OPTIMIZATION PROBLEMS WITH VARIATIONAL INEQUALITY CONSTRAINTS

by

XIANG YANG YE
B.Sc. Xiamen University, 1982

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

We accept this thesis as conforming
to the required standard

[redacted]

Dr. Jane J. Ye, Supervisor (Department of Math. & Stat.)

[redacted]

Dr. Chris Bose, Departmental Member (Department of Math. & Stat.)

[redacted]

Dr. William Reed, Departmental Member (Department of Math. & Stat.)

[redacted]

Dr. Deming Zhuang, External Examiner (Mount St-Vincent University)

University of Victoria

Supervisor: Dr. Jane J. Ye (Department of Mathematics and Statistics)

Abstract

Optimization problems with variational inequality constraints are a class of mathematical programming problems which have variational inequalities as constraints. Due to the implicit hierarchical structure of the constraint region induced by variational inequalities, these problems are usually very difficult to solve. Nevertheless, they have very important applications in the areas such as economics, operations research and engineering. This thesis is devoted to necessary optimality conditions for such problems. Using the penalty methods and nonsmooth analysis technique, two types of necessary optimality conditions are derived.

Examiners:

[redacted]

Dr. Jane J. Ye, Supervisor (Department of Math. & Stat.)

[redacted]

Dr. Chris Bose, Departmental Member (Department of Math. & Stat.)

[redacted]

Dr. William Reed, Departmental Member (Department of Math. & Stat.)

[redacted]

Dr. Deming Zhuang, External Examiner (Mount St-Vincent University)

Contents

Acknowledgement

I would like to express my most sincere appreciation to Dr. Jane J. Ye, my supervisor, for her guidance, encouragement and her great interest in my research work. This thesis would not have been completed without her help and incredible patience.

I would like to thank Dr. Chris Bose, Dr. William Reed and Dr. Deming Zhuang for their time to examine my thesis and provide me with suggestions and comments.

Thanks are due to the Department of Mathematics and Statistics, University of Victoria, for giving me the opportunity to work on this Master's thesis.

Finally I owe a special debt of thanks to my parents and my wife for their love, support and encouragement. I would like to extend heartful thanks to my friends and their family members for their emotional support and prayers.

Chapter 1

Introduction

An optimization problem with variational inequality contraints (OPVIC) is an optimization problem over variables $x \in R^n$ and $y \in R^m$, in which some or all of its constraints are defined by a parametric variational inequality with y as its primary variable and x the parameter. More specifically, this problem is defined as follows:

$$\text{(OPVIC)} \qquad \min\{f(x,y): \quad x \in \Omega_1, y \in S(x)\} \tag{1.1}$$

where $f : R^{n+m} \to R$ is a point-to-point map, Ω_1 is a nonempty closed subset of R^n, and for each $x \in \Omega_1$, $S(x)$ is the solution set of a variational inequality with parameter x, i.e.

$$S(x) = \{y \in \Omega_2(x): \quad \langle F(x,y), z - y\rangle \geq 0 \quad \forall z \in \Omega_2(x)\} \tag{1.2}$$

where $\Omega_2 : R^n \Rightarrow R^m$ is a set-valued map and $F : R^{n+m} \to R^m$ is a point-to-point map.

One can interpret the above problem by a hierarchical decision process where there are two decision makers and the upper level decision maker always has the first choice as follows: Given a decision vector x by the upper level decision maker (hereafter the leader) $S(x)$ can be considered to be the lower level decision maker's (the follower) decision set, i.e., the set of decision vectors that the follower agrees to use. Suppose that the game is co-operative, i.e., if the follower's decision set $S(x)$ is not a singleton, the follower will allow the leader to choose which of them is actually used. Having the complete knowledge of the possible reactions by the follower, the leader then selects a decision vector $x \in \Omega_1$ and $y \in \Omega_2(x)$ to minimize his objective function $f(x, y)$.

If $F(x, y)$ is the gradient of a real-valued differentiable function, i.e., $F(x, y)$
$= \nabla_y g(x, y)$, where $g : R^n \to R$ is differentiable in y and $\Omega_2(x)$ is convex then the variational inequality with parameter x,

$$\langle F(x, y), z - y\rangle \geq 0 \ \forall z \in \Omega_2(x) \tag{1.3}$$

is actually a restatement of the first-order necessary optimality conditions for the following optimization problem with parameter x,

$$(P_x) \qquad \min\{g(x, y) : y \in \Omega_2(x)\} \tag{1.4}$$

(see e.g. [19]). Furthermore, if $g(x, y)$ is pseudo-convex in y (i.e., $\langle \nabla_y g(x, y), z - y\rangle \geq 0$ implies $g(x, z) \geq g(x, y)\ \forall y, z \in \Omega_2(x)$), then a vector $y \in \Omega_2(x)$ is a solution to (1.3) if and only if it is a global optimal solution of (1.4). In this case, (OPVIC) is the following classical bilevel programming problem, or so called Stackelberg game first introduced in an economic model by Von Stackelberg [37]:

$$\text{(BP)} \qquad \min\{f(x,y) : x \in \Omega_1, y \in \Sigma(x)\} \tag{1.5}$$

where $\Sigma(x)$ is the set of solutions of the problem (P_x). There is a rather large literature on this important (but difficult) class of mathematical programs; the recent volume [1] contains a number of interesting articles describing the diverse applications of bilevel optimization in engineering and economics. Some selected references on the bilevel programming problem are [1, 4, 13, 24, 32, 37, 40, 43, 44].

The optimization problem with varitional inequality constraints (OPVIC) is a generalization of the bilevel program (BP) in which the follower's decision problem is modeled by a variational inequality constraint. The last decade has witnessed a growing interest in the theory of (OPVIC). See [1, 5, 15, 16, 17, 21, 18, 25, 26, 27, 31, 35, 36, 38, 41, 42]. Such problems have wide applications in areas such as game theory, economics, design of transportation networks and control of partial differential equations (c.f. [5]) as well as many others.

A natural approach to obtain a necessary optimality condition for (OPVIC) is to reduce (OPVIC) to an ordinary (single level) mathematical programming problem and use the existing necessary optimality condition for the single level problem. There are several equivalent single level formulations for (OPVIC). To illustrate we assume that

$$\Omega_2(x) := \{y \in R^m : \psi(x,y) \leq 0\}$$

where $\psi : R^{n+m} \to R^d$. The Karush-Kuhn-Tucker (KKT) approach is to interpret the varitional inequality constraint $y \in S(x)$ as y being a solution of the following optimization problem:

$$\min\langle F(x,y), z\rangle \qquad \text{s.t. } z \in \Omega(x),$$

and replace it by the KKT necessary optimality conditions for the above optimization problem with is also sufficient if $\Omega(x)$ is convex for each x. Using this approach, under the assumption that $\Omega(x)$ is convex, $F(x,y)$ is differentiable and the usual constraint qualifications such as the Mangasarian-Fromovitz condition hold for the above optimization problem, $(\bar{x}, \bar{y})$ is a solution of (OPVIC) if and only if there exists $\bar{u} \in R^d$ such that $(\bar{x}, \bar{y}, \bar{u})$ is a solution of the following problem:

$$\begin{aligned}
\text{(KS)} \qquad \min \quad & f(x,y) \\
\text{s.t.} \quad & F(x,y) + \nabla_y \psi(x,y)^T u = 0 \\
& \langle u, \psi(x,y)\rangle = 0 \qquad\qquad (1.6) \\
& u \geq 0, \psi(x,y) \leq 0 \\
& x \in X, y \in R^m.
\end{aligned}$$

Using the single level formulation (KS), one can easily derive Fritz-John type nessary optimality conditions for (OPVIC). The Fritz-John type condition however is useless when it is degenerate, i.e., the multiplier corresponding to the objective function is zero. Therefore it is very important to find constraint qualifications which ensure the existence of nondegenerate multipliers, i.e., the multiplier corresponding to the objective function is nonzero. This type of optimality condition is known as Kuhn-Tucker type optimality condition. However it was shown in [41, Proposition 1.1] that the Mangasarian-Fromovitz condition will never hold for problem (KS) as long as the constaint $\psi(x, y) \leq 0$ is binding at $(\bar{x}, \bar{y})$. Other approaches for single level formulation of (OPVIC) include the value function approach and the gap function approach and many others (see [40, 41]). The common characterization for these equivalent single level formulations is that the usual constraint qualification will never be satisfied and the clamness condition is the right constraint qualification (see [40, 41]). For problem (KS) the trouble comes from the constraint (1.6) which reflects the bilevel struction of (OPVIC). In [40, 41, 25] some conditions are given under which (1.6) is the exact penalty function of problem (KS). Since the exact penality formulation moves the troublesome equality constraint (1.6) to the objective function, a Kuhn-Tucker type necessary condition for (KS) (equivalent for (OPVIC)) can then be derived easily.

The purpose of this thesis is to find Kuhn-Tucker type optimality conditions which require less or no constraint qualifications.

In [35, 36], for the case where $\Omega(x)$ is independent of x and convex, a Kuhn-Tucker type necessary optimality condition involving a kind of set-valued map deivatives called paratingent derivative is derived for (OPVIC). This type of condition is almost constraint qualification free. In this thesis, we extend this result to the case where $\Omega(x)$ depends on x which has more applications.

We now illustrate the approach which was suggested by Shi in [35, 36].

Rewrite (OPVIC) as the following single level problem:

$$\begin{array}{lll} (\tilde{P}) & \min & f(x,y) \\ & \text{s.t.} & x \in \Omega_1, y \in \Omega_2(x) \\ & & \| - F(x,y)\|_y^{\Omega_2(x)} = 0, \end{array}$$

where

$$\|y^*\|_y^{\Omega_2(x)} = \sup_{p \in (\Omega_2(x) - y) \cap \bar{B}} \langle y^*, p \rangle.$$

Using the inexact penalty approach, problem $(\tilde{P})$ can be approximated by the following penalized problem:

$$\begin{array}{lll} (P_n) & \min & \{f(x,y) + N_n \| - F(x,y)\|_y^{\Omega_2(x)}\} \\ & \text{s.t.} & x \in R^n, y \in \Omega_2(x), \end{array}$$

where $N_n > 0$ is a "penalty factor" and $N_n \to +\infty$. By applying the Ekeland's Variational Principle on the problem (P_n), we will be able to find an almost minimizer sequence $\{(x_n, y_n)\}$ to problem (OPVIC). Then, by using the lopsided minimax theorem, we will obtain a "Lagrange mlutiplier" sequence $\{p_n\}$ for (P_n). Finally, we will extract a subsequence of $\{(x_n, y_n, p_n)\}$, which will converge to $\{(\bar{x}, \bar{y}, \bar{p})\}$, satisfying the Kuch-Tucker necessary optimality condition involving paratingent derivatives.

In [44], for the classical bilevel programming problems (BP), Zhang derived a Kuhn-Tucker type necessary optimality conditions involving coderivatives. In this thesis, for more general problem (OPVIC) , we derive an optimality condition which is similar but easier to compute than the one derived in [44] with an easier to verify constraint qualification. The main idea is to rewrite (OPVIC) as an optimization problem with generalized equation constaints in the following form:

$$\begin{aligned} \min \quad & f(x,y) \\ \text{s.t.} \quad & (x,y) \in \Omega_1 \times R^m. \\ & 0 \in -F(x,y) + N(y,\Omega_2) \end{aligned}$$

and apply the recent results about the coderivative characterization of a pesudo-Lipschitz continuity of a set-valued map and the calculus of coderivatives of set-valued maps developed in Mordukhovich [29, 28, 30] to obtain the results.

Finally we organize the thesis as follows: Chapter 2 contains the nonsmooth analysis needed in the rest of the thesis. Generalized gradients, multifunction derivatives and related results are reviewed. In chapter 3, (OPVIC) is analyzed using the nonsmooth analysis technique and the inexact penalty method. Existence of the solution to problem (OPVIC) is proved. A Kuhn-Tucker type necessary optimality condition involving paratingent derivatives for (OPVIC) is developed. In Chapter 4 by studying the necessary optimality condition for an optimization problem with generalized equation constraints we establish Kuhn-Tucker type necessary optimality conditions involving coderivatives to problem (OPVIC). An example is given to illustrate the usefulness of the optimality condition derived.

Chapter 2

Nonsmooth Analysis

As discussed in Chapter 1, nonsmooth phenomena, or lack of ordinary differentiability, in mathematics and optimization occur naturally and frequently. The purpose of this chapter is to gather a basic nonsmooth analysis tool kit which will be used frequently throughout this thesis.

The following is an outline of this chapter.

§2.1 Preliminaries. In this section, some basic terminology is presented.

§2.2. Generalized Gradients. This section is elementary. It presents some basic concepts and results of Nonsmooth Analysis.

§2.3. Derivatives of set-valued maps. In this section we introduce two types of derivatives of set-valued maps and some basic results which play an important role in nonsmooth optimization problems.

2.1 Notation and terminology

In this thesis, the following notations and terminologies will be used.

Let Ω be a subset of R^n. We denote by $cl\Omega$ the closure of Ω and $co\Omega$ the convex hull of Ω. Let R^n_+ be the nonnegative orthant of R^n, $\langle\cdot,\cdot\rangle$ be the inner product and $\|x\|$ be the norm of x in a space. $\bar{B}$ is the closure of B, the unit ball in a space.

For any two subsets C, K of R^n, we denote by $C-K := \{c-k : c \in C, k \in K\}$. K is said to be a cone if and only if $tx \in K \ \forall x \in K, \ t \geq 0$. $Cone(\Omega)$ is the conic hull of Ω, the smallest cone which contain Ω. We denote $[x, y]$ by the set $\{u : \quad u = \lambda x + (1-\lambda)y, \quad \forall\lambda \in (0,1)\}$.

The symbol $x(\in \Omega) \to \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$. The adjoint (transpose) matrix to M is denoted by $M^\top$.

Definition 2.1 *A set-valued map Φ from R^n to R^m, denoted by $\Phi : R^n \rightrightarrows R^m$, is a map that associates with any $x \in R^n$ a subet of R^n, called the image or the value of Φ at x.*

We say that a set-valued map is proper if there exists at least one element $x \in R^n$ such that $\Phi(x) \neq \emptyset$, that is, if Φ is not the constant map $\emptyset$. In this case, we say that the subset $Dom(\Phi) = \{x : \Phi(x) \neq \emptyset\}$ is the domain of Φ. The imagine of Φ, the subset of R^m, defined by $Im(\Phi) = \bigcup_{x \in R^n} \Phi(x)$.

The graph of a set-valued map Φ is the subset of R^{n+m} defined by

$$gph\Phi = \{(x, y) : y \in \Phi(x), x \in R^n\}.$$

The inverse Φ^{-1} of Φ is the set valued map from R^m to R^n defined by

$$x \in \Phi^{-1}(y) \text{ if and only if } y \in \Phi(x).$$

We say a set-valued map is convex if its graph is convex.

If the imagines of a set-valued map Φ are closed, convex, bounded, and so on, we say that Φ is a closed-valued map, convex-valued map, bounded-valued map and so on.

Definition 2.2 *An extended-valued function $\phi : R^n \to R \cup \{+\infty\}$ is called lower semicontinuous (l.s.c.) if for each point $x_0 \in R^n$, we have*

$$\liminf_{x \to x_0} \phi(x) \geq \phi(x_0).$$

A function $\phi : R^n \to R \cup \{-\infty\}$ is said to be upper semicontinuous (u.s.c.) if $-\phi$ is l.s.c.

Definition 2.3 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ is said to be upper semicontinuous at $x_0 \in R^n$ if for any neighborhood $O(\Phi(x_0))$ of $\Phi(x_0)$, there exists a neighborhood $O(x_0)$ of x_0 such that*

$$\Phi(x) \subset O(\Phi(x_0)) \text{ for any } x \in O(x_0).$$

Φ is said to u.s.c. if Φ is u.s.c. at every point $x \in R^n$.

Definition 2.4 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ is said to be lower semicontinuous at $x_0 \in R^n$ if for any $y_0 \in \Phi(x_0)$ and any neighborhood $O(y_0)$ of y_0 , there exist a neighborhood $O(x_0)$ of x_0 such that*

$$\Phi(x) \cap O(y_0) \neq \emptyset \text{ for any } x \in O(x_0).$$

Φ is said to be l.s.c. if it is l.s.c. at every $x_0 \in R^n$.

Definition 2.5 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ is said to be continuous at $x_0 \in R^n$ if it is both u.s.c. and l.s.c. at x_0. It is said to be continuous if it is continuous at every point $x_0 \in R^n$.*

An important class of continuous set-valued maps is provided by Lipschitz maps.

Definition 2.6 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ is said to be locally Lipschitz continuous around $x_0 \in R^n$ if there exist a neighborhood $O(x_0)$ and a constant $L_\Phi \geq 0$ such that*

$$\Phi(x) \subset \Phi(y) + L_\Phi \|x - y\| B \quad \forall x, y \in O(x_0).$$

Φ is said to be Lipschitz continuous if there exists a constant $L_\Phi \geq 0$ such that

$$\Phi(x) \subset \Phi(y) + L_\Phi \|x - y\| B \quad \forall x, y \in R^n.$$

Definition 2.7 *For a set-valued map $\Phi : R^n \rightrightarrows R^m$, The set*

$$\limsup_{x \to x_0} \Phi(x) := \{y : \exists x_k \to x_0, y_k \to y \text{ with } y_k \in \Phi(x_k)\}. \tag{2.1}$$

is called the Kuratowski-Painleve upper limit of Φ as $x \to x_0$.

