refer to Appendix B.2 for a short introduction or to [70] for a more detailed survey of splines. Now, Figure 6.5 depicts the current scenario at different time steps. In particular, Figure 6.5d shows that all vehicles reach their target position nearly at the same time. This assumption is confirmed in Figure 6.6, which visualizes the S-stationary Nash equilibrium and we see that the results are basically the same for all players. The latter can be expected, as we have uniform data.*

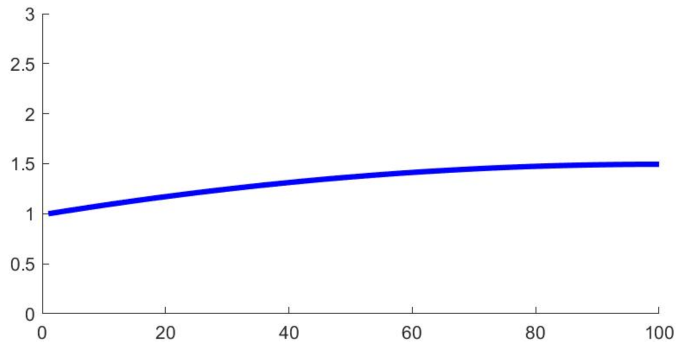


(a) $t = 2$ (b) $t = 30$ (c) $t = 40$ (d) $t = 79$

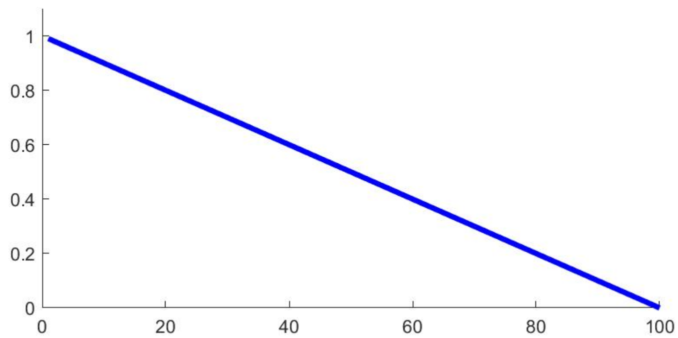
Figure 6.5.: Evaluation of Application 6.3.9 at different time steps $t$.




(a) Position $y_1$




(b) Velocity $y_2$




(c) Control (acceleration) $u$

Figure 6.6.: S-stationary Nash equilibrium of Application 6.3.9.

**Application 6.3.10.** *For the second example, we recall the scenario presented in Section 6.1, which is depicted once again in Figure 6.7.*

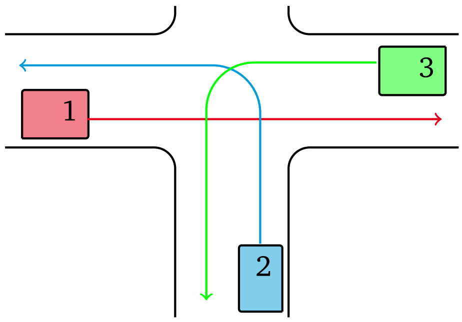



Figure 6.7.: Traffic scenario with three vehicles.

*In contrast to Application 6.3.9, we compute an S-stationary Nash equilibrium in this case for two sets of initial data. On the one hand, we consider individual starting terms (see Table 6.2) and on the other hand again uniform conditions (see Table 6.3) for comparable reasons. Moreover, we use similar parameters as before, i.e.*

$$(M, \delta) = (50, 3h), \qquad \gamma_1 = 5 \cdot 10^{-1}, \qquad \mathrm{tol} = 5 \cdot 10^{-5},$$

*with update scheme* $\gamma_{k+1} = \frac{1}{2}\gamma_k$ *and the same initial values.*

| Name | Parameter | Value |
|---|---|---|
| Weighting parameter | $\alpha$ | $(0.5, 0.5, 0.5)$ |
| Lower bound of the control | $\mathbf{a}$ | $(0, 0, 0)$ |
| Upper bound of the control | $\mathbf{b}$ | $(1, 1.4, 1.2)$ |
| Initial state position | $y_0$ | $(0, 0, 0)$ |
| Initial state velocity | $v_0$ | $(0.8, 1, 1.2)$ |

Table 6.2.: Individual data.

*With respect to individual data (Table 6.2), observe that* $t = 38$ *(Figure 6.8a) denotes the moment when vehicle* $1$ *and* $2$ *have minimum distance and* $t = 45$ *(Figure 6.8c) shows the point when the distance of vehicle* $1$ *and* $3$ *is the smallest. In contrast, Figure 6.8b and Figure 6.8d represent the progress at the same time points with respect to Table 6.3. Finally, Figure 6.8e and Figure 6.8f illustrate the moment, when the first vehicle reaches its target position. In particular, we see that all vehicles arrive at the same time under uniform*

| Name | Parameter | Value |
|---|---|---|
| Weighting parameter | $\alpha$ | $(1,1,1)$ |
| Lower bound of the control | $\mathbf{a}$ | $(0,0,0)$ |
| Upper bound of the control | $\mathbf{b}$ | $(1,1,1)$ |
| Initial state position | $y_0$ | $(0,0,0)$ |
| Initial state velocity | $v_0$ | $(1,1,1)$ |

Table 6.3.: Uniform data.

