583 LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS

Igor V. Konnov
Dinh The Luc
Alexander M. Rubinov
Editors

Generalized Convexity and Related Topics

Springer

Lecture Notes in Economics and Mathematical Systems 583

Igor V. Konnov · Dinh The Luc · Alexander M. Rubinov†

Generalized Convexity and Related Topics

With 11 Figures

Springer

Professor Igor V. Konnov
Department of Applied Mathematics
Kazan University
ul. Kremlevskaya, 18
Kazan 420008
Russia
ikonnov@ksu.ru

Professor Dinh The Luc
Department de Mathematiques
33 rue Louis Pasteur
84000 Avignon
France
dtluc@univ-avignon.fr

Professor Alexander M. Rubinov†
SITMS
University of Ballarat
University Dr. 1
3353 Victoria
Australia
amr@ballarat.edu.au

Library of Congress Control Number: 2006934206

ISBN-10 3-540-37006-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-37006-2 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Typesetting: Camera ready by author
Cover: Erich Kirchner, Heidelberg
Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig

SPIN 11811275 Printed on acid-free paper – 88/3100YL – 5 4 3 2 1 0

Preface

In mathematics generalization is one of the main activities of researchers. It opens up new theoretical horizons and broadens the fields of applications. Intensive study of generalized convex objects began about three decades ago when the theory of convex analysis nearly reached its perfect stage of development with the pioneering contributions of Fenchel, Moreau, Rockafellar and others. The involvement of a number of scholars in the study of generalized convex functions and generalized monotone operators in recent years is due to the quest for more general techniques that are able to describe and treat models of the real world in which convexity and monotonicity are relaxed. Ideas and methods of generalized convexity are now within reach not only in mathematics, but also in economics, engineering, mechanics, finance and other applied sciences.

This volume of referred papers, carefully selected from the contributions delivered at the 8th International Symposium on Generalized Convexity and Monotonicity (Varese, 4-8 July, 2005), offers a global picture of current trends of research in generalized convexity and generalized monotonicity. It begins with three invited lectures by Konnov, Levin and Pardalos on numerical variational analysis, mathematical economics and invexity, respectively. Then come twenty four full length papers on new achievements in both the theory of the field and its applications. The diapason of the topics tackled in these contributions is very large. It encompasses, in particular, variational inequalities, equilibrium problems, game theory, optimization, control, numerical methods in solving multiobjective optimization problems, consumer preferences, discrete convexity and many others.

The volume is a fruit of intensive work of more than hundred specialists all over the world who participated at the latest symposium organized by the Working Group on Generalized Convexity (WGGC) and hosted by the Insubria University. This is the 6th proceedings edited by WGGC, an interdisciplinary research community of more than 300 members from 36 countries (http://www.gencov.org). We hope that it will be useful for students, researchers and practitioners working in applied mathematics and related areas.

Acknowledgement. We wish to thank all the authors for their contributions, and all the referees whose hard work was indispensable for us to maintain the scientific quality of the volume and greatly reduce the publication delay. Special thanks go to the Insubria University for the organizational and financial support of the symposium which has contributed greatly to the success of the meeting and its outcome in the form of the present volume.

Kazan, Avignon and Ballarat
August 2006

I.V. Konnov
D.T. Luc
A.M. Rubinov

Contents

Part I

Invited Papers

Combined Relaxation Methods for Generalized Monotone Variational Inequalities

Igor V. Konnov

Department of Applied Mathematics, Kazan University, Kazan, Russia
ikonnov@ksu.ru

Summary. The paper is devoted to the combined relaxation approach to constructing solution methods for variational inequalities. We describe the basic idea of this approach and implementable methods both for single-valued and for multi-valued problems. All the combined relaxation methods are convergent under very mild assumptions. This is the case if there exists a solution to the dual formulation of the variational inequality problem. In general, these methods attain a linear rate of convergence. Several classes of applications are also described.

Key words: Variational inequalities, generalized monotone mappings, combined relaxation methods, convergence, classes of applications.

1 Introduction

Variational inequalities proved to be a very useful and powerful tool for investigation and solution of many equilibrium type problems in Economics, Engineering, Operations Research and Mathematical Physics. The paper is devoted to a new general approach to constructing solution methods for variational inequalities, which was proposed in [17] and called the *combined relaxation* (CR) approach since it combines and generalizes ideas contained in various relaxation methods. Since then, it was developed in several directions and many works on CR methods were published including the book [29]. The main goal of this paper is to give a simple and clear description of the current state of this approach, its relationships with the known relaxation methods, and its abilities in solving variational inequalities with making an emphasis on generalized monotone problems. Due to the space limitations, we restrict ourselves with simplified versions of the methods, remove some proofs, comparisons with other methods, and results of numerical experiments. Any interested reader can find them in the references.

We first describe the main idea of relaxation and combined relaxation methods.

1.1 Relaxation Methods

Let us suppose we have to find a point of a convex set X^* defined implicitly in the n-dimensional Euclidean space $\mathbb{R}^n$. That is, X^* may be a solution set of some problem. One of possible ways of approximating a point of X^* consists in generating an iteration sequence $\{x^k\}$ in conformity with the following rule:

- *The next iterate x^{k+1} is the projection of the current iterate x^k onto a hyperplane separating strictly x^k and the set X^*.*

Then the process will possess the *relaxation* property:

- *The distances from the next iterate to each point of X^* cannot increase in comparison with the distances from the current iterate.*

This property is also called Fejer-monotonicity. It follows that the sequence $\{x^k\}$ is bounded, hence, it has limit points. Moreover, due to the above relaxation property, if there exists a limit point of $\{x^k\}$ which belongs to X^*, the whole sequence $\{x^k\}$ converges to this point. These convergence properties seem very strong. We now discuss possible ways of implementation of this idea.

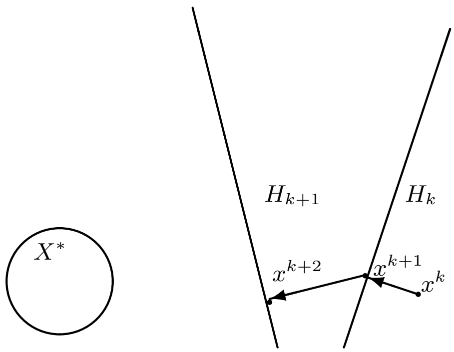

Fig. 1. The relaxation process

First we note that the separating hyperplane H_k is determined completely by its normal vector g^k and a distance parameter ω_k, i.e.

$$H_k = \{x \in \mathbb{R}^n \mid \langle g^k, x^k - x\rangle = \omega_k\}.$$

The hyperplane H_k is strictly separating if

$$\langle g^k, x^k - x^*\rangle \geq \omega_k > 0 \qquad \forall x^* \in X^*. \tag{1}$$

It also means that the half-space

$$\{x \in \mathbb{R}^n \mid \langle g^k, x^k - x\rangle \geq \omega_k\}$$

contains the solution set X^* and represents the image of this set at the current iterate. Then the process is defined by the explicit formula:

$$x^{k+1} = x^k - (\omega_k/\|g^k\|^2)g^k, \tag{2}$$

and the easy calculation confirms the above relaxation property:

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - (\omega_k/\|g^k\|)^2 \qquad \forall x^* \in X^*;$$

see Fig. 1. However, (1) does not ensure convergence of this process in general. We say that the rule of determining a separating hyperplane is *regular*, if the correspondence $x^k \mapsto \omega_k$ possesses the property:

$$(\omega_k/\|g^k\|) \to 0 \qquad \text{implies} \qquad x^* \in X^*$$

for at least one limit point x^* of $\{x^k\}$.

- *The above relaxation process with a regular rule of determining a separating hyperplane ensures convergence to a point of X^*.*

There exist a great number of algorithms based on this idea. For linear equations such relaxation processes were first suggested by S. Kaczmarz [12] and G. Cimmino [7]. Their extensions for linear inequalities were first proposed by S. Agmon [1] and by T.S. Motzkin and I.J. Schoenberg [35]. The relaxation method for convex inequalities is due to I.I. Eremin [8]. A modification of this process for the problem of minimizing a convex function $f : \mathbb{R}^n \to \mathbb{R}$ with the prescribed minimal value f^* is due to B.T. Polyak [40]. Without loss of generality we can suppose that $f^* = 0$. The solution is found by the following gradient process

$$x^{k+1} = x^k - (f(x^k)/\|\nabla f(x^k)\|^2)\nabla f(x^k), \tag{3}$$

which is clearly an implementation of process (2) with $g^k = \nabla f(x^k)$ and $\omega_k = f(x^k)$, since (1) follows from the convexity of f:

$$\langle \nabla f(x^k), x^k - x^*\rangle \geq f(x^k) > 0 \qquad \forall x^* \in X^* \tag{4}$$

for each non-optimal point x^k. Moreover, by continuity of f, the rule of determining a separating hyperplane is regular. Therefore, process (3) generates a sequence $\{x^k\}$ converging to a solution. Note that process (3) can be also viewed as an extension of the Newton method. Indeed, the next iterate x^{k+1} also solves the linearized problem

$$f(x^k) + \langle \nabla f(x^k), x - x^k\rangle = 0,$$

and, in case $n = 1$, we obtain the usual Newton method for the nonlinear equation $f(x^*) = 0$; see Fig. 2. This process can be clearly extended for

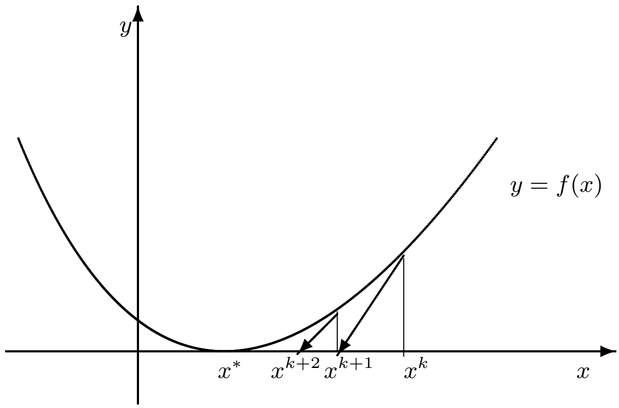

Fig. 2. The Newton method

the non-differentiable case. It suffices to replace $\nabla f(x^k)$ with an arbitrary subgradient g^k of the function f at x^k. Afterwards, it was noticed that the process (3) (hence (2)) admits the additional relaxation parameter $\gamma \in (0, 2)$:

$$x^{k+1} = x^k - \gamma(\omega_k/\|g^k\|^2)g^k,$$

which corresponds to the projection of x^k onto the shifted hyperplane

$$H_k(\gamma) = \{x \in \mathbb{R}^n \mid \langle g^k, x^k - x\rangle = \gamma\omega_k\}. \quad (5)$$

1.2 Combined Relaxation Methods

We now intend to describe the implementation of the relaxation idea in solution methods for variational inequality problems with (generalized) monotone mappings. We begin our considerations from variational inequalities with single-valued mappings.

Let X be a nonempty, closed and convex subset of the space $\mathbb{R}^n$, $G : X \to \mathbb{R}^n$ a continuous mapping. The *variational inequality problem* (VI) is the problem of finding a point $x^* \in X$ such that

$$\langle G(x^*), x - x^*\rangle \geq 0 \quad \forall x \in X. \quad (6)$$

We denote by X^* the solution set of problem (6). Now we recall definitions of monotonicity type properties.

Definition 1. Let Y be a convex set in $\mathbb{R}^n$. A mapping $Q : Y \to \mathbb{R}^n$ is said to be

(a) *strongly monotone* if there exists a scalar $\tau > 0$ such that

$$\langle Q(x) - Q(y), x - y \rangle \geq \tau \|x - y\|^2 \qquad \forall x, y \in Y;$$

(b) *strictly monotone* if

$$\langle Q(x) - Q(y), x - y \rangle > 0 \qquad \forall x, y \in Y, x \neq y;$$

(c) *monotone* if

$$\langle Q(x) - Q(y), x - y \rangle \geq 0 \qquad \forall x, y \in Y;$$

(d) *pseudomonotone* if

$$\langle Q(y), x - y \rangle \geq 0 \quad \Longrightarrow \quad \langle Q(x), x - y \rangle \geq 0 \qquad \forall x, y \in Y;$$

(e) *quasimonotone* if

$$\langle Q(y), x - y \rangle > 0 \quad \Longrightarrow \quad \langle Q(x), x - y \rangle \geq 0 \qquad \forall x, y \in Y;$$

(f) *strongly pseudomonotone* if there exists a scalar $\tau > 0$ such that

$$\langle Q(y), x - y \rangle \geq 0 \quad \Longrightarrow \quad \langle Q(x), x - y \rangle \geq \tau \|x - y\|^2 \qquad \forall x, y \in Y.$$

It follows from the definitions that the following implications hold:

$$(a) \Longrightarrow (b) \Longrightarrow (c) \Longrightarrow (d) \Longrightarrow (e) \quad \text{and} \quad (a) \Longrightarrow (f) \Longrightarrow (d).$$

All the reverse assertions are not true in general.

