Technische Universität München

TUM School of Computation, Information and Technology

On Generalized Nash Equilibrium Problems in Infinite-Dimensional Spaces: Existence, Regularized Nikaido–Isoda Merit Functionals, and Augmented Lagrangian Methods

Julia Fritz, M.Sc.

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Stefan Weltge

Prüfer*innen der Dissertation:

1. Prof. Dr. Michael Ulbrich

2. Prof. Dr. Christian Kanzow

Die Dissertation wurde am 18.10.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 17.04.2023 angenommen.

Abstract

This thesis considers generalized Nash equilibrium problems (GNEPs) in an infinite-dimensional setting with as general assumptions on the underlying functions and spaces as possible.

The theory of and numerical methods for (generalized) Nash equilibrium problems are well-investigated in the finite-dimensional setting under suitable convexity assumptions. Here, we investigate GNEPs in an infinite-dimensional nonconvex setting. In the case of convex constraints, we express the GNEP in terms of the regularized Nikaido–Isoda merit functional and apply a generalized version of the Kakutani fixed point theorem in order to prove the existence of fixed points. We present characterizations of these points by the regularized Nikaido–Isoda merit functional and conclude the existence of variational and normalized equilibria. We emphasize that the Nikaido–Isoda merit functional is not straightaway differentiable in the classical sense. Thus, we develop differentiability results using a generalized version of Danskin's theorem, which enables the use of derivative-based solution methods.

For handling general, possibly nonconvex, constraints, we also present an augmented Lagrangian method for GNEPs in infinite-dimensional spaces. In contrast to previous studies, such as works by Kanzow and Steck, we require weaker assumptions in order to provide convergence results for Karush–Kuhn–Tucker-type equilibria. In particular, we can work with multiplier-penalty terms in more general spaces than Hilbert spaces.

Zusammenfassung

Diese Arbeit betrachtet verallgemeinerte Nash-Gleichgewichtsprobleme (engl. GNEPs) in einem unendlich-dimensionalen Rahmen mit möglichst allgemeinen Annahmen an die zugrundeliegenden Funktionen und Räume.

Die Theorie und die numerischen Methoden für (verallgemeinerte) Nash-Gleichgewichtsprobleme sind im endlich-dimensionalen Rahmen unter geeigneten Konvexitätsannahmen gut erforscht. Hier untersuchen wir GNEPs in einem unendlich-dimensionalen, nicht-konvexen Rahmen. Im Falle von konvexen Nebenbedingungen drücken wir das GNEP mittels des regularisierten Nikaido–Isoda-Funktionals aus und wenden eine verallgemeinerte Version des Kakutani Fixpunkttheorems an, um die Existenz von Fixpunkten zu beweisen. Wir zeigen verschiedene Charakterisierungen dieser Punkte mittels des regulariserten Nikaido–Isoda Funktionals und schließen auf die Existenz variationeller und normalisierter Gleichgewichte. Wir betonen, dass das Nikaido–Isoda-Funktional nicht direkt im klassischen Sinne differenzierbar ist, und entwickeln daher Differenzierbarkeitsergebnisse unter Verwendung einer verallgemeinerten Version des Danskin-Satzes. Dies ermöglicht die Verwendung von Lösungsmethoden, welche Ableitungenen verwenden.

Für den Umgang mit allgemeinen, möglicherweise nicht-konvexen, Nebenbedingungen stellen wir außerdem eine augmentierte Lagrange-Methode zur Bestimmung von Karush-Kuhn-Tucker-basierten Gleichgewichtspunkten vor. Im Gegensatz zu früheren Studien, wie beispielsweise der Arbeiten von Kanzow und Steck, benötigen wir schwächere Annahmen, um Konvergenzergebnisse zu zeigen. Insbesondere können wir mit Multiplikator-Penalty-Termen in allgemeineren unendlich-dimensionalen Räumen als Hilberträumen arbeiten.

Acknowledgments

> *A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between theories. One can imagine that the ultimate mathematician is one who can see analogies between analogies.*
>
> (Stefan Banach)

Finally, I have successfully defended my thesis and my Ph.D. studies are completed.

First and foremost, I would like to thank my supervisor, Prof. Dr. Michael Ulbrich, for the chance to do my Ph.D. studies. I am thankful to Prof. Dr. Christian Kanzow for reviewing my thesis, and I'd also want to thank Prof. Dr. Stefan Weltge for agreeing to serve as the committee chair.

I appreciate the funding provided by the SPP 1962 priority program "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization." The program gave me the opportunity to collaborate with and learn a great deal from Marcel Steinhardt of the Technical University of Darmstadt.

In addition, I would like to thank the former and current members of the M1 group. It was quite enjoyable. I would like to thank Dr. Lukas Hertlein for his assistance and good conversations. I am grateful to Dr. Florian Lindemann for his insightful counsel. As previous Ph.D. students of the chair, I am thankful to Dr. Johannes Haubner, Dr. Johannes Milz and Franziska Neumann for their help throughout my early and middle Ph.D studies. Furthermore, I am thankful for the frequent conversations, discussions, and interesting lunch breaks in both the old and new Mensa with my fellow Ph.D. students. Frauke Bäcker deserves special mention for her assistance with administrative matters. Thank you for creating such a lovely, productive, and welcoming workplace.

Throughout my Master's and Ph.D. studies, Prof. Dr. Christina Kuttler served as my mentor and provided invaluable advice and counsel. I would like to thank her for her support in difficult situations and her uplifting words.

I am grateful to everyone who proofread the present document. Moreover, I would like to thank all of my friends who have assisted me in various ways throughout the years, especially Ina, Martin, Nadine, Sina, and Stefan. In addition, I would want to convey my deepest thanks to my beloved husband Marvin, who has been there for me throughout the years. Without the encouragement and support of my parents, Michael and Susanne, as well as my sister, Sonja, none of this would have been possible. A special thanks to Lilli for being such a lovely and well-behaved daughter.

Thank you everybody, you are incredible.

Contents

Abstract i

Zusammenfassung i

Acknowledgments iii

1 Introduction 1

2 Mathematical Preliminaries 5

2.1 Functional Analysis 6

2.1.1 Topological Spaces 6

2.1.2 Normed Spaces 11

2.1.3 Banach Spaces 14

2.2 Operator Theory 18

2.2.1 Monotone Operators 19

2.2.2 Projection Operator 20

2.2.3 Duality Mapping 21

2.3 Theory of Constrained Optimization 23

2.3.1 Optimality Conditions 23

2.3.2 Equilibrium Concepts for Games 25

2.3.3 Concept of the Augmented Lagrangian Method 29

2.4 Generalized Danskin Theorem 32

3 Mathematical Setting and Assumptions 41

3.1 Topology Assumptions 41

3.2 Continuity and Differentiability Assumptions 43

3.3 Convexity Assumptions 43

4 Existence of Fixed Points of the Solution Map 47

4.1 Fixed Point of the Solution Map 48

4.2 Fixed Point of the Solution Map via Optimality Condition 51

5 Nikaido–Isoda Merit Functionals **59**
5.1 Localization of Merit Functionals 59
5.1.1 Localizing the Feasible Set 60
5.1.2 Penalty-type and Barrier-type Method 60
5.1.3 Combination of Localizing and Penalizing 61
5.2 Merit Functionals and Equilibria 62
5.2.1 Nash Equilibria 62
5.2.2 Normalized Equilibria 71
5.3 Difference of Merit Functionals 80
5.4 Differentiability of Regularized and Localized Nikaido–Isoda Merit Functionals 86
5.4.1 Differentiability of regularized Merit Nikaido–Isoda functionals 87
5.4.2 Application to Localized Nikaido–Isoda Merit Functionals 91

6 Algorithms and Methods for Computing Equilibria **103**
6.1 Descent Method for the Regularized Nikaido–Isoda Merit Functional 104
6.2 Augmented Lagrangian Method 109
6.2.1 Augmented Lagrangian Method for Quasi-Nash Equilibria 111
6.2.2 Augmented Lagrangian Method for Variational Equilibria 124

7 Conclusion and Outlook **131**

Acronyms **133**

Bibliography **135**

Chapter 1

Introduction

> *The concept of a Nash equilibrium n-tuple is perhaps the most important idea in non-cooperative game theory. Whether we are analysing candidates' election strategies, the causes of war, agenda manipulation in legislatures, or the actions of interest groups, predictions about events reduce to a search for and description of equilibria. Put simply, equilibrium strategies are the things that we predict about people.* (Thomas Schelling)

This thesis aims to provide analytical and numerical solutions to generalized Nash equilibrium problems (GNEPs) in infinite-dimensional spaces. Nash equilibrium problems (NEPs) were first described by Nash in his work [84] in 1950 and afterward extended by Arrow and Debreu in 1954, see [5], and Rosen in 1965, see [94]. The concept of Nash equilibria has been discussed and investigated in numerous articles since then, see [40, 42–45, 106]. Facchinei and Kanzow's survey paper [43] summarizes the theory for GNEPs in the convex and finite-dimensional case. In comparison to the well-studied finite-dimensional case, the literature on GNEPs in infinite-dimensional spaces is quite limited, see [58, 85, 86, 109, 112]. The majority of them impose strict requirements on the objective functional, particularly some kind of convexity assumption, e.g., [86, 109, 112]. There is also literature for some specific objective functionals and appropriate infinite spaces, see [57]. Hintermüller has provided a brief overview of an analogous theory as in the finite-dimensional case in [58] for a class of convex objective functionals.

GNEPs have been used in a variety of real-world scenarios, including economics, aerodynamics [102], and electricity [39]. Furthermore, optimization problems with infinite-dimensional spaces have important applications, such as in traffic flow [19–21] and natural gas spot markets [58]. As a result, having the same tools available as in finite dimensions is advantageous. To prove new analytical and numerical results, we combine the methods of finite-dimensional GNEPs with optimality theory in Banach spaces, see, for example [10, 18, 78, 114].

The majority of GNEP research focuses on the case where the objective functionals are convex. Classically, one considers the Nikaido–Isoda functional and the corresponding merit functional, see the original work [86] of Nikaido and Isoda. The properties of this merit functional have been thoroughly investigated, particularly its relationship to variational equilibria and normalized equilibria, see [43, 58, 106]. In the case of convex objective functionals, for example, the concept of variational equilibria is equivalent to normalized equilibria. Descent methods for these Nikaido–Isoda merit functionals, which usually require first-order derivatives, are used in algorithms for computing Nash equilibria and normalized equilibria for convex constraints. It was pointed out in [106] that the numerical methods for quasi-variational inequalitys (QVIs) are not as developed or efficient as those for Nikaido–Isoda merit functionals. Nonetheless, theoretical literature for numerical methods for GNEPs and their convergence properties is still scarce. The GNEP can be reformulated as a constrained minimization problem for the (regularized) Nikaido–Isoda functional. Interesting results for the merit functionals in the jointly convex case have been obtained, for example, see [40, 54, 106, 107]. This merit functional is not twice continuously differentiable in general. However, it is shown in [105] that the gradient of the (regularized) Nikaido–Isoda merit functional is semismooth under additional assumptions. In such a case, there are locally fast convergent Newton-type methods, see the books [62, 104], and they have been applied to GNEPs in [45, 105].

In contrast to the majority of papers, we concentrate on GNEPs with nonconvex objective functionals, resulting in the existence of quasi-Nash equilibria rather than Nash equilibria and variational equilibria rather than normalized equilibria. To that end, we reformulate the GNEP in terms of the Nikaido–Isoda merit functional, as shown in [86], and its regularization, as shown in [106], and using QVIs, as shown in [44, 53, 89]. QVIs were introduced in [9] and have since then become a standard tool for modeling various equilibrium-type scenarios in natural sciences, such as solid and continuum mechanics [75, 88], economics [60], transportation [16, 98], and electrostatics [2, 55, 56, 76, 93]. For more information on QVIs, see the monographs [7, 75, 83] and the references therein.

We consider the solution map, which is based on the regularized Nikaido–Isoda functional and a first-order optimality condition, and we use a generalized version of the Kakutani fixed point theorem to prove the existence of a fixed point. These fixed points can be translated into variational equilibria or normalized equilibria.

Due to the objective functional's lack of strict convexity, the solution map could be multivalued, and thus, the Nikaido–Isoda merit functional is not straightaway differentiable in the classical sense. We compute a continuous derivative of the regularized and localized Nikaido–Isoda merit functional using a generalized version of Danskin's theorem. Further, we use the representation of variational equilibria in terms of the regularized Nikaido–Isoda merit functional and construct a descent method for the corresponding minimization problem to find variational equilibria numerically. Furthermore, we demonstrate its global convergence solely under the assumption of the regularized Nikaido–Isoda merit functional's continuous differentiability.

In order to solve GNEPs with additional (possibly nonconvex) constraints that arise from a solution operator to a partial differential equation, we build an augmented Lagrangian method that employs the descent method for a suitable subproblem. The augmented Lagrangian method was introduced in [89] in 2005 and later improved in [65] in 2016. The augmented Lagrangian method is a well-known method for solving constrained optimization problems, and it is mentioned in textbooks such as [12, 13, 32, 87]. In recent years, the augmented Lagrangian method has been used in the form of safeguarded methods, which have strong global convergence properties and use a different update of the Lagrangian multiplier estimate, see [14]. The work [69] compares the classical augmented Lagrangian method and its safeguarded analogue. In addition, the safeguarded augmented Lagrangian method has been extended to QVIs in finite dimensions, as shown in [65, 70, 89], as well as to constrained optimization problems and variational inequalities in Banach spaces, as shown in [66, 71, 73, 74]. We refer to [24, 25, 66, 68, 101] for results on convergence properties.

We develop and demonstrate the convergence of a safeguarded augmented Lagrangian method for games in this thesis. Unlike the majority of the literature, we consider a GNEP based on minimization problems with objective functionals and constraints that are not necessarily convex. In this case, the convergence theory only provides convergence to Karush–Kuhn–Tucker (KKT) points that satisfy the first-order optimality conditions. These are the corresponding GNEP's quasi-Nash equilibria and variational equilibria, see the work [72] of Kanzow and Steck. In the general assumption of the underlying state space, our results differ. By assuming a uniformly smooth and uniformly convex Banach space, we generalize the Hilbert space assumption. Because the constraints are nonconvex, we use an augmented Lagrangian method whose subproblems approximate a variational inequality (VI) rather than solving a minimization problem based on the augmented Lagrangian functional.

This thesis is structured as follows. We provide an overview of the mathematical preliminaries in Chapter 2. In detail, we present a collection of important functional analysis theorems, generalize Danskin's theorem in Section 2.4, describe game equilibrium concepts in Subsection 2.3.2, and explain the augmented Lagrangian method concept in Subsection 2.3.3.

In Chapter 3, we present the mathematical context and key definitions for our problem. In Chapter 4, we look at two optimization problems with objective functionals of Nikaido–Isoda type. Using a generalized version of the Kakutani fixed point theorem, we show that the corresponding solution maps have a fixed point. These fixed points are linked to Nikaido–Isoda merit functionals, the analytical properties of which are examined in Chapter 5. We then look at the corresponding localized merit functionals. We establish a link between these merit functionals, (quasi-)variational inequalitys ((Q)VIs), (local) Nash equilibria, and (local) normalized equilibria. In Section 5.4, we prove the differentiability of the two Nikaido–Isoda merit functionals and their localized variants. To that end, we employ a suitable version of Danskin's theorem, see Section 2.4.

In Section 6.1, we present a descent method for the regularized Nikaido–Isoda merit functional using the projected gradient. Finally, in Section 6.2, we create a safeguarded augmented Lagrangian method for games by transferring the descent method to a suitable subproblem. Subsection 6.2.1 and Subsection 6.2.2 demonstrate convergence towards QNEs and VEs.

Chapter 2

Mathematical Preliminaries

> *In most sciences one generation tears down what another has built, and what one has established another undoes. In mathematics alone each generation builds a new story to the old structure.* (Hermann Hankel)

This opening chapter establishes several key concepts that are required for the rest of the thesis. The majority of the content offered here is basically a meticulous compilation of results from the literature, structured and presented in such a way that the theory is as apparent as possible.

The following is an outline of the chapter's structure. In Section 2.1, we will primarily focus on the instruments of functional analysis that we require for our findings in infinite-dimensional spaces. The results in this part can be found in classical and newer books such as [3,4,8,23,29,30,33,49,79,82,95,103,111,113]. Subsection 2.1.1 contains preliminary results on topological spaces, including a discussion of distinct definitions of compactness and convergence in the sequential topology. Furthermore, we describe operator properties such as pseudomonotonicity and complete continuity. Subsection 2.1.2 presents some fundamental conclusions on normed spaces, weak convergence, and distinct types of differentiability. The following Subsection 2.1.3 is devoted to Banach spaces, and we discuss several topological properties such as reflexive, uniformly smooth, and uniformly convex spaces. Furthermore, we introduce the concepts of different notions of continuity, such as hemicontinuity and complete continuity. Finally, in Subsection 2.2.3, we present some special operators known as duality mappings, which introduce a connection between the original space and its dual. In particular, we discuss the link between the duality mapping and the derivative of Banach norms.

In Section 2.3, we will look at the theoretical background of constrained optimization in normed spaces. In many ways, we are only scratching the surface of these vast topics and mostly focus on the equilibrium concepts for games. For the general theory of constrained optimization in infinite-dimensional spaces, we refer to [18,59]. Subsection 2.3.1 introduces some essential notions of cones and optimality conditions. Furthermore, constraint qualifications (CQs) and the well-known KKT pairs are discussed. Subsection 2.3.2 provides an introduction to the equilibrium concept of games, including a number of notable examples such as Nash equilibria, as well as the important characterization of equilibria via Nikaido–Isoda merit functionals. Finally, Subsection 2.3.3 contains the augmented Lagrangian method and we provide a first look at the general outline of an algorithm.

We close this chapter with Section 2.4 where we provide a generalized version of the Danskin theorem in order to ensure the differentiability of a specific bidual optimization problem.

2.1 Functional Analysis

This section introduces various functional analysis concepts applicable to spaces with distinct topologies and structures. We begin with topological spaces before introducing the concept of compactness. We present significant results from normed spaces and differentiability, as well as discuss the concept of weak convergence and weak compactness principles.

2.1.1 Topological Spaces

This section is predominately based on the results presented in the book [111], and we begin by defining a topological space and a Hausdorff space.

Definition 2.1 (Topological and Hausdorff space)**.** A *topological space* (X, τ) consists of a set X together with a topology τ, which is a system of subsets such that

(i) $\emptyset, X \subseteq \tau$,

(ii) $\cup_i A_i \subseteq \tau$ for all $A_i \subseteq \tau$, $i \in \mathbb{N}$,

(iii) $\cap_i A_i \subseteq \tau$ for all $A_i \subseteq \tau$, $i \in \mathbb{N}$ finite.

The sets in τ are called *open sets* of the topological space (X, τ). Furthermore, (X, τ) is called a *Hausdorff space* if for all $x, y \in X$ with $x \neq y$ there exist disjoint open sets $A_1, A_2 \in \tau$ such that $x \in A_1$, $y \in A_2$.

In accordance with a common abuse of notation, we write X instead of (X, τ) for the concept of the topological space when the underlying topology is obvious or when it is unnecessary to specify the precise topology. Later on, however, we will equip the underlying set X with different topologies, necessitating a definition for comparing two topologies. This is specified in the following definition.

Definition 2.2 (Finer and coarser topology)**.** Let τ_1, τ_2 be two topologies on the same set X. We say τ_2 is *finer* than τ_1 (or: τ_1 is *coarser* than τ_2) if every subset that is open with respect to τ_1 is also open with respect to τ_2. Shortly, we write $\tau_2 \subseteq \tau_1$.

Since equilibrium problems typically involve multiple players, we must define a product space and discuss its underlying topology.

Definition 2.3 (Product topology)**.** Let $\{X_i\}_{i\in[N]}$, $[N] = \{1,\dots,N\}$, be topological spaces. Then the topology of the *product space*

$$\prod_{i\in[N]} X_i = X_1 \times X_2 \times \cdots X_N$$

is defined by the sets of the form $\prod_{i\in[N]} A_i$ where $\{A_i\}_{i\in[N]}$ are open sets in the corresponding topological space X_i, $i \in [N]$, and coincide with X_i.

Let X be a topological space. In the following, we briefly address some classical terminology in functional analysis. A *neighborhood* of a point $x \in X$ is a set A such that there exists an open set $C \subseteq A$ with $x \in C$. Note that the neighborhood A is not assumed to be open. We call a set A *closed* if and only if its complement $A^c = X \backslash A$ is open in X. The *closure* of A, denoted by $\overline{A}$ or $\mathrm{cl}(A)$, is the intersection of all closed subsets of X which contains A. The *interior* of A, denoted by $\mathrm{int}(A)$, is given by the union of all open subsets of A. A set $A \subseteq X$ is said to be *dense* if $\overline{A} = X$. A set $A \subseteq X$ is *compact* if for every system of open sets $\{O_i\}_{i\in I}$ that covers A, i.e., $A \subseteq \cup_{i\in I} O_i$, there exists a finite set of indices $J \subseteq I$ such that $A \subseteq \cup_{i\in J} O_i$. Furthermore, A is *convex* if for all $x, y \in A$ and $t \in [0,1]$ it holds $(1-t)x + ty \in A$.

The next result demonstrates an important property of compact sets of topological spaces.

Proposition 2.4 (cf. [111, Proposition 1, 2])**.** Let A be a compact set of a topological space. Then A is closed and every closed subset of A is compact.

In metric spaces, the convergence of sequences can express equivalently the topological notions of openness, closedness, and compactness of sets. However, in general topological spaces, these correspondences no longer hold true. For example for the weak topology weakly open sets are weakly sequentially open but weakly sequentially open sets are not weakly open, see [23, Chapter 3.2, Remark 3]. Specifically, there is the notion of nets to ensure the equivalence between the topological and sequential definitions of openness, closedness, and compactness, see [8, Lemma 1.10].

Since nets are challenging to handle, we work with sequences and distinguish between their topological and sequential definitions. Thus, we introduce the essential concept of convergent sequences.

Definition 2.5 (Convergence)**.** A sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq X$ is said to be *convergent* to a point $\bar{x} \in X$, in symbols $x_k \to \bar{x}$ in X as $k \to \infty$, if every neighborhood of $\bar{x}$ contains all but finitely many elements of $\{x_k\}_{k\in\mathbb{N}}$.

We note that the limit point is unique in Hausdorff spaces, see [52, Proposition 2.4]. The usage of sequences and their convergence leads to the sequential definition of openness, closedness, and compactness, as we state in the next definition that is based on [52, Definition 1.5] and [4, Definition 1.47].

Definition 2.6 (Sequential properties)**.** A set $A \subseteq X$ is called

(i) *sequentially open* if whenever $\{x_k\}_{k\in\mathbb{N}} \subseteq X$, $\bar{x} \in A$, $x_k \to \bar{x}$ in X as $k \to \infty$, then $x_k \in A$ for sufficiently large k,

(ii) *sequentially closed* if every limit point of a convergent sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq A$ belongs to A,

(iii) *sequentially compact* if every sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq A$ has a convergent subsequence $\{x_{k_l}\}_{l\in\mathbb{N}} \subseteq \{x_k\}_{k\in\mathbb{N}}$ with a limit point in A.

We note that A is sequentially open if $X\backslash A$ is sequentially closed. Each open set is sequentially open, and each closed set is sequentially closed. A topological space X is called *sequential* if every sequentially open subset is open, see [52, Definition 1.7]. By the result [8, Lemma 1.34], a sequentially compact set is automatically sequentially closed. Furthermore, a sequentially closed subset of a sequentially compact set is sequentially compact, refer to [49, Proposition 15.2.1] and [8, Lemma 1.34].

In case of set-valued mappings, we also mention the following definition of sequential closedness, see [50, Chapter 1, Definition].

Definition 2.7 (Sequentially closed for set-valued mappings)**.** Let X be a Hausdorff linear topological space and $A \subseteq X$ a (nonempty) convex subset of X. Furthermore, let $f : A \rightrightarrows X$ be a set-valued mapping and let $\{a_k\}_{k\in\mathbb{N}} \subseteq A$, $\{x_k\}_{k\in\mathbb{N}} \in f(a_k)$ be two arbitrary converging sequences with $a_k \to \bar{a}$ in A and $x_k \to \bar{x}$ in X as $k \to \infty$ for the limit points $\bar{a} \in A$ and $\bar{x} \in X$. Then the *set-valued mapping f is called sequentially closed* if it holds $\bar{x} \in f(\bar{a})$.

Moreover, the Kakutani fixed point theorem applies to set-valued mappings and is extremely important for proving the existence of fixed points on convex sets.

Theorem 2.8 (Kakutani fixed point theorem, cf. [50, Chapter 1, Theorem])**.** Let X be a Hausdorff locally convex linear topological space and $A \subseteq X$ be a nonempty, convex and compact subset. Let $T : A \rightrightarrows A$ be a set-valued mapping. If T is sequentially closed and $T(a)$ is a nonempty and convex subset of A for any $a \in A$, then there exists a fixed point $x \in A$ with $x \in T(x)$.

Next, we introduce the concept of sequential continuity for mappings, see [23, Chapter 1.4].

Definition 2.9 (Sequential continuity)**.** Let X, Y be two topological spaces.

(i) An operator $T : X \to Y$ is called *(sequentially) continuous* if $T^{-1}(A) \subseteq X$ is (sequentially) open for every open set A of Y.

(ii) A functional $f : X \to \mathbb{R} \cup \{\infty\}$ is called *(sequentially) lower semicontinuous* if the level set $\{x \in X : f(x) \leq c\}$ is (sequentially) closed for all $c \in \mathbb{R}$.

(iii) A functional $f : X \to \mathbb{R} \cup \{\infty\}$ is called *(sequentially) upper semicontinuous* if the level set $\{x \in X : f(x) \geq c\}$ is (sequentially) closed for all $c \in \mathbb{R}$.

Note that a functional $f : X \to \mathbb{R} \cup \{\infty\}$ is (sequentially) upper semicontinuous if and only if $-f$ is (sequentially) lower semicontinuous. The following proposition can be seen as an alternative characterization of sequential continuity.

Proposition 2.10 (cf. [103, page 12, Definition 8.1], [52, Lemma 3.1])**.** Let X, Y be two topological spaces. The operator $T : X \to Y$ is sequentially continuous if and only if for every sequence $\{x_k\}_{k \in N}$ with $x_k \to \bar{x}$ in X as $k \to \infty$ it holds $T(x_k) \to T(\bar{x})$ in Y as $k \to \infty$.

A continuous operator $T : X \to Y$ is sequentially continuous. For the reverse direction, one has to assume a sequential space as first-countable spaces or metric spaces, as we state in the next proposition.

Proposition 2.11 (cf. [52, Lemma 3.1])**.** Let X be a sequential topological space and Y a topological space. Then $T : X \to Y$ is continuous if and only if it is sequentially continuous.

Next, we characterize sequential lower and upper semicontinuity in Hausdorff spaces using the notion of limit inferiors and superiors instead of level sets.

Proposition 2.12 (cf. [8, Lemma 1.36])**.** Let X be a Hausdorff space. Then the functional $f : X \to \mathbb{R} \cup \{\infty\}$ is

(i) sequentially lower semicontinuous if and only if for any sequence $\{x_k\}_{k \in \mathbb{N}}$ such that $x_k \to \bar{x}$ in X as $k \to \infty$ it holds $\liminf_{k \to \infty} f(x_k) \geq f(\bar{x})$,

(ii) sequentially upper semicontinuous if and only if for any sequence $\{x_k\}_{k \in \mathbb{N}}$ such that $x_k \to \bar{x}$ in X as $k \to \infty$ it holds $\limsup_{k \to \infty} f(x_k) \leq f(\bar{x})$.

If $f : X \to \mathbb{R} \cup \{\infty\}$ is lower (upper) semicontinuous, then it is sequentially lower (upper) semicontinuous. The converse direction holds for sequential spaces, see [8, Remark 1.37].

Definition 2.13 (Convex functional)**.** Let X be a vector space. Then the functional $f : X \to \mathbb{R} \cup \{\infty\}$ is called

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

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y),$$

(ii) *strictly convex*, if for $x, y \in X$, $x \neq y$, and $t \in (0, 1)$ it holds

$$f((1-t)x + ty) < (1-t)f(x) + tf(y).$$

In this dissertation, we will examine the concept of the sequential topology, see [15] and [41]. The topology defined by sequentially open sets is called sequential topology, and the topological space that corresponds to it is called sequential. This concept can be applied not only to strongly convergent sequences but also weakly convergent sequences.

We will employ the weak sequential topology, which is induced by weakly sequentially open sets in a topological space. Indeed, it is a topology, compare [15] and [41]. In order to define weak convergence, we first introduce the space $\mathscr{L}(X; Y)$ of all linear and continuous operators from X to Y. For the special choice $Y = \mathbb{R}$, we call the space $\mathscr{L}(X; \mathbb{R})$ the *dual space* of X and shortly write this space as X^*. Next, we define the concept of weak convergence.

Definition 2.14 (Weak and weak-$*$ convergence)**.** A sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq X$ *converges weakly* to a point $\bar{x} \in X$ if it holds

$$\langle x^*, x_k\rangle_{X^*,X} \to \langle x^*, \bar{x}\rangle_{X^*,X} \qquad \forall\, x^* \in X^*,$$

as $k \to \infty$. Moreover, a sequence $\{x_k^*\}_{k\in\mathbb{N}} \subseteq X^*$ *converges weakly-$*$* to a point $\bar{x}^* \in X$ if it holds

$$\langle x_k^*, x\rangle_{X^*,X} \to \langle \bar{x}^*, x\rangle_{X^*,X} \qquad \forall\, x \in X,$$

as $k \to \infty$. Shortly, we write $x_k \rightharpoonup \bar{x}$ in X and $x_k^* \overset{*}{\rightharpoonup} \bar{x}^*$ in X^* as $k \to \infty$, respectively.

We must identify the distinction between weak topology and weak sequential topology. The weak topology is the smallest topology such that each operator $T_i : X \to Y_i$ in the operator family $\{T_i\}_{i\in I}$ with the topological spaces Y_i, $i \in I$, is continuous, see [99, Section 6.6.1]. The weak topology defines a Hausdorff topology for Hausdorff spaces Y_i, $i \in I$, see [4, Proposition 1.8]. Specifically, sequentially open sets with respect to the weak topology are called weakly sequentially open sets, which are equivalent to open sets with respect to the weak sequential topology. We must be cautious because the term "sequentially" is frequently not written explicitly in the literature. Weakly open sets or open sets with respect to the weak topology are weakly sequentially open sets or open sets with respect to the weak sequential topology, but the reverse implication is invalid because the weak topology is not metrizable, as shown in [23, Chapter 3.2, Remark 3]. Consequently, the weak sequential topology is more refined than the weak topology and is also a Hausdorff topology.

In addition, one must be careful when describing compact sets: weakly sequentially compact refers to sequentially compact with respect to the weak topology, not compact with respect to the weak sequential topology. Later, we'll discuss the theorem of Eberlein–Šmulian, which equates weakly compactness and weakly sequentially compactness, see Theorem 2.25 in Subsection 2.1.3. In normed spaces, we will see that these three notions of compactness for a set are equivalent:

- weak compactness,
- weak sequential compactness (or more precisely, sequential compactness with respect to the weak topology),
- compactness with respect to the weak sequential topology.

Since the space with the weak sequential topology is sequential, the topological terminologies in Definition 2.9, Proposition 2.11, and Proposition 2.12 are identical to the following sequential concepts. We refer to [8, p.35] and [33, Definition 3.2].

Definition 2.15 (Continuity in the weak sequential topology)**.** Let $\{x_k\}_{k\in\mathbb{N}}$ be a sequence such that $x_k \rightharpoonup \bar{x}$ in X as $k \to \infty$. An operator $T : X \to Y$ is called

(i) *continuous with respect to the weak sequential topology* if $T(x_k) \rightharpoonup T(\bar{x})$ in Y as $k \to \infty$,

(ii) *completely continuous* if $T(x_k) \to T(\bar{x})$ in Y as $k \to \infty$.

A functional $f : X \to \mathbb{R} \cup \{\infty\}$ is called

(i) *lower semicontinuous with respect to the weak sequential topology* if

$$\liminf_{k\to\infty} f(x_k) \geq f(\bar{x}),$$

(ii) *upper semicontinuous with respect to the weak sequential topology* if

$$\limsup_{k\to\infty} f(x_k) \leq f(\bar{x}).$$

Remark 2.16. Clearly, every completely continuous operator is continuous with respect to the weak sequential topology. In the literature, there is the related notion of compact operators. In fact, a linear operator $T : X \to Y$ is called *compact* if T maps bounded sets to relatively compact sets (i.e., sets with compact closure [103, Definition 6.2]), see also [97, p. 98]. For two Banach spaces X and Y, every compact operator is completely continuous. However, the converse only holds true if X is reflexive, compare the book [33, Proposition 3.3]. In this study, the weaker condition of a completely continuous operator is sufficient. Note that the complete continuity is identical to the continuity between the weak sequential and strong topologies. A completely continuous operator is also known as weak-to-strong sequentially continuous.

2.1.2 Normed Spaces

In optimization theory, topological spaces can be challenging to manipulate, and certain properties require more structure. Following are the primary definitions of normed spaces, including boundedness, differentiability of an operator, dual space, and bidual space. In addition, we will examine the well-known theorem of Eberlein and Šmulian, which allows us to characterize compact sets in terms of the weak sequential topology.

We begin with the definition of a linear and bounded operator.

Definition 2.17 (Bounded operator)**.** Let X and Y be normed spaces. A linear operator $T : X \to Y$ is called *bounded* if $T(A) \subseteq Y$ is bounded for every bounded set $A \subseteq X$.

We note that a linear operator that is bounded is also continuous and vice versa, see [3, Lemma 5.1]. Thus, the set $\mathscr{L}(X;Y)$ contains all the linear and bounded operators that map from X to Y. The operator norm of an operator $T \in \mathscr{L}(X;Y)$ is defined by

$$\|T\|_{\mathscr{L}(X;Y)} = \sup_{\|x\|_X \leq 1} \|Tx\|_Y.$$

Next, we define the concepts of isomorphisms and isometric operators, see [82, Definition 1.4.13].

Definition 2.18 (Isometric isomorphism)**.** Let X and Y be normed spaces and $T : X \to Y$ be linear. Then T is an *isomorphism* if it is injective and continuous and its inverse T^{-1} is continuous. The operator T is called *isometric* if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$. The space X is *(isometrically) embedded* in Y if there exists an (isometric) isomorphism from X to Y. The spaces X and Y are *(isometrically) isomorphic* if there is an (isometric) isomorphism from X to Y, i.e., $X \cong Y$.

Before, we introduced the notion of the dual space $X^* = \mathscr{L}(X;\mathbb{R})$ of a given space X. We can even take the dual of the dual space and call this space the bidual space, see [4, page 62] and the next definition.

Definition 2.19 (Biduality)**.** We define the *bidual space* X^{**} as the dual space of X^*. Furthermore, we define the biduality mapping $\widetilde{J} : X \to X^{**}$ between X and its bidual X^{**} by

$$\langle \widetilde{J}(x), x^* \rangle_{X^{**},X^*} = \langle x^*, x \rangle_{X^*,X} \qquad \text{with} \qquad \|\widetilde{J}(x)\|_{X^{**}} = \|x\|_X.$$

In general, this mapping $\widetilde{J}$ is only continuous. A normed space is called *reflexive* if $\widetilde{J}$ is an isomorphism, see [103, Definition 36.2]. The concept of reflexivity is very useful in the context of the differentiability of the distance functional, see Subsection 2.1.3.

But first, we have to discuss various notions of the differentiability of operators. More precisely, a directional derivative can be nonlinear and discontinuous in contrast to finite dimensions. However, there is the concept of Gâteaux and Fréchet derivatives, see the book [18, Definition 2.44].

Definition 2.20 (Directional, Gâteaux, Fréchet derivative)**.** Let X, Y be normed spaces and consider an operator $T : X \to Y$. Then T is called

(i) *directionally differentiable* at $x \in X$ in direction $h \in X$ if the following limit exists

$$\mathrm{d}T(x,h) = \lim_{t \downarrow 0} \frac{T(x+th) - T(x)}{t},$$

(ii) *Gâteaux differentiable* at $x \in X$ if the directional derivative $\mathrm{d}T(x,h)$ is linear and continuous in h, and the *Gâteaux derivative* is denoted by $T'(x)$, i.e.,

$$T'(x)h = \mathrm{d}T(x,h) \qquad \forall\, h \in X,$$

(iii) *Fréchet differentiable* at $x \in X$ if

$$T(x+h) = T(x) + T'(x)h + \mathrm{o}(\|h\|_X) \qquad \forall\, h \in X.$$

Every Fréchet differentiable operator is also Gâteaux differentiable, and the derivatives in this case are identical. The opposite is not generally true. Provided that it is clear after which variable the derivative is taken, we use the notation $T \mapsto T'$. Otherwise, we emphasize it with $T \mapsto T_x$.

Theorem 2.21 (Mean value theorem, cf., [28, Chapter 3.2, Theorem 1])**.** Let X be a normed space and $f : X \to \mathbb{R}$ be Fréchet differentiable. Furthermore, let $x, y \in X$. Then there exists an element $z \in \{(1-t)x + ty : 0 \leq t \leq 1\}$ such that

$$f(x) - f(y) = \langle f'(z), x - y \rangle_{X^*,X}.$$

Following this, we will introduce pseudoconvex functionals, which are closely related to directionally differentiable functionals. This concept was first introduced by Levi in 1910 in the topic of analytic functionals [77] and later on in 1965 by Mangasarian [80, 81] for optimization problems. Such functionals are included in the class of differentiable quasiconvex functionals, and interestingly, a local property such as a vanishing gradient implies a global optimality condition in the case of pseudoconvex functionals. In fact, every local minimizer corresponds to a global minimizer.

Classically, pseudoconvex functionals need to be differentiable in order to be well-defined. This definition was relaxed to subdifferentials and Dini derivatives in [6, 63]. These articles also relax the underlying spaces to Banach and linear spaces, respectively, and we state our definition also for infinite-dimensional spaces but for differentiable functions nonetheless.

Definition 2.22 (Pseudoconvex functional)**.** Let A be a convex subset of the normed space X. We call the directionally differentiable functional $f : A \to \mathbb{R}$ *pseudoconvex* if for any $x, y \in A$ it holds:

$$\mathrm{d}f(x, y - x) \geq 0 \implies f(y) \geq f(x).$$

In general, sums of pseudoconvex functionals are not pseudoconvex. The sum of a convex and a pseudoconvex functional is also not necessarily pseudoconvex, e.g., consider the sum of the functions $f(x) = x^3 + x$ and $g(x) = -x$.

In the following, we discuss some properties of the norm of a normed space X.

Lemma 2.23 (cf. [8, Lemma 2.35])**.** The norm of a normed space is lower semicontinuous with respect to the weak sequential topology.

Furthermore, we introduce the notion of *closed balls* around some element $x \in X$ with radius R as

$$\overline{B}_R^X(x) = \{y \in X : \|x - y\|_X \leq R\}.$$

We note that every ball in a normed space, open or closed, is convex, see [82, Proposition 1.3.12].

Theorem 2.24. Let X_i be a normed space for all $i \in [N]$. Then the product topology of $X = \Pi_{i=1}^N X_i$ is induced by the norm

$$\|x\|_X = \Big(\sum_{i=1}^N \|x^i\|_{X_i}^2 \Big)^{\frac{1}{2}} \tag{2.1}$$

Next, we discuss the peculiarities of the weak sequential topology and compare it with the weak topology. We have seen that weakly open sets are weakly sequentially open and the other direction does not hold (since the weak topology is not metrizable). Furthermore, the terms of a weakly sequentially open set coincide with an open set with respect to the weakly sequential topology. The theorem of Eberlein–Šmulian provides a characterization of weakly compact sets in the context of weakly sequentially compact sets. Moreover, it can be shown that weakly compact sets are the same as the compact sets with respect to the weakly sequential topology.

Theorem 2.25 (Eberlein–Šmulian, [82, Theorem 2.8.6]). Let A be a subset of a normed space. Then the following are equivalent:

(i) A is weakly compact.

(ii) A is weakly sequentially compact.

Corollary 2.26 (cf. [100]). Let X be a normed space. Then a subset $A \subseteq X$ is compact in the weak topology if and only if it is compact in the weak sequential topology.

Proof. Let A be a weakly compact set or compact set with respect to the weak topology, respectively. By the theorem of Eberlein–Šmulian, see Theorem 2.25, the set A is also weakly sequentially compact. Furthermore, let $C \subseteq A$ be a weakly sequentially closed subset. That means C is sequentially closed with respect to the weak topology or closed with respect to the weak sequential topology. Then C is weakly sequentially compact by the weak sequential compactness of A. Using the theorem of Eberlein–Šmulian once again, we get the weak compactness of C and thus, C is weakly closed, see Proposition 2.4. Since every weakly closed set is weakly sequentially closed, we have shown that $C \subseteq A$ is weakly sequentially closed if and only if it is weakly closed. Hence, on a weakly compact set the weak and the weak sequential topology coincide.

Let A be a compact set with respect to the weak sequential topology. Furthermore, let $\{A_i\}_{i\in I}$ be an open cover of A with respect to the weak topology. Since the weak sequential topology is finer than the weak topology, $\{A_i\}_{i\in I}$ is an open cover of A with respect to the weak sequential topology. Then there exists an finite index $J \subseteq I$ such that $A \subseteq \cup_{i\in J} A_i$. □

All together, we get the equivalence of all three notions of compactness in normed spaces.

2.1.3 Banach Spaces

Next, we add more structure to the topology and end up with Banach spaces. *Banach spaces* are the completion of normed spaces. A metric space (X, d) is called *complete* if any Cauchy sequence in it converges to a point in X. Any sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq X$ satisfying $\lim_{n,m\to\infty} d(x_n, x_m) = 0$ is a *Cauchy sequence*, see [111] and [82, Definition 1.2.12]. This subsection focuses on the various classes of Banach spaces and the properties that give the space more structure.

Definition 2.27 (Hilbert space). A Banach space with a complete inner product is called a *Hilbert space.*

This concept is quite robust, as it requires not only an inner product but also the completion of the Banach space with this inner product. Hilbert spaces have strong properties regarding projections and duality mappings. In contrast to this strong assumption, we construct special Banach spaces such that certain of these properties continue to hold. To this end, we will introduce Banach spaces that are uniformly smooth and uniformly convex.

First, however, we present some fundamental results in Banach and Hilbert spaces.

Theorem 2.28 (cf. [82, Theorem 1.4.8])**.** Let X and Y be normed spaces. Then the space $(\mathscr{L}(X;Y), \|\cdot\|_{\mathscr{L}(X;Y)})$ is a normed space. If Y is a Banach space, then $\mathscr{L}(X;Y)$ is a Banach space.

We note that the Riesz representation theorem guarantees that Hilbert spaces are reflexive, see [23, Proposition 5.1]. In case of Banach spaces, we have the following relationship of reflexivity between X and its dual.

Lemma 2.29 (cf. [3, Chapter 8.8])**.** Let X be a Banach space. If X is reflexive, then X^* is reflexive.

Next, we present some results on the existence of weakly and weakly-$*$ converging subsequences of bounded sequences. But first, we state the following classical result of Banach–Alaoglu on the compactness of the closed unit ball in the dual space.

Theorem 2.30 (Banach–Alaoglu, cf. [23, Theorem 3.16])**.** Let X be a Banach space. Then $\overline{B}_1^{X^*}(0)$ is compact in the weak-$*$ topology of X^*.

Hence, if we have a bounded sequence $\{x_k\}_{k\in\mathbb{N}}$ in X^*, then we obtain the existence of a weakly-$*$ convergent subsequence in X^*. Using reflexivity and a combination of the theorems of Eberlein–Šmulian and Banach–Alaoglu, see Theorem 2.25 and Theorem 2.30, we obtain the following result on weak convergence.

Lemma 2.31 (cf. [23, Theorem 3.18])**.** Let X be a reflexive Banach space and $\{x_k\}_{k\in\mathbb{N}}$ be bounded in X^*. Then there exists a weakly convergent subsequence $\{x_{k_l}\}_{l\in\mathbb{N}}$ in X^*.

Remark 2.32. Since X^* is reflexive, we obtain the existence of a weakly-$*$ convergent subsequence in X^*. In fact, let $\{x_k\}_{k\in\mathbb{N}} \subseteq X^*$ be bounded. Then it holds $x_{k_l} \rightharpoonup x$ in X^* as $l \to \infty$, i.e.,

$$\lim_{l\to\infty} \langle y, x_{k_l}\rangle_{X^{**},X^*} = \langle y, x\rangle_{X^{**},X^*} \qquad \forall\, y \in X^{**}.$$

By the definition of the reflexivity of X, we know that X^{**} is isomorphic to X with the isomorphism $\widetilde{J}_X : X \to X^{**}$, see again Definition 2.19. Thus, it directly yields the convergence

$$\lim_{l\to\infty} \langle x_{k_l}, z\rangle_{X^*,X} = \lim_{l\to\infty} \langle \widetilde{J}_X y, x_{k_l}\rangle_{X^{**},X^*} = \langle \widetilde{J}_X y, x\rangle_{X^{**},X^*} = \langle x, z\rangle_{X^*,X} \qquad \forall\, z \in X.$$

This corresponds to weak-$*$ convergence of the subsequence $\{x_{k_l}\}_{l\in\mathbb{N}}$ in X^*.

We note that the existence of a weakly-$*$ convergent subsequence in X^* is already guaranteed if X is a separable normalized space, as the next theorem states. In particular, no completeness and no reflexivity is needed, which is particularly useful for the spaces $X = L^1(0,1)$ or $X = C([0,1])$.

Theorem 2.33 (Sequential Banach–Alaoglu, cf. [46, Section 5, Exercise 50])**.** Let X be a separable normed space. Then $\overline{B}_1^{X^*}(0)$ is sequentially compact in the weak-$*$ topology of X^*. Equivalently, if $\{x_k\}_{k\in\mathbb{N}}$ is a bounded sequence in X^*, then there exists a weakly-$*$ converging subsequence $\{x_{k_l}\}_{l\in\mathbb{N}}$ in X^*.

We define the important concept of an adjoint operator in normed spaces, see the book [79, Section 6.5], and in the special case if Hilbert spaces are involved.

Definition 2.34 (Adjoint operator)**.** Let X, Y be normed spaces and let $T : X \to Y$ be a linear and bounded operator. Then the *adjoint operator* of T is denoted by $T^* : Y^* \to X^*$ and fulfills

$$\langle T^*y^*, x\rangle_{X^*,X} = \langle y^*, Tx\rangle_{Y^*,Y} \qquad \forall\, x \in X,\ y^* \in Y^*.$$

Furthermore, if X and Y are Hilbert spaces, then $T' : Y \to X$ is called the *Hilbert adjoint* if it holds

$$(T'y, x)_X = (y, Tx)_Y \qquad \forall\, x \in X,\ y \in Y.$$

We note that the Hilbert space adjoint can be related to the Banach space adjoint via

$$T' = R_X^{-1} T^* R_X.$$

In this formula, $R_X : X \to X^*$ denotes the bijective Riesz mapping in Hilbert spaces, which is defined by $R_X x = (\cdot, x)_X \in X^*$.

We mention the following generalization of the open mapping theorem that applies to set-valued functions in Banach spaces.

Theorem 2.35 (Generalized open mapping theorem, cf. [91, 92])**.** Let X and Y be Banach spaces. Furthermore, let $T : X \rightrightarrows Y$ be a convex and closed set-valued operator. Moreover, let $y \in \operatorname{int}(\operatorname{range} T)$. Then for every $x \in T^{-1}(y)$, it follows that $y \in \operatorname{int} T(B_R^X(x))$ for all $R > 0$.

Next, we define the concept of strictly convex spaces, in which the closed unit ball is a strictly convex set, see [30, Chapter II, Definition 1.1]. In other words, given any two distinct points x and y on the unit sphere, the segment connecting x and y only meets the sphere at x and y. We observe that every Hilbert space is strictly convex.

Definition 2.36 (Strictly convex space)**.** A Banach space X is *strictly convex* if for all $x, y \in X$ with $x \neq y$ and $\|x\|_X = \|y\|_X = 1$ there holds

$$\|tx + (1-t)y\|_X < 1$$

for all $t \in (0, 1)$.

The following characterizations apply to strictly convex Banach spaces.

Proposition 2.37 (cf. [30, Chapter II, Proposition 1.2, Proposition 1.6])**.** Let X be a Banach spaces. Then following statements are equivalent:

(i) X is strictly convex.

(ii) For all $x, y \in X$ with $x \neq y$ and $\|x\|_X = \|y\|_X = 1$ there holds $\|x + y\|_X < 2$.

(iii) The mapping $x \mapsto \|x\|_X^2$ is strictly convex.

Next, we will discuss uniformly convex Banach spaces, which serve as examples of typical reflexive Banach spaces. In fact, every Hilbert space and every $L^p(0,1)$ space, $p \in (1,\infty)$, is uniformly convex. This concept was first used by Clarkson in 1936, see [31], and we will see that such spaces are closely related to the notion of strict convexity. This definition is based on [30, Chapter II, Definition 2.1].

Definition 2.38 (Uniformly convex space)**.** A Banach space X is called *uniformly convex* if for all $\{x_k\}_{k\in\mathbb{N}} \subseteq X$, $\{y_k\}_{k\in\mathbb{N}} \subseteq X$ with $\|x_k\|_X = \|y_k\|_X = 1$, $k \in \mathbb{N}$, and $\|x_k + y_k\|_X \to 2$ as $k \to \infty$, it follows $\|x_k - y_k\|_X \to 0$ as $k \to \infty$.

Again, we present some equivalences of uniform convex spaces that can also serve as alternative definitions.

Proposition 2.39 (cf. [30, Chapter II, Proposition 2.3, Proposition 2.11])**.** Let X be a Banach space. Then the following statements are equivalent:

(i) X is uniformly convex.

(ii) For all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in X$ with $\|x\|_X = \|y\|_X = 1$ and $\|y - x\|_X \geq \varepsilon$ there holds $\|y + x\|_X \leq 2(1-\delta)$.

(iii) For all $x \in X$ with $\|x\|_X = 1$ and for all $\varepsilon > 0$, there exists $\delta(x) > 0$ such that for all $y \in X$ with $\|y\|_X = 1$ and $\|y - x\|_X \geq \varepsilon$ there holds $\|y + x\|_X \leq 2(1 - \delta(x))$.

(iv) $f(x) = \frac{1}{2}\|x\|_X^2$ is *uniformly strictly convex*, i.e., f is convex and, for all $\varepsilon > 0$,

$$\inf\{f(x) - 2f(\tfrac{x+y}{2}) + f(y) : \|x\|_X = 1, \|y - x\|_X \geq \varepsilon\} > 0.$$

The following proposition relates uniformly convex Banach spaces to strictly convex Banach spaces.

Proposition 2.40 (cf. [30, Chapter II, Proposition 2.7, Proposition 2.8])**.** Any uniformly convex Banach space X is strictly convex. Furthermore, for a sequence $\{x_k\}_{k\in\mathbb{N}}$ in X the following implication holds:

$$x_k \rightharpoonup \bar{x} \text{ in } X, \quad \|x_k\|_X \to \|\bar{x}\|_X \quad \text{as } k \to \infty \qquad \Longrightarrow \qquad x_k \to \bar{x} \text{ in } X \text{ as } k \to \infty.$$

The famous Milman–Pettis theorem is stated next, which was independently proved by Milman and Pettis.

Theorem 2.41 (Milman–Pettis, cf. [30, Chapter II, Theorem 2.9])**.** A uniformly convex Banach space is reflexive.

We introduce uniformly smooth spaces, which are closely related to uniformly convex spaces through their dual spaces. In fact, we will demonstrate that a space is uniformly smooth only if its dual is uniformly convex. However, we will begin with the definition of uniform smoothness, see [29, Definition 2.4].

Definition 2.42 (Uniformly smooth space)**.** A normed space X is if, given $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x, y \in X$ with $\|x\|_X = 1$ and $\|y\|_X \leq \delta$, then

$$\|x + y\|_X + \|x - y\|_X < 2 + \varepsilon \|y\|_X.$$

An alternative characterization can again be stated, as the following result shows.

Proposition 2.43 (cf. [29, Theorem 2.5])**.** A normed space X is uniformly smooth if and only if

$$\lim_{t \downarrow 0} \sup_{\|x\|_X = \|y\|_X = 1} \frac{\|x + ty\|_X + \|x - ty\|_X - 2}{2t} = 0.$$

One of the main theorems in the theory of uniformly smooth spaces is the following.

Theorem 2.44 (cf. [30, Chapter II, Theorem 2.13], [29, Theorem 2.10], [82, Theorem 5.5.12])**.** Let X be a Banach space. Then X is uniformly convex if and only if X^* is uniformly smooth, and X is uniformly smooth if and only if X^* is uniformly convex.

Now, an easy consequence of this theorem and the Milman–Pettis theorem, see Theorem 2.41, is as follows.

Corollary 2.45 (cf. [30, Chapter II, Corollary 2.15], [29, Corollary 2.11])**.** Every uniformly smooth Banach space is reflexive.

Remark 2.46. We note that the product space X of uniformly convex Banach spaces X_i, $i \in I$ finite, is a uniformly convex Banach space, see [31, Theorem 1]. The product space of reflexive Banach spaces is reflexive. Furthermore, we know that a Banach space is uniformly smooth if the dual space is uniformly convex. And hence, the product space is also uniformly smooth, uniformly convex and reflexive.

We summarize the stated results in the following graph for a given Banach space X.

$$\begin{array}{ccccc} & & X \text{ Hilbert space} & & \\ & & \Downarrow & & \\ X \text{ reflexive} & \Longleftarrow & X \text{ uniformly convex} & \Longleftrightarrow & X^* \text{ uniformly smooth} \\ & & \Downarrow & & \\ & & X \text{ strictly convex} & & \end{array}$$

2.2 Operator Theory

In this section, special operators such as monotone operators, projections, and duality operators are discussed. Typically, such operators are nonlinear and do not belong to the class $\mathscr{L}(X;Y)$, but we nevertheless provide interesting and practical results.

2.2.1 Monotone Operators

Pseudomonotone operators play an important role in the analysis of nonmonotone partial differential equations such as

$$\operatorname{div}(|\nabla u|^{p-2}\nabla u) + su = f$$

for $s < 0$ and a given function f, which is assumed to be sufficiently regular. Such partial differential equations cannot be solved by the theory of monotone operators. But it can be checked that the famous Browder–Minty theorem can be adapted to the more general pseudomonotone operators, see, e.g., [96].

One can introduce pseudomonotone operators on topological vector spaces, see, e.g., the works [22, 61, 113]. Typically, the concept is used for partial differential equations with solution spaces that are reflexive Banach spaces, e.g., see the monograph [48]. For applications of pseudomonotone operators to game theory in the setting of Euclidean spaces, we refer to the book [67]. Additionally, we introduce the concept of hemicontinuous and monotone operators on reflexive Banach space, which are closely connected to pseudomontone operators. In the following definition, all terms are defined on a reflexive Banach space.

Definition 2.47 ((Pseudo)monotonicity and hemicontinuity)**.** Let X be a reflexive Banach space. An operator $T : X \to X^*$ is called

(i) *pseudomonotone* if the conditions $x_k \rightharpoonup \bar{x}$ weakly in X as $k \to \infty$ and

$$\liminf_{k\to\infty} \langle T(x_k), \bar{x} - x_k\rangle_{X^*,X} \geq 0$$

imply

$$\limsup_{k\to\infty} \langle T(x_k), y - x_k\rangle_{X^*,X} \leq \langle T(\bar{x}), y - \bar{x}\rangle_{X^*,X} \qquad \forall\, y \in X,$$

(ii) *hemicontinuous* if the functional $t \mapsto \langle T(x + ty), z\rangle_{X^*,X}$ is continuous on $[0, 1]$ for all $x, y, z \in X$,

(iii) *monotone* if it holds $\langle Tx - Ty, x - y\rangle_{X^*,X} \geq 0$ for all $x, y \in X$.

The next theorem establishes relationships between the previously defined operators. In particular, every completely continuous operator, see Definition 2.15, is pseudomonotone and we will heavily exploit this property later on.

Proposition 2.48 (cf. [113, Proposition 27.6])**.** Let X be a reflexive Banach space and $T : X \to X^*$ and $S : X \to X^*$ be operators. Then it holds:

(i) If T is monotone and hemicontinuous, then T is pseudomonotone.

(ii) If T is completely continuous, then T is pseudomonotone.

(iii) If T and S are pseudomonotone, then $T + S$ is pseudomonotone.

(iv) If T is monotone and hemicontinuous and S is completely continuous, then $T + S$ is pseudomonotone.

Since we have already found a connection between pseudomonotone and completely continuous operators, we also provide the following result on the Fréchet derivative of a completely continuous operator.

Theorem 2.49 (cf. [38, Proposition 8.2])**.** Let X and Y be Banach spaces. Furthermore, let $T \in \mathscr{L}(X;Y)$ be Fréchet differentiable in $x \in X$. If T is completely continuous, then $T'(x) \in \mathscr{L}(X;Y)$ is completely continuous.

Thus, we know that the Fréchet derivative of a completely continuous operator is pseudomonotone.

2.2.2 Projection Operator

We state some properties of the projection onto a convex and closed set. This operator will become particularly useful when studying the projected descent method later on in Section 6.1. But first, we begin with its definition, which we base on [8, p.16] and [23, Exercise 3.32].

Definition 2.50 (Distance functional and projection)**.** Let X be a normed space and $A \subseteq X$ be a closed subset of X. Then, the *distance functional* is defined by

$$\operatorname{dist}(x, A) = \inf_{y \in A} \|x - y\|_X.$$

In case of a uniformly convex Banach space X, the *projection* $P_A(x)$ of some element $x \in X$ to the convex and closed subset A is given by

$$\|x - P_A(x)\|_X = \operatorname{dist}(x, A).$$

We note that the infimum in the definition of the distance functional can be replaced by the minimum if we additionally assume that A is a convex subset of X. Often, the projection is already defined with this assumption, see [59].