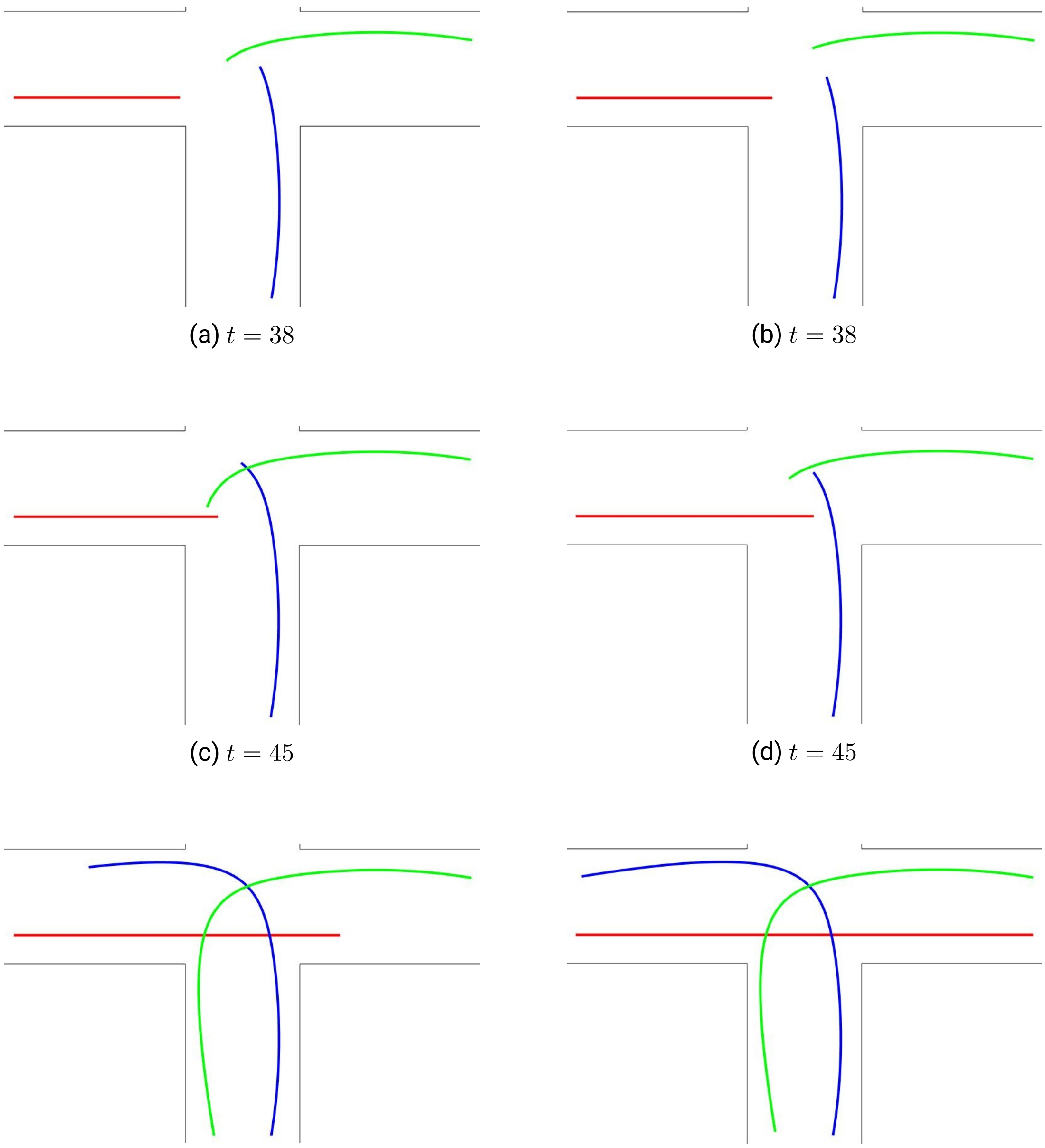
*initial data, which is confirmed in Figure 6.9b. In general, Figure 6.9 shows the S-stationary Nash equilibrium of* $\Gamma^{\text{AD}}_{\text{EPCC}}$ *with respect to Table 6.2 (left) and Table 6.3 (right).*

In our tested examples, the variable $\mathbf{z}^{k,\nu}$ tends to zero for all $\nu = 1,\dots,N$. Recalling the collision avoidance constraint and in particular (6.4), this implies that for two vehicles $\nu$ and $\mu$ with $\nu < \mu$ the condition

$$y_1^{\nu} - \hat{y}_{\mu\nu} \geq \delta$$

has to be satisfied, since the remaining three conditions trivially hold by construction. As a consequence, we conclude that in doubt of a collision the numeral successor (in this case vehicle $\mu$) is favored. With respect to the algorithm, this seems surprising, since Algorithm 4 is based on the Gauß-Seidel method. The latter implies that players with a small numbering have an advantage, as their problem is solved in advance. Hence, the chance of choosing an 'optimal' strategy is bigger. As a consequence, we conclude that the numbering of all vehicles is decisive. Furthermore, we want to emphasize that suitable data is required in order to find a solution. In this context, we observed that the algorithm was not capable to find an S-stationary Nash equilibrium for arbitrary given initial conditions. Therefore, we assume that there is relation between initial data and the existence of such stationary Nash equilibria that encourages further research in this direction.

Finally, we compare the numerical results to [14]. The main differences are on the one hand that the authors considered velocity constraints and on the other hand that they used *model predictive control*. The latter approach means that the solution of the system is computed on a certain interval $[0,t_1] \subseteq [0,1]$, the result is executed and the state at time step $t_1$ is used as the new initial state for the next subinterval $[t_1,t_2] \subseteq [0,1]$ until the final time is reached. As a consequence, this model is more flexible since only collisions in the particular intervals have to be considered. However, the velocity might become nonsmooth, which is not the case in our scenario.