2.2 Generalized Gradients

This section is devoted to reviewing some concepts and results in nonsmooth analysis which will be frequently used in the thesis. Following a geometric approach to the generalized differentiation, we begin with the definition of a normal cone to an arbitrary set in finite dimensions.

Let Ω be a nonempty subset of R^n and let

$$P(x,\Omega)=\{w\in cl\Omega\ :\ \parallel x-w\parallel = dist(x,\Omega)\}$$

be the set of best approximations to x in $cl\Omega$ with respect to the Enclidean distance function $dist(x,\Omega)$.

Definition 2.8 *Given $\bar{x}\in cl\Omega$, the following closed cone*

$$N(\bar{x},\Omega)=\limsup_{x\to\bar{x}}[cone(x-P(x,\Omega)] \tag{2.2}$$

is called the normal cone to the set Ω at the point $\bar{x}$. By convention, $N(\bar{x},\Omega)=\emptyset$ if $\bar{x}\notin\Omega$.

Directly from the definition, one can conclude that the set-valued map $N(\cdot,\Omega)$ always has a closed graph. Furthermore, the condition $N(\bar{x},\Omega)\neq\{0\}$ is necessary and sufficient for $\bar{x}$ being a boundary point $\in cl\Omega$. Observe also that

$$N((\bar{x}_1,\bar{x}_2),\Omega_1\times\Omega_2)=N(\bar{x}_1,\Omega_1)\times N(\bar{x}_2,\Omega_2) \tag{2.3}$$

for any sets Ω_1 and Ω_2 with $(\bar{x}_1,\bar{x}_2)\in cl\Omega_1\times cl\Omega_2$.

In general, the normal cone (2.2) may be nonconvex in the very simple situations, e.g., for the set $\Omega=gph|x|$ at $0\in R^2$ one has

$$N(0,\Omega)=\{(x,y):\quad y\leq -|x|\}\cup\{(x,y):\quad y=|x|\}$$

which is nonconvex.

Definition 2.9 *Given $\bar{x}\in cl\Omega$, the convex closure of $N(\bar{x},\Omega)$ is called the Clarke normal cone to $cl\Omega$ at $\bar{x}$ denoted by $N_C(\bar{x},\Omega)=clcoN(\bar{x},\Omega)$.*

Remark 2.10 *If Ω is a convex set, then both the normal cone and the Clarke normal cone coincide with the normal cone in the sense of the convex analysis, i.e.,* $N(\bar{x}, \Omega) = \{x^* \in R^n : \quad \langle x - \bar{x}, x^* \rangle \geq 0 \quad \forall x \in \Omega\}$

For each $x \in cl\Omega$ and $\epsilon \geq 0$ let us consider the set of the Frechet ϵ-normals to Ω at x defined by

$$\hat{N}_\epsilon(x, \Omega) = \{x^* \in R^n : \limsup_{x'(\in\Omega)\to x} \frac{\langle x^*, x' - x\rangle}{\|x' - x\|} \leq \epsilon\} \tag{2.4}$$

If $\epsilon = 0$, then the set as defined in (2.4) is a closed convex cone which is denoted by $\hat{N}(x, \Omega)$. the following two representations of the normal cone were first obtained by Krugor and Mordukhovich [23]

Proposition 2.11 *For any set $\Omega \subset R^n$ and point $x \in cl\Omega$ one has*

$$N(x, \Omega) = \limsup_{x\to\bar{x}} \hat{N}(x, \Omega) = \limsup_{x\to\bar{x}, \epsilon\downarrow 0} \hat{N}_\epsilon(x, \Omega)$$

We are now ready to define the generalized gradient of extended-valued functions (i.e., those taking values in $R \cup \{+\infty\}$)

Definition 2.12 *Let $f : R^n \to R \cup \{+\infty\}$ be finite at point x. The Clarke generalized gradient of f at x, denoted $\partial_C f(x)$, is the subset of R^n given by*

$$\partial_C f(x) = \{\xi : (\xi, -1) \in N_C((x, f(x)), epi(f))\}.$$

where $epi(f) = \{(x, r) : \quad f(x) \leq r\}$.

As the definition implies, $\partial_C f(x)$ may be empty but not, from the following proposition, if f is locally Lipschitz continuous.

Proposition 2.13 [9, Clarke] *Let f be Lipschitz continuous with modulus L_f near x. Then*

(a) $\partial_C f(x) = \{\xi \in R^n : \langle \xi, v\rangle \leq f^0(x;v) \forall v \in R^n\}$ *where*

$$f^0(x;v) =: \limsup_{y\to x, t\downarrow 0} \frac{f(y+tv) - f(y)}{t},$$

is the Clarke generalized directional derivative of f at x in direction v.

(b) $\partial_C f(x)$ *is a nonempty, convex, compact subset of R^n and $\|\xi\| \leq K$ for every ξ in $\partial_C f(x)$.*

(c) $f^0(x;\cdot)$ *is the support function of* $\partial_C f(x)$. *,i.e.,* $f^0(x;v) = \max\{\langle \xi, v\rangle | \xi \in \partial_C f(x)\}$

(d) *The set-valued map $\partial_C f(x)$ is closed.*

The result below is the characterization of the Clarke generalized gradient. It facilitates greatly the calculation of $\partial_C f(x)$. We recall Rademacher's Theorem which states that a function which is Lipschitz on an open subset of R^n is differentiable almost everywhere (in the sense of Lebesgue measure) on that subset. The set of points at which a given function f fails to be differentiable is denoted Ω_f.

Proposition 2.14 [9, Theorem 2.5.1] *Let f be Lipschitz near x, and suppose S is any set of Lebesgue measure 0 in R^n. Then*

$$\partial_C f(x) = co\{\lim \nabla f(x_i) : x_i \to x, x_i \notin S \cup \Omega_f\}. \tag{2.5}$$

Example 2.15 *Let $f : R \to R$ be given by $f(x) = |x|$. By using the* Proposition 2.14 *it is easy to see that $\partial_C f(0) = co\{-1, 1\} = [-1, 1]$.*

Apparently, if f is continuously differentiable around x then $\partial_C f(x) = \nabla f(x)$. We now proceed to gather an assortment of formulas that faciliate greatly the calculation of $\partial_C f(x)$ when, as is often the case, f is built up from simple function through linear combination, and so on. We still assume that a given function f is Lipschitz continuous near a given point x.

Proposition 2.16 (Scalar Multiples) [9, Proposition 2.3.1] *For any scalar k, one has*

$$\partial_C(kf)(x) = k\partial_C f(x)$$

Proposition 2.17 (Finite Sums) [9, Proposition 2.3.3]

$$\partial_C \sum f_i(x) \subset \sum \partial_C f_i(x).$$

Definition 2.18 *Let $f : R^n \to R \cup \{+\infty\}$ be Lipschitz continuous and finite at x. f is said to be strictly differentiable at x if there exists a continuous mapping $D_s f : R^n \to R^n$ such that*

$$\lim_{x' \to x, t \downarrow 0} \frac{f(x' + tv) - f(x')}{t} = \langle D_s f(x), v \rangle.$$

Corollary 2.19 [9, Corollary1, Proposition 2.3.3] *Equality holds in Proposition 2.17 if all but at most one of the functions f_i are strictly differentiable at x.*

Corollary 2.20 [9, Corollary2, Proposition 2.3.3] *For any scalars s_i, one has*

$$\partial_C \sum s_i f_i(x) \subset \sum s_i \partial_C f_i(x)$$

and equality holds if all but at most one of the f_i are strictly differentiable at x.

As the end of this section, we present below a mean-value theorem which plays an important role in nonsmooth analysis theory and application.

Proposition 2.21 (Mean-valued theorem) [9, Theorem 2.3.7] *Let x and y be points in R^n, and suppose that f is Lipschitz on an open set containing the line segment $[x, y]$. Then there exists a point u in (x, y) such that $f(y) - f(x) \in \langle \partial_C f(u), y - x \rangle$.*

2.3 Derivatives of Set-valued maps

In this section we introduce the concepts of set-valued map derivative which appear naturally and frequently in the area of optimization; see the books of Aubin [3, 2] and the paper of Mordukhovich [23] for many examples and motivations.

Definition 2.22 *Let $\Phi : R^n \rightrightarrows R^m$ be an arbitrary set-valued map and $(\bar{x}, \bar{y}) \in cl(gph\Phi)$. The set-valued map $D^*\Phi(\bar{x}, \bar{y})$ from R^m into R^n defined by*

$$D^*\Phi(\bar{x}, \bar{y})(y^*) = \{x^* \in R^n : (x^*, -y^*) \in N((\bar{x}, \bar{y}), gph\Phi)\} \tag{2.6}$$

is called the coderivative of Φ *at the point* $(\bar{x}, \bar{y})$, *where by convention*

$$D^*\Phi(\bar{x}, \bar{y})(y^*) = \emptyset \text{ if } (\bar{x}, \bar{y}) \notin cl gph\Phi.$$

The symbol $D^*\Phi(\bar{x})$ *is used when* Φ *is single-valued at* $\bar{x}$ *and* $\bar{y} = \Phi(\bar{x})$.

According to this definition, the set-valued map (2.6) is positive homogeneous with respect to y^* ($i.e., D^*\Phi(x, y)(ty^*) = tD^*\Phi(x, y)(y^*)$ for $t > 0$) and the set $D^*\Phi(\bar{x}, \bar{y})(0)$ is a closed cone. Note that the normal cone (2.2) can be expressed as the coderivative (2.6) of the indicator map of the set Ω:

$$\Delta(x, \Omega) = \begin{cases} 0(\in R^m) & \text{if } x \in \Omega \\ \emptyset & \text{otherwise.} \end{cases} \tag{2.7}$$

Indeed,

$$D^*\Delta(\bar{x}, \Omega)(y^*) = N(\bar{x}, \Omega) \ \forall \bar{x} \in \Omega, y^* \in R^m \text{ and } m \geq 1. \tag{2.8}$$

which follows immediately from the definitions and (2.3).

We now consider an extended real-valued function $\varphi : R^n \to R \cup \{+\infty\}$ and its generalized differentiability concepts associated with the coderivative (2.6).

Given φ, we define the epigraphical set-valued map

$$E_\varphi(x) = \{\mu \in R : \mu \geq \varphi(x)\}.$$

It is obvious that $gphE_\varphi = epi\varphi$.

Definition 2.23 *Let* $\bar{x} \in dom\varphi = \{x \in R^n : \varphi(x) \neq +\infty\}$. *The sets*

$$\partial\varphi(\bar{x}) := D^*E_\varphi(\bar{x}, \varphi(\bar{x}))(1) = \{x^* \in R^n | (x^*, -1) \in N(\bar{x}, epi\varphi)\} \tag{2.9}$$

and

$$\partial^\infty\varphi(\bar{x}) := D^*E_\varphi(\bar{x},\varphi(\bar{x}))(0) = \{x^* \in R^n | (x^*,0) \in N(\bar{x}, epi\varphi)\} \quad (2.10)$$

are called, respectively, the subdifferential and the singular subdifferential of φ *at* $\bar{x}$. *If* $\bar{x} \notin dom\varphi$, *we put* $\partial\varphi(\bar{x}) = \partial^\infty\varphi(\bar{x}) = \emptyset$.

The subdifferential (2.9) generalizes the classical concept of subdifferential for nonconvex functions φ and is reduced to the singleton $\{\nabla\varphi(\bar{x})\}$ when φ is strictly defferentiable at $\bar{x}$. When φ is Lipschitz continuous around $\bar{x}$, the set $\partial^\infty\varphi(\bar{x}) = \{0\}$ and $\partial_C\varphi(\bar{x}) = clco\partial\varphi(\bar{x})$. In general, the set (2.9) is nonconvex and

$$\partial_C\varphi(\bar{x}) = clco[\partial\varphi(\bar{x}) + \partial^\infty\varphi(\bar{x})]. \quad (2.11)$$

The following is a finite sum rule for the coderivatives.

Proposition 2.24 (Mordukhovich) *Let* Φ_1 *and* Φ_2 *be set-valued maps from* R^n *to* R^m *with closed graph hand let* $\bar{y} \in (\Phi_1 + \Phi_2)(\bar{x})$. *Assume that the set-valued map* $S : R^{n+m} \rightrightarrows R^{2m}$ *defined by*

$$S(x,y) = \{(y_1,y_2) \in R^{2m} : y_1 \in \Phi_1(x), y_2 \in \Phi_2(x), y_1 + y_2 = y\} \quad (2.12)$$

is locally bounded around $(\bar{x},\bar{y})$ *and the qualification condition*

$$D^*\Phi_1(\bar{x},y_1)(0) \cap (-D^*\Phi_2(\bar{x},y_2)(0)) = \{0\} \quad \forall (y_1,y_2) \in S(\bar{x},\bar{y}) \quad (2.13)$$

is fulfilled. Then for any $y^* \in R^m$ *one has*

$$D^*(\Phi_1+\Phi_2)(\bar{x},\bar{y})(y^*) \subset$$
$$\subset \bigcup_{(y_1,y_2)\in S(\bar{x},\bar{y})} [D^*\Phi_1(\bar{x},y_1)(y^*) + D^*\Phi_2(\bar{x},y_2)(y^*)]. \quad (2.14)$$

The next result follows immediately

Corollary 2.25 [23, Corollary 4.4] *Let $\Phi_1 : R^n \rightrightarrows R^m$ be strictly differentiable at $\bar{x}$ and $\Phi_2 : R^n \rightrightarrows R^m$ be an arbitrary closed set-valued map. Then for any $\bar{y} \in \Phi_1(\bar{x}) + \Phi_2(\bar{x})$ and $y^* \in R^m$ one has*

$$D^*(\Phi_1 + \Phi_2)(\bar{x}, \bar{y})(y^*) = \nabla(\Phi_1(\bar{x}))^\top y^* + D^*\Phi_2(\bar{x}, \bar{y} - \Phi_1(\bar{x}))(y^*) \quad (2.15)$$

Now let us consider some applications of the results obtained above to the first-order subdifferentials of extended-real-valued functions. We always assume that all functions are finite at the points under consideration. We have the following sum formula for subdifferential, which is similar to Proposition 2.17.

Corollary 2.26 [23, Corollary 4.6] *Let the functions $\varphi : R^n \to R, i = 1, 2,$ be l.s.c. around $\bar{x}$. Suppose that the qualification condition*

$$\partial^\infty \varphi_1(\bar{x}) \cap (-\partial^\infty \varphi_2(\bar{x})) = \{0\} \quad (2.16)$$

is fulfilled. Then one has the inclusions

$$\partial(\varphi_1 + \varphi_2)(\bar{x}) \subset \partial\varphi_1(\bar{x}) + \partial\varphi_2(\bar{x}) \quad (2.17)$$

$$\partial^\infty(\varphi_1 + \varphi_2)(\bar{x}) \subset \partial^\infty\varphi_1(\bar{x}) + \partial^\infty\varphi_2(\bar{x}), \quad (2.18)$$

where equalities hold when either φ_1 or φ_2 is strictly differentiable at $\bar{x}$.

The results in Corollary 2.26 imply the following important representation of the normal cone (2.2) to the intersection of sets which was first obtained by Mordukhovich [23]

Corollary 2.27 [23, Corollary 4.7] *Let Ω_1 and Ω_2 be closed sets in R^n and let $\bar{x} \in \Omega_1 \cap \Omega_2$. If $N(\bar{x}, \Omega_1) \cap (-N(\bar{x}, \Omega_2)) = \{0\}$, then one has the inclusion*

$$N(\bar{x}, \Omega_1 \cap \Omega_2) \subset N(\bar{x}, \Omega_1) + N(\bar{x}, \Omega_2).$$

In particular, if both Ω_1 and Ω_2 are closed convex sets, we have the following result.

Proposition 2.28 [3, P.141] *If $0 \in Int(\Omega_1 - \Omega_2)$, then for any $x \in \Omega_1 \cap \Omega_2$ one has*

$$N(x, \Omega_1 \cap \Omega_2) = N(x, \Omega_1) + N(x, \Omega_2).$$

Definition 2.29 *Ω, a subset of R^n is said to be regular at $\bar{x}$ if $N(\bar{x}, \Omega) = \hat{N}(\bar{x}, \Omega)$. $f : R^n \to R$ is said to be regular at $\bar{x}$ if its epi-graph, $epi(f) = \{(x, r) : \quad f(x) \leq r\}$ is regular at $\bar{x}$.*

Remark 2.30 *If Ω is regular at $\bar{x}$, then $N(\bar{x}, \Omega)$ is convex and hence $N(\bar{x}, \Omega) = N_C(\bar{x}, \Omega)$.*

Apparently, if $f(x)$ is regular at $\bar{x}$, then $\partial_C f(\bar{x}) = \partial f(\bar{x})$.

Let $f(x, y)$ be Lipschitz continuous near $(\bar{x}, \bar{y})$. The following proposition is a relationship between generalized gradients and their partial generalized gradients.

Proposition 2.31 [9, Proposition 2.3.15] *Let $f(x, y)$ be Lipschitz continuous near $(\bar{x}, \bar{y})$. If $f(x, y)$ is regular at $(\bar{x}, \bar{y})$ then we have*

$$\partial f(\bar{x}, \bar{y}) \subset \partial_x f(\bar{x}, \bar{y}) \times \partial_y f(\bar{x}, \bar{y}).$$

We now introduce another kind of derivative of set-valued map called paratingent derivative, which will be used in Chapter 3.