First of all we note that the streamlined extension of the above method does not work even for general monotone (but non strictly monotone) mappings. This assertion stems from the fact that one cannot compute the normal vector g^k of a hyperplane separating strictly the current iterate x^k and the set X^* by using only information at the point x^k under these conditions, as the following simple example illustrates.

Example 1. Set $X = \mathbb{R}^n$, $G(x) = Ax$ with A being an $n \times n$ skew-symmetric matrix. Then G is monotone, $X^* = \{\mathbf{0}\}$, but for any $x \notin X^*$ we have

$$\langle G(x), x - x^* \rangle = \langle Ax, x \rangle = 0,$$

i.e., the angle between $G(x^k)$ and $x^k - x^*$ with $x^* \in X^*$ need not be acute (cf.(4)).

Thus, all the previous methods, which rely on the information at the current iterate, are single-level ones and cannot be directly applied to variational inequalities. Nevertheless, we are able to suggest a general relaxation method with the basic property that the distances from the next iterate to each point of X^* cannot increase in comparison with the distances from the current iterate.

The new approach, which is called the *combined relaxation* (CR) approach, is based on the following principles.

- The algorithm has a two-level structure.
- The algorithm involves an auxiliary procedure for computing the hyperplane separating strictly the current iterate and the solution set.
- The main iteration consists in computing the projection onto this (or shifted) hyperplane with possible additional projection type operations in the presence of the feasible set.
- An iteration of most descent methods can serve as a basis for the auxiliary procedure with a regular rule of determining a separating hyperplane.
- There are a number of rules for choosing the parameters of both the levels.

This approach for variational inequalities and its basic principles were first proposed in [17], together with several implementable algorithms within the CR framework. Of course, it is possible to replace the half-space containing the solution set by some other "regular" sets such as an ellipsoid or a polyhedron, but the implementation issues and preferences of these modifications need thorough investigations.

It turned out that the CR framework is rather flexible and allows one to construct methods both for single-valued and for multi-valued VIs, including nonlinearly constrained problems. The other essential feature of all the CR methods is that they are convergent under very mild assumptions, especially in comparison with the methods whose iterations are used in the auxiliary procedure. In fact, this is the case if there exists a solution to the dual formulation of the variational inequality problem. This property enables one to apply these methods for generalized monotone VIs and their extensions.

We recall that the solution of VI (6) is closely related with that of the following problem of finding $x^* \in X$ such that

$$\langle G(x), x - x^* \rangle \geq 0 \quad \forall x \in X. \tag{7}$$

Problem (7) may be termed as the dual formulation of VI (DVI), but is also called the Minty variational inequality. We denote by X^d the solution set of problem (7). The relationships between solution sets of VI and DVI are given in the known Minty Lemma.

Proposition 1. *[34, 13]*
(i) X^d is convex and closed.
(ii) $X^d \subseteq X^*$.
(iii) If G is pseudomonotone, $X^* \subseteq X^d$.

The existence of solutions of DVI plays a crucial role in constructing CR methods for VI; see [29]. Observe that pseudomonotonicity and continuity of G imply $X^* = X^d$, hence solvability of DVI (7) follows from the usual existence results for VI (6). This result can be somewhat strengthened for explicit quasimonotone and properly quasimonotone mappings, but, in the quasimonotone case, problem (7) may have no solutions even on the compact convex feasible sets. However, we can give an example of solvable DVI (7) with the underlying mapping G which is not quasimonotone; see [11] and [29] for more details.

2 Implementable CR Methods for Variational Inequalities

We now consider implementable algorithms within the CR framework for solving VIs with continuous single-valued mappings. For the sake of clarity, we describe simplified versions of the algorithms.

2.1 Projection-based Implementable CR Method

The blanket assumptions are the following.

- *X is a nonempty, closed and convex subset of $\mathbb{R}^n$;*
- *Y is a closed convex subset of $\mathbb{R}^n$ such that $X \subseteq Y$;*
- *$G : Y \to \mathbb{R}^n$ is a continuous mapping;*
- *$X^d \neq \emptyset$.*

The first implementable algorithms within the CR framework for VIs under similar conditions were proposed in [17]. They involved auxiliary procedures for finding the strictly separating hyperplanes, which were based on an iteration of the projection method, the Frank-Wolfe type method, and the symmetric Newton method. The simplest of them is the projection-based procedure which leads to the following method.

Method 1.1. *Step 0 (Initialization)*: Choose a point $x^0 \in X$ and numbers $\alpha \in (0,1)$, $\beta \in (0,1)$, $\gamma \in (0,2)$. Set $k := 0$.

Step 1 (Auxiliary procedure):

Step 1.1 : Solve the auxiliary VI of finding $z^k \in X$ such that

$$\langle G(x^k) + z^k - x^k, y - z^k \rangle \geq 0 \quad \forall y \in X, \tag{8}$$

and set $p^k := z^k - x^k$. If $p^k = 0$, stop.

Step 1.2: Determine m as the smallest number in Z_+ such that

$$\langle G(x^k + \beta^m p^k), p^k \rangle \leq \alpha \langle G(x^k), p^k \rangle, \tag{9}$$

set $\theta_k := \beta^m$, $y^k := x^k + \theta_k p^k$. If $G(y^k) = 0$, stop.

Step 2 (Main iteration): Set

$$g^k := G(y^k), \omega_k := \langle g^k, x^k - y^k \rangle, x^{k+1} := \pi_X[x^k - \gamma(\omega_k / \|g^k\|^2) g^k], \tag{10}$$

$k := k + 1$ and go to Step 1.

Here and below Z_+ denotes the set of non-negative integers and $\pi_X[\cdot]$ denotes the projection mapping onto X.

According to the description, the method finds a solution to VI in the case of its finite termination. Therefore, in what follows we shall consider only the case of the infinite sequence $\{x^k\}$. Observe that the auxiliary procedure in fact represents a simple projection iteration, i.e.

$$z^k = \pi_X[x^k - G(x^k)],$$

and is used for finding a point $y^k \in X$ such that

$$\omega_k = \langle g^k, x^k - y^k \rangle > 0$$

when $x^k \notin X^*$. In fact, (8)–(10) imply that

$$\begin{aligned}\omega_k &= \langle G(y^k), x^k - y^k \rangle = \theta_k \langle G(y^k), x^k - z^k \rangle \\ &\geq \alpha\theta_k \langle G(x^k), x^k - z^k \rangle \geq \alpha\theta_k \|x^k - z^k\|^2.\end{aligned}$$

The point y^k is computed via the simple Armijo-Goldstein type linesearch procedure that does not require a priori information about the original problem (6). In particular, it does not use the Lipschitz constant for G.

The basic property together with (7) then implies that

$$\langle g^k, x^k - x^* \rangle \geq \omega_k > 0 \qquad \text{if} \quad x^k \notin X^d.$$

In other words, we obtain (1) where the normal vector g^k and the distance parameter $\omega_k > 0$ determine the separating hyperplane. We conclude that, under the blanket assumptions, the iteration sequence $\{x^k\}$ in Method 1.1 satisfies the following conditions:

$$\begin{array}{c} x^{k+1} := \pi_X(\tilde{x}^{k+1}), \tilde{x}^{k+1} := x^k - \gamma(\omega_k/\|g^k\|^2)g^k, \gamma \in (0,2), \\ \langle g^k, x^k - x^* \rangle \geq \omega_k \geq 0 \qquad \forall x^* \in X^d; \end{array} \tag{11}$$

therefore $\tilde{x}^{k+1}$ is the projection of x^k onto the shifted hyperplane

$$H_k(\gamma) = \{y \in \mathbb{R}^n \mid \langle g^k, x^k - y \rangle = \gamma\omega_k\},$$

(see (5)) and $H_k(1)$ separates x^k and X^d. Observe that $H_k(\gamma)$, generally speaking, does not possess this property, nevertheless, the distance from $\tilde{x}^{k+1}$ to each point of X^d cannot increase and the same assertion is true for x^{k+1} due to the projection properties because $X^d \subseteq X$. We now give the key property of the above process.

Lemma 1. *If (11) is fulfilled, then*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma(2-\gamma)(\omega_k/\|g^k\|)^2 \qquad \forall x^* \in X^d. \tag{12}$$

Proof. Take any $x^* \in X^d$. By (11) and the projection properties, we have

$$\begin{aligned}\|x^{k+1} - x^*\|^2 \leq \|\tilde{x}^{k+1} - x^*\|^2 &= \|x^k - x^*\|^2 \\ &\quad -2\gamma(\omega_k/\|g^k\|^2)\langle g^k, x^k - x^* \rangle + (\gamma\omega_k/\|g^k\|)^2 \\ &\leq \|x^k - x^*\|^2 - 2\gamma(2-\gamma)(\omega_k/\|g^k\|)^2,\end{aligned}$$

i.e. (12) is fulfilled, as desired.

The following assertions follow immediately from (12).

Lemma 2. *Let a sequence $\{x^k\}$ satisfy (11). Then:*

(i) $\{x^k\}$ is bounded.

(ii) $\sum_{k=0}^{\infty} (\omega_k/\|g^k\|)^2 < \infty$.

(iii) For each limit point x^* of $\{x^k\}$ such that $x^* \in X^d$ we have

$$\lim_{k\to\infty} x^k = x^*.$$

Note that the sequence $\{x^k\}$ has limit points due to (i). Thus, it suffices to show that the auxiliary procedure in Method 1.1 represents a regular rule of determining a separating hyperplane. Then we obtain the convergence of the method. The proof is omitted since the assertion follows from more general Theorem 2.

Theorem 1. *Let a sequence $\{x^k\}$ be generated by Method 1.1. Then:*

(i) There exists a limit point x^* of $\{x^k\}$ which lies in X^*.

(ii) If

$$X^* = X^d, \tag{13}$$

we have

$$\lim_{k\to\infty} x^k = x^* \in X^*.$$

2.2 General CR Methods and Their Modifications

The basic principles of the CR approach claim that an iteration of most descent methods can serve as a basis for the auxiliary procedure with a regular rule of determining a separating hyperplane and that there are a number of rules for choosing the parameters of both the levels. Following these principles, we now indicate ways of creating various classes of CR methods for VI (6).

First we extend the projection mapping in (10).

Definition 2. Let W be a nonempty, convex, and closed set in $\mathbb{R}^n$. A mapping $P : \mathbb{R}^n \to \mathbb{R}^n$ is said to be a *pseudo-projection* onto W, if for every $x \in \mathbb{R}^n$, it holds that

$$P(x) \in W \quad \text{and} \quad \|P(x) - w\| \le \|x - w\| \quad \forall w \in W.$$

We denote by $\mathcal{F}(W)$ the class of all pseudo-projection mappings onto W. Clearly, we can take the projection mapping $\pi_W(\cdot)$ as $P \in \mathcal{F}(W)$. The properties indicated show that the projection mapping in (10) and (11) can be replaced with the pseudo-projection $P \in \mathcal{F}(X)$. Then the assertion of Lemma 1 remains true and so are those of Lemma 2 and Theorem 1. If the definition of the set X includes functional constraints, then the projection onto X cannot be found by a finite procedure. Nevertheless, in that case there exist finite procedures of computation of values of pseudo-projection mappings; see [29] for more details. It means that the use of pseudo-projections may give certain preferences.

Next, Method 1.1 involves the simplest projection-based auxiliary procedure for determining a separating hyperplane. However, we can use more general iterations, which can be viewed as solutions of auxiliary problems approximating the initial problem at the current point x^k. In general, we can replace (8) with the problem of finding a point $z^k \in X$ such that

$$\langle G(x^k) + \lambda^{-1} T_k(x^k, z^k), y - z^k \rangle \geq 0 \quad \forall y \in X, \tag{14}$$

where $\lambda > 0$, the family of mappings $\{T_k : Y \times Y \to \mathbb{R}^n\}$ such that, for each $k = 0, 1, \ldots,$

(A1) *$T_k(x, \cdot)$ is strongly monotone with constant $\tau' > 0$ and Lipschitz continuous with constant $\tau'' > 0$ for every $x \in Y$, and $T_k(x, x) = 0$ for every $x \in Y$.*

The basic properties of problem (14) are given in the next lemma.

Lemma 3. *(i) Problem (14) has a unique solution.*

(ii) It holds that

$$\langle G(x^k), x^k - z^k \rangle \geq \lambda^{-1} \langle T_k(x^k, z^k), z^k - x^k \rangle \geq \lambda^{-1} \tau' \|z^k - x^k\|^2. \tag{15}$$

(iii) $x^k = z^k$ if and only if $x^k \in X^*$.