Moreover, we note that we require the assumption of a uniformly convex Banach space in order to define the projection in the stated manner. Theoretically, one could also define the projection on general Banach spaces, but in this case, the projection of a given element is a set and one has to deal with so-called Chebyshev subsets in order to achieve single-valued projection operators. The neat thing is that any convex and closed subset of a uniformly convex Banach space is automatically Chebyshev, see also [110].

We state the following lemma, which summarizes some useful results of the projection operator that we will require later on when studying the projected gradient method.

Lemma 2.51 (cf. [59, Lemma 1.10])**.** Let $A \subseteq X$ be a nonempty, convex, and closed subset of the Hilbert space X. Then the following holds:

(i) for all $u \in A$ and $d \in X$, the function $t \mapsto \frac{1}{t}\|P_A(u + td) - u\|_X$ is nonincreasing for all $t > 0$.

(ii) for all $u, v \in X$ it holds the equivalency

$$v = P_A(u) \qquad \Longleftrightarrow \qquad v \in A, \quad (u - v, w - v)_X \leq 0 \qquad \forall\, w \in A.$$

Lemma 2.52 (cf. [1, Remark 3.4], [110, Theorem 2], [51, Proposition 3.2])**.** Let X be a uniformly smooth and uniformly convex Banach space. If A is a convex and closed subset of X, then the projection $P_A : X \to A$ is continuous on X, and uniformly continuous on any bounded subset of X.

Results on the (uniform) continuity of the projection P_A can be found in [1]. In particular, if X is uniformly convex, then P_A is continuous, and if X is uniformly smooth and uniformly convex, then P_A is uniformly continuous on bounded sets. By the smoothness properties of the norm and these results, we can obtain continuous differentiability of $\operatorname{dist}^2(\cdot, A)$ and uniform continuity of its derivative on bounded sets under suitable conditions posed on X.

2.2.3 Duality Mapping

This section introduces the duality operator that maps a given space X to its dual X^*, see [30, Definition 4.1]. We discuss several properties of this mapping and we will build a connection to the derivative of the distance functional that we have seen in Definition 2.50. This section is based on [29, Section 3] and [113, Section 32].

Definition 2.53 (Duality mapping)**.** Let X be a normed space. The *duality mapping* of X is denoted by $J_X : X \to X^*$ and it is given by the element $x^* = J_X(x) \in X^*$ for $x \in X$, which is uniquely defined by the properties

$$\langle x^*, x\rangle_{X^*,X} = \|x\|_X^2 \qquad \text{and} \qquad \|x^*\|_{X^*} = \|x\|_X.$$

Even though we have defined the duality mapping on normed spaces, it exhibits several interesting properties if we assume more structure on the underlying space. In fact, we have the following result.

Proposition 2.54 (cf. [113, Proposition 32.22])**.** Let X be a reflexive Banach space with strictly convex dual space X^*. Then it holds:

(i) The duality map J_X is single-valued, monotone, and bounded.

(ii) If X is additionally a strictly convex Banach space, then the duality map J_X is bijective.

(iii) If X^* is uniformly convex, then J_X is hemicontinuous and uniformly continuous on bounded sets.

If we assume that X is a uniformly smooth and uniformly convex Banach space, then the duality map is single-valued, odd, demicontinuous, bounded, bijective, maximal monotone, coercive, positively homogeneous, strictly monotone, and uniformly continuous on bounded sets. These properties can be partially shown under lower assumptions, see [113, Proposition 32.22].

We note that the duality mapping J_X is usually nonlinear and that the single-valuedness and further properties only hold due to higher requirements on X as the following equivalence statement shows.

Proposition 2.55 (cf. [29, Proposition 3.6, Proposition 3.7])**.** Let X be a Banach space. The duality mapping J_X is linear and single-valued if and only if X is a Hilbert space.

In the case of a Hilbert space X the duality mapping J_X corresponds to the Riesz operator R_X, which we introduced in Definition 2.34. The inverse operator of J_X can be expressed by the duality map J_{X^*} of X^*, as we see in the next proposition.

Proposition 2.56 (cf. [113, Proposition 32.22])**.** Let X be a strictly convex reflexive Banach space with a strictly convex dual space. It holds $J_X^{-1} = \widetilde{J}^{-1} J_{X^*}$, where $\widetilde{J} : X \to X^{**}$ is an isometry between X and the bidual space X^{**}. If we identify X^{**} with X, then it reduces to $J_X^{-1} = J_{X^*}$.

Proof. By definition of the isometry, it holds

$$\langle \widetilde{J}x, x^* \rangle_{X^{**},X^*} = \langle x^*, x \rangle_{X^*,X} \qquad \forall\, x \in X, \quad x^* \in X^*,$$

and $\widetilde{J}$ is injective and isometric. Let $x = \widetilde{J}^{-1} J_{X^*}(x^*) \in X$. Then we have

$$\|x^*\|_{X^*} = \|J_{X^*}(x^*)\|_{X^{**}} = \|\widetilde{J}x\|_{X^{**}} = \|x\|_X, \tag{2.2}$$

where we used the definition of the duality mapping $J_{X^*} : X^* \to X^{**}$ in the first equality and the definition of an isometry in the last one. Next, we use the properties of the isometry $\widetilde{J}$ and its inverse $\widetilde{J}^{-1}$ to obtain

$$\langle x^*, x \rangle_{X^*,X} = \langle x^*, \widetilde{J}^{-1} J_{X^*}(x^*) \rangle_{X^*,X} = \langle \widetilde{J}\widetilde{J}^{-1} J_{X^*}(x^*), x^* \rangle_{X^{**},X^*} = \langle J_{X^*}(x^*), x^* \rangle_{X^{**},X^*}.$$

Furthermore, we get

$$\langle J_{X^*}(x^*), x^* \rangle_{X^{**},X^*} = \|x^*\|_{X^*}^2 = \|J_{X^*}(x^*)\|_{X^{**}}^2 = \|x\|_X^2$$

by the definition of the duality mapping J_{X^*} and (2.2). Thus, we arrive at

$$\langle x^*, x \rangle_{X^*,X} = \|x\|_X^2.$$

Finally, we obtain $x^* = J_X(x) = J_X(\widetilde{J}^{-1} J_{X^*}(x^*))$ for all $x^* \in X^*$ and by the bijectivity of J_X it yields the result. □

Proposition 2.57 (cf. [113, Proposition 32.22])**.** If X is a reflexive Banach space with strictly convex dual X^*, then $\|\cdot\|_X$ is Gâteaux differentiable with derivative $x \mapsto \frac{J_X(x)}{\|x\|_X}$ on $X \backslash \{0\}$. If X is additionally uniformly smooth, then the norm $\|\cdot\|_X$ is Fréchet differentiable on $X \backslash \{0\}$ and its derivative is uniformly continuous on bounded subsets that does not contain some neighborhood of 0.

It is clear that if $\|\cdot\|_X$ is Fréchet differentiable outside 0, then $\|\cdot\|_X^2$ is Fréchet differentiable everywhere since $\|\cdot\|_X^2$ is Fréchet differentiable at 0. Moreover, $\|\cdot\|_X^2$ is Gâteaux differentiable with derivative $x \mapsto 2J_X(x)$. If J_X is uniformly continuous on every bounded set, then the same holds true for the derivative of $\|\cdot\|_X^2$, since $(\|\cdot\|_X^2)'(x) = 2J_X(x)$.

Proposition 2.58 (cf. [108, Proposition 5, Proposition 7])**.** If A is a convex and closed subset of the Banach space X, then the squared distance functional $\operatorname{dist}^2(\cdot, A)$ is differentiable on A and it holds for all $x \in A$

$$\begin{aligned} \big\langle \operatorname{dist}_x^2(x, A), h \big\rangle_{X^*,X} &= \langle (\|\cdot\|_X^2)'(x - P_A(x)), h \rangle_{X^*,X} \\ &= 2\big\langle J_X(x - P_A(x)), h \big\rangle_{X^*,X} \qquad \forall\, h \in X. \end{aligned}$$

2.3 Theory of Constrained Optimization

This chapter focuses on the theoretical background of constrained optimization in general Banach spaces and its application to the equilibrium concept of games. In this context, we also introduce the augmented Lagrangian method, which is in the class of penalty methods.

We start with optimization problems in Banach spaces and discuss the well-known KKT conditions and give some results on CQs and their consequences on the existence of minimizers.

2.3.1 Optimality Conditions

In the following, we consider the minimization problem

$$\min_{u\in U} f(u) \qquad \text{s.t.} \qquad G(u) \in K, \quad u \in \mathcal{X}, \tag{2.3}$$

with $f : U \to \mathbb{R}$, $G : U \to X$, $K \subseteq X$ and $\mathcal{X} \subseteq U$. Here, U and X are normed spaces. We say that an element $u \in U$ is *feasible* if $u \in \mathcal{X}$ and $G(u) \in K$. We denote the feasible set by $F = \mathcal{X} \cap G^{-1}(K)$.

But first, we introduce the important concept of tangent and normal cones, see the book [35, Definition 2.1].

Definition 2.59 (Bouligand tangent cone and normal cone)**.** The *Bouligand tangent cone* at $\bar{u} \in U$ is defined as the empty set if $\bar{u} \notin \mathcal{X}$, and otherwise as

$$T_{\mathcal{X}}(\bar{u}) = \big\{d \in U : \exists\, \eta_k > 0,\ \{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X} : \lim_{k\to\infty} u_k = \bar{u}, \lim_{k\to\infty} \eta_k(u_k - \bar{u}) = d \text{ in } U\big\}. \tag{2.4}$$

The *normal cone* $N_{\mathcal{X}}(\bar{u})$ is defined as the polar cone of $T_{\mathcal{X}}(\bar{u})$.

The following lemma lets us rewrite the relevant cones, which makes them easier to handle.

Lemma 2.60 (cf. [35, Proposition 2.1])**.** For a convex and closed set $\mathcal{X} \subseteq U$, the tangent and normal cones at $\bar{u} \in U$ can be written as

$$\begin{aligned} T_{\mathcal{X}}(\bar{u}) &= \operatorname{cl}\big\{d \in U : d = \eta(z - \bar{u}),\ z \in \mathcal{X},\ \eta > 0\big\}, \\ N_{\mathcal{X}}(\bar{u}) &= \big\{d \in U^* : \langle d, z - \bar{u}\rangle_{U^*,U} \le 0 \quad \forall\, z \in \mathcal{X}\big\}. \end{aligned} \tag{2.5}$$

We say that the *first-order necessary optimality condition* holds in a point $\bar{u}$ for the minimization problem (2.3) if the following inequality is fulfilled:

$$\langle f'(\bar{u}), d\rangle_{U^*,U} \ge 0 \qquad \forall\, d \in T_F(\bar{u}).$$

If there exists some $d \in T_F(\bar{u})$, then the tangent cone is nonempty and $\bar{u}$ is feasible. In this case, we will speak about first-order optimality conditions including feasibility. If f is continuously differentiable in $\bar{u}$ and $\bar{u}$ is a local solution to the problem (2.3), then these conditions must be true and are indeed necessary, see [17, Lemma 3.7]. In fact, these conditions build the basis of the famous KKT conditions.

We also state the following sufficient conditions that directly yield a minimum in case of a simplified minimization problem with strict assumptions such as an underlying reflexive Banach space. Nevertheless, we will later on apply it to achieve the existence of a minimizer in a reduced problem.

Lemma 2.61 (cf. [23, Corollary 3.23])**.** Let U be a reflexive Banach space and assume that $\mathcal{X} \subseteq U$ is a nonempty, convex, closed, and bounded subset of U. Furthermore, let $f : \mathcal{X} \to \mathbb{R}$ be a convex and lower semicontinuous functional. Then there exists a $x_0 \in \mathcal{X}$ such that $f(x_0) = \min_{x \in \mathcal{X}} f(x)$.

In order to reformulate the Bouligand tangent cone (2.4) in terms of the tangent cone (2.5) with respect to $\mathcal{X}$ and K, we introduce the concept of the well-known Robinson constraint qualification (RCQ). Moreover, to validate the feasibility of (weak) limit points in our global convergence analysis later on, we define a generalization of the RCQ to consider potentially infeasible points, the so-called extended RCQ, which has been defined in [27, Definition 2.2].

Definition 2.62 (Extended Robinson constraint qualification)**.** Let $u \in U$ be an arbitrary, not necessarily feasible point, and G be continuously differentiable at the point u. We say that the *extended Robinson constraint qualification (ERCQ)* holds for problem (2.3) in u if

$$0 \in \operatorname{int}\big(G(u) + G_v(u)(\mathcal{X} - u) - K\big). \tag{2.6}$$

If u is an admissible point, then we refer to this condition as RCQ.

We note that the ERCQ is equivalent to the existence of some $R > 0$ such that for all $x \in \overline{B}_R^X(0)$ there are elements $y \in K$ and $z \in \mathcal{X}$ with

$$x = G(u) + G_v(u)(z - u).$$

If the RCQ holds in some element u, then the tangent cone of the feasible set $F = \mathcal{X} \cap G^{-1}(K)$ at u reads

$$T_F(u) = \big\{d \in T_{\mathcal{X}}(u) : G_v(u)d \in T_K(G(u))\big\},$$

see [18, Corollary 2.91]. By this result, we can interpret the well-known KKT conditions geometrically.

Definition 2.63 (KKT pair and Lagrangian functional)**.** Let $f : U \to \mathbb{R}$ and $G : U \to X$ be continuously differentiable at the point u. We call a tuple $(u, \lambda) \in U \times X^*$ a *KKT pair* for the optimization problem (2.3) if it holds

$$\begin{aligned} -\big[(L_X)_v(v,\lambda)\big]_{|v=u} &\in N_{\mathcal{X}}(u), \\ \lambda &\in N_K(G(u)), \end{aligned}$$

with the *Lagrangian functional*

$$L_X(v, \lambda) = f(v) + \langle \lambda, G(v) \rangle_{X^*,X}.$$

Note that the Bouligand tangent cone and the normal cone are empty if $u \notin \mathcal{X}$ and thus, KKT points are feasible, see also [47]. Moreover, the KKT conditions are first-order optimality conditions using the Lagrangian functional.

Since we only assume first-order differentiability of our constraints and objective functionals, we are only interested in necessary optimality conditions instead of sufficient conditions. In the case of second-order differentiability and sufficient conditions in infinite-dimensional normed spaces, we refer to the book [90].

2.3.2 Equilibrium Concepts for Games

Let $N \in \mathbb{N}$ be the finite number of players. For each $i \in [N] = \{1, \dots, N\}$ the i-th player's strategy (or control) is denoted by $u^i \in U_i$ and we write $U = \prod_{i\in[N]} U_i$ for the strategy set of all players. Here, the individual control spaces U_i, $i \in [N]$, are normed spaces. In addition, the symbol $u^{-i} \in U_{-i}$ represents the tuple of all strategies excluding that of player i. We employ the notation $u = (u^i, u^{-i})$ to emphasize the i-th player, but we do not reorder the tuple. For each $i \in [N]$, the i-th player solves the minimization problem

$$\min_{v^i \in F_i(u^{-i})} \theta_i(v^i, u^{-i}), \tag{2.7}$$

where $F_i(u^{-i}) \subseteq U_i$ denotes the feasible set of player i that depends on the chosen strategies of the other players. If this feasible set is independent of the selected strategies, the collection of problems (2.7), $i \in [N]$, is called a NEP, while in the general case it is called a GNEP. In the vast majority of published works, the objective functionals $\theta_i : U \to \mathbb{R}$, $i \in [N]$, are assumed to be convex and continuously differentiable in the i-th component. In this thesis, this assumption is relaxed. Specifically, we will employ solution maps with a regularization term $\frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2$ added to the cost functional θ_i. Here, $\iota_{H_i} : U_i \to H_i$ is a linear and continuous operator that maps to some given normed space H_i. Detailed requirements on all spaces and problem data will be given in Chapter 3 in Assumption 3.1, Assumption 3.3 and Assumption 3.4.

We are interested in Nash equilibria and normalized equilibria, as well as points that satisfy their first-order necessary optimality conditions.
The N-tuple $u = (u^1, \dots, u^N) \in U$ is a

- *Nash equilibrium* for the GNEP consisting of (2.7) if and only if, for all $i \in [N]$, given u^{-i}, the vector $v^i = u^i$ solves the i-th player's optimization problem (2.7), i.e., $u \in F(u) = \prod_{i\in[N]} F_i(u^{-i})$ fulfills

$$\theta_i(u^i, u^{-i}) \le \theta_i(v^i, u^{-i}) \qquad \forall\, v^i \in F_i(u^{-i}), \quad i \in [N].$$

  Equivalently, we can state: If $S_i(u^{-i})$ denotes the set of all solutions to the i-th player's optimization problem (2.7), then u is a Nash equilibrium if and only if it holds $u^i \in S_i(u^{-i})$ for all $i \in [N]$.

- *quasi-Nash equilibrium* if and only if, for all $i \in [N]$, given u^{-i}, the vector $v^i = u^i$ solves the first-order optimality conditions of the i-th player's optimization problem

(2.7). If $F_i(u^{-i})$ is convex and θ_i is differentiable with respect to its i-th component for all $i \in [N]$, then the first-order necessary conditions read (which is sufficient if θ_i is additionally convex in the i-th component for all $i \in [N]$)

$$u^i \in F_i(u^{-i}), \qquad \langle (\theta_i)_{v^i}(u^i, u^{-i}), z^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, z^i \in F_i(u^{-i}), \quad i \in [N]. \tag{2.8}$$

If the feasible set $F_i(u^{-i})$ is of the form

$$F_i(u^{-i}) = \{v^i \in U_i : (v^i, u^{-i}) \in \mathcal{X}\}, \tag{2.9}$$

with $\mathcal{X} \subseteq U$ nonempty and convex, then the point $u \in U$ is a quasi-Nash equilibrium if

$$u \in \mathcal{X}, \qquad \langle (\theta_i)_{v^i}(u), z^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, z^i \in U_i \text{ with } (z^i, u^{-i}) \in \mathcal{X}, \quad i \in [N].$$

In the case that $\mathcal{X}$ has product structure, i.e., if $F_i(u^{-i}) = \mathcal{X}_i$, the first-optimality conditions simplify to

$$u \in \mathcal{X}, \qquad \langle (\theta_i)_{v^i}(u), z^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, z^i \in \mathcal{X}_i, \quad i \in [N].$$

If the feasible set $F_i(u^{-i})$ is of the form

$$F_i(u^{-i}) = \{v^i \in U_i : (v^i, u^{-i}) \in \mathcal{X},\ G(v^i, u^{-i}) \in K\}, \tag{2.10}$$

then $v^i \in U_i$ satisfies the optimality conditions if there exists some element $\lambda^i \in X^*$ such that

$$\begin{aligned} &(v^i, u^{-i}) \in \mathcal{X}, \quad G(v^i, u^{-i}) \in K, \quad \lambda^i \in N_K(G(v^i, u^{-i})), \\ &\langle (\theta_i)_{v^i}(v^i, u^{-i}) + G_{v^i}(v^i, u^{-i})^* \lambda^i, z^i - v^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, z^i \in U_i \text{ with } (z^i, u^{-i}) \in \mathcal{X}. \end{aligned} \tag{2.11}$$

Summarizing, one needs to find a point u such that $v^i = u^i$ satisfies (2.11) for all $i \in [N]$. We can characterize a quasi-Nash equilibrium u by the system that we obtain by inserting u^i for v^i into (2.11) and collecting all resulting systems for $i \in [N]$. In fact, one obtains

$$\begin{aligned} &u \in \mathcal{X}, \quad G(u) \in K, \quad \lambda^i \in N_K(G(u)), \\ &\langle (\theta_i)_{v^i}(u) + G_{v^i}(u)^* \lambda^i, z^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, z^i \in U_i \text{ with } (z^i, u^{-i}) \in \mathcal{X}. \end{aligned} \tag{2.12}$$

In the case $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$, the system simplifies to

$$\begin{aligned} &u \in \mathcal{X}, \quad G(u) \in K, \quad \lambda^i \in N_K(G(u)), \\ &\langle (\theta_i)_{v^i}(u) + G_{v^i}(u)^* \lambda^i, z^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, z^i \in \mathcal{X}_i. \end{aligned}$$

- *normalized equilibrium* for GNEP consisting of (2.7) with the feasible set (2.9) if and only if, given u, the vector $v = u$ solves the problem

$$\min_{v \in \mathcal{X}} \sum_{i \in [N]} \theta_i(v^i, u^{-i}), \tag{2.13}$$

or in other words, it holds

$$\sum_{i\in[N]} \theta_i(u^i, u^{-i}) \le \sum_{i\in[N]} \theta_i(v^i, u^{-i}) \qquad \forall\, v \in \mathcal{X}. \tag{2.14}$$

Equivalently, we can formulate it as follows. If, for a given u, $S(u)$ denotes the set of all solutions to (2.13), then u is a normalized equilibrium if and only if it holds $u \in S(u)$.

- *variational equilibrium* if and only if, given u, the vector $v = u$ satisfies the first-order optimality conditions of (2.13). If $\mathcal{X} \subseteq U$ is nonempty and convex and θ_i is differentiable, then u satisfies the necessary first-order optimality conditions for a normalized equilibrium if the following variational inequality (VI) is valid

$$u \in \mathcal{X}, \qquad \sum_{i\in[N]} \langle (\theta_i)_{v^i}(u^i, u^{-i}), z^i - u^i\rangle_{U_i^*,U_i} \ge 0 \qquad \forall\, z \in \mathcal{X}, \tag{2.15}$$

and feasibility is ensured. For a feasible set of the form (2.10), v satisfies the stated first-order optimality conditions of (2.13), given u, if there exists a Lagrangian multiplier $\lambda \in X^*$ such that it holds

$$\begin{aligned} &v \in \mathcal{X}, \quad G(v) \in K, \quad \lambda \in N_K(G(v)),\\ &\sum_{i\in[N]}\left[\langle (\theta_i)_{v^i}(v^i, u^{-i}), z^i - v^i\rangle_{U_i^*,U_i}\right] + \langle G_v(v)^*\lambda, z - v\rangle_{U^*,U} \ge 0 \quad \forall\, z \in \mathcal{X}. \end{aligned} \tag{2.16}$$

Since u is a variational equilibrium if and only if $v = u$ satisfies (2.16), we obtain that u is a variational equilibrium if and only if it holds

$$\begin{aligned} &u \in \mathcal{X}, \quad G(u) \in K, \quad \lambda \in N_K(G(u)),\\ &\sum_{i\in[N]}\left[\langle (\theta_i)_{v^i}(u), z^i - u^i\rangle_{U_i^*,U_i}\right] + \langle G_v(u)^*\lambda, z - u\rangle_{U^*,U} \ge 0 \qquad \forall\, z \in \mathcal{X}. \end{aligned} \tag{2.17}$$

In order to relate such equilibria to KKT points, we define the Lagrangian functionals of (2.7) and (2.13) with the feasible set (2.10) by

$$\begin{aligned} L_X^i(v^i, u^{-i}, \lambda^i) &= \theta_i(v^i, u^{-i}) + \langle \lambda^i, G(v^i, u^{-i})\rangle_{X^*,X},\\ L_X(v, \lambda; u) &= \sum_{i\in[N]} \left[\theta_i(v^i, u^{-i})\right] + \langle \lambda, G(v)\rangle_{X^*,X}, \end{aligned} \tag{2.18}$$

where $u^{-i} \in U_{-i}$ and $u \in U$ are the corresponding parameters. Then the VIs (2.12) and (2.17) can be formulated with the corresponding Lagrangian. Indeed, for all directions $h \in U$ we obtain for the derivatives

$$\begin{aligned} \left\langle \left[(L_X^i)_{v^i}(v^i, u^{-i}, \lambda^i)\right]_{|v^i=u^i}, h^i\right\rangle_{U_i^*,U_i} &= \langle (\theta_i)_{v^i}(u), h^i\rangle_{U_i^*,U_i} + \langle G_{v^i}(u)^*\lambda^i, h^i\rangle_{U_i^*,U_i}\\ \left\langle \left[(L_X)_v(v, \lambda; u)\right]_{|v=u}, h\right\rangle_{U^*,U} &= \sum_{i\in[N]} \left\langle \left[(L_X)_{v^i}(v, \lambda; u)\right]_{|v=u}, h^i\right\rangle_{U_i^*,U_i}\\ &= \sum_{i\in[N]} \left[\langle (\theta_i)_{v^i}(u), h^i\rangle_{U_i^*,U_i}\right]\\ &\quad + \langle G_v(u)^*\lambda, h\rangle_{U^*,U}. \end{aligned}$$

In this setting, quasi-Nash equilibria and variational equilibria can be interpreted as KKT points of the corresponding optimization problem, see Definition 2.63.

For a GNEP with feasible sets $F_i(u^{-i}) = \{v^i \in U_i : (v^i, u^{-i}) \in \mathcal{X}\}$, every normalized equilibrium $u \in \mathcal{X}$ is a Nash equilibrium. In fact, if u is a normalized equilibrium, then for any fixed $i \in [N]$ and for all $v^i \in F_i(u^{-i})$, there holds $v = (v^i, u^{-i}) \in \mathcal{X}$ and

$$\theta_i(v^i, u^{-i}) = \sum_{j\in[N]} \theta_j(v^j, u^{-j}) - \sum_{j\neq i} \theta_j(u^j, u^{-j}) \geq \sum_{j\in[N]} \theta_j(u^j, u^{-j}) - \sum_{j\neq i} \theta_j(u^j, u^{-j}) = \theta_i(u^i, u^{-i}),$$

where the inequality follows directly from the fact that u is a normalized equilibrium.

There are various ways to characterize and compute the introduced types of equilibria (Nash, quasi-Nash, normalized, variational), with the Nikaido–Isoda functional being one of them, see [86]. Indeed, the Nikaido–Isoda functional reads

$$\Psi(u, v) = \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right], \tag{2.19}$$

but regularized versions of the functional have also been studied in the literature, as we will see in the thesis later on. It is evident that $u \in \mathcal{X}$ is a normalized equilibrium if and only if $\Psi(u, v) \leq 0$ for all $v \in \mathcal{X}$. This is equivalent to $u \in \mathcal{X}$ being a solution to $\sup_{v\in\mathcal{X}} \Psi(u, v)$ with optimal value 0, while if $u \in \mathcal{X}$ is not a normalized equilibrium, then the optimal value is strictly positive. Consequently, the so-called *Nikaido–Isoda merit functional* $V : U \to \mathbb{R}$, $V(u) = \sup_{v\in\mathcal{X}} \Psi(u, v)$, is nonnegative on $\mathcal{X}$ and it holds $u \in \mathcal{X}$ and $V(u) = 0$ if and only if u is a normalized equilibrium. A similar construction, where the supremum is taken over $F_i(u^{-i})$ rather than $\mathcal{X}$, yields a Nikaido–Isoda merit functional that represents Nash equilibria. In order to handle the nonconvexity of the objective functional θ_i, we introduce a regularized Nikaido–Isoda functional, see [106], which is given by

$$\Psi_\alpha(u, v) = \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\right]. \tag{2.20}$$

The corresponding *regularized Nikaido–Isoda merit functional* is defined by

$$V_\alpha(u) = \sup_{v\in\mathcal{X}} \Psi_\alpha(u, v). \tag{2.21}$$

Note that solving the problem $\sup_{v\in\mathcal{X}} \Psi_\alpha(u, v)$ is equivalent to solving $\inf_{v\in\mathcal{X}} \widetilde{\Psi}_\alpha(u, v)$, where the associated objective functional is represented by

$$\widetilde{\Psi}_\alpha(u, v) = \sum_{i\in[N]} \left[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\right]. \tag{2.22}$$

One can proceed similarly for Nash equilibria and define the regularized Nikaido–Isoda merit functional by

$$\widetilde{V}_\alpha(u) = \sum_{i\in[N]} \sup_{v^i\in F_i(u^{-i})} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\right]. \tag{2.23}$$

Detailed requirements on all spaces and problem data are given in Chapter 3 in Assumption 3.1 and Assumption 3.3. The parameter $\alpha \geq 0$ is chosen according to Assumption 3.4 in order to guarantee suitable convexity properties. This might result in a large value of α. In order to minimize the requirement for potentially fairly high values of α, which could slow down numerical approaches that employ V_α or $\widetilde{V}_\alpha$, we also investigate local features of the GNEP consisting (2.7).

2.3.3 Concept of the Augmented Lagrangian Method

In the following, we describe the concept of the augmented Lagrangian method for the minimization problem

$$\min_{u \in U} f(u) \qquad \text{s.t.} \qquad G(u) \in K, \quad u \in \mathcal{X}. \tag{2.24}$$

Here, we have introduced the functional $f : U \to \mathbb{R}$, the operator $G : U \to X$, and the subsets $K \subseteq X$, $\mathcal{X} \subseteq U$ of given normed spaces U, X. The exact requirements are discussed in Chapter 3 later on.

The augmented Lagrangian method operates as follows: In order to obtain a differentiable multiplier term, we must utilize a space for which the squared distance functional for convex and closed sets is continuously differentiable. Consequently, we reformulate the problem (2.24) equivalently and assume that there exists an operator $e : X \to Y$ such that $K_Y \subseteq Y$ is a convex and closed set with $e^{-1}(K_Y) = K$. Then the reformulated problem reads

$$\min_{u \in U} f(u) \qquad \text{s.t.} \qquad e(G(u)) \in K_Y, \quad u \in \mathcal{X}. \tag{2.25}$$

The precise specifications for the spaces, sets, and operators are provided in Chapter 3. The canonical choice for Y is some Hilbert space, but we later extend this to a uniformly smooth and uniformly convex Banach space. We denote by P_{K_Y} the metric projection of Y onto K_Y, and by

$$\operatorname{dist}(h, K_Y) = \|h - P_{K_Y}(h)\|_Y$$

the distance between some element $h \in Y$ and the set K_Y, see again Definition 2.50.

The augmented Lagrangian functional for (2.25) reads

$$L_\rho(u, w) = f(u) + \frac{\rho}{2} \operatorname{dist}^2\left(e(G(u)) + \tfrac{J_Y^{-1}(w)}{\rho}, K_Y\right),$$

where $J_Y : Y \to Y^*$ denotes the duality mapping of Y, see again Definition 2.53, and ρ is a positive parameter. In the case of a Hilbert space, the duality mapping corresponds to the Riesz mapping. The general augmented Lagrangian algorithm is given by:

Algorithm 2.64.

0. Choose parameters $\rho_0 > 0$, $\gamma > 1$, $\tau \in (0, 1)$ and some bounded set $B \subseteq Y^*$.

For $k = 0, 1, 2, 3, \ldots$:

1. Choose $w_k \in B$ and compute u_{k+1} as solution to the augmented Lagrangian subproblem:
$$\min_{u \in U} L_{\rho_k}(u, w_k) \qquad \text{s.t.} \qquad u \in \mathcal{X}. \tag{2.26}$$

2. Compute
$$r_{k+1} = \left\| e(G(u_{k+1})) - P_{K_Y}\left(e(G(u_{k+1})) + \tfrac{J_Y^{-1}(w_k)}{\rho_k}\right)\right\|_Y.$$
If $k = 0$ or $r_{k+1} \leq \tau r_k$, set $\rho_{k+1} = \rho_k$. Otherwise, set $\rho_{k+1} = \gamma \rho_k$.

This strategy is challenging to apply to nonconvex problems because it demands global solutions to the general nonconvex problem (2.26). Convexity of (2.26) would necessitate appropriate convexity properties of f and cone-convexity characteristics of G. Consequently, there are versions of the augmented Lagrangian method in which only first-order stationary points of (2.26) are calculated. Then, the first step of Algorithm 2.64 is replaced by:

1^*. Choose $w_k \in B$ and compute u_{k+1} by solving the VI
$$u_{k+1} \in \mathcal{X}, \qquad \left\langle \left[(L_{\rho_k})_v(v, w_k)\right]_{|v=u_{k+1}}, z - u_{k+1}\right\rangle_{U^*,U} \geq 0 \qquad \forall\, z \in \mathcal{X}.$$

Lemma 2.65. Let us define the elements $\lambda_k = e^* \tilde{\lambda}_k$, $k \in \mathbb{N}$, and
$$\tilde{\lambda}_k = \rho_{k-1} J_Y\big(e(G(u_k)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}} - P_{K_Y}\big(e(G(u_k)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}}\big)\big). \tag{2.27}$$
Then there exists a sequence $\zeta_k \downarrow 0$ with
$$\langle \tilde{\lambda}_k, h - e(G(u_k))\rangle_{Y^*,Y} \leq \zeta_k \qquad \forall\, h \in K_Y.$$
Moreover, the following similar result holds for λ_k
$$\langle \lambda_k, y - G(u_k)\rangle_{X^*,X} \leq \zeta_k \qquad \forall\, y \in K.$$

Proof. Let $z_k = e(G(u_k)) + \frac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}}$ and $p_k = P_{K_Y}(z_{k+1}) \in K_Y$. Note that the element $P_{K_Y}(z_{k+1})$ minimizes the functional $h \mapsto \frac{1}{2}\|h - z_{k+1}\|_Y^2$ on the set K_Y. Since the derivative of this functional is equal to $J_Y(h - z_{k+1})$, the first-order optimality condition for this minimizer reads
$$\big\langle J_Y(P_{K_Y}(z_{k+1}) - z_{k+1}), h - P_{K_Y}(z_{k+1})\big\rangle_{Y^*,Y} \geq 0 \qquad \forall\, h \in K_Y.$$
This yields
$$\langle \tilde{\lambda}_{k+1}, h - p_{k+1}\rangle_{Y^*,Y} = \rho_k \big\langle J_Y(z_{k+1} - P_{K_Y}(z_{k+1})), h - P_{K_Y}(z_{k+1})\big\rangle_{Y^*,Y} \leq 0 \quad \forall\, h \in K_Y, \tag{2.28}$$
which shows $\tilde{\lambda}_{k+1} \in N_{K_Y}(p_{k+1})$. Furthermore, solving
$$\tilde{\lambda}_{k+1} = \rho_k J_Y(z_{k+1} - p_{k+1}) = \rho_k J_Y\big(e(G(u_{k+1})) + \tfrac{J_Y^{-1}(w_k)}{\rho_k} - p_{k+1}\big)$$

for $e(G(u_{k+1}))$ we get

$$e(G(u_{k+1})) = \frac{1}{\rho_k}(J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)) + p_{k+1}. \tag{2.29}$$

Now, inserting (2.29) and using the estimate (2.28) in the second step it yields for all $h \in K_Y$,

$$\begin{aligned}\langle \tilde{\lambda}_{k+1}, h - e(G(u_{k+1}))\rangle_{Y^*,Y} &= \langle \tilde{\lambda}_{k+1}, h - p_{k+1}\rangle_{Y^*,Y} - \frac{1}{\rho_k}\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\rangle_{Y^*,Y} \\ &\leq -\frac{1}{\rho_k}\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\rangle_{Y^*,Y}.\end{aligned}$$

Moreover, it holds

$$r_{k+1} = \|e(G(u_{k+1})) - p_{k+1}\|_Y = \frac{1}{\rho_k}\|J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\|_Y.$$

If $\{\rho_k\}_{k\in\mathbb{N}}$ stays bounded, then $r_k \to 0$ as $k \to \infty$ and thus, it holds

$$\|J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\|_Y = \rho_k r_{k+1} \to 0,$$

as $k \to \infty$. We have $\|w_k\|_{Y^*} = \|J_Y(J_Y^{-1}(w_k))\|_{Y^*} = \|J_Y^{-1}(w_k)\|_Y$. Since $\{w_k\}_{k\in\mathbb{N}}$ is bounded in Y^*, we get that $\{J_Y^{-1}(w_k)\}_{k\in\mathbb{N}}$ is bounded in Y and hence, $\{J_Y^{-1}(\tilde{\lambda}_{k+1})\}_{k\in\mathbb{N}}$ is bounded in Y. By the boundedness of J_Y on bounded sets, $\{\tilde{\lambda}_{k+1}\}_{k\in\mathbb{N}}$ is bounded in Y^*. Thus, we have

$$|\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\rangle_{Y^*,Y}| \leq \|\tilde{\lambda}_{k+1}\|_{Y^*}\|J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\|_Y \to 0,$$

as $k \to \infty$. It holds that $\{\rho_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^N_{>0}$ is increasing but bounded and hence $\rho_k \nrightarrow 0$ as $k \to \infty$ and $\{\frac{1}{\rho_k}\}_{k\in\mathbb{N}}$ is bounded. Consequently, in this case, we can choose

$$\zeta_{k+1} = \frac{1}{\rho_k}[\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(w_k) - J_Y^{-1}(\tilde{\lambda}_{k+1})\rangle_{Y^*,Y}]_+.$$

To tackle the case $\rho_k \to \infty$ as $k \to \infty$, we use the representation of the inverse of the duality mapping $J_Y^{-1} = \widetilde{J}^{-1}J_{Y^*}$, see Proposition 2.56. Inserting this fact implies

$$\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(\tilde{\lambda}_{k+1})\rangle_{Y^*,Y} = \langle \tilde{\lambda}_{k+1}, \widetilde{J}^{-1}J_{Y^*}(\tilde{\lambda}_{k+1})\rangle_{Y^*,Y},$$

where $\widetilde{J} : Y \to Y^{**}$ and $J_{Y^*} : Y^* \to Y^{**}$ is the duality mapping on Y^* analogously defined to J_Y. Furthermore, we use that the isometry $\widetilde{J}$ is bijective and obtain

$$\langle \tilde{\lambda}_{k+1}, \widetilde{J}^{-1}J_{Y^*}(\tilde{\lambda}_{k+1})\rangle_{Y^*,Y} = \langle \widetilde{J}\widetilde{J}^{-1}J_{Y^*}(\tilde{\lambda}_{k+1}), \tilde{\lambda}_{k+1}\rangle_{Y^{**},Y^*} = \langle J_{Y^*}(\tilde{\lambda}_{k+1}), \tilde{\lambda}_{k+1}\rangle_{Y^{**},Y^*} = \|\tilde{\lambda}_{k+1}\|^2_{Y^*}.$$

Observe that minimizing

$$\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\rangle_{Y^*,Y} = \|\tilde{\lambda}_{k+1}\|^2_{Y^*} - \langle \tilde{\lambda}_{k+1}, J_Y^{-1}(w_k)\rangle_{Y^*,Y}$$

with respect to $\tilde{\lambda}_{k+1}$ yields $2J_{Y^*}(\tilde{\lambda}_{k+1}) - J_{Y^*}(w_k) = 0$ and thus $\tilde{\lambda}_{k+1} = \frac{w_k}{2}$, since J_Y^{-1} and $\widetilde{J}$ are bijective this property holds also for J_{Y^*}. The corresponding minimum value is $-\frac{1}{4}\|w_k\|_{Y^*}^2$. Hence, we arrive at

$$-\frac{1}{\rho_k}\langle \tilde{\lambda}_{k+1}, J_Y^{-1}(\tilde{\lambda}_{k+1}) - J_Y^{-1}(w_k)\rangle_{Y^*,Y} \leq \frac{1}{4\rho_k}\|w_k\|_{Y^*}^2.$$

For $\rho_k \to \infty$ as $k \to \infty$, the right-hand side tends to 0 as $k \to \infty$. In this case, we thus can choose $\zeta_{k+1} = \frac{1}{4\rho_k}\|w_k\|_{Y^*}^2$.

Finally, for any $y \in K$, it holds $e(y) \in K_Y$ and thus

$$\langle \lambda_k, y - G(u_k)\rangle_{X^*,X} = \langle \tilde{\lambda}_k, e(y) - e(G(u_k))\rangle_{Y^*,Y} \leq \zeta_k. \qquad \square$$

2.4 Generalized Danskin Theorem

Danskin first proved the differentiability of a bilevel optimization problem in 1966, see [36,37]. In fact, he demonstrated that the functional $g : \mathbb{R}^n \to \mathbb{R}$ with

$$g(x) = \max_{z \in Z} f(x, z)$$

admits the directional derivative

$$\nabla g(x) \cdot h = \max_{\substack{z \in Z, \\ f(x,z) = g(x)}} \nabla_x f(x, z) \cdot h$$

in direction $h \in \mathbb{R}^n$, provided that $f : \mathbb{R}^n \times Z \to \mathbb{R}$ is continuous and C^1 in the first component for a given compact topological space Z. Since then, many more general results have been proved, and we refer to [11] for a discussion of several extensions. We highlight the books [18,78], as well as the article [34], which allows for a more general setting of a Hausdorff locally convex topological vector space X as the domain of the functional $g : X \to \mathbb{R}$. Furthermore, the notion of continuity in the second component of f is replaced by upper semicontinuity.

Danskin's theorem is extended in this chapter by introducing two different topologies on the normed space X and assuming the continuity of f in the first component with respect to some coarser topology on X. In contrast to the assumption of a Hausdorff locally convex topological vector space, we assume that X is only a Hausdorff space with respect to this coarser topology. Furthermore, we investigate the directional differentiability of g as well as the Gâteaux and Fréchet derivatives. This was also done in [18], but for a Banach space Z and with the assumption of inf-compactness.

From here on, let (A1)–(A5) of the following assumption be valid.

Assumption 2.66.

(A1) $(X, \|\cdot\|_X)$ is a normed space, (X, τ_X) is a Hausdorff topological space with a (coarser) topology τ_X such that the identity operator on X is $\|\cdot\|_X$-to-τ_X continuous.

(A2) (Z, τ_Z) is a Hausdorff topological space.

(A3) $W \subseteq X$ is an open set.

(A4) $Y \subseteq Z$ is a sequentially compact set with respect to τ_Z.

(A5) $f : W \times Y \to \mathbb{R}$ is sequentially upper semicontinuous with respect to $\tau_X \times \tau_Z$ and $f(\cdot, z) : W \to \mathbb{R}$ is sequentially continuous with respect to τ_X for all $z \in Y$.

(A6) $f : W \times Y \to \mathbb{R}$ is Fréchet differentiable in the first variable and its derivative $f_x : W \times Y \to X^*$ is sequentially continuous from $\tau_X \times \tau_Z$ to $\|\cdot\|_{X^*}$.

Unless otherwise stated, all properties, e.g., compactness, closedness, or convergence, are considered with respect to the norm topology.

We define the functional $g : W \to \mathbb{R}$ by $g(x) = \max_{z \in Y} f(x, z)$ and the set-valued operator $M : W \rightrightarrows Z$ such that $M(x)$ denotes the set of all elements $z \in Y$ for a given points $x \in W$ such that the maximum of g is attained, i.e.,

$$M(x) = \{z \in Y : g(x) = f(x, z)\}. \tag{2.30}$$

The $\|\cdot\|_X$-topology may be the canonical option for τ_X. However, by introducing a topology τ_X that is coarser than the norm topology, it enables us to establish continuity results for g, M, and the derivative of g.

Next, we prove that the set-valued mapping M is sequentially closed as defined in Definition 2.7. But first off, we show that the function g is mathematically well-defined. Moreover, we derive other useful properties of g and M such as local Lipschitz continuity and differentiability.

Lemma 2.67. The functional g is well-defined and $M(x)$ is nonempty for all $x \in W$.

Proof. Let $x \in W$ be arbitrarily fixed. We demonstrate that the supremum $\sup_{z \in Y} f(x, z)$ is indeed attained by some element in Y. Let $\{z_k\}_{k \in \mathbb{N}} \subseteq Y$ be a maximizing sequence, i.e., $f(x, z_k) \to \sup_{z \in Y} f(x, z) \in \mathbb{R} \cup \{\infty\}$ as $k \to \infty$. By assumption (A4), Y is sequentially compact with respect to τ_Z and thus, there exists a subsequence $\{z_{k_l}\}_{l \in \mathbb{N}} \subseteq \{z_k\}_{k \in \mathbb{N}} \subseteq Y$ and an element $\bar{z} \in Y$ such that it holds $z_{k_l} \to \bar{z}$ with respect to τ_Z as $l \to \infty$. Moreover, it yields the estimate

$$\sup_{z \in Y} f(x, z) = \lim_{k \to \infty} f(x, z_k) = \lim_{l \to \infty} f(x, z_{k_l}) = \limsup_{l \to \infty} f(x, z_{k_l}) \leq f(x, \bar{z}),$$

due to the sequential upper semicontinuity of $f(x, \cdot)$ with respect to τ_Z. Consequently, we obtain

$$\sup_{z \in Y} f(x, z) \leq f(x, \bar{z}),$$

from which we conclude that the supremum is attained at $\bar{z} \in Y$, i.e., the maximum exists on Y. Finally, g is well-defined with $g(x) = f(x, \bar{z})$ and $\bar{z} \in M(x) \neq \emptyset$. $\square$

Next, we prove that the set $M(x)$, see again (2.30) for its definition, is sequentially compact in Y and Z with respect to the topology τ_Z.

Lemma 2.68. *$M(x) \subseteq Y \subseteq Z$ is sequentially compact with respect to τ_Z for $x \in W$.*

Proof. Let $x \in W$ be arbitrarily fixed and let $\{z_k\}_{k\in\mathbb{N}}$ be a given sequence in $M(x)$. Since $Y \subseteq Z$ is sequential compact with respect to τ_Z, there exists a subsequence $\{z_{k_l}\}_{l\in\mathbb{N}} \subseteq \{z_k\}_{k\in\mathbb{N}} \subseteq Y$ and an element $\bar{z} \in Y$ such that it holds $z_{k_l} \to \bar{z}$ with respect to τ_Z as $l \to \infty$. We apply the sequential upper semicontinuity of $f(x, \cdot)$ with respect to τ_Z and $z_{k_l} \in M(x)$, which implies $f(x, z_{k_l}) = g(x)$ for all $l \in \mathbb{N}$ and consequently, we obtain

$$g(x) \geq f(x, \bar{z}) \geq \limsup_{l\to\infty} f(x, z_{k_l}) = \limsup_{l\to\infty} g(x) = g(x).$$

This proves $f(x, \bar{z}) = g(x)$ and $\bar{z} \in M(x)$ from which we conclude that $M(x) \subseteq Z$ is sequentially compact with respect to τ_Z for $x \in W$. □

With the compactness of $M(x)$ at hand, we are ready to prove that the set-valued mapping M is indeed sequentially closed.

Lemma 2.69. *The mapping $M : X \rightrightarrows Z$ has a sequentially closed graph with respect to $\tau_X \times \tau_Z$.*

Proof. Let $\{x_k\}_{k\in\mathbb{N}} \subseteq W$ be a converging sequence in W with respect to the topology τ_X, i.e., there exists an element $\bar{x} \in W$ with $x_k \to \bar{x}$ in τ_X as $k \to \infty$. Moreover, we consider another converging sequence $z_k \in M(x_k)$, $k \in \mathbb{N}$, with $z_k \to \bar{z}$ with respect to τ_Z as $k \to \infty$ for some element $\bar{z} \in Z$. We have to verify $\bar{z} \in M(x)$ in order to prove the claim of the lemma. First off, we note that for all $z \in Y$ it holds

$$f(\bar{x}, \bar{z}) \geq \limsup_{k\to\infty} f(x_k, z_k) \geq \limsup_{k\to\infty} f(x_k, z) = f(\bar{x}, z),$$

where we applied the sequential upper semicontinuity of f with respect to $\tau_X \times \tau_Z$ in the first step and the sequential continuity of $f(\cdot, z)$ with respect to τ_X in the last equality. Thus, we arrive at $f(\bar{x}, \bar{z}) \geq f(\bar{x}, z)$ for all $z \in Y$, which shows $\bar{z} \in M(x)$. □

In the next step, we prove several continuity and differentiability aspects of the function g. First, we begin with the sequential continuity of g in the next lemma.

Lemma 2.70. *The functional $g : W \to \mathbb{R}$ is sequentially continuous with respect to τ_X and continuous in the norm topology.*

Proof. We proof this statement by contradiction and assume that g is not sequentially continuous on W with respect to τ_X. Consequently, there is a sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq W$ and $\bar{x} \in W$ such that $x_k \to \bar{x}$ with respect to τ_X as $k \to \infty$ and some $\varepsilon > 0$ such that $|g(x_k) - g(\bar{x})| \geq \varepsilon$ for all $k \in \mathbb{N}$. We assume a sequence $z_k \in M(x_k) \subseteq Y$, $k \in \mathbb{N}$, and since $Y \subseteq Z$ is sequentially compact with respect to τ_Z, there exists a subsequence $\{z_{k_l}\}_{l\in\mathbb{N}} \subseteq \{z_k\}_{k\in\mathbb{N}}$ and an element $\bar{z} \in Z$ such that it holds $z_{k_l} \to \bar{z}$ with respect to τ_Z as $l \to \infty$. We make use of the sequential closedness of the mapping $M : W \rightrightarrows Z$ and the fact that $z_{k_l} \in M(x_{k_l})$ to conclude $\bar{z} \in M(x)$.

Next, we distinguish the sign of the difference $g(x_{k_l}) - g(\bar{x})$ by considering two cases. First, if it holds $g(x_{k_l}) - g(\bar{x}) \geq \varepsilon$, then we obtain by sequential upper semicontinuity of f with respect to $\tau_X \times \tau_Z$ the following estimate

$$g(\bar{x}) + \varepsilon \leq \limsup_{l\to\infty} g(x_{k_l}) = \limsup_{l\to\infty} f(x_{k_l}, z_{k_l}) \leq f(\bar{x}, \bar{z}) = g(\bar{x}).$$

Consequently, we arrive at the contradiction $g(\bar{x}) + \varepsilon \leq g(\bar{x})$ for $\varepsilon > 0$. In the other case, it has to hold $g(x_{k_l}) - g(\bar{x}) \leq -\varepsilon$ and we obtain

$$g(\bar{x}) - \varepsilon \geq \limsup_{l\to\infty} g(x_{k_l}) = \limsup_{l\to\infty} f(x_{k_l}, z_{k_l}).$$

Since it holds $z_{k_l} \in M(x_{k_l})$, we can conclude the estimate

$$\limsup_{l\to\infty} f(x_{k_l}, z_{k_l}) \geq \limsup_{l\to\infty} f(x_{k_l}, \bar{z}),$$

and by the sequential continuity of $f(\cdot, \bar{z})$ with respect to τ_X it yields

$$\limsup_{l\to\infty} f(x_{k_l}, \bar{z}) = \lim_{l\to\infty} f(x_{k_l}, \bar{z}) = f(\bar{x}, \bar{z}) = g(\bar{x}).$$

Finally, we arrive at

$$g(\bar{x}) - \varepsilon \geq g(\bar{x}),$$

which is a contradiction to $\varepsilon > 0$. Altogether, we have proved that g is sequentially continuous with respect to τ_X. □

From here on, let (A6) of Assumption 2.66 be additionally valid. This assumption is key in proving that g is locally Lipschitz continuous, as the next lemma states.

Lemma 2.71. The functional $g : W \to \mathbb{R}$ is locally Lipschitz continuous.

Proof. We consider the functional $\|f_x(\cdot,\cdot)\|_{X^*} : W \times Y \to \mathbb{R}$, which fulfills (A5) of Assumption 2.66. Indeed, $\|f_x(\cdot,\cdot)\|_{X^*} : W \times Y \to \mathbb{R}$ is sequentially upper semicontinuous with respect to $\tau_X \times \tau_Z$ and $\|f_x(\cdot, z)\|_{X^*} : W \to \mathbb{R}$ is continuous since $f_x : W \times Y \to X^*$ is sequentially continuous with respect to $\tau_X \times \tau_Z$. Consequently, we can apply Lemma 2.70 and obtain the continuity of the mapping $x \mapsto \max_{z\in Y} \|f_x(x, z)\|_{X^*}$ with respect to τ_X. By the $\|\cdot\|_X$-to-τ_X continuity of the identity operator on X, this also implies the continuity of $x \mapsto \max_{z\in Y} \|f_x(x, z)\|_{X^*}$ with respect to the norm topology.

By definition of the derived continuity, we know that for any $x \in W$ and for each $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$\Big| \max_{z\in Y} \|f_x(x, z)\|_{X^*} - \max_{z\in Y} \|f_x(y, z)\|_{X^*} \Big| \leq \varepsilon$$

for all $y \in W$ with $\|x - y\|_X < \delta$. We consider some fixed element $x_0 \in W$ and by the last inequality we obtain

$$\max_{z\in Y} \|f_x(y, z)\|_{X^*} \leq \varepsilon + \max_{z\in Y} \|f_x(x_0, z)\|_{X^*}$$

for all $y \in B_\delta(x_0) \subseteq W$. Thus, there exists some constant $L > 0$ with

$$\max_{z \in Y} \|f_x(y, z)\|_{X^*} \leq L \qquad \forall\, y \in B_\delta(x_0) \subseteq W. \tag{2.31}$$

Next, we consider the two elements $x_1 \in B_\delta(x_0)$, $x_2 \in B_\delta(x_0)$ and another two elements $z_1 \in M(x_1)$, $z_2 \in M(x_2)$. Without loss of generality, we assume $g(x_1) \geq g(x_2)$. Since it holds $z_2 \in M(x_2)$, it yields

$$|g(x_1) - g(x_2)| = g(x_1) - g(x_2) = f(x_1, z_1) - f(x_2, z_2) \leq f(x_1, z_1) - f(x_2, z_1),$$

and by the mean value theorem, see Theorem 2.21, there exists some parameter $t \in [0, 1]$ with

$$f(x_1, z_1) - f(x_2, z_1) = \big\langle f_x((1-t)x_1 + tx_2, z_1), x_1 - x_2 \big\rangle_{X^*,X}.$$

Since the ball $B_\delta(x_0)$ is convex, we are allowed to consider the convex combination $(1-t)x_1 + tx_2 \in B_\delta(x_0)$. Applying the bound (2.31) from before, we have shown that it holds $\|f_x(x_1 + t(x_2 - x_1), z_1)\|_{X^*} \leq L$ and we obtain the local Lipschitz continuity of g as follows

$$\begin{aligned} |g(x_1) - g(x_2)| &\leq \big\langle f_x((1-t)x_1 + tx_2, z_1), x_1 - x_2 \big\rangle_{X^*,X} \\ &\leq \|f_x(x_1 + t(x_2 - x_1), z_1)\|_{X^*} \|x_1 - x_2\|_X \\ &\leq L\|x_1 - x_2\|_X. \end{aligned}$$ □

We will heavily exploit the next result later on in the thesis since it states the directional differentiability of g and provides a formula for its derivative.

Lemma 2.72. The functional $g : W \to \mathbb{R}$ is directionally differentiable for all $h \in X$ with

$$\mathrm{d}g(x, h) = \max_{z \in M(x)} \big\langle f_x(x, z), h \big\rangle_{X^*,X}.$$

Proof. We introduce a parameter $T > 0$ that fulfills $x + [0, T]h \subseteq W$. Furthermore, let $\{t_k\}_{k \in \mathbb{N}} \subseteq (0, T]$ be a sequence with $t_k \downarrow 0$. We demonstrate that the limit of $\frac{g(x+th)-g(x)}{t}$ as $t \downarrow 0$ exists and is equal to $\max_{z \in M(x)} \langle f_x(x, z), h \rangle_{X^*,X}$. We prove this statement by considering the limit superior $\limsup_{k\to\infty} \frac{g(x+t_kh)-g(x)}{t_k}$ and the limit inferior $\liminf_{k\to\infty} \frac{g(x+t_kh)-g(x)}{t_k}$ one after another.

Let us denote $x_k = x + t_k h$ and let $z_k \in M(x_k)$, $k \in \mathbb{N}$. Furthermore, let $z \in M(x)$ be arbitrary. Using the assumption $z_k \in M(x_k)$, it yields

$$g(x + t_k h) - g(x) = f(x_k, z_k) - f(x, z) \geq f(x_k, z) - f(x, z).$$

By the mean value theorem, see Theorem 2.21, there exists some $\lambda \in [0, 1]$ such that

$$f(x_k, z) - f(x, z) = \big\langle f_x((1-\lambda)x_k + \lambda x, z), x_k - x \big\rangle_{X^*,X} = \big\langle f_x(x + (1-\lambda)t_k h, z), t_k h \big\rangle_{X^*,X}.$$

Thus, combining these two estimates we obtain the inequality

$$
\begin{aligned}
\liminf_{k\to\infty} \frac{g(x_k)-g(x)}{t_k} &\geq \liminf_{k\to\infty} \big\langle f_x(x+(1-\lambda)t_k h, z), h\big\rangle_{X^*,X} \\
&= \lim_{k\to\infty} \big\langle f_x(x+(1-\lambda)t_k h, z), h\big\rangle_{X^*,X} \\
&= \big\langle f_x(x,z), h\big\rangle_{X^*,X},
\end{aligned}
$$

where we applied the (sequential) continuity of $f_x(\cdot, z)$ with respect to τ_X and by (A1) of Assumption 2.66 with respect to the norm topology. Since this inequality holds for all $z \in M(x)$, it yields

$$
\liminf_{k\to\infty} \frac{g(x_k)-g(x)}{t_k} \geq \max_{z\in M(x)} \big\langle f_x(x,z), h\big\rangle_{X^*,X}. \tag{2.32}
$$

For the other direction, we bound $\limsup_{k\to\infty} \frac{g(x_k)-g(x)}{t_k}$ from above. We prove this by contradiction. Suppose that there exists some $\varepsilon > 0$ such that it holds

$$
\limsup_{k\to\infty} \frac{g(x_k)-g(x)}{t_k} \geq \max_{z\in M(x)} \big\langle f_x(x,z), h\big\rangle_{X^*,X} + \varepsilon. \tag{2.33}
$$

By the sequential compactness of Y with respect to τ_Z, there exists a subsequence $\{z_{k_l}\}_{l\in\mathbb{N}} \subseteq Y$ and some element $\bar{z} \in Z$ such that $z_{k_l} \to \bar{z}$ with respect to τ_Z as $l \to \infty$. In particular, we obtain $z_{k_l} \in M(x_{k_l})$, $l \in \mathbb{N}$. Furthermore, it yields $\bar{z} \in M(x)$ due to the sequential closedness of $M : W \rightrightarrows Z$. Thus, we arrive at the bound

$$
\limsup_{l\to\infty} \frac{g(x_{k_l})-g(x)}{t_{k_l}} \geq \max_{z\in M(x)} \big\langle f_x(x,z), h\big\rangle_{X^*,X} + \varepsilon. \tag{2.34}
$$

Let $z \in M(x)$ be arbitrary in the following. Similarly to before, we compute

$$
g(x_{k_l}) - g(x) = f(x_{k_l}, z_{k_l}) - f(x,z) \leq f(x_{k_l}, z_{k_l}) - f(x, z_{k_l}),
$$

and by the mean value theorem, see Theorem 2.21, there exists some parameter $\lambda \in [0,1]$ such that

$$
\begin{aligned}
f(x_{k_l}, z_{k_l}) - f(x, z_{k_l}) &= \big\langle f_x((1-\lambda)x_{k_l} + \lambda x, z_{k_l}), x_{k_l} - x\big\rangle_{X^*,X} \\
&= \big\langle f_x(x+(1-\lambda)t_{k_l} h, z_{k_l}), t_{k_l} h\big\rangle_{X^*,X}.
\end{aligned}
$$

Again, we make use of the sequential continuity of f_x with respect to $\|\cdot\|_X \times \tau_Z$ and obtain the estimate

$$
\begin{aligned}
\limsup_{l\to\infty} \frac{g(x_{k_l})-g(x)}{t_{k_l}} &\leq \limsup_{l\to\infty} \big\langle f_x(x+(1-\lambda)t_{k_l} h, z_{k_l}), h\big\rangle_{X^*,X} \\
&= \lim_{l\to\infty} \big\langle f_x(x+(1-\lambda)t_{k_l} h, z_{k_l}), h\big\rangle_{X^*,X} \\
&= \big\langle f_x(x,\bar{z}), h\big\rangle_{X^*,X}.
\end{aligned}
$$

Together with (2.34), we obtain the contradiction

$$\max_{z\in M(x)} \langle f_x(x,z),h\rangle_{X^*,X} + \varepsilon \le \limsup_{l\to\infty} \frac{g(x_{k_l})-g(x)}{t_{k_l}} \le \langle f_x(x,\bar z),h\rangle_{X^*,X}.$$

Consequently, the assumption (2.33) was false and it holds an upper bound for the limit superior.