Definition 2.32 *Let Ω be a closed subset of R^n, $\bar{x} \in \Omega$.*

$$P_\Omega(\bar{x}) = \limsup_{h \downarrow 0, x(\in \Omega) \to \bar{x}} (\Omega - x)/h$$

is called the paratingent cone to Ω at $\bar{x} \in \Omega$.

Definition 2.33 *Let $F : R^n \Rightarrow R^m$ be a closed set-valued map. The paratingent derivative of F at $(\bar{x}, \bar{y}) \in gph(F)$ denoted by $D_pF(\bar{x}, \bar{y})$, is a set-valued map from R^n to R^m, defined by*

$$gphD_pF(\bar{x}, \bar{y}) = P_{gphF}(\bar{x}, \bar{y}).$$

By definition, it is easy to see that the following result is true.

Proposition 2.34 *Let F is defined as above, $(\bar{x}, \bar{y}) \in gphF$. Then*

$$\begin{aligned} & v \in D_pF(\bar{x}, \bar{y})(\bar{u}) \Leftrightarrow \\ & \exists h_n \downarrow 0, u_n \to \bar{u}, (x_n, y_n)(\in gphF) \to (\bar{x}, \bar{y}) \\ & \text{and } \exists y_n^* \in F(x_n + h_n u_n) \text{ such that } \lim(y_n^* - y_n)/h_n = v. \end{aligned}$$

Corollary 2.35 *Let Ω be closed subset in R^n and $\varphi : R^n \to R \cup \{+\infty\}$ be Lipschitz continuous. Suppose that $\bar{x}$ minimizes $\psi(x)$ on Ω. Then we have*

$$0 \in \partial\varphi(\bar{x}) + N(\bar{x}, \Omega). \tag{2.19}$$

Note that from Definition 2.9 and equality (2.11) we also have in this case that $0 \in \partial_C\varphi(\bar{x}) + N_C(\bar{x}, \Omega)$. In particular, $0 \in \partial\varphi(\bar{x}) \subset \partial_C\varphi(\bar{x})$ if φ has a local minimum at $\bar{x}$.

As the end of this chapter, we introduce the Ekeland's Variational Principle which is a powerful tool in nonsmooth analysis and its application. Readers who are interested in the details are refered to [14] where lots of examples are provided.

Proposition 2.36 (Ekeland's variational principle) *Let X be a complete metric space, and $f : X \to R \cup \{+\infty\}$ a proper l.s.c. function, bounded from below. For every point $x_0 \in X$ satisfying* $\inf f \le f(x_0) < \inf f + \epsilon$ *and every $\lambda > 0$ there exists some point $x^* \in X$ such that*

(a) $f(x^*) \le f(x_0) < \inf f + \epsilon$

(b) $d(x^*, x_0) \le \lambda$

(c) $\forall x \neq x^*, f(x) > f(x^*) - (\epsilon/\lambda)d(x^*, x)$.

Chapter 3

Necessary Optimality Conditions Involving Paratingent Derivatives

In this chapter, we derive a Kuhn-Tucker type necessary optimality conditions involving paratingent derivatives.

The following is an outline of the chapter.

§3.1. Problem Formulation. In this section, we are concerned with the problem and its reformulation.

§3.2. Preliminaries. In this section, we derive some lemmas for use in subsequent arguments.

§3.3. Main Theorems. In this section, we provide a necessary optimality condition for problem (OPVIC).

3.1 Problem Formulation

In this chapter, we consider problem (OPVIC) with Ω_1 being a closed convex subset and Ω_2 a closed convex Lipschitz set-valued map. We rewrite the problem as follows:

$$\begin{aligned} \text{(OPVIC)} \quad & \min \quad f(x,y) \\ & \text{s.t.} \quad x \in \Omega_1, y \in \Omega_2(x) \\ & \qquad \langle F(x,y), z-y \rangle \geq 0 \ \forall z \in \Omega_2(x). \end{aligned} \tag{3.1}$$

We will assume in this chapter that

(H1) $f : R^{n+m} \to R$ is Lipschitz continous and

$$\liminf_{\|x\|\to\infty, y\in\Omega_2(x)} f(x,y) = \infty, \quad \inf_{x\in\Omega_1, y\in\Omega_2(x)} f(x,y) > -\infty.$$

(H2) $F : R^{n+m} \to R^m$ is continuously differentiable and strongly monotone in y uniformly in $x \in \bar{B}_R := \{x : \|x\| \leq R\}$ for any $R > 0$, i.e., there exists $C_R > 0$ such that

$$\begin{aligned} \langle F(x,y_1) - F(x,y_2), y_1 - y_2 \rangle \geq C_R \|y_1 - y_2\|^2 \\ \forall y_1, y_2 \in \Omega(x), \quad x \in \bar{B}_R. \end{aligned}$$

Under these assumptions, it is not difficult to prove the existence of a solution to problem (OPVIC).

Theorem 3.1 *Suppose that* (H1), (H2) *hold. Then the problem* (OPVIC) *has a solution* $\bar{x} \in \Omega_1, \bar{y} \in \Omega_2(\bar{x})$.

Proof. Using Browder's [8] and Hartman and Stampacchin's [20] theorem, by the assumption (H2), we have that for any $x \in \Omega_1$, there exists a unique $y_x \in \Omega_2(x)$ such that (3.1) holds. Choosing sequence $\{x_n\} \subset \Omega_1$ with the property:

$$\lim_{n\to\infty} f(x_n, y_n) = \inf_{x\in\Omega_1} f(x, y_x), \tag{3.2}$$

where $y_n = y_{x_n}$. From assumption (H2), we know $\{x_n\}$ is bounded, that is, there is a $R > 0$ such that $x_n \in \bar{B}_R$ for all n. Therefore we can suppose that $x_n \to \bar{x} \in \Omega_1$. Now let $\bar{y} = y_{\bar{x}}$, then we have

$$\langle F(x_n, y_n), z - y_n\rangle \geq 0 \ \forall z \in \Omega_2(x_n) \tag{3.3}$$

$$\langle F(\bar{x}, \bar{y}), z - \bar{y}\rangle \geq 0 \ \forall z \in \Omega_2(\bar{x}). \tag{3.4}$$

Note that by the Lipschitz continuity of Ω_2, for any $m > 0$, there exists $\sigma_m > 0$ such that $\Omega_2(\bar{x} + \sigma_m B) \subset \Omega_2(\bar{x}) + \frac{1}{m}B$ and hence there exists N_m such that

$$\Omega_2(x_n) \subset \Omega_2(\bar{x}) + \frac{1}{m}B \ \ \forall n > N_m. \tag{3.5}$$

Therefore we can take $y_n^* \to \bar{y}$ with $y_n^* \in \Omega_2(x_n)$ (due to the continuity of Ω_2) and $z_n \in \Omega_2(\bar{x}), \xi_n \in B$ such that $y_n = z_n + \frac{1}{m}\xi_n$ (from (3.5)). We obtain, when $n > N_m$ and taking $z = y_n^*$ in (3.3) $z = z_n$ in (3.4), that

$$\langle F(\bar{x}, \bar{y}), z - \bar{y}\rangle \geq \langle F(\bar{x}, \bar{y}), \frac{1}{m}\xi_n\rangle \tag{3.6}$$

$$\langle F(x_n, y_n), y_n^* - y_n\rangle \geq 0 \tag{3.7}$$

which implies

$$\langle F(x_n, y_n) - F(x_n, \bar{y}), y_n - \bar{y}\rangle \ \leq \ \langle F(\bar{x}, \bar{y}) - F(x_n, \bar{y}), y_n - \bar{y}\rangle$$

$$+\langle F(x_n, y_n), y_n^* - \bar{y}\rangle - \langle F(\bar{x}, \bar{y}), \frac{1}{m}\xi_n\rangle. \tag{3.8}$$

Our aim is to show that $\{y_n\}$ converges $\bar{y}$. Indeed, if this is not true, then there exists $\delta > 0$ such that $\|y_{n_k} - \bar{y}\| > \delta$ for all k, where $\{y_{n_k}\}$ is a subsequence of $\{y_n\}$. By $y_{n_k}^* \to \bar{y}$ we know for large enough n_k, $\|y_{n_k}^* - \bar{y}\| \le \delta$ and hence $\|y_{n_k}^* - y_{n_k}\| \le 2\|y_{n_k} - \bar{y}\|$. Remembering the assumption (H2) we have

$$\begin{aligned} C_R\|y_{n_k} - \bar{y}\|^2 \le & \quad \langle F(x_{n_k}, y_{n_k}) - F(x_{n_k}, \bar{y}), y_{n_k} - \bar{y}\rangle \\ \le & \quad \langle F(\bar{x}, \bar{y}) - F(x_{n_k}, \bar{y}), y_{n_k} - \bar{y}\rangle \\ & \quad +\langle F(x_{n_k}, y_{n_k}), y_{n_k}^* - \bar{y}\rangle - \frac{1}{m}\langle F(\bar{x}, \bar{y}), \xi\rangle \end{aligned}$$

that is,

$$\begin{aligned} C_R\|y_{n_k} - \bar{y}\| \ \le \ & \langle F(\bar{x}, \bar{y}) - F(x_{n_k}, \bar{y}), \frac{y_{n_k} - \bar{y}}{\|y_{n_k} - \bar{y}\|}\rangle \\ & +\frac{1}{\|y_{n_k} - \bar{y}\|}\langle F(x_{n_k}, y_{n_k}), y_{n_k}^* - \bar{y}\rangle - \frac{1}{m\|y_{n_k} - \bar{y}\|}\langle F(\bar{x}, \bar{y}), \xi\rangle \\ \le \ & \|F(\bar{x}, \bar{y}) - F(x_{n_k})\| \\ & +\frac{1}{\delta}\|F(x_{n_k}, y_{n_k})\|\|y_{n_k}^* - \bar{y}\| + \frac{1}{\delta m}\|F(\bar{x}, \bar{y})\|. \end{aligned} \tag{3.9}$$

Apparently, in the inequality of (3.9), the value of right hand side converges to 0 which contradicts the assumption of $\|y_{n_k} - \bar{y}\| > \delta$. Therefore $\{y_n\} \to \bar{y}$ holds and by assumption (H1) as well as (3.2), (3.4), it is easy to see that $(\bar{x}, \bar{y})$ is a solution to (OPVIC). □

In the following result we prove that problem (OPVIC) can be reformulated as problem $(\tilde{P})$ which is more convenient to deal with by the penalty method.

Lemma 3.2 *The following problem* $(\tilde{P})$ *is equivalent to problem* (OPVIC).

$$(\tilde{P}) \quad \min \quad f(x,y)$$
$$s.t. \quad x \in \Omega_1, y \in \Omega_2(x)$$
$$\| -F(x,y)\|_y^{\Omega_2(x)} = 0, \tag{3.10}$$

where

$$\|y^*\|_y^{\Omega_2(x)} := \sup_{p \in (\Omega_2(x)-y)\cap \bar{B}} \langle y^*, p\rangle.$$

Proof. First note that if $x \in \Omega_1, y \in \Omega_2(x)$ such that $\| -F(x,y)\|_y^{\Omega_2(x)} = 0$, that is,

$$\sup_{p \in (\Omega_2(x)-y)\cap \bar{B}} \langle -F(x,y), p\rangle = 0,$$

then for any $z \in \Omega_2(x) \cap (y + \bar{B})$ we have

$$\langle -F(x,y), z - y\rangle \leq 0. \tag{3.11}$$

Now, for any $z^* \in \Omega_2(x)$ but $\|z^* - y\| > 1$, by convexity of $gph\Omega_2|_{\Omega_1}$ where $gph\Omega_2|_{\Omega_1} = (gph\Omega_2) \cap (\Omega_1 \times R^m)$, we know $\lambda(x,y) + (1-\lambda)(x,z^*) \in gph\Omega_2|_{\Omega_1}$ for all $\lambda \in (0,1)$. Taking λ with $0 < 1 - \frac{1}{\|z^*-y\|} < \lambda < 1$, we obtain $\|\lambda y + (1-\lambda)z^* - y\| \leq 1$ and hence from (3.11)

$$\langle -F(x,y), (1-\lambda)(z^* - y)\rangle = \langle -F(x,y), \lambda y + (1-\lambda)z^* - y\rangle \leq 0 \tag{3.12}$$

Combining (3.11) and (3.12), it is easy to see (x,y) satisfies the contraints (3.1).

Conversely, if $\langle F(x,y), z-y\rangle \geq 0 \; \forall z \in \Omega_2(x)$, then apparently

$$\sup_{p \in (\Omega_2(x)-y)\cap \bar{B}} \langle -F(x,y), p\rangle = 0. \qquad \square$$

3.2 Preliminaries

In this section, we provide some lemmas which will be used as tools in our main theorems. In the following we will always assume that Ω_2 is defined as the one in problem (OPVIC).

Lemma 3.3 *Assume F are defined as in problem* (OPVIC). *Then we have*

(1) $\Phi : gph\Omega_2 \to R$ *defined by*

$$\Phi(x,y) = \| - F(x,y)\|_y^{\Omega_2(x)} \tag{3.13}$$

is continuous.

(2) $M : gph\Omega_2 \rightrightarrows \bar{B}$ *defined by*

$$M(x,y) \quad = \quad \{p \in (\Omega_2(x) - y) \cap \bar{B} : \langle -F(x,y), p\rangle = \| - F(x,y)\|_y^{\Omega_2(x)}\} \tag{3.14}$$

is upper semicontinuous.

Proof: Let $g : \bar{B} \times gph\Omega_2 \to R$ defined as

$$g(p; x, y) = \langle -F(x,y), p\rangle$$

and $W : gph\Omega_2 \rightrightarrows \bar{B}$ defined by

$$W(x,y) = (\Omega_2(x) - y) \cap \bar{B}.$$

Apparently, g is continuous and W is continuous with nonempty compact values. Now notice we have that

$$\Phi(x,y) = \sup_{p \in W(x,y)} g(p; x, y) \text{ and}$$

$$\begin{aligned} M(x,y) &= \{p \in W(x,y) : g(p;x,y) = \Phi(x,y)\} \\ &= \{p \in \bar{B} : g(p;x,y) = \Phi(x,y)\} \cap W(x,y). \end{aligned}$$

Using Berge's Maximum theorem [7] and [2, theorem 8], we have the assertion.□

In fact, it is easy to see that for any $(x,y) \in ghp\Omega_2$, $M(x,y)$ is nonempty, compact and convex.

The following lemma provides a characterization of a convex set-valued map.

Lemma 3.4 *A set-valued map $G : \Omega \Rightarrow R^m$ is convex if and only if for all $x_1, x_2 \in \Omega$ and $\lambda \in (0,1)$ one has*

$$\lambda G(x_1) + (1-\lambda)G(x_2) \subset G(\lambda x_1 + (1-\lambda)x_2) \tag{3.15}$$

where Ω is a convex subset of R^n.

Proof. Suppose G is a convex set-valued map. Then for any $z \in \lambda G(x_1) + (1-\lambda)G(x_2)$, where $x_1, x_2 \in \Omega$, and $\lambda \in (0,1)$, there exist $y_i \in G(x_i), i = 1,2$, such that $z = \lambda y_1 + (1-\lambda)y_2$. Since $(x_i, y_i) \in gphG, i = 1,2$, we have

$$\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) = (\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \in gphG$$

that is $z = \lambda y_1 + (1-\lambda)y_2 \in G(\lambda x_1 + (1-\lambda)x_2)$ and hence $\lambda G(x_1) + (1-\lambda)G(x_2) \subset G(\lambda x_1 + (1-\lambda)x_2)$.

Conversely, if for any $x_1, x_2 \in \Omega$, $\lambda \in (0,1)$ (3.15) is true, then for any $(x_i, y_i) \in gphG, i = 1,2$, we need to show $\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) \in gphG$. In fact, by $y_i \in G(x_i)$, we have

$$\lambda y_1 + (1-\lambda)y_2) \in \lambda G(x_1) + (1-\lambda)G(x_2) \subset G(\lambda x_1 + (1-\lambda)x_2).$$

Therefore

$$\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) = (\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \in ghpG.$$

The proof is complete. □

Lemma 3.5 *Let Ω be a convex subset of R^n, $G : \Omega \rightrightarrows R^m$ a convex set-valued map and $f : D \to R^n$ is a convex function where D is a convex subset of R^m such that $G(y) \subset D$ for all $y \in \Omega$. Then the value function $V(y) = \inf_{x\in G(y)} f(x)$ is convex on Ω.*

Proof. In fact, for any $y_1, y_2 \in \Omega$, and $\lambda \in (0,1)$,

$$\begin{aligned}
V(\lambda y_1 + (1-\lambda)y_2) &= \inf_{x\in G(\lambda y_1+(1-\lambda)y_2)} f(x) \\
&\leq \inf_{x\in \lambda G(y_1)+(1-\lambda)G(y_2)} f(x) \text{ (by lemma 3.4)} \\
&= \inf_{z_1\in G(y_1), z_2\in G(y_2)} f(\lambda z_1 + (1-\lambda)z_2) \\
&\leq \inf_{z_1\in G(y_1), z_2\in G(y_2)} \{\lambda f(z_1) + (1-\lambda) f(z_2)\} \\
&= \inf_{z_1\in G(y_1)} \lambda f(z_1) + \inf_{z_2\in G(y_2)} (1-\lambda) f(z_2) \\
&= \lambda V(y_1) + (1-\lambda)V(y_2). \qquad \square
\end{aligned}$$

The following lemma provides an estimate to a value function.

