

UNIVERSITÉ DE LIMOGES
Faculté des Sciences & Techniques

Bui Trong Kien

FROM NONLINEAR AND VARIATIONAL ANALYSIS
TO OPTIMAL CONTROL PROBLEMS

HABILITATION DISSERTATION IN MATHEMATICS

Limoges - 2026

UNIVERSITÉ DE LIMOGES
Faculté des Sciences & Techniques

Bui Trong Kien

FROM NONLINEAR AND VARIATIONAL ANALYSIS
TO OPTIMAL CONTROL PROBLEMS

HABILITATION DISSERTATION IN MATHEMATICS

Limoges - 2026

To my children Bui Trong Ha and Bui Ngoc Han

CONTENTS

CONTENTS	ii
TABLE OF NOTATIONS	1
PREFACE	4
Chapter 1. RESULTS IN NONLINEAR ANALYSIS 8	
1.1 Duality mappings and metric projections	8
1.1.1 Normalized duality mapping J_p	8
1.1.2 Characteristic properties of uniformly convex Banach spaces	14
1.1.3 Continuity of the metric projection onto a family of closed convex sets	19
1.2 Topological degree for monotone operators	22
1.2.1 Topological degree in finite dimensional spaces . .	22
1.2.2 Browder degree for operators of class $(S)_+$	27
1.2.3 Browder degree for mappings of monotone type .	29
1.2.4 Extension of the Browder degree theory for monotone mappings	34
1.2.5 Proofs of extension results	36
1.3 Variational inequalities	47
1.3.1 Degree theory for generalized variational inequalities	47
1.3.2 Existence of solutions to variational inequality with pseudo-monotone operators in reflexive Banach spaces	59
1.3.3 Stability of solutions to parametric generalized variational inequalities in reflexive Banach spaces . .	73
1.3.4 The Galerkin method and regularization for variational inequalities	87
Chapter 2. SOME RESULTS IN VARIATIONAL ANALYSIS AND VECTOR OPTIMIZATION 102	
2.1 Tangent sets, metric regularity and constraint qualification conditions	102
2.1.1 Tangent sets and subdifferential	102
2.1.2 Metric regularity and constraint qualification conditions	107

2.2	Optimality conditions for vector optimization problems	116
2.2.1	Optimality conditions of KKT-type	116
2.2.2	Optimality conditions of Fritz-John type	125
 Chapter 3. OPTIMAL CONTROL PROBLEMS GOVERNED BY ODE		
3.1	Existence of solutions to a class of optimal control problems	139
3.1.1	Introduction	139
3.1.2	Main result	141
3.2	Optimality conditions for multiobjective optimal control problems	151
3.2.1	First and second-order optimality conditions of KKT-type for multiobjective optimal control problems with free right endpoints	151
3.2.2	Second-order optimality conditions of KKT-type for multiobjective optimal control problems with endpoint and mixed constraints	175
3.2.3	Application to sustainable energy management in smart grids	195
 Chapter 4. OPTIMAL CONTROL PROBLEMS GOVERNED BY PDE		
4.1	Pontryagin's principle for optimal control problem governed by 3D Navier-Stokes equations	201
4.1.1	Introduction	201
4.1.2	Linearized and adjoint equations	203
4.1.3	Pontryagin's Maximum Principle	217
4.2	Regularity of multipliers and second-order optimality conditions for semilinear parabolic optimal control problems	225
4.2.1	Introduction	225
4.2.2	A specific mathematical programming problem	227
4.2.3	State equation and linearized equation	229
4.2.4	Existence of regular multipliers and second-order optimality conditions	233
4.2.5	Hölder continuity of multipliers and optimal solutions	254
4.3	Numerical solutions for semilinear elliptic optimal control with nonlinear and mixed constraints	258
4.3.1	Introduction	258
4.3.2	Discrete problems and optimality conditions	260

4.3.3	Numerical analysis for state equations and adjoint equations	279
4.3.4	Convergence and error estimates	289
4.3.5	Numerical example	300
REFERENCES		304

TABLE OF NOTATIONS

$\mathbb{N} := \{0, 1, 2, \dots\}$	set of natural numbers
$\mathbb{Z}_+ := \{1, 2, \dots\}$	set of positive integers
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of $x \in \mathbb{R}$ with $x \geq 0$
$\mathbb{R}^N$	N -dimensional Euclidean vector space
$ x $	Euclidean norm of vector x in $\mathbb{R}^N$
$\ x\ _X$	norm of vector x in the normed space X
$\ (x, y)\ _{X \times Y}$	norm of vector (x, y) in the normed space $X \times Y$
X^*	topological dual of a normed space X
X^{**}	topological bi-dual of a normed space X
$\langle x^*, x \rangle_{X^*, X}$	dual pair between X^* and X
$\langle x, y \rangle_X$	canonical inner product in a Hilbert space X
$B(x, \delta)$	open ball with center at x and radius δ
$\overline{B}(x, \delta)$	closed ball with center at x and radius δ
B_X	open unit ball in a normed space X
$\overline{B}_X$	closed unit ball in a normed space X
$\text{dist}(x; \Omega)$	distance from x to Ω
$\{x_k\}$	sequence of vectors x_k
$x_k \rightarrow x$	x_k converges strongly to x (in norm topology)
$x_k \rightharpoonup x$	x_k converges weakly to x
$\forall x$	for all x
$\exists x$	there exists x
$A := B$	A is defined by B
$f : X \rightarrow Y$	function from X to Y
$f'(x), \nabla f(x)$	Fréchet derivative of a variable function f at x
$f''(x), \nabla^2 f(x)$	Fréchet second-order derivative of a variable function f at x
$f_{x_i}, \nabla_{x_i} f$	partial derivative of f in x_i
$f_{x_i x_j}, \nabla_{x_i x_j}^2 f$	second-order partial derivative of f in x_i and x_j
$\alpha := (\alpha_1, \alpha_2, \dots, \alpha_N)$	multiindex
D^α	α -order weak partial derivative

Δ	Laplace operator
$\varphi : X \rightarrow \overline{\mathbb{R}}$	extended-real-valued function
$\text{supp } \varphi$	support of φ
$F : X \rightrightarrows Y$	multifunction from X to Y
$\text{Dom } F$	domain of F
$\text{Ran}(F)$	range of F
$\text{Gr}(F)$	graph of F
$\text{Ker } F$	kernel of F
$\text{coKer } F$	cokernel of F
$T(K, x)$	Bouligand tangent cone of the set K at x
$T^\flat(K, x)$	adjacent tangent cone to K at x
$N(K, x)$	normal cone of the set K at x
$\partial\Omega$	boundary of the domain Ω
$\overline{\Omega}$	closure of the set Ω
$\Omega' \subset\subset \Omega$	$\overline{\Omega'} \subset \Omega$ and $\overline{\Omega'}$ is compact.
$L^p(\Omega)$	space of Lebesgue measurable functions f and $\int_{\Omega} f(x) ^p dx < +\infty$
$\ x\ _p$	norm of a functional x in the space $L^p(\Omega)$
$L^\infty(\Omega)$	space of bounded functions almost every Ω
$\ x\ _\infty$	norm of a functional x in the space $L^\infty(\Omega)$
$C(\overline{\Omega})$	space of continuous functions on $\overline{\Omega}$
$C^k(\Omega)$	space of k times continuously differentiable functions on Ω , $k \geq 0$
$C^\infty(\Omega)$	$= \cap_{k \geq 0} C^k(\Omega)$
$C^{0,\alpha}(\Omega)$	$= \left\{ f \in C(\Omega) : \sup_{x,y \in \Omega, x \neq y} \frac{ f(x)-f(y) }{ x-y ^\alpha} < \infty \right\},$ $0 < \alpha \leq 1$
$C^{k,\alpha}(\Omega)$	$= \{ f \in C^k(\Omega) : D^j f \in C^{0,\alpha}(\Omega) \ \forall j, \ j \leq k \}$

$\begin{cases} W^{m,p}(\Omega), W_0^{m,p}(\Omega), \\ W^{s,r}(\Gamma), \\ H^m(\Omega), H_0^m(\Omega) \end{cases}$	Sobolev spaces
$\ x\ _{m,p}$	norm of a functional x in the space $W^{m,p}(\Omega)$
$\ x\ _{2,2}, \ x\ _{1,2}$	norm of a functional x in the space $H^2(\Omega)$, $H_0^1(\Omega)$, respectively
p', q', r', s'	conjugate of $p, q, r, s \in [1, \infty]$, respectively, that is, $\frac{1}{p} + \frac{1}{p'} = \frac{1}{q} + \frac{1}{q'} = \frac{1}{r} + \frac{1}{r'} = \frac{1}{s} + \frac{1}{s'} = 1$
$W^{-m,p'}(\Omega)$	the dual space of $W^{m,p}(\Omega)$
$X \hookrightarrow Y$	X is continuously embedded in Y
$X \hookrightarrow\hookrightarrow Y$	X is compactly embedded in Y
i.e.	id est (that is)
a.e.	almost everywhere
s.t.	subject to
p. 5	page 5
w.r.t.	with respect to
$\square$	The proof is complete

PREFACE

Nonlinear and variational analysis play crucial roles in many fields of applied mathematics and computer science, such as optimization and optimal control, differential equations, mechanics, and machine learning. Their significance can be explained by the following aspects.

In the field of optimization and optimal control theory, nonlinear and variational analysis provides mathematical foundations and tools for optimization problems such as the Ekeland variational principle, tangent cones, normal cones and separation theorems. Based on these tools, we can derive necessary optimality conditions for locally optimal solutions to mathematical programming problems with constraints in Banach spaces. We then solve the system of necessary optimality conditions of KKT-type to find optimal solutions. This framework can be applied for optimal control problems and problems in machine learning.

In the field of differential equations and dynamical systems, nonlinear and variational analysis provides variational principles such as the Lax–Milgram and variational inequalities to solve partial differential equations (PDEs) by finding weak solutions in appropriate function spaces (e.g., Sobolev spaces). Many real-world phenomena are governed by nonlinear differential equations, such as the models of fluid dynamics, biological systems, and electrical circuits. The class of these problems can be solved by using tools and techniques of nonlinear and variational analysis such as the degree theory and fixed-point theorems. Moreover, techniques of monotone operator theory and convex analysis help us to find solutions of operator equations in Hilbert and Banach spaces. These models can be applied for nonlinear partial differential equations.

In the field of machine learning and data science, techniques of nonlinear optimization and numerical methods, such as gradient method and subgradient method, are fundamental for training of deep learning models.

By the importance of nonlinear analysis and variational analysis, I began studying nonlinear analysis in 1998 under the guidance of my supervisor, Professor Nguyen Dong Yen at Hanoi Institute of Mathematics. In this field, I mainly studied properties of duality mappings and uniformly convex Banach spaces, the continuity of metric projections, de-

gree theory, the existence of solutions to variational inequalities, and the Galerkin method for variational inequalities with monotone operators.

I started to study variational analysis in 2009. In this area I focused on properties of subdifferential of non-convex functions, tangent cones, second-order tangent sets, metric regularity, constraint qualification conditions, first- and second-order optimality conditions for optimization problems in Banach spaces. We cite here some important documents which have close connections to our work: [11], [26], [28] and [166].

Nowadays, along with artificial intelligence (AI), automatic control systems (ACS), typically robots, are playing an important role in the Fourth Industrial Revolution. In such a control system, we often need to optimize energy and time. This leads to optimal control problems. To solve such problems, we have to derive optimality conditions and numerical methods of finding approximate optimal solutions to optimal control problems. Recently, these research directions for optimal control problems have been studied by many mathematicians because they have applications in industry, economics, aerospace science and robotics (see, for instance, [157] and [186]).

Motivated by applications of optimal control in the modern technology, recently I have studied optimal control problems governed by ordinary differential equations (ODEs) and partial differential equations (PDEs). In this field, I mainly study the existence of optimal solutions, first- and second-order optimality conditions for locally Pareto solutions to multiobjective optimal control problem governed by ODEs, first and second-order optimality conditions for optimal control problems governed by PDEs and numerical solutions for optimal solutions to semilinear elliptic control problems. To deal with optimal control problems, we often consider the problems under an angle and a view of nonlinear and variational analysis. Namely, we can consider optimal control problems as specific mathematical programming problems with equality constraint and inclusion constraint in Banach spaces. By using tools of nonlinear and variational analysis we derive optimality conditions for locally optimal solutions to the problems in Banach spaces and then apply the obtained results to optimal control problems.

This habilitation dissertation summarizes and presents some obtained results of my work in areas of nonlinear analysis, variational analysis, optimal control and their connections. It has four chapters and a list of references.

Chapter 1 presents results in nonlinear analysis which consists of duality mappings, metric projections, degree theory, and variational inequalities. The content of this chapter is based on our published papers [120], [124], [125], [126], [127] and [134].

Chapter 2 provides results on the variational analysis which consists of subdifferentials, tangent cones, second-order tangent sets, metric regularity, constraint qualification conditions and optimality conditions for vector optimization problems in Banach spaces. The content of this chapter is based on our published papers [2], [129] and [132].

Chapter 3 presents results for optimal control problems governed by ordinary differential equations. The results consist of the existence of solutions, second-order optimality conditions of KKT type for multiobjective optimal control problems with free right endpoints, and second-order optimality conditions of KKT type for multiobjective optimal control problems with endpoint and mixed constraints. The content of this chapter is based on our published papers [2], [132] and [138].

Chapter 4 provides results for optimal control problems governed by partial differential equations. The results consist of Pontryagin's maximum principle for optimal control problems governed by 3D Navier-Stokes equations, second-order optimality conditions and regularity of Lagrange multipliers for semilinear parabolic optimal control problems, numerical solutions for semilinear elliptic optimal control with nonlinear and mixed constraints. The content of this chapter is based on our published papers [130], [118] and [136].

In October 2024, I had a one-month research visit at the XLIM Institute, supported by the Federation de Recherche Margaux. During this stay, I collaborated with Professor Samir Adly on a research project and we had a joint manuscript which has been under review in a mathematical journal.

The research atmosphere at XLIM was highly stimulating, fostering fruitful discussions and high-level collaborations in mathematical optimization and variational analysis. Additionally, during my visit, I had the opportunity to discuss with Professor Samir Adly the possibility of defending my Habilitation à Diriger des Recherches (HDR) at the Université de Limoges. The University of Limoges and other universities in France have many excellent experts in the fields of variational analysis, optimization and applied mathematics. This makes an ideal environment

for research and scientific cooperations in universities of France. Defending my HDR at the University of Limoges would not only strengthen my collaboration with the French mathematical community but also pave a way for future joint research initiatives between my home institution and XLIM. As a result, the habilitation dissertation was started writing in the Autumn 2024.

The author gratefully acknowledges many helpful suggestions of Professor Samir Adly during the preparation of the dissertation.

The author appreciates the assistance and support of the Department of Mathematics, University of Limoges.

The author would like to thank Professor Arnd Rösch for his help and cooperation during the time the author does research from 2013 to 2026 at the Faculty of Mathematics, University of Duisburg-Essen under the support of the Alexander von Humboldt Foundation.

The author gratefully acknowledges to the Alexander von Humboldt Foundation for their financial support during the time the author does research in Germany from 2013 to 2026.

The author would like to express his sincere gratitude to Professor Nguyen Dong Yen and the Institute of Mathematics, Vietnam Academy of Science and Technology for their encouragement and support which facilitate the author defending his habilitation dissertation in mathematics at the University of Limoges, France.

B. T. Kien

January, 2026

Hanoi

CHAPTER 1

RESULTS IN NONLINEAR ANALYSIS

In this chapter we shall present some results in nonlinear analysis. Its content is based on our published papers [120]– [134].

1.1 Duality mappings and metric projections

1.1.1 Normalized duality mapping J_p

The concept of duality mapping was introduced by Beurling and Livingston [36, p. 407] in a geometric form. A slightly extended version of the concept was proposed by Asplund [9] who showed how the duality mappings can be characterized via the subdifferential of convex functions. It is well known that the geometric properties of a Banach space X correspond to the analytic properties of the duality mapping $J : X \rightarrow 2^{X^*}$. For instance, X is strictly convex if and only if J is strictly monotone; X^* is uniformly convex if and only if J is single-valued and uniformly continuous on any bounded subset of X .

The duality mapping has many other applications in the study of metric projections and variational inequalities. For example, in some recent papers (see, for instance, [73], [183], [184]), it has been used as the main tool for studying the continuity of the metric projection.

Together with J , the normalized duality mapping J_p with $1 < p < +\infty$, is also considered (see [184]). In several cases where we have to work with problems in uniformly convex Banach spaces, like $L^p(\Omega)$ and $W_p^m(\Omega)$, J_p is more suitable than J . This is because in the spaces we often have to deal with functions which are expressed via certain exponents of degree p . On the other hand, the metric projection in such spaces can be defined via J_p , so analytical properties of J_p gives us a better understanding of the behavior of the metric projection. Therefore, it is of interest to study various analytical properties of J_p .

In this section, we will give the definition of the normalized duality mapping J_p , derive a formula for computing the map and show how the metric projection can be characterized via J_p . We also obtain the first

characteristic property of uniformly convex Banach spaces by using J_p which is an extended version of a theorem due to J. Prüß [188]. Based on this characterization, we obtain a result of the continuity of the metric projection onto a family of closed convex sets in a uniformly convex Banach space. In particular, it is shown that Lemma 1.1 from [229] on the continuity of the metric projection in Hilbert spaces can be extended to the case of the function spaces $L^p(\Omega, \mu)$, $p > 1$.

Let us assume that X is a normed space or a Banach space with the dual X^* . Symbol $\overline{B}_X$ stands for the closed unit ball in X . Let K be a closed convex set in X and $y \in X$. We denote by $P_K(y)$ the projection of y onto K , that is, $P_K(y) \in K$ and

$$\|y - P_K(y)\| = d(y, K) := \inf_{z \in K} \|y - z\|.$$

The set

$$N_K(x) = \begin{cases} \{x^* \in X^* : \langle x^*, y - x \rangle \leq 0, \forall y \in K\}, & \text{if } x \in K, \\ \emptyset, & \text{if } x \notin K \end{cases} \quad (1.1.1)$$

is called the normal cone to K at x . For a convex function $\varphi : X \rightarrow \mathbb{R} \cup \{+\infty\}$, the set

$$\partial\varphi(x) := \{x^* \in X^* : \varphi(z) - \varphi(x) \geq \langle x^*, z - x \rangle, \forall z \in X\}$$

is called the subdifferential of φ at x . For each $p > 1$, we set

$$\varphi_p(x) = \frac{1}{p} \|x\|^p.$$

Definition 1.1.1. ([184]) *The set-valued map $J_p : X \rightarrow 2^{X^*}$ defined by $J_p(x) = \partial\varphi_p(x)$, $\forall x \in X$, is called the normalized duality mapping of X . In the special case where $p = 2$, the map J_2 is denoted by J .*

Remark 1.1.2. For every $x \in X$ we have $J_p(x) \neq \emptyset$ because $\varphi_p(x)$ is a continuous convex function on X . In particular, the effective domain

$$\text{Dom}(J_p) := \{x : J_p(x) \neq \emptyset\}$$

coincides with X (see [108]). Moreover, for every $x \in X$, $J_p(x)$ is a nonempty, convex, weakly* compact set.

For computing the set $J_p(x)$ one can use the following fact.

Proposition 1.1.3. *For each $x \in X$, it holds*

$$J_p(x) = \{x^* \in X^* : \langle x^*, x \rangle = \|x\| \|x^*\|, \quad \|x^*\| = \|x\|^{p-1}\}. \quad (1.1.2)$$

Proof. For any $x \in X$ we set

$$A(x) = \{x^* \in X^* : \langle x^*, x \rangle = \|x\| \|x^*\|, \quad \|x^*\| = \|x\|^{p-1}\}. \quad (1.1.3)$$

We have to show that $J_p(x) = A(x)$. We first consider the case $x = 0$. From the inequality $p > 1$ it follows that

$$J_p(0) = \partial\varphi_p(0) = \{x^* : \langle x^*, y \rangle \leq \frac{1}{p} \|y\|^p, \quad \forall y \in X\} = \{0\}.$$

Besides, it is obvious that $A(0) = \{0\}$. Hence $J_p(0) = A(0)$.

Next, we consider the case $x \neq 0$. Taking any $x^* \in J_p(x)$, we have

$$\frac{1}{p} \|y\|^p - \frac{1}{p} \|x\|^p \geq \langle x^*, y - x \rangle \quad (1.1.4)$$

for all $y \in X$. Substituting $y = \lambda x$, $\lambda > 1$, into (1.1.4), we get

$$\frac{1}{p} \frac{\lambda^p - 1}{\lambda - 1} \|x\|^p \geq \langle x^*, x \rangle.$$

Letting $\lambda \rightarrow 1$ and noting that $\lim_{\lambda \rightarrow 1} \frac{\lambda^p - 1}{\lambda - 1} = p$, we obtain

$$\|x\|^p \geq \langle x^*, x \rangle. \quad (1.1.5)$$

Similarly, substituting $y = (1 - \lambda)x$, $\lambda > 0$, into (1.1.4) and letting $\lambda \rightarrow 0$, we get

$$\|x\|^p \leq \langle x^*, x \rangle. \quad (1.1.6)$$

Combining (1.1.5) with (1.1.6) yields $\|x\|^p = \langle x^*, x \rangle$ and hence $\|x^*\| \geq \|x\|^{p-1}$. On the other hand, for any y satisfying $\|y\| \leq \|x\|$, by (1.1.4) we have $0 \geq \langle x^*, y - x \rangle$. Therefore, $\langle x^*, x \rangle \geq \langle x^*, y \rangle$. Consequently, $\|x\|^p \geq \langle x^*, y \rangle$. This implies

$$\|x\|^{p-1} \geq \langle x^*, \frac{y}{\|x\|} \rangle.$$

Hence, $\|x\|^{p-1} \geq \langle x^*, z \rangle$ for all z satisfying $\|z\| \leq 1$. It follows that $\|x\|^{p-1} \geq \|x^*\|$. From what has already been said, we deduce that $\|x\|^{p-1} = \|x^*\|$. We have thus proved that $J_p(x) \subset A(x)$. To prove the reverse inclusion, we fix any $x^* \in A(x)$. Consider the function $f(t) = -\ln t$, $t > 0$. Since $f''(t) = \frac{1}{t^2} > 0$ for all $t > 0$, $f(t)$ is a convex function. Therefore, we have

$$-\ln \left(\frac{1}{p} \|y\|^p + \frac{p-1}{p} \|x\|^p \right) \leq -\frac{1}{p} \ln(\|y\|^p) - \frac{p-1}{p} \ln(\|x\|^p).$$

for all $y \in X \setminus \{0\}$. This implies

$$\frac{1}{p}\|y\|^p + \frac{p-1}{p}\|x\|^p \geq \|y\|\|x\|^{p-1}.$$

for all $y \in X \setminus \{0\}$. Hence

$$\begin{aligned} \frac{1}{p}\|y\|^p - \frac{1}{p}\|x\|^p &\geq \|y\|\|x\|^{p-1} - \|x\|^p \\ &\geq \|y\|\|x^*\| - \langle x^*, x \rangle \\ &\geq \langle x^*, y - x \rangle \end{aligned}$$

for all $y \in X$. Thus, $\varphi_p(y) - \varphi_p(x) \geq \langle x^*, y - x \rangle$ for all $y \in X$. This shows that $x^* \in J_p(x)$. The inclusion $A(x) \subset J_p(x)$ has been established, and the proof is complete. $\square$

Remark 1.1.4. From (1.1.2), it follows that $J_p(tx) = t^{p-1}J_p(x)$ for every $p > 1, x \in X$ and $t > 0$. From (1.1.2), it also follows that $J_p(x) = \|x\|^{p-2}J(x)$ for every $p > 1$ and $x \in X$.

The metric projection in Banach spaces can be characterized by using the normalized duality mapping J_p .

Proposition 1.1.5. *Suppose that X is a Banach space, K is a nonempty closed convex set in X , and $y \in X$. Then $x = P_K(y)$ if and only if*

$$0 \in J_p(x - y) + N_K(x). \quad (1.1.7)$$

Proof. It is clear that $x = P_K(y)$ if and only if x is a solution of the problem

$$\begin{cases} f(z) := \frac{1}{p}\|z - y\|^p + i_K(z) \longrightarrow \inf \\ \text{subject to } z \in X, \end{cases}, \quad (1.1.8)$$

where $i_K(z)$ is the indicator function of K , i.e.,

$$i_K(z) = \begin{cases} 0, & \text{if } z \in K, \\ +\infty, & \text{if } z \notin K. \end{cases}$$

Since $f(\cdot)$ is a convex function, x is a solution of (1.1.8) if and only if $0 \in \partial f(x)$. By the Moreau-Rockafellar theorem (see [108, p. 200, Theorem 1]), the inclusion $0 \in \partial f(x)$ is equivalent to the following one:

$$0 \in J_p(x - y) + N_K(x).$$

The proof is complete. $\square$

Definition 1.1.6. ([238, p. 256]) A Banach space X is called *strictly convex* if for all $x, y \in X$, $\|x\| = \|y\| = 1$ and $x \neq y$, and for all $\lambda \in (0, 1)$, it holds $\|\lambda x + (1 - \lambda)y\| < 1$. A Banach space X is called *uniformly convex* if for every ε , $0 < \varepsilon \leq 2$, there exists a $\delta > 0$ for which $\|x\| = 1$, $\|y\| = 1$ and $\|x - y\| \geq \varepsilon$ imply $\|\frac{x+y}{2}\| \leq 1 - \delta$.

Lemma 1.1.7. ([15, p. 42, Proposition 2.13]) A Banach space X is *strictly convex* if and only if the following equivalent properties hold:

(a) If $\|x + y\| = \|x\| + \|y\|$ and $x \neq 0$, there exists $t \geq 0$ such that $y = tx$.

(b) If $\|x\| = \|y\| = 1$ and $x \neq y$, then $\|\frac{x+y}{2}\| < 1$.

Proposition 1.1.8. ([109, p. 251, Theorem 8.1.18]) Let X be a Banach space such that the dual space X^* is strictly convex Banach space. Then, for every $p > 1$, the mapping J_p is single-valued map.

Proof. Taking any $x \in X$ and consider the set $J_p(x)$. If $x = 0$ then $J_p(0) = \{0\}$. Suppose that $x \neq 0$ and $x_1^*, x_2^* \in J_p(x)$. Then $\|x_1^*\| = \|x_2^*\| = \|x\|^{p-1}$ and

$$\begin{aligned} 2\|x_1^*\|\|x\| &= 2\|x\|^p = \langle x_1^* + x_2^*, x \rangle \\ &\leq \|x_1^* + x_2^*\|\|x\|. \end{aligned}$$

Hence $2\|x_1^*\| \leq \|x_1^* + x_2^*\|$. From this, it follows that $\|x_1^*\| + \|x_2^*\| \leq \|x_1^* + x_2^*\|$. This implies that $\|x_1^*\| + \|x_2^*\| = \|x_1^* + x_2^*\|$. By the strict convexity of X^* and the first assertion of Lemma 1.1.7, we get $x_1^* = x_2^*$, as desired. $\square$

It turns out that the uniform convexity of X^* can be characterized via the uniform continuity of the map $J_p(\cdot)$.

Theorem 1.1.9. The dual space X^* of a Banach space X is uniformly convex if and only if J_p is a single-valued map which is uniformly continuous on each bounded subset of X .

To prove this result we shall need the following

Lemma 1.1.10. ([69, p. 36, Theorem 1]) Let X be a Banach space. Then X^* is uniformly convex if and only if the norm of X is uniformly Fréchet differentiable, i.e., the limit $\lim_{\lambda \rightarrow 0} \frac{\|x + \lambda y\| - \|x\|}{\lambda}$ exists uniformly for all $x, y \in S(X)$, where $S(X) = \{x \in X : \|x\| = 1\}$.

Proof of Theorem 1.1.9 (This proof is based on some arguments similar to those of the proof of Theorem 1 in [69, p. 36]).

Necessity. Assume that X^* is uniformly convex. Then X^* is strictly convex. By Proposition 1.1.8, J_p is a single-valued map. It remains to show that J_p is uniformly continuous on each bounded subset of X . According to Remark 1.1.4, we have

$$J_p(x) = \|x\|^{p-2}J(x), \quad \text{for every } x \in X.$$

By Proposition 32.22 of [238, p. 861], J is uniformly continuous on bounded subsets of X . Hence, J_p is uniformly continuous on each bounded subset that lies outside some neighborhood of $x = 0$. On the other hand, $\|J_p(x)\| = \|x\|^{p-1}$, so $J_p(0) = 0$ and J_p is continuous at $x = 0$. Combining these arguments we conclude that J_p is uniformly continuous on each bounded subset of X .

Sufficiency. For any $x, y \in S(X)$ and $\lambda > 0$, using Proposition ??, one has

$$\begin{aligned} \frac{\langle J_p(x), y \rangle}{\|x\|^{p-1}} &= \frac{\langle J_p(x), \lambda y \rangle}{\lambda \|x\|^{p-1}} \\ &= \frac{\langle J_p(x), x \rangle - \|x\|^p + \langle J_p(x), \lambda y \rangle}{\lambda \|x\|^{p-1}} \\ &= \frac{\langle J_p(x), x + \lambda y \rangle - \|x\|^p}{\lambda \|x\|^{p-1}} \\ &\leq \frac{\|x\|^{p-1} \|x + \lambda y\| - \|x\|^p}{\lambda \|x\|^{p-1}} \\ &= \frac{\|x + \lambda y\| - \|x\|}{\lambda} \\ &= \frac{\|x + \lambda y\|^p - \|x\|^p}{\lambda \|x + \lambda y\|^{p-1}} \\ &\leq \frac{\langle J_p(x + \lambda y), x + \lambda y \rangle - |\langle J_p(x + \lambda y), x \rangle|}{\lambda \|x + \lambda y\|^{p-1}} \\ &= \frac{\lambda \langle J_p(x + \lambda y), y \rangle + \langle J_p(x + \lambda y), x \rangle - |\langle J_p(x + \lambda y), x \rangle|}{\lambda \|x + \lambda y\|^{p-1}} \\ &\leq \frac{\langle J_p(x + \lambda y), y \rangle}{\|x + \lambda y\|^{p-1}}. \end{aligned}$$

Hence, for all $x, y \in S(X)$ and $\lambda > 0$, it holds

$$\frac{\langle J_p(x), y \rangle}{\|x\|^{p-1}} \leq \frac{\|x + \lambda y\| - \|x\|}{\lambda} \leq \frac{\langle J_p(x + \lambda y), y \rangle}{\|x + \lambda y\|^{p-1}}.$$

Using the uniform continuity of J_p on each bounded subset of X , from the last property one can obtain that

$$\lim_{\lambda \rightarrow 0} \frac{\|x + \lambda y\| - \|x\|}{\lambda} = \frac{\langle J_p(x), y \rangle}{\|x\|^{p-1}}.$$

According to Lemma 1.1.10, X^* is uniformly convex. The proof is complete. $\square$

1.1.2 Characteristic properties of uniformly convex Banach spaces

In this section, we establish a generalized version of Theorem 1 in [188].

Definition 1.1.11. ([14, p. 708]) *A function $\omega : [0, +\infty) \rightarrow [0, +\infty)$ is said to be firm if $\omega(0) = 0$ and $\omega(\rho) > 0$ for all $\rho > 0$. If ω is nondecreasing and firm, then it is called a gauge. The set of gauges $\omega : [0, +\infty) \rightarrow [0, +\infty)$ is denoted by G .*

Theorem 1.1.12. (Theorem of Prüß, see [188, Theorem 1]) *A Banach space X is uniformly convex if and only if for every $\gamma > 0$, there exists a function $\omega_\gamma \in G$ such that*

$$\langle x^* - y^*, x - y \rangle \geq \omega_\gamma(\|x - y\|)\|x - y\| \quad (1.1.9)$$

for all $x, y \in \overline{B}(0, \gamma)$, $x^* \in J(x)$, and $y^* \in J(y)$.

One may ask whether the conclusion of the above theorem is still true if instead of $J(\cdot)$ one considers the normalized duality mapping $J_p(\cdot)$.

Our main result in this section can be formulated as follows.

Theorem 1.1.13. *If a Banach space X is uniformly convex, then for every $p \geq 2$ and for every $\gamma > 0$, there exists a function $\omega_{p,\gamma} \in G$ such that*

$$\langle x^* - y^*, x - y \rangle \geq \omega_{p,\gamma}(\|x - y\|)\|x - y\| \quad (1.1.10)$$

for all $x, y \in \overline{B}(0, \gamma)$, $x^* \in J_p(x)$, and $y^* \in J_p(y)$. Conversely, if for a fixed $p \geq 2$ and for all $\gamma > 0$, there exists an $\omega_{p,\gamma} \in G$ satisfying (1.1.10), then X is a uniformly convex Banach space.

Remark 1.1.14. *Theorem 1.1.12 follows from Theorem 1.1.13 if we choose $p = 2$.*

To prove Theorem 1.1.13, we shall need the following lemmas.

Lemma 1.1.15. (Bishop-Phelps Theorem, [69, p. 3]) *Let A be a closed bounded convex set in a Banach space X . Then the collection of functionals from X^* that achieve their maximum on A is dense in X^* .*

Lemma 1.1.16. (James Theorem, [69, p. 12]) *A Banach space X is reflexive if and only if for every $f \in X^*$, there exists $x \in X$ such that $f(x) = \|f\|$.*

Lemma 1.1.17. *For all $a \geq 0$, $b \geq 0$, and $p \geq 2$, one has*

$$a^p + b^p - ab^{p-1} - ba^{p-1} \geq |a - b|^p. \quad (1.1.11)$$

Proof. If $a = 0$ or $b = 0$, then (1.1.11) is trivial. We consider the case $a > 0$ and $b > 0$. By symmetry, we assume that $a > b$. Dividing both sides of (1.1.11) by a^p , we obtain the equivalent inequality:

$$1 + \left(\frac{b}{a}\right)^p - \left(\frac{b}{a}\right)^{p-1} - \frac{b}{a} \geq \left(1 - \frac{b}{a}\right)^p. \quad (1.1.12)$$

For $x := \frac{b}{a}$, $x \in (0, 1)$, inequality (1.1.12) is equivalent to

$$\begin{aligned} 1 + x^p - x^{p-1} - x &\geq (1 - x)^p \\ \Leftrightarrow (1 - x)(1 - x^{p-1}) &\geq (1 - x)^p \\ \Leftrightarrow 1 - x^{p-1} &\geq (1 - x)^{p-1} \\ \Leftrightarrow 1 &\geq x^{p-1} + (1 - x)^{p-1} \end{aligned} \quad (1.1.13)$$

Noting that $1 = (x + (1 - x))^{p-1}$ and using the fact that $(u + v)^\alpha \geq u^\alpha + v^\alpha$ for all $u, v \geq 0$ and $\alpha \geq 1$ (see [99, p. 32]), we conclude that (1.1.13) is true. The proof is complete. $\square$

Proof of Theorem 1.1.13.

Necessity. Let X be a uniformly convex Banach space and let $p \geq 2$. For each $\gamma > 0$, we consider the function $\omega_{p,\gamma}$ defined by

$$\omega_{p,\gamma}(0) = 0, \quad \omega_{p,\gamma}(\rho) = \omega_{p,\gamma}(2\gamma)$$

for all $\rho \geq 2\gamma$ and

$$\omega_{p,\gamma}(\rho) = \inf \left\{ \frac{\langle x^* - y^*, x - y \rangle}{\|x - y\|} : x, y \in \overline{B}(0, \gamma), \|x - y\| \geq \rho, \right. \\ \left. x^* \in J_p(x), y^* \in J_p(y) \right\}$$

for all $0 < \rho \leq 2\gamma$. Since $\omega_{p,\gamma}$ is nondecreasing, in order to prove that $\omega_{p,\gamma}(\rho) \in G$, we have only to show that $\omega_{p,\gamma}(\rho) > 0$ for all $\rho \in (0, 2\gamma)$. Suppose on the contrary that $\omega_{p,\gamma}(\rho) = 0$ for some $\rho \in (0, 2\gamma)$. Then there exist sequences $(x_n), (y_n)$ of vectors from $\overline{B}(0, \gamma)$, $x_n^* \in J_p(x_n)$, $y_n^* \in J_p(y_n)$, such that $\|x_n - y_n\| \geq \rho$ and

$$\frac{\langle x_n^* - y_n^*, x_n - y_n \rangle}{\|x_n - y_n\|} \rightarrow 0.$$

This forces $\langle x_n^* - y_n^*, x_n - y_n \rangle \rightarrow 0$. On the other hand, from the inclusions $x_n^* \in J_p(x_n)$, $y_n^* \in J_p(y_n)$ and Lemma 1.1.17, we have

$$\begin{aligned} \langle x_n^* - y_n^*, x_n - y_n \rangle &= \|x_n\|^p + \|y_n\|^p - \langle x_n^*, y_n \rangle - \langle y_n^*, x_n \rangle \\ &\geq \|x_n\|^p + \|y_n\|^p - \|x_n^*\| \|y_n\| - \|y_n^*\| \|x_n\| \\ &= \|x_n\|^p + \|y_n\|^p - \|x_n\|^{p-1} \|y_n\| - \|y_n\|^{p-1} \|x_n\| \\ &\geq |||x_n| - |y_n|||^p. \end{aligned}$$

This implies $|||x_n| - |y_n|| \rightarrow 0$. Since $\|x_n\| \leq \gamma$ and $\|y_n\| \leq \gamma$, without loss of generality we may assume that $\lim_{n \rightarrow \infty} \|x_n\| = \lim_{n \rightarrow \infty} \|y_n\| = a$ for some $a \geq 0$. Since $\|x_n\| + \|y_n\| \geq \|x_n - y_n\| \geq \rho$ for every n , the case $a = 0$ cannot occur, so we have $a > 0$. Hence there is no loss of generality in assuming that $\|x_n\| > 0$, $\|y_n\| > 0$ for all n . We have

$$\begin{aligned} \left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| &= \left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|x_n\|} + \frac{y_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| \\ &\geq \left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|x_n\|} \right\| - \left\| \frac{y_n}{\|y_n\|} - \frac{y_n}{\|x_n\|} \right\|. \end{aligned}$$

Therefore,

$$\begin{aligned} &\left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|x_n\|} \right\| + \left\| \frac{y_n}{\|y_n\|} - \frac{y_n}{\|x_n\|} \right\| \geq \left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|x_n\|} \right\| \\ \Rightarrow &\left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| + \|y_n\| \left| \frac{y_1}{\|y_n\|} - \frac{x_1}{\|x_n\|} \right| \geq \frac{1}{\|x_n\|} \|x_n - y_n\| \\ \Rightarrow &\liminf_{n \rightarrow \infty} \left(\left\| \frac{\|x_n\|}{\|x_n\|} - \frac{\|y_n\|}{\|y_n\|} \right\| + \|y_n\| \left| \frac{y_1}{\|y_n\|} - \frac{x_1}{\|x_n\|} \right| \right) \\ &\geq \liminf_{n \rightarrow \infty} \left(\frac{1}{\|x_n\|} \|x_n - y_n\| \right) \\ \Rightarrow &\liminf_{n \rightarrow \infty} \left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| \geq \frac{\rho}{a}. \end{aligned} \tag{1.1.14}$$

Choose $\varepsilon_1 \in (0, \frac{\rho}{a})$ and put $\varepsilon = \frac{\rho}{a} - \varepsilon_1$. By (1.1.14), there exists n_0 such that

$$\left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| > \varepsilon$$

for all $n \geq n_0$. Since

$$\begin{aligned} \|x_n + y_n\| &= \|x_n\| \left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|x_n\|} \right\| \\ &= \|x_n\| \left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} - \frac{y_n}{\|y_n\|} \right\| \\ &\leq \|x_n\| \left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} \right\| + \|x_n\| \left\| \frac{y_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| \\ &= \|x_n\| \left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} \right\| + \|x_n\| \|y_n\| \left| \frac{1}{\|x_n\|} - \frac{1}{\|y_n\|} \right|, \end{aligned}$$

by (1.1.14) and the uniform convexity of X , there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\begin{aligned} &\limsup_{n \rightarrow \infty} \|x_n + y_n\| \\ &\leq \limsup_{n \rightarrow \infty} \left(\|x_n\| \left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} \right\| + \|x_n\| \|y_n\| \left| \frac{1}{\|x_n\|} - \frac{1}{\|y_n\|} \right| \right) \\ &\leq \limsup_{n \rightarrow \infty} a \left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} \right\| \\ &\leq 2a(1 - \delta). \end{aligned} \tag{1.1.15}$$

Since $\|x_n\|^p + \|y_n\|^p - \langle y_n^*, x_n \rangle - \langle x_n^*, y_n \rangle = \langle x_n^* - y_n^*, x_n - y_n \rangle \rightarrow 0$,

$$\lim_{n \rightarrow \infty} (\langle y_n^*, x_n \rangle + \langle x_n^*, y_n \rangle) = 2a^p.$$

Combining this with (1.1.15), we have

$$\begin{aligned}
2a^p &= \lim_{n \rightarrow \infty} (\langle y_n^*, x_n \rangle + \langle x_n^*, y_n \rangle) \\
&= \limsup_{n \rightarrow \infty} (\langle y_n^*, x_n \rangle + \langle x_n^*, y_n \rangle + \langle x_n^*, x_n \rangle - \langle x_n^*, x_n \rangle) \\
&= \limsup_{n \rightarrow \infty} (\langle y_n^* - x_n^*, x_n \rangle + \langle x_n^*, x_n + y_n \rangle) \\
&\leq \limsup_{n \rightarrow \infty} (\|y_n^*\| \|x_n\| - \|x_n\|^p) + \limsup_{n \rightarrow \infty} \|x_n^*\| \|x_n + y_n\| \\
&\leq \limsup_{n \rightarrow \infty} (\|y_n\|^{p-1} \|x_n\| - \|x_n\|^p) + \limsup_{n \rightarrow \infty} \|x_n\|^{p-1} \|x_n + y_n\| \\
&= \limsup_{n \rightarrow \infty} \|x_n\|^{p-1} \|x_n + y_n\| \\
&\leq a^{p-1} 2a(1 - \delta) = 2a^p(1 - \delta),
\end{aligned}$$

a contradiction.

Sufficiency. Using Proposition 1.1.3, we can show that $J_p(X) := \bigcup_{x \in X} J_p(x)$ is the collection of functionals from X^* that achieve their maximum on $\overline{B}_X$. By Lemma 1.1.15, $\overline{J_p(X)} = X^*$, where $\overline{J_p(X)}$ denotes the closure of $J_p(X)$. We shall show that the range $J_p(X)$ of J_p is closed. Let $(x_n^*) \subset J_p(X)$ and $x_n^* \rightarrow x^*$. Then for each n there exists $x_n \in X$ such that $x_n^* \in J_p(x_n)$. Since (x_n^*) is bounded, (x_n) is bounded. Let $\gamma > 0$ be such that $\|x_n\| \leq \gamma$ for every n . By our hypothesis, there exists $\omega_{p,\gamma} \in G$ such that

$$\begin{aligned}
\omega_{p,\gamma}(\|x_n - x_m\|) \|x_n - x_m\| &\leq \langle x_n^* - x_m^*, x_n - x_m \rangle \\
&\leq \|x_n^* - x_m^*\| \|x_n - x_m\|
\end{aligned}$$

for all m, n . Taking account of this fact, we deduce from the convergence of the sequence (x_n^*) that (x_n) is a Cauchy sequence. Therefore $x_n \rightarrow x$ for some $x \in X$. Since $x_n^* \in J_p(x_n)$ it follows that $x^* \in J_p(x)$. Thus $J_p(X) = \overline{J_p(X)} = X^*$. Consequently, for every $x^* \in X^* \setminus \{0\}$ there exists $x \in X \setminus \{0\}$ such that $x^* \in J_p(x)$. This implies that

$$\left\langle x^*, \frac{x}{\|x\|} \right\rangle = \frac{\|x\|^p}{\|x\|} = \|x\|^{p-1} = \|x^*\|.$$

By Lemma 1.1.16 we conclude that X is reflexive. In this case we get $x^* \in J_p(x)$ if and only if $x \in J_q^*(x^*)$, where $q > 0$ satisfies the relation $\frac{1}{p} + \frac{1}{q} = 1$ and J_q^* denotes the normalized mapping of X^* . Hence $J_p^{-1} = J_q^*$. Fix any $\varrho > 0$. Fix any $x^*, y^* \in \overline{B}^*(0, \varrho)$, $x \in J_q^*(x^*)$, $y \in J_q^*(y^*)$. Choose γ so that $\gamma^{p-1} = \varrho$. Then $x^* \in J_p(x)$, $y^* \in J_p(y)$, and $x, y \in \overline{B}(0, \gamma)$. By

(1.1.10),

$$\omega_{p,\gamma}(\|x - y\|) \leq \|x^* - y^*\|.$$

This implies that the map J_q^* is single-valued and uniformly continuous on each bounded subset of X^* . By Theorem 1.1.9, $X = (X^*)^*$ is uniformly convex. The proof is complete. $\square$

1.1.3 Continuity of the metric projection onto a family of closed convex sets

From now on, (Λ, d) is a metric space and $K : \Lambda \rightarrow 2^X$ is a set-valued map with nonempty closed convex values. Let $(\bar{\lambda}, \bar{x}) \in \Lambda \times X$ be a point satisfying $\bar{x} \in K(\bar{\lambda})$. For each $y \in X$, the projection of y onto $K(\lambda)$ is denoted by $P_{K(\lambda)}(y)$.

Definition 1.1.18. A set-valued map $K : \Lambda \rightarrow 2^X$ is said to be pseudo-Lipschitz around $(\bar{\lambda}, \bar{x})$ if there exist a positive constant l and neighborhoods U, V of $\bar{x}$ and $\bar{\lambda}$, respectively, such that

$$K(\lambda) \cap U \subset K(\lambda') + ld(\lambda, \lambda')\overline{B}(0, 1) \quad (1.1.16)$$

for all $\lambda, \lambda' \in U$.

Theorem 1.1.19. *Let X be a uniformly convex Banach space, and $p \geq 2$. Let $K : \Lambda \rightarrow 2^X$ be a set-valued map, which is pseudo-Lipschitz around $(\bar{\lambda}, \bar{x})$. Then there exist positive constant l_1 , a neighborhood U_1 of $\bar{x}$, a neighborhood V_1 of $\bar{\lambda}$ and a function $\omega_{p,\gamma} \in G$ such that*

$$\omega_{p,\gamma}(\|P_{K(\lambda)}(y) - P_{K(\lambda')}(y)\|)\|P_{K(\lambda)}(y) - P_{K(\lambda')}(y)\| \leq l_1 d(\lambda, \lambda')$$

for all $\lambda, \lambda' \in V_1$ and $y \in U_1$.

Proof. The proof runs similarly as the proof of Theorem 3.1 in [120], where ω_γ and J are replaced by $\omega_{p,\gamma}$ and J_p , respectively. Instead of using Theorem 1.1.12, which is due to J. Prüß, we use Theorem 1.1.13 proved in the preceding section. $\square$

From Theorem 1.1.13 it follows that the function $\omega_{p,\gamma}$ can be computed by the formula:

$$\omega_{p,\gamma}(t) = \begin{cases} 0, & \text{for } t = 0, \\ \inf \left\{ \frac{\langle x_1^* - x_2^*, x_1 - x_2 \rangle}{\|x_1 - x_2\|} : x_1, x_2 \in B(0, \gamma), \right. \\ \left. \|x_1 - x_2\| \geq t, x_1^* \in J_p(x_1), x_2^* \in J_p(x_2) \right\}, & \text{for } t \in (0, 2\gamma], \\ \omega_{p,\gamma}(2\gamma), & \text{for } t > 2\gamma. \end{cases} \quad (1.1.17)$$

In many situations, in order to apply Theorem 1.1.19, it is enough to have a lower estimate for the function $\omega_{p,\gamma}$. Such an estimate in the case $X = L^p(\Omega, \mu)$, $1 < p \leq 2$, has been obtained in [120]. Here and in the sequel, $X = L^p(\Omega, \mu)$ denotes the space of all measurable functions $x(\cdot)$ defined on an open set $\Omega \subset \mathbb{R}^n$ for which $\int_\Omega |x(s)|^p d\mu < \infty$, where μ stands for the Lebesgue measure in $\mathbb{R}^n$. It is well known that for every $1 < p < +\infty$, this space and its dual $X^* = L^q(\Omega, \mu)$, where $q = \frac{p}{p-1}$, are uniformly convex Banach spaces. We now want to use the normalized duality mapping J_p to obtain an estimate for $\omega_{p,\gamma}$ in the case $X = L^p(\Omega, \mu)$ with $p \geq 2$.

Lemma 1.1.20. ([187, Corollary 2.1]) *Suppose that $x, y \in L_p(\Omega, \mu)$ and $1 < p < +\infty$. Then for every $t \in (0, 1)$ one has*

- (a) $\|tx + (1-t)y\|^2 \leq t\|x\|^2 + (1-t)\|y\|^2 - \frac{p-1}{64}t(1-t)\|x-y\|^2$ for all $1 < p < 2$.
- (b) $\|tx + (1-t)y\|^p \leq t\|x\|^p + (1-t)\|y\|^p - \frac{1}{p2^p}(t(1-t)^p + t^p(1-t))\|x-y\|^p$ for all $p \geq 2$.

Theorem 1.1.21. *Suppose that $X = L_p(\Omega, \mu)$, $1 < p < +\infty$. Then the following properties hold:*

- (a) *If $1 < p \leq 2$, then*

$$\langle J_p(x_1) - J_p(x_2), x_1 - x_2 \rangle \geq \frac{p-1}{64} \|x_1 - x_2\|^2,$$

for all $x_1, x_2 \in X$. Besides, for each $\gamma > 0$, there is a function $\omega_{2,\gamma}$ such that

$$\omega_{2,\gamma}(t) \geq \frac{p-1}{64}t, \quad \forall t \in [0, 2\gamma].$$

- (b) *If $p \geq 2$, then*

$$\langle J_p(x_1) - J_p(x_2), x_1 - x_2 \rangle \geq \frac{1}{p2^{2p-1}} \|x_1 - x_2\|^p,$$

for all $x_1, x_2 \in X$. Besides, for each $\gamma > 0$, there exists a function $\omega_{p,\gamma}$ such that

$$\omega_{p,\gamma}(t) \geq \frac{1}{p^2 2^{p-1}} t^{p-1}, \quad \forall t \in [0, 2\gamma].$$

Proof. The proof of (a) runs similarly to the proof of Proposition 4.1 in [120]. It remains to prove (b). Let $\varphi_p(x) = \frac{1}{p} \|x\|^p$, $p \geq 2$. Since $X = L_p(\Omega, \mu)$ and its dual $X^* = L_q(\Omega, \mu)$, where $q = \frac{p}{p-1}$, are uniformly convex Banach spaces, by Theorem 1.1.9, we have $\partial\varphi_p(x) = \{J_p(x)\}$. By Lemma 1.1.20 (b),

$$\varphi_p(y+t(x-y)) - \varphi_p(y) \leq t(\varphi_p(x) - \varphi_p(y)) - \frac{1}{p^2 2^p} (t(1-t)^p + t^p(1-t)) \|x-y\|^p.$$

Hence,

$$\frac{1}{t} (\varphi_p(y+t(x-y)) - \varphi_p(y)) \leq \varphi_p(x) - \varphi_p(y) - \frac{1}{p^2 2^p} ((1-t)^p + t^{p-1}(1-t)) \|x-y\|^p.$$

Letting $t \rightarrow 0$, we obtain

$$\varphi'_p(y; x-y) \leq \varphi_p(x) - \varphi_p(y) - \frac{1}{p^2 2^p} \|x-y\|^p,$$

for all $x, y \in X$, where $\varphi'_p(y; x-y)$ denotes the directional derivative of φ_p at y in direction $x-y$. Since $\langle J_p(y), x-y \rangle = \varphi'_p(y; x-y)$, we have

$$\langle J_p(y), x-y \rangle \leq \varphi_p(x) - \varphi_p(y) - \frac{1}{p^2 2^p} \|x-y\|^p,$$

for all $x, y \in X$. Consequently, for all $x_1, x_2 \in X$ we have

$$\begin{aligned} \langle J_p(x_2), x_1 - x_2 \rangle &\leq \varphi_p(x_1) - \varphi_p(x_2) - \frac{1}{p^2 2^p} \|x_1 - x_2\|^p, \\ \langle J_p(x_1), x_2 - x_1 \rangle &\leq \varphi_p(x_2) - \varphi_p(x_1) - \frac{1}{p^2 2^p} \|x_2 - x_1\|^p. \end{aligned}$$

Combining these inequalities, we obtain

$$\langle J_p(x_2) - J_p(x_1), x_2 - x_1 \rangle \geq \frac{2}{p^2 2^p} \|x_1 - x_2\|^p.$$

For each $\gamma > 0$ and $t \in (0, 2\gamma)$, using (1.1.17), we have

$$\begin{aligned} \omega_{p,\gamma}(t) &= \\ \inf \left\{ \frac{\langle x^* - y^*, x - y \rangle}{\|x - y\|} : x, y \in \overline{B}(0, \gamma), \|x - y\| \geq t, x^* \in J_p(x), y^* \in J_p(y) \right\} \\ &\geq \inf \left\{ \frac{2}{p^2 2^p} \|x - y\|^{p-1} : x_1, x_2 \in \overline{B}(0, \gamma), \|x - y\| \geq t \right\} \geq \frac{2}{p^2 2^p} t^{p-1}. \end{aligned}$$

The proof of property (b) is complete. $\square$

The following theorem describes a Hölder continuity property of the metric projection onto a family of closed convex sets in the space $L_p(\Omega, \mu)$. This result shows that Lemma 1.1 from [229], which was obtained for the case of the metric projection onto closed convex sets in Hilbert spaces, can be extended to the case of the function spaces $L_p(\Omega, \mu)$, $1 < p < \infty$.

Theorem 1.1.22. *Suppose that $X = L_p(\Omega, \mu)$, $1 < p < \infty$, and $K : \Lambda \rightarrow 2^X$ is a set-valued map, which is pseudo-Lipschitz around $(\bar{\lambda}, \bar{x})$. Then the following assertions hold:*

- (a) *If $1 < p \leq 2$ then there exist a positive constant l_0 , a neighborhood U_1 of $\bar{x}$, and a neighborhood V_1 of $\bar{\lambda}$ such that*

$$\|P_{K(\lambda)}(y) - P_{K(\lambda')}(y)\| \leq l_0 d(\lambda, \lambda')^{\frac{1}{2}}$$

for all $\lambda, \lambda' \in V_1$ and $y \in U_1$.

- (b) *If $p \geq 2$ then there exist a positive constant l_1 , a neighborhood U'_1 of $\bar{x}$, and a neighborhood V'_1 of $\bar{\lambda}$ such that*

$$\|P_{K(\lambda)}(y) - P_{K(\lambda')}(y)\| \leq l_1 d(\lambda, \lambda')^{1/p}$$

for all $\lambda, \lambda' \in V'_1$ and $y \in U'_1$.

Proof. Assertion (a) is just the content of Proposition 4.2 in [120], so it remains to prove (b). By Theorem 1.1.19 and Theorem 1.1.21, there exist a constant $\bar{l} > 0$ and neighborhoods U'_1 and V'_1 of $\bar{x}$ and $\bar{\lambda}$, respectively, such that

$$\frac{1}{p^2 2^{p-1}} \|P_K(\lambda)(y) - P_K(\lambda')(y)\|^p \leq \bar{l} d(\lambda, \lambda')$$

for all $\lambda, \lambda' \in V'_1$ and $y \in U'_1$. Hence,

$$\|P_K(\lambda)(y) - P_K(\lambda')(y)\| \leq (\bar{l} p^2 2^{p-1})^{1/p} d(\lambda, \lambda')^{1/p}.$$

Setting $l_1 = (\bar{l} p^2 2^{p-1})^{1/p}$ we obtain the desired conclusion. $\square$

1.2 Topological degree for monotone operators

1.2.1 Topological degree in finite dimensional spaces

Topological degree, also known as degree theory has developed as a means to examine the solution existence and to estimate the number of solutions to a nonlinear equation $f(x) = y$ in a region.

Suppose that X and Y are topological spaces. Let $\mathcal{O}$ be a class of open subsets Ω in X . For each Ω in $\mathcal{O}$, we denote by $\bar{\Omega}$ and $\partial\Omega$ the closure and boundary of Ω , respectively. Let $\mathcal{F}$ be a family of maps $f : \bar{\Omega} \rightarrow Y$ over which the degree theory is to be defined. For each Ω in $\mathcal{O}$, we consider a family of homotopies $\{f_t : 0 \leq t \leq 1\}$ of maps in $\mathcal{F}$, all having the common domain $\bar{\Omega}$. The collection of all such homotopies for the various Ω in $\mathcal{O}$ will be denoted by $\mathcal{H}$, which is called a class of permissible homotopies for the degree theory.

Definition 1.2.1. *By a degree theory over the class $\mathcal{F}$, which is invariant with respect to the homotopies in $\mathcal{H}$ and which is normalized by a given map $f_0 : X \rightarrow Y$, we mean: for each y_0 in Y and for each $f : \bar{\Omega} \rightarrow Y$ in $\mathcal{F}$, with $y_0 \notin f(\partial\Omega)$, an integer $d(f, \Omega, y_0)$ is defined and satisfies the following three conditions:*

(a) *(Normalization) If $d(f, \Omega, y_0) \neq 0$ then there exists x in Ω such that $f(x) = y_0$. For each Ω in $\mathcal{O}$, we have $f_0|_{\Omega} \in \mathcal{F}$, and if $y_0 \in f_0(\Omega)$ then $d(f_0, \Omega, y_0) = +1$.*

(b) *(Additivity) If $f \in \mathcal{F}$, $f : \bar{\Omega} \rightarrow Y$ and that Ω_1, Ω_2 are a pair of disjoint sets in $\mathcal{O}$ contained in Ω such that $y_0 \notin f(\bar{\Omega} \setminus (\Omega_1 \cup \Omega_2))$ then*

$$d(f, \Omega, y_0) = d(f, \Omega_1, y_0) + d(f, \Omega_2, y_0).$$

(c) *(Homotopy invariance) If $\{f_t : 0 \leq t \leq 1\}$ is a homotopy in $\mathcal{H}$ with domain $\bar{\Omega}$ and if $\{y_t : 0 \leq t \leq 1\}$ is a continuous path in Y with $y_t \notin f_t(\partial\Omega)$ for all t in $[0, 1]$ then $d(f_t, \Omega, y_t)$ is independent of t .*

The existence of a degree theory for a given class of mappings is not a trivial assertion. In this subsection we will build a topological degree for a class of continuous mappings in finite dimensional spaces. It is a platform to define the Leray-Schauder degree theory for compact operators in Banach spaces as well as the Browder degree theory for maps of class $(S)_+$ from a bounded open subset of a reflexive Banach space into its dual.

Let us assume that Ω be an open bounded set in $\mathbb{R}^n$, $\phi : \Omega \rightarrow \mathbb{R}^n$ is a mapping of class $C^1(\bar{\Omega})$ and $p \in \mathbb{R}^n$. Let

$$J_\phi(x) := \det \phi'(x) \quad \text{and} \quad Z_\phi = \{x \in \Omega : J_\phi(x) = 0\}$$

be the critical set and $\phi(Z_\phi)$ be the crease of ϕ .

Theorem 1.2.2. ([151, Theorem 1.1.2]) *Suppose $\phi \in C^1(\bar{\Omega})$ and $p \notin \phi(Z_\phi) \cup \phi(\partial\Omega)$. Then $\phi^{-1}(p)$ is finite.*

Proof. Let $S := \phi^{-1}(p)$. We assume that $S \neq \emptyset$. Then S is a closed set. We now claim that any point $x \in S$ is an isolated point, that is, there is a neighborhood V_x such that $V_x \cap S = \{x\}$. Indeed, if this were not true then there exists $x_0 \in S$ and a sequence $\{x_n\} \subset S$ such that x_n converges to x_0 . Since $\phi(x_0) = p$ and $p \notin \phi(Z_\phi) \cup \phi(\partial\Omega)$, $\det(\phi'(x_0)) \neq 0$. Hence

$$r := \min_{u \in S^n} |\phi'(x_0)u| > 0,$$

where S^n is the unit sphere in $\mathbb{R}^n$. This implies that

$$|\phi'(x_0)u| \geq r|u| \quad \forall u \in \mathbb{R}^n. \quad (1.2.1)$$

On the other hand we have

$$0 = \phi(x_n) - \phi(x_0) = \phi'(x_0)(x_n - x_0) + o(|x_n - x_0|).$$

It follows that

$$|\phi'(x_0)(x_n - x_0)| < \frac{r}{2}|x_n - x_0|$$

for sufficiently large n . But this contradicts (1.2.1). Our claim is justified.

Since $S \subset \overline{\Omega}$, it is a compact set. Since $S = \cup_{x \in S} (V_x \cap S)$, we see that, there exist $x_1, x_2, \dots, x_n$ such that $S = \cup_{i=1}^n (V_{x_i} \cap S) = \{x_1, x_2, \dots, x_n\}$. The theorem is proved. $\square$

Definition 1.2.3. Suppose that $\phi \in C^1(\overline{\Omega})$, $p \notin \phi(\partial\Omega) \cup \phi(Z_\phi)$. The degree of ϕ at p relative to Ω is denoted by $d(\phi, \Omega, p)$ and given by

$$d(\phi, \Omega, p) = \sum_{x \in \phi^{-1}(p)} \text{sign} J_\phi(x),$$

where the sum equals to zero when $\phi^{-1}(p) = \emptyset$.

From the definition we see that if $d(\phi, \Omega, p) = m \neq 0$, then the equation $f(x) = p$ has at least $|m|$ solutions in Ω . Moreover, $d(\phi, \Omega, p)$ is normalized by the identity map.

Proposition 1.2.4. Let I denote the identity mapping. If $p \in \Omega$, then $d(I, \Omega, p) = 1$. If $p \notin \overline{\Omega}$, then $d(I, \Omega, p) = 0$.

Next, we define $d(\phi, \Omega, p)$ for the case where $p \in \phi(Z_\phi)$.

Definition 1.2.5. Suppose that $\phi \in C^1(\overline{\Omega})$, $p \notin \phi(\partial\Omega)$ but $p \in \phi(Z_\phi)$. The degree $d(\phi, \Omega, p)$ is given by $d(\phi, \Omega, q)$, where $q \notin \phi(Z_\phi)$ and $|q - p| < \text{dist}(p, \phi(\partial\Omega))$.

To justify this definition, we need to use Theorem 1.2.1 (Sard's theorem) and Theorem 1.3.2 in [151]).

Definition 1.2.6. A C^k homotopy between $\phi, \psi \in C^k(\Omega)$ is a function $H : [0, 1] \times \overline{\Omega} \rightarrow \mathbb{R}^n$ such that $H(0, x) = \phi(x)$, $H(1, x) = \psi(x)$, $H(t, \cdot) \in C^k(\overline{\Omega})$ for all $t \in [0, 1]$ and $\|H(s, \cdot) - H(t, \cdot)\|_k \rightarrow 0$ as $s \rightarrow t$.

For an example $H(t, x) = (1 - t)\phi(x) + t\psi(x)$ is a C^k homotopy whenever $\phi, \psi \in C^k(\Omega)$.

Theorem 1.2.7. ([151, Theorem 1.3.5]) Let $\phi \in C^1(\Omega)$.

- (i) $d(\phi, \Omega, \cdot)$ is constant on connected component of $\mathbb{R}^n \setminus \phi(\partial\Omega)$.
- (ii) If $p \notin \phi(\partial\Omega)$, then there exists $\epsilon > 0$ such that $d(\phi, \Omega, p) = d(\psi, \Omega, p)$ when $\|\psi - \phi\|_1 < \epsilon$.
- (iii) Let $H(x, t)$ be a C^1 homotopy between ϕ and ψ . If $p \notin H(\partial\Omega, t)$ for all $t \in [0, 1]$, then $d(\phi, \Omega, p) = d(\psi, \Omega, p)$.

We are now ready to define a topological degree for continuous mappings.

Definition 1.2.8. Suppose that $\phi \in C(\overline{\Omega})$ and $p \notin \phi(\partial\Omega)$. The degree $d(\phi, \Omega, p)$ is given by $d(\psi, \Omega, p)$, where $\psi \in C^1(\overline{\Omega})$ such that $\|\psi - \phi\|_0 < \text{dist}(p, \phi(\partial\Omega))$.

Justification. Let $\epsilon = \text{dist}(p, \phi(\partial\Omega))$. Then there exists a function $\psi \in C^1(\Omega)$ such that $\|\psi - \phi\|_0 < \epsilon$. Hence for any $x \in \partial\Omega$, we have

$$|p - \phi(x)| \leq |p - \psi(x)| + |\psi(x) - \phi(x)| < |p - \psi(x)| + \text{dist}(p, \phi(\partial\Omega)).$$

This implies that $|p - \psi(x)| > 0$ for all $x \in \partial\Omega$. Hence $p \notin \psi(\partial\Omega)$ and so $d(\psi, \Omega, p)$ is well defined. We now assume that ψ_1, ψ_2 are two mappings of $C^1(\Omega)$ such that $\|\psi_i - \phi\|_0 < \epsilon$ for $i = 1, 2$. Consider C^1 homotopy

$$h_t(x) = t\psi_1(x) + (1 - t)\psi_2(x) \quad x \in \overline{\Omega}, 0 \leq t \leq 1.$$

Then we get

$$\begin{aligned} |h_t(x) - \phi(x)| &= |t(\psi_1(x) - \phi(x)) + (1 - t)(\psi_2(x) - \phi(x))| \\ &< t\epsilon + (1 - t)\epsilon = \epsilon. \end{aligned}$$

Hence if $x \in \partial\Omega$, then

$$|p - h_t(x)| \geq |p - \phi(x)| - |\phi(x) - h_t(x)| > 0.$$

Thus $p \notin h_t(\partial\Omega)$ for all $t \in [0, 1]$. By (iii) of Theorem 1.2.7, we have

$$d(h_0, \Omega, p) = d(h_1, \Omega, p) \Leftrightarrow d(\psi_1, \Omega, p) = d(\psi_2, \Omega, p).$$

We conclude that $d(\psi, \Omega, p)$ is the same for all $\psi \in C^1(\overline{\Omega})$ satisfying $\|\psi - \phi\|_0 < \epsilon$.

The following theorem gives fundamental properties of degree theory in finite dimensional spaces. Its proof can be found in [151].

Theorem 1.2.9. *Let $\phi \in C(\overline{\Omega})$ and $p \in \mathbb{R}^n$.*

(a) (Normalization) *If $p \in \Omega$, then $d(I, \Omega, p) = 1$.*

(b) (Existence) *If $p \notin \phi(\partial\Omega)$ and $d(\phi, \Omega, p) \neq 0$, then there exists $x \in \Omega$ such that $f(x) = p$.*

(c) (Additive) *If $\phi \notin \phi(\partial\Omega)$ and Ω is a union of two disjoint open sets Ω_1, Ω_2 . Then*

$$d(\phi, \Omega, p) = d(\phi, \Omega_1, p) + d(\phi, \Omega_2, p).$$

(d) (Homotopy invariance) *If $H(t, x)$ is a homotopy and $p \notin H(t, \partial\Omega)$ for all $t \in [0, 1]$, then $d(H(t, \cdot), \Omega, p)$ is independent of $t \in [0, 1]$. In particular, $d(H(1, \cdot), \Omega, p) = d(H(0, \cdot), \Omega, p)$.*

(e) (Excision) *If D_0 is a closed set of D and $p \notin \phi(D_0)$ then $d(\phi, D, p) = d(\phi, D \setminus D_0, p)$.*

Here is an application of degree theory in fixed point theorems.

Theorem 1.2.10. ([151, Theorem 4.4.2]) *Let S be a bounded, closed, nonempty convex subset of $\mathbb{R}^n$ with $0 \in \text{int}S$ and $\phi : S \rightarrow S$ is continuous. Then there exists $x_0 \in S$ such that $\phi(x_0) = x_0$.*

Proof. Let $D = \text{int}S$. Then $\bar{D} = S$ and $\partial D = \partial S$. If there exists $x_0 \in \partial D$ such that $\phi(x_0) = x_0$, we obtain the conclusion. Therefore, we may assume that $x - \phi(x) \neq 0$ for all $x \in \partial D$. Consider the homotopy $H(t, x) = x - t\phi(x)$ with $t \in [0, 1]$. For $t = 1$, we have $H(1, x) = x - \phi(x) \neq 0$ for all $x \in \partial D$. For $0 \leq t < 1$, we have $t\phi(x) \in \text{int}S = D$. It follows that $0 \notin H(t, \partial D)$ for all $t \in [0, 1]$. Since degree is independent of t for homotopy, we have

$$d(H(1, \cdot), D, 0) = d(H(0, \cdot), D, 0) = d(I, D, 0) = 1.$$

Hence there exists $x_0 \in D$ such that $H(1, x_0) = 0$ or $x_0 - \phi(x_0) = 0$. $\square$

In the above theorem we required that $0 \in \text{int}S$. To omit this condition, we need the following lemma which is due to Dugundji.

Lemma 1.2.11. ([151, Theorem 4.4.5]) *Suppose that S is closed subset of a metric space X and Y is a normed linear space. Then every continuous mapping $f : S \rightarrow Y$ has a continuous extension $F : X \rightarrow Y$ such that $F(X) \subset \text{cov}f(S)$, where $\text{cov}f(S)$ denotes the convex hull of $f(S)$.*

Theorem 1.2.12. *Let S be a bounded, closed, nonempty convex subset of $\mathbb{R}^n$ and $\phi : S \rightarrow S$ is continuous. Then ϕ has a fixed point.*

Proof. Since S is bounded, it is contained in a closed ball $\bar{B}(0, r)$. Let $X = \bar{B}(0, r)$ and $I_S : S \rightarrow S$ be the identity mapping. By Lemma 1.2.11, there is a continuous mapping $F : X \rightarrow \mathbb{R}^n$ such that $F(X) = F(\bar{B}(0, r)) \subset \text{cov}I_S(S) = S$ and $F|_S = I_S$. Consider mapping $\tilde{\phi} = \phi \circ F$ which maps $\bar{B}$ into $\bar{B}$. By Theorem 1.2.10, there exists $x_0 \in \bar{B}$ such that $\tilde{\phi}(x_0) = x_0$. This means that $\phi(F(x_0)) = x_0$. Since $F(x_0) \in S$, $\phi(F(x_0)) \in S$. Hence $x_0 \in S$. Consequently, $F(x_0) = x_0$ and so $\tilde{\phi}(x_0) = \phi(x_0) = x_0$. The theorem is proved. $\square$

1.2.2 Browder degree for operators of class $(S)_+$

Based on the fact that any compact operator in Banach spaces can be approximated by mappings which have ranges in finite dimensional spaces, in 1934 Leray-Schauder introduced a degree theory for compact operators. For this we refer the readers to Chapter 4 in [151]. Since then there have been various extensions and generalization in different directions for degree theory (see, for instance [35], [50], [53], [88] - [151] and references given therein). One of the most important generalizations is due to Browder [35], where a degree theory was introduced for maps of class $(S)_+$ from a bounded open subset of a reflexive Banach space into its dual. In particular, Browder built the degree theory for more general mappings of monotone type $f + T$, where f is of class $(S)_+$ and T is a maximal monotone operator. This generalization paved the way for applications of the degree theory to a large number of nonlinear partial differential equations (see, e.g., [35] [107]). For this important event, we will outline below some results in [35].

In this subsection we will define a degree theory for demicontinuous operators of class $(S)_+$. For this we assume that X is a reflexive Banach space. By the result which is due to Lindenstrauss, Asplund and Trojanski, it is true that X can be renormed so that X and X^* are both locally uniformly convex (see [56, Theorem 2.11]). In this case the

duality mapping $J = J_2 : X \rightarrow X^*$ defined by

$$\langle J(x), x \rangle = \|x\|^2 = \|J(x)\|^2$$

is of class $(S)_+$, strictly monotone and a homeomorphism (see [35, Proposition 8] and [238, Corollary 32.24]).

Definition 1.2.13. *Let G be a subset of a Banach space X and $f : G \rightarrow X^*$ be a mapping. Then*

- (a) *f is said to be of class $(S)_+$ if for any sequence $\{x_n\}$ in G which converges weakly to x (written $x_n \rightharpoonup x$ briefly) and*

$$\limsup_{n \rightarrow \infty} \langle f(x_n), x_n - x \rangle \leq 0$$

we have $\|x_n - x\| \rightarrow 0$.

- (b) *f is said to be demicontinuous if for any sequence $\{x_n\}$ in G which converges to x we have $f(x_n)$ converges weakly to $f(x)$.*

- (c) *f is said to be pseudo-monotone, if for any sequence $\{x_j\}$ in G for which $x_j \rightharpoonup x$ for some $x \in X$ and $\overline{\lim} \langle f(x_j), x_j - x \rangle \leq 0$, then*

$$\langle f(x), x - y \rangle \leq \underline{\lim} \langle f(x_j), x_j - y \rangle \quad \forall y \in G.$$

Let $\mathcal{S}$ be a family of finite-dimensional subspaces of X , which is ordered by inclusion. For each $L \in \mathcal{S}$, let φ_L be the injection map of L into X and $\varphi_L^* : X^* \rightarrow L^*$ which is defined by setting

$$\langle \varphi_L^* x^*, x \rangle = \langle x^*, \varphi_L x \rangle = \langle x^*, x \rangle \quad \forall x \in L.$$

Let Ω be an open bounded set in X . Then there exists $L_0 \in \mathcal{S}$ such that $\Omega_L := \Omega \cap L \neq \emptyset$ for $L \supseteq L_0$. Define $f_L(x) = \varphi_L^*(f(x))$. Then from the proof of [35, Theorem 4] we have

Theorem 1.2.14. *Suppose that f is demicontinuous on $\overline{\Omega}$ and of class $(S)_+$.*

- (i) *For each $L \in \mathcal{S}$, f_L is continuous on L .*
(ii) *There exists $L_0 \in \mathcal{S}$ such that if $y_0^* \notin f(\partial\Omega)$, then $\varphi_L(y_0^*) \notin f_L(\partial\Omega_L)$ and $d(f_K, \Omega_K, \varphi_K^*(y_0^*)) = d(f_L, \Omega_L, \varphi_L^*(y_0^*))$ for all $K, L \supseteq L_0$.*

Based on this theorem, Browder gave the following definition in 1983.

Definition 1.2.15. *Let Ω be a open bounded set in X and $f : \overline{\Omega} \rightarrow X^*$ be demicontinuous and of class $(S)_+$, and $y_0^* \notin f(\partial\Omega)$. Then the degree $d(f, \Omega, y_0^*) := d(f_L, \Omega_L, \varphi_L^*(y_0^*))$ for $L \in \mathcal{S}$ sufficiently large.*

The following result on the existence and uniqueness of the degree theory for mappings of class $(S)_+$.

Theorem 1.2.16. (*[35, Theorem 4]*) *Let X be a reflexive Banach space, $\mathcal{O}$ be a class of bounded open subsets Ω of X , $\mathcal{F}$ be a class of maps $f : \bar{\Omega} \rightarrow X^*$ such that f is of class $(S)_+$ and demicontinuous. Let $\mathcal{H}$ be the class of affine homotopies in $\mathcal{F}$ and J be the duality mapping from X to X^* corresponding to an equivalent norm on X in which both X and X^* are locally uniformly convex. Then there exists exactly one degree function d in the sense of Definition 1.2.1, which defines on $\mathcal{F}$, to be invariant under $\mathcal{H}$ and normalized by J .*

1.2.3 Browder degree for mappings of monotone type

Let $T : X \rightarrow 2^{X^*}$ be a multifunction. We will call the sets $\text{Gr}T = \{(x, y) : y \in T(x)\}$ and $\text{Dom}(T) = \{x \in X : T(x) \neq \emptyset\}$, the *graph* and the *effective domain* of T , respectively. In the rest of the paper we always assume that $\text{Dom}(T)$ is nonempty.

Recall that the multifunction $T : X \rightarrow 2^{X^*}$ is said to be *monotone* if for any $(x_1, y_1), (x_2, y_2)$ in $\text{Gr}(T)$ one has $\langle y_2 - y_1, x_2 - x_1 \rangle \geq 0$. T is said to be *maximal monotone* if T is monotone and it follows from

$$\langle v - y, u - x \rangle \geq 0, \quad \forall (x, y) \in \text{Gph}(T),$$

that $(u, v) \in \text{Gr}(T)$.

For each $\epsilon > 0$, we consider the generalized *Yosida transformation* T_ϵ corresponding to T , considered in [30] and defined by the formula

$$T_\epsilon = (T^{-1} + \epsilon J^{-1})^{-1}.$$

The following result give some properties of T_ϵ .

Proposition 1.2.17. *Suppose that $T : X \rightarrow 2^{X^*}$ is maximal monotone. Then for each $\epsilon > 0$, T_ϵ is single-valued, monotone and demicontinuous.*

Proof. Let $x \in X$ and $y_1^*, y_2^* \in T_\epsilon(x)$. Then

$$x \in T^{-1}(y_i^*) + \epsilon J^{-1}(y_i^*)$$

or

$$y_i^* \in T(x - \epsilon J^{-1}(y_i^*)), i = 1, 2.$$

Since T is monotone, we have

$$\langle y_1^* - y_2^*, \epsilon(J^{-1}(y_2^*) - J^{-1}(y_1^*)) \rangle \geq 0$$

or equivalently,

$$\langle y_1^* - y_2^*, J^{-1}(y_1^*) - J^{-1}(y_2^*) \rangle \leq 0.$$

By Proposition 32.22 in [238], $J^{-1} : X^* \rightarrow X$ is a duality mapping of X^* , which is strictly monotone. Hence we must have $y_1^* = y_2^*$ and so T_ϵ is single-valued. The monotonicity of T_ϵ follows from monotonicity of T and J^{-1} .

Let $x_n \rightarrow x_0$ and suppose that $y_n^* = T_\epsilon(x_n)$ and $y_0^* = T_\epsilon(x_0)$. We want to show that y_n^* converges weakly to y_0^* . By the same argument as the above, we have

$$\begin{aligned} y_n^* &\in T(x_n - \epsilon J^{-1}(y_n^*)), \\ y_0^* &\in T(x_0 - \epsilon J^{-1}(y_0^*)). \end{aligned}$$

By the monotonicity of T , we have

$$\langle y_n^* - y_0^*, \epsilon(J^{-1}(y_n^*) - J^{-1}(y_0^*)) + x_0 - x_n \rangle \leq 0.$$

Let us claim that $\{\|y_n\|\}$ is bounded. Otherwise, we may assume that $\|y_n\| \rightarrow \infty$ as $n \rightarrow \infty$. Then $\|y_n\| \geq \|y_0^*\|$ for sufficiently large n . It follows that

$$\langle y_n^* - y_0^*, \epsilon(J^{-1}(y_n^*) - J^{-1}(y_0^*)) \rangle \leq \langle y_n^* - y_0^*, x_n - x_0 \rangle. \quad (1.2.2)$$

Using properties of J^{-1} (see Proposition 32.22 and Corollary 32.24 in [238]), we have

$$\begin{aligned} &\epsilon\|y_n\|^2 - \langle y_0^*, \epsilon(J^{-1}(y_n^*)) \rangle - \langle y_n^*, \epsilon J^{-1}(y_0^*) \rangle + \epsilon\|y_0\|^2 \\ &\leq (\|y_n^*\| + \|y_0^*\|)\|x_n - x_0\| \end{aligned}$$

This implies that

$$\epsilon\|y_n\|^2 - \epsilon\|y_0^*\|\|y_n^*\| - \epsilon\|y_n^*\|\|y_0^*\| + \epsilon\|y_0\|^2 \leq 2\|y_n^*\|\|x_n - x_0\|$$

Dividing both sides by $\|y_n\|$, we get

$$\epsilon\|y_n\| - 2\epsilon\|y_0^*\| + \frac{\epsilon\|y_0\|^2}{\|y_n\|} \leq 2\|x_n - x_0\|.$$

By letting $n \rightarrow \infty$, we obtain $\|x_n - x_0\| \rightarrow \infty$ which is absurd. Therefore, the claim is justified.

By the boundedness of $\{y_n\}$, we have from (1.2.2) that

$$\lim_{n \rightarrow \infty} \langle y_n^* - y_0^*, J^{-1}(y_n^*) - J^{-1}(y_0^*) \rangle \leq \lim_{n \rightarrow \infty} \langle y_0^* - y_n^*, x_n - x_0 \rangle = 0.$$

By property of J^{-1} (see (b) of Proposition 32.22 in [238]) we have $y_n \rightharpoonup y_0^*$. The proof of the proposition is complete. $\square$

From Proposition 1.2.17, we see that if $f : \bar{\Omega} \rightarrow X^*$ is demicontinuous and of class $(S)_+$, then $f + T_\epsilon$ is also demicontinuous and of class $(S)_+$. Therefore, we can define $d(f + T_\epsilon, \Omega, y_0^*)$. The following theorem is the basis for building degree theory for mappings of type $f + T$.

Theorem 1.2.18 ([35, Theorem 8]). *Assume $\Omega \in \mathcal{O}$ and conditions:*

- (i) $f : \bar{\Omega} \rightarrow X^*$ is of class $(S)_+$, demicontinuous and bounded.
- (ii) $T : X \rightarrow 2^{X^*}$ is maximal monotone.
- (iii) $y_0 \notin (f + T)(\partial\Omega)$.

Suppose further that

- (iv) $0 \in T(0)$.

Then there exists $\bar{\epsilon} > 0$ such that the following assertions are valid:

- (a) $y_0 \notin (f + T_{\epsilon})(\partial\Omega)$ for all ϵ in $(0, \bar{\epsilon}]$;
- (b) $d(f + T_\epsilon, \Omega, y_0) = d(f + T_{\epsilon'}, \Omega, y_0)$ for all ϵ, ϵ' in $(0, \bar{\epsilon}]$.

Definition 1.2.19. *Let $\Omega \in \mathcal{O}$, $f : \bar{\Omega} \rightarrow X^*$ be of class $(S)_+$, demicontinuous and bounded. Let $T : X \rightarrow 2^{X^*}$ be a maximal monotone operator and y_0 be a point in X^* such that $y_0 \notin (f + T)(\partial\Omega)$. Suppose, in addition, $0 \in T(0)$. The degree function $d(f + T, \Omega, y_0)$ is the common value of $d(f + T_\epsilon, \Omega, y_0)$ for ϵ sufficiently small.*

Definition 1.2.20. *Let $\Omega \in \mathcal{O}$ and $\{f_t, 0 \leq t \leq 1\}$ be a family of maps from $\bar{\Omega}$ into X^* . $\{f_t\}$ is said to be a homotopy of class $(S)_+$ if for any sequence $\{x_n\}$ in $\bar{\Omega}$ converging weakly to x and for any sequence $\{t_n\}$ in $[0, 1]$ converging to t for which*

$$\limsup_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x \rangle \leq 0,$$

then $x_n \rightarrow x$ and $f_{t_n}(x_n) \rightarrow f_t(x)$.

Definition 1.2.21. *Let $\{T_t, 0 \leq t \leq 1\}$ be a family of maximal monotone maps from X into 2^{X^*} . Suppose, in addition, that $0 \in T_t(0)$ for all t in $[0, 1]$. Then $\{T_t\}$ is called a pseudo-monotone homotopy if it satisfies the following mutually equivalent properties ([35, Proposition 15]):*

- (a) *Suppose that $t_n \rightarrow t$ in $[0, 1]$ and $(x_n, y_n) \in \text{Gph}(T_{t_n})$ such that $x_n \rightharpoonup x$ in X , $y_n \rightharpoonup y$ in X^* and $\limsup \langle y_n, x_n \rangle \leq \langle y, x \rangle$. Then $(x, y) \in \text{Gph}(T_t)$ and $\lim \langle y_n, x_n \rangle = \langle y, x \rangle$.*

(b) The mapping φ of $X^* \times [0, 1]$ into X defined by

$$\varphi(w, t) = (T_t + J)^{-1}(w),$$

is continuous, with X and X^* given the norm topologies.

(c) For each w in X^* , the mapping φ_w of $[0, 1]$ into X defined by

$$\varphi_w(t) = (T_t + J)^{-1}(w),$$

is continuous from $[0, 1]$ into X in the norm topology.

(d) For any (x, y) in $\text{Gr}(T_t)$ and a sequence $t_n \rightarrow t$ in $[0, 1]$, there exists a sequence (x_n, y_n) in $\text{Gr}(T_{t_n})$ such that $x_n \rightarrow x$ and $y_n \rightarrow y$.

The following results establish the homotopy invariance of the degree theory for maps of type $f + T$.

Theorem 1.2.22 ([35, Theorem 9]). Assume $\Omega \in \mathcal{O}$ and conditions:

- (i) $\{f_t, 0 \leq t \leq 1\}$ is a homotopy of class $(S)_+$ of maps of $\bar{\Omega}$ into a bounded subset of X^* .
- (ii) $\{T_t : 0 \leq t \leq 1\}$ is a pseudo-monotone homotopy from X into 2^{X^*} .
- (iii) $\{y_t : 0 \leq t \leq 1\}$ is a continuous path in X^* such that

$$y_t \notin (f_t + T_t)(\partial\Omega) \quad \forall t \in [0, 1].$$

Suppose, in addition,

- (iv) $0 \in T_t(0)$ for all t in $[0, 1]$.

Then there exists $\bar{\epsilon} > 0$ such that the following assertions are valid:

- (a) $y_t \notin (f_t + T_{t,\epsilon})(\partial\Omega)$ for all $\epsilon \in (0, \bar{\epsilon}]$ and t in $[0, 1]$.
- (b) $d(f_t + T_{t,\epsilon}, \Omega, y_t)$ is independent of t and ϵ .

Theorem 1.2.23 ([35, Theorem 12]). There exists exactly one degree function in the sense of Definition 1.2.1 on the class $\mathcal{F}$ of maps $f + T$, where f is of class $(S)_+$, demicontinuous and bounded, and T is maximal monotone satisfying $0 \in T(0)$, which is normalized by J and invariant under class $\mathcal{H}$ of all affine homotopies of the form $(1 - t)(T + f) + tg$, where f, g are of class $(S)_+$, demicontinuous and T is maximal monotone.

Theorems 1.2.18, 1.2.22, and 1.2.23 are nice results on the degree theory for mappings of monotone type $f + T$. However, the class of mappings of monotone type $f + T$ given by Browder is rather narrow. Namely, in the above theorems, the conditions

$$0 \in T(0) \quad \text{and} \quad 0 \in T_t(0).$$

were required. Browder used these conditions in an essential way to prove his theorems. Meanwhile, several applications, where the graphs of operators T_t are supposed not to contain $(0, 0)$. It is noted that, if we take any (x_0, x_0^*) in $\text{Gph}(T)$ and put $T'(x) = T(x + x_0) - x_0^*$ and $f'(x) = f(x + x_0)$ then T' is maximal monotone and satisfies $0 \in T'(0)$. It is plausible to define the degree

$$d(f + T, \Omega, y_0)$$

via $d(f' + T', \Omega - x_0, y_0 - x_0^*)$. However, in this case, the degree depends on (x_0, x_0^*) and we will encounter easily a trouble with the homotopy invariance since it might happen that $x_0^* \notin \bigcap_t T_t(x_0)$. The following is a counter example.

Example 1.2.24. *Let K be a nonempty closed convex set in X such that $0 \notin K$ and $2K \subset K$. Assume that $N_K(x)$ is the normal cone to K at x . Put $T_t(x) = (1 - t)(-J(x) + N_K(x))$. Since $0 \notin K$, $0 \notin T_t(0)$. Indeed, we have $\bigcap_t T_t(x) = \emptyset$ for all x in K . Hence $\{T_t, 0 \leq t \leq 1\}$ is not a pseudo-monotone homotopy in the sense of Definition 1.2.21, even after translations.*

In this section, we will extend results of Browder by enlarging the class of pseudo-monotone homotopies $\{T_t\}$ without the assumptions

$$0 \in T(0) \quad \text{and} \quad 0 \in T_t(0).$$

To do this, we will give a new scheme in proving our theorems. The obtained results are enhanced versions of Theorems 1.2.18, 1.2.22, and 1.2.23 under relaxed conditions for a broader class of $\{T_t\}$. We think this is significant, since our results can be applied to a degree theory of variational inequalities in reflexive Banach spaces and solving nonlinear PDEs, for some of which the original Browder's degree theory does not seem to be very handy.

1.2.4 Extension of the Browder degree theory for monotone mappings

Let us keep all the notations of the preceding subsection. The following theorem extends Theorem 1.2.18.

Theorem 1.2.25. *Suppose the following conditions:*

- (i) $f : \bar{\Omega} \rightarrow X^*$ is of class $(S)_+$, demicontinuous and bounded;
- (ii) $T : X \rightarrow 2^{X^*}$ is maximal monotone;
- (iii) $y_0 \notin (f + T)(\partial\Omega)$.

Then there exists $\bar{\epsilon} > 0$ such that the following assertions are valid:

- (a) $y_0 \notin (f + T_\epsilon)(\partial\Omega)$ for all ϵ in $(0, \bar{\epsilon}]$;
- (b) $d(f + T_\epsilon, \Omega, y_0) = d(f + T_{\epsilon'}, \Omega, y_0)$ for all ϵ, ϵ' in $(0, \bar{\epsilon}]$.

Definition 1.2.26. *Let $\Omega \in \mathcal{O}$ and $f : \bar{\Omega} \rightarrow X^*$ be of class $(S)_+$, demicontinuous and bounded. Let $T : X \rightarrow 2^{X^*}$ be a maximal monotone operator and y_0 be a point such that $y_0 \notin (f + T)(\partial\Omega)$. The degree $d_1(f + T, \Omega, y_0)$ is assigned to be the common value of $d(f + T_\epsilon, \Omega, y_0)$ for ϵ sufficiently small.*

The next theorems will be proved for the class of maximal monotone operators and pseudo-monotone homotopies which enlarges the class given by Browder. For such classes, we introduce the following definition.

Definition 1.2.27. *Let $\{T_t, 0 \leq t \leq 1\}$ be a family of maximal monotone maps from X into 2^{X^*} such that their effective domains are nonempty. Then $\{T_t\}$ is called a pseudo-monotone homotopy if for any (x, y) in $\text{Gr}(T_t)$ and a sequence $t_n \rightarrow t$ in $[0, 1]$, there exists a sequence (x_n, y_n) in $\text{Gr}(T_{t_n})$ such that $x_n \rightarrow x$ and $y_n \rightarrow y$.*

It is obvious that if $\{T_t, 0 \leq t \leq 1\}$ is a pseudo-monotone homotopy in the sense of Definition 1.2.21 then it is also a pseudo-monotone homotopy in the sense of Definition 1.2.27. Example 1.2.24 indicates that there exists a pseudo-monotone homotopy in the sense of Definition 1.2.27 but it is not a pseudo-monotone homotopy in the sense of Definition 1.2.21.

Based on the class of pseudo-monotone homotopies in the sense of Definition 1.2.27 we obtain the following results extending Theorems 1.2.22 and 1.2.23.

Theorem 1.2.28. *Suppose conditions:*

- (i) $\{f_t, 0 \leq t \leq 1\}$ is a homotopy of class $(S)_+$ of maps of $\bar{\Omega}$ into a bounded subset of X^* ;
- (ii) $\{T_t : 0 \leq t \leq 1\}$ is a pseudo-monotone homotopy from X into 2^{X^*} ;
- (iii) $\{y_t : 0 \leq t \leq 1\}$ is a continuous path in X^* such that $y_t \notin (f_t + T_t)(\partial\Omega)$ for all t in $[0, 1]$.

Then there exists $\bar{\epsilon} > 0$ such that the following assertions are valid:

- (a) $y_t \notin (f_t + T_{t,\epsilon})(\partial\Omega)$ for all ϵ in $(0, \bar{\epsilon}]$ and t in $[0, 1]$.
- (b) $d(f_t + T_{t,\epsilon}, \Omega, y_t)$ is independent of t and ϵ .

Theorem 1.2.29. *Suppose that $\Omega \in \mathcal{O}$ and $f : \Omega \rightarrow X^*$ is of class $(S)_+$, demicontinuous and bounded. Let $T : X \rightarrow 2^{X^*}$ be a maximal monotone operator and y_0 be a point of X^* such that $y_0 \notin (f + T)(\partial\Omega)$. Then the following assertions hold:*

- (a) (Existence) *If $d_1(f + T, \Omega, y_0) \neq 0$ then there exists x in Ω such that $y_0 \in f(x) + T(x)$.*
- (b) (Additivity) *If Ω_1, Ω_2 are a pair disjoint subsets of $\mathcal{O}$ contained in Ω such that $y_0 \notin (f + T)(\bar{\Omega} \setminus (\Omega_1 \cup \Omega_2))$ then*

$$d_1(f + T, \Omega, y_0) = d_1(f + T, \Omega_1, y_0) + d_1(f + T, \Omega_2, y_0).$$

- (c) (Homotopy invariance) *If $\{f_t, 0 \leq t \leq 1\}$ is a homotopy of class $(S)_+$ of maps of $\bar{\Omega}$ into a bounded set of X^* , $\{T_t, 0 \leq t \leq 1\}$ is a pseudo-monotone homotopy from X into 2^{X^*} and $\{y_t : 0 \leq t \leq 1\}$ is a continuous path in X^* such that $y_t \notin (f_t + T_t)(\partial\Omega)$ for all t in $[0, 1]$, then $d_1(f_t + T_t, \Omega, y_t)$ is independent of t .*

Theorem 1.2.30. *There exists exactly one degree function in the sense of Definition 1.2.26 on the class $\mathcal{F}$ of maps $f + T$, where f is of class $(S)_+$, demicontinuous and bounded, and T is maximal monotone, which is normalized by J and invariant under class $\mathcal{H}$ of all affine homotopies of the form $(1 - t)(f + T) + tg$, where f and g are of class $(S)_+$, demicontinuous and bounded, T is maximal monotone.*

1.2.5 Proofs of extension results

We notice that Theorem 1.2.28 includes Theorem 1.2.25 as a special case when $f_t = f, y_t = y_0$ and $T_t = T$ for all t in $[0, 1]$. Hence it suffices to prove Theorem 1.2.28. We will complete the proof of Theorem 1.2.28 by proving some lemmas.

Lemma 1.2.31. *Let f be of class $(S)_+$ and demicontinuous. Assume that g is demicontinuous and monotone. Then $f + g$ is of class $(S)_+$ and demicontinuous.*

Proof. It is easily seen that $f + g$ is demicontinuous. Let $\{x_n\}$ be a sequence such that $x_n \rightharpoonup x_0$ and $\limsup_{n \rightarrow \infty} \langle f(x_n) + g(x_n), x_n - x_0 \rangle \leq 0$. By the monotonicity of g we get

$$\langle f(x_n) + g(x_n), x_n - x_0 \rangle \geq \langle f(x_n), x_n - x_0 \rangle + \langle g(x_0), x_n - x_0 \rangle.$$

This implies that $\limsup_{n \rightarrow \infty} \langle f(x_n), x_n - x_0 \rangle \leq 0$. Since f belongs to class $(S)_+$, we have $x_n \rightarrow x_0$. Hence $f + g$ is of class $(S)_+$. $\square$

Lemma 1.2.32. *Suppose that f satisfies conditions in Theorem 1.2.28, $x_n \in \Omega$ such that $x_n \rightharpoonup x_0$ and $t_n \in [0, 1]$. Then $\liminf_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x_0 \rangle \geq 0$.*

Proof. Suppose on the contrary that

$$\alpha := \liminf_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x_0 \rangle < 0. \quad (1.2.3)$$

Since $\langle f_{t_n}(x_n), x_n - x_0 \rangle \geq -\|f(x_n)\| \|x_n - x_0\|$, by the boundedness of sequences $\{f_{t_n}(x_n)\}$ and $\{x_n - x_0\}$, we have $\alpha > -\infty$. It follows that there exist subsequence $\{x_{n_k}\}$ and $\{t_{n_k}\}$ such that $\lim_{k \rightarrow \infty} \langle f_{t_{n_k}}(x_{n_k}), x_{n_k} - x_0 \rangle = \alpha < 0$. Hence there exists a number $k_0 > 0$ such that $\langle f_{t_{n_k}}(x_{n_k}), x_{n_k} - x_0 \rangle < 0$ for all $k \geq k_0$. This implies that $\limsup_{k \rightarrow \infty} \langle f_{t_{n_k}}(x_{n_k}), x_{n_k} - x_0 \rangle \leq 0$. Since f is of class $(S)_+$, we get $x_{n_k} \rightarrow x_0$. By using the inequality

$$\langle f_{t_{n_k}}(x_{n_k}), x_{n_k} - x_0 \rangle \geq -\|f_{t_{n_k}}(x_{n_k})\| \|x_{n_k} - x_0\|$$

we obtain $\alpha = \lim_{k \rightarrow \infty} \langle f_{t_{n_k}}(x_{n_k}), x_{n_k} - x_0 \rangle \geq 0$. This contradicts (1). The proof is complete. $\square$

Proof of Theorem 1.2.28

(a) Suppose that the assertion is false. Then we could find a sequence $\epsilon_n \rightarrow 0^+$, a sequence x_n in $\partial\Omega$ and a sequence t_n in $[0, 1]$ such that

$$y_{t_n} = f_{t_n}(x_n) + z_n, \quad (1.2.4)$$

where $z_n = T_{t_n, \epsilon_n}(x_n)$. Since $\{f_{t_n}(x_n)\}$ and $\{y_{t_n}\}$ are bounded, the sequence $\{z_n\}$ is also bounded. Hence we can assume that $x_n \rightharpoonup x_0$, $t_n \rightarrow t$ in $[0, 1]$, $z_n \rightharpoonup z_0$ and $y_{t_n} \rightarrow y_t$. Since

$$\langle z_n, x_n - x_0 \rangle = \langle y_{t_n} - f_{t_n}(x_n), x_n - x_0 \rangle,$$

we have

$$\langle z_n, x_n \rangle = \langle z_n, x_0 \rangle + \langle y_{t_n}, x_n - x_0 \rangle - \langle f_{t_n}(x_n), x_n - x_0 \rangle. \quad (1.2.5)$$

This implies that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ & \leq \limsup_{n \rightarrow \infty} \langle z_n, x_0 \rangle + \limsup_{n \rightarrow \infty} \langle y_{t_n}, x_n - x_0 \rangle - \liminf_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x_0 \rangle. \end{aligned} \quad (1.2.6)$$

By the Lemma 1.2.32, we have $\liminf_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x_0 \rangle \geq 0$. Hence it follows from (1.2.6) that

$$\limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \leq \langle z_0, x_0 \rangle. \quad (1.2.7)$$

Note that the equality $z_n = T_{t_n, \epsilon_n}(x_n)$ means that $z_n \in T_{t_n}(x_n - \epsilon_n J^{-1}(z_n))$. We now take any (u, v) in $\text{Gr}(T_t)$. By the definition of $\{T_t\}$, there exists sequences $u_n \rightarrow u$ and $v_n \rightarrow v$ such that $(u_n, v_n) \in \text{Gr}(T_{t_n})$. Using the monotonicity of T_{t_n} we get

$$\langle z_n - v_n, x_n - \epsilon_n J^{-1}(z_n) - u_n \rangle \geq 0.$$

Hence

$$\langle z_n, x_n \rangle \geq \langle z_n, u_n \rangle + \langle v_n, x_n - u_n \rangle + \epsilon_n \|z_n\|^2 - \epsilon_n \|v_n\| \|z_n\|. \quad (1.2.8)$$

From (1.2.7) and (1.2.8) we have

$$\begin{aligned} \langle z_0, x_0 \rangle & \geq \limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \geq \liminf_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ & \geq \liminf_{n \rightarrow \infty} \langle z_n, u_n \rangle + \liminf_{n \rightarrow \infty} \langle v_n, x_n - u_n \rangle + \liminf_{n \rightarrow \infty} (\epsilon_n \|z_n\|^2 - \epsilon_n \|v_n\| \|z_n\|) \\ & \geq \langle z_0, u \rangle + \langle v, x_0 - u \rangle. \end{aligned} \quad (1.2.9)$$

This implies that $\langle z_0 - v, x_0 - u \rangle \geq 0$ for all (u, v) in $\text{Gr}(T_t)$. By the maximal monotonicity of T_t we obtain $z_0 \in T_t(x_0)$. Substituting $(u, v) = (x_0, z_0)$ into (1.2.9) we get $\lim_{n \rightarrow \infty} \langle z_n, x_n \rangle = \langle z_0, x_0 \rangle$. Hence it follows from (1.2.5) that $\lim_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x_0 \rangle = 0$. Since $\{f_s, 0 \leq s \leq 1\}$ is

homotopy of class $(S)_+$ and $\partial\Omega$ is closed, $x_n \rightarrow x_0 \in \partial\Omega$ and $f_{t_n}(x_n) \rightharpoonup f_t(x_0)$. By letting $n \rightarrow \infty$, from (1.2.4) we obtain $y_t = f_t(x_0) + z_0$, where $z_0 \in T_t(x_0)$ and $x_0 \in \partial\Omega$. This means that $y_t \in (f_t + T_t)(\partial\Omega)$ for some t in $[0, 1]$. We obtain a contradiction.

(b) We first show that for each ϵ in $(0, \bar{\epsilon}]$ and t in $[0, 1]$, the degree $d_1(f_t + T_{t,\epsilon}, \Omega, y_t)$ is well defined. Recall that

$$T_{t,\epsilon} = (T_t^{-1} + \epsilon J^{-1})^{-1}.$$

Since T_t is maximal monotone, T_t^{-1} is also maximal monotone (see [238, Proposition 32.5]). By [56, Proposition 5.5] (see also, [238, Corollary 32.24]), J^{-1} has the same properties of J . According to Proposition 3.10 in [5], $(T_t^{-1} + \epsilon J^{-1})^{-1}$ is a single-valued and demicontinuous maximal monotone operator. Hence, from Lemma 1.2.31, we obtain that $f_t + T_{t,\epsilon}$ is of class $(S)_+$ and demicontinuous. Consequently, the degree $d(f_t + T_{t,\epsilon}, \Omega, y_t)$ is well defined.

We now show that there exists $\bar{\epsilon} > 0$ such that $d(f_t + T_{t,\epsilon}, \Omega, y_t)$ is independent of ϵ in $(0, \bar{\epsilon}]$ and t in $[0, 1]$. Suppose on the contrary that the assertion is false. Then there exist $\epsilon_n \rightarrow 0^+$, $\epsilon'_n \rightarrow 0^+$ and t_n in $[0, 1]$ such that

$$d(f_{t_n} + T_{t_n,\epsilon_n}, \Omega, y_{t_n}) \neq d(f_{t_n} + T_{t_n,\epsilon'_n}, \Omega, y_{t_n}). \quad (1.2.10)$$

For each fixed n , we consider the homotopy H_n defined by

$$H_n(s, x) = sT_{t_n,\epsilon_n}(x) + (1 - s)T_{t_n,\epsilon'_n}(x) + f_{t_n}(x). \quad (1.2.11)$$

If $y_{t_n} \notin H_n(s, \partial\Omega)$ for all s in $[0, 1]$ then

$$d(H_n(1, \cdot), \Omega, y_{t_n}) = d(H_n(0, \cdot), \Omega, y_{t_n}).$$

Hence

$$d(f_{t_n} + T_{t_n,\epsilon_n}, \Omega, y_{t_n}) = d(f_{t_n} + T_{t_n,\epsilon'_n}, \Omega, y_{t_n}),$$

which contradicts (1.2.10). Thus, there exists x_n in $\partial\Omega$ and s_n in $[0, 1]$ such that

$$y_{t_n} = s_n T_{t_n,\epsilon_n}(x_n) + (1 - s_n) T_{t_n,\epsilon'_n}(x_n) + f_{t_n}(x_n). \quad (1.2.12)$$

Put $z_n = s_n T_{t_n,\epsilon_n}(x_n) + (1 - s_n) T_{t_n,\epsilon'_n}(x_n)$. Then $\{z_n\}$ is a bounded sequence because the sequences $\{f_{t_n}(x_n)\}$ and $\{y_{t_n}\}$ are bounded. Without loss of generality we can assume that $t_n \rightarrow t$, $s_n \rightarrow s$, $x_n \rightharpoonup x_0$, $z_n \rightharpoonup z_0$ and $y_{t_n} \rightarrow y_t$. Putting $a_n = T_{t_n,\epsilon_n}(x_n)$, $b_n = T_{t_n,\epsilon'_n}(x_n)$, we have

$$a_n \in T_{t_n}(x_n - \epsilon_n J^{-1}(a_n)), \quad b_n \in T_{t_n}(x_n - \epsilon'_n J^{-1}(b_n)). \quad (1.2.13)$$

As $y_{t_n} = f_{t_n}(x_n) + z_n$, we have

$$\langle z_n, x_n \rangle = \langle z_n, x_0 \rangle + \langle y_{t_n}, x_n - x_0 \rangle - \langle f_{t_n}(x_n), x_n - x_0 \rangle. \quad (1.2.14)$$

By using the same argument as in the proof of part (a) we obtain

$$\limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \leq \langle z_0, x_0 \rangle. \quad (1.2.15)$$

We now consider the following cases:

Case 1. The sequence $\{s_n a_n\}$ is bounded.

Since the sequence $\{s_n a_n\}$ is bounded, $\{(1 - s_n)b_n\}$ is also bounded. Taking any (u, v) in $\text{Gph}(T_t)$, by the definition of the family $\{T_s, 0 \leq s \leq 1\}$ we can find sequences $u_n \rightarrow u$ and $v_n \rightarrow v$ such that $(u_n, v_n) \in \text{Gph}(T_{t_n})$. Using the monotonicity of T_{t_n} we get

$$\langle a_n - v_n, x_n - \epsilon_n J^{-1}(a_n) - u_n \rangle \geq 0.$$

Hence

$$\langle a_n, x_n \rangle \geq \langle a_n, u_n \rangle + \langle v_n, x_n - u_n \rangle + \epsilon_n \|a_n\|^2 - \epsilon_n \|v_n\| \|a_n\|. \quad (1.2.16)$$

Similarly, we also have

$$\langle b_n, x_n \rangle \geq \langle b_n, u_n \rangle + \langle v_n, x_n - u_n \rangle + \epsilon'_n \|b_n\|^2 - \epsilon'_n \|v_n\| \|b_n\|. \quad (1.2.17)$$

Multiplying (1.2.16) with s_n , (1.2.17) with $1 - s_n$ and summing up, we get

$$\begin{aligned} \langle z_n, x_n \rangle &\geq \langle z_n, u_n \rangle + \langle v_n, x_n - u_n \rangle + s_n \epsilon_n (\|a_n\|^2 - \|v_n\| \|a_n\|) \\ &\quad + (1 - s_n) \epsilon'_n (\|b_n\|^2 - \|v_n\| \|b_n\|). \end{aligned} \quad (1.2.18)$$

It follows that

$$\langle z_n, x_n \rangle \geq \langle z_n, u_n \rangle + \langle v_n, x_n - u_n \rangle - \epsilon_n s_n \|v_n\| \|a_n\| - \epsilon'_n (1 - s_n) \|v_n\| \|b_n\|. \quad (1.2.19)$$

From (1.2.15) and (1.2.19) we have

$$\begin{aligned} \langle z_0, x_0 \rangle &\geq \limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ &\geq \liminf_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ &\geq \liminf_{n \rightarrow \infty} \langle z_n, u_n \rangle + \liminf_{n \rightarrow \infty} \langle v_n, x_n - u_n \rangle + \\ &\quad + \liminf_{n \rightarrow \infty} (-\epsilon_n s_n \|v_n\| \|a_n\| - (1 - s_n) \epsilon'_n \|v_n\| \|b_n\|) \end{aligned} \quad (1.2.20)$$

$$\geq \langle z_0, u \rangle + \langle v, x_0 - u \rangle. \quad (1.2.21)$$

This implies that $\langle z_0 - v, x_0 - u \rangle \geq 0$ for all (u, v) in $\text{Gr}(T_t)$. Hence $z_0 \in T_t(x_0)$ by the maximal monotonicity. Substituting $(u, v) = (x_0, z_0)$ into (1.2.21), we get $\lim_{n \rightarrow \infty} \langle z_n, x_n \rangle = \langle z_0, x_0 \rangle$. Hence it follows from (1.2.14) that $\lim_{n \rightarrow \infty} \langle f_{t_n}(x_n), x_n - x_0 \rangle = 0$. Since $\{f_s, 0 \leq s \leq 1\}$ is homotopy of class $(S)_+$, we get $x_n \rightarrow x_0 \in \partial\Omega$ and $f_{t_n}(x_n) \rightarrow f_t(x)$. Hence it yields $y_t = f_t(x_0) + z_0$. This means that $y_t \in (f_t + T_t)(\partial\Omega)$ for some t in $[0, 1]$. We obtain a contradiction.

Case 2. The sequence $\{s_n a_n\}$ is unbounded.

Since $\{s_n a_n\}$ is unbounded, $\{(1 - s_n)b_n\}$ is also unbounded and we have

$$\lim_{n \rightarrow \infty} s_n \|a_n\| = \lim_{n \rightarrow \infty} (1 - s_n) \|b_n\| = +\infty.$$

The next lemma will finish the proof of the theorem.

Lemma 1.2.33.

$$\lim_{n \rightarrow \infty} \epsilon_n \|a_n\| = \lim_{n \rightarrow \infty} \epsilon'_n \|b_n\| = 0.$$

Proof. Fix (u_0, v_0) in $\text{Gr}(T_t)$. Then there exist a sequence $u_n \rightarrow u_0$ and a sequence $v_n \rightarrow v_0$ such that $v_n \in T_{t_n}(u_n)$. Using the monotonicity of T_{t_n} we get

$$\langle a_n - v_n, x_n - \epsilon_n J^{-1}(a_n) - u_n \rangle \geq 0.$$

Hence

$$\langle a_n, x_n - u_n \rangle \geq \langle v_n, x_n - u_n \rangle + \epsilon_n \|a_n\|^2 - \epsilon_n \|v_n\| \|a_n\|.$$

This implies that

$$\|a_n\| \|x_n - u_n\| \geq \langle v_n, x_n - u_n \rangle + \epsilon_n \|a_n\|^2 - \epsilon_n \|v_n\| \|a_n\|.$$

It is equivalent to

$$\epsilon_n \|a_n\|^2 - \|a_n\| (\|x_n - u_n\| + \epsilon_n \|v_n\|) + \langle v_n, x_n - u_n \rangle \leq 0. \quad (1.2.22)$$

Put

$$\Delta_n = (\epsilon_n \|v_n\| + \|x_n - u_n\|)^2 - 4\epsilon_n \langle v_n, x_n - u_n \rangle.$$

We get

$$\Delta_n \geq (\|x_n - u_n\| - \epsilon_n \|v_n\|)^2.$$

Hence it follows from (1.2.22) that

$$\|a_n\| \leq \frac{\epsilon_n \|v_n\| + \|x_n - u_n\| + \sqrt{\Delta_n}}{2\epsilon_n}.$$

Consequently,

$$\epsilon_n \|a_n\| \leq \frac{\epsilon_n \|v_n\| + \|x_n - u_n\| + \sqrt{\Delta_n}}{2}. \quad (1.2.23)$$

Similarly, we have

$$\epsilon'_n \|b_n\| \leq \frac{\epsilon'_n \|v_n\| + \|x_n - u_n\| + \sqrt{\Delta_n}}{2}. \quad (1.2.24)$$

Since $x_n \rightharpoonup x_0$, $u_n \rightarrow u_0$, $v_n \rightarrow v_0$, it follows from (1.2.23) and (1.2.24) that the sequences $\{\epsilon_n \|a_n\|\}$ and $\{\epsilon'_n \|b_n\|\}$ are bounded. We now can assume that

$$\lim_{n \rightarrow \infty} \epsilon_n \|a_n\| = \alpha, \quad \lim_{n \rightarrow \infty} \epsilon'_n \|b_n\| = \beta.$$

It remains to show that $\alpha = \beta = 0$. Conversely, suppose that the assertion is false. Then we must have either $\alpha > 0$ or $\beta > 0$. Fixing (u, v) in $\text{Gph}(T_t)$, we can find out a sequence $u_n \rightarrow u$ and a sequence $v_n \rightarrow v$ such that $v_n \in T_{t_n}(u_n)$. Using the monotonicity of T_{t_n} and the similar arguments, we obtain (1.2.18) again. From (1.2.15) and (1.2.18) we get

$$\begin{aligned} \langle z_0, x_0 \rangle &\geq \limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ &\geq \liminf_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ &\geq \liminf_{n \rightarrow \infty} \langle z_n, u_n \rangle + \liminf_{n \rightarrow \infty} \langle v_n, x_n - u_n \rangle + \\ &\quad + \liminf_{n \rightarrow \infty} (-s_n \epsilon_n \|a_n\| \|v_n\| - (1 - s_n) \epsilon'_n \|b_n\| \|v_n\|) + \\ &\quad + \liminf_{n \rightarrow \infty} (\epsilon_n \|a_n\| s_n \|a_n\| + \epsilon'_n \|b_n\| (1 - s_n) \|b_n\|). \end{aligned} \quad (1.2.25)$$

Note that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \langle z_n, u_n \rangle &= \langle z_0, u \rangle, \\ \liminf_{n \rightarrow \infty} \langle v_n, x_n - u_n \rangle &= \langle v, x_0 - u \rangle, \\ \liminf_{n \rightarrow \infty} (-s_n \epsilon_n \|a_n\| \|v_n\| - (1 - s_n) \epsilon'_n \|b_n\| \|v_n\|) &= -s\alpha \|v\| - (1 - s)\beta \|v\|, \end{aligned}$$

and

$$\liminf_{n \rightarrow \infty} (\epsilon_n \|a_n\| s_n \|a_n\| + \epsilon'_n \|b_n\| (1 - s_n) \|b_n\|) = +\infty.$$

Hence from (1.2.25) we get $\langle z_0, x_0 \rangle \geq +\infty$ which is an absurd. The proof of the lemma is complete. $\square$

We now take any (u, v) in $\text{Gr}(T_t)$. Let $u_n \rightarrow u$, $v_n \rightarrow v$ and $v_n \in T_{t_n}(u_n)$. Using (1.2.25) again we get

$$\begin{aligned}
\langle z_0, x_0 \rangle &\geq \limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \\
&\geq \liminf_{n \rightarrow \infty} \langle z_n, x_n \rangle \\
&\geq \liminf_{n \rightarrow \infty} \langle z_n, u_n \rangle + \liminf_{n \rightarrow \infty} \langle v_n, x_n - u_n \rangle + \\
&\quad + \liminf_{n \rightarrow \infty} (-s_n \epsilon_n \|a_n\| \|v_n\| - (1 - s_n) \epsilon'_n \|b_n\| \|v_n\|) \\
&\geq \langle z_0, x_0 \rangle + \langle v, x_0 - u \rangle.
\end{aligned} \tag{1.2.26}$$

This implies that $\langle z_0 - v, x_0 - u \rangle \geq 0$. Since (u, v) is arbitrary and T_t is maximal monotone, we obtain $z_0 \in T_t(x_0)$. Substituting $(u, v) = (x_0, z_0)$ into (1.2.26) we obtain again $\lim_{n \rightarrow \infty} \langle z_n, x_n \rangle = \langle z_0, x_0 \rangle$. In the same manner of the proof at the end of Case 1 we show that $y_t \in (f_t + T_t)(\partial\Omega)$ for some t in $[0, 1]$. This contradicts our assumptions. The proof of the theorem is now complete. $\square$

Remark 1.2.34. In the proofs of Theorems 8 and 9 in [35], Browder had used heavily the condition $0 \in T(0)$ and the condition $0 \in T_t(0)$ for all t in $[0, 1]$, to show that there exists a constant $M > 0$ such that $s_n \epsilon_n \|a_n\|^2 \leq M$ and $(1 - s_n) \epsilon'_n \|b_n\|^2 \leq M$. Then he concluded that $s_n \epsilon_n \|a_n\| \rightarrow 0$ and $(1 - s_n) \epsilon'_n \|b_n\| \rightarrow 0$. However, this scheme will collapse if those conditions are absent. In our proof above, we do not assume these conditions and give instead a new scheme of reasoning which based on an intrinsic characterization of maximal monotone mappings as well as the structure of the generalized Yosida transformation T_ϵ .