Together with the bound from before on the limit inferior, see (2.32), we obtain

$$\limsup_{k\to\infty} \frac{g(x_k)-g(x)}{t_k} \le \max_{z\in M(x)} \langle f_x(x,z),h\rangle_{X^*,X} \le \liminf_{k\to\infty} \frac{g(x_k)-g(x)}{t_k}.$$

Since $t_k \downarrow 0$ was arbitrary, we can conclude the desired result

$$\mathrm{d}g(x,h) = \lim_{t\downarrow 0} \frac{g(x_k)-g(x)}{t} = \max_{z\in M(x)} \langle f_x(x,z),h\rangle_{X^*,X}. \qquad \square$$

Later on, we are often in a scenario where the correspondence map $M(x)$ is a singleton or, more precisely, it is single-valued near x. Next, we make such an additional assumption and prove additional properties of g.

Lemma 2.73. If the set $M(x) = \{m(x)\}$ is a singleton at the point $x \in W$, then it holds

$$\mathrm{d}g(x,h) = \langle f_x(x,m(x)),h\rangle_{X^*,X}. \tag{2.35}$$

Additionally, it follows that $g : W \to \mathbb{R}$ is Gâteaux differentiable at x.

Proof. By Lemma 2.72, the functional $g : W \to \mathbb{R}$ is directionally differentiable at x where its derivative is given by (2.35). We prove that $g'(x) : X \to \mathbb{R}$, which is defined by $h \mapsto \mathrm{d}g(x,h)$, is linear and bounded. The linearity of $\mathrm{d}g(x,h) = \langle f_x(x,m(x)),h\rangle_{X^*,X}$ with respect to h is obvious and the boundedness follows due to the inequalities

$$\mathrm{d}g(x,h) = \langle f_x(x,m(x)),h\rangle_{X^*,X} \le \|f_x(x,m(x))\|_{X^*}\|h\|_X \le C\|h\|_X. \qquad \square$$

Lemma 2.74. If the correspondence $M : W \rightrightarrows Z$ is single-valued at $x \in W$, i.e., it holds $M(x) = \{m(x)\}$ with $m(x) \in Y$, then M is sequentially τ_X to τ_Z continuous at x.

Proof. Let $\bar x \in W$ be arbitrary but fixed. Suppose that M is not sequentially τ_X to τ_Z continuous. Then there exists some converging sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq W$ with $x_k \to \bar x$ with respect to τ_X as $k \to \infty$ such that $z_k \in M(x_k)$, $k \in \mathbb{N}$, is not converging to the element $z \in M(\bar x) = \{m(\bar x)\}$ with respect to τ_Z as $k \to \infty$. Since z_k does not converge to $m(\bar x)$ with respect to τ_Z, there is some open neighborhood V_Z of $m(\bar x)$ with respect to τ_Z such that it holds $z_k \notin V_Z$ for all $k \in \mathbb{N}$.

Moreover, by assumption, we know that $Y \subseteq Z$ is sequentially compact with respect to τ_Z and therefore, we are allowed to choose a subsequence $\{z_{k_l}\}_{l\in\mathbb{N}} \subseteq \{z_k\}_{k\in\mathbb{N}}$ with $z_{k_l} \to \overline{m}$ with respect to τ_Z as $l \to \infty$. We obtain $z_{k_l} \notin V_Z$ for all $l \in \mathbb{N}$ because it holds $\{z_{k_l}\}_{l\in\mathbb{N}} \subseteq \{z_k\}_{k\in\mathbb{N}}$. Furthermore, we can conclude $\overline{m} \in M(\bar x)$ due to the sequential closedness of M and thus, it follows $\overline{m} = m(\bar x)$. Since V_Z is an open neighborhood of $\overline{m} = m(\bar x)$ with respect to τ_Z and $z_{k_l} \to \overline{m}$ with respect to τ_Z as $l \to \infty$, we are able to find an index L with $z_{k_l} \in V_Z$ for all $l \ge L$. However, this is a contradiction to the supposition in the beginning of the proof. Consequently, M is indeed sequentially τ_X to τ_Z continuous at $\bar x$. $\square$

Our modification of the famous theorem of Danskin stating the continuous differentiability of $x \mapsto g(x) = \max_{z \in Y} f(x, z)$ follows.

Theorem 2.75 (Danskin)**.** If the correspondence $M : W \rightrightarrows Z$ is single-valued in an open set $U_X \subseteq W$, then $g : W \to \mathbb{R}$ is Fréchet differentiable on U_X and its derivative $g' : U_X \to X^*$ is continuous with respect to τ_X and with respect to the norm topology.

Proof. Let $M(\cdot) = \{m(\cdot)\}$ be defined on the set $U_X \subseteq W$. By Lemma 2.72 and Lemma 2.73, we conclude that $g : W \to \mathbb{R}$ is Gâteaux differentiable on U_X with the derivative

$$\langle g'(x), h\rangle_{X^*,X} = \mathrm{d}g(x, h) = \big\langle f_x(x, m(x)), h\big\rangle_{X^*,X} \qquad \forall\, h \in X,\ x \in U_X.$$

Furthermore, $M : W \rightrightarrows Z$ is sequentially τ_X to τ_Z continuous at any $x \in U_X$ due to Lemma 2.74 and thus, the same holds for $m : W \to Z$ on the set U_X. Hence, $f_x(\cdot, m(\cdot)) : W \to X^*$ is continuous on U_X with respect to τ_X and with respect to the norm topology. Finally, we conclude that g is continuously Gâteaux differentiable on U_X, which is equivalent to the continuous Fréchet differentiability on U_X. □

We can generalize Danskin's theorem to a functional $f : W \times Z \to \mathbb{R}$ that can be split into two functionals $f^1 : W \times Z \to \mathbb{R}$ and $f^2 : Z \to \mathbb{R}$ in the following way

$$f(x, z) = f^1(x, z) + f^2(z).$$

Let f^1 be sequentially continuous with respect to $\tau_X \times \tau_Z$ and let f^2 be sequentially upper semicontinuous with respect to τ_Z. Then $f(\cdot, z)$ is sequentially continuous with respect to τ_X and f is sequentially upper semicontinuous with respect to $\tau_X \times \tau_Z$.

Indeed, let $\{x_k\}_{k\in\mathbb{N}} \subseteq X$ be a converging sequence with $x_k \to \bar{x}$ with respect to τ_X as $k \to \infty$ and let $\{z_k\}_{k\in\mathbb{N}} \subseteq Z$ be a sequence such that it holds $z_k \to \bar{z}$ with respect to τ_Z as $k \to \infty$. It holds

$$\limsup_{k\to\infty} f(x_k, z_k) \leq \limsup_{k\to\infty} f^1(x_k, z_k) + \limsup_{k\to\infty} f^2(z_k) \leq f^1(\bar{x}, \bar{z}) + f^2(\bar{z}) = f(\bar{x}, \bar{z}),$$

where we applied the sequential continuity of f^1 with respect to $\tau_X \times \tau_Z$ and the sequential upper semicontinuity of f^2 with respect to τ_Z.

Furthermore, if f^1 is differentiable in x and $(f^1)_x : W \times Z \to X^*$ is sequentially continuous with respect to $\tau_X \times \tau_Z$, then f is differentiable in x with $f_x = (f^1)_x$ where $f_x : W \times Z \to X^*$ is sequentially continuous from $\tau_X \times \tau_Z$ to $\|\cdot\|_{X^*}$. In this situation, (A5) of Assumption 2.66 is satisfied and the above results can be applied to this choice of f.

In order to differentiate localized versions of the Nikaido–Isoda merit functionals later on, it would be helpful if we have some theory for constrained problems with the feasible set depending on x. This type of problem reads

$$\max_{z\in\Phi(x)} f(x, z)$$

for some set-valued correspondence map $\Phi : X \rightrightarrows Z$. We note that some steps of the proof of Danskin's theorem are no longer doable for such general correspondences.

Chapter 3

Mathematical Setting and Assumptions

...the great watershed in optimization isn't between linearity and nonlinearity, but between convexity and nonconvexity. (Ralph Tyrrell Rockafellar)

In the following, we set up the mathematical framework for GNEPs. In the literature, GNEPs have been studied mostly in finite-dimensional spaces, see the review paper [43] and the book [67]. In the case of infinite-dimensional spaces with finer topology than that considered in this thesis, the dissertation [101] is consulted.

Our conclusions, depending on their nature, necessitate different kinds of assumptions for the underlying spaces, continuity and differentiability of the objective functionals, and different sized parameters. It is worth noting that we provide various assumptions, but we only use one at a time while developing particular outcomes. Detailed requirements on all spaces and problem data will be given in this chapter in Assumption 3.1, Assumption 3.3, and Assumption 3.4 in their respective sections.

3.1 Topology Assumptions

We pose the following assumptions on the underlying spaces U, $\widetilde{U}$, H, X, and Y. Here, U denotes the space in which the optimization problem is formulated. As we will see, the auxiliary space $\widetilde{U}$ is strongly connected to U. The space H has a regularizing character and is at least a normed space. In particular, we regularize a nonconvex functional with the squared norm of H. The spaces X and Y are required for augmented Lagrangian method and they model the additional constraints.

Assumption 3.1.

(A1) $(U_i, \|\cdot\|_{U_i})$, $(H_i, \|\cdot\|_{H_i})$ are normed spaces.

(A2) $(U_i, \|\cdot\|_{U_i})$ is a Banach space, $(H_i, \|\cdot\|_{H_i})$ is a uniformly smooth and uniformly convex Banach space with duality map $J_{H_i} : H_i \to H_i^*$.

(A3) $(U_i, \|\cdot\|_{U_i})$, $(H_i, \|\cdot\|_{H_i})$ are uniformly smooth and uniformly convex Banach spaces with duality maps $J_{U_i} : U_i \to U_i^*$ and $J_{H_i} : H_i \to H_i^*$.

(A4) $(U_i, \|\cdot\|_{U_i})$, $(H_i, \|\cdot\|_{H_i})$, and $(\widetilde{U}_i, \|\cdot\|_{\widetilde{U}_i})$ are normed spaces.

(A5) $(U_i, \|\cdot\|_{U_i})$ is a reflexive Banach space, $(H_i, \|\cdot\|_{H_i})$ and $(\widetilde{U}_i, \|\cdot\|_{\widetilde{U}_i})$ are normed spaces.

(A6) $(U_i, \|\cdot\|_{U_i})$ is a uniformly smooth and uniformly convex Banach space with duality map $J_{U_i} : U_i \to U_i^*$, $(H_i, \|\cdot\|_{H_i})$ and $(\widetilde{U}_i, \|\cdot\|_{\widetilde{U}_i})$ are normed spaces.

(A7) $(U_i, \|\cdot\|_{U_i})$, $(H_i, \|\cdot\|_{H_i})$ are uniformly smooth and uniformly convex Banach spaces with duality maps $J_{U_i} : U_i \to U_i^*$ and $J_{H_i} : H_i \to H_i^*$, $(\widetilde{U}_i, \|\cdot\|_{\widetilde{U}_i})$ is a normed space.

(A8) $(U_i, \|\cdot\|_{U_i})$ is a Hilbert space, $(H_i, \|\cdot\|_{H_i})$ is a uniformly smooth and uniformly convex Banach space with duality map $J_{H_i} : H_i \to H_i^*$, and $(\widetilde{U}_i, \|\cdot\|_{\widetilde{U}_i})$ is a normed space.

(A9) $(U_i, \|\cdot\|_{U_i})$, $(X, \|\cdot\|_X)$ are Banach spaces, $(Y, \|\cdot\|_Y)$ is a uniformly smooth and uniformly convex Banach space with duality map $J_Y : Y \to Y^*$, and $e : X \to Y$ is a linear and bounded operator with $e^{-1}(K_Y) = K$ for $K_Y \subseteq Y$ being convex and closed.

(A10) $(U_i, \|\cdot\|_{U_i})$ is a Hilbert space, $(X, \|\cdot\|_X)$ is a Banach spaces, and $(H_i, \|\cdot\|_{H_i})$, $(Y, \|\cdot\|_Y)$ are uniformly smooth and uniformly convex Banach spaces with duality maps $J_{H_i} : H_i \to H_i^*$ and $J_Y : Y \to Y^*$, and $e : X \to Y$ is a linear and bounded operator with $e^{-1}(K_Y) = K$ for $K_Y \subseteq Y$ being convex and closed.

In the case of (A1)–(A8) and (A10), let $\iota_{H_i} : U_i \to H_i$ be a linear, completely continuous, and injective operator. Furthermore, we consider a linear and completely continuous operator $\iota_{\widetilde{U}_i} : U_i \to \widetilde{U}_i$ in the case of (A4)–(A8). As we see in Remark 2.16 for reflexive spaces completely continuous operators are compact. The corresponding product spaces are denoted by $U = \prod_{i\in[N]} U_i$, $\widetilde{U} = \prod_{i\in[N]} \widetilde{U}_i$ and $H = \prod_{i\in[N]} H_i$. Moreover, we define the embeddings $\iota_{\widetilde{U}} : U \to \widetilde{U}$ and $\iota_H : U \to H$ by $\iota_{\widetilde{U}}(u) = (\iota_{\widetilde{U}_i}(u^i), \iota_{\widetilde{U}_{-i}}(u^{-i}))$ and $\iota_H(u) = (\iota_{H_i}(u^i), \iota_{H_{-i}}(u^{-i}))$, respectively. Note that the product spaces have the same properties as their components using the product norms given by (2.1), see Remark 2.46. The duality maps of U, H, and Y are indicated by $J_U : U \to U^*$, $J_H : H \to H^*$, and $J_Y : Y \to Y^*$, respectively. These are defined as usual, see Definition 2.53.

Lastly, we remark that the condition $e^{-1}(K_Y) = K$ of (A9) and (A10) can be equivalently stated as the following lemma shows.

Lemma 3.2. It holds that $e^{-1}(K_Y) = K$ if and only if $x \in K$ is equivalent to $e(x) \in K_Y$.

Proof. We begin the proof with the backward direction "$\Leftarrow$". Let the equivalence statement "$x \in K \Leftrightarrow e(x) \in K_Y$" be valid. Then we have

$$e^{-1}(K_Y) = \{x \in X : e(x) \in K_Y\} = K.$$

Next, we prove the forward direction "⇒". If it holds $e^{-1}(K_Y) = K$, then we observe that $K \subseteq e^{-1}(K_Y)$ is equivalent to $e(K) \subseteq K_Y$. In other words, for any $x \in K$ we obtain $e(x) \in K_Y$. Furthermore, the assumption $e^{-1}(K_Y) \subseteq K$ directly yields $x \in K$ for $e(x) \in K_Y$. □

3.2 Continuity and Differentiability Assumptions

Depending on their nature, our results necessitate varying notions of continuity and differentiability assumptions on the objective functionals θ_i, $i \in [N]$, which we describe below. Note that we provide multiple assumptions, but choose just one at a time when developing specific result. In the following, let U and $\widetilde{U}$ be normed spaces.

Assumption 3.3.

(B1) $\theta_i : U \to \mathbb{R}$ is Fréchet differentiable in the i-th component.

(B2) $\theta_i : U \to \mathbb{R}$ is of the form $\theta_i(u) = \tilde{\theta}_i(\iota_{\widetilde{U}}(u)) + \frac{\gamma}{2}\|u^i\|^2_{U_i}$ and one of the following conditions holds

- (B2a) $\tilde{\theta}_i : \widetilde{U} \to \mathbb{R}$ is continuous.
- (B2b) $\tilde{\theta}_i : \widetilde{U} \to \mathbb{R}$ is continuously differentiable in the i-th component.
- (B2c) $\tilde{\theta}_i : \widetilde{U} \to \mathbb{R}$ is continuously differentiable.

We point out the difference between the derivative with respect to an element in U or to an element in $\widetilde{U}$. As before, we denote the partial derivative with respect to some element $\tilde{v}^i \in \widetilde{U}_i$ by $(\tilde{\theta}_i)_{\tilde{v}^i} : \widetilde{U}_i \to \widetilde{U}_i^*$. By the chain rule, it holds that

$$\begin{aligned}\big\langle \big[\tilde{\theta}_i(\iota_{\widetilde{U}_i}(v^i), \iota_{\widetilde{U}_{-i}}(u^{-i}))\big]_{v^i}, h^i\big\rangle_{U_i^*,U_i} &= \big\langle (\tilde{\theta}_i)_{\tilde{v}^i}(\iota_{\widetilde{U}_i}(v^i), \iota_{\widetilde{U}_{-i}}(u^{-i})), \iota_{\widetilde{U}_i}(h^i)\big\rangle_{\widetilde{U}_i^*,\widetilde{U}_i} \\ &= \big\langle \iota^*_{\widetilde{U}_i}(\tilde{\theta}_i)_{\tilde{v}^i}(\iota_{\widetilde{U}_i}(v^i), \iota_{\widetilde{U}_{-i}}(u^{-i})), h^i\big\rangle_{U_i^*,U_i},\end{aligned} \tag{3.1}$$

for all $h^i \in U_i$. Here, we see that we have transformed the dual pairing in U_i to the one in $\widetilde{U}_i$ by an application of the adjoint operator of $\iota_{\widetilde{U}_i}$. Clearly, under (B2c) of Assumption 3.3, the objective functional $\theta_i : U \to \mathbb{R}$ is continuously differentiable.

3.3 Convexity Assumptions

In this section, we study the choice of the parameter $\alpha \geq 0$ appearing in the definition of the regularized Nikaido–Isoda functional, see (2.20). This parameter is chosen according to Assumption 3.4 in such a way that it guarantees suitable convexity properties. These might result in a large value of α. We want to avoid possibly quite large values of α, which might slow down numerical globally convergent methods that use $\widetilde{V}_\alpha$ or V_α. Hence, we study local properties of the GNEP consisting of (2.7) and to this end, let $\overline{B}_R(u^i) \subseteq U_i$ be the closed ball with radius $R > 0$ centered at $u^i \in F_i(u^{-i})$. Furthermore, let $\overline{B}_R(u) \subseteq U$ be the closed ball with radius R and center $u \in \mathcal{X}$. We define the intersections

$$\widetilde{F}_i(u) = F_i(u^{-i}) \cap \overline{B}_R(u^i), \qquad \widetilde{\mathcal{X}}(u) = \mathcal{X} \cap \overline{B}_R(u).$$

Moreover, we introduce the continuously differentiable penalty-type or barrier-type functional $p : [0,\infty) \to [0,\infty]$ with $p(0) = 0$. The corresponding monotonically increasing or decreasing sequence of parameters are given by $\{\rho_k^i\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_{>0}$ and $\{\rho_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_{>0}$, respectively. Another possibility for localizing is the combination of a further restriction of the feasible set and an addition of a special term. To this end, one defines the convex sets

$$\widehat{F}_i(u) \subseteq F_i(u^{-i}) \ \text{ with } \ F_i(u^{-i}) \cap \overline{B}_R(u^i) \subseteq \widehat{F}_i(u), \qquad \widehat{\mathcal{X}}(u) \subseteq \mathcal{X} \ \text{ with } \ \mathcal{X} \cap \overline{B}_R(u) \subseteq \widehat{\mathcal{X}}(u).$$

For $i \in [N]$, let $q_i : U_i \to [0,\infty]$ be convex, continuously differentiable at the point 0 with $q_i(0) = 0$, and let q_i be finite on $\overline{B}_R^{U_i}(0)$. We define the functional $q : U \to [0,\infty]$ analogously with the same properties.

In the following, we state several convexity assumptions on the regularized versions of θ_i, $i \in [N]$, and on the functionals $\widetilde{\Psi}_\alpha$. Again we collect them here and will require only one of them at a time for obtaining specific results.

Assumption 3.4. Let $\alpha \geq 0$ be such that one of the following conditions holds

(C1) $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2$ is pseudoconvex at u^i on $F_i(u^{-i})$ for all $i \in [N]$.

(C2) $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2$ is pseudoconvex at u^i on $\widetilde{F}_i(u)$ for all $i \in [N]$.

(C3) $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k^i p(\|\cdot - u^i\|_{U_i}^2)$ is pseudoconvex at u^i on $F_i(u^{-i})$ for all $i \in [N]$ and for an arbitrarily fixed $k \in \mathbb{N}$.

(C4) $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2 + q_i(\cdot - u^i)$ is pseudoconvex at u^i on $\widehat{F}_i(u)$ for all $i \in [N]$.

(C5) $\widetilde{\Psi}_\alpha(u, \cdot)$ is convex on $\mathcal{X}$,

(C6) $\widetilde{\Psi}_\alpha(u, \cdot)$ is pseudoconvex at u on $\mathcal{X}$.

(C7) $\widetilde{\Psi}_\alpha(u, \cdot)$ is pseudoconvex at u on $\widetilde{\mathcal{X}}(u)$.

(C8) $\widetilde{\Psi}_\alpha(u, \cdot) + \rho_k p(\|\cdot - u\|_U^2)$ is pseudoconvex at u on $\mathcal{X}$ for an arbitrarily fixed $k \in \mathbb{N}$.

(C9) $\widetilde{\Psi}_\alpha(u, \cdot) + q(\cdot - u)$ is pseudoconvex at u on $\widehat{\mathcal{X}}(u)$.

We note that (C1) implies (C3) and (C4). In fact, for $v^i \in F_i(u^{-i})$ the pseudoconvexity yields

$$\left\langle (\theta_i)_{v^i}(u), v^i - u^i \right\rangle_{U_i^*, U_i} \geq 0 \quad \overset{(C1)}{\Longrightarrow} \quad \theta_i(u) \leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2. \tag{3.2}$$

Furthermore, it holds for all $v^i \in F_i(u^{-i})$ and $v^i \in \widehat{F}_i(u) \subseteq F_i(u^{-i})$, respectively,

$$\theta_i(u) \leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2$$

$$
\begin{aligned}
&\leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k^i p(\|v^i - u^i\|_{U_i}^2), \\
\theta_i(u) &\leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \\
&\leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + q_i(v^i - u^i).
\end{aligned}
$$

Since the penalty-type or barrier-type term $p(\| \cdot - u^i\|_{U_i}^2)$ and $q_i(\cdot - u^i)$ are chosen such that they vanish at the point u^i, the corresponding function values correspond to $\theta_i(u)$. Furthermore, the derivatives of the functionals in (C3) and (C4) at u^i coincide to $(\theta_i)_{v^i}(u)$ due to the representation of the derivatives. Indeed, we see that both (C3) and (C4) hold. Accordingly, the assertions (C8) and (C9) apply under (C6).

Chapter 4

Existence of Fixed Points of the Solution Map

The greatest challenge to any thinker is stating the problem in a way that will allow a solution. (Bertrand Russell)

A drunk man will find his way home, but a drunk bird may get lost forever. (Shizuo Kakutani)

Doppelpunkt def

This chapter is dedicated to the study of the two minimization problems

$$\min_{v \in \mathcal{X}} \widetilde{\Psi}_\alpha(u, v) \qquad \text{and} \qquad \min_{v \in \mathcal{X}} \Phi_\alpha(u, v). \tag{4.1}$$

Here, the Nikaido–Isoda type functional $\widetilde{\Psi}_\alpha$, see (2.22), and the new objective functional Φ_α are given by

$$\widetilde{\Psi}_\alpha(u, v) = \sum_{i \in [N]} \Big[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\Big],$$
$$\Phi_\alpha(u, v) = \sum_{i \in [N]} \Big[\big\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i}, v^i \big\rangle_{U_i^*, U_i} + \gamma \langle J_{U_i}(u^i), v^i \rangle_{U_i^*, U_i}\Big] + \frac{\alpha}{2}\|\iota_H(v) - \iota_H(u)\|_H^2.$$

The objective of this chapter is to demonstrate the existence of a fixed point of the corresponding solution maps of the stated optimization problems. In this respect, we extensively rely on a variant of the Kakutani fixed point theorem, which was discovered by the Japanese mathematician Kakutani in 1941, see [64], and applied to game theory and economic issues by Nash as early as 1950, see [84]. We refer to [101] for applications of some fixed point theorems related to GNEPs in the case of convex functionals in Hilbert spaces. In this chapter, we extend such results to the nonconvex case and coarser topologies.

In Section 4.1, we investigate the left optimization problem in (4.1) and prove that the corresponding solution map admits a fixed point via a generalization of Kakutani's fixed point theorem. In Section 4.2, we proceed similar and study the right optimization problem in (4.1) with regards to its solution map.

4.1 Fixed Point of the Solution Map

In the following, we consider the optimization problem

$$\min_{v \in \mathcal{X}} \widetilde{\Psi}_\alpha(u, v) \tag{4.2}$$

with the objective functional

$$\widetilde{\Psi}_\alpha(u, v) = \sum_{i \in [N]} \big[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2} \|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big].$$

We show that the corresponding solution map $u \mapsto v_{\widetilde{\Psi}_\alpha}(u)$ has a fixed point.

From here on, we assume that (A5) holds for the underlying spaces. Furthermore, let $\theta_i : U \to \mathbb{R}$ fulfill the (B2a) of Assumption 3.3 and let the feasible set $\mathcal{X} \subseteq U$ be nonempty, convex, closed, and bounded. We choose $\alpha \geq 0$ such that $\widetilde{\Psi}_\alpha(u, \cdot) : U \to \mathbb{R}$ is convex. Consequently, (C5) of Assumption 3.4 holds.

The following proposition relates a solution to the optimization problem to the regularized Nikaido–Isoda merit functional $V_\alpha = \sup_{v \in \mathcal{X}} \Psi_\alpha(\cdot, v)$, which was already defined in (2.21) and partly investigated afterwards in Chapter 2.

Proposition 4.1. Let $u \in \mathcal{X}$ and assume that $v = u$ is a global solution to (4.2). Then it holds $V_\alpha(u) = 0$.

Proof. Since $u \in \mathcal{X}$ minimizes $\widetilde{\Psi}_\alpha(u, \cdot)$ on $\mathcal{X}$, it also maximizes $\Psi_\alpha(u, \cdot)$ on $\mathcal{X}$ and thus, we obtain

$$V_\alpha(u) = \Psi_\alpha(u, u) = 0.$$ □

Next up, we investigate the solution map $u \mapsto v_{\widetilde{\Psi}_\alpha}(u)$ for $\widetilde{\Psi}_\alpha$, defined by

$$v_{\widetilde{\Psi}_\alpha}(u) = \Big\{v \in U : v \text{ solves } \min_{v \in \mathcal{X}} \widetilde{\Psi}_\alpha(u, v)\Big\}. \tag{4.3}$$

In particular, we show that the solution map admits a fixed point using the generalization of the Kakutani fixed point theorem, see Theorem 2.8. The statement of the main theorem reads as follows.

Theorem 4.2. The solution map $u \mapsto v_{\widetilde{\Psi}_\alpha}(u)$ has a fixed point in $\mathcal{X}$.

In order to prove this result by an application of the Kakutani fixed point theorem, we verify the theorem's assumptions in the following propositions. To this end, we have to show that $\mathcal{X} \subseteq U$ is convex, closed, and compact with respect to the weak sequential topology. Moreover, we have to verify that the solution map $u \mapsto v_{\widetilde{\Psi}_\alpha}(u)$ is closed and has a convex image. In order to show that the solution mapping is well-defined, we check whether the solution map has a nonempty image.

In order to prove that the solution set contains at least one element, we first check that $\widetilde{\Psi}_\alpha : U \times U \to \mathbb{R}$ is lower semicontinuous with respect to the weak sequential topology.

Lemma 4.3. The functional $\widetilde{\Psi}_\alpha : U \times U \to \mathbb{R}$ is lower semicontinuous with respect to the weak sequential topology.

Proof. In order to show the lower semicontinuity of $\widetilde{\Psi}_\alpha : U \times U \to \mathbb{R}$ with respect to the weak sequential topology, we consider two arbitrary sequences $\{u_k\}_{k\in\mathbb{N}}$ and $\{v_k\}_{k\in\mathbb{N}}$ in U that fulfill $u_k \rightharpoonup \bar{u}$ and $v_k \rightharpoonup \bar{v}$ in U as $k \to \infty$. By the completely continuity of the operators $\iota_{\widetilde{U}} : U \to \widetilde{U}$ and $\iota_H : U \to H$, we obtain the strong convergences

$$\begin{aligned} \iota_{\widetilde{U}}(u_k) &\to \iota_{\widetilde{U}}(\bar{u}) && \text{in } \widetilde{U}, \\ \iota_{\widetilde{U}}(v_k) &\to \iota_{\widetilde{U}}(\bar{v}) && \text{in } \widetilde{U}, \\ \iota_H(u_k) &\to \iota_H(\bar{u}) && \text{in } H, \\ \iota_H(v_k) &\to \iota_H(\bar{v}) && \text{in } H, \end{aligned}$$

as $k \to \infty$. Consequently, we have

$$\lim_{k\to\infty} \iota_{\widetilde{U}}(v_k^i, u_k^{-i}) = \iota_{\widetilde{U}}(\bar{v}^i, \bar{u}^{-i}) \quad \text{in} \quad \widetilde{U},$$

for each $i \in [N]$, and by the sequential continuity of $\tilde{\theta}_i : \widetilde{U} \to \mathbb{R}$, $i \in [N]$, it yields

$$\lim_{k\to\infty} \tilde{\theta}_i(\iota_{\widetilde{U}}(v_k^i, u_k^{-i})) = \tilde{\theta}_i(\iota_{\widetilde{U}}(\bar{v}^i, \bar{u}^{-i})).$$

Furthermore, it holds

$$\lim_{k\to\infty} \|\iota_H(v_k) - \iota_H(u_k)\|_H = \|\iota_H(\bar{v}) - \iota_H(\bar{u})\|_H,$$

and using the lower semicontinuity of $\|\cdot\|_{U_i}^2$ in the weak sequential topology, see Lemma 2.23, we obtain the estimate

$$\begin{aligned} \widetilde{\Psi}_\alpha(\bar{u}, \bar{v}) &= \sum_{i\in[N]} \left[\tilde{\theta}_i(\iota_{\widetilde{U}}(\bar{v}^i, \bar{u}^{-i})) + \frac{\gamma}{2}\|\bar{v}^i\|_{U_i}^2\right] + \frac{\alpha}{2}\|\iota_H(\bar{v}) - \iota_H(\bar{u})\|_H^2 \\ &\le \lim_{k\to\infty} \Big[\sum_{i\in[N]} \left[\tilde{\theta}_i(\iota_{\widetilde{U}}(v_k^i, u_k^{-i}))\right] + \frac{\alpha}{2}\|\iota_H(v_k) - \iota_H(u_k)\|_H^2 \Big] + \frac{\gamma}{2} \sum_{i\in[N]} \left[\liminf_{k\to\infty} \|v_k^i\|_{U_i}^2\right] \end{aligned}$$

Since the limes inferior is superadditive, we get

$$\begin{aligned} \widetilde{\Psi}_\alpha(\bar{u}, \bar{v}) &\le \liminf_{k\to\infty} \Big[\sum_{i\in[N]} \left[\tilde{\theta}_i(\iota_{\widetilde{U}}(v_k^i, u_k^{-i})) + \frac{\gamma}{2}\|v_k^i\|_{U_i}^2\right] + \frac{\alpha}{2}\|\iota_H(v_k) - \iota_H(u_k)\|_H^2 \Big] \\ &= \liminf_{k\to\infty} \widetilde{\Psi}_\alpha(u_k, v_k). \end{aligned}$$

□

We could replace the assumption of the complete continuity of the embedding operator $\iota_{\widetilde{U}} : U \to \widetilde{U}$ by assuming that $\tilde{\theta}_i : U \to \mathbb{R}$ is completely continuous for all $i \in [N]$, i.e.,

$$u_k \rightharpoonup \bar{u} \quad \text{in} \quad U \quad \text{as } k \to \infty \qquad \Longrightarrow \qquad \tilde{\theta}_i(u_k) \to \tilde{\theta}_i(\bar{u}) \quad \text{as } k \to \infty.$$

Alternatively, we could also replace it by assuming that the embedding $\tilde{\iota}_{\widetilde{U}}$ is continuous and $\tilde{\theta}_i : \widetilde{U} \to \mathbb{R}$ completely continuous for all $i \in [N]$. This implies again $\tilde{\theta}_i(\tilde{\iota}_{\widetilde{U}}(u_k)) \to \tilde{\theta}_i(\tilde{\iota}_{\widetilde{U}}(\bar{u}))$ for $u_k \rightharpoonup \bar{u}$ in U as $k \to \infty$ for all $i \in [N]$.

Next, we make sure that a solution to the problem exists, i.e., the solution set is nonempty.

Lemma 4.4. Let $u \in U$ be arbitrary. Then the solution set $v_{\widetilde{\Psi}_\alpha}(u)$ to (4.2) is nonempty and convex.

Proof. We note that U is a reflexive Banach space and $\mathcal{X}$ is assumed to be nonempty, convex, closed, and bounded. Furthermore, we know by Lemma 4.3 that the functional $\widetilde{\Psi}_\alpha(u,\cdot) : U \to \mathbb{R}$ is lower semicontinuous with respect to the weak sequential topology and with respect to the norm topology. Thus, we can apply Lemma 2.61 to conclude that $\min_{v\in\mathcal{X}} \widetilde{\Psi}_\alpha(u,v)$ possesses at least one solution with optimal value $\widetilde{\Psi}^*_\alpha(u)$. Hence, we have shown that $v_{\widetilde{\Psi}_\alpha}(u)$ is nonempty.

The convexity of the solution set $v_{\widetilde{\Psi}_\alpha}(u)$ is a consequence of the convexity of $\widetilde{\Psi}_\alpha(u,\cdot)$ and $\mathcal{X}$. Indeed, for given solutions $x, y \in v_{\widetilde{\Psi}_\alpha}(u)$ with $\lambda x + (1-\lambda)y \in \mathcal{X}$, $\lambda \in [0,1]$ arbitrary, it holds that

$$\widetilde{\Psi}^*_\alpha(u) \le \widetilde{\Psi}_\alpha(u, \lambda x + (1-\lambda)y) \le \lambda\widetilde{\Psi}_\alpha(u,x) + (1-\lambda)\widetilde{\Psi}_\alpha(u,y) = \widetilde{\Psi}^*_\alpha(u).$$

Hence, we obtain $\lambda x + (1-\lambda)y \in v_{\widetilde{\Psi}_\alpha}(u)$. $\square$

Next, we prove that the solution map is closed in the weak sequential topology, which is required for the application of the Kakutani fixed point theorem.

Lemma 4.5. Let $u \in U$ be arbitrary. Then the solution map $u \mapsto v_{\widetilde{\Psi}_\alpha}(u)$ to (4.2) is closed in the weak sequential topology.

Proof. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq U$ and $v_k \in v_{\widetilde{\Psi}_\alpha}(u_k), k \in \mathbb{N}$, be arbitrary sequences such that $u_k \rightharpoonup \bar{u}$ in U and $v_k \rightharpoonup \bar{v}$ in U as $k \to \infty$. We have to verify $\bar{v} \in v_{\widetilde{\Psi}_\alpha}(\bar{u})$. To this end, we apply Lemma 4.3 and obtain

$$\widetilde{\Psi}_\alpha(\bar{u},\bar{v}) \le \liminf_{k\to\infty} \widetilde{\Psi}_\alpha(u_k, v_k).$$

For all $x \in \mathcal{X}$ and $y \in v_{\widetilde{\Psi}_\alpha}(\bar{u})$ it holds $\widetilde{\Psi}_\alpha(\bar{u},y) \le \widetilde{\Psi}_\alpha(\bar{u},x)$ and thus, we get the inequality

$$\begin{aligned}\liminf_{k\to\infty} \widetilde{\Psi}_\alpha(u_k,v_k) &\le \liminf_{k\to\infty} \widetilde{\Psi}_\alpha(u_k,x) \\ &= \liminf_{k\to\infty} \Big[\sum_{i\in[N]} \big[\tilde{\theta}_i(\iota_{\widetilde{U}}(x^i, u_k^{-i})) + \frac{\gamma}{2}\|x^i\|^2_{U_i}\big] + \frac{\alpha}{2}\|\iota_H(x) - \iota_H(u_k)\|^2_H \Big].\end{aligned}$$

Next, the fact that $\iota_{\widetilde{U}}$ and ι_H are completely continuous yields

$$\begin{aligned}\liminf_{k\to\infty} \widetilde{\Psi}_\alpha(u_k,v_k) &\le \sum_{i\in[N]} \big[\lim_{k\to\infty} \tilde{\theta}_i(\iota_{\widetilde{U}}(x^i, u_k^{-i})) + \frac{\gamma}{2}\|x^i\|^2_{U_i}\big] + \lim_{k\to\infty} \big[\frac{\alpha}{2}\|\iota_H(x) - \iota_H(u_k)\|^2_H\big] \\ &= \sum_{i\in[N]} \big[\tilde{\theta}_i(\iota_{\widetilde{U}}(x^i, \bar{u}^{-i})) + \frac{\gamma}{2}\|x^i\|^2_{U_i}\big] + \frac{\alpha}{2}\|\iota_H(x) - \iota_H(\bar{u})\|^2_H \\ &= \widetilde{\Psi}_\alpha(\bar{u},x).\end{aligned}$$

Summarizing, we obtain $\widetilde{\Psi}_\alpha(\bar{u},\bar{v}) \le \widetilde{\Psi}_\alpha(\bar{u},x)$ for all $x \in \mathcal{X}$. By $v_k \in v_{\widetilde{\Psi}_\alpha}(u_k) \subseteq \mathcal{X}$ and the closedness of $\mathcal{X}$ in U in the weak sequential topology, we conclude $\bar{v} \in \mathcal{X}$ and thus, $\bar{v}$ is a minimizer to $\widetilde{\Psi}_\alpha(\bar{u},\cdot)$ on $\mathcal{X}$. Finally, this proves $\bar{v} \in v_{\widetilde{\Psi}_\alpha}(\bar{u})$. $\square$

Proof of Theorem 4.2. Since $\mathcal{X} \subseteq U$ is convex and closed, it is closed in the weak topology of U. By this property and the boundedness of $\mathcal{X}$, we conclude that $\mathcal{X}$ is weakly compact in the reflexive Banach space U. Due to the Eberlein–Šmulian theorem, see Theorem 2.25, we know that weak compactness implies compactness in the weak sequential topology. The closedness of the solution map in the weak sequential topology holds due to Lemma 4.5. Moreover, we have obtained in Lemma 4.4 that this map has a nonempty and convex image. Thus, Glicksberg's version of the Kakutani fixed point theorem, see Theorem 2.8, is applicable and establishes the existence of a fixed point of the solution map. □

Overall, we have proved the existence of a fixed point of the solution map $u \mapsto v_{\widetilde{\Psi}_\alpha}(u)$ corresponding to the problem (4.2), i.e.,

$$\min_{v \in \mathcal{X}} \widetilde{\Psi}_\alpha(u, v)$$

with

$$\widetilde{\Psi}_\alpha(u, v) = \sum_{i \in [N]} \left[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2} \|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right].$$

In other words, there exists an element $u \in \mathcal{X}$ that fulfills both $u \mapsto v_{\widetilde{\Psi}_\alpha}(u) = u$ and $\widetilde{\Psi}_\alpha(u, u) \leq \widetilde{\Psi}_\alpha(u, v)$ for all $v \in \mathcal{X}$. This is equivalent to the property

$$\sum_{i \in [N]} \left[\theta_i(u^i, u^{-i}) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2} \|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] \leq 0$$

for all $v \in \mathcal{X}$. Hence, it holds $V_\alpha(u) = 0$. In Chapter 5, we analyze whether and under what assumptions this condition corresponds to $u \in \mathcal{X}$ being a variational equilibrium or normalized equilibrium, see Theorem 5.4 below.

4.2 Fixed Point of the Solution Map via Optimality Condition

In this section, we consider a solution map based on optimality requirements of the first order. In the end, we derive similar results as we have seen in the preceding section. Particularly, we apply the Kakutani fixed point theorem to obtain the existence of minimizers.

Since we consider solution maps based on a first-order optimality condition, we define the functional $\Phi : U \times U \to \mathbb{R}$ and its regularization $\Phi_\alpha : U \times U \to \mathbb{R}$ by

$$\begin{aligned} \Phi(u, v) &= \sum_{i \in [N]} \left[\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i}, v^i \rangle_{U_i^*, U_i} + \gamma \langle J_{U_i}(u^i), v^i \rangle_{U_i^*, U_i}\right], \\ \Phi_\alpha(u, v) &= \Phi(u, v) + \frac{\alpha}{2} \|\iota_H(v) - \iota_H(u)\|_H^2, \end{aligned} \tag{4.4}$$

where $\alpha \geq 0$ is chosen accordingly. Note that it holds

$$\Phi(u, v) = \sum_{i \in [N]} \left[\langle [\theta_i(u)]_{u^i}, v^i - u^i \rangle_{U_i^*, U_i}\right] + \text{const}(u),$$

and α could be chosen as zero. Here, the symbol $\operatorname{const}(u)$ denotes a constant that may possibly depend on u. In the following, we study the problem

$$\min_{v\in\mathcal{X}} \Phi_\alpha(u,v),$$

and show the existence of a fixed point of the corresponding solution map $u \mapsto v_{\Phi_\alpha}(u)$.

From here on, let (A6) be fulfilled for the underlying spaces. That means, we assume that H_i, $\widetilde{U}_i$ are both normed spaces and that U_i is a uniformly smooth and uniformly convex Banach space. Moreover, we assume that the feasible set $\mathcal{X} \subseteq U$ is nonempty, convex, closed, and bounded. Lastly, we have to set some assumptions on the differentiability of the objective functional $\theta_i : U \to \mathbb{R}$. In particular, let (B2b) of Assumption 3.3 be valid, i.e., it is assumed that it holds $\theta_i(u) = \tilde{\theta}_i(\iota_{\widetilde{U}}(u)) + \frac{\gamma}{2}\|u^i\|^2_{U_i}$ with $\tilde{\theta}_i : \widetilde{U} \to \mathbb{R}$ being a functional that is continuously differentiable in its i-th component. In particular, the notation $(\widetilde{\theta}_i)_{\tilde{u}^i}$ expresses the derivative of $\widetilde{\theta}_i$ in $\widetilde{U}_i^*$, see Section 3.2.

We state the following theorem on the existence of fixed points of the minimization problem.

Theorem 4.6. There exists a fixed point of the mapping $u \mapsto v_{\Phi_\alpha}(u)$ on $\mathcal{X}$.

In order to prove this result, we would like to perform a similar calculation as before. However, this time the objective functional contains the terms

$$\langle J_{U_i}(u^i), v^i\rangle_{U_i^*,U_i},$$

which are not continuous with respect to the weak sequential topology for all $i \in [N]$. For this reason, instead of Φ and Φ_α we consider $\widetilde{\Phi} : U \times U \to \mathbb{R}$ and its regularization $\widetilde{\Phi}_\alpha$ that we define by

$$\begin{aligned} \widetilde{\Phi}(u,v) &= \sum_{i\in[N]} \Big[\big\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i}, v^i\big\rangle_{U_i^*,U_i} + \frac{\gamma}{2}\|v^i\|^2_{U_i}\Big], \\ \widetilde{\Phi}_\alpha(u,v) &= \widetilde{\Phi}(u,v) + \frac{\alpha}{2}\|\iota_H(v) - \iota_H(u)\|^2_H. \end{aligned} \tag{4.5}$$

First of all, we prove the simultaneous existence of a fixed point of the solution maps corresponding to the minimizing problems $\min_{v\in\mathcal{X}} \Phi_\alpha(u,v)$ and $\min_{v\in\mathcal{X}} \widetilde{\Phi}_\alpha(u,v)$. We proceed as in the section before and show both the lower semicontinuity in the weak sequential topology as well the well-posedness of the solution map.

Proposition 4.7. If it holds for all $i \in [N]$ that

$$\langle J_{U_i}(y), x\rangle_{U_i^*,U_i} = \langle J_{U_i}(x), y\rangle_{U_i^*,U_i} \qquad \forall\, x,y \in U_i, \tag{4.6}$$

then for a given $u \in \mathcal{X}$ it yields the following equivalency:

$$\Phi_\alpha(u,u) \le \Phi_\alpha(u,v) \quad \forall\, v\in\mathcal{X} \qquad \Longleftrightarrow \qquad \widetilde{\Phi}_\alpha(u,u) \le \widetilde{\Phi}_\alpha(u,v) \quad \forall\, v\in\mathcal{X}.$$

Proof. First, we begin with the direction "⇒" of the equivalency. Let $\Phi_\alpha(u,u) \leq \Phi_\alpha(u,v)$ be valid for all $v \in \mathcal{X}$. By definition of $\widetilde{\Phi}$, see (4.5), and of the dual product of the product space U, we obtain

$$\begin{aligned}\big\langle(\widetilde{\Phi})_v(u,v),h\big\rangle_{U^*,U} &= \Big\langle\Big(\sum_{i\in[N]}\big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i},v^i\big\rangle_{U_i^*,U_i}+\frac{\gamma}{2}\|v^i\|^2_{U_i}\big]\Big)_v,h\Big\rangle_{U^*,U}\\ &= \sum_{i\in[N]}\big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i},h^i\big\rangle_{U_i^*,U_i}+\gamma\langle J_{U_i}(v^i),h^i\rangle_{U_i^*,U_i}\big].\end{aligned} \tag{4.7}$$

Using the convexity of $\widetilde{\Phi}(u,\cdot)$ on $\mathcal{X}$ and its differentiability in the second component, it yields

$$\begin{aligned}\widetilde{\Phi}_\alpha(u,v)-\widetilde{\Phi}_\alpha(u,u) &= \widetilde{\Phi}(u,v)-\widetilde{\Phi}(u,u)+\frac{\alpha}{2}\|\iota_H(v)-\iota_H(u)\|^2_H\\ &\geq \big\langle(\widetilde{\Phi})_v(u,v)_{|v=u},v-u\big\rangle_{U^*,U}+\frac{\alpha}{2}\|\iota_H(v)-\iota_H(u)\|^2_H.\end{aligned}$$

Applying (4.7) we get

$$\begin{aligned}\widetilde{\Phi}_\alpha(u,v)-\widetilde{\Phi}_\alpha(u,u) \geq& \sum_{i\in[N]}\big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i},v^i-u^i\big\rangle_{U_i^*,U_i}+\gamma\langle J_{U_i}(u^i),v^i-u^i\rangle_{U_i^*,U_i}\big]\\ &+\frac{\alpha}{2}\|\iota_H(v)-\iota_H(u)\|^2_H\\ =& \sum_{i\in[N]}\big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i},v^i\big\rangle_{U_i^*,U_i}+\gamma\langle J_{U_i}(u^i),v^i\rangle_{U_i^*,U_i}\big]+\frac{\alpha}{2}\|\iota_H(v)-\iota_H(u)\|^2_H\\ &-\Big[\sum_{i\in[N]}\big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i},u^i\big\rangle_{U_i^*,U_i}+\gamma\langle J_{U_i}(u^i),u^i\rangle_{U_i^*,U_i}\big]+\frac{\alpha}{2}\|\iota_H(u)-\iota_H(u)\|^2_H\Big].\end{aligned}$$

Thus, we obtain

$$\widetilde{\Phi}_\alpha(u,v)-\widetilde{\Phi}_\alpha(u,u) \geq \Phi_\alpha(u,v)-\Phi_\alpha(u,u) \geq 0.$$

We conclude that u minimizes $\widetilde{\Phi}_\alpha(u,\cdot)$ on $\mathcal{X}$.

Next, we prove the other direction "⇐" of the equivalency. We assume that u is a minimizer to $\widetilde{\Phi}_\alpha(u,\cdot)$ on $\mathcal{X}$. In particular, it holds due to the convexity of $\mathcal{X}$

$$\widetilde{\Phi}_\alpha(u,u) \leq \widetilde{\Phi}_\alpha(u,(1-t)u+tv),$$

for $v \in \mathcal{X}$ and $t \in (0,1]$. Furthermore, the convexity of $\Phi_\alpha(u,\cdot)$ implies

$$\Phi_\alpha(u,(1-t)u+tv) \leq (1-t)\Phi_\alpha(u,u)+t\Phi_\alpha(u,v),$$

and we obtain the following inequality

$$\begin{aligned}\Phi_\alpha(u,v)-\Phi_\alpha(u,u) &\geq \frac{1}{t}[\Phi_\alpha(u,u+t(v-u))-\Phi_\alpha(u,u)]\\ &= \frac{1}{t}\Big[\sum_{i\in[N]}\big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i},u^i+t(v^i-u^i)\big\rangle_{U_i^*,U_i}\end{aligned}$$

$$
\begin{aligned}
&+ \gamma\langle J_{U_i}(u^i), u^i + t(v^i - u^i)\rangle_{U_i^*,U_i}\Big] + \frac{\alpha}{2}\|\iota_H(u + t(v-u)) - \iota_H(u)\|_H^2 \\
&- \Big(\sum_{i\in[N]} \big[\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i}, u^i\rangle_{U_i^*,U_i} + \gamma\langle J_{U_i}(u^i), u^i\rangle_{U_i^*,U_i}\big] \\
&+ \frac{\alpha}{2}\|\iota_H(u) - \iota_H(u)\|_H^2\Big)\Big].
\end{aligned}
$$

Selecting the terms of $\widetilde{\Phi}_\alpha$ we get

$$
\begin{aligned}
\Phi_\alpha(u,v) - \Phi_\alpha(u,u) \geq \frac{1}{t}\Big[&\widetilde{\Phi}_\alpha(u, u + t(v-u)) - \widetilde{\Phi}_\alpha(u,u) - \sum_{i\in[N]}\Big[\frac{\gamma}{2}\|u^i + t(v^i - u^i)\|_{U_i}^2 \\
&- \gamma\langle J_{U_i}(u^i), u^i + t(v^i - u^i)\rangle_{U_i^*,U_i}\Big] + \sum_{i\in[N]}\Big[\frac{\gamma}{2}\|u^i\|_{U_i}^2 - \gamma\langle J_{U_i}(u^i), u^i\rangle_{U_i^*,U_i}\Big]\Big].
\end{aligned}
$$

In the next step, we use that u is a minimizer to $\widetilde{\Phi}_\alpha(u,\cdot)$ on $\mathcal{X}$ and extend the squared norm. Then we arrive at

$$
\begin{aligned}
\Phi_\alpha(u,v) - \Phi_\alpha(u,u) &\geq \frac{1}{t}\Big[\sum_{i\in[N]}\Big[\gamma\langle J_{U_i}(u^i), t(v^i - u^i)\rangle_{U_i^*,U_i} - \frac{\gamma}{2}\|u^i + t(v^i-u^i)\|_{U_i}^2 + \frac{\gamma}{2}\|u^i\|_{U_i}^2\Big]\Big] \\
&= \sum_{i\in[N]}\Big[\gamma\langle J_{U_i}(u^i), v^i - u^i\rangle_{U_i^*,U_i} - \frac{\gamma}{2t}\|u^i\|_{U_i}^2 - \frac{\gamma}{2t}\big[\langle J_{U_i}(u^i), t(v^i-u^i)\rangle_{U_i^*,U_i} \\
&\quad + \langle J_{U_i}(t(v^i - u^i)), u^i\rangle_{U_i^*,U_i}\big] - \frac{\gamma}{2t}\|t(v^i-u^i)\|_{U_i}^2 + \frac{\gamma}{2t}\|u^i\|_{U_i}^2\Big] \\
&= -\sum_{i\in[N]}\frac{\gamma t}{2}\|v^i - u^i\|_{U_i}^2,
\end{aligned}
$$

where we used the assumption (4.6) on the duality map. The right-hand side tends to 0 as $t \to 0$ and therefore, we conclude that $u \in \mathcal{X}$ is the minimizer to $\Phi_\alpha(u,\cdot)$. □

The condition $\langle J_{U_i}(y), x\rangle_{U_i^*,U_i} = \langle J_{U_i}(x), y\rangle_{U_i^*,U_i}$ for any $x, y \in U_i$ is equivalent to $J_{U_i} = J_{U_i}^*\widetilde{J}$ for all $i \in [N]$. In fact, we can make use of the biduality mapping $\widetilde{J}$, see Definition 2.19, to obtain

$$
\langle J_{U_i}(x), y\rangle_{U_i^*,U_i} = \langle \widetilde{J}y, J_{U_i}(x)\rangle_{U_i^{**},U_i^*} = \langle J_{U_i}^*(\widetilde{J}y), x\rangle_{U_i^*,U_i} \qquad \forall\, x,y \in U_i.
$$

Together with the bijectivity of the duality mapping J_{U_i}, it yields $J_{U_i} = J_{U_i}^*\widetilde{J}$. This assumption is satisfied, for example, if U_i is a Hilbert space.

We have demonstrated that fixed points can only exists concurrently for the two minimization issues $\min_{v\in\mathcal{X}}\Phi_\alpha(u,v)$ and $\min_{v\in\mathcal{X}}\widetilde{\Phi}_\alpha(u,v)$. Consequently, we can complete the following proofs using the regularized functional $\widetilde{\Phi}_\alpha$ rather than Φ_α itself. Afterwards, we can still conclude the existence of fixed points of the original minimization problem.

We start the procedure by showing that the regularized functional is, in fact, lower semicontinuous with respect to the weak sequential topology.

Lemma 4.8. The functional $\widetilde{\Phi}_\alpha : U \times U \to \mathbb{R}$ is lower semicontinuous with respect to the weak sequential topology.

Proof. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq U$ and $\{v_k\}_{k\in\mathbb{N}} \subseteq U$ be weakly convergent sequences with $u_k \rightharpoonup \bar{u}$ and $v_k \rightharpoonup \bar{v}$ in U as $k \to \infty$. In the following, let $i \in [N]$ be arbitrary. Since it holds the strong convergence $\iota_{\widetilde{U}}(u_k) \to \iota_{\widetilde{U}}(\bar{u})$ in $\widetilde{U}$ as $k \to \infty$, it implies that $(\tilde{\theta}_i)_{\tilde{u}^i}(\iota_{\widetilde{U}}(u_k)) \to (\tilde{\theta}_i)_{\tilde{u}^i}(\iota_{\widetilde{U}}(\bar{u}))$ in $\widetilde{U}_i^*$ as $k \to \infty$. By the weak convergence of v_k^i in U_i as $k \to \infty$, we obtain $\iota_{\widetilde{U}_i}(v_k^i) \to \iota_{\widetilde{U}_i}(\bar{v}^i)$ in $\widetilde{U}_i$ as $k \to \infty$. Hence, we arrive at the convergence

$$\begin{aligned}\lim_{k\to\infty} \langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u_k))]_{u^i}, v_k^i \rangle_{U_i^*,U_i} &= \lim_{k\to\infty} \langle (\tilde{\theta}_i)_{\tilde{u}^i}(\iota_{\widetilde{U}}(u_k)), \iota_{\widetilde{U}_i}(v_k^i) \rangle_{\widetilde{U}_i^*,\widetilde{U}_i} \\ &= \langle (\tilde{\theta}_i)_{\tilde{u}^i}(\iota_{\widetilde{U}}(\bar{u})), \iota_{\widetilde{U}_i}(\bar{v}^i) \rangle_{\widetilde{U}_i^*,\widetilde{U}_i} = \langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(\bar{u}))]_{u^i}, \bar{v}^i \rangle_{U_i^*,U_i}.\end{aligned}$$

Since the norm functional $\|\cdot\|_{U_i}^2$ is lower semicontinuous in the weak sequential topology, we obtain

$$\begin{aligned}\widetilde{\Phi}_\alpha(\bar{u},\bar{v}) &= \sum_{i\in[N]} \Big[\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(\bar{u}))]_{u^i}, \bar{v}^i \rangle_{U_i^*,U_i} + \frac{\gamma}{2}\|\bar{v}^i\|_{U_i}^2 \Big] + \frac{\alpha}{2}\|\iota_H(\bar{v}) - \iota_H(\bar{u})\|_H^2 \\ &\le \lim_{k\to\infty} \Big[\sum_{i\in[N]} \Big[\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u_k))]_{u^i}, v_k^i \rangle_{U_i^*,U_i} \Big] + \frac{\alpha}{2}\|\iota_H(v_k) - \iota_H(u_k)\|_H^2 \Big] \\ &\quad + \frac{\gamma}{2} \sum_{i\in[N]} \Big[\liminf_{k\to\infty} \|v_k^i\|_{U_i}^2 \Big].\end{aligned}$$

The lower semicontinuity of $\widetilde{\Phi}_\alpha$ follows from the superadditivity of the limes inferior and

$$\begin{aligned}\widetilde{\Phi}_\alpha(\bar{u},\bar{v}) &\le \liminf_{k\to\infty} \Big[\sum_{i\in[N]} \Big[\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u_k))]_{u^i}, v_k^i \rangle_{U_i^*,U_i} + \frac{\gamma}{2}\|v_k^i\|_{U_i}^2 \Big] + \frac{\alpha}{2}\|\iota_H(v_k) - \iota_H(u_k)\|_H^2 \Big] \\ &= \liminf_{k\to\infty} \widetilde{\Phi}_\alpha(u_k, v_k).\end{aligned}$$ ◻

Alternatively, we could have assumed that $\tilde{\theta}_i(\cdot, u^{-i}) : U_i \to \mathbb{R}$ is differentiable for all $i \in [N]$ with $(\tilde{\theta}_i)_{u^i}(u) \in U_i^*$ being continuous with respect to the weak sequential topology. Then we would obtain

$$\langle [\tilde{\theta}_i(u)]_{u^i}, v^i \rangle_{U_i^*,U_i} = \langle (\tilde{\theta}_i)_{u^i}(u), v^i \rangle_{U_i^*,U_i},$$

and $(\tilde{\theta}_i)_{u^i}(u_k) \to (\tilde{\theta}_i)_{u^i}(\bar{u})$ in U_i^* for $u_k \rightharpoonup \bar{u}$ in U as $k \to \infty$ for all $i \in [N]$.