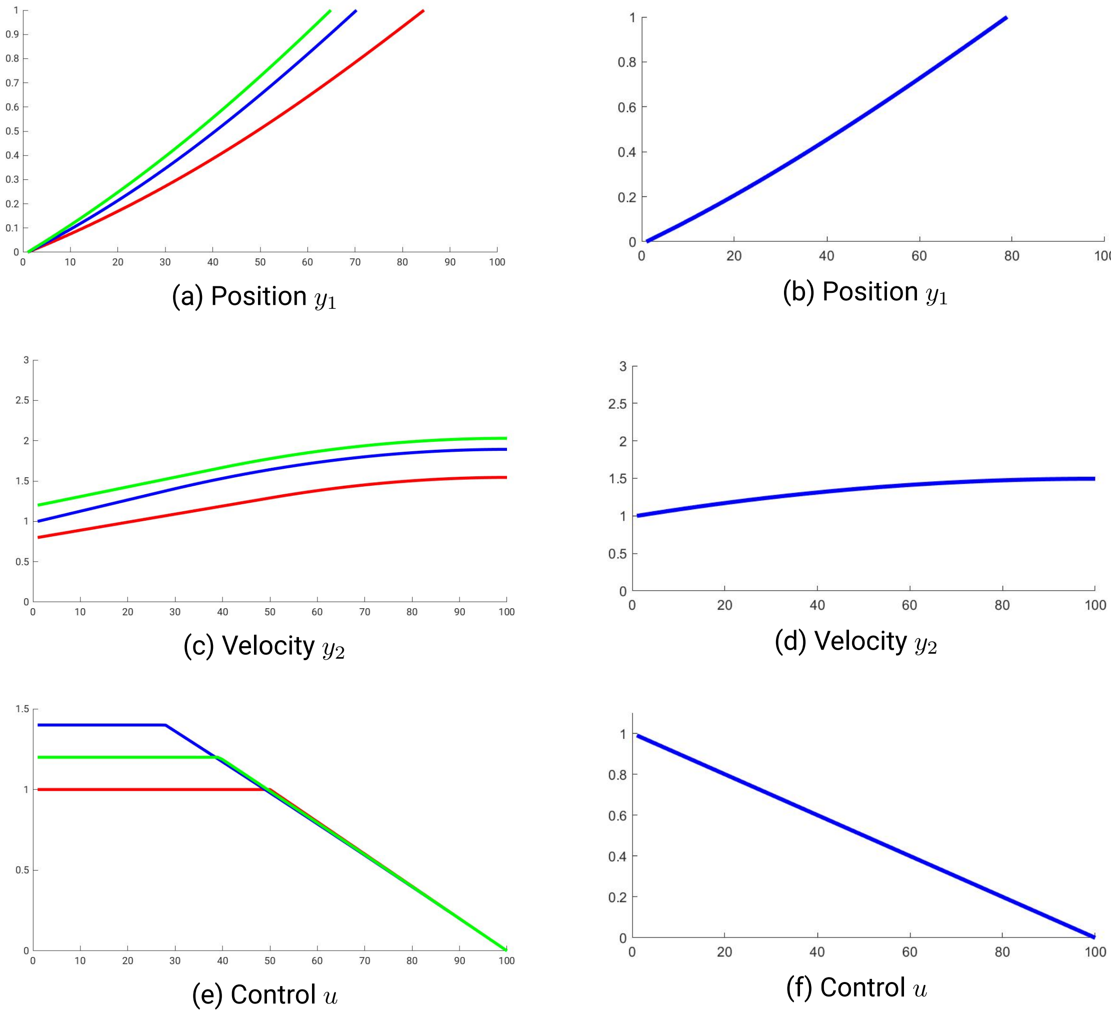


(a) $t = 38$ (b) $t = 38$

(c) $t = 45$ (d) $t = 45$

(e) $t = 65$ (f) $t = 79$

Figure 6.8.: Evaluation at different time steps $t$ w.r.t. individual (left) and uniform data 6.3 (right) of Application 6.3.10.



(a) Position $y_1$ (b) Position $y_1$

(c) Velocity $y_2$ (d) Velocity $y_2$

(e) Control $u$ (f) Control $u$

Figure 6.9.: S-stationary Nash equilibrium w.r.t to individual (left) and uniform data (right) of Application 6.3.10.

# 7. Conclusion

This thesis focused on equilibrium problems with equilibrium constraints in Banach spaces. For this reason, we started with a general problem setting and used the ideas of Gerd Wachsmuth and Patrick Mehlitz, who investigated mathematical programs with complementarity constraints in abstract Banach spaces [56, 94], in order to define weaker forms of stationary Nash equilibria. In addition, we introduced two penalization approaches that were used to find sequences, which converge to one of the latter stationary Nash equilibrium concepts.

In the following, we considered two special problems that could be written as EPECs/EPCCs. The first one were multi-leader multi-follower games. Therein, we concentrated on the relation between a multi-leader single-follower game and the corresponding EPCC and derived MLFG-tailored stationarity concepts. Afterwards, we focused on a special case, where the lower level problem was given by a quadratic program with linear constraints in Hilbert spaces. Under suitable assumptions we derived an equivalent EPEC-formulation on the one hand and an EPCC-formulation on the other hand.

In this context, we were able to show that the penalization approach for the EPEC generated a sequence of auxiliary equilibrium problems that approximated the original one, if all strategy variables converge strongly and the cost functionals are continuously convergent. Both requirements were satisfied for the corresponding EPEC in Chapter 5. However, it would be interesting to see whether the result is still valid for weakly convergent sequences. Furthermore, we emphasize at this point that we required the lower level feasible set to be independent of the leader strategies in order to apply the main theorem on monotone operators. Hence, a natural extension of this approach would be the analysis of parametric feasible sets.

In order to derive the same approximation result for the EPCC-formulation, we needed additionally to the strong convergence of the variables and the continuous convergence of the objective functionals that the feasible set was weakly Mosco-convergent. In particular, it was shown in Chapter 5 that these assumptions are non-trivial as the considered auxiliary equilibrium problem failed to satisfy these requirements. In this context, it is a question of future research, if the approximation is also valid under weaker assumptions, i.e. the weak convergence of the variables.

Despite the different assumptions in both problems, we showed in Chapter 5 that they could be applied to a multi-leader optimal control framework of the obstacle problem in $H_0^1(\Omega)$. Therein, we extended the well-known problem from an MPEC/MPCC-setting towards an equilibrium problem and verified that both approaches generated sequences that converged to an $\epsilon$-almost C-stationary Nash equilibrium. Stronger results were derived for both methods under additional assumptions on the inactive set $\mathcal{Q}_I$ and the Lagrange multiplier $\mu^{k,\nu}$, where we additionally gave a necessary condition for an S-stationary Nash equilibrium in the EPCC-setting. We concluded this part on MLFGs with the numerical analysis. Therein, we used a Gauß-Seidel-type diagonalization algorithm. While the $\epsilon$-almost C-stationary Nash equilibrium of the EPEC was computed with a semismooth Newton's method, we used an active-set strategy for the EPCC. Both methods were implemented on the basis of two academic examples and showed promising results with respect to the fact that it is in general not guaranteed that the Gauß-Seidel method converges towards a solution.
The second class of equilibrium problems was motivated by the work [14] by Axel Dreves and Matthias Gerdts. Therein, the authors considered an autonomous driving scenario that can be modeled by a non-convex game in $L^2(\Omega)$. We showed that the latter can be written as a generalized Nash equilibrium problem with mixed-integer variables, which again is equivalent to an EPCC. We analyzed this equilibrium problem in more general Lebesgue spaces and discussed why all stationarity concepts coincide in this setting. Afterwards, we concentrated again on the traffic scenario and utilized the same penalization technique as for the EPCC that resulted from the MLFG. In contrast to the latter problem, we verified the approximation of the auxiliary equilibrium problems towards the corresponding MINEP. In this context, we used that the objective functionals were independent of the opponents' strategies but had to strengthen the assumptions on the integer variable $z$ and required that its components $z_i^\nu$ were in $H_0^1(\Omega)$ instead of $L^2(\Omega)$. Finally, we considered the MINEP under numerical perspective and derived an algorithm that was also based on a Gauß-Seidel method. In this context, we notice that the algorithm did not find a solution for all possible settings and it would be nice to know, which parameters are the bottlenecks in this game and to derive conditions for the initial data that imply the existence of an S-stationary Nash equilibrium.

# A. The Semismooth Newton Method in Finite and Infinite Dimensional Banach Spaces

This section is mainly based on the survey paper [27]. However, further contributions can be found for instance in [89, 90].
Given two arbitrary Banach spaces $X, Y$ and a mapping $F : X \to Y$, we are looking for a point $\bar{x} \in X$ such that

$$F(\bar{x}) = 0. \tag{A.1}$$

Assuming that $F$ is at least continuously differentiable, the method of choice for solving (A.1) is the Newton's method given below in Algorithm 5.