Proof. Assertion (i) follows directly from strong monotonicity and continuity of $T_k(x, \cdot)$. Next, using (A1) in (14) with $y = x^k$, we have

$$\begin{aligned} &\langle G(x^k), x^k - z^k \rangle \geq \lambda^{-1} \langle T_k(x^k, z^k), z^k - x^k \rangle \\ &= \lambda^{-1} \langle T_k(x^k, z^k) - T_k(x^k, x^k), z^k - x^k \rangle \geq \lambda^{-1} \tau' \|z^k - x^k\|^2, \end{aligned}$$

hence (15) holds, too. To prove (iii), note that setting $z^k = x^k$ in (14) yields $x^k \in X^*$. Suppose now that $x^k \in X^*$ but $z^k \neq x^k$. Then, by (15),

$$\langle G(x^k), z^k - x^k \rangle \leq -\lambda^{-1} \tau' \|z^k - x^k\|^2 < 0,$$

so that $x^k \notin X^*$. By contradiction, we see that assertion (iii) is also true.

There exist a great number of variants of the sequences $\{T_k\}$ satisfying (A1). Nevertheless, it is desirable that there exist an effective algorithm for solving problem (14). For instance, we can choose

$$T_k(x, z) = A_k(z - x) \tag{16}$$

where A_k is an $n \times n$ positive definite matrix. The simplest choice $A_k \equiv I$ in (16) leads to the projection method and has been presented in Method 1.1. Then problem (14) becomes much simpler than the initial VI. Indeed, it coincides with a system of linear equations when $X = \mathbb{R}^n$ or with a linear complementarity problem when $X = \mathbb{R}^n_+$ and, also, reduces to LCP when X

is a convex polyhedron. It is well-known that such problems can be solved by finite algorithms.

On the other hand, we can choose A_k (or $\nabla_z T_k(x^k, z^k)$) as a suitable approximation of $\nabla G(x^k)$. Obviously, if $\nabla G(x^k)$ is positive definite, we can simply choose $A_k = \nabla G(u^k)$. Then problem (14), (16) yields an iteration of the Newton method. Moreover, we can follow the Levenberg–Marquardt approach or make use of an appropriate quasi-Newton update. These techniques are applicable even if $\nabla G(x^k)$ is not positive definite. Thus, the problem (14) in fact represents a very general class of solution methods.

We now describe a general CR method for VI (6) converging to a solution under the blanket assumptions; see [21]. Observe that most of the methods whose iterations are used as a basis for the auxiliary procedure do not provide convergence even under the monotonicity. In fact, they need either G be co-coercive or strictly monotone or its Jacobian be symmetric, etc.

Method 1.2. *Step 0 (Initialization)*: Choose a point $x^0 \in X$, a family of mappings $\{T_k\}$ satisfying (A1) with $Y = X$ and a sequence of mappings $\{P_k\}$, where $P_k \in \mathcal{F}(X)$ for $k = 0, 1, \dots$ Choose numbers $\alpha \in (0, 1)$, $\beta \in (0, 1)$, $\gamma \in (0, 2)$, $\lambda > 0$. Set $k := 0$.

Step 1 (Auxiliary procedure):

Step 1.1 : Solve the auxiliary VI (14) of finding $z^k \in X$ and set $p^k := z^k - x^k$. If $p^k = 0$, stop.

Step 1.2: Determine m as the smallest number in Z_+ such that

$$\langle G(x^k + \beta^m p^k), p^k\rangle \le \alpha \langle G(x^k), p^k\rangle, \tag{17}$$

set $\theta_k := \beta^m$, $y^k := x^k + \theta_k p^k$. If $G(y^k) = 0$, stop.

Step 2 (Main iteration): Set

$$g^k := G(y^k), \omega_k := \langle G(y^k), x^k - y^k\rangle, x^{k+1} := P_k[x^k - \gamma(\omega_k/\|g^k\|^2)g^k],$$

$k := k + 1$ and go to Step 1.

We first show that Method 1.2 is well-defined and that it follows the CR framework.

Lemma 4. *(i) The linesearch procedure in Step 1.2 is always finite.*

(ii) It holds that

$$\langle g^k, x^k - x^*\rangle \ge \omega_k > 0 \qquad \text{if} \quad x^k \notin X^d. \tag{18}$$

Proof. If we suppose that the linesearch procedure is infinite, then (17) holds for $m \to \infty$, hence, by continuity of G,

$$(1 - \alpha)\langle G(x^k), z^k - x^k\rangle \le 0.$$

Applying this inequality in (15) gives $x^k = z^k$, which contradicts the construction of the method. Hence, (i) is true.

Next, by using (15) and (17), we have

$$\begin{aligned} \langle g^k, x^k - x^* \rangle &= \langle G(y^k), x^k - y^k \rangle + \langle G(y^k), y^k - x^* \rangle \\ &\geq \omega_k = \theta_k \langle G(y^k), x^k - z^k \rangle \geq \alpha \theta_k \langle G(x^k), x^k - z^k \rangle \\ &\geq \alpha \theta_k \lambda^{-1} \tau' \| x^k - z^k \|^2, \end{aligned} \tag{19}$$

i.e. (18) is also true.

Thus the described method follows slightly modified rules in (11), where $\pi_X(\cdot)$ is replaced by $P_k \in \mathcal{F}(X)$. It has been noticed that the assertions of Lemmas 1 and 2 then remain valid. Therefore, Method 1.2 will have the same convergence properties.

Theorem 2. *Let a sequence $\{x^k\}$ be generated by Method 1.2. Then:*

(i) If the method terminates at Step 1.1 (Step 1.2) of the kth iteration, $x^k \in X^*$ ($y^k \in X^*$).

(ii) If $\{x^k\}$ is infinite, there exists a limit point x^* of $\{x^k\}$ which lies in X^*.

(iii) If $\{x^k\}$ is infinite and (13) holds, we have

$$\lim_{k \to \infty} x^k = x^* \in X^*.$$

Proof. Assertion (i) is obviously true due to the stopping rule and Lemma 3 (iii). We now proceed to prove (ii). By Lemma 2 (ii), $\{x^k\}$ is bounded, hence so are $\{z^k\}$ and $\{y^k\}$ because of (15). Let us consider two possible cases.
Case 1: $\lim_{k \to \infty} \theta_k = 0$.
Set $\tilde{y}^k = x^k + (\theta_k/\beta) p^k$, then $\langle G(\tilde{y}^k), p^k \rangle > \alpha \langle G(x^k), p^k \rangle$. Select convergent subsequences $\{x^{k_q}\} \to x'$ and $\{z^{k_q}\} \to z'$, then $\{\tilde{y}^{k_q}\} \to x'$ since $\{x^k\}$ and $\{z^k\}$ are bounded. By continuity, we have

$$(1 - \alpha) \langle G(x'), z' - x' \rangle \geq 0,$$

but taking the same limit in (15) gives

$$\langle G(x'), x' - z' \rangle \geq \lambda^{-1} \tau' \| z' - x \|^2,$$

i.e., $x' = z'$ and (14) now yields

$$\langle G(x'), y - x' \rangle \geq 0 \quad \forall y \in X, \tag{20}$$

i.e., $x' \in X^*$.
Case 2: $\limsup_{k \to \infty} \theta_k \geq \tilde{\theta} > 0$.
It means that there exists a subsequence $\{\theta_{k_q}\}$ such that $\theta_{k_q} \geq \tilde{\theta} > 0$. Combining this property with Lemma 2 (ii) and (19) gives

$$\lim_{q \to \infty} \| x^{k_q} - z^{k_q} \| = 0.$$

Without loss of generality we can suppose that $\{x^{k_q}\} \to x'$ and $\{z^{k_q}\} \to z'$, then $x' = z'$. Again, taking the corresponding limit in (14) yields (20), i.e. $x' \in X^*$.

Therefore, assertion (ii) is true. Assertion (iii) follows from Lemma 2 (iii).

In Step 1 of Method 1.2, we first solve the auxiliary problem (14) and afterwards find the stepsize along the ray $x^k + \theta(z^k - x^k)$. Replacing the order of these steps, which corresponds to the other version of the projection method in the simplest case, we can also determine the separating hyperplane and thus obtain another CR method with involves a modified linesearch procedure; see [22]. Its convergence properties are the same as those of Method 1.2.

We now describe another CR method which uses both a modified linesearch procedure and a different rule of computing the descent direction, i.e. the rule of determining the separating hyperplane; see [24].

Method 1.3. *Step 0 (Initialization)*: Choose a point $x^0 \in Y$, a family of mappings $\{T_k\}$ satisfying (A1), and choose a sequence of mappings $\{P_k\}$, where $P_k \in \mathcal{F}(Y)$, for $k = 0, 1, \dots$ Choose numbers $\alpha \in (0,1)$, $\beta \in (0,1)$, $\gamma \in (0,2)$, $\tilde{\theta} > 0$. Set $k := 0$.

Step 1 (Auxiliary procedure):

Step 1.1 : Find m as the smallest number in Z_+ such that

$$\langle G(x^k) - G(z^{k,m}), x^k - z^{k,m} \rangle \leq (1-\alpha)(\tilde{\theta}\beta^m)^{-1} \langle T_k(x^k, z^{k,m}), z^{k,m} - x^k \rangle,$$

where $z^{k,m} \in X$ is a solution of the auxiliary problem:

$$\langle G(x^k) + (\tilde{\theta}\beta^m)^{-1} T_k(x^k, z^{k,m}), y - z^{k,m} \rangle \geq 0 \quad \forall y \in X.$$

Step 1.2: Set $\theta_k := \beta^m \tilde{\theta}$, $y^k := z^{k,m}$. If $x^k = y^k$ or $G(y^k) = 0$, stop.

Step 2 (Main iteration): Set

$$\begin{aligned} g^k &:= G(y^k) - G(x^k) - \theta_k^{-1} T_k(x^k, y^k), \\ \omega_k &:= \langle g^k, x^k - y^k \rangle, \\ x^{k+1} &:= P_k[x^k - \gamma(\omega_k / \|g^k\|^2) g^k], \end{aligned}$$

$k := k + 1$ and go to Step 1.

In this method, g^k and $\omega_k > 0$ are also the normal vector and the distance parameter of the separating hyperplane $H_k(1)$ (see (5)). Moreover, the rule of determining a separating hyperplane is regular. Therefore, the process generates a sequence $\{x^k\}$ converging to a solution. The substantiation is similar to that of the previous method and is a modification of that in [29, Section 1.4]. For this reason, the proof is omitted.

Theorem 3. *Let a sequence $\{x^k\}$ be generated by Method 1.3. Then:*

(i) If the method terminates at the kth iteration, $y^k \in X^*$.

(ii) If $\{x^k\}$ is infinite, there exists a limit point x^* of $\{x^k\}$ which lies in X^*.

(iii) If $\{x^k\}$ is infinite and (13) holds, we have

$$\lim_{k\to\infty} x^k = x^* \in X^*.$$

The essential feature of this method, unlike the previous methods, is that it involves the pseudo-projection onto Y rather than X. Hence one can simply set P_k to be the identity map if $Y = \mathbb{R}^n$ and the iteration sequence $\{x^k\}$ may be infeasible.

The convergence properties of all the CR methods are almost the same. There are slight differences in their convergence rates, which follow mainly from (12). We illustrate them by presenting some convergence rates of Method 1.3.

Let us consider the following assumption.

(A2) *There exist numbers $\mu > 0$ and $\kappa \in [0, 1]$ such for each point $x \in X$, the following inequality holds:*

$$\langle G(x), x - \pi_{X^*}(x)\rangle \geq \mu \, \|x - \pi_{X^*}(x)\|^{1+\kappa}. \tag{21}$$

Observe that Assumption (A2) with $\kappa = 1$ holds if G is strongly (pseudo) monotone and that (A2) with $\kappa = 0$ represents the so-called sharp solution.

Theorem 4. *Let an infinite sequence $\{x^k\}$ be generated by Method 1.3. If G is a locally Lipschitz continuous mapping and (A2) holds with $\kappa = 1$, then $\{\|x^k - \pi_{X^*}(x^k)\|\}$ converges to zero in a linear rate.*

We now give conditions that ensure finite termination of the method.

Theorem 5. *Let a sequence $\{x^k\}$ be constructed by Method 1.3. Suppose that G is a locally Lipschitz continuous mapping and that (A2) holds with $\kappa = 0$. Then the method terminates with a solution.*

The proofs of Theorems 4 and 5 are similar to those in [29, Section 1.4] and are omitted.

Thus, the regular rule of determining a separating hyperplane may be implemented via a great number of various procedures. In particular, an auxiliary procedure may be based on an iteration of the Frank-Wolfe type method and is viewed as a “degenerate” version of the problem (14), whereas a CR method for nonlinearly constrained problems involves an auxiliary procedure based on an iteration of a feasible direction method. However, the projection and the proximal point based procedures became the most popular; their survey can be found e.g. in [48].

3 Variational Inequalities with Multi-valued Mappings

We now consider CR methods for solving VIs which involve multi-valued mappings (or generalized variational inequalities).