Next, we prove that the image of the corresponding solution map is, in fact, nonempty.

Lemma 4.9. The solution map $u \mapsto v_{\widetilde{\Phi}_\alpha}(u)$ of the minimizing problem $\min_{v\in\mathcal{X}} \widetilde{\Phi}_\alpha(u,v)$ is well-defined in the sense that $v_{\widetilde{\Phi}_\alpha}(u)$ is nonempty and convex.

Proof. By Lemma 4.8, we know that the functional $\widetilde{\Phi}_\alpha(u,\cdot)$ is lower semicontinuous with respect to the weak sequential topology. Hence, we can apply Lemma 2.61 that yields the existence of a solution to the minimization problem $\min_{v\in\mathcal{X}} \widetilde{\Phi}_\alpha(u,v)$. We conclude that the set $v_{\widetilde{\Phi}_\alpha}(u)$ is nonempty. The convexity follows analogously to the proof of Lemma 4.4. ◻

We continue with the next auxiliary result that provides us with the closedness of the solution map in the weak sequential topology.

Lemma 4.10. The solution map $u \mapsto v_{\widetilde{\Phi}_\alpha}(u)$ to the minimization problem $\min_{v\in\mathcal{X}} \widetilde{\Phi}_\alpha(u,v)$ is closed in the weak sequential topology.

Proof. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq U$ and $v_k \in v_{\widetilde{\Phi}_\alpha}(u_k)$, $k \in \mathbb{N}$, be two weakly convergent sequences in U. That means, there are elements $\bar{u}, \bar{v} \in U$ such that it holds $u_k \rightharpoonup \bar{u}$ and $v_k \rightharpoonup \bar{v}$ in U as $k \to \infty$. In order to show this lemma, we are going to prove that $\bar{v} \in v_{\widetilde{\Phi}_\alpha}(\bar{u})$. By Lemma 4.8, we obtain the inequality

$$\widetilde{\Phi}_\alpha(\bar{u},\bar{v}) \leq \liminf_{k\to\infty} \widetilde{\Phi}_\alpha(u_k, v_k),$$

and additionally, it holds the estimate

$$\widetilde{\Phi}_\alpha(\bar{u},y) \leq \widetilde{\Phi}_\alpha(\bar{u},x) \qquad \forall\, x \in \mathcal{X}, \quad y \in v_{\widetilde{\Phi}_\alpha}(\bar{u}).$$

Using these two results, we obtain the following inequality

$$\begin{aligned}
\widetilde{\Phi}_\alpha(\bar{u},\bar{v}) &\leq \liminf_{k\to\infty} \widetilde{\Phi}_\alpha(u_k, v_k) \\
&\leq \liminf_{k\to\infty} \widetilde{\Phi}_\alpha(u_k, x) \\
&= \liminf_{k\to\infty} \Big[\sum_{i\in[N]} \big[\big\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(u_k))]_{u^i}, x^i \big\rangle_{U_i^*,U_i} + \frac{\gamma}{2}\|x^i\|_{U_i}^2 \big] + \frac{\alpha}{2}\|\iota_H(x) - \iota_H(u_k)\|_H^2 \Big],
\end{aligned}$$

for any $x \in \mathcal{X}$. Since the operators $\iota_{\widetilde{U}}$ and ι_H are completely continuous, the same holds for the derivative of $\tilde{\theta}_i$. Thus, we get

$$\begin{aligned}
\widetilde{\Phi}_\alpha(\bar{u},\bar{v}) &\leq \sum_{i\in[N]} \big[\big\langle [\tilde{\theta}_i(\iota_{\widetilde{U}}(\bar{u}))]_{u^i}, x^i \big\rangle_{U_i^*,U_i} + \frac{\gamma}{2}\|x^i\|_{U_i}^2 \big] + \frac{\alpha}{2}\|\iota_H(x) - \iota_H(\bar{u})\|_H^2 \\
&= \widetilde{\Phi}_\alpha(\bar{u}, x),
\end{aligned}$$

for any $x \in \mathcal{X}$. By the definition of the solution map, we obtain that $\bar{v} \in \mathcal{X}$ is a minimizer. Hence, we are able to conclude the desired result $\bar{v} \in v_{\widetilde{\Phi}_\alpha}(\bar{u})$. ☐

Proof of Theorem 4.6. Analogously to the proof of Theorem 4.2, we obtain the compactness of $\mathcal{X}$ in the weak sequential topology. By Lemma 4.8, Lemma 4.9 and Lemma 4.10, the assumptions of the fixed point theorem of Kakutani, see Theorem 2.8, are satisfied and we can apply it to the map $u \mapsto v_{\widetilde{\Phi}_\alpha}(u)$. Hence, it yields the existence of a fixed point of $u \mapsto \operatorname{argmin}_{v\in\mathcal{X}} \widetilde{\Phi}_\alpha(u,v)$. By Proposition 4.7, we finally obtain a fixed point of the mapping $u \mapsto \operatorname{argmin}_{v\in\mathcal{X}} \Phi_\alpha(u,v)$. ☐

Let us now summarize this section. We have proved that a fixed point $u \in \mathcal{X}$ exists for the solution mapping corresponding to the minimization problem $\min_{v\in\mathcal{X}} \Phi_\alpha(u,v)$. Specifically, this means that the following inequality holds:

$$\Phi_\alpha(u,u) \leq \Phi_\alpha(u,v) \qquad \forall\, v \in \mathcal{X}.$$

Furthermore, we can derive the inequality

$$\sum_{i\in[N]} \Big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i}, u^i\big\rangle_{U_i^*,U_i} + \gamma\langle J_{U_i}(u^i), u^i\rangle_{U_i^*,U_i}\Big] \\ \leq \sum_{i\in[N]} \Big[\big\langle[\tilde{\theta}_i(\iota_{\widetilde{U}}(u))]_{u^i}, v^i\big\rangle_{U_i^*,U_i} + \gamma\langle J_{U_i}(u^i), v^i\rangle_{U_i^*,U_i}\Big] + \frac{\alpha}{2}\|\iota_H(v) - \iota_H(u)\|_H^2,$$

or equivalently, it holds

$$\sum_{i\in[N]} \Big[\big\langle(\theta_i)_{u^i}(u), u^i\big\rangle_{U_i^*,U_i}\Big] \leq \sum_{i\in[N]} \Big[\big\langle(\theta_i)_{u^i}(u), v^i\big\rangle_{U_i^*,U_i}\Big] + \frac{\alpha}{2}\|\iota_H(v) - \iota_H(u)\|_H^2. \tag{4.8}$$

Especially in the case $\alpha = 0$, one obtains the VI

$$\sum_{i\in[N]} \big\langle(\theta_i)_{u^i}(u), v^i - u^i\big\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v \in \mathcal{X}. \tag{4.9}$$

This VI can also be obtained for arbitrary $\alpha \geq 0$. In fact, for any $v \in \mathcal{X}$ and $t \in (0, 1]$, one inserts $u + t(v - u) \in \mathcal{X}$ into (4.8) and achieves the estimate

$$\sum_{i\in[N]} \big\langle(\theta_i)_{u^i}(u), t(v^i - u^i)\big\rangle_{U_i^*,U_i} \geq -\frac{\alpha t^2}{2}\|\iota_H(v) - \iota_H(u)\|_H^2.$$

From here, we divide this inequality by t and pass the limit $t \downarrow 0$, which finally yields (4.9). In the following section, we study the connection between this VI and normalized equilibria. Particularly, we refer to Theorem 5.3 and Theorem 5.4 below.

Chapter 5

Nikaido–Isoda Merit Functionals

One should avoid by all means to deprive oneself of analytical flexibility by adhering to any single technique. It is essential to choose a method most appropriate to the kind and nature of a problem under study. (Hukukane Nikaido)

This chapter investigates the analytical features of Nikaido–Isoda merit functionals and their local variations. There are a variety of methods for localizing such functionals, and we will explore a few of them here. We shall investigate the origins of merit functionals and determine their relationship with various-types of equilibria. Particularly, we will observe that merit functionals serve as a link between equilibria and optimization problems. Such linkages have been utilized in a number of previous publications, and in the finite-dimensional case, we specifically mention the extensive book [67].

In Section 5.1, we give an overview of the various-types of localized Nikaido–Isoda merit functionals. In Subsection 5.2.1, we demonstrate that Nash equilibria are the zeros of particular Nikaido–Isoda merit functionals that contain the supremum operator within the sum over the players. In contrast to this-type, we will examine other Nikaido–Isoda merit functionals in Subsection 5.2.2 and see that, under certain assumptions, their zeros are in fact normalized equilibrium points. In Section 5.3, we investigate the difference between two regularized Nikaido–Isoda merit functionals and demonstrate that the roots of the difference functional are strongly associated with the roots of the two original merit functionals. The differentiation of regularized and localized merit functionals follows. Subsection 5.4.2 discusses the differentiability of the regularized and localized variants by applying Danskin's theorem, see Section 2.4.

5.1 Localization of Merit Functionals

In this section, we investigate localized versions of the regularized merit functionals V_α and $\widetilde{V}_\alpha$. Recalling their definitions in Chapter 2, see (2.21) and (2.23), they are given by

$$\widetilde{V}_\alpha(u) = \sum_{i\in[N]} \sup_{v^i\in F_i(u^{-i})} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right], \tag{5.1}$$

$$V_\alpha(u) = \sup_{v\in\mathcal{X}} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big]. \tag{5.2}$$

To avoid potentially very large values of the regularization parameter $\alpha \geq 0$, the merit functionals will be localized. There are several techniques for localizing the merit functionals, as indicated in Section 3.3, and we investigate these in this section.

5.1.1 Localizing the Feasible Set

In the first option we will give, the admissible sets $F_i(u^{-i})$ and $\mathcal{X}$ are localized, and the convexity assumptions are imposed only in a neighborhood surrounding u^i or u.

Particularly, the admissible set is restricted to the closed balls centered around u^i and u for a given radius $R > 0$, i.e., the sets $\overline{B}_R(u^i) \subseteq U_i$ and $\overline{B}_R(u) \subseteq U$ are of interest. On the intersections $\widetilde{F_i}(u) = F_i(u^{-i}) \cap \overline{B}_R(u^i)$ and $\widetilde{\mathcal{X}}(u) = \mathcal{X} \cap \overline{B}_R(u)$, we define the regularized and localized Nikaido–Isoda merit functionals as follows:

$$\widetilde{V}_\alpha^{\mathrm{loc}}(u) = \sum_{i\in[N]} \sup_{v^i\in\widetilde{F_i}(u)} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big], \tag{5.3}$$

$$V_\alpha^{\mathrm{loc}}(u) = \sup_{v\in\widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big]. \tag{5.4}$$

5.1.2 Penalty-type and Barrier-type Method

Instead of directly localizing the feasible set, we will also employ penalty-type or barrier-type terms in the merit functionals of Nikaido–Isoda-type. This approach will lead to a localized solution nonetheless. Let $p : [0,\infty) \to [0,\infty]$ be a continuously differentiable penalty-type or barrier-type function with $p(0) = 0$. Furthermore, we introduce the parameter sequences $\{\rho_k^i\}_{k\in\mathbb{N}} \subseteq [0,\infty)$ and $\{\rho_k\}_{k\in\mathbb{N}} \subseteq [0,\infty)$ corresponding to the function p, which is chosen as a penalty-type or a barrier-type function. The difference lies in the region of penalization and monotonicity of the parameter sequence. The penalty-type function p is defined by the key property $p(\|v^i - u^i\|_{U_i}^2) = 0$ for $v^i \in \overline{B}_R(u^i)$, and the corresponding sequence of parameters $\{\rho_k^i\}_{k\in\mathbb{N}}$ is monotonically increasing. In the case of a barrier-type function, we define $p(\|v^i - u^i\|_{U_i}^2) \geq 0$ for $v^i \in B_R(u^i)$, $p(\|v^i - u^i\|_{U_i}^2) = \infty$ for $v \in U_i \backslash B_R(u^i)$, and the sequence $\{\rho_k^i\}_{k\in\mathbb{N}}$ is monotonically decreasing. In case of the feasible set $\mathcal{X}(u)$ and the parameter sequence $\{\rho_k\}_{k\in\mathbb{N}}$, the penalty-type and barrier-type functions are defined analogously.

We introduce the regularized and localized Nikaido–Isoda merit functionals with the addition of a penalty-type or barrier-type term by

$$\widetilde{V}_\alpha^{p_k}(u) = \sum_{i\in[N]} \sup_{v^i\in F_i(u^{-i})} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(\|v^i - u^i\|_{U_i}^2)\big], \tag{5.5}$$

$$V_\alpha^{p_k}(u) = \sup_{v\in\mathcal{X}} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] - \rho_k p(\|v - u\|_U^2)\Big]. \tag{5.6}$$

5.1.3 Combination of Localizing and Penalizing

As the third option, we combine both methods from the previous sections and simultaneously restrict the admissible set and include a penalty-type or barrier-type term in the Nikaido–Isoda merit functionals.
We introduce the restricted feasible sets $\widehat{F}_i(u)$ and $\widehat{\mathcal{X}}(u)$ with the following properties:

- $\widehat{F}_i(u) \subseteq F_i(u^{-i})$ is convex with $F_i(u^{-i}) \cap \overline{B}_R(u^i) \subseteq \widehat{F}_i(u)$ for some $R > 0$.
- $\widehat{\mathcal{X}}(u) \subseteq \mathcal{X}$ is convex with $\mathcal{X} \cap \overline{B}_R(u) \subseteq \widehat{\mathcal{X}}(u)$ for some $R > 0$.

Moreover, we define the penalty-type or barrier-type functionals $q_i : U_i \to [0, \infty]$, $i \in [N]$, and $q : U \to [0, \infty]$ by requiring that they are convex and continuously differentiable at 0 with $q_i(0) = 0$ and $q(0) = 0$. Furthermore, we assume that the functionals q_i and q are finite on $\overline{B}_R^{U_i}(0)$ and $\overline{B}_R^{U}(0)$, respectively.

As the third method, the regularized and localized merit functionals of Nikaido–Isoda-type are defined by

$$\widetilde{V}_\alpha^{\mathrm{lp}}(u) = \sum_{i\in[N]} \sup_{v^i \in \widehat{F}_i(u)} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i)\right], \tag{5.7}$$

$$V_\alpha^{\mathrm{lp}}(u) = \sup_{v\in\widehat{\mathcal{X}}(u)} \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] - q(v - u). \tag{5.8}$$

Since the penalty-type or barrier-type functional $q_i : U_i \to \mathbb{R}$ is assumed to be continuously differentiable with $q_i(0) = 0$ and $q_i(v^i) \geq 0$ for all $v^i \in U_i$, the element 0 is a local minimum and it holds $q_i'(0) = 0$. The same statement applies to q.

We note that as a special case, we can select the functional q as the sum over q_i. In addition, there are various generalizations that weaken the criteria of the parameter α while maintaining the validity of the results in Section 5.2. In fact, we make the following remarks:

- The parameter $\alpha \geq 0$ can be replaced by choosing $\alpha_i \geq 0$ for each $i \in [N]$ separately. Especially, one can interpret $\alpha \geq 0$ either as $\alpha = \max_{i\in[N]} \alpha_i$ or as $\alpha = (\alpha_i)_{i\in[N]}$.
- One can replace the assumption of pseudoconvexity in the closed ball $\overline{B}_R(u^i)$ or $\overline{B}_R(u)$ by the pseudoconvexity in an open neighborhood $B(u^i)$ or $B(u)$, respectively. In fact, one can find some open and closed balls such that $\overline{B}_\delta(u^i) \subseteq B_\varepsilon(u^i) \subseteq B(u^i)$ and $\overline{B}_\delta(u) \subseteq B_\varepsilon(u) \subseteq B(u)$ with $0 < \delta \leq \varepsilon$. Here, it is presumed that the balls are closed for reasons of convention.
- It would be even more general to include the regularization term $\frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2$ and $q_i(v^i - u^i)$ in a newly defined functional that is assumed to be convex, continuously differentiable and vanishes at u^i. We do not pursue this point further because we want to obtain a similar structure to the globally regularized Nikaido–Isoda merit functional, see (5.1).

We note that we do not require a convexity assumption on the functionals q and q_i for any $i \in [N]$. To avoid adding additional difficulties with nonconvexities, it is advantageous to localize and regularize nonetheless using a convex functional. In fact, a nonconvex regularization term may bring undesirable properties to the objective functional and could exacerbate its nonconvexity.

5.2 Merit Functionals and Equilibria

This section examines the regularized and localized functionals (5.1)–(5.8) and discusses their relationship to the various-types of equilibria. Specifically, we demonstrate that these functionals are, in fact, merit functionals to (Q)VIs that express first-order optimality criteria for (local) Nash or (local) normalized equilibria. Under certain pseudoconvexity assumptions, we uncover a relationship between the roots of the regularized merit functional and the roots of the regularized and localized versions. In addition, we link (local) Nash and (local) normalized equilibria to zeros of the respective merit functionals. Let (A3) be satisfied for the underlying spaces moving forward.

We note that only the assertions involving derivatives necessitate the space assumptions on U and H. In other circumstances, these assumptions can be reduced by using normed spaces. In addition, only the penalized functionals $\widetilde{V}_\alpha^{p_k}$ and $V_\alpha^{p_k}$ require that the space U is a uniformly smooth and uniformly convex Banach space.

5.2.1 Nash Equilibria

In this part of the section, we study the connection between the various merit functionals and Nash equilibria. We consider the regularized and localized Nikaido–Isoda merit functionals $\widetilde{V}_\alpha : U \to \mathbb{R}$, $\widetilde{V}_\alpha^{\text{loc}} : U \to \mathbb{R}$, $\widetilde{V}_\alpha^{p_k} : U \to \mathbb{R}$ and $\widetilde{V}_\alpha^{\text{lp}} : U \to \mathbb{R}$ as defined in (5.1), (5.3), (5.5) and (5.7). We relate them to QVIs that express first-order optimality conditions for (local) Nash equilibria.

Theorem 5.1. Let $k \in \mathbb{N}$ be arbitrarily fixed, $\alpha \geq 0$, and let $u \in U$ satisfy $u \in F(u)$. In the case of the assertions (ii)–(vi), we further assume that (B1) of Assumption 3.3 is satisfied and that the feasible set $F_i(u^{-i})$ is convex, $i \in [N]$. In the case of "$\Rightarrow$" of (ii), let the three assumptions (C2), (C3), (C4) of Assumption 3.4 hold for the three statements, respectively. Regarding "$\Rightarrow$" of (iv), (v) and (vi), we assume (C1). In the matter of "$\Rightarrow$" of (vii), we assume (C3) and for the reverse implication "$\Leftarrow$" that $p : [0, \infty) \to [0, \infty]$ is chosen as a penalty-type function. Then it holds:

(i) $\widetilde{V}_\alpha(u) \geq \widetilde{V}_\alpha^{\text{loc}}(u) \geq 0,$

$\widetilde{V}_\alpha(u) \geq \widetilde{V}_\alpha^{p_k}(u) \geq 0,$

$\widetilde{V}_\alpha(u) \geq \widetilde{V}_\alpha^{\text{lp}}(u) \geq 0.$

Moreover, if $p : [0, \infty) \to [0, \infty]$ is chosen as a penalty-type function, it holds $\widetilde{V}_\alpha^{p_k}(u) \geq \widetilde{V}_\alpha^{\text{loc}}(u)$.

(ii) $\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \quad \forall\, v^i \in \widetilde{F}_i(u), \quad i \in [N] \quad \Longleftrightarrow \quad \widetilde{V}_\alpha^{\text{loc}}(u) = 0,$

$\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \quad \forall\, v^i \in F_i(u^{-i}), \quad i \in [N] \quad \Longleftrightarrow \quad \widetilde{V}_\alpha^{p_k}(u) = 0,$

$$\big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in \widehat{F}_i(u), \qquad i \in [N] \qquad \Longleftrightarrow \qquad \widetilde{V}_\alpha^{\mathrm{lp}}(u) \;= 0.$$

(iii) For all $i \in [N]$:

$$\begin{aligned} &\big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}) \\ \Longleftrightarrow\;& \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in \widetilde{F}_i(u) \\ \Longleftrightarrow\;& \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in \widehat{F}_i(u). \end{aligned}$$

(iv) $\widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0 \qquad \Longleftrightarrow \qquad \widetilde{V}_\alpha(u) \;= 0.$

(v) $\widetilde{V}_\alpha^{p_k}(u) \;= 0 \qquad \Longleftrightarrow \qquad \widetilde{V}_\alpha(u) \;= 0.$

(vi) $\widetilde{V}_\alpha^{\mathrm{lp}}(u) \;= 0 \qquad \Longleftrightarrow \qquad \widetilde{V}_\alpha(u) \;= 0.$

(vii) $\widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0 \qquad \Longleftrightarrow \qquad \widetilde{V}_\alpha^{p_k}(u) = 0.$

Proof. We begin by proving that the inequalities of the first assertion are satisfied.

Proof of (i). Let $i \in [N]$ be arbitrary. By the inclusion $\widetilde{F}_i(u) \subseteq F_i(u^{-i})$, we obtain the first inequality of the claimed result, i.e., $\widetilde{V}_\alpha(u) \geq \widetilde{V}_\alpha^{\mathrm{loc}}(u)$. The fact $\widetilde{V}_\alpha(u) \geq \widetilde{V}_\alpha^{p_k}(u)$ follows from the nonnegativity of the penalty-type or barrier-type function p. Moreover, it holds $\widetilde{V}_\alpha(u) \geq \widetilde{V}_\alpha^{\mathrm{lp}}(u)$ since $q_i(u)$ is nonnegative for all $u \in U$ and $\widehat{F}_i(u)$ is a subset of $F_i(u^{-i})$. For the nonnegativity of the respective regularized and localized Nikaido–Isoda merit functionals, we plug $v^i = u^i$ into the respective definitions of $\widetilde{V}_\alpha^{\mathrm{loc}}$, $\widetilde{V}_\alpha^{p_k}$ and $\widetilde{V}_\alpha^{\mathrm{lp}}$, which results in

$$\begin{aligned} \widetilde{V}_\alpha^{\mathrm{loc}}(u) &\geq \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(u^i,u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(u^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] &&= 0, \\ \widetilde{V}_\alpha^{p_k}(u) &\geq \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(u^i,u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(u^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(0)\big] &&= 0, \\ \widetilde{V}_\alpha^{\mathrm{lp}}(u) &\geq \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(u^i,u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(u^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(0)\big] &&= 0. \end{aligned}$$

Regarding a penalty-type function it holds by definition that $p(\|v^i - u^i\|_{U_i}^2) > 0$ for all $v^i \notin \overline{B}_R(u^i)$. Hence, we obtain the inequality

$$\begin{aligned} &\sup_{v^i\in F_i(u^{-i})} \big[\theta_i(u) - \theta_i(v^i,u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] \\ &\qquad \geq \sup_{v^i\in F_i(u^{-i})} \big[\theta_i(u) - \theta_i(v^i,u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(\|v^i - u^i\|_{U_i}^2)\big]. \end{aligned}$$

Furthermore, we can estimate the supremum cutting the feasible set with the set

$$\{v^i \in U_i : p(\|v^i - u^i\|_{U_i}^2) = 0\}.$$

Then using the definition of a penalty-type function we arrive at

$$\begin{aligned}\sup_{v^i\in F_i(u^{-i})}&\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2-\rho_k^i p(\|v^i-u^i\|_{U_i}^2)\big]\\ &\geq \sup_{v^i\in F_i(u^{-i})\cap\{v^i:p(\|v^i-u^i\|_{U_i}^2)=0\}}\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\\ &\qquad\qquad -\rho_k^i p(\|v^i-u^i\|_{U_i}^2)\big]\\ &= \sup_{v^i\in F_i(u^{-i})\cap\overline{B}_R(u^i)}\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big].\end{aligned}$$

Rewriting this with the feasible set $\widetilde{F}_i(u)$ we get

$$\begin{aligned}\sup_{v^i\in F_i(u^{-i})}&\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big]\\ &\geq \sup_{v^i\in\widetilde{F}_i(u)}\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big].\end{aligned}$$

All together, we arrive at

$$\begin{aligned}\sup_{v^i\in F_i(u^{-i})}&\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big]\\ &\geq \sup_{v^i\in\widetilde{F}_i(u)}\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big].\end{aligned}$$

By setting $v^i=u^i\in\widetilde{F}_i(u)\subseteq F_i(u^{-i})$ for all $i\in[N]$ into the derived inequalities, we conclude that

$$\widetilde{V}_\alpha(u)\geq\widetilde{V}_\alpha^{p_k}(u)\geq\widetilde{V}_\alpha^{\mathrm{loc}}(u)\geq 0.$$

Proof of (ii). We begin by proving the direction "⇒" of the equivalency. Let $i\in[N]$ be arbitrary. Since the functional $\theta_i(\cdot,u^{-i})+\frac{\alpha}{2}\|\iota_{H_i}(\cdot)-\iota_{H_i}(u^i)\|_{H_i}^2$ is pseudoconvex at u^i on $\widetilde{F}_i(u)$, we obtain the estimate

$$\theta_i(u)\leq\theta_i(v^i,u^{-i})+\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2$$

for all $v^i\in\widetilde{F}_i(u)$. Here, we have applied the inequality

$$\Big\langle\Big(\big[\theta_i(v^i,u^{-i})+\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big]_{v^i}\Big)_{|_{v^i=u^i}},v^i-u^i\Big\rangle_{U_i^*,U_i}=\big\langle(\theta_i)_{u^i}(u),v^i-u^i\big\rangle_{U_i^*,U_i}\geq 0.$$

Consequently, this yields that $\widetilde{V}_\alpha^{\mathrm{loc}}(u)$ is nonpositive, i.e.,

$$\widetilde{V}_\alpha^{\mathrm{loc}}(u)=\sum_{i\in[N]}\sup_{v^i\in\widetilde{F}_i(u)}\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big]\leq 0.$$

However, we have already seen in the first part of the proof that $\widetilde{V}_\alpha^{\mathrm{loc}}$ is nonnegative and therefore, we conclude that u is a root of $\widetilde{V}_\alpha^{\mathrm{loc}}$.

Next, we prove the direction "⇐" of (ii). Let $i \in [N]$ again be arbitrary. First off, we show that roots of $\widetilde{V}_\alpha^{\mathrm{loc}}$ fulfill the inequality

$$\theta_i(u) \leq \theta_i(w^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(w^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \qquad \forall\, w^i \in \widetilde{F}_i(u). \tag{5.9}$$

Otherwise, there would exist an element $z^j \in \widetilde{F}_j(u)$ for some index $j \in [N]$ with

$$\theta_j(u) > \theta_j(z^j, u^{-j}) + \frac{\alpha}{2}\|\iota_{H_j}(z^j) - \iota_{H_j}(u^j)\|_{H_j}^2.$$

By choosing $v^j = z^j$ and $v^i = u^i$, $i \neq j$, in the definition of $\widetilde{V}_\alpha^{\mathrm{loc}}(u)$, this results in the inequality

$$\widetilde{V}_\alpha^{\mathrm{loc}}(u) \geq \theta_j(u) - \theta_j(z^j, u^{-j}) - \frac{\alpha}{2}\|\iota_{H_j}(z^j) - \iota_{H_j}(u^j)\|_{H_j}^2 > 0.$$

However, this positivity of $\widetilde{V}_\alpha^{\mathrm{loc}}(u)$ contradicts the fact that u is a root of $\widetilde{V}_\alpha^{\mathrm{loc}}$. Thus, the zero u fulfills the inequality (5.9). Exploiting the convexity of $\widetilde{F}_i(u)$, it holds for the convex combination $w^i = u^i + t(v^i - u^i) \in \widetilde{F}_i(u)$ for all $u^i, v^i \in \widetilde{F}_i(u)$ and $t \in (0,1]$. Using Taylor's expansion, it yields

$$\begin{aligned}\theta_i(u) &\leq \theta_i(u^i + t(v^i - u^i), u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(t(v^i - u^i))\|_{H_i}^2 \\ &= \theta_i(u^i, u^{-i}) + \big\langle (\theta_i)_{u^i}(u), t(v^i - u^i)\big\rangle_{U_i^*, U_i} + \mathrm{o}(t).\end{aligned}$$

Dividing by $t \in (0,1]$ and letting $t \downarrow 0$, we arrive at the desired inequality

$$\big\langle (\theta_i)_{u^i}(u), v^i - u^i\big\rangle_{U_i^*, U_i} \geq 0.$$

We note that the proofs for the two other equivalency statements in (ii) follow the preceding lines. Indeed, in the case of the direct implication we exploit the pseudoconvexity of the objective functionals at the point u^i and

$$\begin{aligned}&\big\langle (\theta_i)_{u^i}(u), v^i - u^i\big\rangle_{U_i^*, U_i} \\ &\quad = \Big\langle \Big(\big[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k^i p(\|v^i - u^i\|_{U_i}^2)\big]_{v^i}\Big)_{|_{v^i = u^i}}, v^i - u^i\Big\rangle_{U_i^*, U_i} \\ &\quad = \Big\langle \Big(\big[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + q_i(v^i - u^i)\big]_{v^i}\Big)_{|_{v^i = u^i}}, v^i - u^i\Big\rangle_{U_i^*, U_i}.\end{aligned}$$

Then we make use of the assumptions (C3) and (C4).

The respective reverse directions are based on the fact that the statements $\widetilde{V}_\alpha^{p_k}(u) = 0$ and $\widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0$ imply that

$$\begin{aligned}\theta_i(u) &\leq \theta_i(w^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(w^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k^i p(\|w^i - u^i\|_{U_i}^2) \qquad \forall\, w^i \in F_i(u^{-i}), \\ \theta_i(u) &\leq \theta_i(w^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(w^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + q_i(w^i - u^i) \qquad \forall\, w^i \in \widehat{F}_i(u).\end{aligned}$$

Applying Taylor's expansion at this step and taking the limit $t \downarrow 0$, we arrive at the desired inequality $\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \geq 0$ for all $v^i \in F_i(u^{-i})$ or $v^i \in \widehat{F}_i(u)$, respectively, and all $i \in [N]$.

Proof of (iii). Let $i \in [N]$ be fixed but arbitrary. In the special case of $F_i(u^{-i})$ containing only a single element, it has to hold $u^i \in F_i(u^{-i})$. Hence, it follows $F_i(u^{-i}) = \widetilde{F}_i(u) = \widehat{F}_i(u)$ and the equivalences of (iii) hold trivially.

Generally, we assume that the first VI of (iii) holds for all $v^i \in F_i(u^{-i})$. By the relations $\widetilde{F}_i(u) \subseteq F_i(u^{-i})$ and $\widehat{F}_i(u) \subseteq F_i(u^{-i})$, we directly obtain the desired result

$$\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \geq 0$$

for all $v^i \in \widetilde{F}_i(u)$ and for all $v^i \in \widehat{F}_i(u)$.

Let one of the other two VIs hold, i.e.,

$$\begin{aligned} \langle (\theta_i)_{u^i}(u), \tilde{v}^i - u^i \rangle_{U_i^*,U_i} &\geq 0 \qquad \forall\, \tilde{v}^i \in \widetilde{F}_i(u), \\ \langle (\theta_i)_{u^i}(u), \hat{v}^i - u^i \rangle_{U_i^*,U_i} &\geq 0 \qquad \forall\, \hat{v}^i \in \widehat{F}_i(u). \end{aligned}$$

We consider an arbitrary element $v^i \in F_i(u^{-i})$. If it holds $v^i \in \widetilde{F}_i(u)$ or $v^i \in \widehat{F}_i(u)$ in their respective cases, we can already finish the proof. In the matter of $v^i \in F_i(u^{-i}) \backslash \widetilde{F}_i(u)$ or $v^i \in F_i(u^{-i}) \backslash \widehat{F}_i(u)$, we can construct a suitable element with $\tilde{v}^i \in \widetilde{F}_i(u)$ or $\hat{v}^i \in \widehat{F}_i(u)$, respectively. Indeed, let the auxiliary parameter $t \in (0,1)$ satisfy $t\|v^i - u^i\|_{U_i} \leq R$ and set $\tilde{v}^i = u^i + t(v^i - u^i)$ or $\hat{v}^i = u^i + t(v^i - u^i)$, respectively. Exploiting the convexity of $F_i(u^{-i})$, we obtain $\tilde{v}^i \in F_i(u^{-i}) \cap \overline{B}_R(u^i) = \widetilde{F}_i(u)$ and $\hat{v}^i \in F_i(u^{-i}) \cap \overline{B}_R(u^i) \subseteq \widehat{F}_i(u)$, respectively. This concludes the proof of (iii) by observing

$$\begin{aligned} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} &= \frac{1}{t} \langle (\theta_i)_{u^i}(u), \tilde{v}^i - u^i \rangle_{U_i^*,U_i} \geq 0, \\ \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} &= \frac{1}{t} \langle (\theta_i)_{u^i}(u), \hat{v}^i - u^i \rangle_{U_i^*,U_i} \geq 0. \end{aligned}$$

Proof of (iv). We begin the proof of the equivalency statement by showing the direction "$\Rightarrow$". Let u be a root of $\widetilde{V}_\alpha^{\text{loc}}$. By applying (ii) of this theorem, we know that it has to hold

$$\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in \widetilde{F}_i(u),\ i \in [N].$$

Here, $i \in [N]$ is chosen arbitrarily. Using (iii), we obtain

$$\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}).$$

Since the functional $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2$ is pseudoconvex at u^i on $F_i(u^{-i})$, it holds by definition, see Definition 2.22, that

$$\theta_i(u) \leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \qquad \forall\, v^i \in F_i(u^{-i}).$$

Thus, we arrive at

$$0 \leq \widetilde{V}_\alpha(u) = \sum_{i\in[N]} \sup_{v^i\in F_i(u^{-i})} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] \leq 0.$$

The backwards direction "⇐" is an immediate consequence of both of the results $\widetilde{V}_\alpha(u) = 0$ and $0 \leq \widetilde{V}_\alpha^{\mathrm{loc}}(u) \leq \widetilde{V}_\alpha(u)$, see also (i).

Proof of (v). First, we take a look at the direction "⇒" and assume that it holds $\widetilde{V}_\alpha^{p_k}(u) = 0$. We obtain the variational inequality

$$\left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^i),\ i \in [N]$$

by applying the statements (ii) and (iii). Let $i \in [N]$ be arbitrarily fixed. By the assumed pseudoconvexity of the functional $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2$ at u^i on $F_i(u^{-i})$, we obtain the inequality

$$\theta_i(u) \leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2$$

for all $v^i \in F_i(u^{-i})$ and thus, it holds $\widetilde{V}_\alpha(u) = 0$.

The reverse direction "⇐" follows immediately by $\widetilde{V}_\alpha(u) = 0$ and $0 \leq \widetilde{V}_\alpha^{p_k}(u) \leq \widetilde{V}_\alpha(u)$ for a penalty-type function $p : [0, \infty) \to [0, \infty]$.

Proof of (vi). This proof follows the lines to the proof of the preceding statement (v). In the case of the forward direction "⇒", we assume that it holds $\widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0$. We apply (ii) and (iii) to achieve the variational inequality

$$\left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}),\ i \in [N].$$

Let $i \in [N]$ again be arbitrary. The pseudoconvexity of $\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2$ at u^i on $F_i(u^{-i})$ yields

$$\theta_i(u) \leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \qquad \forall\, v^i \in F_i(u^{-i}).$$

Hence, we arrive at

$$0 \leq \widetilde{V}_\alpha(u) = \sum_{i\in[N]} \sup_{v^i\in F_i(u^{-i})} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] \leq 0.$$

Lastly, the direction "⇐" is an immediate consequence of $\widetilde{V}_\alpha(u) = 0$ and the inequality $0 \leq \widetilde{V}_\alpha^{\mathrm{lp}}(u) \leq \widetilde{V}_\alpha(u)$.

Proof of (vii). We begin with the direction "⇒" and assume that it holds $\widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0$. By the already proved statements (ii) and (iii), we obtain the variational inequality

$$\left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^i),\ i \in [N].$$

Let $i \in [N]$ be arbitrarily fixed. By the assumed pseudoconvexity of the functional

$$\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k^i p(\| \cdot - u^i\|_{U_i}^2)$$

at u^i on $F_i(u^{-i})$, we obtain the inequality

$$\theta_i(u) \leq \theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k^i p(\|v^i - u^i\|_{U_i}^2)$$

for all $v^i \in F_i(u^{-i})$ and thus, it holds $\widetilde{V}_\alpha^{p_k}(u) = 0$.

The reverse direction "$\Leftarrow$" follows immediately by $\widetilde{V}_\alpha^{p_k}(u) = 0$ and $0 \leq \widetilde{V}_\alpha^{\text{loc}}(u) \leq \widetilde{V}_\alpha^{p_k}(u)$.

□

We cannot determine much about the relationship between $\widetilde{V}_\alpha^{\text{loc}}(u)$, $\widetilde{V}_\alpha^{p_k}(u)$, and $\widetilde{V}_\alpha^{\text{lp}}(u)$. The merit functional $\widetilde{V}_\alpha^{\text{lp}}(u)$ combines the localization of the admissible set and the addition of a penalty-type or barrier-type term. Indeed, the admissible sets have the link $\widetilde{F}_i(u) \subseteq \widehat{F}_i(u) \subseteq F_i(u^{-i})$ with one another. However, the penalty-type terms p and q_i restrict us from conceiving any connection between the corresponding merit functionals. In the special case of $\widetilde{F}_i(u) = \widehat{F}_i(u)$, we obtain $\widetilde{V}_\alpha^{\text{loc}}(u) \geq \widetilde{V}_\alpha^{\text{lp}}(u)$.

The parts (i)–(iii) of Theorem 5.1 tell us that $\widetilde{V}_\alpha^{\text{loc}}$ and $\widetilde{V}_\alpha^{p_k}$ can be utilized as merit functionals to the their corresponding systems of QVIs

$$u^i \in F_i(u^{-i}), \qquad \langle(\theta_i)_{u^i}(u), v^i - u^i\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}) \qquad (1 \leq i \leq N), \tag{5.10}$$

$$u^i \in \widetilde{F}_i(u), \qquad \langle(\theta_i)_{u^i}(u), v^i - u^i\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in \widetilde{F}_i(u) \qquad (1 \leq i \leq N). \tag{5.11}$$

if the admissible set $F(u)$ is convex, (B1) of Assumption 3.3 holds, and either (C2) or (C3) of Assumption 3.4 holds. If (C4) of Assumption 3.4 is valid, then $\widetilde{V}_\alpha^{\text{lp}}$ can be served as a merit functional for either of the two subsequent systems of QVIs

$$u^i \in F_i(u^{-i}), \qquad \langle(\theta_i)_{u^i}(u), v^i - u^i\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}) \qquad (1 \leq i \leq N), \tag{5.12}$$

$$u^i \in \widehat{F}_i(u), \qquad \langle(\theta_i)_{u^i}(u), v^i - u^i\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v^i \in \widehat{F}_i(u) \qquad (1 \leq i \leq N). \tag{5.13}$$

Furthermore, if we replace (C2), (C3) and (C4) by the stronger assumption (C1), then $\widetilde{V}_\alpha$ can be utilized as a merit functional for both (5.11) and (5.12). Next, we study the relationship between a (local) Nash equilibrium and the Nikaido–Isoda merit functionals $\widetilde{V}_\alpha$, $\widetilde{V}_\alpha^{\text{loc}}$, $\widetilde{V}_\alpha^{p_k}$, and $\widetilde{V}_\alpha^{\text{lp}}$.

Theorem 5.2. Let $u \in U$ satisfy $u \in F(u)$ and let $F_i(u^{-i})$ be a convex set for all $i \in [N]$. Furthermore, we assume that (B1) holds. Then we conclude the subsequent statements.

(i) If u is a local Nash equilibrium on $\widetilde{F}(u)$ in the sense that it fulfills $\theta_i(u) \leq \theta_i(v^i, u^{-i})$ for all $v^i \in \widetilde{F}_i(u)$ and all $i \in [N]$, then it holds $\widetilde{V}_\alpha^{\text{loc}}(u) = 0$.

If u is a local Nash equilibrium on $\widehat{F}(u)$ in the sense that it fulfills $\theta_i(u) \leq \theta_i(v^i, u^{-i})$ for all $v^i \in \widehat{F}_i(u)$ and all $i \in [N]$, then it holds $\widetilde{V}_\alpha^{\text{lp}}(u) = 0$.

(ii) If $\theta_i(\cdot, u^{-i})$ is pseudoconvex at u^i on $\widetilde{F}_i(u)$ for all $i \in [N]$ and if $\widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0$ holds, then u is a local Nash equilibrium on $\widetilde{F}(u)$.

If $\theta_i(\cdot, u^{-i})$ is pseudoconvex at u^i on $\widehat{F}_i(u)$ for all $i \in [N]$ and if $\widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0$ holds, then u is a local Nash equilibrium on $\widehat{F}(u)$.

(iii) If u is a Nash equilibrium, then it holds

$$\widetilde{V}_\alpha(u) = \widetilde{V}_\alpha^{\mathrm{loc}}(u) = \widetilde{V}_\alpha^{p_k}(u) = \widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0.$$

(iv) If $\theta_i(\cdot, u^{-i})$ is pseudoconvex at u^i on $F_i(u^{-i})$ for all $i \in [N]$, then for arbitrarily fixed $k \in \mathbb{N}$ the following equivalences hold true:

$$u \text{ Nash equilibrium} \iff \widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0 \iff \widetilde{V}_\alpha(u) = 0 \iff \widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0 \iff \widetilde{V}_\alpha^{p_k}(u) = 0.$$

Proof. We first note that the regularized functionals

$$\begin{aligned}
&\frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2 + \rho_k p(\| \cdot - u^i\|_{U_i}^2),\\
&\frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2 + q_i(\cdot - u^i),
\end{aligned}$$

are nonnegative and vanish at u^i. Moreover, their derivatives vanish again at u^i and the presumed pseudoconvexity of $\theta_i(\cdot, u^{-i})$ at u^i on $F_i(u^{-i})$ (or $\widehat{F}_i(u)$) implies (C3) and (C4) of Assumption 3.4.

Proof of (i). By $\alpha \geq 0$, it holds

$$\begin{aligned}
\widetilde{V}_\alpha^{\mathrm{loc}}(u) &= \sum_{i\in[N]} \sup_{v^i \in \widetilde{F}_i(u)} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right]\\
&\leq \sum_{i\in[N]} \sup_{v^i \in \widetilde{F}_i(u)} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right].
\end{aligned}$$

Since u is a local Nash equilibrium on $\widetilde{F}_i(u)$, we obtain $\widetilde{V}_\alpha^{\mathrm{loc}}(u) \leq 0$. Furthermore, we get

$$\begin{aligned}
\widetilde{V}_\alpha^{\mathrm{lp}}(u) &= \sum_{i\in[N]} \sup_{v^i \in \widehat{F}_i(u)} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i)\right]\\
&\leq \sum_{i\in[N]} \sup_{v^i \in \widehat{F}_i(u)} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right]
\end{aligned}$$

by the positivity of the parameter α and the functionals q_i, $i \in [N]$. If u is assumed to be a local Nash equilibrium on $\widehat{F}_i(u)$, then it yields $\widetilde{V}_\alpha^{\mathrm{lp}}(u) \leq 0$.

Proof of (ii). Let either $\widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0$ or $\widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0$ be valid. By (ii) of Theorem 5.1, we obtain the variational inequality

$$\langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, v^i \in \widetilde{F}_i(u) \text{ or } v^i \in \widehat{F}_i(u),$$

respectively. Here, the index $i \in [N]$ is arbitrary. By the pseudoconvexity of $\theta_i(\cdot, u^{-i})$ at u^i on $\widetilde{F}_i(u)$ or $\widehat{F}_i(u)$, we conclude that it holds $\theta_i(u) \leq \theta_i(v^i, u^{-i})$ for all $v^i \in \widetilde{F}_i(u)$ or $v^i \in \widehat{F}_i(u)$, respectively.

Proof of (iii). Since u is a Nash equilibrium, we can prove the fact $\widetilde{V}_\alpha(u) = 0$ similar to (i) by replacing $\widetilde{F}_i(u)$ with $F_i(u^{-i})$. The rest of the proof follows the lines from the inequalities $0 \leq \widetilde{V}_\alpha^{\text{loc}}(u) \leq \widetilde{V}_\alpha(u)$, $0 \leq \widetilde{V}_\alpha^{p_k}(u) \leq \widetilde{V}_\alpha(u)$ and $0 \leq \widetilde{V}_\alpha^{\text{lp}}(u) \leq \widetilde{V}_\alpha(u)$.

Proof of (iv). Since u is a Nash equilibrium, it holds $\widetilde{V}_\alpha^{\text{lp}}(u) = \widetilde{V}_\alpha^{p_k}(u) = \widetilde{V}_\alpha^{\text{loc}}(u) = 0 = \widetilde{V}_\alpha(u)$ by (iii). By (ii) Theorem 5.1, the first-order optimality conditions (2.8) are satisfied. Finally, we conclude that u is a Nash equilibrium by the pseudoconvexity of $\theta_i(\cdot, u^{-i})$ at u^i on $F_i(u^{-i})$. □

We remark that the first assertion (i) is not valid in case of the regularized and localized merit functional $\widetilde{V}_\alpha^{p_k}(u)$. Indeed, since p is nonnegative, we can only achieve

$$\begin{aligned}
\widetilde{V}_\alpha^{p_k}(u) &= \sum_{i\in[N]} \sup_{v^i \in F_i(u^{-i})} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(\|v^i - u^i\|_{U_i}^2)\right] \\
&\leq \sum_{i\in[N]} \sup_{v^i \in F_i(u^{-i})} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right].
\end{aligned}$$

In addition, for the premise of a local Nash equilibrium to hold true, we would need the supremum over $\widetilde{F}_i(u)$. However, this is not possible in the case of penalty-type functions since it would imply that $\widetilde{V}_\alpha^{p_k}$ coincides with $\widetilde{V}_\alpha^{\text{loc}}$.

Since it holds the estimate

$$\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \leq \theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2,$$

for all $\beta \geq \alpha \geq 0$ and all $v^i \in F_i(u^{-i})$, we obtain the following relations between the merit functionals

$$\begin{array}{ccccc}
\widetilde{V}_\beta(u) & \leq & \widetilde{V}_\alpha(u) & \leq & \widetilde{V}_0(u), \\
\widetilde{V}_\beta^{\text{loc}}(u) & \leq & \widetilde{V}_\alpha^{\text{loc}}(u) & \leq & \widetilde{V}_0^{\text{loc}}(u), \\
\widetilde{V}_\beta^{p_k}(u) & \leq & \widetilde{V}_\alpha^{p_k}(u) & \leq & \widetilde{V}_0^{p_k}(u), \\
\widetilde{V}_\beta^{\text{lp}}(u) & \leq & \widetilde{V}_\alpha^{\text{lp}}(u) & \leq & \widetilde{V}_0^{\text{lp}}(u),
\end{array}$$

for all $u \in F(u)$ and $k \in \mathbb{N}$ fixed. Consequently, for any $u \in U$ with $u \in F(u)$ and $F(u)$ convex we conclude the following implications

$$\begin{array}{lclcl}
u \text{ local Nash equilibrium} & \Longrightarrow & \widetilde{V}_0^{\text{loc}}(u) = 0 & \Longrightarrow & \widetilde{V}_\alpha^{\text{loc}}(u) = 0, \\
u \text{ local Nash equilibrium} & \Longrightarrow & \widetilde{V}_0^{\text{lp}}(u) = 0 & \Longrightarrow & \widetilde{V}_\alpha^{\text{lp}}(u) = 0, \\
u \text{ Nash equilibrium} & \Longrightarrow & \widetilde{V}_0^{p_k}(u) = 0 & \Longrightarrow & \widetilde{V}_\alpha^{p_k}(u) = 0, \\
u \text{ Nash equilibrium} & \Longrightarrow & \widetilde{V}_0(u) = 0 & \Longrightarrow & \widetilde{V}_\alpha(u) = 0,
\end{array}$$

for all $\alpha \geq 0$. The reverse implications would require pseudoconvexity assumptions, cf. Theorem 5.2 above.

Altogether, if it holds $u \in U$ with $u \in F(u)$ and $\theta_i(\cdot, u^{-i})$ is pseudoconvex at u^i on $\widetilde{F}_i(u)$ or $\widehat{F}_i(u)$ for all $i \in [N]$, then we have proved that for any $\alpha \geq 0$ it holds

$$\begin{aligned}
\widetilde{V}_0^{\mathrm{loc}}(u) = 0 \iff \widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0 &\iff \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, v^i \in \widetilde{F}_i(u), \quad i \in [N],\\
&\iff \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}), \quad i \in [N],\\
&\iff u \text{ local Nash equilibrium},\\
\widetilde{V}_0^{\mathrm{lp}}(u) = 0 \iff \widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0 &\iff \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, v^i \in \widehat{F}_i(u), \quad i \in [N],\\
&\iff \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, v^i \in F_i(u^{-i}), \quad i \in [N],\\
&\iff u \text{ local Nash equilibrium}.
\end{aligned}$$

Moreover, if $\theta_i(\cdot, u^{-i})$ is pseudoconvex at u^i on $F_i(u^{-i})$ for all $i \in [N]$, we obtain for any $\alpha \geq 0$

$$\begin{aligned}
\widetilde{V}_0^{\mathrm{lp}}(u) = 0 \iff \widetilde{V}_\alpha^{\mathrm{lp}}(u) = 0 &\iff \widetilde{V}_0^{\mathrm{loc}}(u) = 0 \iff \widetilde{V}_\alpha^{\mathrm{loc}}(u) = 0,\\
&\iff \widetilde{V}_0(u) \quad = 0 \iff \widetilde{V}_\alpha(u) \quad = 0,\\
&\iff \widetilde{V}_0^{p_k}(u) \ = 0 \iff \widetilde{V}_\alpha^{p_k}(u) \ = 0,\\
&\iff u \text{ Nash equilibrium}.
\end{aligned}$$

5.2.2 Normalized Equilibria

Now, we investigate the relationship between the various regularized and localized Nikaido–Isoda merit functionals and normalized equilibria. We study the functionals $V_\alpha : U \to \mathbb{R}$, $V_\alpha^{\mathrm{loc}} : U \to \mathbb{R}$, $V_\alpha^{p_k} : U \to \mathbb{R}$ and $V_\alpha^{\mathrm{lp}} : U \to \mathbb{R}$ presented in (5.2), (5.4), (5.6), and (5.8). These functionals are demonstrated to be Nikaido–Isoda merit functionals that correspond to a (Q)VI. Under certain pseudoconvexity assumptions, we arrive at a link between the Nikaido–Isoda merit functionals and (local) normalized equilibria. Hence, we only consider the feasible sets

$$F_i(u^{-i}) = \{v^i \in U_i : (v^i, u^{-i}) \in \mathcal{X}\},$$

such that $\mathcal{X} \subseteq U$ is nonempty and convex.

Theorem 5.3. We assume $\alpha \geq 0$. Moreover, let (B1) be satisfied and $\mathcal{X}$ be convex in case of the assertions (ii)–(vi). Regarding "⇒" in (ii), we additionally assume (C7), (C8) and (C9) of Assumption 3.4, respectively. Addressing the implication "⇒" in (iv), (v), and (vi), we additionally assume (C6) of Assumption 3.4 hold. In the case of "⇒" in (vii), let (C8) of Assumption 3.4 hold. For the reverse direction "⇐" in (vii), we assume that $p : [0, \infty) \to [0, \infty]$ is a penalty-type function. Then for all $u \in \mathcal{X}$ it holds the following statements.

(i) $V_\alpha(u) \geq V_\alpha^{\mathrm{loc}}(u) \geq 0$,

$V_\alpha(u) \geq V_\alpha^{p_k}(u) \geq 0,$

$V_\alpha(u) \geq V_\alpha^{\mathrm{lp}}(u) \geq 0.$

Moreover, it holds $V_\alpha(u) \geq V_\alpha^{p_k}(u) \geq V_\alpha^{\mathrm{loc}}(u) \geq 0$ in the case of a penalty-type function p.

(ii)
$$\begin{aligned}
&\textstyle\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \quad \forall\, v \in \widetilde{\mathcal{X}}(u) &&\iff && V_\alpha^{\mathrm{loc}}(u) = 0,\\
&\textstyle\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \quad \forall\, v \in \mathcal{X} &&\iff && V_\alpha^{p_k}(u) = 0,\\
&\textstyle\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \quad \forall\, v \in \widehat{\mathcal{X}}(u) &&\iff && V_\alpha^{\mathrm{lp}}(u) = 0.
\end{aligned}$$

(iii)
$$\begin{aligned}
& &&\textstyle\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \quad \forall\, v \in \mathcal{X}\\
&\iff &&\textstyle\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \quad \forall\, v \in \widetilde{\mathcal{X}}(u)\\
&\iff &&\textstyle\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \quad \forall\, v \in \widehat{\mathcal{X}}(u)
\end{aligned}$$

(iv) $V_\alpha^{\mathrm{loc}}(u) = 0 \iff V_\alpha(u) = 0.$

(v) $V_\alpha^{p_k}(u) = 0 \iff V_\alpha(u) = 0.$

(vi) $V_\alpha^{\mathrm{lp}}(u) = 0 \iff V_\alpha(u) = 0.$

(vii) $V_\alpha^{\mathrm{loc}}(u) = 0 \iff V_\alpha^{p_k}(u) = 0.$

Proof. Let $u \in \mathcal{X}$ be arbitrary. We will prove the statement individually, beginning with the first claim.

Proof of (i). By the inclusions $\widetilde{\mathcal{X}}(u) \subseteq \mathcal{X}$ and $\widehat{\mathcal{X}}(u) \subseteq \mathcal{X}$, and the nonnegativity of the penalty-type or barrier-type functions p and q, we obtain the first inequalities of the respective statements. Setting the specific element $v = u$ instead of taking the supremum, we obtain the following relationship between the merit functionals

$$\begin{aligned}
V_\alpha(u) \geq V_\alpha^{\mathrm{loc}}(u) &\geq \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(u^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(u^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] &&= 0,\\
V_\alpha(u) \geq V_\alpha^{p_k}(u) &\geq \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(u^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(u^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] - \rho_k p(\|u-u\|_U^2) &&= 0,\\
V_\alpha(u) \geq V_\alpha^{\mathrm{lp}}(u) &\geq \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(u^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(u^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] - q(u-u) &&= 0.
\end{aligned}$$

Lastly, in the case of a penalty-type function it holds by definition $p(\|v-u\|_U^2) = 0$ for all $v \in \overline{B}_R(u)$ and thus, we obtain the inequality

$$\begin{aligned}
&\sup_{v\in\mathcal{X}} \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right]\\
&\qquad \geq \sup_{v\in\mathcal{X}} \Big[\sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] - \rho_k p(\|v-u\|_U^2)\Big].
\end{aligned}$$

Cutting the feasible set with the set $\{v \in U : p(\|v-u\|_U^2) = 0\}$ we get

$$\begin{aligned}
&\sup_{v\in\mathcal{X}} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] - \rho_k p(\|v-u\|_U^2)\Big] \\
&\qquad\geq \sup_{v\in\mathcal{X}\cap\{v: p(\|v-u\|_U^2)=0\}} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] \\
&\qquad\qquad - \rho_k p(\|v-u\|_U^2)\Big] \\
&\qquad= \sup_{v\in\mathcal{X}\cap\overline{B}_R(u)} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big].
\end{aligned}$$

We arrive at the inequality

$$\begin{aligned}
&\sup_{v\in\mathcal{X}} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] \\
&\qquad\geq \sup_{v\in\mathcal{X}} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] - \rho_k p(\|v-u\|_U^2)\Big] \\
&\qquad\geq \sup_{v\in\widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big].
\end{aligned}$$

We can further estimate from below by setting the specific value $v = u \in \widetilde{\mathcal{X}}(u)$ instead of using the supremum. Thus, we obtain the desired result

$$V_\alpha(u) \geq V_\alpha^{p_k}(u) \geq V_\alpha^{\text{loc}}(u) \geq 0.$$

Proof of (ii). First, we consider the forward implication "⇒" in the equivalency statement. By the definitions of $\widetilde{\Psi}_\alpha$ and the dual product of the product space U, it holds

$$\Big\langle \big[(\widetilde{\Psi}_\alpha(u,v))_v\big]_{|v=u}, v-u \Big\rangle_{U^*,U} = \sum_{i\in[N]} \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*,U_i}.$$

Moreover, we make use of the pseudoconvexity of $\widetilde{\Psi}_\alpha(u,\cdot)$ at u on $\widetilde{\mathcal{X}}(u)$ and of the QVI that holds by assumption. Altogether, this yields

$$\sum_{i\in[N]} \theta_i(u) = \widetilde{\Psi}_\alpha(u,u) \leq \widetilde{\Psi}_\alpha(u,v) \qquad \forall\, v \in \widetilde{\mathcal{X}}(u),$$

and we obtain $V_\alpha^{\text{loc}}(u) = 0$.