Proof of Theorem 1.2.29

(a) Choose a sequence $\{\epsilon_n\}$ such that $\epsilon_n \rightarrow 0^+$. By the definition, there exists $n_0 > 0$ such that

$$d_1(f + T, \Omega, y_0) = d(f + T_{\epsilon_n}, \Omega, y_0)$$

for all $n \geq n_0$. Since $d(f + T_{\epsilon_n}, \Omega, y_0) \neq 0$, there exists an x_n in Ω such that $f(x_n) + T_{\epsilon_n}(x_n) = y_0$. Since $\{f(x_n)\}$ is bounded, $\{T_{\epsilon_n}(x_n)\}$ is also bounded. Putting $z_n = T_{\epsilon_n}(x_n)$, we get $z_n \in T(x_n - \epsilon_n J^{-1}(z_n))$ and $f(x_n) + z_n = y_0$. Without loss of generality, we can assume that $x_n \rightharpoonup x_0$, $z_n \rightharpoonup z_0$. By the same argument as in the proof of Theorem 1.2.28 we can show that

$$\limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \leq \langle z_0, x_0 \rangle. \tag{1.2.27}$$

We now utilize a variant of the proof of Theorem 1.2.28. Taking any (u, v) in $\text{Gr}(T)$ and using the monotonicity of T , we have

$$\langle z_n - v, x_n - \epsilon_n J^{-1}(z_n) - u \rangle \geq 0.$$

This implies that

$$\langle z_n, x_n \rangle \geq \langle z_n, u \rangle + \langle v, x_n - u \rangle + \epsilon \|z_n\|^2 - \epsilon_n \|z_n\| \|v\|.$$

By (1.2.27) we get

$$\begin{aligned} \langle z_0, x_0 \rangle &\geq \limsup_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ &\geq \liminf_{n \rightarrow \infty} \langle z_n, x_n \rangle \\ &\geq \liminf_{n \rightarrow \infty} \langle z_n, u \rangle + \liminf_{n \rightarrow \infty} \langle v, x_n - u \rangle + \liminf_{n \rightarrow \infty} (-\epsilon_n \|v\| \|z_n\|) \\ &\geq \langle z_0, u \rangle + \langle v, x_0 - u \rangle. \end{aligned} \tag{1.2.28}$$

Consequently, $\langle z_0 - v, x_0 - u \rangle \geq 0$ for all (u, v) in $\text{Gr}(T)$. By the maximal monotonicity of T , it yields $z_0 \in T(x_0)$. Substituting $(u, v) = (z_0, x_0)$ into (1.2.28), we get $\lim_{n \rightarrow \infty} \langle z_n, x_n \rangle = \langle z_0, x_0 \rangle$. Since

$$\langle z_n, x_n \rangle = \langle z_n, x_0 \rangle + \langle y_0, x_n - x_0 \rangle - \langle f(x_n), x_n - x_0 \rangle,$$

we get $\limsup_{n \rightarrow \infty} \langle f(x_n), x_n - x_0 \rangle = 0$. Hence $x_n \rightarrow x_0$ because f is of class $(S)_+$. On the other hand f is demicontinuous, it follows $f(x_n) \rightarrow f(x_0)$. Using the equality $f(x_n) + z_n = y_0$ we obtain $y_0 = f(x_0) + z_0$, where $z_0 \in T(x_0)$ for some x_0 in Ω . Thus we have showed that $y_0 \in f(x_0) + T(x_0)$ for some x_0 in Ω .

(b) It easy to show that there exists $\bar{\epsilon} > 0$ such that

$$y_0 \notin (f + T_{\epsilon})(\bar{\Omega} \setminus (\Omega_1 \cup \Omega_2))$$

for all ϵ in $(0, \bar{\epsilon}]$. By Theorem 1.2.16, we have

$$d(f + T_{\epsilon}, \Omega, y_0) = d(f + T_{\epsilon}, \Omega_1, y_0) + d(f + T_{\epsilon}, \Omega_2, y_0)$$

for all ϵ in $(0, \bar{\epsilon}]$. Hence

$$d_1(f + T, \Omega, y_0) = d_1(f + T, \Omega_1, y_0) + d_1(f + T, \Omega_2, y_0).$$

(c) The assertion follows directly from Theorem 1.2.28. The proof of the theorem is complete. $\square$

Proof of Theorem 1.2.30

Existence. Let d_1 be the degree function in the Definition 1.2.26. Put $T_t := (1 - t)T$ and $f_t := (1 - t)f + tg$. Then $\{T_t\}$ is a pseudo-monotone homotopy in the sense of Definition 1.2.27 and $\{f_t\}$ is a homotopy of class $(S)_+$ in the sense of Definition 1.2.20. By Theorem 1.2.29, d_1 meets all requirements of the theorem.

Uniqueness. Suppose that there exists two degree functions d_1 and d_2 satisfying the conclusion of our theorem. By Theorem 1.2.16 on the uniqueness of the degree function for mappings of class $(S)_+$, the restrictions of d_1 and d_2 on this class are equal. Let f be of class $(S)_+$, demicontinuous and bounded, T be maximal monotone and y_0 be a point of X^* with $y_0 \notin (f + T)(\partial\Omega)$. Assume that

$$d_1(f + T, \Omega, y_0) \neq d_2(f + T, \Omega, y_0). \quad (1.2.29)$$

By Theorem 1.2.25, there exists $\bar{\epsilon} > 0$ such that

$$y_0 \notin (f + T_\epsilon)(\partial\Omega), \quad \forall \epsilon \in (0, \bar{\epsilon}].$$

Since $f + T_\epsilon$ is of class $(S)_+$ and demicontinuous, we have

$$d_1(f + T, \Omega, y_0) = d_1(f + T_\epsilon, \Omega, y_0) = d_2(f + T_\epsilon, \Omega, y_0).$$

By this and (1.2.29), it yields

$$d_2(f + T, \Omega, y_0) \neq d_2(f + T_\epsilon, \Omega, y_0). \quad (1.2.30)$$

Consider the affine homotopy between $(f + T)$ and $(f + T_\epsilon)$ for $\epsilon \in (0, \bar{\epsilon}]$. In view of (1.2.30), we can find $\epsilon_n \rightarrow 0^+$, $u_n \in \partial\Omega$ and $t_n \in [0, 1]$ such that

$$y_0 \in (1 - t_n)T(u_n) + t_n T_{\epsilon_n}(u_n) + f(u_n). \quad (1.2.31)$$

Hence, there exists w_n in $T_{\epsilon_n}(u_n)$ such that

$$y_0 = (1 - t_n)w_n + t_n T_{\epsilon_n}(u_n) + f(u_n). \quad (1.2.32)$$

Put $z_n = (1 - t_n)w_n + t_n T_{\epsilon_n}(u_n)$ and $v_n = T_{\epsilon_n}(u_n)$. Then $\{z_n\}$ is a bounded sequence because $\{f(u_n)\}$ is a bounded sequence and

$$y_0 = z_n + f(u_n). \quad (1.2.33)$$

Without loss of generality we can assume that $u_n \rightharpoonup u_0$, $z_n \rightharpoonup z_0$ and $t_n \rightarrow t$.

In the sequel, we will use a variant of arguments used in the proof of Theorem 1.2.28. In particular, we can show that

$$\langle z_0, u_0 \rangle \geq \limsup \langle z_n, u_n \rangle. \quad (1.2.34)$$

Fixing any (x, y) in $\text{Grp}(T)$, we claim that

$$\begin{aligned} \langle z_0, u_0 \rangle &\geq \limsup \langle z_n, u_n \rangle \geq \liminf \langle z_n, u_n \rangle \\ &\geq \langle y, u_0 - x \rangle + \langle z_0, x \rangle. \end{aligned} \quad (1.2.35)$$

Since $v_n = T_{\epsilon_n}(u_n)$, $v_n \in T(u_n - \epsilon_n J^{-1}(v_n))$. By the monotonicity of T , we get

$$\langle v_n - y, u_n - \epsilon_n J^{-1}(v_n) - x \rangle \geq 0$$

and

$$\langle w_n - y, u_n - x \rangle \geq 0.$$

It follows that

$$\langle v_n, u_n \rangle \geq \langle v_n, x \rangle + \langle y, u_n - x \rangle + \epsilon_n \|v_n\|^2 - \epsilon_n \|y\| \|v_n\| \quad (1.2.36)$$

and

$$\langle w_n, u_n \rangle \geq \langle w_n, x \rangle + \langle y, u_n - x \rangle. \quad (1.2.37)$$

Multiplying (1.2.35) with t_n , (1.2.36) with $(1 - t_n)$ and summing up, we get

$$\begin{aligned} \langle z_n, u_n \rangle &\geq \langle z_n, x \rangle + \langle y, u_n - x \rangle + \epsilon_n t_n \|v_n\|^2 - \epsilon_n t_n \|y\| \|v_n\| \\ &\geq \langle z_n, x \rangle + \langle y, u_n - x \rangle - \epsilon_n t_n \|y\| \|v_n\|. \end{aligned} \quad (1.2.38)$$

Case 1. The sequence $\{t_n v_n\}$ is bounded.

From (1.2.34) and (1.2.38) we get

$$\begin{aligned} \langle z_0, u_0 \rangle &\geq \limsup \langle z_n, u_n \rangle \\ &\geq \liminf \langle z_n, u_n \rangle \\ &\geq \liminf \langle z_n, x \rangle + \liminf \langle y, u_n - x \rangle + \liminf (-\epsilon_n t_n \|v_n\| \|y\|) \\ &\geq \langle y, u_0 - x \rangle + \langle z_0, x \rangle. \end{aligned}$$

This implies (1.2.35).

Case 2. The sequence $\{t_n v_n\}$ is unbounded.

From (1.2.36) we have

$$\|v_n\| \|u_n - x\| \geq \langle y, u_n - x \rangle + \epsilon_n \|v_n\|^2 - \epsilon_n \|y\| \|v_n\|.$$

It is equivalent to

$$\epsilon_n \|v_n\|^2 - \|v_n\|(\|u_n - x\| + \epsilon_n \|y\|) + \langle y, u_n - x \rangle \leq 0.$$

This implies that

$$\epsilon_n \|v_n\| \leq \frac{(\|u_n - x\| + \epsilon_n \|y\|) + \sqrt{\Delta_n}}{2}, \quad (1.2.39)$$

where

$$\Delta_n = (\|u_n - x\| + \epsilon_n \|y\|)^2 - 4\epsilon \langle y, u_n - x \rangle$$

which satisfies the inequality

$$\Delta_n \geq (\|u_n - x\| - \epsilon_n \|y\|)^2.$$

Since $u_n \rightharpoonup u_0$, (1.2.39) implies that the sequence $\{\epsilon_n \|v_n\|\}$ is bounded. We now can assume that $\lim \epsilon_n \|v_n\| = \alpha$. If $\alpha > 0$ then (1.2.38) implies

$$\begin{aligned} \langle z_0, u_0 \rangle &\geq \limsup \langle z_n, u_n \rangle \geq \liminf \langle z_n, u_n \rangle \\ &\geq \liminf \langle z_n, x \rangle + \liminf \langle y, u_n - x \rangle + \liminf (-\epsilon_n t_n \|v_n\| \|y\|) \\ &\quad + \liminf t_n \|v_n\| \epsilon_n \|v_n\| \\ &\geq \langle y, u_0 - x \rangle + \langle z_0, x \rangle - t\alpha \|y\| + \infty, \end{aligned}$$

which is impossible. Hence we must have $\alpha = 0$. Using (1.2.38) again, it yields

$$\begin{aligned} \langle z_0, u_0 \rangle &\geq \limsup \langle z_n, u_n \rangle \\ &\geq \liminf \langle z_n, u_n \rangle \\ &\geq \liminf \langle z_n, x \rangle + \liminf \langle y, u_n - x \rangle + \liminf (-\epsilon_n t_n \|v_n\| \|y\|) \\ &\geq \langle y, u_0 - x \rangle + \langle z_0, x \rangle - t\|y\|\alpha. \end{aligned}$$

This implies (1.2.35). Thus we have proved that for any (x, y) in $\text{Gr}(T)$, (1.2.35) holds. From (1.2.35) we get $\langle z_0 - y, u_0 - x \rangle \geq 0$. By the maximal monotonicity of T we have $z_0 \in T(u_0)$. Substituting $(x, y) = (u_0, z_0)$ into (1.2.35) we find that $\lim \langle z_n, u_n \rangle = \langle z_0, u_0 \rangle$. Since

$$\langle z_n, u_n \rangle = \langle z_n, u_0 \rangle + \langle y_0, u_n - u_0 \rangle - \langle f(u_n), u_n - u_0 \rangle,$$

we have $\lim \langle f(u_n), u_n - u_0 \rangle = 0$. Since f is of class $(S)_+$ and $\partial\Omega$ is closed, $u_n \rightarrow u_0 \in \partial\Omega$. On the other hand f is demicontinuous, we have $f(u_n) \rightharpoonup f(u_0)$. From (1.2.33), it yields $y_0 = z_0 + f(u_0)$ for some z_0 in $T(u_0)$. Hence $y_0 \in (f + T)(\partial\Omega)$, which contradicts our assumption. Thus, we must have $d_1(f + T, \Omega, y_0) = d_2(f + T, \Omega, y_0)$. The proof is complete. $\square$

In summary, we have proved the existence and uniqueness of the degree function in the sense of Definition 1.2.1 for mappings of type $f+T$. Our results improve and extend those in [35] to a larger class of maximal monotone mappings T and homotopy $\{T_t\}$ without assuming the graph of homotopy $\{T_t; 0 \leq t \leq 1\}$ having a nonempty common intersection. By putting $T = N_K$, where N_K is the normal cone of a closed convex set K , which is a maximal monotone mapping, we will obtain a degree theory for variational inequalities of type $f+N_K$ in reflexive Banach spaces with monotone operators. There are many applications and examples of the above results in this direction, and we refer the readers to [107].

1.3 Variational inequalities

1.3.1 Degree theory for generalized variational inequalities

In the last section, we built a degree theory for variational inequalities of type $f+N_K$ in reflexive Banach spaces with monotone operators. It is natural to ask whether we can build a degree theory for generalized variational inequality (GVI, for short) of the form $F+N_K$, where F is a multifunction. In this section we will focus on building a degree theory for GVI in finite-dimensional spaces.

In the two-volume book [80] dedicated entirely to finite-dimensional variational inequalities (VIs, for brevity), Facchinei and Pang have used the degree theory to obtain existence theorems for variational inequalities (see [80, Proposition 2.2.3 and Theorem 2.3.4]). These results gave a necessary and sufficient condition for a pseudomonotone VI on a general closed convex set to have solution. Also, based on degree theory, Robinson [201] gave a nice result on the stability of solutions to variational conditions; Gowda [94] proved inverse and implicit function theorems for H-differentiable functions, thereby giving a unified treatment of such a theorem for C^1 functions and for locally Lipschitzian functions. In order to obtain these results, the authors have used degree theory as a bridge to joint nonlinear analysis and variational inequality theory under which we can study problems via nonlinear equations.

Let K be a closed convex set in $\mathbb{R}^n$ and $f : K \rightarrow \mathbb{R}^n$ be a continuous mapping. The variational inequality defined by K and f denoted by $VI(f, K)$, is the problem of finding a vector $x \in K$ such that

$$\langle f(x), y - x \rangle \geq 0 \quad \forall y \in K$$

or equivalently

$$0 \in f(x) + N_K(x), \quad (1.3.1)$$

where $N_K(x)$ is the normal cone of K at x defined by formula (1.1.1). We denote by $P_K(x)$ the metric projection of x onto K and put

$$\Phi(x) = x - P_K[x - f(x)]. \quad (1.3.2)$$

Φ is called the *natural map*. It is clear that x is a solution of (1.3.1) if and only if x is a solution of the equation $\Phi(x) = 0$. Let Ω be an open bounded set in $\mathbb{R}^n$ such that $\Omega \cap K \neq \emptyset$. We wish to investigate the number of the solution of (1.3.1) in Ω . Since (1.3.1) is equivalent to the equation $\Phi(x) = 0$, it suggests us to compute the degree $d(\Phi, \Omega, 0)$. By the means, [80] and [201] obtained interesting results on the existence and stability of solutions to VIs.

It is natural to try to study generalized variational inequalities in this direction. Namely, we consider the problem of finding $x \in K$ such that

$$0 \in F(x) + N_K(x), \quad (1.3.3)$$

where $F : K \rightarrow 2^{\mathbb{R}^n}$ is a multifunction. In this case, we will face some difficulties for building a degree theory for GVIs. Namely, we can not apply the degree theory to the natural map Φ directly because $\Phi(x) := x - P_K[x - F(x)]$ may not have convex values and so that the degree of Φ is not well defined. To deal with the problem, we will use approximate selections f_ϵ of F and Φ_ϵ of Φ to build a degree theory for $\text{GVI}(F, K)$.

Note that there have been many papers on the degree theory for multifunctions in the infinite dimensional setting so far (see [35], [50], [53], [107] and [115]). We emphasize that the degree theory for GVIs here is built via the natural map, which is different from the degree theory for upper semicontinuous multifunctions with convex and compact values in previous papers.

• Degree theory for GVIs

Throughout this section, K is a closed convex set in $\mathbb{R}^n$, Ω is an open bounded set in $\mathbb{R}^n$ such that $\Omega \cap K \neq \emptyset$. Let $F : K \rightarrow 2^{\mathbb{R}^n}$ be a multifunction which is upper semicontinuous with compact convex values.

Recall that a map $F : K \rightarrow 2^{\mathbb{R}^n}$ is upper semicontinuous (u.s.c., for brevity) if for all $x \in K$ and for any open set $W \subset \mathbb{R}^n$ satisfying $F(x) \subset W$ there exists an open neighborhood U of x such that $F(y) \subset W$

for all $y \in U \cap K$. If $F(x) \neq \emptyset$ for all $x \in K$ and for any open set $W \subset \mathbb{R}^n$ satisfying $F(x) \cap W \neq \emptyset$, there exists an open neighborhood U of x such that $F(y) \cap W \neq \emptyset$ for all $y \in U \cap K$ then F is said to be lower semicontinuous (l.s.c., for brevity).

We have the following essential lemma.

Lemma 1.3.1. *Suppose that $F : K \rightarrow 2^{\mathbb{R}^n}$ is u.s.c. with closed convex values. Then for any $\epsilon > 0$ there exists a continuous map $f_\epsilon : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that for every $x \in K$ it holds*

$$f_\epsilon(x) \in F((x + \epsilon B) \cap K) + \epsilon B, \quad (1.3.4)$$

where B is the unit ball in $\mathbb{R}^n$.

Proof. By our assumptions and the approximate selection theorem due to A. Cellina (see [10]), for every $\varepsilon > 0$ there exists a continuous map $g_\varepsilon : K \rightarrow \mathbb{R}^n$ such that

$$g_\varepsilon(x) \in F((x + \varepsilon B) \cap K) + \varepsilon B \quad \forall x \in K.$$

By Tietze-Urysohn's theorem (see [78, Theorem 5.1 p. 149]), for each $\epsilon > 0$, there exists a continuous extension $f_\epsilon : \mathbb{R}^n \rightarrow \mathbb{R}^n$ of g_ϵ . As f_ϵ and g_ϵ agree on K , f_ϵ satisfies the conclusion of the theorem. The proof is complete. $\square$

Given a set valued map F and a closed convex set K , for each $\epsilon > 0$ we define a map $\Phi_\epsilon : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by the formula

$$\Phi_\epsilon(x) = x - P_K(x - f_\epsilon(x)), \quad (1.3.5)$$

where f_ϵ is an approximate continuous selection of F which satisfies (1.3.4). By the continuity of the metric projection and Lemma 1.3.1, Φ_ϵ is continuous on $\mathbb{R}^n$ and hence on $\bar{\Omega}$.

Here are some properties of Φ_ϵ .

Lemma 1.3.2. *Suppose that F is u.s.c. with compact convex values and $0 \notin (F + N_K)(\partial\Omega)$. Then the following assertions hold:*

- (a) *there exists $\epsilon_1 > 0$ such that $0 \notin \Phi_\epsilon(\partial\Omega)$ for all $\epsilon \in (0, \epsilon_1]$*
- (b) *there exists $\epsilon_2 > 0$ such that*

$$d(\Phi_\epsilon, \Omega, 0) = d(\Phi_{\epsilon'}, \Omega, 0) \text{ for all } \epsilon, \epsilon' \in (0, \epsilon_2]. \quad (1.3.6)$$

Proof. (a) Suppose the assertion is false. Then there exists a sequence $\epsilon_k \rightarrow 0^+$ and a sequence $x_k \in \partial\Omega$ such that $\Phi_{\epsilon_k}(x_k) = 0$. This means that

$$x_k = P_K(x_k - f_{\epsilon_k}(x_k)). \quad (1.3.7)$$

By the compactness of $\partial\Omega$ we can assume that $x_k \rightarrow x_0 \in \partial\Omega$. Since $x_k \in K \cap \partial\Omega$, by Lemma 1.3.1, there exist $y_k \in K$ and $z_k \in F(x_k)$ such that

$$\|y_k - x_k\| < \epsilon_k; \quad \|z_k - f_{\epsilon_k}(x_k)\| < \epsilon_k.$$

Hence $y_k \rightarrow x_0$. As $F(x_0)$ is a compact set and F is upper semicontinuous at x_0 , by taking a subsequence (if necessary) we can suppose furthermore that $z_k \rightarrow z_0 \in F(x_0)$. Hence $f_{\epsilon_k}(x_k) \rightarrow z_0$. Letting $k \rightarrow \infty$, from (2.4) we obtain $x_0 = P_K(x_0 - z_0)$ with $z_0 \in F(x_0)$. By the property of the metric projection we have

$$0 \in z_0 + N_K(x_0) \subset F(x_0) + N_K(x_0)$$

which contradicts our assumptions. We obtain the proof part (a).

(b) On the contrary, suppose that there exist sequences $0 < \epsilon_k < \epsilon'_k \rightarrow 0$ such that

$$d(\Phi_{\epsilon_k}, \Omega, 0) \neq d(\Phi_{\epsilon'_k}, \Omega, 0). \quad (1.3.8)$$

Put

$$H(t, x) = x - P_K(x - tf_{\epsilon_k}(x) - (1-t)f_{\epsilon'_k}(x)), \quad (t, x) \in [0, 1] \times \bar{\Omega}.$$

We have $H(0, x) = \Phi_{\epsilon'_k}(x)$ and $H(1, x) = \Phi_{\epsilon_k}(x)$. By the homotopy property (see Theorem 1.2.9), if $0 \notin H(t, \partial\Omega)$ for all $t \in [0, 1]$, then

$$d(H(0, \cdot), \Omega, 0) = d(H(1, \cdot), \Omega, 0).$$

But the later contradicts (1.3.8). Hence for each k , there exists $t_k \in [0, 1]$ such that $0 \in H(t_k, \partial\Omega)$. This implies that, for each k there exists $x_k \in \partial\Omega$ such that

$$x_k = P_K(x_k - t_k f_{\epsilon_k}(x_k) - (1 - t_k) f_{\epsilon'_k}(x_k)). \quad (1.3.9)$$

Since $x_k \in K \cap \partial\Omega$, by Lemma 1.3.1, there exist $y_k, y'_k \in K$; $z_k \in F(y_k)$ and $z'_k \in F(y'_k)$ such that

$$\|y_k - x_k\| < \epsilon'_k; \quad \|z_k - f_{\epsilon_k}(x_k)\| < \epsilon'_k;$$

and

$$\|y'_k - x_k\| < \epsilon'_k; \quad \|z'_k - f_{\epsilon'_k}(x_k)\| < \epsilon'_k.$$

By the compactness of $[0, 1] \times \partial\Omega$ we can assume that $(t_k, x_k) \rightarrow (\bar{t}, \bar{x}) \in [0, 1] \times \partial\Omega$. Hence $y_k \rightarrow \bar{x}$ and $y'_k \rightarrow \bar{x}$. Using the standard arguments as the proof of (a) we get $f_{\epsilon_k}(x_k) \rightarrow z_1$ and $f_{\epsilon'_k}(x_k) \rightarrow z_2$ for some $z_1, z_2 \in F(\bar{x})$. By letting $k \rightarrow \infty$, from (1.3.9) we get $\bar{x} = P_K(\bar{x} - \bar{t}z_1 - (1 - \bar{t})z_2)$.

Put $\bar{z} = \bar{t}z_1 + (1 - \bar{t})z_2$ then $\bar{z} \in F(\bar{x})$ and $\bar{x} = P_K(\bar{x} - \bar{z})$. By the property of the metric projection we have

$$0 \in \bar{z} + N_K(\bar{x}) \subset F(\bar{x}) + N_K(\bar{x}).$$

Since $\bar{x} \in \partial\Omega$, we get a contradiction. The proof of the lemma is now complete. $\square$

From Lemma 1.3.2 it follows that there exists $\bar{\epsilon} > 0$ such that $0 \notin \Phi_{\epsilon}(\partial\Omega)$ and $d(\Phi_{\epsilon}, \Omega, 0) = d(\Phi_{\epsilon'}, \Omega, 0)$ for all $\epsilon, \epsilon' \in (0, \bar{\epsilon}]$. It is a basis for the following definition.

Definition 1.3.3. Let $F : K \rightarrow 2^{\mathbb{R}^n}$ be an u.s.c. multifunction with compact convex values and $0 \notin (F + N_K)(\partial\Omega)$. The degree of generalized variational inequality defined by F and K respect to Ω at 0 is the common value $d(\Phi_{\epsilon}, \Omega, 0)$ for $\epsilon > 0$ sufficiently small and denoted by $d(F + N_K, \Omega, 0)$.

Example 1.3.4. Let

$$F(x) = \begin{cases} \{1\} & \text{if } x > 0 \\ [-1, 1] & \text{if } x = 0 \\ \{-1\} & \text{if } x < 0, \end{cases}$$

$K = [-1, 1]$ and $\Omega = (-1/2, 2)$. Then $d(F + N_K, \Omega, 0) = 1$.

Indeed, for each $\epsilon > 0$ we consider the following function

$$f_{\epsilon}(x) = \begin{cases} 1 & \text{if } x \geq \epsilon \\ x/\epsilon & \text{if } x \in (-\epsilon, \epsilon) \\ -1 & \text{if } x \leq -\epsilon. \end{cases}$$

If $x \in K \setminus (-\epsilon, \epsilon)$ then $(x, f_{\epsilon}(x)) = (x, 1) \in \text{Gr}(F)$ and so

$$\text{dist}((x, 1), \text{Gr}(F)) = 0.$$

If $x \in K \cap (-\epsilon, \epsilon)$ then $(x, f_{\epsilon}(x)) = (x, x/\epsilon)$. Since $(0, x/\epsilon) \in \text{Gr}(F)$,

$$\text{dist}((x, x/\epsilon), \text{Gr}(F)) \leq \text{dist}((x, x/\epsilon), (0, x/\epsilon)) = |x| < \epsilon.$$

Thus f_{ϵ} are approximate selections of F . We will compute $\Phi_{\epsilon} = x - P_K(x - f_{\epsilon}(x))$. Choose $\bar{\epsilon} = 1$ and take any $\epsilon \in (0, \bar{\epsilon}]$. We have the following cases:

If $x \in (-\epsilon, \epsilon)$ then $|x - f_{\epsilon}(x)| \leq 1$. So $P_K(x - f_{\epsilon}(x)) = 0$.

If $\epsilon \leq x \leq 2$ then $P_K(x - f_\epsilon(x)) = 0$.

If $x > 2$ then $P_K(x - f_\epsilon(x)) = x - 2$.

If $-2 \leq x \leq -\epsilon$ then $P_K(x - f_\epsilon(x)) = 0$.

If $x < -2$ then $P_K(x - f_\epsilon(x)) = -x - 2$.

From the above we obtain

$$P_K(x - f_\epsilon(x)) = \begin{cases} 0 & \text{if } x \in [-2, 2] \\ x - 2 & \text{if } x > 2 \\ -x - 2 & \text{if } x < -2. \end{cases}$$

Hence

$$\Phi_\epsilon(x) = \begin{cases} x & \text{if } x \in [-2, 2] \\ 2 & \text{if } x > 2 \\ 2x + 2 & \text{if } x < -2. \end{cases}$$

Note that $\partial\Omega = \{-1/2; 2\}$; $F(-1/2) + N_K(-1/2) = \{1\}$ and $F(2) + N_K(2) = \emptyset$. Hence $0 \notin (F + N_K)(\partial\Omega)$. We now compute $d(\Phi_\epsilon, \Omega, 0)$. As Φ_ϵ is differentiable in Ω we get

$$d(\Phi_\epsilon, \Omega, 0) = \sum_{x \in \Phi_\epsilon^{-1}(0)} \text{sign} \Phi'_\epsilon(x) = 1.$$

Since the later equality is true for all $\epsilon \in (0, \bar{\epsilon}]$, we obtain

$$d(F + N_K, \Omega, 0) = 1.$$

We have the following theorem on existence.

Theorem 1.3.5. *Suppose that $0 \notin (F + N_K)(\partial\Omega)$. Then the following assertions hold:*

(a) (Existence) *if $d(F + N_K, \Omega, 0) \neq 0$ then there exists $x \in \Omega \cap K$ such that*

$$0 \in F(x) + N_K(x).$$

(b) *if $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a continuous map such that $f(x) \in F(x)$ for all $x \in K$ then $d(F + N_K, \Omega, 0) = d(\Phi, \Omega, 0)$, where $\Phi(x) = x - P_K(x - f(x))$.*

Proof. (a) By the definition, there exists $\bar{\epsilon} > 0$ such that $d(F + N_K, \Omega, 0) = d(\Phi_\epsilon, \Omega, 0)$ for all $\epsilon \in (0, \bar{\epsilon}]$. Let $\{\epsilon_k\}$ be a sequence such that $\epsilon_k \rightarrow 0^+$. Then $d(\Phi_{\epsilon_k}, \Omega, 0) \neq 0$ for k sufficiently large. By (b) in Theorem 1.2.9, there exists $x_k \in \Omega$ such that $\Phi_{\epsilon_k}(x_k) = 0$. This is equivalent to

$$x_k = P_K(x_k - f_{\epsilon_k}(x_k)). \quad (1.3.10)$$

Since $x_k \in K \cap \Omega$, by Lemma 1.3.1, there exists $y_k \in K$ and $z_k \in F(y_k)$ such that

$$\|y_k - x_k\| < \epsilon_k; \|z_k - f_{\epsilon_k}(x_k)\| < \epsilon_k.$$

By the compactness of $K \cap \bar{\Omega}$ we can assume that $x_k \rightarrow x_0 \in K \cap \bar{\Omega}$. Hence $y_k \rightarrow x_0$. By the standard arguments we get $z_k \rightarrow z_0$ and $f_{\epsilon_k}(x_k) \rightarrow z_0$ for some $z_0 \in F(x_0)$. Letting $k \rightarrow \infty$, from (1.3.10) we obtain $x_0 = P_K(x_0 - z_0)$. The property of the metric projection yields

$$0 \in z_0 + N_K(x_0) \subset F(x_0) + N_K(x_0).$$

Since $0 \notin (F + N_K)(\partial\Omega)$ we have $x_0 \in K \cap \Omega$.

(b) By putting $f_\epsilon = f$ for all $\epsilon > 0$ we get the desired property. The proof of the theorem is complete. $\square$

Example 1.3.6. Consider Example 1.3.4 we have $d(F + N_K, \Omega, 0) = 1$. By the above theorem, $GVI(F, K)$ has a solution $x \in \Omega \cap K$. In this case, $x = 0$ is a solution.

The following theorem contains most usual properties of degree theory.

Theorem 1.3.7. Assume that $0 \notin (F + N_K)(\partial\Omega)$. The following assertions hold:

(a) (Homotopy invariance) If $F_1, F_2 : K \rightarrow 2^{\mathbb{R}^n}$ are u.s.c. multifunctions with compact convex values and $0 \notin (tF_1 + (1-t)F_2 + N_K)(\partial\Omega)$ for all $t \in [0, 1]$ then

$$d(F_1 + N_K, \Omega, 0) = d(F_2 + N_K, \Omega, 0).$$

(b) (Additivity) If Ω_1, Ω_2 are disjoint open subsets of Ω such that $0 \notin (F + N_K)(\bar{\Omega} \setminus (\Omega_1 \cup \Omega_2))$ then

$$d(F + N_K, \Omega, 0) = d(F + N_K, \Omega_1, 0) + d(F + N_K, \Omega_2, 0).$$

(c) (Excision) If $D \subset \Omega$ is a closed set such that $0 \notin (F + N_K)(D)$ then

$$d(F + N_K, \Omega, 0) = d(F + N_K, \Omega \setminus D, 0)$$

Proof. (a) Let $f_\epsilon, g_\epsilon : R \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be approximate selections of F_1 and F_2 , respectively satisfying the conclusion of Lemma 1.3.1. Put

$$\Phi_\epsilon^t(x) = x - P_K(x - tf_\epsilon(x) - (1-t)g_\epsilon(x)).$$

We claim that there is $\bar{\epsilon} > 0$ such that $0 \notin \Phi_\epsilon^t(\partial\Omega)$ for all $\epsilon \in [0, \bar{\epsilon}]$ and $t \in [0, 1]$. In fact, if the claim is false then there exist a sequence $t_k \in [0, 1]$ and a sequence $\epsilon_k \rightarrow 0^+$ such that $0 \in \Phi_{\epsilon_k}^{t_k}(\partial\Omega)$. Hence for each k , there exists $x_k \in \partial\Omega$ such that

$$x_k = P_K(x_k - t_k f_{\epsilon_k}(x_k) - (1 - t_k)g_{\epsilon_k}(x_k)).$$

By the compactness of $[0, 1] \times \partial\Omega$ we can assume that $(t_k, x_k) \rightarrow (t_0, x_0) \in [0, 1] \times \partial\Omega$. By the standard arguments we can show that $f_{\epsilon_k}(x_k) \rightarrow z_1$ for some $z_1 \in F_1(x_0)$ and $g_{\epsilon_k}(x_k) \rightarrow z_2$ for some $z_2 \in F_2(x_0)$. Letting $k \rightarrow \infty$ from the above we obtain

$$x_0 = P_K(x_0 - t_0 z_1 - (1 - t_0)z_2).$$

By the property of the metric projection we get

$$0 \in t_0 z_1 + (1 - t_0)z_2 + N_K(x_0) \subset t_0 F_1(x_0) + (1 - t_0)F_2(x_0) + N_K(x_0).$$

This contradicts the assumption and so our claim is proved. We now can apply (d) of Theorem 1.2.9 for Φ_ϵ^t to get $d(\Phi_\epsilon^0, \Omega, 0) = d(\Phi_\epsilon^1, \Omega, 0)$ for all $\epsilon \in (0, \bar{\epsilon}]$. Hence $d(F_1 + N_K, \Omega, 0) = d(F_2 + N_K, \Omega, 0)$.

(b) We will show that there exists $\bar{\epsilon} > 0$ such that $0 \notin \Phi_\epsilon(\bar{\Omega} \setminus (\Omega_1 \cup \Omega_2))$ for all $\epsilon \in (0, \bar{\epsilon}]$. Indeed, if the assertion is false then there exists a sequence $\epsilon_k \rightarrow 0^+$ and $x_k \in \bar{\Omega} \setminus (\Omega_1 \cup \Omega_2)$ such that $x_k = P_K(x_k - f_{\epsilon_k}(x_k))$. By the compactness of $\bar{\Omega}$ we can assume that $x_k \rightarrow x_0 \in \bar{\Omega}$. If $x_0 \in \Omega_1 \cup \Omega_2$ then $x_k \in \Omega_1 \cup \Omega_2$ for k sufficiently large. This contradicts the fact that $x_k \in \bar{\Omega} \setminus (\Omega_1 \cup \Omega_2)$. Hence $x_0 \in \bar{\Omega} \setminus (\Omega_1 \cup \Omega_2)$. By the standard arguments we have $f_{\epsilon_k}(x_k) \rightarrow z_0$ for some $z_0 \in F(x_0)$. Letting $k \rightarrow \infty$ from the above we obtain $x_0 = P_K(x_0 - z_0)$. This implies that $0 \in F(x_0) + N_K(x_0)$ for some $x_0 \in \bar{\Omega} \setminus (\Omega_1 \cup \Omega_2)$, which is a contradiction.

Thus we have $0 \notin \Phi_\epsilon(\bar{\Omega} \setminus (\Omega_1 \cup \Omega_2))$ for all $\epsilon \in (0, \bar{\epsilon}]$. By (c) of Theorem 1.2.9, we get

$$d(\Phi_\epsilon, \Omega, 0) = d(\Phi_\epsilon, \Omega_1, 0) + d(\Phi_\epsilon, \Omega_2, 0).$$

It follows that

$$d(F + N_K, \Omega, 0) = d(F + N_K, \Omega_1, 0) + d(F + N_K, \Omega_2, 0).$$

(c) By the standard arguments we show that $0 \notin \Phi_\epsilon(D)$ for all $\epsilon \in (0, \bar{\epsilon}]$. Applying (e) of Theorem 1.2.9 for Φ_ϵ we obtain the desired conclusion. The proof of the theorem is complete. $\square$

Definition 1.3.8. A vector $x_0 \in K$ is called an isolated solution of $GVI(F, K)$ if there exists a neighborhood V of x_0 such that x_0 is the unique solution of $GVI(F, K)$ in $\bar{V}$.

Theorem 1.3.9. Suppose that x_0 is an isolated solution of $GVI(F, K)$ and $\mathcal{U}$ is the collection of all open bounded neighborhoods V of x_0 such that $\bar{V}$ does not contain another solution of $GVI(F, K)$. Then

$$d(F + N_K, V_1, 0) = d(F + N_K, V_2, 0)$$

for all $V_1, V_2 \in \mathcal{U}$. The common value $d(F + N_K, V, 0)$ for $V \in \mathcal{U}$ is called the index of $F + N_K$ and denoted by $i(F + N_K, x_0, 0)$.

Proof. We will use some arguments as in [88] for the proof below .

Taking any $V \in \mathcal{U}$ we have $0 \notin (F + N_K)(\partial V)$. Therefore, $d(F + N_K, V, 0)$ is well defined. We now assume that $V_1, V_2 \in \mathcal{U}$. Put $V = V_1 \cup V_2 \in \mathcal{U}$ and $D = \bar{V}_1 \cap V_2^c$, where $V_2^c = \mathbb{R}^n \setminus V_2$. We have that D is a compact set in $\bar{V}$ and $0 \notin (F + N_K)(D)$. By (c) in Theorem 1.3.7, we get

$$d(F + N_K, V, 0) = d(F + N_K, V \setminus D, 0) = d(F + N_K, V_2, 0).$$

Using a similar argument for $D = \bar{V}_2 \cap V_1^c$ we get

$$d(F + N_K, V, 0) = d(F + N_K, V \setminus D, 0) = d(F + N_K, V_1, 0).$$

Thus $d(F + N_K, V_1, 0) = d(F + N_K, V_2, 0)$. □

• Applications

In this section, we shall employ the obtained results in Section 2 to prove some facts on the existence and stability of solutions to GVIs.

The following theorem is an extension of [80, Pr. 2.2.3].

Theorem 1.3.10. Let $K \subset \mathbb{R}^n$ be a nonempty closed convex set and $F : K \rightarrow 2^{\mathbb{R}^n}$ be an u.s.c. multifunction with nonempty compact convex values. Assume that there exist an open bounded set Ω and a vector $\hat{x} \in K \cap \Omega$ such that

$$\inf_{x^* \in F(x)} \langle x^*, x - \hat{x} \rangle > 0 \quad \forall x \in K \cap \partial\Omega. \quad (1.3.11)$$

The $GVI(F, K)$ has a solution in $K \cap \bar{\Omega}$.

Proof. If $0 \in (F + N_K)(\partial\Omega)$, then $GVI(F, K)$ is apparent to have a solution. Otherwise the degree $d(F + N_K, \Omega, 0)$ is well defined. Hence there exists $\epsilon_1 > 0$ such that $0 \notin \Phi_\epsilon(\partial\Omega)$ for all $\epsilon \in (0, \epsilon_1]$. Recall that $\Phi_\epsilon(x) = x - P_K(x - f_\epsilon(x))$, where f_ϵ is approximate selection of F which is continuous on $\mathbb{R}^n$.

We claim that there exists $\epsilon_2 > 0$ such that for every $\epsilon \in (0, \epsilon_2]$ it holds

$$\langle f_\epsilon(x), x - \hat{x} \rangle \geq 0 \quad \forall x \in K \cap \partial\Omega. \quad (1.3.12)$$

Indeed, if the assertion is false then there exist sequences $\epsilon_k \rightarrow 0^+$ and $x_k \in K \cap \partial\Omega$ such that

$$\langle f_{\epsilon_k}(x_k), x_k - \hat{x} \rangle < 0 \quad \forall k \in \mathbb{N}. \quad (1.3.13)$$

By Lemma 1.3.1, there exists $(y_k, z_k) \in \text{Gr}(F)$ such that $\|y_k - x_k\| < \varepsilon_k$ and $\|f_{\epsilon_k}(x_k) - z_k\| < \varepsilon_k$. By the compactness of $K \cap \partial\Omega$ we may suppose that there exists $\bar{x} \in K \cap \partial\Omega$ such that $x_k \rightarrow \bar{x}$. Then $y_k \rightarrow \bar{x}$. As $F(\bar{x})$ is a compact set and F is upper semicontinuous at $\bar{x}$, by taking a subsequence (if necessary) we can suppose furthermore that $z_k \rightarrow \bar{z} \in F(\bar{x})$. Then $f_{\epsilon_k}(x_k) \rightarrow \bar{z}$. Letting $k \rightarrow \infty$, from (1.3.13) we obtain $\langle \bar{z}, \bar{x} - \hat{x} \rangle \leq 0$, hence

$$\inf_{x^* \in F(\bar{x})} \langle x^*, \bar{x} - \hat{x} \rangle \leq 0.$$

This contradicts (1.3.11) and our claim is obtained.

Put $\bar{\epsilon} = \min\{\epsilon_1, \epsilon_2\}$. We now show that $d(\Phi_\epsilon, \Omega, 0) = 1$ for all $\epsilon \in (0, \bar{\epsilon}]$ and so $d(F + N_K, \Omega, 0) = 1$. For this we build a homotopy as in [80].

Fix any $\epsilon \in (0, \bar{\epsilon}]$ and put

$$H(t, x) = x - P_K(t(x - f_\epsilon(x) + (1 - t)\hat{x})), \quad (t, x) \in [0, 1] \times \bar{\Omega}.$$

We have $H(0, x) = x - \hat{x}$ and $H(1, x) = \Phi_\epsilon(x)$. Note that $d(H(0, \cdot), \Omega, 0) = 1$. We now claim that $0 \notin H(t, \partial\Omega)$ for all $t \in [0, 1]$. In fact, it is obvious that $0 \notin H(0, \partial\Omega)$ and $0 \notin H(1, \partial\Omega)$. Assume that there exist $t \in (0, 1)$ and $x \in \partial\Omega$ such that $0 = H(t, x)$. By the property of the metric projection we have

$$\langle x - t(x - f_\epsilon(x) - (1 - t)\hat{x}), y - x \rangle \geq 0 \quad \forall y \in K.$$

In particular, for $y = \hat{x}$ we get

$$\langle tf_\epsilon(x) + (1 - t)(x - \hat{x}), \hat{x} - x \rangle \geq 0.$$

This implies

$$\langle f_\epsilon(x), \hat{x} - x \rangle \geq \frac{1-t}{t} \|x - \hat{x}\|^2 > 0$$

where the last inequality holds because $t \in (0, 1)$ and $x \neq \hat{x}$. But then it follows that $\langle f_\epsilon(x), x - \hat{x} \rangle < 0$ which contradicts (1.3.12). Thus $0 \notin H(t, \partial\Omega)$ for all $t \in [0, 1]$. By the homotopy invariance (see (d) in Theorem 1.2.9) we obtain $d(H(0, \cdot), \Omega, 0) = d(H(1, \cdot), \Omega, 0) = 1$.

In summary, we have proved that $d(\Phi_\epsilon, \Omega, 0) = 1$ for all $\epsilon \in (0, \bar{\epsilon}]$. By the degree definition of GVIs we have $d(F + N_K, \Omega, 0) = 1$. According to Theorem 1.3.5, there exists $x_0 \in \Omega \cap K$ such that $0 \in F(x_0) + N_K(x_0)$. The proof of the theorem is complete. $\square$

In the rest of this subsection we will present a result on the stability of solutions to GVIs. Let us assume that M and Λ are subsets of R^k and R^m respectively; $F : M \times \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ and $K : \Lambda \rightarrow 2^{\mathbb{R}^n}$ be multifunctions. Consider the parametric generalized variational inequality

$$0 \in F(\mu, x) + N_{K(\lambda)}(x), \quad (1.3.14)$$

where $N_{K(\lambda)}(x)$ is the value at x of the normal cone operator associated with the set $K(\lambda)$ and $(\mu, \lambda) \in M \times \Lambda$ are parameters. We denote by $S(\mu, \lambda)$ the solution set of the problem (1.3.14) corresponding to (μ, λ) and suppose that $x_0 \in S(\mu_0, \lambda_0)$ for a given $(\mu_0, \lambda_0) \in M \times \Lambda$.

Our main concern is now to investigate the behaviour of $S(\mu, \lambda)$ when (μ, λ) vary around (μ_0, λ_0) . This problem was interesting to many authors in the last two decades. For the relevant literature of the problem we refer the reader to [121] - [149], [201]- [230] and several references given therein.

The following result gives a sufficient condition for the lower semi-continuity of the solution map of (1.3.14). It is an extension of results in [121] and [201] for the case of GVIs.

Theorem 1.3.11. *Assume that X_0 , Λ_0 and M_0 are neighborhoods of x_0 , λ_0 and μ_0 respectively and the following conditions are satisfied:*

- (i) *$F(\cdot, \cdot)$ is l.s.c. on $M_0 \times X_0$ with closed convex values and $F(\mu_0, \cdot)$ is u.s.c. with compact convex values;*
- (ii) *$K : \Lambda_0 \rightarrow 2^{\mathbb{R}^n}$ has closed convex values and to be pseudo-Lipshitz continuous around (λ_0, x_0) , i.e., there exist neighborhoods V of λ_0 , W of x_0 and a constant $k > 0$ such that*

$$K(\lambda) \cap W \subset K(\lambda') + k\|\lambda - \lambda'\|B(0, 1) \quad \forall \lambda, \lambda' \in V \cap \Lambda;$$

(iii) x_0 is an isolated solution and there exists $\bar{\sigma} > 0$ such that

$$i(\sigma F(\mu_0, \cdot) + N_{K(\lambda_0) \cap X_0}, x_0, 0) \neq 0 \quad \forall \sigma \in (0, \bar{\sigma}].$$

Then there exist a neighborhood $\tilde{M}$ of μ_0 , a neighborhood $\tilde{\Lambda}$ of λ_0 and an open bounded neighborhood $\tilde{X}$ of x_0 such that the solution map $\hat{S} : \tilde{M} \times \tilde{\Lambda} \rightarrow 2^{X_0}$ of (1.3.14) defined by $\hat{S}(\mu, \lambda) = S(\mu, \lambda) \cap \tilde{X}$ has nonempty values and lower semicontinuous at (μ_0, λ_0) .

Proof. By (i) and the continuous selection theorem due to Michael (see [236, Theorem 9G, p. 466]), there exists a continuous mapping $f : M_0 \times X_0 \rightarrow \mathbb{R}^n$ such that $f(\mu, x) \in F(\mu, x)$ for all $(\mu, x) \in M_0 \times X_0$. By Tietze-Urysohn's theorem (see [78, Theorem 5.1 p. 149]) we can assume that f is continuous on $\mathbb{R}^k \times \mathbb{R}^n$.

According to Lemma 1.1 in [229] and (ii), there exist a neighborhood $\Lambda'_0 \subset \Lambda_0 \cap V$ of λ_0 , a neighborhood $X'_0 \subset X_0 \cap W$ of x_0 and a constant $k_0 > 0$ such that

$$\|P_{K(\lambda) \cap X_0}(z) - P_{K(\lambda') \cap X_0}(z)\| \leq k_0 \|\lambda - \lambda'\|^{1/2}$$

for all $\lambda, \lambda' \in \Lambda'_0$ and $z \in X'_0$. Let us define a mapping $\pi : \Lambda'_0 \times X'_0 \rightarrow X_0$ by setting

$$\pi(\lambda, z) = P_{K(\lambda) \cap X_0}(z).$$

Then for any $z, z' \in X'_0$ and $\lambda, \lambda' \in \Lambda'_0$ we have

$$\begin{aligned} \|\pi(\lambda, z) - \pi(\lambda', z')\| &= \|P_{K(\lambda) \cap X_0}(z) - P_{K(\lambda') \cap X_0}(z')\| \leq \\ &\leq \|P_{K(\lambda) \cap X_0}(z) - P_{K(\lambda) \cap X_0}(z')\| + \|P_{K(\lambda) \cap X_0}(z') - P_{K(\lambda') \cap X_0}(z')\| \\ &\leq \|z - z'\| + k_0 \|\lambda - \lambda'\|^{\frac{1}{2}}. \end{aligned}$$

Consequently, $\pi : \Lambda'_0 \times X'_0 \rightarrow X_0$ is uniformly continuous on $\Lambda'_0 \times X'_0$.

Choose $\sigma_0 \in (0, \bar{\sigma}]$ such that $x_0 - \sigma_0 f(\mu_0, x_0) \in X'_0$. By the continuity of f , there exist a neighborhood $X_1 \subset X'_0$ of x_0 , a neighborhood $M'_0 \subset M_0$ of μ_0 such that

$$x - \sigma_0 f(\mu, x) \in X_1 \quad \forall (x, \mu) \in X_1 \times M'_0.$$

Consider the function

$$\Phi_{\sigma_0}(\mu, \lambda, x) = x - P_{K(\lambda) \cap X_0}(x - \sigma_0 f(\mu, x))$$

with $(\mu, \lambda, x) \in M'_0 \times \Lambda'_0 \times X_1$. By the above, Φ_{σ_0} is continuous on $M'_0 \times \Lambda'_0 \times X_1$.

From (iii) and Theorem 1.3.9, there exists an open bounded neighborhood $X_2 \subset X_1$ of x_0 such that x_0 is the unique solution in $\bar{X}_2$ of the inclusion

$$0 \in F(\mu_0, x) + N_{K(\lambda_0)}(x).$$

This means that x_0 is the unique solution in $\bar{X}_2$ of the inclusion

$$0 \in \sigma_0 F(\mu_0, x) + N_{K(\lambda_0)}(x).$$

Since x_0 belongs to the interior of X_0 , it is also the unique solution in $\bar{X}_2$ of the inclusion

$$0 \in \sigma_0 F(\mu_0, x) + N_{K(\lambda_0) \cap X_0}(x).$$

Moreover,

$$d(\sigma_0 F(\mu_0, \cdot) + N_{K(\lambda_0) \cap X_0}, X_2, 0) = i(\sigma_0 F(\mu_0, \cdot) + N_{K(\lambda_0) \cap X_0}, x_0, 0) \neq 0.$$

As $\sigma_0 f(\mu_0, x) \in \sigma_0 F(\mu_0, x)$ for all $x \in K(\lambda_0) \cap X_0$ and Theorem 2.1, we get

$$d(\Phi_{\sigma_0}(\mu_0, \lambda_0, \cdot), X_2, 0) = d(\sigma_0 F(\mu_0, \cdot) + N_{K(\lambda_0) \cap X_0}, X_2, 0) \neq 0. \quad (1.3.15)$$

Note that every solution of equation $\Phi_{\sigma_0}(\mu, \lambda, x) = 0$ are also solutions of $GVI(F(\mu, \cdot), K(\lambda) \cap X_0)$. Hence x_0 is the unique solution of the equation $\Phi_{\sigma_0}(\mu_0, \lambda_0, x) = 0$ in X_2 . Put $\tilde{X} = X_2$. Taking any $w \in \partial \tilde{X}$, we have $\Phi_{\sigma_0}(\mu_0, \lambda_0, w) \neq 0$. This implies that there exist a $\delta_w > 0$ such that $0 \notin B(\Phi_{\sigma_0}(\mu_0, \lambda_0, w), \delta_w) := B_w$. By the continuity of Φ_{σ_0} , there exist a neighborhood $U_w \subset M'_0$ of μ_0 , a neighborhood $\Lambda_w \subset \Lambda'_0$ of λ_0 and a neighborhood Q_w of w such that $\Phi_{\sigma_0}(\mu, \lambda, z) \in B_w$ for all $(\mu, \lambda, z) \in U_w \times \Lambda_w \times Q_w$. Since $\partial \tilde{X}$ is a compact set, there are some $w_1, w_2, \dots, w_n$ such that $\partial \tilde{X} \subset \cup_{i=1}^n Q_{w_i}$. Put $\tilde{M} = \cap_{i=1}^n U_{w_i}$, $\tilde{\Lambda} = \cap_{i=1}^n \Lambda_{w_i}$.

We now use the similar arguments in the proof of Theorem 2.1 in [121] (see also [201, Theorem 3.2]) to show that $\tilde{M}, \tilde{\Lambda}$ and $\tilde{X}$ satisfy the conclusion of the theorem. $\square$

1.3.2 Existence of solutions to variational inequality with pseudo-monotone operators in reflexive Banach spaces

Variational inequalities (VIs, for brevity) and generalized variational inequalities (GVIs) have been recognized as suitable mathematical models for dealing with many problems arising in different fields, such as optimization theory, partial differential equations, economic equilibrium,

mechanics, etc. In the last five decades, since the initial existence results for variational inequalities were established by Stampacchia in 1964 [213], Lion and Stampacchia in 1965 [153] (see, also [89, 101] and [154], Theorem 4.1), solution existence of VIs, GVIs, and other related problems has become a basic research topic which continues to attract attention of researchers in applied mathematics (see, for instance, [139]–[233] and the references therein). Some difficult questions do exist in this field (see, for example, [192] and [232]).

Let us assume that X is a topological vector space which is endowed with weak topology and that X^* is its dual. Suppose that $K \subset X$ is a nonempty closed convex set and $\Phi : K \rightarrow 2^{X^*}$ is a multifunction from K to X^* (which is equipped with the weak* topology).

The *generalized variational inequality* defined by K and Φ , denoted by $\text{GVI}(K, \Phi)$, is the problem of finding a point $x \in K$ such that

$$0 \in \Phi(x) + N_K(x),$$

or equivalently,

$$\exists x^* \in \Phi(x), \quad \langle x^*, y - x \rangle \geq 0, \quad \forall y \in K, \quad (1.3.16)$$

where $N_K(x)$ is the normal cone to K at x , which is given by (1.1.1). Here, $\langle \cdot, \cdot \rangle$ denotes the canonical pairing between X^* and X . The set of all $x \in K$ that satisfy (1.3.16) is denoted by $\text{SOL}(K, \Phi)$. If $\Phi(x) = \{f(x)\}$ for all $x \in K$, where $f : K \rightarrow X^*$ is a single-valued map, then the problem $\text{GVI}(K, \Phi)$ is called a variational inequality, and the abbreviation $\text{VI}(K, f)$ is used instead of $\text{GVI}(K, \Phi)$.

Recall that Φ is a *monotone operator* if for any $x, y \in K$ and $x^* \in \Phi(x), y^* \in \Phi(y)$ one has $\langle x^* - y^*, x - y \rangle \geq 0$. The first substantial results concerning monotone operators were obtained by Minty [165] and Browder [33]. Then the properties of monotone operators were studied systematically by Browder in order to obtain existence theorems for quasi-linear elliptic and parabolic partial differential equations. The existence theorems of Browder were generalized to more general classes of quasi-linear elliptic differential equations by Hartman and Stampacchia (see [101, 139]).

In 1968, Brézis [29] introduced a vast class of operators of pseudomonotone type as follows. An operator $\Phi : K \rightarrow 2^{X^*}$ is called pseudomonotone if, for any $u \in K$ and every net $\{u_i\}$ with $u_i \rightarrow u$, $u_i^* \in \Phi(u_i)$, and $\limsup \langle u_i^*, u_i - u \rangle \leq 0$, then for each $v \in K$, there exists

$v^* \in \Phi(u)$ such that

$$\langle v^*, u - v \rangle \leq \liminf \langle u_i^*, u_i - v \rangle. \quad (1.3.17)$$

The theory of pseudomonotone operators plays an important role in the study of solvability of operator equations and quasi-linear elliptic equations. This class has been studied systematically. We refer the readers to [34] and Chap. 27 in [238] for more information and properties of the class.

Beside the concept of pseudomonotone operators in the sense of Brézis, another concept of pseudomonotone operators was proposed by Karamardian in 1976 (see [114]). An operator $\Phi : K \rightarrow 2^{X^*}$ is said to be pseudomonotone if, for any $x, y \in K$ and $x^* \in \Phi(x), y^* \in \Phi(y)$, the following implication holds:

$$\langle y^*, x - y \rangle \geq 0 \implies \langle x^*, x - y \rangle \geq 0. \quad (1.3.18)$$

It is clear that monotonicity implies pseudomonotonicity in the sense of Karamardian. The converse implication is not true in general (take, for instance, $K = \mathbb{R}$ and $f(x) = x^2 + 1$ for all $x \in K$).

The existence of solutions to variational inequalities with pseudomonotone operators in the sense of Karamardian has been investigated intensively twenty years ago (see for instance [12, 60, 81, 82, 97, 124, 140, 227, 228] and references therein). In particular, by the Minty lemma, the authors in [228] and [124] gave some results on the existence of solutions to VIs for pseudomonotone operators in the reflexive Banach spaces.

Our goal in this section is to give some new results on the existence of solutions to VIs and GVIs with for pseudomonotone operators in the sense of Brézis in reflexive Banach spaces. We derive a new scheme for proofs which is based on the Galerkin method. Using this scheme, we can establish some existence theorems for VIs and GVIs in topological vector spaces and reflexive Banach spaces. The obtained results extend preceding results in [29, 34, 139, 238] and [212]. Besides, the results allow us to retrieve existence results of VIs for the case of finite-dimensional spaces as a special case.

• Existence of solution to VIs

Throughout this section, we denote by $\mathcal{B}_K$ the set of operators $T : K \rightarrow 2^{X^*}$ which are pseudomonotone in the sense of Brézis, and by $\mathcal{K}_K$ the set of operators $\Phi : K \rightarrow 2^{X^*}$ which are pseudomonotone in the sense of Karamardian.

Recall that a single-valued operator $A : K \rightarrow X^*$ belongs to the class $\mathcal{B}_K$ if, for any $u \in K$ and every net $u_i \rightarrow u$ such that $\limsup \langle Au_i, u_i - u \rangle \leq 0$ implies

$$\langle Au, u - w \rangle \leq \liminf \langle Au_i, u_i - w \rangle, \quad \text{for all } w \in K.$$

A is said to be of class $(S)_+$ if, for any net $u_i \rightarrow u$ and $\limsup \langle Au_i, u_i - u \rangle \leq 0$ imply $u_i \rightarrow u$ in the strong topology of X .

A is called *continuous on finite-dimensional subspaces* if, for any finite-dimensional subspace L of X , the map $A : X \cap L \rightarrow X^*$ is continuous. We say that A is *demicontinuous* if $u_i \rightarrow u$ in the strong topology implies $Au_i \rightarrow Au$. It is clear that, if A is demicontinuous, then it is continuous on finite-dimensional subspaces of X . We say that A is *hemicontinuous* if the real function

$$t \mapsto \langle A((1-t)u + tv), w \rangle$$

is continuous on $[0, 1]$ for all $u, v, w \in K$. If the set $A(K)$ is bounded, then A is called *bounded on K* .

Let us give an example showing that there exists an operator $A \in \mathcal{B}_K \setminus \mathcal{K}_K$.

Example 1.3.12. Consider $\text{VI}(K, A)$, where $K \subset \mathbb{R}^2$ is defined by

$$K = \{(x, y) \in \mathbb{R}^2 : x + y \leq 1, x \geq -1, y \geq -1\}$$

and A is defined by $A(x, y) = (-x^3, y^2)$. Since A is continuous and X is a finite-dimensional space, we have $A \in \mathcal{B}_K$ (see [238, Proposition 27.6]). Thus, $\text{VI}(K, A)$ satisfies all the conditions of Theorem 1.2.18. However, $A \notin \mathcal{K}_K$. In fact, taking $u = (1, 0)$ and $v = (-1, 0)$, we have $\langle Av, u - v \rangle = 2$ and $\langle Au, u - v \rangle = -2$. Hence, A is not pseudomonotone in the sense of Karamardian.

The class of pseudomonotone operators in the sense of Brézis is rather large. For example, maximal monotone operators are of class $\mathcal{B}_K$ (see [192, Proposition 8]), monotone and hemicontinuous operators are of class $\mathcal{B}_K$ (see [238, Proposition 27.6]).

The following example shows that there exists a pseudomonotone operator in the sense of Brézis but it is not monotone.

Example 1.3.13. Let G be a bounded open set in $\mathbb{R}^n$ with $n = 1, 2, 3$ and let $X = W_2^{0,1}(G)$ be a closed subspace of the Sobolev space $W_2^1(G)$.

Let $A_n : X \rightarrow X^*$ be an operator defined by

$$\langle A_n u, v \rangle = \alpha \int_G \sum_{i=1}^n (D_i u)(\sin u) v \, dx,$$

where $\alpha < 0$ and $D_i u(x) := \frac{\partial u}{\partial x_i}(x)$. By Proposition 27.11 in [238], A_n is pseudomonotone. However, A_n is not monotone in general. In fact, for $n = 1$ and $G = (0, 1)$, we have

$$\langle A_1 u, v \rangle = \alpha \int_0^1 u \sin u \, v \, dx.$$

Take two functions $u_1(x) \equiv 0$ and $u_2(x) = -x^3 + x^2$. We see that $u_1, u_2 \in C^\infty(0, 1)$ and $u_i(0) = u_i(1) = 0$, $i = 1, 2$. Hence, $u_1, u_2 \in X$. We now have

$$\begin{aligned} \langle A_1 u_1 - A_1 u_2, u_1 - u_2 \rangle &= \langle A_1 u_2, u_2 \rangle \\ &= \alpha \int_0^1 (-3x^2 + 2x)(-x^3 + x^2) \sin(-x^3 + x^2) \, dx. \end{aligned}$$

Since $0 \leq -x^3 + x^2 \leq \frac{4}{27}$ for all $x \in (0, 1)$, we have $\sin(-x^3 + x^2) \geq 0$. Also, $(-3x^2 + 2x)(-x^3 + x^2) \geq 0$ for all $x \in (0, 1)$. Hence, we obtain

$$\int_0^1 (-3x^2 + 2x)(-x^3 + x^2) \sin(-x^3 + x^2) \, dx > 0.$$

Consequently, $\langle A_1 u_1 - A_1 u_2, u_1 - u_2 \rangle < 0$ and so A_1 is not monotone.

Given a closed convex set K in X and an operator $A : K \rightarrow X^*$, we denote by $\mathcal{L}$ the set of all finite-dimensional subspaces L of X such that $L \cap K \neq \emptyset$. Fixing any $L \in \mathcal{L}$, we put $K_L = K \cap L$ and define a map $A_L : K_L \rightarrow L^*$ by

$$\langle A_L x, y \rangle = \langle A(x), y \rangle, \quad \forall y \in L. \quad (1.3.19)$$

Definition 1.3.14. *The operator A is called compatible with finite dimensional subspaces of X if, for any $L \in \mathcal{L}$, $VI(K_L, A_L)$ has a solution.*

Example 1.3.15. Suppose that K is a compact convex set in X and $A : K \rightarrow X^*$ is continuous on each finite-dimensional subspace of X . Then A is compatible with finite-dimensional subspaces of X . Indeed, fix a subspace $L \in \mathcal{L}$ and consider a mapping $\alpha_L : X^* \rightarrow L^*$ defined by

$$\langle \alpha_L x^*, y \rangle = \langle x^*, y \rangle, \quad \forall y \in L. \quad (1.3.20)$$

We then have $A_L = \alpha_L \circ A$ and so $A_L : K_L \rightarrow L^*$ is continuous. Since K_L is a compact set, by the well-known result on the existence of solutions to variational inequalities in finite-dimensional spaces, $\text{VI}(K_L, A_L)$ has a solution. Consequently, A is compatible with finite-dimensional subspaces of X .

We begin with the first existence theorem for class $\mathcal{B}_K$.

Theorem 1.3.16. *Suppose that X is a topological vector space, K is a nonempty compact convex set in X , and $A : K \rightarrow X^*$ is of the class $\mathcal{B}_K$ and compatible with finite-dimensional subspaces of X . Then, $\text{VI}(K, A)$ has a solution.*

Proof. For each $L \in \mathcal{L}$, we consider $\text{VI}(K_L, A_L)$. By the assumptions, there exists a vector $x_L \in K_L$ such that

$$\langle Ax_L, y - x_L \rangle \geq 0, \quad \forall y \in K_L. \quad (1.3.21)$$

For each $Y \in \mathcal{L}$, we denote by S_Y the set of all $\hat{x} \in K$ such that there exists a subspace $L \supseteq Y$ with the property that $\hat{x} \in K_L$ and

$$\langle A\hat{x}, y - \hat{x} \rangle \geq 0, \quad \forall y \in K_L.$$

From (1.3.21), we see that $S_Y \neq \emptyset$ because $x_Y \in S_Y$. Moreover, the family $\{\overline{S_Y}\}_{Y \in \mathcal{L}}$ has a finite intersection property, where $\overline{S_Y}$ is the closure of S_Y (in the weak topology of X). Indeed, taking any $L_1, L_2, \dots, L_n \in \mathcal{L}$ and putting $M = \text{span}\{L_1, L_2, \dots, L_n\}$, we have $M \in \mathcal{L}$ and

$$S_M \subseteq \bigcap_{i=1}^n S_{L_i}.$$

This implies that

$$\emptyset \neq S_M \subseteq \overline{S_M} \subseteq \overline{\bigcap_{i=1}^n S_{L_i}} \subseteq \bigcap_{i=1}^n \overline{S_{L_i}}.$$

We now see that $\overline{S_Y} \subset K$ and K is a compact set. The finite intersection property of $\{\overline{S_Y}\}_{Y \in \mathcal{L}}$ implies

$$\bigcap_{Y \in \mathcal{L}} \overline{S_Y} \neq \emptyset.$$

Thus, there is a point $x_0 \in K$ such that $x_0 \in \overline{S_Y}$ for all $Y \in \mathcal{L}$.

Fix any $y \in K$ and take $Y \in \mathcal{L}$ such that Y contains y and x_0 . Since $x_0 \in \overline{S}_Y$, there exists a net $x_i \in S_Y$ such that $x_i \rightarrow x_0$. By the definition of S_Y , one has

$$\langle Ax_i, v - x_i \rangle \geq 0, \quad \forall v \in K_Y.$$

Particularly,

$$\langle Ax_i, x_i - x_0 \rangle \leq 0 \quad \text{and} \quad \langle Ax_i, x_i - y \rangle \leq 0.$$

Hence,

$$\limsup \langle Ax_i, x_i - x_0 \rangle \leq 0.$$

By the pseudomonotonicity of A , we obtain

$$\langle Ax_0, x_0 - y \rangle \leq \liminf \langle Ax_i, x_i - y \rangle \leq 0.$$

Thus, we have shown that

$$\langle Ax_0, y - x_0 \rangle \geq 0, \quad \text{for all } y \in K.$$

This means that x_0 is a solution of $\text{VI}(K, A)$. The proof is complete. $\square$

Corollary 1.3.17 (Cf. [29, Theorem 24]). *Suppose that X is a topological vector space, K is a nonempty compact convex set in X , and $A : K \rightarrow X^*$ is of the class $\mathcal{B}_K$. Assume further that, for each $y \in K$, the mapping $x \mapsto \langle Ax, x - y \rangle$ is lower bounded on each compact subset of K . Then, for each $f \in X^*$, there exists $x_0 \in K$ such that*

$$\langle Ax_0 - f, y - x_0 \rangle \geq 0,$$

for all $y \in K$.

Proof. Define a mapping $A_f : K \rightarrow X^*$ by the formula $A_f(x) = Ax - f$. It is easy to show that A_f is of class $\mathcal{B}_K$. By Lemma 27 in [29], A_f is compatible with finite-dimensional subspaces of X . The conclusion follows from Theorem 1.3.16. $\square$

Corollary 1.3.18 (Cf. [29, Corollary 30]). *Suppose that X is a reflexive Banach space, K is a nonempty closed convex set in X with $0 \in K$, and $A : K \rightarrow X^*$ is of the class $\mathcal{B}_K$. Assume further that, for each $y \in K$, the mapping $x \mapsto \langle Ax, x - y \rangle$ is lower bounded on each compact subset of K and*

$$\lim_{\|x\| \rightarrow \infty} \frac{\langle Ax, x \rangle}{\|x\|} = +\infty. \quad (1.3.22)$$

Then, for each $f \in X^*$, there exists $x_0 \in K$ such that

$$\langle Ax_0 - f, y - x_0 \rangle \geq 0, \quad (1.3.23)$$

for all $y \in K$.

Proof. Consider A_f as in the proof of Corollary 1.3.17. We see that A_f is of class $\mathcal{B}_K$. Choose a number $\gamma > \|f\|$ and put $\alpha = \gamma - \|f\|$. Then, it follows from (1.3.22) that, for every $x \in K$ with $\|x\| \geq \gamma$, one has

$$\langle Ax - f, x \rangle \geq \alpha \|x\| \geq \alpha \gamma.$$

Since $0 \in K$, it follows that, if x_0 satisfies (1.3.23), then $\|x_0\| < \gamma$. Put $K_\gamma = K \cap \bar{B}(0, \gamma)$ and consider the problem $\text{GVI}(A_f, K_\gamma)$. According to Theorem 2.1, there exists $x_0 \in K_\gamma$ such that

$$\langle Ax_0 - f, x - x_0 \rangle \geq 0, \quad \text{for all } x \in K_\gamma. \quad (1.3.24)$$

Fix any $u \in K$. Since $\|x_0\| < \gamma$, we have $x_0 + \lambda(u - x_0) \in K_\gamma$ for $\lambda > 0$ and sufficiently small. Hence, (1.3.24) implies

$$\langle Ax_0 - f, \lambda(u - x_0) \rangle \geq 0.$$

Consequently,

$$\langle Ax_0 - f, u - x_0 \rangle \geq 0.$$

Since u is arbitrary in K , x_0 satisfies (1.3.23). $\square$

Corollary 1.3.19 ([139], Theorem 1.4, p. 84). *Let K be a closed convex and bounded set in a reflexive Banach space X and let $A : K \rightarrow X^*$ be an operator which is monotone and continuous on finite-dimensional subspaces of X . Then, the problem $\text{VI}(K, A)$ has a solution.*

Proof. Since A is continuous on finite-dimensional subspaces of X , A is hemicontinuous. According to Proposition 27.6 (a) in [238], we have $A \in \mathcal{B}_K$. The conclusion follows from Theorem 2.1. $\square$

Corollary 1.3.20 ([97], Theorem 2.1). *Suppose that K is a nonempty weakly compact convex subset of the reflexive Banach space X and $A : K \rightarrow X^*$. Assume the following conditions:*

- (i) A is of class $(S)_+$.
- (ii) A is continuous on finite-dimensional subspaces.
- (iii) If $x_n \rightarrow x$, then $\{Ax_n\}$ has a weakly convergent subsequence with limit Ax .

Then, $\text{VI}(K, A)$ has a solution.

Proof. It suffices to show that A is of class $\mathcal{B}_K$. In fact, assume that $u_n \rightharpoonup u$ and

$$\limsup_{n \rightarrow \infty} \langle Au_n, u_n - u \rangle \leq 0.$$

Fix any $v \in K$ and put

$$\alpha = \liminf_{n \rightarrow \infty} \langle Au_n, u_n - v \rangle.$$

Then, there exists a subsequence $\{u_{n_k}\}$ such that

$$\alpha = \lim_{k \rightarrow \infty} \langle Au_{n_k}, u_{n_k} - v \rangle.$$

By (i), we have $u_{n_k} \rightarrow u$ as $k \rightarrow \infty$. By (iii), we can assume that $Au_{n_k} \rightharpoonup Au$ as $k \rightarrow \infty$. Hence, we obtain

$$\langle Au, u - v \rangle = \lim_{k \rightarrow \infty} \langle Au_{n_k}, u_{n_k} - v \rangle = \alpha.$$

This implies that A is pseudomonotone. $\square$

Corollary 1.3.21 (Cf. [67], Theorem 5.1). *Suppose that X is a reflexive Banach space, K is a convex set in X and that $A : K \rightarrow X^*$ is an operator such that the following conditions hold:*

- (i) *For each $y \in K$ the function $x \mapsto \langle Ax, y - x \rangle$ is upper semicontinuous.*
- (ii) *There exist a nonempty compact set $C \subset K$ and a compact convex set $D \subset K$ such that, for each $x \in K \setminus C$, there exists $y \in D$ satisfying $\langle Ax, y - x \rangle < 0$.*

Then, $\text{VI}(K, A)$ has a solution.

Proof. By (i), for each $y \in K$, the function $u \mapsto \langle Au, u - y \rangle$ is lower semicontinuous. Hence, it is lower bounded on each compact subset of K . We now show that A is pseudomonotone. Let $x_i \rightarrow x$ in the weak topology of X and $\limsup \langle Ax_i, x_i - x \rangle \leq 0$. Fixing any $y \in K$ and using the lower semicontinuity of the function $u \mapsto \langle Au, u - y \rangle$, we obtain

$$\langle Ax, x - y \rangle \leq \liminf \langle Ax_i, x_i - y \rangle.$$

Hence, A is of class $\mathcal{B}_K$. Thus, we showed that A satisfies the assumptions of Corollary 1.3.17. We choose $r > 0$ such that C and D are contained in $B(0, r)$, where $B(0, r)$ stands for an open ball in X with

center at 0 and radius $r > 0$. Consider $VI(\bar{B}(0, r), A)$. By Corollary 1.3.18, there exists $x_0 \in \bar{B}(0, r)$ such that

$$\langle Ax_0, y - x_0 \rangle \geq 0, \quad \text{for all } y \in \bar{B}(0, r). \quad (1.3.25)$$

By (ii), $x_0 \in C \subset B(0, r)$. Fixing any $x \in K$, we see that

$$x_t := x_0 + t(x - x_0) \in B(0, r)$$

for $t > 0$ sufficiently small. Substituting x_t into (1.3.25), we have $\langle Ax_0, y - x_0 \rangle \geq 0$. As $x \in K$ is arbitrary, x_0 is a solution of $VI(K, A)$. $\square$

• Existence of solution to GVIs

Let us recall that a multifunction $F : X \rightarrow 2^Z$ from a topological space X to a topological space Z is said to be upper semicontinuous (u.s.c. for brevity) on X , if for any $x_0 \in X$ and open set V in Z such that $F(x_0) \subset V$, there exists a neighborhood U of x_0 satisfying $F(x) \subset V$ for all $x \in U$.

As an auxiliary result, we will use the following well-known existence theorem for GVIs in finite-dimensional spaces. For the convenience of the reader, we provide another proof based on approximate selections below.

Theorem 1.3.22. *Suppose that K is a convex and compact set in $\mathbb{R}^n$ and that $F : K \rightarrow 2^{\mathbb{R}^n}$ is a multifunction which is u.s.c. and compact convex valued. Then, problem $GVI(K, F)$ has a solution.*

Proof. By our assumption and the approximate selection theorem due to Cellina (see [10, Theorem 1, p. 84]), for each $\epsilon > 0$ there exists a continuous map $f_\epsilon : K \rightarrow \mathbb{R}^n$ such that, for all $x \in K$, one has

$$f_\epsilon(x) \in F((x + \epsilon B) \cap K) + \epsilon B, \quad (1.3.26)$$

where B is the unit ball of $\mathbb{R}^n$.

We choose a sequence $\{\epsilon_n\} \rightarrow 0^+$ such that $\{f_{\epsilon_n}\}$ satisfy (1.3.26). We now consider $VI(K, f_{\epsilon_n})$. By the well-known result on solution existence of variational inequalities (see, for instance, [139, Theorem 3.1, p. 12]), there exists $x_n \in K$ such that

$$\langle f_{\epsilon_n}(x_n), y - x_n \rangle \geq 0, \quad \forall y \in K. \quad (1.3.27)$$

From (1.3.26), there exist $y_n \in K$ with $\|y_n\| \leq 1$ and $z_n \in F(x_n + \epsilon_n y_n)$ such that $\|f_{\epsilon_n}(x_n) - z_n\| \leq \epsilon_n$. Since K is a compact set, we can assume

that $x_n \rightarrow x_0 \in K$. Consequently, $x_n + \epsilon_n y_n \rightarrow x_0$. Since $F(C)$ is a compact set whenever C is a compact set and $F(x_n + \epsilon_n y_n) \subset F(K + \bar{B}_{\mathbb{R}^n})$, we can assume further that $z_n \rightarrow z_0$ and so $f_{\epsilon_n}(x_n) \rightarrow z_0$. By the upper semicontinuity of F , we have $z_0 \in F(x_0)$. Fixing any $y \in K$ and letting $n \rightarrow \infty$, we obtain from (1.3.27) that

$$\langle z_0, y - x_0 \rangle \geq 0.$$

This implies that x_0 is a solution of $\text{GVI}(K, F)$. The proof is complete. $\square$

In the sequel, we suppose that X is a topological vector space which is endowed with the weak topology, X^* is the dual of X , which is equipped with the weak* topology. Let K be a nonempty convex set of X and let $T : K \rightarrow 2^{X^*}$ be a set-valued map. We say that T is upper semicontinuous on finite-dimensional subspaces of X if, for every $L \in \mathcal{L}$, the map $T : K \cap L \rightarrow 2^{X^*}$ is u.s.c. For each $L \in \mathcal{L}$, we define the mapping $T_L : L \cap K \rightarrow L^*$ by the formula

$$T_L(x) = \{\alpha_L x^* : x^* \in T(x)\}, \quad (1.3.28)$$

where α_L is defined by (1.3.20).

Definition 1.3.23. *The mapping $T : K \rightarrow 2^{X^*}$ is said to be compatible with finite-dimensional subspaces of X if, for each $L \in \mathcal{L}$, $\text{GVI}(K_L, T_L)$ has a solution.*

Example 1.3.24. *Let K be a compact convex set of X and let $T : K \rightarrow 2^{X^*}$ be u.s.c. on finite-dimensional subspaces and compact convex valued. Then, T is compatible with finite-dimensional subspaces of X . Indeed, for each $L \in \mathcal{L}$, $T_L : K_L \rightarrow L^*$ is u.s.c. and compact convex valued. By Theorem 1.3.22, $\text{GVI}(K_L, T_L)$ has a solution.*

We are now ready to state the existence theorem for GVIs.

Theorem 1.3.25. *Suppose that K is a nonempty compact convex set in the topological vector space X and $T : K \rightarrow 2^{X^*}$ is a mapping of class $\mathcal{B}_K$ such that the following conditions hold:*

- (i) *For each $x \in K$, $T(x)$ is a compact convex set.*
- (ii) *T is compatible with finite-dimensional subspaces of X .*

Then, problem $\text{GVI}(K, T)$ has a solution.

Proof. We shall use the scheme in the proof of Theorem 1.3.16.

For each $L \in \mathcal{L}$, we consider problem $\text{GVI}(K_L, T_L)$. By (ii), $\text{GVI}(K_L, T_L)$ has a solution, that is, there exists $x_L \in K_L$ and $z_L^* \in T_L(x_L)$ such that

$$\langle z_L^*, y - x_L \rangle \geq 0, \quad \forall y \in K_L.$$

Since $z_L^* = \alpha_L x^*$ for some $x^* \in T(x_L)$, we get

$$\langle x^*, y - x_L \rangle \geq 0, \quad \forall y \in K_L.$$

Hence, we obtain

$$\sup_{x^* \in T(x_L)} \langle x^*, y - x_L \rangle \geq 0, \quad \forall y \in K_L. \quad (1.3.29)$$

For each $Y \in \mathcal{L}$, we denote by S_Y the set of all $\hat{x} \in K$ such that there exists a subspace $L \supseteq Y$ with the property that $\hat{x} \in K_L$ and

$$\sup_{x^* \in T(\hat{x})} \langle x^*, y - \hat{x} \rangle \geq 0, \quad \forall y \in K_L.$$

We claim that the family $\{S_Y\}$ has the finite intersection property. Indeed, for each $Y \in \mathcal{F}$, by putting $L = Y$, we have from (1.3.29) that $x_Y \in S_Y$. Hence, S_Y is nonempty. Take subspaces $L_1, L_2, \dots, L_n \in \mathcal{L}$ and put $M = \text{span}\{L_1, L_2, \dots, L_n\}$. Then, we have $M \in \mathcal{L}$ and

$$S_M \subset \bigcap_{i=1}^n S_{L_i}.$$

This implies that

$$\emptyset \neq S_M \subseteq \overline{S_M} \subseteq \bigcap_{i=1}^n \overline{S_{L_i}} \subseteq \bigcap_{i=1}^n \overline{S_{L_i}}.$$

The claim is proved.

Since $\overline{S_Y} \subset K$ and K is weakly compact, the finite intersection property of $\{\overline{S_Y}\}$ implies

$$\bigcap_{Y \in \mathcal{F}} \overline{S_Y} \neq \emptyset.$$

This means that there exists a point $x_0 \in K$ such that $x_0 \in \overline{S_Y}$ for all $Y \in \mathcal{L}$. Fix any $y \in K$ and choose $Y \in \mathcal{L}$ such that Y contains y and x_0 . Since $x_0 \in \overline{S_Y}$, there exists a net $x_i \in S_Y$ such that $x_i \rightarrow x_0$. By the definition of S_Y , we have

$$\sup_{x^* \in T(x_i)} \langle x^*, v - x_i \rangle \geq 0, \quad \forall v \in K_Y.$$

Hence,

$$\inf_{v \in K_Y} \sup_{x^* \in T(x_i)} \langle x^*, v - x_i \rangle \geq 0.$$

By the Sion minimax theorem (see [212]), we obtain

$$\sup_{x^* \in T(x_i)} \inf_{v \in K_Y} \langle x^*, v - x_i \rangle = \inf_{v \in K_Y} \sup_{x^* \in T(x_n)} \langle x^*, v - x_i \rangle \geq 0.$$

Since the extended real function $x^* \mapsto \inf_{v \in K_Y} \langle x^*, v - x_n \rangle$ is u.s.c., there exists $x_i^* \in T(x_i)$ such that

$$\inf_{v \in K_Y} \langle x_i^*, v - x_i \rangle \geq 0.$$

Hence,

$$\langle x_i^*, v - x_n \rangle \geq 0, \quad \forall v \in K_Y.$$

In particular,

$$\langle x_i^*, y - x_i \rangle \geq 0 \quad \text{and} \quad \langle x_i^*, x_0 - x \rangle \geq 0. \quad (1.3.30)$$

From this, we get $\limsup_{n \rightarrow \infty} \langle x_i^*, x_i - x_0 \rangle \leq 0$. By the pseudomonotonicity of T and (1.3.30), there exists $x^* \in T(x_0)$ such that

$$\langle x^*, x_0 - y \rangle \leq \liminf \langle x_i^*, x_i - y \rangle \leq 0.$$

This implies that $\langle x^*, y - x_0 \rangle \geq 0$. Thus, we have shown that

$$\inf_{y \in K} \sup_{x^* \in T(x_0)} \langle x^*, y - x_0 \rangle \geq 0.$$

Using the Sion minimax theorem again, we can prove that there exists $x_0^* \in T(x_0)$ such that

$$\langle x_0^*, y - x_0 \rangle \geq 0, \quad \forall y \in K.$$

The proof is complete. □

Let us present some corollaries of Theorem 1.3.22.

Corollary 1.3.26. *Suppose that K is a nonempty compact convex set in the topological vector space X and that $T : K \rightarrow 2^{X^*}$ is a mapping of the class $\mathcal{B}_K$ such that the following conditions hold:*

- (i) *For each $x \in K$, $T(x)$ is a compact convex set.*
- (ii) *T is u.s.c. on each finite-dimensional subspace of X .*

Then, problem $\text{GVI}(K, T)$ has a solution.

Corollary 1.3.27 ([34, Theorem 15]). *Let X be a reflexive Banach space and let K be a nonempty closed convex set of X with $0 \in K$. Let $T : K \rightarrow 2^{X^*}$ be an operator of class $\mathcal{B}_K$ such that conditions (i) and (ii) of Theorem (1.3.25) are fulfilled. Assume further that T is coercive on K , i.e., that there exists a function $c(r)$ from $\mathbb{R}^+$ into $\mathbb{R}$ with $c(r) \rightarrow +\infty$ as $r \rightarrow \infty$, such that, for each $u \in K$, $w \in Tu$, we have $\langle w, u \rangle \geq c(\|u\|)\|u\|$.*

Then for each $f_0 \in X^*$, there exists $u_0 \in K$ and $w_0 \in Tu_0$ such that

$$\langle w_0 - f_0, u - u_0 \rangle \geq 0, \quad \text{for all } u \in K. \quad (1.3.31)$$

Proof. The proof proceeds analogously to the proof of Corollary 1.3.18. $\square$

Corollary 1.3.28 ([97, Theorem 3.1]). *Let K be a nonempty weakly compact convex subset of a reflexive Banach space X and let $T : K \rightarrow 2^{X^*}$. Assume the following conditions:*

- (i) *$T(x)$ is nonempty, closed, and convex for each $x \in K$.*
- (ii) *T is bounded and upper semicontinuous on finite-dimensional subspaces.*
- (iii) *If x_n converges weakly to x , $w_n \in Tx_n$, and $\limsup \langle w_n, x_n - x \rangle \leq 0$, then $x_n \rightarrow x$ and $\{w_n\}$ has a subsequence converging weakly to some $w \in Tx$.*

Then, $\text{GVI}(K, T)$ has a solution.

Proof. It is sufficient to show that T is of class $\mathcal{B}_K$, i.e., if $\{(u_n, u_n^*)\}$ is a sequence such that $u_n^* \in T(u_n)$, $u_n \rightharpoonup u$ and $\limsup_{n \rightarrow \infty} \langle u_n^*, u_n - u \rangle \leq 0$, then for each $v \in K$, there exists $v^* \in T(u)$ satisfying

$$\langle v^*, u - v \rangle \leq \liminf_{n \rightarrow \infty} \langle u_n^*, u_n - v \rangle.$$

Suppose that T is not of class $\mathcal{B}_K$. Then there exists $v \in K$ such that

$$\inf_{v^* \in T(u)} \langle v^*, u - v \rangle > \liminf_{n \rightarrow \infty} \langle w_n, u_n - v \rangle.$$

By passing to a subsequence, we may assume that

$$\inf_{v^* \in T(u)} \langle v^*, u - v \rangle > \lim_{n \rightarrow \infty} \langle w_n, u_n - v \rangle.$$

By (iii), without loss of generality, we can assume that $w_n \rightharpoonup w \in Tu$ and $u_n \rightarrow u$. Hence, we obtain

$$\inf_{v^* \in T(u)} \langle v^*, u - v \rangle > \lim_{n \rightarrow \infty} \langle w_n, u_n - v \rangle = \langle w, u - v \rangle,$$

a contradiction. $\square$

The following theorem gives sufficient conditions for solution existence of GVIs to the case of unbounded sets.

Theorem 1.3.29. *Suppose that X is a reflexive Banach space, let K be a nonempty closed convex set of X and let $T : K \rightarrow 2^{X^*}$ be an operator of the class $\mathcal{B}_K$, which satisfies the following conditions:*

- (i) *For each $x \in K$, $T(x)$ is a closed convex and bounded set.*
- (ii) *T is upper semicontinuous on finite-dimensional subspaces of X .*
- (iii) *There exists an open ball $\Omega \subset \mathbb{R}^n$ and a vector $\hat{x} \in \Omega \cap K$ such that*

$$\inf_{x^* \in T(x)} \langle x^*, x - \hat{x} \rangle > 0, \quad \forall x \in K \cap \partial\Omega,$$

Then problem $\text{GVI}(K, T)$ has a solution.

Proof. Let us denote by $\mathcal{F}$ the set of all finite-dimensional subspaces L of X such that L contains $\hat{x}$. Putting $K_L = K \cap L$, $\Omega_L = \Omega \cap L$, we see that $\partial\Omega_L = \partial\Omega \cap L$, where $\partial\Omega$ is the boundary of Ω in X . It is clear that, for each $L \in \mathcal{F}$, the triplet (K_L, Ω_L, T_L) satisfies all conditions of Theorem 1.3.10. Hence, there exists $x_L \in K \cap \overline{\Omega} \cap L$ such that

$$\sup_{x^* \in T(x_L)} \langle x^*, y - x_L \rangle \geq 0, \quad \forall y \in K_L. \quad (1.3.32)$$

For each $Y \in \mathcal{F}$, we denote by S_Y the set of all $\hat{x} \in K \cap \overline{\Omega}$ such that there exists a subspace $L \supseteq Y$ with the property that $\hat{x} \in \overline{\Omega} \cap K_L$ and

$$\sup_{x^* \in T(\hat{x})} \langle x^*, y - \hat{x} \rangle \geq 0, \quad \forall y \in K_L.$$

As $\overline{\Omega}$ is a weakly compact set and $\overline{S}_Y \subset \overline{\Omega}$, the finite intersection property of $\{S_Y\}$ implies

$$\bigcap_{Y \in \mathcal{F}} \overline{S}_Y \neq \emptyset.$$

We now use similar arguments as in the proof of Theorem 1.3.25 to obtain the desired conclusion. $\square$

1.3.3 Stability of solutions to parametric generalized variational inequalities in reflexive Banach spaces

In the preceding section, we used the degree theory in order to give a result on the stability of solution to parametric variational inequality

in $\mathbb{R}^n$. In this section we continue to study the stability of solution to generalized variational inequality in reflexive Banach spaces.

Let us assume that E is a reflexive Banach space with the dual E^* , (M, d) and (Λ, d) are metric spaces. Let $F : M \times E \rightarrow 2^{E^*}$ be a set-valued operator and $K : \Lambda \rightarrow 2^E$ be a multifunction with closed convex values. The parametric generalized variational inequality involving the set $K(\lambda)$ and the operator $F(\mu, \cdot)$ is the problem of finding $x = x(\mu, \lambda)$ satisfying the generalized equation

$$0 \in F(\mu, x) + N_{K(\lambda)}(x) \quad (1.3.33)$$

where $(\mu, \lambda) \in M \times \Lambda$ are parameters and $N_{K(\lambda)}(x)$ is the value at x of the normal cone operator associated with the set $K(\lambda)$. We denote problem (1.3.33) briefly by $GVI(F(\mu, \cdot), K(\lambda))$. When $F(\mu, x) = f(\mu, x)$ where $f : M \times E \rightarrow E^*$ is a single-valued operator then (1.3.33) is called a parametric variational inequality and denoted by $VI(f(\mu, \cdot), K(\lambda))$. We will denote by $S(\mu, \lambda)$ the solution set of (1.3.33) corresponding to a parameter pair (μ, λ) . Thus $S : M \times \Lambda \rightarrow 2^E$ is a set-valued map which is called the solution map of (1.3.33). In the sequel, we assume that $x_0 \in S(\mu_0, \lambda_0)$, that is

$$0 \in F(\mu_0, x_0) + N_{K(\lambda_0)}(x_0). \quad (1.3.34)$$

Our main concern is to investigate the behavior of $S(\mu, \lambda)$ when (μ, λ) varies around (μ_0, λ_0) . This problem interested many authors in the last two decades. For the results which have a close connection with the present work, we refer the reader to [61]- [124], [146]- [149], [166]- [202], [229]- [233] and the references given therein. It is known that if $F(\mu, \cdot)$ is single-valued and strongly monotone then the solution map S is single-valued. In this case the solution map has many nice properties. Using the metric projection method, Yen [229] showed that if $F(\mu, \cdot)$ is strong monotone and $K(\cdot)$ has the Aubin property then the solution map is Hölder continuous with respect to parameters. This result was extended later by Domokos [73] to the case of reflexive Banach spaces. Also, under the hypothesis that $F(\mu, \cdot)$ is strongly monotone and K is a graph-convex polyhedral multifunction from R^m to R^m , Robinson and Lu [200] showed that the solution map is Lipschitz continuous when parameters vary. The situation will become complicated when the hypothesis on the strong monotonicity is dropped. In this case the solution map is not single-valued in general. When the set $K(\lambda)$ is defined as perturbed system of finitely many nonlinear inequality or equations subject

to additional fixed polyhedral constraints, F is a single valued and C^1 -function, and the system of constraints obeys a certain property called nondegeneracy, Robinson [197] showed that existence, local uniqueness, and Lipschitz continuity, as well as B-differentiability of the solution, can be predicted from a single affine variational inequality that is easily computable in terms of the data of the original problem (see [197, Theorem 5.3]). In 2004, Robinson [201] has introduced a localized version of the so-called normal maps which help him study the existence and continuity of solutions of variational inequalities via existence and continuity of solutions of operator equations. Robinson established homeomorphisms between the solution set of variational inequalities and the solution set of the normal map. Based on the result and degree-theoretic method, Robinson has showed that if the solution of the unperturbed is an isolated solution and the degree of the normal map at this point is different from zero then the solution map is lower semicontinuous at the parameter corresponding to the solution. Thus the degree-theoretic approach and the metric projection method permit him to draw fairly strong conclusion on the existence and continuity of solution of perturbed problems when the original problem satisfy certain regularity requirements. By introducing a localized version of generalized normal maps and using the degree-theoretic method, Kien and Yao [124] extended the results in [201] to the case of reflexive Banach spaces.

We remark that although the degree theory is a strong tool for studying the solution stability of variational inequalities, it has also some handicaps. Namely, in many cases we have to verify that the degree of maps at some point is different from zero. This is the core of applications of the degree theory. However, the computation of the degree of maps, is not very easy. Besides, the approach of the mentioned papers is to use the metric projection which is suitable for variational inequalities rather than for generalized variational inequalities. Therefore the continuity of the solution map of a parametric generalized variational inequality, is more difficult to obtain than that for parametric variational inequalities.

It is noted that although there have been many papers which study solution stability of parametric variational inequalities, very few paper which focus on a such study for parametric generalized variational inequalities. In a recent paper, Mansour and Aussel [159] have obtained a result on lower semicontinuity of the solution map of problem (1) in the case of finite dimensional spaces (see [159, Theorem 4.6]). They showed that if $F(\mu, \cdot)$ is quasi-monotone, $S(\mu, \lambda)$ is nonempty, the sets $K(\lambda)$ have

nonempty interiors, $K(\lambda_0)$ is a convex polyhedron, $K(\lambda)$ int- Mosco converges to $K(\lambda_0)$ and other additional conditions, then the solution map is lower semicontinuous at (μ_0, λ_0) . In our paper, the conditions which are imposed on F and the sets $K(\lambda)$ as well as techniques is different from that in [159].

Below, we shall show that if the data of the unperturbed problem are good enough then the perturbed problems are stable. Namely, if the operator of the unperturbed problem is of the class $(S)_+$ and strictly monotone, the operators of the perturbed problems are pseudo-monotone and demicontinuous, the set-valued map K has the Aubin property, then the solution map is lower semicontinuous. The obtained results are proved without conditions related to the degree theory and the metric projection. Moreover, the conditions imposed on F and K , are quite explicit under which we can directly verify.

Let us recall some notions related to our problem. Assume that C is a nonempty closed convex set in E and $T : C \rightarrow 2^{E^*}$ is an operator. Then

$$\text{Gr}(T) = \{(x, x^*) \in E \times E^* | x^* \in T(x)\}.$$

We say that T is demicontinuous if T is upper semicontinuous with E^* equipped by its weakly star topology. T is called strictly monotone iff for any $(u_1, u_1^*), (u_2, u_2^*) \in \text{Gr}(T)$ and $u_1 \neq u_2$, one has

$$\langle u_1^* - u_2^*, u_1 - u_2 \rangle > 0.$$

T is said to be of class $(S)_+$ if $u_n \rightharpoonup u$ and $\limsup_{n \rightarrow \infty} \langle u_n^*, u_n - u \rangle \leq 0$ for some $u_n^* \in Tu_n$, imply $u_n \rightarrow u$ strongly. If the set $T(C)$ is bounded then T is called bounded on C . T is said to be pseudo-monotone in the sense of Brezis if for any sequence $u_n \rightharpoonup u$, and if $w_n^* \in Tu_n$ is such that $\limsup_{n \rightarrow \infty} \langle w_n^*, u_n - u \rangle \leq 0$, then for each $v \in C$ there exists $w_v^* \in Tv$ with the property that

$$\langle w_v^*, u - v \rangle \leq \liminf_{n \rightarrow \infty} \langle w_n^*, u_n - v \rangle.$$

As an example, if T is monotone and demicontinuous then T is pseudomonotone (see Proposition 27.6 in [238]). There is some pseudomonotone operators which are not monotone (see Example 2.2 in [126]). We refer the reader to [34] and [238] for more properties and applications of pseudo-monotone operators.

The multifunction $K : \Lambda \rightarrow 2^E$ is said to have the Aubin property of order $\alpha > 0$ at a point $(\lambda_0, x_0) \in \text{Gr}(K)$ if there exist positive constants k, ϵ and β such that

$$K(\lambda') \cap (x_0 + \epsilon_0 \bar{B}_E) \subset K(\lambda) + kd(\lambda', \lambda)^\alpha \bar{B}_E \quad \forall \lambda', \lambda \in B(\lambda_0, \beta_0). \quad (1.3.35)$$

Here $\bar{B}_E$ is the closed unit ball of E and $B(\lambda_0, \beta)$ is an open ball centered at λ_0 with radius $\beta > 0$, in the metric space . If K satisfies property (1.3.35) for $\alpha = 1$ then K is said to have the Aubin property at (λ_0, x_0) .

Recall that a multifunction $P : X \rightarrow 2^Y$ from a topological space X to a topological space Y is said to be lower semicontinuous at $x_0 \in X$ if for any open set V in Y satisfying $P(x_0) \cap V \neq \emptyset$, there exists a neighborhood U of x_0 such that $P(x) \cap V \neq \emptyset$ for all $x \in U$. P is called upper semicontinuous at $x_0 \in X$ if for any open set V in Y satisfying $P(x_0) \subset V$, there exists a neighborhood U of x_0 such that $P(x) \subset V$ for all $x \in U$.

Assume that Y is a normed space. We say that P is Hausdorff upper semicontinuous at x_0 provided that for any $\epsilon > 0$, there exists a neighborhood U of x_0 such that $e(P(x), P(x_0)) < \epsilon$ for all $x \in U$, where

$$e(P(x), P(x_0)) := \sup_{y \in P(x)} d(y, P(x_0))$$

and $d(y, P(x_0))$ is the distance from a point y to the set $P(x_0)$.

Here are some results on the stability of solutions to GVIs.

Theorem 1.3.30. *Let E be a reflexive Banach space with the dual E^* . Assume that there exist a bounded open neighborhood X_0 of x_0 and a neighborhood $M_0 \times \Lambda_0$ of (μ_0, λ_0) such that the operator $F : M_0 \times X_0 \rightarrow 2^{E^*}$ and the multifunction $K : \Lambda_0 \rightarrow 2^E$ satisfy the conditions:*

- (i) $F(\mu_0, \cdot) : X_0 \rightarrow 2^{E^*}$ is of class $(S)_+$, bounded and strictly monotone on $X_0 \cap K(\lambda_0)$;
- (ii) $F(\mu, \cdot) : X_0 \rightarrow 2^{E^*}$ is demicontinuous, pseudo-monotone and closed convex valued for all $\mu \in M_0$;
- (iii) F is Hausdorff upper semicontinuous in μ at μ_0 and uniform in x , that is, for any sequence $\mu_n \rightarrow \mu_0$, one has $e(F(\mu_n, x), F(\mu_0, x)) \rightarrow 0$ for all $x \in X_0$;
- (iv) K has the Aubin property of positive order α at (λ_0, x_0) .

Then there exist a neighborhood U_0 of μ_0 , a neighborhood V_0 of λ_0 and an open bounded neighborhood Q_0 of x_0 such that the following assertions hold:

- (a) *The solution map $\hat{S} : U_0 \times V_0 \rightarrow 2^{X_0}$ defined by $\hat{S}(\mu, \lambda) = S(\mu, \lambda) \cap Q_0$, has nonempty values.*

(b) The solution map $\hat{S}$ is lower semicontinuous at (μ_0, λ_0) .

When $F(\mu, x) = \{f(\mu, x)\}$, where $f : M_0 \times X_0 \rightarrow E^*$ is a single-valued operator, we obtain

Theorem 1.3.31. *Let E be a reflexive Banach space with the dual E^* . Assume that there exist a bounded open neighborhood X_0 of x_0 and a neighborhood $M_0 \times \Lambda_0$ of (μ_0, λ_0) such that the operator $f : M_0 \times X_0 \rightarrow 2^{E^*}$ and the multifunction $K : \Lambda_0 \rightarrow 2^E$ satisfy the conditions:*

- (i) $f(\mu_0, \cdot) : X_0 \rightarrow E^*$ is of class $(S)_+$, bounded and strictly monotone on $X_0 \cap K(\lambda_0)$;
- (ii) $F(\mu, \cdot) : X_0 \rightarrow 2^{E^*}$ is demicontinuous and pseudo-monotone for all $\mu \in M_0$;
- (iii) F is continuous in μ at μ_0 and uniform in x , that is, for any sequence $\mu_n \rightarrow \mu_0$, one has $\sup_{x \in \bar{X}_0} \|f(\mu_n, x) - F(\mu_0, x)\| \rightarrow 0$ as $n \rightarrow \infty$;
- (iv) K has the Aubin property of positive order α at (λ_0, x_0) .

Then there exist a neighborhood U_0 of μ_0 , a neighborhood V_0 of λ_0 and an open bounded neighborhood Q_0 of x_0 such that the following assertions hold:

- (a) The solution map $\hat{S} : U_0 \times V_0 \rightarrow 2^{X_0}$ defined by $\hat{S}(\mu, \lambda) = S(\mu, \lambda) \cap Q_0$, has nonempty values.
- (b) The solution map $\hat{S}$ is lower semicontinuous at (μ_0, λ_0) .

In particular for $E = \mathbb{R}^m$, we have the following result.

Corollary 1.3.32. *Suppose that there exist a bounded open neighborhood X_0 of x_0 and a neighborhood $M_0 \times \Lambda_0$ of (μ_0, λ_0) such that the mapping $f : M_0 \times X_0 \rightarrow \mathbb{R}^m$ and the multifunction $K : \Lambda_0 \rightarrow 2^{\mathbb{R}^m}$ satisfy the conditions:*

- (i) $f(\mu_0, \cdot) : X_0 \rightarrow \mathbb{R}^m$ is strictly monotone on $X_0 \cap K(\lambda_0)$;
- (ii) $f(\cdot, \cdot)$ is continuous
- (iii) K has the Aubin property of positive order α at (λ_0, x_0) .

Then there exist a neighborhood U_0 of μ_0 , a neighborhood V_0 of λ_0 and an open bounded neighborhood Q_0 of x_0 such that the following assertions hold:

- (a) The solution map $\hat{S} : U_0 \times V_0 \rightarrow 2^{X_0}$ defined by $\hat{S}(\mu, \lambda) = S(\mu, \lambda) \cap Q_0$, has nonempty values.
- (b) The solution map $\hat{S}$ is lower semicontinuous at (μ_0, λ_0) .

In order to prove the above theorems, we first establish a result on a

relation between the Aubin property and the so-called Hölder property of multifunctions.

Let E be a normed space, $K : \rightarrow 2^E$ be a multifunction, and $(\lambda_0, x_0) \in \text{Gr}K$ be such that property (1.3.35) holds. For each $\epsilon > 0$, we put

$$K_\epsilon(\lambda) = K(\lambda) \cap (x_0 + \epsilon \bar{B}_E). \quad (1.3.36)$$

Lemma 1.3.33. *Suppose that $K : \rightarrow 2^E$ has convex values and satisfies property (1.3.35). Then for any ϵ in $(0, \epsilon_0]$ and any β with $0 < \beta < \min(\beta_0, (\frac{\epsilon}{4k})^{1/\alpha})$, the multifunction K_ϵ defined by (1.3.36), is Hölder continuous with constant $5k$ on the ball $B(\lambda_0, \beta)$, that is*

$$K_\epsilon(\lambda') \subset K_\epsilon(\lambda) + 5kd(\lambda', \lambda)^\alpha \bar{B}_E \quad (1.3.37)$$

for all $\lambda, \lambda' \in B(\lambda_0, \beta)$.

Proof. We mimic the proof of Lemma 2.3 in [23].

Put $\lambda' = \lambda_0$ in (1.3.35), we see that, for all $\lambda \in B(\lambda_0, \beta)$, there exists a point $x_\lambda \in K(\lambda)$ such that

$$\|x_\lambda - x_0\| \leq kd(\lambda, \lambda_0)^\alpha < \frac{\epsilon}{2}. \quad (1.3.38)$$

This implies that $K_\epsilon(\lambda)$ is nonempty for all $\lambda \in B(\lambda_0, \beta)$. Take any $\lambda', \lambda \in B(\lambda_0, \beta)$ and $x' \in K_\epsilon(\lambda')$. We have to show that there exists a point $x'' \in K_\epsilon(\lambda)$ such that

$$\|x' - x''\| \leq 5kd(\lambda, \lambda_0)^\alpha \quad (1.3.39)$$

which proves (1.3.37).

It follows from (1.3.35) that there exists $x \in K(\lambda)$ such that

$$\|x' - x\| \leq kd(\lambda, \lambda_0)^\alpha \leq 5kd(\lambda, \lambda_0)^\alpha.$$

If $x \in x_0 + \epsilon \bar{B}$, then $x \in K_\epsilon(\lambda)$. By putting $x'' = x$, we obtain (1.3.39).

Let us assume that $x \notin x_0 + \epsilon \bar{B}$. This means that

$$r := \|x - x_0\| > \epsilon.$$

Choose $x_\lambda \in K(\lambda)$ such that (1.3.38) holds. By the convexity of $K(\lambda)$, the segment $[x, x_\lambda]$ is contained in the $K(\lambda)$. Take a point $x'' \in [x, x_\lambda]$ belongs to $K(\lambda)$ such that $\|x'' - x_0\| = \epsilon$. Hence $x'' \in K_\epsilon(\lambda)$. Note that such a point x'' always exists. Namely, $x'' = x_\lambda + s_0(x - x_\lambda)$, where

$$s_0 = \max\{s : s \in [0, 1], x_\lambda + s(x - x_\lambda) \in x_0 + \epsilon \bar{B}_E\}$$

It is easy to show that there exists such a number s_0 and $s_0 \in (0, 1)$.

Put $\rho = \|x - x''\|$, $d = r - \epsilon$ and $t = 1 - s_0$. Then we have $x'' = (1 - t)x + tx_\lambda$ with $t \in (0, 1)$ and

$$\epsilon = \|x'' - x_0\| = \|(1 - t)(x - x_0) + t(x_\lambda - x_0)\| \leq (1 - t)r + t\|x_\lambda - x_0\|.$$

This implies that

$$t(r - \|x_\lambda - x_0\|) \leq r - \epsilon.$$

Hence

$$t \leq \frac{r - \epsilon}{r - \epsilon/2}.$$

From this relation we have

$$\begin{aligned} \rho = \|x'' - x\| &= t\|x - x_\lambda\| \leq t(\|x - x_0\| + \|x_\lambda - x_0\|) \\ &\leq t(r + \epsilon/2) \leq (r + \epsilon/2) \frac{r - \epsilon}{r - \epsilon/2}. \end{aligned}$$

It follows that $\|x'' - x\| \leq 4(r - \epsilon) = 4d$. Since $d \leq \|x - x'\|$, we have $\|x'' - x\| \leq 4\|x - x'\|$. Consequently,

$$\|x' - x''\| \leq \|x'' - x\| + \|x' - x\| \leq 5\|x' - x\| \leq 5kd(\lambda', \lambda)^\alpha.$$

We obtaine (1.3.39). The proof is complete. $\square$

Proof of Theorem 1.3.30.

(a). By (iv), there exist positive constants k, ϵ_0 and β_0 such that

$$K(\lambda') \cap (x_0 + \epsilon \bar{B}_E) \subset K(\lambda) + kd(\lambda', \lambda)^\alpha \bar{B}_E \quad \forall \lambda', \lambda \in B(\lambda_0, \beta_0). \quad (1.3.40)$$

Choose positive constants s and δ such that

$$x_0 + s\bar{B}_E \subset (x_0 + \epsilon \bar{B}_E) \cap X_0, \quad kd(\lambda, \lambda_0)^\alpha < s$$

for all $\lambda \in B(\lambda_0, \delta) \subset B(\lambda_0, \beta_0)$. Hence (1.3.40) implies that

$$K(\lambda') \cap (x_0 + s\bar{B}_E) \subset K(\lambda) + kd(\lambda', \lambda)^\alpha \bar{B}_E \quad \forall \lambda', \lambda \in B(\lambda_0, \delta). \quad (1.3.41)$$

Choose a number β such that $0 < \beta < \min(\delta, (\frac{s}{4k})^{1/\alpha})$. By Lemma 1.3.33, we have

$$K(\lambda') \cap (x_0 + s\bar{B}_E) \subset K(\lambda) \cap (x_0 + s\bar{B}_E) + 5kd(\lambda', \lambda)^\alpha \bar{B}_E \quad \forall \lambda', \lambda \in B(\lambda_0, \beta). \quad (1.3.42)$$

Thus for all $\lambda, \lambda' \in B(\lambda_0, \beta)$, (1.3.41) and (1.3.42) are fulfilled.

Putting $\lambda' = \lambda_0$ into (1.3.41), we see that for each $\lambda \in B(\lambda_0, \beta)$, there exists $z_\lambda \in K(\lambda)$ such that $\|z_\lambda - x_0\| \leq kd(\lambda, \lambda_0)^\alpha < s$. Consequently, $K(\lambda) \cap B_E(x_0, s) \neq \emptyset$ for all $\lambda \in B(\lambda_0, \beta)$.

For each $(\mu, \lambda) \in M_0 \times B(\lambda_0, \beta)$, we consider the generalized equation

$$0 \in F(\mu, x) + N_{K(\lambda) \cap B_E(x_0, s)}(x). \quad (1.3.43)$$

To continue, we prove the following lemma.

Lemma 1.3.34. *There exist a neighborhood $U_0 \times V_0$ of (μ_0, λ_0) such that*

$$0 \notin F(\mu, x) + N_{K(\lambda) \cap B_E(x_0, s)}(\partial B_E(x_0, s)) \quad (1.3.44)$$

for all $(\mu, \lambda) \in U_0 \times V_0$, where $\partial B_E(x_0, s)$ is the boundary of $B_E(x_0, s)$.

Proof. Assume that the assertion of the lemma is false. Then we can find sequences $\mu_n \rightarrow \mu_0$, $\lambda_n \rightarrow \lambda_0$ and $x_n \in \partial B(x_0, s)$ such that

$$0 \notin F(\mu_n, x_n) + N_{K(\lambda) \cap B_E(x_0, s)}(x_n). \quad (1.3.45)$$

Since the sequence $\{x_n\}$ is bounded, we can assume that $x_n \rightharpoonup \bar{x}$ for some $\bar{x} \in \partial B_E(x_0, s)$. From (1.3.45), we have $x_n \in K(\lambda) \cap \partial B_E(x_0, s)$ and there exists $x_n^* \in F(\mu_n, x_n)$ such that

$$\langle x_n^*, z - x_n \rangle \geq 0 \quad \forall z \in K(\lambda) \cap \bar{B}_E(x_0, s). \quad (1.3.46)$$

For n sufficiently large, putting $\lambda' = \lambda_n$ and $\lambda = \lambda_0$ in (1.3.42), we see that there exists a point $y_n \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$ such that

$$\|x_n - y_n\| \leq 5kd(\lambda_n, \lambda_0)^\alpha. \quad (1.3.47)$$

Since $K(\lambda_0) \cap \bar{B}_E(x_0, s)$ is a weakly compact set, we can assume that $y_n \rightharpoonup \bar{y} \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$. Fixing any $x^* \in E^*$ and using (1.3.47), we have

$$\begin{aligned} |\langle x^*, x_n \rangle - \langle x^*, \bar{y} \rangle| &= |\langle x^*, x_n \rangle - \langle x^*, y_n \rangle + \langle x^*, y_n \rangle - \langle x^*, \bar{y} \rangle| \\ &\leq \|x^*\| \|x_n - y_n\| + |\langle x^*, y_n - \bar{y} \rangle| \rightarrow 0. \end{aligned}$$

This implies that $x_n \rightharpoonup \bar{y}$ and so $\bar{x} = \bar{y} \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$.

Putting $\lambda' = \lambda_0$ and $\lambda = \lambda_n$ in (1.3.42), we see that for each n , there exists a point $z_n \in K(\lambda_n) \cap \bar{B}_E(x_0, s)$ such that $\|z_n - \bar{x}\| \leq 5kd(\lambda_0, \lambda_n)^\alpha$. Consequently, $z_n \rightarrow \bar{x}$ strongly when $n \rightarrow \infty$.

Substituting $z = z_n$ into (1.3.46), yields

$$\langle x_n^*, z_n - \bar{x} \rangle \geq \langle x_n^*, x_n - \bar{x} \rangle. \quad (1.3.48)$$

Since $F(\mu_0, x_n)$ is a closed convex set in E^* , there exists $u_n^* \in F(\mu_0, x_n)$ such that

$$d(x_n^*, F(\mu_0, x_n)) = \|x_n^* - u_n^*\|.$$

By (iii), we have $\|x_n^* - u_n^*\| \rightarrow 0$. Combining this with the boundedness of $F(\mu_0, \cdot)$, we see that $\{x_n^*\}$ is bounded. Hence from (1.3.48) we obtain

$$\limsup_{n \rightarrow \infty} \langle x_n^*, x_n - \bar{x} \rangle \leq 0.$$

This implies that

$$\limsup_{n \rightarrow \infty} \langle u_n^*, x_n - \bar{x} \rangle = \lim_{n \rightarrow \infty} \langle u_n^* - x_n^*, x_n - \bar{x} \rangle + \limsup_{n \rightarrow \infty} \langle x_n^*, x_n - \bar{x} \rangle \leq 0.$$

Since $F(\mu_0, \cdot)$ is of class $(S)_+$, we conclude that $x_n \rightarrow \bar{x}$ and so $\bar{x} \in \partial B_E(x_0, s) \cap K(\lambda_0)$.

Take any $x \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$. By (1.3.42), for each n sufficiently large, there exists a point $u_n \in K(\lambda_n) \cap B_E(x_0, s)$ such that $u_n \rightarrow x$ as $n \rightarrow \infty$. Substituting $z = u_n$ into (1.3.46), we get

$$\langle x_n^*, u_n - x_n \rangle \geq 0. \quad (1.3.49)$$

Since $F(\mu_0, \cdot)$ is bounded, we can assume that $u_n^* \rightharpoonup u^*$. On the other hand, $F(\mu_0, \cdot)$ is demicontinuous at x and $x_n \rightarrow \bar{x}$, we get $u^* \in F(\mu_0, \bar{x})$. Rewrite

$$\langle x_n^*, u_n - x_n \rangle = \langle x_n^* - u_n^*, u_n - x_n \rangle + \langle u_n^*, u_n - x_n \rangle$$

Letting $n \rightarrow \infty$ and taking into account of (1.3.49) yields

$$\langle u^*, x - \bar{x} \rangle \geq 0 \quad \forall x \in K(\lambda_0) \cap \bar{B}_E(x_0, s).$$

In particular, for $x = x_0$ we get $\langle u^*, x_0 - \bar{x} \rangle \geq 0$. Since $x_0 \in S(\mu_0, \lambda_0)$, there exists $u_0^* \in F(\mu_0, x_0)$ such that

$$\langle u_0^*, x - x_0 \rangle \geq 0 \quad \forall x \in K(\lambda_0).$$

Particularly, $\langle u_0^*, \bar{x} - x_0 \rangle \geq 0$. By the strict monotonicity of $F(\mu_0, \cdot)$ on $X_0 \cap K(\lambda_0)$ and $x_0 \neq \bar{x}$, one has,

$$\langle u^*, \bar{x} - x_0 \rangle > \langle u_0^*, \bar{x} - x_0 \rangle \geq 0$$

which is impossible. The lemma is proved.

We now return to the proof of the theorem. We now choose $U_0 \subset M_0$ and $V_0 \subset B(\lambda_0, \beta)$ such that (1.3.44) in Lemma 1.3.34 is valid. Put $Q_0 = B_E(x_0, s)$ and $\hat{S}(\mu, \lambda) = S(\mu, \lambda) \cap Q_0$ for $(\mu, \lambda) \in U_0 \times V_0$. We will show that U_0 , V_0 and Q_0 satisfy the conclusion of the theorem. Fixing any $(\mu, \lambda) \in U_0 \times V_0$, we consider the generalized equation

$$0 \in F(\mu, x) + N_{K(\lambda) \cap B_E(x_0, s)}(x). \quad (1.3.50)$$

By (ii), $F(\mu, \cdot)$ is pseudomonotone, closed convex valued and upper semi-continuous on finite subspaces of E . Moreover, by the boundedness of $F(\mu_0, \cdot)$ and condition (iii), $F(\mu, \cdot)$ has bounded values. Besides, $K(\lambda) \cap \bar{B}_E(x_0, s)$ is a closed convex and bounded set in E . Thus all conditions of Corollary 1.3.26 are fulfilled. According to Corollary 1.3.26, the generalized equation (1.3.50) has a solution $\hat{x} = \hat{x}(\mu, \lambda) \in K(\lambda) \cap \bar{B}_E(x_0, s)$. By (1.3.44) and the choice of U_0 and V_0 , $\hat{x} \in \text{int} \bar{B}_E(x_0, s)$. Hence

$$N_{K(\lambda) \cap \bar{B}_E(x_0, s)}(x) = N_{K(\lambda)}(x).$$

Consequently, $\hat{x}$ also satisfies the generalized equation

$$0 \in F(\mu, x) + N_{K(\lambda)}(x).$$

This implies that $\hat{x} \in \hat{S}(\mu, \lambda) \cap Q_0$. The assertion (a) of the theorem is proved.

(b) Assume that G is an open set such that $\hat{S}(\mu_0, \lambda_0) \cap G \neq \emptyset$. By the uniqueness of x_0 in $\hat{S}(\mu_0, \lambda_0)$, we have $x_0 \in G$. Choose $\bar{s} \in (0, s)$ such that $B_E(x_0, s) \subset B_E(x_0, s) \cap G$ and (1.3.41) holds. Namely, there exists $\bar{\delta} > 0$ such that $kd(\lambda, \lambda_0)^\alpha < s$ and

$$K(\lambda') \cap (x_0 + \bar{s} \bar{B}_E) \subset K(\lambda) + kd(\lambda', \lambda)^\alpha \bar{B}_E \quad \forall \lambda', \lambda \in B(\lambda_0, \bar{\delta}) \subset B(\lambda_0, \beta_0). \quad (1.3.51)$$

As in the beginning of the proof of assertion (a) we choose a number $\bar{\beta}$ such that $0 < \bar{\beta} < \min(\bar{\delta}, (\frac{\bar{s}}{4k})^{1/\alpha})$. By Lemma 1.3.33, we have

$$K(\lambda') \cap (x_0 + \bar{s} \bar{B}_E) \subset K(\lambda) \cap (x_0 + \bar{s} \bar{B}_E) + 5kd(\lambda', \lambda)^\alpha \bar{B}_E \quad \forall \lambda', \lambda \in B(\lambda_0, \bar{\beta}). \quad (1.3.52)$$

Moreover, for all $\lambda', \lambda \in B(\lambda_0, \bar{\beta})$, (1.3.51) is fulfilled.