3.1 Problem Formulation

Let X be a nonempty, closed and convex subset of the space $\mathbb{R}^n$, $G : X \to \Pi(\mathbb{R}^n)$ a multi-valued mapping. The *generalized variational inequality problem* (GVI for short) is the problem of finding a point $x^* \in X$ such that

$$\exists g^* \in G(x^*), \quad \langle g^*, x - x^* \rangle \geq 0 \quad \forall x \in X. \tag{22}$$

Similarly to the single-valued case, together with GVI (22), we shall consider the corresponding *dual generalized variational inequality problem* (DGVI for short), which is to find a point $x^* \in X$ such that

$$\forall\ x \in X \text{ and } \forall g \in G(x) : \langle g, x - x^* \rangle \geq 0 \tag{23}$$

(cf. (6) and (7)). We denote by X^* (respectively, by X^d) the solution set of problem (22) (respectively, problem (23)).

Definition 3. (see [29, Definition 2.1.1]) Let Y be a convex set in $\mathbb{R}^n$. A multi-valued mapping $Q : Y \to \Pi(\mathbb{R}^n)$ is said to be

(a) a *K-mapping*, if it is upper semicontinuous (u.s.c.) and has nonempty convex and compact values;

(b) *u-hemicontinuous*, if for all $x \in Y$, $y \in Y$ and $\alpha \in [0, 1]$, the mapping $\alpha \mapsto \langle Q(x + \alpha z), z \rangle$ with $z = y - x$ is u.s.c. at 0^+.

Now we give an extension of the Minty Lemma for the multi-valued case.

Proposition 2. *(see e.g. [43, 49])*

(i) The set X^d is convex and closed.

(ii) If G is u-hemicontinuous and has nonempty convex and compact values, then $X^d \subseteq X^*$.

(iii) If G is pseudomonotone, then $X^* \subseteq X^d$.

The existence of solutions to DGVI will also play a crucial role for convergence of CR methods for GVIs. Note that the existence of a solution to (23) implies that (22) is also solvable under u-hemicontinuity, whereas the reverse assertion needs generalized monotonicity assumptions. Again, the detailed description of solvability conditions for (23) under generalized monotonicity may be found in the books [11] and [29].

3.2 CR Method for the Generalized Variational Inequality Problem

We now consider a CR method for solving GVI (22) with explicit usage of constraints (see [18] and [23]). The blanket assumptions of this section are the following:

- *X is a subset of $\mathbb{R}^n$, which is defined by*

$$X = \{x \in \mathbb{R}^n \mid h(x) \le 0\},$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is a convex, but not necessarily differentiable, function;
- *the Slater condition is satisfied, i.e., there exists a point $\bar{x}$ such that $h(\bar{x}) < 0$;*
- *$G : X \to \Pi(\mathbb{R}^n)$ is a K-mapping;*
- *$X^d \neq \emptyset$.*

The method also involves a finite auxiliary procedure for finding the strictly separating hyperplane with a regular rule. Its basic scheme involves the control sequences and handles the situation of a null step, where the auxiliary procedure yields the zero vector, but the current iterate is not a solution of VI (22). The null step usually occurs if the current tolerances are too large, hence they must diminish.

Let us define the mapping $Q : \mathbb{R}^n \to \Pi(\mathbb{R}^n)$ by

$$Q(x) = \begin{cases} G(x) & \text{if } h(x) \le 0, \\ \partial h(x) & \text{if } h(x) > 0. \end{cases}$$

Method 2.1. *Step 0 (Initialization)*: Choose a point $x^0 \in X$, bounded positive sequences $\{\varepsilon_l\}$ and $\{\eta_l\}$. Also, choose numbers $\theta \in (0, 1)$, $\gamma \in (0, 2)$, and a sequence of mappings $\{P_k\}$, where $P_k \in \mathcal{F}(X)$ for $k = 0, 1, \ldots$ Set $k := 0$, $l := 1$.

Step 1 (Auxiliary procedure) :

Step 1.1 : Choose q^0 from $Q(x^k)$, set $i := 0$, $p^i := q^i$, $w^{k,0} := x^k$.

Step 1.2: If

$$\|p^i\| \le \eta_l,$$

set $x^{k+1} := x^k$, $k := k + 1$, $l := l + 1$ and go to Step 1. (*null step*)

Step 1.3: Set $w^{k,i+1} := w^{k,0} - \varepsilon_l p^i / \|p^i\|$, choose $q^{i+1} \in Q(w^{k,i+1})$. If

$$\langle q^{i+1}, p^i \rangle > \theta \|p^i\|^2,$$

then set $y^k := w^{k,i+1}$, $g^k := q^{i+1}$, and go to Step 2. (*descent step*)

Step 1.4: Set

$$p^{i+1} := \mathrm{Nr}\ \mathrm{conv}\{p^i, q^{i+1}\}, \tag{24}$$

$i := i + 1$ and go to Step 1.2.

Step 2 (Main iteration): Set $\omega_k := \langle g^k, x^k - y^k \rangle$,

$$x^{k+1} := P_k[x^k - \gamma(\omega_k / \|g^k\|^2) g^k],$$

$k := k + 1$ and go to Step 1.

Here $\mathrm{Nr}S$ denotes the element of S nearest to origin. According to the description, at each iteration, the auxiliary procedure in Step 1, which is a modification of an iteration of the simple relaxation subgradient method (see [15, 16]), is applied for direction finding. In the case of a null step, the tolerances ε_l and η_l decrease since the point u^k approximates a solution within ε_l, η_l. Hence, the variable l is a counter for null steps. In the case of a descent step we must have $\omega_k > 0$, hence, the point $\tilde{x}^{k+1} = x^k - \gamma(\omega_k/\|g^k\|^2)g^k$ is the projection of the point x^k onto the hyperplane $H_k(\gamma)$, where $H_k(1)$ separates x^k and X^d (see (5) and (11)). Thus, our method follows the general CR framework.

We will call one increase of the index i an inner step, so that the number of inner steps gives the number of computations of elements from $Q(\cdot)$ at the corresponding points.

Theorem 6. *(see e.g. [29, Theorem 2.3.2]) Let a sequence $\{u^k\}$ be generated by Method 2.1 and let $\{\varepsilon_l\}$ and $\{\eta_l\}$ satisfy the following relations:*

$$\{\varepsilon_l\} \searrow 0, \{\eta_l\} \searrow 0. \tag{25}$$

Then:

(i) The number of inner steps at each iteration is finite.

(ii) There exists a limit point x^* of $\{x^k\}$ which lies in X^*.

(iii) If

$$X^* = X^d, \tag{26}$$

we have

$$\lim_{k\to\infty} x^k = x^* \in X^*.$$

As Method 2.1 has a two-level structure, each iteration containing a finite number of inner steps, it is more suitable to derive its complexity estimate, which gives the total amount of work of the method, instead of convergence rates. We use the distance to x^* as an accuracy function for our method, i.e., our approach is slightly different from the standard ones. More precisely, given a starting point x^0 and a number $\delta > 0$, we define the complexity of the method, denoted by $N(\delta)$, as the total number of inner steps t which ensures finding a point $\bar{x} \in X$ such that

$$\|\bar{x} - x^*\|/\|x^0 - x^*\| \leq \delta.$$

Therefore, since the computational expense per inner step can easily be evaluated for each specific problem under examination, this estimate in fact gives the total amount of work. We thus proceed to obtain an upper bound for $N(\delta)$.

Theorem 7. *[29, Theorem 2.3.3] Suppose G is monotone and there exists $x^* \in X^*$ such that*

$$\text{for every } x \in X \text{ and for every } g \in G(x),$$
$$\langle g, x - x^*\rangle \geq \mu\|x - x^*\|,$$

for some $\mu > 0$. *Let a sequence* $\{x^k\}$ *be generated by Method 2.1 where*

$$\varepsilon_l = \nu^l \varepsilon', \eta_l = \eta', l = 0, 1, \dots; \quad \nu \in (0, 1).$$

Then, there exist some constants $\bar{\varepsilon} > 0$ *and* $\bar{\eta} > 0$ *such that*

$$N(\delta) \le B_1 \nu^{-2} (\ln(B_0/\delta)/\ln \nu^{-1} + 1),$$

where $0 < B_0, B_1 < \infty$, *whenever* $0 < \varepsilon' \le \bar{\varepsilon}$ *and* $0 < \eta' \le \bar{\eta}$, B_0 *and* B_1 *being independent of* ν.

It should be noted that the assertion of Theorem 7 remains valid without the additional monotonicity assumption on G if $X = \mathbb{R}^n$ (cf. (21)). Thus, our method attains a logarithmic complexity estimate, which corresponds to a linear rate of convergence with respect to inner steps. We now give a similar upper bound for $N(\delta)$ in the single-valued case.

Theorem 8. *[29, Theorem 2.3.4] Suppose that* $X = \mathbb{R}^n$ *and that* G *is strongly monotone and Lipschitz continuous. Let a sequence* $\{x^k\}$ *be generated by Method 2.1 where*

$$\varepsilon_l = \nu^l \varepsilon', \eta_l = \nu^l \eta', l = 0, 1, \dots; \varepsilon' > 0, \eta' > 0; \quad \nu \in (0, 1).$$

Then,

$$N(\delta) \le B_1 \nu^{-6} (\ln(B_0/\delta)/\ln \nu^{-1} + 1),$$

where $0 < B_0, B_1 < \infty$, B_0 *and* B_1 *being independent of* ν.

3.3 CR Method for Multi-valued Inclusions

To solve GVI (22), we can also apply Method 2.1 for finding stationary points of the mapping P being defined as follows:

$$P(x) = \begin{cases} G(x) & \text{if } h(x) < 0, \\ \operatorname{conv}\{G(x) \bigcup \partial h(x)\} & \text{if } h(x) = 0, \\ \partial h(x) & \text{if } h(x) > 0. \end{cases} \tag{27}$$

Such a method does not include pseudo-projections and is based on the following observations; see [20, 25, 29].

First we note P in (27) is a K-mapping. Next, GVI (22) is equivalent to the multi-valued inclusion

$$0 \in P(x^*). \tag{28}$$

We denote by S^* the solution set of problem (28).

Theorem 9. *[29, Theorem 2.3.1] It holds that*

$$X^* = S^*.$$

In order to apply Method 2.1 to problem (28) we have to show that its dual problem is solvable. Namely, let us consider the problem of finding a point x^* of $\mathbb{R}^n$ such that

$$\forall u \in \mathbb{R}^n, \quad \forall t \in P(u), \quad \langle t, u - u^* \rangle \geq 0,$$

which can be viewed as the dual problem to (28). We denote by $S^*_{(d)}$ the solution set of this problem. Clearly, Proposition 2 admits the corresponding simple specialization.

Lemma 5. *(i)* $S^*_{(d)}$ *is convex and closed.*
(ii) $S^*_{(d)} \subseteq S^*$.
(iii) If P *is pseudomonotone, then* $S^*_{(d)} = S^*$.

Note that P need not be pseudomonotone in general. Nevertheless, in addition to Theorem 9, it is useful to obtain the equivalence result for both the dual problems.

Proposition 3. *[29, Proposition 2.4.1]* $X^d = S^*_{(d)}$.

Combining the above results and Proposition 2 yields a somewhat strengthened equivalence property.

Corollary 1. *If* G *is pseudomonotone, then*

$$X^* = X^d = S^*_{(d)} = S^*.$$

Therefore, we can apply Method 2.1 with replacing G, X, and P_k by P, $\mathbb{R}^n$, and I, respectively, to the multi-valued inclusion (28) under the same blanket assumptions. We call this modification Method 2.2.

Theorem 10. *Let a sequence* $\{x^k\}$ *be generated by Method 2.2 and let* $\{\varepsilon_l\}$ *and* $\{\eta_l\}$ *satisfy (25). Then:*
(i) The number of inner steps at each iteration is finite.
(ii) There exists a limit point x^* *of* $\{x^k\}$ *which lies in* X^*.
(iii) If (26) holds, we have

$$\lim_{k \to \infty} x^k = x^* \in S^* = X^*.$$

Next, the simplest rule (24) in Method 2.1 can be replaced by one of the following:

$$p^{i+1} = \text{Nr conv}\{q^0, \dots, q^{i+1}\},$$

or

$$p^{i+1} = \text{Nr conv}\{p^i, q^{i+1}, S_i\},$$

where $S_i \subseteq \text{conv}\{q^0, \dots, q^i\}$. These modifications may be used for attaining more rapid convergence, and all the assertions of this section remain true. Nevertheless, they require additional storage and computational expenses.

4 Some Examples of Generalized Monotone Problems

Various applications of variational inequalities have been well documented in the literature; see e.g. [36, 29, 9] and references therein. We intend now to give some additional examples of problems which reduce to VI (6) with satisfying the basic property $X^d \neq \emptyset$. It means that they possess certain generalized monotonicity properties. We restrict ourselves with single-valued problems by assuming usually differentiability of functions. Nevertheless, using a suitable concept of the subdifferential, we can obtain similar results for the case of multi-valued GVI (22).