**Algorithm 5** Newton's method for smooth systems.

**Input:** $F : X \to Y$ continuously differentiable, $x^0 \in X$.

1: Set $k := 0$.
2: Unless a stopping rule is satisfied, solve

$$DF(x^k)d^k = -F(x^k).$$

3: Set $x^{k+1} := x^k + d^k$, $k = k + 1$ and go back to line 2.
4: **return** $x^k$

It is well known (see e.g. [91, Satz 10.5]) that, if there exists $\bar{x} \in X$ such that (A.1) is satisfied and $DF(\bar{x})^{-1}$ exists and is bounded, then we find $\epsilon > 0$ such that Algorithm 5 generates a sequence $x^k$ that converges superlinearly to $\bar{x}$ for all initial points $x^0 \in \mathbb{B}_\epsilon(\bar{x})$. Moreover, if $DF(\bar{x})^{-1}$ is Lipschitz continuous on $\mathbb{B}_\epsilon(\bar{x})$, then $x^k$ converges quadratically to $\bar{x}$.
As a first-order method, Algorithm 5 highly depends on the differentiability of $F$. However, if $F$ is not necessarily differentiable anymore, we need to find a generalized derivative concept.
We first focus on the finite dimensional case, where the desired concept is based on *Rademacher*'s theorem (see [10]).

**Theorem A.0.1** (Rademacher's theorem). *Suppose $F : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz continuous. Then $F$ is almost everywhere differentiable, i.e. the set*

$$N_F := \left\{ x \in \mathbb{R}^n \;\middle|\; F(x) \text{ is not differentiable} \right\}$$

*has Lebesgue measure zero.*

Unless otherwise specified, we consider $F : \mathbb{R}^n \to \mathbb{R}^m$, which is locally Lipschitz continuous and define the set $D_F := \{x \in \mathbb{R}^n \mid F(x) \text{ is differentiable}\}$. From nonsmooth analysis we additionally define the *B(oulingand)-subdifferential* of $F$ at $x$ by

$$\partial_B F(x) := \left\{ G \in \mathbb{R}^{n\times m} \;\middle|\; \exists\{x^k\} \subseteq D_F \text{ s.t. } x^k \to x,\ \nabla F(x^k) \to G \right\},$$

and *Clarke's subdifferential* or *generalized Jacobian* by

$$\partial F(x) := \text{co}(\partial_B F(x)),$$

where *co* denotes the convex hull. If $m = 1$ then $\partial F(x)$ is called *generalized gradient*, which can be characterized by

$$\partial F(x) = \left\{ \xi \in \mathbb{R}^n \;\middle|\; \xi^T d \le F^\circ(x; d) \quad \forall d \in \mathbb{R}^n \right\}$$

with generalized directional derivative

$$F^\circ(x; d) := \limsup_{y \to x, t \searrow 0} \frac{F(y + td) - F(y)}{t}.$$

Since it holds $\partial F(x) = \{\nabla F(x)\}$, if $F$ is continuously differentiable we can extend Algorithm 5 with respect to locally Lipschitz continuous functions and obtain Newton's method for nonsmooth systems.

In contrast to Algorithm 5, where we used a linearization of $F$ for an approximation, we can not guarantee the same behavior for the nonsmooth case in Algorithm 6. Additionally, $\partial F$ is set-valued in general, which can cause trouble with respect to the well-definedness of the algorithm. Hence, we need further assumptions on $F$. Firstly stated in [57], a function $F$ is called *semismooth* at $x \in U$, with $U \subseteq \mathbb{R}^n$ open, if it is locally Lipschitz continuous at $x$ and the limit

$$\lim_{\substack{G \in \partial F(x + t\tilde{d}) \\ \tilde{d} \to d,\ t \searrow 0}} G\tilde{d}$$

exists for all $d \in \mathbb{R}^n$. $F$ is called semismooth, if it is semismooth at all $x \in U$. A more appropriate way of the definition above is given in the following result.

**Algorithm 6** Newton's method for nonsmooth systems.

**Input:** $F : X \to Y$ locally Lipschitz continuous, $x^0 \in X$.

1: Set $k := 0$.

2: Unless a stopping rule is satisfied, solve

$$G(x^k)d^k = -F(x^k),$$

where $G(x^k) \in \partial F(x^k)$ is arbitrarily chosen.

3: Set $x^{k+1} := x^k + d^k$, $k = k + 1$ and go back to line 2.

4: **return** $x^k$

**Lemma A.0.2** ([27, Theorem 2.9]). *Let $F : U \to \mathbb{R}^m$ with $U \subseteq \mathbb{R}^n$ open and $x \in U$. Then, the following statements are equivalent:*

1. *$F$ is semismooth at $x$.*
2. *$F$ is locally Lipschitz continuous at $x$, the directional derivative $F'(x; d)$ exists and for any $G \in \partial F(x + d)$ it holds*

   $$\|Gd - F'(x; d)\| = o(\|d\|)$$

   *as $d \to 0$.*
3. *$F$ is locally Lipschitz continuous at $x$, the directional derivative $F'(x; d)$ exists and for any $G \in \partial F(x + d)$ it holds*

   $$\|F(x + d) - F(x) - Gd\| = o(\|d\|)$$

   *as $d \to 0$.*

Now, we can show a local convergence result for Algorithm 6.

**Theorem A.0.3** ([27, Theorem 2.11]). *Suppose that $\bar{x} \in \mathbb{R}^n$ satisfies $F(\bar{x}) = 0$, $F$ is semismooth at $\bar{x}$ and $\partial F(\bar{x})$ is nonsingular. Then there exists $\epsilon > 0$ such that for all $x_0 \in \mathbb{B}_\epsilon(\bar{x})$ the sequence $\{x^k\}$ generated by Algorithm 6 is well-defined, converges to $\bar{x}$ and satisfies*

$$\|x^{k+1} - \bar{x}\| = o(\|x^k - \bar{x}\|)$$

*as $k \to \infty$.*

Summarizing, Theorem A.0.3 shows a local, superlinear convergence of the semismooth Newton method to a solution of (A.1). Note that under 'stronger' semismoothness a quadratic convergence rate can be shown.
Now, we consider the infinite dimensional setting. Therefore, observe that a straight forward generalization is not possible as the Clarke subdifferential relies on Rademacher's theorem, which is only satisfied in finite dimensions. Motivated by Lemma A.0.2, we choose a different access to a generalized derivative concept.