We now use Lemma (1.3.34) to show that there exists a neighborhood $\hat{U}_0 \subset U_0$ of μ_0 , a neighborhood $\hat{V}_0 \subset V_0$ of λ_0 such that for all $(\mu, \lambda) \in \hat{U}_0 \times \hat{V}_0$, one has

$$0 \notin F(\mu, x) + N_{K(\lambda) \cap \bar{B}_E(x_0, \bar{s})}(\partial B_E(x_0, \bar{s})) \quad (1.3.53)$$

Using some arguments similar to those in the proof of part (a), we see that the generalized equation

$$0 \in F(\mu, x) + N_{K(\lambda) \cap B_E(x_0, \bar{s})}(x)$$

has a solution $\bar{u} = \bar{u}(\mu, \lambda)$. By (1.3.53), $\bar{u} \in \text{int} \bar{B}_E(x_0, \bar{s})$. This implies that

$$N_{K(\lambda) \cap B_E(x_0, \bar{s})}(\bar{u}) = N_{K(\lambda)}(\bar{u})$$

and so

$$0 \in F(\mu, x) + N_{K(\lambda)}(\bar{u}).$$

Hence $\bar{u} \in \hat{S}(\mu, \lambda) \cap B_E(x_0, \bar{s}) \subset \hat{S}(\mu, \lambda) \cap G$. In other words, $S(\mu, \lambda) \cap G \neq \emptyset$ for all $(\mu, \lambda) \in \hat{U}_0 \times \hat{V}_0$. This means that the solution map $\hat{S} : U_0 \times V_0 \rightarrow 2^{X_0}$ is lower semicontinuous at (μ_0, λ_0) . The proof of the theorem is complete. $\square$

Proof of Corollary 1.3 As in the proof of Theorem 1.3.30, we can choose positive constants s and β such that (1.3.41) and (1.3.42) hold. In order to show that Lemma 3.1 can be applied in the case $F(\mu, \cdot) = f(\mu, \cdot)$, we assume that the conclusion of the lemma is false. Then we can find sequences $\mu_n \rightarrow \mu_0, \lambda_n \rightarrow \lambda_0$ and $\{x_n\} \subset \partial B_E(x_0, s) \cap K(\lambda_n)$ such that

$$\langle f(\mu_n, x_n), z - x_n \rangle \geq 0 \quad \forall z \in K(\lambda_n) \cap \bar{B}_E(x_0, s). \quad (1.3.54)$$

Since $\partial B_E(x_0, s)$ is a compact set, we can assume that $x_n \rightarrow \bar{x}$. Substituting $\lambda' = \lambda_n, \lambda = \lambda_0$ into (1.3.42) we see that for each n , there exists $y_n \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$ such that

$$\|y_n - x_n\| \leq 5kd(\lambda_n, \lambda_0)^\alpha.$$

Since $K(\lambda_0) \cap \bar{B}_E(x_0, s)$ is compact, without loss of generality, we may assume that $y_n \rightarrow y_0 \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$. From the above, we have $x_n \rightarrow y_0$. Hence $\bar{x} = y_0 \in K(\lambda_0) \cap \bar{B}_E(x_0, s)$. Putting $\lambda' = \lambda_0, \lambda = \lambda_n$ in (1.3.42), we see that for each n there exists a point $z_n \in K(\lambda_n) \cap \bar{B}_E(x_0, s)$ such that $z_n \rightarrow x_0$. Putting $z = z_n$ in (1.3.54) and letting $n \rightarrow \infty$, we obtain $\langle f(\mu_0, x), x_0 - \bar{x} \rangle \geq 0$. Since $f(\mu_0, \cdot)$ is strictly monotone and $x_0 \neq \bar{x}$, we have

$$\langle f(\mu_0, x_0), x_0 - \bar{x} \rangle > \langle f(\mu_0, \bar{x}), x_0 - \bar{x} \rangle \geq 0$$

This contradicts the fact that x_0 is a solution of $VI(f(\mu_0, \cdot), K(\lambda_0))$. Thus the Lemma 1.3.34 is applicable. We now can choose neighborhoods $U_0 \subset M_0$ of μ_0 and $V_0 \subset B(\lambda_0, \beta)$ of λ_0 such that the conclusion of Lemma 1.3.34 is valid. Put $Q_0 = B_E(x_0, s)$. For each $(\mu, \lambda) \in U_0 \times V_0$ we consider the generalized equation

$$0 \in f(\mu, x) + N_{K(\lambda) \cap \bar{B}_E(x_0, s)}(x). \quad (1.3.55)$$

Since $f(\mu, \cdot)$ is continuous and $K(\lambda) \cap \bar{B}_E(x_0, s)$ is compact, (1.3.55) has a solution $\hat{x} = \hat{x}(\mu, \lambda) \in K(\lambda) \cap \bar{B}_E(x_0, s)$. By Lemma 1.3.34, $\hat{x} \in \text{int} \bar{B}(x_0, s)$. Hence

$$N_{K(\lambda) \cap B_E(x_0, s)}(\hat{x}) = N_{K(\lambda)}(\hat{x})$$

and so

$$0 \in f(\mu, \hat{x}) + N_{K(\lambda)}(\hat{x}).$$

This implies that $\hat{x} \in \hat{S}(\mu, \lambda) \cap Q_0$. The assertion (a) is proved. By using the arguments described in the proof of assertion (b) of Theorem 1.3.30 we show that the map $\hat{S}$ is lower semicontinuous at (μ_0, λ_0) . The proof of the corollary is complete. $\square$

Example 1.3.35. To provide some perspective what we have obtained up to now, we consider an illustrative example in the case of finite dimensional spaces.

Let $(\mu_0, \lambda_0) = (-1, 1/8)$ and $M_0 \times \Lambda_0 \subset \mathbb{R}^2$ be a neighborhood of (μ_0, λ_0) and $X_0 = \mathbb{R}^2$. Let $f : M_0 \times X_0 \rightarrow \mathbb{R}^2$ be a mapping which is defined by

$$f(\mu, x) = (x_1, \mu x_2 + x_2^2), \quad x = (x_1, x_2)$$

and $K : \Lambda_0 \rightarrow 2^{\mathbb{R}^2}$ defined by

$$K(\lambda) = \{(x_1, x_2) : x_2 \geq 1/2, x_1 + x_2 = 2\lambda\}. \quad (1.3.56)$$

Then the following assertions hold:

- (a) conditions (i) – (iii) of Corollary 1.3 are satisfied;
- (b) $u_0 = (-1/4, 1/2)$ is a unique solution of $VI(f(\mu_0, \cdot), K(\lambda_0))$;
- (c) the solution map $\hat{S}(\cdot)$ is lower semicontinuous at (μ_0, λ_0) .

In fact, we have $f(\mu_0, x) = (x_1, -x_2 + x_2^2)$ and

$$K(\lambda_0) = \{(x_1, x_2) : x_2 \geq 1/2, x_1 + x_2 = 1/4\}.$$

For any $u = (u_1, u_2), v = (v_1, v_2) \in K(\lambda_0)$, we see that $u_2 + v_2 \geq 1$. Hence

$$\langle f(\mu_0, u) - f(\mu_0, v), u - v \rangle = (u_1 - v_1)2 + (u_2 - v_2)^2(u_2 + v_2 - 1) > 0$$

whenever $u \neq v$. This implies that $f(\mu_0, \cdot)$ is strictly monotone and so condition (i) of Corollary 1.3 is satisfied.

It is clear that $f(\cdot, \cdot)$ is continuous. On the other hand, for each $\lambda \in \Lambda_0$, $K(\lambda)$ is a closed convex set. By [158], the map $K(\cdot)$ is Lipschitz continuous. Hence condition (iii) of Corollary 1.3 is valid. Thus all conditions of Corollary 1.3 are fulfilled. To verify (b) and (c) we note that $u \in K(\lambda)$ is a solution of $VI(f(\mu, \cdot), K(\lambda))$ if and only if u satisfies the equation

$$u = P_{K(\lambda)}(u - \rho f(\mu, u)), \quad \rho > 0.$$

Let $x = (x_1, x_2)$ be any point in $\mathbb{R}^2$. By a simple computation, we obtain

$$P_{K(\lambda)}(x) = \left(\lambda + \frac{x_1 - x_2}{2}, \lambda + \frac{x_2 - x_1}{2} \right).$$

It follows that

$$\begin{aligned} & P_{K(\lambda)}(x - \rho f(\mu, x)) \\ &= 1/2(2\lambda + (1 - \rho)x_1 - (1 - \mu\rho)x_2 + \rho x_2^2, 2\lambda - (1 - \rho)x_1 \\ &+ (1 - \mu\rho)x_2 - \rho x_2^2) \end{aligned}$$

Hence

$$(x_1, x_2) = P_{K(\lambda)}(x - \rho f(\mu, x))$$

if and only if

$$\begin{cases} 2\lambda + (1 - \rho)x_1 - (1 - \mu\rho)x_2 + \rho x_2^2 = 2x_1 \\ 2\lambda - (1 - \rho)x_1 + (1 - \mu\rho)x_2 - \rho x_2^2 = 2x_2. \end{cases}$$

This equivalent to

$$\begin{cases} x_2^2 + (1 + \mu)x_2 - 2\lambda = 0 \\ x_1 + x_2 = 2\lambda \end{cases} \quad (1.3.57)$$

The first equation in (1.3.57) has $\Delta = (\mu + 1)^2 + 8\lambda$. When $(\mu_0, \lambda_0) = (-1, 1/8)$, the system has a unique solution $u_0 = (-1/4, 1/2) \in K(1/8)$. It is a unique solution of $VI(f(\mu_0, \cdot), K(\lambda_0))$. When $\lambda > 0$ the first equation of (1.3.57) gives two solutions

$$\theta_1 = \frac{-(1 + \mu) + \sqrt{\Delta}}{2} \quad \text{and} \quad \theta_2 = \frac{-(1 + \mu) - \sqrt{\Delta}}{2}.$$

Hence $VI(f(\mu, \cdot), K(\lambda))$ has two solutions

$$u(\mu, \lambda) = (2\lambda - \theta_1, \theta_1) \quad \text{and} \quad \hat{u}(\mu, \lambda) = (2\lambda - \theta_2, \theta_2).$$

Thus $S(\mu, \lambda) = \{u(\mu, \lambda), \hat{u}(\mu, \lambda)\}$. It is easy to check that

$$\lim_{(\mu, \lambda) \rightarrow (\mu_0, \lambda_0)} u(\mu, \lambda) = u_0.$$

Hence there exist neighborhoods $U_0 \times V_0 \subset M_0 \times \Lambda_0$ of (μ_0, λ_0) and Q_0 of u_0 such that $\hat{S}$ is lower semicontinuous at (μ_0, λ_0) .

1.3.4 The Galerkin method and regularization for variational inequalities

So far, there have been a lot of papers on the existence of solutions as well as numerical method for variational inequalities, we refer the readers to [25], [93], [98], [120]- [156], [224], [226], [228], [233] and many references given therein. Particularly, in [93] and [224] the authors used the finite element method (FEM, for short), which is an instance of the Galerkin method, to compute solutions of elliptic variational inequalities in Hilbert spaces.

It is known that the Galerkin method and FEM are important in computing solutions of partial differential equations as well as obstacle problems in Sobolev spaces. The Galerkin method for operator equations in separable and reflexive Banach spaces have been studied very thoroughly in [238]. However, it seems that, there is no result on the Galerkin method for variational inequalities in separable and reflexive Banach spaces have been published. In this subsection we will present some new results on the Galerkin method for variational inequalities with monotone operators and pseudomonotone operators in separable and reflexive Banach spaces. Besides, the regularized method for variational inequalities with monotone operators is also investigated.

Throughout the subsection, we assume that $(X, \|\cdot\|)$ is a reflexive and separable Banach space with the dual X^* , K is a nonempty closed and convex set in X and $A : K \rightarrow X^*$ is an operator. As in Subsection 1.3.2, we denote by $\mathcal{B}_K$ the class of operators which are pseudo-monotone in the sense of Brésiz and by $\mathcal{K}_K$ the class of operators which are pseudo-monotone in the sense of Karamardian (see [114]).

The operator A is said to be uniformly monotone if

$$\langle Au - Av, u - v \rangle \geq \omega(\|u - v\|)\|u - v\| \quad \forall u, v \in K,$$

where $\omega : [0, +\infty) \rightarrow [0, +\infty)$ is an increasing function satisfying $\omega(0) = 0$, $\omega(t) > 0$ as $t > 0$ and $\omega(t) \rightarrow +\infty$ as $t \rightarrow +\infty$. It is clear that if A is uniformly monotone on K , then it satisfies condition $(S)_+$, that is, if for any sequence $(x_n) \subset K$, $x_n \rightharpoonup x$ and

$$\limsup_{n \rightarrow \infty} \langle Ax_n - Ax, x_n - x \rangle \leq 0$$

then $\|x_n - x\| \rightarrow 0$.

A is said to be *hemicontinuous* if for any $u, v, w \in X$, the mapping

$$[0, 1] \ni t \mapsto \langle A(u + tv), w \rangle$$

is continuous. A is said to be *demicontinuous* if for any sequence $u_n \rightarrow u$ strongly, then $Au_n \rightharpoonup Au$ in the weak* topology $\sigma(X^*, X)$.

It is noted that if A is monotone, then $A \in \mathcal{K}_K$. If A is hemicontinuous and monotone, then $A \in \mathcal{B}_K$ (see [238, Proposition 27.6]). Therefore, the class of hemicontinuous and monotone operators is included in both classes $\mathcal{K}_K$ and $\mathcal{B}_K$. However, these classes are different. In this paper, we only concentrate on the class of pseudo-monotone operators in the sense of Brézis.

We say that A satisfies a coercive condition if there is a point $w_0 \in K$ such that

$$\frac{\langle Au - Aw_0, u - w_0 \rangle}{\|u - w_0\|} \rightarrow +\infty \quad \text{as} \quad \|u\| \rightarrow +\infty. \quad (1.3.58)$$

The operator A is said to be bounded, if it maps a bounded set into a bounded set. We say A is locally bounded, if for any $u \in K$, there exists $r > 0$ such that $A[\overline{B}_X(u, r)]$ is a bounded set.

Given a functional $f \in X^*$, we consider the following variational inequality:

$$\text{Find } u \in K : \langle Au - f, y - u \rangle \geq 0 \quad \forall y \in K. \quad (1.3.59)$$

We will denote by $VI(F, K)$ the above variational inequality, where $Fu := Au - f$. It is clear that when $K = X$, then $VI(F, K)$ is equivalent to the following operator equation:

$$\text{Find } u \in X : \quad Au = f. \quad (1.3.60)$$

Since X is assumed to be a separable space, K is also a separable metric subspace. Let $\{w_1, w_2, \dots\}$ be a sequence which is dense in K . Then $\{w_0, w_1, w_2, \dots\}$ is also dense in K for any $w_0 \in K$. We define, for each $n \geq 0$, the finite-dimensional space

$$X_n = \text{span}\{w_0, w_1, w_2, \dots, w_n\}$$

if A satisfies coercive condition, where w_0 is given in (1.3.58) and

$$X_n = \text{span}\{w_1, w_2, \dots, w_n\}$$

whenever K is bounded.

We consider the following variational inequalities corresponding to (1.3.59) on X_n :

$$\text{Find } u_n \in X_n \cap K : \langle Au_n - f, y - u_n \rangle \geq 0 \quad \forall y \in X_n \cap K, n = 1, 2, \dots \quad (1.3.61)$$

Here we are looking for $u_n = \sum_{k=0}^n c_{kn} w_k$. We call (1.3.61) *the Galerkin variational inequalities*. When $K = X$ then (1.3.61) becomes the Galerkin equations:

$$\text{Find } u_n \in X_n : Au_n - f = 0 \quad \text{on } X_n, n = 1, 2, \dots \quad (1.3.62)$$

We have the following result on the convergence of solutions to the Galerkin inequalities.

Theorem 1.3.36. *Suppose X is a separable and reflexive Banach space, K is a closed convex set in X , $A : K \rightarrow X^*$ is pseudomonotone and bounded. Assume that A satisfies coercive condition (1.3.58) or K is bounded. Then the following assertions are fulfilled.*

- (i) *For each $n \geq 1$, the Galerkin variational inequality (1.3.61) has a solution u_n and the sequence (u_n) is bounded.*
- (ii) *There exists a subsequence (u_{n_k}) which converges weakly to a solution u of variational inequality (1.3.59). If in addition A satisfies condition $(S)_+$, then $u_{n_k} \rightarrow u$ strongly.*
- (iii) *If A satisfies condition $(S)_+$ and (1.3.59) has a unique solution, then the total sequence (u_n) converges strongly to u .*

Proof. We shall prove the theorem under assumption that A satisfies coercive condition. Without loss of generality, we can assume that $w_0 = 0 \in K$. Otherwise, we can put $K_0 = K - w_0$, $Y_n = X_n - w_0$ and define $A_0 u = A(u + w_0)$ with $u \in K_0$.

(i) Since A is pseudomonotone and bounded, A is demicontinuous (see Proposition 27.7 in [238]). Let us choose a number H such that $H > \|F0\|$. Since A is coercive, there is $R > \|w_0\|$ such that

$$\frac{\langle Fu - F0, u \rangle}{\|u\|} \geq H \quad \text{for } \|u\| \geq R.$$

This implies that

$$\langle Fu, u \rangle > (H - \|F0\|)\|u\| > 0 \quad \text{for } \|u\| \geq R. \quad (1.3.63)$$

Recall that $Fu = Au - f$. Define $j : M \rightarrow X$ to be an injection map and $j^* : X^* \rightarrow M^*$ to be its dual. For each X_n , the mapping

$$F_n := j^* F j : X_n \cap K \cap \overline{B}_X(0, R) \rightarrow M_n^*$$

is continuous. This is because j and j^* are continuous and F is demicontinuous on K (see Theorem 3.10 in [31]). Here

$$\overline{B}_X(0, R) = \{x \in X : \|x\| \leq R\}$$

is a closed ball with radius $R > 0$. Besides, $X_n \cap K \cap \overline{B}_X(0, R)$ is a compact set in X_n , which contains 0. Theorem 3.1 in [139] implies that $VI(F_n, X_n \cap K \cap \overline{B}_X(0, R))$ has a solution $u_n \in X_n \cap K \cap \overline{B}_X(0, R)$, that is,

$$\begin{aligned} & \langle (j^* F j)(u_n), v - u_n \rangle \geq 0 \quad \forall v \in X_n \cap K \cap \overline{B}_X(0, R) \\ \Leftrightarrow & \langle F u_n, j(v - u_n) \rangle \geq 0 \quad \forall v \in X_n \cap K \cap \overline{B}_X(0, R) \\ \Leftrightarrow & \langle F u_n, v - u_n \rangle \geq 0 \quad \forall v \in X_n \cap K \cap \overline{B}_X(0, R). \end{aligned} \quad (1.3.64)$$

We claim that $\|u_n\| < R$ for all $n > 0$. In fact, if there is $n > 0$ such that $\|u_n\| = R$, then from (1.3.63), we have

$$\langle F u_n, 0 - u_n \rangle = -\langle F u_n, u_n - 0 \rangle < 0$$

which contradicts (1.3.64). Hence $\|u_n\| < R$. This implies that for any $y \in X_n \cap K$, there exists $t \in (0, 1)$ such that

$$v_t = u + t(y - u) \in X_n \cap K \cap \overline{B}_X(0, R).$$

Inserting $v = v_t$ into (1.3.64), we get

$$\langle F u_n, y - u_n \rangle \geq 0 \quad \forall y \in X_n \cap K. \quad (1.3.65)$$

It follows that u_n is a solution of (1.3.61) and that $\|u_n\| < R$ for all $n > 0$.

(ii) Since (u_n) is bounded, there exists a subsequence (u_{n_k}) such that $u_{n_k} \rightharpoonup u$. Since A is bounded, $\|A u_n\| \leq \varrho$ for all $n \geq 1$. Note that

$$\overline{\bigcup_{n=1}^{\infty} X_n \cap K} = K.$$

Then for any $\epsilon > 0$, there exists $u' \in X_{n_0}$ such that $\|u' - u\| < \epsilon$. Hence there is $k_0 > 0$ such that $u' \in X_{n_k} \cap K$ for all $k \geq k_0$. From (1.3.65), we have

$$\begin{aligned} \langle A u_{n_k}, u_{n_k} - u \rangle &= \langle A u_{n_k}, u_{n_k} - u' \rangle + \langle A u_{n_k}, u' - u \rangle \\ &\leq \langle f, u_{n_k} - u' \rangle + \varrho \|u' - u\|. \end{aligned}$$

It follows that

$$\limsup_{k \rightarrow \infty} \langle A u_{n_k}, u_{n_k} - u \rangle \leq \langle f, u - u' \rangle + \varrho \epsilon \leq \|f\| \epsilon + \varrho \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we obtain

$$\limsup_{k \rightarrow \infty} \langle A u_{n_k}, u_{n_k} - u \rangle \leq 0. \quad (1.3.66)$$

By the pseudomonotonicity of A , we get

$$\langle Au, u - w \rangle \leq \liminf_{k \rightarrow \infty} \langle Au_{n_k}, u_{n_k} - w \rangle \quad \forall w \in K. \quad (1.3.67)$$

Fixing any $w \in K$, then for any $\epsilon > 0$, there exists $w' \in X_{n_1} \cap K$ such that $\|w' - w\| < \epsilon$ for some $n_1 > 0$. Hence there is $k_1 > 0$ such that $w' \in X_{n_k} \cap K$ for all $k > k_1$. Using (1.3.65) again, we have

$$\begin{aligned} \langle Au_{n_k}, u_{n_k} - w \rangle &= \langle Au_{n_k}, u_{n_k} - w' \rangle + \langle Au_{n_k}, w' - w \rangle \\ &\leq \langle f, u_{n_k} - w' \rangle + \|Au_{n_k}\| \|w' - w\| \\ &\leq \langle f, u_{n_k} - w' \rangle + \varrho \epsilon. \end{aligned}$$

This implies that

$$\begin{aligned} \liminf_{k \rightarrow \infty} \langle Au_{n_k}, u_{n_k} - w \rangle &\leq \langle f, u - w' \rangle + \varrho \epsilon \\ &\leq \langle f, u - w \rangle + \langle b, w - w' \rangle + \varrho \epsilon \\ &\leq \langle f, u - w \rangle + \|f\| \epsilon + \varrho \epsilon. \end{aligned}$$

Consequently

$$\liminf_{k \rightarrow \infty} \langle Au_{n_k}, u_{n_k} - w \rangle \leq \langle f, u - w \rangle.$$

Combining this with (1.3.67), yields

$$\langle Au, u - w \rangle \leq \langle f, u - w \rangle$$

or equivalently,

$$\langle Au - f, u - w \rangle \leq 0.$$

Since w is arbitrary in K , we conclude that u is a solution of (1.3.59). Now we assume that A satisfies condition $(S)_+$. Then from (1.3.66), we have

$$\begin{aligned} \limsup_{k \rightarrow \infty} \langle Au_{n_k} - Au, u_{n_k} - u \rangle &\leq \limsup_{k \rightarrow \infty} \langle Au_{n_k}, u_{n_k} - u \rangle \\ &\quad + \limsup_{k \rightarrow \infty} \langle -Au, u_{n_k} - u \rangle \leq 0. \end{aligned}$$

This implies that $\|u_{n_k} - u\| \rightarrow 0$ as $k \rightarrow \infty$.

(iii). Take any subsequence (u_{m_k}) of (u_n) such that $u_{m_k} \rightharpoonup u_0$. Then by the same argument as in the proof of (ii), we can show that $u_{m_k} \rightarrow u_0$ strongly and u_0 is a solution of (1.3.59). Since (1.3.59) has a unique solution, we have $u_0 = u$. Then by contradiction argument, we can show that $\|u_n - u\| \rightarrow 0$ when $n \rightarrow \infty$. $\square$

Remark 1.3.37. The conclusion of the theorem is still valid for the case where $A : X \rightarrow X^*$ is monotone and hemicontinuous without the boundedness assumption of A . In fact, the hemicontinuous and monotone operators are pseudomonotone operators (see [238, Proposition 27.6]). The monotonicity of A implies that A is locally bounded (see [238, Proposition 26.4]). Based on this fact and the standard argument as in the proof of Theorem 1.3.40, we can show that (Au_n) is bounded and the conclusion is obtained.

From Theorem 1.3.36, we see that the condition $(S)_+$ plays an important role for strong convergence of the sequence (u_n) which are solutions of the Galerkin variational inequalities. However, it is not easy for operators satisfying condition $(S)_+$ even for monotone operators. To surmount this weakness, we need to regularize monotone operators by adding with the normalized duality mapping J_p which satisfies condition $(S)_+$.

Recall that the Banach space $(X, \|\cdot\|)$ is said to be a locally uniformly convex space, if for any $\epsilon > 0$ and $x \in X$ with $\|x\| = 1$, there exists $\delta(\epsilon, x) > 0$ such that

$$[y \in X, \|y\| \leq 1 \text{ and } \|x - y\| > \epsilon] \Rightarrow \left\| \frac{x + y}{2} \right\| < 1 - \delta(\epsilon, x).$$

It is clear that if X is locally uniformly convex, then X is strictly convex.

Lemma 1.3.38. *Suppose X is a reflexive Banach space. Then there exists an equivalent norm such that X and X^* are locally convex and with respect to new norms on X and X^* the following assertions are valid:*

- (i) $J_p : X \rightarrow X^*$ is single-valued and continuous in the strong topologies.
- (ii) For $p \geq 2$, J_p is strictly monotone, that is,

$$\langle J_p(x) - J_p(y), x - y \rangle > 0 \quad \text{for } x \neq y.$$

- (iii) J_p satisfies condition $(S)_+$, that is, if $u_n \rightharpoonup u$ and

$$\limsup_{n \rightarrow \infty} \langle J_p u_n - J_p u, u_n - u \rangle \leq 0$$

then $\|u_n - u\| \rightarrow 0$.

Proof. By the renorming theorem, there exists an equivalent norm such that X and X^* are locally convex (see [238, Proposition 32.23]). By passing to the equivalent norm, we can assume that $(X, \|\cdot\|)$ and $(X^*, \|\cdot\|)$ are locally uniformly convex. It remains to prove (i), (ii) and (iii).

(i) We first show that J_p is a single-valued mapping. If $x = 0$, then $J_p(0) = \{0\}$. Let $x \neq 0$ and $x_1^*, x_2^* \in J_p(x)$. Then $\|x_1^*\|^* = \|x_2^*\|^* = \|x\|^{p-1}$ and

$$2\|x_1^*\|^*\|x\| = 2\|x\|^p = \langle x_1^* + x_2^*, x \rangle \leq \|x_1^* + x_2^*\|^*\|x\|.$$

This implies that $2\|x_1^*\|^* \leq \|x_1^* + x_2^*\|^*$. Hence $\|x_1^*\|^* + \|x_2^*\|^* \leq \|x_1^* + x_2^*\|^*$. Consequently, $\|x_1^*\|^* + \|x_2^*\|^* = \|x_1^* + x_2^*\|^*$. Since X^* is strictly convex, [17, Proposition 1.103] implies that $x_1^* = tx_2^*$ for some $t \geq 0$. Since $\|x_1^*\|^* = \|x_2^*\|^*$, we get $t = 1$ and so $x_1^* = x_2^*$. Hence $J_p(x)$ is singleton.

We now assume that $x_n \rightarrow x$ strongly. Then $x_n \rightharpoonup x$ and $\|J_p x_n\| = \|x_n\|^{p-1} \rightarrow \|x\|^{p-1}$. Hence $(J_p x_n)$ is bounded. Since X^* is reflexive, there is a subsequence $(J_p(x_{n_k}))$ such that $J_p x_{n_k} \rightharpoonup x^* \in X^*$ weakly star. For all $y \in X$, we have

$$\langle x^*, y \rangle = \lim_{k \rightarrow \infty} \langle J_p x_{n_k}, y \rangle \leq \lim_{k \rightarrow \infty} \|x_{n_k}\|^{p-1} \|y\| \leq \|x\|^{p-1} \|y\|.$$

Moreover,

$$\langle x^*, x \rangle = \lim_{k \rightarrow \infty} \langle J_p x_{n_k}, x_{n_k} \rangle = \lim_{k \rightarrow \infty} \|x_{n_k}\|^p = \|x\|^p.$$

Hence $x^* = J_p x$. By the same argument, we see that if $(J_p x_{m_k})$ is any subsequence such that $J_p x_{m_k} \xrightarrow{*} x_1^*$, then $x_1^* = J_p x$. Hence we conclude that $J_p x_n \xrightarrow{*} J_p x$. Since $\|J_p x_n\| \rightarrow \|J_p x\|$ and X^* is locally uniform convex, we have $J_p x_n \rightarrow J_p x$ strongly. Hence J_p is continuous.

(ii) To prove the assertion, we shall use the following inequality (see Lemma 1.1.17):

$$a^p + b^p - a^{p-1}b - ab^{p-1} \geq |a - b|^p \quad \forall a, b \geq 0, p \geq 2. \quad (1.3.68)$$

Taking any $x, y \in X$, using (1.3.68), we have

$$\begin{aligned} \langle J_p x - J_p y, x - y \rangle &= \langle J_p x, x \rangle + \langle J_p y, y \rangle - \langle J_p x, y \rangle - \langle J_p y, x \rangle \\ &\geq \|x\|^p + \|y\|^p - \|x\|^{p-1} \|y\| - \|y\|^{p-1} \|x\| \\ &\geq |||x| - |y|||^p \geq 0. \end{aligned}$$

Let us assume that

$$\langle J_p x - J_p y, x - y \rangle = 0.$$

Then

$$\begin{aligned} 0 &= \langle J_p x - J_p y, x - y \rangle \\ &= \langle J_p x - J_p(\frac{x+y}{2}), \frac{x-y}{2} \rangle + \langle J_p(\frac{x+y}{2}) - J_p y, \frac{x-y}{2} \rangle \\ &\geq |||\frac{x+y}{2} - x|||^p + |||\frac{x+y}{2} - y|||^p. \end{aligned}$$

This implies that

$$\|x\| = \|y\| = \left\| \frac{x+y}{2} \right\|.$$

Since X is strictly convex, we have $x = y$. Hence J_p is strictly monotone.

(iii) Assume that $\limsup_{n \rightarrow \infty} \langle J_p u_n - J_p u, u_n - u \rangle = 0$. Then from

$$\langle J_p u_n - J_p u, u_n - u \rangle \geq \left| \|u_n\| - \|u\| \right|^p,$$

we have $\|u_n\| \rightarrow \|u\|$. On the other hand

$$\begin{aligned} \langle J_p u_n - J_p u, u_n - u \rangle &= \|u_n\|^p + \|u\|^p - \langle J_p u_n, u \rangle - \langle J_p u, u_n \rangle \\ &= \|u_n\|^p + \|u\|^p - \|u_n\|^{p-1} \|u\| - \|u\|^{p-1} \|u_n\| \\ &\quad + [\|u_n\|^{p-1} \|u\| - \langle J_p u_n, u \rangle] + [\|u\|^{p-1} \|u_n\| - \langle J_p u, u_n \rangle] \\ &\geq \left| \|u_n\| - \|u\| \right|^p + [\|u_n\|^{p-1} \|u\| - \langle J_p u_n, u \rangle] + [\|u\|^{p-1} \|u_n\| - \langle J_p u, u_n \rangle]. \end{aligned}$$

Since each term on the right hand side is nonnegative, we conclude that

$$\lim_{n \rightarrow \infty} \langle J_p u, u_n \rangle = \|u\|^p.$$

Since X is reflexive, there is a subsequence u_{n_k} such that $u_{n_k} \rightharpoonup v$. We claim that $v = u$. In fact, we have

$$\|u\|^p = \lim_{n \rightarrow \infty} \langle J_p u, u_n \rangle = \langle J_p u, v \rangle.$$

This implies that $\|u\| \leq \|v\|$. On the other hand, $\|v\| \leq \liminf_{k \rightarrow \infty} \|u_{n_k}\| = \|u\|$. Hence $\|u\| = \|v\|$. Since

$$\|u\|^p = \langle J_p u, v \rangle = \langle J_p u, u \rangle,$$

we have

$$2\|u\|^p = \langle J_p u, u + v \rangle \leq \|u\|^{p-1} \|u + v\|.$$

Hence, $\|u\| + \|v\| = 2\|u\| = \|u + v\|$. Since X is strictly convex, [17, Proposition 1.103] implies that $u = v$. Thus we have shown that $u_{n_k} \rightharpoonup u$ and $\|u_{n_k}\| \rightarrow \|u\|$. Since X is locally convex, we must have $u_{n_k} \rightarrow u$ strongly. Now, by the same argument, we can show that if $\{u_{m_k}\}$ is any subsequence such that $u_{m_k} \rightharpoonup v'$, then $v' = u$ and $u_{m_k} \rightarrow u$. By a contrary argument, we can prove that the total sequence (u_n) converges strongly to u as $n \rightarrow \infty$, that is, $\|u_n - u\| \rightarrow 0$ as $n \rightarrow \infty$. $\square$

Remark 1.3.39. In the case of Hilbert space H , we can take $p = 2$. Then $J_2 u = u$ for all $u \in H$. Although, the duality mapping J_2 is often used for regularization, the normalized duality mapping J_p is more suitable than J in some Banach spaces like $L^p(\Omega)$, $W^{1,p}(\Omega)$ with $p > 1$.

From Lemma 1.3.38, we can assume that $(X, \|\cdot\|)$ and $(X^*, \|\cdot\|^*)$ are locally uniformly convex Banach space so that J_p has properties (i), (ii) and (iii) of the lemma.

For each $\epsilon > 0$, we consider the following *regularized variational inequality*:

$$\text{Find } u \in K : \langle Au - f + \epsilon J_p u, v - u \rangle \geq 0 \quad \forall v \in K. \quad (1.3.69)$$

Here J_p is the normalized duality mapping with $p \geq 2$. Together with (1.3.69), we consider the Galerkin variational inequalities:

$$\text{Find } u_n \in K \cap X_n : \langle Au_n - f + \epsilon J_p u_n, v - u \rangle \geq 0 \quad \forall v \in K \cap X_n. \quad (1.3.70)$$

We now have the following result for the regularized variational inequality (1.3.69).

Theorem 1.3.40. *Suppose X is separable and reflexive, K is a closed convex set in X , $A : X \rightarrow X^*$ is a monotone and hemicontinuous. Assume that A satisfies coercive condition (1.3.58) or K is bounded. Then the following assertions are fulfilled:*

- (i) *For each n , the Galerkin variational inequality (1.3.70) has a unique solution $u_n \in X_n$ and the sequence (u_n) is bounded.*
- (ii) *The sequence (u_n) converges strongly to $u = u_\epsilon$ which is a unique solution of (1.3.69).*
- (iii) *Let $\epsilon = \epsilon_m \rightarrow 0^+$ and u_m is a solution of (1.3.69) corresponding to ϵ_m . Then there exists a subsequence (u_{m_k}) which converges strongly to a solution $\bar{u}$ of variational inequality (1.3.59). If A is strictly monotone, then the total sequence u_m converges strongly to $\bar{u}$.*

Proof. We shall prove the theorem under assumption that A satisfies coercive condition. Without loss of generality, we may assume that $w_0 = 0 \in K$. Define $A'u = Au + \epsilon J_p u$ and $F'u = A'u - f$ for $u \in K$. We want to apply Theorem 1.3.36 for $VI(F', K)$.

Since A is monotone and hemicontinuous, A is pseudomonotone (see [238, Proposition 27.6]). Since J_p is demicontinuous and satisfies $(S)_+$, J_p is pseudomonotone. Consequently, $A' = A + \epsilon J_p$ is pseudomonotone and satisfies $(S)_+$. Moreover, A' is strictly monotone. Since A is coercive, for a fixed number $H > \|A0 - f\|^*$, there exists $R > 0$ such that for if $u \in K$ and $\|u\| \geq R$, then

$$\frac{\langle (Au - f) - (A0 - f), u \rangle}{\|u\|} \geq H.$$

This implies that

$$\langle (Au - f) - (A0 - f), u - 0 \rangle > (H - \|A0 - f\|^*)\|u\| > 0 \quad \text{for} \quad \|u\| \geq R.$$

It follows that

$$\langle (Au - f + \epsilon J_p u) - (A0 - f + \epsilon J_p 0), u \rangle > 0 \quad \text{for} \quad \|u\| \geq R.$$

This is because $\epsilon \langle J_p w - J_p 0, w - 0 \rangle = \epsilon \|w\|^p \geq 0$. Thus we have shown that

$$\langle F'u - F'0, w - 0 \rangle > 0 \quad \text{for} \quad \|u\| \geq R, \quad (1.3.71)$$

where R is independent of ϵ . By the same argument as in the proof of (i) of Theorem 1.3.36, we can show that (1.3.70) has a solution $u_n \in X_n \cap K$. Since A' is strictly monotone, u_n is unique.

Let us now show that $(A'u_n)$ is bounded. In fact, since A is monotone, it is locally bounded (see Proposition 24.6 in [238]). Hence there is $r > 0$ such that the set $A(B_X(0, r))$ is bounded. Hence, there are numbers $r > 0$ and $\gamma > 0$ such that

$$\|Au\| \leq \gamma \quad \forall \|u\| \leq r.$$

Inserting $y = 0$ into the regularized Galerkin inequality (1.3.70), we have

$$\langle Au_n + \epsilon_m J_p u_n, u_n \rangle \leq \langle f, u_n \rangle.$$

Hence

$$\langle Au_n, u_n \rangle + \epsilon_m \|u_n\|^p \leq \|f\|^* \|u_n\| \leq \|f\|^* R.$$

By monotonicity of A , we get for $\|v\| \leq r$ that

$$\begin{aligned} \langle Au_n, v \rangle &\leq \langle Au_n, u_n \rangle - \langle Av, u_n - v \rangle \\ &\leq \|f\|^* R + \|Av\|^* (\|u_n\| + \|v\|) \\ &\leq \|f\|^* R + \gamma(R + r). \end{aligned}$$

This implies that

$$|\langle Au_n, v \rangle| \leq \|f\|^* R + \gamma(R + r) \quad \text{for} \quad \|v\| \leq r.$$

It follows that

$$\|Au_n\|^* = \sup_{\|v\|=r} \frac{1}{r} |\langle Au_n, v \rangle| \leq \frac{1}{r} (\|f\|^* R + \gamma(R + r)) := \varrho.$$

Hence

$$\|A'u_n\|^* \leq \|Au_n\|^* + \epsilon \|J_p u_n\|^* \leq \varrho + \epsilon R^{p-1}.$$

By (ii) and (iii) of Theorem 1.3.36, we conclude that $u_n \rightarrow u = u_\epsilon$ which is a unique solution of (1.3.69). Hence assertions (i) and (ii) are proved. It remains to prove (iii).

Let u_m is a solution of (1.3.69) corresponding to ϵ_m . From (1.3.71), we see that $\|u_m\| \leq R$ for all $m \geq 1$. By the reflexivity of X , there is a subsequence u_{m_k} such that $u_{m_k} \rightharpoonup \bar{u}$ for some $\bar{u} \in K \cap \overline{B}_X(0, R)$. Let us show that $\bar{u}$ is a solution of (1.3.59). By the monotonicity of F' , we have for any $w \in K$ that

$$\langle Aw - f + \epsilon_{m_k} J_p w, u_{m_k} - w \rangle \leq \langle Au_{m_k} - f + \epsilon_{m_k} J_p u_{m_k}, u_{m_k} - w \rangle \leq 0.$$

Hence

$$\langle Aw - f, \bar{u} - w \rangle = \limsup_{k \rightarrow \infty} \langle Aw - f + \epsilon_{m_k} J_p w, u_{m_k} - w \rangle \leq 0.$$

The Minty lemma (see [139, Lemma 1.5]) gives

$$\langle A\bar{u} - f, \bar{u} - w \rangle \leq 0 \quad w \in K. \quad (1.3.72)$$

Hence $\bar{u}$ is a solution of (1.3.59). Since u_{m_k} is a solution of (1.3.69) corresponding to ϵ_{m_k} , we have

$$\langle Au_{m_k} - f + \epsilon_{m_k} J_p u_{m_k}, u_{m_k} - w \rangle \leq 0 \quad \forall w \in K. \quad (1.3.73)$$

Inserting $w = u_{m_k}$ and $w = \bar{u}$ into (1.3.72) and (1.3.73), respectively, we get

$$\begin{aligned} \langle A\bar{u} - f, \bar{u} - u_{m_k} \rangle &\leq 0, \\ \langle Au_{m_k} - f + \epsilon_{m_k} J_p u_{m_k}, u_{m_k} - \bar{u} \rangle &\leq 0. \end{aligned}$$

Adding these inequalities, yields

$$\langle Au_{m_k} - A\bar{u} + \epsilon_{m_k} J_p u_{m_k}, u_{m_k} - \bar{u} \rangle \leq 0.$$

The monotonicity of A implies that

$$\epsilon_{m_k} \langle J_p u_{m_k}, u_{m_k} - \bar{u} \rangle \leq 0.$$

It follows that

$$\langle J_p u_{m_k} - J_p \bar{u}, u_{m_k} - \bar{u} \rangle \leq -\langle J_p \bar{u}, u_{m_k} - \bar{u} \rangle.$$

Passing to the limit, we get

$$\limsup_{k \rightarrow \infty} \langle J_p u_{m_k} - J_p \bar{u}, u_{m_k} - \bar{u} \rangle \leq 0.$$

Using (iii) of Lemma 1.3.38, we conclude that $\|u_{m_k} - \bar{u}\| \rightarrow 0$ as $k \rightarrow \infty$. If A is strictly monotone then $\bar{u}$ is unique. By the same argument, we can show that if u_{i_k} is any subsequence so that $u_{i_k} \rightharpoonup \bar{v}$, then $\bar{v} = \bar{u}$ and $\|u_{i_k} - \bar{v}\| \rightarrow 0$. In this case, by a contrary argument, we can show that $\|u_m - \bar{u}\| \rightarrow 0$ as $m \rightarrow \infty$. The proof of the theorem is complete. $\square$

When $K = X$, we obtain from Theorem 1.3.36 and Remark 1.3.37 the following result.

Corollary 1.3.41. (C.f. [238, Theorem 26.A]) *Suppose X is separable and reflexive, $A : X \rightarrow X^*$ is a monotone, hemicontinuous and coercive operator. Assume that $\{w_1, \dots, w_n, \dots\}$ is a basis in X . Then the following assertions are fulfilled:*

- (i) *For each n , the Galerkin equation (1.3.62) has a solution $u_n \in X_n$ and the sequence (u_n) is bounded.*
- (ii) *There exists a subsequence (u_{n_k}) which converges weakly to a solution u of (1.3.60).*
- (iii) *If A is strictly monotone and satisfies condition $(S)_+$, then the total sequence (u_n) converges strongly to a unique solution u .*

When $X = H$ is a separable Hilbert space we can take $p = 2$ and $J_2 x = x$. In this case the regularized variational inequality becomes

$$\text{Find } u = u_\epsilon \in K : \langle Au - f + \epsilon u, v - u \rangle \geq 0 \quad \forall v \in K \quad (1.3.74)$$

and the Galerkin variational inequalities has the form:

$$\text{Find } u_n \in X_n \cap K : \langle Au_n - f + \epsilon u_n, y - u_n \rangle \geq 0 \quad \forall y \in X_n \cap K, n = 1, 2, \dots \quad (1.3.75)$$

Then from Theorem 1.3.40, we obtain

Corollary 1.3.42. *Suppose H is separable Hilbert space, K is a closed convex set in H , $A : H \rightarrow H^*$ is a monotone and hemicontinuous operator. Assume A satisfies coercive condition or K is bounded. Then the following assertions are fulfilled:*

- (i) *For each n , the Galerkin variational inequality (1.3.75) has a unique solution $u_n \in X_n$ and the sequence (u_n) is bounded.*
- (ii) *The sequence (u_n) converges strongly to the unique solution $u = u_\epsilon$ of (1.3.74).*
- (iii) *Let $\epsilon = \epsilon_m \rightarrow 0^+$ and u_m is a solution of (1.3.74) corresponding to ϵ_m . Then there exists a subsequence (u_{m_k}) which converges strongly*

to a solution $\bar{u}$ of variational inequality (1.3.59). If A is strictly monotone, then $\bar{u}$ is the unique solution and the total sequence (u_m) converges strongly to $\bar{u}$.

Let us present an application of the obtained results for obstacle problem.

Let Ω be a open bounded set in $\mathbb{R}^N$ with smooth boundary, $X = W_0^{1,p}(\Omega)$ be a Sobolev space with $p \geq 2$ and

$$K_\psi = \{u \in X : u(x) \geq \psi(x) \quad \text{a.e. } x \in \Omega\},$$

where $\psi \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega)$. Let $F : K \rightarrow X^*$ be a quasi-linear elliptic operator which is defined by

$$Fu = \sum_{i=1}^N D_i(|D_i u|^{p-2} D_i u) + g(u) - f,$$

where

$$\langle Fu, v \rangle := \int_{\Omega} \left[\sum_{i=1}^N (|D_i u(x)|^{p-2} D_i u(x)) D_i v(x) + g(u(x))v(x) - f(x)v(x) \right] dx$$

for $v \in W_0^{1,p}(\Omega)$. Here $f \in L^q(\Omega)$ with $1/p + 1/q = 1$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ is continuous, nondecreasing and satisfies the following growth condition:

$$|g(u)| \leq C(1 + |u|^{r-1}) \quad (1.3.76)$$

for some constant $C > 0$ and $\frac{1}{r} \leq \frac{1}{p} - \frac{1}{N}$ whenever $2 \leq p < N$,

$$|g(u)| \leq C(1 + |u|^{p-1}) \quad (1.3.77)$$

whenever $p = N$ and

$$|g(u)| \leq C(1 + |u|^{\alpha_0}) \quad (1.3.78)$$

for some $C > 0$ and $\alpha_0 \geq 0$ whenever $p > N$.

We consider the following obstacle problem:

$$\text{Find } u \in K_\psi : \langle Fu, v - u \rangle \geq 0 \quad \forall v \in K_\psi \quad (1.3.79)$$

and the Galerkin variational inequalities:

$$\text{Find } u_n \in X_n \cap K_\psi : \langle Fu_n, v - u_n \rangle \geq 0 \quad \forall v \in X_n \cap K_\psi. \quad (1.3.80)$$

Lemma 1.3.43. *Under assumptions of g , the operator F is uniformly monotone, coercive and hemicontinuous on $W_0^{1,p}(\Omega)$.*

Proof. We use the following facts: there is $\alpha > 0$ such that

$$(|a|^{p-2}a - |b|^{p-2}b)(a - b) \geq \alpha|a - b|^p \quad \forall a, b \in \mathbb{R}, p \geq 2. \quad (1.3.81)$$

This inequality is equivalent to

$$|a|^p + |b|^p - |a|^{p-2}ab - |b|^{p-2}ab \geq \alpha|a - b|^p.$$

If $a, b \geq 0$, then the conclusion follows from (1.3.68) with $\alpha = 1$. If $a \leq 0 \leq b$, then we have

$$b^{p-1} + |a|^{p-1} \geq \frac{1}{2^{p-2}}(b + |a|)^{p-1}$$

Multiplying both sides by $b + |a|$, we obtain the conclusion with $\alpha = \frac{1}{2^{p-2}}$. Besides, the function $\varphi(t) = t^p$ is convex for $x \geq 0$ and $p \geq 2$. Hence we have

$$\frac{1}{N}(t_1^p + t_2^p + \cdots + t_N^p) \geq \left(\frac{t_1 + t_2 + \cdots + t_N}{N}\right)^p \quad (1.3.82)$$

for all $t_i \geq 0$ with $i = 1, 2, \dots, N$ and $p \geq 2$.

Using (1.3.81) and (1.3.82), we have for all $u, w \in W_0^{1,p}(\Omega)$ that

$$\begin{aligned} \langle Fu - Fw, u - w \rangle &= \\ &\int_{\Omega} \left[\sum_{i=1}^N (|D_i u(x)|^{p-2} D_i u(x) - |D_i w(x)|^{p-2} D_i w(x)) (D_i u(x) - D_i w(x)) \right] dx \\ &+ \int_{\Omega} [(g(u(x)) - g(w(x)))[u(x) - w(x)] dx \\ &\geq \alpha \int_{\Omega} \sum_{i=1}^N |D_i u(x) - D_i w(x)|^p dx + 0 \\ &\geq \alpha N^{p-1} \int_{\Omega} |\nabla u(x) - \nabla w(x)|^p dx = \alpha N^{p-1} \|\nabla u - \nabla w\|_p^p. \end{aligned}$$

Using the Poincaré inequality (see [31, Corollary 9.19]), we obtain

$$\langle Fu - Fw, u - w \rangle \geq \alpha N^{p-1} \|\nabla u - \nabla w\|_p^p \geq \beta \|u - w\|_{1,p}^p$$

for some constant $\beta > 0$. This implies that F is uniform monotone and satisfies coercive condition. It remains to show that F is hemicontinuous. For this, we fix any $u, v, w \in X$ and consider the function

$$\begin{aligned} \phi(t) &= \langle F(u + tv), w \rangle = \\ &\int_{\Omega} \left[\sum_{i=1}^N (|D_i(u(x) + tv(x))|^{p-2} D_i(u(x) + tv(x)) D_i w(x)) dx \right] dx \\ &+ \int_{\Omega} g(u(x) + tv(x)) w(x) dx \end{aligned}$$

with $t \in [0, 1]$. If $t_n \in [0, 1]$ and $t_n \rightarrow t_0 \in [0, 1]$, then we have

$$\begin{aligned} & |D_i(u(x) + t_n v(x))|^{p-2} D_i(u(x) + t_n v(x)) D_i w(x) \rightarrow \\ & |D_i(u(x) + t_0 v(x))|^{p-2} D_i(u(x) + t_0 v(x)) D_i w(x) \quad \text{a.e.} \end{aligned}$$

On the other hand

$$\begin{aligned} & |D_i(u(x) + t_n v(x))|^{p-2} |D_i(u(x) + t_n v(x)) D_i w(x)| \leq \\ & (|D_i u(x)| + |D_i v(x)|)^{p-2} (|D_i u(x)| + |D_i v(x)|) |D_i w(x)| \end{aligned}$$

and the term on the right hand side belongs to $L^1(\Omega)$. The dominated convergence theorem implies that

$$\int_{\Omega} \left[\sum_{i=1}^N (|D_i(u(x) + t_n v(x))|^{p-2} D_i(u(x) + t_n v(x)) D_i w(x)) \right] dx$$

converges to

$$\int_{\Omega} \left[\sum_{i=1}^N (|D_i(u(x) + t_0 v(x))|^{p-2} D_i(u(x) + t_0 v(x)) D_i w(x)) \right] dx.$$

Also, we have

$$g(u(x) + t_n v(x)) w(x) \rightarrow g(u(x) + t_0 v(x)) w(x) \quad \text{a.e.}$$

We consider the case $p < N$. The other cases are processed analogously. By (1.3.76), we have

$$|g(u(x) + t_n v(x)) w(x)| \leq C(1 + |u(x)| + |v(x)|)^{r-1} |w(x)|. \quad (1.3.83)$$

By the embedding theorem (see [31, Corollary 9.14]), we have $u, v, w \in L^r(\Omega)$. Hence the term on the right hand side of (1.3.83) belongs to $L^1(\Omega)$. Using the dominated convergence theorem, we have

$$\int_{\Omega} g(u(x) + t_n v(x)) w(x) dx \rightarrow \int_{\Omega} g(u(x) + t_0 v(x)) w(x) dx.$$

In summary, we have show that $\phi(t_n) \rightarrow \phi(t_0)$ as $n \rightarrow \infty$. Hence $\phi(t)$ is continuous on $[0, 1]$ and so F is hemicontinuous on X . $\square$

Proposition 1.3.44. *For each n , the Galerkin variational inequality (1.3.80) has a unique solution u_n and (u_n) converges strongly to the unique solution $\bar{u} \in K_{\psi}$ of the obstacle problem (1.3.79).*

Proof. By Lemma 1.3.43, F is uniformly monotone, coercive and hemicontinuous on X . The conclusion follows from Theorem 1.3.36 and Remark 1.3.37. $\square$

CHAPTER 2

SOME RESULTS IN VARIATIONAL ANALYSIS AND VECTOR OPTIMIZATION

This chapter presents some facts on variational analysis. Its content is based on our published papers [129], [132] and [2].

2.1 Tangent sets, metric regularity and constraint qualification conditions

2.1.1 Tangent sets and subdifferential

In the sequel, Z and E are assumed to be Banach spaces with the dual spaces are Z^* and E^* , respectively. Let T is a metric space. Let $A_t \subset Z$, $t \in T$ be a family of subsets of Z . The Kuratowski limsup and liminf of A_t at t_0 are closed sets given by

$$\begin{aligned}\limsup_{t \rightarrow t_0} A_t &= \{v \in Z \mid \liminf_{t \rightarrow t_0} d(v, A_t) = 0\}, \\ \liminf_{t \rightarrow t_0} A_t &= \{v \in Z \mid \lim_{t \rightarrow t_0} d(v, A_t) = 0\}.\end{aligned}$$

Given a closed set C in Z and a point $z \in C$ the sets

$$\begin{aligned}T^c(C, z) &= \liminf_{\substack{t \rightarrow 0^+ \\ z' \xrightarrow{C} z}} \frac{C - z'}{t} \\ &= \{h \in Z \mid \forall t_n \rightarrow 0^+, \forall z_n \rightarrow z, \exists h_n \rightarrow h, z_n + t_n h_n \in C\},\end{aligned}$$

$$T^b(C, z) = \liminf_{t \rightarrow 0^+} \frac{C - z}{t} = \{h \in Z \mid \forall t_n \rightarrow 0^+, \exists h_n \rightarrow h, z + t_n h_n \in C\}$$

and

$$T(C, z) = \limsup_{t \rightarrow 0^+} \frac{C - z}{t} = \{h \in Z \mid \exists t_n \rightarrow 0^+, \exists h_n \rightarrow h, z + t_n h_n \in C\}$$

are called *the Clarke tangent cone*, *the adjacent tangent cone* and *the contingent cone* to C at z , respectively. It is noted that these cones are

closed and $T^c(C, z)$ is convex. Besides we have the inclusion

$$T^c(C, z) \subset T^b(C, z) \subset T(C, z).$$

Given a convex set D in Z and $z \in D$, we define

$$D(z) = \text{cone}(D - z) = \{\lambda(d - z) | d \in D, \lambda > 0\}.$$

It is well known that when C is convex then

$$T^c(C, z) = T^b(C, z) = T(C, z) = \overline{\text{cone}}(C - z).$$

Given a point $z \in C$ and $v \in Z$, the set

$$\begin{aligned} T^{b2}(C, z, v) &= \liminf_{t \rightarrow 0^+} \frac{C - z - tv}{t^2/2} \\ &= \{w | \forall t_n \rightarrow 0^+, \exists w_n \rightarrow w, z + t_n v + \frac{t_n^2}{2} w_n \in C\} \end{aligned}$$

is called *the inner second-order tangent set* to C at z in the direction $v \in Z$. The set

$$\begin{aligned} T^2(C, z, v) &= \limsup_{t \rightarrow 0^+} \frac{C - z - tv}{t^2/2} \\ &= \{w | \exists t_n \rightarrow 0^+, \exists w_n \rightarrow w, z + t_n v + \frac{t_n^2}{2} w_n \in C\} \end{aligned}$$

is called *the outer second-order tangent set* to C at z in the direction $v \in Z$. Obviously, we have

$$T^{2b}(C, z, v) \subset T^2(C, z, v), \quad T^{2b}(C, z, 0) = T^b(C, z), \quad T^2(C, z, 0) = T(C, z).$$

Note that $T^{2b}(C, z, v)$ is convex whenever C is convex. However, $T^2(C, z, v)$ may not be convex even C is convex. The concepts on tangent sets can be found in [11], [26] and [84].

Definition 2.1.1. Let C be a closed subset in Z and let $z_0 \in C$. The normal cone to C at z_0 is denoted by $N_{co}(C; z_0)$ and defined by

$$N_{co}(C; z_0) = \{z^* \in Z^* : \langle z^*, h \rangle \leq 0 \quad \forall h \in T(C, z_0)\}.$$

From the definition we see that, when C is convex then $N_{co}(C; z_0)$ coincides with the normal cone of C at z_0 in the sense of convex analysis because $T(C, z_0) = \overline{\text{cone}}(C - z_0)$. Precisely, we have

$$N_{co}(C; z_0) = N(C, z_0) = \{z^* \in Z^* \mid \langle z^*, z - z_0 \rangle \leq 0 \quad \forall z \in C\}.$$

Given an extended real function $f : Z \rightarrow (-\infty, +\infty]$, we define

$$\text{dom}(f) = \{z \in Z \mid f(z) < +\infty\}$$

and

$$\text{epi}(f) = \{(z, \alpha) \in Z \times \mathbb{R} \mid f(z) \leq \alpha\}.$$

Definition 2.1.2. Let f be an extended-real function on Z and $x_0 \in \text{dom}(f)$. The subdifferential of f at x_0 is denoted by $\partial_{co}f(x_0)$ and defined by

$$\partial_{co}f(x_0) = \{x^* \in Z^* \mid (x^*, -1) \in N_{co}(\text{epi}(f); (x_0, f(x_0)))\}.$$

Proposition 2.1.3. Let $f : Z \rightarrow (-\infty, +\infty]$ be an extended-real function. If z_0 is a local minimum of f , then

$$0 \in \partial_{co}f(z_0).$$

Proof. Suppose that the conclusion of the proposition is false. Then we get

$$(0; -1) \notin N_{co}(\text{epi}(f), (z_0, f(z_0))).$$

It follows that there exists $(h, \alpha) \in T(\text{epi}(f), (z_0, f(z_0)))$ such that

$$\langle (0, -1), (h, \alpha) \rangle > 0.$$

Hence $\alpha < 0$. By definition of contingent cone, there exist sequences $t_n \rightarrow 0^+$ and $(h_n, \alpha_n) \rightarrow (h, \alpha)$ such that

$$(z_0, f(z_0)) + t_n(h_n, \alpha_n) \in \text{epi}(f) \quad \forall n \geq 1,$$

or equivalently

$$f(z_0 + t_n h_n) \leq f(z_0) + t_n \alpha_n.$$

Since $\alpha_n \rightarrow \alpha$ and $\alpha < 0$, we have $f(z_0 + t_n h_n) < f(z_0)$ for n large enough, a contradiction. The proof is complete. $\square$

In an independence with Penot (see [185, Proposition 4.3]), we showed that

Proposition 2.1.4. Let f be an extend-real function on a Banach space Z and $z_0 \in \text{dom}(f)$. Then

$$z^* \in \partial_{co}f(z_0) \quad \text{iff} \quad \langle z^*, h \rangle \leq f^{co}(z_0, h), \quad \forall h \in Z,$$

where

$$f^{co}(z_0, h) := \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{f(z_0 + th') - f(z_0)}{t}.$$

Remark 2.1.5. It is note that if f is Lipschitz near z_0 , then

$$f^{co}(z_0, h) = \liminf_{t \rightarrow 0^+} \frac{f(z_0 + th) - f(z_0)}{t}. \quad (2.1.1)$$

Indeed, we have

$$f^{co}(z_0, h) \leq \liminf_{t \rightarrow 0^+} \frac{f(z_0 + th) - f(z_0)}{t}.$$

Since f is Lipschitz, there exist a neighborhood U of z_0 and $K > 0$ such that

$$|f(x) - f(y)| \leq K\|x - y\| \quad \text{for all } x, y \in U,$$

we have

$$\begin{aligned} f^{co}(z_0, h) &\geq \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \left\{ \frac{f(z_0 + th) - f(z_0)}{t} - K\|h' - h\| \right\} \\ &= \liminf_{t \rightarrow 0^+} \frac{f(z_0 + th) - f(z_0)}{t}. \end{aligned}$$

Hence (2.1.1) is obtained.

Example 2.1.6. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a function defined by

$$f(x_1, x_2) = x_1^3 + |x_2|.$$

Then

$$\partial_{co}f(0, 0) = \{0\} \times [-1, 1].$$

In fact, it is clear that f is a nonconvex and nonsmooth function. For any $h = (h_1, h_2) \in \mathbb{R}^2$, we get

$$f^{co}((0, 0), h) = \lim_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{f(th'_1, th'_2)}{t} = \lim_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{t^3 h_1'^3 + t|h_2'|}{t} = |h_2|.$$

From this and Proposition 2.1.4 we see that $(x_1^*, x_2^*) \in \partial_{co}f(0, 0)$ iff

$$x_1^* h_1 + x_2^* h_2 \leq |h_2| \quad \text{for all } h_1, h_2 \in \mathbb{R}.$$

This is equivalent to $x_1^* = 0$ and $x_2^* \in [-1, 1]$. Hence $\partial_{co}f(0, 0) = \{0\} \times [-1, 1]$.

It turns out that when f is convex then $\partial_{co}f(z)$ is the subdifferential of f at z in the sense of convex analysis.

Proposition 2.1.7. *If f is a proper convex function and $z_0 \in \text{dom}(f)$ then*

$$\partial_{co}f(z_0) = \partial f(z_0),$$

where

$$\partial f(z_0) := \{z^* \in Z^* : f(z) - f(z_0) \geq \langle z^*, z - z_0 \rangle \quad \forall z \in Z\}.$$

Proposition 2.1.8. *Let C be a closed set in Z and $x_0 \in C$. Then*

$$\partial_{co}\chi_C(x_0) = N_{co}(C; x_0).$$

Proof. For any $x^* \in \partial_{co}\chi_C(x_0)$. By Proposition 2.1.4, we have

$$\langle x^*, h \rangle \leq \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{\chi_C(x_0 + th')}{t} \quad \text{for all } h \in Z.$$

Hence for all $h \in T(C, x_0)$, we get $\langle x^*, h \rangle \leq 0$. Consequently, $x^* \in N_{co}(C; x_0)$ and so $\partial_{co}\chi_C(x_0) \subset N_{co}(C; x_0)$. Conversely, we take any $x^* \in N_{co}(C; x_0)$. We shall show that

$$\langle x^*, h \rangle \leq \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{\chi_C(x_0 + th')}{t} \quad \text{for all } h \in Z. \quad (2.1.2)$$

Indeed, if $h \in T(C, x_0)$ then the inequality (2.1.2) holds. If $h \notin T(C, x_0)$, then the right hand side of (2.1.2) is equal to $+\infty$. Otherwise, $\alpha := \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{\chi_C(x_0 + th')}{t} < +\infty$. Then there exist sequences $t_n \rightarrow 0^+$ and $h_n \rightarrow h$ such that

$$|\chi_C(x_0 + t_n h_n)| \leq t_n(\alpha + 1) < +\infty \quad \text{for all } n \geq 1.$$

It follows that $x_0 + t_n h_n \in C$ for all $n \geq 1$. Consequently, $h \in T(C, x_0)$, which is impossible. Hence $\alpha = \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{\chi_C(x_0 + th')}{t} = +\infty$, so (2.1.2) is satisfied. Thus we have shown that $N_{co}(C; x_0) \subset \partial_{co}\chi_C(x_0)$. The proof is complete. $\square$

The following result helps us to calculate the subdifferential of the sum of two functions.

Proposition 2.1.9. *Let f and g be two extended-real functions on Z and let $x_0 \in \text{dom}(f) \cap \text{dom}(g)$. Assume that f is Gâteaux differentiable at x_0 . Then the following inclusion holds*

$$\partial_{co}(f + g)(x_0) = \nabla f(x_0) + \partial_{co}g(x_0). \quad (2.1.3)$$

Proof. The inclusion $(\supset)$ is obvious. Take any $x^* \in \partial_{co}(f + g)(x_0)$. Then for all $h \in Z$, we have

$$\begin{aligned}
\langle x^*, h \rangle &\leq \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{f(x_0 + th') - f(x_0) + g(x_0 + th') - g(x_0)}{t} \\
&= \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \left[\nabla f(x_0)h' + \frac{o(t)}{t} + \frac{g(x_0 + th') - g(x_0)}{t} \right] \\
&= \nabla f(x_0)h + \liminf_{\substack{h' \rightarrow h \\ t \rightarrow 0^+}} \frac{g(x_0 + th') - g(x_0)}{t} \\
&= \nabla f(x_0)h + g^{co}(x_0, h).
\end{aligned}$$

This implies that $x^* - \nabla f(x_0) \in \partial_{co}g(x_0)$. The proof is complete. $\square$

The following result gives us necessary conditions for programming problems.

Theorem 2.1.10. *Suppose that M is a nonempty closed set in Z , f is an extended-real function on Z and $z_0 \in \text{dom}(f)$ such that f is Gateaux differentiable at z_0 . If z_0 is a local optimum solution of the problem*

$$\begin{cases} f(z) \rightarrow \inf \\ z \in M \end{cases} \quad (2.1.4)$$

then $0 \in \nabla f(z_0) + N_{co}(M, z_0)$.

Proof. Obviously, z_0 is a local optimum solution of the problem

$$f(z) + \chi_M(z),$$

where χ_M is the indicator function of M . By Proposition 2.1.3 we have $0 \in \partial_{co}(f + \chi_M)(z_0)$. By Proposition 2.1.9 we have $0 \in \nabla f(z_0) + N_{co}(M, z_0)$. $\square$

2.1.2 Metric regularity and constraint qualification conditions

Given a multifunction $\Phi : Z \rightrightarrows E$, the set

$$\text{Gr}(\Phi) = \{(z, v) \in Z \times E \mid v \in \Phi(z)\}$$

is called the graph of Φ . Its inverse multifunction $\Phi^{-1} : E \rightrightarrows Z$ is defined by

$$\Phi^{-1}(v) = \{z \in Z \mid v \in \Phi(z)\}.$$

Definition 2.1.11. A multifunction Φ is said to be metrically regular around $(z_0, v_0) \in \text{Gr}(\Phi)$ if there exist positive numbers ϵ_1, ϵ_2 and k such that

$$d(z, \Phi^{-1}(v)) \leq kd(v, \Phi(z))$$

for all $z \in B(z_0, \epsilon_1)$ and $v \in B(v_0, \epsilon_2)$.

It is well known that the metric regularity of Φ around (z_0, v_0) is equivalent to openness of Φ around (z_0, v_0) . For this we can see in Penot [185] and Frankowska [84].

Let $F : Z \rightarrow E$ be single-valued mapping and $D \subset E$ be a nonempty and closed convex set. Define

$$M = F^{-1}(D) = \{z \in Z : F(z) \in D\} = \{z \in Z : 0 \in D - F(z)\}. \quad (2.1.5)$$

The set M plays a role as a constraint set in optimization problems. To derive optimality conditions, we need to compute $T(M, z_0)$ and $T^2(M, z_0, d)$. However, it is not easy to find formulae for $T(M, z_0)$ and $T^2(M, z_0, d)$. In the sequel, we shall show that under constraint qualification conditions, some explicit formulae for $T(M, z_0)$ and $T^2(M, z_0, d)$ can be obtained.

Definition 2.1.12. We say that Robinson's constraint qualification (RCQ) holds at a point $z_0 \in F^{-1}(D)$ with respect to the mapping F and the set D , if the following condition is satisfied:

$$0 \in \text{int}[\nabla F(z_0)(B_Z) - (D - F(z_0)) \cap B_E]. \quad (2.1.6)$$

The Robinson constraint qualification play an important role in the stability of nonlinear constrained systems and optimality conditions of KKT-type.

Let us define a mapping $F_D : Z \rightrightarrows E$ by

$$F_D(z) = -F(z) + D. \quad (2.1.7)$$

The following theorem gives sufficient conditions under which F_D is metrically regular. Its proof can be found in [113, Proposition 3.2] and [26, Proposition 2.89].

Theorem 2.1.13. Suppose that $z_0 \in F^{-1}(D)$, ∇F is continuous at z_0 and the Robinson constraint qualification is valid at z_0 . Then F_D is metrically regular around $(z_0, 0) \in \text{Gr}(F_D)$, that is, there exist positive numbers ϵ_1, ϵ_2 and k such that

$$\text{dist}(z, F^{-1}(D - v)) \leq k \text{dist}(v + F(z), D) \quad (2.1.8)$$

for all $z \in B_Z(z_0, \epsilon_1)$ and $v \in B_E(0, \epsilon_2)$.

The following result gives formulae of $T(M, z_0)$ and $T^2(M, z_0, d)$.

Theorem 2.1.14. (Cf. [26, Proposition 3.33]) *Suppose that $z_0 \in M$, F is of class C^2 around z_0 and RCQ is satisfied at z_0 . Then the following formulas are valid:*

- (i) $T^b(F^{-1}(D), z_0) = T^b(\nabla F(z_0)^{-1}[T^b(D, F(z_0))])$.
- (ii) $T^{2b}(F^{-1}(D), z_0, v) = \nabla F(z_0)^{-1}(T^{2b}(D, F(z_0), \nabla F(z_0)v) - \nabla^2 F(z_0)vv)$.

Proof. Obviously, (i) is a particular case of (ii) when $v = 0$. Hence it is sufficient to prove (ii). Take any $w \in T^{2b}(F^{-1}(D), z_0, v)$. By definition, for any sequence $t_n \rightarrow 0^+$, there exists $w_n \rightarrow w$ such that

$$z_0 + t_n v + \frac{t_n^2}{2} w_n \in F^{-1}(D).$$

It follows that

$$F(z_0 + t_n v + \frac{t_n^2}{2} w_n) \in D.$$

By Taylor expansion, we have

$$\begin{aligned} F(z_0 + t_n v + \frac{t_n^2}{2} w_n) = \\ F(z_0) + t_n \nabla F(z_0)v + \frac{t_n^2}{2} [(\nabla F(z_0)w_n + \nabla^2 F(z_0)vv) + \frac{o(t_n^2)}{t_n^2/2}] \in D. \end{aligned}$$

It follows that

$$\nabla F(z_0)w + \nabla^2 F(z_0)vv \in T^{2b}(D, F(z_0), \nabla F(z_0)v)$$

or equivalently,

$$w \in \nabla F(z_0)^{-1}(T^{2b}(D, F(z_0), \nabla F(z_0)v) - \nabla^2 F(z_0)vv).$$

The inclusion $(\subset)$ in (ii) is proved.

Conversely, suppose that w belongs to the right hand side of (ii). Then for any sequence $t_n \rightarrow 0^+$, there exists $w_n \rightarrow w$ such that

$$d(F(z_0) + t_n \nabla F(z_0)v + \frac{t_n^2}{2} (\nabla F(z_0)w + \nabla^2 F(z_0)vv), D) = o(t_n^2).$$

By Theorem 2.1.13, there exist positive numbers k and ϵ_1 such that

$$d(z, F^{-1}(D)) \leq kd(F(z), D)$$

for all $z \in B_Z(z_0, \epsilon_1)$. It follows that

$$d(z_0 + t_n v + \frac{t_n^2}{2} w_n, F^{-1}(D)) \leq kd(F(z_0 + t_n v + \frac{t_n^2}{2} w_n), D)$$

for n large enough. By Taylor expansion we have

$$\begin{aligned}
& F(z_0 + t_n v + \frac{t_n^2}{2} w_n) \\
&= F(z_0) + t_n \nabla F(z_0) v + \frac{t_n^2}{2} (\nabla F(z_0) w_n + \nabla^2 F(z_0) v v) + o(t_n^2) \\
&= F(z_0) + t_n \nabla F(z_0) v + \frac{t_n^2}{2} (\nabla F(z_0) w + \nabla^2 F(z_0) v v) \\
&+ \frac{t_n^2}{2} (\nabla F(z_0) w_n - \nabla F(z_0) w) + o(t_n^2).
\end{aligned}$$

Hence

$$\begin{aligned}
& d(F(z_0 + t_n v + \frac{t_n^2}{2} w_n), D) \leq \\
& d(F(z_0) + t_n \nabla F(z_0) v + \frac{t_n^2}{2} (\nabla F(z_0) w + \nabla^2 F(z_0) v v), D) + \\
& + \|\frac{t_n^2}{2} (\nabla F(z_0) w_n - \nabla F(z_0) w) + o(t_n^2)\| \\
& \leq o(t_n^2) + \frac{t_n^2}{2} \|\nabla F(z_0) (w_n - w)\| + o(t_n^2) = o(t_n^2).
\end{aligned}$$

Consequently, $d(z_0 + t_n v + \frac{t_n^2}{2} w_n, F^{-1}(D)) = o(t_n^2)$ and so

$$w \in T^{2b}(F^{-1}(D), z_0, v).$$

The proof of the lemma is complete. $\square$

The following result is a primal form of necessary optimality conditions.

Theorem 2.1.15. (Cf. [58, Theorem 4.1]) *Suppose $z_0 \in M = F^{-1}(D)$ is a local optimum solution of (2.1.4). Then the following assertions hold:*

(i) $\langle \nabla f(z_0), v \rangle \geq 0$ for all $v \in T(M, z_0)$.

(ii) If $\langle \nabla f(z_0), v \rangle \leq 0$ then

$$\nabla f(z_0) z + \nabla^2 f(z_0) v^2 \geq 0 \quad \text{for all } z \in T^2(M, z_0, v). \quad (2.1.9)$$

Proof. Assertion (i) follows directly from Theorem 2.1.10. For assertion (ii) we fix any $z \in T^2(M, z_0, v)$. Then there exists $(t_n, z_n) \rightarrow (0^+, z)$ such that

$$z_0 + t_n v + \frac{t_n^2}{2} z_n \in M \quad \text{for all } n \geq 1.$$

This implies that

$$f(z_0 + t_n v + \frac{t_n^2}{2} z_n) \geq f(z_0) \quad \text{for all } n \geq 1.$$

Using Taylor expansion, we get

$$\nabla f(z_0)(t_n v + \frac{t_n^2}{2} z_n) + \frac{1}{2} \nabla^2 f(z_0)(t_n v + \frac{t_n^2}{2} z_n)^2 + o(t_n^2) \geq 0 \quad \text{for all } n \geq 1.$$

Since $\nabla f(z_0)v \leq 0$, we see that

$$\frac{1}{2} \nabla f(z_0) z_n + \frac{1}{2} \nabla^2 f(z_0)(v + \frac{t_n}{2} z_n)^2 + \frac{o(t_n^2)}{t_n^2} \geq 0 \quad \text{for all } n \geq 1.$$

Letting $n \rightarrow +\infty$, we obtain $\nabla f(z_0)z + \nabla^2 f(z_0)v^2 \geq 0$. The proof is complete. $\square$

Remark 2.1.16. Since

$$T^\flat(F^{-1}(D), z_0) \subset T(F^{-1}(D), z_0), \text{ and } T^{2\flat}(M, z_0, v) \subset T^2(M, z_0, v),$$

the conclusion of Theorem 2.1.15 is still valid when $T(F^{-1}(D), z_0)$ and $T^2(M, z_0, v)$ are replaced by $T^\flat(F^{-1}(D), z_0)$ and $T^{2\flat}(M, z_0, v)$, respectively.

Let $H : Z \rightarrow Y$ be a single-valued mapping and

$$A := \{z \in Z : H(z) = 0\}. \quad (2.1.10)$$

To derive a dual form of optimality conditions, we need to compute $T(A, z)$ and $T^2(A, z, v)$ for some $z \in A$ and $v \in Z$. The following results give us such formulae.

Lemma 2.1.17. *Suppose that $z_0 \in A$, $\nabla H(\cdot)$ is continuous at z_0 and $\nabla H(z_0)$ is surjective. Then the following formulae are valid:*

$$(i) \quad T^\flat(A, z_0) = T(A, z_0) = \{v \in Z \mid H'(z_0)v = 0\}.$$

(ii) *If $v \in T(A, z_0)$ then*

$$T^{2\flat}(A, z_0, v) = T^2(A, z_0, v) = \{w \in Z \mid H'(z_0)w + H''(z_0)vv = 0\}.$$

Proof. Obviously, (i) is a particular case of (ii) when $v = 0$. Hence, it is sufficient to prove (ii). To prove, we shall use some techniques in the proof of [177, Theorem 5].

Let $w \in T^{2\flat}(A, z_0, v)$ and sequence $t_n \rightarrow 0^+$. Then there exists a sequence $w_n \rightarrow w$ such that $z_0 + t_n v + \frac{t_n^2}{2} w_n \in A$, that is,

$$H(z_0 + t_n v + \frac{t_n^2}{2} w_n) = H(z_0) = 0.$$

It follows that

$$0 = \frac{H(z_0 + t_n v + \frac{t_n^2}{2} w_n) - H(z_0 + t_n v)}{t_n^2/2} + \frac{H(z_0 + t_n v) - H(z_0) - t_n H'(z_0) v}{t_n^2/2}.$$

Letting $n \rightarrow +\infty$ we see that the first term converges to $H'(z_0)w$, the second term converges to $H''(z_0)vv$. Hence, $H'(z_0)w + H''(z_0)vv = 0$. Consequently,

$$T^{2b}(A, z_0) \subset \{w | H'(z_0)w + H''(z_0)vv = 0\}.$$

Conversely, let w belong to the right hand side. Then for any sequence $t_n \rightarrow 0^+$ we have

$$-H'(z_0)w = H''(z_0)vv = \lim_{n \rightarrow \infty} \frac{H(z_0 + t_n v) - H(z_0) - t_n H'(z_0)v}{t_n^2/2}.$$

Since $H'(z_0)$ is surjective, the Ljusternik theorem (see [108, pp. 30]) implies that, there exist a neighborhood V of z_0 , a number $k > 0$ and a mapping $\theta : V \rightarrow Z$ such that

$$H(\xi + \theta(\xi)) = H(z_0), \quad \|\theta(\xi)\| \leq k \|H(\xi) - H(z_0)\|$$

for all $\xi \in V$. For each n we put $\xi_n = z_0 + t_n v + \frac{t_n^2}{2} w$. Then for n large enough we get

$$0 = H(z_0 + t_n v + \frac{t_n^2}{2} w + \theta(\xi_n)) = H(z_0 + t_n v + \frac{t_n^2}{2} (w + \frac{\theta(\xi_n)}{t_n^2/2})).$$

Define $w_n = w + \frac{\theta(\xi_n)}{t_n^2/2}$. We shall show that $w_n \rightarrow w$ or $\frac{\theta(\xi_n)}{t_n^2/2} \rightarrow 0$. In fact, we have

$$\begin{aligned} \left\| \frac{\theta(\xi_n)}{t_n^2/2} \right\| &\leq k \left\| \frac{H(z_0 + t_n v + \frac{t_n^2}{2} w) - H(z_0)}{t_n^2/2} \right\| = \\ k \left\| \frac{H(z_0 + t_n v + \frac{t_n^2}{2} w) - H(z_0 + t_n v)}{t_n^2/2} + \frac{H(z_0 + t_n v) - H(z_0) - t_n H'(z_0)v}{t_n^2/2} \right\|. \end{aligned}$$

This implies that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left\| \frac{\theta(\xi_n)}{t_n^2/2} \right\| &\leq k \left\| \lim_{n \rightarrow \infty} \frac{H(z_0 + t_n v + \frac{t_n^2}{2} w) - H(z_0 + t_n v)}{t_n^2/2} \right. \\ &\quad \left. + \lim_{n \rightarrow \infty} \frac{H(z_0 + t_n v) - H(z_0) - t_n H'(z_0)v}{t_n^2/2} \right\| \\ &= k \|H'(z_0)w - H'(z_0)w\| = 0. \end{aligned}$$

Thus we show that $w_n \rightarrow w$ and $z_0 + t_n v + \frac{t_n^2}{2} w_n \in A$. Hence $w \in T^{2b}(A, z_0, v)$. Consequently,

$$T^{2b}(A, z_0, v) = \{w | \nabla H(z_0)w + \nabla^2 H(z_0)vv = 0\}.$$

By similar arguments, we can show that

$$T^2(A, z_0, v) = \{w | \nabla H(z_0)w + \nabla^2 H(z_0)vv = 0\}.$$

The proof of the lemma is complete. $\square$

Lemma 2.1.18. *Let $y_0 \in D$ and $v \in D(y_0)$. Then*

$$T^2(D, y_0, v) = T^{2b}(D, y_0, v) = T^b(T^b(D, y_0), v).$$

Proof. Since D is a closed convex cone, we have

$$T^b(D, y_0) = T(D, y_0) = \overline{\text{cone}}(D - y_0) = \overline{\bigcup_{\mu > 0} \mu(D - y_0)}.$$

Also for any $w \in T^b(D, y_0)$ we have

$$T^b(T^b(D, y_0), w) = T(T(D, y_0), w) = \overline{\text{cone}}(\overline{\text{cone}}(D - y_0) - w).$$

If $v \in D(y_0)$. Then there exists $\lambda \geq 0$ such that $v = \lambda(y - y_0)$ for some $y \in D$. By the convexity of D , for any $t_n \rightarrow 0^+$, we have

$$t_n v = t_n \lambda y + (1 - \lambda t_n) y_0 - y_0 \in D - y_0.$$

This implies that $y_0 + t_n v \in D$ and so $0 \in T^{2b}(D, y_0, v)$. By [58, Lemma 3.1], we have

$$T^2(D, y_0, v) \supset T^{2b}(D, y_0, v) = T^b(T^b(D, y_0), v).$$

Conversely, take any $w \in T^2(D, y_0, v)$ then there exist sequences $t_n \rightarrow 0^+$ and $w_n \rightarrow w$ such that

$$y_0 + t_n v + \frac{t_n^2}{2} w_n \in D.$$

It follows that

$$v + \frac{t_n}{2} w_n \in \overline{\text{cone}}(D - y_0).$$

Hence

$$w_n \in \overline{\text{cone}}(\overline{\text{cone}}(D - y_0) - v).$$

Letting $n \rightarrow \infty$ we have

$$w \in \overline{\text{cone}}(\overline{\text{cone}}(D - y_0) - v).$$

Hence

$$T^2(D, y_0, v) \subset \overline{\text{cone}}(\overline{\text{cone}}(D - y_0) - v).$$

The proof is complete. $\square$

The following characterization help us to verify the regularity condition (2.1.6) easier.

Theorem 2.1.19. *Suppose $z_0 \in Z$ and $\nabla F(\cdot)$ is continuous at z_0 . Then the following statements are equivalent:*

- (i) $B_E(0, \rho) \subset [\nabla F(z_0)(B_Z) - (D - F(z_0)) \cap B_E]$ for some $\rho > 0$,
- (ii) $E = [\nabla F(z_0)Z - \text{cone}(D - F(z_0))]$.

Proof. We mimic the proof of [241, Theorem 2.1] for the case where D is a convex cone.

Obviously, (i) implies (ii). Let us show that (ii) implies (i). For this we first claim that

$$\text{cone}(D - F(z_0)) = \bigcup_{n \in \mathbb{N}} n(D - F(z_0)) \cap B_E, \quad (2.1.11)$$

where $\mathbb{N}$ is the set of natural numbers. The inclusions $\supset$ in (2.1.11) are obvious. Take any $w \in \text{cone}(D - F(z_0))$ then $w = \lambda z$ with $z \in (D - F(z_0))$ and $\lambda \geq 0$. We can assume that $\lambda > 0$ and $\|z\| > 1$. By convexity of $D - F(z_0)$ and $0 \in D - F(z_0)$, we have

$$\frac{1}{\|z\|}z + (1 - \frac{1}{\|z\|})0 \in D - F(z_0)$$

It follows that

$$w = \lambda\|z\|\frac{1}{\|z\|}z \in \lambda\|z\|(D - F(z_0)) \cap B_E \subset n(D - F(z_0)) \cap B_E$$

for all $n \geq \lambda\|z\|$. Hence, the inclusion $\subset$ in (2.1.11) is valid and the claim is justified.

For each number $\alpha \geq 0$ we define

$$M_\alpha = \alpha \left[\nabla F(z_0)(B_Z) - \text{cone}(D - F(z_0)) \cap B_E \right].$$

The sets M_α are convex and so their closure $\text{cl}M_\alpha$ are also convex. Moreover,

$$\text{cl}M_\alpha = \alpha \text{cl}M_1. \quad (2.1.12)$$

By (2.1.11) we have

$$\begin{aligned} & [\nabla F(z_0)(Z) - \text{cone}(D - F(z_0))] \\ &= [\nabla F(z_0)(\bigcup_{n \in \mathbb{N}} nB_Z) - \bigcup_{n \in \mathbb{N}} n(D - F(z_0)) \cap B_E] \\ &\subseteq \bigcup_{n \in \mathbb{N}} n[\nabla F(z_0)(B_Z) - (D - F(z_0)) \cap B_E] \\ &= \bigcup_{n \in \mathbb{N}} M_n \subset \bigcup_{n \in \mathbb{N}} \text{cl}M_n. \end{aligned}$$

Hence from (ii) we have

$$E = \bigcup_{n \in \mathbb{N}} \text{cl}M_n. \quad (2.1.13)$$

Since E is a Baire space, a category argument shows that

$$\text{int cl}M_m \neq \emptyset \quad \text{for some } m.$$

Suppose $v \in \text{int cl}M_m$. By (2.1.13), $-v \in \text{cl}M_k$ for some k . By (2.1.12) we have $-\frac{m}{k}v \in \text{cl}M_m$. By property of convex set (see [108, Proposition 4, p.163]), we have

$$\{\lambda(-\frac{m}{k}v) + (1 - \lambda)v \mid 0 \leq \lambda < 1\} \subset \text{int cl}M_m.$$

For $\lambda = \frac{k}{m+k}$ we get $0 \in \text{int cl}M_m$. By (2.1.12) we have $0 \in \text{int cl}A_1$. Hence

$$B_E(0, 2\rho) \subset \text{cl}A_1 \quad \text{for some } \rho > 0.$$

It follows that

$$B_E(0, \rho) = \frac{1}{2}B_E(0, 2\rho) \subset \frac{1}{2}\text{cl}M_1 = \text{cl}M_{1/2} \subset M_{1/2} + B_E(0, \frac{\rho}{2}). \quad (2.1.14)$$

Hence, for $i = 0, 1, 2, \dots$ we have

$$B_E(0, \frac{\rho}{2^i}) = \frac{1}{2^i}B_E(0, \rho) \subset \frac{1}{2^i}[M_{\frac{1}{2}} + B_E(0, \frac{\rho}{2})] = M_{\frac{1}{2^{i+1}}} + B_E(0, \frac{\rho}{2^{i+1}}). \quad (2.1.15)$$

We now take any $w \in B_E(0, \rho)$. By (2.1.14) we have

$$w = \frac{1}{2}v_1 - \frac{1}{2}y_1 + r_1$$

for some $v_1 \in \nabla F(z_0)(B_Z)$, $y_1 \in (D - F(z_0)) \cap B_E$ and $r_1 \in B_E(0, \frac{\rho}{2})$. Then there exists $x_1 \in B_Z$ such that $v_1 = \nabla F(z_0)x_1$ and so

$$w = \frac{1}{2}\nabla F(z_0)x_1 - \frac{1}{2}y_1 + r_1.$$

Apply (2.1.15) with $i = 1$ for r_1 , there exist $v_2 \in \nabla F(z_0)(B_Z)$, $y_2 \in (D - F(z_0)) \cap B_E$ and $r_2 \in B_E(0, \frac{\rho}{2^2})$ such that

$$w = \frac{1}{2}\nabla F(z_0)x_1 + \frac{1}{2^2}v_2 - \frac{1}{2}y_1 - \frac{1}{2^2}y_2 + r_2.$$

Hence there is some $x_2 \in B_Z$ such that $v_2 = \nabla F(z_0)x_2$ and so

$$w = \frac{1}{2}\nabla F(z_0)x_1 + \frac{1}{2^2}\nabla F(z_0)x_2 - \frac{1}{2}y_1 - \frac{1}{2^2}y_2 + r_2.$$

Repeating this procedure for all $n \in \mathbb{N}$ we have

$$w = u_n - p_n + r_n,$$

where

$$u_n = \nabla F(z_0) \left(\sum_{i=1}^n \frac{1}{2^i} x_i \right) \quad \text{with} \quad x_i \in B_Z$$

and

$$p_n = \sum_{i=1}^n \frac{1}{2^i} y_i \quad \text{with} \quad y_i \in (D - F(z_0)) \cap B_E.$$

Obviously, $r_n \rightarrow 0$ as $n \rightarrow \infty$. Putting $\xi_n = \sum_{i=1}^n \frac{1}{2^i} x_i$ we see that $\xi_n \in B_Z$ and

$$\|\xi_n - \xi_{n+m}\| \leq \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{n+m}} \leq \frac{1}{2^n} \quad \forall n, m \geq 1.$$

Hence $\{\xi_n\}$ is a Cauchy sequence. Consequently, $\xi_n \rightarrow \xi$ for some $\xi_z \in B_Z$. It follows that $u_n \rightarrow \nabla F(z_0)(\xi)$. Similarly, $p_n \rightarrow v$ for some $v \in (D - F(z_0)) \cap B_E$. Hence we have

$$w = \nabla F(z_0)(\xi) - v.$$

Consequently,

$$w \in [\nabla F(z_0)(B_Z)] - (D - F(z_0)) \cap B_E,$$

which proves (i). The proof is complete. $\square$

2.2 Optimality conditions for vector optimization problems

2.2.1 Optimality conditions of KKT-type

Let us consider the following vector optimization problem:

$$(VOP1) \quad \begin{cases} \text{Min}_{\mathbb{R}_+^m} & f(z) = (f_1(z), \dots, f_m(z)) \\ \text{s.t.} & G(z) \in Q, \end{cases}$$

where $f: Z \rightarrow \mathbb{R}^m$ and $G: Z \rightarrow E$ are mapping between Banach spaces Z and E , and Q is a nonempty closed convex subset of E . We denote by Σ the feasible set of the problem (VOP1), that is,

$$\Sigma := \{z \in Z \mid G(z) \in Q\}.$$

Recently, there have been some papers dealing with second-order optimality conditions of KKT-type for vector optimization problems. For papers of this topic, we refer the reader to [91,110,111,172] and references given therein. In [91,172], the second-order KKT optimality conditions were derived by scalarization method via the so-called oriented distance function which was used by Ginchev et al. in [92] for the first time. However, this approach has also some certain limits because the oriented distance function is often nonsmooth. In [110,111], using Motzkin's alternative theorem, Jimenez et al. presented some second-order KKT optimality conditions for vector optimization problems under suitable constraint qualification conditions. Although the constraint qualification conditions used in [110,111] seem to be weaker than the Robinson constraint qualification, the “sigma” terms in the second-order conditions do not vanish. In this subsection, we derive second-order KKT optimality conditions for vector optimization directly via separation theorems. The second-order optimality condition of KKT-type is obtained without sigma terms. The obtained results will apply for a class of multiobjective optimal control problems under which the Robinson constraint qualification is fulfilled.

Definition 2.2.1. *We say that $\bar{z} \in \Sigma$ is a local weak Pareto solution of (VOP1) if there exists $\epsilon > 0$ such that for all $z \in B_Z(\bar{z}, \epsilon) \cap \Sigma$, one has $f(z) - f(\bar{z}) \notin -\text{int } \mathbb{R}_+^m$.*

We shall assume that the *Robinson constraint qualification* holds at $\bar{z} \in \Sigma$, that is,

$$0 \in \text{int } [\nabla G(\bar{z})(B_Z) - (Q - G(\bar{z})) \cap B_E].$$

According to Theorem 2.1.19, the Robinson constraint qualification is equivalent to the following condition:

$$E = \nabla G(\bar{z})Z - \text{cone}(Q - G(\bar{z})).$$

When the Robinson constraint qualification holds at $\bar{z}$, we say $\bar{z}$ is a *regular point* of (VOP1). Hereafter, we always assume that $\bar{z}$ is a feasible regular point of (VOP1).

Let us define the following critical cones

$$\begin{aligned} \mathcal{C}_1(\bar{z}) &:= \{d \in Z \mid \nabla f(\bar{z})d \in -\mathbb{R}_+^m, \nabla G(\bar{z})d \in T(Q; G(\bar{z}))\}, \\ \mathcal{C}_{01}(\bar{z}) &:= \{d \in Z \mid \nabla f(\bar{z})d \in -\mathbb{R}_+^m, \nabla G(\bar{z})d \in \text{cone}(Q - G(\bar{z}))\}, \\ \mathcal{C}_{1*}(\bar{z}) &:= \overline{\mathcal{C}_{01}(\bar{z})}. \end{aligned}$$

For each $d \in \mathcal{C}(\bar{z})$, put

$$I(\bar{z}, d) := \{i \in I \mid \nabla f_i(\bar{z})d = 0\},$$

where $I := \{1, \dots, m\}$.

Lemma 2.2.2. *If $\bar{z}$ is a local weak Pareto solution of (VOP1), then*

$$\mathcal{C}_1(\bar{z}) = \{d \in Z \mid \nabla f(\bar{z})d \in -\mathbb{R}_+^m \setminus (-\text{int } \mathbb{R}_+^m), \nabla G(\bar{z})d \in T(Q; G(\bar{z}))\}.$$

Consequently, the following set

$$\{d \in Z \mid \nabla f(\bar{z})d \in -\text{int } \mathbb{R}_+^m, \nabla G(\bar{z})d \in T(Q; G(\bar{z}))\}$$

is empty.

Proof. By Theorem 2.1.14, we have

$$\nabla G(\bar{z})^{-1}(T(Q; G(\bar{z}))) = T(G^{-1}(Q); \bar{z}) = T(\Sigma; \bar{z}).$$

Take $d \in \mathcal{C}_1(\bar{z})$ and get $d \in T(\Sigma; \bar{z})$. Hence, there exists a sequence $\{(t_k, d_k)\}$ converging to $(0^+, d)$ such that $\bar{z} + t_k d_k \in \Sigma$ for all $k \in \mathbb{N}$. Since $\bar{z}$ is a locally weak Pareto solution of the problem (VOP1), we may assume that

$$f(\bar{z} + t_k d_k) - f(\bar{z}) \in \mathbb{R}^m \setminus (-\text{int } \mathbb{R}_+^m)$$

for all $k \in \mathbb{N}$. From the Mean Value Theorem for differentiable functions, we have

$$t_k \nabla f(\bar{z})d_k + o(t_k) \in \mathbb{R}^m \setminus (-\text{int } \mathbb{R}_+^m),$$

or, equivalently,

$$\nabla f(\bar{z})d_k + \frac{o(t_k)}{t_k} \in \mathbb{R}^m \setminus (-\text{int } \mathbb{R}_+^m)$$

for all $k \in \mathbb{N}$. Letting $k \rightarrow \infty$, by the closeness of $\mathbb{R}^m \setminus (-\text{int } \mathbb{R}_+^m)$, we get

$$\nabla f(\bar{z})d \in \mathbb{R}^m \setminus (-\text{int } \mathbb{R}_+^m).$$

The proof is complete. □

Let us denote by $\Lambda_1(\bar{z})$ be the set of normalized Karush–Kuhn–Tucker multipliers of (VOP1) at $\bar{z}$, that is, $\Lambda_1(\bar{z})$ consists of vector $(\lambda, e^*) \in \mathbb{R}_+^m \times E^*$ satisfying the following conditions:

$$|\lambda| = 1, \lambda^T Df(\bar{z}) + DG(\bar{z})^* e^* = 0, e^* \in N(Q; G(\bar{z})).$$

Lemma 2.2.3. *Let $\bar{z}$ be a local weak Pareto solution of (VOP1) and $\bar{z}$ is a regular point. Then $\Lambda_1(\bar{z})$ is a nonempty bounded and compact set of $\mathbb{R}^m \times E^*$ with respect to topology $\tau_{\mathbb{R}^m} \times \tau(E^*, E)$, where $\tau(E^*, E)$ is the weak-star topology in E^* .*

Proof. We first claim that $\Lambda_1(\bar{z})$ is nonempty. Indeed, put

$$\Psi := \{(\nabla f_1(\bar{z})d + r_1, \dots, \nabla f_m(\bar{z})d + r_m, \nabla G(\bar{z})d - v) \mid d \in Z, \\ v \in T(Q; G(\bar{z})), r_i \geq 0, i \in I\}.$$

Then Ψ is a convex subset of $\mathbb{R}^m \times E$. By the Robinson constraint qualification, there exists $\rho > 0$ such that

$$B_E(0, \rho) \subset \nabla G(\bar{z})(B_Z) - (Q - G(\bar{z})) \cap B_E. \quad (2.2.1)$$

This implies that

$$B_E(0, \rho) \subset \nabla G(\bar{z})(B_Z) - T(Q; G(\bar{z})) \cap B_E.$$

For each $i \in I$, put

$$\alpha_i := \|\nabla f_i(\bar{z})\| = \sup_{d \in B_Z} |\nabla f_i(\bar{z})d|.$$

It is easily seen that

$$(\alpha_1, +\infty) \times \dots \times (\alpha_m, +\infty) \times B(0, \rho) \subset \Psi.$$

Thus, Ψ has a nonempty interior. We show that $(0, 0) \notin \text{int } \Psi$. If otherwise, there exist $\epsilon_1 > 0, \dots, \epsilon_m > 0$ such that

$$(-\epsilon_1, \epsilon_1) \times \dots \times (-\epsilon_m, \epsilon_m) \times \{0\} \subset \Psi.$$

This implies that there exist $d \in Z$, $v \in T(Q; G(\bar{z}))$ and $r_i \geq 0$, $i \in I$ such that

$$\begin{cases} \nabla f_i(\bar{z})d + r_i < 0, & i \in I, \\ \nabla G(\bar{z})d - v = 0. \end{cases}$$

Consequently, the system

$$\begin{cases} \nabla f_i(\bar{z})d < 0, & i \in I, \\ \nabla G(\bar{z})d \in T(Q; G(\bar{z})), \end{cases}$$

has at least one solution $d \in Z$, which contradicts the conclusion of (2.2.2). We now can separate $(0, 0)$ from Ψ by a hyperplane, i.e., there exists a functional $(\lambda, e^*) \in (\mathbb{R}^m \times E^*) \setminus \{(0, 0)\}$ such that

$$\sum_{i=1}^m \lambda_i (\nabla f_i(\bar{z})d + r_i) + \langle e^*, \nabla G(\bar{z})d - v \rangle \geq 0 \quad (2.2.2)$$

for all $d \in Z$, $v \in T(Q; G(\bar{z}))$, $r_i \geq 0$, $i \in I$. From (2.2.2) it follows that $\lambda_i \geq 0$ for all $i \in I$, that is, $\lambda \in \mathbb{R}_+^m$. Putting $v = 0$ and $r_i = 0$, $i \in I$ into (2.2.2), we get

$$(\langle \lambda, \nabla f(\bar{z}) \rangle + \nabla G(\bar{z})^* e^*) d \geq 0 \quad \forall d \in Z.$$

Thus, $\langle \lambda, \nabla f(\bar{z}) \rangle + \nabla G(\bar{z})^* e^* = 0$. Putting this equation and $r_i = 0$, $i \in I$, into (2.2.2), one has

$$e^*(v) \leq 0 \quad \forall v \in T(Q; G(\bar{z})).$$

Hence, $e^* \in N(Q; G(\bar{z}))$. We now show that $\lambda \neq 0$. Indeed, if otherwise, then we have

$$\langle e^*, \nabla G(\bar{z}) d - v \rangle \geq 0 \tag{2.2.3}$$

for all $d \in Z$, $v \in T(Q; G(\bar{z}))$. Again by the Robinson constraint qualification, one has

$$E = \nabla G(\bar{z})(Z) - Q(G(\bar{z})).$$

This and (2.2.3) imply that $e^* = 0$, a contradiction. Put $(\bar{\lambda}, \bar{e}^*) := \left(\frac{\lambda}{|\lambda|}, \frac{e^*}{|e^*|} \right)$. Then we have $(\bar{\lambda}, \bar{e}^*) \in \Lambda_1(\bar{z})$, as required.

We now claim that $\Lambda_1(\bar{z})$ is bounded. In fact, fix $(\lambda_0, e_0^*) \in \Lambda_1(\bar{z})$. Then, for any (λ, e^*) belonging to $\Lambda_1(\bar{z})$, we have

$$\begin{cases} \langle \lambda_0, \nabla f(\bar{z}) \rangle + \nabla G(\bar{z})^* e_0^* = 0, & e_0^* \in N(Q; G(\bar{z})), \\ \langle \lambda, \nabla f(\bar{z}) \rangle + \nabla G(\bar{z})^* e^* = 0, & e^* \in N(Q; G(\bar{z})). \end{cases}$$

By (2.2.1), for each $y \in B_E(0, \rho)$, there exist $z \in B_Z$ and $w \in (Q - G(\bar{z})) \cap B_E$ such that $y = \nabla G(\bar{z})z - w$. It follows that

$$\begin{aligned} \langle e_0^* - e^*, y \rangle &= \langle e_0^* - e^*, \nabla G(\bar{z})z - w \rangle \\ &= \langle \nabla G(\bar{z})^* (e_0^* - e^*), z \rangle - \langle e_0^* - e^*, w \rangle \\ &= \langle (\lambda - \lambda_0) \nabla f(\bar{z}), z \rangle - \langle e_0^*, w \rangle + \langle e^*, w \rangle \\ &\leq 2 \|\nabla f(\bar{z})\| + \|e_0^*\|. \end{aligned}$$

Thus,

$$\begin{aligned} \langle -e^*, y \rangle &\leq 2 \|\nabla f(\bar{z})\| + \|e_0^*\| + \|e_0^*\| \|y\| \\ &\leq 2 \|\nabla f(\bar{z})\| + \|e_0^*\| + \|e_0^*\| \rho. \end{aligned}$$

Replacing y by ρy with $\|y\| \leq 1$, we get

$$\|e^*\| \rho \leq 2 \|\nabla f(\bar{z})\| + \|e_0^*\| (1 + \rho).$$

Consequently,

$$\|e^*\| \leq \frac{2}{\rho} \|\nabla f(\bar{z})\| + \frac{1+\rho}{\rho} \|e_0^*\|.$$

Thus, $\Lambda_1(\bar{z})$ is bounded. It is easy to check that the set $\Lambda_1(\bar{z})$ is closed with respect to topology $\tau_{\mathbb{R}^m} \times \tau(E^*, E)$. By [31, Theorem 3.16], $\Lambda_1(\bar{z})$ is compact. $\square$

To derive second-order necessary conditions for (VOP1), we need the following result.

Lemma 2.2.4. *Let $\bar{z} \in \Sigma$ and $d \in \mathcal{C}_1(\bar{z})$. If $\bar{z}$ is a locally weak Pareto solution of (VOP1), then the following system*

$$\nabla f_i(\bar{z})z + \nabla^2 f_i(\bar{z})d^2 < 0, \quad \forall i \in I(\bar{z}, d), \quad (2.2.4)$$

$$\nabla G(\bar{z})z + \nabla^2 G(\bar{z})d^2 \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d), \quad (2.2.5)$$

has no solution $z \in Z$.

Proof. Arguing by contradiction, assume that the system (2.2.4)–(2.2.5) admits a solution, say z . From (2.2.5) it follows that

$$z \in \nabla G(\bar{z})^{-1} \left[T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d) - \nabla^2 G(\bar{z})d^2 \right].$$

By the Robinson constraint qualification and Theorem 2.1.14, we have

$$T^{2b}(\Sigma; \bar{z}, d) = \nabla G(\bar{z})^{-1} \left[T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d) - \nabla^2 G(\bar{z})d^2 \right].$$

Thus, $z \in T^{2b}(\Sigma; \bar{z}, d)$. Let $\{t_k\}$ be an arbitrary sequence converging to 0^+ . Then, there exists a sequence $\{z_k\}$ tending to z such that

$$w_k := \bar{z} + t_k d + \frac{1}{2} t_k^2 z_k \in \Sigma, \quad \forall k \in \mathbb{N}.$$

For each $i \in I(\bar{z}, d)$ and $k \in \mathbb{N}$, we have

$$\begin{aligned} f_i(w_k) - f_i(\bar{z}) &= [f_i(w_k) - f_i(\bar{z} + t_k d)] + [f_i(\bar{z} + t_k d) - f_i(\bar{z}) - t_k \langle \nabla f_i(\bar{z}), d \rangle] \\ &= \frac{1}{2} t_k^2 \langle \nabla f_i(\bar{z} + t_k d), z_k \rangle + \frac{1}{2} t_k^2 \nabla^2 f_i(\bar{z}) d^2 + o(t_k^2). \end{aligned}$$

Therefore,

$$\lim_{k \rightarrow \infty} \frac{f_i(w_k) - f_i(\bar{z})}{\frac{1}{2} t_k^2} = \langle \nabla f_i(\bar{z}), z \rangle + \nabla^2 f_i(\bar{z}) d^2.$$

This and (2.2.4) imply that

$$f_i(w_k) < f_i(\bar{z})$$

for all $i \in I(\bar{z}, d)$ and k large enough. For each $i \in I \setminus I(\bar{z}, d)$, we have $\langle \nabla f_i(\bar{z}), d \rangle < 0$. From this and the fact that

$$\lim_{k \rightarrow \infty} \frac{f_i(w_k) - f_i(\bar{z})}{t_k} = \langle \nabla f_i(\bar{z}), d \rangle,$$

it follows that

$$f_i(w_k) < f_i(\bar{z})$$

for all k large enough. Therefore, there exists k large enough such that

$$f_i(w_k) < f_i(\bar{z}) \quad \forall i \in I,$$

which contradicts the fact that $\bar{z}$ is a locally weak Pareto solution of (VOP1). $\square$

Problem (VOP1) is associated with the Lagrangian

$$\mathcal{L}_1(z, \lambda, e^*) = \lambda^T f(z) + e^* G(z).$$

The following theorem gives some second-order necessary optimality conditions for (VOP1).

Theorem 2.2.5. *Suppose that $\bar{z}$ is a locally weak Pareto solution of (VOP1). Then, for each $d \in \mathcal{C}_{1*}(\bar{z})$, there exists $(\lambda, e^*) \in \Lambda_1(\bar{z})$ such that the following non-negative second-order condition is valid:*

$$\nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d) = \lambda^T D^2 f(\bar{z})(d, d) + e^* D^2 G(\bar{z})(d, d) \geq 0.$$

Proof. We first prove the theorem for the case $d \in \mathcal{C}_{01}(\bar{z})$. We claim that $0 \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$. In fact, since $\nabla G(\bar{z})d \in \text{cone}(Q - G(\bar{z}))$, there exists $\mu > 0$ such that $\mu \nabla G(\bar{z})d \in Q - G(\bar{z})$. From $0 \in Q - G(\bar{z})$ and the convexity of $Q - G(\bar{z})$, for any $0 < \alpha < \mu$, we get

$$\frac{\alpha}{\mu} \mu \nabla G(\bar{z})d + \left(1 - \frac{\alpha}{\mu}\right) \cdot 0 \in Q - G(\bar{z}).$$

This implies that $G(\bar{z}) + \alpha \nabla G(\bar{z})d \in Q$ for all $\alpha \in (0, \mu)$. Then, by definition, we have $0 \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$ as required. We consider the following set

$$\begin{aligned} \Pi := & \{(\nabla^2 f_1(\bar{z})d^2 + \nabla f_1(\bar{z})z + r_1, \dots, \nabla^2 f_m(\bar{z})d^2 + \nabla f_m(\bar{z})z + r_m, \nabla^2 G(\bar{z})d^2 + \nabla G(\bar{z})z - v) \\ & | z \in Z, v \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d), r_i \geq 0, i \in I\}. \end{aligned}$$

By Lemma 2.2.2, the set $I(\bar{z}, d)$ is nonempty. Thus, by choosing $\lambda_i = 0$ and removing the component i -th of Π for $i \in I \setminus I(\bar{z}, d)$, we may assume

that $I(\bar{z}, d) = I$. We claim that Π is a convex set with a nonempty interior and $(0, 0) \notin \text{int } \Pi$. The convexity of Π follows directly from the convexity of $T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$ and $\mathbb{R}_+^m$. By the Robinson constraint qualification, there exists $\rho > 0$ such that

$$B_E(0, \rho) \subset \nabla G(\bar{z})[B_Z] - T^b(Q; G(\bar{z})) \cap B_E.$$

By the inclusion $0 \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$ and [58, Proposition 3.1], we have

$$T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})) = T^b(T^b(Q; G(\bar{z})), \nabla G(\bar{z})).$$

Putting $V := \nabla G(\bar{z}) + B_E(0, \rho)$ and $\rho_1 := 1 + \|\nabla G(\bar{z})d\|$, we have

$$\begin{aligned} V &\subset \nabla G(\bar{z})[B_Z] - \left[T^b(Q; G(\bar{z})) \cap B_E - \nabla G(\bar{z}) \right] \\ &\subset \nabla G(\bar{z})[B_Z] - \left[T^b(Q; G(\bar{z})) - \nabla G(\bar{z}) \right] \cap B_E(0, \rho_1) \\ &\subset \nabla G(\bar{z})[B_Z] - T^b(T^b(Q; G(\bar{z})), \nabla G(\bar{z})) \cap B_E(0, \rho_1) \\ &= \nabla G(\bar{z})[B_Z] - T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})) \cap B_E(0, \rho_1). \end{aligned}$$

By the Robinson constraint qualification,

$$E = \nabla G(\bar{z})[Z] - Q(G(\bar{z})) = \nabla G(\bar{z})[Z] - T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d). \quad (2.2.6)$$

For each $i \in I$, we put

$$\alpha_i := \nabla^2 f_i(\bar{z})d^2 + \sup \{ \nabla f_i(\bar{z})z \mid z \in B_Z(0, \rho_1) \} < +\infty.$$

We then have

$$(\alpha_1, +\infty) \times (\alpha_2, +\infty) \times \dots \times (\alpha_m, +\infty) \times \widehat{V} \subset \Pi,$$

where $\widehat{V} := \nabla^2 G(\bar{z})d^2 + V$. This implies that the interior of Π is nonempty. We now show that $(0, 0) \notin \text{int } \Pi$. If otherwise, there exist $\epsilon_1 > 0, \epsilon_2 > 0, \dots, \epsilon_m > 0$ such that

$$(-\epsilon_1, \epsilon_1) \times (-\epsilon_2, \epsilon_2) \times \dots \times (-\epsilon_m, \epsilon_m) \times \{0\} \subset \Pi.$$

This implies that there exist $z \in Z, v \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d), r_i \geq 0, i \in I$, satisfying

$$\begin{cases} \nabla^2 f_i(\bar{z})d^2 + \nabla f_i(\bar{z})z + r_i < 0, & i \in I, \\ \nabla^2 G(\bar{z})d^2 + \nabla G(\bar{z})z - v = 0. \end{cases}$$

Consequently, z is a solution of the following system

$$\begin{cases} \nabla^2 f_i(\bar{z})d^2 + \nabla f_i(\bar{z})z < 0, & i \in I, \\ \nabla^2 G(\bar{z})d^2 + \nabla G(\bar{z})z \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d), \end{cases}$$

which contradicts 2.2.4. Since $(0, 0) \notin \text{int } \Pi$, we can separate $(0, 0)$ from Π by a hyperplane, i.e., there exists a nonzero functional $(\lambda, e^*) \in (\mathbb{R}^m \times E^*)$ such that

$$\sum_{i=1}^m \lambda_i (\nabla^2 f_i(\bar{z})d^2 + \nabla f_i(\bar{z})z + r_i) + \langle e^*, \nabla^2 G(\bar{z})d^2 + \nabla G(\bar{z})z - v \rangle \geq 0 \quad (2.2.7)$$

for all $r_i \geq 0$, $i \in I$, $z \in Z$, and $v \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$. By (2.2.7), we have $\lambda_i \geq 0$ for all $i \in I$, i.e., $\lambda \in \mathbb{R}_+^m$. We claim that λ is a nonzero vector. If otherwise, then we have

$$\langle e^*, \nabla^2 G(\bar{z})d^2 + \nabla G(\bar{z})z - v \rangle \geq 0$$

for all $z \in Z$, and $v \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$, or, equivalently,

$$\langle e^*, \nabla^2 G(\bar{z})d^2 + w \rangle \geq 0$$

for all $w \in \nabla G(\bar{z})(Z) - T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$. This and (2.2.6) imply that $e^* = 0$, contrary to the fact that $(\lambda, e^*) \neq (0, 0)$. We now rewrite (2.2.7) as follows

$$\langle \nabla_z \mathcal{L}_1(\bar{z}, \lambda, e^*), z \rangle + \nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d) + \sum_{i=1}^m \lambda_i r_i - \langle e^*, v \rangle \geq 0 \quad (2.2.8)$$

for all $r_i \geq 0$, $i \in I$, $z \in Z$, and $v \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$. It follows that $\nabla_z \mathcal{L}_1(\bar{z}, \lambda, e^*) = 0$. Putting $\nabla_z \mathcal{L}_1(\bar{z}, \lambda, e^*) = 0$ and $r_i = 0$, $i \in I$, into (2.2.8), we get

$$\nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d) \geq \langle e^*, v \rangle$$

for all $v \in T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)$. Thus,

$$\nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d) \geq \sigma(e^*, T^{2b}(Q; G(\bar{z}), \nabla G(\bar{z})d)) \geq 0. \quad (2.2.9)$$

By dividing both sides of (2.2.9) by $|\lambda|$, we obtain that

$$\sup_{(\lambda, e^*) \in \Lambda_1(\bar{z})} \nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, e^*)(d, d) \geq 0.$$

We now take any $d \in \mathcal{C}_{1*}(\bar{z})$. Then, there exists a sequence $\{d_k\} \subset \mathcal{C}_{01}(\bar{z})$ converging to d . From what has already been proved, we have

$$\sup_{(\lambda, e^*) \in \Lambda_1(\bar{z})} \nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, e^*)(d_k, d_k) \geq 0. \quad (2.2.10)$$

Since the set $\Lambda_1(\bar{z})$ is compact in topology $\tau_{\mathbb{R}^m} \times \tau(E^*, E)$, the function

$$\psi(d) := \sup_{(\lambda, e^*) \in \Lambda_1(\bar{z})} \nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d)$$

is continuous. Letting $k \rightarrow \infty$ in (2.2.10), we get

$$\sup_{(\lambda, e^*) \in \Lambda_1(\bar{z})} \nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d) \geq 0.$$

Again, by the compactness of $\Lambda_1(\bar{z})$, there exists $(\lambda, e^*) \in \Lambda_1(\bar{z})$ such that

$$\nabla_{zz}^2 \mathcal{L}_1(\bar{z}, \lambda, e^*)(d, d) \geq 0.$$

The proof is complete. □

Remark 2.2.6. From Theorem 3.1 we want to ask whether the conclusion is still true if $\mathcal{C}_{1*}(\bar{z})$ is replaced by $\mathcal{C}_1(\bar{z})$. Clearly, $\mathcal{C}_{1*}(\bar{z}) \subseteq \mathcal{C}_1(\bar{z})$. In the case of single-objective ($m = 1$) under assumptions that Q is *polyhedral* at $G(\bar{z})$ and $\nabla G(\bar{z})$ is surjective, [26, Proposition 3.54] showed that $\mathcal{C}_1(\bar{z}) = \mathcal{C}_{1*}(\bar{z})$. However, when $m > 1$, the proof of Proposition 3.54 in [26] is collapsed. The reason is that the condition $\langle \lambda, \nabla f(\bar{z}) \rangle = 0$ with $\lambda \neq 0$ does not imply $\nabla f(\bar{z}) = 0$. We do not know whether the equality $\mathcal{C}_1(\bar{z}) = \mathcal{C}_{1*}(\bar{z})$ is valid.

2.2.2 Optimality conditions of Fritz-John type

In the previous subsection, we derived first-and second-order optimality conditions of KKT-type for vector optimization problems in Banach spaces. One may ask whether we can derive optimality conditions of Fritz-John type for the problem, and in addition if a constraint qualification condition is satisfied, then the optimality conditions become optimality conditions of KKT-type. Below, we show that if some of constraint sets have nonempty interiors, then we can establish optimality conditions of Fritz-John type for locally optimal solutions to vector optimization problem in Banach spaces.

As before, Z , E and W are Banach spaces with the dual spaces Z^* , E^* and W^* , respectively. We denote by $\tau(\mathbb{R}^m)$ the strong topology in $\mathbb{R}^m$, by $\sigma(E^*, E)$ the weak* topology in E^* and by $\sigma(W^*, W)$ the weak* topology in W^* . Given a Banach space X , $B_X(x_0, r)$ stands for the open ball with center x_0 and radius r . Given a subset M in X . We shall denote by $\text{int}M$ and $\overline{M}$ the interior and the closure of M , respectively.

Let $f = (f_1, f_2, \dots, f_k): Z \rightarrow \mathbb{R}^k$, $F: Z \rightarrow E$, $G: Z \rightarrow W$ and $H = (H_1, H_2, \dots, H_n): Z \rightarrow \mathbb{R}^n$ be mappings and K be a nonempty closed convex set in W . We consider the following vector optimization problem.

$$(VOP2) \quad \begin{cases} f(z) \rightarrow \min \\ \text{s.t.} \\ F(z) = 0, \\ H(z) \leq 0, \\ G(z) \in K. \end{cases}$$

We shall denote by Σ_2 the feasible set of (VOP2), that is,

$$\Sigma_2 = \{z \in Z \mid F(z) = 0, H_i(z) \leq 0 \text{ and } G(z) \in K, i = 1, 2, \dots, n\}.$$

Given a feasible point $z_0 \in \Sigma_2$, we denote by $f'(z_0)$ or $Df(z_0)$ the first order derivative of f at z_0 and by $f''(z_0)$ or $D^2f(z_0)$ the second-order derivative of f at z_0 . Let us impose the following assumptions.

(A1) $\text{int}(K) \neq \emptyset$;

(A2) The mappings f, F and G are of class C^2 around z_0 ;

(A3) The range of $DF(z_0)$ is closed in E .