Lemma 3.6 *Let $(a,b) \in gph\Omega_2$, F be defined as in problem* (OPVIC), *and the value function $V(x,y)$ defined by*

$$V(x,y) = \inf_{p\in S(x,y)} U(p)$$

where $U(p) = \langle F(a,b), p\rangle$, $S(x,y) = N[(\Omega_2(a+x) - b - y) \cap \bar{B}]$, then

$$-V'((0,0);(0,\bar{p})) \leq (N\| - F(a,b)\|_b^{\Omega_2(a)})\|\bar{p}\|$$

.

Proof. First, we want to show

$$S(0,tp) \subset (1+t\|p\|)S(0,0) - tNp. \tag{3.16}$$

Indeed, for $\forall x \in S(0,tp)$, there exists a $z \in \Omega_2(a)-b$ such that $x = N(z-tp)$, that is, $Nz = x + tNp$ which implies $x + tNp \in N(\Omega_2(a)-b)$. Meanwhile, by $x \in N\bar{B}$ we have

$$x + tNp \in N\bar{B} + tNp \subset N(1+t\|p\|)\bar{B}.$$

Therefore

$$\begin{aligned} x + tNp &\in N(\Omega_2(a)-b) \cap N(1+t\|p\|)\bar{B} \\ &\subset N(1+t\|p\|)[(\Omega_2(a)-b) \cap \bar{B}] \\ &= (1+t\|p\|)S(0,0). \end{aligned}$$

It follows that (3.16) holds. Now we turn to estimate the derivative $V'((0,0),(0,\bar{p}))$. Note that we have

$$\begin{aligned} V'((0,0),(0,\bar{p})) &= \lim_{t\downarrow 0} \frac{V(0,t\bar{p}) - V(0,0)}{t} \\ &= \lim_{t\downarrow 0} [\inf_{p\in S(0,t\bar{p})} U(p) - \inf_{p\in S(0,0)} U(p)]/t \\ &\geq \lim_{t\downarrow 0} [\inf_{p\in(1+t\|\bar{p}\|)S(0,0)-tN\bar{p}} U(p) - \inf_{p\in S(0,0)} U(p)]/t \\ &= \lim_{t\downarrow 0} [(1+t\|\bar{p}\|) \inf_{p\in S(0,0)} U(p) - tN\|\bar{p}\|U(\bar{p}) - \inf_{p\in S(0,0)} U(p)]/t \\ &= \|\bar{p}\| \inf_{p\in S(0,0)} U(p) - N\|\bar{p}\|U(\bar{p}) \\ &= -\|\bar{p}\|(N\| - F(a,b)\|_b^{\Omega_2(a)}) + N\langle -F(a,b), \bar{p}\rangle \\ &\geq -\|\bar{p}\|(N\| - F(a,b)\|_b^{\Omega_2(a)}). \end{aligned}$$

It follows that

$$-V'((0,0),(0,p)) \leq \|\bar{p}\|(N\| - F(a,b)\|_b^{\Omega_2(a)}). \qquad \square$$

Lemma 3.7 *Let $\Phi : R^n \Rightarrow R^m$ be a set-valued map with a closed graph. Then*

(1) $(x, y) \in gph\Phi$ *iff* $(y, x) \in gph\Phi^{-1}$

(2) $(v, w) \in N((x, y), gph\Phi)$ *iff* $(w, v) \in N((y, x), gph\Phi^{-1})$.

Proof. Obviously, (1) follows by the definition of Φ^{-1}. As to (2), by definition of $\hat{N}((x', y'), gph\Phi)$ and (1) we have $(v, w) \in \hat{N}((x', y'), gph\Phi)$ if only if $(w, v) \in \hat{N}((y', x'), gph\Phi_{-1})$. Using the equality $N((x, y), gph\Phi) = \limsup_{(x',y')\to(x,y)} \hat{N}((x', y'), gph\Phi)$ (2.11) and definition 2.7 we obtain the result (2). □

Lemma 3.8 *Suppose $f : R^n \to R$ is a Lipschitz continuous function, Ω a closed convex subset of R^n. If $f^0(x; v) \geq \langle g(x), v\rangle$ for all $v \in \Omega - y$ where $y \in \Omega$ and g is a mapping from R^n to R^n. Then one has*

$$g(x) \in \partial_C f(x) + N(y, \Omega).$$

Proof. We show it by contradiction. In fact, if $g(x) \notin \partial_C f(x) + N(y, \Omega)$, then for any $\xi \in \partial_C f(x)$ we have

$$g(x) - \xi \notin N(y, \Omega)$$

which implies that there exists a $v_0 \in \Omega - y$ such that $\langle g(x) - \xi, v_0\rangle > 0$ and hence $\langle g(x), v_0\rangle > \langle \xi, v_0\rangle$ for all $\xi \in \partial_C f(x)$. Therefore, by Proposition 2.13, it follows that $\langle g(x), v_0\rangle > f^0(x; v_0)$ which contradicts the assumption. □

The following result was proved in [2, Corollary 3] .

Lemma 3.9 *Let $g : R^n \to R \cup \{+\infty\}$ be a proper lower semicontinuous convex function and S a set-valued mapping from R^m to R^n with closed convex graph. We assume that*

$$0 \in Int(Im(S) - Dom(g)).$$

Let $V : R^m \to R \cup \{+\infty\}$ be the value function defined by

$$V(y) = \inf_{x \in S(y)} g(x)$$

and let $\bar{x} \in S(\bar{y})$ achieve the minimum of g on $S(\bar{y})$. Then the generalized gradient of the value function V is equal to:

$$\partial_C V(\bar{y}) = \bigcup_{p \in \partial g(\bar{x})} D^* S(\bar{y}, \bar{x})(p).$$

3.3 Main Theorems

The purpose of this section is to derive a necessary optimality condition involving paratingent derivatives for problem (OPVIC). The technique used here is similar to [35, 36, Shuzhong Shi] in which Ω_2 is assumed to be independent of x. In order to simplify our discussion, we first assume that $\Omega_1 = R^n$ and then we will generalize the result to the general case.

Consider problem (OPVIC) in (3.1) with $\Omega_1 = R^n$, we have the following theorem:

Theorem 3.10 *Suppose that assumptions* (H1), (H2) *hold. Then there exist $(\bar{x}, \bar{y})$, a solution of* (OPVIC) *in* (3.1) *and $\bar{p} \in R^m$ such that*

$$\nabla_y F(\bar{x}, \bar{y})^\top \bar{p} \in \partial_y f(\bar{x}, \bar{y}) - D_p \Phi_1((\bar{x}, \bar{y}), -F(\bar{x}, \bar{y}))(0, \bar{p}) \qquad (3.17)$$

$$\begin{aligned} \nabla_x F(\bar{x},\bar{y})^\top \bar{p} &\in \partial_x f(\bar{x},\bar{y}) - \bigcup_{y^*\in\Phi_2(\bar{x},\bar{y})} D_p\Phi_2((\bar{x},\bar{y}),y^*)(0,\bar{p}) \quad (3.18)\\ \langle F(\bar{x},\bar{y}),\bar{p}\rangle &= 0 \quad (3.19)\end{aligned}$$

where $\Phi_1(x,y) = N(y,\Omega_2(x)),\ \Phi_2(x,y) = N(x,\Omega_2^{-1}(y))$.

Proof. Due to the equivalence between problem (OPVIC) and $(\tilde{P})$ (c.f. Lemma 3.2) we know from the inexact penalty method that the problem (OPVIC) can be approximated by

$$\begin{aligned} (P_n) \quad &\min \quad \{f(x,y) + N_n\| - F(x,y)\|_y^{\Omega_2(x)}\}\\ &\text{s.t.} \quad x \in R^n, y \in \Omega_2(x)\end{aligned}$$

where $N_n > 0$ is a 'penalty factor' and $N_n \to +\infty$.

Denote $G_n(x,y) = f(x,y) + N_n\| - F(x,y)\|_y^{\Omega_2(x)}$ for all $(x,y) \in gph\Omega_2$. From Lemma 3.3 and assumption $(H1)$, we know that G_n is l.s.c. and

$$\inf_{(x,y)\in gph\Omega_2} G_n(x,y) > -\infty.$$

Now using Ekeland's Variational Principle (Proposition 2.36), for any $\epsilon_n \downarrow 0$ there exists $(x_n, y_n) \in gph\Omega_2$ and a constant C such that

$$C \le \inf_{(x,y)\in gph\Omega_2} G_n(x,y) \le G_n(x_n,y_n) < \inf_{(x,y)\in gph\Omega_2} G_n(x,y) + \epsilon_n, \quad (3.20)$$

and

$$G_n(x,y) > G_n(x_n,y_n) - \epsilon_n(\|x-x_n\|^2 + \|y-y_n\|^2)^{1/2} \quad (3.21)$$

for any $(x,y) \neq (x_n,y_n)$.

Taking any $(w,s) \in gph\Omega_2 - (x_n,y_n)$, $t_k \downarrow 0$ we have from (3.21)

$$G_n(x_n + t_k w, y_n + t_k s) > G_n(x_n,y_n) - \epsilon_n t_k(\|s\|^2 + \|w\|^2)^{1/2}. \quad (3.22)$$

To complete the proof, we need the following lemmas. Define

$$S_n : R^{n+m} \Rightarrow R^m \text{ by } S_n(x,y) = N_n[(\Omega_2(x_n+x) - y_n - y) \cup \bar{B}],$$
$$U_n : R^m \to R \text{ by } \; U_n(p) = \langle F(x_n, y_n), p\rangle, V_n(x,y) = \inf_{p\in S_n(x,y)} U_n(p).$$

Lemma 3.11 *Assume $M(\cdot,\cdot)$, defined by (3.14) in Lemma 3.3, is a single-valued map. Then one has*

$$f_y^0((x_n,y_n);s) - V_n'((0,0),(0,s)) \geq \langle \nabla_y F(x_n,y_n)^\top p_n, s\rangle - \epsilon_n\|s\|,$$
$$\textit{for any } s \in \Omega_2(x_n) - y_n. \quad (3.23)$$

$$f_x^0((x_n,y_n);w) - V_n'((0,0),(w,0)) \geq \langle \nabla_x F(x_n,y_n)^\top p_n, w\rangle - \epsilon_n\|w\|,$$
$$\textit{for any } w \in \Omega_2^{-1}(y_n) - x_n. \quad (3.24)$$

where $p_n = M_n(x_n,y_n) := N_n M(x_n,y_n)$.

Proof. Let $p_{nk}^{ws} = M_n(x_n + t_k w, y_n + t_k s)$ where $M_n(x,y) := N_n M(x,y)$. We have

$$\begin{aligned} & f(x_n+t_kw, y_n+t_ks) - f(x_n,y_n) + \langle -F(x_n+t_kw, y_n+t_ks), p_{nk}^{ws}\rangle \\ & \quad \geq \langle -F(x_n,y_n), p_{nk}^{ws}\rangle + \langle F(x_n,y_n), p_{nk}^{ws}\rangle + N_n\| - F(x_n,y_n)\|_{y_n}^{\Omega_2(x_n)} - \\ & \quad\quad -\epsilon_n t_k(\|s\|^2 + \|w\|^2)^{1/2}. \end{aligned} \quad (3.25)$$

Note that

$$\langle F(x_n,y_n), p_{nk}^{ws}\rangle \geq \inf_{p\in S_n(t_kw,t_ks)} \langle F(x_n,y_n), p\rangle$$

$$\begin{aligned} N_n\| - F(x_n,y_n)\|_{y_n}^{\Omega_2(x_n)} &= \sup_{p\in S_n(0,0)} \langle -F(x_n,y_n), p\rangle \\ &= -\inf_{p\in S_n(0,0)} \langle F(x_n,y_n), p\rangle. \end{aligned}$$

It is easy to see that S_n is a convex set-valued map. By Lemma 3.5, $V_n(x,y)$ is a convex function. Therefore, from (3.25) we obtain that

$$\begin{aligned} & f(x_n+t_kw, y_n+t_ks) - f(x_n,y_n) + \langle -F(x_n+t_kw, y_n+t_ks), p_{nk}^{ws}\rangle \\ & \geq \langle -F(x_n,y_n), p_{nk}^{ws}\rangle + \inf_{p\in S_n(t_kw,t_ks)} U_n(p) - \inf_{p\in S_n(0,0)} U_n(p) - \\ & \quad -\epsilon_n t_k(\|s\|^2+\|w\|^2)^{1/2} \\ & = \langle -F(x_n,y_n), p_{nk}^{ws}\rangle + V_n(t_kw,t_ks) - V_n(0,0) - \\ & \quad -\epsilon_n t_k(\|s\|^2+\|w\|^2)^{1/2}. \end{aligned} \tag{3.26}$$

That is

$$\begin{aligned} & \{f(x_n+t_kw, y_n+t_ks) - f(x_n,y_n)\}/t_k - \\ & -\langle (F(x_n+t_kw, y_n+t_ks) - F(x_n,y_n))/t_k, p_{nk}^{ws}\rangle \\ & \geq (V_n(t_kw,t_ks) - V_n(0,0))/t_k - \epsilon_n(\|s\|^2+\|w\|^2)^{1/2}. \end{aligned}$$

By assumption of the problem, there exists $\theta_k \in [0,1]$ such that

$$F(x_n+t_kw, y_n+t_ks) - F(x_n,y_n) = \nabla F(x_n+t_k\theta_kw, y_n+t_k\theta_ks)^T(w,s)t_k.$$

And hence,

$$\begin{aligned} & f^0((x_n,y_n);(w,s)) - \langle \nabla F(x_n,y_n)^T(w,s), p_n\rangle \\ & \quad \geq V_n'((0,0),(w,s)) - \epsilon_n(\|s\|^2+\|w\|^2)^{1/2}, \end{aligned} \tag{3.27}$$

where $p_n = M_n(x_n,y_n)$ ($M_n(x_n+t_kw, y_n+t_ks) \to M_n(x_n,y_n)$) for any $(w,s) \in gph\Omega_2 - (x_n,y_n)$. Therefore, for any $s \in \Omega_2(x_n) - y_n, w = 0$ we obtain

$$f_y^0((x_n,y_n);s) - V_n'((0,0),(0,s)) \geq \langle \nabla_y F(x_n,y_n)^\top p_n, s\rangle - \epsilon_n\|s\|$$

and for any $w \in \Omega_2^{-1}(y_n) - x_n, s = 0$

$$f_x^0((x_n, y_n); w) - V_n'((0,0),(w,0)) \geq \langle \nabla_x F(x_n, y_n)^\top p_n, w\rangle - \epsilon_n \|w\|,$$

which are (3.23) and (3.24) respectively. The proof is complete. □

Lemma 3.12 *Suppose the inequalities* (3.23) *and* (3.24) *hold. Then there exists a sub-sequence of* (x_n, y_n, p_n) *which converges to* $(\bar{x}, \bar{y}, \bar{p})$ *such that* $(\bar{x}, \bar{y})$ *is a solution of problem* (OPVIC) *and the condition* (3.17), (3.18) *and* (3.19) *hold.*

Proof. Step 1. We want to show the existence of $(\bar{x}, \bar{y})$.

First,we show $\{x_n\}$ is bounded. In fact, by the inequality (3.20), we have

$$C \leq G_n(x_n, y_n) < G_n(x, y) + \epsilon_n \ \forall (x, y) \in gph\Omega_2.$$

Especially taking $y = y_x$ as defined in Theorem 3.1 we obtain

$$C \leq f(x_n, y_n) + N_n \| - F(x_n, y_n)\|_{y_n}^{\Omega_2(x_n)} < f(x, y_x) + \epsilon_n, \ \forall x \in R^n \quad (3.28)$$

which implies that

(1) $f(x_n, y_n)$ is bounded and therefore by assumption $(H1)$ $\{x_n\}$ is bounded. We then can assume $x_n \to \bar{x}$.

(2) $\delta_n = \| - F(x_n, y_n)\|_{y_n}^{\Omega_2(x_n)} \to 0$.