4.1 Scalar Optimization Problems

We start our illustrations from the simplest optimization problems.

Let us consider the problem of minimizing a function $f : \mathbb{R}^n \to \mathbb{R}$ over the convex and closed set X, or briefly,

$$\min_{x \in X} \to f(x). \tag{29}$$

If f is also differentiable, we can replace (29) by its optimality condition in the form of VI: Find $x^* \in X$ such that

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad \forall x \in X \tag{30}$$

(cf. (6)). The problem is to find conditions which ensure solvability of DVI: Find $x^* \in X$ such that

$$\langle \nabla f(x), x - x^* \rangle \geq 0 \quad \forall x \in X \tag{31}$$

(cf. (7)). It is known that each solution of (31), unlike that of (30), also solves (29); see [14, Theorem 2.2]. Denote by X_f the solution set of problem (29) and suppose that $X_f \neq \emptyset$. We can obtain the solvability of (31) under a rather weak condition on the function f. Recall that $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *quasiconvex* on X, if for any points $x, y \in X$ and for each $\lambda \in [0, 1]$ it holds that

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}.$$

Also, $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *quasiconvex along rays with respect to* X if for any point $x \in X$ we have

$$f(\lambda x + (1 - \lambda)x^*) \leq f(x) \quad \forall \lambda \in [0, 1], \quad \forall x^* \in X_f;$$

see [20]. Clearly, the class of quasiconvex along rays functions strictly contains the class of usual quasiconvex functions since the level sets $\{x \in X \mid f(x) \leq \mu\}$ of a quasiconvex along rays function f may be non-convex.

Proposition 4. *If $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex along rays with respect to X, then the solution set of (31) coincides with X_f.*

Proof. Due to the above observation, we have to show that any solution $x^* \in X_f$ solves (31). Fix $x \in X$ and set $s = x^* - x$. Then we have

$$\begin{aligned}\langle \nabla f(x), s \rangle &= \lim_{\alpha \to 0} \frac{f(x + \alpha s) - f(x)}{\alpha} \\ &= \lim_{\alpha \to 0} \frac{f(\alpha x^* + (1-\alpha)x) - f(x)}{\alpha} \leq 0,\end{aligned}$$

i.e. x^* solves (31) and the result follows.

So, the condition $X^d \neq \emptyset$ then holds.

4.2 Walrasian Price Equilibrium Models

Walrasian equilibrium models describe economies with perfect competition. The economy deals in n commodities and, given a price vector $p = (p_1, \ldots, p_n)$, the demand and supply are supposed to be determined as vectors $D(p)$ and $S(p)$, respectively, and the vector

$$E(p) = D(p) - S(p)$$

represents the excess demand. Then the equilibrium price vector p^* is defined by the following complementarity conditions

$$p^* \in \mathbb{R}^n_+, -E(p^*) \in \mathbb{R}^n_+, \langle p^*, E(p^*) \rangle = 0;$$

which can be equivalently rewritten as VI: Find $p^* \in \mathbb{R}^n_+$ such that

$$\langle -E(p^*), p - p^* \rangle \geq 0 \quad \forall p \in \mathbb{R}^n_+; \tag{32}$$

see e.g. [2, 37]. Here $\mathbb{R}^n_+ = \{p \in \mathbb{R}^n \mid p_i \geq 0 \quad i = 1, \ldots, n\}$ denotes the set of vectors with non-negative components. The properties of E depend on behaviour of consumers and producers, nevertheless, gross substitutability and positive homogeneity are among the most popular. Recall that a mapping $F : P \to \mathbb{R}^n$ is said to be

(i) *gross substitutable*, if for each pair of points $p', p'' \in P$ such that $p' - p'' \in \mathbb{R}^n_+$ and $I(p', p'') = \{i \mid p'_i = p''_i\}$ is nonempty, there exists an index $k \in I(p'_i, p''_i)$ with $F_k(p') \geq F_k(p'')$;

(ii) *positive homogeneous of degree m*, if for each $p \in P$ and for each $\lambda > 0$ such that $\lambda p \in P$ it holds that $F(\lambda p) = \lambda^m F(p)$.

It was shown by K.J. Arrow and L. Hurwicz [3] that these properties lead to a kind of the revealed preference condition. Denote by P^* the set of equilibrium prices.

Proposition 5. *Suppose that E : $\mathrm{int}\mathbb{R}^n_+ \to \mathbb{R}^n$ is gross substitutable, positively homogeneous with degree* 0*, and satisfies the Walras law, i.e.*

$$\langle p, E(p) \rangle = 0 \quad \forall p \in \mathrm{int}\mathbb{R}^n_+;$$

moreover, each function $E_i : \mathrm{int}\mathbb{R}^n_+ \to \mathbb{R}$ *is bounded below, and for every sequence* $\{p^k\} \subset \mathrm{int}\mathbb{R}^n_+$ *converging to* p*, it holds that*

$$\lim_{k\to\infty} E_i(p^k) = \begin{cases} E_i(p) & \text{if} \quad E_i(p) \quad \text{is finite,} \\ +\infty & \text{otherwise.} \end{cases}$$

Then problem (32) is solvable, and

$$\langle p^*, E(p)\rangle > 0 \quad \forall p \in \mathrm{int}\mathbb{R}^n_+ \backslash P^*, \forall p^* \in P^*.$$

Observe that $P^* \subseteq \mathrm{int}\mathbb{R}^n_+$ due to the above conditions, i.e. $E(p^*) = \mathbf{0}$ for each $p^* \in P^*$. It follows that

$$\langle -E(p), p - p^*\rangle \begin{cases} > 0 & \forall p \in \mathrm{int}\mathbb{R}^n_+ \backslash P^*, \\ \geq 0 & \forall p \in P^* \end{cases}$$

for each $p^* \in P^*$, therefore condition $X^d \neq \emptyset$ holds true for VI (32). Similar results can be obtained in the multi-valued case; see [39].

4.3 General Equilibrium Problems

Let $\Phi : X \times X \to \mathbb{R}$ be an equilibrium bifunction, i.e. $\Phi(x, x) = 0$ for each $x \in X$, and let X be a nonempty convex and closed subset of $\mathbb{R}^n$. Then we can consider the general equilibrium problem (EP for short): Find $x^* \in X$ such that

$$\Phi(x^*, y) \geq 0 \quad \forall y \in X. \tag{33}$$

We denote by X^e the solution set of this problem. It was first used by H. Nikaido and K. Isoda [38] for investigation of non-cooperative games and appeared very useful for other problems in nonlinear analysis; see [4, 11] for more details. If $\Phi(x, \cdot)$ is differentiable for each $x \in X$, we can consider also VI (6) with the cost mapping

$$G(x) = \nabla_y \Phi(x, y)|_{y=x}, \tag{34}$$

suggested by J.B. Rosen [41]. Recall that a function $f : X \to \mathbb{R}$ is said to be

(i) *pseudoconvex*, if for any points $x, y \in X$, it holds that

$$\langle \nabla f(x), y - x\rangle \geq 0 \Longrightarrow f(y) \geq f(x);$$

(ii) *explicitly quasiconvex*, if it is quasiconvex and for any point $x, y \in X$, $x \neq y$ and for all $\lambda \in (0, 1)$ it holds that

$$f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\}.$$

Then we can obtain the obvious relationships between solution sets of EP (33) and VI (6), (34).

Proposition 6. *(i) If $\Phi(x, \cdot)$ is differentiable for each $x \in X$, then $X^e \subseteq X^*$.*

(ii) If $\Phi(x, \cdot)$ is pseudoconvex for each $x \in X$, then $X^* \subseteq X^e$.

However, we are interested in revealing conditions providing the property $X^d \neq \emptyset$ for VI (6), (34). Let us consider the dual equilibrium problem: Find $y^* \in X$ such that

$$\Phi(x, y^*) \leq 0 \quad \forall x \in X \tag{35}$$

and denote by X^e_d the solution set of this problem. Recall that $\Phi : X \times X \to \mathbb{R}$ is said to be

(i) *monotone*, if for each pair of points $x, y \in X$ it holds that

$$\Phi(x, y) + \Phi(y, x) \leq 0;$$

(ii) *pseudomonotone*, if for each pair of points $x, y \in X$ it holds that

$$\Phi(x, y) \geq 0 \Longrightarrow \Phi(y, x) \leq 0.$$

Proposition 7. *(see [29, Proposition 2.1.17]) Let $\Phi(x, \cdot)$ be convex and differentiable for each $x \in X$. If Φ is monotone (respectively, pseudomonotone), then so is G in (34).*

Being based on this property, we can derive the condition $X^d \neq \emptyset$ from (pseudo)monotonicity of Φ and Proposition 1. However, it can be deduced from the existence of solutions of problem (35). We recall the Minty Lemma for EPs; see e.g. [4, Section 10.1] and [6].

Proposition 8. *(i) If $\Phi(\cdot, y)$ is upper semicontinuous for each $y \in X$, $\Phi(x, \cdot)$ is explicitly quasiconvex for $x \in X$, then $X^e_d \subseteq X^e$.*

(ii) If Φ is pseudomonotone, then $X^e \subseteq X^e_d$.

Now we give the basic relation between the solution sets of dual problems.

Lemma 6. *Suppose that $\Phi(x, \cdot)$ is quasiconvex and differentiable for each $x \in X$. Then $X^e_d \subseteq X^d$.*

Proof. Take any $x^* \in X^e_d$, then $\Phi(x, x^*) \leq \Phi(x, x) = 0$ for each $x \in X$. Set $\psi(y) = \Phi(x, y)$, then

$$\langle \nabla \psi(x), x^* - x \rangle = \lim_{\alpha \to 0} \frac{\psi(x + \alpha(x^* - x)) - \psi(x)}{\alpha} \leq 0,$$

i.e. $x^* \in X^d$.

Combining these properties, we can obtain relationships among all the problems.

Theorem 11. *Suppose that $\Phi : X \times X \to \mathbb{R}$ is a continuous equilibrium bifunction, $\Phi(x, \cdot)$ is quasiconvex and differentiable for each $x \in X$.*

(i) If holds that $X^e_d \subseteq X^d \subseteq X^*$.

(ii) If $\Phi(x, \cdot)$ is pseudoconvex for each $x \in X$, then

$$X^e_d \subseteq X^d \subseteq X^* = X^e.$$

(iii) If $\Phi(x, \cdot)$ is pseudoconvex for each $x \in X$ and Φ is pseudomonotone, then

$$X^e_d = X^d = X^* = X^e.$$

Proof. Part (i) follows from Lemma 6 and Proposition 1 (ii). Part (ii) follows from (i) and Proposition 6, and, taking into account Proposition 8 (ii), we obtain assertion (iii).

Therefore, we can choose the most suitable condition for its verification.

4.4 Optimization with Intransitive Preference

Optimization problems with respect to preference relations play the central role in decision making theory and in consumer theory. It is well-known that the case of transitive preferences lead to the usual scalar optimization problems and such problem have been investigated rather well, but the intransitive case seems more natural in modelling real systems; see e.g. [10, 44, 46].

Let us consider an optimization problem on the same feasible set X with respect to a binary relation (preference) R, which is not transitive in general, i.e. the implication

$$xRy \text{ and } yRz \Longrightarrow xRz$$

may not hold. Suppose that R is complete, i.e. for any points $x, y \in \mathbb{R}^n$ at least one of the relations holds: xRy, yRx. Then we can define the optimization problem with respect to R: Find $x^* \in X$ such that

$$x^* Ry \quad \forall y \in X. \tag{36}$$

Recall that the strict part P of R is defined as follows:

$$xPy \Longleftrightarrow (xRy \text{ and } \neg(yRx)).$$

Due to the completeness of R, it follows that

$$\neg(yRx) \Longrightarrow xPy,$$

and (36) becomes equivalent to the more usual formulation: Find $x^* \in X$ such that

$$\nexists y \in X, \ yPx^*. \tag{37}$$

Following [46, 42], consider a representation of the preference R by a bifunction $\Phi : X \times X \to \mathbb{R}$:

$$\begin{cases} x'Rx'' \Longleftrightarrow \Phi(x'',x') \leq 0, \\ x'Px'' \Longleftrightarrow \Phi(x'',x') < 0. \end{cases}$$

Note that the bifunction $\Psi(x',x'') = -\Phi(x'',x')$ gives a more standard representation, but the current form is more suitable for the common equilibrium setting. In fact, (37) becomes equivalent to EP (33), whereas (36) becomes equivalent to the dual problem (35).

We now consider generalized monotonicity of Φ.