By defining $\tilde{K} = 0 \times (-\infty, 0]^n \times K$, $T = E \times \mathbb{R}^n \times W$ and $\tilde{G} = (F, H, G)$, the constraint of (VOP2) is equivalent to the constraint $\tilde{G}(z) \in \tilde{K}$. Recall that a feasible point z_0 is said to satisfy the Robinson constraint qualification if

$$0 \in \text{int}\{\tilde{G}(z_0)(B_Z) + (D\tilde{G}(z_0)Z - \tilde{K}) \cap B_T\}, \quad (2.2.11)$$

By Theorem 2.1.19, if K is closed, then condition (2.2.11) is equivalent to

$$D\tilde{G}(z_0)Z - \text{cone}(\tilde{K} - \tilde{G}(z_0)) = E \times \mathbb{R}^n \times W. \quad (2.2.12)$$

This means that for any $(e, \xi, w) \in E \times \mathbb{R}^n \times W$, there exist $z \in Z$, $v \in \text{cone}((-\infty, 0]^n - H(z_0))$ and $v' \in \text{cone}(K - G(z_0))$ such that

$$DF(z_0)z = e, \quad DH(z_0)z - v = \xi \text{ and } DG(z_0)z - v' = w. \quad (2.2.13)$$

The Robinson constraint qualification guarantees that the necessary optimality conditions are of KKT type. Therefore, we require that (VOP2) satisfies the following assumption.

(A4) $z_0 \in K$ satisfies the Robinson constraint qualification (2.2.12).

Note that if (A4) is satisfied, then $DF(z_0)$ is surjective. In this case (A3) is fulfilled. If we define $\hat{K} = (-\infty, 0]^n \times K$, then $\text{int}(\hat{K}) \neq \emptyset$. According to [26, Corollary 2.101, page 70], the following condition is equivalent to condition (2.2.11).

$$(A5) \quad \begin{cases} (i) DF(z_0) \text{ is surjective} \\ (ii) \exists \tilde{z} \in \text{Ker}(DF(z_0)) \text{ such that } H_i(z_0) + DH_i(z_0)\tilde{z} < 0 \\ \text{and } G(z_0) + DG(z_0)\tilde{z} \in \text{int}(K) \forall i = 1, \dots, n. \end{cases}$$

It is worth noting that (A5) is a variant of the Mangasarian-Fromovitz condition.

Definition 2.2.7. *The vector $z_0 \in \Sigma_2$ is a locally weak Pareto solution of problem (VOP) if there exists $\epsilon > 0$ such that $f(z) - f(z_0) \notin -\text{int}(\mathbb{R}_+^k) \quad \forall z \in B_Z(z_0, \epsilon) \cap \Sigma_2$.*

If $f(z) - f(z_0) \notin -\mathbb{R}_+^k \setminus \{0\}$ for all $z \in B_Z(z_0, \epsilon) \cap \Sigma_2$, we say z_0 is a locally Pareto solution to (VOP).

Obviously, a locally Pareto solution is also a locally weak Pareto solution.

To deal with second-order optimality conditions, we need the so-called critical cone, which is defined as follows. Let $\mathcal{C}_0[z_0]$ be the set of vectors $d \in Z$ satisfying the conditions:

$$(a_1) \quad Df(z_0)d \in -\mathbb{R}_+^k$$

$$(a_2) \quad DF(z_0)d = 0$$

$$(a_3) \quad DH_i(z_0)d \leq 0 \text{ for } i \in I_0, \text{ where } I_0 := \{i \in \{1, 2, \dots, n\} : H_i(z_0) = 0\}$$

$$(a_4) \quad DG(z_0)d \in \text{cone}(K - G(z_0))$$

Then the closure of $\mathcal{C}_0[z_0]$ in Z is called the critical cone at z_0 and denoted by $\mathcal{C}[z_0]$. Each vector $d \in \mathcal{C}[z_0]$ is called a critical direction.

Lemma 2.2.8. *Suppose that $\bar{z} \in \Sigma_2$ is a locally weak Pareto solution of (VOP) under which (A1), (A2) and (A3) are valid, and $DF(z_0)$ is*

surjective. Then for each $d \in \mathcal{C}_0[z_0]$, the system

$$f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2 < 0, \quad \forall i \in \{1, 2, \dots, k\}, \quad (2.2.14)$$

$$H'_i(z_0)z + \frac{1}{2}2H''_i(z_0)d^2 < 0, \quad \forall i \in I_0, \quad (2.2.15)$$

$$DF(z_0)z + \frac{1}{2}D^2F(z_0)(d, d) = 0, \quad (2.2.16)$$

$$DG(z_0)z + \frac{1}{2}D^2G(z_0)(d, d) \in \text{cone}(\text{int}(K) - G(z_0)) \quad (2.2.17)$$

has no solution $z \in Z$.

Proof. On the contrary, we suppose that the system (2.2.14)-(2.2.17) has a solution $z \in Z$. For each $\epsilon > 0$, we define $z_\epsilon = z_0 + \epsilon d + \epsilon^2 z$. By Taylor expansion, we have

$$F(z_\epsilon) = F(z_0) + \epsilon DF(z_0)d + \epsilon^2(DF(z_0)z + \frac{1}{2}D^2F(z_0)(d, d)) + o(\epsilon^2) = o(\epsilon^2).$$

Since $DF(z_0)Z = E$, the Ljusternik theorem (see [108], page 30) implies that, there exist a neighborhood $B(z_0, \gamma)$ of z_0 and a mapping $h : B(z_0, \gamma) \rightarrow Z$ such that for all $\hat{z} \in B(z_0, \gamma)$ we have

$$F(\hat{z} + h(\hat{z})) = 0, \quad \|h(\hat{z})\| \leq C\|F(\hat{z})\|.$$

Taking $\hat{z} = z_\epsilon$ and $h_\epsilon = h(z_\epsilon)$, we see that, there exists $\epsilon_0 > 0$ such that

$$F(z_\epsilon + h_\epsilon) = 0, \quad \|h_\epsilon\| = o(\epsilon^2) \quad \text{for all } 0 < \epsilon < \epsilon_0. \quad (2.2.18)$$

Putting

$$z'_\epsilon = z_\epsilon + h_\epsilon = z_0 + \epsilon d + \epsilon^2(z + \frac{1}{\epsilon^2}h_\epsilon)$$

and using a Taylor's expansion, we get

$$f_i(z'_\epsilon) = f_i(z_0) + \epsilon f'_i(z_0)d + \epsilon^2(f'_i(z_0)z + \frac{1}{2}f''_i(z_0)(d, d) + \frac{o(\epsilon^2)}{\epsilon^2})$$

for $0 < \epsilon < \epsilon_0$. Since

$$f'_i(z_0)z + \frac{1}{2}f''_i(z_0)(d, d) < 0,$$

there exists $\epsilon_i \in (0, \epsilon_0)$ such that for all $\epsilon \in (0, \epsilon_i)$, we have

$$f'_i(z_0)z + \frac{1}{2}f''_i(z_0)(d, d) + \frac{o(\epsilon^2)}{\epsilon^2} < 0.$$

From this and the fact that $f'_i(z_0)d \leq 0$, we have

$$f_i(z'_\epsilon) - f_i(z_0) < 0 \quad \forall \epsilon < \epsilon_i, \quad i = 1, 2, \dots, m.$$

Let $\epsilon' = \min\{\epsilon_1, \dots, \epsilon_m\}$. It follows that

$$f(z'_\epsilon) - f(z_0) \in -\text{int}\mathbb{R}_+^m \quad \text{for } 0 < \epsilon < \epsilon'. \quad (2.2.19)$$

Also by a Taylor's expansion, for each $i \in \{1, 2, \dots, n\}$ we have

$$H_i(z'_\epsilon) = H_i(z_0) + \epsilon H'_i(z_0)d + \epsilon^2(H'_i(z_0)z + \frac{1}{2}H''_i(z_0)(d, d) + \frac{o(\epsilon^2)}{\epsilon^2})$$

for $0 < \epsilon < \epsilon'$. If $i \in I_0$ then

$$H'_i(z_0)z + \frac{1}{2}H''_i(z_0)(d, d) < 0,$$

and $H_i(z_0) = 0$. Therefore, there exists $\epsilon'_i \in (0, \epsilon')$ such that for all $\epsilon \in (0, \epsilon'_i)$ one has

$$H'_i(z_0)z + \frac{1}{2}H''_i(z_0)(d, d) + \frac{o(\epsilon^2)}{\epsilon^2} < 0.$$

Hence

$$H_i(z'_\epsilon) \leq +\epsilon H'_i(z_0)d \leq 0 \quad \forall \epsilon \in (0, \epsilon'_i).$$

If $i \notin I_0$ then $H_i(z_0) < 0$. By the continuity of H_i , there exists $\epsilon''_i < \epsilon'_i$ such that $H_i(z'_\epsilon) < 0$ for all $\epsilon \in (0, \epsilon''_i)$. Hence $H_i(z'_\epsilon) < 0$ for all $i \in \{1, 2, \dots, n\}$ and $\epsilon \in (0, \epsilon'')$ with $\epsilon'' = \min(\epsilon''_1, \dots, \epsilon''_n)$. Besides, (2.2.19) is valid for all $\epsilon \in (0, \epsilon'')$. Using a Taylor's expansion again, we have

$$G(z'_\epsilon) = G(z_0) + \epsilon DG(z_0)d + \epsilon^2(DG(z_0)z + \frac{1}{2}D^2G(z_0)(d, d) + \frac{o(\epsilon^2)}{\epsilon^2}).$$

Since $DG(z_0)d \in \text{cone}(K - G(z_0))$, there exist $\lambda_1 > 0$ and $v_1 \in K$ such that $\nabla G(z_0)d = \lambda_1(v_1 - G(z_0))$ for some $v_1 \in K$. Also, since

$$DG(z_0)z + \frac{1}{2}D^2G(z_0)(d, d) \in \text{cone}(\text{int}(K) - G(z_0)),$$

there exist $\lambda_2 > 0$ and $v_2 \in \text{int}(K)$ such that

$$DG(z_0)z + \frac{1}{2}D^2G(z_0)(d, d) = \lambda_2(v_2 - G(z_0)).$$

Hence,

$$\begin{aligned} G(z'_\epsilon) &= G(z_0) + \epsilon\lambda_1(v_1 - G(z_0)) + \epsilon^2\lambda_2(v_2 - G(z_0)) + o(\epsilon^2) \\ &= (1 - \epsilon\lambda_1 - \epsilon^2\lambda_2)G(z_0) + \epsilon\lambda_1v_1 + \epsilon^2\lambda_2(v_2 + \frac{o(\epsilon^2)}{\lambda_2\epsilon^2}). \end{aligned}$$

Since $v_2 \in \text{int}(K)$, there exists $\epsilon_* \in (0, \epsilon'')$ such that $v_2 + \frac{o(\epsilon^2)}{\lambda_2 \epsilon^2} \in \text{int}(K)$ and $1 - \epsilon\lambda_1 - \epsilon^2\lambda_2 > 0$ for all $\epsilon \in (0, \epsilon_*)$. This implies that

$$\begin{aligned} G(z'_\epsilon) &\in (1 - \epsilon\lambda_1 - \epsilon^2\lambda_2)K + \epsilon\lambda_1 K + \epsilon^2\lambda_2 K \\ &\subset (1 - \epsilon\lambda_1 - \epsilon^2\lambda_2)K + (\epsilon\lambda_1 + \epsilon^2\lambda_2)\left(\frac{\epsilon\lambda_1}{\epsilon\lambda_1 + \epsilon^2\lambda_2}K + \frac{\epsilon^2\lambda_2}{\epsilon\lambda_1 + \epsilon^2\lambda_2}K\right) \\ &\subset (1 - \epsilon\lambda_1 - \epsilon^2\lambda_2)K + (\epsilon\lambda_1 + \epsilon^2\lambda_2)K \subset K. \end{aligned}$$

Combining this with (2.2.18) and the fact $H(z'_\epsilon) \leq 0$, we have $z'_\epsilon \in \Sigma$ for all $\epsilon \in (0, \epsilon_*)$. By this and (2.2.19), we conclude that z_0 is not a locally weak Pareto solution of (VOP) which is absurd. The lemma is proved. $\square$

Let us assume that

$$\mathcal{L}(z, \lambda, e^*, l, w^*) = \lambda^T f(z) + \langle e^*, F(z) \rangle + l^T H(z) + \langle w^*, G(z) \rangle \quad (2.2.20)$$

is the Lagrange function associated with the problem (VOP2), where $\lambda \in \mathbb{R}^k$, $e^* \in E^*$, $l \in \mathbb{R}^n$ and $w^* \in W^*$. We say that vector $(\lambda, e^*, l, w^*) \in \mathbb{R}^k \times E^* \times \mathbb{R}^n \times W^*$ are Lagrange multipliers at z_0 if the following conditions are fulfilled:

$$\begin{aligned} D_z \mathcal{L}(z_0, \lambda, e^*, l, w^*) &= \\ \lambda^T Df(z_0) + DF(z_0)^* e^* + DH(z_0)^* l^T + DG(z_0)^* w^* &= 0, \end{aligned} \quad (2.2.21)$$

$$\lambda \geq 0, \quad (2.2.22)$$

$$l_i \geq 0, \quad l_i H_i(z_0) = 0, \quad i = 1, 2, \dots, n, \quad (2.2.23)$$

$$w^* \in N(K, G(z_0)), \quad (2.2.24)$$

where $l = (l_1, l_2, \dots, l_n)$. We denote by $\Lambda[z_0]$ the set of Lagrange multipliers at z_0 . In addition, if $|\lambda| = 1$, then we say (λ, e^*, l, w^*) are normal. We shall denote by $\Lambda_*[z_0]$ the set of normal Lagrange multipliers at z_0 .

We now have the following result on the necessary first and second order optimality conditions for (VOP2).

Theorem 2.2.9. *Suppose that z_0 is a local weak Pareto solution to (VOP2). Then the following assertions are fulfilled:*

(a) *If (A1) – (A3) are satisfied, then $\Lambda[z_0]$ is nonempty and*

$$\sup_{(\lambda, e^*, l, w^*) \in \Lambda[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)[d, d] \geq 0, \quad \forall d \in \mathcal{C}_0[z_0]. \quad (2.2.25)$$

(b) If (A1), (A2) and (A4) or (A1), (A2) and (A5) are satisfied, then $\Lambda_*[z_0]$ is nonempty and compact in the topology $\sigma(\mathbb{R}^k, \mathbb{R}^k) \times \sigma(E^*, E) \times \sigma(W^*, W)$, and

$$\max_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)[d, d] \geq 0, \quad \forall d \in \mathcal{C}[z_0]. \quad (2.2.26)$$

Proof. (a). We take any $d \in \mathcal{C}_0[z_0]$ and consider two cases.

Case 1. $DF(z_0)Z \neq E$. Then there is a point $e_0 \in E$ and $e_0 \notin DF(z_0)Z$. Since $DF(z_0)Z$ is a closed subspace, the separation theorem (see [209, Theorem 3.4]) implies that there exists a nonzero functional $e^* \in E^*$ which separate e_0 and $DF(z_0)Z$. By a simple argument, we see that $DF(z_0)^* e^* = 0$. Putting $\lambda = 0, l = 0, w^* = 0$, we get $(0, e^*, 0, 0) \in \Lambda(z_0)$. If $e^* DF(z_0)d^2 \leq 0$, then we replace e^* by $-e^*$. Then $(0, -e^*, 0, 0)$ satisfies the conclusion of (a).

Case 2. $DF(z_0)Z = E$.

Let $n_0 = |I_0|$. We define a set S which consists of vectors $(\mu, \gamma, e, w) \in \mathbb{R}^k \times \mathbb{R}^{n_0} \times E \times W$ such that there exists $z \in Z$ satisfying

$$f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2 < \mu_i, i = 1, 2, \dots, k \quad (2.2.27)$$

$$H'_j(z_0)z + \frac{1}{2}H''_j(z_0)d^2 < \gamma_j, j \in I_0, \quad (2.2.28)$$

$$DF(z_0)z + \frac{1}{2}D^2F(z_0)d^2 = e \quad (2.2.29)$$

$$DG(z_0)z + \frac{1}{2}D^2G(z_0)d^2 - w \in \text{cone}(\text{int}(K) - G(z_0)). \quad (2.2.30)$$

It is clear that S is convex. We now show that S is open. In fact, take $(\hat{\mu}, \hat{\gamma}, \hat{e}, \hat{w}) \in S$ corresponding to $\hat{z}$. Choose $\epsilon > 0$ such that

$$f'_i(z_0)\hat{z} + \frac{1}{2}f''_i(z_0)d^2 < \hat{\mu}_i - \epsilon, \quad H'_j(z_0)\hat{z} + \frac{1}{2}H''_j(z_0)d^2 < \hat{\gamma}_j - \epsilon$$

for all $i = 1, 2, \dots, k$ and $j \in I_0$. Then there is $\delta > 0$ such that for all $z \in B(\hat{z}, \delta)$ one has

$$f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2 < \hat{\mu}_i - \epsilon, \quad H'_j(z_0)z + \frac{1}{2}H''_j(z_0)d^2 < \hat{\gamma}_j - \epsilon$$

for all $i = 1, 2, \dots, k$ and $j \in I_0$. It follows that

$$\begin{aligned} f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2 &< \mu_i, \quad \forall z \in B(\hat{z}, \delta) \text{ and } |\mu_i - \hat{\mu}_i| < \epsilon, i = 1, 2, \dots, k, \\ H'_j(z_0)z + \frac{1}{2}H''_j(z_0)d^2 &< \gamma_j, \quad \forall z \in B(\hat{z}, \delta) \text{ and } |\gamma_j - \hat{\gamma}_j| < \epsilon, j \in I_0. \end{aligned}$$

Since $\text{cone}(\text{int}K - G(z_0))$ is an open convex cone and

$$DG(z_0)\hat{z} + \frac{1}{2}D^2G(z_0)d^2 - \hat{w} \in \text{cone}(\text{int}(K) - G(z_0)),$$

the continuity of $DG(z_0)$ implies that there exist balls $B(\hat{z}, \delta)$ and $B(\hat{w}, r)$ such that

$$DG(z_0)z + \frac{1}{2}D^2G(z_0)d^2 - w \in \text{cone}(\text{int}(K) - G(z_0)) \quad \forall (z, w) \in B(\hat{z}, \delta) \times B(\hat{w}, r).$$

Since $DF(z_0)$ is surjective, $DF(z_0)[B(\hat{z}, \delta)] + \frac{1}{2}D^2F(z_0)d^2$ is open. Hence, there exists a number $\alpha > 0$ such that

$$B(\hat{e}, \alpha) \subset DF(z_0)[B(\hat{z}, \delta)] + \frac{1}{2}D^2G(z_0)d^2.$$

Thus, for any $(\mu, \gamma, e, w) \in B_{\mathbb{R}^k}(\hat{\mu}, \epsilon) \times B_{\mathbb{R}^{n_0}}(\hat{\gamma}, \epsilon) \times B_E(\hat{e}, \alpha) \times B_W(\hat{w}, r)$, there exists $z \in B(\hat{z}, \delta)$ such that

$$\left\{ \begin{array}{l} f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2 < \mu_i, \quad H'_j(z_0)z + \frac{1}{2}H''_j(z_0)d^2 < \gamma_j, \\ e = DF(z_0)z + \frac{1}{2}D^2F(z_0)d^2, \\ DG(\bar{z})z + \frac{1}{2}D^2G(z_0)d^2 - w \in \text{cone}(\text{int}(K) - G(\bar{z})), \\ \forall i = 1, 2, \dots, k; \forall j \in I_0. \end{array} \right.$$

This means that

$$B_{\mathbb{R}^k}(\hat{\mu}, \epsilon) \times B_{\mathbb{R}^{n_0}}(\hat{\gamma}, \epsilon) \times B_E(\hat{e}, \alpha) \times B_W(\hat{w}, r) \subset S.$$

Hence S is open. By Lemma 2.2.8, we have $(0, 0, 0, 0) \notin S$. By the separation theorem (see [108, Theorem 1, p. 163], there exists a nonzero vector $(\lambda, l, e^*, w^*) \in \mathbb{R}^k \times \mathbb{R}^{n_0} \times E^* \times W^*$ such that

$$\lambda\mu^T + l\gamma\langle e^*, e \rangle + \langle w^*, w \rangle \geq 0 \quad \forall (\mu, \gamma, e, w) \in S. \quad (2.2.31)$$

Fix any $z \in Z$, $w' \in \text{cone}(\text{int}K - G(z_0))$, $r_i > 0$ and $r'_j > 0$. Set

$$\begin{aligned} \mu_i &= r_i + f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2, \quad \gamma_j = r'_j + H'_j(z_0)z + \frac{1}{2}H''_j(z_0)d^2, \\ e &= DF(z_0)z + \frac{1}{2}D^2F(z_0)d^2, \quad w = DG(z_0)z + \frac{1}{2}D^2G(z_0)d^2 - w' \end{aligned}$$

Then $(\mu, \gamma, e, w) \in S$. From this and (2.2.31), we get

$$\begin{aligned} & \sum_{i=1}^k \lambda_i(r_i + f'_i(z_0)z + \frac{1}{2}f''_i(z_0)d^2) + \sum_{j \in I_0} l_j(r'_j + H'_j(z_0)z + \frac{1}{2}H''_j(z_0)d^2) \\ & + \langle e^*, DF(z_0)z + \frac{1}{2}D^2F(z_0)d^2 \rangle + \langle w^*, DG(z_0)z + \frac{1}{2}D^2G(z_0)d^2 \rangle \\ & - \langle w^*, w' \rangle \geq 0. \end{aligned}$$

If there exists i_0 such that $\lambda_{i_0} < 0$ then by letting $r_{i_0} \rightarrow +\infty$, the term on the left hand side approach to $-\infty$ which is impossible. Hence we must have $\lambda_i \geq 0$ for all $i = 1, 2, \dots, k$. Similarly, we have $l_j \geq 0$ for all $j \in I_0$.

By letting $r_i \rightarrow 0$ and $r'_j \rightarrow 0$, we get

$$\sum_{i=1}^k \lambda_i f'_i(z_0) + \sum_{j \in I_0} l_j H'_j(z_0) + DF(z_0)^* e^* + DG(z_0)^* w^* = 0 \quad (2.2.32)$$

and

$$\begin{aligned} & \frac{1}{2} \left(\sum_{i=1}^k \lambda_i f''_i(z_0) d^2 + \sum_{j \in I_0} l_j H''_j(z_0) d^2 + \langle e^*, D^2 F(z_0) d^2 \rangle + \langle w^*, D^2 G(z_0) d^2 \rangle \right) \\ & \geq \langle w^*, w' \rangle \quad \forall w' \in \text{cone}(\text{int}(K) - G(z_0)). \end{aligned}$$

It follows that

$$\langle w^*, w' \rangle \leq 0, \forall w' \in \text{cone}(\text{int}(K) - G(z_0)).$$

This implies that $w^* \in N(K, G(z_0))$ because $K \subseteq \bar{K} = \overline{\text{int}(K)}$. By letting $w' \rightarrow 0$, we get

$$\sum_{i=1}^k \lambda_i f''_i(z_0) d^2 + \sum_{j \in I_0} l_j H''_j(z_0) d^2 + \langle e^*, D^2 F(z_0) d^2 \rangle + \langle w^*, D^2 G(z_0) d^2 \rangle \geq 0. \quad (2.2.33)$$

We now take $l_j = 0$ for $j \in \{1, 2, \dots, n\} \setminus I_0$. Then we have $l_j \geq 0$ and $l_j H_j(z_0) = 0$ for all $j = 1, 2, \dots, n$ and (2.2.32) and (2.2.33) become

$$\sum_{i=1}^k \lambda_i f'_i(z_0) + \sum_{j=1}^n l_j H'_j(z_0) + DF(z_0)^* e^* + DG(z_0)^* w^* = 0 \quad (2.2.34)$$

and

$$\sum_{i=1}^k \lambda_i f''_i(z_0) d^2 + \sum_{j=1}^n l_j H''_j(z_0) d^2 + \langle e^*, D^2 F(z_0) d^2 \rangle + \langle w^*, D^2 G(z_0) d^2 \rangle \geq 0. \quad (2.2.35)$$

Hence assertion (a) is proved.

(b). We claim that $\lambda \neq 0$. For this we consider two cases.

Case 1. (A1), (A2) and (A4) are satisfied.

If $\lambda = 0$, then we have

$$DH(z_0)^* l^T + DF(z_0)^* e^* + DG(z_0)^* w^* = 0.$$

Take any $\xi \in \mathbb{R}^n$, $e \in E$ and $w \in W$. By (A4), there exist $z \in Z$, $v \in \text{cone}((-\infty, 0]^n - H(z_0))$ and $v' \in \text{cone}(K - G(z_0))$ satisfying (2.2.13). From the above, we have

$$\begin{aligned} 0 &= l^T DH(z_0)z + \langle e^*, DF(z_0)z \rangle + \langle w^*, DG(z_0)z \rangle \\ &= l^T \xi + l^T v + \langle e^*, e \rangle + \langle w^*, w \rangle + \langle w^*, v' \rangle \\ &\leq l^T \xi + \langle e^*, e \rangle + \langle w^*, w \rangle. \end{aligned}$$

Here we used the fact that $w^* \in N(K, G(z_0))$ and $l \in N((-\infty, 0]^n, H(z_0))$. As ξ, e, w are arbitrary, we obtain $l = 0$, $e^* = 0$ and $w^* = 0$ which is impossible. Hence we must have $\lambda \neq 0$ and so $\Lambda_*[z_0] \neq \emptyset$.

Case 2. (A1), (A2) and (A5) are fulfilled.

If $\lambda = 0$, then we have

$$DH(z_0)^* l^T + DF(z_0)^* e^* + DG(z_0)^* w^* = 0.$$

Let $\tilde{z}$ be a vector satisfying (A5). Then we have

$$l^T DH(z_0)\tilde{z} + \langle e^*, DF(z_0)\tilde{z} \rangle + \langle w^*, DG(z_0)\tilde{z} \rangle = 0.$$

This implies that

$$l^T DH(z_0)\tilde{z} + \langle w^*, DG(z_0)\tilde{z} \rangle = 0,$$

where

$$DH(z_0)\tilde{z} \in (-\infty, 0]^n - H(z_0) = \text{int}((-\infty, 0]^n - H(z_0))$$

and

$$DG(z_0)\tilde{z} \in \text{int}(K) - G(z_0) = \text{int}(K - G(z_0)).$$

Fixing any $\xi \in \mathbb{R}^n$ and $w \in W$, we see that, there exist $\tau > 0$ and $s > 0$ small enough such that

$$\tau\xi + DH(z_0)\tilde{z} \in (-\infty, 0]^n - H(z_0) \text{ and } sw + DG(z_0)\tilde{z} \in K - G(z_0).$$

Since $l \geq 0$, $l^T H(z_0) = 0$ and $w^* \in N(K, G(z_0))$, we have

$$\tau l^T \xi + s \langle w^*, w \rangle = l^T (\tau\xi + DH(z_0)\tilde{z}) + \langle w^*, sw + DG(z_0)\tilde{z} \rangle \leq 0.$$

Hence $l^T \xi \leq 0$ for all $\xi \in \mathbb{R}^n$ and $\langle w^*, w \rangle \leq 0$ for all $w \in W$. This implies that $l = 0$ and $w^* = 0$. Consequently, $DF(z_0)^* e^* = 0$. Since $DF(z_0)$

is surjective, we get $e^* = 0$. Hence $(\lambda, e^*, l, w^*) = (0, 0, 0, 0)$ which is absurd. The claim is justified.

The boundedness of $\Lambda_*[z_0]$ is proved analogously with Lemma 2.2.3. It remains to prove (2.2.26).

Let $d \in \mathcal{C}[z_0]$. Then there exists a sequence $(d_j) \subset \mathcal{C}_0[z_0]$ such that $d_j \rightarrow d$. By assertion (a), there exists $(\lambda_j, e_j^*, l_j, w_j^*) \in \Lambda_*[z_0]$ such that

$$D_z^2 \mathcal{L}(z_0, \lambda_j, e_j^*, l_j, w_j^*)(d_j, d_j) \geq 0.$$

By [37, Theorem 13.13, p.207] (see also [87, Theorem 4.29, p. 129] and [78, Theorem 1.3, Chapter XI, p. 223]), a set in a topological space is compact if and only if any net in it, has a cluster point (also called accumulation point). Since $\Lambda_*[z_0]$ is compact in the topology $\tau(\mathbb{R}^k) \times \sigma(E^*, E) \times \tau(\mathbb{R}^m) \times \sigma(W^*, W)$, there exists $(\lambda_0, e_0^*, l_0, w_0^*) \in \Lambda_*[z_0]$ such that $(\lambda_j, e_j^*, l_j, w_j^*)$ accumulates to $(\lambda_0, e_0^*, l_0, w_0^*)$. This means that, for any neighborhood $V_0(\lambda_0) \times V_1(e_0^*) \times V_2(l_0) \times V_3(w_0^*)$ in the topology $\tau(\mathbb{R}^k) \times \sigma(E^*, E) \times \tau(\mathbb{R}^m) \times \sigma(W^*, W)$ and for any $s \geq 1$, there exists a number $j_s \geq s$ such that

$$(\lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*) \in V_0(\lambda_0) \times V_1(e_0^*) \times V_2(l_0) \times V_3(w_0^*).$$

Recall that

$$\begin{aligned} D_z^2 \mathcal{L}(z_0, \lambda_j, e_j^*, l_j, w_j^*)(d_j, d_j) &= \lambda_j^T f''(z_0)[d_j, d_j] + e_j^* D_z^2 F(z_0)[d_j, d_j] \\ &+ l_j^T D_z^2 H(z_0)[d_j, d_j] + w_j^* D_z^2 G(z_0)[d_j, d_j]. \end{aligned}$$

Define

$$\begin{aligned} r_j &= f''(z_0)[d_j, d_j], \quad e_j = D_z^2 F(z_0)[d_j, d_j], \\ \xi_j &= D_z^2 H(z_0)[d_j, d_j], \quad w_j = D_z^2 G(z_0)[d_j, d_j]. \end{aligned}$$

Then we have the following strong convergences

$$\begin{aligned} r_j &\rightarrow f''(z_0)[d, d] := r \\ e_j &\rightarrow D_z^2 F(z_0)[d, d] := e, \\ \xi_j &\rightarrow D_z^2 H(z_0)[d, d] := \xi, \\ w_j &\rightarrow D_z^2 G(z_0)[d, d] := w \end{aligned}$$

when $k \rightarrow \infty$. Let us set

$$V_s(\lambda_0) = \{\lambda \in \mathbb{R}^k : |\lambda - \lambda_0| < \frac{1}{s}\},$$

$$V_s(e_0^*) = \{e^* \in E^* : |\langle e^* - e_0^*, e \rangle| < \frac{1}{s}\},$$

$$V_s(l_0) = \{l \in \mathbb{R}^m : |l - l_0| < \frac{1}{s}\},$$

$$V_s(w_0^*) = \{w^* \in W^* : |\langle w^* - w_0^*, w \rangle| < \frac{1}{s}\},$$

where e, w are as above and $s \geq 1$. Then $V_s(\lambda_0) \times V_s(e_0^*) \times V_s(l_0) \times V_s(w_0^*)$ is a neighborhood of $(\lambda_0, e_0^*, l_0, w_0^*)$ in the topology $\tau(\mathbb{R}^k) \times \sigma(E^*, E) \times \tau(\mathbb{R}^m) \times \sigma(W^*, W)$. Therefore, for each $s \geq 1$, there exists $j_s \geq s$ such that

$$(\lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*) \in V_s(\lambda_0) \times V_s(e_0^*) \times V_s(l_0) \times V_s(w_0^*).$$

This implies that

$$|\lambda_{j_s} - \lambda_0| < \frac{1}{s}, |\langle e_{j_s}^* - e_0^*, e \rangle| < \frac{1}{s}, \quad (2.2.36)$$

$$|l_{j_s} - l_0| < \frac{1}{s}, \quad |\langle w_{j_s}^* - w_0^*, w \rangle| < \frac{1}{s}, \quad \forall s \geq 1. \quad (2.2.37)$$

We now have

$$\begin{aligned} & D_z^2 \mathcal{L}(z_0, \lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*)(d_{j_s}, d_{j_s}) - D_z^2 \mathcal{L}(z_0, \lambda_0, e_0^*, l_0, w_0^*)(d, d) \\ &= \lambda_{j_s}^T r_{j_s} - \lambda_0^T r + \langle e_{j_s}^*, e_{j_s} \rangle - \langle e_0^*, e \rangle + l_{j_s}^T \xi_{j_s} - l_0^T \xi + \langle w_{k_n}^*, w_{k_n} \rangle - \langle w_0^*, w \rangle \\ &= \lambda_{j_s}^T (r_{j_s} - r) + (\lambda_{j_s} - \lambda_0)^T r + \langle e_{j_s}^*, e_{j_s} - e \rangle + \langle e_{j_s}^* - e_0^*, e \rangle \\ &+ l_{j_s}^T (\xi_{j_s} - \xi) + (l_{j_s} - l_0)^T \xi + \langle w_{k_n}^*, w_{k_n} - w \rangle + \langle w_{k_n}^* - w_0^*, w \rangle. \end{aligned}$$

From this and (2.2.36) and 2.2.37, we get

$$\begin{aligned} & |D_z^2 \mathcal{L}(z_0, 1, e_{j_s}^*, l_{j_s}, w_{j_s}^*)(d_{j_s}, d_{j_s}) - D_z^2 \mathcal{L}(z_0, 1, e_0^*, l_0, w_0^*)(d, d)| \\ &\leq |\lambda_{j_s}| |r_{j_s} - r| + |\lambda_{j_s} - \lambda_0| |r| + \|e_{j_s}^*\| \|e_{j_s} - e\| + |\langle e_{j_s}^* - e_0^*, e \rangle| \\ &+ |l_{j_s}| \|\xi_{j_s} - \xi\| + |l_{j_s} - l_0| \|\xi\| + \|w_{j_s}^*\| \|w_{j_s} - w\| + |\langle w_{j_s}^* - w_0^*, w \rangle| \\ &\leq |\lambda_{j_s}| |r_{j_s} - r| + \frac{1}{s} |r| + \|e_{j_s}^*\| \|e_{j_s} - e\| + \frac{1}{s} \\ &+ |l_{j_s}| \|\xi_{j_s} - \xi\| + \frac{1}{s} \|\xi\| + \|w_{j_s}^*\| \|w_{j_s} - w\| + \frac{1}{s} \rightarrow 0 \quad \text{as } s \rightarrow \infty. \end{aligned}$$

Here we used the fact that the sequence $\{(\lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*)\}$ is bounded. It follows that

$$\lim_{n \rightarrow \infty} D_z^2 \mathcal{L}(z_0, \lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*)(d_{j_s}, d_{j_s}) = D_z^2 \mathcal{L}(z_0, \lambda_0, e_0^*, l_0, w_0^*)(d, d).$$

Since

$$D_z^2 \mathcal{L}(z_0, \lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*)(d_{j_s}, d_{j_s}) \geq 0,$$

we obtain

$$D_z^2 \mathcal{L}(z_0, \lambda_0, e_0^*, l_0, w_0^*)(d, d) \geq 0.$$

This also implies that

$$\sup_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d) \geq 0.$$

It remains to show that for each $d \in \mathcal{C}[z_0]$,

$$\begin{aligned} & \sup_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d) \\ &= \max_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d). \end{aligned}$$

Since $\Lambda_*[z_0]$ is bounded, then there exists a number $M > 0$ such that

$$\|(\lambda, e^*, l, w^*)\| \leq M, \quad \forall (\lambda, e^*, l, w^*) \in \Lambda_*[z_0].$$

Therefore, we have

$$\begin{aligned} & |D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d)| = \\ & |\lambda^T f''(z_0)[d, d] + e^* D_z^2 F(z_0)[d, d] + l^T D_z^2 H(z_0)[d, d] + w^* D_z^2 G(z_0)[d, d]| \\ & \leq M(\|f''(z_0)[d, d]\| + \|D_z^2 F(z_0)[d, d]\| + \|D_z^2 H(z_0)[d, d]\| \\ & \quad + \|D_z^2 G(z_0)[d, d]\|) \end{aligned}$$

for all $d \in \mathcal{C}[z_0]$. Consequently, for each $d \in \mathcal{C}[z_0]$, we have

$$\sup_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d) < +\infty.$$

Fix $d \in \mathcal{C}[z_0]$. By definition of supremum, there exists a sequence $(\lambda_j, e_j^*, l_j, w_j^*) \in \Lambda_*[z_0]$ such that

$$\sup_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d) = \lim_{j \rightarrow \infty} D_z^2 \mathcal{L}(z_0, \lambda_j, e_j^*, l_j, w_j^*)(d, d).$$

Since $\Lambda_*[z_0]$ is compact, $(\lambda_j, e_j^*, l_j, w_j^*)$ has an accumulation point

$$(\hat{\lambda}, \hat{e}^*, \hat{l}, \hat{w}^*) \in \Lambda_*[z_0].$$

By the same argument as the above, we see that, there exists a subsequence $(\lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*)$ such that

$$\begin{aligned} & \sup_{(\lambda, e^*, l, w^*) \in \Lambda_*[z_0]} D_z^2 \mathcal{L}(z_0, \lambda, e^*, l, w^*)(d, d) = \lim_{j \rightarrow \infty} D_z^2 \mathcal{L}(z_0, \lambda_j, e_j^*, l_j, w_j^*)(d, d) \\ &= \lim_{s \rightarrow \infty} D_z^2 \mathcal{L}(z_0, \lambda_{j_s}, e_{j_s}^*, l_{j_s}, w_{j_s}^*)(d, d) = D_z^2 \mathcal{L}(z_0, \hat{\lambda}, \hat{e}^*, \hat{l}, \hat{w}^*)(d, d). \end{aligned}$$

The proof of Theorem 2.2.9 is completed. $\square$

Remark 2.2.10. In comparison with the obtained results in [21], Theorem 2.2.9 is an extension of [21, Theorem 8.2]. In [21] the author considered problem (VOP2) without constraint $H(z) \leq 0$ and required that K is a cone. Here we consider problem (VOP2) with constraint $H(z) \leq 0$ and K is a convex set. Besides, our techniques for the proof of Theorem 2.2.9 is very different from those in [21].

CHAPTER 3

OPTIMAL CONTROL PROBLEMS GOVERNED BY ODE

This chapter presents some results for optimal control problems governed by ordinary differential equations. Its content is based on our published papers [132], [138] and [2].

3.1 Existence of solutions to a class of optimal control problems

3.1.1 Introduction

In this section we consider the following optimal control problem of finding a control function $u \in L^\infty([0, 1], \mathbb{R}^m)$ and the corresponding state $x \in W^{1,\infty}([0, 1], \mathbb{R}^n)$ which solve

$$J(x, u) = \int_0^1 L(t, x(t), u(t))dt \rightarrow \min \quad (3.1.1)$$

s.t.

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(0) = x_0, \quad (3.1.2)$$

$$x(1) \in D, \quad (3.1.3)$$

$$x(t) \in X \text{ a.e. } t \in [0, 1], \quad (3.1.4)$$

$$u(t) \in U \text{ a.e. } t \in [0, 1], \quad (3.1.5)$$

where $x_0 \in \mathbb{R}^n$, D and X are nonempty closed sets in $\mathbb{R}^n$, U is nonempty and compact set in $\mathbb{R}^m$, $L : [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ and $f : [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ are given mappings. From now on we denote by Φ the feasible set of problem (3.1.1)-(3.1.4), that is, Φ consists of couples $(x, u) \in W^{1,\infty}([0, 1], \mathbb{R}^n) \times L^\infty([0, 1], \mathbb{R}^m)$ which satisfies constraints (3.1.2)-(3.1.5). Our main concern is to prove that under certain conditions, there exists a control function $\bar{u} \in U$ which steers x_0 to a point $x_1 \in D$ in an optimal way.

The study topic on the existence of solution to optimal control problems studied deeply in [51], [86], [108], [208] and references given therein. However, in [51], the authors mainly studied problems where the state

equations are linear in control variable u . Meanwhile, Roxin [208] and Filippov [86] studied the time-optimal control problem $T \rightarrow \min$. In [208], Roxin required that if U is compact, the set $f(t, x, U)$ is convex and $f(t, \cdot, u)$ is global Lipschitz in x for all $(t, u) \in I \times U$, then the time-optimal control problem has a solution. It is noted that the global Lipschitz condition is quite strictly because the functions $f(t, x, u)$ which are polynomial of x , fails to apply except that $f(t, x, u)$ is linear in x . More general, in [86] Fillippov showed that if for each $(t, x) \in I \times \mathbb{R}^n$, $U(t, x)$ is a compact set in $\mathbb{R}^m$ and the set

$$R(t, x) = \{f(t, x, u) | u \in U(t, x)\}$$

is convex, and that

$$x \cdot f(t, x, u) \leq C(|x|^2 + 1) \quad \forall (t, x) \in I \times \mathbb{R}^n, u \in U(t, x)$$

for some constant $C > 0$, then the time-optimal control problem has an optimal solution. Thus the convexity of the set $f(t, x, U)$ and the compactness of U are important for the existence of solutions to optimal control. In this section we will develop this idea to prove the existence of solution to problem (3.1.1)-(3.1.5), where the state equation may be nonlinear in both variable x and u .

Let us talk about the idea of proofs for the case where $f(t, x, u)$ is linear in u . Let $(x_k, u_k) \in \Phi$ be a minimal sequence. We somehow show that $\|u_k\|_p$ is bounded and so $u_k \rightharpoonup \bar{u}$ and $x_k \rightarrow \bar{x}$. Then we try to show that $f(\cdot, x_k, u_k)$ converges weakly to $f(\cdot, \bar{x}, \bar{u})$. By taking the limit both sides of the state equation $\dot{x}_k = f(t, x_k, u_k)$, we get $\dot{\bar{x}} = f(t, \bar{x}, \bar{u})$ and so $(\bar{x}, \bar{u}) \in \Phi$, which is an optimal solution. Unfortunately, when f is nonlinear in u , the above standard argument may not be correct. For an example when $n = 1$ and $f(t, x, u) = x + u^2$. It is known that the sequence $u_k(t) = \sin(k\pi t) \rightharpoonup 0$ in $L^2([0, 1], \mathbb{R})$ as $k \rightarrow \infty$ but $u_k^2(t)$ does not converges weakly to 0 in $L^2([0, 1], \mathbb{R})$. This motivate us to study the existence of solutions for optimal control problems where the state equations are nonlinear in control variable. For this case in [108, Theorem 3, p. 373], the authors showed that if the following condition

(H) *The function*

$$h(t, x, y) := \inf\{L(t, x, u) | u \in U, f(t, x, u) = y\} \quad \text{is convex in } y$$

is valid and some addition assumptions are satisfied, then problem (3.1.1)-(3.1.5) has an optimal solution $(\bar{x}, \bar{u}) \in W^{1,1}([0, T], \mathbb{R}^n) \times L^\infty([0, 1], \mathbb{R}^m)$.

In this section we shall characterize assumption (H) and give a result on the existence of optimal solutions to problem (3.1.1)-(3.1.5) under new assumptions (see assumption (A3)). To obtain the result, we shall provide a new proof which is based on the measurable selection theorem in order to build optimal control function. It is worth noting that in our obtained result, the assumptions are easy to verify and the regularity of optimal solutions is also given, that is, $\bar{x} \in W^{1,\infty}([0, 1], \mathbb{R}^n)$.

3.1.2 Main result

Let $F : \mathbb{R}^k \rightarrow 2^{\mathbb{R}^l}$ be a multifunction. If F has closed images and satisfies the property that for any open set O of $\mathbb{R}^l$, the set

$$F^{-1}(O) = \{w \in \mathbb{R}^k | F(w) \cap O \neq \emptyset\}$$

is measurable, then we say F is measurable. A single-valued mapping $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^l$ is said to be a selection of F if $\varphi(w) \in F(w)$ for all $w \in \mathbb{R}^k$.

In the sequel, we shall need the following generalization of the Gronwall inequality.

Lemma 3.1.1. *Suppose $v(t) \geq 0$ and $w(t) \geq 0$ for all $t \geq 0$, v is continuous and w is integrable. Let $v_0 \geq 0$ and $\alpha \geq 0$ such that*

$$v(t) \leq v_0 + \int_0^t w(s)v(s)^\alpha ds. \quad (3.1.6)$$

Then the following assertion are valid.

(i) *If $\alpha = 1$, then*

$$v(t) \leq v_0 \exp\left(\int_0^t w(s)ds\right) \quad \forall t \geq 0. \quad (3.1.7)$$

(ii) *If $0 \leq \alpha < 1$ then*

$$v(t) \leq [v_0^{1-\alpha} + (1-\alpha) \int_0^t w(s)ds]^{\frac{1}{1-\alpha}} \quad (3.1.8)$$

Proof. Assertion (i) is only the Gronwall inequality. Let us prove assertion (ii). For this we define

$$\psi(t) = v_0 + \int_0^t w(s)v(s)^\alpha ds.$$

Then we have

$$\psi'(t) = w(t)v(t)^\alpha \leq w(t)\psi(t)^\alpha.$$

By [144, Theorem 1], ψ is bounded by a maximal solution of the equation

$$r'(t) = w(t)r(t)^\alpha, \quad r(0) = v_0.$$

This is a Bernoulli equation. To solve it we put $z(t) = r^{1-\alpha}$. Then z satisfies equation

$$z'(t) = (1 - \alpha)w(t), \quad z(0) = v_0^{1-\alpha}$$

By formula of solutions of Bernoulli equations (see [142, Section 1.5, p. 27]), we have

$$r(t)^{1-\alpha} = z(t) = [v_0^{1-\alpha} + (1 - \alpha) \int_0^t w(s)ds].$$

It follows that

$$v(t) \leq \psi(t) \leq r(t) \leq [v_0^{1-\alpha} + (1 - \alpha) \int_0^t w(s)ds]^{\frac{1}{1-\alpha}}.$$

The proof of the lemma is complete. $\square$

We will make the following assumptions:

(A1) $L : [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ is a Carathéodory function and there exist a function $\phi_0 \in L^1([0, 1], \mathbb{R})$, non-negative constants α_0, α_1 and real numbers c_0, c_1 such that

$$\phi_0(t) + c_0|x|^{\alpha_0} + c_1|u|^{\alpha_1} \leq L(t, x, u) \quad \forall (t, x, u) \in [0, 1] \times \mathbb{R}^n \times U. \quad (3.1.9)$$

Moreover, for each $M > 0$ there exists a number $k_{L,M} > 0$ such that

$$|L(t, x, u)| \leq K_{L,M} \quad (3.1.10)$$

for a.e. $t \in [0, 1]$, $|x| \leq M$ and $u \in U$.

(A2) $f : [0, 1] \times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ is continuous and there exist a function $r(\cdot) \in L^\infty([0, 1], \mathbb{R})$ and a number $k_f > 0$ such that

$$|f(t, x, u)| \leq r(t) + k_f|x|^\alpha \quad (3.1.11)$$

for a.e. $t \in [0, 1]$, for all $x \in \mathbb{R}^n$ and $u \in U$. Moreover, we require that if $\alpha > 1$, then $\alpha \leq \alpha_0$ and $c_0 > 0$ under which (3.1.9) is satisfied.

(A3) For each $(t, x) \in [0, 1] \times \mathbb{R}^n$, the mapping $L(t, x, \cdot)$ and $f(t, x, \cdot)$ satisfy the convex condition: for any $u, v \in U$ and $\lambda \in [0, 1]$, there exists $w \in U$ such that

$$\begin{aligned} L(t, x, w) &\leq \lambda L(t, x, u) + (1 - \lambda)L(t, x, v) \\ f(t, x, w) &= \lambda f(t, x, u) + (1 - \lambda)f(t, x, v). \end{aligned}$$

Among the above assumptions, (A3) is important for the existence of optimal solutions to problem (3.1.1)-(3.1.5). A variant of (A3) was also used by [51, Theorem 9.3.i] in order to establish optimality conditions. It turn out that (A3) is equivalent to (H) by the following characterization.

Lemma 3.1.2. *Suppose that U is compact. Then assumption (H) is equivalent to assumption (A3).*

Proof. Suppose that assumption (H) is valid. We take any $u_1, u_2 \in U$ and put $y_i = f(t, x, u_i)$ with $i = 1, 2$. Since $h(t, x, y)$ is convex in y , we have

$$h(t, x, \lambda y_1 + (1 - \lambda)y_2) \leq \lambda h(t, x, y_1) + (1 - \lambda)h(t, x, y_2) \quad \forall \lambda \in [0, 1].$$

Since U is compact and $f(t, x, \cdot), L(t, x, \cdot)$ are continuous, there exists $w \in U$ such that $f(t, x, w) = \lambda f(t, x, u_1) + (1 - \lambda)f(t, x, u_2)$ and

$$\begin{aligned} L(t, x, w) = h(t, x, \lambda y_1 + (1 - \lambda)y_2) &\leq \lambda h(t, x, y_1) + (1 - \lambda)h(t, x, y_2) \\ &\leq \lambda L(t, x, u_1) + (1 - \lambda)L(t, x, u_2). \end{aligned}$$

Hence (A3) is followed.

Conversely, suppose that (A3) is valid. Take any $y_1, y_2 \in \mathbb{R}^n$. By the compactness of U and continuity of $L(t, x, \cdot)$ and $f(t, x, \cdot)$, there exist $u_i \in U$ and $f(t, x, u_i) = y_i$ such that $h(t, x, y_i) = L(t, x, u_i)$ with $i = 1, 2$. By (A3), there exists $w \in U$ such that

$$f(t, x, w) = \lambda f(t, x, u_1) + (1 - \lambda)f(t, x, u_2) = \lambda y_1 + (1 - \lambda)y_2$$

and

$$L(t, x, w) \leq \lambda L(t, x, u_1) + (1 - \lambda)L(t, x, u_2)$$

This implies that

$$\begin{aligned} h(t, x, \lambda y_1 + (1 - \lambda)y_2) &\leq L(t, x, w) \leq \lambda L(t, x, u_1) + (1 - \lambda)L(t, x, u_2) \\ &= \lambda h(t, x, y_1) + (1 - \lambda)h(t, x, y_2). \end{aligned}$$

Hence (H) is satisfied. The proof of the lemma is complete. $\square$

Let us give an example for L and f satisfying (A3).

Example 3.1.3. (i) $L(t, x, u)$ is convex w.r.t. u and

$$f(t, x, u) = A(t, x) + B(t, x)u$$

and U is convex. In this case for any $u, v \in U$ and $\lambda \in [0, 1]$, we choose $w = \lambda u + (1 - \lambda)v$.

(ii) Let $m = n = 1$ and $L(t, x, u) = \varphi(t, x) + u^6$, $f(t, x, u) = A(t, x) + u^3$ and $U = [a, b]$. Then for any $u, v \in [a, b]$ and $\lambda \in [0, 1]$, we choose $w = (\lambda u^3 + (1 - \lambda)v^3)^{1/3}$.

We now have the following result on the existence of optimal solutions. The idea of its proof is to use measurable selections of set-valued maps in order to build an optimal control function.

Theorem 3.1.4. *Suppose that assumptions (A1)–(A3) are satisfied and there is $(x_*, u_*) \in \Phi$ such that $J(x_*, u_*) < +\infty$. Then problem (3.1.1)–(3.1.5) has at least an optimal solution.*

Proof. Let us set $V = \inf_{(x,u) \in \Phi} J(x, u)$. By (A1) and (A2), we have $-\infty < V \leq J(x_*, u_*) < +\infty$. Let $(x_k, u_k) \in \Phi$ be a sequence so that $V = \lim_{k \rightarrow \infty} J(x_k, u_k)$. By (3.1.9), there exists $k_0 > 0$ such that

$$\int_0^1 \phi_0(t) dt + c_0 \int_0^1 |x_k(t)|^{\alpha_0} dt + c_1 \int_0^1 |u_k(t)|^{\alpha_1} dt < V + 1 \quad (3.1.12)$$

for all $k \geq k_0$. Hence if $c_0 > 0$, then there exists a number $M_1 > 0$ such that

$$\int_0^1 |x_k(t)|^{\alpha_0} dt \leq M_0 \quad \forall k \geq k_0. \quad (3.1.13)$$

Let us show that $\{x_k\}$ is uniformly equicontinuous and $x_k \rightarrow \bar{x}$ in norm of $C([0, 1], \mathbb{R}^n)$ for some $\bar{x} \in C([0, 1], \mathbb{R})$. Since $(x_k, u_k) \in \Phi$, they must satisfy equation (3.1.2), that is

$$x_k(t) = x_0 + \int_0^t f(s, x_k(s), u_k(s)) ds. \quad (3.1.14)$$

From this and (A2), we have

$$|x_k(t)| \leq |x_0| + \int_0^t (r(s) + k_f |x_k(s)|^\alpha) ds \leq |x_0| + k_0 + k_f \int_0^t |x_k(s)|^\alpha ds, \quad (3.1.15)$$

where $k_0 = \int_0^1 |r(s)| ds$. We now consider three cases. If $\alpha = 1$, then we have from the Gronwall inequality that

$$|x_k(t)| \leq (|x_0| + k_0) \exp(k_f t).$$

If $0 \leq \alpha < 1$ then from (3.1.8), we have

$$|x_k(t)| \leq [(|x_0| + k_0)^{1-\alpha} + (1 - \alpha)k_f t]^{\frac{1}{1-\alpha}}. \quad (3.1.16)$$

If $\alpha > 1$ then $c_0 > 0$ and so (3.1.13) is valid. Since $\alpha \leq \alpha_0$, $L^{\alpha_0}([0, 1], \mathbb{R}^n) \hookrightarrow L^\alpha([0, 1], \mathbb{R}^n)$. Hence there exists $M_1 > 0$ such that

$$\int_0^1 |x_k(t)|^\alpha dt \leq M_1 \quad \forall k \geq 1.$$

Combining this with (3.1.15), we have

$$|x_k(t)| \leq |x_0| + k_0 + k_f M_1 \quad \forall k \geq k_0. \quad (3.1.17)$$

In a summary, we showed that there exists $M_2 > 0$ such that

$$\|x_k\|_0 \leq M_2, \quad \forall k > 0.$$

Let us show that $\{x_k\}$ is uniformly equicontinuous. Taking any $t_1, t_2 \in [0, 1]$ with $t_1 < t_2$, we have from (3.1.14), (A2) and Hölder inequality that

$$\begin{aligned} |x_k(t_1) - x_k(t_2)| &= \left| \int_{t_1}^{t_2} f(s, x_k(s), u_k(s)) ds \right| \\ &\leq \int_{t_1}^{t_2} (|r(s)| + k_f |x_k(s)|^\alpha) ds \\ &\leq \int_{t_1}^{t_2} (|r(s)| + k_f M_2^\alpha) ds \\ &\leq (t_2 - t_1) \| |r(\cdot)| + k_f M_2^\alpha \|_\infty. \end{aligned}$$

Hence $\{x_k\}$ is uniformly equicontinuous. By the Arzelà-Ascoli theorem, $\{x_k\}$ is a relatively compact set in $C([0, 1], \mathbb{R}^n)$. Passing to a subsequence, we can assume that $x_k \rightarrow \bar{x}$ uniformly in $C([0, 1], \mathbb{R}^n)$. The claim is justified.

Since D and X are closed sets, we have $\bar{x}(1) \in D$ and $\bar{x}(t) \in X$ for all $t \in [0, 1]$. By (3.1.10), there exists $K_{L, M_2} > 0$ such that

$$|L(t, x_k(t), u_k(t))| \leq K_{L, M_2} := M_4$$

for a.e. $t \in [0, 1]$. Also, from (3.1.11), there exist a number $k'_f > 0$ such that

$$|f(t, x_k(t), u_k(t))| \leq k'_f (1 + M_1^\alpha) := M_5$$

for a.e. $t \in [0, 1]$. Let us define $\xi_k(t) = L(t, x_k(t), u_k(t))$ and $\zeta_k(t) = f(t, x_k(t), u_k(t))$. Then $\|\xi_k\|_\infty \leq M_4$ and $\|\zeta_k\|_\infty \leq M_5$ for all $k > 0$. By passing to subsequences, we may assume that $\xi_k(\cdot)$ converges weakly star

to a function $\xi_0(\cdot) \in L^\infty$ and $\zeta_k(\cdot)$ converges weakly star to a function $\zeta_0(\cdot) \in L^\infty$ as $k \rightarrow \infty$. Then $\|\xi_0\|_\infty \leq M_4$, $\|\zeta_0\|_\infty \leq M_5$ and

$$\begin{aligned}\lim_{k \rightarrow \infty} \int_0^1 \varphi(t) \xi_k(t) dt &= \int_0^1 \varphi(t) \xi_0(t) dt \\ \lim_{k \rightarrow \infty} \int_0^1 h(t) \zeta_k(t) dt &= \int_0^1 h(t) \zeta_0(t) dt\end{aligned}$$

for all $\varphi, h \in L^1$. Particularly, for any measurable set $E \subset [0, 1]$ and $\varphi, h \in L^1(E, \mathbb{R})$, we have

$$\lim_{k \rightarrow \infty} \left[\int_E \varphi(t) \xi_k(t) dt + \int_E h(t) \zeta_k(t) dt \right] = \int_E \varphi(t) \xi_0(t) dt + \int_E h(t) \zeta_0(t) dt. \quad (3.1.18)$$

Let us now define a set-valued map $Q : [0, 1] \times B(0, M_2) \rightarrow \mathbb{R}^{n+1}$ by setting

$$Q(t, x) = \{(y, z) | M_4 \geq y \geq L(t, x, u), z = f(t, x, u) \text{ for some } u \in U\}, \quad (3.1.19)$$

where $B(0, M_2)$ is a closed ball in $\mathbb{R}^n$.

By (A3) and the compactness of U , we can show that $Q(t, x)$ is a nonempty, compact and convex set. We now prove the following lemma.

Lemma 3.1.5. *The set-valued map Q is a Caratheödory multi-function, that is, for each fixed $x \in B(0, M_2)$, $Q(\cdot, x)$ is measurable in t , and for a.e. $t \in [0, 1]$, $Q(t, \cdot)$ is continuous.*

Proof. We mimic the proof of Lemma 3.1 in [135].

- Showing that $Q(t, x)$ is measurable in t .

In fact, let C be a closed set in $\mathbb{R}^{1+n}$, we want to show that the set

$$Q(\cdot, x)^{-1}(C) = \{t \in [0, 1] | Q(t, x) \cap C \neq \emptyset\}$$

is closed. Let $t_n \in Q(\cdot, x)^{-1}(C)$ and $t_n \rightarrow t_0$. Then there exist $(y_n, z_n) \in C$ and $M_4 \geq y_n \geq L(t_n, x, u_n)$ and $z_n = f(t_n, x, u_n)$ for some $u_n \in U$. We can assume that $u_n \rightarrow u_0 \in U$. Hence $z_n \rightarrow z_0 = f(t_0, x, u_0)$, $L(t_n, x, u_n) \rightarrow L(t_0, x, u_0)$ and so y_n is bounded. By passing to a subsequence, we may assume that $y_n \rightarrow y_0$. By letting $n \rightarrow \infty$, we get $M_4 \geq y_0 \geq L(t_0, x, u_0)$. Since C is closed, $(y_0, z_0) \in C$. Hence $(y_0, z_0) \in Q(t_0, x) \cap C$ or $t_0 \in Q(\cdot, x)^{-1}(C)$. This means $Q(\cdot, x)^{-1}(C)$ is a closed set and so it is a measurable set. By Theorem 8.1.4 in [11], $Q(\cdot, x)$ is measurable.

- Showing that $Q(t, x)$ is lower semicontinuous with respect to x .

On the contrary, suppose that there is $x \in B(0, M_2)$ such that $Q(t, \cdot)$ is not lower semicontinuous at x . Then there exists an open set V in $\mathbb{R}^2$ such that $Q(t, x) \cap V \neq \emptyset$ and a sequence $x_n \in B(0, M_2), x_n \rightarrow x$ satisfying $Q(t, x_n) \cap V = \emptyset$. Then there exist $(y, v) \in V$ such that $M_4 \geq y \geq L(t, x, u)$ and $z = f(t, x, u)$ for some $u \in U$. Choose a sequence $u_n \in U$ such that $u_n \rightarrow u$. If $y = L(t, x, u)$, then we put $y_n = L(t, x_n, u_n)$ and $z_n = f(t, x_n, u_n)$. This implies that $(y_n, z_n) \rightarrow (y, z)$ and so $(y_n, z_n) \in Q(t, x_n) \cap V$ which is absurd. If $y > L(t, x, u)$ then we have $y > L(t, x_n, u_n)$ for n large enough. Define $y_n = y$ and $z_n = f(t, x_n, u_n)$ we have $(y_n, z_n) \rightarrow (y, u)$ and $(y_n, u_n) \in Q(t, x_n) \cap V$ for n large enough. It lead to a contradiction. By definition, $Q(t, \cdot)$ is lower semicontinuous at x .

- Showing that $Q(t, x)$ is upper semicontinuous with respect to x .

By a contradiction, we suppose that there is $x \in B(0, M_2)$ under which $Q(t, \cdot)$ is not upper semicontinuous at x . Then there exists an open set W in $\mathbb{R}^{n+1}$ such that $Q(t, x) \subset W$ and a sequence $x_n \rightarrow x$ such that $Q(t, x_n) \not\subset W$. Then there exists a sequence $(y_n, z_n) \in Q(t, x_n) \setminus W$. Hence $M_4 \geq y_n \geq L(t, x_n, u_n)$ and $z_n = f(t, x_n, u_n)$ for $u_n \in U$. By passing to a subsequence, we can assume that $u_n \rightarrow u \in U$. Hence $L(t, x_n, u_n) \rightarrow L(t, x, u)$ and $z_n \rightarrow z = f(t, x, u)$. It follows that $\{y_n\}$ is bounded. So we may assume that $y_n \rightarrow y$ as $n \rightarrow \infty$. By letting $n \rightarrow \infty$, we get $M_4 \geq y \geq L(t, x, u)$. Thus we have $(y, z) \in Q(t, x) \subset W$. Therefore $(y_n, z_n) \subset W$ for n large enough. This leads to a contradiction. The lemma is proved.

Let us define a set-valued map $\bar{Q} : [0, 1] \rightarrow \mathbb{R}^{1+n}$ by setting

$$\begin{aligned} \bar{Q}(t) &= Q(t, \bar{x}(t)) = \\ &= \{(y, z) \in \mathbb{R}^{1+n} \mid M_4 \geq y \geq L(t, \bar{x}(t), u), z = f(t, \bar{x}(t), u) \text{ for some } u \in U\}. \end{aligned}$$

Since Q is a Caratheodory multi-function, Theorem 8.2.8 in [11] implies that $\bar{Q}$ is a measurable set-valued map. Moreover, it has nonempty, compact and convex values. It is clear that $(\xi_k(t), \zeta_k(t)) \in Q(t, x_k(t))$ for a.e. $t \in [0, 1]$.

Lemma 3.1.6. $(\xi_0(t), \zeta_0(t)) \in \bar{Q}(t)$ for a.e. $t \in [0, 1]$.

Proof. On the contrary, suppose that there exists a measurable set $E \subset [0, 1]$ with positive measure such that

$$w(t) = (\xi_0(t), \zeta_0(t)) \notin \bar{Q}(t) \quad \forall t \in E.$$

Then $d(w(t), \bar{Q}(t)) > 0$ for all $t \in E$ because $\bar{Q}(t)$ are closed sets. Here $d(w(t), \bar{Q}(t))$ denotes the distance from $w(t)$ to $\bar{Q}(t)$. Since $\bar{Q}(t)$ is convex, for each t , there exists a unique $w'(t) \in \bar{Q}(t)$ such that $|w(t) - w'(t)|_e = d(w(t), \bar{Q}(t))$, where $|\cdot|_e$ denotes the Euclid norm in $\mathbb{R}^{1+n}$. In other words, $w'(t)$ is the metric projection of $w(t)$ on $\bar{Q}(t)$. Let us show that $w'(t)$ is measurable. Define function

$$g(t, w'') = |w(t) - w''|_e.$$

Then g is a Caratheödory function and

$$d(w(t), \bar{Q}(t)) = \inf_{w'' \in \bar{Q}(t)} g(t, w'').$$

Define the set-valued map

$$P(t) = \{w' \in \mathbb{R}^{1+n} | g(t, w') = \inf_{w'' \in \bar{Q}(t)} g(t, w'')\}.$$

Since $\bar{Q}(t)$ is compact, $P(t)$ has nonempty closed values. By Theorem 8.2.11 in [11], the mapping $t \mapsto d(w(t), \bar{Q}(t))$ and $P(t)$ are measurable. By the measurable selection theorem (see Theorem 8.1.3 in [11]), P has a measurable selection, that is there exists a measurable function $w_0(t)$ such that $w_0(t) \in P(t)$. This implies that

$$|w(t) - w_0(t)|_e = d(w(t), \bar{Q}(t)) = |w(t) - w'(t)|_e.$$

By the uniqueness, we have $w'(t) = w_0(t)$. Hence $w'(t)$ is measurable too. Let $w'(t) = (w'_1(t), w'_2(t)) \in \mathbb{R} \times \mathbb{R}^n$. Then $w'_1 \in L^\infty(E, \mathbb{R})$ and $w'_2 \in L^\infty(E, \mathbb{R}^n)$. In fact, there exist $u \in U$ such that $M_4 \geq w'_1(t) \geq L(t, \bar{x}(t), u)$ and $w'_2(t) = f(t, \bar{x}(t), u)$. Let $M_0 := \|\bar{x}\|_0$. By (3.1.10) and (3.1.11), there exist $K_{L, M_0} > 0$ and $K_{f, M_0} > 0$ such that

$$M_4 \geq w'_1(t) \geq L(t, \bar{x}(t), u) \geq -K_{L, M_0}$$

and $|w'_2(t)| = |f(t, \bar{x}(t), u)| \leq K_{f, M_0}$ for all $t \in E$. By a property of metric projection (see Theorem 5.2 in [31]), we have

$$\langle w(t) - w'(t), v - w'(t) \rangle \leq 0 \quad \forall v \in \bar{Q}(t).$$

Here $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^{1+n}$. This implies that

$$\begin{aligned} 0 < |w(t) - w'(t)|_e^2 &= \langle w(t) - w'(t), w(t) - w'(t) \rangle \\ &\leq \langle w(t) - w'(t), w(t) - v \rangle \quad \forall v \in \bar{Q}(t). \end{aligned}$$

Hence

$$\langle w'(t) - w(t), w(t) \rangle < \langle w'(t) - w(t), v \rangle \quad \forall v \in \bar{Q}(t).$$

Let us put $w'(t) - w(t) = (\varphi(t), h(t))$ and $v = (v_1, v_2)$. Then we have

$$\varphi(t)\xi_0(t) + h(t)\zeta_0(t) \leq \varphi(t)v_1 + h(t)v_2 \quad \forall (v_1, v_2) \in \bar{Q}(t).$$

By taking $v_1 = L(t, \bar{x}(t), u)$ and $v_2 = f(t, \bar{x}(t), u)$ with $u \in U$, we have

$$\varphi(t)\xi_0(t) + h(t)\zeta_0(t) < \varphi(t)L(t, \bar{x}(t), u) + h(t)f(t, \bar{x}(t), u) \quad \forall u \in U. \quad (3.1.20)$$

Hence

$$\varphi(t)\xi_0(t) + h(t)\zeta_0(t) < \min_{u \in U} [\varphi(t)L(t, \bar{x}(t), u) + h(t)f(t, \bar{x}(t), u)]. \quad (3.1.21)$$

Note that $\varphi \in L^\infty(E, \mathbb{R})$, $h \in L^\infty(E, \mathbb{R})^n$, $x_k(t) \rightarrow \bar{x}(t)$ as $k \rightarrow \infty$ and the mapping

$$x \mapsto \min_{u \in U} [\varphi(t)L(t, x, u) + h(t)f(t, x, u)]$$

is continuous. It follows that

$$\begin{aligned} \varphi(t)\xi_0(t) + h(t)\zeta_0(t) &< \min_{u \in U} [\varphi(t)L(t, \bar{x}(t), u) + h(t)f(t, \bar{x}(t), u)] \\ &= \lim_{k \rightarrow \infty} \min_{u \in U} [\varphi(t)L(t, x_k(t), u) + h(t)f(t, x_k(t), u)] \leq \varrho \end{aligned}$$

for some $\varrho > 0$. This and the dominated convergence theorem imply that

$$\begin{aligned} &\int_E [\varphi(t)\xi_0(t) + h(t)\zeta_0(t)] dt \\ &< \int_E \left(\lim_{k \rightarrow \infty} \min_{u \in U} [\varphi(t)L(t, x_k(t), u) + h(t)f(t, x_k(t), u)] \right) dt \\ &= \lim_{k \rightarrow \infty} \int_E \left(\min_{u \in U} [\varphi(t)L(t, x_k(t), u) + h(t)f(t, x_k(t), u)] \right) dt \\ &\leq \lim_{k \rightarrow \infty} \int_E (\varphi(t)L(t, x_k(t), u_k(t)) + h(t)f(t, x_k(t), u_k(t))) dt \end{aligned}$$

which contradicts (3.1.18). The lemma is proved.

According to Lemma 3.1.6, we have $(\xi_0(t), \zeta_0(t)) \in \bar{Q}(t)$ for a.e. $t \in [0, 1]$. Therefore, for each $t \in [0, 1]$ there exists $u \in U$ such that

$$M_4 \geq \xi_0(t) \geq L(t, \bar{x}(t), u), \quad \zeta_0(t) = f(t, \bar{x}(t), u). \quad (3.1.22)$$

Define set-valued maps $Q_1 : [0, 1] \rightarrow 2^{\mathbb{R}^{1+n}}$ and $Q_2 : [0, 1] \rightarrow 2^U$ by setting

$$Q_1(t) := (-\infty, \xi_0(t)] \times \{\zeta_0(t)\}$$

and

$$Q_2(t) = \{u \in U | (L(t, \bar{x}(t), u), f(t, \bar{x}(t), u)) \in Q_1(t)\}.$$

Then $Q_2(t)$ has nonempty and closed values. Besides, Q_1 is measurable. In fact, let $O = (a, b) \times \prod_{i=1}^n (a_i, b_i)$ be a open set in $\mathbb{R}^{1+n}$. Then

$$Q_1^{-1}(O) = \{t \in [0, 1] | P_1(t) \cap O \neq \emptyset\} = \xi_0^{-1}((a, +\infty)) \cap \zeta_0^{-1}\left(\prod_{i=1}^n (a_i, b_i)\right) \quad (3.1.23)$$

which is a measurable set because ξ_0 and ζ_0 are measurable functions. Hence Q_1 is a measurable set-valued map. By Theorem 8.2.9 in [11], Q_2 is a measurable set-valued map with closed images. By Theorem 8.1.3 in [11], there exists a measurable function $\bar{u}(t)$ such that $\bar{u}(t) \in Q_2(t)$ for a.e. $t \in [0, 1]$. It follows that $\bar{u}(t) \in U$ for a.e. $t \in [0, 1]$ and

$$\xi_0(t) \geq L(t, \bar{x}(t), \bar{u}(t)), \quad \zeta_0(t) = f(t, \bar{x}(t), \bar{u}(t)).$$

In a summary, we have shown that $\xi_k = L(\cdot, x_k, u_k)$ converges weakly star to ξ_0 with $\xi_0(t) \geq L(t, \bar{x}(t), \bar{u}(t))$, and $\zeta_k = f(\cdot, x_k, u_k)$ converges weakly star to $\zeta_0 = f(\cdot, \bar{x}, \bar{u})$.

Since $(x_k, u_k) \in \Phi$, we have

$$x_k(t) = x_0 + \int_0^t f(s, x_k(s), u_k(s))ds = x_0 + \int_0^1 \chi_{[0,t]}(s) \zeta_k(s)ds,$$

where $\chi_{[0,t]}(s) = 1$ if $s \in [0, t]$ and $\chi_{[0,t]}(s) = 0$ if $s \notin [0, t]$. By letting $k \rightarrow \infty$, we obtain

$$\bar{x}(t) = x_0 + \int_0^t f(s, \bar{x}(s), \bar{u}(s))ds. \quad (3.1.24)$$

This implies that

$$\dot{\bar{x}}(t) = f(t, \bar{x}(t), \bar{u}(t)), \quad \bar{x}(0) = x_0.$$

It easy to show that $f(\cdot, \bar{x}, \bar{u}) \in L^\infty([0, 1], \mathbb{R}^n)$. Hence $\bar{x} \in W^{1,\infty}([0, 1], \mathbb{R}^n)$. Since $\bar{x}(1) \in D$ and $\bar{x}(t) \in X$ for all $t \in [0, 1]$, we have $(\bar{x}, \bar{u}) \in \Phi$. Finally, we have

$$\begin{aligned} V &= \lim_{k \rightarrow \infty} J(x_k, u_k) = \lim_{k \rightarrow \infty} \int_0^1 L(t, x_k(t), u_k(t))dt \\ &= \int_0^1 \xi_0(t)dt \geq \int_0^1 L(t, \bar{x}(t), \bar{u}(t))dt. \end{aligned}$$

Hence $(\bar{x}, \bar{u})$ is an optimal solution of problem (3.1.1)-(3.1.3). The proof of the theorem is complete. $\square$

3.2 Optimality conditions for multiobjective optimal control problems

3.2.1 First and second-order optimality conditions of KKT-type for multiobjective optimal control problems with free right endpoints

Let

$$\begin{aligned} L_j &: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}, \quad j = 1, 2, \dots, m \\ \varphi &: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}, \quad g: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R} \end{aligned}$$

be given functions. Hereafter, we denote by $\mathbb{R}_+^m$ the nonnegative orthant of $\mathbb{R}^m$, that is,

$$\mathbb{R}_+^m = \{(a_1, a_2, \dots, a_m) \in \mathbb{R}^m : a_i \geq 0, i = 1, 2, \dots, m\}.$$

The interior of $\mathbb{R}_+^m$ is denoted by $\text{int } \mathbb{R}_+^m$.

We consider the multi-objective optimal control problem of finding a control vector $u \in L^\infty([0, 1], \mathbb{R}^l)$ and the corresponding state $x \in C([0, 1], \mathbb{R}^n)$ which solve

$$\text{Min}_{\mathbb{R}_+^m} I(x, u) \tag{3.2.1}$$

$$\text{s.t.} \quad x(t) = x_0 + \int_0^t \varphi(s, x(s), u(s)) ds \quad \text{for a.e. } t \in [0, 1], \tag{3.2.2}$$

$$g(t, x(t), u(t)) \leq 0 \quad \text{for a.e. } t \in [0, 1]. \tag{3.2.3}$$

Here x_0 is a given vector in $\mathbb{R}^n$ and the multi-objective function I is given by

$$I(x, u) = (I_1(x, u), I_2(x, u), \dots, I_m(x, u)),$$

where

$$I_j(x, u) := \int_0^1 L_j(t, x(t), u(t)) dt.$$

We denote by (MCP1) the problem (3.2.1)–(3.2.3) and by Φ its feasible set, that is, Φ consists of couples $(x, u) \in C([0, 1], \mathbb{R}^n) \times L^\infty([0, 1], \mathbb{R}^l)$ which satisfy constraints (3.2.2)–(3.2.3).

The multi-objective optimal control problems are important in mechanics and economics. For example, when we want to minimize energy and time of a system, we need to use two-objective optimal control which has a form (3.2.1)–(3.2.3) (see for instance [117]). Recently, problem (MCP1) has been studied by several mathematicians. For papers which have a close connection to the present work, we refer the readers to [20], [27], [95], [117], [128], [170], [173], [174], [211], [240] and the references therein. However, in these papers, the authors mainly studied numerical methods and first-order necessary optimality conditions for multi-objective optimal control problems. To the best of our knowledge, so far there have been no papers investigating second-order optimality conditions for multi-objective optimal control problems. The study of second-order optimality conditions for optimization problems as well as for multi-objective optimal control problems is a fundamental topic in optimization theory. The second-order optimality conditions play an important role in solution stability and numerical methods of finding optimal solutions.

In this subsection, we shall establish second-order conditions of KKT-type for locally Pareto solutions to the multi-objective optimal control problem with free right endpoints. To do this, we first establish optimality conditions for an abstract multi-objective optimal control problem under the Robinson constraint qualification. This can be done by using Theorem 2.2.5. We then apply the obtained results to problem (3.2.1)–(3.2.3).

• Abstract multi-objective optimal control problem

Let E_0, E, X and U be Banach spaces and Q be a nonempty closed convex set in E . Define $Z := X \times U$ and assume that

$$\begin{aligned} I: X \times U &\rightarrow \mathbb{R}^m, \\ F: X \times U &\rightarrow E_0, \\ G: X \times U &\rightarrow E \end{aligned}$$

are given mappings. We consider the following multi-objective optimal control problem of finding a control $u \in U$ and the corresponding state $x \in X$ which solve

$$(MP2) \quad \begin{cases} \text{Min}_{\mathbb{R}_+^m} & I(x, u), \\ \text{s.t.} & F(x, u) = 0, \\ & G(x, u) \in Q. \end{cases}$$

We denote by Φ the feasible of the (MP2) and put

$$D := \{(x, u) \in Z \mid F(x, u) = 0\}.$$

Fix $z_0 = (y_0, u_0) \in \Phi$. We denote by $\Lambda_2(z_0)$ the set of multipliers $(\lambda, v^*, e^*) \in \mathbb{R}_+^m \times E_0^* \times E^*$ with $|\lambda| = 1$, which satisfies the following conditions

$$\nabla_z \mathcal{L}_2(z, \lambda, v^*, e^*) = 0, \quad e^* \in N(Q, G(z, w_0)),$$

where $\mathcal{L}_2(z, \lambda, v^*, e^*)$ is the Lagrangian given by

$$\mathcal{L}_2(z, \lambda, v^*, e^*) := \langle \lambda, I(z) \rangle + \langle v^*, F(z) \rangle + \langle e^*, G(z) \rangle.$$

Let $\mathcal{C}_2(z_0)$ consist of directions $d \in Z$ satisfying conditions:

$$\begin{cases} \nabla I(z_0)d \in -\mathbb{R}_+^m, \\ \nabla F(z_0)d = 0, \\ \nabla G(z_0)d \in \text{cone}(Q - G(z_0)). \end{cases}$$

Let $\mathcal{C}_2(z_0)$ be the closure of $\mathcal{C}_2(z_0)$ in Z . Then $\mathcal{C}_2(z_0)$ is called the critical cone of the (3.2.1) at z_0 .

Let us introduce the following assumptions:

- (A1) There exist positive numbers r_1, r'_1 such that the mapping $I(\cdot, \cdot)$, $F(\cdot, \cdot)$ and $G(\cdot, \cdot)$ are twice continuously Fréchet differentiable on $B_X(x_0, r_1) \times B_U(u_0, r'_1)$;
- (A2) The mapping $F_x(z_0)$ is bijective;
- (A3) $E = \nabla G(z_0)(T(D; z_0)) - \text{cone}(Q - G(z_0))$.

From assumptions (A1) and (A3), we have that $F(\cdot, \cdot)$ is continuously differentiable on $B_X(x_0, r_1) \times B_U(u_0, r'_1)$ and $F_x(z_0)$ is bijective. By the implicit function theorem (see [?, Theorem 4.E]), there exist balls $B_X(x_0, r_2)$, $B_U(u_0, r'_2)$ with $r_2 < r_1, r'_2 < r'_1$ such that for each $u \in B_U(u_0, r'_2)$, the equation

$$F(x, u) = 0$$

has a unique solution $x = \zeta(u) \in B_X(x_0, r_2)$. Moreover, the mapping

$$\zeta: B_U(u_0, r'_2) \rightarrow B_X(x_0, r_2)$$

is of class C^2 and $\zeta(u_0) = x_0$. Thus,

$$F(\zeta(u), u) = 0 \quad \forall u \in B_U(u_0, r'_2). \quad (3.2.4)$$

We now define the following mappings:

$$\begin{aligned} J: U &\rightarrow \mathbb{R}^m, & J(u) &:= J(\zeta(u), u), \\ H: U &\rightarrow E, & H(u) &:= G(\zeta(u), u). \end{aligned} \quad (3.2.5)$$

Then we can show that (x_0, u_0) is a local weak Pareto solution of the (3.2.1) if and only if u_0 is a local weak Pareto solution of the following problem:

$$\begin{aligned} &\text{Min } \mathbb{R}_+^m J(u) \\ &\text{s.t. } H(u) \in Q. \end{aligned} \quad (\text{MP3})$$

Problem (MP3) is associated with the Lagrangian

$$\mathcal{L}_3(u, \lambda, e^*) := \lambda J(u) + e^* H(u).$$

Given a feasible point u_0 of the (MP3), we define

$$\mathcal{C}_{03}(u_0) := \{u \in U \mid \nabla J(u_0)u \in -\mathbb{R}_+^m, \nabla H(u_0)u \in \text{cone}(Q - H(u_0))\}$$

and $\mathcal{C}_3(u_0) := \overline{\mathcal{C}_{03}(u_0)}$ the interior critical cone and the critical cone at u_0 , respectively.

The following theorem provides second-order necessary optimality conditions for the (3.2.1).

Theorem 3.2.1. *Suppose that z_0 is a feasible point of the (3.2.1) and assumptions (A1)–(A3) are satisfied. If z_0 is a local weak Pareto solution of the (3.2.1), then, for each $d \in \mathcal{C}_2(z_0)$, there exists a nonzero triple $(\lambda, v^*, e^*) \in \Lambda_2(z_0)$ such that*

$$\nabla_z^2 \mathcal{L}_2(z_0, e^*, v^*)(d, d) = [\lambda J_{zz}(z_0) + v^* F_{zz}(z_0) + e^* G_{zz}(z_0)](d, d) \geq 0.$$

Proof. Since $z_0 = (x_0, u_0)$ is a local weak Pareto solution to the (3.2.1), u_0 is a local weak Pareto solution of the (MP3). By assumption (A2), $\nabla F(z_0)$ is surjective. Indeed, for any $v \in E_0$, there exists $x \in X$ such that $F_x(\bar{z})x = v$. Hence, $(x, 0) \in X \times U$ and $\nabla F(\bar{z})(x, 0) = v$. This means that $\nabla F(\bar{z})$ is surjective. From this and Lemma 2.1.17, we have

$$T(D; z_0) = \{(x, u) \in Z \mid F_x(z_0)x + F_u(z_0)u = 0\} = \{(\zeta'(u_0)u, u) \mid u \in U\}.$$

Combining this with (A3), we get

$$\begin{aligned} E &\subseteq \{\nabla_x G(z_0)\zeta'(u_0)u + \nabla_u G(z_0)u \mid u \in U\} - \text{cone}(Q - G(z_0)) \\ &\subseteq \nabla H(u_0)(U) - \text{cone}(Q - H(u_0)), \end{aligned}$$

where H is defined by (3.2.5). Hence the Robinson constraint qualification for the (MP3) is satisfied at u_0 .

Fix any $d = (x, u) \in \mathcal{C}_2(z_0)$. Then there exists a sequence $\{d_k\} = \{(x_k, u_k)\} \subset \mathcal{C}_{02}(z_0)$ such that $d_k \rightarrow d$. Since $F_x(z_0)x_k + F_u(z_0)u_k = 0$, we have $x_k = \zeta'(u_0)u_k$ and so $u_k \in \mathcal{C}_{03}(u_0)$. Consequently, $u \in \mathcal{C}_3(u_0)$. By Theorem 2.2.5, there exists a multipliers $(\lambda, e^*) \in \Lambda_*(u_0)$ such that the following conditions hold:

- (a) $\lambda \nabla J(u_0) + \nabla e^* H(u_0) = 0$, $e^* \in N(Q, H(u_0))$,
- (b) $\langle \lambda \nabla^2 J(u_0)u, u \rangle + \langle e^* \nabla^2 H(u_0)u, u \rangle \geq 0$.

Note that from (3.2.4), we have

$$F(\zeta(v), v) = 0 \quad \forall v \in B_U(u_0, r'_3). \quad (3.2.6)$$

Taking first-order derivative on both sides, we get

$$F_x(z_0)\zeta'(u_0) + F_u(z_0) = 0$$

and so

$$\zeta'(u_0)^* = -F_u(z_0)^*(F_x^*(z_0))^{-1}. \quad (3.2.7)$$

From (a), we have

$$\lambda I_x(z_0)\zeta'(u_0) + \lambda I_u(z_0) + e^* G_x(z_0)\zeta'(u_0) + e^* G_u(z_0) = 0,$$

or, equivalently,

$$\zeta'(u_0)^*(\lambda I_x(z_0) + e^* G_x(z_0)) = -(\lambda I_u(z_0) + e^* G_u(z_0)). \quad (3.2.8)$$

Let us put

$$\phi := (\lambda I_x(z_0) + e^* G_x(z_0)), \quad v^* := -(F_x^*(z_0))^{-1}\phi. \quad (3.2.9)$$

Then, from (3.2.7) and (3.2.8), we have

$$-F_u(z_0)^*(F_x^*(z_0))^{-1}\phi = -(\lambda I_u(z_0) + e^* G_u(z_0)).$$

Consequently,

$$\begin{cases} F_u(z_0)^*v^* + \lambda I_u(z_0) + e^* G_u(z_0) = 0 \\ F_x(z_0)^*v^* = -(\lambda I_x(z_0) + e^* G_x(z_0)). \end{cases}$$

This is equivalent to

$$\nabla I(z_0)^* \lambda^T + \nabla F(z_0)^* v^* + \nabla G(z_0)^* e^* = 0.$$

Hence we have $(\lambda, v^*, e^*) \in \Lambda_2(z_0)$.

Given $u \in \mathcal{C}_3(u_0)$, we define the following function

$$\psi(t) := \mathcal{L}_3(u_0 + tu, \lambda, e^*) = \lambda J(u_0 + tu) + e^* H(u_0 + tu), \quad -1 < t < 1.$$

Then by (b), we have

$$\psi''(0) = \nabla_u^2 \mathcal{L}_3(u_0)(u, u) = \langle \lambda \nabla^2 J(u_0)u, u \rangle + \langle e^* \nabla^2 H(u_0)u, u \rangle \geq 0.$$

On the other hand, by simple calculation, we get

$$\begin{aligned} \psi''(0) &= \langle \lambda I_{xx}(z_0) \zeta'(u_0)u, \zeta'(u_0)u \rangle + \langle \lambda I_{xu}(z_0) \zeta'(u_0)u, u \rangle \quad (3.2.10) \\ &\quad + \langle \lambda I_{ux}(z_0)u, \zeta'(u_0)u \rangle + \langle \lambda I_{uu}(z_0)u, u \rangle + \langle e^* G_{xx}(z_0) \zeta'(u_0)u, \zeta'(u_0)u \rangle \\ &\quad + \langle e^* G_{xu}(z_0) \zeta'(u_0)u, u \rangle + \langle e^* G_{ux}(z_0)u, \zeta'(u_0)u \rangle + \langle e^* G_{uu}(z_0)u, u \rangle \\ &\quad + \langle (\lambda I_x(z_0) + e^* G_x(z_0)) \zeta''(u_0)u, u \rangle \\ &= \langle \lambda I_{xx}(z_0)x, x \rangle + \langle \lambda I_{xu}(z_0)x, u \rangle + \langle \lambda I_{ux}(z_0)u, x \rangle + \langle \lambda I_{uu}(z_0)u, u \rangle \\ &\quad + \langle e^* G_{xx}(z_0)x, x \rangle + \langle e^* G_{xu}(z_0)x, u \rangle + \langle e^* G_{ux}(z_0)u, x \rangle \\ &\quad + \langle e^* G_{uu}(z_0)u, u \rangle + \langle (\zeta''(u_0)u)^* (\lambda I_x(z_0) + e^* G_x(z_0)), u \rangle. \end{aligned}$$

Taking second-order derivatives on both sides of (3.2.6) at u_0 and acting on $u \in \mathcal{C}''(u_0)$ and $v \in U$, we obtain

$$\begin{aligned} \langle F_x(z_0) \zeta''(u_0)u, v \rangle + \langle F_{xx}(z_0) \zeta'(u_0)u, \zeta'(u_0)v \rangle + \langle F_{xu}(z_0) \zeta'(u_0)u, v \rangle \\ + F_{ux}(z_0)(u, \zeta'(z_0)v) + F_{uu}(z_0)(u, v) = 0. \end{aligned}$$

This is equivalent to

$$\begin{aligned} \langle F_x(z_0) \zeta''(u_0)u, v \rangle &= -[\langle F_{xx}(z_0)x, \zeta'(u_0)v \rangle + \langle F_{xu}(z_0)x, v \rangle \\ &\quad + \langle F_{ux}(z_0)u, \zeta'(z_0)v \rangle + \langle F_{uu}(z_0)u, v \rangle]. \end{aligned}$$

It follows that

$$\begin{aligned} (\zeta''(u_0)u)^* &= -[F_{xx}(z_0)x \zeta'(u_0) + F_{xu}(z_0)x + F_{ux}(z_0)u \zeta'(z_0) + F_{uu}(z_0)u]^* (F_x(z_0)^*)^{-1}. \end{aligned}$$

Combining this with formula (3.2.9), we have

$$\begin{aligned} (\zeta''(u_0)u)^* (\lambda I_x(z_0) + e^* G_x(z_0)) &= -(\zeta''(u_0)u)^* \phi \\ &= -[F_{xx}(z_0)x \zeta'(u_0) + F_{xu}(z_0)x + F_{ux}(z_0)u \zeta'(z_0) + F_{uu}(z_0)u]^* (F_x(z_0)^*)^{-1} \phi \\ &= -[F_{xx}(z_0)x \zeta'(u_0) + F_{xu}(z_0)x + F_{ux}(z_0)u \zeta'(z_0) + F_{uu}(z_0)u]^* v^*. \end{aligned}$$

Hence,

$$\begin{aligned}
& \langle (\zeta''(u_0)u)^*(\lambda I_x(z_0) + e^* G_x(z_0)), u \rangle = \\
& = \langle v^* F_{xx}(z_0)x, \zeta'(u_0)u \rangle + \langle v^* F_{xu}(z_0)x, u \rangle + \langle F_{ux}(z_0)u, \zeta'(z_0)u \rangle \\
& + \langle v^* F_{uu}(z_0)u, u \rangle \\
& = \langle v^* F_{xx}(z_0)x, x \rangle + \langle v^* F_{xu}(z_0)x, u \rangle + \langle F_{ux}(z_0)u, x \rangle + \langle v^* F_{uu}(z_0)u, u \rangle.
\end{aligned}$$

Inserting this term into (3.2.10), we obtain

$$\begin{aligned}
& \langle \lambda \nabla^2 J(u_0)u, u \rangle + \langle e^* \nabla^2 H(u_0)u, u \rangle = \\
& = \langle \lambda \nabla^2 I(z_0)d, d \rangle + \langle v^* \nabla^2 F(z_0)d, d \rangle + \langle e^* \nabla^2 G(z_0)d, d \rangle \geq 0.
\end{aligned}$$

The proof is complete. $\square$

• Second-order optimality conditions for (MCP1)

In what follows, $C([0, 1], \mathbb{R}^n)$ is the Banach space of continuous vector-valued functions $x: [0, 1] \rightarrow \mathbb{R}^n$ with the norm $\|x\|_0 = \max_{t \in [0, 1]} |x(t)|$ and $\mathbb{R}^n$ is the Euclidean space of n -tuples $\xi = (\xi_1, \dots, \xi_n)$ with the norm $|\xi| = (\sum_{i=1}^n \xi_i^2)^{\frac{1}{2}}$. For each $1 \leq p \leq \infty$, $L^p([0, 1], \mathbb{R}^l)$ stands for the Lebesgue spaces with the norms $\|\cdot\|_p$. Let $W^{1,2}([0, 1], \mathbb{R}^n)$ be Sobolev space which consists of absolutely continuous functions x such that $\dot{x} \in L^2([0, 1], \mathbb{R}^n)$. Its norm is given by

$$\|x\|_{1,2} = \|x\|_2 + \|\dot{x}\|_2.$$

For convenience, we put $X := C([0, 1], \mathbb{R}^n)$ and $U := L^\infty([0, 1], \mathbb{R}^l)$. In the sequel, L and ϕ stand for $(L_1, \dots, L_m)$ and (φ, g) , respectively. Define

$$Q := \{v \in L^\infty([0, 1], \mathbb{R}) \mid v(t) \leq 0, \text{ a.e. } t \in [0, 1]\}.$$

Let us impose the following assumptions on L and ϕ .

(H1) The function L is a Carathéodory function and ϕ is a continuous mapping. For a.e. $t \in [0, 1]$, $L(t, \cdot, \cdot)$ is of class C^2 while $\phi(t, \cdot, \cdot)$ is of class C^2 for all $t \in [0, 1]$. Besides, for each $M > 0$, there exist numbers $k_{LM} > 0$ and $k_{\phi M} > 0$ such that

$$\begin{aligned}
& |L(t, x_1, u_1) - L(t, x_2, u_2)| + |\nabla L(t, x_1, u_1) - \nabla L(t, x_2, u_2)| + \\
& + |\nabla^2 L(t, x_1, u_1) - \nabla^2 L(t, x_2, u_2)| \leq k_{LM}(|x_1 - x_2| + |u_1 - u_2|)
\end{aligned}$$

for a.e. $t \in [0, 1]$ and for all $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^l$ satisfying $|x_i| \leq M$, $|u_i| \leq M$ with $i = 1, 2$, and

$$\begin{aligned}
& |\phi(t, x_1, u_1) - \phi(t, x_2, u_2)| + |\nabla \phi(t, x_1, u_1) - \nabla \phi(t, x_2, u_2)| + \\
& + |\nabla^2 \phi(t, x_1, u_1) - \nabla^2 \phi(t, x_2, u_2)| \leq k_{\phi M}(|x_1 - x_2| + |u_1 - u_2|)
\end{aligned}$$

for all $t \in [0, 1]$, $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^l$ satisfying $|x_i| \leq M$, $|u_i| \leq M$ with $i = 1, 2$. Moreover, we require that the functions

$L(t, 0, 0), |\nabla L(t, 0, 0)|, |\nabla^2 L(t, 0, 0)|, \phi(t, 0, 0), |\nabla \phi(t, 0, 0)|, |\nabla^2 \phi(t, 0, 0)|$
belong to $L^\infty([0, 1], \mathbb{R})$.

(H2) Given a couple $(\bar{x}, \bar{u}) \in \Phi$, there exist $i_0 \in \{1, 2, \dots, l\}$ and $\alpha > 0$ such that

$$|g_{u_{i_0}}(t, \bar{x}(t), \bar{u}(t))| \geq \alpha \text{ for a.e. } t \in [0, 1].$$

Hereafter, $L[t], \varphi[t], g[t], L_x[t], \varphi_u[t], g_u[t]$ and so on, stand for

$$\begin{aligned} &L(t, \bar{x}(t), \bar{u}(t)), \varphi(t, \bar{x}(t), \bar{u}(t)), g(t, \bar{x}(t), \bar{u}(t)), \\ &L_x(t, \bar{x}(t), \bar{u}(t)), \varphi_u(t, \bar{x}(t), \bar{u}(t)), g_u(t, \bar{x}(t), \bar{u}(t)), \dots \end{aligned}$$

Definition 3.2.2. Assume that $\bar{z} = (\bar{x}, \bar{u})$ is a feasible point of the (MCP1). We say that:

(i) $\bar{z}$ is a local weak Pareto solution of the (MCP1) if there exists $\epsilon > 0$ such that for all $(x, u) \in (B(\bar{x}, \epsilon) \times B(\bar{u}, \epsilon)) \cap \Phi$, one has

$$I(x, u) - I(\bar{x}, \bar{u}) \notin -\text{int } \mathbb{R}_+^m.$$

(ii) $\bar{z}$ is a local Pareto solution of the (MCP1) if there exists $\epsilon > 0$ such that for all $(x, u) \in (B(\bar{x}, \epsilon) \times B(\bar{u}, \epsilon)) \cap \Phi$, one has

$$I(x, u) - I(\bar{x}, \bar{u}) \notin -\mathbb{R}_+^m \setminus \{0\}.$$

Let us denote by $\mathcal{C}_0(\bar{z})$ the set of vectors

$$z = (x, u) \in C([0, 1], \mathbb{R}^n) \times L^\infty([0, 1], \mathbb{R}^l)$$

such that the following conditions hold:

- (c1) $\int_0^1 (L_x[t]x(t) + L_u(t)u(t)) dt \in -\mathbb{R}_+^m$;
- (c2) $x(\cdot) = \int_0^{(\cdot)} (\varphi_x[s]x(s) + \varphi_u[s]u(s)) ds$;
- (c3) $g_x[\cdot]x + g_u[\cdot]u \in \text{cone}(Q - g(\cdot, \bar{x}, \bar{u}))$.

Let $\mathcal{C}(\bar{z})$ be the closure of $\mathcal{C}_0(\bar{z})$ in $C([0, 1], \mathbb{R}^m) \times L^\infty([0, 1], \mathbb{R}^l)$. We call $\mathcal{C}(\bar{z})$ the *critical cone* of the (MCP1) at $\bar{z}$. Each vector $z \in \mathcal{C}(\bar{z})$ is called a *critical direction* to the (MCP1) at $\bar{z}$. It is easily seen that $\mathcal{C}(\bar{z})$ is a closed convex cone containing 0.

The following theorem gives necessary optimality conditions for the (MCP1).

Theorem 3.2.3. *Suppose that assumptions (H1) and (H2) are valid and $\bar{z}$ is a local weak Pareto solution of the (MCP1). Then, for each $z \in \mathcal{C}(\bar{z})$, there exist a vector $\lambda \in \mathbb{R}_+^m$ with $|\lambda| = 1$, an absolutely continuous function $\bar{p}: [0, 1] \rightarrow \mathbb{R}^n$ and a function $\theta \in L^1([0, 1], \mathbb{R})$ such that the following conditions are fulfilled:*

(i) $\theta \in N(Q, g[\cdot]);$

(ii) (the adjoint equation)

$$\begin{cases} \dot{\bar{p}}(t) = -\lambda^T L_x[t] - \varphi_x[t]\bar{p}(t) - \theta(t)g_x[t] & \text{a.e. } t \in [0, 1], \\ \bar{p}(1) = 0; \end{cases}$$

(iii) (the stationary condition in u)

$$\lambda^T L_u[t] + \bar{p}^T \varphi_u[t] + \theta(t)g_u[t] = 0 \quad \text{a.e. } t \in [0, 1];$$

(iv) (the non-negative condition)

$$\begin{aligned} & \int_0^1 \left(\sum_{j=1}^m \lambda_j \nabla^2 L_j[t] z(t), z(t) \right) dt + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t] z(t), z(t)) dt \\ & + \int_0^1 (\theta(t)^T \nabla^2 g[t] z(t), z(t)) dt \geq 0. \end{aligned}$$

Proof. For the proof, we first put

$$\begin{aligned} X &:= C([0, 1], \mathbb{R}^n), \quad U := L^\infty([0, 1], \mathbb{R}^l), \\ E_0 &:= C([0, 1], \mathbb{R}^n), \quad E := L^\infty([0, 1], \mathbb{R}), \end{aligned}$$

and define the following mappings

$$\begin{aligned} F: X \times U &\rightarrow E_0, \quad F(x, u) := x - x_0 - \int_0^{(\cdot)} \varphi(s, x(s), u(s)) ds, \\ G: X \times U &\rightarrow E, \quad G(x, u) := g(\cdot, x(\cdot), u(\cdot)). \end{aligned}$$

The problem (MCP) can be formulated in the form of the problem (MP3). Therefore, we can apply Theorem 3.2.1 for problem (MCP1) in order to derive necessary optimality conditions.

Step 1. Verification of assumptions (A1)–(A3).

• Verification of (A1). From (H1) we see that the mapping I, F and G are of class C^2 around $\bar{z}$. Hence, (A1) is valid. Here $\nabla I_j(\bar{z}), \nabla^2 I(\bar{z})$,

$\nabla F(\bar{z})$, $\nabla^2 F(\bar{z})$, $\nabla G(\bar{z})$ and $\nabla^2 G(\bar{z})$ are defined by:

$$\begin{aligned} I_{jx}(\bar{z})x &= \int_0^1 L_{jx}[s]x(s)ds, \quad I_{ju}(\bar{z})u = \int_0^1 L_{ju}[s]u(s)ds, \\ F_x(\bar{z})x &= x - \int_0^{(\cdot)} \varphi_x[s]x(s)ds, \quad F_u(\bar{z})u = - \int_0^{(\cdot)} \varphi_u[s]u(s)ds, \\ \nabla G(\bar{z}) &= (G_x(\bar{z}), G_u(\bar{z})) = (g_x[\cdot], g_u[\cdot]), \\ \langle \nabla^2 I_j(\bar{z})z, z \rangle &= \int_0^1 (\nabla^2 L_j[s]z(s), z(s))ds, \\ \langle \nabla^2 F(\bar{z})z, z \rangle &= - \int_0^{(\cdot)} (\nabla^2 \varphi[s]z(s), z(s))ds \end{aligned}$$

for all $z = (x, u) \in Z$, and

$$\nabla^2 G(\bar{z}) = \begin{bmatrix} g_{xx}[\cdot] & g_{xu}[\cdot] \\ g_{ux}[\cdot] & g_{uu}[\cdot] \end{bmatrix}.$$

• Verification of (A2). Taking any $v \in E_0$, we consider equation $F_x(\bar{z})x = v$. This equation is equivalent to

$$x = \int_0^{(\cdot)} \varphi_x[s]x(s)ds + v.$$

By assumption (H1), we have $\varphi_x[\cdot] \in L^\infty([0, 1], \mathbb{R}^n)$. By [108, Lemma 1, p. 51], the equation has a unique solution $x \in X$. Hence (A2) is valid.

• Verification of (A3). Let $D := \{(x, u) \in Z \mid F(z) = 0\}$. Under assumption (A2), the mapping $\nabla F(\bar{z}): X \times U \rightarrow E_0$ is surjective. This implies that

$$T(D; \bar{z}) = \{(x, u) \in Z \mid F_x(\bar{z})x + F_u(\bar{z})u = 0\}.$$

Therefore, assumption (A3) is satisfied if we can show that for each $v \in E$, there exists $(x, u) \in Z$ such that

$$F_x(\bar{z})x + F_u(\bar{z})u = 0, \tag{3.2.11}$$

$$v = G_x(\bar{z})x + G_u(\bar{z})u - 0. \tag{3.2.12}$$

We will find u in the form $u = (0, 0, \dots, u_{i_0}, 0, \dots, 0)$. Consider the following equation

$$F_x(\bar{z})x + F_u(\bar{z})\frac{v - G_x(\bar{z})x}{g_{i_0u}[\cdot]} = 0.$$

This equation is equivalent to

$$x = \int_0^{(\cdot)} \left(\varphi_x[s] + \varphi_u[s] \frac{g_x[s]}{g_{i_0 u}[\cdot]} \right) x(s) ds + \int_0^{(\cdot)} \varphi_u[s] \frac{v}{g_{i_0 u}[s]} ds.$$

By (H2), $\varphi_u[\cdot] \frac{g_x[\cdot]}{g_{i_0 u}[\cdot]}$ and $\varphi_u[\cdot] \frac{v}{g_{i_0 u}[\cdot]}$ belong to $L^\infty([0, 1], \mathbb{R}^n)$. By [108, Lemma 1, p. 51], the above equation has a unique solution $x \in X$. Choosing $u = (0, 0, \dots, u_{i_0}, 0, \dots, 0)$ with

$$u_{i_0} = \frac{v - G_x(\bar{z})x}{g_{i_0 u}[\cdot]}.$$

We see that (x, u) satisfies equations (3.2.11)-(3.2.12). Hence assumption (A3) is fulfilled.

Step 2. Deriving optimality conditions.

Let $\mathcal{L}(z, \lambda, v^*, e^*) := \lambda I(x, u) + v^* F(x, u) + e^* G(x, u)$ be the Lagrangian associated with the (MCP). According to Theorem 3.2.1, for each $z = (\tilde{x}, \tilde{u}) \in \mathcal{C}(\bar{z})$, there exist multipliers $\lambda \in \mathbb{R}_+^m$ with $|\lambda| = 1$, $v^* \in E_0^*$ and $e^* \in E^*$ such that the following conditions are valid:

$$e^* \in N(Q, G(\bar{z})), \quad (3.2.13)$$

$$\lambda I_x(\bar{z}) + v^* F_x(\bar{z}) + e^* G_x(\bar{z}) = 0, \quad (3.2.14)$$

$$\lambda I_u(\bar{z}) + v^* F_u(\bar{z}) + e^* G_u(\bar{z}) = 0, \quad (3.2.15)$$

$$\langle \lambda \nabla^2 I(\bar{z})z, z \rangle + \langle v^* \nabla^2 F(\bar{z})z, z \rangle + \langle e^* \nabla^2 G(\bar{z})z, z \rangle \geq 0. \quad (3.2.16)$$

Here v^* is a signed Radon measure and e^* is a signed and finite additive measure on $[0, 1]$ which is absolutely continuous w.r.t the Lebesgue measure $|\cdot|$ on $[0, 1]$. By Riesz's Representation (see [108, Chapter 01, p. 19] and [102, Theorem 3.8, p. 73]), there exists a vector function of bounded variation ν , which is continuous from the right and vanish at zero such that

$$\langle v^*, y \rangle = \int_0^1 y(t) d\nu(t) \quad \forall y \in E_0,$$

where $\int_0^1 y(t) d\nu(t)$ is the Riemann-Stieltjes integral.

Define $\bar{p}: [0, 1] \rightarrow \mathbb{R}^n$ by setting

$$\bar{p}(t) := \nu((t, 1]) = \nu(t) - \nu(1).$$

Clearly, $\bar{p}(1) = 0$ and the function $\bar{p}$ is of bounded variation. By the

Fubini Theorem, for each $x \in C([0, 1], \mathbb{R}^n)$, we have

$$\begin{aligned}
\langle v^* F_x(\bar{z}), x \rangle &= \left\langle v^*, x - \int_0^{\cdot} \varphi_x[s] x(s) ds \right\rangle \\
&= \int_0^1 x^T(t) d\nu(t) - \int_0^1 \int_0^t \varphi_x[s] x(s) ds d\nu(t) \\
&= \int_0^1 x^T(t) d\nu(t) - \int_0^1 \varphi_x[s] x(s) ds \int_s^1 d\nu(t) \\
&= \int_0^1 x^T(t) d\nu(t) + \int_0^1 \varphi_x[s] x(s) \bar{p}(s) ds.
\end{aligned} \tag{3.2.17}$$

Similarly, for any $u \in L^\infty([0, 1], \mathbb{R}^l)$, we get

$$\langle v^* F_u(\bar{z}), u \rangle = \int_0^1 \int_0^t \varphi_u[s] u(s) ds d\nu(t) = \int_0^1 \bar{p}(s)^T \varphi_u[s] u(s) ds \tag{3.2.18}$$

and

$$\langle v^* \nabla^2 F(\bar{z}) z, z \rangle = \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[s] z(s), z(s)) ds. \tag{3.2.19}$$

From (3.2.15) and (3.2.18), we have

$$\int_0^1 \lambda L_u[s] u(s) ds + \int_0^1 \bar{p}(s)^T \varphi_u[s] u(s) ds + \langle e^*, G_u(\bar{z}) u \rangle = 0 \quad \forall u \in U. \tag{3.2.20}$$

Let us claim that e^* can be represented by a density in $L^1([0, 1], \mathbb{R})$. Indeed, take any $e \in E$. By (H2), there exists $u \in U$ such that $G_u(\bar{z})u = e$. Namely, $u = G_u(\bar{z})^{-1}e = (0, 0, \dots, u_{i_0}, 0, \dots, 0)$ with $u_{i_0} = \frac{e}{g_{u_{i_0}}[\cdot]}$. Inserting $G_u(\bar{z})u = e$ into (3.2.20), we have

$$\langle e^*, e \rangle = - \int_0^1 \sum_{j=1}^m \lambda_j L_j u_{i_0}[s] \frac{1}{g_{u_{i_0}}[\cdot]} e ds - \int_0^1 \sum_{j=1}^n p_j \varphi_{j, u_{i_0}} \frac{1}{g_{u_{i_0}}[\cdot]} e ds.$$

Since $\lambda_j L_j u_{i_0}[s] \frac{1}{g_{u_{i_0}}[\cdot]}$ and $p_j \varphi_{j, u_{i_0}} \frac{1}{g_{u_{i_0}}[\cdot]}$ belong to $L^1([0, 1], \mathbb{R})$, we conclude that, there exists a function $\theta \in L^1([0, 1], \mathbb{R})$ such that

$$\langle e^*, v \rangle = \int_0^1 \theta(t) v(t) dt \quad \forall v \in E. \tag{3.2.21}$$

Therefore, the claim is justified.

Based on the representation of e^* , (3.2.13), and (3.2.20), we obtain assertions (i) and (iii). Also, from (3.2.21), (3.2.17) and (3.2.14), we get

$$\begin{aligned} & \int_0^1 \lambda^T L_x[s] x(s) ds + \int_0^1 x^T(t) d\nu(t) + \int_0^1 \bar{p}(t)^T \varphi_x[s] x(s) ds \\ & + \int_0^1 \theta(s)^T g_x[s] x(s) ds = 0 \quad \forall x \in X. \end{aligned}$$

This is equivalent to

$$- \int_0^1 x^T(t) d\nu(t) = \int_0^1 (\lambda^T L_x[s] + \bar{p}(s)^T \varphi_x[s] + \theta(s)^T g_x[s]) x(s) ds \quad \forall x \in X. \quad (3.2.22)$$

We now fix any vector $\xi \in \mathbb{R}^n$ and $t \in [0, 1]$. Define $x_t(s) := \xi \chi_{(t,1]}(s)$, where $\chi_{(t,1]}(\cdot)$ is the indicator function of $(t, 1]$. Let us define

$$\vartheta(s) := \lambda^T L_x[s] + \bar{p}(t)^T \varphi_x[s] + \theta(s)^T g_x[s].$$

Then, $\vartheta(\cdot) \in L^1([0, 1], \mathbb{R}^n)$ and so are the functions $\vartheta(s)$ and $s\vartheta(s)$. By the Lebesgue Differentiation Theorem (see [225, Theorem 7.15]), these functions have Lebesgue points for a.e. on $[0, 1]$. Let us denote by P and P' the sets of Lebesgue points of $\vartheta(s)$ and $s\vartheta(s)$, respectively. Then $|P \cap P'| = 1$ and we have the following key lemma.

Lemma 3.2.4. *For each $t \in P \cap P'$, the following equality is valid:*

$$- \int_0^1 x_t^T(s) d\nu(s) = \int_0^1 (\lambda^T L_x[s] + \bar{p}(s)^T \varphi_x[s] + \theta(s)^T g_x[s]) x_t(s) ds.$$

Proof. Note that any function with bounded variation as well as any signed Radon measure can be represented as the difference of two increasing functions, and the difference of two positive Radon measures, respectively (see [225, Corollary 2.7] and [215, Lemma 13.6]). Therefore, we can assume that ν is increasing, right continuous and of bounded variation.

For each ϵ with $t < \epsilon < 1$, we define a function x_ϵ as follows:

$$x_\epsilon(s) := \begin{cases} \xi & \text{if } s \in [\epsilon, 1], \\ \frac{\xi(s-t)}{\epsilon-t} & \text{if } s \in [t, \epsilon], \\ 0 & \text{if } s \in [0, t]. \end{cases}$$

Then, $x_\epsilon \in C([0, 1], \mathbb{R}^n)$. By (3.2.22), we have

$$- \int_0^1 x_\epsilon^T(s) d\nu(s) = \int_0^1 (\lambda^T L_x[s] + \bar{p}(t)^T \varphi_x[s] + \theta(s)^T g_x[s]) x_\epsilon(s) ds,$$

or, equivalently,

$$-\int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} d\nu(s) - \int_\epsilon^1 \xi d\nu(s) = \int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} \vartheta(s) ds + \int_\epsilon^1 \xi \vartheta(s) ds. \quad (3.2.23)$$

By the Mean Value Theorem (see [225, Theorem 2.27, p. 33]), there is a point $t' \in [t, \epsilon]$ such that

$$\int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} d\nu(s) = \frac{\xi(t'-t)}{\epsilon-t} (\nu(\epsilon) - \nu(t)).$$

Hence

$$\begin{aligned} \left| \int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} d\nu(s) \right| &= \frac{|\xi|(t'-t)}{\epsilon-t} (\nu(\epsilon) - \nu(t)) \\ &\leq \frac{|\xi|(\epsilon-t)}{\epsilon-t} (\nu(\epsilon) - \nu(t)) \leq |\xi|(\nu(\epsilon) - \nu(t)). \end{aligned}$$

By letting $\epsilon \rightarrow t^+$ and using the right continuity of ν , we see that

$$\left| \int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} d\nu(s) \right| \rightarrow 0 \quad \text{as } \epsilon \rightarrow t^+. \quad (3.2.24)$$

Also, we have

$$\left| \int_t^1 \xi d\nu(s) ds - \int_\epsilon^1 \xi d\nu(s) \right| = \left| \int_t^\epsilon \xi d\nu(s) \right| \leq |\xi|(\nu(\epsilon) - \nu(t)) \rightarrow 0 \quad \text{as } \epsilon \rightarrow t^+.$$

Consequently,

$$\int_\epsilon^1 \xi d\nu(s) ds \rightarrow \int_t^1 \xi d\nu(s) \quad \text{as } \epsilon \rightarrow t^+.$$

For the first term of (3.2.23), we have from the Lebesgue Differentiation Theorem (see [225, Theorem 7.16]) that

$$\begin{aligned} \left| \int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} \vartheta(s) ds \right| &\leq |\xi| \frac{1}{\epsilon-t} \int_t^\epsilon |(s-t) \vartheta(s)| ds \\ &\leq |\xi| \frac{1}{\epsilon-t} \int_t^\epsilon |s \vartheta(s) - t \vartheta(t)| ds + |\xi| \frac{t}{\epsilon-t} \int_t^\epsilon |\vartheta(t) - \vartheta(s)| ds \rightarrow 0 \end{aligned}$$

as $\epsilon \rightarrow t^+$. Hence

$$\left| \int_t^\epsilon \frac{\xi(s-t)}{\epsilon-t} \vartheta(s) ds \right| \rightarrow 0 \quad \text{as } \epsilon \rightarrow t^+.$$

The convergence

$$\int_\epsilon^1 \xi^T \vartheta(s) ds \rightarrow \int_t^1 \xi^T \vartheta(s) ds \quad \text{as } \epsilon \rightarrow t^+ \quad (3.2.25)$$

is straightforward. Passing the limit both sides of (3.2.23) and using (3.2.24)–(3.2.25), we obtain

$$-\int_t^1 \xi^T d\nu(s) = \int_t^1 \xi^T \vartheta(s) ds.$$

The proof of the lemma is complete. $\square$

From Lemma 3.2.4, we have for a.e. $t \in [0, 1]$ that

$$-\int_t^1 \xi^T d\nu(s) = \int_t^1 (\lambda^T L_x[t] + \bar{p}(t)^T \varphi_x[t] + \theta(t) g_x[t]) \xi ds,$$

or, equivalently,

$$\xi^T \bar{p}(t) = \int_t^1 \xi^T (\lambda^T L_x[t] + \bar{p}(t)^T \varphi_x[t] + \theta(t) g_x[t]) ds.$$

Since ξ is arbitrary, we obtain

$$\begin{aligned} \dot{\bar{p}}(t) &= -\lambda^T L_x[t] - \bar{p}(t)^T \varphi_x[t] - \theta(t) g_x[t] \quad \text{a.e. } t \in [0, 1], \\ \bar{p}(1) &= 0, \end{aligned}$$

which is assertion (ii) of Theorem 3.2.3. Finally, from (3.2.19) and (3.2.16), we have

$$\begin{aligned} \int_0^1 \left(\sum_{j=1}^m \lambda_j \nabla^2 L_j[t] z(t), z(t) \right) dt + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t] z(t), z(t)) dt \\ + \int_0^1 (\theta(t) \nabla^2 g[t] z(t), z(t)) dt \geq 0, \end{aligned}$$

which is assertion (iv) of the theorem. The proof of Theorem 3.2.3 is complete. $\square$

In multi-objective optimization problems, the critical cone for second-order sufficient conditions is often required bigger than the one for second-order necessary conditions. Therefore, we need to enlarge $\mathcal{C}(\bar{z})$ to deal with second-order sufficient conditions. We denote by $\mathcal{C}'(\bar{z})$ the set of vectors $(x, u) \in C([0, 1], \mathbb{R}^n) \times L^2([0, 1], \mathbb{R}^l)$ which satisfy the following conditions:

$$\begin{aligned} (\mathcal{C}'_1) \quad & \int_0^1 (L_x[t]x(t) + L_u[t]u(t)) dt \in -\mathbb{R}_+^m; \\ (\mathcal{C}'_2) \quad & x(\cdot) = \int_0^{(\cdot)} (\varphi_x[s]x(s) + \varphi_u[s]u(s)) ds; \end{aligned}$$

(c'_3) $g_x[t]x(t) + g_u[t]u(t) \in T((-\infty, 0]; g[t])$ for a.e. $t \in [0, 1]$.

Obviously, $\mathcal{C}'(\bar{z})$ is a closed convex cone and $\mathcal{C}(\bar{z}) \subset \mathcal{C}'(\bar{z})$.

We now introduce the concept of local strong Pareto solution for the multi-objective optimal control problem (MCP1).