Now we show $y_n \to \bar{y}$. First, we have

$$\langle F(x_n, y_n), z - y_n\rangle \geq -\delta_n \max\{1, \|z - y_n\|\} \ \forall z \in \Omega_2(x_n). \qquad (3.29)$$

Indeed, by definition, $\forall p \in (\Omega_2(x_n) - y_n) \cap \bar{B}$ we have $\langle F(x_n, y_n), p \rangle \geq -\delta_n$ which implies for any $z \in \Omega_2(x_n)$ if $\|z - y_n\| \leq 1$ then

$$\langle F(x_n, y_n), z - y_n \rangle \geq -\delta_n. \tag{3.30}$$

For the cace of $\|z - y_n\| \geq 1$, since $(x_n, \lambda z + (1-\lambda)y_n) \in gph\Omega_2, \forall \lambda \in (0,1)$, we can choose λ such that

$$\|\lambda z + (1-\lambda)y_n - y_n\| = \lambda \|z - y_n\| = 1.$$

Hence by (3.30) one has

$$\langle F(x_n, y_n), \lambda z + (1-\lambda)y_n - y_n \rangle \geq -\delta_n,$$

that is,

$$\langle F(x_n, y_n), z - y_n \rangle \geq -\delta_n/\lambda = -\delta_n \|z - y_n\|. \tag{3.31}$$

Therefore, combining (3.30) and (3.31), the inequality (3.29) follows. Second, by definition of $\bar{y} = y_{\bar{x}}$ we know

$$\langle F(\bar{x}, \bar{y}), z - \bar{y} \rangle \geq 0 \quad \forall z \in \Omega_2(\bar{x}). \tag{3.32}$$

Third, from the Lipschitz continuity of Ω_2, we have for any $m > 0$, there exists $\sigma_m > 0$ such that $\Omega_2(\bar{x} + \sigma_m B) \subset \Omega_2(\bar{x}) + \frac{1}{m}B$ and hence there exists N_m such that

$$\Omega_2(x_n) \subset \Omega_2(\bar{x}) + \frac{1}{m}B \quad \forall n > N_m. \tag{3.33}$$

Therefore we can take $y_n^* \to \bar{y}$ with $y_n^* \in \Omega_2(x_n)$ (due to continuity of Ω_2) and $z_n \in \Omega_2(\bar{x}), \xi_n \in B$ such that $y_n = z_n + \frac{1}{m}\xi_n$ (from (3.33)). By taking

$z = y_n^*$ in (3.29) and $z = z_n$ in (3.32), we obtain that for $n > N_m$,

$$\langle F(x_n, y_n), y_n - \bar{y}\rangle \leq \delta_n \max\{1, \|y_n - y_n^*\|\} - \langle F(x_n, y_n), \bar{y} - y_n^*\rangle \tag{3.34}$$

$$\langle F(\bar{x}, \bar{y}), y_n - \bar{y}\rangle \geq \langle F(\bar{x}, \bar{y}), \frac{1}{m}\xi_n\rangle. \tag{3.35}$$

Combining (3.35) and (3.34) we have

$$\begin{aligned}\langle F(x_n, y_n) - F(x_n, \bar{y}), y_n - \bar{y}\rangle \leq \quad & \delta_n \max\{1, \|y_n - y_n^*\|\} - \\ & -\langle F(x_n, \bar{y}) - F(\bar{x}, \bar{y}), y_n - \bar{y}\rangle - \\ & -\langle F(x_n, y_n), \bar{y} - y_n^*\rangle - \\ & -\langle F(\bar{x}, \bar{y}), \frac{1}{m}\xi_n\rangle. \end{aligned} \tag{3.36}$$

What we want to show is that $\{y_n\}$ converges $\bar{y}$. In fact, if it is not true, then there exists $\delta > 0$ such that $\|y_{n_k} - \bar{y}\| > \delta$ for all k, where $\{y_{n_k}\}$ is a subsequence of $\{y_n\}$. By $y_{n_k}^* \to \bar{y}$ we know for large enough n_k, $\|y_{n_k}^* - \bar{y}\| \leq \delta$ and hence $\|y_{n_k}^* - y_{n_k}\| \leq 2\|y_{n_k} - \bar{y}\|$. Remembering the assumption (H2) we have

$$\begin{aligned} C_R\|y_{n_k} - \bar{y}\|^2 \quad & \leq \langle F(x_{n_k}, y_{n_k}) - F(x_{n_k}, \bar{y}), y_{n_k} - \bar{y}\rangle \\ & \leq \delta_{n_k} \max\{1, 2\|y_{n_k} - \bar{y}\|\} - \\ & -\langle F(x_{n_k}, \bar{y}) - F(\bar{x}, \bar{y}), y_{n_k} - \bar{y}\rangle - \\ & -\langle F(x_{n_k}, y_{n_k}), \bar{y} - y_{n_k}^*\rangle - \\ & -\frac{1}{m}\langle F(\bar{x}, \bar{y}), \xi\rangle. \end{aligned}$$

Therefore

$$C_R\|y_{n_k} - \bar{y}\| \leq \quad \delta_{n_k} \max\{1/\delta, 2\} - \langle F(x_{n_k}, \bar{y}) - F(\bar{x}, \bar{y}), \frac{y_{n_k} - \bar{y}}{\|y_{n_k} - \bar{y}\|}\rangle -$$

$$\begin{aligned}
& \frac{1}{\|y_{n_k}-\bar{y}\|}\langle F(x_{n_k},y_{n_k}),\bar{y}-y^*_{n_k}\rangle - \frac{1}{m\|y_{n_k}-\bar{y}\|}\langle F(\bar{x},\bar{y}),\xi\rangle \\
\leq\ & \delta_{n_k}\max\{1/\delta,2\} + \|F(\bar{x},\bar{y})-F(x_{n_k},\bar{y})\| + \\
& +\frac{1}{\delta}\|F(x_{n_k},y_{n_k})\|\|y^*_{n_k}-\bar{y}\| + \frac{1}{\delta m}\|F(\bar{x},\bar{y})\|. \qquad (3.37)
\end{aligned}$$

Apparently, in the inequality (3.37), the value of right hand side converges to 0 as $k \to \infty$ which contradicts the assumption of $\|y_{n_k}-\bar{y}\| > \delta$. By the penalty method, it is easy to see that $(\bar{x},\bar{y})$ is a solution of the problem (P).

Step 2. Now we turn to consider $\{p_n\}$. First, by $N_n\|-F(x_n,y_n)\|^{\Omega_2(x_n)}_{y_n}$ being bounded and Lemma 3.6 we have

$$-V_n'((0,0),(0,p_n)) \leq O(\|p_n\|). \qquad (3.38)$$

Now taking $s = p_n/N_n \in \Omega_2(x_n) - y_n$ in (3.23) one has

$$f^0_y((x_n,y_n);p_n) - V_n'((0,0),(0,p_n)) \geq \langle \nabla_y F(x_n,y_n)^\top p_n, p_n\rangle - \epsilon_n\|p_n\|. (3.39)$$

Remembering that

$$\langle F(x,y+tp)-F(x,y),tp\rangle \geq C_R t^2\|p\|^2$$

we know $\langle \nabla_y F(x,y)^\top p,p\rangle C_R t^2\|p\|^2$ and hence from (3.39) we obtain

$$\begin{aligned}
L_f\|p_n\| + O(\|p_n\|) + \epsilon_n\|p_n\| \ &\geq\ \langle \nabla_y F(x_n,y_n)^\top p_n, p_n\rangle \\
&\geq\ C_R\|p_n\|^2
\end{aligned}$$

which implies that $\{p_n\}$ is bounded. Assume $p_n \to \bar{p}$, by $\delta_n \to 0$, it is easy to see $(\bar{x},\bar{y},\bar{p})$ satisfies (3.19).

Step 3. Our tasks remained now is to show that $(\bar{x},\bar{y},\bar{p})$ satisfies the condition (3.17) and (3.18). First we prove the condition (3.17) holds. Using

Lemma 3.8 and the inequality (3.23) we have

$$\nabla_y F(x_n, y_n)^\top p_n \in \partial_y f(x_n, y_n) + N(y_n, \Omega_2(x_n)) - \partial_y V_n(0,0) + \epsilon_n B, \quad (3.40)$$

Denote

$$S_{1n}(y) = S_n(0, y) = N_n[(\Omega_2(x_n) - y_n - y) \cap \bar{B}]$$

and $V_{1n}(y) = \inf_{p \in S_{1n}(y)} U_n(p) = V_n(0, y)$. We have $V_n'((0,0);(0,s)) = V_{1n}'(0;s)$ which means that $\partial_y V_n(0,0) = \partial V_{1n}(0)$.

By Lemma 3.9, we have

$$\partial V_{1n}(0) = D^* S_{1n}(0, p_n) \partial U_n(p_n) = D^* S_{1n}(0, p_n)(F(x_n, y_n)).$$

(3.40) becomes:

$$\begin{aligned} \nabla_y F(x_n, y_n)^\top p_n \;\; \in \;\; & \partial_y f(x_n, y_n) + N(y_n, \Omega_2(x_n)) \\ & - D^* S_{1n}(0, p_n)(F(x_n, y_n)) + \epsilon_n B. \qquad (3.41) \end{aligned}$$

Since $\partial_y f(x_n, y_n)$ is compact, by extracting a subsequence , we can assume that there are two sequences $u_n^* \in N(y_n, \Omega_2(x_n))$ and

$$v_n^* \in D^* S_{1n}(0, p_n)(F(x_n, y_n))$$

such that

$$\lim(u_n^* - v_n^*) = w^*.$$

Our aim is to show that $w^* \in -D_p \Phi_1((\bar{x}, \bar{y}), -F(\bar{x}, \bar{y}))(0, \bar{p})$. First we show that for n large enough

$$v_n^* \in N_n N(y_n + (1/N_n) p_n, \Omega_2(x_n)) \quad (3.42)$$

In fact, from the definition we have

$$(v_n^*, -F(x_n, y_n)) \in N((0, p_n), gphS_{1n}).$$

Let $\hat{S}_{1n}(y) = N_n(\Omega_2(x_n) - y_n - y)$ and $\bar{B}_n(y) = N_n\bar{B}$. Then $S_{1n}(y) = \hat{S}_{1n}(y) \cap \bar{B}_n(y)$. By 2.28 we obtain

$$(v_n^*, -F(x_n, y_n)) \in N((0, p_n), gph\hat{S}_{1n}) + N((0, p_n), gph\bar{B}_n).$$

But $\{p_n\}$ is bounded and so $(0, p_n) \in Int(gph\bar{B}_n)$, the interior of $gph\bar{B}_n$ for n large enough which implies $N((0, p_n), gph\bar{B}_n) = \{0\}$. Therefore for large enough n,

$$(v_n^*, -F(x_n, y_n)) \in N((0, p_n), gph\hat{S}_{1n}),$$

that is, for any $v \in R^n$, $\forall w \in \hat{S}_{1n}(v)$, we have

$$\langle v_n^*, v\rangle + \langle -F(x_n, y_n), w - p_n\rangle \leq 0. \tag{3.43}$$

Therefore $\langle v_n^*, v\rangle \leq \langle F(x_n, y_n), w - p_n\rangle, \forall (v, w) \in gph\hat{S}_{1n}$. Especially, we take $w = p_n - N_n v$ then the above inequality become

$$\begin{aligned} \langle v_n^*, v\rangle &\leq \langle F(x_n, y_n), -N_n v\rangle \\ &= \langle -N_n F(x_n, y_n), v\rangle \ \forall v \in R^n. \end{aligned} \tag{3.44}$$

It follows that

$$v_n^* = -N_n F(x_n, y_n). \tag{3.45}$$

Now let $v = 0$ in (3.43) we know that

$$\langle -F(x_n, y_n), w - p_n\rangle \leq 0 \ \forall w \in \hat{S}_{1n}(0). \tag{3.46}$$

Since $\hat{S}_{1n}(0) = N_n(\Omega(x_n) - y_n)$, for any $z \in \Omega_2(x_n)$ there exists a $w \in \hat{S}_{1n}(0)$ such that $(w - p_n) = N_n(z - (y_n + (1/N_n)p_n))$ and hence

$$\langle -F(x_n, y_n), (1/N_n)(w - p_n)\rangle = \langle -F(x_n, y_n), z - (y_n + (1/N_n)p_n)\rangle \leq 0 \quad \forall z \in \Omega_2(x_n). \tag{3.47}$$

Therefore, combining (3.45) and (3.47), (3.42) follows.

Note that $\lim(u_n^*/N_n) = \lim(v_n^*/N_n) = -F(\bar{x}, \bar{y})$. By (3.47) and the Lipschitz continuity of Ω_2, it is easy to show that $-F(\bar{x}, \bar{y}) \in N(\bar{y}, \Omega_2(\bar{x}))$. Moreover, by $u_n^*/N_n \in N(y_n, \Omega_2(x_n))$ and the assumption we have

$$\lim[\frac{v_n^*}{N_n} - \frac{u_n^*}{N_n}]N_n = -w^*.$$

Using Proposition 2.34 we have

$$-w^* \in D_p\Phi_1((\bar{x}, \bar{y}), -F(\bar{x}, \bar{y}))(0, \bar{p}).$$

Therefore condition (3.17) holds.

We now turn back to prove condition (3.18). Similarly we can get

$$\begin{aligned} \nabla_x F(x_n, y_n)^\top p_n \in \quad & \partial_x f(x_n, y_n) + N(x_n, \Omega_2^{-1}(y_n)) \\ & -D^* S_{2n}(0, p_n)(F(x_n, y_n)) + \epsilon_n B \end{aligned} \tag{3.48}$$

where

$$S_{2n}(x) := S_n(x, 0) = N_n[(\Omega_2(x_n + x) - y_n) \cap \bar{B}].$$

and $V_{2n}(x) := \inf_{p \in S_{2n}(x)} U_n(p) = V_n(x, 0)$. Suppose $u_n^* \in N(x_n, \Omega_2^{-1}(y_n))$, $v_n^* \in D^* S_{2n}(0, p_n)(F(x_n, y_n))$ with

$$\lim(u_n^* - v_n^*) = w^*.$$

The idea is to show

$$-w^* \in \bigcup_{y^* \in \Phi_2(\bar{x},\bar{y})} D_p\Phi_2((\bar{x},\bar{y}),y^*)(0,\bar{p}).$$

Note that we have for n large enough,

$$v_n^* \in N(x_n, \Omega_2^{-1}(y_n + (1/N_n)p_n)). \tag{3.49}$$

In fact, similar to the proof of (3.42), we have, for n large enough,

$$(v_n^*, -F(x_n,y_n)) \in N((0,p_n), gph\hat{S}_{2n}),$$

where $\hat{S}_{2n}(x) = N_n[\Omega_2(x_n + x) - y_n]$. Note that

$$\hat{S}_{2n}^{-1}(y) = \Omega_2^{-1}(y_n + (1/N_n)y) - x_n.$$

Using Lemma 3.7, it follows that $(-F(x_n,y_n), v_n^*) \in N((0,p_n), gph\hat{S}_{2n}^{-1})$. Therefore for any $(v,w) \in gph\hat{S}_{2n}^{-1}$ one has

$$\langle -F(x_n,y_n), v - p_n\rangle + \langle v_n^*, w\rangle \leq 0,$$

that is, $\langle v_n^*, w\rangle \leq \langle F(x_n,y_n), v - p_n\rangle$ for any $v \in R^n$, and

$$w \in \Omega_2^{-1}(y_n + (1/N_n)v) - x_n, \ \forall v \in R^n.$$

In particular, let $v = p_n$ we have

$$\langle v_n^*, w\rangle \leq 0, \ \forall w \in \Omega_2^{-1}(y_n + (1/N_n)p_n) - x_n. \tag{3.50}$$

and (3.49) holds.

Therefore by

$$\lim \frac{[(v_n^*/N_n) - (u_n^*/N_n)]}{1/N_n} = -w^*$$

where

$$u_n^*/N_n \in N(x_n, \Omega_2^{-1}(y_n)), v_n^*/N_n \in N(x_n, \Omega_2^{-1}(y_n + (1/N_n)p_n)),$$

condition (3.18) follows. The proof is complete. □

Until now, by the Lemma 3.11 and Lemma 3.12 we have learnt that the theorem is true under the singleton assumption of $M_n(\cdot,\cdot)$. For the general case, we consider

$$L_n(x,y) = N_n \| - F(x,y)\|_y^{\Omega_2(x)}.$$

By [2, Proposition3.2.24] and (3.21) we know $L_n(x,y)$ is locally Lipschitz and

$$f^0((x_n,y_n);(w,s)) + L_n^0((x_n,y_n);(w,s)) \geq -\epsilon_n(\|s\|^2 + \|w\|^2)^{1/2} \quad (3.51)$$

for any $(w,s) \in gph\Omega_2 - (x_n, y_n)$. Therefore there exists

$$q_n^{ws*} \in \partial L_n(x_n, y_n)$$

such that $\langle q_n^{ws*}, (w,s)\rangle = L_n^0((x_n,y_n);(w,s))$. Hence we have

$$f^0((x_n,y_n);(w,s)) + \langle q_n^{ws*}, (w,s))\rangle \geq -\epsilon_n(\|s\|^2 + \|w\|^2)^{1/2}. \quad (3.52)$$

Now consider the function $H : [gph\Omega_2 - (x_n,y_n)] \times \partial L_n(x_n,y_n) \to R$ defined by

$$H(w,s;q) = f^0((x_n,y_n);(w,s)) + \langle q, (w,s))\rangle + \epsilon_n(\|s\|^2 + \|w\|^2)^{1/2}$$

Using the Lopsided Minimax theorem [2, p.319], we can conclude that there exists $q_n \in \partial L_n(x_n,y_n)$ such that

$$\inf_{(w,s)} H(w,s;q_n) = \inf_{(w,s)} \max_{q \in \partial L_n(x_n,y_n)} H(w,s;q) \geq 0$$

Therefore, for any $(w, s) \in gph\Omega_2 - (x_n, y_n)$, one has

$$f^0((x_n, y_n); (w, s)) + \langle q_n, (w, s)) \rangle \geq -\epsilon_n(\|s\|^2 + \|w\|^2)^{1/2}. \qquad (3.53)$$

Our aim now is to estimate the generalized gradient of $L_n(x, y)$. We have the following lemma.