In the setting of the backward implication "⇐", we have $V_\alpha^{\text{loc}}(u) = 0$. This leads to the conclusion that it holds

$$\sum_{i\in[N]} \theta_i(u) \leq \widetilde{\Psi}_\alpha(u,w) \qquad \forall\, w \in \widetilde{\mathcal{X}}(u).$$

Since the set $\widetilde{\mathcal{X}}(u)$ is assumed to be convex, we are allowed to consider the convex combination $w = u + t(v-u) \in \widetilde{\mathcal{X}}(u)$ for $u, v \in \widetilde{\mathcal{X}}(u)$ and $t \in (0,1]$. We apply Taylor's expansion of $\widetilde{\Psi}_\alpha(u,\cdot)$ around w and we obtain

$$0 \leq \widetilde{\Psi}_\alpha(u, u + t(v-u)) - \sum_{i\in[N]} \theta_i(u) = \sum_{i\in[N]} \left[\left\langle (\theta_i)_{u^i}(u), t(v^i - u^i)\right\rangle_{U_i^*,U_i}\right] + \mathrm{o}(t).$$

We divide this inequality by t and take the limit $t \downarrow 0$, which directly yields the statement's desired inequality.

The proof of the other equivalences in (ii) follows the lines above. The direct implication is based on the corresponding pseudoconvexity assumptions at the point u, (C8) and (C9), and the fact that we can write

$$\begin{aligned}\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i\right\rangle_{U_i^*,U_i} &= \left\langle \left[\left(\widetilde{\Psi}_\alpha(u,v) + \rho_k p(\|v-u\|_U^2)\right)_v\right]_{|v=u}, v-u\right\rangle_{U^*,U} \\ &= \left\langle \left[\left(\widetilde{\Psi}_\alpha(u,v) + q(v-u)\right)_v\right]_{|v=u}, v-u\right\rangle_{U^*,U}.\end{aligned}$$

For the reverse implication we exploit that $V_\alpha^{p_k}(u) = 0$ and $V_\alpha^{\mathrm{lp}}(u) = 0$ imply the following

$$\begin{aligned}\sum_{i\in[N]} \theta_i(u) &\leq \widetilde{\Psi}_\alpha(u,w) + \rho_k p(\|w-u\|_U^2) \qquad &&\forall\, w \in \mathcal{X}, \\ \sum_{i\in[N]} \theta_i(u) &\leq \widetilde{\Psi}_\alpha(u,w) + q(w-u) &&\forall\, w \in \widehat{\mathcal{X}}(u).\end{aligned}$$

After applying Taylor's expansion, dividing by t, and taking the limit $t \downarrow 0$, it yields the desired inequalities.

Proof of (iii). If $\mathcal{X}$ contains only one point, then it follows $\mathcal{X} = \{u\} = \widetilde{\mathcal{X}}(u) = \widehat{\mathcal{X}}(u)$ and the claim holds trivially.

In the general case, we assume that the VI holds for all $v \in \mathcal{X}$. By the known subset relationships $\widetilde{\mathcal{X}}(u) \subseteq \mathcal{X}$ and $\widehat{\mathcal{X}}(u) \subseteq \mathcal{X}$, we directly obtain

$$\sum_{i\in[N]} \left\langle (\theta_i)_{v^i}(u), v^i - u^i\right\rangle_{U_i^*,U_i} \geq 0$$

for all $v \in \widetilde{\mathcal{X}}(u)$ and for all $v \in \widehat{\mathcal{X}}(u)$, respectively.

Next, we assume that one of the following VIs holds:

$$\begin{aligned}\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), \tilde{v}^i - u^i\right\rangle_{U_i^*,U_i} &\geq 0 \qquad &&\forall\, \tilde{v} \in \widetilde{\mathcal{X}}(u), \\ \sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), \hat{v}^i - u^i\right\rangle_{U_i^*,U_i} &\geq 0 &&\forall\, \hat{v} \in \widehat{\mathcal{X}}(u).\end{aligned}$$

We consider any $v \in \mathcal{X}$ and we note that in the cases with either $v \in \widetilde{\mathcal{X}}(u)$ or $v \in \widehat{\mathcal{X}}(u)$, the first inequality clearly holds. Consequently, we consider an element v with either $v \in \mathcal{X}\backslash\widetilde{\mathcal{X}}(u)$

or $v \in \mathcal{X} \setminus \widehat{\mathcal{X}}(u)$. Furthermore, we choose a parameter $t \in (0,1]$ with $t\|v-u\|_U \leq R$ and by convexity of $\mathcal{X}$, it follows $\tilde{v} = u + t(v-u) \in \widetilde{\mathcal{X}}(u)$ and $\hat{v} = u + t(v-u) \in \widehat{\mathcal{X}}(u)$. Thus, we obtain the desired VIs, i.e.,

$$\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} = \frac{1}{t} \sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), \tilde{v}^i - u^i \right\rangle_{U_i^*,U_i} \geq 0,$$
$$\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} = \frac{1}{t} \sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), \hat{v}^i - u^i \right\rangle_{U_i^*,U_i} \geq 0.$$

Proof of (iv). First, we consider the direction "⇒" and consequently, assume that it holds $V_\alpha^{\mathrm{loc}}(u) = 0$. By (ii) and (iii) of this theorem, we obtain

$$\left\langle \left[(\widetilde{\Psi}_\alpha)_v(u,v)\right]_{|v=u}, v-u \right\rangle_{U^*,U} = \sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v \in \mathcal{X}.$$

Moreover, the pseudoconvexity of $\widetilde{\Psi}_\alpha(u,\cdot)$ at u on $\mathcal{X}$ yields

$$\sum_{i\in[N]} \theta_i(u) = \widetilde{\Psi}_\alpha(u,u) \leq \widetilde{\Psi}_\alpha(u,v) \qquad \forall\, v \in \mathcal{X},$$

which implies the desired result $V_\alpha(u) = 0$.

The direction "⇐" trivially follows from $0 \leq V_\alpha^{\mathrm{loc}}(u) \leq V_\alpha(u)$.

Proof of (v). We begin with the implication "⇒" and assume that it holds $V_\alpha^{p_k}(u) = 0$. Then we obtain the inequality

$$\sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v \in \mathcal{X}$$

by the statements (ii) and (iii) of this theorem. Furthermore, we apply the pseudoconvexity of $\widetilde{\Psi}_\alpha(u,\cdot)$ at u on $\mathcal{X}$ and arrive at the result

$$\sum_{i\in[N]} \theta_i(u) \leq \widetilde{\Psi}_\alpha(u,v) \qquad \forall\, v \in \mathcal{X}.$$

Thus, we have $V_\alpha(u) = 0$.

The reverse direction "⇐" follows immediately by $0 \leq V_\alpha^{p_k}(u) \leq V_\alpha(u)$.

Proof of (vi). In the forward direction "⇒", we assume that it holds $V_\alpha^{\mathrm{lp}}(u) = 0$. Again, by (ii) and (iii) we can conclude

$$\left\langle \left[(\widetilde{\Psi}_\alpha)_v(u,v)\right]_{|v=u}, v-u \right\rangle_{U^*,U} = \sum_{i\in[N]} \left\langle (\theta_i)_{u^i}(u), v^i - u^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, v \in \mathcal{X}.$$

Furthermore, we exploit the pseudoconvexity of $\widetilde{\Psi}_\alpha(u,\cdot)$ at u on $\mathcal{X}$ and obtain

$$\sum_{i\in[N]} \theta_i(u) = \widetilde{\Psi}_\alpha(u,u) \leq \widetilde{\Psi}_\alpha(u,v) \qquad \forall\, v \in \mathcal{X},$$

which finally implies $V_\alpha(u) = 0$.

The implication "⇐" can be concluded as before by the inequality $0 \le V_\alpha^{\text{lp}}(u) \le V_\alpha(u)$.

Proof of (vii). We prove the implication "⇒" and assume that it holds $V_\alpha^{\text{loc}}(u) = 0$. Then we obtain by (ii) and (iii) of this theorem the inequality

$$\sum_{i\in[N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \ge 0 \qquad \forall\, v \in \mathcal{X}.$$

Due to the pseudoconvexity of $\widetilde{\Psi}_\alpha(u,\cdot) + \rho_k p(\|\cdot - u\|_U^2)$ at u on $\mathcal{X}$, we arrive at the result

$$\sum_{i\in[N]} \theta_i(u) \le \widetilde{\Psi}_\alpha(u,v) + \rho_k p(\|v-u\|_U^2) \qquad \forall\, v \in \mathcal{X}.$$

Thus, we have $V_\alpha^{p_k}(u) = 0$.

The reverse direction "⇐" follows immediately by $V_\alpha^{p_k}(u) = 0$ and $0 \le V_\alpha^{\text{loc}}(u) \le V_\alpha^{p_k}(u)$ in the case of a penalty-type function $p : [0,\infty) \to [0,\infty]$. □

Under the assumptions (B1) of Assumption 3.3, (C7) and (C8) of Assumption 3.4, respectively, and that $\mathcal{X}$ is convex, Theorem 5.3, (i)–(iii), yields that V_α^{loc} and $V_\alpha^{p_k}$, respectively, can be utilized as a merit functional for any of the two (Q)VIs

$$u \in \mathcal{X}, \qquad \sum_{i\in[N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \ge 0 \qquad \forall\, v \in \mathcal{X}, \tag{5.14}$$

$$u \in \widetilde{\mathcal{X}}(u), \qquad \sum_{i\in[N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \ge 0 \qquad \forall\, v \in \widetilde{\mathcal{X}}(u). \tag{5.15}$$

Moreover, if (C9) of Assumption Assumption 3.4 holds, then V_α^{lp} can be used as a merit functional for any of the two (Q)VIs

$$u \in \mathcal{X}, \qquad \sum_{i\in[N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \ge 0 \qquad \forall\, v \in \mathcal{X}, \tag{5.16}$$

$$u \in \widehat{\mathcal{X}}(u), \qquad \sum_{i\in[N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*,U_i} \ge 0 \qquad \forall\, v \in \widehat{\mathcal{X}}(u). \tag{5.17}$$

Furthermore, if we replace (C7), (C8) and (C9) by the stronger assumption (C6), then V_α can be used as a merit functional for (5.14) and (5.16).

In the following, we consider the connection between these merit functionals and a (local) normalized equilibrium.

Theorem 5.4. Let $\alpha \ge 0$ and $\mathcal{X} \subseteq U$ be convex, and assume that (B1) holds true. Moreover, we assume $u \in \mathcal{X}$.

(i) If u is a local normalized equilibrium on $\widetilde{\mathcal{X}}(u)$, i.e., if it holds

$$\sum_{i\in[N]} \theta_i(u) \le \sum_{i\in[N]} \theta_i(v^i, u^{-i}) \qquad \forall\, v \in \widetilde{\mathcal{X}}(u),$$

then it implies $V_\alpha^{\text{loc}}(u) = 0$. On the other hand, if u is a local normalized equilibrium on $\widehat{\mathcal{X}}(u)$, i.e., if it holds

$$\sum_{i\in[N]} \theta_i(u) \leq \sum_{i\in[N]} \theta_i(v^i, u^{-i}) \qquad \forall\, v \in \widehat{\mathcal{X}}(u),$$

then it implies $V_\alpha^{\text{lp}}(u) = 0$.

(ii) If $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$ is pseudoconvex at u on $\widetilde{\mathcal{X}}(u)$ and if $V_\alpha^{\text{loc}}(u) = 0$ holds, then u is a local normalized equilibrium on $\widetilde{\mathcal{X}}(u)$.

If $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$ is pseudoconvex at u on $\widehat{\mathcal{X}}(u)$ and if $V_\alpha^{\text{lp}}(u) = 0$ holds, then u is a local normalized equilibrium on $\widehat{\mathcal{X}}(u)$.

(iii) If u is a normalized equilibrium, then it holds

$$V_\alpha(u) = V_\alpha^{\text{loc}}(u) = V_\alpha^{p_k}(u) = V_\alpha^{\text{lp}}(u) = 0.$$

(iv) If $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$ is pseudoconvex at u on $\mathcal{X}$, then for arbitrarily fixed $k \in \mathbb{N}$ the following equivalences hold true:

$$u \text{ normalized eq.} \iff V_\alpha^{\text{lp}}(u) = 0 \iff V_\alpha(u) = 0 \iff V_\alpha^{\text{loc}}(u) = 0 \iff V_\alpha^{p_k}(u) = 0.$$

Proof. We begin the proof by showing the first statement.

Proof of (i). Let u be a local normalized equilibrium on $\widetilde{\mathcal{X}}(u)$. By definition of a local normalized equilibrium, it holds $\sum_{i\in[N]} \theta_i(u) \leq \sum_{i\in[N]} \theta_i(v^i, u^{-i})$ for all $v \in \widetilde{\mathcal{X}}(u)$ and thus,

$$\sup_{v\in\widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right] \leq 0.$$

Since the parameter $\alpha \geq 0$ is nonnegative we get

$$\begin{aligned} V_\alpha^{\text{loc}}(u) &= \sup_{v\in\widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right] \\ &\leq \sup_{v\in\widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right]. \end{aligned}$$

All together, we obtain $0 \leq V_\alpha^{\text{loc}}(u) \leq 0$ by (i) of Theorem 5.3.

In the case of a local normalized equilibrium on $\widehat{\mathcal{X}}(u)$, we proceed similarly. If u is a local normalized equilibrium on $\widehat{\mathcal{X}}(u)$, it yields

$$\sup_{v\in\widehat{\mathcal{X}}(u)} \sum_{i\in[N]} \left[\theta_i(u) - \theta_i(v^i, u^{-i})\right] \leq 0.$$

Once again by the nonnegativity of $\alpha \geq 0$ and using that $q : U \to \mathbb{R}$ is nonnegative, we obtain the estimate

$$\begin{aligned} V_\alpha^{\mathrm{lp}}(u) &= \sup_{v \in \widehat{\mathcal{X}}(u)} \Big[\sum_{i \in [N]} [\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2} \| \iota_{H_i}(v^i) - \iota_{H_i}(u^i) \|_{H_i}^2] - q(v-u) \Big] \\ &\leq \sup_{v \in \widehat{\mathcal{X}}(u)} \sum_{i \in [N]} [\theta_i(u) - \theta_i(v^i, u^{-i})]. \end{aligned}$$

Finally, we arrive at $0 \leq V_\alpha^{\mathrm{lp}}(u) \leq 0$.

Proof of (ii). Let $V_\alpha^{\mathrm{loc}}(u) = 0$ or $V_\alpha^{\mathrm{lp}}(u) = 0$ be valid, respectively. By Theorem 5.3, (ii), we have

$$\sum_{i \in [N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \tag{5.18}$$

for all $v \in \widetilde{\mathcal{X}}(u)$ or $v \in \widehat{\mathcal{X}}(u)$, respectively. Since $\sum_{i \in [N]} \theta_i(\cdot, u^{-i}) : U \to \mathbb{R}$ is pseudoconvex at the point u on the sets $\widetilde{\mathcal{X}}(u)$ or $\widehat{\mathcal{X}}(u)$, using (5.18) we get

$$\sum_{i \in [N]} \theta_i(u) \leq \sum_{i \in [N]} \theta_i(v^i, u^{-i})$$

for all $v \in \widetilde{\mathcal{X}}(u)$ or $v \in \widehat{\mathcal{X}}(u)$, respectively. Consequently, $u \in \mathcal{X}$ is a local normalized equilibrium on $\widetilde{\mathcal{X}}(u)$ or $\widehat{\mathcal{X}}(u)$.

Proof of (iii). The assertion $V_\alpha(u) = 0$ is proved like in (i), replacing $\widetilde{\mathcal{X}}(u)$ by $\mathcal{X}$. The rest follows from the facts

$$0 \leq V_\alpha^{\mathrm{loc}}(u) \leq V_\alpha(u), \qquad 0 \leq V_\alpha^{p_k}(u) \leq V_\alpha(u) \qquad \text{and} \qquad 0 \leq V_\alpha^{\mathrm{lp}}(u) \leq V_\alpha(u),$$

see (i) of Theorem 5.3.

Proof of (iv). If u is a normalized equilibrium, then

$$V_\alpha^{\mathrm{lp}}(u) = V_\alpha^{p_k}(u) = V_\alpha^{\mathrm{loc}}(u) = V_\alpha(u) = 0$$

follows from (iii). For the last implication, let $V_\alpha^{\mathrm{lp}}(u) = 0$. Then we obtain

$$\sum_{i \in [N]} \langle (\theta_i)_{u^i}(u), v^i - u^i \rangle_{U_i^*, U_i} \geq 0 \qquad \forall\, v \in \mathcal{X} \tag{5.19}$$

by Theorem 5.3, (ii) and (iii). The pseudoconvexity of $\sum_{i \in [N]} \theta_i(\cdot, u^{-i})$ at u on $\mathcal{X}$ and (5.19) yield

$$\sum_{i \in [N]} \theta_i(u) \leq \sum_{i \in [N]} \theta_i(v^i, u^{-i})$$

for all $v \in \mathcal{X}$. Thus, u is a normalized equilibrium. □

We note that we obtain for all $\beta \geq \alpha \geq 0$ and $v \in \mathcal{X}$ the following estimate

$$\sum_{i\in[N]} \Big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\Big] \leq \sum_{i\in[N]} \Big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\Big],$$

which allows us to conclude the following useful relationships between the merit functionals

$$\begin{array}{ccccc} V_\beta(u) & \leq & V_\alpha(u) & \leq & V_0(u), \\ V_\beta^{\mathrm{loc}}(u) & \leq & V_\alpha^{\mathrm{loc}}(u) & \leq & V_0^{\mathrm{loc}}(u), \\ V_\beta^{p_k}(u) & \leq & V_\alpha^{p_k}(u) & \leq & V_0^{p_k}(u), \\ V_\beta^{\mathrm{lp}}(u) & \leq & V_\alpha^{\mathrm{lp}}(u) & \leq & V_0^{\mathrm{lp}}(u), \end{array}$$

for $u \in \mathcal{X}$ and $k \in \mathbb{N}$ fixed. Furthermore, for any $u \in \mathcal{X}$ with $\mathcal{X}$ convex, we conclude

$$\begin{array}{lclcl} u \text{ local normalized equilibrium} & \Longrightarrow & V_0^{\mathrm{loc}}(u) = 0 & \Longrightarrow & V_\alpha^{\mathrm{loc}}(u) = 0, \\ u \text{ local normalized equilibrium} & \Longrightarrow & V_0^{\mathrm{lp}}(u) = 0 & \Longrightarrow & V_\alpha^{\mathrm{lp}}(u) = 0, \\ u \text{ normalized equilibrium} & \Longrightarrow & V_0^{p_k}(u) = 0 & \Longrightarrow & V_\alpha^{p_k}(u) = 0, \\ u \text{ normalized equilibrium} & \Longrightarrow & V_0(u) = 0 & \Longrightarrow & V_\alpha(u) = 0, \end{array}$$

for all $\alpha \geq 0$. In the second line, the first implications can be converted to an equivalence if it holds $q = 0$. The reverse directions require suitable pseudoconvexity assumptions, see Theorem 5.4.

Summarizing, if $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$ is pseudoconvex at $u \in \mathcal{X}$ on $\widetilde{\mathcal{X}}(u)$ or $\widehat{\mathcal{X}}(u)$, respectively, we have shown that it holds for any $\alpha \geq 0$

$$\begin{array}{lclcll} V_0^{\mathrm{loc}}(u) = 0 & \Longleftrightarrow & V_\alpha^{\mathrm{loc}}(u) = 0 & \Longleftrightarrow & \sum_{i\in[N]} \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*, U_i} \geq 0 & \forall\, v \in \widetilde{\mathcal{X}}(u), \\ & & & \Longleftrightarrow & \sum_{i\in[N]} \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*, U_i} \geq 0 & \forall\, v \in \mathcal{X}, \\ & & & \Longleftrightarrow & u \in \mathcal{X} \text{ local normalized equilibrium}, & \\ V_0^{\mathrm{lp}}(u) = 0 & \Longleftrightarrow & V_\alpha^{\mathrm{lp}}(u) = 0 & \Longleftrightarrow & \sum_{i\in[N]} \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*, U_i} \geq 0 & \forall\, v \in \widehat{\mathcal{X}}(u), \\ & & & \Longleftrightarrow & \sum_{i\in[N]} \big\langle (\theta_i)_{u^i}(u), v^i - u^i \big\rangle_{U_i^*, U_i} \geq 0 & \forall\, v \in \mathcal{X}, \\ & & & \Longleftrightarrow & u \in \mathcal{X} \text{ local normalized equilibrium}. & \end{array}$$

Moreover, if $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$ is pseudoconvex at u on $\mathcal{X}$, we obtain for any $\alpha \geq 0$

$$\begin{array}{lclclcl} V_0^{\mathrm{lp}}(u) = 0 & \Longleftrightarrow & V_\alpha^{\mathrm{lp}}(u) = 0 & \Longleftrightarrow & V_0^{\mathrm{loc}}(u) = 0 & \Longleftrightarrow & V_\alpha^{\mathrm{loc}}(u) = 0, \\ & & & \Longleftrightarrow & V_0(u) = 0 & \Longleftrightarrow & V_\alpha(u) = 0, \end{array}$$

$$\iff \quad V_0^{p_k}(u) = 0 \quad \iff \quad V_\alpha^{p_k}(u) = 0,$$
$$\iff \quad u \text{ normalized equilibrium.}$$

Lastly, if $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$ is convex on $\mathcal{X}$, then we have shown the existence of a fixed point $u \in \mathcal{X}$ of the corresponding solution maps in Chapter 4. Consequently, we obtain $V_\alpha(u) = 0$ and we conclude that $u \in \mathcal{X}$ is a normalized equilibrium due to Theorem 5.4, (iv).

5.3 Difference of Merit Functionals

In this part of the thesis, we consider the difference between two regularized Nikaido–Isoda merit functionals, which are denoted by

$$V_{\alpha\beta}(u) = V_\alpha(u) - V_\beta(u),$$

and the corresponding localized modifications $V_{\alpha\beta}^{\text{loc}}$, $V_{\alpha\beta}^{p_k}$, and $V_{\alpha\beta}^{\text{lp}}$ for the parameters $\beta > \alpha \geq 0$, see also the works [45, 106]. We prove analytical properties of $V_{\alpha\beta}$, $V_{\alpha\beta}^{\text{loc}}$, $V_{\alpha\beta}^{p_k}$, and $V_{\alpha\beta}^{\text{lp}}$. In particular, we show a relation between roots of the regularized and localized Nikaido–Isoda merit functional and the roots of the difference of two of these.

In the following, we only require that (A1) of Assumption 3.1 is fulfilled, i.e., U and H are normed spaces. Furthermore, let $\mathcal{X} \subseteq U$ be nonempty, convex, closed, and bounded and $\widehat{\mathcal{X}}(u)$ be nonempty and closed. By definition of $\widehat{\mathcal{X}}(u)$ is also convex and bounded. Then $\mathcal{X}$ and $\widehat{\mathcal{X}}(u)$ are compact with respect to the weak sequential topology. Furthermore, note that $\widetilde{\mathcal{X}}(u)$ is compact with respect to the weak sequential topology as an intersection of two compact sets with respect to the weak sequential topology. Moreover, $\widetilde{\mathcal{X}}(u)$ is nonempty, convex, closed and bounded. To ensure the well-posedness of the definitions of the difference of the regularized and localized Nikaido–Isoda merit functionals, we assume that the supremum of $\Psi_\beta(u,\cdot)$ exists on $\mathcal{X}$.

An option for demonstrating this existence is to presume that $\theta_i : U \to \mathbb{R}$ is continuously differentiable in the i-th component, i.e., (B1) of Assumption 3.3 holds, and that $\alpha \geq 0$ is chosen such that

$$\sum_{i\in[N]} \left[\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2\right]$$

is convex on $\mathcal{X}$. In this setting, we are able to prove the existence of the supremum of $\Psi_\beta(u,\cdot)$ on $\mathcal{X}$. In fact, for $\beta > \alpha$ the convexity and differentiability of the functional

$$\sum_{i\in[N]} \left[\theta_i(\cdot, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(\cdot) - \iota_{H_i}(u^i)\|_{H_i}^2\right]$$

yields

$$\begin{aligned}\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\big] \\ \leq \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\big] \\ \leq -\sum_{i\in[N]} \langle(\theta_i)_{u^i}(u), v^i - u^i\rangle_{U_i^*,U_i}.\end{aligned} \tag{5.20}$$

Next, we apply standard estimates to (5.20) and get

$$\begin{aligned}\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\big] \\ \leq \sum_{i\in[N]} \|(\theta_i)_{u^i}(u)\|_{U_i^*}\|v^i - u^i\|_{U_i} \\ \leq \max_{i\in[N]}\big[\|(\theta_i)_{u^i}(u)\|_{U_i^*}\big] \sum_{i\in[N]} \big[\|v^i\|_{U_i} + \|u^i\|_{U_i}\big].\end{aligned}$$

In addition, by $\max_{i\in[N]} \|(\theta_i)_{u^i}(u)\|_{U_i^*} \leq C(u) < \infty$ and Young's inequality, we obtain

$$\begin{aligned}\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\big] \leq C(u) \sum_{i\in[N]} \big[\|v^i\|^2_{U_i} + \frac{1}{4} + \|u^i\|^2_{U_i} + \frac{1}{4}\big] \\ \leq C(u)\big[\|v\|^2_U + \|u\|^2_U + \frac{1}{2}N\big],\end{aligned}$$

where we recall that N denotes the number of players. Moreover, it holds $\|u\|^2_U \leq K(u) < \infty$, and due to the boundedness of the set $\mathcal{X} \subseteq U$, there exists some constant $K < \infty$ such that $\|v\|_U \leq K$ for all $v \in \mathcal{X}$. Hence, we arrive at

$$\sup_{v\in\mathcal{X}} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\big] < \infty.$$

Due to $\widetilde{\mathcal{X}}(u) \subseteq \mathcal{X}$ it also holds

$$\sup_{v\in\widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\beta}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|^2_{H_i}\big] < \infty.$$

Consequently, a maximizing sequence of $\Psi_\beta(u, \cdot)$ on $\mathcal{X}$ and $\widetilde{\mathcal{X}}(u)$ respectively, exists. By theorem Theorem 5.3, (i), it holds $V_\beta^{p_k}(u) \leq V_\beta(u) < \infty$ and $V_\beta^{\text{lp}}(u) \leq V_\beta(u) < \infty$. Thus, there exist maximizing sequences of $\Psi_\beta(u, \cdot) - \rho_k p(\| \cdot - u\|^2_U)$ on $\mathcal{X}$ and of $\Psi_\beta(u, \cdot) - q(\cdot - u)$ on $\widehat{\mathcal{X}}(u)$.

Note that the continuous differentiability of the objective functional $\theta_i : U \to \mathbb{R}$ is only assumed to demonstrate the existence of the supremum of $\Psi_\beta(u, \cdot)$ on the set $\mathcal{X}$. This assumption is not needed for the following results.

Theorem 5.5. For assertions (iii) and (iv), let $\sum_{i\in[N]} \theta_i(\cdot, u^{-i}) : U \to \mathbb{R}$ be lower semicontinuous with respect to the weak sequential topology. For the first equivalence in (iv) let $p(\|\cdot - u\|_U^2)$ be lower semicontinuous on $\mathcal{X}$ with respect to the weak sequential topology and for the second characterization therein we assume that $q(\cdot - u)$ is lower semicontinuous on $\widehat{\mathcal{X}}(u)$ with respect to the weak sequential topology. Then it holds:

(i) $V_{\alpha\beta}(u) \geq 0$, $V_{\alpha\beta}^{\text{loc}}(u) \geq 0$, $V_{\alpha\beta}^{p_k}(u) \geq 0$, and $V_{\alpha\beta}^{\text{lp}}(u) \geq 0$ for all $u \in U$.

(ii) $V_{\alpha\beta}(u) > 0$ and $V_{\alpha\beta}^{p_k}(u) > 0$ for all $u \notin \mathcal{X}$, $V_{\alpha\beta}^{\text{loc}}(u) > 0$ for all $u \notin \widetilde{\mathcal{X}}(u)$, and $V_{\alpha\beta}^{\text{lp}}(u) > 0$ for all $u \notin \widehat{\mathcal{X}}(u)$.

(iii)
$$V_{\alpha\beta}(u) = 0 \iff u \in \mathcal{X} \text{ and } V_\alpha(u) = 0,$$
$$V_{\alpha\beta}^{\text{loc}}(u) = 0 \iff u \in \widetilde{\mathcal{X}}(u) \text{ and } V_\alpha^{\text{loc}}(u) = 0.$$

(iv)
$$V_{\alpha\beta}^{p_k}(u) = 0 \iff u \in \mathcal{X} \text{ and } V_\alpha^{p_k}(u) = 0,$$
$$V_{\alpha\beta}^{\text{lp}}(u) = 0 \iff u \in \widehat{\mathcal{X}}(u) \text{ and } V_\alpha^{\text{lp}}(u) = 0.$$

Proof. Proof of (i). Due to the assumption that the supremum of $\Psi_\beta(u, \cdot)$ on $\mathcal{X}$ is finite, the implications $\widetilde{\mathcal{X}}(u) \subseteq \mathcal{X}$ and $\widehat{\mathcal{X}}(u) \subseteq \mathcal{X}$, the nonnegativity of $p : \mathbb{R} \to \mathbb{R}$ and $q : U \to \mathbb{R}$ yield

$$\begin{aligned}
&& \sup_{v\in\widetilde{\mathcal{X}}(u)} \Psi_\beta(u, v) &< \infty, \\
\sup_{v\in\mathcal{X}} \left[\Psi_\beta(u, v) - \rho_k p(\|v - u\|_U^2)\right] &\leq & \sup_{v\in\mathcal{X}} \Psi_\beta(u, v) &< \infty, \\
\sup_{v\in\widehat{\mathcal{X}}(u)} \left[\Psi_\beta(u, v) - q(v - u)\right] &\leq & \sup_{v\in\mathcal{X}} \Psi_\beta(u, v) &< \infty.
\end{aligned} \tag{5.21}$$

Hence, there exist some maximizing sequences $\{v_l\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ of $\Psi_\beta(u, \cdot)$, $\{v_l^{\text{loc}}\}_{l\in\mathbb{N}} \subseteq \widetilde{\mathcal{X}}(u)$ of $\Psi_\beta(u, \cdot)$, $\{v_l^{p_k}\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ of $\Psi_\beta(u, \cdot) - \rho_k p(\|\cdot - u\|_U^2)$, and $\{v_l^{\text{lp}}\}_{l\in\mathbb{N}} \subseteq \widehat{\mathcal{X}}(u)$ of $\Psi_\beta(u, \cdot) - q(\cdot - u)$. For the maximizing sequence $\{v_l\}_{l\in\mathbb{N}}$ we get

$$0 \leq \frac{\beta - \alpha}{2}\|\iota_H(v_l) - \iota_H(u)\|_H^2 = \Psi_\alpha(u, v_l) - \Psi_\beta(u, v_l) \leq \sup_{v\in\mathcal{X}} \left[\Psi_\alpha(u, v)\right] - \Psi_\beta(u, v_l), \tag{5.22}$$

using $\beta > \alpha$ in the first estimate and the definition of a supremum in the last equation. Thus, we obtain the inequalities for the corresponding maximizing sequences by a similar reasoning

$$\begin{aligned}
0 &\leq \Psi_\alpha(u, v_l^{\text{loc}}) - \Psi_\beta(u, v_l^{\text{loc}}) \leq \sup_{v\in\widetilde{\mathcal{X}}(u)} \left[\Psi_\alpha(u, v)\right] - \Psi_\beta(u, v_l^{\text{loc}}), \\
0 &\leq \Psi_\alpha(u, v_l^{p_k}) - \Psi_\beta(u, v_l^{p_k}) \leq \sup_{v\in\mathcal{X}} \left[\Psi_\alpha(u, v) - \rho_k p(\|v - u\|_U^2)\right] - \Psi_\beta(u, v_l^{p_k}) \\
&\qquad + \rho_k p(\|v_l^{p_k} - u\|_U^2), \\
0 &\leq \Psi_\alpha(u, v_l^{\text{lp}}) - \Psi_\beta(u, v_l^{\text{lp}}) \leq \sup_{v\in\widehat{\mathcal{X}}(u)} \left[\Psi_\alpha(u, v) - q(v - u)\right] - \Psi_\beta(u, v_l^{\text{lp}}) + q(v_l^{\text{lp}} - u).
\end{aligned} \tag{5.23}$$

Taking the limit $l \to \infty$ on the right side of the inequalities (5.22) and (5.23) results in

$$\begin{aligned}
V_{\alpha\beta}(u) &= \sup_{v\in\mathcal{X}} \big[\Psi_\alpha(u,v)\big] - \sup_{v\in\mathcal{X}} \Psi_\beta(u,v) && \geq 0,\\
V_{\alpha\beta}^{\text{loc}}(u) &= \sup_{v\in\widetilde{\mathcal{X}}(u)} \big[\Psi_\alpha(u,v)\big] - \sup_{v\in\widetilde{\mathcal{X}}(u)} \Psi_\beta(u,v) && \geq 0,\\
V_{\alpha\beta}^{p_k}(u) &= \sup_{v\in\mathcal{X}} \big[\Psi_\alpha(u,v) - \rho_k p(\|v-u\|_U^2)\big] - \sup_{v\in\mathcal{X}} \big[\Psi_\beta(u,v) - \rho_k p(\|v-u\|_U^2)\big] && \geq 0,\\
V_{\alpha\beta}^{\text{lp}}(u) &= \sup_{v\in\widehat{\mathcal{X}}(u)} \big[\Psi_\alpha(u,v) - q(v-u)\big] - \sup_{v\in\widehat{\mathcal{X}}(u)} \big[\Psi_\beta(u,v) - q(v-u)\big] && \geq 0.
\end{aligned}$$

Proof of (ii). Let $u \notin \mathcal{X}$ and $\{v_l\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ be a maximizing sequence. By compactness of $\mathcal{X}$ with respect to the weak sequential topology, it yields the existence of some subsequence $\{v_{l_m}\}_{m\in\mathbb{N}} \subseteq \{v_l\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ such that $v_{l_m} \rightharpoonup \bar{v} \in \mathcal{X}$ as $m \to \infty$ holds true and hence, it yields $\bar{v} \neq u$. Thus, we obtain $\iota_H(v_{l_m}) \to \iota_H(\bar{v})$ in H as $m \to \infty$ due to the complete continuity of the embedding operator $\iota_H : U \to H$. Since ι_H is also injective, we have $\iota_H(\bar{v}) \neq \iota_H(u)$. Hence, there exists some $\varepsilon > 0$ such that $\|\iota_H(v_{l_m}) - \iota_H(u)\|_H \geq \varepsilon$. Moreover, for any sequence $\{x_l\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ it holds

$$\limsup_{l\to\infty} \Psi_\alpha(u,x_l) = \lim_{n\to\infty} \sup\{\Psi_\alpha(u,x_l) : l \geq n\}$$

by the definition of the limes superior. Since the maximizing sequence $\{v_{l_m}\}_{m\in\mathbb{N}} \subseteq \{v_l\}_{l\in\mathbb{N}}$ lies in the set $\mathcal{X}$ we get

$$\begin{aligned}
\limsup_{m\to\infty} \Psi_\alpha(u,v_{l_m}) &\leq \lim_{n\to\infty} \sup\{\Psi_\alpha(u,v) : v \in \mathcal{X}\}\\
&= \sup_{v\in\mathcal{X}} \Psi_\alpha(u,v).
\end{aligned} \tag{5.24}$$

Hence, by definition of a maximizing sequence and applying the estimate (5.24) we obtain

$$\begin{aligned}
V_{\alpha\beta}(u) &= \sup_{v\in\mathcal{X}} \big[\Psi_\alpha(u,v)\big] - \sup_{v\in\mathcal{X}} \Psi_\beta(u,v)\\
&= \sup_{v\in\mathcal{X}} \big[\Psi_\alpha(u,v)\big] - \lim_{m\to\infty} \Psi_\beta(u,v_{l_m})\\
&\geq \limsup_{m\to\infty} \big[\Psi_\alpha(u,v_{l_m})\big] - \lim_{m\to\infty} \Psi_\beta(u,v_{l_m}).
\end{aligned}$$

Using the superaddivity of the limes superior we arrive at

$$\begin{aligned}
V_{\alpha\beta}(u) &\geq \limsup_{m\to\infty} \big[\Psi_\alpha(u,v_{l_m}) - \Psi_\beta(u,v_{l_m})\big]\\
&= \limsup_{m\to\infty} \frac{\beta-\alpha}{2} \|\iota_H(v_{l_m}) - \iota_H(u)\|_H^2\\
&\geq \frac{\beta-\alpha}{2}\varepsilon^2,
\end{aligned} \tag{5.25}$$

and $V_{\alpha\beta}(u) > 0$ for $u \notin \mathcal{X}$.

An analogous calculation holds for $V^{\mathrm{loc}}_{\alpha\beta}(u)$, $V^{p_k}_{\alpha\beta}(u)$, and $V^{\mathrm{lp}}_{\alpha\beta}(u)$. To this end, we consider $u \notin \widetilde{\mathcal{X}}(u)$, $u \notin \mathcal{X}$ or $u \notin \widehat{\mathcal{X}}(u)$, respectively. By (5.21), the corresponding maximizing sequence $\{v^{\mathrm{loc}}_{l_m}\}_{m\in\mathbb{N}} \subseteq \widetilde{\mathcal{X}}(u)$ of $\Psi_\beta(u,\cdot)$, $\{v^{p_k}_{l_m}\}_{m\in\mathbb{N}} \subseteq \mathcal{X}$ of $\Psi_\beta(u,\cdot) - \rho_k p(\|\cdot - u\|^2_U)$, and $\{v^{\mathrm{lp}}_{l_m}\}_{m\in\mathbb{N}} \subseteq \widehat{\mathcal{X}}(u)$ of $\Psi_\beta(u,\cdot) - q(\cdot - u)$ exists. Thus, there are some constants $\varepsilon^{\mathrm{loc}} > 0$, $\varepsilon^{p_k} > 0$, and $\varepsilon^{\mathrm{lp}} > 0$ such that $\|\iota_H(v^{\mathrm{loc}}_{l_m}) - \iota_H(u)\|_H > \varepsilon^{\mathrm{loc}}$, $\|\iota_H(v^{p_k}_{l_m}) - \iota_H(u)\|_H > \varepsilon^{p_k}$, and $\|\iota_H(v^{\mathrm{lp}}_{l_m}) - \iota_H(u)\|_H > \varepsilon^{\mathrm{lp}}$, respectively. Furthermore, using the estimate (5.25) it yields

$$
\begin{aligned}
V^{\mathrm{loc}}_{\alpha\beta}(u) &\geq \limsup_{m\to\infty} \left[\Psi_\alpha(u, v^{\mathrm{loc}}_{l_m})\right] - \lim_{m\to\infty} \Psi_\beta(u, v^{\mathrm{loc}}_{l_m}) \\
&\geq \frac{\beta-\alpha}{2}(\varepsilon^{\mathrm{loc}})^2, \\
V^{p_k}_{\alpha\beta}(u) &\geq \limsup_{m\to\infty} \left[\Psi_\alpha(u, v^{p_k}_{l_m}) - \rho_k p(\|v^{p_k}_{l_m} - u\|^2_U)\right] - \lim_{m\to\infty} \left[\Psi_\beta(u, v^{p_k}_{l_m}) - \rho_k p(\|v^{p_k}_{l_m} - u\|^2_U)\right] \\
&\geq \frac{\beta-\alpha}{2}(\varepsilon^{p_k})^2, \\
V^{\mathrm{lp}}_{\alpha\beta}(u) &\geq \limsup_{m\to\infty} \left[\Psi_\alpha(u, v^{\mathrm{lp}}_{l_m}) - q(v^{\mathrm{lp}}_{l_m} - u)\right] - \lim_{m\to\infty} \left[\Psi_\beta(u, v^{\mathrm{lp}}_{l_m}) - q(v^{\mathrm{lp}}_{l_m} - u)\right] \\
&\geq \frac{\beta-\alpha}{2}(\varepsilon^{\mathrm{lp}})^2,
\end{aligned}
$$

and hence, we obtain $V^{\mathrm{loc}}_{\alpha\beta}(u) > 0$ for $u \notin \widetilde{\mathcal{X}}(u)$, $V^{p_k}_{\alpha\beta}(u) > 0$ for $u \notin \mathcal{X}$, and $V^{\mathrm{lp}}_{\alpha\beta}(u) > 0$ for $u \notin \widehat{\mathcal{X}}(u)$.

Proof of (iii). For the forward implication "⇒", let $V_{\alpha\beta}(u) = 0$ be valid. Since $V_{\alpha\beta}(u) > 0$ holds for $u \notin \mathcal{X}$ by (ii), we have $u \in \mathcal{X}$. Let $\{v_l\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ be a maximizing sequence of $\Psi_\beta(u,\cdot)$ on $\mathcal{X}$. By compactness of $\mathcal{X}$ with respect to the weak sequential topology, there exists a subsequence $\{v_{l_m}\}_{m\in\mathbb{N}} \subseteq \{v_l\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ such that $v_{l_m} \rightharpoonup \bar{v} \in \mathcal{X}$ as $m \to \infty$ and by the complete continuity of the embedding operator ι_H it holds $\iota_H(v_{l_m}) \to \iota_H(\bar{v})$ in H as $m \to \infty$. Applying the equation (5.25) we obtain

$$V_{\alpha\beta}(u) \geq \limsup_{m\to\infty} \frac{\beta-\alpha}{2}\|\iota_H(v_{l_m}) - \iota_H(u)\|^2_H \geq 0.$$

By the assumption $V_{\alpha\beta}(u) = 0$, it follows

$$\limsup_{m\to\infty} \frac{\beta-\alpha}{2}\|\iota_H(v_{l_m}) - \iota_H(u)\|^2_H = 0.$$

Thus, it yields the strong convergences $\iota_H(v_{l_m}) \to \iota_H(u)$ in H and $v_{l_m} \rightharpoonup u$ in U as $m \to \infty$ by uniqueness of limits and injectivity of ι_H. Using the definition

$$\Psi_\beta(u,v) = \sum_{i\in[N]} [\theta_i(u)] - \widetilde{\Psi}_\beta(u,v),$$

the lower semicontinuity of $\sum_{i\in[N]} \theta_i(\cdot, u^{-i}) : U_i \to \mathbb{R}$ with respect to the weak sequential topology and Lemma 4.3, we obtain the upper semicontinuity with respect to the weak sequential topology of $\Psi_\beta(u,\cdot) : U \to \mathbb{R}$. Then it follows

$$V_\beta(u) = \limsup_{m\to\infty} \Psi_\beta(u, v_{l_m}) \leq \Psi_\beta(u,u) = 0,$$

and thus $V_\alpha(u) = V_\beta(u) + V_{\alpha\beta}(u) = 0$.

In case of the other direction "⇐", we assume $V_{\alpha\beta}(u) > 0$ for $u \in U$. Then it yields $V_\alpha(u) = V_\beta(u) + V_{\alpha\beta}(u) > V_\beta(u)$ and either $u \in \mathcal{X}$ or $u \notin \mathcal{X}$. In the case $u \in \mathcal{X}$, we additionally have $V_\beta(u) \geq 0$ and thus $V_\alpha(u) > 0$.

Next, we move on to the second equivalence. For the direct implication "⇒", we assume $V^{\text{loc}}_{\alpha\beta}(u) = 0$. Then we have $u \in \widetilde{\mathcal{X}}(u)$ by assertion (ii). Denote the weakly converging maximizing subsequence of $\Psi_\beta(u,\cdot)$ on $\widetilde{\mathcal{X}}(u)$ by $\{v^{\text{loc}}_{l_m}\}_{m\in\mathbb{N}}$ with $v^{\text{loc}}_{l_m} \rightharpoonup \bar{v}^{\text{loc}} \in \widetilde{\mathcal{X}}(u)$ as $m \to \infty$. By a similar reasoning as in the first equivalence, we get the strong convergence of $\iota_H(v^{\text{loc}}_{l_m}) \to \iota_H(u)$ in H as $m \to \infty$ and therefore, it yields $v^{\text{loc}}_{l_m} \rightharpoonup u$ in U as $m \to \infty$ by the injectivity of the operator ι_H. Thus, we obtain

$$V^{\text{loc}}_\alpha(u) = V^{\text{loc}}_\beta(u) + V^{\text{loc}}_{\alpha\beta}(u) = V^{\text{loc}}_\beta(u) = \limsup_{m\to\infty} \Psi_\beta(u, v^{\text{loc}}_{l_m}) \leq 0$$

by the assumption of $V^{\text{loc}}_{\alpha\beta}(u) = 0$ and the upper semicontinuity of $\Psi_\beta(u,\cdot) : U \to \mathbb{R}$ with respect to the weak sequential topology.

To demonstrate the other direction "⇐", we assume $V^{\text{loc}}_{\alpha\beta}(u) > 0$ holds for all $u \in U$. Then either $u \in \widetilde{\mathcal{X}}(u)$ or $u \notin \widetilde{\mathcal{X}}(u)$. Furthermore, we get $V^{\text{loc}}_\alpha(u) > V^{\text{loc}}_\beta(u) \geq 0$ for $u \in \widetilde{\mathcal{X}}(u)$.

Proof of (iv). For the functionals $V^{p_k}_{\alpha\beta}(u)$ and $V^{\text{lp}}_{\alpha\beta}(u)$ we argue analogously using the corresponding lower semicontinuity with respect to the weak sequential topology.

For the direction "⇒", let it hold either $V^{p_k}_{\alpha\beta}(u) = 0$ or $V^{\text{lp}}_{\alpha\beta}(u) = 0$. Then we have $u \in \mathcal{X}$ and $u \in \widehat{\mathcal{X}}(u)$ by (ii), respectively. Denoting the corresponding maximizing sequences of $\Psi_\beta(u,\cdot) - \rho_k p(\|\cdot - u\|^2_U)$ on $\mathcal{X}$ and of $\Psi_\beta(u,\cdot) - q(\cdot - u)$ on $\widehat{\mathcal{X}}(u)$ by $\{v^{p_k}_l\}_{l\in\mathbb{N}}$ and $\{v^{\text{lp}}_l\}_{l\in\mathbb{N}}$, respectively, the compactness of $\mathcal{X}$ and $\widehat{\mathcal{X}}(u)$ with respect to the weak sequential topology yields the existence of subsequences $\{v^{p_k}_{l_m}\}_{m\in\mathbb{N}}$ and $\{v^{\text{lp}}_{l_m}\}_{l\in\mathbb{N}}$ such that it holds $v^{p_k}_{l_m} \rightharpoonup \bar{v}^{p_k} \in \mathcal{X}$ and $v^{\text{lp}}_{l_m} \rightharpoonup \bar{v}^{\text{lp}} \in \widehat{\mathcal{X}}(u)$ as $m \to \infty$, respectively. Following the argumentation in (iii), we obtain the strong convergences $\iota_H(v^{p_k}_{l_m}) \to \iota_H(u)$ in H and $\iota_H(v^{\text{lp}}_{l_m}) \to \iota_H(u)$ in H as $m \to \infty$ and therefore, it holds $v^{p_k}_{l_m} \rightharpoonup u$ and $v^{\text{lp}}_{l_m} \rightharpoonup u$ in U as $m \to \infty$ by the injectivity of ι_H. By the assumption of $V^{p_k}_{\alpha\beta}(u) = 0$, we have

$$V^{p_k}_\alpha(u) = V^{p_k}_\beta(u) + V^{p_k}_{\alpha\beta}(u) = V^{p_k}_\beta(u).$$

By arguments as in Lemma 4.3 and the assumption that $\sum_{i\in[N]} \theta_i(\cdot, u^{-i})$, $p(\|\cdot - u\|^2_U)$ and $q(v - \cdot)$ are lower semicontinuous with respect to the weak sequential topology, we get the lower semicontinuity of $\Psi_\beta(u,\cdot) - p(\|\cdot - u\|^2_U)$ on $\mathcal{X}$ and $\Psi_\beta(u,\cdot) - q(v - \cdot)$ on $\widehat{\mathcal{X}}(u)$ with respect to the weak sequential topology. Using this fact, we obtain

$$\begin{aligned} V^{p_k}_\alpha(u) &= \limsup_{m\to\infty} \left[\Psi_\beta(u, v^{p_k}_{l_m}) - \rho_k p(\|v^{p_k}_{l_m} - u\|^2_U)\right] \\ &\leq \Psi_\beta(u,u) - \rho_k p(\|u - u\|^2_U). \end{aligned}$$

Hence, we get $V_\alpha^{p_k}(u) = 0$ for $u \in \mathcal{X}$. We present the similar argument for the difference of two merit functionals $V_{\alpha\beta}^{\mathrm{lp}}$. The assumption $V_{\alpha\beta}^{\mathrm{lp}}(u) = 0$ yields

$$V_\alpha^{\mathrm{lp}}(u) = V_\beta^{\mathrm{lp}}(u) + V_{\alpha\beta}^{\mathrm{lp}}(u) = V_\beta^{\mathrm{lp}}(u).$$

Since $\Psi_\beta(u,\cdot) - q(u-\cdot)$ is lower semicontinuous on $\widehat{\mathcal{X}}(u)$ we arrive at

$$\begin{aligned} V_\alpha^{\mathrm{lp}}(u) &= \limsup_{m\to\infty} \left[\Psi_\beta(u, v_{l_m}^{\mathrm{lp}}) - q(v_{l_m}^{\mathrm{lp}} - u)\right] \\ &\leq \Psi_\beta(u,u) - q(u-u). \end{aligned}$$

All together, we obtain $V_\alpha^{\mathrm{lp}}(u) = 0$ for $u \in \widehat{\mathcal{X}}(u)$.

For the direction "$\Leftarrow$" we assume that $V_{\alpha\beta}^{p_k}(u) > 0$ or $V_{\alpha\beta}^{\mathrm{lp}}(u) > 0$ holds for all $u \in U$. Then either $u \in \mathcal{X}$ or $u \notin \mathcal{X}$ and $u \in \widehat{\mathcal{X}}(u)$ or $u \notin \widehat{\mathcal{X}}(u)$, respectively. Furthermore, we get $V_\alpha^{p_k}(u) > V_\beta^{p_k}(u) \geq 0$ for $u \in \mathcal{X}$ and $V_\alpha^{\mathrm{lp}}(u) > V_\beta^{\mathrm{lp}}(u) \geq 0$ for $u \in \widehat{\mathcal{X}}(u)$, respectively. $\square$

By the nonnegativity of $V_{\alpha\beta}(u)$ and $V_\alpha(u)$ on U, the third and fourth properties of Theorem 5.5 are equivalent to

$$\begin{aligned} V_{\alpha\beta}(u) > 0 &\iff u \notin \mathcal{X} \quad \text{or} \quad V_\alpha(u) > 0, \\ V_{\alpha\beta}^{\mathrm{loc}}(u) > 0 &\iff u \notin \widetilde{\mathcal{X}}(u) \quad \text{or} \quad V_\alpha^{\mathrm{loc}}(u) > 0, \\ V_{\alpha\beta}^{p_k}(u) > 0 &\iff u \notin \mathcal{X} \quad \text{or} \quad V_\alpha^{p_k}(u) > 0, \\ V_{\alpha\beta}^{\mathrm{lp}}(u) > 0 &\iff u \notin \widehat{\mathcal{X}}(u) \quad \text{or} \quad V_\alpha^{\mathrm{lp}}(u) > 0. \end{aligned}$$

In the proof of the second assertion, the compactness of $\mathcal{X}$ is needed, because the subsequences $\{v_{l_m}\}$, $\{v_{l_m}^{\mathrm{loc}}\}$, $\{v_{l_m}^{p_k}\}$, and $\{v_{l_m}^{\mathrm{lp}}\}$ have to lie in $\mathcal{X}$, $\widetilde{\mathcal{X}}(u)$, $\mathcal{X}$, or $\widehat{\mathcal{X}}(u)$, respectively.

Note that we have no assumptions on the regularization parameters $\alpha \geq 0$ or $\beta \geq 0$. Thus, we do not use convexity of the cost functionals except to show the existence of supremum of $\Psi_\beta(u,\cdot)$ on the set $\mathcal{X}$.

The essential statement is the equivalence between $V_{\alpha\beta}(u) = 0$ and $V_\alpha(u) = 0$ for $u \in \mathcal{X}$. Thus, one could search for an element $u \in U$ such that it holds $V_{\alpha\beta}(u) = 0$, $V_{\alpha\beta}^{\mathrm{loc}}(u) = 0$, $V_{\alpha\beta}^{p_k}(u) = 0$ or $V_{\alpha\beta}^{\mathrm{lp}}(u) = 0$ and get a root of the corresponding regularized and localized Nikaido–Isoda merit functionals V_α, V_α^{loc}, $V_\alpha^{p_k}$ or V_α^{lp}, respectively. Furthermore, in Subsection 5.2.2 we have established assumptions under which these conditions give a (local) normalized equilibrium.

5.4 Differentiability of Regularized and Localized Nikaido–Isoda Merit Functionals

In this section, we study the differentiability of the regularized Nikaido–Isoda merit functionals $\widetilde{V}_\alpha$, V_α, and their localized modifications $\widetilde{V}_\alpha^{\mathrm{loc}}$, $\widetilde{V}_\alpha^{p_k}$, $\widetilde{V}_\alpha^{\mathrm{lp}}$, V_α^{loc}, $V_\alpha^{p_k}$ and V_α^{lp}. These different Nikaido–Isoda merit functionals are defined in (5.1)–(5.8). We have already seen these functionals in Subsection 5.2.1 and Subsection 5.2.2 where we studied their connection to equilibria. In this section, we will differentiate this kind of merit functionals. First of all, we consider the corresponding merit functionals based on the regularized Nikaido–Isoda functional on the fixed sets $\mathcal{X}_i$ and $\mathcal{X}$, respectively. We show that we can apply a suitable version of Danskin's theorem, see Section 2.4 above, on these functionals. In Subsection 5.4.2, we discuss the differentiability of the regularized and localized merit functionals $\widetilde{V}_\alpha^{\text{loc}}$, $\widetilde{V}_\alpha^{p_k}$, $\widetilde{V}_\alpha^{\text{lp}}$, V_α^{loc}, $V_\alpha^{p_k}$ and V_α^{lp}. For feasible sets dependent on u itself we need additional assumptions in order to use Danskin's theorem, and for independent ones we apply Danskin's theorem directly. In the following, let (A7) of Assumption 3.1 hold for the underlying spaces. Furthermore, let $\mathcal{X}_i \subseteq U_i$ and $\mathcal{X} \subseteq U$ be nonempty, convex, closed, and bounded. Consequently, they are compact with respect to the respective weak sequential topology. We assume that (B2c) of Assumption 3.3 for the objective functional $\theta_i : U \to \mathbb{R}$ is valid.

5.4.1 Differentiability of regularized Merit Nikaido–Isoda functionals

In this subsection, we consider the regularized Nikaido–Isoda merit functional $\widetilde{V}_\alpha$, defined in (5.1), on the nonempty, convex, closed, and bounded set $\mathcal{X} = \Pi_i \mathcal{X}_i$. Moreover, we investigate the merit functional V_α, see (5.2), which is defined on the nonempty, convex, closed, and bounded set $\mathcal{X}$. These sets are independent of u^i and u, respectively. Thus, we can apply Danskin's theorem, see Theorem 2.75.

First, we consider the regularized Nikaido–Isoda merit functional $\widetilde{V}_\alpha$, which simplifies to

$$\widetilde{V}_\alpha(u) = \sum_{i\in[N]} \max_{v^i\in\mathcal{X}_i} \left[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\right]. \tag{5.26}$$

Theorem 5.6. Let $\alpha \geq 0$ and the nonempty, open set A be such that the solution mapping $u \mapsto v_{f_i}(u)$ to the maximization problem $\max_{v^i\in\mathcal{X}_i} f_i(u, v^i)$ where $f_i : U \times U_i \to \mathbb{R}$ is given by

$$f_i(u, v^i) = -\theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2,$$

is single-valued on A for each $i \in [N]$. Then $\widetilde{V}_\alpha$, defined in (5.26), is continuously differentiable on the set A. For the derivative there holds

$$\begin{aligned} \mathrm{d}\widetilde{V}_\alpha(u)h = \sum_{i\in[N]} \Big[&\big\langle(\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}} + \gamma\langle J_{U_i}(u^i), h^i\rangle_{U_i^*,U_i} \\ &- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle(\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_{f_i}(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j)\big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j}\Big] \\ &+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_{f_i}(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}\Big], \end{aligned} \tag{5.27}$$

and the derivative is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. We prove the differentiability by Danskin's theorem, see Section 2.4. To this end, we show that each term of the sum in (5.26) is differentiable. Since the term $\theta_i(u)$ is independent of v^i, we split up the maximum term and get

$$\begin{aligned}\max_{v^i\in\mathcal{X}_i} &\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big]\\ &=\theta_i(u)+\max_{v^i\in\mathcal{X}_i}\big[-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big].\end{aligned}$$

By this characterization, we consider the terms $\theta_i(u)$ and

$$\max_{v^i\in\mathcal{X}_i} f_i(u,v^i) \tag{5.28}$$

of the sum separately where the objective functional is given by

$$f_i(u,v^i)=-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2.$$

The mapping $u\mapsto\theta_i(u)$ is continuously differentiable with

$$\big\langle(\theta_i)_u(u),h\big\rangle_{U^*,U}=\big\langle(\tilde\theta_i)_{\tilde u}(\iota_{\widetilde U}(u)),\iota_{\widetilde U}(h)\big\rangle_{\widetilde U^*,\widetilde U}+\gamma\langle J_{U_i}(u^i),h^i\rangle_{U_i^*,U_i}$$

for all $h\in U$. For the second term (5.28) we apply Danskin's theorem in the setting such that U represents X, U_i represents Z, and the corresponding topologies τ_X and τ_Z are the weak sequential topologies of U and U_i, respectively, see Section 2.4. Then we have to satisfy Assumption 2.66 and apply Theorem 2.75 to obtain the differentiability of the maximization problem (5.28). To this end, we define $f_i^1: U\times U_i\to\mathbb{R}$ and $f_i^2: U_i\to\mathbb{R}$ by

$$\begin{aligned}f_i^1(u,v^i)&=-\tilde\theta_i(\iota_{\widetilde U}(v^i,u^{-i}))-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2,\\ f_i^2(v^i)&=-\frac{\gamma}{2}\|v^i\|_{U_i}^2.\end{aligned}$$

Here, f_i^1 is continuous in both arguments with respect to the weak sequential topology and f_i^2 is upper semicontinuous with respect to the weak sequential topology. Thus, $f_i=f_i^1+f_i^2$ is continuous in u and upper semicontinuous in both variables with respect to the weak sequential topology.