Definition 3.2.5. Let $\bar{z} = (\bar{x}, \bar{u}) \in \Phi$ be a feasible point of the (MCP1). We say that $\bar{z}$ is a local strong Pareto solution of the (MCP1) if there exist a number $\epsilon > 0$ and a vector $c \in \text{int } \mathbb{R}_+^m$ such that for all $(x, u) \in (B_X(\bar{x}, \epsilon) \times B_U(\bar{u}, \epsilon)) \cap \Phi$, one has

$$I(x, u) - I(\bar{x}, \bar{u}) - c\|u - \bar{u}\|_2^2 \notin -\text{int } \mathbb{R}_+^m.$$

By a standard argument, we can show that every local strong Pareto solution of the (MCP1) is also a local Pareto solution of this problem. Note that in (3.2.5) we use two norms to define local strong Pareto solutions. Here $B_U(\bar{u}, \epsilon)$ is a ball in $L^\infty([0, 1], \mathbb{R}^l)$ while $\|u - \bar{u}\|_2$ is the norm in $L^2([0, 1], \mathbb{R}^l)$.

The following theorem provides sufficient conditions for local strong Pareto solutions.

Theorem 3.2.6. Suppose that assumptions (H1) and (H2) are valid and $\bar{z} \in \Phi$. Suppose that there exist a vector $\lambda \in \mathbb{R}_+^m$ with $|\lambda| = 1$, an absolutely continuous function $\bar{p}: [0, 1] \rightarrow \mathbb{R}^n$ and a function $\theta \in L^2([0, 1], \mathbb{R})$ satisfy conditions (i), (ii) and (iii) of (3.2.3) and the strict second-order condition

$$\begin{aligned} & \int_0^1 (\lambda^T \nabla^2 L[t]z(t), z(t)) dt + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t]z(t), z(t)) dt \\ & + \int_0^1 (\theta(t) \nabla^2 g[t]z(t), z(t)) dt > 0, \quad \forall z = (x, u) \in \mathcal{C}'(\bar{z}) \setminus \{0\}. \end{aligned} \quad (3.2.26)$$

Furthermore, there is a number $\gamma_0 > 0$ such that

$$(\lambda^T L_{uu}[t] + \bar{p}(t)^T \varphi_{uu}[t] + \theta(t)^T g_{uu}[t])(\xi, \xi) \geq \gamma_0 |\xi|^2 \quad \forall \xi \in \mathbb{R}^l. \quad (3.2.27)$$

Then, $(\bar{x}, \bar{u})$ is a local strong Pareto solution of the (MCP1).

Proof. Let us define the Lagrangian $\mathcal{L}(z, \lambda, \bar{p}, \theta)$ associated with the (MCP) by setting

$$\mathcal{L}(z, \lambda, \bar{p}, \theta) := \lambda^T I(z) + \bar{p}^T F(z) + \theta G(z),$$

where

$$\begin{aligned}\lambda^T I(z) &:= \int_0^1 \lambda^T L(s, x(s), u(s)) ds, \\ \bar{p}^T F(z) &:= \int_0^1 \dot{\bar{p}}(s) x(s) ds + \int_0^1 \bar{p}(s) \varphi(s, x(s), u(s)) ds, \\ \theta G(z) &:= \int_0^1 \theta(s) g(s, x(s), u(s)) ds.\end{aligned}$$

Then, from conditions (ii) and (iii) of Theorem 3.2.3, we can show that

$$\nabla_z \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta) = 0. \quad (3.2.28)$$

We now return to the proof of the theorem. Suppose the theorem was false. Then, we could find sequences $\{(x_k, u_k)\} \subset \Phi$ and $\{c_k\} \subset \text{int } \mathbb{R}_+^m$ such that $(x_k, u_k) \rightarrow (\bar{x}, \bar{u})$, $c_k \rightarrow 0$ and

$$I(x_k, u_k) - I(\bar{x}, \bar{u}) - c_k \|u_k - \bar{u}\|_2^2 \in -\text{int } \mathbb{R}_+^m. \quad (3.2.29)$$

Clearly, $(x_k, u_k) \neq (\bar{x}, \bar{u})$ for all $k \in \mathbb{Z}_+$. By replacing the sequence $\{(x_k, u_k)\}$ by a subsequence we may assume that $u_k = \bar{u}$ or $u_k \neq \bar{u}$ for all $k \in \mathbb{Z}_+$. If $u_k = \bar{u}$ for all $k \in \mathbb{Z}_+$, then we have

$$x_k(t) = x_0 + \int_0^t \varphi(s, x_k(s), \bar{u}(s)) ds$$

and

$$\bar{x}(t) = x_0 + \int_0^t \varphi(s, \bar{x}(s), \bar{u}(s)) ds.$$

Hence,

$$x_k(t) - \bar{x}(t) = \int_0^t (\varphi(s, x_k(s), \bar{u}(s)) - \varphi(s, \bar{x}(s), \bar{u}(s))) ds.$$

From this and (H1), there exist numbers $M > 0$ and $k_{\varphi M} > 0$ such that for k large enough, we have

$$|\varphi(s, x_k(s), \bar{u}(s)) - \varphi(s, \bar{x}(s), \bar{u}(s))| \leq k_{\varphi M} |x_k(t) - \bar{x}(t)|.$$

Hence,

$$|x_k(t) - \bar{x}(t)| \leq \int_0^t k_{\varphi M} |x_k(s) - \bar{x}(s)| ds.$$

Using the Gronwall Inequality (see [51, 18.1.i, p. 503]), we get $x_k = \bar{x}$, a contradiction. Therefore, we have that $u_k \neq \bar{u}$ for all $k \in \mathbb{N}$.

Define $t_k := \|u_k - \bar{u}\|_2$, $\hat{x}_k := \frac{x_k - \bar{x}}{t_k}$ and $\hat{u}_k := \frac{u_k - \bar{u}}{t_k}$. Then $t_k \rightarrow 0^+$ and $\|\hat{u}_k\|_2 = 1$. Since $L^2([0, 1], \mathbb{R}^l)$ is reflexive, we may assume that $\hat{u}_k \rightharpoonup \hat{u}$. From the above, we have

$$\lambda^T I(z_k) - \lambda^T I(\bar{z}) \leq t_k^2 \lambda^T c_k \leq t_k^2 |\lambda| |c_k| \leq o(t_k^2). \quad (3.2.30)$$

We claim that $\hat{x}_k$ converges uniformly to some $\hat{x}$ in $C([0, 1], \mathbb{R}^n)$. In fact, since $(x_k, u_k) \in \Phi$, we have

$$x_k(t) = x_0 + \int_0^t \varphi(s, x_k(s), u_k(s)) ds.$$

Since $x_k = \bar{x} + t_k \hat{x}_k$, we have

$$t_k \hat{x}_k(t) = \int_0^t (\varphi(s, x_k(s), u_k(s)) - \varphi(s, \bar{x}(s), \bar{u}(s))) ds. \quad (3.2.31)$$

Since $x_k \rightarrow \bar{x}$ uniformly and $u_k \rightarrow \bar{u}$ in $L^\infty([0, 1], \mathbb{R}^l)$, there exists a constant $\varrho > 0$ such that $\|x_k\|_0 \leq \varrho$, $\|u_k\|_\infty \leq \varrho$. By assumption (H1), there exists $k_{\varphi, \varrho} > 0$ such that

$$|\varphi(s, x_k(s), u_k(s)) - \varphi(s, \bar{x}(s), \bar{u}(s))| \leq k_{\varphi, \varrho} (|x_k(s) - \bar{x}(s)| + |u_k(s) - \bar{u}(s)|)$$

for a.e. $s \in [0, 1]$. Hence we have from (3.2.31) that

$$|\hat{x}_k(t)| \leq \int_0^t k_{\varphi, \varrho} (|\hat{x}_k(s)| + |\hat{u}_k(s)|) ds$$

and

$$|\dot{\hat{x}}_k(t)| \leq k_{\varphi, \varrho} (|\hat{x}_k(t)| + |\hat{u}_k(t)|). \quad (3.2.32)$$

It follows that

$$\begin{aligned} |\hat{x}_k(t)| &\leq \int_0^t k_{\varphi, \varrho} |\hat{x}_k(s)| ds + \int_0^1 k_{\varphi, \varrho} |\hat{u}_k(s)| ds \\ &\leq \int_0^t k_{\varphi, \varrho} |\hat{x}_k(s)| ds + k_{\varphi, \varrho} \left(\int_0^1 |\hat{u}_k(s)|^2 ds \right)^{1/2} \\ &\leq \int_0^t k_{\varphi, \varrho} |\hat{x}_k(s)| ds + k_{\varphi, \varrho}, \quad (\|\hat{u}_k\|_2 = 1). \end{aligned}$$

Using the Gronwall Inequality, we have

$$|\hat{x}_k(t)| \leq k_{\varphi, \varrho} \exp(k_{\varphi, \varrho}).$$

From this and (3.2.32), we see that

$$|\dot{\hat{x}}_k(t)|^2 \leq 2k_{\varphi, \varrho}^2 (|\hat{x}_k(t)|^2 + |\hat{u}_k(t)|^2) \leq 2k_{\varphi, \varrho}^2 (k_{\varphi, \varrho}^2 \exp(2k_{\varphi, \varrho}) + |\hat{u}_k|^2).$$

Hence,

$$\int_0^1 |\dot{\hat{x}}_k(t)|^2 dt \leq 2k_{\varphi,\varrho}^2 (k_{\varphi,\varrho}^2 \exp(2k_{\varphi,\varrho}) + 1).$$

Consequently, $\{\hat{x}_k\}$ is bounded in $W^{1,2}([0, 1], \mathbb{R}^n)$. By passing subsequence, we can assume that $\hat{x}_k \rightharpoonup \hat{x}$ weakly in $W^{1,2}([0, 1], \mathbb{R}^n)$. By [31, Theorem 8.8], the embedding $W^{1,2}([0, 1], \mathbb{R}^n) \hookrightarrow C([0, 1], \mathbb{R}^n)$ is compact. Hence we have that $\hat{x}_k \rightarrow \hat{x}$ uniformly on $[0, 1]$. The claim is justified. The remains of the proof is divided into some steps.

Step 1. Showing that $(\hat{x}, \hat{u}) \in \mathcal{C}'(\bar{z})$.

By the Taylor expansion, we have from (3.2.29) that

$$I_x(\bar{z})\hat{x}_k + I_u(\bar{z})\hat{u}_k + \frac{o(t_k)}{t_k} \in -\mathbb{R}_+^m. \quad (3.2.33)$$

Note that $L_{ju}[\cdot] \in L^\infty([0, 1], \mathbb{R}^l)$ and $I_{ju}(\bar{z}): L^2([0, 1], \mathbb{R}^l) \rightarrow \mathbb{R}$ is a continuous linear mapping, where

$$\langle I_{ju}(\bar{z}), u \rangle := \int_0^1 L_{ju}[s]u(s)ds \quad \forall u \in L^2([0, 1], \mathbb{R}^l).$$

By [31, Theorem 3.10], $I_{ju}(\bar{z})$ is weakly continuous on $L^2([0, 1], \mathbb{R}^l)$. By letting $k \rightarrow \infty$ in (3.2.33), we get

$$I_x(\bar{z})\hat{x} + I_u(\bar{z})\hat{u} \in -\mathbb{R}_+^m. \quad (3.2.34)$$

Since $F(\bar{z}) = 0$, $F(x_k, u_k) = 0$ and by the Taylor expansion, we have

$$F_x(\bar{z})\hat{x}_k + F_u(\bar{z})\hat{u}_k + \frac{o(t_k)}{t_k} = 0.$$

By the same arguments as the above and letting $k \rightarrow \infty$, we obtain

$$F_x(\bar{z})\hat{x} + F_u(\bar{z})\hat{u} = 0. \quad (3.2.35)$$

Since $G(x_k, u_k) - G(\bar{x}, \bar{u}) \in Q - G(\bar{x}, \bar{u})$ and by the Taylor expansion, we have

$$G_x(\bar{z})\hat{x}_k + G_u(\bar{z})\hat{u}_k + \frac{o(t_k)}{t_k} \in \text{cone}(Q - G(\bar{x}, \bar{u})) \subset T(Q; G(\bar{x}, \bar{u})),$$

where $T(Q; G(\bar{x}, \bar{u}))$ is the tangent cone to Q at $G(\bar{x}, \bar{u})$ in $L^\infty([0, 1], \mathbb{R})$. It is easily seen that

$$\begin{aligned} T(Q; G(\bar{x}, \bar{u})) &\subseteq \{v \in L^\infty([0, 1], \mathbb{R}) \mid v(t) \in T((-\infty, 0], g[t]) \quad \text{a.e.}\} \\ &\subseteq \{v \in L^2([0, 1], \mathbb{R}) \mid v(t) \in T((-\infty, 0], g[t]) \quad \text{a.e.}\}. \end{aligned}$$

Hence

$$G_x(\bar{z})\hat{x}_k + G_u(\bar{z})\hat{u}_k + \frac{o(t_k)}{t_k} \in \{v \in L^2([0, 1], \mathbb{R}) \mid v(t) \in T((-\infty, 0], g[t]) \text{ a.e.}\} . \quad (3.2.36)$$

Note that

$$\{v \in L^2([0, 1], \mathbb{R}) \mid v(t) \in T((-\infty, 0], g[t]) \text{ a.e.}\} = T_{L^2}(Q; G(\bar{x}, \bar{u})),$$

where $T_{L^2}(Q, G(\bar{x}, \bar{u}))$ is the tangent cone to the set Q at $G(\bar{x}, \bar{u})$ in $L^2([0, 1], \mathbb{R})$. Since $T_{L^2}(Q; G(\bar{x}, \bar{u}))$ is a closed convex set of $L^2([0, 1], \mathbb{R})$, it is also a weakly closed set of $L^2([0, 1], \mathbb{R})$. Since

$$G_u(\bar{z}): L^2([0, 1], \mathbb{R}^l) \rightarrow L^2([0, 1], \mathbb{R})$$

is a continuous linear mapping, [31, Theorem 3.10] implies that it is continuous from weak topology of $L^2([0, 1], \mathbb{R}^l)$ to weak topology of $L^2([0, 1], \mathbb{R})$. By passing the limit in (3.2.36) when $k \rightarrow \infty$, we obtain

$$G_x(\bar{z})\hat{x} + G_u(\bar{z})\hat{u} \in \{v \in L^2([0, 1], \mathbb{R}) \mid v(t) \in T((-\infty, 0]; g[t]) \text{ a.e.}\} .$$

Combining this with (3.2.34) and (3.2.35), we get $(\hat{x}, \hat{u}) \in \mathcal{C}'(\bar{z})$.

Step 2. Showing that $(\hat{x}, \hat{u}) = 0$.

By the second-order Taylor expansion for $\mathcal{L}$ and (3.2.28), we get

$$\mathcal{L}(z_k, \lambda, \bar{p}, \theta) - \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta) = \frac{t_k^2}{2} \nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta)(\hat{z}_k, \hat{z}_k) + o(t_k^2),$$

where $\hat{z}_k := (\hat{x}_k, \hat{u}_k)$. On the other hand from (3.2.30), we have

$$\mathcal{L}(z_k, \lambda, \bar{p}, \theta) - \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta) = \lambda^T (I(z_k) - I(\bar{z})) + \langle \theta, G(z_k) - G(\bar{z}) \rangle \leq o(t_k^2).$$

Here we used the fact that $\theta \in N(Q, G(\bar{z}))$ and $F(z_k) = F(\bar{z}) = 0$. Therefore, we have

$$\frac{t_k^2}{2} \nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta)(\hat{z}_k, \hat{z}_k) + o(t_k^2) \leq o(t_k^2),$$

or, equivalently,

$$\nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta)(\hat{z}_k, \hat{z}_k) \leq \frac{o(t_k^2)}{t_k^2}. \quad (3.2.37)$$

By letting $k \rightarrow \infty$, we obtain

$$\nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta)(\hat{z}, \hat{z}) \leq 0.$$

By a simple calculation, we have

$$\begin{aligned} & \int_0^1 (\lambda^T \nabla^2 L[t] \hat{z}(t), \hat{z}(t)) dt + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t] \hat{z}(t), \hat{z}(t)) dt \\ & + \int_0^1 (\theta(t) \nabla^2 g[t] \hat{z}(t), \hat{z}(t)) dt = \nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta)(\hat{z}, \hat{z}) \leq 0. \end{aligned}$$

Combining this with (3.2.26), we must have $\hat{z} = 0$.

Step 3. Showing a contradiction.

From (3.2.27) and (3.2.37), we have

$$\begin{aligned} & \frac{o(t_k^2)}{t_k^2} \geq \nabla_{zz}^2 \mathcal{L}(\bar{z}, \lambda, \bar{p}, \theta)(\hat{z}_k, \hat{z}_k) \\ & = \int_0^1 (\lambda^T \nabla^2 L[t] \hat{z}_k(t), \hat{z}_k(t)) dt + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t] \hat{z}_k(t), \hat{z}_k(t)) dt \\ & \quad + \int_0^1 (\theta(t) \nabla^2 g[t] \hat{z}_k(t), \hat{z}_k(t)) dt \\ & = \int_0^1 \lambda^T L_{uu}[t] \hat{u}_k^2(t) dt + 2 \int_0^1 \lambda^T L_{xu}[t] \hat{x}_k(t) \hat{u}_k(t) dt + \int_0^1 \lambda^T L_{xx}[t] \hat{x}_k^2(t) dt + \\ & \quad + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t] \hat{z}_k(t), \hat{z}_k(t)) dt + \int_0^1 (\theta(t)^T \nabla^2 g[t] \hat{z}_k(t), \hat{z}_k(t)) dt \\ & \geq \gamma_0 + \int_0^1 2 (\lambda^T L_{xu}[t] \hat{x}_k(t) \hat{u}_k(t) + \lambda^T L_{xx}[t] \hat{x}_k^2(t)) dt + \\ & \quad + \int_0^1 (2\bar{p}(t)^T \varphi_{xu}[t] \hat{x}_k(t) \hat{u}_k(t) + \bar{p}(t)^T \varphi_{xx}[t] \hat{x}_k^2(t)) dt \\ & \quad + \int_0^1 (2\theta(t)^T g_{xu}[t] \hat{x}_k(t) \hat{u}_k(t) + \theta(t)^T g_{xx}[t] \hat{x}_k^2(t)) dt. \end{aligned}$$

By letting $k \rightarrow \infty$ and using the fact $\hat{x}_k \rightarrow 0$ uniformly, we obtain $0 \geq \gamma_0$, which is impossible. The proof of Theorem 3.2.6 is complete. $\square$

Next, we give some examples to illustrate the main results. The first example shows us how to use Theorem 3.2.3 and Theorem 3.2.6 to obtain solutions of the (MCP1). The second one indicates the important role of the second-order necessary optimality conditions in checking optimal solutions.

Example 3.2.7. Consider the problem (MCP1), where

$$\begin{aligned} L(t, x(t), u(t)) &= (x_1^2(t) + u_1^2(t), x_2^2(t) + u_2^2(t)), \\ \varphi(t, x(t), u(t)) &= (u_1(t), u_2(t)), \\ x_0 &= (0, 0), \\ g(t, x(t), u(t)) &= x_1(t) + x_2(t) - u_1(t) - u_2(t) \end{aligned}$$

for all $x(t) = (x_1(t), x_2(t))$, $u(t) = (u_1(t), u_2(t))$ and $t \in [0, 1]$. Then, the feasible solution set of the (MCP) is

$$\Phi = \left\{ (x, u) \in C([0, 1], \mathbb{R}^2) \times L^\infty([0, 1], \mathbb{R}^2) \mid x(t) = \int_0^t u(s) ds, \right. \\ \left. x_1(t) + x_2(t) - u_1(t) - u_2(t) \leq 0 \text{ a.e. } t \in [0, 1] \right\}.$$

It is easy to see that conditions (H1) and (H2) are valid. We will use conditions (i)–(iii) of Theorem 3.2.3 to find out KKT points of the (MCP1) which are good candidates for optimal solutions. Assume that $\bar{z} = (\bar{x}, \bar{u})$ is a feasible solution of the (MCP1) and satisfies conditions (i)–(iii) of Theorem 3.2.3 with respect to $(\lambda, \bar{p}, \theta)$. By simple computations, we have

$$\begin{aligned} L_{1x}[t] &= (2\bar{x}_1, 0)^T, L_{2x}[t] = (0, 2\bar{x}_2)^T, \varphi_{1x}[t] = \varphi_{2x}[t] = (0, 0)^T, g_x[t] = (1, 1)^T, \\ L_{1u}[t] &= (2\bar{u}_1, 0)^T, L_{2u}[t] = (0, 2\bar{u}_2)^T, \varphi_{1u}[t] = (1, 0)^T, \varphi_{2u}[t] = (0, 1)^T, \\ g_u[t] &= (-1, -1)^T. \end{aligned}$$

From conditions (ii) and (iii), we get

$$\begin{cases} \dot{\bar{p}}_1 = -2\lambda_1\bar{x}_1 - \theta, \\ \dot{\bar{p}}_2 = -2\lambda_2\bar{x}_2 - \theta, \\ \bar{p}_1(1) = \bar{p}_2(1) = 0, \end{cases} \quad (3.2.38)$$

and

$$\begin{cases} 2\lambda_1\bar{u}_1 + \bar{p}_1 - \theta = 0, \\ 2\lambda_2\bar{u}_2 + \bar{p}_2 - \theta = 0. \end{cases} \quad (3.2.39)$$

Combining (3.2.38) and (3.2.39) yields

$$\begin{cases} \dot{\bar{p}}_1 = -2\lambda_1\bar{x}_1 - 2\lambda_1\bar{u}_1 - \bar{p}_1, \\ \dot{\bar{p}}_2 = -2\lambda_2\bar{x}_2 - 2\lambda_2\bar{u}_2 - \bar{p}_2. \end{cases} \quad (3.2.40)$$

By condition (3.1.2), one has

$$\begin{cases} \dot{\bar{x}}_1 = \bar{u}_1, \\ \dot{\bar{x}}_2 = \bar{u}_2. \end{cases} \quad (3.2.41)$$

Then inserting these equations into (3.2.40), we obtain

$$\begin{cases} \dot{\bar{p}}_1 = -2\lambda_1\bar{x}_1 - 2\lambda_1\dot{\bar{x}}_1 - \bar{p}_1, \\ \dot{\bar{p}}_2 = -2\lambda_2\bar{x}_2 - 2\lambda_2\dot{\bar{x}}_2 - \bar{p}_2, \end{cases} \Leftrightarrow \begin{cases} \dot{\bar{p}}_1 + 2\lambda_1\dot{\bar{x}}_1 = -\bar{p}_1 - 2\lambda_1\bar{x}_1, \\ \dot{\bar{p}}_2 + 2\lambda_2\dot{\bar{x}}_2 = -\bar{p}_2 - 2\lambda_2\bar{x}_2. \end{cases}$$

This implies that

$$\begin{cases} \bar{p}_1 + 2\lambda_1\bar{x}_1 = c_1 \exp(-t), \\ \bar{p}_2 + 2\lambda_2\bar{x}_2 = c_2 \exp(-t), \end{cases} \quad (3.2.42)$$

where $c_1, c_2 \in \mathbb{R}$ are constants. Hence,

$$\bar{p}_1 - \bar{p}_2 + 2\lambda_1\bar{x}_1 - 2\lambda_2\bar{x}_2 = c_3 \exp(-t), \quad (3.2.43)$$

where $c_3 := c_1 - c_2$. From (3.2.39) and (3.2.41), we have

$$\bar{p}_1 - \bar{p}_2 = -2\lambda_1\dot{\bar{x}}_1 + 2\lambda_2\dot{\bar{x}}_2.$$

Inserting this equation into (3.2.43), we get

$$-2\lambda_1\dot{\bar{x}}_1 + 2\lambda_2\dot{\bar{x}}_2 + 2\lambda_1\bar{x}_1 - 2\lambda_2\bar{x}_2 = c_3 \exp(-t), \quad (3.2.44)$$

Let $\alpha = 2\lambda_1\bar{x}_1 - 2\lambda_2\bar{x}_2$. Then equation (3.2.44) becomes

$$-\dot{\alpha} + \alpha = c_3 \exp(-t).$$

From this equation and $\alpha(0) = 0$, it is easily seen that

$$\alpha(t) = \frac{1}{2}c_3 \exp(-t) - \frac{1}{2}c_3 \exp(t) \quad \forall t \in [0, 1].$$

This and (3.2.41) imply that

$$\begin{cases} 2\lambda_1\bar{x}_1 - 2\lambda_2\bar{x}_2 = \frac{1}{2}c_3 \exp(-t) - \frac{1}{2}c_3 \exp(t), \\ 2\lambda_1\bar{u}_1 - 2\lambda_2\bar{u}_2 = -\frac{1}{2}c_3 \exp(-t) - \frac{1}{2}c_3 \exp(t). \end{cases} \quad (3.2.45)$$

We see that, for each $\lambda \in \mathbb{R}_+^2 \setminus \{0\}$, every solution $(\bar{x}, \bar{u})$ of system (3.2.45) is a KKT point of the (MCP1). To illustrate Theorem 3.2.3, let us verify condition (iv) at a solution of the system (3.2.45). Let $\bar{x} = (0, 0)$, $\bar{u} = (0, 0)^T$. By (3.2.42) and $\bar{p}(1) = (0, 0)$, we have $c_1 = 0$, $c_2 = 0$, and $\bar{p}(t) = (0, 0)$ for all $t \in [0, 1]$. Consequently, $c_3 = 0$ and

$\theta(t) = 0$ for all $t \in [0, 1]$. Thus, $(\bar{x}, \bar{u})$ is a solution of the system (3.2.45) for every $\lambda \in \mathbb{R}_+^2 \setminus \{0\}$. By simple calculation, we get

$$\begin{aligned} \int_0^1 \left(\sum_{j=1}^m \lambda_j \nabla^2 L_j[t] z(t), z(t) \right) dt &= 2\lambda_1 \int_0^1 (x_1^2(t) + u_1^2(t)) dt \\ &\quad + 2\lambda_2 \int_0^1 (x_2^2(t) + u_2^2(t)) dt \geq 0 \end{aligned}$$

for all $z = (x, u) \in X \times U$ and $\lambda \in \mathbb{R}_+^2 \setminus \{0\}$. Hence, condition (iv) is satisfied.

We now use Theorem 3.2.6 to show that $(\bar{x}, \bar{u})$ is a local strong Pareto solution of the (MCP1). Let $\lambda = (\frac{1}{2}, \frac{1}{2})$, $\bar{p} = (0, 0)$ and $\theta = 0$. Then, we have

$$\begin{aligned} &\int_0^1 \left(\sum_{j=1}^m \lambda_j \nabla^2 L_j[t] z(t), z(t) \right) dt \\ &= \int_0^1 (x_1^2(t) + u_1^2(t)) dt + \int_0^1 (x_2^2(t) + u_2^2(t)) dt > 0 \end{aligned}$$

for all $z = (x, u) \in X \times U \setminus \{(0, 0)\}$. Hence condition (3.2.26) is satisfied. Furthermore, we see that

$$\lambda^T L_{uu}[t](\xi, \xi) = |\xi|^2$$

for all $\xi \in \mathbb{R}^2$. This implies that condition (3.2.27) holds at $\bar{z}$ with respect to $\gamma_0 = 1$. By Theorem 3.2.6, $\bar{z}$ is a local strong Pareto solution of the (MCP1).

Example 3.2.8. Let φ and g be as in Example 3.2.7 and L be defined by

$$L(t, x(t), u(t)) = (x_1^2(t) - u_1^2(t), x_2^2(t) - u_2^2(t))$$

for all $x(t) = (x_1(t), x_2(t))$, $u(t) = (u_1(t), u_2(t))$ and $t \in [0, 1]$. Clearly, $(\bar{x}, \bar{u}) = ((0, 0), (0, 0))$ is a feasible point of the (MCP1). We claim that $(\bar{x}, \bar{u})$ is not a local weak Pareto solution of the (MCP1). Indeed, if otherwise, then due to Theorem 3.2.3, for each critical direction $z \in \mathcal{C}(\bar{z})$, there exist multipliers λ , $\bar{p}$ and θ such that conditions (i)–(iv) are fulfilled. Let $\tilde{x}(t) = (t, t)^T$ and $\tilde{u}(t) = (1, 1)^T$ for all $t \in [0, 1]$. It is easy to check that $z := (\tilde{x}, \tilde{u})$ is a critical direction of the (MCP1) at $\bar{z}$. By conditions (ii) and (iii) of Theorem 2.1, we get

$$\begin{cases} \dot{\bar{p}} = -\theta(1, 1)^T, \\ \bar{p}(1) = 0, \\ \bar{p} + \theta(-1, -1)^T = 0, \end{cases} \Leftrightarrow \begin{cases} \dot{\bar{p}}_1 = \dot{\bar{p}}_2 = -\theta, \\ \bar{p}(1) = 0, \\ \bar{p}_1 = \bar{p}_2 = \theta. \end{cases}$$

This implies that $\bar{p}_1(t) = \bar{p}_2(t) = \theta(t) = 0$ for all $t \in [0, 1]$. Since $\lambda \in \mathbb{R}_+^2 \setminus \{0\}$, $\bar{p} = (0, 0)^T$ and $\theta = 0$, we have

$$\begin{aligned} & \int_0^1 (\lambda_1 \nabla^2 L_1[t]z(t) + \lambda_2 \nabla^2 L_2[t]z(t), z(t)) dt + \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[t]z(t), z(t)) dt \\ & + \int_0^1 (\theta(t) \nabla^2 g[t]z(t), z(t)) dt \\ & = \int_0^1 (\lambda_1 \nabla^2 L_1[t]z(t) + \lambda_2 \nabla^2 L_2[t]z(t), z(t)) dt \\ & = \int_0^1 (\lambda_1 + \lambda_2)(2t^2 - 2)dt = -\frac{4}{3}(\lambda_1 + \lambda_2) < 0, \end{aligned}$$

which does not satisfy condition (iv) of Theorem 3.2.3. Hence, $(\bar{x}, \bar{u})$ is not a local weak Pareto solution of the (MCP1).

3.2.2 Second-order optimality conditions of KKT-type for multiobjective optimal control problems with endpoint and mixed constraints

We investigate the following multiobjective optimal control problem: find a control vector $u \in L^\infty([0, 1], \mathbb{R}^m)$ and the corresponding state $x \in C([0, 1], \mathbb{R}^n)$ which solve

$$\min_{\mathbb{R}_+^k} J(x, u) \tag{3.2.46}$$

s.t.

$$x(t) = x_0 + \int_0^t \varphi(s, x(s), u(s)) ds, \quad \text{for all } t \in [0, 1], \tag{3.2.47}$$

$$h(x(1)) \leq 0, \tag{3.2.48}$$

$$g(t, x(t), u(t)) \leq 0, \quad \text{for almost all } t \in [0, 1], \tag{3.2.49}$$

where the objective function $J(x, u)$ is defined component-wise as

$$J(x, u) = \left(J_1(x, u), J_2(x, u), \dots, J_k(x, u) \right),$$

with each component $J_i(x, u)$ given by

$$J_i(x, u) := \ell_i(x(1)) + \int_0^1 L_i(t, x(t), u(t)) dt, \quad i = 1, 2, \dots, k,$$

where $L_i: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$, $\ell_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $\varphi: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $g_j: [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, for $i = 1, 2, \dots, k$ and

$j = 1, 2, \dots, r$. The functions $h(x(1))$ and $g(t, x, u)$ are also defined component-wise as:

$$\begin{aligned} h(x(1)) &= (h_1(x(1)), \dots, h_n(x(1))), \\ g(t, x, u) &= (g_1(t, x, u), \dots, g_r(t, x, u)). \end{aligned}$$

For vectors $y, z \in \mathbb{R}^d$, the notation $y \leq z$ implies that $z - y \in \mathbb{R}_+^d$, where

$$\mathbb{R}_+^d = \{(\xi_1, \xi_2, \dots, \xi_d) \in \mathbb{R}^d : \xi_i \geq 0 \text{ for } i = 1, 2, \dots, d\}.$$

We denote by (MCP2) the problem (3.2.46)-(3.2.49). Our primary objective is to determine a control vector function $u \in L^\infty([0, 1], \mathbb{R}^m)$ that steers the state $x \in C([0, 1], \mathbb{R}^n)$ from an initial state $x(0) = x_0$ to a terminal state $x(1)$ optimally, in the sense of Pareto optimality.

Multi-objective optimal control problems (MOCPs) have garnered significant attention due to their numerous applications in modern technology (see, for instance, [181]). As a result, MOCPs have been extensively studied by various researchers in recent years. For works related to the present study, we direct readers to [71, 72, 117, 128, 132, 133, 136, 170, 173–175, 240] and the references therein. The main research topics in optimal control encompass the existence of optimal solutions, optimality conditions, and numerical methods for computing optimal solutions. In particular, the study of optimality conditions for multi-objective optimal control problems, as well as optimization problems in general, is a fundamental topic in optimization theory. First- and second-order optimality conditions of KKT type play a crucial role in numerical methods and the stability analysis of optimal solutions (see, for instance, [32] and [90]).

In previous section we established first- and second-order optimality conditions of KKT type for locally optimal solutions in the sense of Pareto for a class of multi-objective time optimal problems with free right endpoints (see [136]). It is well-known that in optimal control problems without endpoint constraints, the Robinson constraint qualification is relatively easy to verify, allowing for the establishment of KKT-type optimality conditions. However, when problems include endpoint constraints, the situation becomes more complex, and verifying the Robinson constraint qualification becomes challenging. Consequently, establishing KKT-type optimality conditions becomes more difficult.

In the scalar case, several papers have addressed the normality of multipliers for problems with endpoint constraints. In [85, Theorem 2],

H. Frankowska presented a constraint qualification under which the necessary optimality conditions are normal. This constraint qualification is significant and has been studied in [21, Theorem 9.3] for mathematical programming problems. Additionally, V. Zeidan [235] provided sufficient conditions for the normality of multipliers, demonstrating that if the linearized equation is controllable, then normality of multipliers is achieved.

The aim of this section is to derive first- and second-order necessary optimality conditions of KKT type, as well as second-order sufficient conditions for (MCP2) with endpoint constraints and mixed constraints. To address this problem, we first reduce the problem into a specific vector optimization problem (VOP2) in Chapter 2. We then derive optimality conditions for problem (MCP2) by utilizing the results obtained for (VOP2). We demonstrate that if the linearized equation is controllable, then the Robinson constraint qualification is valid, ensuring that the multipliers are normal and the optimality conditions are of KKT type.

It is worth noting that although the abstract VOP serves as a good model for establishing optimality conditions for MOCPs, we cannot directly derive optimality conditions for MOCPs from those of VOP. Specifically, in MOCPs, the control variables typically belong to L^∞ , and consequently, the Lagrange multipliers corresponding to mixed constraints (3.1.4) belong to $(L^\infty)^*$, which are additive measures rather than functions. Therefore, we require specific techniques and conditions to represent a functional on $(L^\infty)^*$ by an L^1 -function. In [71] and [72], A.V. Dmitruk and N.P. Osmolovskii imposed so-called regularity conditions on mixed constraints and employed the Yosida-Hewitt Theorem to represent a functional on $(L^\infty)^*$ by an L^1 -function. However, in our opinion, these conditions are not easily verifiable. In our paper, we impose alternative conditions on mixed constraints that are more readily checkable. We then introduce a new technique to represent a functional on L^∞ by an L^1 -function and a novel method to derive adjoint equations.

Throughout this section, we assume that Φ_2 is the feasible set of (MCP2) and let

$$X = C([0, 1], \mathbb{R}^n), \quad U = L^\infty([0, 1], \mathbb{R}^m), \quad (3.2.50)$$

$$K = \{v \in L^\infty([0, 1], \mathbb{R}^r) | v(t) \leq 0 \text{ a.a. } t \in [0, 1]\}. \quad (3.2.51)$$

Let ψ stands for L, φ and g , where $L = (L_1, \dots, L_k)$, $g = (g_1, \dots, g_r)$ and $\varphi = (\varphi_1, \dots, \varphi_n)$. Let Ω be an open set in $\mathbb{R}^n \times \mathbb{R}^m$ such that $(\bar{x}(t), \bar{u}(t)) \in \Omega$ for all $t \in [0, 1]$. Given $(\bar{x}, \bar{u}) \in \Phi_2$, the symbols

$L[t], \varphi[t], g[t], L_x[t], \varphi_u[t], g_u[t]$ and so on, stand for

$$\begin{aligned} &L(t, \bar{x}(t), \bar{u}(t)), \varphi(t, \bar{x}(t), \bar{u}(t)), g(t, \bar{x}(t), \bar{u}(t)), \\ &L_x(t, \bar{x}(t), \bar{u}(t)), \varphi_u(t, \bar{x}(t), \bar{u}(t)), g_u(t, \bar{x}(t), \bar{u}(t)), \dots \end{aligned}$$

We shall denote by $\ell'(\bar{x}(1))$ and $h'(\bar{x}(1))$ the first-order derivatives of ℓ and h at $\bar{x}(1)$, respectively. Also, by $\ell''(\bar{x}(1))$ and $h''(\bar{x}(1))$ the second-order derivatives of ℓ and h at $\bar{x}(1)$, respectively. We impose the following assumptions.

(H1) The function ψ is a Carathéodory function on $[0, 1] \times \Omega$ and $\psi(t, \cdot, \cdot)$ is of class C^2 on Ω for a.e. $t \in [0, 1]$. Besides, for each $M > 0$, there exist numbers $k_{\psi K} > 0$ such that

$$\begin{aligned} &|\psi(t, x_1, u_1) - \psi(t, x_2, u_2)| + |\nabla_z \psi(t, x_1, u_1) - \nabla_z \psi(t, x_2, u_2)| + \\ &+ |\nabla_z^2 \psi(t, x_1, u_1) - \nabla_z^2 \psi(t, x_2, u_2)| \leq k_{\psi M}(|x_1 - x_2| + |u_1 - u_2|) \end{aligned}$$

for all $(x_i, u_i) \in \Omega$ and $t \in [0, 1]$ satisfying $|x_i|, |u_i| \leq M$ with $i = 1, 2$ and $z = (x, u) \in Z$.

Moreover, we require that the functions $\psi(t, 0, 0), |\nabla_z \psi(t, 0, 0)|, |\nabla_z^2 \psi(t, 0, 0)|$ belong to $L^\infty([0, 1], \mathbb{R})$.

[(H2)] The mappings $h(\cdot)$ and $\ell(\cdot)$ are of class C^2 around $\bar{x}(1)$ and $\det(h'(\bar{x}(1))) \neq 0$.

[(H3)] There exists a number $\gamma > 0$ such that

$$|\det R[t]| \geq \gamma \quad \text{for a.a. } t \in [0, 1], \quad (3.2.52)$$

where matrix $R[t]$ is defined by setting

$$R[t] = g_u[t]g_u[t]^T. \quad (3.2.53)$$

Among the above assumptions, (H1) and (H2) make sure that ψ is of class C^2 around $(\bar{x}, \bar{u})$ while (H3) together with rank of $g_u[t]$ play an important role for normality of multipliers. To derive KKT optimality conditions for (P), we will consider two cases: $\text{rank}(g_u[\cdot]) < m$ and $\text{rank}(g_u[t]) = m$.

(H4) for a.a. $t \in [0, 1]$, $\text{rank}(g_u[t]) < m$ and $\dim(T[t]) = m_*$ with the basis $\{w_1, \dots, w_{m_*}\} \subset L^\infty([0, 1], \mathbb{R}^m)$, where the subspace $T[t] := \{w \in \mathbb{R}^m | g_u[t]w = 0\}$.

Let us define matrices

$$S[t] = [w_1, w_2, \dots, w_{m_*}], \quad (3.2.54)$$

$$A[t] = \varphi_x[t] - \varphi_u[t]g_u[t]^T R[t]^{-1} g_x[t], \quad (3.2.55)$$

$$B[t] = \varphi_u[t]S[t] \quad (3.2.56)$$

and require that the linear equation $\dot{x} = A[t]x + B[t]u$ is controllable. Namely, we need the assumption

(H5) The mapping $H : L^\infty([0, 1], \mathbb{R}^{m_*}) \rightarrow \mathbb{R}^n$ is surjective, where

$$H(v) := \int_0^1 \Omega(1, s)B[s]v(s)ds \quad (3.2.57)$$

with $\Omega(t, \tau)$ is the principal matrix solution to the equation $\dot{x} = A[t]x$ (see [4, p.123]).

When $\text{rank}(g_u[t]) = m$, then $T[t] = \{0\}$ and $S = 0$. Hence $B[t] = 0$ and the system $\dot{x} = A[t]x + B[t]u$ is not controllable. In this situation, instead of (H4) and (H5) we need the following assumption.

(H4)' There exists $(\hat{x}, \hat{u}) \in X \times U$ such that the following conditions are fulfilled:

- (a) $\hat{x} = \int_0^{(\cdot)} (\varphi_x[s]\hat{x} + \varphi_u[s]\hat{u})ds$;
- (b) $h_i(\bar{x}(1)) + h'_i(\bar{x}(1))\hat{x}(1) < 0$ for $i = 1, 2, \dots, n$;
- (c) $\text{esssup}(g_j[t] + g_{jx}[t]\hat{x} + g_{ju}[t]\hat{u}) < 0$ for $j = 1, 2, \dots, r$.

Next we define a critical cone. For this we put $I_* = \{i \in \{1, 2, \dots, n\} : h_i(\bar{x}(1)) = 0\}$ and denote by $\mathcal{C}_*[\bar{z}]$ the set of vectors $(x, u) \in X \times U$ satisfying the following conditions:

- (b₁) $\ell'(\bar{x}(1))x(1) + \int_0^1 (L_x[t]x(t) + L_u[t]u(t)) dt \in -\mathbb{R}_+^k$;
- (b₂) $x = \int_0^{(\cdot)} (\varphi_x[s]x + \varphi_u[s]u(s))ds$;
- (b₃) $h'_i(\bar{x}(1))x(1) \leq 0$ for $i \in I_*$;
- (b₄) $g_u[\cdot]x + g_u[\cdot]u \in \text{cone}(K - g[\cdot])$.

The closure of $\mathcal{C}_*[\bar{z}]$ in $X \times U$ is called a *critical cone* to the (P) at $\bar{z}$ and denoted by $\mathcal{C}_\infty[\bar{z}]$. Each vector $(x, u) \in \mathcal{C}_\infty[\bar{z}]$ is called a critical direction.

The following theorem gives first-and second-order necessary optimality conditions of KKT-type for locally weak Pareto solutions to (P).

Theorem 3.2.9. *Suppose that $\bar{z}$ is a locally weak Pareto solution of (P) under which assumptions (H1) – (H5) or (H1) – (H3) and (H4)' are satisfied. Then for each vector $(\tilde{x}, \tilde{u}) \in \mathcal{C}_\infty[\bar{z}]$, there exist a vector $\lambda \in \mathbb{R}_+^k$ with $|\lambda| = 1$, a vector $l \in \mathbb{R}^n$, an absolutely continuous function*

$p: [0, 1] \rightarrow \mathbb{R}^n$ and a vector function $\theta = (\theta_1, \theta_2, \dots, \theta_r) \in L^\infty([0, 1], \mathbb{R}^r)$ such that the following conditions are fulfilled:

(i) (the adjoint equation)

$$\begin{cases} \dot{p}(t) = -\varphi_x[t]^T p(t) + L_x[t]^T \lambda + g_x[t]^T \theta(t), & a.e. \quad t \in [0, 1] \\ p(1) = -(\lambda^T \ell'(\bar{x}(1)) + l^T h'(\bar{x}(1))); \end{cases}$$

(ii) (the stationary condition in u)

$$L_u[t]^T \lambda - \varphi_u[t]^T p(t) + g_u[t]^T \theta(t) = 0 \quad a.e. \quad t \in [0, 1];$$

(iii) (complementary conditions)

$$l_i \geq 0, \quad l_i h_i(\bar{x}(1)) = 0, \quad i = 1, 2, \dots, n, \quad (3.2.58)$$

$$\theta_j(t) \geq 0, \quad \theta_j(t) g_j[t] = 0 \quad a.a. \quad t \in [0, 1], \quad j = 1, 2, \dots, r; \quad (3.2.59)$$

(iv) (the nonnegative second-order condition)

$$\begin{aligned} & \lambda^T \ell''(\bar{x}(1)) \tilde{x}(1)^2 + \int_0^1 \lambda^T [L_{xx}[t] \tilde{x}^2 + 2L_{ux}[t] \tilde{x} \tilde{u} + L_{uu}[t] \tilde{u}^2] dt \\ & + l^T h''(\bar{x}(1)) \tilde{x}(1)^2 - \int_0^1 p^T(t) [\varphi_{xx}[t] \tilde{x}^2 + 2\varphi_{ux}[t] \tilde{x} \tilde{u} + \varphi_{uu}[t] \tilde{u}^2] dt \\ & + \int_0^1 \theta^T(t) [g_{xx}[t] \tilde{x}^2 + 2g_{ux}[t] \tilde{x} \tilde{u} + g_{uu}[t] \tilde{u}^2] dt \geq 0. \end{aligned}$$

Proof. We first formulate (MCP2) in the form of (VOP). Put

$$Z = X \times U, \quad E = C([0, 1], \mathbb{R}^n), \quad W = L^\infty([0, 1], \mathbb{R}^r)$$

and define mappings $F : Z \rightarrow E$, $H : Z \rightarrow \mathbb{R}^n$, $H : Z \rightarrow \mathbb{R}^n$ and $G : Z \rightarrow W$ by setting

$$F(x, u) := x - x_0 - \int_0^{(\cdot)} \varphi(s, x(s), u(s)) ds \quad (3.2.60)$$

$$H(x, u) := h(x(1)) \text{ and } G(x, u) := g(\cdot, x, u). \quad (3.2.61)$$

Then (MCP2) can be written in the following form of (VOP):

$$(P) \quad \begin{cases} J(x, u) \rightarrow \min, \\ F(x, u) = 0, \quad H(x, u) \leq 0 \text{ and } G(x, u) \in K. \end{cases}$$

We shall show that (P) satisfies assumptions (A1), (A2) and (A4) or (A1), (A2) and (A5) of Theorem 2.2.9. By (3.2.51), we have $\text{int}(K) \neq \emptyset$. Hence (A1) is valid. By (H1) and (H2), J , F , G and H are of class C^2

around $(\bar{x}, \bar{u})$ and $\bar{x}(1)$, respectively. Hence (A2) is fulfilled. By a simple calculation, we have

$$DF(\bar{x}, \bar{u})[(x, u)] = \left(x - \int_0^{(\cdot)} (\varphi_x[s]x(s) + \varphi_u[s]u(s))ds \right), \quad (3.2.62)$$

$$DH(\bar{z})[(x, u)] = h'(\bar{x}(1))x(1) \text{ and } DG(\bar{x}, \bar{u})[(x, u)] = g_x[t]x(t) + g_u[t]u(t). \quad (3.2.63)$$

The following lemma shows that (P) satisfies (A4).

Lemma 3.2.10. *Under assumptions (H4) and (H5), $(\bar{x}, \bar{u})$ satisfies condition (A4).*

Proof. To verify condition (2.2.13) we prove that for any $e \in E$, $\xi \in \mathbb{R}^n$ and $w \in W$, there exist $(x, u) \in Z$, $\eta \in \text{cone}((-\infty, 0]^n - H(z_0))$ and $\zeta \in \text{cone}(K - g[\cdot])$ such that

$$DF(\bar{x}, \bar{u})[(x, u)] = e, \quad DH(z_0)[(x, u)] - \eta = \xi, \quad DG(\bar{x}, \bar{u})[(x, u)] - \zeta = w. \quad (3.2.64)$$

These equations is equivalent to the system:

$$x - \int_0^{(\cdot)} (\varphi_x[s]x(s) + \varphi_u[s]u(s))ds = e, \quad (3.2.65)$$

$$h'(\bar{x}(1))x(1) - \eta = \xi \text{ and } g_x[t]x + g_u[t]u - \zeta = w. \quad (3.2.66)$$

Note that $g_x[t]$ is a matrix with size $r \times n$, $g_u[t]$ is a matrix with size $r \times m$ and $S[t]$ is a matrix with size $m \times m_*$, which is given by (3.2.54). We take $\eta = 0, \zeta = 0$ and find u in the form $u = S[t]v + g_u[t]^T R^{-1}[t]\hat{u}$, where $\hat{u} \in L^\infty([0, 1], \mathbb{R}^r)$ and $v \in L^\infty([0, 1], \mathbb{R}^{m_*})$. Then the system (3.2.65)-(3.2.66) becomes

$$\begin{aligned} x &= \int_0^{(\cdot)} (\varphi_x[s] - \varphi_u[s]g_u[s]^T R^{-1}[s]g_x[s])x(s) + \varphi_u[s]S[s]v(s))ds \\ &\quad + \int_0^{(\cdot)} \varphi_u[s]g_u[s]^T R^{-1}[s]w(s)ds + e \end{aligned} \quad (3.2.67)$$

$$x(1) = h'(x(1))^{-1}\xi \text{ and } \hat{u} = w - g_x[\cdot]x. \quad (3.2.68)$$

Here we used the fact that $g_u[t]S[t]v = 0$. The above system can be written in the form

$$x = \int_0^{(\cdot)} (A[s]x(s) + B[s]v(s))ds + \int_0^{(\cdot)} \varphi_u[s]g_u[s]^T R^{-1}[s]w(s)ds + e, \quad (3.2.69)$$

$$x(1) = h'(\bar{x}(1))^{-1}\xi \text{ and } \hat{u} = w - g_x[\cdot]x, \quad (3.2.70)$$

where A and B are defined by (3.2.55) and (3.2.56), respectively. By setting $y = x - e$, the system becomes

$$\begin{aligned} y &= \int_0^{(\cdot)} (A[s]y(s) + B[s]v(s))ds \\ &\quad + \int_0^{(\cdot)} (A[s]e(s) + \varphi_u[s]g_u[s]^T R^{-1}[s]w(s))ds, \end{aligned} \quad (3.2.71)$$

$$y(1) = h'(\bar{x}(1))^{-1}\xi - e(1) \text{ and } \hat{u} = w - g_x[\cdot]x. \quad (3.2.72)$$

Let $\Omega(t, \tau)$ be the principal matrix solution of the system $\dot{y} = Ay$. Then for each $v \in L^\infty([0, 1], \mathbb{R}^{m_*})$, the solution of equations (3.2.71) is given by

$$y(t) = \int_0^t (\Omega(t, s)B[s]v(s) + \Omega(t, s)(A[s]e(s) + \varphi_u[s]g_u[s]^T R^{-1}[s]w(s)))ds. \quad (3.2.73)$$

By (H4), the mapping

$$H(v) = \int_0^1 \Omega(1, s)B[s]v(s)ds$$

is surjective. Hence, there exists $\hat{v} \in L^\infty([0, 1], \mathbb{R}^{m_*})$ such that

$$\begin{aligned} h'(\bar{x}(1))^{-1}\eta - e(1) &= \int_0^1 \Omega(1, s)B[s]\hat{v}(s)ds + \int_0^1 \Omega(1, s)(A[s]e(s) \\ &\quad + \varphi_u[s]g_u[s]^T R^{-1}[s]w(s))ds. \end{aligned}$$

Let us define

$$\hat{y}(t) = \int_0^t \Omega(t, s)B[s]\hat{v}(s)ds + \int_0^t \Omega(t, s)(A[s]e(s) + \varphi_u[s]g_u[s]^T R^{-1}[s]w(s))ds.$$

Then $\hat{y}$ is a solution of (3.2.71) and (3.2.72) corresponding to $\hat{v}$. Let $\tilde{x} = \hat{y} + e$. Then $\tilde{x}$ is a solution of (3.2.69)-(3.2.70) corresponding to $\hat{v}$. Set $\hat{u} = w - g_x[\cdot]\tilde{x}$ and $\tilde{u} = S[t]\hat{v} + g_u[t]^T M^{-1}[t]\hat{u}$. Then $(\tilde{x}, \tilde{u})$ and $(\eta, \zeta) = (0, 0)$ satisfies (3.2.65)-(3.2.66). The lemma is proved.

Lemma 3.2.11. *Under (H4)', $(\bar{x}, \bar{u})$ satisfies (A5).*

Proof. It is easy to show that $DF(\bar{x}, \bar{u})[X \times U] = E$. By (a) of (H4)', we have $DF(\bar{x}, \bar{u})[(\hat{x}, \hat{u})] = 0$. By (b) of (H4)', we have

$$H_i(\bar{x}, \bar{u}) + DH_i(\bar{x}, \bar{u})[(\hat{x}, \hat{u})] < 0.$$

By (3.2.51), we get

$$\text{int}(K) = \{(v_1, \dots, v_r) \in L^\infty([0, 1], \mathbb{R}) : \text{esssup} v_j < 0, j = 1, 2, \dots, r.\}$$

Therefore, from (c) in $(H4)'$, we have

$$G(\bar{x}, \bar{u}) + DG(\bar{x}, \bar{u})[(\hat{x}, \hat{u})] \in \text{int}(K).$$

Hence (A5) is valid. The lemma is proved.

We now derive first and second-order optimality conditions of KKT-type for (P). Let

$$\mathcal{L}(z, \lambda, l, v^*, w^*) = \lambda^T J(x, u) + v^* F(z) + l^T H(z) + w^* G(z)$$

be the Lagrangian associated with the (P), where $z = (x, u) \in Z$. According to (b) of Theorem 2.2.9, for each critical direction $\tilde{z} = (\tilde{x}, \tilde{u}) \in \mathcal{C}_\infty[\bar{z}]$, there exist multipliers $\lambda \in \mathbb{R}_+^k$ with $|\lambda| = 1$, $l \in \mathbb{R}^n$, $v^* \in C([0, 1], \mathbb{R}^n)^*$ and $w^* \in L^\infty([0, 1], \mathbb{R}^r)^*$ such that the following conditions are valid:

$$\lambda^T J_x(\bar{x}, \bar{u}) + v^* F_x(\bar{z}) + l^T H_x(\bar{z}) + w^* g_x[\cdot] = 0, \quad (3.2.74)$$

$$\lambda^T J_u(\bar{x}, \bar{u}) + v^* F_u(\bar{z}) + w^* g_u[\cdot] = 0, \quad (3.2.75)$$

$$l_i \geq 0, \quad l_i H_i(\bar{z}) = 0, i = 1, 2, \dots, n, \quad (3.2.76)$$

$$w^* \in N(K, G(\bar{z})), \quad (3.2.77)$$

$$\lambda^T \nabla^2 J(\bar{z})[\tilde{z}, \tilde{z}] + v^* D^2 F(\bar{z})[\tilde{z}, \tilde{z}] + l^T D^2 H(\bar{z})[\tilde{z}, \tilde{z}] + w^* D^2 G(\bar{z})[\tilde{z}, \tilde{z}] \geq 0. \quad (3.2.78)$$

Here v^* is a signed Radon measure and w^* is a signed and finite additive measure on $[0, 1]$ which is absolutely continuous w.r.t the Lebesgue measure $|\cdot|$ on $[0, 1]$. By Riesz's Representation (see [108, Chapter 01, p. 19] and [102, Theorem 3.8, p. 73]), there exists a vector function of bounded variation ν , which is continuous from the right and vanish at zero such that

$$\langle v^*, y \rangle = \int_0^1 y(t)^T d\nu(t) \quad \forall y \in E,$$

where $\int_0^1 y(t) d\nu(t)$ is the Riemann-Stieltjes integral.

Define $\bar{p}: [0, 1] \rightarrow \mathbb{R}^n$ by setting

$$\bar{p}(t) = \nu((t, 1]) = \int_t^1 d\nu(s) = \nu(1) - \nu(t).$$

Clearly, $\bar{p}(1) = 0$ and the function $\bar{p}$ is of bounded variation.

By [225, Theorem 2.21], for any $a \in C([0, 1], \mathbb{R}^n)$, we have

$$\int_0^1 a(s)^T d\nu(s) = a(1)^T \nu(1) - a(0)^T \nu(0) - \int_0^1 \nu(s) da(s)^T. \quad (3.2.79)$$

Applying formula (3.2.79) for $a(t) = \int_0^t \varphi_x[s]x(s)ds$, we get

$$\begin{aligned} \langle v^* F_x(\bar{z}), x \rangle &= \left\langle v^*, x - \int_0^{(\cdot)} \varphi_x[s]x(s)ds \right\rangle \\ &= \int_0^1 x^T(t)d\nu(t) - \int_0^1 \int_0^t (\varphi_x[s]x(s))^T ds d\nu(t) \\ &= \int_0^1 x^T(t)d\nu(t) - \left[\int_0^1 (\varphi_x[s]x(s))^T \nu(1)ds - \int_0^1 (\varphi_x[s]x(s))^T \nu(s)ds \right] \\ &= \int_0^1 x^T(t)d\nu(t) - \int_0^1 x^T(t)\varphi_x[t]^T \bar{p}(t)dt. \\ &= \int_0^1 x^T(t)d\nu(t) - \int_0^1 \bar{p}^T(t)\varphi_x[t]x(t)dt. \end{aligned} \quad (3.2.80)$$

Similarly, for any $u \in L^\infty([0, 1], \mathbb{R}^m)$, we get

$$\langle v^* F_u(\bar{z}), u \rangle = - \int_0^1 \int_0^t (\varphi_u[s]u(s))^T ds d\nu(t) = - \int_0^1 u(s)^T \varphi_u[s]^T \bar{p}(s)ds \quad (3.2.81)$$

and

$$\langle v^* \nabla^2 F(\bar{z})z, z \rangle = - \int_0^1 (\bar{p}(t)^T \nabla^2 \varphi[s]z(s), z(s)) ds. \quad (3.2.82)$$

From (3.2.75) and (3.2.81), we have

$$\int_0^1 \lambda^T L_u[s]u(s)ds - \int_0^1 \bar{p}(s)^T \varphi_u[s]u(s)ds + \langle w^*, g_u[\cdot]u \rangle = 0 \quad \forall u \in U. \quad (3.2.83)$$

Let us claim that w^* can be represented by a density in $L^\infty([0, 1], \mathbb{R}^r)$. Indeed, by the formula of inverse matrix (see [145, Theorem 8.1, p. 176]), we have

$$R^{-1}[t] = \left(\frac{e_{ij}(t)}{\det R[t]} \right),$$

where $e_{ij}(\alpha)$ are cofactors. By (H3), $|\det R[t]| \geq \gamma_0$ for a.a $t \in [0, 1]$. Hence $R^{-1}[\cdot] \in L^\infty([0, 1], \mathbb{R}^{r \times r})$. Taking any $w \in L^\infty([0, 1], \mathbb{R}^r)$, we put $u(t) = g_u[t]^T R^{-1}[t]w(t)$. Inserting u into (3.2.83), we have

$$\begin{aligned} &\int_0^1 \lambda^T L_u[t]g_u[t]^T R^{-1}[t]w(t)dt - \int_0^1 \bar{p}^T(t)\varphi_u[t]g_u[t]^T R^{-1}[t]w(t)dt \\ &+ \langle w^*, w \rangle = 0. \end{aligned} \quad (3.2.84)$$

This implies that

$$\langle w^*, w \rangle = \int_0^1 (-\lambda^T L_u[t] g_u[t]^T R^{-1}[t] + \bar{p}^T(t) \varphi_u[t] g_u[t]^T R^{-1}[t]) w(t) dt.$$

Let us define a function

$$\theta(t) = -\lambda^T L_u[t] g_u[t]^T R^{-1}[t] + \bar{p}^T(t) \varphi_u[t] g_u[t]^T R^{-1}[t].$$

Then $\theta \in L^\infty([0, 1], \mathbb{R})$ and

$$\langle w^*, w \rangle = \int_0^1 \theta(t) w(t) dt \quad \forall w \in L^\infty([0, 1], \mathbb{R}^r). \quad (3.2.85)$$

The claim is justified. Combining this with (3.2.83), we get

$$\int_0^1 \lambda^T L_u[t] u(t) dt - \int_0^1 \bar{p}(t)^T \varphi_u[t] u(t) dt + \int_0^1 \theta(t)^T g_u[t] u(t) dt = 0$$

for all $u \in L^\infty([0, 1], \mathbb{R}^m)$. It follows that

$$\lambda^T L_u[t] - \bar{p}(t)^T \varphi_u[t] + \theta(t)^T g_u[t] = 0 \quad \text{a.a.} \quad t \in [0, 1]. \quad (3.2.86)$$

We obtain assertion (ii) of the theorem. Let us define

$$p(t) = \begin{cases} \bar{p}(t) & \text{if } t \in [0, 1) \\ \lim_{t \rightarrow 1^-} \bar{p}(t) & \text{if } t = 1. \end{cases} \quad (3.2.87)$$

Note that although ν is right continuous, it may not be continuous at 1 and so is $\bar{p}$.

We now have from (3.2.74) and (3.2.80) that

$$\begin{aligned} & \int_0^1 x^T(s) d\nu(s) - \int_0^1 \bar{p}^T(s) \varphi_x[s] x(s) ds + \int_0^1 \lambda^T L_x[s] x(s) ds \\ & + \int_0^1 \theta(s)^T g_x[s] x(s) ds + \lambda^T \ell_{x_1}(\bar{x}(1)) x(1) + l^T h_{x_1}(\bar{x}(1)) x(1) = 0 \quad \forall x \in X. \end{aligned}$$

This is equivalent to

$$\begin{aligned} & \int_0^1 x^T(s) d\nu(s) = \int_0^1 \bar{p}^T(s) \varphi_x[s] x(s) ds - \int_0^1 \lambda^T L_x[s] x(s) ds \\ & - \int_0^1 \theta(s)^T g_x[s] x(s) ds - \lambda^T \ell_{x_1}(\bar{x}(1)) x(1) - l^T h_{x_1}(\bar{x}(1)) x(1) \quad \forall x \in X. \end{aligned}$$

By definition of p , we have

$$\begin{aligned} \int_0^1 x^T(s) d\nu(s) &= \int_0^1 p^T(s) \varphi_x[s] x(s) ds - \int_0^1 \lambda^T L_x[s] x(s) ds \\ &\quad - \int_0^1 \theta(s)^T g_x[s] x(s) ds - \lambda^T \ell_{x_1}(\bar{x}(1)) x(1) - l^T h_{x_1}(\bar{x}(1)) x(1) \quad \forall x \in X. \end{aligned} \quad (3.2.88)$$

By Lemma 3.2.4, the above equality is also valid for $x_t(s) = \xi \chi_{(t,1]}(s)$ with $0 \leq t < 1$, where $\xi \in \mathbb{R}^n$ and $\chi_{(t,1]}$ is the indicator function of the set $(t, 1] \subset [0, 1]$. Inserting $x_t(s)$ into (3.2.88), we get

$$\begin{aligned} \int_t^1 \xi^T d\nu(s) &= \int_t^1 \xi^T (p^T(s) \varphi_x[s] - \lambda^T L_x[s] - \theta(s)^T g_x[s]) ds \\ &\quad - \xi^T (\lambda^T \ell_{x_1}(\bar{x}(1)) - l^T h_{x_1}(\bar{x}(1))) \end{aligned}$$

for all $t \in [0, 1)$. Since ξ is arbitrary, we obtain

$$\begin{aligned} \int_t^1 d\nu(s) &= \int_t^1 (p^T(s) \varphi_x[s] - \lambda^T L_x[s] - \theta(s)^T g_x[s]) ds \\ &\quad - (\lambda^T \ell_{x_1}(\bar{x}(1)) + l^T h_{x_1}(\bar{x}(1))). \end{aligned}$$

This means that

$$\begin{aligned} \bar{p}(t) &= \int_t^1 (p^T(s) \varphi_x[s] - \lambda^T L_x[s] - \theta(s)^T g_x[s]) ds \\ &\quad - (\lambda^T \ell_{x_1}(\bar{x}(1)) + l^T h_{x_1}(\bar{x}(1))) \end{aligned} \quad (3.2.89)$$

for all $t \in [0, 1)$. It follows that

$$\dot{p}(t) = -p^T(t) \varphi_x[t] + \lambda^T L_x[t] + \theta(t)^T g_x[t] \quad \text{a.a. } t \in (0, 1). \quad (3.2.90)$$

Taking the limit both sides of (3.2.89) when $t \rightarrow 1^-$, we obtain

$$p(1) = \lim_{t \rightarrow 1^-} \bar{p}(t) = -(\lambda^T \ell_{x_1}(\bar{x}(1)) + l^T h_{x_1}(\bar{x}(1))).$$

Hence assertion (i) is followed. Since $H_i(\bar{z}) = h_i(\bar{x}(1))$, condition (3.2.76) implies that $l_i h_i(\bar{x}(1)) = 0$. From condition (3.2.77), we have

$$\int_0^1 \theta^T(w(t) - g[t]) dt = \langle w^*, w - G(\bar{z}) \rangle \leq 0 \quad \forall w \in K.$$

This and [178, Corolary 4.4] imply that $\theta(t) \in N((-\infty, 0]^r, g[t])$ for a.e. $t \in [0, 1]$. Hence $\theta_j(t) g_j[t] = 0$ and $\theta_j(t) \geq 0$ for a.a. $t \in [0, 1]$ and for all $j = 1, 2, \dots, r$. We obtain (3.2.59) in assertion (iii).

Finally, by combining (3.2.78), (3.2.82) and (3.2.85), we derive assertion (iv). The proof of the theorem is complete. $\square$

To derive sufficient conditions, we need to enlarge the critical cone. For this we define $\mathcal{C}_2[\bar{z}]$ which consists of couples $(x, u) \in X \times L^2([0, 1], \mathbb{R}^m)$ such that

$$(b'_1) \quad \ell'(\bar{x}(1))x(1) + \int_0^1 (L_x[t]x(t) + L_u[t]u(t)) dt \in -\mathbb{R}_+^k;$$

$$(b'_2) \quad x = \int_0^1 (\varphi_x[s]x + \varphi_u[s]u(s)) ds;$$

$$(b'_3) \quad h'_i(\bar{x}(1))x(1) \leq 0 \quad \text{for } i \in I_*;$$

$$(b'_4) \quad g_u[\cdot]x + g_u[\cdot]u \in \overline{\text{cone}}(K - g[\cdot]).$$

Clearly, $\mathcal{C}_\infty[\bar{z}] \subseteq \mathcal{C}_2[\bar{z}]$.

Definition 3.2.12. Let $\bar{z} = (\bar{x}, \bar{u}) \in \Phi_2$ be a feasible point of (P). We say that $\bar{z}$ is a locally strong Pareto solution of (P) if there exist a number $\epsilon > 0$ and a vector $c \in \text{int}\mathbb{R}_+^k$ such that for all $(x, u) \in (B_X(\bar{x}, \epsilon) \times B_U(\bar{u}, \epsilon)) \cap \Phi_2$, one has

$$J(x, u) - J(\bar{x}, \bar{u}) - c\|u - \bar{u}\|_2^2 \notin -\text{int}(\mathbb{R}_+^k).$$

It is easy to show that every locally strong Pareto solution of (P) is also a locally Pareto solution of this problem. The following theorem provides sufficient conditions for locally strong Pareto solutions to (P).

Theorem 3.2.13. Suppose $(\bar{x}, \bar{u}) \in \Phi_2$, assumptions (H1) – (H3), multipliers $\lambda \in \mathbb{R}_+^k$ with $|\lambda| = 1$, $l \in \mathbb{R}^n$, the absolutely continuous function $p : [0, 1] \rightarrow \mathbb{R}^n$ and a function $\theta \in L^\infty([0, 1], \mathbb{R})$ such that conditions (i) – (iii) of Theorem 3.1 are fulfilled. Furthermore, assume that (iv)' (the strictly second-order condition)

$$\begin{aligned} & \lambda^T \ell''(\bar{x}(1))\tilde{x}(1)^2 + \int_0^1 \lambda^T [L_{xx}[t]\tilde{x}^2 + 2L_{ux}[t]\tilde{x}\tilde{u} + L_{uu}[t]\tilde{u}^2] dt \\ & + l^T h''(\bar{x}(1))\tilde{x}(1)^2 - \int_0^1 p^T(t) [\varphi_{xx}[t]\tilde{x}^2 + 2\varphi_{ux}[t]\tilde{x}\tilde{u} + \varphi_{uu}[t]\tilde{u}^2] dt \\ & + \int_0^1 \theta^T(t) [g_{xx}[t]\tilde{x}^2 + 2g_{ux}[t]\tilde{x}\tilde{u} + g_{uu}[t]\tilde{u}^2] dt > 0 \end{aligned}$$

for all $(\tilde{x}, \tilde{u}) \in \mathcal{C}_2[\bar{z}] \setminus \{0\}$;

(v)' there exists a number $\gamma_0 > 0$ such that for a.a. $t \in [0, 1]$ one has

$$(\lambda^T L_{uu}[t] - p(t)^T \varphi_{uu}[t] + \theta(t)^T g_{uu}[t])(v, v) \geq \gamma_0 |v|^2 \quad \forall v \in \mathbb{R}^m.$$

Then $(\bar{x}, \bar{u})$ is a locally strong Pareto solution of (P).

Proof. Suppose the conclusion of the theorem was false. Then, we could find sequences $\{(x_j, u_j)\} \subset \Phi_2$ and $\{c_j\} \subset \text{int } \mathbb{R}_+^k$ such that $(x_j, u_j) \rightarrow (\bar{x}, \bar{u})$ as $j \rightarrow \infty$, $c_j \rightarrow 0$ and

$$J(x_j, u_j) - J(\bar{x}, \bar{u}) - c_j \|u_j - \bar{u}\|_2^2 \in -\text{int } \mathbb{R}_+^k. \quad (3.2.91)$$

It follows that $(x_j, u_j) \neq (\bar{x}, \bar{u})$ for all $j \geq 1$. By a simple argument, we can show that if $u_j = \bar{u}$ then $x_j = \bar{x}$, which is absurd. Hence we must have $u_j \neq \bar{u}$ for all $j \geq 1$. Let

$$\mathcal{L}(z, \lambda, p, l, \theta) := \lambda^T J(z) + p^T F(z) + l^T H(z) + \theta^T G(z)$$

be the Lagrange function associated with (P), where

$$\begin{aligned} \lambda^T J(z) &= \lambda^T \ell(x(1)) + \int_0^1 \lambda^T L(s, x(s), u(s)) ds, \\ p^T F(z) &= - \int_0^1 \dot{p}^T(s) x(s) ds - \int_0^1 p^T(s) \varphi(s, x(s), u(s)) ds + p^T(1) x(1), \\ l^T H(z) &= l^T h(x(1)), \\ \theta^T G(z) &= \int_0^1 \theta^T(s) g(s, x(s), u(s)) ds. \end{aligned}$$

Then, from conditions (i) (ii) and (iii) of Theorem 3.1, we can show that

$$\nabla_z \mathcal{L}(\bar{z}, \lambda, p, l, \theta) = 0. \quad (3.2.92)$$

Define $t_j = \|u_j - \bar{u}\|_2$, $\hat{x}_j = \frac{x_j - \bar{x}}{t_j}$ and $\hat{u}_j = \frac{u_j - \bar{u}}{t_j}$. Then $t_j \rightarrow 0^+$ and $\|\hat{u}_j\|_2 = 1$. By reflexivity of $L^2([0, 1], \mathbb{R}^m)$, we may assume that $\hat{u}_j \rightharpoonup \hat{u}$. Since $\lambda \in \mathbb{R}_+^k$, we have from (3.2.91) that

$$\lambda^T J(z_j) - \lambda^T J(\bar{z}) \leq t_j^2 \lambda^T c_j \leq t_j^2 |\lambda| |c_j| \leq o(t_j^2). \quad (3.2.93)$$

For clarification, we divide the remain of the proof into some steps.

Step 1. Showing hat $\hat{x}_j$ converges uniformly to some $\hat{x}$ in $C([0, 1], \mathbb{R}^n)$.

Indeed, since $(\bar{x}, \bar{u}), (x_j, u_j) \in \Phi_2$, we get

$$x_j(t) = x_0 + \int_0^t \varphi(s, x_j(s), u_j(s)) ds, \quad \bar{x}(t) = x_0 + \int_0^t \varphi(s, \bar{x}(s), \bar{u}(s)) ds.$$

Hence

$$t_j \hat{x}_j(t) = \int_0^t (\varphi(s, x_j(s), u_j(s)) - \varphi(s, \bar{x}(s), \bar{u}(s))) ds. \quad (3.2.94)$$

Since $x_j \rightarrow \bar{x}$ uniformly and $u_j \rightarrow \bar{u}$ in $L^\infty([0, 1], \mathbb{R}^m)$, there exists a constant $\varrho > 0$ such that $\|x_j\|_0 \leq \varrho, \|u_j\|_\infty \leq \varrho$. By assumption (H1), there exists $k_{\varphi, \varrho} > 0$ such that

$$|\varphi(s, x_j(s), u_j(s)) - \varphi(s, \bar{x}(s), \bar{u}(s))| \leq k_{\varphi, \varrho}(|x_j(s) - \bar{x}(s)| + |u_j(s) - \bar{u}(s)|)$$

for a.e. $s \in [0, 1]$. Hence we have from (3.2.94) that

$$|\hat{x}_j(t)| \leq \int_0^t k_{\varphi, \varrho}(|\hat{x}_j(s)| + |\hat{u}_j(s)|)ds \text{ and } |\dot{\hat{x}}_j(t)| \leq k_{\varphi, \varrho}(|\hat{x}_j(t)| + |\hat{u}_j(t)|). \quad (3.2.95)$$

Since $\|\hat{u}_j\|_2 = 1$, we get

$$\begin{aligned} |\hat{x}_j(t)| &\leq \int_0^t k_{\varphi, \varrho}|\hat{x}_j(s)|ds + \int_0^1 k_{\varphi, \varrho}|\hat{u}_j(s)|ds \\ &\leq \int_0^t k_{\varphi, \varrho}|\hat{x}_j(s)|ds + k_{\varphi, \varrho} \left(\int_0^1 |\hat{u}_j(s)|^2 ds \right)^{1/2} \\ &\leq \int_0^t k_{\varphi, \varrho}|\hat{x}_j(s)|ds + k_{\varphi, \varrho}. \end{aligned}$$

By the Gronwall Inequality (see [51, 18.1.i, p. 503]), we have

$$|\hat{x}_j(t)| \leq k_{\varphi, \varrho} \exp(k_{\varphi, \varrho}).$$

Combining this with (3.2.95) yields

$$|\dot{\hat{x}}_j(t)|^2 \leq 2k_{\varphi, \varrho}^2 (|\hat{x}_j(t)|^2 + |\hat{u}_j(t)|^2) \leq 2k_{\varphi, \varrho}^2 (k_{\varphi, \varrho}^2 \exp(2k_{\varphi, \varrho}) + |\hat{u}_j|^2).$$

Hence

$$\int_0^1 |\dot{\hat{x}}_j(t)|^2 dt \leq 2k_{\varphi, \varrho}^2 (k_{\varphi, \varrho}^2 \exp(2k_{\varphi, \varrho}) + 1).$$

Consequently, $\{\hat{x}_j\}$ is bounded in $W^{1,2}([0, 1], \mathbb{R}^n)$. By passing subsequence, we may assume that $\hat{x}_j \rightharpoonup \hat{x}$ weakly in $W^{1,2}([0, 1], \mathbb{R}^n)$. As the embedding $W^{1,2}([0, 1], \mathbb{R}^n) \hookrightarrow C([0, 1], \mathbb{R}^n)$ is compact (see [31, Theorem 8.8]) we have $\hat{x}_j \rightarrow \hat{x}$ in norm of $C([0, 1], \mathbb{R}^n)$.

Step 2. Showing that $(\hat{x}, \hat{u}) \in \mathcal{C}_2[\bar{z}]$.

Using the mean value theorem for the left hand side of (3.2.91) and dividing both sides by t_j , we get

$$\begin{aligned} &\ell'(\bar{x}(1) + \gamma_j(x_j(1) - \bar{x}(1)))\hat{x}_j(1) + \int_0^1 L_x(\bar{x} + \xi_j(x_j - \bar{x}), u_j)\hat{x}_j dt \\ &+ \int_0^1 L_u(\bar{x}, \bar{u} + \xi_j(u_j - \bar{u}))\hat{u}_j dt + \frac{o(t_j^2)}{t_j} \in -\text{int}(\mathbb{R}_+^k), \end{aligned}$$

where $0 \leq \gamma_j \leq 1$ and $0 \leq \xi_j(t) \leq 1$. Letting $j \rightarrow \infty$ and using the dominated convergence theorem, we obtain

$$\ell'(\bar{x}(1))\hat{x}(1) + \int_0^1 (L_x[t]\hat{x}(t) + L_u[t]\hat{u}(t))dt \in -\mathbb{R}_+^k. \quad (3.2.96)$$

Hence $(\hat{x}, \hat{u})$ satisfies condition (b'_1) . Also, from (3.2.94) we have

$$\hat{x}_j(t) = \int_0^t (\varphi_x(\bar{x} + \xi_j(x_j - \bar{x}), u_j)\hat{x}_j + \varphi_u(\bar{x}, \bar{u} + \eta_j(u_j - \bar{u}))\hat{u}_j)dt,$$

where $0 \leq \xi_j(s), \eta_j(s) \leq 1$. Letting $j \rightarrow \infty$ and using the dominated convergence theorem, we obtain

$$\hat{x}(t) = \int_0^t (\varphi_x[s]\hat{x}(s) + \varphi_u[s]\hat{u}(s))ds.$$

Hence $(\hat{x}, \hat{u})$ satisfies condition (b'_2) . Obviously, we have

$$h(x_j(1)) - h(\bar{x}(1)) \in ((-\infty, 0]^n - h(\bar{x}(1))).$$

By a Taylor's expansion, we get

$$\begin{aligned} h'(\bar{x}(1) + \tau(x_j(1) - \bar{x}(1)))\hat{x}_j(1) &\in \frac{1}{t_j}((-\infty, 0]^n - h(\bar{x}(1))) \\ &\subset T((-\infty, 0]^n, h(\bar{x}(1))). \end{aligned}$$

By letting $j \rightarrow \infty$, we obtain

$$h'(\bar{x}(1))\hat{x}(1) \in T((-\infty, 0]^n, h(\bar{x}(1))).$$

Hence $(\hat{x}, \hat{u})$ satisfies condition (b'_3) . It remains to verify condition (b'_4) .

Since $G(x_j, u_j) - G(\bar{x}, \bar{u}) \in K - G(\bar{x}, \bar{u})$, we have

$$\frac{1}{t_j}(G(x_j, u_j) - G(\bar{x}, \bar{u})) \in \text{cone}(K - G(\bar{z})) \subset \overline{\text{cone}}(K - G(\bar{z})). \quad (3.2.97)$$

Note that

$$\begin{aligned} \overline{\text{cone}}(K - G(\bar{z})) &= T(K; G(\bar{z})) \\ &= \{v \in L^2([0, 1], \mathbb{R}^r) \mid v(t) \in T((-\infty, 0]^r, g[t]) \text{ a.e.}\}. \end{aligned}$$

Using the mean value theorem, we have from (3.2.97) that

$$g_x(t, \bar{x} + \eta_j(x_j - \bar{x}), \bar{u})\hat{x}_j + g_u(t, x_j, \bar{u} + \zeta_j(u_j - \bar{u}))\hat{u}_j \in T(K; G(\bar{z})), \quad (3.2.98)$$

where $0 \leq \eta_j(t), \zeta_j(t) \leq 1$ a.a. $t \in [0, 1]$. Since $T(K; G(\bar{z}))$ is a closed convex set in $L^2([0, 1], \mathbb{R}^r)$, it is also a weakly closed set in $L^2([0, 1], \mathbb{R}^r)$. By (H1), we can show that

$$g_x(\cdot, \bar{x} + \eta_j(x_j - \bar{x}), \bar{u})\hat{x}_j \rightarrow g_x[\cdot]\hat{x}$$

strongly in $L^2([0, 1], \mathbb{R}^r)$ and

$$g_u(t, x_j, \bar{u} + \zeta_j(u_j - \bar{u}))\hat{u}_j \rightarrow g_u[\cdot]\hat{u}$$

weakly in $L^2([0, 1], \mathbb{R}^r)$. Letting $j \rightarrow \infty$, we obtain from that (3.2.98) that

$$g_x[\cdot]\hat{x} + g_u[\cdot]\hat{u} \in T(K, g[\cdot]).$$

Hence $(\hat{x}, \hat{u})$ satisfies condition (b'_4) and so $(\hat{x}, \hat{u}) \in \mathcal{C}_2[\bar{z}]$.

Step 3. Showing that $(\hat{x}, \hat{u}) = (0, 0)$.

By (3.2.58), we have $l^T(h(x_j(1)) - h(\bar{x}(1))) \leq 0$. By (3.2.59),

$$\int_0^1 \theta^T(t)(g(t, x_j, u_j) - g[t])dt \leq 0.$$

Besides, we have

$$\begin{aligned} & \int_0^1 (-\dot{p}^T(x_j - \bar{x}) - p(\varphi(t, x_j, u_j) - \varphi(t, \bar{x}, \bar{u})))dt + p^T(1)(x_j(1) - \bar{x}(1)) \\ &= \int_0^1 p((\dot{x}_j - \dot{\bar{x}}) - (\varphi(t, x_j, u_j) - \varphi(t, \bar{x}, \bar{u})))dt = 0. \end{aligned}$$

From the above, (3.2.93) and definition of $\mathcal{L}$, we have

$$\begin{aligned} & \mathcal{L}(x_j, u_j, \lambda, p, l, \theta) - \mathcal{L}(\bar{x}, \bar{u}, \lambda, p, l, \theta) \\ &= \lambda^T(\ell(x_j(1)) - \ell(\bar{x}(1))) + \lambda^T \int_0^1 (L(t, x_j, u_j) - L(t, \bar{x}, \bar{u}))dt \\ &+ l^T(h(x_j(1)) - h(\bar{x}(1))) - \int_0^1 (\dot{p}^T(x_j - \bar{x}) + p(\varphi(t, x_j, u_j) - \varphi(t, \bar{x}, \bar{u})))dt \\ &+ \int_0^1 \theta^T(t)(g(t, x_j, u_j) - g(t, \bar{x}, \bar{u}))dt \\ &\leq \lambda^T(\ell(x_j(1)) - \ell(\bar{x}(1))) + \lambda^T \int_0^1 (L(t, x_j, u_j) - L(t, \bar{x}, \bar{u}))dt \leq o(t_j^2). \end{aligned}$$

Using a second-order Taylor expansion for $\mathcal{L}$ and (3.2.92), we get

$$o(t_j^2) \geq \mathcal{L}(z_k, \lambda, \bar{p}, l, \theta) - \mathcal{L}(\bar{z}, \lambda, \bar{p}, l, \theta) =$$

$$\begin{aligned}
& \frac{t_j^2}{2} [\lambda^T \ell''(\bar{x}(1) + \xi_{1j}(x_j(1) - \bar{x}(1))) \hat{x}_j(1)^2 + l^T h''(\bar{x}(1) + \xi_{2j}(x_j(1) - \bar{x}(1))) \hat{x}_j(1)^2] \\
& + \frac{t_j^2}{2} \lambda^T \int_0^1 [L_{xx}(t, \bar{x} + \eta_{1j}(x_j - \bar{x}), \bar{u}) \hat{x}_j^2 + 2L_{ux}(t, \bar{x} + \eta_{2j}(x_j - \bar{x}), \bar{u}) \hat{x}_j \hat{u}_j] dt \\
& + \frac{t_j^2}{2} \lambda^T \int_0^1 L_{uu}(t, x_j, \bar{u} + \eta_{3j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& - \frac{t_j^2}{2} \int_0^1 p^T(t) [\varphi_{xx}(t, \bar{x} + \eta_{4j}(x_j - \bar{x}), \bar{u}) \hat{x}_j^2 + 2\varphi_{ux}(t, \bar{x} + \eta_{5j}(x_j - \bar{x}), \bar{u}) \hat{x}_j \hat{u}_j] dt \\
& - \frac{t_j^2}{2} \int_0^1 p^T(t) \varphi_{uu}(t, x_j, \bar{u} + \eta_{6j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& + \frac{t_j^2}{2} \int_0^1 \theta^T(t) [g_{xx}(t, \bar{x} + \eta_{7j}(x_j - \bar{x}), \bar{u}) \hat{x}_j^2 + 2g_{ux}(t, \bar{x} + \eta_{8j}(x_j - \bar{x}), \bar{u}) \hat{x}_j \hat{u}_j] dt \\
& + \frac{t_j^2}{2} \int_0^1 \theta^T(t) g_{uu}(t, x_j, \bar{u} + \eta_{9j}(u_j - \bar{u})) \hat{u}_j^2 dt,
\end{aligned}$$

where $0 \leq \xi_{ij}, \eta_{ij}(t) \leq 1$. This implies that

$$\begin{aligned}
& \frac{o(t_j^2)}{t_j^2} \geq \\
& \lambda^T \ell''(\bar{x}(1) + \xi_{1j}(x_j(1) - \bar{x}(1))) \hat{x}_j(1)^2 + l^T h''(\bar{x}(1) + \xi_{2j}(x_j(1) - \bar{x}(1))) \hat{x}_j(1)^2 \\
& + \lambda^T \int_0^1 [L_{xx}(t, \bar{x} + \eta_{1j}(x_j - \bar{x}), \bar{u}) \hat{x}_j^2 + 2L_{ux}(t, \bar{x} + \eta_{2j}(x_j - \bar{x}), \bar{u}) \hat{x}_j \hat{u}_j] dt \\
& + \lambda^T \int_0^1 L_{uu}(t, x_j, \bar{u} + \eta_{3j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& - \int_0^1 p^T(t) [\varphi_{xx}(t, \bar{x} + \eta_{4j}(x_j - \bar{x}), \bar{u}) \hat{x}_j^2 + 2\varphi_{ux}(t, \bar{x} + \eta_{5j}(x_j - \bar{x}), \bar{u}) \hat{x}_j \hat{u}_j] dt \\
& - \int_0^1 p^T(t) \varphi_{uu}(t, x_j, \bar{u} + \eta_{6j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& + \int_0^1 \theta^T(t) [g_{xx}(t, \bar{x} + \eta_{7j}(x_j - \bar{x}), \bar{u}) \hat{x}_j^2 + 2g_{ux}(t, \bar{x} + \eta_{8j}(x_j - \bar{x}), \bar{u}) \hat{x}_j \hat{u}_j] dt \\
& + \int_0^1 \theta^T(t) g_{uu}(t, x_j, \bar{u} + \eta_{9j}(u_j - \bar{u})) \hat{u}_j^2 dt. \tag{3.2.99}
\end{aligned}$$

Since $\|u_j - \bar{u}\|_\infty \rightarrow 0$, $\|x_j - \bar{x}\|_0 \rightarrow 0$ and (H1), we have

$$\|\lambda^T L_{uu}(\cdot, x_j, \bar{u} + \eta_{3j}(u_j - \bar{u})) \hat{u}_j^2 - \lambda^T L[\cdot]\|_\infty \rightarrow 0 \quad \text{as } j \rightarrow \infty.$$

By (v)', the functional

$$u \mapsto \int_0^1 (\lambda^T L_{uu}[t] - p(t)^T \varphi_{uu}(t) + \theta(t) g_{uu}[t]) u^2 dt$$

is convex and sequentially lower semicontinuous. We now write

$$\begin{aligned}
& \lambda^T \int_0^1 L_{uu}(t, x_j, \bar{u} + \eta_{3j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& - \int_0^1 p^T(t) \varphi_{uu}(t, x_j, \bar{u} + \eta_{6j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& + \int_0^1 \theta^T(t) g_{uu}(t, x_j, \bar{u} + \eta_{9j}(u_j - \bar{u})) \hat{u}_j^2 dt. \\
& = \int_0^1 (\lambda^T L_{uu}(t, x_j, \bar{u} + \eta_{3j}(u_j - \bar{u})) - \lambda^T L_{uu}[t]) \hat{u}_j^2 dt + \int_0^1 \lambda^T L_{uu}[t] \hat{u}_j^2 dt. \\
& - \int_0^1 p^T(t) (\varphi_{uu}(t, x_j, \bar{u} + \eta_{6j}(u_j - \bar{u})) - \varphi_{uu}[t]) \hat{u}_j^2 dt - \int_0^1 p(t)^T \varphi_{uu}[t] \hat{u}_j^2 dt \\
& + \int_0^1 \theta^T(t) (g_{uu}(t, x_j, \bar{u} + \eta_{9j}(u_j - \bar{u})) - g_{uu}[t]) \hat{u}_j^2 dt + \int_0^1 \theta(t)^T g_{uu}[t] \hat{u}_j^2 dt.
\end{aligned}$$

From this and (v)', we have

$$\begin{aligned}
& \lim_{j \rightarrow \infty} \left(\lambda^T \int_0^1 L_{uu}(t, x_j, \bar{u} + \eta_{3j}(u_j - \bar{u})) \hat{u}_j^2 dt \right. \\
& \quad - \int_0^1 p^T(t) \varphi_{uu}(t, x_j, \bar{u} + \eta_{6j}(u_j - \bar{u})) \hat{u}_j^2 dt \\
& \quad \left. + \int_0^1 \theta^T(t) g_{uu}(t, x_j, \bar{u} + \eta_{9j}(u_j - \bar{u})) \hat{u}_j^2 dt \right) \\
& \geq \max \left(\gamma_0, \int_0^1 (\lambda^T L_{uu}[t] - p(t)^T \varphi_{uu}[t] + \theta(t)^T g_{uu}[t]) \hat{u}^2(t) dt \right).
\end{aligned} \tag{3.2.100}$$

Using this fact and taking limits both sides of (3.2.99), we obtain

$$\begin{aligned}
0 & \geq \lambda^T \ell''(\bar{x}(1)) \hat{x}(1)^2 + l^T h''(\bar{x}(1)) \hat{x}(1)^2 \\
& + \int_0^1 \lambda^T [L_{xx}[t] \hat{x}^2 + 2L_{ux}[t] \hat{x} \hat{u} + L_{uu}[t] \hat{u}^2] dt \\
& - \int_0^1 p^T(t) [\varphi_{xx}[t] \hat{x}^2 + 2\varphi_{ux}[t] \hat{x} \hat{u} + \varphi_{uu}[t] \hat{u}^2] dt \\
& + \int_0^1 \theta^T(t) [g_{xx}[t] \hat{x}^2 + 2g_{ux}[t] \hat{x} \hat{u} + g_{uu}[t] \hat{u}^2] dt.
\end{aligned}$$

Combining this with (iv)', yields $(\hat{x}, \hat{u}) = (0, 0)$.

Finally, using (3.2.100) and taking limits both sides of (3.2.99) again, we get

$$0 \geq \lambda^T \ell''(\bar{x}(1)) \hat{x}(1)^2 + l^T h''(\bar{x}(1)) \hat{x}(1)^2 + \int_0^1 \lambda^T [L_{xx}[t] \hat{x}^2 + 2L_{ux}[t] \hat{x} \hat{u}] dt$$

$$\begin{aligned}
& + \gamma_0 - \int_0^1 p^T(t) [\varphi_{xx}[t]\hat{x}^2 + 2\varphi_{ux}[t]\hat{x}\hat{u} + \varphi_{uu}[t]\hat{u}^2] dt \\
& + \int_0^1 \theta^T(t) [g_{xx}[t]\hat{x}^2 + 2g_{ux}[t]\hat{x}\hat{u} + g_{uu}[t]\hat{u}^2] dt.
\end{aligned}$$

Since $(\hat{x}, \hat{u}) = (0, 0)$, we get $0 \geq \gamma_0$ which is absurd. This finishes the proof. $\square$

We will now present an illustrative example that satisfies conditions (H1)-(H5) or (H1)-(H3), (H4)' of Theorem 3.2.9.

Example 3.2.14. Consider the problem of finding $x \in C([0, 1], \mathbb{R}^2)$ and $u \in L^\infty([0, 1], \mathbb{R}^3)$ which solve

$$\left\{ \begin{array}{l} J(x, u) = (J_1(x, u), J_2(x, u), \dots, J_k(x, u)) \rightarrow \min, \\ \text{s.t.} \\ x_1(t) = x_{10} + \int_0^t (sx_1 + u_1 + u_3) ds \\ x_2(t) = x_{20} + \int_0^t (-\frac{1}{3}x_1 + sx_2 + u_2) ds \\ h_1(x(1)) = x_1(1) - x_2^2(1) - 1 \leq 0, \\ h_2(x(1)) = x_2(1) - 1 \leq 0, \\ g(x, u) = x_1 - x_2 + u_1 + u_2 - u_3 - 1 \leq 0, \end{array} \right.$$

where L_i and ℓ_i are assumed to satisfy (H1) and (H2), respectively. In this problem we have $n = 2, m = 3$ and $r = 1$.

Let $(\bar{x}, \bar{u})$ be a feasible point of the problem. We show that the problem satisfies conditions (H1) – (H5). Assumption (H1) is straightforward. For (H2), we have

$$h'(\bar{x}(1)) = \begin{pmatrix} 1 & -2\bar{x}_2(1) \\ 0 & 1 \end{pmatrix}$$

and $\det h'(\bar{x}(1)) = 1$. We have $g_u[t] = [1, 1, -1]$ and so $R[t] = g_u[t]g_u[t]^T = 3$. Hence $\det R[t] = 3$ and $R[t]^{-1} = 1/3$. Thus the problem satisfies (H1) – (H3). For (H4), we have

$$\begin{aligned}
T[t] &= \{(w_1, w_2, w_3) \in \mathbb{R}^3 | w_1 + w_2 - w_3 = 0\} = \{(w_1, w_2, w_1 + w_2)\} \\
&= \text{span}([1, 0, 1]^T, [0, 1, 1]^T)
\end{aligned}$$

Hence $\dim T[t] = 2$ and so (H4) is valid. It remains to check (H5). From the above, we have

$$S[t] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \varphi_u[t] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B[t] = \varphi_u[t]S[t] = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

By (3.2.55), we have

$$\begin{aligned} A[t] &= \varphi_x[t] - \varphi_u[t]g_u[t]^T R[t]^{-1}g_x[t] \\ &= \begin{pmatrix} t & 0 \\ -1/3 & t \end{pmatrix} - \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} [1, 1, -1]^T \frac{1}{3} [1, -1] \\ &= \begin{pmatrix} t & 0 \\ 0 & t - \frac{1}{3} \end{pmatrix}. \end{aligned}$$

Let $\Omega(t, \tau)$ be the principal matrix solution to the system $\dot{x} = Ax$, By definition, each column of Ω is a solution of $\dot{x} = Ax$ and for $t = \tau$ these columns become $e_1 = [1, 0]$ and $e_2 = [0, 1]^T$. Hence we have

$$\Omega(t, \tau) = \begin{pmatrix} \exp(\frac{t^2}{2} - \frac{\tau^2}{2}) & 0 \\ 0 & \exp(\frac{t^2}{2} - \frac{t}{3} - \frac{\tau^2}{2} + \frac{\tau}{3}) \end{pmatrix}$$

and the mapping $H : L^\infty([0, 1], \mathbb{R}) \rightarrow \mathbb{R}^2$ is given by

$$\begin{aligned} H(v) &= \int_0^1 \Omega(1, \tau) B[\tau] v(\tau) d\tau \\ &= [\int_0^1 \alpha(\tau)(2v_1 + v_2) d\tau, \int_0^1 \beta(\tau)(v_1 + v_2) d\tau]^T, \end{aligned}$$

where $v = [v_1, v_2]^T$ and $\alpha(\tau) := \exp(\frac{1}{2} - \frac{\tau^2}{2})$ and $\beta(\tau) := \exp(\frac{1}{6} - \frac{\tau^2}{2} + \frac{\tau}{3})$. Let $\alpha_0 = \int_0^1 \alpha(\tau) d\tau$ and $\beta_0 = \int_0^1 \beta(\tau) d\tau$. Then $\alpha_0 > 0$ and $\beta_0 > 0$. For any $\xi = [\xi_1, \xi_2]^T \in \mathbb{R}^2$, we consider the equation

$$\begin{cases} \alpha_0(2v_1 + v_2) = \xi_1 \\ \beta_0(v_1 + v_2) = \xi_2. \end{cases}$$

Its solution is given by

$$v_1 = \frac{\xi_1}{\alpha_0} - \frac{\xi_2}{\beta_0}, \quad v_2 = \frac{2\xi_2}{\beta_0} - \frac{\xi_1}{\alpha_0} \quad (3.2.101)$$

which also satisfies the equation $Hv = \xi$. Hence H is onto. Consequently $(H5)$ is valid and Theorem 3.2.9 is applicable for the problem.

3.2.3 Application to sustainable energy management in smart grids

The transition to sustainable energy systems is a critical challenge of our time, requiring a delicate balance between multiple, often conflicting objectives. Smart grids, with their integration of various energy

sources and advanced control capabilities, offer a promising platform for addressing these challenges [83]. However, the complexity of these systems necessitates sophisticated optimization techniques [220]. In this section, we present and analyze a multiobjective optimal control problem for sustainable energy management in smart grids. Our model incorporates four key objectives: cost minimization, maximization of renewable energy usage, minimization of environmental impact, and maintenance of grid stability. By applying our theoretical results on KKT optimality conditions for multiobjective optimal control problems, we aim to provide insights into the structure of Pareto optimal solutions for this complex system. This application not only demonstrates the practical relevance of our theoretical framework but also contributes to the development of effective strategies for managing the intricate trade-offs in sustainable energy systems.

We now consider the following multiobjective optimal control problem:

$$\text{Min}_{\mathbb{R}_+^4} J(x, u) \quad (3.2.102)$$

$$\text{s.t.} \quad (3.2.103)$$

$$x(t) = x_0 + \int_0^t \varphi(s, x(s), u(s)) ds \quad \text{for all } t \in [0, 1], \quad (3.2.104)$$

$$h(x(1)) \leq 0, \quad (3.2.105)$$

$$g(t, x(t), u(t)) \leq 0 \quad \text{for a.a. } t \in [0, 1]. \quad (3.2.106)$$

Here, $x(t) = (x_1(t), x_2(t), x_3(t)) \in \mathbb{R}^3$ represents the state variables: energy stored in batteries (MWh), grid load (MW), and cumulative CO₂ emissions (tons), respectively. The control variables

$$u(t) = (u_1(t), u_2(t), u_3(t)) \in \mathbb{R}^3$$

represent solar, wind, and conventional energy inputs (MW). The specific components are defined as follows:

1. Objective functions:

$$J_1(x, u) = \int_0^1 [c_1 u_1^2(t) + c_2 u_2^2(t) + c_3 u_3^2(t)] dt, \quad J_2(x, u) = - \int_0^1 [u_1(t) + u_2(t)] dt,$$

$$J_3(x, u) = x_3(1), \quad J_4(x, u) = \int_0^1 [x_2(t) - x_2^{\text{target}}]^2 dt$$

where c_1 , c_2 , and c_3 are the unit costs of solar, wind, and conventional energy, respectively, and x_2^{target} is the target grid load.

2. System dynamics:

$$\varphi(t, x, u) = \begin{bmatrix} \eta(u_1 + u_2 + u_3 - x_2) \\ \varphi_2(t, x_2, u_1, u_2, u_3) \\ \alpha_3 u_3, \end{bmatrix}$$

where $\eta > 0$ is the energy storage efficiency, φ_2 is a function describing the grid load dynamics, and α_3 is the emission rate of conventional energy. Here

$$\varphi_2(t, x_2, u_1, u_2, u_3) = f(t, x_2) + b_1(t)u_1 + b_2(t)u_2 + b_3(t)u_3, \quad (3.2.107)$$

where f is a continuous function on $[0, 1] \times \mathbb{R}$ and $f(t, \cdot)$ is of class C^2 for each $t \in [0, 1]$, and $b_i(\cdot) \in L^\infty([0, 1], \mathbb{R})$ with $i = 1, 2, 3$.

3. Terminal constraints:

$$h(x(1)) = x_1(1) - x_1^{\max} \leq 0,$$

where $x_1^{\max}$ is the maximum battery capacity.

4. Path constraints:

$$\begin{aligned} g_1(t, x, u) &= u_1 - u_1^{\max} \leq 0 \\ g_2(t, x, u) &= u_2 - u_2^{\max} \leq 0 \\ g_3(t, x, u) &= tu_1 + tu_2 + u_3 - c(t) \leq 0 \end{aligned}$$

where $u_1^{\max}$ and $u_2^{\max}$ are the maximum production capacities for energy sources 1 and 2 respectively, and $c(t)$ is a time-varying function representing the upper bound of the combined energy input.

The objective functions, system dynamics, and constraints are defined as in the problem formulation. We now verify that this problem satisfies the necessary assumptions for our main theorem.

Verification of the Assumptions:

(H1) Smoothness and Lipschitz Continuity: The functions L , φ , and g are Carathéodory functions and C^2 in x and u . The cost functions, system dynamics, and constraints are well-behaved, satisfying the Lipschitz condition and boundedness of derivatives.

(H2) Smoothness of Terminal Conditions: The function $h(x(1))$ includes battery capacity constraints, which are linear and thus C^2 . The terminal cost $\ell(x(1))$ is linear in $x_3(1)$ and thus C^2 . Moreover,

$$\det h'(\bar{x}(1)) \neq 0$$

is satisfied as h is a linear function of $x(1)$.

(H3) Determinant Condition: Define $g_u[t] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t & t & 1 \end{bmatrix}$.

Then, $R[t] = g_u[t]g_u[t]^T = \begin{bmatrix} 1+t^2 & t^2 & t \\ t^2 & 1+t^2 & t \\ t & t & 1+t^2 \end{bmatrix}$. We have $\det(R[t]) = 1 + t^2 + 2t^4 \geq 1$ for all real t . Thus, (H3) is satisfied with $\gamma = 1$.

(H4) Rank Condition: With the given $g_u[t]$, we have $\text{rank}(g_u[t]) = 3 = m$, so (H4) is not applicable in this case.

(H4)' Alternative Condition: Since (H3) is satisfied and (H4) is not applicable, we verify (H4)'. To check (H4)' we need to show that, there exist $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ and $(\hat{u}_1, \hat{u}_2, \hat{u}_3)$ satisfying the following conditions:

$$\hat{x}_1(t) = \int_0^t \eta(\hat{u}_1 + \hat{u}_2 + \hat{u}_3 - \hat{x}_2)ds \quad (3.2.108)$$

$$\hat{x}_2(t) = \int_0^t (f_{x_2}[s]\hat{x}_2 + b_1\hat{u}_1 + b_2\hat{u}_2 + b_3\hat{u}_3)ds \quad (3.2.109)$$

$$\hat{x}_3(t) = \int_0^t \alpha_3\hat{u}_3ds, \quad (3.2.110)$$

$$h(\bar{x}_1(1)) + h'(\bar{x}_1(1))x(1) = \bar{x}_1(1) - x_1^{\max} + \hat{x}_1(1) < 0 \quad (3.2.111)$$

and

$$\bar{u}_1 - u^{\max} + \hat{u}_1 < -\epsilon_3, \quad (3.2.112)$$

$$\bar{u}_2 - u_2^{\max} + \hat{u}_2 < -\epsilon_2, \quad (3.2.113)$$

$$g_3[t] + t\hat{u}_1 + t\hat{u}_2 + \hat{u}_3 < -\epsilon_3 \quad (3.2.114)$$

for some $\epsilon_i > 0$ with $i = 1, 2, 3$.

Since $f_{x_2}[\cdot]$ is continuous on $[0, 1]$, there exists a number $a > 0$ such that $|f_{x_2}[t]| \leq a$ for all $t \in [0, 1]$. This is equivalent to $-a \leq -f_{x_2}[t] \leq a$ for all $t \in [0, 1]$.

Let us take $\hat{u}_i = -\gamma_i e^{\frac{t}{\gamma_i}}$ with $\gamma_i > 0$ for $i = 1, 2, 3$ and assume that $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ is a solution of (3.2.108)-(3.2.110) corresponding to $(\hat{u}_1, \hat{u}_2, \hat{u}_3)$.

Put $\hat{u} = \sum_{i=1}^3 b_i \hat{u}_i$ and $\mu(t) := \exp(-\int_0^t f_{2x_2}[s]ds)$. Then we have

$$e^{-a} \leq e^{-at} \leq \mu(t) \leq e^{at} \leq e^a \quad \forall t \in [0, 1]. \quad (3.2.115)$$

From equation (3.2.109), we have

$$\dot{\hat{x}}_2 - f_{x_2}[t]\hat{x}_2 = \hat{u}.$$

Multiplying both sides by $\mu(t)$, we get

$$\mu(t)\dot{\hat{x}}_2(t) - f_{x_2}[t]\hat{x}_2(t)\mu(t) = \hat{u}(t)\mu(t).$$

This implies that

$$\frac{d}{dt}(\mu(t)\hat{x}_2(t)) = \hat{u}(t)\mu(t).$$

Integrating on $[0, t]$ with $t \in [0, 1]$, we obtain

$$\hat{x}_2(t) = \frac{\int_0^t \hat{u}(s)\mu(s)ds}{\mu(t)}.$$

Inserting $\hat{x}_2$ into (3.2.108) and using (3.2.115), we have

$$\begin{aligned} \hat{x}_1(1) &= -\eta \sum_{i=1}^3 \gamma_i^2 (e^{1/\gamma_i} - 1) - \eta \int_0^1 \frac{\int_0^t \mu(s)\hat{u}(s)ds}{\mu(t)} dt \\ &= -\eta \sum_{i=1}^3 \gamma_i^2 (e^{1/\gamma_i} - 1) + \eta \int_0^1 \frac{\int_0^t \sum_{i=1}^3 b_i \gamma_i e^{s/\gamma_i} \mu(s) ds}{\mu(t)} dt \\ &\leq -\eta \sum_{i=1}^3 \gamma_i^2 (e^{1/\gamma_i} - 1) + \eta \int_0^1 \frac{\int_0^t \sum_{i=1}^3 \|b_i\|_\infty \gamma_i e^{s/\gamma_i} e^a ds}{e^{-a}} dt \\ &= -\eta \sum_{i=1}^3 \gamma_i^2 (e^{1/\gamma_i} - 1) + \eta e^{2a} \int_0^1 \sum_{i=1}^3 \|b_i\|_\infty \gamma_i^2 (e^{t/\gamma_i} - 1) dt \\ &= -\eta \sum_{i=1}^3 \gamma_i^2 (e^{1/\gamma_i} - 1) + \eta e^{2a} \sum_{i=1}^3 \|b_i\|_\infty \gamma_i^2 [\gamma_i (e^{1/\gamma_i} - 1) - 1] \\ &= -\eta \sum_{i=1}^3 \gamma_i^2 (e^{1/\gamma_i} - 1) (1 - e^{2a} \|b_i\|_\infty \gamma_i) - \eta e^{2a} \sum_{i=1}^3 \|b_i\|_\infty \gamma_i^2. \end{aligned}$$

By choosing $\gamma_i > 0$ small enough such that $1 - e^{2a} \|b_i\|_\infty \gamma_i \geq 0$, we see that $\hat{x}_1(1) < 0$. Hence (3.2.111) is valid. Since $e^{\frac{t}{\gamma_i}} \geq 1$ for all $t \in [0, 1]$, we have $\hat{u}_i < -\gamma_i$. This implies that 3.2.112-(3.2.114) are valid with $\epsilon_i = \gamma_i$. Consequently, $(H4)'$ is fulfilled.

Thus our problem satisfies assumptions $(H1)$ – $(H3)$ and $(H4)'$. Therefore, we can apply our main theorem to derive necessary optimality conditions for Pareto optimal solutions.

Theorem 3.2.15 (KKT Conditions for Sustainable Energy Management). *Let $\bar{z} = (\bar{x}, \bar{u})$ be a locally weak Pareto optimal solution of the sustainable energy management problem.*

Then there exist a vector $\lambda \in \mathbb{R}_+^4$ with $|\lambda| = 1$, a number $l \in \mathbb{R}$, an absolutely continuous function $p : [0, 1] \rightarrow \mathbb{R}^3$, and a vector function $\theta = (\theta_1, \theta_2, \theta_3) \in L^\infty([0, 1], \mathbb{R}^3)$, such that the following conditions are satisfied:

(i) *Adjoint equation:*

$$\begin{aligned} \dot{p}_1(t) &= 0, \quad \dot{p}_2(t) = \eta p_1(t) - f_{x_2}[t]p_2(t) + 2\lambda_4(\bar{x}_2(t) - x_2^{target}), \quad \dot{p}_3(t) = 0, \\ p_1(1) &= -l, \quad p_2(1) = 0, \quad p_3(1) = -\lambda_3 \end{aligned}$$

(ii) *Stationary condition in u :*

$$\begin{aligned} 2\lambda_1 c_1 \bar{u}_1 - \lambda_2 - \eta p_1 - b_1 p_2 + \theta_1 + t\theta_3 &= 0 \\ 2\lambda_1 c_2 \bar{u}_2 - \lambda_2 - \eta p_1 - b_2 p_2 + \theta_2 + t\theta_3 &= 0 \\ 2\lambda_1 c_3 \bar{u}_3 - \eta p_1 - b_2 p_2 - \alpha_3 p_3 + \theta_3 &= 0 \end{aligned}$$

(iii) *Complementary conditions:*

$$\begin{aligned} l &\geq 0, \quad l(\bar{x}_1(1) - x_1^{\max}) = 0 \\ \theta_1(t) &\geq 0, \quad \theta_1(t)(\bar{u}_1(t) - u_1^{\max}) = 0 \\ \theta_2(t) &\geq 0, \quad \theta_2(t)(\bar{u}_2(t) - u_2^{\max}) = 0 \\ \theta_3(t) &\geq 0, \quad \theta_3(t)(t\bar{u}_1(t) + t\bar{u}_2(t) + \bar{u}_3(t) - c(t)) = 0 \end{aligned}$$

(iv) *Nonnegative second-order condition:*

$$\lambda_1 \int_0^1 (c_1 \tilde{u}_1^2 + c_2 \tilde{u}_2^2 + c_3 \tilde{u}_3^2) dt + 2\lambda_4 \int_0^1 \tilde{x}_2^2 dt - \int_0^1 p_2 f_{x_2 x_2}[t] \tilde{x}_2^2 dt \geq 0$$

Proof. The proof follows directly from Theorem 3.2.9. □

CHAPTER 4

OPTIMAL CONTROL PROBLEMS GOVERNED BY PDE

This chapter provides some results of optimal control problems governed by partial differential equations. Its content is based on our published papers [130], [118] and [136].

4.1 Pontryagin's principle for optimal control problem governed by 3D Navier-Stokes equations

4.1.1 Introduction

Better understanding of optimality conditions of optimal control problems governed by Navier–Stokes equations is an ongoing topic of research. This question is of great importance in theory and in applications. It has been studied by many mathematicians recently. For the papers which have a close connection to the topic, we refer the readers to [16]–[223] and references given therein. Let us give some short view on these papers. Among them, [16] studied necessary optimality conditions for the problem with 2D Navier–Stokes equation without control constraints, [24] established the existence and uniqueness of solutions for optimal control problems for the 2D Navier–Stokes equations in a 2D domain, [43, 62, 63] gave optimality conditions for the problems governed by stationary Navier–Stokes equations, [217] established the existence of a weak solution of a free boundary problem for a 3D non-stationary Navier–Stokes equations with measure data, and [223] gave necessary optimality conditions for optimal control problems governed by 3D Navier–Stokes equations with some types of state constraints but without pointwise control constraint. Particularly, Casas [38] gave a result on the existence of solution and first-order necessary optimality conditions for the problem with a special cost function. Besides, the proof of [38] is also applicable to functionals that assert strong solutions.

To the best of our knowledge, the establishment of Pontryagin's maximum principle for optimal control problems governed by 3D Navier–Stokes

equations has been an open problem so far. In this section we will address this problem.

As usual we use the diffuse variation method (see for examples [44]–[189]), techniques from optimal control theory and partial differential equations in order to establish the Pontryagin maximum principle. However, in the case of optimal control problems governed by the 3D Navier–Stokes equations we will meet some difficulties. One of the difficulties is to expand the control-to-state mapping in the first-order form of Taylor’s expansion with respect to variation of optimal control. To overcome this obstacle, we have to use results on compact imbeddings, prove the existence of global strong solutions to the linearized Navier–Stokes equations and give some bounds for these solutions (see Lemma 2.4).

Let Ω be a bounded domain in $\mathbb{R}^3$ with boundary Γ of class C^2 . We consider the problem of finding a control $u \in L^2(0, T; L^2(\Omega)^3)$ and a state $y \in W^{1,2}(0, T; W, H)$ which minimize the cost function

$$J(y, u) = \int_Q L(x, t, y(x, t), u(x, t)) \, dxdt \quad (4.1.1)$$

with the state equations

$$\frac{dy}{dt} - \nu \Delta y + (y \cdot \nabla)y + \nabla \phi = Eu \quad \text{in } Q := \Omega \times (0, T), \quad (4.1.2)$$

$$y = 0 \quad \text{on } \Sigma := \Gamma \times (0, T), \quad (4.1.3)$$

$$\operatorname{div} y = 0 \quad \text{in } Q, \quad (4.1.4)$$

$$y(0) = y_0 \quad \text{in } \Omega, \quad (4.1.5)$$

and the pointwise constraint

$$u(x, t) \in \mathcal{M} \quad \text{a.a. } (x, t) \in Q. \quad (4.1.6)$$

Here, ν is a positive constant, $y_0 \in V$, $\mathcal{M}$ is a closed set in $\mathbb{R}^3$,

$$L : Q \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$$

is a given function, and

$$E : L^2(0, T; L^2(\Omega)^3) \rightarrow L^2(0, T; H)$$

is a linear bounded operator. Hereafter, we shall use the following func-

tion spaces:

$$\mathbf{L}^2(\Omega) := (L^2(\Omega))^3 = L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega),$$

$$\mathbf{H}_0^1(\Omega) := (H_0^1(\Omega))^3 = H_0^1(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega),$$

$$\mathbf{H}^m(\Omega) := (H^m(\Omega))^3, \quad \mathbf{W}_0^{m,s}(\Omega) := (W_0^{m,s}(\Omega))^3,$$

$$\mathcal{V} := \{y \in D(\Omega)^3 \mid \operatorname{div} y = 0\},$$

$$V := \text{the closure of } \mathcal{V} \text{ in } \mathbf{H}_0^1(\Omega) = \{y \in \mathbf{H}_0^1(\Omega) \mid \operatorname{div} y = 0\},$$

$$H := \text{the closure of } V \text{ in } \mathbf{L}^2(\Omega) = \{y \in \mathbf{L}^2(\Omega) \mid \operatorname{div} y = 0, y \cdot \vec{n} = 0 \text{ on } \Gamma\},$$

$$H^\perp := \{\phi \in \mathbf{L}^2(\Omega) \mid \phi = \nabla p, \quad p \in H^1(\Omega)\},$$

$$W := V \cap \mathbf{H}^2(\Omega), \quad L_0^2(\Omega) := L^2(\Omega)/\mathfrak{R} = \{p \in L^2(\Omega) \mid \int_\Omega p(x) dx = 0\},$$

V^* and W^* are the dual spaces of V and W , respectively,

$$W^{1,2}(0, T; W, H) := \{y \in L^2(0, T; W) \mid \frac{dy}{dt} \in L^2(0, T; H)\},$$

$$W^{1,2}(0, T; H, W^*) := \{y \in L^2(0, T; H) \mid \frac{dy}{dt} \in L^2(0, T; W^*)\}.$$

Here, $\vec{n}$ is the unit outward normal vector on Γ .

Problem (4.1.1)–(4.1.6) comes from fluid mechanics, which minimizes the flow in a 3D channel Ω . A standard example for the choice of L is the quadratic function

$$L(x, t, y, u) = \frac{1}{2} \int_Q |y(x, t) - y_d(x, t)|^2 dxdt + \frac{1}{2} \int_Q |u(x, t)|^2 dxdt,$$

where $y_d \in L^2(0, T; L^2(\Omega)^3)$ is given.

4.1.2 Linearized and adjoint equations

Throughout this section, we shall denote by $(\cdot, \cdot)$ and $|\cdot|$ the scalar product and the norm in H , respectively. Also, we denote by $((\cdot, \cdot))$ and $\|\cdot\|$ the scalar product and the norm in V , respectively, where

$$((v, w)) = \sum_{i=1}^3 (D_i v, D_i w), \quad \forall v, w \in V.$$

It is well known that the embeddings

$$W \hookrightarrow V \hookrightarrow H$$

are compact and each space is dense in the following one. Let us denote by $\mathbb{P} : \mathbf{L}^2(\Omega) \rightarrow H$ the Leray projection on H , and by $A : W \rightarrow H$,

$A = -\mathbb{P}\Delta$ the Stokes operator. Define

$$B : W \times W \rightarrow H, \quad B(v, w) = \mathbb{P}((v \cdot \nabla)w).$$

By applying the Leray projection on both sides of (4.1.2), we obtain the following equivalent equation of state Eqs. (4.1.2)–(4.1.5):

$$\frac{dy}{dt} + \nu Ay + B(y, y) = Eu, \quad y(0) = y_0. \quad (4.1.7)$$

Recall that given a vector function $u \in L^2(0, T; L^2(\Omega)^3)$, the vector function $y \in W^{1,2}(0, T; W, H)$ satisfying Eq. (4.1.7) for a.e. $(x, t) \in Q$, is said to be a strong solution of the Navier–Stokes equations. It is known that under smallness conditions of y_0 and Eu , Eq (4.1.7) has a unique strong solution (see [59, Theorem 9.3]).