Lemma 3.13

$$\partial L_n(x_n, y_n) \subset \bigcup_{p \in M_n(x_n, y_n)} (-\nabla_x F(x_n, y_n)^\top p, -\nabla_y F(x_n, y_n)^\top p)$$
$$-(\partial_x V_n(0,0), \partial_y V_n(0,0)). \qquad (3.54)$$

Proof. To show the estimate (3.54), we choose $t_k \downarrow 0$, $z_k (\in R^n) \to 0$ and $w_k (\in R^m) \to 0$ such that

$$L_n^0((x_n, y_n); (w, s)) = \lim_{k \to \infty} [L_n(x_n + z_k + t_k w, y_n + w_k + t_k s) - L_n(x_n + z_k, y_n + w_k)]/t_k.$$

Since

$$\begin{aligned}
& L_n(x_n + z_k + t_k w, y_n + w_k + t_k s) - L_n(x_n + z_k, y_n + w_k) \\
= \quad & N_n \| - F(x_n + z_k + t_k w, y_n + w_k + t_k s)\|_{y_n + w_k + t_k s}^{\Omega_2(x_n + z_k + t_k w)} \\
& -N_n \| - F(x_n + z_k, y_n + w_k)\|_{y_n + w_k}^{\Omega_2(x_n + z_k)} \\
= \quad & \langle -F(x_n + z_k + t_k w, y_n + w_k + t_k s), p_{nk}^{ws} \rangle \\
& -\langle -F(x_n + z_k, y_n + w_k), p_{nk}^{ws} \rangle \\
& +\langle -F(x_n + z_k, y_n + w_k), p_{nk}^{ws} \rangle + \bar{V}_n(z_k, 0; w_k, 0) \\
\leq \quad & -\langle F(x_n + z_k + t_k w, y_n + w_k + t_k s) - F(x_n + z_k, y_n + w_k), p_{nk}^{ws} \rangle \\
& -\bar{V}_n(z_k, t_k w; w_k, t_k s) + \bar{V}_n(z_k, 0; w_k, 0) \qquad (3.55)
\end{aligned}$$

where $p_{nk}^{ws} \in M_n(x_n + z_k + t_k w, y_n + w_k + t_k s)$ and

$$S_n(x, y; u, v) = N_n[(\Omega_2(x_n + x + y) - y_n - u - v) \cap \bar{B}]$$
$$\bar{V}_n(x, y; u, v) = \inf_{p \in S_n(x,y;u,v)} \langle F(x_n + x, y_n + y), p \rangle.$$

Therefore we have

$$\begin{aligned} & L_n(x_n + z_k + t_k w, y_n + w_k + t_k s) - L_n(x_n + z_k, y_n + w_k) \\ \leq \quad & -\langle F(x_n + z_k + t_k w, y_n + w_k + t_k s) - F(x_n + z_k, y_n + w_k), p_{nk}^{ws} \rangle \\ & -\{\bar{V}_n(z_k, t_k w; w_k, t_k s) - \bar{V}_n(z_k, 0; w_k, 0)\}. \end{aligned} \tag{3.56}$$

Now note that

$$\lim_{k \to \infty} [\langle F(x_n + z_k + t_k w, y_n + w_k + t_k s) - F(x_n + z_k, y_n + w_k), p_{nk}^{ws} \rangle] / t_k$$
$$= \langle \nabla F(x_n, y_n)^T (w, s), p_{nk}^{ws} \rangle$$

where $p_{nk}^{ws} \in M_n(x_n, y_n)$. On the other hand, by the [2, Proposition3.2.24], we know $\bar{V}_n(x, y; u, v)$ is Lipschitz and hence using the Mean-value theorem (Proposition 2.21) we conclude that there exist $\theta_k \in (0, 1)$ and

$$r_{nk} \in \partial_{y,v} \bar{V}_n(z_k, \theta_k t_k w; w_k, \theta_k t_k s)$$

such that

$$\bar{V}_n(z_k, t_k w; w_k, t_k s) - \bar{V}_n(z_k, 0; w_k, 0) = \langle r_{nk}, t_k(w, s) \rangle.$$

By [9, Corollary 1,P87] we can assume that

$$\lim_{k \to \infty} r_{nk} = r_n \in \partial_{y,v} \bar{V}(0, 0; 0, 0) = \partial V_n(0, 0)$$

and hence

$$\lim_{k\to\infty}[\bar{V}_n(z_k, t_k w; w_k, t_k s) - \bar{V}_n(z_k, 0; w_k, 0)]/t_k = \langle r_n, (w,s)\rangle.$$

By $-r_n \in \partial(-V_n)(0,0)$ we obtain that

$$\begin{aligned}\langle r_n, (w,s)\rangle &\leq (-V_n)'((0,0);(w,s)) \\ &\leq -V_n'((0,0);(w,s)).\end{aligned}$$

Therefore from (3.56) we have that

$$L_n^0((x_n,y_n);(w,s)) \leq \langle -\nabla F(x_n,y_n)^T(w,s), p_{nk}^{ws}\rangle - V_n'((0,0);(w,s)).$$

Hence

$$L_n^0((x_n,y_n);(w,s)) \leq \sup_{p\in M_n(x_n,y_n)}\{\langle -\nabla F(x_n,y_n)^T(w,s),p\rangle\} - V_n'((0,0);(w,s))$$

from which (3.54) follows. The proof is complete. □

We are now ready to complete the proof of the theorem. In fact, by Lemma 3.13, for any $q_n \in \partial L_n(x_n,y_n)$ there exists a $p_n \in M_n(x_n,y_n)$ such that

$$q_n \in (-\nabla_x F(x_n,y_n)^\top p_n, -\nabla_y F(x_n,y_n)^\top p_n) - (\partial_x V_n(0,0), \partial_y V_n(0,0))$$

Combining the above inclusion with (3.53), it deduces that (3.23) and (3.24) hold again. And hence by Lemma 3.12 the assertion follow. □

So far we have known that for problem (OPVIC) with assumptions (H1),(H2), there must exists a solution $(\bar{x},\bar{y})$ which satisfies conditions (3.17)-(3.19). But what we really interest is whether conditions (3.17)-(3.19) are a necessary optimality condition for any other solutions of problem (OPVIC). The positive answer is provided in the following theorem.

Theorem 3.14 *Suppose that problem* (OPVIC) *satisfies the same conditions as that of theorem* 3.10 *and* $(\bar{x}, \bar{y})$ *is any solution of* (OPVIC), *then there exists* $\bar{p} \in R^m$ *such that conditions* (3.17)-(3.19) *hold at* $(\bar{x}, \bar{y})$.

Proof. Let $f^*(x, y) = f(x, y) + \|x - \bar{x}\|$. We consider the following problem:

$$\begin{array}{lll} (P^*) & \min & f^*(x,y) \\ & \text{s.t.} & x \in R^n, \; y \in \Omega_2(x) \\ & & \| - F(x,y)\|_y^{\Omega_2(x)} = 0. \end{array}$$

Apparently, $f^*(x, y)$ satisfies the condition $(H1)$ and hence there is a solution of (P^*), say, (x^*, y^*) which satisfies (3.17)-(3.19) where f is replaced by f^*. Now we need to show $(x^*, y^*) = (\bar{x}, \bar{y})$. In fact, noting that both problem (OPVIC) and (P^*) have the same feasible solution, we have

$$f(\bar{x}, \bar{y}) \le f(x^*, y^*) \le f^*(x^*, y^*) \le f^*(\bar{x}, \bar{y}) \le f(\bar{x}, \bar{y}).$$

Therefore $x^* = \bar{x}$ which implies $y^* = \bar{y}$. On the other hand, note that $\partial_x f^*(\bar{x}, \bar{y}) = \partial_x f(\bar{x}, \bar{y})$, $\partial_y f^*(\bar{x}, \bar{y}) = \partial_y f(\bar{x}, \bar{y})$ and hence the assertion follows.
□

By the same way we can prove the following general theorem where Ω_1 may not be R^n.

Theorem 3.15 *Suppose that (H1), (H2) hold and* $(\bar{x}, \bar{y})$ *is a solution of (OPVIC). Then there exists* $\bar{p} \in R^m$ *such that*

$$\begin{array}{lll} \nabla_y F(\bar{x}, \bar{y})^\top \bar{p} & \in & \partial_y f(\bar{x}, \bar{y}) - D_p \Phi_1((\bar{x}, \bar{y}), -F(\bar{x}, \bar{y}))(0, \bar{p}) \\ \nabla_x F(\bar{x}, \bar{y})^\top \bar{p} & \in & \partial_x f(\bar{x}, \bar{y}) - \bigcup_{y^* \in \Phi_2(\bar{x}, \bar{y})} D_p \Phi_2((\bar{x}, \bar{y}), y^*)(0, \bar{p}) \end{array}$$

$$
\begin{aligned}
&+ \quad N(\bar{x}, \Omega_1) \\
\langle F(\bar{x}, \bar{y}), \bar{p} \rangle &= \quad 0.
\end{aligned}
$$

Chapter 4

Necessary Optimality Conditions Involving Coderivatives

We focus in this chapter on necessary optimality conditions involving coderivatives for the optimization problem with variational inequality constraints (OPVIC). The following is the outline of the chapter.

§4.1. Preliminaries. This section is devoted to reviewing some concepts and results which are used in the main body of the chapter.

§4.2. Optimality conditions. This is the main body of the chapter. Some necessary optimality conditions are derived. An example is given.

4.1 Preliminaries

This section contains some background material which will be used in the next section.

Definition 4.1 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ with a closed graph is said to be upper-Lipschitz continuous at $\bar{x} \in R^n$ if there exist a neighborhood U of $\bar{x}$ and a constant $L_\Phi \geq 0$ such that*

$$\Phi(x) \subset \Phi(\bar{x}) + L_\Phi \|x - \bar{x}\| \bar{B}, \ \forall x \in U \tag{4.1}$$

The constant L_Φ is called the modulus of Φ at $\bar{x}$.

Definition 4.2 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ with a closed graph is said to be pseudo-Lipschitz continuous around $(\bar{x}, \bar{y}) \in gph\Phi$ if there exist a neighborhood U of $\bar{x}$, a neighborhood V of $\bar{y}$, and a constant $L_\Phi \geq 0$ such that*

$$\Phi(x') \cap V \subset \Phi(x) + L_\Phi \|x' - x\| \bar{B}, \ \forall x', x \in U. \tag{4.2}$$

Remark 4.3 (1) *In general, pseudo-Lipschitz continuity of Φ around $(\bar{x}, y)$ for all $y \in \Phi(\bar{x})$ does not imply upper-Lipschitz continuity of Φ at $\bar{x}$ and vice versa.*

(2) *If $\Phi : R^n \rightrightarrows R^m$ is Lipschitz continuous around $\bar{x}$, then Φ is upper-Lipschitz continuous at $\bar{x}$, pseudo-Lipschitz continuous around $(\bar{x}, y)$ for all $y \in \Phi(\bar{x})$.*

Definition 4.4 *A set-valued map $\Phi : R^n \rightrightarrows R^m$ with a closed graph is said to be locally upper-Lipschitz continuous at $(\bar{x}, \bar{y}) \in gph\Phi$ if there exist a neighborhood U of $\bar{x}$, a neighborhood V of $\bar{y}$, and a constant $L_\Phi \geq 0$ such that*

$$\Phi(x) \cap V \subset \Phi(\bar{x}) + L_\Phi \|x - \bar{x}\| \bar{B}, \quad \forall x \in U. \tag{4.3}$$

Directly from the definition we have the following lemma

Lemma 4.5 (1) *If Φ is upper-Lipschitz continuous at $\bar{x}$ then Φ is locally upper Lipschitz continuous at $(\bar{x}, y)$ for all $y \in \Phi(\bar{x})$.*

(2) *If Φ is pseudo-Lipschitz continuous around $(\bar{x}, \bar{y}) \in gph\Phi$, then Φ is locally upper-Lipschitz continuous at $(\bar{x}, \bar{y})$.*

The following results were given by Mordukhovich in [30, Proposition 3.5].

Proposition 4.6 *Let $\Phi : R^n \rightrightarrows R^m$ be a set-valued map with a closed graph. Then Φ is pseudo-Lipschitz continuous around $(\bar{x}, \bar{y})$ if and only if*

$$D^*\Phi(\bar{x}, \bar{y})(0) = \{0\}. \tag{4.4}$$

Corollary 4.7 *Let Φ be a set-valued map with a closed graph and $(\bar{x}, \bar{y}) \in gph\Phi$. Then $\Phi^{-1} : R^m \rightrightarrows R^n$ is pseudo-Lipschitz continuous around $(\bar{y}, \bar{x}) \in gph\Phi^{-1}$ if and only if*

$$Ker(D^*\Phi(\bar{x}, \bar{y})) = \{0\}, \tag{4.5}$$

where $Ker(Q)$ is the kernal of a set-valued map $Q : R^n \rightrightarrows R^m$ defined by

$$Ker(Q) = \{x \in R^n : 0 \in Q(x)\}.$$

4.2 Optimality conditions

The purpose of this section is to explore the necessary optimality condition involving coderivatives for the optimization problem with variational inequlity constraints (OPVIC). First, we consider the following optimization problem with a generalized equation constraint:

$$\begin{array}{lll} (P_0) & \min & f(z) \\ & \text{s.t.} & z \in \Phi(0), \end{array} \tag{4.6}$$

where $f : R^{n+m} \to R$ is Lipschitz continuous with modulus L_f, $\Phi : R^q \rightrightarrows R^{n+m}$ is a set-valued map with a closed graph. It is obvious that (P_0) can be rewritten as the following problem:

$$\begin{array}{lll} (P_0') & \min & f(z) \\ & \text{s.t.} & z \in \Phi(v) \\ & & v = 0. \end{array} \tag{4.7}$$

In the following Lemma, we show that the problem (P_0') is actually equivalent to its penalized problem (P_τ) when the set-valued map in the generalized equation constriant of (P_0) is locally upper Lipschitz continuous.

Lemma 4.8 *Assume that $\bar{z}$ solves (P_0) and Φ is locally upper-Lipschitz continuous with modulus $L_\Phi \geq 0$ around $(0, \bar{z}) \in gph\Phi$, then there exist a neighborhood U of 0, a neighborhood V of $\bar{z}$ such that $(0, \bar{z})$ solves the following penalized problem of (P_0'):*

$$\begin{array}{lll} (P_r) & \min & f(z) + r\|v\| \\ & s.t. & z \in \Phi(v) \cap V \\ & & v \in U \end{array}$$

where $r > L_f L_\Phi$.

Proof. Since $\bar{z}$ solves (P_0), we have that

$$f(\bar{z}) = f(\bar{z}) + r\|0\| \leq f(z) \quad \forall z \in \Phi(0).$$

By virtue of locally upper-Lipschitz continuity of Φ around $(0, \bar{z})$, there exist U neighborhoods of 0 and V neighborhoods of $\bar{z}$ such that (4.3) is satisfied. Therefore, for any $v \in V$ and $z \in \Phi(v) \cap V$ there exists $z^* \in \Phi(0)$ such that

$$\|z - z^*\| \leq L_\Phi \|v\|.$$

Thus we have for any $z \in \Phi(v) \cap V, \quad v \in U$

$$\begin{aligned} f(\bar{z}) \leq f(z^*) &= f(z) + (f(z^*) - f(z)) \leq f(z) + L_f\|z^* - z\| \\ &\leq f(z) + L_\Phi L_f \|v\| < f(z) + r\|v\|. \end{aligned}$$

The proof is complete. □

We now give a Kuhn-Tucker type optimality condition for problem (P_0).

Lemma 4.9 *Let $\bar{z}$ be a solution of (P_0) and Φ be locally upper-Lipschitz continuous around $(0, \bar{z})$. Then for any $r > L_f L_\Phi$ there exists $\eta \in r\bar{B}_q$ such that*

$$0 \in \partial f(\bar{z}) + D^*\Phi^{-1}(\bar{z}, 0)(\eta), \tag{4.8}$$

where $\bar{B}_q$ denotes the closed unit ball in R^q.

Proof. By lemma 4.8 we know that $(0, \bar{z})$ is a solution of (P_r). Rewrite (P_r) in the following form:

$$\begin{array}{lll} (\bar{P}_r) & \min & f(z) + r\|v\| \\ & \text{s.t.} & (v, z) \in gph\Phi \cap (U \times V). \end{array}$$

By Corollary 2.35 we have

$$\begin{aligned} (0,0) &\in (r\bar{B}_q) \times \partial f(\bar{z}) + N((0,\bar{z}), gph\Phi \cap (U \times V)) \\ &= (r\bar{B}_q) \times \partial f(\bar{z}) + N((0,\bar{z}), gph\Phi), \end{aligned}$$

since U and V are neighborhood of 0, and $\bar{z}$ respectively. That is, there exist $(p_1, q_1) \in (r\bar{B}_q) \times \partial f(\bar{z})$, $(p_2, q_2) \in N((0, \bar{z}), gph\Phi)$ such that

$$p_1 + p_2 = 0, \quad q_1 + q_2 = 0. \tag{4.9}$$

By the definition of coderivatives, $(p_2, q_2) \in N((0, \bar{z}), gph\Phi)$ implies that $p_2 \in D^*\Phi(0, \bar{z})(-q_2)$. By Lemma 3.7, that is $q_2 \in D^*\Phi^{-1}(\bar{z}, 0)(-p_2)$. Therefore by (4.9)

$$q_2 \in D^*\Phi^{-1}(\bar{z}, 0)(p_1).$$

And hence there exists $p_1 \in r\bar{B}_q$ such that

$$0 = q_1 + q_2 \in \partial f(\bar{z}) + D^*\Phi^{-1}(\bar{z}, 0)(p_1).$$

The proof is complete. □

Now we consider the following problem (P).

$$\begin{array}{lll} (P) & \min & f(z) \\ & \text{s.t.} & 0 \in -h(z) + Q(z) \\ & & z \in \Omega, \end{array} \tag{4.10}$$

where $z \in R^{n+m}$, Ω is a closed subset of R^{n+m}, $f : R^{n+m} \to R$, $h : R^{n+m} \to R^q$ are Lipschitz continuous and continuously differentiable respectively and $Q : R^{n+m} \rightrightarrows R^q$ is a set-valued map with a closed graph.