**Definition A.0.4** ([27, Definition 3.2]). Let $X, Y$ be arbitrary Banach spaces. Then the mapping $F : D \subseteq X \to Y$ is *generalized (slantly or Newton) differentiable* on the open set $U \subseteq D$, if there exists a family of mappings $G : U \to \mathcal{L}(X, Y)$ such that it holds

$$\lim_{d \to 0} \frac{1}{\|d\|_X} \|F(x+d) - F(x) - G(x+d)d\|_Y = 0$$

for every $x \in U$. $G$ is called *generalized* (or *Newton*) *derivative* of $F$.

As a result, we can define Algorithm 7, the infinite dimensional analogon to Algorithm 6. In particular, we obtain the same results that are given in the subsequent theorem.

**Algorithm 7** Newton's method for semismooth operator equations

**Input:** $F : D \to Y$ generalized differentiable in $U \subseteq D$ and $x^0 \in X$.
1: Set $k := 0$.
2: Unless a stopping rule is satisfied, solve

$$G(x^k)d^k = -F(x^k),$$

3: where $G(x^k)$ is an arbitrary generalized derivative of $F$ at $x^k$.
4: Set $x^{k+1} := x^k + d^k$, $k = k+1$ and go back to line 2.
5: **return** $x^k$

**Theorem A.0.5** ([27, Theorem 3.2]). *Suppose that $\bar{x} \in X$ is a solution of $F(x) = 0$ and that $F$ is generalized differentiable in an open neighborhood $U$ containing $\bar{x}$ with Newton derivative $G$. Moreover, let $G(x)$ be nonsingular for all $x \in U$ and assume that the set $\{\|G(x)\|^{-1} \mid x \in U\}$ is bounded. Then there exists $\epsilon > 0$ such that for $x^0 \in \mathbb{B}_\epsilon(\bar{x})$, Algorithm 7 is well-defined and converges superlinearly to $\bar{x}$, provided that $\|x^0 - \bar{x}\|$ is sufficiently small.*

# B. Autonomous Driving

This section is dedicated to additional information with respect to Section 6.3. In particular, we show in Appendix B.1 the reformulation of the non-convex collision avoidance constraint (6.3) to a linear inequality with additional binary variables. Finally, Appendix B.2 gives a short introduction into the theory of B-spline curves that are used to model the preassigned paths.

## B.1. From a Non-Convex to a Mixed-Integer Problem

In the following section, we verify that (6.3) is equivalent to (6.4), i.e. we consider a non-convex inequality constraint of type

$$\max\{|x-a|, |y-b|\} - d \geq 0 \tag{B.1}$$

with $a, b, x, y \in \mathbb{R}$ and $d > 0$ and show that this can be written as

$$\begin{aligned} H \begin{bmatrix} x-a \\ y-b \end{bmatrix} + \tilde{\mathcal{M}} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - c \geq 0&, \\ z \in \{0,1\}^2&. \end{aligned} \tag{B.2}$$

Here, $H, \tilde{\mathcal{M}} \in \mathbb{R}^{4\times 2}$ and $c \in \mathbb{R}^4$ are given by

$$H := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \; \tilde{\mathcal{M}} := \begin{pmatrix} M & M \\ M & -M \\ -M & M \\ -M & -M \end{pmatrix} \text{ and } c := \begin{pmatrix} d \\ d-M \\ d-M \\ d-2M \end{pmatrix},$$

where $M >> 0$ is assumed to be sufficiently large.
We start by observing that (B.1) is equivalent to the logical conditions

$$x - a \geq d \quad \text{OR} \quad y - b \geq d \quad \text{OR} \quad a - x \geq d \quad \text{OR} \quad b - y \geq d. \tag{B.3}$$

In order to get rid of the 'OR'-expression, we introduce a binary variable $\tilde{z} \in \{0,1\}^4$ and a positive constant $M >> 0$ sufficiently large. Then observe that (B.3) can be written as the following system, which is given by

$$\begin{array}{r} x - a \geq d - M\tilde{z}_1 \\ y - b \geq d - M\tilde{z}_2 \\ a - x \geq d - M\tilde{z}_3 \\ b - y \geq d - M\tilde{z}_4 \end{array}, \qquad \sum_{i=1}^{4} \tilde{z}_i \leq 3, \qquad \tilde{z} \in \{0,1\}^4. \tag{B.4}$$

In this context, the second condition ensures that at least one component of $z$ is zero and hence, one original inequality constraint of (B.3) is satisfied. However, (B.4) can further be simplified in the sense that we can reduce on the one hand the number of binary variables $\tilde{z}_i$ and on the other hand that the sum inequality condition vanishes. For this reason, observe that (B.3) can also be written as

$$\begin{aligned} x - a &\geq d - M(z_1 + z_2), \\ y - b &\geq d - M(1 + z_1 - z_2), \\ a - x &\geq d - M(1 - z_1 + z_2), \\ b - y &\geq d - M(2 - z_1 - z_2), \\ z &\in \{0,1\}^2 \end{aligned}$$

and hence, (B.2) holds. We immediately see that at least one inequality in (B.3) is satisfied for all possible configurations $z$ can take.

## B.2. B-spline Curves and the Description of Preassigned Paths

In this section, we focus on the description of curves in the plane, i.e. a representation of a function $\gamma : [t_0, t_F] \to \mathbb{R}^2$. At this point, we have decided to use B-spline curves and refer to [70] for a more detailed analysis. The main motivation for choosing this parametrization is the flexibility with respect to almost all kind of curves, which can occur in our autonomous driving scenario. However, we want to emphasize that other approaches, e.g. polynomials or nonuniform rational B-splines (NURBS), also work.
In order to define B-splines, we need some basic definitions. Therefore, let $\tau \in [t_0, t_F]$. Then we call

$$\gamma : [t_0, t_F] \to \mathbb{R}^2, \tau \to \gamma(\tau) := \sum_{i=1}^{m} B_i^p(\tau) P_i$$

the *Bézier curve* of order $m$ (see also [70, Section 1.3]), where $P_i \in \mathbb{R}^2$ denote *control points* and the set $\{B_i^m(\tau)\} \subseteq \mathbb{R}$ defines a basis of Bernstein polynomials given by

$$B_i^m(\tau) := \frac{1}{(t_F - t_0)^m} \begin{pmatrix} m \\ i \end{pmatrix} (\tau - t_0)^i (t_F - \tau)^{m-i}$$

for all $i = 1, \dots, m$. Now, B-splines can be seen as a first natural extension of Bézier curves. While both are mainly used in the field of computer aided designs, i.e. to approximate geometrical lines and surfaces, the latter parametrization has the critical drawback that the polynomial degree increases proportionally to the complexity of the corresponding form, which can lead to numerical instability. Hence, B-splines are used to decrease the order of the polynomial. For this reason, a vector $\hat{T} := \big(\hat{t}_1, \dots, \hat{t}_n\big)$ with $t_0 \leq \hat{t}_i \leq \hat{t}_{i+1} \leq t_F$ for all $i = 1, \dots, n$ is called *knot vector*, where the elements are called *knots*. Then we can define B-spline basis functions.