Let us define a trilinear mapping

$$\mathbf{b} : \mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega) \rightarrow \mathbb{R}$$

by setting

$$\mathbf{b}(y, z, w) = \sum_{i,j=1}^3 \int_{\Omega} y_i D_i z_j w_j dx, \quad \forall (y, z, w) \in \mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega).$$

According to Lemma 1.2 and Lemma 1.3 in Chapter II of [216], $\mathbf{b}$ is a trilinear continuous mapping, and there exists an absolute constant $C_0 > 0$ such that the following properties are satisfied:

$$\begin{aligned} |\mathbf{b}(y, z, w)| &\leq C_0 \|y\|_{\mathbf{H}_0^1(\Omega)} \|z\|_{\mathbf{H}_0^1(\Omega)} \|w\|_{\mathbf{H}_0^1(\Omega)}, \quad \forall y, z, w \in \mathbf{H}_0^1(\Omega), \\ \mathbf{b}(y, v, v) &= 0, \quad \forall y \in V, v \in \mathbf{H}_0^1(\Omega), \\ \mathbf{b}(y, v, w) &= -\mathbf{b}(y, w, v), \quad \forall y \in V, v, w \in \mathbf{H}_0^1(\Omega). \end{aligned}$$

From [59, Theorem 4.4], the operator $A^{-1} : H \rightarrow W$ is self-adjoint, injective, and compact. By a well-known theorem of Hilbert (see [31, Theorem 6.11]), there exists a sequence of positive numbers $\mu_j > 0$ with $\mu_{j+1} < \mu_j$ and an orthogonal basis of H , (w_j) such that $A^{-1}(w_j) = \mu_j w_j$. We denote by $\lambda_j = \mu_j^{-1}$. Since A^{-1} has range in W , we obtain that

$$\begin{aligned} Aw_j &= \lambda_j w_j, \quad w_j \in W, \\ 0 &< \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots, \\ \lim_{j \rightarrow \infty} \lambda_j &= +\infty, \\ (w_j)_{j=1, \dots} &\text{ are forms an orthogonal basis of } H. \end{aligned}$$

Let $\alpha > 0$ be a real number. We define the operator A^α by

$$A^\alpha v = \sum_{j=1}^{\infty} \lambda_j^\alpha \xi_j w_j, \quad v = \sum_{j=1}^{\infty} \xi_j w_j, \quad y \in \mathcal{D}(A^\alpha) \quad (4.1.8)$$

$$\mathcal{D}(A^\alpha) = \left\{ v \in H \mid v = \sum_{j=1}^{\infty} \xi_j w_j, \quad \sum_{j=1}^{\infty} \lambda_j^{2\alpha} |\xi_j|^2 < +\infty \right\}. \quad (4.1.9)$$

The space $\mathcal{D}(A^\alpha)$ carries a natural scalar product

$$\langle u, v \rangle_\alpha = \sum_{j=1}^{\infty} \lambda_j^{2\alpha} \xi_j \eta_j, \quad \text{if } u = \sum_{j=1}^{\infty} \xi_j w_j, \quad v = \sum_{j=1}^{\infty} \eta_j w_j.$$

In particular, for $\alpha = \frac{1}{2}$, we have $\mathcal{D}(A^{1/2}) = \mathbf{V}$ and $\langle u, v \rangle_{1/2} = ((u, v))$. Moreover, the vectors $v_j := \lambda_j^{-1/2} w_j$ form a complete orthogonal system in V . When $\alpha = 1$, we also have

$$\mathcal{D}(A) = W, \quad \|w\|_W^2 = \sum_{j=1}^{\infty} \lambda_j^2 \xi_j^2, \quad \text{for } w = \sum_{j=1}^{\infty} \xi_j w_j.$$

Hence, from [155, Proposition 2.1], we have

$$[\mathcal{D}(A), H]_{1/2} = [W, H]_{1/2} = V.$$

where $[\mathcal{D}(A), H]_{1/2}$ is an interpolation space. By [155, Theorem 3.1], the embedding

$$W^{1,2}(0, T; W, H) \hookrightarrow C(0, T; [W, H]^{1/2}) = C(0, T; V) \quad (4.1.10)$$

is continuous. Based on this embedding, we have

Lemma 4.1.1. *If $y \in W^{1,2}(0, T; W, H)$ then $y_t + \nu Ay + B(y, y) \in L^2(0, T; H)$.*

Proof. Since $|Ay| = |P\Delta y| \leq \|y\|_W$, we have $\int_0^T |Ay|^2 dt < \int_0^T \|y\|_W^2 dt < +\infty$. Hence $Ay \in L^2(0, T; H)$. From the embedding $\mathbf{H}_0^1(\Omega) \hookrightarrow L^6(\Omega)^3$, there exists a constant $C > 0$ such that

$$|B(y, y)|^2 \leq \sum_{i=1}^3 \int_{\Omega} (y_i D_i y)^2 dx \leq \sum_{i=1}^3 \left(\int_{\Omega} y_i^6 dx \right)^{1/3} \left(\int_{\Omega} (D_i y)^3 dx \right)^{2/3}.$$

Combining this with embedding (4.1.10) yields

$$\int_0^T |B(y, y)|^2 dt \leq \|y\|_{C(0, T; V)}^2 \int_0^T \|y\|_W^2 dt < +\infty.$$

We obtain the conclusion. □

Let us define a mapping

$$\Phi : W^{1,2}(0, T; W, H) \times L^2(0, T; L^2(\Omega)^3) \rightarrow L^2(0, T; H) \times V$$

by setting

$$\Phi(y, u) = \left(\frac{dy}{dt} + \nu Ay + B(y, y) - Eu, y(0) - y_0 \right).$$

The following proposition gives us some properties of Φ .

Proposition 4.1.2. *The mapping Φ is of class C^1 and for each $(\hat{y}, \hat{u}) \in W^{1,2}(0, T; W, H) \times L^2(0, T; L^2(\Omega)^3)$,*

$$\nabla_y \Phi(\hat{y}, \hat{u}) : W^{1,2}(0, T; W, H) \rightarrow L^2(0, T; H) \times V$$

is an isomorphism.

Proof. It is easy to see that the mapping Φ is of class C^1 and its derivatives are given by

$$\nabla \Phi(y, u) = \begin{bmatrix} \frac{d}{dt} + \nu A + B(\delta y_0, \cdot) + B(\cdot, y) & -E \\ \delta_0 & 0 \end{bmatrix}. \quad (4.1.11)$$

where δ_0 is a continuous linear mapping which is defined by $\langle \delta_0, y \rangle = y(0)$. To prove that $\nabla_y \Phi(\hat{y}, \hat{u})$ is bijective, we show that for each $(f, v_0) \in L^2(0, T; H) \times V$, the equation

$$\nabla_y \Phi(\hat{y}, \hat{u})y = (f, v_0)$$

has a unique solution. This equation is equivalent to

$$\frac{dy}{dt} + \nu Ay + B(\hat{y}, y) + B(y, \hat{y}) = f, \quad y(0) = v_0.$$

By Lemma 4.1.3, we obtain the desired conclusion. $\square$

Lemma 4.1.3. *Let $\hat{y} \in W^{1,2}(0, T; W, H)$. Then there exists a unique solution $y \in W^{1,2}(0, T; W, H)$ of the equation*

$$\frac{dy}{dt} + \nu Ay + B(\hat{y}, y) + B(y, \hat{y}) = f, \quad y(0) = v_0. \quad (4.1.12)$$

Moreover, there exist positive constants C, C_1, C_2 , and C_3 depending on ν, Ω such that

$$\begin{aligned} & \|y\|_{L^2(0, T; W)}^2 + \|\dot{y}\|_{L^2(0, T; H)}^2 \\ & \leq (\exp(K_1)K_2K_1 + K_2) \left(\frac{2C}{\nu} + \frac{8CC_3}{\nu} \|\hat{y}\|_{C^2(0, T; V)}^{C_2} + 8\nu \right) \\ & + 4C_3 \exp(K_1)K_2 \int_0^T \|\hat{y}\|_W^2 dt + 4 \int_0^T |f|^2 dt, \end{aligned} \quad (4.1.13)$$

where K_1 and K_2 are given by

$$K_1 = C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4, \quad K_2 = \|v_0\|^2 + \frac{2}{\nu} \int_0^T \|f\|^2 ds.$$

Proof. The proof of the lemma proceeds analogously to the proof of Lemma 4.1.7 with $\rho = 0$, where the proof is based on the Galerkin method. Estimation (4.1.13) is standard in the theory of Navier–Stokes equations. $\square$

Remark 4.1.4. *It is possible to define a linear, coercive, and continuous mapping $\mathcal{A}(t) = \nu A + B(\bar{y}, y) + B(y, \bar{y}) : V \rightarrow V^*$ and reduce the linear equation (4.1.12) to the Cauchy problem*

$$\frac{dy}{dt} + \mathcal{A}(t)y = f, \quad y(0) = y_0,$$

where $f \in L^2(0, T; V^*)$ and $V \subset H \subset V^*$. Then, by a classical existence result of J.L. Lions (see Theorem 10.9 of [31]), there exists a unique weak solution y satisfying $y \in L^2(0, T; V) \cap C([0, T], H)$, $\frac{dy}{dt} \in L^2(0, T; V^*)$. However, by this way, we shall meet some difficulty in showing that y is a strong solution, i.e., $y \in W^{1,2}(0, T; W, H)$.

Let $\hat{y} \in W^{1,2}(0, T; W, H)$ and $f_* \in L^2(0, T; W^*)$. We consider the adjoint equation

$$-\dot{\vartheta} + \nu A^* \vartheta + B(\hat{y}, \cdot)^* \vartheta + B(\cdot, \hat{y})^* \vartheta = f_*, \quad \vartheta(T) = 0. \quad (4.1.14)$$

Here, A^* denotes the adjoint mapping of A .

Definition 4.1.5. *A vector function $\vartheta \in W^{1,2}(0, T; H, W^*)$ is said to be a solution of the adjoint equation if (4.1.14) is satisfied for almost every $t \in [0, T]$.*

Lemma 4.1.6. *For each $f_* \in L^2(0, T; W^*)$, Eq. (4.1.14) has a unique solution $\vartheta \in W^{1,2}(0, T; H, W^*)$. Moreover, the following integration by parts formula is valid:*

$$\int_0^T (\dot{\vartheta}(t), z(t)) dt = - \int_0^T (\vartheta(t), \dot{z}(t)) dt - (\vartheta(0), z(0)), \quad \forall z \in W^{1,2}(0, T; W, H). \quad (4.1.15)$$

Proof. Since $\nabla_y \Phi(\hat{y}, \hat{u}) : W^{1,2}(0, T; W, H) \rightarrow L^2(0, T; H) \times V$ is an isomorphism, its adjoint mapping

$$\nabla_y \Phi(\hat{y}, \hat{u})^* : L^2(0, T; H) \times V^* \rightarrow W^{1,2}(0, T; W, H)^*$$

is also an isomorphism. From (4.1.11), we have

$$\nabla_y \Phi(\hat{y}, \hat{u}) = \left(\frac{d}{dt} + \nu A + B(\hat{y}, \cdot) + B(\cdot, \hat{y}), \quad \delta_0 \right).$$

By some standard arguments, we find that

$$\nabla_y \Phi(\hat{y}, \hat{u})^*(\varphi, v^*) = \left(\frac{d}{dt} \right)^* \varphi + \nu A^* \varphi + B(\hat{y}, \cdot)^* \varphi + B(\cdot, \hat{y})^* \varphi + \delta_0^* v^*, \quad (4.1.16)$$

for $(\varphi, v^*) \in L^2(0, T; H) \times V^*$. Note that the embedding $L^2(0, T; W^*) \hookrightarrow W^{1,2}(0, T; W, H)^*$ is continuous. Hence, if $f_* \in L^2(0, T; W^*)$, then $f_* \in W^{1,2}(0, T; W, H)^*$. Since $\nabla_y \Phi(\hat{y}, \hat{u})^*$ is an isomorphism, there exists a unique $(\vartheta, v^*) \in L^2(0, T; H) \times V^*$ such that

$$\left(\frac{d}{dt} \right)^* \vartheta + \nu A^* \vartheta + B(\hat{y}, \cdot)^* \vartheta + B(\cdot, \hat{y})^* \vartheta + \delta_0^* v^* = f_*.$$

This means that for any $z \in W^{1,2}(0, T; W, H)$, one has

$$\begin{aligned} \int_0^T (\vartheta(t), \dot{z}(t)) dt + \int_0^T (\vartheta(t), \nu A z(t) + B(\hat{y}(t), z(t)) + B(z(t), \hat{y}(t))) dt \\ + \langle v^*, z(0) \rangle = \int_0^T \langle f_*(t), z(t) \rangle dt. \end{aligned} \quad (4.1.17)$$

Taking any $w \in W$, $\varphi \in C_0^\infty(0, T)$ and substituting $z = \varphi w$ into (4.1.17), we have

$$\begin{aligned} \int_0^T (\vartheta(t), w) \dot{\varphi}(t) dt + \int_0^T (\vartheta(t), \nu A w + B(\hat{y}(t), w) + B(w, \hat{y}(t))) \varphi(t) dt \\ = \int_0^T \langle f_*(t), w \rangle \varphi(t) dt. \end{aligned} \quad (4.1.18)$$

Note that $\varphi w \in W^{1,2}(0, T; W, H)$. It follows that

$$\frac{d}{dt} (\vartheta(t), w)_H = \langle \dot{\vartheta}(t), w \rangle = (\nu A^* \vartheta + B(\hat{y}, \cdot)^* \vartheta + B(\cdot, \hat{y})^* \vartheta, w) - \langle f_*(t), w \rangle. \quad (4.1.19)$$

Since $\nu A^* \vartheta + B(\hat{y}, \cdot)^* \vartheta + B(\cdot, \hat{y})^* \vartheta$ and f_* belong to $L^2(0, T; W^*)$, we have $\dot{\vartheta} \in L^2(0, T; W^*)$ and the first equation of (4.1.14) is valid. We now take a continuously differentiable φ satisfying $\varphi(0) = 0$ and $\varphi(T) = 1$. Substituting $z = \varphi(t)w$ into (4.1.17) and integrating by parts, we get

$$\begin{aligned} - \int_0^T \frac{d}{dt} (\vartheta(t), w) \varphi(t) dt + (\vartheta(T), w) + \int_0^T (\vartheta(t), \nu A w + B(\hat{y}(t), w) \\ + B(w, \hat{y}(t))) \varphi(t) dt = \int_0^T \langle f_*(t), w \rangle \varphi(t) dt. \end{aligned}$$

Combining this with (4.1.19) yields $\langle \vartheta(T), w \rangle = (\vartheta(T), w) = 0$. Since W is dense in H , we have $\langle \vartheta(T), h \rangle = 0$ for all $h \in H$. Hence, $\vartheta(T) = 0$. Thus, $\vartheta \in W^{1,2}(0, T; H, W^*)$, which is a unique solution of (4.1.14). It remains to show that (4.1.15) is fulfilled. Take any $w \in W$ and a continuously differentiable function ψ on $[0, T]$ satisfying $\psi(T) = 0$ and $\psi(0) = 1$. Substituting $\psi(t)w$ into (4.1.17) and integrating by parts, we get

$$\begin{aligned} & - \int_0^T \frac{d}{dt} (\vartheta(t), w) \psi(t) dt - (\vartheta(0), w) + \int_0^T (\vartheta(t), \nu Aw + B(\hat{y}(t), w) \\ & + B(w, \hat{y}(t))) \psi(t) dt + \langle v^*, w \rangle = \int_0^T \langle f_*(t), w \rangle \psi(t) dt. \end{aligned}$$

Combining this with (4.1.19) yields $\langle v^* - \vartheta(0), w \rangle = 0$. Since W is dense in V , we get $\langle v^* - \vartheta(0), v \rangle = 0$ for all $v \in V$. Hence, $v^* = \vartheta(0)$. Substituting this into (4.1.17) and using (4.1.14), we obtain

$$\int_0^T (\vartheta(t), \dot{z}(t)) dt + \int_0^T \langle \dot{\vartheta}(t), z(t) \rangle dt + (\vartheta(0), z(0)) = 0,$$

which yields formula (4.1.15). $\square$

Lemma 4.1.7. *Let $\hat{y} \in W^{1,2}(0, T; W, H)$ and $f \in L^2(0, T; H)$. Then there exist constants C , C_1 , and $C_2 > 0$ depending on ν and Ω such that for any $\rho \geq 0$ satisfying*

$$\rho^2 \left(\|f\|_{L^2(0, T; H)}^2 + 1 \right) < \frac{\nu^{3/2} \lambda_1}{4C_2^{1/2}} \exp \left(-C_1 \|\hat{y}\|_{L^2(0, T; W)}^2 - C_2 T \|\hat{y}\|_{C(0, T; V)}^4 \right), \quad (4.1.20)$$

the equation

$$\frac{dy}{dt} + \nu Ay + \rho B(y, y) + B(\hat{y}, y) + B(y, \hat{y}) = f, \quad y(0) = 0 \quad (4.1.21)$$

has a unique solution $y \in W^{1,2}(0, T; W, H)$, which satisfies the following estimation:

$$\begin{aligned} & \|y\|_{L^2(0, T; W)}^2 + \|\dot{y}\|_{L^2(0, T; H)}^2 \leq \\ & M_2(\rho, \hat{y}, f) \left(\frac{20\rho^2 C}{\nu^2} \exp(M_1(\hat{y})) (\|f\|_{L^2(0, T; H)}^2 + 1) + \frac{5C}{\nu} \|\hat{y}\|_{C(0, T; V)}^2 + \frac{C}{\nu} + 5\nu \right) \\ & + \frac{20}{\nu} \|\hat{y}\|_{L^2(0, T; W)}^2 \exp(M_1(\hat{y})) (\|f\|_{L^2(0, T; H)}^2 + 1) + 5\|f\|_{L^2(0, T; H)}^2, \end{aligned} \quad (4.1.22)$$

where

$$\begin{aligned} M_1(\hat{y}) &:= C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4, \\ M_2(\rho, \hat{y}, f) &:= C_2 T \rho^4 \exp(3M_1(\hat{y})) \frac{4^3}{\nu^3} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)^3 \\ &\quad + T M_1(\hat{y}) \exp(M_1(\hat{y})) \frac{4}{\nu} (\|f\|_{L^2(0,T;H)}^2 + 1) + \frac{4}{\nu} \|f\|_{L^2(0,T;H)}^2. \end{aligned}$$

Proof. We first show that there exists a constant $C > 0$ depending on Ω such that

$$|\mathbf{b}(z, \hat{y}, Az)| \leq C \|z\| \|\hat{y}\|_W |Az|, \quad \forall z \in W, \quad (4.1.23)$$

$$|\mathbf{b}(v, z, Az)| \leq C \|v\| \|z\|^{1/2} |Az|^{3/2}, \quad \forall v \in V, z \in W. \quad (4.1.24)$$

In fact, using the embedding $H^1(\Omega) \hookrightarrow L^6(\Omega)$, we have

$$\begin{aligned} |\mathbf{b}(z, \hat{y}, Az)| &= \sum_{i,j=1}^3 \int_{\Omega} (z_i D_i \hat{y}_j (Az)_j) dx \leq \sum_{i,j=1}^3 \|z_i D_i \hat{y}_j\|_{L^2(\Omega)} \|(Az)_j\|_{L^2(\Omega)} \\ &\leq \sum_{i,j=1}^3 \|z_i\|_{L^6(\Omega)} \|D_i \hat{y}_j\|_{L^3(\Omega)} \|(Az)_j\|_{L^2(\Omega)} \\ &\leq C \|z\| \|\hat{y}\|_W |Az|, \end{aligned}$$

which proves (4.1.23). For (4.1.24), we notice that $C^\infty(\bar{\Omega})^3$ is dense in $H^1(\Omega)^3$ and $H^2(\Omega)^3$ (see [79, Theorem 3, p. 252]). Therefore, from inequality (61) with $s_1 = 1$, $s_2 = 1/2$, and $s_3 = 0$, we obtain

$$|\mathbf{b}(v, z, Az)| \leq C \|v\| \|z\|^{1/2} \|z\|_{\mathbf{H}^2(\Omega)}^{1/2} |Az|, \quad \forall v \in V, z \in W.$$

By [59, Proposition 4.7], we have

$$\|z\|_{\mathbf{H}^2(\Omega)} \leq C |Az|, \quad \forall z \in W. \quad (4.1.25)$$

It follows that

$$|\mathbf{b}(v, z, Az)| \leq C \|v\| \|z\|^{1/2} |Az|^{3/2}, \quad \forall v \in V, z \in W,$$

for some constant $C > 0$. We obtain (4.1.24). Let us put

$$C_1 = \frac{4C^2}{\nu}, \quad C_2 = \frac{108C^4}{\nu^3} \quad (4.1.26)$$

and fix any $\rho \geq 0$ satisfying (4.1.20).

Existence. Since $C([0, T]; H)$ is dense in $L^2(0, T; H)$ (see [237, Proposition 23.2, p. 407]), we can assume that $f \in C([0, T]; H)$. Let $\{w_1, w_2, \dots\}$

be an orthogonal basis of H , where w_j is the j' th eigenvector of A corresponding to eigenvalue λ_j for $j = 1, 2, \dots$. Consider the subspace $H_m = \text{span}\{w_1, w_2, \dots, w_m\}$, which is spanned by the first m eigenvectors of A . Let $P_m : H \rightarrow H_m$ be the orthogonal projection onto H_m . We consider the Galerkin system

$$\frac{dy_m}{dt} + \nu A y_m + \rho P_m B(y_m, y_m) + P_m B(\hat{y}, y_m) + P_m B(y_m, \hat{y}) = g_m, \quad (4.1.27)$$

$$y_m(0) = 0, \quad (4.1.28)$$

where $y_m \in H_m$ with $y_m = \sum_{j=1}^m \xi_j(t) w_j$ and

$$g_m = P_m f = \sum_{j=1}^m \eta_j(t) w_j.$$

Then (4.1.27) and (4.1.28) are equivalent to

$$\begin{aligned} \frac{d\xi_j}{dt} + \nu \lambda_j \xi_j + \rho \sum_{k,l=1}^m b(w_k, w_l, w_j) \xi_k \xi_l + \sum_{l=1}^m (b(\hat{y}, w_l, w_j) \\ + \mathbf{b}(w_l, \hat{y}, w_j)) \xi_l = \eta_j, \quad \xi_j(0) = 0, \quad j = 1, 2, \dots, m, \end{aligned}$$

or equivalently,

$$\dot{\xi}_j(t) = F_j(t, \xi), \quad \xi_j(0) = 0, \quad j = 1, 2, \dots, m, \quad (4.1.29)$$

where $\xi = (\xi_1, \dots, \xi_m)$ and

$$\begin{aligned} F_j(t, \xi) = \eta_j - \\ - \left(\nu \lambda_j \xi_j + \rho \sum_{k,l=1}^m \mathbf{b}(w_k, w_l, w_j) \xi_k \xi_l + \sum_{l=1}^m (\mathbf{b}(\hat{y}, w_l, w_j) + \mathbf{b}(w_l, \hat{y}, w_j)) \xi_l \right) \end{aligned}$$

Since $f \in C([0, T]; H)$, $\eta_j(t)$ is a continuous function. Let $\{t_k\} \subset [0, T]$ be a sequence such that $t_k \rightarrow t$. Using inequality (61) with $s_1 = 1$, $s_2 = 1/2$, $s_3 = 0$ and a property of $\mathbf{b}$, we have:

$$\begin{aligned} |\mathbf{b}(w_l, \hat{y}(t_k), w_j) - \mathbf{b}(w_l, \hat{y}(t), w_j)| &= |\mathbf{b}(w_l, \hat{y}(t_k) - \hat{y}(t), w_j)| \\ &= |-\mathbf{b}(w_l, w_j, \hat{y}(t_k) - \hat{y}(t))| \leq C \|w_l\| \|w_j\|^{1/2} \|w_j\|_W^{1/2} |\hat{y}(t_k) - \hat{y}(t)| \end{aligned}$$

for some constant $C > 0$. Since $\hat{y} \in C([0, T]; V) \hookrightarrow C([0, T]; H)$, we obtain $|\hat{y}(t_k) - \hat{y}(t)| \rightarrow 0$ as $k \rightarrow \infty$. Hence,

$$|\mathbf{b}(w_l, \hat{y}(t_k), w_j) - \mathbf{b}(w_l, \hat{y}(t), w_j)| \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Thus the function $t \mapsto \mathbf{b}(w_l, \hat{y}(t), w_j)$ is continuous. Similarly, the function $\mathbf{b}(\hat{y}(t), w_l, w_j)$ is continuous. Consequently, $F_j(t, \xi)$ is a continuous function of (t, ξ) in $[0, T] \times \mathbb{R}^m$. By [100, Corollary 3.1, p.14], system (4.1.29) has a maximal solution defined on some interval $[0, t_m]$. If $t_m < T$ then $|\xi(t)| = |y_m(t)|$ must tend to $+\infty$ as $t \rightarrow t_m$. However, the a priori estimate we shall prove later shows that this does not happen and therefore $t_m = T$. In fact, taking the scalar product of (4.1.27) with Ay_m yields

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|y_m\|^2 + \nu |Ay_m|^2 + \rho \mathbf{b}(y_m, y_m, Ay_m) + \mathbf{b}(\hat{y}, y_m, Ay_m) + \mathbf{b}(y_m, \hat{y}, Ay_m) \\ = \langle f, Ay_m \rangle. \end{aligned} \quad (4.1.30)$$

Using (4.1.23), we have

$$|\mathbf{b}(y_m, \hat{y}, Ay_m)| \leq C \|y_m\| \|\hat{y}\|_W |Ay_m| \quad (4.1.31)$$

for some constant $C > 0$. Also, using (4.1.24), we get

$$|\mathbf{b}(\hat{y}, y_m, Ay_m)| \leq C \|\hat{y}\| \|y_m\|^{1/2} |Ay_m|^{3/2}$$

and

$$|\mathbf{b}(y_m, y_m, Ay_m)| \leq C \|y_m\| \|y_m\|^{1/2} |Ay_m|^{3/2} = C \|y_m\|^{3/2} |Ay_m|^{3/2} \quad (4.1.32)$$

for some constant $C > 0$. Combining (4.1.30) with (4.1.31)–(4.1.32), we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|y_m\|^2 + \nu |Ay_m|^2 &\leq C \rho \|y_m\|^{3/2} |Ay_m|^{3/2} + C \|y_m\| \|\hat{y}\|_W |Ay_m| + \\ &+ C \|\hat{y}\| \|y_m\|^{1/2} |Ay_m|^{3/2} + |(f, Ay_m)| \\ &\leq \frac{\nu}{8} |Ay_m|^2 + \left(\frac{24}{\nu}\right)^3 \left(\frac{\rho C}{4}\right)^4 \|y_m\|^6 + \frac{\nu}{8} |Ay_m|^2 + \frac{2C^2}{\nu} \|\hat{y}\|_W^2 \|y_m\|^2 \\ &+ \frac{\nu}{8} |Ay_m|^2 + \left(\frac{24}{\nu}\right)^3 \left(\frac{C}{4}\right)^4 \|\hat{y}\|^4 \|y_m\|^2 + \frac{\nu}{8} |Ay_m|^2 + \frac{2}{\nu} |f|^2. \end{aligned}$$

It follows that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|y_m\|^2 + \frac{\nu}{2} |Ay_m|^2 &\leq \left(\frac{24}{\nu}\right)^3 \left(\frac{\rho C}{4}\right)^4 \|y_m\|^6 + \frac{2C^2}{\nu} \|\hat{y}\|_W^2 \|y_m\|^2 \\ &+ \left(\frac{24}{\nu}\right)^3 \left(\frac{C}{4}\right)^4 \|\hat{y}\|^4 \|y_m\|^2 + \frac{2}{\nu} |f|^2. \end{aligned}$$

or equivalently,

$$\begin{aligned}
& \frac{d}{dt} \|y_m\|^2 + \nu |Ay_m|^2 \\
& \leq \frac{108C^4}{\nu^3} \rho^4 \|y_m\|^6 + \left(\frac{4c^2}{\nu} \|\hat{y}\|_W^2 + \frac{108C^4}{\nu^3} \|\hat{y}\|_{C(0,T;V)}^4 \right) \|y_m\|^2 + \frac{4}{\nu} |f|^2 \\
& \leq C_2 \rho^4 \|y_m\|^6 + \left(C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4 \right) \|y_m\|^2 + \frac{4}{\nu} |f|^2, \quad (4.1.33)
\end{aligned}$$

where C_1 and C_2 are defined by (4.1.26).

Claim: For all $t \in [0, T]$ one has

$$\|y_m(t)\|^2 \leq \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right).$$

Indeed, obviously we have

$$\|y_m(0)\|^2 = 0 < \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right).$$

By the continuity of $\|y_m(\cdot)\|$, there exists $t_1 > 0$ such that

$$\|y_m(t)\|^2 < \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)$$

for all $t \in [0, t_1]$. Since $|Ay_m| \geq \lambda_1 \|y_m\|$ and the choice of ρ , we have

$$\begin{aligned}
\nu \|Ay_m\|^2 - C_2 \rho^4 \|y_m\|^6 & \geq \|y_m(t)\|^2 (\nu \lambda_1^2 - C_2 \rho^4 \|y_m(t)\|^4) \\
& \geq \|y_m(t)\|^2 \left(\nu \lambda_1^2 - C_2 \rho^4 \exp \left(2C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + \right. \right. \\
& \quad \left. \left. 2C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{16}{\nu^2} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)^2 \right) \geq 0
\end{aligned}$$

for all $t \in [0, t_1]$. Hence from (4.1.33), we get

$$\frac{d}{dt} \|y_m\|^2 \leq \left(C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4 \right) \|y_m\|^2 + \frac{4}{\nu} \|f\|^2$$

for all $t \in [0, t_1]$. The Gronwall inequality implies that

$$\begin{aligned}
\|y_m(t)\|^2 & \leq \exp \left(\int_0^t (C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4) dt \right) \frac{4}{\nu} \int_0^t |f(t)|^2 dt \\
& \leq \exp \left(\int_0^T (C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4) dt \right) \frac{4}{\nu} \int_0^T |f(t)|^2 dt \\
& < \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)
\end{aligned}$$

for all $t \in [0, t_1]$. Let us denote by $\hat{T}$ the least upper bound of the set of $t_* \in [0, T]$ such that

$$\|y_m(t)\|^2 < \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)$$

for all $t \in [0, t_*]$. By definition of $\hat{T}$, we have

$$\|y_m(t)\|^2 \leq \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)$$

for all $t \in [0, \hat{T}]$. If $\hat{T} = T$, then the claim is proved. Let us assume that $\hat{T} < T$. Then from the choice of ρ , we get from (4.1.33) that

$$\frac{d}{dt} \|y_m\|^2 \leq \left(C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4 \right) \|y_m\|^2 + \frac{4}{\nu} \|f\|^2$$

for all $t \in [0, \hat{T}]$. Using the Gronwall inequality again, we get

$$\begin{aligned} \|y_m(\hat{T})\|^2 &\leq \exp \left(\int_0^{\hat{T}} \left(C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4 \right) dt \right) \frac{4}{\nu} \int_0^{\hat{T}} |f(t)|^2 dt \\ &\leq \exp \left(\int_0^T \left(C_1 \|\hat{y}\|_W^2 + C_2 \|\hat{y}\|_{C(0,T;V)}^4 \right) dt \right) \frac{4}{\nu} \int_0^T |f(t)|^2 dt \\ &< \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right). \end{aligned}$$

By the continuity of y_m , there exists $T_1 > \hat{T}$ such that

$$\|y_m(t)\|^2 < \exp \left(C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)$$

for all $t \in [0, T_1]$. This contradicts to the definition of $\hat{T}$. We obtain the claim.

Let us define

$$M_1(\hat{y}) = C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 + C_2 T \|\hat{y}\|_{C(0,T;V)}^4.$$

By the claim, the sequence $\{y_m\}$ is bounded in $L^\infty(0, T; V)$ and we have

$$\|y_m\|_{L^\infty(0,T;V)}^2 \leq \exp(M_1(\hat{y})) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right). \quad (4.1.34)$$

Integrating two sides of (4.1.33) yields

$$\begin{aligned} \|y_m(T)\|^2 + \nu \int_0^T |Ay_m|^2 dt &\leq C_2 T \rho^4 \exp(3M_1(\hat{y})) \frac{4^3}{\nu^3} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right)^3 \\ &+ T M_1(\hat{y}) \exp(M_1(\hat{y})) \frac{4}{\nu} \left(\|f\|_{L^2(0,T;H)}^2 + 1 \right) + \frac{4}{\nu} \|f\|_{L^2(0,T;H)}^2. \end{aligned} \quad (4.1.35)$$

Let us denote by $M_2(\rho, \hat{y}, f)$ the left-hand side of (4.1.35). Combining (4.1.35) with (4.1.25), we get

$$\frac{\nu}{C} \int_0^T \|y_m\|_W^2 dt \leq \nu \int_0^T |Ay_m|^2 dt \leq M_2(\rho, \hat{y}, f).$$

Hence

$$\|y_m\|_{L^2(0,T;W)}^2 \leq \frac{C}{\nu} M_2(\rho, \hat{y}, f). \quad (4.1.36)$$

Consequently, the sequence y_m is bounded in $L^2(0, T; W)$. From (4.1.27) we get

$$\begin{aligned} \left\| \frac{dy_m}{dt} \right\|^2 &\leq 5 \left(\nu^2 |Ay_m|^2 + \rho^2 |P_m B(y_m, y_m)|^2 + |P_m B(\hat{y}, y_m)|^2 \right. \\ &\quad \left. + |P_m B(y_m, \hat{y})|^2 + |g_m|^2 \right) \\ &\leq 5 \left(\nu^2 |Ay_m|^2 + \rho^2 |(y_m \cdot \nabla) y_m|^2 + |(\hat{y} \cdot \nabla) y_m|^2 + |(y_m \cdot \nabla) \hat{y}|^2 + |f|^2 \right) \\ &\leq 5 \left(\nu^2 |Ay_m|^2 + \rho^2 \|y_m\|^2 \|y_m\|_W^2 + \|\hat{y}\|^2 \|y_m\|_W^2 + \|y_m\|^2 \|\hat{y}\|_W^2 + |f|^2 \right) \\ &\leq 5 \left(\nu^2 |Ay_m|^2 + \rho^2 \|y_m\|_{L^\infty(0,T;V)}^2 \|y_m\|_W^2 + \|\hat{y}\|_{C(0,T;V)}^2 \|y_m\|_W^2 \right. \\ &\quad \left. + \|y_m\|_{L^\infty(0,T;V)}^2 \|\hat{y}\|_W^2 + |f|^2 \right). \end{aligned}$$

Combining this with (4.1.35) and (4.1.36) yields

$$\begin{aligned} \int_0^T |\dot{y}_m(t)|^2 dt &\leq 5 \left(\nu M_2(\rho, \hat{y}, f) + \rho^2 \|y_m\|_{L^\infty(0,T;V)}^2 \frac{C}{\nu} M_2(\rho, \hat{y}, f) \right. \\ &\quad \left. + \|\hat{y}\|_{C(0,T;V)}^2 \frac{C}{\nu} M_2(\rho, \hat{y}, f) + \|y_m\|_{L^\infty(0,T;V)}^2 \|\hat{y}\|_{L^2(0,T;W)}^2 + \|f\|_{L^2(0,T;H)}^2 \right). \end{aligned} \quad (4.1.37)$$

Moreover, adding (4.1.36) and (4.1.37) and using (4.1.34), we obtain

$$\begin{aligned}
& \|y_m\|_{L^2(0,T;W)}^2 + \|\dot{y}_m\|_{L^2(0,T;H)}^2 \\
& \leq 5 \left[\nu M_2(\rho, \hat{y}, f) + \rho^2 \|y_m\|_{L^\infty(0,T;V)}^2 \frac{C}{\nu} M_2(\rho, \hat{y}, f) \right. \\
& \quad \left. + \|\hat{y}\|_{C^2(0,T;V)}^2 \frac{C}{\nu} M_2(\rho, \hat{y}, f) + \|y_m\|_{L^\infty(0,T;V)}^2 \|\hat{y}\|_{L^2(0,T;W)}^2 + \|f\|_{L^2(0,T;H)}^2 \right] \\
& \quad + \frac{C}{\nu} M_2(\rho, \hat{y}, f) \\
& = M_2(\rho, \hat{y}, f) \left[5 + \rho^2 \frac{5C}{\nu} \|y_m\|_{L^\infty(0,T;V)}^2 + \frac{5C}{\nu} \|\hat{y}\|_{C^2(0,T;V)}^2 + \frac{C}{\nu} \right] \\
& \quad + 5 \|\hat{y}\|_{L^2(0,T;W)}^2 \|y_m\|_{L^\infty(0,T;V)}^2 + 5 \|f\|_{L^2(0,T;H)}^2 \\
& \leq M_2(\rho, \hat{y}, f) \left[\frac{20\rho^2 C}{\nu^2} \exp(M_1(\hat{y})) (\|f\|_{L^2(0,T;H)}^2 + 1) \right. \\
& \quad \left. + \frac{5C}{\nu} \|\hat{y}\|_{C^2(0,T;V)}^2 + \frac{C}{\nu} + 5\nu \right] \\
& \quad + \frac{20}{\nu} \|\hat{y}\|_{L^2(0,T;W)}^2 \exp(M_1(\hat{y})) (\|f\|_{L^2(0,T;H)}^2 + 1) + 5 \|f\|_{L^2(0,T;H)}^2.
\end{aligned} \tag{4.1.38}$$

Thus

$$\text{sequence } y_m \text{ is bounded in } W^{1,2}(0, T; W, H). \tag{4.1.39}$$

From (4.1.34) and (4.1.39), by passing subsequences, we can find an element $y \in L^\infty(0, T; V) \cap W^{1,2}(0, T; W, H)$ such that

$$\begin{aligned}
y_m & \rightharpoonup y \quad \text{in } W^{1,2}(0, T; W, H) \text{ weakly,} \\
y_m & \xrightarrow{*} y \quad \text{in } L^\infty(0, T; V) \text{ weakly star.}
\end{aligned}$$

Since $W \hookrightarrow V$ is compact, we have compactness of the embedding $W^{1,2}(0, T; W, H) \hookrightarrow L^2(0, T; V)$ by the Aubin theorem. Hence, we get

$$y_m \rightarrow y \quad \text{in } L^2(0, T; V) \text{ strongly.}$$

Note that since y_m converges weakly to y in $W^{1,2}(0, T; W, H)$, we have

$$\frac{dy_m}{dt} \rightharpoonup \frac{dy}{dt} \quad \text{in } L^2(0, T; H) \text{ weakly.}$$

By using the procedure as in the proof of [216, Theorem 3.1, p. 282] and passing the limit when $m \rightarrow \infty$, we see that $y \in W^{1,2}(0, T; W, H)$ is a

strong solution of Eq. (4.1.21). Also, by letting $m \rightarrow \infty$ in estimation (4.1.38), we obtain estimation (4.1.22).

Uniqueness For this, we use the Gronwall inequality and standard arguments as in [59]. We obtain the desired conclusion. $\square$

4.1.3 Pontryagin's Maximum Principle

In this subsection, we shall establish the Pontryagin maximum principle for problem (4.1.1)–(4.1.6). Let us denote by $\mathcal{A}$ the feasible set of problem (4.1.1)–(4.1.6), that is, $\mathcal{A}$ consists of $(y, u) \in W^{1,2}(0, T; W, H) \times L^2(0, T; L^2(\Omega)^3)$ satisfying constraints (4.1.2)–(4.1.6). We denote by $\langle \cdot, \cdot \rangle$ and $|\cdot|_e$ the scalar product and Euclidean norm in $\mathbb{R}^3$, respectively. Let us define

$$\mathcal{U} = \{u \in L^2(Q, \mathbb{R}^3) \mid u(x, t) \in M \text{ a.e.}\} \quad (4.1.40)$$

and impose the following assumptions on function L .

(A1) L is a Carathéodory function, and $L_y(x, t, \cdot, u(x, t))$ is differentiable around $\hat{y}(x, t)$ for a.e. $(x, t) \in Q$ and $u \in \mathcal{U}$.

(A2) The mapping $|L_y(\cdot, \cdot, \hat{y}(\cdot, \cdot), \hat{u}(\cdot, \cdot))|_e$ belongs to $L^2(Q; \mathbb{R})$, and there exists a function $l \in L^\infty(Q; \mathbb{R})$ such that

$$|L_y(x, t, y_1, u_1) - L_y(x, t, y_2, u_2)|_e \leq l(x, t)(|y_1 - y_2|_e + |u_1 - u_2|_e) \quad (4.1.41)$$

for a.e. $(x, t) \in Q$ and for all $(y_1, u_1), (y_2, u_2)$ in a neighborhood of $(\hat{y}(x, t), \hat{u}(x, t))$.

Definition 4.1.8. A couple $(\hat{y}, \hat{u}) \in \mathcal{A}$ is said to be an optimal solution of (4.1.1)–(4.1.6) if there exists $\epsilon > 0$ such that the following implication holds:

$$(y, u) \in \mathcal{A}, \quad \|y - \hat{y}\|_{L^2(0, T; V)} + \|u - \hat{u}\|_{L^2(0, T; L^2(\Omega)^3)} < \epsilon \Rightarrow J(y, u) \geq J(\hat{y}, \hat{u}).$$

Remark 4.1.9. In Definition 4.1.8, the requirement

$$(\hat{y}, \hat{u}) \in W^{1,2}(0, T; W, H) \times L^2(0, T; L^2(\Omega)^3)$$

ensures that $\hat{y}$ is a global strong solution to the Navier–Stokes equation corresponding to $\hat{u}$. It is known that under a smallness condition on $\hat{u}$, such a solution $\hat{y}$ always exists (see Theorem 9.3 in [59]). However, proving the existence of a global strong/smooth solution to the Navier–Stokes equation without a smallness condition on $\hat{u}$ remains a challenge for mathematicians.

The following lemmas play an important role in establishing Pontryagin's principles.

Lemma 4.1.10 (Cf. [189, Lemma 4.2]). *Let $v \in \mathcal{U}$ and $(\hat{y}, \hat{u}) \in \mathcal{A}$. Then for every $0 < \rho < 1$, there exists a sequence of Lebesgue measurable sets $\{E_\rho^k\}$ of Q such that*

$$|E_\rho^k| = \rho|Q|, \quad (4.1.42)$$

$$\frac{1}{\rho} \chi_{E_\rho^k} \rightharpoonup 1 \quad \text{in } L^\infty(Q; \mathbb{R}) \text{ weakly star as } k \rightarrow \infty, \quad (4.1.43)$$

$$\begin{aligned} & \int_{E_\rho^k} (L(x, t, \hat{y}, v) - L(x, t, \hat{y}, \hat{u})) \, dx dt \\ &= \rho \int_Q (L(x, t, \hat{y}, v) - L(x, t, \hat{y}, \hat{u})) \, dx dt, \end{aligned} \quad (4.1.44)$$

where $|\cdot|$ denotes the Lebesgue measure, and $\chi_{E_\rho^k}$ is the characteristic function of E_ρ^k .

Lemma 4.1.11. *Suppose $(\hat{y}, \hat{u})$ is an admissible pair and $v \in \mathcal{U}$. Then there exists $\rho_0 \in (0, 1)$ such that for each $\rho \in (0, \rho_0)$, there exists a measurable set $E_\rho \subset Q$ with $|E_\rho| = \rho|Q|$ and if*

$$u_\rho(x, t) := (\hat{u} + \chi_{E_\rho}(v - \hat{u}))(x, t) = \begin{cases} v(x, t), & (x, t) \in E_\rho, \\ \hat{u}(x, t), & (x, t) \in Q \setminus E_\rho \end{cases} \quad (4.1.45)$$

and let y_ρ be the state corresponding to u_ρ , then $(y_\rho, u_\rho) \in \mathcal{A}$ and the following formulae hold:

$$y_\rho = \hat{y} + \rho z + r_\rho, \quad \lim_{\rho \rightarrow 0^+} \frac{\|r_\rho\|_{L^2(0, T; V)}}{\rho} = 0, \quad (4.1.46)$$

$$J(y_\rho, u_\rho) = J(\hat{y}, \hat{u}) + \rho \Delta J + o(\rho), \quad (4.1.47)$$

where z is a solution of the equation

$$\frac{dy}{dt} + \nu A y + B(\hat{y}, y) + B(y, \hat{y}) = E(v - \hat{u}), \quad y(0) = 0.$$

and

$$\begin{aligned} \Delta J = & \int_Q \left[L_y(x, t, \hat{y}(x, t), \hat{u}(x, t)) z(x, t) \right. \\ & \left. + L(x, t, \hat{y}(x, t), v(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) \right] dx dt. \end{aligned}$$

Proof. Let v be given in $\mathcal{U}$. We put

$$3\varepsilon = \frac{\nu^{3/2}\lambda_1}{4C_2^{1/2}} \exp \left(-C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 - C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right),$$

where C_1, C_2 are constants as in Lemma 4.1.7. Since $|E(v - \hat{u})(\cdot, \cdot)|_e^2 \in L^1(Q; \mathbb{R})$, there exists $\delta > 0$ such that for all measurable subsets $\mathcal{O}$ of Q with $|\mathcal{O}| < \delta$, one has

$$\int_{\mathcal{O}} |E(v - \hat{u})(x, t)|_e^2 dx dt < \varepsilon.$$

Let us choose

$$\rho_0 = \min \left\{ \frac{\delta}{|Q|}, \sqrt{\varepsilon}, 1 \right\}.$$

By Lemma 4.1.10, for each $\rho \in (0, \rho_0)$, there exists a sequence of measurable sets E_ρ^k of Q such that (4.1.42)–(4.1.44) are valid. By the choice of ρ_0 , we have $|E_\rho^k| = \rho|Q| < \rho_0|Q| < \delta$. Hence,

$$\begin{aligned} \rho^2 \left\| \frac{1}{\rho} \chi_{E_\rho^k} E(v - \hat{u}) \right\|_{L^2(0,T;H)}^2 &= \int_0^T \int_\Omega \chi_{E_\rho^k} |E(v - \hat{u})|_e^2 dx dt \\ &= \int_{E_\rho^k} |E(v - \hat{u})|_e^2 dx dt < \varepsilon. \end{aligned} \quad (4.1.48)$$

This implies that

$$\rho^2 \left(\left\| \frac{1}{\rho} \chi_{E_\rho^k} E(v - \hat{u}) \right\|_{L^2(0,T;H)}^2 + 1 \right) < 2\varepsilon. \quad (4.1.49)$$

Hence,

$$\begin{aligned} \rho^2 \left(\left\| \frac{1}{\rho} \chi_{E_\rho^k} E(v - \hat{u}) \right\|_{L^2(0,T;H)}^2 + 1 \right) \\ < \frac{\nu^{3/2}\lambda_1}{4C_2^{1/2}} \exp \left(-C_1 \|\hat{y}\|_{L^2(0,T;W)}^2 - C_2 T \|\hat{y}\|_{C(0,T;V)}^4 \right). \end{aligned} \quad (4.1.50)$$

Let us define $u_\rho^k = \hat{u} + \chi_{E_\rho^k}(v - \hat{u})$, and let y_ρ^k be the state corresponding to the control u_ρ^k . By (4.1.43), we have

$$\|u_\rho^k - \hat{u}\|_{L^2(Q; \mathbb{R})}^2 = \rho \int_Q \frac{1}{\rho} \chi_{E_\rho^k} |(v - \hat{u})(x, t)|_e^2 dx dt \rightarrow \rho \int_Q |(v - \hat{u})(x, t)|_e^2 dx dt$$

as $k \rightarrow \infty$. Hence for $\rho > 0$ small enough and k big enough, u_ρ^k belongs to a neighborhood of $\hat{u}$ in $L^2(Q, \mathbb{R})$. Consequently, $y_\rho^k \in W^{1,2}(0, T; W, H)$. Also, since $|E(v - \hat{u})(\cdot, \cdot)|_e \in L^2(Q, \mathbb{R})$ and by (4.1.43), we have

$$\frac{1}{\rho} \chi_{E_\rho^k} |E(v - \hat{u})(x, t)|_e \rightharpoonup |E(v - \hat{u})(x, t)|_e \quad \text{in } L^2(Q, \mathbb{R}) \text{ as } k \rightarrow \infty.$$

Hence, there is a constant $M_\rho > 0$ such that for all $k \geq 1$ one has

$$\left\| \frac{1}{\rho} \chi_{E_\rho^k} E(v - \hat{u}) \right\|_{L^2(0, T; H)}^2 = \int_Q \left(\frac{1}{\rho} \chi_{E_\rho^k} \right)^2 |E(v - \hat{u})(x, t)|_e^2 dx dt \leq M_\rho^2. \quad (4.1.51)$$

Put $z_\rho^k = \frac{1}{\rho}(y_\rho^k - \hat{y})$ and $f_k = \frac{1}{\rho} \chi_{E_\rho^k} E(v - \hat{u})$. Then $z_\rho^k \in W^{1,2}(0, T; W, H)$ and satisfies the equation

$$\dot{z}_\rho^k + \nu A z_\rho^k + \rho B(z_\rho^k, z_\rho^k) + B(\hat{y}, z_\rho^k) + B(z_\rho^k, \hat{y}) = f_k, \quad z_\rho^k(0) = 0. \quad (4.1.52)$$

From (4.1.50) and Lemma 4.1.7, there exist constants $M_1(\hat{y})$ and $M_2(\rho, \hat{y}, f_k)$ such that

$$\begin{aligned} & \|z_\rho^k\|_{L^2(0, T; W)}^2 + \|\dot{z}_\rho^k\|_{L^2(0, T; H)}^2 \leq \\ & M_2(\rho, \hat{y}, f_k) \left(\frac{20\rho^2 C}{\nu^2} \exp(M_1(\hat{y})) (\|f_k\|_{L^2(0, T; H)}^2 + 1) \right. \\ & \quad \left. + \frac{5C}{\nu} \|\hat{y}\|_{C^2(0, T; V)}^2 + \frac{C}{\nu} + 5\nu \right) \\ & \quad + \frac{20}{\nu} \|\hat{y}\|_{L^2(0, T; W)}^2 \exp(M_1(\hat{y})) (\|f_k\|_{L^2(0, T; H)}^2 + 1) + 5\|f_k\|_{L^2(0, T; H)}^2, \end{aligned} \quad (4.1.53)$$

where

$$M_1(\hat{y}) = C_1 \|\hat{y}\|_{L^2(0, T; W)}^2 + C_2 T \|\hat{y}\|_{C(0, T; V)}^4, \quad (4.1.54)$$

$$\begin{aligned} M_2(\rho, \hat{y}, f_k) &= C_2 T \rho^4 \exp(3M_1(\hat{y})) \frac{4^3}{\nu^3} (\|f_k\|_{L^2(0, T; H)}^2 + 1)^3 \\ & \quad + T M_1(\hat{y}) \exp(M_1(\hat{y})) \frac{4}{\nu} (\|f_k\|_{L^2(0, T; H)}^2 + 1) + \frac{4}{\nu} \|f_k\|_{L^2(0, T; H)}^2. \end{aligned} \quad (4.1.55)$$

Combining (4.1.51) with (4.1.53)–(4.1.55), we conclude that z_ρ^k is bounded in $W^{1,2}(0, T; W, H)$. Without loss of generality, we can assume that $z_\rho^k \rightharpoonup z_\rho$ weakly in $W^{1,2}(0, T; W, H)$. Since $W \hookrightarrow V$ is compact, the Aubin theorem (see [216, Theorem 2.1, p. 271]) implies that the embedding

$$W^{1,2}(0, T; W, H) \hookrightarrow L^2(0, T; V)$$

is compact. Consequently, since $z_\rho^k \rightarrow z_\rho$ strongly in $L^2(0, T; V)$, it follows that there exists $k(\rho) > 0$ such that

$$\|z_\rho^{k(\rho)} - z_\rho\|_{L^2(0, T; V)} < \rho. \quad (4.1.56)$$

Passing the limit as $k \rightarrow \infty$ and using similar arguments as in the proof of Lemma 4.1.7, we obtain from (4.1.52) that

$$\dot{z}_\rho + \nu A z_\rho + \rho B(z_\rho, z_\rho) + B(\hat{y}, z_\rho) + B(z_\rho, \hat{y}) = E(v - \hat{u}), \quad z_\rho(0) = 0. \quad (4.1.57)$$

Also, from (4.1.49) we get

$$\begin{aligned} \rho^2 (\|E(v - \hat{u})\|_{L^2(0, T; H)}^2 + 1) &\leq \liminf_{k \rightarrow \infty} \rho^2 \left(\left\| \frac{1}{\rho} \chi_{E_\rho^k} E(v - \hat{u}) \right\|_{L^2(0, T; H)}^2 + 1 \right) \\ &< 3\varepsilon. \end{aligned}$$

This implies that

$$\rho^2 (\|E(v - \hat{u})\|_{L^2(0, T; H)}^2 + 1) < \frac{\nu^{3/2} \lambda_1}{4C_2^{1/2}} \exp(-C_1 \|\hat{y}\|_{L^2(0, T; W)}^2 - C_2 T \|\hat{y}\|_{C(0, T; V)}^4).$$

Hence, Lemma 4.1.7 is applicable to Eq. (4.1.57). According to Lemma 4.1.7, there exist constants $M_1(\hat{y})$ and $M_2(\rho, \hat{y}, f)$ with $f = E(v - \hat{u})$ such that

$$\begin{aligned} &\|z_\rho\|_{L^2(0, T; W)}^2 + \|\dot{z}_\rho\|_{L^2(0, T; H)}^2 \leq \\ &M_2(\rho, \hat{y}, f) \left(\frac{20\rho^2 C}{\nu^2} \exp(M_1(\hat{y})) (\|f\|_{L^2(0, T; H)}^2 + 1) + \frac{5C}{\nu} \|\hat{y}\|_{C(0, T; V)}^2 + \frac{C}{\nu} + 5\nu \right) \\ &+ \frac{20}{\nu} \|\hat{y}\|_{L^2(0, T; W)}^2 \exp(M_1(\hat{y})) (\|f\|_{L^2(0, T; H)}^2 + 1) + 5\|f\|_{L^2(0, T; H)}^2, \end{aligned}$$

where $M_1(\hat{y})$ and $M_2(\rho, \hat{y}, f)$ are given by (4.1.54) and (4.1.55) with f_k replaced by f . It follows that $\{z_\rho\}$ is bounded in $W^{1,2}(0, T; W, H)$. Therefore, we can assume that $z_\rho \rightharpoonup z$ in $W^{1,2}(0, T; W, H)$. Since the embedding $W^{1,2}(0, T; W, H) \hookrightarrow L^2(0, T; V)$ is compact, it follows that $z_\rho \rightarrow z$ strongly in $L^2(0, T; V)$. Letting $\rho \rightarrow 0^+$ and using similar arguments as in the proof of Lemma 4.1.7, we obtain from (4.1.57) that

$$\dot{z} + \nu A z + B(\hat{y}, z) + B(z, \hat{y}) = E(v - \hat{u}), \quad z(0) = 0. \quad (4.1.58)$$

Let us now define

$$E_\rho = E_\rho^k(\rho), \quad u_\rho = u_\rho^{k(\rho)} = \hat{u} + \chi_{E_\rho}(v - \hat{u}), \quad y_\rho = y_\rho^{k(\rho)}.$$

It is clear that $(y_\rho, u_\rho) \in \mathcal{A}$. From (4.1.56), we have

$$\begin{aligned} \|z_\rho^{k(\rho)} - z\|_{L^2(0, T; V)} &\leq \|z_\rho^{k(\rho)} - z_\rho\|_{L^2(0, T; V)} + \|z_\rho - z\|_{L^2(0, T; V)} \\ &\leq \rho + \|z_\rho - z\|_{L^2(0, T; V)} \rightarrow 0 \quad \text{as } \rho \rightarrow 0. \end{aligned}$$

Putting $r_\rho = \rho(z_\rho^{k(\rho)} - z)$, we get $y_\rho = \hat{y} + \rho z + r_\rho$, which is formula (4.1.46). It remains to show that (4.1.47) is valid. Define

$$L_\rho(x, t) = \int_0^1 L_y(x, t, \hat{y}(x, t) + \theta(y_\rho(x, t) - \hat{y}(x, t)), u_\rho(x, t)) d\theta.$$

Then from Lemma 4.1.10, we have

$$\begin{aligned} \frac{J(y_\rho, u_\rho) - J(\hat{y}, \hat{u})}{\rho} &= \\ &= \int_Q \frac{L(x, t, y_\rho(x, t), u_\rho(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t))}{\rho} dx dt \\ &= \int_Q \frac{L(x, t, y_\rho(x, t), u_\rho(x, t)) - L(x, t, \hat{y}(x, t), u_\rho(x, t))}{\rho} dx dt \\ &+ \int_Q \frac{L(x, t, \hat{y}(x, t), u_\rho(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t))}{\rho} dx dt \\ &= \int_Q L_\rho(x, t) z_\rho(x, t) dx dt \\ &+ \int_Q \frac{1}{\rho} \chi_{E_\rho}(x, t) [L(x, t, \hat{y}(x, t), v(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t))] dx dt \\ &= \int_Q L_\rho(x, t) z_\rho(x, t) dx dt \\ &+ \int_Q [L(x, t, \hat{y}(x, t), v(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t))] dx dt. \end{aligned}$$

From (4.1.41) and the dominated convergence theorem, we have

$$\int_Q L_\rho(x, t) z_\rho(x, t) dx dt \rightarrow \int_Q L_y(x, t, \hat{y}(x, t), \hat{u}(x, t)) z(x, t) dx dt.$$

Consequently,

$$\begin{aligned} \frac{J(y_\rho, u_\rho) - J(\hat{y}, \hat{u})}{\rho} &\rightarrow \int_Q L_y(x, t, \hat{y}(x, t), \hat{u}(x, t)) z(x, t) dx dt + \\ &+ \int_Q [L(x, t, \hat{y}(x, t), v(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t))] dx dt, \end{aligned}$$

as $\rho \rightarrow 0$. Hence (4.1.47) is valid. $\square$

We are now ready to state and prove the Pontryagin principle for optimal solution to problem (4.1.1)-(4.1.6).

Theorem 4.1.12. *Suppose that assumptions (A1) and (A2) are fulfilled and $(\hat{y}, \hat{u})$ is an optimal solution of problem (4.1.1)-(4.1.6). Then there*

exists a vector function $\vartheta \in W^{1,2}(0, T; H, W^*)$ such that the following conditions are satisfied:

(a₁) (the adjoint equation)

$$-\dot{\vartheta} + \nu A^* \vartheta + B(\hat{y}, \cdot)^* \vartheta + B(\cdot, \hat{y})^* \vartheta = -L_y(\cdot, \cdot, \hat{y}, \hat{u}), \quad \vartheta(T) = 0.$$

(a₂) (the maximum principle)

$$\begin{aligned} & \langle (E\hat{u})(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) \\ &= \max_{\omega \in \mathcal{M}} \langle (E\omega)(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), \omega) \end{aligned} \quad (4.1.59)$$

for a.e. $(x, t) \in Q$.

Proof. Fix any $v \in \mathcal{U}$. By Lemma 4.1.11, there exists $\rho_0 \in (0, 1)$ such that for each $\rho \in (0, \rho_0)$, there exists a measurable set $E_\rho \subset Q$ with $|E_\rho| = \rho|Q|$, which has a property that if

$$u_\rho(x, t) := \hat{u} + \chi_{E_\rho}(v - \hat{u})$$

and y_ρ is the corresponding state to u_ρ , then (4.1.46) and (4.1.47) hold. Since $(y_\rho, u_\rho) \in \mathcal{A}$, $(y_\rho, u_\rho) \rightarrow (\hat{y}, \hat{u})$, and $(\hat{y}, \hat{u})$ is an optimal solution, we have

$$J(y_\rho, u_\rho) \geq J(\hat{y}, \hat{u}).$$

It follows that

$$0 \leq \frac{1}{\rho}(J(y_\rho, u_\rho) - J(\hat{y}, \hat{u})) \rightarrow \Delta J.$$

Hence,

$$\begin{aligned} 0 \leq \int_Q & \left[L_y(x, t, \hat{y}(x, t), \hat{u}(x, t))z(x, t) + \right. \\ & \left. L(x, t, \hat{y}(x, t), v(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) \right] dx dt \end{aligned} \quad (4.1.60)$$

for all $v \in \mathcal{U}$, where z is a solution of the equation

$$\dot{z} + \nu Az + B(\hat{y}, z) + B(z, \hat{y}) = E(v - \hat{u}), \quad z(0) = 0. \quad (4.1.61)$$

Since $L_y(\cdot, \cdot, \hat{y}, \hat{u}) \in L^2(0, T; L^2(\Omega)^3) \hookrightarrow L^2(0, T; W^*)$, Lemma 4.1.6 implies that the equation

$$-\dot{\vartheta} + \nu A^* \vartheta + B(\hat{y}, \cdot)^* \vartheta + B(\cdot, \hat{y})^* \vartheta = -L_y(\cdot, \cdot, \hat{y}, \hat{u}), \quad \vartheta(T) = 0 \quad (4.1.62)$$

has a unique solution $\vartheta \in W^{1,2}(0, T; H, W^*)$, in which formula (4.1.15) is valid. Hence, assertion (a₁) is obtained. Taking the scalar product

of (4.1.61) with ϑ , of (4.1.62) with z , integrating, and using (4.1.15), we get

$$\int_0^T (\dot{z}, \vartheta) dt + \int_0^T (\nu Az + B(\hat{y}, z) + B(z, \hat{y}), \vartheta) dt = \int_0^T (E(v - \hat{u}), \vartheta) dt$$

and

$$\int_0^T (\vartheta, \dot{z}) dt + \int_0^T (\vartheta, \nu Az + B(\hat{y}, z) + B(z, \hat{y})) dt = - \int_0^T (L_y(t, \hat{y}, \hat{u}), z) dt.$$

It follows that $\int_0^T (E(v - \hat{u}), \vartheta) dt = - \int_0^T (L_y(t, \hat{y}, \hat{u}), z) dt$. Substituting this into (4.1.60), we get

$$0 \leq \int_Q \left[- \langle E(v - \hat{u})(x, t), \vartheta(x, t) \rangle + L(x, t, \hat{y}(x, t), v(x, t)) - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) \right] dx dt,$$

or equivalently,

$$0 \leq \int_Q \left[\langle (E\hat{u})(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) - (\langle (Ev)(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), v(x, t))) \right] dx dt, \quad \forall v \in U. \quad (4.1.63)$$

Since $\mathcal{M}$ is closed in $\mathbb{R}^3$, we see that $\mathcal{M}$ is a separable metric subspace (see [70, Proposition 3.10.9]). Hence, there exists a countable dense set $\mathcal{M}_0 = \{v_i, i \geq 1\} \subset \mathcal{M}$. For each $v_i \in \mathcal{M}_0$, we define

$$H_{v_i}(x, t) = \langle (E\hat{u})(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) - (\langle (Ev_i)(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), v_i)).$$

By the Lebesgue differentiation theorem, for each i , there exists a subset $\mathcal{O}_i \subset Q$ with $|\mathcal{O}_i| = |Q|$ such that for all $z_0 = (x_0, t_0) \in \mathcal{O}_i$, we have

$$\lim_{r \rightarrow 0} \frac{1}{|B(z_0, r)|} \int_{B(z_0, r)} H_{v_i}(x, t) dx dt = H_{v_i}(z_0).$$

Define

$$v_r(t) = \begin{cases} v_i, & \text{if } (x, t) \in B(z_0, r), \\ \hat{u}(x, t), & \text{otherwise.} \end{cases}$$

Then $v_r \in \mathcal{M}$. Substituting v_r into (4.1.63), we get

$$0 \leq \int_{B(z_0, r)} H_{v_i}(x, t) dx dt.$$

Hence

$$0 \leq \frac{1}{|B(z_0, r)|} \int_{B(z_0, r)} H_{v_i}(x, t) dx dt.$$

Letting $r \rightarrow 0^+$, we obtain $0 \leq H_{v_i}(x_0, t_0)$. We now put $\mathcal{O} = \bigcap_{i \geq 1} \mathcal{O}_i$. Then $|\mathcal{O}| = |Q|$. Fix a $(x_0, t_0) \in \mathcal{O}$ and take any $\omega \in \mathcal{M}$. Then there exists a subsequence $v_{i_n} \rightarrow \omega$. From the above, we have $0 \leq H_{v_{i_n}}(t_0)$. Letting $n \rightarrow \infty$, we obtain $0 \leq H_\omega(x_0, t_0)$. Thus, we have shown that for a.e. $(x, t) \in Q$,

$$\begin{aligned} 0 \leq & \langle (E\hat{u})(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), \hat{u}(x, t)) \\ & - (\langle (E\omega)(x, t), \vartheta(x, t) \rangle - L(x, t, \hat{y}(x, t), \omega)), \end{aligned}$$

for all $\omega \in \mathcal{M}$. Therefore, assertion (a₂) is established. The proof of the theorem is complete. $\square$

4.2 Regularity of multipliers and second-order optimality conditions for semilinear parabolic optimal control problems

4.2.1 Introduction

Let Ω be a bounded domain in $\mathbb{R}^N$ with $N = 2$ or 3 and its boundary $\Gamma = \partial\Omega$ is regular enough. Let $D = H^2(\Omega) \cap H_0^1(\Omega)$ and $H = L^2(\Omega)$. We consider the problem of finding control function $u \in L^\infty(Q)$ and the corresponding state $y \in L^\infty(Q) \cap W_2^{1,1}(0, T; D, H)$ which solve

$$J(y, u) := \int_Q L(x, t, y(x, t), u(x, t)) dx dt \rightarrow \inf \quad (4.2.1)$$

s.t.

$$\frac{\partial y}{\partial t} + Ay + f(y) = u \quad \text{in } Q = \Omega \times (0, T) \quad (4.2.2)$$

$$y(x, t) = 0 \quad \text{on } \Sigma = \Gamma \times [0, T], \quad y(x, 0) = y_0(x) \quad \text{on } \Omega \quad (4.2.3)$$

$$g(x, t, y(x, t)) + \epsilon u(x, t) \leq 0 \quad \text{a.a. } (x, t) \in Q \quad (4.2.4)$$

$$a \leq u(x, t) \leq b \quad \text{a.a. } (x, t) \in Q \quad (4.2.5)$$

where $y_0 \in H_0^1(\Omega) \cap L^\infty(\Omega)$, $\epsilon > 0$, $a, b \in \mathbb{R}$ with $a < b$, A denotes a second-order elliptic operator of the form

$$Ay = - \sum_{i,j=1}^N D_j (a_{ij}(x) D_i y),$$

$L : Q \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \Omega \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ are of class C^2 . In problem (4.2.1)-(4.2.5), constraint (4.2.4) may be seen as a regularization of the constraint $g(x, t, y(x, t)) \leq 0$ which never satisfies regular conditions. Throughout this section, we denote by Φ_3 the feasible set to problem (4.2.1)-(4.2.5).

The theory of optimal control governed by parabolic equations has been interested by many mathematicians so far. The mainly research topics in this area consist of the existence of optimal solutions, optimality conditions and numerical methods of computing optimal solutions. In order to compute optimal solutions, we need optimality conditions of KKT type. Based on first-order optimality conditions and regularity of multipliers we can find approximate solutions while the second-order sufficient conditions help us to give error estimates between approximate solutions and exact solutions. Therefore, the optimality conditions, regularity of multipliers and optimal solutions play an important role in numerical methods such as the finite element method (FEM).

In the theory of optimal control problems governed by partial differential equations, we often require that the control variables belong to L^∞ . This requirement make the objective functional and constraint mappings easily to be differentiable with respect to u . However, in this situation, the Lagrange multipliers corresponding to constraints are finitely additive measures rather than L^p - functions. This motivates us to study the regularity of Lagrange multipliers, that is to find conditions under which Lagrange multipliers are functions in L^p -spaces.

Recently, there have been many papers dealing with optimality conditions and regularity of multipliers for parabolic optimal control problems. For this we refer the reader to [5], [6], [7], [39], [163], [168], [191], [203], [204] and [218] and many references given therein. Among these papers, Arada and Raymond [5]- [7] and Rösch and Tröltzsch [204] focused on first-order optimality conditions and regularity of multipliers as well as optimal solutions for semilinear parabolic optimal control problems with mixed pointwise constraints. As usual, in order to derive optimality conditions of KKT-type, the authors need a condition of Mangasarian-Fromowitz type. Meanwhile, to obtain the regularity of multipliers, they required that the so-called separation conditions are satisfied. The separation conditions make sure that the Lagrange multipliers corresponding to mixed constraints can be represented by densities in L^1 . However, it is not easy to verify the Mangasarian-Fromowitz condition for the class of problems.

In this subsection we continue develop results in [5] and [204] to derive first-and second order optimality conditions of KKT-type and to study regularity of multipliers and optimal solutions to problem (4.2.1)-(4.2.5). We will give simple conditions under which the Mangasarian-Fromowitz condition and the separation conditions are satisfied. This condition is a form of *generalized separation condition* (see assumption (H4)), which is a key assumption for the existence of regular Lagrange multipliers and second-order optimality conditions of KKT-type.

In order to establish such results, we first embed problem (4.2.1)-(4.2.5) into a specific mathematical programming problem in Banach spaces. We then show that under the generalized separation condition, the Robinson constraint qualification is satisfied. From this we can derive first and second-order optimality conditions of KKT-type. Also, from the generalized separation condition, some techniques in [204] and results of Yosida-Hewitt in [234], we obtain the regularity of Lagrange multipliers.

4.2.2 A specific mathematical programming problem

Let E, W and Z be Banach spaces with the dual spaces E^*, W^* and Z^* . Let Q be a nonempty and closed subset in Z and $\bar{z} \in Q$. Let $\psi : Z \rightarrow \mathbb{R}$; $F : Z \rightarrow E$ and $G : Z \rightarrow W$ be given mappings and K be a nonempty closed convex subset in W . We now consider the optimization problem

$$\psi(z) \rightarrow \inf \quad (4.2.6)$$

s.t.

$$F(z) = 0, \quad (4.2.7)$$

$$G(z) \in K. \quad (4.2.8)$$

By defining $K_1 = \{0\} \times K$ and $G_1 = (F, G)$, we can formulate the problem (4.2.6)–(4.2.8) in the following form

$$\begin{cases} \psi(z) \rightarrow \inf \\ \text{s.t.} \\ G_1(z) \in K_1. \end{cases}$$

Given a feasible point z_0 of (4.2.6)–(4.2.8). Then z_0 is said to satisfy the Robinson constraint qualification if

$$0 \in \text{int}\{G_1(z_0) + DG_1(z_0)Z - K_1\}. \quad (4.2.9)$$

According to [26, condition (2.190), page 71], if $\text{int}(K) \neq \emptyset$ then (4.2.9) is equivalent to

$$\begin{cases} (i) & DF(z_0) \text{ is surjective} \\ (ii) & \exists \tilde{z} \in \text{Ker} DF(z_0) \text{ such that } G(z_0) + DG(z_0)\tilde{z} \in \text{int}(K). \end{cases} \quad (4.2.10)$$

Note that the second condition in (4.2.10) is of Mangasarian–Fromowitz type. It was used in [5] and [218, (3.6), p. 20] to deal with first-order optimality conditions.

Let us impose the following assumptions.

- (A_1^k) The mappings ψ, F and G are of class C^k around z_0 with $k = 1, 2$;
- (A_2) K is a closed convex set with nonempty interior;
- (A_3) $DF(z_0)Z$ is a closed subspace.

The Lagrange function associated with the problem is given by the formula

$$\mathcal{L}(z, \lambda, e^*, w^*) = \lambda\psi(z) + \langle e^*, F(z) \rangle + \langle w^*, G(z) \rangle,$$

where $z \in Z$, $\lambda \in \mathbb{R}$, $e^* \in E^*$ and $w^* \in W^*$. We say that a triple (λ, e^*, w^*) are Lagrange multipliers at z_0 if

$$D\mathcal{L}(z_0, \lambda, e^*, w^*) = \lambda D\psi(z_0) + DF(z_0)^* e^* + DG(z_0)^* w^* = 0, \quad (4.2.11)$$

$$\lambda \geq 0, \quad w^* \in N(K, G(z_0)). \quad (4.2.12)$$

We denote by $\Lambda[z_0]$ the set of Lagrange multipliers (λ, e^*, w^*) at z_0 and by $\Lambda_1[z_0]$ the set of Lagrange multipliers (e^*, w^*) corresponding to $\lambda = 1$.

We have the following result on the first-order optimality conditions. Its proof follows from Theorem 2.2.9 for the scalar case $k = 1$ (see also [136]).

Proposition 4.2.1. *Suppose that z_0 is a locally optimal solution to under which (A_1^1) , (A_2) and (A_3) are satisfied. Then the set $\Lambda(z_0)$ is nonempty. In addition, if the Robinson constraint qualification (4.2.10) is fulfilled, then $\Lambda_1(z_0)$ is nonempty and weakly star compact.*

To deal with the second-order optimality conditions, we need to define the critical cone. Let us denote by $\mathcal{C}_0[z_0]$ the set of vectors $d \in Z$ satisfying the following conditions:

$$(a_1) \quad D\psi(z_0)d \leq 0,$$

$$(a_2) \quad DF(z_0)d = 0,$$

$$(a_3) \quad DG(z_0)d \in \text{cone}(K - G(z_0))$$

and by $\mathcal{C}[z_0]$ the closure of $\mathcal{C}_0[z_0]$ in Z , which is called the critical cone at z_0 . Each vector $d' \in \mathcal{C}[z_0]$ is called a critical direction.

We have the following result on first and second-order optimality conditions.

Proposition 4.2.2. *Suppose that $(A_1^2) - (A_3)$ are satisfied and the Robinson constraint qualification (4.2.10) is valid. Then if z_0 is a locally optimal solution, then $\Lambda_1[z_0]$ is nonempty and*

$$\begin{aligned} \max_{(e^*, w^*) \in \Lambda_1[z_0]} D_z^2 \mathcal{L}(z_0, e^*, w^*)[d, d] = \\ \max_{(e^*, w^*) \in \Lambda_1[z_0]} (D^2 \psi(z_0)d^2 + e^* D^2 F(z_0)d^2 + w^* D^2 G(z_0)d^2) \geq 0 \quad \forall d \in \mathcal{C}[z_0]. \end{aligned}$$

Proof. The proof follows from assertion (b) in Theorem 2.2.9 with $m = 1$ and $\Lambda_1[z_0]$ is a weakly star compact set. $\square$

4.2.3 State equation and linearized equation

We assume that $\partial\Omega$ is of class C^2 . Let $H := L^2(\Omega)$, $V := H_0^1(\Omega)$ and $D := H^2(\Omega) \cap H_0^1(\Omega)$. The norm and the scalar product in H are denoted by $|\cdot|$ and $(\cdot, \cdot)_H$, respectively. It is known that the embeddings

$$D \hookrightarrow V \hookrightarrow H$$

are compact and each space is dense in the following one.

Given $0 < \alpha \leq 1$, we denote by $C^{0,\alpha}(\bar{\Omega})$ and $C^{0,\alpha}(\bar{Q})$ the space of Hölder continuous functions of order α , on $\bar{\Omega}$ and $\bar{Q}$, respectively.

Let $H^{-1}(\Omega)$ be the dual of $H_0^1(\Omega)$. We define the following function

spaces

$$\begin{aligned}
H^1(Q) &= W_2^{1,1}(Q) = \{y \in L^2(Q) : \frac{\partial y}{\partial x_i}, \frac{\partial y}{\partial t} \in L^2(Q)\}, \\
V_2(Q) &= L^\infty(0, T; H) \cap L^2(0, T; V), \\
W(0, T) &= \{y \in L^2(0, T; H_0^1(\Omega)) : y_t \in L^2(0, T; H^{-1}(\Omega))\}, \\
W_2^{1,1}(0, T; V, H) &= \{y \in L^2([0, T], V) : \frac{\partial y}{\partial t} \in L^2([0, T], H)\}, \\
W_2^{1,1}(0, T; D, H) &= \{y \in L^2([0, T], D) : \frac{\partial y}{\partial t} \in L^2([0, T], H)\}, \\
U &= L^\infty(Q), \\
K_\infty &= \{u \in U : u(x, t) \leq 0 \text{ a.a. } (x, t) \in Q\}, \\
Y &= \left\{ y \in W_2^{1,1}(0, T; D, H) \cap L^\infty(Q) : \frac{\partial y}{\partial t} + Ay \in L^p(0, T; H) \right\},
\end{aligned}$$

where

$$\frac{4}{4-N} < p < 4 \quad \text{whenever } N = 2, \quad (4.2.13)$$

$$\frac{4}{4-N} < p < \frac{2N}{N-2} \quad \text{whenever } N = 3. \quad (4.2.14)$$

Y is endowed with the graph norm

$$\|y\|_Y := \|y\|_{W_2^{1,1}(0, T; D, H)} + \|y\|_{L^\infty(Q)} + \left\| \frac{\partial y}{\partial t} + Ay \right\|_{L^p(0, T; H)}.$$

Under the graph norm Y is a Banach space. Besides, we have the following embedding:

$$Y \hookrightarrow W_2^{1,1}(0, T; D, H) \hookrightarrow C(0, T; V), \quad (4.2.15)$$

$$W(0, T) \hookrightarrow C([0, T], H), \quad (4.2.16)$$

where $C([0, T], X)$ stands for the space of continuous mappings $v : [0, 1] \rightarrow X$ with X is a Banach space. Moreover, for all $v, w \in W(0, T)$, we have the following integration by part formula (see Proposition 23.23, p.422 in [237]) :

$$(v(t_1), w(t_1))_H - (v(0), w(0))_H = \int_0^{t_1} \langle v_t, w \rangle dt + \int_0^{t_1} \langle w_t, v \rangle dt, \quad (4.2.17)$$

where $t_1 \in [0, T]$. Recall that given $y_0 \in H$ and $u \in L^2(0, T; H)$, a function $y \in W(0, T)$ is said to be a weak solution of the semilinear

parabolic equation (4.2.2)-(4.2.3) if

$$\begin{aligned}
& \int_{\Omega} y(x, t_1) \eta(x, t_1) dx - \int_0^{t_1} \int_{\Omega} y(x, t) \eta_t(x, t) dx dt \\
& + \sum_{i,j=1}^N \int_0^{t_1} \int_{\Omega} a_{ij} D_i y(x, t) D_j \eta(x, t) dx dt + \int_0^{t_1} \int_{\Omega} f(y(x, t)) \eta(x, t) dx dt \\
& = \int_{\Omega} y_0(x) \eta(x, 0) dx + \int_0^{t_1} \int_{\Omega} u(x, t) \eta(x, t) dx dt
\end{aligned} \tag{4.2.18}$$

for all test functions $\eta \in W_2^{1,1}(0, T; V, H)$ and $t_1 \in [0, T]$. By standard arguments and formular (4.2.17), we can show that if $f(y) \in L^2(0, T; H)$ whenever $y \in L^2(0, T; V)$ then y is a weak solution of equations (4.2.2)-(4.2.3) if and only if for a.a. $t \in [0, T]$ one has

$$\begin{cases} \langle y_t, v \rangle + \sum_{i,j=1}^n \int_{\Omega} a_{ij} D_i y D_j v dx + (f(y), v)_H = (u, v)_H & \forall v \in H_0^1(\Omega), \\ y(0) = y_0. \end{cases} \tag{4.2.19}$$

If a weak solution y such that $y \in W_2^{1,1}(0, T; D, H)$ and $f(y) \in L^2(0, T; H)$ then we have $y_t + Ay + f(y) \in L^2(0, T; H)$ and

$$\langle y_t, v \rangle + (Ay, v) + (f(y), v) = (u, v) \quad \forall v \in H_0^1(\Omega).$$

Since $H_0^1(\Omega)$ is dense in $H = L^2(\Omega)$, we have

$$(y_t, v) + (Ay, v) + (f(y), v) = (u, v) \quad \forall v \in H.$$

Hence

$$y_t + Ay + f(y) = u \quad \text{a.a. } t \in [0, T], \quad y(0) = y_0.$$

In this case we say y is a *strong solution* of (4.2.2)-(4.2.3). From now on a solution to (4.2.2)-(4.2.3) is understood a strong solution.

Let us make the following assumptions which are related to the state equation.

(H1) Coefficients $a_{ij} = a_{ji} \in L^\infty(\Omega)$ for every $1 \leq i, j \leq N$, satisfy the uniform ellipticity conditions, i.e., there exists a number $\alpha > 0$ such that

$$\alpha |\xi|^2 \leq \sum_{i,j=1}^N a_{ij}(x) \xi_i \xi_j \quad \forall \xi \in \mathbb{R}^N \text{ for a.a. } x \in \Omega. \tag{4.2.20}$$

(H2) The mapping $f : \mathbb{R} \rightarrow \mathbb{R}$ is of class C^2 such that $f(0) = 0$ and $f'(y) \geq 0$. Furthermore, for each $M > 0$, there exists $k_M > 0$ such that

$$|f'(y)| \leq k_M, \quad (4.2.21)$$

$$|f(y_1) - f(y_2)| + |f'(y_1) - f'(y_2)| + |f''(y_1) - f''(y_2)| \leq k_M |y_1 - y_2| \quad (4.2.22)$$

for all y, y_1, y_2 satisfying $|y|, |y_1|, |y_2| \leq M$.

Lemma 4.2.3. *Suppose that (H1) and (H2) are satisfied and $y_0 \in L^\infty(\Omega) \cap H_0^1(\Omega)$. Then for each $u \in L^p(0, T; H)$ with $p > \frac{4}{4-N}$, the state equation (4.2.2)-(4.2.3) has a unique solution $y \in Y$ and there exist positive constants $C_1 > 0$ and $C_2 > 0$ such that*

$$\|y\|_{L^\infty(Q)} + \|f(y)\|_{L^2(0,T;H)} \leq C_1(\|u\|_{L^p([0,T],H)} + \|y_0\|_{L^\infty(\Omega)}) \quad (4.2.23)$$

and

$$\left\| \frac{\partial y}{\partial t} \right\|_{L^2([0,T],H)} + \|y\|_{L^2([0,T],D)} \leq C_2(\|u\|_{L^2([0,T],H)} + \|y_0\|_{L^\infty(Q)} + \|y_0\|_V). \quad (4.2.24)$$

Proof. By [49, Theorem 2.1] and [49, Lemma 2.2.III], for each $u \in L^p(0, T; H)$ with $p > \frac{4}{4-N}$, the state equation has a unique solution $y \in L^\infty(Q) \cap W_2^{1,1}(0, T; V, H)$ such that $f(y) \in L^2(0, T; H)$ and inequalities (4.2.23) is fulfilled. By (H2), we have $f(y) \in L^\infty(Q)$. Let us claim that $y \in W_2^{1,1}(0, T; D, H)$. In fact, we have

$$y_t + Ay = u - f(y), \quad y(0) = y_0.$$

Since $u - f(y) \in L^2(0, T; H)$, [79, Theorem 5, p. 360], implies that $y \in W_2^{1,1}(0, T; D, H) \cap L^\infty(0, T; V)$ and there exists a constant $C > 0$ such that

$$\left\| \frac{\partial y}{\partial t} \right\|_{L^2([0,T],H)} + \|y\|_{L^2([0,T],D)} \leq C(\|u\|_{L^2([0,T],H)} + \|f(y)\|_{L^2(0,T;H)} + \|y_0\|_V).$$

Combining this with (4.2.23), we obtain (4.2.24). $\square$

Let us consider the linearized equation

$$y_t + Ay + c(x, t)y = u, \quad y(0) = y_0. \quad (4.2.25)$$

Lemma 4.2.4. *Suppose that $u \in L^p(0, T; H)$ with $p > \frac{4}{4-N}$, $y_0 \in L^\infty(\Omega) \cap H_0^1(\Omega)$ and $c \in L^\infty(Q)$. Then equation (4.2.25) has a unique solution $y \in Y$.*

Proof. Since $c \in L^\infty(Q)$, there exists a constant $c_0 > 0$ such that $|c(x, t)| \leq c_0$ for a.a. $(x, t) \in Q$. By Theorem 5 in [79, Chapter 7], equation (4.2.25) has a unique solution $y \in W_2^{1,1}(0, T; D, H) \cap L^\infty(0, T; H_0^1(\Omega))$. It follows that $c(x, t)y \in L^p(0, T; H)$. Let us consider the equation

$$z_t + Az + c_0 z = u + (c_0 - c)y, \quad z(0) = y_0. \quad (4.2.26)$$

By [79, Theorem 5, Chapter 7] and Lemma 4.2.3, the equation has a unique solution $z \in Y$. From (4.2.25) and (4.2.26), we have

$$z_t - y_t + A(z - y) + c_0(z - y) = 0, \quad (z - y)(0) = 0. \quad (4.2.27)$$

Taking the scalar product both side with $z - y$ and using (H1), we get

$$\frac{1}{2} \frac{d}{dt} |z - y|^2 + \alpha |\nabla(z - y)|^2 + c_0 |z - y|^2 \leq 0.$$

Integrating on $[0, T]$, we obtain

$$\sup_{t \in [0, T]} |z - y|^2 + \alpha \|z - y\|_{L^2(0, T; D)}^2 + c_0 \|z - y\|_{L^2(0, T; H)}^2 \leq 0.$$

This implies that $z = y$ a.a. in Q . The conclusion is followed. $\square$

4.2.4 Existence of regular multipliers and second-order optimality conditions

Let $\phi : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a mapping which stands for L and g . We impose the following hypotheses.

(H3) ϕ is a Carathéodory function and for each $(x, t) \in \Omega \times [0, T]$, $\phi(x, t, \cdot, \cdot)$ is of class C^2 and satisfies the following property: for each $M > 0$, there exists $l_{\phi, M} > 0$ such that

$$\begin{aligned} & |\phi(x, t, y_1, u_1) - \phi(x, t, y_2, u_2)| + |\phi_y(x, t, y_1, u_1) - \phi_y(x, t, y_2, u_2)| \\ & + |\phi_u(x, t, y_1, u_1) - \phi_u(x, t, y_2, u_2)| + |\phi_{yy}(x, t, y_1, u_1) - \phi_{yy}(x, t, y_2, u_2)| \\ & + |\phi_{yu}(x, t, y_1, u_1) - \phi_{yu}(x, t, y_2, u_2)| + |\phi_{uu}(x, t, y_1, u_1) - \phi_{uu}(x, t, y_2, u_2)| \\ & \leq l_{\phi, M} (|y_1 - y_2| + |u_1 - u_2|) \end{aligned}$$

for all $(x, t, y_i, u_i) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}$ satisfying $|y_i|, |u_i| \leq M$ with $i = 1, 2$. Furthermore, we require that the functions $\phi_y(\cdot, \cdot, 0, 0)$, $\phi_u(\cdot, \cdot, 0, 0)$, $\phi_{yy}(\cdot, \cdot, 0, 0)$, $\phi_{yu}(\cdot, \cdot, 0, 0)$ and $\phi_{uu}(\cdot, \cdot, 0, 0)$ belong to $L^\infty(Q)$.

(H4) $f_y[x, t] \geq -\frac{1}{\epsilon} g_y[x, t] \geq 0$ for a.a. $(x, t) \in Q$ and there exists $\gamma > 0$ such that

$$a - \bar{u}(x, t) + \epsilon \bar{u}(x, t) + g[x, t] \leq -\gamma \quad \text{a.a. } (x, t) \in Q. \quad (4.2.28)$$

Hereafter, given $\bar{z} = (\bar{y}, \bar{u}) \in \Phi_3$,

$$\phi[x, t] := \phi(x, t, \bar{y}(x, t), \bar{u}(x, t)).$$

Note that hypothesis (H3) makes sure that J and g are of class C^2 on $Y \times U$. Hypothesis (H4) is a *generalized separation condition* which guarantees that the Robinson constraint qualification is satisfied and the Lagrange multipliers belong to $L^1(\Omega)$. Namely, the first part of (H4) makes sure that $f_y[x, t] + \frac{1}{\epsilon}g_y[x, t] \geq 0$. Under this condition, the maximum principle (see [143, Theorem 7.2, p. 188]) is applicable for the linearized equation

$$z_t + Az + (f_y[x, t] + \frac{1}{\epsilon}g_y[x, t])z = v.$$

This fact is important for the Robinson constraint qualification and the existence of the normalized Lagrange multipliers (see Lemma 2.5 in [130]). Meanwhile the second part of (H4) guarantees the regularity of Lagrange multipliers. We will see the significant role of (H4) in the proof of Lemma 4.2.7 and Theorem 4.2.8. Let us give an example under which (H4) is satisfied.

Example 4.2.5. Let $\epsilon = 1$, $a = 0$, $f(x, t, y) = y^3 + y$, $g(x, t, y) = -y - \gamma$ and $y_0(x) \geq 0$ for a.a. $x \in \Omega$. Then we have $f_y(x, t, y) = 3y^2 + 1$ and $g_y(x, t, y) = -1$. Hence $f_y(x, t, y) \geq -g_y(x, t, y) \geq 0$. Let $u \in [0, b]$. Then $0 \leq u \in L^\infty(Q)$. The maximum principle implies that the solution y of the state equation corresponding to u satisfies the property that $y(x, t) \geq 0$ for a.a. $(x, t) \in Q$. It follows that

$$a - u(x, t) + g(x, t, y) + u(x, t) \leq g(x, t, y) = -y(x, t) - \gamma \leq -\gamma. \quad (4.2.29)$$

Let us define Banach spaces

$$Z = Y \times U, \quad W = W_1 \times W_2, \quad E = L^\infty(Q),$$

$$W_1 = L^p(0, T; H), \quad W_2 = L^\infty(\Omega) \cap H_0^1(\Omega) \quad \text{with} \quad p > \frac{4}{4 - N}$$

and

$$K_\infty = \{v \in L^\infty(Q) : v(x, t) \leq 0 \text{ a.a. } (x, t) \in Q\} \quad (4.2.30)$$

$$K = K_\infty \times K_\infty \times K_\infty. \quad (4.2.31)$$

Define mappings $F : Y \times U \rightarrow W$ and $G : Y \times U \rightarrow E \times E \times E$ by setting

$$F(y, u) = (F_1(y, u), F_2(y, u)) = \left(\frac{\partial y}{\partial t} + Ay + f(y) - u, y(0) - y_0 \right) \quad (4.2.32)$$

and

$$G(y, u) = (G_1(y, u), G_2(y, u), G_3(y, u)), \quad (4.2.33)$$

where

$$G_1(y, u) = u - b, \quad (4.2.34)$$

$$G_2(y, u) = -u + a \quad (4.2.35)$$

$$G_3(y, u) = \epsilon u + g(y). \quad (4.2.36)$$

By definition of space Y , if $(y, u) \in Y \times U$ then $\frac{\partial y}{\partial t} + Ay \in L^p(0, T; H)$, $f(y) \in L^\infty(Q) \subset L^p(0, T; H)$ (since $y \in L^\infty(Q)$) and $u \in L^\infty(Q) \subset L^p(0, T; H)$. Hence

$$\frac{\partial y}{\partial t} + Ay + f(y) - u \in L^p(0, T; H) = W_1$$

and F_1 is well defined. Let us show that F_2 is also well defined. In fact, take any $y \in Y$. By (4.2.15), we have $y \in C([0, T], V)$. Hence $y(0) \in H_0^1(\Omega)$. Since $y \in L^\infty(Q)$, there exists a constant $M > 0$ such that $|y(t, x)| \leq M$ for a.a. $(t, x) \in Q$. Particularly, if $t_n \in [0, T]$ such that $t_n \rightarrow 0$, we have $|y(t_n, x)| \leq M$ for all n and a.a. $x \in \Omega$ and

$$y(0, \cdot) = \lim_{n \rightarrow \infty} y(t_n, \cdot) \quad \text{in } V.$$

This implies that $\|y(0, \cdot) - y(t_n, \cdot)\|_{L^2(\Omega)} \rightarrow 0$ as $n \rightarrow \infty$. By passing to a subsequence, we may assume that $y(t_n, x) \rightarrow y(0, x)$ for a.a. $x \in \Omega$. It follows that

$$|y(0, x)| \leq |y(0, x) - y(t_n, x)| + |y(t_n, x)| \leq |y(0, x) - y(t_n, x)| + M.$$

By passing to the limit when $n \rightarrow \infty$, we get $|y(0, x)| \leq M$. Hence $y(0, \cdot) \in L^\infty(Q)$. Consequently, $y \in H_0^1(\Omega) \cap L^\infty(Q)$. Hence F_2 is well defined. Consequently, F is well defined.

Then problem (4.2.1)-(4.2.5) can be formulated in the form

$$J(y, u) \rightarrow \min, \quad (4.2.37)$$

s.t.

$$F(y, u) = 0, \quad (4.2.38)$$

$$G(y, u) \in K. \quad (4.2.39)$$

Therefore, we can apply Proposition 4.2.2 for the problem (4.2.1)-(4.2.5) to derive necessary optimality conditions.

Let $\mathcal{C}_0[\bar{z}]$ be a set of all $z = (y, u) \in Y \times U$ such that the following conditions hold:

$$(c_1) \quad \int_Q (L_y[x, t]y(x, t) + L_u[x, t]u(x, t)) dx dt \leq 0;$$

$$(c_2) \quad \frac{\partial y}{\partial t} + Ay + f'(\bar{y})y = u, y(0) = 0;$$

$$(c_3) \quad (u, -u, g_y[\cdot, \cdot]y + \epsilon u) \in \text{cone}[(K_\infty - (\bar{u} - b)) \times (K_\infty - (a - \bar{u})) \times (K_\infty - (g[\cdot, \cdot] + \epsilon \bar{u}))].$$

We denote by $\mathcal{C}[\bar{z}]$ the closure of $\mathcal{C}_0[\bar{z}]$ in $Y \times U$. The set $\mathcal{C}[\bar{z}]$ is called a critical cone to problem (4.2.1)-(4.2.5) at $\bar{z}$. Each vector $(y, u) \in \mathcal{C}[\bar{z}]$ is called a critical direction.

Firstly, we have:

Lemma 4.2.6. *For each $\bar{z} \in \Phi_3$, the operator $DF(\bar{z})$ is surjective.*

Proof. It is sufficient to show that $F_y(\bar{z})$ is bijective. In fact, under assumptions (H1) and (H2), the linear mapping $F_y(\bar{z}) : Y \rightarrow W$, defined by

$$F_y(\bar{z})y = (F_{1y}(\bar{z})y, F_{2y}(\bar{z})y) = \left(\frac{\partial y}{\partial t} + Ay + f'(\bar{y})y, y(0) \right).$$

Taking any $v = (u, y_0) \in W$, $u \in L^p(0, T; H)$ and $y_0 \in L^\infty(\Omega) \cap H_0^1(\Omega)$, we consider equation $F_y(\bar{z})y = v$. This equation is equivalent to

$$\frac{\partial y}{\partial t} + Ay + f'(\bar{y})y = u, \quad y(0) = y_0.$$

Since $u \in L^p(0, T; H)$ with $p > \frac{4}{4-N}$, Lemma 4.2.4 implies that the above parabolic equation has a unique solution $y \in Y$. Thus $F_y(\bar{z})$ is bijective, and so, $DF(\bar{z})$ is surjective. $\square$

Lemma 4.2.7. *Suppose that $\bar{z} \in \Phi_3$ and (H4) is satisfied. Then the Robinson constraint qualification is fulfilled at $\bar{z}$.*

Proof. It is easy to see that

$$\text{int}K_\infty = \{v \in L^\infty(Q) \text{ and } \text{essup}v < 0\}$$

and

$$\text{int}K = (\text{int}K_\infty) \times (\text{int}K_\infty) \times (\text{int}K_\infty).$$

By Lemma 4.2.6, condition (i) of (4.2.10) is satisfied. It remains to show that the condition (ii) of (4.2.10) is satisfied, that is

$$\bar{u} - b + \tilde{u} \leq -\delta, \quad (4.2.40)$$

$$a - \bar{u} - \tilde{u} \leq -\delta, \quad (4.2.41)$$

$$g[x, t] + \epsilon \bar{u} + g_y[x, t] \tilde{y} + \epsilon \tilde{u} \leq -\delta \quad (4.2.42)$$

for some $\delta > 0$ and for some $\tilde{z} = (\tilde{y}, \tilde{u}) \in Y \times U$ satisfying the equation

$$\frac{\partial \tilde{y}}{\partial t} + A\tilde{y} + f'(\bar{y})\tilde{y} = \tilde{u} \quad \text{in } Q, \quad \tilde{y} = 0 \quad \text{on } \Sigma, \quad \tilde{y}(0) = 0 \quad \text{in } \Omega. \quad (4.2.43)$$

To show the existence of $\tilde{z}$, we do this as follows. For $0 < \rho < 1$, we define

$$\alpha = \alpha(x, t) := \frac{1}{\epsilon} g_y[x, t] \leq 0 \quad \text{a.a.} \quad (x, t) \in Q$$

$$w_\rho := \bar{u} - \rho + \alpha z_{\bar{u}}.$$

Here we denote by z_ζ the solution of equation $D_{(y,u)}F(\bar{z})(y, u) = 0$ with $u = \zeta$. Namely, z_ζ is a solution of the equation

$$z_t + Az + f'(\bar{y})z = \zeta, \quad z(0) = 0.$$

By Lemma 4.2.4, the equation

$$\frac{\partial \xi}{\partial t} + A\xi + [f'(\bar{y}) + \alpha]\xi = w_\rho \quad \text{in } Q, \quad \xi = 0 \quad \text{on } \Sigma, \quad \xi(0) = 0 \quad \text{in } \Omega. \quad (4.2.44)$$

has a unique solution ξ_ρ . This implies that

$$\frac{\partial \xi_\rho}{\partial t} + A\xi_\rho + f'(\bar{y})\xi_\rho = w_\rho - \alpha\xi_\rho.$$

Therefore, if we set

$$u_\rho = w_\rho - \alpha\xi_\rho = \bar{u} - \rho - \alpha(\xi_\rho - z_{\bar{u}}) \quad (4.2.45)$$

then $\xi_\rho = z_{u_\rho}$. Subtracting the equation satisfied by $z_{\bar{u}}$ from the equation satisfied by z_{u_ρ} , we get

$$\begin{cases} \frac{\partial(z_{u_\rho} - z_{\bar{u}})}{\partial t} + A(z_{u_\rho} - z_{\bar{u}}) + f'(\bar{y})(z_{u_\rho} - z_{\bar{u}}) = u_\rho - \bar{u} & \text{in } Q, \\ z_{u_\rho} - z_{\bar{u}} = 0 & \text{on } \Sigma, \\ (z_{u_\rho} - z_{\bar{u}})(0) = 0 & \text{in } \Omega. \end{cases} \quad (4.2.46)$$

On the other hand, from (4.2.45) we have

$$u_\rho - \bar{u} = -\rho - \alpha(\xi_\rho - z_{\bar{u}}) = -\rho - \alpha(z_{u_\rho} - z_{\bar{u}}). \quad (4.2.47)$$

Inserting (4.2.47) into (4.2.46), we obtain

$$\begin{cases} \frac{\partial(z_{u_\rho} - z_{\bar{u}})}{\partial t} + A(z_{u_\rho} - z_{\bar{u}}) + [f'(\bar{y}) + \alpha(x, t)](z_{u_\rho} - z_{\bar{u}}) = -\rho & \text{in } Q, \\ z_{u_\rho} - z_{\bar{u}} = 0 & \text{on } \Sigma, \\ (z_{u_\rho} - z_{\bar{u}})(0) = 0 & \text{in } \Omega. \end{cases} \quad (4.2.48)$$

In this equation, we have $-\rho < 0$ and $f'(\bar{y}) + \alpha(x, t) = f'(\bar{y}) + \frac{1}{\epsilon}g_y[x, t] \geq 0$ because of (H4). Then the maximum principle (see [143, Theorem 7.2, p. 188]) implies that

$$\xi_\rho = z_{u_\rho} \leq z_{\bar{u}} \quad \text{and} \quad \|z_{u_\rho} - z_{\bar{u}}\|_{L^\infty(Q)} \leq C_1\rho. \quad (4.2.49)$$

Combining this and (4.2.47), we get

$$\|u_\rho - \bar{u}\|_{L^\infty(Q)} \leq C_2\rho \quad (4.2.50)$$

for some constant $C_2 > 0$. Define

$$\begin{aligned} Q_a &= \{(x, t) \in Q : \bar{u}(x, t) = a\}, \\ Q_\rho &= \{(x, t) \in Q : a < \bar{u}(x, t) \leq a + 2C_2\rho\}. \end{aligned}$$

Obviously, we have $Q_a \cap Q_\rho = \emptyset$. Set

$$\hat{u}_\rho = (a + \rho)\chi_a + (\bar{u} + \rho)\chi_\rho + u_\rho(1 - \chi_a - \chi_\rho), \quad (4.2.51)$$

where χ_a is a characteristic function of the set Q_a and χ_ρ is a characteristic function of Q_ρ . We now claim that for ρ small enough, the couple

$$\tilde{z} := (z_{\hat{u}_\rho} - z_{\bar{u}}, \quad \hat{u}_\rho - \bar{u}) \in Y \times U \quad (4.2.52)$$

satisfies (4.2.40)-(4.2.43). Indeed, the definition of $\tilde{z}$ in (4.2.52) implies that (4.2.43) is satisfied. We now have

$$\hat{u}_\rho = a + \rho \text{ on } Q_a$$

and

$$\hat{u}_\rho = \bar{u} + \rho \geq a + \rho \text{ on } Q_\rho.$$

Since (4.2.50),

$$\hat{u}_\rho = u_\rho \geq \bar{u} - |u_\rho - \bar{u}| \geq a + 2C_2\rho - C_2\rho = a + C_2\rho \quad \text{on } Q \setminus (Q_a \cup Q_\rho)$$

. Hence, we obtain

$$\widehat{u}_\rho \geq \min(a + \rho, a + C_2\rho) = a + \rho \min(1, C_2) \quad \text{a.a. } (x, t) \in Q. \quad (4.2.53)$$

• Verification of (4.2.40). Firstly, on Q_a , we have

$$\bar{u} - b + (\widehat{u}_\rho - \bar{u}) = \widehat{u}_\rho - b = a + \rho - b = \rho - (b - a) \leq -\frac{b - a}{2} < 0 \quad (4.2.54)$$

for ρ small enough. Secondly, on Q_ρ , we have

$$\begin{aligned} \bar{u} - b + (\widehat{u}_\rho - \bar{u}) &= \widehat{u}_\rho - b = \bar{u} + \rho - b \leq a + 2C_2\rho + \rho - b \\ &= \rho(1 + 2C_2) - (b - a) \leq -\frac{b - a}{2} < 0 \end{aligned} \quad (4.2.55)$$

for ρ small enough. Thirdly, on $Q \setminus (Q_a \cup Q_\rho)$, we have from the definition of α and (4.2.49) that

$$\alpha(z_{u_\rho} - z_{\bar{u}}) \geq 0.$$

Hence

$$\bar{u} - b + (\widehat{u}_\rho - \bar{u}) = \widehat{u}_\rho - b = u_\rho - b = (\bar{u} - b) - \alpha(z_{u_\rho} - z_{\bar{u}}) - \rho \leq -\rho. \quad (4.2.56)$$

Thus, (4.2.40) is satisfied with $\tilde{z}$ defined by (4.2.52) for ρ small enough.

• Verification of (4.2.41). From (4.2.53) we have

$$a - \bar{u} - (\widehat{u}_\rho - \bar{u}) = a - \widehat{u}_\rho \leq a - [a + \rho \min(1, C_2)] = -\rho \min(1, C_2) < 0. \quad (4.2.57)$$

This implies that (4.2.41) is satisfied with $\tilde{z}$ defined in (4.2.52) and for ρ small enough.

• Verification of (4.2.42). Use (4.2.49) and (4.2.50), we get

$$\begin{aligned} \|z_{\widehat{u}_\rho} - z_{\bar{u}}\|_{L^\infty(Q)} &\leq \|z_{\widehat{u}_\rho} - z_{u_\rho}\|_{L^\infty(Q)} + \|z_{u_\rho} - z_{\bar{u}}\|_{L^\infty(Q)} \\ &\leq C_5 \|\widehat{u}_\rho - u_\rho\|_{L^\infty(Q)} + C_1\rho \\ &= C_5 \|(\bar{u} - u_\rho + \rho)(\chi_a + \chi_\rho)\|_{L^\infty(Q)} + C_1\rho \leq C_4\rho. \end{aligned} \quad (4.2.58)$$

On the other hand, on $Q_a \cup Q_\rho$, we have the following estimations

$$\begin{aligned} \bar{u} &\leq a + 2C_2\rho, \quad \widehat{u}_\rho - \bar{u} = \rho, \quad g_y[x, t] \leq 0, \\ - (z_{\widehat{u}_\rho} - z_{\bar{u}}) &\leq \|z_{\widehat{u}_\rho} - z_{\bar{u}}\|_{L^\infty(Q)} \leq C_4\rho. \end{aligned}$$

Combining these facts with (H4), we obtain the following estimations on $Q_a \cup Q_\rho$:

$$\begin{aligned}
& g[x, t] + \epsilon \bar{u} + g_y[x, t](z_{\hat{u}_\rho} - z_{\bar{u}}) + \epsilon(\hat{u}_\rho - \bar{u}) \\
& \leq -\gamma + (\bar{u} - a) - C_4 g_y[x, t]\rho + \epsilon\rho \\
& \leq -\gamma + 2C_2\rho - C_4 g_y[x, t]\rho + \epsilon\rho \\
& \leq -\gamma + \rho(2C_2 - C_4 g_y[x, t] + \epsilon) \leq \frac{-\gamma}{2}
\end{aligned} \tag{4.2.59}$$

for ρ small enough. It remains to show that the condition is satisfied on $Q \setminus (Q_a \cup Q_\rho)$. Indeed, since $g_y[\cdot, \cdot] \leq 0$, $\alpha = \frac{1}{\epsilon} g_y[x, t] \leq 0$. Hence

$$\alpha(z_{u_\rho} - z_{\bar{u}}) = \frac{1}{\epsilon} g_y[x, t](z_{u_\rho} - z_{\bar{u}}) \geq 0.$$

On the other hand

$$\hat{u}_\rho - u_\rho = \begin{cases} 2\rho + \alpha(z_{u_\rho} - z_{\bar{u}}) > 0 & \text{on } Q_a \cup Q_\rho, \\ 0 & \text{on } Q \setminus (Q_a \cup Q_\rho). \end{cases} \tag{4.2.60}$$

This implies that $\hat{u}_\rho - u_\rho \geq 0$. The maximum principle implies that $z_{\hat{u}_\rho} \geq z_{u_\rho}$. Hence

$$g_y[x, t](z_{\hat{u}_\rho} - z_{u_\rho}) \leq 0.$$

Combining these facts with (4.2.47), we have on $Q \setminus (Q_a \cup Q_\rho)$ that

$$\begin{aligned}
& (g[x, t] + \epsilon \bar{u}) + g_y[x, t](z_{\hat{u}_\rho} - z_{\bar{u}}) + \epsilon(\hat{u}_\rho - \bar{u}) \\
& = (g[x, t] + \epsilon \bar{u}) + g_y[x, t](z_{\hat{u}_\rho} - z_{\bar{u}}) + \epsilon(u_\rho - \bar{u}) \\
& \leq g_y[x, t](z_{\hat{u}_\rho} - z_{\bar{u}}) + \epsilon(u_\rho - \bar{u}) \\
& = g_y[x, t](z_{\hat{u}_\rho} - z_{u_\rho}) + g_y[x, t](z_{u_\rho} - z_{\bar{u}}) + \epsilon(u_\rho - \bar{u}) \\
& = g_y[x, t](z_{\hat{u}_\rho} - z_{u_\rho}) + g_y[x, t](z_{u_\rho} - z_{\bar{u}}) - \epsilon\rho - g_y[x, t](z_{u_\rho} - z_{\bar{u}}) \\
& = g_y[x, t](z_{\hat{u}_\rho} - z_{u_\rho}) - \epsilon\rho \leq -\epsilon\rho.
\end{aligned}$$

Hence (4.2.42) is satisfied with $\tilde{z}$ defined in (4.2.52) and for ρ small enough. The lemma is proved. $\square$

In the sequel, we will use the following sets:

$$Q_a := \{(x, t) \in Q : \bar{u}(x, t) = a\}, \tag{4.2.61}$$

$$Q_b := \{(x, t) \in Q : \bar{u}(x, t) = b\}, \tag{4.2.62}$$

$$Q_{ab} := \{(x, t) \in Q : a < \bar{u}(x, t) < b\}, \tag{4.2.63}$$

$$Q_0 := \{(x, t) \in Q : g[x, t] + \epsilon \bar{u}(x, t) = 0\}. \tag{4.2.64}$$

The following theorem gives first and second-order necessary optimality conditions and the regularity of multipliers.

Theorem 4.2.8. Suppose that $(\bar{y}, \bar{u}) \in \Phi_3$ is a locally optimal solution of the problem (4.2.1)-(4.2.5) under which assumptions (H1) – (H4) are satisfied. Then there exist a function $\varphi \in L^\infty(Q) \cap W_2^{1,1}(0, T; D, H)$ and functions $e, \hat{e} \in L^\infty(Q)$ such that the following conditions are fulfilled:

(i) (the adjoint equation)

$$-\frac{\partial \varphi}{\partial t} + A^* \varphi + f'(\bar{y})\varphi = -L_y[\cdot, \cdot] - e g_y[\cdot, \cdot], \quad \varphi(\cdot, T) = 0, \quad (4.2.65)$$

where A^* is the adjoint operator of A , which is defined by

$$A^* \varphi = - \sum_{i,j=1}^N D_i(a_{ij}(x) D_j \varphi);$$

(ii) (the stationary condition in u)

$$L_u[x, t] - \varphi(x, t) + \epsilon e(x, t) + \hat{e}(x, t) = 0 \quad \text{a.a.} \quad (x, t) \in Q; \quad (4.2.66)$$

(iii) (the complimentary condition) $e(x, t) \geq 0$ and $e(x, t)(g[x, t] + \epsilon \bar{u}(x, t)) = 0$ a.a. $(x, t) \in Q$. $\hat{e}(x, t)$ has the property that

$$\hat{e}(x, t) \begin{cases} \leq 0 & \text{a.a.} \quad (x, t) \in Q_a, \\ \geq 0 & \text{a.a.} \quad (x, t) \in Q_b, \\ 0 & \text{a.a.} \quad (x, t) \in Q_{ab}. \end{cases} \quad (4.2.67)$$

(iv) (the non-negative second-order condition)

$$\begin{aligned} & \max_{(\varphi, e, \hat{e}) \in \Lambda_0[\bar{z}]} \left(\int_Q (L_{yy}[x, t] \zeta^2 + 2L_{yu}[x, t] \zeta v + L_{uu}[x, t] v^2) dx dt \right. \\ & \left. + \int_Q e(x, t) g_{yy}[x, t] \zeta^2 dx dt + \int_Q \varphi f''(\bar{y}) \zeta^2 dx dt \right) \geq 0 \quad \forall (\zeta, v) \in \mathcal{C}[\bar{z}]. \end{aligned} \quad (4.2.68)$$

Here $\Lambda_0[\bar{z}]$ is the set of Lagrange multipliers $(\varphi, e, \hat{e})$ satisfying conditions (i), (ii) and (iii).