Theorem 4.10 *Suppose $\bar{z}$ solves problem (P). Further we assume*

(1) *f is locally Lipschitz continuous around $\bar{z}$.*

(2) *h is continuously differentiable around $\bar{z}$.*

(3) $D^*Q(\bar{z}, h(\bar{z}))(0) \cap \{-N(\bar{z}, \Omega)\} = \{0\}$.

(4) $\nabla h(\bar{z})^\top p \in D^*Q(\bar{z}, h(\bar{z}))(p) + N(\bar{z}, \Omega)$ *implies* $p = 0$.

Then there exist $r > 0, \eta \in r\bar{B}_q$ such that

$$\nabla h(\bar{z})^\top \eta \in \partial f(\bar{z}) + D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega). \tag{4.11}$$

Proof. Denote $G(z) = -h(z) + Q(z)$ and $\Phi(v) = G^{-1}(v) \cap \Omega$. Then problem (P) can be represented as

$$\begin{aligned} \min \quad & f(z) \\ \text{s.t.} \quad & z \in \Phi(0). \end{aligned}$$

First we show that

$$N((0,\bar{z}), gphG^{-1}) \cap (-N((0,\bar{z}), R^n \times \Omega)) = \{0\}. \tag{4.12}$$

Indeed, for any $(0,q) \in N((0,\bar{z}), gphG^{-1})$ where $q \in -N(\bar{z},\Omega)$, by lemma 3.7, one has $(q,0) \in N((\bar{z},0), gphG)$ and hence from Corollary 2.25 we know

$$q \in D^*G(\bar{z},0)(0) = -\nabla h(\bar{z})^\top 0 + D^*Q(\bar{z},h(z))(0) = D^*Q(\bar{z},h(z))(0).$$

The assertion follows from the assumption (3).

We now prove that $\Phi(v)$ is pseudo-Lipschitz continuous around $(0,\bar{z})$. By Proposition 4.6 it is sufficient to show $D^*\Phi(0,\bar{z})(0) = \{0\}$. Suppose that $p \in D^*\Phi(0,\bar{z})(0)$, which means by the definition of coderivatives that $(p,0) \in N((0,\bar{z}), gph\Phi)$. It is easy to see that $gph\Phi = gphG^{-1} \cap (R^n \times \Omega)$. Since (4.12) holds, we can apply Corollary 2.27 and obtain

$$\begin{aligned} (p,0) &\in N((0,\bar{z}), gphG^{-1}) + N((0,\bar{z}), R^n \times \Omega) \\ &= N((0,\bar{z}), gphG^{-1}) + \{0\} \times N(\bar{z},\Omega). \end{aligned}$$

That is, there exist

$$(p_1,q_1) \in N((0,\bar{z}), gphG^{-1}), \quad q_2 \in N(\bar{z},\Omega). \tag{4.13}$$

such that

$$p = p_1, \quad q_1 + q_2 = 0. \tag{4.14}$$

By virtue of lemma 3.7,

$$(p_1, q_1) \in N((0, \bar{z}), gphG^{-1})$$

implies $(q_1, p_1) \in N((\bar{z}, 0), gphG)$. By the definition of coderivatives, the last inclusion implies that $q_1 \in D^*G(\bar{z}, 0)(-p_1)$, which implies by (4.14) and Corollary 2.25 that

$$q_1 \in D^*G(\bar{z}, 0)(-p_1) = \nabla h(\bar{z})^{\mathsf{T}} p + D^*Q(\bar{z}, h(\bar{z}))(-p).$$

Since $-q_1 = q_2 \in N(\bar{z}, \Omega)$ by (4.13) and (4.14) the above inclusion becomes

$$\nabla h(\bar{z})^{\mathsf{T}}(-p) \in D^*Q(\bar{z}, h(\bar{z}))(-p) + N(\bar{z}, \Omega)$$

By assumption (4) we deduce from above inclusion that p=0. That is $D^*\Phi(0, \bar{z})(0) = \{0\}$. Hence Φ is pseudo-Lipschitz continuous around $(0, \bar{z})$ by Proposition 4.6.

By lemma 4.5, pseudo-Lipschitz continuity implies locally upper-Lipschitz continuity. Therefore, by lemma 4.9, there exists $r > 0$, $\eta \in r\bar{B}_q$ such that

$$0 \in \partial f(\bar{z}) + D^*\Phi^{-1}(\bar{z}, 0)(\eta). \tag{4.15}$$

We now try to express $D^*\Phi^{-1}(\bar{z}, 0)(\eta)$ in terms of coderivatives of h and Q. For this purpose let $p \in D^*\Phi^{-1}(\bar{z}, 0)(\eta)$. By definition, $(p, -\eta) \in N((\bar{z}, 0), gph\Phi^{-1})$. By lemma 3.7, we have $(-\eta, p) \in N((0, \bar{z}), gph\Phi)$. Hence from (4.12) and Corollary 2.27 we obtain

$$(-\eta, p) \in N((0, \bar{z}), gph\Phi) \subset N((0, \bar{z}), gphG^{-1}) + \{0\} \times N(\bar{z}, \Omega).$$

Thus there exist $(p_1, q_1) \in N((0,\bar{z}), gphG^{-1})$, $q_2 \in N(\bar{z}, \Omega)$ such that

$$p_1 = -\eta, \quad q_1 + q_2 = p. \tag{4.16}$$

Moreover, $(p_1, q_1) \in N((0,\bar{z}), gphG^{-1})$ implies that $(q_1, p_1) \in N((\bar{z}, 0), gphG)$. By lemma 3.7. That is $q_1 \in D^*G(\bar{z}, 0)(-p_1)$. By virtue of Corollary 2.25 and (4.16) we have

$$q_1 \in D^*G(\bar{z}, 0)(-p_1) = -\nabla h(\bar{z})^\top \eta + D^*Q(\bar{z}, h(\bar{z}))(\eta).$$

Hence $p = q_1 + q_2 \in -\nabla h(\bar{z})^\top \eta + D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega)$ by (4.16). That is

$$D^*\Phi^{-1}(\bar{z}, 0)(\eta) \subset -\nabla h(\bar{z})^\top \eta + D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega).$$

Therefore (4.15) becomes

$$0 \in \partial f(\bar{z}) - \nabla h(\bar{z})^\top \eta + D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega),$$

which is (4.11). The proof of the theorem is complete. □

Now we consider the optimization problems with variational inequality constraints (OPVIC):

$$\begin{aligned} \text{(OPVIC)} \quad \min \quad & f(x, y) \\ \text{s.t.} \quad & x \in \Omega_1, \ y \in \Omega_2. \\ & \langle F(x, y), z - y \rangle \geq 0 \ \forall z \in \Omega_2, \end{aligned} \tag{4.17}$$

where Ω_1, Ω_2 are closed convex subset of R^n, R^m respectively.

Since Ω_2 is a convex set, by the definition of normal cone it is easy to see that problem (OPVIC) can be rewritten as the following optimization problem with generalized equation constraints:

$$\begin{aligned} \min \quad & f(x,y) \\ \text{s.t.} \quad & (x,y) \in \Omega_1 \times R^m. \\ & 0 \in -F(x,y) + N(y,\Omega_2). \end{aligned} \tag{4.18}$$

We have following result.

Corollary 4.11 *Let* $(\bar{x},\bar{y})$ *be a solution of problem* (OPVIC). *Further suppose that*

(1) f *is Lipschitz continuous and regular at* $(\bar{x},\bar{y})$.

(2) F *is continuously differentiable around* $(\bar{x},\bar{y})$.

(3) $\nabla_x F(\bar{x},\bar{y})^\top p \in N(\bar{x},\Omega_1)$ *and* $\nabla_y F(\bar{x},\bar{y})^\top p \in D^* N((\bar{y},F(\bar{x},\bar{y})),\Omega_2)(p)$ *implies* $p = 0$.

Then there exist $r > 0$, $\eta \in r\bar{B}_m$ *such that*

$$\begin{aligned} \nabla_x F(\bar{x},\bar{y})^\top \eta &\in \partial_x f(\bar{x},\bar{y}) + N(\bar{x},\Omega_1) \\ \nabla_y F(\bar{x},\bar{y})^\top \eta &\in \partial_y f(\bar{x},\bar{y}) + D^* N((\bar{y},F(\bar{x},\bar{y})),\Omega_2)(\eta). \end{aligned} \tag{4.19}$$

Proof. Problem (4.18) is a special case of (P) where $Q(x,y) = N(y,\Omega_2)$, $h(x,y) = F(x,y), q = m$ and $\Omega = \Omega_1 \times R^m$. Denote $0_k := 0 \in R^k$ for any $k = 1,2,\cdots$. It is easy to see that $gphQ = R^n \times gphN(\cdot,\Omega_2)$. Now we want to show the condition $(1) - (4)$ in Theorem 4.10 are satisfied so that we can apply the theorem.

First, note that

$$D^*Q((\bar{x},\bar{y}),h(\bar{x},\bar{y}))(p) = \{0_n\} \times D^*N((\bar{y},h(\bar{x},\bar{y})),\Omega_2)(p). \qquad (4.20)$$

In fact, for any $w = (w_1, w_2) \in D^*Q((\bar{x},\bar{y}),h(\bar{x},\bar{y}))(p)$, we have by the definition of coderivatives that

$$\begin{aligned} &(w_1, w_2, -p) \in N(((\bar{x},\bar{y}),h(\bar{x},\bar{y})), gphQ) \\ &= N(((\bar{x},\bar{y}),h(\bar{x},\bar{y})), R^n \times gphN(\cdot,\Omega_2)) \\ &= \{0_n\} \times N((\bar{y},h(\bar{x},\bar{y})), gphN(\cdot,\Omega_2)). \end{aligned}$$

Therefore, $w_1 = 0$ and $(w_2, -p) \in N((\bar{y},h(\bar{x},\bar{y})), gphN(\cdot,\Omega_2))$ which implies that $w_2 \in D^*N((\bar{y},h(\bar{x},\bar{y})),\Omega_2)(p)$.

We now check the condition (3) in Theorem 4.10. Note that

$$\begin{aligned} &D^*Q((\bar{x},\bar{y}),h(\bar{x},\bar{y}))(0) \cap \{-N((\bar{x},\bar{y}),\Omega)\} \\ &= \{\{0_n\} \times D^*N((\bar{y},h(\bar{x},\bar{y})),\Omega_2)(0)\} \cap \{-N(\bar{x},\Omega_1) \times \{0_m\}\} \\ &= \{0_{n+m}\}, \end{aligned}$$

which implies the condition (3) in Theorem 4.10 is satisfied.

As to condition (4) in Theorem 4.10, by virtue of (4.20),

$$\nabla h(\bar{x},\bar{y})^\top p \in D^*Q((\bar{x},\bar{y}),h(\bar{x},\bar{y}))(p) + N((\bar{x},\bar{y}),\Omega)$$

implies that

$$\nabla_x F(\bar{x},\bar{y})^\top p \in N(\bar{x},\Omega_1)$$

and

$$\nabla_y F(\bar{x},\bar{y})^\top p \in D^*N((\bar{y},F(\bar{x},\bar{y})),\Omega_2)(p).$$

Thus by assumption (3) the constraint qualification (4) in Theorem 4.10 is satisfied. By Theorem 4.10, there exist $r > 0, \eta \in r\bar{B}_q$ such that

$$\begin{aligned}
\nabla F(\bar{x},\bar{y})^\top \eta &\in \partial f(\bar{x},\bar{y}) + D^*Q((\bar{x},\bar{y}), h(\bar{x},\bar{y}))(\eta) + N((\bar{x},\bar{y}),\Omega) \\
&\subset \partial_x f(\bar{x},\bar{y}) \times \partial_y f(\bar{x},\bar{y}) + \{0_n\} \times D^*N((\bar{y}, h(\bar{x},\bar{y})),\Omega_2)(\eta) \\
&\quad + N(\bar{x},\Omega_1) \times \{0_m\},
\end{aligned}$$

where the last inclusion follows from the regularity of f (c.f. Proposition 2.31) and equation (4.20). Therefore (4.19) hold. □

The proof of the following corollary is straightforward.

Corollary 4.12 *Assume that* $(\bar{x},\bar{y})$ *is a solution of problem* (OPVIC) *in which* $\Omega_1 = R^n$. f, F *satisfy the condition* (1), (2) *in* Corollary 4.11. *Moreover suppose that* $rank(\nabla_x F(\bar{x},\bar{y})) = m$. *Then there exist* $r > 0$ $\eta \in r\bar{B}_m$ *such that*

$$\begin{aligned}
\nabla_x F(\bar{x},\bar{y})^\top \eta &\in \partial_x f(\bar{x},\bar{y}) \\
\nabla_y F(\bar{x},\bar{y})^\top \eta &\in \partial_y f(\bar{x},\bar{y}) + D^*N((\bar{y}, F(\bar{x},\bar{y})),\Omega_2)(\eta). \qquad (4.21)
\end{aligned}$$

Now consider the bilevel programming problem.

$$\begin{aligned}
(BP) \quad \min \quad & f(x,y) \\
\text{s.t.} \quad & x \in R^n \\
& y \in \arg\min_{y \in \Omega(x)} g(x,y) \qquad (4.22)
\end{aligned}$$

where $\Omega(x) = \{y \in R^m : \psi(x,y) \le 0\}$, $g, f : R^{n+m} \to R$ and ψ are mappings from R^{n+m} to R^q . The Kuhn-Tucker conditions for the lower problem of

(BP),

$$\nabla_y g(x,y) + u\nabla_y \psi(x,y) = 0, \psi(x,y) \leq 0$$
$$u \geq 0, \langle u, \psi(x,y)\rangle = 0$$

where $u\nabla_y\psi = \sum u_k \nabla_y \psi_k(x,y)$, can be written as the generalized equation

$$0 \in -F(x,z) + N(z,\Omega) \tag{4.23}$$

where $\Omega = R^m \times R^q_+$ $z = (y,u,v) \in R^{m+q}$, $F : R^{n+m+q} \to R^{m+q}$ given by

$$F(x,z) = \begin{bmatrix} -[\nabla_y g + u\nabla_y \psi]^\top (x,y) \\ \psi(x,y) \end{bmatrix}. \tag{4.24}$$

Denote

$$A(x,z)^\top = \begin{bmatrix} -[\nabla^2_{xy} g + u\nabla^2_{xy} \psi](x,y) \\ \nabla_x \psi(x,y) \end{bmatrix}. \tag{4.25}$$

where $u\nabla^2_{xy}\psi = \sum_k u_k \nabla^2_{xy}\psi_k$, and

$$\nabla^2_{xy} = \begin{bmatrix} \frac{\partial^2}{\partial x_1 \partial y_1} & \cdots & \frac{\partial^2}{\partial x_n \partial y_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2}{\partial x_1 \partial y_m} & \cdots & \frac{\partial^2}{\partial x_n \partial y_m} \end{bmatrix}$$

And therefore, a necessary optimality conditions of (BP) can be derived from Corollary 4.11 and 4.12. We have the following results.

Corollary 4.13 *Suppose f is Lipschiz continuous, g, ψ are twice continuously differentiable. Further assume that g is pseudoconvex in y, ψ is quasiconvex in y. Let $(\bar{x}, \bar{y})$ solve the problem (BP). Suppose that $\nabla_y g_i(\bar{x}, \bar{y})$ for*

$i \in I = \{i : g_i(\bar{x}, \bar{y}) = 0\}$ are linearly independent and $\bar{u}$ is a corresponding multiplier. If

$$A(\bar{x}, \bar{z})^\top \eta = 0, \quad \nabla_z F(\bar{x}, \bar{z})^\top \eta \in D^* N((\bar{z}, F(\bar{x}, \bar{z})), \Omega)(\eta)$$

implies that $\eta = 0$, then there exist $r > 0$, $\eta = (\eta_1, \eta_2) \in rB_{m+q}$ such that

$$0 \in \partial_x f(\bar{x}, \bar{y}) - \nabla_x \psi(\bar{x}, \bar{y})^\top \eta_2 + (\nabla^2_{xy} g + \bar{u} \nabla^2_{xy} \psi)(\bar{x}, \bar{y})^\top \eta_1 \tag{4.26}$$

$$0 \in \partial_y f(\bar{x}, \bar{y}) - \nabla_y \psi(\bar{x}, \bar{y})^\top \eta_2 + (\nabla^2_{yy} g + \bar{u} \nabla^2_{yy} \psi)(\bar{x}, \bar{y})^\top \eta_1 \tag{4.27}$$

$$(-\nabla_y \psi(\bar{x}, \bar{y}) \eta_1, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), gphN(\cdot, R^q_+)) \tag{4.28}$$

where $F(x, z)$ and $A(x, z)$ are defined as in (4.24) and (4.25) respectively.