**Definition B.2.1** ([70, Section 2.2])**.** Let $\hat{T}$ be a knot vector. Then the $i$th *B-spline basis function* of order $p$ is defined recursively by

$$N_i^0(\tau) := \begin{cases} 1 & \text{if } \hat{t}_i \leq \tau \leq \hat{t}_{i+1}, \\ 0 & \text{else}, \end{cases}$$

$$N_i^p(\tau) := \frac{\tau - \hat{t}_i}{\hat{t}_{i+p} - \hat{t}_i} N_i^{p-1} + \frac{\hat{t}_{i+1+1} - \tau}{\hat{t}_{i+p+1} - \hat{t}_{i+1}} N_{i+1}^{p-1}(\tau).$$

Notice that the naming is appropriate since the set $\{N_i^p\}_{i=1}^m$ with $m \in \mathbb{N}$ indeed defines a basis (see [70, Section 2.4]). Now, we are able to introduce B-spline curves.

**Definition B.2.2** ([70, Section 3.2])**.** Consider the knot vector

$$\hat{T} = (\underbrace{t_0, \dots, t_0}_{p+1 \text{ times}}, \hat{t}_{p+2}, \dots, \hat{t}_m, \underbrace{t_F, \dots, t_F}_{p+1 \text{ times}})$$

and let $\{P_i\}_{i=1}^m \subseteq \mathbb{R}^2$ be a set of control points and $\{N_i^p(\tau)\}_{i=1}^m$ the B-spline basis functions of order $p$. Then the function

$$\gamma : [t_0, t_F] \to \mathbb{R}^2, \gamma(\tau) := \sum_{i=1}^m N_i^p(\tau) P_i$$

is called *B-spline curve of order $p$*.

Note that due to the special structure of $\hat{T}$, it holds $\gamma(t_0) = P_1$ and $\gamma(t_F) = P_m$, i.e. the first and the last control point are on the curve, while the intermediate points can only be seen as a limiting polygon (see also Figure B.1). Given suitable control points, we are now able to model each vehicle's path in a sufficient way.
The next natural extension in this context are NURBS. The main advantage of the latter curves are the possibility to approximate conic forms, i.e. ellipses, parabolas and hyperbolas, respectively, which also includes circles. Since these curves do not appear in our context that often, we do not consider the extension here, but refer again to [70] for more information.

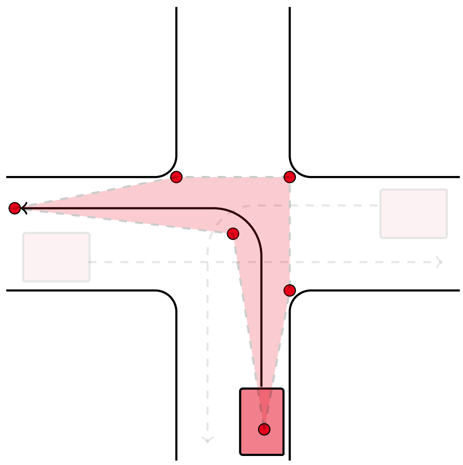


Figure B.1.: Description of preassigned path.

# Bibliography


[1] R. A. Adams and J. JF Fournier. *Sobolev Spaces*. Vol. 140. Elsevier Science, 2003.

[2] H. W. Alt. *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. Springer Berlin Heidelberg, 2006.

[3] J.-P. Aubin and H. Frankowska. *Set-valued Analysis*. Birkhäuser Boston, 2009.

[4] V. Barbu. “Optimal control of variational inequalities”. In: *Research Notes in Math.* 100 (1984).

[5] V. Barbu and T. Precupanu. *Convexity and Optimization in Banach Spaces*. Springer Netherlands, 2012.

[6] P. Belotti et al. “Mixed-integer nonlinear optimization”. In: *Acta Numerica* 22 (2013).

[7] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer New York, 2000.

[8] J. C. Butcher and N. Goodwin. *Numerical methods for ordinary differential equations*. Wiley Online Library, 2008.

[9] M. Červinka. “Hierarchical Structures in Equilibrium Problems”. PhD thesis. Univerzita Karlova, 2008.

[10] F. H Clarke. *Optimization and Nonsmooth Analysis*. SIAM, 1990.

[11] G. Dal Maso. *An Introduction to* Γ*-Convergence*. Vol. 8. Springer Science & Business Media, 1993.

[12] S. Dempe. “Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints”. In: (2003).

[13] S. Dempe and J. Dutta. “Is bilevel programming a special case of a mathematical program with complementarity constraints?” In: *Mathematical programming* 131.1-2 (2012), pp. 37–48.

[14] A. Dreves and M. Gerdts. “A generalized Nash equilibrium approach for optimal control problems of autonomous cars”. In: *Optimal Control Applications and Methods* 39.1 (2018), pp. 326–342.

[15] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998.

[16] F. Facchinei, V. Piccialli, and M. Sciandrone. “Decomposition algorithms for generalized potential games”. In: *Computational Optimization and Applications* 50.2 (2011), pp. 237–262.

[17] K. Fan. “A minimax inequality and applications”. In: *Inequalities* 3 (1972), pp. 103–113.

[18] M. L. Flegel and C. Kanzow. “On M-stationary points for mathematical programs with equilibrium constraints”. In: *Journal of Mathematical Analysis and Applications* 310.1 (2005), pp. 286–302.

[19] S. Franke, P. Mehlitz, and M. Pilecka. “Optimality conditions for the simple convex bilevel programming problem in Banach spaces”. In: *Optimization* 67.2 (2017), pp. 237–268.

[20] M. Gerdts. “Optimal control of ordinary differential equations and differential-algebraic equations”. Habilitation. University of Bayreuth, 2006.

[21] R. Glowinski, J.-L. Lions, et al. *Numerical analysis of variational inequalities*. Elsevier, 1981.

[22] G. Gürkan and J.-S. Pang. “Approximations of Nash equilibria”. In: *Mathematical Programming* 117.1-2 (2009), pp. 223–253.

[23] F. Harder and G. Wachsmuth. “Comparison of optimality systems for the optimal control of the obstacle problem”. In: *GAMM-Mitteilungen* 40.4 (2018), pp. 312–338.

[24] R. Henrion, J. Outrata, and T. Surowiec. “Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 18.2 (2012), pp. 295–317.