Proof. Let K , F and G be defined by (4.2.31), (4.2.32) and (4.2.33), respectively. Then problem (4.2.1)-(4.2.5) is formulated in the form of problem (4.2.37)-(4.2.39). From (H2) and (H3) we see that the mappings J, F and G are of class C^2 around $\bar{z} = (\bar{y}, \bar{u})$. Here $D_{(y,u)} J(\bar{z})$, $D_{(y,u)}^2 J(\bar{z})$, $D_{(y,u)} F(\bar{z})$, $D_{(y,u)}^2 F(\bar{z})$, $D_{(y,u)} G(\bar{z})$ and $D_{(y,u)}^2 G(\bar{z})$ are given

by

$$D_{(y,u)}J(\bar{z})(y, u) = \int_Q (L_y[x, t]y + L_u[x, t]u)dxdt, \quad (4.2.69)$$

$$D_{(y,u)}^2J(\bar{z})(y, u)^2 = \int_Q (L_{yy}[x, t]y^2 + 2L_{yu}[x, t]yu + L_{uu}[x, t]u^2)dxdt, \quad (4.2.70)$$

$$D_{(y,u)}F(\bar{z})(y, u) = (D_{(y,u)}F_1(\bar{z})(y, u), D_{(y,u)}F_2(\bar{z})(y, u)) \quad (4.2.71)$$

$$D_{(y,u)}^2F(\bar{z})(y, u)^2 = \left(D_{(y,u)}^2F_1(\bar{z})(y, u)^2, D_{(y,u)}^2F_2(\bar{z})(y, u)^2 \right), \quad (4.2.72)$$

where

$$\begin{aligned} D_{(y,u)}F_1(\bar{z})(y, u) &= \frac{\partial y}{\partial t} + Ay + f'(\bar{y})y - u \\ D_{(y,u)}F_2(\bar{z})(y, u) &= y(0), \\ D_{(y,u)}^2F_1(\bar{z})(y, u)^2 &= f''(\bar{y})y^2 \\ D_{(y,u)}^2F_2(\bar{z})(y, u)^2 &= 0. \end{aligned}$$

Also

$$\begin{aligned} D_{(y,u)}G(\bar{z})(y, u) &= (D_{(y,u)}G_1(\bar{z})(y, u), D_{(y,u)}G_2(\bar{z})(y, u), D_{(y,u)}G_3(\bar{z})(y, u)) \\ &= (u, -u, g_y[x, t]y + \epsilon u), \end{aligned} \quad (4.2.73)$$

$$D_{(y,u)}^2G(\bar{z})(y, u)^2 = (0, 0, g_{yy}[x, t]y^2) \quad \text{for all } (y, u) \in Z. \quad (4.2.74)$$

Let q be a conjugate number of p with $p > \frac{4}{4-N}$, that is, $\frac{1}{p} + \frac{1}{q} = 1$. Since $H = L^2(\Omega)$, $H = H^*$. Hence dual of $L^p(0, T; H)$ is $L^q(0, T; H)$. We recall that $E^* = (L^\infty(Q))^*$ and $W^* = W_1^* \times W_2^*$, where $W_1^* = L^p(0, T; H)^* = L^q(0, T; H) \subset L^1(0, T; H)$ and $W_2^* = (L^\infty(\Omega) \cap H_0^1(\Omega))^*$. Then the Lagrange function which is associated with the problem is defined as follows

$$\begin{aligned} \mathcal{L} : Y \times U \times L^q(0, T; H) \times (L^\infty(\Omega) \cap H_0^1(\Omega))^* \times ((L^\infty(Q))^*)^3 &\rightarrow \mathbb{R}, \\ \mathcal{L}(y, u, \phi_1, \phi_2^*, e_1^*, e_2^*, e_3^*) &= \\ \int_Q L(x, t, y, u)dxdt + \int_Q \left[\phi_1 \left(\frac{\partial y}{\partial t} + Ay + f(y) - u \right) \right] dxdt &+ \\ + \langle \phi_2^*, y(0) - y_0 \rangle_{W_2^*, W_2} + \sum_{i=1}^3 \langle e_i^*, G_i(y, u) \rangle_{L^\infty(Q)^*, L^\infty(Q)}. \end{aligned}$$

By Lemma 4.2.6, $DF(\bar{z})$ is surjective. By Lemma 4.2.7 the Robinson constraint qualification is valid at $\bar{z}$. Hence Proposition 4.2.2 is applicable. According to Proposition 4.2.2, there exist $\phi_1 \in L^q(0, T; H)$, $\phi_2^* \in$

$(L^\infty(\Omega) \cap H_0^1(\Omega))^*$ and $e^* = (e_1^*, e_2^*, e_3^*) \in L^\infty(Q)^* \times L^\infty(Q)^* \times L^\infty(Q)^*$ such that the following conditions hold:

$$D_{(y,u)}\mathcal{L}(\bar{z}, \phi_1, \phi_2^*, e^*) = 0, \quad (4.2.75)$$

$$e_1^* \in N(K_\infty, \bar{u} - b), \quad (4.2.76)$$

$$e_2^* \in N(K_\infty, -\bar{u} + a), \quad (4.2.77)$$

$$e_3^* \in N(K_\infty, g[\cdot, \cdot] + \epsilon\bar{u}), \quad (4.2.78)$$

$$\max_{(\phi_1, \phi_2^*, e^*) \in \Lambda_1[\bar{z}]} D_{(y,u)}^2 \mathcal{L}(\bar{z}, \phi_1, \phi_2^*, e^*)[d, d] \geq 0 \quad \forall d = (\zeta, v) \in \mathcal{C}[\bar{z}]. \quad (4.2.79)$$

The condition (4.2.75) is equivalent to the two following relations

$$\begin{aligned} & \int_Q L_y[x, t] y dx dt + \int_Q \left[\phi_1 \left(\frac{\partial y}{\partial t} + Ay + f'(\bar{y})y \right) \right] dx dt + \\ & + \langle \phi_2^*, y(0) \rangle_{W_2^*, W_2} + \langle e_3^*, g_y[x, t]y \rangle_{L^\infty(Q)^*, L^\infty(Q)} = 0, \quad \forall y \in Y \end{aligned} \quad (4.2.80)$$

and

$$\int_Q L_u[x, t] u dx dt - \int_Q \phi_1 u dx dt + \langle e_1^* - e_2^* + \epsilon e_3^*, u \rangle_{L^\infty(Q)^*, L^\infty(Q)} = 0, \quad \forall u \in U. \quad (4.2.81)$$

From (4.2.76)-(4.2.78), we see that e_i^* are non-negative functional. By Theorem 2.3 in [234], e_i^* with $i = 1, 2, 3$ are finitely additive measures on Q . Let us claim that each e_i^* can be represented by densities in $L^1(Q)$. This can be done as follows. From Theorem 1.23 in [234], we have representations

$$e_i^* = e_{ic}^* + e_{ip}^*, \quad e_{ic}^* \geq 0, \quad e_{ip}^* \geq 0, \quad (4.2.82)$$

where e_{ic}^* is countably additive and e_{ip}^* is purely finitely additive. Our goal is to show that $e_{ip}^* = 0$. Then e_i^* is countably additive which can be represented by densities in $L^1(Q)$. According to [234, Theorem 1.22], there exists a sequence $(Q_n)_{n \in \mathbb{N}}$ of Lebesgue measurable sets such that $Q \supset Q_1 \supset Q_2 \supset \dots \supset Q_n \supset \dots$, $\lim_{n \rightarrow +\infty} |Q_n| = 0$ and $e_{1p}^*(Q_n) = e_{1p}^*(Q)$ for all $n \in \mathbb{N}$. This means, that $|Q_n| = \text{meas}(Q_n) \rightarrow 0$ as $n \rightarrow +\infty$, but

$$\|e_{1p}^*\|_{L^\infty(Q)^*} = \int_Q de_{1p}^* = \int_{Q_n} de_{1p}^* = \langle e_{1p}^*, \chi_{Q_n} \rangle_{L^\infty(Q)^*, L^\infty(Q)} \quad \text{for all } n \in \mathbb{N}. \quad (4.2.83)$$

We now define

$$\begin{aligned} A_1(\delta) &= \{(x, t) \in Q : \bar{u}(x, t) - b \geq -\delta\}, \\ A_2(\delta) &= \{(x, t) \in Q : a - \bar{u}(x, t) \geq -\delta\}, \\ A_3(\delta) &= \{(x, t) \in Q : g[x, t] + \epsilon\bar{u}(x, t) \geq -\delta\}. \end{aligned}$$

Then it is easy to check that $A_1(\delta) \cap A_2(\delta) = \emptyset$ for all $\delta \in (0, \frac{b-a}{2})$. Moreover, condition (H4) guarantees that $A_3(\delta) \cap A_2(\delta) = \emptyset$ for δ small enough. Hence, we can choose $\delta = \delta^* \in (0, \frac{b-a}{2})$ such that $A_1(\delta^*) \cap A_2(\delta^*) = A_3(\delta^*) \cap A_2(\delta^*) = \emptyset$. Also, we have

$$\int_{Q \setminus A_i(\delta^*)} de_i^* = 0, \quad i = 1, 2, 3 \quad (4.2.84)$$

because the support of e_i^* is a subset of $A_i(\delta^*)$. Next, we define

$$\widehat{u}(x, t) = \begin{cases} \max(\frac{1}{\epsilon}, 1) & \text{on } A_1(\delta^*) \cup A_3(\delta^*) \\ -1 & \text{on } A_2(\delta^*) \\ 0 & \text{otherwise.} \end{cases} \quad (4.2.85)$$

Then the function $v_n := \chi_{Q_n} \widehat{u} \in L^\infty(Q)$ for $n = 1, 2, \dots$. Without loss of generality, we can assume $Q_n \subset A_1(\delta^*)$. Inserting $u = v_n$ into (4.2.81), we get

$$- \int_{Q_n} L_u[x, t] \widehat{u} dx dt + \int_{Q_n} \phi_1 \widehat{u} dx dt = \langle e_1^* + \epsilon e_3^*, v_n \rangle_{L^\infty(Q)^*, L^\infty(Q)}. \quad (4.2.86)$$

Here we use the fact that $\text{supp}(e_2^*) \subset Q \setminus A_1(\delta^*)$ which implies that $\langle e_2^*, v_n \rangle_{L^\infty(Q)^*, L^\infty(Q)} = 0$. Moreover, since $e_3^* \geq 0$ and $v_n \geq 0$ on $\text{supp}(e_3^*)$, $(e_1^*)^c \geq 0$ and $v_n = \chi_{Q_n} \max(\frac{1}{\epsilon}, 1) \geq \chi_{Q_n}$ on $\text{supp}(e_1^*)$, deduce from (4.2.86), we have

$$\begin{aligned} - \int_{Q_n} L_u[x, t] \widehat{u} dx dt + \int_{Q_n} \phi_1 \widehat{u} dx dt &\geq \langle e_1^*, v_n \rangle_{L^\infty(Q)^*, L^\infty(Q)} \\ &= \int_Q v_n de_{1c}^* + \int_Q v_n de_{1p}^* \geq \int_Q v_n de_{1p}^* \\ &\geq \int_{Q_n} de_{1p}^* = \int_Q de_{1p}^* = \|e_{1p}^*\|_{L^\infty(Q)^*}. \end{aligned}$$

By letting $n \rightarrow +\infty$, we get $0 \geq \|e_{1p}^*\|_{L^\infty(Q)^*}$. Hence $e_{1p}^* = 0$. By the same arguments, we can show that $e_{2p}^* = 0$ and $e_{3p}^* = 0$. Therefore, the measures e_1^*, e_2^* and e_3^* are countably additive. By the Radon-Nikodym theorem (see Theorem 1.47 in [1]), e_i^* can be represented by L^1 -density, that is, there exists $e_i \in L^1(Q)$ such that

$$\int_Q v(x, t) de_i^* = \int_Q e_i(x, t) v(x, t) dx dt \quad \forall v \in L^\infty(Q), \quad i = 1, 2, 3.$$

For simplification, we can identify e_i^* by e_i with $i = 1, 2, 3$. Then from (4.2.76), (4.2.77) and (4.2.78), we have

$$\begin{aligned} \int_Q e_1(x, t)[\eta(x, t) - (\bar{u}(x, t) - b)]dxdt &\leq 0 \quad \forall \eta \in K_\infty; \\ \int_Q e_2(x, t)[\eta(x, t) - (-\bar{u}(x, t) + a)]dxdt &\leq 0 \quad \forall \eta \in K_\infty; \\ \int_Q e_3(x, t)[\eta(x, t) - (\epsilon \bar{u}(x, t) + g[x, t])]dxdt &\leq 0 \quad \forall \eta \in K_\infty. \end{aligned}$$

By [178, Corollary 4], we have

$$\begin{aligned} e_1(x, t) &\in N((-\infty, 0], \bar{u}(x, t) - b) \quad \text{a.a.} \quad (x, t) \in Q; \\ e_2(x, t) &\in N((-\infty, 0], -\bar{u}(x, t) + a) \quad \text{a.a.} \quad (x, t) \in Q; \\ e_3(x, t) &\in N((-\infty, 0], g[x, t] + \epsilon \bar{u}(x, t)) \quad \text{a.a.} \quad (x, t) \in Q. \end{aligned}$$

This implies that $e_i(x, t) \geq 0$ with $i = 1, 2, 3$, $e_1(x, t)(\bar{u}(x, t) - b) = 0$, $e_2(x, t)(-\bar{u}(x, t) + a) = 0$ and $e_3(x, t)(g[x, t] + \epsilon \bar{u}(x, t)) = 0$ for a.a. $(x, t) \in Q$. By setting

$$e := e_3 \quad \text{and} \quad \widehat{e} := e_1 - e_2, \quad (4.2.87)$$

we see that $e, \widehat{e} \in L^1(Q)$ and satisfy the complementary condition (iii) of the theorem. Next we derive the assertion (ii) of the theorem. For this we rewrite (4.2.80) and (4.2.81) in the following equivalent forms:

$$\begin{aligned} \int_Q L_y[x, t]ydxdt + \int_Q \left[\phi_1 \left(\frac{\partial y}{\partial t} + Ay + f'(\bar{y})y \right) \right] dxdt + \\ + \langle \phi_2^*, y(0) \rangle_{W_2^*, W_2} + \int_Q e g_y[x, t]ydxdt = 0, \quad \forall y \in Y \end{aligned} \quad (4.2.88)$$

and

$$\int_Q L_u[x, t]udxdt - \int_Q \phi_1 udxdt + \int_Q (\epsilon e + \widehat{e})udxdt = 0, \quad \forall u \in U. \quad (4.2.89)$$

The latter implies that

$$L_u[\cdot, \cdot] - \phi_1 + \epsilon e + \widehat{e} = 0 \quad \text{a.a.} \quad (x, t) \in Q. \quad (4.2.90)$$

Define $\zeta = \phi_1 - L_u[\cdot, \cdot]$. Then we have $\zeta \in L^q(0, T; H)$ and

$$\epsilon e + \widehat{e} = \zeta.$$

Let us show that $\widehat{e} \in L^q(0, T; H)$. In fact, for $(x, t) \in Q_b$, we have

$$0 \leq \widehat{e}(x, t) = \zeta(x, t) - \epsilon e(x, t) \leq \zeta(x, t);$$

for $(x, t) \in Q_a$, we have from (H4) that

$$\epsilon \bar{u}(x, t) + g[x, t] \leq -\gamma.$$

The condition $e(x, t)[\epsilon \bar{u}(x, t) + g[x, t]] = 0$ implies that $e(x, t) = 0$ on Q_a and so $\widehat{e}(x, t) = \zeta(x, t)$ on Q_a . Note that $\widehat{e}(x, t) = 0$ for $(x, t) \in Q \setminus (Q_a \cup Q_b)$. For each $t \in [0, T]$, we define

$$\begin{aligned}\Omega_{at} &= \{x \in \Omega : (x, t) \in Q_a\}, \\ \Omega_{bt} &= \{x \in \Omega : (x, t) \in Q_b\}.\end{aligned}$$

Then we have

$$\begin{aligned}\int_0^T \left(\int_{\Omega} \widehat{e}(x, t)^2 dx \right)^{q/2} dt &= \int_0^T \left(\int_{\Omega_{at}} \widehat{e}(x, t)^2 dx + \int_{\Omega_{bt}} \widehat{e}(x, t)^2 dx \right)^{q/2} dt \\ &\leq \int_0^T \left(\int_{\Omega_{at}} \zeta(x, t)^2 dx + \int_{\Omega_{bt}} \zeta(x, t)^2 dx \right)^{q/2} dt \\ &\leq \int_0^T \left(\int_{\Omega} \zeta(x, t)^2 dx \right)^{q/2} dt < +\infty\end{aligned}$$

Hence $\widehat{e} \in L^q(0, T; H)$ and so is e . This implies that $eg_y[\cdot, \cdot] \in L^q(0, T; H)$.

Let us consider the following equation

$$\begin{aligned}-\frac{\partial \varphi}{\partial t} + A^* \varphi + f'(\bar{y}) \varphi \\ = -L_y[\cdot, \cdot] - eg_y[\cdot, \cdot], \quad \varphi(\cdot, T) = 0,\end{aligned}\tag{4.2.91}$$

where $-L_y[\cdot, \cdot] - eg_y[\cdot, \cdot] \in L^q(0, T; H)$. By changing variable $\widetilde{\varphi}(t) = \varphi(T - t)$, the equation (4.2.91) becomes

$$\begin{aligned}\frac{\partial \widetilde{\varphi}(t)}{\partial t} + A^* \widetilde{\varphi}(t) + f'(\bar{y}(x, T - t)) \widetilde{\varphi}(t) \\ = -L_y[x, T - t] - e(x, T - t)g_y[x, T - t], \\ \widetilde{\varphi}(\cdot, 0) = 0.\end{aligned}$$

Let us choose $r > 0$ such that

$$\frac{1}{r} + \frac{N}{2q} = 1 + \frac{N}{4}.$$

Then for $N = 2$, we have from (4.2.13) that $q \in (\frac{4}{3}, 2)$. Hence $r \in (\frac{3}{2}, 2)$. When $N = 3$, we have from (4.2.13) that $q \in (\frac{6}{5}, \frac{4}{3}) = (\frac{2N}{N+2}, \frac{4}{3})$. Hence $r \in (\frac{8}{5}, 2)$. It follows that

$$L^q(0, T; H) \hookrightarrow L^q(0, T; L^r(\Omega)).$$

Hence $-L_y[\cdot, \cdot] - eg_y[\cdot, \cdot] \in L^q(0, T; L^r(\Omega))$. By Theorem 4.1 in [143] (page 153, Chapter 3), the above equation has a unique weak solution $\tilde{\varphi} \in V_2(Q) = L^\infty(0, T; H) \cap L^2(0, T; V)$. Hence $\varphi(t) = \tilde{\varphi}(T - t)$ is a weak solution to the equation (4.2.91), that is

$$\begin{aligned} & \int_{\Omega} \varphi(x, t_1) \eta(x, t_1) dx - \int_{\Omega} \varphi(x, 0) \eta(x, 0) dx + \int_Q \varphi \eta_t dx dt + \\ & + \int_Q \left(\sum_{i,j=1}^N a_{ij} D_i \eta D_j \varphi + f'(\bar{y}) \eta \varphi \right) dx dt = - \int_Q (L_y[\cdot, \cdot] + eg_y[\cdot, \cdot]) \eta dx dt \end{aligned}$$

for all test functions $\eta \in W_2^{1,1}(0, T; V, H)$ and $t_1 \in [0, T]$. By taking $\eta = y \in Y$ with $y(0) = 0$ and $t_1 = T$, we have

$$\begin{aligned} & \int_Q (y_t \varphi + \sum_{i,j=1}^N a_{ij} D_i y D_j \varphi + f'(\bar{y}) y \varphi) dx dt = \\ & \int_Q (-L_y[x, t] - e(x, t) g_y[x, t]) y dx dt. \end{aligned}$$

Using the integration by part formula (see Theorem 1.5.3.1 in [96]), we get

$$\int_Q (y_t + Ay + f'(\bar{y})y) \varphi dx dt = \int_Q (-L_y[x, t] - e(x, t) g_y[x, t]) y dx dt. \quad (4.2.92)$$

Subtracting (4.2.88) from (4.2.92), we get

$$\int_Q (\varphi - \phi_1) \left(y_t + Ay + f'(\bar{y})y \right) dx dt = 0, \quad \text{for all } y \in Y, \quad y(0) = 0. \quad (4.2.93)$$

On the other hand, as it was shown that for every $\vartheta \in L^p(0, T; H)$, the parabolic equation

$$y_t + Ay + f'(y)y = \vartheta, \quad y(0) = 0$$

has a unique solution $y \in Y$. From this and (4.2.93), we deduce that

$$\int_Q (\varphi - \phi_1) \vartheta dx dt = 0 \quad \text{for all } \vartheta \in L^p(0, T; H).$$

This implies that

$$\phi_1 = \varphi \in L^\infty(0, T; H) \cap L^2(0, T; V).$$

Then (4.2.90) becomes

$$L_u[x, t] - \varphi + \epsilon e + \widehat{e} = 0, \quad \text{a.a. } (x, t) \in Q. \quad (4.2.94)$$

Hence assertion (ii) of the theorem is derived. Repeating the procedure in the proof of claim $e, \widehat{e} \in L^p(0, T; H)$ and using the fact $L_u[x, t] - \varphi \in L^\infty(0, T; H)$ and equality (4.2.94), we can show that $e, \widehat{e} \in L^\infty(0, T; H)$. Hence $-L_y[\cdot, \cdot] - g_y[\cdot, \cdot]e \in L^\infty(0, T; H)$. By the Lemma 4.2.4, the solution φ of the adjoint equation belongs to $L^\infty(Q) \cap W_2^{1,1}(0, T; D, H)$. Again, we obtain from (4.2.94) that $e, \widehat{e} \in L^\infty(Q)$. Hence assertion (i) of the theorem is established. Assertion (iv) of the theorem follows from condition (4.2.79) and formulae (4.2.70), (4.2.72) and (4.2.74). The proof of the theorem is complete. $\square$

To deal with second-order sufficient optimality conditions, we need to enlarge the critical cone $\mathcal{C}[\bar{z}]$ by the cone $\mathcal{C}_2[\bar{z}]$ which consists of couples $(y, u) \in W_2^{1,1}(0, T; D, H) \times L^2(Q)$ satisfying the following conditions:

$$(c'_1) \quad \int_Q (L_y[x, t]y(x, t) + L_u[x, t]u(x, t)) dx dt \leq 0;$$

$$(c'_2) \quad \frac{\partial y}{\partial t} + Ay + f'(\bar{y})y = u, y(0) = 0;$$

$$(c'_3) \quad u(x, t) \in T([a, b], \bar{u}(x, t)) \text{ for a.a. } (x, t) \in Q;$$

$$(c'_4) \quad g_y[x, t]y(x, t) + \epsilon u(x, t) \in T((-\infty, 0], g[x, t] + \epsilon \bar{u}(x, t)) \text{ for a.a. } (x, t) \in Q.$$

It easy to see that conditions (c'_3) and (c'_4) can be written in the following form:

$$(c''_3) \quad (u(x, t), -u(x, t), g_y[x, t]y(x, t) + \epsilon u(x, t)) \in T((-\infty, 0], \bar{u}(x, t) - b) \times T((-\infty, 0], a - \bar{u}(x, t)) \times T((-\infty, 0], g[x, t] + \epsilon \bar{u}(x, t))$$

for a.a. $(x, t) \in Q$. Hence $\mathcal{C}_2[\bar{z}]$ contains the closure of $\mathcal{C}_0[\bar{z}]$ in $W_2^{1,1}(0, T, D, H) \times L^2(0, T; H)$.

Let us rewrite the Lagrange function in the form

$$\begin{aligned} \mathcal{L}(y, u, \varphi, e, \widehat{e}) = & \int_Q L(x, t, y, u) dx dt + \int_Q \left[y \left(-\frac{\partial \varphi}{\partial t} + A^* \varphi \right) + (f(y) - u) \varphi \right] dx dt \\ & - \int_\Omega \varphi(0) y_0 dx + \int_Q e [g(y) + \epsilon u] dx dt + \int_Q \widehat{e} u dx dt. \end{aligned} \quad (4.2.95)$$

Then conditions (i) and (ii) of Theorem 4.2.8 is equivalent to condition $D_{(y,u)} \mathcal{L}(\bar{y}, \bar{u}, \varphi, e, \widehat{e}) = 0$ and condition (iv) is of the form

$$\max_{(\varphi, e, \widehat{e}) \in \Lambda_0[\bar{z}]} D_{(y,u)}^2 \mathcal{L}(\bar{y}, \bar{u}, \varphi, e, \widehat{e})[d, d] \geq 0 \quad \forall d \in \mathcal{C}[\bar{z}].$$

The following theorem gives second-order sufficient conditions for locally optimal solutions to problem (4.2.1)–(4.2.5).

Theorem 4.2.9. *Suppose that (H1), (H2), (H3) and (H4) are valid, there exist the multipliers $\varphi \in W_2^{1,1}(0, T; D, H) \cap L^\infty(Q)$ and $e, \widehat{e} \in L^\infty(Q)$ corresponding to $(\bar{y}, \bar{u}) \in \Phi_3$ satisfies conclusions (i), (ii) and (iii) of Theorem 4.2.8 and the following strictly second-order condition:*

$$\int_Q \left[2L_{yu}[x, t] \zeta v + \left(L_{yy}[x, t] + e g_{yy}[x, t] + \varphi f''(\bar{y}) \right) \zeta^2 + L_{uu}[x, t] v^2 \right] dx dt > 0 \quad (4.2.96)$$

for all $(\zeta, v) \in \mathcal{C}_2[(\bar{y}, \bar{u})] \setminus \{(0, 0)\}$. Furthermore, there exists a number $\Lambda > 0$ such that

$$L_{uu}[x, t] \geq \Lambda \quad \text{a.e. } (x, t) \in Q. \quad (4.2.97)$$

Then there exist number $\epsilon > 0$ and $\kappa > 0$ such that

$$J(y, u) \geq J(\bar{y}, \bar{u}) + \kappa \|u - \bar{u}\|_{L^2(Q)}^2 \quad \forall (y, u) \in \Phi_3 \cap [B_Y(\bar{y}, \epsilon) \times B_U(\bar{u}, \epsilon)]. \quad (4.2.98)$$

Here $B_Y(\bar{y}, \epsilon)$ and $B_U(\bar{u}, \epsilon)$ are balls in Y and $L^\infty(Q)$, respectively.

Proof. Suppose to the contrary that the conclusion were false. Then, we could find sequences $\{(y_n, u_n)\} \subset \Phi_3$ and $\{\gamma_n\} \subset \mathbb{R}_+$ such that $y_n \rightarrow \bar{y}$ in Y , $u_n \rightarrow \bar{u}$ in $L^\infty(Q)$, $\gamma_n \rightarrow 0^+$ and

$$J(y_n, u_n) < J(\bar{y}, \bar{u}) + \gamma_n \|u_n - \bar{u}\|_{L^2(Q)}^2. \quad (4.2.99)$$

If $u_n = \bar{u}$ then by the uniqueness we have $y_n = \bar{y}$. This leads to $J(\bar{y}, \bar{u}) < J(\bar{y}, \bar{u}) + 0$ which is absurd. Therefore, we can assume that $u_n \neq \bar{u}$ for all $n \geq 1$.

Define $t_n = \|u_n - \bar{u}\|_{L^2(Q)}$, $\zeta_n = \frac{y_n - \bar{y}}{t_n}$ and $v_n = \frac{u_n - \bar{u}}{t_n}$. Then $t_n \rightarrow 0^+$ and $\|v_n\|_{L^2(Q)} = 1$. By the reflexivity of $L^2(Q)$, we may assume that $v_n \rightharpoonup v$ in $L^2(Q)$. From the above, we have

$$J(y_n, u_n) - J(\bar{y}, \bar{u}) \leq t_n^2 \gamma_n \leq o(t_n^2). \quad (4.2.100)$$

Now we split the proof into several steps.

Step 1. Showing that ζ_n converges weakly to some ζ in $W_2^{1,1}(0, T; D, H)$.

Since $(y_n, u_n) \in \Phi_3$ and $(\bar{y}, \bar{u}) \in \Phi_3$, we have

$$\begin{aligned} \frac{\partial y_n}{\partial t} + Ay_n + f(y_n) &= u_n, \quad y_n(0) = y_0 \\ \frac{\partial \bar{y}}{\partial t} + A\bar{y} + f(\bar{y}) &= \bar{u}, \quad \bar{y}_n(0) = y_0. \end{aligned}$$

This implies that

$$\frac{\partial(y_n - \bar{y})}{\partial t} + A(y_n - \bar{y}) + f(y_n) - f(\bar{y}) = u_n - \bar{u}, \quad (y_n - \bar{y})(0) = 0. \quad (4.2.101)$$

By Taylor's expansion, we have

$$f(y_n) - f(\bar{y}) = f'(\bar{y} + \theta_n(y_n - \bar{y}))(y_n - \bar{y}), \quad 0 \leq \theta_n(x, t) \leq 1.$$

Such a function θ_n does exist and it is measurable. Indeed, we consider the set-valued map $H_n : Q \rightarrow 2^{\mathbb{R}}$ by setting

$$\begin{aligned} H_n(x, t) &= \{\theta \in [0, 1] : f(y_n(x, t)) - f(\bar{y}(x, t)) \\ &= f'(\bar{y}(x, t) + \theta(y_n(x, t) - \bar{y}(x, t)))(y_n(x, t) - \bar{y}(x, t))\}. \end{aligned}$$

By [11, Theorem 8.2.9], H_n is measurable and has a measurable selection, that is, there exists a measurable mapping $\theta_n : Q \rightarrow [0, 1]$ such that $\theta_n(x, t) \in H_n(x, t)$ for a.a. $(x, t) \in Q$.

Note that $y_n \rightarrow \bar{y}$ in $L^\infty(Q)$ so there exists $M > 0$ such that $\|\bar{y} + \theta_n(y_n - \bar{y})\|_{L^\infty(Q)} \leq M$. By (H2), there exists $k_M > 0$ such that

$$|f'(\bar{y} + \theta_n(y_n - \bar{y}))| \leq k_M |\bar{y} + \theta_n(y_n - \bar{y})| + |f'(0)|.$$

Hence $f'(\bar{y} + \theta_n(y_n - \bar{y})) \in L^\infty(Q)$. From (4.2.101), we get

$$\frac{\partial \zeta_n}{\partial t} + A\zeta_n + f'(\bar{y} + \theta_n(y_n - \bar{y}))\zeta_n = v_n, \quad \zeta_n(0) = 0. \quad (4.2.102)$$

By [79, Theorem 5, p.360], there exists a constant $C > 0$ such that

$$\|\zeta_n\|_{W_2^{1,1}(0, T; D, H)} \leq C \|v_n\|_{L^2(Q)} = C.$$

Hence $\{\zeta_n\}$ is bounded in $W_2^{1,1}(0, T; D, H)$. Without loss of generality, we may assume that $\zeta_n \rightharpoonup \zeta$ in $W_2^{1,1}(0, T, D, H)$.

Step 2. Showing that $(\zeta, v) \in \mathcal{C}_2[(\bar{y}, \bar{u})]$.

By Aubin's lemma, the embedding $W_2^{1,1}(0, T; D, H) \hookrightarrow L^2([0, T], V)$ is compact. Therefore, we may assume that $\zeta_n \rightarrow \zeta$ in norm of $L^2([0, T], V)$. Particularly, $\zeta_n \rightarrow \zeta$ in norm of $L^2(Q)$. We now use the procedure in the proof of [79, Theorem 3, p. 356]. By passing to the limit, we obtain from (4.2.102) that

$$\frac{\partial \zeta}{\partial t} + A\zeta + f'(\bar{y})\zeta = v, \quad \zeta(0) = 0. \quad (4.2.103)$$

By a Taylor expansion, we have from (4.2.100) that

$$\int_Q L_y(x, t, \bar{y} + \theta(y_n - \bar{y}), u_n) \zeta_n dx dy + \int_Q L_u(x, t, \bar{y}, \bar{u} + \xi(u_n - \bar{u})) v_n dx dt \leq \frac{o(t_n^2)}{t_n}, \quad (4.2.104)$$

where $0 \leq \theta, \xi \leq 1$. Since $y_n \rightarrow \bar{y}$ and $u_n \rightarrow \bar{u}$ in norm of $L^\infty(Q)$ and (H3), we have

$$\|L_y(\cdot, \cdot, \bar{y} + \theta(y_n - \bar{y}), u_n) - L_y(\cdot, \cdot, \bar{y}, \bar{u})\|_{L^\infty(Q)} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

and

$$\|L_u(\cdot, \cdot, \bar{y}, \bar{u} + \xi(u_n - \bar{u})) - L_u(\cdot, \cdot, \bar{y}, \bar{u})\|_{L^\infty(Q)} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Letting $n \rightarrow \infty$ we deduce from (4.2.104) that

$$\int_Q (L_y[x, t]\zeta(x, t) + L_u[x, t]v(x, t)) dx dt \leq 0. \quad (4.2.105)$$

Hence conditions (c'_1) and (c'_2) of $\mathcal{C}_2[(\bar{y}, \bar{u})]$ are valid. It remains to verify (c'_3) and (c'_4) . Let

$$G(y, u)(x, t) = \left(G_1(y, u)(x, t), G_2(y, u)(x, t) \right),$$

where $G_1(y, u) = u$ and $G_2(y, u) = g(\cdot, \cdot, y) + \epsilon u$. Since $(y_n, u_n) \in \Phi_3$, we have $G_i(y_n, u_n) \in K_i$ with $i = 1, 2$ and

$$\begin{aligned} K_1 &:= \{v \in L^2(Q) : a \leq v(x, t) \leq b \quad \text{a.a. } (x, t) \in Q\}, \\ K_2 &:= \{v \in L^2(Q) : v(x, t) \leq 0 \quad \text{a.a. } (x, t) \in Q\}. \end{aligned}$$

By the Taylor expansion and the definitions of ζ_n, v_n , we have

$$\begin{aligned} G_{iy}(\bar{z})\zeta_n + G_{iu}(\bar{z})v_n + \frac{o(t_n)}{t_n} &\in \frac{1}{t_n}(K_i - G_i(\bar{y}, \bar{u})) \subseteq \text{cone}(K_i - G_i(\bar{y}, \bar{u})) \\ &\subseteq T_{L^2(Q)}(K_i; G_i(\bar{y}, \bar{u})) \end{aligned}$$

where $T_{L^2(Q)}(K_i; G_i(\bar{y}, \bar{u}))$ is the tangent cone to K_i at $G_i(\bar{y}, \bar{u})$ in $L^2(Q)$. Since $T_{L^2}(K_i; G_i(\bar{y}, \bar{u}))$ is a closed convex subset of $L^2(Q)$, it is a weakly closed set of $L^2(Q)$ and so $T_{L^2}(K_i; G_i(\bar{y}, \bar{u}))$ is sequentially weakly closed. By a simple argument, we can show that $G_{iy}(\bar{z})\zeta_n + G_{iu}(\bar{z})v_n \rightharpoonup G_{iy}(\bar{z})\zeta + G_{iu}(\bar{z})v$ in $L^2(Q)$. Letting $n \rightarrow \infty$, we obtain from the above that

$$G_{iy}(\bar{z})\zeta + G_{iu}(\bar{z})v \in T_{L^2}(K_i; G_i(\bar{y}, \bar{u})). \quad (4.2.106)$$

By Lemma 2.4 in [130], we have

$$T_{L^2}(K_1, G_1(\bar{y}, \bar{u})) = \{v \in L^2(Q) : v(x, t) \in T([a, b], \bar{u}(x, t)) \text{ a.a. } (x, t) \in Q\}.$$

and

$$\begin{aligned} T_{L^2}(K_2; G_2(\bar{y}, \bar{u})) = \\ \{v \in L^2(Q) : v(x, t) \in T((-\infty, 0]; g[x, t] + \epsilon \bar{u}(x, t)) \text{ a.a. } (x, t) \in Q\}. \end{aligned}$$

It follows that

$$v(x, t) \in T([a, b]; \bar{u}(x, t)) \quad \text{a.a. } (x, t) \in Q$$

and

$$g_y[x, t]\zeta(x, t) + \epsilon v(x, t) \in T((-\infty, 0]; g[x, t] + \epsilon \bar{u}(x, t)) \quad \text{a.a. } (x, t) \in Q$$

Hence (ζ, v) satisfies (c'_3) and (c'_4) . Consequently, $(\zeta, v) \in \mathcal{C}_2[(\bar{y}, \bar{u})]$ and the claim is justified.

Step 3. Proving that $(\zeta, v) = (0, 0)$.

By assertion (iii) of Theorem 4.2.8, we get

$$\int_Q e(g(y_n) + \epsilon u_n - g(\bar{y}) - \epsilon \bar{u}) dx dt \leq 0 \quad \text{and} \quad \int_Q \widehat{e}(u_n - \bar{u}) dx dt \leq 0.$$

From this and definition of $\mathcal{L}$ in (4.2.95), we have

$$\begin{aligned} & \mathcal{L}(y_n, u_n, \varphi, e, \widehat{e}) - \mathcal{L}(\bar{y}, \bar{u}, \varphi, e, \widehat{e}) \\ & \leq J(y_n, u_n) - J(\bar{y}, \bar{u}) \\ & + \int_Q ((y_n - \bar{y})\left(-\frac{\partial \varphi}{\partial t} + A^* \varphi\right) + (f(y_n) - f(\bar{y}) - (u_n - \bar{u}))\varphi) dx dt \\ & = J(y_n, u_n) - J(\bar{y}, \bar{u}) \\ & + \int_Q \left(\frac{\partial(y_n - \bar{y})}{\partial t} + A(y_n - \bar{y}) + f(y_n) - f(\bar{y}) - (u_n - \bar{u})\right)\varphi dx dt \\ & = J(y_n, u_n) - J(\bar{y}, \bar{u}) \leq o(t_n^2). \end{aligned}$$

Using a Taylor expansion for $\mathcal{L}$ and noting that $D_{(y,u)}\mathcal{L}(\bar{y}, \bar{u}, \varphi, e, \hat{e}) = 0$, we get

$$\begin{aligned}
o(t_n^2) &\geq \mathcal{L}(y_n, u_n, \varphi, e) - \mathcal{L}(\bar{y}, \bar{u}, \varphi, e) = \\
&\int_Q \left(t_n L_y[x, t] \zeta_n + t_n^2 L_{yu}(x, t, \bar{y}, \bar{u} + \theta_2(u_n - \bar{u})) \zeta_n v_n \right) dx dt \\
&+ \int_Q \left(\frac{t_n^2}{2} L_{yy}(x, t, \bar{y} + \theta_3(y_n - \bar{y}), u_n) \zeta_n^2 \right) dx dt \\
&+ \int_Q \left(t_n L_u[x, t] v_n + \frac{t_n^2}{2} L_{uu}(x, t, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) v_n^2 \right) dx dt \\
&+ \int_Q \left[t_n \left(-\varphi_t + A^* \varphi + f'(\bar{y}) \varphi \right) \zeta_n + \frac{t_n^2}{2} \varphi f''(\bar{y} + \theta_4(y_n - \bar{y})) \zeta_n^2 - t_n \varphi v_n \right] dx dt \\
&+ \int_Q e \left(t_n g_y[x, t] \zeta_n + \frac{t_n^2}{2} g_{yy}(x, t, \bar{y} + \theta_5(y_n - \bar{y}), u_n) \zeta_n^2 \right) dx dt \\
&+ \int_Q (t_n \epsilon e v_n + t_n \hat{e} v_n) dx dt \\
&= \frac{t_n^2}{2} \int_Q \left(2L_{yu}(x, t, \bar{y}, \bar{u} + \theta_2(u_n - \bar{u})) \zeta_n v_n + L_{yy}(x, t, \bar{y} + \theta_3(y_n - \bar{y}), u_n) \zeta_n^2 \right) dx dt \\
&+ \frac{t_n^2}{2} \int_Q \left(e g_{yy}(x, t, \bar{y} + \theta_5(y_n - \bar{y}), u_n) + \varphi f''(\bar{y} + \theta_4(y_n - \bar{y})) \right) \zeta_n^2 dx dt \\
&+ \frac{t_n^2}{2} \int_Q L_{uu}(x, t, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) v_n^2 dx dt,
\end{aligned}$$

where $0 \leq \theta_i, \xi_j \leq 1$. This implies that

$$\frac{o(t_n^2)}{t_n^2} \geq \int_Q \left(2L_{yu}(x, t, \bar{y}, \bar{u} + \theta_2(u_n - \bar{u})) \zeta_n v_n \right) dx dt \quad (4.2.107)$$

$$\begin{aligned}
&+ \int_Q \left(L_{yy}(x, t, \bar{y} + \theta_3(y_n - \bar{y}), u_n) \zeta_n^2 \right) dx dt \\
&+ \int_Q \left(e g_{yy}(x, t, \bar{y} + \theta_5(y_n - \bar{y}), u_n) + \varphi f''(\bar{y} + \theta_4(y_n - \bar{y})) \right) \zeta_n^2 dx dt \\
&+ \int_Q L_{uu}(x, t, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) v_n^2 dx dt. \quad (4.2.108)
\end{aligned}$$

Since $\|u_n - \bar{u}\|_{L^\infty(Q)} \rightarrow 0$, $\|y_n - \bar{y}\|_{L^\infty(Q)} \rightarrow 0$ and (H3), we get

$$\|L_{uu}(\cdot, \cdot, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) - L_{uu}[\cdot, \cdot]\|_{L^\infty(Q)} \rightarrow 0.$$

From this and (4.2.97), we see that there exists $n_0 > 0$ such that

$$L_{uu}(x, t, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) > \frac{\Lambda}{2} \quad (4.2.109)$$

for all $n > n_0$ and a.a. $(x, t) \in Q$. Also, by (4.2.97), the functional

$$\xi \mapsto \int_Q L_{uu}[x, t] \xi^2 dx dt$$

is convex and so it is sequentially lower semicontinuous. Therefore, we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_Q L_{uu}(x, t, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) v_n^2 dx dt \\ &= \lim_{n \rightarrow \infty} \int_Q (L_{uu}(x, t, \bar{y}, \bar{u} + \xi_2(u_n - \bar{u})) - L_{uu}[x, t]) v_n^2 dx dt \\ &+ \lim_{n \rightarrow \infty} \int_Q L_{uu}[x, t] v_n^2 dx dt \geq \int_Q L_{uu}[x, t] v^2 dx dt. \end{aligned}$$

From this and letting $n \rightarrow \infty$, we obtain from (4.2.107) that

$$0 \geq \int_Q \left(2L_{yu}[x, t] \zeta v + \left[L_{yy}[x, t] + eg_{yy}[x, t] + \varphi f''(\bar{y}) \right] \zeta^2 + L_{uu}[x, t] v^2 \right) dx dt.$$

Combining this with (4.2.96), we conclude that $(\zeta, v) = (0, 0)$.

Step 4. Showing a contradiction.

From (4.2.107), (4.2.109) and $\|v_n\|_{L^2(Q)} = 1$, we have

$$\begin{aligned} & \frac{o(t_n^2)}{t_n^2} \geq \\ & \int_Q \left(2L_{yu}(x, t, \bar{y}, \bar{u} + \theta_2(u_n - \bar{u})) \zeta_n v_n + L_{yy}(x, t, \bar{y} + \theta_3(y_n - \bar{y}), u_n) \zeta_n^2 \right) dx dt \\ & + \int_Q \left(eg_{yy}(x, t, \bar{y} + \theta_5(y_n - \bar{y}), u_n) + \varphi f''(\bar{y} + \theta_4(y_n - \bar{y})) \right) \zeta_n^2 dx dt + \frac{\Lambda}{2}. \end{aligned}$$

By letting $n \rightarrow \infty$ and using the equality $(\zeta, v) = (0, 0)$, we get $0 \geq \frac{\Lambda}{2}$ which is impossible. The proof of theorem is complete. $\square$

4.2.5 Hölder continuity of multipliers and optimal solutions

In this section we show that under strengthen hypotheses on y_0 , L , g and $\partial\Omega$, then the multipliers and optimal solution are Hölder continuous. We require the following assumptions.

(H3') L is a continuous function and for each $(x, t) \in \Omega \times [0, T]$, $L(x, t, \cdot, \cdot)$ is of class C^2 and satisfies the following property: for each $M > 0$, there

exists $k_{L,M} > 0$ such that

$$\begin{aligned} & |L(x_1, t_1, y_1, u_1) - L(x_2, t_2, y_2, u_2)| + |L_y(x_1, t_1, y_1, u_1) - L_y(x_2, t_2, y_2, u_2)| \\ & + |L_u(x_1, t_1, y_1, u_1) - L_u(x_2, t_2, y_2, u_2)| \\ & \leq k_{L,M}(|x_1 - x_2| + |t_1 - t_2| + |y_1 - y_2| + |u_1 - u_2|), \end{aligned}$$

and

$$\begin{aligned} & |L_{yy}(x, t, y_1, u_1) - L_{yy}(x, t, y_2, u_2)| + |L_{yu}(x, t, y_1, u_1) - L_{yu}(x, t, y_2, u_2)| \\ & + |L_{uu}(x, t, y_1, u_1) - L_{uu}(x, t, y_2, u_2)| \leq k_{L,M}(|y_1 - y_2| + |u_1 - u_2|) \end{aligned}$$

for all $(x, t), (x_i, t_i) \in Q$ and $(y_i, u_i) \in \mathbb{R} \times \mathbb{R}$ satisfying $|y_i|, |u_i| \leq M$ with $i = 1, 2$. g is a continuous function and for each $(x, t) \in \Omega \times [0, T]$, $g(x, t, \cdot)$ is of class C^2 and satisfies the following property: for each $M > 0$, there exists $k_{g,M} > 0$ such that

$$\begin{aligned} & |g(x_1, t_1, y_1) - g(x_2, t_2, y_2)| + |g_y(x_1, t_1, y_1) - g_y(x_2, t_2, y_2)| + \\ & + |g_u(x_1, t_1, y_1) - g_u(x_2, t_2, y_2)| \leq k_{g,M}(|x_1 - x_2| + |t_1 - t_2| + |y_1 - y_2|), \end{aligned}$$

and

$$\begin{aligned} & |g_{yy}(x, t, y_1) - g_{yy}(x, t, y_2)| + |g_{yu}(x, t, y_1) - g_{yu}(x, t, y_2)| + \\ & + |g_{uu}(x, t, y_1) - g_{uu}(x, t, y_2)| \leq k_{g,M}|y_1 - y_2| \end{aligned}$$

for all $(x, t), (x_i, t_i) \in Q$ and $y_i \in \mathbb{R}$ satisfying $|y_i| \leq M$ with $i = 1, 2$.

(H4') $f_y[x, t] \geq -\frac{1}{\epsilon}g_y[x, t] \geq 0$ for a.a. $(x, t) \in Q$ and there exists a number $\gamma > 0$ such that

$$a - \bar{u}(x, t) + g[x, t] + \epsilon \bar{u}(x, t) \leq -\gamma \quad \text{a.a. } (x, t) \in Q. \quad (4.2.110)$$

Moreover, the following condition is satisfied

$$g[x, t] + \epsilon b < 0 \quad \text{on } \{(x, t) \in Q : \bar{u}(x, t) = b\}. \quad (4.2.111)$$

(H5') There exists a number $\gamma_1 > 0$ such that

$$L_{uu}(x, t, y, u) \geq \gamma_1 \quad \forall (x, t, y, u) \in Q \times \mathbb{R} \times \mathbb{R}. \quad (4.2.112)$$

Recall that the boundary $\partial\Omega$ satisfies the property of positive geometric density, that is, there exist number $\alpha^* \in (0, 1)$ and $R_0 > 0$ such that for all $x_0 \in \partial\Omega$ and all $R \leq R_0$, one has

$$|\Omega \cap B(x_0, R)| \leq (1 - \alpha^*)|B(x_0, R)|,$$

where $|C|$ denote Lebesgue measure of a measurable set C in $\mathbb{R}^N$. It is known that if $\partial\Omega$ is of class $C^{1,1}$ or Ω is convex, then this property

is valid. The property of positive geometric density is important for the Hölder continuity of solutions to the semilinear parabolic equations (see [65]).

The following theorem gives Hölder continuity of $(\bar{y}, \bar{u})$ and multipliers.

Theorem 4.2.10. *Suppose $y_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, hypotheses $(H1)$, $H(2)$, $(H3')$, $(H4')$, $(H5')$, $(\bar{y}, \bar{u}) \in \Phi_3$, and $\partial\Omega$ is of class $C^{1,1}$ or Ω is convex. Assume that there exist a function $\varphi \in L^\infty(Q) \cap W_2^{1,1}(0, T; D, H)$ and functions $e, \hat{e} \in L^\infty(Q)$ which satisfy conditions (i) – (iv) of Theorem 4.2.8. Then $\bar{y}, \bar{u}, \varphi, e$ and $\hat{e}$ are Hölder continuous on $\bar{Q}$.*

Proof. By the Sobolev inequality (see [79, Theorem 6, p. 270]) we have $y_0 \in C^{0,\alpha}(\Omega) \cap H_0^1(\Omega)$ for some $\alpha \in (0, 1]$. Since implications $(H3') \Rightarrow (H3)$ and $(H4') \Rightarrow (H4)$ are valid, the existence of multipliers φ and e follows from Theorem 4.2.8. It remains to show that $\bar{y}, \bar{u}, \varphi$ and e are Hölder continuous.

Step 1. Showing that $\bar{y}$ and φ are Hölder continuous functions.

We have $\bar{y} \in L^\infty(Q)$, hence, $\bar{y}$ is as a bounded solution of the parabolic (1.2)-(1.3). Since $\partial\Omega$ satisfies the property of positive geometric density, according to [65, Corollary 0.1], $\bar{y}$ is α -Hölder continuous on $\bar{Q}$ with exponent $\alpha \in (0, 1)$. By setting $\hat{\varphi}(t) = \varphi(T - t)$, then $\hat{\varphi}$ is unique solution of equation

$$\begin{aligned} \frac{\partial \hat{\varphi}(t)}{\partial t} + A^* \hat{\varphi}(t) + f'(\bar{y}(x, T - t)) \hat{\varphi}(t) = \\ - L_y[x, T - t] - e(x, T - t) g_y[x, T - t], \quad \hat{\varphi}(\cdot, 0) = 0. \end{aligned} \quad (4.2.113)$$

By the same argument, we have from [65, Corollary 0.1] that $\hat{\varphi}$ is also α -Hölder continuous on $\bar{Q}$. It follows that φ is α -Hölder continuous.

Step 2. Showing that e and $\hat{e}$ are Hölder continuous.

Let Q_a, Q_b and Q_0 be defined by (4.2.61), (4.2.62) and (4.2.64), respectively. Since $e(x, t) = \hat{e}(x, t) = 0$ on $Q \setminus (Q_a \cup Q_b \cup Q_0)$, we only prove that e and $\hat{e}$ are Hölder continuous on Q_a, Q_b and Q_0 .

• On Q_a we have $\bar{u}(x, t) = a$. By (4.2.110), we have $g[x, t] + \epsilon \bar{u}(x, t) \leq -\gamma < 0$. Hence $e(x, t) = 0$ and so e is Hölder continuous on Q_a . From (ii) of Theorem 4.2.8 we have

$$\hat{e} = \varphi - L_u[\cdot, \cdot]. \quad (4.2.114)$$

By Step 1, φ is Hölder continuous function on Q . Let us claim that $L_u[\cdot, \cdot]$ is Hölder continuous function on Q_a . Indeed, by $(H3')$, for any

$(x_1, t_1), (x_2, t_2) \in Q_a$, we have

$$\begin{aligned} & |L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_1, t_1)) - L_u(x_2, t_2, \bar{y}(x_2, t_2), \bar{u}(x_2, t_2))| \\ &= |L_u(x_1, t_1, \bar{y}(x_1, t_1), a) - L_u(x_2, t_2, \bar{y}(x_2, t_2), a)| \\ &\leq k_{L,M}(|x_1 - x_2| + |t_1 - t_2| + |\bar{y}(x_1, t_1) - \bar{y}(x_2, t_2)|). \end{aligned} \quad (4.2.115)$$

Therefore, $L_u[\cdot, \cdot]$ is Hölder continuous function on Q_a because $\bar{y}$ is Hölder continuous function on Q . The claim is justified. Consequently, Thus $\hat{e}$ is Hölder continuous function on Q_a .

• On Q_b we have $\bar{u}(x, t) = b$. By (4.2.111), we have $g[x, t] + \epsilon \bar{u}(x, t) < 0$ on Q_b . Hence $e(x, t) = 0$ on Q_b and so e is Hölder continuous on Q_b . Again, we have from (ii) of Theorem 4.2.8 that $\hat{e}(x, t) = \varphi(x, t) - L_u[x, t]$ on Q_b . This implies that $\hat{e}$ is Hölder continuous function on Q_b .

• On Q_0 we have $\epsilon \bar{u}(x, t) + g[x, t] = 0$. This and (4.2.110) imply that $\bar{u}(x, t) > a$. By (4.2.111), we get $\bar{u}(x, t) < b$. Hence $\hat{e}(x, t) = 0$ on Q_0 and so it is Hölder continuous on Q_0 . Again, from (ii) of Theorem 4.2.8 we have

$$e(x, t) = \frac{1}{\epsilon}(\varphi(x, t) - L_u[x, t]). \quad (4.2.116)$$

It follows that e is Hölder continuous on Q_0 .

Step 3. Showing that $\bar{u}$ is Hölder continuous on $\bar{Q}$.

Put $\Psi(x, t) := \epsilon e(x, t) + \hat{e}(x, t)$, $(x, t) \in Q$, then from the step 2, $\Psi \in C^{0,\alpha}(\bar{Q})$. Fixing any $(x_i, t_i) \in Q$ with $i = 1, 2$ and using a Taylor' expansion, we have

$$\begin{aligned} & |L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_1, t_1)) - L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_2, t_2))| = \\ & |L_{uu}(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_2, t_2) + \theta(\bar{u}(x_1, t_1) - \bar{u}(x_2, t_2)))|(\bar{u}(x_1, t_1) - \bar{u}(x_2, t_2))| \\ & \geq \gamma_1|(\bar{u}(x_1, t_1) - \bar{u}(x_2, t_2))|, \end{aligned}$$

where $\theta \in [0, 1]$ and the last inequality follows from (H6'). Combining this with equality $L_u[\cdot, \cdot] = \varphi - \Psi$, we get

$$\begin{aligned} & \gamma_1|(\bar{u}(x_1, t_1) - \bar{u}(x_2, t_2))| \\ & \leq |L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_1, t_1)) - L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_2, t_2))| \\ & \leq |L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_1, t_1)) - L_u(x_2, t_2, \bar{y}(x_2, t_2), \bar{u}(x_2, t_2))| \\ & + |L_u(x_2, t_2, \bar{y}(x_2, t_2), \bar{u}(x_2, t_2)) - L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_2, t_2))| \\ & = |\varphi(x_1, t_1) - \varphi(x_2, t_2) - \Psi(x_1, t_1) + \Psi(x_2, t_2)| \\ & + |L_u(x_2, t_2, \bar{y}(x_2, t_2), \bar{u}(x_2, t_2)) - L_u(x_1, t_1, \bar{y}(x_1, t_1), \bar{u}(x_2, t_2))| \\ & \leq |\varphi(x_1, t_1) - \varphi(x_2, t_2)| + |\Psi(x_1, t_1) - \Psi(x_2, t_2)| \\ & + k_{L,M}(|x_1 - x_2| + |t_1 - t_2| + |\bar{y}(x_1, t_1) - \bar{y}(x_2, t_2)|). \end{aligned} \quad (4.2.117)$$

Since φ, Ψ and $\bar{y}$ are Hölder continuous function on Q , the estimation (4.2.117) shows that $\bar{u}$ is Hölder continuous on $\bar{Q}$. The proof of the theorem is complete. $\square$

4.3 Numerical solutions for semilinear elliptic optimal control with nonlinear and mixed constraints

4.3.1 Introduction

In this section we study a numerical method for the following optimal control problem governed by semilinear elliptic equations with nonlinear and mixed pointwise constraints. Let Ω be a convex and bounded domain in $\mathbb{R}^N$ with $N = 2, 3$ with a polygonal or polyhedral boundary Γ . We determine a control function $u \in L^2(\Omega)$ and the corresponding state function $y \in H^2(\Omega) \cap H_0^1(\Omega)$ which solve

$$(P) \quad \begin{cases} J(y, u) = \int_{\Omega} (\phi(x, y(x)) + \frac{1}{2}u^2(x))dx \rightarrow \min, \\ \text{s.t.} \\ Ay + f(x, y) = u \text{ in } \Omega, \quad y = 0 \quad \text{on } \Gamma, \\ g(y(x)) + \epsilon u(x) \leq 0 \text{ a.e. } x \in \Omega, \end{cases}$$

where $\phi, f, g : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ are Carathéodory functions with g is a non-convex function and ϵ is a positive constant, and the operator A is defined by

$$Ay = - \sum_{i,j=1}^N D_j(a_{ij}(x))D_i y$$

with functions $a_{ij} \in C(\bar{\Omega})$ and D_i denotes the partial derivative with respect to x_i . We call (P) the original problem and denote by Φ_4 the feasible set of (P) .

We are going to discretize the original problem by the finite element method (FEM for short) in order to get a sequence of mathematical programming problems (P_h) in finite dimensional spaces, which are called approximate optimal problems. The aim of this paper is to study the convergence of optimal solutions of (P_h) to the optimal solution of problem (P) and to give their error estimates. This topic has been interested

by some mathematicians recently. For papers which have close connections to our work, we refer the readers to [8], [40], [41], [45], [130], [161], [169], [206] and many references therein. In these papers the authors mainly considered problems with pure state constraints or with pure control constraints. Among these papers, Casas and Mateos [40] studied convergence for semilinear elliptic optimal control with pure state constraint; Casas and Tröltzsch [45] proved the convergence of optimal control problem governed by quasilinear elliptic equations with the control constraint of box form; Meyer and Rösch [18] gave superconvergence properties of approximation solutions for linear elliptic optimal control problems with control constraint of box form.

It is natural to ask whether these results are still true for problem (P) with the appearance of nonlinear, non-convex and mixed pointwise constraints. One of the motivations for considering optimal control problems with mixed control-state constraints is the Lavrentiev regularization which frequently is used to relax pure state constraints. In [104], M. Hinze and Meyer showed that under certain conditions, the locally optimal solutions to the Lavrentiev regularized problems converge to a optimal solution of the original problem.

There have been some papers working with optimal control problems with mixed constraints so far. Indeed, in [54], Cherednichenko and Rösch gave error estimates for the discretization of linear elliptic optimal control problems with mixed pointwise constraints in the form of linear inequalities. in [104], Hinze and Meyer considered linear elliptic optimal control problem with constraints of linear inequalities and gave error estimates for approximate solution.

Rösch and Wachsmuth [205] considered a class of optimal control problems for a linear parabolic partial differential equation with control and mixed control-state constraints in the form of linear inequalities. Also, Hoppe and Kieweg [106] studied adaptive finite element methods for optimal control problems governed by linear elliptic equations with mixed control-state constraints of linear inequalities. In a summary, the problems considered by [54], [104], [106] and [205] have the following common features: the cost functions are of quadratic forms which are strongly convex functions; the state equations are linear and the mixed constraints are of linear inequalities. In such settings, they are strongly convex optimization problems. Therefore, they have a unique solution. Moreover, when the constraints are linear, the approximate optimal problems have finitely many constraints.

In this subsection we shall develop the results obtained in [54] for problem (P) which is more general. We shall prove that under certain conditions, the optimal solutions of discrete problems converge to an optimal solution of the original problem. In particular, we show that if problem (P) satisfies the so-called no-gap second-order conditions, then some error estimates of approximate solutions are obtained.

Unlike optimal control problem with pure control constraints, the optimal control problem with nonlinear mixed control-state constraints is more complicated. The reason is that the discrete optimal problems may have infinitely many constraints. Besides, it is difficult to build a sequence of admissible points of discrete optimal problems, which converges strongly to an optimal solution of the original problem. To overcome these difficulties, we choose piecewise linear functions in order to approximate both control variable and state variable. We first transform (P) to problem (P') with pure control constraint by changing variable $v = g(y) + \epsilon u$. We then discretize (P') by (P'_h) . After that, we establish error estimates between solutions of state equations as well as solutions of adjoint equations to (P') and (P'_h) . Based on the the results and some techniques from [46], [54] and [104], we show that the sequence of optimal solutions $(\bar{y}_h, \bar{v}_h)$ of (P'_h) converges strongly to an optimal solution $(\bar{y}, \bar{v})$ of (P') which implies that $(\bar{y}_h, \frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h)))$ converges strongly to an optimal solution $(\bar{y}, \frac{1}{\epsilon}(\bar{v} - g(\bar{y})))$ of (P) . Moreover, we show that if (P) satisfies the strictly second-order condition, then some error estimates are obtained.

Thus the novelty of the subsection consists of results on the convergence and error estimates of approximate optimal solutions of approximate optimal problems which are generated by the semilinear elliptic optimal control problem with non-convex cost function and nonlinear inequality constraints via FEM.

4.3.2 Discrete problems and optimality conditions

Let E be a Banach space with the dual E^* . The notation $B_E(e, r)$ stands for an open ball with center $e \in E$ and radius $r > 0$. Let K be a closed set in E .

In the sequel, we define $U = L^2(\Omega)$, $Y = H^2(\Omega) \cap H_0^1(\Omega)$ and $Z = Y \times U$. The norm in $L^p(\Omega)$ with $1 \leq p \leq +\infty$ will be denoted by $\|\cdot\|_p$. Meanwhile $\|\cdot\|_{2,2}$ and $\|\cdot\|_{1,2}$ denote norms in $H^2(\Omega)$ and $H_0^1(\Omega)$, respectively. The norm $\|\cdot\|_Z$ in Z is defined by $\|(y, u)\|_Z = \|y\|_{2,2} + \|u\|_2$. The

notations $\phi_y, \phi_{yy}, f_y, f_{yy}$ are used to denote $\partial_y \phi, \partial_{yy}^2 \phi, \partial_y f, \partial_{yy}^2 f$, respectively. Given an admissible couple $(\bar{y}, \bar{u}) \in \Phi_4$, symbols $\phi[x], \phi_y[x], f[\cdot]$, etc., stand respectively for $\phi(x, \bar{y}(x)), \phi_y(x, \bar{y}(x)), f(\cdot, \bar{y}(\cdot))$, etc. We now make the following assumptions for (P) .

- (A1) $\phi : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable and satisfies the property that for each $M > 0$, there exist positive numbers $k_{\phi, M}$ and $L_{\phi, M}$ such that

$$\begin{aligned} |\phi_y(x, y)| &\leq L_{\phi, M}, \\ |\phi(x, y_1) - \phi(x, y_2)| + |\phi_y(x, y_1) - \phi_y(x, y_2)| + \\ &+ |\phi_{yy}(x, y_1) - \phi_{yy}(x, y_2)| \leq k_{\phi, M} |y_1 - y_2| \end{aligned}$$

for a.e. $x \in \Omega$, all $y, y_i \in \mathbb{R}$ satisfying $|y|, |y_i| \leq M$ with $i = 1, 2$. Moreover, $\phi(\cdot, 0) \in L^1(\Omega)$ and there exists a function $\zeta \in L^1(\Omega)$ such that

$$\phi(x, y) \geq \zeta(x) \quad \forall (x, y) \in \Omega \times \mathbb{R}.$$

- (A2) $a_{ij} = a_{ji} \in C^{0,1}(\bar{\Omega})$ and there exists a number $\bar{\alpha} > 0$ such that

$$\sum_{i,j=1}^N a_{ij}(x) \xi_i \xi_j \geq \bar{\alpha} |\xi|^2 \quad \forall \xi = (\xi_1, \dots, \xi_N) \in \mathbb{R}^N.$$

- (A3) $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a Carathéodory functions of class C^2 with respect to the second variable and satisfies the property that for each $M > 0$, there exist positive numbers $k_{f, M}$ and $L_{f, M}$ such that

$$\begin{aligned} |f_y(x, y)| &\leq L_{f, M}, \\ |f(x, y_1) - f(x, y_2)| + |f_y(x, y_1) - f_y(x, y_2)| + \\ &+ |f_{yy}(x, y_1) - f_{yy}(x, y_2)| \leq k_{f, M} |y_1 - y_2| \end{aligned}$$

for a.e. $x \in \Omega$, all $y, y_i \in \mathbb{R}$ satisfying $|y|, |y_i| \leq M$ with $i = 1, 2$. Besides, we require that $f_y(x, y) \geq 0$ for all $(x, y) \in \Omega \times \mathbb{R}$ and $f(x, 0) = 0$ for a.e. $x \in \Omega$.

- (A4) $g : \mathbb{R} \rightarrow \mathbb{R}$ is a function of class C^2 and satisfies the property that for each $M > 0$, there exist a positive number $k_{g, M}$ such that

$$|g(y_1) - g(y_2)| + |g'(y_1) - g'(y_2)| + |g''(y_1) - g''(y_2)| \leq k_{g, M} |y_1 - y_2|$$

for all $y, y_i \in \mathbb{R}$ satisfying $|y|, |y_i| \leq M$ with $i = 1, 2$. Furthermore, we require that $g(0) \leq 0$ and $g'(y) \geq 0$ for all $y \in \mathbb{R}$.

Among the above conditions, (A3) assures that the state equation has a unique solution for each $u \in U$. Meanwhile (A4) makes sure that the so-called Robinson constraint qualification is valid and the Lagrange function corresponding to (P) is normal. Moreover, (A4) implies that for each $M > 0$ there exist a positive number $L_{g,M}$ such that $|g(y)| + |g'(y)| \leq L_{g,M}$ for all $y \in \mathbb{R}$ satisfying $|y| \leq M$.

Recall that a couple $\bar{z} = (\bar{y}, \bar{u}) \in \Phi_4$ is said to be a local minimizer of (P) if there exist a ball $B_Z(\bar{z}, r) \subset Z$ such that

$$J(y, u) \geq J(\bar{y}, \bar{u}) \quad \forall (y, u) \in \Phi_4 \cap B_Z(\bar{z}, r).$$

When

$$J(y, u) \geq J(\bar{y}, \bar{u}) \quad \forall (y, u) \in \Phi_4$$

we say $(\bar{y}, \bar{u})$ is a global minimizer or a minimizer to (P).

Proposition 4.3.1. *Suppose that assumptions (A2) and (A3) are valid. Then for each $u \in L^2(\Omega)$, there exists a unique solution $y = y(u) \in H_0^1(\Omega) \cap H^2(\Omega)$ of the equation*

$$Ay + f(x, y) = u \text{ in } \Omega, \quad y = 0 \quad \text{on } \Gamma \quad (4.3.1)$$

and there exists a constant $C_0 = C(N, \Omega) > 0$ such that

$$\|y\|_{2,2} \leq C_0 \|u\|_2. \quad (4.3.2)$$

Moreover, if $\|u_1\|_2 + \|u_2\|_2 \leq M$ and $y(u_1)$ and $y(u_2)$ are solutions corresponding to u_1 and u_2 , then there exists a constant $C = C(M, N, \Omega) > 0$ such that

$$\|y(u_1) - y(u_2)\|_{2,2} \leq C \|u_1 - u_2\|_2. \quad (4.3.3)$$

Proof. The proof of the proposition can be found in [204] and Theorem 4.7 in [219]. $\square$

Remark 4.3.2. If $f(x, 0) \neq 0$, then we can replace $f(x, y)$ by $f(x, y) - f(x, 0)$ and (4.3.2) becomes

$$\|y\|_{2,2} \leq C_0 \|u - f(\cdot, 0)\|_2.$$

By a standard argument, we have the following result on the existence of optimal solutions to (P).

Proposition 4.3.3. *Suppose that (A1)–(A4) are fulfilled. Then problem (P) has at least an optimal solution $(\bar{y}, \bar{u})$.*

Let us define mappings $F : Y \times U \rightarrow U$ and $G : Y \times U \rightarrow U$ by setting

$$F(y, u) = Ay + f(\cdot, y) - u, \quad G(y, u) = g(y) + \epsilon u.$$

Then (P) can be reformulated in the form

$$(P) \quad \begin{cases} J(y, u) \rightarrow \min, \\ \text{s.t.} \\ F(y, u) = 0, \\ G(y, u) \in K, \end{cases}$$

where $K := \{v \in L^2(\Omega) | v(x) \leq 0 \text{ a.e. } x \in \Omega\}$. Let

$$\mathcal{L}(y, u, \varphi, e) := J(y, u) + \langle \varphi, F(y, u) \rangle + \langle e, G(y, u) \rangle$$

be the Lagrange function associated to (P) , where $\varphi, e \in U$. Under assumptions (A1) – (A4), J , F and G are of class C^2 around $(\bar{y}, \bar{u}) \in \Phi_4$ and so is $\mathcal{L}(\cdot, \cdot, \varphi, e)$.

Recall that given a vector $(\bar{y}, \bar{u}) \in \Phi_4$, the critical cone to (P) at $(\bar{y}, \bar{u})$ is denoted by $\mathcal{C}[(\bar{y}, \bar{u})]$ which consists of vectors (z, v) satisfying the following conditions:

$$(c_1) \quad \langle \nabla J(\bar{y}, \bar{u}), (z, v) \rangle = \int_{\Omega} (\phi_y[x]z(x) + \bar{u}(x)v(x))dx \leq 0;$$

$$(c_2) \quad Az + f_y[x]z = v \text{ in } \Omega \text{ and } z = 0 \text{ on } \Gamma;$$

$$(c_3) \quad g'(\bar{y}(x))z(x) + \epsilon v(x) \leq 0 \text{ whenever } g(\bar{y}(x)) + \epsilon \bar{u}(x) = 0.$$

Each vector $(z, v) \in \mathcal{C}[(\bar{y}, \bar{u})]$ is called a critical direction. Note that (c_3) is equivalent to saying that

$$g'(\bar{y}(x))z(x) + \epsilon v(x) \in T((-\infty, 0], g(\bar{y}(x)) + \epsilon \bar{u}(x)).$$

By a simple calculation, we have $T((-\infty, 0], \xi) = \mathbb{R}$ if $\xi < 0$ and $T((-\infty, 0], \xi) = \mathbb{R}_-$ if $\xi = 0$.

The following proposition gives first- and second-order optimality conditions for locally optimal solutions to (P) .

Proposition 4.3.4. *Suppose that (A1) – (A4) are satisfied and $(\bar{y}, \bar{u})$ is a local minimizer of (P) . Then there exist a unique function $\bar{\varphi} \in Y$ and a unique function $\bar{e} \in U$ such that the following conditions are fulfilled:*

(i) *(the adjoint equation)*

$$\begin{cases} A^* \bar{\varphi} + f_y(\cdot, \bar{y}) \bar{\varphi} = -\phi_y(\cdot, \bar{y}) - g'(\bar{y}) \bar{e} & \text{in } \Omega \\ \bar{\varphi} = 0 & \text{on } \Gamma; \end{cases} \quad (4.3.4)$$

(ii) (the stationary condition in u)

$$\bar{u} - \bar{\varphi} + \epsilon \bar{e} = 0; \quad (4.3.5)$$

(iii) (the complementarity condition) For a.e. $x \in \Omega$ one has

$$\bar{e}(x) \begin{cases} \geq 0 & \text{if } g(\bar{y}(x)) + \epsilon \bar{u}(x) = 0, \\ = 0 & \text{if } g(\bar{y}(x)) + \epsilon \bar{u}(x) < 0; \end{cases} \quad (4.3.6)$$

(iv) (the non-negative second-order condition)

$$\begin{aligned} & D^2 \mathcal{L}(\bar{y}, \bar{u}, \bar{\varphi}, \bar{e})[(z, v), (z, v)] \\ &= \int_{\Omega} [(\phi_{yy}[x] + f_{yy}[x] \bar{\varphi}(x) + \bar{e}(x) g''(\bar{y}(x))) z^2(x) + v(x)^2] dx \geq 0 \end{aligned} \quad (4.3.7)$$

for all $(z, v) \in \mathcal{C}[(\bar{y}, \bar{u})]$. Moreover, $\bar{y}, \bar{\varphi}, \bar{e}$ and $\bar{u}$ are of class $C^{0,\nu}$ with $\nu = 2 - N/2$, and $\bar{u} \in H^1(\Omega)$.

Conversely, suppose that there exist $\bar{\varphi} \in Y$ and $\bar{e} \in U$ such that conditions (i) – (iii) are fulfilled and the strictly second-order condition (SSC):

$$D^2 \mathcal{L}(\bar{y}, \bar{u}, \bar{\varphi}, \bar{e})[(z, v), (z, v)] > 0 \quad \forall (z, v) \in \mathcal{C}[(\bar{y}, \bar{u})] \setminus \{(0, 0)\}, \quad (4.3.8)$$

then $(\bar{y}, \bar{u})$ satisfies the quadratic growth condition in the sense that there exist a number $\alpha_0 > 0$, a ball $B(\bar{y}, r) \subset H_0^1(\Omega) \cap H^2(\Omega)$ and a ball $B(\bar{u}, r') \subset L^2(\Omega)$ such that

$$J(y, u) \geq J(\bar{y}, \bar{u}) + \alpha_0 \|u - \bar{u}\|_2^2 \quad \forall (y, u) \in (B(\bar{y}, r) \times B(\bar{u}, r')) \cap \Phi_4. \quad (4.3.9)$$

In this case we say $(\bar{y}, \bar{u})$ is a locally strong solution to (P).

Proof. The optimality conditions follow from Theorem 2.1 and Theorem 3.1 in [130]. Let us show that $\bar{u}, \bar{e} \in C^{0,\nu}(\bar{\Omega})$ with $\nu = 2 - \frac{N}{2}$. Since $\bar{y}, \bar{\varphi} \in H^2(\Omega)$ and the embedding theorem, we have $\bar{y}, \bar{\varphi} \in C^{0,\nu}(\bar{\Omega})$. From (4.3.5) and (4.3.6) we get

$$\bar{\varphi}(x) - \bar{u}(x) \in N((-\infty, 0], g(\bar{y}(x)) + \epsilon \bar{u}(x)) \quad \text{a.e. } x \in \Omega.$$

It follows that

$$(\epsilon \bar{\varphi}(x) + g(\bar{y}(x))) - (g(\bar{y}(x)) + \epsilon \bar{u}(x)) \in N((-\infty, 0], g(\bar{y}(x)) + \epsilon \bar{u}(x)) \quad \text{a.e. } x \in \Omega,$$

which is equivalent to

$$g(\bar{y}(x)) + \epsilon \bar{u}(x) = P_{(-\infty, 0]}(\epsilon \bar{\varphi}(x) + g(\bar{y}(x))) \quad \text{a.e. } x \in \Omega \quad (4.3.10)$$

or

$$u(x) = \frac{1}{\epsilon} [P_{(-\infty, 0]}(\epsilon \bar{\varphi}(x) + g(\bar{y}(x))) - g(\bar{y}(x))] \quad \text{a.e. } x \in \Omega, \quad (4.3.11)$$

where $P_{K'}(a)$ denotes the metric projection of a onto a closed convex set K' in a finite-dimensional space. Here $K' = (-\infty, 0]$ and $P_{(-\infty, 0]}(\alpha) = \min\{0, \alpha\}$. Since g is of class C^1 and locally Lipschitz continuous, and $\bar{y} \in H^2(\Omega)$, we have $g(\bar{y}) \in H^1(\Omega)$ (see Theorem 3.1 in [171]). Also, the metric projection is Lipschitz continuous (see Theorem 5.2 in [31]). Hence we conclude that $\bar{u} \in H^1(\Omega)$. We now choose $M > 0$ such that $\|\bar{y}\|_{C(\bar{\Omega})} \leq M$. By assumption (A4), there is a number $k_{g,M} > 0$ such that

$$|g(\bar{y}(x_1)) - g(\bar{y}(x_2))| \leq k_{g,M} |\bar{y}(x_1) - \bar{y}(x_2)| \quad \text{for all } x_1, x_2 \in \Omega.$$

Using non-expansive property of the metric projection and the fact $0 = P_{(-\infty, 0]}(0)$, we have from (4.3.11) that

$$\begin{aligned} \epsilon |\bar{u}(x_1) - \bar{u}(x_2)| &\leq |g(\bar{y}(x_1)) - g(\bar{y}(x_2))| \\ &\quad + |P_{(-\infty, 0]}(\epsilon \bar{\varphi}(x_1) + g(\bar{y}(x_1))) - P_{(-\infty, 0]}(\epsilon \bar{\varphi}(x_2) + g(\bar{y}(x_2)))| \\ &\leq 2|g(\bar{y}(x_1)) - g(\bar{y}(x_2))| + \epsilon |\bar{\varphi}(x_1) - \bar{\varphi}(x_2)| \\ &\leq 2k_{g,M} (|\bar{y}(x_1) - \bar{y}(x_2)|) + \epsilon |\bar{\varphi}(x_1) - \bar{\varphi}(x_2)|. \end{aligned}$$

This implies that $\bar{u} \in C^{0,\nu}(\bar{\Omega})$. Since $\bar{e} = \frac{1}{\epsilon}(\bar{\varphi} - \bar{u})$, we also have $\bar{e} \in C^{0,\nu}(\bar{\Omega})$. $\square$

Next we will set up approximate optimal control problems to problem (P) by using the finite element method for semilinear elliptic equation. Let $T_h = \{K_1, K_2, \dots, K_{q_h}\}$ be a triangulation of Ω , i.e.,

$$\bar{\Omega} = \bigcup_{i=1}^{q_h} K_i,$$

where

$$K_i \cap K_j = \emptyset, \text{ one point, one common side or one common face } \forall i \neq j.$$

For each $K \in T_h$, we shall denote by $\rho(K)$ the diameter of K and $\sigma(K)$ by the diameter of the largest ball contained in K . Let $h = \max_{K \in T_h} \rho(K)$. Throughout of this paper, we assume that there exist positive numbers γ_1 and γ_2 such that

$$\frac{\rho(K)}{\sigma(K)} \leq \gamma_1, \quad \frac{h}{\rho(K)} \leq \gamma_2 \quad \forall K \in T_h.$$

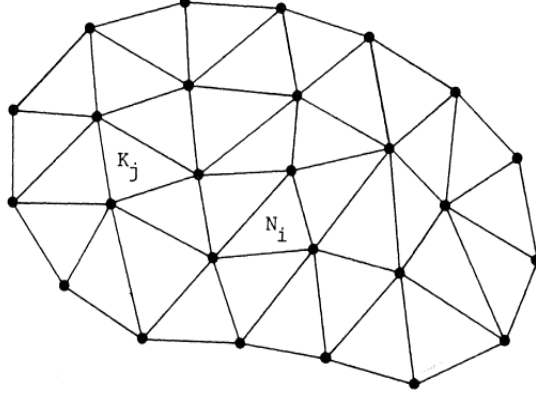


Figure 4.1: Finite element triangulation

We define finite element spaces

$$Y_h = \{y_h \in C(\bar{\Omega}) : y_h|_T \text{ is affine for all } T \in T_h, y_h = 0 \text{ on } \Gamma\}, \quad (4.3.12)$$

$$U_h = \{u_h \in C(\bar{\Omega}) : u_h|_T \text{ is affine for all } T \in T_h\}. \quad (4.3.13)$$

Here we choose the type of piecewise linear control approximation because it is frequently used in numerical computations and its implementation is fairly easy. For the other control approximations, we refer the readers to [46], [103] and [104].

Fix $h > 0$ and let $T_h = \{K_1, K_2, \dots, K_q\}$ be a triangulation of Ω . Let $N_1, N_2, \dots, N_{m_h}, N_{m_h+1}, \dots, N_{k_h}$ be nodes of T_h , where $\{N_1, N_2, \dots, N_{m_h}\} \subset \Omega$ and $\{N_{m_h+1}, \dots, N_{k_h}\} \subset \Gamma$. Let $\varphi_i \in Y_h$ be functions which are defined by

$$\varphi_i(N_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases} \quad (4.3.14)$$

where $i, j \in \{1, 2, \dots, k_h\}$. Then $\{\varphi_1, \dots, \varphi_{m_h}\}$ forms a basis of Y_h and $\{\varphi_1, \dots, \varphi_{m_h}, \varphi_{m_h+1}, \dots, \varphi_{k_h}\}$ forms a basis of U_h . Under these bases, each $y_h \in Y_h$ and $v_h \in U_h$ can be represented by

$$\begin{aligned} y_h(x) &= \alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) + \dots + \alpha_{m_h} \varphi_{m_h}(x), \\ v_h(x) &= \beta_1 \varphi_1(x) + \beta_2 \varphi_2(x) + \dots + \beta_{m_h} \varphi_{m_h}(x) + \dots + \beta_{k_h} \varphi_{k_h}(x). \end{aligned}$$

We can verify that $\alpha_i = y_h(N_i)$ with $i = 1, 2, \dots, m_h$ and $\beta_j = u_h(N_j)$ with $j = 1, 2, \dots, k_h$. Therefore, we can write $y_h = (\alpha_1, \alpha_2, \dots, \alpha_{m_h})$ and $v_h = (\beta_1, \beta_2, \dots, \beta_{k_h})$.

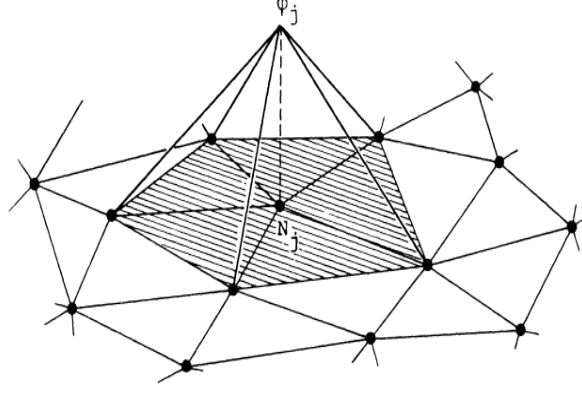


Figure 4.2: The basis functions φ_j

We first consider the following discrete version (P_h) of (P) .

$$\left\{ \begin{array}{l} J(y_h, u_h) = \int_{\Omega} (\phi(x, y_h(x)) + \frac{1}{2}u_h^2(x))dx \rightarrow \min, \\ \text{s.t. } (y_h, u_h) \in Y_h \times U_h, \\ \int_{\Omega} (\sum_{i,j=1}^N a_{ij}(x) D_i y_h D_j z_h + f(x, y_h) z_h(x)) dx \\ = \int_{\Omega} u_h(x) z_h(x) dx \quad \forall z_h \in Y_h, \\ g(y_h(x)) + \epsilon u_h(x) \leq 0 \quad \text{for all } x \in \Omega. \end{array} \right.$$

Obviously, (P_h) is a mathematical programming problem in finite dimensional space $Y_h \times U_h$. The following proposition showing that the pointwise constraint of (P_h) is actually equivalent to finitely many constraints.

Proposition 4.3.5. *If g is convex, then the constraint*

$$g(y_h(x)) + \epsilon u_h(x) \leq 0 \text{ for all } x \in \Omega \quad (4.3.15)$$

is equivalent to the constraints

$$g(\alpha_i) + \epsilon \beta_i \leq 0 \quad \text{for } i = 1, 2, \dots, k_h, \quad (4.3.16)$$

where $\alpha_i = 0$ for $i = m_h + 1, \dots, k_h$.

Proof. By taking $x \in \{N_1, N_2, \dots, N_{k_h}\}$, we see that that (4.3.15) implies (4.3.16). Conversely, for simplification, we consider the case $\Omega \subset \mathbb{R}^2$. Then each triangle is a convex hull of its three vertices. Take any $x \in \Omega$. Then there exists $T \in T_h$ such that $x \in T$. We can assume that $T = N_1 N_2 N_3$. Then $x = \lambda_1 N_1 + \lambda_2 N_2 + \lambda_3 N_3$ with $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $\lambda_i \geq 0$ with $i = 1, 2, 3$. Since y_h and u_h are affine on T and g is convex,

we have

$$\begin{aligned}
g(y_h(x)) + \epsilon u_h(x) &= g(\lambda_1 y_h(N_1) + \lambda_2 y_h(N_2) + \\
&\quad + \lambda_3 y_h(N_3)) + \epsilon (\lambda_1 u_h(N_1) + \lambda_2 u_h(N_2) + \lambda_3 u_h(N_3)) \\
&\leq \sum_{i=1}^3 \lambda_i (g(y_h(N_i)) + \epsilon u_h(N_i)) = \sum_{i=1}^3 \lambda_i (g(\alpha_i) + \epsilon \beta_i) \leq 0.
\end{aligned}$$

The proposition is proved. $\square$

From the above result, we see that if g is convex, then (P_h) is a really mathematical programming with finitely many constraints on finite dimensional spaces $Y_h \times U_h$. However, when g is not convex, the point-wise constraint of (P_h) may have infinitely many constraints. This causes some difficulties in obtaining error estimates for solutions of (P) and (P_h) . To overcome the difficulties, we need to transform (P) to problem (P') with pure control constraint and then discretize it by (P'_h) . In the sequel, instead of working directly with problems (P) and (P_h) , we shall concentrate on (P') and (P'_h) and give error estimate for solutions of (P') and (P'_h) .

Let us change the control variable by setting $v = g(y) + \epsilon u$. Then (P) is transformed to the following problem.

$$(P') \quad \begin{cases} J_0(y, v) = \int_{\Omega} [\phi(x, y(x)) + \frac{1}{2}(\frac{1}{\epsilon}(v(x) - g(y(x))))^2] dx \rightarrow \min, \\ \text{s.t.} \\ Ay + f(x, y) + \frac{1}{\epsilon}g(y) = \frac{1}{\epsilon}v \text{ in } \Omega, \quad y = 0 \quad \text{on } \Gamma, \\ v(x) \leq 0 \text{ a.e. } x \in \Omega, \end{cases}$$

Let us define

$$L(x, y, v) = \phi(x, y) + \frac{1}{2}\left(\frac{1}{\epsilon}(v - g(y))\right)^2 \quad (4.3.17)$$

and denote by Φ' the feasible set of (P') . Then we have

$$L_y[x] = \phi_y[x] - \frac{1}{\epsilon^2}(\bar{v} - g[x])g'[x], \quad (4.3.18)$$

$$L_{yy}[x] = \phi_{yy}[x] + \frac{1}{\epsilon^2}(g'[x]^2 + g[x]g''[x] - \bar{v}(x)g''[x]) \quad (4.3.19)$$

$$L_v[x] = \frac{1}{\epsilon^2}(\bar{v} - g[x]), \quad L_{vv}[x] = \frac{1}{\epsilon^2}, \quad (4.3.20)$$

$$L_{yv} = L_{vy} = -\frac{1}{\epsilon^2}g'[x]. \quad (4.3.21)$$

By the monotonicity of g and $f(x, \cdot)$, and Proposition 4.3.1, for each $v \in L^2(\Omega)$, the equation

$$Ay + f(x, y) + \frac{1}{\epsilon}g(y) = \frac{1}{\epsilon}v \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma \quad (4.3.22)$$

has a unique solution $y = y(v) \in H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$\|y(v)\|_{2,2} \leq C \left\| \frac{1}{\epsilon}v - \frac{1}{\epsilon}g(0) \right\|_2 \quad (4.3.23)$$

and

$$\|y(v_1) - y(v_2)\|_{2,2} \leq C \|v_1 - v_2\|_2 \quad (4.3.24)$$

for all $v_1, v_2 \in L^2(\Omega)$. Let us define $S(v) = y(v)$. Then we can show that $S(v)$ is of class C^2 . By Lemma 2.2 in [106] for each $v_0, v \in L^2(\Omega)$, $z_{v_0,v} := S'(v_0)v$ is a unique solution of the equation

$$Az_{v_0,v} + (f_y(\cdot, y(v_0)) + \frac{1}{\epsilon}g'(y(v_0)))z_{v_0,v} = \frac{1}{\epsilon}v \text{ in } \Omega, \quad z_{v_0,v} = 0 \text{ on } \Gamma. \quad (4.3.25)$$

Moreover, for all $v_1, v_2 \in L^2(\Omega)$, $z_{v_0,v_1,v_2} := S''(v_0)[v_1, v_2]$ is the unique solution of the equation

$$\begin{cases} Az_{v_0,v_1,v_2} + (f_y(\cdot, y(v_0)) + \frac{1}{\epsilon}g'(y(v_0)))z_{v_0,v_1,v_2} + (f_{yy}(\cdot, y(v_0)) + \frac{1}{\epsilon}g''(y(v_0)))z_{v_0,v_1}z_{v_0,v_2} = 0 \text{ in } \Omega \\ z_{v_0,v_1,v_2} = 0 \text{ on } \Gamma. \end{cases} \quad (4.3.26)$$

By using the mapping $S(v)$, problem (P') can be formulated in the following simple form:

$$(P'') \quad \begin{cases} \hat{J}(v) = J_0(y(v), v) = \int_{\Omega} L(x, y(v)(x), v(x)) dx \rightarrow \min, \\ \text{s.t. } v \in U, \\ v(x) \leq 0 \text{ a.e. } x \in \Omega. \end{cases}$$

We shall denote by Φ'' the feasible set of (P'') , that is,

$$\Phi'' = \{v \in L^2(\Omega) | v(x) \leq 0 \text{ a.e.}\}.$$

Proposition 4.3.6. *If $(\bar{y}, \bar{u}) \in \Phi_4$ is a local minimizer to (P) , then $(\bar{y}, \bar{v})$ with $\bar{v} := g(\bar{y}) + \epsilon \bar{u}$ is a local minimizer to (P') and $\bar{v}$ is a local minimizer to (P'') . Conversely, if $\bar{v}$ is a local minimizer to (P'') , then $(\bar{y}, \bar{v})$ with $\bar{y} = S(\bar{v})$ is a local minimizer to (P') and $(\bar{y}, \bar{u})$ with $\bar{u} := \frac{1}{\epsilon}(\bar{v} - g(\bar{y}))$ is a local minimizer to (P) .*

Proof. Assume that $(\bar{y}, \bar{u})$ is a local minimizer to (P) . Then there exists $\delta > 0$ such that for all $(y, u) \in \Phi_4$ satisfying $\|(y, u) - (\bar{y}, \bar{u})\| \leq \delta$ one has $J(y, u) \geq J(\bar{y}, \bar{u})$. Putting $\bar{v} = g(\bar{y}) + \epsilon \bar{u}$, we see that $(\bar{y}, \bar{v}) \in \Phi'$. Let C be the constant so that $\|y\|_\infty \leq C\|y\|_{2,2}$ for all $y \in H^2(\Omega)$. Put $M = C(\|\bar{y}\|_{2,2} + 1)$. Then (A4) implies that there exists $k_{g,M} > 0$ such that

$$|g(y(x)) - g(\bar{y}(x))| \leq k_{g,M}|y(x) - \bar{y}(x)| \leq k_{g,M}\|y - \bar{y}\|_\infty.$$

for all $y \in H^2(\Omega)$ satisfying $\|y - \bar{y}\|_{2,2} < \delta' < 1$. This is because

$$\|y\|_\infty \leq C\|y\|_{2,2} \leq C(\|y - \bar{y}\|_{2,2} + \|\bar{y}\|_{2,2}) \leq C(\delta' + \|\bar{y}\|_{2,2}) < M.$$

Let $\delta'' < \min(\delta', \frac{\epsilon\delta}{\epsilon + k_{g,M} + 1})$. Taking any $(y, v) \in \Phi' \cap B_Z((\bar{y}, \bar{v}), \delta'')$ and putting $u = \frac{1}{\epsilon}(v - g(y))$, we have $(y, u) \in \Phi_4$ and

$$\begin{aligned} \|(y, u) - (\bar{y}, \bar{u})\| &= \|y - \bar{y}\|_{2,2} + \|u - \bar{u}\|_2 \\ &= \|y - \bar{y}\|_{2,2} + \frac{1}{\epsilon}\|v - g(y) - \bar{v} + g(\bar{y})\|_2 \\ &\leq \|y - \bar{y}\|_{2,2} + \frac{k_{g,M}}{\epsilon}\|y - \bar{y}\|_2 + \frac{1}{\epsilon}\|v - \bar{v}\|_2 \\ &\leq (1 + \frac{k_{g,M}}{\epsilon})\|y - \bar{y}\|_{2,2} + \frac{1}{\epsilon}\|v - \bar{v}\|_2 \\ &\leq (\frac{\epsilon + k_{g,M} + 1}{\epsilon})(\|y - \bar{y}\|_{2,2} + \|v - \bar{v}\|_2) < \delta. \end{aligned}$$

Hence $J_0(y, v) = J(y, u) \geq J(\bar{y}, \bar{u}) = J_0(\bar{y}, \bar{v})$. Hence $(\bar{y}, \bar{v})$ is a local minimizer to (P') . Let us choose $r < \frac{\delta''}{1+C}$, where $C > 0$ is a constant which is given by (4.3.24). Then for all $v \in B_U(\bar{v}, r) \cap \Phi''$, we have $(y(v), v) \in \Phi'$ and

$$\begin{aligned} \|(y(v), v) - (y(\bar{v}), \bar{v})\| &= \|y(v) - y(\bar{v})\|_{2,2} + \|v - \bar{v}\|_2 \leq (C + 1)\|v - \bar{v}\|_2 \\ &< (C + 1)r < \delta''. \end{aligned}$$

Hence

$$\hat{J}(v) = J_0(y(v), v) \geq J_0(\bar{y}, \bar{v}) = \hat{J}(\bar{v})$$

and so $\bar{v}$ is a local minimizer to (P'') .

By the same argument, we can show that if $\bar{v}$ is a local minimizer to (P'') , then $(y(\bar{v}), \bar{v}) \in \Phi'$ is a local minimizer to (P') and then $(\bar{y}, \bar{u})$ with $\bar{u} = \frac{1}{\epsilon}(\bar{v} - g(\bar{y}))$ is a local minimizer to (P) . $\square$

Corollary 4.3.7. *Problems (P') and (P'') have at least one minimizer.*

Let $(\bar{y}, \bar{v}) \in \Phi'$. Then the critical cone to (P) at $(\bar{y}, \bar{v})$ is denoted by $\mathcal{C}'[(\bar{y}, \bar{v})]$ which consists of vectors (y, v) satisfying the following conditions:

$$(c'_1) \quad \langle \nabla J_0(\bar{y}, \bar{v}), (y, v) \rangle = \int_{\Omega} (L_y[x]y(x) + L_v[x]v(x))dx \leq 0;$$

$$(c'_2) \quad Ay + (f_y[x] + \frac{1}{\epsilon}g'[x])y = \frac{1}{\epsilon}v \text{ in } \Omega \text{ and } y = 0 \text{ on } \Gamma;$$

$$(c'_3) \quad v(x) \leq 0 \text{ whenever } \bar{v}(x) = 0.$$

Also, we denote by $\mathcal{C}''_{v_0}$ the set of critical directions $v \in U$ satisfying the following conditions:

$$(c''_1) \quad \hat{J}'(v_0)v = \int_{\Omega} [L_y(x, y_0, v_0)S'(v_0)v + L_v(x, y_0, v_0)v]dx \leq 0;$$

$$(c''_2) \quad v(x) \leq 0 \text{ whenever } v_0(x) = 0. \text{ Here } y_0 = S(v_0).$$

It is easy to show that if $v \in \mathcal{C}''_{\bar{v}}$, then $(y, v) \in \mathcal{C}'[(\bar{y}, \bar{v})]$ with $y = S'(\bar{v})v$ and $\bar{y} = S(\bar{v})$, and $(y, u) \in \mathcal{C}[(\bar{y}, \bar{u})]$ with $u = \frac{1}{\epsilon}(v - g'[\bar{y}]y)$ and $\bar{u} = \frac{1}{\epsilon}(\bar{v} - g(\bar{y}))$. Let

$$\mathcal{L}_0(y, v, \psi, w) = J_0(y, v) + \langle \psi, Ay + f(\cdot, y) + \frac{1}{\epsilon}(g(y) - v) \rangle + \langle w, v \rangle$$

be the Lagrange function corresponding to (P') , where $\psi, w \in L^2(\Omega)$. Along with Proposition 4.3.4, we have the following proposition on the first- and second-order optimality conditions for locally optimal solutions to (P') .

Proposition 4.3.8. *Suppose that (A1) – (A4) are satisfied and $(\bar{y}, \bar{v})$ is a local minimizer of (P') . Then there exist a unique function $\bar{\psi} \in Y$ and a unique function $\bar{w} \in U$ such that the following conditions are fulfilled:*

(i) *(the adjoint equation)*

$$\begin{cases} A\bar{\psi} + (f_y(\cdot, \bar{y}) + \frac{1}{\epsilon}g'(\bar{y}))\bar{\psi} = -L_y[\cdot] & \text{in } \Omega \\ \bar{\psi} = 0 & \text{on } \Gamma; \end{cases} \quad (4.3.27)$$

(ii) *(the stationary condition in v)*

$$\bar{w} = \frac{1}{\epsilon}\bar{\psi} + \frac{1}{\epsilon^2}(g(\bar{y}) - \bar{v}) = \frac{1}{\epsilon}\bar{\psi} - L_v[\cdot]; \quad (4.3.28)$$

(iii) *(the complementarity condition) For a.e. $x \in \Omega$ one has*

$$\bar{w}(x) \begin{cases} \geq 0 & \text{if } \bar{v}(x) = 0, \\ = 0 & \text{if } \bar{v}(x) < 0; \end{cases} \quad (4.3.29)$$

(iv) (the non-negative second-order condition)

$$\begin{aligned} D_{(y,v)}^2 \mathcal{L}_0(\bar{y}, \bar{v}, \bar{\psi}, \bar{w})[(y, v), (y, v)] = \\ \int_{\Omega} [L_{yy}[x]y^2(x) + 2L_{yv}[x]y(x)v(x) + L_{vv}[x]v^2(x)] dx + \\ + \int_{\Omega} [\bar{\psi}(f_{yy}[x] + \frac{1}{\epsilon}g''[x])y^2(x)] dx \geq 0 \quad \forall (y, v) \in \mathcal{C}'[(\bar{y}, \bar{v})]. \end{aligned}$$

Moreover, $\bar{y}, \bar{\psi}, \bar{w}$ and $\bar{v}$ belong to $C^{0,\nu}(\bar{\Omega})$ with $\nu = 2 - \frac{N}{2}$.

Conversely, suppose that there exists $\bar{\psi} \in Y$ and $\bar{w} \in U$ such that conditions (i) – (iii) are fulfilled and the strictly second-order condition (SSC):

$$D_{(y,v)}^2 \mathcal{L}_0(\bar{y}, \bar{v}, \bar{\psi}, \bar{w})[(y, v), (y, v)] > 0 \quad \forall (y, v) \in \mathcal{C}'[(\bar{y}, \bar{v})] \setminus \{(0, 0)\}, \quad (4.3.30)$$

then $(\bar{y}, \bar{v})$ is a locally strong minimizer to (P') , that is, there exist a number $\gamma_0 > 0$, a ball $B(\bar{y}, r) \subset H_0^1(\Omega) \cap H^2(\Omega)$ and a ball $B(\bar{v}, r') \subset L^2(\Omega)$ such that

$$J_0(y', v') \geq J_0(\bar{y}, \bar{v}) + \gamma_0 \|v' - \bar{v}\|_2^2 \quad \forall (y', v') \in (B(\bar{y}, r) \times B(\bar{v}, r')) \cap \Phi'. \quad (4.3.31)$$

Let $v_0 \in \Phi''$ and $y_0 = S(v_0)$. Then there exists a unique $\psi_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ which solves

$$A\psi_0 + (f_y(\cdot, y_0) + \frac{1}{\epsilon}g'(y_0))\psi_0 = -L_y(\cdot, y_0, v_0) \quad \text{in } \Omega, \quad \psi_0 = 0 \quad \text{on } \Gamma. \quad (4.3.32)$$

By a simple calculation, we have

$$\begin{aligned} \hat{J}'(v_0)v &= \int_{\Omega} (L_y(x, y_0, v_0)S'(v_0)v(x) + L_v(x, y_0, v_0)v(x)) dx \\ &= \int_{\Omega} (L_v(x, y_0, v_0) - \frac{1}{\epsilon}S'(v_0)(-\epsilon L_y(x, y_0, v_0)))v(x) dx \\ &= \int_{\Omega} (L_v(x, y_0, v_0) - \frac{1}{\epsilon}\psi_0)v(x) dx, \end{aligned} \quad (4.3.33)$$

where ψ_0 satisfies the adjoint equation (4.3.32). By Lemma 2.3 in [16], we have

$$\begin{aligned} \hat{J}''(v_0)v^2 &= \int_{\Omega} (L_{yy}(x, y_0, v_0)z_{v_0,v}^2 + 2L_{yv}(x, y_0, v_0)z_{v_0,v}v + L_{vv}(x, y_0, v_0)v^2 \\ &\quad + \psi_0(f_{yy}(x, y_0, v_0) + \frac{1}{\epsilon}g''(y_0))z_{v_0,v}^2(x)) dx, \end{aligned} \quad (4.3.34)$$

where $z_{v_0, v}$ is a solution of (4.3.25) corresponding to v and $y_0 = S(v_0)$.

The following proposition gives optimality conditions for (P'') .

Proposition 4.3.9. *If $v_0 \in \Phi''$ is a local minimizer of (P'') , then there exists a function $\psi_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ such that the following conditions hold:*

(i) *(The adjoint equation)*

$$\begin{cases} A\psi_0 + (f_y(x, y_0) + \frac{1}{\epsilon}g'(x, y_0))\psi_0 = -L_y(x, y_0, v_0) & \text{in } \Omega, \\ \psi_0 = 0 & \text{on } \Gamma, \text{ where } y_0 = S(v_0); \end{cases}$$

(ii) *(Variational inequality)*

$$\hat{J}'(v_0)(v - v_0) = \int_{\Omega} (L_v(x, y_0, v_0) - \frac{1}{\epsilon}\psi_0)(v(x) - v_0(x))dx \geq 0 \quad \forall v \in \Phi''.$$

This is equivalent to saying that

$$-(L_v(x, y_0(x), v_0(x)) - \frac{1}{\epsilon}\psi_0(x)) \in N((-\infty, 0], v_0(x)) \quad \text{a.e. } x \in \Omega; \quad (4.3.35)$$

(iii) $\hat{J}''(v_0)d^2 \geq 0$ for all $d \in \mathcal{C}_{v_0}''$. Moreover, $v_0, \psi_0 \in C^{0, \nu}$ with $\nu = 2 - N/2$.

Conversely, if (i), (ii) and the strictly second-order sufficient condition

$$\hat{J}''(v_0)d^2 > 0 \quad \forall d \in \mathcal{C}_{v_0}'' \setminus \{0\} \quad (4.3.36)$$

are fulfilled, then there exist numbers $r_0 > 0$ and $\varrho > 0$ such that

$$\hat{J}(v) \geq \hat{J}(v_0) + \varrho\|v - v_0\|_2^2 \quad \forall v \in B_U(v_0, r_0) \cap \Phi''.$$

Proof. The first part of the proposition follows from Theorem 5.1 in [44].

The second part of the proposition is proved by contradiction. It is similar to the proof of Theorem 3.1 in [130]. $\square$

Remark 4.3.10. From the proof of Proposition 4.3.9, we have

$$\begin{aligned} \mathcal{C}_{v_0}'' &= T(\Phi'', v_0) \cap (\hat{J}'(v_0))^\perp, \\ T(\Phi'', v_0) &= \{w \in L^2(\Omega) | w(x) \in T((-\infty, 0], v_0(x)) \text{ a.e. } x \in \Omega\} \\ &= \{w \in L^2(\Omega) | w(x) \leq 0 \text{ whenever } v_0(x) = 0\}. \end{aligned}$$

Since $N((-\infty, 0], \xi_0) = 0$ if $\xi_0 < 0$ and $N((-\infty, 0], \xi_0) = [0, +\infty)$ if $\xi_0 = 0$, we have

$$L_v(x, y_0(x), v_0(x)) - \frac{1}{\epsilon}\psi_0(x) \begin{cases} = 0 & \text{if } v_0(x) < 0, \\ < 0 & \text{if } v_0(x) = 0. \end{cases} \quad (4.3.37)$$

The following result give a relation between strictly second-order sufficient conditions.

Proposition 4.3.11. *If the strictly second-order sufficient condition (4.3.8) is valid at $(\bar{y}, \bar{u})$ with multiplier $(\bar{\varphi}, \bar{e})$, then (4.3.30) is fulfilled at $(\bar{y}, \bar{v})$ with multiplier $(\bar{\varphi}, \bar{e})$ and $\bar{v} = g(\bar{y}) + \epsilon \bar{u}$; and (4.3.36) is satisfied at $\bar{v}$ corresponding to the multiplier $\bar{\varphi}$.*

Proof. Assume that (4.3.8) is valid at $(\bar{y}, \bar{u})$ with the corresponding multiplier $(\bar{\varphi}, \bar{e})$. Let $\bar{v} = g(\bar{y}) + \epsilon \bar{u}$ and $(\bar{\psi}, \bar{w})$ be the Lagrange multipliers corresponding to $(\bar{y}, \bar{v})$. Take any $(y, v) \in \mathcal{C}'[(\bar{y}, \bar{v})] \setminus \{(0, 0)\}$. By putting $u = \frac{1}{\epsilon}(v - g'(\bar{y})y)$, we see that $(y, u) \in \mathcal{C}[(\bar{y}, \bar{u})] \setminus \{(0, 0)\}$. From (4.3.5), we have $\bar{e} = \frac{1}{\epsilon}\bar{\varphi} + \frac{1}{\epsilon^2}(g(\bar{y}) - \bar{v})$. By the uniqueness of Lagrange multipliers, we obtain $\bar{\varphi} = \bar{\psi}$ and $\bar{e} = \bar{w}$. It follows that

$$\begin{aligned} & D_{(y,v)}^2 \mathcal{L}_0(\bar{y}, \bar{v}, \bar{\psi}, \bar{w})[(y, v), (y, v)] \\ &= \int_{\Omega} [(\phi_{yy}[x] + \frac{1}{\epsilon^2}g'[x]^2 + \bar{w}(x)g''[x] + \bar{\psi}(x)f_{yy}[x])y^2(x)] dx + \\ &+ \int_{\Omega} [-\frac{2}{\epsilon^2}g'(\bar{y})y(x)v(x) + \frac{1}{\epsilon^2}v^2(x)] dx \\ &= \int_{\Omega} [(\phi_{yy}[x] + \bar{\psi}(x)f_{yy}[x] + \bar{w}(x)g''[x])y^2(x) + \frac{1}{\epsilon^2}(g'[x]y(x) - v(x))^2] dx \\ &= \int_{\Omega} [(\phi_{yy}[x] + \bar{\psi}(x)f_{yy}[x] + \bar{w}(x)g''[x])y^2(x) + u^2(x)] dx \\ &= D_{(y,u)}^2 \mathcal{L}(\bar{y}, \bar{u}, \bar{\psi}, \bar{w})[(y, u), (y, u)] > 0. \end{aligned}$$

Hence (4.3.30) is fulfilled at $(\bar{y}, \bar{v})$ with $\bar{v} = g(\bar{y}) + \epsilon \bar{u}$. Finally, suppose that there exists ψ such that

$$A\psi + (f_y[x] + \frac{1}{\epsilon}g'[x])\psi = -L_y[x] \quad \text{in } \Omega, \quad \psi = 0 \quad \text{on } \Gamma,$$

where $\bar{y} = S(\bar{v})$ and

$$\hat{J}'(\bar{v})(v - \bar{v}) = \int_{\Omega} (L_v[x] - \frac{1}{\epsilon}\psi)(v(x) - \bar{v}(x))dx \geq 0 \quad \forall v \in \Phi''.$$

Define $w = -(L_v[x] - \frac{1}{\epsilon}\psi)$. Then we have $w \in N(\Phi'', \bar{v})$. Hence $w(x) \in N((-\infty, 0], \bar{v}(x))$. This implies that $w(x) \geq 0$ when $\bar{v}(x) = 0$ and $w(x) = 0$ whenever $\bar{v}(x) < 0$. By the uniqueness of Lagrange multipliers, we have $w = \bar{w} = \bar{e}$ and $\psi = \bar{\psi} = \bar{\varphi}$. Let $v \in \mathcal{C}_v''$. Then

$(y, v) \in \mathcal{C}'[(\bar{y}, \bar{v})]$ with $y = z_{\bar{v}, v}$. By (4.3.34), we have

$$\begin{aligned} \hat{J}''(\bar{v})v^2 &= \\ & \int_{\Omega} (L_{yy}[x]z_{\bar{v}, v}^2 + 2L_{yv}[x]z_{\bar{v}, v}v + L_{vv}[x]v^2 + \psi(f_{yy}[x] + \frac{1}{\epsilon}g''[x])z_{\bar{v}, v}^2(x))dx \\ &= \int_{\Omega} (L_{yy}[x]y^2 + 2L_{yv}[x]yv + L_{vv}[x]v^2 + \bar{\psi}(f_{yy}[x] + \frac{1}{\epsilon}g''[x])y^2)dx \\ &= D_{(y, v)}^2 \mathcal{L}(\bar{y}, \bar{v}, \bar{\psi}, \bar{w})[(y, v), (y, v)] > 0. \end{aligned}$$

The proof of the proposition is complete. $\square$

In the sequel, we will discretize problem (P') into a sequence of approximately optimization problems (P'_h) . This can be done by the variational discretization method in [103] and [104], which discretizes the state variable alone. Then the convergence and error estimates of approximate solutions to (P'_h) can be proved. However, the problem (P'_h) which is obtained by the method in [103] and [104] is a mathematical programming problem in the infinite dimensional space with respect to control variable $u \in U_{ad}$. Here we want to discretize problem (P') for both state variable and control variable in order to obtain mathematical programming problems in the finite dimensional spaces. Namely, we consider the following discrete version (P'_h) for (P') :

$$\left\{ \begin{array}{l} J_0(y_h, v_h) = \\ \int_{\Omega} (\phi(x, y_h(x)) + \frac{1}{2\epsilon^2}g(y_h(x))^2 - \frac{1}{\epsilon^2}v_h(x)g(y_h(x)) + \frac{1}{2\epsilon^2}v_h^2(x))dx \rightarrow \min, \\ \text{s.t. } (y_h, v_h) \in Y_h \times U_h, \\ \int_{\Omega} (\sum_{i,j=1}^N a_{ij}(x)D_i y_h D_j \xi_h + (f(x, y_h) + \frac{1}{\epsilon}g(y_h))\xi_h(x))dx = \\ \int_{\Omega} \frac{1}{\epsilon}v_h(x)\xi_h(x)dx \quad \forall \xi_h \in Y_h, \\ v_h(x) \leq 0 \quad \text{for all } x \in \Omega. \end{array} \right.$$

Obviously, (P'_h) is a mathematical programming problem with finitely many constraints in the finite dimensional space $Y_h \times U_h$. Let us denote by Φ'_h the feasible sets of (P'_h) . By techniques of fixed point theorem (see Lemma 1.4 and the proof of Theorem 1.2 in [216]), we can show that for each $v \in U$, there exists a unique $y_h = y_h(v) \in Y_h$ such that

$$\int_{\Omega} (\sum_{i,j=1}^N a_{ij}(x)D_i y_h D_j \xi_h + (f(x, y_h) + \frac{1}{\epsilon}g(y_h))\xi_h(x))dx = \int_{\Omega} \frac{1}{\epsilon}v(x)\xi_h(x)dx$$

for all $\xi_h \in Y_h$. Let us define the solution map S_h by setting $S_h(v) = y_h(v)$. Then S_h is of class C^1 on U and $\xi_{h, \bar{v}, v} := S'_h(\bar{v})v$ is a solution of

the equation

$$\int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i y_h D_j \xi_h + (f_y(x, \bar{y}_h) + \frac{1}{\epsilon} g'(\bar{y}_h(x))) \xi_h(x) \right) dx = \int_{\Omega} \frac{1}{\epsilon} v(x) \xi_h(x) dx \quad \forall \xi_h \in Y_h.$$

Then the mapping $\hat{J}_h(v) = \int_{\Omega} L(x, S_h(v), v) dx$ is of class C^1 on U . It is easy to show that $(\bar{y}_h, \bar{v}_h) \in \Phi'_h$ is a local minimizer to (P'_h) if and only if $\bar{v}_h$ is a local minimizer of the following problem (P''_h) :

$$(P''_h) \quad \begin{cases} \hat{J}_h(v_h) = \int_{\Omega} L(x, S_h(v_h)(x), v_h(x)) dx \rightarrow \min, \\ \text{s.t. } v_h \in U_h, \\ v_h(x) \leq 0 \quad \text{for all } x \in \Omega. \end{cases} \quad (4.3.38)$$

We denote by Φ''_h the feasible set of (P''_h) , that is,

$$\Phi''_h = \{v_h \in U_h | v_h(x) \leq 0\}.$$

Proposition 4.3.12. *Suppose that (A1) – (A4) are valid. Then for each $h > 0$, problem (P'_h) has at least one minimizer $(\bar{y}_h, \bar{v}_h) \in Y_h \times U_h$ and so (P''_h) has a minimizer $\bar{v}_h$.*

Proof. By (A4), we see that $(0, g(0)) \in \Phi'_h$ for all $h > 0$. Besides, by (A1), we have $J_0(y_h, v_h) \geq \int_{\Omega} \zeta(x) dx$. It follows that $\gamma_h := \inf_{(y_h, v_h) \in \Phi'_h} J_0(y_h, v_h)$ is finite. Then there is a sequence $\{(y_h^l, v_h^l)\}_{l \geq 1} \subset \Phi'_h$ such that

$$\gamma_h = \lim_{l \rightarrow \infty} J_0(y_h^l, v_h^l).$$

Since $(0, g(0)) \in \Phi'_h$, we have

$$J_0(0, g(0)) = \int_{\Omega} \phi(x, 0) dx \geq \lim_{l \rightarrow \infty} J_0(y_h^l, v_h^l).$$

By setting $u_h^l = \frac{1}{\epsilon}(v_h^l - g(y_h^l))$ and using (A1), there exists $l_0 > 0$ such that for all $l \geq l_0$ one has

$$1 + \int_{\Omega} \phi(x, 0) dx > \int_{\Omega} (\zeta(x) + \frac{1}{2} u_h^l(x)^2) dx.$$

This implies that there is a number $M > 0$ such that $\|u_h^l\|_2 \leq M$ for all $l > 0$. It holds that $\{y_h^l\}$ is also bounded in $H_1^0(\Omega)$. In fact, since $(y_h^l, v_h^l) \in \Phi'_h$, we have

$$\int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i y_h^l D_j z_h + f(x, y_h^l) z_h \right) dx = \int_{\Omega} u_h^l z_h dx \quad \forall z_h \in Y_h. \quad (4.3.39)$$

Taking $z_h = y_h^l$ and using (A2), (A3), we get

$$\|y_h^l\|_{H_0^1(\Omega)} \leq C\|u_h^l\|_{L^2(\Omega)} \leq CM \quad (4.3.40)$$

for some constant $C > 0$ which is independent of l . Let

$$y_h^l = \alpha_1^l \varphi_1 + \alpha_2^l \varphi_2 + \cdots + \alpha_{m_h}^l \varphi_{m_h}.$$

By Lemma 2.4.1 in [141], there exist constant $c_1 > 0$ such that

$$c_1(|\alpha_1^l| + |\alpha_2^l| + \cdots + |\alpha_{m_h}^l|) \leq \|y_h^l\|_{1,2} \leq CM.$$

Hence $\{\alpha_i^l\}$ are bounded. Let

$$v_h^l = \beta_1^l \varphi_1 + \beta_2^l \varphi_2 + \cdots + \beta_{k_h}^l \varphi_{k_h}.$$

Since $v_h^l = g(y_h^l) + \epsilon u_h^l$ and (A4), we can show that there exists a constant $M' > 0$ such that

$$\|v_h^l\|_2 \leq M' \quad \forall l > 0.$$

Using Lemma 2.4.1 in [130] again, there exist constant $c_2 > 0$ such that

$$c_2(|\beta_1^l| + |\beta_2^l| + \cdots + |\beta_{k_h}^l|) \leq \|v_h^l\|_2 \leq M'.$$

Hence $\{\beta_i^l\}$ are bounded. By passing to subsequences, we may assume that $\alpha_i^l \rightarrow \alpha_i^0$ and $\beta_i^l \rightarrow \beta_i^0$ as $l \rightarrow \infty$. Put $\bar{v}_h = \sum_{i=1}^{k_h} \beta_i^0 \varphi_i$ and $\bar{y}_h = \sum_{i=1}^{m_h} \alpha_i^0 \varphi_i$. Then $v_h^l \rightarrow \bar{v}_h$ strongly in $L^2(\Omega)$ and $y_h^l \rightarrow \bar{y}_h$ strongly in $H_0^1(\Omega)$. Since $(y_h^l, v_h^l) \in \Phi'_h$, we have

$$\begin{aligned} & \int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i y_h^l D_j z_h + (f(x, y_h^l) + \frac{1}{\epsilon} g(y_h^l)) z_h \right) dx \\ &= \int_{\Omega} \frac{1}{\epsilon} v_h^l z_h dx \quad \forall z_h \in Y_h. \end{aligned} \quad (4.3.41)$$

By letting $l \rightarrow \infty$ in both sides of (4.3.41), we obtain

$$\begin{aligned} & \int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i \bar{y}_h D_j z_h + (f(x, \bar{y}_h) + \frac{1}{\epsilon} g(\bar{y}_h)) z_h \right) dx \\ &= \int_{\Omega} \frac{1}{\epsilon} \bar{v}_h z_h dx \quad \forall z_h \in Y_h. \end{aligned}$$

Since $v_h^l(x) \leq 0$ and $v_h^l(x) \rightarrow \bar{v}_h(x)$ uniformly when $l \rightarrow \infty$, we have

$$\bar{v}_h(x) \leq 0, \quad \forall x \in \Omega.$$

Consequently, $(\bar{y}_h, \bar{v}_h) \in \Phi'_h$. It is easy to show that

$$\gamma_h = \lim_{l \rightarrow \infty} \int_{\Omega} [\phi(x, y_h^l(x)) + \frac{1}{2} u_h^l(x)^2] dx = \int_{\Omega} [\phi(x, \bar{y}_h(x)) + \frac{1}{2} \bar{u}_h(x)^2] dx,$$

where $\bar{u}_h := \frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h))$. Hence $(\bar{y}_h, \bar{v}_h)$ is an optimal solution of (P'_h) . The proof of the proposition is complete. $\square$

To end this section, we present optimality conditions for discrete problems (P'_h) and (P''_h) . Their proofs are similar to the case of continuous problems.

Proposition 4.3.13. *Suppose that (A1)–(A4) are satisfied and $(\bar{y}_h, \bar{v}_h) \in \Phi'_h$ is a local minimizer of (P'_h) . Then there exist a unique function $\bar{\psi}_h \in Y_h$ and a unique function $\bar{w}_h \in U_h$ such that the following conditions are fulfilled:*

(i) *(The adjoint equation)*

$$\begin{aligned} \int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i \bar{\psi}_h D_j \xi_h + (f_y(x, \bar{y}_h) + \frac{1}{\epsilon} g'(\bar{y}_h(x))) \bar{\psi}_h(x) \xi_h(x) \right) dx = \\ - \int_{\Omega} L_y(x, \bar{y}_h(x), \bar{v}_h(x)) \xi_h(x) dx \quad \forall \xi_h \in Y_h; \end{aligned} \quad (4.3.42)$$

(ii) *(The stationary condition in v_h)*

$$\bar{w}_h = \frac{1}{\epsilon} \bar{\psi}_h - L_v(\cdot, \bar{y}_h, \bar{v}_h) = \frac{1}{\epsilon} \bar{\psi}_h - \frac{1}{\epsilon^2} (\bar{v}_h - g(\bar{y}_h)); \quad (4.3.43)$$

(iii) *(The complementarity condition) For a.e. $x \in \Omega$ one has*

$$\bar{w}_h(x) \begin{cases} \geq 0 & \text{if } \bar{v}_h(x) = 0, \\ = 0 & \text{if } \bar{v}_h(x) < 0. \end{cases} \quad (4.3.44)$$

Proposition 4.3.14. *Suppose that (A1)–(A4) are satisfied. If $\bar{v}_h \in \Phi''_h$ is a local minimizer of (P''_h) , then there exist a unique function $\bar{\psi}_h \in Y_h$ such that the following conditions are fulfilled:*

(i) *(The adjoint equation)*

$$\begin{aligned} \int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i \bar{\psi}_h D_j \xi_h + (f_y(x, \bar{y}_h) + \frac{1}{\epsilon} g'(\bar{y}_h(x))) \bar{\psi}_h(x) \xi_h(x) \right) dx = \\ - \int_{\Omega} L_y(x, \bar{y}_h(x), \bar{v}_h(x)) \xi_h(x) dx \end{aligned} \quad (4.3.45)$$

for all $\xi_h \in Y_h$, where $\bar{y}_h = y_h(\bar{v}_h)$;

(ii) (Variational inequality)

$$\hat{J}'_h(\bar{v}_h)(v - \bar{v}_h) = \int_{\Omega} (L_v(x, \bar{y}_h(x), \bar{v}_h(x)) - \frac{1}{\epsilon} \bar{\psi}_h(x))(v_h(x) - \bar{v}_h(x)) dx \geq 0 \quad (4.3.46)$$

for all $v_h \in \Phi''_h$.

4.3.3 Numerical analysis for state equations and adjoint equations

In this subsection, we shall need the following important lemma. It is a particular case of Theorem 3.1.5 in [55] when $s = 0$.

Lemma 4.3.15. *Let $k \geq 0, m \geq 0$, and $p, q \in [1, +\infty]$. If the embeddings*

$$\begin{aligned} W^{k+1,p}(T) &\hookrightarrow C^0(T) \\ W^{k+1,p}(T) &\hookrightarrow W^{m,q}(T) \end{aligned}$$

hold, then there exists a constant $C > 0$ independent of h such that

$$\|y - \Pi_h y\|_{W^{m,q}(T)} \leq C h^{N(1/q-1/p)+k+1-m} \|y\|_{W^{k+1,p}(T)}, \quad (4.3.47)$$

where $\Pi_h : C(\bar{\Omega}) \rightarrow U_h$ denotes the nodal interpolation operator, that is, $\Pi_h y(N_i) = y(N_i)$ for any node N_i of a triangulation.

In the sequel we shall use the following inequality. Its proof can be found in [11, Theorem 17.2].

$$\|y_h\|_{W^{m,q}(\Omega)} \leq C \frac{1}{h^{N \max(0, 1/p-1/q)} h^{m-l}} \|y_h\|_{W^{l,p}(\Omega)} \quad \forall y_h \in Y_h, l \leq m, \quad (4.3.48)$$

where $C > 0$ being independent of h .

• Error estimates for solutions of state equations

Given $v \in L^2(\Omega)$, we shall denote by $y(v)$ and $y_h(v)$ the solutions of the state equations of (P') and (P'_h) corresponding to v , respectively. Namely, $y(v) = S(v) \in Y$, $y_h(v) = S_h(v) \in Y_h$ such that

$$\int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i y(v) D_j \xi + (f(x, y(v)) + \frac{1}{\epsilon} g(y(v)) \xi) \right) dx = \int_{\Omega} \frac{1}{\epsilon} v \xi dx \quad (4.3.49)$$

for all $\xi \in H_0^1(\Omega)$, and

$$\int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i y_h(v) D_j \xi_h + (f(x, y_h(v)) + \frac{1}{\epsilon} g(y_h(v))) \xi_h \right) dx = \int_{\Omega} \frac{1}{\epsilon} v \xi_h dx \quad (4.3.50)$$

for all $\xi_h \in Y_h$.

Proposition 4.3.16. (C.f. [5, Theorem 2]) *Suppose assumptions (A2) and (A3) are satisfied. Then for each v with $\|v\|_2 \leq M$, there exist $C_M > 0$ and $h_M > 0$ which depend on ϵ such that*

$$\|y(v) - y_h(v)\|_{1,2} \leq C_M h \|y(v)\|_{2,2}, \quad (4.3.51)$$

$$\|y(v) - y_h(v)\|_2 \leq C_M h^2 \|y(v)\|_{2,2}, \quad (4.3.52)$$

$$\|y(v) - y_h(v)\|_{\infty} \leq C_M h^{2-\frac{N}{2}} \|y(v)\|_{2,2} \quad (4.3.53)$$

for $0 < h < h_M$.

Below, we shall assume that $0 < h < h_M$.