Proof. The linear independence assumption of $\nabla_y g_i(\bar{x}, \bar{y})$ for $i \in I$ implies that there exists a multiplier $\bar{u}$ such that the Kuhn-Tucker condition is satisfied. Since the objective function of the lower level problem g is pseudoconvex in y and the constraint ϕ is quasiconvex in y, by [6, Theorem 4.2.11] the Kuhn-Tucker condition is a necessary and sufficient condition for optimality. Therefore we know that $(\bar{x}, \bar{y})$ is a solution of the following problem:

$$\begin{aligned} \min \quad & \tilde{f}(x, z) \\ \text{s.t.} \quad & 0 \in -F(x, z) + N(z, \Omega) \end{aligned}$$

where $z = (y, u) \in R^{m+q}$, $\Omega = R^m \times R^q_+$ and $\tilde{f}(x, z) = f(x, y)$.

It is straightforward to show

$$\nabla_x F(\bar{x}, \bar{y}) = A(\bar{x}, \bar{y})$$

and

$$\nabla_z F(\bar{x}, \bar{z})^\top = \begin{bmatrix} -(\nabla^2_{yy} g + u \nabla^2_{yy} \psi)(\bar{x}, \bar{y}) & \nabla_y \psi(\bar{x}, \bar{y})^\top \\ -\nabla_y \psi(\bar{x}, \bar{y}) & 0 \end{bmatrix}. \tag{4.29}$$

Apparently, the assumptions of the orollary imply that the conditions (1) – (3) in Corollary 4.11 (for the case of $\Omega = R^m$) hold. And hence applying Corollary 4.11, we obtain

$$\nabla_x F(\bar{x},\bar{z})^\top \eta \in \partial_x \tilde{f}(\bar{x},\bar{z}) = \partial_x f(\bar{x},\bar{y}) \tag{4.30}$$

and

$$\nabla_z F(\bar{x},\bar{z})^\top \eta \in \partial_z \tilde{f}(\bar{x},\bar{z}) + D^*N((\bar{z},F(\bar{x},\bar{z})),\Omega)(\eta) \tag{4.31}$$

where $\eta \in rB_{m+q}$ for some $r > 0$.

Denote $\eta = (\eta_1,\eta_2) \in R^m \times R^q$. (4.30) become

$$-(\nabla^2_{xy} g + u\nabla^2_{xy}\psi)(\bar{x},\bar{y})\eta_1 + \nabla_x\psi(\bar{x},\bar{y})\eta_2 \in \partial_x f(\bar{x},\bar{y}).$$

Thus (4.26) hold.

As to (4.27) and (4.28), note that we have

$$\partial_z \tilde{f}(\bar{x},\bar{z}) = (\partial_y f(\bar{x},\bar{y}),0).$$

One the other hand, for any $(v,w) \in D^*N((\bar{z},F(\bar{x},\bar{z})),\Omega)(\eta)$, by the definition of coderivatives we have

$$(v,w,-\eta) = (v,w,-\eta_1,-\eta_2) \in N((\bar{z},F(\bar{x},\bar{z})), gphN(\cdot,\Omega))$$

Due to

$$\begin{aligned} gphN(\cdot,\Omega) &= \{((y,u),(w_1,w_2)) : y \in R^m, w_1 = 0, w_2 \in N(u,R^q_+)\} \\ &= R^m \times \{(u,0,w_2) : w_2 \in N(u,R^q_+)\} \end{aligned}$$

we have

$$v = 0, \quad (w, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), gphN(\cdot, R^q_+)). \tag{4.32}$$

Therefore combining (4.29),(4.31) and (4.32), we obtain

$$-(\nabla^2_{yy} g + \nabla^2_{yy}\psi)(\bar{x}, \bar{y})\eta_1 + \nabla_y \psi(\bar{x}, \bar{y})\eta_2 \in \partial_y f(\bar{x}, \bar{y})$$
$$(-\nabla_y \psi(\bar{x}, \bar{y})\eta_1, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), gphN(\cdot, R^q_+)).$$

The proof is complete. □

Example 4.14 *Consider the following classical bilevel problem:*

$$\begin{aligned} \min \quad & x^2 - 2y \\ s.t. \quad & x \in R \\ & y \in \arg\min\{y^2 - 2xy : y - 1 \le 0, -y \le 0\} \end{aligned}$$

Suppose that $(\bar{x}, \bar{y})$ solves the problem. Using the same notation as that of Corollary 4.13, *we have*

$$F(\bar{x}, \bar{z}) = \begin{bmatrix} -2\bar{y} + 2\bar{x} - \bar{u}_1 + \bar{u}_2 \\ \bar{y} - 1 \\ -\bar{y} \end{bmatrix}$$

$$0 \in F(\bar{x}, \bar{z}) + N(\bar{z}, \Omega)$$

where $z \in (\bar{z}, \bar{u}_1, \bar{u}_2)$, $\quad \Omega = R \times R^2_+$.

Therefore

$$A(\bar{x}, \bar{z})^\top = \nabla_x F(\bar{x}, \bar{z})^\top = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \quad \nabla_z F(\bar{x}, \bar{z})^\top = \begin{bmatrix} -2 & 1 & -1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Denote $p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$. $\nabla_z F(\bar{x},\bar{z})^\top p = \begin{bmatrix} -2p_1 + p_2 - p_3 \\ -p_1 \\ p_1 \end{bmatrix}$.

Note that

$$N((\bar{u}_1,\bar{u}_2),R^2_+) = \begin{cases} \{(0,0)\}, & \bar{u}_1 > 0, \bar{u}_2 > 0 \\ (-\infty,0]^2, & \bar{u}_1 = 0, \bar{u}_2 = 0 \\ (-\infty,0] \times \{0\}, & \bar{u}_1 = 0, \bar{u}_2 > 0 \\ \{0\} \times (-\infty,0], & \bar{u}_1 > 0, \bar{u}_2 = 0 \end{cases}$$

$$\begin{aligned} gphN(\cdot,\Omega) = R \quad \times \quad & \{[(0,\infty) \times (0,\infty) \times \{0\} \times \{0\} \times \{0\}] \\ & \cup[\{0\} \times \{0\} \times \{0\} \times (-\infty,0] \times (-\infty,0]] \\ & \cup[\{0\} \times (0,\infty) \times \{0\} \times (-\infty,0] \times \{0\}] \\ & \cup[(0,\infty) \times \{0\} \times \{0\} \times \{0\} \times (-\infty,0]]\}. \end{aligned}$$

Thus

$$\begin{aligned} gphN(\cdot,R^2_+) \quad = \quad & [(0,\infty) \times (0,\infty) \times \{0\} \times \{0\}] \\ & \cup[\{0\} \times \{0\} \times (-\infty,0] \times (-\infty,0]] \\ & \cup[\{0\} \times (0,\infty) \times (-\infty,0] \times \{0\}] \\ & \cup[(0,\infty) \times \{0\} \times \{0\} \times (-\infty,0]]. \end{aligned}$$

Suppose $A(\bar{x},\bar{z})^\top p = 0$ *and* $\nabla_z F(\bar{x},\bar{z})^\top p \in D^* N((\bar{z},F(\bar{x},\bar{z})),\Omega)(p)$. *We obtain from the equality that* $p_1 = 0$, *while the inclusion implies that*

$$(\nabla_z F(\bar{x},\bar{z})^\top p, -p) \in N((\bar{z},F(\bar{x},\bar{z})), gphN(\cdot,\Omega)),$$

that is,

$$(-2p_1 + p_2 - p_3, -p_1, p_1, -p_1, -p_2, -p_3) \in N((\bar{z}, -F(\bar{x}, \bar{z})), gphN(\cdot, \Omega)).$$

It is easy to see that $p_2 = p_3 = 0$.

Therefore the condition of Corollary 4.13 *are satisfied. Thus, by necessary condition* (4.26)-(4.28), *there exist* $r > 0$, $\eta = (\eta_1, \eta_2^1, \eta_2^2) \in rB_3$ *such that*

(1) $\eta_1 = \bar{x}$.

(2) $2\bar{x} - 2 = \eta_2^1 - \eta_2^2$.

(3) $(-\eta_1, \eta_1, -\eta_2^1, -\eta_2^2) \in N((\bar{u}_1, \bar{u}_2, \bar{y} - 1, -\bar{y}), gphN(\cdot, R_+^2))$,

(4) $(\bar{u}_1, \bar{u}_2)$ *satisfies* $0 \in F(\bar{x}, \bar{z}) + N(\bar{z}, \Omega)$, *that is,*

$$2\bar{y} - 2\bar{x} + \bar{u}_1 - \bar{u}_2 = 0$$

$$\bar{u}_1 \geq 0, \quad \bar{u}_2 \geq 0, \quad \bar{u}_1(\bar{y} - 1) = 0, \quad \bar{u}_2(-\bar{y}) = 0.$$

We now discuss possible cases:

Case (1) *Suppose* $\bar{y} \in (0, 1)$, *then* $\bar{u}_1 = \bar{u}_2 = 0$ *by* (4). *Thus* $\bar{y} = \bar{x}$. *On the other hand, by* (3) *we have* $\eta_2^1 = \eta_2^2 = 0$ *and hence* $\bar{x} = 1$ *by* (2). *Contradiction.*

Case (2) *Suppose* $\bar{y} = 0$. *Then* $\bar{u}_1 = 0$ *by* (4). *By* (3) *we know* $\eta_2^1 = 0$. *Therefore, if* $\bar{u}_2 > 0$, *using* (3) *one has* $\eta_1 = 0$ *and hence* $\bar{x} = 0$ *which contradicts* (4). *If* $\bar{u}_2 = 0$, *by* (4) *we have* $\bar{x} = \bar{y} = 0$ *which by* (1) *and* (2) *implies* $\eta_1 = 0$, $\eta_2^2 = 2$. *This contradicts that* $-\eta_2^2 \in [0, +\infty)$ *by* (3).

Case (3) $\bar{y} = 1$. *Then by* (4) *and* (3) *we have* $\bar{u}_2 = 0$ *and* $\eta_2^2 = 0$. *Suppose* $\bar{u}_1 > 0$, *then* $\eta_1 = 0$,*that is,* $\bar{x} = 0$. *By* (4), *it is easy to see that it is impossible. Therefore,* $\bar{u}_1 = 0$ *and by* (4) *one has* $\bar{x} = \bar{y} = 1$. *By* (1) *and* (2) *we have* $\eta_1 = 1$, $\eta_2^1 = 0$.

Therefore we conclude that $\bar{x} = \bar{y} = 1, \quad \eta = (1, 0, 0)$.

Bibliography

[1] G. Anandalingam and T.L. Friesz, Editors, Hierarchical Optimization, *Annals of Operations Research* **34**, 1992.

[2] J.P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.

[3] J.P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhauser, Boston, 1990.

[4] J.F. Bard, Some properties of the bilevel programming problem, *Journal of Optimization Theory and Applications* **68**(1991), 371-378.

[5] V. Barbu, *Optimal Control of Variational Inequalities* Pitman, Boston, 1984.

[6] M.S. Bazaraa and C.M. Shetty, *Nonlinear Programming Theory and Algorithms*, John Wiley & Sons, Inc. 1979.

[7] C. Berge, *Espaces Topologiques et Fonctions Multivoques*, Dunod, Paris, 1959.

[8] F.E. Browder, Nonlinear monotone operators and convex sets in Banach spaces, *Bull. Amer. Math. Soc.,* **71**(1965), 780-785.

[9] F.H. Clarke, *Optimation and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.

[10] R.W. Cottle, Nonlinear programs with positively bounded Jacobians, *SIAM Journal on Applied Mathematics* **14**(1966), 147-158.

[11] R.W. Cottle, J. Kyparisis, and J.S. Pang, Foreword to a special issue devoted to variational inequality problems, *Math. Programming* **48**(1990), 3-4.

[12] G.B. Dantzig and R.W.Cottle, Positive (Semi-)definite programming, in: J, Abadie, ed., *Nonlinear Programming*(1968), 55-73.

[13] S. Dempe, A necessary and sufficient optimality condition for bilevel programming problems, *Optimization* **25**(1992), 314-354.

[14] I. Ekeland, Nonconvex minimization problems, *Bull. Am. Math. Soc.* **1**(1979), 443-474.

[15] T.L. Friesz, R.L. Tobin, H.J. Cho, and N.J. Mehta, Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints, *Math. Programming,* **48**(1990), 265-284.

[16] T.L. Friesz, R.L. Tobin,and T. Miller, Existence theory for spatially competitive network facility location models, *Ann. Oper. Res.*,**18**(1989), 267-276.

[17] T.L. Friesz, Hsun-Jun Cho, N.J. Mehta, R.L. Tobin, and G. Anandalingam, A Simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Sci.*, **26**(1992), 18-26.

[18] P.T. Harker and S.C. Choi, A penalty function approach for mathematical programs with variational inequality constraints, *WP 87-09-08*, Decision Sciences Department, Univ. of Pennsylvania, Philadelphia, PA, 1987.

[19] P.T. Harker and J.S. Pang, Finite-dimensional variational inequality and nonlinear complimentarity problems: a survey of theory, algorithms and applications, *Mathematical Programming* 48(1990), 161-220.

[20] P. Hartman and G. Stampacchia, On some nonlinear elliptic differential functional equations, *Acta Math.,* **115**(1966), 153-188.

[21] J. Haslinger and P. Neittaanmaki, Finite element approximation for optimal shape design: *Theory and Applications*, J. Wiley & Sons, Chichester, UK, 1988.

[22] A. J. King and R. T. Rockafellar, Sensitivity analysis for nonsmooth generalized equations, *Mathematical Programming* **55**(1992), 193-212.

[23] A.Y. Kruger and B.S. Mordukhovich, Extremal points and Euler equations in nonsmooth optimization, *Dokl. Akad. Nauk BSSR* **24**(1980), 684-687.

[24] P. Loridan and J. Morgan, A theoretical approximation scheme for Stackelberg problems, *Journal of Optimization Theory and Applications* **11**(1989), 95-110.

[25] Z.Q. Luo, J.S. Pang, D. Ralph and S.Q. Wu, Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints, Preprint.

[26] P. Marcotte, Network design problem with congestion effects: A case of bilevel programming, *Math. Programming*, **34**(1986), 142-162.

[27] F. Mignot and J.P. Puel, Optimal control in some variational inequalities, *SIAM J. Control Optim.*, **22**(1984), 466-476.

[28] B.S. Mordukhovich, Generalized Differential Calculus for Nonsmooth and Set-Valued Mappings, *Journal of Mathematical Analysis and Applications* **183**(1994), 250-288.

[29] B.S. Mordukhovich, Lipschitzian stability of constraint systems and generalized equations, *Nonlinear Anal.* **22**(1994), 173-206.

[30] B.S. Mordukhovich, Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis, *Transactions of the American Mathematical Society* **343**(1994), 609-655.

[31] J. V. Outrata, On optimization problems with variational inequality constraints, *SIAM J. Optimization.* **2**(1994), 340-357.

[32] J.V. Outrata, Necessary optimality conditions for Stackelberg problems, *Journal of Optimization Theory and Applications* **76** (1993), 305-320.

[33] S.M. Robinson, Strongly regular generalized equations, *Mathematics of Operations Research* **5**(1980), 43-62.

[34] S.M. Robinson, Generalized equations and their solutions, Part I: base theory, *Math. Programming Study* **10**(1979), 128-141.

[35] Shuzhong Shi, Optimal control of strongly monotone variational inequalities,*SIAM J. Control Optim.,* **26**(1988), 274-290.

[36] Shuzhong Shi, Erratum: Optimal control of strongly monotone variational inequalities, *SIAM J. Control Optim.,* **28**(1990), 243-249.

[37] H.V. Stackelberg, *The Theory of the Market Economy*, Oxford University Press, Oxford, England, 1952.

[38] R. Tobin and T.L. Friesz, Spatial competition facility location models: definition, formulation and solution approach, *Ann. Oper. Res.*, **6**(1986), 49-74.

[39] J.J. Ye, New uniform parametric error bounds, Preprint.

[40] J.J. Ye and D.L. Zhu, Optimality conditions for bilevel programming problems, *Optimization,* **33**(1995), 9-27.

[41] J.J. Ye and D.L. Zhu and Q.J. Zhu, Exact penalization and necessary optimality conditions for generalized bilevel programming problems, Preprint.

[42] J.J. Ye, Necessary conditions for bilevel dynamic optimization problems, *SIAM J. Control and Optimization* 4(1995), 1208-1223

[43] R. Zhang, Problems of hierarchical optimization: Nonsmoothness and analysis of solutions, Ph.D. thesis, Dept. of Applied Mathematics, University of Washingto, Seattle, 1990.

[44] R. Zhang, Problems of hierarchical optimization in finite dimensions, To appear in *SIAM J. Optimization.*

VITA

Surname: Ye	Given Names: Xiang Yang
Place of Birth: P.R.China	Date of Birth: March 2, 1961

Educational Institutions Attended:

University of Victoria:	1994 to 1995
Shanghai Jiao Tong University:	1985 to 1987
Xiamen University:	1978 to 1982

Degree Awarded:

B.Sc.	Xiamen University 1982
Graduate certificate	Shanghai Jiao Tong University 1987

Honours and Awards:

University of Victoria Fellowship	1994-1995
Victoria Canada-China Friendship Association Bursaries	1994-1995
People's Prize	1985-1987
People's Prize	1978-1982

PARTIAL COPYRIGHT LICENSE

Title of Thesis: **OPTIMIZATION PROBLEMS WITH VARIATIONAL INEQUALITY CONTRAINTS**

Author [redacted]

XIANG YANG YE

(Name in Block Letter)

Sept. 26, 1995

(Date)