[25] M. Herty, S. Steffensen, and A. Thünen. “Solving quadratic multi-leader-follower games by smoothing the follower’s best response”. In: *Optimization Methods and Software* (2020), pp. 1–28.

[26] A. von Heusinger and C. Kanzow. “Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions”. In: *Computational Optimization and Applications* 43.3 (2009), pp. 353–377.

[27] M. Hintermüller. “Semismooth Newton methods and applications”. In: *Department of Mathematics, Humboldt-University of Berlin* (2010).

[28] M. Hintermüller, K. Ito, and K. Kunisch. “The primal-dual active set strategy as a semismooth Newton method”. In: *SIAM Journal on Optimization* 13.3 (2002), pp. 865–888.

[29] M. Hintermüller and I. Kopacka. “A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs”. In: *Computational Optimization and Applications* 50.1 (2011), pp. 111–145.

[30] M. Hintermüller and I. Kopacka. “Mathematical programs with complementarity constraints in function space: C-and strong stationarity and a path-following algorithm”. In: *SIAM Journal on Optimization* 20.2 (2009), pp. 868–902.

[31] M. Hintermüller, C. Löbhard, and M. H. Tber. “An l1-penalty scheme for the optimal control of elliptic variational inequalities”. In: *Numerical Analysis and Optimization* (2015), pp. 151–190.

[32] M. Hintermüller, B. S. Mordukhovich, and T. M. Surowiec. “Several approaches for the derivation of stationarity conditions for elliptic MPECs with upper-level control constraints”. In: *Mathematical Programming* 146.1-2 (2014), pp. 555–582.

[33] M. Hintermüller and T. M. Surowiec. “A PDE-constrained generalized Nash equilibrium problem with pointwise control and state constraints”. In: *Pacific J. on Optimization* 9.2 (2013), pp. 251–273.

[34] M. Hintermüller, T. M. Surowiec, and A. Kämmler. “Generalized Nash Equilibrium Problems in Banach Spaces: Theory, Nikaido–Isoda-Based Path-Following Methods, and Applications”. In: *SIAM Journal on Optimization* 25.3 (2015), pp. 1826–1856.

[35] M. Hinze et al. *Optimization with PDE Constraints*. Vol. 23. Springer Science & Business Media, 2008.

[36] A. Hori and M. Fukushima. “Gauss–Seidel Method for Multi-leader–follower Games”. In: *Journal of Optimization Theory and Applications* 180.2 (2019), pp. 651–670.

[37] M. Hu and M. Fukushima. “Multi-leader-follower games: models, methods and applications”. In: *Journal of the Operations Research Society of Japan* 58.1 (2015), pp. 1–23.

[38] M. Hu and M. Fukushima. “Smoothing approach to Nash equilibrium formulations for a class of equilibrium problems with shared complementarity constraints”. In: *Computational Optimization and Applications* 52.2 (2012), pp. 415–437.

[39] M. Hu and M. Fukushima. "Variational inequality formulation of a class of multi-leader-follower games". In: *Journal of optimization theory and applications* 151.3 (2011), p. 455.

[40] A. N. Iusem and W. Sosa. "New existence results for equilibrium problems". In: *Nonlinear Analysis: Theory, Methods & Applications* 52.2 (2003), pp. 621–635.

[41] J. Jahn. *Introduction to the Theory of Nonlinear Optimization*. Springer Science & Business Media, 2007.

[42] J. Jahn. *Vector Optimization*. Springer, 2011.

[43] C. Kanzow and A. Schwartz. *Spieltheorie*. Springer, 2018.

[44] C. Kanzow and D. Steck. "Quasi-variational inequalities in Banach spaces: theory and augmented Lagrangian methods". In: *SIAM Journal on Optimization* 29.4 (2019), pp. 3174–3200.

[45] C. Kanzow et al. "The multiplier-penalty method for generalized Nash equilibrium problems in Banach spaces". In: *SIAM Journal on Optimization* 29.1 (2019), pp. 767–793.

[46] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and their Applications*. Vol. 31. Siam, 1980.

[47] A. A. Kulkarni and U. V. Shanbhag. "A shared-constraint approach to multi-leader multi-follower games". In: *Set-valued and variational analysis* 22.4 (2014), pp. 691–720.

[48] K. Kunisch and D. Wachsmuth. "Sufficient optimality conditions and semi-smooth Newton methods for optimal control of stationary variational inequalities". In: *ESAIM: Control, Optimisation and Calculus of Variations* 18.2 (2012), pp. 520–547.

[49] G. Leoni. *A First Course in Sobolev Spaces*. American Mathematical Soc., 2017.

[50] S. Leyffer and T. Munson. "Solving multi-leader-common-follower games". In: *Optimisation Methods & Software* 25.4 (2010), pp. 601–623.

[51] M. B. Lignola and J. Morgan. "Convergence of solutions of quasi-variational inequalities and applications". In: *Topological Methods in Nonlinear Analysis* 10.2 (1997), pp. 375–385.

[52] J.-L. Lions. "Various topics in the theory of optimal control of distributed systems". In: *Optimal Control Theory and its Applications*. Springer, 1974, pp. 166–309.

[53] P. Mehlitz. "Contributions to Complementarity and Bilevel Programming in Banach spaces". PhD thesis. Technische Universität Bergakademie Freiberg, 2017.

[54] P. Mehlitz. “Necessary optimality conditions for a special class of bilevel programming problems with unique lower level solution”. In: *Optimization* 66.10 (2017), pp. 1533–1562.

[55] P. Mehlitz and G. Wachsmuth. “The limiting normal cone to pointwise defined sets in Lebesgue spaces”. In: *Set-Valued and Variational Analysis* 26.3 (2018), pp. 449–467.

[56] P. Mehlitz and G. Wachsmuth. “Weak and strong stationarity in generalized bilevel programming and bilevel optimal control”. In: *Optimization* 65.5 (2016), pp. 907–935.

[57] R. Mifflin. “Semismooth and semiconvex functions in constrained optimization”. In: *SIAM Journal on Control and Optimization* 15.6 (1977), pp. 959–972.

[58] F. Mignot. “Contrôle dans les inéquations variationelles elliptiques”. In: *Journal of Functional Analysis* 22.2 (1976), pp. 130–185.

[59] F. Mignot and J. P. Puel. “Optimal control in some variational inequalities”. In: *SIAM Journal on Control and Optimization* 22.3 (1984), pp. 466–476.

[60] B. S. Mordukhovich. “Equilibrium problems with equilibrium constraints via multiobjective optimization”. In: *Optimization Methods and Software* 19.5 (2004), pp. 479–492.