Proposition 9. *For each pair of points $x', x'' \in X$ it holds that*

$$\begin{array}{l} \Phi(x',x'') > 0 \Longleftrightarrow \Phi(x'',x') < 0, \\ \Phi(x',x'') = 0 \Longleftrightarrow \Phi(x'',x') = 0. \end{array} \tag{38}$$

Proof. Fix $x', x'' \in X$. If $\Phi(x',x'') > 0$, then $\neg(x''Rx')$ and $x'Px''$, i.e. $\Phi(x'',x') < 0$, by definition. The reverse implication $\Phi(x',x'') < 0 \Longrightarrow \Phi(x'',x') > 0$ follows from the definition of P. It means that the first equivalence in (38) is true, moreover, we have

$$\Phi(x',x'') \leq 0 \Longrightarrow \Phi(x'',x') \geq 0.$$

Hence, $\Phi(x',x'') = 0$ implies $\Phi(x'',x') \geq 0$, but $\Phi(x'',x') > 0$ implies $\Phi(x',x'') < 0$, a contradiction. Thus, $\Phi(x',x'') = 0 \Longleftrightarrow \Phi(x'',x') = 0$, and the proof is complete.

Observe that (38) implies

$$\Phi(x,x) = 0 \quad \forall x \in X,$$

i.e. Φ is an equilibrium bifunction and R is reflexive. Next, on account of Proposition 9, both Φ and $-\Phi$ are pseudomonotone, which yields the equivalence of (33) and (35) because of Proposition 8 (ii). The relations in (38) hold if Φ is skew-symmetric, i.e.

$$\Phi(x',x'') + \Phi(x'',x') = 0 \quad \forall x', x'' \in X;$$

cf. Example 1.

In order to find a solution of problem (36) (or (37)), we have to impose additional conditions on Φ; see [20] for details. Namely, suppose that Φ is continuous and that $\Phi(x,\cdot)$ is quasiconvex for each $x \in X$. Then R is continuous and also convex, i.e. for any points $x', x'', y \in X$, we have

$$x'Ry \text{ and } x''Ry \Longrightarrow [\lambda x' + (1-\lambda)x'']Ry \quad \forall \lambda \in [0,1].$$

If Φ is skew-symmetric, it follows that $\Phi(\cdot,y)$ is quasiconcave for each $y \in X$, and there exists a CR method for finding a solution of such EPs; see [19]. However, this is not the case in general, but then we can solve EP via its reducing to VI, as described in Section 4.3. In fact, if $\Phi(x,\cdot)$ is differentiable, then (36) (or (37)) implies VI (6), (34) and DVI (7), (34), i.e., the basic condition $X^d \neq \emptyset$ holds true if the initial problem is solvable, as Theorem 11 states. Then the CR methods described are also applicable for finding its solution.

4.5 Quasi-concave-convex Zero-sum Games

Let us consider a zero-sum game with two players. The first player has the strategy set X and the utility function $\Phi(x, y)$, whereas the second player has the utility function $-\Phi(x, y)$ and the strategy set Y. Following [5, Section 10.4], we say that the game is *equal* if $X = Y$ and $\Phi(x, x) = 0$ for each $x \in X$. If Φ is continuous, $\Phi(\cdot, y)$ is quasiconcave for each $y \in X, \Phi(x, \cdot)$ is quasiconvex for each $x \in X$, and X is a nonempty, convex and closed set, then this equal game will have a saddle point $(x^*, y^*) \in X \times X$, i.e.

$$\Phi(x, y^*) \leq \Phi(x^*, y^*) \leq \Phi(x^*, y) \quad \forall x \in X, \forall y \in X$$

under the boundedness of X because of the known Sion minimax theorem [45]. Moreover, its value is zero, since

$$0 = \Phi(y^*, y^*) \leq \Phi(x^*, y^*) \leq \Phi(x^*, x^*) = 0.$$

Thus, the set of optimal strategies of the first player coincides with the solution set X^e of EP (33), whereas the set of optimal strategies of the second player coincides with X^e_d, which is the solution set of the dual EP (35). Unlike the previous sections, Φ may not possess generalized monotonicity properties. A general CR method for such problems was proposed in [19]. Nevertheless, if $\Phi(x, \cdot)$ is differentiable, then Theorem 11 (i) gives $X^e_d \subseteq X^d \subseteq X^*$, where X^d (respectively, X^*) is the solution set of DVI (7), (34) (respectively, VI (6), (34), i.e. existence of saddle points implies $X^d \neq \emptyset$. However, by strengthening slightly the quasi-concavity-convexity assumptions, we can obtain additional properties of solutions. In fact, replace the quasiconcavity (quasiconvexity) of $\Phi(x, y)$ in x (in y) by the explicit quasiconcavity (quasiconvexity), respectively. Then Proposition 8 (i) yields $X^e = X^e_d$, i.e., the players have the same solution sets. Hence, $X^e \neq \emptyset$ implies $X^d \neq \emptyset$ and this result strengthens a similar property in [47, Theorem 5.3.1].

Proposition 10. *If the utility function $\Phi(x, y)$ in an equal game is continuous, explicitly quasiconcave in x, explicitly quasiconvex and differentiable in y, then*

$$X^e = X^e_d \subseteq X^d \subseteq X^*.$$

If, additionally, $\Phi(x, y)$ is pseudoconvex in y, then

$$X^e = X^e_d = X^d = X^*.$$

Proof. The first assertion follows from Theorem 11 (i) and Proposition 8 (i). The second assertion now follows from Theorem 11 (ii).

This general equivalence result does not use pseudomonotonicity of Φ or G, nevertheless, it also enables us to develop efficient methods for finding optimal strategies.

Therefore, many optimization and equilibrium type problems possess required generalized monotonicity properties.

Further Investigations

The CR methods presented can be extended and modified in several directions. In particular, they can be applied to extended VIs involving additional mappings (see [30, 31]) and to mixed VIs involving non-linear convex functions (see [29, 31, 33]).

It was mentioned that the CR framework is rather flexible and admits specializations for each particular class of problems. Such versions of CR methods were proposed for various decomposable VIs (see [27, 28, 29, 32]). In this context, CR methods with auxiliary procedures based on an iteration of a suitable splitting method seem very promising (see [26, 29, 31, 33]).

References

1. Agmon S (1954): The relaxation method for linear inequalities. Canadian Journal of Mathematics 6: 382–392
2. Arrow KJ, Hahn FH (1972): General competitive analysis. Holden Day, San Francisco
3. Arrow KJ, Hurwicz L (1960): Competitive stability under weak gross substitutability: The “Euclidean distance” approach, International Economic Review 1: 38–49
4. Baiocchi C, Capelo A (1984): Variational and quasivariational inequalities. Applications to free boundary problems. John Wiley and Sons, New York
5. Belen’kii VZ, Volkonskii VA (eds) (1974): Iterative methods in game theory and programming. Nauka, Moscow (in Russian)
6. Bianchi M, Schaible S (1996): Generalized monotone bifunctions and equilibrium problems. Journal of Optimization Theory and Applications 90: 31–43
7. Cimmino G (1938): Calcolo approsimato per le soluzioni dei sistemi de equazioni lineari. La Ricerca Scientifica XVI **1**, 326–333
8. Eremin II (1965): The relaxation method of solving systems of inequalities with convex functions on the left-hand sides. Soviet Mathematics Doklady 6: 219–222
9. Facchinei F, Pang JS (2003): Finite-dimensional variational inequalities and complementarity problems. Springer-Verlag, Berlin
10. Fishburn PC (1970): Utility theory for decision making. Wiley, New York
11. Hadjisavvas N, Komlósi S, Schaible S (eds) (2005): Handbook of generalized convexity and generalized monotonicity. “Nonconvex optimization and applications”, Vol.76. Springer, New York
12. Kaczmarz S (1937): Angenäherte Auflösung von Systemen linearer Gleichungen. Bulletin Internationel de l’Académie Polonaise des Sciences et des Lettres. Classe des Sciences Mathématiques et Naturelles. Séries A 60: 596–599
13. Karamardian S (1976): Complementarity over cones with monotone and pseudomonotone maps. Journal of Optimization Theory and Applications 18: 445–454
14. Komlósi S (1999): On the Stampacchia and Minty variational inequalities. In: Giorgi G, Rossi F (eds) Generalized convexity and optimization in economic and financial decisions. Pitagora Editrice, Bologna, 231–260
15. Konnov IV (1982): A subgradient method of successive relaxation for solving optimization problems. Preprint VINITI No. 531-83, Kazan University, Kazan (in Russian)

16. Konnov IV (1984): A method of the conjugate subgradient type for minimization of functionals. Issledovaniya po Prikladnoi Matematike 12: 59–62 (Engl. transl. in Journal of Soviet Mathematics 45: 1026–1029, 1989).
17. Konnov IV (1993): Combined relaxation methods for finding equilibrium points and solving related problems, Russian Mathematics (Iz. VUZ) 37 (2): 44–51
18. Konnov IV (1993): On combined relaxation method's convergence rates. Russian Mathematics (Iz. VUZ) 37(12): 89–92
19. Konnov IV (1994): Applications of the combined relaxation method to finding equilibrium points of a quasi-convex-concave function. Russian Mathematics (Iz. VUZ) 38 (12): 66–71
20. Konnov IV (1996): A general approach to finding stationary points and the solution of related problems. Computational Mathematics and Mathematical Physics 36: 585–593
21. Konnov IV (1997): A class of combined iterative methods for solving variational inequalities. Journal of Optimization Theory and Applications **94**, 677–693
22. Konnov IV (1997): Combined relaxation methods having a linear rate of convergence. News-Letter of the Mathematical Programming Association 7: 135–137 (in Russian)
23. Konnov IV (1998): A combined relaxation method for variational inequalities with nonlinear constraints. Mathematical Programming 80: 239–252
24. Konnov IV (1998): On the convergence of combined relaxation methods for variational inequalities. Optimization Methods and Software 9: 77–92
25. Konnov IV (1998): An inexact combined relaxation method for multivalued inclusions. Russian Mathematics (Iz. VUZ) 42 (12): 55–59
26. Konnov IV (1999): A combined method for variational inequalities with monotone operators. Computational Mathematics and Mathematical Physics 39:1051–1056
27. Konnov IV (1999): Combined relaxation methods for variational inequality problems over product sets. Lobachevskii Journal of Mathematics 2: 3–9
28. Konnov IV (1999): Combined relaxation method for decomposable variational inequalities. Optimization Methods and Software 10: 711–728
29. Konnov IV (2001): Combined relaxation methods for variational inequalities. Springer-Verlag, Berlin
30. Konnov IV (2001): Combined relaxation method for extended variational inequalities. Russian Mathematics (Iz. VUZ) 45 (12): 46–54
31. Konnov IV (2002): A combined relaxation method for nonlinear variational inequalities. Optimization Methods and Software 17: 271–292
32. Konnov IV (2002): A class of combined relaxation methods for decomposable variational inequalities. Optimization 51: 109–125
33. Konnov IV (2002): A combined relaxation method for a class of nonlinear variational inequalities. Optimization 51: 127–143
34. Minty G (1967): On the generalization of a direct method of the calculus of variations. Bulletin of the American Mathematical Society 73: 315–321
35. Motzkin TS, Schoenberg IJ (1954): The relaxation method for linear inequalities. Canadian Journal of Mathematics 6: 393–404
36. Nagurney A (1999): Network economics: a variational inequality approach. Kluwer Academic Publishers, Dordrecht
37. Nikaido H (1968): Convex structures and economic theory. Academic Press, New York

38. Nikaido H, Isoda K (1955): Note on noncooperative convex games. Pacific Journal of Mathematics 5: 807–815
39. Polterowich VM, Spivak VA (1983): Gross substitutability of point-to-set correspondences. Journal of Mathematical Economics 11: 117–140
40. Polyak BT (1969): Minimization of unsmooth functionals. USSR Computational Mathematics and Mathematical Physics 9: 14–29
41. Rosen JB (1965): Existence and uniqueness of equilibrium points for concave n-person games. Econometrica 33: 520–534
42. Shafer W (1974): The non-transitive consumer. Econometrica 42: 913–919
43. Shih MH, Tan KK (1988): Browder-Hartmann-Stampacchia variational inequalities for multivalued monotone operators. Journal of Mathematical Analysis and Applications 134: 431–440
44. Sen AK (1970): Collective choice and social welfare. Holden Day, San Francisco
45. Sion M (1958): On general minimax theorems. Pacific Journal of Mathematics 8: 171–176
46. Tversky A (1969): Intransitivity of preferences. Psychological Review 76: 31–48
47. Uryas'ev SP (1990): Adaptive algorithms of stochastic optimization and game theory. Nauka, Moscow (in Russian)
48. Xiu N, Zhang J (2003): Some recent advances in projection-type methods for variational inequalities. Journal of Computational and Applied Mathematics 152: 559–585
49. Yao JC (1994): Multivalued variational inequalities with K - pseudomonotone operators. Journal of Optimization Theory and Applications 83: 391–403

Abstract Convexity and the Monge–Kantorovich Duality

Vladimir L. Levin *

Central Economics and Mathematics Institute of the Russian Academy of Sciences, 47 Nakhimovskii Prospect, 117418 Moscow, Russia
vl_levin@cemi.rssi.ru

Summary. In the present survey, we reveal links between abstract convex analysis and two variants of the Monge–Kantorovich problem (MKP), with given marginals and with a given marginal difference. It includes: (1) the equivalence of the validity of duality theorems for MKP and appropriate abstract convexity of the corresponding cost functions; (2) a characterization of a (maximal) abstract cyclic monotone map $F : X \to L \subset \mathbb{R}^X$ in terms connected with the constraint set

$$Q_0(\varphi) := \{u \in \mathbb{R}^Z : u(z_1) - u(z_2) \le \varphi(z_1, z_2) \quad \forall z_1, z_2 \in Z = \operatorname{dom} F\}$$

of a particular dual MKP with a given marginal difference and in terms of L-subdifferentials of L-convex functions; (3) optimality criteria for MKP (and Monge problems) in terms of abstract cyclic monotonicity and non-emptiness of the constraint set $Q_0(\varphi)$, where φ is a special cost function on $X \times X$ determined by the original cost function c on $X \times Y$. The Monge–Kantorovich duality is applied then to several problems of mathematical economics relating to utility theory, demand analysis, generalized dynamics optimization models, and economics of corruption, as well as to a best approximation problem.