Due to the assumption that H_i is a uniformly smooth and uniformly convex Banach space, the mapping $\|\iota_{H_i}(v^i)-\iota_{H_i}(\cdot)\|_{H_i}^2: U_i\to\mathbb{R}$ is differentiable in direction $h^i\in U_i$ with the continuous derivative

$$-2\big\langle J_{H_i}(\iota_{H_i}(v^i)-\iota_{H_i}(\cdot)),\iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}$$

from the weak sequential topology of U_i to the norm topology of U_i^*. Since $\tilde\theta_i\circ\iota_{\widetilde U}: U\to\mathbb{R}$ is continuously differentiable with respect to the weak sequential topology, we obtain the continuity of $(f_i^1)_u$ in both arguments from the weak sequential topology of $U\times U_i$ to the norm topology of U^*. Furthermore, f_i^2 is independent of u which yields

$$\big\langle(f_i)_u(u,v^i),h\big\rangle_{U^*,U}=\big\langle(f_i^1)_u(u,v^i),h\big\rangle_{U^*,U}=\sum_{j\in[N]}\big\langle(f_i^1)_{u^j}(u,v^i),h^j\big\rangle_{U_j^*,U_j}.$$

Then the derivative of f_i with respect to u reads

$$\begin{aligned}\langle (f_i)_u(u,v^i),h\rangle_{U^*,U} &= -\sum_{j\in[N]}\Big[\big\langle(\tilde\theta_i(\iota_{\widetilde U}(v^i,u^{-i})))_{u^j},h^j\big\rangle_{U_j^*,U_j}\\ &\quad+\frac{\alpha}{2}\big\langle(\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|^2_{H_i})_{u^j},h^j\big\rangle_{U_j^*,U_j}\Big]\\ &= -\sum_{j\in[N]\setminus\{i\}}\Big[\big\langle(\tilde\theta_i)_{\tilde u^j}(\iota_{\widetilde U}(v^i,u^{-i})),\iota_{\widetilde U_j}(h^j)\big\rangle_{\widetilde U_j^*,\widetilde U_j}\Big]\\ &\quad+\alpha\big\langle J_{H_i}(\iota_{H_i}(v^i)-\iota_{H_i}(u^i)),\iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}.\end{aligned}\tag{5.29}$$

We see that $(f_i)_u$ is continuous in both arguments from the weak sequential topology of $U\times U_i$ to the norm topology of U^*. Thus, Assumption 2.66 is fulfilled and we can apply Danskin's theorem to the regularized Nikaido–Isoda merit functional $\widetilde V_\alpha$. If one chooses $\alpha\geq 0$ sufficiently large such that the solution map $u\mapsto v_{f_i}(u)$ to (5.28) is single-valued on some open set A for each $i\in[N]$, then the mapping $u\mapsto\max_{v^i\in\mathcal X_i} f_i(u,v^i)$ is continuously differentiable on A by Danskin's theorem, see Theorem 2.75. For the derivative we apply (5.29) and obtain

$$\begin{aligned}\mathrm d_u\big[\max_{v^i\in\mathcal X_i} f_i(u,v^i)\big]h &= \mathrm d_u\max_{v^i\in\mathcal X_i}\big[-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|^2_{H_i}\big]h\\ &= -\sum_{j\in[N]\setminus\{i\}}\Big[\big\langle(\tilde\theta_i)_{\tilde u^j}(\iota_{\widetilde U}(v_{f_i}(u),u^{-i})),\iota_{\widetilde U_j}(h^j)\big\rangle_{\widetilde U_j^*,\widetilde U_j}\Big]\\ &\quad+\alpha\big\langle J_{H_i}(\iota_{H_i}(v_{f_i}(u))-\iota_{H_i}(u^i)),\iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}.\end{aligned}\tag{5.30}$$

and the derivative is continuous from the weak sequential topology of U to the norm topology of U^*. Altogether, we have shown that $\widetilde V_\alpha(u)$, defined as in equation (5.26), is continuously differentiable on A. □

Note, we do not achieve the continuous differentiability of the regularized Nikaido–Isoda merit functional, defined in (2.23), unless we assume that $\mathcal X=\prod_{i\in[N]}\mathcal X_i$ and thus, it holds $F_i(u^{-i})=\mathcal X_i$.

Remark 5.7. Under weaker assumptions, some results of Section 2.4 are still applicable. For instance, if $\tilde\theta_i:\widetilde U\to\mathbb R$ is only continuous, then, without requiring the single-valuedness of the solution map $u\mapsto v_{f_i}(u)$ for $f_i(u,v^i)=-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|^2_{H_i}$, Lemma 2.70 still yields that (5.28) is continuous with respect to the weak sequential topology and thus, $\widetilde V_\alpha(u)$ is continuous with respect to the norm topology and lower semicontinuous with respect to the weak sequential topology.

Next, we move on to the regularized Nikaido–Isoda merit functional V_α, see (5.2). Recall the definition of the merit functional

$$V_\alpha(u)=\sup_{v\in\mathcal X}\sum_{i\in[N]}\big[\theta_i(u)-\theta_i(v^i,u^{-i})-\frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|^2_{H_i}\big].\tag{5.31}$$

Theorem 5.8. Let $\alpha \geq 0$ and the nonempty, open set A be such that the solution mapping $u \mapsto v_f(u)$ to the maximization problem $\max_{v\in\mathcal{X}} f(u,v)$ where $f : U \times U \to \mathbb{R}$ is given by

$$f(u,v) = -\sum_{i\in[N]} [\theta_i(v^i, u^{-i})] - \frac{\alpha}{2}\|\iota_H(v) - \iota_H(u)\|_H^2,$$

is single-valued on A. Then V_α, defined in (5.31), is continuously differentiable on the set A. For the derivative there holds

$$\begin{aligned} \mathrm{d}V_\alpha(u)h = \sum_{i\in[N]} \Big[&\big\langle(\tilde\theta_i)_{\tilde u}(\iota_{\widetilde U}(u)), \iota_{\widetilde U}(h)\big\rangle_{\widetilde U^*,\widetilde U} + \gamma\langle J_{U_i}(u^i), h^i\rangle_{U_i^*,U_i} \\ &- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle(\tilde\theta_i)_{\tilde u^j}(\iota_{\widetilde U}(v_f^i(u), u^{-i})), \iota_{\widetilde U_j}(h^j)\big\rangle_{\widetilde U_j^*,\widetilde U_j}\Big] \\ &+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_f^i(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}\Big], \end{aligned} \tag{5.32}$$

and the derivative is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. We proceed as above and split the supremum in (5.31) into the terms $\sum_{i\in[N]} \theta_i(u)$ and

$$\sup_{v\in\mathcal{X}} \sum_{i\in[N]} \big[-\theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] =: f(u,v).$$

The mapping $u \mapsto \sum_{i\in[N]} \theta_i(u)$ is continuously differentiable with

$$\Big\langle\Big[\sum_{i\in[N]} \theta_i(u)\Big]_u, h\Big\rangle_{U^*,U} = \sum_{i\in[N]} \Big[\big\langle(\tilde\theta_i)_{\tilde u}(\iota_{\widetilde U}(u)), \iota_{\widetilde U}(h)\big\rangle_{\widetilde U^*,\widetilde U}\Big] + \gamma\langle J_U(u), h\rangle_{U^*,U}.$$

Furthermore, we apply Danskin's theorem, see Theorem 2.75, to achieve the continuous differentiability of $u \mapsto \max_{v\in\mathcal{X}} f(u,v)$. To this end, we define $f^1 : U \times U \to \mathbb{R}$ and $f^2 : U \to \mathbb{R}$ by

$$\begin{aligned} f^1(u,v) &= -\sum_{i\in[N]} [\tilde\theta_i(\iota_{\widetilde U}(v^i, u^{-i}))] - \frac{\alpha}{2}\|\iota_H(v) - \iota_H(u)\|_H^2, \\ f^2(v) &= -\frac{\gamma}{2}\|v\|_U^2. \end{aligned}$$

By an analogous reasoning as in the proof of Theorem 5.6, the functional $f : U \times U \to \mathbb{R}$, $f(u,v) = f^1(u,v) + f^2(v)$, is continuous in u and upper semicontinuous in both arguments with respect to the weak sequential topology. Additionally, f is differentiable in u with continuous derivative $f_u = (f^1)_u$ from the weak sequential topology of $U \times U$ to the norm topology of U^*. Hence, the Assumption 2.66 is fulfilled and the theorem of Danskin, Theorem 2.75, is applicable. For $\alpha \geq 0$ chosen large enough to ensure that the solution map of $\max_{v\in\mathcal{X}} f(u,v)$ is single-valued on some open set A, one obtains the continuous differentiability of

$$u \mapsto \max_{v\in\mathcal{X}} \sum_{i\in[N]} \big[-\theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big]$$

and its derivative is continuous from the weak sequential topology of U to the norm topology of U^*. Thus, V_α is continuously differentiable on A and it holds (5.32) with the solution map $u \mapsto v_f(u)$. □

5.4.2 Application to Localized Nikaido–Isoda Merit Functionals

In this subsection, we consider the differentiability of regularized and localized Nikaido–Isoda merit functionals $\widetilde{V}_\alpha^{\mathrm{loc}}$, $\widetilde{V}_\alpha^{p_k}$, $\widetilde{V}_\alpha^{\mathrm{lp}}$, V_α^{loc}, $V_\alpha^{p_k}$, and V_α^{lp}, given in (5.3)–(5.8). We recall the definitions of these merit functionals

$$\widetilde{V}_\alpha^{\mathrm{loc}}(u) = \sum_{i\in[N]} \sup_{v^i \in \widetilde{F}_i(u)} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big],$$
$$\widetilde{V}_\alpha^{p_k}(u) = \sum_{i\in[N]} \sup_{v^i \in \mathcal{X}_i} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(\|v^i - u^i\|_{U_i}^2)\big],$$
$$\widetilde{V}_\alpha^{\mathrm{lp}}(u) = \sum_{i\in[N]} \sup_{v^i \in \widehat{F}_i(u)} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i)\big],$$

and

$$V_\alpha^{\mathrm{loc}}(u) = \sup_{v \in \widetilde{\mathcal{X}}(u)} \sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big],$$
$$V_\alpha^{p_k}(u) = \sup_{v \in \mathcal{X}} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] - \rho_k p(\|v - u\|_U^2)\Big],$$
$$V_\alpha^{\mathrm{lp}}(u) = \sup_{v \in \widehat{\mathcal{X}}(u)} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] - q(v - u)\Big].$$

We would like to apply the theory of Danskin's theorem, see Section 2.4, to these kinds of regularized and localized Nikaido–Isoda merit functionals. Considering the various merit functionals, we notice that we have two different-types of feasible sets. In fact, the constraint is either dependent on u itself or independent of any variable. For $\widetilde{V}_\alpha^{p_k}$ and $V_\alpha^{p_k}$, we can apply the generalized Danskin theorem, see Theorem 2.75, directly since the feasible sets are fixed. In the case that the constraints are dependent on u^i and u, i.e., for $\widetilde{V}_\alpha^{\mathrm{loc}}$, $\widetilde{V}_\alpha^{\mathrm{lp}}$, V_α^{loc} and V_α^{lp}, we have to find another way to differentiate and apply Danskin's theorem.

We begin with the part that is more straight-forward and investigate the differentiability of the merit functionals $\widetilde{V}_\alpha^{p_k}$ and $V_\alpha^{p_k}$ in the next two theorems.

Theorem 5.9. For the penalty-type or barrier-type function $p : [0, \infty) \to [0, \infty]$ we assume the following for each $i \in [N]$:

- $(u^i, v^i) \mapsto p(\|v^i - u^i\|_{U_i}^2)$ is continuous with respect to the weak sequential topology of $U_i \times U_i$.
- $(u^i, v^i) \mapsto (p(\|v^i - u^i\|_{U_i}^2))_{u_i}$ is continuous from the weak sequential topology of $U_i \times U_i$ to the norm topology of U_i^*.

Let α be nonnegative and the nonempty, open set $A \subseteq U$ be chosen such that the solution mapping $u \mapsto v_{f_i}(u)$, $i \in [N]$, to the problem $\max_{v^i \in \mathcal{X}_i} f_i(u, v^i)$ with $u \in A$ is single-valued where the objective functional $f_i : U \times U_i \to \mathbb{R}$ is given by

$$f_i(u, v^i) = -\theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(\|v^i - u^i\|_{U_i}^2).$$

Then $\widetilde{V}_\alpha^{p_k}$, defined in (5.5), is continuously differentiable on A with the derivative

$$\begin{aligned} d\widetilde{V}_\alpha^{p_k}(u)h = \sum_{i\in[N]} \Big[&\big\langle (\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}} + \gamma\langle J_{U_i}(u^i), h^i\rangle_{U_i^*,U_i} \\ &- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_{f_i}(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j)\big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j}\Big] \\ &+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_{f_i}(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i} \\ &+ 2\rho_k^i p'(\|v_{f_i}(u) - u^i\|_{U_i}^2)\big\langle J_{U_i}(v_{f_i}(u) - u^i), h^i\big\rangle_{U_i^*,U_i}\Big], \end{aligned} \tag{5.33}$$

which is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. As in the proof of Theorem 5.6, we study the terms $\theta_i(u)$ and $\max_{v^i\in\mathcal{X}_i} f_i(u, v^i)$ of the sum separately. Here, $i \in [N]$ is arbitrary but fixed. We split $f_i : U \times U_i \to \mathbb{R}$ into the sum of the functionals $f_i^1 : U \times U_i \to \mathbb{R}$ and $f_i^2 : U_i \to \mathbb{R}$ with

$$\begin{aligned} f_i^1(u, v^i) &= -\tilde{\theta}_i(\iota_{\widetilde{U}}(v^i, u^{-i})) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - \rho_k^i p(\|v^i - u^i\|_{U_i}^2) \\ f_i^2(v^i) &= -\frac{\gamma}{2}\|v^i\|_{U_i}^2. \end{aligned}$$

Making use of the assumptions on the penalty-type or barrier-type function p, we conclude that $f_i(\cdot, v^i)$ is continuous for any $v^i \in U_i$ and f_i is upper semicontinuous in both arguments with respect to the weak sequential topology. Furthermore, the continuous derivative regarding u from the weak sequential topology of U to the norm topology of U of the functional f_i is given by $(f_i)_u = (f_i^1)_u$. Hence, Assumption 2.66 is satisfied and we are allowed to apply Danskin's theorem, see Theorem 2.75, to the map $u \mapsto \max_{v^i\in\mathcal{X}_i} f_i(u, v^i)$ with the setting $X = U$, $Z = U_i$, $Y = \mathcal{X}_i$, and τ_X, τ_Z as the corresponding weak sequential topologies. Consequently, we obtain its continuous differentiability from the weak sequential topology of $U \times U_i$ to the norm topology of U^* on the open set A on which the solution map $u \mapsto v_{f_i}(u)$ is single-valued. It follows

$$\begin{aligned} d_u\big[\max_{v^i\in\mathcal{X}_i} f_i(u, v^i)\big]h = &- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_{f_i}(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j)\big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j}\Big] \\ &+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_{f_i}(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i} \\ &+ 2\rho_k^i p'(\|v_{f_i}(u) - u^i\|_{U_i}^2)\big\langle J_{U_i}(v_{f_i}(u) - u^i), h^i\big\rangle_{U_i^*,U_i}, \end{aligned}$$

and thus, the functional $\widetilde{V}_\alpha^{p_k}$ is continuously differentiable from the weak sequential topology of U to the norm topology of U^* on the open set A with desired derivative (5.33). ☐

Having examined the differentiability of $\widetilde{V}_\alpha^{p_k}$, we now turn to $V_\alpha^{p_k}$. The distinction lies in whether the supremum or the sum is taken first.

Theorem 5.10. For the penalty-type or barrier-type function $p : [0,\infty) \to [0,\infty]$, we assume the following:

- $(u,v) \mapsto p(\|v-u\|_U^2)$ is continuous with respect to the weak sequential topology of $U \times U$.
- $(u,v) \mapsto (p(\|v-u\|_U^2))_u$ is continuous from the weak sequential topology of $U \times U$ to the norm topology of U^*.

Let α be nonnegative and $A \subseteq U$ be a nonempty, open set such that the solution mapping $u \mapsto v_f(u)$ to the problem $\max_{v\in\mathcal{X}} f(u,v)$ with $u \in A$ is single-valued where $f : U \times U \to \mathbb{R}$ is given by

$$f(u,v) = \sum_{i\in[N]} \big[-\theta_i(v^i,u^{-i}) - \frac{\alpha}{2}\|\iota_H(v)-\iota_H(u)\|_H^2\big] - \rho_k p(\|v-u\|_U^2).$$

Then $V_\alpha^{p_k}$, defined in (5.6), is continuously differentiable on A and its continuous derivative from the weak sequential topology of U to the norm topology of U^* reads as follows:

$$\begin{aligned} \mathrm{d}V_\alpha^{p_k}(u)h = \sum_{i\in[N]} \Big[&\big\langle (\tilde\theta_i)_{\tilde u}(\iota_{\widetilde U}(u)), \iota_{\widetilde U}(h)\big\rangle_{\widetilde U^*,\widetilde U} + \gamma\langle J_{U_i}(u^i), h^i\rangle_{U_i^*,U_i} \\ &- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle (\tilde\theta_i)_{\tilde u^j}(\iota_{\widetilde U}(v_f^i(u),u^{-i})), \iota_{\widetilde U_j}(h^j)\big\rangle_{\widetilde U_j^*,\widetilde U_j}\Big] \\ &+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_f^i(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}\Big] \\ &+ 2\rho_k p'(\|v_f(u)-u\|_U^2)\big\langle J_U(v_f(u)-u), h\big\rangle_{U^*,U}. \end{aligned} \tag{5.34}$$

Proof. We apply Danskin's theorem, see Theorem 2.75, to the second term of the regularized and localized Nikaido–Isoda merit functional

$$V_\alpha^{p_k}(u) = \sum_{i\in[N]} [\theta_i(u)] + \max_{v\in\mathcal{X}} f(u,v),$$

where the objective functional $f : U \times U \to \mathbb{R}$ is defined by

$$f(u,v) = \sum_{i\in[N]} \big[-\theta_i(v^i,u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i)-\iota_{H_i}(u^i)\|_{H_i}^2\big] - \rho_k p(\|v-u\|_U^2).$$

Analogously to the proof of Theorem 5.9, we write f as the sum of the two functionals $f^1 : U\times U \to \mathbb{R}$ and $f^2 : U \to \mathbb{R}$ with

$$\begin{aligned} f^1(u,v) &= -\sum_{i\in[N]} \big[\tilde\theta_i(\iota_{\widetilde U}(v^i,u^{-i}))\big] - \frac{\alpha}{2}\|\iota_H(v)-\iota_H(u)\|_H^2 - \rho_k p(\|v-u\|_U^2), \\ f^2(v) &= -\frac{\gamma}{2}\|v\|_U^2. \end{aligned}$$

We observe that Assumption 2.66 is fulfilled and by an application of Danskin's theorem, see Theorem 2.75, it implies the continuous differentiability of $V_\alpha^{p_k}$ from the weak sequential topology of U to the norm topology of U^* on the set A. Moreover, it yields the derivative

$$\begin{aligned}
\mathrm{d}V_\alpha^{p_k}(u)h = \sum_{i\in[N]} \Big[&\big\langle(\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}} + \gamma\langle J_{U_i}(u^i), h^i\rangle_{U_i^*,U_i} \\
&- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle(\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_f^i(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j)\big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j}\Big] \\
&+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_f^i(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}\Big] \\
&+ 2\rho_k p'(\|v_f(u) - u\|_U^2)\big\langle J_U(v_f(u) - u), h\big\rangle_{U^*,U}. \qquad \square
\end{aligned}$$

Next, we move on to the regularized and localized Nikaido–Isoda merit functionals whose feasible sets are dependent on u itself. First, we look at the merit functionals $\widetilde{V}_\alpha^{\mathrm{lp}}$ and V_α^{lp}. In general, the u-dependent constraints $v^i \in \widehat{F}_i(u)$ and $v \in \widehat{\mathcal{X}}(u)$ would cause nonsmoothness of $\widetilde{V}_\alpha^{\mathrm{lp}}$ and V_α^{lp}, respectively. This is the reason for introducing a C^1-penalty or barrier term, denoted by q_i and q, to enforce locality without ruining differentiability. To apply Danskin's theorem, we require additional assumptions on the sets $\widehat{F}_i(u)$ and $\widehat{\mathcal{X}}(u)$ and on the functionals $q_i : U_i \to \mathbb{R}$ and $q : U \to \mathbb{R}$.

First, we consider the merit functional $\widetilde{V}_\alpha^{\mathrm{lp}}$, defined in (5.7), and prove its continuous differentiability under some additional assumptions.

Theorem 5.11. For the penalty-type or barrier-type functionals $q_i : U_i \to [0,\infty]$, $i \in [N]$, we assume the following:

- q_i is convex with $q_i(0) = 0$ and continuously differentiable in $\mathrm{int}(\mathrm{dom}(q_i))$.
- q_i and $q_i' : U_i \to U_i^*$ are completely continuous in $\mathrm{int}(\mathrm{dom}(q_i))$.

Let $A \subseteq U$ be nonempty, open and $C_i \subseteq U_i$, $i \in [N]$, be convex and closed sets with $0 \in C_i$. Moreover, we assume that there exist the radii $r, R > 0$ with

$$\begin{aligned}
\overline{B}_R^{U_i}(0) \cup \mathrm{cl}(C_i + B_r^{U_i}(0)) &\subseteq \mathrm{int}(\mathrm{dom}(q_i)), \\
B_r(u) &\subseteq A \qquad \forall\, u \in A, \\
F_i(u^{-i}) \cap \overline{B}_R(u^i) \subseteq \widehat{F}_i(u) &\subseteq F_i(u^{-i}) \qquad \forall\, u \in A.
\end{aligned} \tag{5.35}$$

The set A is chosen such that the solution map $u \mapsto v_{f_i}(u)$ to $\max_{v^i\in\widehat{F}_i(u)} f_i(u, v^i)$ with $u \in A$ is single-valued where $f_i : U \times U_i \to \mathbb{R}$ is defined by

$$f_i(u, v^i) = -\theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i). \tag{5.36}$$

If it holds $v_{f_i}(u) \in u^i + C_i$ for all $u \in A$, then the functional $\widetilde{V}_\alpha^{\mathrm{lp}}$ is differentiable on the set

A with the derivative

$$\begin{aligned}\mathrm{d}\widetilde{V}_\alpha^{\mathrm{lp}}(u)h = \sum_{i\in[N]} \Big[&\big\langle(\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}} + \gamma\big\langle J_{U_i}(u^i), h^i\big\rangle_{U_i^*,U_i} \\ &- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle(\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_{f_i}(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j)\big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j}\Big] + \big\langle q_i'(v_{f_i}(u)-u^i), h^i\big\rangle_{U_i^*,U_i} \\ &+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_{f_i}(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}\Big],\end{aligned} \tag{5.37}$$

which is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. Let $i \in [N]$ be arbitrary. The i-th summand of $\widetilde{V}_\alpha^{\mathrm{lp}}$ reads

$$\sup_{v^i\in\widehat{F}_i(u)} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i)\big].$$

Since $\theta_i(u)$ is independent of v^i, we can pull it out of the supremum and consider the merit functional

$$\widetilde{V}_\alpha^{\mathrm{lp}}(u) = \sum_{i\in[N]} \theta_i(u) + \sum_{i\in[N]} \sup_{v^i\in\widehat{F}_i(u)} f_i(u, v^i),$$

where we already inserted the definition of the functional $f_i : U \times U_i \to \mathbb{R}$, see (5.36), in the second sum. In the following, we investigate the differentiability of both sums separately. The first sum is continuously differentiable with

$$\Big\langle\Big[\sum_{i\in[N]} \theta_i(u)\Big]_u, h\Big\rangle_{U^*,U} = \sum_{i\in[N]} \Big[\big\langle(\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}}\Big] + \gamma\langle J_U(u), h\rangle_{U^*,U}$$

and its derivative is continuous from the weak sequential topology of U to the norm topology of U^*. In the case of the second sum, we consider the i-th summand $u \mapsto \max_{v^i\in\widehat{F}_i(u)} f_i(u, v^i)$ for an arbitrary $i \in [N]$ and we will apply Danskin's theorem, see Theorem 2.75, to achieve its continuous differentiability.

Let $\hat{u} \in A$ be arbitrarily fixed. We define the parameter $\rho = \frac{r}{2}$ and we introduce the convex, closed, and bounded set $Y_i = \mathcal{X}_i \cap \mathrm{cl}(C_i + B_\rho(\hat{u}^i))$. By assumption of the solution mapping v_{f_i}, we obtain

$$v_{f_i}(u) \in u^i + C_i \subseteq B_\rho(\hat{u}^i) + C_i \qquad \forall\, u \in B_\rho(\hat{u}).$$

Since it holds $v_{f_i}(u) \in \mathcal{X}_i$ anyway, we can conclude that $v_{f_i}(u) \in Y_i$ for all $u \in B_\rho(\hat{u})$. Hence, we obtain the equality of the maximization problems

$$\max_{v^i\in\widehat{F}_i(u)} f_i(u, v^i) = \max_{v^i\in Y_i} f_i(u, v^i) \qquad \forall\, u \in B_\rho(\hat{u}),$$

with the unique, single-valued solution operator $v_{f_i} : B_r(\hat{u}) \to Y_i$.

To achieve the continuous differentiability of $u \mapsto \max_{v^i \in Y_i} f_i(u, v^i)$ on the set $B_\rho(\hat{u})$, we are going to apply Danskin's theorem, see Theorem 2.75, with the choices

$$\begin{gathered} X = U, \qquad W = B_\rho(\hat{u}), \qquad Z = \hat{u}^i + \operatorname{dom}(q_i), \qquad Y = Y_i, \qquad f = f_i, \\ \tau_X = \text{weak sequential topology of } U, \\ \tau_Z = \text{topology induced by weak sequential topology of } U_i. \end{gathered} \tag{5.38}$$

Let us check whether this choice does indeed fulfill the assumptions of Danskin's theorem.

We note that $\hat{u}^i + \operatorname{dom}(q_i) \subseteq U_i$ inherits the weak sequential topology of U_i and $\mathcal{X}_i$ is compact in U_i with respect to the weak sequential topology by the assumption that we stated at the beginning of Section 5.4. Consequently, it yields that $Y_i \subseteq \mathcal{X}_i$ is compact in U_i regarding the weak sequential topology. Moreover, we note that it holds

$$Y_i \subseteq \operatorname{cl}(C_i + B_\rho(\hat{u}^i)) \subseteq \operatorname{cl}(C_i + B_\rho^{U_i}(0) + u^i),$$

and using the fact that it holds $\operatorname{cl}(Z_1 + Z_2) = \operatorname{cl}(Z_1) + \operatorname{cl}(Z_2)$ for Z_1 compact and Z_2 arbitrary, it yields

$$Y_i \subseteq \operatorname{cl}(C_i + B_\rho^{U_i}(0)) + u^i.$$

However, this is just a subset of $\operatorname{cl}(C_i + B_\rho^{U_i}(0)) \cup \overline{B}_R^{U_i}(0) + u^i$ and we see that we are in the setting to apply assumption (5.35) on the radius R. Consequently, we obtain $Y_i \subseteq \operatorname{dom}(q_i) + \hat{u}^i$, which implies the compactness of Y_i in $Z = \hat{u}^i + \operatorname{dom}(q_i)$ with respect to the weak sequential topology.

For any $(u, v^i) \in B_\rho(\hat{u}) \times Y_i$ it holds $v^i - u^i \in v^i - \hat{u}^i + B_\rho^{U_i}(0)$. Furthermore, we make use of the definition of Y_i to conclude $v^i \in \operatorname{cl}(C_i + B_\rho(\hat{u}^i))$, which implies

$$v^i - u^i \in \operatorname{cl}(C_i + B_\rho(\hat{u}^i)) - \hat{u}^i + B_\rho^{U_i}(0).$$

Next, we employ the fact that $\operatorname{cl}(Z_1) + \operatorname{cl}(Z_2) \subseteq \operatorname{cl}(Z_1 + Z_2)$ for any two sets Z_1 and Z_2, and conclude that $v^i - u^i \in \operatorname{cl}(C_i + B_r^{U_i}(0))$ due to $r = 2\rho$. However, the set on the right-hand side is a subset of the interior of the domain of q_i, see (5.35), and consequently, we achieve

$$v^i - u^i \in \operatorname{int}(\operatorname{dom}(q_i)) \qquad \forall\, (u, v^i) \in B_\rho(\hat{u}) \times Y_i. \tag{5.39}$$

We split the functional f_i, see (5.36), into the functionals f_i^1 and f_i^2, which we define by

$$\begin{aligned} f_i^1(u, v^i) &= -\tilde{\theta}_i(\iota_{\widetilde{U}}(v^i, u^{-i})) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i), \\ f_i^2(v^i) &= -\frac{\gamma}{2}\|v^i\|_{U_i}^2. \end{aligned}$$

The functional f_i^1 is continuous with respect to the weak sequential topology in both its arguments in $B_\rho(\hat{u}) \times Y_i$ due to the assumption (B2c) on the functional $\theta_i : U \to \mathbb{R}$ and the complete continuity of q_i on $\operatorname{int}(\operatorname{dom}(q_i))$, compare the second assumption of this theorem and (5.39). Moreover, the norm operator is known to be lower semicontinuous regarding the weak sequential topology and thus, we conclude the upper semicontinuity of the functional f_i^2 on U with respect to the weak sequential topology. Consequently, it yields that the functional f_i is upper semicontinuous with respect to the weak sequential topology in both its variables on $B_\rho(\hat{u}) \times Y_i$ and $f_i(\cdot, v^i) : U \to \mathbb{R}$ is continuous regarding the weak sequential topology on $B_\rho(\hat{u})$ for any $v^i \in Y_i$. Lastly, the u-derivative of f_i is completely continuous on $B_\rho(\hat{u}) \times Y_i$ due to the complete continuity of q_i' on $\operatorname{int}(\operatorname{dom}(q_i))$, compare the second assumption in this theorem and (5.39).

Hence, we allowed to apply Danskin's theorem, see Theorem 2.75, with the choice of the spaces and topologies as stated in (5.38). Thus, we obtain the continuous differentiability of $u \mapsto \max_{v^i \in Y_i} f_i(u, v^i)$ on $B_\rho(\hat{u})$ with the derivative $(f_i)_u(u, v_{f_i}(u))$. Moreover, the functional $u \mapsto \max_{v^i \in Y_i} f_i(u, v^i)$ is continuously differentiable from the weak sequential topology of U to the norm topology of U^* on the whole set A since $\hat{u} \in A$ was chosen arbitrarily. This implies that

$$\widetilde{V}_\alpha^{\mathrm{lp}}(u) = \sum_{i \in [N]} \Big[\theta_i(u) + \max_{v^i \in Y_i} f_i(u, v^i)\Big]$$

is continuously differentiable with the desired formula for its derivative, see (5.37), by recalling Lemma 2.73 it holds

$$\Big\langle \big[\max_{v^i \in Y_i} f_i(u, v^i)\big]_u, h\Big\rangle_{U^*,U} = \big\langle (f_i^1)_u(u, v_{f_i}(u)), h\big\rangle_{U^*,U} \qquad \forall\, u \in A, \quad h \in U. \qquad \square$$

Next, we prove an analogous result for the differentiability of the merit functional V_α^{lp}, see (5.41). The proof is comparable to what we have just witnessed, and we will only highlight key differences and steps.

Theorem 5.12. For the penalty-type or barrier-type functional q we assume the following:

- $q : U \to [0, \infty]$ is convex with $q(0) = 0$ and continuously differentiable in $\operatorname{int}(\operatorname{dom}(q))$.
- $q : U \to [0, \infty]$ and $q' : U \to U^*$ are completely continuous in $\operatorname{int}(\operatorname{dom}(q))$.

Let $A \subseteq U$ be nonempty, open and $C \subseteq U$ be a convex and closed set with $0 \in C$. Moreover, we assume that there exist the radii $r, R > 0$ with

$$\begin{aligned} \overline{B}_R^U(0) \cup \operatorname{cl}(C + B_r^U(0)) &\subseteq \operatorname{int}(\operatorname{dom}(q)), \\ B_r(u) &\subseteq A \qquad && \forall\, u \in A, \\ \mathcal{X} \cap \overline{B}_R(u) &\subseteq \widehat{\mathcal{X}}(u) \qquad && \forall\, u \in A. \end{aligned}$$

The set A is chosen such that the solution map $u \mapsto v_f(u)$ to $\max_{v \in \widehat{\mathcal{X}}(u)} f(u, v)$ with $u \in A$ is single-valued where $f : U \times U \to \mathbb{R}$ is defined by

$$f(u, v) = -\sum_{i \in [N]} \Big[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) + \iota_{H_i}(u^i)\|_{H_i}^2\Big] - q(v - u). \tag{5.40}$$

If the solution mapping satisfies $v_f(u) \in u + C$ for all $u \in A$, then the functional V_α^{lp} is differentiable on the set A with the continuous derivative as follows:

$$
\begin{aligned}
\mathrm{d}V_\alpha^{\mathrm{lp}}(u)h = \sum_{i\in[N]} \Big[&\big\langle(\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}} + \gamma\big\langle J_{U_i}(u^i), h^i\big\rangle_{U_i^*,U_i} \\
&- \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle(\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_f^i(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j)\big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j}\Big] \\
&+ \alpha\big\langle J_{H_i}(\iota_{H_i}(v_f^i(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i)\big\rangle_{H_i^*,H_i}\Big] \\
&+ \big\langle q'(v_f(u) - u), h\big\rangle_{U^*,U}.
\end{aligned} \tag{5.41}
$$

Furthermore, the derivative is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. As in the proof of Theorem 5.11, we first split the supremum in the definition of V_α^{lp}. We note that the sum of $u \mapsto \theta_i(u)$ is independent of v and considering the definition of the functional $f : U \times U \to \mathbb{R}$, see (5.40), we investigate the mappings $u \mapsto \sum_{i\in[N]} \theta_i(u)$ and $u \mapsto \max_{v\in\widehat{\mathcal{X}}(u)} f(u, v)$ separately. The first mapping is continuously differentiable with the derivative

$$
\sum_{i\in[N]} \Big[\big\langle(\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h)\big\rangle_{\widetilde{U}^*,\widetilde{U}}\Big] + \gamma\langle J_U(u), h\rangle_{U^*,U},
$$

which is continuous from the weak sequential topology of U to the norm topology of U^*. In the case of the second mapping $u \mapsto \max_{v\in\widehat{\mathcal{X}}(u)} f(u, v)$, we will apply Danskin's theorem, see Theorem 2.75, after proving that the assumptions of the theorem are fulfilled.

Let $\hat{u} \in A$ be arbitrarily fixed. Furthermore, we define the parameter $\rho = \frac{r}{2}$ and the convex, closed, and bounded set $Y = \mathcal{X} \cap \mathrm{cl}(C + B_\rho(\hat{u}))$. Since it holds $v_f(u) \in B_\rho(\hat{u}) + C$ for any $u \in B_\rho(\hat{u})$ and $v_f(u) \in \mathcal{X}$, we conclude $v_f(u) \in Y$ for $u \in B_\rho(\hat{u})$ and

$$
\max_{v\in\widetilde{\mathcal{X}}(u)} f(u, v) = \max_{v\in Y} f(u, v)
$$

with the unique, single-valued solution operator $v_f : B_r(\hat{u}) \to Y$. Moreover, the mapping $u \mapsto \max_{v\in Y} f(u, v)$ is continuously differentiable on the set $B_\rho(\hat{u})$ due to Danskin's theorem, see Theorem 2.75. Indeed, we are in the setting

$$
\begin{gathered}
X = U, \qquad W = B_\rho(\hat{u}), \qquad Z = \hat{u} + \mathrm{dom}(q), \qquad Y = Y, \qquad f = f, \\
\tau_X = \text{weak sequential topology of } U, \\
\tau_Z = \text{topology induced by weak sequential topology of } U,
\end{gathered}
$$

and using a similar line of reasoning as in the proof of Theorem 5.11, we can conclude the compactness of Y in Z with respect to the weak sequential topology of U. We split the functional f, see (5.40) again for its definition, into the functionals f^1 and f^2, which are defined by

$$
\begin{aligned}
f^1(u, v) &= -\sum_{i\in[N]} \Big[\tilde{\theta}_i(\iota_{\widetilde{U}}(v^i, u^{-i})) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\Big] - q(v - u), \\
f^2(v) &= -\frac{\gamma}{2}\|v\|_U^2.
\end{aligned}
$$

We note that it holds $v - u \in \operatorname{int}(\operatorname{dom}(q))$ for any $(u, v) \in B_\rho(\hat{u}) \times Y$. As in the preceding proof, we obtain that f is upper semicontinuous with respect to the weak sequential topology in both its variables on $B_\rho(\hat{u}) \times Y$ and $f(\cdot, v) : U \to \mathbb{R}$ is continuous with respect to the weak sequential topology on $B_\rho(\hat{u})$ for any $v \in Y$. Moreover, the u-derivative of f is completely continuous on $B_\rho(\hat{u}) \times Y$ due to the assumptions of q'. Hence, the assumptions of Danskin's theorem, see Assumption 2.66, are fulfilled and we obtain the continuous differentiability of $u \mapsto \max_{v\in Y} f(u, v)$ on $B_\rho(\hat{u})$ with the derivative $(f^1)_u(u, v_f(u))$. Since $\hat{u} \in A$ was chosen arbitrarily, we achieve the continuous differentiability of $u \mapsto \max_{v\in Y} f(u, v)$ on the set A. This implies the continuous differentiability of V_α^{lp} as claimed in the theorem. □

Remark 5.13. Another way to differentiate the merit functional $\widetilde{V}_\alpha^{\mathrm{lp}}$ and V_α^{lp} is to apply Danskin's theorem directly. To this end, we have to consider the special case, where

$$\widehat{F}_i(u) = \mathcal{X}_i \cap \overline{B}_R(u^i), \qquad \widehat{\mathcal{X}}(u) = \mathcal{X} \cap \overline{B}_R(u), \qquad \forall\, u \in U,$$

with nonempty, convex, closed, and bounded sets $\mathcal{X}_i \subseteq U_i$ and $\mathcal{X} \subseteq U$. Moreover, we express the localization with the functionals $q_i : U_i \to \mathbb{R}$ and $q : U \to \mathbb{R}$. These functionals are chosen such that they encode these u-dependent feasible sets by setting $q_i(z^i) = \infty$ if $u^i + z^i \in U_i \backslash \widehat{\mathcal{X}}_i(u)$ or $q(z) = \infty$ if $u + z \in U \backslash \widehat{\mathcal{X}}(u)$, respectively. Then the merit functionals read

$$\widetilde{V}_\alpha^{\mathrm{lp}}(u) = \sum_{i\in[N]} \sup_{v^i\in\mathcal{X}_i} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 - q_i(v^i - u^i)\big],$$
$$V_\alpha^{\mathrm{lp}}(u) = \sup_{v\in\mathcal{X}} \Big[\sum_{i\in[N]} \big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\big] - q(v - u)\Big].$$

We are able to apply Danskin's theorem to these two merit functionals by considering similar assumptions as in the preceding theorems, e.g., see Theorem 5.9 and Theorem 5.10. Consequently, we obtain their continuous differentiability from the weak sequential topology of U to the norm topology of U^* and their derivatives can be computed in the same manner.

Now, we move to the differentiability of the merit functional $\widetilde{V}_\alpha^{\mathrm{loc}}$, see again (5.3) for its definition.

Theorem 5.14. Let $A \subseteq U$ be a nonempty and open set such that the solution map $u \mapsto v_{f_i}(u)$ to $\max_{v^i\in\widetilde{F}_i(u)} f_i(u, v^i)$ is single-valued for all $u \in A$ where the objective functional $f_i : U \times U_i \to \mathbb{R}$ is defined by

$$f_i(u, v^i) = -\theta_i(v^i, u^{-i}) - \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2. \tag{5.42}$$

Furthermore, let $C_i \subseteq U_i$, $i \in [N]$, be convex and closed sets with $0 \in C_i$. Moreover, we assume that there exists the radius $r > 0$ with $B_r(u) \in A$ for any $u \in A$. If it holds $v_{f_i}(u) \in u^i + C_i$ for all $u \in A$, then the functional $\widetilde{V}_\alpha^{\mathrm{loc}}$ is differentiable on the set A with the derivative

$$\begin{aligned} \mathrm{d}\widetilde{V}_\alpha^{\mathrm{loc}}(u)h = \sum_{i\in[N]} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h) \big\rangle_{\widetilde{U}^*,\widetilde{U}} + \gamma \big\langle J_{U_i}(u^i), h^i \big\rangle_{U_i^*,U_i} \\ - \sum_{j\in[N]\setminus\{i\}} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_{f_i}(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j) \big\rangle_{\widetilde{U}_j^*,\widetilde{U}_j} \Big] \\ + \alpha \big\langle J_{H_i}(\iota_{H_i}(v_{f_i}(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i) \big\rangle_{H_i^*,H_i} \Big], \end{aligned} \tag{5.43}$$

which is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. Let $i \in [N]$ be arbitrary but fixed. The i-th summand of $\widetilde{V}_\alpha^{\mathrm{loc}}$ reads

$$\sup_{v^i \in \widetilde{F}_i(u)} \Big[\theta_i(u) - \theta_i(v^i, u^{-i}) - \frac{\alpha}{2} \|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \Big].$$

Since $\theta_i(u)$ is independent of v^i, we pull it out of the supremum and consider the merit functional

$$\widetilde{V}_\alpha^{\mathrm{loc}}(u) = \sum_{i\in[N]} \theta_i(u) + \sum_{i\in[N]} \sup_{v^i\in\widetilde{F}_i(u)} f_i(u, v^i),$$

where we already inserted the definition of the functional $f_i : U \times U_i \to \mathbb{R}$, see (5.42), in the second sum. The first sum is continuously differentiable with

$$\Big\langle \Big[\sum_{i\in[N]} \theta_i(u) \Big]_u, h \Big\rangle_{U^*,U} = \sum_{i\in[N]} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h) \big\rangle_{\widetilde{U}^*,\widetilde{U}} \Big] + \gamma \langle J_U(u), h \rangle_{U^*,U}$$

and its derivative is continuous from the weak sequential topology of U to the norm topology of U^*. In the case of the second sum, we consider the i-th summand $u \mapsto \max_{v^i\in\widetilde{F}_i(u)} f_i(u, v^i)$ for an arbitrary $i \in [N]$ and we apply Danskin's theorem, see Theorem 2.75, to achieve its continuous differentiability.

We proceed similar to the preceding proofs to show that the assumptions to Danskin's theorem are satisfied. Indeed, we define the parameter $\rho = \frac{r}{2}$ and the convex, closed, and bounded set $Y_i = \mathcal{X}_i \cap \mathrm{cl}(C_i + B_\rho(\hat{u}))$ for an arbitrarily fixed $\hat{u} \in A$. We conclude $v_{f_i}(u) \in Y_i$ for all $u \in B_\rho(\hat{u})$ and obtain the equality of the maximization problems

$$\max_{v^i\in\widetilde{F}_i(u)} f_i(u, v^i) = \max_{v^i\in Y_i} f_i(u, v^i) \qquad \forall\, u \in B_\rho(\hat{u}),$$

with the unique, single-valued solution operator $v_{f_i} : B_r(\hat{u}) \to Y_i$. We will apply Danskin's theorem to achieve the continuous differentiability of $u \mapsto \max_{v^i\in Y_i} f_i(u, v^i)$ on the set $B_\rho(\hat{u})$.

We are in the setting

$$\begin{gathered} X = U, \qquad W = B_\rho(\hat{u}), \qquad Z = \hat{u}^i + U_i, \qquad Y = Y_i, \qquad f = f_i, \\ \tau_X = \text{weak sequential topology of } U, \\ \tau_Z = \text{topology induced by weak sequential topology of } U_i. \end{gathered} \tag{5.44}$$

Making use of similar arguments as in the preceding proofs, we conclude that Y_i is compactly embedded in $Z = \hat{u}^i + U_i$ with respect to the weak sequential topology. Furthermore, we observe that the functional f_i is upper semicontinuous regarding the weak sequential topology on $B_\rho(\hat{u}) \times Y_i$ and $f_i(\cdot, v^i) : U_i \to \mathbb{R}$ is continuous with respect to the weak sequential topology on $B_\rho(\hat{u})$ for any $v^i \in Y_i$. Lastly, the u-derivative of f_i is completely continuous on $B_\rho(\hat{u}) \times Y_i$. Here, we used the assumption (B2c) of the functional $\theta_i : U \to \mathbb{R}$.

Consequently, we are allowed to apply Danskin's theorem, see Theorem 2.75, and obtain the continuous differentiability of $u \mapsto \max_{v^i \in Y_i} f_i(u, v^i)$ on $B_\rho(\hat{u})$ with the derivative $(f_i)_u(u, v_{f_i}(u))$. Since $\hat{u} \in A$ was chosen arbitrarily, we can extend its continuous differentiability from the weak sequential topology of U to the norm topology of U^* to the whole set A. This implies that $\widetilde{V}_\alpha^{\text{loc}}(u)$ is continuously differentiable for all $u \in A$ with the desired formula for its derivative, see (5.43). □

Now, we move on to the differentiability of the merit functional V_α^{loc}, see again (5.4) for its definition. In contrast to the preceding theorem, the sum over $i \in [N]$ is inside the supremum over $v \in \widetilde{X}(u)$.

Theorem 5.15. Let $C \subseteq U$ be a convex and closed set with $0 \in C$. The nonempty and open set $A \subseteq U$ is chosen such that the solution map $u \mapsto v_f(u)$ to $\max_{v \in \widetilde{X}(u)} f(u, v)$ is single-valued for any $u \in A$ where the objective functional $f : U \times U \to \mathbb{R}$ is defined by

$$f(u, v) = -\sum_{i \in [N]} \Big[\theta_i(v^i, u^{-i}) + \frac{\alpha}{2}\|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2\Big]. \tag{5.45}$$

Moreover, we assume the existence of a radius $r > 0$ such that $B_r(u) \in A$ for any $u \in A$. If it holds $v_f(u) \in u + C$ for all $u \in A$, then the functional V_α^{loc} is differentiable on A with the derivative

$$\begin{aligned} \mathrm{d}V_\alpha^{\text{loc}}(u)h = \sum_{i \in [N]} \Big[&\big\langle (\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h) \big\rangle_{\widetilde{U}^*, \widetilde{U}} + \gamma \big\langle J_{U_i}(u^i), h^i \big\rangle_{U_i^*, U_i} \\ &- \sum_{j \in [N] \setminus \{i\}} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}^j}(\iota_{\widetilde{U}}(v_f^i(u), u^{-i})), \iota_{\widetilde{U}_j}(h^j) \big\rangle_{\widetilde{U}_j^*, \widetilde{U}_j} \Big] \\ &+ \alpha \big\langle J_{H_i}(\iota_{H_i}(v_f^i(u)) - \iota_{H_i}(u^i)), \iota_{H_i}(h^i) \big\rangle_{H_i^*, H_i} \Big], \end{aligned} \tag{5.46}$$

which is continuous from the weak sequential topology of U to the norm topology of U^*.

Proof. As in the proof of Theorem 5.14, we split the supremum in the definition of V_α^{loc}. We note that the sum of $u \mapsto \theta_i(u)$ is independent of v and considering the definition of the functional $f : U \times U \to \mathbb{R}$, see (5.45), we investigate the mappings $u \mapsto \sum_{i \in [N]} \theta_i(u)$ and $u \mapsto \max_{v \in \widetilde{\mathcal{X}}(u)} f(u, v)$ separately. The first mapping is continuously differentiable with the derivative

$$\sum_{i \in [N]} \Big[\big\langle (\tilde{\theta}_i)_{\tilde{u}}(\iota_{\widetilde{U}}(u)), \iota_{\widetilde{U}}(h) \big\rangle_{\widetilde{U}^*, \widetilde{U}} \Big] + \gamma \langle J_U(u), h \rangle_{U^*, U},$$

which is continuous from the weak sequential topology of U to the norm topology of U^*. In the case of the second mapping $u \mapsto \max_{v \in \widetilde{\mathcal{X}}(u)} f(u, v)$, we apply Danskin's theorem, see Theorem 2.75, after proving that the assumptions of the theorem are fulfilled.

As in the preceding proofs, we define the parameter $\rho = \frac{r}{2}$ and the convex, closed, and bounded set $Y = \mathcal{X} \cap \mathrm{cl}(C + B_\rho(\hat{u}))$ for an arbitrarily fixed element $\hat{u} \in A$. Since it holds $v_f(u) \in Y$ for any $u \in B_\rho(\hat{u})$, it yields

$$\max_{v \in \widetilde{\mathcal{X}}(u)} f(u, v) = \max_{v \in Y} f(u, v)$$

with the unique, single-valued solution operator $v_f : B_r(\hat{u}) \to Y$. We prove the differentiability $u \mapsto \max_{v \in Y} f(u, v)$ on the set $B_\rho(\hat{u})$ by applying Danskin's theorem to the setting

$$X = U, \qquad W = B_\rho(\hat{u}), \qquad Z = \hat{u} + U, \qquad Y = Y, \qquad f = f,$$
$$\tau_X = \text{weak sequential topology of } U,$$
$$\tau_Z = \text{topology induced by weak sequential topology of } U.$$

Analogously, we obtain the compactness of Y in Z with respect to τ_Z. Moreover, the objective functional f is upper semicontinuous with respect to the weak sequential topology in both variables on $B_\rho(\hat{u}) \times Y$ and $f(\cdot, v) : U \to \mathbb{R}$ is continuous with respect to the weak sequential topology on $B_\rho(\hat{u})$ for any $v \in Y$. Furthermore, its u-derivative is completely continuous on $B_\rho(\hat{u}) \times Y$. Thus, the assumptions of Danskin's theorem hold and we obtain the continuous differentiability of $u \mapsto \max_{v \in Y} f(u, v)$ on A with the derivative $f_u(u, v_f(u))$. Here, we took advantage of the fact that $\hat{u} \in A$ was arbitrarily chosen to expand the domain of differentiability. Finally, we conclude that V_α^{loc} is continuously differentiable. $\square$

Chapter 6

Algorithms and Methods for Computing Equilibria

We have already various treatises on Mechanics, but the plan of this one is entirely new. I intend to reduce the theory of this science, and the art of solving problems relating to it, to general formulae, the simple development of which provides all the equations necessary for the solution of each problem. I hope that the manner in which I have tried to attain this object will leave nothing to be desired. The methods that I explain require neither geometrical, nor mechanical, constructions or reasoning, but only algebraical operations in accordance with regular and uniform procedure. Those who love Analysis will see with pleasure that Mechanics has become a branch of it, and will be grateful to me for having thus extended its domain. (Joseph-Louis Lagrange)

In this chapter, numerical methods for equilibria are discussed. We have demonstrated the presence of fixed points of the solution map, which can be understood as normalized equilibria and variational equilibria, and we have investigated the relationship between the regularized and localized Nikaido–Isoda merit functional and various forms of equilibria. Using these characterizations of equilibria, we then build numerical approaches employing the regularized Nikaido–Isoda merit functional. This chapter is separated into three sections. Consider a GNEP with convex constraints first. In this instance, we define a globally convergent descent method for obtaining the roots of the regularized Nikaido–Isoda merit functional. Under some situations, these roots can be read as normalized equilibria, see Subsection 5.2.2. In Section 6.2, a nonconvex restriction is added to the GNEP. This type of nonconvex constraint can be used to define constraints arising from a partial differential equation. This nonconvex restriction is managed via a safeguarded augmented Lagrangian method. The augmented Lagrangian method's subproblems consist of appropriate VIs that embody first-order optimality criteria.

6.1 Descent Method for the Regularized Nikaido–Isoda Merit Functional

We consider a GNEP with convex constraints in this section. We are looking for normalized equilibria, so we employ the regularized Nikaido–Isoda merit functional, as described in Section Subsection 5.2.2. In Theorem 5.4, we demonstrated that acceptable candidates for a normalized equilibrium are the zeros of the regularized Nikaido–Isoda merit functional. The following is an examination of the minimization problem

$$\min_{u \in \mathcal{X}} V_\alpha(u), \tag{6.1}$$

whose solutions correspond to the zeros of the functional V_α. Here, we create a descent method for the regularized Nikaido–Isoda merit functional V_α and use the projected gradient

$$p(w) = w - P_\mathcal{X}(w - \nabla V_\alpha(w)) \tag{6.2}$$

as a measure of criticality to demonstrate its convergence. In reality, the optimality conditions of the first order for the minimization problem (6.1) is equivalent to the condition $p(u) = 0$. This theory derives from [59]. In the following, assume that (A8) of Assumption 3.1 is satisfied for the spaces U, H, and $\widetilde{U}$. Let $\mathcal{X} \subseteq U$ be a nonempty, convex, closed, and bounded subset of U. Furthermore, let (B2c) of Assumption 3.3 be valid for the objective functional $\theta_i : U \to \mathbb{R}$. Moreover, we suppose that the parameter $\alpha \geq 0$ is selected so that the solution map

$$u \mapsto \sup_{v \in \mathcal{X}} \sum_{i \in [N]} \big[-\theta_i(v^i, u^{-i}) - \frac{\alpha}{2} \|\iota_{H_i}(v^i) - \iota_{H_i}(u^i)\|_{H_i}^2 \big]$$

is single-valued. The functional $V_\alpha : U \to \mathbb{R}$ is therefore continuously differentiable according to Danskin's theorem, see Theorem 2.75.

In this case, the descent method follows a projected gradient, as described in (6.2). The formula

$$\nabla f(x) = R_U^{-1} f_x(x) \in U$$

with $h \in U$ defines the gradient of a continuously differentiable functional $f : U \to \mathbb{R}$, where R_U specifies the Riesz operator from U to U^* for a given Hilbert space U. We relate the gradient to the Fréchet derivative via

$$(\nabla f(x), h)_U = \langle f_x(x), h \rangle_{U^*,U} = \mathrm{d}f(x, h).$$

Using the projected gradient, which we defined in (6.2), we can now formulate the algorithm as follows:

Algorithm 6.1.

0. Choose $\tilde{u}_0 \in U$ and set $u_0 = P_\mathcal{X}(\tilde{u}_0)$.

For $k = 0, 1, 2, 3, \ldots$:

1. If u_k satisfies stopping criterion: STOP
2. Set $s_k = -\nabla V_\alpha(u_k)$.
3. Choose $\sigma_k \in \{1, \frac{1}{2}, \frac{1}{4}, \dots\}$ maximally such that the projected Armijo rule
$$V_\alpha(P_\mathcal{X}(u_k + \sigma_k s_k)) - V_\alpha(u_k) \leq -\frac{\mu}{\sigma_k} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U^2 \tag{6.3}$$
with $\mu \in (0, 1)$ is satisfied.
4. Set $u_{k+1} = P_\mathcal{X}(u_k + \sigma_k s_k)$.

First, we note that the Armijo condition's selection of the step size is well-defined, as the next proposition states. Often, we say that $\{u_k\}_{k\in\mathbb{N}}$ is generated by Algorithm 6.1 and indirectly include the corresponding parameter sequence with this phrase, i.e., $\{\sigma_k\}_{k\in\mathbb{N}}$ is simultaneously generated by the algorithm without mentioning it directly.

Proposition 6.2 (cf., [59, Lemma 2.5])**.** If the Nikaido–Isoda merit functional $V_\alpha : U \to \mathbb{R}$ is continuously differentiable in some neighborhood of the convex, closed set $\mathcal{X}$, then the projected Armijo rule (6.3) terminates successfully for any sequence $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ with $p(u_k) \neq 0$.

Next, we demonstrate the algorithm's convergence using the projected gradient as the criticality measure. In this regard, we require the particular properties of the projection operator P_X, which are outlined in Lemma 2.51 of Chapter 2.

Theorem 6.3. It holds $\liminf_{k\to\infty} \|p(u_k)\|_U = 0$.

Proof. We argue by contradiction and assume that it holds $\liminf_{k\to\infty} \|p(u_k)\|_U > 0$. Consequently, there exists the parameters $\varepsilon > 0$ and $K \in \mathbb{N}$ such that $\|p(u_k)\|_U \geq \varepsilon$ for all $k \geq K$.