[61] B. S. Mordukhovich. “Optimization and equilibrium problems with equilibrium constraints in infinite-dimensional spaces”. In: *Optimization* 57.5 (2008), pp. 715–741.

[62] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory*. Vol. 330. Springer Science & Business Media, 2006.

[63] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation II: Applications*. Springer Science & Business Media, 2006.

[64] U. Mosco. “Convergence of convex sets and of solutions of variational inequalities”. In: *Advances in Mathematics* 3.4 (1969), pp. 510–585.

[65] J.F. Nash. “Equilibrium points in n-person games”. In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49.

[66] J.F. Nash. “Non-cooperative games”. In: *Annals of mathematics* (1951), pp. 286–295.

[67] J. v. Neumann. “Zur Theorie der Gesellschaftsspiele”. In: *Mathematische Annalen* 100.1 (1928), pp. 295–320.

[68] M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT press, 1994.

[69] J.-S. Pang and M. Fukushima. "Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games". In: *Computational Management Science* 2.1 (2005), pp. 21–56.

[70] L. Piegl and W. Tiller. *The NURBS Book*. Springer Science & Business Media, 2012.

[71] I. Prodan et al. "Collision avoidance and path following for multi-agent dynamical systems". In: *ICCAS 2010*. IEEE. 2010, pp. 1930–1935.

[72] A. M. Ramos and T. Roubicek. "Nash equilibria in noncooperative predator-prey games". In: *Applied Mathematics and Optimization* 56.2 (2007), pp. 211–241.

[73] A.-T. Rauls and G. Wachsmuth. "Generalized derivatives for the solution operator of the obstacle problem". In: *Set-Valued and Variational Analysis* (2018), pp. 1–27.

[74] R. T. Rockafellar and R. J-B Wets. *Variational Analysis*. Vol. 317. Springer Science & Business Media, 2009.

[75] W. Rudin. *Functional analysis*. McGraw-Hill, 1991.

[76] S. Sagratella. "Algorithms for generalized potential games with mixed-integer variables". In: *Computational Optimization and Applications* 68.3 (2017), pp. 689–717.

[77] S. Sagratella. "On generalized Nash equilibrium problems with linear coupling constraints and mixed-integer variables". In: *Optimization* 68.1 (2019), pp. 197–226.

[78] H. Scheel and S. Scholtes. "Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity". In: *Mathematics of Operations Research* 25.1 (2000), pp. 1–22.

[79] A. Schiela and D. Wachsmuth. "Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 47.3 (2013), pp. 771–787.

[80] S. Scholtes. "Convergence properties of a regularization scheme for mathematical programs with complementarity constraints". In: *SIAM Journal on Optimization* 11.4 (2001), pp. 918–936.

[81] T. Schouwenaars et al. "Mixed integer programming for multi-vehicle path planning". In: *2001 European Control Conference (ECC)*. IEEE. 2001, pp. 2603–2608.

[82] A. Schwartz. "Mathematical Programs with Complementarity Constraints: Theory, Methods and Applications". PhD thesis. Julius-Maximilians-Universität Würzburg, 2011.

[83] D. Steck. “Lagrange Multiplier Methods for Constrained Optimization and Variational Problems in Banach Spaces”. PhD thesis. Julius-Maximilians-Universität Würzburg, 2018.

[84] C.-L. Su. “Equilibrium Problems with Equilibrium Constraints: Stationarities, Algorithms, and Applications”. PhD thesis. Stanford University, 2005.

[85] T. M. Surowiec. “Explicit Stationarity Conditions and Solution Characterization for Equilibrium Problems with Equilibrium Constraints”. PhD thesis. Humboldt-Universität Berlin, 2010.

[86] T. M. Surowiec. “Numerical Optimization Methods for the Optimal Control of Elliptic Variational Inequalities”. In: *Frontiers in PDE- Constrained Optimization*. Springer, 2018, pp. 123–170.

[87] Z. Tang, J-A Désidéri, and J. Périaux. “Multicriterion aerodynamic shape design optimization and inverse problems using control theory and nash games”. In: *Journal of Optimization Theory and Applications* 135.3 (2007), pp. 599–622.

[88] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*. Vol. 112. American Mathematical Soc., 2010.

[89] M. Ulbrich. “Semismooth Newton methods for operator equations in function spaces”. In: *SIAM Journal on Optimization* 13.3 (2002), pp. 805–841.

[90] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. SIAM, 2011.

[91] M. Ulbrich and S. Ulbrich. *Nichtlineare Optimierung*. Springer-Verlag, 2012.

[92] H. Von Stackelberg. *Marktform und Gleichgewicht*. J. springer, 1934.

[93] G. Wachsmuth. “A guided tour of polyhedric sets”. In: *Preprint, TU Chemnitz* (2016).

[94] G. Wachsmuth. “Mathematical programs with complementarity constraints in Banach spaces”. In: *Journal of Optimization Theory and Applications* 166.2 (2015), pp. 480–507.

[95] G. Wachsmuth. “Strong stationarity for optimal control of the obstacle problem with control constraints”. In: *SIAM Journal on Optimization* 24.4 (2014), pp. 1914–1932.

[96] G. Wachsmuth. “Towards M-stationarity for optimal control of the obstacle problem with control constraints”. In: *SIAM Journal on Control and Optimization* 54.2 (2016), pp. 964–986.

[97] D. Werner. *Funktionalanalysis*. Springer, 2018.

[98] J.-P. Yvon. “Optimal control of systems governed by variational inequalities”. In: *IFIP Technical Conference on Optimization Techniques*. Springer. 1973, pp. 265–275.

[99] E. Zeidler. *Nonlinear Functional Analysis and its Applications: II/B: Nonlinear Monotone Operators*. Springer Science & Business Media, 1990.

[100] J. Zowe and S. Kurcyusz. “Regularity and stability for the mathematical programming problem in Banach spaces”. In: *Applied mathematics and Optimization* 5.1 (1979), pp. 49–62.

# Curriculum Vitae

## Jan Becker

| | |
|---|---|
| 03/2021 | Promotion in Mathematik (Dr. rer. nat.)<br>Technische Universität Darmstadt |
| 02/2018 – 01/2021 | Assoziation im Rahmen der Graduiertenschule Computational Engineering<br>Technische Universität Darmstadt |
| 11/2017 – 03/2021 | Projektstelle im Rahmen des DFG-SPP 1962, Teilprojekt P21<br>Technische Universität Darmstadt |
| 10/2017 | M.Sc. in Computational Mathematics<br>Julius-Maximilians-Universität Würzburg |
| 11/2015 | B.Sc. in Computational Mathematics<br>Julius-Maximilians-Universität Würzburg |