Key words: H-convex function, infinite linear programs, duality relations, Monge-Kantorovich problems (MKP) with given marginals, MKP with a given marginal difference, abstract cyclic monotonicity, Monge problem, utility theory, demand analysis, dynamics models, economics of corruption, approximation theory

1 Introduction

Abstract convexity or convexity without linearity may be defined as a theory which deals with applying methods of convex analysis to non-convex objects.

*Supported in part by the Russian Leading Scientific School Support Grant NSh-6417.2006.6.

Today this theory becomes an important fragment of non-linear functional analysis, and it has numerous applications in such different fields as non-convex global optimization, various non-traditional duality schemes for particular classes of sets and functions, non-smooth analysis, mass transportation problems, mathematical economics, approximation theory, and measure theory; for history and references, see, e.g., [15], [30], [41], [43], [53], [54] [59], [60], [62]...[2]

In this survey, we'll dwell on connections between abstract convexity and the Monge—Kantorovich mass transportation problems; some applications to mathematical economics and approximation theory will be considered as well.

Let us recall some basic notions relating to abstract convexity. Given a nonempty set Ω and a class H of real-valued functions on it, the *H-convex envelope* of a function $f : \Omega \to \mathbb{R} \cup \{+\infty\}$ is defined to be the function $co_H(f)(\omega) := \sup\{h(\omega) : h \in H(f)\}$, $\omega \in \Omega$, where $H(f)$ comprises functions in H majorized by f, $H(f) := \{h \in H : h \leq f\}$. Clearly, $H(f) = H(co_H(f))$. A function f is called *H-convex* if $f = co_H(f)$.

In what follows, we take $\Omega = X \times Y$ or $\Omega = X \times X$, where X and Y are compact topological spaces, and we deal with H being a convex cone or a linear subspace in $C(\Omega)$. The basic examples are $H = \{h_{uv} : h_{uv}(x,y) = u(x) - v(y), (u,v) \in C(X) \times C(Y)\}$ for $\Omega = X \times Y$ and $H = \{h_u : h_u(x,y) = u(x) - u(y), u \in C(X)\}$ for $\Omega = X \times X$. These examples are closely connected with two variants of the Monge—Kantorovich problem (MKP): with given marginals and with a given marginal difference.

Given a cost function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ and finite positive regular Borel measures, σ_1 on X and σ_2 on Y, $\sigma_1 X = \sigma_2 Y$, the MKP with marginals σ_1 and σ_2 is to minimize the integral

$$\int_{X \times Y} c(x,y)\, \mu(d(x,y))$$

subject to constraints: $\mu \in C(X \times Y)^*_+$, $\pi_1\mu = \sigma_1$, $\pi_2\mu = \sigma_2$, where $\pi_1\mu$ and $\pi_2\mu$ stand for the marginal measures of μ.[3]

A different variant of MKP, the MKP with a given marginal difference, relates to the case $X = Y$ and consists in minimizing the integral

$$\int_{X \times X} c(x,y)\, \mu(d(x,y))$$

subject to constraints: $\mu \in C(X \times X)^*_+$, $\pi_1\mu - \pi_2\mu = \sigma_1 - \sigma_2$.

Both variants of MKP were first posed and studied by Kantorovich [17, 18] (see also [19, 20, 21]) in the case where $X = Y$ is a metric compact space with

[2]Abstract convexity is, in turn, a part of a broader field known as generalized convexity and generalized monotonicity; see [14] and references therein.

[3]For any Borel sets $B_1 \subseteq X$, $B_2 \subseteq Y$, $(\pi_1\mu)(B_1) = \mu(B_1 \times Y)$, $(\pi_2\mu)(B_2) = \mu(X \times B_2)$.

its metric as the cost function c. In that case, both variants of MKP are equivalent but, in general, the equivalence fails to be true.

The MKP with given marginals is a relaxation of the Monge 'excavation and embankments' problem [49], a non-linear extremal problem, which is to minimize the integral

$$\int_X c(x, f(x))\, \sigma_1(dx)$$

over the set $\Phi(\sigma_1, \sigma_2)$ of measure-preserving Borel maps $f : (X, \sigma_1) \to (Y, \sigma_2)$. Of course, it can occur that $\Phi(\sigma_1, \sigma_2)$ is empty, but in many cases it is non-empty and the measure μ_f on $X \times Y$,

$$\mu_f B = \sigma_1\{x : (x, f(x)) \in B\}, \quad B \subset X \times Y,$$

is positive and has the marginals $\pi_1\mu_f = \sigma_1$, $\pi_2\mu_f = \sigma_2$. Moreover, if μ_f is an optimal solution to the MKP then f proves to be an optimal solution to the Monge problem.

Both variants of MKP may be treated as problems of infinite linear programming. The dual MKP problem with given marginals is to maximize

$$\int_X u(x)\sigma_1(dx) - \int_Y v(y)\sigma_2(dy)$$

over the set

$$Q'(c) := \{(u, v) \in C(X) \times C(Y) : \, u(x) - v(y) \le c(x, y) \quad \forall (x, y) \in X \times Y\},$$

and the dual MKP problem with a given marginal difference is to maximize

$$\int_X u(x)\, (\sigma_1 - \sigma_2)(dx)$$

over the set

$$Q(c) := \{u \in C(X) : \, u(x) - u(y) \le c(x, y) \quad \forall x, y \in X\}.$$

As is mentioned above, in the classical version of MKP studied by Kantorovich, X is a metric compact space and c is its metric. In that case, $Q(c)$ proves to be the set of Lipschitz continuous functions with the Lipschitz constant 1, and the Kantorovich optimality criterion says that a feasible measure μ is optimal if and only if there exists a function $u \in Q(c)$ such that $u(x) - u(y) = c(x, y)$ whenever the point (x, y) belongs to the support of μ. This criterion implies the duality theorem asserting the equality of optimal values of the original and the dual problems.

Duality for MKP with general continuous cost functions on (not necessarily metrizable) compact spaces is studied since 1974; see papers by Levin [24, 25, 26] and references therein. A general duality theory for arbitrary compact spaces and continuous or discontinuous cost functions was developed by Levin and Milyutin [47]. In that paper, the MKP with a given marginal difference is studied, and, among other results, a complete description of all cost functions, for which the duality relation holds true, is given. Further generalizations (non-compact and non-topological spaces) see [29, 32, 37, 38, 42].

An important role in study and applications of the Monge—Kantorovich duality is played by the set $Q(c)$ and its generalizations such as

$$Q(c; E(X)) := \{u \in E(X) : u(x) - u(y) \leq c(x,y) \quad \forall x, y \in X\},$$

where $E(X)$ is some class of real-valued functions on X. Typical examples are the classes: $\mathbb{R}^X$ of all real-valued functions on X, $l^\infty(X)$ of bounded real-valued functions on X, $U(X)$ of bounded universally measurable real-valued functions on X, and $\mathcal{L}^\infty(\mathbb{R}^n)$ of bounded Lebesgue measurable real-valued functions on $\mathbb{R}^n$ (Lebesgue equivalent functions are not identified).

Notice that the duality theorems and their applications can be restated in terms of abstract convexity of the corresponding cost functions. In that connection, mention an obvious equality $Q(c; E(X)) = H(c)$ where $H = \{h_u : u \in E(X)\}$. Conditions for $Q(c)$ or $Q_0(c) = Q(c; \mathbb{R}^Z)$ to be nonempty are some kinds of abstract cyclic monotonicity, and for specific cost functions c, they prove to be crucial in various applications of the Monge—Kantorovich duality. Also, optimality criteria for solutions to the MKP with given marginals and to the corresponding Monge problems can be given in terms of non-emptiness of $Q(\varphi)$ where φ is a particular function on $X \times X$ connected with the original cost function c on $X \times Y$.

The paper is organized as follows. Section 2 is devoted to connections between abstract convexity and infinite linear programming problems more general than MKP. In Section 3, both variants of MKP are regarded from the viewpoint of abstract convex analysis (duality theory; abstract cyclic monotonicity and optimality conditions for MKP with given marginals and for a Monge problem; further generalizations). In Section 4, applications to mathematical economics are presented, including utility theory, demand analysis, dynamics optimization, and economics of corruption. Finally, in Section 5 an application to approximation theory is given.

Our goal here is to clarify connections between the Monge - Kantorovich duality and abstract convex analysis rather than to present the corresponding duality results (and their applications) in maximally general form.

2 Abstract Convexity and Infinite Linear Programs

Suppose Ω is a compact Hausdorff topological space, and $c : \Omega \to \mathbb{R} \cup \{+\infty\}$ is a bounded from below universally measurable function on it. Given a convex cone $H \subset C(\Omega)$ such that $H(c) = \{h \in H : h \leq c\}$ is nonempty, and a measure $\mu_0 \in C(\Omega)^*_+$, we consider two infinite linear programs, the original one, I, and the dual one, II, as follows.

The original program is to maximize the linear functional $\langle h, \mu_0 \rangle := \int_\Omega h(\omega)\, \mu_0(d\omega)$ subject to constraints: $h \in H$, $h(\omega) \le c(\omega)$ for all $\omega \in \Omega$. The optimal value of this program will be denoted as $v_I(c; \mu_0)$.

The dual program is to minimize the integral functional

$$c(\mu) := \int_\Omega c(\omega)\, \mu(d\omega)$$

subject to constraints: $\mu \ge 0$ (i.e., $\mu \in C(\Omega)^*_+$) and $\mu \in \mu_0 - H^0$, where H^0 stands for the conjugate (polar) cone in $C(\Omega)^*_+$,

$$H^0 := \{\mu \in C(\Omega)^* : \langle h, \mu \rangle \le 0 \ \text{ for all } \ h \in H\}.$$

The optimal value of this program will be denoted as $v_{II}(c; \mu_0)$.

Thus, for any $\mu_0 \in C(\Omega)^*_+$, one has

$$v_I(c; \mu_0) = \sup\{\langle h, \mu_0 \rangle : h \in H(c)\}, \tag{1}$$

$$v_{II}(c; \mu_0) = \inf\{c(\mu) : \mu \ge 0, \mu \in \mu_0 - H^0\}. \tag{2}$$

In what follows, we endow $C(\Omega)^*$ with the weak* topology and consider $v_I(c; \cdot)$ and $v_{II}(c; \cdot)$ as functionals on the whole of $C(\Omega)^*$ by letting $v_I(c; \mu_0) = v_{II}(c; \mu_0) = +\infty$ for $\mu_0 \in C(\Omega)^* \setminus C(\Omega)^*_+$.

Clearly, both functionals are sublinear that is semi-additive and positive homogeneous. Furthermore, it is easily seen that the subdifferential of v_I at 0 is exactly the closure of $H(c)$,

$$\partial v_I(c; 0) = \operatorname{cl} H(c). \tag{3}$$

Note that

$$v_I(c; \mu_0) \le v_{II}(c; \mu_0). \tag{4}$$

Also, an easy calculation shows that the conjugate functional $v^*_{II}(c; u) := \sup\{\langle u, \mu_0 \rangle - v_{II}(c; \mu_0) : \mu_0 \in C(\Omega)^*\}$, $u \in C(\Omega)$, is the indicator function of $\operatorname{cl} H(c)$,

$$v^*_{II}(c; u) = \begin{cases} 0, & u \in \operatorname{cl} H(c); \\ +\infty, & u \notin \operatorname{cl} H(c); \end{cases} \tag{5}$$

therefore, the second conjugate functional $v^{**}_{II}(c; \mu_0) := \sup\{\langle u, \mu_0 \rangle - v^*_{II}(c; u) : u \in C(\Omega)\}$ is exactly $v_I(c; \mu_0)$,

$$v^{**}_{II}(c; \mu_0) = v_I(c; \mu_0),\ \mu_0 \in C(\Omega)^*. \tag{6}$$

As is known from convex analysis (e.g., see [47] where a more general duality scheme was used), the next result is a direct consequence of (6).