First, we introduce the auxiliary function $\Phi : \mathbb{R} \to \mathbb{R}$ that we define by
$$\Phi(\sigma_k) = \frac{1}{\sigma_k} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U.$$
Since Φ is nonincreasing by Lemma 2.51, we trivially obtain $\Phi(\sigma_k) \geq \Phi(1)$. In other words, we conclude
$$\frac{1}{\sigma_k} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \geq \|P_\mathcal{X}(u_k - \nabla V_\alpha(u_k)) - u_k\|_U = \|p(u_k)\|_U. \tag{6.4}$$
We subtract $V_\alpha(u_l)$ from $V_\alpha(u_K)$ and use the technique of telescopic sums to achieve for all $l \geq K$
$$V_\alpha(u_K) - V_\alpha(u_l) = \sum_{k=K}^{l-1} \big[V_\alpha(u_k) - V_\alpha(u_{k+1})\big].$$

Then we can apply the projected Armijo rule (6.3) as follows

$$\begin{aligned}\sum_{k=K}^{l-1} \left[V_\alpha(u_k) - V_\alpha(u_{k+1})\right] &\geq \sum_{k=K}^{l-1} \frac{\mu}{\sigma_k} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U^2 \\ &= \sum_{k=K}^{l-1} \frac{\mu}{\sigma_k} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \|u_{k+1} - u_k\|_U,\end{aligned}$$

and by the inequality (6.4) on the monotonicity of Φ we can directly obtain for all $l \geq K$ the following estimates

$$\begin{aligned}V_\alpha(u_K) - V_\alpha(u_l) &\geq \sum_{k=K}^{l-1} \frac{\mu}{\sigma_k} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \|u_{k+1} - u_k\|_U \\ &\geq \mu \sum_{k=K}^{l-1} \|P_\mathcal{X}(u_k - \nabla V_\alpha(u_k)) - u_k\|_U \|u_{k+1} - u_k\|_U.\end{aligned}$$

By the assumption of the proof via contradiction, we know that it holds $\|p(u_k)\|_U \geq \varepsilon$. We insert this lower bound of p into the last inequality to achieve

$$V_\alpha(u_K) - V_\alpha(u_l) \geq \mu\varepsilon \sum_{k=K}^{l-1} \|u_{k+1} - u_k\|_U \geq 0,$$

for all $l \geq K$. Hence, $V_\alpha(u_k)$ is monotonically decreasing with respect to the sequence index k and consequently, we know that the l-limit of $V_\alpha(u_l)$ is bounded from below by the infimum of $V_\alpha(u)$ over all $u \in \mathcal{X}$. Moreover, $V_\alpha(u_k)$ is bounded from below by 0 in $\mathcal{X}$ and consequently, we obtain

$$\mu\varepsilon \sum_{k=K}^{\infty} \|u_{k+1} - u_k\|_U \leq V_\alpha(u_K) - \lim_{l\to\infty} V_\alpha(u_l) \leq V_\alpha(u_K) - \inf_{u\in\mathcal{X}} V_\alpha(u) < \infty$$

as $l \to \infty$. In other words, $\|u_{k+1} - u_k\|_U$ is summable and for all $m > l$, we achieve

$$\|u_m - u_l\|_U \leq \sum_{k=l}^{m-1} \|u_{k+1} - u_k\|_U \leq \sum_{k=l}^{\infty} \|u_{k+1} - u_k\|_U \to 0$$

as $l \to \infty$. We conclude that $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ is a Cauchy sequence and since $\mathcal{X} \subseteq U$ is closed, we obtain that its limit $\bar{u}$ is an element in $\mathcal{X}$.

Next up, we exploit again the inequality (6.4) on the nonincreasing character of Φ and use the assumption of the contradiction proof to obtain

$$\|u_{k+1} - u_k\|_U = \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \geq \sigma_k \|P_\mathcal{X}(u_k - \nabla V_\alpha(u_k)) - u_k\|_U \geq \sigma_k \varepsilon$$

for $k > K$. Since the left-hand side of the inequality converges to zero, we can conclude that it has to hold both $\sigma_k \to 0$ as $k \to \infty$ and

$$\lim_{k\to\infty} \|P_\mathcal{X}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U = 0. \tag{6.5}$$

Thus, there exists some value $\widetilde{K} \geq K$ such that $\sigma_k \leq \frac{1}{2}$ is valid for all $k \geq \widetilde{K}$. Together with Lemma 2.51, we conclude

$$\begin{aligned} \frac{1}{\sigma_k}\|P_{\mathcal{X}}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U &\geq \frac{1}{2\sigma_k}\|P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \\ &\geq \|P_{\mathcal{X}}(u_k - \nabla V_\alpha(u_k)) - u_k\|_U, \end{aligned} \tag{6.6}$$

and thus, we obtain as $k \to \infty$

$$\|P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \leq 2\|P_{\mathcal{X}}(u_k - \sigma_k \nabla V_\alpha(u_k)) - u_k\|_U \to 0.$$

Moreover, the value $2\sigma_k$ does not fulfill the Armijo condition (6.3) in the k-th step because σ_k in the Armijo condition is already chosen maximal. The negation of (6.3) provides us with the estimate

$$-\frac{\mu}{2\sigma_k}\|P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k\|_U^2 \leq V_\alpha(P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k))) - V_\alpha(u_k). \tag{6.7}$$

In the following, we derive an estimate for the right-hand side. We apply the mean value theorem, see Theorem 2.21, and obtain the existence of some element w_k in the set

$$\{(1-t)u_k + tP_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) : 0 \leq t \leq 1\} \tag{6.8}$$

such that it holds

$$V_\alpha(P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - V_\alpha(u_k) = (\nabla V_\alpha(w_k), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k)_U.$$

Furthermore, we add and subtract a term on the right-hand side to create a difference term $\nabla V_\alpha(w_k) - \nabla V_\alpha(u_k)$ in the scalar product. Indeed, we achieve the equality

$$\begin{aligned} V_\alpha(P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - V_\alpha(u_k) &= (\nabla V_\alpha(u_k), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k)_U \\ &\quad + (\nabla V_\alpha(w_k) - \nabla V_\alpha(u_k), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k)_U, \end{aligned} \tag{6.9}$$

and in the following, we further manipulate each of the present terms separately.

We multiply this equality by $-2\sigma_k$ and note that we can manipulate the first scalar product on the right-hand side of (6.9) by adding and subtracting $P_{\mathcal{X}}(u_k) = u_k$ in the scalar product. Following this procedure, we achieve

$$\begin{aligned} &-2\sigma_k(\nabla V_\alpha(u_k), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - u_k)_U \\ &\quad = (u_k - 2\sigma_k \nabla V_\alpha(u_k) - P_{\mathcal{X}}(u_k), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - P_{\mathcal{X}}(u_k))_U \\ &\quad = (P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - P_{\mathcal{X}}(u_k), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - P_{\mathcal{X}}(u_k))_U \\ &\qquad + (u_k - 2\sigma_k \nabla V_\alpha(u_k) - P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)), P_{\mathcal{X}}(u_k - 2\sigma_k \nabla V_\alpha(u_k)) - P_{\mathcal{X}}(u_k))_U. \end{aligned}$$

We recall Lemma 2.51, which provides the nonnegativitiy of the last scalar product on the right-hand side. Furthermore, we observe that the first scalar product on the right-hand side has the same terms in each component of the scalar product and therefore, it can be written as a squared norm in U. Concluding these observations, we obtain an estimate for the first term on the right-hand side of (6.9) in the form of

$$-2\sigma_k(\nabla V_\alpha(u_k), P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k)_U \geq \|P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - P_{\mathcal{X}}(u_k)\|_U^2.$$

We divide this estimate by $-2\sigma_k$ and insert it back into (6.9) to get the following result:

$$\begin{aligned}V_\alpha(P_{\mathcal{X}}(u_k-2\sigma_k\nabla V_\alpha(u_k))-V_\alpha(u_k) \leq &-\frac{1}{2\sigma_k}\|P_{\mathcal{X}}(u_k-2\sigma_k\nabla V_\alpha(u_k))-u_k\|_U^2\\ &+\|\nabla V_\alpha(w_k)-\nabla V_\alpha(u_k)\|_U\|P_{\mathcal{X}}(u_k-2\sigma_k\nabla V_\alpha(u_k))-u_k\|_U.\end{aligned}$$

We apply the lower bound (6.7) on the left-hand side of the inequality and add the negation of the first term on the right-hand side on both sides of the inequality to obtain

$$\frac{1-\mu}{2\sigma_k}\|P_{\mathcal{X}}(u_k-2\sigma_k\nabla V_\alpha(u_k))-u_k\|_U^2 \leq \|\nabla V_\alpha(w_k)-\nabla V_\alpha(u_k)\|_U\|P_{\mathcal{X}}(u_k-2\sigma_k\nabla V_\alpha(u_k))-u_k\|_U.$$

We already used the inequality (6.6) in this proof, and using it once more on the left-hand side, it yields

$$\begin{aligned}&(1-\mu)\|p^k\|_U\|P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k\|_U\\ &\qquad\leq \frac{1-\mu}{2\sigma_k}\|P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k\|_U\|P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k\|_U\\ &\qquad\leq \|\nabla V_\alpha(w_k) - \nabla V_\alpha(u_k)\|_U\|P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k\|_U.\end{aligned}$$

We divide this inequality by the norm of $P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k$ and thus, we get

$$(1-\mu)\varepsilon \leq \|\nabla V_\alpha(w_k)-\nabla V_\alpha(u_k)\|_U \leq \|\nabla V_\alpha(w_k)-\nabla V_\alpha(\bar{u})\|_U+\|\nabla V_\alpha(\bar{u})-\nabla V_\alpha(u_k)\|_U. \quad (6.10)$$

We recall that we have derived the convergences $\|P_X(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k\|_U \to 0$ and $u_k \to \bar{u}$ in $\mathcal{X}$ as $k \to \infty$ in (6.5). Together with the definition of $\{w_k\}_{k\in\mathbb{N}}$, see (6.8), we conclude

$$\|w_k - \bar{u}\|_U \leq \|u_k - \bar{u}\|_U + \|P_{\mathcal{X}}(u_k - 2\sigma_k\nabla V_\alpha(u_k)) - u_k\|_U \to 0,$$

as $k \to \infty$. Finally, taking the limit $k \to \infty$ in (6.10) and using the continuity of ∇V_α, we arrive at the contradiction

$$0 < (1-\mu)\varepsilon \leq 0. \qquad \square$$

The descent method can be formulated for any continuously differentiable functional $f : U \to \mathbb{R}$. In addition, it is sufficient to assume that the feasible set is convex and closed. Consequently, the boundedness of the feasible set is not needed for proving the convergence theorem, see as well Theorem 6.3. However, the descent technique for the merit functional V_α uses Danskin's theorem, see Section 2.4, which requires the assumption of a convex, closed, and bounded set.

6.2 Augmented Lagrangian Method

In this section, we move on to GNEPs with nonconvex constraints. In the following, we consider a N-players game where the i-th player's optimization problem reads

$$\min_{v^i \in U_i} \theta_i(v^i, u^{-i}) \qquad \text{s.t.} \qquad G(v^i, u^{-i}) \in K, \quad (v^i, u^{-i}) \in \mathcal{X}, \tag{6.11}$$

and study GNEPs that consist of these minimization problems. In contrast to Section 6.1 from before, we additionally consider the nonconvex constraints $G(v^i, u^{-i}) \in K$. Regarding the involved infinite-dimensional spaces, we assume that U and X are Banach spaces. In the optimization problem, the objective function $\theta_i : U \to \mathbb{R}$ is considered to be differentiable in its i-th component, and the constraints are assumed to be expressed by a completely continuous operator $G : U \to X$ with a completely continuous derivative. Furthermore, let $K \subseteq X$ be a convex and closed set, and let $\mathcal{X} \subseteq U$ also be convex and closed. We note that we make no assumptions about the convexity of the nonlinear constraints. Speaking from an application point of view, such nonconvex constraints can arise from a solution operator to a partial differential equation, for example.

In the section that follows, we propose an augmented Lagrangian approach and discuss its convergence with respect to the constraints. As stated in Subsection 2.3.3, the augmented Lagrangian technique necessitates a differentiable multiplier term. To obtain a differentiable multiplier term, it is necessary to select a space in which the squared distance functional for a convex and closed set is continuously differentiable. The canonical choice is a Hilbert space, but in this section we expand it to a uniformly smooth and uniformly convex Banach space Y. In this regard, we reformulate the problem by introducing an operator e that maps the original function space X to Y. In the following, we will concentrate on the GNEP corresponding to the reformed i-th player's problem

$$\min_{v^i \in U_i} \theta_i(v^i, u^{-i}) \qquad \text{s.t.} \qquad e(G(v^i, u^{-i})) \in K_Y, \quad (v^i, u^{-i}) \in \mathcal{X}. \tag{6.12}$$

Regarding the reformulated problem, let $e : X \to Y$ be a linear and bounded operator, and let $K_Y \subseteq Y$ be a convex and closed set such that $x \in K$ is equivalent to $e(x) \in K_Y$. The latter equivalence holds if and only if we are in the situation $e^{-1}(K_Y) = K$, see Lemma 3.2. Assumption wise, we require (A9) for the underlying spaces and (B1) for the objective functional.

We note that one could have also considered a GNEP with differing constraints G_i for each $i \in [N]$ and in the same manner differing $e_i : X_i \to Y_i$, K_i, K_{Y_i}, and $\mathcal{X}_i$. Then the optimization problem for the i-th player reads

$$\min_{v^i \in U_i} \theta_i(v^i, u^{-i}) \qquad \text{s.t.} \qquad G_i(v^i, u^{-i}) \in K_i, \quad v^i \in \mathcal{X}_i.$$

We address the nonconvex constraints in (6.12) in a suitable subproblem of the augmented Lagrangian method's algorithm. Here, instead of minimization problems, we examine subproblems comprised of suitable VIs. To approximately solve the VIs, we could employ the descent method presented in Section 6.1.

Due to the lack of convexity, convergence only occurs towards the respective GNEP's KKT points, i.e., quasi-Nash equilibria and variational equilibria. Consequently, the convergence theory examines the convergence to KKT points. Specifically, we demonstrate that the weak limit of the sequence created by the augmented Lagrangian method is a stationary point of the squared distance functional of the relevant constraints and a KKT point, i.e., a quasi-Nash equilibrium or a variational equilibrium for the corresponding optimization problems. Consequently, the demonstration of convergence consists of two phases: feasibility and optimality. These considerations are grounded in the works [27, 72].

For the proof of feasibility, we require an ERCQ assumption, see again Definition 2.62 for its definition. In general, we must be cautious about whatever problem is selected for the ERCQ. In this regard, we typically assume that a point satisfies the ERCQ with respect to the original problem (6.11) instead of the reformulated problem as stated in (6.12).

Before examining the augmented Lagrangian approach for quasi-Nash equilibria and variational equilibria, we elaborate on the assumptions of the sets K and K_Y and on the requirements of the duality mapping J_Y.

First, we note that K is a convex and closed set due to the condition $K = e^{-1}(K_Y)$. In fact, if we have $x, y \in K$ and $\mu \in [0, 1]$, then we obtain by the linearity of e

$$e(\mu x + (1-\mu)y) = \mu e(x) + (1-\mu)e(y) \in K_Y,$$

and thus, it follows that $\mu x + (1-\mu)y \in K$. If K_Y is closed, then its complement K_Y^c is open and consequently, $e^{-1}(K_Y^c) = (e^{-1}(K_Y))^c$ is again open by the continuity of the operator e. Thus, we obtain that $K = e^{-1}(K_Y)$ is closed. Overall, we have shown that K is a convex and closed set. Hence, in order to ensure the existence of such an operator, we should assume that K defines a convex and closed set. In any case, one has to assume that K_Y is a convex and closed set. This cannot be inferred from possible properties of K, unless one interprets $e^{-1} : Y \to X$ as an inverse operator.

In addition, the condition that K_Y is convex and closed is necessary for differentiating the squared distance functional $\operatorname{dist}^2(\cdot, K_Y)$ that we have examined in Proposition 2.58 with respect to its differentiability. For the formulation of first-order optimality conditions using the tangent cone, it is necessary for $\mathcal{X}$ to be convex and closed.

So far, we have assumed that Y is a uniformly smooth and uniformly convex Banach space up to this point. We might easily simplify this assumption by demanding only some duality map attributes. For the convergence theory, it is sufficient to suppose that Y is a Banach space in which the duality map J_Y exists, is continuous, bijective, bounded on bounded sets, and its inverse J_Y^{-1} is also bounded on bounded sets. Additionally, we stipulate that the projection is continuous. This is the case, for instance, for a Banach space Y that is uniformly smooth and uniformly convex, see Lemma 2.52, Proposition 2.54, Proposition 2.56, and Proposition 2.57.

6.2.1 Augmented Lagrangian Method for Quasi-Nash Equilibria

In this subsection, we develop an augmented Lagrangian method and focus on its convergence behavior to quasi-Nash equilibria. We assume that $\mathcal{X} \subseteq U$ admits the product structure $\mathcal{X} = \prod_{i\in[N]} \mathcal{X}_i$ with given nonempty, convex, and closed sets $\mathcal{X}_i \subseteq U_i$. Thus, the reformulated GNEP consisting of the problems (6.12) reads

$$\min_{v^i\in U_i} \theta_i(v^i, u^{-i}) \qquad \text{s.t.} \qquad v^i \in \mathcal{X}_i, \quad e(G(v^i, u^{-i})) \in K_Y, \quad i \in [N]. \tag{6.13}$$

The assumption of the product structure $\mathcal{X} = \prod_{i\in[N]} \mathcal{X}_i$ will be shown to be crucial for proving convergence results. In Remark 6.10, we specify in which steps this assumption is critical.

In Subsection 2.3.2, we have studied the Lagrangian $L^i_Y : U \times Y^* \to \mathbb{R}$, which is defined by

$$L^i_Y(v^i, u^{-i}, \tilde{\lambda}^i) = \theta_i(v^i, u^{-i}) + \langle \tilde{\lambda}^i, e(G(v^i, u^{-i}))\rangle_{Y^*,Y}. \tag{6.14}$$

We introduce its augmented version $L^i_{\rho^i} : U \times Y^* \to \mathbb{R}$ corresponding to the i-th optimization problem in (6.13) for each player $i \in [N]$ and for some parameter $\rho^i > 0$ by

$$\begin{aligned} L^i_{\rho^i}(v^i, u^{-i}, w^i) &= \theta_i(v^i, u^{-i}) + \frac{\rho^i}{2}\Big\|e(G(v^i, u^{-i})) + \tfrac{J_Y^{-1}(w^i)}{\rho^i} - P_{K_Y}\big(e(G(v^i, u^{-i})) + \tfrac{J_Y^{-1}(w^i)}{\rho^i}\big)\Big\|^2_Y \\ &= \theta_i(v^i, u^{-i}) + \frac{\rho^i}{2}\operatorname{dist}^2\big(e(G(v^i, u^{-i})) + \tfrac{J_Y^{-1}(w^i)}{\rho^i}, K_Y\big). \end{aligned}$$

Here, $w^i \in Y^*$ denotes a bounded and safeguarded version of the Lagrangian multiplier.

We note that one can view w^i as a Lagrangian multiplier estimate for the constraint $e(G(u)) \in K_Y$ and one can interpret $e^* w^i \in X^*$ as an estimate for the Lagrangian multiplier λ^i for the original constraint $G(u) \in K$. Indeed, instead of $G(v) \in K$ we consider the augmented Lagrangian term for $e(G(v)) \in K_Y$. Formally, the Lagrangian term for $G(v) \in K$ would be $\langle \lambda^i, G(v)\rangle_{X^*,X}$, while the Lagrangian term for $e(G(v)) \in K_Y$ would correspond to

$$\langle \tilde{\lambda}^i, e(G(v))\rangle_{Y^*,Y} = \langle e^*\tilde{\lambda}^i, G(v)\rangle_{X^*,X}.$$

Thus, $e^*\tilde{\lambda}^i$ can indeed be viewed as an approximation of λ^i. Since e^* is linear and bounded from Y^* to X^*, we conclude that $e^*(Y^*)$ is a subspace of X^*. Hence, we consider w^i to be an approximation of $\tilde{\lambda}^i$ and $e^*\tilde{\lambda}^i$ to be an approximation of λ^i, resulting in $e^* w^i$ being an approximation of λ^i.

In our setting, the norm $\|\cdot\|_Y$ is Fréchet differentiable on $Y\backslash\{0\}$ and its derivative is uniformly continuous on bounded subsets that do not contain some neighborhood of 0, see Proposition 2.57. Moreover, by Proposition 2.58 the squared distance functional $\operatorname{dist}^2(\cdot, A)$ is differentiable for any convex, closed set $A \subseteq Y$ and its derivative reads

$$\big\langle \operatorname{dist}^2_y(y, A), h\big\rangle_{Y^*,Y} = \big\langle (\|\cdot\|^2_Y)'(y - P_A(y)), h\big\rangle_{Y^*,Y} = 2\langle J_Y(y - P_A(y)), h\rangle_{Y^*,Y}.$$

Here, the projection $P_A : Y \to A$ is continuous on Y and uniformly continuous on any bounded subset of Y, see Lemma 2.52.

Next, we state the algorithm of the augmented Lagrangian approach:

Algorithm 6.4.

0. Choose a bounded set $B \subseteq Y^*$ and the parameters $\rho_0 \in \mathbb{R}^N_{>0}$, $\gamma > 1$ and $\tau \in (0,1)$ arbitrarily.

For $k = 0, 1, 2, 3, \ldots$:

1. If u_k satisfies stopping criterion: STOP

2. Choose $w^i_k \in B$, $i \in [N]$, and compute an approximate solution u_{k+1} to the system of VIs

$$u_{k+1} \in \mathcal{X}, \quad \Big\langle \big[(L^i_{\rho^i_k})_{v^i}(v^i, u^{-i}_{k+1}, w^i_k)\big]\big|_{v^i = u^i_{k+1}}, z^i - u^i_{k+1} \Big\rangle_{U^*_i, U_i} \geq 0 \quad \forall\, z^i \in \mathcal{X}_i, \quad i \in [N]. \tag{6.15}$$

3. Compute for any $i \in [N]$

$$r^i_{k+1} = \Big\| e(G(u_{k+1})) - P_{K_Y}\Big(e(G(u_{k+1})) + \tfrac{J_Y^{-1}(w^i_k)}{\rho^i_k}\Big) \Big\|_Y.$$

4. If $k = 0$, or $k \geq 1$ and $r^i_{k+1} \leq \tau r^i_k$, then set $\rho^i_{k+1} = \rho^i_k$. Otherwise, set $\rho^i_{k+1} = \gamma \rho^i_k$ for all $i \in [N]$.

In the second step of the algorithm, a typical selection for the bounded version of the Lagrangian multiplier is

$$w^i_k = P_B\Big(\rho^i_{k-1} J_Y\Big(e(G(u_k)) + \tfrac{J_Y^{-1}(w^i_{k-1})}{\rho^i_{k-1}} - P_{K_Y}\Big(e(G(u_k)) + \tfrac{J_Y^{-1}(w^i_{k-1})}{\rho^i_{k-1}}\Big)\Big)\Big), \qquad i \in [N],$$

see [72]. This choice is equal to $w^i_k = P_B(\tilde{\lambda}^i_k)$ for $\tilde{\lambda}^i_k$ as defined in (2.27). We note that one can simplify this selection by applying the same updates for all w^i_k and using only a single w_k and a single ρ_k for all $i \in [N]$. Furthermore, an approximate solution $u_{k+1} \in \mathcal{X}$ to the VIs (6.15) is characterized by the representation

$$\Big\langle \big[(L^i_{\rho^i_k})_{v^i}(v^i, u^{-i}_{k+1}, w^i_k)\big]\big|_{v^i = u^i_{k+1}}, z^i - u^i_{k+1} \Big\rangle_{U^*_i, U_i} \geq \langle \varepsilon^i_k, z^i - u^i_{k+1} \rangle_{U^*_i, U_i} \qquad \forall\, z^i \in \mathcal{X}_i, \ i \in [N]. \tag{6.16}$$

To demonstrate the convergence of the algorithm, we assume that a sequence $\{\varepsilon^i_k\}_{k \in \mathbb{N}} \subseteq U^*_i$ is chosen such that it holds $\varepsilon^i_k \to 0$ as $k \to \infty$ and assume that the computation of an approximate solution in the second step is well-posed.

The solution u_{k+1} to the system (6.15) cannot be regarded as a quasi-Nash equilibrium of the associated game consisting of (6.13) since the solution's feasibility cannot be guaranteed in general.

Often, we say that $\{u_k\}_{k \in \mathbb{N}}$ is generated by Algorithm 6.4 and indirectly include the corresponding parameter sequences with this phrase, i.e., $\{\rho^i_k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}^N_{>0}$, $\{r^i_k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}^N_{\geq 0}$, and $\{w^i_k\}_{k \in \mathbb{N}} \subseteq B$, $i \in [N]$, are simultaneously generated by the algorithm without mentioning them directly.

Next, we examine the convergence of the augmented Lagrangian approach, as described in Algorithm 6.4. Due to the nonconvexity of the constraints, we can only demonstrate the first-order optimality conditions and not the convergence to a Nash equilibrium but merely a quasi-Nash equilibrium. The convergence proof consists of two components: feasibility and optimality.

Lemma 6.5. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.4 and let the i-th derivative of θ_i be bounded on bounded sets. If $\{u_{k_l}\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ is a weakly convergent subsequence, then its weak limit $\bar{u}$ is a stationary point to

$$\min_{u\in U} \operatorname{dist}^2(e(G(u)), K_Y),$$

i.e., the first-order optimality condition holds at $\bar{u}$.

Proof. Let $i \in [N]$ be arbitrary. Since $\mathcal{X} \subseteq U$ is closed with respect to the weak sequential topology, it holds $\bar{u} \in \mathcal{X}$. Moreover, by Proposition 2.58, we can rewrite the derivative of the squared distance functional via the duality mapping J_Y as

$$\operatorname{dist}^2_v(v, K_Y) = 2J_Y\big(v - P_{K_Y}(v)\big) \qquad \forall\, v \in Y,$$

and thus, we obtain

$$\begin{aligned}&\Big\langle \operatorname{dist}^2_v\big(e(G(\cdot)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}}, K_Y\big)(v), h\Big\rangle_{U^*,U}\\ &\qquad = 2\Big\langle G_v(v)^* e^* J_Y\Big(e(G(v)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}} - P_{K_Y}\big(e(G(v)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}}\big)\Big), h\Big\rangle_{U^*,U},\end{aligned}$$

for all $v \in Y$ and $h \in U$. Considering the corresponding subsequences in this equality, we distinguish between two cases where the parameter sequence $\{\rho^i_{k_l-1}\}_{l\in\mathbb{N}}$ is either bounded or unbounded.

In the case of a bounded subsequence $\{\rho^i_{k_l-1}\}_{l\in\mathbb{N}}$, there exists some $m \in \mathbb{N}$ with $\rho^i_{k_l-1} = \rho^i_{k_m-1}$ for all $l \geq m$. By the fourth step of Algorithm 6.4, we obtain

$$r^i_{k_l-1} \leq \tau^{k_l-1-(k_{l-1}-1)} r^i_{k_{l-1}-1} = \tau^{k_l-k_{l-1}} r^i_{k_{l-1}-1} \qquad \text{for all } l \geq m.$$

Since it holds $P_{K_Y}\big(e(G(u_{k_{l+1}})) + \frac{J_Y^{-1}(w^i_{k_{l+1}-1})}{\rho^i_{k_{l+1}-1}}\big) \in K_Y$ by the definition of the projection operator, we can conclude the estimate

$$\operatorname{dist}\big(e(G(u_{k_{l+1}})), K_Y\big) \leq \Big\|e(G(u_{k_{l+1}})) - P_{K_Y}\big(e(G(u_{k_{l+1}})) + \tfrac{J_Y^{-1}(w^i_{k_{l+1}-1})}{\rho^i_{k_{l+1}-1}}\big)\Big\|_Y = r^i_{k_{l+1}}. \tag{6.17}$$

Combining the last two inequalities that we have derived, we obtain

$$0 \leq \operatorname{dist}\big(e(G(u_{k_{l+1}})), K_Y\big) \leq r^i_{k_{l+1}} \leq r^i_{k_{l+1}-1} \leq \tau^{k_{l+1}-k_m} r^i_{k_m-1} \to 0,$$

as $l \to \infty$. By the continuity of the functional

$$\operatorname{dist}^2(e(G(\cdot)), K_Y) = \big\|e(G(\cdot)) - P_{K_Y}\big(e(G(\cdot))\big)\big\|^2_Y$$

with respect to the weak sequential topology, we achieve that $\bar{u}$ is a root of this functional, i.e., it holds $\operatorname{dist}^2(e(G(\bar{u})), K_Y) = 0$. Hence, $\bar{u}$ is a global minimizer and a stationary point to the minimization problem over $\operatorname{dist}^2(e(G(\cdot)), K_Y)$.

Next, we are in the case where the subsequence of parameters $\{\rho^i_{k_l-1}\}_{l\in\mathbb{N}}$ is unbounded, i.e., it holds $\rho^i_{k_l-1} \to \infty$ as $l \to \infty$. We are going to show that it implies

$$\left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(\bar{u}), h^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, h^i \in T_{\mathcal{X}_i}(\bar{u}^i).$$

We argue by contradiction and assume that there exists an element $y^i \in \mathcal{X}_i$ with the property

$$\left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(\bar{u}), y^i - \bar{u}^i \right\rangle_{U_i^*,U_i} < 0. \tag{6.18}$$

Since $\{w^i_{k_l-1}\}_{l\in\mathbb{N}} \subseteq Y^*$ is bounded and $\rho^i_{k_l-1} \to \infty$ as $l \to \infty$ in the current case, it holds

$$\lim_{l\to\infty} \left[e(G(u_{k_{l+1}})) + \tfrac{J_Y^{-1}(w^i_{k_{l+1}-1})}{\rho^i_{k_{l+1}-1}} \right] = e(G(\bar{u})).$$

Thus, we obtain the convergence

$$\left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)) + \tfrac{J_Y^{-1}(w^i_{k_{l+1}-1})}{\rho^i_{k_{l+1}-1}}, K_Y)(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} \to \left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(\bar{u}), y^i - \bar{u}^i \right\rangle_{U_i^*,U_i},$$

as $l \to \infty$. Since the limit is negative by (6.18), it implies the existence of a positive constant $c_1 > 0$ with

$$\left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)) + \tfrac{J_Y^{-1}(w^i_{k_{l+1}-1})}{\rho^i_{k_{l+1}-1}}, K_Y)(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} < -c_1, \tag{6.19}$$

for sufficiently large l. We recall the definition of an approximate solution and the introduction of the sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ in (6.16). Since $u_{k_{l+1}}$ is an approximate solution to the system of VIs (6.15), we obtain

$$\begin{aligned}
\langle \varepsilon^i_{k_{l+1}-1}, y^i - u^i_{k_{l+1}} \rangle_{U_i^*,U_i} &\leq \left\langle \left[(L^i_{\rho^i_{k_{l+1}-1}})_{v^i}(v, w^i_{k_{l+1}-1}) \right]_{|v=u_{k_{l+1}}}, y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} \\
&= \left\langle (\theta_i)_{v^i}(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} \\
&\quad + \frac{\rho^i_{k_{l+1}-1}}{2} \left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)) + \tfrac{J_Y^{-1}(w^i_{k_{l+1}-1})}{\rho^i_{k_{l+1}-1}}, K_Y)(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i},
\end{aligned}$$

and inserting the estimate (6.19) into this inequality, we achieve

$$\langle \varepsilon^i_{k_{l+1}-1}, y^i - u^i_{k_{l+1}} \rangle_{U_i^*,U_i} < \left\langle (\theta_i)_{v^i}(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} - \frac{\rho^i_{k_{l+1}-1}}{2} c_1. \tag{6.20}$$

Since the derivative $(\theta_i)_{v^i}$ is assumed to be bounded on bounded sets, it implies the existence of a constant $c_2 \in \mathbb{R}$ such that it holds

$$\left\langle (\theta_i)_{v^i}(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} \leq c_2$$

for sufficiently large l. Finally, we insert this bound into (6.20) and arrive at the conclusion

$$\lim_{l\to\infty} \langle \varepsilon^i_{k_{l+1}-1}, y^i - u^i_{k_{l+1}} \rangle_{U_i^*,U_i} \leq c_2 - \lim_{l\to\infty} \frac{\rho^i_{k_{l+1}-1}}{2} c_1 = -\infty,$$

which contradicts $\varepsilon^i_{k_{l+1}-1} \to 0$ in U_i^* as $l \to \infty$. Altogether, we obtain

$$\big\langle \operatorname{dist}^2_v(e(G(\cdot)), K_Y)(\bar{u}), h \big\rangle_{U^*,U} = \sum_{i\in[N]} \big\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(\bar{u}), h^i \big\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, h \in T_{\mathcal{X}}(\bar{u}),$$

and thus, $\bar{u}$ is a stationary point of $\min_{u\in U} \operatorname{dist}^2(e(G(u)), K_Y)$. □

We have already assumed in the lemma that such a weakly convergent sequence of $\{u_k\}_{k\in\mathbb{N}}$ exists. We note that the existence of a weakly convergent subsequence is established if we assume that $\mathcal{X} \subseteq U$ is bounded and the Banach space U is reflexive. Indeed, if $\mathcal{X}$ is bounded, then Algorithm 6.4 creates a sequence $\{u_k\}_{k\in\mathbb{N}}$ that is bounded. In the case of a reflexive Banach space U, a combination of the Eberlein–Šmulian and Banach–Alaoglu theorems, see Lemma 2.31, implies the existence of a weakly convergent subsequence. We conclude that the weak limit point resides in $\mathcal{X}$ due to the closedness of $\mathcal{X}$.

Next, we compare [27, Lemma 5.2] and [72, Lemma 7.3] to determine the assumptions for the weak limit $\bar{u} \in \mathcal{X}$ to be a feasible point.

Proposition 6.6. Let $e(X)$ be dense in Y. If $u \in \mathcal{X}$ is a stationary point to the optimization problem $\min_{u\in U} \operatorname{dist}^2(e(G(u)), K_Y)$ and satisfies the ERCQ, then it holds $e(G(u)) \in K_Y$ and $G(u) \in K$.

Proof. Let $i \in [N]$ be arbitrary. We want to prove $G(u) \in e^{-1}(K_Y) = K$, i.e., $e(G(u)) = P_{K_Y}(e(G(u))) \in K_Y$. In other words, we want to show

$$e(G(u)) - P_{K_Y}(e(G(u))) = 0. \tag{6.21}$$

The ERCQ ensures by Definition 2.62 that there exists some radius $R > 0$ such that it holds

$$\overline{B}^X_R(0) \subseteq G(u) + G_{v^i}(u)(\mathcal{X}_i - u^i) - K.$$

We conclude that for all $x \in X$ with $\|x\|_X \leq R$ there exists elements $y \in K$ and $z^i \in \mathcal{X}_i$ with

$$x = G(u) + G_{v^i}(u)z^i - u^i - y. \tag{6.22}$$

We apply J_Y to the operator in (6.21). This results in a functional in Y^*, to which we can apply the element $e(x) \in Y$ with x having the representation (6.22). This procedure yields

$$\begin{aligned} &\Big\langle J_Y\big(e(G(u)) - P_{K_Y}(e(G(u)))\big), e(x)\Big\rangle_{Y^*,Y} \\ &\quad = \Big\langle J_Y\big(e(G(u)) - P_{K_Y}(e(G(u)))\big), e\big(G_{v^i}(u)(z^i - u^i)\big)\Big\rangle_{Y^*,Y} \\ &\quad\quad + \Big\langle J_Y\big(e(G(u)) - P_{K_Y}(e(G(u)))\big), e(G(u) - y)\Big\rangle_{Y^*,Y}. \end{aligned} \tag{6.23}$$

First off, we investigate the first term in (6.23). The stationarity assumption of u, the formula Proposition 2.58 for the derivative of the squared distance functional, and z^i being an element in $\mathcal{X}_i$ implies

$$\begin{aligned}\Big\langle J_Y\big(e(G(u)) - P_{K_Y}\big(e(G(u))\big)\big), e\big(G_{v^i}(u)(z^i - u^i)\big)\Big\rangle_{Y^*,Y} \\ = \frac{1}{2}\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(u), z^i - u^i\rangle_{U_i^*,U_i} \\ \geq 0.\end{aligned} \tag{6.24}$$

Next, we consider the case of the second term in the right-hand side of (6.23). For any $h_1 \in Y$ and $h_2 \in K_Y$ we make use of the following trick by adding and subtracting $P_{K_Y}(h_1)$ in the right side of the following dual product

$$\begin{aligned}\big\langle J_Y(h_1 - P_{K_Y}(h_1)), h_1 - h_2\big\rangle_{Y^*,Y} &= \big\langle J_Y(h_1 - P_{K_Y}(h_1)), h_1 - P_{K_Y}(h_1)\big\rangle_{Y^*,Y} \\ &\quad + \big\langle J_Y(h_1 - P_{K_Y}(h_1)), P_{K_Y}(h_1) - h_2\big\rangle_{Y^*,Y}.\end{aligned} \tag{6.25}$$

We can easily rewrite the first dual product on the right-hand side by Definition 2.53 of the duality mapping J_Y. Consequently, we will investigate the second one. Note that the element $P_{K_Y}(h_1)$ minimizes the functional $h \mapsto \frac{1}{2}\|h - h_1\|_Y^2$ on the set K_Y. Since the derivative of this functional is equal to $J_Y(h - h_1)$, the first-order optimality condition for this minimizer reads

$$\big\langle J_Y(P_{K_Y}(h_1) - h_1), h_2 - P_{K_Y}(h_1)\big\rangle_{Y^*,Y} \geq 0 \qquad \forall\, h_2 \in K_Y.$$

We insert this estimate into (6.25) and make use of Definition 2.53 to conclude

$$\big\langle J_Y(h_1 - P_{K_Y}(h_1)), h_1 - h_2\big\rangle_{Y^*,Y} \geq \|h_1 - P_{K_Y}(h_1)\|_Y^2 \geq 0. \tag{6.26}$$

Since $e(y)$ is an element of K_Y because of $y \in K$, we obtain for the second term in (6.23)

$$\Big\langle J_Y\big(e(G(u)) - P_{K_Y}\big(e(G(u))\big)\big), e(G(u)) - e(y)\Big\rangle_{Y^*,Y} \geq 0. \tag{6.27}$$

Altogether, we insert (6.24) and (6.27) into (6.23) and get

$$\Big\langle J_Y\big(e(G(u)) - P_{K_Y}\big(e(G(u))\big)\big), e(x)\Big\rangle_{Y^*,Y} \geq 0 \qquad \forall\, x \in \overline{B}_R^X(0).$$

Moreover, by virtue of scaling and the operator's linearity, this inequality holds for all $x \in X$. By density of $e(X)$ in Y and by continuity of e, it implies

$$J_Y\big(e(G(u)) - P_{K_Y}\big(e(G(u))\big)\big) = 0.$$

This condition is equivalent to

$$\Big\|J_Y\big(e(G(u)) - P_{K_Y}\big(e(G(u))\big)\big)\Big\|_{Y^*} = 0,$$

and since it holds $\|J_Y(h)\|_{Y^*} = \|h\|_Y$, see Definition 2.53, it yields the desired result (6.21).
□

In the following, we demonstrate the boundedness of the corresponding Lagrangian multipliers $\{\lambda_k^i\}_{k\in\mathbb{N}} \subseteq X^*$, $i \in [N]$ and prove that the weak-$*$ limit point forms a KKT point with $\bar{u}$ of the respective problems (6.11) under appropriate assumptions, see Proposition 6.7 and Proposition 6.8 below. Furthermore, if we assume that the ERCQ for the reformulated problem (6.13), $i \in [N]$, holds true, then we are able to prove the convergence to a KKT point $(\bar{u}, \tilde{\lambda})$ of the GNEP consisting of (6.13), $i \in [N]$, see Proposition 6.9 below. In this case, we can extract a weakly convergent subsequence of the bounded sequence $\{\tilde{\lambda}_k^i\}_{k\in\mathbb{N}} \subseteq Y^*$ since Y is reflexive due to the Milman–Pettis theorem, see Theorem 2.41 and Lemma 2.31.

Considering the proof's strategy of the next result, we apply the generalized open mapping theorem, see Theorem 2.35, to a suitable functional. We follow the approach in [26, Theorem 5.2] and use the generalized open mapping theorem on the functional

$$(x, y) \mapsto G(\bar{u}) + G_{u^i}(\bar{u})(x - \bar{u}^i) - y$$

on the domain $\mathcal{X}_i \times K$. However, we want to emphasize that this is not the only possibility and in the works [27, Theorem 5.5] and [72, Theorem 7.4], the open mapping theorem is applied to $x \mapsto G(\bar{u}) + G_{u^i}(\bar{u})x - K$ on the domain $\mathcal{X}_i - \bar{u}^i$.

Proposition 6.7. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.4 and let the i-th derivative of θ_i be bounded on bounded sets. We assume that $\{u_k\}_{k\in\mathbb{N}}$ has a weakly convergent subsequence $\{u_{k_l}\}_{l\in\mathbb{N}}$ in U. If the ERCQ holds in the weak limit $\bar{u}$, then $\{\lambda_{k_l}^i\}_{l\in\mathbb{N}}$ is bounded in X^*.

Proof. We show the boundedness of the Lagrangian multiplier $\{\lambda_{k_l}^i\}_{l\in\mathbb{N}}$ in X^*. To this end, we use the ERCQ and apply the generalized open mapping theorem to a suitable operator. During this procedure, we will also exploit the assumption that the operator G and its derivative are completely continuous.

First, we note that it holds $\bar{u} \in \mathcal{X}$ because $\mathcal{X} \subseteq U$ is closed with respect to the weak sequential topology. Since the ERCQ (2.6) holds in $\bar{u}$, we have by definition

$$0 \in \operatorname{int}\big(G(\bar{u}) + G_{v^i}(\bar{u})(\mathcal{X}_i - \bar{u}^i) - K\big). \tag{6.28}$$

Next, we will apply the generalized open mapping theorem, see Theorem 2.35, to some suitable operator in order to demonstrate the existence of some radius $R > 0$ with

$$\overline{B}_R^X(0) \subseteq G(\bar{u}) + G_{v^i}(\bar{u})((\mathcal{X}_i - \bar{u}^i) \cap \overline{B}_1^{U_i}(0)) - K. \tag{6.29}$$

To this end, we define the operator $\Psi : U_i \times X \to X$ by

$$\Psi(x, y) = G_{v^i}(\bar{u})(x - \bar{u}^i) - y + G(\bar{u}),$$

for any $x \in \mathcal{X}_i$ and $y \in K$. The graph of Ψ is given by the following set

$$\operatorname{graph}(\Psi) = \Big\{(x, y, z) \in U_i \times X \times X : x \in \mathcal{X}_i,\ y \in K,\ z = G_{v^i}(\bar{u})(x - \bar{u}^i) - y + G(\bar{u})\Big\}.$$

However, this is the intersection of the convex, closed set $\mathcal{X}_i \times K \times X$ and the linear and bounded operator's null space $(x, y, z) \mapsto z - G_{v^i}(\bar{u})(x - \bar{u}^i) + y - G(\bar{u})$. Consequently, the graph is convex and closed. Furthermore, the ERCQ implies the property $0 \in \text{int}(\text{range}\Psi)$ and we note that the point $(\bar{u}^i, G(\bar{u})) \in \mathcal{X}_i \times K$ is an element of $\Psi^{-1}(0)$ since $\Psi(\bar{u}^i, G(\bar{u})) = 0$ holds true. The generalized open mapping theorem, see Theorem 2.35, yields the following result

$$0 \in \text{int}\Psi\big(B_{R_1}^{U_i}(\bar{u}^i) \times B_{R_2}^{X}(G(\bar{u}))\big),$$

for some radii $R_1 > 0$ and $R_2 > 0$. Hence, there exists some radius $R > 0$ with

$$\overline{B}_R^X(0) \subseteq G_{v^i}(\bar{u})\big(\mathcal{X}_i \cap B_{R_1}^{U_i}(\bar{u}^i) - \bar{u}^i\big) - K \cap B_{R_2}^{X}(G(\bar{u})) + G(\bar{u}).$$

For $y \in (\mathcal{X}_i \cap B_R^{U_i}(\bar{u}^i) - \bar{u}^i)$, we get that y has the form $y = z - \bar{u}^i$ with some $z \in \mathcal{X}_i \cap B_R^{U_i}(\bar{u}^i)$. Moreover, we obtain $y \in B_{R_1}^{U_i}(0)$ and $y \in (\mathcal{X}_i - \bar{u}^i)$, thus, we arrive at $y \in (\mathcal{X}_i - \bar{u}^i) \cap B_{R_1}^{U_i}(0)$. For the second part, let $y \in (K \cap B_{R_2}^X(G(\bar{u})) - G(\bar{u})$. Then there exists an element $z \in K \cap B_{R_2}^X(G(\bar{u}))$ such that $y = z - G(\bar{u})$. Furthermore, it holds $z \in K$ and by the representation of y, that means $y \in K - G(\bar{u})$. Specifically, we choose the value $R_1 = 1$, and the generalized open mapping theorem yields the existence of some radius $R > 0$ with

$$\overline{B}_R^X(0) \subseteq G(\bar{u}) + G_{v^i}(\bar{u})\big((\mathcal{X}_i - \bar{u}^i) \cap \overline{B}_1^{U_i}(0)\big) - K,$$

as we have claimed in (6.29).

By the definition of the dual norm, we are able to choose a sequence $\{b_{k_l}^i\}_{l\in\mathbb{N}} \subseteq \overline{B}_R^X(0)$ that satisfies the properties $\|b_{k_l}^i\|_X = 1$ and $\langle \lambda_{k_l}^i, b_{k_l}^i \rangle_{X^*,X} \geq \frac{1}{2}\|\lambda_{k_l}^i\|_{X^*}$.

Consequently, there are elements $y_{k_l}^i \in K$ and $z_{k_l}^i \in \mathcal{X}_i$ with $\|z_{k_l}^i - \bar{u}^i\|_{U_i} \leq 1$ and

$$-Rb_{k_l}^i = G(\bar{u}) + G_{v^i}(\bar{u})(z_{k_l}^i - \bar{u}^i) - y_{k_l}^i.$$

Moreover, we obtain the estimate

$$\begin{aligned}
&\big\|G(u_{k_l}) + G_{v^i}(u_{k_l})(z_{k_l}^i - u_{k_l}^i) - \big(G(\bar{u})) + G_{v^i}(\bar{u})(z_{k_l}^i - \bar{u}^i)\big)\big\|_X \\
&\quad\leq \|G(u_{k_l}) - G(\bar{u})\|_X + \big\|G_{v^i}(u_{k_l}) - G_{v^i}(\bar{u})\big\|_{\mathscr{L}(U_i;X)}\|z_{k_l}^i - u_{k_l}^i\|_{U_i} \\
&\qquad + \big\|G_{v^i}(\bar{u})(\bar{u}^i - u_{k_l}^i)\big\|_X.
\end{aligned} \tag{6.30}$$

The first term on the right-hand side of (6.30) converges to zero since G is assumed to be completely continuous, i.e., it follows $G(u_{k_l}) \to G(\bar{u})$ in X as $l \to \infty$ due to the weak convergence $u_{k_l} \rightharpoonup \bar{u}$ in U as $l \to \infty$. Furthermore, Theorem 2.49 yields that the Fréchet derivative $G_{v^i}(\bar{u})$ of $G(\cdot)$ at $\bar{u}$ is again completely continuous and therefore, it holds

$$G_{v^i}(\bar{u})u_{k_l}^i \to G_{v^i}(\bar{u})\bar{u}^i,$$

as $l \to \infty$. Thus, the third term in (6.30) converges to zero. Moreover, by assumption we know that $G_{v^i} : U \to \mathscr{L}(U_i; X)$ is completely continuous and thus, we obtain the strong convergence $G_{v^i}(u_{k_l}) \to G_{v^i}(\bar{u})$ in $\mathscr{L}(U_i; X)$ as $l \to \infty$. In case of the second term in (6.30), we apply the boundedness of $\|z_{k_l}^i - u_{k_l}^i\|_{U_i}$, which follows from

$$\|z_{k_l}^i - u_{k_l}^i\|_{U_i} \leq \|z_{k_l}^i - \bar{u}^i\|_{U_i} + \|\bar{u}^i - u_{k_l}^i\|_{U_i} \leq 1 + \|\bar{u}^i - u_{k_l}^i\|_{U_i}$$

and the weak convergence of $\{u_{k_l}\}_{l\in\mathbb{N}}$. This shows the convergence of the second term towards zero. All together, we have as $l \to \infty$

$$\delta^i_{k_l} = \left\| Rb^i_{k_l} + G(u_{k_l}) + G_{v^i}(u_{k_l})(z^i_{k_l} - u^i_{k_l}) - y^i_{k_l} \right\|_X \to 0.$$

Furthermore, note that u_{k_l} is an approximate solution to the VIs (6.15). Then it holds

$$\begin{aligned}
\langle \varepsilon^i_{k_l-1}, z^i - u^i_{k_l} \rangle_{U_i^*,U_i} \leq{}& \left\langle (\theta_i)_{v^i}(u_{k_l}), z^i - u^i_{k_l} \right\rangle_{U_i^*,U_i} \\
&+ \rho^i_{k_l-1} \Big\langle e^* J_Y \Big(e(G(u_{k_l})) + \tfrac{J_Y^{-1}(w^i_{k_l-1})}{\rho^i_{k_l-1}} - P_{K_Y}\big(e(G(u_{k_l})) + \tfrac{J_Y^{-1}(w^i_{k_l-1})}{\rho^i_{k_l-1}} \big) \Big), \\
&\qquad G_{v^i}(u_{k_l})(z^i - u^i_{k_l}) \Big\rangle_{X^*,X} \\
={}& \left\langle (\theta_i)_{v^i}(u_{k_l}), z^i - u^i_{k_l} \right\rangle_{U_i^*,U_i} + \left\langle \lambda^i_{k_l}, G_{v^i}(u_{k_l})(z^i - u^i_{k_l}) \right\rangle_{X^*,X}
\end{aligned} \tag{6.31}$$

for all $z^i \in \mathcal{X}_i$ by definition of $\lambda^i_{k_l}$ via $\lambda^i_{k_l} = e^* \tilde{\lambda}^i_{k_l}$, see (2.27). Adding a zero we obtain

$$\begin{aligned}
\frac{R}{2} \|\lambda^i_{k_l}\|_{X^*} \leq{}& \langle \lambda^i_{k_l}, Rb^i_{k_l} \rangle_{X^*,X} \\
={}& \left\langle \lambda^i_{k_l}, Rb^i_{k_l} + G(u_{k_l}) + G_{v^i}(u_{k_l})(z^i_{k_l} - u^i_{k_l}) - y^i_{k_l} \right\rangle_{X^*,X} \\
&+ \left\langle \lambda^i_{k_l}, y^i_{k_l} - G(u_{k_l}) - G_{v^i}(u_{k_l})(z^i_{k_l} - u^i_{k_l}) \right\rangle_{X^*,X},
\end{aligned}$$

and inserting $z^i = z^i_{k_l}$ in (6.31), we can estimate

$$\begin{aligned}
\frac{R}{2} \|\lambda^i_{k_l}\|_{X^*} \leq{}& \delta^i_{k_l} \|\lambda^i_{k_l}\|_{X^*} + \langle \lambda^i_{k_l}, y^i_{k_l} - G(u_{k_l}) \rangle_{X^*,X} + \left\langle (\theta_i)_{v^i}(u_{k_l}), z^i_{k_l} - u^i_{k_l} \right\rangle_{U_i^*,U_i} \\
&- \langle \varepsilon^i_{k_l-1}, z^i_{k_l} - u^i_{k_l} \rangle_{U_i^*,U_i}.
\end{aligned}$$

We apply Lemma 2.65 to arrive at the inequality

$$\frac{R}{2} \|\lambda^i_{k_l}\|_{X^*} \leq \delta^i_{k_l} \|\lambda^i_{k_l}\|_{X^*} + \zeta^i_{k_l} + \left\langle (\theta_i)_{v^i}(u_{k_l}) - \varepsilon^i_{k_l-1}, z^i_{k_l} - u^i_{k_l} \right\rangle_{U_i^*,U_i} \tag{6.32}$$

We derive the following estimate for the last term in (6.32)

$$\left\langle (\theta_i)_{v^i}(u_{k_l}) - \varepsilon^i_{k_l-1}, z^i_{k_l} - u^i_{k_l} \right\rangle_{U_i^*,U_i} \leq \left\| (\theta_i)_{v^i}(u_{k_l}) - \varepsilon^i_{k_l-1} \right\|_{U_i^*} \|z^i_{k_l} - u^i_{k_l}\|_{U_i} \leq C,$$

where we used that the i-th derivative of θ_i is bounded on bounded sets and weakly convergent sequences are bounded. Here, $C > 0$ denotes some constant. We note that the sequences $\{\delta^i_{k_l}\}_{l\in\mathbb{N}}$ and $\{\zeta^i_{k_l}\}_{l\in\mathbb{N}}$ in the right-hand side of (6.32) converge to zero as $l \to \infty$ and we select l large enough so that it holds $|\delta^i_{k_l}| \leq \frac{R}{4}$ and $|\zeta^i_{k_l}| \leq C$. Altogether, we obtain $\frac{R}{4}\|\lambda^i_{k_l}\|_{X^*} \leq 2C$ by inserting the bounds into (6.32) and we conclude that $\{\lambda^i_{k_l}\}_{l\in\mathbb{N}}$ is bounded in X^*. □

In the next result, we are in the same situation of the previous proposition. Again, we assume that $\{u_k\}_{k\in\mathbb{N}}$ has a weakly convergent subsequence and the ERCQ holds in the weak limit. By the result from above, we know that the sequence of Lagrangian multipliers is bounded in X^*. In the next result, we further assume that the sequence of Lagrangian multipliers converges weakly-$*$ in X^* (e.g. if X is reflexive or separable) and $e(X)$ is dense in Y. We prove that the weak limit $\bar{u}$ is a quasi-Nash equilibrium.

Proposition 6.8. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.4 and the i-th derivative of θ_i be bounded on bounded sets and pseudomonotone. We assume that $\{u_k\}_{k\in\mathbb{N}}$ admits a weakly convergent subsequence in U, i.e., $u_{k_l} \rightharpoonup \bar{u}$ in U as $l \to \infty$, and the weak limit $\bar{u}$ fulfills the ERCQ. If there exist weakly-$*$ convergent subsequences of the bounded multipliers $\{\lambda^i_{k_l}\}_{l\in\mathbb{N}} \subseteq X^*$ with the limit point $\lambda^i \in X^*$ for all $i \in [N]$, then it holds

$$\begin{aligned} \left\langle \left[(L^i_X)_{v^i}(v^i, \bar{u}^{-i}, \lambda^i)\right]\big|_{v^i=\bar{u}^i}, z^i - \bar{u}^i \right\rangle_{U_i^*,U_i} &\geq 0 \qquad \forall\, z^i \in \mathcal{X}_i, \\ \langle \lambda^i, y - G(\bar{u})\rangle_{X^*,X} &\leq 0 \qquad \forall\, y \in K. \end{aligned} \tag{6.33}$$

If $e(X)$ is additionally dense in Y, then it yields $G(\bar{u}) \in K$ and $(\bar{u}, \lambda) \in U \times (X^*)^N$ satisfies the KKT conditions, i.e., $\bar{u}$ is a quasi-Nash equilibrium.