Proposition 4.3.17. *Suppose that $\|v_h\|_2 + \|v\|_2 \leq M$, $y(v)$ and $y_h(v_h)$ are solutions of (4.3.49) and (4.3.50) corresponding to v and v_h . Then there exists a constant $C = C(M, N, \Omega, \epsilon)$ such that the following estimations are valid.*

$$\|y_h(v_h) - y(v)\|_{1,2} \leq C(h \|y(v_h)\|_{2,2} + \|v_h - v\|_2), \quad (4.3.54)$$

$$\|y_h(v_h) - y(v)\|_2 \leq C(h^2 \|y(v_h)\|_{2,2} + \|v_h - v\|_2), \quad (4.3.55)$$

$$\|y_h(v_h) - y(v)\|_{\infty} \leq C(h^{2-\frac{N}{2}} \|y(v_h)\|_{2,2} + \|v_h - v\|_2). \quad (4.3.56)$$

Proof. By (4.3.3) and Proposition 4.3.16, we have

$$\begin{aligned} \|y_h(v_h) - y(v)\|_{1,2} &\leq \|y_h(v_h) - y(v_h)\|_{1,2} + \|y(v_h) - y(v)\|_{1,2} \\ &\leq C(h \|y(v_h)\|_{2,2} + \|v_h - v\|_2). \end{aligned}$$

We obtain inequality (4.3.54). The other estimates are proved similarly. $\square$

• Error estimates for solutions of adjoint equations

Given $v \in U$ and $v_h \in U_h$, we consider the following adjoint equations

for (P') and (P'_h) :

$$\begin{aligned} a(\psi, \xi) + \langle (f_y(\cdot, y(v)) + \frac{1}{\epsilon} g'(y(v))) \psi, \xi \rangle = \\ \langle -\phi_y(\cdot, y(v)) - \frac{1}{\epsilon^2} g(y(v)) g'(y(v)) + \frac{1}{\epsilon^2} g'(y(v)) v, \xi \rangle, \forall \xi \in H_0^1(\Omega), \end{aligned} \quad (4.3.57)$$

$$\begin{aligned} a(\psi_h, \xi_h) + \langle (f_y(\cdot, y_h(v_h)) + \frac{1}{\epsilon} g'(y_h(v_h))) \psi_h, \xi_h \rangle = \\ \langle -\phi_y(\cdot, y_h(v_h)) - \frac{1}{\epsilon^2} (g(y_h(v_h)) g'(y_h(v_h)) + g'(y_h(v_h)) v_h), \xi_h \rangle, \forall \xi_h \in Y_h, \end{aligned} \quad (4.3.58)$$

where

$$a(y, \xi) := \int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_j y(x) D_i \xi(x) \right) dx,$$

$y(v)$ is the solution to (4.3.49) corresponding to v and $y_h(v_h)$ is the solution to (4.3.50) corresponding to v_h . For convenience, we shall put

$$\gamma(y) = \phi_y(\cdot, y) + \frac{1}{\epsilon^2} g(y) g'(y). \quad (4.3.59)$$

By (A2), a has symmetric property, that is, $a(y, \xi) = a(\xi, y)$ for all $y, \xi \in H_0^1(\Omega)$. In this subsection, we shall give error estimates of $\|\psi_h(v_h) - \psi(v)\|$. Since

$$\|\psi_h(v_h) - \psi(v)\| \leq \|\psi_h(v_h) - \psi(v_h)\| + \|\psi(v_h) - \psi(v)\|,$$

we need to estimate the terms $\|\psi_h(v_h) - \psi(v_h)\|$ and $\|\psi(v_h) - \psi(v)\|$.

Proposition 4.3.18. *Suppose that there exists $M > 0$ such that $\|v\|_2 + \|v_h\|_2 \leq M$, then there exists a constant $C_M = C_M(\epsilon) > 0$ such that*

$$\|\psi_h(v_h) - \psi(v_h)\|_{1,2} \leq C_M h, \quad (4.3.60)$$

$$\|\psi_h(v_h) - \psi(v_h)\|_2 \leq C_M h^2, \quad (4.3.61)$$

$$\|\psi_h(v_h) - \psi(v_h)\|_{\infty} \leq C_M h^{2-\frac{N}{2}}. \quad (4.3.62)$$

Proof. (i). Since $\|v\|_2 + \|v_h\|_2 \leq M$ and Proposition 4.3.16, there exists $M_1 > 0$ such that

$$\|y(v)\|_{\infty} + \|y_h(v_h)\|_{\infty} \leq M_1.$$

Hence

$$\|\psi(u_h)\|_{\infty} \leq C \|\psi(u_h)\|_{2,2} \leq C \left\| -\zeta(y(v_h)) + \frac{1}{\epsilon^2} g'(y(v_h)) v_h \right\|_2 \leq M_2$$

for constant M_2 which is independent of h . Inserting $v = v_h$ and $\xi = \xi_h$ into (4.3.57), and subtracting from (4.3.58), we obtain

$$\begin{aligned}
& a(\psi - \psi_h, \xi_h) + \langle f_y(x, y(v_h))\psi - f_y(x, y_h(v_h))\psi_h + \\
& \frac{1}{\epsilon}(g'(y(v_h))\psi - g'(y_h(v_h)))\psi_h, \xi_h \rangle \\
& = \langle \gamma(y_h(v_h)) - \gamma(y(v_h)) + \frac{v_h}{\epsilon^2}(g'(y(v_h)) - g'(y_h(v_h))), \xi_h \rangle \quad \forall \xi_h \in Y_h.
\end{aligned} \tag{4.3.63}$$

From this, (A2) and the fact

$$\langle (f_y(x, y_h(u_h)) + \frac{1}{\epsilon}g'(y_h(u_h)))(\varphi - \varphi_h), \varphi - \varphi_h \rangle \geq 0,$$

we have

$$\begin{aligned}
& \bar{\alpha} \|\psi - \psi_h\|_{1,2}^2 \leq a(\psi - \psi_h, \psi - \psi_h) \\
& = a(\psi - \psi_h, \psi - \xi'_h) + a(\psi - \psi_h, \xi'_h - \psi_h) = \\
& = a(\psi - \psi_h, \psi - \xi'_h) + \langle \gamma(y_h(v_h)) - \gamma(y(v_h)) + \\
& + \frac{v_h}{\epsilon^2} (g'(y(v_h)) - g'(y_h(v_h))), \xi'_h - \psi_h \rangle \\
& - \langle (f_y(x, y(v_h))\psi - f_y(x, y_h(v_h))\psi_h + \frac{1}{\epsilon} (g'(y(v_h))\psi - g'(y_h(v_h))\psi_h), \xi'_h - \psi_h \rangle \\
& = a(\psi - \psi_h, \psi - \xi'_h) + \langle \gamma(y_h(v_h)) - \gamma(y(v_h)) + \\
& + \frac{v_h}{\epsilon^2} (g'(y(v_h)) - g'(y_h(v_h))), \xi'_h - \psi_h \rangle \\
& - \langle (f_y(x, y(v_h))\psi - f_y(x, y_h(v_h))\psi_h + \frac{1}{\epsilon} (g'(y(v_h))\psi - g'(y_h(v_h))\psi_h), \xi'_h - \psi \rangle \\
& - \langle (f_y(x, y(v_h))\psi - f_y(x, y_h(v_h))\psi_h + \frac{1}{\epsilon} (g'(y(v_h))\psi - g'(y_h(v_h))\psi_h), \psi - \psi_h \rangle \\
& \leq \|\psi - \psi_h\|_{1,2} \|\psi - \xi'_h\|_{1,2} + L_{g,M_1} k_{\phi,M_1} \|y(v_h) - y_h(v_h)\|_{\infty} \|\xi'_h - \psi_h\|_2 + \\
& + \frac{M}{\epsilon^2} k_{g,M_1} \|y(v_h) - y_h(v_h)\|_{\infty} \|\xi'_h - \psi_h\|_2 \\
& - \langle (f_y(x, y(v_h))\psi - f_y(x, y_h(v_h))\psi_h + \frac{1}{\epsilon} (g'(y(v_h))\psi - g'(y_h(v_h))\psi_h), \xi'_h - \psi \rangle \\
& - \langle (f_y(x, y(v_h))\psi - f_y(x, y_h(v_h))\psi_h + \frac{1}{\epsilon} (g'(y(v_h))\psi - g'(y_h(v_h))\psi_h), \psi - \psi_h \rangle \\
& \leq \left(\|\psi - \psi_h\|_{1,2} \|\psi - \xi'_h\|_{1,2} + L_{g,M_1} k_{\phi,M_1} \|y(v_h) - y_h(v_h)\|_2 \|\xi'_h - \psi_h\|_2 + \right. \\
& \left. + \frac{M}{\epsilon^2} k_{g,M_1} \|y(v_h) - y_h(v_h)\|_{\infty} \|\xi'_h - \psi_h\|_2 \right) \\
& + \left(k_{f,M_1} \|y(v_h) - y_h(v_h)\|_{\infty} \|\psi\|_2 + L_{f,M_1} \|\psi - \psi_h\|_2 + \right. \\
& \left. + \frac{k_{g,M_1}}{\epsilon} \|y(v_h) - y_h(v_h)\|_{\infty} \|\psi\|_2 + \frac{L_{g,M_1}}{\epsilon} \|\psi - \psi_h\|_2 \|\xi'_h - \psi\|_2 \right) \\
& + \left((k_{f,M_1} + \frac{k_{g,M_1}}{\epsilon}) \|y(v_h) - y_h(v_h)\|_2 \|\psi\|_{\infty} \|\psi - \psi_h\|_2 \right) \\
& =: T_1 + T_2 + T_3, \tag{4.3.64}
\end{aligned}$$

where $\xi'_h = \Pi_h \psi$. Using Lemma 4.3.15 and Proposition 4.3.16, we have

$$\begin{aligned}
T_1 &= \left(\|\psi - \xi'_h\|_{1,2} \|\psi - \psi_h\|_{1,2} + L_{g,M_1} k_{\phi,M_1} \|y(v_h) - y_h(v_h)\|_2 \|\xi'_h - \psi_h\|_2 + \right. \\
&\quad \left. + \frac{M}{\epsilon^2} k_{g,M_1} \|y(u_h) - y_h(u_h)\|_\infty \|\xi'_h - \psi_h\|_2 \right) \\
&\leq Ch \|\psi(u_h)\|_{2,2} \|\psi - \psi_h\|_{1,2} + L_{g,M_1} k_{\phi,M_1} Ch^2 \|y(v_h)\|_{2,2} h^2 \|\psi(v_h)\|_{2,2} + \\
&\quad + \frac{M}{\epsilon^2} k_{g,M} Ch^{2-\frac{N}{2}} \|y(v_h)\|_{2,2} h^2 \|\psi(u_h)\|_{2,2} \\
&\leq C_1(h \|\psi - \psi_h\|_{1,2} + h^2). \tag{4.3.65}
\end{aligned}$$

Here we used the fact that $\|y(v_h)\|_{2,2}$ and $\|\psi(v_h)\|_{2,2}$ are bounded and

$$\begin{aligned}
\|\xi'_h - \psi_h\|_2 &\leq \|\xi'_h - \psi(v_h)\|_2 + \|\psi(v_h) - \psi_h(v_h)\|_2 \\
&\leq Ch^2 \|\psi(v_h)\|_{2,2} + \|\psi - \psi_h\|_2 \leq C(h^2 \|\psi(v_h)\|_{2,2} + \|\psi - \psi_h\|_{1,2}).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
T_2 &= \left(k_{f,M_1} \|y(v_h) - y_h(v_h)\|_\infty \|\Psi\|_2 + L_{f,M_1} \|\psi - \psi_h\|_2 \right. \\
&\quad + \frac{k_{g,M_1}}{\epsilon} \|y(v_h) - y_h(v_h)\|_\infty \|\psi\|_2 \\
&\quad \left. + \frac{L_{g,M_1}}{\epsilon} \|\psi - \psi_h\|_2 \|\xi'_h - \psi\|_2 \right) \leq C_2(\|\psi - \psi_h\|_{1,2} + 1)h^2. \tag{4.3.66}
\end{aligned}$$

and

$$T_3 = (k_{f,M_1} + \frac{k_{g,M_1}}{\epsilon}) \|y(v_h) - y_h(v_h)\|_2 \|\psi\|_\infty \|\psi - \psi_h\|_2 \leq C_3 h \|\psi - \psi_h\|_{1,2}. \tag{4.3.67}$$

Plugging estimates of T_1, T_2 and T_3 into (4.3.64), we get

$$\begin{aligned}
\bar{\alpha} \|\psi - \psi_h\|_{1,2}^2 &\leq C_1(h \|\psi - \psi_h\|_{1,2} + h^2) + C_2(\|\psi - \psi_h\|_{1,2} + 1)h^2 \\
&\quad + C_3 h \|\psi - \psi_h\|_{1,2} \leq C_M(\|\psi - \psi_h\|_{1,2} h + h^2).
\end{aligned}$$

Hence $\|\psi - \psi_h\|_{1,2} \leq C_M h$ for constant $C_M > 0$ which is independent of h . We obtain estimate (4.3.60).

(ii). To prove (4.3.61), we put

$$\begin{aligned}
\alpha(x) &= f_y(x, y(v_h)(x)) + \frac{1}{\epsilon} g'(y(v_h)(x)) \\
\beta(x) &= f_y(x, y_h(v_h)(x)) + \frac{1}{\epsilon} g'(y_h(v_h)(x)).
\end{aligned}$$

Then $\alpha(x) \geq 0$ and $\beta(x) \geq 0$ and so for each $v \in L^2(\Omega)$, there exists $w \in H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$a(w, \xi) + \langle \beta w, \xi \rangle = \langle v, \xi \rangle \quad \forall \xi \in H_0^1(\Omega)$$

and $\|w\|_{2,2} \leq C\|v\|_2$. Also, there exists $w_h \in Y_h$ such that

$$a(w_h, \xi_h) + \langle \beta w_h, \xi_h \rangle = \langle v, \xi_h \rangle \quad \forall \xi_h \in Y_h$$

and $\|w_h\|_{1,2} \leq C\|v\|_2$. Let us claim that

$$\|w - w_h\|_{1,2} \leq Ch\|w\|_{2,2} \leq Ch\|v\|_2 \quad (4.3.68)$$

$$\|w - w_h\|_2 \leq Ch^2\|w\|_{2,2} \leq Ch^2\|v\|_2. \quad (4.3.69)$$

In fact, (4.3.68) follows from Cea's lemma in [55] and (4.3.69) follows

from Theorem 3.2.5 in [55]. Using (4.3.63), we have

$$\begin{aligned}
\langle v, \psi - \psi_h \rangle &= a(w, \psi - \psi_h) + \langle \beta w, \psi - \psi_h \rangle \\
&= a(w - w_h, \psi - \psi_h) + a(w_h, \psi - \psi_h) + \langle \beta w, \psi - \psi_h \rangle \\
&= a(w - w_h, \psi - \psi_h) - \langle \alpha \psi - \beta \psi_h, w_h \rangle + \langle \gamma(y_h(v_h)) - \gamma(y(v_h)) \rangle + \\
&\quad + \frac{v_h}{\epsilon^2} (g'(y(v_h)) - g'(y_h(v_h)), w_h) + \langle \beta(\psi - \psi_h), w \rangle \\
&= a(w - w_h, \psi - \psi_h) + \langle (\beta - \alpha)\psi, w_h \rangle - \langle \beta \psi_h, w - w_h \rangle - \langle \beta \psi, w_h - w \rangle \\
&\quad + \langle \gamma(y_h(v_h)) - \gamma(y(v_h)) \rangle + \frac{1}{\epsilon^2} (g'(y(v_h)) - g'(y_h(v_h)))v_h, w_h \rangle \\
&\leq \|w - w_h\|_{1,2} \|\psi - \psi_h\|_{1,2} + (k_{f,M_1} + \frac{k_{g,M_1}}{\epsilon}) \|\psi(u_h)\|_\infty \|y(v_h) - y_h(v_h)\|_2 \|w_h\|_2 \\
&\quad + (L_{f,M_1} + \frac{L_{g,M_1}}{\epsilon}) \|\psi_h\|_2 \|w - w_h\|_2 + (L_{f,M_1} + \frac{L_{g,M_1}}{\epsilon}) \|\psi\|_2 \|w - w_h\|_2 \\
&\quad + M(L_{g,M_1} k_{\phi,M_1} + \frac{k_{g,M_1}}{\epsilon^2}) \|y(v_h) - y_h(v_h)\|_2 \|w_h\|_2 + \\
&\quad + \frac{1}{\epsilon^2} (g'(y(v_h)) - g'(y_h(v_h)))v_h, w \rangle \\
&\leq Ch \|w\|_{2,2} h + (k_{f,M_1} + \frac{k_{g,M_1}}{\epsilon}) \|\psi(v_h)\|_\infty Ch^2 \|y(v_h)\|_{2,2} \|v\|_2 \\
&\quad + (L_{f,M_1} + \frac{L_{g,M_1}}{\epsilon}) Ch^2 \|w\|_{2,2} + (L_{f,M_1} + \frac{L_{g,M_1}}{\epsilon}) Ch^2 \|w\|_{2,2} \\
&\quad + M(L_{g,M_1} k_{\phi,M_1} + \frac{k_{g,M_1}}{\epsilon^2}) Ch^2 \|y(v_h)\|_{2,2} \|v\|_2 + \frac{k_{g,M_1}}{\epsilon} \|w\|_\infty \|y(u_h) - y_h\|_2 \|v_h\|_2 \\
&\leq C' h^2 \|v\|_2 + (k_{f,M_1} + \frac{k_{g,M_1}}{\epsilon}) C' h^2 \|v\|_2 + (L_{f,M_1} + \frac{L_{g,M_1}}{\epsilon}) C' h^2 \|v\|_2 + \\
&\quad + (L_{f,M_1} + \frac{L_{g,M_1}}{\epsilon}) C' h^2 \|v\|_2 \\
&\quad + M(k_{\phi,M_1} + \frac{k_{g,M_1}}{\epsilon^2}) Ch^2 \|y(v_h)\|_{2,2} \|v\|_2 + \frac{k_{g,M_1}}{\epsilon^2} Ch^2 \|y(u_h)\|_{2,2} \|v\|_2 \\
&\leq C'' h^2 \|v\|_2.
\end{aligned}$$

Here $C'' > 0$ is independent of h and v . This implies that $\|\psi - \psi_h\|_2 \leq C'' h^2$. Hence (4.3.61) is established.

(iii) To prove (4.3.62) we mimic the proof of Theorem 1 in [4] as follows. We first have that

$$\|\psi(v_h) - \psi_h(v_h)\|_\infty \leq \|\psi(v_h) - \Pi_h \psi(v_h)\|_\infty + \|\Pi_h \psi(v_h) - \psi_h(v_h)\|_\infty. \quad (4.3.70)$$

Due to Lemma 4.3.15, taking $k = 1, m = 0, q = 1$ and $p = 2$, we have that

$$\|\psi(v_h) - \Pi_h \psi(v_h)\|_\infty \leq Ch^{2-N/2} \|\psi(v_h)\|_{2,2} \quad (4.3.71)$$

Using (4.3.48), we have

$$\|\Pi_h \psi(v_h) - \psi_h(v_h)\|_\infty \leq Ch^{-N/2} \|\Pi_h \psi(v_h) - \psi_h(v_h)\|_2. \quad (4.3.72)$$

Applying Lemma 4.3.15 for $k = 1, m = 0, q = 2, p = 2$, we get

$$\|\Pi_h \psi(v_h) - \psi(v_h)\|_2 \leq Ch^2 \|\psi(v_h)\|_{2,2}.$$

It follows that

$$\begin{aligned} \|\Pi_h \psi(v_h) - \psi_h(v_h)\|_\infty &\leq Ch^{-N/2} (\|\Pi_h \psi(v_h) - \psi(v_h)\| + \|\psi(v_h) - \psi_h(v_h)\|_2) \\ &\leq Ch^{2-N/2} (Ch^2 \|\psi(v_h)\|_{2,2} + Ch^2) \leq C'' h^{2-N/2}. \end{aligned}$$

Combining this with (4.3.70) and (4.3.71), we get

$$\|\psi(v_h) - \psi_h(v_h)\|_\infty \leq Ch^{2-N/2} \|\psi(v_h)\|_{2,2} + C'' h^{2-N/2} \leq C_M h^{2-N/2}. \quad (4.3.73)$$

The proof is complete. $\square$

Proposition 4.3.19. *Suppose that there exists $M > 0$ such that $\|v\|_2 + \|v_h\|_2 \leq M$, then there exists constant $C'_M = C'_M(\epsilon) > 0$ such that*

$$\|\psi_h(v_h) - \psi(v)\|_{1,2} \leq C'_M (h + \|v_h - v\|_2), \quad (4.3.74)$$

$$\|\psi_h(v_h) - \psi(v)\|_2 \leq C'_M (h^2 + \|v_h - v\|_2), \quad (4.3.75)$$

$$\|\psi_h(v_h) - \psi(v)\|_\infty \leq C'_M (h^{2-\frac{N}{2}} + \|v_h - v\|_2). \quad (4.3.76)$$

Proof. We first estimate $\|\psi(v_h) - \psi(v)\|_{2,2}$. Note that $\psi(v_h)$ and $\psi(v)$ are solutions of

$$A^* \psi(u_h) + (f_y(\cdot, y(v_h)) + \frac{1}{\epsilon} g'(y(v_h))) \varphi(u_h) = -\gamma(y(v_h)) + \frac{1}{\epsilon^2} g'(y(v_h)) v_h \quad (4.3.77)$$

and

$$A^* \psi(v) + (f_y(\cdot, y(v)) + \frac{1}{\epsilon} g'(y(v))) \psi(v) = -\gamma(y(v)) + \frac{1}{\epsilon^2} g'(y(v)) v, \quad (4.3.78)$$

respectively, where γ is defined by (4.3.59). Recall that $y(u_h)$ and $y = y(u)$ are unique solutions of equation (4.3.49) corresponding to v_h and v , respectively. By subtracting both sides of the equations, we have

$$\begin{aligned} A^* (\psi(v_h) - \psi(v)) + (f_y(\cdot, y(v_h)) + \frac{1}{\epsilon} g'(y(v_h))) (\psi(v_h) - \psi(v)) = \\ (f_y(\cdot, y(v)) - f_y(\cdot, y(v_h)) + \frac{1}{\epsilon} g'(y) - \frac{1}{\epsilon} g'(y(v_h))) \psi(v) + \gamma(y(v)) - \gamma(y(v_h)) + \\ + \frac{1}{\epsilon^2} (g'(y(v_h)) v_h - g'(y(v)) v). \end{aligned}$$

Since $\|v\|_2 + \|v_h\|_2 \leq M$, Propositions 4.3.1 implies that

$$\|y(v_h)\|_\infty + \|y(v)\|_{2,2} \leq M_1$$

for some constant $M_1 > 0$ independent of h . By (A3) and (A4), we see that $f_y(\cdot, y(v_h)) + \frac{1}{\epsilon}g'(y(v_h))$ is non-negative and belongs to $L^\infty(\Omega)$. Hence, there exists a constant $C > 0$ such that

$$\begin{aligned} & \|\psi(v_h) - \psi(v)\|_{2,2} \leq \\ & C\left(\|(f_y(\cdot, y(v)) - f_y(\cdot, y(v_h)) + \frac{1}{\epsilon}g'(y(v)) - \frac{1}{\epsilon}g'(y(v_h)))\psi(v)\|_2 + \right. \\ & \left. \|\gamma(\cdot, y(v)) - \gamma(\cdot, y(v_h))\|_2 + \frac{1}{\epsilon^2}\|g'(y(v_h))v_h - g'(y(v))v\|_2\right). \end{aligned} \quad (4.3.79)$$

By (A3) and (A4), we have

$$\begin{aligned} & \|(f_y(\cdot, y(v)) - f_y(\cdot, y(v_h)) + \frac{1}{\epsilon}g'(y(v)) - \frac{1}{\epsilon}g'(y(v_h)))\psi(v)\|_2 \\ & \leq \|(f_y(\cdot, y(v)) - f_y(\cdot, y(v_h)) + \frac{1}{\epsilon}g'(y(v)) - \frac{1}{\epsilon}g'(y(v_h)))\|_\infty \|\psi(v)\|_2 \\ & \leq (k_{f,M_1} + \frac{1}{\epsilon}k_{g,M_1})\|y(v) - y(v_h)\|_\infty C\|v\|_2 \\ & \leq CM(k_{f,M_1} + \frac{1}{\epsilon}k_{g,M_1})\|y(v) - y(v_h)\|_{2,2} \\ & \leq M_2(k_{f,M_1} + \frac{1}{\epsilon}k_{g,M_1})\|v - v_h\|_2. \end{aligned}$$

By (A1) and (A4), we have

$$\begin{aligned} \|\gamma(y(v)) - \gamma(y(v_h))\|_2 & \leq (L_{g,M_1}k_{\phi,M_1} + \frac{k_{g,M_1}}{\epsilon^2})\|y(v) - y(v_h)\|_2 \\ & \leq C(L_{g,M_1}k_{\phi,M_1} + \frac{k_{g,M_1}}{\epsilon^2})\|v - v_h\|_2. \end{aligned}$$

Also, we get

$$\begin{aligned} \frac{1}{\epsilon^2}\|g'(y(v_h))v_h - g'(y(v))v\|_2 & = \frac{1}{\epsilon^2}\|(g'(y(v_h)) - g'(y(v)))v_h + g'(y(v))(v_h - v)\|_2 \\ & \leq \frac{1}{\epsilon^2}(k_{g,M}\|y(v_h) - y(v)\|_\infty\|v_h\|_2 + L_{g,M}\|v - v_h\|_2) \leq C\|v - v_h\|_2. \end{aligned}$$

Combining these with (4.3.79), yields

$$\begin{aligned} \|\psi(v_h) - \psi(v)\|_{2,2} & \leq C(M_2(k_{f,M_1} + \frac{1}{\epsilon}k_{g,M_1})\|v - v_h\|_2 + C(L_{g,M_1}k_{\phi,M_1} + \\ & + \frac{k_{g,M_1}}{\epsilon^2})\|v - v_h\|_2 + C\|v - v_h\|_2) \leq C'\|v - v_h\|_2 \end{aligned}$$

for some constant $C' > 0$ independent of h . From this and Proposition 4.3.18, we have

$$\begin{aligned}\|\psi_h(v_h) - \psi(v)\|_{1,2} &\leq \|\psi_h(v_h) - \psi(v_h)\|_{1,2} + \|\psi(v_h) - \psi(v)\|_{1,2} \\ &\leq C_M(h + \|v - v_h\|_2), \\ \|\psi_h(v_h) - \psi(v)\|_2 &\leq \|\psi_h(v_h) - \psi(v_h)\|_2 + \|\psi(v_h) - \psi(v)\|_2 \\ &\leq C_M(h^2 + \|v - v_h\|_2), \\ \|\psi_h(v_h) - \psi(v)\|_\infty &\leq \|\psi_h(v_h) - \psi(v_h)\|_\infty + \|\psi(v_h) - \psi(v)\|_{2,2} \\ &\leq C_M(h^{2-N/2} + \|v - v_h\|_2),\end{aligned}$$

which are estimations (4.3.74), (4.3.75) and (4.3.76). The proof of the proposition is completed. $\square$

4.3.4 Convergence and error estimates

The following theorem shows that the solutions of (P'_h) converge to a solution of (P') and (P) .

Theorem 4.3.20. *Suppose (A1) – (A4) are satisfied, $\{(\bar{y}_h, \bar{v}_h)\}$ is a sequence of optimal solutions to (P'_h) . Then there exists a subsequence (y'_h, v'_h) of $\{(\bar{y}_h, \bar{v}_h)\}$ converging strongly to an optimal solution $(\bar{y}, \bar{v})$ of (P') in $(H_0^1(\Omega) \cap L^\infty(\Omega)) \times L^\infty(\Omega)$. Consequently, $(\bar{y}, \frac{1}{\epsilon}(\bar{v} - g(\bar{y})))$ is an optimal solution to (P) and the sequence $(y'_h, \frac{1}{\epsilon}(v'_h - g(y'_h)))$ converges strongly to $(\bar{y}, \frac{1}{\epsilon}(\bar{v} - g(\bar{y})))$ in $(H_0^1(\Omega) \cap L^\infty(\Omega)) \times L^\infty(\Omega)$.*

In addition, if (P) has a unique optimal solution $(\bar{y}, \bar{u})$, then the total sequence $(\bar{y}_h, \frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h)))$ converges strongly to $(\bar{y}, \bar{u})$ in $(H_0^1(\Omega) \cap L^\infty(\Omega)) \times L^\infty(\Omega)$.

Proof. Since $(\bar{y}_h, \bar{v}_h)$ is an optimal solution to (P'_h) and $(0, g(0)) \in \Phi'_h$ for all h , we have

$$J_0(0, g(0)) = \int_{\Omega} \phi(x, 0) dx \geq J_0(\bar{y}_h, \bar{v}_h) = \int_{\Omega} (\phi(x, \bar{y}_h) + \frac{1}{2}[\frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h))]^2) dx$$

By putting $\bar{u}_h = \frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h))$ and using (A1), we get

$$\int_{\Omega} \phi(x, 0) dx \geq \int_{\Omega} (\zeta(x) + \frac{1}{2}\bar{u}_h(x)^2) dx.$$

This implies that $\|\bar{u}_h\|_2$ is bounded. Since $(\bar{y}_h, \bar{v}_h) \in \Phi'_h$, we have

$$\begin{aligned}&\int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i \bar{y}_h D_j \xi_h + (f(x, \bar{y}_h) + \frac{1}{\epsilon} g(\bar{y}_h)) \xi_h(x) \right) dx \\ &= \int_{\Omega} \frac{1}{\epsilon} \bar{v}_h(x) \xi_h(x) dx \quad \forall \xi_h \in Y_h.\end{aligned}$$

This is equivalent to

$$\int_{\Omega} \left(\sum_{i,j=1}^N a_{ij}(x) D_i \bar{y}_h D_j \xi_h + f(x, \bar{y}_h) \xi_h(x) \right) dx = \int_{\Omega} \bar{u}_h(x) \xi_h(x) dx \quad \forall \xi_h \in Y_h.$$

It follows that

$$\|\bar{y}_h\|_{1,2} \leq C \|\bar{u}_h\|_2 \quad \forall h > 0,$$

where $C > 0$ is independent of h . Hence $\{\bar{y}_h\}$ is bounded in $H_0^1(\Omega)$. Therefore, there exists a subsequence $\{(y'_h, u'_h)\}$ of $\{(\bar{y}_h, \bar{u}_h)\}$ such that $y'_h \rightharpoonup \bar{y}$ in $H_0^1(\Omega)$ and $u'_h \rightharpoonup \bar{u}$ in $L^2(\Omega)$. Let $y = y(\bar{u})$ be a solution of the equation

$$Ay + f(x, y) = \bar{u}$$

and $y'_h(u'_h)$ be a solution of the equation

$$\langle Ay'_h + f(\cdot, y'_h), \xi_h \rangle = \langle u'_h, \xi_h \rangle \quad \forall \xi_h \in Y_h.$$

By Proposition 4.3.17, we have

$$\begin{aligned} & \|y(\bar{u}) - y'_h(u'_h)\|_{\infty} \\ & \leq C(h^{2-N/2} \|y(u'_h)\|_{2,2} + \|\bar{u} - u'_h\|_2) \leq C(h^{2-N/2} \|u'_h\|_2 + \|\bar{u} - u'_h\|_2). \end{aligned}$$

This implies that $\{y'_h\}$ is bounded in $L^{\infty}(\Omega)$. Thus, there exists $M > 0$ such that $\|y'_h\|_{\infty} + \|u'_h\|_2 \leq M$. By (A3), one has

$$|v'_h(x)| = |g(y'_h(x)) + \epsilon u'_h(x)| \leq M k_{g,M} + |g(0)| + \epsilon |u'_h(x)| \quad \text{a.e. } x \in \Omega.$$

Hence $\|v'_h\|_2 \leq M_1$ for all $h > 0$. Consequently, $\{(y'_h, v'_h)\}$ is bounded in $(H_0^1(\Omega) \cap L^{\infty}(\Omega)) \times L^2(\Omega)$. Without loss of generality, we may assume that y'_h converges weakly to a function $\bar{y} \in H_0^1(\Omega) \cap L^{\infty}(\Omega)$ and v'_h converges weakly to $\bar{v} \in L^2(\Omega)$.

By the compact embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$, we have $y'_h \rightarrow \bar{y}$ in $L^2(\Omega)$. By (A3) and (A4), we can show that $f(\cdot, y'_h) + \frac{1}{\epsilon} g(y'_h)$ converges strongly to $f(\cdot, \bar{y}) + \frac{1}{\epsilon} g(\bar{y})$ in $L^2(\Omega)$.

We now take any $z \in C_0^{\infty}(\Omega)$. By Lemma 4.3.15 (see also [164, Theorem 9.10]), there exists $z_h \in Y_h$ such that $z_h \rightarrow z$ in $H_0^1(\Omega)$. It follows that

$$\int_{\Omega} [f(x, y'_h(x)) + \frac{1}{\epsilon} g(y'_h(x))] z_h(x) dx \rightarrow \int_{\Omega} [f(x, \bar{y}(x)) + \frac{1}{\epsilon} g(\bar{y}(x))] z(x) dx$$

when $h \rightarrow 0$. Since $(y'_h, v'_h) \in \Phi_h$, we have

$$a(y'_h, z_h) + (f(x, y'_h) + \frac{1}{\epsilon} g(y'_h), z_h) = (\frac{1}{\epsilon} v'_h, z_h) \quad \forall z_h \in Y_h$$

and

$$v'_h(x) \leq 0 \text{ a.e. } x \in \Omega.$$

Note that $\nabla y'_h \rightharpoonup \nabla \bar{y}$ in $L^2(\Omega)$, $\nabla z_h \rightarrow \nabla z$ in $L^2(\Omega)$ and the set $\{v \in L^2(\Omega) : v(x) \leq 0\}$ is weakly closed. Passing to the limit the above when $h \rightarrow 0$, we obtain

$$a(\bar{y}, z) + (f(x, \bar{y}) + \frac{1}{\epsilon}g(\bar{y}))z = (\frac{1}{\epsilon}\bar{v}, z) \quad \forall z \in C_0^\infty(\Omega), \quad (4.3.80)$$

$$a(\bar{y}, z) + f(x, \bar{y})z = (\bar{u}, z) \quad \forall z \in C_0^\infty(\Omega), \quad (4.3.81)$$

and

$$\bar{v}(x) \leq 0 \text{ a.e. } x \in \Omega.$$

Since $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, we see that (4.3.81) is valid for all $z \in H_0^1(\Omega)$. Hence we have $(\bar{y}, \bar{v}) \in \Phi'$. The remain of the proof is divided into some steps.

Step 1. Showing that $(\bar{y}, \bar{v})$ is a minimizer of (P') .

By Proposition 4.3.3, (P) has an optimal solution (y_0, u_0) . By Proposition 4.3.4, we have $u_0 \in H^1(\Omega) \cap C(\bar{\Omega})$ and $y_0 \in H^2(\Omega) \cap H_0^1(\Omega)$. Let $v_0 = g(y_0) + \epsilon u_0$. Then $v_0 \in H^1(\Omega)$ because g is of class C^1 and locally Lipschitz continuous (see Theorem 3.1 in [171]). By Proposition 4.3.6, (y_0, v_0) is a minimizer to (P') . Let $v_h = \Pi_h v_0$ and $y_h = y_h(v_h) = S_h(v_h)$. Then $v_h(x) = \sum_{i=1}^{k_h} v_0(N_i)\zeta_i(x) \leq 0$ and so $(y_h, v_h) \in \Phi'_h$. Applying Lemma 4.3.15 for $k = 0, m = 0, p = q = 2$, we have

$$\|v_h - v_0\|_2 \leq Ch\|v_0\|_{1,2}.$$

By Proposition 4.3.17, we have

$$\|y_h(v_h) - y_0(v_0)\|_\infty \leq C(h^{2-N/2}\|y(v_h)\|_{2,2} + \|v_h - v_0\|_2).$$

Hence (y_h, v_h) converges strongly to (y_0, v_0) in $L^\infty(\Omega) \times L^2(\Omega)$.

Using (A1), $y'_h \rightarrow \bar{y}$, $v'_h \rightharpoonup \bar{v}$ in $L^2(\Omega)$, the weakly lower continuity of the mapping $v \mapsto \int_\Omega \frac{1}{2}v^2(x)dx$ and $g(y'_h) \rightarrow g(\bar{y})$ strongly in $L^2(\Omega)$, we can show that

$$J_0(\bar{y}, \bar{v}) \leq \liminf_{h \rightarrow 0} J_0(y'_h, v'_h) \leq \lim_{h \rightarrow 0} J_0(y_h, v_h) = J_0(y_0, v_0).$$

Hence $(\bar{y}, \bar{v})$ is a minimizer of (P') .

Step 2. Showing that $\lim_{h \rightarrow 0} \|v'_h - \bar{v}\|_2 = 0$.

By the same argument as Step 1, we can show that there exists a sequence $\{(\hat{y}_h, \hat{v}_h)\} \subset \Phi'_h$ such that $(\hat{y}_h, \hat{v}_h) \rightarrow (\bar{y}, \bar{v})$ strongly in $L^\infty(\Omega) \times L^2(\Omega)$. It follows that

$$J_0(\bar{y}, \bar{v}) \leq \liminf_{h \rightarrow 0} J_0(y'_h, v'_h) \leq \limsup_{h \rightarrow 0} J(y'_h, v'_h) \leq \limsup_{h \rightarrow 0} J_0(\hat{y}_h, \hat{v}_h) = J_0(\bar{y}, \bar{v}).$$

Hence

$$J_0(\bar{y}, \bar{v}) = \lim_{h \rightarrow 0} J_0(y'_h, v'_h).$$

This implies that $\|v'_h\|_2^2 \rightarrow \|\bar{v}\|_2^2$. Combining this with the fact that $v'_h \rightharpoonup \bar{v}$, we conclude that $\|v'_h - \bar{v}\|_2^2 \rightarrow 0$ as $h \rightarrow 0^+$.

Step 3. Showing that $\|v'_h - \bar{v}\|_\infty \rightarrow 0$ and $\|y'_h - \bar{y}\|_\infty \rightarrow 0$.

Let $(\bar{\psi}_h, \bar{w}_h)$ be the Lagrange multiplier corresponding to (y'_h, v'_h) and $(\bar{\psi}, \bar{w})$ be the Lagrange multiplier corresponding to $(\bar{y}, \bar{v})$. By Proposition 4.3.17, we see that $y'_h(v'_h)$ converges strongly to $\bar{y} = \bar{y}(\bar{v})$ in $H_0^1(\Omega) \cap L^\infty(\Omega)$. Since y'_h is bounded in $C(\bar{\Omega})$ and v'_h is bounded in $L^2(\Omega)$, there exists a constant $M_1 > 0$ such that $\|v'_h\|_2 + \|\bar{v}\|_2 \leq M_1$ for all $h > 0$. Combining this with Propositions 4.3.1, 4.3.18 and 4.3.19, we find out a positive constant M_2 and h_0 such that

$$\|y'_h\|_\infty, \|\bar{y}\|_\infty, \|\bar{\psi}_h\|_\infty, \|\bar{\psi}\|_\infty \leq M_2$$

for all $h < h_0$. This implies that

$$\|\epsilon \bar{\psi} + g(\bar{y})\|_\infty, \|\epsilon \bar{\psi}_h + g(y'_h)\|_\infty \leq M_3 \quad \forall h < h_0, \quad (4.3.82)$$

where $M_3 := \epsilon M_2 + k_{g, M_2} M_2 + \|g(0)\|_\infty$. By (4.3.44), we have

$$\epsilon^2 \bar{w}_h(x) \in N((-\infty, 0], v'_h(x)).$$

From this and (4.3.43) and we get

$$\epsilon \bar{\psi}_h(x) + g(y'_h(x)) - v'_h(x) \in N((-\infty, 0], v'_h(x)) \quad \text{a.e. } x \in \Omega.$$

Hence

$$v'_h(x) = P_{(-\infty, 0]}(\epsilon \bar{\psi}_h(x) + g(y_h(x))) \quad \text{a.e. } x \in \Omega. \quad (4.3.83)$$

Similarly, from (4.3.28) and (4.3.29), we have

$$\bar{v}(x) = P_{(-\infty, 0]}(\epsilon \bar{\psi}(x) + g(\bar{y}(x))) \quad \text{a.e. } x \in \Omega. \quad (4.3.84)$$

Since the metric projections are non-expansive, we have from (4.3.83) and (4.3.84) that

$$\begin{aligned} |v'_h(x) - \bar{v}(x)| &= |P_{(-\infty, 0]}(\epsilon \bar{\psi}_h(x) + g(y_h(x))) - P_{(-\infty, 0]}(\epsilon \bar{\psi}(x) + g(\bar{y}(x)))| \\ &\leq \epsilon |\bar{\psi}_h(x) - \bar{\psi}(x)| + |g(y'_h(x)) - g(\bar{y}(x))| \\ &\leq \epsilon |\bar{\psi}_h(x) - \bar{\psi}(x)| + k_{g, M_2} |y'_h(x) - \bar{y}(x)| \end{aligned}$$

for a.e. $x \in \Omega$ and for all $h < h_0$. Hence

$$\|v'_h - \bar{v}\|_\infty \leq \epsilon \|\bar{\psi}_h - \bar{\psi}\|_\infty + k_{g,M_2} \|y'_h - \bar{y}\|_\infty,$$

for all $h < h_0$. From this, Propositions 4.3.17 and 4.3.19, we get

$$\|v'_h - \bar{v}\|_\infty \leq C\epsilon(h^{2-\frac{N}{2}} + \|v'_h - \bar{v}\|_2) + k_{g,M_2} C(h^{2-N/2} \|y(v'_h)\|_{2,2} + \|v'_h - \bar{v}\|_2). \quad (4.3.85)$$

Since $\lim_{h \rightarrow 0} \|v'_h - \bar{v}\|_2 = 0$, we obtain

$$\lim_{h \rightarrow 0} \|v'_h - \bar{v}\|_\infty = 0.$$

It is clear that $\frac{1}{\epsilon}(v'_h - g(y'_h)) \rightarrow \frac{1}{\epsilon}(\bar{v} - g(\bar{y}))$ strongly in $L^\infty(\Omega)$. By Proposition 4.3.6, $(\bar{y}, \frac{1}{\epsilon}(\bar{v} - g(\bar{y})))$ is a minimizer of (P) .

Finally, if $(\bar{y}, \bar{u})$ is a unique solution of (P) then Proposition 4.3.6 implies that $(\bar{y}, g(\bar{y}) + \epsilon \bar{u})$ is a unique solution to (P') and $\bar{v} = g(\bar{y}) + \epsilon \bar{u}$. By a contradiction argument, we can show that the total sequence $(\bar{y}_h, \bar{v}_h)$ converges strongly to $(\bar{y}, \bar{v})$ in $(H_0^1(\Omega) \cap L^\infty(\Omega)) \times L^\infty(\Omega)$. Consequently, the total sequence $(\bar{y}_h, \frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h))) \rightarrow (\bar{y}, \bar{u})$ in $(H_0^1(\Omega) \cap L^\infty(\Omega)) \times L^\infty(\Omega)$. The theorem is proved. $\square$

To obtain error estimates, we need the following lemma which is a version of Theorem 2.14 in [46] for problems (P'') and (P''_h) . Note that Theorem 2.14 in [46] was established under assumptions (A2), (A3) (A7) and the constraint set is bounded. However, problems (P'') and (P''_h) do not satisfy assumptions (A2) and (A7) in [46] because the constraint set Φ'' is unbounded.

Lemma 4.3.21. *Suppose that (A1) – (A4) are satisfied, $\{\bar{v}_h\}$ is a sequence of local minimizers of (P''_h) , which converges strongly in $L^2(\Omega)$ to a local minimizer $\bar{v}$ of (P'') under which the strictly second-order condition (4.3.36) is valid. Then there exists constants $M > 0$, $C_M = C_M(\epsilon) > 0$, $C > 0$ and $h_0 > 0$ such that*

$$\epsilon_h := \sup_{v \in B_U(0, M)} \|\hat{J}'_h(v) - \hat{J}'(v)\| \leq C_M h^2 \quad (4.3.86)$$

and

$$\|\bar{v}_h - \bar{v}\|_2 \leq C[\epsilon_h^2 + \|\bar{v} - v_h\|_2^2 + \hat{J}'(\bar{v})(v_h - \bar{v})]^{1/2} \quad \forall v_h \in \Phi''_h, h < h_0. \quad (4.3.87)$$

Proof. We mimic the proof of Theorem 2.14 in [46]. Since $\bar{v}_h \rightarrow \bar{v}$ strongly, there exists $M > 0$ and h_0 such that $\|\bar{v}_h\|_2 \leq M$ for all $h < h_0$. Let $\epsilon_h := \sup_{v \in B_U(0, M)} \|\hat{J}'_h(v) - \hat{J}'(v)\|$. Fix any $v \in B_U(0, M)$. Let $y_h(v)$ and $y(v)$ are solutions of (4.3.49) and (4.3.50) corresponding to v , respectively; $\psi(v)$ and $\psi_h(v)$ are solutions of (4.3.57) and (4.3.58) corresponding to v , respectively. By Proposition 4.3.16, we have

$$\|y_h(v) - y(v)\|_\infty \leq C_M h^{2-N/2} \|y(v)\|_{2,2} \leq C_M h^{2-N/2} \|v\|_2.$$

Hence $y(v)$ and $y_h(v)$ are bounded in $L^\infty(\Omega)$. This implies that there is a number $M_0 > 0$ such that $\|y(v)\|_\infty + \|y_h(v)\|_\infty \leq M_0$. By (A3), we have

$$|g(y(v))(x) - g(y_h(v))(x)| \leq k_{g, M_0} |y(v)(x) - y_h(v)(x)|.$$

Hence

$$\|g(y(v)) - g(y_h(v))\|_2 \leq k_{g, M} \|y(v) - y_h(v)\|_2 \leq C' k_{g, M_0} \|y(v)\|_{2,2} h^2. \quad (4.3.88)$$

By Proposition 4.3.18, we have

$$\|\psi(v) - \psi_h(v)\|_2 \leq C'_M h^2. \quad (4.3.89)$$

Besides, for any $w \in L^2(\Omega)$, we have

$$\begin{aligned} & |(\hat{J}'_h(v) - \hat{J}'(v))w| \\ &= \left| \int_\Omega (L_v(x, y_h(v), v) - \frac{1}{\epsilon} \psi_h(v) - L_v(x, y(v), v) + \frac{1}{\epsilon} \psi(v)) w dx \right| \\ &\leq \int_\Omega \left| \frac{1}{\epsilon^2} (g(y_h(v)) - g(y(v))) + \frac{1}{\epsilon} (\psi(v) - \psi_h(v)) \right| |w| dx, \\ &\leq \left(\frac{1}{\epsilon^2} \|g(y_h(v)) - g(y(v))\|_2 + \frac{1}{\epsilon} \|\psi(v) - \psi_h(v)\|_2 \right) \|w\|_2. \end{aligned}$$

Combining this with (4.3.88) and (4.3.89) yields

$$\epsilon_h \leq \frac{1}{\epsilon^2} \|g(y_h(v)) - g(y(v))\|_2 + \frac{1}{\epsilon} \|\psi(v) - \psi_h(v)\|_2 \leq C_M h^2 \quad (4.3.90)$$

for some constant $C_M = C_M(\epsilon) > 0$ which is independent of h . We obtain (4.3.86).

From (4.3.24) and (4.3.57), we can show that, the map $v \mapsto \psi(v)$ is Lipschitz continuous. Hence there exists a constant $C > 0$ such that

$$|\hat{J}'(\bar{v}_h) - \hat{J}'(\bar{v})|(v_h - \bar{v})| \leq C \|\bar{v}_h - \bar{v}\|_2 \|v_h - \bar{v}\|_2.$$

Let $\bar{\psi}$ and $\bar{\psi}_h$ be Lagrange multipliers corresponding to $\bar{v}$ and $\bar{v}_h$. Since $\bar{v}_h$ is a local minimizer, we have

$$\hat{J}'_h(\bar{v}_h)(v_h - \bar{v}_h) \geq 0 \quad \forall v_h \in \Phi''_h. \quad (4.3.91)$$

This implies that

$$\begin{aligned} & \hat{J}'(\bar{v}_h)(\bar{v} - \bar{v}_h) + [\hat{J}'_h(\bar{v}_h) - \hat{J}'(\bar{v}_h)](\bar{v} - \bar{v}_h) + \\ & + [\hat{J}'_h(\bar{v}_h) - \hat{J}'(\bar{v})](v_h - \bar{v}) + \hat{J}'(\bar{v})(v_h - \bar{v}) \geq 0. \end{aligned} \quad (4.3.92)$$

Since $\bar{v}$ is a local minimizer of (P'') , we have

$$\hat{J}'(\bar{v})(\bar{v}_h - \bar{v}) \geq 0. \quad (4.3.93)$$

Adding (4.3.92) and (4.3.93) yields

$$\begin{aligned} & [\hat{J}'(\bar{v}_h) - \hat{J}'(\bar{v})](\bar{v}_h - \bar{v}) \\ & \leq [\hat{J}'_h(\bar{v}_h) - \hat{J}'(\bar{v}_h)](\bar{v} - \bar{v}_h) + [\hat{J}'_h(\bar{v}_h) - \hat{J}'(\bar{v})](v_h - \bar{v}) + \hat{J}'(\bar{v})(v_h - \bar{v}) \\ & = [\hat{J}'_h(\bar{v}_h) - \hat{J}'(\bar{v}_h)](\bar{v} - \bar{v}_h) + [\hat{J}'_h(\bar{v}_h) - \hat{J}'(\bar{v}_h)](v_h - \bar{v}) + \\ & + [\hat{J}'(\bar{v}_h) - \hat{J}'(\bar{v})](v_h - \bar{v}) + \hat{J}'(\bar{v})(v_h - \bar{v}). \end{aligned}$$

This implies that

$$\begin{aligned} & [\hat{J}'(\bar{v}_h) - \hat{J}'(\bar{v})](\bar{v}_h - \bar{v}) \\ & \leq \epsilon_h \|\bar{v} - \bar{v}_h\|_2 + \epsilon_h \|v_h - \bar{v}\|_2 + C \|\bar{v}_h - \bar{v}\|_2 \|v_h - \bar{v}\|_2 + \hat{J}'(\bar{v})(v_h - \bar{v}) \end{aligned}$$

Since $\hat{J}$ is of class C^2 , there exists $\hat{v}_h \in [\bar{v}_h, \bar{v}]$ such that

$$[\hat{J}'(\bar{v}_h) - \hat{J}'(\bar{v})](\bar{v}_h - \bar{v}) = \hat{J}''(\hat{v}_h)(\bar{v}_h - \bar{v})^2.$$

Since $\|\bar{v}_h - \bar{v}\|_2 \rightarrow 0$, we have $\|\hat{v}_h - \bar{v}\|_2 \rightarrow 0$ and so $\|\hat{v}_h\|_2 \leq M$ for all $h < h_0$. Consequently,

$$\begin{aligned} & \hat{J}''(\hat{v}_h)(\bar{v}_h - \bar{v})^2 \\ & \leq \epsilon_h \|\bar{v} - \bar{v}_h\|_2 + \epsilon_h \|v_h - \bar{v}\|_2 + C \|\bar{v}_h - \bar{v}\|_2 \|v_h - \bar{v}\|_2 + \hat{J}'(\bar{v})(v_h - \bar{v}). \end{aligned} \quad (4.3.94)$$

Let us claim that, there exists $C > 0$ such that (4.3.87) is fulfilled. By contradiction, there exists a subsequence $\{h_k\}$ converging to 0 and $v_{h_k} \in \Phi''_{h_k}$ such that

$$\|\bar{v}_{h_k} - \bar{v}\|_2^2 > k[\epsilon_{h_k}^2 + \|\bar{v} - v_{h_k}\|_2^2 + \hat{J}'(\bar{v})(v_{h_k} - \bar{v})]$$

or equivalently,

$$\frac{1}{k} > \frac{\epsilon_{h_k}^2}{\|\bar{v}_{h_k} - \bar{v}\|_2} + \frac{\|\bar{v} - v_{h_k}\|_2^2}{\|\bar{v}_{h_k} - \bar{v}\|_2^2} + \frac{\hat{J}'(\bar{v})(v_{h_k} - \bar{v})}{\|\bar{v} - v_{h_k}\|_2^2}. \quad (4.3.95)$$

Put $w_{h_k} = \frac{\bar{v}_{h_k} - \bar{v}}{\|\bar{v}_{h_k} - \bar{v}\|_2}$. Then $\|w_{h_k}\|_2 = 1$. We may assume that w_{h_k} converges weakly to w . We want to show that $w \in \mathcal{C}''_{\bar{v}}$. Since

$$w_{h_k} \in \text{cone}(\Phi'' - \bar{v}),$$

we have $w \in T(\Phi'', \bar{v})$. It remain to show that $\hat{J}'(\bar{v})w = 0$. Since $\bar{v}$ is a local minimizer, we have $\hat{J}'(\bar{v})w \geq 0$. On the other hand,

$$\begin{aligned} \hat{J}'(\bar{v})w &= \lim_{k \rightarrow \infty} \hat{J}'(\bar{v})w_{h_k} = \\ &= \lim_{k \rightarrow \infty} \hat{J}'_{h_k}(\bar{v}_{h_k})w_{h_k} + \lim_{k \rightarrow \infty} [\hat{J}'(\bar{v}_{h_k}) - \hat{J}'_{h_k}(\bar{v}_{h_k})]w_{h_k} + \lim_{k \rightarrow \infty} [\hat{J}'(\bar{v}) - \hat{J}'(\bar{v}_{h_k})]w_{h_k}. \end{aligned}$$

The first limits exists because $w_{h_k} \rightharpoonup w$. The second limit and the third limit equal to zero as follows.

$$\lim_{k \rightarrow \infty} |[\hat{J}'(\bar{v}_{h_k}) - \hat{J}'_{h_k}(\bar{v}_{h_k})]w_{h_k}| \leq \lim_{k \rightarrow \infty} \epsilon_{h_k} \|w_{h_k}\|_2 = 0$$

and

$$\lim_{k \rightarrow \infty} |[\hat{J}'(\bar{v}) - \hat{J}'(\bar{v}_{h_k})]w_{h_k}| \leq \lim_{k \rightarrow \infty} M_1 \|\bar{v} - \bar{v}_{h_k}\|_2 \|w_{h_k}\|_2 = 0.$$

Thus we have shown that $\hat{J}'(\bar{v})w \leq \lim_{k \rightarrow \infty} \hat{J}'_{h_k}(\bar{v}_{h_k})w_{h_k}$. Using

$$w_{h_k} = \frac{\bar{v}_{h_k} - \bar{v}}{\|\bar{v}_{h_k} - \bar{v}\|_2},$$

(4.3.91) and (4.3.95), we have

$$\begin{aligned} \hat{J}'(\bar{v})w &\leq \lim_{k \rightarrow \infty} \hat{J}'_{h_k}(\bar{v}_{h_k})w_{h_k} \\ &= \lim_{k \rightarrow \infty} \frac{1}{\|\bar{v}_{h_k} - \bar{v}\|_2} [\hat{J}'_{h_k}(\bar{v}_{h_k})(v_{h_k} - \bar{v}) + \hat{J}'_{h_k}(\bar{v}_{h_k})(\bar{v}_{h_k} - v_{h_k})] \\ &\leq \lim_{k \rightarrow \infty} \frac{1}{\|\bar{v}_{h_k} - \bar{v}\|_2} [\hat{J}'_{h_k}(\bar{v}_{h_k})(v_{h_k} - \bar{v})] \\ &\leq \lim_{k \rightarrow \infty} (M_2 + \epsilon_{h_k}) \frac{\|v_{h_k} - \bar{v}\|_2}{\|\bar{v}_{h_k} - \bar{v}\|_2} = 0. \end{aligned}$$

Here we used the fact that

$$\|\hat{J}'_{h_k}(\bar{v}_{h_k})\| \leq \|\hat{J}'_{h_k}(\bar{v}_{h_k}) - \hat{J}'(\bar{v}_{h_k})\| + \|\hat{J}'(\bar{v}_{h_k})\| \leq \epsilon_{h_k} + M_2.$$

In a summary, we have shown that $w \in \mathcal{C}''_{\bar{v}}$. We now have

$$\begin{aligned} \liminf_{k \rightarrow \infty} \hat{J}''(\hat{v}_{h_k})w_{h_k}^2 &= \liminf_{k \rightarrow \infty} [\hat{J}''(\bar{v})w_{h_k}^2 + (\hat{J}''(\hat{v}_{h_k}) - \hat{J}''(\bar{v}))w_{h_k}^2] \\ &\geq \liminf_{k \rightarrow \infty} \hat{J}''(\bar{v})w_{h_k}^2 \geq \hat{J}''(\bar{v})w^2. \end{aligned}$$

Combining this with (4.3.94) and (4.3.95), we obtain

$$\begin{aligned} \hat{J}''(\bar{v})w^2 &\leq \liminf_{k \rightarrow \infty} \hat{J}''(\hat{v}_{h_k})w_{h_k}^2 \leq \limsup_{k \rightarrow \infty} \hat{J}''(\hat{v}_{h_k})w_{h_k}^2 \leq \\ \lim_{k \rightarrow \infty} \left\{ \frac{\epsilon_{h_k}}{\|\bar{v}_{h_k} - \bar{v}\|_2} \left(1 + \frac{\|v_{h_k} - \bar{v}\|_2}{\|\bar{v}_{h_k} - \bar{v}\|_2}\right) + C \frac{\|v_{h_k} - \bar{v}\|_2}{\|\bar{v}_{h_k} - \bar{v}\|_2} + \frac{\hat{J}''(\bar{v})(v_{h_k} - \bar{v})}{\|\bar{v}_{h_k} - \bar{v}\|_2^2} \right\} &= 0. \end{aligned} \quad (4.3.96)$$

Here we used the fact that the map $v \mapsto \hat{J}''(v)$ is continuous and the map $w \mapsto \hat{J}''(\bar{v})w^2$ is weakly lower semicontinuous.

By the strictly second-order condition (4.3.36), we must have $w = 0$. By (4.3.34), we get

$$\begin{aligned} \hat{J}''(\bar{v})w_{h_k}^2 &= \int_{\Omega} \left(L_{yy}[x]z_{\bar{v},w_{h_k}}^2 + 2L_{yv}[x]z_{\bar{v},w_{h_k}}w_{h_k} + L_{vv}[x]w_{h_k}^2 + \psi_0(f_{yy}[x] + \right. \\ &\quad \left. + \frac{1}{\epsilon}g''[x])z_{\bar{v},w_{h_k}}^2(x) \right) dx \\ &= \int_{\Omega} \left((\phi_{yy}[x] + \frac{1}{\epsilon^2}(g'[x]^2 + g[x]g''[x])z_{\bar{v},w_{h_k}}^2 - g'[x]z_{\bar{v},w_{h_k}}w_{h_k} + \right. \\ &\quad \left. + \frac{1}{\epsilon^2}w_{h_k}^2 + \psi_0(f_{yy}[x] + \frac{1}{\epsilon}g''[x])z_{\bar{v},w_{h_k}}^2(x) \right) dx, \end{aligned}$$

where $z_{\bar{v},w_{h_k}}$ is a solution of the equation:

$$Az_{\bar{v},w_{h_k}} + [f_y[x] + \frac{1}{\epsilon}g'[x]]z_{\bar{v},w_{h_k}} = \frac{1}{\epsilon}w_{h_k}.$$

Since $\|z_{\bar{v},w_{h_k}}\|_{2,2} \leq C\|w_{h_k}\|_2$, we may assume that $z_{\bar{v},w_{h_k}} \rightharpoonup z$ in $H^2(\Omega)$. It is easy to see that $z = z_{\bar{v},w}$. Since $w = 0$, $z = 0$. By the compact embedding $H^2(\Omega) \hookrightarrow C(\bar{\Omega})$, $z_{\bar{v},w_{h_k}} \rightarrow 0$ strongly in $C(\bar{\Omega})$. Thus we have

$$\lim_{k \rightarrow \infty} \hat{J}''(\bar{v})w_{h_k}^2 = \frac{1}{\epsilon^2}.$$

However, by using the fact $\|\hat{v}_{h_k} - \bar{v}\|_2 \rightarrow 0$ and (4.3.96), we get

$$\begin{aligned} \frac{1}{\epsilon^2} &= \lim_{k \rightarrow \infty} \hat{J}''(\bar{v})w_{h_k}^2 = \lim_{k \rightarrow \infty} \hat{J}''(\hat{v}_{h_k})w_{h_k}^2 + \lim_{k \rightarrow \infty} [\hat{J}''(\bar{v}) - \hat{J}''(\hat{v}_{h_k})]w_{h_k}^2 \\ &= \lim_{k \rightarrow \infty} \hat{J}''(\hat{v}_{h_k})w_{h_k}^2 = 0 \end{aligned}$$

which is impossible. The lemma is proved. $\square$

Theorem 4.3.22. Suppose that $\{(\bar{y}_h, \bar{v}_h)\}$ is a sequence of local minimizers to (P'_h) and $\bar{u}_h := \frac{1}{\epsilon}(\bar{v}_h - g(\bar{y}_h))$ such that $(\bar{y}_h, \bar{u}_h)$ converges to a local minimizer $(\bar{y}, \bar{u})$ to (P) in $(H_0^1(\Omega) \cap L^\infty(\Omega)) \times L^\infty(\Omega)$ under which the strictly second-order condition (4.3.8) is satisfied. Then there exist positive constants C'_5 and C'_6 and h^* which are independent of h such that

$$\|\bar{u}_h - \bar{u}\|_\infty \leq C'_5 h^{2-\frac{N}{2}}, \quad (4.3.97)$$

$$\|\bar{y}_h - \bar{y}\|_\infty \leq C'_6 h^{2-\frac{N}{2}} \quad (4.3.98)$$

for all $0 < h < h^*$.

Proof. Put $\bar{v} = g(\bar{y}) + \epsilon \bar{u}$. Then $\bar{v}_h = g(\bar{y}_h) + \epsilon \bar{u}_h \rightarrow \bar{v}$ in $L^\infty(\Omega)$. By Proposition 4.3.6, $(\bar{y}, \bar{v})$ is a minimizer to (P') and $\bar{v}$ is a minimizer of (P'') . Besides, $\{\bar{v}_h\}$ are minimizers to (P'_h) . Let $(\bar{\varphi}, \bar{e})$ be the Lagrange multipliers corresponding to $(\bar{y}, \bar{u})$ under which (4.3.8) is satisfied. By Proposition 4.3.11, (P') satisfies the strictly second-order condition (4.3.30) at $(\bar{y}, \bar{v})$ corresponding to $(\bar{\psi}, \bar{w}) = (\bar{\varphi}, \bar{e})$ and (P'') satisfies the strictly second-order condition (4.3.36) at $\bar{v}$ corresponding to $(\bar{\psi}, \bar{w})$.

By Lemma 4.3.21, there exists $h_0 > 0$, $C_M > 0$ and $C > 0$ such that

$$\|\bar{v}_h - \bar{v}\|_2 \leq C[C_M h^2 + \|\bar{v} - v_h\|_2^2 + \hat{J}'(\bar{v})(v_h - \bar{v})]^{1/2} \quad \forall v_h \in \Phi''_h, h < h_0. \quad (4.3.99)$$

Let $v_h = \Pi_h \bar{v}$. Then for any $x \in \Omega$, there exists a triangle/tetrahedron $N_1 N_2 \dots N_m$ containing x , where $m = N + 1$. So $x = \sum_{i=1}^m \lambda_i N_i$ with $\sum_{i=1}^m \lambda_i = 1$ and $\lambda_i \geq 0$ for $i = 1, \dots, m$. Since v_h interpolates $\bar{v}$, we have

$$v_h(x) = \sum_{i=1}^m \lambda_i v_h(N_i) = \sum_{i=1}^m \lambda_i \bar{v}(N_i).$$

Since $\bar{v}$ is Hölder continuous, we have

$$|v_h(x) - \bar{v}(x)| = \left| \sum_{i=1}^m [\lambda_i (\bar{v}(N_i) - \bar{v}(x))] \right| \leq k_{\bar{v}} \sum_{i=1}^m |N_i - x|^\nu \leq m k_{\bar{v}} h^\nu,$$

where $\nu = 2 - N/2$. This implies that $\|v_h - \bar{u}\|_\infty \leq m k_{\bar{v}} h^\nu$ and $\|v_h - \bar{v}\|_2 \leq m k_{\bar{v}} |\Omega| h^\nu$. Inserting $v_h = \Pi_h \bar{v}$ into (4.3.99), we get

$$\|\bar{v}_h - \bar{v}\|_2 \leq C[C_M h^2 + C h^{2\nu} + \hat{J}'(\bar{v})(\Pi_h \bar{v} - \bar{v})]^{1/2} \quad \forall h < h_0. \quad (4.3.100)$$

Let us estimate $\hat{J}'(\bar{v})(\Pi_h \bar{v} - \bar{v})$. Recall that

$$\begin{aligned}\hat{J}'(\bar{v})(\Pi_h \bar{v} - \bar{v}) &= \int_{\Omega} [L_v[x] - \frac{1}{\epsilon} \bar{\varphi}(x)] (\Pi_h \bar{v} - \bar{v})(x) dx \\ &= \int_{\Omega} [\frac{1}{\epsilon^2} (\bar{v} - g[x]) - \frac{1}{\epsilon} \bar{\varphi}(x)] (\Pi_h \bar{v}(x) - \bar{v}(x)) dx\end{aligned}$$

Let us put

$$d(x) = L_v[x] - \frac{1}{\epsilon} \bar{\varphi}(x) = \frac{1}{\epsilon^2} (\bar{v}(x) - g[x]) - \frac{1}{\epsilon} \bar{\varphi}(x).$$

Then d is Hölder continuous with order $\nu = 2 - N/2$. Let $T_h = \{K_1, K_2, \dots, K_{q_h}\}$ be a triangulation of Ω and

$$T_h^0 = \{T \in T_h \mid d(x) \neq 0 \quad \forall x \in T\}.$$

Take $T \in T_h^0$. By Remark 4.3.10, we see that $d(x) < 0$ for all $x \in T$. Hence $\bar{v}(x) = 0$ for all $x \in T$. Consequently, $\Pi_h \bar{v}(x) = 0$ for all $x \in T$. Let $x_T \in T$ such that $d(x_T) = 0$ for $T \in T_h \setminus T_h^0$. Then we have

$$\begin{aligned}\hat{J}'(\bar{v})(\Pi_h \bar{v} - \bar{v}) &= \sum_{T \in T_h} \int_T d(x) (\Pi_h \bar{v}(x) - \bar{v}(x)) dx \\ &= \sum_{T \in T_h \setminus T_h^0} \int_T d(x) (\Pi_h \bar{v}(x) - \bar{v}(x)) dx \\ &= \sum_{T \in T_h \setminus T_h^0} \int_T (d(x) - d(x_T)) (\Pi_h \bar{v}(x) - \bar{v}(x)) dx \\ &\leq \sum_{T \in T_h \setminus T_h^0} k_d h^\nu \int_T |\Pi_h \bar{v}(x) - \bar{v}(x)| dx \leq k_d h^\nu \|\Pi_h \bar{v} - \bar{v}\|_2 \\ &\leq k_d h^\nu m_{\bar{v}} h^\nu \leq Ch^{2\nu}.\end{aligned}\tag{4.3.101}$$

Plugging this into (4.3.100), we obtain

$$\|\bar{v}_h - \bar{v}\|_2 \leq C[C_M h^2 + Ch^{2\nu} + Ch^{2\nu}]^{1/2} \leq Ch^\nu, \quad \forall h < h_0. \tag{4.3.102}$$

By Proposition 4.3.17, one gets

$$\begin{aligned}\|\bar{y}_h(\bar{v}_h) - \bar{y}(\bar{v})\|_\infty &\leq C(h^{2-N/2} \|\bar{y}(\bar{v}_h)\|_{2,2} + \|\bar{v}_h - \bar{v}\|_2) \\ &\leq Ch^{2-N/2} \|\bar{v}_h\|_2 + Ch^{2-N/2} \leq C'_6 h^{2-N/2}.\end{aligned}$$

From (4.3.28) and (4.3.29), we have

$$\bar{v}(x) = P_{(-\infty, 0]}(\epsilon \bar{\psi}(x) + g(\bar{y}(x))).$$

From (4.3.43) and (4.3.44), we have

$$\bar{v}_h(x) = P_{(-\infty, 0]}(\epsilon \bar{\psi}_h(x) + g(\bar{y}_h(x))).$$

It follows that

$$\begin{aligned} |\bar{v}_h(x) - \bar{v}(x)| &\leq \epsilon |\bar{\psi}(x) - \bar{\psi}_h(x)| + |g(\bar{y}(x)) - g(\bar{y}_h(x))| \\ &\leq \epsilon \|\bar{\psi} - \bar{\psi}_h\|_\infty + k_{g,M} \|\bar{y}(\bar{u}) - \bar{y}_h(\bar{v}_h)\|_\infty \leq C''' h^{2-N/2}. \end{aligned}$$

Hence $\|\bar{v}_h - \bar{v}\|_\infty \leq C''' h^{2-N/2}$. This implies that

$$\begin{aligned} \|\bar{u}_h - \bar{u}\|_\infty &= \frac{1}{\epsilon} \|\bar{v}_h - g(\bar{y}_h) - \bar{v} + g(\bar{y})\|_\infty \\ &\leq \frac{1}{\epsilon} (\|\bar{v}_h - \bar{v}\|_\infty + k_{g,M} \|\bar{y}_h - \bar{y}\|_\infty) \leq C'_5 h^{2-N/2}. \end{aligned}$$

The proof of the theorem is complete. $\square$

Remark 4.3.23. When $N = 2$, $g = 0$, $\phi(x, y) = |y - y_0(x)|^2$ and the state equation is linear, (P') is a strongly convex problem. In this case Meyer and Rösch [162] showed that $\|\bar{u}_h - \bar{u}\|_2 = O(h^2)$. To do this they choose finite element space U_h consists of piecewise constant functions u_h . After solving a fully discretized optimal control problem, an optimal control $\bar{u}_h$ is calculated by the projection of the adjoint state p_h (see Lemma 3.5 in [162] and Section 3.1 in [104]). However, in the present setting, problem (P') is non-convex and has a nonlinear state equation. Besides, the space U_h is different from the finite element space in [162]. Therefore, $\bar{u}_h$ is not calculated by the projection of the adjoint state. Consequently, we do not know whether or not $\|\bar{u}_h - \bar{u}\|_2 = O(h^2)$. Nevertheless, we hope that $\|\bar{u}_h - \bar{u}\|_2 = O(h^2)$ as in the following numerical example.

4.3.5 Numerical example

In this section, we give a numerical example that illustrates our results. Let $\Omega := (0, 1)^2 = \{(x_1, x_2) : 0 < x_1, x_2 < 1\}$ be a unit square in $\mathbb{R}^2$. Consider the following optimal control problem

$$(P) \quad \begin{cases} J(y, u) = \frac{1}{2} \int_\Omega ((y(x) - y_d(x))^2 + u(x)^2) dx \rightarrow \min, \\ \text{s.t.} \\ -\Delta y + y = u \text{ in } \Omega, \quad y = 0 \quad \text{on } \Gamma, \\ y^3(x) + \epsilon u(x) \leq 0 \text{ a.e. } x \in \Omega, \end{cases}$$

where $y_d(x) := -10 \sin(\pi x_1) \sin(\pi x_2)$. It is easy to check that the problem satisfies assumptions (A1)–(A4) and so it has at least one minimizer. By changing variable $v = y^3 + \epsilon u$, we obtain the problem

$$(P') \quad \begin{cases} J_0(y, u) = \frac{1}{2} \int_{\Omega} [(y(x) - y_d(x))^2 + \frac{1}{\epsilon^2} (v(x) - y^3(x))^2] dx \rightarrow \min, \\ \text{s.t.} \\ -\Delta y + y + \frac{1}{\epsilon} y^3 = \frac{1}{\epsilon} v \text{ in } \Omega, \quad y = 0 \quad \text{on } \Gamma, \\ v(x) \leq 0 \text{ a.e. } x \in \Omega. \end{cases}$$

Then the discrete version of (P') has the following form:

$$(P'_h) \quad \begin{cases} J_0(y_h, u_h) = \frac{1}{2} \int_{\Omega} [(y_h(x) - y_d(x))^2 + \frac{1}{\epsilon^2} (v_h(x) - y_h^3(x))^2] dx \rightarrow \min \\ \text{s.t.} \\ (y_h, v_h) \in Y_h \times U_h, \\ \int_{\Omega} [\nabla y_h \nabla \xi_h + (y_h + \frac{1}{\epsilon} y_h^3) \xi_h] dx = \frac{1}{\epsilon} \int_{\Omega} v_h \xi_h dx \quad \forall \xi_h \in Y_h, \\ v_h(x) \leq 0 \text{ for all } x \in \Omega. \end{cases}$$

Let $(\bar{y}_h, \bar{v}_h)$ be an optimal solution to (P'_h) . By Theorem 4.3.20, the approximate solution to (P) is $(\bar{y}_h, \frac{1}{\epsilon}(\bar{v}_h - \bar{y}_h^3))$.

We now create a uniform triangular Friedrich–Keller triangulation with $n_h \times n_h$ vertices which is the basis for a finite element discretization of the discrete optimal control problem (P'_h) and its optimality system. By virtue of Proposition 4.3.12 and 4.3.13, for each $h > 0$, $\bar{v}_h \in L^\infty(\Omega)$, $\bar{y}_h \in Y_h$, $\bar{\psi}_h \in Y_h$, and $\bar{w}_h \in L^2(\Omega)$ exist and satisfy

$$\begin{cases} \int_{\Omega} \nabla \bar{y}_h \cdot \nabla \xi_h + (\bar{y}_h + \frac{1}{\epsilon} \bar{y}_h^3) \xi_h dx = \frac{1}{\epsilon} \int_{\Omega} \bar{v}_h \xi_h dx, \\ \bar{v}_h(x) \leq 0, \\ \int_{\Omega} \nabla \bar{\psi}_h \cdot \nabla \xi_h + (1 + \frac{1}{\epsilon} \bar{y}_h^2) \bar{\varphi}_h \xi_h dx = - \int_{\Omega} [\bar{y}_h - y_d - \frac{3}{\epsilon^2} \bar{y}_h^2 (\bar{v}_h - \bar{y}_h^3)] \xi_h dx, \\ \bar{w}_h = \frac{1}{\epsilon} \bar{\psi}_h + \frac{1}{\epsilon^2} (\bar{y}_h^3 - \bar{v}_h), \\ \bar{w}_h(x) \in N((-\infty, 0], \bar{v}_h(x)) \end{cases} \quad (4.3.103)$$

for all $\xi_h \in Y_h$ and for all $x \in \Omega$. The two last conditions in (4.3.103) imply that

$$0 \geq \bar{w}_h(x)(s - \bar{w}_h(x)) = \frac{1}{\epsilon^2} (\epsilon \bar{\psi}_h(x) + \bar{y}_h^3(x) - \bar{v}_h(x))(s - \bar{v}_h(x))$$

for a.e. $x \in \Omega$ and for all $s \leq 0$. This, together with the projection theorem, yields

$$\bar{v}_h(x) = P_{(-\infty, 0]}(\epsilon \bar{\varphi}_h(x) + \bar{y}_h^3(x)) = \mathbb{1}_{\{\epsilon \bar{\varphi}_h(x) + \bar{y}_h^3(x) \geq 0\}} \quad (4.3.104)$$

and so

$$\bar{w}_h(x) = \frac{1}{\epsilon} \bar{\psi}_h(x) + \frac{1}{\epsilon^2} \bar{y}_h^3(x) - \frac{1}{\epsilon^2} \mathbb{1}_{\{\epsilon \bar{\varphi}_h(x) + \bar{y}_h^3(x) \geq 0\}}. \quad (4.3.105)$$

Here $\mathbb{1}_M$ denotes the characteristic function of a set M , i.e., $\mathbb{1}_M(x) = 1$ if $x \in M$ and $\mathbb{1}_M(x) = 0$ if $x \notin M$; $\{f \geq 0\} := \{x \in \Omega \mid f(x) \geq 0\}$ for a function f defined on Ω . Inserting (4.3.104) and (4.3.105) into (4.3.103), we derive the system

$$\begin{cases} \int_{\Omega} \nabla \bar{y}_h \cdot \nabla \xi_h + \bar{y}_h \xi_h dx = \frac{1}{\epsilon} \int_{\Omega} \mathbb{1}_{\{\epsilon \bar{\varphi}_h(x) + \bar{y}_h^3(x) \geq 0\}} \xi_h dx, \\ \int_{\Omega} \nabla \bar{\psi}_h \cdot \nabla \xi_h + (1 + \frac{1}{\epsilon} \bar{y}_h^2) \bar{\varphi}_h \xi_h dx = \\ - \int_{\Omega} [\bar{y}_h - y_d + \frac{3}{\epsilon^2} \bar{y}_h^3 - \frac{3}{\epsilon^2} \bar{y}_h^2 \mathbb{1}_{\{\epsilon \bar{\varphi}_h(x) + \bar{y}_h^3(x) \geq 0\}}] \xi_h dx, \end{cases} \quad (4.3.106)$$

for all $\xi_h \in Y_h$ and for a.e. $x \in \Omega$.

Since the coefficients of the above system are semi-smooth, it is reasonable to solve this system via a semi-smooth Newton method (SSN). The SSN is terminated when the L^2 -norm of the difference between of two consecutive iterations $(\bar{y}_{h,k}, \bar{\varphi}_{h,k})$ and $(\bar{y}_{h,k+1}, \bar{\varphi}_{h,k+1})$ is not greater than 10^{-14} . We used the Python implementation using DOLFIN to generate the numerical results. In all cases of different values of h , the parameter $\epsilon = 10^{-2}$ and the SSN solving (4.3.106) was started with the starting guess $(y_{h,0}, \psi_{h,0}) := (0, 0)$. Since we do not know the exact optimal control $\bar{v}$ and $\bar{u}$, we will take as reference control the numerical solution $\bar{v}_{h^*}$ calculated on the grid with mesh size $h^* = \frac{\sqrt{2}}{800}$. Therefore, instead of estimating the error $\|\bar{v}_h - \bar{v}\|_X$ with $X = L^2(\Omega)$ or $X = L^\infty(\Omega)$ as $h \rightarrow 0$, we study the error $\|\bar{v}_h - \bar{v}_{h^*}\|_X$. Note that $\|\bar{v}_h - \bar{v}\|_X = O(h)$ and $\|\bar{u}_h - \bar{u}\|_X = O(h)$. Therefore, we only to estimate $\|\bar{v}_h - \bar{v}\|_X$. The numerical values of this error for different values of n_h are displayed in Table 4.1, where we list the L^2 - and L^∞ -error estimates for the computed discrete control $\bar{u}_h$ as well as the number of SSN iterations. Setting the mesh sizes $h_n := \frac{\sqrt{2}}{100n}$ with $1 \leq n \leq 7$, we denote by $EOC(n)$ the L^∞ -experimental order of convergence (EOC) for the numerical control $\bar{u}_{h^*}$ with respect to two consecutive mesh sizes h_n and h_{n+1} , that is,

$$EOC(n) := \frac{\log(\|\bar{u}_{h^*} - \bar{u}_{h_{n+1}}\|_{L^\infty(\Omega)}) - \log(\|\bar{u}_{h^*} - \bar{u}_{h_n}\|_{L^\infty(\Omega)})}{\log(h_{n+1}) - \log(h_n)}.$$

Table 4.2 represents the EOC for the control $\bar{u}$.

According to Theorem 4.2, we expect that the convergence rate of optimal controls in the L^∞ -norms is of $O(h)$. However, a close look

n_h	$\ \bar{u}_h - \bar{u}_{h^*}\ _{L^2(\Omega)}$	$\ \bar{u}_h - \bar{u}_{h^*}\ _{L^\infty(\Omega)}$	#SSN
100	5.62×10^{-5}	1.11×10^{-4}	5
200	1.34×10^{-5}	2.65×10^{-5}	5
300	4.89×10^{-6}	1.08×10^{-5}	5
400	2.68×10^{-6}	5.31×10^{-6}	5
500	8.41×10^{-7}	2.76×10^{-6}	5
600	3.96×10^{-7}	1.65×10^{-6}	5
700	5.34×10^{-7}	1.82×10^{-6}	5

Table 4.1: numerical results: L^2 - and L^∞ -errors for control $\bar{u}$ in dependence of n_h , and the number of SSN iterations

h_n	$\rightarrow$	h_{n+1}	$EOC(n)$
1.41×10^{-2}	$\rightarrow$	7.07×10^{-3}	2.07
7.07×10^{-3}	$\rightarrow$	4.71×10^{-3}	2.21
4.71×10^{-3}	$\rightarrow$	3.54×10^{-3}	2.47
3.54×10^{-3}	$\rightarrow$	2.83×10^{-3}	2.93
2.83×10^{-3}	$\rightarrow$	2.36×10^{-3}	2.81
2.36×10^{-3}	$\rightarrow$	2.02×10^{-3}	-6.29×10^{-1}

Table 4.2: numerical results of $EOC(n)$

at Table 2 reveals that the convergence rate in the L^∞ - norms is Ch^2 which is much better than the predicted conclusion in Theorem 4.2. This also shows that the error estimate stated in Theorem 4.2 might not be optimal and it should be Ch^2 . The problem of finding the optimal rate of convergence for optimal control with mixed pointwise constraints is an interesting open question.

REFERENCES

- [1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
- [2] S. Adly and B.T. Kien, *KKT optimality conditions for multiobjective optimal control problems with endpoint and mixed constraints*, J. Optim. Theor. Appl., 2026, <https://doi.org/10.1007/s10957-026-03093-0>
- [3] W. Alt, R. Griesse, N. Metla and A. Rösch, *Lipschitz stability for elliptic optimal control problems with mixed control-state constraints*, Optimization, 59(2010), 833-849.
- [4] V.M. Alekseev, V.M. Tikhomirov and S.V. Formin, *Optimal Control*, Springer Science + Business Media LLC, 1987
- [5] N. Arada and J.-P. Raymond, *Control problems with mixed control-state constraints*, SIAM J. Control. Optim., 39(2000) 1391-1407.
- [6] N. Arada and J.-P. Raymond, *Dirichlet boundary control of semilinear parabolic equations. Part 1: Problems with no state constraints*, Appl. Math. Optim., 45(2002) 125-143.
- [7] N. Arada and J.-P. Raymond, *Dirichlet boundary control of semilinear parabolic equations. Part 2: Problems with pointwise state constraints*, Appl. Math. Optim., 45(2002) 145-167.
- [8] N. Arada, E. Casas, F. Tröltzsch, *Error estimate for the numerical approximation of a semilinear elliptic control problem*. Comp. Optim. Appl., 23(2002), 201-229.
- [9] E. Asplund, *Positivity of duality mappings*, Bull. Amer. Math. Soc. Vol., 73(1967), 200–203.
- [10] J.-P. Aubin, A. Cellina, *Differential Inclusions*, Springer-Verlag, 1984.

- [11] J.-P. Aubin and H. Frankowska, *Set-valued Analysis*, Birkhauser, 1990.
- [12] D. Aussel, N. Hadjisavvas, *On quasimonotone variational inequalities*. J. Optim. Theory Appl., 121(2004), 445–450.
- [13] D. Aussel and N. Hadjisavvas, *On quasimonotone variational inequalities*, J. Optim. Theory Appl. 121(2004), 445–450.
- [14] D. Azé and J.-P. Penot, *Uniformly convex and uniformly smooth convex functions*, Annales Fac. Sci. Toulouse, 6(1995), 705–730.
- [15] V. Barbu and Th. Precupanu, *Convexity and Optimization in Banach Spaces*, Editura Academiei, Bucuresti, Romania, 1986.
- [16] V. Barbu, *Optimal control of Navier–Stokes equations with periodic inputs*. Nonlinear Anal. Theory Methods Appl., 31(1988), 15–31.
- [17] V. Barbu and T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer, 2010.
- [18] A. Barbagallo, L. Mallozzi and M. Théra, *Preface: Special Issue on Recent Advances in Variational Calculus: Principles, Tools and Applications*, J. Optim. Theory Appl. 182 (2019), 1–3.
- [19] A. Barbagallo, R. Bot, C. Sagastizábal and M. Théra, *Preface to variational analysis for beginners*, Set-Valued Var. Anal. 28 (2020), no. 4, 583–584.
- [20] S. Bellaassali and A. Jourani, *Necessary optimality conditions in dynamic optimization*, SIAM J. Control Optim., 42(2004), 2043–261.
- [21] A. Ben-Tal and J. Zowe, *A unified theory of first and second order conditions for extremum problems in topological vector spaces*, Math. Program. Study, 19(1982), 39–76.
- [22] C. Berge, *Topological Spaces*, Oliver and Boyd LTD, 1963.

- [23] D.N. Bessis, Yu.S. Ledyayev, R.B. Vinter, *Dualization of the Euler and Hamiltonian inclusions*. Nonlinear Anal. 43(2001), 861–882.
- [24] T. Bewley, R. Temam, M. Ziane, *Existence and uniqueness of optimal control to the Navier–Stokes equations*. C. R. Acad. Sci. Paris, 30(2000), 1007–1011.
- [25] A. Bnouhachem, M.A. Noor, M. Khalfaoui and Z. Sheng, *On a new numerical method for solving general variational inequalities*, Internat. J. Modern Phys. B 25(2011), 4443–4455.
- [26] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
- [27] H. Bonnel and N.S. Pham, *Nonsmooth optimization over the weakly or properly Pareto set of a linear-quadratic multi-objective control problem: Explicit optimality conditions*, J. Indust. Manag. Optim., 7 (2011), 789–809.
- [28] J.M. Borwein and Q.J. Zhu, *Techniques of Variational Analysis*, Springer, 2005.
- [29] H. Brézis, *Equations et inéquations non linéaires dans les espaces vectoriel en dualité*, Ann. Inst. Fourier 18(1968), 115–175.
- [30] H. Brezis, M.G. Crandall and A. Pazy, *Perturbations of nonlinear maximal monotone sets in Banach spaces*, Comm. Pure Appl. Math. 23(1970), 123–144.
- [31] H. Brézis, *Functional Analysis, Sobolev spaces and Partial Differential Equations*, Springer, 2010.
- [32] T.D. Binh, B.T. Kien and N.H. Son, *Local stability of solutions to a parametric multiobjective optimal control problem*, Optimization, 73(2024), Pages 2247–2275.

- [33] F.E. Browder, *Nonlinear elliptic boundary value problems*. Bull. Am. Math. Soc., 69 (1963), 862–874.
- [34] F.E. Browder, P. Hess, *Nonlinear mappings of monotone type in Banach spaces*. J. Funct. Anal., 11(1972), 251–294.
- [35] F.E. Browder, *Fixed point theory and nonlinear problems*. Bull. Am. Math. Soc., 9(1983), 1–39.
- [36] A. Beurling and A.E. Livingston, *A theorem on duality mappings in Banach spaces*, Arkiv für Matematik, 4(1962), 405–411.
- [37] G. Buskes and A.V. Rooij, *Topological Spaces. From Distance to Neighborhood*, Springer, 1996.
- [38] E. Casas, *An optimal control problem governed by the evolution Navier–Stokes equations, Chapter 4*. In: S. Sritharan (ed.), *Optimal Control of Viscous Flow*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1998).
- [39] E. Casas, J.-P. Raymond and H. Zidani, *Pontryagin’s principle for local solutions of control problems with mixed control-state constraints*, SIAM J. Control Optim., 39(2000) 1182–1203.
- [40] E. Casas, F. Tröltzsch, *Error estimates for the finite-element approximation of a semilinear elliptic control problem*. Control and Cybernetics, 31(2002), 695–712.
- [41] E. Casas, M. Mateos, *Uniform convergence of the FEM. Applications to state constrained control problems*. Comput. Appl. Math., 21(2002), 67–100.
- [42] E. Casas, J.-P. Raymond, *Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations*. SIAM J. Control Optim., 45(2006), 1586–1611.
- [43] E. Casas, M. Mateos, J.-P. Raymond, *Error estimates for the numerical approximation of a distributed control problem for the*

- steady-state Navier–Stokes equations*. SIAM J. Control Optim., 46(2007), 952–982.
- [44] E. Casas, F. Tröltzsch, *First- and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations*. SIAM J. Control Optim., 48(2009), 688–718.
 - [45] E. Casas, F. Tröltzsch, *Numerical analysis of some optimal control problems governed by a class of quasilinear elliptic equations*. ESAIM: COCV., 17(2011), 771–800.
 - [46] E. Casas, F. Tröltzsch, *A general theorem on error estimates with application to a quasilinear elliptic optimal control problem*. Comput. Optim. Appl., 53(2012), 173–206.
 - [47] E. Casas and F. Tröltzsch, *Second-order optimality conditions for weak and strong local solutions of parabolic optimal control problems*, Vietnam J. Math. 44(2016), pages 181–202.
 - [48] E. Casas and K. Kunisch, *Optimal control of semilinear parabolic equations with non-smooth pointwise-integral control constraints in time-space*, Appl. Math. Optim., 12(2022).
 - [49] E. Casas and D. Wachsmuth, *A Note on existence of solutions to control problems of semilinear partial differential equations*, SIAM J. Control Optim., 61(2023) 1095–1112.
 - [50] A. Cellina and A. Lasota, *A new approach to the definition of topological degree for multi-valued mappings*, Atti Acca d. Naz. Lincei Rend. Cl. Sci. Fis. Math. Natur., 47(1969), 434–440.
 - [51] L. Cesari, *Optimization Theory and Applications, Problems with Ordinary Differential Equations*, Springer-Verlage, New York Inc., 1983.
 - [52] D. Chang, J.S. Pang, *The generalized quasi-variational inequality problem*. Math. Oper. Res., 2(1982), 211–222.

- [53] Y. Chen and Y.J. Cho, *Topological degree theory for multivalued mapping of class $(S_+)_L$* , Arch. Math., 84(2005), 325-333.
- [54] S. Cherednichenko, A. Rösch, *Error estimates for the discretization of elliptic control problems with pointwise control and state constraints*. Comput. Optim. Appl., 44(2009), 27-77.
- [55] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Company, 1978.
- [56] I. Cioranescu, *Geometry of Banach spaces duality mappings and nonlinear problems*, Kluwer Academic Publishers, 1990.
- [57] F.H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia, 1990.
- [58] R. Cominetti, *Metric regularity, tangent sets, and second-order optimality conditions*, Appl. Math. Optim., 21(1990), 265-287.
- [59] P. Constantin, C. Foias, *Navier-Stokes Equations*. University of Chicago Press, Chicago, 1988.
- [60] J.-P. Crouzeix, *Pseudomonotone variational inequality problems: Existence of solutions*, Math. Program., 78(1997), 305-314.
- [61] S. Dafermos, *Sensitivity analysis in variational inequalities*. Math. Oper. Res., 13(1988), 421-434.
- [62] J.C. De Los Reyes, F. Tröltzsch, *Optimal control of the stationary Navier-Stokes equations with mixed control-state constraints*. SIAM J. Control Optim., 46(2007), 604-629.
- [63] J.C. De Los Reyes, R. Griesse, *State-constrained optimal control of the three-dimensional stationary Navier-Stokes equations*. J. Math. Anal. Appl., 343(2008), 257-272.
- [64] M. Desai, K. Ito, *Optimal controls of Navier-Stokes equations*. SIAM J. Control Optim., 32(1994), 1428-1446.

- [65] E. Di Benedetto, *On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients*. Ann. Scuola Sup. Pisa, Ser. I, 13(1986) 487-535.
- [66] J. Diestel, *Geometry of Banach Spaces-Selected Topics*, Springer-Verlag, 1975.
- [67] P. Daniele, A. Maugeri, W. Oettli, *Time-dependent traffic equilibria*. J. Optim. Theory Appl., 103(1999), 543–555.
- [68] K. Deimling, *Nonlinear Functional Analysis*, Springer-Verlag, 1985.
- [69] J. Diestel, *Geometry of Banach Spaces - Selected Topics, Lecture Notes in Mathematics 485*, Springer-Verlag, Berlin-Heidelberg-New York, 1975.
- [70] J. Dieudonné, *Foundation of Modern Analysis*. Academic Press, New York, 1969.
- [71] A.V. Dmitruk, *Maximum principle for the general optimal control problem with phase and regular mixed constraints*, Computational Mathematics and Modeling 4(1993), 364-377.
- [72] A.V. Dmitruk and N.P. Osmolovskii, *Necessary conditions for a weak minimum in optimal control problems with integral equations subject to state and mixed constraints*, SIAM J. Control Optim., 52(2014), 3437-3462.
- [73] A. Domokos, *Solution sensitivity of variational inequalities*, J. Math. Anal. Appl., 230(1999), 381–389.
- [74] A.L. Dontchev, W. W. Hager, *Implicit functions, Lipschitz maps, and stability in optimization*. Math. Oper. Res., 19(1994), 753–768.
- [75] A.L. Dontchev, *Implicit function theorems for generalized equations*. Math. Programming 70(1995), 91–106.

- [76] A.Ya. Dubovitskii and A.A. Milyutin, *Second variations in extremal problems with constarints*, Dokl. Akad. Nauk SSSR, 160(1965), 18-21.
- [77] N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory*, Interscience Publisher, Inc., New York, 1957.
- [78] J. Dugundji, *Topology*, Allyn and Bacon, Inc., Boston, 1968.
- [79] L.C. Evans, *Partial Differential Equations*, AMS, Providence Rhode Island, 2010.
- [80] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. I, II, Springer, 2003.
- [81] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. I, Springer, 2003.
- [82] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. II, Springer, 2003.
- [83] X. Fang, S. Misra, G. Xue, and D. Yang, *Smart Grid — The New and Improved Power Grid: A Survey*, IEEE Communications Surveys & Tutorials, vol. 14(2012), 944-980.
- [84] H. Frankowska, *Some inverse mapping theorems*, Ann. Inst. Henri Poincaré, 7(1990), 183-234.
- [85] H. Frankowska, *On second-order necessary conditions in optimal control of problems with mixed final point constraints*, IEEE (2019), 1456-1461.
- [86] A.F. Filippov, On certain question in the theory of optimal control, J. SIAM. Control, 1(1962), 76-84.
- [87] G.B. Folland, *Real Analysis. Modern Technique and Their Applications*, John Willey & Son, INC., 1999.

- [88] I. Fonseca and W. Gangbo, *Degree Theory in Analysis and Applications*, Oxford, 1995.
- [89] Giannessi, F., Maugeri, A.: Variational analysis and applications. In: Proceedings of the 38th Conference of the School of Mathematics “G. Stampacchia” in memory of Stampacchia and J.-L. Lions held in Erice, June 20–July 1, 2003. Nonconvex Optimization and Its Applications, vol. 79. Springer, New York, 2005.
- [90] I.V. Girsanov, *Lecture Notes in Economics and Mathematical Systems, Operations Research, Computer Science, Social Science*, Springer-Verlag, Berlin· Heidelberg· New York, 1972
- [91] H. Gfrerer, *Second-order optimality conditions for scalar and vector optimization problems in Banach spaces*, SIAM J. Control Optim., 45(2006), 972–997.
- [92] I. Ginchev, A. Guerraggio and M. Rocca, *Second-order condition in $C^{1,1}$ constrained vector optimization*, Math. Program. Ser. B., 104(2005), 389–405.
- [93] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York, 1984.
- [94] M.S. Gowda *Inverse and implicit function theorems for H -differentiable and semismooth functions*, Optim. Meth. Software, 19(2004), 443–461.
- [95] W. Grecksch, F. Heyde, G. Isac and Chr. Tammer, *A characterization of approximate solutions of multiobjective stochastics optimal control problems*, Optimization, 52 (2003), 153–170.
- [96] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman Publishing Inc., 1985.
- [97] J.S. Guo and J.C. Yao, *Variational inequalities with nonmonotone operators*. J. Optim. Theory Appl., 80(1994), 63–74.

- [98] W. Han and S. Zeng, *On convergence of numerical methods for variational-hemivariational inequalities under minimal solution regularity*, Appl. Math. Lett., 93(2019), 105–110.
- [99] G.H. Hardy, J.E. Littlewood and G. Polya, *Inequalities*, Cambridge University Press, 1988.
- [100] P. Hartman, *Ordinary Differential Equations*, Wiley, 1964.
- [101] P. Hartmann, G. Stampacchia, *On some nonlinear elliptic differential functional equations*. Acta Math., 115(1966), 153–188.
- [102] F. Hirsch and G. Lacombe, *Elements of Functional Analysis*, Springer-Verlage, 1999.
- [103] M. Hinze, *A variational discretization concept in control constrained optimization: The linear-quadratic case*. Comput. Optim. Appl., 30(2005), 45–61.
- [104] M. Hinze, C. Meyer, *Variational discretization of Lavrentiev-regularized state constrained elliptic optimal control problems*. Comput. Optim. Appl., 46 (2010), 487–510.
- [105] M. Hinze, C. Meyer, *Stability of semilinear elliptic optimal control problems with pointwise state constraints*. Comput. Optim. Appl., 52 (2012), 87–114.
- [106] R.H.W. Hoppe, M. Kieweg, *Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems*. Comput. Optim. Appl., 46(2010), 511–533.
- [107] S. Hu and N.S. Parageorgiou, *Generalization of Browder’s degree theory*, Trans. Amer. Math. Soc., 347(1995), 233–259.
- [108] A.D. Ioffe and V.M. Tihomirov, *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.

- [109] V.I. Istratescu, *Fixed Point Theory. An Introduction*, D. Reidel Publishing Company, Dordrecht, 1981.
- [110] B. Jiménez and V. Novo, *Second order necessary conditions in set constrained differentiable vector optimization*, Math. Methods Oper. Res., 58(2003), 299–317.
- [111] B. Jiménez and V. Novo, *Optimality conditions in differentiable vector optimization via second-order tangent sets*, Appl. Math. Optim., 49(2004), 123–144.
- [112] A. Jourani, *Regularity and strong sufficient optimality conditions in differentiable optimization problems*, Numer. Funct. Anal. Optim., 14(1993), 69-87.
- [113] A. Jourani, *Metric regularity and second-order necessary optimality conditions for minimization problems under inclusion constraints*, J. Optim. Theor. Appl., 81(1994), 97-120.
- [114] S. Karamardian, *Complementarity problems over cones with monotone and pseudomonotone maps*. J. Optim. Theory Appl., 18(1976), 445–454.
- [115] A.G. Kartsatos and I. V. Skrypnik, *Topological degree theories for densely defined mapping involving operators of type (S_+)* , Adv. Differential Equ. 4(1999), 413-456.
- [116] A.G. Kartsatos and I. V. Skrypnik, *The index of a critical point for densely defined operators of type $(S_+)_L$ in Banach spaces*, Trans. Amer. Math. Society 354(2001), 1601-1630.
- [117] C.Y. Kaya and H. Maurer, *A numerical method for nonconvex multi-objective optimal control problems*, Comput. Optim. Appl., 57(2014), 685–702.
- [118] H. Khanh and B.T. Kien, *On the regularity of multipliers and second-order optimality conditions of KKT-type for semilinear*

- parabolic control problems*, J. Math. Anal. Appl., 538 (2024), no. 1, Paper No. 128436, 30 pp.
- [119] B.T. Kien, *On the metric projection onto a family of closed convex sets in a uniformly convex Banach space*, Nonlinear Anal. Forum 7(2002), no. 1, 93–102.
 - [120] B.T. Kien, *The normalized duality mapping and two related characteristic properties of a uniformly convex Banach space*, Acta Mathematica Vietnamica 27(2002), no. 1, 53–67.
 - [121] B.T. Kien and M.-M. Wong, *On the solution stability of variational inequalities*, J. Global. Optim., 39(2007), 101–111.
 - [122] B.T. Kien, M. M. Wong and N. C. Wong, *On the degree theory for general mappings of monotone type*, J. Math. Anal. Appl., 340(2008), 707–720.
 - [123] B.T. Kien, J. C. Yao, *Localization of generalized normal maps and stability of variational inequalities in reflexive Banach spaces*. Set-Valued Anal., 16(2008), 399–412.
 - [124] B.T. Kien, J.C. Yao and N.D. Yen, *On the solution existence of pseudomonotone variational inequalities*, J. Global Optim., 41(2008), 135–145.
 - [125] B.T. Kien, *Lower semicontinuity of the solution map to a parametric generalized variational inequality in reflexive Banach spaces*, Set-Valued Analysis, 16(2008), 1089–1105.
 - [126] B.T. Kien, M.M. Wong, N.C. Wong and J.-C. Yao, *Solution existence of variational inequalities with pseudomonotone operators in the sense of Brézis*, J. Optim. Theory Appl., 140(2009), 249–263.
 - [127] B.T. Kien, N.-C. Wong and J.C.-Yao, *Degree theory for generalized variational inequalities and applications*, European J. Oper. Res., 192(2009), 730–736.

- [128] B.T. Kien, N.-C. Wong and J.-C. Yao, *Necessary conditions for multi-objective optimal control problem with free end-time*, SIAM J. Control Optim., 47(2010), 2251–2274.
- [129] B.T. Kien and V.H. Nhu, *Second-order necessary optimality conditions for a class of semilinear elliptic optimal control problems with mixed pointwise constraints*, SIAM J. Control Optim., 52(2014), 1166-1202.
- [130] B.T. Kien, A. Rösh and D. Wachsmuth, *Pontryagin’s maximum principle for optimal control problem governed by 3-dimensional Navier-Stokes equations with pointwise constraints*, J. Optim. Theor. Appl., 173(2017), 30-55.
- [131] B.T. Kien, V.H. Nhu and N.H. Son, *Second-order optimality conditions for a semilinear elliptic optimal control problem with mixed pointwise constraints*, Set-Valued Var. Anal., 25(2017) 177-210.
- [132] B.T. Kien, N.V. Tuyen and J.C. Yao, *Second-order KKT optimality conditions for multiobjective optimal control problems*, SIAM J. Control Optim., 56(2018), 4069-4097.
- [133] B.T. Kien, X. Quin, C.F. Wen and J.-C. Yao, *Second-order optimality conditions for multi-objective optimal control problems with mixed pointwise constraints and free right end-point*, SIAM J. Control Optim., 58(2020), 2658–2677.
- [134] B.T. Kien, C.F. Wen and J. C. Yao, *The Galerkin method and regularization for variational inequalities in reflexive Banach spaces*, J. Optim. Theory Appl., 189(2021), no. 2, 578-596.
- [135] B.T. Kien, V.E. Fedorov and T.D. Phuong, *Optimal control problems governed by fractional differential equations with control constraints*, SIAM J. Control Optim., 60(2022), pp. 1732 - 1762.

- [136] B.T. Kien, A. Rosch, N.H. Son and N.V. Tuyen, *FEM for semilinear elliptic optimal control with nonlinear and mixed constraints*, J. Optim. Theory Appl., 197(2023), no. 1, 130-173.
- [137] B.T. Kien and T.D. Binh, *On the second-order optimality conditions for multi-objective optimal control problems with mixed pointwise constraints*, J. Glob. Optim., 85(2023), 155–183.
- [138] B.T. Kien and N.T. Qui, *On the existence of solutions to a class of optimal control problems*, Journal of Nonlinear and Convex Analysis, 25(2024), 2949-2958.
- [139] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, San Diego, 1980.
- [140] I. Konnov, *Generalized monotone equilibrium problems and variational inequalities*. In: Hadjisavvas, N., Komlósi, S., Schaible, S. (eds.) *Handbook of Generalized Convexity and Generalized Monotonicity*, pp. 559–618. Springer, Berlin, 2005.
- [141] E. Kreyszig, *Introductory Functional Analysis with Applications*. John Wiley & Sons, 1989.
- [142] E. Kreyszig, *Advanced Engineering Mathematics*, John Wiley and Sons, INC, 2006.
- [143] O. Ladyzhenskaya, V. Solonnikov, N. Ural'tseva, *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, Providence, 1988.
- [144] V. Lakshmikantham, *Upper and lower bound of the norm of solutions of differential equations*, Proc. Amer. Math. Soc., 13(1962), 615–616.
- [145] S. Lang, *Linear Algebra*, Springer, 1987.

- [146] A.B. Levy and R.T. Rockafellar, *Sensitivity analysis of solutions to generalized equations*. Trans. Amer. Math. Soc., 345(1994), 661–671.
- [147] A.B. Levy, *Sensitivity of solutions to variational inequalities on Banach spaces*, SIAM J. Control Optim., 38(1999), 50-60.
- [148] A.B. Levy and R. A. Poliquin and R. Rockafellar, *Stability of locally optimal solutions*, SIAM J. Control Optim., 10(2000), 580-604.
- [149] A.B. Levy and B. S. Mordukhovich, *Coderivatives in parametric optimization*, Mathematical Programming, 99(2004), 311-327.
- [150] X. Li, J. Yong, *Optimal Control Theory for Infinite Dimensional Systems*. Birkhäuser, Boston, 1995.
- [151] N.G. LLOYD, *Degree Theory*, Cambridge University Press, 1978.
- [152] N.G. LLOYD, *A survey of degree theory: basis and development*, IEEE Trans. Circuits and Systems 30(1983), 607–616.
- [153] J.L. Lions, G. Stampacchia, *Inéquations Variationnelles non Coercives*. C. R. Acad. Sci. Paris, 261(1965), 25–27.
- [154] J.L. Lions, G. Stampacchia, *Variational inequalities*. Commun. Pure Appl. Math., 20(1967), 493–519.
- [155] J.L. Lions, E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*. Springer, Berlin (1972).
- [156] J.L. Lions, *Numerical methods for variational inequalities—applications in physics and in control theory*, Information processing 77 (Proc. IFIP Congr., Toronto, Ont., 1977), pp. 917–924. IFIP Congr. Ser., Vol. 7, North-Holland, Amsterdam, 1977.
- [157] J.M. Longuski, J.J. Guzmán and J.E. Prussing, *Optimal Control with Aerospace Applications*, Springer Science+Business Media New York 2014.

- [158] O.L. Mangasarian, T.-H. Shiau, *Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems*. SIAM J. Control Optim., 25(1987), 583–595.
- [159] M.A. Mansour, D. Aussel, *Quasimonotone variational inequalities and quasiconvex programming: qualitative stability*. Pac. J. Optim., 2(2006), 611–626.
- [160] H. Maurer and J. Zowe, *First- and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Program., 16(1979), 98–110.
- [161] P. Merino, F. Tröltzsch, B. Vexler, *Error estimates for finite element approximation of a semilinear elliptic control problem with state constraints and finite dimensional control space*. ESAIM: M2AN, 44 (2010), 167–188.
- [162] C. Meyer, A. Rösch, *Superconvergence properties of optimal control problems*. SIAM J. Control Optim., 43(2004), 970–985.
- [163] C. Meyer, U. Prüfert and F. Tröltzsch, *On two numerical methods for state-constrained elliptic control problems*, Optimisation Methods and Software, 22(2007) 871–899.
- [164] C. Meyer, *Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints*. Control Cybern., 37(2008), 53–83.
- [165] G. Minty, *Monotone operators in Hilbert spaces*. Duke Math. J., 29(1962), 341–346.
- [166] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I, Basis Theory*, Springer, 2006.
- [167] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation II, Applications*, Springer, 2006.

- [168] I. Neitzel, Tröltzsch, *On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints*, ESAIM: Control, Optimisation and Calculus of Variations, 15(2009) 426-453.
- [169] I. Neitzel, J. Pfefferer, A. Rösch, *Finite element discretization of state-constraint elliptic optimal control problem with semilinear state equation*. SIAM J. Control Optim., 53(2015), 874–904.
- [170] T.-N. Ngo and N. Hayek, *Necessary conditions of Pareto optimality for multiobjective optimal control problems under constraints*, Optimization, 66(2017), 149–177.
- [171] V.H. Nhu, N.Q. Tuan, N.B. Giang, N.T.T. Huong, *Continuity regularity of optimal control solutions to distributed and boundary semilinear elliptic optimal control problems with mixed pointwise control-state constraints*. J. Math. Anal. Appl., 512(2022), 126139.
- [172] E. Ning, W. Song and Y. Zhang, *Second-order sufficient optimality conditions in vector optimization*, J. Glob. Optim., 54 (2012), 537–549.
- [173] V.A. de Oliveira, G.N. Silva and M.A. Rojas-Medar, *A class of multi-objective control problems*, Optim. Control Appl. Meth., 30(2009), 77–86.
- [174] V.A. de Oliveira and G.N. Silva, *On sufficient optimality condition for multiobjective control problems*, J. Glob. Optim., 64(2016), 721–744.
- [175] N.P. Osmolovskii, *Necessary second-order conditions for a weak local minimum in problem with endpoint and control constraints*, J. Math. Anal. Appl., 457(2018), 1613-1633.
- [176] Z. Páles and V. Zeidan, *Optimum problems with certain lower semicontinuous set-valued constraints*, SIAM J. Control and Optimization, 8(1988), 707-727.

- [177] Z. Páles and V. Zeidan, *Nonsmooth optimum problems with constraints*, SIAM J. Control and Optimization, 32(1994), 1476-1502.
- [178] Z. Páles and V. Zeidan, *Characterization of L^1 -closed decomposable sets in L^∞* , J. Math. Anal. Appl., 238(1999), 491-515.
- [179] Z. Páles and V. Zeidan- *Optimal control problems with set-valued control and state constraints*, SIAM J. Optim., 14(2003), 334-358.
- [180] J.-S. Pang and J.-C. Yao, *On a generalization of a normal map and equation*, SIAM J. Control Optim., 33(1995), 401-426.
- [181] S. Peitz, K. Schafer, S.O. Boebaun, J. Echstein, U. Koehler and M. Dellnitz, *A multi-objective MPC approach for autonomously driven electric vehicles*, IFAC PaperOnline, 50(1)(2017), 8674-8679.
- [182] J.-P. Penot, *Metric regularity, openness and Lipschitzian behavior of multifunctions*, Nonlinear Anal. Th. Meth. Appl., 13(1989), 629-643.
- [183] J.-P. Penot and R. Ratsimahalo, *Characterizations of metric projections in Banach spaces and applications*, Abstract and Applied Analysis, 3(1998), 85-103.
- [184] J.-P. Penot and R. Ratsimahalo, *Uniform continuity of projection mappings in Banach spaces*, Communications on Applied Nonlinear Analysis, 6(1999), 1-17.
- [185] J.-P. Penot, *Calculus Without Derivatives*, Springer, 2013.
- [186] V. Pesce, A. Colagrossi and S. Silvestrini, *Modern Spacecraft Guidance, Navigation and Control*, Elsevier Inc, 2023.
- [187] B. Prus and R. Smarzewski, *Strongly unique best approximations and centers in uniformly convex spaces*, J. Math. Anal. Appl., 121(1987), 10-21.
- [188] J. Prüß, *A characterization of uniform convexity and applications to accretive operators*, Hiroshima Math. J., 11(1981), 229-234.

- [189] J.-P. Raymond, H. Zidani, *Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls*. SIAM J. Control Optim., 36(1998), 1853–1879.
- [190] J.-P. Raymond, H. Zidani, *Pontryagin's principle for time-optimal problems*. J. Optim. Theory Appl., 101(1999), 375–402.
- [191] J.-P. Raymond and H. Zidani, *Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations*, Appl. Math. Optim., 39(1999) 143–177.
- [192] B. Ricceri, *Basic existence theorems for generalization variational and quasi-variational inequalities*. In: Giannessi, F., Maugeri, A. (eds.) *Variational Inequalities and Network Equilibrium Problems*, pp. 251–255. Plenum, New York (1995).
- [193] S.M. Robinson, *Regularity and stability for convex multivalued functions*. Math. Oper. Res., 1(1976), 130–143.
- [194] S.M. Robinson, *Stability theory for systems of inequalities, part II: Differentiable nonlinear systems*, SIAM J. Numer. Anal., 12(1976), 497–513.
- [195] S.M. Robinson, *An implicit-function theorem for a class of nonsmooth functions*. Math. Oper. Res., 16(1991), 292–309.
- [196] S.M. Robinson, *Normal maps induced by linear transformations*. Math. Oper. Res., 17(1992), 691–714.
- [197] S.M. Robinson, *Constraint nondegeneracy in variational analysis*. Math. Oper. Res., 28(2003), 201–232.
- [198] S.M. Robinson, *Localized normal maps and the stability of variational conditions*. Set-Valued Anal., 12(2004), 259–274.
- [199] S.M. Robinson, *Solution continuity affine variational inequalities*. SIAM J. Optim., 18(2007), 1046–1060.

- [200] S.M. Robinson, S. Lu, *Solution continuity in variational conditions*. J. Glob. Optim., 40(2008), 405-415.
- [201] S.M. Robinson, *Localized normal maps and the stability of variational conditions*, Set-valued Analysis, 12(2004), 259-274.
- [202] R.T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer 1997.
- [203] A. Rösch and F.Tröltzsch, *Existence of regular Lagrange multiplier for a nonlinear elliptic optimal control problem with pointwise control-state constraints*, SIAM J Control Optim., 45(2006) 548-564.
- [204] A. Rösch and F.Tröltzsch, *On regularity of solutions and Lagrange multipliers of optimal control problems for semilinear elliptic equations with pointwise control-state constraints*, SIAM J Control Optim., 46(2007) 1098-1115.
- [205] A. Rösch, D. Wachsmuth, *Semi-smooth Newton method for an optimal control problem with control and mixed control-state constraints*. Optim. Meth. Sof., 26 (2011), 169–186.
- [206] A. Rösch, S. Steinig, *A priori error estimates for a state-constrained elliptic optimal control problem*. ESAIM: M2AN, 46(2012), 1107–1120.
- [207] T. Roubicek, *Nonlinear Partial Differential Equations with Applications*. Birkhäuser, Basel, 2013.
- [208] E. Roxin, *The existence of optimal controls*, Michigan Math. J. 9(1962), 109-119.
- [209] W. Rudin, *Functional Analysis*, McGraw-Hill, Inc., 1973.
- [210] A. Shapiro, *Sensitivity analysis of parameterized variational inequalities*, Math. Oper. Res, 30(2005), 100-126.

- [211] Y.-H. Shao and K. Tsujioka, *On proper-efficiency for nonsmooth multiobjective optimal control problems*, Bull. Infor. Cyber., 32(2000), 139–155.
- [212] M. Sion, *On general minimax theorems*. Pacific J. Math., 8(1958), 171–176.
- [213] G. Stampacchia, *Formes Biliéaires Coercitives sur les Ensembles Convexes*. C. R. Acad. Sci. Paris, 258(1964), 4413–4416.
- [214] O. Steinbach, *Numerical Approximation Methods for Elliptic Boundary Value Problems, Finite and Boundary Elements*. Springer (2008).
- [215] M. Taylor, *Measure Theory and Integration*, American Mathematical Society, 2006.
- [216] R. Temam, *Navier–Stokes Equations, Theory and Numerical Analysis*. North-Holland Publishing, Amsterdam (1979).
- [217] B.A. Ton, *An optimal control free boundary problem for the Navier–Stokes equations*. Nonlinear Anal., 63(2005), 831–839.
- [218] F. Tröltzsch, *Optimality Conditions for Parabolic Control Problems and Applications*, Teubner-Texte Math. 62, B.G. Teubner Verlagsgesellschaft, Leipzig, 1984.
- [219] F. Tröltzsch, *Optimal Control of Partial Differential Equations, Theory, Method and Applications*, American Mathematical Society, Providence Rhode Island, 2010.
- [220] M.L. Tuballa, M.L. Abundo, *A review of the development of Smart Grid technologies*. Renewable and Sustainable Energy Reviews, 59(2016), 710–725.
- [221] H. Tuy, *Convex Analysis and Global Optimization (Second edition)*, Springer, 2017.

- [222] R. Vinter, *Optimal Control*, Birkhäuser, Boston, Basel, Berlin, 2010.
- [223] G. Wang, *Optimal control of 3-dimensional Navier–Stokes equations with state constraints*. SIAM J. Control Optim. 41(2002), 583–606.
- [224] F. Wang, W. Hang and X.-L. Cheng, *Discontinuous Galerkin methods for solving elliptic variational inequalities*, SIAM J. Numer. Anal., 48(2010), 708–733.
- [225] R.L. Wheeden and A. Zygmund, *Measure and Integral, An Introduction to Real Analysis*, CRC Press, Taylor and Francis Group, 2015.
- [226] B. Xiao, P.T. Harker, *A nonsmooth Newton method for variational inequalities. II. Numerical results*, Math. Programming, 65(1994), Ser. A, 195–216.
- [227] J.C. Yao, *Multi-valued variational inequalities with K -pseudomonotone operators*. J. Optim. Theory Appl., 80(1994), 63–74.
- [228] J.C. Yao, *Variational inequalities with generalized monotone operators*, Math. Oper. Res., 19(1994), 691–705.
- [229] N.D. Yen, *Hölder continuity of solution to a parametric variational inequality*, Appl. Math. Optim., 31(1995) 245–255.
- [230] N.D. Yen, *Lipshitz continuity of solutions of variational inequalities with a parametric polyhedral constraint*, Math. Oper. Res., 20(1995), 695–707.
- [231] N.D. Yen, *A result related to Ricceri’s conjecture on generalized quasi-variational inequalities*, Arch. Math., 69(1997), 507–514.
- [232] N.D. Yen, *On a problem of B. Ricceri on variational inequalities*. In: Cho, Y.J., Kim, J.K., Kang, S.M. (eds.) *Fixed Point Theory and Applications*, vol. 5(2004), pp. 163–173. Nova Science, New York.

- [233] N.D. Yen and G.M. Lee, *Solution sensitivity of a class of variational inequalities*. J. Math. Anal. Appl., 215(1997), 48–55.
- [234] K. Yosida and E. Hewitt, *Finitely additive measures*. Trans. Am. Math. Soc., 72(1952) 46-66.
- [235] V. Zeidan, *The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency*, SIAM J. Control. Optim., 32(1994), 1297-1321.
- [236] E. Zeidler, *Nonlinear Functional Analysis and Its Applications, I. Fixed-Point Theorems*, Springer-Verlag, 1986.
- [237] E. Zeidler, *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*, Springer-Verlag, Berlin, 1990.
- [238] E. Zeidler, *Nonlinear Functional Analysis and its Applications, II/B: Nonlinear Monotone Operators*, Springer-Verlag, Berlin, 1990.
- [239] E. Zeidler, *Applied Functional Analysis, Main Principles and Their Applications*, Springer-Verlag, 1995.
- [240] Q.J. Zhu, *Hamiltonian necessary conditions for a multiobjective optimal control problems with endpoint constraints*, SIAM J. Control Optim., 39(2000), 97–112.
- [241] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5(1979), 49-62.