Proposition 1. *Given $\mu_0 \in \operatorname{dom} v_I(c; \cdot) := \{\mu \in C(\Omega)^*_+ : v_I(c; \mu) < +\infty\}$, the following assertions are equivalent:*

(a) $v_I(c; \mu_0) = v_{II}(c; \mu_0)$;

(b) the functional $v_{II}(c; \cdot)$ is weakly lower semi-continuous (lsc) at μ_0.*

Say c is *regular* if it is lsc on Ω and, for every $\mu_0 \in \operatorname{dom} v_I(c;\cdot)$,

$$v_{II}(c;\mu_0) = \inf\{c(\mu) : \mu \geq 0, \mu \in \mu_0 - H^0, \|\mu\| \leq M\|\mu_0\|\}, \tag{7}$$

where $M = M(c;H) > 0$. Note that if $\mu_0 \notin \operatorname{dom} v_I(c;\cdot)$ then, by (4), $v_{II}(c;\mu_0) = +\infty$; therefore, for such μ_0, (7) is trivial. Thus, for a regular c, (7) holds true for all $\mu_0 \in C(\Omega)^*$.

Proposition 2. *(i) If c is regular, then $v_{II}(c;\cdot)$ is weakly* lsc on $C(\Omega)^*_+$ hence both statements of Proposition 1 hold true whenever $\mu_0 \in C(\Omega)^*_+$.*

(ii) If, in addition, $\mu_0 \in \operatorname{dom} v_I(c;\cdot)$ then there exists an optimal solution to program II.

Proof. (i) It suffices to show that for every real number C the Lebesgue sublevel set $L(v_{II}(c;\cdot);C) := \{\mu_0 \in C(\Omega)^*_+ : v_{II}(c;\mu_0) \leq C\}$ is weakly* closed. According to the Krein–Shmulian theorem (see [11, Theorem V.5.7]), this is equivalent to that the intersections of $L(v_{II}(c;\cdot);C)$ with the balls $B_{C_1}(C(\Omega)^*) := \{\mu_0 \in C(\Omega)^* : \|\mu_0\| \leq C_1\}$, $C_1 > 0$, are weakly* closed. Since c is regular, one has

$$L(v_{II}(c;\cdot);C) \cap B_{C_1}(C(\Omega)^*) = \{\mu_0 : (\mu_0,\mu) \in L'(C,C_1)\}, \tag{8}$$

where

$$\begin{aligned} L'(C,C_1) := \{(\mu_0,\mu) \in C(\Omega)^*_+ \times C(\Omega)^*_+ : \|\mu_0\| \leq C_1, \|\mu\| \leq M\|\mu_0\|, \\ c(\mu) \leq C, \mu \in \mu_0 - H^0\}. \end{aligned} \tag{9}$$

Note that the functional $\mu \mapsto c(\mu)$ is weakly* lcs on $C(\Omega)^*_+$ because of lower semi-continuity of c as a function on Ω, and it follows from here that $L'(C,C_1)$ is weakly* closed hence weakly* compact in $C(\Omega)^* \times C(\Omega)^*$. Being a projection of $L'(C,C_1)$ onto the first coordinate, the set $L(v_{II}(c;\cdot);C) \cap B_{C_1}(C(\Omega)^*)$ is weakly* compact as well, and the result follows.

(ii) This follows from the weak* compactness of the constraint set of (7) along with the weak* lower semi-continuity of the functional $\mu \mapsto c(\mu)$. □

We say that the *regularity assumption* is satisfied if every H-convex function is regular.

The next result is a direct consequence of Proposition 2.

Corollary 1. *Suppose the regularity assumption is satisfied, then the duality relation $v_I(c;\mu_0) = v_{II}(c;\mu_0)$ holds true whenever c is H-convex and $\mu_0 \in C(\Omega)^*_+$. If, in addition, $\mu_0 \in \operatorname{dom} v_I(c;\cdot)$, then these optimal values are finite, and there exists an optimal solution to program II.*

We now give three examples of convex cones H, for which the regularity assumption is satisfied. In all the examples, $\Omega = X \times Y$, where X, Y are compact Hausdorff spaces.

Example 1. Suppose $H = \{h = h_{uv} : h_{uv}(x, y) = u(x) - v(y),\ u \in C(X),\ v \in C(Y)\}$. Since H is a vector subspace and $\mathbf{1}_\Omega \in H$, one has $\|\mu\| = \langle \mathbf{1}_\Omega, \mu\rangle = \langle \mathbf{1}_\Omega, \mu_0\rangle = \|\mu_0\|$ whenever $\mu \in \mu_0 - H^0$, $\mu \geq 0$, $\mu_0 \geq 0$; therefore, (7) holds with $M = 1$, and the regularity assumption is thus satisfied.

Remark 1. As follows from [42, Theorem 1.4, (b)⇔(c)] (see also [43, Theorem 10.3]), a function $c : \Omega = X \times Y \to \mathbb{R} \cup \{+\infty\}$ is H-convex relative to H from Example 1 if and only if it is bounded below and lsc. (Note that, since Ω is compact, every lsc function c is automatically bounded below.)

Example 2. Let $X = Y$ and $H = \{h = h_u : h_u(x, y) = u(x) - u(y),\ u \in C(X)\}$, then $H^0 = \{\nu \in C(\Omega)^* : \pi_1\nu - \pi_2\nu = 0\}$, where $\pi_1\nu$ and $\pi_2\nu$ are (signed) Borel measures on X as given by $\langle u, \pi_1\nu\rangle = \int_{X\times X} u(x)\, \nu(d(x, y))$, $\langle u, \pi_2\nu\rangle = \int_{X\times X} u(y)\, \nu(d(x, y))$ for all $u \in C(X)$. Observe that any H-convex function $c : \Omega = X \times X \to \mathbb{R} \cup \{+\infty\}$ is lsc (hence, bounded from below), vanishes on the diagonal ($c(x, x) = 0 \quad \forall x \in X$), and satisfies the triangle inequality $c(x, y) + c(y, z) \geq c(x, z)$ whenever $x, y, z \in X$. Moreover, it follows from [47, Theorem 6.3] that every function with such properties is H-convex. Let $\mu_0, \mu \in C(\Omega)^*_+$ and $\mu \in \mu_0 - H^0$. Then $\nu = \mu - \mu_0 \in -H^0 = H^0$, hence $\pi_1\mu - \pi_2\mu = \pi_1\mu_0 - \pi_2\mu_0$, and (2) is rewritten as

$$v_{II}(c; \mu_0) = \inf\{c(\mu) : \mu \geq 0,\ \pi_1\mu - \pi_2\mu = \pi_1\mu_0 - \pi_2\mu_0\}. \tag{10}$$

Furthermore, since c is lsc, vanishes on the diagonal, and satisfies the triangle inequality, it follows from [47, Theorem 3.1] that (10) is equivalent to

$$v_{II}(c; \mu_0) = \inf\{c(\mu) : \mu \geq 0,\ \pi_1\mu = \pi_1\mu_0,\ \pi_2\mu = \pi_2\mu_0\}. \tag{11}$$

Therefore,

$$\|\mu\| = \langle \mathbf{1}_\Omega, \mu\rangle = \langle \mathbf{1}_X, \pi_1\mu\rangle = \langle \mathbf{1}_X, \pi_1\mu_0\rangle = \langle \mathbf{1}_\Omega, \mu_0\rangle = \|\mu_0\|$$

whenever μ satisfies the constraints of (11); therefore, (7) holds with $M = 1$, and the regularity assumption is thus satisfied.

Example 3. Let $X = Y$ and $H = \{h = h_{u\alpha} : h_{u\alpha}(x, y) = u(x) - u(y) - \alpha,\ u \in C(X),\ \alpha \in \mathbb{R}_+\}$, then $(-\mathbf{1}_\Omega) \in H$, and for any $\mu \in \mu_0 - H^0$ one has $\|\mu\| - \|\mu_0\| = \langle \mathbf{1}_\Omega, \mu - \mu_0\rangle \leq 0$. Therefore, (7) holds with $M = 1$, and the regularity assumption is satisfied.

Remark 2. Taking into account Example 2, it is easily seen that any function $c : \Omega = X \times X \to \mathbb{R} \cup \{+\infty\}$ of the form $c(x, y) = \varphi(x, y) - \alpha$, where $\alpha \in \mathbb{R}_+, \varphi$ is lsc, vanishes on the diagonal, and satisfies the triangle inequality, is H-convex relative to H from Example 3. On the other hand, it is clear that any H-convex function c satisfies the condition $c(x, x) = \mathit{const} \leq 0 \quad \forall x \in X$.

Now suppose that $\mu_0 = \delta_\omega$ is the Dirac measure (delta function) at some point $\omega \in \Omega$, $\langle u, \delta_\omega\rangle := u(\omega)$ whenever $u \in C(\Omega)$. We shall show that in this case some duality results can be established without the regularity assumption.

Observe that for all $\omega \in \Omega$ one has $v_I(c; \delta_\omega) = v_I(co_H(c); \delta_\omega) = co_H(c)(\omega)$.

Proposition 3. *Two statements hold as follows:*

(i) If c is H-convex, then the duality relation $v_I(c; \delta_\omega) = v_{II}(c; \delta_\omega)$ is valid whenever $\omega \in \Omega$;

(ii) If, for a given $\omega \in \Omega$, $v_I(c; \delta_\omega) = v_{II}(c; \delta_\omega)$, then $v_I(co_H(c); \delta_\omega) = v_{II}(c; \delta_\omega) = v_{II}(co_H(c); \delta_\omega)$.

Proof. (i) By using the definition of v_I and taking into account that c is H-convex, one gets $v_I(c; \delta_\omega) = co_H(c)(\omega) = c(\omega)$. Further, since $\mu = \delta_\omega$ satisfies constraints of the dual program, it follows that $v_{II}(c; \delta_\omega) \leq c(\omega)$; hence $v_I(c; \delta_\omega) \geq v_{II}(c; \delta_\omega)$, and applying (4) completes the proof.

(ii) Since $c \geq co_H(c)$, it follows that $v_{II}(c; \delta_\omega) \geq v_{II}(co_H(c); \delta_\omega)$; therefore, $v_I(co_H(c); \delta_\omega) = v_I(c; \delta_\omega) = v_{II}(c; \delta_\omega) \geq v_{II}(co_H(c); \delta_\omega)$, and taking into account (4), the result follows. □

Let us define a function

$$c_\#(\omega) := v_{II}(c; \delta_\omega). \tag{12}$$

Clearly, $c_\# \leq c$.

Lemma 1. $H(c) = H(c_\#)$.

Proof. If $h \in H(c)$, then, for every $\mu \geq 0, \mu \in \delta_\omega - H^0$, one has $c(\mu) \geq \langle h, \mu \rangle \geq h(\omega)$, hence $c_\#(\omega) = \inf\{c(\mu) : \mu \geq 0, \mu \in \delta_\omega - H^0\} \geq h(\omega)$, that is $h \in H(c_\#)$.

If now $h \in H(c_\#)$, then $h \in H(c)$ because $c_\# \leq c$. □

The next result is a direct consequence of Lemma 1.

Corollary 2. *For every $\omega \in \Omega$, $c(\omega) \geq c_\#(\omega) \geq co_H(c)(\omega)$.*

It follows from Corollary 2 that if c is H-convex, then $c_\# = c$.

Corollary 3. *$c_\#$ is H-convex if and only if $c_\# = co_H(c)$.*

Proof. If $c_\#$ is H-convex, then $c_\#(\omega) = \sup\{h(\omega) : h \in H(c_\#)\}$, and applying Lemma 1 yields $c_\#(\omega) = \sup\{h(\omega) : h \in H(c)\} = co_H(c)(\omega)$. If $c_\#$ fails to be H-convex, then there is a point $\omega \in \Omega$ such that $c_\#(\omega) > \sup\{h(\omega) : h \in H(c_\#)\}$, and applying Lemma 1 yields $c_\#(\omega) > \sup\{h(\omega) : h \in H(c)\} = co_H(c)(\omega)$. □

Proposition 4. *The following statements are equivalent:*

(a) $c_\#$ is H-convex;

(b) the duality relation $v_I(c; \delta_\omega) = v_{II}(c; \delta_\omega)$ holds true whenever $\omega \in \Omega$;

(c) for all $\omega \in \operatorname{dom} co_H(c) := \{\omega \in \Omega : co_H(c)(\omega) < +\infty\}$, the functional $v_{II}(c; \cdot)$ is weakly lsc at δ_ω.*