Proof. Let $\{\lambda^i_{k_{l_m}}\}_{m\in\mathbb{N}} \subseteq X^*$ be a weakly-$*$ convergent subsequence with $\lambda^i_{k_{l_m}} \overset{*}{\rightharpoonup} \lambda^i$ in X^* as $m \to \infty$. By Lemma 2.65, it holds for all $y \in X$

$$\langle \lambda^i, y - G(\bar{u})\rangle_{X^*,X} = \lim_{m\to\infty} \left[\langle \lambda^i_{k_{l_m}}, y - G(u_{k_{l_m}})\rangle_{X^*,X} - \zeta^i_{k_{l_m}}\right] \leq 0.$$

Thus, the second inequality of the KKT conditions is satisfied. In order to prove the first KKT condition, we show the convergence of both terms of the Lagrangian's derivative. Let $z^i \in \mathcal{X}_i$ be arbitrarily fixed. As in equation (6.30) we have as $m \to \infty$

$$\left\| G_{v^i}(u_{k_{l_m}})(z^i - u^i_{k_{l_m}}) - G_{v^i}(\bar{u})(z^i - \bar{u}^i) \right\|_X \to 0.$$

Since $u_{k_{l_m}}$ is an approximate solution to the VIs (6.15), it holds

$$\begin{aligned} \langle \varepsilon^i_{k_{l_m}-1}, z^i - u^i_{k_{l_m}}\rangle_{U_i^*,U_i} \leq{}& \left\langle (\theta_i)_{v^i}(u_{k_{l_m}}), z^i - u^i_{k_{l_m}} \right\rangle_{U_i^*,U_i} \\ &+ \left\langle \lambda^i_{k_{l_m}}, G_{v^i}(u_{k_{l_m}})(z^i - u^i_{k_{l_m}}) \right\rangle_{X^*,X} \end{aligned} \tag{6.34}$$

for all $z^i \in \mathcal{X}_i$. Thus, applying $\limsup_{m\to\infty}$ in (6.34) and using $\varepsilon^i_{k_{l_m}-1} \to 0$ in U_i^* as $m \to \infty$, it yields

$$\begin{aligned} 0 = \limsup_{m\to\infty} \langle \varepsilon^i_{k_{l_m}-1}, z^i - u^i_{k_{l_m}}\rangle_{U_i^*,U_i} \leq{}& \limsup_{m\to\infty} \left\langle (\theta_i)_{v^i}(u_{k_{l_m}}), z^i - u^i_{k_{l_m}} \right\rangle_{U_i^*,U_i} \\ &+ \left\langle \lambda^i, G_{v^i}(\bar{u})(z^i - \bar{u}^i) \right\rangle_{X^*,X}. \end{aligned} \tag{6.35}$$

For the first term on the right-hand side of (6.35), we exploit the pseudomonotonicity of $(\theta_i)_{v^i}$ for all $i \in [N]$. Similarly, considering $z^i = \bar{u}^i$ and applying the limit inferior as $m \to \infty$ in (6.34) gives

$$0 \leq \liminf_{m\to\infty} \left\langle (\theta_i)_{v^i}(u_{k_{l_m}}), \bar{u}^i - u^i_{k_{l_m}} \right\rangle_{U_i^*,U_i}. \tag{6.36}$$

The pseudomonotonicity of $(\theta_i)_{v^i}$ for any $i \in [N]$ and (6.36) yields

$$\limsup_{m\to\infty} \left\langle (\theta_i)_{v^i}(u_{k_{l_m}}), z^i - u^i_{k_{l_m}} \right\rangle_{U_i^*,U_i} \leq \left\langle (\theta_i)_{v^i}(\bar{u}), z^i - \bar{u}^i \right\rangle_{U_i^*,U_i}$$

for all $z^i \in \mathcal{X}_i$. Then we arrive at

$$\begin{aligned} 0 &\leq \big\langle (\theta_i)_{v^i}(\bar{u}), z^i - \bar{u}^i \big\rangle_{U_i^*,U_i} + \big\langle \lambda^i, G_{v^i}(\bar{u})(z^i - \bar{u}^i) \big\rangle_{X^*,X} \\ &= \Big\langle \big[(L_X^i)_{v^i}(v^i, \bar{u}^{-i}, \lambda^i) \big]_{|_{v^i=\bar{u}^i}}, z^i - \bar{u}^i \Big\rangle_{U_i^*,U_i}. \end{aligned}$$

The feasibility follows from Lemma 6.5 and Proposition 6.6 and thus, $(\bar{u}, \lambda) \in U \times (X^*)^N$ is a KKT pair. □

Next, we state an example for which the pseudomonotonicity of the i-th derivative of θ_i is satisfied as we have assumed in Proposition 6.8. Indeed, it is fulfilled if $\theta_i : U \to \mathbb{R}$ is of the form $\theta_i(u) = \tilde{\theta}_i(\iota_{\widetilde{U}}(u)) + \frac{\gamma}{2}\|u^i\|_{U_i}^2$ for a given reflexive Banach space U that admits a uniformly convex dual space U^*. Since $(\theta_i)_{u^i}$ is the sum of a completely continuous operator and a monotone, hemicontinuous operator, it is pseudomonotone by Proposition 2.48. Indeed, it holds

$$\big\langle (\theta_i)_{u^i}(u), z^i \big\rangle_{U_i^*,U_i} = \langle \iota_{\widetilde{U}_i}^*(\tilde{\theta}_i)_{\tilde{u}^i}(\iota_{\widetilde{U}}(u)), z^i \rangle_{U_i^*,U_i} + \gamma \langle J_{U_i}(u^i), z^i \rangle_{U_i^*,U_i}. \tag{6.37}$$

The first operator is completely continuous because $\iota_{\widetilde{U}} : U \to \widetilde{U}$ is assumed to be completely continuous. Furthermore, by definition of the duality mapping $J_{U_i} : U_i \to U_i^*$, see Definition 2.53, we obtain the nonnegativity of the following term

$$\gamma \langle J_{U_i}(u^i) - J_{U_i}(v^i), u^i - v^i \rangle_{U_i^*,U_i} = \gamma \|u^i - v^i\|_{U_i}^2 \geq 0,$$

which yields the monotonicity of the second term of (6.37). Moreover, the duality mapping $J_{U_i} : U_i \to U_i^*$ is hemicontinuous by Proposition 2.54. Hence, the operator $(\theta_i)_{u^i}$ is indeed pseudomonotone, conferring Proposition 2.48.

Next, we formulate a similar result to Proposition 6.7 and Proposition 6.8, but this time we investigate the Lagrangian L_Y^i instead of L_X^i, see again (6.14) and (2.18) for their respective definitions. In this case, we require that the ERCQ holds for the reformulated problems (6.13), $i \in [N]$, and we are able to show the boundedness of the Lagrangian multipliers in Y^* and not only in X^*. Therefore, we obtain the existence of a weakly convergent subsequence of $\{\tilde{\lambda}_k^i\}_{k\in\mathbb{N}} \subseteq Y^*$ since Y is reflexive by the Milman–Pettis theorem. Moreover, we prove its weak convergence to the Lagrangian multipliers $\tilde{\lambda}^i \in Y^*$.

Proposition 6.9. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.4 and let the i-th derivative of θ_i be bounded on bounded sets and pseudomonotone. Furthermore, we assume that $\{\tilde{\lambda}_k^i\}_{k\in\mathbb{N}}$ is defined as in (2.27) and the sequence $\{u_k\}_{k\in\mathbb{N}}$ contains a weakly convergent subsequence $\{u_{k_l}\}_{l\in\mathbb{N}}$ in U. If the ERCQ holds in the weak limit $\bar{u}$ with respect to the reformulated problem (6.13) for all $i \in [N]$, then $\{\tilde{\lambda}_{k_l}^i\}_{l\in\mathbb{N}}$ is bounded in Y^* and it contains a further subsequence that converges weakly to an element $\tilde{\lambda}^i \in Y^*$ for all $i \in [N]$ that fulfills

$$\begin{aligned} \Big\langle \big[(L_Y^i)_{v^i}(v^i, \bar{u}^{-i}, \tilde{\lambda}^i) \big]_{|_{v^i=\bar{u}^i}}, z^i - \bar{u}^i \Big\rangle_{U_i^*,U_i} &\geq 0 \qquad \forall\, z^i \in \mathcal{X}_i, \\ \langle \tilde{\lambda}^i, y - e(G(\bar{u})) \rangle_{Y^*,Y} &\leq 0 \qquad \forall\, y \in K_Y. \end{aligned} \tag{6.38}$$

Additionally, if $\bar{u}$ is feasible, then the tuple $(\bar{u}, \tilde{\lambda}) \in U \times (Y^*)^N$ satisfies the KKT conditions of the GNEP consisting of (6.13), i.e., $\bar{u}$ is a quasi-Nash equilibrium.

Proof. We follow the steps in the proofs of Proposition 6.7 and Proposition 6.8. Namely, we separate the proof into two parts and begin by showing the boundedness of the Lagrangian multiplier $\{\tilde{\lambda}^i_{k_l}\}_{l\in\mathbb{N}}$ in Y^*. To this end, we use the ERCQ and apply the generalized open mapping theorem. Afterwards, we show in the proof's second step that the KKT conditions are satisfied for a feasible limit point $\bar{u}$.

Since the ERCQ holds in $\bar{u} \in \mathcal{X}$ with respect to the reformulated problems (6.13) for all $i \in [N]$, we have by definition

$$0 \in \operatorname{int}\big(e(G(\bar{u})) + e\big(G_{v^i}(\bar{u})(\mathcal{X}_i - \bar{u}^i)\big) - K_Y\big).$$

Next, we apply the generalized open mapping theorem, see Theorem 2.35, to the operator $\Psi : U_i \times Y \to Y$ that is defined by

$$\Psi(x, y) = e\big(G_{v^i}(\bar{u})(x - \bar{u}^i)\big) - y + e\big(G(\bar{u})\big),$$

for any $x \in \mathcal{X}_i$ and $y \in K_Y$. The generalized open mapping theorem implies the existence of some radius $R > 0$ with

$$\overline{B}^Y_R(0) \subseteq e\big(G(\bar{u})\big) + e\big(G_{v^i}(\bar{u})\big((\mathcal{X}_i - \bar{u}^i) \cap \overline{B}^{U_i}_1(0)\big)\big) - K_Y.$$

By the definition of the dual norm, we are able to choose a sequence $\{b^i_{k_l}\}_{l\in\mathbb{N}} \subseteq Y$ that satisfies $\|b^i_{k_l}\|_Y = 1$ and $\langle \tilde{\lambda}^i_{k_l}, b^i_{k_l}\rangle_{Y^*,Y} \geq \frac{1}{2}\|\tilde{\lambda}^i_{k_l}\|_{Y^*}$. Consequently, there are elements $y^i_{k_l} \in K_Y$ and $z^i_{k_l} \in \mathcal{X}_i$ with $\|z^i_{k_l} - \bar{u}^i\|_{U_i} \leq 1$ and

$$-Rb^i_{k_l} = e(G(\bar{u})) + e\big(G_{v^i}(\bar{u})(z^i_{k_l} - \bar{u}^i)\big) - y^i_{k_l}.$$

Moreover, passing the limit $l \to \infty$ yields

$$\delta^i_{k_l} = \Big\|Rb^i_{k_l} + e(G(u_{k_l})) + e\big(G_{v^i}(u_{k_l})(z^i_{k_l} - u^i_{k_l})\big) - y^i_{k_l}\Big\|_Y \to 0.$$

This fact follows by an analogous estimate as in (6.30) and the same reasoning as in the proof of Proposition 6.7.

Since u_{k_l} is an approximate solution to the VIs (6.15), it holds

$$\langle \varepsilon^i_{k_l-1}, z^i - u^i_{k_l}\rangle_{U_i^*,U_i} \leq \big\langle (\theta_i)_{v^i}(u_{k_l}), z^i - u^i_{k_l}\big\rangle_{U_i^*,U_i} + \Big\langle \tilde{\lambda}^i_{k_l}, e\big(G_{v^i}(u_{k_l})(z^i - u^i_{k_l})\big)\Big\rangle_{Y^*,Y} \tag{6.39}$$

for all $z^i \in \mathcal{X}_i$. Adding a zero and inserting $z^i = z^i_{k_l}$ in (6.39), we can estimate

$$\begin{aligned}\frac{R}{2}\|\tilde{\lambda}^i_{k_l}\|_{Y^*} \leq \delta^i_{k_l}\|\tilde{\lambda}^i_{k_l}\|_{Y^*} + \langle \tilde{\lambda}^i_{k_l}, y^i_{k_l} - e(G(u_{k_l}))\rangle_{Y^*,Y} + \big\langle (\theta_i)_{v^i}(u_{k_l}), z^i_{k_l} - u^i_{k_l}\big\rangle_{U_i^*,U_i} \\ - \langle \varepsilon^i_{k_l-1}, z^i_{k_l} - u^i_{k_l}\rangle_{U_i^*,U_i}.\end{aligned}$$

We apply Lemma 2.65 to arrive at the inequality

$$\frac{R}{2}\|\tilde{\lambda}^i_{k_l}\|_{Y^*} \leq \delta^i_{k_l}\|\tilde{\lambda}^i_{k_l}\|_{Y^*} + \zeta^i_{k_l} + \big\langle (\theta_i)_{v^i}(u_{k_l}) - \varepsilon^i_{k_l-1}, z^i_{k_l} - u^i_{k_l}\big\rangle_{U_i^*,U_i}. \tag{6.40}$$

In this inequality, we exploit that the sequences $\{\delta^i_{k_l}\}_{l\in\mathbb{N}}$ and $\{\zeta^i_{k_l}\}_{l\in\mathbb{N}}$ on the right-hand side converge to zero as $l \to \infty$ and that the i-th derivative of θ_i is bounded on bounded sets. Therefore, we are able to select l large enough such that it holds $\frac{R}{4}\|\tilde{\lambda}^i_{k_l}\|_{Y^*} \leq 2C$ for some constant $C > 0$ and we conclude that $\{\tilde{\lambda}^i_{k_l}\}_{l\in\mathbb{N}}$ is bounded in Y^*.

By the Milman–Pettis theorem, see Theorem 2.41, we know that Y^* is reflexive and we conclude by Lemma 2.31 that there exists a weakly convergent subsequence $\tilde{\lambda}^i_{k_{l_m}} \rightharpoonup \tilde{\lambda}^i$ in Y^* as $m \to \infty$. Moreover, $\{\tilde{\lambda}^i_{k_{l_m}}\}_{m\in\mathbb{N}}$ also converges in the weak-$*$ topology of Y^* to $\tilde{\lambda}^i$ as $m \to \infty$, see Remark 2.32. By Lemma 2.65, we get for all $y \in Y$

$$\langle \tilde{\lambda}^i, y - e(G(\bar{u}))\rangle_{Y^*,Y} = \lim_{m\to\infty} [\langle \tilde{\lambda}^i_{k_{l_m}}, y - e(G(u_{k_{l_m}}))\rangle_{Y^*,Y} - \zeta^i_{k_{l_m}}] \leq 0,$$

which shows the second inequality of the KKT conditions.

Next, we choose an arbitrarily fixed element $z^i \in \mathcal{X}_i$ and apply the limit superior in (6.39), which yields

$$0 \leq \limsup_{m\to\infty} \big\langle (\theta_i)_{v^i}(u_{k_{l_m}}), z^i - u^i_{k_{l_m}}\big\rangle_{U_i^*,U_i} + \big\langle \tilde{\lambda}^i, e\big(G_{v^i}(\bar{u})(z^i - \bar{u}^i)\big)\big\rangle_{Y^*,Y}, \tag{6.41}$$

where we used $\varepsilon^i_{k_{l_m}-1} \to 0$ in U_i^* as $m \to \infty$. Similarly, we consider $z^i = \bar{u}^i$ and apply the limit inferior as $m \to \infty$ in (6.39) to obtain

$$0 \leq \liminf_{m\to\infty} \big\langle (\theta_i)_{v^i}(u_{k_{l_m}}), \bar{u}^i - u^i_{k_{l_m}}\big\rangle_{U_i^*,U_i}. \tag{6.42}$$

Then exploiting the pseudomonotonicity of $(\theta_i)_{v^i}$ for any $i \in [N]$ and (6.42) yields

$$\begin{aligned} 0 &\leq \big\langle (\theta_i)_{v^i}(\bar{u}), z^i - \bar{u}^i\big\rangle_{U_i^*,U_i} + \big\langle \tilde{\lambda}^i, e\big(G_{v^i}(\bar{u})(z^i - \bar{u}^i)\big)\big\rangle_{Y^*,Y} \\ &= \Big\langle \big[(L^i_Y)_{v^i}(v^i, \bar{u}^{-i}, \tilde{\lambda}^i)\big]\big|_{v^i=\bar{u}^i}, z^i - \bar{u}^i\Big\rangle_{U_i^*,U_i}. \end{aligned}$$

for all $z^i \in \mathcal{X}_i$. In the case of a feasible point $\bar{u}$, we follow that $\bar{u}$ is a quasi-Nash equilibrium. □

Up to now, we have shown in this section that the weak limit $\bar{u} \in \mathcal{X}$ is admissible under the ERCQ assumption with respect to the original problem. If we assume that $\{\lambda^i_{k_l}\}_{l\in\mathbb{N}} \subseteq X^*$ has a weakly-$*$ convergent subsequence with limit λ^i and that the ERCQ holds for the GNEP consisting of (6.13), $i \in [N]$, then it holds $\lambda^i = e^*\tilde{\lambda}^i$. Indeed, for any $x \in X$ we can make the following limit computation

$$\langle \lambda^i, x\rangle_{X^*,X} = \lim_{l\to\infty} \langle \lambda^i_{k_l}, x\rangle_{X^*,X} = \lim_{l\to\infty} \langle \tilde{\lambda}^i_{k_l}, ex\rangle_{Y^*,Y} = \langle \tilde{\lambda}^i, ex\rangle_{Y^*,Y} = \langle e^*\tilde{\lambda}^i, x\rangle_{X^*,X}.$$

Remark 6.10. In the proof of convergence, we required the assumption $\mathcal{X} = \prod_{i\in[N]} \mathcal{X}_i$. This assumption is indispensable in the following steps of the proof:

- First, the tangent cone $T_{\mathcal{X}_i}(u^i)$ is only defined if and only if $\mathcal{X}_i$ itself exists. If we do not have a product structure, then the tangent cone cannot be calculated in a point $u^i \in U_i$ but only in $u \in \mathcal{X} \subseteq U$.

- By definition of the tangent cone, see Definition 2.59, given an arbitrary convex, closed set $\mathcal{X}$, the first-order optimality condition to the problem $\min_{u\in U} \operatorname{dist}^2(e(G(u)), K_Y)$ in $u \in \mathcal{X}$ reads

$$\left\langle \operatorname{dist}^2_v(e(G(\cdot)), K_Y)(u), h \right\rangle_{U^*,U} \geq 0 \qquad \forall\, h \in T_{\mathcal{X}}(u), \tag{6.43}$$

which is equivalent to

$$\left\langle \operatorname{dist}^2_v(e(G(\cdot)), K_Y)(u), z-u \right\rangle_{U^*,U} \geq 0 \qquad \forall\, z \in \mathcal{X}.$$

Then it yields

$$\left\langle \operatorname{dist}^2_v(e(G(\cdot)), K_Y)(u), z-u \right\rangle_{U^*,U} = \sum_{i\in[N]} \left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(u), z^i-u^i \right\rangle_{U_i^*,U_i} \tag{6.44}$$

for all $z \in \mathcal{X}$. However, we only have $z^i \in U_i$ and the terms in the sum of the right-hand side do not express the first-order optimality condition. In the case of $\mathcal{X} = \prod_{i\in[N]} \mathcal{X}_i$, we claim that we obtain the equivalence of (6.43) and

$$\left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(u), h^i \right\rangle_{U_i^*,U_i} \geq 0 \qquad \forall\, h^i \in T_{\mathcal{X}_i}(u^i). \tag{6.45}$$

In fact, if (6.45) holds, then we can make use of (6.44) and conclude that $u \in \mathcal{X}$ is a stationary point of $\min_{v\in\mathcal{X}} \operatorname{dist}^2(e(G(v)), K_Y)$. In the case of the other direction of the equivalency, we assume that (6.43) is valid for all $h \in T_{\mathcal{X}}(u) = T_{\mathcal{X}_1}(u^1)\times\cdots\times T_{\mathcal{X}_N}(u^N)$. Since $0 \in T_{\mathcal{X}_i}(u^i)$ holds true for any $i \in [N]$, we obtain for $h = (h^i, 0, \ldots, 0) \in T_{\mathcal{X}}(u)$ the following result

$$\left\langle \operatorname{dist}^2_{v^i}(e(G(\cdot)), K_Y)(u), h^i \right\rangle_{U_i^*,U_i} = \left\langle \operatorname{dist}^2_v(e(G(\cdot)), K_Y)(u), h \right\rangle_{U^*,U} \geq 0.$$

- We cannot infer $(z^i, u^{-i}) \in \mathcal{X}$ for $u \in \mathcal{X}$ by $z \in \mathcal{X}$ and vice versa. For example, one can consider $\mathcal{X} = \overline{B}_1(0)$, $u = (0,1) \in \mathcal{X}$, and $z = (1,0) \in \mathcal{X}$. In this example, it holds $(z^1, u^{-1}) = (1,1) \notin \mathcal{X}$.

- For proving the boundedness of the Lagrangian multiplier in the proof of the KKT conditions, an appropriate ERCQ assumption is essential. In fact, the ERCQ provides a representation of elements in $\mathcal{X}$. If it holds $\mathcal{X} \neq \prod_{i\in[N]} \mathcal{X}_i$, however, the representation cannot be reconnected with an $\mathcal{X}_i$ element.

6.2.2 Augmented Lagrangian Method for Variational Equilibria

In the following subsection, we will discuss an augmented Lagrangian method for finding variational equilibria to the GNEP consisting of (6.13). Due to the nonconvex nature of the constraints, we can only expect variational equilibria and not normalized equilibria. Recall that normalized equilibria of the problem are characterized by a solution to

$$\min_{v\in U} \sum_{i\in[N]} \theta_i(v^i, u^{-i}) \qquad \text{s.t.} \qquad v \in \mathcal{X}, \quad e(G(v)) \in K_Y, \tag{6.46}$$

but variational equilibria satisfy this problem's first-order optimality conditions. In this context, the convergence behavior is of interest. Specifically, we demonstrate that the weak limit of the augmented Lagrangian method is a KKT point of the associated optimization problem.

We introduce the augmented Lagrangian functional $L_\rho : U \times Y^* \to \mathbb{R}$ for some given positive parameter $\rho > 0$ and an element $u \in U$ by

$$L_\rho(v, w; u) = \sum_{i\in[N]} [\theta_i(v^i, u^{-i})] + \frac{\rho}{2}\Big\|e(G(v)) + \tfrac{J_Y^{-1}(w)}{\rho} - P_{K_Y}\big(e(G(v)) + \tfrac{J_Y^{-1}(w)}{\rho}\big)\Big\|_Y^2.$$

Recalling the derivative of the distance functional, see Proposition 2.58, the derivative of L_ρ with respect to its first component can be computed as follows

$$\begin{aligned}
&\big\langle L_\rho(v,w;\, u)_v, h\big\rangle_{U^*,U} \\
&\quad = \Big\langle\Big[\sum_{i\in[N]} [\theta_i(v^i, u^{-i})] + \frac{\rho}{2}\Big\|e(G(v)) + \tfrac{J_Y^{-1}(w)}{\rho} - P_{K_Y}\big(e(G(v)) + \tfrac{J_Y^{-1}(w)}{\rho}\big)\Big\|_Y^2\Big]_v, h\Big\rangle_{U^*,U} \\
&\quad = \sum_{i\in[N]} [\langle(\theta_i)_{v^i}(v^i, u^{-i}), h^i\rangle_{U_i^*,U_i}] \\
&\qquad + \rho\Big\langle J_Y\big(e(G(v)) + \tfrac{J_Y^{-1}(w)}{\rho} - P_{K_Y}\big(e(G(v)) + \tfrac{J_Y^{-1}(w)}{\rho}\big)\big), \big(e(G(\cdot))\big)_v(v)h\Big\rangle_{Y^*,Y},
\end{aligned} \tag{6.47}$$

for all $h \in U$. Next, we state the augmented Lagrangian method for the optimization problem (6.46) with a subproblem that utilizes first-order stationary points.

Algorithm 6.11.

0. Choose a bounded set $B \subseteq Y^*$ and the parameters $\rho_0 > 0$, $\gamma > 1$, $\tau \in (0, 1)$.

For $k = 0, 1, 2, 3, \ldots$:

1. If u_k satisfies a given stopping criterion: STOP
2. Choose $w_k \in B$ and compute an approximate solution u_{k+1} to the VI

$$u_{k+1} \in \mathcal{X}, \qquad \Big\langle\big[(L_{\rho_k})_v(v, w_k; u_{k+1})\big]_{|v=u_{k+1}}, z - u_{k+1}\Big\rangle_{U^*,U} \geq 0 \qquad \forall\, z \in \mathcal{X}. \tag{6.48}$$

3. Compute

$$r_{k+1} = \Big\|e(G(u_{k+1})) - P_{K_Y}\big(e(G(u_{k+1})) + \tfrac{J_Y^{-1}(w_k)}{\rho_k}\big)\Big\|_Y.$$

4. If $k = 0$, or $k \geq 1$ and $r_{k+1} \leq \tau r_k$, then set $\rho_{k+1} = \rho_k$. Otherwise, set $\rho_{k+1} = \gamma\rho_k$.

A typical choice for the element $w_k \in B$ in the second step of Algorithm 6.11 is given in [72] and reads as follows

$$w_k = P_B\Big(\rho_{k-1} J_Y\big(e(G(u_k)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}} - P_{K_Y}\big(e(G(u_k)) + \tfrac{J_Y^{-1}(w_{k-1})}{\rho_{k-1}}\big)\big)\Big).$$

As discussed previously, $w \in Y^*$ can be viewed as a Lagrangian multiplier estimate for the constraint $e(G(u)) \in K_Y$ and can be regarded as a safeguarded version of the Lagrangian multiplier $\tilde{\lambda}$. Looking at the VI (6.48) in the second step of the algorithm, one can interpret an exact solution u_{k+1} to the VI (6.48) as a variational equilibrium to the game that is connected with the minimization problem

$$\min_{v \in U} L_{\rho_k}(v, w_k; u_{k+1}) \qquad \text{s.t.} \qquad v \in \mathcal{X}. \tag{6.49}$$

In fact, we consider a game where player i's problem is given by

$$\min_{v^i \in U_i} \left[\theta_i(v^i, u_{k+1}^{-i}) + \frac{\rho_k}{2N} \text{dist}^2\left(e(G(v)) + \tfrac{J_Y^{-1}(w_k)}{\rho_k}, K_Y\right)\right] \qquad \text{s.t.} \qquad (v^i, u_{k+1}^{-i}) \in \mathcal{X}. \tag{6.50}$$

Comparing this minimization problem with the original i-th players problem in the beginning of the section, see (6.11), we observe that we have a different cost functional, u is replaced by u_{k+1} and in (6.50) we have no constraints involving G and K. If we write the counterpart to the problem (2.13) whose solutions are normalized equilibria for the game formed by (6.50), we must add up all cost functionals, which yields

$$\min_{v \in U} \sum_{i \in [N]} \left[\theta_i(v^i, u_{k+1}^{-i}) + \frac{\rho_k}{2N} \text{dist}^2\left(e(G(v)) + \tfrac{J_Y^{-1}(w_k)}{\rho_k}, K_Y\right)\right] \qquad \text{s.t.} \qquad v \in \mathcal{X}.$$

Simplifying the cost functional demonstrates that this problem is equivalent to the one presented in (6.49). Transferring the notion of a variational equilibrium from game consisting of (6.11) to the game connected to (6.50) reveals that solving (6.48) is the exact condition for a variational equilibrium u_{k+1} to the game formed by (6.50).

We only need to solve the VI (6.48) approximately for the algorithm, implying that we are interested in searching for a solution $u_{k+1} \in \mathcal{X}$ to the VI

$$\left\langle \left[(L_{\rho_k})_v(v, w_k; u_{k+1})\right]_{|v=u_{k+1}}, z - u_{k+1} \right\rangle_{U^*,U} \geq \langle \varepsilon_k, z \rangle_{U^*,U} \qquad \forall\, z \in \mathcal{X}. \tag{6.51}$$

In order to prove that Algorithm 6.11 converges, we consider a sequence $\{\varepsilon_k\}_{k \in \mathbb{N}} \subseteq U^*$ with $\varepsilon_k \to 0$ in U^* as $k \to \infty$. We can compute an approximate solution in the second step of Algorithm 6.11 by applying the descent method for continuously differentiable objective functionals to a suitable minimization problem, see Section 6.1 where we considered a similar procedure.

The remaining part of this subsection is dedicated to the convergence analysis of Algorithm 6.11. Regarding this topic, we investigate the weak limit of the proposed sequence for feasibility, stationarity, and optimality requirements. Since the minimization problem (6.46) in this section involves nonconvex constraints, we cannot prove that the limit is a normalized equilibrium.

However, we are able to compute KKT points of the problem (6.46) and we will verify that the first-order optimality criterion hold. The latter is discussed in the following lemma that provides similar results as stated in [27, Lemma 5.2], [72, Lemma 7.3], and [101, Lemma 6.20]. Moreover, the proof strategy is outlined in [66, Lemma 4.3] and we will adapt the same procedure to our problem. For proving the following statements we need some kind of ERCQ, see Definition 2.62. If nothing is mentioned the ERCQ is taken with respect to (6.52) below.

Lemma 6.12. Let the sequence $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.11 and let $\{u_{k_l}\}_{l\in\mathbb{N}} \subseteq \mathcal{X}$ be a weakly convergent subsequence with the weak limit $\bar{u}$. Furthermore, let the derivative $\left(\sum_{i\in[N]} \theta_i(v^i, u^{-i})\right)_v : U \to U^*$ be bounded on bounded sets. Then $\bar{u} \in \mathcal{X}$ is a stationary point of $\min_{v\in\mathcal{X}} \operatorname{dist}^2(e(G(v)), K_Y)$, i.e., the first-order optimality condition holds at $\bar{u}$.

Proof. This proof follows the lines of the proof of Lemma 6.5 from above. We split the proof in two different cases. Either, the parameter subsequence $\{\rho_{k_l-1}\}_{l\in\mathbb{N}}$ in the fourth step of Algorithm 6.11 is bounded or unbounded.

In the case of a bounded subsequence $\{\rho_{k_l-1}\}_{l\in\mathbb{N}}$, there exists some $m \in \mathbb{N}$ with $\rho_{k_l-1} = \rho_{k_m-1}$ for all $l \geq m$ and therefore, it has to hold $r_{k_l-1} \leq \tau^{k_l-k_{l-1}} r_{k_{l-1}-1}$ for all $l \geq m$ by the update rule of ρ_{k_l-1}, see the fourth step in Algorithm 6.11. We estimate the distance functional by the value $r_{k_{l+1}}$, see (6.17), and obtain as $l \to \infty$

$$0 \leq \operatorname{dist}(e(G(u_{k_{l+1}})), K_Y) \leq r_{k_{l+1}} \leq r_{k_{l+1}-1} \leq \tau^{k_{l+1}-k_m} r_{k_m-1} \to 0.$$

By the continuity of the functional $\operatorname{dist}^2(e(G(\cdot)), K_Y)$ with respect to the weak sequential topology, we conclude $\operatorname{dist}^2(e(G(\bar{u})), K_Y) = 0$, implying that $\bar{u}$ is a stationary point.

In the case of an unbounded subsequence $\{\rho_{k_l-1}\}_{l\in\mathbb{N}}$, we argue by contradiction and assume that there are constants $c_1 > 0$, $c_2 \in \mathbb{R}$, and an element $y \in \mathcal{X}$ with

$$\left\langle \left[(\operatorname{dist}^2 \left(e(G(v)) + \tfrac{J_Y^{-1}(w_{k_{l+1}-1})}{\rho_{k_{l+1}-1}}, K_Y \right) \right]_v)_{|v=u_{k_{l+1}}}, y - u_{k_{l+1}} \right\rangle_{U^*,U} < -c_1,$$

and

$$\sum_{i\in[N]} \left\langle (\theta_i)_{v^i}(u_{k_{l+1}}), y^i - u^i_{k_{l+1}} \right\rangle_{U_i^*,U_i} \leq c_2,$$

for an index l being sufficiently large. Since $u_{k_{l+1}}$ is an approximate solution to (6.48), it fulfills the VI (6.51) and plugging the derivative of the Lagrangian (6.47) into the VI, we arrive at the contradiction

$$\lim_{l\to\infty} \langle \varepsilon_{k_{l+1}-1}, y - u_{k_{l+1}} \rangle_{U^*,U} \leq c_2 - c_1 \lim_{l\to\infty} \frac{\rho_{k_{l+1}-1}}{2} = -\infty. \qquad \square$$

Next, we examine the conditions under which the weak limit $\bar{u} \in \mathcal{X}$ is a feasible point. We refer to the works [27, Lemma 5.2] and [72, Lemma 7.3] for similar results.

Proposition 6.13. Let $e(X)$ be dense in Y. If $u \in \mathcal{X}$ is a stationary point of $\min_{v\in\mathcal{X}} \operatorname{dist}^2(e(G(v)), K_Y)$ and satisfies the ERCQ, then it follows $e(G(u)) \in K_Y$ and $G(u) \in K$, i.e., u is feasible.

Proof. The proof proceeds in the same manner as the proof of Proposition 6.6. However, this time the ERCQ holds in terms of the problem

$$\min_{v\in U} \sum_{i\in[N]} \theta_i(v^i, u^{-i}) \qquad \text{s.t.} \qquad v \in \mathcal{X}, \quad G(v) \in K, \tag{6.52}$$

whereas in Proposition 6.6 we studied the ERCQ for the i-th minimization problem (6.11). Consequently, for any $x \in \overline{B}_R^X(0)$, we obtain the existence of elements $y \in K$ and $z \in \mathcal{X}$ such that it holds

$$x = G(u) + G_v(u)(z - u) - y.$$

Using an analogous equation as (6.23), we get

$$\left\langle e^* J_Y\big(e(G(u)) - P_{K_Y}(e(G(u)))\big), x\right\rangle_{X^*,X} = \left\langle G_v(u)^* e^* J_Y\big(e(G(u)) - P_{K_Y}(e(G(u)))\big), z - u\right\rangle_{U^*,U} + \left\langle e^* J_Y\big(e(G(u)) - P_{K_Y}(e(G(u)))\big), G(u) - y\right\rangle_{X^*,X}.$$

The first term corresponds to the derivative of $\frac{1}{2}\mathrm{dist}^2(e(G(\cdot)), K_Y)$, see Proposition 2.58, and since $u \in \mathcal{X}$ is stationary, this term is nonnegative. The nonnegativity of the second term follows from the estimate (6.26) that we have already derived in the proof of Proposition 6.6. Altogether, using the density of $e(X)$ in Y, it yields $e(G(u)) \in K_Y$ and $G(u) \in K$. □

In the following propositions, we show that the approximate solutions to the games (6.46) and (6.52) converge to KKT points under given assumptions. In other words, they converge to variational equilibria of the corresponding GNEP. We refer to [27, Theorem 5.4] and [72, Theorem 7.4] for similar statements and we follow their proof strategy. As the main tool, we apply again the generalized open mapping theorem, see Theorem 2.35, to a suitable functional.

In particular, we prove in Proposition 6.14 the boundedness of $\{\lambda_{k_l}\}_{l\in\mathbb{N}}$ in the space X^* under given assumptions. If there exists a weakly-$*$ convergent subsequence, its limit fulfills the KKT conditions, see Proposition 6.15. For example, this assumption is fulfilled if X is a separable or reflexive space. Afterwards, we prove in Proposition 6.16 the boundedness of the subsequence $\{\tilde{\lambda}_{k_l}\}_{l\in\mathbb{N}}$ in the reflexive space Y^*. To this end, we need the ERCQ assumption for the reformulated problem (6.46). In this setting, we are allowed to extract a weakly converging subsequence and show that the limit point is a KKT point of the corresponding GNEP.

Proposition 6.14. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.11 and let the derivative $\big(\sum_{i\in[N]} \theta_i(v^i, u^{-i})\big)_v : U \to U^*$ of the objective functional be bounded on bounded sets. Furthermore, we assume that there is a subsequence of $\{u_k\}_{k\in\mathbb{N}}$ with $u_{k_l} \rightharpoonup \bar{u}$ in U as $l \to \infty$. If the ERCQ holds in the weak limit $\bar{u}$, then the sequence $\{\lambda_{k_l}\}_{l\in\mathbb{N}}$ is bounded in X^*.

Proof. We follow the procedure of Proposition 6.7. However, we work with problem (6.52) instead of the i-th problem (6.11). Consequently, we use the ERCQ with respect to the optimization issue (6.52) and the derivative of G with respect to v.

As in the referred proof, we make use of the generalized open mapping theorem, see Theorem 2.35. We conclude the existence of some radius $R > 0$, an element $z_{k_l} \in \mathcal{X}$ with $\|z_{k_l} - \bar{u}\|_U \leq 1$, and an element $y_{k_l} \in K_Y$ with

$$-Rb_{k_l} = G(\bar{u}) + G_v(\bar{u})(z_{k_l} - \bar{u}) - y_{k_l},$$

where $b_{k_l} \in X$ is chosen with $\|b_{k_l}\|_X = 1$ and $\langle \lambda_{k_l}, b_{k_l} \rangle_{X^*,X} \geq \frac{1}{2}\|\lambda_{k_l}\|_{X^*}$. We have assumed that G and its derivative are completely continuous and therefore, it yields as $l \to \infty$

$$\delta_{k_l} = \left\| Rb_{k_l} + G(u_{k_l}) + G_v(u_{k_l})(z_{k_l} - u_{k_l}) - y_{k_l} \right\|_X \to 0.$$

Furthermore, we exploit that u_{k_l} is an approximate solution to the VI (6.48) and fulfills (6.51), from which we conclude

$$\frac{R}{2}\|\lambda_{k_l}\|_{X^*} \leq \delta_{k_l}\|\lambda_{k_l}\|_{X^*} + \zeta_{k_l} + \sum_{i\in[N]} \left\langle (\theta_i)_{v^i}(u_{k_l}) - \varepsilon_{k_l-1}, z^i_{k_l} - u^i_{k_l} \right\rangle_{U_i^*,U_i},$$

where the sequence $\{\zeta_{k_l}\}_{l\in\mathbb{N}}$ appeared on the right-hand side due to Lemma 2.65. Consequently, the subsequence $\{\lambda_{k_l}\}_{l\in\mathbb{N}}$ is bounded in X^*. □

Proposition 6.15. Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ be generated by Algorithm 6.11 and let the derivative $\left(\sum_{i\in[N]} \theta_i(v^i, u^{-i})\right)_v : U \to U^*$ of the objective functional be bounded on bounded sets and pseudomonotone. Furthermore, we assume that there is a subsequence of $\{u_k\}_{k\in\mathbb{N}}$ with $u_{k_l} \rightharpoonup \bar{u}$ in U as $l \to \infty$. If there exists a weakly-$*$ convergent subsequence of $\{\lambda_{k_l}\}_{l\in\mathbb{N}}$ with weak-$*$ limit point λ, then it holds

$$\begin{aligned} \left\langle \left[(L_X)_v(v, \lambda; \bar{u})\right]_{|v=\bar{u}}, z - \bar{u} \right\rangle_{U^*,U} &\geq 0 \qquad \forall\, z \in \mathcal{X}, \\ \langle \lambda, y - G(\bar{u}) \rangle_{X^*,X} &\leq 0 \qquad \forall\, y \in K. \end{aligned}$$

Moreover, if $e(X)$ is dense in Y, then $(\bar{u}, \lambda) \in U \times X^*$ is a KKT pair, i.e., a variational equilibrium to the game consisting of (6.11).

Proof. This proof is based on the proof of Proposition 6.8. Let $\{\lambda_{k_{l_m}}\}_{m\in\mathbb{N}} \subseteq X^*$ be a weakly-$*$ convergent subsequence with $\lambda_{k_{l_m}} \overset{*}{\rightharpoonup} \lambda$ in X^* as $m \to \infty$. Since G is completely continuous and $\zeta_{k_{l_m}}$ converges to zero due to Lemma 2.65, we obtain

$$\langle \lambda, y - G(\bar{u}) \rangle_{X^*,X} = \lim_{m\to\infty} \left[\langle \lambda_{k_{l_m}}, y - G(u_{k_{l_m}}) \rangle_{X^*,X} - \zeta_{k_{l_m}}\right] \leq 0 \qquad \forall\, y \in X.$$

Let $z \in \mathcal{X}$ be arbitrarily fixed. Then the pseuomonotonicity of $\left(\sum_{i\in[N]} \theta_i(v^i, u^{-i})\right)_v$ yields

$$\begin{aligned} 0 &\leq \sum_{i\in[N]} \left[\left\langle (\theta_i)_{v^i}(\bar{u}), z^i - \bar{u}^i \right\rangle_{U_i^*,U_i}\right] + \left\langle \lambda, G_v(\bar{u})(z - \bar{u}) \right\rangle_{X^*,X} \\ &= \left\langle \left[(L_X)_v(v, \lambda; \bar{u})\right]_{|v=\bar{u}}, z - \bar{u} \right\rangle_{U^*,U}. \end{aligned}$$

Together with Lemma 6.12 and Proposition 6.13, we conclude that $\bar{u}$ is feasible. Thus, the tuple $(\bar{u}, \lambda) \in U \times X^*$ is a KKT pair. □

The next proposition is similar to Proposition 6.16. Its proof follows the lines of the proof of Proposition 6.16, but uses the ERCQ with respect to the reformulated problem (6.46) instead of the reformulated i-th problems (6.13). We emphasize that we derive the boundedness of the Lagrangian multiplier in Y^* in this result and not only the boundedness of the corresponding multiplier in X^* as in the results Proposition 6.14 and Proposition 6.15.

Proposition 6.16. Let $\{u_k\}_{k\in\mathbb{N}}$ be generated by Algorithm 6.11 and let the derivative $\big(\sum_{i\in[N]}\theta_i(v^i,u^{-i})\big)_v : U \to U^*$ be bounded on bounded sets and pseudomonotone. Furthermore, we assume that $u_{k_l} \rightharpoonup \bar{u}$ in U as $l \to \infty$. If the ERCQ for the reformulated problem (6.46) holds in $\bar{u}$, then the sequence $\{\tilde{\lambda}_{k_l}\}_{l\in\mathbb{N}}$ is bounded in Y^* and it possesses a subsequence that converges weakly to $\tilde{\lambda} \in Y^*$ such that

$$\Big\langle \big[(L_Y)_v(v,\tilde{\lambda};\bar{u})\big]_{|v=\bar{u}}, z-\bar{u}\Big\rangle_{U^*,U} \geq 0 \qquad \forall\, z \in \mathcal{X},$$
$$\langle \tilde{\lambda}, y - e(G(\bar{u}))\rangle_{Y^*,Y} \leq 0 \qquad \forall\, y \in K_Y.$$

Furthermore, if $\bar{u}$ is feasible, then $(\bar{u},\tilde{\lambda}) \in U\times Y^*$ is a KKT pair, i.e., a variational equilibrium.

Overall, we have shown that the ERCQ assumption with respect to (6.46) implies that there exists a subsequence such that it holds $\tilde{\lambda}_{k_{l_m}} \rightharpoonup \tilde{\lambda}$ in Y^* as $m \to \infty$. This fact follows from the reflexivity of Y. Furthermore, if there exists a weakly-$*$ convergent subsequence $\{\lambda_{k_{l_m}}\}_{m\in\mathbb{N}}$ in X^*, then we have proved that the limit λ satisfies $\lambda = e^*\tilde{\lambda}$.

Chapter 7

Conclusion and Outlook

> *We have not succeeded in answering all our problems. The answers we have found only serve to raise a whole set of new questions. In some ways we feel we are as confused as ever, but we believe we are confused on a higher level and about more important things.* (Bernt Øksendal)

In this thesis, we have discussed GNEPs in infinite-dimensional spaces in the sense that we have investigated the existence of equilibria, characterized those points using the regularized Nikaido–Isoda merit functional, and studied algorithms and methods for computing them. More precisely, we considered GNEPs without the assumption of convexity on the problem's objective functional. We applied a generalized version of the Kakutani theorem to a solution mapping, which is based on the regularized Nikaido–Isoda functional and the first-order optimality conditions. This allowed us to prove the existence of a fixed point of the corresponding solution map. Furthermore, we characterize different types of equilibria using regularized and localized Nikaido–Isoda functionals. Thus, we were able to prove that these fixed points are related to the variational and normalized equilibria of the GNEP. Based on a generalized version of Danskin's theorem, which is a classic result on the differentiability of min-max problems, we developed a derivative-based framework for deriving continuity and differentiability results for the regularized and localized Nikaido–Isoda merit functionals. Moreover, we presented an augmented Lagrangian method for approximating a GNEP with nonconvex constraints and proved several statements regarding the convergence to KKT points that characterize first-order equilibria.

One open task in the study of GNEPs with nonconvex constraints and objective functionals is the application of the derived methods with an example from the real world. In such a scenario, one can consider for the nonconvex constraint a solution operator to a partial differential equation such as a hyperbolic equation describing traffic flow. Typically, one cannot expect that a solution operator to a possibly ill-behaving partial differential equation admits very convenient properties for the analysis of the problem such as smoothness or convexity.

Acronyms

CQ	constraint qualification	5
ERCQ	extended Robinson constraint qualification	24
GNEP	generalized Nash equilibrium problem	1
KKT	Karush–Kuhn–Tucker	3
NEP	Nash equilibrium problem	1
QVI	quasi-variational inequality	2
(Q)VI	(quasi-)variational inequality	3
RCQ	Robinson constraint qualification	24
VI	variational inequality	3

Bibliography

[1] Y. I. Alber. A bound for the modulus of continuity for metric projections in a uniformly convex and uniformly smooth Banach space. *Journal of Approximation Theory*, 85(3):237–249, 1996.

[2] A. Alphonse, M. Hintermüller, and C. N. Rautenberg. Directional differentiability for elliptic quasi-variational inequalities of obstacle type. *Calculus of Variations and Partial Differential Equations*, 58(1):39, 2019.

[3] H. W. Alt. *Linear Functional Analysis: An Application-Oriented Introduction*. Springer, 2016.

[4] A. B. Amar and D. O'Regan. *Topology and Approximate Fixed Points*. Springer, 2022.

[5] K. J. Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3):265–290, 1954.

[6] D. Aussel. Subdifferential properties of quasiconvex and pseudoconvex functions: Unified approach. *Journal of Optimization Theory and Applications*, 97(1):29–45, 1998.

[7] C. Baiocchi and A. Capelo. *Variational and Quasivariational Inequalities: Applications to Free-Boundary Problems*. Wiley, 1984.

[8] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2011.

[9] A. Bensoussan and J.-L. Lions. *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, 1984.

[10] C. Berge. *Topological Spaces: Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity*. Courier Corporation, 1963.

[11] P. Bernhard and A. Rapaport. On a theorem of Danskin with an application to a theorem of von Neumann–Sion. *Nonlinear Analysis: Theory, Methods & Applications*, 24(8):1163–1181, 1995.

[12] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1996.

[13] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2016.

[14] E. G. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization*. SIAM, 2014.

[15] G. Birkhoff. On the combination of topologies. *Fundamenta Mathematicae*, 26(1):156–166, 1936.

[16] M. C. Bliemer and P. H. Bovy. Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem. *Transportation Research Part B: Methodological*, 37(6):501–519, 2003.

[17] J. F. Bonnans and A. Shapiro. Optimization problems with perturbations: A guided tour. *SIAM Review*, 40(2):228–264, 1998.

[18] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, 2013.

[19] A. Bressan and K. Han. Nash equilibria for a model of traffic flow with several groups of drivers. *ESAIM: Control, Optimisation and Calculus of Variations*, 18(4):969–986, 2012.

[20] A. Bressan and K. Han. Existence of optima and equilibria for traffic flow on networks. *Networks and Heterogeneous Media*, 8(3):627–648, 2013.

[21] A. Bressan and K. T. Nguyen. Optima and equilibria for traffic flow on networks with backward propagating queues. *Networks and Heterogeneous Media*, 10(4):717–748, 2015.

[22] H. Brezis. Equations et inéquations non linéaires dans les espaces vectoriels en dualité. *Annales de l'institut Fourier*, 18(1):115–175, 1968.

[23] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.

[24] L. F. Bueno, G. Haeser, F. Lara, and F. N. Rojas. An augmented Lagrangian method for quasi-equilibrium problems. *Computational Optimization and Applications*, 76:737–766, 2020.

[25] L. F. Bueno, G. Haeser, and F. N. Rojas. Optimality conditions and constraint qualifications for generalized Nash equilibrium problems and their practical implications. *SIAM Journal on Optimization*, 29(1):31–54, 2019.

[26] E. Börgens, C. Kanzow, P. Mehlitz, and G. Wachsmuth. New constraint qualifications for optimization problems in Banach spaces based on asymptotic KKT conditions. *SIAM Journal on Optimization*, 30(4):2956–2982, 2020.

[27] E. Börgens, C. Kanzow, and D. Steck. Local and global analysis of multiplier methods for constrained optimization in Banach spaces. *SIAM Journal on Control and Optimization*, 57(6):3694–3722, 2019.

[28] W. Cheney. *Analysis for Applied Mathematics*. Springer, 2001.

[29] C. Chidume. *Geometric Properties of Banach Spaces and Nonlinear Iterations*. Springer, 2009.

[30] I. Cioranescu. *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*. Springer, 1990.

[31] J. A. Clarkson. Uniformly convex spaces. *Transactions of the American Mathematical Society*, 40(3):396–414, 1936.

[32] A. R. Conn, N. I. Gould, and P. L. Toint. *Trust-Region Methods*. SIAM, 2000.

[33] J. B. Conway. *A Course in Functional Analysis*. Springer, 2019.

[34] R. Correa and A. Seeger. Directional derivative of a minimax function. *Nonlinear Analysis: Theory, Methods & Applications*, 9(1):13–22, 1985.

[35] S. Dancs. Generalized tangent cone and an optimization problem in a normed space. *Journal of Optimization Theory and Applications*, 67(1):43–55, 1990.

[36] J. M. Danskin. The theory of max-min, with applications. *SIAM Journal on Applied Mathematics*, 14(4):641–664, 1966.

[37] J. M. Danskin. *The Theory of Max-Min and its Application to Weapons Allocation Problems*. Springer, 2012.

[38] K. Deimling. *Nonlinear Functional Analysis*. Springer, 1985.

[39] P. Di Barba. *Multiobjective Shape Design in Electricity and Magnetism*. Springer, 2009.

[40] A. Dreves, C. Kanzow, and O. Stein. Nonsmooth optimization reformulations of player convex generalized Nash equilibrium problems. *Journal of Global Optimization*, 53(4):587–614, 2012.

[41] R. Dudley. On sequential convergence. *Transactions of the American Mathematical Society*, 112(3):483–507, 1964.

[42] F. Facchinei, A. Fischer, and V. Piccialli. On generalized Nash games and variational inequalities. *Operations Research Letters*, 35(2):159–164, 2007.

[43] F. Facchinei and C. Kanzow. Generalized Nash equilibrium problems. *Annals of Operations Research*, 175(1):177–211, 2010.

[44] F. Facchinei and J.-S. Pang. Nash equilibria: The variational approach. In D. P. Palomar and Y. C. Eldar, editors, *Convex Optimization in Signal Processing and Communications*, pages 443–498. Cambrdige, 2010.

[45] A. Fischer, M. Herrich, and K. Schönefeld. Generalized Nash equilibrium problems-recent advances and challenges. *Pesquisa Operacional*, 34(3):521–558, 2014.

[46] G. B. Folland. *Real Analysis: Modern Techniques and their Applications.* Wiley, 1999.

[47] S. Franke, P. Mehlitz, and M. Pilecka. Optimality conditions for the simple convex bilevel programming problem in Banach spaces. *Optimization*, 67(2):237–268, 2018.

[48] M. Galewski. *Basic Monotonicity Methods with some Applications.* Springer, 2021.

[49] D. J. Garling. *A Course in Mathematical Analysis Volume II: Metric and Topological Spaces, Functions of a Vector Variable.* Cambridge University Press, 2013.

[50] I. L. Glicksberg. A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.

[51] K. Goebel and S. Reich. *Uniform convexity, Hyperbolic Geometry, and Nonexpansive Mappings.* Taylor & Francis, 1984.

[52] A. Goreham. Sequential convergence in topological spaces. *arXiv:math/0412558 [math.GN]*, 2004.

[53] P. T. Harker. Generalized Nash games and quasi-variational inequalities. *European Journal of Operational Research*, 54(1):81–94, 1991.

[54] N. Harms, C. Kanzow, and O. Stein. Smoothness properties of a regularized gap function for quasi-variational inequalities. *Optimization Methods and Software*, 29(4):720–750, 2014.

[55] M. Hintermüller and C. N. Rautenberg. A sequential minimization technique for elliptic quasi-variational inequalities with gradient constraints. *SIAM Journal on Optimization*, 22(4):1224–1257, 2012.

[56] M. Hintermüller and C. N. Rautenberg. Parabolic quasi-variational inequalities with gradient-type constraints. *SIAM Journal on Optimization*, 23(4):2090–2123, 2013.

[57] M. Hintermüller and T. Surowiec. A PDE-constrained generalized Nash equilibrium problem with pointwise control and state constraints. *Pacific Journal of Optimization*, 9(2):251–273, 2013.

[58] M. Hintermüller, T. Surowiec, and A. Kämmmler. Generalized Nash equilibrium problems in Banach spaces: Theory, Nikaido–Isoda-based path-following methods, and applications. *SIAM Journal on Optimization*, 25(3):1826–1856, 2015.

[59] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints.* Springer, 2008.

[60] T. Ichiishi. *Game Theory for Economic Analysis.* Elsevier, 2014.

[61] D. Inoan and J. Kolumbán. On pseudomonotone set-valued mappings. *Nonlinear Analysis: Theory, Methods & Applications*, 68(1):47–53, 2008.

[62] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. SIAM, 2008.

[63] V. I. Ivanov. Optimization and variational inequalities with pseudoconvex functions. *Journal of Optimization Theory and Applications*, 146(3):602–616, 2010.

[64] S. Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Mathematical Journal*, 8(3):457–459, 1941.

[65] C. Kanzow. On the multiplier-penalty-approach for quasi-variational inequalities. *Mathematical Programming*, 160(1):33–63, 2016.

[66] C. Kanzow, V. Karl, D. Steck, and D. Wachsmuth. The multiplier-penalty method for generalized Nash equilibrium problems in Banach spaces. *SIAM Journal on Optimization*, 29(1):767–793, 2019.

[67] C. Kanzow and A. Schwartz. *Spieltheorie*. Springer, 2008.

[68] C. Kanzow and D. Steck. Augmented Lagrangian methods for the solution of generalized Nash equilibrium problems. *SIAM Journal on Optimization*, 26(4):2034–2058, 2016.

[69] C. Kanzow and D. Steck. An example comparing the standard and safeguarded augmented Lagrangian methods. *Operations Research Letters*, 45(6):598–603, 2017.

[70] C. Kanzow and D. Steck. Augmented Lagrangian and exact penalty methods for quasi-variational inequalities. *Computational Optimization and Applications*, 69(3):801–824, 2018.

[71] C. Kanzow and D. Steck. On error bounds and multiplier methods for variational problems in Banach spaces. *SIAM Journal on Control and Optimization*, 56(3):1716–1738, 2018.

[72] C. Kanzow and D. Steck. Quasi-variational inequalities in Banach spaces: Theory and augmented Lagrangian methods. *SIAM Journal on Optimization*, 29(4):3174–3200, 2019.

[73] C. Kanzow, D. Steck, and D. Wachsmuth. An augmented Lagrangian method for optimization problems in Banach spaces. *SIAM Journal on Control and Optimization*, 56(1):272–291, 2018.

[74] V. Karl and D. Wachsmuth. An augmented Lagrange method for elliptic state constrained optimal control problems. *Computational Optimization and Applications*, 69(3):857–880, 2018.

[75] A. S. Kravchuk and P. J. Neittaanmäki. *Variational and Quasi-Variational Inequalities in Mechanics*. Springer, 2007.

[76] M. Kunze and J.-F. Rodrigues. An elliptic quasi-variational inequality with gradient constraints and some of its applications. *Mathematical Methods in the Applied Sciences*, 23(10):897–908, 2000.

[77] E. E. Levi. Studii sui punti singolari essenziali delle funzioni analitiche di due o più variabili complesse. *Annali di Matematica Pura ed Applicata (1898-1922)*, 17(1):61–87, 1910.

[78] B. Luderer, L. Minchenko, and T. Satsura. *Multivalued Analysis and Nonlinear Programming Problems with Perturbations*. Springer, 2013.

[79] D. G. Luenberger. *Optimization by Vector Space Methods*. Wiley, 1997.

[80] O. L. Mangasarian. Pseudo-convex functions. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, 3(2):281–290, 1965.

[81] O. L. Mangasarian. *Nonlinear Programming*. SIAM, 1994.

[82] R. E. Megginson. *An Introduction to Banach Space Theory*. Springer, 1998.

[83] U. Mosco. Implicit variational problems and quasi variational inequalities. In J. P. Gossez, E. J. Lami Dozo, J. Mawhin, and L. Waelbroeck, editors, *Nonlinear Operators and the Calculus of Variations*, pages 83–156. Springer, 1976.

[84] J. F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

[85] R. Nessah and G. Tian. On the existence of Nash equilibrium in discontinuous games. *Economic Theory*, 61(3):515–540, 2016.

[86] H. Nikaidô and K. Isoda. Note on non-cooperative convex games. *Pacific Journal of Mathematics*, 5(1):807–815, 1955.

[87] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 2006.

[88] J. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*. Springer, 2013.

[89] J.-S. Pang and M. Fukushima. Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. *Computational Management Science*, 2(1):21–56, 2005.

[90] J. Peypouquet. *Convex Optimization in Normed Spaces: Theory, Methods and Examples*. Springer, 2015.

[91] S. M. Robinson. Normed convex processes. *Transactions of the American Mathematical Society*, 174:127–140, 1972.

[92] S. M. Robinson. Regularity and stability for convex multivalued functions. *Mathematics of Operations Research*, 1(2):130–143, 1976.

[93] J. F. Rodrigues and L. Santos. A parabolic quasi-variational inequality arising in a superconductivity model. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 29(1):153–169, 2000.

[94] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, 33:520–534, 1965.

[95] W. Rudin. *Functional Analysis*. McGraw-Hill, 1991.

[96] M. Ružička. *Nichtlineare Funktionalanalysis*. Springer, 2004.

[97] H. H. Schaefer. *Topological Vector Spaces*. Springer, 1999.

[98] L. Scrimali. Quasi-variational inequalities in transportation networks. *Mathematical Models and Methods in Applied Sciences*, 14(10):1541–1560, 2004.

[99] T. B. Singh. *Introduction to Topology*. Springer, 2019.

[100] D. Steck. Covering compactness in the weak sequential topology. MathOverflow. https://mathoverflow.net/q/285141.

[101] D. Steck. *Lagrange Multiplier Methods for Constrained Optimization and Variational Problems in Banach Spaces*. PhD thesis, Universität Würzburg, 2018.

[102] Z. Tang, J.-A. Désidéri, and J. Périaux. Multicriterion aerodynamic shape design optimization and inverse problems using control theory and Nash games. *Journal of Optimization Theory and Applications*, 135(3):599–622, 2007.

[103] F. Treves. *Topological Vector Spaces Distributions and Kernels*. Pergamon Press, 1967.

[104] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. SIAM, 2011.

[105] A. von Heusinger and C. Kanzow. SC^1 optimization reformulations of the generalized Nash equilibrium problem. *Optimisation Methods & Software*, 23(6):953–973, 2008.

[106] A. von Heusinger and C. Kanzow. Optimization reformulations of the generalized Nash equilibrium problem using Nikaido–Isoda-type functions. *Computational Optimization and Applications*, 43(3):353–377, 2009.

[107] A. von Heusinger and C. Kanzow. Relaxation methods for generalized Nash equilibrium problems with inexact line search. *Journal of Optimization Theory and Applications*, 143(1):159–183, 2009.

[108] J. C. Wells. Differentiable functions on Banach spaces with Lipschitz derivatives. *Journal of Differential Geometry*, 8(1):135–152, 1973.

[109] R. Williams. Sufficient conditions for Nash equilibria in n-person games over reflexive Banach spaces. *Journal of Optimization Theory and Applications*, 30(3):383–394, 1980.

[110] Z. Xu and G. Roach. On the uniform continuity of metric projections in Banach spaces. *Approximation Theory and its Applications*, 8(3):11–20, 1992.

[111] K. Yosida. *Functional Analysis*. Springer, 2012.

[112] J. Yu. On Nash equilibria in n-person games over reflexive Banach spaces. *Journal of Optimization Theory and Applications*, 73(1):211–214, 1992.

[113] E. Zeidler. *Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators*. Springer, 1990.

[114] J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.